

Institut für Parallele und Verteilte Systeme
Universität Stuttgart
Universitätsstraße 38
D-70569 Stuttgart

Diplomarbeit Nr. 3192

Lokalisierung von Unterwasserrobotern mit Hilfe von elektrischen Feldern

Christian Stürmer

Studiengang:	Informatik
Prüfer:	Prof. Dr. rer. nat. habil. Paul Levi
Betreuer:	Dipl.-Ing. Tobias Dipper
begonnen am:	17. Mai 2011
beendet am:	16. November 2011
CR-Klassifikation:	C.3, I.2.9, J.2, J.7



Zusammenfassung

In dieser Arbeit wird ein Verfahren vorgestellt, mit dessen Hilfe es möglich ist, elektrische Spannungsquellen unter Wasser anhand im Wasser gemessener elektrischer Potentialunterschiede, sowie weiterer Informationen, die einem Roboter zur Verfügung stehen könnten – wie z.B. der Umgebungsdruck – zu lokalisieren. Grundlage des Verfahrens ist die mathematische Modellierung der Ausbreitung von elektrischen Feldern unter Wasser, sowie die Kombination der Inversen Problemtheorie mit dem Downhill-Optimierungsverfahren nach Nelder und Mead. In der Arbeit werden verschiedene Ansätze von physikalischen Modellen betrachtet, in einer Simulationsumgebung in Matlab implementiert und verglichen. Abschließend wird eine Implementierung des Verfahrens anhand eines praktischen Versuches getestet und die Versuchsergebnisse ausgewertet, sowie ein Ausblick auf mögliche Verbesserungen gegeben.

Abstract

This thesis describes a method of locating electrical current sources under water, based on measuring the differences of the electrical potentials as well as other information, which could be available to a swarm robot. The method is based on a mathematical model of the electrical field distribution under water and the combination of the inverse problem theory and the downhill-optimization algorithm of Nelder and Mead. Different physical models are examined and implemented as well as compared in a simulation environment in Matlab. A real-world implementation of the complete method is tested in an experiment and its results are evaluated. An outlook on possible improvements is given.



D

Inhaltsverzeichnis

Formelverzeichnis	v
1. Einleitung	1
1.1. Aufgabenstellung	2
2. Lokalisierung von Ladungen	5
2.1. Vorbedingungen	5
2.1.1. Reduktion auf die Elektrostatik	5
2.1.2. Die relative Permittivität im betrachteten Raum ist homogen . . .	6
2.1.3. Messungen haben keinen Einfluss auf das Messobjekt	6
2.1.4. Ladungsträger und Messelektroden sind im Modell punktförmig . .	6
2.2. Einführung	6
2.2.1. Physikalische Grundlagen	6
2.2.2. Der geometrische Ansatz	8
2.2.3. Ansatz: Gleichungssystem mit Unbekannten	11
3. Inverse Probleme	15
3.1. Einführung	15
3.2. Modellraum und Datenraum	16
3.2.1. Modellraum	16
3.2.2. Datenraum	16
3.2.3. Definition der Wahrscheinlichkeit	17
3.2.4. Die homogene Wahrscheinlichkeitsverteilung	18
3.2.5. Verknüpfung von Wahrscheinlichkeitsverteilungen	19
3.2.6. Bedingte Wahrscheinlichkeit	20
3.2.7. Rand- und abhängige Wahrscheinlichkeitsdichte	22
3.2.8. Das vorwärtsgerichtete Problem	24
3.3. Messungen und a-priori-Information	25
3.3.1. Messergebnisse	25
3.3.2. a-priori-Information über Modellparameter	27
3.3.3. Verbund-a-priori-Information	27
3.4. Die Lösung des inversen Problems	27
3.4.1. Lösen von inversen Problemen	28
4. Die Lokalisierung eines Dipols als inverses Problem	31

4.1. Modellraum	31
4.2. a-priori-Information Modellraum	32
4.3. Datenraum	33
4.4. a-priori-Information Datenraum	33
4.5. Theoretisches Modell	34
4.6. Zusammensetzen der Lösungsbestandteile	35
5. Umsetzung	37
5.1. Einleitung	37
5.2. Simplex-Verfahren nach Nelder und Mead	39
5.2.1. Übersicht	39
5.2.2. Ablauf des Nelder-Mead-Algorithmus	39
6. Implementierung	45
6.1. Implementierung Simulationsumgebung	45
6.2. Verwendung der Simulationsumgebung	45
6.2.1. Verwendete Parameter	46
6.3. Implementierung in Hardware	48
6.3.1. Messhardware	48
6.3.2. Software	51
7. Versuch	53
7.1. Versuchsaufbau	53
7.2. Erfasste Messdaten	56
7.3. Anwendung des Lokalisierungsverfahrens auf die erfassten Messdaten	63
7.3.1. a-priori-Information	63
7.3.2. Verwendete Simulationsmodelle	63
7.3.3. Überblick über Verarbeitung und Auswertung der Messdaten	65
7.3.4. Auswertung des Versuchs	67
7.4. Bewertung des Verfahrens anhand der Versuchsergebnisse	71
8. Abschließende Bewertung und Ausblick	73
8.1. Zusammenfassung	73
8.2. Mögliche Verbesserungen	74
8.2.1. Physikalisches Modell	74
8.2.2. Messhardware	74
8.2.3. Einbezug der Zeit	75
8.2.4. Nutzung mehrerer Roboter	75
8.2.5. Parallelisierbarkeit zur Beschleunigung	75
A. Simulationsbeispiel	77
B. Messergebnisse und Simulationsdaten	79

B.1. Verwendete Startpunkte für Nelder-Mead	79
B.2. Diagramme und Daten zu den einzelnen Messdatensätzen	79
B.2.1. Messdatensatz 1	79
B.2.2. Messdatensatz 2	80
B.2.3. Messdatensatz 3	81
B.2.4. Messdatensatz 4	82
B.2.5. Messdatensatz 5	83
B.2.6. Messdatensatz 6	84
B.2.7. Messdatensatz 7	85
B.2.8. Messdatensatz 8	86
B.2.9. Messdatensatz 9	87
B.2.10. Messdatensatz 10	88
B.2.11. Messdatensatz 11	89
B.2.12. Messdatensatz 12	90
B.2.13. Messdatensatz 13	91
B.2.14. Messdatensatz 14	92
B.2.15. Messdatensatz 15	93
Glossar	95
Nomenklatur Mathematik	97
Nomenklatur Physik	101
Tabellenverzeichnis	103
Abbildungsverzeichnis	105
Literatur	106
Literaturverzeichnis	107

Formelverzeichnis

2.1.	Gauß'sches Gesetz in Integralform	7
2.2.	Elektrische Flusstärke durch eine geschlossene Kugel	7
2.3.	Elektrische Feldstärke einer Punktladung an einem Punkt in der Entfernung r	8
2.4.	Positionsbestimmung einer Ladung Q mit Hilfe des elektrischen Feldes	9
2.5.	Stärke des elektrischen Feldes einer geladenen Kugel	11
2.6.	Stärke des elektrischen Feldes einer geladenen Kugel mit Faktor k	12
2.7.	Stärke des elektrischen Feldes eines Dipols	12
2.8.	Aus Dipol resultierende Spannung zwischen Elektroden	12
3.1.	Kolmogorovaxiom 1	17
3.2.	Kolmogorovaxiom 2	17
3.3.	Die Wahrscheinlichkeit des leeren Ereignisses ist gleich Null.	17
3.4.	Wahrscheinlichkeit der Vereinigung zweier nicht notwendigerweise disjunkte Ereignisse \mathcal{A} und \mathcal{B}	17
3.5.	Definition: Wahrscheinlichkeitsdichtefunktion	18
3.6.	Definition: Volumen $V(\mathcal{A})$	18
3.7.	Definition: normalisierte Wahrscheinlichkeitsdichte $\mu(\mathbf{x})$	18
3.8.	Definition: homogene Wahrscheinlichkeit $M(\mathcal{A})$	19
3.9.	Wahrscheinlichkeit leeres Volumen	19
3.10.	Definition: volumenbezogene Wahrscheinlichkeit	19
3.11.	Definition: Verknüpfung \wedge von Wahrscheinlichkeitsverteilungen	19
3.12.	Definition: Wahrscheinlichkeitsdichteverteilungen	20
3.13.	Lösung Verknüpfung Wahrscheinlichkeitsdichten	20
3.14.	Definition: Wahrscheinlichkeiten über \mathcal{A}	20
3.15.	Definition: Wahrscheinlichkeitsdichte $\mu_{\mathcal{A}}$	21
3.16.	Verknüpfung von Wahrscheinlichkeitsdichten	21
3.17.	Definition bedingte Wahrscheinlichkeit $P(\mathcal{A} \mathcal{B})$	21
3.18.	Äquivalenz: Wahrscheinlichkeitsdichte verknüpft mit Ereignis – bedingte Wahrscheinlichkeit	21
3.19.	Definition: Koordinaten des Modellraumes \mathfrak{M}	22
3.20.	Definition: Koordinaten des Datenraumes \mathfrak{D}	22
3.21.	Definition: Koordinaten der Mannigfaltigkeit \mathfrak{X}	22
3.22.	Normalisierung Wahrscheinlichkeitsdichtefunktion	22
3.23.	Definition: Randwahrscheinlichkeitsdichten über \mathfrak{D} und \mathfrak{M}	23

3.24. Definition: Normalisierung der Randwahrscheinlichkeitsdichten über \mathfrak{D} und \mathfrak{M}	23
3.25. Unabhängigkeit von \mathbf{d} und \mathbf{m}	23
3.26. Definition: Bedingte Wahrscheinlichkeitsdichte $f_{\mathfrak{D} \mathfrak{M}}(\mathbf{d} \mathbf{m}(\mathbf{d}))$	23
3.27. Definition: Bedingte Wahrscheinlichkeit mit konstanter Bedingung $f_{\mathfrak{D} \mathfrak{M}}(\mathbf{d} \mathbf{m}(\mathbf{d}))$	23
3.28. Definition: Bedingte Wahrscheinlichkeit $f_{\mathfrak{D} \mathfrak{M}}(\mathbf{d} \mathbf{m})$	24
3.29. Definition: Verbundwahrscheinlichkeit $f(\mathbf{d}, \mathbf{m})$	24
3.30. Definition: Vorwärtsgerichtetes Problem	24
3.31. Definition: Koordinaten des Modellraumes \mathfrak{M}	24
3.32. Definition: Koordinaten des Datenraumes \mathfrak{D}	24
3.33. Definition: Koordinaten der Mannigfaltigkeit \mathfrak{X}	25
3.34. Definition: Homogene Wahrscheinlichkeitsdichte \mathfrak{X}	25
3.35. Definition: Verbundwahrscheinlichkeit von Datenraum \mathfrak{D} und Modellraum \mathfrak{M} .	25
3.36. Definition: Wahrscheinlichkeitsdichtefunktion Messinstrument ohne weitere Informationen über die Messdaten	26
3.37. Definition: Wahrscheinlichkeitsdichtefunktion Messinstrument bei linearem Datenraum mit Konstante k	26
3.38. Definition: Wahrscheinlichkeitsdichtefunktion über Datenraum bei gegebenen Eigenschaften $v(\mathbf{d}_{\text{obs}} \mathbf{d})$ des Messinstruments	26
3.39. Definition: Wahrscheinlichkeitsdichtefunktion über Daten bei perfektem Instrument	26
3.40. Definition: Delta-Funktion	26
3.41. Definition: Wahrscheinlichkeitsdichtefunktion über Daten bei Gauß-verteilter Abweichung	27
3.42. Modellierung der „nicht vorhandenen“ a-priori-Information über die Modellparameter	27
3.43. Modellierung der Verbund-a-priori-Information über $\mathfrak{D} \times \mathfrak{M}$	27
3.44. Definition: Lösung des inversen Problems als Wahrscheinlichkeitsdichtefunktion	28
3.45. Definition: a-posteriori-Information über die Modellparameter	28
3.46. Definition: a-posteriori-Information über die Daten	28
3.47. Definition: Theoretische Verbundwahrscheinlichkeit von Datenraum \mathfrak{D} und Modellraum \mathfrak{M} (physikalisches Modell)	28
3.48. Modellierung der Verbund-a-priori-Information über $\mathfrak{D} \times \mathfrak{M}$	28
3.49. Homogener Grenzwert der a-priori-Verbundwahrscheinlichkeit $\rho_D(\mathbf{d}, \mathbf{m})$ über $\mathfrak{D} \times \mathfrak{M}$	29
3.50. Definition: a-posteriori-Information über den Modellraum	29
3.51. Definition: a-posteriori-Information über den Modellraum mit likelihood-Funktion	29
3.52. Definition: likelihood-Funktion	29
3.53. Definition: modellabhängige Wahrscheinlichkeitsdichte $\theta(\mathbf{d} \mathbf{m})$ für geringe Modellierungsunsicherheit	29
3.54. Definition: a-posteriori-Information über den Modellraum für geringe Modellierungsunsicherheit	30

3.55.	Definition: a-posteriori-Information über den Modellraum für geringe Modellierungsunsicherheit mit Proportionalitätsfaktor	30
4.1.	Definition: Spezialfall konstante homogene Wahrscheinlichkeitsdichte über kartesischem Modellraum	32
4.2.	Modellierung der a-priori-Information durch eine homogene Wahrscheinlichkeitsdichte über die Modellparameter bei kartesischem Koordinatensystem.	32
4.3.	Modellierung der a-priori-Information über die Modellparameter bei der Verwendung von Höheninformationen.	33
4.4.	Modellierung der a-priori-Information über die Daten mit Gauß-verteilter Abweichung durch eine Wahrscheinlichkeitsdichtefunktion	34
4.5.	Elektrisches Feld einer geladenen Kugel	34
4.6.	Elektrisches Feld einer geladenen Kugel mit Faktor k	34
4.7.	Durch Dipol erzeugte Spannung zwischen Elektroden	35
4.8.	Definition: $di_3(a, b)$, $a, b \in \mathbb{R}^3$, Abstand in dreidimensionalem kartesischem Raum	35
4.9.	Lösung des inversen Problems als a-posteriori-Wahrscheinlichkeitsdichte $\sigma_M(pQ)$	36
5.1.	Definition von \bar{x}	39
5.2.	Bedingungen der Parameter des Nelder-Mead-Algorithmus	40
5.3.	Belegung der Parameter des Nelder-Mead-Algorithmus	40
5.4.	Definition des Reflexionspunktes x_r	41
5.5.	Definition des Expansionspunktes x_e	41
5.6.	Definition des inneren Kontraktionspunktes x_{cc}	42
5.7.	Definition des äußeren Kontraktionspunktes x_c	43
5.8.	Definition der für die Schrumpfung verwendeten Punktmenge V_i	43
7.1.	Standardabweichung	67

1. Einleitung

Der Hintergrund dieser Arbeit ist die Forschung mit unter Wasser schwimmenden Schwarmrobotern. Aus diesen Voraussetzungen ergeben sich mehrere Einschränkungen, die die Lokalisierung erschweren: Die Roboter bewegen sich in einem dreidimensionalen Raum, was die Navigation im Vergleich zu Robotern, die sich auf einer Ebene bewegen, erschwert. Das Medium Wasser verhindert den Einsatz von GPS (Navstar Global Positioning System)[15, S. 1]. Da die Schwarmroboter relativ einfach aufgebaut sind, verfügen sie nur über einfache Sensorik und keine leistungsstarke Hardware für aufwändige Berechnungen. Allerdings ist die Lokalisierung ein wichtiges Hilfsmittel für viele Aufgaben. Zum Beispiel ist es mittels der Lokalisierung einer feststehenden Marke für einen Roboter möglich, seine eigene Position festzustellen. Falls die Roboter miteinander interagieren sollen, ist es nötig, dass sie ihre relative Position zueinander bestimmen. Da die eingesetzte Sensorik möglichst einfach sein soll, Wasser elektrischen Strom leitet und einer der einfachsten Sensoren ein Sensor ist, der elektrische Spannungen misst, ist eine naheliegende Idee, unter Wasser elektrische Felder zu erzeugen, diese passiv zu messen und aufgrund der erfassten Daten die Senderposition zu bestimmen.

Die Elektrolokation, also die Lokalisierung von Objekten anhand von elektrischen Feldern, ist ein Teilgebiet der Elektrorezeption, die aus der Tierwelt bekannt ist. Genauer untersucht wurde dieses Themengebiet bei schwach elektrischen Fischen [8] [3]. Aufgrund der physikalischen Zusammenhänge lassen sich einige Eigenschaften ableiten. Die elektrische Feldstärke nimmt mit zunehmender Entfernung zur Quelle quadratisch ab [8, S. 1217]. Da die elektrischen Felder außerdem nicht als elektromagnetische Welle abgestrahlt werden, ist die Reichweite stark begrenzt. Eine Verdoppelung der Reichweite würde entweder die vierfache Energiemenge zur Erzeugung des elektrischen Feldes oder eine entsprechende Verbesserung bei der Erfassung der Messdaten und deren Verarbeitung nötig machen.

Die Elektrolokation bietet aber auch einige Vorteile. Der Aufbau der Sensoren ist relativ einfach im Vergleich zu z.B. optischen Sensoren. Die Ausrichtung der Sensoren hat zwar einen Einfluss, aber er ist nicht so gravierend wie bei optischen Sensoren, bei denen eine Sichtverbindung sinnvoll ist. Da das Verfahren auf den Eigenschaften der Leitfähigkeit des Wassers und der Objekte darin basiert, hat es auch den Vorteil, dass es z.B. wenig durch Trübung im Wasser oder Objekten, die ähnliche elektrische Eigenschaften wie Wasser besitzen, beeinträchtigt wird.

Die Eigenschaft des Verfahrens, dass die elektrischen Felder nicht als elektromagnetische Wellen abgestrahlt werden, verringert zwar einerseits die Reichweite, andererseits vereinfacht sie auch die Auswertung der Messdaten, da die speziellen Eigenschaften,

die elektromagnetische Wellen ausmachen und die Komplexität der in der Modellierung verwendeten Elektrostatik übersteigen, nicht berücksichtigt werden müssen.

Fische nutzen passive Elektrolokation ebenfalls, z.B. um Beute zu finden oder um ihr Revier gegen Eindringlinge zu verteidigen. In beiden Fällen ist es nötig, das Objekt zu lokalisieren, um zu ihm zu schwimmen.

Um das Verhalten von Haien bei der Beutesuche zu erklären, wurde folgender Algorithmus postuliert[9] [3, S. 271]: Der Hai schwimmt geradeaus, bis er ein elektrisches Feld registriert. Wenn er dann ein elektrisches Feld eines möglichen Beutetiers registriert, schwimmt er so, dass sich der Richtungsvektor des elektrischen Feldes nicht ändert. Da die Feldlinien gekrümmt sind und am elektrischen Dipol beginnen und enden, schwimmt er auf den Dipol zu.

Bei Versuchen mit kleineren schwach elektrischen Fischen¹, die ihr Revier verteidigen, haben die Fische ein anderes Verhalten als der Hai gezeigt. Sobald die Fische ein Feld eines sendenden Dipols registrierten, auf den sie zuschwimmen wollten, richteten sie ihren Körper parallel zu den Feldlinien aus, und fingen an, in Richtung der Quelle zu schwimmen. Allerdings nicht auf dem direkten und kürzeren Weg, sondern entlang der gekrümmten Feldlinien. Das deutet darauf hin, dass die Fische nicht zu Beginn die Position abschätzen und sich dann in die richtige Richtung bewegen, sondern dass sie darauf angewiesen sind, ständig ihre Schwimmrichtung zu korrigieren. In weiteren Experimenten wurde diese Vermutung bestätigt. So erreichten sie ihr Ziel meistens nicht, wenn der Dipol aufhörte zu senden. Außerdem schwammen sie auch nach 80 bzw. 160 Wiederholungen nicht auf dem direkten Weg zum sendenden Dipol, sondern folgten den Feldlinien[3, S. 273–277].

Die beiden Algorithmen, die das Verhalten der Fische erklären, sind relativ robust und lösen das Lokalisierungsproblem. Auf die Lösung der Aufgabenstellung dieser Arbeit können sie allerdings nicht übertragen werden, denn das würde bedeuten, dass es nötig wäre, den Roboter zum zu lokalisierenden Objekt fahren zu lassen, bis er physischen Kontakt hat, womit dann die Lokalisierung mit dem Ergebnis beendet wäre, dass sich das zu lokalisierende Objekt direkt vor dem Roboter befindet. Gesucht ist aber eine Lösung, die aufgrund von Messwerten die Position des sendenden Dipols versucht zu berechnen, ohne dass es nötig ist, den Roboter zu bewegen.

1.1. Aufgabenstellung

Projektbeschreibung

Das ANGELS (Anguliform Robot with Electric Sense) Projekt beschäftigt sich mit der Untersuchung und Entwicklung von bio–inspirierten, rekonfigurierbaren Unterwasserrobotern, deren Sensor– und Kommunikationssysteme durch die von verschiedenen schwach elektrischen Fische erzeugten elektrischen Feldern inspiriert sind.

¹Schwach elektrische Fische sind Spezies, die zwar aktiv ein elektrisches Feld z.B. zur Erkundung ihrer Umgebung oder zur Kommunikation erzeugen können, aber keines, das stark genug ist, um Tiere zu betäuben, wie etwa Zitteraale es tun.

Aufgabenbeschreibung

Der in dieser Arbeit verwendete Roboter besitzt eine Reihe von Elektrodenpaaren. An den Elektroden eines dieser Paare kann mit Hilfe eines Digital–Analog–Wandlers (DAC) eine Analogspannung angelegt werden, wodurch im Wasser ein elektrisches Feld erzeugt wird. Das Feld kann mit Hilfe des DAC fast beliebig moduliert werden. Die übrigen Elektrodenpaare sind über Differenzverstärker mit Analog–Digital–Wandlern (ADC) verbunden. Mit deren Hilfe kann der Potentialunterschied im elektrischen Feld zwischen zwei Elektroden gemessen werden.

Mit Hilfe dieser Messungen kann die Position des von einem anderen Roboter erzeugten elektrischen Feldes bestimmt werden. Für eine eingeschränkte Umgebung (2D) wurde hierfür bereits ein Verfahren entwickelt [10].

Die Aufgabe dieser Diplomarbeit besteht darin, ein Verfahren zu finden, das die Lokalisierung auch in einer echten dreidimensionalen Umgebung durchführen kann. Hierfür sollen zuerst mit Hilfe von Matlab–Simulationen Algorithmen untersucht und entwickelt werden. Diese sollen für den späteren Einsatz auf der Roboterplattform geeignet sein und wenn möglich auch auf der Plattform implementiert werden.

2. Lokalisierung von Ladungen

2.1. Vorbedingungen

Um die Entwicklung eines Modells und die Simulation zu vereinfachen, werden vorerst die folgenden Annahmen getroffen.

2.1.1. Reduktion auf die Elektrostatik

Der Einbezug der Zeit führt zu vielen Effekten, die zur Lösung der Problemstellung wenig beitragen, die aber aufwändig zu modellieren und zu berechnen wären. Daher beschränkt sich die Betrachtung zunächst auf die Elektrostatik. Einige der angesprochenen Effekte sind beispielsweise:

Elektrische Felder benötigen Zeit zur Ausbreitung

Die Effekte einer Veränderung der Quelle eines elektrischen Feldes pflanzen sich mit Lichtgeschwindigkeit im Vakuum fort. Wenn ein elektrisches Feld ein Paraelektrikum – wie z.B. Wasser – durchdringt, benötigen die Moleküle des Paraelektrikums Zeit, um sich entlang der elektrischen Feldlinien auszurichten [6, S. 292–294].

Die Permittivität eines Mediums kann sich mit der Frequenz des elektrischen Feldes ändern

Die Moleküle eines Paraelektrikums benötigen Zeit, um sich bei einer Änderung der Richtung der elektrischen Feldlinien zu drehen. Da hierbei auch Reibungsverluste entstehen, wird die Ausbreitung des Feldes nicht nur verlangsamt, ein Wechselfeld wird zusätzlich noch gedämpft und zwar umso stärker, je höher die Frequenz eines elektrischen Feldes ist [6, S. 294]. Diese Drehbewegung der Moleküle führt dazu, dass die so genannte „Dielektrizitätskonstante“ nicht unbedingt konstant ist. Zum Beispiel kann sich die relative Permittivität von Eis im Bereich 5–81 bewegen, wenn die Frequenz des elektrischen Feldes von 1 MHz auf 1 kHz gesenkt wird [5, S. 19].

Interferenz

Das elektrische Feld an einem Punkt ist die Summe aller elektrischen Felder an diesem Punkt [16, S. 29]. Das würde bei einer Modellierung der Elektrodynamik dazu führen, dass die Möglichkeit in Betracht gezogen werden muss, dass sich im

Zeitraum einer Messung verschiedene dynamische elektrische Felder ausgelöscht haben. Die Aufsummierung der elektrischen Felder findet natürlich auch bei statischen elektrischen Feldern statt, allerdings ist hierbei nur ein Fall zu betrachten, während bei dynamischen Feldern die zeitliche Verschiebung in Betracht gezogen werden muss.

Diese Einschränkung kann getroffen werden, da bei den verwendeten Frequenzen von weniger als 1 kHz die Wellenlänge mehrere Kilometer betragen würde und der Wellencharakter daher keinen großen Effekt hat [3, S. 265].

2.1.2. Die relative Permittivität im betrachteten Raum ist homogen

In der Modellierung bedeutet diese Einschränkung konkret, dass die elektrischen Felder in einem unendlichen, mit Wasser gefüllten Raum berechnet werden, der messende Roboter also z.B. das elektrische Feld nicht verändert.

2.1.3. Messungen haben keinen Einfluss auf das Messobjekt

In der Realität wird ein elektrisches Feld durch eine Messung abgeschwächt, da ihm Energie entzogen wird. Im Folgenden wird angenommen, dass das elektrische Feld unverändert bleibt.

2.1.4. Ladungsträger und Messelektroden sind im Modell punktförmig

In der Realität haben sowohl die Sende- als auch die Messelektroden eine räumliche Ausdehnung, die in der Berechnung nicht modelliert werden.

2.2. Einführung

Zur besseren Nachvollziehbarkeit des Verfahrens und der Arbeit werden im Folgenden die zugrundeliegende Physik und die daraus resultierenden möglichen Ansätze zur Problemlösung und deren Nachteile betrachtet.

2.2.1. Physikalische Grundlagen

Die Aufgabenstellung besteht darin, einen elektrischen Dipol aufgrund seines elektrischen Feldes zu lokalisieren. Da ein elektrischer Dipol vereinfacht als zwei einzelne Punktladungen betrachtet werden kann [16, S. 45–49], wird im Folgenden zunächst nur die Lokalisierung einer einzelnen Punktladung betrachtet. Die Elektrodynamik wird mit Hilfe der Maxwellgleichungen beschrieben. Da die Betrachtung außerdem auf die Elektrostatik beschränkt wird, ist das Gauß'sche Gesetz über elektrische Felder relevant. Es gibt zwei Formen des Gauß'schen Gesetzes, die Integralform [5, S. 1–2] und die Differentialform [5, S. 29–30].

$$\oint_S \vec{E} \circ \hat{n} \, da = \frac{q_{enc}}{\epsilon_0} \quad (2.1)$$

Formel 2.1: Gauß'sches Gesetz in Integralform.

\oint_S : Integral über die geschlossene Oberfläche S (von „surface“),

\vec{E} : gerichtetes elektrisches Feld,

\hat{n} : Normalenvektor, der senkrecht auf der Oberfläche steht,

q_{enc} : eingeschlossene („enclosed“) Ladung q ,

ϵ_0 : elektrische Permittivität des Vakuums.

Die Bedeutung der Integralform kann in natürlicher Sprache wie folgt definiert werden, siehe [5, S. 1]:

„Eine elektrische Ladung erzeugt ein elektrisches Feld. Die Stärke des durch das Feld erzeugten Flusses durch eine beliebige geschlossene Fläche ist proportional zum Betrag der Ladung, die von der Fläche umschlossen wird.“

In den Vorbedingungen wurde festgelegt, dass zur Vereinfachung in der mathematischen Modellierung nur Punkte betrachtet werden. In der integralen Form werden allerdings Aussagen über Flächen gemacht. Es ist also nötig, die Aussagen, die das Gauß'sche Gesetz über Flächen macht, auf Aussagen über einzelne Punkte zu übertragen.

Je weiter ein Ort von einer Punktladung entfernt ist, desto schwächer wird das von der Ladung erzeugte elektrische Feld an diesem Ort. Dieser Effekt wird Divergenz genannt [5, S. 32]. Eine mögliche Visualisierung der Divergenz ist die Vorstellung, dass die Dichte der elektrischen Feldlinien proportional mit der Abnahme der Feldstärke abnimmt. Offensichtlich ist die Divergenz ausgehend von der Punktladung in jede Richtung identisch. Wenn nun ein Punkt p mit dem Abstand r zur Punktladung c betrachtet wird, muss es ein Punkt auf einer Kugel mit dem Radius r sein. Die elektrische Felddichte ist überall auf der Kugel identisch. Da eine Kugel eine Oberfläche besitzt, ist es nun möglich, die elektrische Feldstärke an einem bestimmten Punkt im Raum mit Hilfe der integralen Form des Gauß'schen Gesetzes zu berechnen. Dazu wird zuerst die elektrische Flussstärke durch die gesamte Kugeloberfläche berechnet [5, S. 22, 26–27]:

$$\Phi_E = \oint_S \vec{E} \circ \hat{n} \, da = \oint_S |\vec{E}| |\hat{n}| \cos(0^\circ) \, da = \oint_S |\vec{E}| \, da = |\vec{E}| \oint_S \, da = |\vec{E}| (4\pi r^2) \quad (2.2)$$

Formel 2.2: Vom Gauß'schen Gesetz in Integralform zur elektrischen Flussstärke durch die Kugeloberfläche.

\hat{n} : der senkrecht auf der Kugeloberfläche stehende Normalenvektor,

r : Radius der Kugel.

Formel 2.2 zusammen mit dem Gauß'schen Gesetz, Formel 2.1 ergibt dann eine Formel für die elektrische Feldstärke an einem Punkt in der Entfernung r von der Punktladung [5, S. 26–27]:

$$|\vec{E}| = \frac{q_{enc}}{4\pi\epsilon_0 r^2} \quad (2.3)$$

Formel 2.3: Elektrische Feldstärke einer Punktladung an einem Punkt.

r : Distanz des Punktes zur Ladung q ,
 q_{enc} : eingeschlossene (enclosed) Ladung q ,
 ϵ_0 : elektrische Permittivität des Vakuums.

An dieser Stelle noch nicht ausreichend betrachtet wurde die Richtung der Feldlinien. Die übliche Konvention ist folgende: Wenn die Punktladung positiv ist, zeigen die Feldlinien im rechten Winkel von der Punktladung weg, wenn die Punktladung negativ ist, zeigen die Feldlinien im rechten Winkel zur Punktladung hin [5, S. 3–4].

Unter der Voraussetzung, dass die Position und Stärke einer Ladung bekannt ist, ist es also möglich, Stärke und Richtung des von der Ladung erzeugten elektrischen Feldes an einem beliebigen Punkt zu berechnen.

Gesucht wird allerdings eine Lösung für das inverse Problem: Aus den Informationen über das elektrische Feld, die mit Hilfe bestimmter Sensoren durch Messung gewonnen werden, sollen bestimmte Eigenschaften – wie z.B. Position und Stärke – der Ladung berechnet oder zumindest gut abgeschätzt werden.

2.2.2. Der geometrische Ansatz

Zu Beginn wird der einfachste Fall betrachtet, bei dem lediglich eine Punktladung und ein Sensor verwendet werden. Der Sensor liefert absolute Informationen über Stärke und Richtung des elektrischen Feldes an seiner Position. Nun gibt es zwei mögliche Fälle. Entweder, als Ergebnis der Messung wird eine Feldstärke und eine Richtung des elektrischen Feldes bestimmt, oder es wird kein elektrisches Feld gemessen. Im zweiten Fall gibt es wieder zwei Möglichkeiten. Entweder, die elektrische Ladung existiert nicht, ist also gleich null, oder der Sensor befindet sich an exakt derselben Position wie die Ladung selbst. Wenn das elektrische Feld eine Richtung besitzt, kann daraus geschlossen werden, dass es entweder direkt von der Ladung weg, oder zur Ladung hin ausgerichtet ist. Daraus folgt, dass sich die Ladung auf einer zur Richtung des elektrischen Feldes parallelen Linie durch den Sensor befinden muss. Wenn die Polarität der Ladung bekannt ist, kann auf die Richtung geschlossen werden. Falls außerdem die Stärke der Ladung, sowie die Permittivität des Mediums bekannt ist, ist es möglich, mit der Formel 2.3 die Entfernung der Punktladung zu berechnen. Mit dieser Information ist es dann möglich, die relative Position der Punktladung zum Sensor zu bestimmen:

$$\vec{Q} = \vec{S} + \vec{V}_F \cdot D \quad (2.4)$$

Formel 2.4: Positionsbestimmung der Ladung Q .

\vec{Q} : Vektor, der vom Ursprung des Koordinatensystems auf die Position der Ladung Q zeigt,

\vec{S} : Vektor, der vom Ursprung auf die Sensorposition zeigt,

\vec{V}_F : ein zur gemessenen Feldrichtung paralleler Vektor der Länge 1,

D : die aufgrund der Feldstärke berechnete Entfernung.

Dieser direkte Ansatz kann in der Realität allerdings nicht ohne Modifikation genutzt werden. Es ist technisch nicht möglich, direkt das absolute elektrische Feld an einem Punkt im Raum zu messen. Was gemessen werden kann, ist die Differenz der elektrischen Potentiale an zwei „Punkten“ im Raum, also eine Spannung. Eine Möglichkeit, den direkten Ansatz mit Modifikationen in der Realität umzusetzen, wird im folgenden Abschnitt besprochen.

