R. Hildenbrandt

# DA Stochastic Dynamic Programming, Stochastic Dynamic Distance Optimal Partitioning Problems and Partitions-Requirements-Matrices

Cuvillier Verlag Göttingen
Internationaler wissenschaftlicher Fachverlag

# DA Stochastic Dynamic Programming,

# Stochastic Dynamic Distance Optimal Partitioning Problems and

# Partitions-Requirements-Matrices

R. Hildenbrandt

**Bibliografische Information der Deutschen Nationalbibliothek**
Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über http://dnb.d-nb.de abrufbar.
1. Aufl. - Göttingen: Cuvillier, 2010

978-3-86955-608-6


Nonnenstieg 8, 37075 Göttingen
Telefon: 0551-54724-0
Telefax: 0551-54724-21
www.cuvillier.de


1. Auflage, 2010
Gedruckt auf säurefreiem Papier

978-3-86955-608-6

# DA Stochastic Dynamic Programming, Stochastic Dynamic Distance Optimal Partitioning Problems and Partitions-Requirements-Matrices

R. Hildenbrandt

*Technical University Ilmenau, Institute of Mathematics*

*For my family*

I want to thank Mary Donahue for her patent and careful language corrections of this book.

# Contents

# Chapter 1

# Introduction

The **S**tochastic **D**ynamic **D**istance Optimal **P**artitioning (SDDP) problem - an Operations Research problem - was the motivation for the investigations presented in this book.

As evident from the name of the problem, investigations in two different mathematical fields were necessary for its treatment, i.e. in stochastic dynamic programming and in combinatorics ("Partitioning").

This book therefore, apart from the introduction, covers the following three chapters

2 DA Stochastic Dynamic Programming with Random Disturbances,

3 The Problem of Stochastic Dynamic Distance Optimal Partitioning (SDDP),

4 Partitions-Requirements-Matrices (PRMs).

DA ("decision after") stochastic dynamic programming with random disturbances is characterized by the fact that these random disturbances are observed before the decision is made at each stage.

In the past only very moderate attention was given to problems with this characteristic (see also Section 1.1).

Examples of DA models are SDDP problems and certain inspection-replacement problems. (Also refer to connections with k-server problems and metric task systems at the end of Section 1.2.)

In Chapter 2 specific properties of DA stochastic dynamic programming problems are worked out for theoretical characterization and for more efficient solution strategies of such problems.

In order to understand this chapter, and the book altogether, previous knowledge about stochastic dynamic programming and Markov decision processes (MDPs) is useful, however not absolutely necessary since the concerned models are developed from scratch. (Basic knowledge can be found in [7], [12], [28], [16] or [31].)

In Chapter 3 we formulate and discuss in detail the problem of **S**tochastic **D**ynamic **D**istance Optimal **P**artitioning (**SDDP**).

SDDP problems are extremely complex.

Superordinately regarded, SDDP problems are DA stochastic dynamic programming problems (**S**tochastic **D**ynamic DP).

It requires a certain initial effort, however, in order to compute the real input data for the DA stochastic dynamic programming problem (SD **D**istance optimal P).

Furthermore, the problem shows combinatorial aspects (SDD **P**artitioning).

The understanding for the formulation of the problem and the basic methods of its solution requires knowledge from Section 2.1 (at least from the beginning of this section) and absolutely from Section 2.3.

However, an important statement concerning certain SDDP problems is proven at the end of Chapter 4, only after several combinatorial considerations.

Partitions-requirements-matrices (PRMs) (Chapter 4) are matrices of transition probabilities of SDDP problems which are formulated as Markov decision processes (MDPs).

PRMs "in the strict meaning" include optimal decisions of certain SDDP problems, as is shown toward the end of Chapter 4.

PRMs (in the strict meaning) themselves represent interesting (almost self-evident) combinatorial structures, which are not otherwise found in literature.

We therefore ensure that the treatise of Chapter 4 can essentially be understood independent of Chapters 2 and 3.

Relationships to Chapter 3 specifically marked and they can be omitted if one is only interested in PRMs.

Retrospectively, in relation to the topic of "optimal dominant policies" of MDPs, PRMs in the strict meaning include policies of certain SDDP problems for which the "condition of dominance" is typically infringed on, however only to a slight extent such that a generalization of the concept of "dominant policies" seems possible.

We now discuss the contents of the chapters in more detail.

## 1.1 Chapter 2 Contents

In Section 2.1 we introduce the DA model of stochastic dynamic programming with random disturbances and give the corresponding functional equation.

In Section 2.2 a "certainty equivalence principle" is formulated and also proven in cases of DA models with linear dynamics and quadratic criteria.

Markov decision processes which result from DA models under appropriate assumptions (DA MDPs) are investigated in Section 2.3.

In literature the state space, which is used for DA MDPs, is the cross product set of the origin state space and the disturbance space.

However, such a state space is markedly larger than the original state space.

Moreover, corresponding matrices of transition probabilities would have many zeros, in general. An analogous situation is found in linear programming: the classical transportation problem which can be solved by the Simplex algorithm. Special solution methods for this transportation problem have been developed (for example the "MODI-method", refer to [30], Section 2.8.9).

In Section 2.3 we keep the origin state space when modelling DA models as MDPs. In this way special structures of decisions follow.

Here, the corresponding decisions are characterized by a "simple" structure. The transition probability matrices differ by only two elements for corresponding "neighbouring" decisions.

An effect of this structure of decisions is that optimal decisions imply an "almost-partial order" of the states, if the underlying average one-step reward functions do not depend on the decisions.

Thus, the solution of a DA MDP by solving a corresponding parameterized DA MDP in terms of a continuation of the solutions of the parameterized problem arises as one variant for solving DA MDPs, for which the Howard algorithm (policy iteration) is adapted (Section 2.3.4). For this, the underlying internal costs and hence the average one-step reward functions are considered in dependence on one parameter such that these costs do not depend on the decisions for the initial parameter. Then, the adapted Howard algorithm yields a purposeful computation for the solution. Furthermore, under certain additional conditions, this solution method is a greedy algorithm.

Section 2.3.3 includes special considerations of DA MDPs with "distance properties" and "dominant policies".

"Distance properties" can also be found in flow problems, metric task system or k-server problems. In particular, we use the statements of this section for SDDP problems.

The "dominance of Markov chains" can be seen in Daley 68 (see [10]).
We can apply this concept to Markov chains which correspond to policies of MDPs. However, if we want to transfer this concept to the MDPs themselves then convenient properties are also required for the average one-step reward functions (and for the corresponding policies).
If dominant policies should also be optimal, further strong conditions (which contain comparisons of any feasible policies with the dominant policy) are required.
The question which follows is: can we find (useful) MDPs which fulfil all of these conditions?

A certain kind of equipment replacement models with dominant policies can be found in Puterman [31]. However, in these models only two different decisions are possible.

The chance of finding MDPs with more than two decisions which fulfil these conditions is better for MDPs which are based on DA models, due to their decision structures.

Some SDDP problems have optimal dominant policies (Section 4.6.2.2).
For other SDDP problems we will consider the above-mentioned interesting effect in which the conditions of dominance are infringed on, however only to a slight extent.

The state spaces of SDDP problems are inherently finite. Therefore, we will also concentrate our efforts on finite-state models in Chapter 2. Notes on countable-state models can be found in Puterman [31]; more information can be found here at the beginning of Section 2.3.

## 1.2 Chapter 3 Contents

In Chapter 3 the "Problem of Stochastic Dynamic Distance Optimal Partitioning (SDDP)" is described in detail. Possibilities and methods of its exact or approximate solution are discussed.

A problem in industry, which contains an optimal conversion of moulds, supplied the origin of investigations.
Essentially, SDDP problems include the following practical facts:

- A fixed number of machines is given. (∗)
  (Moulds are also conceivable.)
- Different types of parts can be produced by these machines. For this purpose the machines have to be converted to states, which in accordance with the types of the parts. Costs are incurred. (∗∗)
- The production takes place in successive stages (periods).
- In a single stage, one part (at most) can be produced by one machine.
- At each stage a requirement of parts (of several types) is to be met.
  Initially, probability functions of the requirements are given.

The realizations of the requirements are known at the beginning of the stages (before decisions about conversions of machines have to be made).

- The objective is to minimize the expected cost of the conversions over all stages (or the average expected cost per stage). (To accomplish this we must decide which machine should be converted into which state in each stage.)

Thus, SDDP problems are DA stochastic dynamic programming problems.

More specifically, from a mathematical view point, we could designate this practical problem as a stochastic dynamic transportation problem, since throughout the stages feasible solutions of transportation problems must be determined (see (∗∗)). (We have also used this designation in previous papers.)

Here, however designating this problem as a stochastic dynamic distance optimal partitioning problem (SDDP) seems more appropriate. Partitioning means partitions of the number of machines into numbers of machines which are in the same state. The number of machines is therefore constant (see (∗)).

We will thus use this designation in the future.

(In this way we also emphasize the conceptual distinguishment of the designation of our problem from the typical stochastic dynamic transportation problems, see Arnold [4].) [1]

In this mathematical model, partitions of integers are the "states" of the DA stochastic dynamic programming problems (ordered partitions in general and unordered partitions in the case of certain reduced SDDP problems).

Partitioning the integers as "states" involves the combinatorial aspects of SDDP problems, which can also be observed in "matrices of transition probabilities" and "average one-step reward functions" of SDDP problems, modelled as DA MDPs.

It can therefore, only in Chapter 4 by means of combinatorial consideration, be shown that decisions for feasible states with least square sums of

[1] Further comments in connection with transportation problems and corresponding references can be found in the preface of [22].

their parts are in every case optimal for special SDDP problems.

Partitions of integers as states of DA MDPs require an enormous amount of storage space for the corresponding computer programs.

Furthermore, many transportation problems have to be solved (see (∗∗)) in order to compute "average one-step reward functions" for the SDDP problems, modelled as DA MDPs.

Thus, investigations of inherent characteristic structures of SDDP problems are also important as a basis for heuristics.

Finally, we refer to connections of SDDP problems with other problems in operations research and informatics such as stochastic dynamic facility location problems (refer to [27]) or metric task systems and more specific k-server problems, see [8], Chapter 10 and [5], for instance.

Since the current request, which is to be fulfilled, is known (and without knowing the future requests) k-server problems can also be initially labeled as a certain kind of DA model. Furthermore, distance properties are also assumed for k-server problems. However, on-line algorithms are often the center of attraction for consideration of k-server problems.

In contrast, we assume probability functions for requirements of SDDP problems and consider SDDP problems as stochastic dynamic programming problems with the aim to minimize the expected cost or the average expected cost per stage. Typical characteristics of SDDP problems as stochastic dynamic programming problems, in particular Markov decision process, are worked out.

Furthermore, let us note that we consider a number of machines which are in the same state (in the terms of k-server problems, on the same point), in general, and many machines must convert at the beginning of each equidistant stage.

## 1.3 Chapter 4 Contents

Partitions-Requirements-Matrices (PRMs) are the main topic of Chapter 4.

If SDDP problems are modelled as DA MDPs, then the matrices of transition probabilities are called "general PRMs".

The strict meaning of PRMs assumes that the costs of converting the machines into different types are identical and the requirements are identically distributed. Then in every case decisions for feasible states with least square sums of their components lead to PRMs (in the strict meaning).

The definition of PRMs (in the strict meaning) includes that PRMs can be initially computed by means of simple enumeration, however a laborious method. In addition, there is a main difficulty to deal with: No formulas are known for most of the elements in PRMs. Due to this lack of formulas, PRMs themselves represent interesting (almost self-evident) combinatorial structures.

Properties which are associated with SDDP problems (modelled as DA MDPs), besides the search for effective methods to compute the elements of PRMs, are in the realm of investigation of PRMs (in the strict meaning) in this chapter.

Thus in Section 4.6 so-called "Poisson equations" are considered. That their solutions are "monotone" is shown in many cases. This means that, in every case, decisions for feasible states with least square sums of their components are optimal for the corresponding SDDP problems.

The above-mentioned SDDP problems, for which the "condition of dominance" is infringed on, however only to a slight extent, are also in this set of SDDP problems.

A more detailed specification of the content of Chapter 4 can be found at the beginning of this chapter.

# Chapter 2

# DA Stochastic Dynamic Programming with Random Disturbances

It is assumed for many concepts in the theory of stochastic dynamic programming that random disturbances are observed after the decision is made at each stage. (For instance, refer to Bertsekas [7], Schneeweiss [33], Dinkelbach [11].)

We denote problems for which this is assumed as "Decision Before" models (DB models).

Conversely, we call problems where random disturbances are observed before the decision is made at each stage "Decision After" models (DA models).

We began to take notice of DA models with our investigation of **S**tochastic **D**ynamic **D**istance Optimal **P**artitioning (SDDP) problems [1] (see [19], [20], [22]).

In general, not much information exists dealing only with DA modelled problems.
We can find some, however, included in a book by Sebastian and Sieber [34]. Here, situations in which incomplete information is given are described by

[1] In previous papers, SDDP problems were termed stochastic dynamic transportation problems, see also Section 1.2.

means of operators as starting points for further investigations (see [34], 2.7 with $n = 1$).

Dreyfus and Law give an example in relation to certainty equivalence and also an example of a stochastic equipment inspection and replacement model, where some components of the random vector are observed after the decision is made (as usual) but some components are observed before (see [12], pages 189 and 137).

(The k-server problems mentioned at the end of Section 1.3 also show the "DA" property.)

On the one hand, DA models belong to the extensive group of stochastic dynamic programming problems, but on the other hand DA models show peculiarities.

The complexity of such problems (refer here also to the inspection/replacement problem by Dreyfus and Law) is one aspect of the motivation for the further consideration of DA models.

An introduction to the extended content of Chapter 2 has already been given in Section 1.1.

## 2.1 The DA Model

In the following we use

| | |
|---|---|
| $N \in \mathbb{N} \cup \{\infty\}$ | the horizon |
| $t \in \{1, 2, ..., N\}$ | numbers of stages |
| $S$ | state space |
| $s \in S$ | states |
| $B$ | disturbance space |
| $w \in B$ | random disturbances |
| $A$ | decision space |
| $x \in A$ | decisions (or controls) |

(Questions of measurability are skipped for the most part. In the beginning, let $S$ and $A$ be Borel spaces and let the values of w be elements of a Borel space. Afterward we often assume $S \subseteq \mathbb{Z}^n$ (or $\mathbb{R}^n$) and so on. We will use the same notations for the random vectors and their realizations.)

The above data are written with the subscript $t$ in order to attach the time to the stages $t$.

Furthermore,

$$K_t : S_t \times B_t \times A_t \to \mathbb{R}_+ \qquad \text{stage - cost (or - return) functions}$$

$$G_t : S_t \times B_t \times A_t \to S_{t+1} \qquad \text{transition functions}$$

denote (measurable) functions.

Decision spaces $A_t$ can depend on previous states and disturbances.

We now introduce the basic problem of the DA model:

**(DAP)**:

Let DA models be closed-loop optimization problems (i.e. feedback control, refer to [7], I, page 4 or [27], Section 2.4): More precisely, this means that we postpone making the decision $x_t$ until the last possible moment (time t) when the current state $s_t$ and (in the case of a DA model) the realization of the random vector $w_t$ will be known. We assume that an initial state $s_1 \in S_1$ and an initial realization $w_1$ of the random disturbances are given.

A policy

$$F = \{x_1(s_1, w_1), x_2(s_2, w_2), \ldots, x_N(s_N, w_N)\}$$

is to be found so that

$$\underset{w_2,\ldots,w_N}{E}\left(\sum_{t=1}^{N} K_t(s_t, w_t, x_t)|s_1, w_1\right) \to min$$

$$= K_1(s_1, w_1, x_1) + \underset{w_2,\ldots,w_N}{E}\left(\sum_{t=2}^{N} K_t(s_t, w_t, x_t)|w_1, s_2\right) \to min$$

subject to the constraints

$$s_t \in S_t, t = 2, \cdots, N,$$

$$x_t \in A_t(s_t, w_t), t = 1, \cdots, N$$

(dependences $A_t(\overline{s_t}, w_t)$ with $\overline{s_t} = \{s_1, \ldots, s_t\}$ are also conceivable),

$$s_{t+1} = G_t(s_t, w_t, x_t), t = 1, \ldots, N-1 \text{ (dynamic constraints).}$$

(The objective function always exists when $K_t \geq 0$, but it may have the value $\infty$ without some additional assumptions.) We assume that the distribution functions and the densities of the sequence of disturbances $\{w_t : t = 1, \ldots, N\}$ are known and that all (following) conditional expected values exist.

**Remarks 2.1.1.** *The dependence of $A_t$ on $w_t$ is a peculiarity of DA models. In DA models more information is known before the decisions are made at each stage than in the usual DB models, namely* $x_t \in A_t(s_t, \mathbf{w_t})$.

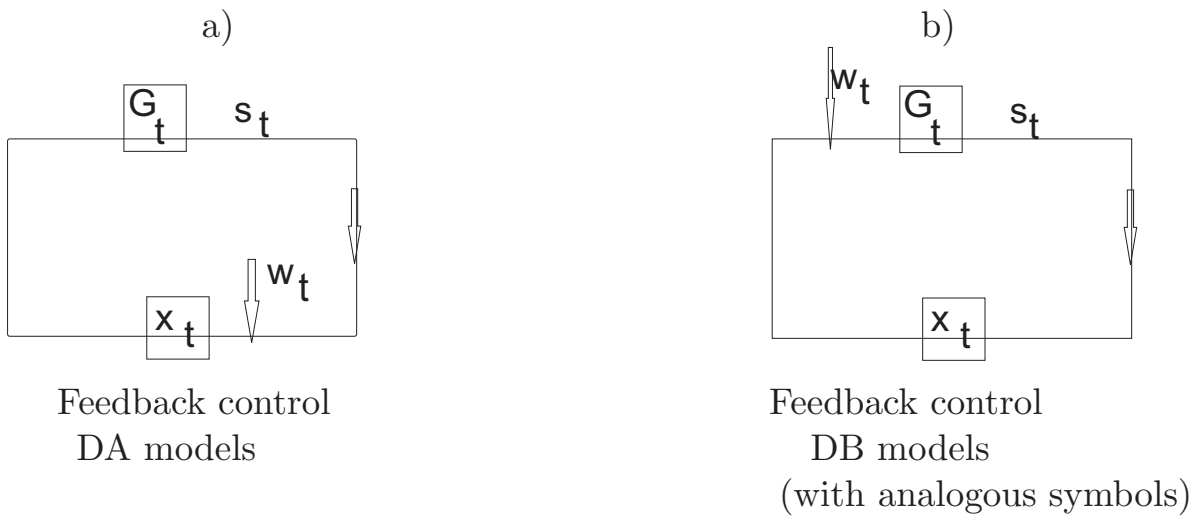



**Figure 2.1.1.**

Of course DA models are also stochastic dynamic programming problems. When a decision $x_t$ is made, then the realizations $w_{t+1}, w_{t+2}, \cdots$ of the disturbances at the next stages are not known. The cost of the next stages also depends on $s_{t+1} = G_t(s_t, w_t, x_t)$.

**The Optimal Value Function for the Remaining Periods and the Functional Equation**

We use $F_t = \{x_t(s_t, w_t), x_{t+1}(s_{t+1}, w_{t+1}), \ldots, x_N(s_N, w_N)\}$, t = 1, ...,N for any admissible policy F and the symbol $\overline{w_t} := (s_1,\ w_1,\ \ldots,\ w_t)$. (An admissible policy F = $\{x_1(s_1, w_1), x_2(s_2, w_2), \ldots, x_N(s_N, w_N)\}$ means $x_{t'} \in A_{t'}(s_{t'}, w_{t'})\ \forall\ s_{t'} \in S_{t'},\ \forall\ t' \in \{1, ..., N\}$.)

The optimal value function for the remaining periods $t, \ldots, N$ is

$$f_t(s_t\ , \overline{w_t}) = \min_{F_t} \quad \underset{w_{t+1},\ldots,w_N}{E} \quad \left(\sum_{t'=t}^{N} K_{t'}(s_{t'}, w_{t'}, x_{t'})|\overline{w_t}\right)$$

$$= \min_{F_t}\left(K_t(s_t, w_t, x_t) + \underset{w_{t+1},\ldots,w_N}{E}\left(\sum_{t'=t+1}^{N} K_{t'}(s_{t'}, w_{t'}, x_{t'})|\overline{w_t}\right)\right) \tag{2.1.1}$$

for $t = 1, ..., N-1$,

$$f_N(s_N, \overline{w_N}) = \min_{F_N} K_N(s_N, w_N, x_N)$$

for DA models.

We define

$$f_{N+1} \equiv 0. \tag{2.1.2}$$

The functional equation

$$f_t(s_t, \overline{w_t}) = \min_{x_t \in A_t(s_t, w_t)} \left(K_t(s_t, w_t, x_t) + \underset{w_{t+1}}{E}\ (f_{t+1}(s_{t+1}, \overline{w_{t+1}})|\overline{w_t})\right), \tag{2.1.3}$$

$t = N, \ldots, 1$

follows.

In the case that an optimal policy exists the functional equation can be proved directly by means of mathematical induction (refer also to Sebastian

and Sieber [34], general formula (2.188) and the upper remarks on page 147):

***Proof.***

$f_N(s_N, \overline{w_N}) := \min_{F_N} K_N(s_N, w_n, x_N)$ (for $t = N$).

**Step 1.**

(beginning of mathematical induction $t = N - 1$)

$$f_{N-1}(s_{N-1}, \overline{w_{N-1}})$$

$$:= \min_{F_{N-1}} \Big( K_{N-1}\left(s_{N-1}, w_{N-1}, x_{N-1}\right) + \underset{w_N}{E}\left(K_N(s_N, w_N, x_N) \mid \overline{w_{N-1}}\right)\Big)$$

(see (2.1.1) for $t = N - 1$)

$$= \min_{\substack{x_{N-1} \in A_{N-1}(s_{N-1}, w_{N-1}) \\ x_N \in A_N(s_N, w_N)}} \Big( K_{N-1}(s_{N-1}, w_{N-1}, x_{N-1}) +$$

$$\underset{w_N}{E}\left(K_N(s_N, w_N, x_N) \mid \overline{w_{N-1}}\right)\Big)$$

$$= \min_{x_{N-1} \in A_{N-1}(s_{N-1}, w_{N-1})} \Big\{ K_{N-1}(s_{N-1}, w_{N-1}, x_{N-1}) +$$

$$\min_{x_N \in A_N(s_N, w_N)} \Big( \underset{w_N}{E}\left(K_N(s_N, w_N, x_N) \mid \overline{w_{N-1}}\right)\Big)\Big\}.$$

(Here $\min_{x_N \in A_N(s_N, w_N)} \ldots$ means, in detail, $\min_{x_N(w_N) \in A_N(s_N, w_N)} \ldots$ $\forall\ w_N \in B_N$.)

We now use the relation $\min_x E\{\phi(x)\} = E\left\{\min_x \phi(x)\right\}$.

$$= \min_{x_{N-1} \in A_{N-1}(s_{N-1}, w_{N-1})} \Big\{ K_{N-1}(s_{N-1}, w_{N-1}, x_{N-1}) +$$

$$\underset{w_N}{E}\Big( \min_{x_N \in A_N(s_N, w_N)} K_N(s_N, w_N, x_N) \mid \overline{w_{N-1}}\Big)\Big\}$$

$$= \min_{x_{N-1}\in A_{N-1}(s_{N-1},w_{N-1})} \Big( K_{N-1}(s_{N-1}, w_{N-1}, x_{N-1}) + \underset{w_N}{E}\,(f_N(s_N, \overline{w_N}) \mid \overline{w_{N-1}}) \Big).$$

**Step** $N - t^*$ :

Let us now assume

$$f_t(s_t, \overline{w_t}) = \min_{x_t\in A(s_t,w_t)} \left( K_t(s_t, w_t, x_t) + \underset{w_{t+1}}{E}\,(f_{t+1}(s_{t+1}, \overline{w_{t+1}}) \mid \overline{w_t}) \right) \qquad (*)$$

$$\text{for } t = N, N-1, \ldots, t^*+1 \ (t^*+1 > 1).$$

We will then prove the functional equation for $t = t^*$ :

$$f_{t^*}(s_{t^*}, \overline{w_{t^*}}) := \min_{F_{t^*}} \Big( K_{t^*}(s_{t^*}, w_{t^*}, x_{t^*}) +$$

$$\underset{w_{t^*+1},\ldots,w_N}{E} \Big( \sum_{t'=t^*+1}^{N} K_{t'}(s_{t'}, w_{t'}, x_{t'}) \mid \overline{w_{t^*}} \Big) \Big)$$

(see (2.1.1))

$$= \min_{\substack{x_{t^*}\in A_{t^*}(s_{t^*},w_{t^*}) \\ \vdots \\ x_N\in A_N(s_N,w_N)}} \left( K_{t^*}(s_{t^*}, w_{t^*}, x_{t^*}) + \sum_{t'=t^*+1}^{N} \underset{w_{t'},\ldots,w_N}{E} (K_{t'}(s_{t'}, w_{t'}, x_{t'}) \mid \overline{w_{t^*}}) \right)$$

$$= \min_{\substack{x_{t^*}\in A_{t^*}(s_{t^*},w_{t^*}) \\ \vdots \\ x_N\in A_N(s_N,w_N)}} \Big\{ K_{t^*}(s_{t^*}, w_{t^*}, x_{t^*}) + \underset{w_{t^*+1},\ldots,w_N}{E} (K_{t^*+1}(s_{t^*+1}, w_{t^*+1}, x_{t^*+1})$$

$$+ \underset{w_{t^*+2},\ldots,w_N}{E} (K_{t^*+2}(s_{t^*+2}, w_{t^*+2}, x_{t^*+2})$$

$$+ \cdots + \underset{w_N}{E}\,(K_N(s_N, w_N, x_N) \mid \overline{w_{N-1}}) \mid \ \ldots \ \mid \overline{w_{t^*+1}}) \mid \overline{w_{t^*}}) \Big\}$$

$$= \min_{x_{t^*}\in A_{t^*}(s_{t^*},w_{t^*})} \Big\{ K_{t^*}(s_{t^*}, w_{t^*}, x_{t^*}) +$$

$$\underset{w_{t^*+1},\ldots,w_N}{E} \Big( \min_{x_{t^*+1}\in A_{t^*+1}(s_{t^*+1},w_{t^*+1})} (K_{t^*+1}(s_{t^*+1}, w_{t^*+1}, x_{t^*+1})$$

$$+\cdots+\underset{w_N}{E}\Big(\min_{x_N\in A_N(s_N,w_N)} K_N(s_N,w_N)\mid\overline{w_{N-1}})\mid\ \ldots)\mid\overline{w_{t^*}}\Big)\Big\}.$$

Now, we use (*) for $t = N, N-1, \ldots, t^*+1$.

$$=\min_{x_{t^*}\in A_{t^*}(s_{t^*},w_{t^*})}\Big(K_{t^*}(s_{t^*},w_{t^*},x_{t^*})+\underset{w_{t^*+1}}{E}(f_{t^*+1}(s_{t^*+1},\overline{w_{t^*+1}})\mid\overline{w_{t^*}})\Big). \qquad \blacksquare$$

For subsequent sections we introduce here:

**The "DA Decision Functions" and Additional Definitions**
(which are based on DA models)

In DA models the state $s_{t+1}$ is (for given $s_t, w_t$) completely determined by the decision (in contrast to DB models). Thus, we can introduce:
the DA decision sets

$$\hat{A}_t(s_t,w_t) := \{s' \mid s' = G_t(s_t,w_t,x_t) \text{ with } x_t \in A_t(s_t,w_t)\} \tag{2.1.4}$$

for given $s_t \in S_t$ , $w_t \in B_t$,
where $s' \in \hat{A}_t(s_t,w_t)$ are called feasible states,

internal costs

$$\begin{aligned}&\hat{c}_t(s_t,w_t,s') := min\left\{K_t(s_t,w_t,x_t)|x_t : s' = G(s_t,w_t,x_t)\right\}\\ &\text{with } s' \in \hat{A}_t(s_t,w_t)\end{aligned} \tag{2.1.5}$$

and DA decision functions

$$\begin{aligned}\hat{d}_t :\ & S_t \times B_t \to S_{t+1}\\ \text{with}\quad \hat{d}_t(s_t,w_t) &= s' \in \hat{A}_t(s_t\ ,w_t).\end{aligned} \tag{2.1.6}$$

Finally, we use

**Definition 2.1.1.** *The set of DA decision functions is the set*

$$\hat{D}_t := \{\hat{d}_t|\ \hat{d}_t :\ S_t \times B_t \to S_{t+1} \text{ with } \hat{d}_t(s_t,w_t) \in \hat{A}_t(s_t\ ,w_t)\}$$

*for given $S_t, B_t, S_{t+1}$ and DA decision sets $\hat{A}_t$.*

*In addition, single maps*

$$\left.\begin{array}{l}(s_t, w_t) \to s' \ (by\ \hat{d}) \\ \\ \text{this means } \hat{d}_t(s_t, w_t) = s'\end{array}\right\} \tag{2.1.7}$$

*for* $s_t \in S_t$ , $w_t \in B_t$ *are called single decisions.*

If $S_t$ and $B_t$ are finite sets, then $\hat{d}_t$ will include $|S_t| \cdot |B_t|$ single decisions (where $|S_t|$ and $|B_t|$ denote the numbers of elements in the sets $S_t$ and $B_t$, respectively).

With this in mind Figure 2.1.1 a) can be replaced by

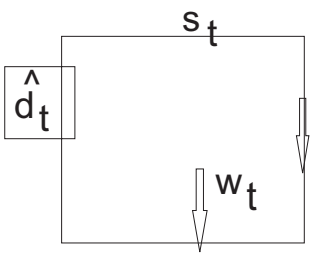



**Figure 2.1.2.**

We can see that $x_t$ and $G_t$ are combined into $\hat{d}_t$.

(DAP) can then be represented in the following way:

**(DAPa)**:

A policy

$$\{\hat{d}_1(s_1, w_1),\ \hat{d}_2(s_2, w_2),\ \ldots, \hat{d}_N(s_N, w_N)\}$$

is to be found so that

$$\underset{w_2,\ldots,w_N}{E}\left(\sum_{t=1}^{N} \hat{c}_t(s_t, w_t, s_{t+1})|s_1, w_1\right) \ \to\ \min$$

subject to the constraints

$$s_t \in S_t,\ t = 2, \cdots, N,$$

$$\hat{d}_t(s_t, w_t) \in \hat{A}_t(s_t, w_t),\ t = 1, \cdots, N,$$

$$s_{t+1} = \hat{d}_t(s_t, w_t),\ t = 1, \ldots, N-1.$$

If (DAPa) exists under the following assumptions, we use the symbol: (**DA$\bar{\mathbf{P}}$a**).

This indicates (DAPa) with

- stationary properties: the sets and functions $B_t, S_t, \hat{A}_t, \hat{d}_t, \hat{c}_t$ are the same at each stage and will be written as $B$, $S$ and so on,
- $B$ and $S$ are finite sets,
- $q(w)(q:\ B \to (0,1))$ denote the probabilities of random disturbances and these $q(\cdot)$ are also the same at every stage.

## 2.2 The Certainty Equivalence Principle

For many DB models with quadratic cost functionals and linear dynamics (so-called quadratic linear problems) it is possible to replace the random disturbances with their expected values and to then solve the yielded deterministic problems. The solutions are the same (certainty equivalence principle). We have found a similar statement for DA models.

Let us begin by considering the following example.

**Example 2.2.1.** *We contemplate the stochastic dynamic programming problems*

$$E\left(\sum_{t=1}^{N=3}\left((x_t)^2 + (s_t)^2\right)\right) \to min,$$

*where* $s_1 \in \mathbb{R}$ *or* $s_1 \in \mathbb{R}$ *and* $w_1 \in \mathbb{R}$ *are given*

*and* $\quad s_{t+1} = s_t + w_t + x_t,$

$x_t \in \mathbb{R}$.

*Here,* $\{w_t\}_{t=1,2,3}$ *is a sequence of independent random disturbances with realizations* $w_t \in \mathbb{R}$.

*Since the decision spaces* $(A_t(s_t, w_t) =)\mathbb{R}$ *(at each stage) are independent of* $w_t$*, we can classify such stochastic dynamic programming problems as DA models or as DB models (with the same data, but* $x_t(s_t, \mathbf{w_t})$ *for DA models and* $x_t(s_t)$ *for DB models).*

*The optimal solution of the DB modeled problem is*

$$x_N = x_3 = 0$$

$$x_{N-1} = x_2 = \frac{-s_2 - E(w_2)}{2}$$

$$x_{N-2} = x_1 = \frac{-3s_1 - E(w_2) - 3E(w_1)}{5}.$$

*(We can calculate this by means of the Bellman-principle or the certainty equivalence principle.)*

*The optimal solution of the DA modeled problem is*

$$x_N = x_3 = 0$$

$$x_{N-1} = x_2 = \frac{-s_2 - w_2}{2}$$

$$x_{N-2} = x_1 = \frac{-3s_1 - E(w_2) - 3w_1}{5}.$$

*(At the beginning we have calculated this by means of the Bellman-principle, see* (2.1.3)*.)*

Obviously, the minimal expected cost for the DA model are not greater than the cost for the DB model since every policy of the DB model is also possible for the DA model ($A_t(s_t, w_t)$ are independent of $w_t$).

Example 2.2.1 demonstrates the strong relationship between the solutions of the DB and DA models.

We will now generalize the results of the example.

**Quadratic-Linear-Problems**

Let us assume for (DAP) that

$$S_t = \mathbb{R}^n,\ t = 1, \ldots, N,$$

$$A_t = \mathbb{R}^q,\ t = 1, \ldots, N.$$

The dynamic constraints are

$$s_{t+1} = \Phi_t s_t + \Gamma_t x_t + \Pi_t w_t \ \ for \ \ t = 1, \ldots, N \tag{2.2.1}$$

with given matrices $\Phi_t, \Gamma_t$ and $\Pi_t$ and a given $s_1$ or given $s_1$ and $w_1$. (These symbols are taken from the model in Schneeweiss [33], Section 11.3.) The types of these matrices are determined by the types of the states, disturbances and decisions.

If $z_t = \binom{w_t}{1}, v_t = \binom{s_t}{z_t}, y_t = \binom{x_t}{v_t}$ and $T_t = (\Gamma_t, \Phi_t, \Pi_t, 0)$

are used, then (2.2.1) has the form

$$s_{t+1} = T_t y_t.$$

The cost functional is

$$E\left\{\sum_{t=1}^{N} y_t^T W_{t,yy} y_t\right\} \to min,$$

where the matrices $W_{t,yy}$ have the following structure

$$W_{t,yy} = \begin{pmatrix} W_{t,xx} & W_{t,xv} \\ W_{t,vx} & W_{t,vv} \end{pmatrix} = \begin{pmatrix} W_{t,xx} & W_{t,xs} & W_{t,xz} \\ W_{t,sx} & W_{t,ss} & W_{t,sz} \\ W_{t,zx} & W_{t,zs} & W_{t,zz} \end{pmatrix} =$$

$$= \begin{pmatrix} W_{t,xx} & W_{t,xs} & W_{t,xw} & W_{t,x1} \\ W_{t,sx} & W_{t,ss} & W_{t,sw} & W_{t,s1} \\ W_{t,wx} & W_{t,ws} & W_{t,ww} & W_{t,w1} \\ W_{t,1x} & W_{t,1s} & W_{t,1w} & W_{t,11} \end{pmatrix}$$

with regard to $v_t, s_t$ and $y_t$.

Let $W_{t,yy}$ be symmetric matrices (without loss of generality) and let $W_{t,xx}$ be positive definite. Furthermore, let all matrices $V_{xx}$ which are calculated by means of the backward dynamic programming procedure be positive definite.

Quadratic-linear-problems can be classified as DA models or as DB models with the same data, however $x_t(s_t)$ is used for DB models and $x_t(s_t, \mathbf{w_t})$ for DA models (compare Example 2.2.1).

**Theorem 2.2.1.** *(Certainty equivalence principle)*
*Let a quadratic-linear DB model and a quadratic-linear DA model with the same data be given.*

*In addition, let*

$$x_N = 0,$$

$$x_t = \varphi(E(w_t), E(w_{t+1}), \cdots, E(w_{N-1})), t = N-1, \cdots, 1$$

*be a representation of an optimal solution of the quadratic-linear DB model. Then*

$$x_N = 0,$$

$$x_t = \varphi(w_t, E(w_{t+1}), \cdots, E(w_{N-1})), t = N-1, \cdots, 1$$

*is an optimal solution of the quadratic-linear DA model.*

***Proof.*** The above symbols and the following representations are taken from the model in Schneeweiss [33] (see Section 11.3) and they are applied to the DA models here.

The functional equation for this DA problem is

$$f_t(s_t, \overline{w_t}) = \min_{x_t}\Big\{y_t^T W_{t,yy} y_t + \underset{w_{t+1}}{E}\Big\{f_{t+1}(s_t, \overline{w_{t+1}})|\overline{w_t}\Big\}\Big\}$$
$$t = N, \cdots, 1, \qquad (*1)$$

$$f_{N+1} \equiv 0$$

(see (2.1.3)).

**Step 1.**

(beginning of mathematical induction, $t = N$)

$$\begin{aligned} f_N(s_N, \overline{w_N}) &= \min_{x_N} \left( y_N^T W_{N,yy} y_N \right) \\ &= \min_{x_N} \left( x_N^T W_{N,xx} x_N + 2x_N^T W_{N,xv} v_N + v_N^T W_{N,vv} v_N \right), \quad (*2) \\ x_N^* &= -(W_{N,xx})^{-1} W_{N,xv} v_N \quad (*3) \\ &= -\left(W_{N,xx}\right)^{-1} \left(W_{N,xs} s_N + W_{N,xw} w_N + W_{N,x1}\right) \quad (*3a) \end{aligned}$$

is the optimal $x_N$ for (*2), since $W_{N,xx}$ is positive definite.

If we (*3) use in (*2), it follows that

$$f_N(s_N, \overline{w_N})$$

$$= -v_N^T W_{N,xv}^T (W_{N,xx})^{-1} W_{N,xv} v_N + v_N^T W_{N,vv} v_N$$

$$= s_N^T Q_N s_N + 2s_N^T \bar{\beta}_N + GA_N(w_N)$$

with

$$Q_N = W_{N,ss} - W_{N,xs}^T (W_{N,xx})^{-1} W_{N,xs},$$

$$\tilde{\beta}_N = (W_{N,sz} - W_{N,xs}^T (W_{N,xx})^{-1} W_{N,xz}) z_N,$$

$$GA_N(w_N) = \tilde{\gamma}_N = z_N^T W_{N,zz} z_N - z_N^T W_{N,xz}^T (W_{N,xx})^{-1} W_{N,xz} z_N.$$

Additionally, let $\beta_N$ and $\gamma_N$ denote the following:

$$\beta_N = E\{\tilde{\beta}_N|\overline{w_{N-1}}\}$$

$$= (W_{N,sz} - W_{N,xs}^T(W_{N,xx})^{-1}W_{N,xz})E\{z_N|\overline{w_{N-1}}\},$$

$$\gamma_N = E\{z_N^T W_{N,zz} z_N|\overline{w_{N-1}}\} - \hat{z}_N^T W_{N,xz}^T(W_{N,xx})^{-1}W_{N,xz}\hat{z}_N$$

where

$$\hat{z}_N = E\{z_N|\overline{w_{N-1}}\}.$$

**Step** $N-t+2$ :

Now, let us assume

$$f_t(s_t, \overline{w_t}) = s^T Q_t s_{t-1} + 2s_t^T\tilde{\beta}_t + GA_t(w). \tag{*4}$$

On the one hand we will prove

$$f_{t-1}(s_{t-1}, \overline{w_{t-1}}) = s_{t-1}^T Q_{t-1} s_{t-1} + 2s_{t-1}^T\tilde{\beta}_{t-1} + GA_{t-1}(w) \tag{*5}$$

for the optimal expected value function at stage $t-1$, where

$$Q_t := V_{t,ss}V_{t,xs}^T(V_{t,xx})^{-1}V_{t,xs},$$

$$\tilde{\beta}_t := (V_{t,sz} - V_{t,xs}^T(V_{t,xx})^{-1}V_{t,xs})z_t.$$

In addition, let $\beta_t, \tilde{\gamma}_t$ and $\gamma_t$ denote the following:

$$\beta_t = E\{\tilde{\beta}_t|\overline{w_{t-1}}\},$$

$$\tilde{\gamma}_t := z_t^T V_{t,zz} z_t - z_t^T V_{t,xz}^T(V_{t,xx})^{-1}V_{t,xz}z_t,$$

$$\gamma_t := E\{z_t^T V_{t,zz} z_t|\overline{w_{t-1}}\} - \hat{z}_t^T V_{t,xz}^T(V_{t,xx})^{-1}V_{t,xz}\hat{z}_t,\ \hat{z}_t = E\{z_t|\overline{w_{t-1}}\}.$$

Here the sub-matrices $V_{t,ij}(i,j = x, v, z, w, s, 1)$ are calculated from

$$y_t^T V_{t,yy} y_t = y_t^T W_{t,yy} y_t + y_t^T T_t^T Q_{t+1} T_t y_t + 2y_t^T \beta_{t+1} + \gamma_{t+1},$$

where it is initially given

$$Q_{N+1} = 0, \tilde{\beta}_{N+1} = 0, \gamma_{N+1} = 0, \tilde{\gamma}_{N+1} = 0.$$

We will then show that

$$\begin{aligned} x_{t-1}^* &= - \ (V_{t-1,xx})^{-1} V_{t-1,xv} v_{t-1} \\ &= - \ (V_{t-1,xx})^{-1} (V_{t-1,xs} s_{t-1} + V_{t-1,xz} z_{t-1}) \end{aligned}$$

is an optimal decision at stage $t-1$:

$(*1)$ for $t-1$ and (*4) yield

$$\begin{aligned} & f_{t-1}(s_{t-1}, w_{t-1}) \\ &= \min_{x_{t-1}} \{y_{t-1}^T \ W_{t-1,yy} y_{t-1} + E\{f_t(T_{t-1} y_{t-1}, \overline{w_t}) | \overline{w_{t-1}}\}\} \\ &= \min_{x_{t-1}} \{y_{t-1}^T \ W_{t-1,yy} \ y_{t-1} + \\ &\qquad E\{y_{t-1}^T \ T_{t-1}^T \ Q_t \ T_{t-1} \ y_{t-1} \ + \ 2y_{t-1}^T \ T_{t-1}^T \ \tilde{\beta}_t + GA_t(w) \mid \overline{w_{t-1}}\}\} \\ &= \min_{x_{t-1}} \{y_{t-1}^T \ W_{t-1,yy} \ y_{t-1} \ + y_{t-1}^T T_{t-1}^T \ Q_t \ T_{t-1} \ y_{t-1} \ + \ 2y_{t-1}^T \ T_{t-1} \ \beta_t \\ &\qquad + E\{GA_t(w) \mid \overline{w_{t-1}}\}\} \\ &= \min_{x_{t-1}} \{y_{t-1}^T \ V_{t-1,yy} \ y_{t-1} \ - \ \gamma_t + E\{GA_t \ (w) \mid \overline{w_{t-1}}\}\} \end{aligned}$$

$$= \min_{x_{t-1}} \{x_{t-1}^T \; V_{t-1,xx} \; x_{t-1} + 2x_{t-1}^T \; V_{t-1,xv} \; v_{t-1} + v_{t-1} \; V_{t-1,vv} \; v_{t-1}$$

$$-\gamma_t + E\{GA_t(w) \mid \overline{w_{t-1}}\}\}. \quad (*6)$$

$$x_{t-1}^* = -(V_{t-1,xx})^{-1} V_{t-1,xv} \; v_{t-1}$$

$$= -(V_{t-1,xx})^{-1}(V_{t-1,xs} z_{t-1} + V_{t-1,xs} \; s_{t-1}) \quad (*7)$$

follows for positive definite $V_{t-1,xx}$.

(*7) plugged into (*6) results in

$$f_{t-1}(s_{t-1}, w_{t-1})$$

$$= \; - \; v_{t-1} \; V_{t-1,xv}^T \; (V_{t-1,xx})^{-1} \; V_{t-1,xv} \; v_{t-1} \; + \; v_{t-1}^T \; V_{t-1,vv} \; v_{t-1}$$

$$- \; \gamma_t \; + \; E\{GA_t(w) \mid \overline{w_{t-1}}\}$$

$$= s_{t-1}^T \; Q_{t-1} \; s_{t-1} + 2s_{t-1}^T \; \tilde{\beta}_{t-1} + \gamma_{t-1} - \gamma_t + E\{GA_t(w) \mid \overline{w_{t-1}}\}$$

$$= s_{t-1}^T \; Q_{t-1} \; s_{t-1} + 2s_{t-1}^T \; \tilde{\beta}_{t-1} + GA_{t-1}(w).$$

Now, we compare the optimal decisions (*7) of the quadratic-linear DA models with the optimal decisions of quadratic-linear DB models, see Schneeweiss [33], Section 11.3.

The above matrices $V_{t,yy}$ are the same as the corresponding matrices in Schneeweiss.

$x_t^*$ corresponds to $u_k^*$ in Schneeweiss with the exception of $w_t$ (in $v_t$ and $z_t$). In Schneeweiss we find there $E\{w_t \mid \overline{w_{t-1}}\}(\hat{=}\hat{r}^k)$ (refer to [33], pages 162 and 163).

This concludes the proof of Theorem 2.2.1. ■

**Remarks 2.2.1.** .

- *We use the denotation "certainty equivalence principle" above (for DA models) due to of the relationship between the solutions of DB and DA models.*

- *Dreyfus and Law have a differing opinion about certainty equivalence principle for DA models (see [12], pages 275, 276). The remarks on page 276 are, however, very short. The calculations and considerations are not given in sufficient detail.*

- *Sebastian and Sieber, [34], also deal with quadratic-linear-problems (refer to 2.8.3.3). However, the disturbances do not take place in the cost functional and an interpretation of the calculations is not given.*

## 2.3 DA Models as Markov Decision Processes

In Section 2.3.1 some known basic results to Markov decision processes (MDPs) with average reward criterion which are needed in the following sections are to be found.

For this we confine ourselves to consider finite-state models since the state spaces of SDDP problems are finite. Corresponding notes on countable-state models can be found in Sections 8.10 and 8.11 by Puterman [31] and also in his Bibliographic Remarks, pages 430 and 431.

In Section 2.3.2 some characteristic properties of the structure of decisions within DA MDPs are considered. There DA MDPs are modelled in such a way that the original state space is maintained (compare also to the Introduction 1.1).

Furthermore a specification of the Howard algorithm (policy iteration) for MDPs (with average reward criterion) leads to an optimality criterion for DA MDPs.

An "almost-partial order" of the states, which is based on optimal decisions of DA MDPs with special internal cost, is observed in Section 2.3.2.2.

In what follows considerations are given to DA MDPs with special properties.
DA MDPs with distance properties are investigated. (SDDP problems fulfill distance properties.)

Then, DA MDPs with (optimal) dominant policies are considered.

Finally, "cost-parametric DA MDPs" are discussed in Section 2.3.4. In this process, the Howard algorithm can be used to solve DA MDPs, where the starting decisions for the iterations are the above-mentioned optimal decisions of DA MDPs with special internal costs.

### 2.3.1 Markov Decision Processes: Models and Properties

To begin we consider discrete-time MDPs with finite state and decision spaces. We assume infinite horizons. The average expected cost per stage will be minimized. In addition we demand stationary properties.

(Detailed representations of MDPs can be found, for example, in books by Müller and Nollau [28] (see Sections 2.1 and 2.4), by Puterman [31] (see Chapters 2, 5 and 8), by Hernández-Lerma [17] (see Sections 1.1, 1.2, 3.1, 3.2 and 3.3), by Bertsekas [7] (see Chapter 4), by Neumann [29] (see Section 3.3.2), by Girlich [15] (see Sections 5.2.6 and 5.3.4) or by Girlich, Köchel and Küenle [16] (see Chapter 5).)

We use the terminology:

| | |
|---|---|
| $N = \infty$ | horizon |
| $t \in \{1, 2, \dots\}$ | numbers of stages |
| $S$ | finite state space<br>with $m$ elements $s$ ($S = \{s^1, \dots, s^m\}$) |
| $A^M$ | sets of finite decision spaces [2] $A^M(s)$ ($s \in S$)<br>with elements (decision functions) $d$,<br>where $d : S \to A^M$ (with $d(s) \in A^M(s)$) [3] |
| $P^d = (p(s^l \vert s^f, d))_{\substack{f=1,\dots,m \\ l=1,\dots,m}} =: \left(p^d_{fl}\right)_{\substack{f=1,\dots,m \\ l=1,\dots,m}}$ | matrix of transition probabilities for any d<br>(or markov kernel) |

$\gamma^d = (\gamma(s^f, d))_{f=1,\dots,m} =: (\gamma^d_f)_{f=1,\dots,m}$
vector of average (one-step) reward functions [4] for any d with $\gamma(s^f, d) \in \mathbb{R}_+ \; \forall \, s^f \in S, \; \forall \, d \in A^M$.

At every stage $t$ $(t = 1, 2, \dots)$ with a state $s_t$ at the beginning of the stage such decisions $d_{(t)}(s^t) \in A^M(s^t)$ are to be found so that altogether the average expected cost per stage will be minimal.

The sequence $(d_{(t=1)}, d_{(t=2)}, \cdots)$ is called policy or strategy.
If $d_{(t=1)} = d_{(t=2)} = \cdots = d$ then we have a stationary policy. In this case the policy is completely determined by the decision function $d$.

For such MDPs we use the notation: MDP$(N = \infty, S, A^M, P, \gamma)$.

Under the assumption that the stationary distributions $(p^{d,\infty}_f)_{f=1,\dots,m}$ with

$$\lim_{t\to\infty} \left( \; (p^d_{fl})_{\binom{f=1,\dots,m}{l=1,\dots,m}} \; \right)^t = \begin{pmatrix} p^{d,\infty}_1 & \dots & p^{d,\infty}_m \\ \vdots & & \vdots \\ p^{d,\infty}_1 & \dots & p^{d,\infty}_m \end{pmatrix} \tag{2.3.1}$$

exist, an optimal stationary policy $d$ is to be found so that

$$g^d := \gamma(s^1, d) p^{d,\infty}_1 + \cdots + \gamma(s^m, d) p^{d,\infty}_m \to \quad \min \tag{2.3.2}$$

(refer to Theorem 2.4.7 by Müller and Nollau [28]).

If the condition

$$p^d_{fl} \neq 0 \; \forall \; f \in \{1, \dots, m\}, \; \forall \; l \in \{1, \dots, m\}^5 \tag{2.3.3}$$

[2] As by Müller and Nollau [28] (and by Hernández-Lerma [17], page 2) we use for our model only decision spaces, and not decision and action spaces.

[3] In order to put emphasis on the condition $d(s) \in A^M(s)$ we sometimes speak about **feasible** decision (functions) $d$.

[4] More detailed: $\gamma(s^f, d(s^f))$.

[5] Remark 2.3.7 in Appendix A.2.3 by Müller and Nollau, [28] includes the weaker condition that all elements of at least one column of a matrix $(P^d)^y$ $(y \in \mathbb{N})$ have to be greater than 0. However (2.3.3) will be fulfilled for SDDP problems (see Chapter 3).

is fulfilled, then an ergodic Markov chain is implied by $P^d$ and the stationary distribution exists (see Remark 2.3.7 in Appendix A.2.3 by Müller and Nollau, [28]). If (2.3.3) is valid for all stationary policies of the MDP, then an optimal stationary policy can be found (see Remark 2.3.7 in Appendix A.2.3 and Theorem 2.4.8 by Müller and Nollau, [28]).

Thus the simple property

$$\gamma(s^f, d) > 0 \text{ for at least one } s^f \;\Rightarrow\; g^d > 0 \tag{2.3.4}$$

follows.

## The Poisson Equation and Properties

If (2.3.3) is valid, then the average expected cost per stage $g^d$ satisfies the linear equation system, which is called the Poisson equation:

$$g \cdot \epsilon + \nu = \gamma^d + P^d \nu, \quad \epsilon = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \tag{2.3.5}$$

$$(\text{or } -g \cdot \epsilon + (P^d - I)\nu = -\gamma^d)$$

where I is the identity matrix and

$g$ and $(\nu_f)_{f=1,\ldots,m}$ are the variables.

If one variable $\nu_{f_0}$ is fixed in any way, then the remaining equation system with the variable $g$ and $m-1$ variables $\nu_f$ has an unique solution (refer to the proof of Theorem 2.4.8 by Müller and Nollau [28]).

Furthermore, a property of the Howard algorithm implies the following Lemma (see Theorem 2.4.9b) by Müller and Nollau [28]).

**Lemma 2.3.1.** *Let (2.3.3) be valid for all stationary policies of a given MDP($N = \infty, S, A^M, P, \gamma$). In addition, let $d^{*1}$ and $d^{*2}$ be optimal decision functions.*

*Let $(g^{*l}, \nu^{*l})$ represent solutions of the Poisson equations*

$$g\,\epsilon + \nu = P^{d^{*l}}\nu + \gamma^{d^{*l}} \quad \textit{for } l = 1; 2 \textit{ with } \nu^{*l}_m = 0,$$

*then $\nu^{*1}_f = \nu^{*2}_f$ for $f = 1, 2, \cdots, m$ and obviously $g^{*2} = g^{*1}$.*

In the following sections we also need the simple statement:

**Lemma 2.3.2.** *Let (2.3.3) be valid for any stationary policy $d$ of a given MDP$(N = \infty, S, A^M, P, \gamma)$. Furthermore, let $\gamma'$ be any affine transformation of $\gamma^d$ with*

$$\gamma' = \alpha\, \gamma^d + \beta \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}, \ \alpha \neq 0.$$

*Then, $(g_0, \nu_0) \in \mathbb{R} \times \mathbb{R}^m$ is a solution of the Poisson equation (2.3.5) if and only if $(g' = \alpha g_0 + \beta, \nu' = \alpha\nu_0) \in \mathbb{R} \times \mathbb{R}^m$ is a solution of the linear equation system*

$$g \begin{pmatrix} -1 \\ \vdots \\ -1 \end{pmatrix} + (P^d - I)\nu = -\gamma'. \qquad \textit{(2.3.5')}$$

(Remark: If $\alpha > 0$ then it obviously follows that $\nu_{0_i} < \nu_{0_j} \ \Leftrightarrow \ \nu'_i < \nu'_j$.)

### 2.3.2 DA Models as Markov Decision Processes under Appropriate Assumptions

If the stationary DA model $(DA\bar{P}a)$ with finite sets $B$ and $S$ (see Section 2.1) is transformed into a corresponding model $(DA\bar{P}b)$ with an infinite horizon $N = \infty$, where average expected cost per stage will be minimized

then the problem

(**DA$\bar{\mathbf{P}}$b**):

for given $s_1$ and $w_1$, a policy

$$\{\hat{d}_1(s_1, w_1),\ \hat{d}_2(s_2, w_2),\ \ldots\}$$

is to be found so that

$$\overline{\lim_{n\to\infty}} \frac{1}{n} E\left(\sum_{t=1}^{n} \hat{c}(s_t, w_t, s_{t+1})|s_1, w_1\right) \to inf,$$

which is subject to the constraints

$$s_{t+1} = \hat{d}(s_t, w_t),\ t = 1, 2 \cdots$$

$$\hat{d}(s_t, w_t) \in \hat{A}(s_t, w_t)(\subseteq S),\ t = 1, 2, \cdots$$

remains to be solved.

We want to represent this problem as a Markov decision process. We can do this in two ways.

We could use a method, which can be found by Neumann and Morlock [30], page 618 or Girlich, Köchel and Kuenle [16], page 36. A state space $S \times B$ would follow. Obviously, this state space is markedly larger than the original state space $S$. Furthermore, corresponding matrices of transition probabilities would have many zeros, in general. (An analogous situation is known in linear programming: If the classical transportation problems are solved by the Simplex algorithm, then many zeros exist in the coefficient matrices of the restrictions. Special solution methods for the transportation problem have been developed (for example the "MODI-method", which is based on the Simplex algorithm, see [30], Section 2.8.9).)

Thus, we will directly convert $(DA\bar{P}b)$ into a MDP. The corresponding state space is then $S$ and in addition peculiarities of the DA model can be better characterized. In this way special structures of decisions follow (refer to Sections 2.3.2.1 and 2.3.2.2).

In order to convert $(DA\bar{P}b)$ into a MDP with the original state space $S$ we use DA decision functions $\hat{d}$, the set of DA decision functions $\hat{D}$ (see (2.1.6)), Definition 2.1.1) and the internal costs $\hat{c}$ (see (2.1.5)). The disturbance space $B$ and the probability $q$ merely serve to calculate the transition probabilities of such a MDP.

$$A^M(s) = \{d(s) := (\hat{d}(s, w^1), \hat{d}(s, w^2), \cdots, \hat{d}(s, w^{IBI}))|\hat{d} \in \hat{D}\}, s \in S \tag{2.3.6}$$

(where $|B|$ denotes the number of elements in the set $B$)

$$p(s^l|s^f,d)(=p^d_{fl}) = \sum_{w:s^l=\hat{d}(s^f,w)} q(w) \tag{2.3.7}$$

$$\gamma(s^f,d)(=\gamma^d_f) = \sum_{s^l\in S}\sum_{w:s^l=\hat{d}(s^f,w)} \hat{c}(s^f,w,s^l)q(w)$$

$$= \sum_{s^l\in S}\left(\sum_{w:s^l=\hat{d}(s^f,w)} \hat{c}(s^f,w,s^l)\frac{q(w)}{p(s^l|s^f,d)}\right)p(s^l|s^f,d). \tag{2.3.8}$$

We define the cost

$$c^d(s^f,s^l) = \sum_{w:s^l=\hat{d}(s^f,w)} \hat{c}(s^f,w,s^l)\frac{q(w)}{p(s^l|s^f,d)} =: c^d_{fl}. \tag{2.3.9}$$

The relation

$$\gamma(s^f,d) = \sum_{s^l} c^d(s^f,s^l)p(s^l|s^f,d) \tag{2.3.10}$$

follows.

From this point on, DA MDP($N=\infty, S, A^M, P, \gamma$) denotes a MDP which is derived from the $(DA\bar{P}b)$ in the above way.

Under the assumption that the stationary distributions (see (2.3.1)) exist, an optimal policy $d$ is to be found so that

$$g^d = \gamma(s^1,d)p_1^{d,\infty} + \cdots + \gamma(s^m,d)p_m^{d,\infty} \to min. \tag{2.3.11}$$

We will now characterize:

**Two Special Cases for the Cost**

((a) is valid for the SDDP problem. (b) includes a more specific case, however it also leads to initial decisions for policy iteration (Howard algorithm) in the case of more general DA MDPs.)

(a) $\hat{c}(s^f, w, s^l)$ do not depend on $w$.
That means

$$\hat{c}(s^f, w, s^l) = \hat{c}(s^f, s^l) =: \hat{c}_{fl} \text{ for each w with } s^l \in \hat{A}(s^f, w). \quad (2.3.12)$$

(2.3.9) then yields

$$\begin{aligned} c^d(s^f, s^l) &= \hat{c}(s^f, s^l) \left( \sum_{w: s^l = \hat{d}(s^f, w)} \frac{q(w)}{p(s^l | s^f, d))} \right) \\ &= \hat{c}(s^f, s^l). \end{aligned}$$

(b) $\hat{c}(s^f, w, s^l)$ do not depend on $s^l$.
Thus

$$\hat{c}(s^f, w, s^l) =: \hat{c}(s^f, w) \text{ for each } s^l \in \hat{A}(s^f, w). \quad (2.3.13)$$

In other words: $\hat{c}(s^f, w, \cdot)$ are the same for all feasible decisions.

Therefore,

$$\gamma(s^f, d) = \sum_w \hat{c}(s^f, w) q(w) \text{ also do not depend on } d. \quad (2.3.14)$$

#### 2.3.2.1 The Structure of Decisions within DA Markov Decision Processes

**Definition 2.3.1.** *$d^1 \in A^M, d^2 \in A^M$ will be called neighbouring if a unique $s^{f_0} \in S$ and a unique $w^0 \in B$ exist with*

*$d^1(s) \equiv d^2(s)$ for each $s \in S$ and $s \neq s^{f_0}$,*

*$\hat{d}^1(s^{f_0}, w) = \hat{d}^2(s^{f_0}, w)$ for each $w \in B$ and $w \neq w^0$ and*

*$\hat{d}^1(s^{f_0}, w^0) \neq \hat{d}^2(s^{f_0}, w^0)$.*

*(This means $d^1$ and $d^2$ are only different in one single decision (see (2.1.7))).*

**Lemma 2.3.3.** .

**I** *Let $d^0 \in A^M$ and $d^{0_1} \in A^M$. Then a sequence $d^0, d^1, d^2, \cdots, d^v = d^{0_1}$ of neighbouring decisions $d^i, d^{i+1}$ $(0 \leq i \leq v-1)$ exists with $d^i \in A^M$.*

**II** *Now, let* $d \in A^M, \bar{d} \in A^M$ *be neighbouring with the different single decisions:*

$$\begin{aligned} \hat{d}\left(s^f, w\right) &= s^l, \\ \hat{\bar{d}}\left(s^f, w\right) &= s^{\bar{l}} \quad \left(l \neq \bar{l}\right). \end{aligned} \tag{2.3.15}$$

*The following relations then hold (in regard to the transition probabilities, the average reward functions and the cost):*

*a)*

$$\begin{aligned} p_{fl}^{\bar{d}} &= p_{fl}^{d} - q(w), \\ p_{f\bar{l}}^{\bar{d}} &= p_{f\bar{l}}^{d} + q(w), \\ p_{rv}^{\bar{d}} &= p_{rv}^{d} \text{ for } (f,l) \neq (r,v) \neq \left(f,\bar{l}\right) \quad \text{(see (2.3.7))}. \end{aligned} \tag{2.3.16}$$

*Therefore, the transition probability matrices differ only in two elements (of a row) for corresponding "neighbouring" decisions!*

*b)*

$$\begin{aligned} \gamma\left(s^f, \bar{d}\right) &= \gamma\left(s^f, d\right) + q\left(w\right)\left(\hat{c}\left(s^f, w, s^{\bar{l}}\right) - \hat{c}\left(s^f, w, s^l\right)\right), \\ \gamma\left(s^l, \bar{d}\right) &= \gamma\left(s^l, d\right) \quad \text{for } l \neq f \end{aligned} \tag{2.3.17}$$

*and*

*c)*

$$\begin{aligned} c^{\bar{d}}\left(s^f, s^l\right) &= \left(c^d\left(s^f, s^l\right) - \frac{\hat{c}\left(s^f, w, s^l\right)}{p_{fl}^{d}} q\left(w\right)\right) \frac{p_{fl}^{d}}{p_{fl}^{\bar{d}}}, \\ c^{\bar{d}}\left(s^f, s^{\bar{l}}\right) &= \left(c^d\left(s^f, s^{\bar{l}}\right) + \frac{\hat{c}\left(s^f, w, s^{\bar{l}}\right)}{p_{f\bar{l}}^{d}} q\left(w\right)\right) \frac{p_{f\bar{l}}^{d}}{p_{f\bar{l}}^{\bar{d}}}, \\ c^{\bar{d}}\left(s^r, s^v\right) &= c^d\left(s^r, s^v\right) \text{ for } (f,l) \neq (r,v) \neq \left(f,\bar{l}\right). \end{aligned} \tag{2.3.18}$$

Of course, the computation of the stationary distributions of neighbouring decisions is more complicated. In general $p_r^{d,\infty} \neq p_r^{\bar{d},\infty}$ for all $r$, however the differences between $p_r^{d,\infty}$ and $p_r^{\bar{d},\infty}$ are "greater" for $r = l$ and $r = \bar{l}$ than for other $r \in \{1, \cdots, m\}$. Moreover, we have:

**Theorem 2.3.4.** *Let $P^d$ and $P^{\bar{d}}$ be two stochastic matrices with positive elements. Let $P^d$ differ from $P^{\bar{d}}$ for only two elements in the following manner*

$$p_{fl}^d > p_{fl}^{\bar{d}} \quad \textit{and} \quad p_{f\bar{l}}^d < p_{f\bar{l}}^{\bar{d}}.$$

*Then, corresponding relations are true for the stationary distributions $p^{d,\infty}$ and $p^{\bar{d},\infty}$ belonging to $P^d$ and $P^{\bar{d}}$:*

$$p_l^{d,\infty} > p_l^{\bar{d},\infty} \quad \textit{and} \quad p_{\bar{l}}^{d,\infty} < p_{\bar{l}}^{\bar{d},\infty}.$$

(See proof and remarks in [22], Section 3.2.3.)

**Figure 2.3.1.** *If $d$ and $\bar{d}$ are different as in (2.3.15), then the average expected cost per stage are different, especially, in the bold terms*

$$\begin{aligned}
\sum_{v=1}^{m} \underbrace{\gamma(s^v, d)}\, p_v^{d,\infty} &= \cdots + \gamma(\mathbf{s^f}, \mathbf{d}) \cdot p_f^{d,\infty} + \cdots + \underbrace{\gamma(s^l, d)} \cdot \mathbf{p_l^{d,\infty}} + \cdots + \underbrace{\gamma(s^{\bar{l}}, d)} \cdot \mathbf{p_{\bar{l}}^{d,\infty}} + \cdots \\
= \ \textit{for}\, w \neq f \qquad & \qquad\qquad\qquad\qquad = \qquad\qquad\qquad\qquad = \\
\sum_{v=1}^{m} \overbrace{\gamma(s^v, \bar{d})}\, p_v^{\bar{d},\infty} &= \cdots + \gamma(\mathbf{s^f}, \mathbf{\bar{d}}) \cdot p_f^{\bar{d},\infty} + \cdots + \overbrace{\gamma(s^l, \bar{d})} \cdot \mathbf{p_l^{\bar{d},\infty}} + \cdots + \overbrace{\gamma(s^{\bar{l}}, \bar{d})} \cdot \mathbf{p_{\bar{l}}^{\bar{d},\infty}} + \cdots
\end{aligned}$$

Algorithms for approximate solutions of DA MDPs from this section which are based on the above structures are conceivable.

Finally, we want to point out a property which is naturally fulfilled for DA MDPs.

**Lemma 2.3.5.** *Let a DA MDP$(N = \infty, S, A^M, P, \gamma)$ be given. And let $s^1, s^2, \cdots, s^m$ be any ordering of the states.*

*Then, a decision function $\bar{d} \in A^M$ exists, such that*

$$\sum_{l=1}^{\bar{m}} p_{fl}^{\bar{d}} \leq \sum_{l=1}^{\bar{m}} p_{fl}^{d} \quad \textit{for} \quad \bar{m} = 1, \cdots, m$$
$$f = 1, \cdots, m \quad \textit{and} \ \forall\, d \in A^M.$$

***Proof.*** Obviously, $\bar{d}$ with

$\hat{\bar{d}}(s^f, w) = s^{\bar{l}}$, where $\bar{l} = \max\{l \mid s^l \in \hat{A}(s^f, w)\}\ \forall\, s^f \in S,\ \forall\, w \in B$

satisfies the condition. ■

**Remarks 2.3.1.** *In the above proof, feasible states with largest indices are chosen by single decisions $\hat{\bar{d}}$ in each case.*

#### 2.3.2.2 An Optimality Criterion for DA MDP and an "Almost-Partial Order" of the States

To begin, an optimality criterion for DA MDP will be derived from the Howard algorithm (policy iteration). This optimality criterion is associated with the single decisions.

Then, we show that optimal (single) decisions imply an "almost-partial order" of the states if the internal cost do not depend on the decisions.

A detailed description of the Howard algorithm can be found in [28], Section 2.4.2.1 (see also [15], Section 5.3.5). In particular, we use the addition to the Howard algorithm at the end of Section 2.4.2.1 in [28].

In relation to the MDP($N = \infty, S, A^M, P, \gamma$), where (2.3.3) is valid for all stationary policies of the MDP, the Howard algorithm includes:

If $(g^d, \nu^d)$ is a solution of the Poisson equation (2.3.5) for a decision $d$, then a better decision can be found if a state $s^f \in S$ and a decision $\bar{d}(s^f) \in A^M(s^f)$ exist such that

$$\sum_{l=1}^{m} p_{fl}^{\bar{d}} \nu_l^d + \gamma(s^f, \bar{d}) < \sum_{l=1}^{m} p_{fl}^{d} \nu_l^d + \gamma(s^f, d) \tag{2.3.19}$$

(The algorithm terminates after a finite number of iterations.)

$d$ is a optimal decision if

$$\sum_{l=1}^{m} p_{fl}^{\bar{d}} \nu_l^d + \gamma(s^f, \bar{d}) \geq \sum_{l=1}^{m} p_{fl}^{d} \nu_l^d + \gamma(s^f, d) \;\; \forall\, s^f \in S,\; \forall\, \bar{d} \in A^M. \quad (2.3.19\text{ a})$$

Now we want to specify (2.3.19) for a DA MDP.

For this reason we assume that the following single decisions of $d$ and $\bar{d}$ are different:

$$\hat{d}(s^f, w^{y_1}) = s^{l_1}, \;\; \hat{\bar{d}}(s^f, w^{y_1}) = s^{\bar{l}_1}$$

$$\vdots \qquad\qquad\qquad \vdots$$

$$\hat{d}(s^f, w^{y_z}) = s^{l_z}, \;\; \hat{\bar{d}}(s^f, w^{y_z}) = s^{\bar{l}_z},$$

Then

$$\begin{aligned}
&\sum_{l=1}^{m} p_{fl}^{\bar{d}} \nu_l^d + \gamma(s^f, \bar{d}) \\
&= \sum_{l=1}^{m} p_{fl}^{d} \nu_l^d + \gamma(s^f, d) + \sum_{i=1}^{z} q(w^{y_i}) \left[ \hat{c}(s^f, w^{y_i}, s^{\bar{l}_i}) - \hat{c}(s^f, w^{y_i}, s^{l_i}) + \nu_{\bar{l}_i}^d - \nu_{l_i}^d \right] \\
&< \sum_{l=1}^{m} p_{fl}^{d} \nu_l^d + \gamma(s^f, d)
\end{aligned}$$

follows from (2.3.19), (2.3.8), Lemma 2.3.3 a), b) and (2.3.7).

Thus

$$\sum_{i=1}^{z} q(w^{y_i}) \left[ \hat{c}(s^f, w^{y_i}, s^{\bar{l}_i}) - \hat{c}(s^f, w^{y_i}, s^{l_i}) + \nu_{\bar{l}_i}^d - \nu_{l_i}^d \right] < 0. \qquad (2.3.20)$$

Since $q(w^{y_i}) > 0$, single decisions $\hat{d}(s^f, w^{y_i}) = s^{l_i}$ and $\hat{\bar{d}}(s^l, w^{y_i}) = s^{\bar{l}_i}$ exist such that

$$\hat{c}(s^f, w^{y_i}, s^{\bar{l}_i}) - \hat{c}(s^f, w^{y_i}, s^{l_i}) + \nu_{\bar{l}_i}^d - \nu_{l_i}^d < 0.$$

We can then modify the Howard algorithm for the DA MDP.
Namely, a better decision can be found if a state $s^f \in S$ and a **single** decision $\hat{\bar{d}}(s^f, w)$ exist such that

$$\triangle H^d(s^f, w, s^{\bar{l}} - s^l) := \hat{c}(s^f, w, s^{\bar{l}}) - \hat{c}(s^f, w, s^l) + \nu_{\bar{l}}^d - \nu_l^d < 0 \quad (\text{where } s^{\bar{l}} = \hat{\bar{d}}(s^f, w),\ s^l = \hat{d}(s^f, w)). \tag{2.3.21}$$

(This means (2.3.19) is replaced by (2.3.21).)

The modified Howard algorithm for DA MDPs can be found in Section 2.3.5.

Continuing, an optimality criterion for DA MDPs is established in the following Lemma.

**Lemma 2.3.6.** *Let (2.3.3) be valid for all stationary policies of a given DA MDP($N = \infty, S, A^M, P, \gamma$). Furthermore, let $(g^d, \nu^d)$ be solutions of the Poisson equations (2.3.5) in relation to a decision $d$.*

*$d$ is a optimal decision if and only if the equalities*

$$\triangle H^d(s, w, \bar{s}' - s') = \hat{c}(s, w, \bar{s}') - \hat{c}(s, w, s') + \nu^d(\bar{s}') - \nu^d(s') \geq 0 \tag{2.3.22}$$

*(where $s' = \hat{d}(s, w)$ and $\bar{s}' = \hat{\bar{d}}(s, w)$)*

*are valid for all $s \in S, w \in B$ and all (feasible) single decisions $\hat{\bar{d}}(s, w)$.*

*If , additionally, the internal costs satisfy (2.3.13), then the inequalities (2.3.22) are simplified*

$$\triangle H^d(s, w, \bar{s}' - s') = \nu^d(\bar{s}') - \nu^d(s') \geq 0. \tag{2.3.22a}$$

**Theorem 2.3.7.** *Let (2.3.3) be valid for all stationary policies of a given DA MDP($N = \infty, S, A^M, P, \gamma$).*

*a) In addition, let $d^{*1}$ and $d^{*2}$ be optimal decision functions and*

*$(g^{*l}, \nu^{*l})$ are solutions of the Poisson equations*
*$g\,\epsilon + \nu = P^{d^{*l}}\nu + \gamma^{d^{*l}}$ for $l = 1; 2$ with $\nu_m^{*l} = 0$.*

*If*

$$\hat{d}^{*1}(s^f, w) = s^l \neq \hat{d}^{*2}(s^f, w) = s^{\bar{l}} \text{ (for some } s^f \in S \text{ and some } w \in B)$$

*then*

$$\triangle H^{d^{*1}}(s^f, w, s^{\bar{l}} - s^l) = 0 \text{ and } \triangle H^{d^{*2}}(s^f, w, s^l - s^{\bar{l}}) = 0.$$

*(where* $\triangle H^d(s^f, w, s^{\bar{l}} - s^l)$ *is defined as in (2.3.21).)*

b) *Let* $d^*$ *be an optimal decision function and* $(g^*, \nu^*)$ *be a solution of the Poisson equation*

$$g\,\epsilon + \nu = P^{d^*}\nu + \gamma^{d^*}. \qquad (*1)$$

*If*

$$\triangle H^{d^*}(s^f, w, s^{\bar{l}} - s^l) = 0 \quad \text{for some } s^f \in S, \text{ some } w \in B \text{ and some } s^{\bar{l}} = \hat{d}(s^f, w) \neq s^l = \hat{d}^*(s^f, w),$$

*then*

*d with*

$$\hat{d}(\bar{s}, \bar{w}) = \begin{cases} s^{\bar{l}} & \text{if } (\bar{s}, \bar{w}) = (s^f, w), \\ \hat{d}^*(\bar{s}, \bar{w}) & \text{otherwise} \end{cases}$$

*is also an optimal decision function.*

**Proof:**

a) Assumption: $\triangle H^{d^{*1}}(s^f, w, s^{\bar{l}} - s^l) \neq 0$.

In addition, as a consequence of the optimality criterion (2.3.22),

$$\triangle H^{d^{*1}}(s^f, w, s^{\bar{l}} - s^l) > 0$$

follows.

Since $d^{*1}$ and $d^{*2}$ are optimal decisions, Lemma 2.3.1 implies

$$\nu^{*1} = \nu^{*2},$$

hence

$$\triangle H^{d^{*2}}(s, w, s^l - s^{\bar{l}}) = \hat{c}(s^f, w, s^l) - \hat{c}(s^f, w, s^{\bar{l}}) + \nu_l^{*2} - \nu_{\bar{l}}^{*2}$$
$$= -(\hat{c}(s^f, w, s^{\bar{l}}) - \hat{c}(s^f, w, s^l) + \nu_{\bar{l}}^{*1} - \nu_l^{*1}) = -\triangle H^{d^{*1}}(s, w, s^{\bar{l}} - s^l) < 0.$$

This is a contradiction to the optimality of $d^{*2}$.

(Similarly, $\triangle H^{d^{*2}}(s, w, s^l - s^{\bar{l}}) = 0$ can be shown.)

b) We show that $(g^*, \nu^*)$ also satisfies the Poisson equation

$$g\,\epsilon + \nu = P^d\nu + \gamma^d. \qquad (*2)$$

The Poisson equations (*1) and (*2) differ only in row $f$.

According to (2.3.16) and (2.3.17), row $f$ of (*2) can be represented as

$$\begin{aligned} g + \nu_f &= p^d_{f1}\nu_1 + \cdots + p^d_{fl-1}\nu_{l-1} + (p^d_{fl} - q(w))\nu_l + p^d_{fl+1}\nu_{l+1} + \cdots \\ &\quad + p^d_{f\bar{l}-1}\nu_{\bar{l}-1} + (p^d_{f\bar{l}} + q(w))\nu_{\bar{l}} + p^d_{f\bar{l}+1}\nu_{\bar{l}+1} + \cdots + p^d_{fm}\nu_m \\ &\quad + \gamma^d_f + q(w)[\hat{c}(s^f, w, s^{\bar{l}}) - \hat{c}(s^f, w, s^l)] \\ &= p^d_{f1}\nu_1 + \cdots + p^d_{fm}\nu_m + \gamma^d_f \\ &\quad + q(w)[\hat{c}(s^f, w, s^{\bar{l}}) - \hat{c}(s^f, w, s^l) + \nu_{\bar{l}} - \nu_l]. \end{aligned} \tag{*3}$$

The assumption

$$\triangle H^{d^*}(s^f, w, s^{\bar{l}} - s^l) = \hat{c}(s^f, w, s^{\bar{l}}) - \hat{c}(s^f, w, s^l) + \nu^*_{\bar{l}} - \nu^*_l = 0$$

yields that $(g^*, \nu^*)$ satisfies (*3), thus also the Poisson equation (*2).

This means the value $g^*$ of the objective function for $d^*$ is the same as for $d$. Hence, $d$ is also an optimal decision function.

■

Finally, the relation

$$\begin{aligned} &\triangle H^d(s^f, w, s^{l_1} - s^{l_2}) = \triangle H^d(s^f, w, s^{l_1} - s^{l_3}) + \triangle H^d(s^f, w, s^{l_3} - s^{l_2}) \\ &\text{(where } s^{l_i} = \hat{d}^{\,l_i}(s^f, w) \text{ for } i = 1, 2, 3) \end{aligned} \tag{2.3.23}$$

follows easily from the Definition of $\triangle H^d(.,.,.)$ (see (2.3.21)).

### An "Almost-Partial Order" of the States

**Definition 2.3.2.** *Let a set $M$ and a corresponding binary relation $<$ be given.*

*(i) The relation is called almost-transitive, if*

$$\{\{x_1, x_2, x_3\} \subset M \ \wedge \ x_1 < x_2 \ \wedge \ x_2 < x_3\} \Rightarrow \{x_3 \not< x_1\}.$$

*($x_3 \not< x_1$ means either $x_1 < x_3$ or $x_1$ and $x_3$ may not be related to each other in this way [6].)*

[6] This possibility leads to the terminology: "**almost-**" transitive.

(ii) *The relation $<$ is called an (strict) almost-partial order (and the set $M$ an almost-partially ordered set), if $<$ is a irreflexive, asymmetric and almost-transitive relation.*

**Theorem 2.3.8.** *Let a DA MDP$(N = \infty, S, A^M, P, \gamma)$ be given where the underlying internal costs do not depend on decisions (this means the internal cost satisfy (2.3.13)). Furthermore, let (2.3.3) be valid for all stationary policies of this DA MDP.*

(i) *Then any optimal decision $d^*$ implies an almost-partial order $<$ of the states in the following way*

$$\begin{aligned} \left\{s^l < s^f\right\} := \Big\{ & \exists s \in S, w \in B : \hat{d}^*(s,w) = s^f \\ & \wedge\, s^l \in \hat{A}(s,w) \wedge \ \nexists \text{ optimal } d'^* : \hat{d}'^*(s,w) = s^l \Big\} \end{aligned} \tag{2.3.24}$$

*(In addition, $s^l$ is called costlier than $s^f$.)*

(ii) *In regard to the solutions $(g^*, \nu^*)$ of the Poisson equation $g\,\epsilon + \nu = P^{d^*}\nu + \gamma^{d^*}$ the relations*

$$s^l < s^f \ \Rightarrow\ \nu_l^* > \nu_f^* \tag{2.3.25}$$

*are valid.*

(iii) *If $s^l \in S, s^f \in S$, $s \in S, w \in B$ and optimal decisions $d^*, d'^*$ exist so that $\hat{d}^*(s,w) = s^f, \hat{d}'^*(s,w) = s^l$, then most importantly,*

$$\nu_l^* \ = \ \nu_f^*. \tag{2.3.26}$$

*follows for the solutions of the Poisson equation.*

***Proof.***

(ii) Suppose $s^l < s^f$ as in (2.3.24).

This also includes

$\exists s \in S, w \in B : \{s^f, s^l\} \subset \hat{A}(s,w)$

and $\hat{d}^*(s,w) = s^f$ for an optimal decision $d^*$.

Initially, the optimality criterion (2.3.22 a) from Lemma 2.3.6 leads to

$\triangle H^{d^*}(s,w,s^l - s^f) = \nu_l^* - \nu_f^* \geq 0,$

hence $\nu_l^* \geq \nu_f^*$.

Assumption: $\nu_l^* = \nu_f^*$.

We then see that the optimality criterion (2.3.22 a) from Lemma 2.3.6 is also satisfied for the decision $d'$ with

$$\begin{cases} \hat{d}'(s',w') = \hat{d}^*(s',w') \text{ for } (s',w') \neq (s,w), \\ \hat{d}'(s,w) = s^l \end{cases}.$$

Thus, the decision $d'$ with $d'(s,w) = s^l$ is an optimal decision contradicting (2.3.24).

(i) 1. Asymmetry:

Assumption: $s^l < s^f$ and $s^f < s^l$.

Then (ii) yields

$\nu_l^* > \nu_f^*$ and $\nu_l^* < \nu_f^*$. This is a contradiction to the order of real numbers.

2. Almost-transitivity:

Assumption: $s^l < s^f$, $s^f < s^y$ and $s^l > s^y$.

Then (ii) yields

$\nu_l^* > \nu_f^*, \nu_f^* > \nu_y^*$ and $\nu_l^* < \nu_y^*$.

This is also a contradiction to the order of real numbers.

(iii) $\nu_l^* = \nu_f^*$ follows from Theorem 2.3.7a) and (2.3.22a) under the condition (2.3.13). ■

**Corollary 2.3.9.** *Suppose that the assumptions from Theorem 2.3.8 hold.*

*(i) If $\hat{A}(s,w) = \hat{A}(s',w')$ for $\{s,s'\} \subseteq S$; $\{w,w'\} \subset B$, then an optimal decision $d$ with $\hat{d}(s,w) = s^f$ ($s^f \in \hat{A}(s,w)$) exists if and only if an optimal decision $d'$ with $\hat{d}'(s',w') = s^f$ exists.*

*(ii)* *If* $\{s^f, s^l\} \subseteq \hat{A}(s,w)$, $\{s^f, s^l\} \subseteq \hat{A}(s',w')$ *for* $\{s, s', s^f, s^l\} \subseteq S$, $\{w, w'\} \subseteq B$, $d^*$ *is an optimal decision with* $\hat{d}^*(s,w) = s^f$ *and no optimal decision* $d$ *with* $\hat{d}(s,w) = s^l$ *exists it then follows that no optimal decision* $d'$ *with* $\hat{d}'(s',w') = s^l$ *exists.*

***Proof.***

(i) If $d$ is optimal it then follows from Lemma 2.3.6, (2.3.22a) that $\triangle H^d(s,w,s^l - s^f) = \nu_l^d - \nu_f^d \geq 0$ for each $s^l \in \hat{A}(s,w)$.

Furthermore, this Lemma yields $\triangle H^d(s,w,s^l - s^f) = \triangle H^d(s',w',s^l - s^f) = \nu_l^d - \nu_f^d$ (under the assumption (2.3.13)).

Thus $d'$ with

$$\begin{cases} \hat{d}'(s,w) = \hat{d}(s,w) \text{ for } (s,w) \neq (s',w'), \\ \hat{d}'(s',w') = s^f \end{cases}$$

is also an optimal decisions.

(ii) $s^l < s^f$ and $\nu_l^* - \nu_f^* > 0$ follow from Theorem 2.3.8(i) and (ii), respectively.

Lemma 2.3.6 in connection with the latter inequality implies that no optimal decision $d'$ with $\hat{d}'(s',w') = s^l$ can exist. ■

We additionally now establish the following Definition 2.3.3 and Corollary 2.3.10.

**Definition 2.3.3.** *Let a DA MDP*$(N = \infty, S, A^M, P, \gamma)$ *be given where the underlying internal costs do not depend on decisions (this means the internal cost satisfy (2.3.13)). In addition, let (2.3.3) be valid for all stationary policies of this DA MDP.*

*(a)* *If a partial (almost-partial) order* $<$ *of the state space* $S$ *is given, then a solution of a Poisson equation (2.3.5) or a solution of a linear equation system (2.3.5') is called monotone (in* $\nu$*) with respect to the partial (almost-partial) order if*

$s^l < s^f \Rightarrow \nu_l > \nu_f$.

(b) *Let $d$ be a (feasible) decision.* [7]

*Then a solution of the corresponding Poisson equation (2.3.5) or a solution of a linear equation system (2.3.5') is called monotone (in $\nu$) in relation to $d$ if*

$$\begin{aligned}\Big\{\exists s \in S, w \in B : \hat{d}(s,w) = s^f \ \wedge \ s^l \in \hat{A}(s,w) \\ \wedge \ \nexists\, d' \ mit\ g^d = g^{d'} : \hat{d}'(s,w) = s^l\Big\}\end{aligned} \tag{2.3.27}$$

[8]

$$\Rightarrow \nu_l^d > \nu_f^d.$$

**Corollary 2.3.10.** *Suppose that the assumptions from Definition 2.3.3 hold. In addition, let $d$ be a (feasible) decision.*

*Then the solutions of the corresponding Poisson equation (2.3.5) or the solutions of the linear equation systems (2.3.5') are monotone (in $\nu$) in relation to $d$ if and only if $d$ is an optimal decision.*

***Proof.***

1. Let $d$ be an optimal decision. Then Theorem 2.3.8(i) and (ii) imply that the solutions of (2.3.5) or (2.3.5') are monotone (in $\nu$) in relation to $d$
   (also in relation to the almost-partial order which is induced by $d$ as in Theorem 2.3.8(i)).

