

Jörn Frank

# Klassifikationsmethode für ein Gasdetektorenarray mit Chromatographen zur Identifikation von Gefahrstoffen



Cuvillier Verlag Göttingen  
Internationaler wissenschaftlicher Fachverlag



**Klassifikationsmethode für ein Gasetektorenarray mit Chromatographen  
zur Identifikation von Gefahrstoffen**





# **Klassifikationsmethode für ein Gasdetektorenarray mit Chromatographen zur Identifikation von Gefahrstoffen**

**Vom Promotionsausschuss der  
Technischen Universität Hamburg-Harburg  
zur Erlangung des akademischen Grades  
Doktor-Ingenieur (Dr.-Ing.)**

**genehmigte Dissertation**

von

Jörn Frank

aus

Neustadt i.H.

2016



## **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, 2016

Zugl.: (TU) Hamburg, Univ., Diss., 2016

1. Gutachter: Prof. Dr.-Ing. G. Matz

2. Gutachter: Prof. Dr.-Ing. R.-R. Grigat

Tag der mündlichen Prüfung: 3. August 2016

© CUVILLIER VERLAG, Göttingen 2016

Nonnenstieg 8, 37075 Göttingen

Telefon: 0551-54724-0

Telefax: 0551-54724-21

[www.cuvillier.de](http://www.cuvillier.de)

Alle Rechte vorbehalten. Ohne ausdrückliche Genehmigung des Verlages ist es nicht gestattet, das Buch oder Teile daraus auf fotomechanischem Weg (Fotokopie, Mikrokopie) zu vervielfältigen.

1. Auflage, 2016

Gedruckt auf umweltfreundlichem, säurefreiem Papier aus nachhaltiger Forstwirtschaft

ISBN 978-3-7369-9375-4

eISBN 978-3-7369-8375-5



## Danksagung

Die vorliegende Arbeit entstand während meiner Tätigkeit als wissenschaftlicher Mitarbeiter im Institut für Messtechnik der Technischen Universität Hamburg-Harburg.

Ganz besonderen Dank möchte ich Herrn Prof. Dr.-Ing. G. Matz für seine fachliche und persönliche Unterstützung bei der Durchführung dieser Arbeit aussprechen. Danken möchte ich zudem Herrn Prof. Dr.-Ing. R.-R. Grigat für die Übernahme des Koreferates sowie Herrn Prof. Dr.-Ing. W. Krautschneider für seine Bereitschaft zur Übernahme des Vorsitzes im Promotionsausschuss.

Ich bedanke mich bei allen Mitarbeitern des Instituts für Messtechnik für die kollegiale interdisziplinäre Zusammenarbeit und die interessanten fachlichen und nichtfachlichen Diskussionen. Ganz besonderer Dank gilt meinem geschätzten Kollegen Herrn Dipl.-Ing. Hendrik Fischer für die enge Zusammenarbeit in den dieser Arbeit zu Grunde liegenden Projekten. Sehr wertvoll waren stets die offenen Diskussionen mit Ihm, durch welche meist die richtigen Wege in den Forschungsprojekten eingeschlagen wurden. Danken möchte ich auch Herrn Dr.-Ing. Jens Eichmann, Herrn Dipl.-Ing. Matthias Feindt und Herrn Dipl.-Ing. Andreas Behn für die erquickenden informellen Gespräche, die mir häufig einen zusätzlichen Blickwinkel auf das Forschungsthema ermöglicht haben.

Ebenso gilt mein Dank allen Studenten, die an den Projekten mitgewirkt haben. Besonderer Dank gilt Herrn M.Sc. Tobias Woldert und Herrn B.Sc. Sebastian Kronenwerth.

Weiterhin danke ich den Mitarbeitern der Firma Airsense GmbH, Herrn Dr.-Ing. Bert Ungethüm und Herrn Dr.-Ing. Andreas Walte, für die fachliche Zusammenarbeit in dem Projektverbund. Den Projektmitarbeitern des Bundesamtes für Bevölkerungsschutz und Katastrophenhilfe und der Analytischen Taskforces der Feuerwehren Hamburg und Mannheim danke ich für die Bereitstellung des Labors sowie für die interessanten Einblicke in die Anwendungsseite.

Abschließend möchte ich meiner Familie und Freunden für die nötige Motivation danken.





## Kurzfassung

Das Risiko für Freisetzungen von Gefahrstoffen in Form von toxischen Gasen wächst mit der zunehmenden Anzahl verarbeiteter und transportierter gefährlicher Güter. Hilfskräfte benötigen zur Gefahrenabwehr und zum Eigenschutz schnelle und nachweisstarke Vor-Ort-Analysemethoden.

Bisher eingesetzte Verfahren sind entweder zielstoffspezifisch, wie z. B. Gasprüfröhrchen oder aber wenig bzw. nicht selektiv, wie z. B. ein Photoionisationsdetektor. Gasprüfröhrchen erlauben nur qualifizierende Messungen bei bekanntem Gefahrstoff, da sie nur für ihren Zielstoff selektiv sind. Ein einzelner breitbandiger Detektor hingegen gestattet keine Identifikation eines Stoffes. Diese Lücke in der Nachweisbarkeit von unbekanntem gasförmigen Gefahrstoffen kann durch Kombinationsinstrumente und der automatisierten Interpretation von Signalmustern geschlossen werden.

Mit dem im Rahmen eines Forschungsvorhabens entwickelten Gasdetektorenarray mit Chromatographen können geringe Konzentrationen verschiedener toxischer Substanzen detektiert werden. Das integrierte Ionenmobilitätsspektrometer mit einer gaschromatographischen Kopplung erlaubt eine Auftrennung der Moleküle der Gasprobe nach Ionisierbarkeit, Gewicht und effektiver Größe sowie nach Flüchtigkeit und Polarität. Über die sich hieraus ergebenden Kennwerte der Ionenmobilität und Retentionszeit sowie den Signalmustern zusätzlicher Gassensoren, können Substanzen unterschieden werden.

Das Ziel dieser Arbeit ist es, ein optimales Verfahren der Stoffidentifikation für das vorgestellte Messsystem zu finden. Dazu werden zunächst die Signale unterschiedlicher Detektor- und Sensortechnologien betrachtet und individuell gegen Störeinflüsse gefiltert. Aus den entstornten Signalen werden Merkmale für eine Mustererkennung abgeleitet. Mithilfe der extrahierten Merkmale werden über einen Messdatensatz bekannter Substanzen Klassifikationsverfahren trainiert. Das Training des Klassifikators erfolgt anhand von Labormessungen unter unterschiedlichen Probengasbedingungen. Die regelbasierte Klassifikationsmethode wird für den schnellen Vor-Ort-Einsatz als Algorithmus auf dem entwickelten Messinstrument ressourcensparend umgesetzt. Anhand von Erprobungsmessungen wird die hervorragende Klassifikationsrate der optimierten Methode im Vergleich zu anderen Verfahren gezeigt.





# Inhaltsverzeichnis

|  |           |
|--|-----------|
| <b>1. Einleitung</b>   | <b>1</b>  |
| <b>2. Grundlagen</b>   | <b>5</b>  |
| 2.1. Gassensoren . . . . .   | 5         |
| 2.1.1. Halbleitergassensoren . . . . .                                     | 6         |
| 2.1.2. Photoionisationsdetektor . . . . .                                  | 9         |
| 2.1.3. Elektrochemische Zellen . . . . .                                   | 12        |
| 2.2. Ionenmobilitätsspektrometer . . . . .                                 | 16        |
| 2.2.1. Ionenmobilität . . . . .  | 17        |
| 2.2.2. Prinzip eines Flugzeit-Ionenmobilitätsspektrometers . . . . .       | 18        |
| 2.2.3. Prinzip eines Doppelröhren-Flugzeit-Ionenmobilitätsspektrometers    | 20        |
| 2.2.4. Ionisierungsprozess . . . . .                                       | 21        |
| 2.2.5. Eigenschaften der Spektren . . . . .                                | 23        |
| 2.3. Gaschromatographie . . . . .  | 24        |
| 2.3.1. Einflussgrößen der Gaschromatographie . . . . .                     | 25        |
| 2.3.2. Methoden der Injektion von Probengas . . . . .                      | 26        |
| 2.4. Verfahren der Mustererkennung und Klassifikation . . . . .            | 30        |
| 2.4.1. Gestaltung von Klassengrenzen . . . . .                             | 33        |
| 2.4.2. Klassifikationsbäume . . . . .                                      | 37        |
| 2.4.3. Random Forest . . . . .   | 43        |
| <b>3. Aufbau von mobilen GC-Multisensor-Systemen</b>                       | <b>46</b> |
| 3.1. Realisierung eines GC-PID-Systems mit kurzer Multikapillare . . . . . | 46        |
| 3.1.1. Fluidik der analytischen Komponenten . . . . .                      | 46        |
| 3.1.2. Detektor . . . . .  | 49        |
| 3.1.3. Datenaufnahme und -verarbeitung . . . . .                           | 51        |



|  |           |
|--|-----------|
| 3.2. Realisierung eines Multi-Sensoren-Arrays . . . . .  | 51        |
| 3.2.1. Gasführung . . . . .  | 52        |
| 3.2.2. Sensorik . . . . .  | 55        |
| 3.2.3. Elektronik . . . . .  | 60        |
| 3.2.4. Datenaufnahme und -verarbeitung . . . . .   | 64        |
| 3.2.5. Umsetzung als Handgerät . . . . .   | 64        |
| <b>4. Analyse des Signalverhaltens</b>   | <b>66</b> |
| 4.1. Merkmalsextraktion der Einzelsensoren . . . . .   | 68        |
| 4.1.1. Relative Sensorantwort . . . . .  | 70        |
| 4.1.2. Anstiegszeit . . . . .  | 70        |
| 4.1.3. Auswahl der Merkmale . . . . .  | 71        |
| 4.2. Merkmalsextraktion von GC-Verläufen . . . . .   | 72        |
| 4.2.1. Filterung nach Savitzky-Golay . . . . .   | 73        |
| 4.2.2. Peakerkennung bei Chromatogrammen . . . . .   | 76        |
| 4.3. Merkmalsextraktion von IM-Spektren . . . . .  | 76        |
| 4.3.1. Signalaufbereitung . . . . .  | 77        |
| 4.3.2. Peakerkennung bei Ionenmobilitäts-Spektren . . . . .                                    | 79        |
| 4.3.3. K0-Wert-Bestimmung . . . . .  | 80        |
| 4.4. Eingangsmerkmale des maschinellen Lerner . . . . .  | 81        |
| 4.5. Signaldarstellung für Vor-Ort-Interpretation durch Expertennutzer . . . . .               | 82        |
| <b>5. Implementierung einer Datenbasis zur Identifikation von Substanzen</b>                   | <b>85</b> |
| 5.1. Messdatenaufnahme . . . . .   | 85        |
| 5.1.1. Messaufbau . . . . .  | 86        |
| 5.1.2. Kalibriergasgenerator mit Permeationsgefäßen . . . . .                                  | 88        |
| 5.1.3. Datenformat . . . . .   | 89        |
| 5.2. Maschinelles Lernen mit Hilfe realer Messdaten . . . . .                                  | 89        |
| 5.2.1. Überwachtes Lernen mit einem Entscheidungsbaum nach CART . . . . .                      | 90        |
| 5.2.2. Überwachtes Lernen mit Random Forest . . . . .  | 92        |
| 5.2.3. Vergleich der Verfahren . . . . .   | 94        |
| 5.3. Implementierung eines Random-Forest Klassifikators . . . . .                              | 95        |
| 5.3.1. Datenaustauschformat zwischen überwachtem Lerner und Klassifikations-Programm . . . . . | 96        |



|  |            |
|--|------------|
| 5.3.2. Implementierung des Klassifikators auf dem Zielsystem . . . . . | 97         |
| <b>6. Evaluation und Ergebnisse</b>                                    | <b>99</b>  |
| 6.1. Test des Klassifikators mit unabhängigen Testdaten . . . . .      | 102        |
| 6.1.1. Bewertung der Klassifikationsraten . . . . .                    | 105        |
| 6.1.2. Diskussion zur Übertragbarkeit des Klassifikators . . . . .     | 109        |
| 6.2. Klassifikation von GC-Läufen . . . . .                            | 110        |
| 6.3. Untersuchung von GC-IMS-Läufen . . . . .                          | 113        |
| <b>7. Zusammenfassung</b>  | <b>116</b> |
| <b>A. Anhang</b>   | <b>120</b> |
| A.1. Liste relevanter gefährlicher Stoffe . . . . .                    | 120        |
| A.2. Dateiformat der Musterdatenbank . . . . .                         | 122        |
| A.3. Fehlerfortpflanzung der K0-Wert-Berechnung . . . . .              | 123        |
| A.4. Bewertung der Random Forest Methode zur Chromatogramm-Auswertung  | 123        |





# Tabellenverzeichnis

|  |     |
|--|-----|
| 2.1. Klassen von Gassensoren . . . . .   | 6   |
| 3.1. Schaltzustände der Injektions- und Verdünnungseinheit . . . . .   | 53  |
| 4.1. Extrahierbare Parameter (Auszug) [53] . . . . .   | 69  |
| 4.2. Extrahierte Merkmale nach gewählter Messmethode eines Multisensor-<br>Arrays mit GC . . . . .   | 82  |
| 6.1. Evaluation eines Messdatenverlaufs durch den Random Forest Klassifikator  | 100 |
| 6.2. Konfusionsmatrix eines Zwei-Klassen-Problems . . . . .  | 102 |
| 6.3. Konfusionsmatrix eines Mehr-Klassen-Problems . . . . .  | 102 |
| A.1. Stoffliste nach vfdb 10/05 mit Einsatztoleranzwerten und Bewertung nach<br>AEGL ( <i>Acute Exposure Guideline Levels</i> ) (Stoffe A-D) [5] [6] . . . . . | 120 |
| A.2. Stoffliste nach vfdb 10/05 mit Einsatztoleranzwerten und Bewertung nach<br>AEGL ( <i>Acute Exposure Guideline Levels</i> ) (Stoffe E-Z) . . . . .         | 121 |



# Abbildungsverzeichnis

|   |    |
|---|----|
| 2.1. Zinnoxid-Halbleiter in Umgebungsluft . . . . .   | 6  |
| 2.2. Zinnoxid-Halbleiter bei Vorhandensein von reduzierendem Gas . . . . .  | 7  |
| 2.3. Zusammensetzung des Sensorleitwerts . . . . .  | 7  |
| 2.4. Physikalischer Aufbau eines Halbleitersensors . . . . .  | 9  |
| 2.5. Physikalisches Prinzip eines Gassensors zur Messung ionisierbarer Gase . .   | 10 |
| 2.6. Aufbau eines Photoionisationsdetektors . . . . .   | 11 |
| 2.7. Allgemeiner Aufbau einer Elektrochemischen Zelle zur Gasetektion . . .   | 13 |
| 2.8. Schematische Darstellung der Strom-Potentialkurve . . . . .  | 14 |
| 2.9. Diffusionsgrenzströme bei unterschiedlichen Stoffkonzentrationen . . . . .   | 15 |
| 2.10. Aufbau einer Drei-Elektroden-Zelle . . . . .  | 16 |
| 2.11. Aufbau der Driftstrecke eines Ionenmobilitätsspektrometers . . . . .  | 19 |
| 2.12. Aufbau eines Doppelröhren-IMS . . . . .   | 21 |
| 2.13. Möglichkeit der Darstellung von Ionen-Mobilitäts-Spektren . . . . .   | 24 |
| 2.14. Prinzip der gaschromatographischen Trennung ( $I =$ Injektor, $D =$ Detektor)   | 24 |
| 2.15. Loop Injektor auf Basis eines Multiport-Ventils . . . . .   | 27 |
| 2.16. Anreicherung über thermische Ad- und Desorption . . . . .   | 28 |
| 2.17. Gegenstrom Injektor . . . . .   | 30 |
| 2.18. Die Kombination mehrerer breitbandiger Sensoren erlaubt eine Verbesserung der Selektivität . . . . .                          | 31 |
| 2.19. Klassifikation von Sensorsignalen . . . . .   | 32 |
| 2.20. Streudiagramm eines zweidimensionalen Merkmalsvektors . . . . .   | 33 |
| 2.21. Optimierung einer Klassengrenze. Ein Klassifikationsalgorithmus sollte eine gute Generalisierungsfähigkeit besitzen . . . . . | 34 |
| 2.22. Festlegen einer Entscheidungsgrenze nach der Verteilungsdichte von Messwerten . . . . .                                       | 35 |



|  |    |
|--|----|
| 2.23. Klasseneinteilung eines zweidimensionalen Merkmalsraums über einen Entscheidungsbaum . . . . .           | 38 |
| 2.24. Klassifikationsbaum . . . . .  | 39 |
| 2.25. Aufteilung des Merkmalsraums mithilfe der Random Forest Methode . .                                      | 44 |
| 2.26. Vergleich von Klasseneinteilung nach Random Forest und CART . . . . .                                    | 45 |
| 3.1. Prinzipschaltbild des Kurzsäulen GC-PID . . . . .   | 47 |
| 3.2. Multikapillar-Gaschromatograph . . . . .  | 48 |
| 3.3. Verwendeter Photoionisationsdetektor Alphasense PID-AH . . . . .  | 49 |
| 3.4. Bestimmung der Nachweisgrenze des eingesetzten Photoionisationsdetektors                                  | 51 |
| 3.5. Flussplan des realisierten Detektoren Arrays mit Gaschromatograph . . .                                   | 52 |
| 3.6. Trägermodul des Kurzsäulen-Gaschromatographen . . . . .   | 55 |
| 3.7. Im DACHS-System eingesetzte Sensoren . . . . .  | 56 |
| 3.8. Foto der Eingangsseite des verwendeten Doppelröhren-Ionenmobilitätsspektrometers . . . . .                | 57 |
| 3.9. Realisierte Gasführung und Sensoren-Aufnahme . . . . .  | 59 |
| 3.10. Blockschaltbild der Elektronik . . . . .   | 60 |
| 3.11. Benutzerschnittstellen . . . . .   | 64 |
| 3.12. Aussenansicht des Messsystems . . . . .  | 65 |
| 4.1. Signalpfad . . . . .  | 66 |
| 4.2. Zeitlicher Verlauf einer direkten Messung ohne Datenvorverarbeitung . .                                   | 67 |
| 4.3. Merkmale eines Signalsprungs . . . . .  | 69 |
| 4.4. Chromatogramm von Aceton, Toluol, Xylol . . . . .   | 73 |
| 4.5. Merkmale eines GC-Peaks . . . . .   | 74 |
| 4.6. Vergleich von Savitzky-Golay-Filter und gleitendem Mittelwert . . . . .                                   | 75 |
| 4.7. Klassische Methode der Peakerkennung durch Bildung der ersten Ableitung des Signalverlaufs [54] . . . . . | 76 |
| 4.8. Datenverarbeitung der Ionenmobilitätsspektren . . . . .   | 77 |
| 4.9. Erkennung von Peaks im Ionenmobilitätsspektrum . . . . .  | 79 |
| 4.10. Zusammensetzung des Wertes der reduzierten Ionenmobilität $K_0$ . . . . .                                | 80 |
| 4.11. Balkendarstellung der Messwerte zur visuellen Interpretation . . . . .                                   | 83 |
| 5.1. Messvarianten zum Aufbau einer Stoffdatenbank für zwei Funktionsmuster des DACHS . . . . .                | 86 |



## Abbildungsverzeichnis

|      |  |     |
|------|--|-----|
| 5.2. | Schematische Darstellung des Messaufbaus . . . . .   | 87  |
| 5.3. | Prinzipschaltbild des Kalibriergasgenerators . . . . .   | 88  |
| 5.4. | Vergleich von zwei Funktionen der Trenngrenzenbestimmung bei der Bau-<br>minduktion . . . . .  | 91  |
| 5.5. | Ablauf des Random-Forest Klassifikationsprozesses . . . . .  | 92  |
| 5.6. | Klassifikationsrate abhängig von Baumanzahl und minimaler Blattgröße .   | 94  |
| 5.7. | Innere Knoten einzelner Entscheidungsbäume in der Klassifikatordatei . .   | 96  |
| 5.8. | Speicherstruktur des Random-Forest-Klassifikators . . . . .  | 98  |
| 6.1. | Evaluation eines Messdatenverlaufs durch den Random Forest Klassifikator   | 99  |
| 6.2. | Zwei IM-Spektren im Messverlauf mit Bezugszeitpunkt der Reaktantionen<br>(senkrechte Linie) . . . . .  | 101 |
| 6.3. | Konfusionsmatrix . . . . .   | 104 |
| 6.4. | Beschreibung des Klassifikationsverhaltens eines Random Forest Klassifi-<br>kators anhand von Konfusionsmatrizen der Sensitivität und Präzision . .                      | 105 |
| 6.5. | Vergleich zweier Funktionsmuster des Messsystems DACHS hinsichtlich<br>der $K_0$ -Werte bei Messung von Essigsäure . . . . .   | 109 |
| 6.6. | Retentionszeit bei unterschiedlichen Druckdifferenzen . . . . .  | 111 |
| 6.7. | Streudiagramme von GC-Läufen bei unterschiedlichen Säulenparametern.<br>Deutlich zu erkennen in der Ebene sind ausgeprägte Bereiche der Klas-<br>sengültigkeit . . . . . | 112 |
| 6.8. | GC-Lauf mit IMS-Auswertung . . . . .   | 114 |
| A.1. | Kopf der Klassifikatordatei . . . . .  | 122 |
| A.2. | Ergbnisknoten einzelner Entscheidungsbäume in der Klassifikatordatei . .   | 122 |
| A.3. | Reklassifizierungsfähigkeit abhängig von Größe des Random Forest . . . .   | 123 |



# Formelzeichenverzeichnis

|                        |       |  |
|------------------------|-------|--|
| $\Delta H$             | ..... | Molare Enthalpie eines Eluenten                                  |
| $\frac{\alpha}{\beta}$ | ..... | Molare Entropie eines Eluenten                                   |
| $\mathcal{L}$          | ..... | Datensatz mit Messwerten zum Anlernen eines Klassifikators       |
| $\mathcal{L}^{(B)}$    | ..... | Zufallsauswahl von Messwerten aus einem Trainingsdatensatz       |
| $\mu$                  | ..... | Erwartungswert   |
| $\Omega_D$             | ..... | Stoßquerschnitt  |
| $\Omega^{(1,1)*}$      | ..... | Kollisionsintegral   |
| $\omega_{1..k}$        | ..... | Klassenbezeichner  |
| $\bar{k}$              | ..... | Retentionszeitindex  |
| $\bar{v}_C$            | ..... | mittlere Trägergasgeschwindigkeit                                |
| $\Sigma$               | ..... | Kovarianzmatrix  |
| $\sigma^2$             | ..... | Varianz  |
| $\sigma_{hl}$          | ..... | spezifische Leitfähigkeit eine Halbleitersensors                 |
| $\theta_p$             | ..... | Komplexitätsparameter als Abbruchkriterium bei der Bauminduktion |
| $\varphi_0$            | ..... | Standardpotential  |
| $\vec{v}$              | ..... | Merkmalsvektor   |
| $A$                    | ..... | Fläche   |
| $a$                    | ..... | Aktivität  |
| $c$                    | ..... | Stoffkonzentration   |
| $c_M$                  | ..... | Stoffmengenkonzentration   |
| $D$                    | ..... | Diffusionskoeffizient  |



## Abbildungsverzeichnis

|                  |   |
|------------------|---|
| $d$              | Abstand   |
| $E$              | elektrische Feldstärke  |
| $e^-$            | Elektron  |
| $F$              | Faradaykonstante  |
| $h$              | Peak-Höhe   |
| $I_m$            | Entropiegehalt des Knotens $m$  |
| $i_{\text{neg}}$ | Strom negativer Ladungsträger am Detektor eines IMS   |
| $i_{\text{pos}}$ | Strom positiver Ladungsträger am Detektor eines IMS   |
| $i_{\text{gr}}$  | Diffusionsgrenzstrom  |
| $K$              | Ionenmobilität  |
| $K_0$            | reduzierte Ionenmobilität   |
| $k_B$            | Boltzmann-Konstante   |
| $l_C$            | Länge der chromatographischen Trennsäule  |
| $l_D$            | Länge der Driftstrecke  |
| $M$              | Masse der Driftgasmoleküle  |
| $m$              | Bezeichner des Entscheidungsknotens   |
| $m_{\text{Ion}}$ | Masse eines Ions  |
| $N$              | Teilchendichte  |
| $n_e$            | elektrochemische Wertigkeit   |
| $N_m$            | Gesamtzahl von Merkmalsvektoren im Eingaberaum zum Knoten $m$                                 |
| $N_m^i$          | Anzahl von Merkmalsvektoren im Eingaberaum zum Knoten $m$ , die zur Klasse $\omega_i$ gehören |
| $p$              | Anzahl von Merkmalen  |
| $P_0$            | Normaldruck, $P_0 = 101,325$ kPa  |
| $p_a$            | Partialdruck  |
| $P_D$            | Absolutdruck im Driftraum [kPa]   |
| $q$              | Ladung eines Ions   |



## Abbildungsverzeichnis

|                     |                                    |
|---------------------|------------------------------------|
| $R$ .....           | allgemeine Gaskonstante            |
| $r_m$ .....         | Querschnitt eines Ions             |
| $s_{1..q}(t)$ ..... | Sensorsignalverläufe               |
| $T$ .....           | Temperatur [K]                     |
| $T^*$ .....         | dimensionslose Temperatur          |
| $T_0$ .....         | Normaltemperatur, $T_0 = 273,15$ K |
| $T_D$ .....         | Temperatur im Driftraum            |
| $t_d$ .....         | Driftzeit                          |
| $t_i$ .....         | Injektionszeitpunkt                |
| $t_R$ .....         | Retentionszeit                     |
| $T_{Ads}$ .....     | Adsorptionstemperatur              |
| $T_{Des}$ .....     | Desorptionstemperatur              |
| $v_d$ .....         | Driftgeschwindigkeit               |
| $v_{1..p}$ .....    | Merkmal                            |





# 1. Einleitung

Die Produktion, Verarbeitung und der Transport toxischer Substanzen stellen große potentielle Gefahren für Mensch und Umwelt dar. Es kommt immer wieder zu Unfällen bei denen es zu einer Freisetzung von gefährlichen Stoffen kommt. Zur Abwehr dieser Gefahren gibt es in Deutschland ein starkes Netzwerk von Fachkräften, welches durch Bund und Länder durch lokale Fachgruppen der Feuerwehren und überregional durch die Analytischen Task Forces bereitgestellt wird [1]. Über zweidrittel der meldepflichtigen Störfälle in verfahrenstechnischen Anlagen auf dem Bundesgebiet waren 2010 mit einer Stofffreisetzung verbunden [2]. Da es sich häufig um flüchtige Stoffe handelt, wird aus dem lokalen Störfall schnell eine sich ausbreitende Gefahrenquelle für die Bevölkerung. Neben Störfällen in stationären Anlagen gibt es eine große Anzahl mobiler gefährlicher Güter auf verschiedenen Transportwegen.

Gerade Mischtransporte stellen aufgrund ihrer Vielzahl verschiedener Gefahrstoffe ein großes Gefährdungspotential dar. So wurden zum Beispiel auf der 2012 verunglückten MSC Flaminia 149 Seefracht-Container mit unterschiedlichsten Gefahrgütern transportiert. Ein Problem für die Einsatzkräfte ist die häufige Fehldeklaration von gefährlichen Gütern: Im betreffenden Laderaum des verunglückten Schiffes waren 38 der 397 Güter falsch oder unzureichend als Gefahrgut gekennzeichnet. [3]

Insgesamt wurden allein 2011 in Deutschland 292 Millionen Tonnen Gefahrgut transportiert [4]. Den bei einem Störfall herbeigerufenen Einsatzkräften stellen sich die wichtigen Fragen, ob ein gefährlicher Stoff freigesetzt wurde (*Detektion*), um welchen Stoff es sich handelt (*Identifikation*) und ob die Konzentration eine Gefährdung darstellt (*Quantifikation*). Für diese Fragestellung benötigen die eingesetzten Fachkräfte schnelle, nachweisstarke und mobile Vor-Ort-Messtechnik.



## 1. Einleitung

Es existiert eine große Anzahl von Stoffen, die eine Gefährdung für Leben und Umwelt darstellen. Diese Stoffe nach ihrer Toxizität zu beurteilen ist Aufgabe öffentlicher Institutionen, wie der US-amerikanischen *EPA*<sup>1</sup> oder dem deutschen *vfdb*<sup>2</sup>. Diese Einrichtungen haben Listen von Substanzen mit dem größten Gefährdungspotential herausgegeben. Hierfür wurden die Stoffe neben der Toxizität auch nach ihrer Häufigkeit des Vorkommens bewertet. Die vom *vfdb* veröffentlichte Liste enthält 44 Stoffe und eine Konzentrationsangabe die als *Einsatztoleranzwert (ETW)*[5] bezeichnet wird. Diese Werte gelten für den Einsatz von Hilfskräften. Bei Überschreitung des substanzspezifischen Grenzwertes müssen Einsatzkräfte Maßnahmen des Atemschutzes ergreifen. Die von der EPA publizierte Grenzwerte (*Acute Exposure Guideline Levels - AEGL*)[6] gelten für eine Störfallbewertung. So werden fünf Expositionsdauern (10 Minuten bis 8 Stunden) und drei Schweregrade (AEGL1 - AEGL3) der Auswirkung auf den Menschen mit einem Konzentrationswert belegt.

- AEGL1 - Unwohlsein
- AEGL2 - dauerhaft gesundheitsgefährdend
- AEGL3 - tödlich

Der Umfang der Stoffliste wird ständig erweitert. Bis zum Jahr 2015 sind 176 Chemikalien erfasst worden. In den Tabellen A.1 und A.2 im Anhang dieser Arbeit sind die Stoffe der ETW-Liste mit relevanten Werten der AEGL angegeben.

Den Kräften des Katastrophenschutzes stehen bislang für die drei Aufgaben Detektion, Identifikation und Quantifikation der Stoffe dieser Liste unterschiedliche Systeme zur Verfügung. Für den flächendeckenden Vor-Ort-Einsatz werden in den CBRN-Erkundungswagen des Bundes zur chemischen Detektion ein Photoionisationsdetektor (PID) und ein Ionenmobilitätsspektrometer (IMS) bereitgehalten [7]. Das IMS besitzt gute Eigenschaften zur Identifikation aber nur unzureichende Quantifikationseigenschaften. Der PID hingegen ist nur in der Lage zu detektieren und zu quantifizieren. Es existieren zwei Messtrategien für Störfallbeurteilung: Das *Spüren* und das *Messen*. Der verbreitete Erkundungskraftwagen des Bundes ist mit seiner analytischen Ausrüstung für die Aufgabe des Spürens geeignet. In diesem Anwendungsfall soll ein kontaminiertes Gebiet durch das Aufspüren geringer Konzentrationen am Gebietsrand eingegrenzt

---

<sup>1</sup>US Environmental Protection Agency

<sup>2</sup>Vereinigung zur Förderung des Deutschen Brandschutzes (vfdb) e.V.



## 1. Einleitung

werden. Beim Messen hingegen geht es um den Stoffnachweis an einer Emissionsquelle. Hier treten also mitunter hohe Konzentrationen auf. Die Identifikations- und Quantifikationsfähigkeit von Detektorsystemen wird dabei aufgrund von Sensorsättigungen beeinträchtigt. Beim *Messen* ist daher eine definierte Verdünnung der Probe nötig, will man eine gute Nachweisfähigkeit erreichen. Das von Albrecht und Hunte [8] [9] vorgeschlagene Kombinationsinstrument verfügt im Gegensatz zur flächendeckend eingesetzten Spürtechnik über eine Verdünnungseinheit und besitzt damit gute Eigenschaften in der Identifikation und Quantifikation von Stoffen, wenn diese nicht in Gemischen vorliegen.

Im Forschungsvorhaben DACHS sollte ein solches Kombinationssystem realisiert und für eine verbesserte Nachweisfähigkeit um zusätzliche Sensoren erweitert werden [10][11]. Eine gaschromatographische Vortrennung soll zu einer verbesserten Identifikationsfähigkeit führen [12].

Der Mehrwert einer Kombination mehrerer Sensortechnologien entsteht durch die Fusionierung der Daten. Bereits die Disjunktion der Sensordaten führt zu einer Erweiterung der Detektierbarkeit von Gasen, da Lücken der Nachweisfähigkeit einzelner Sensoren für eine bestimmte Substanz durch andere Sensoren aufgefüllt werden. Berücksichtigt eine Datenauswertung zusätzlich die Relationen der Einzelsignale, so ist über Mustererkennungsverfahren eine optimierte Stoffidentifikation möglich. Diesem Ansatz folgen bereits die olfaktorischen Rezeptoren aller Säugetiere [13]. Ein Gas reizt einen Teil der mehreren tausend Sinneszellen der Nase. Die Geruchswahrnehmung erfolgt über die Interpretation des Musters durch die Synapsen des olfaktorischen Cortexes [14].

Für die Klassifikation der Muster eines künstlichen olfaktorischen Systems sind verschiedene Verfahren möglich. So bietet die Disziplin der Mustererkennung eine ganze Reihe an Methoden, wie die lineare Diskriminanzanalyse, künstliche Neuronale Netze, Support Vector Machines etc. [15] [16]. Mit einer solchen elektronischen Nase können Gerüche bzw. Gasgemische durch eine Zusammenstellung chemischer Sensoren unterschiedlicher Selektivitäten und einem geeigneten Verfahren der Mustererkennung [17] unterschieden werden.

Genauso wie das Geruchsempfinden einer menschlichen Nase, ist ein Sensorenarray allerdings nur unzureichend in der Lage Einzelkomponenten eines Gasgemisches selektiv zu identifizieren. Die Kopplung mit einer gaschromatographischen Trennsäule bietet



## 1. Einleitung

zusätzlich zum Muster der Sensoren, durch eine zeitliche Auftrennung der Gas-Probe, die Möglichkeit auch komplexe Gemische zu analysieren [18].

### Zielsetzung und Gliederung der Arbeit

Ziel der Arbeit ist es für das Messinstrument DACHS ein optimales Verfahren der Stoffidentifikation zu finden und dieses für den Vor-Ort-Einsatz auf einem eingebetteten Mikroprozessorsystem ressourcensparend zu implementieren.

In Kapitel 2 werden für den Aufbau eines Sensorenarrays zunächst verschiedene in Frage kommende Sensortechnologien beschrieben, damit die optimalen Sensoren selektiert werden können. Der zweite Teil dieses Theoriekapitels beschreibt Grundlagen der Mustererkennung, das Verfahren zum Aufbau von Klassifikationsbäumen und die Ensemblemethode des Random Forest. Aus diesen Grundlagen wird später eine Methode für diesen Anwendungsfall entwickelt.

Die Umsetzung von Systemen zur Vermessung von Gasen wird in Kapitel 3 beschrieben. Es werden jeweils zunächst schematisch die fluidischen Verschaltungen für die prinzipielle Funktionsweise gezeigt. Darauf aufbauend wird die Realisierung der analytischen Komponenten untersucht, sowie die Hard- und Software entworfen und implementiert.

Kapitel 4 behandelt die Signaleigenschaften aller eingesetzten Sensoren. Es werden Methoden zur Rauschunterdrückung und zur Extraktion von Merkmalen aus Signalverläufen, Chromatogrammen und Ionenmobilitäts-Spektren untersucht.

Anhand von realen Messdaten wird in Kapitel 5 gezeigt, wie aus gewonnenen Merkmalen Klassifikatoren abgeleitet werden. Es wird die optimale Klassifikationsmethode mit ihren besten Parametern bestimmt. Abschließend wird in diesem Kapitel die praktische Implementierung des Klassifikators auf dem Zielsystem beschrieben.

In Kapitel 6 wird der gefundene Klassifikator mit unabhängigen Testdaten für verschiedene Anwendungsfälle evaluiert.

Abschließend werden die Inhalte und Ergebnisse der Arbeit in Kapitel 7 zusammengefasst.

## 2. Grundlagen

In diesem Kapitel werden die Grundlagen für eine maschinengestützte Identifikation von Gasen mithilfe von Sensorenarrays mit einem gekoppelten Trennverfahren beschrieben. In den ersten beiden Abschnitten werden die verschiedenen verwendeten Sensortechnologien mit ihren Funktionsprinzipien aufgezeigt. Diese bilden die Grundlage für das Verständnis welche Merkmale aus den jeweiligen Sensorsignalverläufen für eine Mustererkennung extrahiert werden können. Der dritte Abschnitt erläutert die Prinzipien der nachfolgend angewendeten Gaschromatographie. Am Ende dieses Kapitels soll ein Grundverständnis für die Mustererkennung vermittelt werden. Dabei wird gezeigt, wie für zwei ausgewählte Klassifikationsmethoden die Klassengrenzen in einem Merkmalsraum anhand von Trainingsmessdaten ermittelt werden.

### 2.1. Gassensoren

Sensoren bilden physikalische oder chemische Messgrößen auf elektrische Signale ab [19]. Diese können dann digitalisiert und für einen Informationsgewinn weiterverarbeitet werden. Gassensoren, welche der Erfassung chemischer Eigenschaften eines Gases dienen, basieren auf unterschiedlichen Messprinzipien. Sie können in physikalische und chemische Messmethoden kategorisiert werden (siehe Tabelle 2.1) [20].

In dieser Arbeit wird eine Auswahl von Sensoren betrachtet, welche für den vorliegenden Anwendungsfall eines tragbaren Vor-Ort-Analysesystems relevant sind. Das Ziel ist es, möglichst sämtliche für Feuerwehren einsatztaktisch relevanten Gefahrstoffe detektieren zu können [5].

Tabelle 2.1.: Klassen von Gassensoren [20]

| PHYSIKALISCHE MESSMETHODEN |             |            |              | CHEMISCHE MESSMETHODEN |             |            |
|----------------------------|-------------|------------|--------------|------------------------|-------------|------------|
| Paramagnetismus            | Photometrie | Ionisation | Wärmeleitung | Elektrochemisch        | Wärmetönung | Halbleiter |
|                            | UV          | PID        |              | amperometrisch         |             |            |
|                            | NDIR        | FID        |              | potentiometrisch       |             |            |
|                            |             |            |              | polarografisch         |             |            |

### 2.1.1. Halbleitergassensoren

Bei der Interaktion von Gasmolekülen mit einem Metalloxid kommt es an der Grenzschicht zur Oxidation. Dies hat zur Folge, dass sich im Halbleitermaterial die Anzahl freier Ladungsträger verändert und sich somit auch die Leitfähigkeit des Materials ändert. Abhängig von der Dotierungsart kann das Material leitfähiger werden oder sich der Materialwiderstand erhöhen. Bei n-dotierten Halbleitern führt die Oxidation zu einer Leitfähigkeitserhöhung, während p-dotierte Halbleiter ihre Leitfähigkeit verringern.

In Umgebungsluft werden auf der Halbleiteroberfläche Sauerstoffmoleküle chemisorbiert. Die hier gebundenen Sauerstoffionen  $O^-$  entziehen dem Material jeweils ein Elektron  $e^-$ . Durch das Unterangebot freier Elektronen im Halbleitermaterial, ist die Sensorschicht in normaler Atmosphäre in ihrer Leitfähigkeit herabgesetzt. Die Abbildung 2.1 veranschaulicht

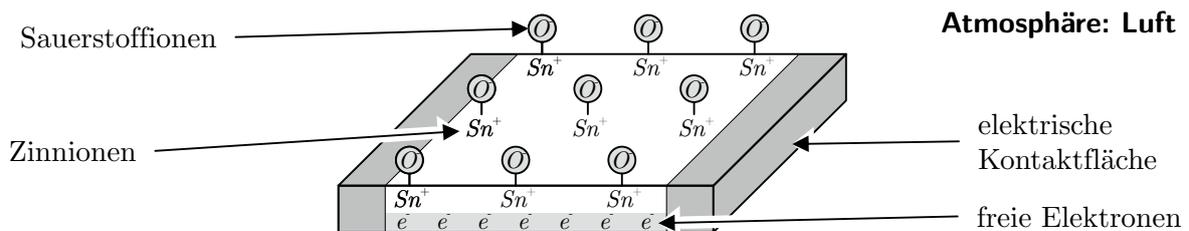


Abbildung 2.1.: Zinnoxid-Halbleiter in Umgebungsluft [9]

licht das Verhalten eines Zinnoxid-Halbleiters bei Umgebungsluft. Durch diese Konfiguration wird der Sensor empfindlich für reduzierende Gase.

## 2. Grundlagen

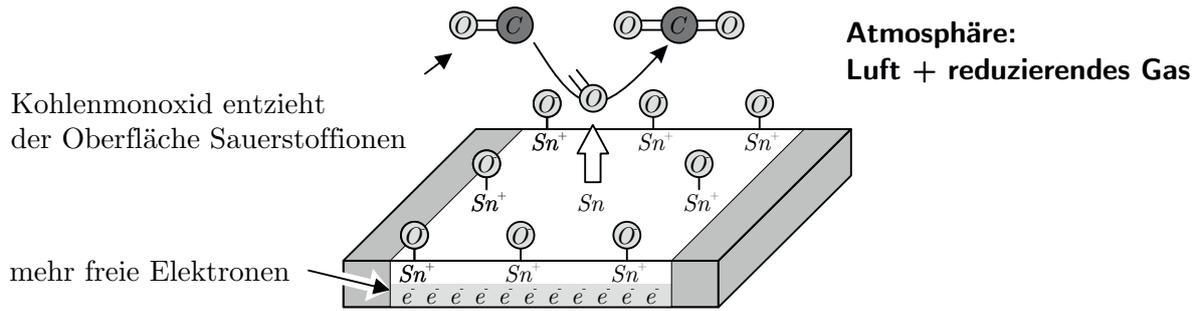
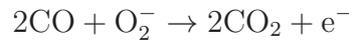


Abbildung 2.2.: Zinnoxid-Halbleiter bei Vorhandensein von reduzierendem Gas

In Abbildung 2.2 ist die Reaktion des Sensors auf Kohlenmonoxid gezeigt. Die CO-Moleküle lösen Sauerstoffionen von der sensitiven Schicht und reagieren zu CO<sub>2</sub>:



Allgemein kann dieser Prozess beschrieben werden als:



wobei R das reduzierende Gas beschreibt. Die entstehenden freien Elektronen  $e^-$  wandern ins Leitungsband und verringern so den Widerstand des Sensors. Diese hervorgerufene Leitfähigkeitsänderung  $\Delta G$  kann elektrisch bestimmt werden. Die Gesamtleitfähigkeit des Sensors kann als Zusammensetzung von Einzelleitwerten betrachtet werden [21]. Wie in Abbildung 2.3 gezeigt, besteht die Gesamtimpedanz aus Teilwiderständen des Halbleitermaterials und der Kontaktstellen. Verschiedene Ladungstransportmechanismen finden in unterschiedlichen Bereichen des Sensors statt. Der Halbleiter erfährt seine Widerstandsänderung durch Reaktionen an der Oberfläche und durch Diffusionsprozesse im Materialvolumen. Da es sich typischerweise um einen polykristallinen

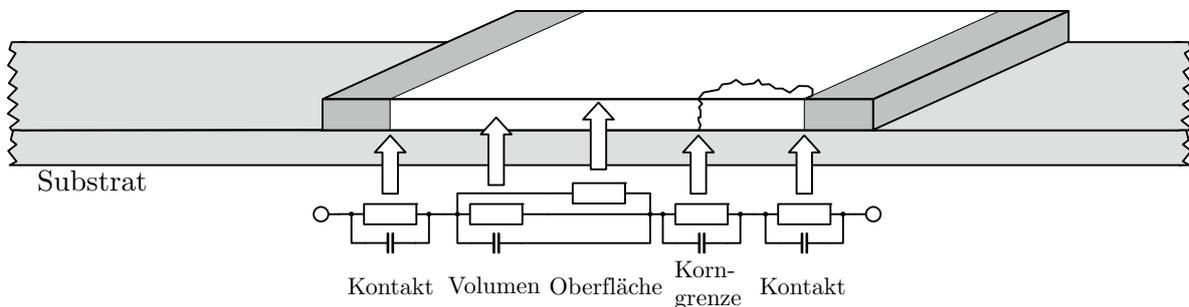


Abbildung 2.3.: Zusammensetzung des Sensorleitwerts [21]

## 2. Grundlagen

Materialaufbau handelt, sind auch die Korngrenzen separat zu betrachten. Neben der ohmschen Leitfähigkeitsänderung, findet eine Veränderung der Permittivität  $\varepsilon = \varepsilon_0 \cdot \varepsilon_r$ , und damit der materialspezifischen Kapazitäten, statt.

Die freien Ladungsträger aus dem durch Formel 2.1 beschriebenen Prozess, führen zu einer momentanen Leitfähigkeitsänderung des Materials (spezifische Leitfähigkeit  $\sigma_{hl} = \frac{1}{\rho_{hl}}$  abhängig von der Substanzkonzentration  $c$  nach folgendem Verhalten [22][23]:

$$\sigma_{hl} = k \cdot c^\kappa \quad (2.2)$$

Die beiden Konstanten  $k$  und  $\kappa$  werden empirisch bestimmt. Typische Werte für  $\kappa$  liegen bei  $-1/4$ ,  $-1/5$ , etc. [21]. Neben der beschriebenen Ladungsträgerfreisetzung können durch direkte Adsorption von Substanzmolekülen Ladungsträger lokal verarmt oder angereichert werden. Dies geschieht an der Halbleiteroberfläche oder den Korngrenzen. Hier findet somit eine Verbiegung des Leitungs- und des Valenzbandes statt. Dieser Zielstoff-abhängige Effekt kann durch die Auswahl des Sensormaterials begünstigt oder reduziert werden. In Dünnschicht-Technologie hergestellte Sensoren erfahren eine höhere Beeinflussung durch die Chemisorption als Dickschichtsensoren.

Ein weiterer, die Leitfähigkeit beeinflussender Effekt, ist die Reaktion von Wassermolekülen auf der Materialoberfläche. Bei Messungen von Gaskonzentrationen in Umgebungsluft, ist die Luftfeuchtigkeit ein bestimmender Einflussfaktor. Eine hohe Betriebstemperatur des Sensors verringert den Einfluss der Luftfeuchtigkeit. Gleichzeitig können durch Oberflächenkatalyse von freigesetztem Wasserstoff und Sauerstoff, Wassermoleküle gebildet werden, welche wiederum einen Einfluss auf die Leitfähigkeit haben [24]:



Je nach Halbleitermaterial finden die Reaktionen an der sensitiven Schicht bei höheren Temperaturen statt. Unbehandeltes Zinn-Oxid ( $\text{SnO}_2$ ) zeigt seine Nachweisfähigkeit bei Temperaturen zwischen  $400\text{ °C}$  und  $500\text{ °C}$ . Durch Dotierung des Halbleiters mit einem Edelmetall, wie z.B. Platin, Palladium oder Silber kann die Ansprechtemperatur des Sensors herabgesetzt werden. Hierdurch werden Betriebstemperaturen zwischen  $200\text{ °C}$  und  $400\text{ °C}$  erreicht. [25]

Die Betriebstemperatur eines Halbleitersensors wird in der Praxis durch einen Heizmäander unterhalb des Sensorsubstrates erzeugt. In Abbildung 2.4 ist der typische phy-

## 2. Grundlagen

sikalische Aufbau gezeigt. In der untersten Ebene ist ein Träger aus Keramik zu se-

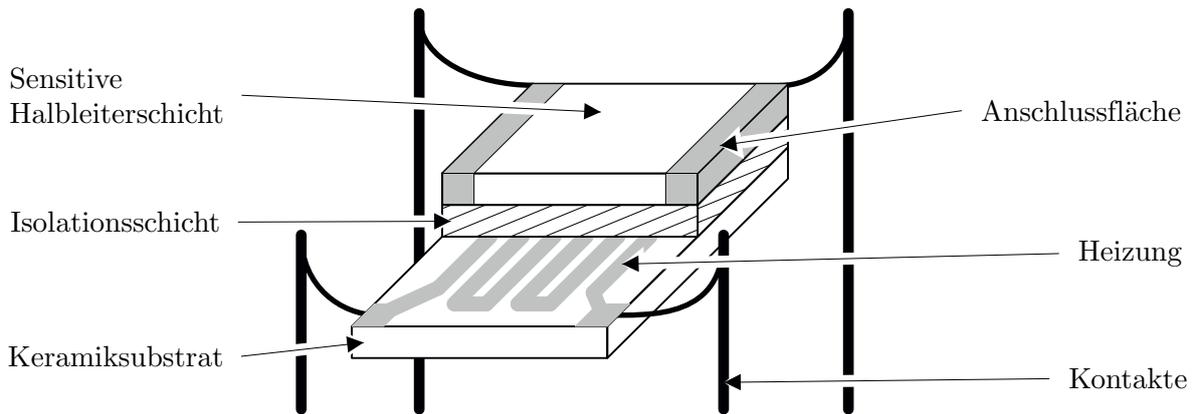


Abbildung 2.4.: Physikalischer Aufbau eines Halbleitergassensors [10]

hen. Hierauf aufgebracht ist eine mäanderförmige leitende Struktur, welche als resistive Heizung des Sensors fungiert. Häufig besteht diese Struktur aus Platin, da dieses Widerstandsmaterial eine sehr lineare Temperaturabhängigkeit aufweist. Über die Widerstandsänderung wird die Betriebstemperatur des Sensors mittels einer Ansteuerelektronik geregelt.

In Dickschicht-Technologie hergestellte Sensoren haben eine typische Oberfläche von weniger als 5 mm mal 5 mm. Sie werden häufig in genormten Transistor-Gehäusen eingesetzt. Dünnschicht-Sensoren sind als Strukturen der Mikrosystemtechnik etabliert.

### 2.1.2. Photoionisationsdetektor

Es existieren verschiedene Methoden um Gasmolekülen eine Ladung aufzuprägen. Entstandene Molekül-Ionen werden allgemein bei einem Ionisationsdetektor durch Elektroden aufgefangen. In einem angeschlossenen Stromkreis stellt sich durch das Einfangen der ionisierten Moleküle ein Stromfluss ein, der abhängig zur Anzahl der aufgenommenen Ladungsträger ist. Als Energiequellen der Ionisation kommen dabei Verfahren in Frage, wie Flammen-Ionisation, Ionisation mit radioaktiven Isotopen oder die hier beschriebene Photoionisation. Die Art und der Energiegehalt der Ionisationsquelle beeinflusst die Ionenausbeute des Analyten, eine energiereiche Quelle kann jedoch zu Fragmentierung

## 2. Grundlagen

der Moleküle führen. Die vorgestellte Photoionisation ist vergleichsweise energiearm. Sie zählt zu den weichen Ionisierungsarten [26].