Der direkte geometrische Ansatz in der Realität

Mit Hilfe von Elektroden ist es möglich, eine Spannung zwischen zwei Punkten zu messen. Da die Position der Elektroden bekannt ist, folgt daraus, dass mit einem Paar Elektroden die elektrische Feldstärke in einer bestimmten Richtung und zwar von der einen zur anderen Elektrode bestimmt werden kann. Das ist gleichbedeutend mit der Bestimmung der Feldstärke in einer Dimension. In der Geometrie kann die Position eines Punktes im dreidimensionalen Raum beschrieben werden, indem drei linear unabhängige Vektoren als Basis des Vektorraumes definiert werden. Um die Berechnung und die Vorstellung möglichst nicht unnötig zu verkomplizieren, werden die Basisvektoren so gewählt, dass sie jeweils die Länge 1 haben und rechtwinklig aufeinander stehen. Die Sensoren, die jeweils aus einem Elektrodenpaar bestehen, werden nun in einer ähnlichen Weise positioniert, wie die Basisvektoren eines Koordinatensystems. Also z.B. eine Elektrode jedes Sensors in den Koordinatenursprung $(0, 0, 0)$ und die jeweils andere Elektrode jedes Sensors jeweils eine Längeneinheit davon entfernt und rechtwinklig zueinander, also z.B. an den Positionen $(0, 0, 1)$, $(0, 1, 0)$, $(1, 0, 0)$. Mit dieser Anwendung wäre es möglich, die Stärke des elektrischen Feldes in jeder der drei Dimensionen zu messen und damit die Stärke und Richtung des elektrischen Feldes in allen drei Dimensionen durch eine Vektoraddition zu rekonstruieren.

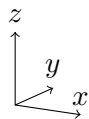


Abbildung 2.1.: Aufeinander rechtwinklig stehende Vektoren als Basis für eine Sensorkonfiguration

Mit dieser Anordnung sollte es möglich sein, falls die Ladung und die Permittivität des Mediums bekannt ist und die Sensoren entsprechend kalibriert werden können, die Position einer Punktladung in einem dreidimensionalen Raum zu berechnen. Dies gilt allerdings nur für die Position einer einzigen Punktladung, denn da das elektrische Feld an einem Punkt die Summe aller Einzelfelder darstellt, würde das Hinzufügen weiterer Ladungen, während in der Berechnung weiter von einer einzelnen Ladung ausgegangen wird, zur Berechnung der Position einer „Phantomladung“ führen. Um diesen Fehler zu bemerken, wäre es nötig, weitere Sensoren hinzuzufügen.

Es gibt ein weiteres technisches Problem. Eine Punktladung unter Wasser zu erzeugen ist schwieriger, als einen elektrischen Dipol zu erzeugen, der mathematisch wie zwei Punktladungen betrachtet werden kann. Während zur Erzeugung eines elektrischen Dipols lediglich eine Stromquelle benötigt wird, deren Pole nach außen geführt werden, wäre die Erzeugung eines Monopols in einem nach außen isolierten System, wie einem Roboter, z.B. durch eine Ladungstrennung, deutlich aufwändiger.

Feldlinien beginnen an positiven Ladungen und enden an negativen Ladungen. Ein Dipol wird als eine positive und eine negative Ladung mit einem bekannten Abstand zwischen ihnen modelliert. Das entstehende elektrische Feld ist die Summe der beiden Felder, die von den Ladungen erzeugt würden, wenn sie einzeln betrachtet würden [6, S. 45–46].

Daraus folgt das Problem, dass, wenn die Messungen lediglich mit einem der vorgestellten Sensoren durchgeführt würden, die Stärke und Ausrichtung des elektrischen Feldes an nur einem Punkt festgestellt werden kann. Die Information ist schlichtweg nicht ausreichend, um auch nur die Richtung, in welcher der Dipol sich befindet, bestimmen zu können.

Eine Hilfe zur Vorstellung: Der Dipol ist vertikal im Raum ausgerichtet, also so, dass sich beide Punktladungen direkt übereinander befinden. Wenn der Sensor nun auf einem Kreis um den Dipol bewegt und passend rotiert wird, ändert sich das von ihm gemessene elektrische Feld nicht, weder in der Stärke, noch im Richtungsvektor. Siehe dazu Abbildung 2.2.

Falls genug über die Stärke der Ladung und die Permittivität bekannt ist, ist es vielleicht möglich, Aussagen über die Entfernung des Sensors zum Dipol zu treffen. Um die Position des Dipols zu triangulieren, also anpeilen zu können, wären weitere Sensoren nötig. Mindestens zwei zusätzliche Sensoren würden benötigt werden, um den Spezialfall auszuschließen, dass sich alle Sensoren in einer Äquipotentialebene befinden und die gleichen Messwerte liefern.

Bis hierhin wurden bei diesem Lösungsansatz einige Kompromisse eingegangen, die ihn als nicht praktikabel erscheinen lassen. Die Anzahl der benötigten Messelektroden ist relativ groß, die Möglichkeit ihrer Positionierung eingeschränkt und die Positionsbestimmung verwendet im einfachsten Fall nur die Daten über die Ausrichtung des elektrischen Feldes. Wünschenswert wäre ein Verfahren, bei dem es möglich ist, die Elektroden beliebig zu positionieren und welches sich nicht auf eine geometrische Konstruktion zur Lokalisierung verlässt, die vielleicht zusätzliche, in den Sensordaten enthaltene Informationen nicht verwendet. Außerdem wäre es bei mehr als einem zu lokalisierenden Dipol nicht mehr möglich, die Positionsbestimmung mit der vorgestellten einfachen geometrischen

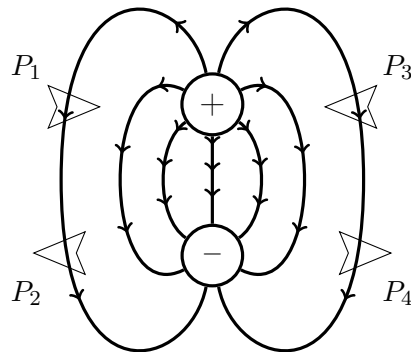


Abbildung 2.2.: Feldlinien eines Dipols. Sensoren an den Positionen P_1 und P_3 , sowie P_2 und P_4 messen dieselbe Feldstärke und denselben Richtungsvektor. Betrachtet man noch die Möglichkeit, Sensoren nicht nur zu verschieben, sondern auch 180° um die Dipolachse zu drehen, lassen sich alle Sensoren so platzieren, dass sie dieselben Daten messen.

Konstruktion durchzuführen.

Auf der Basis der physikalischen Grundlagen ist ein weiterer, relativ direkter Ansatz denkbar, der im nächsten Abschnitt vorgestellt wird.

2.2.3. Ansatz: Gleichungssystem mit Unbekannten

Ausgehend von den physikalischen Grundlagen, die für den geometrischen Lösungsansatz verwendet wurden, ist es möglich, ein Gleichungssystem aufzustellen, das auf folgender Formel basiert:

$$\vec{E} = \frac{1}{4\pi\epsilon_0\epsilon_r} \frac{Q}{r^2} \hat{r} \quad (2.5)$$

Formel 2.5: Stärke des elektrischen Feldes, das von einer gleichmäßig geladenen Kugel in der Entfernung r vom Kugelmittelpunkt erzeugt wird.

\vec{E} : gerichtetes elektrisches Feld,

ϵ_0 : elektrische Permittivität des Vakuums,

ϵ_r : relative elektrische Permittivität des Mediums,

Q : elektrische Ladung,

r : Distanz zwischen dem gewählten Punkt auf der Kugeloberfläche und der Punktladung Q ,

\hat{r} : Vektor von der Ladung Q zu dem Punkt auf der Kugeloberfläche.

2. Lokalisierung von Ladungen

Der Faktor $k_1 = \frac{1}{4\pi\epsilon_0\epsilon_r} \frac{Q}{1}$ enthält keine Modellparameter, kann daher als konstant angenommen werden und muss nicht jedes Mal neu berechnet werden.

Durch die entsprechendes Ersetzen in der Formel 2.5 ergibt sich folgende Formel:

$$\vec{E} = \frac{k_1}{r^2} \hat{r} \quad (2.6)$$

Formel 2.6: Stärke des elektrischen Feldes, das von einer gleichmäßig geladenen Kugel in der Entfernung r vom Kugelmittelpunkt erzeugt wird.

\vec{E} : gerichtetes elektrisches Feld,

$$k_1 = \frac{1}{4\pi\epsilon_0\epsilon_r} \frac{Q}{1},$$

\hat{r} : Vektor von der Ladung Q zu dem Punkt auf der Kugeloberfläche

Der zu lokalisierende Dipol wird durch die zwei Punktladungen P (positiv) und Q (negativ) modelliert, wobei sich die Punktladungen lediglich in ihrer Position und der Polarität ihrer Ladung unterscheiden. Der Betrag der beiden Ladungsmengen ist identisch.

Damit lässt sich die Stärke des elektrischen Feldes am Punkt $a \in \mathbb{R}^3$ mit folgender Formel berechnen:

$$\vec{E}(a, p, q) = \frac{k}{di_3(p, a)} + \frac{-k}{di_3(q, a)} \quad (2.7)$$

Formel 2.7: An Position a messbare Stärke des elektrischen Feldes, das von den beiden Dipolladungen an den Positionen p bzw. q erzeugt wird. Mit $di_3(p, a)$ wird dabei der Abstand der beiden Punkte p und a im dreidimensionalen Raum \mathbb{R}^3 berechnet.

In der Anwendung können die Potentialunterschiede zwischen den beiden Sensorelektroden gemessen werden, die mathematische Formulierung davon wäre folgende Formel:

$$\begin{aligned} U(a, b, p, q) &= \vec{E}(a, p, q) - \vec{E}(b, p, q) \\ &= \frac{k}{di_3(p, a)} + \frac{-k}{di_3(q, a)} - \left(\frac{k}{di_3(p, b)} + \frac{-k}{di_3(q, b)} \right) \end{aligned} \quad (2.8)$$

Formel 2.8: Gemessene Spannung zwischen den beiden Elektroden an den Positionen a und b , die aus dem von den beiden Dipolladungen an den Positionen p und q erzeugten Feld resultiert.

Es ist möglich, für jeden Sensor mit jeweils zwei Elektroden eine entsprechende Gleichung aufzustellen und daraus ein Gleichungssystem zu bilden. Mit Hilfe einer allgemeinen Lösung dieses Gleichungssystems, wäre es zur Positionsbestimmung des Dipols nur noch nötig, die Messwerte in eine Formel einzusetzen, d.h. der Berechnungsaufwand wäre $O(1)$, ebenso wäre der Speicherplatzverbrauch konstant. Es ist im Rahmen dieser Arbeit – auch unter Zuhilfenahme von Algebra-Systemen wie Mathematica und Maxima – nicht gelungen, eine Lösung zu finden, mit der es möglich wäre, bei gegebenen Elektrodenpositionen und Spannungen die Position des gesuchten Dipols direkt zu berechnen.

Zusätzlich zu diesem Problem hätte diese Lösung auch noch den Nachteil, dass sie sich nicht einfach auf die Lokalisierung mehrerer Dipole erweitern lässt und möglicherweise vorhandenes Vorwissen über die Dipolposition sich nicht einfach nutzen lässt.

Eine „Standardlösung“ zur Positionsbestimmung von elektrischen Dipolen wurde im Bereich der Medizin entwickelt. Hierbei wird ein Messelektrodenetz verwendet, um die Potentialunterschiede auf der Kopfoberfläche eines Patienten zu messen. Zur Auswertung dieser Messdaten wird eine Vorwärtssimulation der elektrischen Felder im Kopf benutzt, bei der die Modellparameter, wie z.B. die Dipolanzahl und deren Position, so lange verändert werden, bis die simulierten Sensoren „möglichst ähnliche“ Messdaten – was z.B. mit einer Gütefunktion wie der Methode der kleinsten Quadrate [17, S. 128] bestimmt werden kann – erfassen wie die realen Sensoren [20].

Diese Art von Problemen, also, dass die „Vorwärtsrichtung“ des Problems bekannt und im Vergleich zur Gegenrichtung relativ einfach zu formulieren ist, während die Lösung der „Rückwärtsrichtung“ nicht offensichtlich ist, wird **inverse Probleme** genannt.

2. Lokalisierung von Ladungen

3. Inverse Probleme

Das Problem, eine elektrische Energiequelle aufgrund gemessener Spannungen zu lokalisieren, ist ein inverses Problem. Eines der Standardwerke zu diesem Thema ist das Buch *Geophysical data analysis: discrete inverse theory* von William Menke [13]. Einen allgemeineren und damit auch mächtigeren Ansatz zur Lösung von inversen Problemen stellt Albert Tarantola im Buch *Inverse Problem Theory and Methods for Model Parameter Estimation* [19] vor und zwar die Modellierung von inversen Problemen als Verknüpfung von Wahrscheinlichkeitsdichtefunktionen. Die kurze Einführung in das Thema inverse Probleme im nächsten Abschnitt orientiert sich stark an diesem Buch.

3.1. Einführung

Die grundlegende Idee der inversen Theorie ist die Annahme, dass sich die Information über einen Zustand durch eine Menge von Parametern beschreiben lässt. Um keine möglichen Problemlösungen auszuschließen, wurde eine sehr allgemeine Möglichkeit der Beschreibung der Parameter gewählt: Eine Wahrscheinlichkeitsdichtefunktion [19, S. 1].

Mit Wahrscheinlichkeitsdichtefunktionen werden drei Mengen von Informationen beschrieben:

a-priori-Information über Modellparameter und Daten beschreiben im Voraus vorhandene Informationen über den Zustand des Systems, das untersucht werden soll. Beispiele hierfür wären die Position eines gesuchten Dipols, die Permittivität von Wasser oder die Position der Elektroden auf der Hülle des messenden Roboters. Ein Beispiel für die a-priori-Information über die Daten wäre der Spannungsbereich, in dem die Messwerte liegen können.

Daten oder Messergebnisse sind Informationen über die beobachtbaren Parameter des Systems. Also z.B. die gemessene Spannung zwischen zwei Elektroden.

Informationen über den Zusammenhang von Modellparametern und Messergebnissen beschreiben die Gesetze und Zusammenhänge, die im verwendeten Modell betrachtet werden, z.B. das Gauß'sche Gesetz (Formel 2.1, S.7).

Das allgemeine inverse Problem ist die Kombination dieser Informationsmengen. Der für die Lösung der Problemstellung in dieser Arbeit relevante Fall ist die inverse Richtung, also mit Hilfe der bekannten Zusammenhänge und den Messergebnissen auf die Modellparameter zu schließen, insbesondere auf die Dipolposition.

Parameter	Beschreibung	Einheit	Datentyp
per_w	Permittivität von Wasser	$\frac{F}{m}$	Float
h_s	Höhe des Ursprungs des Koordinatensystems über dem Grund. Der Ursprung des Koordinatensystems liegt im Roboter.	m	Float
h_s	Wasserstand im Becken	m	Float
$coord$	Koordinaten des Dipols	$[m, m, m]$	Float
Q_D	Ladung eines Dipol-Pols	C	Float
$coord E_i$	Koordinaten der Elektrode \mathbf{i}^a	$[m, m, m]$	Float

^aJeder Sensor besteht aus 2 Elektroden, zwischen denen die Spannung gemessen wird.

Tabelle 3.1.: Verwendete Modellparameter und deren Parametrisierung. Verwendete Einheiten: F: Farad, m: Meter, C: Coulomb.

3.2. Modellraum und Datenraum

Um ein inverses Problem zu beschreiben, ist es notwendig, den Modellraum \mathfrak{M} und den Datenraum \mathfrak{D} zu definieren.

3.2.1. Modellraum

Die Parameter, die den Modellraum beschreiben, können willkürlich gewählt werden. Ziel bei der Auswahl der Parameter ist es, eine minimale Anzahl von Parametern auszuwählen, die alle nötigen Informationen über das System erfasst, um alle für den Anwendungsfall wichtigen Eigenschaften zu beschreiben [19, S. 2–5]. Die Tabelle 3.2.1, S. 16 listet die verwendeten Parameter auf.

Der nächste Schritt ist die Parametrisierung. Hierbei werden die Parameter messbar gemacht, d.h. es wird ihnen z.B. eine physikalische Einheit zugewiesen. Obwohl die Tabelle 3.2.1 suggeriert, dass die Daten in Feldern abgelegt werden und die Implementierung auch so geschehen kann, ist es auch möglich, sämtliche Parameter m_i des Modells durchnummerieren und $\mathcal{M}_j = (m_1, m_2, \dots, m_i)$ als einen Punkt im i -dimensionalen Modellraum zu betrachten [19, S. 2–3].

3.2.2. Datenraum

Der Datenraum \mathfrak{D} wird analog zum Modellraum \mathfrak{M} definiert [19, S. 5]. Das bedeutet, jede mögliche Kombination von Messdaten entspricht jeweils einem bestimmten Punkt in \mathfrak{D} . Wenn z.B. die Spannungen an den Sensoren 1–4, 100 mV, 200 mV, 300 mV und 400 mV betragen, wäre es möglich, diese Messung als einen Punkt $\mathcal{D} = (100, 200, 300, 400) \in \mathfrak{D}$ darzustellen.

Eine Übersicht über die verwendeten Daten befindet sich in Tabelle 3.2.2, S. 17.

Datum	Beschreibung	Einheit	Datentyp
U_{ab}	Gemessene Spannung zwischen den Elektroden ^a a und b	V	Float
p_S	Gemessener Druck des Drucksensors	mbar	Float

^aJeder Sensor besteht aus 2 Elektroden, zwischen denen die Spannung gemessen wird.

Tabelle 3.2.: Verwendete Daten und deren Parametrisierung. Einheiten: V: Volt, mbar: Millibar

3.2.3. Definition der Wahrscheinlichkeit

Sei \mathfrak{X} eine endlichdimensionale *Mannigfaltigkeit*, also ein Raum, der zumindest lokal \mathbb{R}^n gleicht. In der Anwendung kann das der Modellraum \mathfrak{M} , der Datenraum \mathfrak{D} oder eine Vereinigung der beiden Räume sein.

Die Teilmengen $\mathcal{A}, \mathcal{B} \subseteq \mathfrak{X}$, die z.B. einzelne Punkte oder auch größere Bereiche von \mathfrak{X} sein können, werden Ereignisse genannt. Sei P ein Wahrscheinlichkeitsmaß, also eine Abbildung, welche folgende Kolmogorovaxiome erfüllt [19, S. 6]:

$$\forall \mathcal{A} \subseteq \mathfrak{X} : P(\mathcal{A}) \rightarrow [0, 1] \quad (3.1)$$

Formel 3.1: Kolmogorovaxiom 1: Wahrscheinlichkeiten von Ereignissen liegen im Intervall der reellen Zahlen zwischen einschließlich 0 und 1.

$$\mathcal{A} \cap \mathcal{B} = \emptyset \Rightarrow P(\mathcal{A} \cup \mathcal{B}) = P(\mathcal{A}) + P(\mathcal{B}) \quad (3.2)$$

Formel 3.2: Kolmogorovaxiom 2: Die Wahrscheinlichkeit des Eintretens einer Vereinigung zweier inkompatibler (oder disjunkter) Ereignisse ist gleich der Summe der Wahrscheinlichkeiten des Eintretens der Einzelereignisse.

Aus den Axiomen folgt [19, S. 6]:

$$P(\emptyset) = 0 \quad (3.3)$$

Formel 3.3: Die Wahrscheinlichkeit des leeren Ereignisses ist gleich Null.

$$P(\mathcal{A} \cup \mathcal{B}) = P(\mathcal{A}) + P(\mathcal{B}) - P(\mathcal{A} \cap \mathcal{B}) \quad (3.4)$$

Formel 3.4: Wahrscheinlichkeit der Vereinigung zweier nicht notwendigerweise disjunkte Ereignisse \mathcal{A} und \mathcal{B}

In der Anwendung kann nicht vorausgesetzt werden, dass die Modellparameter ausschließlich diskrete Werte annehmen. Wäre dies der Fall, wäre es möglich, alle Fälle abzuzählen und eine Summe über die Eintrittswahrscheinlichkeiten aller Ereignisse zu

bilden. Da die Verteilung stetig ist, ist es nötig und möglich, die Funktion P mit einer Wahrscheinlichkeitsdichtefunktion auszudrücken [19, S. 7].

Sei $x = \{x_1, x_2, \dots\}$ ein bestimmtes Koordinatensystem und P eine Wahrscheinlichkeitsverteilung über einer Mannigfaltigkeit \mathfrak{X} , dann gibt es nach dem Satz von Radon–Nikodym, (siehe [4, S. 139–142], sowie [1, S. 167–169]) für jede Wahrscheinlichkeitsverteilung P eine positive Funktion $f(x)$, so dass gilt:

$$\forall \mathcal{A} \subseteq \mathfrak{X} : P(\mathcal{A}) = \int_{\mathcal{A}} dx f(x), \tag{3.5}$$

mit $\int_{\mathcal{A}} dx = \int_{\mathcal{A}} dx_1 \int_{\mathcal{A}} dx_2 \int_{\mathcal{A}} dx_3 \dots$

Formel 3.5: Definition: Wahrscheinlichkeitsdichtefunktion

Die Funktion f wird als Wahrscheinlichkeitsdichtefunktion bezeichnet.

3.2.4. Die homogene Wahrscheinlichkeitsverteilung

Es wird angenommen, dass es möglich ist, der Mannigfaltigkeit \mathfrak{X} ein Volumen zuzuordnen. Diese Bedingung trifft auf den betrachteten Modellraum \mathfrak{M} zu, da er alle möglichen Positionen des zu lokalisierenden Dipols darstellen können soll. Daher ist es möglich, jeder Teilmenge $\mathcal{A} \subseteq \mathfrak{X}$ ebenfalls ein Volumen $V(\mathcal{A})$ zuzuordnen [19, S. 10].

Wenn für den Fall $\mathbf{x} = \{x_1, x_2, \dots\}$ $dV(\mathbf{x}) = v(\mathbf{x}) d\mathbf{x}$ gesetzt wird, ist es möglich, das Volumen $V(\mathcal{A})$ wie folgt zu definieren:

$$V(\mathcal{A}) = \int_{\mathcal{A}} d\mathbf{x} v(\mathbf{x}). \tag{3.6}$$

Formel 3.6: Definition: Volumen $V(\mathcal{A})$

Die Funktion $v(\mathbf{x})$, wird als Volumendichte von \mathfrak{X} an der Stelle \mathbf{x} bezeichnet [19, S. 10].

Das Volumen V des verwendeten Modellraumes \mathfrak{M} ist endlich, denn zum einen werden die Versuche in einem geschlossenen Aquarium durchgeführt. Zum anderen nehmen die messbaren Effekte des sendenden Dipols mit der Entfernung stark ab, weshalb es in der Praxis nicht nötig ist, ein unendliches Volumen zu modellieren. Daher kann angenommen werden, dass das Volumen $V(\mathfrak{X}) = \int_{\mathfrak{X}} d\mathbf{x} v(\mathbf{x})$ endlich ist. Dann ist es möglich, eine normalisierte Wahrscheinlichkeitsdichte $\mu(\mathbf{x})$ zu definieren [19, S. 10].

$$\mu(\mathbf{x}) = \frac{v(\mathbf{x})}{V} \tag{3.7}$$

Formel 3.7: Definition: normalisierte Wahrscheinlichkeitsdichte $\mu(\mathbf{x})$

Sie definiert für jedes Teilvolumen $\mathcal{A} \subseteq \mathfrak{X}$ die homogene Wahrscheinlichkeit $M(\mathcal{A})$, die proportional zum Volumen $V(\mathcal{A})$ ist [19, S. 10]:

$$M(\mathcal{A}) = \int_{\mathcal{A}} d\mathbf{x} \mu(\mathbf{x}) \quad (3.8)$$

Formel 3.8: Definition: homogene Wahrscheinlichkeit $M(\mathcal{A})$

Dieser Fall tritt in der Anwendung z.B. dann ein, wenn keinerlei a-priori-Information über die Position des zu lokalisierenden Dipols zur Verfügung steht und davon ausgegangen werden muss, dass er sich mit gleicher Wahrscheinlichkeit an einer beliebigen von sämtlichen möglichen Positionen befinden könnte.¹

Es wird verlangt, dass für jede Wahrscheinlichkeitsverteilung P gilt, dass die Wahrscheinlichkeit eines Ereignisses $P(\mathcal{A})$, $\mathcal{A} \subseteq \mathfrak{X}$ gleich Null ist, wenn das Volumen $V(\mathcal{A})$ gleich Null ist. Sei $f(\mathbf{x})$ eine Wahrscheinlichkeitsdichtefunktion, dann gilt [19, S. 10]:

$$\mu(\mathbf{x}) = 0 \Rightarrow f(\mathbf{x}) = 0 \quad (3.9)$$

Formel 3.9: Wahrscheinlichkeit leeres Volumen

Anschaulich auf den verwendeten Modellraum übertragen bedeutet das: Die Wahrscheinlichkeit, dass der sendende Dipol sich in einem leeren Volumen befindet, ist gleich Null.

Wenn $f(\mathbf{x})$ eine Wahrscheinlichkeitsdichtefunktion und $\mu(\mathbf{x})$ die zugehörige homogene Wahrscheinlichkeitsdichtefunktion ist, dann ist es möglich, eine volumenbezogene Wahrscheinlichkeit $\varphi(\mathbf{x})$ zu definieren [19, S. 12]:

$$\varphi(\mathbf{x}) = \frac{f(\mathbf{x})}{\mu(\mathbf{x})} \quad (3.10)$$

Formel 3.10: Definition: volumenbezogene Wahrscheinlichkeit

3.2.5. Verknüpfung von Wahrscheinlichkeitsverteilungen

Sei \mathfrak{X} eine endlichdimensionale Mannigfaltigkeit, $\mathcal{A} \subseteq \mathfrak{X}$ ein Ereignis, sowie P_1 und P_2 zwei auf \mathfrak{X} definierte Wahrscheinlichkeitsverteilungen. Für diese Wahrscheinlichkeitsverteilungen wird nun eine Verknüpfung \wedge definiert, so dass gilt [19, S. 13]:

$$\begin{aligned} \forall \mathcal{A} \subseteq \mathfrak{X} : P_1(\mathcal{A}) = 0 \text{ oder } P_2(\mathcal{A}) = 0 \\ \Rightarrow (P_1 \wedge P_2)(\mathcal{A}) = 0 \end{aligned} \quad (3.11)$$

$$P_1 \wedge P_2 = P_2 \wedge P_1$$

Formel 3.11: Definition: Verknüpfung \wedge von Wahrscheinlichkeitsverteilungen

¹Diese Annahme ist natürlich immer eine Vereinfachung, da bekannt ist, dass sich der zu lokalisierende Dipol nicht innerhalb des Roboters befinden kann.

Außerdem muss die homogene Wahrscheinlichkeitsverteilung M neutrales Element der Verknüpfung sein, also die Bedingung $P \wedge M = P$ gelten.

Es werden folgende Wahrscheinlichkeitsdichteverteilungen analog zu 3.2.4 definiert [19, S. 13]:

$$\begin{aligned} P_1(\mathcal{A}) &= \int_{\mathcal{A}} d\mathbf{x} f_1(\mathbf{x}), \\ P_2(\mathcal{A}) &= \int_{\mathcal{A}} d\mathbf{x} f_2(\mathbf{x}), \\ M(\mathcal{A}) &= \int_{\mathcal{A}} d\mathbf{x} \mu(\mathbf{x}) \end{aligned} \tag{3.12}$$

Formel 3.12: Definition: Wahrscheinlichkeitsdichteverteilungen

Wenn $(f_1 \wedge f_2)$ die Wahrscheinlichkeitsdichten von $P_1 \wedge P_2$ darstellt, also $(P_1 \wedge P_2)(\mathcal{A}) = \int_{\mathcal{A}} d\mathbf{x} (f_1 \wedge f_2)(\mathbf{x})$ gilt, dann lassen sich die Axiome folgendermaßen lösen [19, S. 14]:

$$\begin{aligned} (f_1 \wedge f_2)(\mathbf{x}) &= \frac{1}{v} \frac{f_1(\mathbf{x}) f_2(\mathbf{x})}{\mu(\mathbf{x})}, \\ \text{mit } v &= \int_{\mathfrak{X}} d\mathbf{x} \frac{f_1(\mathbf{x}) f_2(\mathbf{x})}{\mu(\mathbf{x})} \end{aligned} \tag{3.13}$$

Formel 3.13: Lösung der aufgestellten Axiome für die Verknüpfung von Wahrscheinlichkeitsdichtefunktionen

3.2.6. Bedingte Wahrscheinlichkeit

Sei $\mathcal{A} \subseteq \mathfrak{X}$ ein Ereignis und seien P, Q durch die Funktionen $f(\mathbf{x}), g(\mathbf{x})$ definierte Wahrscheinlichkeitsverteilungen über \mathfrak{X} , dann sind die Wahrscheinlichkeiten über \mathcal{A} wie folgt definiert [19, S. 15]:

$$\begin{aligned} P(\mathcal{A}) &= \int_{\mathcal{A}} d\mathbf{x} f(\mathbf{x}) \\ Q(\mathcal{A}) &= \int_{\mathcal{A}} d\mathbf{x} g(\mathbf{x}) \end{aligned} \tag{3.14}$$

Formel 3.14: Definition: Wahrscheinlichkeiten über \mathcal{A}

Es ist somit möglich, jedem Ereignis \mathcal{A} eine Wahrscheinlichkeitsdichte $M_{\mathcal{A}}$ zuzuordnen. Dazu wird die Wahrscheinlichkeitsdichte $\mu_{\mathcal{A}}$ mit einem Normalisierungsfaktor k definiert [19, S. 16]:

$$\mu_{\mathcal{A}}(\mathbf{x}) = \begin{cases} k\mu(\mathbf{x}) & \text{wenn } \mathbf{x} \in \mathcal{A}, \\ 0 & \text{sonst} \end{cases} \quad (3.15)$$

Formel 3.15: Definition: Wahrscheinlichkeitsdichte $\mu_{\mathcal{A}}$

Mit Hilfe dieser Definition von $M_{\mathcal{A}}$ ist es möglich, jedem Ereignis $\mathcal{B} \subseteq \mathfrak{X}$ eine Wahrscheinlichkeit $M_{\mathcal{A}}(\mathcal{B}) = \int_{\mathcal{B}} d\mathbf{x} \mu_{\mathcal{A}}(\mathbf{x})$ zuzuordnen. $M_{\mathcal{A}}(\mathcal{B})$ ist dabei proportional zum Volumen $V(\mathcal{A} \cap \mathcal{B})$, da die homogene Dichtefunktion $\mu(\mathbf{x})$ verwendet wurde. $M_{\mathcal{A}}$ wird als Wahrscheinlichkeitsereignis bezeichnet [19, S. 16].

Sei $\mathcal{B} \subseteq \mathfrak{X}$ ein Ereignis und $M_{\mathcal{B}}$ das zugehörige Wahrscheinlichkeitsereignis. Die Wahrscheinlichkeitsverteilung $(P \wedge M_{\mathcal{B}})$. Für dieses \mathcal{B} und ein beliebiges $\mathcal{A} \subseteq \mathfrak{X}$ ergibt sich die Lösung [19, S. 16]:

$$(P \wedge M_{\mathcal{B}})(\mathcal{A}) = \frac{P(\mathcal{A} \cap \mathcal{B})}{P(\mathcal{B})}, \text{ wenn } P(\mathcal{B}) \neq 0 \quad (3.16)$$

Formel 3.16: Verknüpfung von Wahrscheinlichkeitsdichten

Der Ausdruck auf der rechten Seite entspricht der üblichen Definition der bedingten Wahrscheinlichkeit², dem Bayes-Theorem([11, S. 13–14]) [19, S. 16]:

$$P(\mathcal{A}|\mathcal{B}) \equiv \frac{P(\mathcal{A} \cap \mathcal{B})}{P(\mathcal{B})} \quad (3.17)$$

Formel 3.17: Definition bedingte Wahrscheinlichkeit $P(\mathcal{A}|\mathcal{B})$

Daraus folgt, dass für beliebige \mathcal{A} und \mathcal{B} gilt [19, S. 17]:

$$(P \wedge M_{\mathcal{B}})(\mathcal{A}) = P(\mathcal{A}|\mathcal{B}) \quad (3.18)$$

Formel 3.18: Äquivalenz: Wahrscheinlichkeitsdichte verknüpft mit Ereignis – bedingte Wahrscheinlichkeit

Die bedingte Wahrscheinlichkeitsverteilung ist also ein Spezialfall der Verknüpfung von Wahrscheinlichkeitsverteilungen. Die Umkehrung gilt allerdings nicht, denn die Verknüpfung von Wahrscheinlichkeitsverteilungen lässt mehr Freiheiten in der Formulierung von Wahrscheinlichkeitsverteilungen zu. Der Unterschied ist folgender: Das Ereignis $\mathcal{B} \subseteq \mathfrak{X}$ beschreibt ein gewisses Volumen. Ein Punkt \mathbf{x} kann entweder innerhalb oder außerhalb dieses Volumens liegen, es gibt keine Zwischenwerte. Das heißt, dass bei der Verwendung der bedingten Wahrscheinlichkeit eine „scharfe“ Grenze eingeführt wird, denn über die

² $\frac{P(\mathcal{A} \cap \mathcal{B})}{P(\mathcal{B})}$ definiert die Wahrscheinlichkeit, dass das Ereignis \mathcal{A} eintritt, unter der Voraussetzung, dass auch Ereignis \mathcal{B} eintritt.

Bereiche außerhalb von \mathcal{B} macht die bedingte Wahrscheinlichkeit keine Aussage mehr [19, S. 17].