2. If the solutions of (2.3.5) or (2.3.5') are monotone (in $\nu$) in relation to $d$, then the optimality of $d$ follows from Lemma 2.3.6 (under the assumption (2.3.13)) (and the Howard algorithm).

■

**Remarks 2.3.2.** *Suppose that the assumptions of Theorem 2.3.8 hold. Frequently the relation $s^l < s^f$ implies the relation $\gamma(s^l) > \gamma(s^f)$.*

[7] Contrary to Theorem 2.3.8 the decision $d$ is not optimal, at first.
[8] (2.3.27) is analogous to (2.3.24).

*However, counter-examples also exist:*

*Let* $\gamma^T = (5, 2, 1)^T$ *be a vector of average one-step reward functions (which fulfills (2.3.14)) and* $p^{d_i}$, $i = 1, 2, 3$ *three (feasible) decisions with the following corresponding matrices of transition probabilities:*

$$P^{d_i} = \frac{1}{6}\begin{pmatrix} 4 & 1 & 1 \\ 1+\delta_{1,i} & 1+\delta_{2,i} & 1+\delta_{3,i} \\ 4 & 1 & 1 \end{pmatrix} \text{ where } \delta_{j,i} = \begin{cases} 3, & \text{if } i = j \\ 0, & \text{otherwise} \end{cases}$$

$$\text{for } i = 1, 2, 3, j = 1, 2, 3.$$

*Obviously, (2.3.3) is valid in relation to* $d_1$, $d_2$ *and* $d_3$.

*The average expected costs per stage can be computed by solving of the Poisson equations (2.3.5). The results are:*

$$g^{d_1} = 23/6, \quad g^{d_2} = 10/3, \quad g^{d_3} = 7/2.$$

*Thus,* $d_2$ *is the optimal decision.*

*If the considered MDP is a DA MDP, which means the foundation is a* $(DA\bar{P}b)$ *model,* [9] *then the relation* $s^3 < s^2$ *follows according to (2.3.24). However, for the average one-step reward functions the relation* $\gamma(s^3) < \gamma(s^2)$ *is valid.*

### 2.3.3 DA Models with Special Properties

#### 2.3.3.1 Notes on DA Models with Distance Properties

In this section we will assume "distance properties" for DA models. Such properties can be found, for instance, in flow problems (for example see [1], near Theorem 3.4), metric task system or k-server problems (see [8], Chapter 10, for instance).

For the SDDP problems (Chapter 3) these distance properties are fulfilled.

The "distance properties" include a "triangle-inequality" and an internal cost of the value 0, if the state does not change at this stage. Here, we do not assume the "commutativity of distances" since, for SDDP problems, the costs can be different if a machine of type $i$ is converted to type $j$ or a machine of type $j$ is converted to type $i$, (see Section 1.2 and Chapter 3).

For stochastic dynamic DA models the distance properties have to be formulated in connection with the DA decisions sets (contrary to deterministic

[9] We do not discuss this here in detail.

flow problems, k-server problems and so on).

The purpose of the procedural method of this section:
initially, we give the triangle-inequality with an additional equals sign, which later will do not play a role for optimal decisions, according to Theorem 2.3.13 will become clear in Chapter 3.
(In Section 3.3 certain decisions, which are not better than others can be excluded from the outset according to Theorem 2.3.13, see Lemma 3.3.8 and Example 3.3.1. [10])

For DA MDPs a transition probability of a state transitioning to itself will prove to be constant under the assumption of the distance properties.

In the following consideration we assume $(DA\bar{P}a)$ or $(DA\bar{P}b)$ models for which

**The Distance Properties**

1. $\hat{c}(s^f, w, s^l) = 0$ if and only if $s^f \in \hat{A}(s^f, w)$ and $s^l = s^f$, (2.3.28)

2. $$\left.\begin{array}{l} \text{if } s^v \in \hat{A}(s^l, w^2), s^v \in \hat{A}(s^f, w^1) \text{ and } s^l \in \hat{A}(s^f, w^1), \\ \text{then } \hat{c}(s^l, w^2, s^v) + \hat{c}(s^f, w^1, s^l) \geq \hat{c}(s^f, w^1, s^v) \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{must follow,} \end{array}\right\} \quad (2.3.29)$$

*(triangle-inequality)*

are additionally satisfied.

We will show that states $s^v$ are not essential for optimal single decisions $\hat{d}^*\left(s^f, w^1\right)$ if we have equality in (2.3.29).
For this purpose we define the smaller DA decision sets.

**Definition 2.3.4.** *Let $s^f \in S, w^1 \in B$ be given.*

$$\hat{\hat{A}}\left(s^f, w^1\right) = \{s^v \in \hat{A}\left(s^f, w^1\right) \mid \exists\, w^2 \in B :$$
$$\hat{c}\left(s^l, w^2, s^v\right) + \hat{c}\left(s^f, w^1, s^l\right) > \hat{c}\left(s^f, w^1, s^v\right) \;\forall\, s^l \in \hat{A}(s^f, w^1)$$

[10] In other words, using "lazy algorithms" is sufficient in order to compute optimal solutions of SDDP problems.

*with* $s^v \in \hat{A}\left(s^l, w^2\right), s^l \neq s^f$ *and* $s^l \neq s^v\}$

*is called the smaller DA decision set of feasible states for the given state* $s^f$ *and the realized disturbance* $w^1$.

**Lemma 2.3.11.** *Let* $(DA\bar{P}a)$ *or* $(DA\bar{P}b)$ *be a DA model for which the distance properties (2.3.28) and (2.3.29) are additionally satisfied. Then the properties*

*(i)* $\hat{\hat{A}}(s^f, w^1) = \{s^v\}$ *if* $\hat{A}(s^f, w^1) = \{s^v\}$,

*(ii)* $s^v \in \hat{\hat{A}}(s^f, w^1)$ *if* $\hat{c}(s^f, w^1, s^v) = \min\limits_{s^l \in \hat{A}(s^f, w^1)} \hat{c}(s^f, w^1, s^l)$

*and*

*(iii)* $\hat{\hat{A}}(s^f, w^1) \neq \emptyset$ *for any* $s^f \in S, w \in B$

*are valid.*

***Proof.***

(i) We set $w^2 = w^1$ in relation to Definition 2.3.4 and see that there exists no $s^l \neq s^v$ with $s^l \in \hat{A}(s^f, w^1)$ according to (i). Thus, the inequality in Definition 2.3.4 is insignificant, however true, and $s^v \in \hat{\hat{A}}(s^f, w^1)$ follows.

(ii) Assume $s^l \in \hat{A}(s^f, w^1)$, $s^l \neq s^v$ and $s^v \in \hat{A}(s^l, w^2)$.

According to $\hat{c}(s^f, w^1, s^v) = \min\limits_{s^l \in \hat{A}(s^f, w^1)} \hat{c}(s^f, w^1, s^l)$ (see (ii))

$$\hat{c}(s^f, w^1, s^v) \leq \hat{c}(s^f, w^1, s^l) \tag{*1}$$

follows.

Additionally, from (2.3.28) we find that

$$\hat{c}(s^l, w^2, s^v) \neq 0 \text{ for } s^l \neq s^v. \tag{*2}$$

(*1) and (*2) together result in

$$\hat{c}(s^f, w^1, s^v) < \hat{c}(s^f, w^1, s^l) + \hat{c}(s^l, w^2, s^v),$$

which means $s^v \in \hat{\hat{A}}(s^f, w^1)$.

(iii) Since $(DA\bar{P}a)$ and $(DA\bar{P}b)$ include finite sets $S$ and $B$, $s^v \in \hat{\hat{A}}(s^f, w^1)$ as in (ii) exists. ■

We will now establish the following lemma and use it in the proof of Theorem 2.3.13.

**Lemma 2.3.12.** *Let $(DA\bar{P}a)$ or $(DA\bar{P}b)$ be a DA model for which the distance properties (2.3.28) and (2.3.29) are additionally satisfied. Furthermore, let the condition*

$$\begin{aligned} &\left\{ \{s^f, s^l\} \subseteq \hat{A}(s,w) \textit{ for } s \in S,\ w \in B \right\} \\ &\qquad \Rightarrow \quad \left\{ \hat{A}(s^f, w') = \hat{A}(s^l, w') \textit{ for each } w' \in B \right\} \end{aligned} \tag{2.3.30}$$

*be valid.*

*If $s^v \in \hat{A}(s^f, w^1)$ however $s^v \notin \hat{\hat{A}}(s^f, w^1)$ then*

$$\forall\, w^2 \in B\ \ \exists\, s^l (\neq s^v) \in \hat{\hat{\mathbf{A}}}(s^f, w^1) \textit{ with } s^v \in \hat{A}\left(s^l, w^2\right) :$$
$$\hat{c}\left(s^l, w^2, s^v\right) + \hat{c}\left(s^f, w^1, s^l\right) = \hat{c}\left(s^f, w^1, s^v\right).$$

***Proof.*** (2.3.29) and Definition 2.3.4 yield

$$\left\{ s^v \in \hat{A}(s^f, w^1), s^v \notin \hat{\hat{A}}(s^f, w^1) \right\} \Rightarrow$$
$$\left\{ \begin{array}{c} \text{for all } w^2 \in B\ \exists\, s^{l_1} (\neq s^v) \in \hat{\mathbf{A}}(s^f, w^1) \text{ with } s^v \in \hat{A}\left(s^{l_1}, w^2\right) : \\ \hat{c}\left(s^{l_1}, w^2, s^v\right) + \hat{c}\left(s^f, w^1, s^{l_1}\right) = \hat{c}\left(s^f, w^1, s^v\right). \end{array} \right\} \tag{*1}$$

Since $\hat{c}\left(s^{l_1}, w^2, s^v\right) \neq 0$ according to (2.3.28)

$$\hat{c}\left(s^f, w^1, s^{l_1}\right) < \hat{c}\left(s^f, w^1, s^v\right) \tag{*1 a}$$

follows from (*1).

If we suppose that $s^{l_1} \notin \hat{\hat{A}}(s^f, w^1)$, then (2.3.29) and Definition 2.3.4 yield

$$\left\{ \begin{array}{c} \text{for all } \ w^2 \in B \ \exists\, s^{l_2} (\neq s^{l_1}) \in \ \hat{\mathbf{A}}(s^f, w^1) \text{ with } \ s^{l_1} \in \hat{A}\left(s^{l_2}, w^2\right) : \\ \hat{c}\left(s^{l_2}, w^2, s^{l_1}\right) + \hat{c}\left(s^f, w^1, s^{l_2}\right) = \hat{c}\left(s^f, w^1, s^{l_1}\right). \end{array} \right\} \tag{*2}$$

Since $\hat{c}\left(s^{l_2}, w^2, s^{l_1}\right) \neq 0$ according to (2.3.28)

$$\hat{c}\left(s^f, w^1, s^{l_2}\right) < \hat{c}\left(s^f, w^1, s^{l_1}\right) < \hat{c}\left(s^f, w^1, s^v\right) \tag{*2 a}$$

follows from (*2) and (*1 a).

If we iteratively suppose that $s^{l_2} \notin \hat{\hat{A}}(s^f, w^1)$, then (2.3.29) and Definition 2.3.4 yield

$$\left\{ \begin{array}{c} \text{for all } \ w^2 \in B \ \exists\, s^{l_3} (\neq s^{l_2}) \in \ \hat{\mathbf{A}}(s^f, w^1) \text{ with } \ s^{l_2} \in \hat{A}\left(s^{l_3}, w^2\right) : \\ \hat{c}\left(s^{l_3}, w^2, s^{l_2}\right) + \hat{c}\left(s^f, w^1, s^{l_3}\right) = \hat{c}\left(s^f, w^1, s^{l_2}\right). \end{array} \right\} \tag{*3}$$

Since $\hat{c}\left(s^{l_3}, w^2, s^{l_2}\right) \neq 0$ according to (2.3.28)

$$\hat{c}\left(s^f, w^1, s^{l_3}\right) < \hat{c}\left(s^f, w^1, s^{l_2}\right) < \hat{c}\left(s^f, w^1, s^{l_1}\right) < \hat{c}\left(s^f, w^1, s^v\right) \tag{*3 a}$$

follows from (*3) and (*2 a).

If we suppose that $s^{l_3} \notin \hat{\hat{A}}(s^f, w^1)$, then we can iteratively pursue the above procedural method to:

$s^{l_{j-1}} \notin \hat{\hat{A}}(s^f, w^1)$, then (2.3.29) and Definition 2.3.4 yield

$$\left\{ \begin{array}{c} \text{for all } \ w^2 \in B \ \exists\, s^{l_j} (\neq s^{l_{j-1}}) \in \ \hat{\mathbf{A}}(s^f, w^1) \text{ with } \ s^{l_{j-1}} \in \hat{A}\left(s^{l_j}, w^2\right) : \\ \hat{c}\left(s^{l_j}, w^2, s^{l_{j-1}}\right) + \hat{c}\left(s^f, w^1, s^{l_j}\right) = \hat{c}\left(s^f, w^1, s^{l_{j-1}}\right). \end{array} \right\} \tag{*j}$$

Since $\hat{c}\left(s^{l_j}, w^2, s^{l_{j-1}}\right) \neq 0$ according to (2.3.28)

$$\hat{c}\left(s^f, w^1, s^{l_j}\right) < \hat{c}\left(s^f, w^1, s^{l_{j-1}}\right) < \cdots < \hat{c}\left(s^f, w^1, s^v\right) \tag{*j a}$$

follows from (*j) and (*(j-1) a).

Since $S$ is finite, chains of inequalities such as (*j a) cannot be infinite.

Thus, $j \geq 1$ has to exist, so that $\mathbf{s^{l_j}} \in \hat{\hat{\mathbf{A}}}(\mathbf{s^f}, \mathbf{w^1})$.

We now substitute successively:

(*2) into (*1):

$$\hat{c}\left(s^{l_1}, w^2, s^v\right) + c\left(s^{l_2}, w^2, s^{l_1}\right) + \hat{c}\left(s^f, w^1, s^{l_2}\right) = \hat{c}\left(s^f, w^1, s^v\right). \tag{*2 b}$$

$s^{l_1} \in \hat{A}(s^f, w^1)$, $s^{l_2} \in \hat{A}(s^f, w^1)$ (see (*1), (*2)) and (2.3.30) imply that $\hat{A}(s^{l_1}, w^2) = \hat{A}(s^{l_2}, w^2)$. Since $s^v \in \hat{A}(s^{l_1}, w^2)$ (see (*1))

$$s^v \in \hat{A}(s^{l_2}, w^2) \tag{*2 c}$$

follows.

Now we are in a position to apply (2.3.29) to the first two summands in (*2 b).

The inequality

$$\hat{c}\left(s^{l_2}, w^2, s^v\right) + \hat{c}\left(s^f, w^1, s^{l_2}\right) \leq \hat{c}\left(s^f, w^1, s^v\right) \tag{*2 d}$$

follows.

Since $s^v \in \hat{A}(s^f, w^1)$ we can apply (2.3.29) to the left side of (*2 d):

$$\hat{c}\left(s^f, w^1, s^v\right) \leq \hat{c}\left(s^f, w^1, s^v\right).$$

Hence equal signs can only be possible in the last two inequalities.

This means

$$\hat{c}\left(s^{l_2}, w^2, s^v\right) + \hat{c}\left(s^f, w^1, s^{l_2}\right) = \hat{c}\left(s^f, w^1, s^v\right). \tag{*2 e}$$

Substituting (*3) into (*2 e) yields

$$\hat{c}\left(s^{l_2}, w^2, s^v\right) + c\left(s^{l_3}, w^2, s^{l_2}\right) + \hat{c}\left(s^f, w^1, s^{l_3}\right) = \hat{c}\left(s^f, w^1, s^v\right). \tag{*3 b}$$

$s^{l_2} \in \hat{A}(s^f, w^1)$, $s^{l_3} \in \hat{A}(s^f, w^1)$ (see (*2), (*3)) and (2.3.30) imply that $\hat{A}(s^{l_2}, w^2) = \hat{A}(s^{l_3}, w^2)$. Since $s^v \in \hat{A}(s^{l_2}, w^2)$ (see (*2 c))

$$s^v \in \hat{A}(s^{l_3}, w^2) \tag{*3 c}$$

follows.

We can now apply (2.3.29) to the first two summands in (*3 b).
The inequality

$$\hat{c}\left(s^{l_3}, w^2, s^v\right) + \hat{c}\left(s^f, w^1, s^{l_3}\right) \leq \hat{c}\left(s^f, w^1, s^v\right) \tag{*3 d}$$

then follows.

Since $s^v \in \hat{A}(s^f, w^1)$ we can also apply (2.3.29) to the left side of (*3 d):

$$\hat{c}\left(s^f, w^1, s^v\right) \leq \hat{c}\left(s^f, w^1, s^v\right).$$

Hence equal signs are only possible in the last two inequalities.
This means

$$\hat{c}\left(s^{l_3}, w^2, s^v\right) + \hat{c}\left(s^f, w^1, s^{l_3}\right) = \hat{c}\left(s^f, w^1, s^v\right). \tag{*3 e}$$

We can iteratively pursue this procedural method until:

$$s^v \in \hat{A}(s^{l_j}, w^2) \tag{*j c}$$

$$\hat{c}\left(s^{l_j}, w^2, s^v\right) + \hat{c}\left(s^f, w^1, s^{l_j}\right) = \hat{c}\left(s^f, w^1, s^v\right) \tag{*j e}$$

(where $s^{l_j} \in \hat{\hat{A}}(s^f, w^1)$).

Thus, the assertion of the lemma is shown for $l = l_j$. ■

The statement of the following theorem seems plausible, however, the corresponding proof is not trivial.

**Theorem 2.3.13.** *Let $(DA\bar{P}a)$ or $(DA\bar{P}b)$ be a DA model for which the distance properties (2.3.28) and (2.3.29) are additionally satisfied. In addition let the condition (2.3.30) be valid.*

*Then the minimum will not increase, when smaller DA decision sets $\hat{\hat{A}}(s,w)$ are used instead of $\hat{A}(s,w)$.*

***Proof.*** Let $s_1(=\bar{s}_1)$ be some initial state. Furthermore let any sequences $w_t, t = 1,2,\ldots$ with $w_t \in B$ and $\bar{s}_t, t = 2,3\ldots$ with $\bar{s}_{t+1} \in \hat{A}(\bar{s}_t, w_t)$ for $t = 1,2,\ldots$ be given.

We will construct a sequence $s_t, t = 2,3,\ldots$ with $s_{t+1} \in \hat{\hat{A}}(s_t, w_t)$ such that

$$\sum_{t'=1}^{t} \hat{c}(\bar{s}_{t'}, w_{t'}, \bar{s}_{t'+1}) \geq \sum_{t'=1}^{t} \hat{c}(s_{t'}, w_{t'}, s_{t'+1}) \tag{2.3.31}$$

for $t = 1,2,\ldots$.

(It is possible that the original policy is not stationary. But it is well known that in the case of an infinite horizon an optimal stationary policy with the same set of decision spaces also exists.)

Now, we sequentially construct $s_{t+1} \in \hat{\hat{A}}(s_t, w_t)$ by

$$s_{t+1} = \begin{cases} \bar{s}_{t+1} & \text{if } \bar{s}_{t+1} \in \hat{\hat{A}}(s_t, w_t), \\ s_{t+1} \in \hat{\hat{A}}(s_t, w_t): & \\ \quad \hat{c}(s_{t+1}, w_{t+1}, \bar{s}_{t+1}) + \hat{c}(s_t, w_t, s_{t+1}) = \hat{c}(s_t, w_t, \bar{s}_{t+1}) & \\ \quad \text{with } \bar{s}_{t+1} \in \hat{A}(s_{t+1}, w_{t+1}) & \\ & \text{if } \bar{s}_{t+1} \notin \hat{\hat{A}}(s_t, w_t). \end{cases} \tag{2.3.32}$$

Such a $s_{t+1} \in \hat{\hat{A}}(s_t, w_t)$ exists in the second case according to Lemma 2.3.12.

(2.3.30) and $s_1 = \bar{s}_1$ yield the identity of the sets

$$\hat{A}(\bar{s}_t, w_t) = \hat{A}(s_t, w_t) \text{ for t = 1, 2, ... .}$$

We show that $s_t, t = 1, 2, \ldots$ fulfil (2.3.31) by means of mathematical induction:

Obviously the inequality

$$\hat{c}\left(s_1, w_1, \bar{s}_2\right) \geq \hat{c}\left(s_1, w_1, s_2\right)$$

is valid (see (2.3.32)).

Now, we assume that the inequality (2.3.31) is correct for $1, 2, \ldots, t$.

In the case that $\bar{s}_{t+1} = s_{t+1}$, the inequality (2.3.31) for $t+1$ follows from

$$\hat{c}\left(\bar{s}_{t+1}\left(= s_{t+1}\right), w_{t+1}, \bar{s}_{t+2}\right) \geq c\left(s_{t+1}, w_{t+1}, s_{t+2}\right),$$

compare (2.3.32).

Finally, we consider cases with

$$\bar{s}_{t_1} = s_{t_1},\ t_1 < t+1$$

and

$$\bar{s}_{t''} \neq s_{t''},\ t_1 < t'' \leq t+1.$$

In the following computations we alternately use **(2.3.32)** and the <u>triangle-inequality</u>:

$$\begin{aligned}
&\hat{\mathbf{c}}(\mathbf{s_{t_1}}, \mathbf{w_{t_1}}, \bar{\mathbf{s}}_{\mathbf{t_1+1}}) + \hat{c}(\bar{s}_{t_1+1}, w_{t_1+1}, \bar{s}_{t_1+2}) + \hat{c}(\bar{s}_{t_1+2}, w_{t_1+2}, \bar{s}_{t_1+3}) + \cdots + \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \hat{c}(\bar{s}_{t+1}, w_{t+1}, \bar{s}_{t+2}) \\
&= \underline{\hat{\mathbf{c}}(\mathbf{s_{t_1+1}}, \mathbf{w_{t_1+1}}, \bar{\mathbf{s}}_{\mathbf{t_1+1}})} + \hat{\mathbf{c}}(\mathbf{s_{t_1}}, \mathbf{w_{t_1}}, \mathbf{s_{t_1+1}}) + \underline{\hat{c}(\bar{s}_{t_1+1}, w_{t_1+1}, \bar{s}_{t_1+2})} \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad + \hat{c}(\bar{s}_{t_1+2}, w_{t_1+2}, \bar{s}_{t_1+3}) + \cdots + \hat{c}(\bar{s}_{t+1}, w_{t+1}, \bar{s}_{t+2}) \\
&\geq \underline{\hat{\mathbf{c}}(\mathbf{s_{t_1+1}}, \mathbf{w_{t_1+1}}, \bar{\mathbf{s}}_{\mathbf{t_1+2}})} + \hat{c}(s_{t_1}, w_{t_1}, s_{t_1+1}) + \hat{c}(s_{t_1+2}, w_{t_1+2}, s_{t_1+3}) + \cdots + \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \hat{c}(\bar{s}_{t+1}, w_{t+1}, \bar{s}_{t+2}) \\
&\geq \underline{\hat{\mathbf{c}}(\mathbf{s_{t_1+2}}, \mathbf{w_{t_1+2}}, \bar{\mathbf{s}}_{\mathbf{t_1+2}})} + \hat{\mathbf{c}}(\mathbf{s_{t_1+1}}, \mathbf{w_{t_1+1}}, \mathbf{s_{t_1+2}}) + \hat{c}(s_{t_1}, w_{t_1}, s_{t_1+1}) + \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad + \underline{\hat{c}(\bar{s}_{t_1+2}, w_{t_1+2}, \bar{s}_{t_1+3})} + \cdots + \hat{c}(\bar{s}_{t+1}, w_{t+1}, \bar{s}_{t+2})
\end{aligned}$$

$$\geq \underline{\hat{\mathbf{c}}(\mathbf{s_{t_1+2}}, \mathbf{w_{t_1+2}}, \bar{\mathbf{s}}_{\mathbf{t_1+3}})} + \hat{c}(s_{t_1+1}, w_{t_1+1}, s_{t_1+2}) + \hat{c}(s_{t_1}, w_{t_1}, s_{t_1+1}) + \cdots + \hat{c}(\bar{s}_{t+1}, w_{t+1}, \bar{s}_{t+2})$$

$$\vdots$$

$$\geq \underline{\hat{\mathbf{c}}(\mathbf{s_{t+1}}, \mathbf{w_{t+1}}, \bar{\mathbf{s}}_{\mathbf{t+1}})} + \hat{\mathbf{c}}(\mathbf{s_t}, \mathbf{w_t}, \mathbf{s_{t+1}}) + \cdots + \hat{c}(s_{t_1+1}, w_{t_1+1}, s_{t_1+2}) + \hat{c}(s_{t_1}, w_{t_1}, s_{t_1+1}) + \underline{\hat{c}(\bar{s}_{t+1}, w_{t+1}, \bar{s}_{t+2})}$$

$$\geq \underline{\hat{\mathbf{c}}(\mathbf{s_{t+1}}, \mathbf{w_{t+1}}, \bar{\mathbf{s}}_{\mathbf{t+2}})} + \hat{c}(s_t, w_t, s_{t+1}) + \cdots + \hat{c}(s_{t_1}, w_{t_1}, s_{t_1+1})$$

$$\geq \hat{\mathbf{c}}(\mathbf{s_{t+2}}, \mathbf{w_{t+2}}, \bar{\mathbf{s}}_{\mathbf{t+2}}) + \hat{\mathbf{c}}(\mathbf{s_{t+1}}, \mathbf{w_{t+1}}, \mathbf{s_{t+2}}) + \hat{c}(s_t, w_t, s_{t+1}) + \cdots + \hat{c}(s_{t_1}, w_{t_1}, s_{t_1+1})$$

$$\geq \hat{c}(s_{t+1}, w_{t+1}, s_{t+2}) + \hat{c}(s_t, w_t, s_{t+1}) + \cdots + \hat{c}(s_{t_1}, w_{t_1}, s_{t_1+1}).$$

Since (2.3.31) is valid for any $s_1$ and any sequence $w_t, t = 1, 2, \ldots$:

$$\underset{w_2,\ldots,w_N}{E}\left(\sum_{t=1}^{N} \hat{c}_t(s_t, w_t, s_{t+1}) \mid s_1, w_1\right) \leq \underset{w_2,\ldots,w_N}{E}\left(\sum_{t=1}^{N} \hat{c}_t(\bar{s}_t, w_t, \bar{s}_{t+1}) \mid s_1, w_1\right) \quad \text{for } (DA\bar{P}a)$$

$$\text{and } \overline{\lim_{n\to\infty}} \frac{1}{n} E\left(\sum_{t=1}^{n} \hat{c}(s_t, w_t, s_{t+1})\right) \leq \overline{\lim_{n\to\infty}} \frac{1}{n} E\left(\sum_{t=1}^{n} \hat{c}(\bar{s}_t, w_t, \bar{s}_{t+1})\right) \quad \text{for } (DA\bar{P}b)$$

follow. ■

If we want to apply Theorem 2.3.13, then we must first check condition (2.3.30). This is, however, laborious in general. Therefore, in the following Theorem 2.3.13a additional properties are assumed so that property (2.3.30) is satisfied in an effective way.

If $(DA\bar{P}b)$ is represented as a DA MDP$(N = \infty, S, A^M, P, \gamma)$ then it is useful to give:

**A Note on Transition Probabilities**

**Lemma 2.3.14.** *Let a DA MDP$(N = \infty, S, A^M, P, \gamma)$ be given where the underlying internal costs additionally satisfy the distance properties (2.3.28)*

*and (2.3.29).*

*Then the properties*

i) $s \in \hat{A}(s,w) \Rightarrow \hat{\hat{A}}(s,w) = \{s\}$,

ii) $p(s|s,d) = \sum_{w:s\in\hat{A}(s,w)} q(w)$ *for* $d$ *with*

$$\hat{d}(s',w') \in \hat{\hat{A}}(s',w') \;\forall\, s' \in S, w' \in B$$

*(this means* $p(s|s,d) =: p(s|s)$ *do not depend on* $d$*)*

*are valid.*

***Proof.***

i) Suppose that $s \in \hat{A}(s,w)$ and $s' \in \hat{A}(s,w)$ with $s' \neq s$.

For $s^f = s^l = s,\ s^v = s',\ w^1 = w$ and $w^2 \in B$ with $s' \in \hat{A}(s,w^2)$ the triangle-inequality (2.3.29) has the representation

$$\hat{c}\left(s,w^2,s'\right) + \hat{c}\left(s,w,s\right) \geq \hat{c}\left(s,w,s'\right). \tag{2.3.33}$$

$\hat{c}(s,w^2,s') + \hat{c}(s,w,s) = \hat{c}(s,w,s')$ follows from $\hat{c}(s,w,s) = 0$ (see (2.3.28)). According to Definition 2.3.4 $s' \notin \hat{\hat{A}}(s,w)$ and hence $\{s\} = \hat{\hat{A}}(s,w)$.

ii) Equation (2.3.7) and property i) yield

$$p(s|s,d) = \sum_{w:s=\hat{d}(s,w)} q(w) = \sum_{w:s\in\hat{A}(s,w)} q(w) =: p(s|s).$$

■

We now consider the distance properties, the smaller DA decision set and Theorem 2.3.13 under the additional assumptions:

$$\hat{c}(s^f,w,s^l)(=\hat{c}(s^f,s^l)) \text{ do not depend on } w \qquad \text{(see (2.3.12))},$$
$$\hat{A}(s,w)(=\hat{A}(w)) \text{ do not depend on s.} \tag{2.3.34}$$

It is reasonable for us to assume $\hat{A}(w) \neq \emptyset$ for all $w \in B$.

Obviously, (2.3.34) implies (2.3.30).

(SDDP problems will satisfy these conditions, see Chapter 3.)

Initially we specify the triangle-inequality under these assumptions.

$$\left.\begin{array}{r} \text{Let be } s^v \in \hat{A}(w^2), s^v \in \hat{A}(w^1),\ s^l \in \hat{A}(w^1) \\ \text{then } \hat{c}(s^l, s^v) + \hat{c}(s^f, s^l) \geq \hat{c}(s^f, s^v)\ \forall s^f \in S \end{array}\right\} \tag{2.3.35}$$

has to follow.

The representation of the smaller DA decision sets can be simplified in the following way (in particular, we set $w^2 = w^1$):

**Definition 2.3.4 a.** *Let $s^f \in S, w^1 \in B$ be given.*

$$\hat{\hat{A}}\left(s^f, w^1\right) = \{s^v \in \hat{A}\left(w^1\right) \mid \hat{c}\left(s^l, s^v\right) + \hat{c}\left(s^f, s^l\right) > \hat{c}\left(s^f, s^v\right)$$
$$\forall\ s^l \in \hat{A}\left(w^1\right) \text{ with } s^l \neq s^v \text{ and } s^l \neq s^f\}$$

*is called the smaller DA decision set of feasible states for a given state $s^f$ and a realized disturbance $w^1$ under the additional assumptions (2.3.12) and (2.3.34).*

Finally, we can replace (2.3.30) by (2.3.34) in Theorem 2.3.13:

**Theorem 2.3.13 a.** *Let $(DA\bar{P}a)$ or $(DA\bar{P}b)$ be DA models for which the distance properties (2.3.28) and (2.3.29) are additionally satisfied. Furthermore, let the condition (2.3.30) be valid.*

*Then the minimum will not increase, when smaller DA decision sets $\hat{\hat{A}}(s, w)$ are used instead of $\hat{A}(s, w)$.*

The following condition corresponds to (2.3.13).

$$\begin{array}{r} \hat{c}(s, s^{l_1}) = \cdots = \hat{c}(s, s^{l_v}) \quad \text{for any } s \in S, w \in B \\ \text{and } \{s^{l_1}, \cdots, s^{l_v}\} = \hat{\hat{A}}\left(s, w\right). \end{array} \tag{2.3.36}$$

Under this special assumption we show:

**Lemma 2.3.15.** *Let* $(DA\bar{P}a)$ *or* $(DA\bar{P}b)$ *be DA models for which the distance properties (2.3.28) and (2.3.29) are additionally satisfied. Furthermore, let the conditions (2.3.12), (2.3.34) and (2.3.36) be valid. Then the relations*

(i) $\hat{c}(s^f, s^l) < \hat{c}(s^f, \bar{s}^l)$ *for any* $w \in B$, $s^l \in \hat{\hat{A}}(s^f, w)$ *and* $\bar{s}^l \in \hat{A}(w)$ *however* $\bar{s}^l \notin \hat{\hat{A}}(s^f, w)$,

(ii) $c(s^f, s^l) = min\left\{\hat{c}(s^f, \bar{s}^l) \mid \bar{s}^l \in \hat{A}(w)\right\} \ \forall\, s^l \in \hat{\hat{A}}(s^f, w)$,

(iii) $\gamma(s, d)$ *do not depend on* $d$ *for* $d$ *with* $\hat{d}(s', w') \in \hat{\hat{A}}(s', w') \ \forall s' \in S, \ \forall w' \in B$

*hold.*

***Proof.***

(i): Assumption: $\hat{c}(s^f, \bar{s}^l) \leq \hat{c}(s^f, s^l)$ for $\bar{s}^l \notin \hat{\hat{A}}(s^f, w)$ and $s^l \in \hat{\hat{A}}(s^f, w)$.

According to Lemma 2.3.11(ii), (iii)

$$\bar{s} \in \hat{\hat{A}}(s^f, w) \ \text{ with } \ \hat{c}(s^f, \bar{s}) = \min_{s' \in \hat{A}(w)} \hat{c}(s^f, s') \ \text{ exists and}$$

furthermore, $\hat{c}(s^f, \bar{s}) < \hat{c}(s^f, \bar{s}^l)$ follows.

$\hat{c}(s^f, \bar{s}) < (\hat{c}(s^f, \bar{s}^l) \leq)\ \hat{c}(s^f, s^l)$, contradicting (2.3.36).

(ii): (ii) follows from (i) and (2.3.36).

(iii): Substituting (2.3.36) into the first equation of (2.3.8) yields (iii).

■

In [24] surrogate problems are given (these surrogate problems are a kind of two-stage-problems), which can be used as approximate solutions for the DA MDP with distance properties. These surrogate problems are to be used, above all, if the state spaces of the MDPS are very immense.

#### 2.3.3.2 The Dominant Policy

In this Section we consider "dominant policies" of MDPs. We will see that it is easy to analyze MDPs which are based on DA models in the case that dominant policies exist.

The dominance of Markov chains can be found in Daley 68 (see [10]).

We can apply this denotation to Markov chains which correspond to policies of MDPs. However, if we want to transfer this denotation to MDPs themselves, then convenient properties are also required for the corresponding average (one-step) reward functions (Puterman [31], Theorem 8.11.3 and Hildenbrandt [22], Theorem 3.2 or [20]).

Therefore, the dominance of policies of MDPs includes numerous strong conditions. Hence, the question is: can we find (useful) MDPs which fulfil the conditions of dominance?

Specific equipment replacement models with dominant policies can be found in Puterman [31]. Although in these models only two different decisions are possible (see [31], Example 8.11.1, Section 6.10.4. and Section 8.10.4).

Due to there decision structures the chance of finding MDPs with more than two decisions which fulfil the conditions of dominance is better for MDPs which are based on DA models (compare the Sections 2.3.2.1 and 2.3.2.2).

In Section 4.6.2.2 we will characterize SDDP problems with optimal dominant policies.

For other SDDP problems we will consider the interesting effect that conditions of dominance are only partially violated. It then seems possible to modify these conditions of dominance.

Since in Section 3.3 certain SDDP problems will be reduced, we now include such a possibility in the following definitions and theorems.

**Definition 2.3.5.** *Let a MDP*$(N = \infty, S, A^M, P, \gamma)$ *with states* $s^1, \cdots, s^m$ *be given and let* $\bar{d} \in A^M$ *be a (stationary) policy for which the stationary distribution* $p^{\bar{d},\infty}$ *exists. Furthermore, the following conditions are assumed to be fulfilled in relation to the sets of indices*

$I_v = \{h_{v-1}+1, h_{v-1}+2, \cdots, h_v\},\ v = 1, 2, \cdots, r$ *where*
$h_0 = 0,\ h_{v-1} < h_v,\ h_r = m$ :

(C1) *(domination)*

$$\sum_{v=1}^{\bar{r}} \sum_{l \in I_v} p_{h_1 l}^{\bar{d}} \geq \sum_{v=1}^{\bar{r}} \sum_{l \in I_v} p_{h_2 l}^{\bar{d}} \geq \cdots \geq \sum_{v=1}^{\bar{r}} \sum_{l \in I_v} p_{h_r l}^{\bar{d}}, \quad \forall\, \bar{r} = 1, \cdots, r,$$

(C2) *(reduction)*

$$\sum_{\bar{y} \in I_y} p_{h_{q-1}+1\ \bar{y}}^{\bar{d}} = \sum_{\bar{y} \in I_y} p_{h_{q-1}+2\ \bar{y}}^{\bar{d}} = \cdots = \sum_{\bar{y} \in I_y} p_{h_q\ \bar{y}}^{\bar{d}},$$
$$\forall\, q = 1, \cdots, r\ \forall\, y = 1, \cdots, r,$$

(C3) *(comparison)*

$$\sum_{v=1}^{\bar{r}} \sum_{l \in I_v} p_{\bar{y} l}^{\bar{d}} \leq \sum_{v=1}^{\bar{r}} \sum_{l \in I_v} p_{\bar{y} l}^{d} \quad \forall\, \bar{y} \in I_y,\ \forall\, \bar{r} = 1, \cdots, r$$
$$\forall\, y = 1, \cdots, r \quad \textit{and} \quad \forall\, d \in A^M.$$

*Additionally, let the one-step reward functions belonging to $\bar{d}$ fulfil the conditions*

(Cr1) $\gamma(s^{h_1}, \bar{d}) \geq \gamma(s^{h_2}, \bar{d}) \geq \cdots \geq \gamma(s^{h_r}, \bar{d})$,

(Cr2) $\gamma(s^{h_{v-1}+1}, \bar{d}) = \gamma(s^{h_{v-2}+2}, \bar{d}) = \cdots = \gamma(s^{h_v}, \bar{d})\ \forall\, v = 1, \cdots, r$,

(Cr3) $\gamma(s^l, \bar{d}) \ \leq\ \gamma(s^l, d)\ \forall\, l = 1, \cdots, m$ *and* $\forall\, d \in A^M$.

*Then $\bar{d} \in A^M$ is called a dominant policy (in relation to the sets of indices).*

**Remarks 2.3.3.** *If (C2) and (Cr2) are fulfilled in relation to the sets of indices $I_v$, this is used for the reduction of the corresponding MDPs. Of course, $I_v = \{v\},\ v = 1, \cdots, m$ is also possible in Definition 2.3.5, in the following Theorem 2.3.17 (and so on).*

**Lemma 2.3.16.** *Let a MDP($N = \infty, S, A^M, P, \gamma$) with states $s^1, \cdots, s^m$ be given and let $\bar{d} \in A^M$ be a dominant policy in relation to the sets of indices*

$I_v = \{h_{v-1}+1, h_{v-1}+2, \cdots, h_v\},\ v = 1, 2, \cdots, r$ *where*

$h_0 = 0,\ h_{v-1} < h_v,\ h_r = m.$

*(i) Then (C1), (C2) and (C3) (from Definition (2.3.5)) are valid for each of the powers $(P^{\bar{d}})^t$ of $P^{\bar{d}}$ and $(P^d)^t$ of $P^d$ $(t = 1, 2, \cdots)$.*
*(Here, $P^{\bar{d}}$ and $P^d$ are matrices of transition probabilities for $\bar{d}$ and $d$, respectively.)*

*(ii) Additionally, the condition*

$$p^{\bar{d},\infty}_{h_{q-1}+1} = p^{\bar{d},\infty}_{h_{q-1}+2} = \cdots = p^{\bar{d},\infty}_{h_q}\ \forall\, q = 1, \cdots, r$$

*which corresponds to (C2) is valid for the stationary distribution $p^{\bar{d},\infty}$.*

(Refer to the proofs of Lemma 3.14 and Theorem 3.17(a) in [22] for (i) and the proof of Corollary 3.15 for (ii).)

A dominant policy is always an optimal policy:

**Theorem 2.3.17.** . *Let (2.3.3) be valid for all stationary policies of a given MDP($N = \infty, S, A^M, P, \gamma$). In addition, let $\bar{d} \in A^M$ be a dominant policy in relation to the sets of indices*

$I_v = \{h_{v-1}+1, h_{v-1}+2, \cdots, h_v\},\ v = 1, 2, \cdots, r$ *where*

$h_0 = 0,\ h_{v-1} < h_v,\ h_r = m.$

*Then the dominant policy $\bar{d}$ is an optimal policy.*

(Refer to the proof of Theorem 3.17(b) in [22].)

In order to check whether a dominant policy exists for a given MDP($N = \infty, S, A^M, P, \gamma$) (where (2.3.3) is valid for all stationary policies), we first consider the average one-step reward functions in relation to condition (Cr3):

(In Algorithm 2.3.1 we do not include the possibility of reduction of MDPs.)

**Algorithm 2.3.1.** .

1. *If we can find $d^0$ such that*

$$\gamma(s^f, d^0) = min\{\gamma(s^f, d) \mid d \in A^M\} \;\; for\ any \;\; s^f \in S$$

*then condition (Cr3) of Definition 2.3.5 is fulfilled.*

2. *We then number the states in a new way such that*

$$\gamma(s^{\lambda_1}, d^0) \geq \gamma(s^{\lambda_2}, d^0) \geq \cdots \geq \gamma(s^{\lambda_m}, d^0), \tag{2.3.37}$$

$$\{\lambda_1, \cdots, \lambda_m\} = \{1, \cdots, m\} \quad \textit{(Condition (Cr1)!).}$$

3. *Lastly, we check the conditions (C1) and (C3) of Definition 2.3.5.*

   *Either not all conditions are valid or $d^0$ is an optimal dominant policy ($\bar{d} = d^0$).*

(Clearly, if equals signs exist in (2.3.37) then the numbering of the states is not unique in step 2.)

Now, let a Markov decision process DA MDP($N = \infty, S, A^M, P, \gamma$) be given, which results from a DA model (where (2.3.3) is valid for all stationary policies). On the whole we can investigate the existence of (and construct) a dominant policy as in Algorithm 2.3.1. In step 1 we determine $d^0$ by means of the internal costs:

(The possibility of reduction of MDPs is included in a supplement after Algorithm 2.3.1a.)

**Algorithm 2.3.1a**

1. *We compute $d^0$ such that*

$$\left.\begin{aligned} &\hat{d}^0(s^f, w) = s^l, \quad where \\ &\hat{c}(s^f, w, s^l) = min\{\hat{c}(s^f, w, s^{l'}) | s^{l'} \in \hat{A}(s^f, w)\} \\ &\qquad\qquad\qquad\qquad for\ f = 1, \cdots, m,\ w \in B \end{aligned}\right\} \tag{2.3.38}$$

   *and $\gamma(s^f, d^0)$ according to the first equation of (2.3.8).*

   *Then, condition (Cr3) of Definition 2.3.5 is fulfilled (see the first equations of (2.3.8) and (2.3.17)).*

2. *as in Algorithm 2.3.1.*

3. *as in Algorithm 2.3.1.*

*We must, however, take into consideration that $d^0$ does not have to be unique and that equals signs can also be found in (2.3.37).*

*Therefore we supplement:*

(a) *If $d^0$ is not unique in the following way:*

$$\left.\begin{array}{l} \hat{c}(s^f, w, s^{l_1}) = \hat{c}(s^f, w, s^{l_2}) = min\ \{\hat{c}(s^f, w, s^{l'}) \mid s^{l'} \in \hat{A}(s^f, w)\} \\ and\ \gamma(s^{l_1}, d^0) < \gamma(s^{l_2}, d^0) \\ then\ d^0(s^f)\ with\ \hat{d}^0(s^f, w) = s^{l_1}\ must\ be\ used. \end{array}\right\} \tag{2.3.39}$$

*This is necessary for the validity of (Cr3).*

(b) *If equals signs are in (2.3.37) the possibility of formation of sets of indices $I_v$ must be considered. That means that several cases have to be taken into account when checking the conditions of dominance. (See also papers pertaining to lumpings of Markov chains such as [6]). ((b) additionally requires that $d^0$ is not unique.)*

**Lemma 2.3.18.** *Let a DA MDP$(N = \infty, S, A^M, P, \gamma)$ be given.*

*Then condition (Cr3) is valid only for decision functions $d^0$ which are computed as in (2.3.38).*

***Proof.***

1. Let $d^0$ be computed as in (2.3.38). From the first equation of (2.3.8) it follows that (Cr3) is valid.

2. Let $(Cr3)$ be valid.

   Assumption: $\hat{d}^0(s^f, w) = s^v$ and $\exists\ s^l \in \hat{A}(s^f, w) : \hat{c}(s^f, w, s^l) < \hat{c}(s^f, w, s^v)$.

   $\bar{d}$ with

$$\hat{\bar{d}}(s', w') := \begin{cases} \hat{d}^0(s', w') & \text{if } (s', w') \neq (s^f, w), \\ s^l & \text{if } (s', w') = (s^f, w) \end{cases}$$

   yields $\gamma(s^f, \bar{d}) < \gamma(s^f, d^0)$ according to (2.3.17). This inequality is a contradiction of $(Cr3)$. ■

**Remarks 2.3.4.** *For a DA MDP($N = \infty, S, A^M, P, \gamma$) the condition (C3) includes (single) decisions $\hat{\bar{d}}$ for the feasible states whose indices are elements of indices sets with the largest possible indices.*

Hence, an almost-partial order of the states follows:

**Lemma 2.3.19.** *Let a DA MDP($N = \infty, S, A^M, P, \gamma$) with states $s^1, \cdots, s^m$ be given and let the (stationary) policy $\bar{d} \in A^M$ satisfy the conditions (C2) and (C3) of Definition 2.3.5 in relation to the following sets of indices*

$I_v = \{h_{v-1} + 1, h_{v-1} + 2, \cdots, h_v\},\ v = 1, 2, \cdots, r$ *where*

$h_0 = 0,\ h_{v-1} < h_v,\ h_r = m.$

*(i) If $\hat{\bar{d}}(s^f, w) = s^l$ for $s^l \in \hat{A}(s^f, w)$ with $l \in I_v$*
*then no $s^{l'} \in \hat{A}(s^f, w)$ with $l' \in I_{v'}$ and $v' > v$ exists.*

*(ii) $\bar{d}$ implies an almost-partial order $<$ of the states in the way*

$$\left\{s^l < s^f,\ l \in I_{v(l)},\ f \in I_{v(f)}\right\} :=$$
$$\left\{\exists\, s^y \in S,\ w \in B \text{ and } s^{l'} \in \hat{A}(s^y, w),\ l' \in I_{v(l)},\ s^{f'} \in \hat{A}(s^y, w),\ \ f' \in I_{v(f)} \right.$$
$$\left. \text{so that } \hat{\bar{d}}(s^y, w) = s^{f'}.\right\}$$

***Proof.***

(i) Assumption: $\exists s^{l'} \in \hat{A}(s^f, w)$ with $l' \in I_{v'}$ and $v' > v$.

We set $\bar{d}'$ with:

$$\hat{\bar{d}}'(s', w') := \begin{cases} \hat{\bar{d}}(s', w') & \text{if } (s', w') \neq (s^f, w), \\ s^{l'} & \text{if } (s', w') = (s^f, w). \end{cases}$$

From Lemma 2.3.3 IIa)

$$\sum_{\bar{v}=1}^{v} \sum_{\bar{l} \in I_{\bar{v}}} p_{f\bar{l}}^{\bar{d}'} = \sum_{\bar{v}=1}^{v} \sum_{\bar{l} \in I_{\bar{v}}} p_{f\bar{l}}^{\bar{d}} - q(w) < \sum_{\bar{v}=1}^{v} \sum_{\bar{l} \in I_{\bar{v}}} p_{f\bar{l}}^{\bar{d}}$$

follows contradicting (C3).

(ii) (See Definition 2.3.2 for the almost-partial order.)

Asymmetry:

Let $s^l < s^f$ as in (ii) and (w.l.o.g) $v(l) < v(f)$ .

Assumption: $s^l > s^f$. This means

$$\exists\, s^{y'} \in S, w' \in B \text{ and } s^{l''} \in \hat{A}(s^{y'}, w'), l'' \in I_{v(l)}, s^{f''} \in \hat{A}(s^{y'}, w'), f'' \in I_{v(f)}$$

$$\text{so that } \hat{\bar{d}}(s^{y'}, w') = s^{l''}.$$

Analogously to the proof of (i), we set $\bar{d}'$ with:

$$\hat{\bar{d}}'(s'', w'') := \begin{cases} \hat{\bar{d}}(s'', w'') & \text{if } (s'', w'') \neq (s^{y'}, w'), \\ s^{f''} & \text{if } (s'', w'') = (s^{y'}, w') \end{cases}$$

and from Lemma 2.3.3 IIa)

$$\sum_{\bar{v}=1}^{v(l)} \sum_{\bar{l} \in I_{\bar{v}}} p^{\bar{d}'}_{y'\bar{l}} = \sum_{\bar{v}=1}^{v(l)} \sum_{\bar{l} \in I_{\bar{v}}} p^{\bar{d}}_{y'\bar{l}} - q(w') < \sum_{\bar{v}=1}^{v(l)} \sum_{\bar{l} \in I_{\bar{v}}} p^{\bar{d}}_{y'\bar{l}}$$

follows contradicting (C3).

Almost-transitivity:

Let $s^l < s^y$ and $s^y < s^f$.

$s^l \not> s^f$ can be proven by contradiction, similar to the proof of asymmetry.

■

**Remarks 2.3.5.** .

a) *For a DA MDP($N = \infty, S, A^M, P, \gamma$) it is not difficult to verify whether a dominant policy exists. In particular, in step 3 of Algorithm 2.3.1, the extensive condition (C3) is easily analyzed for a DA MDP (see Lemma 2.3.19).*

b) *Lemma 2.3.19(ii) and Theorem 2.3.8(i) imply that dominant policies could exist particularly for DA MDPs if internal costs with special properties (for example (2.3.13)) are underlain.*

*c) MDPs (with more than two decisions) which are not based on DA models with optimal dominant policies are rarely found.*

### 2.3.4 Cost-Parametric Analysis of DA Markov Decision Processes

In the main part of this section we consider DA MDPs where the underlying internal costs and hence the average one-step reward functions depend linearly on one deterministic parameter.

We assume that the internal costs and the average one-step reward functions do not depend on the decisions in regard to the initial parameter (see (2.3.13) and (2.3.14) in Section 2.3.2).

In this case optimal decisions imply almost-partial orders of the states (compare Section 2.3.2.2). These almost-partial orders of the states mean that the complexity of computing optimal decisions can be reduced.

If the parameter increases, then in general the optimality criterion (2.3.22) is violated for single decisions.

Optimal decisions can therefore be purposefully computed for the increasing parameter (by means of policy iteration, for example). This also means that an adapted Howard algorithm is a greedy algorithm for cost-parametric DA Markov decision processes.

(Furthermore, a finite number of optimal policies exist for an infinite set of parameter values).

In a later part of this section conditions (AC1), (AC2) and (AC3) are given. If they are valid the complexity of computation of optimal decisions can be reduced further.

(A remaining problem is to what extent these conditions are fulfilled for (parametric) SDDP problems.)

On one hand as we carry out the investigation in this Section we will realize the solutions of the DA MDPs themselves (by a continuation of the solutions of the parametric DA MDPs) and the motivation of heuristics for DA MDPs.

On the other hand, we will work out specific properties in relation to the solution behavior of DA MDPs.

#### 2.3.4.1 Some Properties of Parametric Markov Decision Processes

Let $\vartheta \in \bar{I} \subseteq \mathbb{R}$ be a deterministic parameter, where $\bar{I}$ is one of the following intervals $[\vartheta_1, \vartheta_2]$ or $[\vartheta_1, \infty)$ or $(-\infty, \vartheta_2]$ or $(-\infty, \infty)$ (with $\vartheta_1 \in \mathbb{R}, \vartheta_2 \in \mathbb{R}$).

Now, we introduce a set of MDPs of the type MDP$(N = \infty, S, A^M, P, \gamma)$ (see Section 2.3.1), where

$$\begin{gathered} S, s \text{ and } A^M, d \text{ do not depend on } \vartheta, \\ \gamma(s, d, \vartheta) \text{ and } p^d_{fl}(\vartheta) \text{ are continuous at } \vartheta \text{ for each} \\ s \in S, f \in \{1, \cdots, m\}, l \in \{1, \cdots, m\}, d \in A^M. \end{gathered} \tag{2.3.40}$$

For the set of such MDPs (which satisfy (2.3.40)) we use the notation

$$MDP_c(N = \infty, S, A^M, P(\vartheta), \gamma(\vartheta)),$$

where the subscript index $c$ means that $P^d(\vartheta)$ and $\gamma^d(\vartheta)$ are "**c**ontinuous" at $\vartheta$.

**Lemma 2.3.20.** *Let a $MDP_c(N = \infty, S, A^M, P(\vartheta), \gamma(\vartheta))$ on $\bar{I}$ be given. In addition, let $p^d_{fl}(\vartheta) \neq 0 \; \forall \; f, l \in \{1, \ldots, m\}, \; \forall \; d \in A^M, \forall \; \vartheta \in \bar{I}$.*

*(i) The average expected cost per stage $g^d(\vartheta)$ and the other solutions $\nu^d(\vartheta)$ of the Poisson equations*

$$g(\vartheta)\epsilon + \nu(\vartheta) = P^d(\vartheta)\nu(\vartheta) + \gamma^d(\vartheta), \quad \epsilon = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$$

*are continuous at $\vartheta$, where (w. l. o. g.) $\nu^d_m(\vartheta) = 0 \; \forall \; \vartheta \in \bar{I}$ (see (2.3.5)).*

*(ii) If $d$ is an optimal decision on $(\vartheta', \vartheta'') \subseteq \bar{I}$, then $d$ is also an optimal decision on $[\vartheta', \vartheta'']$.*

***Proof.***

(i) If $\nu_m^d(\vartheta)(=0)$ are fixed then the remaining equation systems have unique solutions and fundamental rules of arithmetic can therefore be used to solve these equation systems. Hence, the solutions are continuous at $\vartheta$ (if $P^d(\vartheta)$ and $\gamma^d(\vartheta)$ are continuous at $\vartheta$).

(ii) $$\triangle H^d(s^f,\bar{d},\vartheta) := \sum_{l=1}^{m} p_{fl}^{\bar{d}}\nu_l^d(\vartheta)+\gamma(s^f,\bar{d},\vartheta)-(\sum_{l=1}^{m} p_{fl}^{d}\nu_l^d(\vartheta)+\gamma(s^f,d,\vartheta)).$$

According to (2.3.19) and (2.3.19 a), respectively,
a decision $d$ is optimal if and only if

$$\triangle H^d(s^f,\bar{d},\vartheta) \geq 0 \;\; \forall\, s^f \in S,\; \forall\, \bar{d} \in A^M.$$

(i) implies that $\triangle H^d(s^f,\bar{d},\vartheta)$ are also continuous at $\vartheta$.

If $\triangle H^d(s^f,\bar{d},\vartheta) \geq 0 \;\; \forall\, s^f \in S,\; \forall\, \bar{d} \in A^M$ and $\forall\, \vartheta \in (\vartheta',\vartheta'')$ (optimality criterion) then

$$\triangle H^d(s^f,\bar{d},\vartheta') \geq 0 \;\text{ and }\; \triangle H^d(s^f,\bar{d},\vartheta'') \geq 0 \;\text{ follows}$$

since $\triangle H^d(s^f,\bar{d},\vartheta)$ are continuous at $\vartheta$.

■

When the optimal decisions change at $\vartheta_0$ the solutions of the corresponding Poisson equations fulfil the following property:

**Theorem 2.3.21.** *Let a $MDP_c(N=\infty, S, A^M, P(\vartheta), \gamma(\vartheta))$ on $\bar{I}$ be given. In addition, let $p_{fl}^d(\vartheta) \neq 0\; \forall\, f,l \in \{1,\ldots,m\},\; \forall\, d \in A^M, \forall\, \vartheta \in \bar{I}$ be given.*

*If $d^{*'}$ is an optimal decision on $[\vartheta',\vartheta_0]$ and $d^{*''}$ is an optimal decision on $\vartheta \in [\vartheta_0,\vartheta'']$ where $\vartheta' \leq \vartheta_0 \leq \vartheta''$ then*

$$\nu^{*'}(\vartheta_0) = \nu^{*''}(\vartheta_0)$$

*for the solutions $(g^{*'}(\vartheta),\; \nu^{*'}(\vartheta))$ with $\nu_m^{*'}(\vartheta) = 0$ and $(g^{*''}(\vartheta), \nu^{*''}(\vartheta))$ with $\nu_m^{*''}(\vartheta) = 0$ of the corresponding Poisson equations.*

The theorem follows immediately from Lemma 2.3.1.

For further considerations we assume that only the average one-step reward functions $\gamma$ depend on the parameter $\vartheta$ and that $\gamma(\vartheta)$ are linear functions at $\vartheta$.

Furthermore, $\gamma$ also depend on (fixed) $\xi$.
(For DA MDPs $\xi$ will be related to the internal costs):

$(LPC_l)$ [11] *Let* $\bar{I} = [0, \infty)$, $\{\xi_0,\ \xi\} \subseteq \mathbb{R}^{\zeta}_{+}$, [12] $\Theta_{\xi}(\vartheta) := \xi_0 + \vartheta\xi$ *for* $\vartheta \in \bar{I}$.

*Additionally, let* $\gamma$ *be a linear function at* $\vartheta$*:*

$$\gamma^d(\vartheta) = \gamma^d(\Theta_{\xi}(\vartheta)) = \gamma^d(\xi_0 + \vartheta\xi) = \gamma^d(\xi_0) + \vartheta\gamma^d(\xi)\ (\forall\ d \in A^M). \tag{2.3.41}$$

Since $\gamma^d \in \mathbb{R}^m_+$ (see Section 2.3), $\gamma^d(\vartheta)$ is an increasing function according to (2.3.41).

For the set of corresponding cost-parametric MDPs (which satisfy $(LPC_l)$) we use the notation

$MDP_l(N = \infty, S, A, P, \gamma(\xi_0) + \vartheta\gamma(\xi))$,

where the subscript index $l$ means that $\gamma^d(\vartheta)$ are "**l**inear" at $\vartheta$ $(\forall\ d \in A^M)$.

**Lemma 2.3.22.** *Let a* $MDP_l(N = \infty, S, A, P, \gamma(\xi_0) + \vartheta\gamma(\xi))$ *on* $\bar{I} = [0, \infty)$ *be given. In addition, let (2.3.3) be valid for all stationary policies.*

*Then the average expected cost per stage* $g^d(\xi_0 + \vartheta\xi)$ *and the other solutions* $\nu^d(\xi_0 + \vartheta\xi)$ *of the Poisson equations*

$$g(\xi_0 + \vartheta\xi)\epsilon + \nu(\xi_0 + \vartheta\xi) = P^d\nu(\xi_0 + \vartheta\xi) + \gamma^d(\xi_0 + \vartheta\xi), \quad \epsilon = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$$

*where (w. l. o. g.)* $\nu^d_m(\xi_0) = 0, \nu^d_m(\xi) = 0$

*are linear functions at* $\vartheta$*:*

$g^d(\xi_0 + \vartheta\xi) = g^d(\xi_0) + \vartheta g^d(\xi)$

$\nu^d_f(\xi_0 + \vartheta\xi) = \nu^d_f(\xi_0) + \vartheta\nu^d_f(\xi),\ f = 1, \cdots, m-1(, m)$,

*where* $g^d(\xi_0), \nu^d(\xi_0)$ *are solutions of the Poisson equation*

$$g(\xi_0)\epsilon + \nu(\xi_0) = P^d\nu(\xi_0) + \gamma^d(\xi_0), \quad \epsilon = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$$

*and* $g^d(\xi), \nu^d(\xi)$ *are solutions of the Poisson equation*

[11] **P**arametric **c**ase where $\gamma$ is a **l**inear function at $\vartheta$.
[12] Initially, the type of the set which contains $\xi_0$ and $\xi$ does not play a significant role.

$$g(\xi)\epsilon + \nu(\xi) = P^d\nu(\xi) + \gamma^d(\xi), \quad \epsilon = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}.$$

*$g^d(\xi_0 + \vartheta\xi)$ is an increasing function at $\vartheta$ and if $l$ with $\gamma_l^d(\xi) > 0$ exists then $g^d(\xi_0 + \vartheta\xi)$ is a strictly increasing function at $\vartheta$ (see (2.3.4)).*

Obviously, this Lemma follows from properties of linear equation systems in connection with fundamental rules of arithmetic.

**Lemma 2.3.23.** *Let a $MDP_l(N = \infty, S, A, P, \gamma(\xi_0) + \vartheta\gamma(\xi))$ on $\bar{I} = [0, \infty)$ be given. In addition, let (2.3.3) be valid for all stationary policies.*

*(i) If $d$ is an optimal decision at $\vartheta_0 \in I$ and*

$$\exists\, \vartheta' \neq \vartheta_0 \text{ with } \left\{ \begin{array}{l} \vartheta' < \vartheta_0, \\ \vartheta_0 < \vartheta' \end{array} \right\} : \ d \text{ is optimal at } \vartheta',$$

$$\text{then } d \text{ is optimal on } \left\{ \begin{array}{l} [\vartheta', \vartheta_0]\,, \\ [\vartheta_0, \vartheta'] \end{array} \right\}.$$

*(ii) An optimal decision $d^*$ at $\vartheta_0 = 0$ and an $\varepsilon > 0$ exist such that $d^*$ is also optimal on $[0,\ \varepsilon]$.*

***Proof.***

(i): From Lemma 2.3.22 it follows that $\triangle H^d(s^f, \bar{d}, \vartheta)$ ($s^f \in S,\ \bar{d} \in A^M$) are also linear functions at $\vartheta$.

Hence, $\triangle H^d(s^f, \bar{d}, \vartheta_0) \geq 0$ and $\triangle H^d(s^f, \bar{d}, \vartheta') \geq 0$ imply that the optimality criterion $\triangle H^d(s^f, \bar{d}, \vartheta) \geq 0$ is also fullfiled on the complete interval $\left\{ \begin{array}{l} [\vartheta', \vartheta_0]\,, \\ \text{or } [\vartheta_0, \vartheta'] \end{array} \right\}$.

(ii): From Lemma 2.3.22 it follows that

$\triangle H^d(s^f, \bar{d}, \vartheta) = \triangle H^d(s^f, \bar{d}, \xi_0) + \vartheta\ \triangle H^d(s^f, \bar{d}, \xi)$ with

$$\triangle H^d(s^f, \bar{d}, \xi_0) = \sum_{l=1}^{m} p_{fl}^{\bar{d}} \nu_l^d(\xi_0) + \gamma(s^f, \bar{d}, \xi_0) - (\sum_{l=1}^{m} p_{fl}^{d} \nu_l^d(\xi_0) + \gamma(s^f, d, \xi_0))$$

$$\triangle H^d(s^f, \bar{d}, \xi) = \sum_{l=1}^{m} p_{fl}^{\bar{d}} \nu_l^d(\xi) + \gamma(s^f, \bar{d}, \xi) - (\sum_{l=1}^{m} p_{fl}^{d} \nu_l^d(\xi) + \gamma(s^f, d, \xi)).$$

Since $A^M$ is a finite set of decision functions, a decision $d^*$ and an infinite sequence $\{\vartheta_i\}_{i=1,2,\cdots}$ with $0 < \vartheta_i \ \forall\ i = 1, 2, \cdots$ and $\lim\limits_{i\to\infty} \vartheta_i = 0$ exist such that $d^*$ is optimal at $\vartheta_i$ for $i = 1, 2, \cdots$.

Then, from $\triangle H^{d^*}(s^f, \bar{d}, \vartheta_i) = \triangle H^{d^*}(s^f, \bar{d}, \xi_0) + \vartheta_i \triangle H^{d^*}(s^f, \bar{d}, \xi) \geq 0$ for $i = 1, 2, \cdots$, the inequalities $\triangle H^{d^*}(s^f, \bar{d}, 0) = \triangle H^{d^*}(s^f, \bar{d}, \xi_0) \geq 0$ follow. According to (i) we can chose $\varepsilon = \vartheta_1$ in order to complete the proof for (ii).

■

Since $g^d(\xi_0 + \vartheta\xi)$ is an increasing function at $\vartheta$ for any $d$ (see Lemma 2.3.22) and the decision space is a finite set the following Lemma is valid (refer also to Lemma 2.3.20(ii) and Lemma 2.3.23):

**Lemma 2.3.24.** *Let a* $MDP_l(N = \infty, S, A^M, P, \gamma(\xi_0) + \vartheta\gamma(\xi))$ *on* $\bar{I} = [0, \infty)$ *be given. In addition, let (2.3.3) be valid for all stationary policies.*

*Then a finite sequence of real numbers* $0 = \bar{\vartheta}_1 < \bar{\vartheta}_2 < \cdots < \bar{\vartheta}_{y-1}$ *and a corresponding sequence of decisions* $d^{*1} \in A^M, \ldots, d^{*(y-1)} \in A^M$ *exist such that*

$d^{*f}$ *are optimal on* $I_f := [\bar{\vartheta}_f, \bar{\vartheta}_{f+1}]$ *for* $f = 1, 2, \cdots, y-2$

*and* $d^{*(y-1)}$ *is optimal on* $I_{y-1} = [\bar{\vartheta}_{y-1}, \infty)$.

*If* $d^{*f}$ *are optimal on* $I_f$ $(f < y-1)$ *however* $d^{*f}$ *are not optimal on* $I_{f+1} \backslash \{\bar{\vartheta}_{f+1}\}$ *then* $d^{*f}$ *are also not optimal on* $I_l$ *for* $l = f+1, f+2, \cdots, y-1$.

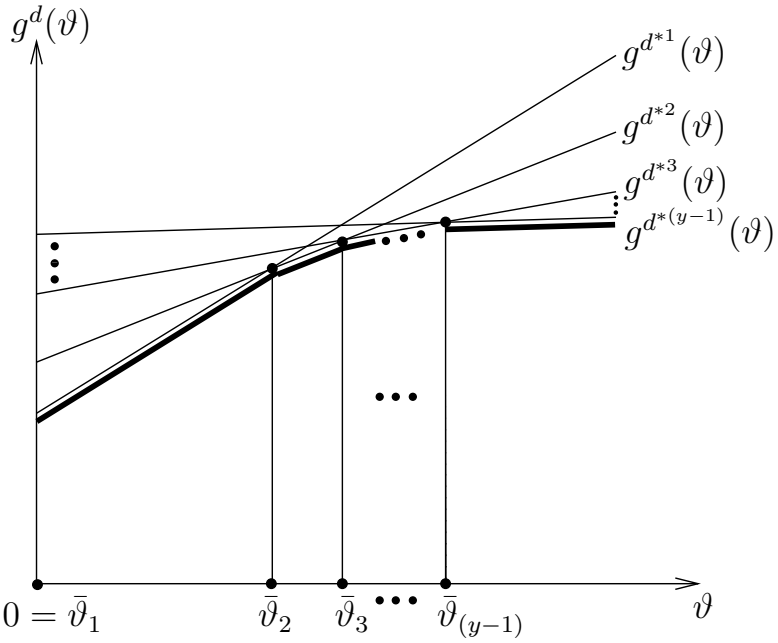



**Figure 2.3.2.**

#### 2.3.4.2 Cost-Parametric DA Markov Decision Processes

Cost-parametric DA MDPs mean that the underlying internal costs and thus the average one-step reward functions (see (2.3.8)) depend linearly on a deterministic parameter.

We also assume that internal costs and thus the average one-step reward functions do not depend on the decisions in regard to the initial parameter (refer to (2.3.13) and (2.3.14) in Section 2.3.2):

(LPC) *Let* $\bar{I} = [0, \infty)$, $\{\kappa_0,\ \kappa\} \subseteq \mathbb{R}^n_+ \times \mathbb{R}^{|B|}_+ \times \mathbb{R}^n_+$ *(or* $\{\kappa_0,\ \kappa\} \subseteq \mathbb{R}^n_+ \times \mathbb{R}^n_+$ *if (2.3.12) is valid),* $\vartheta \in \bar{I}$ *and* $\Theta_\kappa(\vartheta) := \kappa_0 + \vartheta\kappa$ *for* $\vartheta \in \bar{I}$.

*Furthermore, let the internal costs be linear functions at* $\vartheta$ *and let the internal costs not depend on the decisions in regard to the initial parameter (see (2.3.13):*

$$\hat{c}(s, w, s', \vartheta) = \hat{c}(s, w, s', \Theta_\kappa(\vartheta)) = \hat{c}(s, w, s', \kappa_0 + \vartheta\kappa)$$
$$= \hat{c}(s, w, s', \kappa_0) + \vartheta\ \hat{c}(s, w, s', \kappa) \text{ for any } s, s' \in S,\ w \in B, \vartheta \in \bar{I},$$

*thus*

$$\begin{aligned} &= \hat{c}(s, w, \kappa_0) + \vartheta\ \hat{c}(s, w, s', \kappa) \\ &\text{for any } s, s' \in S, w \in B, \text{ and given } \{\kappa_0,\ \kappa\}. \end{aligned} \tag{2.3.42}$$

(**Remark:** *In Section 3.5 the internal costs for SDDP problems have to be calculated as solutions of classical transportation problems with costs* $\kappa = (k_{ij})_{\substack{i=1,\dots,n \\ j=1,\dots,n}} \in \mathbb{R}^n_+ \times \mathbb{R}^n_+$.*)*

In order to calculate the average one-step reward functions according to (2.3.8) we set

$$\xi_0 = (\hat{c}(s,w,\kappa_0))_{\{s,s'\}\subseteq S, w\in B} \text{ and } \xi = (\hat{c}(s,w,s',\kappa))_{\{s,s'\}\subseteq S, w\in B}.$$
$$\text{(Or } \xi_0 = (\hat{c}(s,s',\kappa_0))_{\{s,s'\}\subseteq S} \text{ and } \xi = (\hat{c}(s,s',\kappa))_{\{s,s'\}\subseteq S} \tag{2.3.43}$$
$$\text{if (2.3.12) is valid.)}$$

According to the first equation of (2.3.8)

$$\gamma^d(s,\xi) \;=\; \sum_{s'\in S} \sum_{w:s'=\hat{d}(s,w)} \hat{c}(s,w,s',\kappa)\; q(w),$$
$$\gamma^d(s,\xi_0) \;=\; \sum_{w} \hat{c}(s,w,\kappa_0)\; q(w)$$

and

$$\gamma^d(s,\vartheta) := \gamma^d(s,\xi_0+\vartheta\xi) = \gamma^d(s,\xi_0) + \vartheta\,\gamma^d(s,\xi)$$

follow for any $s \in S,\ d \in A^M$ and $\vartheta \in [0,\infty)$.

Hence $\gamma^d(s,\vartheta)$ satisfy $(LPC_l)$, if $\hat{c}(s,w,s',\vartheta)$ satisfy (LPC).

For the set of corresponding cost-parametric DA MDPs (which satisfy (LPC) and (2.3.43)) we use the notation

$$DA\ MDP_l(N=\infty, S, A^M, P, \gamma(\vartheta)).$$

In relation to the initial parameter and the increasing parameter the following Lemma is valid for such DA MDPs:

**Lemma 2.3.25.** *Let a* $DA\ MDP_l(N=\infty, S, A^M, P, \gamma(\vartheta))$ *on* $\bar{I} = [0,\infty)$ *be given. And let (2.3.3) be valid for all stationary policies.*

*(i) If (2.3.43) is assumed, then* $\gamma^d(s,\xi_0)$ *do not depend on* $d$.

*(ii)* $\triangle H^d(s^f, w, s^{\bar{l}} - s^l, \vartheta) = \hat{c}(s^f, w, s^{\bar{l}}, \vartheta) - \hat{c}(s^f, w, s^l, \vartheta) + \nu^d_{\bar{l}}(\vartheta) - \nu^d_l(\vartheta)$ *are linear functions at* $\vartheta$ *for any* $\{s^f, s^{\bar{l}}, s^l\} \subseteq S, w \in B, d \in A^M$ *(*$\nu^d(\vartheta)$ *are solutions of the Poisson equations as in Lemma 2.3.22). (Refer to (2.3.21) and also the optimality criterion (2.3.22).)*

*(iii) Let* $d^{*'}$ *be optimal on* $[\vartheta', \vartheta_0]$ *and* $d^{*''}$ *optimal on* $[\vartheta_0, \vartheta'']$ *where* $0 \leq \vartheta' \leq \vartheta_0 \leq \vartheta''$.

*Then*

$$\left\{\hat{d}^{*'}(s^f, w) = s^l \neq \hat{d}^{*''}(s^f, w) = s^{\bar{l}} \text{ for an } s^f \in S \text{ and a } w \in B\right\}$$

$$\Rightarrow \left\{\triangle H^{d^{*'}}(s^f, w, s^{\bar{l}} - s^l, \vartheta_0) = 0 \text{ and } \triangle H^{d^{*''}}(s^f, w, s^l - s^{\bar{l}}, \vartheta_0) = 0\right\}.$$

*(iv) Let* $d^{*'}$ *be optimal on* $[\vartheta', \vartheta_0]$ *and* $d^{*''}$ *optimal on* $[\vartheta_0, \vartheta'']$ *where* $0 \leq \vartheta' < \vartheta_0 < \vartheta''$.
*In addition let* $\hat{d}^{*'}(s^f, w) = s^l \neq \hat{d}^{*''}(s^f, w) = s^{\bar{l}}$ *for an* $s^f \in S$ *and a* $w \in B$.

*Then*

*a)* $\triangle H^{d^{*'}}(s^f, w, s^{\bar{l}} - s^l, \vartheta)$ *is decreasing at* $\vartheta$ *and* $\triangle H^{d^{*''}}(s^f, w, s^l - s^{\bar{l}}, \vartheta)$ *is increasing at* $\vartheta$.

*b) If, additionally,*

$$\left\{\begin{array}{l} \triangle H^{d^{*''}}(s^f, w, s^l - s^{\bar{l}}, \vartheta) \text{ is strictly increasing at } \vartheta, \\ \triangle H^{d^{*'}}(s^f, w, s^{\bar{l}} - s^l, \vartheta) \text{ is strictly decreasing at } \vartheta \end{array}\right\}$$

*then optimal decisions*

$$\left\{\begin{array}{l} d^{*1} \text{ with } \hat{d}^{*1}(s^f, w) = s^l \text{ on } (\vartheta_0, \vartheta_0 + \varepsilon], \\ d^{*2} \text{ with } \hat{d}^{*2}(s^f, w) = s^{\bar{l}} \text{ on } [\vartheta_0 - \varepsilon, \vartheta_0) \end{array}\right\} \text{ do not exist for any } \varepsilon > 0$$

*thus*

$$\left\{\begin{array}{l} d^{*'} \text{ is not optimal on } (\vartheta_0, \vartheta_0 + \varepsilon], \\ d^{*''} \text{ is not optimal on } [\vartheta_0 - \varepsilon, \vartheta_0) \end{array}\right\} \text{ for any } \varepsilon > 0.$$

***Proof.***

(i) follows from (2.3.14) under the assumption of (LPC).

(ii) According to the assumptions of (LPC) and Lemma 2.3.22 together

with (2.3.43) the relations

$$\begin{aligned}
&\triangle H^d(s^f, w, s^{\bar{l}} - s^l, \vartheta) \\
&\quad = \hat{c}(s^f, w, s^{\bar{l}}, \vartheta) - \hat{c}(s^f, w, s^l, \vartheta) + \nu_{\bar{l}}^d(\vartheta) - \nu_l^d(\vartheta) \\
&\quad = \hat{c}(s^f, w, s^{\bar{l}}, \kappa_0) + \vartheta\hat{c}(s^f, w, s^{\bar{l}}, \kappa) - (\hat{c}(s^f, w, s^l, \kappa_0) + \vartheta\hat{c}(s^f, w, s^l, \kappa)) \\
&\qquad + \nu_{\bar{l}}^d(\xi_0) + \vartheta\ \nu_{\bar{l}}^d(\xi) - (\nu_l^d(\xi_0) + \vartheta\ \nu_l^d(\xi)) \\
&\quad = \vartheta\ [\hat{c}(s^f, w, s^{\bar{l}}, \kappa) - \hat{c}(s^f, w, s^l, \kappa) + \ \nu_{\bar{l}}^d(\xi) - \nu_l^d(\xi)] + \nu_{\bar{l}}^d(\xi_0) - \nu_l^d(\xi_0)
\end{aligned} \tag{2.3.44}$$

are valid. $H^d(s^f, w, s^{\bar{l}} - s^l, \vartheta)$ are therefore linear functions at $\vartheta$.

(iii) follows from Theorem 2.3.7.

(iv) a) $\triangle H^{d^{*'}}(s^f, w, s^{\bar{l}} - s^l, \vartheta)$ and $\triangle H^{d^{*''}}(s^f, w, s^l - s^{\bar{l}}, \vartheta)$ are increasing or decreasing at $\vartheta$ on $[0, \infty)$ since these terms are linear functions at $\vartheta$ (according to (ii)).

In addition, since $d^{*'}$ is optimal on $[\vartheta', \vartheta_0]$ and according to (2.3.22) and (iii) the relations

$\triangle H^{d^{*'}}(s^f, w, s^{\bar{l}} - s^l, \vartheta) \geq 0$ for all $\vartheta \in [\vartheta', \vartheta_0]$
and $H^{d^{*'}}(s^f, w, s^{\bar{l}} - s^l, \vartheta_0) = 0$ are valid.

Hence $\triangle H^{d^{*'}}(s^f, w, s^{\bar{l}} - s^l, \vartheta)$ is decreasing at $\vartheta$.

That $\triangle H^{d^{*''}}(s^f, w, s^l - s^{\bar{l}}, \vartheta)$ is increasing follows similarly.

b) $\triangle H^{d^{*'}}(s^f, w, s^{\bar{l}} - s^l, \vartheta_0) = 0$ and $\triangle H^{d^{*''}}(s^f, w, s^l - s^{\bar{l}}, \vartheta_0) = 0$ is valid according to (iii).

If, additionally, $\triangle H^{d^{*''}}(s^f, w, s^l - s^{\bar{l}}, \vartheta)$ is strictly increasing at $\vartheta$ then

$$\triangle H^{d^{*''}}(s^f, w, s^l - s^{\bar{l}}, \vartheta) > 0 \text{ on } (\vartheta_0, \infty) \tag{*1}$$

follows.

Assumption: $d^{*1}$ with $\hat{d}^{*1}(s^f, w) = s^l$ is an optimal decision at $\vartheta_0 + \varepsilon'$ for some $\varepsilon' \in (0, \varepsilon)$.

$$(\triangle H^{d^{*1}}(s^f, w, s^{\bar{l}} - s^l, \vartheta_0 + \varepsilon') = 0 \text{ and})\ H^{d^{*''}}(s^f, w, s^l - s^{\bar{l}}, \vartheta_0 + \varepsilon') = 0$$

according to (iii) (in relation to $\vartheta_0 + \varepsilon'$) is in contradiction to (*1).
(The other case can be shown similarly.) ■

**Definition 2.3.6.** *Let a DA* $MDP_l(N = \infty, S, A^M, P, \gamma(\vartheta))$ *on* $[0, \infty)$ *be given. In addition, let (2.3.3) be valid for all stationary policies.*

*Furthermore, let* $d^*$ *be an optimal decision at* $\vartheta_0$ *however not on the interval* $(\vartheta_0, \vartheta_0 + \varepsilon)$ *for an* $\varepsilon > 0$.

*If unique* $s^l \in S, s^{\bar{l}} \in S$ *and corresponding* $s^f \in S$, $w \in B$ *(not necessarily unique) with* $\hat{d}^*(s^f, w) = s^l$ *exist so that*

$$\triangle H^{d^*}(s^f, w, s^{\bar{l}} - s^l, \vartheta_0) = 0 \qquad (*)$$

*and if, additionally, in the case that* $\vartheta_0 = 0$ *the differences* $\hat{c}(s^f, w, s^{\bar{l}}, \kappa) - \hat{c}(s^f, w, s^l, \kappa)$ *for* $s^f \in S$ *and* $w \in B$ *from (*) are equal to each other, then the violation of the optimality on* $(\vartheta_0, \vartheta_0 + \varepsilon)$ *by* $d^*$ *is called a single violation.*

**Lemma 2.3.26.** *Let a DA* $MDP_l(N = \infty, S, A^M, P, \gamma(\vartheta))$ *on* $[0, \infty)$ *be given. Also, let (2.3.3) be valid for all stationary policies.*

*Furthermore, a single violation of the optimality on* $(\vartheta_0, \vartheta_0 + \varepsilon)$ *by* $d^*$ *with*

$$\triangle H^{d^*}(s^f, w, s^{\bar{l}} - s^l, \vartheta_0) = 0 \qquad (*)$$

*is given.*

*(i) In the case* $\vartheta_0 > 0$ *the differences* $\hat{c}(s^f, w, s^{\bar{l}}, \kappa) - \hat{c}(s^f, w, s^l, \kappa)$ *for* $s^f \in S$ *and* $w \in B$ *from (*) are equal to each other (as in the case* $\vartheta_0 = 0$*, according to Definition 2.3.6).*

*(ii)* $\varepsilon' > 0$ *exists such that* $d^{*'}$ *with*

$$\hat{d}^{*'}(s^{f'}, w') = \begin{cases} s^{\bar{l}} & \text{if } \triangle H^{d^*}(s^{f'}, w', s^{\bar{l}} - s^l, \vartheta_0) = 0, \\ d^*(s^{f'}, w') & \text{otherwise} \end{cases}$$

*is optimal on* $[\vartheta_0, \vartheta_0 + \varepsilon']$.

***Proof.***

(i) In the case $\vartheta_0 > 0$

$$\begin{aligned}
&\triangle H^{d^*}(s^f, w, s^{\bar{l}} - s^l, \vartheta_0) \\
&\quad = \hat{c}(s^f, w, s^{\bar{l}}, \vartheta_0) - \hat{c}(s^f, w, s^l, \vartheta_0) + \nu_{\bar{l}}^{d^*}(\vartheta_0) - \nu_l^{d^*}(\vartheta_0) \\
&\quad = \hat{c}(s^f, w, s^{\bar{l}}, \kappa_0) + \vartheta_0\, \hat{c}(s^f, w, s^{\bar{l}}, \kappa) - (\hat{c}(s^f, w, s^l, \kappa_0) + \vartheta_0\, \hat{c}(s^f, w, s^l, \kappa)) \\
&\qquad + \nu_{\bar{l}}^{d^*}(\vartheta_0) - \nu_l^{d^*}(\vartheta_0) \\
&\quad = \vartheta_0\, [\hat{c}(s^f, w, s^{\bar{l}}, \kappa) - \hat{c}(s^f, w, s^l, \kappa)] + \nu_{\bar{l}}^{d^*}(\vartheta_0) - \nu_l^{d^*}(\vartheta_0) = 0
\end{aligned}$$

yields

$$\hat{c}(s^f, w, s^{\bar{l}}, \kappa) - \hat{c}(s^f, w, s^l, \kappa) = \tfrac{1}{\vartheta_0}(\nu_{\bar{l}}^{d^*}(\vartheta_0) - \nu_l^{d^*}(\vartheta_0)).$$

(i) follows since $l$ and $\bar{l}$ are unique.

(ii) According to Lemma 2.3.25(iii) $d^*$ (optimal at $\vartheta_0$) and an optimal decision $d^{*''}$ on $[\vartheta_0, \vartheta_0 + \varepsilon']$ (see also Lemma 2.3.24) can differ only in single decisions $\hat{d}^*(s^f, w) = s^l$ and $\hat{d}^{*''}(s^f, w) = s^{\bar{l}}$ if $\triangle H^{d^*}(s^f, w, s^{\bar{l}} - s^l, \vartheta_0) = 0$.

Since $d^*$ is not optimal on $(\vartheta_0, \vartheta_0 + \varepsilon')$, $f^0$ and $w^0$ with $\triangle H^{d^*}(s^{f^0}, w^0, s^{\bar{l}} - s^l, \vartheta_0) = 0$ and $\hat{d}^{*''}(s^{f^0}, w^0) = s^{\bar{l}} \neq s^l = \hat{d}^*(s^{f^0}, w^0)$ exist.

Moreover, that $d^{*''}$ is optimal on $[\vartheta_0, \vartheta_0 + \varepsilon']$ means

$$\begin{aligned}
&\triangle H^{d^{*''}}(s^{f^0}, w^0, s^l - s^{\bar{l}}, \vartheta) \\
&\quad = \vartheta\, (\hat{c}(s^{f^0}, w^0, s^l, \kappa) - \hat{c}(s^{f^0}, w^0, s^{\bar{l}}, \kappa)) + \nu_l^{d^{*''}}(\vartheta) - \nu_{\bar{l}}^{d^{*''}}(\vartheta) \\
&\quad \geq 0 \;\; \forall\, \vartheta \in [\vartheta_0, \vartheta_0 + \varepsilon'].
\end{aligned}$$

According to (i)

$$\hat{c}(s^{f^0}, w^0, s^l, \kappa) - \hat{c}(s^{f^0}, w^0, s^{\bar{l}}, \kappa) = c(s^{f'}, w', s^l, \kappa) - \hat{c}(s^{f'}, w', s^{\bar{l}}, \kappa)$$

is valid for any $s^{f'}, w'$ with $\triangle H^{d^*}(s^{f'}, w', s^{\bar{l}} - s^l, \vartheta_0) = 0$.

Thus,

$$H^{d^{*''}}(s^{f^0}, w^0, s^l - s^{\bar{l}}, \vartheta) = H^{d^{*''}}(s^{f'}, w', s^l - s^{\bar{l}}, \vartheta) \geq 0 \;\; \forall\, \vartheta \in [\vartheta_0, \vartheta_0 + \varepsilon']$$

follows.

Hence, $d^{*'}(\equiv d^{*''})$ from (ii) is optimal on $[\vartheta_0, \vartheta_0 + \varepsilon']$.