Ein einfacher Aufbau zur Messung ionisierbarer Gase ist in Abbildung 2.5 gezeigt. Innerhalb einer Messkammer befinden sich zwei gegenüberliegende Elektroden, zwischen welchen eine Spannung angelegt ist.

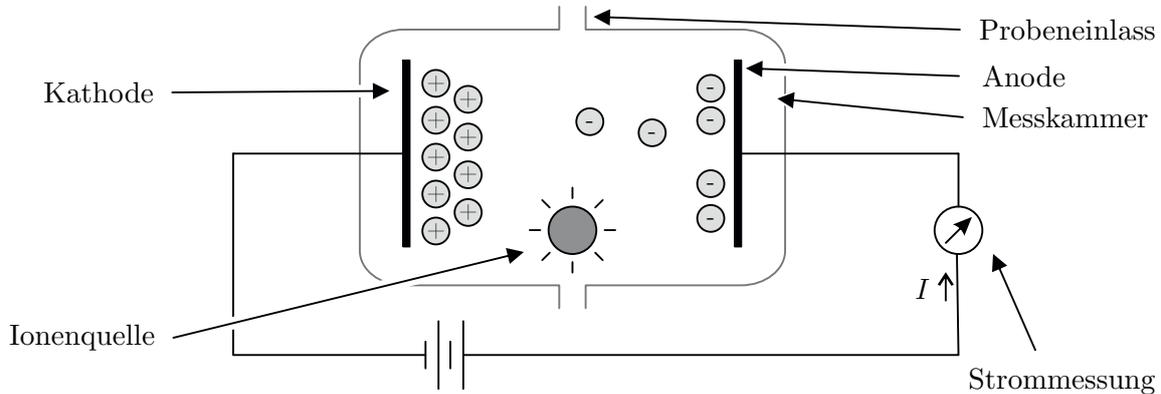
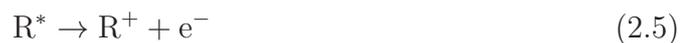


Abbildung 2.5.: Physikalisches Prinzip eines Gassensor zur Messung ionisierbarer Gase (in Anlehnung an [27])

Die Gasprobe wird an einer Ionenquelle vorbeigeführt. Sind die Moleküle mit der Energie der Quelle ionisierbar, so bilden sich freie Elektronen und Substanz-Ionen. Diese werden aufgrund des vorhandenen elektrostatischen Feldes getrennt. Die negativen Ladungsträger wandern zur Anode und erzeugen damit einen Stromfluss durch den verbundenen Stromkreis. An der Gegenelektrode rekombinieren die Stoffmoleküle mit den aus der Kathode austretenden Elektronen.

Bei der Photoionisation wird als Energiequelle ein durch eine UV-Lampe emittiertes Lichtquant  $h\nu$  genutzt. Das Analytmolekül  $R$  wird hierdurch, wie in Formel 2.5 gezeigt, in einen angeregten Zustand  $R^*$  versetzt.



Aus der Anregung gehen ein Elektron ( $e^-$ ) und ein positiv geladenes Ion ( $R^+$ ) hervor. Zusammen mit den Trägergasmolekülen ( $C$ ) kommt es an der Kathode wieder zur

## 2. Grundlagen

Rekombination nach Formel 2.6. [28]



Die Energie des Lichtquants  $h\nu$  ist abhängig von der eingesetzten UV-Quelle. Als Lichtquelle werden häufig kapazitiv angeregte Gasentladungslampen eingesetzt. Die Ionisierungsenergie ist hierbei abhängig von dem Leuchtgas und dem Fenstermaterial der Entladungslampe. Gebräuchliche Gase sind Argon, Deuterium, Krypton und Xenon mit Strahlungsenergien von 11,8 eV, 10,2 eV, 10,6 eV und 9,6 eV [29].

Das Fenstermaterial der Lampe muss für den gewünschten Spektralbereich durchlässig sein, daher wird Magnesiumfluorid oder Lithiumfluorid verwendet [30]. Die abgestrahlten Photonen der Lampe regen die Gasmoleküle an, deren Ionisierungspotentiale unterhalb der Nennenergie der Entladungslampe liegen. Eine Übersicht ionisierbarer Moleküle und deren Ionisierungsenergien findet sich im „CRC handbook of chemistry and physics“ [31].

Der praktische Aufbau eines Photoionisationsdetektors ist in der nachfolgenden Abbildung 2.6 zu sehen. Der linke Bildteil beschreibt die Außenansicht eines Sensorkörpers,

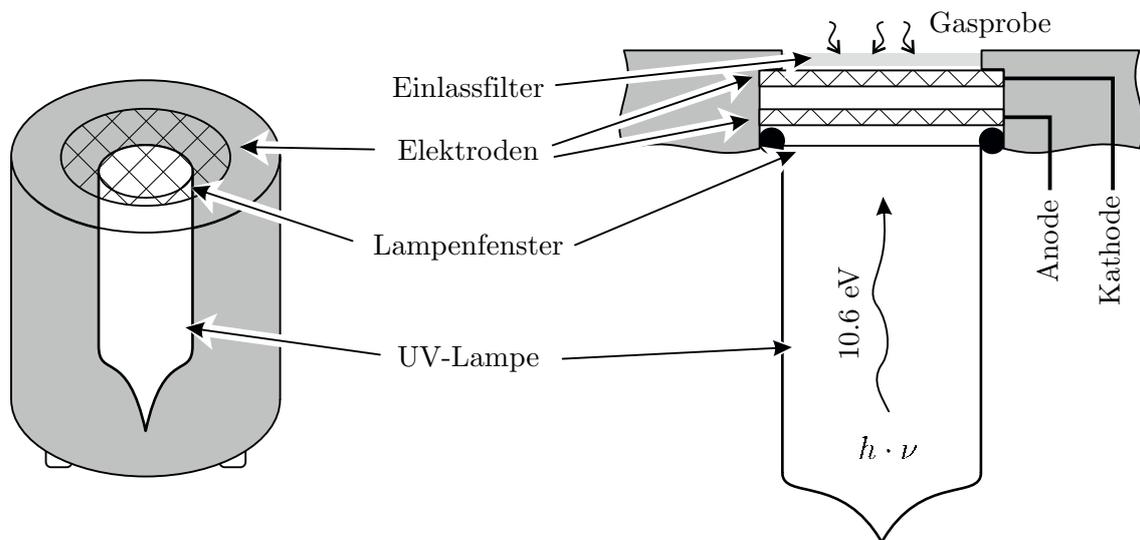


Abbildung 2.6.: Aufbau eines Photoionisationsdetektors

wie er von verschiedenen Herstellern produziert wird. Mittig befindet sich die UV-Lampe. Der am häufigsten verwendete Lampentyp ist eine Hochfrequenz-angeregte Krypton-Lampe mit Magnesiumfluorid-Fenster. Dieser Typ hat ein Ionisierungspotential von

## 2. Grundlagen

10,6 eV und womit die meisten aromatischen Verbindungen ionisiert werden können. Gleichzeitig ist die Lampe durch ihr stabiles Fenstermaterial für den mobilen Einsatz stoßunempfindlich und leicht zu reinigen [32]. Über dem Lampenkörper befindet sich der Stapel aus Gitterelektroden. Wie im rechten Teilbild zu erkennen, kann die Gasprobe durch einen Einlassfilter und die obenliegende Kathode hindurch in den Reaktionsraum gelangen. Der übrige Bauraum des Sensors wird für die elektronische Ansteuerung der Lampe, sowie für die Messelektronik des Ionenstroms genutzt.

Der zu messende Ionenstrom liegt im Bereich von Pico-Ampere und erfordert somit eine sehr rauscharme Verstärkerkette [33]. Da die Signalhöhe nicht nur von der Konzentration, sondern auch von der Ionisierbarkeit der Zielsubstanz abhängig ist, hat man einen stoffspezifischen Korrekturfaktor eingeführt. Dieser Responsefaktor ist zudem sensorabhängig und wird von den Herstellern bereitgestellt. Der Responsefaktor gibt die Steigung der linearen Gerätefunktion bei dem gemessenen Stoff bezogen auf Isobuten an. Damit besitzt ein Photoionisationsdetektor sehr gute Quantifikationseigenschaften. Durch diesen Sensortyp können primär Organika über einen Bereich von circa 10 ppb bis 1000 ppm nachgewiesen werden.

### 2.1.3. Elektrochemische Zellen

Bei elektrochemischen Zellen gehen die Analyt-Moleküle Redox-Reaktionen mit den in einem Elektrolyt eingebetteten Elektroden ein [20]. Wie in Abbildung 2.7 gezeigt, kann das Probengas durch eine Membran oder eine Lochblende in den Reaktionsraum, in dem sich auch der Elektrolyt befindet, gelangen. An dem speziell für die Zielsubstanz gewählten Elektrodenmaterial kommt es zu einer Reaktion bei welcher Ladungsträger freiwerden. An der Gegenelektrode rekombinieren die aus dem angeschlossenen Schaltkreis austretenden Ladungsträger wieder.

Dieser Prozess findet für die anorganischen Moleküle statt, sodass mit Elektrochemischen Zellen Gefahrstoffe, wie Phosgen, Phosphin oder luftgängige Schwefelverbindungen mit hoher Nachweisstärke erfasst werden können.

Je nach Aufbau des elektrochemischen Sensors, ist dieser für die jeweilige Zielsubstanz optimiert. So unterscheidet man nach anodisch oxidierenden und kathodisch reduzierenden Stoffen [34]. Die chemischen Reaktionen der Elektroden transferieren dabei entweder

## 2. Grundlagen

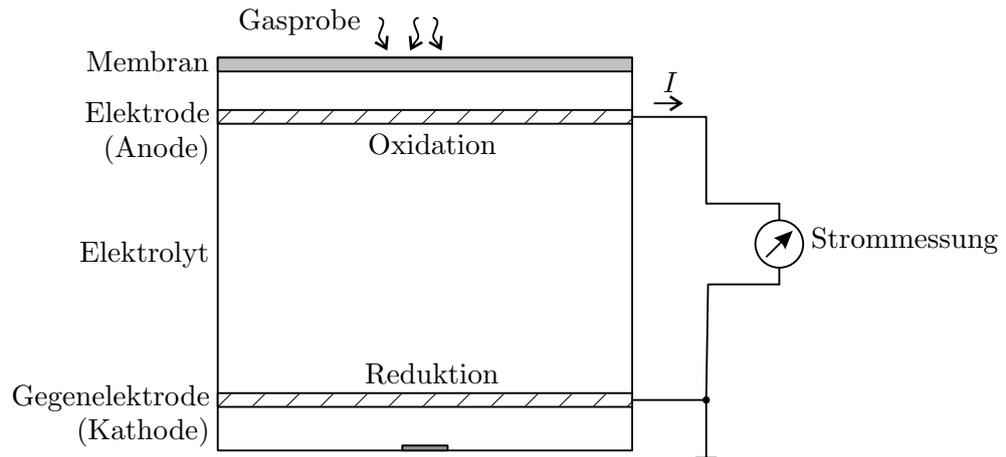
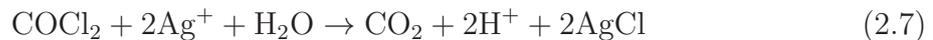


Abbildung 2.7.: Allgemeiner Aufbau einer Elektrochemischen Zelle zur Gasdetektion [10]

Elektronen direkt oder sie besitzen eine weitere vorgelagerte Reaktion. Einige Reaktionsgleichungen sind bei [35] zu finden.

Als Beispiel sei die zweistufige Reaktion von Phosgen ( $\text{COCl}_2$ ) mit Silberionen einer angepassten Anode genannt:



Eine solche Zelle mit Silber-Elektrodenmaterial weist Querempfindlichkeiten auf Phosphin ( $\text{PH}_3$ ) und Arsenwasserstoff ( $\text{AsH}_3$ ) auf, da diese Moleküle mit den Silberionen und dem Wasser des Elektrolyten oxidieren:



Um solche Querempfindlichkeiten zu reduzieren, wird neben dem Elektrodenmaterial auch das Membranmaterial für die Zielsubstanz angepasst. Über eine Filterwirkung der Membran können bestimmte Moleküle vom Reaktionsraum ausgeschlossen werden [36]. Bei einer Zelle mit offenen Klemmen kommt es durch die Ladungstrennung zur Polarisierung. Das sich einstellende Potential ist bis zu einem gewissen Grad konzentrationsabhängig und wird über die Nernst-Gleichung beschrieben nach:

$$\varphi = \varphi_0 + \frac{R \cdot T}{n_e \cdot F} \cdot \ln \frac{a_{\text{ox}}}{a_{\text{red}}} \quad (2.11)$$

## 2. Grundlagen

- $\varphi_0$  : Standardpotential
- $R$  : allgemeine Gaskonstante
- $T$  : Temperatur [K]
- $F$  : Faradaykonstante
- $a \propto p_a$  : Aktivität, proportional zum Partialdruck
- $n_e$  : elektrochemische Wertigkeit  
(Zahl der Elektronen in der Redoxgleichung)

Die Abhängigkeit der Elektrodenspannung von der Konzentration ist in dem in Gleichung 2.11 gezeigten Zusammenhang leicht störbar. Die Störung kann durch Reaktionskomponenten  $a_x$  oder aber auch durch das nicht vollständige Erreichen eines Potentialgleichgewichtszustandes hervorgerufen werden [34]. Daher werden die Elektroden extern niederohmig abgeschlossen, so dass ein Ladungsträgerausgleich hierüber stattfinden kann. Der Stoffumsatz an den Elektroden findet schneller statt, als der Stofftransport zu den Elektroden, bedingt durch die Diffusionseigenschaften des Elektrolyten. Es stellt sich damit ein Diffusionsgrenzstrom ( $i_{gr}$ ) ein, welcher proportional zur an der Zelle anliegenden Zielstoffkonzentration ist. Der Grenzstrom gilt für einen bestimmten Potentialbereich. Dieser Bereich ist in Abbildung 2.8 grau hinterlegt. Der sich einstellende Grenz-

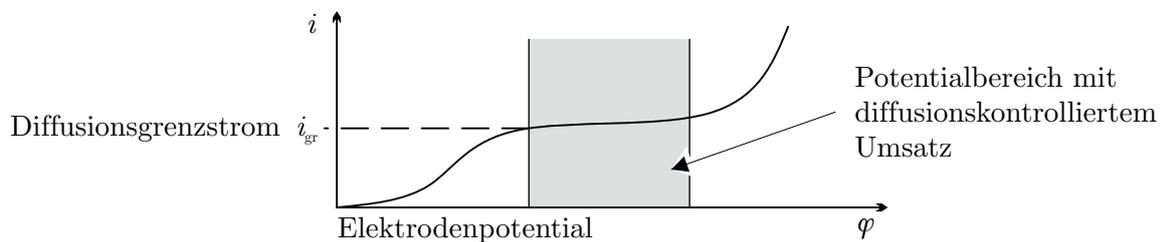


Abbildung 2.8.: Schematische Darstellung der Strom-Potentialkurve (grau unterlegt: Diffusionsgrenzstrombereich)

strom wird über Gleichung 2.12 formal beschrieben [35]. Hier gehen die geometrischen Parameter der Zelle ( $A$ : Fläche,  $d$ : Elektrodenabstand) und die Diffusionseigenschaften ( $D$ : Diffusionskoeffizient) des Elektrolyten ein.

$$i_{gr} = \frac{n_e \cdot F \cdot A \cdot D \cdot c_M}{d} \quad (2.12)$$

Der von der Stoffmengenkonzentration abhängige Diffusionsgrenzstrom  $i_{gr}(c_M)$  kann amperometrisch erfasst werden. Insgesamt verschieben sich die polarographischen Kurven,

## 2. Grundlagen

welche den Sensorstrom abhängig von der Elektroden-Potentialdifferenz darstellen, bei höheren Konzentrationen. Idealerweise bleibt der Potentialbereich mit diffusionskontrolliertem Umsatz über die Konzentration gleich. Die Abbildung 2.9 veranschaulicht den

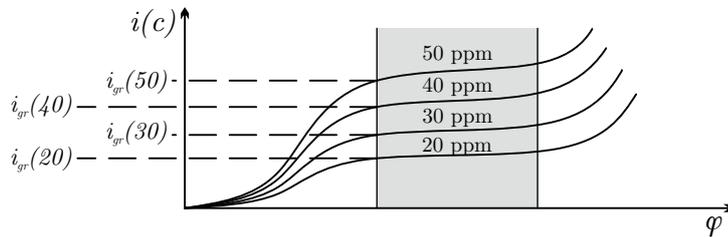


Abbildung 2.9.: Diffusionsgrenzströme bei unterschiedlichen Stoffkonzentrationen (schematisch)

Zusammenhang des Diffusionsgrenzstroms von der am Sensor anliegenden Konzentration: Die Kurvenform des Sensorstroms  $i$  bleibt erhalten, das Plateau des Grenzstroms  $i_{gr}$  verschiebt sich.

Bei manchen Reaktionspaarungen von Zielgas und Messelektrode ist der sich einstellende Stromfluss schon bei mittleren Konzentrationen in einer Größenordnung, bei der es zu einem Abfall der Sensorvorspannung kommen kann. Dies hätte zur Folge, dass der Sensor nicht mehr in seinem elektrochemischen Betriebspunkt arbeitet und so den Diffusionsgrenzstrom nicht erreichen kann. Lokale Verarmungs- und Anreicherungszone im Sensor führen dann zu Signaldrift und schließlich zur allmählichen Zersetzung der Zelle.

Diesem Verhalten kann durch einen erweiterten Sensoraufbau entgegengewirkt werden (siehe Abbildung 2.10). Über eine Referenzelektrode wird ständig die Spannung der Zelle gemessen. Eine Potentiostatschaltung regelt die an der Gegenelektrode anliegende Spannung. Dazu wird die Spannungsdifferenz zwischen Arbeits- und Referenzelektrode herangezogen. Bei manchen Zellen muss der Gegenelektrode zusätzlich eine Bias-Spannung aufgeprägt werden um die Zelle in ihrem Diffusionsgrenzstrombereich betreiben zu können.

Aufgrund des häufig auf Wasser basierenden Elektrolyten, ist die Lebensdauer von elektrochemischen Zellen maßgeblich durch mögliche Austrocknung begrenzt. Mess-Zellen mit einem Festkörper-Elektrolyten besitzen eine längere Lebensdauer [37].

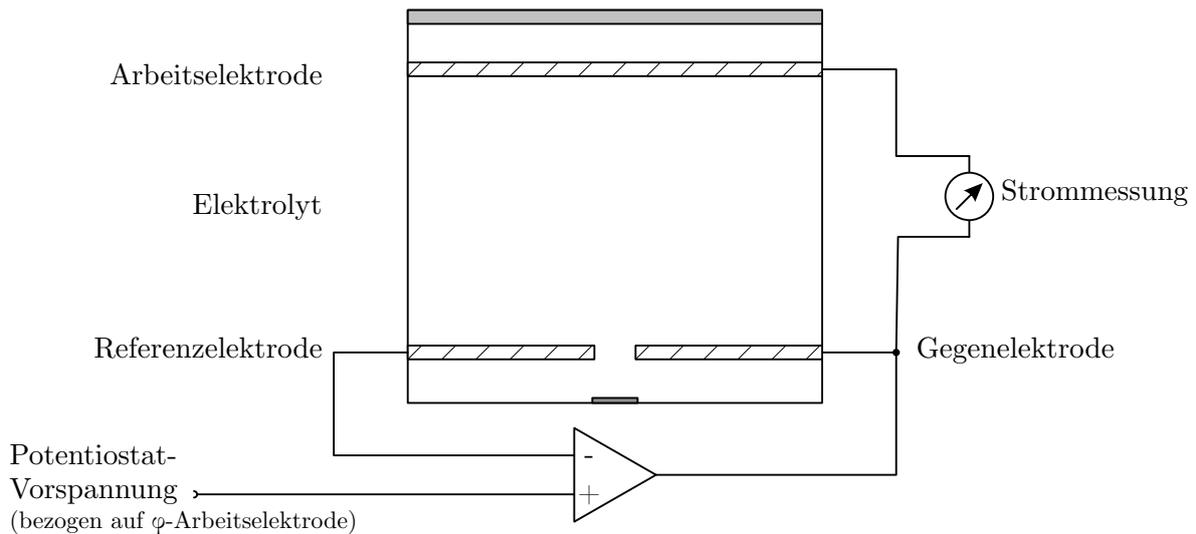


Abbildung 2.10.: Aufbau einer Drei-Elektroden-Zelle

## 2.2. Ionenmobilitätsspektrometer

Die Ionen-Mobilitäts-Spektrometrie ist ein Analyseverfahren, bei welchem die Laufzeit von Ionen durch eine Driftröhre gemessen wird. Diese erfasste Driftzeit ist abhängig von der stoffspezifischen Ionenmobilität. Über dieses Verfahren können Substanzen mit extrem hoher Nachweisstärke unterschieden und damit identifiziert werden.

Durch ein Ionenmobilitätsspektrometer nachweisbar sind Moleküle, die eine Protonenaffinität oder Elektronegativität aufweisen. In diese Klasse gehören die meisten Gefahrstoffe und besonders auch die chemischen Kampfstoffe, weshalb diese Detektionstechnologie auch im militärischen Bereich Anwendung findet.

Ein Ionenmobilitätsspektrometer besteht aus drei Hauptkomponenten: Einer *Ionenquelle*, einer *Driftstrecke* und einem *Ionendetektor*.

Im Bereich der Quelle werden den eintretenden Ziel-Molekülen Ladungen aufgeprägt. Über ein angelegtes elektrisches Feld haben die gebildeten Ionen das Bestreben sich in Richtung des Detektors zu bewegen. Aufgrund eines Haltemechanismus bleiben die Ionen zunächst im Bereich der Ionenquelle und werden erst durch einen Startimpuls auf die Driftstrecke freigegeben.

## 2. Grundlagen

Die Strecke besitzt ein homogenes elektrisches Feld, welches die Ionen konstant in Richtung des Detektors beschleunigt. Auf ihrem Weg kollidieren die Ionenpakete immer wieder mit den Molekülen des Driftgases. Abhängig von Masse, Ladung und auch effektiver Molekülgröße, benötigen die Substanzen eine unterschiedliche Zeit um die Driftröhre zu passieren. Der Detektor erfasst den sich ergebenden Strom der eintreffenden Ionen.

Neben dem beschriebenen Flugzeit-Ionen-Mobilitäts-Spektrometer (TOF-IMS) gibt es noch weitere Varianten: Ein Beispiel ist das *Ballistik-Ionenmobilitätsspektrometer*, bei welchem die Ionen über ein orthogonal zur Flugbahn angelegtes elektrisches Feld abgelenkt werden. Je nach Mobilität der Ionen verschiebt sich ihr Auftreffpunkt. Über die Länge der Zielfläche sind mehrere Detektoren angeordnet. Somit ist jeder Faraday-Detektor spezifisch für einen bestimmten Ionen-Mobilitätsbereich [11].

### 2.2.1. Ionenmobilität

Allen IMS-Typen gemein ist die Abhängigkeit des analytischen Trennergebnisses von der Ionenmobilität. Der Begriff *Mobilität* beschreibt in diesem Zusammenhang, wie schnell sich ein Ion innerhalb des Driftgases aufgrund des vorhandenen elektrischen Feldes fortbewegen kann. Der Quotient aus Driftgeschwindigkeit  $v_d$  und elektrischem Feld  $E$  beschreibt damit die Ionenmobilität  $K$ . Sie wird in der Einheit  $[\text{cm}^2/(\text{V} \cdot \text{s})]$  angegeben.

$$K = \frac{v_d}{E} \quad (2.13)$$

Die Ionenmobilität und damit die Driftgeschwindigkeit kann ebenfalls analytisch aus den Eigenschaften der Moleküle bestimmt werden. Dabei gehen die Teilchendichte  $N$  und Masse der Analyt-Ionen  $m$  sowie deren Stoßquerschnitt  $\Omega_D$  ein [38].

$$K = \frac{3q}{16N} \cdot \sqrt{\frac{1}{m_{\text{Ion}}} + \frac{1}{M}} \cdot \sqrt{\frac{2\pi}{k_B T}} \cdot \frac{1}{\Omega_D} \quad (2.14)$$

## 2. Grundlagen

$q$  : Ladung des Ions

$N$  : Teilchendichte

$m_{\text{Ion}}$  : Masse des Ions

$M$  : Masse der Driftgasmoleküle

$k_B$  : Boltzmann-Konstante

$\Omega_D$  : Stoßquerschnitt

Wie die Gleichung 2.14 zeigt, wird die Mobilität bei schweren Ionen maßgeblich vom Stoßquerschnitt  $\Omega_D$  bestimmt. Diese abgeleitete Größe ist stoff- und temperaturabhängig und wird bestimmt durch:

$$\Omega_D = \pi r_m^2 \Omega^{(1,1)*}(T^*) \quad (2.15)$$

Der Querschnitt des Ions  $r_m$  und das Kollisionsintegral  $\Omega^{(1,1)*}$  sind Stoffparameter. Eine Auflistung dieser Parameter ist zu finden bei Klein et. al. [39].

Da die Driftgeschwindigkeit und damit die Ionenmobilität jedoch immer abhängig sind von Temperatur und Druck im Driftraum, werden für eine Vergleichbarkeit der Stoffe die *reduzierten Ionenmobilitäten*  $K_0$  herangezogen. Dazu wird die Ionenmobilität auf Normbedingungen von Temperatur und Druck bezogen ( $T_0 = 273,15$  K,  $P_0 = 101,325$  kPa):

$$K_0 = K \cdot \frac{P_D}{P_0} \cdot \frac{T_0}{T_D} \quad (2.16)$$

Beim TOF-IMS wird die Ionenmobilität in der Praxis über die Driftröhrengometrie (Länge:  $l_D$ ) und die gemessene Driftzeit  $t_d$  mit  $v_d = \frac{l_D}{t_d}$  berechnet. Die elektrische Feldstärke ist ebenfalls abhängig von der Länge der Driftröhre, sowie von der angelegten Beschleunigungsspannung  $U$ :  $E = \frac{U}{l_D}$ . Zusammen mit Gleichung 2.13 erhält man folgende Berechnungsvorschriften für  $K$  und  $K_0$ :

$$K = \frac{l_D^2}{t_d \cdot U} \quad (2.17)$$

$$K_0 = \frac{l_D^2}{t_d \cdot U} \cdot \frac{P_D}{101,325 \text{ kPa}} \cdot \frac{273,15 \text{ K}}{T_D} \quad (2.18)$$

### 2.2.2. Prinzip eines Flugzeit-Ionenmobilitätsspektrometers

In der nachfolgenden Abbildung 2.11 ist das Schema eines TOF-IMS gezeigt. Hierbei

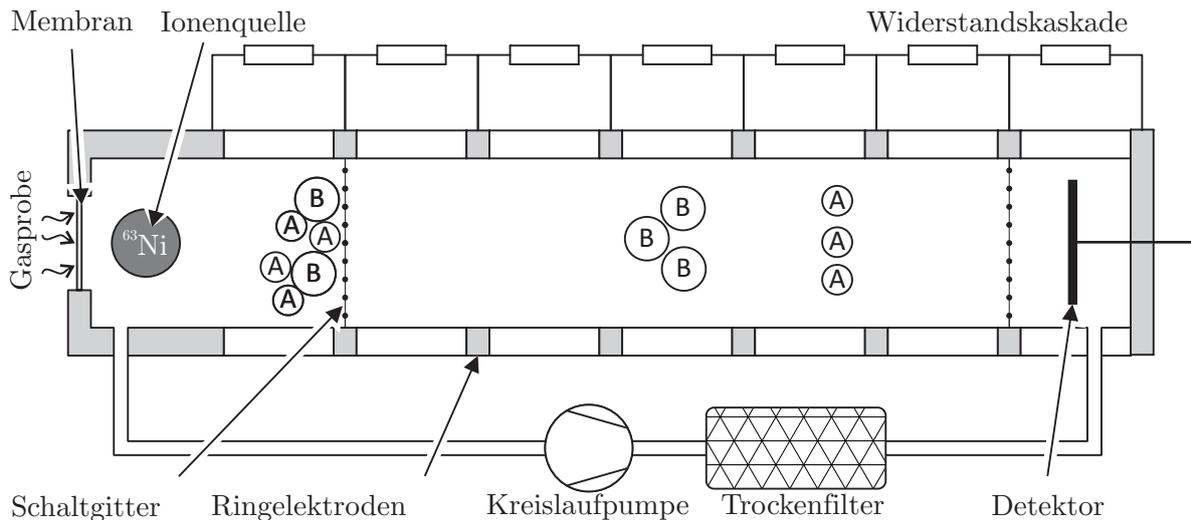


Abbildung 2.11.: Aufbau der Driftstrecke eines Ionenmobilitätsspektrometers [10]

handelt es sich um Ionenmobilitätsspektrometer mit Membraneinlass und einem Betastrahler als Quelle der Ionisierung. Typische Driftröhren haben eine Länge von etwa 10 cm. Auf der linken Seite der Abbildung befindet sich der Einlass. Hier gelangen die Analyt-Moleküle in den Ionisationsraum. Eine Strahlungsquelle mit dem Isotop Nickel-63 erzeugt über einen Sekundärionisierungsprozess sowohl positive als auch negative Produktionen eines Analyten. Das elektrische Feld liegt über der gesamten Röhre an und wird in dem gezeigten Aufbau durch Ringelektroden und deren Beschaltung über eine Widerstandskaskade homogenisiert. Ein Ionentor in Form eines Bradbury-Nielsen-Gitters verhindert zunächst eine Bewegung der Ionen in Richtung des Detektors [40]. Durch Abschalten der Gitterspannung für eine kurze Zeitspanne kann ein Ionenpaket in die Driftstrecke gelangen. Hier kollidieren die Ionen mit dem Neutralgas. Wie in Kapitel 2.2.1 beschrieben, werden die Ionen über die Flugzeit getrennt. Im gezeigten Aufbau ist dies gereinigte und getrocknete Luft. Am Faraday-Detektor rekombinieren die Ladungsträger wieder und erzeugen so einen messbaren Strom. Dem Detektor vorgeschaltet, befindet sich an der letzten Ringelektrode ein Aperturgitter. Dies dient zur Verbesserung des Signal-Rausch-Verhältnisses.

Je nach Polarität der angelegten Driftspannung werden entweder Anionen oder Kationen in Richtung des Detektors beschleunigt. Sollen mit dem Aufbau beide Ionenspezies erfasst werden können, so wird die über die Driftröhre anliegende Hochspannung zyklisch umgepolt. Die zur Beschleunigung verwendete Spannung bei dieser Geometrie liegt bei

zwei bis drei Kilovolt. Dies entspricht einer Feldstärke von 200 V/cm bis 300 V/cm [41]. Der Wechsel der Polarität der Hochspannung benötigt einige Zeit, da man versucht zu hohe Ströme zu vermeiden. Für die Messung bedeutet dies, dass Spektren nur mit Zeitabständen von einigen Sekunden aufgenommen werden können. In einigen Anwendungen ist dies von Nachteil: Im Zusammenspiel mit einer vorgeschalteten gaschromatographischen Trennung beispielsweise, muss der Säulenausgang kontinuierlich überwacht werden. Eine Möglichkeit dies zu realisieren, ist der Einsatz zweier analytisch parallelgeschalteter Ionenmobilitätsspektrometer. Das nachfolgend beschriebene Doppelröhren-IMS, das im Rahmen des DACHS-Projektes von Ungethüm [11] entwickelt worden ist, ist ein optimierter Aufbau zur Detektion beider Ionenspezies. Es kommt aufgrund seiner Injektionstechnik mit nur einer radioaktiven Quelle aus.

### 2.2.3. Prinzip eines

#### Doppelröhren-Flugzeit-Ionenmobilitätsspektrometers

Der in Abbildung 2.12 dargestellte Aufbau gestattet das gleichzeitige Erfassen positiver und negativer Ionen. Aufgrund der parallelen Anordnung beider Driftröhren müssen die Ionen zunächst nach ihren Ladungen getrennt in eine Warteposition vor die jeweiligen Driftstrecken positioniert werden. Beide Driftröhren teilen sich einen Einlassbereich mit Membran und Ionenquelle. Hier findet der in Kapitel 2.2.4 beschriebene Ionisierungsprozess statt. Die Ionen werden in diesem Aufbau jedoch nicht durch ein Schaltgitter zurückgehalten, sondern durch eine Verzerrung des elektrischen Feldes im Einlassbereich. Die in der Abbildung als Injektionselektroden gekennzeichneten Flächen besitzen Potentiale, welche die jeweiligen Kat- bzw. Anionen ablenken. Es stellt sich eine Ionenwanderung vom Reaktionsraum in der Nähe der Ionenquelle zu den entsprechenden Pull-Elektroden ein. Hier kommt es zur Rekombination der Ladungsträger. Bei aktiven Pull-Elektroden findet somit keine Drift der Ionen in Richtung der Detektoren statt. Für eine Injektion werden die Pull-Elektroden deaktiviert und die Push-Elektroden, mit einem für die jeweilige Ionenart abstoßenden Potential, aktiviert. In diesem Potentialzustand der Driftröhre kann ein Ionenpaket in Richtung des Detektors wandern. Die Laufzeit wird ab diesem Zeitpunkt bis zum Auftreffen gemessen. [11]

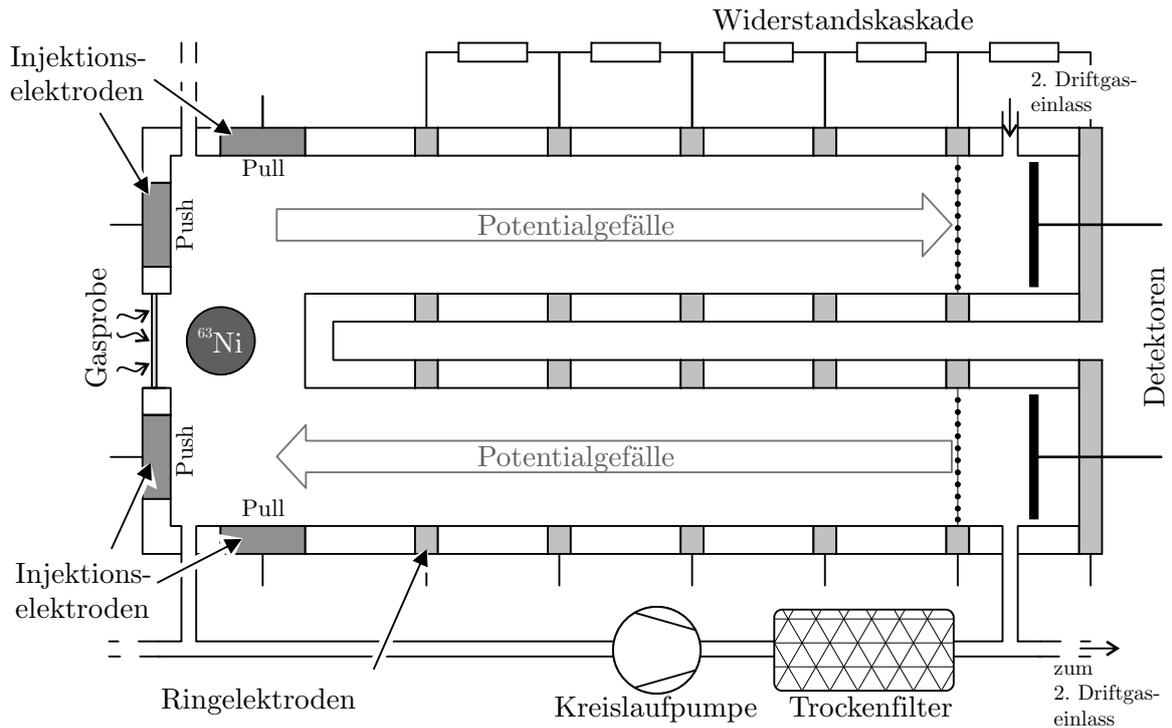


Abbildung 2.12.: Aufbau eines Doppelröhren-IMS [42]

### 2.2.4. Ionisierungsprozess

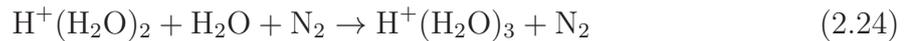
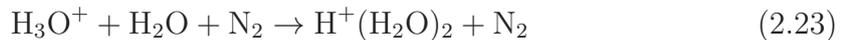
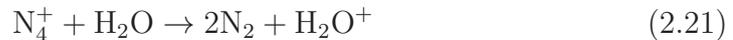
In den beiden gezeigten IMS-Prinzipien dient ein  $\beta$ -Strahler als Ionenquelle. Das eingesetzte Nickel-63-Isotop zerfällt unter Freisetzung eines Elektrons mit einer maximalen Energie von 66.94 keV [43]. Durch die freiwerdenden Elektronen wird ein Sekundärionisationsprozess in Gang gesetzt. Es werden über eine Reaktionskette *Reaktantionen* gebildet, welche aus der Restfeuchte des Driftgases hervorgehen.

**Positive Reaktantionen** Zur Bildung von positiven Reaktantionen werden zunächst Stickstoffmoleküle der Trägerluft ionisiert. Diese übertragen wiederum Ladung an Was-



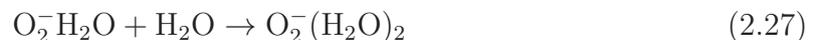
## 2. Grundlagen

Sermoleküle, wie in den folgenden Gleichungen gezeigt [44]:



Die Reaktionsprodukte  $\text{H}^+(\text{H}_2\text{O})_2$  und  $\text{H}^+(\text{H}_2\text{O})_3$  aus 2.23 und 2.24 sind hierbei stabile Reaktantionen.

**Negative Reaktantionen** Die negativen Reaktantionen werden ebenfalls durch eine Reaktionsfolge gebildet. Die durchs Isotop freigesetzten Elektronen ionisieren zunächst Sauerstoffmoleküle des Driftgases. In der Wasserchemie durch die nachfolgend beschriebene Reaktionskette beschrieben [11]:



Am Ende der Reaktionskette steht damit nach Gleichung 2.27 das negative Reaktantion  $\text{O}_2^-(\text{H}_2\text{O})_2$ .

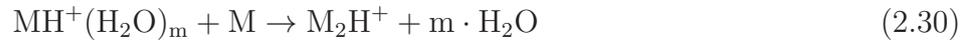
**Produktionen** Die Zielsubstanzmoleküle (im Folgenden als M bezeichnet) bilden durch Ladungstransfer von den zuvor gebildeten Reaktant-Ladungsträgern die Produktionen. Die Neigung einer Zielsubstanz von den Reaktantionen eine Ladung aufgeprägt zu bekommen hängt von der Protonenaffinität und der Elektronegativität des Moleküls ab [11]. Positive Produktionen werden durch eine protolytische Reaktion, also durch Transfer eines Protons aus dem positiven Reaktantion  $\text{H}^+(\text{H}_2\text{O})_n$  gebildet. Es werden vornehmlich Monomere ( $\text{MH}^+$ ) nach Gleichung 2.28 gebildet [11]:



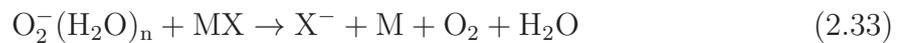
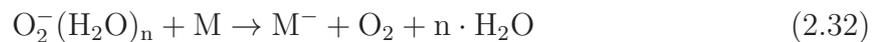
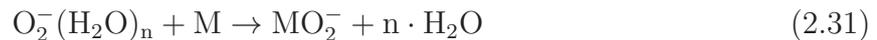


## 2. Grundlagen

Bei höheren Konzentrationen der Zielsubstanz können auch positive Dimere ( $M_2H^+$ ) nach folgender Reaktionskette gebildet werden [11]:



Zielsubstanzmoleküle (M), welche negativ ionisierbar sind, bilden durch Sekundärionisation mithilfe der Reaktantionen aus Gleichung 2.27 ihre Produktionen ( $MO_2^-$ ,  $M^-$  oder  $X^-$ ).



Bei den gezeigten drei Möglichkeiten handelt es sich um Produktionen die über Ionen-transfer (Gleichung 2.31), Ladungstransfer (Gleichung 2.32) oder dissoziativen Ladungstransfer (Gleichung 2.33) gebildet werden können [41].

### 2.2.5. Eigenschaften der Spektren

Die Flugzeit der Ionen liegt in dem gezeigten Aufbau im zweistelligen Millisekundenbereich. Als Spektrum wird die Darstellung des Detektorstroms über die Flugzeit bezeichnet. Die wichtigsten Eigenschaften innerhalb dieser Darstellung sind die Höhen und Lagen der sich ergebenden lokalen Maxima. Es werden die Signale beider Detektoren separat dargestellt. Bei einer Messung ohne Vorliegen eines zu analysierenden Probengases sind neben dem Driftgas in der Strecke nur die Reaktantionen nach Gleichung 2.23 und 2.24 beziehungsweise nach Gleichung 2.27 in der Driftstrecke vorhanden. Somit werden auch nur diese Ionen vom Detektor erfasst. In den sich ergebenden Nullspektren sind nur die Peaks der Reaktantionen sichtbar. In dem Diagramm (siehe Abbildung 2.13) ist die Driftzeit des RIP bzw. RIN ablesbar. Diese ist abhängig von Druck und Temperatur in der Driftrohre und kann nach Gleichung (2.16) auf die reduzierte Ionenmobilität  $K_0$  überführt werden.

## 2. Grundlagen

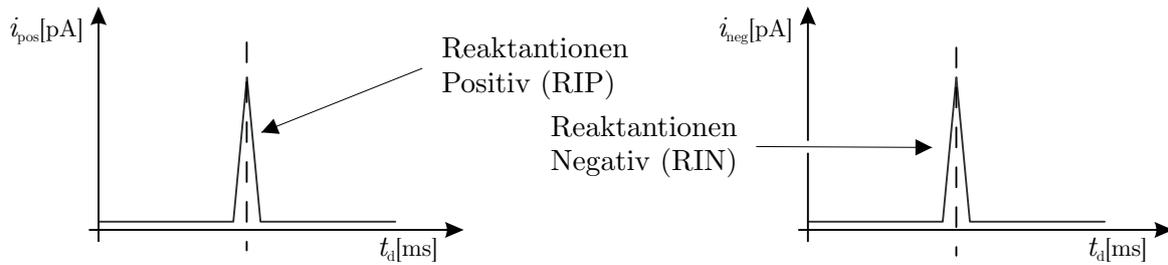


Abbildung 2.13.: Möglichkeit der Darstellung von Ionen-Mobilitäts-Spektren

### 2.3. Gaschromatographie

Die Gaschromatographie ist ein Verfahren, durch welches Stoffgemische aufgrund der Löslichkeit ihrer Einzelkomponenten über die Zeit getrennt werden. Das Gemisch wird mithilfe eines Trägergases durch eine Trennkapillare transportiert. Dabei benötigt jede Komponente des Gemisches eine bestimmte Zeit zum Passieren der Trennstrecke. Die Wandergeschwindigkeit durch die gaschromatographische Trennstrecke (GC-Säule) hängt von der Sorptions- und Wechselwirkungsfähigkeit zwischen Kapillarwand und Analyt ab. Der Analyt wechselt in dem Trennprozess immer wieder von der mobile Phase in die stationäre Phase der Kapillarinnenwand. Weitere Einflussfaktoren sind die Temperatur und der Volumenstrom. In Abbildung 2.14 ist die Trennung von zwei Sub-

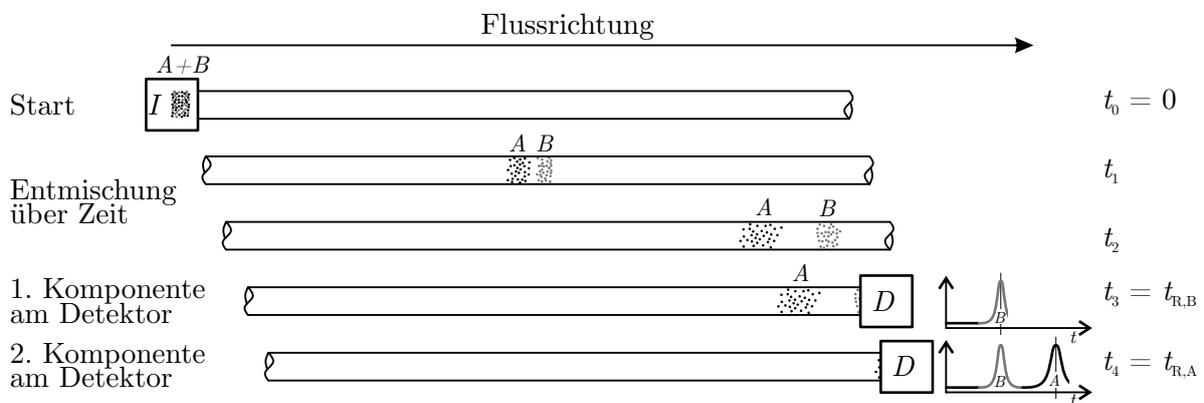


Abbildung 2.14.: Prinzip der gaschromatographischen Trennung ( $I =$  Injektor,  $D =$  Detektor)

stanzen durch eine Kapillar-Trennsäule gezeigt. Die Stoffe  $A$  und  $B$  liegen beim Start der Chromatographie ( $t_0$ ) zunächst als Gemisch vor. Mithilfe eines Trägergases wird

## 2. Grundlagen

das Stoffpaket durch die Kapillare transportiert. Die Wandergeschwindigkeit der Einzelkomponenten wird auf ihrem Weg durch wiederkehrende Adsorption und Desorption an der Kapillarwand bestimmt. Über die Wanderzeit entmischen sich die Komponenten, in der Abbildung ist zum Zeitpunkt  $t_1$  eine Auftrennung beider Substanzen erkennbar. Durch Diffusion (Eddy-Diffusion und Längsdiffusion) werden die Pakete im Zeitverlauf auseinandergezogen. Zum Zeitpunkt  $t_2$  ist es beim Paket  $A$  mit der längeren Retentionszeit bereits zu einer merklichen Verbreiterung gekommen. Zum Zeitpunkt  $t_3$  wird Substanz  $B$  vom Detektor registriert. Die Dauer von Injektion bis zur Registrierung des Signalmaximums wird als Retentionszeit  $t_R$  bezeichnet. Mit der Detektion der letzten Gemischkomponente ist das Chromatogramm vollständig. Wie im Diagramm im rechten Bildteil von Abbildung 2.14 dargestellt, sind beide Substanzen zeitliche voneinander getrennt.

Gaschromatographische Trennsäulen werden mit unterschiedlichen Injektionseinheiten und Detektoren kombiniert. Häufig werden sehr breitbandige Detektoren wie Flammenionisationsdetektoren, Elektroneneinfangdetektoren oder auch Massenspektrometer eingesetzt. Die Trennkapillare kann in der Laboranalytik Einsatzlängen von einigen zehn Metern besitzen.

### 2.3.1. Einflussgrößen der Gaschromatographie

Die Abhängigkeit der Retentionszeit von den Geräteparametern kann formal beschrieben werden. Haupteinflussfaktoren sind Säulenlänge, Trägergasgeschwindigkeit ( $v_C$ ), Säulentemperatur und die Stoffeigenschaften. Die Gleichung 2.34 beschreibt diesen Zusammenhang [45][46]. Die Temperatur geht über den Retentionszeitindex  $k$  ein.

$$t_R = \frac{l_C}{v_C}(1 + k) \quad (2.34)$$

Der Retentionszeitindex kann bestimmt werden über die nachfolgende Gleichung:

$$\ln k = \ln \frac{\alpha}{\beta} + \frac{\Delta H}{RT} \quad (2.35)$$

Hierbei können die Enthalpie- und Entropie-Terme  $\frac{\alpha}{\beta}$  und  $\Delta H$  praktisch nach Guan et. al [47] ermittelt werden. Sie beschreiben die temperaturabhängige Wechselwirkungsfähigkeit zwischen Eluent und stationärer Phase. Die Säulentemperatur wird mit dem

## 2. Grundlagen

Quotienten der Temperatur ( $T$ ) zusammen mit der allgemeinen Gaskonstante  $R$  in der Gleichung berücksichtigt.

Bei der Auswertung von Chromatogrammen wird die ermittelte Retentionszeit genutzt um Stoffe zu bestimmen. Die beiden Gleichungen zeigen, dass diese Zeit  $t_R$  abhängig ist von der Temperatur und von der Trägergasgeschwindigkeit. Die weiteren Terme der Gleichungen sind substanz- und säulenspezifisch. Für eine Substanzidentifikation mit Hilfe eines Chromatogramms sind daher neben der Retentionszeit die beiden Kennwerte der Säulentemperatur  $T$  und der Trägergasgeschwindigkeit  $v_C$  nötig. Die Trägergasgeschwindigkeit wird dabei üblicherweise als abgeleitete Größe bestimmt. Dies kann über den gemessenen Massenstrom oder den Säulenvordruck geschehen.

### 2.3.2. Methoden der Injektion von Probengas

Es haben sich mehrere Verfahren der Aufbringung der Gasprobe auf gaschromatographische Trennsysteme etabliert. Der Injektionszeitpunkt ist ein wichtiger Parameter für die Laufzeitmessung. Daher muss dieser analytisch klar definiert sein und elektrisch genau erfasst werden können. Für eine Vergleichbarkeit von Messläufen muss vor allem die Wiederholbarkeit einer Injektionsmethode sichergestellt sein. In dieser Arbeit werden drei Verfahren vorgestellt: Loop-Injektion, Thermische Desorption mit Anreicherung und Gegenstrom-Injektion.

#### Loop-Injektion

Bei dieser Methode wird eine kleine Menge Probengas als Stoffpaket aus einem Speichervolumen zum Injektionszeitpunkt in die Säule injiziert. Ein häufiger Aufbau ist in Abbildung 2.15 zu sehen. Hier kommt ein Multiport-Ventil mit mehreren gleichzeitig geschalteten Ein- und Ausgängen zum Einsatz. Im, in Teilbild a) gezeigten Schaltzustand, wird die Loop mit der Probe durchspült. Hierdurch wird sichergestellt, dass immer ein definiertes Volumen der Probe vorgehalten wird. Gleichzeitig wird ein Trägergas durch die gaschromatographische Trennsäule geleitet. Die dauerhafte Durchspülung mit Inertgas sorgt dafür, dass möglichst keine Stoffreste eines vorherigen Chromatographie-Laufs in der Säule persistent bleiben.