Dagegen ist es bei der Verknüpfung von Wahrscheinlichkeitsverteilungen möglich, „unscharfe“ Grenzen zu verwenden, indem als „Filter“ z.B. nicht eine Wahrscheinlichkeitsverteilung verwendet wird, die dem Ereignis \mathcal{B} entspricht, sondern eine, deren Wahrscheinlichkeit im Zentrum von \mathcal{B} maximal ist und zum Rand von \mathcal{B} abnimmt [19, S. 17].

Bei der Problemstellung, die in dieser Arbeit behandelt wird, wäre es z.B. möglich, dass die Wahrscheinlichkeit, dass sich der zu suchende Dipol in der Mitte zwischen Wasseroberfläche und Grund befindet, maximal ist, während die Wahrscheinlichkeit jeweils in Richtung Wasseroberfläche und Grund abnimmt.

3.2.7. Rand- und abhängige Wahrscheinlichkeitsdichte

Sei \mathfrak{M} der Modellraum mit α Koordinaten und $\mathbf{m} \in \mathfrak{M}$ [19, S. 18]:

$$\mathbf{m} = \{m_1, m_2, \dots, m_\alpha\} \quad (3.19)$$

Formel 3.19: Definition: Koordinaten des Modellraumes \mathfrak{M}

Sei $\mu_M(\mathbf{m})$ eine auf \mathfrak{M} definierte homogene Wahrscheinlichkeitsdichte [19, S. 18].

Sei \mathfrak{D} der Datenraum mit β Koordinaten und $\mathbf{d} \in \mathfrak{D}$ [19, S. 18]:

$$\mathbf{d} = \{d_1, d_2, \dots, d_\beta\} \quad (3.20)$$

Formel 3.20: Definition: Koordinaten des Datenraumes \mathfrak{D}

Sei $\mu_D(\mathbf{d})$ eine auf \mathfrak{D} definierte homogene Wahrscheinlichkeitsdichte [19, S. 18].

Sei \mathfrak{X} die Verbundmannigfaltigkeit, definiert als das kartesische Produkt $\mathfrak{D} \times \mathfrak{M}$ mit den Koordinaten [19, S. 18]:

$$\mathbf{x} = \{\mathbf{d}, \mathbf{m}\} = \{d_1, d_2, \dots, d_\beta, m_1, m_2, \dots, m_\alpha\} \quad (3.21)$$

Formel 3.21: Definition: Koordinaten der Mannigfaltigkeit \mathfrak{X}

Eine Wahrscheinlichkeit über \mathfrak{X} kann dann als normalisierte Wahrscheinlichkeitsdichtefunktion $f(\mathbf{x}) = f(\mathbf{d}, \mathbf{m})$ modelliert werden, d.h. es gilt [19, S. 18]:

$$\int_{\mathfrak{D}} d\mathbf{d} \int_{\mathfrak{M}} d\mathbf{m} f(\mathbf{d}, \mathbf{m}) = \int_{\mathfrak{M}} d\mathbf{m} \int_{\mathfrak{D}} d\mathbf{d} f(\mathbf{d}, \mathbf{m}) = 1 \quad (3.22)$$

Formel 3.22: Normalisierung Wahrscheinlichkeitsdichtefunktion

Hiermit ist es nun möglich, folgende Randwahrscheinlichkeitsdichten zu definieren [19, S. 18]:

$$f_{\mathfrak{M}}(\mathbf{m}) = \int_{\mathfrak{D}} d\mathbf{d}f(\mathbf{d}, \mathbf{m}) \quad , \quad f_{\mathfrak{D}}(\mathbf{d}) = \int_{\mathfrak{M}} d\mathbf{m}f(\mathbf{d}, \mathbf{m}) \quad (3.23)$$

Formel 3.23: Definition: Randwahrscheinlichkeitsdichten über \mathfrak{D} und \mathfrak{M}

$$\int_{\mathfrak{M}} d\mathbf{m}f_{\mathfrak{M}}(\mathbf{m}) = 1 \quad , \quad \int_{\mathfrak{D}} d\mathbf{d}f_{\mathfrak{D}}(\mathbf{d}) = 1 \quad (3.24)$$

Formel 3.24: Definition: Normalisierung der Randwahrscheinlichkeitsdichten über \mathfrak{D} und \mathfrak{M}

\mathbf{d} und \mathbf{m} werden unabhängig genannt, wenn gilt [19, S. 19]:

$$f(\mathbf{d}, \mathbf{m}) = f_{\mathfrak{D}}(\mathbf{d})f_{\mathfrak{M}}(\mathbf{m}) \quad (3.25)$$

Formel 3.25: Unabhängigkeit von \mathbf{d} und \mathbf{m}

Mit Hilfe von $f(\mathbf{d}, \mathbf{m})$ wird im Folgenden eine bedingte Wahrscheinlichkeitsdichte definiert. Sei $\mathbf{d} \mapsto \mathbf{m}(\mathbf{d})$ eine Abbildung von \mathfrak{D} nach \mathfrak{M} . Die zugrundeliegende Idee bei der Definition der bedingten Wahrscheinlichkeitsdichte ist es, nur die Werte von $f(\mathbf{d}, \mathbf{m})$ zu betrachten, für welche die Bedingung $\mathbf{m} = \mathbf{m}(\mathbf{d})$ gilt [19, S. 19].

Um die Werte entsprechend zu „filtern“ wird die allgemeine Definition der bedingten Wahrscheinlichkeit verwendet, (Formel 3.17, s.S. 21) wobei der benutzte Grenzwert gegen $\mathbf{m} = \mathbf{m}(\mathbf{d})$ strebt [19, S. 19].

Unter den Bedingungen, dass $\mathbf{m} = \mathbf{m}(\mathbf{d})$ wenig von der Linearität abweicht und die Koordinaten $\mathbf{m} = \{m_1, m_2, \dots, m_\alpha\}$, und $\mathbf{d} = \{d_1, d_2, \dots, d_\beta\}$ als kartesische Koordinaten betrachtet werden können, ist es möglich, die bedingte Wahrscheinlichkeit wie folgt zu definieren [19, S. 19]:

$$f_{\mathfrak{D}|\mathfrak{M}}(\mathbf{d}|\mathbf{m}(\mathbf{d})) = \frac{f(\mathbf{d}, \mathbf{m}(\mathbf{d}))}{\int_{\mathfrak{D}} d\mathbf{d}'f(\mathbf{d}', \mathbf{m}(\mathbf{d}'))} \quad (3.26)$$

Formel 3.26: Definition: Bedingte Wahrscheinlichkeitsdichte $f_{\mathfrak{D}|\mathfrak{M}}(\mathbf{d}|\mathbf{m}(\mathbf{d}))$

Ein Spezialfall dieser Definition ist der Fall, dass \mathbf{m} nicht durch die Funktion $\mathbf{m}(\mathbf{d})$ definiert wird, sondern durch eine Konstante, also $\mathbf{m} = \mathbf{m}_0$ gilt [19, S. 20].

Dann wird aus der Formel 3.26:

$$f_{\mathfrak{D}|\mathfrak{M}}(\mathbf{d}|\mathbf{m}_0) = \frac{f(\mathbf{d}, \mathbf{m}_0)}{\int_{\mathfrak{D}} d\mathbf{d}'f(\mathbf{d}', \mathbf{m}_0)} \quad (3.27)$$

Formel 3.27: Definition: Bedingte Wahrscheinlichkeit mit konstanter Bedingung $f_{\mathfrak{D}|\mathfrak{M}}(\mathbf{d}|\mathbf{m}(\mathbf{d}))$

Mit der Formel 3.23, S. 23, ergibt sich [19, S. 20]:

$$f_{\mathcal{D}|\mathfrak{M}}(\mathbf{d}|\mathbf{m}) = \frac{f(\mathbf{d}, \mathbf{m})}{f_{\mathfrak{M}}(\mathbf{m})} \quad (3.28)$$

Formel 3.28: Definition: Bedingte Wahrscheinlichkeit $f_{\mathcal{D}|\mathfrak{M}}(\mathbf{d}|\mathbf{m})$

Daher kann die Verbundwahrscheinlichkeit wie folgt definiert werden [19, S. 20]:

$$f(\mathbf{d}, \mathbf{m}) = f_{\mathcal{D}|\mathfrak{M}}(\mathbf{d}|\mathbf{m})f_{\mathfrak{M}}(\mathbf{m}) \quad (3.29)$$

Formel 3.29: Definition: Verbundwahrscheinlichkeit $f(\mathbf{d}, \mathbf{m})$

3.2.8. Das vorwärtsgerichtete Problem

Die grundlegende Idee der inversen Theorie ist die Annahme, dass es einen Zusammenhang zwischen den Modellparametern, also den Punkten $m \in \mathcal{M}$ und den Daten, also den Punkten $d \in \mathcal{D}$ gibt. Dieser Zusammenhang ist das Modell. Die abstrakte Formulierung des vorwärtsgerichteten Problems ist [19, S. 20]:

$$\mathbf{d} = \mathbf{g}(\mathbf{m}) \quad (3.30)$$

Formel 3.30: Definition: Vorwärtsgerichtetes Problem

Das bedeutet, dass durch die Anwendung des Vorwärtsoperators \mathbf{g} , der die mathematische Repräsentation des benutzten physikalischen Modells darstellt, auf die Modellparameter die zu erwartenden (Mess-)Daten berechnet werden können [19, S. 20].

Da aufgrund experimenteller Unsicherheit nicht zu erwarten ist, dass die berechneten Daten \mathbf{d} exakt mit den gemessenen übereinstimmen, ist es nötig, die Unsicherheit zu modellieren. Eine Möglichkeit hierfür ist die Verwendung von Wahrscheinlichkeitsdichten [19, S. 21].

Sei \mathfrak{M} der Modellraum mit α Koordinaten:

$$\mathbf{m} = \{m_1, m_2, \dots, m_\alpha\} \quad (3.31)$$

Formel 3.31: Definition: Koordinaten des Modellraumes \mathfrak{M}

Sei $\mu_M(\mathbf{m})$ eine auf \mathfrak{M} definierte homogene Wahrscheinlichkeitsdichte [19, S. 20].

Sei \mathcal{D} der Datenraum mit β Koordinaten:

$$\mathbf{d} = \{d_1, d_2, \dots, d_\beta\} \quad (3.32)$$

Formel 3.32: Definition: Koordinaten des Datenraumes \mathcal{D}

Sei $\mu_D(\mathbf{d})$ eine auf \mathfrak{D} definierte homogene Wahrscheinlichkeitsdichte [19, S. 20].

Sei \mathfrak{X} die Verbundmannigfaltigkeit, definiert als das kartesische Produkt $\mathfrak{D} \times \mathfrak{M}$ mit den Koordinaten [19, S. 20]:

$$\mathbf{x} = \{\mathbf{d}, \mathbf{m}\} = \{d_1, d_2, \dots, d_\beta, m_1, m_2, \dots, m_\alpha\} \quad (3.33)$$

Formel 3.33: Definition: Koordinaten der Mannigfaltigkeit \mathfrak{X}

Auf \mathfrak{X} ist dann folgende homogene Wahrscheinlichkeitsdichte definiert [19, S. 20]:

$$\mu(\mathbf{x}) = \mu(\mathbf{d}, \mathbf{m}) = \mu_D(\mathbf{d})\mu_M(\mathbf{m}) \quad (3.34)$$

Formel 3.34: Definition: Homogene Wahrscheinlichkeitsdichte \mathfrak{X}

Um die Unsicherheit der Vorwärtsmodellierung – also die „Unschärfe“ der Daten $\mathbf{d} = \mathbf{g}(\mathbf{m})$ – auszudrücken, ist es möglich, eine von \mathbf{m} bedingte Wahrscheinlichkeitsdichtefunktion $\theta(\mathbf{d}|\mathbf{m})$ zu definieren, die diese Unsicherheit modelliert. Bildlich gesprochen stellt $\theta(\mathbf{d}|\mathbf{m})$ einen unscharfen Korridor um die Funktion $\mathbf{d} = \mathbf{g}(\mathbf{m})$ dar. D.h. die von der Funktion $\theta(\mathbf{d}|\mathbf{m})$ berechnete Wahrscheinlichkeit für einen bestimmten Wert \mathbf{m} , nimmt ab, je weiter man sich vertikal von dem entsprechenden Funktionswert der Funktion $\mathbf{d} = \mathbf{g}(\mathbf{m})$ an der gleichen Stelle \mathbf{m} entfernt [19, S. 22–23].

Mit dieser Definition ist es möglich, die theoretische Verbundwahrscheinlichkeit $\Theta(\mathbf{d}, \mathbf{m})$ über Datenraum und Modellraum als das Produkt der bedingten und der Randwahrscheinlichkeit zu definieren [19, S. 22]:

$$\Theta(\mathbf{d}, \mathbf{m}) = \theta(\mathbf{d}|\mathbf{m})\mu_M(\mathbf{m}) \quad (3.35)$$

Formel 3.35: Definition: Verbundwahrscheinlichkeit von Datenraum \mathfrak{D} und Modellraum \mathfrak{M} .

Die erwähnte Unsicherheit des Vorwärtsmodells kann z.B. mit Hilfe einer Gaußverteilung modelliert werden [19, S. 23].

3.3. Messungen und a-priori-Information

3.3.1. Messergebnisse

Sei $\mathbf{d} = \{d_1, d_2, \dots, d_\beta\}$ die Menge der messbaren Parameter, dann können Messergebnisse als Wahrscheinlichkeitsdichtefunktion $\rho_D(\mathbf{d})$ über \mathfrak{D} definiert werden [19, S. 24].

Die allgemeinste Definition der Informationen über die Eigenschaften der Daten ist eine Wahrscheinlichkeitsdichtefunktion, die jedem möglichen realen Datenwert \mathbf{d} und jedem Messergebnis \mathbf{d}_{out} eines Messinstruments eine Wahrscheinlichkeit zuordnet. Sei $v(\mathbf{d}_{\text{out}}|\mathbf{d})$

diese bedingte Wahrscheinlichkeit. Wenn $f(\mathbf{d}_{\text{out}}, \mathbf{d})$ die Verbundwahrscheinlichkeit von \mathbf{d} und \mathbf{d}_{out} ist und keine weiteren Informationen über die Messwerte zur Verfügung stehen, kann zur Modellierung die homogene Wahrscheinlichkeitsdichte nach Formel 3.29 verwendet werden, was zur folgenden Formel führt [19, S. 25–26]:

$$f(\mathbf{d}_{\text{out}}, \mathbf{d}) = v(\mathbf{d}_{\text{out}}|\mathbf{d})\mu_D(\mathbf{d}) \quad (3.36)$$

Formel 3.36: Definition: Wahrscheinlichkeitsdichtefunktion Messinstrument ohne weitere Informationen über die Messdaten

Falls der Datenraum linear ist, ist die homogene Wahrscheinlichkeitsdichtefunktion $\mu_D(\mathbf{d})$ konstant und es ergibt sich [19, S. 26]:

$$f(\mathbf{d}_{\text{out}}, \mathbf{d}) = k v(\mathbf{d}_{\text{out}}|\mathbf{d}) \quad (3.37)$$

Formel 3.37: Definition: Wahrscheinlichkeitsdichtefunktion Messinstrument bei linearem Datenraum mit Konstante k

Das Messergebnis einer Messung wird mit \mathbf{d}_{obs} bezeichnet. Hiermit wird eine abhängige Wahrscheinlichkeitsdichtefunktion $\rho_D(\mathbf{d})$ definiert [19, S. 26]:

$$\begin{aligned} \rho_D(\mathbf{d}) = f(\mathbf{d}|\mathbf{d}_{\text{out}} = \mathbf{d}_{\text{obs}}) &= \frac{f(\mathbf{d}_{\text{obs}}, \mathbf{d})}{\int_{\mathbf{D}} d\mathbf{d} f(\mathbf{d}_{\text{obs}}, \mathbf{d})} \\ &= kv(\mathbf{d}_{\text{obs}}|\mathbf{d}) \end{aligned} \quad (3.38)$$

Formel 3.38: Definition: Wahrscheinlichkeitsdichtefunktion über Datenraum bei gegebenen Eigenschaften $v(\mathbf{d}_{\text{obs}}|\mathbf{d})$ des Messinstruments

Ein Extrem für die Wahrscheinlichkeitsdichtefunktion wäre ein perfektes Messinstrument, also ein Instrument, bei dem für den Messwert \mathbf{d}_{obs} eines realen, gemessenen Parameters \mathbf{d} , $\mathbf{d}_{\text{obs}} = \mathbf{d}$ gilt. Es ist daher möglich, die Wahrscheinlichkeitsdichtefunktion $\rho_D(\mathbf{d})$ wie folgt zu definieren [19, S. 26]:

$$\rho_D(\mathbf{d}) = \delta(\mathbf{d} - \mathbf{d}_{\text{obs}}) \quad (3.39)$$

Formel 3.39: Definition: Wahrscheinlichkeitsdichtefunktion über Daten bei perfektem Instrument

Dabei ist δ die Delta-Funktion:

$$\delta(\mathbf{x}) = \begin{cases} 1 & \text{wenn } \mathbf{x} = 0, \\ 0 & \text{sonst} \end{cases} \quad (3.40)$$

Formel 3.40: Definition: Delta-Funktion

Eine realistischere Annahme ist es, dass die Messergebnisse eine gewisse Streuung um den realen Parameter aufweisen, was sich z.B. mit einer Gaußverteilung wie folgt modellieren lässt [19, S. 254]:

$$\rho_D(\mathbf{d}_1, \mathbf{d}_2, \dots, \mathbf{d}_\alpha) = k \exp \left(-\frac{1}{2} \sum_{i=1}^{\alpha} \left(\frac{(\mathbf{d}_i - \mathbf{d}_{\text{obsi}})}{\sigma^2} \right)^2 \right) \quad (3.41)$$

Formel 3.41: Definition: Wahrscheinlichkeitsdichtefunktion über Daten bei Gauß-verteilter Abweichung

3.3.2. a-priori-Information über Modellparameter

a-priori-Informationen sind die Informationen, die sich nicht aus Messungen ergeben, sondern unabhängig davon „im Voraus“ verfügbar sind. Die a-priori-Informationen über die Modellparameter werden mit der Wahrscheinlichkeitsdichtefunktion $\rho_M(\mathbf{m})$ über \mathfrak{M} modelliert. Wenn keinerlei Vorwissen verfügbar ist, ist die Wahrscheinlichkeitsdichtefunktion gleich der homogenen Wahrscheinlichkeitsdichte, also gilt [19, S. 27]:

$$\rho_M(\mathbf{m}) = \mu_M(\mathbf{m}) \quad (3.42)$$

Formel 3.42: Modellierung der „nicht vorhandenen“ a-priori-Information über die Modellparameter

Falls Vorwissen verfügbar ist, ist es möglich, eine geeignete Wahrscheinlichkeitsdichtefunktion anzugeben, die für jeden Punkt $\mathbf{m} \in \mathfrak{M}$ einen Funktionswert als Ergebnis hat, der angibt, wie wahrscheinlich diese Modellparameterkombination ist.

3.3.3. Verbund-a-priori-Information

Da die a-priori-Information unabhängig von Messungen ist, ist es möglich das Vorwissen über Daten- und Modellparameter mit einer Verbundwahrscheinlichkeitsdichte über $\mathfrak{D} \times \mathfrak{M}$ zu modellieren [19, S. 31–32]:

$$\rho_D(\mathbf{d}, \mathbf{m}) = \rho_D(\mathbf{d})\rho_M(\mathbf{m}) \quad (3.43)$$

Formel 3.43: Modellierung der Verbund-a-priori-Information über $\mathfrak{D} \times \mathfrak{M}$

3.4. Die Lösung des inversen Problems

Mit Hilfe der Verbund-a-priori-Information $\rho_D(\mathbf{d}, \mathbf{m})$ über $\mathfrak{D} \times \mathfrak{M}$ und der theoretischen Wahrscheinlichkeitsdichte $\Theta(\mathbf{d}, \mathbf{m})$, die das Wissen über das physikalische Modell, also den Zusammenhang zwischen den Messdaten und den Modellparametern modelliert, ist es nun möglich, die a-posteriori-Information wie folgt zu definieren [19, S. 32]:

$$\sigma(\mathbf{d}, \mathbf{m}) = k \frac{\rho(\mathbf{d}, \mathbf{m})\Theta(\mathbf{d}, \mathbf{m})}{\mu(\mathbf{d}, \mathbf{m})} \quad (3.44)$$

Formel 3.44: Definition: Lösung des inversen Problems als Wahrscheinlichkeitsdichtefunktion

Aus der a-posteriori-Information über $\mathfrak{D} \times \mathfrak{M}$ lassen sich die einzelnen a-posteriori-Informationen, $\sigma_M(\mathbf{m})$ über \mathfrak{M} und $\sigma_D(\mathbf{d})$ über \mathfrak{D} als Randwahrscheinlichkeitsdichten definieren [19, S. 32]:

$$\sigma_M(\mathbf{m}) = \int_{\mathfrak{D}} d\mathbf{d}\sigma(\mathbf{d}, \mathbf{m}) \quad (3.45)$$

Formel 3.45: Definition: a-posteriori-Information über die Modellparameter

$$\sigma_D(\mathbf{d}) = \int_{\mathfrak{M}} d\mathbf{m}\sigma(\mathbf{d}, \mathbf{m}) \quad (3.46)$$

Formel 3.46: Definition: a-posteriori-Information über die Daten

3.4.1. Lösen von inversen Problemen

Die Formel 3.44 ist eine allgemeine Lösung des inversen Problems, da sie die gesamte a-posteriori-Information umfasst. Um diese allgemeine Lösung auf das Problem anwenden zu können, ist es sinnvoll, folgende Annahmen zu treffen:

Es wird angenommen, dass das physikalische Modell mit Hilfe einer Wahrscheinlichkeitsdichte beschrieben werden kann (Formel 3.44) [19, S. 33]:

$$\Theta(\mathbf{d}, \mathbf{m}) = \theta(\mathbf{d}|\mathbf{m})\mu_M(\mathbf{m}) \quad (3.47)$$

Formel 3.47: Definition: Theoretische Verbundwahrscheinlichkeit von Datenraum \mathfrak{D} und Modellraum \mathfrak{M} (physikalisches Modell)

Weiterhin wird angenommen, dass die Verbund-a-priori-Information über den Modell- und den Datenraum wie folgt definiert werden kann (Formel 3.43) [19, S. 33]:

$$\rho_D(\mathbf{d}, \mathbf{m}) = \rho_D(\mathbf{d})\rho_M(\mathbf{m}) \quad (3.48)$$

Formel 3.48: Modellierung der Verbund-a-priori-Information über $\mathfrak{D} \times \mathfrak{M}$

Da die Messinformation unabhängig von der a-priori-Information gewonnen wurde, ist der homogene Grenzwert $\mu(\mathbf{d}, \mathbf{m})$ der Formel 3.48 [19, S. 33]:

$$\mu(\mathbf{d}, \mathbf{m}) = \mu_D(\mathbf{d})\mu_M(\mathbf{m}) \quad (3.49)$$

Formel 3.49: Homogener Grenzwert der a-priori-Verbundwahrscheinlichkeit $\rho_D(\mathbf{d}, \mathbf{m})$ über $\mathfrak{D} \times \mathfrak{M}$

Damit kann unter der Verwendung der Formeln 3.44 und 3.45 die a-posteriori-Information über den Modellraum wie folgt definiert werden [19, S. 34]:

$$\sigma_M(\mathbf{m}) = k\rho_M(\mathbf{m}) \int_{\mathfrak{D}} d\mathbf{d} \frac{\rho_D(\mathbf{d}) \theta(\mathbf{d}|\mathbf{m})}{\mu_D(\mathbf{d})} \quad (3.50)$$

Formel 3.50: Definition: a-posteriori-Information über den Modellraum

Es ist möglich, diese Formel folgendermaßen umzuschreiben [19, S. 34]:

$$\sigma_M(\mathbf{m}) = k\rho_M(\mathbf{m}) L(\mathbf{m}) \quad (3.51)$$

Formel 3.51: Definition: a-posteriori-Information über den Modellraum mit likelihood-Funktion

Wobei k ein konstanter Proportionalitätsfaktor und $L(\mathbf{m})$, die likelihood-Funktion ist, die folgendermaßen definiert wird [19, S. 34]:

$$L(\mathbf{m}) = \int_{\mathfrak{D}} d\mathbf{d} \frac{\rho_D(\mathbf{d}) \theta(\mathbf{d}|\mathbf{m})}{\mu_D(\mathbf{d})} \quad (3.52)$$

Formel 3.52: Definition: likelihood-Funktion

Geringe Modellierungsunsicherheit

Wenn der Datenraum \mathfrak{D} linear, die homogene Wahrscheinlichkeitsdichte über \mathfrak{D} konstant und die Modellierungsunsicherheit verglichen mit der Messunsicherheit gering ist, kann die vom Modell abhängige Wahrscheinlichkeitsdichtefunktion für \mathfrak{D} , $\theta(\mathbf{d}|\mathbf{m})$ wie folgt definiert werden [19, S. 34]:

$$\theta(\mathbf{d}|\mathbf{m}) = \delta(\mathbf{d} - \mathbf{g}(\mathbf{m})) \quad (3.53)$$

Formel 3.53: Definition: modellabhängige Wahrscheinlichkeitsdichte $\theta(\mathbf{d}|\mathbf{m})$ für geringe Modellierungsunsicherheit

Mit der Definition für $\sigma_M(\mathbf{m})$ aus Formel 3.50 ergibt sich dann [19, S. 34]:

$$\sigma_M(\mathbf{m}) = \frac{1}{v} \rho_M(\mathbf{m}) \rho_D(\mathbf{g}(\mathbf{m})),$$

$$\text{mit } v = \int_{\mathfrak{M}} d\mathbf{m} \rho_M(\mathbf{m}) \rho_D(\mathbf{g}(\mathbf{m}))$$
(3.54)

Formel 3.54: Definition: a-posteriori-Information über den Modellraum für geringe Modellierungsunsicherheit

Diese Formel kann auch folgendermaßen geschrieben werden, wobei k ein Proportionalitätsfaktor ist [19, S. 35]:

$$\sigma_M(\mathbf{m}) = k \rho_M(\mathbf{m}) L(\mathbf{m}),$$

$$\text{mit } L(\mathbf{m}) = \rho_D(\mathbf{g}(\mathbf{m}))$$
(3.55)

Formel 3.55: Definition: a-posteriori-Information über den Modellraum für geringe Modellierungsunsicherheit mit Proportionalitätsfaktor

4. Die Lokalisierung eines Dipols als inverses Problem

In dem Buch *Inverse Problem Theory and Methods for Model Parameter Estimation* von Albert Tarantola [19] wird als Beispielproblem die Lokalisierung einer seismischen Erschütterung beschrieben, deren Druckwelle von Sensoren erfasst wird [19, S. 253–256]. Aufgrund der Verwandtheit der Problemstellungen diente das Lösungsverfahren im Buch als Anregung zur hier vorgestellten Lösung.

4.1. Modellraum

Verwendet werden die bereits in der Einführung zur Theorie der inversen Probleme aufgeführten Modellparameter, siehe Tabelle 3.2.1, S. 16.

Erwähnenswert ist hierbei, dass die Dipolposition als ein Punkt definiert wird, statt die allgemeinere Möglichkeit zu wählen und jeweils die Positionen der beiden Elektroden zu erfassen, aus denen der Dipol besteht. Diese Vorgehensweise resultiert aus der vorhandenen a-priori-Information, auf die im nächsten Abschnitt genauer eingegangen wird.

Desweiteren ist zu beachten, dass für die Positionsangabe des zu lokalisierenden Dipols ein kartesisches Koordinatensystem verwendet wird. Dies ist insofern wichtig, als dass sich die Definition und Berechnung der volumenbezogenen Wahrscheinlichkeitsdichte $\varphi(\mathbf{x})$, siehe Formel 3.10, Seite 19, einfacher gestaltet.

Falls z.B. ein Kugelkoordinatensystem verwendet werden würde, hätte ein Kugelabschnitt, der näher am Mittelpunkt liegt, ein geringeres Volumen als ein Kugelabschnitt über den gleichen Winkeln und mit dem gleichen Entfernungsinkrement, der weiter entfernt vom Mittelpunkt liegt. D.h., wenn davon ausgegangen wird, dass jede beliebige Dipolposition im Raum gleich wahrscheinlich ist, darf daraus nicht einfach geschlossen werden, dass jeder Kugelabschnitt definiert mit den gleichen Winkeln und dem gleichen Entfernungsinkrement mit der gleichen Wahrscheinlichkeit den Dipol enthält. Das bedeutet, die homogene Wahrscheinlichkeitsdichte wäre im Fall der Kugelkoordinaten im Gegensatz zu dem Fall mit einem kartesischen Koordinatensystem keine Konstante.

Eine ausführlichere Behandlung des Themas volumenbezogene Wahrscheinlichkeit wird in *Probability Density versus Volumetric Probability* [18] vollzogen.

Zusammengefasst: Da ein kartesisches Koordinatensystem verwendet wird, ist die homogene Wahrscheinlichkeitsdichtefunktion $\mu_M(\mathbf{m})$ über dem Modellraum \mathfrak{M} konstant:

$$\mu_M(\mathbf{m}) = k_{hws} \quad (4.1)$$

Formel 4.1: Definition: Spezialfall konstante homogene Wahrscheinlichkeitsdichte über kartesischem Modellraum

4.2. a-priori-Information Modellraum

Folgende a-priori-Information über den Modellraum ist bekannt:

- Der zu lokalisierende sendende Dipol steht senkrecht im Wasser.
- Der Abstand *dis* der beiden sendenden Dipolelektroden.
- Die Position der Messelektroden relativ zum Mittelpunkt des Roboters.
- Eventuell die Höhe des Sendedipols im Wasser.

Die zu lösende Aufgabe ist nun, diese a-priori-Information $\rho_M(\mathbf{m})$ über den Modellraum mit Hilfe einer Wahrscheinlichkeitsdichtefunktion zu modellieren.

Die ersten Modellparameter, die betrachtet werden, sind die, welche die Dipolposition definieren. Da am Schluss die Berechnung von einem Mikrocontroller durchgeführt werden soll, der nur begrenzte Ressourcen besitzt, ist ein Teil der a-priori-Information in die Definition des Modellraums eingeflossen. Die freiere und bei der Veränderung des Modells einfacher modifizierbare Modellierung wäre gewesen, die Position der beiden einzelnen Dipolelektroden zu Modellparametern zu erfassen und mit Hilfe der Funktion $\rho_M(\mathbf{m})$ später den Modellparameterkombinationen, für die nicht gilt, dass die beiden Elektroden übereinander liegen oder die den falschen Abstand haben, eine Wahrscheinlichkeit von null zuzuweisen. Das hätte allerdings den Nachteil, dass sich der zu durchsuchende Parameterraum stark vergrößert und für die meisten Parameterkombinationen \mathbf{m} $\rho_M(\mathbf{m}) = 0$ gelten würde. Daher ist dieser Teil der a-priori-Information über die Modellparameter in die Definition des Modellraumes eingeflossen und es wird nur ein Modellparameter, bestehend aus drei Koordinaten, für die Dipolposition verwendet.

Es wird im Folgenden davon ausgegangen, dass jede mögliche Dipolposition gleich wahrscheinlich ist, daher ist es möglich, die Formel 3.42, S. 27 zu verwenden und $\rho_M(\mathbf{m})$ mit Hilfe einer homogenen Wahrscheinlichkeitsdichte zu definieren. Da außerdem ein kartesisches Koordinatensystem verwendet wird, lässt sich die homogene Wahrscheinlichkeitsdichte für die Dipolposition durch eine Konstante k_{prioM} definieren:

$$\rho_M(\mathbf{m}) = \mu_M(\mathbf{m}) = k_{prioM} \quad (4.2)$$

Formel 4.2: Modellierung der a-priori-Information durch eine homogene Wahrscheinlichkeitsdichte über die Modellparameter bei kartesischem Koordinatensystem.

Falls a-priori-Information über den Wasserstand im Becken zur Verfügung steht und bekannt ist, dass sich der zu lokalisierende Dipol ungefähr auf der Höhe h_d über dem Grund befindet, ist es möglich, eine von h_d , h_s und dem die Höhe der Dipolposition bestimmende Koordinatenparameterdimensionskomponente $coord_{d3}$ abhängige Wahrscheinlichkeitsdichtefunktion zu definieren. Sie soll dann maximal werden, wenn sich der Dipol auf der Höhe h_d befindet. Da der Nullpunkt des verwendeten Koordinatensystems allerdings auf der Roboterposition liegt, ist das der Fall, wenn die Bedingung $h_d - h_s = coord_{d3}$ wahr ist. Zur Modellierung kann die Gaußverteilung verwendet werden, für den Erwartungswert μ gilt $\mu = h_d$ und für den Parameter x gilt $x = h_s + coord_{d3}$:

$$\rho_M(\mathbf{h}_d, \mathbf{h}_s, \mathbf{coord}_{d3}) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2} \left(\frac{(\mathbf{h}_s + \mathbf{coord}_{d3}) - \mathbf{h}_d}{\sigma}\right)^2\right) \quad (4.3)$$

Formel 4.3: Modellierung der a-priori-Information über die Modellparameter bei der Verwendung von Höheninformationen.

Die Standardabweichung σ sollte im Bereich weniger Zentimeter gewählt werden, z.B. $\sigma = 0,01$, da als Einheit Meter verwendet werden.

Desweiteren sind Informationen über die Positionen der Messelektroden innerhalb des verwendeten Koordinatensystems bekannt. Bei den Messelektroden wird davon ausgegangen, dass ihre Position exakt definiert ist und sich ihre Position auch nicht mit dem Lauf der Zeit verändert.

4.3. Datenraum

Für den Datenraum werden die bereits in der Einführung zur Theorie der inversen Probleme aufgeführten Datenparameter verwendet, siehe Tabelle 3.2.2, S. 17. Da der Datenraum ebenfalls linear ist, kann die homogene Wahrscheinlichkeitsdichte $\mu_D(\mathbf{d})$ durch eine Konstante ersetzt werden, siehe 3.37, Seite 26.