■

We can now solve the cost-parametric $DA\ MDP_l(N = \infty, S, A^M, P, \gamma(\vartheta))$, where (2.3.3) is valid, in the following way:

**Algorithm 2.3.2.** .

1. *Solve $DA\ MDP_l(N = \infty, S, A^M, P, \gamma(\vartheta))$ for $\vartheta = 0$ (initial parameter) (by means of the common Howard algorithm, if necessary [13]).*

   *Let $d^{*0}$ be an optimal decision at $\bar{\vartheta}_1 = 0$ and*

   *let, in addition, $d^{*0}$ satisfy (ii) from Lemma 2.3.23.*

2. *Solve $DA\ MDP_l(N = \infty, S, A^M, P, \gamma(\vartheta))$ for $\vartheta \in (0, \infty)$ (or, if sufficient for $\vartheta \in (0, 1]$):*

   *For $i = 0, 1, 2, \cdots$:*

   2.1 *Compute $\bar{\vartheta}_{i+1} \in \mathbb{R}_+ \cup \{\infty\}$, so that $d^{*i}$ is an optimal decision on $[\bar{\vartheta}_i, \bar{\vartheta}_{i+1}]$ however not on $(\bar{\vartheta}_{i+1}, \bar{\vartheta}_{i+1} + \varepsilon)$ for any $\varepsilon > 0$ in the case $\bar{\vartheta}_{i+1} \neq \infty$ (see also Lemma 2.3.24):*

   *In greater detail:*

   *Calculate*

   $\triangle H^{d^{*i}}(s^f, w, s^{\bar{l}} - s^l, \vartheta)$ *according to (2.3.44)*

   *where $\nu^{d^{*i}}(\xi_0)$, $\nu^{d^{*i}}(\xi)$ are solutions of the corresponding Poisson equations (see Lemma 2.3.22),*

   $$\bar{\vartheta}_{i+1} = \min\{\vartheta \mid \triangle H^{d^{*i}}(s^f, w, s^{\bar{l}} - s^l, \vartheta) = 0 \text{ and } \triangle H^{d^{*i}}(s^f, w, s^{\bar{l}} - s^l, \vartheta') < 0 \text{ for } \vartheta' > \vartheta, \; s^f \in S,\ s^{\bar{l}} \in S,\ s^l \in S,\ w \in B\}.$$

   *($\bar{\vartheta}_{i+1} > \bar{\vartheta}_i$ follows from step 2.2.)*

   2.2 *Compute $d^{*(i+1)}$ in the case $\bar{\vartheta}_{i+1} \neq \infty$, so that $d^{*(i+1)}$ is an optimal decision on $[\bar{\vartheta}_{i+1}, \bar{\vartheta}_{i+1} + \varepsilon)$ for some $\varepsilon > 0$:*

   *In greater detail:*

[13] Also see the following remarks.

*Case 1: Single violation of the optimality on* $(\bar{\vartheta}_{i+1}, \bar{\vartheta}_{i+1}+\varepsilon)$ *by* $d^{*i}$ *(refer to Definition 2.3.6)*

*Set*

$$\hat{d}^{*(i+1)}(s^f, w) := \begin{cases} s^{\bar{l}} & \text{if } \triangle H^{d^{*i}}(s^f, w, s^{\bar{l}} - s^l, \bar{\vartheta}_{i+1}) = 0, \\ \hat{d}^{*i}(s^f, w) & \text{otherwise} \end{cases}$$

*(refer to Lemma 2.3.26 (ii)).*

*Case 2: No single violation of the optimality on* $(\bar{\vartheta}_{i+1}, \bar{\vartheta}_{i+1} + \varepsilon)$ *by* $d^{*i}$ [14]

*Use the common Howard algorithm in relation to* $\bar{\vartheta}_{i+1} + \varepsilon$ *for sufficiently small* $\varepsilon > 0$.

**Remarks on Algorithm 2.3.2:**

*Step 1* *We have assumed that the internal costs and the average one-step reward functions do not depend on the decisions in regard to the initial parameter* $\vartheta = 0$. *The optimal decisions therefore imply an almost-partial order of the states (see Section 2.3.2.2).*

- *This almost-partial order of the states means that the complexity of computation of optimal decisions in regard to the initial parameter can be reduced (refer also to Corollary 2.3.9).*
- *In some cases theoretical investigations yield the almost-partial order of the states and the optimal decisions in regard to the initial parameter. As a result computation is then not required.*
  *(For instance, see Sections 3.4, 3.5 and 4.7.)*
- *It is possible that the consideration of the objective function (scalar product) (refer to (2.3.2) or (2.3.11)) can help to find optimal decisions if the average one-step reward functions do not depend on these decisions.*
  *(Increase* $p^d_{fl}$ *and thus* $p^{d,\infty}_l$ *(Theorem 2.3.4) if* $\gamma(s^l)$ *is small.)*

[14] See also the following remarks.

*Step 2*

- *If the parameter increases then the violations of optimality are single violations (Definition 2.3.6) in general and the optimal decisions can be purposefully computed for the increasing parameter (see Case 1 in the above algorithm). Step 2 of the algorithm is a greedy algorithm for the cost-parametric DA Markov decision process in this case.*
- *Only single decisions* $\hat{d}^{*i}(s^f, w)$ *with* $\triangle H^{d^{*i}}(s^f, w, s^{\bar{l}} - s^l, \bar{\vartheta}_{i+1}) = 0$ *would have to be perhaps changed in Case 2 (refer to Lemma 2.3.25 (iii)).*
- *A finite number of optimal policies exists for an infinite set of parameter values (refer to Lemma 2.3.24).*

In addition to step 2 of Algorithm 2.3.2, we will deal with the following two topics in relation to an increasing parameter:

- The meaning of the internal cost in relation to the optimality criterion,
- Changes of optimal single decisions and monotonicity of the functions $\triangle H^d(s^f, w, s^{\bar{l}} - s^l, \vartheta)$ at $\vartheta$.

**Meaning of the Internal Cost in relation to the Optimality Criterion**

**Remarks 2.3.6.** *Violations of the optimality criterion (see Lemma 2.3.6 and (2.3.44))*

$$\begin{aligned}
&\triangle H^d(s^f, w, s^{\bar{l}} - s^l, \vartheta) \\
&\quad = \vartheta\ [\hat{c}(s^f, w, s^{\bar{l}}, \kappa) - \hat{c}(s^f, w, s^l, \kappa)] + \\
&\qquad \nu_{\bar{l}}^d(\xi_0) - \nu_l^d(\xi_0) + \vartheta\ [\ \nu_{\bar{l}}^d(\xi) - \nu_l^d(\xi)] \\
&\quad < 0 \qquad (\vartheta > 0)
\end{aligned}$$

*seem to occur especially if*

$\hat{c}(s^f, w, s^{\bar{l}}, \kappa) - \hat{c}(s^f, w, s^l, \kappa) < 0.$

Therefore we formulate a corresponding **a**dditional **c**ondition for $DA\ MDP_l(N=\infty, S, A^M, P, \gamma(\vartheta))$ on $[0,\infty)$, where (2.3.3) is valid for all stationary policies:

(AC1) *Let an almost-partial order of the states be implied by the optimal decision at the initial parameter $\vartheta = 0$.*
*Optimal decisions (for $\vartheta > 0$) for costlier states (refer to Theorem 2.3.8 (i)) can occur only if the corresponding internal cost are lower.*

(AC1) is always valid at least for sufficiently small values:

**Theorem 2.3.27.** *Let a $DA\ MDP_l(N=\infty, S, A^M, P, \gamma(\vartheta))$ on $[0,\infty)$ (or on $[0,1]$) be given. In addition, let (2.3.3) be valid for all stationary policies.*

*Then, some $\vartheta' > 0$ exists such that optimal decisions at any $\vartheta \in [0,\vartheta']$ satisfy the additional condition (AC1).*

***Proof.*** Let $d^*$ be an optimal decision with respect to $\vartheta = 0$.

If a state $s^{\bar{l}}$ is costlier than a state $s^l$ (meaning $s^{\bar{l}} < s^l$) (see Theorem 2.3.8(i)), then

$$\nu_{\bar{l}}^{d^*}(\vartheta = 0) - \nu_l^{d^*}(\vartheta = 0) > 0 \qquad (*1)$$

is valid for solutions of the corresponding Poisson equation (see Theorem 2.3.8(ii)).

We show that (AC1) is true on $[0,\vartheta']$ for some $\vartheta' > 0$ in such a way that no change of an optimal single decision for a costlier state for any $\vartheta \in (0,\vartheta']$ occurs.

Since

$$\hat{c}(s^f, w, s^{\bar{l}}, \vartheta = 0) - \hat{c}(s^f, w, s^l, \vartheta = 0) = 0$$

for $s^f, w$ with $\{s^l, s^{\bar{l}}\} \subseteq \hat{A}(s^f, w)$ according to (LPC) and (*1)

$$\begin{aligned}&\triangle H^{d^*}(s^f, w, s^{\bar{l}} - s^l, \vartheta = 0)\\ &\quad = \hat{c}(s^f, w, s^{\bar{l}}, \vartheta = 0) - \hat{c}(s^f, w, s^l, \vartheta = 0) + \nu_{\bar{l}}^{d^*}(\vartheta = 0) - \nu_l^{d^*}(\vartheta = 0) > 0\end{aligned} \qquad (*2)$$

follows.

If $d^*$ is not optimal on $(0, \varepsilon)$ for any $\varepsilon > 0$, then some real number $\vartheta' > 0$ and a decision function $d^{*'}$ exist such that $d^{*'}$ is optimal on $[0, \vartheta']$ according to Lemma 2.3.23(ii).

$d^{*'}$ and $d^*$ can be different only in relation to single decisions $d^{*'}(s^{f'}, w') = s^{\bar{l}'}$ and $d^*(s^{f'}, w') = s^{l'}$ if $\triangle H^{d^{*'}}(s^{f'}, w', s^{l'} - s^{\bar{l}'}, \vartheta = 0) = 0$ and $\triangle H^{d^*}(s^{f'}, w', s^{\bar{l}'} - s^{l'}, \vartheta = 0) = 0$ (see Lemma 2.3.25(iii))).

Thus, according to (*2) $s^{\bar{l}'}$ cannot be costlier than $s^{l'}$. ■

**Changes of Optimal Single Decisions and Monotonicity of the Functions $\triangle \mathbf{H^d}(\mathbf{s^f}, \mathbf{w}, \mathbf{s^{\bar{l}}} - \mathbf{s^l}, \vartheta)$ at $\vartheta$**

If a change of an optimal single decision at $\vartheta_0 > 0$ occurs, then property (iv), a) from Lemma 2.3.25 is valid.

A stronger property is contained in the following **a**dditional **c**ondition for the $DA\ MDP_l(N = \infty, S, A^M, P, \gamma(\vartheta))$ on $I = [0, \infty)$ (or $\vartheta \in I = [0, 1]$), where (2.3.3) is valid for all stationary policies:

(AC2) *In accordance with Lemma 2.3.24, let a finite sequence of real numbers* $0 = \bar{\vartheta}_1 < \bar{\vartheta}_2 < \cdots < \bar{\vartheta}_{y-1}$ *(with* $\vartheta_{y-1} < 1$, *if* $I = [0, 1]$*) and a corresponding sequence of decisions* $d^{*1} \in A^M, \ldots, d^{*(y-1)} \in A^M$ *be given so that*

$d^{*i}$ *are optimal on* $I_i := [\bar{\vartheta}_i, \bar{\vartheta}_{i+1}]$ *for* $i = 1, 2, \cdots, y-2$

*and* $d^{*(y-1)}$ *is optimal on* $I_{y-1} = [\bar{\vartheta}_{y-1}, \infty)$ *(or* $I_{y-1} = [\bar{\vartheta}_{y-1}, 1]$*).*

*If*

$\hat{d}^{*(i_0-1)}(s^f, w) = s^l$

*for certain* $i_0 \in \{2, 3, \cdots, y\}$, $\{s^f, s^l\} \subseteq S$, $w \in B$ *and*

$\hat{d}^{*i_0}(s^f, w) = s^{\bar{l}}$ *if* $i_0 < y$ $(s^{\bar{l}} \in S, s^{\bar{l}} \neq s^l)$

*then*

$\triangle H^{d^{*i}}(s^f, w, s^{\bar{l}} - s^l, \vartheta)$ *are decreasing at* $\vartheta$ *for all* $i = 1, \cdots, i_0 - 1$

*and*

$\triangle H^{d^{*i}}(s^f, w, s^l - s^{\bar{l}}, \vartheta)$ *are increasing at* $\vartheta$ *for all* $i = i_0, \cdots, y-1$, *if* $i_0 < y$.

If (AC2) is valid, then the complexity of computation in step 2 of Algorithm 2.3.2 can be estimated in the following way:

**Theorem 2.3.28.** *Let a DA* $MDP_l(N = \infty, S, A^M, P, \gamma(\vartheta))$ *on* $[0, \infty)$ *(or on* $[0, 1]$*) be given where (2.3.3) is valid for all stationary policies.*

*If (AC2) is valid, then at most* $\sum_{s^f \in S; w \in B} |\hat{A}(s^f, w)|$ *single decisions have to be changed (in relation to* $d^{*1}$*) in step 2 of Algorithm 2.3.2 in order to compute optimal decisions on* $(0, \infty)$ *(or* $(0, 1]$*).*

***Proof.*** Let us note that $d^{*1}$ optimal at $\vartheta = 0$.

1. Initially, we construct $d'^{*i}$ for $i = 1, 2, \cdots, y-1$ such that

   $d'^{*1} := d^{*1}$

   and

   $$\hat{d}'^{*i}(s^f, w) := \begin{cases} \hat{d}'^{*(i-1)}(s^f, w) & \text{if } \triangle H^{d^{*i}}(s^f, w, s^{\bar{l}} - s^{l'}, \vartheta) \equiv 0, \\ \hat{d}^{*i}(s^f, w) & \text{otherwise} \end{cases} \qquad (*1)$$

   (where $s^{l'} = \hat{d}'^{*(i-1)}(s^f, w)$ and $s^{\bar{l}} = \hat{d}^{*i}(s^f, w)$)

   for $i = 2, 3, \cdots, y-1,\ s^f \in S,\ w \in B$.

   Successively using Theorem 2.3.7b) yields that $d'^{*i}$ are also optimal on $I_i := [\bar{\vartheta}_i, \bar{\vartheta}_{i+1}]$ for $i = 2, 3, \cdots, y-1$.

2. Next, we will show that $\triangle H^{d^{*i}}(s^f, w, \mathbf{s^{l'}} - s^{\bar{l}}, \vartheta)$ are also increasing at $\vartheta$ for $i \geq i_0$. (The proof that $\triangle H^{d^{*i}}(s^f, w, s^{\bar{l}} - s^{l'}, \vartheta)$ are decreasing at $\vartheta$ for $i < i_0$ is analogous.)

   Let

   $\hat{d}^{*i}(s^f, w) = s^{l_1},\ \ \hat{d}'^{*i}(s^f, w) = s^{l_1}$ for $i = 1, \cdots, i_1 - 1$,

   $\hat{d}^{*i}(s^f, w) = s^{l_2},\ \ \hat{d}'^{*i}(s^f, w) = s^{l_1}$ for $i = i_1, \cdots, i_2 - 1$, $(s^{l_2} \neq s^{l_1})$,

   case a)

   $\hat{d}^{*i}(s^f, w) = s^{l_3},\ \ \hat{d}'^{*i}(s^f, w) = s^{l_3}$ for $i = i_2$, $(s^{l_1} \neq s^{l_3} \neq s^{l_2})$

case b)

$\hat{d}^{*i}(s^f, w) = s^{l_2}, \ \ \hat{d}'^{*i}(s^f, w) = s^{l_2}$ for $i = i_2$.

Then,

$\triangle H^{d^{*i}}(s^f, w, s^{l_1} - s^{l_2}, \vartheta)$ are increasing at $\vartheta$ for $i \geq i_1$ and

$\triangle H^{d^{*i}}(s^f, w, s^{l_2} - s^{l_3}, \vartheta)$ are increasing at $\vartheta$ for $i \geq i_2$ (case a)

according to (AC2).

Since

$$\triangle H^{d^{*i}}(s^f, w, s^{l_1} - s^{l_3}, \vartheta)$$
$$= \triangle H^{d^{*i}}(s^f, w, s^{l_1} - s^{l_2}, \vartheta) + \triangle H^{d^{*i}}(s^f, w, s^{l_2} - s^{l_3}, \vartheta)$$

according to (2.3.23)

$\triangle H^{d^{*i}}(s^f, w, s^{l_1} - s^{l_3}, \vartheta)$ is also increasing at $\vartheta$ for $i \geq i_2$.

If we use the denotation $l_1 = l'$, $l_2 = l$ and $l_3 = \bar{l}$, we then see that

$\triangle H^{d^{*i}}(s^f, w, s^{l'} - s^{\bar{l}}, \vartheta)$ is increasing at $\vartheta$ for $i \geq i_2$.

(In case b it follows immediately, according to (AC2) that $\triangle H^{d^{*i}}(s^f, w, s^{l'} - s^{\bar{l}}, \vartheta)$ are increasing at $\vartheta$ for $i \geq i_2 (\geq i_1)$ (where $l_1 = l'$, $l_2 = \bar{l}$).)

If we successively continue the above consideration the more general relationship follows:

If

$\hat{d}^{*(i_0-1)}(s^f, w) = s^l, \ \ \hat{d}'^{*(i_0-1)}(s^f, w) = s^{l'}$,

$\hat{d}^{*i_0}(s^f, w) = s^{\bar{l}}$ and $\hat{d}'^{*i_0}(s^f, w) = s^{\bar{l}}$

then

$$\triangle H^{d^{*i}}(s^f, w, s^{l'} - s^{\bar{l}}, \vartheta) \text{ are increasing at } \vartheta \text{ for } i \geq i_0. \qquad (*2)$$

3. Now, let $\hat{d}'^{*(i_0-1)}(s^f, w) = s^{l'}$

$$\text{and } \hat{d}'^{*(i_0)}(s^f, w) = d^{*(i_0)}(s^f, w) = s^{\bar{l}}. \qquad (*3)$$

We can then show that

$$\triangle H^{d^{*i}}(s^f, w, s^{l'} - s^{\bar{l}}, \vartheta) > 0 \quad \text{on } I_i \setminus \{\bar{\vartheta}_{i_0}\} \text{ for } i = i_0 \text{ or} \qquad (*4)$$

$$\text{on } I_i \text{ for } i = i_0+1, i_0+2, \cdots, y-1.$$

Initially,

$$\triangle H^{d'^{*(i_0-1)}}(s^f, w, s^{\bar{l}} - s^{l'}, \bar{\vartheta}_{i_0}) = \triangle H^{d^{*i_0}}(s^f, w, s^{l'} - s^{\bar{l}}, \bar{\vartheta}_{i_0}) = 0 \text{ (*5)}$$

is valid according to Lemma 2.3.25(iii).

Furthermore,

$$\triangle H^{d^{*i_0}}(s^f, w, s^{l'} - s^{\bar{l}}, \vartheta) \geq 0 \text{ on } I_{i_0}$$

according to the optimality criterion (Lemma 2.3.6).

$\triangle H^{d^{*i_0}}(s^f, w, s^{l'} - s^{\bar{l}}, \vartheta)$ is a linear function at $\vartheta$ (refer to Lemma 2.3.25(ii)) and together with (*2)

$$\triangle H^{d^{*i_0}}(s^f, w, s^{l'} - s^{\bar{l}}, \vartheta) \geq 0 \text{ on } [\bar{\vartheta}_{i_0}, \infty) \text{ (or } \vartheta \in I = [\bar{\vartheta}_{i_0}, 1])$$

follows.

If $\triangle H^{d^{*i_0}}(s^f, w, s^{l'} - s^{\bar{l}}, \vartheta) = 0$ for some $\vartheta > \vartheta_{i_0}$,

then $\triangle H^{d^{*i_0}}(s^f, w, s^{l'} - s^{\bar{l}}, \vartheta) \equiv 0$ (refer to (*5) and Lemma 2.3.25(ii)).

$\hat{d}'^{*i_0}(s^f, w) = s^{l'}$ would then follow according to (*1) which is in contradiction to (*3).

Thus,

$$\triangle H^{d^{*i_0}}(s^f, w, s^{l'} - s^{\bar{l}}, \vartheta) > 0 \text{ for } \vartheta > \vartheta_{i_0}.$$

and in particular,

$$\triangle H^{d^{*i_0}}(s^f, w, s^{l'} - s^{\bar{l}}, \bar{\vartheta}_{i_0+1}) > 0.$$

Since $d^{*i_0}$ and $d^{*(i_0+1)}$ are optimal at $\bar{\vartheta}_{i_0+1}$, the relation

$$\triangle H^{d^{*i_0}}(s^f, w, s^{l'} - s^{\bar{l}}, \bar{\vartheta}_{i_0+1}) = H^{d^{*(i_0+1)}}(s^f, w, s^{l'} - s^{\bar{l}}, \bar{\vartheta}_{i_0+1}) (> 0)$$

is valid according to the definition of $\triangle H^d(.,.,.)$ (see (2.3.21)) and Theorem 2.3.21.

$$H^{d^{*(i_0+1)}}(s^f, w, s^{l'} - s^{\bar{l}}, \vartheta) > 0 \text{ for } \vartheta \geq \bar{\vartheta}_{i_0+1}$$

follows since $H^{d^{*(i_0+1)}}(s^f, w, s^{l'} - s^{\bar{l}}, \vartheta)$ is increasing at $\vartheta$ according to (*2).

If we continue these considerations, successively, in relation to $i_0+2, i_0+3, \cdots, y-1$, then (*4) is shown.

4. We now show that if

$$\hat{d}'^{*(i_0-1)}(s^f, w) = s^{l'} \ (i_0 < y) \text{ and } \hat{d}'^{*i_0}(s^f, w) = s^{\bar{l}},\ s^{\bar{l}} \neq s^{l'} \qquad (*6)$$

then there exists no decision $d^*$ with $\hat{d}^*(s^f, w) = s^{l'}$ which is optimal at some $\vartheta > \bar{\vartheta}_{i_0}$:

Assumption: $d^*$ with $\hat{d}^*(s^f, w) = s^{l'}$ is optimal at some $\vartheta' \in I_{i'}$ where $i' \geq i_0, \vartheta' \neq \bar{\vartheta}_{i_0}$.

(*6) together with (*1) means

$$\hat{d}^{*i_0}(s^f, w) = s^{\bar{l}} \qquad (*7)$$

and

$$\triangle H^{d^{*i'}}(s^f, w, s^{l'} - s^{\bar{l}}, \vartheta') > 0$$

follows from (*4).

However,

$$\triangle H^{d^*}(s^f, w, s^{\bar{l}} - s^{l'}, \vartheta')$$
$$= \triangle H^{d^{*i'}}(s^f, w, s^{l'} - s^{\bar{l}}, \vartheta') = 0$$

follows from (*7), the previous assumption and Lemma 2.3.25(iii).

The statement of 4 means that after a change of a single decision $\hat{d}'^{*i}(s^f, w) = s^{l'}$, this single decision can not be repeatedly optimal.

5. In step 2 of Algorithm 2.3.2 single decisions are changed only if violations of the optimality occur in $d'^{*i}$. Thus, at most, $\sum_{s^f \in S; w \in B} |\hat{A}(s^f, w)|$ single decisions have to be changed (in relation to $d^{*1}$) in order to compute optimal decisions on $(0, \infty)$ (or on $(0, 1]$).

■

The complexity of computation of optimal decisions on $(0, \infty)$ (or on $(0, 1]$) can be reduced further if the following strong additional condition is satisfied:

(AC3) *Let $d^{*0}$ be an optimal decision at $\vartheta = 0$ (refer to step 1 of Algorithm 2.3.2).*

*Single decisions $\hat{d}^{*0}(s^f, w)$ have to be changed at most once in order to compute optimal decisions on $(0, \infty)$ (or on $(0, 1]$).*

Thus, at most $|S| * |B|$ single decisions have to be changed (in relation to $d^{*0}$) in order to compute optimal decisions on $(0, \infty)$ (or on $(0, 1]$) if (AC3) is valid.

In addition, step 2 of Algorithm 2.3.2 is a greedy algorithm in relation to the DA MDP itself, which corresponds to the cost-parametric DA MDP at $\vartheta = 1$.

The SDDP problem from Example 3.5.1, Section 3.5 satisfies (AC1), (AC2) and (AC3).

#### 2.3.4.3 Remarks on the Parametrization of DA Markov Decision Processes

If we want to attempt to use Algorithm 2.3.2 for solving of a given DA MDP (in terms of a continuation of the solutions of the parameterized problem) then the internal cost should be parameterized in a logical way.

The parameterized internal costs at $\vartheta = 0$ and the given internal costs of the DA MDP should not differ greatly (*)

so that the almost-partial order of the states at $\vartheta = 0$ (refer to Theorem 2.3.8) does not completely change.

(Since the parameterized internal costs at $\vartheta = 0$ have to satisfy (2.3.13) (see (LPC)), the demand (*) can be met only to a certain degree).

**Example 2.3.1.** *Let a DA MDP be given where, in particular,*
$S = \{s^1, s^2\}$, $B = \{w^1, w^2, w^3\}$, $q(w^i) = \frac{1}{3}$ *for* $i = 1, 2, 3$,
*the DA decision sets* $\hat{A}(s^1, w^1) = \hat{A}(s^2, w^2) = \{s^1, s^2\}$,
$\hat{A}(s^1, w^2) = \hat{A}(s^2, w^3) = \{s^1\}$ *and* $\hat{A}(s^2, w^1) = \hat{A}(s^1, w^3) = \{s^2\}$
*(hence (2.3.3) is valid for all stationary policies of this DA MDP)*
*and the internal costs* $(\hat{c}(s^f, s^l))_{\substack{f=1;2 \\ l=1;2}} = \begin{pmatrix} 25 & 26 \\ 30 & 14 \end{pmatrix}$ *(which satisfy (2.3.12)).*

*Then, the parametrization*

*P1:* $(\hat{c}(s^f, s^l, \vartheta))_{\substack{f=1;2\\ l=1;2}} = \begin{pmatrix} 25(+0\vartheta) & 25+\vartheta \\ 14+16\vartheta & 14(+0\vartheta) \end{pmatrix}, \ \vartheta \in [0,1]$

*is better than*

*P2:* $(\hat{c}(s^f, s^l, \vartheta))_{\substack{f=1;2\\ l=1;2}} = \begin{pmatrix} 1+24\vartheta & 1+25\vartheta \\ 2+28\vartheta & 2+12\vartheta \end{pmatrix}, \ \vartheta \in [0,1].$

*($s^1 < s^2$ is valid (at $\vartheta = 0$ under the terms of Theorem 2.3.8) in relation to P1 contrary to $s^2 < s^1$ in relation to P2.)*

For certain DA MDP a parametrization of the internal costs arises in "natural" way (compare SDDP problems, Section 3.5).

### 2.3.5 Remarks on the Solutions of DA Markov Decision Processes

- In principle, DA MDP$(N = \infty, S, A^M, P, \gamma)$ can be solved by means of the Howard algorithm (policy iteration).

  Here, the computation of a better decision at a step of the iteration can be formulated using the definition of single decisions (see (2.3.21)):

  **Algorithm 2.3.3.** .

  1. *Let a (feasible) decision $d^0$ be given.*

     *Set* $\nu_m^i := 0$ *for* $i = 0, 1, \cdots,$

     $i := 0.$

  2. *Compute the solution $(g^i, (\nu_f^i)_{f=1,\dots,m})$ (with $\nu_m^i := 0$) of the Poisson equation*

     $$g \begin{pmatrix} -1 \\ \vdots \\ -1 \end{pmatrix} + (P^{d^i} - I)\, \nu = -\gamma^{d^i}$$

     *(where $I$ is the identity matrix).*

  3. *Compute single decisions $\hat{d}^{*i}(s^f, w) = s^{\bar{l}}$ according to*

     $$\min_{s^{\bar{l}} \epsilon \hat{A}(s^f, w)} \triangle H^{d^i}(s^f, w, s^{\bar{l}} - s^l)$$
     $$(= \min_{s^{\bar{l}} \epsilon \hat{A}(s^f, w)} (\hat{c}(s^f, w, s^{\bar{l}}) - \hat{c}(s^f, w, s^l) + \nu_{\bar{l}}^i - \nu_l^i).$$

*(where* $s^l = \hat{d}^i(s^f, w)$*)*
*for* $s^f \in S, w \in B$*.*

*Set*

$$\hat{d}^{i+1}(s^f, w) = \begin{cases} s^{\bar{l}} & \text{if } \triangle H^{d^i}(s^f, w, s^{\bar{l}} - s^l) < 0, \\ s^l = d^i(s^f, w) & \text{otherwise.} \end{cases}$$

*4. If* $d^{i+1} \neq d^i$ *then set* $i := i + 1$ *and go to step 2, otherwise* $d^{i+1}(= d^i)$ *is an optimal decision function.*

If DA MDP$(N = \infty, S, A^M, P, \gamma)$ can be parameterized in sensible way (*)

(refer the remarks at beginning of Section 2.3.4.3), then the optimal decision $d^{*0}$ of the parameterized problem at $\vartheta = 0$ can be recommended as $d^0$ in the above Howard algorithm.

- On the other hand, we can compute an optimal decision of a DA MDP itself (under the assumption (*)) by solving the parameterized DA MDP on $[0, 1]$ by way of a continuation of the solutions of the parameterized problem.

  Often the violations of the optimality, if the parameter is increasing, are single violations (Definition 2.3.6):

  – advantage: purposeful computation,
  – disadvantage: Poisson equations have to be solved, repeatedly if only single decisions are changed.

  (However, in Algorithm 2.3.3 it can also occur that changes of single decisions are, in the end, not optimal and further changes of such single decisions are necessary.)

### Consideration of Simple Heuristic Solutions of DA MDPs

Let us assume that a DA MDP can be parameterized in sensible way.

The optimal decision $d^{*0}$ of the parameterized problem at $\vartheta = 0$ implies an almost-partial orders of the states (see Section 2.3.2.2).

(Sometimes theoretical investigations yield the almost-partial order and no computation is required.)

If the parameter increases, the (different) internal costs have an effect on optimal decisions.

Often, cheaper internal costs imply cheaper average expected cost per stage (see Remarks 2.3.6), especially if (AC1) is valid.

If almost-partial order and internal costs in relation to single decisions are "balanced", simple heuristic solutions $d^*$ follow. In greater detail this means:

If $\hat{d}^{*0}(s^f, w) = s^l$, in principle, we set

$$\hat{d}^*(s^f, w) = \begin{cases} s^{\bar{l}} \quad \text{if} \quad s^{\bar{l}} \in \hat{A}(s^f, w) \quad \text{and} \\ \qquad\qquad \hat{c}(s^f, w, s^{\bar{l}}) \ll \hat{c}(s^f, w, s^l), \\ s^l = d^{*0}(s^f, w) \quad \text{otherwise.} \end{cases}$$

Such considerations are very important if the state space of a DA MDP is extremely large and exact methods of stochastic dynamic programming are not practical.

# Chapter 3

# The Problem of Stochastic Dynamic Distance Optimal Partitioning (SDDP)

In Section 3.1, we initially make a few statements about non-balanced transportation problems, which we will then need in the following sections.

The SDDP problem is introduced in Section 3.2 from the view of a possible practical application.

Formulation of the mathematical model then follows as a DA stochastic dynamic programming problem with random disturbances and afterwards more specific as a DA MDP (see Sections 2.1 and 2.3).

Additional properties of the SDDP problem and its characteristic parameters are addressed in Section 3.3.

In particular the SDDP problem proves to be a DA MDP with distance properties.

In this section, we will also introduce the idea of the "conversion number" and show that optimal decisions with a minimal conversion number always exist.

Special cases of the SDDP problem are the focus of interest in Section 3.4.

Under the assumption of identical "basic costs" (in other words of "unit

distances"), formulas for the average one-step reward functions are derived and a conjecture of optimal solutions of corresponding SDDP problems are given.

Under the assumptions of identical "basic costs" and independent and identically distributed requirements, the corresponding SDDP problems can be reduced. However, the conjecture for optimal decisions in certain cases can only be proven after several combinatorial considerations, which are discussed in Section 4.6.

In Section 3.5 we will discuss possibilities of exact and approximate solution methods for SDDP problems. For this purpose we orientate ourselves in addition by Section 2.3.5, whereby we also use the special cases from Section 3.3 for the initial solutions.

## 3.1 Preliminary Notes on Non-Balanced Transportation Problems

In this section we make a few statements about non-balanced transportation problems (TPs) in order to understand the following SDDP problems.

We use the following variables and symbols

$a$ - availabilities, $a \in \mathbb{Z}_+^n, \sum\limits_{i=1}^{n} a_i = su,\ su \in \mathbb{Z}_+,$

$b$ - requirements, $b \in \mathbb{Z}_+^n,$

$(k_{ij})_{\substack{i=1,\dots,n \\ j=1,\dots,m}}$ - basic cost, $(k_{ij})_{\substack{i=1,\dots,n \\ j=1,\dots,m}} \in \mathbb{R}_+^n \times \mathbb{R}_+^n$

(or in other words "distances")

and

$C[b, su]$ - denotes the case $\sum\limits_{j=1}^{n} b_j \leq \sum\limits_{i=1}^{n} a_i = su$ (surplus-situation, the requirements can be completely fulfilled),

$C[su, b]$ - denotes the case $\sum\limits_{j=1}^{n} b_j \geq \sum\limits_{i=1}^{n} a_i = su$ (scarcity-situation, the requirements can not be completely met, however they should be met as much as possible)

and

$X_{f(easible)}(a,b)$ - set of feasible solutions:

$$X_f(a,b) = \left\{ x \in \mathbb{Z}_+^n \times \mathbb{Z}_+^n \;\middle|\; \begin{array}{l} \sum\limits_{j=1}^{n} x_{ij} \leq a_i \;\forall i, \sum\limits_{i=1}^{n} x_{ij} = b_j \forall j \;\text{ in } C[b,su], \\ \sum\limits_{i=1}^{n} x_{ij} \leq b_j \;\forall j, \sum\limits_{j=1}^{n} x_{ij} = a_i \forall i \;\text{ in } C[su,b] \end{array} \right\} \tag{3.1.1}$$

[1]

$$\left( \begin{array}{c} \text{hence if } C[b,su] \text{ and } C[su,b] \text{ are valid it is immediately seen that} \\ X_f(a,b) = \left\{ x \in \mathbb{Z}_+^n \times \mathbb{Z}_+^n \;\middle|\; \begin{array}{c} \sum\limits_{j=1}^{n} x_{ij} = a_i \;\forall i, \sum\limits_{i=1}^{n} x_{ij} = b_j \forall j \\ \text{if } \sum\limits_{i=1}^{n} b_i = \sum\limits_{i=1}^{n} a_i = su \end{array} \right\} \end{array} \right). \tag{3.1.1a}$$

Then,

$TP^*(a,b)$ - denotes the transportation problem:

$\sum\limits_{i=1}^{n} \sum\limits_{j=1}^{n} k_{ij}\, x_{ij} \;\rightarrow\; min,\; x \in X_f(a,b),$ where

the basic costs $(k_{ij})_{\substack{i=1,\ldots,n \\ j=1,\ldots,m}}$ still have to satisfy the distance properties:

$$\begin{aligned} &k_{ii} = 0 \;\forall\, i,\; k_{ij} > 0 \;\forall\, i \neq j, \\ &k_{ij} + k_{jl} > k_{il} \;\forall i \neq j \neq l \text{ (triangle-inequality).} \end{aligned} \tag{3.1.2}$$

Established statements are that $X_f(a,b) \neq \varnothing$ and $TP^*(a,b)$ is solvable (see Sections 2.8.7 and 3.1.2 in [30], for example).

If $x \in X_f(a,b)$, then we can also define slack-variables

[1] In case $C[su,b]$, the scarcity-situation, the relations secure that in no case are the availabilities dissipated.

$$\begin{cases} x_{i\,n+1} := a_i - \sum\limits_{j=1}^{n} x_{ij},\ i = 1, 2, \cdots, n \ \text{ in case } C[b, su], \\ x_{j\,n+1} := -(b_j - \sum\limits_{i=1}^{n} x_{ij}),\ j = 1, 2, \cdots, n \ \text{ in case } C[su, b], \end{cases} \tag{3.1.3}$$
$$x_{n+1} := (x_{j\,n+1})_{j=1,\cdots,n}.$$

We will later use the following definition:

**Definition 3.1.1.** *Let a transportation problem $TP^*(a, b)$ and a feasible solution $x \in X_f(a, b)$ with the corresponding slack-vector $x_{n+1}$ be given.*

*Then, $TP^*(a, b + x_{n+1})$ is called balanced transportation problem with respect to $TP^*(a, b)$ and $x \in X_f(a, b)$.*

**Lemma 3.1.1.** *Let a transportation problem $TP^*(a, b)$ and a feasible solution $x \in X_f(a, b)$ be given.*

*(i) If the relationship*

$$x_{ii} = min\{a_i, b_i\} \ \ for \ \ i = 1, 2, \cdots, n \tag{3.1.4}$$

*is not valid, then a feasible solution $\tilde{x} \in X_f(a, b)$ with lower cost: $\sum\limits_{i=1}^{n}\sum\limits_{j=1}^{n} k_{ij}\ \tilde{x}_{ij} < \sum\limits_{i=1}^{n}\sum\limits_{j=1}^{n} k_{ij}\ x_{ij}$ (and with $b + \tilde{x}_{n+1} = b + x_{n+1}$ in case $C[b, su]$) exists.*

*(ii) The relation (3.1.4) is true for optimal solutions $x$ of $TP^*(a, b)$.*

*(iii) Let $TP^*(a, b+x_{n+1})$ be the balanced transportation problem with respect to $TP^*(a, b)$ and $x \in X_f(a, b)$.*

*Then $\hat{x}$ with*

$$\hat{x}_{ij} = \begin{cases} x_{ii} + x_{i\,n+1} = a_i - \sum\limits_{l:i\neq l} x_{il} \ \ if\, i = j \\ x_{ij} \ \ if\, i \neq j \end{cases} \qquad in\ case\ C[b, su]$$

*and*

$$\hat{x} = x \qquad in\ case\ C[su, b]$$

*is a feasible solution of $TP^*(a, b + x_{n+1})$.*

***Proof.***

(i)

Case $C[b, su]$ and $b_{i_0} \leq a_{i_0}$:

Assumption: $x_{i_0 i_0} \neq b_{i_0}$.

$$x_{i_0 i_0} < b_{i_0}$$

follows from (3.1.1).

If $b_{i_0} \leq a_{i_0}$, then

$$\exists\, j_0 \neq i_0 : \; x_{i_0 j_0} > 0 \;\; (j_0 \in \{1, \cdots, n+1\}).$$

Furthermore, (3.1.1) yields:

$$\exists\, y_0 \neq i_0 : \; x_{y_0 i_0} > 0 \;\; (y_0 \in \{1, \cdots, n\}).$$

We now define $\tilde{x}$ as:

$$\tilde{x}_{ij} = \begin{cases} x_{ij} + 1 & \text{if } (i,j) = (i_0, i_0) \text{ or } (i,j) = (y_0, j_0), \\ x_{ij} - 1 & \text{if } (i,j) = (i_0, j_0) \text{ or } (i,j) = (y_0, i_0), \\ x_{ij} & \text{otherwise.} \end{cases}$$

Obviously, $\tilde{x} \in X_f(a, b)$ and $\tilde{x}_{n+1} = x_{n+1}$ thus $b + \tilde{x}_{n+1} = b + x_{n+1}$.

Using the distance properties (3.1.2)

$$\begin{aligned} \sum_{i,j=1}^{n} k_{ij}\, \tilde{x}_{ij} &= \sum_{i,j=1}^{n} k_{ij}\, x_{ij} + k_{i_0 i_0} + k_{y_0 j_0} - k_{i_0 j_0} - k_{y_0 i_0} \\ &= \sum_{i,j=1}^{n} k_{ij}\, x_{ij} + k_{y_0 j_0} - k_{i_0 j_0} - k_{y_0 j_0} \\ &< \sum_{i,j=1}^{n} k_{ij}\, x_{ij} \end{aligned}$$

follows.

If $b_{i_0} > a_{i_0}$ in case $C[b, su]$, then (i) can be proven analogously to the above statement.

It is now necessary to discuss a set of subcases corresponding to case $C[su, b]$. These are either carried out exactly as case $C[b, su]$, or with slight modifications, as shown in the following subcase example.

Case $C[su, b]$ with $b_{i_0} \leq a_{i_0}$:

Assumption: $x_{i_0 i_0} \neq b_{i_0}$.
Assumption (subcase): $x_{i_0 i_0} < b_{i_0}$ and $\nexists\, y_0 \neq i_0$ with $x_{y_0 i_0} > 0$.

Since $x_{i_0 i_0} < b_{i_0} \leq a_{i_0}$ (3.1.1) yields:
$\exists\, j_0 \neq i_0 :\ x_{i_0 j_0} > 0$.

We now define $\tilde{x}$ as:

$$\tilde{x}_{ij} = \begin{cases} x_{ij} + 1 & \text{if } (i,j) = (i_0, i_0), \\ x_{ij} - 1 & \text{if } (i,j) = (i_0, j_0), \\ x_{ij} & \text{otherwise.} \end{cases}$$

Obviously, $\tilde{x} \in X_f(a, b)$.

Using the distance properties (3.1.2)

$$\begin{aligned} \sum_{i,j=1}^{n} k_{ij}\, \tilde{x}_{ij} &= \sum_{i,j=1}^{n} k_{ij}\, x_{ij} + k_{i_0 i_0} - k_{i_0 j_0} \\ &= \sum_{i,j=1}^{n} k_{ij}\, x_{ij} - k_{i_0 j_0} \\ &< \sum_{i,j=1}^{n} k_{ij}\, x_{ij} \end{aligned} \tag{3.1.5}$$

follows.

(ii) follows from (i).

(iii)

In case $C[b, su]$:

$$\sum_{j=1}^{n} \hat{x}_{ij} = \sum_{j:j\neq i}^{n} x_{ij} + x_{ii} + x_{i\,n+1} = \sum_{j:j\neq i}^{n} x_{ij} + a_i - \sum_{j:i\neq j} x_{ij} = a_i,$$

$$\sum_{i=1}^{n} \hat{x}_{ij} = \sum_{i:i\neq j}^{n} x_{ij} + x_{jj} + x_{j\,n+1} = b_j + x_{j\,n+1},$$

is valid. This means $\hat{x} \in X_f(a, b + x_{n+1})$.

In case $C[su, b]$, $\hat{x} \in X_f(a, b + x_{n+1})$ follows directly from $\hat{x} = x$. ■

**Lemma 3.1.2.** *Let a transportation problem $TP^*(a, b)$ and $\tilde{b} \in \mathbb{Z}_+^n$ be given.*

*$x \in X_f(a, b)$, which additionally satisfies (3.1.4), with $\tilde{b} = b + x_{n+1}$ exists if and only if*

*$b_i \leq \tilde{b}_i \leq max\{a_i, b_i\}$ for $i = 1, \cdots, n$ in case $C[b, su]$ and*

*$min\{a_i, b_i\} \leq \tilde{b}_i \leq b_i$ for $i = 1, \cdots, n$ in case $C[su, b]$.*

***Proof.***

1. $(\Rightarrow)$ : Let an $x \in X_f(a, b)$, which additionally satisfies (3.1.4), be given.

   Case $C[b, su]$: $x_{n+1} \geq 0$ is valid according to (3.1.3),

   thus $b_i \leq \tilde{b}_i = b_i + x_{i\,n+1}$ for $i = 1, 2, \cdots, n$.

   Subcase $a_i \leq b_i$: $x_{ii} = a_i$ follows according to (3.1.4).

   Utilizing (3.1.3) we find:

$$\tilde{b}_i = b_i + x_{i\,n+1} = b_i + a_i - \sum_{j=1}^{n} x_{ij}, \tag{3.1.6}$$

$$\tilde{b}_i = b_i + x_{i\,n+1} = b_i - \sum_{j:j \neq i} x_{ij} \leq b_i = max\{a_i, b_i\}.$$

   Subcase $a_i \geq b_i$: $x_{ii} = b_i$ follows according to (3.1.4).

   According to (3.1.6)

$$\tilde{b}_i = b_i + x_{i\,n+1} = a_i - \sum_{j:j \neq i} x_{ij} \leq a_i = max\{a_i, b_i\}.$$

   Case $C[su, b]$: $x_{n+1} \leq 0$ is valid according to (3.1.3),

   thus $b_i \geq \tilde{b}_i = b_i + x_{i\,n+1}$ for $i = 1, 2, \cdots, n$ is also valid.

   Subcase $a_i \leq b_i$: $x_{ii} = a_i$ follows according to (3.1.4).

   Utilizing (3.1.3) we find:

$$\tilde{b}_i = b_i + x_{i\,n+1} = b_i - (b_i - \sum_{j=1}^{n} x_{ji}) = \sum_{j=1}^{n} x_{ji}, \tag{3.1.7}$$

$$\tilde{b}_i = b_i + x_{i\,n+1} = a_i + \sum_{j:j\neq i} x_{ji} \geq a_i = min\{a_i, b_i\}.$$

Subcase $a_i \geq b_i$: $x_{ii} = b_i$ follows according to (3.1.4).

According to (3.1.7)

$$\tilde{b}_i = b_i + x_{i\,n+1} = b_i + \sum_{j:j\neq i} x_{ji} \geq b_i = min\{a_i, b_i\}.$$

2. $(\Leftarrow)$ : Let $\tilde{b}$, which satisfies the corresponding relationships of the Lemma, and an $\tilde{x} \in X_f(a, \tilde{b})$, which additionally satisfies (3.1.4), be given.

Case $C[b, su]$: ( $b \leq \tilde{b}$ is implied in this case):

We set $x$, so that

$$x_{ij} = \begin{cases} \tilde{x}_{ii} - (\tilde{b}_i - b_i) & \text{if } i = j, \\ \tilde{x}_{ij} & \text{if } i \neq j. \end{cases}$$

Together with $\tilde{x} \in X_f(a, \tilde{b})$, the relations

$$\sum_{j=1}^{n} x_{ij} \leq \sum_{j=1}^{n} \tilde{x}_{ij} \leq a_i \quad \text{and}$$

$$\sum_{i=1}^{n} x_{ij} = \sum_{i=1}^{n} \tilde{x}_{ij} - (\tilde{b}_i - b_i) = \tilde{b}_i - (\tilde{b}_i - b_i) = b_i,$$

follow. Thus, $x \in X_f(a, b)$.

Furthermore,

$max\{a_i, b_i\} = a_i$ and, according to (3.1.4), $\tilde{x}_{ii} = min\{a_i, \tilde{b}_i\} = \tilde{b}_i$ follow

if $\tilde{b}_i > b_i$.

Thus, $x_{ii} = \tilde{x}_{ii} - (\tilde{b}_i - b_i) = \tilde{b}_i - (\tilde{b}_i - b_i) = b_i$.

$x_{ii} = \tilde{x}_{ii} = min\{a_i, \tilde{b}_i\} = min\{a_i, b_i\}$ follows directly if $\tilde{b}_i = b_i$.

$x$ also additionally satisfies (3.1.4).

In case $C[su, b]$ we set $x = \tilde{x}$, then it is yielded in a simple way that $x \in X_f(a, b)$ and $x$ additionally satisfies (3.1.4).

■

In the subsequent sections we use the following extension of the set of feasible solutions $X_{f(easible)}(a,b)$:

$$X_{f^e}(a,b) = \left\{ x \in \mathbb{Z}_+^n \times \mathbb{Z}_+^n \;\middle|\; \begin{array}{l} \sum\limits_{j=1}^{n} x_{ij} \leq a_i \;\forall i, \sum\limits_{i=1}^{n} x_{ij} \geq b_j \;\forall j \text{ in } C[b, su], \\ \sum\limits_{i=1}^{n} x_{ij} \leq b_j \;\forall j, \sum\limits_{j=1}^{n} x_{ij} = a_i \;\forall i \text{ in } C[su, b] \end{array} \right\} \tag{3.1.8}$$

From a theoretical point of view it is necessary to consider SDDP problems with extended sets of feasible solutions initially, in order to not lose the generality.

(In relation to the objective function of the $TP^*(a,b)$, obviously, the inequality

$\sum\limits_{i,j=1}^{n} k_{ij}\, x_{ij} \geq \sum\limits_{i,j=1}^{n} k_{ij}\, x_{ij}^*$ is valid,

where $x^*$ is an optimal solution of $TP^*(a,b)$ and $x \in X_{f^e}(a,b) \setminus X_f(a,b)$.)

## 3.2 Model Formulation

Before we formulate the mathematical models of SDDP problems, we discuss the SDDP problems from a practical point of view. Nevertheless, we will proceed with the introduction of the symbols of the mathematical models. (Refer also to [22] or [20].)

Parts of different types are produced by means of machines. For this purpose, the machines have to be converted into states, which are in accordance with the types of the parts. [2] Thereby, costs are incurred.

(It would also be conceivable to repeatedly place a fixed number of workers in different factories or buildinge sites.)

Each machine can be converted into each state.

[2] One may think, for example, of concrete moulds, see [19] and [22], Section 1.2.

The number of machines is denoted by $su$ and the number of types of parts by $n$.

If a machine is converted from a state $i$ ($i \in \{1, 2, \cdots, n\}$) into a state $j$ ($j \in \{1, 2, \cdots, n\}$) (which are in accordance with the types $i$ and $j$ of the parts), then the incurred cost are denoted by $k_{ij}$. Without loss of generality we can suppose that $(k_{ij})_{\substack{i=1,\ldots,n \\ j=1,\ldots,n}}$ satisfy the distance properties (see (3.1.2)).

Finally, $x_{ij}$ is the number of the machines which are converted from state $i$ into state $j$.

The production takes place in successive (equidistant) stages (periods).

In one stage one part can be produced (at most) by one machine.

In each stage a requirement of parts (of several types) is to be met.

At first, probability functions (denoted by $q_i$ ($i \in \{1, 2, \cdots, n\}$)) of the requirements are given.

The realizations of the requirements for a stage are known at the beginning of the stages (before the decision of conversions of machines has to be made).

The maximum number of produced parts in a stage of type $i$ ($i \in \{1, 2, ..., n\}$) is denoted by $k_{0_i}$.

The stages are numbered by subscript $t$, $t = 1, 2, \cdots$ (as in Section 2.1).

We use the same notations $w_t$, $w_t \in \mathbb{Z}_+^n$ for the random requirements of parts of different types and their realizations at stage $t$.

Principally, two cases in relation to the requirements (which correspond to the cases $C[b, su]$ and $C[su, b]$ from Section 3.1) have to be considered at each stage:

$$- C[w_t, su] : \text{ denotes the case } \sum_{i=1}^{n} w_{t,i} \leq su, \tag{3.2.1}$$

which means the requirements can be completely fulfilled.

$$- C[su, w_t] : \text{ denotes the case } \sum_{i=1}^{n} w_{t,i} \geq su \text{ [3]} \tag{3.2.2}$$

[3]In this instance it is not necessary to distinguish the case $\sum_{i=1}^{n} w_{t,i} = su$.

in which the requirement $w_t$ cannot be completely met, if $'' >''$ is valid. In this case the machines $su$ are not sufficient. However, in no case should availabilities be dissipated. [4]

The objective is to minimize the expected cost of the conversions of the machines over the stages altogether $(E(\sum_{t=1}^{N}\sum_{i,j=1}^{n} k_{ij}x_{t,ij}))$ (or in the case of an infinite horizon, the average expected cost per stage $\overline{\lim_{N\to\infty}}\frac{1}{N}E\left(\sum_{t=1}^{N}\sum_{i,j=1}^{n} k_{ij}x_{t,ij}\right)$).

(Therefore it must be decided which machine is to be converted to which state in each stage.)

Thus, SDDP problems are DA stochastic dynamic programming problems. [5]

The $su$ machines in the different states correspond to a partition of $su$.

Let $\tilde{s}_i$ denote the number of machines in state $i$.

Thus, the states of SDDP problems modelled as DA stochastic dynamic programming problems are (in general restricted ordered) partitions of integers which are written as vectors:

$\tilde{s} = (\tilde{s}_1, \tilde{s}_2, \cdots, \tilde{s}_n)$ where $\sum_{i=1}^{n} \tilde{s}_i = su$. [6]

The cost $\sum_{i,j=1}^{n} k_{ij}\, x_{t,ij}$, which are accrued if the machines in states $\tilde{s}_t$ at the beginning of stage $t$ are converted into states $\tilde{s}_{t+1}$ at the end of stage $t$ (and at the beginning of stage $t+1$) in order to satisfy the requirement $w_t$ of stage $t$, are denoted by $\hat{c}(\tilde{s}_t, w_t, \tilde{s}_{t+1})$ (adapted from Sections 2.1 and 2.3.2).

Obviously, $\hat{x}_t$, $\tilde{s}_t$ and $\tilde{s}_{t+1}$ (with $\hat{x}_{t,ij} = x_{t,ij}$ for $i \neq j$ and

[4] If a possibility for the storage of parts would be given, then the mathematical model could be extended as follows. In periods where case C[w, su] (with a "<"-sign) is present, additional parts could be produced as reserve for periods where the production capacity is not sufficient. However, the extension of the mathematical model would be determined by the concrete aims of the production process and the detailed storage possibilities.

[5] For dynamic models, where the requirements are deterministic, see [35].

[6] $s$ as a denotation of the states is used in adaption of Chapter 2. The additional symbol " ˜ " is attached to $s$ in order later to differentiate ordered from unordered partitions. Unordered partitions will be states of certain reduced SDDP problems (compare Section 3.4.2).

$\hat{x}_{t,ii} = min\,\{\tilde{s}_{t,i},\ \tilde{s}_{t+1,i}\}$) are correlated in such a manner that $\hat{x}_t$ is an optimal solution of $TP^*(\tilde{s}_t, \tilde{s}_{t+1})$.

We therefore use the phrase "distance optimal partitioning".

If $\tilde{s}_{t+1}$ has been chosen (such that the requirement $w_t$ has been met) then $\hat{c}(\tilde{s}_t, w_t, \tilde{s}_{t+1})$ itself does not directly depend on $w_t$. This means $\hat{c}(\tilde{s}_t, w_t, \tilde{s}_{t+1}) = \hat{c}(\tilde{s}_t, \tilde{s}_{t+1})$ (thus, (2.3.12) from Section 2.3.2 is valid).

Contrary to $(x_{t,ij})_{\substack{i=1,\dots,n\\ j=1,\dots,n}} \in X_{f^e}(\tilde{s}_t, w_t))$ (see (3.1.8)), the conditions $(x_{t,ij})_{\substack{i=1,\dots,n\\ j=1,\dots,n}} \in X_f(\tilde{s}_t, w_t)$ (see (3.1.1)) and (3.1.4) with $t = 1, 2, \cdots$ mean that the machines are not converted if they are not needed to produce parts for the requirements of stage $t$. [7] It seems plausible that under these conditions the expected cost of the stages altogether will not become higher.

However, we do not show the proof for this until later (see Lemma 3.3.8).

We will consider stationary models. This means $(k_{ij})_{\substack{i=1,\dots,n\\ j=1,\dots,n}}$ and the probability functions $q_i$ of the requirements are the same in all stages.

The corresponding **mathematical models** are now formulated as DA stochastic dynamic programming problems with random disturbances (as already mentioned above):

At first, let

$$\begin{aligned} &n \in \mathbb{N},\ n \geq 3,\ k_0 = (k_{0_1}, k_{0_2}, \dots, k_{0_n}) \text{ with } k_{0_i} \in \mathbb{N} \text{ and} \\ &su \in \mathbb{N} \text{ with } su < \sum_{i=1}^{n} k_{0_i} \end{aligned} \tag{3.2.3}$$

and

$$(k_{ij})_{\substack{i=1,\dots,n\\ j=1,\dots,n}} \in \mathbb{R}_+^n \times \mathbb{R}_+^n \text{ - the so-called "basic costs"} \tag{3.2.4}$$

(or in other words "distances")

be given. Thereby, let the basic costs satisfy the distance properties (3.1.2):

$$\begin{aligned} &k_{ii} = 0\ \forall\ i,\ k_{ij} > 0\ \forall\ i \neq j, \\ &k_{ij} + k_{jl} > k_{il}\ \forall i \neq j \neq l \text{ (triangle-inequality).} \end{aligned}$$

[7] Corresponding algorithms with this property are called "lazy algorithms", see Section 10.2.3 in [8].

(Refer also later to (3.3.1), (3.3.1a) and Lemma 3.3.1.)

**SDDP Problems as DA Stochastic Dynamic Programming Problems with Random Disturbances as in Section 2.1**

Since we now want to consider a stationary model, we do not need the subscript $t$ to indicate the state space, the disturbance space and the decision space.

Instead, we use the inferior indices $n, k_0$ and $su$ as characteristics of the concrete SDDP problems.

Furthermore, we use the following symbols:

$$\tilde{S}_{n;su;k_0} = \left\{ \tilde{s} \in \mathbb{Z}_+^n \mid 0 \leq \tilde{s} \leq k_0, \sum_{i=1}^{n} \tilde{s}_i = su \right\} \text{ - the state space} \tag{3.2.5}$$

(sets of the ordered restricted partitions with at most $n$ parts, whereby the elements are written as vectors),

$$\tilde{r} := |\tilde{S}_{n;su;k_0}|, \tag{3.2.5a}$$

$$B_{n;k_0} = \left\{ w \in \mathbb{Z}_+^n \mid 0 \leq w \leq k_0 \right\} \text{ - the disturbance space,} \tag{3.2.6}$$

[8]

$q : \mathbb{Z}_+^n \rightarrow [0,1]$ - probability functions

with

$$q(w) \neq 0 \Leftrightarrow w \in B_{n;k_0} \tag{3.2.7}$$

and

[8] We use the same notation for random vectors and their realizations.

$$A_{n;su;k_0}(\tilde{s},w)=$$

$$\left\{ x \in \mathbb{Z}_+^n \times \mathbb{Z}_+^n \left| \begin{array}{l} \sum_{j=1}^{n} x_{ij} \leq \tilde{s}_i \ \forall_i, \ \sum_{i=1}^{n} x_{ij} = w_j \ \forall_j \text{ in } C[w,su], \\ \sum_{j=1}^{n} x_{ij} = \tilde{s}_i \ \forall_i \ , \sum_{i=1}^{n} x_{ij} \leq w_j \ \forall_j \text{ in } C[su,w], \end{array} \right. \quad x_{ii} = \min\{\tilde{s}_i, w_i\} \ \forall_i \right\}$$

- the decision sets. (3.2.8)

($x \in A_{n;su;k_0}(\tilde{s},w)$ are thus feasible solutions of the transportation problems $TP^*(s,w)$ which additionally satisfy (3.1.4).)

We now consider the DA model (refer to Section 2.1):

A stationary policy

$$F = \{x_1(s_1,w_1), x_2(s_2,w_2), \ldots, x_N(s_N,w_N)\}$$

(or

$$F = \{x_1(s_1,w_1), x_2(s_2,w_2), \ldots\})$$

is to be found so that

$$E\left\{\sum_{t=1}^{N}\sum_{i,j=1}^{n} k_{ij}x_{t,ij}\right\} \to min$$

in the case of a finite horizon or

$$\overline{\lim_{N\to\infty}} \frac{1}{N} E\left\{\sum_{t=1}^{N}\sum_{i,j=1}^{n} k_{ij}x_{t,ij}\right\} \to inf,$$

in the case of an infinite horizon

subject to the constraints

$$x_t \in A_{n;su;k_0}(\tilde{s}_t, w_t),$$

$$\tilde{s}_{t+1,i} = \tilde{s}_{t,i} - \sum_{j=1}^{n} x_{t,ij} + w_{t,i} \text{ for } i = 1, \cdots, n \quad \text{in } C[w_t, su] \tag{3.2.9}$$

$$\tilde{s}_{t+1,j} = \sum_{i=1}^{n} x_{t,ij} \text{ for } j = 1, \cdots, n \quad \text{in } C[su, w_t]. \tag{3.2.10}$$

(In reference to (3.2.9) and (3.2.10) compare (3.1.6) and (3.1.7) in the proof of Lemma 3.1.2, where $\tilde{b}_i = \tilde{s}_{t+1,i}$, $a_i = \tilde{s}_{t,i}$ and $b_i = w_{t,i}$.

According to (3.2.8) and (3.2.9) or (3.2.10):

$$\sum_{i=1}^{n} \tilde{s}_{t+1,i} = \sum_{i=1}^{n} (\tilde{s}_{t,i} - \sum_{j=1}^{n} x_{t,ij} + w_{t,i}) = su - \sum_{j=1}^{n} \sum_{i=1}^{n} x_{t,ij} + \sum_{i=1}^{n} w_{t,i}$$

$$= su - \sum_{j=1}^{n} w_{t,j} + \sum_{i=1}^{n} w_{t,i} = su \quad \text{in the case } C[w_t, su] \text{ or}$$

$$\sum_{j=1}^{n} \tilde{s}_{t+1,j} = \sum_{j=1}^{n} (\sum_{i=1}^{n} x_{t,ij}) = \sum_{i=1}^{n} (\sum_{j=1}^{n} x_{t,ij}) = \sum_{i=1}^{n} \tilde{s}_{t,i} = su \quad \text{in the case } C[su, w_t]$$

are valid, of course.)

**Remarks 3.2.1.** *The "certainty equivalence principle" (see Section 2.2, Theorem 2.2.1) is not valid for SDDP problems. The objective function is not quadratic and the different kind of restrictions (3.2.9) and (3.2.10) in the cases $C[w_t, su, w_t]$ and $C[su, w_t]$ cannot be brought into a standardized form (2.2.1).*

We can convert the above problem into a stationary DA Markov decision process for $N = \infty$ as in Section 2.3.2:

**SDDP Problems as DA Markov Decision Processes as in Section 2.3.2**

The state space is of course $\tilde{S}_{n;su;k_0}$ from (3.2.5).

In order to construct $A^M, P$ and $\gamma$ of the DA MDP the DA decision sets $\hat{A}_{n;su;k_0}(\tilde{s}, w)$, the DA decision functions $\hat{d}$ and the internal costs $\hat{c}(\tilde{s}, w, \tilde{s}')$

(see (2.1.4), (2.1.5), (2.1.6) at the end of Section 2.1) have to be used (as in Section 2.3.2).

The DA decision sets $\hat{A}_{n;su;k_0}(\tilde{s}, w)$ include the states $\tilde{s}'$ (see (2.1.4)), which follow from (3.2.9) or (3.2.10) for $x \in A_{n;su;k_0}(\tilde{s}, w)$ and given $\tilde{s} \in \tilde{S}_{n;su;k_0}$, $w \in B_{n;k_0}$.

Lemma 3.1.2 (with $a = \tilde{s},\ b = w$ and $\tilde{b} = \tilde{s}'$) yields:

$$\hat{A}_{n;su;k_0}(\tilde{s}, w) = \left\{ \tilde{s}' \in \tilde{S}_{n;su;k_0} \;\middle|\; \begin{array}{l} w_i \le \tilde{s}'_i \le max\{\tilde{s}_i, w_i\},\ i = 1, \cdots, n, \text{ in } C[w, su] \\ min\{\tilde{s}_i, w_i\} \le \tilde{s}'_i \le w_i,\ i = 1, \cdots, n, \text{ in } C[su, w] \end{array} \right\} \tag{3.2.11}$$

$$\text{for } \tilde{s} \in S_{n;su;k_0},\ w \in B_{n;k_0}.$$

(In Section 3.3 we will also see investigations of the DA decision sets.)

DA decision functions $\hat{d}_t$, which are based on the DA decision sets, follow according to (2.1.6). (In addition, the set of DA decision functions is the set $\hat{D}_t$ according to Definition 2.1.1.)

The internal costs (see (2.1.5)) are calculated in the following way:

Let $\tilde{s}' \in \hat{A}_{n;su;k_0}(\tilde{s}, w)$ be determined by a DA decision function for given $\tilde{s} \in \tilde{S}_{n;su;k_0}$ and realized $w \in B_{n;k_0}$.

Then, as mentioned above, the internal cost can be compute as the optimal value of the (balanced) transportation problem $TP^*(\tilde{s}, \tilde{s}')$:

$$\hat{c}(\tilde{s}, \tilde{s}') \;=\; min \left\{ \sum_{i,j=1}^{n} k_{ij}\hat{x}_{ij} \,\middle|\, \sum_{j=1}^{n} \hat{x}_{ij} = \tilde{s}_i, \sum_{i=1}^{n} \hat{x}_{ij} = \tilde{s}'_j, \hat{x}_{ij} \in \mathbb{Z}_+ \right\}. \tag{3.2.12}$$

(The symbol " ˆ " is used in order to formally differentiate between (optimal) solutions $\hat{x}$ of the balanced transportation problems $TP^*(\tilde{s}, \tilde{s}')$ (see (3.2.12)) and (feasible) solutions $x$ of the non-balanced transportation problems $TP^*(\tilde{s}, w)$, which additionally satisfy (3.1.4) (see also (3.2.8)).

Thereby, $\hat{x}_{ij}$ and $x_{ij}$ for $i \neq j$ express the number of machines which are converted from state $i$ into state $j$ for the production of $w$ parts of type $j$. That means $x_{ij} = \hat{x}_{ij}$ for $i \neq j$.

For the rest, as is easily seen, $\hat{x}$ yields a feasible $x$ in the following way:

$$x_{ij} = \begin{cases} \hat{x}_{ij} & \text{if } i \neq j \\ \hat{x}_{ii} - \max\{(\tilde{s}'_i - w_i), 0\} & \text{if } i = j \end{cases} \quad \text{in the case } C[w, su] \text{ or}$$

$$x = \hat{x} \qquad \text{in the case } C[su, w].)$$

(Further properties of the internal costs can be found in Section 3.3.)

Now, the sets of decision spaces $A^M(\tilde{S})$, the matrices of transition probabilities $P$ and the average (one-step) reward functions $\gamma$ of the DA Markov decision processes can be constructed by means of $\hat{d}$ and $\hat{c}$ - as in Section 2.3.2 (see (2.3.6), (2.3.7) and (2.3.8)).

Whereby the matrices of transition probabilities are called (general) partitions-requirements-matrices (PRM).

For SDDP problems as DA MDPs optimal policies always exist, since (2.3.3) is valid:

If $w = \tilde{s}'$ $(\in \tilde{S}_{n;su;k_0} \subseteq B_{n;k_0})$ then, obviously, $\hat{A}_{n;su;k_0}(\tilde{s}, \tilde{s}') = \{\tilde{s}'\}$ according to (3.2.11) and

$$p(\tilde{s}'|\tilde{s}, d) = \sum_{w:\tilde{s}'=\hat{d}(\tilde{s},w)} q(w) \geq q(\tilde{s}') \neq 0 \tag{3.2.13}$$

$$(q(\tilde{s}') \neq 0 \text{ according to (3.2.7)})$$

follows according to (2.3.7) for any $d$.

The computation of average (one-step) reward functions $\gamma$ is also laborious since transportation problems have to be solved in order to calculate the internal costs $\hat{c}(\tilde{s}, \tilde{s}')$ $(\tilde{s} \in \tilde{S}_{n;su;k_0},\ \tilde{s}' \in \tilde{S}_{n;su;k_0})$, which are needed for the computation of $\gamma$ (see the first equation of (2.3.8)).

In Section 3.3 (Lemma 3.2.10) we will see that SDDP problems are DA MDPs with distance properties.
[9]

[9] A representation of SDDP problems as linear programming problems can be found in [22], Appendix A 1.

## 3.3 Properties of SDDP Problems and their Characteristic Quantities and Conversion Numbers

**In reference to the Possibility of the Restriction of the Basic Costs**

Without loss of generality, it is sufficient for SDDP problems modelled as DA MDPs to assume basic costs $(k_{ij})_{\substack{i=1,\dots,n\\ j=1,\dots,n}}$ which satisfy (3.1.2) with either

$$0 < k_{ij} \leq 1 \quad \text{for } i \neq j \tag{3.3.1}$$

or

$$k_{ij} \geq 1 \text{ for } i \neq j. \tag{3.3.1a}$$

This statement follows from Lemma 3.3.1 with $\alpha = \frac{1}{\max\limits_{i,j:i\neq j} k_{ij}}$ or $\alpha' = \frac{1}{\min\limits_{i,j:i\neq j} k_{ij}}$.

**Lemma 3.3.1.** *If two SDDP problems (modelled as DA MDPs) are only different in the basic cost $k'$ and $k''$ in such a way that $k'' = \alpha \cdot k'$ for some $\alpha > 0$, then the optimal policies of the SDDP problems are the same.*

***Proof.*** Obviously, $\hat{c}''_{fl} = \alpha\, \hat{c}'_{fl}$ for any $f$ and $l$ and furthermore $\gamma''(\tilde{s}^f, d) = \alpha\, \gamma'(\tilde{s}^f, d)$ for any $\tilde{s}^f$ and $d$ according to (2.3.8) (first equation).

Consideration of the objective function (2.3.2) completes the proof. ■

**The Conversion Number**

Carrying on,

$$U(\tilde{s}^f, \tilde{s}^l) := \sum_{\substack{i,j=1\\ i\neq j}}^{n} \hat{x}_{ij} \tag{3.3.2}$$

denote the numbers of all (real) conversions (in a stage), if $\tilde{s}^f$ are converted into $\tilde{s}^l$ by means of $\hat{x} \in X_f(\tilde{s}^f, \tilde{s}^l)$, which also satisfy (3.1.4).

The following Lemma shows that $U(\tilde{s}^f, \tilde{s}^l)$ for all such $\hat{x}$ are identical.

**Lemma 3.3.2.** *Let $\{\tilde{s}^f, \tilde{s}^l\} \subseteq \tilde{S}_{n;su;k_0}$, basic costs $(k_{ij})_{\substack{i=1,\dots,n\\ j=1,\dots,n}}$ (satisfying the distance properties (3.1.2)) and any $\hat{x} \in X_f(\tilde{s}^f, \tilde{s}^l)$ which additionally*

*satisfies (3.1.4), be given. (In particular, an optimal $\hat{x} \in X_f(\tilde{s}^f, \tilde{s}^l)$ satisfies (3.1.4), according to Lemma 3.1.1(ii)).*

*Then,*

$$(a)\ \ U(\tilde{s}^f, \tilde{s}^l) = \tfrac{1}{2} \sum_{i=1}^{n} \mid \tilde{s}_i^f - \tilde{s}_i^l \mid = \sum_{i=1}^{n} \max\{0, \tilde{s}_i^l - \tilde{s}_i^f\}$$
$$= -\sum_{i=1}^{n} \min\{0, \tilde{s}_i^f - \tilde{s}_i^l\},$$

$$(b)\ \ U(\tilde{s}^f, \tilde{s}^l) = U(\tilde{s}^l, \tilde{s}^f).$$

***Proof.*** Initially,

$$U(\tilde{s}^f, \tilde{s}^l) = \sum_{\substack{i,j=1 \\ i \neq j}}^{n} \hat{x}_{ij} = \sum_{i,j=1}^{n} \hat{x}_{ij} - \sum_{i=1}^{n} \hat{x}_{ii} = \sum_{i=1}^{n} \tilde{s}_j^f - \sum_{i=1}^{n} \hat{x}_{ii} = su - \sum_{i=1}^{n} \hat{x}_{ii},$$

and

$$\tfrac{1}{2} \sum_{i=1}^{n} \mid \tilde{s}_i^f - \tilde{s}_i^l \mid = \tfrac{1}{2} \sum_{i=1}^{n} (\tilde{s}_i^f - \min\{\tilde{s}_i^f, \tilde{s}_i^l\} + \tilde{s}_i^l - \min\{\tilde{s}_i^f, \tilde{s}_i^l\})$$
$$= \tfrac{1}{2} \sum_{i=1}^{n} (\tilde{s}_i^f + \tilde{s}_i^l) - \sum_{i=1}^{n} \min\{\tilde{s}_i^f, \tilde{s}_i^l\} = su - \sum_{i=1}^{n} \min\{\tilde{s}_i^f, \tilde{s}_i^l\}$$

are valid. Then (a) follows from (3.1.4)

$$\sum_{i=1}^{n} \hat{x}_{ii} = \sum_{i=1}^{n} \min\{\tilde{s}_i^f, \tilde{s}_i^l\}.$$

$(b)$ is shown by $U(\tilde{s}^f, \tilde{s}^l) = \frac{1}{2} \sum_{i=1}^{n} \mid \tilde{s}_i^f - \tilde{s}_i^l \mid = U(\tilde{s}^l, \tilde{s}^f)$. ■

**Lemma 3.3.3.** *Let $\{\tilde{s}^f, \tilde{s}^l\} \subseteq \tilde{S}_{n;su;k_0}$ be given.*

*Then*

$$U(\tilde{s}^f, \tilde{s}^l) = \sum_i \max\{0, w_i - \tilde{s}_i^f\} \ \ \forall\ \tilde{s}^l \in \hat{A}_{n;su;k_o}(\tilde{s}^f, w) \text{ in the case } C[w, su],$$

$$U(\tilde{s}^f, \tilde{s}^l) = \sum_i \max\{0, \tilde{s}_i^f - w_i\} \ \ \forall\ \tilde{s}^l \in \hat{A}_{n;su;k_o}(\tilde{s}^f, w) \text{ in the case } C[su, w].$$

***Proof.*** In the case $C[w, su]$

$$\begin{aligned} U(\tilde{s}^f, \tilde{s}^l) &= \sum_i \max\{0, \tilde{s}_i^l - \tilde{s}_i^f\} = \sum_{i:\tilde{s}_i^f < w_i} (\tilde{s}_i^l - \tilde{s}_i^f) \\ &= \sum_{i:\tilde{s}_i^f < w_i} (w_i - \tilde{s}_i^f) = \sum_i \max\{0, w_i - \tilde{s}_i^f\} \end{aligned}$$

follows from Lemma 3.3.2 and (3.2.11).

The proof in case $C[su, w]$ is analogous. ■

**Regarding the Internal Costs**

- The internal costs of SDDP problems fulfil (2.3.12) from Section 2.3.2 as is discussed in Section 3.2 (see also (3.2.12)).

**Theorem 3.3.4.** *Let $\tilde{S}_{n;su;k_0} = \{\tilde{s}^1, \tilde{s}^2, \cdots, \tilde{s}^r, \}$ and the basic costs $(k_{ij})_{\substack{i=1,\dots,n \\ j=1,\dots,n}}$ which satisfy the distance properties (3.1.2) be given.*

*Then, the internal costs $(\hat{c}_{fl})_{\substack{f=1,\dots,r \\ l=1,\dots,r}} := (\hat{c}(s^f, s^l)_{\substack{f=1,\dots,r \\ l=1,\dots,r}}$ (see (3.2.12)) fulfil the following corresponding distance properties (with an additional equals sign)*

$$\begin{aligned} &\hat{c}_{ff} = 0 \text{ for } f = 1, ..., r, \\ &\hat{c}_{fl} + \hat{c}_{lv} \geq \hat{c}_{fv} \text{ for } f \neq l \neq v. \end{aligned}$$

***Proof.*** Initially, let $TP^*(a, b)$ be a balanced transportation problem with availabilities $a$, requirements $b$ and $\sum_{i=1}^n b_i = \sum_{i=1}^n a_i$.

Along with the symbol

$$X_f(a, b) = \left\{ x \in \mathbb{Z}_+^n \times \mathbb{Z}_+^n \;\middle|\; \sum_{j=1}^n x_{ij} = a_i \; \forall i, \sum_{i=1}^n x_{ij} = b_j \forall j \right\} \tag{3.3.3}$$

for the set of feasible solutions, as in Section 3.1, we use also the symbol

$$X_{opt}(a, b) = \left\{ x \in X_f(a, b) \;\middle|\; \sum_{i,j=1}^n k_{ij} x_{ij} \leq \sum_{i,j=1}^n k_{ij} x'_{ij} \text{ for any } x' \in X_f(a, b) \right\} \tag{3.3.4}$$

for the set of optimal solutions.

Now, let $\hat{x}^1 \in X_{opt}(\tilde{s}^f, \tilde{s}^l)$ and $\hat{x}^2 \in X_{opt}(\tilde{s}^l, \tilde{s}^v)$. (Thus, $\hat{c}_{fl} = \sum\limits_{i,j=1}^{n} k_{ij}\hat{x}^1_{ij}$ and $\hat{c}_{lw} = \sum\limits_{i,j=1}^{n} k_{ij}\hat{x}^2_{ij}$.)

Then,

$$\tilde{s}^f_i = \sum_y \hat{x}^1_{iy}, \tag{*1}$$

$$\tilde{s}^v_j = \sum_y \hat{x}^2_{yj}, \tag{*2}$$

$$\text{and } \tilde{s}^l_y = \sum_i \hat{x}^1_{iy} = \sum_j \hat{x}^2_{yj}. \tag{*3}$$

Furthermore, optimal solutions $(\triangle_{iyj})_{\substack{i=1,\dots,n\\ j=1,\dots,n}}$ of the $n$ transportation problems with availabilities $\left(\hat{x}^1_{iy}\right)_{i=1,\dots,n}$ and requirements $\left(\hat{x}^2_{yj}\right)_{j=1,\dots,n}$ $(y = 1, ..., n)$ exist because of $(*3)$.

These solutions also fulfil

$$\sum_j \triangle_{iyj} = \hat{x}^1_{iy} \tag{*4}$$

$$\sum_i \triangle_{iyj} = \hat{x}^2_{yj}. \tag{*5}$$

We set

$$\hat{x}^3_{ij} := \sum_y \triangle_{iyj}.$$

$$\sum_j \hat{x}^3_{ij} = \sum_y \sum_j \triangle_{iyj} = \sum_y \hat{x}^1_{iy} = \tilde{s}^f_i,$$
$$\sum_i \hat{x}^3_{ij} = \sum_y \sum_i \triangle_{iyj} = \sum_y \hat{x}^2_{yj} = \tilde{s}^v_j$$

follow from $(*4)$, $(*1)$, $(*5)$ and $(*2)$.

This means $\hat{x}^3 \in X_f(\tilde{s}^f, \tilde{s}^v)$.

(3.1.2) then leads to

$$\begin{aligned} & k_{iy} + k_{yj} \;>\; k_{ij} \quad (i \neq y \neq j) \\ \sum_{i,y,j} \; & k_{iy} \triangle_{iyj} + \sum_{i,y,j} k_{yj} \triangle_{iyj} \;\geq\; \sum_{i,y,j} k_{ij} \triangle_{iyj} \\ \sum_{i,y} \; & k_{iy} \hat{x}^1_{iy} + \sum_{j,y} k_{yj} \hat{x}^2_{yj} \;\geq\; \sum_{i,j} k_{ij} \hat{x}^3_{ij} \end{aligned} \tag{$*6$}$$

$$\hat{c}_{fl} + \hat{c}_{lv} \;\geq\; \sum_{i,j} k_{ij} \hat{x}^3_{ij} \;\geq\; \hat{c}_{fv}. \tag{$*7$}$$

Thereby, the equals signs in (*6) and in the first inequality of (*7) are valid if and only if $\triangle_{iyj} > 0$ with $i \neq y \neq j$ do not exist (refer also to the following Example 3.3.1). ■

Since the internal costs of SDDP problems fulfil (2.3.12), Lemma 3.3.4 implies the following statement.

**Lemma 3.3.5.** *SDDP problems (modelled as DA MDPs) satisfy the distance properties (2.3.28) and (2.3.29).*

The following Lemma is needed in Section 3.5 with regard to cost-parametric considerations of SDDP problems.

**Lemma 3.3.6.** *Let $\tilde{S}_{n;su;k_0}$ and basic costs with (w. l. o. g.) $k_{ij} \geq 1$ for $i \neq j$ (see (3.3.1a)) be given. Let $(k_{ij})_{\substack{i=1,\ldots,n \\ j=1,\ldots,n}}$ satisfy (3.1.2) and let $(k_{ij} - \delta_{ij})_{\substack{i=1,\ldots,n \\ j=1,\ldots,n}}$, where $\delta_{ij} = \begin{cases} 1 \text{ if } i \neq j, \\ 0 \text{ if } i = j \end{cases}$ satisfy modified distance properties in such a way that additional equals signs are also allowed in the inequalities of (3.1.2).*

*Then,*

$$\begin{aligned} & \hat{c}(\tilde{s}, \tilde{s}', (\delta_{ij} + \vartheta\,(k_{ij} - \delta_{ij}))_{\substack{i=1,\ldots,n \\ j=1,\ldots,n}}) \\ & \quad = \hat{c}(\tilde{s}, \tilde{s}', (\delta_{ij})_{\substack{i=1,\ldots,n \\ j=1,\ldots,n}}) + \vartheta\, \hat{c}(\tilde{s}, \tilde{s}', (k_{ij} - \delta_{ij})_{\substack{i=1,\ldots,n \\ j=1,\ldots,n}}) \\ & \quad = U(\tilde{s}, \tilde{s}') + \vartheta\, \hat{c}(\tilde{s}, \tilde{s}', (k_{ij} - \delta_{ij})_{\substack{i=1,\ldots,n \\ j=1,\ldots,n}}) \end{aligned}$$

*are valid for $\vartheta \in [0; 1]$, where $\hat{c}(\tilde{s}, \tilde{s}', (k_{ij})_{\substack{i=1,\ldots,n \\ j=1,\ldots,n}})$ denotes the optimal value*

*of the transportation problem with availabilities $\tilde{s}$, requirements $\tilde{s}'$ and with costs $(k_{ij})_{\substack{i=1,\dots,n\\ j=1,\dots,n}}$ and so on.*

***Proof.*** At first, we show that an optimal solution $x^0$ of the transportation problem $TP^*(\tilde{s},\tilde{s}')$ with costs $(k_{ij}-\delta_{ij})_{\substack{i=1,\dots,n\\ j=1,\dots,n}}$, which additionally satisfies (3.1.4), exists:

If $x$ is an optimal solution of this transportation problem which does not satisfy (3.1.4), then we apply the construction step from the proof of Lemma 3.1.1(i) (w.l.o.g. case $C[b,su]$) to $x$, where an equation instead of the inequality in (3.1.5) is right. If neccessary, repeatedly applying this construction step will finally yield the optimal solution $x^0$ of the transportation problem, which satisfies (3.1.4).

Now, let $x$ be a feasible solution of the transportation problem with availabilities $\tilde{s}$, requirements $\tilde{s}'$ and costs $(k_{ij}-\delta_{ij})_{\substack{i=1,\dots,n\\ j=1,\dots,n}}$, which additionally satisfies (3.1.4).

$$\begin{aligned}
&\hat{c}(\tilde{s},\tilde{s}',(k_{ij}-\delta_{ij})_{\substack{i=1,\dots,n\\ j=1,\dots,n}}) = \sum_{i,j} x^0_{ij}\,(k_{ij}-\delta_{ij}) \le \sum_{i,j} x_{ij}\,(k_{ij}-\delta_{ij}),\\
&U(\tilde{s},\tilde{s}')+\vartheta\,\sum_{i,j} x^0_{ij}\,(k_{ij}-\delta_{ij}) \le U(\tilde{s},\tilde{s}')+\vartheta\,\sum_{i,j} x_{ij}\,(k_{ij}-\delta_{ij})
\end{aligned}$$

follow and furthermore, by means of (3.3.2),

$$\begin{aligned}
&\sum_{i,j} x^0_{ij}\,\delta_{ij}+\vartheta\,\sum_{i,j} x^0_{ij}\,(k_{ij}-\delta_{ij}) \le \sum_{i,j} x_{ij}\,\delta_{ij}+\vartheta\,\sum_{i,j} x_{ij}\,(k_{ij}-\delta_{ij}),\\
&\sum_{i,j} x^0_{ij}\,(\delta_{ij}+\vartheta\,(k_{ij}-\delta_{ij})) \le \sum_{i,j} x_{ij}\,(\delta_{ij}+\vartheta\,(k_{ij}-\delta_{ij}))\,.
\end{aligned}$$

Hence, (and with attention to Lemma 3.1.1(ii)) $x^0$ is also an optimal decision of transportation problems $TP^*(\tilde{s},\tilde{s}')$ with cost $(\delta_{ij}+\vartheta\,(k_{ij}-\delta_{ij}))$ where $\vartheta\in[0;1]$. Since the objective function of a transportation problem is also linear in the costs

$$\begin{aligned}
&\hat{c}(\tilde{s},\tilde{s}',(\delta_{ij}+\vartheta\,(k_{ij}-\delta_{ij}))_{\substack{i=1,\dots,n\\ j=1,\dots,n}})\\
&\qquad = \hat{c}(\tilde{s},\tilde{s}',(\delta_{ij})_{\substack{i=1,\dots,n\\ j=1,\dots,n}})+\vartheta\,\hat{c}(\tilde{s},\tilde{s}',(k_{ij}-\delta_{ij})_{\substack{i=1,\dots,n\\ j=1,\dots,n}})\\
&\qquad = U(\tilde{s},\tilde{s}')+\vartheta\,\hat{c}(\tilde{s},\tilde{s}',(k_{ij}-\delta_{ij})_{\substack{i=1,\dots,n\\ j=1,\dots,n}})
\end{aligned}$$

is valid (for the last equation see also (3.3.2)). ■

**Regarding the DA Decisions Sets**

As already mentioned in Section 3.2, contrary to $(x_{t,ij})_{\substack{i=1,\dots,n\\ j=1,\dots,n}} \in X_{f^e}(\tilde{s}_t, w_t)$ (the extended set, see (3.1.8)), $(x_{t,ij})_{\substack{i=1,\dots,n\\ j=1,\dots,n}} \in X_f(\tilde{s}_t, w_t)$ and additionally satisfying (3.1.4), means that machines are not converted if they are not needed to produce parts for the requirements of a stage $t$ (see also the following Lemma 3.3.7 and (3.2.11)).

In Theorem 3.3.8 we will show that the expected cost of the conversions of the machines over the stages altogether (or in the case of an infinite horizon, the average expected cost per stage) do not become smaller if $(x_{t,ij})_{\substack{i=1,\dots,n\\ j=1,\dots,n}} \in X_{f^e}(\tilde{s}_t, w_t),\ t = 1, 2, \cdots$ are allowed in place of $(x_{t,ij})_{\substack{i=1,\dots,n\\ j=1,\dots,n}} \in X_f(\tilde{s}_t, w_t)$ and which additionally satisfy (3.1.4).

Initially, we make relationships with corresponding decision sets (see (3.2.11)):

Let $\tilde{s} \in \tilde{S}_{n;su;k_0}$ and $w \in B_{n;k_0}$. Obviously, for $\tilde{s}' \in \tilde{S}_{n;su;k_0}$ an

$$(x_{ij})_{\substack{i=1,\dots,n\\ j=1,\dots,n}} \in X_{f^e}(\tilde{s}, w) \text{ with } \tilde{s}'_i = \begin{cases} \sum\limits_{i=1}^{n} x_{ij} + \tilde{s}_i - \sum\limits_{j=1}^{n} x_{ij} & \text{in } C[w, su], \\ \sum\limits_{j=1}^{n} x_{ji} & \text{in } C[su, w] \end{cases}$$

exists if and only if $\tilde{s}'$ is an element of the following set:

$$\hat{A}'_{n;su;k_0}(w) = \left\{ \tilde{s}' \in \tilde{S}_{n;su;k_0} \;\middle|\; \begin{array}{l} \tilde{s}' \geq w \text{ in } C[w, su], \\ \tilde{s}' \leq w \text{ in } C[su, w] \end{array} \right\}. \tag{3.3.5}$$

[10]

Together with (3.2.11)

$$\hat{A}_{n;su;k_0}(\tilde{s}, w) \subseteq \hat{A}'_{n;su;k_0}(w) \text{ for any } \tilde{s} \in \tilde{S}_{n;su;k_0},\ w \in B_{n;k_0} \tag{3.3.6}$$

follows.