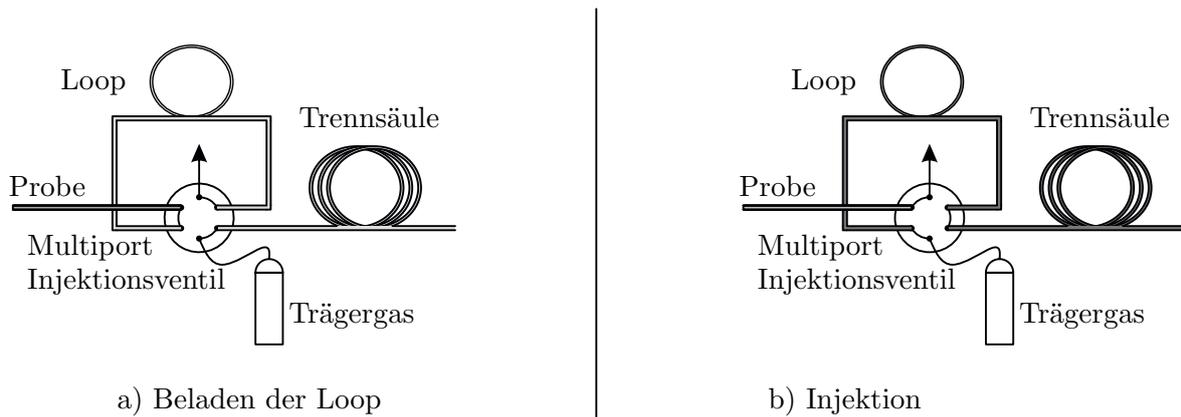


Abbildung 2.15.: Loop Injektor auf Basis eines Multiport-Ventils

Die Injektion des in der Loop gehaltenen Probengasvolumens ist in der Teilabbildung b) dargestellt. Das Multiport-Ventil verbindet in diesem Schaltzustand das Trärgas mit dem Eingang der Loop und deren Ausgang mit der Trennsäule. Durch den Vordruck des Trärgases wird das Stoffpaket auf die Trennstrecke aufgebracht. Der Probeneingang des Systems wird auf einen „waste“-Ausgang geschaltet.

### Thermische Desorption mit Anreicherung

Beim Prinzip der thermischen Desorption findet im ersten Schritt eine Anreicherung eines Analyten durch Adsorption statt. Im zweiten Schritt folgt die Desorption, welche den Injektionszeitpunkt des Trennverfahrens bestimmt.

Ein Adsorptionsmaterial wird mit der zu untersuchenden Gasprobe durchspült. Im kalten Zustand werden Analytmoleküle an der Oberfläche eines Substrats adsorbiert. Das als Adsorptionsmittel verwendete Substrat besteht aus kleinen hochporösen Polymerkugeln. Durch die Porosität wird die, für die Chemisorption aktive Oberfläche, maximiert. Die Adsorption, also die Bindung von Analytmolekülen an das Adsorbens erfolgt bei der Chemisorption auf Grund von Van-der-Waals-Kräften. Durch Erhitzen werden die Van-der-Waals Bindungen wieder gelöst. [48]

Dieser Prozess wird dazu benutzt eine geringe Stoffkonzentration eines Probengases anzureichern. Durch die Dauer der Adsorption kann bestimmt werden, wie viele Analyt-

## 2. Grundlagen

moleküle auf dem Adsorbens gebunden werden. Durch längeres Bespülen erhöht sich die Beladung des Adsorptionsmittels.

Nach einem bestimmten Volumen des zu analysierenden Gases, mit dem das Adsorbens durchspült wurde, sind alle freien Positionen zu Bindung an das Adsorptionsmittel belegt. Es können keine weiteren Moleküle mehr durch das Adsorbens aufgenommen werden. Diese substanzabhängige Volumengrenze wird als Durchbruchvolumen bezeichnet, da jedes weitere Molekül keinen Bindungspartner mehr findet und die Anreicherungsrichtung passieren kann.

Bei der Desorption durch Temperaturerhöhung wird dann eine höhere Konzentration des Analyten freigegeben.

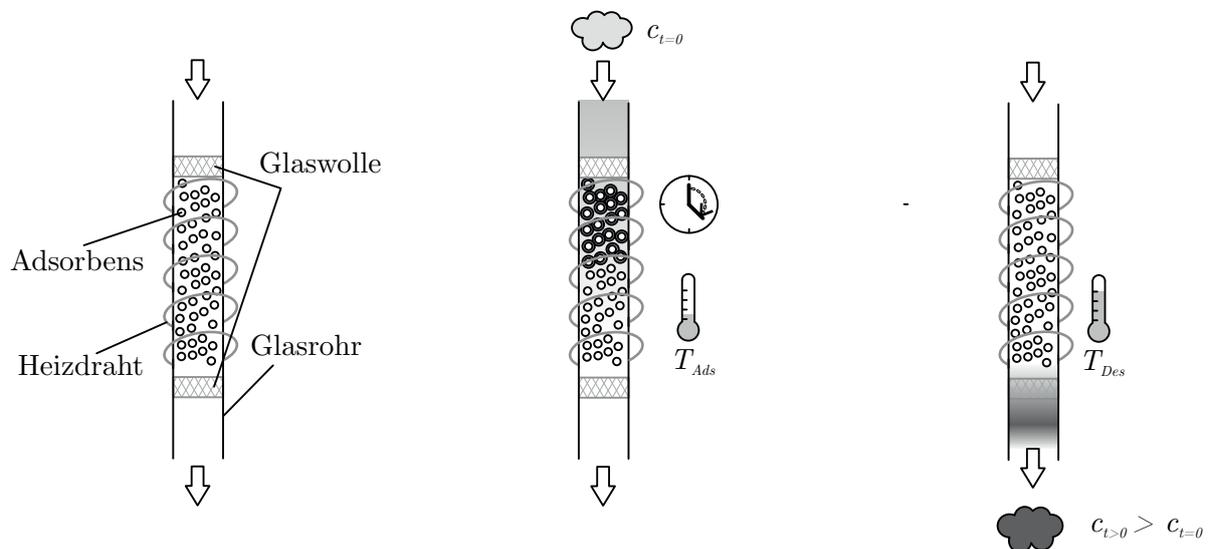


Abbildung 2.16.: Anreicherung über thermische Ad- und Desorption [10]

In der Abbildung 2.16 sind die Adsorptions- und Desorptionsvorgänge anhand eines, mit Adsorptionsmittel gefüllten, Glasrohres gezeigt. Das linke Teilbild zeigt den Aufbau, das mittlere die Adsorptionsphase und das rechte Teilbild die Desorptionsphase.

Das Rohr ist mit einem Widerstandsdraht umwickelt, welcher die für die Desorption nötige thermische Energie bereitstellt.

## 2. Grundlagen

Während der Adsorption zum Zeitpunkt  $t = 0$  wird die Gasprobe mit der Konzentration  $c$  für eine gewisse Dauer durch die Anreicherungsrohre geleitet. Die Adsorptionstemperatur  $T_{Ads}$  ist geringer als die Desorptionstemperatur und liegt bei gängigen Adsorbentien im Bereich der Zimmertemperatur. Zur Desorption wird das Stoffpaket bei hoher Desorptionstemperatur  $T_{Des} \gg T_{Ads}$  mit potenziierter Konzentration wieder freigegeben. Dies kann, wie hier dargestellt in gleicher Flussrichtung, wie die Beladung erfolgen. Optimal ist eine Flussumkehr zur Desorption, da die Beladung vom Einlass her beginnt. Die Probe muss in diesem Fall nicht den gesamten Anreicherungsbereich durchqueren und es ist eine höhere Ausbeute zu erwarten.

Das Adsorptionsmittel wird abhängig von den Zielsubstanzen gewählt. Für flüchtige organische Komponenten kann z.B. Tenax eingesetzt werden [49]. Eine häufige Nebenbedingung ist eine geringe Affinität zum in der Umgebungsluft vorhandenen Wasser [50].

### **Gegenstrom-Injektion**

Bei der Analytik kleinster Konzentrationen ist die Kontaminationsfreiheit gasführender Komponenten von besonderer Bedeutung. Um eine Kontamination und ein damit einhergehendes chemisches Memory im Gaspfad eines analytischen Instruments zu reduzieren, werden möglichst inerte Materialien verwendet. Ebenso wird die mit dem Analytgas in Kontakt stehende Oberfläche möglichst gering gehalten. Polymerdichtungen und Ventile mit ihren Dichtflächen besitzen nur bedingte inerte Eigenschaften. Ventile im analytischen Pfad sind daher meist von Nachteil für die Stoffanalyse.

Ein Gegenstrom-Injektor ist eine Methode der Probenaufbringung, welche ohne schaltende Komponenten in der Einlassfluidik auskommt [51]. Die Abbildung 2.17 beschreibt den Aufbau dieser Injektor-Art. Der Einlass wird permanent mit einem neutralen Trägergas gespült. Gleichzeitig erzeugt eine Pumpe am Ende der gaschromatographischen Trennstrecke einen Fluss durch die Kapillare und den damit verbundenen Injektor. Der Volumenstrom des Neutralgases ist etwas höher eingestellt, als der Volumenstrom durch die analytischen Komponenten. Somit gelangt im Standby-Betrieb nur das Trägergas durch die Trennstrecke zum Detektor. Alternativ zum dargestellten Druckbehälter kann

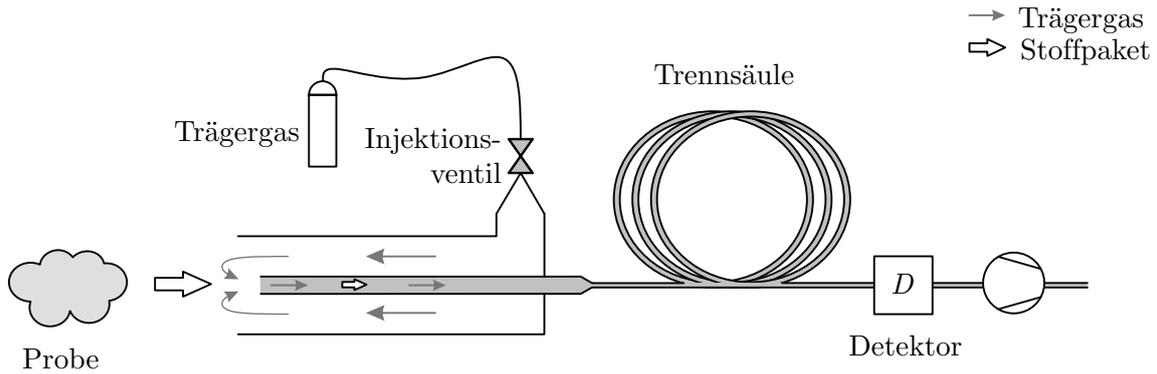


Abbildung 2.17.: Gegenstrom Injektor

das Neutralgas auch durch eine Pumpe und ein dahinter geschaltetes Molekularsieb aus Umgebungsluft erzeugt werden.

Für den Injektionsmoment wird der Neutralgasstrom kurzzeitig durch ein Ventil unterbrochen. Zu diesem Zeitpunkt wird die Probe vom Injektor-Einlass angesaugt. Ein Stoffpaket wird dann in die Säule injiziert und über die Trennstrecke gaschromatographisch aufgetrennt. Die Dauer der Trägergasunterbrechung bestimmt abzüglich einer Totzeit die Injektionsdauer.

## 2.4. Verfahren der Mustererkennung und Klassifikation

Die im Kapitel 2.1 vorgestellten Sensoren besitzen Querempfindlichkeiten, welche die Aussagekraft eines Einzelsensors für eine Stoffidentifikation verschlechtert. Erst die Kombination von mehreren wenig spezifischen Sensoren kann eine gute Identifikationsfähigkeit bei unbekannter Probe hervorbringen. Dies kann erreicht werden durch die Bildung von Schnittmengen der Sensorantworten (siehe Abbildung 2.18).

Im gezeigten Beispiel kann Sensortyp A die zwei Stoffe *Substanz 1* und *Substanz 2* detektieren. Sensortyp B kann ebenfalls auf *Substanz 2* und einen weiteren Stoff eine Reaktion zeigen. Zeigen nun beide Sensoren gleichzeitig eine Antwort auf eine unbekannte Probe, so handelt es sich wahrscheinlich um *Substanz 2*.

## 2. Grundlagen

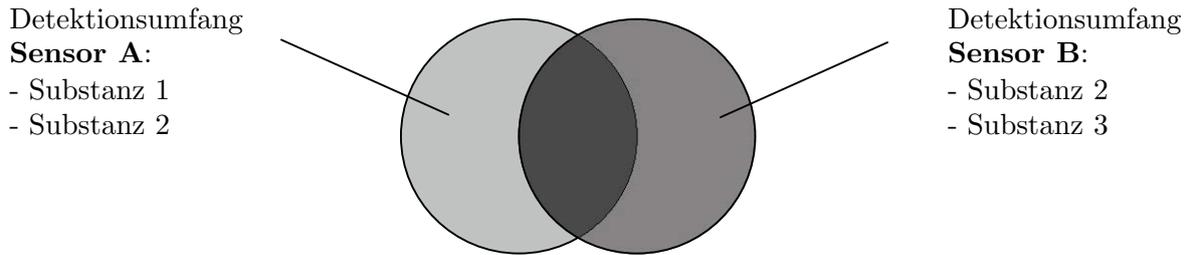


Abbildung 2.18.: Die Kombination mehrerer breitbandiger Sensoren erlaubt eine Verbesserung der Selektivität

Werden mehrere Sensorsignale für eine Vielzahl von zu detektierenden Substanzen eingesetzt, handelt es sich um ein multivariates Problem [52]. Die Methoden zur Lösung der Klassifikationsaufgabe werden als Mustererkennung zusammengefasst. Die Eingangsgrößen einer Mustererkennung müssen nicht nur, wie im bebilderten Beispiel (Abb. 2.18) der Mengenanschauung, ordinalskalierte Aussagen eines Detektors (schlägt an oder nicht an) enthalten, sondern können auch die quantitativen Informationen eines Sensors nutzen. Ebenso können im Zusammenhang mit dem in dieser Arbeit betrachteten Sensorsystem, weitere Informationen wie die Ionenmobilitäten und Retentionszeiten für Klassifikation berücksichtigt werden. Die in einer Mustererkennungsmethode verwendeten Informationen werden als *Merkmale* einer Messung bezeichnet.

Die aufeinanderfolgenden Schritte einer Mustererkennung sind in Abbildung 2.19 aufgeführt. Aus den rohen Signalen der Sensoren werden zunächst die zu betrachtenden Merkmale extrahiert. Im einfachen Fall kann dies zum Beispiel die Höhe des Signals zu einem bestimmten Zeitpunkt sein.

Die betrachteten Merkmale spannen einen Raum auf, welcher nicht zwangsläufig ebenso viele Dimensionen besitzt, wie es Signalquellen  $q$  gibt. Sind Sensoren beispielsweise in ihrem Antwortverhalten sehr ähnlich, so kann diese Redundanz zu einer Dimensionsreduktion führen. In dem Signalverlauf eines Sensors können neben der Signalhöhe ggf. noch mehr Merkmale zur Stoffunterscheidung enthalten sein. Dies können Informationen sein, wie An- oder Abstiegszeiten.

Eine umfangreiche Liste von extrahierbaren Parametern eines Halbleitersensordesigns ist bei Eklöv et. al [53] zu finden. Die Dimensionalität des Merkmalsvektors ist

## 2. Grundlagen

hierdurch häufig höher als die Anzahl von Sensorsignalen. Die Merkmalsextraktion ist damit mathematisch eine Abbildung  $q \rightarrow p$  der Signale in den Merkmalsraum. Die Zusammenstellung der Merkmale eines Messpunktes wird im weiteren als Merkmalsvektor  $\vec{v}$  bezeichnet.

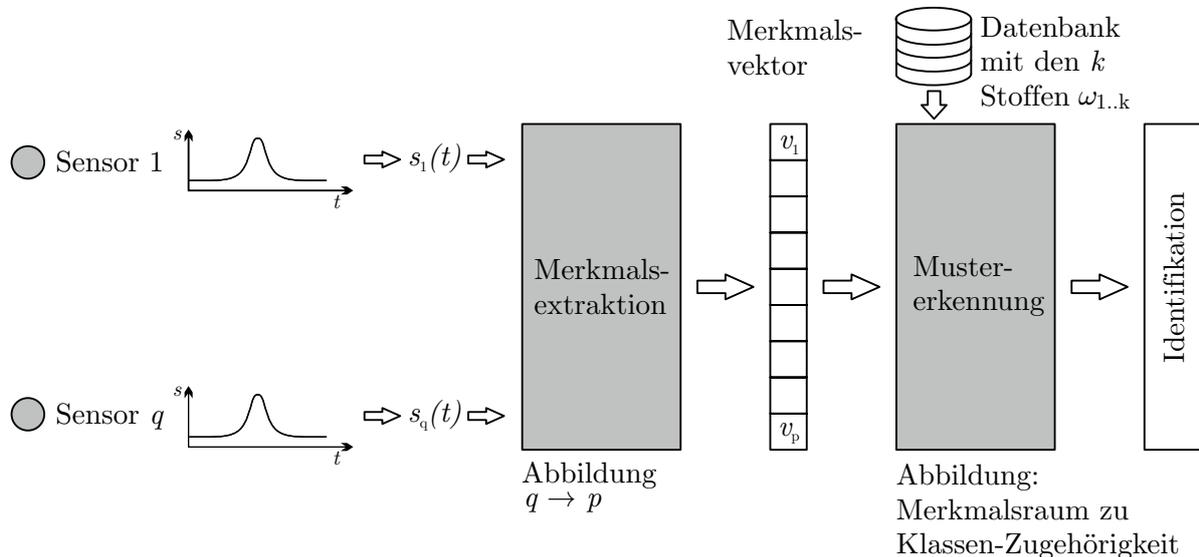


Abbildung 2.19.: Klassifikation von Sensorsignalen [54]

Das sich einstellende Muster, also der Wert des Merkmalsvektors wird für eine Substanzidentifikation mit den zuvor gelernten Informationen einer Datenbank verglichen und dann einer Klasse mit dem Bezeichner  $\omega$  zugeordnet. Da Messungen streuen und jedes Element des Merkmalsvektors ein individuelles Vertrauensintervall besitzt, sind Gültigkeitsbereiche der Klassen zu definieren. Dies kann durch a-priori-Wissen oder durch maschinelles Lernen geschehen. Veranschaulicht werden kann dieser Sachverhalt durch Streudiagramme, wie in Abbildung 2.20 a) dargestellt. Hier ist beispielhaft das Diagramm eines zweidimensionalen Merkmalsvektors  $\vec{v}$  gezeigt. Die durch die Elemente  $v_1$  und  $v_2$  des Merkmalsvektors bestimmten Messungen werden im  $p = 2$  dimensional kartesischen Koordinatensystem lokalisiert.

Zum Anlernen eines Klassifikators wird ein Trainingsdatensatz  $\mathcal{L}$  mit Messungen benötigt, deren Klassenzuordnung bekannt ist. In der Abbildung des exemplarischen Problems ist die Zuordnung durch Quadrate sowie Dreiecke veranschaulicht. Die verschiede-

## 2. Grundlagen

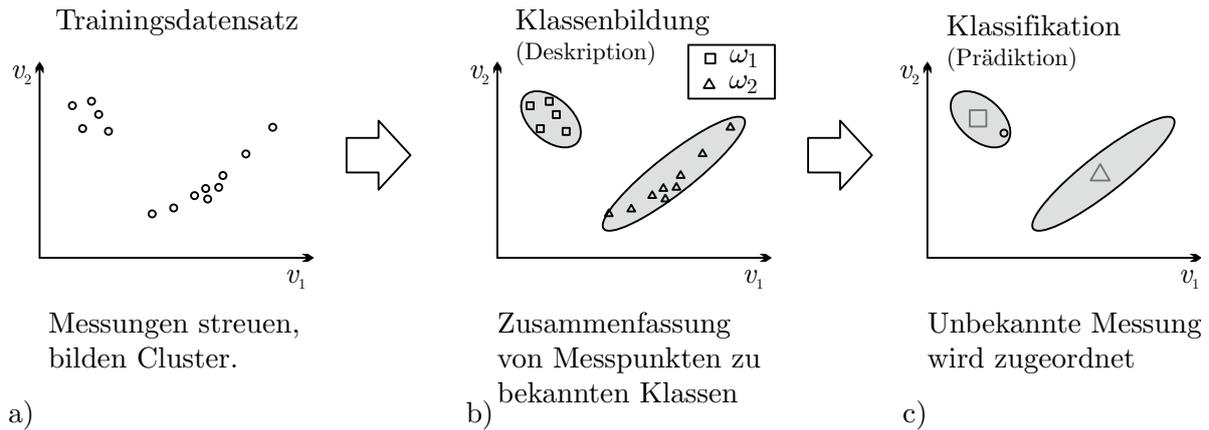


Abbildung 2.20.: Streudiagramm eines zweidimensionalen Merkmalsvektors

nen Verfahren der Mustererkennung unterscheiden sich maßgeblich durch die Methoden, wie die Klassengrenzen definiert werden. Die Klassifikation schließlich ordnet eine Messung mit unbekannter Herkunftsclass einer möglichen Klasse zu. Im Rahmen dieser Arbeit entspricht die Klasse einer Substanzzuordnung. Der in der Teilabbildung c) gezeigte Messwert eines unbekanntes Stoffes würde in diesem Fall der Klasse „Quadrat“ zugeordnet werden.

### 2.4.1. Gestaltung von Klassengrenzen

Aus einem Trainingsdatensatz werden Regeln abgeleitet, welche die Klassengrenzen innerhalb des Merkmalsraums beschreiben. Ein Algorithmus zum Erzeugen eines deskriptiven Modells aus den Trainingsdaten kann dies sehr genau machen. Eine Reklassifizierung der Trainingsdaten ordnet dann jedem Messpunkt auch genau der Ausgangsklasse wieder zu. Ein solcher genauer Lern-Algorithmus berücksichtigt dabei jedoch auch jeden Ausreißer der Messdaten. Die Grenze des Klassifikationsbereichs kann durch diesen Umstand eine komplexe Form annehmen. Ein Deskriptionsalgorithmus kann auf der anderen Seite auch dahingehend parametrisiert sein, möglichst einfache Klassengrenzen zu erzeugen und ggf. Ausreißer in Trainingsdaten zu ignorieren. Durch eine zu starke Vereinfachung von Klassifikationsbereichen kann es dann bei Testdaten auch zur Fehlklassifikation kommen.

**Überanpassung und Unteranpassung** Eine zu präzise Abbildung der Klassengrenze auf die Trainingsdaten wird auch als *Überanpassung* (engl. Overfitting) und eine zu lockere Abbildung wird als *Unteranpassung* (engl. Underfitting) bezeichnet. Die Kunst bei der Auswahl des Verfahrens und dem Training einer Musterdatenbank besteht darin, möglichst Über- und Unteranpassung eines Klassifikators zu vermeiden.

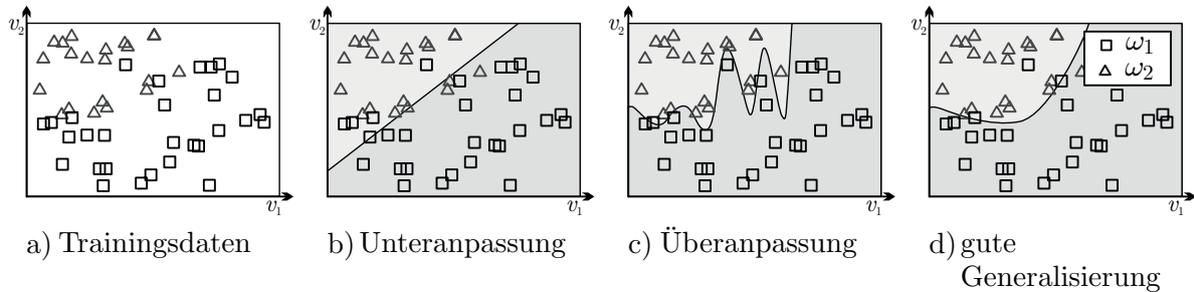


Abbildung 2.21.: Optimierung einer Klassengrenze. Ein Klassifikationsalgorithmus sollte eine gute Generalisierungsfähigkeit besitzen [55]

In den Streudiagrammen aus Abbildung 2.21 sind die beiden Anpassungsfehler zweier Klassen dargestellt. Im Streubild in Abbildung 2.21 a) ist ersichtlich, dass sich anhand der Trainingsdaten nur schwerlich disjunkte Bereiche im Merkmalsraum ausmachen lassen. Es sind Ausreißer beider Klassen erkennbar.

In 2.21 b) wird der Merkmalsraum in zwei Klassifikationsbereiche durch eine Gerade aufgeteilt. Der hellgraue Bereich oben links beschreibt die Klassifizierungsregion der Klasse  $\omega_2$ . Im dunkleren Bereich des Raums werden Messungen der Klasse  $\omega_1$  zugeordnet. Es ist zu erkennen, dass aufgrund der Simplifizierung der Klassengrenze mehrere Fehlklassifikationen auftreten.

In dem in 2.21 c) gezeigten überangepassten Klassifikationsintervall werden zwar alle Messungen des Trainingsdatensatzes korrekt reklassifiziert, neue Testdaten werden allerdings wegen der sprunghaften Grenzziehung oft falsch zugeordnet.

Von einem Mustererkennungsverfahren wird erwartet weder über- noch unterangepasst an die Trainingsdaten zu sein. Die eingesetzte Methode soll aus den gegebenen Trainingsdaten möglichst gut *generalisieren* können. In Teilabbildung 2.21 d) ist eine mögliche Klassifikationsgrenze gezeigt die einen guten Kompromiss darstellt.

## 2. Grundlagen

**Diskriminanzfunktion** Ein Modell zur Festlegung einer Klassengrenze ist die Diskriminanzanalyse unter Berücksichtigung des *Maximum-Likelihood* (maximale Wahrscheinlichkeit). Eine mehrfache Erhebung von Messdaten einer Klasse, also durch wiederholtes Messen einer Substanz, erhält man eine statistische Verteilung der aufgenommenen Daten.

Bezugnehmend auf das Streudiagramm aus Abbildung 2.20 wird in folgender Abbildung 2.22 das Merkmal  $v_1$  einzeln betrachtet. Teilbild a) zeigt die Projektion der Mess-

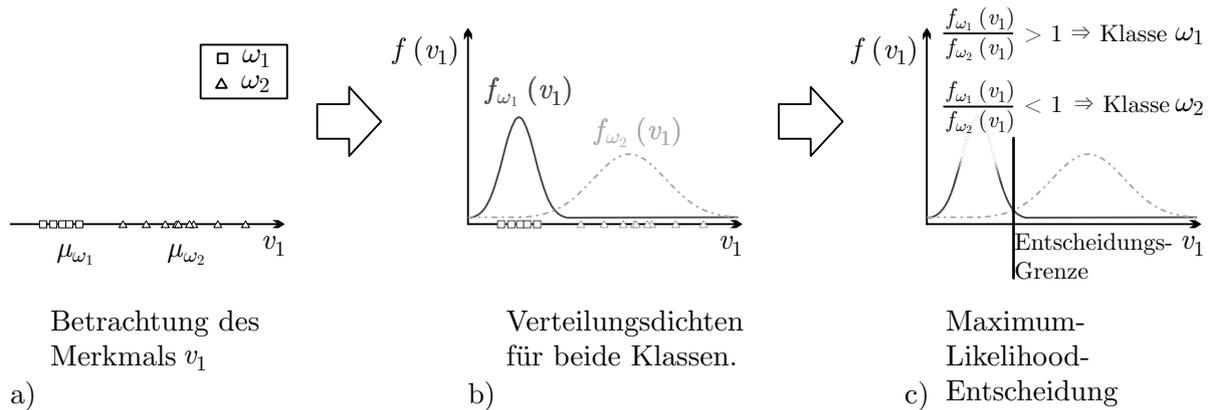


Abbildung 2.22.: Festlegen einer Entscheidungsgrenze nach der Verteilungsdichte von Messwerten [54]

werte auf die  $v_1$ -Achse mit den vorher bekannten Klassen *Quadrat* und *Dreieck*. Aus der Lage der Messpunkte einer Klasse kann eine Verteilungsdichtefunktion  $f(v_1)$  abgeleitet werden. In b) sind die geschätzten Dichtefunktionen beider Klassen des Beispiels dargestellt. Sie werden auch als a posteriori Wahrscheinlichkeit bezeichnet [56]. Durch eine ausreichende Fallzahl ergeben sich Normalverteilungen. Die Wiederholungsmessung kann durch die statistischen Parameter Mittelwert  $\mu$  und Varianz  $\sigma^2$  beschrieben werden (siehe Gleichung 2.36).

$$f(v) = \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-\frac{(v-\mu)^2}{2\sigma^2}} \quad (2.36)$$

Man erkennt in dem gezeigten Beispiel eine breitere Basis der Verteilungsdichtefunktion der Klasse  $\omega_1$  gegenüber  $\omega_2$  und damit ein größeres  $\sigma_{\omega_1}^2$ . Beide Verteilungsdichtefunktionen überlagern sich in ihren Randbereichen. Durch die Unschärfe wird eine Aussage über die Klassenzugehörigkeit einzelner Messpunkte in diesen Bereichen erschwert. Die Maximum-Likelihood-Entscheidungsregel (siehe Gleichungen 2.37-2.39) erlaubt eine



## 2. Grundlagen

systematische Zuordnung von Messungen, auch bei sich überschneidenden Messwertdistributionsen [52].

$$f_{\omega_1}(v) > f_{\omega_2}(v) \Rightarrow \text{Messwert } v \text{ wird zugeordnet zu Klasse } \omega_1 \quad (2.37)$$

$$f_{\omega_1}(v) < f_{\omega_2}(v) \Rightarrow \text{Messwert } v \text{ wird zugeordnet zu Klasse } \omega_2 \quad (2.38)$$

$$f_{\omega_1}(v) = f_{\omega_2}(v) \Rightarrow \text{Messwert } v \text{ kann willkürlich zugeordnet werden} \quad (2.39)$$

In Teilabbildung 2.22 c) wurde diese Entscheidungsregel angewendet. Alle Messungen mit einer Merkmalsausprägung  $v_1$  rechts der Entscheidungsgrenze werden der Klasse  $\omega_1$  zugeordnet, alle anderen der Klasse  $\omega_2$ .

**Falsch positiv und falsch negativ** In den Grenzbereichen kann es zu Zuordnungsfehlern kommen. Ein solcher Fehler kann je nach Auswirkung in „falsch positiv“ und „falsch negativ“ unterschieden werden.

Ein *falsch negatives* Ergebnis tritt ein, wenn die Messung nicht der Klasse zugeordnet wird, aus der die Probe eigentlich stammt. Im Falle der hier betrachteten Gasanalyse trifft dies zu, wenn ein schädliches Gas in der Luft vorhanden ist, es aber von der Klassifikation nicht als solches erkannt wird. Dieser Fall sollte nicht auftreten. Der gegensätzliche Fall, also eine *falsch positive* Zuordnung ist der klassische Fehlalarm. Eine Messung wird einer Klasse zugeordnet, aus der sie aber in Wirklichkeit nicht stammt.

Da ein falsch negatives Ergebnis im Falle der Schadstoffdetektion von schlimmeren Folgen geprägt ist als ein Fehlalarm, ist eine Verschiebung der Klassengrenze in diesem Fall erwünscht. Die Schwere einer Falschzuordnung kann über eine Kostenfunktion abgebildet werden, welche die Entscheidungsgrenze verschiebt. Formal wird sie beschrieben als  $L(\omega_j|\omega_i)$  für Kosten, die entstehen, wenn eine Messung, die zur Klasse  $\omega_i$  gehört, der falschen Klasse  $\omega_j$  zugeordnet wird.  $L(\omega_j|\omega_i)$  kann Werte zwischen 0 und 1 annehmen. Die Entscheidungsregel aus Gleichungen  $L(\omega_j|\omega_i)$  2.37-2.39 kann nun mit den Kosten verknüpft werden und man erhält:

$$L(\omega_2|\omega_1)f_{\omega_1}(v) > L(\omega_1|\omega_2)f_{\omega_2}(v) \Rightarrow v \text{ wird zugeordnet zu Klasse } \omega_1 \quad (2.40)$$

$$L(\omega_2|\omega_1)f_{\omega_1}(v) < L(\omega_1|\omega_2)f_{\omega_2}(v) \Rightarrow v \text{ wird zugeordnet zu Klasse } \omega_2 \quad (2.41)$$

$$L(\omega_2|\omega_1)f_{\omega_1}(v) = L(\omega_1|\omega_2)f_{\omega_2}(v) \Rightarrow v \text{ kann willkürlich zugeordnet werden} \quad (2.42)$$

## 2. Grundlagen

Neben der Kostenfunktion kann noch eine *a priori Wahrscheinlichkeit* eingeführt werden. Diese schätzt ab, wie häufig eine Klasse insgesamt bei allen Messungen vorkommen wird. Bei der Messung von Spurengasen kann die a priori Wahrscheinlichkeit zum Beispiel auf Statistiken und Erfahrungen über die Häufigkeit transportierter Gefahrgüter basieren. Es ist beispielsweise wesentlich wahrscheinlicher, dass saubere Luft gemessen wird, als ein seltener Gefahrstoff. Die a priori Wahrscheinlichkeit mit dem Formelzeichen  $\pi_{\omega_i}$  kann als weiterer Faktor in den Gleichungen  $L(\omega_j|\omega_i)$  2.37-2.39 berücksichtigt werden.

$$\pi_{\omega_1} L(\omega_2|\omega_1) f_{\omega_1}(v) > \pi_{\omega_2} L(\omega_1|\omega_2) f_{\omega_2}(v) \Rightarrow v \text{ gehört zu Klasse } \omega_1 \quad (2.43)$$

$$\pi_{\omega_1} L(\omega_2|\omega_1) f_{\omega_1}(v) < \pi_{\omega_2} L(\omega_1|\omega_2) f_{\omega_2}(v) \Rightarrow v \text{ gehört zu Klasse } \omega_2 \quad (2.44)$$

$$\pi_{\omega_1} L(\omega_2|\omega_1) f_{\omega_1}(v) = \pi_{\omega_2} L(\omega_1|\omega_2) f_{\omega_2}(v) \Rightarrow v \text{ kann willkürlich zugeordnet werden} \quad (2.45)$$

Unter Berücksichtigung der a priori Wahrscheinlichkeit wird diese Regel zur Festlegung der Entscheidungsgrenze auch als Bayes-Entscheidungsregel bezeichnet.

In der bisherigen Betrachtung wurde nur von dem univariaten Fall ausgegangen, bei dem nur ein Merkmal, also in Bezug auf die Gassensorik nur ein Sensor berücksichtigt wird. Für eine multivariate Betrachtung mit einer Vielzahl von Eingangssignalen müssen die Gleichungen auf einen mehrdimensionalen Merkmalsvektor  $\vec{v}$  erweitert werden. Die Erweiterung der Normalverteilung (Gleichung 2.36) für den mehrdimensionalen Fall lautet dann:

$$f(\vec{v}) = \frac{1}{|\Sigma|^{\frac{1}{2}} \sqrt{(2\pi)^p}} \cdot e^{-\frac{(\vec{v}-\vec{\mu})' \Sigma^{-1} (\vec{v}-\vec{\mu})}{2}} \quad (2.46)$$

Hier beschreibt  $\Sigma$  die Kovarianzmatrix und  $|\Sigma|$  deren Determinante. Die Potenz der Exponentialfunktion beschreibt eine Distanz zwischen dem Merkmalsvektor  $\vec{v}$  und dem Erwartungswert für alle Messungen der Klasse ( $\vec{\mu}$ ).

### 2.4.2. Klassifikationsbäume

Klassifikationsbäume teilen ein multivariates Klassifizierungsproblem mit mehreren Merkmalen in kleinere Einzelprobleme, bei denen jeweils nur ein Merkmal betrachtet wird,

## 2. Grundlagen

auf. Der CART-Algorithmus (Classification And Regression Trees) nach Breiman et al. [57] legt methodisch die Klassifikationsgrenzen der Einzelprobleme anhand eines Trainingsdatensatzes fest. Neben der CART-Methode existieren noch weitere Entscheidungsbaummodelle zur Klassifikation von mehrdimensionalen Entscheidungsräumen. Verbreitete Methoden sind *ID3* und deren Erweiterung *C4.5*. Im Gegensatz zum nachfolgend vorgestellten CART-Algorithmus handelt es sich bei diesen Methoden nicht um reine Binärbaum-Implementationen. Es wurden bereits Untersuchungen durchgeführt, bei denen unterschiedliche Verfahren des maschinellen Lernens für die Dateninterpretation von Elektronischen Nasen verglichen wurden. Die CART Methode hat sich dabei mit Blick auf Klassifikationsgenauigkeit als sehr geeignet herausgestellt [58].

In der Abbildung 2.23 ist beispielhaft die Vorgehensweise zur Aufteilung eines zweidimensionalen Merkmalsraums in binäre Entscheidungen beschrieben. Wie im linken

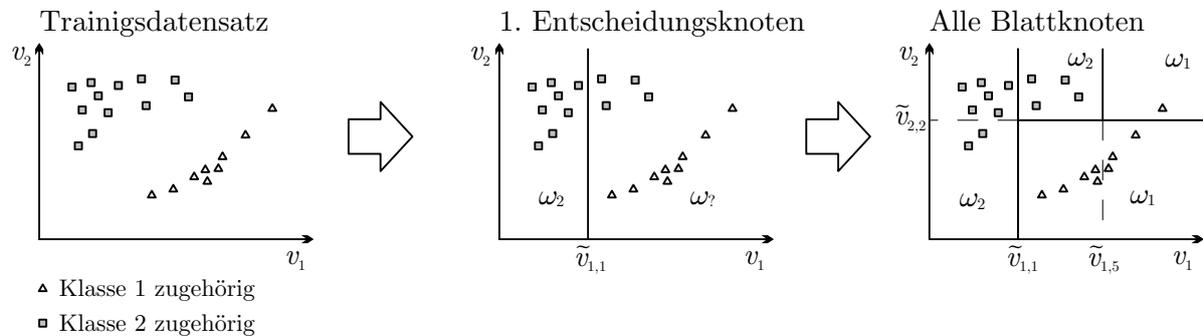


Abbildung 2.23.: Klasseneinteilung eines zweidimensionalen Merkmalsraums über einen Entscheidungsbaum

Teilbild zu erkennen, streuen die Messwerte des Trainingssets über beide Merkmalsdimensionen. Es lassen sich jedoch die zwei Klassen  $\omega_1$  und  $\omega_2$  optisch bereits sehr gut trennen.

Für einen ersten Entscheidungsknoten mit der Bezeichnung  $m_1$  wird der Merkmalsraum unter Betrachtung von Merkmal  $v_1$  in zwei Bereiche unterteilt (mittleres Teilbild). Die gefundene Klassifikationsgrenze dieses Knotens wird als  $\tilde{v}_{1,1}$  markiert. Alle Messwerte links dieser Grenze können in diesem Beispiel bereits der Klasse  $\omega_2$  zugeordnet werden. Für eine Zuordnung der rechtsseitigen Messwerte muss der Merkmalsraum noch feiner unterteilt werden.

## 2. Grundlagen

Im rechten Teilbild ist der gesamte aufgespannte Merkmalsraum in rechteckförmige Klassifikationsbereiche aufgeteilt worden (Hyperrechtecke). Es ergeben sich die drei Grenzen  $\tilde{v}_{1,1}$ ,  $\tilde{v}_{2,2}$  und  $\tilde{v}_{1,5}$ .

Bei den Klassifikationsbäumen nach CART handelt es sich um reine Binärbäume. Dies bedeutet, jeder Knoten des Baumes darf genau zwei Nachfolger besitzen. Der Entscheidungsbaum wird bei der Klassifizierung vom Wurzelknoten bis zum Ergebnisblatt durchlaufen. An jedem Knoten wird dabei eine binäre *True*- und *False*-Entscheidung getroffen. Das Ergebnis eines Entscheidungsknotens führt zu dem nächsten zu durchlaufenden Kindbaum. Die Kinder eines Entscheidungsknotens haben als Eingangsdatensatz die Messungen des jeweils übergeordneten Hyperrechtecks im Entscheidungsraum.

In Abbildung 2.24 ist die Baumstruktur der Klassifikationsbereiche des beschriebenen Merkmalsraums aus Abbildung 2.23 dargestellt. Der Ursprungsknoten führt den Ver-

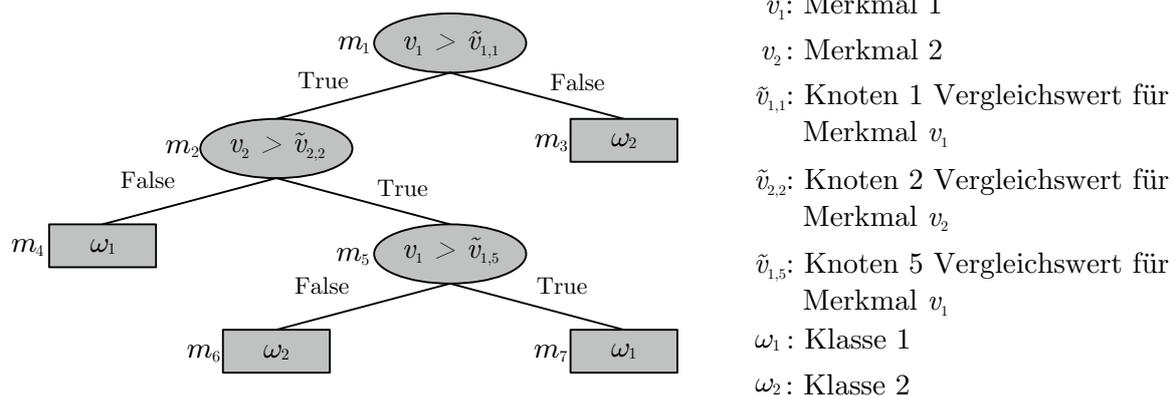


Abbildung 2.24.: Klassifikationsbaum

gleich des Merkmals  $v_1$  mit der Diskriminante  $\tilde{v}_{1,1}$  durch. Ist dieses Merkmal größer als die festgelegte Grenze, so wird der linksseitige Teilbaum weiter betrachtet. Ist dies nicht der Fall, so verweist der Knoten bereits auf einen Endknoten, auch als Blatt bezeichnet, welcher ein Klassifikationsergebnis der Messung  $\vec{v}$  mit der Klasse  $\omega_1$  markiert.

Ein entscheidender Vorteil von Entscheidungsbäumen gegenüber anderen Auswertungsalgorithmen ist der geringe Rechenaufwand bei der Klassifikation. Es müssen nur binäre

Entscheidungen durch *if-then-else*-Abfragen getroffen werden. Es sind nur maximal so viele Abfragen bis zum Klassifikationsergebnis erforderlich, wie durch die *Höhe* des Baumes bestimmt.

### Erzeugung von Klassifikationsbäumen

So ressourcenschonend die Klassifizierung mithilfe eines solchen Entscheidungsbaumes ist, so rechenaufwendig ist die Induktion eines Entscheidungsbaumes. Der hierfür verwendete CART-Algorithmus legt rekursiv die Diskriminanten des Baumes anhand des Entropiegehaltes eines zu betrachtenden Eingaberaumes fest. Für die Entropiebestimmung wird die Wahrscheinlichkeit benötigt, die widerspiegelt, ob ein Merkmalsvektor  $\vec{v}$  im Knoten  $m$  zur Klasse  $\omega_i$  gehört. Dieses Wahrscheinlichkeitsmaß  $p_m^i$  wird definiert als [59]:

$$\text{prob}(\omega_i|\vec{v}, m) = p_m^i = \frac{N_m^i}{N_m} \quad (2.47)$$

Der Wert  $N_m^i$  beschreibt die Anzahl der Merkmalsvektoren des Eingaberaums (am Knoten  $m$ ), die zur Klasse  $\omega_i$  gehören. Der Eingaberaum besteht dabei aus den noch verbleibenden Merkmalsvektoren, nachdem schon vorher durch Elternknoten der Datenumfang verkleinert wurde. Für den Wurzelknoten ist der Eingaberaum also noch der gesamte Trainingsdatensatz. Der Eingaberaum wird zu der Gesamtanzahl aller Merkmalsvektoren im Eingaberaum  $N_m$  in Relation gesetzt. Über das Maß  $p_m^i$  kann nun bestimmt werden, wie gut der jeweilige Knoten den Datensatz des Eingaberaums zu trennen vermag. Das beste Trennvermögen ist bei  $p_m^i = 1$  oder  $p_m^i = 0$  erreicht. Dies bedeutet, dass alle bzw. keinerlei Messwerte  $\vec{v}$  des Eingangsraums am Knoten  $m$  der Klasse  $\omega_i$  zugeschlagen werden.

Zur Festlegung der Diskriminante wird mit Hilfe des berechneten Wahrscheinlichkeitsmaßes  $p_m^i$  die Entropie für diesen Knoten nach Gleichung 2.48 berechnet:

$$I_m = - \sum_{i=1}^k p_m^i \log_2(p_m^i) \quad (2.48)$$

Die Unreinheit (*impurity*)  $I_m$  ist also eine Kennzahl für den zu untersuchenden Knoten  $m$ , die besagt wie gut eine gerade ausgewählte Trennfunktion den Eingaberaum unterteilt. Neben der Berechnung der Entropie nach Gleichung 2.48 als Maß für die „impurity“

## 2. Grundlagen

eines Knotens wird in einigen Implementationen des Baum-Induktionsalgorithmus auch der Gini-Index verwendet:  $I_{m,gini} = \sum_{i \neq j} p_i p_j = 1 - \sum_{j=1}^J p_j^2$ .

Die Trennfunktion hat dabei zwei Freiheitsgrade: Die Auswahl des zu nutzenden Merkmals des multivariaten Raums ( $v_p, p = ?$ ) und die Lage der Trenngrenze  $\tilde{v}_{p,m} = ?$ . Bei der Aufgabe die richtige Trenngrenze  $\tilde{v}_{p,m}$  zu finden, handelt es sich also um ein Optimierungsproblem. Es gilt die Entropie für einen Knoten durch geschicktes Festlegen der Diskriminante so gering wie möglich zu halten:

$$I_{m,opt} = \min(I_m) \quad (2.49)$$

Der Algorithmus testet verschiedene Diskriminanten und berechnet jeweils die verbleibende Entropie  $I'_m$ . In der Gleichung 2.50 wird durch  $j$  die Nummer der getesteten Merkmal- und der Trenngrenzenkombination als Laufindex angegeben:

$$I'_m = - \sum_{j=1}^n \frac{N_{mj}}{N_m} \sum_{i=1}^k p_{mj}^i \log_2(p_{mj}^i) \quad (2.50)$$

Es werden nur solche Kandidaten für mögliche Trenngrenzen der Entropieberechnung unterworfen, welche zunächst die notwendige Bedingung erfüllen, zwischen zwei Messpunkten unterschiedlicher Klassen zu liegen.

Das Abbruchkriterium zur Erzeugung des Baumes kann sein, dass alle Knoten vollständig rein sind und keine Entropie mehr aufweisen müssen (alle  $p_m^i = 1$  oder  $p_m^i = 0$ ). Dies kann jedoch zu sehr komplexen und überangepassten Baumstrukturen führen. Daher wird häufig eine tolerierbare Unreinheit oder auch der Komplexitätsparameter  $\theta_p$  als Schwellwert für die a posteriori Wahrscheinlichkeit am Knoten definiert. [59]

**Eigenschaften von Entscheidungsbäumen nach CART** Ein in dieser Form induzierter Binärbaum mit der Merkmalsraumteilung in einzelne Hyperrechte wie in Abbildung 2.23 kann bei bestimmten Streubildern der Messungen von Nachteil sein. Durch die zu den Merkmalsachsen  $v_p$  orthogonal verlaufenden Klassengrenzen  $\tilde{v}_p$  kann es zu Treppenbildung bei Messwerten einer Klasse mit diagonalen Hauptachsen kommen. Solche Merkmalsausprägungen kommen vor allem dann vor, wenn zwei oder mehr Elemente des Merkmalsvektors korrelieren. Am Beispiel der Gassensorik ist dies der Fall bei Sensoren mit ähnlichem Antwortverhalten auf bestimmte Stoffe.

## 2. Grundlagen

Sind mit der beschriebenen Optimierungsstrategie ( $\min(I_m)$ ) alle Entscheidungsgrenzen festgelegt worden, so bildet der für die Trainingsdaten optimale Baum das gesamte Trainingsset in die a priori bekannten Klassen ab. Ein solcher Entscheidungsbaum neigt dabei zur Überanpassung an die Eingangsdaten. Ein Beispiel für die Tendenz zur Überanpassung ist dieses Verhalten: Liegt ein einzelner Merkmalsvektor des Trainingssets der Klasse  $\omega_1$  innerhalb eines Cluster von Vektoren einer anderen Klasse  $\omega_2$ , so würde der CART-Algorithmus diesen Ausreißer in den Entscheidungsbaum aufnehmen. Der Baum wäre damit nur für die Trainingsdaten sehr gut angepasst, würde aber neue Testvektoren fehlklassifizieren.

Die Streubilder von Klassen im Merkmalsraum von realen Messdaten erfüllen im Sinne der Mengenlehre die Bedingung der Kontrahierbarkeit, dies bedeutet, dass deren Gültigkeitsbereiche frei von Löchern sind. Bei derartigen Ausreißern würde es sich also offenbar um Fehler der Eingangsdaten handeln, die ein gut parametrierter Algorithmus der Bauminduktion ignorieren sollte. Für eine weitere Bewertung des Klassifikationsergebnisses des Baumes ist es sinnvoll, die Wahrscheinlichkeiten der Korrektuzuordnung ( $p_{m_j}^i$ ) der Blätter zu speichern [59]. Diese kann für eine finale Zuordnung eines Messwertes bei einem Klassifikationslauf mit berücksichtigt werden.

Eine bessere Generalisierung kann auch erreicht werden durch Einführung eines Abbruchkriteriums während der Bauminduktion. Dieses *Prepruning* - also ein Beschneiden des Baumes schon beim Wachstum, kann zum Beispiel anhand der Anzahl der Merkmalsvektoren, die zu einer neuen Verzweigung führen würden, geschehen. Eine kleine Datenbasis ist mit einer großen Ergebnisunsicherheit verbunden.