4.4. a-priori-Information Datenraum

Die a-priori-Information über den Datenraum, also die Messdaten, besteht zum einen aus dem Wissen, in welchem Bereich sich die gemessenen Spannungen befinden können und zum anderen, auf welche Weise die Messwerte d_{obs} von den tatsächlichen Parametern \mathbf{d} abweichen, die von einem perfekten Messinstrument erfasst würden.

Es existieren keine genauen Informationen darüber, wie die Messsensoren die Daten verfälschen. Allerdings ist davon auszugehen, dass die Sensoren nicht perfekt sind. Wegen der einfachen Berechnung und weil Messdaten allgemein oft um den realen Wert schwanken, wird zur Modellierung der Verteilung der Messdaten ebenfalls die Gaußverteilung genutzt, Formel 3.41:

$$\rho_D(\mathbf{d}_1, \mathbf{d}_2, \dots, \mathbf{d}_\alpha) = \rho_D(\mathbf{U}_1, \mathbf{U}_2, \dots, \mathbf{U}_n) = k \exp \left(-\frac{1}{2} \sum_{i=1}^n \left(\frac{(\mathbf{U}_i - \mathbf{U}_{\text{obsi}})}{\sigma^2} \right)^2 \right) \quad (4.4)$$

Formel 4.4: Modellierung der a-priori-Information über die Daten mit Gauß-verteilter Abweichung durch eine Wahrscheinlichkeitsdichtefunktion

4.5. Theoretisches Modell

Der Zusammenhang zwischen den Messdaten und den Modellparametern, die bestimmt werden sollen, wird mit einem mathematischen Modell hergestellt, dem die physikalischen Gesetze zugrunde liegen und das die Funktion $\mathbf{d}=\mathbf{g}(\mathbf{m})$ definiert, die das vorwärts gerichtete Problem, siehe Abschnitt 3.2.8, S. 24 beschreibt. Mit Hilfe dieser Funktion ist es möglich, die theoretische Wahrscheinlichkeitsdichtefunktion $\Theta(\mathbf{d}, \mathbf{m})$ zu definieren.

Jeder Satz von Messwerten, also die vom Roboter an den Sensoren $S_{ab}, S_{cd}, \dots, S_{yz}$ gemessenen Spannungen U_1, U_2, \dots, U_n sowie dessen über den Drucksensor gemessenen Umgebungsdruck Pr stellt einen Punkt \mathbf{d} im Datenraum dar, z.B.:

$$(U_1, U_2, \dots, U_n, Pr) = (d_1, \dots, d_{n+1}) = \mathbf{d} \in \mathfrak{D}$$

Dabei wird, wie bereits erwähnt, angenommen, dass die gemessenen Werte aufgrund der experimentellen Unsicherheit nicht exakt sind, wobei sich die Wahrscheinlichkeitsverteilung um den eigentlich korrekten Wert mit Hilfe der Gaußverteilung modellieren lässt.

Um die Stärke des elektrischen Feldes an einem bestimmten Punkt zu berechnen, wird folgende Formel aus Abschnitt 2.2.1, S. 6 verwendet, die aus dem Gauß'schen Gesetz abgeleitet werden kann [5, S. 5]:

$$\vec{E} = \frac{1}{4\pi\epsilon_0\epsilon_r} \frac{Q}{r^2} \hat{r} \quad (4.5)$$

Formel 4.5: Stärke des elektrischen Feldes, das von einer gleichmäßig geladenen Kugel in der Entfernung r vom Kugelmittelpunkt erzeugt wird.

$\frac{1}{4\pi\epsilon_0\epsilon_r} \frac{Q}{r^2}$ enthält keine Modellparameter und kann durch eine Konstante k_1 ersetzt werden:

$$\vec{E} = \frac{k_1}{r^2} \hat{r} \quad (4.6)$$

Formel 4.6: Stärke des elektrischen Feldes, das von einer gleichmäßig geladenen Kugel in der Entfernung r vom Kugelmittelpunkt erzeugt wird mit Faktor k_1

Der sendende Dipol wird als zwei Einzelladungen mit den Bezeichnungen Q_1 und Q_2 , die sich an den Positionen $P_{Q_1} = (pQ_{11}, pQ_{12}, pQ_{13})$ bzw. $P_{Q_2} = (pQ_{21}, pQ_{22}, pQ_{23})$ befinden, modelliert. Da als gegeben angenommen wird, dass der Dipol senkrecht steht und die Ladungen den Abstand dis voneinander haben, wird festgelegt, dass $pQ_{13} > pQ_{23}$ gilt, woraus sich folgende Beziehung der Ladungspositionen ergibt:

$(pQ_{11}, pQ_{12}, pQ_{13}) = (pQ_{21}, pQ_{22}, pQ_{23} + dis_3)$. Damit ist es, wie im Abschnitt 4.2 erwähnt, möglich, die Information über die Position des Dipols mit lediglich einem Modellparameter aus drei Komponenten, bestehend aus den Koordinaten im dreidimensionalen Raum, abzubilden.

Da die Information über die Polarität der Ladung in den Faktor k_1 codiert wird, ist es nötig, das Vorzeichen bei der Berechnung jeweils anzupassen.

Die an einem Sensor S_{ab} mit Elektroden an den Positionen a und b gemessene Spannung $U_i = U_{ab}$, wenn sich der sendende Dipol an der Position pQ befindet, lässt sich daher wie folgt, analog zum Ansatz in Abschnitt 2.2.3, berechnen:

$$U_i = g_i(pQ) = U(a, b, pQ) = \frac{k_1}{di_3(pQ, a)} + \frac{-k_1}{di_3(pQ - dis, a)} - \left(\frac{k_1}{di_3(pQ, b)} + \frac{-k_1}{di_3(pQ - dis, b)} \right) \quad (4.7)$$

Formel 4.7: Gemessene Spannung zwischen den beiden Elektroden an den Positionen a und b , die aus dem von der Dipolladung an Position pQ erzeugten Feld resultiert. Die Funktion di_3 berechnet dabei den Abstand von zwei Punkten in einem dreidimensionalen Koordinatensystem nach dem Satz des Pythagoras.

$$di_3(a, b) = \sqrt{(a_1 - b_1)^2 + (a_2 - b_2)^2 + (a_3 - b_3)^2} \quad (4.8)$$

Formel 4.8: Definition: $di_3(a, b)$, $a, b \in \mathbb{R}^3$, Abstand in dreidimensionalem kartesischem Raum

Formel 4.7 löst also das Vorwärtsproblem $\mathbf{d}=\mathbf{g}(\mathbf{m})$, womit der Zusammenhang zwischen Messdaten und Modellparametern hergestellt ist.

Es sei $\Theta(\mathbf{d}, \mathbf{m}) = \theta(\mathbf{d}|\mathbf{m})\mu_M(\mathbf{m})$ die theoretische Wahrscheinlichkeitsdichtefunktion. Da der verwendete Modellraum kartesisch ist, lässt sich die homogene Wahrscheinlichkeitsdichte $\mu_M(\mathbf{m})$ durch einen konstanten Faktor ersetzen. Eine noch zu behandelnde Frage ist, inwieweit das theoretische Modell der Wirklichkeit entspricht und wie diese Unsicherheit eventuell modelliert werden sollte, oder anders gesagt, wie der Faktor $\theta(\mathbf{d}|\mathbf{m})$ definiert wird.

Um die Implementierung einfacher zu halten, wird angenommen, dass keine Modellierungsungenauigkeit existiert und die bereits definierte Funktion $\mathbf{d}=\mathbf{g}(\mathbf{m})$ korrekt ist. Um die vorhandenen Modellierungsfehler bei der Berechnung zu berücksichtigen, wird bei der Verarbeitung der Daten von einer höheren möglichen Varianz ausgegangen.

4.6. Zusammensetzen der Lösungsbestandteile

Durch die Modellierung der a-priori-Information über die Modellparameter durch die Funktion 3.42,

$$\rho_M(\mathbf{m}) = \mu_M(\mathbf{m}) = k_{prioM} \quad (4.9)$$

und die Modellierung der a-priori-Information über die Daten durch die Formel 4.4,

$$\rho_D(\mathbf{d}_1, \mathbf{d}_2, \dots, \mathbf{d}_\alpha) = \rho_D(\mathbf{U}_1, \mathbf{U}_2, \dots, \mathbf{U}_n) = k \exp \left(-\frac{1}{2} \sum_{i=1}^n \left(\frac{(\mathbf{U}_i - \mathbf{U}_{obs\ i})}{\sigma^2} \right)^2 \right) \quad (4.10)$$

sowie die Definition des theoretischen Modells durch die Formel 4.7,

$$\begin{aligned} g_i(pQ) &= U_i(pQ) = U(a, b, pQ) \\ &= \frac{k_1}{di_3(pQ, a)} + \frac{-k_1}{di_3(pQ - dis, a)} - \left(\frac{k_1}{di_3(pQ, b)} + \frac{-k_1}{di_3(pQ - dis, b)} \right) \end{aligned} \quad (4.11)$$

ist es möglich, die gesuchte Lösung zu formulieren, die a-posteriori-Information über den Modellraum, σ_M , Formel 3.55:

$$\sigma_M(\mathbf{m}) = k\rho_M(\mathbf{m})\rho_D(\mathbf{g}(\mathbf{m})) \quad (4.12)$$

Diese relativ einfache Formulierung ist möglich, weil die Entscheidung getroffen wurde, die Modellierungsunsicherheiten zu vernachlässigen, siehe dazu auch den Abschnitt 3.4.1, „Geringe Modellierungsunsicherheit“, Seite 29.

Wenn die Faktoren k_1 und k_{prioM} zu k_2 zusammengefasst werden und die Formeln 4.9, 4.10, 4.11, 4.12 ineinander eingesetzt werden, ergibt sich:

$$\begin{aligned} \sigma_M(pQ) &= k_2 \exp \left(-\frac{1}{2} \sum_{i=1}^n \left(\frac{(\mathbf{U}_{cal\ i}(pQ) - \mathbf{U}_{obs\ i})}{\sigma^2} \right)^2 \right), \text{ mit} \\ U_{cal\ i}(pQ) &= U_{ab}(pQ) \\ &= \frac{k_1}{di_3(pQ, a)} + \frac{-k_1}{di_3(pQ - dis, a)} - \left(\frac{k_1}{di_3(pQ, b)} + \frac{-k_1}{di_3(pQ - dis, b)} \right) \end{aligned} \quad (4.13)$$

Formel 4.9: Lösung des inversen Problems als a-posteriori-Wahrscheinlichkeitsdichte $\sigma_M(pQ)$

Diese Wahrscheinlichkeitsdichte beschreibt die Information, die über die Dipolposition bekannt ist, nachdem alle zur Verfügung stehenden Informationen genutzt wurden.

5. Umsetzung

5.1. Einleitung

Die verwendete Theorie der inversen Probleme ist sehr mächtig in der Beschreibung von Problemen, allerdings hat sie auch Nachteile, insbesondere bei der Verwendung von Hardware mit vergleichsweise geringer Speicherkapazität und Rechenleistung und der Anforderung, die Berechnung in Echtzeit auszuführen.

Zunächst einmal ist die berechnete Lösung eine Wahrscheinlichkeitsdichte. Entsprechend visualisiert kann sie von einem Menschen relativ einfach durch „Betrachten“ ausgewertet werden. Ein Menschen erkennt z.B., dass ein Torus hoher Wahrscheinlichkeit um den Roboter herum bedeutet, dass sich der gesuchte Dipol sehr sicher im Abstand des Torusradius befindet und die Höhe relativ genau bestimmt ist; die Richtung, in der sich der Dipol befindet, allerdings nicht. Die Interpretation einer Wahrscheinlichkeitsdichte durch einen Algorithmus dagegen ist deutlich aufwändiger.

Der zweite Nachteil ist mit dem vorhergehenden verknüpft. Berechnet wird die Lösung durch die „Kartographierung“ der Wahrscheinlichkeitsdichtefunktion. Das heißt, um zur Lösung zu kommen, müssen nach einem bestimmten Verfahren mögliche Modellparameter in die Funktion eingesetzt werden. Das beste Ergebnis würde natürlich ein Verfahren liefern, das alle möglichen Werte durchrechnet. Da die verwendeten Modellparameter allerdings nicht diskret sondern kontinuierlich sind, wäre es nur möglich, eine endliche Anzahl von Punkten zu berechnen, z.B. ein Gitter. Falls mit diesem Verfahren mehrere Dipole lokalisiert werden sollen, der Datenraum also nicht nur drei Dimensionen, sondern $3 \cdot n$ Dimensionen enthielte, würde der Aufwand exponentiell mit der Anzahl der Dipole zunehmen. Außerdem wird das Verhältnis vom Volumen mit relativ hoher Positionswahrscheinlichkeit des Dipols im Vergleich zum Volumen mit niedriger Positionswahrscheinlichkeit des Dipols immer schlechter. Es wäre also nötig, mit jeder weiteren Dimension mehr Punkte zu berechnen, um genau so viele Punkte mit höherer Wahrscheinlichkeit zu finden, wie bei einer niedrigeren Dimensionszahl.

Hierzu ein Beispiel aus dem Buch *Inverse Problem Theory* [18, S. 43]: Das Verhältnis des Volumens eines Hyperwürfels $((2r)^n)$ zum Volumen einer in den Hyperwürfel einbeschriebenen Hyperkugel $(\frac{2\pi^{\frac{n}{2}} r^n}{n\Gamma(\frac{n}{2})})$ bei wachsender Dimensionszahl. Hierbei sei der Hyperwürfel das zu durchsuchende Volumen und die Hyperkugel das Volumen mit „hoher Wahrscheinlichkeit“. Mit jeder weiteren Dimension ist es weniger wahrscheinlich, dass ein zufällig ausgewählter Punkt innerhalb der Hyperkugel liegt, siehe Abbildung 5.1, S. 38.

Auf diese Nachteile sollte auf jeder Ebene der Umsetzung eingegangen werden. Zum

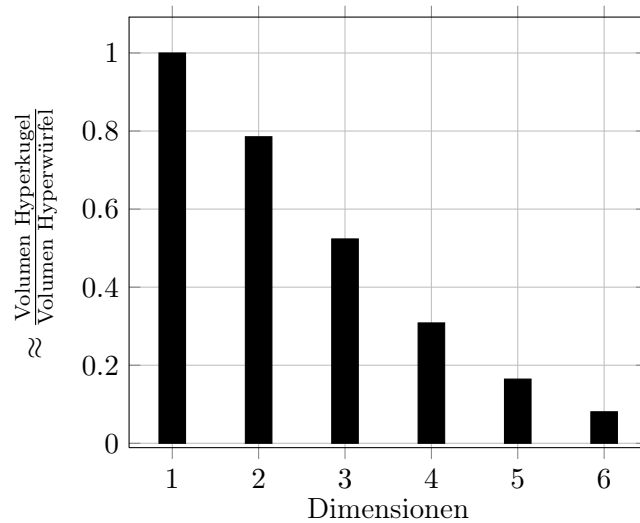


Abbildung 5.1.: Verhältnis des Volumens eines Hyperwürfels zum Volumen der in den Würfel eingeschriebenen Hyperkugel.

Beispiel wurde im Theorieteil eine relativ einfache und damit weniger rechenintensive Modellierung verwendet, in dem die Ungenauigkeit des theoretischen Modells als vernachlässigbar eingestuft wurde. Auf der Implementierungsebene ist es sinnvoll, eine Kurzschlussauswertung durchzuführen und z.B. die Berechnung abubrechen, wenn die Wahrscheinlichkeit einer bestimmten Dipolposition bereits aufgrund der Wahrscheinlichkeitsdichte $\rho_M(\mathbf{m})$ für die a-priori-Information über die Modellparameter unter eine gewisse Schwelle fällt. Auf der konzeptionellen Ebene ist es sinnvoll, nicht strikt die Wahrscheinlichkeitsdichte etwa anhand eines Gitters auszurechnen und sich dann den Punkt – also eine bestimmte Modellparameterkombination – mit der höchsten Wahrscheinlichkeit auszusuchen, sondern die Ressourcen sparsam einzusetzen und ein Verfahren einzusetzen, das möglichst schnell gegen den Punkt mit der höchsten Wahrscheinlichkeit konvergiert. Ein Verfahren, das in niedrigen Dimensionen diese Aufgabe erfüllen kann, ist das Simplex-Verfahren nach Nelder und Mead [14].

Die Grundidee dabei ist, die Modellparameter, also die gesuchte Dipolposition, als Eingabeparameter und die a-posteriori-Wahrscheinlichkeitsdichte $\sigma_M(pQ)$ als Teil der Bewertungsfunktion, für die ein Minimum gesucht wird, für den Nelder-Mead-Algorithmus zu verwenden.

5.2. Simplex-Verfahren nach Nelder und Mead

5.2.1. Übersicht

Das Simplex-Verfahren nach Nelder und Mead [14] ist auch unter den Namen Downhill-Simplex-Verfahren oder Amoeba (siehe [2]) bekannt.

Der Algorithmus wird dazu verwendet, das Minimum¹ einer Funktion mit n Eingabeparametern (oder Variablen) zu bestimmen.

Das Flussdiagramm des Algorithmus ist in Abbildung 5.7, S. 44 dargestellt.

Die Namen des Verfahrens illustrieren die Grundidee des Algorithmus. Er verwendet ein n -Simplex, also ein Polytop mit der Eigenschaft, dass jeder seiner Punkte es in eine andere Dimension erweitert. Wenn aus einem $(n-1)$ -Simplex ein n -Simplex gebildet werden soll, wird ein Punkt in einer weiteren Dimension ausgewählt und alle bisherigen niedrigerdimensionalen Punkte mit dem neu gewählten Punkt verbunden. Wenn diese Regel angewendet wird, ergibt sich jeweils um eine Dimension zunehmend die Reihe *Punkt* (*1-dimensional*), *Strecke* (*2-dimensional*), *Dreieck* (*3-dimensional*), *Tetraeder* (*4-dimensional*).

Dieses Simplex wird im Idealfall im Definitionsbereich der zu minimierenden Funktion so verschoben, dass es das Minimum „umschließt“. Dabei „fließt“ das Simplex ähnlich wie eine Amöbe, was den zweiten Namen erklärt.

5.2.2. Ablauf des Nelder-Mead-Algorithmus

Die hier vorgestellte Variante des Nelder-Mead-Algorithmus basiert auf dem Paper *Convergence Properties of the Nelder-Mead Simplex Method in Low Dimensions*[12].

Der Algorithmus hat das Ziel, das globale Minimum einer Funktion $f(x) : \mathbb{R}^n \rightarrow \mathbb{R}$ zu bestimmen, wobei x aus n Parametern besteht.

Bei jeder Iteration wird der Punkt, der den „schlechtesten“ – wenn das Minimum gesucht wird, also den höchsten – Funktionswert als Ergebnis hat, mit Hilfe bestimmter Operationen durch einen Punkt ersetzt, der ein besseres Ergebnis liefert.

Zu Beginn der k -ten Iteration, wobei $k \geq 0$, existieren die Punkte $x_1^k, x_2^k, \dots, x_{n+1}^k$, von denen jeder als Eingabe für die Funktion $f(x)$ dienen kann. Diese Punkte bilden das $n+1$ -Simplex.

Sei $f_i^k = f(x_i^k), \forall i : 1 \leq i \leq n+1$. Die Punkte x_i^k werden sortiert und umbenannt, so dass gilt $\forall y, z \in \{1, 2, \dots, n+1\}, y \neq z : f_y^k \leq f_z^k \Rightarrow y < z$. Es gilt also $f(x_1^k) = f_1^k = \min_i(f_i)$, sowie $f(x_{n+1}^k) = f_{n+1}^k = \max_i(f_i)$. Außerdem wird \bar{x} als der Schwerpunkt der Punkte x_i , für $i < n+1$, definiert:

$$\bar{x} = \sum_{i=1}^n \frac{x_i}{n} \quad (5.1)$$

Formel 5.1: Definition von \bar{x}

¹Es ist möglich, ein Maximum oder andere Werte zu bestimmen, indem die zu optimierende Funktion sinnvoll so umdefiniert wird, dass sie an der entsprechenden Stelle ein Minimum besitzt.

Die Entfernung zwischen den Punkten x_i und x_j wird mit $[x_i x_j]$ bezeichnet.

Bei jedem Iterationsschritt wird der Punkt x_{n+1} durch einen neuen Punkt ersetzt, dies geschieht durch vier Operationen:

- Reflexion
- Expansion
- innere Kontraktion
- äußere Kontraktion

Für das Verfahren werden vier skalare Parameter definiert:

- ρ , der Reflexionskoeffizient
- χ , der Expansionskoeffizient
- γ , der Kontraktionskoeffizient
- σ , der Schrumpfkoeffizient

Für die Parameter gelten die folgenden Bedingungen:

$$\rho > 0, \quad \chi > 1, \quad \chi > \rho, \quad 0 < \gamma < 1, \quad \text{und} \quad 0 < \sigma < 1, \quad (5.2)$$

[12, S. 115]

Formel 5.2: Bedingungen der Parameter des Nelder–Mead–Algorithmus

Als günstige Parameterwahl für die meisten Anwendungsfälle werden die Parameter folgendermaßen belegt:

$$\rho = 1, \quad \chi = 2, \quad \gamma = \frac{1}{2}, \quad \text{und} \quad \sigma = \frac{1}{2}, \quad (5.3)$$

[12, S. 115]

Formel 5.3: Belegung der Parameter des Nelder–Mead–Algorithmus

Ein Iterationsschritt des Nelder–Mead–Algorithmus

Im Folgenden wird Schritt für Schritt der Algorithmus erklärt, die Buchstaben in eckigen Klammern verweisen dabei jeweils auf den Knoten mit der gleichen Bezeichnung im Flussdiagramm, Abbildung 5.7, S. 44. Jede Operation, die ein neues Simplex erzeugt, wird durch eine Abbildung illustriert, wobei das aktuelle Simplex durch die mit gestrichelten Linien verbundenen Punkte und das neue Simplex durch die mit durchgezogenen Linien verbundenen Punkte dargestellt wird. Der Algorithmus wird in den Abbildungen auf

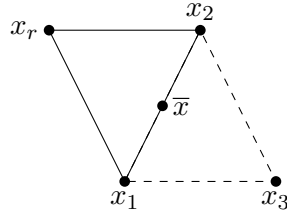


Abbildung 5.2.: Reflexion: Altes (gestrichelt) und neues Simplex mit $\bar{x}, x_r, \rho = 1$

eine Funktion mit zwei unabhängigen Variablen angewendet, daher besteht das Simplex aus drei Punkten: x_1, x_2 , und x_3 . Hierbei ist x_1 der beste Punkt mit dem niedrigsten Funktionswert und $x_3 = x_{n+1}$ der schlechteste Punkt mit dem höchsten Funktionswert. Die Werte f_i werden nicht dargestellt.

1. **Sortieren** Das Feld der Punkte x_i wird anhand der Funktionswerte f_i sortiert, um x_1, x_n und x_{n+1} zu bestimmen.² (Knoten [a]) Außerdem wird der Schwerpunkt \bar{x} der n besten Punkte, also aller x_i bis auf x_{n+1} berechnet.
2. **Reflektieren** Die Reflexion von x_{n+1} am Schwerpunkt \bar{x} wird mit x_r bezeichnet. Sie wird durch die folgende Formel definiert:

$$x_r = \bar{x} + \rho(\bar{x} - x_{n+1}) = (1 + \rho)\bar{x} - \rho x_{n+1} \quad (5.4)$$

Formel 5.4: Definition des Reflexionspunktes x_r

x_r befindet sich auf der Geraden, die x_{n+1} und \bar{x} verbindet, auf der abgewandten Seite von \bar{x} von x_{n+1} aus, mit $[x_r \bar{x}] = \rho[x_{n+1} \bar{x}]$, siehe Abbildung 5.2.2. Der Punkt x_{n+1} wird genau dann durch x_r ersetzt und der Iterationsschritt beendet, wenn die Bedingung $f_1 \leq f_r < f_n$ wahr ist. (Knoten [c], [e], [d])

3. **Expansion** Falls die Bedingung $f_r < f_1$ zutrifft, wird der Expansionspunkt x_e nach der Formel 5.5 berechnet, siehe Abbildung 5.3. Falls die Bedingung $f_e = f(x_e) < f_r$ zutrifft, der Funktionswert also in Richtung des Expansionspunktes näher an einem Minimum liegt, wird $x_{n+1} = x_e$ gesetzt und der Iterationsschritt beendet. (Knoten [d], [e], [f], [g], [h])

$$x_e = \bar{x} + \chi(x_r - \bar{x}) = \bar{x} + \rho\chi(\bar{x} - x_{n+1}) = (1 + \rho\chi)\bar{x} - \rho\chi x_{n+1} \quad (5.5)$$

Formel 5.5: Definition des Expansionspunktes x_e

²Im Originalartikel von Nelder und Mead [14] werden lediglich x_1 und x_{n+1} bestimmt. In [12] werden alle Punkte sortiert um ein definiertes Verhalten des Algorithmus zu erzwingen, falls zwei Funktionswerte gleich sind.

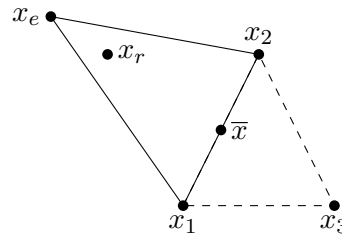


Abbildung 5.3.: Expansion: Altes (gestrichelt) und neues Simplex, mit $\bar{x}, x_r, x_e, \rho = 1, \chi = 1, 5$

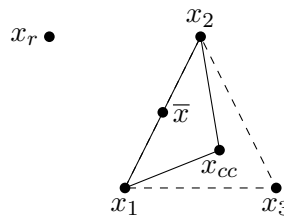


Abbildung 5.4.: Innere Kontraktion: Altes (gestrichelt) und neues Simplex, mit $\bar{x}, x_r, x_{cc}, \rho = 1, \chi = \frac{3}{2}, \gamma = \frac{1}{2}$

4. Kontraktion Bei einer Kontraktion wird das Simplex verkleinert. Dabei kann das neue Simplex entweder in Richtung des Reflexionspunktes x_r wandern, falls dieser günstiger ist – was äußere Kontraktion genannt wird – oder innerhalb des Gebietes des alten Simplex liegen, was als innere Kontraktion bezeichnet wird.

a. Innere Kontraktion Falls die Bedingung $f_{n+1} \leq f_r$ zutrifft, der Funktionswert des bisher schlechtesten Punktes x_{n+1} also mindestens genauso gut wie der Funktionswert des Reflexionspunktes x_r ist, wird eine innere Kontraktion durchgeführt, wobei $x_{n+1} = x_{cc}$ (siehe Formel 5.6) gesetzt und der aktuelle Iterationsschritt beendet wird. (Knoten [i], [m], [n], [o], [p])

$$x_{cc} = \bar{x} - \gamma(\bar{x} - x_{n+1}) = (1 - \gamma)\bar{x} + \gamma x_{n+1} \quad (5.6)$$

Formel 5.6: Definition des inneren Kontraktionspunktes x_{cc}

b. Äußere Kontraktion Falls die Bedingung $f_r < f_{n+1}$ wahr ist, der Funktionswert des Reflexionspunktes f_r also besser als der des bisher schlechtesten Punktes x_{n+1} – aber dennoch schlechter als alle anderen Funktionswerte $f_i, i \leq n$, sonst wäre eine Reflexion durchgeführt worden – ist, wird der äußere Kontraktionspunkt x_c berechnet, Definition siehe Formel 5.7. Nun wird überprüft, ob die Bedingung $f_c \leq f_r$ zutrifft, also der Funktionswert $f_c = f(x_c)$ mindestens

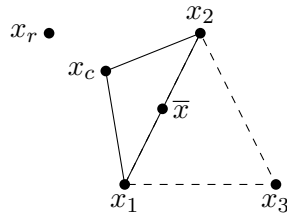


Abbildung 5.5.: Äußere Kontraktion: Altes (gestrichelt) und neues Simplex, mit $\bar{x}, x_r, x_{cc}, \rho = 1, \chi = \frac{3}{2}, \gamma = \frac{1}{2}$

genauso gut wie der des Reflexionspunktes f_r ist. Falls die Bedingung zutrifft, wird $x_{n+1} = x_c$ gesetzt. Ansonsten wird ein Schrumpfschritt durchgeführt. (Knoten [i], [j], [k], [l], [p])

$$x_c = \bar{x} + \gamma(x_r - \bar{x}) = \bar{x} + \gamma\rho(\bar{x} - x_{n+1}) = (1 + \rho\gamma)\bar{x} - \rho\gamma x_{n+1} \quad (5.7)$$

Formel 5.7: Definition des äußeren Kontraktionspunktes x_c

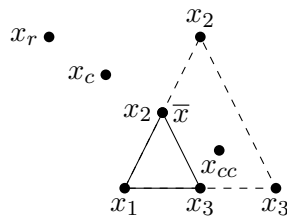


Abbildung 5.6.: Schrumpfen: Altes (gestrichelt) und neues Simplex, mit $\bar{x}, x_r, x_{cc}, \rho = 1, \chi = \frac{3}{2}, \gamma = \frac{1}{2}, \sigma = \frac{1}{2}$

5. Schrumpfen Da weder der Reflexionspunkt x_r noch einer der Kontraktionspunkte x_c oder x_{cc} einen besseren Funktionswert liefert, wird das Simplex in keine dieser Richtungen ausgedehnt, sondern der Suchbereich verkleinert.

Die Punktmenge V_i , Definition siehe Formel 5.8, wird berechnet und die Punkte x_2, \dots, x_{n+1} werden durch V_i ersetzt, siehe Abbildung 5.4. (Knoten [p])

$$V_i = \{v_i = x_1 + \sigma(x_i - x_1), i \in \{2, \dots, n+1\}\} \quad (5.8)$$

Formel 5.8: Definition der für die Schrumpfung verwendeten Punktmenge V_i

In der Simulation hat es sich als bessere Vorgehensweise herausgestellt, mehrere verteilte Startpunkte, um die jeweils ein kleineres Startsimplex aufgespannt wird, zu verwenden, als ein großes Simplex, das alle diese Startpunkte umschreibt.

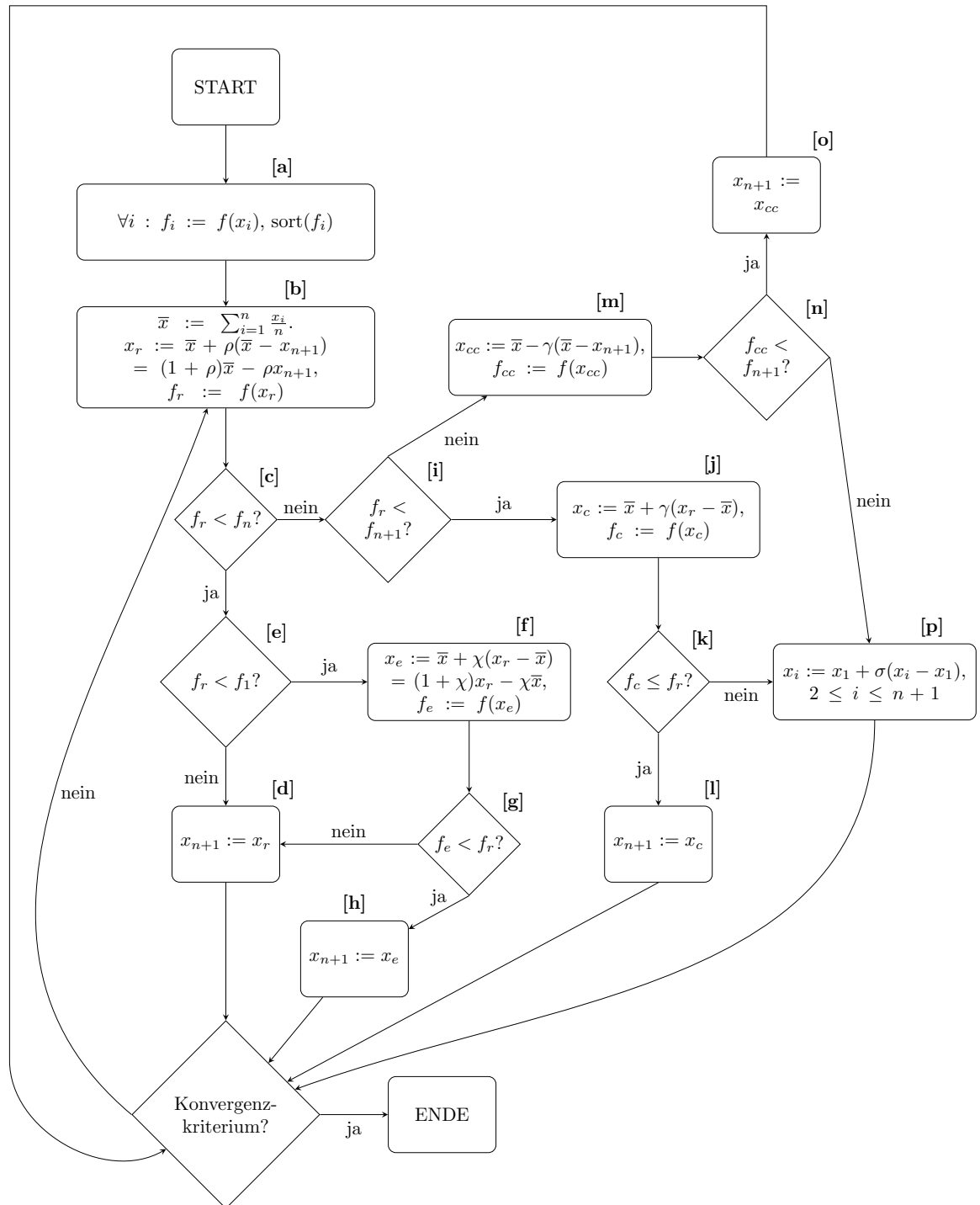


Abbildung 5.7.: Flussdiagramm des Nelder–Mead–Algorithmus

6. Implementierung

6.1. Implementierung Simulationsumgebung

Die Simulationsumgebung wurde mit Hilfe der objektorientierten Erweiterungen von Matlab erstellt. Ein vereinfachtes Klassendiagramm der Umgebung ist auf Seite 47 abgebildet. Sämtliche Objekte, die an der Simulation beteiligt sind, sind entweder direkt oder transitiv mit einem **World**-Objekt assoziiert, das die globalen Simulationsparameter enthält, wie z.B. die physische Ausdehnung des Simulationsraumes, die Permittivität des Mediums, oder wie elektrische Feldstärken an bestimmten Punkten berechnet werden. Das **World**-Objekt enthält eine Liste von **Driver**-Objekten, die jeweils eine Liste von **SimObject**-Objekten enthalten, sowie Funktionen für die Berechnung von elektrischen Spannungen zwischen Elektroden bereitstellen.¹

Von der **SimObject**-Klasse abgeleitet werden die Objekte, die an der Simulation teilnehmen. Sie können wiederum eine Liste von **SimObject**-Objekten enthalten. Zum Beispiel ist der Roboter in der Simulation von der Klasse **SimObject** abgeleitet und ist über seine **SimObject**-Liste mit seinen Messelektroden assoziiert, die eine Position relativ zu ihm besitzen. Der sendende Dipol, dessen Position bestimmt werden soll, ist ebenfalls ein **SimObject**. Die Klasse **Electrode** modelliert ein Elektrodenpaar, das sowohl aktiv senden, als auch Potentialunterschiede messen kann.