Thus, $\hat{A}'_{n;su;k_0}(w)$ are extensions of DA decision sets $\hat{A}_{n;su;k_0}(\tilde{s}, w)$.

[10] $\tilde{s}' = w$ follows in the special case that $w \in \tilde{S}_{n;su;k_0}$.

**Lemma 3.3.7.** *Let* $\tilde{s} \in \tilde{S}_{n;su;k_0}$ *and* $w \in B_{n;k_o}$ *be given.*

*If* $\tilde{s}'' \in \hat{A}'_{n;su;k_0}(w) \setminus \hat{A}_{n;su;k_0}(\tilde{s}, w)$ *then* $U(\tilde{s}, \tilde{s}'') > U(\tilde{s}, \tilde{s}')$ *for any* $\tilde{s}' \in \hat{A}_{n;su;k_0}(\tilde{s}, w)$.

***Proof.*** Initially, we consider the case $\sum w_j < su$.

For $\tilde{s}'' \in \hat{A}'_{n;su;k_0}(w) \setminus \hat{A}_{n;su;k_0}(\tilde{s}, w)$ an $i_0$ exists so that $\tilde{s}''_{i_0} > \max\{\tilde{s}_{i_0}, w_{i_0}\}$, according to (3.2.11) and (3.3.5).

$\tilde{s}''_{i_0} - \tilde{s}_{i_0} > \max\{0, w_{i_0} - \tilde{s}_{i_0}\}$ follows.

Lemma 3.3.2 and Lemma 3.3.3 then yield:

$$U(\tilde{s}, \tilde{s}'') = \sum_{i=1}^{n} \max\{0, \tilde{s}''_i - \tilde{s}_i\} > \sum_{i=1}^{n} \max\{0, w_i - \tilde{s}_i\} = U(\tilde{s}, \tilde{s}')$$

for any $\tilde{s}' \in \hat{A}_{n;su;k_0}(\tilde{s}, w)$.

The case $\sum w_j > su$ is proved analogously.

In the case that $\sum w_j = su$, no $\tilde{s}'' \in \hat{A}'_{n;su;k_0}(w) \setminus \hat{A}_{n;su;k_0}(\tilde{s}, w)$ exists since $\hat{A}'_{n;su;k_0}(w) = \hat{A}_{n;su;k_0}(\tilde{s}, w) = \{w\}$ (see (3.2.11) and (3.3.5)). ■

We now consider the following simple example:

**Example 3.3.1.** *Let a SDDP problem with*

$$n = 2,\ k_0 = (3,3),\ su = 4,\ (k_{ij})_{\substack{i=1,2\\ j=1,2}} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

*be given.*

$$\hat{A}_{2;4;3}(\tilde{s}, w) = \{\tilde{s}' = \begin{pmatrix} 2 \\ 2 \end{pmatrix}\} \text{ for } \tilde{s} = \begin{pmatrix} 1 \\ 3 \end{pmatrix},\ w = \begin{pmatrix} 2 \\ 1 \end{pmatrix} \text{ (see (3.2.11)).}$$

*Furthermore,* $\tilde{s}'' = \begin{pmatrix} 3 \\ 1 \end{pmatrix} \in \hat{A}'_{2;4;3}(w)$ *(see (3.3.5)), for example.*

$U(\tilde{s}, \tilde{s}') = 1$ *and* $U(\tilde{s}, \tilde{s}'') = 2$ *are the corresponding conversions numbers (see Lemma 3.3.2).*

*It seems intuitively evident that a transition from* $\tilde{s}$ *and* $w$ *to* $\tilde{s}''$ *could not be a decision of an optimal solution. More is done for the concerned stage than is necessary. (This statement is proved in the following Lemma.)*

*If we compute the internal cost (according to (3.2.12)):* $\hat{c}(\tilde{s}, \tilde{s}') = 1$, $\hat{c}(\tilde{s}, \tilde{s}'') = 2$ *and* $\hat{c}(\tilde{s}', \tilde{s}'') = 1$, *then*

$$\hat{c}(\tilde{s}', \tilde{s}'') + \hat{c}(\tilde{s}, \tilde{s}') \geq \hat{c}(\tilde{s}, \tilde{s}'')$$

*is valid in conformance with Theorem 3.3.4 and moreover*

$$\hat{c}(\tilde{s}', \tilde{s}'') + \hat{c}(\tilde{s}, \tilde{s}') = \hat{c}(\tilde{s}, \tilde{s}'').$$

*This shows that a transition from* $\tilde{s}$ *and* $w$ *to* $\tilde{s}''$ *is in reality not necessary for optimal solutions according to Theorem 2.3.13.* [11]

*(Thus, the exclusion of the "=" sign after Theorem 2.3.13a leads to the exclusion of an entire set of non-essential decisions.)*

This is also generally correct for SDDP problems:

**Theorem 3.3.8.** *Let a SDDP problem, as in Section 3.2, be given and let SDDP' be the corresponding "extended" SDDP problem. This means that the extended decision sets* $\hat{A}'_{n;su;k_0}(w_t)$ *are used in place of* $\hat{A}_{n;su;k_0}(\tilde{s}_t, w_t)$ *for any* $\tilde{s}_t$, $w_t$ *and* $t$.

*(a) The extended SDDP problem satisfies the distance properties (2.3.28) and (2.3.35).*
$\hat{A}_{n;su;k_0}(\tilde{s}, w)$ *are smaller DA decision sets (see Definition 2.3.4).*

*(b) An optimal policy of the SDDP problem is also an optimal policy of the corresponding SDDP'.*

***Proof.***
(a) That the distance properties (2.3.28) and (2.3.35) are fulfilled follows directly from Theorem 3.3.4.

If we apply Definition 2.3.4a to $\hat{A}'_{n;su;k_0}(w)$, $(w \in B_{n;k_0})$ then

$$\hat{\hat{A}}_{n;su;k_0}(\tilde{s}^f, w^1) = \{\tilde{s}^v \in \hat{A}'_{n;su;k_0}(w^1) \mid \hat{c}_{fl} + \hat{c}_{lv} > \hat{c}_{fv} \ \forall\ \tilde{s}^l \in \hat{A}'_{n;su;k_0}\left(w^1\right) \text{ with } \tilde{s}^l \neq \tilde{s}^v\}$$

are the smaller DA decision sets, where $\tilde{s}^f \in \tilde{S}_{n;su;k_0}, w^1 \in B_{n;k_0}$.

We will now prove:

[11] Corresponding algorithms are so-called "lazy algorithms", see Section 10.2.3 in [8].

$\hat{A}_{n;su;k_0}(\tilde{s^f}, w^1) = \hat{\hat{A}}_{n;su;k_0}(\tilde{s}^f, w^1)$ for any $\tilde{s}^f \in \tilde{S}_{n;su;k_0}, w^1 \in B_{n;k_0}$.

1. We show

$$\tilde{s}^v \in \hat{A'}_{n;su;k_0}\left(w^1\right) \setminus \hat{A}_{n;su;k_0}(\tilde{s}^f, w^1) \;\Rightarrow\; \tilde{s}^v \notin \hat{\hat{A}}_{n;su;k_0}(\tilde{s}^f, w^1).$$

Case: $\sum\limits_{i=1}^{n} w_i^1 < su$:

Let $\hat{x}^v \in X_{opt}(\tilde{s}^f, \tilde{s}^v)$.

Since $\tilde{s}^v \in \hat{A'}_{n;su;k_0}\left(w^1\right) \setminus \hat{A}_{n;su;k_0}(\tilde{s}^f, w^1)$

$\exists\, j_0 :\; \tilde{s}^v_{j_0} > max\,\{\tilde{s}^f_{j_0}, w^1_{j_0}\}$ (see (3.2.11) and (3.3.5)), thus

$\exists\, i_0 \neq j_0 :\; \hat{x}^v_{i_0 j_0} > 0.$

We set

$$\hat{x}^l :\quad \hat{x}^l_{ij} = \begin{cases} \hat{x}^v_{ij} - 1 & \text{if } (i,j) = (i_0, j_0), \\ \hat{x}^v_{ij} + 1 & \text{if } (i,j) = (i_0, i_0), \\ \hat{x}^v_{ij} & \text{otherwise,} \end{cases}$$

$$\tilde{s}^l :\quad \tilde{s}^l_i = \begin{cases} \tilde{s}^v_i - 1 & \text{if } i = j_0, \\ \tilde{s}^v_i + 1 & \text{if } i = i_0, \\ \tilde{s}^v_i & \text{otherwise.} \end{cases}$$

Obviously,

$\tilde{s}^l \in \hat{A'}_{n;su;k_0}\left(w^1\right),\; \tilde{s}^l \neq \tilde{s}^v$ and furthermore

$\hat{c}_{fl} \leq \hat{c}_{fv} - k_{i_0 j_0}$.

Together with $\hat{c}_{lv} = k_{i_0 j_0}$

$\hat{c}_{fl} + \hat{c}_{lv} \leq \hat{c}_{fv}$

follows.

According to Theorem 3.3.4 that means

$\hat{c}_{fl} + \hat{c}_{lv} = \hat{c}_{fv}$, hence

$\tilde{s}^v \notin \hat{\hat{A}}_{n;su;k_0}(\tilde{s}^f, w^1)$.

The case $\sum_{i=1}^{n} w_i^1 > su$ is proved analogously.

2. We show by contradiction:

$$\tilde{s}^v \in \hat{A}_{n;su;k_0}(\tilde{s}^f, w^1) \;\Rightarrow\; \tilde{s}^v \in \hat{\hat{A}}_{n;su;k_0}(\tilde{s}^f, w^1).$$

Assumption:

$\tilde{s}^v \in \hat{A}_{n;su;k_0}(\tilde{s^f}, w^1)$ and $\tilde{s}^l \in \hat{A}'_{n;su;k_0}(w^1)$ with $\tilde{s}^l \neq \tilde{s}^v$ exist,

so that $\hat{c}_{fl} + \hat{c}_{lv} = \hat{c}_{fv}$.

(That means $\tilde{s}^v \notin \hat{\hat{A}}_{n;su;k_0}(\tilde{s}^f, w^1)$.)

Let $\hat{x}^1 \in X_{opt}(\tilde{s}^f, \tilde{s}^l)$ and $\hat{x}^2 \in X_{opt}(\tilde{s}^l, \tilde{s}^v)$ (see (3.3.4)).

Case: $\sum_{i=1}^{n} w_i^1 < su$:

$$\begin{aligned} \tilde{s}^l \neq \tilde{s}^v \;&\Rightarrow\; \exists\, j_0 :\; \tilde{s}^v_{j_0} > \tilde{s}^l_{j_0}, \\ &\Rightarrow\; \exists\, y_0 \neq j_0 :\; \hat{x}^2_{y_0 j_0} > 0. \end{aligned} \tag{*1}$$

Since $\hat{x}^2$ is optimal

$$\Rightarrow\; \tilde{s}^v_{y_0} = \hat{x}^2_{y_0 y_0} \leq \tilde{s}^l_{y_0} - \hat{x}^2_{y_0 j_0}$$

$$\text{thus } \tilde{s}^v_{y_0} < \tilde{s}^l_{y_0}.$$

In the considered case

$$w^1_{y_0}\ (\leq \tilde{s}^v_{y_0}) \;<\; \tilde{s}^l_{y_0} \tag{*2}$$

$$\text{and } 0 < \tilde{s}^l_{y_0} \tag{*2a}$$

follow.

Since $\hat{x}^1$ is optimal (see also Lemma 3.1.1(ii))

$$\Rightarrow\; 0 < \hat{x}^1_{y_0 y_0} \;\text{ if }\; \tilde{s}^f_l > 0; \tag{*3}$$

otherwise, $y_2 (\neq y_0)$ exists such that $0 < \hat{x}^1_{y_2 y_0}$.

Since $\tilde{s}^v_{j_0} > \tilde{s}^l_{j_0}$, $\tilde{s}^l \in \hat{A}'_{n;su;k_0}(w^1)$ and $\tilde{s}^v \in \hat{A}_{n;su;k_0}(\tilde{s^f}, w^1)$

(that means $\tilde{s}^v_{j_0} \leq max\ \{w^1_{j_0}, \tilde{s^f}_{j_0}\}$, see (3.2.11))

$$w^1_{j_0} \leq \tilde{s}^l_{j_0} < \tilde{s}^v_{j_0} \leq \tilde{s}^f_{j_0}$$

$$\text{and thus } \tilde{s}^l_{j_0} < \tilde{s}^f_{j_0} \tag{*3a}$$

follow. Furthermore,

$$\Rightarrow \quad \exists\, y_1 \neq j_0 : \ \hat{x}^1_{j_0 y_1} > 0, \tag{*4}$$

$$\Rightarrow \quad \hat{x}^1_{j_0 j_0} < \tilde{s}^f_{j_0}. \tag{*5}$$

Now, we go on to distinguish between the three cases: a) $y_1 \neq y_0$ and $0 < \hat{x}^1_{y_0 y_0}$, b) $y_1 = y_0$ and c) $y_1 \neq y_0$ and $0 < \hat{x}^1_{y_2 y_0}$ where $y_2 \neq y_0$.

Case: a) $y_1 \neq y_0$ and $0 < \hat{x}^1_{y_0 y_0}$:

See Figure 3.3.1.

We set (see also (*3a) and (*2a)):

$$\tilde{s}'^l : \quad \tilde{s}'^l_y = \begin{cases} \tilde{s}^l_y + 1 & \text{if } y = j_0, \\ \tilde{s}^l_y - 1 & \text{if } y = y_0, \\ \tilde{s}^l_y & \text{otherwise,} \end{cases}$$

$$\hat{x}'^1 : \quad \hat{x}'^1_{iy} = \begin{cases} \hat{x}^1_{iy} + 1 & \text{if } (i,y) = (j_0, j_0) \text{ or } (i,y) = (y_0, y_1), \\ \hat{x}^1_{iy} - 1 & \text{if } (i,y) = (j_0, y_1) \text{ or } (i,y) = (y_0, y_0), \\ \hat{x}^1_{iy} & \text{otherwise} \end{cases} \tag{*6}$$

(see also (*5), (*3) and (*4)).

Obviously, $\hat{x}'^1 \in X_f(\tilde{s}^f, \tilde{s}'^l)$.

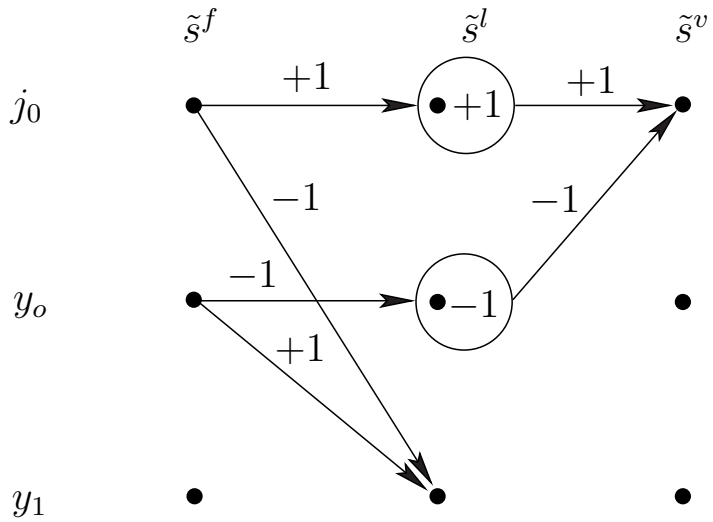



**Figure 3.3.1.**

$$\hat{x}'^2: \quad \hat{x}'^2_{yj} = \begin{cases} \hat{x}^2_{yj} + 1 & \text{if } (y,j) = (j_0, j_0), \\ \hat{x}^2_{yj} - 1 & \text{if } (y,j) = (y_0, j_0), \\ \hat{x}^2_{iy} & \text{otherwise} \end{cases}$$

(see also (*6) and (*1)).

Obviously, $\hat{x}'^2 \in X_f(\tilde{s}'^l, \tilde{s}^v)$.

Using Theorem 3.3.4 and the distance property $k_{y_0 y_1} < k_{j_0 y_1} + k_{y_0 j_0}$, see (3.1.2),

$$\hat{c}_{fv} \le \hat{c}(\tilde{s}^f, \tilde{s}'^l) + \hat{c}_(\tilde{s}'^l, \tilde{s}^v) \le \sum_{i,y} k_{iy}\ \hat{x}'^1_{iy} + \sum_{y,j} k_{yj}\ \hat{x}'^2_{yj} =$$

$$\sum_{i,y} k_{iy}\ \hat{x}^1_{iy} + \sum_{y,j} k_{yj}\ \hat{x}^2_{yj} + k_{y_0 y_1} - k_{j_0 y_1} - k_{y_0 j_0} < \hat{c}_{fl} + \hat{c}_{lv}$$

follows.

$\hat{c}_{fv} < \hat{c}_{fl} + \hat{c}_{lv}$ means that the assumption was false in case a).

Case: b) $y_1 = y_0$:

We set $s'^l$ and $\hat{x}'^2$ as in case a) but:

$$\hat{x}'^1: \quad \hat{x}'^1_{iy} = \begin{cases} \hat{x}^1_{iy} + 1 & \text{if } (i,y) = (j_0, j_0), \\ \hat{x}^1_{iy} - 1 & \text{if } (i,y) = (j_0, y_0), \\ \hat{x}^1_{iy} & \text{otherwise} \end{cases}$$

(see also (*5) and (*4) with $y_1 = y_0$).

Obviously, $\hat{x}'^1 \in X_f(\tilde{s}^f, \tilde{s}'^l)$ and $\hat{x}'^2 \in X_f(\tilde{s}'^l, \tilde{s}^v)$.

Using Theorem 3.3.4 and $0 < k_{j_0 y_0} + k_{y_0 j_0}$, see (3.1.2),

$$\hat{c}_{fv} \le \hat{c}(\tilde{s}^f, \tilde{s}'^l) + \hat{c}_(\tilde{s}'^l, \tilde{s}^v) \le \sum_{i,y} k_{iy}\ \hat{x}'^1_{iy} + \sum_{y,j} k_{yj}\ \hat{x}'^2_{yj} =$$

$$\sum_{i,y} k_{iy}\ \hat{x}^1_{iy} + \sum_{y,j} k_{yj}\ \hat{x}^2_{yj} - k_{j_0 y_0} - k_{y_0 j_0} < \hat{c}_{fl} + \hat{c}_{lv}$$

follows.

$\hat{c}_{fv} < \hat{c}_{fl} + \hat{c}_{lv}$ means that the assumption was false in case b).

Case: c) $y_1 \neq y_0$ and $0 < \hat{x}^1_{y_2 y_0}$ where $y_2 \neq y_0$:

We set $s'^l$ and $\hat{x}'^2$ as in case a) and

$$\hat{x}'^1 : \quad \hat{x}'^1_{iy} = \begin{cases} \hat{x}^1_{iy} + 1 & \text{if } (i,y) = (j_0, j_0) \text{ or } (i,y) = (y_2, y_1), \\ \hat{x}^1_{iy} - 1 & \text{if } (i,y) = (j_0, y_1) \text{ or } (i,y) = (y_2, y_0), \\ \hat{x}^1_{iy} & \text{otherwise.} \end{cases}$$

Similarly to case a)

$$\hat{c}_{fv} \leq \hat{c}(\tilde{s}^f, \tilde{s}'^l) + \hat{c}_(\tilde{s}'^l, \tilde{s}^v) \leq \sum_{i,y} k_{iy}\ \hat{x}'^1_{iy} + \sum_{y,j} k_{yj}\ \hat{x}'^2_{yj} =$$

$$\sum_{i,y} k_{iy}\ \hat{x}^1_{iy} + \sum_{y,j} k_{yj}\ \hat{x}^2_{yj} + k_{y_2 y_1} - k_{y_2 y_0} - k_{j_0 y_1} - k_{y_0 j_0} < \hat{c}_{fl} + \hat{c}_{lv}$$

follows.

The case $\sum\limits_{i=1}^{n} w^1_i > su$ is proved analogously.

(The case $\sum\limits_{i=1}^{n} w^1_i = su$ is not relevant
since $\hat{A}'_{n;su;k_0}\left(w^1\right) = \hat{A}_{n;su;k_0}(\tilde{s}^f, w^1) = w^1\ (\in \tilde{S}_{n;su;k_0})$).

(b) follows directly from Theorem 2.3.13a. ■

Finally, conditions that two states are elements of one decision set are given in the following theorem.

**Theorem 3.3.9.** .

*(i) For given states $\tilde{s}^1 \in \tilde{S}_{n;su;k_0}$ and $\tilde{s}^2 \in \tilde{S}_{n;su;k_0}$, a state $\tilde{s} \in \tilde{S}_{n;su;k_0}$ and requirements $w \in B_{n;k_0}$ exist so that $\{\tilde{s}^1, \tilde{s}^2\} \subseteq \hat{A}_{n;su;k_0}(\tilde{s}, w)$ if and only if at least one of the following inequalities is valid:*

$$(a) \quad \sum_{i:\tilde{s}^1_i = \tilde{s}^2_i} \tilde{s}^1_i \geq \frac{1}{2} \sum_{i:\tilde{s}^1_i \neq \tilde{s}^2_i} \mid \tilde{s}^1_i - \tilde{s}^2_i \mid = \frac{1}{2} \sum_i \mid \tilde{s}^1_i - \tilde{s}^2_i \mid,$$

$$(b) \quad \sum_{i:\tilde{s}^1_i = \tilde{s}^2_i} (k_{0_i} - \tilde{s}^1_i) \geq \frac{1}{2} \sum_{i:\tilde{s}^1_i \neq \tilde{s}^2_i} \mid \tilde{s}^1_i - \tilde{s}^2_i \mid = \frac{1}{2} \sum_i \mid \tilde{s}^1_i - \tilde{s}^2_i \mid .$$

*(ii) For given $\tilde{s}^1$ and $\tilde{s}^2$, an $\tilde{s} \in \tilde{S}_{n;su;k_0}$ and a $w \in B_{n;k_0}$ exist so that $\{\tilde{s}^1, \tilde{s}^2\} = \hat{A}_{n;su;k_0}(\tilde{s}, w)$ if and only if $U(\tilde{s}^1, \tilde{s}^2) = 1$ and $\exists j : \tilde{s}^1_j = \tilde{s}^2_j$.*

***Proof.***

(i) :

1. ($\Rightarrow$): Let $\{\tilde{s}^1, \tilde{s}^2\} \subseteq \hat{A}_{n;su;k_0}(\tilde{s}, w)$.

   In case $C[w, su]$, the conditions of (3.2.11) and $\sum_i \tilde{s}^1_i = \sum_i \tilde{s}^2_i = \sum_i \tilde{s}_i$ yield

$$\sum_{i:w_i \geq \tilde{s}_i} (\tilde{s}^l_i - \tilde{s}_i) = \sum_{i:w_i \leq \tilde{s}_i} (\tilde{s}_i - \tilde{s}^l_i) \text{ for } l = 1; 2.$$

   According to (3.2.11)

$$\sum_{i:\tilde{s}^1_i = \tilde{s}^2_i} \tilde{s}^l_i \geq \sum_{i:w_i \geq \tilde{s}_i} \tilde{s}^l_i \geq \sum_{i:w_i \geq \tilde{s}_i} (\tilde{s}^l_i - \tilde{s}_i) = \sum_{i:w_i < \tilde{s}_i} (\tilde{s}_i - \tilde{s}^l_i) = \sum_{i:w_i < \tilde{s}_i} \mid \tilde{s}_i - \tilde{s}^l_i \mid$$

   for $l = 1, 2$ follows.

   Finally, using the last inequality for $l = 1$ and $l = 2$ and $\mid \tilde{s}_i - \tilde{s}^1_i \mid + \mid \tilde{s}_i - \tilde{s}^2_i \mid \geq \mid \tilde{s}^1_i - \tilde{s}^2_i \mid$,

$$\sum_{i:\tilde{s}^1_i = \tilde{s}^2_i} \tilde{s}^1_i = \frac{1}{2}\Big(\sum_{i:\tilde{s}^1_i = \tilde{s}^2_i} (\tilde{s}^1_i + \tilde{s}^2_i)\Big) \geq \frac{1}{2} \sum_{i:w_i < \tilde{s}_i} \mid \tilde{s}^1_i - \tilde{s}^2_i \mid = \frac{1}{2} \sum_i \mid \tilde{s}^1_i - \tilde{s}^2_i \mid$$

   can easily be seen.

   In the case $C[su, w]$ we come to a similar conclusion. The inequality $(b)$ results in:

$$\begin{aligned} \sum_{i:\tilde{s}^1_i = \tilde{s}^2_i} (k_{0_i} - \tilde{s}^l_i) \geq \sum_{i:\tilde{s}_i \geq w_i} (k_{0_i} - \tilde{s}^l_i) &\geq \sum_{i:\tilde{s}_i \geq w_i} (\tilde{s}_i - \tilde{s}^l_i) \\ &= \sum_{i:\tilde{s}_i < w_i} (\tilde{s}^l_i - \tilde{s}_i) \geq \frac{1}{2} \sum_i \mid \tilde{s}^1_i - \tilde{s}^2_i \mid . \end{aligned}$$

2. ($\Leftarrow$): Let $(a)$ be valid in relation to $\tilde{s}^1 \in \tilde{S}_{n;su;k_0}$ and $\tilde{s}^2 \in \tilde{S}_{n;su;k_0}$.

   We set

$$w : \quad w_i \begin{cases} = \tilde{s}^1_i \text{ if } \tilde{s}^1_i = \tilde{s}^2_i, \\ \leq \min\{\tilde{s}^1_i, \tilde{s}^2_i\} \text{ otherwise} \end{cases}$$

and

$$\tilde{s}: \quad \begin{cases} \tilde{s}_i \leq \tilde{s}_i^1 \ \text{ if } \tilde{s}_i^1 = \tilde{s}_i^2, \\ \tilde{s}_i \geq \max\{\tilde{s}_i^1, \tilde{s}_i^2\} \ \text{ if } \tilde{s}_i^1 \neq \tilde{s}_i^2, \\ \sum\limits_i \tilde{s}_i = su. \end{cases}$$

Thereby, $w$ satisfies the condition $\sum_i w_i \leq su$. In addition, an $\tilde{s}$ which fulfils the above conditions exists, since on the one hand

$$\sum_{i:\tilde{s}_i^1 \neq \tilde{s}_i^2} \max\{\tilde{s}_i^1, \tilde{s}_i^2\} = \sum_{i:\tilde{s}_i^1 \neq \tilde{s}_i^2} (\tilde{s}_i^1 + \frac{1}{2} \mid \tilde{s}_i^1 - \tilde{s}_i^2 \mid) \leq \sum_i \tilde{s}_i^1 = su$$

because $\sum\limits_i \tilde{s}_i^1 = \sum\limits_i \tilde{s}_i^2$ and $(a)$ and on the other hand

$$\sum_{i:\tilde{s}_i^1 \neq \tilde{s}_i^2} \max\{\tilde{s}_i^1, \tilde{s}_i^2\} + \sum_{i:\tilde{s}_i^1 = \tilde{s}_i^2} \tilde{s}_i^1 \geq su.$$

According to the construction of $w$ and $s$:

$$\begin{aligned} &\tilde{s}_i^1 = \tilde{s}_i^2 = w_i \geq \tilde{s}_i && \text{for } i \text{ with } \tilde{s}_i^1 = \tilde{s}_i^2, \\ &w_i \leq \min\{\tilde{s}_i^1, \tilde{s}_i^2\} < \max\{\tilde{s}_i^1, \tilde{s}_i^2\} \leq \tilde{s}_i && \text{for } i \text{ with } \tilde{s}_i^1 \neq \tilde{s}_i^2 \end{aligned}$$

are valid. Hence,

$$\tilde{s}^1 \text{ and } \tilde{s}^2 \text{ are elements of } \hat{A}_{n;su;k_0}(\tilde{s}, w) \text{ (see (3.2.11))}.$$

Now, let $(b)$ be valid in relation to $\tilde{s}^1 \in \tilde{S}_{n;su;k_0}$ and $\tilde{s}^2 \in \tilde{S}_{n;su;k_0}$.

We set

$$w: \quad w_i \begin{cases} = \tilde{s}_i^1 \ \text{ if } \tilde{s}_i^1 = \tilde{s}_i^2, \\ \geq \max\{\tilde{s}_i^1, \tilde{s}_i^2\} \text{ otherwise} \end{cases}$$

and

$$\tilde{s}: \quad \begin{cases} k_{0_i} \geq \tilde{s}_i \geq \tilde{s}_i^1 \ \text{ if } \tilde{s}_i^1 = \tilde{s}_i^2, \\ \tilde{s}_i \leq \min\{\tilde{s}_i^1, \tilde{s}_i^2\} \ \text{ if } \tilde{s}_i^1 \neq \tilde{s}_i^2, \\ \sum\limits_i \tilde{s}_i = su. \end{cases}$$

Thus, $w$ satisfies the condition $\sum_i w_i > su$. The continuation of this proof for (b) is analogous to the proof of case (a).

(ii) :

1. ($\Leftarrow$): Let $U(\tilde{s}^1, \tilde{s}^2) = 1$ and $\exists j : \tilde{s}^1_j = \tilde{s}^2_j$.

   According to (3.3.2) and Lemma 3.3.2(a) indices $i_0$ and $i_1$ exist so that

$$\tilde{s}^1 = \tilde{s}^2[i_0;\ i_1] = \begin{cases} \tilde{s}^2_i + 1 & \text{for} \qquad i = i_0, \\ \tilde{s}^2_i - 1 & \text{for} \qquad i = i_1, \\ \tilde{s}^2_i & \text{otherwise.} \end{cases}$$

$$\tilde{s} \text{ with } \tilde{s}_i = \begin{cases} \min\{\tilde{s}^1_i, \tilde{s}^2_i\} \ \text{ if } i = i_0, i = i_1, \\ \tilde{s}^1_i + 1 = \tilde{s}^2_i + 1 \ \text{ if } i = j, \\ \tilde{s}^1_i = \tilde{s}^2_i \ \text{ if } i \notin \{i_0, i_1, j\} \end{cases}$$

$$\text{and } w \text{ with } w_i = \begin{cases} \max\{\tilde{s}^1_i, \tilde{s}^2_i\} \ \text{ if } i = i_0, i = i_1, \\ \tilde{s}^1_i = \tilde{s}^2_i \ \text{ if } i_0 \neq i \neq i_1 \end{cases}$$

   satisfy the condition $\{\tilde{s}^1, \tilde{s}^2\} = \hat{A}_{n;su;k_0}(\tilde{s}, w)$ if $\tilde{s}^1_j = \tilde{s}^2_j < k_0$.

$$\tilde{s} \text{ with } \tilde{s}_i = \begin{cases} \max\{\tilde{s}^1_i, \tilde{s}^2_i\} \ \text{ if } i = i_0, i = i_1, \\ \tilde{s}^1_i - 1 = \tilde{s}^2_i - 1 \ \text{ if } i = j, \\ \tilde{s}^1_i = \tilde{s}^2_i \ \text{ if } i \notin \{i_0, i_1, j\} \end{cases}$$

$$\text{and } w \text{ with } w_i = \begin{cases} \min\{\tilde{s}^1_i, \tilde{s}^2_i\} \ \text{ if } i = i_0, i = i_1, \\ \tilde{s}^1_i = \tilde{s}^2_i \ \text{ if } i_0 \neq i \neq i_1 \end{cases}$$

   can be used for the proof if $k_0 = \tilde{s}^1_j = \tilde{s}^2_j (> 0)$.

2. ($\Rightarrow$): The other direction of the proof for (ii) can easily be seen by using the definition of $\hat{A}_{n;su;k_0}(\tilde{s}, w)$ as proof by contradiction.

■

(For example, $\tilde{s}^1 = (0,1,0,1)$ and $\tilde{s}^2 = (0,1,1,0)$, where $n = 4$, $k_0 = (1,1,1,1)$ satisfy $(a)$ as well as $(b)$.)

**Regarding the Transition Probabilities**

As already mentioned, if the SDDP problems are modelled as MDPs the corresponding matrices of transition probabilities are called (general) partitions-requirements-matrices (PRMs) and the condition (2.3.3), which requires that $p^d_{fl} > 0$ for all $f, l$ and $d$, is fulfilled (see (3.2.13)).

We now show that the elements of the main diagonal $p^d_{ff}$ of PRMs are independent of $d$ (see also Lemma 2.3.14) and that they can be computed in a straight forward way:

**Lemma 3.3.10.** *Let $\tilde{s}^l \in \tilde{S}_{n;su;k_0}, \tilde{s}^f \in \tilde{S}_{n;su;k_0}$ with $\tilde{s}^l \neq \tilde{s}^f$ be given.*

*Then,*

*a)* $| \{w \mid \hat{A}_{n;su;k_o}(\tilde{s}^l, w) = \tilde{s}^l\} | = | \{w \mid \hat{A}_{n;su;k_o}(\tilde{s}^l, w) \supseteq \{\tilde{s}^l\}\} |$
$= \prod_i (\tilde{s}^l_i + 1) + \prod_i (k_0 - \tilde{s}^l_i + 1) - 1,$

*b)* $p_{ll} = \sum_{w:0 \leq w \leq \tilde{s}^l} q(w) + \sum_{w:\tilde{s}^l \leq w \leq k_0} q(w) - q(\tilde{s}^l)$

*(independent of d).*

***Proof.***

a) $\tilde{s}^l \subseteq \hat{A}_{n;su;k_0}(\tilde{s}^l, w)$ is valid if and only if the inequality $0 \leq w_i \leq \tilde{s}^l_i$ for all $i$ or the inequality $k_{0_i} \geq w_i \geq \tilde{s}^l_i$ for all $i$ are fulfilled (see (3.2.11)). Then, $\tilde{s}^l = \hat{A}_{n;su;k_0}(\tilde{s}^l, w)$ and the equalities from a) follow.

b) results from (2.3.7) and the proof of a). ■

## 3.4 Characterization of Special Cases of SDDP Problems

Throughout this Section we investigate SDDP problems where "identical basic costs" (in other words, unit distances) are supposed. Without loss of

generality (see Lemma 3.3.1) we can set $k_{ij} = 1$ for $i \neq j$ ($k_{ii} = 0$ have to be valid according to (3.1.2)).

Optimal decisions of such problems can be used as approximate solutions of corresponding SDDP problems, in which the basic costs differ only slightly from each other. Such optimal decisions can also be used as starting decisions if corresponding SDDP problems are solved by iterative methods, such as the Howard algorithm, for example.

However, we will also see that is not simple to solve even such special cases of SDDP problems or to prove that optimally conjectured decisions of these problems are infact optimal.

Moreover, the corresponding PRMs (in the strict meaning) themselves lead to interesting combinatorial problems (see Chapter 4). Some conjectures of special SDDP problems can then only be proven in Chapter 4 (see Section 4.7).

Throughout this section we suppose, as discussed above,

$$\left.\begin{array}{l} k_{ij} = 1 \ \ \forall\, i \neq j, \\ k_{ii} = 0 \ \ \forall\, i \end{array}\right\}. \tag{3.4.1}$$

Then, the internal costs correspond to the conversion numbers and

$$\left.\begin{array}{l} k_{ij} = 1 \ \ \forall\, i \neq j, \\ k_{ii} = 0 \ \ \forall\, i \end{array}\right\} \Rightarrow \hat{c}_{fl} = \left\{\begin{array}{ll} \sum\limits_i \max\{0, w_i - \tilde{s}_i^f\} & \text{in case } C[w, su], \\ \sum\limits_i \max\{0, \tilde{s}_i^f - w_i\} & \text{in case } C[su, w] \end{array}\right. \tag{3.4.2}$$

$$\text{for all } \tilde{s}^l \in \hat{A}_{n;su;k_0}(\tilde{s}^f, w)$$

follows (see Lemma 3.3.3).

Thus, the internal costs fulfill not only (2.3.12) (as discussed in Sections 3.2) but also (2.3.13) (see Section 2.3.2) in the case of identical basic costs.

Hence, the corresponding average one-step reward functions are independent of the decisions (see (2.3.14)).

(Comparisons of the average reward functions are possible simply by using the formulas from the following theorem 3.4.1.)

According to Theorem 2.3.8, an almost-partial order of the states exists which means that the complexity of computing optimal decisions can be reduced (refer to Corollary 2.3.9).

Furthermore, the average one-step reward functions are equivalent to the expected conversion numbers of a stage.

We can now make the following conjecture:

Decisions for feasible states $s$ with minimum $\gamma(s)$ are optimal, if (3.4.1) is supposed (Con3.1)

(without further conditions for the probability functions.)

(Such solutions would not be obvious for MDPs. We can think of the demanding conditions of the matrices of transition probabilities in the case of dominance (see Section 2.3.3.2).)

Weaker conjectures follow with additional conditions for the probability functions.

Let $w_i,\ i = 1, 2, \cdots, n$ be independent and let

$q_i : \mathbb{Z}_+ \ \rightarrow\ [0, 1],\ i = 1, \cdots, n$

with

$$q_i(z) \neq 0 \ \Leftrightarrow\ z \in \{0, 1, \cdots, k_{0_i}\} \ \text{ and } \sum_{z=0}^{k_{0_i}} q_i(z) = 1$$

be the corresponding probability functions.

Then, $q(z) = \prod_{i=1}^{n} q_i(z_i)$ for $z \in \mathbb{Z}_+^n$ follows (see (3.2.7)).

In particular, we designate the two cases:

$$q_i(w_i) = \tfrac{1}{k_{0_i}+1},\ i = 1, 2, \cdots n \tag{3.4.3}$$

(this means $w_i$ are discrete uniformly distributed)

and

$w_i,\ i = 1, 2, \cdots n$ are independent and identically distributed, where $k_{0_1} = k_{0_2} = \cdots = k_{0_n}$ is assumed. (3.4.4)

Then, the corresponding conjectures are:

Decisions for feasible states $s$ with minimum $\gamma(s)$ are optimal if (3.4.1) and (3.4.3) are assumed (Con3.2)

and

decisions for feasible states $s$ with minimum $\gamma(s)$ are optimal if (3.4.1) and (3.4.4) are supposed. (Con3.3)

Furthermore, we will see that conjecture (Con3.3) is equivalent to:

Decisions for feasible states with least square sums of their components are optimal if (3.4.1) and (3.4.4) are supposed. (Con3.3a)

(See Lemma 4.2.2 and Definition 4.1.1(a), (c) in Chapter 4.)

At this point, we have dealt in detail with conjecture (Con3.3) and have proven this conjecture for many cases.

SDDP problems involve combinatorial aspects (in particular) if (3.4.1) and (3.4.4) are assumed. Therefore, several considerations for special SDDP problems can only be found in Chapter 4.

Whether the conjectures (Con3.1) and (Con3.2) are true seems more questionable. (They are true for specific examples.) Otherwise, these conjectures have not been further investigated.

At the very least, the almost-partial order of the states mentioned above exists, which means that the complexity of computing of optimal decisions can be reduced if (3.4.1) is supposed (see the following Section 3.5 also).

In Section 3.4.1. we give formulas for computation of the average one-step reward functions under condition (3.4.1).

In addition, in Section 3.4.2 we will see that special SDDP problems can be reduced. This will also prepare us for Chapter 4.

### 3.4.1 Average Reward Functions of SDDP Problems with Identical Basic Costs

**Theorem 3.4.1.** *Let a SDDP problem (modelled as MDP) with identical basic costs ((3.4.1) is therefore fulfilled) be given, (where* $\tilde{S}_{n;su;k_0} = \{\tilde{s}^1, \tilde{s}^2, \cdots, \tilde{s}^r\}$*).*

*Then, the average one-step reward functions* $\gamma = (\gamma_1, \cdots, \gamma_r)^T$ *are independent of (feasible) decisions and the following formulas are valid:*

*(a)* $\gamma(\tilde{s}^f) = \sum\limits_{w \in B_{n;k_0}} \left( \sum\limits_{i=1}^{n} \max\{0, \tilde{s}_i^f - w_i\} \right) q(w) + R(n, su, k_0, q),$

*(b) if, in addition,* $w_i$, $i = 1, 2, \cdots n$ *are independent (this also means* $q(w) = \prod\limits_{i=1}^{n} q_i(w_i)$*)*

$$\gamma(\tilde{s}^f) = \sum_{i=1}^{n} \sum_{w_i=0}^{\tilde{s}_i^f} (\tilde{s}_i^f - w_i)\, q_i(w_i) + R(n, su, k_0, q),$$

*(c) if, in addition, (3.4.4) is supposed and* $q_0 :\equiv q_1 \equiv q_2 \equiv \cdots \equiv q_n$

$$\gamma(\tilde{s}^f) = \sum_{i=1}^{n} \sum_{w_i=0}^{\tilde{s}_i^f} (\tilde{s}_i^f - w_i)\, q_0(w_i) + R(n, su, k_0, q),$$

*(d) if, in addition, (3.4.3) is supposed*

$$\gamma(\tilde{s}^f) = \frac{1}{2} \sum_{i=1}^{n} \frac{\tilde{s}_i^f(\tilde{s}_i^f + 1)}{k_{0_i} + 1} + R(n, su, k_0, q) \text{ and}$$

*(e) if (3.4.3) and (3.4.4) are additionally supposed* $k_0 := k_{0_1} = k_{0_2} = \cdots = k_{0_n}$

$$\gamma(\tilde{s}^f) = \frac{1}{2} \frac{1}{k_0 + 1} \sum_{i=1}^{n} (\tilde{s}_i^f)^2 + R_1(n, su, k_0)$$

*where* $R(n, su, k_0, q) = \sum\limits_{w \in B_n;k_0 : C[w, su]} q(w) (\sum\limits_{i=1}^{n} w_i - su)$ *and*

$$R_1(n; su; k_0) = \frac{1}{(k_0+1)^n} \sum_{w \in B_n;k_0 : C[w, su]} (\sum_{i=1}^{n} w_i - su) + \frac{su}{2(k_0+1)}$$

*are independent of* $s^f$.

**Remarks 3.4.1.** *In Section 4.2.2 we will see that*

$$\gamma(\tilde{s}^f) > \gamma(\tilde{s}^l) \;\Leftrightarrow\; \sum_{i=1}^{n}(\tilde{s}_i^f)^2 > \sum_{i=1}^{n}(\tilde{s}_i^l)^2$$

*if* $\tilde{s}^l$ *is a successor of* $\tilde{s}^f$ *(see Definition 4.1.1(a), (c)) and (3.4.1) and (3.4.4) are assumed (see Lemma 4.2.2(a)).*

***Proof*** of Theorem 3.4.1: According to (2.3.8) (first equation), the average one-step reward functions of DA MDPs are computed as follows

$$\gamma(\tilde{s}^f, d) = \sum_{\tilde{s}^l} \sum_{w:\tilde{s}^l = \hat{d}(\tilde{s}^f, w)} \hat{c}_{fl}\; q(w).$$

Together with (3.4.2)

$$\begin{aligned}\gamma(\tilde{s}^f) := \gamma(\tilde{s}^f, d) = &\sum_{w \in B_{n;k_0}: C[w,su]} \left(\sum_{i=1}^{n} \max\{0, w_i - \tilde{s}_i^f\}\right) q(w) \\ &+ \sum_{w \in B_{n;k_0}: \sum_i w_i > su} \left(\sum_{i=1}^{n} \max\{0, \tilde{s}_i^f - w_i\}\right) q(w)\end{aligned}$$

for any $d$ with $d(\tilde{s}^f, w) \in \hat{A}_{n;su;k_0}(\tilde{s}^f, w)$

and

$$\begin{aligned}\gamma(\tilde{s}^f) = &\sum_{w \in B_{n;k_0}} \left(\sum_{i=1}^{n} \max\{0, \tilde{s}_i^f - w_i\}\right) q(w) \\ &+ \sum_{w \in B_{n;k_0}: C[w,su]} \left(\sum_{i=1}^{n} \Big(\max\{0, w_i - \tilde{s}_i^f\} - \max\{0, \tilde{s}_i^f - w_i\}\Big)\right) q(w)\end{aligned} \tag{*1}$$

follow.

Since

$$\sum_{i=1}^{n} \Big(\max\{0, w_i - \tilde{s}_i^f\} - \max\{0, \tilde{s}_i^f - w_i\}\Big) = \sum_{i=1}^{n} \Big(w_i - \tilde{s}_i^f\Big) = \sum_{i=1}^{n} w_i - su,$$

the relation

$$\gamma(\tilde{s}^f) = \sum_{w \in B_{n;k_0}} \left( \sum_{i=1}^{n} \max\{0, \tilde{s}_i^f - w_i\} \right) q(w) + R(n, su, k_0, q) \qquad (*2)$$

where

$$R(n, su, k_0, q) = \sum_{w \in B_{n;k_0} : C[w,su]} q(w) (\sum_{i=1}^{n} w_i - su) \qquad (*3)$$

is valid and (a) is thus proved.

If, in addition, $w_i,\ i = 1, 2, \cdots n$ are independent, then using the definition of $B_{n;k_0}$ the equation (*2) can be transformed into:

$$\gamma(\tilde{s}^f) = \sum_{w_1=0}^{k_{0_1}} \sum_{w_2=0}^{k_{0_2}} \cdots \sum_{w_n=0}^{k_{0_n}} \left( \sum_{i=1}^{n} \max\{0, \tilde{s}_i^f - w_i\} \right) \prod_{i=1}^{n} q_i(w_i) + R(n, su, k_0, q).$$

Step by step, we now move the sum "$\sum_{i=1}^{n}$" to the left (outside):

$$\gamma(\tilde{s}^f) = \sum_{w_1=0}^{k_{0_1}} \sum_{w_2=0}^{k_{0_2}} \cdots \sum_{w_{n-1}=0}^{k_{0_{n-1}}} \left[ \left( \sum_{i=1}^{n-1} \max\{0, \tilde{s}_i^f - w_i\} \right) \prod_{i=1}^{n-1} q_i(w_i) \left( \sum_{w_n=0}^{k_{0_n}} q_n(w_n) \right) + \prod_{i=1}^{n-1} q_i(w_i) \sum_{w_n=0}^{k_{0_n}} \Big( \max\{0, \tilde{s}_n^f - w_n\}\ q_n(w_n) \Big) \right] + R(n, su, k_0, q).$$

Since $\sum_{w_n=0}^{k_{0_n}} q_n(w_n) = 1$, the last equation can be slightly simplified.

We then continue the transposition of the sums:

$$\gamma(\tilde{s}^f) = \sum_{w_1=0}^{k_{0_1}} \sum_{w_2=0}^{k_{0_2}} \cdots \sum_{w_{n-2}=0}^{k_{0_{n-2}}} \Bigg[ \left( \sum_{i=1}^{n-2} \max\{0, \tilde{s}_i^f - w_i\} \right) \prod_{i=1}^{n-2} q_i(w_i) \left( \sum_{w_{n-1}=0}^{k_{0_{n-1}}} q_{n-1}(w_{n-1}) \right) + \prod_{i=1}^{n-2} q_i(w_i) \sum_{w_{n-1}=0}^{k_{0_{n-1}}} \Big( \max\{0, \tilde{s}_{n-1}^f - w_{n-1}\} \, q_{n-1}(w_{n-1}) \Big) + \prod_{i=1}^{n-2} q_i(w_i) \left( \sum_{w_{n-1}=0}^{k_{0_{n-1}}} q_{n-1}(w_{n-1}) \right) \sum_{w_n=0}^{k_{0_n}} \Big( \max\{0, \tilde{s}_n^f - w_n\} \, q_n(w_n) \Big) \Bigg] + R(n, su, k_0$$

$\sum_{w_{n-1}=0}^{k_{0_{n-1}}} q_{n-1}(w_{n-1}) = 1$ can be repeatedly used to simplify the equation:

$$\gamma(\tilde{s}^f) = \sum_{w_1=0}^{k_{0_1}} \sum_{w_2=0}^{k_{0_2}} \cdots \sum_{w_{n-2}=0}^{k_{0_{n-2}}} \Bigg[ \left( \sum_{i=1}^{n-2} \max\{0, \tilde{s}_i^f - w_i\} \right) \prod_{i=1}^{n-2} q_i(w_i) + \prod_{i=1}^{n-2} q_i(w_i) \sum_{w_{n-1}=0}^{k_{0_{n-1}}} \Big( \max\{0, \tilde{s}_{n-1}^f - w_{n-1}\} \, q_{n-1}(w_{n-1}) \Big) + \prod_{i=1}^{n-2} q_i(w_i) \sum_{w_n=0}^{k_{0_n}} \Big( \max\{0, \tilde{s}_n^f - w_n\} \, q_n(w_n) \Big) \Bigg] + R(n, su, k_0, q)$$

The summation of the middle two terms yield:

$$\gamma(\tilde{s}^f) = \sum_{w_1=0}^{k_{0_1}} \sum_{w_2=0}^{k_{0_2}} \cdots \sum_{w_{n-2}=0}^{k_{0_{n-2}}} \Bigg[ \left( \sum_{i=1}^{n-2} \max\{0, \tilde{s}_i^f - w_i\} \right) \prod_{i=1}^{n-2} q_i(w_i) + \prod_{i=1}^{n-2} q_i(w_i) \sum_{i=n-1}^{n} \sum_{w_i=0}^{k_{0_i}} \Big( \max\{0, \tilde{s}_i^f - w_i\} \, q_i(w_i) \Big) \Bigg] + R(n, su, k_0, q)$$

Iteratively, we move the "$\sum_{i=1}^{n-2}$" to the left (outside):

$$\gamma(\tilde{s}^f) = \sum_{i=1}^{n} \sum_{w_i=0}^{k_{0_i}} \Big( \max\{0, \tilde{s}_i^f - w_i\} \, q_i(w_i) \Big) + R(n, su, k_0, q).$$

The following is equivalent to this equation:

$$\gamma(\tilde{s}^f) = \sum_{i=1}^{n} \sum_{w_i=0}^{\tilde{s}_i^f} (\tilde{s}_i^f - w_i)\, q_i(w_i) + R(n, su, k_0, q)$$

and (b) is thus proved.

(c) follows directly from (b).

In the case of discrete uniformly distributed requirements, we set $q_i(w_i) = \frac{1}{k_{0_i}+1}$ $(i \in \{1, 2, \cdots, n\},\ w_i \in \{0, 1, \cdots, k_{0_i}\})$ in the last equation:

$$\begin{aligned}
\gamma(\tilde{s}^f) &= \sum_{i=1}^{n} \sum_{w_i=0}^{\tilde{s}_i^f} (\tilde{s}_i^f - w_i)\, \tfrac{1}{k_{0_i}+1} + R(n, su, k_0, q) \\
&= \sum_{i=1}^{n} \tfrac{1}{k_{0_i}+1} \sum_{w_i=0}^{\tilde{s}_i^f} (\tilde{s}_i^f - w_i) + R(n, su, k_0, q) \\
&= \sum_{i=1}^{n} \tfrac{1}{k_{0_i}+1} (\tilde{s}_i^f (\tilde{s}_i^f + 1) - \tfrac{1}{2} \tilde{s}_i^f (\tilde{s}_i^f + 1)) + R(n, su, k_0, q) \\
&= \tfrac{1}{2} \sum_{i=1}^{n} \tfrac{\tilde{s}_i^f (\tilde{s}_i^f + 1)}{k_{0_i}+1} + R(n, su, k_0, q).
\end{aligned}$$

Thus, (d) is proved.

In order to prove (e) we use $k_{0_i} = k_0$ for $i = 1, 2, \cdots, n$ and:

$$\begin{aligned}
\gamma(\tilde{s}^f) &= \tfrac{1}{2} \sum_{i=1}^{n} \tfrac{\tilde{s}_i^f (\tilde{s}_i^f + 1)}{k_0+1} + R(n, su, k_0, q) \\
&= \tfrac{1}{2} \tfrac{1}{k_0+1} \sum_{i=1}^{n} \left( (\tilde{s}_i^f)^2 + \tilde{s}_i^f \right) + R(n, su, k_0, q) \\
&= \tfrac{1}{2} \tfrac{1}{k_0+1} \sum_{i=1}^{n} (\tilde{s}_i^f)^2 + \tfrac{su}{2(k_0+1)} + R(n, su, k_0, q) \\
&= \tfrac{1}{2} \tfrac{1}{k_0+1} \sum_{i=1}^{n} (\tilde{s}_i^f)^2 + R_1(n; su; k_0)
\end{aligned}$$

where $R_1(n; su; k_0) = \frac{1}{(k_0+1)^n} \sum_{w \in B_{n;k_0} : C[w, su]} (\sum_{i=1}^{n} w_i - su) + \frac{su}{2(k_0+1)}$

follows. ■

### 3.4.2 Reduction of SDDP Problems with Identical Basic Costs and Independent and Identically Distributed Requirements

To begin, we discuss the possibility of reduction of DA MDPs in general.

**Definition 3.4.1.** *Let a matrix* $A = (a_{fl})_{\substack{f=1,\cdots,m \\ l=1,\cdots,m}}$ *be given. If sets of indices* $I_v = \{h_{v-1}+1,\ h_{v-1}+2,\ \cdots,\ h_v\},\ v = 1, 2, \cdots, r$ *with* $h_0 = 0$, $h_{v-1} < h_v,\ h_r = m$ *and a matrix* $A_R$ *exist so that*

$$\begin{aligned} &A_R = (\varrho_{yv})_{\substack{y=1,\cdots,r \\ v=1,\cdots,r}} \text{ and} \\ &\sum_{l \in I_v} a_{fl} = \varrho_{yv}\ \forall\ f \in I_y\ \forall y = 1, \cdots, r,\ \forall\ v = 1, \cdots, r, \end{aligned} \tag{3.4.5}$$

*then* $A$ *is called reducible (in relation to the sets of indices) and* $A_R$ *is the reduced matrix corresponding to* $A$.

**Lemma 3.4.2.** *Let a MDP*$(N = \infty, S, A^M, P, \gamma)$ *(as in Section 2.3.1), a decision function* $d$ *with the corresponding matrix of transition probabilities* $P^d$ *and the vector of average (one-step) reward functions* $\gamma^d$ *be given, where* $P^d$ *satisfies (2.3.3). Furthermore, let a set of indices as in Definition 2.1.1 and a reduced matrix* $P^d_R$ *corresponding to* $P^d$ *in relation to these sets of indices exist and let*

$$\gamma^d_{h_{v-1}+1} = \gamma^d_{h_{v-1}+2} = \cdots = \gamma^d_{h_v}\ \forall\ v = 1, \cdots, r \tag{3.4.6}$$

*be valid for the average (one-step) reward functions* $\gamma^d$.

$\gamma^d_R \in \mathbb{R}^r_+$ *with* $\gamma^d_{R_v} := \gamma^d_{h_v}\ \forall\ v = 1, \cdots, r$ *is the reduced vector corresponding to* $\gamma^d$ *in relation to the sets of indices.*

*Then,* $P^d_R$ *also satisfies (2.3.3). The solution* $(g_R, \nu_R) \in \mathbb{R} \times \mathbb{R}^r$ *with* $\nu_{R_r} = 0$ *of the Poisson equation (2.3.5) for* $P^d_R$ *and* $\gamma^d_R$ *then yields the solution* $(g, \nu) \in \mathbb{R} \times \mathbb{R}^m$ *of the Poisson equation (2.3.5) for* $P^d$ *and* $\gamma^d$ *in the following way*

$$\begin{aligned}&g = g_R \ \textit{and}\\&\nu_{h_{v-1}+1} = \nu_{h_{v-1}+2} = \cdots = \nu_{h_v} = \nu_{R_v} \ \forall\, v = 1, \cdots, r\\&(\textit{with}\ \nu_m = 0).\end{aligned} \tag{3.4.7}$$

This Lemma can be simply shown using (3.4.5).

**Definition 3.4.2.** *Let a DA MDP*$(N = \infty, \tilde{S}, A^M, P, \gamma)$ *(as in Section 2.3.2) be given, which is constructed by means of DA decision functions* $\hat{d}$, *DA decision sets* $\hat{A}(\tilde{s}, w)$ *and internal costs* $\hat{c}$ *which satisfy (2.3.13), (where* $\tilde{S} = \{\tilde{s}^1, \tilde{s}^2, \cdots, \tilde{s}^m\}$*). Furthermore, let* $I_v = \{h_{v-1}+1,\ h_{v-1}+2,\ \cdots,\ h_v\}$, $v = 1, 2, \cdots, r$ *with* $h_0 = 0,\ h_{v-1} < h_v,\ h_r = m$ *be sets of indices.*

*Then,*

$$\tilde{S}^v := \{\tilde{s}^{h_{v-1}+1}, \tilde{s}^{h_{v-1}+2}, \cdots, \tilde{s}^{h_v}\}, \quad v = 1, 2, \cdots, r \tag{3.4.8}$$

*are subspaces of states related to the sets of indices.*

$S = \{s^1, s^2, \cdots, s^r\}$ *is called reduced state space where*

$$s^v := \tilde{s}^{h_{v-1}+1} \tag{3.4.9}$$

*are representatives of the sets* $\tilde{S}^v \quad (v = 1, 2, \cdots, r)$.

$\hat{A}(s^v, w) = \{s^{y_1}, s^{y_2}, \cdots, s^{y_{z(v)}}\}$ *is called the reduced DA decision set (in relation to the sets of indices), if* $w^\alpha \in B$ *for* $\alpha = 1, 2, \cdots, h_v - h_{v-1}$ *with* $w^1 = w$ *and* $\tilde{S}^{y_\beta}_\alpha \subseteq \tilde{S}^{y_\beta}$ *for* $\beta(= \beta(v)) = 1, 2, \cdots, z(v)$ *exist so that*

$$\left.\begin{aligned}&\hat{A}(\tilde{s}^{h_{v-1}+\alpha}, w^\alpha) = \{\tilde{S}^{y_1}_\alpha, \tilde{S}^{y_2}_\alpha, \cdots, \tilde{S}^{y_{z(v)}}_\alpha\},\\&q(w^1) = q(w^2) = \cdots = q(w^{h_v - h_{v-1}})\ \textit{and}\\&\hat{c}(\tilde{s}^{h_{v-1}+1}, w^1) = \hat{c}(\tilde{s}^{h_{v-1}+2}, w^2) = \cdots = \hat{c}(\tilde{s}^{h_v}, w^{h_v - h_{v-1}})\end{aligned}\right\}. \tag{3.4.10}$$

$$\hat{c}(s^v, w) := \hat{c}(\tilde{s}^{h_v}, w^{h_v - h_{v-1}}) \tag{3.4.11}$$

*is defined.*

*The set of DA decision sets* $\{\hat{A}(\tilde{s}, w) \mid \tilde{s} \in \tilde{S}, w \in B\}$ *is called reducible if reduced DA decision sets exist for all* $v = 1, 2, \cdots, r,\ w \in B$.

*If the set of DA decision sets is reducible, then a DA decision function* $\hat{d}$ *with*

$$\hat{d}(\tilde{s}^{h_{v-1}+\alpha}, w^{\alpha}) \in \tilde{S}^{y_{\beta(v)}},\ \forall\ \alpha = 1, 2,\ \cdots,\ h_v - h_{v-1},\ (v = 1, 2, \cdots, r) \tag{3.4.12}$$

*is also called reducible (in relation to given the sets of indices).*

*In addition,*

$$\hat{d}(s^v, w = w^1) = s^y_{\beta(v)}(:= \tilde{s}^{h_{y_{\beta(v)}-1}+1}),\ v = 1, 2, \cdots, r,\ w \in B \tag{3.4.13}$$

*is the reduced DA decision function.*

**Lemma 3.4.3.** *Let a DA MDP*$(N = \infty, \tilde{S}, A^M, P, \gamma)$ *(as in Definition 3.4.2) be given where the set of DA decision sets is reducible with respect to the sets of indices* $I_v = \{h_{v-1} + 1,\ h_{v-1} + 2,\ \cdots,\ h_v\},\ v = 1, 2, \cdots, r$ *with* $h_0 = 0,\ h_{v-1} < h_v,\ h_r = m$.

*If a DA decision function* $\hat{d}$ *is reducible (in relation to the sets of indices), then a reduced matrix of transition probabilities* $P^d_R$ *which corresponds to* $P^d$ *and a reduced vector of average (one-step) reward functions* $\gamma^d_R$ *corresponding to* $\gamma^d$ *exist.*

***Proof.*** That $P^d$ is reducible follows simply from (2.3.7) together with (3.4.12) and the second equations of (3.4.10).

This along with the first equation from (2.3.8) and the third equations from (3.4.10) show the reducibility of the average (one-step) reward functions. ■

**Lemma 3.4.4.** *Let a DA MDP*$(N = \infty, \tilde{S}, A^M, P, \gamma)$ *(as in Definition 3.4.2) be given where the set of DA decision sets is reducible with respect to the set of indices* $I_v = \{h_{v-1} + 1,\ h_{v-1} + 2,\ \cdots,\ h_v\},\ v = 1, 2, \cdots, r$ *with* $h_0 = 0, h_{v-1} < h_v,\ h_r = m$. *Furthermore, let (2.3.3) be valid for all stationary policies.*

*Then, an optimal DA decision function exists which is reducible.*

***Proof.*** Based on the given DA MDP we construct a new corresponding reduced DA MDP, whose state space is the reduced state space (see Definition 3.4.2).

Further characteristic quantities of the reduced DA MDP are constructed by means of the reduced DA decision sets, the reduced DA decision functions and the internal costs from (3.4.11), as in Section 2.3.2.

According to Lemma 3.4.3, a reduced matrix of transition probabilities $P_R^d$ corresponding to $P^d$ and a reduced vector of average (one-step) reward functions $\gamma_R^d$ which corresponds to $\gamma^d$ (in relation to the given sets of indices) exist for any reducible DA decision function $\hat{d}$.

Now, let $\hat{d}^*$ be a reducible DA decision function whose reduced DA decision function is an optimal decision function of the reduced DA MDP.

Then the optimality criterion (2.3.22) is satisfied in relation to this reduced DA decision function and the reduced DA MDP.

Solutions of the Poisson equation (2.3.5) with respect to the reducible DA decision function follow from solutions of the Poisson equation for the optimal reduced DA decision function, as in Lemma 3.4.2.

Since the third equations of (3.4.10) are valid for the internal costs and (3.4.7) for solutions of the Poisson equations, the optimality criterion (2.3.22) is also satisfied in relation to the reducible DA decision function $\hat{d}^*$ and the given DA MDP. ■

We now discuss:

**The Reduction of SDDP Problems with Identical Basic Costs and Independent and Identically Distributed Requirements**

Under the conditions of (3.4.1), the internal costs satisfy (2.3.13) (as discussed at the beginning of Section 3.4).

If $\tilde{s}$ (ordered partition of $su$) is a state of the state space $\tilde{S}_{n;su;k_0}$, then all permutations of $\tilde{s}$ are also elements of this state space since $k_{0_1} = k_{0_2} = \cdots = k_{0_n}$ (see (3.4.4)).

Hence, we use the set $S_{n;su;k_0}$ of the unordered partitions of $su$ with at most $n$ parts with summands not greater than $k_0$, (where the elements of $S_{n;su;k_0}$ are expressed as $n$-dimensional vectors, meaning $s = (s_1, s_2, \cdots, s_n)$) as the reduced state space corresponding to $\tilde{S}_{n;su;k_0}$ if (3.4.1) and (3.4.4) are supposed.

(Let $r := |S_{n;su;k_0}|$ be the number of the restricted partitions.)

That implies that the sets of indices (see Definition 3.4.2) are determined such that each corresponding subspace of states which is related to the set of indices includes all permutations of a state from the original state spaces.

Since $w_i,\ i = 1, 2, \cdots n$ are independent and identically distributed, the probabilities of permutations $w_\pi$ of $w$ are equal: $q(w_\pi) = q(w)$.
(Thus, the second equations of (3.4.10) are fulfilled.)

If

$$\hat{A}_{n;su;k_0}(\tilde{s}, w) = \{\tilde{s}^{y_1}, \tilde{s}^{y_2}, \cdots, \tilde{s}^{y_{\tilde{z}}}\} \tag{3.4.14}$$

is a DA decision set, $\tilde{s}_\pi$ a permutation of $\tilde{s}$ and $w_\pi$ an analogous permutation of $w$ then the DA decision set $\hat{A}_{n;su;k_0}(\tilde{s}_\pi, w_\pi)$ is:

$$\hat{A}_{n;su;k_0}(\tilde{s}_\pi, w_\pi) = \{\tilde{s}_\pi^{y_1}, \tilde{s}_\pi^{y_2}, \cdots, \tilde{s}_\pi^{y_{\tilde{z}}}\},$$

where $\tilde{s}_\pi^{y_\beta}$ $(\beta = 1, 2, \cdots, \tilde{z})$ are permutations of $\tilde{s}^{y_\beta}$, which are analogous to the permutation $\tilde{s}_\pi$ of $\tilde{s}$.

Furthermore, (3.4.1) yields

$$\hat{c}(\tilde{s}, w) = \hat{c}(\tilde{s}_\pi, w_\pi)$$

for a permutation $\tilde{s}_\pi$ of $\tilde{s}$ and an analogous permutation $w_\pi$ of $w$.

Hence, $\hat{A}_{n;su;k_0}(\tilde{s}, w)$ are reducible for any $\tilde{s} \in \tilde{S}_{n;su;k_0},\ w \in B_{n;k_0}$ (see Definition 3.4.2).

Corresponding reduced DA decision sets are $\hat{A}_{n;su;k_0}(s, w) = \{s^{y_1}, s^{y_2}, \cdots, s^{y_z}\}$ $(s \in S_{n;su;k_0},\ w \in B_{n;k_0})$ if (3.4.14) is assumed.

Together with (3.2.11) this implies

$$\hat{A}_{n;su;k_0}(s, w)$$
$$= \left\{ s' \in S_{n;su;k_0} \;\middle|\; \begin{array}{l} \exists\ s'_\pi \text{ permutation of } s' : \\ w_i \le \tilde{s}'_{\pi i} \le max\{s_i, w_i\},\ i = 1, \cdots, n, \text{ in } C[w, su], \\ \\ min\{s_i, w_i\} \le s'_{\pi i} \le w_i,\ i = 1, \cdots, n, \text{ in } C[su, w] \end{array} \right\}. \tag{3.4.15}$$

According to Lemma 3.4.4, optimal reducible DA decision functions for SDDP problems exist if (3.4.1) and (3.4.4) are supposed. These optimal reducible DA decision functions can be completely described by means of optimal DA decision functions of the corresponding reduced SDDP problems and corresponding permutations.

Average one-step reward functions for reduced SDDP problems can be computed by means of the formula from Theorem 3.4.1(c) (consider also (3.4.6) and Lemma 3.4.3). [12]

In Section 4.7 we will show that the conjecture (Con3.3) is true for SDDP problems (with identical basic costs and discrete uniformly distributed requirements) with arbitrary but fixed number of states, with exeption of perhaps a finite number of such problems (see Corollary 4.7.1).

Here we also introduce the symbol $d^*$:

Under the conditions (3.4.1) and (3.4.4), let a reduced SDDP problem be given.

$d^*$ denotes the DA decision functions with decisions for feasible states with minimum average one-step reward functions. (3.4.16)

## 3.5 Notes on the Solution Methods of SDDP Problems

To begin we discuss possibilities to compute exact solutions of SDDP problems.

However, if the number of states is very large, then the computation of exact solutions is hardly realizable. Therefore, we deal in the final part of

[12] This cannot be more simply computed for the following reasons: The internal costs of (not reduced) SDDP problems fulfil (2.3.12) (as discussed in Sections 3.2 and 3.3), which means $\hat{c}(\tilde{s}^f, w, \tilde{s}^l) = \hat{c}(\tilde{s}^f, \tilde{s}^l)$ where $\tilde{s}^l \in \hat{A}_{n;su;k_0}(\tilde{s}^f, w)$.

However, (2.3.12) is not satisfied for reduced SDDP problems (with identical basic costs and independent and identically requirements) since different permutations of $s^l$ fulfil the inequalities of (3.4.15) for differing $w$, in general.

this section with heuristics.

In principle, we can use all of this for the solution methods of SDDP problems (modelled as DA MDPs), as was generally stated in Section 2.3.5 for the solution methods of DA MDPs.

SDDP problems can therefore be exactly solved by means of the Howard algorithm prepared for DA MDPs (see Section 2.3.5).

In addition, all decision sets $\hat{A}_{n;su;k_0}(\tilde{s}, w)$ must be determined according to (3.2.11) and the internal costs $(\hat{c}(\tilde{s^f}, \tilde{s^l}))_{\substack{f=1,\dots,r \\ l=1,\dots,r}}$ (as a basis for average (one-step) reward functions, see (2.3.8)) must first be computed according to (3.2.12).

Decisions as in Section 3.4, thus also optimal decisions of corresponding SDDP problems with identical basic costs, may come into consideration as initial decisions for the Howard algorithm.

If one would like to solve SDDP problems as cost-parametric DA MDPs according to Algorithm 2.3.2 (Section 2.3.4.2), or theoretically investigate in this way, then the following specifications should be applied to Algorithm 2.3.2:

Regarding the step 1 of Algorithm 2.3.2, suggesting a parameterization of the costs (as described in Section 2.3.4.3) is to be made here on the basis of the basic cost, for instance in the following way:

$k_{i_0j_0} + \vartheta(k_{ij} - k_{i_0j_0})$ with

$$k_{i_0j_0} = min\{k_{ij} \mid i, j \in \{1, 2, ..., n\},\ i \neq j\},\ \vartheta \in [0, 1]\ .$$

Without loss of generality, according to Lemma 3.3.5 and Lemma 3.3.1, we can set:

$$\begin{aligned} &1 = min\{k_{ij} \mid i, j \in \{1, 2, ..., n\},\ i \neq j\}\ (= k_{i_0j_0}) \\ &\text{and the parameterization } 1 + \vartheta(k_{ij} - 1) \end{aligned} \tag{3.5.1}$$

follows.

This means (3.4.1) and thus (2.3.13) are valid in relation to the initial parameter $\vartheta = 0$.

Then, (considering Lemma 3.3.6) the assumptions for (LPC) and (2.3.42)

are also fulfilled for the resulting cost-parametric DA Markov decision process, where

$$\{\kappa_0,\ \kappa\} \subseteq \mathbb{R}^n_+ \times \mathbb{R}^n_+,\ \kappa_0 = \left(\delta_{ij} = \begin{cases} 1 & \text{if } i \neq j, \\ 0 & \text{if } i = j \end{cases}\right)_{\substack{i=1,\ldots,n \\ j=1,\ldots,n}},$$

$$\kappa = (k_{ij} - \delta_{ij})_{\substack{i=1,\ldots,n \\ j=1,\ldots,n}}.$$

In relation to step 1 of Algorithm 2.3.2 (where $\vartheta = 0$) and its application to SDDP problems with the above parameterization of the costs, we find that:

- either decisions for states with minimal average (one-step) reward functions are already optimal (refer to the Conjectures (Con3.1), (Con3.2) and (Con3.3) and Corollary 4.7.1 in Section 4.7)

- or the the complexity of computing the optimal decisions can be reduced (see Theorem 2.3.8 and Corollary 2.3.9) since in this case an almost-partial order of the states exists (as discussed in detail following Algorithm 2.3.2 in Section 2.3.4.2).

If the parameter $\vartheta$ increases, then the violations of the optimality are single violations (Definition 2.3.6) in general and the optimal decisions can be purposefully computed for the increasing parameter $\vartheta > 0$ (as discussed in detail in Algorithm 2.3.2).

It is possible that the additional conditions (AC1), (AC2) and (AC3) from Section 2.3.4.2 are satisfied for SDDP problems or for certain subsets of such problems. If (AC3) is valid, then step 2 of Algorithm 2.3.2 is a greedy algorithm as discussed in Section 2.3.4.2.

For the following Example 3.5.1 this is indeed the case.

Proofs of such statements have, however, not yet been completed. These could be just as complicated as the proof of Conjecture (Con3.3) in Chapter 4.

**Example 3.5.1.** *We consider SDDP problems with*

$n = 3,\ su = 5,\ k_0 = (3, 3, 3),$

$w_i\ (i = 1, 2, 3)$ *discrete uniformly distributed random variables,*

$$(k_{ij})_{\substack{i=1,2,3\\ j=1,2,3}} = \begin{pmatrix} 0 & 1 & k \\ 1 & 0 & k \\ k & k & 0 \end{pmatrix}, \quad k \in [1, \infty)$$

*(thus the distance properties (3.1.2) are valid for all* $k \in [1, \infty)$*).*

*Let the SDDP problems be modelled as DA MDPs (for* $k \in [1, \infty)$*).*

*Initially, the state spaces include 12 states. However the SDDP problems can be reduced, so that the state spaces of the reduced problems include only 7 states:*

$$s^1 = \begin{pmatrix} 3 \\ 2 \\ 0 \end{pmatrix}, s^2 = \begin{pmatrix} 3 \\ 0 \\ 2 \end{pmatrix}, s^3 = \begin{pmatrix} 2 \\ 0 \\ 3 \end{pmatrix}, s^4 = \begin{pmatrix} 3 \\ 1 \\ 1 \end{pmatrix}, s^5 = \begin{pmatrix} 1 \\ 1 \\ 3 \end{pmatrix},$$

$$s^6 = \begin{pmatrix} 2 \\ 2 \\ 1 \end{pmatrix}, s^7 = \begin{pmatrix} 2 \\ 1 \\ 2 \end{pmatrix},$$

*by reason of the given basic costs* $(k_{ij})_{\substack{i=1,2,3\\ j=1,2,3}}$ *and the discrete uniformly distributed requirements* $w_i$ $(i = 1, 2, 3)$.

*We can consider the SDDP problems (with* $k \in [1, \infty)$*) as cost-parametric SDDP problems (where* $k = 1 + \vartheta$*).*

*Calculations by means of a computer led to the following results:*

*Decisions for feasible states with minimal average one-step reward functions are indeed optimal for* $k = 1$ *(thus for* $\vartheta = 0$*).*

*If the parameter* $\vartheta$ *increases, then the violations of the optimality are single violations (see Definition 2.3.6) and the following single decisions have to be changed for optimal solutions:*

*A:* $\hat{d}(s^f, w) = s^7$ *to* $\hat{d}(s^f, w) = s^4$

*for* $f = 1, 4, 6$ *and certain calculated* $w \in B$,

*B:* $\hat{d}(s^f, w) = s^6$ *to* $\hat{d}(s^f, w) = s^1$

*for* $f = 1$ *and certain certain calculated* $w \in B$,

*C:* $\hat{d}(s^f, w) = s^4$ *to* $\hat{d}(s^f, w) = s^2$

*for* $f = 2, 3, 5, 7$ *and certain certain calculated* $w \in B$,

*D:* $\hat{d}(s^f, w) = s^7$ *to* $\hat{d}(s^f, w) = s^5$

*for* $f = 3$ *and certain certain calculated* $w \in B$,

*E:* $\hat{d}(s^f, w) = s^7$ *to* $\hat{d}(s^f, w) = s^3$

*for* $f = 3, 5$ *and certain certain calculated* $w \in B$,

*more detailed:*

| | | | | | |
|---|---|---|---|---|---|
| $\vartheta$ : | *0, 272* | *0, 466* | *1, 276* | *2,393* | *5, 101* |
| *Change:* | *C* | *A* | *D* | *E* | *B.* |

*(Single decisions with corresponding minimal internal costs are optimal for* $\vartheta \geq 5,101$*.)*

*In addition, we note initially that*

$$\sum_{i=1}^{3}(s_i^1)^2 = \sum_{i=1}^{3}(s_i^2)^2 = \sum_{i=1}^{3}(s_i^3)^2 = 13 > \sum_{i=1}^{3}(s_i^4)^2 = \sum_{i=1}^{3}(s_i^5)^2 = 11$$

$$> \sum_{i=1}^{3}(s_i^6)^2 = \sum_{i=1}^{3}(s_i^7)^2 = 9$$ *and the changes* $C, A, D$ *(for smaller* $\vartheta$*)*

*include the differences of square sums of the corresponding states:*

$2 = 13 - 11$ *for* $C$ *and* $2 = 11 - 9$ *for* $A$ *and* $D$ *contrary to the changes* $B$ *and* $E$ *(for greater* $\vartheta$*) with differences* $4 = 13 - 9$.

*Furthermore, the additional conditions (AC1), (AC2) and (AC3) from Section 2.3.4.2 are definitely valid for this example.*

*However the validity of (AC3) seems to only be based on the the small variation between the states* $s^1, s^2, \cdots, s^9$.

A great difficulty of computing exact solutions of SDDP problems is:

The number of states of the state spaces can grow rather large for increasing $n$ and $k_0$. For example, state spaces with more than fifty billion states exist only for $n = 10, k_{0_1} = \cdots = k_{0_{10}} = 19$.

Then, obviously, $B_{n;k_0}$ has $20^{10}$ elements. Integers from 1 to $10 * 19 \approx 200$ are possible for $su$ (see (3.2.3)). Thus, the average size of corresponding state spaces is approximately $\frac{20^{10}}{200} \approx 5 * 10^{10}$ states.

These considerations demonstrate the usefulness of theoretical investiga-

tions of SDDP problems and of the possibilities for rough estimations of the quality of feasible solutions and of approximation methods.

We will now discuss a few different approaches for approximation methods.

The first approach includes only small reductions of state spaces. Parts which are to be produced and are nearly equal, will be regarded as equal for the mathematical modelling.

Moreover, the basic costs matrix $(k_{ij})_{\substack{i=1,\dots,n\\ j=1,\dots,n}}$ can be simplified in such a manner that it only includes low and high costs $k_{ij}$, $i \neq j$. (Thus, without loss of generality, $k_{ij} = 1$ (low) and $k_{ij} = k > 1$ (high).) Then, (as in Example 3.5.1) the state space can be reduced to a certain extent. For the resulting problem an exact solution must then be found.

Further theoretical investigation, for instance with regard to the validity of the additional conditions (AC1), (AC2) and (AC3) from Section 2.3.4.2, seems interesting for these approximation methods.

Another approach for heuristics diverts from the deductive conception which includes all feasible transitions immediately.

In each stage such approaches use the present state $s^f$, the present realization $w$ of requirements, the DA decision set $\hat{A}(s^f, w)$, the internal costs $\hat{c}(s^f, w, s^l)$ with $s^l \in \hat{A}(s^f, w)$ and quantities, which are depend on probabilities in a simple way.

The method at the end of Section 2.3.5 can be used in order to find roughly approximate solutions (see this section).

Additionally, the relation $s^{\bar{l}} < s^l$ (which is implied by the almost-partial order, see Theorem 2.3.8) can be replaced by relations such as

$$\gamma(s^{\bar{l}}) = \sum_{w' \in B_{n;k_0}} \left( \sum_{i=1}^{n} \max\{0, s_i^{\bar{l}} - w_i'\} \right) q(w') + R(n, su, k_0, q)$$

$$> \gamma(s^l) = \sum_{w' \in B_{n;k_0}} \left( \sum_{i=1}^{n} \max\{0, s_i^{l} - w_i'\} \right) q(w) + R(n, su, k_0, q),$$

thus

$$\sum_{w' \in B_{n;k_0}} \left( \sum_{i=1}^{n} \max\{0, s_i^{\bar{l}} - w_i'\} \right) q(w') > \sum_{w' \in B_{n;k_0}} \left( \sum_{i=1}^{n} \max\{0, s_i^{l} - w_i'\} \right) q(w)$$

(see Section 3.4, Theorem 3.4.1).

Furthermore, the condition $\hat{c}(s^f, w, s^{\bar{l}}) \ll \hat{c}(s^f, w, s^l)$ (see the method at the end of Section 2.3.5) for changes of the single decision $d^*(s^f, w)$ could be connected with the differences of the states $s^l$ and $s^{\bar{l}}$ expressed by the corresponding square sums of these states (see also the note corresponding to Example 3.5.1).

Another heuristic of this type can be found in [22], Chapter 2 or [24], Section 4.1. The probabilities of "optimum domains" are maximized as surrogate problems for SDDP problems. These surrogate problems are a kind of two stage problem and are suitable for SDDP problems with log-concave distributed disturbances.

Finally, we note that another origin for heuristics is the on-line optimization. At the end of Section 1.2 we have referred to connections with k-server problems. Several ideas from on-line algorithms for k-server problems could be included in heuristics for SDDP problems. Among other things and also concerning probabilities, $k_{ij}$ themselves would be used in order to determine heuristic solutions. $c_{ij}$ would play a role if the states of SDDP problems modelled as DA MDPs (which means partitions) are considered as states of metric task systems. (However, keep in mind that probability functions are given for the requirements of SDDP problems as noted in Section 1.2.)

Investigations with regard to useful heuristics for SDDP problems are far from being finished.

# Chapter 4

# Partitions-Requirements-Matrices

Partitions-requirements-matrices (PRMs) are on the one hand matrices of transition probabilities for certain SDDP problems which are modelled as DA MDPs. On the other hand, PRMs themselves represent interesting (almost self-evident) combinatorial structures, which have not yet been discussed in literature.

General PRMs are constructed on the basis of ordered restricted partitions of integers and PRMs ("in the strict meaning") on the basis of unordered restricted partitions of integers.

PRMs (in the strict meaning) are matrices of transition probabilities for reduced SDDP problems (see Section 3.4.2) and for decisions for feasible states with least square sums of their components. In Section 4.7 it is shown that these decisions are optimal for a great (infinite) number of SDDP problems.

In Chapter 4 we dealt primarily with PRMs in the strict meaning.

Special consideration was taken to ensure that the treatise of Chapter 4 can be essentially understood independent of Chapters 2 and 3. *Relationships to Chapter 3 are marked extra.* They can be omitted if one is only interested in PRMs.

As a partial order on (un/)ordered restricted partitions of integers we use dominance (or majorization) ordering (see Section 4.1) which can, for instance, be found by Marshall, A.W. and Olkin, I. (see [26]).

The definition of PRMs (in the strict meaning) includes that PRMs can initially be computed by means of simple enumeration. However this is a laborious method. To date no formulas are known for most of the elements of PRMs.

In Section 4.3 it is demonstrated that by making use of the definition of "perturbed partitions" elements of PRMs can be computed more effectively (in comparison to the enumeration). However, permutations with certain characteristics must additionally be determined. This section, as also Section 4.5 may be skipped by readers, which are only interested in results that have meaning for SDDP problems.

Limits of elements of certain PRMs are computed in Section 4.4. In order to accomplish this, sets of (unordered) restricted partitions of integers must be classified before hand.

In Section 4.5 (partial) results of PRM elements are given.

A polynomial, and sometimes an exponential, dependence of the elements of PRMs on the parameters, which determines the restrictions of the partitions, will be shown in the case of discrete uniformly distributed requirements by means of perturbed partitions.

Formulas for the elements of the last row and the last column of certain PRMs are also given.

Poisson equations which are based on PRMs are considered in Section 4.6. The "monotonicity of their solutions" is proven in many cases. *This implies that the decisions for feasible states with least square sums of their components are optimal for the corresponding reduced SDDP problems.*

The solutions of the Poisson equations, with regard to the limits of PRMs, have an elegant structure, in contrast to the formulas for the limits of PRMs. These solutions include in relation to the distribution of requirements generalized harmonic numbers.

We use the symbol "$s$" from here on for the denotation of the partitions

in agreement with the denotation of the **s**tates in the previous chapters. An $s$ with the additional symbol ˜ signifies an ordered partition and serves as distinction from the unordered partitions.

## 4.1 Arrangement of the Partitions

This section serves as basis for what follows.

As a partial order on (un/)ordered restricted partitions of integers we use the dominance or majorization ordering (see [26], Marshall, A.W. and Olkin, I.,Chapter 1. A., B. and Chapter 5. D.).

*In relation to the reduced SDDP problems (see Section 3.4.2), we will see later that the above mentioned partial order includes the almost-partial order of the states of the reduced SDDP problems. According to Theorem 2.3.8 and Corollary 2.3.9 this is implied by optimal solutions. (This will be proven for most cases).*

The dominance or majorization (ordering) includes that

- the "difference" between neighbouring partitions is as small as possible,

- the square sums of the parts of the partitions become smaller. (See also the following Definition 4.2.3 of PRMs (in the strict meaning).)

The dominance (ordering) implies that sets of restricted partitions of integers are lattices.

In particular the terminology "main minimum chain" will be introduced, *which is important for the dominance of corresponding SDDP problems (see Section 4.6.2.2).*

As usual, we introduce ordered and unordered partitions first. Then we will consider only the unordered (in general restricted) partitions.

A more detailed representation of the topic in this section can be found in [21].

Initially, for $y \in \mathbb{Z}^n$ we define the vector $y[i_1; i_2]$ with

$$y_i[i_1; i_2] = \begin{cases} y_i + 1 & \text{for} \quad i = i_1, \\ y_i - 1 & \text{for} \quad i = i_2, \\ y_i & \text{otherwise.} \end{cases} \tag{4.1.1}$$

Furthermore, let

- $n \in \mathbb{N}, \ \ n \geq 3$ (or 2), $su \in \mathbb{N}$
- either $k_0 \in \mathbb{Z}_+$ with

$$k_0 \leq su < n\, k_0 \tag{4.1.2}$$

  or $k_0 \in \mathbb{Z}_+^n$ with

$$(\max_{i \in \{1, \cdots, n\}} k_{0_i} \leq)\ su < \sum_{i=1}^{n} k_{0_i} \tag{4.1.3}$$

- $\tilde{S}_{n;su;k_0} = \{\tilde{s} \in \mathbb{Z}_+^n \mid 0 \leq \tilde{s}_i \leq k_{0_i} \text{ for } i = 1, \cdots, n, \sum_{i=1}^{n} \tilde{s}_i = su\}$ [1]
  - the set of the ordered partitions of $su$ into at most $n$ parts, each not greater than $k_{0_i}$:
  In this book the elements of $\tilde{S}_{n;su;k_0}$ are expressed as $n$-dimensional vectors, meaning that $\tilde{s} = (\tilde{s}_1, \tilde{s}_2, \cdots, \tilde{s}_n)$ (with $\tilde{s}_i \leq k_{0_i}$).
  An element of $\tilde{S}_{n;su;k_0}$ is called an ordered restricted partition (of $su$) and $\tilde{s}_i$ a part or component.
  $\tilde{r} := |\tilde{S}_{n;su;k_0}|$ is the number of the restricted ordered partitions.