Ein weiterer Ansatz zur Erhöhung der Generalisierungsfähigkeit eines Baumes ist das *Postpruning*. Hier wird zunächst der Baum vollständig induziert. Danach werden die Verzweigungen und Blätter dahingehend überprüft wie groß ihr jeweiliger Anteil für Klassifikation ist. In der praktischen Umsetzung wird der Trainingsdatensatz hierzu in zwei Teilmengen  $a$  und  $b$  aufgeteilt. Mit der Teilmenge  $a$  wird ein Baum induziert, welcher keine Unreinheit mehr besitzt, dessen Blattknoten also alle eine a posteriori Wahrscheinlichkeit von Eins ( $p_m^i = 1$ ) besitzen. Jeder Entscheidungsknoten erhält nach der Induktion ein Klassenlabel, welches dem Ergebnis der Entscheidungen seines Kind-Baumes entspricht. Mit der Teilmenge  $b$  der Trainingsdaten wird dann überprüft, ob eine Richtig-Klassifikation schon an dem Entscheidungsknoten eintritt. Ist das Er-

gebnis besser, als das der verknüpften Unterbäume, so kann auf die Komplexität der weiteren Entscheidungen verzichtet werden. Der Baum wird somit beschnitten und generalisiert dadurch besser.

### 2.4.3. Random Forest

Die Trenngrenzen von CART-Entscheidungsbäumen verlaufen als scharfe Diskriminanten zwischen den Klassen. Durch ihre Orthogonalität zu den Merkmalsachsen bilden sie die Klassifizierungsbereiche oft nur unzureichend ab. Wie schon im vorherigen Abschnitt beschrieben, fehlt es einem einzelnen Entscheidungsbaum auch an Robustheit gegenüber Ausreißern innerhalb der Trainings-Messdaten. Eine Kombination von mehreren Klassifikatoren zu einem „*multiple classifier system*“ verspricht nach Tim Kam Ho et. al [60] eine starke Verbesserung des Klassifikationsergebnisses. Die Methode *Random Forest* erlaubt eine optimierte Abbildung der Klassifikationsbereiche und besitzt eine gesteigerte Robustheit bei verrauschten Eingangsmessdaten [61][62]. Für diese Methode werden mehrere unbeschnittene Entscheidungsbäume generiert, welche dann für die Klassifikation von Testdaten alle durchlaufen werden müssen. Die individuellen Ergebnisse der Entscheidungsbäume werden dann zusammenfassend bewertet und hieraus ein Gesamtergebnis abhängig von der Ergebnismajorität aller Klassifikatoren gebildet [63].

**Aufteilung des Trainingsdatensatzes** Beim Verfahren des *Bootstrap Aggregating* wird der Datensatz, mit dem der Klassifikator angelehrt wird ( $\mathcal{L}$ ), in Untermengen ( $\mathcal{L}^{(B)}$ ) unterteilt. Die Auswahl der Messungen für die Teilmengen der Trainingsdaten wird zufällig getroffen. Diese Zufallsauswahl geschieht im Sinne der Wahrscheinlichkeitstheorie mit Zurücklegen der Stichprobe. Der Stichprobenumfang kann je nach verwendeter Methode ein fester Wert (beispielsweise 50 Prozent der Trainingsdaten) oder als zufällige variable Anzahl definiert werden.

In der Abbildung 2.25 ist in Anlehnung an das bekannte Beispiel aus Kapitel 2.4.2 eine Aufteilung der Trainingsdaten gezeigt. Aus dem Gesamtumfang der Lerndaten werden drei Untermengen  $\mathcal{L}^{(1)}$ ,  $\mathcal{L}^{(2)}$ ,  $\mathcal{L}^{(3)}$  zufällig gewählt. Aus den 22 Messungen sind Stichproben mit unterschiedlicher Anzahl hervorgegangen. In den drei Teilbildern ist zu erkennen,

## 2. Grundlagen

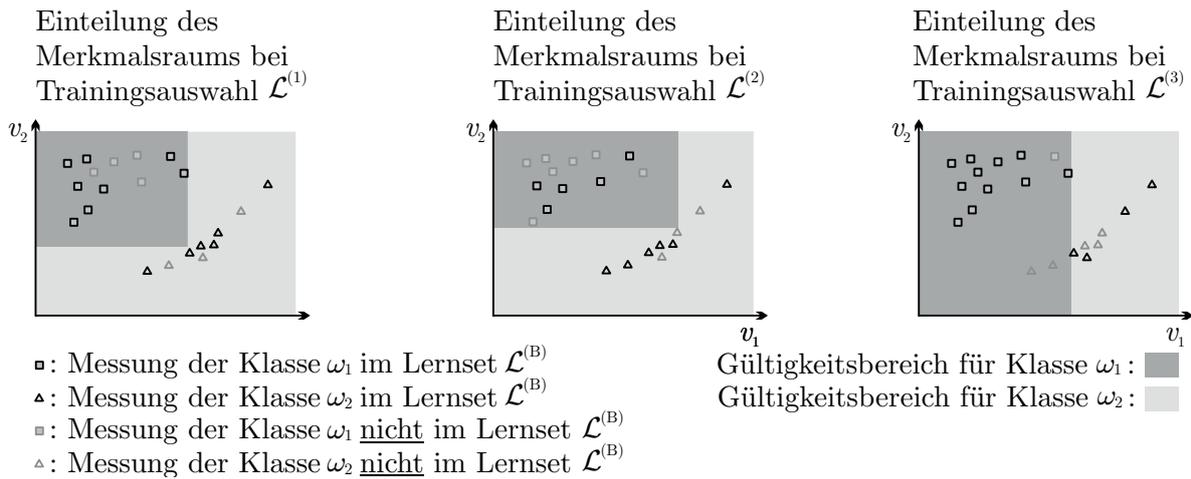


Abbildung 2.25.: Aufteilung des Merkmalsraums mithilfe der Random Forest Methode

dass das *Bagging*-Verfahren einige Messungen in alle Trainingsstichproben aufgenommen hat. Dies ist auf die Stichprobenwahl *mit Zurücklegen* zurückzuführen. Aus den jeweiligen Stichproben sind Entscheidungsgrenzen nach dem CART-Verfahren festgelegt worden. Die Gültigkeitsbereiche sind mit unterschiedlichen Grautönen veranschaulicht. Messungen, deren Merkmale im dunkelgrauen Bereich liegen werden der Klasse  $\omega_1$  zugeordnet und Messungen mit Merkmalsvektoren im hellgrauen Bereich gehören zu Klasse  $\omega_2$ . Wie zu erkennen ist, verursacht beispielsweise der Klassifikator, welcher aus dem Lernset  $\mathcal{L}^{(3)}$  gebildet wurde, bereits Fehlklassifizierungen für die Gesamtheit der Trainingsdaten: Die beiden untersten Messwerte mit dem kleinsten Merkmal  $v_2$  würden Klasse  $\omega_1$  zugeschlagen werden obwohl sie aus  $\omega_2$  stammen. Für die Gesamtbetrachtung ist dies kein Problem, da die Ergebnismajorität für das Klassifikationsergebnis bestimmend ist.

**Klassifikation mit Random Forest** In Abbildung 2.26 sind die Gültigkeitsbereiche der drei Entscheidungsbäume überlagert dargestellt. Das linke Teilbild fasst die Bäume aus  $\mathcal{L}^{(1)}$  und  $\mathcal{L}^{(3)}$  zusammen. Für die beiden Fehlklassifikations-Kandidaten besteht nun noch eine fünfzigprozentige Chance falsch zugeordnet zu werden. Ein Baum ordnet sie der Klasse  $\omega_1$  zu, der andere  $\omega_2$ . Zusammen mit dem Entscheidungsbaum aus Lernset  $\mathcal{L}^{(2)}$  fällt die Entscheidung klar auf Klasse  $\omega_2$ . Im rechten Teilbild ist die Klasseneinteilung nach dem CART-Verfahren auf Basis des gesamten Lernsets abgebildet. Man erkennt eine etwas andere Struktur. Im direkten Vergleich zeigt sich der Vorteil der Kombination

## 2. Grundlagen

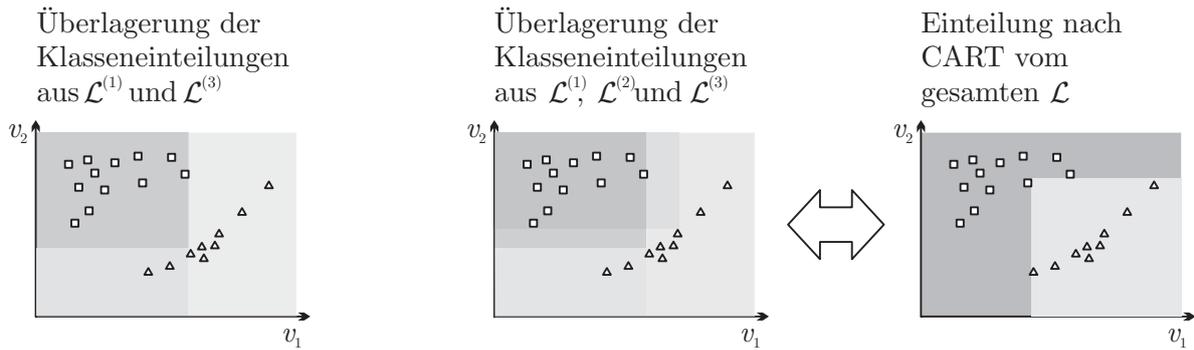


Abbildung 2.26.: Vergleich von Klasseneinteilung nach Random Forest und CART

mehrerer schwacher Lerner gegenüber dem Baum basierend auf dem gesamten Trainingsdatensatz. Gerade für verrauschte Daten mit einer geringen Präzision können in den Randbereichen die Messunsicherheiten mit berücksichtigt werden. Ensemble Methoden wie Random Forest können so dem Klassifikationsergebnis auf einfache Weise einen Wert für die Unsicherheit mitgeben. Es sind damit zum Beispiel auch Verschiebungen der Entscheidungsgrenzen möglich, um hiermit die Zuordnungsfehler in Richtung „falsch positiv“ oder „falsch negativ“ zu steuern. Gerade in der Gasmesstechnik als Schutztechnologie sind Fehlalarme (falsch positiv) eher zu tolerieren als die Nicht-Detektion eines Gefahrstoffs.

## 3. Aufbau von mobilen GC-Multisensor-Systemen

In dieser Arbeit werden die Ergebnisse zweier entwickelter Messsysteme dargelegt. Zunächst wird ein GC-PID-Messgerät beschrieben. Anhand dieses, auf nur einem Detektor basierenden Systems, wird später in Kapitel 4 die Datenverarbeitung und Merkmalsextraktion von Gaschromatographie-Läufen erläutert. Danach wird das realisierte Multisensorsystem DACHS vorgestellt. In dem entwickelten Messsystem finden die im Grundlagenkapitel vorgestellten Mustererkennungsverfahren zur Klassifikation chemischer Substanzen Anwendung.

### 3.1. Realisierung eines GC-PID-Systems mit kurzer Multikapillare

Das GC-PID Messsystem dient als Testgerät für die Entwicklung von GC-Algorithmen. Es wird im Folgeprojekt als Handgerät für die Erfassung von organischen Schadstoffen weiterentwickelt. [64]

#### 3.1.1. Fluidik der analytischen Komponenten

Die Analytik des Messsystems basiert auf der Kopplung eines breitbandigen Photoionisationsdetektors mit einer kurzen Multikapillar-Chromatographiesäule. Für die Probengasinjektion kommt ein Gegenstrom-Injektor zur Anwendung. In der Abbildung 3.1 ist die Fluidik des Messgerätes dargestellt. Durch eine Drehschieber-Pumpe wird die



Probenkammer in welcher der Photoionisationsdetektor eingebracht ist, permanent be-  
saugt. Die sich einstellende Druckreduktion wird durch einen Absolutdrucksensor erfasst.  
Durch die Druckdifferenz wird die Gaschromatographie-Säule kontinuierlich durchspült.  
Im Ruhezustand stellt der Injektor-Aufbau genügend Spülluft zur Verfügung, sodass  
kein Probengas in den Einlass gelangt. Die Spülluft bzw. das Trägergas, wird aus der  
Umgebungsluft gewonnen. Hierfür fördert die Gegenstrom-Pumpe die Luft durch einen  
Aktivkohlefilter, welcher mögliche Verunreinigungen adsorbiert.

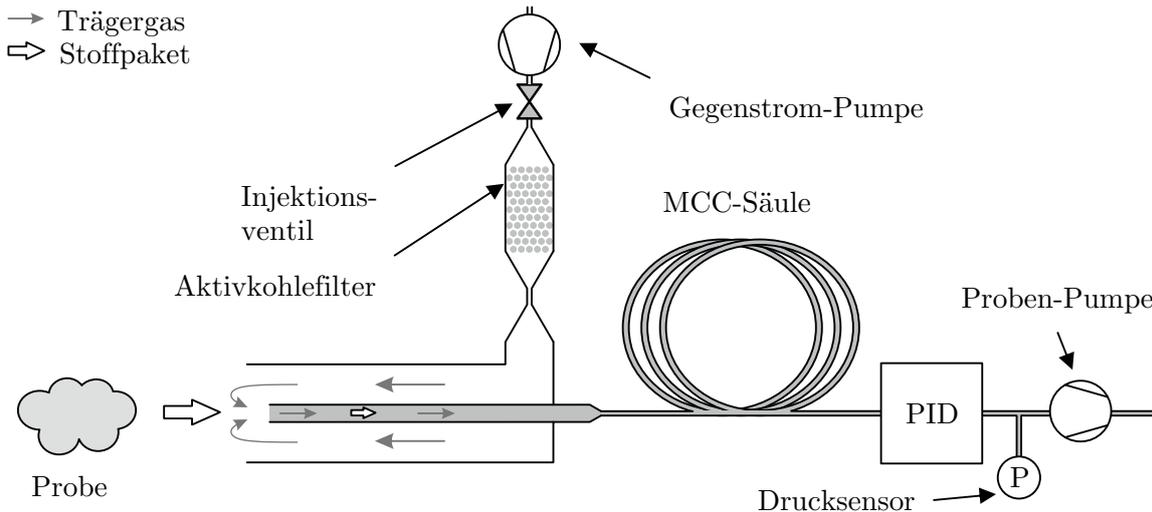


Abbildung 3.1.: Prinzipschaltbild des Kurzsäulen GC-PID [65]

Für eine Injektion eines Probengas-Paketes wird ein Injektionsventil kurzzeitig geschlos-  
sen. Die Probe gelangt in die Trennstrecke und wird dann, zeitlich aufgetrennt, vom  
Photoionisationsdetektor erfasst. Eine Auswerte-Software misst die Zeit zwischen Injek-  
tion (Injektionsventil schließt) und Peak-Maximum und gibt so Rückschlüsse auf die  
Zusammensetzung einer Gasprobe.

### GC-Trennsäule

Die gaschromatographische Trennsäule wird durch eine umlaufende elektrische Heiz-  
wendel auf eine definierte Temperatur beheizt. In den durchgeführten Versuchen zur  
Systemevaluation wurde die Temperatur konstant gehalten.



Durch den Einsatz einer kurzen analytischen Säule stellt sich für die Zielsubstanzen eine befriedigende Trennleistung ein. Gleichzeitig ist die Analysezeit kurz gehalten. Bei einer isothermen und isobaren Analyse werden Analysedauern von unter einer Minute erreicht [65]. Die kurze Messzeit ist insbesondere für den mobilen Feldeinsatz eine wichtige Bedingung. Die Trennstrecke wurde als Bündel mehrerer parallel geschalteter Einzelkapillaren ausgeführt. Hierdurch ist ein höherer Volumenstrom möglich, welcher zu einer besseren Durchspülung des Detektorvolumens führt [66]. Multikapillar-Säulen sind im Vergleich zu gepackten Säulen sowie Einzelkapillaren in ihrer Trennwirkung besser bei optimal eingestelltem Fluss [67].

Das in Abbildung 3.2 gezeigte Kapillarbündel visualisiert den Aufbau des Multikapillar-Gaschromatographen. Es handelt sich um bis zu 38 einzelne GC-Kapillaren mit einer Länge von 290 Millimetern. Die Glaskapillaren weisen jeweils einen Innendurchmesser von 0,25 Millimetern auf. Sie sind mit einer 1,4 Mikrometer dicken, mittelpolaren stationären Phase vom Typ SPB-624 beschichtet [68]. An den Enden des Bündels werden die

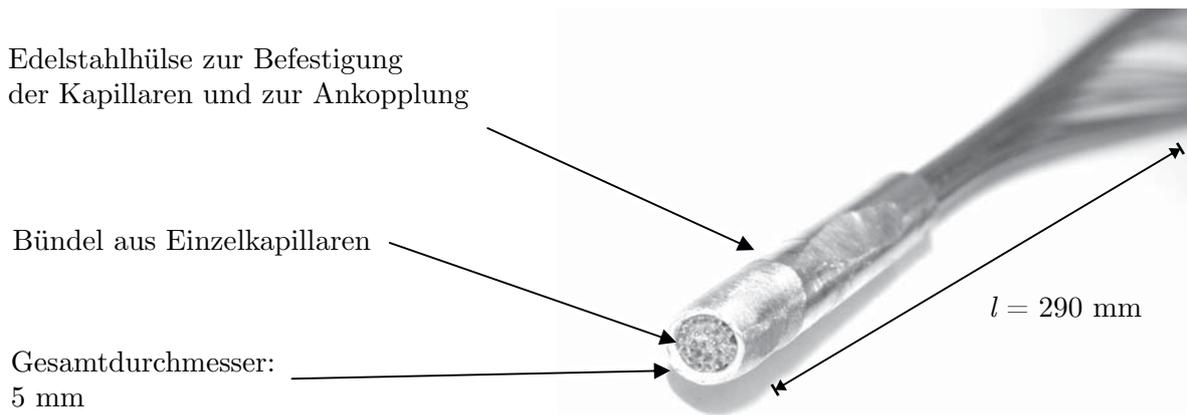


Abbildung 3.2.: Multikapillar-Gaschromatograph

Kapillaren durch eine Edelstahlhülse zusammengehalten. Die Hülse dient auch gleichzeitig zur fluidischen Ankopplung der Säule.

### Herstellung des Multikapillar-Bündels

Die Kapillaren werden in der Metallhülse durch einen auf Polyimidharz basierenden Klebstoff fixiert. Zur Montage werden die Kapillaren zunächst durch die Hülse geführt



### 3. Aufbau von mobilen GC-Multisensor-Systemen

und mit etwas Materialüberstand verklebt. Nach Aushärtung des Klebstoffs wird das Bündel abgelängt. Dabei wird kontinuierlich Luft durch die Kapillaren gefördert, so dass während des Schneidvorgangs kein Glasspan die Kapillaren verblocken kann.

#### 3.1.2. Detektor

Der eingesetzte Sensor ist ein Photoionisationsdetektor der Firma Alphasense mit einem Ionisierungspotential von 10,6 eV. Dieser Typ hat laut Hersteller eine untere Nachweisgrenze von 5 ppb bezogen auf das Prüfgas Isobuten in Stickstoff [69]. Dieses Gasgemisch besitzt für den eingesetzten Sensor einen Responsefaktor von eins [70]. Der technisch komplexe Aufbau eines Photoionisationsdetektors (siehe Kapitel 2.1.2) ist bei diesem Miniatursensor vollkommen integriert. So erfolgen sowohl die Lampenansteuerung als auch die Signalverstärkung innerhalb des 20 mm × 16,5 mm messenden Sensorgehäuses. Der innere Aufbau gleicht dem aus Abbildung 2.6 im Grundlagenkapitel. Durch den



Abbildung 3.3.: Verwendeter Photoionisationsdetektor Alphasense PID-AH

kompakten Aufbau ist dieser Sensor gut geeignet für eine Integration in handgehaltene Messinstrumente. Die Ankopplung an einen Gasstrom erfolgt durch eine mit einem Filterpapier versehene Diffusionsöffnung an der Oberseite des Sensors. Darunter befindet sich unmittelbar eine Anordnung aus Gitterelektroden. Die zur Ionisierung erforderliche ultraviolette Strahlung wird durch eine über ein elektrisches Feld angeregte Lampe unterhalb der Elektrodenanordnung erzeugt. Sowohl die Lampe als auch der Elektrodenstapel mit seinen Dichtflächen sind austauschbare Elemente des Sensors. Das Lampenfenster wird mithilfe einer Druckfeder gegen den Reaktionsraum gedichtet. Der gesamte Sensor wiederum wird in das Messinstrument über eine Stirndichtung aus Fluor-Kautschuk

### 3. Aufbau von mobilen GC-Multisensor-Systemen

an den analytischen Pfad hinter der GC-Trennsäule angeschlossen. Die Zu- und Abluftöffnungen sind so angeordnet, dass eine gute Durchspülung der Sensorkammer erreicht wird.

Elektrisch besitzt der Sensor auf der Unterseite drei Anschlusspins: Über einen wird die Versorgungsspannung von 3,3 Volt bezogen. Ein Masseanschluss sorgt für das Bezugspotential für die Spannungsversorgung und den Signalausgang. Das über den dritten Anschlusspin ausgegebene Sensorsignal repräsentiert analog den detektierten Ionenstrom und reicht von 50 Millivolt bis 3,3 Volt.

**Feststellung der Nachweisgrenze** Zur Bestimmung der Nachweisgrenze des Photoionisationsdetektors in der verwendeten Kombination aus Sensor und entwickelter Datenaufnahme-Elektronik wurde ein Testaufbau mit einer geringen Konzentration eines Prüfgasgemisches beaufschlagt. Zur Probenbereitung wurde Toluol in flüssiger Phase mit einer Mikroliterspritze in ein definiertes Luftvolumen gegeben. In einem Verdünnungsschritt wurde von dem Probenvolumen ein Teil in einen weiteren mit Luft gefüllten Probenbeutel gegeben. Über diese Methode wurde ein Prüfgas mit einer Massenkonzentration von 100 ppb Toluol hergestellt. Mithilfe einer Miniatur-Membranpumpe wurde diese durch die aus Sensorkammer, Sensor und Pumpe bestehende Messanordnung gefördert.

Das Messergebnis ist in der Abbildung 3.4 dargestellt. In dem Diagramm ist die einheitenlose Signalthöhe des Photoionisationsdetektors über die Zeit aufgetragen. In den ersten zehn Sekunden der Messzeit wird der Detektor nur mit sauberer Luft umspült. Es ist also das Grundrauschen der Messanordnung zu beobachten. Zum Zeitpunkt  $t = 10\text{ s}$  wird der Sensor mit dem Prüfgasgemisch beaufschlagt. Über den sich ergebenden Signalhub von der Basislinie bis zum Plateau kann die Ordinate skaliert werden. Zur Bestimmung der Nachweisgrenze ( $s_{NWG}$ ) wird nach DIN 32645 der dreifache Wert der Standardabweichung des Grundrauschens herangezogen [71].

$$s_{NWG} = 3 \cdot \sigma \quad (3.1)$$

Für das Experiment dieser Anordnung aus Sensor und Datenaufnahme mit dem Kalibriergas Toluol ergibt sich aus der genannten Norm eine Nachweisgrenze von unter 20 ppb. Toluol besitzt einen Responsefaktor von 0,5 [70].

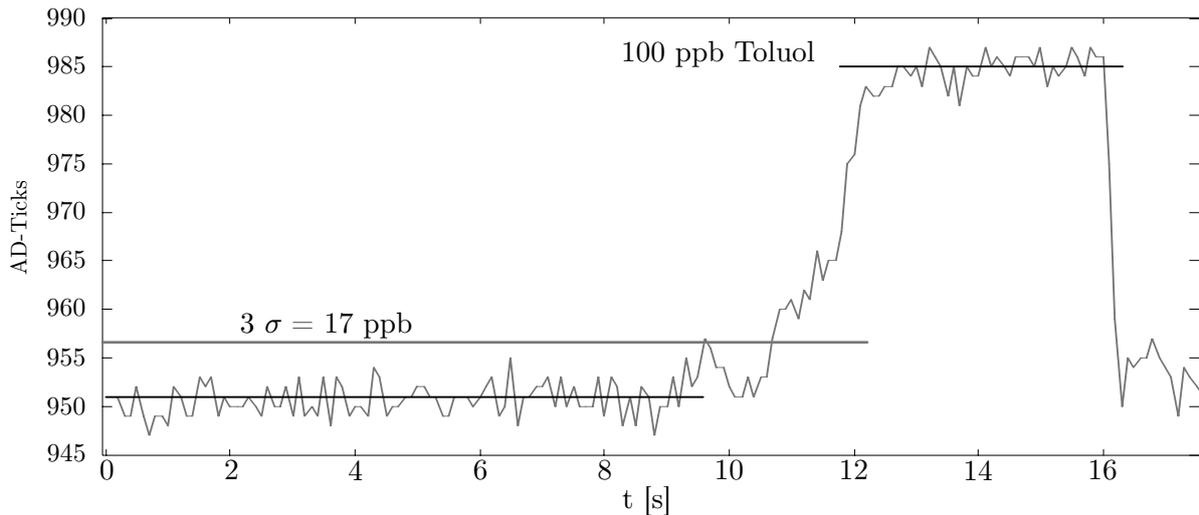


Abbildung 3.4.: Bestimmung der Nachweisgrenze des eingesetzten Photoionisationsdetektors

### 3.1.3. Datenaufnahme und -verarbeitung

Der verwendete Photoionisationsdetektor gibt die Signalhöhe als analoge Spannung zurück. Für die Weiterverarbeitung wird das Signal zunächst bandbegrenzt und über einen 14-Bit Analog-Digital-Wandler digitalisiert. Die Werte werden von einem Mikrocontroller aufgenommen und an eine Auswertesoftware übertragen. Der Controller übernimmt zudem die Steuerung des Injektionsventils, der Pumpen und des Thermostaten zur Trennsäulenbeheizung. Die Auswertung und die damit verbundene Merkmalsextraktion geschehen in der externen Software. Für die Anzeige der Ergebnisse im Feld wird ein Tabletcomputer genutzt.

## 3.2. Realisierung eines Multi-Sensoren-Arrays

Das im Forschungsprojekt DACHS umgesetzte Sensorenarray kombiniert die in dem Grundlagenkapitel vorgestellten Sensortechnologien. Damit wird über die individuellen Signale der Detektoren und der hier zu entwickelnden Klassifikation die Analyse von Gasproben durchgeführt. Mit dem integrierten Kurzsäulen-Gaschromatograph kann die

Identifikation von Stoffgemischen optimiert werden. In den nachfolgenden Kapiteln werden zunächst das fluidische Konzept und die notwendigen Anforderungen an die Sensorik bei einer GC-Analyse beschrieben.

### 3.2.1. Gasführung

Das Messsystem ist für zwei Einsatzarten konstruiert worden: Die *direkte Messung* und die *Messung mit gaschromatographischer Vortrennung*. Um beide Anwendungsfälle optimal bedienen zu können, ist eine Injektionseinheit für den Gaschromatographen nötig, welche den analytischen Gaspfad für die direkte Messung nicht beeinflusst. Dies wird

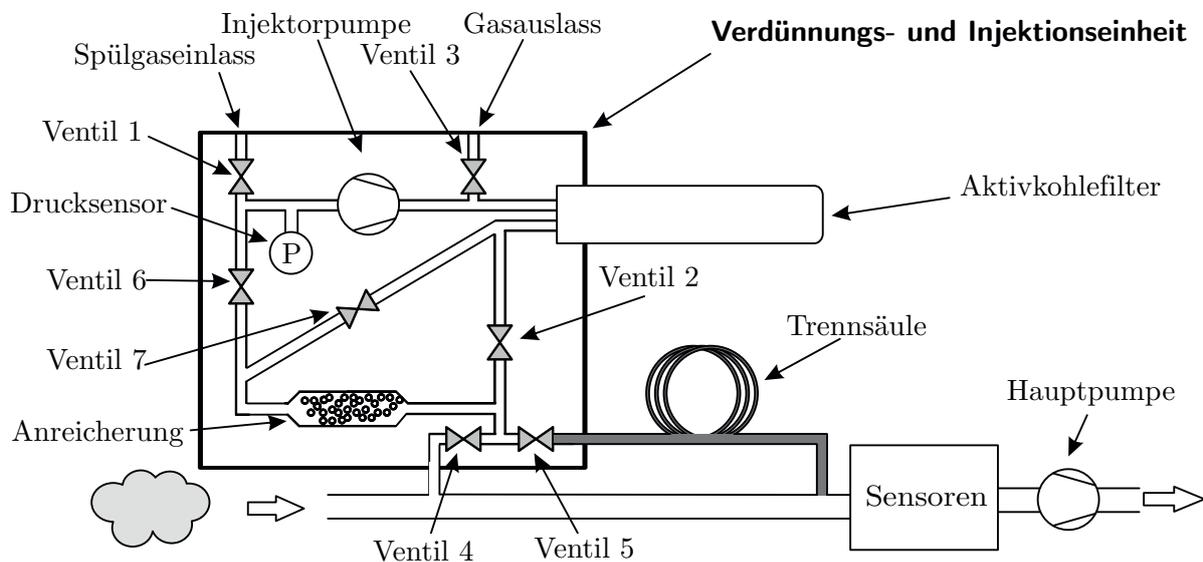


Abbildung 3.5.: Flussplan des realisierten Detektoren Arrays mit Gaschromatograph [10]

durch die in Abbildung 3.5 gezeigte Verschaltung erreicht. Bei einer direkten Messung gelangt die Probe durch den Einlass auf direktem Wege zu den Sensoren. In dem analytischen Pfad befinden sich keine Ventile. Dies reduziert die Möglichkeiten für eine Kontamination der gasführenden Elemente durch hohe Probenkonzentrationen. Der Gaspfad selbst besteht aus inertem Material. Die in der Abbildung rechts von den Sensoren angeordnete Hauptpumpe erzeugt den nötigen Volumenstrom zur Förderung der Probe durch den Sensor-Block.



Tabelle 3.1.: Schaltzustände der Injektions- und Verdünnungseinheit

| Zustand                | V1 | V2 | V3 | V4 | V5 | V6 | V7 |
|------------------------|----|----|----|----|----|----|----|
| <i>Direkte Messung</i> | 1  | 1  | 0  | 1  | 1  | 0  | 0  |
| <i>GC-Sammeln</i>      | 0  | 0  | 1  | 1  | 0  | 1  | 0  |
| <i>GC-Injektion</i>    | 1  | 0  | 0  | 0  | 1  | 0  | 1  |
| <i>GC-Spülen</i>       | 1  | 1  | 0  | 0  | 1  | 0  | 0  |

Legende: 1 offen  
0 geschlossen

**Direkte Messung** Für die beschriebene direkte Messung ohne den Gaschromatographen werden die Ventile der Verdünnungs- und Injektionseinheit in den Schaltzustand *direkte Messung* gebracht. In der Tabelle 3.1 findet sich eine Auflistung aller möglichen Ventilzustände. Bei der *direkten Messung* ist ein Verdünnungspfad durch den Schaltzustand zum Einlass freigegeben. Die Injektorpumpe kann so Spülgas über den Aktivkohlefilter erzeugen und diesen über das Ventil 4 dem Volumenstrom des Probengases beimischen. Das Verdünnungsverhältnis wird über die Drehzahl der Pumpe gesteuert. Gleichzeitig wird über das geöffnete Ventil 5 ein Teil der Verdünnungsluft durch die GC-Trennsäule gefördert. Mithilfe der Durchspülung des Gaschromatographen mit gereinigter Luft, werden persistente Anreicherungen von Substanzen auf der stationären Phase verhindert.

**Gaschromatographie** Ein gaschromatographischer Analyselauf gliedert sich in drei Phasen: Zunächst wird über den Schaltzustand *GC-Sammeln* ein Adsorbens, bzw. eine Loop mit Probengas durchspült. *GC-Injektion* befördert die Gasprobe auf die Trennsäule und *GC-Spülen* transportiert das Stoffpaket durch die Trennstrecke.

Während der *GC-Sammeln*-Phase ist der Ventilweg über V4, V6 und V3 geschaltet. Die Injektionspumpe wirkt in dieser Konfiguration als Vakuumpumpe und saugt die Probe über das Adsorbens bzw. durch die Loop. Wird ein Adsorptionsmittel, wie zum Beispiel Tenax zur Anreicherung verwendet, so wird das Adsorbens von der Probenseite her mit der Substanz beladen. In der Abbildung 3.5 befindet sich die höhere Konzentration rechts auf der dem Ventil 4 zugewandten Seite.

### 3. Aufbau von mobilen GC-Multisensor-Systemen

Zur Injektion des Probengases auf die Trennstrecke wird bei Verwendung der Anreicherung ein Zwischenschritt eingefügt. Hier muss das Adsorbens zunächst ohne Durchfluss aufgeheizt werden. Zur Desorption wird das Adsorptionsmittel auf 120 °C durch einen umlaufenden Widerstandsheizdraht erhitzt. Bei Einsatz einer Loop, also nur eines geringen Puffervolumens, entfällt dieser Zwischenschritt: Es kann direkt in den Schaltzustand *GC-Injektion* gewechselt werden. Wie in Tabelle 3.1 aufgelistet, ist hier der Ventilweg so geschaltet, dass die Injektorpumpe einen Volumenstrom durch den Aktivkohlefilter und das Ventil 7 auf die Anreicherung/Loop fördert. Die Probe wird auf diese Weise in die Gegenrichtung auf die Chromatographiesäule gedrückt. Im Falle der Anreicherung müssen die desorbierten Moleküle (von der rechten Seite der Adsorptionskapillare) also nicht mehr das gesamte (linksseitige) Adsorbens passieren.

Nachdem für ein bis drei Sekunden die Probe in diesem Ventilzustand injiziert wurde, werden die Ventile auf *GC-Spülen* umgeschaltet. Durch das Molsieb gereinigte Luft wird als Trägergas durch Ventil 2 und Ventil 5 zur Trennsäule gepumpt.

#### **Konstruktive Umsetzung der Fluidik**

Die Umsetzung des in Abbildung 3.5 gezeigten Flussplans erfolgte integrativ, angepasst an ein pistolenförmiges Gehäuselayout für die handgehaltene Anwendung.

Die Fluidik kann in zwei Bereiche unterschieden werden: Gasführende Pfade, welche für die Bereitstellung von Nullluft als Verdünnungsluft genutzt werden und analytische Pfade die mit der Probe in Kontakt kommen. Die Nulllufterzeugung wird durch eine kombinierte Einheit aus Aktivkohlefilter und Injektorpumpe sichergestellt. Die stabförmige Kartusche des Aktivkohlefilters ist so positioniert, dass sie von dem Pistolen-Griff des Messsystems umschlossen wird (vgl. Abbildung 3.5). Diese ist austauschbar, da sie über die Nutzungsdauer des Systems mit Spurengasen und Feuchtigkeit der Umgebungsluft beladen wird.

Fluidik-Wege, welche mit der Probe in Kontakt stehen, wurden geometrisch kurz realisiert um aktive Oberflächen so gering wie möglich zu halten. Die um den Probengaseinlass konstruierte Verschaltung der analytischen Komponenten beinhaltet die Anschlüsse für die temperaturgeregelte GC-Säule und eine zylinderförmige Anreicherungs-einheit.

### 3. Aufbau von mobilen GC-Multisensor-Systemen

Die Anreicherungseinheit besteht aus einem koaxial angeordneten Aufbau zweier Zylinder. Das äußere Rohr ist dabei mit einem Heizdraht und einem Thermoelement zur Temperaturbestimmung versehen. Bei dem innen liegenden Zylinder handelt es sich um eine Edelstahlkapillare, welche mit dem Adsorbtionsmittel gefüllt ist. Dieser Adsorbens-Träger ist von außen entnehmbar, sodass das Adsorbens bei zu starker Kontamination durch die Probe ausgetauscht werden kann. Gerade schwerflüchtige Substanzen führen zu einer persistenten Beladung des Adsorbtionsmittels.

Die Multikapillar-Säule für die chromatographische Trennung ist in ein U-förmiges Modul eingebracht. Das GC-Modul (siehe Abbildung 3.6) beinhaltet eine Temperaturregelung und kann hierdurch auf bis zu 60 °C isotherm beheizt werden. Die Gasan-

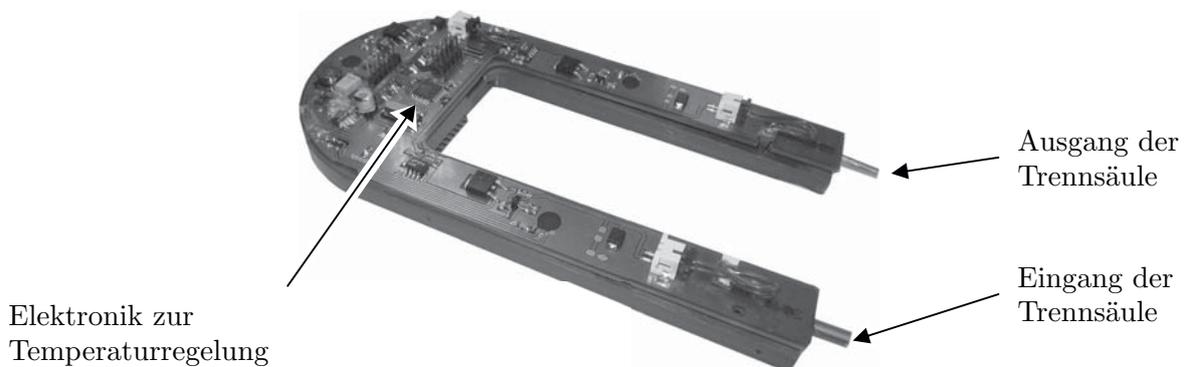


Abbildung 3.6.: Trägermodul des Kurzsäulen-Gaschromatographen [10]

schlüsse der Trennsäulen werden mit umlaufenden Dichtungen in die Aufnahmen der beiden vorderen Ventilblöcke eingefasst. Das gesamte Trägermodul ist auswechselbar. Zum einen kann damit Säulenmaterial mit verschiedenen Belegungen eingesetzt werden und zum anderen kann eine durch Betrieb gealterte Trennkapillare schnell ausgetauscht werden.

#### 3.2.2. Sensorik

Das realisierte Funktionsmuster eines Multi-Sensoren-Arrays kombiniert die bereits vorgestellten Sensortechnologien. Im Einzelnen wurden für die Realisierung drei Halbleitersensoren, zwei Elektrochemische Zellen, ein Photoionisationsdetektor und ein Doppelröhren-Ionenmobilitätsspektrometer eingesetzt. Die Auswahl der Sensoren erfolgte

### 3. Aufbau von mobilen GC-Multisensor-Systemen

anhand der Nachweisfähigkeit von Substanzen der Einsatztoleranzwert-Liste [5]. In der Arbeit von Torsten Hunte [9] wurde die Kombination von Ionenmobilitätsspektrometer und Photoionisationsdetektor sowie zweier Halbleitersensoren und einer Elektrochemischen Zelle hinsichtlich der Detektionsleistung untersucht. Durch diese Kombination wurde eine beinahe umfassende Abdeckung aller ETW-Substanzen erreicht. Aufbauend auf die Sensorenauswahl der Arbeit wurde eine weitere Elektrochemische Zelle als Lückenschluss zum Nachweis von Formaldehyd hinzugefügt. Zusätzlich erhält der Messaufbau einen weiteren Halbleitersensor dessen Nachweisfähigkeit zu untersuchen ist.

In der Abbildung 3.7 sind die eingesetzten Einzelsensoren des Messsystems abgebildet. Bei den Sensoren (A) und (B) handelt es sich um Elektrochemische Zellen, Sensor (C) ist der aus Kapitel 3.1.2 bekannte Photoionisationsdetektor und bei (D) handelt es sich um die Halbleitersensoren mit unterschiedlichem Sensormaterial.

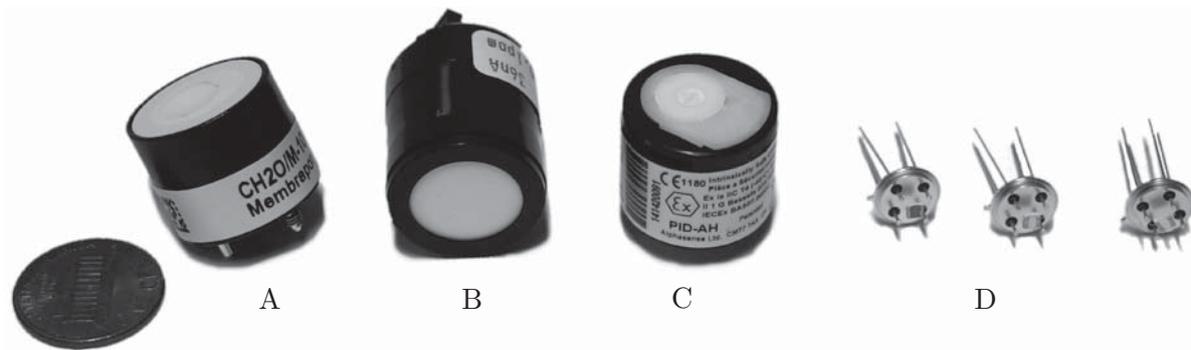


Abbildung 3.7.: Im DACHS-System eingesetzte Sensoren

**Elektrochemische Zelle (A)** Dieser Zellentyp ist laut Hersteller für den Nachweis von Formaldehyd geeignet [72]. Sie besitzt laut Datenblatt zusätzlich eine Querempfindlichkeit auf Ethanol. Es handelt sich um eine dreipolige Zelle ohne Bias-Spannung.

**Elektrochemische Zelle (B)** Der zweite eingesetzte Zellentyp stammt von der Firma Compur Monitors und ist für die Detektion von Phosgen vorgesehen [73]. Der Sensor besitzt einige Querempfindlichkeiten, wie zum Beispiel auf Phosphin oder Blausäure.



**Halbleitergassensoren (D)** Die Halbleitergassensoren der Firma Umweltsensortechnik sind für den breitbandigen Nachweis bestimmt [74]. Durch ihre individuellen Konfigurationen der sensitiven Halbleiterschicht unterscheiden sich die Ausgangssignale hinsichtlich ihres Ansprechverhaltens auf die Probe. Im Einzelnen werden Sensoren mit den Herstellerbezeichnungen *2000*, *5000* und *W000* genutzt [75][76].

### Eingesetztes Ionenmobilitätsspektrometer

Das Ionenmobilitätsspektrometer ist eine Neuentwicklung der Firma Airsense Analytics GmbH [42]. Durch seinen Aufbau als Doppelröhren-IMS eignet es sich besonders für den Einsatz in Kombination mit einer gaschromatographischen Vortrennung, da sowohl positive als auch negative Ionenspezies gleichzeitig erfasst werden können. Das Funktionsprinzip dieses Ionenmobilitätsspektrometers folgt dem im Grundlagenkapitel 2.2.3 dargelegten Prinzip. Die Realisierung der Driftröhreneinheit zur Flugzeitmessung ist in der Abbildung 3.8 zu sehen. Die Driftröhre besteht aus Macor (Abk. engl. : machinable ceramic), also spanend modifizierbarem Keramik-Halbzeug. Daran angeschlossen sind die Kreislauf-Fluidik des IMS sowie die elektrischen Wirkelemente für die elektrischen Felder. Durch den angepassten Aufbau des Probeneinlasses kann das Spektrometer optimal in die Gasführung, welche in Grafik 3.5 gezeigt wurde, eingepasst werden. Das Proben-

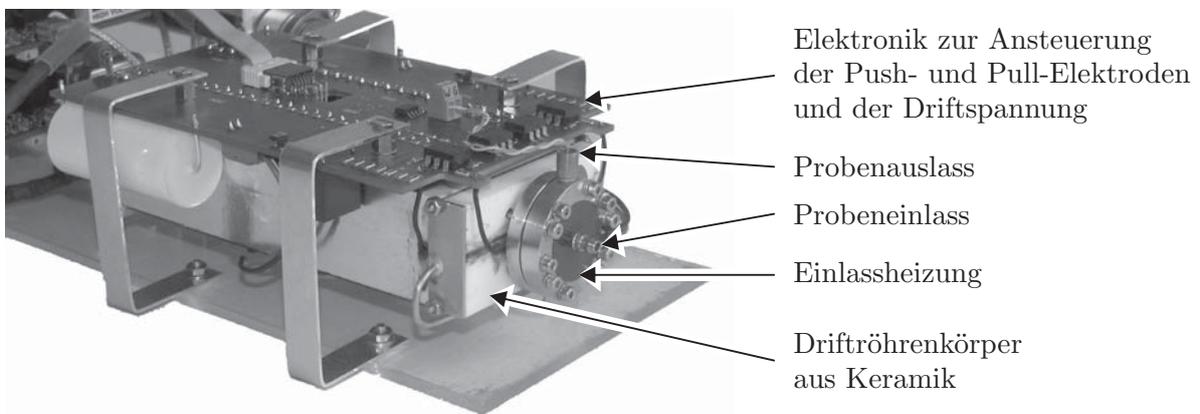


Abbildung 3.8.: Foto der Eingangsseite des verwendeten Doppelröhren-Ionenmobilitätsspektrometers

gas wird in die hervorstehende Einlassöffnung in eine beheizte Kammer eingeleitet. Hier strömt es an einer ebenfalls beheizten Membran vorbei in Richtung der obenliegenden

### 3. Aufbau von mobilen GC-Multisensor-Systemen

Gasaustrittsöffnung. Aufgrund des Konzentrationsgradienten zwischen Einlass-Kammer und Reaktionsraum im Driftröhrenkörper, findet eine Permeation der Analytmoleküle in Richtung des Reaktionsraums statt. Als Membranmaterial findet Polydimethylsiloxan (PDMS) mit einer Stärke von 10  $\mu\text{m}$  hier Anwendung [11]. Der Driftraum wird durch die Membran pneumatisch von der Umgebung abgetrennt. Die beheizte Einlassmembran reduziert das Eindringen von Wassermolekülen.

Der Reaktionsraum enthält eine aus Nickel-63-Folie bestehende Ionenquelle. Im Bereich hinter dem zylinderförmigen Einlassblock befinden sich die, für eine Ionen-Injektion notwendigen Push- und Pullelektroden. In der Abbildung sind die zugehörigen elektrischen Anschlussleitungen sichtbar (schwarz). Die entsprechende impulsgebende Elektronik ist auf der Leiterplatte direkt über der weiß erscheinenden Driftröhre angeordnet. Hier befindet sich auch die Widerstandskaskade zur Aufteilung von elektrischen Potentialen auf die im Inneren des Driftraums befindlichen Ringelektroden. Im Bild nicht zu sehen sind die beiden Faraday-Auffänger zur Detektion der Ionen. Sie befinden sich im Inneren der Driftstrecke und sind mit einer hochempfindlichen Vorverstärkerschaltung am hinteren Ende der Röhre verbunden. Die Detektoren befinden sich elektrisch auf dem jeweiligen Hochspannungs-Potential der Beschleunigungsspannung.

#### Gasdistribution

Die sechs Einzelsensoren und das Ionenmobilitätsspektrometer benötigen eine definierte Gas-Zuführung. Wie im vorherigen Abschnitt dargelegt, wird der Einlass des Ionenmobilitätsspektrometers mit der Probe durchspült. Ein Teil permeiert für die Analyse durch eine Membran, der andere größere Teil wird durch den Auslass zu den Einzelsensoren geleitet. In der Abbildung 3.9 ist die realisierte Gasdistribution dargestellt. Der als *Sensorhalter* bezeichnete Gasverteilungsblock ist ein lasergenerierter Edelstahlblock [77][78]. Hier werden die Einzelsensoren mechanisch montiert und fluidisch an den Gasfluss angekoppelt. Die von oben zugänglichen Sensoren werden über O-Ringe axial gegen den Block gedichtet. Das kleine Volumen zwischen Sensorblock und Wirkfläche des Sensors wird über je eine Einlass- und Auslassöffnung mit der Probe durchspült. Das Herstellungsverfahren des Lasergenerierens erlaubt innenliegende Kanäle zur Gasverteilung. In Teilabbildung 3.9 b) sind die im Block eingebrachten Kanäle zu sehen. Der Volumenstrom durch den Sensorhalter wird auf drei Kanäle aufgeteilt. Durch einen werden die

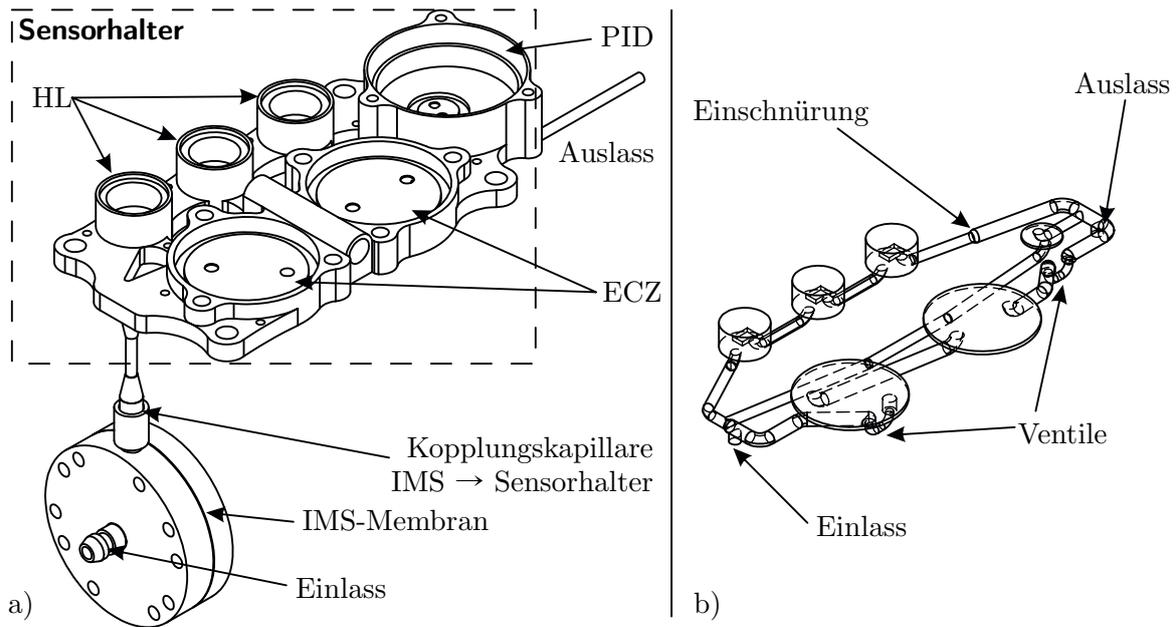


Abbildung 3.9.: Realisierte Gasführung und Sensoren-Aufnahme [10]

drei Halbleitersensoren umspült, der mittlere Kanal versorgt den Photoionisationsdetektor und in den Dritten sind die Elektrochemischen Zellen eingebracht. Durch diese Aufteilung werden gegenseitige Beeinflussungen der Sensoren verhindert. Gerade die Elektrochemischen Zellen neigen dazu Substanzen bei hohen Konzentrationen auf ihren Oberflächen anzureichern und im Betrieb allmählich wieder freizugeben. Durch die parallelisierte Anordnung wird mögliches analytisches Memory in Richtung des Auslasses gespült und beeinflusst nicht die anderen Sensoren. Aufgrund der in den Elektrochemischen Zellen vorhandenen Elektrolyten, sind diese einer Alterung durch Austrocknung unterworfen. Um der frühzeitigen Degradation des Sensors entgegenzuwirken befinden sich im Kanal der Zellen zwei Ventile zur Absperrung. Im ausgeschalteten Systemzustand werden die Ventile verschlossen, so dass kein Gasaustausch in diesem Pfad mehr stattfinden kann. Die Halbleitersensoren besitzen heiße Oberflächen. Durch die punktuell hohen Temperaturen von circa 400 °C können die Moleküle des Probengases gegebenenfalls pyrolytisch zersetzt werden. Dies hätte dann Einfluss auf die Detektion durch andere Sensoren.