6.2. Verwendung der Simulationsumgebung

Um die Simulationsumgebung zu verwenden, müssen zuerst jeweils mindestens ein **World**- und ein zugehöriges **Driver**-Objekt erstellt werden, da sie die grundlegenden Daten und Funktionen für die Simulation bereitstellen. Mit dem Attribut *model* eines **World**-Objektes ist es möglich, die Art und Weise, wie elektrische Potentialunterschiede zwischen zwei Punkten berechnet werden, zu beeinflussen. Ist das Attribut auf 's' (= static) gesetzt, wird zur Berechnung das Gauß-Modell verwendet, das auf der Theorie von statischen elektrischen Feldern basiert. Ein weiteres Berechnungsmodell ist die Modellierung durch ein Widerstandsnetzwerk.² Falls zur Berechnung ein Widerstandsnetzwerk genutzt werden

¹Für den Einsatz in dieser Arbeit wäre ein weniger komplexes Design ausreichend gewesen, etwa, in dem **World**- und **Driver**-Objekte mit den **SimObject**-Objekten verschmolzen werden, aber mit dem gewählten Design besteht die Möglichkeit, die Simulationsumgebung erweitern zu können – etwa um eine zeitliche Komponente – ohne eine Umstrukturierung vornehmen zu müssen.

²Bei der Verwendung eines Widerstandsnetzwerkes zur Modellierung wird der simulierte Raum in ein gleichmäßiges Gitter unterteilt. Die Punkte des Gitters sind durch Ohm'sche Widerstände miteinander

soll, muss das Attribut *model* auf 'r' (= resistance) gesetzt, sowie dem Attribut *res_matrix* des **World**-Objektes eine Matrix zugewiesen werden, die die elektrischen Potentiale im Widerstandsnetzwerk repräsentiert. Bei der Verwendung des Widerstandsnetzwerkes ist zu beachten, dass sämtliche aktiven sendenden Elektroden, die bei der Verwendung des Gauß-Modelles die Simulation beeinflussen würden, ignoriert werden.

Objekte, die Teil der Simulationsumgebung sein sollen, also entweder elektrische Felder erzeugen oder messen können sollen, müssen entweder eine Instanz der Klasse **SimObject** sein, oder von dieser abgeleitet werden. Das **SimObject** dient als Verwaltungsobjekt, das eine Liste von Elektrodenpaaren der Klasse **Electrode** verwaltet, die jeweils aus zwei Elektrodenpolen bestehen. Dabei werden die Positionen jeweils relativ zum Elternobjekt angegeben. Wenn sich z.B. das oberste **SimObject** an der absoluten Position $[1, 1, 0]$, ein **Electrode**-Objekt an der (relativen) Position $[0, 1, 0]$ und ein Pol dieses Objekts an der (relativen) Position $[1, 1, 1]$ befindet, dann hat dieser Pol die absolute Position $[1, 1, 0] + [0, 1, 0] + [1, 1, 1] = [3, 2, 1]$. Diese Definition hat den Vorteil, dass es möglich ist, interaktiv ein ganzes **SimObject** mit allen Elektroden zu verschieben, indem lediglich die Position des **SimObject**s geändert wird. Natürlich ist es, z.B. in Skripten, auch möglich, alle Positionen bis auf die unterste Ebene auf $[0, 0, 0]$ zu setzen und den Elektrodenpolen direkt eine absolute Position zuzuweisen.

Wie im letzten Abschnitt erwähnt, können Elektroden sowohl senden, als auch messen. Zum Umschalten zwischen den beiden Modi kann das Attribut *mode* von 'r' (= receive) auf 't' (= transmit) gesetzt werden. Zusätzlich besitzt ein Objekt vom Typ **Electrode** noch das Attribut *isActive*, das, wenn es auf '0' gesetzt wird, den Effekt hat, dass die Elektrode auf den Rest der Simulation keinen Einfluss hat und auch von ihr nicht beeinflusst wird.

Ein kommentiertes Beispiel, wie mit dem Paket eine einfache Simulation durchgeführt werden kann, befindet sich im Anhang als Quelltext A.1, S. 77.

6.2.1. Verwendete Parameter

Name	Wert	Datentyp	Beschreibung
k	1	Float	Proportionalitätskonstante für die Berechnung der a-posteriori-Wahrscheinlichkeitsdichte
standardDeviation	0.6	Float	Standardabweichung der Gaußverteilung
sensorNum	6	Integer	Anzahl der Sensoren

Tabelle 6.2.: Allgemeine Parameter

verbunden. Innerhalb des Gitters hat ein Punkt sechs Nachbarpunkte, an einer Grenzfläche fünf und in einer Ecke drei. Der Sendedipol wird als Stromquelle bzw. -senke in jeweils einem Knoten modelliert. Mit Hilfe des Knotenpotentialverfahrens werden die elektrischen Potentiale aller Knoten des Netzwerkes berechnet und in einer Matrix gespeichert. Aus dieser Matrix kann dann das elektrische Potential an einem beliebigen Punkt ausgelesen werden. Diese noch experimentelle Methode der Modellierung wird am Institut für Parallele und Verteilte Systeme erforscht.

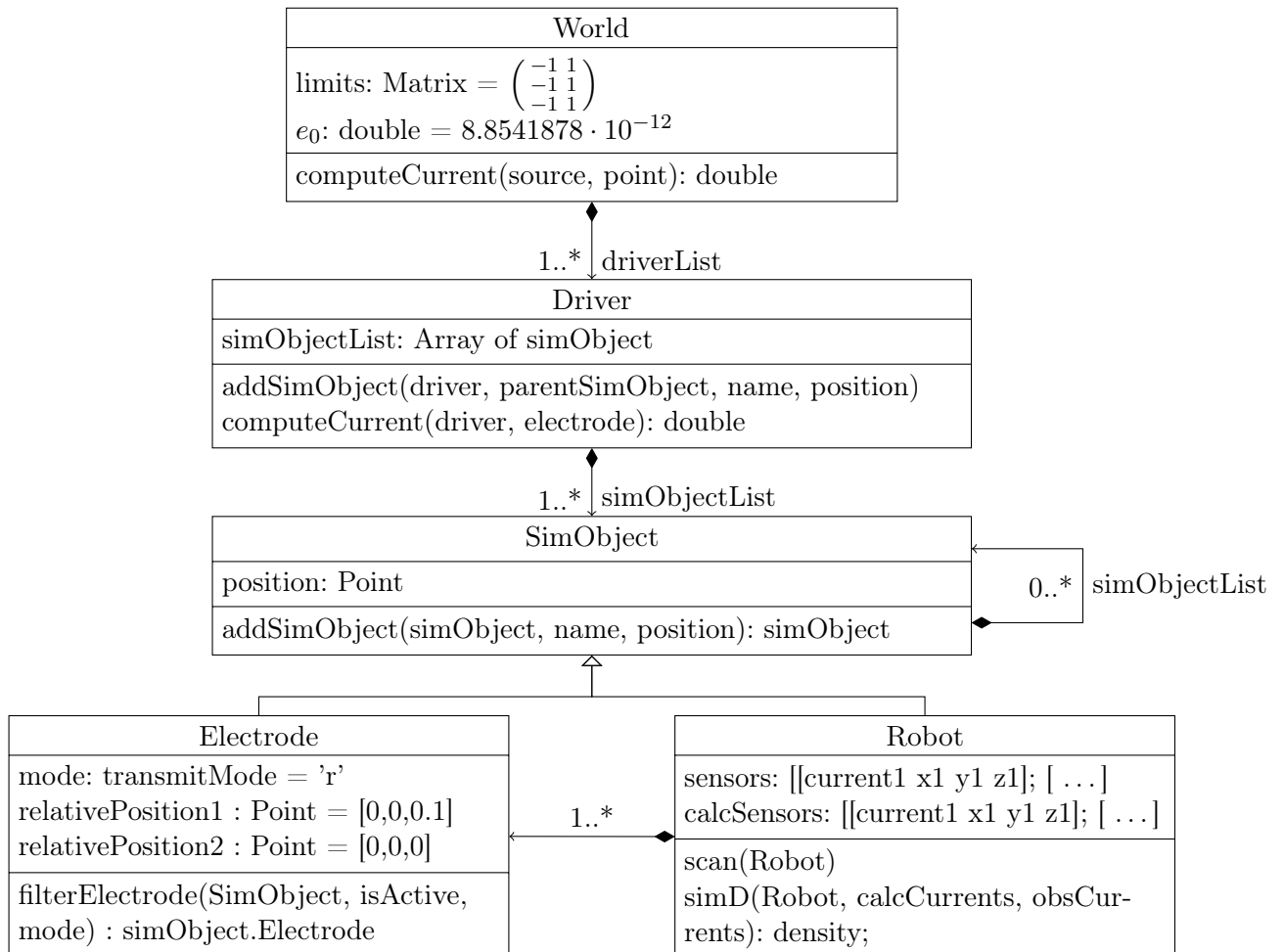


Tabelle 6.1.: UML-Klassendiagramm der Klassen **World**, **Driver**, **SimObject**, **Electrode** und **Robot**. Das Klassendiagramm soll nur einen groben Überblick vermitteln, ist daher vereinfacht und enthält nur eine Auswahl der in den Klassen definierten Attribute und Operationen.

Name	Wert	Datentyp	Beschreibung
rho	1	Float	Reflexionskoeffizient
chi	2	Float	Expansionskoeffizient
gamma	0.5	Float	Kontraktionskoeffizient
sigma	0.5	Float	Schrumpfkoeffizient
itmax	100	Integer	Maximale Anzahl der Iterationen
fmax	600	Integer	Maximale Anzahl der Funktionsaufrufe

Tabelle 6.3.: Parameter für den Nelder–Mead–Algorithmus

Name	Wert	Datentyp	Beschreibung
e_med	7.0834e-010	Float	Permittivität des Wassers
diCapacity	0.000001	Float	Kapazität einer Dipolelektrode
diCurrent	0.1	Float	Spannung einer Dipolelektrode
dipolNum	1	Integer	Anzahl der Dipole
dipolDist	0.4	Float	Abstand der Dipolelektroden in Metern

Tabelle 6.4.: Parameter für das physikalische Modell

6.3. Implementierung in Hardware

6.3.1. Messhardware

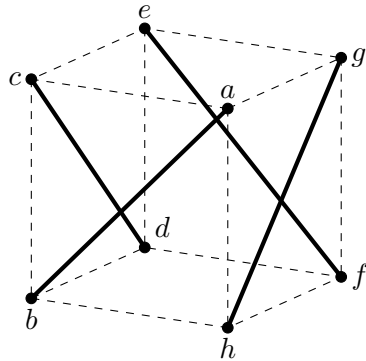
Für die Versuche wird eine kugelförmige Messsonde aus Kunststoff mit einem Radius von 0,05 m verwendet. In ihr befindet sich die Messhardware inklusive der Verstärker und der Analog–Digital–Wandler. Die Elektroden werden durch die Außenhülle gesteckt und durch Verkleben abgedichtet. An der Oberseite der Messsonde ist ein hohler Stab angebracht, der sowohl als Führung für die Stromversorgung und die Kommunikationsleitung zum Controllerboard, als auch als Griff für die Positionierung der Messsonde unter Wasser dient.

Die physische Anordnung der Elektroden der Messsonde ist zwar prinzipiell veränderbar, allerdings nur mit erheblichem Aufwand, daher werden die Versuche mit einer einzigen Elektrodenanordnung durchgeführt.

Mögliche Anordnungen der Elektroden

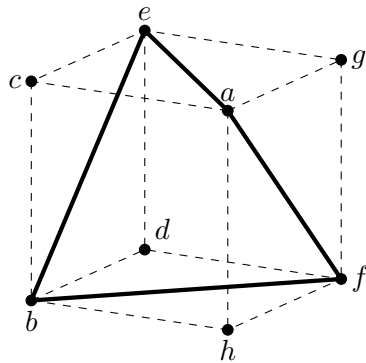
Würfecken Die acht Eckpunkte eines gedachten Würfels, die auf der Oberfläche der kugelförmigen Messsonde liegen, können als Elektrodenpolpositionen benutzt werden, z.B.:

- „Senkrechte Diagonalen im Kreis, gleiche Orientierung“



Sensor	Elektrode 1	Elektrode 2
a-b	a: (0, 0, 1)	b: (0, 1, 0)
c-d	c: (0, 1, 1)	d: (1, 1, 0)
e-f	e: (1, 1, 1)	f: (1, 0, 0)
g-h	g: (1, 0, 1)	h: (0, 0, 0)

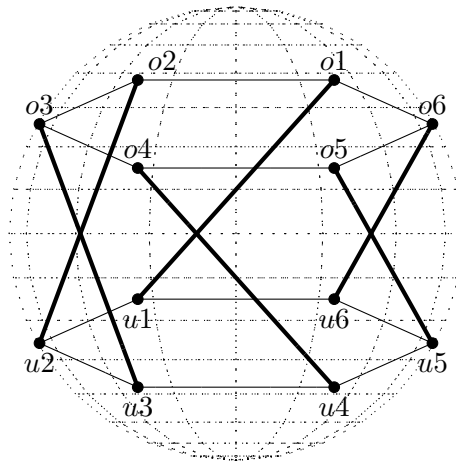
- „Verbundene Diagonalen einmal um den Würfel, von der oberen Diagonale aus“



Sensor	Elektrode 1	Elektrode 2
a-e	a: (0, 0, 1)	e: (1, 1, 1)
e-b	e: (1, 1, 1)	b: (0, 1, 0)
b-f	b: (0, 1, 0)	f: (1, 0, 0)
f-a	f: (1, 0, 0)	a: (0, 0, 1)

Ecken eines Prismas mit einem regelmäßigen Hexagon als Grundfläche Der Ursprung des Kugelkoordinatensystems, mit dem die Elektrodenpositionen angegeben werden, befindet sich am Mittelpunkt der Kugel. r bezeichnet den Abstand eines Punktes P zum Koordinatenursprung. φ bezeichnet den Winkel zwischen der positiven X-Achse und der Projektion des Vektors, der auf P zeigt, auf die X-Y-Ebene. ϑ bezeichnet den Winkel zwischen der positiven Z-Achse und des Vektors, der auf P zeigt.

Dies ist die im Versuch verwendete Elektrodenanordnung.



Sensor	Elektrode 1			Elektrode 2		
	r	φ	ϑ	r	φ	ϑ
o1-u1	0.05	0	$\frac{\pi}{6}$	0.05	$\frac{5\pi}{3}$	$-\frac{\pi}{6}$
o2-u2	0.05	$\frac{\pi}{3}$	$\frac{\pi}{6}$	0.05	0	$-\frac{\pi}{6}$
o3-u3	0.05	$\frac{2\pi}{3}$	$\frac{\pi}{6}$	0.05	$\frac{\pi}{3}$	$-\frac{\pi}{6}$
o4-u4	0.05	π	$\frac{\pi}{6}$	0.05	$\frac{2\pi}{3}$	$-\frac{\pi}{6}$
o5-u5	0.05	$\frac{4\pi}{3}$	$\frac{\pi}{6}$	0.05	π	$-\frac{\pi}{6}$
o6-u6	0.05	$\frac{5\pi}{3}$	$\frac{\pi}{6}$	0.05	$\frac{4\pi}{3}$	$-\frac{\pi}{6}$

Sendedipol

Als Sendedipol wird ein hohler Kunststoffstab mit zwei ringförmigen Sendelektroden in einem Abstand von 0,4 m verwendet. Um den Sendedipol senkrecht und in der richtigen Position zu halten, ist an der Oberseite ein Schwimmer mit quadratischer Grundfläche befestigt.

Analog-Digital-Wandler

Erfasst werden die gemessenen Spannungen, nachdem das Signal durch einen einfachen Hochpassfilter mit einer Grenzfrequenz von ca. 30 Hz gefiltert und mit Hilfe von Differenzverstärkern des Typs AD8226 um den Faktor 1000 verstärkt wurde, durch einen Analog-Digital-Wandler des Typs AD7656. Dieser digitalisiert Eingangsspannungen im Bereich von -5 V bis +5 V mit einer Genauigkeit von 16 Bit. Der Analog-Digital-Wandler kommuniziert mit dem Mikrocontrollerboard über einen SPI-Bus. Über den SPI-Bus kann der Digitalisierungsvorgang gestartet und gestoppt werden. Außerdem werden über den Bus die Daten durch den Mikrocontroller ausgelesen.

Mikrocontrollerboard

Das verwendete Mikrocontrollerboard ist das Entwicklungsboard CY8CKIT-014 mit einem PSoC 5 (Programmable-System-on-Chip) von Cypress Semiconductor. Das PSoC 5 hat einen Cortex-M3 Prozessor, sowie 256 KiB Flashspeicher und 64 KiB SRAM. Außerdem verfügt das Board über Steckerleisten, mit Hilfe derer die GPIOs kontaktiert werden können und Peripherie, wie z.B. LEDs und eine USB-Schnittstelle. Über die USB-Schnittstelle kann das PSoC sowohl programmiert, als auch das Board mit Strom versorgt werden.

Zur Kommunikation mit dem PC, der die Messdaten auswertet, werden während der Versuche GPIO-Pins des Boards als serielle Schnittstelle konfiguriert.

Das Mikrocontrollerboard fragt ein Mal pro Millisekunde über den SPI-Bus die Messdaten aller Kanäle des Analog-Digital-Wandlers – und damit die Spannungen an allen Sensoren – ab. Diese Daten werden zwischengespeichert und über eine RS232-Schnittstelle mit einer Geschwindigkeit von 921.600 Bits/s an einen PC übertragen, der die Messdaten speichert. Ein Datensatz hat folgendes Format:

Bytes Nr.	Daten
1–2	Zeitstempel
3–4	Spannung Sensor 1
5–6	Spannung Sensor 2
7–8	Spannung Sensor 3
9–10	Spannung Sensor 4
11–12	Spannung Sensor 5
13–14	Spannung Sensor 6
15–16	Zeitstempel

Tabelle 6.5.: Aufbau der vom Mikrocontroller an den PC übertragenen Datensätze

Die relativ niedrige Abtastrate wurde aus zwei Gründen gewählt. Zum einen führt ein zu häufiges Abfragen des Analog-Digital-Wandlers zu verfälschten Daten, zum anderen traten Übertragungsfehler über die RS232-Schnittstelle auf. Auch aus diesem Grund wird sowohl zu Beginn als auch am Ende eines Datensatzes ein Zeitstempel übertragen. Der Zeitstempel ist ein vorzeichenloser 16-Bit-Integer, der bei jedem Datensatz um 1 erhöht wird.

6.3.2. Software

Zur Speicherung und späteren Verarbeitung der vom Mikrocontrollerboard erfassten Messdaten wird die Software Matlab verwendet, die eine Anbindung an die serielle Schnittstelle bietet. Eine genauere Darstellung der Verarbeitung erfolgt im nächsten Kapitel anhand von Beispieldaten.

7. Versuch

7.1. Versuchsaufbau

Der Versuch wird in einem mit Wasser gefüllten Glasbecken mit den Abmessungen $2,0\text{ m} \times 1,4\text{ m} \times 0,9\text{ m}$ (Breite \times Tiefe \times Höhe), wobei der Wasserspiegel $0,6\text{ m}$ Höhe beträgt, durchgeführt. Zur Bezeichnung von Positionen wird folgendes Koordinatensystem verwendet: Der Nullpunkt ($x = 0, y = 0, z = 0$) liegt in der „vorderen linken unteren Ecke“ des Beckens, beziehungsweise der hier abgebildeten Schaubilder des Beckens. Die X-Achse verläuft vom Nullpunkt aus entlang der längsten Kante des Beckens. Eine Darstellung des Versuchsaufbaus zeigt Abbildung 7.3.

An jedem der in der Tabelle 7.1 aufgeführten Messpunkte werden ca. 30 s lang Messdaten aufgezeichnet.

Die untere Sendelektrode befindet sich auf der Position $(0,13, 0,13, 0,15)$, die obere Sendelektrode befindet sich auf der Position $(0,13, 0,13, 0,55)$. Alle 500 ms erzeugt der Sendedipol einen Rechteckimpuls nach folgendem Schema: $+5\text{ V}$ für 10 ms , -5 V für 10 ms , 480 ms Pause, siehe folgende Abbildung:

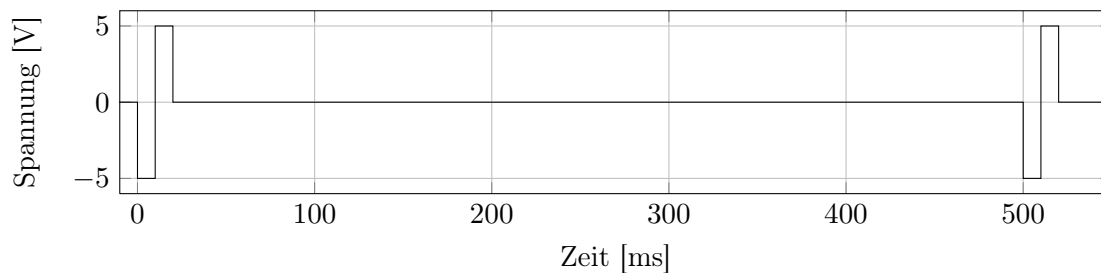


Abbildung 7.1.: Vom Sender ohne Last erzeugtes Signal

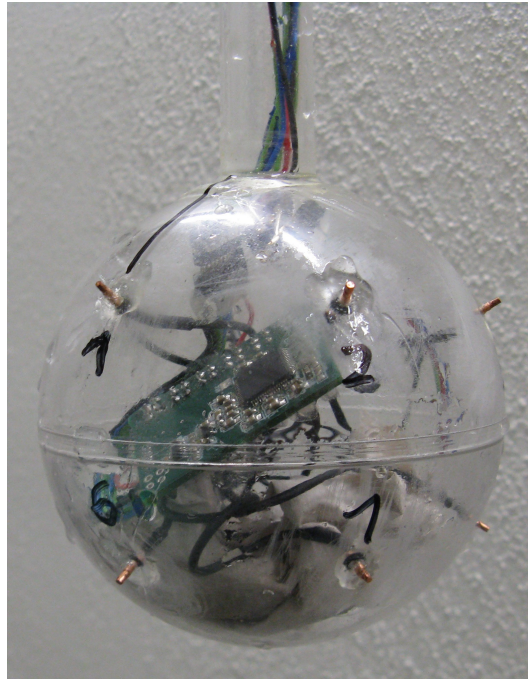


Abbildung 7.2.: Messsonde

Messung Nr.	Breite in m	Tiefe in m	Höhe in m	Ausrichtung obere Elektrode 1 in Richtung Achse
1,2	0,20	0,60	0,40	X
3	0,20	0,80	0,40	X
4	0,20	1,20	0,30	X
5	0,50	0,50	0,30	X
6	0,50	0,80	0,50	X
7	0,50	1,10	0,20	X
8	0,80	0,10	0,40	X
9	0,80	0,60	0,30	X
10	1,20	0,70	0,40	X
11	1,50	0,50	0,30	X
12	1,50	0,50	0,50	X
13	0,50	1,10	0,20	Y
14	0,50	0,80	0,50	Y
15	0,80	0,10	0,40	Y

Tabelle 7.1.: Positionen der Messsonde

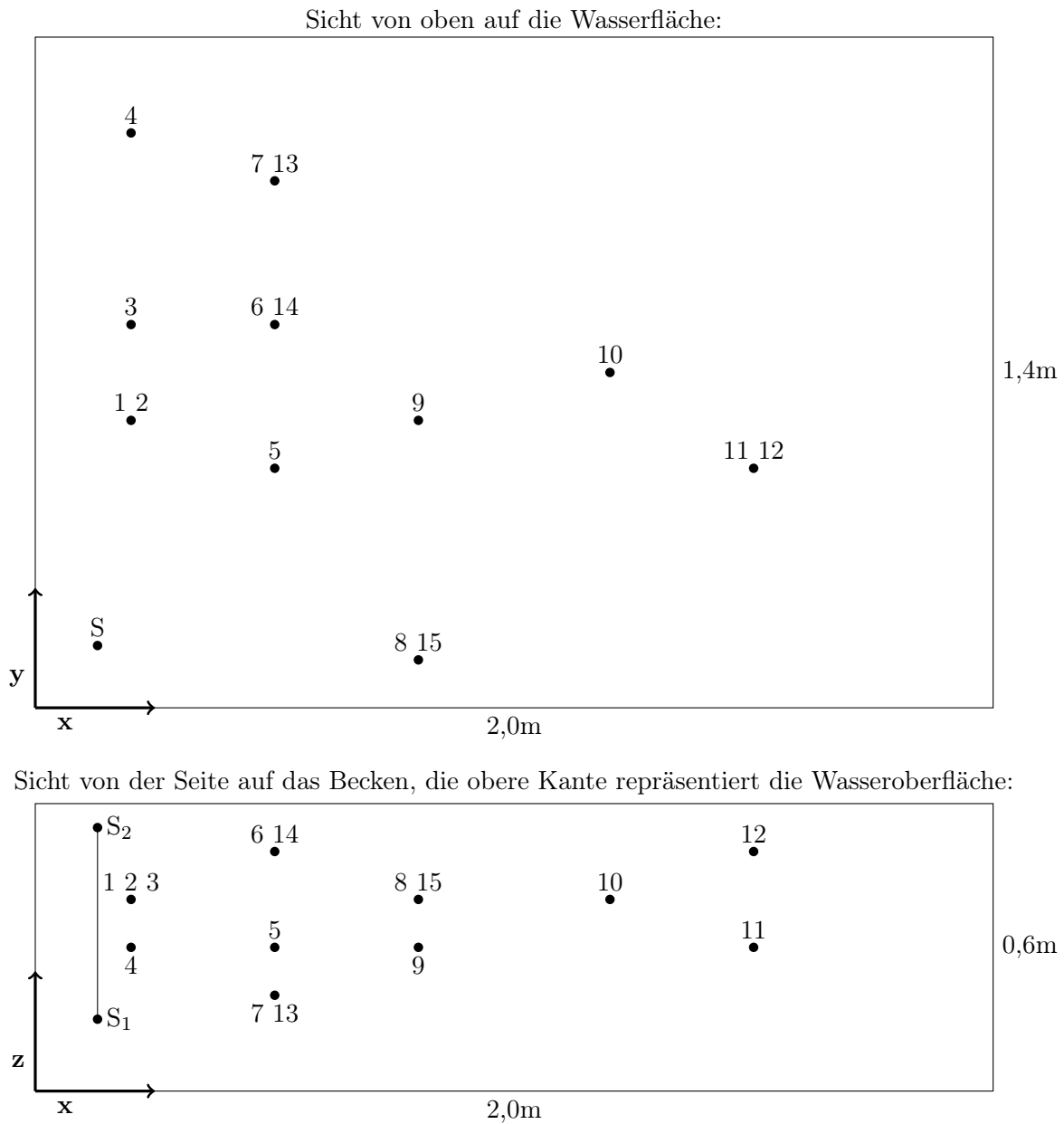


Abbildung 7.3.: Versuchsaufbau. Die Punkte 1–15 bezeichnen die verschiedenen Positionen der Messsonde, siehe Tabelle 7.1, Seite 54. Die Punkte S, S₁ und S₂ bezeichnen die Positionen der Elektroden des Sendedipols. $S_1 = [0,13 \ 0,13 \ 0,15]$, $S_2 = [0,13 \ 0,13 \ 0,55]$

7.2. Erfasste Messdaten

In diesem Abschnitt werden Beispiele der während des Versuchs erfassten Messdaten dargestellt. Desweiteren wird diskutiert, welche Folgen sich aus den Eigenschaften der Messdaten für die Weiterverarbeitung ergeben und welche Verbesserungsmöglichkeiten es geben könnte.

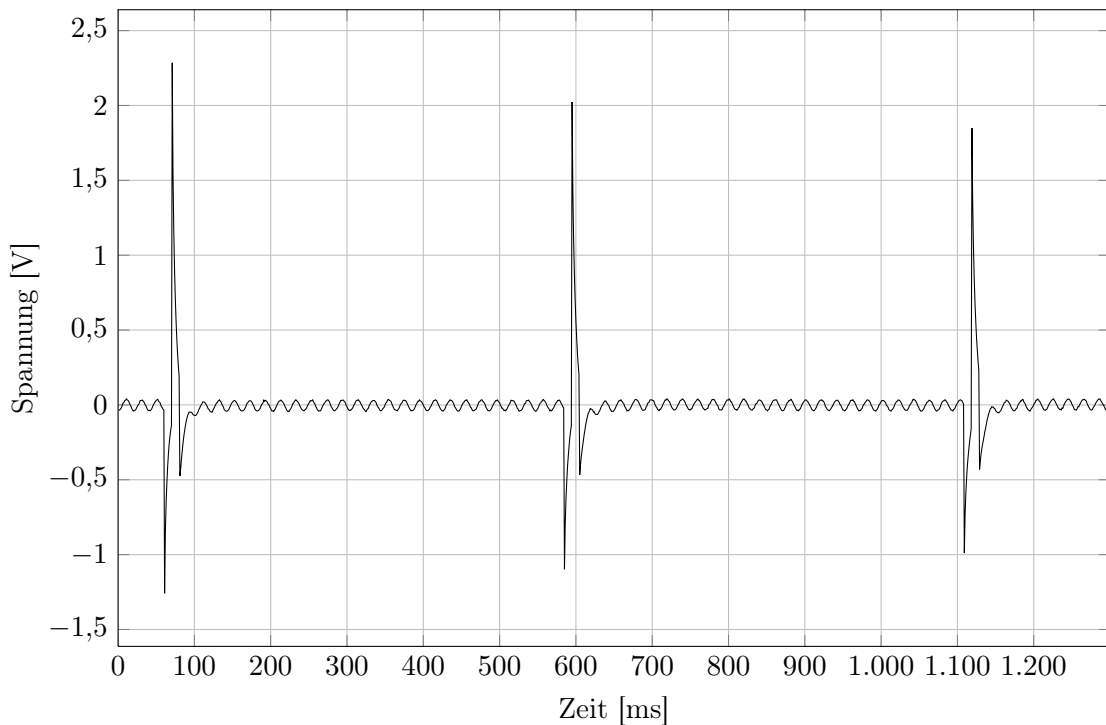


Abbildung 7.4.: Ausschnitt aus den Messdaten aus Messung Nr. 3, erfasst von Sensor 1. Gutes Verhältnis von Nutzsignal und Rauschen.

Abbildung 7.4 stellt einen Ausschnitt der von Sensor 1 erfassten Messdaten an der Position 3 dar. Die Nutzsignale des Senders heben sich deutlich vom Rauschen ab, aber die maximale und minimale Spannung ist nicht in einem Bereich, in dem der Verstärker oder der Analog-Digital-Wandler übersteuert wird. Das bedeutet, dass sich aus Messdaten dieser Art am Besten die nötigen Informationen gewinnen lassen. Anhand der Abbildung lässt sich aber auch ein Problem erkennen: Obwohl die Messsonde nicht bewegt wird, verändert sich die maximale und minimale erfasste Spannung bei jedem Ausschlag. Der Grund hierfür liegt wahrscheinlich in der im Verhältnis zum Nutzsignal relativ niedrigen Abtastrate. Der Sprung von der Nulllinie zu einem Minimum oder einem Maximum liegt zeitlich zwischen zwei Messvorgängen. Daher wird bei der Messung eine Spannung registriert, die womöglich durch den Hochpassfilter zu einem Teil bereits

wieder abgeklungen ist. Hieraus folgt für die Verwendung der gemessenen Daten, dass nicht davon ausgegangen werden kann, dass eine erfasste Spannung der real an den Elektroden anliegenden Spannung entspricht. Die Verhältnisse der zu einem Zeitpunkt gemessenen Spannungen untereinander entspricht allerdings dem Verhältnis der realen Spannungen untereinander, da alle Spannungen zum selben Zeitpunkt erfasst werden. Eine Möglichkeit, diesem Problem zu begegnen, die auch in der Auswertung angewendet wird, ist die Normalisierung der Messdaten. D.h., die gemessenen Spannungen werden unter Beibehaltung ihrer reaktiven Größenverhältnisse untereinander z.B. in das Intervall $[0, 1] \in \mathbb{R}$ abgebildet, wobei die kleinste Spannung nicht auf 0 gesetzt, sondern im korrekten Verhältnis zur größten Spannung skaliert wird. Hierbei gehen Informationen über die absolute gemessene Spannung verloren. Allerdings wäre eine Verwendung des Wissens, dass die realen Spannungen höher als die gemessenen sein könnten, deutlich aufwändiger.