- $S_{n;su;k_0}$ - the set of the (unordered) partitions of $su$ into at most $n$ parts, each not greater than $k_0$:
  The elements $s$ of $S_{n;su;k_0}$ are also expressed as $n$-dimensional vectors, meaning that $s = (s_1, s_2, \cdots, s_n)$. ((Unordered) partitions $s$

[1] We will continue to use the symbol "s" for the denotation of the partitions in agreement with the denotation of the states in the previous chapters.

and $s'$ are equal if $s'$ is a permutation of $s$.)

An element of $S_{n;su;k_0}$ is called a restricted partition (of $su$).

$r := |S_{n;su;k_0}|$ is the number of the restricted partitions. [2]

**Definition 4.1.1.** .

(a) *The restricted partitions $s^f \in S_{n;su;k_0}$ and $s^l \in S_{n;su;k_0}$ are called neighbouring if there permutations $s^f_\pi$ of $s^f$ and $s^l_\pi$ of $s^l$ exist such that* $\frac{1}{2}\sum_i \mid s^f_{\pi_i} - s^l_{\pi_i} \mid = 1$.

*If $\sum_i (s^f_i)^2 > \sum_i (s^l_i)^2$, then $s^f$ is called a direct predecessor of $s^l$ and $s^l$ a direct successor of $s^f$ (with symbols: $s^f \dot{\rightarrow} s^l$).*

(b) *A sequence of partitions $s^{f_1}, s^{f_2}, \cdots, s^{f_z}$ is called a chain [3] if $s^{f_j}$ is a direct predecessor of $s^{f_{j+1}}$ for each $j \in \{1, \cdots, z-1\}$. In this case $s^{f_1}$ is called the least element and $s^{f_z}$ the greatest element of the chain.*

(c) *$s^f$ is called a predecessor of $s^l$ and $s^l$ successor of $s^f$ if there is a chain with least element $s^f$ and greatest element $s^l$ (with symbols: $s^f \rightarrow s^l$).*

(d) *Let $SP$ be a subset of $S_{n;su;k_0}$. A restricted partition $s^f \in SP$ is called the least element of $SP$ if $s^f \rightarrow s$ for each $s \in SP$ $(s \neq s^f)$. Moreover, $s^l \in SP$ is called the greatest element of $SP$ if $s \rightarrow s^l$ for each $s \in SP$ $(s \neq s^l)$.*

An equivalent definition of neighbouring partitions is yielded by

**Lemma 4.1.1.** *Let $s^f \in S_{n;su;k_0}$, $s^l \in S_{n;su;k_0}$ be given with $s^y_1 \geq s^y_2 \geq \cdots \geq s^y_n$ for $y = 1, 2$.*

*Then, $s^f$ is a direct predecessor of $s^l$ if and only if $\exists\, i_1, i_2 :\ (s^l_{i_1} \geq s^l_{i_2}) \wedge (s^l = s^f[i_2; i_1])$.*

[2] The theory of (unordered) partitions is complicated and presents a number of interesting problems, see [3].

[3] In the following we use the above definition of the chain, which varies slightly from the usual definition of a chain.

***Proof.***

1. ($\Rightarrow$): Let $s_1^f \geq s_2^f \geq \cdots \geq s_n^f$ and let $s^f$ be a direct predecessor of $s^l$ according to Definition 4.1.1(a).

   Since $\frac{1}{2}\sum_i |\, s_i^f - s_{\pi_i}^l \,| = 1$ for a permutation $s_\pi^l$ of $s^l$, the partitions $s^f$ and $s^l$ are different in only two parts:

   $s_{\pi_i}^l = s^f[j_0; i_0]$ with $i_0 > j_0$ and $s_{i_0}^f > s_{j_0}^f + 1$ because
   $$\sum_i (s_i^f)^2 > \sum_i (s_i^l)^2.$$

   If $s_{i_0}^f = s_{i_0+1}^f = \cdots = s_{i_0+\alpha}^f > s_{i_0+\alpha+1}^f$ then $s_{\pi_{i_0}}^l = s_{i_0}^f - 1$ should be exchanged with $s_{\pi_{i_0}+\alpha}^l (= s_{i_0+\alpha}^f)$.

   If $s_{j_0}^f = s_{j_0-1}^f = \cdots = s_{j_0-\beta}^f < s_{j_0-\beta-1}^f$ then $s_{\pi_{j_0}}^l = s_{j_0}^f + 1$ should be exchanged with $s_{\pi_{j_0}-\beta}^l\ (= s_{j_0}^f)$.

   Then, the coordinates of $s_\pi^l$ are also ordered monotonically increasing and the condition from Lemma 4.1.1 is satisfied.

2. ($\Leftarrow$): If the condition from Lemma 4.1.1 is fulfilled, then it immediately follows from Definition 4.1.1(a) that $s^f$ is a direct predecessor of $s^l$.

■

**Lemma 4.1.2.** *$S_{n:su;k_0}$ has a least and a greatest element ($s^1$ and $s^r$, respectively). Furthermore, for any $s \in S_{n;su;k_0}$ a chain with $s$, the least element being $s^1$ and the greatest element $s^r$, exists.*

(The proof can be found in [20], Lemma 4.6.)

**Example 4.1.1.** *Let $S_{n;su;k_0}$ with $n = 3$, $su = 9$ and $k_0 = 5$ be given:*

$$S_{3;9;5} = \left\{ s^1 = \begin{pmatrix} 5 \\ 4 \\ 0 \end{pmatrix},\ s^2 = \begin{pmatrix} 5 \\ 3 \\ 1 \end{pmatrix},\ s^3 = \begin{pmatrix} 5 \\ 2 \\ 2 \end{pmatrix},\ s^4 = \begin{pmatrix} 4 \\ 4 \\ 1 \end{pmatrix},\right.$$
$$\left. s^5 = \begin{pmatrix} 4 \\ 3 \\ 2 \end{pmatrix},\ s^6 = \begin{pmatrix} 3 \\ 3 \\ 3 \end{pmatrix} \right\}.$$

*The relations: direct predecessor* $\dot{\rightarrow}$ *direct successor are represented in the following diagram:*

$$\begin{tikzcd}
 & \begin{pmatrix}5\\3\\1\end{pmatrix} \arrow[r, "\bullet"] \arrow[dr, "\bullet"] \arrow[drr, "\bullet"] & \begin{pmatrix}5\\2\\2\end{pmatrix} \arrow[ddr, "\bullet"] & \\
\begin{pmatrix}5\\4\\0\end{pmatrix} \arrow[ur, "\bullet"] \arrow[r, "\bullet"] & \begin{pmatrix}4\\4\\1\end{pmatrix} \arrow[r, "\bullet"] & \begin{pmatrix}4\\3\\2\end{pmatrix} \arrow[r, "\bullet"] & \begin{pmatrix}3\\3\\3\end{pmatrix}
\end{tikzcd}$$

$s^1$ *is the least element and* $s^6$ *is the greatest element of* $S_{3;9;5}$.

**Definition 4.1.2.** *A chain with the least element* $s^f$ *and the greatest element* $s^l$ *is called a minimal chain if no other chain with the same least and greatest elements and fewer partitions exits.*

**Definition 4.1.3.** *A chain* $s^{f_1}, s^{f_2}, \cdots, s^{f_z}$ *(where, without loss of generality* $s_1^{f_y} \geq s_2^{f_y} \geq \cdots \geq s_n^{f_y}$ *for* $y = 1, \cdots, z$*) with the property*

$$s^{f_y} = s^{f_{y-1}}[j_y; i_y] \text{ where } s_{i_y}^{f_{y-1}} = \max\{s_i^{f_{y-1}} \mid s_i^{f_{y-1}} > s_i^{f_z}\}$$

$$\text{and } s_{j_y}^{f_{y-1}} = \min_i\{s_i^{f_{y-1}} \mid s_i^{f_{y-1}} < s_i^{f_z}\}$$

*for* $y = 2, \cdots, z$ *is called main minimal chain.*

In Example 4.1.1 the chains $s^1, s^2, s^5$ and $s^1, s^4, s^5$ are minimal chains with the least element $s^1$ and the greatest element $s^5$. The latter chain is also a main minimal chain.

**Lemma 4.1.3.** .

*a) A main minimal chain is also a minimal chain.*

*b) A minimal chain with the least element* $s^{f_1}$ *and the greatest element* $s^{f_z}$ *includes*

$\frac{1}{2}\sum_{j=1}^{n} \mid s_j^{f_1} - s_j^{f_z} \mid +1$ *partitions, where* $s_1^l \geq s_2^l \geq \cdots \geq s_n^l$ *for* $l = f_1, f_z$.

$s^{f_y}$ *is an element of a minimal chain with the least element* $s^{f_1}$ *and*

*the greatest element* $s^{f_z}$ *if and only if a permutation* $s_\pi^{f_y}$ *of* $s^{f_y}$ *exists so that*

$$\begin{array}{lll} & s_j^{f_1} \le s_{\pi_j}^{f_y} \le s_j^{f_z} & \text{if} \quad s_j^{f_1} \le s_j^{f_z} \\ \text{and} & s_j^{f_1} \ge s_{\pi_j}^{f_y} \ge s_j^{f_z} & \text{if} \quad s_j^{f_1} \ge s_j^{f_z}. \end{array}$$

***Proof.*** Obviously, the number $\frac{1}{2}\sum_j \mid s_j^{f_2} - s_j^{f_z} \mid +1$ corresponding to $s^{f_2}$, a direct successor of $s^{f_1}$, is at most one less than the number $\frac{1}{2}\sum_j \mid s_j^{f_1} - s_j^{f_r} \mid +1$ corresponding to $s^{f_1}$. Analogous to Lemma 4.1.1 we can suppose here that the components of $s^{f_2}$ are ordered monotonically increasing. Hence, consideration can be successively applied.

Thus, chains with the least element $s^{f_1}$ and the greatest element $s^{f_z}$ include at least $\frac{1}{2}\sum_j \mid s_j^{f_1} - s_j^{f_z} \mid +1$ partitions. Clearly, the main minimal chain includes this number of partitions. Therefore, the main minimal chain is a minimal chain and a) and the first statement of b) are proven.

Now, let $s^{f_y}$ satisfy the condition from Lemma 4.1.3b) and without loss of generality let $s_1^{f_y} \ge s_2^{f_y} \ge \cdots \ge s_n^{f_y}$. If we combine a minimal chain with the least element $s^{f_1}$ and the greatest element $s^{f_y}$ and a minimal chain with the least element $s^{f_y}$ and the greatest element $s^{f_z}$ then a chain with

$$\frac{1}{2}\sum_j \mid s_j^{f_1} - s_j^{f_y} \mid +1 + \frac{1}{2}\sum_j \mid s_j^{f_y} - s_j^{f_z} \mid = \frac{1}{2}\sum_j \mid s_j^{f_1} - s_j^{f_z} \mid +1 \qquad (*)$$

partitions follows which is thus a minimal chain with the least element $s^{f_1}$ and the greatest element $s^{f_z}$.

If $s^{f_y}$ does not satisfy the conditions from Lemma 4.1.3b), then simple computations with absolute values show that in place of the equals sign a $>$-sign is correct in (*). Thus no chain with the least element $s^{f_1}$, the greatest element $s^{f_z}$, the element $s^{f_y}$ and $\frac{1}{2}\sum_j \mid s^{f_1} - s^{f_z} \mid +1$ partitions exists.

■

Whether a partition is a predecessor or a successor of another can also be confirmed by the following Lemma in place of Definition 4.1.1.

**Lemma 4.1.4.** *(Muirhead 1903, see Section 5.D. in [26])*

*Let* $s^{f_1}, s^{f_2}$ *be different partitions where (without loss of generality)*

$s_1^l \geq s_2^l \geq \cdots \geq s_n^l$ *for* $l = f_1, f_2$.

*Then* $s^{f_1} \to s^{f_2}$ *is valid if and only if*

$$\sum_{j=1}^{\bar{n}} s_j^{f_1} \geq \sum_{j=1}^{\bar{n}} s_j^{f_2} \quad \textit{for} \quad \bar{n} = 1, \cdots, n$$

*(thus* $\sum_{j=1}^{\bar{n}} (s_j^{f_1} - s_j^{f_2}) \geq 0 \quad \textit{for} \quad \bar{n} = 1, \cdots, n$*).*

**Lemma 4.1.5.** *(See [26], Section 1.B.)*

*If the relation from Definition 4.1.1(c) is supplemented in such a way that each partition* [4] *bears the relation "→" to itself, then the relation "→" implies a partial order* [5] *on sets of restricted partitions (with symbols:* $(S_{n;su;k_0}, \to)$*).*

**Theorem 4.1.6.** *Partially ordered sets* $(S_{n;su;k_0}, \to)$ *are lattices.* [6] [7]

***Proof.***

(See, for instance, [36] for detailed explanations of the terms lattice, infimum and supremum.)

Let $s^{f_1}, \cdots, s^{f_z}$ be given, where $s_1^{f_l} \geq s_2^{f_l} \geq \cdots \geq s_n^{f_l}$ for $l = 1, \cdots, z$.

Obviously, $\underline{s}$ with

$\underline{s}_1 := \max\{s_1^{f_l} \mid l = 1, \cdots, z\}$ and

$$\underline{s}_j := \max\{\sum_{k=1}^{j} s_k^{f_l} \mid l = 1, \cdots, z\} - \sum_{k=1}^{j-1} \underline{s}_j \quad \text{for} \quad j = 2, \cdots, n$$

is a predecessor of $s^{f_l}$ for $l = 1, 2, \cdots, z$ (see Lemma 4.1.4).

Since any predecessor $s$ (where $s_1 \geq s_2 \geq \cdots \geq s_n$) of all partitions $s^{f_l}$ $(l = 1, \cdots, z)$ must satisfy the condition

$$\sum_{k=1}^{j} s_k \leq \max\{\sum_{k=1}^{j} s_k^{f_l} \mid l = 1, \cdots, z\} = \sum_{k=1}^{j} \underline{s}_j \quad \text{for } j = 1, \cdots, n$$

[4] A single partition is also a chain.

[5] This is the "ordering of dominance or majorization", see [26], Section 1.B.

[6] For the term lattice, see [36], for instance.

[7] In [9] Brylawski has shown such a statement for sets of partitions, whose parts are not restricted by a $k_0$.

the partition $\underline{s}$ is the infimum of $\{s^{f_1}, \cdots, s^{f_z}\}$.

Analogously, $\bar{s}$ with
$\bar{s}_1 := \min\{s_1^{f_l} \mid l = 1, \cdots, z\}$ and
$\bar{s}_j := \min\{\sum_{k=1}^{j} s_k^{f_l} \mid l = 1, \cdots, z\} - \sum_{k=1}^{j-1} \bar{s}_j \quad \text{for} \quad j = 2, \cdots, n.$
is the supremum of $\{s^{f_1}, \cdots, s^{f_z}\}$. ■

**Properties of the Lattices** $(\mathbf{S_{n;su;k_0}}, \rightarrow)$

a) First characteristic properties of these lattices can be found in Lemma 4.1.3.

b) The Jordan-Dedekind-Condition (see [36] or [14], Definition 2.1.12) is not valid for these lattices, in general (see Example 4.1.1).

c)

If the infimum $\underline{s}$ of two partitions $s^{f_1}$ and $s^{f_2}$ is a direct predecessor of $s^{f_1}$ and $s^{f_2}$, then the supremum of $s^{f_1}$ and $s^{f_2}$ is a direct successor of $s^{f_1}$ and $s^{f_2}$, and vice versa. (4.1.4)

***Proof*** of property c): Let the components of $\underline{s}$, $s^{f_1}$ and $s^{f_2}$ be ordered monotonically increasing (see also Lemma 4.1.1).

Since $\underline{s}$ is a direct predecessor of $s^{f_1}$ and $s^{f_2}$, indices $\alpha, \beta, \gamma, \delta$ exist so that $\underline{s}_\alpha > \underline{s}_\beta + 1$, $\underline{s}_\gamma > \underline{s}_\delta + 1$ and $s^{f_1} = \underline{s}[\beta; \alpha]$, $s^{f_2} = \underline{s}[\delta; \gamma]$.

In the case that the indices $\alpha, \beta, \gamma, \delta$ are pairwise different, the partition $\bar{s}$ with $\bar{s} = s^{f_1}[\delta; \gamma] = s^{f_2}[\beta; \alpha]$ is the supremum and obviously a direct successor of $s^{f_1}, s^{f_2}$.

The cases $\alpha = \gamma$ $(\beta \neq \delta)$ and $\beta = \delta$ $(\alpha \neq \gamma)$ are not possible, since in these cases $s^{f_1}$ or $s^{f_2}$ itself, however not $\underline{s}$, is the infimum of $s^{f_1}$ and $s^{f_2}$.

In the case $\beta = \gamma$ (analogously, $\delta = \alpha$) the partition $\bar{s}$ with $\bar{s} = \underline{s}[\delta; \alpha] = s^{f_1}[\delta; \beta] = s^{f_2}[\beta; \alpha]$ is the supremum and a direct succes-

sor of $s^{f_1}$ and $s^{f_2}$ (compare with the method of computation of $\bar{s}$ in the proof of Theorem 4.1.6).

If the supremum is a direct successor of $s^{f_1}$ and $s^{f_2}$, the statement that the infimum is a direct predecessor of these partitions can be proved analogously. ■

## 4.2 Definitions of Partitions-Requirements-Matrices and Initial Properties and Results

General PRMs are constructed on the basis of ordered restricted partitions of integers and PRMs (in the strict meaning) on the basis of unordered restricted partitions of integers.

PRMs are uniquely determined contrary to general PRMs.

*As noted in the introduction of Chapter 4, PRMs are matrices of transition probabilities of certain SDDP problems, which are modelled as MDPs (see Chapter 3).*

To ensure that the treatise of Chapter 4 can be, essentially, understood independent of Chapters 2 and 3, the definitions of PRMs from Chapter 3 are restated in Section 4.2.1 and Section 4.2.2.

In Section 4.2 initial simple statements are noted beside the definitions of PRMs.

In particular, PRMs for "equivalent sets of partitions" are considered in Section 4.2.3.

Let us use:

- $n,\ k_0,\ su,\ \tilde{S}_{n;su;k_0},\ S_{n;su;k_0}$ - as in Section 4.1
- $B_{n;k_0} := \{w \in \mathbb{Z}_+^n \mid 0 \leq w \leq k_0\}$, which is called the set of requirements.

- We assume that the requirements $w$ are random vectors [8] with a probability function $q$ (with $\sum\limits_{w \in B_{n;k_0}} q(w) = 1$).

  From Section 4.2.2 on we assume, in addition, that the requirements $w_i$, $(i = 1, \cdots, n)$ are independent and identically distributed for PRMs (in the strict meaning). This implies

$$q(w) = \prod_{i=1}^{n} q_0(w_i), \tag{4.2.1}$$

  where the marginal or "single" probabilities $q_0(w_i)$ are such that

$$q_0(w_i) > 0 \text{ for } w_i \in \{0, 1, \cdots, k_0\} \quad \text{and} \quad \sum_{j=0}^{k_0} q_0(j) = 1. \tag{4.2.2}$$

- Let $C[w, su]$ denote the case $\sum\limits_{i=1}^{n} w_i \leq su$ and $C[su, w]$ the case $\sum\limits_{i=1}^{n} w_i \geq su$. [9]

### 4.2.1 General Partitions-Requirements-Matrices

**Definition 4.2.1.** *Let $\tilde{s} \in \tilde{S}_{n;su;k_0}$ and $w \in B_{n;k_0}$.*
*Then,*

$$\hat{A}_{n;su;k_0}(\tilde{s}, w) = \left\{ \tilde{s}' \in \tilde{S}_{n;su;k_0} \;\middle|\; \begin{array}{l} w_i \leq \tilde{s}'_i \leq max\{\tilde{s}_i, w_i\},\ i = 1, \cdots, n,\ in\, C[w, su] \\ min\{\tilde{s}_i, w_i\} \leq \tilde{s}'_i \leq w_i,\ i = 1, \cdots, n,\ in\, C[su, w] \end{array} \right\} \tag{4.2.3}$$

*is called the set of* **feasible (ordered) partitions with respect to $\tilde{s}$ and** *$w$ and $\tilde{s}' = \tilde{s}'(\tilde{s}, w) \in \hat{A}_{n;su;k_0}(\tilde{s}, w)$ are feasible (ordered) partitions with*

[8] We use the same notation for random vectors and their realizations.

[9] In this way it is not necessary to distinguish the case $\sum\limits_{i=1}^{n} w_i = su$.

*Regarding the meaning of the cases for SDDP problems refer to the explanations in (3.2.1) and (3.2.2).*

*respect to $\tilde{s}$ and $w$.*

*Furthermore $\tilde{s}' = \tilde{s}'(\tilde{s}, w)$ is also called a feasible transition from $\tilde{s}$ to $\tilde{s}'$ for $w$.*

*((4.2.3) corresponds to (3.2.11) from Section 3.2. Regarding SDDP problems refer also to the explanation above (3.2.11).)*

The elements $p_{fl}$ of general PRMs are now defined as the sum of the probabilities of the $w$, for which feasible transitions from $\tilde{s}^f$ to $\tilde{s}^l$ are given:

**Definition 4.2.2.** *Let $\tilde{S}_{n;su;k_0} = \{\tilde{s}^1, \tilde{s}^2, \cdots, \tilde{s}^{\tilde{r}}\}$, $B_{n;k_0}$ and feasible partitions $\tilde{s}'(\tilde{s}, w) \in \hat{A}_{n;su;k_0}(\tilde{s}, w)$ with respect to every $f = 1, 2, ..., \tilde{r}$ and $w \in B_{n;k_0}$ be given.*

*$P = P_{n;su;k_0} = (p_{fl})_{\substack{f=1,\ldots,\tilde{r} \\ l=1,\ldots,\tilde{r}}}$ with elements*

$$p_{fl} = p(\tilde{s}^l | \tilde{s}^f) = \sum_{w:\tilde{s}^l = \tilde{s}'(\tilde{s}^f, w)} q(w) \text{ for } f = 1, \cdots, \tilde{r},\ l = 1, \cdots, \tilde{r} \qquad (4.2.4)$$

*is called a* **general partitions-requirements-matrix**.

**Computation of Maximal Values of Elements of General PRMs**

Computation of maximal values (max $p_{fl}$) of $p_{fl}$ is simple. (Clearly, not all elements of general PRMs can be simultaneously equal to their maximal values.)

For the computation of $\max p_{fl}$, let ordered partitions $\tilde{s}^f \in \tilde{S}_{n;su;k_0}$ and $\tilde{s}^l \in \tilde{S}_{n;su;k_0}$ be given.

In the case $C[w, su]$

$w_i \in \{0, 1, \cdots, \tilde{s}_i^l\}$ if $\tilde{s}_i^f \geq \tilde{s}_i^l$ and

$w_i = \tilde{s}_i^l$ if $\tilde{s}_i^f < \tilde{s}_i^l$

follows from (4.2.3) for requirements $w$, which satisfy (4.2.3).

The set of requirements $w$, which fulfil these relations is denoted by $B_1$.

In case $C[su, w]$

$$w_i \in \{\tilde{s}_i^l, \tilde{s}_i^l + 1, \cdots, k_{0_i}\} \quad \text{if} \quad \tilde{s}_i^f \leq \tilde{s}_i^l \quad \text{and}$$
$$w_i = \tilde{s}_i^l \quad \text{if} \quad \tilde{s}_i^f > \tilde{s}_i^l$$

follows from (4.2.3) and the corresponding set of requirements is denoted by $B_2$.

From this

$$\max p_{fl} = \sum_{w \in B_1} q(w) + \sum_{w \in B_2} q(w) - q(\tilde{s}^l)$$

follows if we note that $w = \tilde{s}^l$ satisfies (4.2.3) in case $C[w, su]$ and in case $C[su, w]$.

### 4.2.2 Partitions-Requirements-Matrices

PRMs (in the strict meaning) are constructed on the basis of lattices of unordered restricted partitions of integers and feasible partitions with least square sums of their parts (later referred to as "feasible balanced partitions").

PRMs corresponding to given lattices of partitions are uniquely determined.

Following the definition of feasible balanced partitions an iterative method of their computation will be given. From this it will be clear that PRMs corresponding to given lattices of partitions are uniquely determined.

At the end of this section, difficulties with effective computations of PRMs will be mentioned in an example.

*Relationships to Chapter 3: PRMs (in the strict meaning) are matrices of transition probabilities for reduced SDDP problems if (3.4.1) and (3.4.3) are supposed (see Section 3.4.2) and for decisions for feasible states (unordered restricted partitions of integers) with minimum average one-step reward functions. These decisions are identical to transitions into feasible partitions with least square sums of their parts (later referred to as feasible balanced partitions), which is shown in the following Lemma 4.2.2. See also Lemma 4.2.3.*

**Definition 4.2.3.** *Let $s \in S_{n;su;k_0}$ with (without loss of generality) $s_1 \geq s_2 \geq \cdots \geq s_n$ and $w \in B_{n;k_0}$ be given.*

*(a) Then*

$$\hat{A}_{n;su;k_0}(s,w) = \left\{ s' \in S_{n;su;k_0} \;\middle|\; \begin{array}{l} \exists\, s'_\pi \text{ permutation of } s' : \\ w_i \leq s'_{\pi i} \leq max\{s_i, w_i\},\ i = 1, \cdots, n,\ in\, C[w, su], \\ \\ min\{s_i, w_i\} \leq s'_{\pi i} \leq w_i,\ i = 1, \cdots, n,\ in\, C[su, w] \end{array} \right\} \tag{4.2.5}$$

[10]

*is called the set of* **feasible (unordered) partitions with respect to** $s$ **and** $w$ *and $s' = s'(s,w) \in \hat{A}_{n;su;k_0}(s,w)$ are feasible (unordered) partitions with respect to $s$ and $w$.*

*Furthermore, $s' = s'(s,w)$ is also called a feasible transition from $s$ to $s'$ for $w$.*

*(b) A feasible partition $s^* = s^*(s,w) \in S_{n;su;k_0}$ with respect to $s$ and $w$ is called the* **feasible balanced partition with respect to** $s$ **and** $w$ *if $s^*$ is an optimal solution of the problem*

$$\sum_{i=1}^{n} (s_i^*)^2 \to min$$

*subject to*

$$s^* \in \hat{A}_{n;su;k_0}(s,w)$$

*and $s^* = s^*(s,w)$ is also called a feasible balanced transition from $s$ to $s^*$ for $w$.*

*In addition, $B^*_{n,k_0}(s,s^*) = \{w \in B_{n,k_0} \mid s^* = s^*(s,w)\}$ is the set of balancing requirements.*

**An Iterative Method for the Computation of Feasible Balanced Partitions - Enumeration**

Feasible balanced partitions for given $s \in S_{n;su;k_0}$ and $w \in B_{n,k_0}$ can be

[10] *(4.2.5) corresponds to (3.4.15) from Section 3.4.*

computed by the following iterative method.[11]

Case $C[w, su]$ :

Set $s^{*'} = s$.

If (4.2.5) is satisfied, then $s^{*'}$ is the desired partition (end) else: (*)

Determine a component $s_j^{*'}$ of $s^{*'}$ with $s_j^{*'} = max\{s_i^{*'} \mid s_i^{*'} > w_i\}$ and a component $s_y^{*'}$ with $s_y^{*'} < w_y$.
Set $s_j^{*'} = s_j^{*'} - 1$ and $s_y^{*'} = s_y^{*'} + 1$.

Go to (*).

Case $C[su, w]$ :

Set $s^{*'} = s$.

If (4.2.5) is satisfied, then $s^{*'}$ is the desired partition (end) else: (**)

Determine a component $s_j^{*'}$ of $s^{*'}$ with $s_j^{*'} = min\{s_i^{*'} \mid s_i^{*'} < w_i\}$ and a component $s_y^{*'}$ with $s_y^{*'} > w_y$.
Set $s_j^{*'} = s_j^{*'} + 1$ and $s_y^{*'} = s_y^{*'} - 1$.

Go to (**).

(For the proof see [18].)

**Lemma 4.2.1.** *Let $S_{n;su;k_0} = \{s^1, s^2, \cdots, s^r\}$ and $B_{n;k_0}$ be given.*

*(a) The feasible balanced partition $s^* = s^*(s, w)$ with respect to $s$ and $w$ is uniquely determined.*

*(b) If $\bar{s} \in \hat{A}_{n;su;k_0}(s, w)$ (feasible partition), then the feasible balanced partition $s^* = s^*(s, w)$ is a successor of $\bar{s}$.*

[11] This section is taken from [18], Section 2.3.

***Proof.***

(a) follows from the iterative method for the considered unordered partitions.

(b) We consider the case $C[w, su]$. (In the other case the proof is analogous.)

According to (4.2.5) a permutation $\bar{s}_\pi$ of $\bar{s}$ exists so that

$$s_i \geq \bar{s}_{\pi i} \geq w_i \text{ for } i \text{ with } s_i \geq w_i. \tag{$*1$}$$

$s^*$ is constructed step-by-step over $s = s^{*1}, s^{*2}, \cdots, s^{*l} = s^*$ by the iterative method (above).

It remains to be shown that $s^*$ is a successor of $\bar{s}$.

For this purpose unordered partitions $\bar{s}^f \in S_{n;su;k_0}$ are to be determined for each $s^{*f}$ $(f = 1, \cdots, l)$ so that $\bar{s}^f = \bar{s}^{f-1}$ or $\bar{s}^f$ is a direct successor of $\bar{s}^{f-1}$ (see Definition 4.1.1) and $\bar{s}^l = s^*$ in the following way:

Without loss of generality, let $\bar{s}_i = \bar{s}_{\pi i}$, for $i = 1, \cdots, n$.

$\bar{s}^1 := \bar{s}$,

$\bar{s}^f$ with

$$\bar{s}_i^f = \begin{cases} \bar{s}_{i_0}^{f-1} - 1 & \text{if } s_{i_0}^{*f} < s_{i_0}^{*(f-1)} \text{ and } s_{i_0}^{*(f-1)} = \bar{s}_{i_0}^{f-1} \text{ for } i = i_0, \\ \bar{s}_{i_1}^{f-1} + 1 & \text{for } i = i_1 \text{ with } \bar{s}_{i_1}^{f-1} = \min\{\bar{s}_i^{f-1} \mid \bar{s}_i^{f-1} < s_i^{*(f-1)}\} \\ & \text{if an } i_0 \text{ exists which satisfies the above condition,} \\ \bar{s}_i^{f-1} & \text{otherwise} \end{cases} \tag{$*2$}$$

for $f = 2, \cdots, l$.

Let $i$ be given with $s_i \geq w_i$, then $s_i^{*1} = s_i \geq \bar{s}_i = \bar{s}_i^1$ according to $(*1)$ and $(*2)$.

If, in addition, $i = i_0$ satisfies the first condition from $(*2)$ then

$$s_{i_0}^{*f} \geq \bar{s}_{i_0}^f \text{ for } i_0 \text{ with } s_{i_0} \geq w_{i_0} \tag{$*3$}$$

follows successively from $(*2)$.

$$\sum_{i:s_i > w_i} \bar{s}_i^f = su - \sum_{i:s_i \leq w_i} w_i < \sum_{i:s_i > w_i} s_i^{*f} (\leq \sum_{i:s_i > w_i} s_i)$$

is then valid for any $f \in \{1, \cdots, l-1\}$ because of $\bar{s}^1 = \bar{s}$, $(*2)$ and the iterative method.

Thus $i_1$, which satisfies the desired condition in $(*2)$, in fact exists.

Since $s_{i_0}^{*(f-1)} = \max\{s_j^{*(f-1)} \mid s_j^{*(f-1)} > w_j\}$ for $s_{i_0}^{*(f-1)} > s_{i_0}^{*f}$ according to the iterative method

$$s_{i_1}^{*(f-1)} \leq s_{i_0}^{*(f-1)}$$

follows and furthermore

$$\bar{s}_{i_1}^{f-1} < s_{i_1}^{*(f-1)} \leq s_{i_0}^{*(f-1)} = \bar{s}_{i_0}^{f-1}$$

(this means $(\bar{s}_{i_1}^{f-1})^2 + (\bar{s}_{i_0}^{f-1})^2 \geq (\bar{s}_{i_1}^{f})^2 + (\bar{s}_{i_0}^{f})^2$). Consequently $\bar{s}^f$ is a successor of $\bar{s}^{f-1}$ or $\bar{s}^f = \bar{s}^{f-1}$ (if $s_{i_1}^{*(f-1)} = s_{i_0}^{*(f-1)}$, $\bar{s}_{i_1}^{f-1} + 1 = \bar{s}_{i_0}^{f-1}$ or $\bar{s}_{i_1}^{f} = \bar{s}_{i_0}^{f-1}$).

Finally, $\bar{s}^l = s^*$ follows from $s^{*l} = s^*$ and $(*3)$. ■

*In the following Lemma connections with the average one-step reward functions and the decision function $d^*$ of the reduced SDDP problem from Section 3.4 are shown.*

**Lemma 4.2.2.** *Let $S_{n;su;k_0} = \{s^1, s^2, \cdots, s^r\}$ and $B_{n;k_0}$ be given and let the requirements $w_i$, $(i = 1, \cdots, n)$ be independent and identically distributed, where (4.2.2) is additionally assumed.*

*(a) If $s^l \in S_{n;su;k_0}$ is a successor of $s^f \in S_{n;su;k_0}$ $(s^f \neq s^l)$, then the inequality $\gamma(s^f) > \gamma(s^l)$ is valid for the average one-step reward functions (which can be computed according to the formula from Theorem 3.4.1(c)).*

*(b) If $\bar{s} \in \hat{A}_{n;su;k_0}(s, w)$ (feasible partition) then $\gamma(\bar{s}) > \gamma(s^*)$ follows for the feasible balanced partition $s^* = s^*(s, w)$ with respect to $s$ and $w$ if $\bar{s} \neq s^*$.*

*Thus $d^*$ (see (3.4.16)) is identical to decisions for feasible balanced partitions.*

***Proof.***

(a) We initially show the conjecture for a direct successor $s^l$ of $s^f$ from which (a) then follows.

See Lemma 4.1.1:

Let $s^l$ be such that $s_i^l[i_1;i_2] = \begin{cases} s_i^f + 1 & \text{for} \quad i = i_1, \\ s_i^f - 1 & \text{for} \quad i = i_2, \\ s_i^f & \text{otherwise,} \end{cases}$

where $s_{i_2}^f > s_{i_1}^f + 1$.

Let $s \in S_{n;su;k_0}$. The corresponding average one-step reward function $\gamma(s)$ can be computed according to Theorem 3.4.1(c)) for independent and identically distributed requirements:

$$\gamma(s) = \sum_{i=1}^{n} \sum_{w_i=0}^{s_i} (s_i - w_i)\, q_0(w_i) + R(n, su, k_0, q).$$

And

$$\begin{aligned}
\gamma(s^f) - \gamma(s^l) &= \sum_{w_{i_2}=0}^{s_{i_2}^f} (s_{i_2}^f - w_{i_2})\, q_0(w_{i_2}) + \sum_{w_{i_1}=0}^{s_{i_1}^f} (s_{i_1}^f - w_{i_1})\, q_0(w_{i_1}) \\
&\quad - \left( \sum_{w_{i_2}=0}^{s_{i_2}^f - 1} (s_{i_2}^f - 1 - w_{i_2})\, q_0(w_{i_2}) + \sum_{w_{i_1}=0}^{s_{i_1}^f + 1} (s_{i_1}^f + 1 - w_{i_1})\, q_0(w_{i_1}) \right) \\
&= \sum_{w_{i_2}=0}^{s_{i_2}^f - 1} q_0(w_{i_2}) - \sum_{w_{i_1}=0}^{s_{i_1}^f} q_0(w_{i_1}) \\
&= \sum_{w=s_{i_1}^f + 1}^{s_{i_2}^f - 1} q_0(w) > 0
\end{aligned}$$

follow.

(b) is yielded by Lemma 4.2.1(b) and (a) of this Lemma.

■

Now, the elements of PRMs $p_{fl}^*$ are defined as sum of the probabilities of the requirements $w$, for which feasible balanced transitions from $s^f$ to $s^l$ are given:

**Definition 4.2.4.** *Let* $S_{n;su;k_0} = \{s^1, s^2, \cdots, s^r\}$ *and* $B_{n;k_0}$ *be given and let the requirements* $w_i$, $(i = 1, \cdots, n)$ *be independent and identically distributed, where (4.2.2) is additionally assumed.*

$P^* = P^*_{n;su;k_0} = (p_{fl}^*)$ *with elements*

$$p^*_{fl} = p^*(s^l|s^f) = \sum_{w:s^l=s^*(s^f,w)} q(w) \text{ for } f = 1, \cdots, r,\ l = 1, \cdots, r \tag{4.2.6}$$

*is called the* **partitions-requirements-matrix**.

Obviously, since $p^*_{fl} > q(w = s^l)$,

$$p^*_{fl} > 0. \tag{4.2.7}$$

*Relationships to reduced SDDP problems:*

**Lemma 4.2.3.** *PRMs are matrices of transition probabilities of reduced SDDP problems for decisions* $d^*$ *(see (3.4.16) in Section 3.4.2).*

***Proof.*** PRMs are matrices of transition probabilities of reduced SDDP problems, since sets of feasible partitions (see (4.2.5)) and DA decisions sets of reduced SDDP problems (see (3.4.15)) are the same, and also due to Definition 4.2.4 (and to (2.3.7)). That PRMs are matrices of transition probabilities for decisions $d^*$ follows from (3.4.16), Lemma 4.2.2(b) and Definition 4.2.3(b). ■

**Example 4.2.1.** *Let* $n = 3, k_0 = 3, su = 6$ *and* $q_0(w_i) = \frac{1}{1+k_0} = \frac{1}{4}$ *for* $w_i = 0, 1, 2, 3$ *be given (hence* $q(w) = \left(\frac{1}{4}\right)^3 = \frac{1}{64}$ *for all* $w \in B_{3;3}$*).*

*Then the set* $S_{3;6;3}$ *includes the elements*

$$s^1 = \begin{pmatrix} 3 \\ 3 \\ 0 \end{pmatrix},\ s^2 = \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix},\ s^3 = \begin{pmatrix} 2 \\ 2 \\ 2 \end{pmatrix}.$$

*Feasible balanced partitions are, for instance,*

$$s^*\left(s^1 = \begin{pmatrix} 3 \\ 3 \\ 0 \end{pmatrix}, w = \begin{pmatrix} 0 \\ 0 \\ 2 \end{pmatrix}\right) = \begin{pmatrix} 2 \\ 2 \\ 2 \end{pmatrix} \left(= s^3\right),$$

$$s^*\left(s^1 = \begin{pmatrix} 3 \\ 3 \\ 0 \end{pmatrix}, w = \begin{pmatrix} 3 \\ 2 \\ 3 \end{pmatrix}\right) = \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix} \left(= s^2\right).$$

*However, in order to compute the partitions-requirements-matrix we need to know all of the* $r \cdot (k_0+1)^n$ *feasible balanced partitions (with respect to all* $s \in S_{n;su;k_0}, w \in B_{n;k_0}$*).*

*For example,* $p^*_{13}$ *can be computed by means of* $s^3 = s^*(s^1, w)$ *for*

$$w \in B^*_{3;3}(s^1, s^3) = \left\{ \begin{pmatrix} 0 \\ 0 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \\ 2 \end{pmatrix}, \begin{pmatrix} 0 \\ 2 \\ 2 \end{pmatrix}, \right.$$

$$\left. \begin{pmatrix} 2 \\ 2 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 2 \\ 2 \\ 3 \end{pmatrix} \right\}$$

*and we obtain*

$$p^*_{13} = \sum_{w \in B^*_{3;3}(s^1, s^3)} q(w) = \tfrac{10}{64}.$$

*Further enumerations yield*

$$P^* = \tfrac{1}{64} \begin{pmatrix} 21 & 33 & 10 \\ 3 & 49 & 12 \\ 3 & 27 & 34 \end{pmatrix}.$$

### 4.2.3 Partitions-Requirements-Matrices for Equivalent Lattices of Partitions

In this section relationships between certain lattices of partitions are given and equalities of corresponding PRMs are shown if the probability functions of the requirements are suitable.

**Definition 4.2.5.** *Let $S_{n;su;k_0}$, $S_{n;\overline{su};\overline{k}_0}$ and $k_c \in \mathbb{Z}_+$ be given such that $k_c \geq k_0$, $k_c \geq \overline{k}_0$ and $\overline{su} = n\, k_c - su$.*

*(a) $s \in S_{n;su;k_0}$ and $\bar{s} \in S_{n;\overline{su};\overline{k}_0}$ are called complementary partitions if permutations $s_\pi$ of $s$ and $\bar{s_\pi}$ of $\bar{s}$ exist with $s_\pi + \bar{s_\pi} = (k_c, ..., k_c)^T$.*

*(b) If $\mid S_{n;su;k_0} \mid \; = \; \mid S_{n;\overline{su};\overline{k}_0} \mid$, then $S_{n;su;k_0}$ and $S_{n;\overline{su};\overline{k}_0}$ are called equivalent (with respect to the partial order) (see also the following Lemma 4.2.4).*

*(c) If $k_c = k_0 = \overline{k}_0$, (then (b) is fulfilled, see the following Lemma 4.2.4(ii)) and $S_{n;su;k_0}$ and $S_{n;\overline{su};\overline{k}_0}$ are called equivalent with regard to the Poisson equation.* [12]

**Lemma 4.2.4.** *Let $S_{n;su;k_0}$, $S_{n;\overline{su};\overline{k}_0}$ and $k_c \in \mathbb{Z}_+$ be given such that $k_c \geq k_0$, $k_c \geq \overline{k}_0$ and $\overline{su} = n\, k_c - su$.*

*(i) If $\mid S_{n;su;k_0} \mid \; = \; \mid S_{n;\overline{su};\overline{k}_0} \mid$, then a one-to-one correspondence of the partitions of $S_{n;su;k_0}$ and $S_{n;\overline{su},\overline{k}_0}$ is yielded by $s_\pi + \bar{s_\pi} = (k_c, ..., k_c)^T$ for permutations $s_\pi$ of $s \in S_{n;su;k_0}$ and $\bar{s_\pi}$ of $\bar{s} \in S_{n;\overline{su};\overline{k}_0}$.*

*$s^f \in S_{n;su;k_0}$ is a direct predecessor of $s^l \in S_{n;su;k_0}$ if and only if $\bar{s^f} = (k_c, ..., k_c)^T - s^f \;\; (\in S_{n;\overline{su};\overline{k}_0})$ is a direct predecessor of $\bar{s^l} = (k_c, ..., k_c)^T - s^l \;\; (\in S_{n;\overline{su};\overline{k}_0})$.*

*(ii) If $k_c = k_0 = \overline{k}_0$ then $\mid S_{n;su;k_0} \mid \; = \; \mid S_{n;\overline{su};\overline{k}_0} \mid$ follows and (i) is valid.* [13]

***Proof.***

(i) The one-to-one correspondence of the partitions is obvious.

The relationship concerning the direct predecessors follows from the one-to-one correspondence $s_\pi + \bar{s_\pi} = (k_c, ..., k_c)^T$, (4.1.1) and Lemma 4.1.1.

[12] This denotation follows from Lemma 4.6.2 in the subsequent Section 4.6.2.1.

[13] $\mid S_{n;su;k_0} \mid \; = \; \mid S_{n;\overline{su};\overline{k}_0} \mid$ can also be found in Andrews [3], Theorem 3.10, page 47. The corresponding proof is time consuming.

(ii) Clearly, $\bar{s} = (k_0, ..., k_0)^T - s$ implies a one-to-one correspondence between all elements of the set $S_{n;su;k_0}$ and all elements of the set $S_{n;\overline{su};k_0}$, from which the statement directly follows.

■

**Theorem 4.2.5.** *Let $S_{n;su;k_0}$, $S_{n;\overline{su};k_0}$ with $\overline{su} = n\ k_0 - su$ and $B_{n;k_0}$ be given and let the requirements $w_i$, $(i = 1, \cdots, n)$ be independent and identically distributed where, (4.2.2) is additionally assumed. Furthermore, let*

$$q_0(w_i) = q_0(k_0 - w_i) \text{ for } i = 1, 2, ..., n,\ w \in B_{n;k_0} \tag{4.2.8}$$

*also be valid. Then, the corresponding partitions-requirements-matrices $P^*_{n;su;k_0}$ and $P^*_{n;\overline{su};k_0}$ are equal.*

***Proof.*** Initially we show:

$s \in S_{n;su;k_0}$ satisfies (4.2.5) with respect to a given $s' \in S_{n;su;k_0}$ and $w \in B_{n;k_0}$ if and only if $\bar{s} \in S_{n;\bar{s}u;k_0}$ satisfies (4.2.5) with respect to $\bar{s}' \in S_{n;\bar{s}u;k_0}$ and $\bar{w} \in B_{n;k_0}$ where $s$ and $\bar{s}$ are complementary partitions as well as $s'$ and $\bar{s}'$ and $w + \bar{w} = (k_0, ..., k_0)^T$:

Obviously, if $C[w, su]$ is present for $w \in B_{n;k_0}$, then $C[\bar{s}u, \bar{w}]$ is present for $\bar{w} = (k_0, ..., k_0)^T - w$ (and vice versa).

In the case $C[w, su]$

$w_i \leq s_{\pi_i} \leq \max\{s'_i, w_i\}$ (see (4.2.5)) implies

$\min\{\bar{s}'_i, \bar{w}_i\} = \min\{k_0 - s'_i, k_0 - w_i\} \leq \bar{s}_{\pi_i} = k_0 - s'_{\pi_i} \leq k_0 - w_i = \bar{w}_i$ (and vice versa).

Case $C[su, w]$ is handled analogously.

Simple computations with respect to the square sums of components of $s$ and $\bar{s}$ then yield: $s$ is the feasible balanced partition with respect to $s'$ and $w$ if and only if $\bar{s}$ is the feasible balanced partition with respect to $s$ and $w$ (refer to Definition 4.2.3).

That the PRMs $P^*_{n;su;k_0}$ and $P^*_{n;\overline{su};k_0}$ are equal then follows from Definition 4.2.4 and (4.2.8). ■

**Example 4.2.2.** .

*a)* $S_{3;4;3}$ *and* $S_{3;5;3}$ *are equivalent with regard to the Poisson equation (thereby* $k_c = k_0 = \bar{k}_0 = 3$*).*

*b)* $S_{3;4;4}$ *and* $S_{3;17;7}$ *are equivalent (with respect to the partial order) (thereby* $k_c = 7$*).*

## 4.3 The Computation of PRMs by means of Permutations of Perturbed Partitions

Based on the definition of "perturbed partitions" elements of PRMs can be computed more effectively than by enumeration.

If we use the iterative method from Section 4.2.2 for the computation of feasible balanced partitions with respect to a given $s^f \in S_{n;su;k_0}$ and $w \in B_{n;k_0}$, then the complete row $f$ of the corresponding PRM is computed by these enumerations.

In contrast single elements of PRMs can be computed by means of the method of "perturbed partitions". However, permutations with certain characteristics must additionally be determined.

Perturbed partitions are otherwise used in Section 4.5 in order to show a polynomial and sometimes an exponential dependence of the elements of PRMs on the variables $n$ and $k_0$ in the case of discrete uniformly distributed requirements (and similar relationships for other distributions).

This section can be skipped by readers, which are interested only in Sections 4.4 and 4.6.

If a single element $p^*_{fl}$ of a PRM is to be computed according to Definition 4.2.4 as sum

$$p^*_{fl} = \sum_{w:s^l=s^*(s^f,w)} q(w)$$

then a substantial difficulty is the fact that a $w$ together with various permutations $s^l_\pi$ of $s^l$ can satisfy the inequalities (4.2.5):

If

$$s^f = \begin{pmatrix} 6 \\ 4 \\ 2 \\ 2 \end{pmatrix} \text{ and } s^l = \begin{pmatrix} 5 \\ 4 \\ 3 \\ 2 \end{pmatrix}$$

are given, for instance, then the inequalities in (4.2.5) are satisfied by

$$w = \begin{pmatrix} 3 \\ 2 \\ 5 \\ 5 \end{pmatrix} \text{ and the permutations } s_\pi^{l_{(1)}} = \begin{pmatrix} 3 \\ 2 \\ 5 \\ 4 \end{pmatrix} \text{ or } s_\pi^{l_{(2)}} = \begin{pmatrix} 3 \\ 2 \\ 4 \\ 5 \end{pmatrix},$$

for example (and $s^l = s^*(s^f, w)$ is of course valid).

If, now, $w$ are determined for each permutation $s_\pi^l$ of $s^l$ so that (4.2.5) is fulfilled, then one $w$ could satisfy (4.2.5) together with different permutations of $s^l$ (see above). However, the corresponding $q(w)$ may only include in the sum $\sum\limits_{w:s^l=s^*(s^f,w)} q(w)$ one time in order to compute $p_{fl}^*$!

Instead of the permutations $s_\pi^{l_{(1)}}$ and $s_\pi^{l_{(2)}}$ we will use the "perturbed partition".

$$\begin{pmatrix} 3 \\ 2 \\ \mathbf{4} \\ \mathbf{4} \end{pmatrix}$$

in this section (see the following Definitions 4.3.3 and 4.3.7).

The reversed situation considered:

If $s^*(s^f, w)$ is computed by the iterative method from Section 4.2.2 for given $s^f$ and $w$, then different permutations of $s^*(s^f, w)$ may be possible in the last iterations if $i_0$ with $s_{i_0}^{*'} = max\{s_i^{*'} \mid s_i^{*'} > w_i\}$ is not unique in case C[w, su] or $i_0$ with $s_{i_0}^{*'} = min\{s_i^{*'} \mid s_i^{*'} < w_i\}$ is not unique in case C[su,w].

In addition, we use the following **terminology**:

Let us assume in this section that (without loss of generality) the components of the partitions $s^f \in S_{n;su;k_0}$ and $s^l \in S_{n;su;k_0}$ are initially ordered

monotonically decreasing:

$$s_1^f \geq s_2^f \geq \cdots \geq s_n^f \quad \text{and} \quad s_1^l \geq s_2^l \geq \cdots \geq s_n^l$$

and furthermore we use the notation and symbols:

$$\begin{aligned} &\text{F: the number of components of } s^f \text{ which are not equal to 0,} \\ &\text{L: the number of components of } s^l \text{ which are not equal to 0.} \end{aligned} \tag{4.3.1}$$

$$\begin{aligned} \mathbf{s^f}: \ & s_1^f = \cdots = s_{F_1}^f > s_{F_1+1}^f = \cdots = s_{F_2}^f > \cdots > s_{F_{z-1}+1}^f = \cdots = s_{F_z}^f > 0 \\ & \qquad \left(s_i^f = 0 \text{ for } i \geq F_z + 1 \text{ if } F_z < n\right) \\ & (\text{with } \ F_1 < F_2 < \cdots < F_z = F(< F_{z+1} = n \ for \ F_z < n)), \end{aligned} \tag{4.3.2}$$

$$\begin{aligned} \mathbf{s^l}: \ & s_1^l = \cdots = s_{L_1}^l > s_{L_1+1}^l = \cdots = s_{L_2}^l > \cdots \\ & > \boldsymbol{s^l_{L_{J_o-2}+1}} = \cdots = \boldsymbol{s^l_{L_{J_o-1}}} > \boldsymbol{s^l_{L_{J_o-1}+1}} = \cdots = \boldsymbol{s^l_{L_{J_o}}} > \boldsymbol{s^l_{L_{J_o}+1}} = \\ & \qquad \cdots = \boldsymbol{s^l_{L_{J_o}+1}} > \cdots > s_{L_{y-1}+1}^l = \cdots = s_{L_y}^l > 0 \\ & \qquad (0 = s_{L_y+1}^l = \cdots = s_{L_{y+1}}^l = s_n^l \text{ if } L_y < n) \end{aligned} \tag{4.3.3}$$

(with $L_1 < L_2 < \cdots < L_y = L(< L_{y+1} = n \ for \ L_y < n)$, furthermore $L_0 := 0$).

Moreover, we define

$$\begin{aligned} & \sigma_J^l := s_{L_J}^l \ \text{ for } \ J = 1, 2, \ldots, y \ (or \ y+1 \ for \ L_y < n). \\ & \sigma_1^l > \sigma_2^l > \cdots > \boldsymbol{\sigma^l_{J_o-1}} > \boldsymbol{\sigma^l_{J_o}} > \boldsymbol{\sigma^l_{J_o+1}} > \cdots > \sigma_y^l \ (> \sigma_{y+1}^l = 0 \\ & \qquad \text{for } L_y < n) \end{aligned} \tag{4.3.4}$$

follows.

$$\delta(s^f, s^l) = \delta_{fl} := \begin{cases} 1 & \text{if } s^f = s^l, \\ 0 & \text{if } s^f \neq s^l. \end{cases} \tag{4.3.5}$$

We now compute the requirements $w$ with $s^*(s^f, w) = s^l$ in the cases C[su, w] and C[w, su] by means of sets of perturbed partitions.

The elements $p^*_{fl}$ of the PRMs are then calculated in the following way

$$p^*_{fl} = p^{*2}_{fl} + p^{*1}_{fl} - p^{*1,2}_{fl}$$

$$\text{with } p^{*2}_{fl} = \sum_{C[su,w],w:s^l=s^*(s^f,w)} q(w),\ p^{*1}_{fl} = \sum_{C[w,su],w:s^l=s^*(s^f,w)} q(w)$$

$$\text{and } p^{*1,2}_{fl} = \sum_{C[su,w]\bigcap C[su,w],w:s^l=s^*(s^f,w)} q(w) = \sum_{s^l_\pi:\ permutation\ of\ s^l} q(s^l_\pi). \tag{4.3.6}$$

**Case C[su, w]** *(the requirements cannot be completely fulfilled)*:
Let a partition $s^f \in S_{n;su;k_0}$ and a permutation $s^l_\pi$ of a partition $s^l \in S_{n;su;k_0}$ be given. We then compare the components of $s^f$ with the components of $s^l_\pi$ in order of decreasing $s^l_{\pi_i}$.

**Definition 4.3.1.** *Let $J_o \in \{1, 2, \ldots, y\}$, $j_o \in \{1, 2, \ldots, L_{J_o} - L_{J_o-1}\}$.*
*If*

$$s^f_i \geq s^l_{\pi_i} \quad \textit{for any } s^l_{\pi_i} \geq \sigma^l_{J_o-1}, \tag{d1}$$

$$s^f_i \geq s^l_{\pi_i} \quad \textit{for } L_{J_o} - L_{J_o-1} - j_o \textit{ of the } s^l_{\pi_i} = \sigma^l_{J_o}, \tag{d2}$$

$$\textit{and } s^f_i < s^l_{\pi_i} \quad \textit{for } j_o \textit{ of the } s^l_{\pi_i} = \sigma^l_{J_o}, \tag{d3}$$

*then we refer to a* $(\mathbf{J_o}, \mathbf{j_o})$**-perturbation of the relation "$\geq$" between $\mathbf{s^f}$ and $\mathbf{s^l_\pi}$.**

Formally, we define the $(J_o, j_o)$-perturbed partition $\hat{s}^l$ of $s^l$.

Contrary to $s^l$, the $(J_o, j_o)$-perturbed partition $\hat{s}^l$ has exactly $j_o$ components which are reduced from $\sigma^l_{J_o}$ by 1 to $\sigma^l_{J_o} - 1$:

**Definition 4.3.2.** *Let $J_o$, $j_o$ be given with $J_o \in \{1, 2, \ldots, y\}$, $j_o \in \{1, 2, \ldots, L_{J_o} - L_{J_o-1}\}$.*

$$\hat{s}^l : \begin{cases} \hat{s}^l_j = s^l_j & \textit{for} \quad j \in \{1, 2, \cdots, L_{J_o} - j_o\} \\ & \textit{and for} \quad j \in \{L_{J_o} + 1, \cdots, n\} \quad (d4) \\ \hat{s}^l_j = s^l_j - 1 (= \sigma^l_{J_o} - 1) & \textit{for} \quad j \in \{L_{J_o} - j_o + 1, \cdots, L_{J_o}\} \quad (d5) \end{cases}$$

*is called the* $(\mathbf{J_o}, \mathbf{j_o})$**-perturbed partition of $\mathbf{s^l}$.**

$$\text{Thus, } \sum_{j=1}^{n} \hat{s}_j^l = su - j_o \text{ follows.} \tag{4.3.7}$$

**Definition 4.3.3.** *Let a $(J_o, j_o)$-perturbation of the relation "$\geq$" between $s^f$ and $s_\pi^l$ be given. Furthermore, let $j_1$ denote the number of $i$'s with:*

$$s_i^f \leq s_{\pi_i}^l = \sigma_{J_o}^l - 1. \qquad (d6)$$

*(Obviously, $j_1 \in \{0, 1, \cdots, L_{J_o+1} - L_{J_o}\}$ if $\sigma_{J_o}^l - 1 = \sigma_{J_o+1}^l$ and*

$$j_1 = 0 \qquad \text{if } \sigma_{J_o}^l - 1 > \sigma_{J_o+1}^l.) \qquad (d7)$$

*Then $\hat{s}_\pi^l$ with*

$$\hat{s}_{\pi_i}^l = \begin{cases} s_{\pi_i}^l - 1 & \text{for} \quad s_i^f < s_{\pi_i}^l = \sigma_{J_o}^l \quad (d8) \\ & \text{(see (d3) from Definition 4.3.1),} \\ s_{\pi_i}^l & \text{otherwise} \quad (d9) \end{cases}$$

*is called a* $(\mathbf{J_o}, \mathbf{j_o}, \mathbf{j_1})$**-perturbed permutation of the** $(\mathbf{J_o}, \mathbf{j_o})$**-perturbed partition** $\hat{\mathbf{s}}^{\mathbf{l}}$ **with respect to** $\mathbf{s^f}$.

$\hat{\mathbf{S}}_\pi^{\mathbf{f,l}}(\mathbf{J_o}, \mathbf{j_o}, \mathbf{j_1})$ *is the set of all $(J_o, j_o, j_1)$-perturbed permutations $\hat{s}_\pi^l$ of permutations $s_\pi^l$ of $s^l$, for which a $(J_o, j_o)$-perturbation of the relation "$\geq$" between $s^f$ and $s_\pi^l$ is present.*

(See also Remarks 4.3.1 following Definition 4.3.4.)

**Lemma 4.3.1.** *A permutation $\hat{s}_\pi^l$ of a $(J_o, j_o)$-perturbed partition $\hat{s}^l$ is an element of a set $\hat{S}_\pi^{f,l}(J_o, j_o, j_1)$ if and only if $\hat{s}_\pi^l$ fulfils the following conditions regarding $s^f$:*

$$s_i^f \geq \hat{s}_{\pi_i}^l \quad \text{if } \hat{s}_{\pi_i}^l \geq \sigma_{J_o}^l, \qquad (d10)$$

$$s_i^f > \hat{s}_{\pi_i}^l \quad \text{for } L_{J_o+1} - L_{J_o} - j_1 \text{ components } \hat{s}_{\pi_i}^l = \sigma_{J_o}^l - 1 \quad \text{if } \sigma_{J_o}^l - 1 = \sigma_{J_o+1}^l, \qquad (d11)$$

$$s_i^f \leq \hat{s}_{\pi_i}^l \quad \text{for } j_o + j_1 \text{ components } \hat{s}_{\pi_i}^l = \sigma_{J_o}^l - 1. \qquad (d12)$$

***Proof.***

1. ($\Rightarrow$): Let $s_\pi^l \in \hat{S}_\pi^{f,l}(J_o, j_o, j_1)$ be given.

(d10) is valid for $\hat{s}_{\pi_i}^l \geq \sigma_{J_o-1}^l$ according to (d1)

and for $\hat{s}^l_{\pi_i} = \sigma^l_{J_o}$ according to (d2) and (d8).

The condition (d11) is fulfilled for the remaining $L_{J_o+1} - L_{J_o} - j_1$ components $\hat{s}^l_{\pi_i} = \sigma^l_{J_o} - 1$ according to the Definition of $j_1$ (see (d6)).

(d12) is valid according to (d3) together with (d8) ($j_0$ components for this case),

and to (d6) ($j_1$ components).

2. ($\Leftarrow$): Now, let $\hat{s}^l_\pi$ be a permutation of $\hat{s}^l$ satisfying (d10), (d11) and (d12).

A permutation $s^l_\pi$ of $s^l$ may then be constructed in the following way:

$$s^l_{\pi_i} = \begin{cases} \hat{s}^l_{\pi_i} + 1 & \text{for} \quad j_o \text{ components } \hat{s}^l_{\pi_i} = \sigma^l_{J_o} - 1 \geq s^f_i \\ & \quad \text{(which thus also satisfies (d12)),} \\ \hat{s}^l_{\pi_i} & \text{otherwise.} \end{cases} \tag{4.3.8}$$

We show that $s^l_\pi$ fulfils the conditions from Definition 4.3.1:

(d1) follows from (d10) (specifically for $\hat{s}^l_{\pi_i} \geq \sigma^l_{J_o-1} (> \sigma^l_{J_o})$).

(d2): According to Definition 4.3.2 (and (4.3.3)) $\hat{s}^l_{\pi_i} = \sigma^l_{J_o}$ is valid for $L_{J_o} - L_{J_o-1} - j_0$ components. Only (d10) can be present in Lemma 4.3.1 for these components, which means $s^f_i \geq \hat{s}^l_{\pi_i}$ (see also (4.3.8)). (d2) then follows.

(d12) (and (4.3.3)), Definition 4.3.2 (see (d5)) and (4.3.8) yield (d3).

Vice versa, the permutation $s^l_\pi$ leads to $\hat{s}^l_\pi$, according to (d8) and (d9). Keep in mind (d12) and (d11) then, $\hat{s}^l_\pi$ is an element of the set $\hat{S}^{f,l}_\pi(J_o, j_o, j_1)$. ■

Definition 4.3.2 and Lemma 4.3.1 obviously yield:

**Lemma 4.3.2.** . *Let $\hat{S}^{f,l}_\pi(J^1_o, j^1_o, j^1_1)$ and $\hat{S}^{f,l}_\pi(J^2_o, j^2_o, j^2_1)$ (with respect to $s^f$) be given with $J^1_o \neq J^2_o$ or $j^1_o \neq j^2_o$ or $j^1_1 \neq j^2_1$.*

*Then, $\hat{S}^{f,l}_\pi(J^1_o, j^1_o, j^1_1) \bigcap \hat{S}^{f,l}_\pi(J^2_o, j^2_o, j^2_1) = \varnothing$ follows.*

***Proof.***

Case $J_o^1 \neq J_o^2$ : $\sigma^l_{J_o^1}$ and $\sigma^l_{J_o^2}$, which were reduced by 1 for perturbed permutations of $\hat{S}^{f,l}_\pi(J_o^1, j_o^1, j_1^1)$ or $\hat{S}^{f,l}_\pi(J_o^2, j_o^2, j_1^2)$ (see Definition 4.3.2), are different.

Case $J_o^1 = J_o^2$ and $j_o^1 \neq j_o^2$ : Perturbed permutations of $\hat{S}^{f,l}_\pi(J_o^1, j_o^1, j_1^1)$ and $\hat{S}^{f,l}_\pi(J_o^2, j_o^2, j_1^2)$ then have different numbers of components with the value $\sigma^l_{J_o^{1=2}}$, since different numbers of components were reduced by 1 (see Definition 4.3.2).

Case $J_o^1 = J_o^2,\ j_o^1 = j_o^2$ and $j_1^1 \neq j_1^2$ : Perturbed permutations of $\hat{S}^{f,l}_\pi(J_o^1, j_o^1, j_1^1)$ and $\hat{S}^{f,l}_\pi(J_o^2, j_o^2, j_1^2)$ have different numbers of components for which (d11) or (d12) from Lemma 4.3.1 is valid.

■

**Definition 4.3.4.** *Let a permutation $\hat{s}^l_\pi$ of a $(J_o, j_o)$-perturbed partition $\hat{s}^l$ from the set $\hat{S}^{f,l}_\pi(J_o, j_o, j_1)$ (with respect to $s^f$) be given.*

*The set of requirements $w \in B_{n,k_0}$ which fulfils the properties:*

$$w_i \in \{\hat{s}^l_{\pi_i}, \hat{s}^l_{\pi_i} + 1, \ldots, k_o\} \qquad \text{if } s_i^f = \hat{s}^l_{\pi_i} \geq \sigma^l_{J_o} \qquad (d13)$$

$$\begin{cases} w_i \in \{\hat{s}^l_{\pi i} = \sigma^l_{J_o} - 1, \hat{s}^l_{\pi_i} + 1, \ldots, k_0\} \\ \quad \text{with at most } \boldsymbol{j_1} \text{ coordinates } w_i = \sigma^l_{J_o} - 1, \\ \qquad\qquad \text{if } s_i^f \leq \hat{s}^l_{\pi_i} = \sigma^l_{J_o} - 1, \end{cases} \qquad (d14)$$

$$w_i = \hat{s}^l_{\pi_i} \qquad \text{otherwise} \qquad (d15)$$

*is denoted by $B^2_{n;k_0}(s^f, \hat{s}^l_\pi)$.*

**Remarks 4.3.1.** *(d14) (and (d13)) shows that the reduction of components of the value $\sigma^l_{J_o}$ by 1 in order to determine a $(J_o, j_o)$-perturbed partition and corresponding permutations for $B^2_{n;k_0}(s^f, \hat{s}^l_\pi)$ is in fact not necessary in the case $\sigma^l_{J_o} - 1 > \sigma^l_{J_o+1}$, which means $\mathbf{j_1} = 0$ (see (d7)). However, this method leads to clearer and more uniform representations of the Definitions 4.3.2, 4.3.3, 4.3.4 and so on, where then distinctions in certain cases for the representations are not necessary.*

**Lemma 4.3.3.** *Let a permutation $\hat{s}^l_\pi$ of a $(J_o, j_o)$-perturbed partition $\hat{s}^l$ from a set $\hat{S}^{f,l}_\pi(J_o, j_o, j_1)$ (with respect to $s^f$) be given. In addition, let $w \in B^2_{n;k_0}(s^f, \hat{s}^l_\pi)$.*

*Then, in Definition 4.3.4 the case "otherwise" with $w_i = \hat{s}^l_{\pi_i}$ is valid if*

$$s^f_i > \hat{s}^l_{\pi_i} \text{ or} \tag{d16}$$

$$s^f_i \leq \hat{s}^l_{\pi_i} < \begin{cases} \sigma^l_{J_o+1} & \text{if } \sigma^l_{J_o} - 1 = \sigma^l_{J_o+1}, \\ \sigma^l_{J_o} & \text{if } \sigma^l_{J_o} - 1 > \sigma^l_{J_o+1}. \end{cases} \tag{d17}$$

***Proof.***

Case $\hat{s}^l_{\pi_i} \geq \sigma^l_{J_o}$: The equality $s^f_i \geq \hat{s}^l_{\pi_i}$ follows according to (d10) from Lemma 4.3.1.

$s^f_i = \hat{s}^l_{\pi_i} (\geq \sigma^l_{J_o})$ can be found in (d13) of Definition 4.3.4.

$s^f_i > \hat{s}^l_{\pi_i}$ belongs to "otherwise" in this definition.

Case $\hat{s}^l_{\pi_i} = \sigma^l_{J_o} - 1$: $s^f_i \leq \hat{s}^l_{\pi_i} (= \sigma^l_{J_o} - 1)$ can be found in (d14) of Definition 4.3.4.

$s^f_i > \hat{s}^l_{\pi_i}$ belongs to "otherwise" in this definition.

Case $\hat{s}^l_{\pi_i} < \sigma^l_{J_o} - 1$: This means $\hat{s}^l_{\pi_i} < \begin{cases} \sigma^l_{J_o+1} & \text{if } \sigma^l_{J_o} - 1 = \sigma^l_{J_o+1}, \\ \sigma^l_{J_o} & \text{if } \sigma^l_{J_o} - 1 > \sigma^l_{J_o+1}. \end{cases}$

If in addition $s^f_i \leq \hat{s}^l_{\pi_i}$, then (d17) is valid and if $s^f_i > \hat{s}^l_{\pi_i}$, then (d16).

■

**Lemma 4.3.4.** *Let a set $B^2_{n;k_0}(s^f, \hat{s}^l_\pi)$ as in Definition 4.3.4 be given.*

*Then exactly $\binom{j_o + j_1}{j_o}$ requirements $w \in B^2_{n;k_0}(s^f, \hat{s}^l_\pi)$ exist, which satisfy the cases $C[su, w]$ and $C[w, su]$ simultaneously.*

(Probabilities of $w$ from Lemma 4.3.4 are added in order to compute $p^{*1}_{fl}$ and also $p^{*2}_{fl}$. Therefore, these probabilities must be subtracted once from $p^{*1}_{fl} + p^{*2}_{fl}$ for the determination of $p^*_{fl}$ (see (4.3.6)).

***Proof.***
**Case** $\sigma^l_{J_o} - 1 = \sigma^l_{J_o+1}$ : ($j_1 > 0$ is possible in this case, see Definition 4.3.3).

If the components of $s^l_\pi$ and of the permutation $\hat{s}^l_\pi$ of a $(J_o, j_o)$-perturbed partition $\hat{s}^l$ from $\hat{S}^{f,l}_\pi(J_o, j_o, j_1)$ (with respect to $s^f$) are compared, then

$$\hat{s}^l_{\pi_i} < s^l_{\pi_i} \text{ may only be possible if } s^f_i \le \hat{s}^l_{\pi_i} = \sigma^l_{J_o} - 1 \qquad (*1)$$
(see also Definition 4.3.4)

is valid according to (d3) of Definition 4.3.1 and (d5) of Definition 4.3.2.

The condition (d14) from Definition 4.3.4:

$$s^f_i \le \hat{s}^l_{\pi_i} = \sigma^l_{J_o} - 1 \text{ is valid for exactly } j_o + j_1 \text{ components } \hat{s}^l_{\pi_i} \qquad (*2)$$

according to Lemma 4.3.1, (d12).

In relation to (d14) from Definition 4.3.4 let:

$$w_i = \sigma^l_{J_o} - 1 \text{ for } j_2 \text{ coordinates } w_i \text{ where } j_2 \le j_1. \qquad (*3)$$

From (*3) and (*1) (refer also to (*2))

$$\sum_{i=1}^{n} w_i \ge j_2(\sigma^l_{J_o} - 1) + [j_o + (j_1 - j_2)]\sigma^l_{J_o} + \sum_{i:\, if\ not\ s^f_i \le \hat{s}^l_{\pi_i} = \sigma^l_{J_o} - 1} s^l_{\pi_i}$$

$$= \sum_{i=1}^{n} s^l_{\pi_i} + (j_1 - j_2) = su + (j_1 - j_2) \ge su. \qquad (*4)$$

follows. In (*4) the left side of "$= su$" is only correct, if $j_2 = j_1$ and $w_i$ are as small as possible, according to Definition 4.3.4. This means that, in relation to (d14) from Definition 4.3.4, $j_1$ coordinates $w_i = \sigma^l_{J_o} - 1$ and $j_0$ coordinates $w_i = \sigma^l_{J_o}$ (see also (*2)).

Thus, exactly $\binom{j_o + j_1}{j_o}$ different requirements $w$ satisfy the cases $C[su, w]$ and $C[w, su]$ simultaneously.

**Case** $\sigma^l_{J_o} - 1 > \sigma^l_{J_o+1}$ : $j_1 = 0$ follows according to Definition 4.3.3.

Hence, in relation to (d14) from Definition 4.3.4, the possibility that $w_i = \sigma^l_{J_o} - 1$ (for any $i$ with $s^f_i \le \hat{s}^l_{\pi_i} = \sigma^l_{J_o} - 1$) does not exist.

Since $\binom{j_o + 0 = j_o}{j_o} = 1$, there is only one possibility, in which all coordinates $w_i$ from Definition 4.3.4 are as small as possible, which then implies $\sum_{i=1}^{n} w_i = su$. ■

**Theorem 4.3.5.** *Let $s^f \in S_{n;su;k_0}$ and $s^l \in S_{n;su;k_0}$ be given. In the case C[su, w] the following relationship is valid:*

$$s^*(s^f, w) = s^l \;\Leftrightarrow\; \begin{cases} w \in B^2_{n;k_0}(s^f, \hat{s}^l_\pi) \;\; \textit{where } \hat{s}^l_\pi \textit{ is an element of a} \\ \qquad\qquad \textit{set } \hat{S}^{f,l}_\pi(J_o, j_o, j_1) \; (\textit{with respect to } s^f), \\ \textit{or } \; w \in \{w \in B_{n;k_0} |\; w_i \geq s^f_i, \; i = 1, ..., n\} \\ \qquad\qquad \textit{if } s^f = s^l. \end{cases}$$

***Proof.***

1. ($\Rightarrow$): Let a requirement $w \in B_{n;k_0}$ with $\sum\limits_{i=1}^{n} w_i > su$ and $s^*(s^f, w) = s^l$ be given.

The trivial case in which $w_i \geq s^f_i$ for $i = 1, ..., n$, so that a permutation $s^l_\pi$ of $s^l$, $s^f$ and $w$ satisfy (4.2.5), is only possible if $s^f_i = s^l_{\pi_i}$ for $i = 1, ..., n$.

Now, let $s^l_\pi \neq s^f$ be a permutation of $s^l$ such that $s^l_\pi, s^f$ and $w$ satisfy (4.2.5).

$$w_i \leq s^f_i \;\Rightarrow\; s^l_{\pi_i} = w_i \;\; (\text{thus also } s^l_{\pi_i} \leq s^f_i) \tag{*1}$$

follow from (4.2.5), case $C[su, w]$.

We next consider the iterative method from Section 4.2.2, case C[su, w]:

Since $s^{*'}_j = s^{*'}_j + 1$ for $s^{*'}_j = min\{s^{*'}_i \mid s^{*'}_i < w_i\}$,

components of different permutations of $s^l$ which, together with $s^f$ and $w$, satisfy (4.2.5), differ by at most 1.

(This can be possible only if $j$ with $s^{*'}_j = min\{s^{*'}_i \mid s^{*'}_i < w_i\}$ is not unique in the final iteration steps).

Explained in greater detail, $s^l_{\pi_{i_0}}$ can be different if:

$$s^l_{\pi_{i_0}} = max\{s^l_{\pi_i} \mid s^f_i < s^l_{\pi_i} \leq w_i\} \tag{*2}$$

and if $i_1$ exists so that

$$s^f_{i_1} \leq s^l_{\pi_{i_1}} = s^l_{\pi_{i_0}} - 1 < w_{i_1}. \tag{*3}$$

(Then $s^{l'}_{\pi_{i_1}}$ could be increased by 1 instead of $s^{l'}_{\pi_{i_0}}$ in the final iteration steps if the iterative method from Section 4.2.2 (case C[su, w]) is used.)

Other relationships between $s^f_i, w_i, s^l_{\pi_i}$ and $s^l_{\pi_{i_0}}$ from (*2) (in relation to

(*1)) can be:

$$s_i^f = s_{\pi_i}^l = s_{\pi_{i_0}}^l < w_i, \tag{*4}$$

$$s_{i_2}^f < s_{\pi_{i_2}}^l = s_{\pi_{i_0}}^l - 1 = w_{i_2} \text{ for } i = i_2. \tag{*5}$$

If $s_i^f < w_i$, then the relationship

$$s_i^f < s_{\pi_i}^l = w_i < s_{\pi_{i_0}}^l - 1 \tag{*6}$$

between $s_i^f, w_i, s_{\pi_i}^l$ and $s_{\pi_{i_0}}^l$ as in (*2) remains possible, in addition to the previous (*2),(*3),(*4) or (*5), according to the iterative method from Section 4.2.2, case C[su, w].

With regard to Definition 4.3.1, we now use

$s_{\pi_{i_0}}^l$ (from (*2)) as $\sigma_{J_o}^l$,

and the number of $i_0$, for which (*2) is satisfied, as $j_0$.

$(J_o, j_o)$ is then a perturbation of the relation "$\geq$" between $s^f$ and $s_\pi^l$, according to (*1), ..., (*5) :

(d1): $s_{\pi_i}^l \geq \sigma_{J_o-1}^l (> \sigma_{J_o}^l)$ can only be possible if (*1) is valid,
from which (d1) follows,

(d2) and (d3): $s_{\pi_i}^l = \sigma_{J_o}^l$ is only valid if (*2) and (*4) are valid,
then (d3) follows from (*2) and (d2) from (*4).

Now, $\hat{s}_\pi^l$ with

$$\hat{s}_{\pi_i}^l = \begin{cases} s_{\pi_i}^l - 1 & \text{if (*2) is satified for } \ i = i_0, \\ s_{\pi_i}^l & \text{otherwise} \end{cases}$$

is a permutation of the $(J_o, j_o)$-perturbed partition $\hat{s}^l$ of $s^l$.

With regard to Definition 4.3.3, we set

$j_1$ equal to the number of $i_2$'s for which (*5) is satisfied.

Obviously, $\hat{s}_\pi^l \in \hat{S}_\pi^{f,l}(J_o, j_o, j_1)$.

Finally, we show that the considered $w$ is an element of $B_{n;k_0}^2(s^f, \hat{s}_\pi^l)$ (see Definition 4.3.4):

(d13): follows, in particular, from (*4) and (*1) with $s_{\pi_i}^l = s_i^f \geq s_{\pi_{i_0}}^l$,

(d14): follows, in particular, from (*2), (*3) and (*5)

(considering the previous determination of $\hat{s}^l_\pi$),

(d15): follows, in particular, from (*1) and (*6).

2. ($\Leftarrow$): Let $w \in B^2_{n;k_0}(s^f, \hat{s}^l_\pi)$.

If $w_i \geq s^f_i$ for $i = 1, ..., n$, then $s^*(s^f, w) = s^f (= s^l)$ follows immediately.

Now, let $i$ exist with $w_i < s^f_i$.

We will show that the iterative method from Section 4.2.2, case C[su, w], initially leads to $\hat{s}^l_\pi$ (from this theorem), and then to a $s^l_\pi$ (as in (4.3.8)). This means that $s^*(s^f, w) = s^l$.

At first, we note that

$$\sum_{i:s^f_i > w_i} (s^f_i - w_i) = \sum_{i:s^l_{\pi_i} > s^f_i} (s^l_{\pi_i} - s^f_i) \tag{*7}$$

is a necessary condition for $s^*(s^f, w)(= s^l_\pi) = s^l$ since according to the iterative method from Section 4.2.2, case C[su, w] (together with (4.2.5)), differences between $s^f_i$ and $w_i$, in the cases that $s^f_i > w_i$, are used in order to increase $s^f_i$ to a certain $s^l_{\pi_i}$ in the cases that $s^f_i < w_i$ (where $s^l_\pi$ is a permutation of $s^l = s^*(s^f, w)$).

We prove that the necessary condition is valid:

$s^f_i > w_i$ is only possible in the case (d15) of Definition 4.3.4 where $w_i = \hat{s}^l_{\pi_i}$. Then

$$\sum_{i:s^f_i > w_i} (s^f_i - w_i) = \sum_{i:s^f_i > \hat{s}^l_{\pi_i}} (s^f_i - \hat{s}^l_{\pi_i}) \tag{*8}$$

follows.

According to (d8) (together with (d3)), and since $\sum_{i=1}^n s^f_i = \sum_{i=1}^n s^l_i = su$,

$$\sum_{i:s^f_i > \hat{s}^l_{\pi_i}} (s^f_i - \hat{s}^l_{\pi_i}) = \sum_{i:\hat{s}^l_{\pi_i} > s^f_i} (\hat{s}^l_{\pi_i} - s^f_i) + j_o \tag{*9}$$

is valid.

$$\sum_{i:\hat{s}^l_{\pi_i}>s^f_i} (\hat{s}^l_{\pi_i} - s^f_i) + j_o = \sum_{i:s^l_{\pi_i}>s^f_i} (\, s^l_{\pi_i} - s^f_i) \qquad (*10)$$

follows for $s^l_\pi$ as in (4.3.8).

Now, (*8), (*9) and (*10) imply (*7).

Lastly, the consideration of the following cases shows that the iterative method from Section 4.2.2, case C[su, w], indeed initially leads to $\hat{s}^l_\pi$ (from this theorem) and then to a $s^l_\pi$ (as in (4.3.8)):

Case $s^f_i > w_i$:

According to the iterative method (and also according to (4.2.5)) $w_i = \hat{s}^l_{\pi_i} (= s^l_{\pi_i})$ follows in this case which corresponds to (d15).

Case $s^f_i = w_i$:

The relationship $s^f_i = s^l_{\pi_i}$ remains according to the iterative method and also as in the relevant case of Definition 4.3.4 ($s^f_i = s^l_{\pi_i} (= \hat{s}^l_{\pi_i})$).

Case $s^f_i < w_i < \begin{cases} \sigma^l_{J_o+1} & \text{if } \sigma^l_{J_o} - 1 = \sigma^l_{J_o+1}, \\ \sigma^l_{J_o} & \text{if } \sigma^l_{J_o} - 1 > \sigma^l_{J_o+1} \end{cases}$ :

$s^f_i$ are increased up to $w_i$ using the iterative method. This means $s^l_{\pi_i} = w_i (= \hat{s}^l_{\pi_i})$ which corresponds to (d17) (partial case of (d15)).

Case $s^f_i < w_i$ and $w_i \geq \begin{cases} \sigma^l_{J_o+1} & \text{if } \sigma^l_{J_o} - 1 = \sigma^l_{J_o+1}, \\ \sigma^l_{J_o} & \text{if } \sigma^l_{J_o} - 1 > \sigma^l_{J_o+1} \end{cases} \geq s^f_i$:

Initially, $s^f_i$ are increased up to $\sigma^l_{J_o} - 1$ using the iterative method (which corresponds to $\hat{s}^l_{\pi_i}$ from (d14)) of Definition 4.3.4. $j_0$ units then remain, which can be used in order to further increase the $j_0$ parts of the value $\sigma^l_{J_o} - 1$ by 1 (if $w_i \geq \sigma^l_{J_o}$) (which corresponds to $\hat{s}^l_{\pi_i}$ from Definition 4.3.4, (d13) with $s^f_i = \sigma^l_{J_o}$), according to the iterative method.

Case $s^f_i < w_i$ and $w_i > s^f_i > \begin{cases} \sigma^l_{J_o+1} & \text{if } \sigma^l_{J_o} - 1 = \sigma^l_{J_o+1}, \\ \sigma^l_{J_o} & \text{if } \sigma^l_{J_o} - 1 > \sigma^l_{J_o+1} \end{cases}$ :

A further increase of parts is no longer possible according to the iterative method (see also (*7), (*8), (*9) and (*10)). Only the relationship $s^l_\pi = s^f_i (= \hat{s}^l_\pi)$ remains possible, as also in (d13). ■

**Theorem 4.3.6.** *Let* $\hat{s}_{\pi}^{l,1} \in \hat{S}_{\pi}^{f,l}(J_o^1, j_o^1, j_1^1)$ *and* $\hat{s}_{\pi}^{l,2} \in \hat{S}_{\pi}^{f,l}(J_o^2, j_o^2, j_1^2)$ *(with respect to* $s^f$*) be given with*

$$\hat{s}_{\pi}^{l,1} \neq \hat{s}_{\pi}^{l,2}. \tag{*1}$$

*(*$J_o^1 = J_o^2,\ j_o^1 = j_o^2,\ j_1^1 = j_1^2$ *are also possible. In the case* $\hat{S}_{\pi}^{f,l}(J_o^1, j_o^1, j_1^1) \neq \hat{S}_{\pi}^{f,l}(J_o^2, j_o^2, j_1^2)$*, the relation* $\hat{s}_{\pi}^{l,1} \neq \hat{s}_{\pi}^{l,2}$ *follows from Lemma 4.3.2.)*

*Then,*

$$B_{n;k_0}^2(s^f, \hat{s}_{\pi}^{l,1}) \bigcap B_{n;k_0}^2(s^f, \hat{s}_{\pi}^{l,2}) = \varnothing.$$

*(Furthermore,* $B_{n;k_0}^2(s^f, \hat{s}_{\pi}^{l}) \bigcap \{w \in B_{n;k_0} |\ w_i \geq s_i,\ i = 1, ..., n\} = \varnothing$ *is valid in the case* $s^f = s^l$ *where* $\hat{s}_{\pi}^{l} \in \hat{S}_{\pi}^{f,l}(J_o, j_o, j_1)$ *(with respect to* $s^f$*).)*

***Proof.*** Let $w^1$ be any element of the set $B_{n;k_0}^2(s^f, \hat{s}_{\pi}^{l,1})$ and $w^2$ of $B_{n;k_0}^2(s^f, \hat{s}_{\pi}^{l,2})$.

Furthermore, let

$$\hat{s}_{\pi_{i_o}}^{l,1} \neq \hat{s}_{\pi_{i_o}}^{l,2}$$

according to (*1).

In order to show that $w^1 \neq w^2$ we consider 3 cases.

**Case 1:** Let $\hat{s}_{\pi_{i_o}}^{l,1} < s_{i_o}^f$.

According to Definition 4.3.4

$$w_{\pi_{i_o}}^1 = \hat{s}_{\pi_{i_o}}^{l,1} \qquad \text{(in particular, see (d15)and (d10))},$$

$$w_{\pi_{i_o}}^2 \begin{cases} = \hat{s}_{\pi_{i_o}}^{l,2} & \text{if } \hat{s}_{\pi_{i_o}}^{l,2} < s_{i_o}^f, \\ \geq \hat{s}_{\pi_{i_o}}^{l,2} & \text{if } \hat{s}_{\pi_{i_o}}^{l,2} \geq s_{i_o}^f \end{cases}$$

follows. Thus, $w_{\pi_{i_o}}^1 \neq w_{\pi_{i_o}}^2$.

**Case 2:**

$$\begin{aligned} &\text{Let } s_{i_o}^f \leq \hat{s}_{\pi_{i_o}}^{l,2} < \hat{s}_{\pi_{i_o}}^{l,1} \\ &\text{and } \sigma_{J_o^1}^l \leq \sigma_{J_o^2}^l. \end{aligned} \tag{*2}$$

The relationship

$\hat{s}^{l,1}_{\pi_{i_o}} \geq \sigma^l_{J^1_o}$

is not possible according to (*2), Definition 4.3.4 and Lemma 4.3.1 (in particular, see (d10)).

Hence the relationship

$\hat{s}^{l,1}_{\pi_{i_o}} \leq \sigma_{J^1_o} - 1$

remains to consider.