### 3.2.3. Elektronik

Die sechs Sensoren wandeln in der Messkette die physikalische Größe der Stoffkonzentration mit ihrer jeweiligen Gerätefunktion in elektrische Signale um. Im Falle der Halbleitersensoren ist elektrische Sensorantwort eine Widerstandsänderung, bei den Elektrochemischen Zellen ein sich einstellender Strom und im Fall des vorgestellten integrierten Photoionisationsdetektors handelt es sich um eine Spannung am Sensorausgang. Die elektrische Größe muss für die Digitalisierung einheitlich in einen gut messbaren Spannungsbereich gewandelt werden. Zur Aufnahme von Ionenmobilitäts-Spektren sind neben der Messwerterfassung noch umfangreiche Filterung der Signale sowie eine aufwendige Ansteuerung der Drift-Spannungen notwendig. Zudem müssen in dem vorgestellten System verschiedene Aktorik Elemente wie die Pumpen zur Gasförderung und die Ventile gesteuert werden. Die Temperaturen von Gaschromatograph und Anreicherungseinheit müssen auf ihre Sollwerte geregelt werden. Da es sich um ein handgehaltenes Messinstrument handelt, ist zudem ein Akku-Management nötig. Diese zu erfüllenden Aufgaben sind als Hardware-Module realisiert worden. Jedes Modul wird von einem zentralen Mikroprozessor abgefragt und gesteuert. In der Abbildung 3.10 ist der modulare Aufbau der Elektronik visualisiert. Jedes der Module enthält einen Mikrocontroller und kann seine

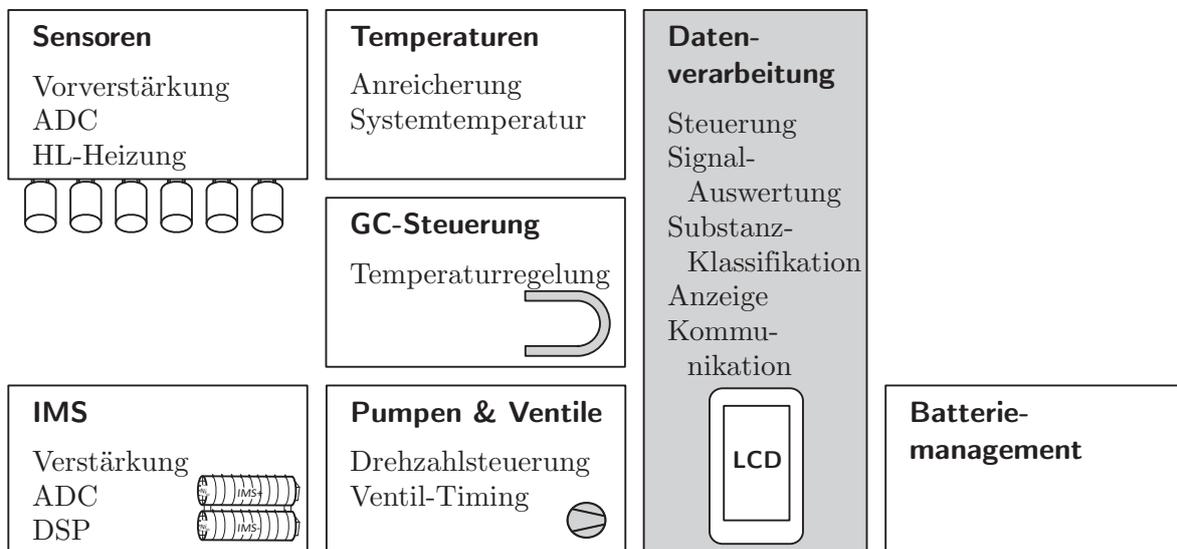


Abbildung 3.10.: Blockschaltbild der Elektronik

Aufgaben autonom ausführen. Steuerbefehle werden dabei von dem grau dargestellten Modul „Datenverarbeitung“ über einen internen seriellen Bus übertragen. Das Protokoll

zum Datenaustausch folgt dem Master-Slave-Prinzip. Für die Messdaten-Aufnahme bedeutet dies, dass der Master jeden Datensatz aktiv von dem jeweiligen Modul anfordern muss.

#### Sensoren

Die Elektronik zur Abfrage der sechs Einzelsensoren beinhaltet die individuellen Ansteuerungs-Schaltungen der Sensoren, die analoge Vorverarbeitung sowie Digitalisierung der Signale.

**Halbleitergassensoren** Für den Betrieb der Halbleitergassensoren wird, wie in den Grundlagen beschrieben, eine Thermostatisierung des Sensormaterials benötigt. Wechselnde Volumenströme am Sensor verursachen im ungeregelten Fall eine Abkühlung des Sensors und somit eine Veränderung des Widerstandes der sensitiven Schicht. Eine Temperaturregelung geschieht auf Basis des Wertes des Platin-Heizwiderstands. Der Widerstandswert beeinflusst eine Messbrücke, deren Brückenspannung die Heizer-Spannung steuert. In diesem analogen Aufbau wird die Führungsgröße über ein digitales Potentiometer vorgegeben. Es ist somit kein weiterer digitaler Regelaufwand nötig. Die Wandlung des von der Stoffkonzentration abhängigen Widerstandswertes (Gleichung 2.2) in eine direkt messbare Spannung erfolgt durch einen angepassten ohmschen Spannungsteiler mit einer Referenzspannung am Kopf des Teilers.

**Elektrochemische Zellen** Bei der eingesetzten Elektrochemischen Zelle zum Nachweis von Phosgen handelt es sich um eine Zwei-Elektroden-Zelle. Diese benötigt keine weitere Ansteuerungselektronik. Die Formaldehydzelle ist als Drei-Elektroden-Zelle ausgeführt und bedarf daher eine Potentiostatschaltung, welche die Gegenelektrode konstant auf einem Nullpotential hält. Diese wurde nach Herstellerangaben dimensioniert [79]. Da Elektrochemische Zellen Substanzen amperometrisch detektieren (siehe Gleichung 2.12), ist eine Schaltung notwendig, welche den sich einstellenden Strom verstärkt und in eine Spannung wandelt. Die implementierten Transimpedanz-Verstärkerschaltungen wandeln die zu erwartenden Ströme von einigen Mikroampere in einen Spannungsbereich von 0 bis 4 Volt.



**Photoionisationsdetektor** Die Ausgangsspannung des integrierten Sensors wird direkt digitalisiert. Die Elektronik stellt dem Sensor eine präzise und rauscharme Betriebsspannung zur Verfügung.

Die analogen Signale der drei genannten Sensorklassen werden mit einer Abtastrate von 10 Hertz digitalisiert. Die Information der Signalhöhe wird taktsynchron an das Modul Datenverarbeitung (vgl. Abb. 3.10) weitergeleitet.

#### **Ionenmobilitätsspektrometer**

Das Elektronikmodul des Ionenmobilitätsspektrometers gliedert sich in zwei Komponenten: Dies ist zum einen die Hochspannungsverteilung auf der Oberseite der Driftröhre mit den beiden unmittelbar an die Detektoren angekoppelten Vorverstärkern. Zum anderen existiert ein Hardwaremodul zur Generierung der Hochspannung und zur Aufnahme und Filterung der Spektren. Da sich die Detektoren jeweils auf dem Hochspannungspotential der zugehörigen Driftröhre befinden, findet die weitere analoge Verstärkung in einer zweiten Stufe als auch die Digitalisierung auf dem Hochspannungspotential statt. Die digitale Kommunikation zwischen Mikroprozessor und Analog-Digital-Wandler erfolgt galvanisch getrennt über hochspannungsfeste Optokoppler. Die Spektren werden mit einer Abtastrate von 44  $\mu$ s aufgenommen. Nach der Abtastung erfolgt eine Filterung des Signals mithilfe eines im Mikroprozessor integrierten Digitalen Signalprozessors (DSP) (siehe Kapitel 4.3). Die so vorverarbeiteten Spektren mit jeweils 1024 Stützstellen werden durch das Modul *Datenverarbeitung* wiederkehrend abgefragt.

Das Ionenmobilitätsspektrometer besitzt Temperaturregelungen für die Einlassmembran und den Driftröhren-Körper. Bei der Pumpe zur Erzeugung des Driftstroms handelt es sich um eine Drehschieberpumpe, deren Förderleistung über ihre Betriebsspannung von der IMS-Hardware gesteuert wird. Die Hochspannungen werden über steuerbare Gleichspannungsnetzteile generiert. Die Steuerspannung wird über einen Digital-Analog-Wandler vorgegeben. Gleichzeitig verfügt die Elektronik über eine Schaltung zum Rücklesen der anliegenden Hochspannung. Diese wird über das unterste Element der Widerstandskaskade abgegriffen. Die Temperaturregelung und die Hochspannungsüberwachung geschehen autonom von diesem Hardwaremodul.



#### **GC-Steuerung und Temperaturen**

Diese beiden Elektronikmodule führen Temperaturregelungen für die Anreicherungseinheit und für die GC-Säule durch. Die Heizung der Anreicherung besteht aus einem Thermosteuelement und einem Heizleiter. Die Regelung erfolgt durch einen digitalen PI-Regler. Das U-förmige GC-Modul besitzt eine in drei Heizzonen unterteilte Silikon-Heizmatte mit digitalen Temperatursensoren. Die PI-Regelung erfolgt über pulsweitenmodulierte Steuerspannungen für jede Heizzone. Die Elektronik zur Regelung befindet sich auf dem austauschbaren GC-Modul und wird elektrisch über einen Steckkontakt an den internen Steuerbus angekoppelt.

#### **Pumpen & Ventile**

Die Injektorpumpe und die Hauptpumpe zur Probengasförderung müssen in ihrer Förderleistung aufeinander abgestimmt sein. Die elektronische Ansteuerung gibt hierzu die Betriebsspannung für bestimmte Arbeitspunkte vor. Dies geschieht ebenfalls über eine Pulsweitenmodulation und dahinter geschaltete Leistungstreiber. Bei den sieben Ventilen der *Verdünnungs- und Injektionseinheit* und den zwei Ventilen zur Abschaltung des Gaspfades der Elektrochemischen Zellen handelt es sich um bistabile Ventile. Diese werden mithilfe einer H-Brückenschaltung durch Polaritätsumkehr geschaltet.

#### **Elektronik zur Datenverarbeitung**

Dieses Elektronikmodul steuert die anderen Module durch Sollwertvorgaben und fragt die Messdaten über den internen Bus ab. Hier wird durch den eingesetzten *ARM9* Mikroprozessor ein Klassifikationsverfahren auf die multivariaten Signale angewandt. Das Ergebnis der Auswertung wird über verschiedene Schnittstellen dem Benutzer zur Verfügung gestellt (siehe Abbildung 3.11). Die aufgenommenen Messdaten und die Ergebnisse des Auswertalgorithmus werden zudem auf einem Massenspeicher in Form einer SD-Karte gespeichert.

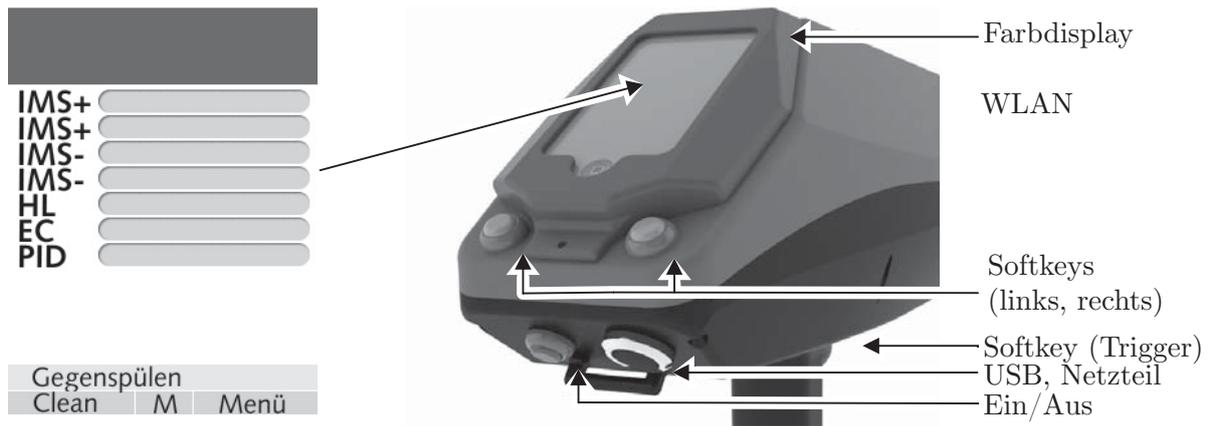


Abbildung 3.11.: Benutzerschnittstellen

### 3.2.4. Datenaufnahme und -verarbeitung

Durch die verteilten Mikroprozessoren kann das System einen großen Teil der Steuerungs- und Datenaufnahme-Aufgaben nebenläufig durchführen. Die Software des Hauptprozessors greift die vorverarbeiteten Messdaten der Co-Prozessoren Timer-gesteuert ab. Im Programmablauf werden die Daten fusioniert und im gemeinsamen Kontext einer Auswertung unterzogen. Der eingesetzte Prozessor unterstützt die Funktion des Speicherdirektzugriffs (engl. Direct Memory Access - DMA) und kann so gleichzeitig Daten über seine Peripherie aufnehmen und ausgeben. Die im Speicher gehaltenen Messdaten können daher verarbeitet werden, während neue Daten abgefragt und die Ergebnisse der Datenverarbeitung auf einem Display und über die digitalen Kommunikationsschnittstellen USB und WLAN ausgegeben werden. Die Datenverarbeitung erfolgt über einen Prozess-Scheduler mit Echtzeit-Eigenschaften. Die Routine zum Abruf von Messdaten erfolgt zu festen Zeitpunkten im Programmablauf. Direkt darauffolgend wird ein schon im Hauptspeicher vorhandener Datensatz für eine direkte Merkmalserkennung verarbeitet. Die eingesetzten Filterverfahren und die Methoden der Merkmalserkennung, sowie Klassifikation werden im Detail im nachfolgenden Kapitel 4 beschrieben.

### 3.2.5. Umsetzung als Handgerät

Das entworfene und realisierte Messsystem besitzt durch seinen Aufbau Eigenschaften für die Verwendung durch Fachpersonal für die Gefahrstoffanalyse. Die nachweisstarke

### 3. Aufbau von mobilen GC-Multisensor-Systemen

Detektortechnologie ist in einem robusten Kunststoffgehäuse untergebracht. In Abbildung 3.12 ist ein Funktionsmuster des umgesetzten Messsystems gezeigt. Das Gehäuse



Abbildung 3.12.: Aussenansicht des Messsystems

besteht aus zwei Schalen, welche über eine umlaufende Dichtschnur, die analytischen Komponenten im Inneren gegen Staub und Spritzwasser schützen. Gleichzeitig kann das Gerät hierdurch einfach geöffnet werden, um die Sensoren für Wartung und Austausch erreichen zu können. Das Einsatzfeld für ein solches System sieht vor, dass es in kontaminierten Bereichen zur Anwendung kommen wird. In solchen Lagen wird der Bediener durch Schutzkleidung und Atemschutzausrüstung gegen Umwelteinflüsse geschützt. Die Ausrüstung schränkt den Nutzer in der Ausführung filigraner Arbeiten stark ein. Das Bedienkonzept setzt daher auf wenige haptische Eingabemöglichkeiten. Messungen können über eine Ein-Knopf-Bedienung an der Zeigefingerposition des Handgriffs gestartet werden. Zwei weitere Bedientaster im Sichtfeld des Nutzers dienen einer rudimentären menügeführten Parametrierung.

Eine Beleuchtung von schwer einsichtigen Messstellen ist über Leuchtmittel im Einlassbereich vorgesehen.

Der Probengaseinlass selbst besitzt einen groben Filter zur Vermeidung des Eindringens von Fremdkörpern.

## 4. Analyse des Signalverhaltens

Für ein Verständnis der Messsignale und deren Interpretation ist es notwendig, zunächst die vielfältigen Einflüsse der Signalkette zu untersuchen. Der Weg von der Gas-Probe bis zu einem vom Instrument angezeigten Messergebnis ist in der Abbildung 4.1 veranschaulicht. Die verschiedenen Komponenten der Messkette haben jeweils einen Einfluss

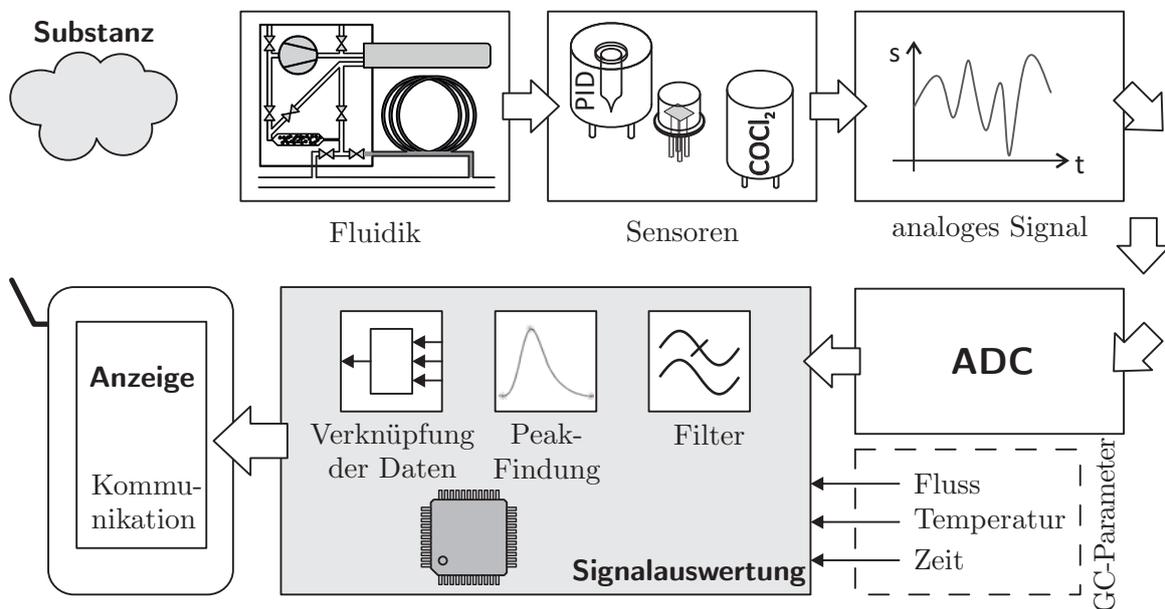


Abbildung 4.1.: Signalpfad

auf das Signalverhalten. So durchwandert das Probengas mit einem bestimmten Massenstrom den Gaspfad. Verdünnungs- und Oberflächeneffekte sorgen dabei für eine Verschleifung der Probengaskonzentration. Nachfolgend haben die Sensoren selbst, durch ihr individuelles Ansprechverhalten, einen erheblichen Einfluss auf den Signalverlauf. Der darauffolgende analoge Messumformer besitzt ebenfalls eine Charakteristik, welche im Idealfall als linear zu betrachten ist. Bei dieser Komponente wird die Störgröße

#### 4. Analyse des Signalverhaltens

des Messrauschens besonders deutlich. Der für die Digitalisierung zuständige Analog-Digital-Umsetzer wirkt durch seine Quantisierungsstufen und mögliche Linearisierungs- und Offset-Fehler auf das Signal ein. Diesen Elementen der klassischen Messketten [80] mit ihren einzelnen Kennlinien folgt die digitale Signalverarbeitung. Hier kommen verschiedene Filtermethoden zur Rauschunterdrückung die nachfolgend in den Kapiteln 4.2.1 und 4.3.1 beschrieben werden, zum Einsatz.

Der Einstiegspunkt zur Messwertinterpretation in dieser Arbeit ist der Funktionsblock *Signalauswertung* der Abbildung 4.1 nach der Digitalisierung. In der nachfolgenden Abbildung 4.2 ist exemplarisch ein Signalverlauf von sechs Sensoren mit der beschriebenen Datenaufnahme gezeigt. Dargestellt ist der zeitliche Verlauf der sich bei der Messung

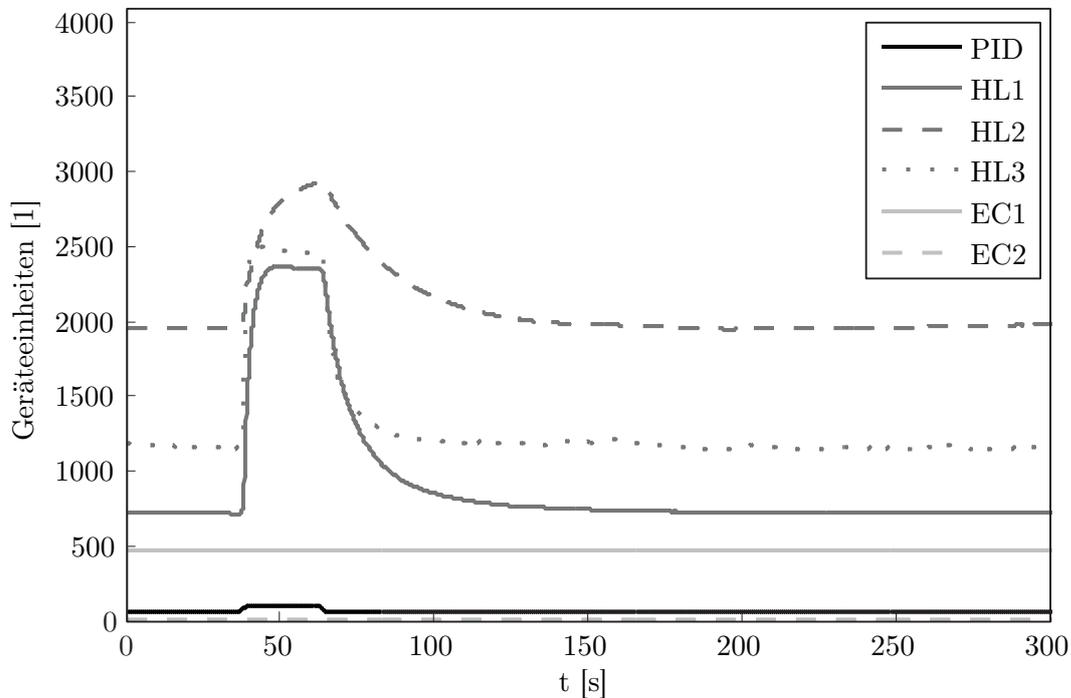


Abbildung 4.2.: Zeitlicher Verlauf einer direkten Messung ohne Datenvorverarbeitung

eines Konzentrationssprungs vom Gaseingang des Systems bei einer Probe mit 25 ppm Aceton einstellt. Die Skalierung der Ordinate stellt den gesamten Messbereich der Sensoren dar. Da es sich bei der Digitalisierungsstufe um einen 12-Bit Analog-Digital-Umsetzer handelt, kann ein Wertebereich bis 4096 Geräteeinheiten aufgenommen werden. In der Darstellung ist zu erkennen, dass die Sensoren sich in ihrem Ansprechverhalten unter-

scheiden. So besitzt der zweite Halbleitersensoren (*HL2*) beispielsweise eine flachere Anstiegsflanke, als der erste Halbleitersensoren (*HL1*). Der Photoionisationsdetektor reagiert in dem Diagramm schwächer als die anderen Sensoren, da dieser Sensor im Gegensatz zu den Halbleitern ein sehr Konzentrations-lineares Verhalten besitzt. Der Wertebereich des PIDs ist in dieser Rohdatendarstellung zudem aus technischen Gründen auf 3300 Geräteeinheiten beschränkt. Weitere zu berücksichtigende Eigenschaften der Rohdaten sind die Höhen der Basislinien und das Signalrauschen.

Diese Art von Signalverläufen ist sowohl zur Merkmalsextraktion von Einzelsensoren eines Sensorenarrays zu interpretieren als auch bei einer analytischen Betrachtung von GC-Läufen. Signale eines Ionenmobilitätsspektrometers besitzen eine viel kleinere Zeitbasis und ein höheren Rauschanteil. In den drei nachfolgenden Abschnitten werden diese drei Fälle der Merkmalsextraktion eingehend betrachtet.

### 4.1. Merkmalsextraktion der Einzelsensoren

Das bedeutendste Merkmal eines Sensorsignals ist die Signalthöhe. Sie ist linear oder exponentiell verknüpft mit der Stoffkonzentration. Aus der Signalform können jedoch noch weitere Merkmale für eine Interpretation abgeleitet werden.

Bei dem Sensortyp der Halbleitersensoren wird die Form des Signals durch Eigenschaften der Proben-Moleküle beeinflusst. Bestimmend sind hierbei diese Einflussfaktoren: Der Haftkoeffizient zwischen Proben-Molekülen und Sensor-Oberfläche, die Anzahl freier Adsorptionsplätze, und die Möglichkeit von katalytischen Reaktionen auf dem Sensor [53][81]. Diese Einflüsse spielen auch bei den anderen Sensoren eine Rolle. Die Elektrochemischen Zellen sind zudem auf eine Diffusion der Proben-Moleküle in den Elektrolyten angewiesen. Die resultierende Permeationsdauer führt zu einer flachen Anstiegskurve.

In der Abbildung 4.3 ist eine typische Signalform eines Halbleitersensors bei Anlegen eines Konzentrationssprungs gezeigt. Die Sensorantwort auf den beaufschlagten Konzentrationssprung wird definiert als der Mittelwert des Signal-Plateaus während des Konzentrationssprungs. Weiterhin lässt sich die Höhe der Basislinie aus dem Diagramm ablesen. Anstiegs- und Abfallzeit kann als Dauer zwischen dem Erreichen von 10% und

#### 4. Analyse des Signalverhaltens

90 % des Signalhubs angesehen werden. Neben diesen vier offensichtlichen Merkmalen,

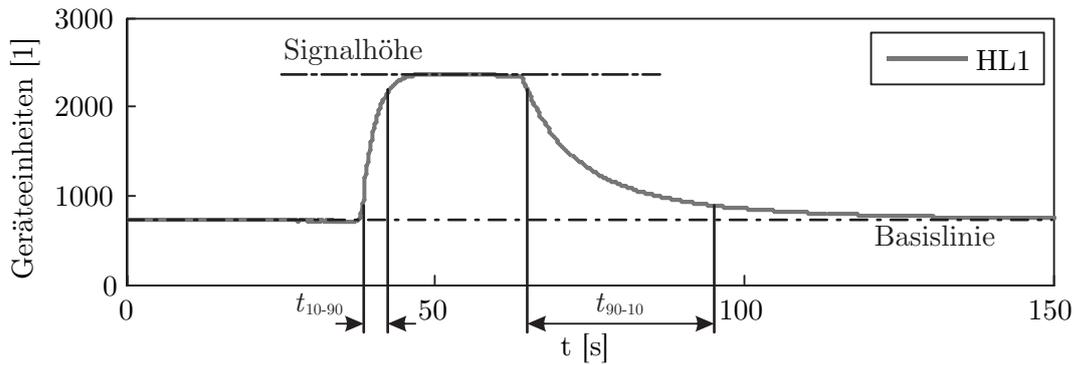


Abbildung 4.3.: Merkmale eines Signalsprungs

wie in Tabelle 4.1 aufgelistet, können noch weitere interpretationsfähige Parameter aus dem gezeigten Verlauf abgeleitet werden.

| PARAMETER              | BESCHREIBUNG   | FORMELZEICHEN   |
|------------------------|--|---|
| Basislinie             | Mittelwert des Sensorsignals bei Inertgas  | $\bar{s}_{\text{null}}$   |
| Signalmaximum          | Maximalwert nach stationärem Zustand bei anliegender Probe                         | $\hat{s}_{\text{probe}}$  |
| relative Sensorantwort | Mittelwert nach stationärem Zustand bei anliegender Probe anzüglich der Basislinie | $\tilde{s}_{\text{rel}}(t) = \hat{s}_{\text{probe}}(t) - \bar{s}_{\text{null}}$ |
| Anstiegszeit           | Zeit für einen Signalsprung von 10 % bis 90 % der relativen Sensorantwort          | $t_{10-90}$   |
| Abfallzeit             | Zeit für einen Signalabfall von 90 % bis 10 % der relativen Sensorantwort          | $t_{90-10}$   |

Tabelle 4.1.: Extrahierbare Parameter (Auszug) [53]

So berücksichtigt Eklöv et. al [53] zur Auswertung von Halbleitergassensoren 46 weitere Parameter, wie die erste Ableitung des Plateaus, das Integral von steigender und fallender Signalflanke und per Auto-Regression parametrisierte angepasste Kurven.

### 4.1.1. Relative Sensorantwort

Um die relative Sensorantwort zu erhalten, muss das zeitabhängige Signal des Sensors um den Bias der Signalbasis reduziert werden. Alle eingesetzten Sensoren sind über die Zeit einer Alterung unterworfen, welche zu einer Drift des Ausgangssignals führt. Elektrochemische Zellen besitzen deshalb eine vom Hersteller vorgegebene begrenzte Haltbarkeit. Bei Halbleitersensoren sorgen neben der Degradation der sensitiven Schicht auch Schwankungen der Betriebstemperatur für eine Veränderung der Basislinie [82] [83]. Das Signal des Photoionisationsdetektors ist dagegen weniger einer Drift unterworfen. Ein geringer Effekt tritt durch eine allmähliche Verschmutzung des Lampenfensters und der damit einhergehenden verringerten Ionenausbeute aus. Gerade für die Halbleitersensoren ist es daher wichtig, möglichst häufig die Höhe der Signalbasis festzustellen.

Praktisch erfolgt die Bestimmung des Nullpunktes durch die Bereitstellung von gereinigter Luft. Entsprechend des in Kapitel 3.2 dargelegten Aufbaus geschieht dies mithilfe einer mit Aktivkohle gefüllten Filterkartusche. Bei der Beprobung mit Null-Luft wird eine gemittelte Höhe des Signals bestimmt. Für die Dauer der eigentlichen Messung wird davon ausgegangen, dass die Basislinie nicht variiert. Daher wird das Relativsignal als Differenz vom Momentanwert und festgestelltem Basislinienwert berechnet (Gleichung 4.1):

$$\tilde{s}(t) = s(t) - \hat{s}_{\text{null}} \quad (4.1)$$

Dieser von der Basisliniendrift befreite Wert kann für eine Dateninterpretation mit Werten anderer Messreihen verglichen werden. Um den sich ergebenden Sensorhub für eine bestimmte Stoffkonzentration zu bestimmen, ist ein stationärer Endwert des Signals erforderlich. Ein Punkt des sich ergebenden Plateaus, bzw. ein Mittelwert der Werte im stationären Zustand mit angelegter Probe kann dann als Merkmal für ein Klassifikationsverfahren genutzt werden.

### 4.1.2. Anstiegszeit

Zur Bestimmung der  $t_{10-90}$ -Zeit an der steigenden Signalfanke werden die Durchgänge bei 10 % und 90 % der relativen Sensorantwort betrachtet. Das Signal wird hierfür

## 4. Analyse des Signalverhaltens

zunächst von der Basislinie nach Gleichung 4.1 befreit. Zur Berechnung von  $\tilde{s}_{(t_{10})} = 0,1 \cdot \tilde{s}$  und  $\tilde{s}_{(t_{90})} = 0,9 \cdot \tilde{s}_{\text{rel}}$  muss zunächst die relative Sensorantwort  $\tilde{s}$  im stationären Zustand bestimmt worden sein. Das Anstiegszeit-Merkmal kann daher nicht unmittelbar zur Signallaufzeit abgeleitet werden. Zusätzlich ist es für dieses Merkmal erforderlich, dass es sich bei dem Messsignal immer um einen Sprung von der Nullkonzentration auf den Konzentrations-Endwert handelt. Die gleichen Bedingungen gelten für eine Bestimmung der Abfallzeit  $t_{90-10}$ . Wie aus Abbildung 4.3 zu entnehmen, ist diese Zeit, das *Tailing*, jedoch üblicherweise erheblich länger als die Anstiegszeit. Die Anstiegs- und Abfallzeit sowie die Signalthöhe sind Merkmale, welche von der Konzentration einer Probe abhängig sind.

### 4.1.3. Auswahl der Merkmale

Zur Bestimmung von Merkmalen, wie sie in Tabelle 4.1 und in der Arbeit von Eklöv et. al. [53] beschrieben sind, ist ein definiertes Verlaufsmuster der Messung erforderlich. Ein Signalverlauf, wie er in Abbildung 4.3 gezeigt wurde, ist als ideal anzusehen. Eine Messung muss also diesem Ablauf folgen:

1. Nullluft für eine ausreichende Zeit auf Sensorik geben
2. Zu einem definierten Zeitpunkt das Analytgas anlegen
3. Beprobung für eine bestimmte Zeit andauern lassen (Mindestzeit richtet sich nach der Dauer bis zum Eintreten eines stationären Zustandes)
4. Nullluft für eine gewisse Zeit auf Sensorik geben (Mindestzeit richtet sich nach der Dauer bis zum Eintreten eines stationären Zustandes)

Für eine praktische Messaufgabe im Feld ist dies nicht immer realisierbar. Durch eine Verdünnungseinheit mit einer Gegenspülfunktion, wie sie in Kapitel 3.2 beschrieben wurde, ist pulsformige Probenaufgabe zwar möglich, jedoch muss idealerweise in solch einem Fall ein Quasi-Unendlicher Vorrat an Probengas zur Verfügung stehen. Gleichzeitig darf die Konzentration der Probe während der Messung nicht variieren. Praktisch treten bei einem unbegrenzten Volumen bei einem Feldeinsatz zur Gefahrstoffdetektion jedoch immer Verdünnungseffekte auf, welche kaum stationäre Konzentrationsverhältnisse zulassen. Zudem ist eine repetierende Messung nur eingeschränkt möglich, da nach Anlegen

## 4. Analyse des Signalverhaltens

des Konzentrationssprunges das Sensorsignal zunächst bis zur Basislinie abgeklungen sein muss. Bei der kontinuierlichen Messung in einer *Monitoring*-Applikation können permanente Signalveränderungen auftreten, z. B. durch einen allmählichen Anstieg der Umgebungskonzentration auf. Damit sind einige Merkmale, wie die Anstiegszeit für die Identifikation durch die große Variabilität nicht nutzbar.

In dieser Arbeit wurde für die Auswertung der Einzelsensoren bei einer kontinuierlichen Direktmessung daher nur das Merkmal der Signalthöhe in Bezug auf das Merkmal der Basislinienhöhe betrachtet. Der Einfluss der Anstiegszeit wurde zusätzlich untersucht.

Ausblickend können weitere Merkmale in Betracht gezogen werden, wenn in bestimmten Anwendungsfällen auf eine kontinuierliche Messung verzichtet wird. Durch eine intelligente Ablaufsteuerung, bei der im Wechsel Nullluft bis zum vollständigen Abklingen auf die Signalebasislinie aufgegeben wird, und dann Probe für eine bestimmte Zeit beaufschlagt wird, könnten weitere Signalparameter zur Auswertung herangezogen werden.

### 4.2. Merkmalsextraktion von GC-Verläufen

Im Gegensatz zur Auswertung von kontinuierlichen Signalen der Einzelsensoren, ist die Chromatographie ein diskontinuierlicher Prozess. Das Ergebnis dieses Vorgangs ist ein Chromatogramm in dem die lokalen Extrema des zeitabhängigen Detektorsignals die Komponenten der Substanzprobe repräsentieren. In Abbildung 4.4 ist beispielhaft ein mit dem in Kapitel 3.1 beschriebenes GC-PID-System gemessenes Chromatogramm dargestellt. Es wurde eine Probe mit einem Gasgemisch aus Aceton, Toluol und Xylol vermessen. Dargestellt ist das digitalisierte und skaliertes Signal des Photoionisationsdetektors. Die drei ausgeprägten lokalen Maxima sind die Detektionsergebnisse der drei nacheinander aus der Trennsäule eluierten Einzelsubstanzen. So beschreibt das zeitlich erste Signalmaximum bei etwa 3 Sekunden die Substanz Aceton, nach 16 Sekunden folgt Toluol und das letzte lokale Maximum bei 48 Sekunden ist Xylol. Die drei registrierten Peaks sind in ihrer Gestalt unterschiedlich. Die aus der Form ableitbaren Parameter sind in der Abbildung 4.5 beschrieben. Allgemein folgt die Peakform der Eluenten einer

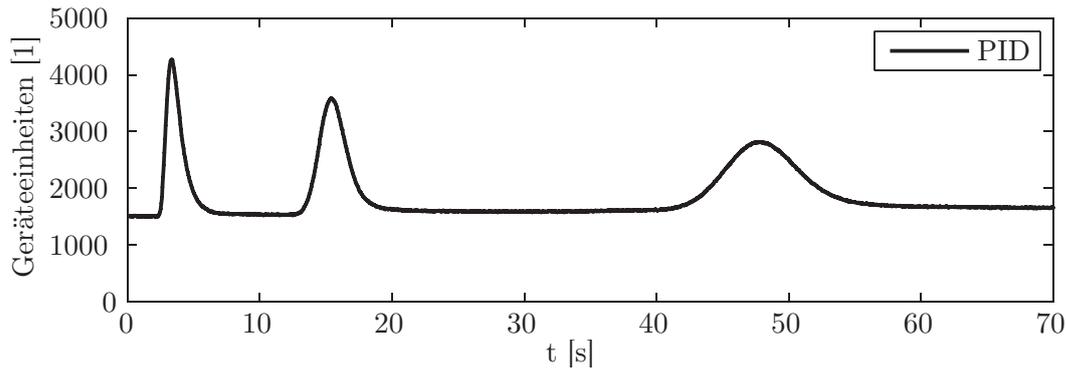


Abbildung 4.4.: Chromatogramm von Aceton, Toluol, Xylol

Poissonschen Verteilungskurve [84]. Für die Parameterbeschreibung wird eine symmetrische Kurvenform angenommen. In der Praxis können zudem *Fronting* und *Tailing* im Signalverlauf vorhanden sein. Diese Eigenschaften werden für die hier beschriebene Merkmalsbestimmung nicht weiter herangezogen. Für eine Auswertung werden die in der Abbildung markierten Merkmale des Peaks bestimmt. Das für eine Identifikation bestimmende Merkmal ist die Brutto-Retentionszeit  $t_R$ , die ab dem Injektionszeitpunkt  $t_i$  gemessen wird. Für eine Quantifikation der Stoffmenge kann die Höhe des Peaks oder die Fläche unter der Kurve, bis zur Basislinie herangezogen werden. Die Breite des Peaks wird für eine Berechnung der chromatographischen Auflösung herangezogen.

### 4.2.1. Filterung nach Savitzky-Golay

Zur Bestimmung der beschriebenen Merkmale ist ein differenzierbarer Signalverlauf erforderlich. Die nachfolgend beschriebenen Methoden zur Merkmalsbestimmung nutzen numerische Ableitungen des Detektorsignals. Daher ist es nötig, den Verlauf der realen Messdaten hinreichend zu glätten. Der gebräuchliche Filter des *gleitenden Mittelwertes* ist nur bedingt geeignet:

$$\bar{s}_k = \frac{1}{2m+1} \sum_{j=-m}^m s_{k+j}$$

Ein Fenster mit der Breite  $2m+1$  wird über das Rohsignal geschoben. Über den Umfang der durch das Fenster bestimmten Datenauswahl wird der Mittelwert berechnet und dem zentralen Punkt  $k$  zugewiesen. Für die Filterung von Chromatogrammen ist dieses

#### 4. Analyse des Signalverhaltens

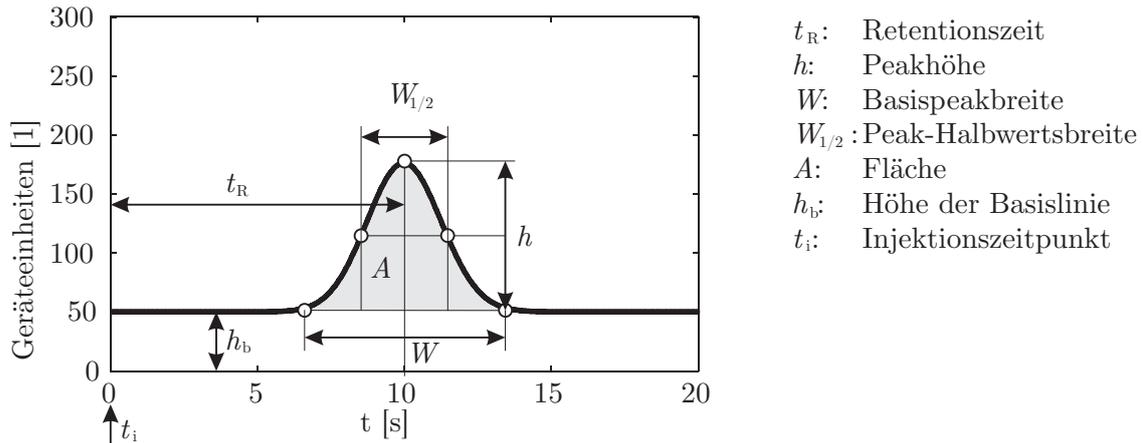


Abbildung 4.5.: Merkmale eines GC-Peaks

Filter jedoch nachteilig, da bei groß gewählter Fensterbreite, die Peaks geebnet werden. Eine Erweiterung des gleitenden Mittelwertes ist das Filter nach *Savitzky-Golay* [85] (Gleichung 4.2):

$$s_k^* = \sum_{j=-m}^m P_j \cdot s_{k+j} \quad (4.2)$$

Hier werden die Werte des Rohsignals innerhalb des Glättungsfensters vor der Mittelung mit einem Faktor  $P_j$  gewichtet. Diese Gewichtungsfaktoren werden aus einem Polynom  $n$ -ter Ordnung abgeleitet, welches über die Methode der kleinsten Quadrate ein Nutzsignal für die  $(2m + 1)$  gewählten Datenpunkte approximiert. Für symmetrische Glättungsfenster mit konstanter Abtastrate können die Gewichtungsfaktoren  $P_j$  als geschlossene Lösung mathematisch über Polynomgleichungen bestimmt werden [86]. Über folgende Gleichung werden die Gewichtungsfaktoren für eine Approximation zweiter bzw. dritter Ordnung berechnet:

$$P_j = \frac{3(3m^2 + 3m - 1 - 5j^2)}{(2m + 3)(2m + 1)(2m - 1)}$$

Der Parameter  $m$  gibt die Fensterbreite vor und  $j$  beschreibt den jeweiligen Index des Faktors für die Summationsgleichung 4.2.

Ein Vergleich der beiden Filtermethoden ist anhand eines praktischen Beispiels in Abbildung 4.6 gezeigt. Ein mit einem Kurzsäulen-GC aufgenommenes Chromatogramm mit drei Substanzpeaks ist in der linken Teilabbildung als Übersicht dargestellt. Die Spitze des zweiten Substanzpeaks ist in der rechten Teilabbildung detailliert zu sehen. Auf die

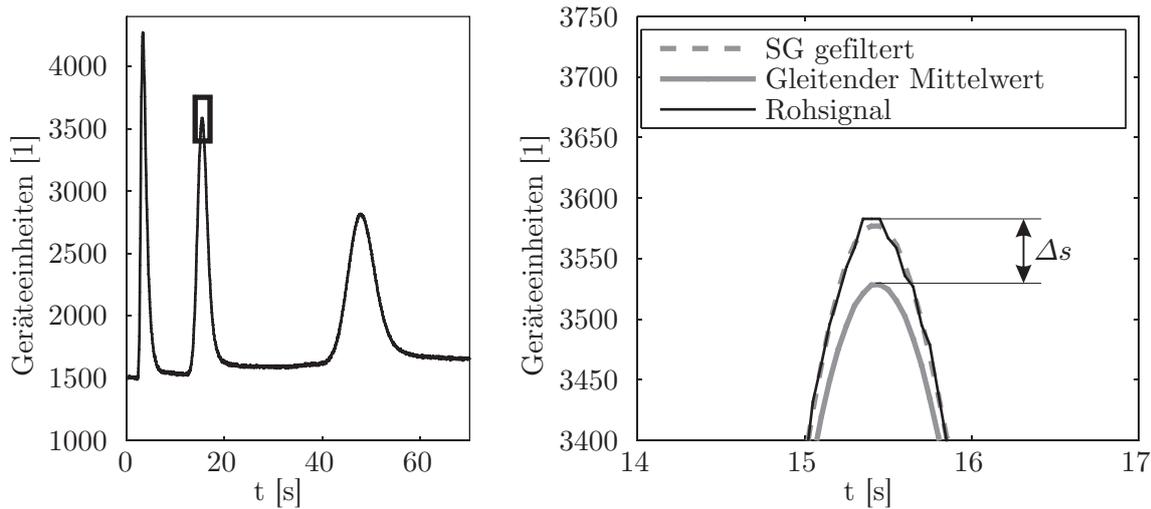


Abbildung 4.6.: Vergleich von Savitzky-Golay-Filter und gleitendem Mittelwert (Links: Überblick eines Beispiel-Chromatogramms; Rechts: Detail des zweiten Peaks)

rohen Sensorwerte wurden das Savitzky-Golay-Filter, sowie ein Tiefpassfilter in Form des gleitenden Mittelwertes angewendet. Für beide Methoden wurde eine Fensterbreite von 15 Messwerten, d.h. 1,5 Sekunden gewählt. Diese Auswahl basiert auf der Breite des schmalsten Peaks - die Fensterbreite sollte zwischen dem 0,7-fachen und dem doppelten der Peak-Halbwertsbreite ( $W_{1/2}$ ) liegen [86]. Aus der Abbildung ist zu entnehmen, dass beide Verfahren mit dieser Parametrierung gute Glättungsergebnisse für diese Anwendung ergeben. Wie jedoch ebenfalls zu erkennen ist, führt die Anwendung des gleitenden Mittelwertes zu einer Reduktion der Peakhöhe. Das Savitzky-Golay-Filter hingegen bewahrt die Peakform und -höhe des Signals.

Zur Anwendung des gewählten Filters ist eine Überabtastung des Signals erforderlich. Die Abtastrate von zehn Datenpunkten pro Sekunde erfasst die schmalsten Peaks mit mehr als zwanzig Datenpunkten und genügt so den Anforderungen der chromatographischen Auswertung [87].

Ein nach dem Savitzky-Golay-Filterverfahren geglättetes Signal kann nachfolgend einer Merkmalsextraktion unterworfen werden. Maßgeblich ist die Peaklage zur Bestimmung von Retentionszeit  $t_R$  und Peakhöhe  $h$ .

### 4.2.2. Peakerkennung bei Chromatogrammen

In der Praxis werden verschiedene Verfahren der Peakerkennung angewandt, zwei Beispiele sind die *klassische Methode* und die *Wendetangenten Methode*. In der implementierten Auswertesoftware wird das *klassische Verfahren* der Peakerkennung angewandt. Es wird die numerische erste Ableitung des Signalverlaufs gebildet. Ein Nulldurchgang der Ableitungsfunktion ist die notwendige Bedingung für ein lokales Extremum. Da

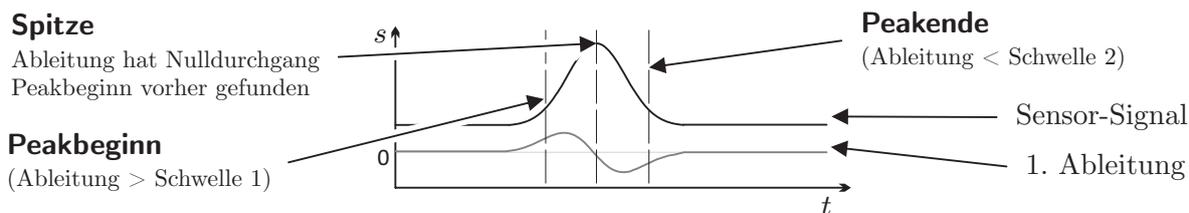


Abbildung 4.7.: Klassische Methode der Peakerkennung durch Bildung der ersten Ableitung des Signalverlaufs [54]

jedoch auch nach Anwendung von Filterfunktionen ein Rausch-Residuum im Signal verbleiben kann, werden Schwellwerte für die Flanken des Peaks eingeführt. In der Abbildung 4.7 sind ein idealer Signalverlauf und die zugehörige erste Ableitung gezeigt. Der Algorithmus durchläuft alle Datenpunkte des Signals  $s(t)$  und prüft dabei das Überschreiten eines positiven Schwellwertes  $\hat{s}$  der Ableitung  $\dot{s}(t)$ . Ist dies der Fall, so wird dieser Punkt als möglicher Peakbeginn gespeichert. Daraufhin werden die folgenden Stützstellen des Signals auf den Nulldurchgang der Ableitung geprüft und eine mögliche Spitze markiert. Ein Peakende wird ebenfalls über das Unterschreiten eines negativen Schwellwertes  $\check{s}$  von  $\dot{s}(t)$  gefunden. Erst wenn Beginn, Spitze und Ende zusammen vorliegen, so wird der Peak als gültig befunden. Die Retentionszeit und die Peakhöhe aus den Rohdaten von  $s(t)$  werden daraufhin für die Identifikation genutzt.

### 4.3. Merkmalsextraktion von IM-Spektren

Die Messdaten des Ionenmobilitätsspektrometers werden als zeitlicher Verlauf des Ionen-Detektorsignals dargestellt. Dieses Spektrum wird, ähnlich wie die beschriebenen Chromatogramme, auf Peaks und deren Driftzeit untersucht. Da es sich bei dem Signal

des Faraday-Detektors um die verstärkten Ströme einzelner auftreffender Ladungsträger handelt, wird starkes Signalrauschen am Ende der Messkette mitregistriert. Es werden daher zunächst Maßnahmen zur Verbesserung des Signal-zu-Rauschverhältnisses durchgeführt. Erst danach werden Driftzeiten von Substanz-Peaks bestimmt.