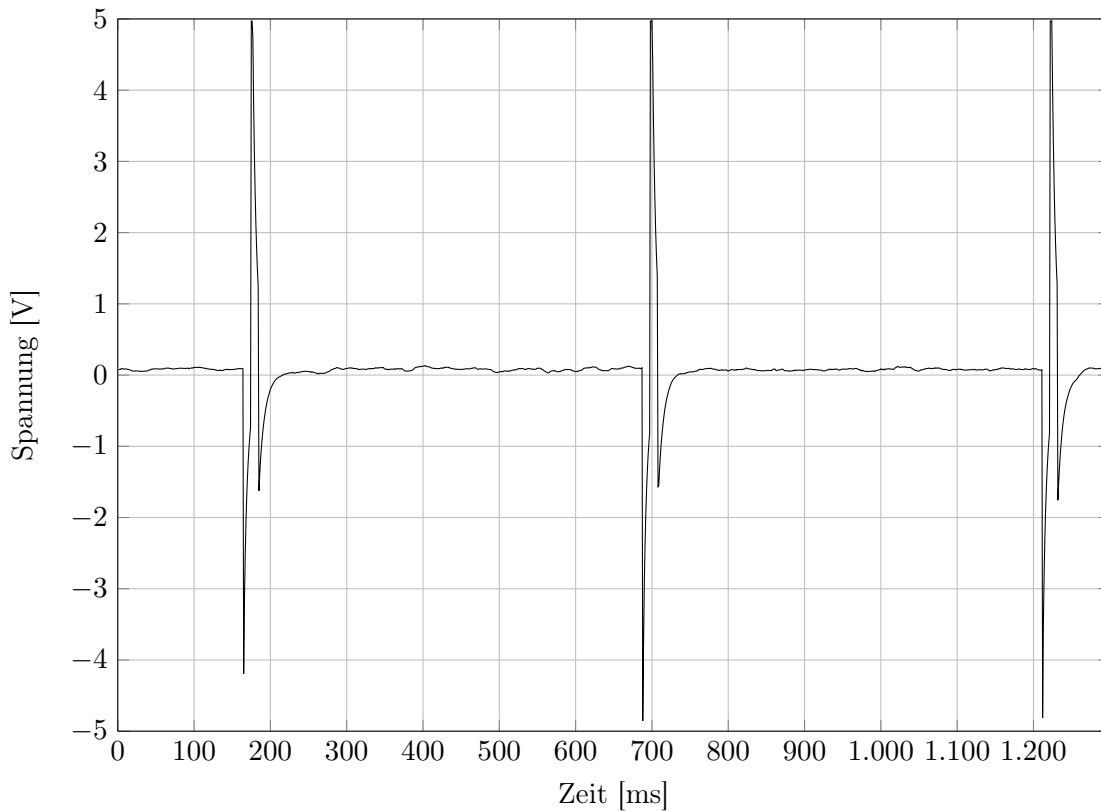


Abbildung 7.5.: Ausschnitt aus den Messdaten aus Messung Nr. 2, erfasst von Sensor 5. Da der Analog-Digital-Wandler einen Eingangsspannungsbereich von -5 V bis +5 V hat, lässt sich aus dem Signal ablesen, dass der Sensor vermutlich übersteuert ist.

Abbildung 7.5 illustriert ein weiteres Problem der Messdaten. Wenn sich die Messsonde in der Nähe des Senders befindet, kann es passieren, dass der Verstärker übersteuert und nicht die gesamte Dynamik des Nutzsignales erfasst wird. In der Abbildung 7.5 ist außerdem, analog zur Abbildung 7.4, zu erkennen, dass die erfassten Minima aufgrund der Erfassungszeitpunkte nicht die gleichen Werte aufweisen.

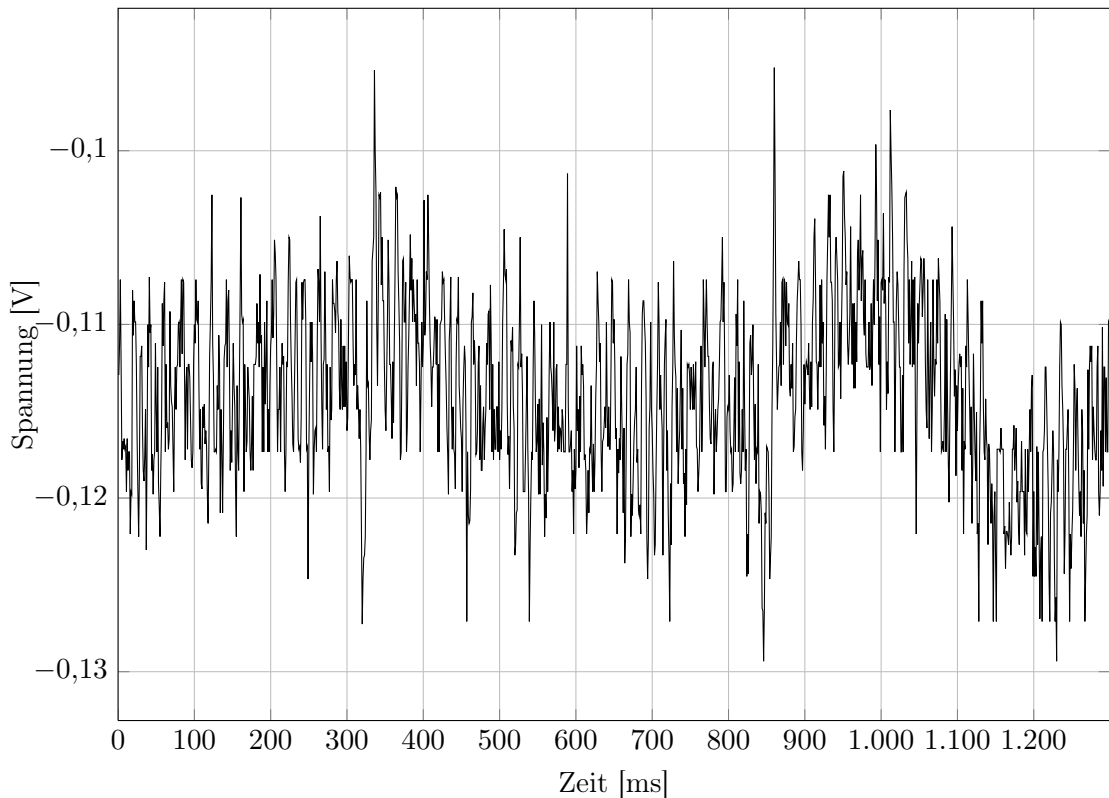


Abbildung 7.6.: Ausschnitt aus den Messdaten aus Messung Nr. 12, erfasst von Sensor 2. Es ist aufwändiger, das von Rauschen überlagerte Nutzsignal herauszufiltern und die Spannungsdifferenzen betragen nur noch wenige Vielfache der Auflösung des Analog-Digital-Wandlers.

Abbildung 7.6 zeigt einen Ausschnitt der während der Messung Nr. 12 erfassten Messwerte des Sensors 2. Bei dieser Messung befand sich die Messsonde in relativ großer Entfernung zum Sendedipol. Daraus resultiert das relativ schwache erfasste Nutzsignal, das vom Rauschen überlagert wird. Dadurch steigt bei zunehmender Entfernung der Einfluss eventueller Fehlerquellen der Messelektronik, wie z.B. die nicht perfekte Linearität der eingesetzten Verstärker oder die Auflösung des Analog-Digital-Wandlers.

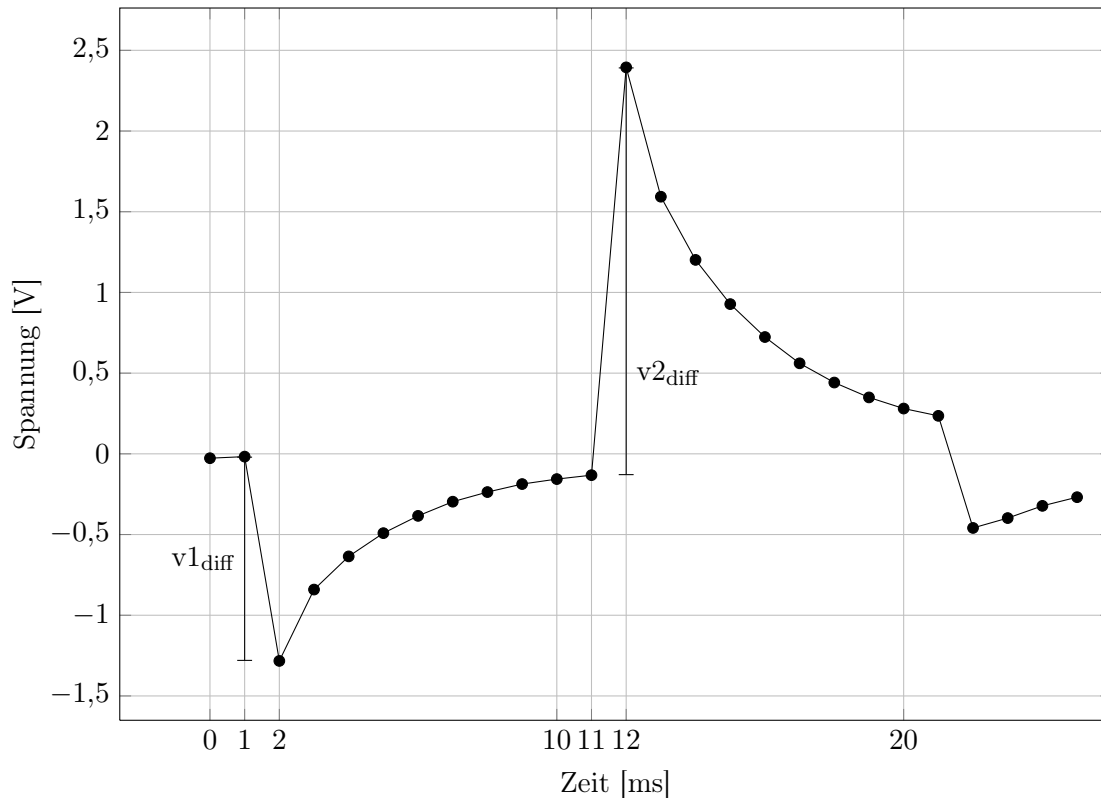


Abbildung 7.7.: Ausschnitt aus erfassten Messdaten an Sensor 1 in Messung Nr. 3. Für die Berechnung verwendet wird die Differenz v_{diff} der Spannungen der Zeitpunkte 1 und 2.

Abbildung 7.7 zeigt einen Ausschnitt der während der Messung Nr. 3 erfassten Messdaten des Sensors 1, die ein Nutzsignal des Senders enthalten. Die Punkte stellen die im zeitlichen Abstand von 1 ms erfassten Spannungen dar. Die Kurve zeigt eine Verformung, die vor allem durch den eingesetzten Hochpassfilter erzeugt wird.

Es ist anzunehmen, dass das Verhältnis der Differenz $v1_{\text{diff}}$ der Spannungswerte der Zeitpunkte 1 ms und 2 ms und der Differenz $v2_{\text{diff}}$ der Spannungswerte der Zeitpunkte 11 ms und 12 ms für alle Messungen annähernd gleich bleibt. Da die Differenz $v1_{\text{diff}}$ außerdem proportional zur Differenz des Maximums und des Minimums eines Peaks ist, kann für die Auswertung $v1_{\text{diff}}$ verwendet werden. Dies hat mehrere Vorteile. Die Suche eines Minimums ist algorithmisch einfacher zu handhaben. Da die Spannung von der „Nulllinie“ innerhalb eines Zeitraumes von weniger als 1 ms zum Minimum absinkt, kann $v1_{\text{diff}}$ einfach berechnet werden, indem die Differenz der Spannung des erkannten Minimums zum vorhergehenden Messwert gebildet wird. Da außerdem $v1_{\text{diff}} < v2_{\text{diff}}$ gilt, ist die Wahrscheinlichkeit höher, dass der Verstärker bei der Erfassung des Minimums nicht

übersteuert wurde.

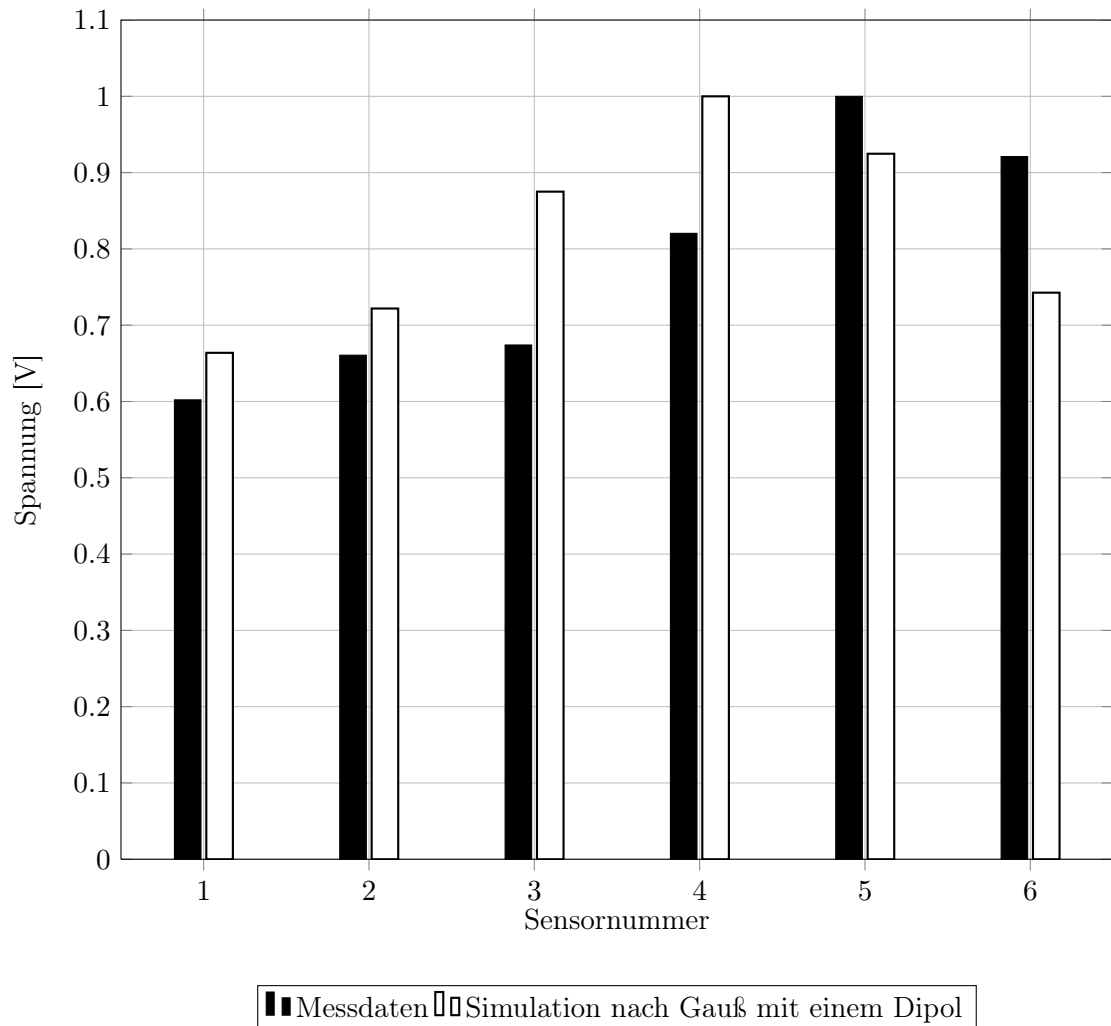


Abbildung 7.8.: Arithmetisches Mittel der in Position 4 gemessenen Spannungen, sowie simulierte Spannungen an derselben Position, jeweils so skaliert, dass die maximale Spannung 1 V beträgt.

Die Abbildung 7.8 zeigt das arithmetische Mittel der in Position 4 von der Messsonde erfassten Spannungen, sowie die in der nach dem Gauß'schen Gesetz modellierten Simulation berechneten Spannungen an derselben Position, jeweils so skaliert, dass die maximale Spannung 1 V beträgt. Die Standardabweichung der Differenz der real gemessenen Spannungen sowie der Spannungsverteilung, die sich als Ergebnis einer Simulation ergibt, stellt die Basis für die Berechnung der Aufenthaltswahrscheinlichkeit der Messsonde an einer

bestimmten Position dar.

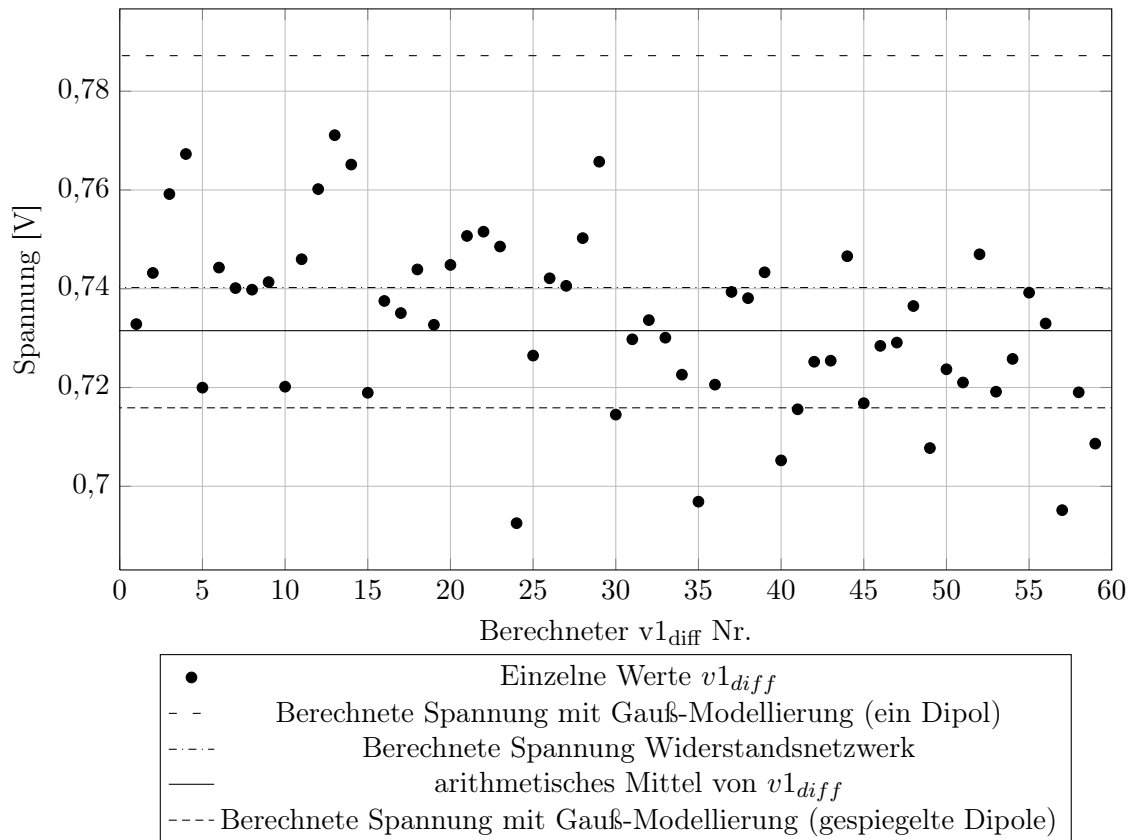


Abbildung 7.9.: Diagramm aller berechneter $v_{1_{diff}}$ -Werte von Sensor 1 aus der Messung an Position 3, sowie der Durchschnitt dieser Werte und die berechneten Spannungen aus verschiedenen Simulationsmodellen.

Die Abbildung 7.9 stellt alle berechneten $v_{1_{diff}}$ -Werte des Sensors 1 aus der Messung an Position 3, sowie den Durchschnitt dieser Werte und die berechneten Spannungen, die ein simulierter Sensor an derselben Position in verschiedenen Simulationsmodellen gemessen hat, dar. Alle Spannungen werden dabei so skaliert, dass die maximale Spannung aus jedem Datensatz 1 V beträgt, wobei die Spannungswerte der übrigen Sensoren um den gleichen Faktor skaliert werden.

An der Abbildung 7.9 lässt sich erkennen, dass die einzelnen Messwerte um den Durchschnitt herum streuen, was zu einer Streuung der $v_{1_{diff}}$ -Werte führt. Außerdem wird deutlich, dass die verschiedenen Modelle verschieden stark von der real gemessenen Spannungsverteilung abweichen. Zum Vergleich ist das entsprechende Diagramm für die Messung an Position 4 in Abbildung 7.10 zu sehen.

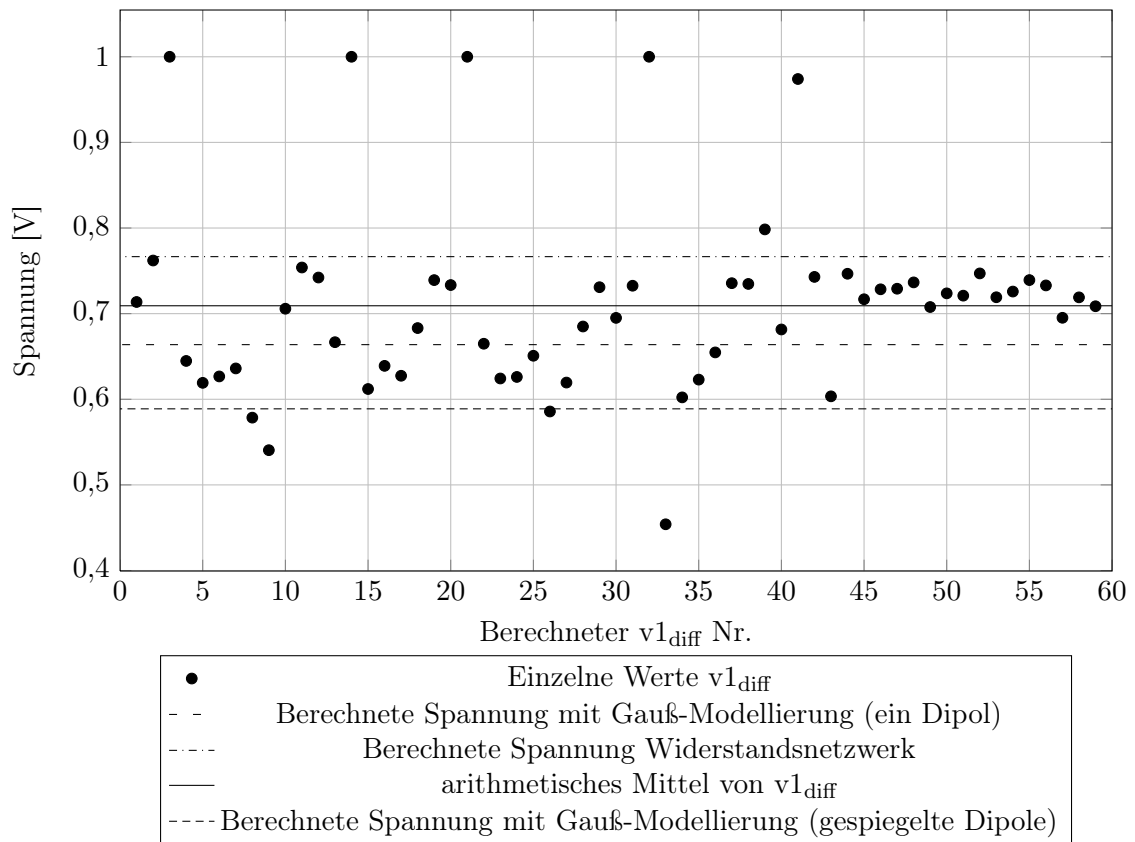


Abbildung 7.10.: Diagramm aller berechneter $v_{1\text{diff}}$ -Werte von Sensor 1 aus der Messung an Position 4, sowie der Durchschnitt dieser Werte und die berechneten Spannungen aus verschiedenen Simulationsmodellen.

Das Problem der niedrigen Abtastrate, die dazu führt, dass nicht genug Messwerte erfasst werden, um die absoluten Maxima- und Minimawerte zu bestimmen lässt sich vermutlich mit einem neuen Design der Verstärkerschaltung erreichen. Hierbei könnte es allerdings notwendig sein, die Minimaerkennung bereits im Mikrocontroller durchzuführen, um die Menge der an den PC zu übertragenden Daten zu reduzieren. Spätestens wenn die Lokalisierung auf der Roboterplattform durchgeführt werden soll, wird dieser Schritt sowieso notwendig sein.

Das Problem der Übersteuerung bei zu geringer Entfernung zum Sender bzw. der zu geringen Auflösung bei zu großer Entfernung zum Sender ließe sich vielleicht durch eine Messhardware lösen, bei der der Verstärkungsfaktor vom Mikrocontroller verändert werden kann, damit der Analog-Digital-Wandler in einem günstigen Bereich betrieben wird.

7.3. Anwendung des Lokalisierungsverfahrens auf die erfassten Messdaten

7.3.1. a-priori-Information

Da der Versuch nicht mit der ursprünglich geplanten Roboterplattform durchgeführt werden konnte, ergeben sich einige Änderungen in der zur Lösung des Problems verwendbaren a-priori-Information und der variablen Modellparameter sowie der zur Verfügung stehenden Daten.

In der Theorie wurde bisher von einem unbegrenzten Raum ausgegangen, in dem ein Roboter einen Sendedipol lokalisieren sollte. Daher wurde der Ursprung des Koordinatensystems in den Roboter gelegt. Bei der Anwendung des Verfahrens auf den Versuch ist es allerdings sinnvoller, das bereits eingeführte Koordinatensystem für das Wasserbecken zu verwenden.

Da die Maße des Beckens, oder genauer gesagt, des Wasserquaders, bekannt sind, können diese in die Simulation einfließen. Die Begrenzungen des Wasserquaders werden hierbei mit Hilfe von Wahrscheinlichkeitsdichtefunktionen modelliert. In der Implementierung wurde hierbei für Messsondenpositionen innerhalb des Wassers eine konstante Wahrscheinlichkeit angenommen, für Positionen außerhalb des Wassers eine Wahrscheinlichkeit von 0.

Die Wände des Beckens bewirken nicht nur eine Einschränkung der möglichen Positionen der Messsonde, sondern haben auch einen Einfluss auf das vom Sendedipol erzeugte elektrische Feld [16, S. 44–55] [6, S. 68–70]. Da die Energie das Wasser größtenteils nicht verlässt, werden die Feldlinien verdichtet, das elektrische Feld wird also verstärkt und verformt. Auch aus diesem Grund werden bei der Anwendung verschiedene Simulationsmodelle verwendet.

Bei diesem Versuch ist die Position des Senders, im Gegensatz zur Position der Messsonde, bekannt, kann also als a-priori-Information verwendet werden.

Anders als die Roboterplattform verfügt die Messsonde nicht über einen Drucksensor, daher kann aus den Messdaten keine direkte Information über die Tiefe, in der sich die Messsonde befindet, gewonnen werden.

7.3.2. Verwendete Simulationsmodelle

Der Erfolg des Lokalisierungsverfahrens hängt in hohem Maße von der Korrektheit der Modellierung ab, da die relativen Spannungsverhältnisse der realen Messdaten mit den simulierten relativen Spannungsverhältnissen an einer möglichen Position der Messsonde verglichen wird. Aus diesem Grund werden zur Auswertung der Messdaten mehrere denkbare Simulationsmodelle verwendet und die Ergebnisse gegenübergestellt.

Modell 1: Modellierung als Widerstandsnetzwerk

Bisher wurde in dieser Arbeit als Modellierung die Annahme verwendet, dass der Strom entlang der elektrischen Feldlinien fließt und daher die Feldlinien statischer elektrischer

Felder zur Modellierung genutzt werden können. Eine andere Betrachtungsweise ist die, das Wasser als einen elektrischen Leiter zu sehen, ihn in ein dreidimensionales Gitter aus Widerständen abzubilden und zu berechnen, wie sich die Ströme auf dieses Widerstandsnetzwerk verteilen. Diese Modellierung führt zu einem lösbaeren linearen Gleichungssystem. Mit Hilfe der Lösung lässt sich eine Matrix bilden, die bestimmten Punkten im Wasserquader ein elektrisches Potential zuweist.

Der Vorteil dieser Methode ist, dass die Begrenzungen des Wasserquaders modelliert werden. An den Begrenzungen existiert kein weiterer Knoten, der Strom kann also nur parallel zur Wand weiterfließen.

Der Nachteil dieser Methode ist, dass sich das Gleichungssystem nicht auf klassische Weise von einem Mikrocontroller lösen lässt, da er nicht über die entsprechenden Ressourcen verfügt. Im Falle eines Senders, der seine Position nicht verändert, wäre es möglich, die entsprechende Matrix auf einem PC zu berechnen und den Mikrocontroller auf die fertig berechnete Matrix zugreifen zu lassen.

Modell 2: Gauß'sches Modell, ein Sendedipol

Die einfachste Modellierungsvariante ist die direkte Anwendung der Formeln aus dem theoretischen Abschnitt 4.6. Der Sendedipol wird mit Hilfe zweier Punktladungen modelliert, von denen sich jeweils eine an einer der Elektrodenpositionen befindet.

Der herausragende Vorteil dieser Modellierung ist der geringe Rechenaufwand, da sich die elektrische Feldstärke an einem Punkt lediglich als Summe der elektrischen Felder beider Punktladungen an diesem Punkt berechnen lässt.

Der Nachteil dieser Modellierungsvariante ist, dass sie nicht die Einflüsse, die sich aus der Begrenzung des Wasserquaders ergeben, modelliert.

Gauß'sches Modell, Sendedipol mit Spiegelungen

Aufbauend auf der Modellierung des Sendedipols als zwei einzelne Punktladungen ist es möglich, die Begrenzungen des Wasserquaders zu modellieren. Hierzu werden zusätzliche Ladungen in das Modell eingeführt, sogenannte Bild- oder Spiegelladungen. Der Name deutet die zugrundeliegende Idee an: Wenn eine Lichtquelle vor einen Spiegel gestellt wird, wird eine bestimmte Menge des Lichtes am Spiegel reflektiert, im Spiegel ist eine gespiegelte Lichtquelle zu sehen. Um im Bereich vor dem Spiegel die gleichen Lichtverhältnisse ohne den Spiegel zu erhalten, ist es möglich, eine weitere Lichtquelle an der Position der gespiegelten Lichtquelle aufzustellen und den Spiegel zu entfernen. Für elektrische Ladungen lässt sich das gleiche Prinzip anwenden [16, S. 44–55] [6, S. 68–70]:

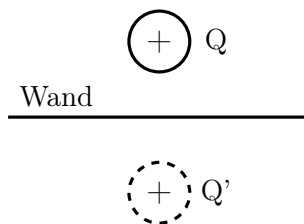


Abbildung 7.11.: Ladung vor isolierender Wand mit zugehöriger Bildladung

Um die Wirkung einer isolierenden Wand auf das von einer Punktladung Q in der Entfernung d erzeugte elektrische Feld zu modellieren, wird eine Ladung Q' mit derselben Polarität „hinter“ der Wand im gleichen Abstand in das Modell aufgenommen.

Der Vorteil dieser Methode ist die bessere Modellierung der Begrenzungen des Wasserquaders, der Nachteil ist der erhöhte Rechenaufwand mit jedem weiteren hinzugefügten Dipol. Da die Begrenzungen des Wasserquaders parallel zueinander stehen, wäre es theoretisch auch nötig, eine unendliche Anzahl von Spiegelbildern zu berechnen, da sich die Spiegelbilder wieder spiegeln.

Im zweidimensionalen Fall entstehen an einer rechtwinkligen Ecke, an der zwei Wände aneinandertreffen, insgesamt drei Spiegelungen. Im dreidimensionalen Fall entstehen, wenn man nur die Spiegelungen in der „Nähe“ des echten Dipols betrachtet, in der Ecke des Beckens, in der sich der Dipol befindet, elf Spiegelungen. Drei auf der Höhe des Dipols und jeweils vier über der Wasseroberfläche und unterhalb des Beckenbodens. Bei vier Ecken würde diese Herangehensweise also die Berechnung von mindestens 48 Dipolen benötigen, je nachdem, welche Spiegelladungen miteinbezogen werden. Da der Berechnungsaufwand linear mit der Anzahl der Dipole steigt, werden im Folgenden zwei Fälle betrachtet:

Modell 3: 4 Dipole in der Senderecke auf der Höhe des Sendedipols: Alle gespiegelten Dipole, die sich in der Ecke und auf Höhe des Senders befinden. Hiermit wird die Verdichtung des elektrischen Feldes durch die Nähe zur Ecke modelliert.

Modell 4: 8 Dipole auf der Höhe des Sendedipols: Die gespiegelten Dipole in der Ecke des Sendedipols und der nächstgelegenen Ecke in Verlängerung der X-Achse, die sich auf derselben Höhe wie der Sendedipol befinden. Hiermit wird die Verdichtung der Feldlinien an der nächsten gegenüberliegenden Wand modelliert.

7.3.3. Überblick über Verarbeitung und Auswertung der Messdaten

Aus jedem Messdurchgang an einer Sensorposition resultiert ein Messdatensatz, der 30720 Spannungsdatensätze (siehe Tabelle 6.5, S. 51) enthält. Für jeden Messdatensatz werden folgende Verarbeitungsschritte durchgeführt.

1. Überprüfung der Zeitstempel

Anhand der Zeitstempel wird überprüft, ob bei der Datenübertragung einzelne Spannungsdatensätze oder Teile von Spannungsdatensätzen nicht übertragen wurden.