$w^1_{\pi_{i_o}} \geq \hat{s}^{l,1}_{\pi_{i_o}}$

follows for requirements $w^1 \in B^2_{n;k_0}(s^f, \hat{s}^{l,1}_\pi)$

and with regard to (*2)

$\hat{s}^{l,2}_{\pi_{i_o}} < (\hat{s}^{l,1}_{\pi_{i_o}} \leq \sigma_{J^1_o} - 1 \leq)\ \sigma_{J^2_o} - 1$

is valid.
Thus,

$w^2_{\pi_{i_o}} = \hat{s}^{l,2}_{\pi_{i_o}}$

is valid according to Definition 4.3.4 (see (d15) together with (d17)), hence $w^1_{\pi_{i_o}} \neq w^2_{\pi_{i_o}}$.

**Case 3:**

$$\begin{aligned} &\text{Let } s^f_{i_o} \leq \hat{s}^{l,2}_{\pi_{i_o}} < \hat{s}^{l,1}_{\pi_{i_o}} \\ &\text{and } \sigma^l_{J^1_o} > \sigma^l_{J^2_o}. \end{aligned} \tag{*3}$$

Now,

$$\exists\, i_1 :\ s^f_{i_1} < s^{l,1}_{\pi_{i_1}} = \hat{s}^{l,1}_{\pi_{i_1}} + 1 = \sigma^l_{J^1_o} \tag{*4}$$

according to (d3) and (d8).

In relation to $\hat{s}^{l,2}_\pi$ and $i_1$ either

$$s^{l,2}_{\pi_{i_1}} \leq \sigma^l_{J^2_o} \tag{*5a}$$

or

$$\sigma^l_{J^2_o} < s^{l,2}_{\pi_{i_1}} \leq s^f_{i_1} \text{ (see also (d1))} \tag{*5b}$$

is valid.

(*4),(*5a) and $\sigma^l_{J^1_o} > \sigma^l_{J^2_o}$ or (*4) and (*5b), respectively yield

$$s^{l,2}_{\pi_{i_1}} \leq \sigma^l_{J^2_o} < \sigma^l_{J^1_o} = \hat{s}^{l,1}_{\pi_{i_1}} + 1 = s^{l,1}_{\pi_{i_1}} \tag{*6a}$$

or

$$s^{l,2}_{\pi_{i_1}} \leq s^f_{i_1} < \hat{s}^{l,1}_{\pi_{i_1}} + 1 = s^{l,1}_{\pi_{i_1}} = \sigma^l_{J^1_o}. \tag{*6b}$$

From (*6a) and (*6b) together

$$s^{l,2}_{\pi_{i_1}} < s^{l,1}_{\pi_{i_1}} = \sigma^l_{J^1_o} \tag{*7}$$

follows (where $i_o = i_1$ is possible).

Since $s^{l,2}_\pi$ is a permutation of $s^{l,1}_\pi$

$$\exists\, i_2\ (i_2 \neq i_1):\ s^{l,2}_{\pi_{i_2}} = s^{l,1}_{\pi_{i_1}} (= \sigma^l_{J^1_o} \geq \sigma^l_{J^2_o} + 1). \tag{*8}$$

Furthermore,

$$s^f_{i_2} \geq \hat{s}^{l,2}_{\pi_{i_2}} = s^{l,2}_{\pi_{i_2}} \tag{*9}$$

is valid (see also (d1)).

If $\hat{s}^{l,1}_{\pi_{i_2}} < \hat{s}^{l,2}_{\pi_{i_2}} (\leq s^f_{i_2})$ then $w^1_{i_2} = \hat{s}^{l,1}_{\pi_{i_2}} < \hat{s}^{l,2}_{\pi_{i_2}} \leq w^2_{i_2}$ follows according to Definition 4.3.4 and Lemma 4.3.3. Thus

$$w^1_{i_2} \neq w^2_{i_2}. \tag{*10}$$

In addition, if $\hat{s}^{l,1}_{\pi_{i_2}} \geq \hat{s}^{l,2}_{\pi_{i_2}} (= s^{l,2}_{\pi_{i_2}} = s^{l,1}_{\pi_{i_1}})$ (see also (*8) and (*9)), then we can conclude similarly in the two possible subcases $s^{l,1}_{\pi_{i_2}} = s^{l,2}_{\pi_{i_2}}$ and $s^{l,1}_{\pi_{i_2}} > s^{l,2}_{\pi_{i_2}}$ again, with $s^{l,2}_\pi$ as a permutation of $s^{l,1}_\pi$:

$$\exists\, i_3\ (i_3 \neq i_2 \wedge i_3 \neq i_1):\ s^{l,2}_{\pi_{i_3}} = s^{l,1}_{\pi_{i_2}} (\geq \sigma^l_{J^1_o} \geq \sigma^l_{J^2_o} + 1) \text{ and so on.}$$

Since the numbers of parts of $s^{l,2}_\pi$ and $s^{l,1}_\pi$ are finite, $w^1_{i_m} \neq w^2_{i_m}$ follows any times for a certain $m$ analogous to (*10).

(Finally, the Definition 4.3.4 of $B^2_{n;k_0}(s^f, \hat{s}^l_\pi)$ and Lemma 4.3.3 directly yields $B^2_{n;k_0}(s^f, \hat{s}^l_\pi) \bigcap \{w \in B_{n;k_0} |\ w_i \geq s_i,\ i = 1, ..., n\} = \varnothing$ if $s^f = s^l$.) ■

**Example 4.3.1.** *Let* $n = 11$, $su = 37$, $k_0 = 8$ *and*

$s^f = (6,\ 5,\ 5,\ 4,\ 4,\ 4,\ 3,\ 2,\ 2,\ 2,\ 0)$,

$s^l = (5,\ 5,\ 4,\ 4,\ 4,\ 3,\ 3,\ 3,\ 2,\ 2,\ 2)$

*(this means* $\sigma_1^l = 5$, $\sigma_2^l = 4$, $\sigma_3^l = 3$, $\sigma_4^l = 2$ *(compare to the beginning of Section 4.3, (4.3.4))*

*and (for example)*

$s_\pi^l = (5,\ 5,\ 4,\ 4,\ 3,\ 2,\ 3,\ 4,\ 3,\ 2,\ 2)$

*be given.*

(a) • *According to Definition 4.3.1 a* $(J_o = 2, j_o = 1)$*-perturbation (with* $\sigma_{J_o}^l = \sigma_2^l = 4$*) of the relation* $''\geq''$ *between*

$$\left.\begin{array}{l} s^f = (6,\ 5,\ 5,\ 4,\ 4,\ 4,\ 3,\ \mathbf{2},\ 2,\ 2,\ 0) \\ \textit{and} \\ s_\pi^l = (5,\ 5,\ 4,\ 4,\ 3,\ 2,\ 3,\ \mathbf{4},\ 3,\ 2,\ 2) \end{array}\right\} \qquad (*1)$$

*is valid, since*

$s_i^f \geq s_{\pi_i}^l$ *for any* $s_{\pi_i}^l \geq 5$,

$s_3^f \geq s_{\pi_3}^l$ *and* $s_4^f \geq s_{\pi_4}^l$ *(where* $s_{\pi_3}^l = s_{\pi_4}^l = 4 = \sigma_2^l = \sigma_{J_o}^l$*)*

*and* $s_8^f < s_{\pi_8}^l = 4 = \sigma_{J_o}^l$.

• *Formally,*

$\hat{s}^l = (5,\ 5,\ 4,\ 4,\ \mathbf{3},\ 3,\ 3,\ 3,\ 2,\ 2,\ 2)$

*is a (2, 1)-perturbed partition of* $s^l$ *according to Definition 4.3.2.*

• *In relation to Definition 4.3.3 we find that*

$j_1 = 2$, *since* $s_i^f \leq s_{\pi_i}^l = 3 = \sigma_2^l - 1$ *for the two indices* $i = 7$ *and* $i = 9$.

*Then, according to (d8) and (d9)*

$\hat{s}_\pi^l = (5,\ 5,\ 4,\ 4,\ 3,\ 2,\ 3,\ \mathbf{3},\ 3,\ 2,\ 2)$

*is an element of the set* $\hat{S}_\pi^{f,l}(2, 1, 2)$.

- *Continuing our example the set* $B^2_{11,8}(s^f, \hat{s}^l_\pi)$ *is (see Definition 4.3.4):*

$$B^2_{11,8}(s^f, \hat{s}^l_\pi)$$
$$= \{w|\ w = (5,\ w_2,\ 4,\ w_4,\ 3,\ 2,\ w_7,\ w_8,\ w_9,\ 2,\ 2)$$
$$\textit{with } w_2 \in \{5,6,7,8\},\ w_4 \in \{4,5,\ldots,8\},\ w_j \in \{3,4,\ldots,8\}$$
$$\textit{for } j = 7,\ 8,\ 9 \textit{ and with at most 2 coordinates}$$
$$w_j = 3, j \in \{7,8,9\}\}.$$

(b) • *Altogether* $\hat{S}^{f,l}_\pi(2,1,2)$ *includes* $360 = 3 \cdot 12 \cdot 10$ *elements:*
*According to Lemma 4.3.1 the components*

- 5; 5 *(of* $\hat{s}^l$*) must be under the first 3 components of the elements of* $\hat{S}^{f,l}_\pi(2,1,2)$ *(see (d10)),* $(= \binom{3}{2} = 3$ *possibilities),*
- 4; 4; 3 *(of* $\hat{s}^l$*) must be under the first 6 components of the elements of* $\hat{S}^{f,l}_\pi(2,1,2)$ *(see (d10) and (d11)),*
  *(this means:* $* \binom{6-2}{3} \cdot 3 = *12$ *possibilities),*
- 3; 3; 3 *(of* $\hat{s}^l$*) must be under the last 5 components of the elements of* $\hat{S}^{f,l}_\pi(2,1,2)$ *(see (d12))* $(= \binom{5}{3} = 10$ *possibilities).*

(c) • *We now return to our discussion about the motivation for the concept of perturbed partitions based on this example:*

*If we use the iterative method from Section 4.2.2, case C[su, w], in relation to the given partition* $s^f$ *and a requirement* $w \in B^2_{11,8}(s^f, \hat{s}^l_\pi)$ *(with* $w_{7,8,9} > 3$*), then this method yields exactly one* $s^{*'}$ *in the second to last iteration step:*

$$s^{*'} = \hat{s}^l_\pi.$$

*In the last iteration step three possibilities for* $s^{*'}$ *follow :*

$$(5,\ 5,\ 4,\ 4,\ 3,\ 2,\ 4,\ 3,\ 3,\ 2,\ 2)$$

*or* $\quad (5,\ 5,\ 4,\ 4,\ 3,\ 2,\ 3,\ 4,\ 3,\ 2,\ 2)$

*or* $\quad (5,\ 5,\ 4,\ 4,\ 3,\ 2,\ 3,\ 3,\ 4,\ 2,\ 2)$

*(as permutations of $s^l$).*

*Sets $B^2_{n,k_0}(s^f, \hat{s}^l_\pi)$ of requirements, which imply transitions from $s^f$ to $s^l$ (see Theorem 4.3.5), are disjunct for different permutations $\hat{s}^l_\pi$ of perturbed partitions (see Theorem 4.3.6). An analogon for permutations of partitions $s^l$ themselves would not be valid.*

Now, we want to determine all requirements $w$ with

$$\sum_{i=1}^{n} w_i \geq su \text{ and } s^l = s^*(s^f, w) \qquad (*)$$

for given $s^f$ and $s^l$ in order initially to compute $p^{*2}_{fl}$ and later the elements $p^*_{fl} = \sum\limits_{w:s^l=s^*(s^f,w)} q(w)$ of PRMs.

For this purpose we can determine the perturbed permutations $\hat{s}^l_\pi$ of all permutations $s^l_\pi$ of $s^l$ (see Definition 4.3.3). If identical $\hat{s}^l_\pi$ follow for different $s^l_\pi$ then these $\hat{s}^l_\pi$ are naturally used only once in order to compute the corresponding $B^2_{n,k_0}(s^f, \hat{s}^l_\pi)$, according to Definition 4.3.4.

Then the set

$$\bigcup_{\hat{s}^l_\pi} B^2_{n,k_o}(s^f, \hat{s}^l_\pi) \ \ (\bigcup \{w \in B_{n;k_0} |\ w_i \geq s_i,\ i = 1, ..., n\} \text{ if } s^f = s^l)$$

includes all requirements $w$, which satisfy (*) (see Theorems 4.3.5 and 4.3.6).

A slightly different method would be for $J_o \in \{1, 2, \ldots, y\}$, $j_o \in \{1, 2, \ldots, L_{J_o} - L_{J_o-1}\}$ and

$$j_1 \in \begin{cases} \{0, 1, \ldots, L_{J_o+1} - L_{J_o}\} & \text{when } \sigma^l_{J_o} - 1 = \sigma^l_{J_o+1}, \\ \{0\} & \text{when } \sigma^l_{J_o} - 1 > \sigma^l_{J_o+1} \end{cases}$$

to determine all $\hat{s}^l_\pi$ from sets $\hat{S}^{f,l}_\pi(J_o, j_o, j_1)$ (as in Example 4.3.1(b)) if it is possible.

(Sets $\hat{S}^{f,l}_\pi(J_o, j_o, j_1)$ are disjunct for various $(J_o, j_o, j_1)$, see Lemma 4.3.2.)

According to Definition 4.3.4, Theorem 4.3.5 and Theorem 4.3.6

$$p_{fl}^{*2} = \sum_{\hat{S}_\pi^{f,l}(J_o,j_o,j_1)(\neq\varnothing)} \quad \sum_{\hat{s}_\pi^l \in \hat{S}_\pi^{f,l}(J_o,j_o,j_1)} \quad \sum_{w\in B^2_{n;k_0}(s^f,\hat{s}_\pi^l)} q(w)$$
$$+ \ \delta_{fl} \ * \sum_{w\in\{w\in B_{n;k_0} \mid w_i\geq s_i,\ i=1,\dots,n\}} q(w) \tag{4.3.9}$$

or, in more detail

$$p_{fl}^{*2} = \sum_{J_o=1}^{y} \sum_{j_o=1}^{L_{J_o}-L_{J_o-1}} \sum_{j_1\in\begin{cases} \{0,1,\dots,L_{J_o+1}-L_{J_o}\} & \text{when } \sigma_{J_o}^l - 1 = \sigma_{J_o+1}^l, \\ \{0\} & \text{when } \sigma_{J_o}^l - 1 > \sigma_{J_o+1}^l \end{cases}}$$
$$\sum_{\hat{s}_\pi^l\in\hat{S}_\pi^{f,l}(J_o,j_o,j_1)} \quad \sum_{w\in B^2_{n;k_0}(s^f,\hat{s}_\pi^l)} q(w) + \ \delta_{fl} \ * \sum_{w\in\{w\in B_{n;k_0} \mid w_i\geq s_i,\ i=1,\dots,n\}} q(w) \tag{4.3.10}$$

follows.

In the case of discrete uniformly distributed requirements, Definition 4.3.4 yields:

$$\sum_{w\in B^2_{n;k_0}(s^f,\hat{s}_\pi^l)} q(w) = \tfrac{1}{(k_0+1)^n} \left[ \prod_{i:\hat{s}_{\pi_i}^l = s_i^f \geq \sigma_{J_o}^l} (k_0+1-\hat{s}_{\pi_i}^l) \right.$$
$$\left. \left( (k_0-\sigma_{J_o}^l+2)^{j_o+j_i} - \binom{j_o+j_1}{j_1+1}(k_0-\sigma_{J_o}^l+1)^{j_o-1} - \dots - \binom{j_o+j_1}{j_o+j_1}(k_o-\sigma_{J_o}^l+1)^0 \right) \right],$$
$$\delta_{fl} \ * \sum_{w\in\{w\in B_{n;k_0} \mid w_i\geq s_i,\ i=1,\dots,n\}} q(w) = \ \delta_{fl} \ * \ \tfrac{1}{(k_0+1)^n} \prod_{i=1}^{n} (k_0+1-s_i^f). \tag{4.3.11}$$

**Remarks 4.3.2.** *If $\hat{s}_{\pi_i}^l \neq 0 \Rightarrow s_i^f = 0$, then $\hat{S}_\pi^{f,l}(J_o,j_o,j_1) \neq \emptyset$ at most for $J_o = 1, j_o = L_2 - L_1$ and $j_1 \leq L_3 - L_2$, if $\sigma_1 - 1 = \sigma_2$ (otherwise $j_1 = 0$).*

**Case C[w, su]** (*the requirements can be completely fulfilled*):

The considerations are analogous to the case C[su, w]. By that reason we present only definitions, lemmas and theorems in this part of the section however no corresponding proofs, which would be very similar to the proofs in the other case.

Let a partition $s^f \in S_{n;su;k_0}$ and a permutation $s^l_\pi$ of a partition $s^l \in S_{n;su;k_0}$ be given. We then compare the components of $s^f$ with the components of $s^l_\pi$ in order of increasing $s^l_{\pi_i}$.

**Definition 4.3.5.** *Let $J_o \in \{1, 2, \ldots, y$ (or $y+1$ for $L_y < n$)$\}$ (see (4.3.3)) and $j_3 \in \{1, 2, \ldots, L_{J_o} - L_{J_o-1}\}$.*

*If*

$$s^l_{\pi_i} \geq s^f_i \quad \textit{for any } s^l_{\pi_i} \leq \sigma^l_{J_o+1}, \tag{d18}$$

$$s^f_i \leq s^l_{\pi_i} \quad \textit{for } L_{J_o} - L_{J_o-1} - j_3 \textit{ of the } s^l_{\pi_i} = \sigma^l_{J_o} \tag{d19}$$

$$\textit{and } s^f_i > s^l_{\pi_i} \quad \textit{for } j_3 \textit{ of the } s^l_{\pi_i} = \sigma^l_{J_o}, \tag{d20}$$

*then we refer to a* $(\mathbf{J_o}, \mathbf{j_3})$**-perturbation of the relation "$\leq$" between $\mathbf{s^f}$ and $\mathbf{s^l_\pi}$.**

Formally, we define the $(J_o, j_3)$-perturbed partition $\hat{s}^l$ of $s^l$.

Contrary to $s^l$ the $(J_o, j_3)$-perturbed partition $\hat{s}^l$ has exactly $j_3$ components which are increased from $\sigma^l_{J_o}$ by 1 to $\sigma^l_{J_o} + 1$:

**Definition 4.3.6.** *Let $J_o$ and $j_3$ be given with $J_o \in \{1, 2, \ldots, y$ (or $y+1$ for $L_y < n$)$\}$ and $j_3 \in \{1, 2, \ldots, L_{J_o} - L_{J_o-1}\}$.*

$$\hat{s}^l : \begin{cases} \hat{s}^l_j = s^l_j & \textit{for} \quad j \in \{1, 2, \cdots, L_{J_o-1}\} \\ & \textit{and for} \quad j \in \{L_{J_o} + j_3 + 1, \cdots, n\} \quad (d21) \\ \hat{s}^l_j = s^l_j + 1 (= \sigma^l_{J_o} + 1) & \textit{for} \quad j \in \{L_{J_o-1} + 1, \cdots, L_{J_o-1} + j_3\} \quad (d22) \end{cases}$$

*is called the* $(\mathbf{J_o}, \mathbf{j_3})$**-perturbed partition of $\mathbf{s^l}$.**

$$\text{Thus, } \sum_{j=1}^{n} \hat{s}^l_j = su + j_3 \text{ follows.} \tag{4.3.12}$$

**Definition 4.3.7.** *Let a $(J_o, j_3)$-perturbation of the relation "$\leq$" between $s^f$ and $s^l_\pi$ be given. Furthermore, let $j_4$ denote the number of $i$'s with:*

$$s^f_i \geq s^l_{\pi_i} = \sigma^l_{J_o} + 1. \qquad (d23)$$

*(Obviously, $j_4 \in \{0, 1, \cdots, L_{J_o-1} - L_{J_o-2}\}$ if $\sigma^l_{J_o} = \sigma^l_{J_o-1} - 1$ and*

$$j_4 = 0 \qquad \text{if } \sigma^l_{J_o} < \sigma^l_{J_o-1} - 1.) \qquad (d24)$$

*Then $\hat{s}^l_\pi$ with*

$$\hat{s}^l_{\pi_i} = \begin{cases} s^l_{\pi_i} + 1 & \text{for} \quad s^f_i > s^l_{\pi_i} = \sigma^l_{J_o} \quad (d25) \\ & \text{(see (d20) from Definition 4.3.5),} \\ s^l_{\pi_i} & \text{otherwise} \quad (d26) \end{cases}$$

*is called a* $(\mathbf{J_o}, \mathbf{j_3}, \mathbf{j_4})$**-perturbed permutation of the** $(\mathbf{J_o}, \mathbf{j_3})$**-perturbed partition** $\hat{\mathbf{s}}^{\mathbf{l}}$ **with respect to** $\mathbf{s^f}$.

$\hat{\mathbf{S}}^{\mathbf{f,l}}_\pi(\mathbf{J_o}, \mathbf{j_3}, \mathbf{j_4})$ *is the set of all $(J_o, j_3, j_4)$-perturbed permutations $\hat{s}^l_\pi$ of permutations $s^l_\pi$ of $s^l$, for which a $(J_o, j_3)$-perturbation of the relation "$\leq$" between $s^l_\pi$ and $s^f$ is present.*

(See also Remarks 4.4.1 following Definition 4.3.5.)

**Lemma 4.3.7.** *A permutation $\hat{s}^l_\pi$ of a $(J_o, j_3)$-perturbed partition $\hat{s}^l$ is an element of a set $\hat{S}^{f,l}_\pi(J_o, j_3, j_4)$ if and only if $\hat{s}^l_\pi$ fulfils the following conditions regarding to $s^f$:*

$$s^f_i \leq \hat{s}^l_{\pi_i} \quad \text{if } \hat{s}^l_{\pi_i} \leq \sigma^l_{J_o}, \qquad (d27)$$

$$s^f_i < \hat{s}^l_{\pi_i} \quad \text{for } L_{J_o-1} - L_{J_o-2} - j_4 \text{ components } \hat{s}^l_{\pi_i} = \sigma^l_{J_o} + 1 \quad \text{if } \sigma^l_{J_o} + 1 = \sigma^l_{J_o-1}, \qquad (d28)$$

$$s^f_i \geq \hat{s}^l_{\pi_i} \quad \text{for } j_3 + j_4 \text{ components } \hat{s}^l_{\pi_i} = \sigma^l_{J_o} + 1, \qquad (d29)$$

Definition 4.3.6 and Lemma 4.3.7 obviously yield:

**Lemma 4.3.8.** *Let $\hat{S}^{f,l}_\pi(J^1_o, j^1_3, j^1_4)$ and $\hat{S}^{f,l}_\pi(J^2_o, j^2_3, j^2_4)$ (with respect to $s^f$) be given with $J^1_o \neq J^2_o$ or $j^1_3 \neq j^2_3$ or $j^1_4 \neq j^2_4$.*

*Then, $\hat{S}^{f,l}_\pi(J^1_o, j^1_3, j^1_4) \bigcap \hat{S}^{f,l}_\pi(J^2_o, j^2_3, j^2_4) = \varnothing$ follows.*

**Definition 4.3.8.** *Let a permutation $\hat{s}^l_\pi$ of a $(J_o, j_3)$-perturbed partition $\hat{s}^l$ from a set $\hat{S}^{f,l}_\pi(J_o, j_3, j_4)$ (with respect to $s^f$) be given.*

*The set of requirements $w \in B_{n,k_0}$ which fulfils the properties:*

$$w_i \in \{0, 1, \ldots, \hat{s}^l_{\pi_i}\} \qquad \text{if } s^f_i = \hat{s}^l_{\pi_i} \leq \sigma^l_{J_o} \qquad (d30)$$

$$\begin{cases} w_i \in \{0, 1, \ldots, \hat{s}^l_{\pi_i} = \sigma^l_{J_o} + 1\} \\ \quad \textit{with at most } \boldsymbol{j_4} \textit{ coordinates } w_i = \sigma^l_{J_o} + 1 \\ \qquad\qquad\qquad\qquad\qquad \textit{if } s^f_i \geq \hat{s}^l_{\pi_i} = \sigma^l_{J_o} + 1, \end{cases} \quad (d31)$$

$$w_i = \hat{s}^l_{\pi_i} \qquad \textit{otherwise} \qquad (d32)$$

*is denoted by* $B^1_{n;k_0}(s^f, \hat{s}^l_\pi)$.

**Remarks 4.3.3.** *(d31) (and (30)) shows that the increase of components of the value* $\sigma^l_{J_o}$ *by 1 in order to determine a* $(J_o, j_3)$*-perturbed partition and corresponding permutations for* $B^1_{n;k_0}(s^f, \hat{s}^l_\pi)$ *is in fact not necessary in the case* $\sigma^l_{J_o} + 1 < \sigma^l_{J_o-1}$*, which means* $\mathbf{j_4} = 0$ *(see (d24)). However, this method leads to clearer and more uniform representations of the Definitions 4.3.6, 4.3.7, 4.3.8 and so on, where then distinctions in certain cases in the representations are not necessary.*

**Lemma 4.3.9.** *Let a permutation* $\hat{s}^l_\pi$ *of a* $(J_o, j_3)$*-perturbed partition* $\hat{s}^l$ *from a set* $\hat{S}^{f,l}_\pi(J_o, j_3, j_4)$ *(with respect to* $s^f$*) be given. In addition, let* $w \in B^1_{n;k_0}(s^f, \hat{s}^l_\pi)$.

*Then, in Definition 4.3.8 the case "otherwise" with* $w_i = \hat{s}^l_{\pi_i}$ *is valid if*

$$s^f_i < \hat{s}^l_{\pi_i} \textit{ or} \qquad (d33)$$

$$s^f_i \geq \hat{s}^l_{\pi_i} > \begin{cases} \sigma^l_{J_o-1} & \textit{if } \sigma^l_{J_o} + 1 = \sigma^l_{J_o-1}, \\ \sigma^l_{J_o} & \textit{if } \sigma^l_{J_o} + 1 < \sigma^l_{J_o-1}. \end{cases} \qquad (d34)$$

**Theorem 4.3.10.** *Let* $s^f \in S_{n;su;k_0}$ *and* $s^l \in S_{n;su;k_0}$ *be given. In the case C[w, su] the following relationship is valid:*

$$s^*(s^f, w) = s^l \;\Leftrightarrow\; \begin{cases} w \in B^1_{n;k_0}(s^f, \hat{s}^l_\pi) \textit{ where } \hat{s}^l_\pi \textit{ is an element of a} \\ \qquad\qquad \textit{set } \hat{S}^{f,l}_\pi(J_o, j_3, j_4) \textit{ (with respect to } s^f\textit{)}, \\ \textit{or } w \in \{w \in B_{n;k_0} |\, w_i \leq s^f_i,\ i = 1, ..., n\} \\ \qquad\qquad\qquad\qquad \textit{if } s^f = s^l. \end{cases}$$

**Theorem 4.3.11.** *Let* $\hat{s}^{l,1}_\pi \in \hat{S}^{f,l}_\pi(J^1_o, j^1_3, j^1_4)$ *and* $\hat{s}^{l,2}_\pi \in \hat{S}^{f,l}_\pi(J^2_o, j^2_3, j^2_4)$ *(with respect to* $s^f$*) be given with*

$$\hat{s}^{l,1}_\pi \neq \hat{s}^{l,2}_\pi.$$

*($J_o^1 = J_o^2,\ j_3^1 = j_3^2,\ j_4^1 = j_4^2$ are the possible. In the case $\hat{S}_\pi^{f,l}(J_o^1, j_3^1, j_4^1) \neq \hat{S}_\pi^{f,l}(J_o^2, j_3^2, j_4^2)$, the relation $\hat{s}_\pi^{l,1} \neq \hat{s}_\pi^{l,2}$ follows from Lemma 4.3.8.)*

*Then,*
$B_{n;k_0}^1(s^f, \hat{s}_\pi^{l,1}) \bigcap B_{n;k_0}^1(s^f, \hat{s}_\pi^{l,2}) = \varnothing.$

*(Furthermore, $B_{n;k_0}^1(s^f, \hat{s}_\pi^l) \bigcap \{w \in B_{n;k_0} |\ w_i \leq s_i,\ i = 1, ..., n\} = \varnothing$ is valid in the case $s^f = s^l$ where $\hat{s}_\pi^l \in \hat{S}_\pi^{f,l}(J_o, j_3, j_4)$ (with respect to $s^f$).)*

We now compute the probability of requirements $w$ with $s^*(s^f, w) = s^l$ in the case C[w, su]:

$$p_{fl}^{*1} = \sum_{\hat{S}_\pi^{f,l}(J_o,j_3,j_4)(\neq\varnothing)} \ \sum_{\hat{s}_\pi^l \in \hat{S}_\pi^{f,l}(J_o,j_3,j_4)} \ \sum_{w \in B_{n;k_0}^1(s^f,\hat{s}_\pi^l)} q(w) + \delta_{fl} * \sum_{w \in \{w \in B_{n;k_0} |\ w_i \leq s_i,\ i=1,...,n\}} q(w) \tag{4.3.13}$$

or, in greater detail:

$$p_{fl}^{*1} = \sum_{J_o=1}^{y} \ \sum_{j_3=1}^{L_{J_o} - L_{J_o-1}} \ \sum_{j_4 \in \begin{cases} \{0, 1, \ldots, L_{J_o-1} - L_{J_o-2}\} & \text{when } \sigma_{J_o}^l + 1 = \sigma_{J_o-1}^l, \\ \{0\} & \text{when } \sigma_{J_o}^l + 1 < \sigma_{J_o-1}^l \end{cases}} \ \sum_{\hat{s}_\pi^l \in \hat{S}_\pi^{f,l}(J_o,j_3,j_4)} \ \sum_{w \in B_{n;k_0}^1(s^f,\hat{s}_\pi^l)} q(w) + \delta_{fl} * \sum_{w \in \{w \in B_{n;k_0} |\ w_i \leq s_i,\ i=1,...,n\}} q(w) \tag{4.3.14}$$

follows according to Definition 4.3.8, Theorem 4.3.10 and Theorem 4.3.11.

In the case of discrete uniformly distributed requirements, Definition 4.3.8 yields:

$$\sum_{w\in B^1_{n;k_0}(s^f,\hat{s}^l_\pi)} q(w) \;=\; \tfrac{1}{(k_0+1)^n}\Bigg[\prod_{i:\hat{s}^l_{\pi_i}=s^f_i\le\sigma^l_{J_o}} (\hat{s}^l_{\pi_i}+1)$$

$$\left((\sigma^l_{J_o}+2)^{j_3+j_4} - \binom{j_3+j_4}{j_4+1}(\sigma^l_{J_o}+1)^{j_3-1} - \cdots - \binom{j_3+j_4}{j_3+j_4}(\sigma^l_{J_o}+1)^0\right)\Bigg],$$

$$\delta_{fl} \;*\; \sum_{w\in\{w\in B_{n;k_0}\,|\,w_i\le s_i,\; i=1,\ldots,n\}} q(w) \;=\; \delta_{fl} \;*\; \tfrac{1}{(k_0+1)^n}\prod_{i=1}^{n}(s^f_i+1). \tag{4.3.15}$$

**Theorem 4.3.12.** *Elements $p^*_{fl}$ of PRMs can be calculated by:*

$$p^*_{fl} = p^{*1}_{fl} + p^{*2}_{fl} - p^{*1,2}_{fl},$$

*where $p^{*1}_{fl}$ is computed as in (4.3.13) or in (4.3.14), $p^{*2}_{fl}$ as in (4.3.9) or in (4.3.10) and* $p^{*1,2}_{fl} = \sum_{s^l_\pi:\ permutations\ of\ s^l} q(s^l_\pi)$.

*In the case of discrete uniformly distributed requirements (4.3.15) and (4.3.11) can be used in particular to compute elements $p^*_{fl}$.*

## 4.4 Limits of Partitions-Requirements-Matrices

In this section limits of elements of PRMs are determined as $n\to\infty$ and $k_0\to\infty$ if the numbers of rows and columns of the corresponding PRMs are fixed. (Only limits as $su\to\infty$ are not possible since $su$ is limited by $n\,k_0$, see (4.1.2).)

At first, we classify the sets of (restricted) partitions in order to find classes with an infinite number of sets $S_{n;su;k_0}$ where all sets $S_{n;su;k_0}$ have the same number of partitions so that the corresponding PRMs all have the same numbers of rows and columns. We use such classes in this section and in Section 4.6.

### 4.4.1 Classification of the Lattices of the Restricted Partitions

Let the sets of restricted partitions be arranged in a 3-dimensional lattice with respect to $n, su, k_0$ with

$$n = 2, 3, \cdots ;\ su = 1, 2, \cdots ;\ k_0 = \lceil \frac{su}{n} \rceil, \lceil \frac{su}{n} \rceil + 1, \cdots , su. \tag{4.4.1}$$

(See Figure 4.4.1). The arrangement in relation to $k_0$ (dimension 3) is also found in the plane:

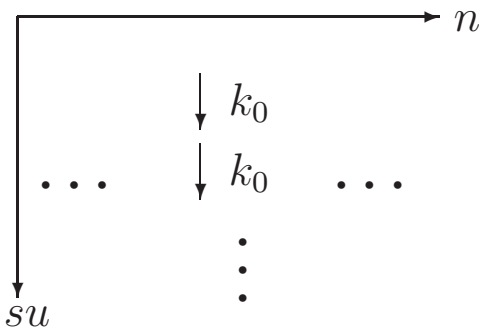
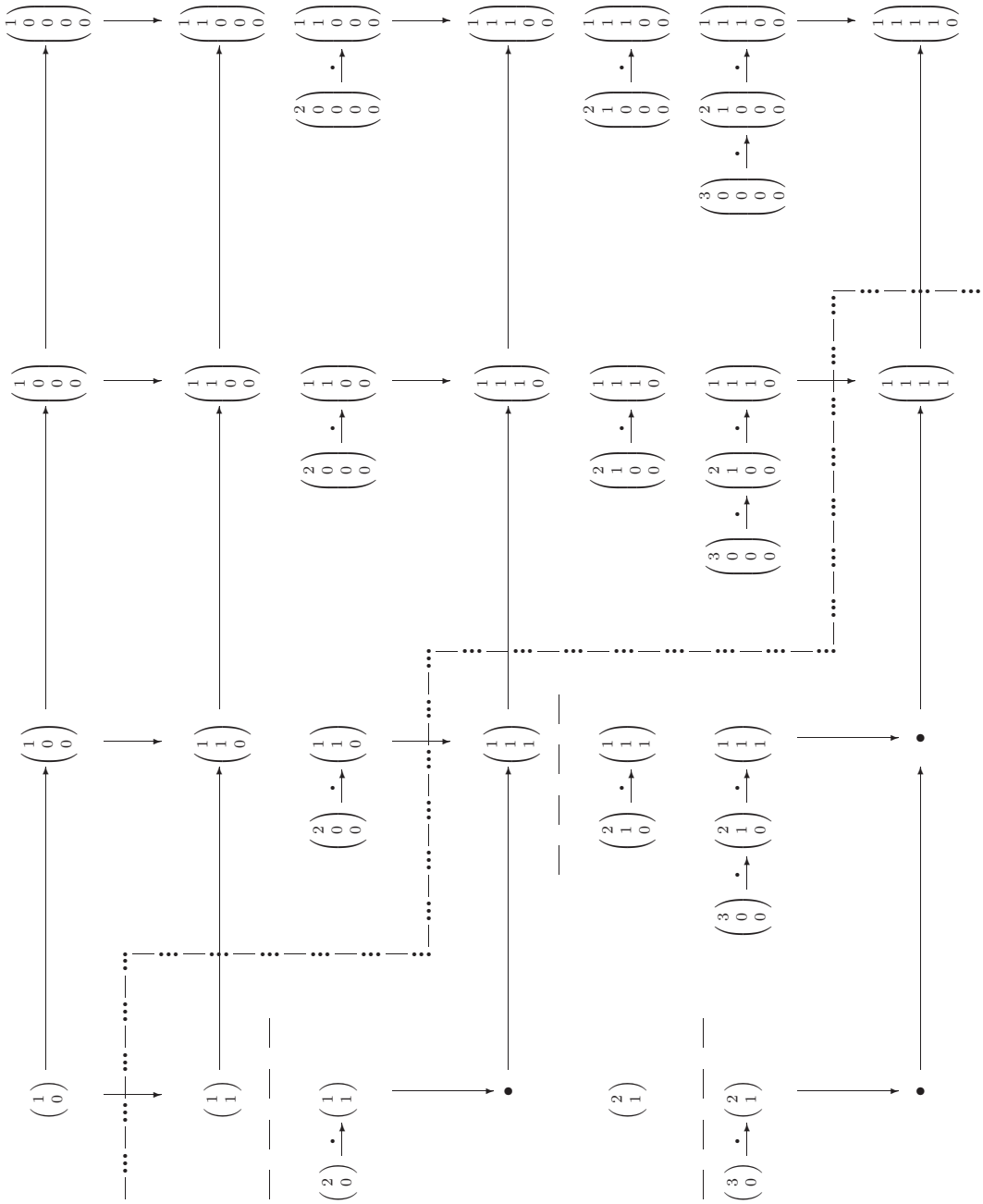
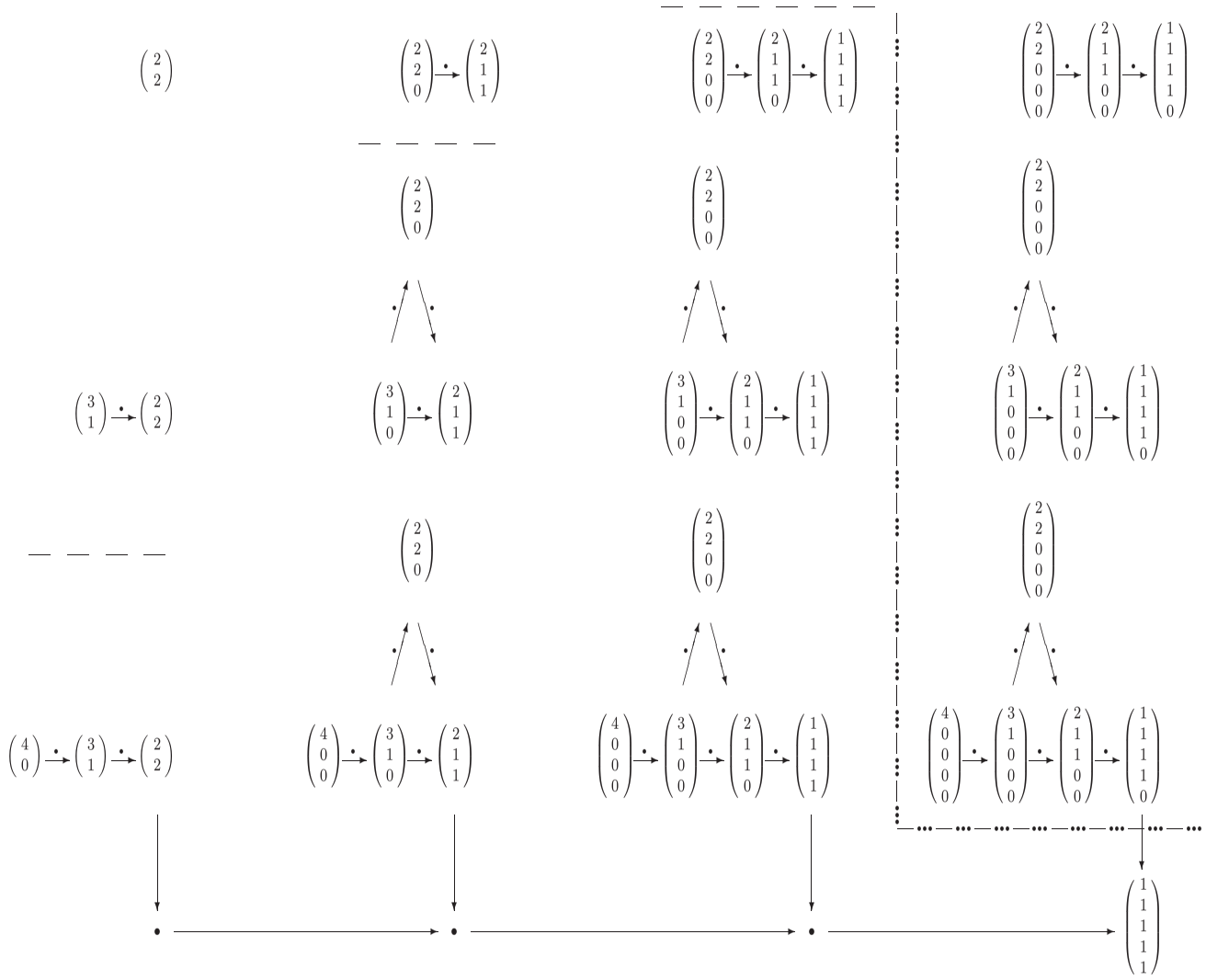


**Figure 4.4.1.** .

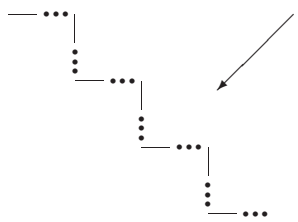



$\begin{pmatrix}2\\2\\0\\0\\0\end{pmatrix} \to \begin{pmatrix}2\\1\\1\\0\\0\end{pmatrix} \to \begin{pmatrix}1\\1\\1\\1\\0\end{pmatrix}$
$\begin{pmatrix}2\\2\\0\\0\\0\end{pmatrix}$
$\begin{pmatrix}3\\1\\0\\0\\0\end{pmatrix} \to \begin{pmatrix}2\\1\\1\\0\\0\end{pmatrix} \to \begin{pmatrix}1\\1\\1\\1\\0\end{pmatrix}$
$\begin{pmatrix}2\\2\\0\\0\\0\end{pmatrix}$
$\begin{pmatrix}4\\0\\0\\0\\0\end{pmatrix} \to \begin{pmatrix}3\\1\\0\\0\\0\end{pmatrix} \to \begin{pmatrix}2\\1\\1\\0\\0\end{pmatrix} \to \begin{pmatrix}1\\1\\1\\1\\0\end{pmatrix}$
$\begin{pmatrix}1\\1\\1\\1\\1\end{pmatrix}$
$\begin{pmatrix}2\\2\\0\\0\\0\end{pmatrix} \to \begin{pmatrix}2\\1\\1\\0\\0\end{pmatrix} \to \begin{pmatrix}1\\1\\1\\1\\1\end{pmatrix}$
$\begin{pmatrix}2\\2\\0\\0\end{pmatrix}$
$\begin{pmatrix}3\\1\\0\\0\end{pmatrix} \to \begin{pmatrix}2\\1\\1\\0\end{pmatrix} \to \begin{pmatrix}1\\1\\1\\1\end{pmatrix}$
$\begin{pmatrix}2\\2\\0\\0\end{pmatrix}$
$\begin{pmatrix}4\\0\\0\\0\end{pmatrix} \to \begin{pmatrix}3\\1\\0\\0\end{pmatrix} \to \begin{pmatrix}2\\1\\1\\0\end{pmatrix} \to \begin{pmatrix}1\\1\\1\\1\end{pmatrix}$
$\begin{pmatrix}2\\2\\0\end{pmatrix} \to \begin{pmatrix}2\\1\\1\end{pmatrix}$
$\begin{pmatrix}2\\2\\0\end{pmatrix}$
$\begin{pmatrix}3\\1\\0\end{pmatrix} \to \begin{pmatrix}2\\1\\1\end{pmatrix}$
$\begin{pmatrix}2\\2\\0\end{pmatrix}$
$\begin{pmatrix}4\\0\\0\\0\end{pmatrix} \to \begin{pmatrix}3\\1\\0\\0\end{pmatrix} \to \begin{pmatrix}2\\1\\1\\0\end{pmatrix}$
$\begin{pmatrix}2\\2\end{pmatrix}$
$\begin{pmatrix}3\\1\end{pmatrix} \to \begin{pmatrix}2\\2\end{pmatrix}$
$\begin{pmatrix}4\\0\end{pmatrix} \to \begin{pmatrix}3\\1\end{pmatrix} \to \begin{pmatrix}2\\2\end{pmatrix}$

$$\begin{tikzcd}
 & & & \begin{pmatrix}2\\2\\1\\0\\0\end{pmatrix} \arrow[r, "\bullet"] & \begin{pmatrix}2\\1\\1\\1\\0\end{pmatrix} \arrow[r, "\bullet"] & \begin{pmatrix}1\\1\\1\\1\\1\end{pmatrix} \\
 & & \begin{pmatrix}2\\2\\1\end{pmatrix} & \begin{pmatrix}2\\2\\1\\0\end{pmatrix} \arrow[r, "\bullet"] & \begin{pmatrix}2\\1\\1\\1\end{pmatrix} & \\
 & \begin{pmatrix}3\\1\\1\end{pmatrix} & & \begin{pmatrix}3\\1\\1\\0\end{pmatrix} & & \begin{pmatrix}3\\1\\1\\0\\0\end{pmatrix} \\
 & \bullet \nearrow \quad \searrow \bullet & & \bullet \nearrow \quad \searrow \bullet \quad \searrow \bullet & & \bullet \nearrow \quad \searrow \bullet \quad \searrow \bullet \\
\begin{pmatrix}3\\2\end{pmatrix} & \begin{pmatrix}3\\2\\0\end{pmatrix} \arrow[r, "\bullet"] & \begin{pmatrix}2\\2\\1\end{pmatrix} & \begin{pmatrix}3\\2\\0\\0\end{pmatrix} \arrow[r, "\bullet"] & \begin{pmatrix}2\\2\\1\\0\end{pmatrix} \arrow[r, "\bullet"] & \begin{pmatrix}2\\1\\1\\1\end{pmatrix} & \begin{pmatrix}3\\2\\0\\0\\0\end{pmatrix} \arrow[r, "\bullet"] & \begin{pmatrix}2\\2\\1\\0\\0\end{pmatrix} \arrow[r, "\bullet"] & \begin{pmatrix}2\\1\\1\\1\\0\end{pmatrix} \arrow[r, "\bullet"] & \begin{pmatrix}1\\1\\1\\1\\1\end{pmatrix} \\
 & \begin{pmatrix}3\\2\\0\end{pmatrix} & & \begin{pmatrix}3\\2\\0\\0\end{pmatrix} \arrow[r, "\bullet"] & \begin{pmatrix}2\\2\\1\\0\end{pmatrix} & & \begin{pmatrix}3\\2\\0\\0\\0\end{pmatrix} \arrow[r, "\bullet"] & \begin{pmatrix}2\\2\\1\\0\\0\end{pmatrix} \\
 & \bullet \nearrow \quad \searrow \bullet \quad \searrow \bullet & & \bullet \nearrow \quad \searrow \bullet \quad \bullet \nearrow \quad \searrow \bullet & & \bullet \nearrow \quad \searrow \bullet \quad \bullet \nearrow \quad \searrow \bullet \\
\begin{pmatrix}4\\1\end{pmatrix} \arrow[r, "\bullet"] & \begin{pmatrix}3\\2\end{pmatrix} & \begin{pmatrix}4\\1\\0\end{pmatrix} \arrow[r, "\bullet"] & \begin{pmatrix}3\\1\\1\end{pmatrix} \arrow[r, "\bullet"] & \begin{pmatrix}2\\2\\1\end{pmatrix} & \begin{pmatrix}4\\1\\0\\0\end{pmatrix} \arrow[r, "\bullet"] & \begin{pmatrix}3\\1\\1\\0\end{pmatrix} \arrow[r, "\bullet"] & \begin{pmatrix}2\\1\\1\\1\end{pmatrix} & \begin{pmatrix}4\\1\\0\\0\\0\end{pmatrix} \arrow[r, "\bullet"] & \begin{pmatrix}3\\1\\1\\0\\0\end{pmatrix} \arrow[r, "\bullet"] & \begin{pmatrix}2\\1\\1\\1\\0\end{pmatrix} \arrow[r, "\bullet"] & \begin{pmatrix}1\\1\\1\\1\\1\end{pmatrix} \\
 & \begin{pmatrix}3\\2\\0\end{pmatrix} & & \begin{pmatrix}3\\2\\0\\0\end{pmatrix} \arrow[r, "\bullet"] & \begin{pmatrix}2\\2\\1\\0\end{pmatrix} & & \begin{pmatrix}3\\2\\0\\0\\0\end{pmatrix} \arrow[r, "\bullet"] & \begin{pmatrix}2\\2\\1\\0\\0\end{pmatrix} \\
 & \bullet \nearrow \quad \searrow \bullet \quad \searrow \bullet & & \bullet \nearrow \quad \searrow \bullet \quad \bullet \nearrow \quad \searrow \bullet & & \bullet \nearrow \quad \searrow \bullet \quad \bullet \nearrow \quad \searrow \bullet \\
\begin{pmatrix}5\\0\end{pmatrix} \arrow[r, "\bullet"] & \begin{pmatrix}4\\1\end{pmatrix} \arrow[r, "\bullet"] & \begin{pmatrix}3\\2\end{pmatrix} & \begin{pmatrix}5\\0\\0\\0\end{pmatrix} \arrow[r, "\bullet"] & \begin{pmatrix}4\\1\\0\\0\end{pmatrix} \arrow[r, "\bullet"] & \begin{pmatrix}3\\1\\1\\0\end{pmatrix} \arrow[r, "\bullet"] & \begin{pmatrix}2\\1\\1\\1\end{pmatrix} & \begin{pmatrix}5\\0\\0\\0\\0\end{pmatrix} \arrow[r, "\bullet"] & \begin{pmatrix}4\\1\\0\\0\\0\end{pmatrix} \arrow[r, "\bullet"] & \begin{pmatrix}3\\1\\1\\0\\0\end{pmatrix} \arrow[r, "\bullet"] & \begin{pmatrix}2\\1\\1\\1\\0\end{pmatrix} \arrow[r, "\bullet"] & \begin{pmatrix}1\\1\\1\\1\\1\end{pmatrix}
\end{tikzcd}$$

We now classify the sets (lattices) of partitions in the following way:
(See also the remarks at the end of this section.)

C1) Sets of **"sparse" partitions**: $S_{n;su;k_0}$ with $n > su$.



(See Figure 4.1: These sets can be found above the staircase-line.
In general, many 0's are in the vectors of these sets.
In this case $2 \geq \lceil \frac{2su}{n} \rceil$ follows. Hence $k_0 \geq \lceil \frac{2su}{n} \rceil$ is valid for $2 \leq k_0$.)

C2) Sets: $S_{n;su;k_0}$ with $n \leq su$ and $k_0 \geq \lceil \frac{2su}{n} \rceil$.

(See Figure 4.1: These sets can be found in each case below the interrupted line $(\overset{-----}{\nearrow})$.)

Only a **finite number** of such sets with any fixed number r of partitions exists! (See the following Theorem 4.4.4.)

C3) Sets: $S_{n;su;k_0}$ with $n \leq su$ and $k_0 < \lceil \frac{2su}{n} \rceil$.

(See Figure 4.1: These sets can be found in each case above the interrupted line $(\_\_\_\overset{\swarrow}{\_}\_\_\_)$ and below the stair-case-line.)
These sets can be further classified as follows:

C3a) Sets: $S_{n;su;k_0}$ with $n \leq su,\ k_0 < \lceil \frac{2su}{n} \rceil$ **and** $\exists\, i : s_i^1 = 0$ (where $s^1$ is the least element of $S_{n;su;k_0}$).
These sets are, with regard to the Poisson equation, **equivalent** to sets of sparse partitions (see C1)) or to sets from C2), see Definition 4.2.5 and the following Lemma 4.4.2(ii), (iii).

C3b) Sets of **"heavy" partitions**:
$S_{n;su;k_0}$ with $n \leq su,\ k_0 < \lceil \frac{2su}{n} \rceil$ **and** $s_i^1 > 0$ for $i = 1, \cdots, n$.

(An equivalent characterization can be found in the following Lemma 4.4.1.)

These sets are subsequently subclassified as follows:

C3b.1) Sets of **"non-truncated" heavy partitions**:
$S_{n;su;k_0}$ with $n \leq su,\ k_0 < \lceil \frac{2su}{n} \rceil,\ s_i^1 > 0$ for $i = 1, \cdots, n$

$$\textbf{and } s^r = \begin{pmatrix} k_0 \\ \vdots \\ k_0 \\ k_0 - 1 \\ \vdots \\ k_0 - 1 \end{pmatrix} \text{ (that is, } \exists\, i : s_i^r = k_0)$$

(where $s^r$ is the greatest element of $S_{n;su;k_0}$).

C3b.2) Sets of **"truncated" heavy partitions**:
$S_{n;su;k_0}$ with $n \leq su,\ k_0 < \lceil \frac{2su}{n} \rceil,\ s_i^1 > 0$ for $i = 1, \cdots, n$ **and** $s_i^r < k_0$ for $i = 1, \cdots, n$.

(Example: $S_{2;2k_0-3;k_0}$ with $k_0 > 3$.)

In the last example $su$ is given as $2k_0 - 3$. It can also be useful, in general, to present $su$ for sets of heavy partitions as $su = nk_0 - \overline{su}$ where

$$0 < \overline{su} \leq nk_0 - k_0 = (n-1)k_0 \tag{4.4.2}$$

according to (4.1.2):

**Lemma 4.4.1.** *Let $S_{n;su=nk_0-\overline{su};k_0}$ be a set of (restricted) partitions.*

*(a) $S_{n;su;k_0}$ is a set of heavy partitions if and only if $\overline{su} < k_0$.*

*(b) $S_{n;su;k_0}$ is a set of non-truncated heavy partitions if and only if $\overline{su} < k_0$ and $\overline{su} < n$.*

*(c) $S_{n;su;k_0}$ is a set of truncated heavy partitions if and only if $\overline{su} < k_0$ and $\overline{su} \geq n$.*

***Proof.*** We initially consider the first two inequalities from C3b):
$su = nk_0 - \overline{su}$ implies that $n \leq su$ is equivalent to

$$\overline{su} \leq n(k_0 - 1) \tag{*1}$$

and $k_0 < \lceil \frac{2su}{n} \rceil$ to

$$k_0 < \lceil \frac{2(nk_0 - \overline{su})}{n} \rceil = 2k_0 - \lfloor \frac{2\overline{su}}{n} \rfloor,$$

and $\lfloor \frac{2\overline{su}}{n} \rfloor < k_0.$ (*2)

(a) $s_i^1 > 0$ for $i = 1, \cdots, n$ from C3b) is equivalent to $\lceil \frac{su}{k_0} \rceil \geq n$

and furthermore (together with $su = nk_0 - \overline{su}$) to:

$$\lceil \frac{nk_0 - \overline{su}}{k_0} \rceil \geq n,$$

$$\lceil n - \frac{\overline{su}}{k_0} \rceil \geq n,$$

$$n - \lfloor \frac{\overline{su}}{k_0} \rfloor \geq n,$$

$$0 \geq \lfloor \frac{\overline{su}}{k_0} \rfloor,$$

$$\overline{su} < k_0. \qquad (*3)$$

If we conversely suppose that (*3) is valid, then

(*1) and (*2) clearly (for $n \geq 2$) follow.

(b) The condition: $\exists\, i :\ s_i^r = k_0$ from C3b.1) is equivalent to $n(k_0 - 1) < su$ and furthermore to

$$\overline{su} < n.$$

(c) The condition: $s_i^r < k_0$ for $i = 1, \cdots, n$ from C3b.2) is equivalent to $n(k_0 - 1) \geq su$ and additionally to

$$\overline{su} \geq n.$$

■

Now, we will prove (see (iii) from the following Lemma) that sets of partitions $S_{n;su;k_0}$ which satisfy the conditions from C3a) are equivalent with regard to the Poisson equation to sets of sparse partitions (see C1)) or to sets which satisfy the conditions from C2).

For this purpose we will use the denotation $S_{n;su;k_0}$ for sets which satisfy the conditions from C3a) and then show that corresponding sets $S_{n;\overline{su};k_0}$ with $\overline{su} = n \cdot k_0 - su$ fulfil the conditions from C1) and C2) according to Definition 4.2.5(c).

**Lemma 4.4.2.** *Let $S_{n;su;k_0}$ with $n \leq su$ and $k_0 < \lceil \frac{2su}{n} \rceil$ (which also means $2su > n \cdot k_0$) be given (see classification C3)). In addition let $\overline{su} = n \cdot k_0 - su$.*

*(i) Then $S_{n;su;k_0}$ and $S_{n;\overline{su};\overline{k}_0}$ with $\overline{k}_0 = k_0 - \max\{0, su-(n-1)k_0\}$ are equivalent with respect to the partial order (see Definition 4.2.5(b)) and the inequality $\overline{k}_0 \geq \lceil \frac{2\overline{su}}{n} \rceil$ is correct.*

*(ii) Furthermore, the relationship*
$\overline{k_0} = k_0 \Leftrightarrow \exists\, i : s_i^1 = 0$
*in relation to the least element $s^1$ of $S_{n;su;k_0}$ is valid.*
*If $\overline{k_0} = k_0$, then $S_{n;su;k_0}$ and $S_{n;\overline{su};k_0}$ are also equivalent with regard to the Poisson equation.*

*(iii) If $S_{n;su;k_0}$ satisfies the conditions from C3a), then $S_{n;\overline{su};k_0}$ is a set from C1) or C2).*

***Proof.***

(i) Initially, we note that
$$\overline{k}_0 = k_0 - \max\{0, su-(n-1)k_0\} \leq k_0.$$
We set $k_c = k_0$ (see Def. 4.2.5(b))
and $\bar{s} = (k_0, ..., k_0)^T - s$ for $s \in S_{n;su;k_0}$.
Then
$$\sum_{i=1}^{n} \bar{s}_i = \sum_{i=1}^{n} (k_0 - s_i) = nk_0 - su = \bar{su}$$
is valid and also
$$\bar{s}_i = k_0 - s_i \leq k_0 - \max\{0, su-(n-1)k_0\} = \bar{k}_0$$
since $s_i \geq \max\{0, su-(n-1)k_0\}$ for $s \in S_{n;su;k_0}$ and $i \in \{1, 2, \cdots, n\}$.

Thus, $\bar{s} \in S_{n;\overline{su};\overline{k}_0}$.

Analogously $s = ((k_0, ..., k_0)^T - \bar{s}) \in S_{n;su;k_0}$ is correct for $\bar{s} \in S_{n;\overline{su};\overline{k}_0}$.

Furthermore, $\bar{s} = (k_0, ..., k_0)^T - s$ and $s = (k_0, ..., k_0)^T - \bar{s} \in S_{n;su;k_0}$ are one-to-one maps such that $\mid S_{n;su;k_0} \mid = \mid S_{n;\overline{su},\overline{k}_0} \mid$ follows.

Hence $S_{n;su;k_0}$ and $S_{n;\overline{su};\overline{k}_0}$ with $\overline{su} = n \cdot k_0 - su$ are equivalent with respect to the partial order.

Now, we show the inequality $\overline{k}_0 \geq \lceil \frac{2\overline{su}}{n} \rceil$ :
$n \leq su$ and $k_0 < \lceil \frac{2su}{n} \rceil$ (which means $2su > n \cdot k_0$) imply

$$\bar{su} = n \cdot k_0 - su < 2su - su = su$$

at first and additionally

$$\bar{k}_0 = \begin{cases} k_0 & \text{if } su-(n-1)k_0 \le 0, \\ nk_0 - su & \text{if } su-(n-1)k_0 \ge 0 \end{cases} = \begin{cases} k_0 \text{ if } su-(n-1)k_0 \le 0, \\ \bar{su} \text{ if } su-(n-1)k_0 \ge 0 \end{cases}$$

$$\ge \begin{cases} \lceil \frac{\bar{su}+nk_0-su}{n} \rceil & \text{if } su-(n-1)k_0 \le 0, \\ \bar{su} & \text{if } su-(n-1)k_0 \ge 0 \end{cases} \ge \lceil \frac{2\bar{su}}{n} \rceil.$$

(ii) $\bar{k}_0 = k_0$ is equivalent to $su-(n-1)k_0 \le 0$

since $\bar{k}_0 = k_0 - \max\{0, su-(n-1)k_0\}$.

$su-(n-1)k_0 \le 0$ if and only if the least element $s^1$ of $S_{n;su;k_0}$ (see Definition 4.1.1(b)) has at least one part $s_i^1 = 0$ and vice versa .

(iii) follows from $\bar{k}_0 = k_0 \ge \lceil \frac{2\bar{su}}{n} \rceil$ (see (i)) together with (ii). ■

We will use the simple facts from the following lemma in the subsequent discussions.

**Lemma 4.4.3.** *Let $n$ and $su$ be given. Then*

*(i)* $|S_{n;su;k_0}| < |S_{n;su;k_0+1}|$ *for* $k_0 = \lceil \frac{su}{n} \rceil, \lceil \frac{su}{n} \rceil + 1, \cdots, su-1$ *and*

*(ii)* $|S_{n;su;k_0}| < |S_{n+1;su;k_0}|$ *for* $n = 1, 2, \cdots, su-1$ .

(Note: In dependence on the variable $su$ (if $n; k_0$ are given) it seems that the number $|S_{n;su;k_0}|$ initially increases and decreases again later. See [3], Theorem 3.10: $|S_{n;su;k_0}| \ge |S_{n;su-1;k_0}|$ for $0 < su \le \frac{nk_0}{2}$.)

***Proof.***

(i) Obviously, $s \in S_{n;su;k_0+1}$ follows for partitions $s \in S_{n;su;k_0}$.

Furthermore, $S_{n;su;k_0+1}$ includes partitions (with components of the value $k_0+1$) which are not elements of $S_{n;su;k_0}$.

(ii) $(s_1, \ldots, s_n) \in S_{n;su;k_0}$ imply that $(s_1, \ldots, s_n, 0) \in S_{n+1;su;k_0}$.

Furthermore, all parts of the greatest element of $S_{n+1;su;k_0}$ ($n \in \{1, 2, \cdots, su-1\}$)(see Definition 4.1.1(d)) are nonzero. Thus, this greatest element can not be included in $S_{n;su;k_0}$. ■

The following Theorem implies that only a **finite number** of sets which satisfy the conditions from C2) **with any fixed** number $r$ of partitions exists.

**Theorem 4.4.4.** *For a given $\rho \in \mathbb{Z}_+$ a $su_0$ exists so that $\mid S_{n;su;k_0} \mid > \rho$ for all $S_{n;su;k_0}$ with $su > su_0$ and $(su \geq)\ k_0 \geq \lceil \frac{2su}{n} \rceil$.*

We use $su_0 = 6(\rho - 1)$ for the following proof in the case $n \geq 3$. A sharp bound would most likely be $su_0 = 2(\rho - 1)$.

***Proof.*** Initially, $n = 2$ which implies $k_0 = su$ and $\mid S_{2;su;su} \mid = su + 1$ is a simple case and $su_0 = \rho - 1$ satisfies the demand of the theorem.

Now let $su_0 = 6(\rho - 1)$.

We first show that $|S_{n;su;k_0=\lceil \frac{2su}{n} \rceil}| > \rho$ for $n$ with $2 < n \leq su$.

$|S_{n;su;k_0'}| > \rho$ for $k_0' > k_0$ or/and $n > su$ follow from Lemma 4.4.3i) and the obvious relation $|S_{su;su;k_0}| = |S_{n;su;k_0}|$ for $n > su$.

$$\text{The partition } s^1 = \begin{pmatrix} s_1^1 \\ \vdots \\ s^1_{\lceil \frac{n}{2} \rceil - 1} \\ s^1_{\lceil \frac{n}{2} \rceil} \\ s^1_{\lceil \frac{n}{2} \rceil + 1} \\ \vdots \\ s_n^1 \end{pmatrix} = \begin{pmatrix} \lceil \frac{2su}{n} \rceil \\ \vdots \\ \lceil \frac{2su}{n} \rceil \\ su - \lceil \frac{2su}{n} \rceil (\lceil \frac{n}{2} \rceil - 1) \\ 0 \\ \vdots \\ 0 \end{pmatrix} \begin{matrix} \left.\vphantom{\begin{matrix} a \\ a \\ a \end{matrix}}\right\} \lceil \frac{n}{2} \rceil - 1 \\ \left.\vphantom{\begin{matrix} a \\ a \\ a \end{matrix}}\right\} \lfloor \frac{n}{2} \rfloor \end{matrix}$$

is the least element (see Definition 4.1.1(d)) of $S_{n;su;k_0}$ with $k_0 = \lceil \frac{2su}{n} \rceil$.

There are also, at least, the following partitions, each of which are elements of $S_{n;su;k_0}$ with $k_0 = \lceil \frac{2su}{n} \rceil$:

Uniformly 1's are initially subtracted from the first $\lceil \frac{n}{2} \rceil - 1$ (or $\lceil \frac{n}{2} \rceil$) upper components of the least element and with the value $k_0 = \lceil \frac{2su}{n} \rceil$ and one by one the remaining $\lceil \frac{n}{2} \rceil$ lower components are increased (with these 1's) until equality with the upper components is reached. The procedure is continued analogously up to the greatest element. This demonstrated here in detail:

$$\begin{pmatrix} s_1^1-\mathbf{1} \\ s_2^1 \\ \vdots \\ s^1_{\lceil \frac{n}{2}\rceil-1} \\ s^1_{\lceil \frac{n}{2}\rceil}+\mathbf{1} \\ s^1_{\lceil \frac{n}{2}\rceil+1}=0 \\ \vdots \\ s_n^1=0 \end{pmatrix}, \begin{pmatrix} s_1^1-\mathbf{1} \\ s_2^1-\mathbf{1} \\ \vdots \\ s^1_{\lceil \frac{n}{2}\rceil-1} \\ s^1_{\lceil \frac{n}{2}\rceil}+\mathbf{2} \\ s^1_{\lceil \frac{n}{2}\rceil+1}=0 \\ \vdots \\ s_n^1=0 \end{pmatrix}, \dots, \begin{pmatrix} s_1^{l_2} \\ s_2^{l_2} \\ \vdots \\ s^{l_2}_{\lceil \frac{n}{2}\rceil-1} \\ s^{l_2}_{\lceil \frac{n}{2}\rceil} \\ s^{l_2}_{\lceil \frac{n}{2}\rceil+1}=0 \\ \vdots \\ s_n^{l_2}=0 \end{pmatrix},$$

with $s_i^{l_2} = \lfloor \frac{su}{\lceil \frac{n}{2}\rceil} \rfloor$ or $s_i^{l_2} = \lfloor \frac{su}{\lceil \frac{n}{2}\rceil} \rfloor + 1$ for $i = 1, 2, ..., \lceil \frac{n}{2} \rceil$ and (without loss of generality) $s_1^{l_2} \geq s_2^{l_2} \geq \cdots \geq s^{l_2}_{\lceil \frac{n}{2}\rceil} > 0$,

or

$$\begin{pmatrix} s_1^{l_2}-\mathbf{1} \\ s_2^{l_2} \\ \vdots \\ s^{l_2}_{\lceil \frac{n}{2}\rceil} \\ s^{l_2}_{\lceil \frac{n}{2}\rceil+1}=\mathbf{1} \\ s^{l_2}_{\lceil \frac{n}{2}\rceil+2}=0 \\ \vdots \\ s_n^{l_2}=0 \end{pmatrix}, \begin{pmatrix} s_1^{l_2}-\mathbf{1} \\ s_2^{l_2}-\mathbf{1} \\ \vdots \\ s^{l_2}_{\lceil \frac{n}{2}\rceil} \\ s^{l_2}_{\lceil \frac{n}{2}\rceil+1}=\mathbf{2} \\ s^{l_2}_{\lceil \frac{n}{2}\rceil+2}=0 \\ \vdots \\ s_n^{l_2}=0 \end{pmatrix}, \dots, \begin{pmatrix} s_1^{l_3} \\ s_2^{l_3} \\ \vdots \\ s^{l_3}_{\lceil \frac{n}{2}\rceil} \\ s^{l_3}_{\lceil \frac{n}{2}\rceil+1} \\ s^{l_3}_{\lceil \frac{n}{2}\rceil+2}=0 \\ \vdots \\ s_n^{l_3}=0 \end{pmatrix},$$

with $s_i^{l_3} = \lfloor \frac{su}{\lceil \frac{n}{2}\rceil+1} \rfloor$ or $s_i^{l_3} = \lfloor \frac{su}{\lceil \frac{n}{2}\rceil+1} \rfloor + 1$ for $i = 1, 2, ..., \lceil \frac{n}{2} \rceil + 1$

and so on.

There are (with $s^1$) at least

$$\rho' = 1+\left(\lfloor \frac{su}{\lceil \frac{n}{2}\rceil} \rfloor - \left(su - \lceil \frac{2su}{n} \rceil \left(\lceil \frac{n}{2} \rceil - 1\right)\right)\right)+\lfloor \frac{su}{\lceil \frac{n}{2}\rceil + 1} \rfloor+\lfloor \frac{su}{\lceil \frac{n}{2}\rceil + 2} \rfloor+\cdots+\lfloor \frac{su}{n} \rfloor$$

partitions at least. A rough estimation yields

$$\rho' \geq 1 + \lfloor \frac{su}{n} \rfloor \cdot \lfloor \frac{n}{2} \rfloor \geq 1 + \lfloor \frac{su}{n} \rfloor \frac{n-1}{2}. \qquad (*1)$$

Now, let $j \in \mathbb{N}$ so that

$$j \cdot n \leq su < (j+1)n.$$

($j \geq 1$ is valid since $n \leq su$ are assumed above.)

$$\frac{su}{n} \geq j \text{ and } su \cdot \frac{1}{2}\frac{j}{j+1} \cdot \frac{n-1}{n} < \frac{n-1}{2} \cdot j \tag{*2}$$

follow.

Combining (*1) and (*2) we get

$$\rho' \geq 1 + j\frac{n-1}{2} > 1 + \frac{su}{2} \cdot \frac{j}{j+1} \cdot \frac{n-1}{n} \geq \frac{su}{6} + 1 \text{ for } n > 2.$$

Thus,

$$|S_{n;su;k_0=\lceil \frac{2su}{n} \rceil}| \geq \rho' > \frac{su}{6} + 1 > \frac{su_0}{6} + 1 = \rho$$

for $su_0 = 6(\rho - 1)$.

$$|S_{n;su;k_0}| \geq \rho' > \frac{su}{6} + 1 > \frac{su_0}{6} + 1 = \rho$$

follows for $su > su_0$ and $k_0 \geq \lceil \frac{2su}{n} \rceil$ according to Lemma 4.4.3i) (and since $|S_{su;su;k_0}| = |S_{n';su;k_0}|$ for $n' > su$).

■

**Remarks 4.4.1.** *According to Lemma 4.4.2(iii) and Theorem 4.4.4 sets of*

- *sparse partitions (see C1)),*
- *non-truncated heavy partitions (see C3b.1)) and*
- *truncated heavy partitions (see C3b.2))*

*are of interest when dealing with the computation of limits of PRMs.*

*These sets include an* **infinite** *number of sets of restricted partitions* **with arbitrary but fixed number r of partitions**.

*This means that the corresponding PRMs are all of the same size, $r$ by $r$.*

### 4.4.2 Limits of Partitions-Requirements-Matrices with regard to Sets of Sparse Partitions

Sets of sparse partitions with $k_0 = 1$ have only one partition. As this is a trivial situation, the solution is straightforward. If $k_0 \geq 2$ then $k_0 \geq \lceil \frac{2su}{n} \rceil$ (see classification C1)). We can then use Theorem 4.4.4.

An infinite number of sets of sparse partitions with a given fixed number of partitions r is obtained only when $n \to \infty$ since, obviously, $|S_{su;su;k_0}| = |S_{n;su;k_0}|$ for $n \geq su$ and according to Theorem 4.4.4.

Let us begin by considering an example:

**Example 4.4.1.** *PRMs with regard to sets of partitions*

$$S_{n;4;3} = \left\{ \begin{pmatrix} 3 \\ 1 \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 2 \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \\ 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \right\} \ (n > 4) \ \textit{and with}$$

*discrete uniformly distributed requirements are:*

$$P^*(n) = \frac{1}{4^n} * \begin{pmatrix} \mathbf{4^{n-1}} + n^2 - n + 6 & 2n^2 - 4n + 4 & \mathbf{4^{n-1}} + \frac{13}{3}n^3 - 21n^2 + \frac{110}{3}n - 28 & \mathbf{2 \cdot 4^{n-1}} - \frac{13}{3}n^3 + 18n^2 - \frac{95}{3}n + 18 \\ n^2 & \mathbf{4^{n-1}} + 2n^2 - 2n + 4 & \mathbf{2 \cdot 4^{n-1}} + \frac{13}{3}n^3 - 30n^2 + \frac{200}{3}n - 56 & \mathbf{4^{n-1}} - \frac{13}{3}n^3 + 27n^2 - \frac{194}{3}n + 52 \\ n^2 & 2n^2 - 2n & \mathbf{2 \cdot 4^{n-1}} + \frac{13}{3}n^3 - 21n^2 + \frac{89}{3}n - 6 & \mathbf{2 \cdot 4^{n-1}} - \frac{13}{3}n^3 + 18n^2 - \frac{83}{3}n + 6 \\ n^2 & 2n^2 - 2n & \frac{13}{3}n^3 - 12n^2 + \frac{26}{3}n & \mathbf{4^n} - \frac{13}{3}n^3 + 9n^2 - \frac{20}{3}n \end{pmatrix}.$$

*The enumeration of $P^*(n)$ is carried out laboriously and it yields:*

$$\lim_{n\to\infty} P^*(n) = \frac{1}{4}\begin{pmatrix} 1 & 0 & 1 & 2 \\ 0 & 1 & 2 & 1 \\ 0 & 0 & 2 & 2 \\ 0 & 0 & 0 & 4 \end{pmatrix}.$$

In this section we compute limits of PRMs $P^*(n)$ with regard to sets of sparse partitions as $n \to \infty$ in the general case.

This will yield triangular matrices with supplementary zeroes (see also $\lim_{n\to\infty} p^*_{12}(n) = 0$, Example 4.4.1).

With this in mind, we introduce the concepts of monotone successors, and principle parts of partitions.

**Definition 4.4.1.** .

(a) *Let $s \in S_{n;su;k_0}$ be a partition with $n \underset{(=)}{>} su$ and (w. l. o. g.) $s_1 \geq s_2 \geq \cdots \geq s_n$.*
*Then $s_H = (s_1, \cdots, s_\eta)$ with $s_i > 1$ for $i = 1, \cdots, \eta$ and $s_i \in \{0, 1\}$ for $i = \eta + 1, \cdots, n$ is called the* **principal part** *of $s$.*

(b) *Let $s^l_H$ be the principal parts of the partitions $s^l \in S_{n_l;su_l;k_{0l}}$ (where $n_l \underset{(=)}{>} su_l$, $s^l_1 \geq s^l_2 \geq \cdots \geq s^l_{n_l}$) for $l = 1, 2$.*
*Then $s^2$ is called a* **monotone successor** *of $s^1$ if $\eta_1 \geq \eta_2$ and $s^1_i \geq s^2_i$ for $i = 1, \cdots, \eta_2(\eta_1)$.*

**Lemma 4.4.5.** .

(a) *If $\{s^f, s^l\} \subseteq S_{n;su;k_0}$ (where $n \underset{(=)}{>} su$) and $s^l$ is a monotone successor of $s^f$, then $s^l$ is also a successor of $s^f$ (see Definition 4.1.1).*

(b) *Let $s'$ be any monotone successor of $s$, $s \in S_{n;su;k_0}$ and $s' \in S_{n';su';k_{0'}}$.*
*Then $s'' \in S_{n;su;k_0}$ exists so that the principal parts of $s'$ and $s''$ are equal. (This means $s''$ is a monotone successor of $s$.)*

(The proof is obvious.)

**Theorem 4.4.6.** *(Limits of PRMs with regard to sets of sparse partitions)*

*Let* $n \underset{(=)}{>} su$ *and* $\{s^f(n), s^l(n)\} \subseteq S_{n;su;k_0}$ *with (w. l. o. g.)* $s^\lambda(n)_1 \geq s^\lambda(n)_2 \geq \cdots \geq s^\lambda(n)_n$ *for* $\lambda = f, l$.

*(This means, in particular,* $s^\lambda(n)_{su+1} = s^\lambda(n)_{su+2} = \cdots = s^\lambda(n)_n = 0$ *for* $\lambda = f, l$*.)*

*Furthermore, let* $s^\lambda := (s^\lambda(n)_1, \cdots, s^\lambda(n)_{su})^T \in S_{su;su;k_0}$ *for* $\lambda = f, l$. *Then,*

$$\lim_{n\to\infty} p^*(s^l(n)|s^f(n))$$

$$= \begin{cases} 0 & \text{if } s^l \text{ is not a monotone successor of } s^f, \\ (q_0(0)+q_0(1))^{\eta_f-\eta_l} \sum\limits_{s^l_\pi \in S^l_{\eta_f}} \prod\limits_{i:s^f_i > s^l_{\pi_i} \geq 2} q_0(s^l_{\pi_i}) \prod\limits_{i:s^f_i = s^l_{\pi_i} \geq 2} (q_0(s^f_i) + \cdots + q_0(k_0)) & \\ & \text{if } s^l \text{ is a monotone successor of } s^f \end{cases}$$

*where* $\eta_f$ *and* $\eta_l$ *are defined as in Definition 4.4.1(b) and*

$$S^l_{\eta_f} = \Big\{ s_\pi \in \mathbb{Z}^{\eta_f}_+ \mid s_\pi \text{ is a permutation of } (s^l_1, s^l_2, \cdots, s^l_{\eta_l}, 0, \cdots, 0)^T \in \mathbb{Z}^{\eta_f}_+ \text{ with } s^f_i \geq s_{\pi_i} \text{ for } i = 1, \cdots, \eta_f \Big\}.$$

***Proof.*** Let us denote

$$p^*(s^l(n)|s^f(n)) = \sum_{w:s^*(s^f(n),w)=s^l(n)} q(w) = \sum_{w\in B^1} q(w) + \sum_{w\in B^2} q(w)$$

where

$B^1 = \{w \in B_{n;k_0} \mid C[w; su] \ \wedge \ s^*(s^f(n), w) = s^l(n)\}$
and
$B^2 = \{w \in B_{n;k_0} \mid C[su; w] \ \wedge \ s^*(s^f(n), w) = s^l(n)\}$.

Now, we consider $w \in B^1$ :

C [w; su] implies that at least $(n - su)$ coordinates of $w$ are equal to 0. The simple inequality relation

$$\sum_{w \in B^1} q(w) \leq \binom{n}{su} q_0(0)^{n-su} \tag{*1}$$

follows and furthermore

$$\lim_{n\to\infty} \sum_{w \in B^1} q(w) \leq \lim_{n\to\infty} \binom{n}{su} q_0(0)^{n-su} = 0 \tag{*2}$$

since $\binom{n}{su}$ is a polynomial (of the degree $su$) and $q_0(0) < 1$ (see (4.2.2)).

Thus,

$\lim\limits_{n\to\infty} p^*(s^l(n) \mid s^f(n)) = \lim\limits_{n\to\infty} \sum\limits_{w \in B^2} q(w)$

remains to be considered:

**Case 1:** Let $s^l$ not be a monotone successor of $s^f$.
This means $\exists i : s^l(n)_i > s^f(n)_i \ \wedge \ s^l(n)_i > 1$.

Thus, at least $(n - su)$ coordinates of $w \in B^2$ must be equal to 0 according to the iterative method from Section 4.2.2, case C[su, w]. Otherwise in certain iteration steps 0's would needlessly be increased to 1 and the above $s^l(n)_i$ could then not be obtained by the iterative method.

Analogously to (*1) and (*2)

$\lim\limits_{n\to\infty} p^*(s^l(n)|s^f(n)) = \lim\limits_{n\to\infty} \sum\limits_{w \in B^2} q(w) = 0$

follows in case 1.

**Case 2:** Let $s^l$ be a monotone successor of $s^f$.
We partition $B^2 = B^{2a} \cup B^{2b}$ where

$$B^{2a} = \{w \in B^2 \mid \forall \text{ permutations } s^l_\pi(n) \text{ of } s^l(n) \text{ satisfying (4.2.5)} \\ \exists i : \ s^l_\pi(n)_i > s^f(n)_i \ \wedge \ s^l_\pi(n)_i > 1\},$$

$$B^{2b} = \{w \in B^2 \mid \exists \text{ a permutation } s^l_\pi(n) \text{ of } s^l(n) \text{ satisfying (4.2.5)} : \\ s^l_\pi(n)_i \leq s^f(n)_i \text{ for } i = 1, \cdots, \eta_f \text{ and} \\ s^l_\pi(n)_i \in \{0, 1\} \text{ for } i = \eta_f + 1, \cdots, n\}.$$

We see at first, analogous to case 1, that $\lim\limits_{n\to\infty} \sum\limits_{w \in B^{2a}} q(w) = 0$ follows.

Finally, let $s^l_\pi(n)$ be a permutation of $s^l(n)$ with $s^l_\pi(n)_i \leq s^f(n)_i$ for $i = 1, \cdots, \eta_f$ and $s^l_\pi(n)_i \in \{0, 1\}$ for $i = \eta_f + 1, \cdots, n$.

Then $s^*(s^f(n), w) = s^l(n)$, where $s^l_\pi(n)$ together with $w$ and $s^f(n)$ fullfil (4.2.5), is only valid for $w \in B_{n;k_0}$ with

$$\left.\begin{array}{lll} s^f_i \le w_i \le k_0 & if & s^l_\pi(n)_i = s^f_i (= s^f(n)_i) \ge 2, \\ w_i = s^l_\pi(n)_i & if & s^f_i > s^l_\pi(n)_i \ge 2, \\ w_i \in \{0,1\} & if & s^f_i \ge 2 \wedge s^l_\pi(n)_i \in \{0,1\}, \\ 0 \le w_i \le k_0 & if & i > \eta_f \end{array}\right\} \tag{4.4.3}$$

where, additionally, $\sum\limits_{i=1}^{n} w_i \ge su$.

If $w$ fullfils (4.4.3) however $\sum\limits_{i=1}^{n} w_i < su$ that implies that at least $(n - su)$ coordinates of $w$ are equal to 0. Analogous to (*1) and (*2) the relationship

$$\lim_{n\to\infty} \sum_{w:\sum\limits_i w_i < su} q(w) \le \lim_{n\to\infty} \binom{n}{su} q_0(0)^{n-su} = 0 \tag{*3}$$

follows.

If $s^l_\pi(n)$ corresponds to $w$ as in $B^{2b}$ then, obviously,

$$s^l_\pi \in S^l_{\eta_f}, \text{ where } s^l_{\pi_i} = \begin{cases} s^l_\pi(n)_i & \text{if } s^l_\pi(n)_i \ge 2, \\ 0 & \text{otherwise} \end{cases} \quad \text{for } i = 1, \cdots, \eta_f.$$

Different $s^l_\pi$ imply different $w$ which together with the corresponding $s^l_\pi(n)$ satisfy (4.4.3). This in conjunction with (*3) yields

$$\begin{aligned} &\lim_{n\to\infty} \sum_{w\in B^{2b}} q(w) \\ &= \sum_{s^l_\pi \in S^l_{\eta_f}} (q_0(0)+q_0(1))^{\eta_f - \eta_l} \prod_{i: s^f_i > s^l_{\pi_i} \ge 2} q_0(s^l_{\pi_i}) \prod_{i: s^f_i = s^l_{\pi i} \ge 2} (q_0(s^f_i) + \cdots + q_0(k_0)) \\ &= \lim_{n\to\infty} p^*(s^l(n)|s^f(n)). \end{aligned}$$

■

(See also Example 4.4.1.)

**Corollary 4.4.7.** *Let $su$ and $k_0$ ($su \ge k_0$) be given and let the numbering of the partitions of the sets $S_{n;su;k_0} = \{s^1(n), \cdots, s^r(n)\}$, $n \underset{(=)}{>} su$ be so that $\{s^f(n) \to s^l(n)\} \Rightarrow \{f < l\}$.*

*Then the matrix $\lim\limits_{n\to\infty} P^*(n)$ is a triangular matrix.*

### 4.4.3 Limits of Partitions-Requirements-Matrices with regard to Sets of Non-Truncated Heavy Partitions

Sets of non-truncated heavy partitions can be represented in the form of $S_{n;nk_0-\bar{su};k_0}$ with $\bar{su} < k_0, \bar{su} < n, (n \geq 2, k_0 \geq 2)$ (see Lemma 4.4.1(b)). In particular, this means that we use the representation

$$su = nk_0 - \bar{su} \tag{4.4.4}$$

for $su$.

**Example 4.4.2.** *Example sets of non-truncated heavy partitions with one or two partitions are*

$S_{n;nk_0-1;k_0},\ n \geq 2, k_0 \geq 2$ *and* $S_{n;nk_0-2;k_0},\ n \geq 3, k_0 \geq 3.$

*In greater detail,*

*(a)* $S_{n;nk_0-1;k_0},\ n \geq 2, k_0 \geq 2$

$$\left\{\begin{pmatrix} 2 \\ 1 \end{pmatrix}\right\}, \cdots, \left\{\begin{pmatrix} k_0 & \\ k_0 & -1 \end{pmatrix}\right\}, \cdots$$

$$\vdots \qquad\qquad \vdots$$

$$\left\{\begin{pmatrix} 2 \\ \vdots \\ 2 \\ 1 \end{pmatrix}\right\}, \cdots, \left\{\begin{pmatrix} k_0 & \\ \vdots & \\ k_0 & \\ k_0 & -1 \end{pmatrix}\right\}, \cdots$$

$$\vdots \qquad\qquad \vdots$$

*(b)* $(S_{n;nk_0-2,k_0}, \rightarrow),\ n \geq 3, k_0 \geq 3$

$$\left\{\begin{pmatrix} 3 \\ 3 \\ 1 \end{pmatrix} \rightarrow \begin{pmatrix} 3 \\ 2 \\ 2 \end{pmatrix}\right\}, \cdots, \left\{\begin{pmatrix} k_0 & \\ k_0 & \\ k_0 & -2 \end{pmatrix} \rightarrow \begin{pmatrix} k_0 & \\ k_0 & -1 \\ k_0 & -1 \end{pmatrix}\right\}, \cdots$$

$$\vdots \qquad\qquad \vdots$$

$$\left\{\begin{pmatrix} 3 \\ \vdots \\ 3 \\ 3 \\ 1 \end{pmatrix} \rightarrow \begin{pmatrix} 3 \\ \vdots \\ 3 \\ 2 \\ 2 \end{pmatrix}\right\}, \quad \cdots, \quad \left\{\begin{pmatrix} k_0 & \\ \vdots & \\ k_0 & \\ k_0 & \\ k_0 & -2 \end{pmatrix} \rightarrow \begin{pmatrix} k_0 & \\ \vdots & \\ k_0 & \\ k_0 & -1 \\ k_0 & -1 \end{pmatrix}\right\}, \cdots$$

$\vdots \qquad\qquad\qquad \vdots$

**Lemma 4.4.8.** *Sets of non-truncated heavy partitions*

$S_{n;nk_0-\bar{su};k_0}$ $(\bar{su} < k_0, \bar{su} < n, n \geq 2, k_0 \geq 2)$ *are equivalent to sets of sparse partitions* $S_{n;\bar{su};\bar{k_0}}$ *with* $\bar{k_0} = \bar{su}$ *with respect to to the partial order.*

*A corresponding relation between the partitions is*

$$\bar{s} = \begin{pmatrix} k_0 \\ \vdots \\ k_0 \end{pmatrix} - s, \ s \in S_{n;su;k_0}(su = nk_0 - \bar{su}), \ \bar{s} \in S_{n;\bar{su};\bar{k_0}}.$$

*Furthermore, the least element of* $S_{n;\bar{su};\bar{k_0}}$ *is* $(\bar{su}, 0, \ldots, 0)^T$.