### 4.3.1. Signalaufbereitung

Nachfolgend werden Daten betrachtet, welche vom Ionenmobilitätsspektrometer des in Kapitel 3.2 beschriebenen Systems aufgenommen wurden. Da es sich um ein Doppelröhren-IMS handelt, werden sowohl das positive als auch das negative Spektrum gleichermaßen betrachtet. In Abbildung 4.8 wird ein Vergleich von einem ungefilterten und einem gefilterten Spektrum angestellt. Es handelt sich hierbei um zwei beispielhafte unabhängige *Nullmessungen*, d.h. es wurde keine Substanzprobe vermessen, sondern das sich ergebende Spektrum von gereinigter Luft aufgenommen. Das ungefil-

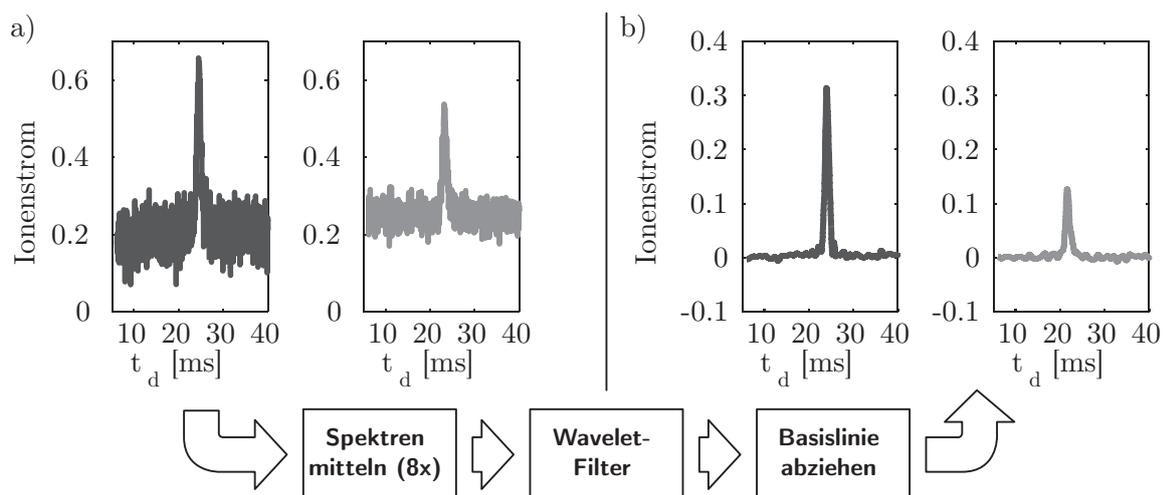


Abbildung 4.8.: Datenverarbeitung der Ionenmobilitätsspektren (ungefiltertes Spektrum in a) und gefiltertes Spektrum in b)), schwarzer Kurvenverlauf repräsentiert das positive Spektrum, grauer Verlauf das negative Spektrum

terte positive Spektrum in Abbildung 4.8 a) (schwarz eingefärbt) besitzt ein Signal-zu-Rauschverhältnis von 13 *dB*. Bei Anwendung einer Filterkette auf die Rohdaten des Ionenmobilitätsspektrometers erhält man Spektren, wie in Abbildung 4.8 b) gezeigt.

#### 4. Analyse des Signalverhaltens

Hier beträgt das Signal-zu-Rausch-Verhältnis  $28 \text{ dB}$ . Durchgeführte Maßnahmen sind: Die Mittelung mehrerer aufeinanderfolgender Spektren, Filterung des gemittelten Datensatzes mithilfe eines Wavelet-Filters und die Befreiung des Signalverlaufs von der Basislinie.

##### **Mittelung**

Die Aufnahme eines einzelnen Spektrums dauert 45 Millisekunden. Um auch chromatographisch aufgetrennte Stoffe mit dem Spektrometer erfassen zu können, ist es erforderlich eine Repetitionsrate von mindestens einem Spektrum pro Sekunde zu erreichen. Innerhalb dieses Zeitfensters ist es also möglich, Spektren überabzutasten. Zur Verbesserung des Signal-zu-Rauschverhältnisses wird die Möglichkeit der Überabtastung für eine Spektrenmittelung genutzt. Es werden dazu bis zu acht aufeinanderfolgende Spektren abgetastet und diese temporär gespeichert. Aus diesem Spektren-Satz wird daraufhin ein arithmetisches Mittel gebildet.

##### **Wavelet-Filterung**

Die Wavelet-Filterung ist ein Verfahren, das sich speziell für Signale mit zeitlich begrenzten Änderungen eignet. Dies liegt an der Eigenschaft der Wavelet-Transformation, Frequenz- und Zeitinformationen gleichzeitig aus dem Signale zu gewinnen [88]. Aus den Koeffizienten der Wavelet-Transformation lässt sich das Signal durch Rücktransformation wieder rekonstruieren. Der zur IMS-Spektrenfilterung verwendete Filteralgorithmus beinhaltet eine mehrstufige Kombination aus Wavelet-Transformation und diskretem Tief-/Hochpassfilter [89].

Im Gegensatz zum Savitzky-Golay-Filter benötigt der in dieser Weise implementierte Wavelet-Filter mehr Rechenzeit. Für die Auswertung von Ionenmobilitätsspektren führt dieses Verfahren jedoch zu wesentlich besseren Filterergebnissen als der Savitzky-Golay-Filter [11].

## Basislinie

Durch die linearen Komponenten der Verstärkerelektronik ergibt sich aufgrund von Bauteiltoleranzen ein Versatz der Detektorbasislinie. Dieser Basislinienoffset wird in der Signalverarbeitung nach den beschriebenen Schritten der Filterung vom Spektrum abgezogen. Von den spektralen Stützstellen werden die ersten 140 und die letzten 40 abgeschnitten, weil sich hier die Verzerrung des Injektionsimpulses sichtbar zeigt. Da am Beginn des Spektrums keine Produktions-Peaks zu erwarten sind, wird der Bereich zwischen den Abtastpunkten 140 und 180 in der Firmware als Grundlage zur Bestimmung der Basislinie genutzt. Aus diesen 40 Werten wird der arithmetische Mittelwert gebildet und dieser von jeder der 1024 Stützstellen abgezogen. [90]

### 4.3.2. Peakerkennung bei Ionenmobilitäts-Spektren

Ein in dieser Weise vorverarbeitetes IMS-Signal wird einer Peakerkennung unterworfen. Es wird die *klassische Methode* der Peak-Detektion genutzt. Der für diese Signalart angepasste Algorithmus ist in der Abbildung 4.9 als Flussdiagramm dargestellt. Die in

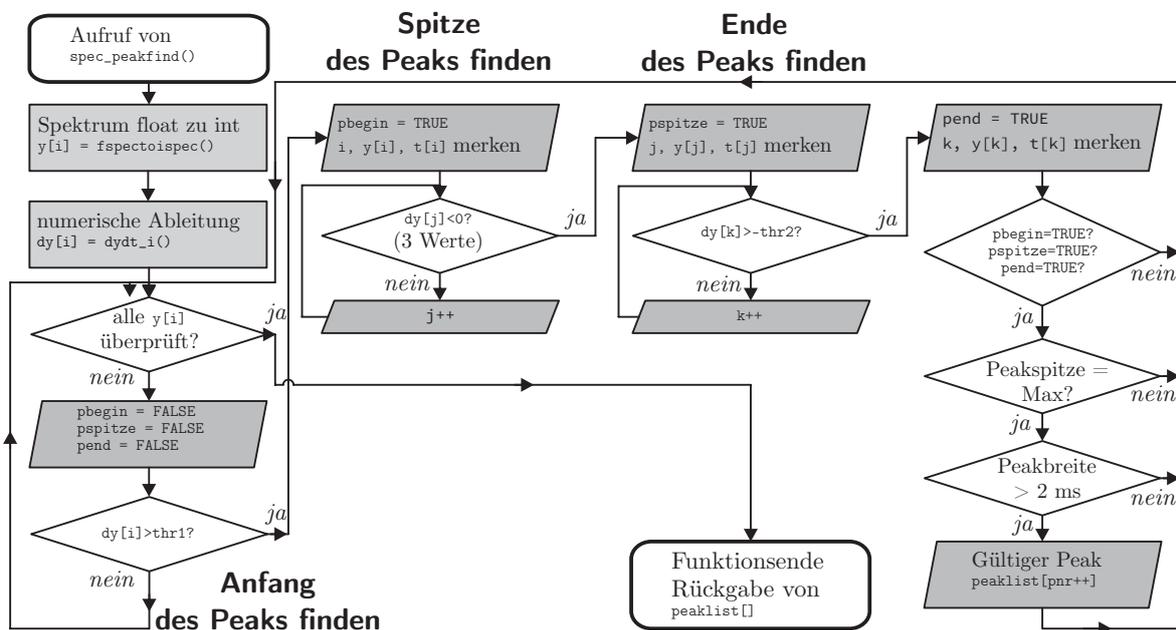


Abbildung 4.9.: Erkennung von Peaks im Ionenmobilitätsspektrum

#### 4. Analyse des Signalverhaltens

der Software als Datenfeld abgelegten Stützstellen des vorverarbeiteten Spektrums werden dem Algorithmus übergeben. Es wird die numerische Ableitung ( $dy[i]$ ) gebildet und diese dann sukzessive auf das Überschreiten eines vorher definierten Schwellwertes überprüft. Wie schon in Kapitel 4.2.2 beschrieben, werden über das Überschreiten einer Schwelle, den Nulldurchgang und das Unterschreiten einer Schwelle die drei Eigenschaften *Peak-Beginn*, *Peak-Spitze* und *Peak-Ende* bestimmt. Wenn alle drei Elemente vorhanden sind, wird zusätzlich kontrolliert, ob es sich bei der vermeintlichen Spitze um den Maximalwert des Eingangsdatensatzes ( $y[i]$ ) innerhalb des untersuchten Intervalls handelt. Eine letzte Plausibilitätskontrolle prüft die Fußpunktbreite des Peaks, um so lokale Stör-Transienten auszuschließen. Das Ergebnis der Peakssuche ist der relative Zeiger auf das Datenfeld der Peakspitze. Der Index wird in einer wachsenden Ergebnisliste abgelegt, die am Ende der Suche von der Subroutine zurückgegeben wird. Die interessierende Driftzeit in Millisekunden wird aus dem Ergebnis-Index linear errechnet, da es sich um äquidistante Abtastpunkte handelt. Die Driftzeiten und zugehörigen Peakhöhen werden der  $K_0$ -Wert-Bestimmung unterworfen.

#### 4.3.3. $K_0$ -Wert-Bestimmung

Für eine Vergleichbarkeit der Driftzeiten von Substanzen in Ionenmobilitätsspektren ist eine Umrechnung, wie sie im Grundlagenkapitel 2.2.1 beschrieben wurde, nötig. Für den  $K_0$ -Wert werden Geräteparameter wie die Driftraumgeometrie und das elektrische Feld berücksichtigt, sowie die Zustandsgrößen des Gases in Form von Druck und Temperatur. In Abbildung 4.10 ist die Zusammensetzung der Ionenmobilität grafisch dargelegt. Hier

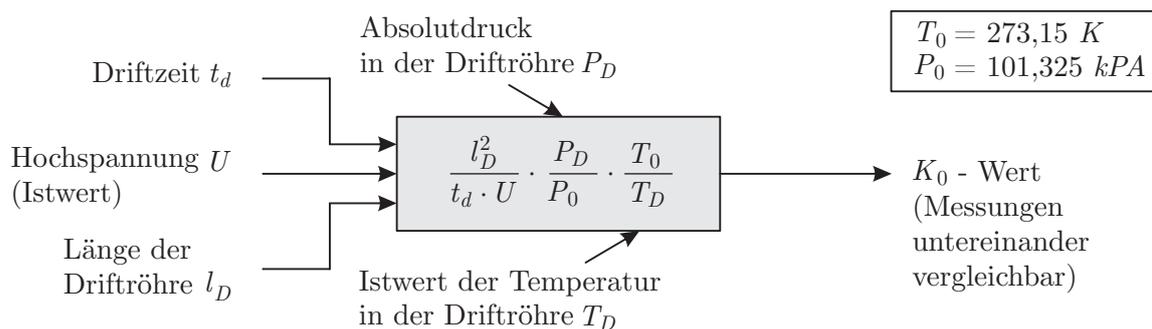


Abbildung 4.10.: Zusammensetzung des Wertes der reduzierten Ionenmobilität  $K_0$

handelt es sich nur bei der Driftraumlänge  $l_D$  um einen konstanten Parameter. Die

## 4. Analyse des Signalverhaltens

Parameter der Driftzeit  $t_d$ , Driftraumspannung  $U$ , Absolutdruck der Driftstrecke  $P_D$  und Temperatur der Driftstrecke  $T_D$  sind Momentanwerte. Diese Istwerte werden vom System kontinuierlich gemessen. So werden Druck und Temperatur von in die Driftröhrenfluidik eingebrachten Sensoren erfasst. Die Driftröhrenspannung wird am unteren Ende der Widerstandskaskade gemessen und über den sich ergebenden Spannungsteiler auf die gesamte Driftstrecke hochgerechnet. Jeder dieser Parameter kann fehlerbehaftet sein und beeinflusst den  $K_0$ -Wert über das Fehlerfortpflanzungsgesetz.

Eine Substanz wird im Spektrum durch die Ausprägung eines Peaks repräsentiert. Dieser beruht dann auf den gebildeten Produktionen der Reaktionskette in Gleichung 2.28 folgend. Bei dem Ionisationsvorgang können bei hohen Konzentrationen jedoch auch Ionen-Cluster in Form von Dimeren gebildet werden [11]. Bei Betrachtung des Spektrums resultiert dieses Verhalten in der Ausprägung eines weiteren Produktionen-Peaks. Für die Ergebnisausgabe werden die drei Ionenmobilitäten  $K_0$  des Spektrums mit der höchsten Intensität ausgewählt. Bei geringen Konzentrationen enthält die Ergebnisliste dabei neben den Produkt- $K_0$ -Werten ebenfalls den  $K_0$ -Wert der Reaktantionen.

### 4.4. Eingangsmerkmale des maschinellen Lernalers

In den letzten drei Kapiteln 4.1 - 4.3 wurde beschrieben, wie verschiedene Merkmale aus den Sensoren, der Gaschromatographie und der Spektrometrie extrahiert werden. Bei einem Kombinationsinstrument, wie dem im Projekt *DACHS* realisierten Funktionsmuster, werden diese drei Merkmalsgruppen vereint. Das beschriebene System unterscheidet jedoch zwischen zwei Anwendungsfällen (siehe Kapitel 3.2). Die *direkte Messung* ermöglicht eine kontinuierliche Probenanalyse, wohingegen die *Messung mit GC-Vortrennung* diskontinuierlich Komponenten von Gas-Gemischen diskriminieren kann. Insgesamt stehen für die Daten-Interpretation sechs Merkmale von Einzelsensoren in Form von relativen Sensorantworten zur Verfügung. Hinzu kommen aus der Peakerkennung der IM-Spektren (Kapitel 4.3) die reduzierten Ionenmobilitäten  $K_0$  für positive und negative Ionenspezien. Es werden aus beiden Spektren die drei höchsten Peaks ausgewählt. Der Sonderfall der GC-Vortrennung berücksichtigt zusätzlich eine stoffspezifische Retentionszeit. Diese Zeit ist abhängig von der gewählten *Methode* im Sinne der



#### 4. Analyse des Signalverhaltens

Gaschromatographie. Eine solche Methode wird definiert über die Parameter Injektionsdauer, Temperatur der Trennstrecke und Volumenstrom durch die Strecke. In einer Datenbank können Retentionszeiten für drei GC-Methoden zur Analyse eines Stoffes hinterlegt werden. In der Tabelle 4.2 ist eine Übersicht der für eine Stoffklassifikation

|                 | 6 relative Sensorwerte |     |     |     |     |     | 6 Ionenmobilitäten |      |      |      |      |      | Retentionszeiten |           |           |
|-----------------|------------------------|-----|-----|-----|-----|-----|--------------------|------|------|------|------|------|------------------|-----------|-----------|
|                 | HL1                    | HL2 | HL3 | EC1 | EC2 | PID | K01+               | K02+ | K03+ | K01- | K02- | K03- | $t_{R,1}$        | $t_{R,2}$ | $t_{R,3}$ |
| direkte Messung | x                      | x   | x   | x   | x   | x   | x                  | x    | x    | x    | x    | x    |                  |           |           |
| GC-Messung      | x                      | x   | x   | x   | x   | x   | x                  | x    | x    | x    | x    | x    | x                | x         | x         |

Tabelle 4.2.: Extrahierte Merkmale nach gewählter Messmethode eines Multisensor-Arrays mit GC

eingesetzten Merkmale gegeben. Der Eingangsraum der Datenauswertung für eine direkte Messung besitzt also insgesamt 12 Größen, dies entspricht der Dimensionalität des Merkmalsraums. Eine GC-Analyse erweitert den Merkmalsraum um die Größe der Retentionszeit für die gewählte GC-Methode.

### 4.5. Signaldarstellung für Vor-Ort-Interpretation durch Expertennutzer

Der Mensch besitzt die Fähigkeit durch die visuelle Wahrnehmung Muster zu erkennen und hierdurch beispielsweise Objekte zu identifizieren. An diese Interpretationsfähigkeit durch einen geschulten System-Benutzer knüpft eine Darstellungsmethode des Messsystems DACHS an. In der Arbeit von Hunte [9] wurde bereits die Darstellung von Sensorsignalen eines Gefahrstoff-Detektoren-Arrays durch ein Balkendiagramm vorgestellt. Die relativen Sensorwerte der Einzelsensoren werden hierbei normiert durch die Länge eines Diagramm-Balkens veranschaulicht. Die Ionenmobilitätsspektren werden jeweils in zwei Quadranten aufgeteilt. Die Lage des Reaktantenpeaks dient dabei als Bereichs-Grenze. Das Integral der Signalstärke innerhalb der Quadranten wird dann ebenfalls normiert als Diagramm-Balken ausgegeben. Die Bestimmung der Integrations-

#### 4. Analyse des Signalverhaltens

grenze geschieht durch eine Driftzeitbestimmung der Reaktantionenpeaks während einer Nullmessung.

Um die gefundene Peakmitte wird ein Fenster mit einer festen Breite gelegt. Dieses wird durch die Driftzeiten  $t_{d,wl}$  für die linke Fenstergrenze und  $t_{d,wr}$  für die rechte Grenze bestimmt. Außerhalb des Fensters befinden sich Integrationsbereiche, über welche die Signalintensität jedes Abtastpunktes des Spektrums aufsummiert wird (in nachfolgenden Gleichungen 4.3- 4.4 mit Indizierung für das Spektrum der positiven Ionen).

$$I_1^+ = \int_0^{t_{d,wl}} i_{\text{pos}} dt_d, \quad I_r^+ = \int_{t_{d,wr}}^{\infty} i_{\text{pos}} dt_d, \quad I_{\text{RI}}^+ = \int_{t_{d,wr}}^{t_{d,wl}} i_{\text{pos}} dt_d \quad (4.3)$$

Für die Normierung wird zusätzlich das Integral über das Fenster selbst berechnet und repräsentiert die Intensität der Reaktantionenpeaks.

$$I_{\text{rel l}}^+ = \frac{I_1^+}{I_{\text{RI}}^+} \quad I_{\text{rel r}}^+ = \frac{I_r^+}{I_{\text{RI}}^+} \quad (4.4)$$

Der Relativwert für eine Balkendarstellung errechnet sich als Quotient zwischen dem jeweiligen Quadranten und der Intensität des Reaktantions.

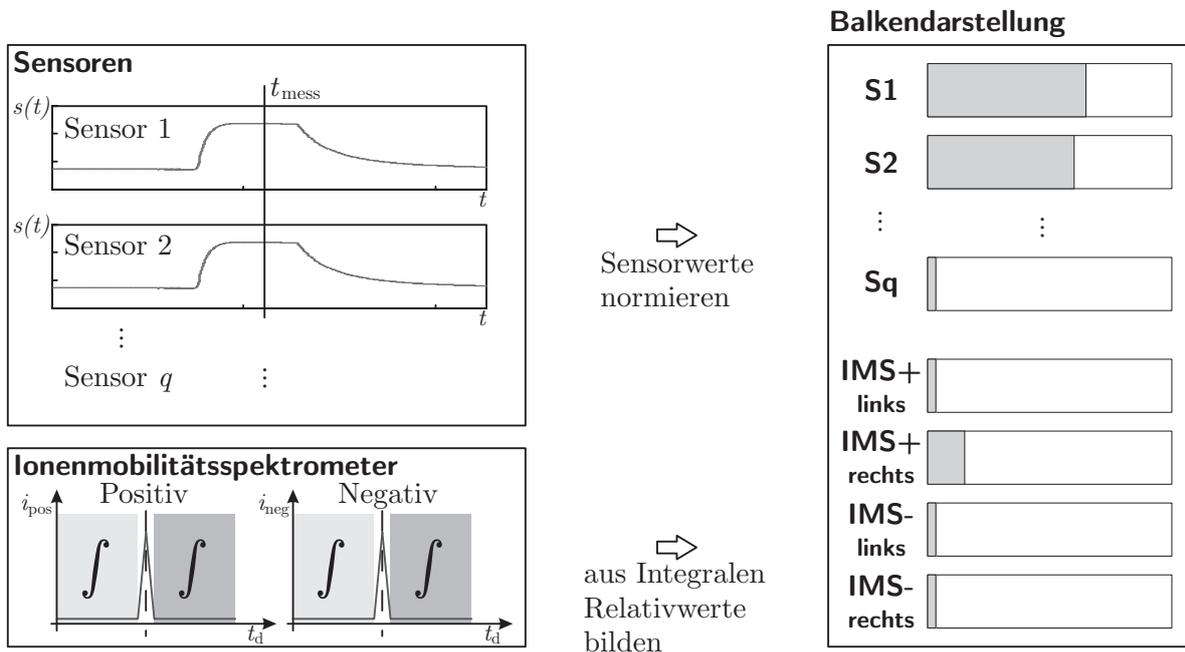


Abbildung 4.11.: Balkendarstellung der Messwerte zur visuellen Interpretation

#### 4. Analyse des Signalverhaltens

In der Abbildung 4.11 ist die Zusammensetzung der grafischen Signaldarstellung veranschaulicht. Die Anzeige wird zu jedem Zeitpunkt  $t_{\text{mess}}$  neu berechnet. Ein erfahrener Benutzer kann diese Darstellung interpretieren und eine Stoffgruppe identifizieren. Mithilfe einer Entscheidungstafel wird je nach qualitativer Anzeige der Messkanäle die Probe klassifiziert [9]. Für das Messsystem DACHS wurde diese Art der Anzeige angepasst. Die drei verwendeten Halbleitersensoren haben für die visuelle Interpretation nur geringen zusätzlichen Identifikationswert. Diese Sensorwerte werden für eine Darstellung mit einer logischen *Oder*-Verknüpfung als ein einzelner Kanal zur Anzeige gebracht.

# 5. Implementierung einer Datenbasis zur Identifikation von Substanzen

Für den Aufbau einer Datenbasis zur automatischen Identifikation von Substanzen wurden die Zielstoffe im Labor vermessen. Aufgenommen wurden möglichst unverfälschte Rohdaten der im System integrierten Sensoren. Diese sind dann einer rechnergestützten Mustererkennung unterworfen worden, aus welcher sich die Klassifikationsregeln ableiten. Diese Regeln wurden nachfolgend auf das Mikroprozessorsystem des entwickelten Messinstrumentes implementiert. Nachfolgende Kapitel beschreiben die Schritte von der Messung bis zur Evaluation der Klassifikationsmethode.

## 5.1. Messdatenaufnahme

In Zusammenarbeit mit dem chemischen Labor des Bundesamtes für Bevölkerungsschutz und Katastrophenhilfe wurden mehrere Substanzen durch die im Rahmen des DACHS-Projektes realisierten Funktionsmuster messtechnisch erfasst. Exemplarisch werden nachfolgend die Messdaten dieser Substanzen betrachtet:

- Toluol
- Aceton
- Essigsäure
- Chloroform

Diese Substanzen wurden jeweils unter unterschiedlichen Bedingungen gemessen. Es ergeben sich Variationen des Signalmusters durch unterschiedliche Stoffkonzentrationen. Hohe Konzentrationen werden vom Messsystem durch die Verdünnungseinheit (siehe

## 5. Implementierung einer Datenbasis zur Identifikation von Substanzen

Kapitel 3.2) auf eine *scheinbare Konzentration* reduziert. Zudem hat die Feuchtigkeit des Trägergases einen Einfluss auf die Nachweisfähigkeit einiger Sensoren [91]. Aufgrund dieser Einflussfaktoren ergibt sich ein umfangreicher Datensatz mit Trainingsdaten. Die

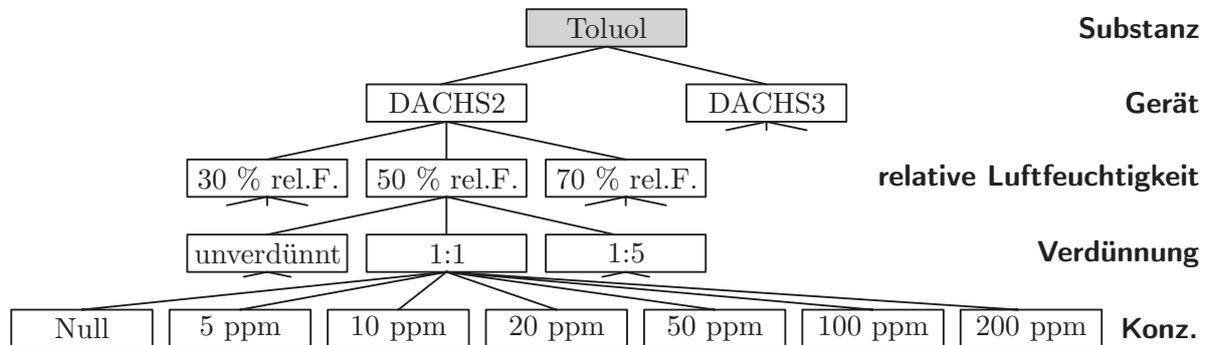


Abbildung 5.1.: Messvarianten zum Aufbau einer Stoffdatenbank für zwei Funktionsmuster des DACHS

Abbildung 5.1 stellt die verschiedenen Messvarianten grafisch als Baumdiagramm dar. Wird ein Zielstoff von beiden Systemen in den gezeigten Varianten vermessen, so ergeben sich pro Substanz 63 einzelne Messungen allein für das Verfahren der *direkten Messung* ohne chromatographische Vortrennung. Dazu kommen Wiederholungsmessungen um eine langzeitliche Sensordrift in den Messdaten zu berücksichtigen.

Durch diese Vielzahl an Varianten kann eine gute Generalisierung des Mustererkennungsverfahrens erreicht werden.

### 5.1.1. Messaufbau

Zur Erzeugung einer definierten Stoffkonzentration werden die in flüssiger Phase vorliegenden Substanzen verdampft und mit konditionierter Laborluft gemischt. Der Messaufbau ist in Abbildung 5.2 schematisch gezeigt. Der Aufbau teilt sich in einen Bereich zur Probengenerierung und einen zur Messdatenerfassung. Das Trägermedium wird als getrocknete und gereinigte Labordruckluft bereitgestellt. Ein Befeuchtungsgerät regelt den Feuchtigkeitsgehalt der Verdünnungsluft und stellt so die drei gewünschten Feuchtestufen 30 %, 50 % und 70 % ein. Ein Dosiergerät (HovaCAL) fördert einen definierten Volumenstrom der flüssigen Probe mithilfe einer Push-Pull-Spritzenpumpe zu einem beheizten Verdampfer. Zwei, über elektrische Massendurchflussregler eingestellte

## 5. Implementierung einer Datenbasis zur Identifikation von Substanzen

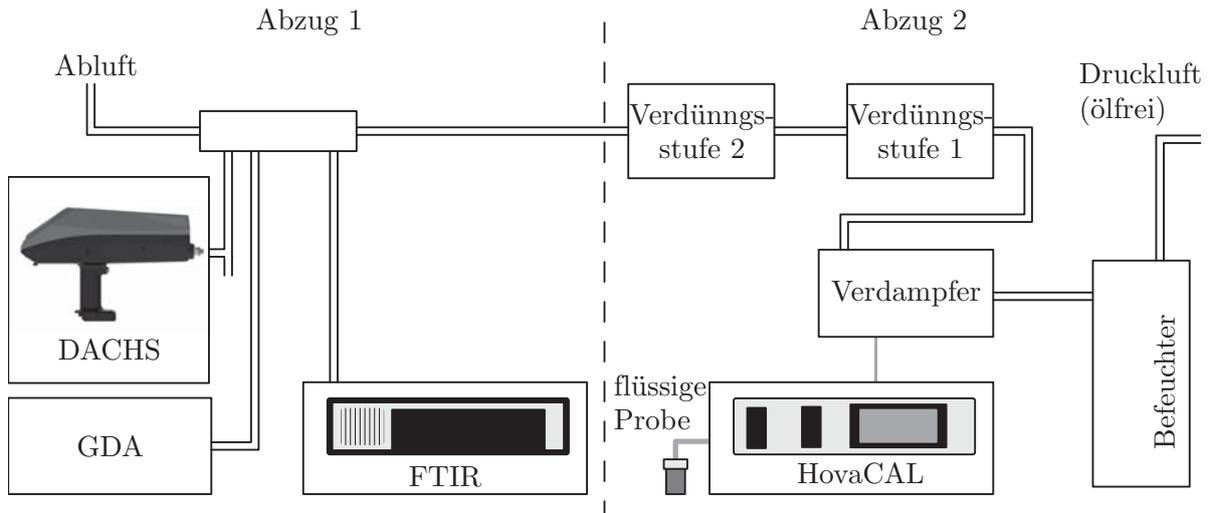


Abbildung 5.2.: Schematische Darstellung des Messaufbaus

Verdünnungsstufen, stellen die schließlich gewünschte Stoffkonzentration anhand der beigemischten Inertgasmenge ein.

Der Gasfluss wird auf die Messinstrumente verteilt. Überschüssiges Probegas wird abgeleitet. Ein Fourier-Transform-Infrarot-Spektrometer dient als quantifizierendes Referenzgerät. Ein zweiter Abgang wird zum Einlass des DACHS-Systems geführt. Die Messdaten werden sowohl auf einer im DACHS befindlichen SD-Karte abgelegt, als auch von einem Messrechner aufgezeichnet. Über diesen erfolgt auch die Steuerung der Verdünnungsstufen des DACHS.

Diese Art der Probenbereitung bietet nur begrenzte Möglichkeiten die Konzentration zu steuern. Durch die direkte Verdampfung können nur relativ hohe Konzentration im Bereich von mehreren *ppm* generiert werden. Die Erzeugung von Konzentrationen darunter ist mit diesem Aufbau nicht möglich. Für die Untersuchung von Nachweisgrenzen des Sensorsystems wurde daher ein anderes Verfahren gewählt.

Für den Aufbau von Datenbanken zur Stoffidentifikation ist diese Methode jedoch ausreichend.

### 5.1.2. Kalibriergasgenerator mit Permeationsgefäßen

Mit diesem System können auch sehr geringer Substanzkonzentrationen erzeugt werden. Es werden hierfür Substanzen als flüssige Probe in, mit einer Permeationsbarriere ausgestatteten, Probenbehältern bereitgehalten. Im isothermen Zustand der Probe und bei gleichbleibendem Konzentrationsgradienten über der Permeationsbarriere stellt sich ein konstanter Massenfluss durch das Permeationsmaterial ein. Zusammen mit einem konstanten Massenfluss von Nullluft (proportional zum Volumenstrom bei konstanten Bedingungen), der die Permeationskammer durchspült, ergibt sich eine definierte Stoffkonzentration.

In Abbildung 5.3 ist der schematische Aufbau des Kalibriergasgenerators veranschaulicht. Die als Trägergas genutzte Nullluft wird über eine schwingungsarme Drehschieber-

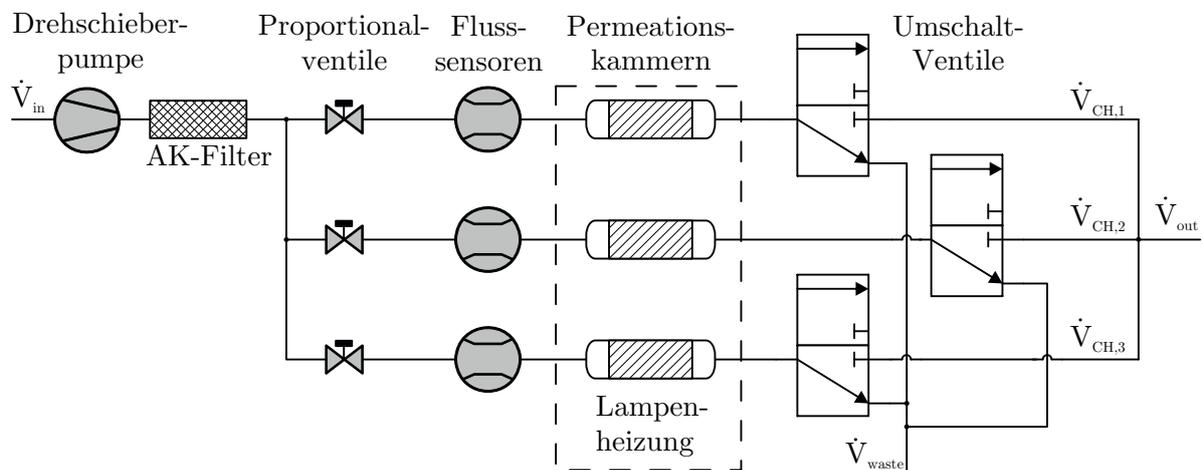


Abbildung 5.3.: Prinzipschaltbild des Kalibriergasgenerators

pumpe und ein mit Aktivkohle gefülltes Molsieb bereitgestellt. Mithilfe von elektrisch angesteuerten Proportionalventilen und Durchflusssensoren wird der Volumenstrom geregelt. Der auf drei Permeationskammern aufgeteilte Trägergasstrom mischt sich kontinuierlich mit dem Massenstrom der Prüfsubstanzen.

Der die Permeationsbarriere der Probenbehälter durchdringende Massenstrom wird vor der Applikation im Kalibriergasgenerator gravimetrisch bestimmt. Die erzeugte Stoffkonzentration kann in diesem Aufbau entweder als Kalibriergas am Ausgang zur Verfügung gestellt werden oder wird abgeleitet.

Dieses Verfahren wird zur Bestimmung von Nachweisgrenzen und zur Kalibrierung genutzt.

### 5.1.3. Datenformat

Das Instrument nimmt die vorverarbeiteten Daten auf und speichert sie in einem bestimmten Datenformat auf der integrierten SD-Karte und auf dem Messrechner. Es werden pro Messlauf fünf zusammengehörige Dateien erzeugt. In allen Dateien wird ein bestimmter Teil der Messwerte in Form von kommaseparierten Listen gespeichert. Eine Zeile innerhalb der Messdateien enthält alle Werte zu einem Zeitpunkt der Datenaufnahme. Zur Synchronisation der Messwerte enthält jede Zeile zudem einen Zeitstempel und eine fortlaufende Nummer. Im Einzelnen werden Dateien mit diesen Informationsgehalten erzeugt:

- Endung `_raw.dachs`: Messwerte der Einzelsensoren, Systemparameter wie Temperaturen, Drücke, Spannungen, ausgewertete Peaks der IM-Spektren.
- Endung `.txt`: Messwerte und Parameter, jedoch zeitlich reduzierte Abtastung.
- Endung `_IMS_pos.dachs`: Abgetastetes, gefiltertes Spektrum positiver Ionen.
- Endung `_IMS_neg.dachs`: Abgetastetes, gefiltertes Spektrum negativer Ionen.
- Endung `_Orders.dachs`: Ausgeführte Systemzustandsänderungen (Verdünnung, Anreichern, Online-Messung).

Diese Daten werden für den Aufbau der Klassifizierungsregeln von der Numerik-Software Matlab eingelesen und skriptbasiert ausgewertet.

## 5.2. Maschinelles Lernen mit Hilfe realer Messdaten

Wie bereits in Abbildung 5.1 auf Seite 86 gezeigt, ergeben sich mehrere verschiedene Varianten einer Substanzmessung. Zusammen mit Wiederholungsmessungen und der Vielzahl an Abtastpunkten innerhalb eines Messlaufs ergibt sich ein umfangreicher Datensatz für das Training eines Klassifikators. Das Ziel ist die Identifikation der Stoffe mit Hilfe der extrahierbaren Merkmale.

## 5. Implementierung einer Datenbasis zur Identifikation von Substanzen

Damit ein Verfahren des maschinellen Lernens bewertet werden kann, sind neben den Messdaten, welche für den Aufbau der Diskriminanzfunktionen benötigt werden, zusätzlich Testdaten nötig. Bei dem vorliegenden Umfang von Messungen wurde dies durch eine zufällige Aufteilung der, in mehreren Messkampagnen aufgenommenen Daten, in einen Trainingsdatensatz und einen Testdatensatz erreicht.

### 5.2.1. Überwachtes Lernen mit einem Entscheidungsbaum nach CART

Eine untersuchte Methode zur Klassifikation von Merkmalsvektoren welche mit dem entwickelten Detektorenarray erzeugt wurden, ist die des CART-Entscheidungsbaums nach Breiman [57]. Zur Generierung des Entscheidungsbaumes wurde die Implementierung der Numerik-Software Matlab genutzt. Diese baut vollständig auf dem CART-Algorithmus auf, welcher in Kapitel 2.4.2 dargelegt wurde. Mit Hilfe der aufgenommenen Trainings- und Testdaten werden Entscheidungsbäume nach diesem Verfahren induziert und evaluiert.

In den Grundlagen wurden bereits zwei Bestimmungsmaße für die enthaltene Unreinheit eines möglichen Splits beschrieben: Die Entropie und der Gini-Index (siehe Gleichung 2.48 und folgend). Diese sind in der verwendeten Implementierung ebenfalls Auswahlparameter. Ein weiterer Einflussfaktor des Verhaltens des Entscheidungsbaums ist der Parameter zur Festlegung der minimalen Blattgröße. Dieser bestimmt, wie viele Messwerte mindestens zu einem Ergebnisknoten zusammengefasst werden. Ist dieser Parameter zum Beispiel mit dem Wert Eins belegt, so wird aus jedem der Merkmalsvektoren aus  $\mathcal{L}$  auch ein Blattknoten mit einem Klassifikationsergebnis gebildet. Anderenfalls werden mehrere Blattknoten beim *Prepruning* zusammengefasst.

Zur Parameterfindung werden aus den Trainingsdaten Klassifikationsbäume mit variierendem Induktionsparameter der minimalen Blattgröße generiert. Darauffolgend werden die induzierten Bäume durch zweifache Kreuzvalidierung ausgewertet und die Korrekturklassifikationsrate betrachtet. In der Abbildung 5.4 wurde die Gesamtheit der Trainingsdaten aus dem Abschnitt 5.1 in dieser Weise ausgewertet. Die beiden Kurven zeigen die

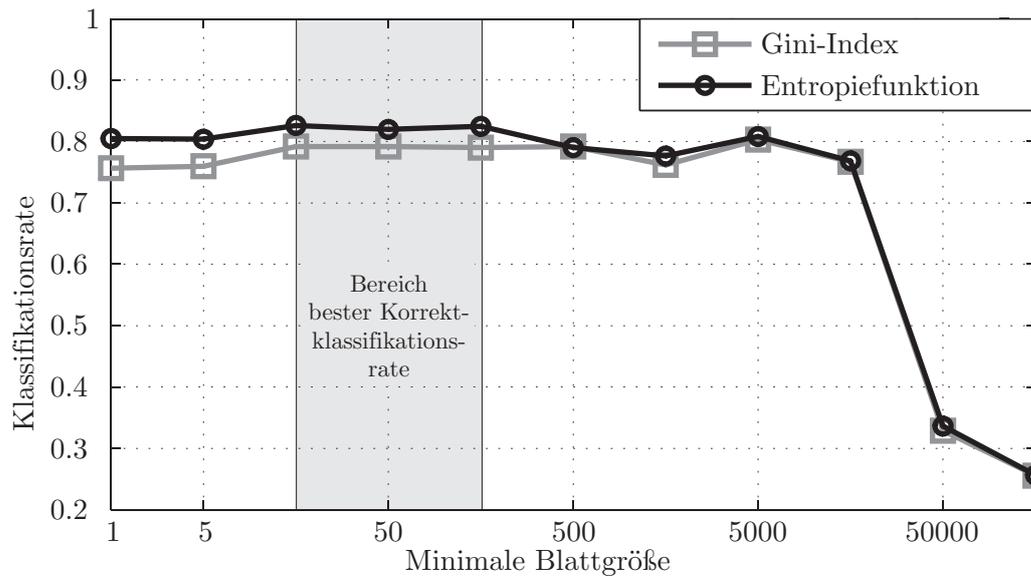


Abbildung 5.4.: Vergleich von zwei Funktionen der Trenngrenzenbestimmung bei der Bauminduktion [55]

Klassifikationsraten abhängig vom Parameter der minimalen Blattgröße. Die Klassifikationsbäume wurden einmal unter Zuhilfenahme des Unreinheitsmaßes der Entropie und einmal des Gini-Indexes induziert.

Es ist zu erkennen, dass die Entropiefunktion als Unreinheitsmaß in diesem Fall insgesamt eine leicht bessere Korrektklassifikationsrate erreicht. In Bezug auf den Parameter der minimalen Blattgröße ist ein Optimum der Klassifikationsrate im Bereich zwischen 10 und 100 zusammengefassten Trainingsinstanzen auszumachen (grau unterlegt).

Neben der Klassifikationsrate wurde zudem die Komplexität der induzierten Bäume untersucht. Ein auf den 40000 Instanzen von Trainingsdaten basierender Entscheidungsbaum bei dem keine Blätter zusammengefasst werden, enthält circa 400 bzw. 300 Entscheidungsknoten bei der Nutzung des Gini-Indexes bzw. der Entropiefunktion. Die Komplexität ist entscheidend für die Rechenzeit eines Klassifikationslaufs. Auch hier ist die Entropiefunktion aufgrund der sich ergebenden geringeren Baumkomplexität zu bevorzugen.

### 5.2.2. Überwachtes Lernen mit Random Forest

Die zweite untersuchte Methode zur Stoffidentifikation ist die Random-Forest-Klassifikation. Der Aufbau der Regeln findet ebenfalls in Form des *Überwachten Lernens* anhand ausgewählter Daten statt. Der rechenintensive Aufbau der Klassifikationsregeln erfolgt daher PC-basiert. In Abbildung 5.5 sind die Schritte dieser Methode dargestellt. Die

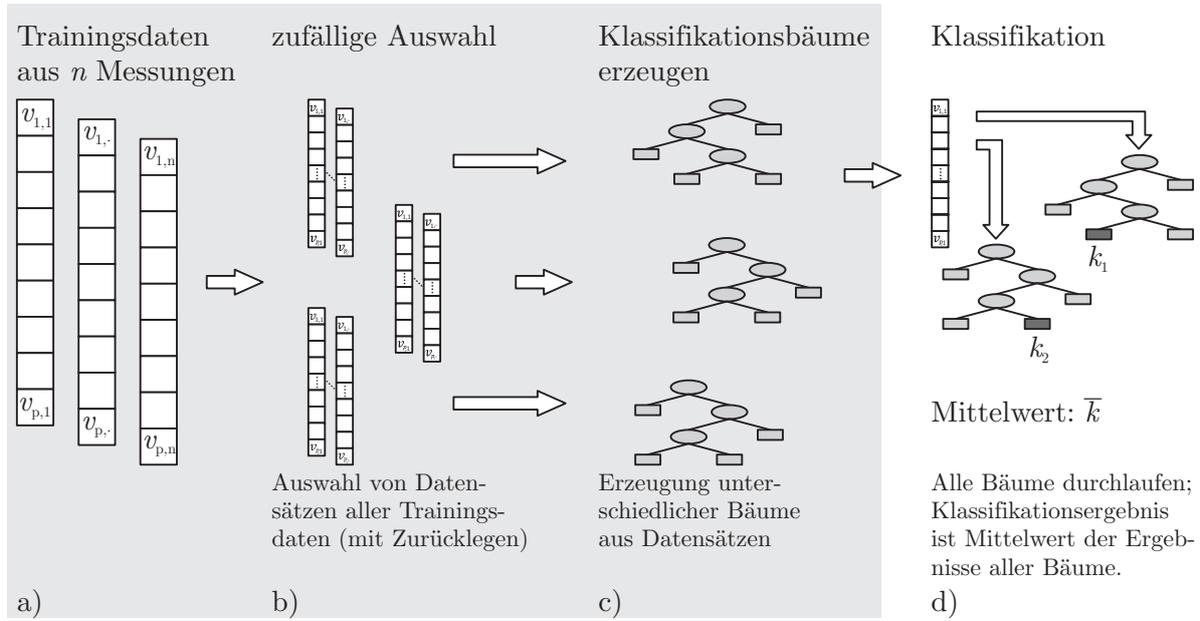


Abbildung 5.5.: Ablauf des Random-Forest Klassifikationsprozesses

Schritte des Lernens, welche offline anhand aufgenommener Messdaten erfolgen, sind hier grau unterlegt. Die Klassifikation geschieht auf dem eingesetzten eingebetteten System (Abbildung 5.5 d)).

Bei den Trainingsdaten der Teilabbildung a) handelt es sich um Merkmalsvektoren mit der in Tabelle 4.2 gezeigten Ausprägung. Für die direkte Messung hat dieser also die Dimensionalität  $\dim(v_{.,n}) = p = 12$ .

#### Genutztes Softwarepaket zum Bootstrap Aggregating

Wie in der Abbildung 5.5 b) zu sehen, nutzt die Random-Forest-Methode mehrere zufällig ausgewählte Untermengen des Trainingssets ( $\mathcal{L}^{(B)} \subset \mathcal{L}$ ) zur Erzeugung einzelner Entscheidungsbäume. Dieses *Bagging*-Verfahren nach Breiman [63] (siehe Kapitel 2.4.3) wird

## 5. Implementierung einer Datenbasis zur Identifikation von Substanzen

in der verwendeten Numerik-Software Matlab durch die Klasse `TreeBagger` bereitgestellt. Der Algorithmus wird maßgeblich über zwei Angaben parametrisiert:

- minimale Blattgröße
- Anzahl der zu generierenden Bauminstanzen des Ensembles

Wie schon im vorhergehenden Abschnitt erläutert, bestimmt der Parameter der minimalen Blattgröße, wie viele Trainingsvektoren mindestens zu einem Ergebnisblatt zusammengefasst werden.

Die Anzahl der Bauminstanzen gibt an, wie viele Lernsets ( $\mathcal{L}^{(B)}$ ) aus der Gesamtheit der Trainingsdaten ( $\mathcal{L}$ ) gebildet werden. Aus jedem der ( $B$ ) Datensätze wird ein Entscheidungsbaum für das Ensemble generiert. Je größer die Anzahl der Bauminstanzen in dem *Forest* ist, desto stärker konvergiert der Klassifikationsfehler bezogen auf die Trainingsdaten gegen Null. Gleichzeitig steigen jedoch auch die benötigte Rechenleistung und der Speicherbedarf.

Ausgehend von diesen beiden Parametern muss ein guter Kompromiss zwischen guter Anpassung an die Trainingsdaten und Generalisierungsfähigkeit der Klassifizierungsmethode gefunden werden. Zusätzlich ist die verfügbare Rechen- und Speicherleistung des eingebetteten Zielsystems eine entscheidende Randbedingung.

In Abbildung 5.6 ist ein Diagramm der Korrekt-Klassifikationsrate abhängig von diesen beiden Parametern dargestellt. Als Grundlage für diese Auswertung dient der in Abschnitt 5.1 dargelegte Messdatenumfang. Aus diesem Datensatz wurden für 48 Kombinationen der beiden Parameter, Random Forest Klassifikatoren gebildet. Diese wurden dann mit einem Testdatensatz bekannter Klasse untersucht. Betrachtet man zunächst den Parameter der minimalen Blattgröße, so fällt auf, dass bei Werten bis zu 5000 die Korrekt-Klassifikationsrate über 95 % liegt. Wenn erst eine Menge von mehr als 5000 Merkmalsvektoren innerhalb eines Bereiches zu einem Klassifikationsblatt führt, neigt das Klassifikations-Ensemble zur Unteranpassung. Es werden nur noch ca. 40 % des Testdatensatzes korrekt klassifiziert.

Die zweite Achse in der Abszissen-Ebene gibt die Anzahl der Baum-Instanzen des Random Forest an. Es ist der Grafik zu entnehmen, dass die Klassifikationsrate ab der Anzahl von circa 30 Bäumen im Klassifikationsmodell nicht mehr signifikant steigt.

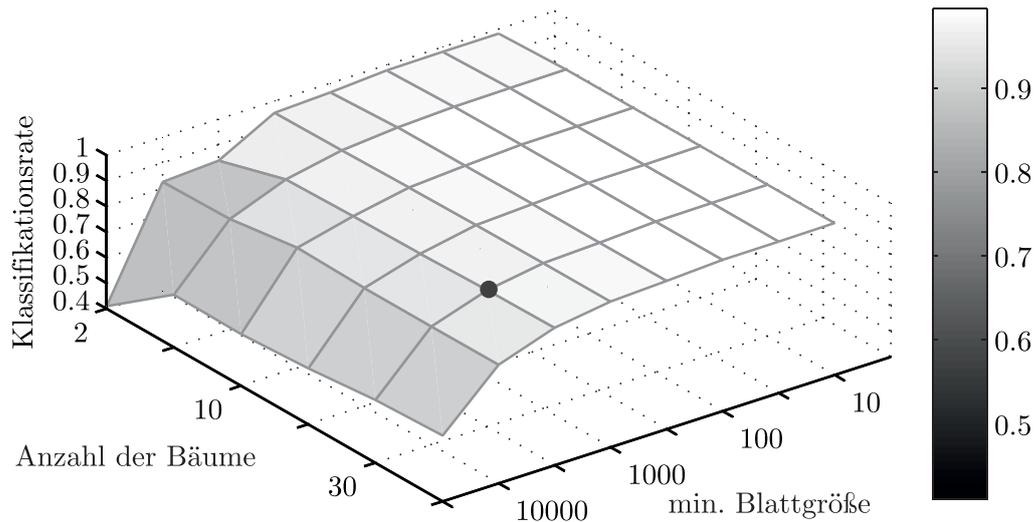


Abbildung 5.6.: Klassifikationsrate abhängig von Baumanzahl und minimaler Blattgröße [55]

Zur weiteren Auswahl der optimalen Parameter sollte die benötigte Rechenzeit betrachtet werden. Die Rechenzeit ist abhängig von der Gesamtanzahl von Knoten innerhalb der Klassifikations-Ensembles. Die Knotenanzahl kann maßgeblich über die minimale Blattzahl gesteuert werden. Bei einer Zusammenfassung von 1000 Instanzen des Trainingssets ergeben sich 100 Knoten innerhalb eines Baumes. Bei kleineren „Zusammenfassungsraten“ steigt die Knotenanzahl jedoch stark an. So enthält ein gut angepasster Baum, bei welchem nur 10 Messwertinstanzen zusammengefasst werden, beispielsweise bereits über 500 Knoten. Ein guter Kompromiss ist bei dieser Merkmalsauswahl für ein Klassifikationsmodell mit *30 Entscheidungsbäumen* mit einer *minimalen Blattgröße von 1000* Trainingsinstanzen gegeben (siehe Markierung in Abbildung 5.6).