Da bei keinem Messdatensatz ein Datenverlust auftrat, wird im Folgenden nicht mehr auf diesen Punkt eingegangen.

2. „Sichtprüfung“ des Messdatensatzes

Die Spannungsverläufe eines Messdatensatzes werden als Diagramm dargestellt und manuell gesichtet, um die Qualität des Spannungsdatensatzes beurteilen zu können. Zum Beispiel, ob alle Sensoren plausibel erscheinende Spannungen erfasst haben, ob der Verstärker übersteuert wurde, welchen Einfluss das Hintergrundrauschen auf die Messdaten hat, etc.

3. Extrahieren der $v_{1,\text{diff}}$ -Werte

Für jeden Messdatensatz werden die $v_{1,\text{diff}}$ -Werte, also die Differenz zwischen dem vorangegangenen Messwert eines Minimums und dem Minimum selbst eines vom Sender empfangenen Nutzsignales, berechnet. Siehe hierzu auch S. 59. Der Durchschnitt über alle $v_{1,\text{diff}}$ -Werte eines Messdatensatzes wird gebildet.

4. Vergleich der Simulationsdaten mit den Messdaten

Da, wie bereits erwähnt, von den Messdaten nicht sicher auf die tatsächlich an den Messelektroden anliegenden Spannungen geschlossen werden kann, sondern lediglich auf die Verhältnisse der Spannungen untereinander, werden sowohl die $v_{1,\text{diff}}$ -Werte, als auch die aus den Simulationen gewonnen Spannungsverteilungen normalisiert. Das bedeutet, dass jeder Spannungsdatensatz, bestehend aus sechs Spannungen, so skaliert wird, dass die maximale Spannung 1 V beträgt; die übrigen Spannungen eines Datensatzes werden mit dem selben Faktor skaliert.

5. Anwendung des Lokalisierungsverfahrens auf die Messdaten

Im letzten Schritt wird das vorgestellte Lokalisierungsverfahren auf die erfassten Messdaten angewendet. Hierbei wird die Abweichung zwischen realer Messsondenposition und der durch das Verfahren berechneten Messsondenposition betrachtet.

Um beurteilen zu können, wie gut die Optimierung nach dem Nelder-Mead-Ansatz funktioniert, werden mehrere Startpunkte verwendet und auch ein brute-force-Gitter, dessen einzelne Punkte einen Abstand von 0,05 m aufweisen, der Aufenthaltswahrscheinlichkeit der Messsonde berechnet.

Nr.	bestes Modell	min. Entf.	Ø Entf.	max. Entf.	b-f-Entf.	Funkt.aufr. N-M
2	WS-Netzwerk	0,03 m	0,34 m	0,64 m	0,32 m	152
3	WS-Netzwerk	0,10 m	0,26 m	0,42 m	0,49 m	140
4	WS-Netzwerk	0,23 m	0,56 m	1,34 m	0,56 m	154
5	Gauß 1 Dipol	0,09 m	0,31 m	0,82 m	0,09 m	226
6	Gauß 4 Dipole	0,08 m	0,35 m	1,58 m	0,21 m	236
7	WS-Netzwerk	0,19 m	0,31 m	0,53 m	0,20 m	136
8	Gauß 4 Dipole	0,04 m	0,52 m	0,80 m	0,56 m	262
9	WS-Netzwerk	0,21 m	0,50 m	0,81 m	0,35 m	157
10	WS-Netzwerk	0,17 m	0,49 m	0,88 m	0,81 m	124
11	WS-Netzwerk	0,14 m	0,84 m	1,19 m	1,41 m	128
12	WS-Netzwerk	0,29 m	0,83 m	1,15 m	1,44 m	117
13	Gauß 1 Dipol	0,18 m	0,18 m	0,18 m	0,29 m	186
14	Gauß 8 Dipole	0,29 m	1,00 m	1,28 m	1,25 m	212
15	WS-Netzwerk	0,04 m	0,76 m	1,16 m	0,55 m	163

Tabelle 7.2.: Bestes Modell pro Messdatensatz nach geringster berechneter Entfernung zum Sender.

b-f: brute-force-Verfahren nach bestem Modell,
 N-M: Nelder-Mead-Verfahren

7.3.4. Auswertung des Versuchs

Ein Schaubild, sowie eine Tabelle der Messsondenpositionen ist auf S. 55 dargestellt.

In der Abbildung 7.12 ist die Standardabweichung (siehe Formel 7.1) der simulierten und normierten Spannungsverteilungen von den real gemessenen und normierten Spannungsverteilungen über alle Sensoren, jeweils für jedes Modell und jeden Datensatz abgebildet. Sie dient als Indiz dafür, wie sehr die simulierten Spannungsverteilungen von den realen Messwerten abweichen, je nachdem, an welcher Position im Becken gemessen wurde.

$$\sigma = \sqrt{\frac{1}{n} \sum_{i=1}^n (v_{\text{diff } i} - \overline{v_{\text{diff}}})^2} \quad (7.1)$$

Formel 7.1: Standardabweichung, $\overline{v_{\text{diff}}}$ bezeichnet den Durchschnitt der Werte.

In den beiden Tabellen 7.2 und 7.3 sind die wichtigsten Ergebnisse des Versuchs zusammengefasst. Die jeweils erste Spalte jeder Tabelle enthält die Nummer des Messdatensatzes. Die zweite Spalte enthält das für diesen Datensatz „beste Modell“, d.h. das Modell, das die beste Lokalisierungsgenauigkeit aufweist.

Die beiden Tabellen unterscheiden sich in der Definition dieses Kriteriums. Bei der Tabelle 7.2 wurde von allen Startpunkten des Nelder-Mead-Verfahrens derjenige ausge-

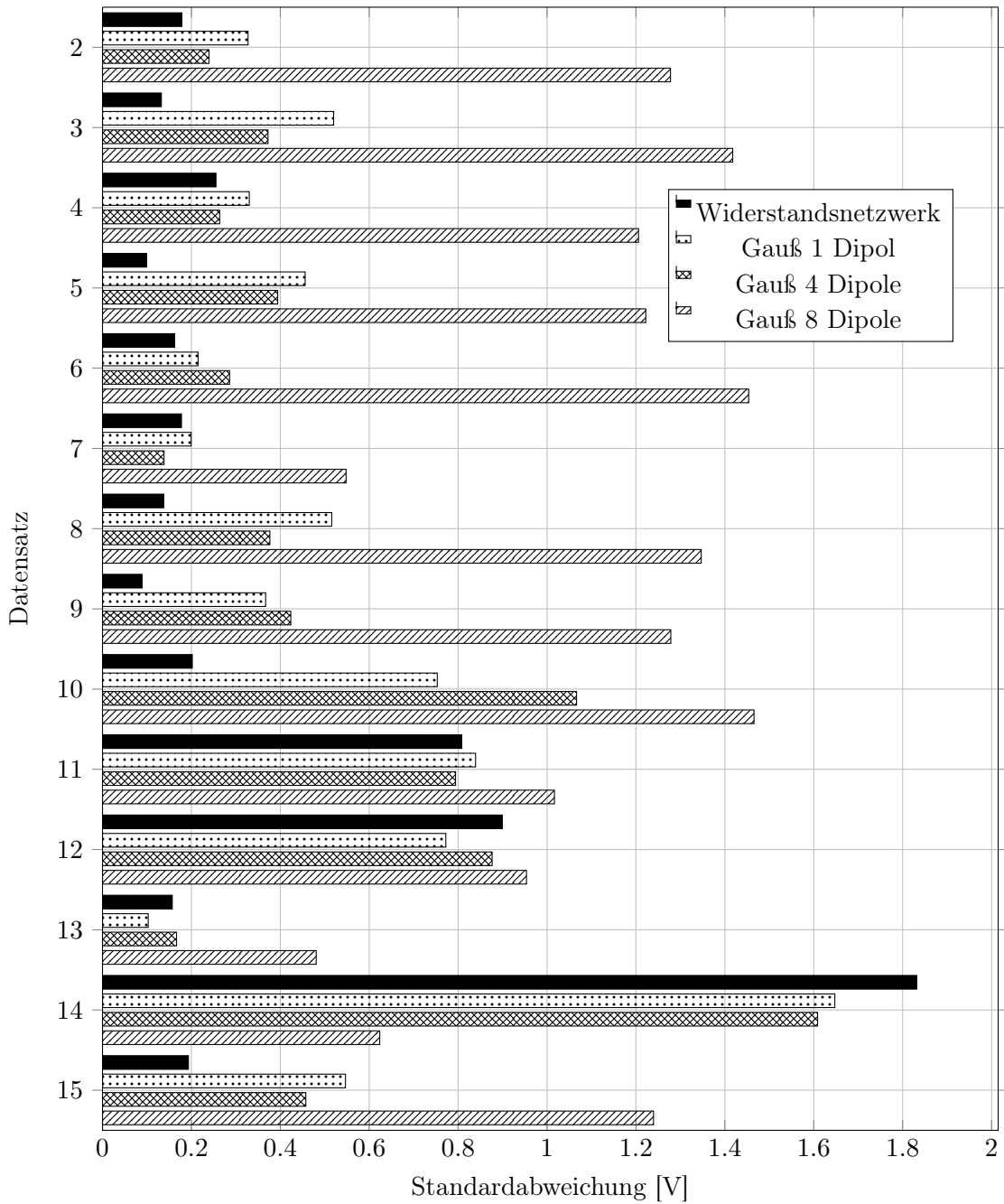


Abbildung 7.12.: Die Standardabweichung der simulierten und normierten Spannungsverteilungen von den real gemessenen und normierten Spannungsverteilungen über alle Sensoren, jeweils für jedes Modell und jeden Datensatz.

7.3. Anwendung des Lokalisierungsverfahrens auf die erfassten Messdaten

Nr.	bestes Modell	min. Entf.	ø Entf.	max. Entf.	b-f-Entf.	Funkt.aufr. N-M
2	WS-Netzwerk	0,03 m	0,34 m	0,64 m	0,32 m	152
3	WS-Netzwerk	0,10 m	0,26 m	0,42 m	0,49 m	140
4	WS-Netzwerk	0,23 m	0,56 m	1,34 m	0,56 m	154
5	Gauß 1 Dipol	0,09 m	0,31 m	0,82 m	0,09 m	226
6	Gauß 1 Dipol	0,18 m	0,22 m	0,46 m	0,21 m	216
7	Gauß 4 Dipole	0,20 m	0,20 m	0,20 m	0,21 m	187
8	WS-Netzwerk	0,06 m	0,52 m	0,99 m	1,05 m	152
9	Gauß 1 Dipol	0,22 m	0,29 m	0,81 m	0,31 m	217
10	WS-Netzwerk	0,17 m	0,49 m	0,88 m	0,81 m	124
11	Gauß 1 Dipol	0,37 m	0,61 m	1,13 m	1,46 m	130
12	Gauß 4 Dipole	0,45 m	0,68 m	0,97 m	1,47 m	152
13	Gauß 1 Dipol	0,18 m	0,18 m	0,18 m	0,29 m	186
14	Gauß 1 Dipol	0,46 m	0,49 m	0,52 m	0,50 m	201
15	Gauß 1 Dipol	0,44 m	0,57 m	0,99 m	0,51 m	208

Tabelle 7.3.: Bestes Modell pro Messdatensatz nach geringster durchschnittlicher Entfernung zum Sender über alle Startpunkte des Nelder-Mead-Verfahrens.

b-f: brute-force-Verfahren nach bestem Modell,
N-M: Nelder-Mead-Verfahren

wählt, der auf die simulierte Messsondenposition mit dem geringsten Abstand zur realen Messsondenposition führt. Dabei wurde nicht einbezogen, wie schlecht die Positionsbestimmung mit der gleichen Modellierung bei anderen Startpunkten war. Bei der Tabelle 7.3 wurde die Modellierung als beste Modellierung bewertet, bei der über alle berechneten Messsondenpositionen aller Startpunkte des Nelder-Mead-Verfahrens der Abstand im Mittel am geringsten ist.

In den weiteren Spalten werden für die jeweils als beste aufgeführte Modellierung folgende Daten aufgeführt:

Die minimale Entfernung zur realen Messsondenposition, die über alle Startpunkte des Nelder-Mead-Verfahrens für diese Modellierung bestimmt wurde. Sie ist ein Maß dafür, wie gut die Positionsbestimmung mit diesem Modell bei einer idealen Startpunktwahl im besten Fall sein kann.

Die durchschnittliche Entfernung zur realen Messsondenposition, die über alle Startpunkte des Nelder-Mead-Verfahrens für diese Modellierung bestimmt wurde.

Die maximale Entfernung zur realen Messsondenposition, die über alle Startpunkte des Nelder-Mead-Verfahrens für diese Modellierung bestimmt wurde. Zusammen mit der durchschnittlichen Entfernung ist es möglich, abzuschätzen, wie weit das

Verfahren mit der jeweiligen Modellierung bei der Positionsbestimmung streut oder im schlechtesten Fall sein kann.

Die minimale per brute-force Verfahren bestimmte Entfernung, wobei die gleiche Modellierung zum Einsatz kam. Hierbei wurde versucht, die Entfernung per brute-force zu bestimmen, wobei der Punkt mit der höchsten Wahrscheinlichkeit als Messsondenposition ausgewählt wurde.

Die Anzahl der Funktionsaufrufe, die während des gesamten Ablaufs des Nelder-Mead-Verfahrens zur Bestimmung der minimalen Entfernung notwendig waren. Ein Funktionsaufruf im Nelder-Mead-Verfahren steht hierbei für die Bestimmung der Aufenthaltswahrscheinlichkeit der Messsonde an einem Punkt. Das bedeutet, dass der Aufwand der Verfahren nicht direkt vergleichbar ist. Während bei der Gauß-Modellierung mit einem Dipol die elektrischen Feldverteilungen von zwei Einzelladungen berechnet werden müssen, sind es bei der Verwendung von acht Dipolen 16 Feldverteilungen, die berechnet werden müssen. Dagegen ist der Rechenaufwand bei der Verwendung des Widerstandsnetzwerkes gering, da das elektrische Potential lediglich aus einer Matrix ausgelesen werden muss.

Aus der Tabelle 7.2 lässt sich ablesen, dass das Widerstandsnetzwerk in neun von 14 Fällen eine Messsondenposition mit der geringsten Abweichung zur realen Position bestimmt hat. Desweiteren zeigt sich, dass das Nelder-Mead-Verfahren bei der Verwendung des Widerstandsnetzwerkes schneller konvergiert als bei der Modellierung nach dem Gauß-Modell. Auch scheint die Güte der Modellierung mit Hilfe des Widerstandsnetzwerkes nicht abhängig von der Entfernung der Messsonde zum Sendedipol zu sein. Aufgrund der geringen Anzahl von Fällen, in denen andere Verfahren am besten die Position bestimmen haben, lässt sich kaum eine Aussage treffen, welches Verfahren unter welchen Umständen geeigneter ist. Das brute-force-Verfahren ist in den meisten Fällen bei der Positionsbestimmung nicht besser als der Durchschnitt der jeweiligen Nelder-Mead-Durchläufe bei der Verwendung des gleichen Modells.

Bei der Verwendung des geringsten durchschnittlichen Fehlers bei der Positionsbestimmung als Kriterium zur Auswahl der besten Modellierung, siehe Tabelle 7.3, bestimmt das Modell nach Gauß unter Verwendung eines Dipols in den meisten Fällen die Position mit der geringsten Abweichung. In der Nähe des Senders und der Wand scheint die Positionsbestimmung unter Verwendung des Widerstandsnetzwerkes im Mittel allerdings genauer zu sein. Auffällig ist, dass die Verwendung des Gauß-Modelles auch hier mehr Punkte bei der Verwendung des Nelder-Mead-Verfahrens berechnen muss, als es bei der Verwendung des Widerstandsnetzwerkes nötig ist.

Diagramme und Daten zu den berechneten Senderpositionen für jeden Messdatensatz befinden sich im Anhang B.

7.4. Bewertung des Verfahrens anhand der Versuchsergebnisse

Die Genauigkeit der Positionsbestimmung variiert stark mit der Güte der jeweiligen Modellierung. Das Modell, das aufgrund der Stichproben die im Durchschnitt beste Lokalisierungsleistung zu ermöglichen scheint, ist das Widerstandsnetzwerk.

Wenn die Umgebung der tatsächlichen Position der Messsonde hinreichend gut durch das Modell beschrieben wird, ist zumindest eine ungefähre Positionsbestimmung aufgrund der gemessenen relativen Spannungsverhältnisse möglich. Hierbei kann sowohl der brute-force-Ansatz, wie auch das Nelder-Mead-Verfahren zum Einsatz kommen. Bei der Verwendung des Nelder-Mead-Verfahrens zur Positionsbestimmung erscheint es sinnvoll, mehrere über das Suchgebiet verteilte Startpunkte zu verwenden, da das Verfahren dazu neigt, auf ein lokales Minimum hin zu konvergieren. Da das Nelder-Mead-Verfahren relativ schnell konvergiert – im Versuch in den meisten Fällen mit deutlich unter 300 Funktionsaufrufen zur Wahrscheinlichkeitsberechnung – ist der Rechenaufwand auch bei mehreren Startpunkten geringer als es die Verwendung des brute-force-Verfahrens unter Nutzung des gleichen Modellierungsverfahrens wäre. Das bei diesem Vergleich genutzte brute-force-Verfahren berechnet die Wahrscheinlichkeit an 8325 Punkten im Abstand von 0,05 m, ist also deutlich aufwändiger.

Zur Auswahl der wahrscheinlichsten Messsondenposition aus der eventuell aufgrund der beim Nelder-Mead-Verfahren verwendeten unterschiedlichen Startpunkte verschiedenen Messsondenpositionen, kann die berechnete Aufenthaltswahrscheinlichkeit der Messsonde an den verschiedenen Positionen dienen.

8. Abschließende Bewertung und Ausblick

8.1. Zusammenfassung

In dieser Arbeit wurde gezeigt, dass auf Grundlage der Formulierung der Aufgabe der Elektrolokation als inverses Problem prinzipiell eine Lösung möglich ist. Die zugrundeliegende Theorie lässt durch die Verwendung von Wahrscheinlichkeitsdichtefunktionen dabei große Freiheiten bei der Modellierung des physikalischen Modells, sowie der Modellparameter und der Eigenschaften der verwendeten Messhardware. Auch können weitere Informationen, die nicht nur durch die Erfassung von elektrischen Feldern gewonnen werden, in die Positionsbestimmung mit einfließen, ohne dass Änderungen an der Modellierung der elektrischen Feldausbreitung nötig sind. Somit ist das Verfahren stark modular.

Das Downhill-Verfahren nach Nelder und Mead kann verwendet werden, um den Rechenaufwand gegenüber der Verwendung eines brute-force-Ansatzes, bei dem alle Positionen innerhalb eines Gitters berechnet werden, stark zu reduzieren. Hierbei ist der Einsatz von mehreren Startpunkten für das Nelder-Mead-Verfahren sinnvoll, da die berechnete Wahrscheinlichkeit nicht als einziges Kriterium ausreicht, um beurteilen zu können, wie groß der Fehler der berechneten Position ist. Bei der Verwendung mehrerer Startpunkte können sowohl die berechneten Positionen, als auch die berechneten Wahrscheinlichkeiten dieser Positionen miteinander verglichen werden, um eventuelle Fehler zu bestimmen.

Deutliches Verbesserungspotential gibt es bei der physikalischen Modellierung. Da die Lokalisierung auf der Übereinstimmung der Realität mit der Simulation basiert und die verschiedenen Modelle Unterschiede in den simulierten elektrischen Potentialen sowohl untereinander, als auch im Vergleich mit den realen Messwerten aufweisen, ist anzunehmen, dass sich die Lokalisierungsgenauigkeit bei einer realistischeren Modellierung ebenfalls verbessert.

Bei der Betrachtung der verschiedenen Modelle wurde die höchste Lokalisierungsgenauigkeit mit dem Widerstandsnetzwerk und der Verwendung eines Dipols mit der elektrostatischen Modellierung nach Gauß erzielt. Dies muss nicht bedeuten, dass die Verwendung von Spiegelladungen ein falscher Ansatz ist, da nur ein kleiner Teil der denkbaren Kombinationen von gespiegelten Dipolen betrachtet wurde.

Die Verwendung des Widerstandsnetzwerkes zur Modellierung bietet die Vorteile der genauesten Lokalisierung im besten Fall und des geringen Rechenaufwandes, wenn die verwendete Matrix bereits berechnet wurde und lediglich darauf zugegriffen und Werte, die zwischen den Punkten des Gitters liegen, interpoliert werden müssen. Die Berechnung der

Matrix über ein lineares Gleichungssystem ist so aufwändig, dass sie nur schwer während des Einsatzes von einem Mikrocontroller durchgeführt werden kann. Entweder muss eine Möglichkeit gefunden werden, die Matrix zumindest näherungsweise auf andere Weise zu berechnen, oder sie muss im Mikrocontroller abgespeichert werden. Falls die Matrix im voraus berechnet wird, folgen daraus gewisse Einschränkungen: Entweder, der Sender muss sich an einer festen Position befinden, da sich bei einer Positionsänderung die elektrische Potentialverteilung im Wasser ändert. Oder, der Roboter, der die hinterlegte Matrix verwendet, muss sich in einer Umgebung befinden, bei der eine Positionsänderung keine wesentliche Änderung der Potentialverteilung ergibt. Das bedeutet, dass sich sämtliche Begrenzungen des Wassers in ausreichender Entfernung befinden müssen.

Die Verwendung der Modellierung nach Gauß mit nur einem Dipol unter Kombination mit dem Nelder–Mead–Verfahren ist vermutlich die beste Wahl zur Implementierung auf einem Mikrocontroller, da der Rechenaufwand im Vergleich zu den anderen Modellen, mit Ausnahme des Widerstandsnetzwerkes, im Einsatz geringer ist und die Lokalisierungs-genauigkeit im Mittel die beste war.

Da auf der Roboterplattform vermutlich ein Drucksensor zur Tauchtiefenbestimmung verwendet werden kann, ist anzunehmen, dass sich die Lokalisierungs-genauigkeit durch Verwendung dieser Informationen noch verbessert. Desweiteren wird es möglich und sinnvoll sein, Informationen aus vergangenen Lokalisierungsberechnungen, wie z.B. die letzte bekannte Position, zu verwenden.

8.2. Mögliche Verbesserungen

8.2.1. Physikalisches Modell

Bei den reinen Simulationen, bei denen auch die Messdaten, anhand derer die Lokalisierung durchgeführt wird, mit Hilfe der Simulation erzeugt werden, ist das Lokalisierungsverfahren genauer als bei der Verwendung von realen Messdaten. Da die verschiedenen physikalischen Modelle außerdem verschiedene Potentialverteilungen im Raum vorhersagen, ist anzunehmen, dass die Verbesserung des physikalischen Modells die Lokalisierungs-genauigkeit des Verfahrens ebenfalls verbessert.

8.2.2. Messhardware

Wie bereits angesprochen wäre es gut, den Spannungsbereich, in dem gemessen werden kann veränderbar zu gestalten, damit der Verstärker nicht übersteuert wird oder auch schwächere Spannungen noch gemessen werden können.

Da das Verfahren auf der Verwendung von Wahrscheinlichkeitsdichten aufbaut, ist es möglich zur Positionsbestimmung auch Informationen anderer Sensoren, wie z.B. die eines Drucksensors oder eines Kompasses, zu verwenden, ohne dass es nötig ist, die grundlegende Logik des Verfahrens zu ändern. Die Informationen eines Drucksensors z.B. würden bei der

Berechnung lediglich die Wahrscheinlichkeit, dass sich der Roboter in einer bestimmten Tiefe befindet, beeinflussen.

8.2.3. Einbezug der Zeit

Da sich Objekte unter Wasser nur mit einer begrenzten Geschwindigkeit bewegen, wäre es sinnvoll, diesen Aspekt bei der Modellierung zu nutzen und bei der Lokalisierung z.B. eine „Aufenthaltswahrscheinlichkeitskugel“ um eine ältere bekannte Position zu legen, bei der die Wahrscheinlichkeit zur Kugeloberfläche hin abnimmt.

8.2.4. Nutzung mehrerer Roboter

Bei der Nutzung mehrerer Roboter, die untereinander Informationen austauschen können, sind mehrere Verbesserungen denkbar.

Eine Möglichkeit wäre, dass ein Roboter andere Roboter bei seiner Lokalisierung unterstützt, indem er einen Code sendet, der ihn eindeutig indentifiziert. Desweiteren könnte er Informationen, die er über seine eigene Position bezüglich eines gemeinsamen Koordinatensystems hat, entsprechend kodieren und ebenfalls senden. Eine Dimension, die alle Roboter miteinander teilen, könnte die Tiefe unter der Wasseroberfläche sein, die sie mit einem Drucksensor registrieren. Diese Informationen könnten dann z.B. in die Wahrscheinlichkeitsdichtefunktion $\rho_M(\mathbf{m})$, die die a-priori-Information über die Modellparameter darstellt, einfließen.

Eine weitere Möglichkeit wäre das Pooling von Informationen. Roboter könnten eine Liste der ihnen bekannten Objekte und der Position, die sie zuletzt für sie berechnet haben, sowie einer Einschätzung, wie genau die Lokalisierung vermutlich war, broadcasten. Außerdem könnten Roboter, die ihre relativen Positionen zueinander kennen, synchron messen und somit die Menge der verfügbaren Messdaten vergrößern. Dies würde allerdings die Komplexität des Systems deutlich erhöhen, sobald Roboter an den Messungen teilnehmen, die nicht mit allen anderen Robotern direkt kommunizieren können, was die Einführung eines Verfahrens zur Uhrensynchronisation, Routingverfahren zur Kommunikation, etc. nötig machen würde.

Außerdem wäre es möglich, Berechnungen vom messenden Roboter an einen anderen Roboter zu übertragen und diesen die Berechnungen durchführen zu lassen, da die zu übertragende Menge an Messdaten bzw. berechneter Positionsdaten relativ gering ist. Desweiteren ist es denkbar, bestimmte Berechnungen parallel auszuführen.

8.2.5. Parallelisierbarkeit zur Beschleunigung

Eine Möglichkeit zur Beschleunigung der Lokalisierung ist die Ausnutzung der Parallelisierbarkeit von Berechnungen. Da eine Roboterplattform oft mehrere Prozessoren zur Verfügung hat, die Corezahl im Augenblick auch bei eingebetteten Systemen steigt und bei einer Gruppe kooperierender Roboter theoretisch ebenfalls mehrere Prozessoren be-

nutzt werden können, wird im Folgenden angenommen, dass mehrere Threads gleichzeitig ausgeführt werden können.

Da das behandelte Verfahren auf einem Downhill-Algorithmus basiert, ist eine relativ einfache Möglichkeit der Parallelisierung, das Verfahren für mehrere verschiedene Startpunkte parallel durchzuführen. Diese Variante hätte auch den Vorteil, dass relativ wenige Daten ausgetauscht werden müssen, also zu Beginn die Messdaten und die möglichen Startpositionen und als Ergebnis der Berechnung eine Position und die Wahrscheinlichkeit dieser Position. Aufgrund der relativ hohen Kommunikationskosten bei verteilten Systemen ist allerdings anzunehmen, dass sich der Aufwand – zumindest zur Lokalisierung eines einzigen Dipols – nicht lohnt.

Falls es möglich ist, Threads auf die gleichen Arrayelemente zugreifen zu lassen, z.B. bei einem Mehrkernsystem, wäre es denkbar, Funktionswerte für mehrere Punkte des Simplex parallel zu berechnen.

A. Simulationsbeispiel

Quelltext: A.1: Beispiel zur Verwendung der definierten Klassen für eine einfache Simulation

```
% erzeuge World- und Driver-Objekte
world = ludwig.World();
driver = world.addDriver();

% erzeuge den Dipol, der lokalisiert werden soll
dipole = driver.addSimObject({}, 'dipole', ludwig.Point((rand(1,3)*2*1)-1));

% füge dem zu lokalisierenden Dipolobjekt eine Sendeelektrode hinzu
el_dip = ludwig.Electrode(dipole, 'el', ludwig.Point([0.0, 0.0, 0.0]));
el_dip.relativePosition1 = ludwig.Point([ 0.0, 0.0, 0.0]);
el_dip.relativePosition2 = ludwig.Point([ 0.0, 0.0, -0.1]);
el_dip.mode = 't';

% erzeuge den Dipol, mit dem probe.analyzer später die Berechnungen
% durchführt
c_dipole = driver.addSimObject({}, 'calculation_dipole',
    ludwig.Point([0.0, 0.0, 0.0]));
el_c_dip = ludwig.Electrode(c_dipole, 'el_c',
    ludwig.Point([0.0, 0.0, 0.0]));
el_c_dip.relativePosition1 = ludwig.Point([ 0.0, 0.0, 0.0]);
el_c_dip.relativePosition2 = ludwig.Point([ 0.0, 0.0, -0.1]);
el_c_dip.mode = 't';
el_c_dip.isActive = 0;

% erzeuge die Messsonde
probe = driver.addSimObject({}, 'probe', ludwig.Point([0,0,0]));

% erzeuge Messelektroden der Messsonde
el_ab = ludwig.Electrode(probe, 'ab', ludwig.Point([0.0, 0.0, 0.1]));
el_ab.relativePosition1 = ludwig.Point([ 0.0, 0.0, 0.0]);
el_ab.relativePosition2 = ludwig.Point([ 0.0, 0.1, -0.1]);
```

A. Simulationsbeispiel

```
% erzeuge den Analyzer der Messsonde
probe.analyzer = ludwig.Analyzer(probe);

% berechne die an allen messenden Elektroden anliegende Spannung
driver.computeSetAllCurrents();

% Normalabweichung, die der Analyzer verwendet
div = 0.000001;

% Startpunkt für den Nelder-Mead-Algorithmus der vom Analyzer
% verwendet wird
nmVal = [1,1,1];

% berechne mit Hilfe des Analyzers eine mögliche Position
% des sendenden Dipols
calcVal = probe.analyzer.nelderMeadSimd(dipole, c_dipole, nmVal, 1,
div);
```

B. Messergebnisse und Simulationsdaten

B.1. Verwendete Startpunkte für Nelder-Mead

Startpunkt Nr.	x	y	z
1	1.0	0.7	0.3
2	0.666	0.466	0.2
3	0.666	0.466	0.4
4	0.666	0.933	0.2
5	0.666	0.933	0.4
6	1.333	0.466	0.2
7	1.333	0.466	0.4
8	1.333	0.933	0.2
9	1.333	0.933	0.4

Tabelle B.1.: Verwendete Startpunkte für Nelder-Mead

B.2. Diagramme und Daten zu den einzelnen Messdatensätzen

B.2.1. Messdatensatz 1

Während der ersten Messung drang Wasser in die Messsonde ein, was zu offensichtlich fehlerhaften Messdaten führte. Da die zweite Messung an derselben Position durchgeführt wurde, wird der erste Messdatensatzes nicht ausgewertet.

B.2.2. Messdatensatz 2

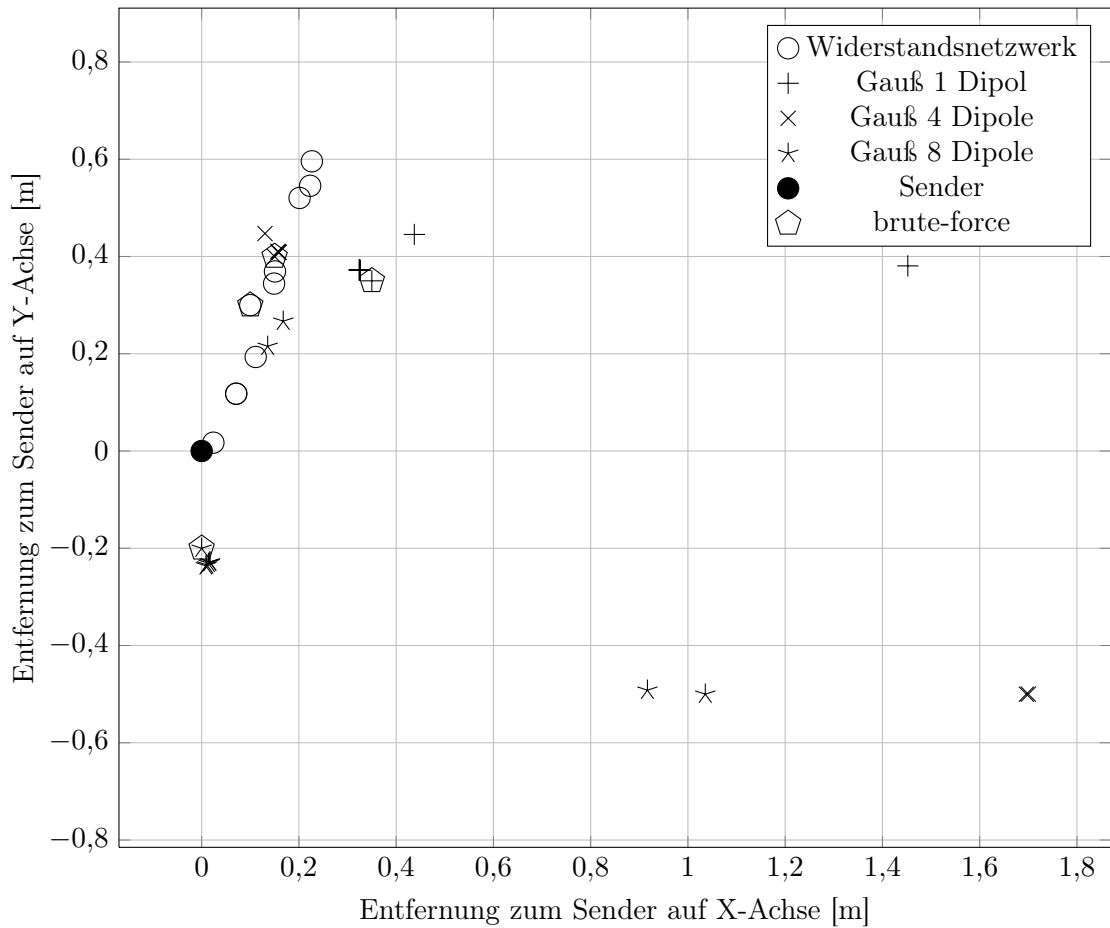


Abbildung B.1.: Messdatensatz 2, berechnete Senderpositionen.