The proof is obvious. (See also Lemma 4.4.2.)

In Example 4.4.2 $\left\{\begin{pmatrix} 2 \\ 1 \end{pmatrix}\right\}, \cdots, \left\{\begin{pmatrix} k_0 \\ k_0 - 1 \end{pmatrix}\right\}$ are equivalent to $\left\{\begin{pmatrix} 1 \\ 0 \end{pmatrix}\right\}$ and so on.

Lemma 4.4.8 and Theorem 4.4.4 lead to the following remarks:

**Remarks 4.4.2.** *An infinite number of sets of non-truncated heavy partitions with a given fixed number of partitions is yielded by* $k_0 \to \infty$ *or* $n \to \infty$.

*(See also Example 4.4.2.)*

**The case** $n \to \infty$

can be discussed analogously to Section 4.4.2 if we pay attention to the equivalence from Lemma 4.4.8!

Here we consider **the case** $k_0 \to \infty$.

This requires additional properties of the probability functions $q$: Let $q^{k_0}$ denote probability functions corresponding to $B_{n;k_0}$ (where $n$ is fixed). Then we assume

$$\lim_{k_0 \to \infty} q_0^{k_0}(w_i) = 0 \text{ for } w_i = 0, 1, \ldots \tag{4.4.5}$$

and

$$\begin{aligned}&\exists\, c(k_0) \text{ (real numbers with } 1 > c(k_0) > 0) :\\ &\exists \lim_{k_0\to\infty} \tfrac{q_0^{k_0}(w_i)}{c(k_0)} \neq 0 \text{ for } w_i = 0, 1, \ldots .\end{aligned} \tag{4.4.6}$$

$$(\lim_{k_0\to\infty} \frac{(q_0^{k_0}(w_i))^l}{c(k_0)} = 0 \text{ for } l \geq 2 \text{ and } w_i = 0, 1, \ldots . \tag{4.4.7}$$

follows from (4.4.5) and (4.4.6).)

(The limits $\lim_{k_0\to\infty} \frac{q_0^{k_0}(\cdot)}{c(k_0)}$ are unique, however can differ by a constant multiple (in relation to $c(k_0)$).)

**Definition 4.4.2.** .

(a) *Let* $s \in S_{n;su;k_0}$ *be a non-truncated heavy partition with (w. l. o. g.)* $s_1 \geq s_2 \geq \cdots \geq s_n$.

*Then* $s_H = (s_\eta, \cdots, s_n)$ *with* $s_i < k_0 - 1$ *for* $i = \eta, \cdots, n$ *and* $s_i \in \{k_0, k_0 - 1\}$ *for* $i = 1, \cdots, \eta - 1$ *is called the* **principal part** *of* $s$.

(b) *Let* $s_H^l$ *be the principal parts of non-truncated heavy partitions* $s^l \in S_{n_l;su_l;k_{0_l}}$ *(where* $s_1^l \geq s_2^l \geq \cdots \geq s_{n_l}^l$*) for* $l = 1, 2$.

*Then* $s^2$ *is called a* **restricted monotone successor** *of* $s^1$ *if* $n_1 - \eta_1 \geq n_2 - \eta_2$, *and* $s_{n_1-i}^1 \leq s_{n_2-i}^2$ *for* $i = 0, \cdots, (n_1 - \eta_1)n_2 - \eta_2$ *and*

$$\begin{aligned}&\textit{a permutation } s_\pi^2 \textit{ of } s^2 \textit{ exists such that}\\ &s_{n_1-i}^1 \neq s_{\pi(n_2-i)}^2 \textit{ for } \textbf{at most} \textit{ one } i \in \{0, \cdots, n_1 - \eta_1\}.\end{aligned} \tag{4.4.8}$$

*(This also implies* $n_1 - \eta_1 = n_2 - \eta_2$ *or* $n_1 - \eta_1 = n_2 - \eta_2 + 1$.*)*

**Remarks 4.4.3.** *We cannot derive Definition 4.4.2(b) completely from Definition 4.4.1(b) by using the equivalence of sets with non-truncated heavy partitions and certain sets with sparse partitions (see Lemma 4.4.8) because of condition (4.4.8)!*

**Lemma 4.4.9.** .

(a) *Let* $S_{n;su;k_0}$ *be a set of non-truncated heavy partitions and* $\{s^f, s^l\} \subseteq S_{n;su;k_0}$. *If* $s^l$ *is a restricted monotone successor of* $s^f$ *then* $s^l$ *is also a successor of* $s^f$ *(see Definition 4.1.1).*

(b) *Let* $S_{n;su;k_0}, S_{n',su',k_0'}$ *be sets of non-truncated heavy partitions and* $s \in S_{n;su;k_0}$, $s' \in S_{n';su';k_0'}$.

*If* $s'$ *is a restricted monotone successor of* $s$, *then* $s'' \in S_{n;su;k_0}$ *exists such that the principal parts of* $s'$ *and* $s''$ *are equal.*

*(This means,* $s''$ *is a restricted monotone successor of* $s$*.)*

(The proof is obvious.)

**Theorem 4.4.10.** *(Limits of PRMs with regard to sets of non-truncated heavy partitions)*

*Let* $S_{n;su;k_0}$ *be sets of non-truncated heavy partition where* $su$ *is represented by* $su = nk_0 - \bar{su}$ *(see Lemma 4.4.2) with fixed* $n$ *and* $\bar{su}$ *(and* $n \geq 2,\ k_0 \geq 2$*).*

*Furthermore, let* $\{s^f(k_0), s^l(k_0)\} \subseteq S_{n;su;k_0}$ *with (w. l. o. g.)* $s^\lambda(k_0)_1 \geq s^\lambda(k_0)_2 \geq \cdots \geq s^\lambda(k_0)_n$ *for* $\lambda = f; l$.

*(In particular, that means* $s^\lambda(k_0)_1 = \cdots = s^\lambda(k_0)_{n-\bar{su}} = k_0$ *for* $\lambda = f; l$*.)*

*In addition let* $s^\lambda = (s_1^\lambda, \cdots, s_{\bar{su}}^\lambda)^T := (s^\lambda(k_0)_{n-\bar{su}+1}, \cdots, s^\lambda(k_0)_n)^T \in S_{\bar{su},su,k_0}$ *(where* $\sum_i s_i^\lambda = \bar{su}k_0 - \bar{su}$*).*

*Finally, let given probability functions* $q^{k_0}$ *fulfill (4.4.5) and (4.4.6) for certain* $c(k_0)$ *and*

$$q_0^0(w_i) := \lim_{k_0 \to \infty} \frac{q_0^{k_0}(w_i)}{c(k_0)}.$$

*Then,*

$$\lim_{k_0 \to \infty} \frac{1}{c(k_0)} (p^*(s^l(k_0)|s^f(k_0)) - \delta(s^l(k_0), s^f(k_0)))$$

$$= \begin{cases} 0 & \textit{if } s^l \textit{ is not a restricted monotone successor of } s^f, \\ -\sum\limits_{i:s_i^f \le k_0-2} (q_0^0(s_i^f+1)+\cdots+q_0^0(k_0)) & \textit{if } s^f = s^l, \\ \sum\limits_{s_\pi^l \in S_{\eta_f}^l} |\{j|s_j^f = s_{i_0}^f < s_{\pi_{i_0}}^l| \begin{cases} q_0^0(s_{\pi_{i_0}}^l) & \textit{for } s_{\pi_{i_0}}^l < k_0-1, \\ (q_0^0(k_0-1)+q_0^0(k_0)) & \textit{for } s_{\pi_{i_0}}^l = k_0-1 \end{cases} \\ \qquad \textit{if } s^l \textit{ is a restricted monotone successor of } s^f \textit{ and } s^l \ne s^f \end{cases}$$

*where* $\delta(s^l(k_0), s^f(k_0)) = \begin{cases} 1 & \textit{if } s^f(k_0) = s^l(k_0) \\ 0 & \textit{if } s^f(k_0) \ne s^l(k_0) \end{cases}$,

$$S_{\eta_f}^l = \Big\{ s_\pi \in \mathbb{Z}_+^{n-\eta_f+1} \mid s_\pi \textit{ is a permutation}$$
$$\textit{of} \begin{cases} (s_{\eta_l}^l, \cdots, s_n^l)^T & \textit{if } \eta_f = \eta_l, \\ (k_0-1, s_{\eta_l}^l, \cdots, s_n^l)^T & \textit{if } \eta_f = \eta_l - 1 \end{cases} \textit{with } s_i^f \le s_{\pi_i} \textit{ for } i = \eta_f, \cdots, n \Big\}$$

*and* $\eta_f$, $\eta_l$ *are as in Definition 4.4.2(b).*

Formally, the proof is analogous to the proof of Theorem 4.4.6, however the roles of the cases $C[w, su]$ and $C[su, w]$ are reversed. In detail the considerations are somewhat different.

We demonstrate this for $\sum\limits_{w \in B^2} q^{k_0}(w)$

(where $B^2 = \{w \in B_{n;k_0} \mid C[su; w] \ \wedge \ s^*(s^f(k_0), w) = s^l(k_0)\}$).

(4.2.5), case C[su;w] implies that an $s_\pi^l(k_0)$, a permutation of $s^l(k_0)$, exists with $w \ge s_\pi^l(k_0)$. Hence, at least one coordinate of $w$ is equal to $k_0$ and at least one is either also equal to $k_0$ or is equivalent to $k_0 - 1$ (since $s_\pi^l(k_0)$ is a non-truncated heavy partition).

The simple inequality relation

$$\tfrac{1}{c(k_0)} \sum_{w \in B^2} q^{k_0}(w) \le \tfrac{1}{c(k_0)} \binom{n}{2} (q_0^{k_0}(k_0) + q_0^{k_0}(k_0-1))^2$$

follows and furthermore (4.4.7) yields

[13] This is a finite number of i since $\bar{su}$ is fixed.

$$\sum_{w\in B^2} q_0^0(w) \le \lim_{k_0\to\infty} \tfrac{1}{c(k_0)} \begin{pmatrix} n \\ 2 \end{pmatrix} (q_0^{k_0}(k_0) + q_0^{k_0}(k_0-1))^2 = 0.$$

**Corollary 4.4.11.** *Let sets of non-truncated heavy partition* $S_{n;su;k_0}$ *as in Theorem 4.4.10 be given and let the numbering of the partitions of the sets* $S_{n;su;k_0} = \{s^1(k_0), \cdots, s^r(k_0)\}$ *be such that* $\{s^f(k_0) \to s^l(k_0)\} \Rightarrow \{f < l\}$.

*Then, the matrix* $\lim\limits_{k_0\to\infty} \frac{1}{c(k_0)}(P^*(k_0) - I)$ *is a triangular matrix.*

**Example 4.4.3.** *Let us begin with*

$S_{n;4;4}$ $(n > 4)$ – *sets of sparse partitions and*
$S_{5;5k_0-4;k_0}$ $(k_0 > 4)$ – *sets of non-curtailed heavy partitions.*

*Such sets are equivalent with respect to the partial order (see Lemma 4.4.8).*

*Now, we compare some limits of elements of the corresponding PRMs:*

$$\lim_{n\to\infty} p^*(s^3(n)|s^3(n)) - 1 = (q_0(2) + \cdots + q_0(k_0))^2 - 1 = (1 - q_0(0) - q_0(1))^2 - 1$$

$$(s^3(n) = (2, 2, 0, \cdots, 0)^T;\ \textit{see Theorem 4.4.6})$$

*with*

$$\lim_{k_0\to\infty} \tfrac{1}{c(k_0)}(p^*(s^3(k_0)|s^3(k_0)) - 1) = -2(q_0^0(k_0) + q_0^0(k_0-1))$$

$$(s^3(k_0) = (k_0, k_0, k_0, k_0-2, k_0-2)^T;\ \textit{see Theorem 4.4.10))}$$

*and*

$$\lim_{n\to\infty} p^*(s^4(n)|s^3(n)) = (q_0(0) + q_0(1))\ 2\ (q_0(2) + \cdots + q_0(k_0))$$

$$= 2(q_0(0) + q_0(1))(1 - q_0(0) - q_0(1))$$

$$(s^4(n) = (2, 1, 1, 0, \cdots, 0)^T;\ \textit{see Theorem 4.4.6})$$

*with*

$$\lim_{k_0\to\infty} \tfrac{1}{c(k_0)} p^*(s^4(k_0)|s^3(k_0)) = 2(q_0^0(k_0-1) + q_0^0(k_0))$$

$$(s^4(k_0) = (k_0, k_0, k_0-1, k_0-1, k_0-2)^T,\ \textit{see Theorem}\ 4.4.10)).$$

*These limits as well as the formulas for limits of PRMs with regard to sets of sparse partitions (Theorems 4.4.6) and of PRMs with regard to sets of*

*non-truncated heavy partitions (Theorems 4.4.10) are very different. In contrast to that, solutions of Poisson equations with regard to limits of corresponding PRMs will have similar structures, see the Theorems 4.6.8 and 4.6.12 in Section 4.6.)*

## 4.5 Further Results from Elements of Partitions-Requirements-Matrices

PRMs depend on the variables $n$, $su$ and $k_0$. Of these, it seems most difficult to find formulas which place the dependence on $su$.

In Section 4.5.1 we will show, by means of the concept of perturbed partitions, that with regard to the variables $n$ and $k_0$, elements of PRMs are sums of probabilities over subsets of requirements whose numbers of elements are described by polynomials in $k_0$ and/or either polynomials or sums of exponential functions and polynomials in $n$.

Hence, the elements of PRMs themseves (multiplied by $(k_0+1)^n$) are polynomials or sums of such exponential functions and polynomials in the case of discrete uniformly distributed requirements. (See also Example 4.4.1.)

Thereby the exponential functions can be determined similarly to the limits of elements of PRMs in Section 4.4. The determination of the polynomials however appears more difficult. The degree of such polynomials and the corresponding leading terms are given in Section 4.5.1.

Remarks on the methods of the corresponding proofs follow, in particular for the use of the concept of disturbed partitions.

Finally, in Section 4.5.2 the elements of the last row and the last column of PRMs (in the case that $n \geq su$) are computed.

In this section we also use the terminology from Section 4.3.

### 4.5.1 Elements of PRMs in Dependence on Variables $n$ and $k_0$

**Theorem 4.5.1.** *Let* $\bar{s}^f \in S_{\bar{n};su;k_0}$ *and* $\bar{s}^l \in S_{\bar{n};su;k_0}$ *be given with (w. l. o. g.)* $\bar{s}_1^f \geq \bar{s}_2^f \geq \cdots \geq \bar{s}_{\bar{n}}^f$ *and* $\bar{s}_1^l \geq \bar{s}_2^l \geq \cdots \geq \bar{s}_{\bar{n}}^l$ *and where* $\bar{n} \geq \max\{F, L\}$ *(F, L as in (4.3.1)) and (su* $\geq$*)* $k_0 \geq \max\{s_{max}^f, s_{max}^l\}$, $s_{max}^f = \bar{s}_1^f, s_{max}^l = \bar{s}_1^l$.

*Furthermore, let* $s^\lambda = (\bar{s}_1^\lambda, \cdots, \bar{s}_\Lambda^\lambda, 0, \cdots, 0) \in S_{n;su;k_0}$ *for* $\lambda = f, l$, $\Lambda = F, L$ *and any* $n \geq \max\{F, L\}$.

*Then the following statements are valid:*

*(i) The number of elements in the sets*

$B_{n;k_0}^{*(f,l)} := B_{n;k_0}^*(s^f, s^l) = \{w \in B_{n;k_0} \mid s^l = s^*(s^f, w)\}$ *(see Definition 4.2.3(b))*

*(with* $n \geq \max\{F, L\}$ *and* $su \geq k_0 \geq \max\{s_{max}^f, s_{max}^l\}$*)*

*is a polynomial in* $k_0$ *and/or either a polynomial or a sum of an exponential function and a polynomial in* $n$.

*(ii) In the case of discrete uniformly distributed requirements, the elements of the corresponding PRMs multiplied by* $(k_0+1)^n$ *are polynomials in* $k_0$ *and/or either polynomials or sums of exponential functions and polynomials in* $n$.

*The degree of the polynomial part of* $(k_0+1)^n * p_{fl}^*$, *as function of* $n$, *is* $L$ *(the number of parts of* $s^l$ *which are not equal to 0) for* $s^l \neq (1, ..., 1, 0, ..., 0)$ *and* $su - 1$ *for* $s^l = (1, ..., 1, 0, ..., 0)$.

*The coefficient of the leading term* [14] *of the polynomial in the case that* $s^l \neq (1, ..., 1, 0, ..., 0)$ *is:*

$$\frac{1}{L!} * \begin{pmatrix} L \\ L_2^* \end{pmatrix} * C^* * \left[ \sum_{j=L_1}^{L_2^*} \begin{pmatrix} L_2^* \\ j \end{pmatrix} (k_0 + 1 - s_{max}^l)^j \right]$$

*with* $s_{max}^l := s_1^l = ... = s_{L_1}^l$ *(*$L_1$ *as in (4.3.3)) the maximum of the components of* $s^l$,

$$L_2^* = \begin{cases} L_2 & \text{if } s_{L_1}^l = s_{L_2}^l + 1, \\ L_1 & \text{otherwise} \end{cases}$$

[14] See also [18].

*and $C^*$ is the number of permutations of the $L - L_2^*$ components of $s^l$ which are smaller than $s^l_{max} - 1$.*

*In the case that $s^l = (1, ..., 1, 0, ..., 0)$, the coefficient of the leading term of the polynomial is:*

$-\frac{1}{(su-1)!} * (k_0^{su-1} - 1)$.

*For $n \underset{(=)}{>} su$ all corresponding PRMs are of the same type and the polynomial parts of the elements multiplied by $(k_0 + 1)^n$ as functions of $n$, which are in the same column, must have the same leading term.*

***Proof.*** We use "perturbed partitions" (see Section 4.3) in the following proof.

The results with respect to $k_0$ follow directly from Section 4.3:

Thereby we can confine ourselves to the case $C[su, w]$, since $k_0$ does not play a role in the case $C[w, su]$.

With regard to determination of corresponding sets of perturbed partitions, the possibilities of $(J_o, j_o, j_1)$ with $\hat{S}^{f,l}_\pi(J_o, j_o, j_1) \neq \varnothing$ and the elements of $\hat{S}^{f,l}_\pi(J_o, j_o, j_1)$ depend directly on $s^f$ and $s^l$ however not on $k_0$. (See Definitions 4.3.1 and 4.3.3 and also Lemma 4.3.1.)

Only sets $B^2_{n;k_0}(s^f, \hat{s}^l_\pi)$ depend on $k_0$ (see Definition 4.3.4). That $|B^2_{n;k_0}(s^f, \hat{s}^l_\pi)|$ are polynomials in $k_0$ follows from (4.3.11) and also that the elements of the corresponding PRMs multiplied by $(k_0+1)^n$ are polynomials in $k_0$ in the case of discrete uniformly distributed requirements.

In relation to variable $n$ we also use perturbed partitions for the proof.

The formulas (4.3.9) and (4.3.10) or (4.3.13) and (4.3.14) are therefore not directly suitable for a proof since in these formulas the sums $\cdots \sum_{\hat{s}^l_\pi \in \hat{S}^{f,l}_\pi(J_o, j_o, j_1)} \cdots$ or $\cdots \sum_{\hat{s}^l_\pi \in \hat{S}^{f,l}_\pi(J_o, j_3, j_4)} \cdots$ are included and the numbers of elements $\hat{s}^l_\pi$ in the sets $\hat{S}^{f,l}_\pi(J_o, j_o, j_1)$ and $\hat{S}^{f,l}_\pi(J_o, j_3, j_4)$ can depend on $n$.

For this reason we will construct certain subsets of $\hat{S}^{f,l}_\pi(J_o, j_o, j_1)$ and $\hat{S}^{f,l}_\pi(J_o, j_3, j_4)$.

Furthermore, we will have to consider four cases, or five cases if $s^f = s^l$.

For this proof we also define that products over sets are equal to 1 if the

sets are empty!

**Case**: In $C[su;w]$ a $(J_o, j_o)$-perturbation of the relation "$\geq$" between $s^f$ and $s^l_\pi$ is considered. (see Definition 4.3.1)

Furthermore, let $\hat{s}^l$ be a $(J_o, j_o)$-perturbed partition of $s^l$ (see Definition 4.3.2).

**Subcase:** $\sigma^l_{J_o} > 1$. **($\sigma^l_{J_o}$ as from Definition 4.3.1)**

We split non-empty sets $\hat{S}^{f,l}_\pi(J_o, j_o, j_1)$ (see Definition 4.3.3) into the disjunct subsets:

$\hat{S}^{2,(f,l)}_{\pi_\xi}(J_o, j_o, j_1), \ \ \xi = 1, \cdots, \Xi(= \Xi(J_o, j_o, j_1))$

where $\hat{S}^{2,(f,l)}_{\pi_\xi}$ include elements $\hat{s}^l_\pi$ with certain fixed components $\hat{s}^l_{\pi_i}$. In greater detail:

$L_{J_o} - j_o$ components $\hat{s}^l_{\pi_i} \geq \sigma^l_{J_o}$ of $\hat{s}^l_\pi$ are fixed for $i$ where

$$\hat{s}^l_i \geq \sigma^l_{J_o} \text{ so that } s^f_i \geq \hat{s}^l_{\pi_i} \tag{4.5.1}$$

(see (d10) in Lemma 4.3.1),

and if $\sigma^l_{J_o} - 1 = \sigma^l_{J_o+1}$ then in addition $L_{J_o+1} - L_{J_o} - j_1$ components $\hat{s}^l_{\pi_i} = \sigma^l_{J_o} - 1$ of $\hat{s}^l_\pi$ are fixed for $i$ where

$$\hat{s}^l_i = \sigma^l_{J_o} - 1 \ \text{ so that } s^f_i > \hat{s}^l_{\pi_i} \tag{4.5.2}$$

(see (d11) in Lemma 4.3.1).

Furthermore, we use the symbols

$$\begin{array}{l} I^{2,(f,l)}_{\pi_\xi}(J_o, j_o, j_1) \ \text{ for the set of indices } i \text{ for which} \\ \text{the components } \hat{s}^l_{\pi_i} \text{ of } \hat{s}^l_\pi \text{ are fixed above,} \end{array} \tag{4.5.3}$$

$$n1 = n1(J_o, j_o, j_1) = \begin{cases} L_{J_o+1} - j_o - j_1 & \text{if } \sigma^l_{J_o} - 1 = \sigma^l_{J_o+1}, \\ L_{J_o} - j_o & \text{if } \sigma^l_{J_o} - 1 > \sigma^l_{J_o+1}, \end{cases} \tag{4.5.4}$$

$$n2 = n2(J_o, j_o, j_1) = n - n1, \tag{4.5.5}$$

$$n3 = n3_{\xi}^{f,l}(J_o, j_o, j_1) = n2 - \left|\left\{ s_i^f \mid s_i^f \geq \sigma_{J_o}^l \ \wedge \ i \notin I_{\pi_\xi}^{2,(f,l)}(J_o, j_o, j_1) \right\}\right|. \tag{4.5.6}$$

(Obviously, $\left|I_{\pi_\xi}^{2,(f,l)}(J_o, j_o, j_1)\right| = n1$ and $n - \left|I_{\pi_\xi}^{2,(f,l)}(J_o, j_o, j_1)\right| = n2$.)

We now want to compute the numbers of elements in the sets $\hat{S}_{\pi_\xi}^{2,(f,l)}(J_o, j_o, j_1),\ \xi = 1, \cdots, \Xi$. These numbers are equal to the numbers of possibilities that the remaining $n2$ components $\hat{s}_i^l,\ i \notin I_{\pi_\xi}^{2,(f,l)}(J_o, j_o, j_1)$ of $\hat{s}^l$ can be permuted under certain restrictions.

In greater detail, at first, the $j_o + j_1$ components $\hat{s}_i^l$ with size $\sigma_{J_o}^l - 1$, which were not taken into consideration in (4.5.2), can only be permuted in such a way that

$s_i^f \leq \hat{s}_{\pi_i}^l$

(see (d8) in Lemma 4.3.1).

That means, these components $\hat{s}_i^l$ can only be permuted with respect to $n3$ positions.

Thus, $\begin{pmatrix} n3 \\ j_o + j_1 \end{pmatrix}$ possibilities follow.

Then, $n2 - (j_o + j_1)$ positions remain for the $L - n1 - (j_o + j_1)$ components $\hat{s}_i^l$ of $\hat{s}^l$ with $0 < \hat{s}_i^l < \sigma_{J_o}^l - 1$.

That yields $\begin{pmatrix} n2 - (j_o + j_1) \\ L - n1 - (j_o + j_1) \end{pmatrix} \frac{(L-n1-(j_o+j_1))!}{\prod\limits_{i:\ 1 \leq \sigma_i^l < \sigma_{J_o}^l - 1} (L_i - L_{i-1})!}$ possibilities.

Finally, the positions still remaining are occupied with $\hat{s}_i^l = 0$.

In summary, a set $\hat{S}_{\pi_\xi}^{2,(f,l)}(J_o, j_o, j_1),\ \xi \in \{1, \cdots, \Xi\}$ has

$$Z1_{\xi}^{f,l}(J_o,j_o,j_1) := \begin{pmatrix} n3 \\ j_o+j_1 \end{pmatrix} \begin{pmatrix} n2-(j_o+j_1) \\ L-n1-(j_o+j_1) \end{pmatrix} \frac{(L-n1-(j_o+j_1))!}{\prod\limits_{i:\ 1\le\sigma_i^l<\sigma_{J_o}^l-1}(L_i-L_{i-1})!}$$

$$= \begin{cases} \begin{pmatrix} n3 \\ j_o+j_1 \end{pmatrix} \begin{pmatrix} n-L_{J_o+1} \\ L-L_{J_o+1} \end{pmatrix} \frac{(L-L_{J_o+1})!}{\prod\limits_{i:\ 1\le\sigma_i^l<\sigma_{J_o}^l-1}(L_i-L_{i-1})!} \\ \qquad\qquad\qquad\qquad\qquad\qquad \text{if } \sigma_{J_o}^l-1=\sigma_{J_o+1}^l, \\ \begin{pmatrix} n3 \\ j_o \end{pmatrix} \begin{pmatrix} n-L_{J_o} \\ L-L_{J_o} \end{pmatrix} \frac{(L-L_{J_o})!}{\prod\limits_{i:\ 1\le\sigma_i^l<\sigma_{J_o}^l-1}(L_i-L_{i-1})!} \\ \qquad\qquad\qquad\qquad\qquad\qquad \text{if } \sigma_{J_o}^l-1>\sigma_{J_o+1}^l, \end{cases} \tag{4.5.7}$$

elements, where $Z1$ only depends on $\xi$ as $n3$.

Now, we consider the sets

$$B_{\xi}^{2,(f,l)}(J_o,j_o,j_1) := \bigcup_{\hat{s}_{\pi}^l\in\hat{S}_{\pi_{\xi}}^{2,(f,l)}(J_o,j_o,j_1)} B_{n;k_0}^2(s^f,\hat{s}_{\pi}^l),\ \xi=1,\cdots,\Xi, \tag{4.5.8}$$

with $B_{n;k_0}^2(s^f,\hat{s}_{\pi}^l)$ as in Definition 4.3.4.

According to (4.5.7) and Definition 4.3.4, a set $B_{\xi}^{2,(f,l)}(J_o,j_o,j_1)$ contains the following number of elements:

$$\left|B_{\xi}^{2,(f,l)}(J_o,j_o,j_1)\right| = Z1_{\xi}^{f,l}(J_o,j_o,j_1) * Z2_{\xi}^{f,l}(J_o,j_o,j_1) * Z3^{f,l}(J_o,j_o,j_1), \tag{4.5.9}$$

where

$$Z2_{\xi}^{f,l}(J_o,j_o,j_1) = \prod_{i:s_i^f=\hat{s}_{\pi_i}^l\ge\sigma_{J_o}^l}(k_0+1-\hat{s}_{\pi_i}^l) \tag{4.5.10}$$

(see also Definition 4.3.4, (d13) and (4.5.1))

and

$$Z3^{f,l}(J_o,j_o,j_1) = \left(k_0+2-\sigma^l_{J_o}\right)^{j_o+j_1} - \sum_{j=0}^{j_o-1}\left(\begin{array}{c} j_o+j_1 \\ j \end{array}\right)\left(k_0+1-\sigma^l_{J_o}\right)^j$$
$$= \sum_{j=j_o}^{j_o+j_1}\left(\begin{array}{c} j_o+j_1 \\ j \end{array}\right)\left(k_0+1-\sigma^l_{J_o}\right)^j \tag{4.5.11}$$

(see Definition 4.3.4, (d14)) and thereby

„$-\sum_{j=0}^{j_o-1}\left(\begin{array}{c} j_o+j_1 \\ j \end{array}\right)\left(k_0+1-\sigma^l_{J_o}\right)^j$" includes that $w$ with more than $j_1$ coordinates $w_i = \sigma^l_{J_o}-1$ are not feasible according to Definition 4.3.4, (d14).

Thus, $\left|B^{2,(f,l)}_\xi(J_o,j_o,j_1)\right|$ are polynomials in $n$!

The consideration, in particular, of $Z1^{f,l}_\xi(J_o,j_o,j_1)$ implies that the degree of such a polynomial is

$$\begin{array}{ll} j_o+j_1+L-L_{J_o+1} & \text{if } \sigma^l_{J_o}-1=\sigma^l_{J_o+1}, \\ \text{and } j_o+L-L_{J_o} & \text{if } \sigma^l_{J_o}-1>\sigma^l_{J_o+1}. \end{array} \tag{4.5.12}$$

Now, let

$$p^{2,(J_o,j_o,j_1,\xi)}_{fl} := \sum_{w\in B^{2,(f,l)}_\xi(J_o,j_o,j_1)} q(w). \tag{4.5.13}$$

If the coordinates $w_i,\ i=1,\cdots,n$ of the requirements $w\in B_{n;k_0}$ are independent and identically distributed, where (4.2.1) and (4.2.2) are fulfilled, then

$$p^{2,(J_o,j_o,j_1,\xi)}_{fl}$$
$$= Z1^{f,l}_\xi(J_o,j_o,j_1) * ZP2^{f,l}_\xi(J_o,j_o,j_1) * ZP3^{f,l}(J_o,j_o,j_1) * ZP1^{f,l}(J_o,j_o,j_1), \tag{4.5.14}$$

follows, where

$$ZP2_{\xi}^{f,l}(J_o, j_o, j_1) = \prod_{i:s_i^f = \hat{s}_{\pi_i}^l \geq \sigma_{J_o}^l} \left( \sum_{w_i = \hat{s}_{\pi_i}^l}^{k_0} q_0(w_i) \right)$$

$$\prod_{i:s_i^f > \hat{s}_{\pi_i}^l \geq \sigma_{J_o}^l} q_0(\hat{s}_{\pi_i}^l))\; q_0(\sigma_{J_o}^l - 1)^{((L_{J_o+1} - L_{J_o} - j_1)*\delta)} \tag{4.5.15}$$

with $\delta = \begin{cases} 1 & \text{if } \sigma_{J_o}^l - 1 = \sigma_{J_o+1}^l, \\ 0 & \text{otherwise} \end{cases}$

(see (4.5.1), (4.5.2) and Definition 4.3.4, (d13), (d15)),

$$ZP3^{f,l}(J_o, j_o, j_1) = \left( \sum_{w_i = \sigma_{J_o}^l - 1}^{k_0} q_0(w_i) \right)^{j_o + j_1}$$

$$- \sum_{j=0}^{j_o - 1} \binom{j_o + j_1}{j} q_0(\sigma_{J_o}^l - 1)^{j_o + j_1 - j} \left( \sum_{w_i = \sigma_{J_o}^l}^{k_0} q_0(w_i) \right)^j$$

$$= \sum_{j=j_o}^{j_o + j_1} \binom{j_o + j_1}{j} q_0(\sigma_{J_o}^l - 1)^{j_o + j_1 - j} \left( \sum_{w_i = \sigma_{J_o}^l}^{k_0} q_0(w_i) \right)^j \tag{4.5.16}$$

(see Definition 4.3.4, (d14)),

(thereby "$- \sum_{j=0}^{j_o - 1} \binom{j_o + j_1}{j} q_0(\sigma_{J_o}^l - 1)^{j_o + j_1 - j} \left( \sum_{w_i = \sigma_{J_o}^l}^{k_0} q_0(w_i) \right)^j$" includes that $w$ with more than $j_1$ coordinates $w_i = \sigma_{J_o}^l - 1$ are not feasible according to Definition 4.3.4, (d14))

and

$$ZP1^{f,l}(J_o, j_o, j_1) = \prod_{i:\hat{s}_{\pi_i}^l < \sigma_{J_o}^l - 1} q_0(\hat{s}_{\pi_i}^l)) \tag{4.5.17}$$

(see Definition 4.3.4, (d15)).

Finally, in the case of discrete uniformly distributed requirements, (4.5.9) or (4.5.14) implies

$$p_{fl}^{2,(J_o,j_o,j_1,\xi)} = \frac{1}{(k_0+1)^n} \left| B_\xi^{2,(f,l)}(J_o,j_o,j_1) \right|$$

$$= \frac{1}{(k_0+1)^n} * Z1_\xi^{f,l}(J_o,j_o,j_1) * Z2_\xi^{f,l}(J_o,j_o,j_1) * Z3^{f,l}(J_o,j_o,j_1). \tag{4.5.18}$$

**Subcase: $\sigma_{J_o}^l = 1$ ($\sigma_{J_o}^l$ as from Definition 4.3.1) which means $J_o = y$, see (4.3.3)**

(Then and only then is $s^l$ a monotone successor of $s^f$, according to Definition 4.4.1(b)!)

Initially,

$$j_1 = n - \left| \left\{ i \mid s_i^f \neq 0 \right\} \right| - j_o = n - F - j_o \tag{4.5.19}$$

follows for $\sigma_{J_o}^l = 1$ according to Definition 4.3.3, (d6), where $F(= F_z)$ as in (4.3.1) and (4.3.2).

We now use in principle methods similar to those in the subcase $\sigma_{J_o}^l > 1$:

$\hat{S}_{\pi_\xi}^{2,(f,l)}(y,j_o,j_1)$, $B_\xi^{2,(f,l)}(y,j_o,j_1)$, $p_{fl}^{2,(y,j_o,j_1,\xi)}$ $\xi = 1,\cdots,\Xi$ have the same meaning as in the subcase $\sigma_{J_o}^l > 1$ (see (4.5.8) and (4.5.13) among other things).

However, we will see that a set $\hat{S}_{\pi_\xi}^{2,(f,l)}(y,j_o,j_1)$ contains only one element.

Sets $\hat{S}_{\pi_\xi}^{2,(f,l)}(y,j_o,j_1)$ are characterized by certain fixed components $\hat{s}_{\pi_i}^l$ of their elements $\hat{s}_\pi^l$. Similar to (4.5.1) and (4.5.2) and using $\sigma_y^l = 1$ and (4.5.19):

$L - j_o$ components $\hat{s}_{\pi_i}^l \geq 1$ of $\hat{s}_\pi^l$ are fixed for $i$ where

$$\hat{s}_i^l \geq 1 \text{ so that } s_i^f \geq \hat{s}_{\pi_i}^l \tag{4.5.20}$$

(see (d10) in Lemma 4.3.1),

and if $n > L$ then $n - L - j_1 = F + j_o - L$ (see (4.5.19)) additional components of $\hat{s}^l_{\pi_i}$ of $\hat{s}^l_\pi$ are fixed for $i$ where

$$s^f_i > 0 \text{ so that } \hat{s}^l_{\pi_i} = 0 \tag{4.5.21}$$

(see (d11) in Lemma 4.3.1).

For the remaining positions $i$, the relationship $s^f_i = 0$ is valid and, hence, $s^f_i \leq \hat{s}^l_{\pi_i} = 0$ (see (d8) in Lemma 4.3.1).

Thus, all components $\hat{s}^l_{\pi_i}$ have been fixed, $\hat{S}^{2,(f,l)}_{\pi_\xi}(y, j_o, j_1)$ only include one element and

$$B^{2,(f,l)}_\xi(y, j_o, j_1) = B^2_{n;k_0}(s^f, \hat{s}^l_\pi), \text{ with } \left\{\hat{s}^l_\pi\right\} = \hat{S}^{2,(f,l)}_{\pi_\xi}(y, j_o, j_1)$$

follows, where $B^2_{n;k_0}(s^f, \hat{s}^l_\pi)$ is defined as in Definition 4.3.4.

Specifying $Z2^{f,l}_\xi(J_o, j_o, j_1)$ and $Z3^{f,l}(J_o, j_o, j_1)$ from (4.5.10) and (4.5.11), the number of elements of such a set of requirements is

$$\begin{aligned}
&\left|B^{2,(f,l)}_\xi(y, j_o, j_1)\right| \\
&= \prod_{i:s^f_i = \hat{s}^l_{\pi_i} \geq 1} (k_0 + 1 - \hat{s}^l_{\pi_i}) * \left[(k_0+1)^{n-F} - \sum_{j=0}^{j_o-1} \binom{n-F}{j} k_0^j\right] \\
&\left(= \prod_{i:s^f_i = \hat{s}^l_{\pi_i} \geq 1} (k_0 + 1 - \hat{s}^l_{\pi_i}) * \left[\sum_{j=j_o}^{n-F} \binom{n-F}{j} k_0^j\right]\right)
\end{aligned} \tag{4.5.22}$$

(where $\left\{\hat{s}^l_\pi\right\} = \hat{S}^{2,(f,l)}_{\pi_\xi}(y, j_o, j_1)$).

Obviously, $\left|B^{2,(f,l)}_\xi(y, j_o, j_1)\right|$ is a sum of an exponential function and a polynomial in $n$. The degree of the polynomial is then

$$j_o - 1. \tag{4.5.23}$$

If the coordinates $w_i,\ i = 1, \cdots, n$ of the requirements $w \in B_{n;k_0}$ are independent and identically distributed, where (4.2.1) and (4.2.2) are fulfilled, then a specification of $ZP2^{f,l}_\xi(J_o, j_o, j_1)$ and $ZP3^{f,l}(J_o, j_o, j_1)$ from (4.5.15)

and (4.5.16) yields

($ZP1^{f,l}(J_o, j_o, j_1)$ is not relevant here since $\hat{s}^l_{\pi_i} < \sigma^l_{J_o} - 1 = 0$ is not possible):

$$p_{fl}^{2,(y,j_o,j_1,\xi)} = \prod_{i:s_i^f=\hat{s}^l_{\pi_i}\geq 1} \Big( \sum_{w_i=\hat{s}^l_{\pi_i}}^{k_0} q_0(w_i)\Big) \prod_{i:s_i^f>\hat{s}^l_{\pi_i}\geq 1} q_0(\hat{s}^l_{\pi_i})\; q_0(0)^{((F+j_o-L_y)}$$

$$* \left[ \sum_{j=j_o}^{n-F} \binom{n-F}{j} \; q_0(0)^{n-F-j} \; (1-q_0(0))^j \right], \tag{4.5.24}$$

(where $\{\hat{s}^l_\pi\} = \hat{S}^{2,(f,l)}_{\pi_\xi}(y, j_o, j_1)$).

Finally, in the case of discrete uniformly distributed requirements, (4.5.22) implies the following analogous to (4.5.18)

$$\begin{aligned} & p_{fl}^{2,(y,j_o,j_1,\xi)} \\ &= \tfrac{1}{(k_0+1)^n} * \prod_{i:s_i^f=\hat{s}^l_{\pi_i}\geq 1} (k_0+1-\hat{s}^l_{\pi_i}) * \left[ (k_0+1)^{n-F} - \sum_{j=0}^{j_o-1} \binom{n-F}{j} k_0^j \right] \\ &= \tfrac{1}{(k_0+1)^n} * \prod_{i:s_i^f=\hat{s}^l_{\pi_i}\geq 1} (k_0+1-\hat{s}^l_{\pi_i}) * \left[ \sum_{j=j_o}^{n-F} \binom{n-F}{j} k_0^j \right]. \end{aligned} \tag{4.5.25}$$

**Case:** In $C[w; su]$ a $(J_o, j_3)$-perturbation of the relation "$\leq$" between $s^f$ and $s^l_\pi$ is considered. (see Definition 4.3.5)

Furthermore, let $\hat{s}^l$ be a $(J_o, j_3)$-perturbed partition of $s^l$ (see Definition 4.3.6).

**Subcase:** $\sigma^l_{J_o} > 0$. **($\sigma^l_{J_o}$ as from Definition 4.3.5)**

We later split non-empty sets $\hat{S}^{f,l}_\pi(J_o, j_3, j_4)$ (see Definition 4.3.7) into disjunct subsets:

$\hat{S}^{1,(f,l)}_{\pi_\xi}(J_o, j_3, j_4), \;\; \xi = 1, \; \cdots \; , \; \Xi(= \Xi(J_o, j_3, j_4))$.

At first, the $n-L$ components of $s^l$ and $\hat{s}^l$ with value 0 can only be permuted in such a way that

$\hat{s}^l_{\pi_i} = 0 \;\Rightarrow\; s^f_i = 0$ and
$s^l_{\pi_i} = 0 \;\Rightarrow\; s^f_i = 0$

since $s^f_i \leq \hat{s}^l_{\pi_i} = 0$ follows from Lemma 4.3.7, (d27) for $\hat{s}^l_{\pi_i} = 0$ and $\sigma^l_{J_o} > 0$. This yields

$$\begin{pmatrix} n-F \\ n-L \end{pmatrix} = \begin{pmatrix} n-F \\ L-F \end{pmatrix} \tag{4.5.26}$$

possibilities.

$\hat{s}^l_{\pi_i} = 0$ into consideration, certain other components $\hat{s}^l_{\pi_i}$ of the elements $\hat{s}^l_\pi$ of $\hat{S}^{1,(f,l)}_{\pi_\xi}(J_o, j_3, j_4)$ are fixed. More specifically:

$L - L_{J_o-1} - j_3$ components $0 < \hat{s}^l_{\pi_i} \leq \sigma^l_{J_o}$ of $\hat{s}^l_\pi$ are fixed for $i$ where

$$\mathbf{0} < \hat{s}^l_i \leq \sigma^l_{J_o} \text{ so that } s^f_i \leq \hat{s}^l_{\pi_i} \tag{4.5.27}$$

(see (d27) in Lemma 4.3.7),

and if $\sigma^l_{J_o} + 1 = \sigma^l_{J_o-1}$ then $L_{J_o-1} - L_{J_o-2} - j_4$ additional components $\hat{s}^l_{\pi_i} = \sigma^l_{J_o} + 1$ of $\hat{s}^l_\pi$ are fixed for $i$ where

$$\hat{s}^l_i = \sigma^l_{J_o} + 1 \;\text{ so that } s^f_i < \hat{s}^l_{\pi_i} \tag{4.5.28}$$

(see (d28) in Lemma 4.3.7).

Furthermore, we use the symbols

$$\begin{aligned} &I^{1,(f,l)}_{\pi_\xi}(J_o, j_3, j_4) \;\text{ for the set of indices } i \text{ for which the} \\ &\text{components } \hat{s}^l_{\pi_i} \text{ of } \hat{s}^l_\pi \;(\hat{s}^l_{\pi_i} = 0 \text{ included}) \;\text{ are fixed above,} \end{aligned} \tag{4.5.29}$$

$$n4 = n4(J_o, j_3, j_4) = \begin{cases} n - L_{J_o-2} - j_3 - j_4 & \text{if } \sigma^l_{J_o} + 1 = \sigma^l_{J_o-1}, \\ n - L_{J_o-1} - j_3 & \text{if } \sigma^l_{J_o} + 1 < \sigma^l_{J_o-1}, \end{cases} \tag{4.5.30}$$

$$n5 = n5(J_o, j_3, j_4) = n - n4 = \begin{cases} L_{J_o-2} + j_3 + j_4 & \text{if } \sigma^l_{J_o} + 1 = \sigma^l_{J_o-1}, \\ L_{J_o-1} + j_3 & \text{if } \sigma^l_{J_o} + 1 < \sigma^l_{J_o-1}, \end{cases} \tag{4.5.31}$$

$$n6 = n6^{f,l}_{\xi}(J_o, j_3, j_4) = n5 - \left|\left\{ s^f_i \mid s^f_i \le \sigma^l_{J_o} \ \wedge \ i \notin I^{1,(f,l)}_{\pi_\xi}(J_o, j_o, j_1) \right\}\right|. \tag{4.5.32}$$

(Obviously, $\left| I^{1,(f,l)}_{\pi_\xi}(J_o, j_3, j_4) \right| = n4$.)

We now want to compute the numbers of elements in the sets $\hat{S}^{1,(f,l)}_{\pi_\xi}(J_o, j_3, j_4)$, $\xi = 1, \cdots, \Xi$. These numbers are equal to $\begin{pmatrix} n-F \\ L-F \end{pmatrix}$ (see (4.5.26)) multiplied by the numbers of the possibilities the remaining $n5$ components $\hat{s}^l_i$, $i \notin I^{1,(f,l)}_{\pi_\xi}(J_o, j_3, j_4)$ of $\hat{s}^l$ have to permute under certain restrictions.

More specifically, at first, the $j_3 + j_4$ components $\hat{s}^l_i$ with value $\sigma^l_{J_o} + 1$, which were not taken into consideration in (4.5.28), can only be permuted in such a way that

$s^f_i \ge \hat{s}^l_{\pi_i}$

(see (d29) in Lemma 4.3.7).

This means, these parts $\hat{s}^l_i$ can only be permuted with regard to $n6$ positions.

Thus, $\begin{pmatrix} n6 \\ j_3 + j_4 \end{pmatrix}$ possibilities follow.

Then $n5 - (j_3 + j_4) = n - n4 - (j_3 + j_4)$ positions remain for the $n - n4 - (j_3 + j_4)$ parts $\hat{s}^l_i$ of $\hat{s}^l$ with $\hat{s}^l_i > \sigma^l_{J_o} + 1$.

That yields $\frac{(n5-(j_3+j_4))!}{\prod\limits_{i:\ \sigma^l_i > \sigma^l_{J_o}+1} (L_i - L_{i-1})!}$ possibilities.

In summary, it follows that a set $\hat{S}^{1,(f,l)}_{\pi_\xi}(J_o, j_3, j_4)$, $\xi \in \{1, \cdots, \Xi\}$ has

$$Z4_{\xi}^{f,l}(J_o, j_3, j_4) := \begin{pmatrix} n-F \\ L-F \end{pmatrix} \begin{pmatrix} n6 \\ j_3+j_4 \end{pmatrix} \frac{(n5-(j_3+j_4))!}{\prod\limits_{i:\ \sigma_i^l > \sigma_{J_o}^l+1} (L_i - L_{i-1})!}$$

$$= \begin{cases} \begin{pmatrix} n-F \\ L-F \end{pmatrix} \begin{pmatrix} n6 \\ j_3+j_4 \end{pmatrix} \frac{L_{J_o-2}!}{\prod\limits_{i:\ \sigma_i^l > \sigma_{J_o}^l+1} (L_i - L_{i-1})!} \\ \qquad\qquad\qquad\qquad\qquad\qquad \text{if } \sigma_{J_o}^l + 1 = \sigma_{J_o-1}^l, \\ \begin{pmatrix} n-F \\ L-F \end{pmatrix} \begin{pmatrix} n6 \\ j_3 \end{pmatrix} \frac{L_{J_o-1}!}{\prod\limits_{i:\ \sigma_i^l > \sigma_{J_o}^l+1} (L_i - L_{i-1})!} \\ \qquad\qquad\qquad\qquad\qquad\qquad \text{if } \sigma_{J_o}^l + 1 < \sigma_{J_o-1}^l \end{cases} \tag{4.5.33}$$

elements.

Now we consider the sets

$$B_{\xi}^{1,(f,l)}(J_o, j_3, j_4) := \bigcup_{\hat{s}_{\pi}^l \in \hat{S}_{\pi_{\xi}}^{1,(f,l)}(J_o, j_3, j_4)} B_{n;k_0}^1(s^f, \hat{s}_{\pi}^l),\ \xi = 1, \cdots, \Xi \tag{4.5.34}$$

with $B_{n;k_0}^1(s^f, \hat{s}_{\pi}^l)$ as in Definition 4.3.8.

According to (4.5.33) and Definition 4.3.8, a set $B_{\xi}^{1,(f,l)}(J_o, j_3, j_4)$ includes the following number of elements:

$$\left| B_{\xi}^{1,(f,l)}(J_o, j_3, j_4) \right| = Z4_{\xi}^{f,l}(J_o, j_3, j_4) * Z5_{\xi}^{f,l}(J_o, j_3, j_4) * Z6^{f,l}(J_o, j_3, j_4) \tag{4.5.35}$$

where

$$Z5_{\xi}^{f,l}(J_o, j_3, j_4) = \prod_{i: s_i^f = \hat{s}_{\pi_i}^l \le \sigma_{J_o}^l} (\hat{s}_{\pi_i}^l + 1) \tag{4.5.36}$$

(see also Definition 4.3.8, (d30)) and (4.5.27))
and

$$Z6^{f,l}(J_o, j_3, j_4) = \left(\sigma^l_{J_o} + 2\right)^{j_3+j_4} - \sum_{j=0}^{j_3-1} \left( \begin{array}{c} j_3 + j_4 \\ j \end{array} \right) \left(\sigma^l_{J_o} + 1\right)^j$$
$$= \sum_{j=j_3}^{j_3+j_4} \left( \begin{array}{c} j_3 + j_4 \\ j \end{array} \right) \left(\sigma^l_{J_o} + 1\right)^j \tag{4.5.37}$$

(see Definition 4.3.8, (d31)) and thereby

„$-\sum_{j=0}^{j_3-1} \left( \begin{array}{c} j_3 + j_4 \\ j \end{array} \right) \left(\sigma^l_{J_o} + 1\right)^j$" includes that $w$ with more than $j_4$ coordinates $w_i = \sigma^l_{J_o} + 1$ are not feasible according to Definition 4.3.8, (d31).

Obviously, $\left| B^{1,(f,l)}_{\xi}(J_o, j_3, j_4) \right|$ is a polynomial in $n$ (for any $k_0$) of the degree $L - F$. (4.5.38)

Now, let

$$p^{1,(J_o,j_3,j_4,\xi)}_{fl} := \sum_{w \in B^{1,(f,l)}_{\xi}(J_o,j_3,j_4)} q(w). \tag{4.5.39}$$

If the coordinates $w_i,\ i = 1, \cdots, n$ of the requirements $w \in B_{n;k_0}$ are independent and identically distributed, where (4.2.1) and (4.2.2) are fulfilled, then

$$p^{1,(J_o,j_3,j_4,\xi)}_{fl} = Z4^{f,l}_{\xi}(J_o, j_3, j_4) * ZP5^{f,l}_{\xi}(J_o, j_3, j_4) * ZP6^{f,l}(J_o, j_3, j_4)$$
$$* ZP4^{f,l}(J_o, j_3, j_4) \tag{4.5.40}$$

follows, where

$$ZP5^{f,l}_{\xi}(J_o, j_3, j_4) = \prod_{i: s^f_i = \hat{s}^l_{\pi_i} \le \sigma^l_{J_o}} \left( \sum_{w_i=0}^{\hat{s}^l_{\pi_i}} q_0(w_i) \right)$$
$$\prod_{i: s^f_i < \hat{s}^l_{\pi_i} \le \sigma^l_{J_o}} q_0(\hat{s}^l_{\pi_i})\ q_0(\sigma^l_{J_o} + 1)^{((L_{J_o-1} - L_{J_o-2} - j_4) * \delta)} \tag{4.5.41}$$

$$\text{with } \delta = \begin{cases} 1 & \text{if } \sigma^l_{J_o} + 1 = \sigma^l_{J_o-1}, \\ 0 & \text{otherwise,} \end{cases}$$

(see (4.5.27), (4.5.28) and Definition 4.3.8, (d30), (d32)),

$$ZP6^{f,l}(J_o, j_3, j_4)$$

$$= \left( \sum_{w_i=0}^{\sigma^l_{J_o}+1} q_0(w_i) \right)^{j_3+j_4} - \sum_{j=0}^{j_3-1} \binom{j_3+j_4}{j} q_0(\sigma^l_{J_o}+1)^{j_3+j_4-j} \left( \sum_{w_i=0}^{\sigma^l_{J_o}} q_0(w_i) \right)^j$$

$$= \sum_{j=j_3}^{j_3+j_4} \binom{j_3+j_4}{j} q_0(\sigma^l_{J_o}+1)^{j_3+j_4-j} \left( \sum_{w_i=0}^{\sigma^l_{J_o}} q_0(w_i) \right)^j \tag{4.5.42}$$

(see Definition 4.3.8, (d31)),

$$\text{„}- \sum_{j=0}^{j_3-1} \binom{j_3+j_4}{j} q_0(\sigma^l_{J_o}+1)^{j_3+j_4-j} \left( \sum_{w_i=0}^{\sigma^l_{J_o}} q_0(w_i) \right)^j \text{"}$$

includes that $w$ with more than $j_4$ coordinates $w_i = \sigma^l_{J_o} + 1$ are not feasible according to Definition 4.3.8, (d31))

and

$$ZP4^{f,l}(J_o, j_3, j_4) = \prod_{i:\hat{s}^l_{\pi_i} > \sigma^l_{J_o}+1} q_0(\hat{s}^l_{\pi_i}) \tag{4.5.43}$$

(see Definition 4.3.8, (d32)).

Finally, in the case of discrete uniformly distributed requirements, (4.5.35) or (4.5.40) implies

$$p_{fl}^{1,(J_o,j_3,j_4,\xi)} = \frac{1}{(k_0+1)^n} \left| B_\xi^{1,(f,l)}(J_o, j_3, j_4) \right|$$

$$= \frac{1}{(k_0+1)^n} * Z4_\xi^{f,l}(J_o, j_3, j_4) * Z5_\xi^{f,l}(J_o, j_3, j_4) * Z6^{f,l}(J_o, j_3, j_4). \tag{4.5.44}$$

**Subcase: $\sigma^l_{J_o} = 0$ ($\sigma^l_{J_o}$ as from Definition 4.3.5) which means $J_o = y+1$, see (4.3.3)**

We now use similar methods, in principle, to those in the subcase $\sigma^l_{J_o} > 0$:

$\hat{S}^{1,(f,l)}_{\pi_\xi}(y+1, j_3, j_4)$, $B^{1,(f,l)}_{\xi}(y+1, j_3, j_4)$, $p^{1,(y+1,j_3,j_4,\xi)}_{fl}$ $\xi = 1, \cdots, \Xi$ have the same meaning as in the subcase $\sigma^l_{J_o} > 0$ (see (4.5.34) and (4.5.39) among other things).

We will later also split non-empty sets $\hat{S}^{f,l}_{\pi}(y+1, j_3, j_4)$ (see Definition 4.3.7) into disjunct subsets: $\hat{S}^{1,(f,l)}_{\pi_\xi}(y+1, j_3, j_4)$, $\xi = 1, \cdots, \Xi(= \Xi(y+1, j_3, j_4))$.

Initially, the $n - L - j_3$ components with value 0 of $s^l$ and $\hat{s}^l$ can only be permuted in such a way that

$\hat{s}^l_{\pi_i} = 0 \Rightarrow s^f_i = 0$ and
$s^l_{\pi_i} = 0 \Rightarrow s^f_i = 0$

since $s^f_i \leq \hat{s}^l_{\pi_i} = 0$ follows from Lemma 4.3.7, (d27) together with (d29) for $\hat{s}^l_{\pi_i} = 0$ and $\sigma^l_{J_o} = 0$. This yields

$$\begin{pmatrix} n-F \\ n-L-j_3 \end{pmatrix} = \begin{pmatrix} n-F \\ L-F+j_3 \end{pmatrix} \tag{4.5.45}$$

possibilities.

If now additional components $\hat{s}^l_{\pi_i}$ are set analogous to (4.5.27) and (4.5.28), then (4.5.27) is not relevant here since $\sigma^l_{y+1} = 0$. It remains:
if $\sigma^l_{J_o} + 1 = \sigma^l_{y+1} + 1 = \sigma^l_y = \sigma^l_{J_o-1}$, which means $1 = \sigma^l_{L_y}$, then
$L_y - L_{y-1} - j_4$ additional components $\hat{s}^l_{\pi_i} = 1$ of $\hat{s}^l_\pi$ are fixed for $i$ where

$$\hat{s}^l_i = 1 \text{ so that } s^f_i < \hat{s}^l_{\pi_i} = 1, \text{ hence } s^f_i = 0 \tag{4.5.46}$$

(see (d28) in Lemma 4.3.7).

Specifying $\sigma^l_{J_o} = 0$ and $J_o = y+1$ in (4.5.29) to (4.5.33), the number of elements in a set $\hat{S}^{1,(f,l)}_{\pi_\xi}(y+1, j_3, j_4)$, $\xi \in \{1, \cdots, \Xi\}$ is equal to

$$Z4_{\xi}^{f,l}(y+1,j_3,j_4) := \begin{pmatrix} n-F \\ L-F+j_3 \end{pmatrix} \begin{pmatrix} n6 \\ j_3+j_4 \end{pmatrix} \frac{(n5-(j_3+j_4))!}{\prod\limits_{i:\,\sigma_i^l>1} (L_i-L_{i-1})!}$$

$$= \begin{cases} \begin{pmatrix} n-F \\ L-F+j_3 \end{pmatrix} \begin{pmatrix} n6 \\ j_3+j_4 \end{pmatrix} \frac{L_{y-1}!}{\prod\limits_{i:\,\sigma_i^l>1} (L_i-L_{i-1})!} \\ \qquad\qquad\qquad\qquad\qquad\qquad \text{if } 1=\sigma_L^l, \\ \begin{pmatrix} n-F \\ L-F+j_3 \end{pmatrix} \begin{pmatrix} n6 \\ j_3 \end{pmatrix} \frac{L_y!}{\prod\limits_{i:\,\sigma_i^l>1} (L_i-L_{i-1})!} \\ \qquad\qquad\qquad\qquad\qquad\qquad \text{if } 1<\sigma_L^l. \end{cases} \tag{4.5.47}$$

Further specifications in (4.5.35) to (4.5.37) yield the number of elements in a set $B_{\xi}^{1,(f,l)}(y+1,j_3,j_4)$ to be:

$$\left|B_{\xi}^{1,(f,l)}(y+1,j_3,j_4)\right| = Z4_{\xi}^{f,l}(y+1,j_3,j_4) * \left(2^{j_3+j_4} - \sum_{j=0}^{j_3-1} \begin{pmatrix} j_3+j_4 \\ j \end{pmatrix}\right)$$

$$= Z4_{\xi}^{f,l}(y+1,j_3,j_4) * \sum_{j=j_3}^{j_3+j_4} \begin{pmatrix} j_3+j_4 \\ j \end{pmatrix}. \tag{4.5.48}$$

Hence, (see $Z4_{\xi}^{f,l}(y+1,j_3,j_4)$, in particular)
$\left|B_{\xi}^{1,(f,l)}(y+1,j_3,j_4)\right|$ is a polynomial in $n$ (for any $k_0$) (4.5.49)
of the degree $L-F+j_3$.

If the coordinates $w_i,\ i=1,\cdots,n$ of the requirements $w \in B_{n;k_0}$ are independent and identically distributed, where (4.2.1) and (4.2.2) are fulfilled, then a specification of $ZP5_{\xi}^{f,l}(J_o,j_3,j_4)$, $ZP6^{f,l}(J_o,j_3,j_4)$ and $ZP4^{f,l}(J_o,j_3,j_4)$ (see (4.5.40) to (4.5.43)) yields:

$$p_{fl}^{1,(y+1,j_3,j_4,\xi)} = Z4_{\xi}^{f,l}(y+1,j_3,j_4) * ZP5_{\xi}^{f,l}(y+1,j_3,j_4) \\ *ZP6^{f,l}(y+1,j_3,j_4) * ZP4^{f,l}(y+1,j_3,j_4), \tag{4.5.50}$$

where

$$ZP5_{\xi}^{f,l}(y+1,j_3,j_4) = q_0(0)^{\left|\left\{i:s_i^f=\hat{s}_{\pi_i}^l=0\right\}\right|}, \tag{4.5.51}$$

$$ZP6^{f,l}(y+1,j_3,j_4)$$

$$= (q_0(0)+q_0(1))^{j_3+j_4} - \sum_{j=0}^{j_3-1} \binom{j_3+j_4}{j} q_0(1)^{j_3+j_4-j}\, q_0(0)^j \tag{4.5.52}$$

$$= \sum_{j=j_3}^{j_3+j_4} \binom{j_3+j_4}{j} q_0(1)^{j_3+j_4-j}\, q_0(0)^j,$$

$$ZP4^{f,l}(y+1,j_3,j_4) = \prod_{i:\hat{s}_{\pi_i}^l>1} q_0(\hat{s}_{\pi_i}^l)). \tag{4.5.53}$$

Finally, in the case of discrete uniformly distributed requirements, (4.5.48) or (4.5.50) implies

$$p_{fl}^{1,(y+1,j_3,j_4,\xi)} = \tfrac{1}{(k_0+1)^n} \left| B_{\xi}^{1,(f,l)}(y+1,j_3,j_4) \right|$$

$$= \tfrac{1}{(k_0+1)^n} * Z4_{\xi}^{f,l}(y+1,j_3,j_4) * \left( 2^{j_3+j_4} - \sum_{j=0}^{j_3-1} \binom{j_3+j_4}{j} \right) \tag{4.5.54}$$

$$= \tfrac{1}{(k_0+1)^n} * Z4_{\xi}^{f,l}(y+1,j_3,j_4) * \left( \sum_{j=j_3}^{j_3+j_4} \binom{j_3+j_4}{j} \right).$$

If $s^f = s^l$ then the

**Case:** $s^f = s^l_\pi (= s^l)$

remains to be considered since in this case neither a $(J_o, j_1)$-perturbation of the relation
"$\geq$" nor a $(J_o, j_3)$-perturbation of the relation "$\leq$" between $s^f$ and $s^f_\pi$, respectively, are given.

Obviously, $s^f = s^*(s^f, w)$ is valid for $w \in B_{n;k_0}$ where $(0, \cdots, 0)^T \leq w \leq s^f$ (case $C[w, su]$; see also Theorem 4.3.10) or where $s^f \leq w \leq (k_0, \cdots, k_0)^T$ (case $C[su, w]$; see Theorem 4.3.5).
(See also the Theorems 4.3.11 and 4.3.6.)

Thus, we set

$$B^{(f,f)}_{spec} = \left\{ w \in B_{n;k_0} \mid (0, \cdots, 0)^T \leq w \leq s^f \ \vee \ s^f \leq w \leq (k_0, \cdots, k_0)^T \right\} \tag{4.5.55}$$

and

$$\begin{aligned} |B^{(f,f)}_{spec}| &= \prod_{i=1}^{n} (s^f_i + 1) + \prod_{i=1}^{n} (k_0 - s^f_i + 1) - 1 \\ &= \prod_{i=1}^{F} (s^f_i + 1) + \prod_{i=1}^{F} (k_0 - s^f_i + 1)(k_0 + 1)^{n-F} - 1 \end{aligned} \tag{4.5.56}$$

follows (refer to Lemma 3.3.10).

Obviously, $|B^{(f,f)}_{spec}|$ is a sum of an exponential function and a constant in $n$.

If the coordinates $w_i,\ i = 1, \cdots, n$ of the requirements $w \in B_{n;k_0}$ are independent and identically distributed, then we compute

$$
\begin{aligned}
p_{ll}^{spec} &:= \sum_{w \in B_{spec}^{(f,f)}} q(w) \\
&= \sum_{w:(0,\cdots,0)^T \le w \le s^f} q(w) + \sum_{w: s^f \le w \le (k_0,\cdots,k_0)^T} q(w) - q(s^f) \\
&= \prod_{i=1}^{F} \left( \sum_{w_j=0}^{s_i^f} q_0(w_j) \right) q_0(0)^{n-F} + \prod_{i=1}^{F} \left( \sum_{w_j=s_i^f}^{k_0} q_0(w_j) \right) \\
&\qquad - \prod_{i=1}^{F} q_0(s_i^f) q_0(0)^{n-F}.
\end{aligned}
\tag{4.5.57}
$$

In the case of discrete uniformly distributed requirements

$$
\begin{aligned}
p_{ll}^{spec} &= \tfrac{1}{(k_0+1)^n} \left| B_{spec}^{(f,f)} \right| \\
&= \tfrac{1}{(k_0+1)^n} \left( \prod_{i=1}^{F} (s_i^f + 1) + \prod_{i=1}^{F} (k_0 - s_i^f + 1)(k_0+1)^{n-F} - 1 \right)
\end{aligned}
\tag{4.5.58}
$$

follows.

Finally, (4.5.56) in the case that $s^f = s^l$ together with the summation of (4.5.9) and (4.5.22) over $(J_o, j_o, j_1)$ and $\xi$ and with the summation of (4.5.35) and (4.5.48) over $(J_o, j_3, j_4)$ and $\xi$ yield the number of all elements in a set $B_{n;k_0}^{*(f,l)} := \left\{ w \in B_{n;k_0} \mid s^l = s^*(s^f, w) \right\}$.

Since the possibilities for $(J_o, j_o, j_1)$, $\xi$ and $(J_o, j_3, j_4)$, $\xi$ do not depend on $n$ (and $k_0$) (for sufficiently large $n$), $|B_{n;k_0}^{*(f,l)}|$ is either a polynomial or a sum of an exponential function and a polynomial in $n$ according to (4.5.12), (4.5.23), (4.5.38) and (4.5.49).

Hence, (i) and the first part of (ii) have been proven.

Now, we will determine the degrees and the corresponding leading terms of elements of PRMs multiplied by $(k_0+1)^n$ as polynomials in $n$.

Initially, (4.5.12), (4.5.23), (4.5.38) and (4.5.49) imply that the degree of a $(k_0+1)^n * p^*_{fl}$ as a polynomial (in $n$) is not greater than $L$.

**Case** $s^l \neq (1,...,1,0,...,0)$**:**

According to (4.5.12) and (4.5.49) (only)

the **subcase:** $C[su;w]$**,** $\sigma^l_{J_o} > 1$
if the additional condition
$$\begin{cases} j_o + j_1 = L_{J_o+1} & \text{if } \sigma^l_{J_o} - 1 = \sigma^l_{J_o+1}, \\ \text{or } j_o = L_{J_o} & \text{if } \sigma^l_{J_o} - 1 > \sigma^l_{J_o+1} \end{cases}$$
is valid or

the **subcase:** $C[w;su]$**,** $\sigma^l_{J_o} = 0$**,** $J_o = y+1$
if the additional condition $F = j_3$
is valid

imply that the degree of an element $p^*(s^f|s^l) = p^*_{fl},\ s^l \neq (1,...,1,0,...,0)$ of a PRM multiplied by $(k_0+1)^n$ as a polynomial in $n$ is equal to $L$.

In order to compute the corresponding leading term we must determine the corresponding set of requirements.

At first, we will see that it is sufficient to consider only the above first subcase:

In the (second) subcase, $C[w;su]$, $\sigma^l_{J_o} = 0$, the additional condition $F = j_3$ implies that

$s^f_i > 0 \ \Rightarrow\ s^l_{\pi_i} (= \hat{s}^l_{\pi_i}) = 0$

(see Definition 4.3.5, (d20), in particular);
(this also implies $n \geq F + L$).

Furthermore, $j_4 = 0$ follows
(see Definition 4.3.7, (d23), in particular).

Then only $w = s^l_\pi$ (for certain $s^l_\pi$) are elements of the sets $B^{1,(f,l)}_\xi(y+1, j_3 = F, j_4 = 0)$ according to Definition 4.3.8.

Of course, $w = s^l_\pi$ also satisfy the case $C[su;w]$. Thus, it is sufficient to only consider the above first subcase.

In the (first) subcase: $C[su;w]$, $\sigma^l_{J_o} > 1$ the additional condition

$$\begin{cases} j_o + j_1 = L_{J_o+1} & \text{if } \sigma^l_{J_o} - 1 = \sigma^l_{J_o+1}, \\ \text{or } j_o = L_{J_o} & \text{if } \sigma^l_{J_o} - 1 > \sigma^l_{J_o+1} \end{cases}$$

implies

$$J_o = 1,\ j_o = L_1$$

$$\text{and} \begin{cases} j_1 = L_2 - L_1 & \text{if } \sigma^l_{J_o} - 1 = \sigma^l_{J_o+1} \\ j_1 = 0 & \text{if } \sigma^l_{J_o} - 1 > \sigma^l_{J_o+1} \end{cases}$$

(refer, in particular, to Definitions 4.3.1, (d3) and 4.3.3, (d7)).

Furthermore,

$$s^l_{\pi_i} = \sigma^l_{J_o=1} \ \Rightarrow\ s^f_i < s^l_{\pi_i} (= \sigma^l_{J_o=1}) \tag{4.5.59}$$

$$\begin{aligned} &\text{and if } \ \sigma^l_1 - 1 = \sigma^l_2, \\ &s^l_{\pi_i} = \sigma^l_2 \ \Rightarrow\ s^f_i \le s^l_{\pi_i} (= \sigma^l_2) \end{aligned} \tag{4.5.60}$$

follows (for sufficiently large $n$)
(refer, in particular, to Definitions 4.3.1, (d3) and 4.3.3, (d6)).

If we now apply the methods from the case $C[su;w]$, $\sigma^l_{J_o} > 1$ at the beginning of the proof in the case that the additional condition is supposed then

(4.5.1) and (4.5.2) are initially not relevant

according to (4.5.59) and (4.5.60), respectively.

Hence

$$\begin{aligned} &\Xi(1, L_1, j_1) = 1, \\ &I^{2,(f,l)}_{\pi_{\xi=1}}(1, L_1, j_1) = \varnothing, \\ &n1 = 0, \\ &n2 = n, \\ &n3 = n - \left| \left\{ s^f_i \mid s^f_i \ge \sigma^l_1 = s^l_1 \right\} \right|. \end{aligned}$$

Using the symbol

$$L_2^* = \begin{cases} L_2 & \text{if } s^l_{L_1} = s^l_{L_2} + 1, \text{ meaning } \sigma^l_1 - 1 = \sigma^l_2, \\ L_1 & \text{otherwise} \end{cases}$$

specifications from (4.5.7) and some corresponding subsequent relationships yield

$$Z1 := \begin{pmatrix} n3 \\ L_2^* \end{pmatrix} \begin{pmatrix} n - L_2^* \\ L - L_2^* \end{pmatrix} \frac{(L-L_2^*)!}{\prod\limits_{i:\ 1 \le \sigma^l_i < \sigma^l_1 - 1} (L_i - L_{i-1})!} \ ,$$

$Z2$ is not relevant since (4.5.1) and (4.5.2) are not relevant under the additional condition,

$$Z3 = \sum_{j=L_1}^{L_2^*} \begin{pmatrix} L_2^* \\ j \end{pmatrix} \left(k_0 + 1 - s^l_1\right)^j \ .$$

Thus (see (4.5.9)),

$$\left| B^{2,(f,l)}_{\xi=1}(1, L_1, j_1) \right|$$

$$= \begin{pmatrix} n3 \\ L_2^* \end{pmatrix} \begin{pmatrix} n - L_2^* \\ L - L_2^* \end{pmatrix} \frac{(L-L_2^*)!}{\prod\limits_{i:\ 1 \le \sigma^l_i < \sigma^l_1 - 1} (L_i - L_{i-1})!} \left[ \sum_{j=L_1}^{L_2^*} \begin{pmatrix} L_2^* \\ j \end{pmatrix} \left(k_0 + 1 - s^l_1\right)^j \right].$$

Thereby $\begin{pmatrix} n3 \\ L_2^* \end{pmatrix} \begin{pmatrix} n - L_2^* \\ L - L_2^* \end{pmatrix} = \begin{pmatrix} n - \left| \left\{ s^f_i \mid s^f_i \ge \sigma^l_1 = s^l_1 \right\} \right| \\ L_2^* \end{pmatrix} \begin{pmatrix} n - L_2^* \\ L - L_2^* \end{pmatrix}$ is also a polynomial in n with the degree $L_2^* + L - L_2^* = L$ and the coefficient $\frac{1}{L_2^*!} \frac{1}{(L-L_2^*)!} = \frac{1}{L!} \frac{L!}{L_2^*!\,(L-L_2^*)!} = \frac{1}{L!} \begin{pmatrix} L \\ L_2^* \end{pmatrix}$ of its leading term.

Hence, the degree of $\left| B^{2,(f,l)}_{\xi=1}(1, L_1, j_1) \right|$ as a polynomial in $n$ is also equal to $L$ and the coefficient of the corresponding leading term is

$$\frac{1}{L!} \begin{pmatrix} L \\ L_2^* \end{pmatrix} \frac{(L-L_2^*)!}{\prod\limits_{i:\ 1 \le \sigma^l_i < \sigma^l_1 - 1} (L_i - L_{i-1})!} \left[ \sum_{j=L_1}^{L_2^*} \begin{pmatrix} L_2^* \\ j \end{pmatrix} \left(k_0 + 1 - s^l_1\right)^j \right]$$

where $C^* := \frac{(L-L_2^*)!}{\prod\limits_{i:\ 1 \le \sigma^l_i < \sigma^l_1 - 1} (L_i - L_{i-1})!}$ is the number of permutations of the $L - L_2^*$ components $s^l_i$ which are smaller than $s^l_1 - 1$.

The corresponding formula in relation to $p^*(s^l|s^f)$ follows by means of the first equation from (4.5.18) in the case of discrete uniformly distributed requirements.

That the degrees and the leading terms for every $p^*(s^l|s^f), f = 1, 2, ..., r$ multiplied by $(k_0 + 1)^n$ as polynomials in $n$ ( $n \underset{(=)}{>} su$ ) in a column $l$ are the same is now easily seen from the above formulas.

**Case** $s^l = (1, ..., 1, 0, ..., 0)$**:**

$S_{n;su;k_0}$ includes, for $n \underset{(=)}{>} su$,

only one partition $s^l$ with $L = su$, namely $s^l = s^r = (1, ..., 1, 0, ..., 0)$, and

only one partition $s^{l'}$ with $L' = su-1$, namely $s^{l'} = s^{r-1} = (2, 1, ..., 1, 0, ..., 0)$.

According to the above results for $s^l \neq (1, ..., 1, 0, ..., 0)$, the degrees of $(k_0 + 1)^n * p^*_{fr-1}, f = 1, 2, ..., r$ as polynomials in $n$ are $su - 1$ and the coefficients of the corresponding leading terms are

$$\frac{1}{(su-1)!} \begin{pmatrix} su-1 \\ su-1 \end{pmatrix} \left[ \sum_{j=1}^{su-1} \begin{pmatrix} su-1 \\ j \end{pmatrix} (k_0 + 1 - 2)^j \right]$$

$$= \frac{1}{(su-1)!} * (k_0^{su-1} - 1).$$

The above considerations together with the fact that the sum of the elements of a row of a PRM is equal to 1 $(= n^0)$ yield that the degrees of $(k_0 + 1)^n * p^*_{fr}, f = 1, 2, ..., r$ as polynomials in $n$ are also $su - 1$ and the coefficients of the corresponding leading terms are

$$-\frac{1}{(su-1)!} * (k_0^{su-1} - 1).$$ ■

**Remarks 4.5.1.** *The use of perturbed partitions is not necessary in order to compute the exponential functions mentioned in Theorem 4.5.1 or the leading terms of the corresponding polynomials in n. For example, the exponential functions follow from the limits of elements of PRMs computed in Section 4.4.2. Using such direct methods, the computations are shorter and not as technical as when the concept of the perturbed partitions would be applied.*

*However, in order to prove that subsets of requirements* $B^{*(f,l)}_{n;k_0}$*, as in Theorem 4.5.1, and elements of the corresponding PRMs multiplied by* $(k_0+1)^n$ *in the case of discrete uniformly distributed requirements are polynomials in*

*$k_0$ and/or either polynomials or sums of exponential functions and polynomials in $n$, it seems inevitable that we must use a concept such as that of the perturbed partitions!*

### 4.5.2 The Elements of the Last Row and the Last Column of PRMs in the Case that $n \geq su$

**Elements of the Last Row**

**Theorem 4.5.2.** *Let a set of partitions* $S_{n;su;k_0} = \{s^1, s^2, \cdots, s^r\}$ [15] *with* $n \geq su$ *be given where* $s^r : s^r_1 = 1 = \cdots 1 = s^r_{su} > s^r_{su+1} = 0 = \cdots 0 = s^r_n$.

*Furthermore, let* $s^l \in S_{n;su;k_0}$, $l \neq r$ [16] *with (w. l. o. g.)* $s^l_1 \geq s^l_2 \geq \cdots \geq s^l_n$.

*Then,*

$$s^l = s^*(s^r, w) \quad \Leftrightarrow \quad \begin{cases} w \in B^{1,(r,l)}_{n;k_0} & \text{in case } C[w, su], \\ w \in B^{2,(r,l)}_{n;k_0} & \text{in case } C[su, w], \end{cases}$$

*where*

$B^{1,(r,l)}_{n;k_0}$ *is the subset of requirements* $w$ *with coordinates*

$$\begin{cases} s^l_1, \ldots, s^l_L \text{ and } n - L \text{ zeroes } (L \text{ as in } (4.3.1)) & \text{if } s^l_L > 1 \\ \quad \text{or} & \\ s^l_1, \ldots, s^l_{L_{y-1}} \text{ and } n - L_{y-1} \text{ coordinates equal to } 0 \text{ or } 1, & \\ \quad \text{where at most } L - L_{y-1} \text{ coordinates can} & \\ \quad \text{be equal to } 1 \ (L_{y-1} \text{ as in } (4.3.3)) & \text{if } s^l_L = 1 \end{cases}$$

*and*

$B^{2,(r,l)}_{n;k_0}$ *is the subset of requirements* $w$ *with coordinates*

[15] The partitions are arranged here in order of (partial) dominance, see Section 4.1. However, only the last element $s^r$ is important for this section, see below.

[16] The case $l = r$ will be considered in Theorem 4.5.4.

$$\begin{cases} s^l_{L_1+1}, \dots, s^l_L, s^l_{L-1} \text{ and } n-L \text{ zeroes} \\ \quad \text{and } L_1 \text{ coordinates } w_i : s^l_{L_1} \le w_i \le k_0 & \text{if } \sigma^l_1 > \sigma^l_2 + 1 \\ & (\sigma^l \text{ as in (4.3.4))} \\ \quad \text{or} \\ s^l_{L_2+1}, \dots, s^l_L, s^l_{L-1} \text{ and } n-L \text{ zeroes} \\ \quad \text{and } L_2 \text{ coordinates } w_i : s^l_{L_2} \le w_i \le k_0, \text{ where} \\ \quad \text{at most } L_2 - L_1 \text{ coordinates can} \\ \quad \text{be equal to } s^l_{L_2} & \text{if } \sigma^l_1 = \sigma^2_l + 1. \end{cases}$$

*Furthermore,*

$$p^*_{rl} = \sum_{w \in B^{1,(r,l)}} q(w) + \sum_{w \in B^{2,(r,l)}} q(w) - \sum_{w:\, w \text{ permutation of } s^l} q(w), \; l \ne r$$

*are the elements of the last rows in the corresponding PRMs.*

***Concerning the proof.*** Using the iterative method for the computation of feasible balanced partitions from Section 4.2.2 and the sets $B^{1,(r,l)}$ and $B^{2,(r,l)}$ given in the theorem, the statements of the theorem are simple to prove. We omit the detailed proof here.

In the case of discrete uniformly distributed requirements the above theorem yields the following formulas:

**Corollary 4.5.3.** *Let a set of partitions* $S_{n;su;k_0} = \{s^1, s^2, \cdots, s^r\}$ *be given where* $s^r : s^r_1 = 1 = \cdots 1 = s^r_{su} > s^r_{su+1} = 0 = \cdots 0 = s^r_n$ *and let the requirements* $w \in B_{n;k_0}$ *be discrete uniformly distributed.*

*Furthermore, let* $s^l \in S_{n;su;k_0}$, $l \ne r$ *with (w. l. o. g.)* $s^l_1 \ge s^l_2 \ge \cdots \ge s^l_n$.

*Then the elements of the last row of the corresponding PRM can be computed by means of*

$$p^*_{rl} = p^{*1}_{rl} + p^{*2}_{rl} - p^{*1,2}_{rl} \; (l \ne r),$$

*where, using the symbols from (4.3.3) and (4.3.4),*

$$p^{*1}_{rl} = \frac{1}{(k_0+1)^n} \begin{cases} \binom{n}{L} \frac{L!}{L_1!(L_2-L_1)!\dots(L-L_{y-1})!} & \text{if } s^l_L > 1, \\ \sum_{j=0}^{L-L_{y-1}} \binom{n}{L_{y-1}+j} \frac{(L_{y-1}+j)!}{L_1!(L_2-L_1)!\dots(L_{y-1}-L_{y-2})!\, j!} & \text{if } s^l_L = 1 \end{cases}$$

*(here $j$ is the number of $w_i$ from the corresponding $w$ which are equal to 1, if $s^l_L = 1$),*

$$p^{*2}_{rl} = \frac{1}{(k_0+1)^n} \begin{cases} \binom{n}{L} \frac{L!}{L_1!(L_2-L_1)!\ldots(L-L_{y-1})!} \left(k_0+1-\sigma^l_1\right)^{L_1} & \text{if } \sigma^l_1 > \sigma^l_2 + 1 \\ \binom{n}{L} \frac{L!}{L_2!(L_3-L_2)!\ldots(L-L_{y-1})!} * \\ \qquad \sum_{j=0}^{L_2-L_1} \binom{L_2}{j} (k_0+1-\sigma^l_1)^{L_2-j} & \text{if } \sigma^l_1 = \sigma^l_2 + 1 \end{cases}$$

*(here $j$ is the number of $w_i$ from the corresponding $w$ which are equal to $s^l_{L_2}$, if $\sigma^l_1 = \sigma^l_2 + 1$)*

*and*

$$p^{*1,2}_{rl} = \frac{1}{(k_0+1)^n} \binom{n}{L} \frac{L!}{L_1!(L_2-L_1)!\ldots(L-L_{y-1})!}.$$

**Elements of the Last Column**

**Theorem 4.5.4.** *Let a set of partitions $S_{n;su;k_0} = \{s^1, s^2, \cdots, s^r\}$ with $n \geq su$ be given, where $s^r$ : $s^r_1 = 1 = \cdots 1 = s^r_{su} > s^r_{su+1} = 0 = \cdots 0 = s^r_n$.*

*Furthermore, let $s^f \in S_{n;su;k_0}$ with (w. l. o. g.) $s^f_1 \geq s^f_2 \geq \cdots \geq s^f_n$.*

*Then*

$$s^r = s^*(s^f, w) \quad \Leftrightarrow \quad w \in B^a \setminus B^b$$

*where*

$$B^a = \left\{ w \in B_{n;k_0} \,\middle|\, w : \begin{cases} w_i \in \{0,1\} & \text{for } i \leq \eta^f, \\ w_i \in \{0,1,\cdots,k_0\} & \text{for } i > \eta^f \end{cases} \right\}$$

*($\eta_f$ as in Definition 4.4.1(a)),*

$B^b = \{w \in B^a \mid w$ *with more than $n - su$ coordinates equal to 0, except for the $w$ of this type consisting of only zeroes and ones and with at least $H = \sum_{i=1}^{\eta^f} s^f_i - \eta^f$ coordinates $w_i = 1$ where $s^f_i = 0$* $\}$.

*Furthermore,*

$$p^*_{fr} = \sum_{w \in B^a} q(w) - \sum_{w \in B^b} q(w)$$

*are the elements of the last columns of corresponding PRMs.*

***Concerning the proof.*** Using the iterative method for the computation of feasible balanced partitions from Section 4.2.2 and the sets $B^a$ and $B^b$ given in the theorem, the statements of the theorem are simple to prove. We omit the detailed proof here.

**Corollary 4.5.5.** *Let a set of partitions* $S_{n;su;k_0} = \{s^1, s^2, \cdots, s^r\}$ *be given, where* $s^r : s^r_1 = 1 = \cdots 1 = s^r_{su} > s^r_{su+1} = 0 = \cdots 0 = s^r_n$ *and let the requirements* $w \in B_{n;k_0}$ *be discrete uniformly distributed.*

*Furthermore, let* $s^f \in S_{n;su;k_0}$ *with (w. l. o. g.)* $s^f_1 \geq s^f_2 \geq \cdots \geq s^f_n$.

*Then the elements of the last column of the corresponding PRM can be computed by means of*

$$p^*_{fr} = \tfrac{1}{(k_0+1)^n} \left[ 2^{\eta^f} (k_0+1)^{n-\eta^f} - |B^b| \right]$$

$$(\eta_f \text{ as in Definition 4.4.1(a)}),$$

*where, using F from (4.3.1) and H from Theorem 4.5.4,*

$$|B^b| = \sum_{\beta=0}^{\eta^f} \binom{\eta^f}{\beta} \left[ \sum_{\epsilon=0}^{su-1-\beta} \sum_{\mu=\max\{0,\epsilon-(n-F)\}}^{\min\{F-\eta^f,\epsilon\}} \binom{F-\eta^f}{\mu} \binom{n-F}{\epsilon-\mu} * \left( (k_0)^\epsilon - \begin{cases} 1 & \text{if } \epsilon-\mu \geq H \\ 0 & \text{otherwise} \end{cases} \right) \right]$$

$$= \sum_{\beta=0}^{\eta^f} \binom{\eta^f}{\beta} \left[ \sum_{\epsilon=0}^{su-1-\beta} \binom{n-\eta^f}{\epsilon} (k_0)^\epsilon - \sum_{\epsilon=H}^{su-1-\beta} \sum_{\mu=\max\{0,\epsilon-(n-F)\}}^{\min\{F-\eta^f,\epsilon-H\}} \binom{F-\eta^f}{\mu} \binom{n-F}{\epsilon-\mu} \right].$$

***Concerning the proof.*** It is obvious that $|B^a| = 2^{\eta^f} (k_0+1)^{n-\eta^f}$.

That the number of elements of $B^b$ can be computed according to the formula from the corollary will be evident by an explanation of the indices and bounds of the summations and the corresponding binomial coefficients. Here, however, we only discuss the fundamental ideas.