### 5.2.3. Vergleich der Verfahren

Im Rahmen dieser Arbeit werden die Verfahren CART-Klassifikationsbaum und Random Forest betrachtet. In diesem Abschnitt werden die optimalen Parameter der jeweiligen Methode ermittelt und es werden die Korrektklassifikationsraten bestimmt. Zur Bestimmung der Raten wird die zweifache Kreuzvalidierung angewendet. Bei der Kreuzvalidierung wird eine Zufallsauswahl von Messvektoren für das Training und für

## 5. Implementierung einer Datenbasis zur Identifikation von Substanzen

den Test durchgeführt. Aufgrund der Zufälligkeit handelt es sich bei den untersuchten Datensätzen nicht um unabhängige Daten. Test- und Trainingsvektoren können aus demselben Messverlauf stammen. Bei unabhängigen Messungen sind geringere Korrekturklassifikationsraten zu erwarten. Für eine Bewertung der Methoden ist ein Vergleich der so ermittelten Kenngröße geeignet.

Die bestimmten Korrekturklassifikationsraten betragen:

- 79 % für den CART-Entscheidungsbaum
- 98 % für das Entscheidungsbaum-Ensemble nach der Random Forest Methode

Es zeigt sich also, dass für diese Fragestellung eine Kombination vieler schwacher Lerner im Gegensatz zu einem einzelnen komplexen Entscheidungsbaum zu einer besseren Klassifizierung führt. In einer gesonderten Untersuchung wurde zudem die Korrekturklassifikationsrate dieser Daten durch das Verfahren der Support Vector Machine (SVM) bestimmt. Die Rate unter Anwendung dieser Methode liegt mit 92 % zwischen dem CART-Entscheidungsbaum und dem Random Forest [55]. Betrachtet man die erforderlichen Rechenschritte zur Klassifikation eines unbekanntes Merkmalsvektors, so benötigt der Random Forest etwa 20-mal mehr Takte als der CART-Entscheidungsbaum. Die SVM hingegen benötigt circa den Faktor 1000 mehr an Rechenschritten.

Da das Verfahren auf einem embedded System mit einer geringen Rechenleistung umgesetzt werden soll, ist ein Algorithmus mit einer geringen Komplexität zu bevorzugen. Das Random Forest Verfahren erscheint aufgrund seines mittleren Rechenbedarfs und der sehr guten Korrekturklassifikationsrate als geeignet.

### 5.3. Implementierung eines Random-Forest Klassifikators

Das Ergebnis des überwachten Lernens durch die Bootstrap-Aggregation-Implementierung der Random Forest Methode in der Numeriksoftware Matlab, ist zunächst eine interne Datenstruktur. Diese weist ein proprietäres Format auf, welches zunächst für die Anwendung in einem Eingebetteten-System angepasst werden muss. Generell handelt es

sich bei der Struktur um eine Ansammlung von Binärbäumen, welche als Objekt in der Software abgelegt ist.

### 5.3.1. Datenaustauschformat zwischen überwachtem Lerner und Klassifikations-Programm

Als Austauschformat zwischen der Computernumeriksoftware und der Firmware des Messinstruments wurde eine XML-Struktur gewählt [92]. Hier werden alle Bäume des Random Forest in einer Datei abgelegt. Die Daten bleiben in dieser Klartext-Datei menschenlesbar. Zur Erstellung der Klassifikator-Datei wird ein Skript genutzt, welches das Random-Forest-Objekt in die XML-Struktur schreibt. In dieser Struktur werden in einem Dateikopf allgemeine Informationen zu dem Klassifikations-Ensemble abgelegt (Abbildung A.1 im Anhang auf Seite 122 ff.).

Die wichtigsten Informationen sind die Anzahl der enthaltenen identifizierbaren Stoffe und die Anzahl der Entscheidungsbäume im Ensemble. Auf den Kopf folgen die durchnummerierten Entscheidungsbäume (Abbildung 5.7), unter Angabe der Anzahl der enthaltenen Knoten. Jeder Knoten besitzt strukturiert seine Informationen, wie in Kapitel 2.4.2 beschrieben: Zu vergleichendes Merkmal, den Vergleichsoperator („gt“: Größer als), Vergleichswert und die Nachfolger für ein logisch-falsches und ein logisch-wahres Ergebnis der Vergleichsoperation. Der in dem XML-Quellcode beschriebene Knoten würde

```
<Trees>
  <Tree1>
    <TreeInfo>
      <nodes> 3 </nodes>
    </TreeInfo>
    <node1>
      <feature> 1 </feature>
      <operator> gt </operator>
      <value> 54.000000 </value>
      <false_follower> 3 </false_follower>
      <true_follower> 2 </true_follower>
    </node1>
```

Abbildung 5.7.: Innere Knoten einzelner Entscheidungsbäume in der Klassifikatordatei

## 5. Implementierung einer Datenbasis zur Identifikation von Substanzen

als das Merkmal 1 (Signalhöhe des Photoionisationsdetektors) mit dem Wert 54,0 vergleichen. Ist der Messwert höher als der im Knoten hinterlegte Wert, so wird als nächstes der Knoten 2 im Baum angesteuert. Bei logisch-falschem Ergebnis des Vergleichs folgt der Baumknoten Nummer 3.

Auf diese *inneren Knoten*, welche die Entscheidungen repräsentieren, folgen die Ergebnisblätter der Bäume (Abbildung A.2 im Anhang).

Diese enthalten den Namen der Substanz und eine Identifikationsnummer als Verweis auf eine Stoffdatenbank mit weiteren Informationen zur gefundenen Substanz. Zusätzlich ist die Wahrscheinlichkeit jedes Blattknoten, basierend auf den Trainingsdaten, angegeben.

### 5.3.2. Implementierung des Klassifikators auf dem Zielsystem

Die erzeugte XML-Datei mit der Beschreibung des Klassifikators nach der Random Forest Methode wird in dem verwendeten embedded System zur Online-Auswertung von direkten Messungen genutzt. Hierfür wird die XML-Beschreibung des Klassifikators zunächst in eine Binärbaumstruktur gewandelt. Durch die byte-weise Repräsentation im flüchtigen Speicher des Prozessorsystems, ist eine schnelle Verarbeitung der Entscheidungsbäume möglich.

Das Firmwareprogramm des Messsystems analysiert die, auf einer SD-Karte hinterlegte, XML-Datei für jeden hinterlegten Entscheidungsbaum einzeln. Aus den Informationen der Knoten wird jeweils eine verkettete Datenstruktur im Speicher des Zielsystems erzeugt [93]. Diese besitzt als Datenwerte das zu vergleichende Merkmal und den zugehörigen Vergleichswert. Die beiden Nachfolgeknoten sind als Zeiger auf die Speicherstellen des Arbeitsspeichers ausgeführt.

Die Abbildung 5.3.2 veranschaulicht den Aufbau der Entscheidungsbäume im Speicher des Zielsystems. Die Zylinder stellen die Instanzen der Entscheidungsknoten im Speicher dar. Grau unterlegt sind die Adress-Zeiger, welche auf die Nachfolge-Elemente verweisen. Die Entscheidungsbäume eines Random Forest Ensembles haben eine große Varianz an Knotenanzahl. Die hier aus den Trainingsdaten erstellten Baum-Instanzen besitzen beispielsweise zwischen 200 und 1000 Knoten. Aufgrund der unterschiedlichen

## 5. Implementierung einer Datenbasis zur Identifikation von Substanzen

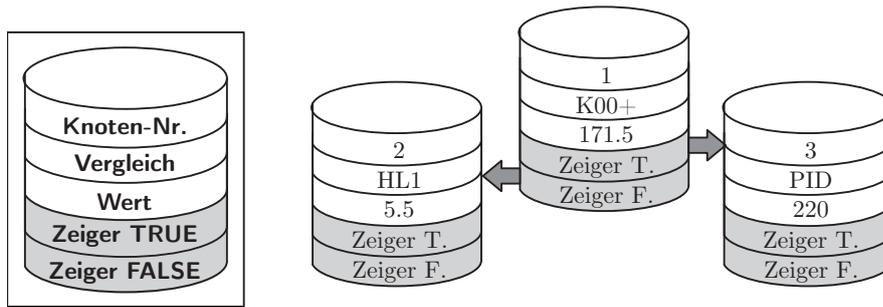


Abbildung 5.8.: Speicherstruktur des Random-Forest-Klassifikators

Komplexitäten der Bäume wurde eine dynamische Speicherreservierung für die Baum-Instanzierung auf dem Zielsystem gewählt.

## 6. Evaluation und Ergebnisse

In diesem Kapitel wird ein erstelltes Ensemble an Entscheidungsbäumen evaluiert. Zunächst soll anhand eines realen Messlaufs untersucht werden, wie sich der Klassifikator verhält. Der Verlauf im oberen Teil von Abbildung 6.1 zeigt die Rohdaten des PID und

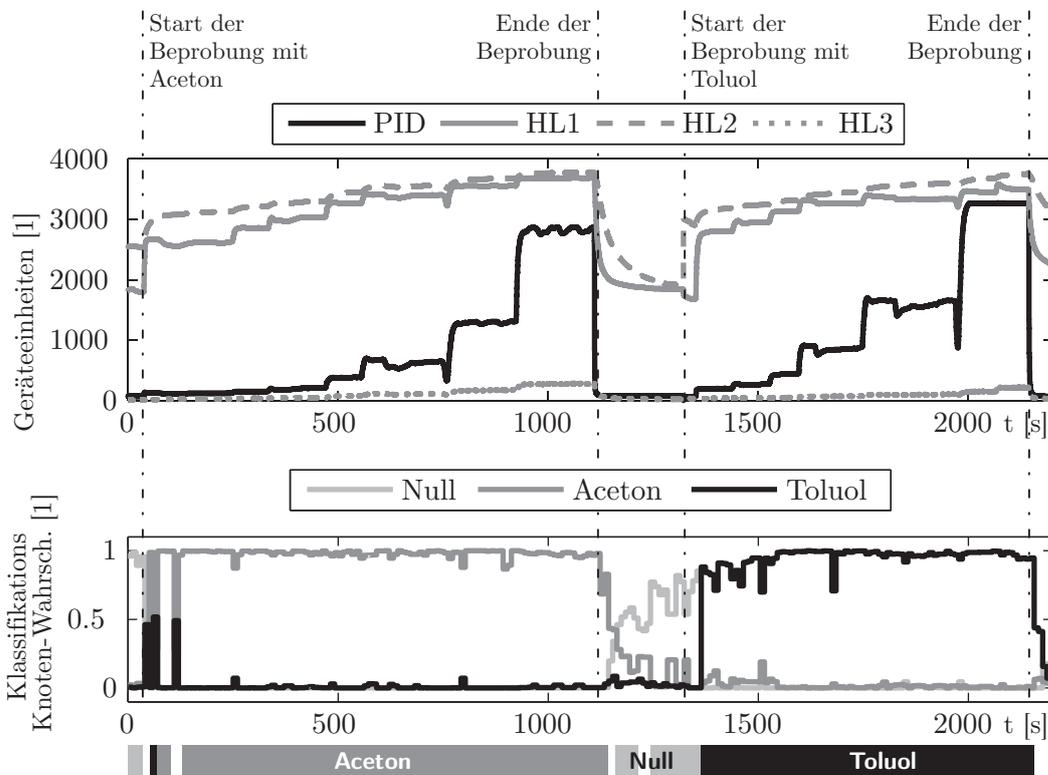


Abbildung 6.1.: Evaluation eines Messdatenverlaufs durch den Random Forest Klassifikator

der Halbleitersensoren. Zur Verbesserung der Übersichtlichkeit sind die Verläufe der Elektrochemischen Zellen und der IM-Spektren nicht dargestellt. Es ist der zeitliche Verlauf von *zwei* Labormessungen gezeigt. Die Messverläufe wurden in dieser Darstellung für



## 6. Evaluation und Ergebnisse

die Veranschaulichung hintereinander gehängt. Bis Sekunde 1390 handelt es sich um eine Messung von aufsteigenden Konzentrationsstufen von Aceton. Darauf folgen, in einer unabhängigen Messung, Konzentrationsstufen von Toluol. Sie wurden mit der Messanordnung aus Abschnitt 5.1.1 erzeugt. Bei den Messungen wurden die Konzentrationen aus Tabelle 6.1 durch die Verdampfer-Einheit erzeugt. Die Konzentration wurde durch

| Aceton    |               | Toluol    |               |
|-----------|---------------|-----------|---------------|
| Zeitpunkt | Konzentration | Zeitpunkt | Konzentration |
| 245 s     | 5 ppm         | 1416 s    | 5 ppm         |
| 330 s     | 20 ppm        | 1503 s    | 10 ppm        |
| 465 s     | 50 ppm        | 1592 s    | 20 ppm        |
| 555 s     | 100 ppm       | 1662 s    | 50 ppm        |
| 750 s     | 200 ppm       | 1812 s    | 100 ppm       |
| 920 s     | 500 ppm       | 2034 s    | 200 ppm       |
|           |               | 2130 s    | 500 ppm       |

Tabelle 6.1.: Evaluation eines Messdatenverlaufs durch den Random Forest Klassifikator eine parallele Vergleichsmessung mithilfe eines FTIR bestimmt.

Die in der Tabelle angegebenen Zeiten geben die Zeitpunkte an, bei denen die Konzentration laut Referenzmessung sicher anlag. Die in den zeitlichen Verläufen zu beobachtenden Schwankungen rühren von Umschaltvorgängen des Aufbaus zur Probenbereitung her, so wurde z. B. beim Konzentrationsprung von 200 ppm auf 500 ppm eine zusätzliche Verdünnungsstufe abgeschaltet. Die Oszillationen des Signalverlaufs bei der hohen Aceton-Konzentration stammt ebenfalls von der Probenbereitung. Hier hat eine der beiden Spritzenpumpen, welche die Testsubstanz in flüssiger Phase fördern, Luftblasen angesaugt. Somit wurde periodisch weniger Substanzmenge verdampft. Insgesamt ist zu erkennen, dass sich die Signalhöhen der Sensoren in ihrem zeitlichen Verlauf für beide Substanzen stark ähneln.

Neben den Einzelsensoren wurden zur Klassifikation auch die Merkmale der IM-Spektren herangezogen. In Abbildung 6.2 sind die Signaturen des IMS für zwei Zeitpunkte im Verlauf dargestellt. Zu erkennen sind die Unterschiede im positiven Spektrum (schwarze Darstellung) - Aceton verursacht ein Produktionenpeak rechts des Reaktantionenpeaks. Dies verursacht einen niedrigeren  $K_0$ -Wert für das Merkmal mit dem Bezeichner K01+ (siehe

## 6. Evaluation und Ergebnisse

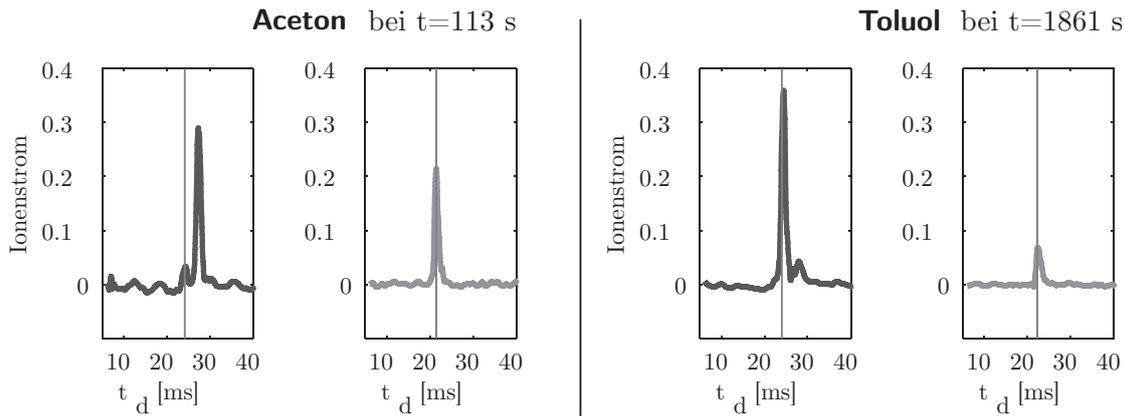


Abbildung 6.2.: Zwei IM-Spektren im Messverlauf mit Bezugszeitpunkt der Reaktantionen (senkrechte Linie)

Tabelle 4.2, S. 82). Zur Auswertung des Messverlaufs aus Abbildung 6.1 wurde jeder einhundertste Messzeitpunkt durch einen Random Forest Klassifikator evaluiert. Bei dem genutzten Klassifikator handelt es sich um ein Ensemble aus 30 Entscheidungsbäumen mit einer minimalen Blattgröße von 1000 Trainingsinstanzen. Die Entscheidungsbäume werten 14 Merkmale der Online-Messung aus (sechs Sensorwerte und acht  $K_0$ -Werte des Ionenmobilitätsspektrums). Die evaluierten Messdaten sind nicht vollständig unabhängig von den Trainingsdaten.

Im unteren Bereich der Abbildung 6.1 ist das Ergebnis der Klassifikation als Verlauf über die Messzeit aufgetragen. Die Graphen zeigen den Durchschnitt der Knoten-Wahrscheinlichkeiten der entsprechenden Klasse für alle Entscheidungsbäumergebnisse des Ensembles. Neben den beiden Stoffklassen Toluol und Aceton und der *Null*-Klasse für kontaminationsfreie Messungen, enthält der Random Forest Klassifikator noch weitere Zielsubstanzen (siehe Abschnitt 5.1). Im Verlauf der Klassifikations-Knotenwahrscheinlichkeit kann man bereits eine gute Korrektklassifikation erkennen.

Vor dem Start der Beprobung ist die *Null*-Klasse (hellgrau dargestellt) am stärksten ausgeprägt. Sobald die Beprobung mit einer geringen Testkonzentration von 1 ppm Aceton startet, geht die Wahrscheinlichkeit der *Null*-Klasse sofort auf unter 50 %. Es findet bei dieser Auswertung also keine Falsch-Negative Einschätzung statt. Nach einer kurzen Einschwingphase bei welcher kurzzeitig keine Substanz, aber auch keine Nullluft identifiziert wird, geht die Klassifikationsbewertung mit fast 100 % auf die Stoffklasse Aceton



(mittelgrau dargestellt). Nach Ende der Beprobung mit Aceton sinkt die Klassenwahrscheinlichkeit, bis schließlich wieder die Null-Klasse am dominantesten auftritt. Bei der Applikation der Testsubstanz Toluol erfolgt daraufhin ebenfalls eine sofortige Korrekturklassifikation nahe 100 % (schwarze Darstellung). Unter dem Diagramm ist die Identifikation der Zeitbereiche als Balken visualisiert. Es wurde ein Klassifikationsschwellwert von 50 % als Entscheidungsgrundlage genommen. Die identifizierten Zeitbereiche decken sich sehr gut mit den Beprobungsbereichen.

## 6.1. Test des Klassifikators mit unabhängigen Testdaten

Zur weiteren Untersuchung der Klassifikationseigenschaften der gewählten Methode wurde das Modell mit unabhängigen Testdaten beaufschlagt. Die nachfolgenden Ergebnisse basieren auf einem Random Forest Klassifikator, welcher mit  $10^5$  Messdaten aus 39 Messreihen trainiert wurde. Dies entspricht einem Viertel der Grundgesamtheit der Messungen. Der Forest besteht aus 30 Entscheidungsbäumen mit einer minimalen Blattgröße von 1000 Instanzen. Getestet wurde mit einem Testset, dessen Messungen nicht aus den Messreihen und damit nicht aus dem Trainingsdatensatz stammen. Hier wurde eine Stichprobe von 10 000 Messungen gewählt.

In den Tabellen 6.2 und 6.3 wird zunächst der allgemeine Aufbau der nachfolgend betrachteten Konfusionsmatrizen veranschaulicht. In den Spalten der beiden Tabellen

| wahre Klasse | Ergebnis-Klasse      |                      | wahre Klasse | Ergebnis-Klasse |          |          |
|--------------|----------------------|----------------------|--------------|-----------------|----------|----------|
|              | A                    | B                    |              | A               | B        | C        |
| A            | Richtig-Positiv $TP$ | Falsch-Negativ $FN$  | A            | $TP_A$          | $E_{AB}$ | $E_{AC}$ |
| B            | Falsch-Positiv $FP$  | Richtig-Negativ $TN$ | B            | $E_{BA}$        | $TP_B$   | $E_{BC}$ |
|              |                      |                      | C            | $E_{CA}$        | $E_{CB}$ | $TP_C$   |

Tabelle 6.2.: Konfusionsmatrix eines Zwei-Klassen-Problems

Tabelle 6.3.: Konfusionsmatrix eines Mehr-Klassen-Problems

wird eingetragen, welches Ergebnis der auf Messdaten angewendete Klassifikator liefert. Dem gegenüber stehen in den Zeilen die bekannten wahren Klassenbezeichnungen der Test-Messdaten [94]. Im Zweiklassen-Problem liefert ein Prädiktor nur eine

## 6. Evaluation und Ergebnisse

binäre Aussage. Im Sinne der Gefahrenabwehr von chemischen Substanzen sind dies beispielsweise die Aussagen *gefährlicher Stoff vorhanden* und *Luft ist frei von Gefahrstoffen*.

In der Tabelle 6.2 entspräche Klasse A damit der Aussage *keine Gefahr* und Klasse B der Aussage *Gefahrstoff vorhanden*. Die vier Quadranten bewerten nun die vier möglichen Aussagenkombinationen. So beschreiben die beiden Elemente der Hauptdiagonale die Korrektklassifikationen. Die Falsch-Klassifikationen werden nach ihrer Aussage weiter bewertet (falsch-positiv: Fehllarm, falsch-negativ: kein Alarm, trotz Gefahr) (siehe auch Kapitel 2.4.1). Die Tabelle 6.3 zeigt die Erweiterung der Konfusionsmatrix auf eine Problemstellung mit drei möglichen Klassen. Die korrekten Ergebnisse des Klassifikators befinden sich auch hier auf der Hauptdiagonalen der Matrix. Die übrigen Matrix-Elemente beinhalten die individuellen Fehlklassifikationen aller Klassen-Kombinationen ( $E_{**}$ ).

Die Abbildung 6.3 zeigt eine Konfusionsmatrix bei welcher die vom vorgestellten Random Forest Klassifikator identifizierten Substanznamen den a priori bekannten Stoffbezeichnungen gegenübergestellt werden. Hier angegeben sind die absoluten Zahlen der getesteten Stichprobe von Messdaten (10 000 Instanzen). Da nicht für alle Substanzen die gleiche Anzahl an Messreihen aufgenommen wurde, ergibt sich beim Blick auf die absoluten Klassenzahlen eine Varianz. Es ist zu erkennen, dass der Klassifikator nicht frei von Fehlklassifizierungen ist. So gibt es eine Häufung bei der Identifikation von Ethanol, dieser Stoff wird auch als Aceton klassifiziert. Die Fehlklassifikation von Aceton als Ethanol ist dagegen nicht so stark ausgeprägt. Ein Vergleich mit der Anzahl der Korrektklassifikationen von Aceton und Ethanol zeigt jedoch auch, dass diese Messungen in großer Zahl in der Test-Stichprobe vorhanden sind. Für eine weitere Bewertung des Klassifikators eignet sich die Beobachtung der relativen Daten in Form der *Sensitivität* und der *Präzision* durch Konfusionsmatrizen. Die Sensitivität (engl. recall) des Klassifikators beschreibt die *richtig positiven* Ergebnisse eines Klassifikators in Bezug auf die falsch negative Missklassifikation. Die Sensitivität ist damit eine Maß für die Trefferquote. In Gleichung 6.1 ist die Berechnung für das dreiklassige Beispiel in Tabelle 6.3 dargelegt.

$$\text{recall}_A = \text{prob}(\text{richtig erkannt}|\text{wahr positiv}) = \frac{TP_A}{TP_A + E_{AB} + E_{AC}} \quad (6.1)$$



## 6. Evaluation und Ergebnisse

|              |                       | Klassifikationsergebnis |             |             |            |             |        |
|--------------|-----------------------|-------------------------|-------------|-------------|------------|-------------|--------|
|              |                       | Null                    | Aceton      | Chloroform  | Essigsäure | Ethanol     | Toluol |
| wahre Klasse | korrekt klassifiziert |                         |             |             |            |             |        |
|              | Null                  | <b>593</b>              | 3           | 36          | 1          | 86          | 1      |
|              | Aceton                | 12                      | <b>2349</b> | 0           | 0          | 49          | 5      |
|              | Chloroform            | 37                      | 0           | <b>1676</b> | 0          | 96          | 12     |
|              | Essigsäure            | 0                       | 48          | 0           | <b>782</b> | 35          | 26     |
|              | Ethanol               | 44                      | 428         | 1           | 1          | <b>1658</b> | 88     |
| Toluol       | 140                   | 124                     | 250         | 1           | 10         | <b>1408</b> |        |

Abbildung 6.3.: Konfusionsmatrix absoluter Anzahl von klassifizierten Beobachtungen [55]

Es wird die Anzahl der richtigen Ereignisse also über die Summe der Spalteneinträge normiert.

Die Präzision dagegen setzt die *richtig positiven* Ergebnisse ins Verhältnis zu den falsch positiven Klassifikationsergebnissen (Gl. 6.2).

$$\text{precision}_A = \text{prob}(\text{richtig erkannt} | \text{erkannt positiv}) = \frac{TP_A}{TP_A + E_{BA} + E_{CA}} \quad (6.2)$$

Zur Bestimmung der Präzision wird die Anzahl der richtigen Ereignisse damit über die Summe der Zeileneinträge normiert. Die Normierung kann nicht nur für die richtig klassifizierten Stichprobenelemente durchgeführt werden, sondern ebenfalls für falsch klassifizierte.

Aus den Daten der hier betrachteten Stichprobe ergeben sich damit die relativen Raten der Sensitivität und Präzision aus Abbildung 6.4. Die Elemente der Matrix sind abhängig von ihrem Wert der Sensitivität bzw. der Präzision stärker oder schwächer eingefärbt.



## 6. Evaluation und Ergebnisse

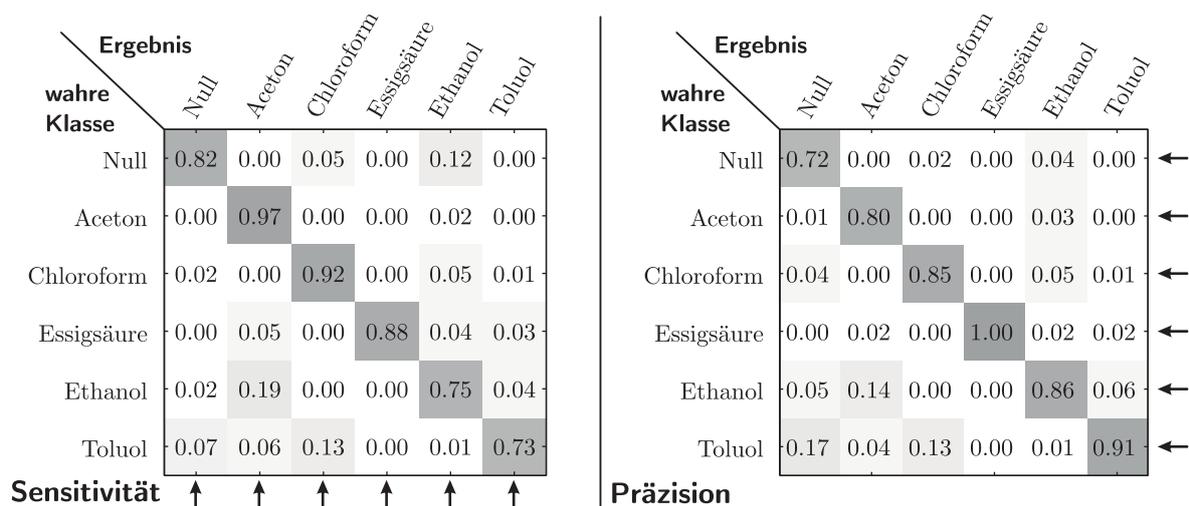


Abbildung 6.4.: Beschreibung des Klassifikationsverhaltens eines Random Forest Klassifikators anhand von Konfusionsmatritzen der Sensitivität und Präzision

Es ist zu erkennen, dass der Anteil der richtig klassifizierten Substanzen bezogen auf die Gesamtheit der Klassifikationsergebnisse eines Stoffes, also die Sensitivität, im Mittel bei 85 % liegt. Die Präzision der richtig klassifizierten Messungen befindet sich ebenfalls im Mittel in diesem Bereich mit 86 %.

Unterschiede in Präzision und Sensitivität lassen sich bei den Werten für die Klassifikation von Toluol erkennen. So wird diese Substanz in Bezug auf die Anzahl der wahren Klasse mit einer Präzision von 91 % erkannt. In Bezug auf die Ergebnisklassen jedoch mit einer Sensitivität von 73 %.

### 6.1.1. Bewertung der Klassifikationsraten

In den vorangegangenen Darstellungen der Klassifikationsergebnisse wurde festgestellt, dass eine gewisse Rate an Fehlklassifikationen gegeben ist. Die Fehler rühren in erster Linie aus der Unsicherheit der Messergebnisse her. Betrachtet man die Signalkette aus Abbildung 4.1 auf Seite 66, so ist jeder funktionelle Block mit einer Messunsicherheit behaftet, welche sich bis zur Stoff-Identifikation fortpflanzt.

### Unsicherheit der Probe

So hat bereits die Substanzprobe eine große Unsicherheit hinsichtlich ihrer Konzentration und zum Teil auch hinsichtlich ihrer Reinheit. Die Verdampfungseinheit des Messaufbaus (Abbildung 5.2) besitzt einen statischen Fehler in der Konzentrationsbereitung aufgrund von Toleranzen der Massendurchflussregler und der Spritzenpumpe. Dieser Fehler wird durch die Regelung mithilfe eines Referenz-Messgerätes - dem FTIR kompensiert. Das Referenzmessgerät selbst besitzt jedoch ebenfalls eine Unsicherheit von bis zu 2% in der Quantifizierung der Konzentration [95]. Zudem wird durch Lufteinschlüsse bei der Förderung flüssiger Phase durch die Spritzenpumpe in einem Teil der Messungen ein dynamischer Fehler erzeugt. Dieser führt dann zu einer Konzentrations-Oszillation, wie sie schon in Abbildung 6.1 zu beobachten war. Diese Oszillation ist jedoch nicht durch das Referenz-FTIR feststellbar, da dieses Instrument prinzipbedingt eine Gaszelle mit einem relativ großen Volumen von 0,45 Litern beinhaltet. Aufgrund der sich einstellenden Konzentrationsvermischung innerhalb der Messzelle ergibt sich eine träge Dynamik des Referenzmesssystems. Neben den statischen und dynamischen Fehlern der Substanzkonzentration sind ebenso Verunreinigungen der Gas-Probe möglich. So handelt es sich bei den gasführenden Komponenten des Versuchsaufbaus größtenteils um unbeheizte Glas und Inert-Kunststoff-Verbindungen. Hier kann es bei hohen Probenkonzentrationen zu analytischem Memory kommen. Wird die Probensubstanz gewechselt, kann es zur Vermischung der neuen Substanz und des an den Wandungen haftenden Substanz-Residuums kommen.

### Einfluss der Systemfluidik

Diese Klasse der Unsicherheit setzt sich auch hinter der Systemgrenze fort. Die Fluidik des evaluierten Messgerätes ist ebenfalls von analytischem Memory betroffen. Dies wird z. B. durch eine beheizte Einlassmembran des Ionenmobilitätsspektrometers und durch die Wahl von möglichst inerten Materialien weitestgehend gedämpft. Zwischen den Messungen wurde das System mit gereinigter Luft permanent gespült. Das Substanz-Residuum an den Oberflächen der System-Fluidik kann dadurch in das Spülgas übergehen. Das analytische Memory wird durch diese Maßnahme auf ein Minimum reduziert.

### Einfluss der Sensorik

Bei weiterer Betrachtung der Signalkette ist der Fehler der Sensoren zu berücksichtigen. Hier fällt insbesondere bei den elektrochemischen Zellen die Sensordegradation ins Gewicht. Über den gesamten Zeitraum der Messkampagnen wurden die Zellen mit ihrer Drift von 0,5 %-2 % pro Monat entsprechend ihrem Degradations-Grad getauscht [79]. Die Halbleitersensoren werden stark von Temperatur- und Fluss-Schwankungen beeinflusst. Thermische Schwankungen werden durch Regelung der Temperatur gedämpft, kurzfristige Änderungen können jedoch nicht vollständig vom Regler abgefangen werden.

Zudem kann es über eine längere Standzeit hinweg zu einer Basisliniendrift kommen. Die Signalqualität des Ionenmobilitätsspektrometers wird maßgeblich durch die Konditionierung des Driftgases beeinflusst. So spielt vor allem der Feuchtegehalt eine große Rolle, da dieser den Ladungstransfer auf die Zielmoleküle bestimmt. Eine hohe Feuchte reduziert den Transfer von Ladung auf Analytmoleküle, welche eine nur wenig höhere Protonenaffinität als die Wassercluster besitzen [96]. Da der eingesetzte Trocken-Filter jedoch einer Alterung unterworfen ist, nimmt der Feuchtigkeitsgehalt des Driftgases mit der Zeit zu. Dies hat eine Reduktion der Nachweisstärke zur Folge.

**Fehlerbetrachtung des  $K_0$ -Wertes** Das Ionenmobilitätsspektrometer ist durch die stoffspezifischen  $K_0$ -Werte ein sehr identifikationsstarker Detektor. Allerdings wird dieser, in der Theorie vom System unabhängige Vergleichswert für die Ionenmobilität aus vielen Messwerten zusammengesetzt, welche jeweils eine eigene Messunsicherheit besitzen. Nach dem Gaußschen Fehlerfortpflanzungsgesetz wird für die Gesamtunsicherheit die Euklidische Norm der approximierten Einzelunsicherheiten gebildet [97]. In der Gleichung 6.3 ist die Gesamtunsicherheit der reduzierten Ionenmobilitäten (Gl. 2.16) aus der Veranschaulichung von Abbildung 4.10 (S. 80) angedeutet:

$$\Delta K_0 = \sqrt{\left(\frac{\partial K_0}{\partial P_D} \Delta P_D\right)^2 + \left(\frac{\partial K_0}{\partial T_D} \Delta T_D\right)^2 + \left(\frac{\partial K_0}{\partial t_d} \Delta t_d\right)^2 \dots} \quad (6.3)$$

Die in Abschnitt 4.3.3 beschriebenen vier gemessenen Momentanwerte sind nach Gleichungen A.2-A.5 fehlerbehaftet. Dies sind bei den Druck- und Temperaturmesswerten  $P_D$  und  $T_D$  die Unsicherheiten der Messaufnehmer, bei der Hochspannung  $U$  die der

Widerstandskaskade und bei der Driftzeit  $t_d$  kann dies ein Fehler des Abtastzeitpunktes der Datenaufnahme sein. Hinzu kommt ein statischer Fehler bei der Bestimmung der Driftröhrenlänge, welche quadratisch eingeht (Gl. A.6).

### Qualität der Datensätze

Zum Aufbau und zur Evaluation des Klassifikators wurden 39 Messreihen über einen Zeitraum von einigen Monaten aufgenommen. Jede Messreihe enthält einige 1000 Messvektoren, so dass sich eine umfangreiche Grundgesamtheit ergibt. Diese Menge wurde nach der Bootstrapping-Methode in Trainings- und Testdaten aufgeteilt [98][99]. Um eine Unabhängigkeit der Test- von den Trainingsdaten zu gewährleisten, wurden zuvor ganze Messreihen als Testsätze ausgewählt. Es hat sich gezeigt, dass einige Messreihen aufgrund der Sensordegradation eine schlechtere Nachweisfähigkeit besitzen. Dies wirkt sich negativ auf die Klassifikation aus, sei es in der Verwendung als Quelle für Trainings- oder für Testdaten. Dieser Umstand betrifft gerade die Substanzen mit einem geringeren Datenumfang. Das Bootstrap-Verfahren unterscheidet nicht zwischen Datensätzen mit schwach ausgeprägter Signatur und starken Signalmustern. So werden für den Lerner auch Messungen verwendet, bei denen eine Substanzkonzentration nahe der Nachweisgrenze anliegt. Wie schon in Abbildung 6.1 zu sehen, kommt es gerade bei diesen niedrigen Konzentrationen zur Fehlidentifikation.

Es ist die Kombination aller Unsicherheiten der Signalkette, die bei dieser Auswahl von 39 Messreihen zu einer Fehlklassifikationsrate führt. Durch die Extraktion mehrerer Tausend Merkmalsvektoren wurde bereits eine breite Datenbasis erzeugt, bei denen die Instanzen innerhalb einer Konzentrationsstufe eine Abhängigkeit untereinander besitzen. Um die Generalisierungsfähigkeit noch weiter zu optimieren sind daher weitere Messungen mit unterschiedlichen Sensor-Alterungen empfehlenswert. Um die Korrekt-klassifikationsrate weiter zu verbessern, könnten Datensätze mit Probengaskonzentrationen nahe der Nachweisgrenze ausgespart werden. Dies hätte zur Folge, dass bei Vorliegen von geringen Konzentrationen um die Nachweisgrenze nur qualitative Aussagen getroffen werden können, eine Identifikation des Stoffes jedoch erst bei Konzentrationen etwas oberhalb der Nachweisgrenze der jeweiligen Substanz möglich ist.

### 6.1.2. Diskussion zur Übertragbarkeit des Klassifikators

Im Rahmen des Forschungsprojekts DACHS wurden mehrere Funktionsmuster des Messsystems aufgebaut. Als Evaluationssysteme wurden die Demonstratoren mit den Bezeichnungen DACHS2 und DACHS3 untersucht. Die in Kapitel 6 und 5 zu Grunde liegenden Messdaten stammen vom System DACHS2. Der Aufbau der Fluidik und Elektronikkomponenten beider Funktionsmuster ist identisch. Unterschiede ergeben sich hinsichtlich der Degradation der eingesetzten Sensoren. So kommen in den beiden Systemen aufgrund unterschiedlicher Fertigstellungszeitpunkte nicht vollständig chargengleiche Sensoren zur Anwendung. Dieser Umstand zeigt sich besonders bei den beiden Elektrochemischen Zellen und den Halbleitergassensoren. Die integrierten Ionenmobilitätsspektrometer sind prototypisch aufgebaute Module der Gesamtsysteme. Hier zeigen sich Fertigungstoleranzen in den auszuwertenden Spektren. Bei Betrachtung der reduzierten Ionenmobilitäten ( $K_0$ ) stellt man einen Unterschied der Spektrometer fest (siehe Abbildung 6.5). Es sind die sich ergebenden  $K_0$ -Werte beider Ionenspezies für ei-

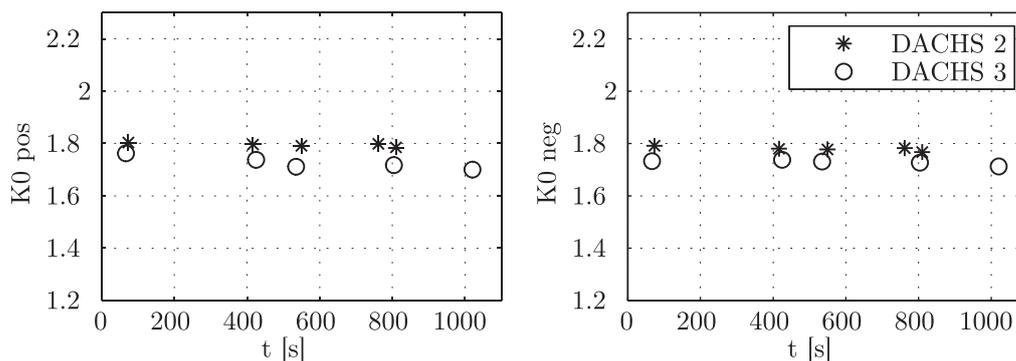


Abbildung 6.5.: Vergleich zweier Funktionsmuster des Messsystems DACHS hinsichtlich der  $K_0$ -Werte bei Messung von Essigsäure

ne Essigsäure-Probe über eine Messdauer abgebildet. Diese Substanz wurde mit beiden Funktionsmustern über eine Zeit von circa 20 Minuten vermessen. Der Grafik ist zu entnehmen, dass zwischen den sich ergebenden  $K_0$ -Werten der beiden Systeme eine statische Abweichung vorherrscht. Zurückzuführen ist dieser Unterschied auf den konstanten Parameter der Driftröhrenlänge für die Berechnung der Kenngröße  $K_0$  nach Gleichung 2.18. Dieser konstante, quadratische Faktor kann genutzt werden, um durch eine Kalibrierung die statische Abweichung zu kompensieren. Auf eine Kalibrierung aufbauend kann der

Klassifikator auch auf weitere Demonstratoren übertragen werden.

## 6.2. Klassifikation von GC-Läufen

Anhand des in Kapitel 3.1 beschriebenen GC-PID-Systems wurde die Anwendbarkeit eines Random Forest Klassifikators auf Chromatographie-Läufe eines Kurzsäulen-Gaschromatographen untersucht. Bei einer Trennstrecke mit einem angeschlossenen Breitband-Detektor ist das entscheidende Merkmal die Retentionszeit  $t_R$ . Diese besitzt jedoch ihre stoffspezifische Gültigkeit nur für bestimmte Bedingungen der Trennstrecke, der sogenannten *Methode*. Diese beinhaltet eine genaue Definition der Säulentemperatur und des Volumenstroms durch die Strecke (meist angegeben über den Vordruck der Kapillarsäule). Es ist möglich, ähnlich wie bei der reduzierten Ionenmobilität  $K_0$  eines Ionenmobilitätsspektrometers, die Retentionszeit  $t_d$  auf die Parameter der gewählten Methode zu normieren (siehe auch Gleichung 2.35). In der Praxis hat sich jedoch die direkte Verwendung der Nettoretentionszeiten zur Stoffidentifikation durchgesetzt. Dazu werden dann die Parameter der Methode definiert und die Retentionszeiten von Zielstoffen empirisch durch Messungen von Referenzsubstanzen ermittelt. Eine unbekannte Substanz kann dann, aufgrund ihrer für diese Methode gültigen Retentionszeit, identifiziert werden.

Häufig werden die Methoden abhängig von der Zusammensetzung des Zielstoffgemisches durch einen Expertennutzer für eine optimale Trennleistung angepasst. Dabei werden vor allem die Parameter des Volumenstroms durch die Trennsäule und die Temperatur der Strecke variiert. Dies führt zu einem multivariaten Klassifikationsproblem, bei welchem die Parameter der Methode berücksichtigt werden müssen.

Das im heterogenen Sensorenarray angewendete Klassifikationsverfahren des Random Forest wurde für die Identifikation von Substanzen eines Kurzsäulen-Chromatogramms mit variablen Temperatur- und Flussparametern evaluiert. Mithilfe des in Kapitel 3.1 beschriebenen GC-PID-Systems wurden die drei aromatischen Verbindungen Aceton, Toluol und Xylol bei unterschiedlichen GC-Methoden vermessen. Die in Abbildung 6.6 gezeigten Chromatogramme stellen das Verhalten des Kurzsäulen-GC bei Variation des Säulendrucks und damit des Volumenstroms durch die Säule dar [65]. Die Trennsäule

## 6. Evaluation und Ergebnisse

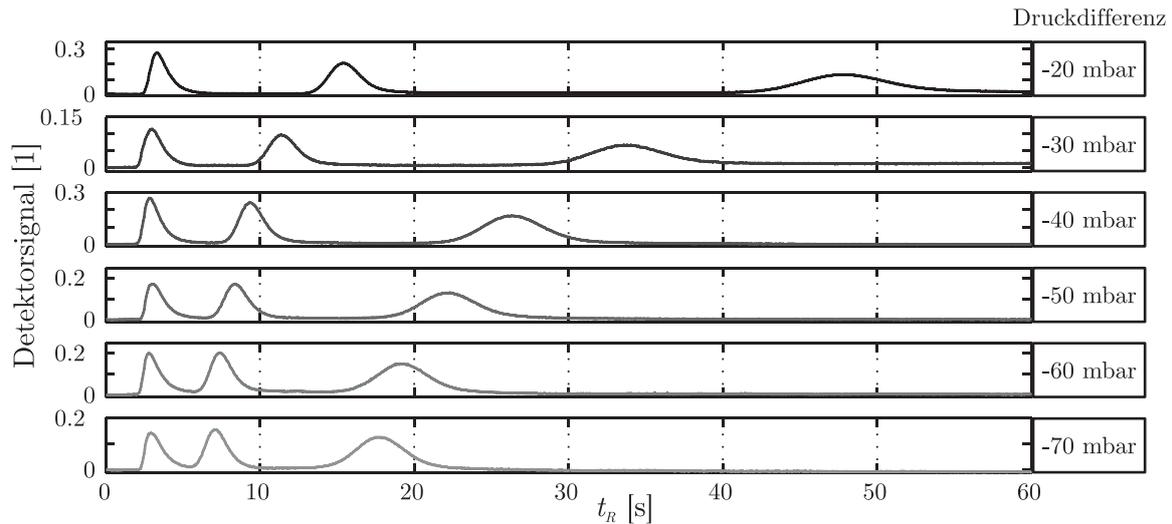


Abbildung 6.6.: Retentionszeit bei unterschiedlichen Druckdifferenzen

wurde in diesem Experiment isotherm bei 32 °C betrieben. Bei der Probe handelt es sich um ein Gemisch der drei Substanzen mit einer Konzentration von jeweils 50 ppm. Zu erkennen sind die Peaks der Substanzen in dem Gemisch in der Reihenfolge abhängig ihrer Löslichkeit: Aceton - Toluol - Xylol. Weiterhin ist zu sehen, dass die Peaks mit zunehmendem Säulendruck (angegeben ist der Differenzdruck der Saugseite der Multikapillarsäule) zusammenrücken. Für eine automatisierte Peakidentifikation ist ein wichtiger Umstand die Kontraktion zwischen dem obersten Chromatogramm mit -20 mbar Druck an der Saugseite und dem untersten Chromatogramm mit einer Methode bei welche -70 mbar am Kapillarausgang anliegen. Der Toluol-Peak des ersten Laufs liegt in einem Retentionszeitbereich des Xylol-Peaks des letzten Laufs. Hierdurch wird eine Klassifikation bei Unkenntnis des Drucks erschwert.

In weiteren Messungen wurde der Druck bzw. ein lineares Druckprogramm konstant gehalten und die Säulentemperatur variiert. Die, durch das in Kapitel 4.2.2 beschriebene Verfahren erkannten Peaks, wurden hinsichtlich ihrer Retentionszeit notiert. Für eine Visualisierung der Streuung der Retentionszeiten sind in Abbildung 6.7 zwei Diagramme dargestellt. Das linke Diagramm stellt Messwerte der drei Substanzen über die beiden Merkmale Retentionszeit und Temperatur dar. Es ist zu erkennen, dass sich bereits über diese Achsen drei disjunkte Gültigkeits-Bereiche ausprägen. Gleichzeitig nimmt mit sinkender Temperatur die Streuung der Messwerte zu. Für den Temperaturbereich zwischen 30 °C und 35 °C ist eine breitere Datenbasis verfügbar.

## 6. Evaluation und Ergebnisse

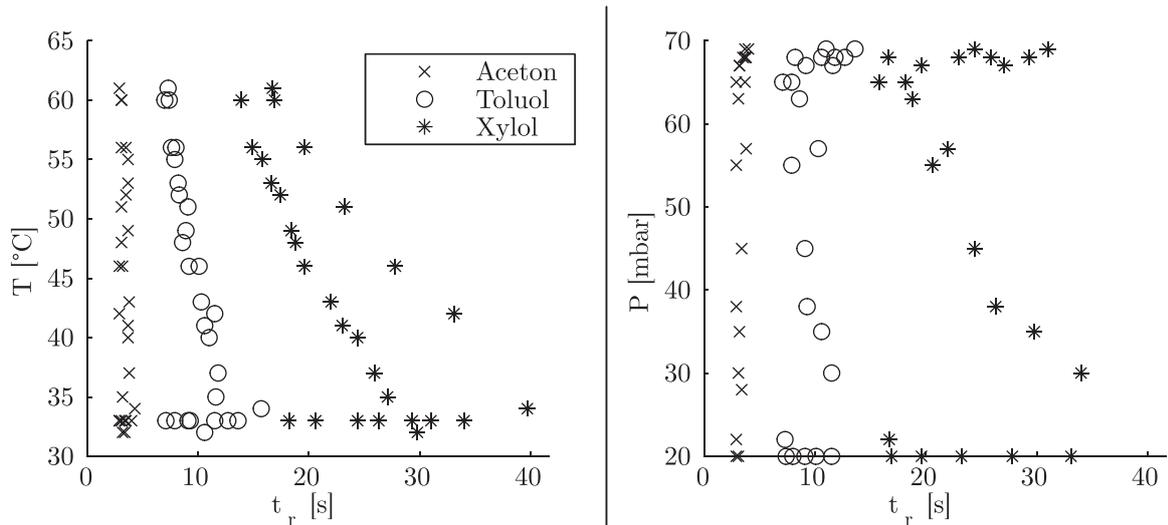


Abbildung 6.7.: Streudiagramme von GC-Läufen bei unterschiedlichen Säulenparametern. Deutlich zu erkennen in der Ebene sind ausgeprägte Bereiche der Klassengültigkeit

Im rechten Streudiagramm, bei dem die Datensätze im Kontext des Differenzdrucks der Trennstrecke aufgetragen sind, zeigen sich ebenfalls klar abgrenzbare Bereiche.

Aus einem Teil des Datensatzes wurde ein Random Forest Klassifikator trainiert. Es wurde die Hälfte der Punkte als Zufallsstichprobe für das Anlernen genutzt. Zur Feststellung der optimalen Baumanzahl ist die Fehlklassifikation als Scree-Plot in Abbildung A.3 aufgezeigt [100]. Aufgrund des kleinen Datenumfangs von 85 Messpunkten, wurde für die dargestellte Evaluierung die Grundgesamtheit an Messdaten verwendet. Nach dem Ellenbogenkriterium sollte der Klassifikator für diesen Datenumfang eine Baumanzahl von fünf bis zehn enthalten. Die Fehlklassifikationsrate liegt mit zehn Bäumen bei circa 2%.

Diese Methode der Substanzidentifikation basierend auf den Messdaten eines Kurzsäulen-GC-PID-Systems mit variierenden Säulenbedingungen hat sich für dieses Klassifikationsproblem als geeignet erwiesen. Sind die Säulenbedingungen durch eine definierte *Methode* für die Messungen fixiert, so eignet sich bereits ein eindimensionales Klassifikationsverfahren.

### 6.3. Untersuchung von GC-IMS-Läufen

In den vorhergehenden Kapiteln wurde das vorgestellte Random Forest Verfahren zur Substanzidentifikation für zwei Anwendungen evaluiert. Zum einen wurde es zur Online-Auswertung von Signalen aus einem Array heterogener Sensoren, bei welchem auch Spektren eines Ionenmobilitätsspektrometers ausgewertet wurden, getestet. Zum anderen kam es zur Auswertung von Chromatogrammen eines breitbandigen Detektors zur Anwendung. Im Folgenden soll gezeigt werden, wie sich Chromatogramme eines Kurzsäulen-Gaschromatographen mit einem Ionenmobilitätsspektrometer als Detektor verhalten.

Sowohl bei der Gaschromatographie als auch bei der Ionenmobilitätsspektrometrie werden Verbindungen aus Gemischen bzw. Ionencluster über die Zeit getrennt. Bei gemeinsamer Berücksichtigung beider Verfahren ergeben sich somit zwei zeitliche Skalen. Mit der Ionenmobilitätsspektrometrie werden Ionen über eine maximale Driftzeit von 45 ms detektiert, wohingegen mit der hier eingesetzte Gaschromatographie Gemische über eine Retentionsdauer von 120 s getrennt werden. Für eine Darstellung von GC-IMS-Messdaten eignet sich daher ein Wasserfallplot, wie in Abbildung 6.8 gezeigt. Das Diagramm ist zweigeteilt, mit dem Spektrum der positiven Ionen im oberen Teil, und der Spektrum der negativen Ionen in der unteren Hälfte. Die vertikalen Achsen beschreiben die Driftzeit der Ionen. Über eine gemeinsame horizontale Zeitachse wird die Laufzeit der Gaschromatographie aufgetragen. Die Signalintensität am Detektor des IMS wird über die Graustufen dargestellt. Die Abstufungen wurden für diese Darstellung stark vereinfacht und die Basis-Ebene ausgeblendet. So sind nur die erhöhten Intensitäten als graue Konturen sichtbar. Die hier gezeigte Messung beinhaltet drei Substanzen in einem Probengemisch. Es wurden drei der für die Gefahrenabwehr wichtigen chlorierten Verbindungen für diese Messung ausgewählt [5]: 1,1,1-Trichlorethan, Trichlorethen und Tetrachlorethen.

Das Diagramm zeigt für beide Ionenspezies die Reaktantionen zu den Driftzeiten  $t_{d,pos} \approx 25$  ms und  $t_{d,neg} \approx 24$  ms als Streifen über die gesamte GC-Laufzeit. Zum Laufzeit-Zeitpunkt  $t = 98$  s wird das Gemisch auf den Kurzsäulen-GC injiziert. Daraufhin ergeben sich Signaturen im IM-Spektrum zu drei Retentionszeitpunkten. Zunächst erscheint die Signatur der Produktionen von 1,1,1-Trichlorethan mit ihrem Maximum

## 6. Evaluation und Ergebnisse

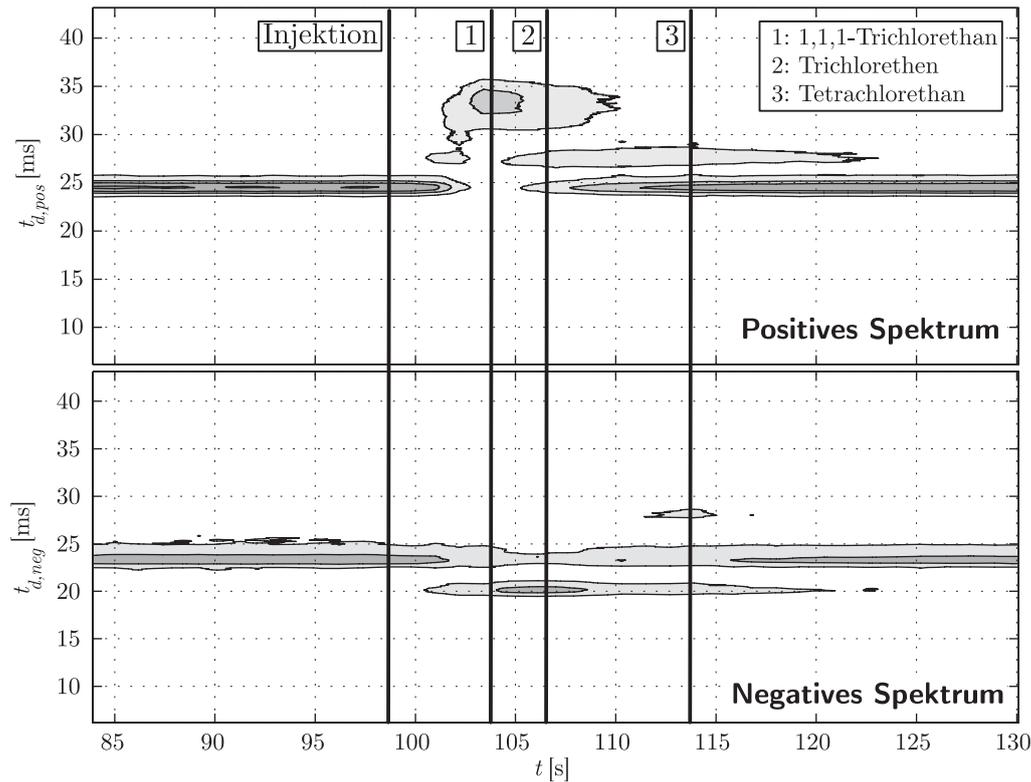


Abbildung 6.8.: GC-Lauf mit IMS-Auswertung [101]

zum Zeitpunkt  $t = 104$  s. Direkt darauf folgt Trichlorethen mit einem Abstand von etwa 2 Sekunden. Die Substanz Tetrachlorethan besitzt einen weiteren Peak im negativen Spektrum und eluiert bei  $t = 114$  s.

Für einen solchen dreidimensionalen Darstellungsraum gibt es mehrere Methoden der Peakerkennung. Bei Hauschild et. al. [102] wurden vier gängige Prinzipien für eine MCC-IMS Auswertung verglichen. Das in dieser Arbeit verwendete Prinzip der *klassischen Methode* nach Kapitel 4.3.2 schneidet bei dem Vergleich besonders gut ab.

Innerhalb eines IMS-Chromatogramms werden zu jedem Zeitpunkt  $t$  die Spektren nach dem Verfahren aus Abbildung 4.9 ausgewertet. Nach Abschluss des Chromatographie-laufs sind alle einzelnen Spektren hinsichtlich ihrer Peaks ausgewertet. Daraufhin werden die Peaks der Spektren in einem Driftzeitintervall von  $\pm 0,3$  ms zusammengefasst über die Retentionszeitachse auf lokale Maxima über das Verfahren aus Kapitel 4.2.2 untersucht. [103] [104]



## 6. Evaluation und Ergebnisse

Die gewonnenen  $K_0$ -Werte sowie Retentionszeiten  $t_r$  werden gemeinsam für eine Substanzidentifikation herangezogen (siehe Tabelle 4.2). Aus einer Stichprobe von Messungen kann analog der Auswertung von Messdaten der *direkten Messung* und der GC-Trennung unter Anwendung eines Breibanddetektors ein Klassifikationsverfahren, wie die vorgestellte Random Forest Methode, trainiert werden.

## 7. Zusammenfassung

In dieser Arbeit wurde der Aufbau zweier Messsysteme zur Detektion und Identifikation toxischer Gase beschrieben. Auf Basis der verwendeten Sensorik wurde eine Methode zur Substanzidentifikation mithilfe von bekannten Mustererkennungsverfahren entwickelt und implementiert.

Die für den Aufbau eines Gasdetektorenarrays für die Gefahrstoffidentifikation in Frage kommenden Sensor- und Detektortechnologien wurden zunächst theoretisch betrachtet. Ein wichtiger Aspekt war dabei das jeweils der Messmethode zugrunde liegende physikalische Wirkprinzip. Aufgrund der Heterogenität der vorgeschlagenen Sensor-Prinzipien ist eine Vielzahl an Zielsubstanzen nachweisbar. So eignet sich der Photoionisationsdetektor besonders zum Nachweis von vielen aromatischen Kohlenwasserstoffen. Die beschriebenen Halbleitersensoren detektieren mit unterschiedlichem Antwortverhalten Alkohole und brennbare Gase. Das Ionenmobilitätsspektrometer registriert Probengasmoleküle abhängig von ihren Protonenaffinitäten bzw. Elektronegativitäten und kann diese zeitlich nach Ionengeometrie, Ionengewicht und Ladung auftrennen. Das Wirkprinzip der Elektrochemischen Zellen basiert auf der Ladungsträgerfreisetzung durch Reduktions- und Oxidationsprozesse des Zielgasmoleküls mit einer Edelmetallelektrode. Durch die angepasste Materialwahl von Elektrode und Elektrolyt sind diese Sensoren sehr selektiv.

Ein weiteres Augenmerk bei der Beschreibung der Sensorgrundlagen wurde auf die zu erwartenden Signaleigenschaften gelegt. So ergeben sich bei Photoionisationsdetektoren, Halbleitersensoren und Elektrochemischen Zellen einzelne Signale mit individuellem Anstiegsverhalten. Das Ionenmobilitätsspektrometer hingegen erzeugt Driftzeit-Spektren. Es wurde gezeigt, wie aus den Wanderungszeiten der Ionen die stoffspezifischen Ionenmobilitäten als Kennwerte bestimmt werden können. Weiter wurde das chemische Trennverfahren der Gaschromatografie erläutert und dabei verschiedene

## 7. Zusammenfassung

Methoden der Probeninjektion verglichen. Dabei wurden die ableitbaren Kenngrößen eines Chromatogramms beschrieben. Diese sind von Bedeutung für eine Klassifikation und eine Identifikation von Substanzen.

Die im Rahmen dieser Arbeit eingesetzten Methoden der Klassifizierung durch die maschinengestützte Auswertung von Signalmustern, sind im zweiten Teil der Grundlagen beschrieben worden. Dazu wurde zunächst erläutert, nach welchen Bedingungen die Grenzen im Merkmalsraum zwischen den Klassen gelegt werden. Danach wurde die Entscheidungsbaummethode CART von den Algorithmen zur Bauminduktion bis zur eigentlichen Klassifikation von Messungen dargestellt. Die Kombination mehrerer Entscheidungsbäume zu einem Ensemble erlaubt eine gesteigerte Robustheit gegen fehlerbehaftete Messdaten. Die daraus abgeleitete Random Forest Methode wurde daher im Theorieteil ebenfalls beschrieben.

Ausgehend von diesen Grundlagen wurden zwei Systeme zur mobilen Gasanalyse praktisch umgesetzt. Dabei handelte es sich um einen Kurzsäulen Gaschromatographen mit einem einzelnen PID als Detektor.

Das zweite realisierte Messsystem war als Kernsystem dieser Arbeit ein Detektorenarray mit GC-Kopplung, welches im Rahmen des Forschungsprojekts „DACHS“ entwickelt wurde. Die Realisierung des Messinstrumentes wurde umfassend dargestellt. Es wurde erläutert, wie die Messdaten der Sensoren zur Gasdetektion gegen Störungen gefiltert werden können. Dabei wurde besonders auf die Filtermethode nach Savitzky-Golay eingegangen, welche Form und Höhe von Signalpeaks bei der Gaschromatographie möglichst wenig beeinflusst. Dieses Filterverfahren wurde für beide Messinstrumente angewendet. Für die gefilterten Chromatogramme und Ionenmobilitätsspektren wurde eine Methode zur numerischen Peakerkennung umgesetzt. Hierüber lassen sich für die Chromatogramme die Retentionszeiten der Einzelkomponenten eines Gases ableiten, sowie die Driftzeiten von Ionen des Ionenmobilitätsspektrometers. Aus den Driftzeiten wiederum können reduzierte Ionenmobilitäten, die  $K_0$ -Werte, berechnet werden. Diese so bestimmten Kenngrößen dienten darauffolgend als Eingangsmerkmale für eine automatische Mustererkennung.

Anhand von umfangreichen Labormessungen wurde ein großer Messdatensatz für einige der wichtigsten toxischen Industriechemikalien aufgebaut. Mit Hilfe des Datensatzes des Detektorenarrays wurden verschiedene Klassifikationsmethoden trainiert. Für den

## 7. Zusammenfassung

CART- und den Random-Forest-Algorithmus wurden die optimalen Parameter für die beste Korrektklassifikation ermittelt. Das sich als geeignet erwiesene Random-Forest-Klassifikationsverfahren wurde auf dem Zielsystem ressourcenschonend als verkettete Binärbaumstruktur implementiert. Als Übertragungsformat zwischen rechnergestütztem Lernalgorithmus und Entscheidungsbaum-Ensemble des eingebetteten Zielsystems wurde die Objektbeschreibungssprache XML genutzt und ein Datenformat beschrieben. Mit Hilfe dieser Messdaten wurde eine Datenbank zur Identifikation von fünf Substanzen bei variierenden Umgebungsbedingungen und Konzentrationen erstellt.

Ein Teil der Labormessungen wurde nicht für das Training des Klassifikators genutzt. Mit dieser unabhängigen Teilmenge wurde die Klassifikationsfähigkeit des Detektorenarrays bei der direkten Messung evaluiert. Die Sensitivität und die Präzision der Stoffidentifikation wurden bewertet. Es wurde gezeigt, dass für diese Art der Messung das gewählte Verfahren des Random-Forest mit 30 Entscheidungsbäumen geeignet ist und gute Klassifikationsergebnisse auch bei unabhängigen Testdaten bietet. Dabei wurde auch der mögliche Fehler der Messwerte durch unterschiedliche Einflussfaktoren bei der Datenaufnahme betrachtet.

Zur Untersuchung von Chromatogrammen wurden Messdaten des entwickelten GC-PID-Systems verwendet. Die Einbeziehung von Temperatur- und Flussverhältnissen, welche innerhalb einer Trennsäule herrschen als Merkmale für eine Mustererkennung, erlaubt eine bessere Generalisierung der Stoffidentifikation auf mehrere GC-Methoden. Es wurde dargestellt, wie das Verfahren der Merkmalsextraktionen auch für Chromatogramm-Spektren bei einem GC-gekoppelten IMS anzuwenden ist.

Mit Abschluss dieser Arbeit steht eine Methode zur Verfügung, gasförmige Gefahrstoffe mit Hilfe eines Detektorenarrays zu identifizieren. Die Methode ist als Softwaremodul in dem Messsystem DACHS umgesetzt. Eine Musterdatenbank mit fünf der wichtigsten Gefahrstoffe ist einsatzbereit. Die Musterdatenbank ist durch weitere Messungen erweiterbar um das gesamte Spektrum der nach ETW und AEGL bewerteten Gefahrstoffe identifizieren zu können. Mit der Random-Forest-Methode wurde ein optimales Verfahren zur automatischen Stoffbestimmung gefunden, da es zugleich gute Klassifikationsergebnisse liefert und wenig Rechenleistung benötigt.

Es ist zu erwarten, dass neue Sensortechnologien in die weitere Entwicklung von heterogenen Sensorenarrays einfließen. Mit diesen können durch die zu erwartende höhere Spe-



## 7. Zusammenfassung

zifität und Nachweisstärke der Sensoren die Identifikations- und Detektionseigenschaften eines Detektorenarrays gesteigert werden. Gleichzeitig können neue Elektrochemische Zellen z.B. durch den Einsatz von Festkörperelektrolyten zukünftig eine geringere Degradation aufweisen. Durch die damit einhergehende verbesserte Langzeitstabilität der Sensorsignale kann somit auch das Klassifikationsverfahren, aufgrund der verringerten Unsicherheit der Eingangsmerkmale, besser arbeiten.

Es ist ebenfalls absehbar, dass die Stabilität des eingesetzten Ionenmobilitätsspektrometers optimiert wird. Für den in der Arbeit dargelegten Systemaufbau stand ein Funktionsmuster zur Verfügung, dessen Signalqualität durch Oberflächenladungen über die Zeit abgenommen hat. Für zukünftige Arbeiten stehen Doppelröhren-Ionenmobilitätsspektrometer bereit, welche diesen Effekt durch eine optimierte Materialauswahl der Driftstrecke, nicht mehr aufweisen.

Sowohl die in dieser Arbeit entwickelte Methode der Merkmalsextraktion und der Klassifikation durch das Random-Forest-Verfahren als auch das umgesetzte Messgerät sind Fortschritte für die Weiterentwicklung von Messinstrumenten zur Detektion und Identifikation von Gefahrstoffen. Die untersuchten Prinzipien und Systeme können für den Einsatz bei Feuerwehren und im Katastrophenschutz ertüchtigt werden. Damit stehen optimierte nachweisstarke Messinstrumente für die Gefahrenabwehr zur Verfügung.

# A. Anhang

## A.1. Liste relevanter gefährlicher Stoffe

Tabelle A.1.: Stoffliste nach vfdb 10/05 mit Einsatztoleranzwerten und Bewertung nach AEGL (*Acute Exposure Guideline Levels*) (Stoffe A-D) [5] [6]

| Substanzbezeichnung          | Konzentration in ppm |        |      |      |        |      |
|------------------------------|----------------------|--------|------|------|--------|------|
|                              | ETW                  | AEGL 1 |      |      | AEGL 2 |      |
|                              |                      | 1 h    | 4 h  | 8 h  | 4 h    | 8 h  |
| Aceton                       | 500                  | 200    | 200  | 3200 | 1400   | 95   |
| Acrolein                     | 0,1                  | 0,03   | 0,03 | 0,1  | 0,1    | 0,1  |
| Acrylnitril                  | 20                   |        |      | 1,7  | 0,48   | 0,26 |
| Ammoniak                     | 110                  | 30     | 30   | 160  | 110    | 110  |
| Anilin                       | 3                    | 8      | 1    | 12   | 3      | 1,5  |
| Arsenwasserstoff             | 0,04                 |        |      | 0,17 | 0,04   | 0,02 |
| Benzol                       | 20                   | 52     | 9    | 800  | 400    | 200  |
| Carbonylchlorid (Phosgen)    | 0,08                 |        |      | 0,3  | 0,08   | 0,04 |
| Chlor                        | 1                    | 0,5    | 0,5  | 2    | 1      | 0,71 |
| Chlorbenzol                  | 100                  | 10     | 10   | 150  | 150    | 150  |
| Chlorcyan                    | 0,3                  |        |      |      |        |      |
| Chloroform                   | 44                   |        |      | 64   | 40     | 29   |
| Chlorwasserstoff (Salzsäure) | 5,4                  | 1,8    | 1,8  | 22   | 11     | 11   |
| Cyanwasserstoff (Blausäure)  | 3,5                  | 2      | 1    | 7,1  | 3,5    | 2,5  |
| Cyclohexylamin               | 5,4                  | 1,8    | 1,8  | 8,6  | 5,4    | 2,7  |



## A. Anhang

Tabelle A.2.: Stoffliste nach vfdb 10/05 mit Einsatztoleranzwerten und Bewertung nach AEGL (*Acute Exposure Guideline Levels*) (Stoffe E-Z)

| Substanzbezeichnung                       | Konzentration in ppm |        |      |      |        |      |
|---|----------------------|--------|------|------|--------|------|
|   | ETW                  | AEGL 1 |      |      | AEGL 2 |      |
|   |                      | 1 h    | 4 h  | 8 h  | 4 h    | 8 h  |
| Epichlorhydrin                            | 16                   | 1,7    | 1,7  | 24   | 14     | 6,7  |
| Essigsäure                                | 20                   |        |      |      |        |      |
| Ethanol                                   | 3000                 |        |      |      |        |      |
| Ethylendiamin                             | 6,1                  |        |      | 9,7  | 6,1    | 4,8  |
| Ethylenoxid                               | 14                   |        |      | 45   | 14     | 7,9  |
| Fluor                                     | 2                    | 1,7    | 1,7  | 5    | 2,3    | 2,3  |
| Fluorwasserstoff                          | 12                   | 1      | 1    | 24   | 12     | 12   |
| Formaldehyd                               | 1                    | 0,9    | 0,9  | 14   | 14     | 14   |
| n-Hexan                                   | 200                  |        |      | 2900 | 2900   | 2900 |
| Hydrazin                                  | 3,1                  | 0,1    | 0,1  | 13   | 3,1    | 1,6  |
| Kohlenstoffdioxid                         | 10000                |        |      |      |        |      |
| Kohlenstoffdisulfid (Schwefelkohlenstoff) | 10                   | 13     | 6,7  | 160  | 100    | 50   |
| Kohlenstoffmonoxid                        | 33                   |        |      | 83   | 33     | 27   |
| Methanol                                  | 720                  | 530    | 270  | 2100 | 730    | 520  |
| Methylmercaptan                           | 3                    |        |      | 23   | 14     | 7,3  |
| Phosphorwasserstoff (Phosphin)            | 0,5                  |        |      | 2    | 0,5    | 0,25 |
| Salpetersäure                             | 3                    | 0,16   | 0,16 | 24   | 6      | 3    |
| Schwefeldioxid                            | 1                    | 0,2    | 0,2  | 0,75 | 0,75   | 0,75 |
| Schwefelwasserstoff                       | 20                   | 0,51   | 0,33 | 27   | 20     | 17   |
| Stickstoffdioxid                          | 8,2                  | 0,5    | 0,5  | 12   | 8,2    | 6,7  |
| Styrol                                    | 40                   | 20     | 20   | 130  | 130    | 130  |
| Tetrachlorethen                           | 120                  | 35     | 35   | 230  | 120    | 81   |
| Tetrachlorkohlenstoff                     | 39                   |        |      | 13   | 7,6    | 5,8  |
| Toluol                                    | 94                   | 67     | 67   | 560  | 310    | 250  |
| Toluoldiisocyanat                         | 0,021                |        |      |      |        |      |
| 1, 1, 1-Trichlorethan                     | 380                  | 230    | 230  | 600  | 380    | 310  |
| 1, 1, 2-Trichlorethan                     | 25                   |        |      |      |        |      |
| Trichlorethen                             | 100                  | 130    | 77   | 450  | 270    | 240  |
| Vinylchlorid                              | 100                  | 250    | 70   | 1200 | 820    | 820  |

## A.2. Dateiformat der Musterdatenbank

```
<Root>
  <DatabaseInfo>
    <name> Musterdatenbank 1 </name>
    <version> 1.0 </version>
    <trees> 2 </trees>
  </DatabaseInfo>
  <Stoffe>
    <anzahl> 2 </anzahl>
    <stoff1> BeispielsubstanzA </stoff1>
    <stoff2> Null </stoff2>
  </Stoffe>
```

Abbildung A.1.: Kopf der Klassifikatordatei

```
    <node10>
      <substance> BeispielsubstanzA </substance>
      <substance_id> 1 </substance_id>
      <prob> 71.000000 </prob>
    </node10>
  </Tree1>
  ...
</Trees>
</Root>
```

Abbildung A.2.: Ergebnisknoten einzelner Entscheidungsbäume in der Klassifikatordatei

### A.3. Fehlerfortpflanzung der K<sub>0</sub>-Wert-Berechnung

$$\Delta K_0 = \sqrt{A^2 + B^2 + C^2 + D^2 + E^2} \quad (\text{A.1})$$

$$A = \frac{\partial K_0}{\partial P_D} \cdot \Delta P_D = \frac{k l_d^2}{T_D U t_d} \cdot \Delta P_D \quad (\text{A.2})$$

$$B = \frac{\partial K_0}{\partial T_D} \cdot \Delta T_D = -\frac{P_D k l_d^2}{T_D^2 U t_d} \cdot \Delta T_D \quad (\text{A.3})$$

$$C = \frac{\partial K_0}{\partial U} \cdot \Delta U = -\frac{P_D k l_d^2}{T_D U^2 t_d} \cdot \Delta U \quad (\text{A.4})$$

$$D = \frac{\partial K_0}{\partial t_d} \cdot \Delta t_d = -\frac{P_D k l_d^2}{T_D U t_d^2} \cdot \Delta t_d \quad (\text{A.5})$$

$$E = \frac{\partial K_0}{\partial l_d} \cdot \Delta l_d = \frac{2 P_D k l_d}{T_D U t_d} \cdot \Delta l_d \quad (\text{A.6})$$

### A.4. Bewertung der Random Forest Methode zur Chromatogramm-Auswertung

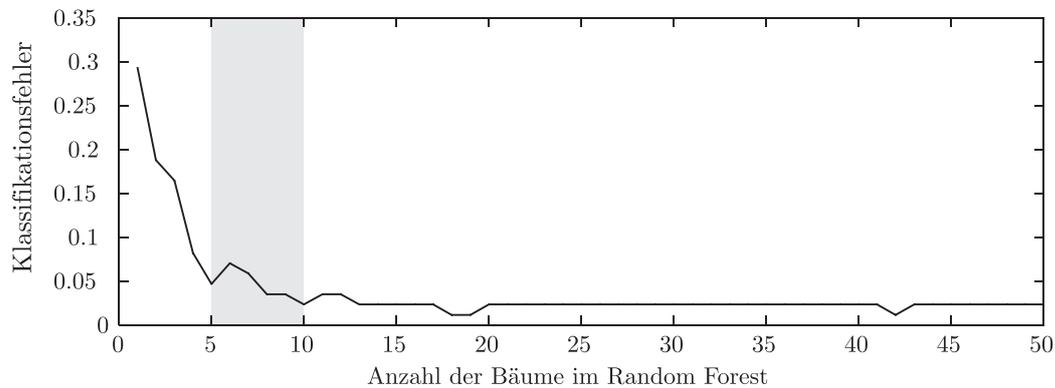


Abbildung A.3.: Reklassifizierungsfähigkeit abhängig von Größe des Random Forest



# Literatur

- [1] G. Matz, A. Schillings und P. Rechenbach. *Task Force für die Schnellanalytik bei großen Chemieunfällen und Bränden*. Bd. N.F., 49. Zivilschutzforschung. Bonn: Bundesverwaltungsamt Zentralstelle für Zivilschutz, 2003.
- [2] *Jahresbericht / Zentrale Melde- und Auswertestelle für Störfälle und Störungen in verfahrenstechnischen Anlagen, ZEMA: Jahresbericht 2010*. Dessau-Roßlau: Umweltbundesamt, 2012.
- [3] Gefahrgutunfälle mit Personenschaden. *Brand und Explosion an Bord der MSC FLAMINIA am 14. Juli 2012 auf dem Atlantik und die sich daran anschließenden Ereignisse: Untersuchungsbericht 255/12*. Hrsg. von Gefahrgutunfälle mit Personenschaden. Hamburg, 28.02.2014.
- [4] Statistisches Bundesamt, Hrsg. *Gefahrguttransporte*. Bd. Fachserie 8 Reihe 1.4. Verkehr. Wiesbaden, 2011.
- [5] *Bewertung von Schadstoffkonzentrationen im Feuerwehreinsatz*. Bd. 10,1. vfdb-Richtlinien. Köln: VdS Schadenverhütung-Verlag, 2005.
- [6] US Environmental Protection Agency. *Acute Exposure Guideline Levels (Compiled Version): Final AEGLs*. Hrsg. von US Environmental Protection Agency. Washington, D.C., 2015.
- [7] Zentralstelle für Zivilschutz. *Technische Kurzbeschreibung des CBRN-Erkunders: Der ABC - Erkundungs-Kraftwagen: Eine technische Kurzbeschreibung*. Hrsg. von Zentralstelle für Zivilschutz. Bonn, 2005.



- [8] T. Albrecht, G. Matz und J. Müller. „Gassensoren-Array mit Mustererkennung zur Bestimmung von Gefahrstoffen in Luft“. Diss. Aachen und Hamburg-Harburg: Technische Universität Hamburg-Harburg, 1995.
- [9] T. Hunte. „Gas-Detektoren-Arrays mit Meßbereichsanpassung zur Erfassung von Gefahrstoffen bei Chemieunfällen und Emissionen“. Diss. Düsseldorf und Hamburg-Harburg: Techn. Univ., Arbeitsbereich Meßtechnik und Technische Universität Hamburg, 1999.
- [10] G. Matz, H. Fischer und J. Frank. *Detektoren Array mit Gaschromatograph zur Identifikation toxischer Substanzen: DACHS ; Abschlussbericht ; Laufzeit des Vorhabens 01.12.2007 bis 31.12.2011*. Hamburg-Harburg: TUHH, 2012.
- [11] B. Ungethüm. „Gitterloses Bipolar-Ionen-Mobilitäts-Spektrometer für die Spurengas-Analyse“. Diss. Wismar und Hamburg-Harburg: Techn. Univ., Institut für Messtechnik und Institut für Messtechnik, 2012.
- [12] M. Drobig. *Schlussbericht Detektoren-Array mit Gaschromatograph zur Identifikation toxischer Substanzen ( DACHS ) Förderkennzeichen 13N9532*. Bonn, 2012.
- [13] G. H. Dodd und D. J. Squirrel, Hrsg. *Structure and mechanism in the mammalian olfactory system*. Bd. 45. 1980.
- [14] K. Persaud und G. Dodd. „Analysis of discrimination mechanisms in the mammalian olfactory system using a model nose“. In: *Nature* 299.5881 (1982), S. 352–355. DOI: 10.1038/299352a0.
- [15] Rolf-Rainer Grigat. „Pattern Recognition“. Hamburg, 09.04.2014.
- [16] C. M. Bishop. *Pattern Recognition and Machine Learning // Pattern recognition and machine learning*. New York, NY: Springer Science + Business Media LLC, 2006.
- [17] J. W. Gardner und P. N. Bartlett. „A brief history of electronic noses“. In: *Sensors and Actuators B* 1994.18-19 (1994), S. 211–220.



- [18] J. A. Ragazzo-Sanchez, P. CHALIER, D. CHEVALIER, M. Calderon-Santoyo und C. GHOMMIDH. „Identification of different alcoholic beverages by electronic nose coupled to GC“. In: *Sensors and Actuators B: Chemical* 134.1 (2008), S. 43–48. DOI: 10.1016/j.snb.2008.04.006.
- [19] *Grundbegriffe: Basic terminology*. Januar 1995. Bd. ; Teil 1. Deutsche Norm. Berlin: Beuth, 1995.
- [20] W. Jessel. *Gase - Dämpfe - Gasmesstechnik: Ein Kompendium für die Praxis*. Lübeck: Dräger Safety AG, 2001.
- [21] C.-D. Kohl und T. Wagner, Hrsg. *Gas Sensing Fundamentals*. Aufl. 2014. Bd. 15. Springer Series on Chemical Sensors and Biosensors. Berlin: Springer Berlin, 2014.
- [22] P. K. Clifford und D. T. Tuma. „Characteristics of semiconductor gas sensors I. Steady state gas response“. In: *Sensors and Actuators* 3 (1982), S. 233–254. DOI: 10.1016/0250-6874(82)80026-7.
- [23] P. Gründler. *Chemical sensors: An introduction for scientists and engineers ; with 25 tables*. Berlin, Heidelberg: Springer, 2007.
- [24] P. T. Moseley. *New trends and future prospects of thick- and thin-film gas sensors*. Hrsg. von Sensors and Actuators: B. Chemical, 1991, Vol.3, pp.167-174. 1991. DOI: 10.1016/0925-4005(91)80002-2.
- [25] J. F. McAleer, P. T. Moseley, Norris, John O. W., D. E. Williams und B. C. Tofield. „Tin dioxide gas sensors. Part 2.—The role of surface additives“. In: *Journal of the Chemical Society, Faraday Transactions 1: Physical Chemistry in Condensed Phases* 84.2 (1988), S. 441. DOI: 10.1039/F19888400441.
- [26] R. H. Zimmermann, E. Schramm und M. Sklorz. „Method for detection of organic trace components on surfaces of carrier, involves detecting molecules ion-spectrometrically, where molecules of trace components of surface are laser desorped with laser source“. DE102007045361 (A1). 2009.
- [27] J. E. Lovelock. „Ionization Methods for the Analysis of Gases and Vapors“. In: *Analytical Chemistry* 33.2 (1961), S. 162–178. DOI: 10.1021/ac60170a003.



- [28] J. N. Driscoll und J. B. Clarici. „Ein neuer Photoionisationsdetektor für die Gas-Chromatographie“. In: *Chromatographia* 9.11 (1976), S. 567–570. DOI: 10.1007/BF02275964.
- [29] msscientific Chromatographie-Handel GmbH. *PID-Lampen (Photoionisationslampen) von Heraeus Noblelight*. Hrsg. von msscientific Chromatographie-Handel GmbH. Berlin, 2011.
- [30] Heraeus Noblelight GmbH. *Photoionisation Detector Lamps (PID): High Quality PID Lamps from Heraeus Noblelight*. Hrsg. von Heraeus Noblelight GmbH. Hanau, 2012.
- [31] D. R. Lide, Hrsg. *CRC handbook of chemistry and physics: A ready-reference book of chemical and physical data*. 86. ed. Boca Raton: CRC Press, 2005.
- [32] J. Chou. *Hazardous gas monitors: A practical guide to selection, operation and applications*. New York, NY: McGraw-Hill, 2000.
- [33] Price, John G. W., D. C. Fenimore, P. G. Simmonds und A. Zlatkis. „Design and operation of a photoionization detector for gas chromatography“. In: *Analytical Chemistry* 40.3 (1968), S. 541–547. DOI: 10.1021/ac60259a013.
- [34] H. Böhm. „Elektrochemische Gassensoren und Gasanalysengeräte / Electrochemical sensors for gas analysis“. In: *tm - Technisches Messen* 50.JG (1983). DOI: 10.1524/teme.1983.50.jg.399.
- [35] D. Kitzelmann und C. Gottschalk. „Elektrochemische Gassensoren - Wirkungsweisen und Möglichkeiten zur Funktionsüberwachung“. In: *tm - Technisches Messen* 62.JG (1995). DOI: 10.1524/teme.1995.62.jg.152.
- [36] H. Schaumburg und D. H. Althen, Hrsg. *Sensoranwendungen: Mit 37 Tabellen*. Bd. 8. Werkstoffe und Bauelemente der Elektrotechnik. Stuttgart: Teubner, 1995.
- [37] W. Göpel. „Entwicklung chemischer Sensoren: Empirische Kunst oder systematische Forschung?/ Development of chemical sensors: empirical art or systematic research?“. In: *tm - Technisches Messen* 52.3 (1985). DOI: 10.1524/teme.1985.52.3.92.



- [38] H. E. Revercomb und E. A. Mason. „Theory of plasma chromatography/gaseous electrophoresis. Review“. In: *Analytical Chemistry* 47.7 (1975), S. 970–983. DOI: 10.1021/ac60357a043.
- [39] M. Klein, Hanley, H. J. M. und F. J. Smith. *Tables of collision integrals and second virial coefficients for the (m,6,8,) intermolecular potential function*. Bd. 47. National standard reference data series. Washington, D.C.: U.S. Gov. Print. Off, 1974.
- [40] N. Bradbury und R. Nielsen. „Absolute Values of the Electron Mobility in Hydrogen“. In: *Physical Review* 49.5 (1936), S. 388–393. DOI: 10.1103/PhysRev.49.388.
- [41] G. A. Eiceman und Z. Karpas. *Ion mobility spectrometry*. Boca Raton: CRC Press, 1993.
- [42] A. Walte, W. Münchmeyer und B. Ungethüm. „Verfahren und Vorrichtung zur Detektion und Identifizierung von Gasen mittels Ionenmobilitätsspektrometrie“. US2013161509 (A1). 2013.
- [43] Norman E. Holden. „Table of the Isotopes“. In: *CRC handbook of chemistry and physics*. Hrsg. von D. R. Lide. Boca Raton: CRC Press, 2005, 11.50–11.201.
- [44] A. Good, D. A. Durden und P. Kebarle. *Ion–Molecule Reactions in Pure Nitrogen and Nitrogen Containing Traces of Water at Total Pressures 0.5–4 torr. Kinetics of Clustering Reactions Forming  $H+(H_2O)_n$* . 2003-09-08.
- [45] H. Snijders, H.-G. Janssen und C. Cramers. „Optimization of temperature-programmed gas chromatographic separations I. Prediction of retention times and peak widths from retention indices“. In: *Journal of Chromatography A* 718.2 (1995), S. 339–355. DOI: 10.1016/0021-9673(95)00692-3.
- [46] J. Curvers, J. Rijks, C. Cramers, K. Knauss und P. Larson. „Temperature programmed retention indices: Calculation from isothermal data. Part 1: Theory“. In: *Journal of High Resolution Chromatography* 8.9 (1985), S. 607–610. DOI: 10.1002/jhrc.1240080926.



- [47] Y. Guan, J. Kiraly und J. A. Rijks. „Interactive retention index database for compound identification in temperature-programmed capillary gas chromatography“. In: *Journal of Chromatography A* 472 (1989), S. 129–143. DOI: 10.1016/S0021-9673(00)94101-3.
- [48] K. Dettmer-Wilde und W. Engewald, Hrsg. *Practical Gas Chromatography: A Comprehensive Reference*. Berlin, Heidelberg und s.l.: Springer Berlin Heidelberg, 2014. DOI: 10.1007/978-3-642-54640-2.
- [49] A. R. Mastrogiacomo, E. Pierini und L. Sampaolo. „A comparison of the critical parameters of some adsorbents employed in trapping and thermal desorption of organic pollutants“. In: *Chromatographia* 41.9-10 (1995), S. 599–604. DOI: 10.1007/BF02269726.
- [50] K. Figge, W. Rabel und A. Wieck. „Adsorptionsmittel zur Anreicherung von organischen Luftinhaltsstoffen“. In: *Z. Anal. Chem. (Fresenius' Zeitschrift für Analytische Chemie)* 327.3-4 (1987), S. 261–278. DOI: 10.1007/BF00491828.
- [51] J. Frank, H. Fischer, I. Radev, A. Baermann und G. Matz. „GC-PID for In-Situ Soil Investigation“. In: Pittcon 2014. Chicago, 2014.
- [52] A. Handl. *Multivariate Analysemethoden: Theorie und Praxis multivariater Verfahren unter besonderer Berücksichtigung von S-Plus*. 2. Aufl. Statistik und ihre Anwendungen. Heidelberg u. a.: Springer-Verlag Berlin Heidelberg, 2010.
- [53] T. Eklöv, P. Mårtensson und I. Lundström. „Enhanced selectivity of MOSFET gas sensors by systematical analysis of transient parameters“. In: *Analytica chimica acta : an international journal devoted to all branches of analytical chemistry* 353.2-3 (1997), S. 291–300.
- [54] G. Matz, H. Fischer und J. Frank. *Ertüchtigung eines Detektoren Arrays mit Gaschromatograph für Spezialeinheiten zur mobilen Analyse: FP 395 „Weiterentwicklung GDA“*. Hamburg-Harburg: TUHH, 2014.
- [55] T. Woldert. *Untersuchung und Implementierung von Mustererkennungsverfahren für die Datenauswertung eines Gassensorenarrays: Masterarbeit*. Hamburg, 17.02.2015.



- [56] L. Fahrmeier und A. Hamerle. *Multivariate statistische Verfahren*. 1. Aufl. Berlin: Walter de Gruyter, 1984.
- [57] L. Breiman. *Classification and regression trees*. Repr. Boca Raton: Chapman & Hall, 1998.
- [58] J. H. Cho und P. U. Kurup. „Decision tree approach for classification and dimensionality reduction of electronic nose data“. In: *Sensors and Actuators B* 2011.160 (2011), S. 542–548.
- [59] E. Alpaydm und S. Linke. *Maschinelles Lernen*. München: Oldenbourg, 2008.
- [60] T. K. Ho, J. J. Hull und S. N. Srihari. „Decision combination in multiple classifier systems“. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 16.1 (1994), S. 66–75. DOI: 10.1109/34.273716.
- [61] L. Breiman. „Random Forests“. In: *Machine learning* 45.1 (2001), S. 5–32.
- [62] T. K. Ho. „Random decision forests“. In: *3rd International Conference on Document Analysis and Recognition*. 14-16 Aug. 1995, S. 278–282. DOI: 10.1109/ICDAR.1995.598994.
- [63] L. Breiman. „Bagging Predictors“. In: *Machine learning* 24.2 (1996), S. 123–140.
- [64] W. Kördel. *Untersuchungen zum Vollzug und zur Weiterentwicklung der Bundes-Bodenschutz- und Altlastenverordnung - Teil 1*. Bd. 2011,68. Texte / Umweltbundesamt. Dessau-Roßlau: Umweltbundesamt, 2011.
- [65] J. Frank, H. Fischer und G. Matz. „Effect of Temperature and Flow variation for a short Multi Capillary Column - PID“. In: *Pittcon 2013*. Philadelphia, PA, 2013.
- [66] W. Schröder, G. Matz und J. Kübler. „Fast detection of preservatives on waste wood with GC/MS, GC-ECD and ion mobility spectrometry“. In: *Field Analytical Chemistry & Technology* 2.5 (1998), S. 287–297.
- [67] W. S. Cooke. „Multicapillary columns: An idea whose time has come“. In: *Today's Chemist at Work* 5 (1996), S. 14–22.



- [68] Sigma-Aldrich Co. LLC. *GC Column Selection Guide: Achieve Optimal Method Performance*. Hrsg. von Sigma-Aldrich Co. LLC. St. Louis, 2013.
- [69] Alphasense Ltd. *PID-AH Photo Ionisation Detector: Datenblatt*. Hrsg. von Alphasense Ltd. Great Notley, UK, 2009.
- [70] Alphasense Ltd. *VOC Correction Factor (10.6eV)*. Hrsg. von Alphasense Ltd. Great Notley, UK, 2010.
- [71] Deutsches Institut für Normung. *Chemische Analytik: Nachweis-, Erfassungs- und Bestimmungsgrenze unter Wiederholbedingungen - Begriffe, Verfahren, Auswertung = Chemical analysis : Decision limit, detection limit and determination limit under repeatability conditions - terms, methods, evaluation*. Berlin.
- [72] Membrapor AG. *Specifications Sheet for Formaldehyde-Sensor Type CH20/M-10*. Hrsg. von Membrapor AG. Zürich, 2008.
- [73] Compur Monitors GmbH & Co KG. *Statorx 501 / S Messkopf: Technische Dokumentation*. Hrsg. von Compur Monitors GmbH & Co KG. München, 2012.
- [74] O. Kiesewetter, K. Roth und A. Reinholz. „Semiconductor gas sensor“. DE19606272 (A1). 1997.
- [75] UST Umweltsensortechnik GmbH. *MOX gas sensor element GGS 5430 T*. Hrsg. von UST Umweltsensortechnik GmbH. Geschwenda, 2013.
- [76] UST Umweltsensortechnik GmbH. *MOX gas sensor element GGS 2530 T*. Hrsg. von UST Umweltsensortechnik GmbH. Geschwenda, 2009.
- [77] P. L. Dimatteo. „Method of generating and constructing three-dimensional bodies“. US3932923 (A). 1976.
- [78] J. Kruth, L. Froyen, J. Van Vaerenbergh, P. Mercelis, M. Rombouts und B. Lauwers. „Selective laser melting of iron-based powder“. In: *Journal of materials processing technology* 149.1 (2004), S. 616–622.
- [79] Membrapor AG. *Application Notes*. Hrsg. von Membrapor AG. Zürich, 2004.



- [80] H.-R. Tränkler und G. Fischerauer. „Das Ingenieurwissen Messtechnik“. In: *HÜTTE : das Ingenieurwissen*, 34. Aufl. (2014). DOI: 10.1007/978-3-662-44030-8.
- [81] I. Lundström, M. Armgarth und L.-G. Petersson. „Physics with catalytic metal gate chemical sensors“. In: *Critical Reviews in Solid State and Materials Sciences* 15.3 (1989), S. 201–278. DOI: 10.1080/10408438908243446.
- [82] D. Vlachos, D. Fragoulis und J. Avaritsiotis. „An adaptive neural network topology for degradation compensation of thin film tin oxide gas sensors“. In: *Sensors and Actuators B: Chemical* 45.3 (1997), S. 223–228. DOI: 10.1016/S0925-4005(97)00309-2.
- [83] H. Matthiessen. „Method for sensitizing and stabilizing the operating condition of a semiconductor gas sensor and a semiconductor gas sensor device therefor“. US4663958 (A). 1987.
- [84] E. Kováts. „Gas-chromatographische Charakterisierung organischer Verbindungen. Teil 1: Retentionsindices aliphatischer Halogenide, Alkohole, Aldehyde und Ketone“. In: *Helvetica Chimica Acta* 41.7 (1958), S. 1915–1932. DOI: 10.1002/hlca.19580410703.
- [85] A. Savitzky und Golay, M. J. E. „Smoothing and Differentiation of Data by Simplified Least Squares Procedures“. In: *Analytical Chemistry* 36.8 (1964), S. 1627–1639. DOI: 10.1021/ac60214a047.
- [86] H. H. Madden. „Comments on the Savitzky-Golay convolution method for least-squares-fit smoothing and differentiation of digital data“. In: *Analytical Chemistry* 50.9 (1978), S. 1383–1386. DOI: 10.1021/ac50031a048.
- [87] S. Kromidas und H.-J. Kuss, Hrsg. *Chromatogramme richtig integrieren und bewerten: Ein Praxishandbuch für die HPLC und GC*. Weinheim: Wiley-VCH, 2008.
- [88] S. Bader, W. Urfer und J. Baumbach. „Preprocessing of ion mobility spectra by lognormal detailing and wavelet transform“. In: *International Journal for Ion*



- Mobility Spectrometry* 11.1-4 (2008), S. 43–49. DOI: 10.1007/s12127-008-0005-6.
- [89] P. Núñez Díaz. *Noise Reduction of Ion Mobility Spectrometer Signals Using the Wavelet Transform: Project Work*. Hamburg, 2009.
- [90] J. Dohrn. *Mikrocontroller gestützte Ansteuerung eines neuartigen Ionenmobilitätsspektrometers: Master Thesis*. 2011.
- [91] J. Frank, H. Fischer, B. Ungethüm, A. Walte und G. Matz. „Effect of Sample Gas Humidity on Detector Arrays“. In: Pittcon 2014. Chicago, 2014.
- [92] S. St. Laurent und M. Fitzgerald. *XML pocket reference*. 3rd ed. Pocket Reference (O’Reilly). Sebastopol, Calif.: O’Reilly, 2005.
- [93] H. Reß und G. Viebeck. *Datenstrukturen und Algorithmen in C++*. 2., überarb. Aufl. München: Hanser, 2003.
- [94] T. Fawcett. „An introduction to ROC analysis“. In: *ROC Analysis in Pattern Recognition* 27.8 (2006), S. 861–874. DOI: 10.1016/j.patrec.2005.10.010.
- [95] I. Törmä. *DX4000 FTIR gas analyzer: Technical Data*. Hrsg. von Gaset Technologies Oy. Helsinki.
- [96] T. Wròblewski, L. Ziemczonek und G. P. Karwasz. „Proton transfer reactions for ionized water clusters“. In: *Czechoslovak Journal of Physics* 54.S3 (2004), S. C747–C752.
- [97] *Auswertung von Messungen, Meßunsicherheit: Evaluation of measurements, uncertainty of measurement = Exploitation des mesurages, incertitude de mesure : DIN 1319-4*. Februar 1999. Bd. ; Teil 4. Grundbegriffe der Meßtechnik. Berlin: Beuth, 1999.
- [98] B. Efron und R. J. Tibshirani. *An introduction to the bootstrap*. Bd. 57. Monographs on statistics and applied probability. New York: Chapman & Hall, 1993.



- [99] M. R. Chernick. *Bootstrap methods: A guide for practitioners and researchers*. 2. ed. Wiley series in probability and statistics. Hoboken, NJ: Wiley-Interscience, 2008. DOI: 10.1002/9780470192573.
- [100] R. B. Cattell. „The Scree Test For The Number Of Factors“. In: *Multivariate Behavioral Research* 1.2 (1966), S. 245–276. DOI: 10.1207/s15327906mbr0102{ }10.
- [101] J. Frank, H. Fischer, B. Ungethüm und G. Matz. „Data Analysis of a short Multi-Capillary-Column coupled Ion Mobility Spectrometer“. In: Pittcon 2013. Philadelphia, PA, 2013.
- [102] A.-C. Hauschild, D. Kopczynski, M. D’Addario, J. I. Baumbach, S. Rahmann und J. Baumbach. „Peak detection method evaluation for ion mobility spectrometry by using machine learning approaches“. In: *Metabolites* 3.2 (2013), S. 277–293. DOI: 10.3390/metabo3020277.
- [103] M. Manke. *Mehrdimensionale Signalanalyse für GC-Systeme: Diplomarbeit*. Hamburg, 28. Oktober 2010.
- [104] B. Bödeker, W. Vautz und J. Baumbach. „Peak finding and referencing in MCC/IMS-data“. In: *International Journal for Ion Mobility Spectrometry* 11.1-4 (2008), S. 83–87. DOI: 10.1007/s12127-008-0012-7.



## Lebenslauf

|                          |   |
|--------------------------|---|
| <b>Name</b>              | Frank   |
| <b>Vorname</b>           | Jörn  |
| <b>Geburtsdatum</b>      | 03.12.1982  |
| <b>Geburtsort, -land</b> | Neustadt i.H., Deutschland  |
| <b>08.1989 – 07.1993</b> | Grundschule Lensahn   |
| <b>08.1993 – 07.1999</b> | Realschule Lensahn  |
| <b>08.1999 – 07.2002</b> | Fachgymnasium Oldenburg i.H.<br>Abschluss: Allgemeine Hochschulreife  |
| <b>10.2002 – 02.2008</b> | Studium der Elektrotechnik an der<br>Technischen Universität Hamburg-Harburg<br>Abschluss: Diplom                                       |
| <b>02.2008 – 11.2015</b> | wissenschaftlicher Mitarbeiter und Gastwissenschaftler am<br>Institut für Messtechnik an der Technischen Universität<br>Hamburg-Harburg |
| <b>seit 11.2015</b>      | Laser-Produktingenieur, Alltec GmbH   |