The indices of the summations initially include the following numbers of certain coordinates of $w \in B^b$:

$\beta$: the number of $w_i = 1$ where $i \in \{1,\ 2, \cdots,\ \eta^f\}$,

$\epsilon$: the number of $w_i \geq 1$ where $i \in \{\eta^f + 1,\ \eta^f + 2,\ \cdots,\ n\}$,

$\mu$: the number of $w_i \geq 1$ where $i \in \{\eta^f + 1,\ \eta^f + 2,\ \cdots,\ F\}$.

The upper bound $su - 1 - \beta$ of the summation with the index $\epsilon$ implies that $\epsilon + \beta \leq su - 1$, which means,

fewer than $su$ coordinates of $w$ are not $0$.

Hence, more than $n - su$ coordinates of $w$ are equal to 0.

Thus, the first condition in the definition of $B^b$ (see Theorem 4.5.4) is satisfied.

$\epsilon - \mu$ is the number of $w_i \geq 1$ where $i > F$ which means $s_i^f = 0$.

If $\epsilon - \mu \geq H$ then the $(k_0)^\epsilon$ possibilities corresponding to the $\epsilon$ coordinates $w_i : 1 \leq w_i \leq k_0$ are reduced by one possibility with all such $w_i = 1$:

$$(k_0)^\epsilon - \begin{cases} 1 & \text{if } \epsilon - \mu \geq H \\ 0 & \text{otherwise} \end{cases} .$$

Thus, the second condition in the definition of $B^b$ (see Theorem 4.5.4) is also satisfied.

The lower bound $\min\{F - \eta^f, \epsilon - H\}$ and the upper bound $\max\{0, \epsilon - (n - F)\}$ of the sum over $\mu$ guarantee the compatibility between the indices $\mu$ and $\epsilon$ and the characteristic quantities $F, n, \eta^f$ of $s^f$.

The binomial coefficients in the formula for $|B^b|$ include, of course, all choices of $i$ where $w_i \geq 1$.

The second equation for $|B^b|$ follows from the first equation (among other

things) by means of the Vandermonde's identity [17] for binomial coefficients.

## 4.6 Poisson Equations for PRMs and the Monotonicity of their Solutions

In this section we will initially give the central definition of Poisson equations for PRMs and the definition of the monotonicity of their solutions (independent of Chapters 2 and 3).

In relation to PRMs as matrices of transition probabilities of reduced SDDP problems, the monotonicity of the solutions of the corresponding Poisson equations means that the decisions for feasible states with least square sums of their components are optimal for the corresponding reduced SDDP problems. (See also the *notes* on the connections with Chapters 2 and 3, below.)

We conjecture that the solutions of all Poisson equations for PRMs are monotone.

It is simple to prove the conjecture for a small number of Poisson equations where the corresponding PRMs and the right sides of the equations satisfy the conditions of dominance (see also Section 2.3.3.2).

This proof and considerations of several other special cases can be found in Section 4.6.2.

The main results of this section are the proofs of the conjecture in relation to PRMs, which are based on sets of sparse partitions with sufficiently great $n$ or on sets of non-truncated heavy partitions with sufficiently great $k_0$, given in the Subsections 4.6.3 and 4.6.4. Limits of PRMs with regard to sets of sparse partitions and sets of non-truncated heavy partitions from Sections 4.4.2 are 4.4.3 are used for these proofs.

The solutions of the Poisson equations with regard to limits of PRMs

[17] Vandermonde's identity: $\binom{r+s}{n} = \sum_{k=0}^{n} \binom{r}{k}\binom{s}{n-k}$ where n is a non-negative integer.

have an elegant structure, in contrast to the formulas of the limits of PRMs themselves. The solutions are vectors which include, in relation to the distribution of requirements, generalized harmonic numbers.

*Notes on the connections with Chapters 2 and 3:*

*The Poisson equations were initially introduced in (2.3.5) in Section 2.3.1. Such an equation include a matrix of transition probabilities and average (one-step) reward functions, which corresponds to a decision function of a MDP.*

*For example, Poisson equations are used in the Howard algorithm, by which optimal solutions of MDPs are computed (see the last part of Section 2.3.2.2).*

*Monotonicity of the solutions of Poisson equations was introduced in Definition 2.3.3 for DA MDPs, where the underlying internal costs and the average (one-step) reward functions did not depend on the (feasible) decisions. Also, in Corollary 2.3.10 the relationship between the monotonicity of the solutions of Poisson equations and the optimality of the corresponding decisions has been shown.*

*PRMs are matrices of transition probabilities of reduced SDDP problems (where identical basic cost and independent and identically distributed requirements are assumed) for decisions $d^*$ for feasible states (unordered partitions) with least square sums of their components (see the second part of Section 3.4.2).*

*The Poisson equations for PPMs are the Poisson equations of reduced SDDP problems and decisions $d^*$ where, in addition, certain affine transformations of the right side of the equations are allowed. Such affine transformations have no effect on the monotonicity of the solutions of the Poisson equations (see also Lemma 2.3.2). However, such affine transformations are useful since, among other things, no useful formulas are known for the additive remainder terms in the formulas of the average (one-step) reward functions for reduced SDDP problems (see Theorem 3.4.1).*

*If a solution of a Poisson equation for a PRM is monotone, then decisions $d^*$ for feasible states (unordered partitions) with the least square sums of*

*their components are optimal for the corresponding reduced SDDP problems (see also the proof of Corollary 4.7.1)!*

*Another aspect of the importance of the results of this section involves the dominance of MDPs (see Section 2.3.3.2).*

*As mentioned above, the conditions of dominance are only satisfied for a small number of PRMs (as matrices of transition probabilities of reduced SDDP problems). The conditions are typically infringed on for "most" of the PRMs - however only to a slight extent. This also induces the questions, whether and in which way the concept of dominance could be generalized.*

### 4.6.1 Poisson Equations for PRMs

**Definition 4.6.1.** *Let a partially ordered set of restricted partitions* $(S_{n;su;k_0}, \rightarrow) = (\{s^1, s^2, \cdots, s^r\}, \rightarrow)$, *as in Definition 4.1.1(a), (c) and Lemma 4.1.5, and a corresponding set of requirements* $B_{n;k_0}$ *be given and let the requirements* $w_i$, $(i = 1, \cdots, n)$ *be independent and identically distributed, where (4.2.2) is additionally assumed.*

*Furthermore, let* $P^* = (p^*_{fl})$ *be the the corresponding PRM.*

*The vector equations (with the variables* $(g, \nu) \in \mathbb{R} \times \mathbb{R}^r$*)*

$$g \begin{pmatrix} -1 \\ \vdots \\ -1 \end{pmatrix} + (P^* - I)\nu = -\gamma' \tag{4.6.1}$$

*are called the Poisson equations for the PRM* $P^*$,

*where* $I$ *is the identity matrix,*

$$\gamma_f (= \gamma(s^f)) = \sum_{i=1}^{n} \sum_{w_i=0}^{s_i^f} (s_i^f - w_i)\, q_0(w_i)^{18} \text{ for } f = 1, 2, \cdots, r \tag{4.6.2}$$

*and* $\gamma'$ *any affine transformation of* $\gamma$ :

$$\gamma' = \alpha\, \gamma + \beta \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \quad \text{with } \alpha > 0. \tag{4.6.3}$$

[18] See Theorem 3.4.1(c).

It is well-known that if one variable $\nu_{f_0}$ is fixed in any way, then under the condition (4.2.7), for example, the remaining equation system has an unique solution. (See the proof of Theorem 2.4.8 by Müller and Nollau [28], for instance, and Lemma 2.3.2).

**Definition 4.6.2.** *A solution of a Poisson equation for a PRM, as in Definition 4.6.1, is called monotone (in $\nu$) (with respect to the partial order) if*

$$s^f \dot{\to} s^l \;\Rightarrow\; \nu_f > \nu_l$$

*(consequently $s^f \to s^l \;\Rightarrow\; \nu_f > \nu_l$).*

(If a solution of a Poisson equation is monotone, then the solutions of a the corresponding Poisson equations with another affine transformation (with $\alpha > 0$ ) of $\gamma$ are also monotone according to Lemma 2.3.2.)

We expect that the statement of following conjecture is true:

**Conjecture:** Solutions of all Poissons equation for PRMs are monotone!

*(Referring to Chapter 3, this conjecture means that decisions for feasible unordered partitions with least square sums of their components are optimal for the corresponding reduced SDDP problems!)*

## 4.6.2 Partial Results

### 4.6.2.1 PRMs for Sets of Restricted Partitions which are Equivalent with regard to the Poisson Equation

In Section 4.2.3 PRMs (among other things) are considered for sets of restricted partitions $S_{n;su;k_0}$ and $S_{n;\bar{su},k_0}$ with $\bar{su} = n\ k_0 - su$, which are equivalent with regard to the Poisson equation (see Definition 4.2.5(c)).

According to Theorem 4.2.5 such PRMs are identical if $q_0$ satisfies (4.2.8).

The following Lemma 4.6.1 also implies the equivalence of the corresponding $\gamma_f$ from (4.6.2) (except for perhaps certain additive remainder terms), so that the solutions of the corresponding Poisson equations are either the same or are affine transformations of each other (see the following Lemma 4.6.2).

**Lemma 4.6.1.** *Let sets of restricted partitions* $S_{n;su;k_0} = \{s^1, s^2, \cdots, s^r\}$ *and* $S_{n;\bar{s}u;k_0} = \{\bar{s}^1, \bar{s}^2, \cdots, \bar{s}^r\}$ *be given, where* $\bar{s}u = n\ k_0 - su$, *and let* $s^f$ *and* $\bar{s}^f$ *be complementary partitions for* $f = 1, \cdots, r$ *as in Definition 4.2.5(a) with* $k_c = k_0$.

*Furthermore, let* $B_{n;k_0}$ *be the corresponding set of requirements and let* $w_i$, $(i = 1, \cdots, n)$ *be independent and identically distributed. In addition, let the conditions (4.2.2) and (4.2.8)*

$$q_0(w_i) = q_0(k_0 - w_i) \text{ for } i = 1, 2, ..., n,\ w \in B_{n;k_0}$$

*be satisfied. Then*

$$\gamma(s^l) - \gamma(s^f) = \gamma(\bar{s}^l) - \gamma(\bar{s}^f) \text{ for any } \{f, l\} \subseteq \{1, 2, .., r\}.$$

***Proof.*** According to (4.6.2) we have

$$\gamma(s) = \sum_{i=1}^{n} \sum_{w_i=0}^{s_i} (s_i - w_i)\ q_0(w_i),\ s \in S_{n;su;k_0}$$

and

$$\gamma(\bar{s}) = \sum_{i=1}^{n} \sum_{\bar{w}_i=0}^{\bar{s}_i} (\bar{s}_i - \bar{w}_i)\ q_0(\bar{w}_i),\ \bar{s} \in S_{n;\bar{s}u;k_0}.$$

Now, we replace $\bar{s}$ by $(k_0, k_0, \cdots, k_0)^T - s$ in the second equation above and according to Definition 4.2.5(a) where $k_c = k_0$ :

$$\gamma(\bar{s}) = \sum_{i=1}^{n} \sum_{\bar{w}_i=0}^{k_0-s_i} (k_0 - s_i - \bar{w}_i)\ q_0(\bar{w}_i)$$

results.

The substitution $\bar{w}_i = k_0 - w_i$ yields

$$\gamma(\bar{s}) = \sum_{i=1}^{n} \sum_{k_0-w_i=0}^{k_0-s_i} (k_0 - s_i - k_0 + w_i)\ q_0(k_0 - w_i).$$

Using (4.2.8)

$$\gamma(\bar{s}) = \sum_{i=1}^{n} \sum_{w_i=s_i}^{k_0} (-s_i + w_i)\ q_0(w_i)$$

follows.

Finally, the computation of the difference:

$$\gamma(s) - \gamma(\bar{s}) = \sum_{i=1}^{n} \sum_{w_i=0}^{k_0} (s_i - w_i)\, q_0(w_i)$$

$$= \sum_{i=1}^{n} s_i \sum_{w_i=0}^{k_0} q_0(w_i) - \sum_{i=1}^{n} \sum_{w_i=0}^{k_0} w_i\, q_0(w_i)$$

$$= su - n\, E(w_i) \quad \text{(since } w_i \text{ are independent and identically distributed)}$$

shows that such a difference does not depend on $s$ or $\bar{s}$.

Hence,

$$\gamma(s^l) - \gamma(s^f) = \gamma(\bar{s}^l) - \gamma(\bar{s}^f).$$

■

The above lemma together with Theorem 4.2.5 yields:

**Lemma 4.6.2.** *Let the same assumption as in Lemma 4.6.1 be valid and let the corresponding PRMs with regard to $S_{n;su;k_0}$ and $S_{n;\bar{s}u;k_0}$ be given.*

*Then, the solutions of the corresponding Poisson equations are either the same or are affine transformations of each other.*

#### 4.6.2.2 The Dominance Condition and PRMs with regard to m-Totally Ordered Sets of Partitions

In this subsection we will see that solutions of the Poisson equations are monotone if the corresponding PRMs satisfy the following dominance condition.

We will also see that PRMs with regard to "m-totally" ordered sets of partitions fulfil this dominance condition.

**Definition 4.6.3.** *Let the numbering of the elements of a given set of restricted partitions $S_{n;su;k_0} = \{s^1, s^2, \cdots, s^r\}$ be such that*
$l < f \;\Rightarrow\; \gamma(s^l) \geq \gamma(s^f)$.

*(This numbering also implies partial order as in Definition 4.1.1(a), (c) and Lemma 4.1.5 according to Lemma 4.2.2(a).)*

*Furthermore, let* $P^* = (p^*_{fl})$ *be the corresponding PRM.*

*Then the condition*

$$\sum_{l=1}^{\bar{l}} p^*_{1l} \geq \sum_{l=1}^{\bar{l}} p^*_{2l} \geq \cdots \geq \sum_{l=1}^{\bar{l}} p^*_{rl} \ \ \textit{for} \ \ \bar{l} = 1, 2, \cdots, r \tag{4.6.4}$$

*is called the dominance condition.*

**Corollary 4.6.3.** *If a PRM satisfies the dominance condition (4.6.4), then the solutions of the corresponding Poisson equations are monotone (with respect to the partial order).*

Definition 4.6.3 and the above corollary are formulated independent of Chapters 2 and 3. However, this corollary is proven most simply by means of statements from Chapters 2 and 3.

***Proof of Corollary* 4.6.3.** We consider reduced SDDP problems (see the second part of Section 3.4.2) with decisions $d^*$ (keep (3.4.16) and Lemma 4.2.2(b) in mind) for feasible unordered partitions with least square sums of their parts where the corresponding matrices of transition probabilities (which are PRMs) satisfy the dominance property (4.6.4).

Below we show that the conditions in Definition 2.3.5 are fulfilled for the reduced SDDP problems mentioned above. The decisions $d^*$ are then optimal according to Theorem 2.3.17.

Note that (2.3.13) and (2.3.14) are valid for the reduced SDDP problems (see Section 3.4.1) and then Theorem 2.3.8 can be applied with regard to $d^*$.

If $s^f \dot{\rightarrow} s^l$ then $s^f < s^l$, as with the almost-partial order from Theorem 2.3.8(i) based on $d^*$, and Theorem 2.3.8(ii) hence yields $\nu^f > \nu^l$ which results in the monotonicity of the solutions of the corresponding Poisson equations.

Checking of the conditions from Definition 2.3.5:

(C1), (Cr1): are valid according to (4.6.4),

(C2), (Cr2): follow from the construction of reduced SDDP problems (Section 3.4.2),

(C3), (Cr3): follow from the Definition of $d^*$ (see (3.4.16)) together with the numbering of the elements of $S_{n;su;k_0}$ in Definition 4.6.3.

■

In general, PRMs do not satisfy the dominance property (4.6.4). (See also the 3rd relationship of the following Lemma 4.6.4.)

However, in relation to main minimal chains (Definition 4.1.3) the dominance property is valid for all PRMs (see the following Theorem 4.6.5).

Hence, the dominance property is satisfied by PRMs whose corresponding sets of restricted partitions $S_{n;su;k_0}$ are themselves main minimal chains (see the following Definition 4.6.4 and Corollary 4.6.6).

**Lemma 4.6.4.** *Let* $w \in B_{n;k_0}$, $s^{f_1} \in S_{n;su;k_0}$ *and* $s^{f_2} \in S_{n;su;k_0}$ *be given so that* $s^{f_2}$ *is a direct successor of* $s^{f_1}$. *Furthermore, let* $s^{l_1} = s^*(s^{f_1}, w)$ *and* $s^{l_2} = s^*(s^{f_2}, w)$.

*Then, in general, the following relationships are possible:*

1. $s^{l_2}$ *is a direct successor of* $s^{l_1}$ *or*
2. $s^{l_2} = s^{l_1}$. *Regardless of 1. and 2.,*
3. $s^{l_2}$ *is a direct predecessor of* $s^{l_1}$, *is also possible.*

(See Lemma 3.25 and the proof in [22], pages 107, 108 for the proof.)

**Theorem 4.6.5.** *Let a partially ordered set of restricted partitions* $(S_{n;su;k_0}, \rightarrow) = (\{s^1, s^2, \cdots, s^r\}, \rightarrow)$ *and the corresponding PRM* $P^* = (p^*_{fl})_{\substack{l=1,\cdots,r \\ f=1,\cdots,r}}$ *be given. Furthermore, let* $s^{f_1}, s^{f_2}, \cdots, s^{f_q}$ *be a main minimal chain (see Definition 4.1.3).*

*Then the dominance condition is fulfilled in relation to the main minimal chain, which means*

$$\sum_{l=1}^{\bar{r}} p^*_{f_1 l} \geq \sum_{l=1}^{\bar{r}} p^*_{f_2 l} \geq \cdots \geq \sum_{l=1}^{\bar{r}} p^*_{f_q l} \quad \forall\, \bar{r} = 1, \cdots, r.$$

(See Lemma 3.26 and the proof in [22], page 109 for the proof. Thereby, the 3rd relationship from Lemma 4.6.4 is not possible in relation to main minimal chains.)

**Definition 4.6.4.** *If a main minimal chain of a set of restricted partitions* $(S_{n;su;k_0}, \rightarrow) = (\{s^1, s^2, \cdots, s^r\}, \rightarrow)$ *includes all partitions, then this set of restricted partitions is called m-totally ordered.*

Theorem 4.6.5 and Corollary 4.6.3 then imply:

**Corollary 4.6.6.** *Let* $S_{n;su;k_0}$ *be a m-totally ordered set of partitions.*

*Then the dominance condition is fulfilled for the corresponding PRM and the solutions of the corresponding Poisson equations are monotone (with respect to the partial order).*

**Examples:**

- The dominance property is fulfilled for $P^*$ from Example 4.2.1.
- All $S_{2;su;k_0}$ are m-totally ordered sets.

#### 4.6.2.3 PRMs with regard to Sets of Restricted Partitions with at most 4 Partitions

We initially list sets of restricted partitions with at most 4 partitions (without detailed proofs):

Sets with 2 partitions:

- are, obviously, m-totally ordered sets (see Definition 4.6.4).

  (Thus, the dominance condition is fulfilled for corresponding PRMs according to Corollary 4.6.6.)

Sets with 3 partitions:

- m-totally sets,

- $S_{3;4;3}$,

- $S_{3;5;3}$ this set is equivalent to $S_{3;4;3}$ with regard to the Poisson equation (if the probability functions of $w$ are in correspondence).

Sets with 4 partitions:

- m-totally ordered sets,

- $S_{3;4;4}$, $S_{3;5;4}$, $S_{4;5;3}$, $S_{4;4;3}$ (see C2) of the classification in Section 4.4.1),

- $S_{n;4;3}$ with $n > 4$ (see C1) of the classification in Section 4.4.1),

- $S_{3;3k_0-4;k_0}$ with $k_0 > 4$ (see C3) of the classification in Section 4.4.1),

- sets which are equivalent to $S_{3;4;4}$ or to other above mentioned sets with regard to the Poisson equations (where the probability functions of $w$ are in correspondence).

Since we have computed PRMs for

- $S_{3;4;3}$ $(S_{3;5;3})$ for any probability function of $w$ and

- $S_{3;4;4}, \cdots$ for discrete uniformly distributed requirements

and the solutions of corresponding the Poisson equations, we can give the following partial result (here without detailed computations).
(See also Example 4.4.1.)

**Theorem 4.6.7.** *Solutions of Poisson equations are in general monotone for PRMs with regard to sets of restricted partitions with at most 3 partitions and with regard to sets with 4 partitions in the case of discrete uniformly distributed requirements.*

### 4.6.3 The Poisson Equations with regard to Sets of Sparse Partitions with Sufficiently Large $n$

Sets of sparse partitions were characterized in Section 4.4.1, see classification C1). Limits of corresponding PRMs were computed in Section 4.4.2.

**Example 4.6.1.** *(Continuation of Example 4.4.1)*
*Let the sets of partitions* $S_{n;4;3}$ $(n > 4)$ *and the corresponding sets of requirements* $B_{n;3}$ *be given and let* $w_i$, $(i = 1, \cdots, n)$ *be independent and identically distributed. The corresponding PRMs* $P^*(n)$ *can be found in Example 4.4.1.*

$\gamma'$ *can be calculated by means of (4.6.2) and (4.6.3) (see also Theorem 3.4.1(e)).* $\gamma' = (3, 2, 1, 0)^T$ *results from such a calculation for any* $n > 4$ .

*It should be noted that the Poisson equations have monotone solutions. The limits of the variables* $\nu = \nu(n)$ *are:*

$$\begin{array}{ll} v_1 & \rightarrow \frac{14}{3} \\ v_2 & \rightarrow 4 \\ v_3 & \rightarrow 2 \\ v_4 & \rightarrow 0 \end{array}$$

$for\ \ n \rightarrow \infty,\ if\ v_4(n) := 0,\ for\ all\ n > 4.$

In this section we consider equation systems similar to the Poisson equations where PRMs with regard to sets of sparse partitions are initially replaced by the limits of such matrices, as $n$ approaches infinity. The solutions of such equation systems will be vectors which include, in relation to the distribution of requirements, generalized harmonic numbers. The solutions are also monotone, from which the monotonicity of the solutions of the Poisson equations themselves, with regard to sets of sparse partitions for sufficiently great $n$, follows.

In the considerations in this subsection, we also use the definition of monotone successors and principle parts of partitions (Definition 4.4.1) as we have used these for the computation of the limits of PRMs with regard to sets of sparse partitions in Section 4.4.2. Since monotone successors and principle parts do not directly depend on $su$ and $k_0$ we can include all equation systems (in relation to different $su$ and $k_0$) in one proof.

Now, we give a affine transformation of $\gamma$, which directly depends only on the principal parts of the partitions and the distribution of the requirements:

$$\begin{aligned}\gamma(s) &= \sum_{i=1}^{n} \sum_{w_i=0}^{s_i} (s_i - w_i)\, q_0(w_i) \qquad \text{(see (4.6.2))} \\ &= \sum_{i=1}^{\eta} \sum_{w_i=0}^{s_i} (s_i - w_i)\, q_0(w_i) + (su - \sum_{i=1}^{\eta} s_i)\, q_0(0) \qquad (4.6.5) \\ &= \sum_{i:s_i\geq 2} \sum_{w_i=\mathbf{1}}^{s_i} (s_i - w_i)\, q_0(w_i) + su\, q_0(0),\end{aligned}$$

with $\eta$ as in Definition 4.4.1(a) and $su\, q_0(0)$ is independent of s.
We define

$$\begin{aligned}\gamma'(s) &:= \sum_{i:s_i\geq 2} \sum_{w_i=1}^{s_i} (s_i - w_i)\, q_0(w_i) \\ &= \sum_{i:s_i\geq 2} \sum_{w_i=1}^{s_i-1} (s_i - w_i)\, q_0(w_i),\end{aligned} \qquad (4.6.6)$$

in particular,

$$\gamma'((1,\cdots,1,0,\cdots,0)^T) := 0$$

and

$$\gamma'(s_i) := \sum_{w_i=1}^{s_i} (s_i - w_i) q_0(w_i) \quad \text{for } s_i \geq 2. \qquad (4.6.6a)$$

If we now consider the equation systems

$$g \begin{pmatrix} -1 \\ \vdots \\ -1 \end{pmatrix} + (\lim_{n\to\infty} P^*(n) - I)\, \nu = -\gamma' \qquad (4.6.7)$$

with $\gamma'$ from (4.6.6),

then $g = 0$ initially follows from the last equation (see also Corollary 4.4.7), and we see in the following theorem that the solutions
$\nu$ $(= \nu(s),\, s \in S_{n;su;k_0})$ only depend on the principal parts of the partitions

and the distributions of the requirements. The sets $S_{n;su;k_0}$ themselves are therefore insignificant for the calculations.

**Theorem 4.6.8.** *Let $S_{n;su;k_0}$, $n = (su,)\ su+1,\ su+2, \cdots$ be sets of sparse partitions, each of these sets with $r$ partitions*[19]*, and with the partial order as in Definition 4.1.1(a), (c) and Lemma 4.1.5.*

*Let $B_{n;k_0}$ be corresponding sets of requirements with the same marginal probability functions $q_0$ for all $n$ and where the requirements $w_i$, $(i = 1, \cdots, n)$ are independent and identically distributed and where (4.2.2) is additionally assumed. Furthermore, let $P^*(n)$ be the the corresponding PRMs.*

*Then,*

$$\nu((1,1,\cdots,1,0,0,\cdots,0)^T) := 0$$

$$\nu(s) = \sum_{i:s_i\geq 2} \left( \frac{q_0(1)}{q_0(0)+q_0(1)} + \frac{q_0(1)+q_0(2)}{q_0(0)+q_0(1)+q_0(2)} + \cdots + \frac{q_0(1)+q_0(2)+\cdots+q_0(s_i-1)}{q_0(0)+q_0(1)+\cdots+q_0(s_i-1)} \right),$$

$$s \in S_{n;su;k_0},\ s \neq (1,1,\cdots,1,0,0,\cdots,0)^T$$

$$= \sum_{i:s_i\geq 2} \left( s_i - q_0(0) \left( \frac{1}{Q_0(0)} + \frac{1}{Q_0(1)} + \cdots + \frac{1}{Q_0(s_i-1)} \right) \right) \text{ where}$$

$$Q_0(\omega) = q_0(0) + q_0(1) + \cdots + q_0(\omega)$$

*together with a corresponding value of $g$ are solutions of the equation system (4.6.7).*

Briefly, we note that $\frac{1}{Q_0(0)} + \frac{1}{Q_0(1)} + \cdots + \frac{1}{Q_0(S_i-1)}$ are, in relation to the distribution of requirements, generalized harmonic numbers.

***Proof.***

1. We use the following notations and definitions

$$Q_0(\omega) := q_0(0) + q_0(1) + \cdots + q_0(\omega),$$

$$Q_1(\omega) := \qquad q_0(1) + \cdots + \quad q_0(\omega)\ (= Q_0(\omega) - q_0(0)),$$

$$\bar{Q}_0(\omega) := q_0(\omega) + q_0(\omega+1) + \cdots + q_0(k_0)\ (= 1 - Q_0(\omega-1)),$$

$$\nu(s_i) := \sum_{\omega=1}^{s_{i-1}} \frac{Q_1(\omega)}{Q_0(\omega)} \text{ for } s_i \geq 2,\ \nu(0) := 0,\ \nu(1) := 0$$

[19] See Remarks 4.4.1.

(hence $\nu(s) = \sum_{i:s_i \geq 2} \nu(s_i) = \sum_{i=1}^{n} \nu(s_i)$),

$\wp\,(M) -$ the power set of a set $M$

$\wp := \wp(\{1, 2, \cdots, \eta\})$,

$\wp^- := \wp \backslash \{1, 2, \cdots \eta\}$,

$\wp_j := \wp(\{1, 2, \cdots, \eta\} \backslash \{j\})$,

$MS(s) := \{s^2 \in S_{n;su;k_0} \mid s^2$ is a monotone successor of $s\}$.

Furthermore, we note that
if $\bar{I} \in \wp$ then $\{1, 2, \cdots, \eta\} \backslash \bar{I} := \{j_1, j_2, \cdots, j_h\}$ with $h = \eta - |\bar{I}|$.

2. Now, we give 4 relationships which are employed in the following proof:
Without loss of generality, let be $s_i \geq 2$ for $i = 1, 2, \cdots, \eta$
and $s_j \in \{0, 1\}$ for $j = \eta + 1, \eta + 2, \cdots, n$.

$$1 = \sum_{l=1}^{r} \lim_{n \to \infty} p^*(s^l|s)$$

$$= \sum_{s^2 \in MS(s)} (q_0(0) + q_0(1))^{\eta - \eta_2} \sum_{s^2_\pi \in S^2_\pi} \Big( \prod_{i:s_i > s^2_{\pi_i} \geq 2} q_0(s^2_{\pi_i}) \Big) \Big( \prod_{i:s_i = s^2_{\pi_i} \geq 2} \bar{Q}_0(s_i) \Big)$$

(see Theorem 4.4.6)

These summations over **all** monotone successors of $s$ can however be realized in the following way:

$$= \sum_{\bar{I} \in \wp} \Big( \prod_{i \in \bar{I}} \bar{Q}_0(s_i) \Big) \sum_{s^2_{\pi_{j_1}} = 0}^{s_{j_1} - 1} \cdots \sum_{s^2_{\pi_{j_h}} = 0}^{s_{j_h} - 1} \prod_{h'=1}^{h} q_0(s^2_{\pi_{j_{h'}}})$$

(which means $s_i = s^2_{\pi_i} \geq 2$ for $i \in \bar{I}$ and
$s_j > s^2_{\pi_j} \wedge s_j \geq 2$ for $j \in \{1, 2, \cdots, \eta\} \backslash \bar{I}$)

$$= \sum_{\bar{I} \in \wp} \Big( \prod_{i \in \bar{I}} \bar{Q}_0(s_i) \Big) \Big( \prod_{j \notin \bar{I}} Q_0(s_j - 1) \Big) \qquad (*1)$$

where $\prod_{i \in \bar{I}} \bar{Q}_0(s_i) \quad := 1$ for $\bar{I} = \varnothing$,

$\prod_{j \notin \bar{I}} Q_0(s_j - 1) := 1$ for $\bar{I} = \{1, 2, \cdots, \eta\}$.

$$\begin{aligned}
&\sum_{\bar{I}\in\wp^-} \left(\prod_{i\in\bar{I}} \bar{Q}_0(s_i)\right) \left[\sum_{j\notin\bar{I}} \left(\prod_{j'\notin\bar{I}\cup\{j\}} Q_0(s_{j'}-1)\right)\gamma'(s_j)\right] \\
&= \sum_{j=1}^{\eta} \left[\gamma'(s_j) \sum_{\bar{I}\in\wp_j} \left(\prod_{i\in\bar{I}} \bar{Q}_0(s_i)\right) \left(\prod_{j'\notin\bar{I}\cup\{j\}} Q_0(s_{j'}-1)\right)\right] \\
&= \sum_{j=1}^{\eta} \gamma'(s_j)\cdot 1 \qquad \text{(see (*1), applied to } s \text{ with the principal parts } (s_1,\cdots,s_{j-1},s_{j+1},\cdots,s_\eta)) \\
&= \gamma'(s). \qquad (*2)
\end{aligned}$$

$$\lim_{n\to\infty} p^*(s|s) = \prod_{i:s_i\geq 2} (q_0(s_i)+\cdots+q_0(k_0)) = \prod_{i:s_i\geq 2} \bar{Q}_0(s_i) \qquad (*3)$$

(see Theorem 4.4.6).

$$\begin{aligned}
\gamma'(s_i) &= \sum_{\omega=1}^{s_i} (s_i-\omega)\, q_0(\omega) \qquad (\textit{see } (4.6.6a)) \\
&= \sum_{\beta=1}^{s_i-1} \sum_{\omega=1}^{\beta} q_0(\omega) \qquad \text{(reorganization of the sum)} \\
&= \sum_{\beta=1}^{s_i-1} \left(\sum_{\omega=0}^{\beta} q_0(\omega)\right) \frac{Q_1(\beta)}{Q_0(\beta)} \\
&= \left(\sum_{\omega=0}^{s_i-1} q_0(\omega)\right) \sum_{\beta=1}^{s_i-1} \frac{Q_1(\beta)}{Q_0(\beta)} - \sum_{\beta=1}^{s_i-2} \left(\sum_{\omega=\beta+1}^{s_i-1} q_0(\omega)\right) \frac{Q_1(\beta)}{Q_0(\beta)} \\
&= Q_0(s_i-1)\,\nu(s_i) - \sum_{\omega=2}^{s_i-1} q_0(\omega)\,\nu(\omega) \qquad (*4)
\end{aligned}$$

(reorganization of the last sums; $\nu(\cdot)$ as in this theorem).

## 3. The proof of the Theorem

(4.6.7) together with Corollary 4.4.7 yields

$$\nu(s) = \frac{1}{1-\lim_{n\to\infty} p^*(s|s)} \left(\gamma'(s) + \sum_{s^2\in MS(s)\setminus\{s\}} \lim_{n\to\infty} p^*(s^2|s)\,\nu(s^2)\right).$$

An equivalent equation is

$$0 = \gamma'(s) - \nu(s) + \sum_{s^2 \in MS(s)} \lim_{n\to\infty} p^*(s^2|s)\, \nu(s^2).$$

Further equivalent transformations, using the summation over **all** monotone successors as in the derivation of (*1), yield

$$0 = \gamma'(s) - \nu(s) + \sum_{\bar{I}\in\wp} \left(\prod_{i\in\bar{I}} \bar{Q}_0(s_i)\right) \left[\sum_{s^2_{j_1}=0}^{s_{j_1}-1} \cdots \sum_{s^2_{j_{h-1}}=0}^{s_{j_{h-1}}-1} \sum_{s^2_{j_h}=0}^{s_{j_h}-1} \left(\prod_{h'=1}^{h} q_0(s^2_{j_{h'}})\right) \left(\sum_{i\in\bar{I}} \nu(s_i) + \sum_{h'=1}^{h} \nu(s^2_{j_{h'}})\right)\right]$$

$((q_0(0)+q_0(1))^{\eta-\eta_2}$ from Theorem 4.4.6 is realized by $s^2_{j_{h'}} = 0; 1$ for $h' = 1, \cdots, h)$

where $\prod_{i\in\bar{I}} \bar{Q}_0(s_i) := 1$ for $I = \varnothing$.

We now apply (*4) in the form of

$$\sum_{s^2_{j_{h'}}=0}^{s_{j_{h'}}-1} q_0(s^2_{j_{h'}})\, \nu\,(s^2_{j_{h'}}) = -\gamma'(s_{j_{h'}}) + Q_0(s_{j_{h'}} - 1)\, \nu\,(s_{j_{h'}})$$

(with $\nu(\cdot)$ as in this theorem)

to the above equation (at first for $h' = h$):

$$0 = \gamma'(s) - \nu(s) + \sum_{\bar{I}\in\wp} \left(\prod_{i\in\bar{I}} \bar{Q}_0(s_i)\right) \left[\sum_{s^2_{j_1}=0}^{s_{j_1}-1} \cdots \sum_{s^2_{j_{h-1}}=0}^{s_{j_{h-1}}-1} \left[\left(\prod_{h'=1}^{h-1} q_0(s^2_{j_{h'}})\right) Q_0(s_{j_h} - 1) \left(\sum_{i\in\bar{I}} \nu(s_i) + \sum_{h'=1}^{h-1} \nu(s^2_{j_{h'}})\right) + \prod_{h'=1}^{h-1} q_0(s^2_{j_{h'}})(-\gamma'(s_{j_h}) + Q_0(s_{j_h} - 1)\nu(s_{j_h}))\right]\right]$$

$$0 = \gamma'(s) - \nu(s) + \sum_{\bar{I}\in\wp}\left(\prod_{i\in\bar{I}}\bar{Q}_0(s_i)\right)\left[\sum_{s^2_{j_1}=0}^{s_{j_1}-1}\cdots\sum_{s^2_{j_{h-1}}=0}^{s_{j_{h-1}}-1}\right.$$
$$\left[\left(\prod_{h'=1}^{h-1} q_0(s^2_{j_{h'}})\right) Q_0(s_{j_h}-1)\left(\sum_{i\in\bar{I}U\{j_h\}}\nu(s_i) + \sum_{h'=1}^{h-1}\nu(s^2_{j_{h'}})\right)\right.$$
$$\left.\left. - \prod_{h'=1}^{h-1} q_0(s^2_{j_{h'}})\gamma'(s_{j_h})\right]\right].$$

The repeated application of (*4) (for $h' = h-1$) yields

$$0 = \gamma'(s) - \nu(s) + \sum_{\bar{I}\in\wp}\left(\prod_{i\in\bar{I}}\bar{Q}_0(s_i)\right)\left[\sum_{s^2_{j_1}=0}^{s_{j_1}-1}\cdots\sum_{s^2_{j_{h-2}}=0}^{s_{j_{h-2}}-1}\left[\left(\prod_{h'=1}^{h-2} q_0(s^2_{j_{h'}})\right)\right.\right.$$
$$\prod_{h'=h-1}^{h} Q_0(s_{j_{h'}}-1)\left(\sum_{i\in\bar{I}U\{j_h\}}\nu(s_i) + \sum_{h'=1}^{h-2}\nu(s^2_{j_{h'}})\right)$$
$$- \prod_{h'=1}^{h-2} q_0(s^2_{j_{h'}})Q_0(s_{j_{h-1}}-1)\gamma'(s_{jh}) + \prod_{h'=1}^{h-2} q_0(s^2_{j_{h'}})Q_0(s_{j_h}-1)(-\gamma'(s_{j_{h-1}})$$
$$\left.\left. + Q_0(s_{j_{h-1}}-1)\nu(s_{j_{h-1}}))\right]\right]$$

$$0 = \gamma'(s) - \nu(s) + \sum_{\bar{I}\in\wp}\left(\prod_{i\in\bar{I}}\bar{Q}_0(s_i)\right)\left[\sum_{s^2_{j_1}=0}^{s_{j_1}-1}\cdots\sum_{s^2_{j_{h-2}}=0}^{s_{j_{h-2}}-1}\left[\left(\prod_{h'=1}^{h-2} q_0(s^2_{j_{h'}})\right)\right.\right.$$
$$\prod_{h'=h-1}^{h} Q_0(s_{j_{h'}}-1)\left(\sum_{i\in\bar{I}U\{j_h,j_{h-1}\}}\nu(s_i) + \sum_{h'=1}^{h-2}\nu(s^2_{j_{h'}})\right)$$
$$\left.\left. - \prod_{h'=1}^{h-2} q_0(s^2_{j_{h'}})Q_0(s_{j_{h-1}}-1)\gamma'(s_{j_h}) - \prod_{h'=1}^{h-2} q_0(s^2_{j_{h'}})Q_0(s_{j_h}-1)\gamma'(s_{j_{h-1}})\right]\right]$$

$$\vdots$$

$$0 = \gamma'(s) - \nu(s) + \sum_{\bar{I}\in\wp}\left(\prod_{i\in\bar{I}}\bar{Q}_0(s_i)\right)\prod_{j\notin\bar{I}}Q_0(s_j-1)\nu(s) -$$

$$\sum_{\bar{I}\in\wp^-}(\prod_{i\in\bar{I}}\bar{Q}_0(s_i))\left[\sum_{j\notin\bar{I}}(\prod_{j'\notin\bar{I}U\{j\}}Q_0(s_{j'}-1))\gamma'(s_j)\right].$$

Finally, (*1) implies

$$0 = \gamma'(s) - \sum_{\bar{I}\in\wp^-}(\prod_{i\in\bar{I}}\bar{Q}_0(s_i))\left[\sum_{j\notin\bar{I}}(\prod_{j'\notin\bar{I}U\{j\}}Q_0(s_{j'}-1))\gamma'(s_j)\right].$$

This is a valid equation, as seen by (*2).

Thus, the assertion of the Theorem is confirmed since we have only used equivalent transformations of the equations. ■

**Corollary 4.6.9.** *Let the same assumptions as in Theorem 4.6.8 be valid. Furthermore, let $s^l$ be a successor of $s^f$, where $\{s^f, s^l\} \subseteq S_{n;su;k_0}$ for any $n \underset{(=)}{>} su$ and $s^f \neq s^l$.*

*Then, $\nu(s^l) < \nu(s^f)$ is valid for solutions of the equation system (4.6.7). This means that the solutions of the equation systems (4.6.7) are monotone (with respect to the partial order).*

This statement results from simple computations using the formulas from Theorem 4.6.8 for $\nu(s^l)$ and $\nu(s^f)$, where the inequality is initially shown if $s^l$ is a direct successor of $s^f$.

**Corollary 4.6.10.** *Let the same assumptions as in Theorem 4.6.8 be valid.*

*Then, the solutions of the Poisson equations (see (4.6.1)) with regards to sets $S_{n;su;k_0}$ of sparse partitions are monotone for sufficiently large $n$.*

This result follows from the solution behavior of linear equation systems for cases of passing to the limits in the coefficient matrices and Corollary 4.6.9 (where it is used in particular that the inequalities $\nu(s^l) < \nu(s^f)$ are strict inequalities).

**Corollary 4.6.11.** *Let the same assumptions as in Theorem 4.6.8 be valid. This means, in particular, that sets $S_{n;su;k_0}$, $n \underset{(=)}{>} su$ of sparse partitions, each of these sets with $r$ partitions, are considered.*

*The corresponding Poisson equations, with exception of perhaps a finite number of them, then have monotone solutions.*

### 4.6.4 The Poisson Equations with regard to Sets of Non-Truncated Heavy Partitions with Sufficiently Large $n$ or $k_0$

Sets of non-truncated heavy partitions were characterized in Section 4.4.1, see classification C3b.1) and Lemma 4.4.1(b). Limits of corresponding PRMs were computed in Section 4.4.3.

Formally, we proceed as in Section 4.6.3. Differences are found in the following:

In Section 4.6.3 we used $\lim\limits_{n\to\infty} P^*(n)$ in order to show the monotonicity of $\nu_f$ (for large n), which are solutions of the equation systems (4.6.1) (see Theorem 4.6.8 and Corollary 4.6.10).

We now begin with

$$g' \begin{pmatrix} -1 \\ \vdots \\ -1 \end{pmatrix} + \lim_{k_0\to\infty} \frac{1}{c(k_0)} \left(P^*(k_0) - I\right) \nu = - \lim_{k_0\to\infty} \frac{1}{c(k_0)} \gamma' \tag{4.6.8}$$

(with $g' = \lim\limits_{k_0\to\infty} \frac{1}{c(k_0)} g,\ g = g(k_0)$)

by reason of (4.4.5).

The definition of the restricted monotone successor (Definition 4.4.2(b)) includes an additional property in comparison with Definition 4.4.1(b).

Although $\lim\limits_{n\to\infty} P^*(n) - I \neq \lim\limits_{k_0\to\infty} \frac{1}{c(k_0)}(P^*(k_0) - I)$ (see Example 4.4.3)

the solutions $\nu$ of the equations systems (4.6.7) and (4.6.8) are analogous (see Theorem 4.6.8 and Theorem 4.6.12).

As in Section 4.6.3, we initially give a affine transformation of $\gamma$ (see (4.6.2)):

$$\gamma(s) = \sum_{i=1}^{n} \sum_{w_i=0}^{s_i} (s_i - w_i)\, q_0(w_i)$$

$$= \sum_{i=1}^{n} \sum_{w_i=0}^{k_0} (s_i - w_i)\, q_0(w_i)$$

$$- \sum_{i=\eta}^{n} \sum_{w_i=s_i+1}^{k_0} (s_i - w_i)\, q_0(w_i) + (\bar{su} - \sum_{i=\eta}^{n} (k_0 - s_i))\, q_0(k_0)$$

(see Definition 4.4.2 for $\eta$; $\bar{su} := n\, k_0 - su$),

$$= \sum_{i=1}^{n} s_i \sum_{w_i=0}^{k_0} q_0(w_i) - \sum_{i=1}^{n} \sum_{w_i=0}^{k_0} w_i\, q_0(w_i)$$

$$+ \sum_{i=\eta}^{n} \sum_{w_i=s_i+1}^{k_0-1} (w_i - s_i)\, q_0(w_i) + \sum_{i=\eta}^{n} (k_0 - s_i)\, q_0(k_0)$$

$$+ \bar{su}\, q_0(k_0) - \sum_{i=\eta}^{n} (k_0 - s_i)\, q_0(k_0)$$

$$= su - \sum_{i=1}^{n} \sum_{w_i=0}^{k_0} w_i\, q_0(w_i)$$

$$+ \sum_{i=\eta}^{n} \sum_{w_i=s_i+1}^{k_0-1} (w_i - s_i)\, q_0(w_i) + \bar{su}\, q_0(k_0)$$

$$= \sum_{i:s_i \le k_0-2} \sum_{w_i=s_i+1}^{k_0-1} (w_i - s_i)\, q_0(w_i) + \bar{R}(n, su, k_0, q),$$

where $\bar{R}(n, su, k_0, q) = su - \sum_{i=1}^{n} \sum_{w_i=0}^{k_0} w_i\, q_0(w_i) + \bar{su}\, q_0(k_0)$ is independent of s.

We define

$$\gamma'(s) := \sum_{i:s_i \le k_0-2} \sum_{w_i=s_i+1}^{k_0-1} (w_i - s_i)\, q_0(w_i) \tag{4.6.9}$$

$$(= \sum_{i:\bar{s_i} \ge 2} \sum_{\bar{w_i}=1}^{\bar{s_i} \text{ or } (\bar{s_i}-1)} (\bar{s_i} - \bar{w_i})\, q_0(k_0 - \bar{w_i}) \tag{4.6.9a}$$

where $\bar{s_i} := k_0 - s_i)$

and in particular,

$$\gamma'((k_0, \cdots, k_0, k_0 - 1, \cdots, k_0 - 1)^T) := 0.$$

Using $\gamma'(s)$ from (4.6.9) in the equation system (4.6.8), $g' = 0$ follows from the last equation (the equation with $\gamma'(s^r = (k_0, \cdots, k_0, k_0 - 1, \cdots, k_0 - 1)^T)$ on the right side) where Corollary 4.4.11 is kept in mind.

If we fix $\nu(s^r) = 0$, the following system remains

$$\sum_{l=1}^{r-1} \lim_{k_0 \to \infty} \frac{1}{c(k_0)} (p^*(s^l|s^f) - \delta(s^l, s^f))\, \nu(s^l) = - \lim_{k_0 \to \infty} \frac{1}{c(k_0)}\, \gamma'(s^f), \tag{4.6.10}$$

$$f = 1, \cdots, r-1.$$

**Example 4.6.2.** *(Continuation of Example 4.4.3)*
*Let us consider*

$S_{n;4;4}$ $(n > 4)$ – *sets of sparse partitions and*

$S_{5;5k_0-4;k_0}$ $(k_0 > 4)$ – *sets of non-truncated heavy partitions.*

*Limits of elements of the corresponding PRMs can be completely different (see Example 4.4.3).*

*However corresponding $\nu$ - solutions of the equation systems (4.6.7) and (4.6.8) - are analogous:*

*For instance,*

$\nu(s^4(n)) = \frac{q_0(1)}{q_0(0)+q_0(1)}$ *and* $\nu(s^4(k_0)) = \frac{q_0^0(k_0-1)}{q_0^0(k_0)+q_0^0(k_0-1)}$

*and*

$\nu(s^3(n)) = \frac{2q_0(1)}{q_0(0)+q_0(1)}$ *and* $\nu(s^4(k_0)) = \frac{2q_0^0(k_0-1)}{q_0^0(k_0)+q_0^0(k_0-1)}$

*(where* $q_0^0(w_i) := \lim_{k_0\to\infty} \frac{q_0^{k_0}(w_i)}{c(k_0)}$ *for corresponding* $c(k_0)$*, see (4.4.6)).*

*This is yielded by Theorem 4.6.8 and the following Theorem 4.6.12.*

**Theorem 4.6.12.** *Let* $S_{n;su=nk_0-\bar{su};k_0}$, $k_0 = (\bar{su},)\ \bar{su}+1, \bar{su}+2, \cdots$ $(\bar{su} < n)$ *be sets of non-truncated heavy partitions*[20]*, each of these sets with* $r$ *partitions, and with the partial order as in Definition 4.1.1(a), (c) and Lemma 4.1.5.*

*Furthermore, let* $B_{n;k_0}$ *be the corresponding sets of requirements, where for any* $k_0$*, the requirements* $w_i$, $(i = 1, \cdots, n)$ *are independent and identically distributed and where (4.2.2) is assumed. In addition, let corresponding given probability functions* $q^{k_0}$ *fulfill (4.4.5) and (4.4.6) for certain* $c(k_0)$*. Finally, let* $P^*(k_0)$ *be the the corresponding PRMs.*

*Then,*

$$\nu((k_0, \cdots, k_0, k_0-1, \cdots, k_0-1)^T) := 0$$

$$\nu(s) = \sum_{i:s_i\le k_0-2} \Big(\frac{q_0^0(k_0-1)}{q_0^0(k_0)+q_0^0(k_0-1)} + \frac{q_0^0(k_0-1)+q_0^0(k_0-2)}{q_0^0(k_0)+q_0^0(k_0-1)+q_0^0(k_0-2)} + \cdots$$
$$+ \frac{q_0^0(k_0-1)+q_0^0(k_0-2)+\cdots+q_0^0(s_i+1)}{q_0^0(k_0)+q_0^0(k_0-1)+\cdots+q_0^0(s_i+1)}\Big)$$

$$\text{(where } q_0^0(w_i) := \lim_{k_0\to\infty} \frac{q_0^{k_0}(w_i)}{c(k_0)}\text{)}$$

$$= \sum_{i:s_i\le k_0-2} \Big(k_0 - s_i - q_0^0(k_0)\big(\frac{1}{Q_0^0(k_0)} + \frac{1}{Q_0^0(k_0-1)} + \cdots + \frac{1}{Q_0^0(s_i+1)}\big)\Big)$$

$$\text{where } Q_0^0(w) = q_0^0(k_0) + q_0^0(k_0-1) + \cdots + q_0^0(w)$$

*are solutions of the equation systems (4.6.8) with* $\gamma'$ *from (4.6.9).*

***Proof.*** We use the following notations in relation to partitions:

$$s(k_0) = (k_0, \cdots, k_0, k_0-1, \cdots, k_0-1, k_0-\bar{s}_\eta, \cdots, k_0-\bar{s}_n)$$

$$\text{with } 2 \le \bar{s}_\eta \le \bar{s}_{\eta+1} \le \cdots \le \bar{s}_n \le k_0-1 \qquad (*1)$$

(see Definition 4.4.2(a) for $\eta$)

and additionally

$$\bar{s}_{h_j} = \bar{s}_{h_{j+1}} = \cdots = \bar{s}_{h_{j+1}-1},\ j = 1, 2, \cdots, \alpha$$

[20]See Remarks 4.4.1.

$$\text{where } \eta = h_1 < h_2 < \cdots < h_\alpha \le n := h_{\alpha+1} - 1.$$

Furthermore, we note

$$\nu(s_i) := \frac{q_0^0(k_0-1)}{q_0^0(k_0)+q_0^0(k_0-1)} + \cdots + \frac{q_0^0(k_0-1)+q_0^0(k_0-2)+\cdots+q_0^0(s_i+1)}{q_0^0(k_0)+q_0^0(k_0-1)+\cdots+q_0^0(s_i+1)}$$
$$\text{for } \ s_i \le k_0 - 2, \ \nu(k_0-1) := 0, \ \nu(k_0) := 0,$$

$$\gamma'(s_i) := \sum_{w_i=s_i+1}^{k_0-1} (w_i - s_i)\, q_0(w_i) \ \text{ for } \ s_i \le k_0 - 2$$
$$\text{(this means } i \ge \eta) \text{ (}see\text{ (4.6.9))}.$$

Since $\nu((k_0, \cdots, k_0, k_0-1, \cdots, k_0-1)^T) = 0$, we have to prove the identity (see (4.6.10))

$$\sum_{l=1}^{r-1} \lim_{k_0\to\infty} \frac{1}{c(k_0)} \left[p^*(s^l(k_0)|s(k_0)) - \delta(s^l(k_0), s(k_0))\right] \nu(s^l(k_0))$$
$$= -\lim_{k_0\to\infty} \frac{1}{c(k_0)}\, \gamma'(s(k_0)) \qquad (*2)$$

(where $\gamma'(\cdot)$ is from (4.6.9) and $\nu(\cdot)$ from this theorem)

for any partitions $s(k_0) \ne (k_0, \cdots, k_0, k_0-1, \cdots, k_0-1)^T$.

Here, $\lim_{k_0\to\infty} \frac{1}{c(k_0)} p^*(s^l(k_0)|s(k_0)) = 0$ if $s^l(k_0)$ is not a restricted monotone successor of $s(k_0)$ (see Theorem 4.4.10).

Hence, we can replace the sum over $l = 1, \cdots, r-1$ in (*2) with a sum over the narrow monotone successors of $s(k_0)$.

$s^l(k_0)$ are restricted monotone successors of $s(k_0)$

if $s^l(k_0)_i \in \{k_0, \ k_0 - 1\}$ for $i = 1, \cdots, \eta - 1$,

$s^l(k_0)_{i_0} = k_0 - \bar{s'}_{i_0} \ge k_0 - \bar{s}_{i_0} = s(k_0)_{i_0}$ for one $i_0 \in \{\eta, \cdots, n\}$ and

$s^l(k_0)_i = s(k_0)_i$ for $i \in \{\eta, \cdots, n\} \setminus \{i_0\}$

(see Definition 4.4.2(b)).

This implies the summation $\sum_{i=1}^{\alpha} \sum_{\bar{s'}_{h_i}=1}^{\bar{s}_{h_i}} \cdots$ as realization of the summation over the restricted monotone successors of $s(k_0)$.

Here the summand

$$\lim_{k_0\to\infty} \tfrac{1}{c(k_0)}\ (p^*(s(k_0)|s(k_0)) - 1)\ \nu(s(k_0))$$

$$= -\sum_{i=\eta}^{n} (q_0^0(k_0 - \bar{s}_i + 1) + \cdots + q_0^0(k_0))\ \nu(s(k_0)) \quad \text{(see Theorem 4.4.10)}$$

is separately noted down.

The summations $\sum_{i=1}^{\alpha} \sum_{\bar{s'}_{h_i}=1}^{\bar{s}_{h_i}-\mathbf{1}} \cdots$ then remain.

With this method, we can then give an equation equivalent to equation (*2) (where, additionally, the formula from Theorem 4.4.10 is used for the remaining limits):

$$-\sum_{i=\eta}^{n} (q_0^0(k_0 - \bar{s}_i + 1) + \cdots + q_0^0(k_0))\ \nu(s(k_0))$$

$$+\sum_{i=1}^{\alpha} (h_{i+1} - h_i) \left[ \sum_{\bar{s'}_{h_i}=1}^{s_{h_i}-1} q_0^0(k_0 - \bar{s'}_{h_i})\ \big(\nu(s(k_0)) - \nu(k_0 - \bar{s}_{h_i}) + \nu(k_0 - \bar{s'}_{h_i})\big) \right.$$

$$\left. + q_0^0(k_0)\ (\nu(s(k_0)) - \nu(k_0 - \bar{s}_{h_i}) + 0) \right]$$

$$= -\lim_{k_0\to\infty} \tfrac{1}{c(k_0)}\ \gamma'(s(k_0)), \qquad (*\ 3)$$

where according to the formulas for $\nu(s)$ (from this theorem) and $\nu(s_i)$ (from the above proof):

$$\nu(s(k_0)) - \nu(k_0 - \bar{s}_{h_i}) + \nu(k_0 - \bar{s'}_{h_i}) = \nu(s^l(k_0)) \ \text{ for } s^l(k_0) \text{ with}$$

$$s^l(k_0)_{h_i} = k_0 - \bar{s'}_{h_i} \text{ and } s^l(k_0)_j = s(k_0) \text{ for } j \neq h_i$$

and thus, in particular,

$$\nu(s(k_0)) - \nu(k_0 - \bar{s}_{h_i}) = \nu(s^l(k_0)) \ \text{ if } \bar{s'}_{h_i} = k_0 - 1.$$

A simple reorganization of the sums yields

$$\left\{ -\sum_{i=\eta}^{n} (q_0^0(k_0 - \bar{s}_i + 1) + \cdots + q_0^0(k_0)) \right.$$

$$+\sum_{i=1}^{\alpha}(h_{i+1}-h_i)\sum_{\bar{s'}_{h_i}=\mathbf{0}}^{\bar{s}_{h_i}-1} q_0^0(k_0-\bar{s'}_{h_i})\Bigg\}\,\nu(s(k_0))$$

$$+\sum_{i=1}^{\alpha}(h_{i+1}-h_i)\sum_{\bar{s'}_{h_i}=1}^{\bar{s}_{h_i}-1} q_0^0(k_0-\bar{s'}_{h_i})\,(-\nu(k_0-\bar{s}_{h_i})+\nu(k_0-\bar{s'}_{h_i}))$$

$$+\sum_{i=1}^{\alpha}(h_{i+1}-h_i)\,q_0^0(k_0)\,(-\nu(k_0-\bar{s}_{h_i}))$$

$$=-\lim_{k_0\to\infty}\tfrac{1}{c(k_0)}\,\gamma'(s(k_0)). \qquad (*\,4)$$

Since $\sum_{i=1}^{\alpha}(h_{i+1}-h_i)\cdots$ can be replaced with $\sum_{i=\eta}^{n}\cdots$, the term in the parentheses $\{\cdots\}$ yields 0 and the following equivalent equation remains:

$$\sum_{i=\eta}^{n}\sum_{\bar{s'}_{h_i}=1}^{\bar{s}_{h_i}-1} q_0^0(k_0-\bar{s'}_{h_i})\,(-\nu(k_0-\bar{s}_{h_i})+\nu(k_0-\bar{s'}_{h_i}))+\sum_{i=\eta}^{n} q_0^0(k_0)\,(-\nu(k_0-\bar{s}_{h_i}))$$

$$=-\lim_{k_0\to\infty}\tfrac{1}{c(k_0)}\,\gamma'(s(k_0)).$$

Finally,

$$-\sum_{i=\eta}^{n}\left\{\sum_{\bar{s'}_{h_i}=\mathbf{0}}^{\bar{s}_{h_i}-1} q_0^0(k_0-\bar{s'}_{h_i})\,\nu(k_0-\bar{s}_{h_i})-\sum_{\bar{s'}_{h_i}=1}^{\bar{s}_{h_i}-1} q_0^0(k_0-\bar{s'}_{h_i})\,\nu(k_0-\bar{s'}_{h_i})\right\}$$

$$=-\sum_{i=\eta}^{n}\left\{\lim_{k_0\to\infty}\tfrac{1}{c(k_0)}\,\gamma'(s(k_0)_i)\right\}.$$

Here $\{\cdots\}=\{\cdots\}$ is valid. This relationship is analogous to (*4) from the proof of Theorem 4.6.8 and could be proven in an analogous way.

Thus, the assertion of the Theorem is confirmed since we have only used equivalent transformations of the equations. ■

Similar to Corollary 4.6.9, in relation to sets of sparse partitions, Theorem 4.6.12 implies here:

**Corollary 4.6.13.** *Let the same assumptions as in Theorem 4.6.12 be valid. Furthermore, let $s^l$ be a successor of $s^f$, where $\{s^f, s^l\}\subseteq S_{n;su;k_0}$ for any*

$k_0 \underset{(=)}{>} s\bar{u}$ *and* $s^f \neq s^l$.

*Then,* $\nu(s^l) < \nu(s^f)$ *is valid for solutions of the equation system (4.6.8). This means that the solutions of the equation systems (4.6.8) are monotone (with respect to the partial order).*

**Corollary 4.6.14.** *Let the same assumptions as in Theorem 4.6.12 be valid.*

*Then, the solutions of the Poisson equations (see (4.6.1)) with regard to sets* $S_{n;su;k_0}$ *of non-truncated heavy partitions, are monotone for sufficiently large* $n$ *and* $k_0$.

(See also Remarks 4.4.2.)

In order to prove this corollary we use considerations analogous to those for the case of sets of sparse partitions, refer to Corollary 4.6.10.

We must additionally note that $\gamma'$ (see (4.6.9)), the limits of PRMs $P^*(k_0)$, as $k_0$ tends to infinity, and the solutions $\nu$ of the corresponding equation systems depend on the principal parts of the partitions, however not on $n$.

**Corollary 4.6.15.** *Let the same assumptions as in Theorem 4.6.12 be valid. This means, in particular, that sets* $S_{n;su=nk_0-s\bar{u};k_0}$, $k_0 \underset{(=)}{>} s\bar{u}$ *of non-truncated heavy partitions, each of these sets with* $r$ *partitions, are considered.*

*The corresponding Poisson equations, with exception of perhaps a finite number of them, then have monotone solutions.*

**Remarks:** *We have answered several important questions regarding PRMs and the corresponding Poisson equations in Chapter 4. There are, however, other problems which still remain, for instance:*

- *formulas for (most of) the elements of PRMs,*
- *the proof of the monotonicity of solutions of the Poisson equations with regard to sets of truncated heavy partitions for large* $k_0$,
- *the proof of the the monotonicity of the solutions of Poisson equations for small* $n$ *and* $k_0$ *in the case where the PRMs do not fulfil the dominance condition.*

## 4.7 Conclusion for SDDP Problems with Identical Basic Cost and Independent and Identically Distributed Requirements

**Corollary 4.7.1.** .

*(i) Let reduced SDDP problems, as in Section 3.4.2, with*

*a) m-totally ordered state spaces or*

*b) state spaces with at most 4 partitions and discrete uniformly distributed requirements in the case of exactly 4 partitions*

*be given.*

*Then, decisions $d^*$ for feasible states (unordered partitions) with least square sums of their components are optimal for such reduced SDDP problems.*

*(ii) Let reduced SDDP problems, as in Section 3.4.2, with state spaces*

$$\begin{cases} S_{n;su;k_0},\ n = (su,)\ su+1, su+2, \cdots & \textit{(sets of sparse partitions, } su \textit{ and } k_0 \textit{ arbitrary but fixed) or} \\ S_{n;su=nk_0-\bar{su};k_0},\ k_0 = (\bar{su};)\ \bar{su}+1, \bar{su}+2, \cdots\ (\bar{su} < n) & \textit{(sets of non-truncated heavy partitions, } \bar{su} \textit{ arbitrary but fixed)} \end{cases}$$

*each of these sets with* $\begin{cases} r_1 \\ r_2 \end{cases}$ *partitions, be given. Furthermore, let the probability functions of the requirements of the corresponding sets of requirements be as in Theorem 4.6.8 or Theorem 4.6.12, respectively.*

*Then, decisions $d^*$ for feasible states with least square sums of their components are optimal for such reduced SDDP problems, with exception of perhaps a finite number of them.*

***Proof.*** Matrices of transition probabilities of reduced SDDP problems for decisions $d^*$ are PRMs (see Lemma 4.2.3).

Thus, the Poisson equations for PRMs (see Definition 4.6.1 and (2.3.5)) are the Poisson equations for reduced SDDP problems and decisions $d^*$, where in addition certain affine transformations of the right hand side of the equations are allowed (see Theorem 3.4.1(c) and (4.6.2)).

However, such affine transformations (with $\alpha > 0$) have no effect on the monotonicity of the solutions of the Poisson equations (see Lemma 2.3.2).

According to Corollary 4.6.6 or to Theorem 4.6.15 the Poisson equations then have monotone solutions (with respect to the partial order, see Definition 4.6.2) in case (i).

In case (ii), the Poisson equations, with exception of perhaps a finite number of them, have monotone solutions, which here follow from Corollaries 4.6.11 and 4.6.15.

If $\hat{d}^*(s,w) = s^*$ and $s' \in \hat{A}_{n;su;k_0}(s,w)$, $s' \neq s^*$ (see (3.4.15) and (4.2.5)) then $s^*$ is a successor of $s'$ according to Lemma 4.2.1(b) and monotone solutions imply $\nu(s') < \nu(s^*)$ (see Definition 4.6.2). Hence, the optimality criterion (2.3.22a) from Lemma 2.3.6 is fulfilled and $d^*$ are optimal decisions for the corresponding reduced SDDP problems. ■

# Final observations on SDDP problems

SDDP problems are extremely complex.

The formulation of distance properties for DA stochastic dynamic programming problems in Chapter 2 and corresponding statements in this chapter and Chapter 3 led to the fact that the use of *lazy algorithms* is sufficient in order to compute optimal solutions of SDDP problems (Theorem 3.3.8).

Under the assumptions of *identical basic costs* (in other words, of *unit distances*), the average one-step reward functions of SDDP problems modelled as DA MDPs do not depend on the decisions. According to Section 2.3.2, *optimal decisions* then imply an *"almost-partial order"* of the states and the complexity of computing optimal decisions can be reduced.

In the case that in addition to *identical basic costs, identically distributed requirements* are also assumed, decisions for feasible *states with least square sums of their components are optimal* for "most" of such SDDP problems. This was shown in Chapter 4 by means of combinatorial considerations.

Based on the last facts, the question of how the optimal decisions vary if the (initially identical) basic cost change is of interest. This means a *parametric analysis* should be utilized in relation to variable cost. For this the cost are considered to be linearly dependent on one parameter.

If the parameter increases, then the violations of the optimality are *single violations* (see Definition 2.3.6) in general. This also means that an adapted Howard algorithm is a *greedy algorithm* for *cost-parametric* SDDP problems.

Moreover, the *additional conditions (AC1), (AC2) and (AC3)* have been formulated in Section 2.3.4. They are satisfied for Example 3.5.1. Whether the additional conditions are valid for SDDP problems in general, or only

for certain subsets of SDDP problems, remains to be investigated. Nevertheless, they substantiate basic *heuristic methods* for finding of approximate solutions of SDDP problems.

However, *investigations* with regard to SDDP problems and useful heuristics for SDDP problems are *far from being finished*.

## References


[1] R.K. Ahuja, T.L. Magnanti and J.B. Orlin, *Networks Flows*. In: Handbooks in Operations Research and Management Science, Vol. I. Ed. by Neuhauser, G.L.; Rinnooy Kan, A.H.G.; Todd, M.J. Elsevier Science Pub. B.V.; Amsterdam 1989.

[2] M. Aigner, *Kombinatorik, 1. Grundlagen und Zähltheorie*. Springer-Verlag, Heidelberg 1975.

[3] G.E. Andrews, *The Theory of partitions.*
Addison-Wesley Publishing Company, London, $\cdots$ 1976.

[4] K.-P. Arnold, *Stochastische Transportprobleme*. Verlag Dr. Kovač, Hamburg 1987.

[5] Y. Bartal and E. Grove, The Harmonic k-Server Algorithm Is Competive. *Journal of the ACM*, Vol. 47, No. 1 (2000) pp. 1 - 15.

[6] N. Beaumont, Some results pertaining to lumpings of a Markov chain. *Asia-Pac. J. Oper. Res. 9*, No. 1 (1992) pp. 1-8.

[7] D.P. Bertsekas, *Dynamic Programming and Optimal Control, I and II*, Athena Scientific, Belmont, Massachusetts. New York-San Francisco-London 1995.

[8] A. Borodin and R. El-Yaniv, *Online computation and competitive analysis*, University Press, Cambrigde 1998.

[9] T. Brylawski, The lattice of integer partitions. *Discrete Mathematics 6* (1973) pp. 201 - 219.

[10] D.J. Daley, Stochastic monotone Markov processes, *Z. Wahrsch. Verw. Gebiete*, **10** (1968), 305–317.

[11] W. Dinkelbach, *Entscheidungsmodelle*, Springer-Verlag, Berlin-New York 1982.

[12] S.E. Dreyfus and A.M. Law, *The Art and Theory of Dynamic Programming*, Academic Press New York-San Francisco-London 1977.

[13] V.A. Emelicev, M.M. Kovalev and M.K. Kravcov, *Polyeder-Graphen-Optimierung*. VEB Deutscher Verlag der Wissenschaften, Berlin 1985.

[14] E. Girlich und M.M. Kowaljew, *Nichtlineare diskrete Optimierung*, Akademie-Verlag, Berlin 1981.

[15] H.-J. Girlich, *Diskrete stochastische Entscheidungsprozesse*, Teubner Verlagsgesellschaft, Leipzig 1973.

[16] H.-J. Girlich, P. Köchel and H.-U. Küenle, *Steuerung dynamischer Systeme.* Birkhäuser, Basel 1990.

[17] O. Hernández-Lerma, *Adaptive Markov control Processe*, Springer-Verlag, New York-Berlin 1989.

[18] A. Hildenbrandt, Computergestützte Berechnungen und neue Erkenntnisse zu Partitionen-Bedarfs-Matrizen. *Facharbeit - Mathematik*, Goethegymnasium, Ilmenau 2002.

[19] R. Hildenbrandt, Kostengünstige Belegung von Formen im Hinblick auf Bedarfsschwankungen, *Wiss. Zeitschrift der TH Ilmenau*, **34**, Heft 4 (1988).

[20] R. Hildenbrandt, *A special stochastic decision problem*, Optimization **28** (1993) pp. 95-110.

[21] R. Hildenbrandt, Eine Verbandsstruktur für Partitionen ganzer Zahlen, *Preprint* No. M 14/94, TU Ilmenau 1994.
*http://www.db−thueringen.de/servlets/DerivateServlet/Derivate−8547/IfM_Preprint_M_94_14.PDF*

[22] R. Hildenbrandt, *Methoden aus ganzzahliger Optimierung und Verbandstheorie zur Behandlung eines stochastischen dynamischen Transportproblems*, Habilitationsschrift, TU Ilmenau 1995, (Libri BoD 2000).

[23] R. Hildenbrandt, Partitions-requirements-matrices, Symposium on Operations Research, Duisburg 9/2001 (*Proceedings*, Springer, Berlin, Heidelberg, ..., (2002) pp. 303 - 311).

[24] R. Hildenbrandt, Stochastic dynamic programming with random disturbances, *Discussiones Mathematicae Probability and Statistics*, **23** (2003) pp. 5-44.

[25] P. Kumar and R.P. Varaiya, *Stochastic Systems: Estimation, Identification, and Adaptive Control*, Prentice Hall, New Jersey 1986.

[26] A.W. Marshall and I. Olkin, *Inequalities: Theory of majorization and its applications*, Mathematics in Science and Engineering, Vol. 143. Academic Press, New York 1979.

[27] Maßmann, Matthias, *Ein Ansatz zur kapazitierten stochastisch-dynamischen Facility-Location-Planung*, Dissertation, Deutscher Universitäts-Verlag, 2006.

[28] P.H. Müller and V. Nollau, *Steuerung stochastischer Prozesse*, Akademie-Verlag, Berlin 1984.

[29] K. Neumann, *Operations-Research-Verfahren*, Band II Carl Hanser, München 1977.

[30] K. Neumann and M. Morlock, *Operations Research*, Carl Hanser Verlag München, Wien 1993.

[31] M. Puterman, *Markov decision processes.* Wiley, New York, .... (1994).

[32] C. D. Savage, Gray code sequences of partitions. *Journal of Algorithms* **10** (1989), pp. 577-595.

[33] C. Schneeweiss, *Dynamisches Programmieren*, Physica-Verlag Würzburg-Wien 1974.

[34] H.-J. Sebastian and N. Sieber, *Diskrete dynamische Optimierung*, Geest & Porting K.-G.; Leipzig 1981.

[35] W. Szwarc, The Dynamic Transportation Problem. *Mathematica 13* (1971) pp. 335-345.

[36] D.J.A. Welsh, *Matroid theory.* Academic Press, New York 1976.

[37] R. Wille, Aspects of finite Lattices. *Proceedings of the NATO Advanced Study Institute held in Berlin* (West Germany) 1976, (edited by Aigner, M.(1976). *Higher Combinatories.* D. Reidel Publishing Company, Dordrecht-Holland.)

## Glossary of Symbols and Abbreviations

| | |
|---|---|
| $\mathbb{N}$ | the set of positive integers |
| $\mathbb{R}$ | the set of real numbers |
| $\mathbb{R}^n_+$ | the set of n-dimensional real vectors - with nonnegative coordinates |
| $\mathbb{Z}$ | the set of integers |
| $\mathbb{Z}_+$ | the set of nonnegative integers |
| $\mathbb{Z}^n_+$ | the set of n-dimensional integer vectors - with nonnegative coordinates |
| $x \leq y$ for $x \in \mathbb{Z}^n_+$, $y \in \mathbb{Z}^n_+$ | means $x_i \leq y_i$ for $i = 1, ..., n$ |
| $\delta$ | Kronecker's symbol |
| $\lvert M \rvert$ | number of elements in the finite set $M$ |
| $\lfloor \frac{x}{y} \rfloor$ $(x \in \mathbb{Z}_+,\ y \in \mathbb{N})$ | the integer with $\frac{x}{y} - 1 < \lfloor \frac{x}{y} \rfloor \leq \frac{x}{y}$, |
| $\lceil \frac{x}{y} \rceil$ $(x \in \mathbb{Z}_+,\ y \in \mathbb{N})$ | the integer with $\frac{x}{y} \leq \lceil \frac{x}{y} \rceil < \frac{x}{y} + 1$ |
| $y[i_1; i_2]$ $(y \in \mathbb{Z}^n)$ | $y_i[i_1; i_2] = \begin{cases} y_i + 1 & \text{for } i = i_1, \\ y_i - 1 & \text{for } i = i_2, \\ y_i & \text{otherwise} \end{cases}$ |
| w. l. o. g. | without loss of generality |

| | | *essential locations* |
|---|---|---|
| $\dot{\rightarrow}$, $\rightarrow$ | | Section 4.1 |
| $a$ | availabilities | Section 3.1 |
| $A$ | decision space | Section 2.1 |

| | | |
|---|---|---|
| $A^M$ | set of finite decision spaces | Section 2.3<br>Section 2.3.2<br>Section 3.2 |
| $\hat{A}_t(s_t, w_t)$ | DA decision set | Section 2.1<br>Section 2.3.2<br>Section 2.3.3.1 |
| $A_{n;su;k_0}(\tilde{s}, w)$ | decision set | Section 3.2 |
| $\hat{A}_{n;su;k_0}(\tilde{s}, w)$ | DA decision set/ set of feasible ordered partition | Section 3.2<br>Section 3.3<br>Section 4.2.1 |
| $\hat{A}_{n;su;k_0}(s, w)$ | reduced DA decision set/ set of feasible (unordered) partitions | Section 3.4.2<br>Section 4.2.2 |
| $\hat{A}'_{n;su;k_0}(w)$ | extended DA decision set | Section 3.3 |
| $\hat{\hat{A}}(s, w)$ | smaller DA decision set | Section 2.3.3.1 |
| $\hat{\hat{A}}_{n;su;k_0}(\tilde{s}, w)(= \hat{A}_{n;su;k_0}(\tilde{s}, w))$ | smaller DA decision set | Section 3.3 |
| $\hat{A}(w)$ | | Section 2.3.3.1 |
| (AC1) | additional condition | Section 2.3.4.2 |
| (AC2) | additional condition | Section 2.3.4.2 |
| (AC3) | additional condition | Section 2.3.4.2 |
| $b$ | requirements | Section 3.1 |
| $B$ | disturbance space | Section 2.1<br>Section 3.2 |
| $B_{n;k_0}$ | disturbance space/set of requirements | Section 3.2<br>Section 4.2 |
| $B^*_{n,k_0}(s, s^*)$, $B^{*(f,l)}_{n;k_0}$ | set of balancing requirements | Section 4.2.2<br>Section 4.5.1 |
| $B^{1/2}_{n;k_0}(s^f, \hat{s}^l_\pi)$ | certain subsets of $B_{n;k_0}$ | Section 4.3 |
| $B^{1/2,(r,l)}_{n;k_0}$ | certain subsets of $B_{n;k_0}$ | Section 4.5.2 |

| | | |
|---|---|---|
| $\hat{c}_t(s_t, w_t, s')$ | internal costs | Section 2.1<br>Section 3.2 |
| $c^d(s^f, s^l) = c^d_{fl}$ | | Section 2.3.2 |
| $\hat{c}(s^f, s^l) = \hat{c}_{fl}$ | internal costs | Section 2.3.2<br>Section 3.2 |
| $\hat{c}(s^f, w)$ | | Section 2.3.2 |
| $C[b, su]$ | surplus-situation | Section 3.1 |
| $C[w, su]$ | surplus-situation | Section 3.2<br>Section 4.2.1 |
| $C[su, b]$ | scarcity-situation | Section 3.1 |
| $C[su, w]$ | scarcity-situation | Section 3.2<br>Section 4.2.1 |
| $d$ | decision (function) | Section 2.3 |
| $\hat{d}_t$ | DA decision function | Section 2.1<br>Section 2.3.2<br>Section 3.2 |
| $\hat{d}_t(s_t, w_t)$ | single decision | Section 2.1<br>Section 2.3.2<br>Section 3.2 |
| $d^*$ | DA decision function with decisions for feasible states with minimum average one-step reward functions/ for feasible partitions with least square sums of their parts/ optimal decision | Section 3.4.2<br>Section 4.2.2<br>Section 4.6 |
| $\hat{D}_t$ | the set of DA decision functions | Section 2.1<br>Section 3.2 |
| DA | "decision after" | Section 2.1 |
| (DAP) | basic problem of a DA model | Section 2.1 |
| (DAPa) | basic problem of a DA model in another representation | Section 2.1 |
| $(DA\bar{P}a)$ | (DAPa) under appropriate assumptions | Section 2.1 |

| | | |
|---|---|---|
| $(DA\bar{P}b)$ | DA model with infinite horizon and under appropriate assumptions | Section 2.3.2 |
| DA MDP$(N = \infty, S, A^M, P, \gamma)$ | DA Markov decision process | Section 2.3.2 |
| $DAMDP_l(N = \infty, S, A^M, P, \gamma(\vartheta))$ | cost-parametric DA Markov decision process | Section 2.3.4.2 |
| DB | "decision before" | Section 2.1 |
| $E$ | expected value | |
| $\eta$ | determines the dimension of the "principal part" | Section 4.4.2<br>Section 4.4.3<br>Section 4.5.2 |
| $F$ | policy | Section 2.1<br>Section 3.2 |
| $F$ (Sections 4.3/5) | the number of components of $s^f$ which are nonequal 0 | Section 4.3<br>Section 4.5 |
| $F_z$ | | Section 4.3<br>Section 4.5 |
| $f_t(s_t\ , \overline{w_t})$ | the optimal value function for the remaining periods | Section 2.1 |
| $G_t$ | transition function | Section 2.1 |
| $g^d$ | average expected cost per stage | Section 2.3 |
| $\gamma^d/\gamma$ | vector of average (one-step) reward functions/ right hand side of the Poisson equations in Section 4.6 | Section 2.3<br>Section 3.4.1<br>Section 4.2.2<br>Section 4.6 |
| $\triangle H^d(s^f, w, s^{\bar{l}} - s^l)$ | | Section 2.3.2.2 |
| $\triangle H^d(s^f, \bar{d}, \vartheta)$ | | Section 2.3.4.1 |
| $\triangle H^d(s^f, w, s^{\bar{l}} - s^l, \vartheta)$ | | Section 2.3.4.2 |
| $I$ | identity matrix | Section 2.3.1 |
| $I_v$ | sets of indices | Section 2.3.3.2 |

| | | |
|---|---|---|
| | | Section 3.4.2 |
| $I_i$ | intervals | Section 2.3.4.2 |
| $(J_o, j_o)$ -perturbation | | Section 4.3<br>Section 4.5.1 |
| $(J_o, j_3)$ -perturbation | | Section 4.3<br>Section 4.5.1 |
| $(J_o, j_o)$-perturbed partition | | Section 4.3<br>Section 4.5.1 |
| $(J_o, j_3)$-perturbed partition | | Section 4.3<br>Section 4.5.1 |
| $(J_o, j_o, j_1)$-perturbed permutation | | Section 4.3<br>Section 4.5.1 |
| $(J_o, j_3, j_4)$-perturbed permutation | | Section 4.3<br>Section 4.5.1 |
| $k_0$ | bounds | Section 3.2<br>Chapter 4 |
| $(k_{ij})_{\substack{i=1,\ldots,n \\ j=1,\ldots,m}}$ | basic costs (or distances) | Chapter 3 |
| $K_t$ | stage - cost (or return) function | Section 2.1 |
| $L$ | the number of components of $s^l$ which are nonequal 0 | Section 4.3<br>Section 4.5 |
| $L_y$ | | Section 4.3<br>Section 4.5 |
| (LPC) | | Section 2.3.4.2 |
| $(LPC_1)$ | | Section 2.3.4.1 |
| MDP | Markov decision process | Section 2.3 |
| MDP$(N = \infty, S, A^M, P, \gamma)$ | Markov decision process with average reward criterion | Section 2.3.2 |
| $MDP_c(N = \infty, S, A^M, P(\vartheta), \gamma(\vartheta))$ | | |

| | | |
|---|---|---|
| | parametric Markov decision process | Section 2.3.4.1 |
| $MDP_l(N=\infty, S, A, P, \gamma(\xi_0)+\vartheta\gamma^d(\xi))$ | | |
| | cost-parametric Markov decision process | Section 2.3.4.1 |
| $n$ | number of types/components | Section 3.2<br>Chapter 4 |
| $N \in \mathbb{N} \cup \{\infty\}$ | horizon | Section 2.1<br>Section 2.3.1<br>Section 3.2 |
| $P^d = (p(s^l \vert s^f, d))_{\substack{f=1,\dots,m \\ l=1,\dots,m}} = \left(p^d_{fl}\right)_{\substack{f=1,\dots,m \\ l=1,\dots,m}}$ | | |
| | matrix of transition probabilities for d | Section 2.3<br>Section 3.2 |
| $(p_f^{d,\infty})_{f=1,\dots,m}$ | stationary distribution | Section 2.3 |
| $P_{n;su;k_0}$ | general partitions-requirements-matrix | Section 4.2.1 |
| $P^* = P^*_{n;su;k_0} = (p^*_{fl})$ | | |
| | partitions-requirements-matrix | Section 4.2.2<br>to Section 4.6 |
| PRM | partitions-requirements-matrix | Section 3.3<br>Chapter 4 |
| $q$ | probabilities of the random disturbances | Section 2.3.2<br>Section 3.2<br>Section 3.4.1<br>Section 4.2<br>... |
| $q_0(w_i)$ | single probabilities | Section 4.2<br>Section 4.4 |
| $q_0^0$ | limit with regard to probabilities | Section 4.4.3<br>Section 4.6.4 |
| $r$ | number of states/ of (unordered) restricted partitions | Section 3.4.2<br>Section 4.1<br>Section 4.4 |
| $\tilde{r}$ | number of states/ of ordered restricted partitions | Section 3.2<br>Section 4.1 |

| | | |
|---|---|---|
| $R(n, su, k_0, q)$, $R_1(n; su; k_0)$ | additive remainder terms | Section 3.4.1 |
| $S$ | state space | Section 2.1<br>Section 2.3.1<br>Section 3.4.2 |
| $s$ | state/(unordered) partition | Section 2.1<br>Section 3.4.2<br>Section 4.1 |
| $\tilde{s}$ | state; number of machines in a state/ ordered partitions | Section 3.2<br>Section 4.1 |
| $\tilde{s}[i_0;\ i_1]$ | direct successor | Section 3.3<br>Section 4.2.1 |
| $S_{n;su;k_0}$ | state space/ set of (unordered) partitions | Section 3.4.2<br>Section 4.1<br>Section 4.2<br>Section 4.4 |
| $(S_{n;su;k_0}, \rightarrow)$ | lattice of restricted partitions | Section 4.1<br>Section 4.4 |
| $\tilde{S}_{n;su;k_0}$ | state space/ set of ordered partitions | Section 3.2<br>Section 4.1<br>Section 4.2 |
| $s^*(s, w)$ | feasible/feasible balanced transition | Section 4.2.2<br>Section 4.3 |
| $\tilde{s}'(\tilde{s}, w)$ | feasible transition | Section 4.2.1<br>Section 4.2.2 |
| $\sigma_y^l$ | representatives of components of states | Section 4.3<br>Section 4.5 |
| SDDP problem | problem of stochastic dynamic distance optimal partitioning | Chapter 3<br>Section 4.6 |
| SDDP' problem | extended SDDP problem | Section 3.3 |

| | | |
|---|---|---|
| $\hat{S}_{\pi}^{f,l}(J_o, j_o, j_1)$ | the set of $(J_o, j_o, j_1)$-perturbed permutations | Section 4.3<br>Section 4.5.1 |
| $\hat{S}_{\pi}^{f,l}(J_o, j_3, j_4)$ | the set of $(J_o, j_3, j_4)$-perturbed permutations | Section 4.3<br>Section 4.5.1 |
| $s_H = (s_1, \cdots, s_\eta)$ | principal part of a sparse partition | Section 4.4.2 |
| $s_H = (s_\eta, \cdots, s_n)$ | principal part of a non-truncated heavy partition | Section 4.4.3 |
| $su$ | sum/ number of machines /integer which should be partitioned | Section 3.2<br>Chapter 4 |
| $t$ | numbers of stages | Section 2.1<br>Section 2.3<br>Section 3.2 |
| $\vartheta$ | parameter | Section 2.3.4<br>Section 3.5 |
| TP | transportation problem | Section 3.1 |
| $TP^*(a, b)$ | transportation problem with distance properties | Section 3.1<br>Section 3.2 |
| $U(\tilde{s}^f, \tilde{s}^l)$ | conversion number | Section 3.3 |
| $w \in B$ | random disturbances or their realizations | Section 2.1<br>Section 3.2<br>Section 4.2 |
| $\overline{w_t} := (s_1, w_1, \ldots, w_t)$ | | Section 2.1 |
| $x \in A$ | decisions (or controls) | Section 2.1 |
| $x \in X_f(a, b)$ | feasible solutions (of the TP) | Section 3.1 |
| $x, \hat{x}$ | | Section 3.1<br>Section 3.2 |
| $x_{i\ n+1}, x_{j\ n+1}$ | slack-variables | Section 3.1 |

| | | |
|---|---|---|
| $X_f(a,b)$ | set of feasible solutions (of the TP) | Section 3.1<br>Section 3.2<br>Section 3.3 |
| $X_{f^e}(a,b)$ | extended set of feasible solutions | Section 3.1<br>Section 3.2<br>Section 3.3 |
| $X_{opt}(a,b)$ | set of optimal solutions (of the TP) | Section 3.3 |

# Index