Modell	beste Lokalisierung mit Nelder–Mead				
	Funktionsaufrufe	min. Entf.	ø Entf.	max. Entf.	brute-force-Entf.
WS-Netzwerk	124	0,030	0,343	0,638	0,316
Gauß 1 Dipol	293	0,503	0,629	1,505	0,505
Gauß 4 Dipole	308	0,449	0,747	1,775	0,439
Gauß 8 Dipole	202	0,250	0,484	1,189	0,224

Tabelle B.2.: Messdatensatz Nr. 2: Beste Lokalisierung mit geringster Entfernung zum Sender über alle Startpunkte des Nelder–Mead–Verfahrens pro Modell.

B.2.3. Messdatensatz 3

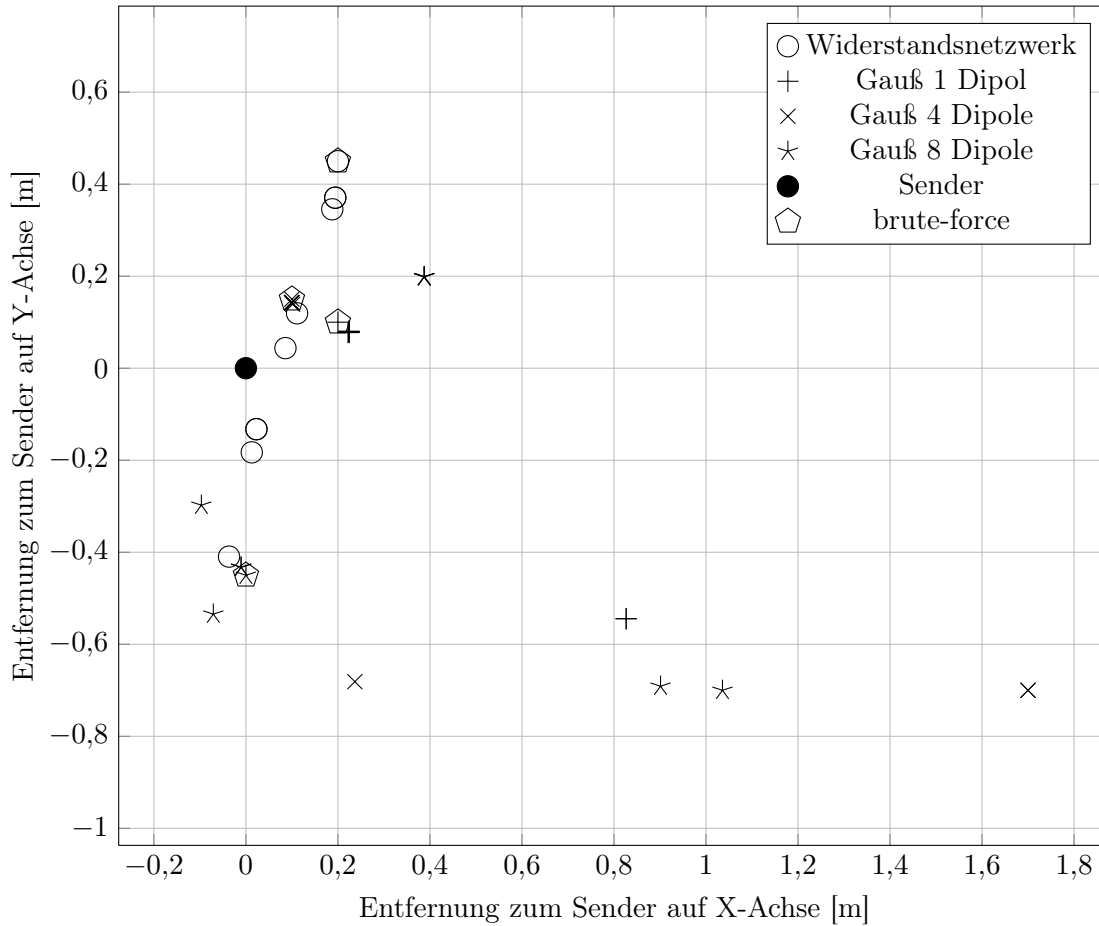


Abbildung B.2.: Messdatensatz 3, berechnete Senderpositionen.

Modell	beste Lokalisierung mit Nelder–Mead				
	Funktionsaufrufe	min. Entf.	ø Entf.	max. Entf.	brute-force-Entf.
WS-Netzwerk	127	0,098	0,263	0,420	0,495
Gauß 1 Dipol	303	0,255	0,421	0,993	0,245
Gauß 4 Dipole	209	0,199	0,624	1,844	0,206
Gauß 8 Dipole	146	0,329	0,645	1,286	0,461

Tabelle B.3.: Messdatensatz Nr. 3: Beste Lokalisierung mit geringster Entfernung zum Sender über alle Startpunkte des Nelder–Mead–Verfahrens pro Modell.

B.2.4. Messdatensatz 4

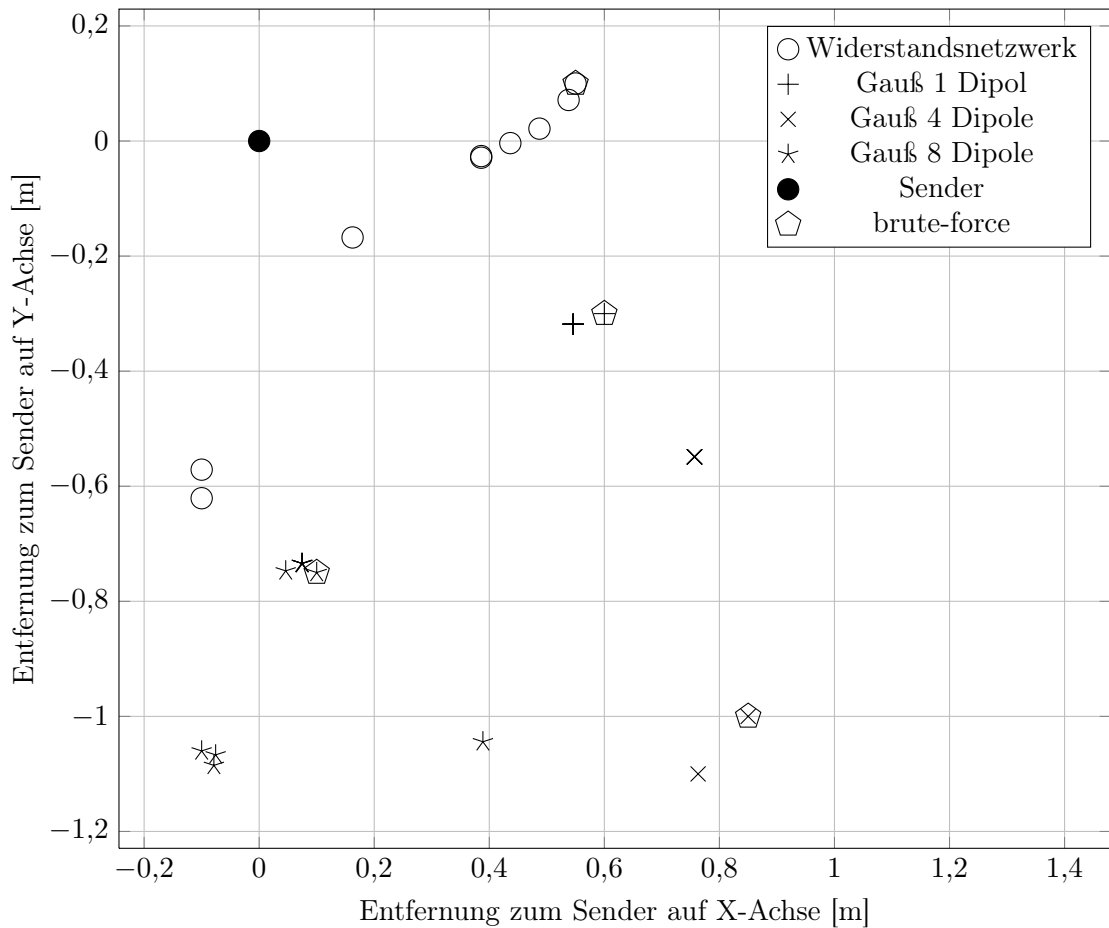


Abbildung B.3.: Messdatensatz 4, berechnete Senderpositionen.

Modell	beste Lokalisierung mit Nelder–Mead				
	Funktionsaufrufe	min. Entf.	ø Entf.	max. Entf.	brute-force-Entf.
WS-Netzwerk	161	0,234	0,559	1,338	0,561
Gauß 1 Dipol	133	0,639	0,639	0,639	0,678
Gauß 4 Dipole	323	0,945	0,989	1,341	1,316
Gauß 8 Dipole	200	0,764	0,911	1,132	0,783

Tabelle B.4.: Messdatensatz Nr. 4: Beste Lokalisierung mit geringster Entfernung zum Sender über alle Startpunkte des Nelder–Mead–Verfahrens pro Modell.

B.2.5. Messdatensatz 5

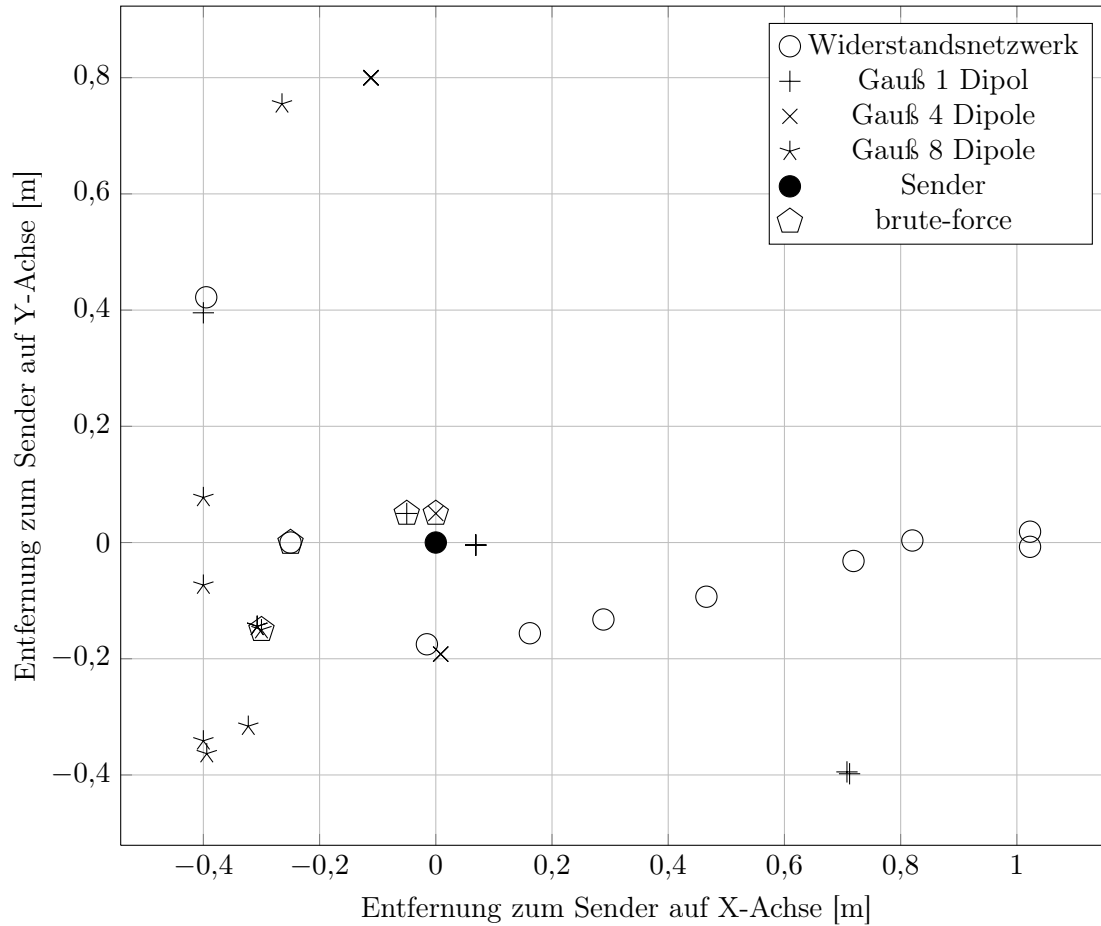


Abbildung B.4.: Messdatensatz 5, berechnete Senderpositionen.

Modell	beste Lokalisierung mit Nelder–Mead				
	Funktionsaufrufe	min. Entf.	ø Entf.	max. Entf.	brute-force-Entf.
WS-Netzwerk	190	0,178	0,596	1,023	0,250
Gauß 1 Dipol	173	0,091	0,306	0,823	0,087
Gauß 4 Dipole	141	0,201	0,606	0,808	0,071
Gauß 8 Dipole	259	0,354	0,482	0,825	0,350

Tabelle B.5.: Messdatensatz Nr. 5: Beste Lokalisierung mit geringster Entfernung zum Sender über alle Startpunkte des Nelder–Mead–Verfahrens pro Modell.

B.2.6. Messdatensatz 6

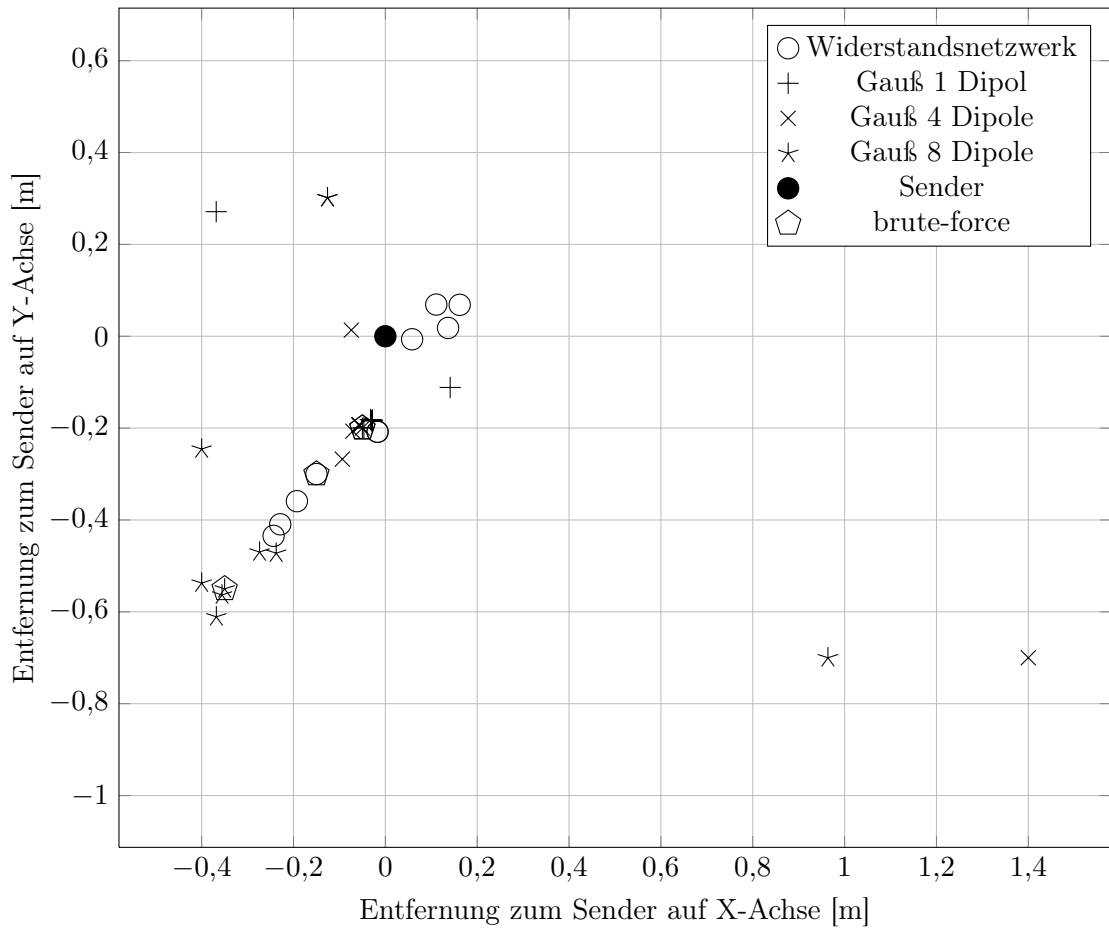


Abbildung B.5.: Messdatensatz 6, berechnete Senderpositionen.

Modell	beste Lokalisierung mit Nelder–Mead				
	Funktionsaufrufe	min. Entf.	ø Entf.	max. Entf.	brute-force-Entf.
WS-Netzwerk	145	0,077	0,259	0,500	0,339
Gauß 1 Dipol	215	0,180	0,217	0,457	0,206
Gauß 4 Dipole	248	0,075	0,352	1,582	0,206
Gauß 8 Dipole	146	0,470	0,650	1,211	0,654

Tabelle B.6.: Messdatensatz Nr. 6: Beste Lokalisierung mit geringster Entfernung zum Sender über alle Startpunkte des Nelder–Mead–Verfahrens pro Modell.

B.2.7. Messdatensatz 7

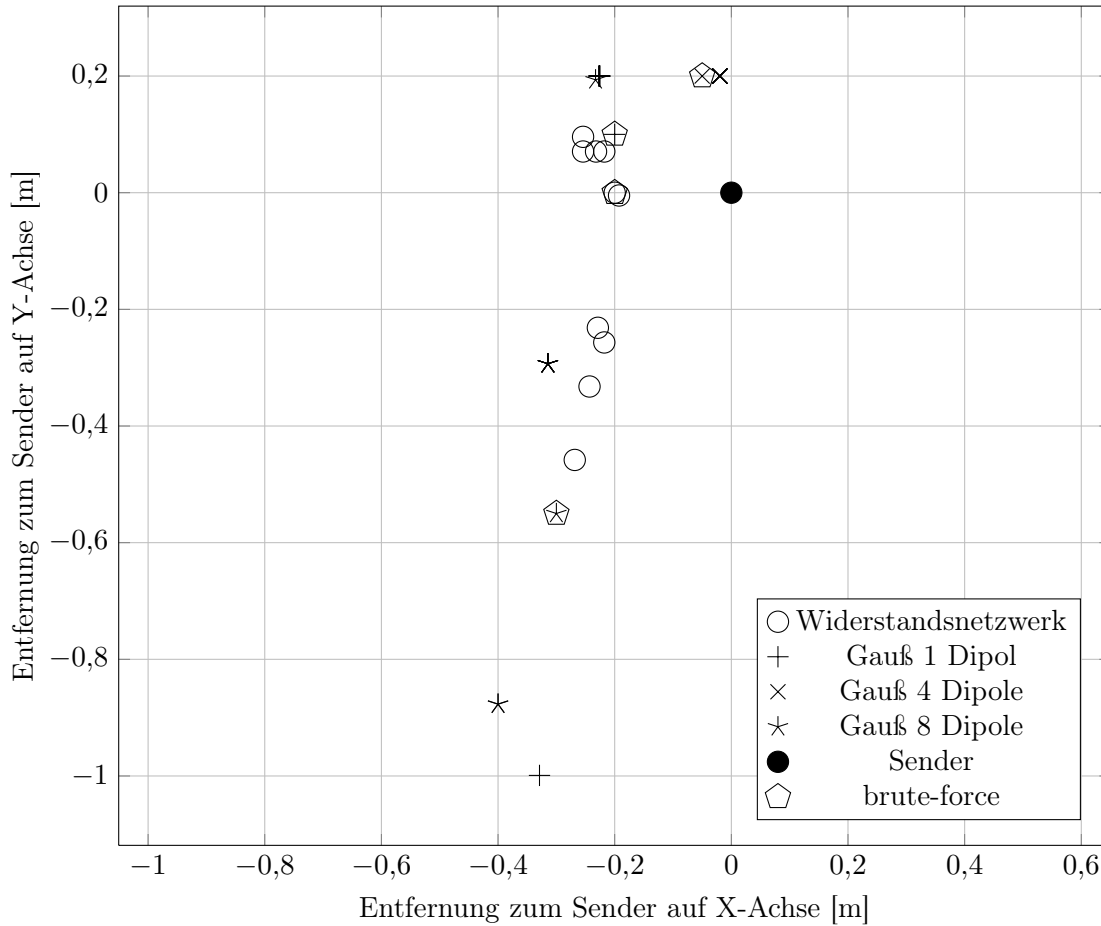


Abbildung B.6.: Messdatensatz 7, berechnete Senderpositionen.

Modell	beste Lokalisierung mit Nelder–Mead				
	Funktionsaufrufe	min. Entf.	ø Entf.	max. Entf.	brute-force-Entf.
WS-Netzwerk	111	0,193	0,312	0,531	0,200
Gauß 1 Dipol	192	0,302	0,388	1,075	0,224
Gauß 4 Dipole	221	0,201	0,201	0,201	0,206
Gauß 8 Dipole	127	0,307	0,589	0,966	0,658

Tabelle B.7.: Messdatensatz Nr. 7: Beste Lokalisierung mit geringster Entfernung zum Sender über alle Startpunkte des Nelder–Mead–Verfahrens pro Modell.

B.2.8. Messdatensatz 8

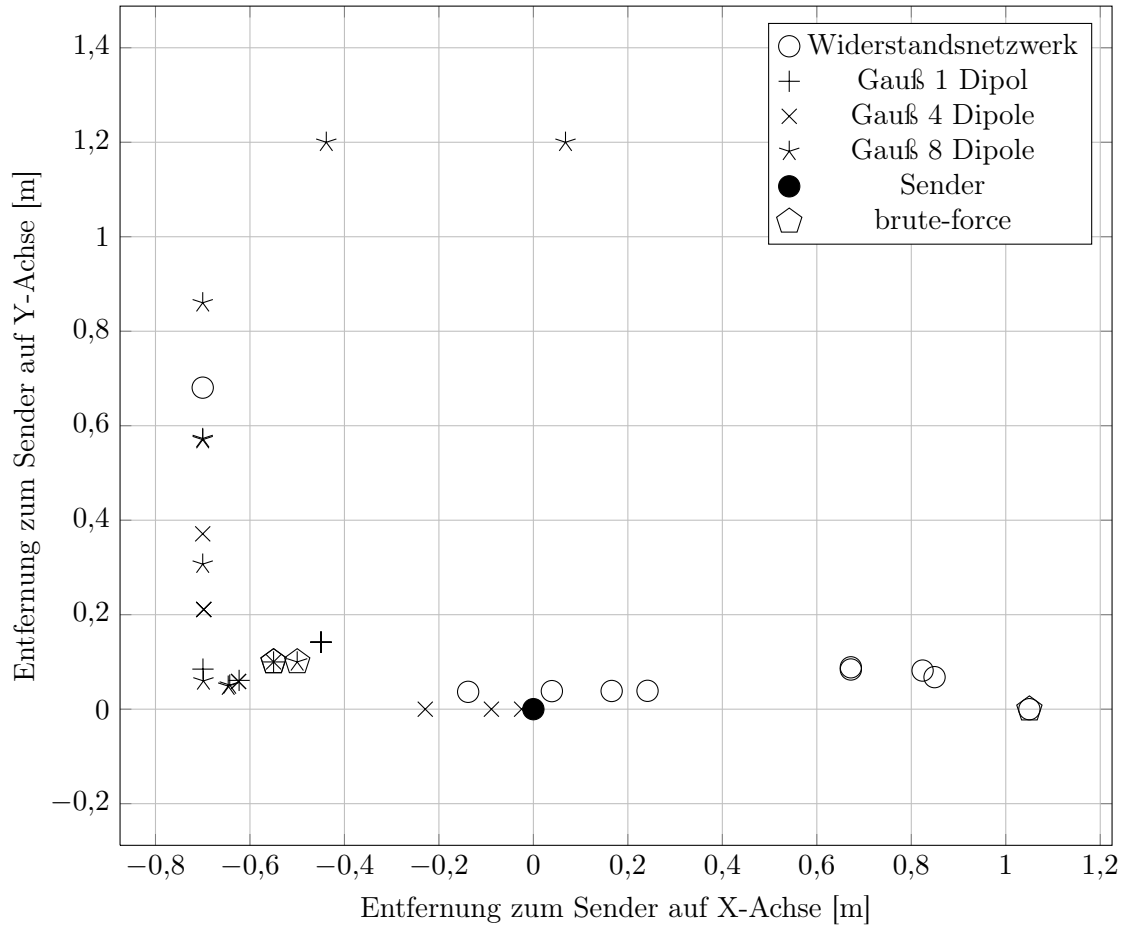


Abbildung B.7.: Messdatensatz 8, berechnete Senderpositionen.

Modell	beste Lokalisierung mit Nelder–Mead				
	Funktionsaufrufe	min. Entf.	ø Entf.	max. Entf.	brute-force-Entf.
WS-Netzwerk	115	0,056	0,516	0,989	1,051
Gauß 1 Dipol	195	0,471	0,516	0,715	0,559
Gauß 4 Dipole	232	0,045	0,516	0,799	0,559
Gauß 8 Dipole	197	0,648	0,911	1,281	0,512

Tabelle B.8.: Messdatensatz Nr. 8: Beste Lokalisierung mit geringster Entfernung zum Sender über alle Startpunkte des Nelder–Mead–Verfahrens pro Modell.

B.2.9. Messdatensatz 9

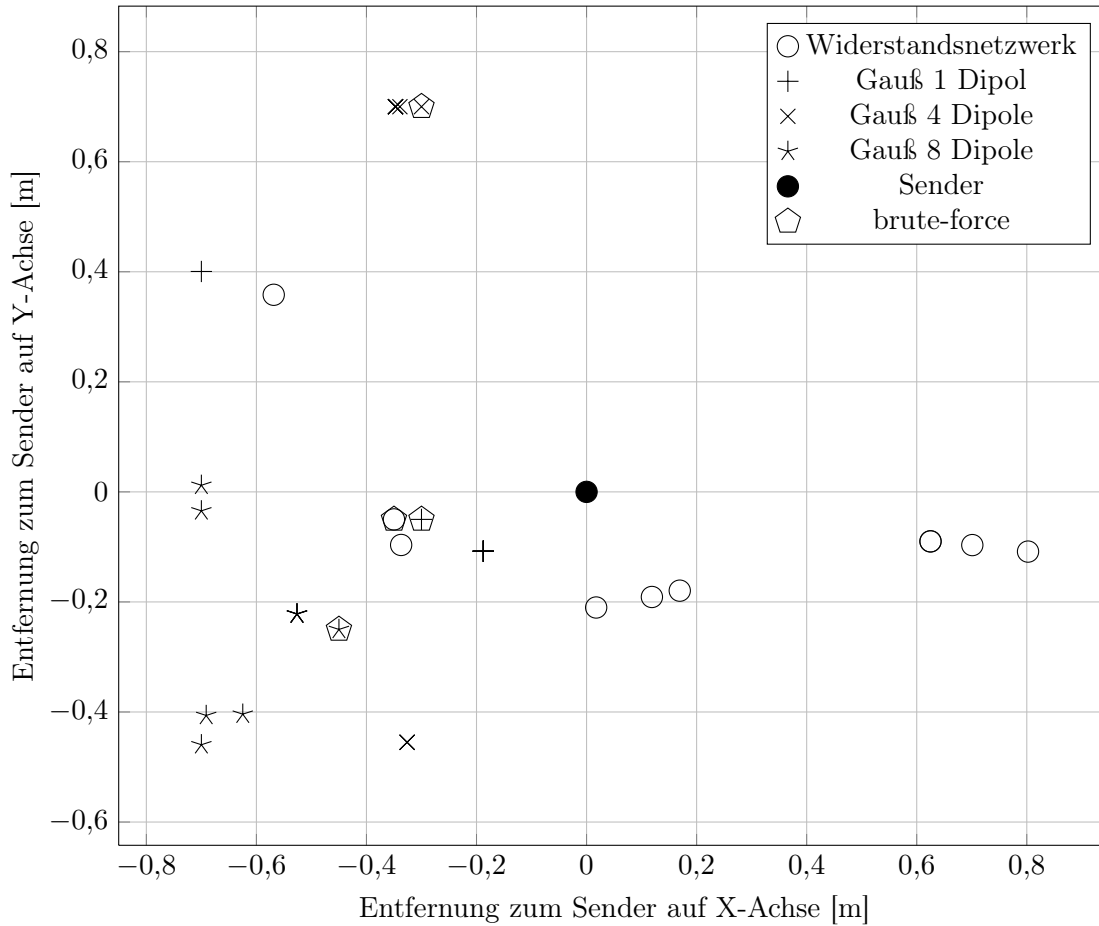


Abbildung B.8.: Messdatensatz 9, berechnete Senderpositionen.

Modell	beste Lokalisierung mit Nelder–Mead				
	Funktionsaufrufe	min. Entf.	ø Entf.	max. Entf.	brute-force-Entf.
WS-Netzwerk	175	0,212	0,499	0,810	0,354
Gauß 1 Dipol	255	0,224	0,289	0,807	0,308
Gauß 4 Dipole	230	0,563	0,709	0,783	0,763
Gauß 8 Dipole	189	0,585	0,690	0,851	0,536

Tabelle B.9.: Messdatensatz Nr. 9: Beste Lokalisierung mit geringster Entfernung zum Sender über alle Startpunkte des Nelder–Mead–Verfahrens pro Modell.

B.2.10. Messdatensatz 10

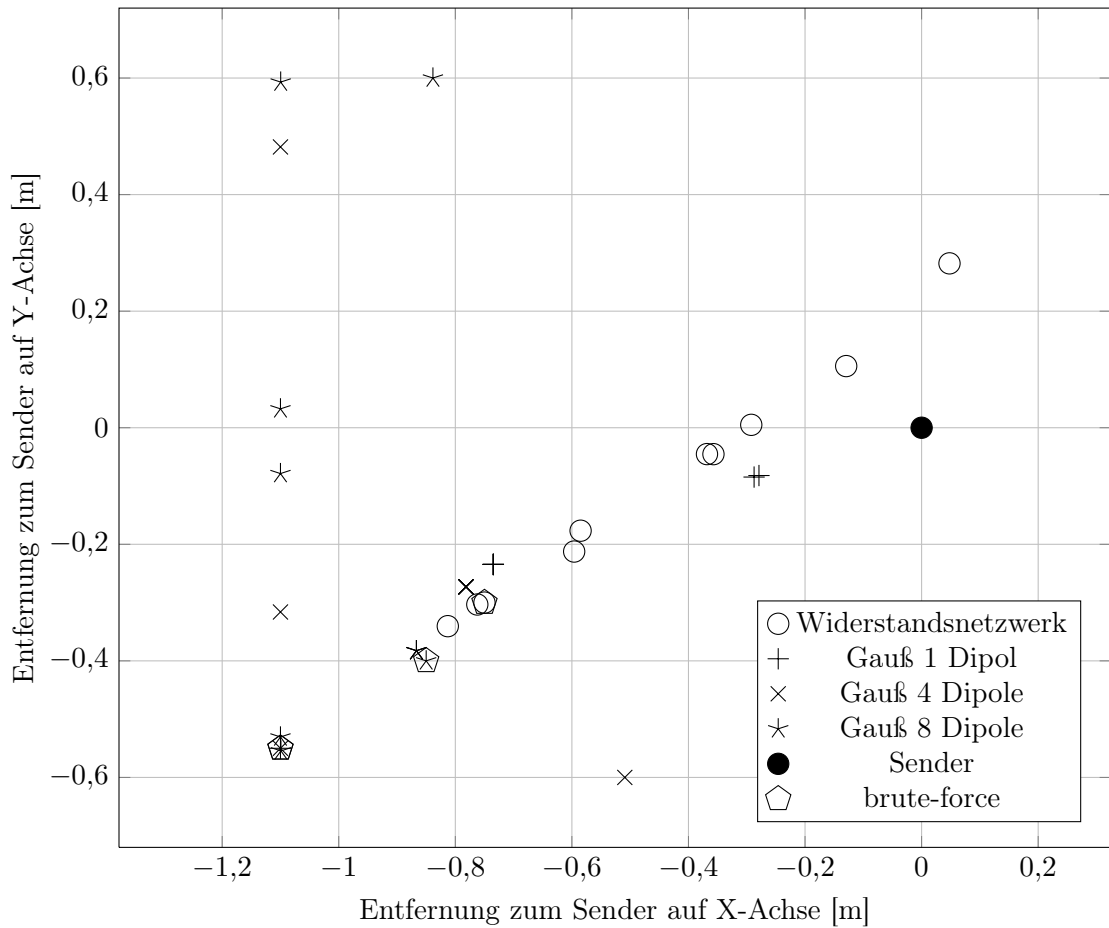


Abbildung B.9.: Messdatensatz 10, berechnete Senderpositionen.

Modell	beste Lokalisierung mit Nelder–Mead				
	Funktionsaufrufe	min. Entf.	ø Entf.	max. Entf.	brute-force-Entf.
WS-Netzwerk	123	0,167	0,491	0,881	0,808
Gauß 1 Dipol	277	0,307	0,669	0,772	1,239
Gauß 4 Dipole	334	0,795	0,901	1,202	1,239
Gauß 8 Dipole	148	0,953	1,093	1,253	0,945

Tabelle B.10.: Messdatensatz Nr. 10: Beste Lokalisierung mit geringster Entfernung zum Sender über alle Startpunkte des Nelder–Mead–Verfahrens pro Modell.

B.2.11. Messdatensatz 11

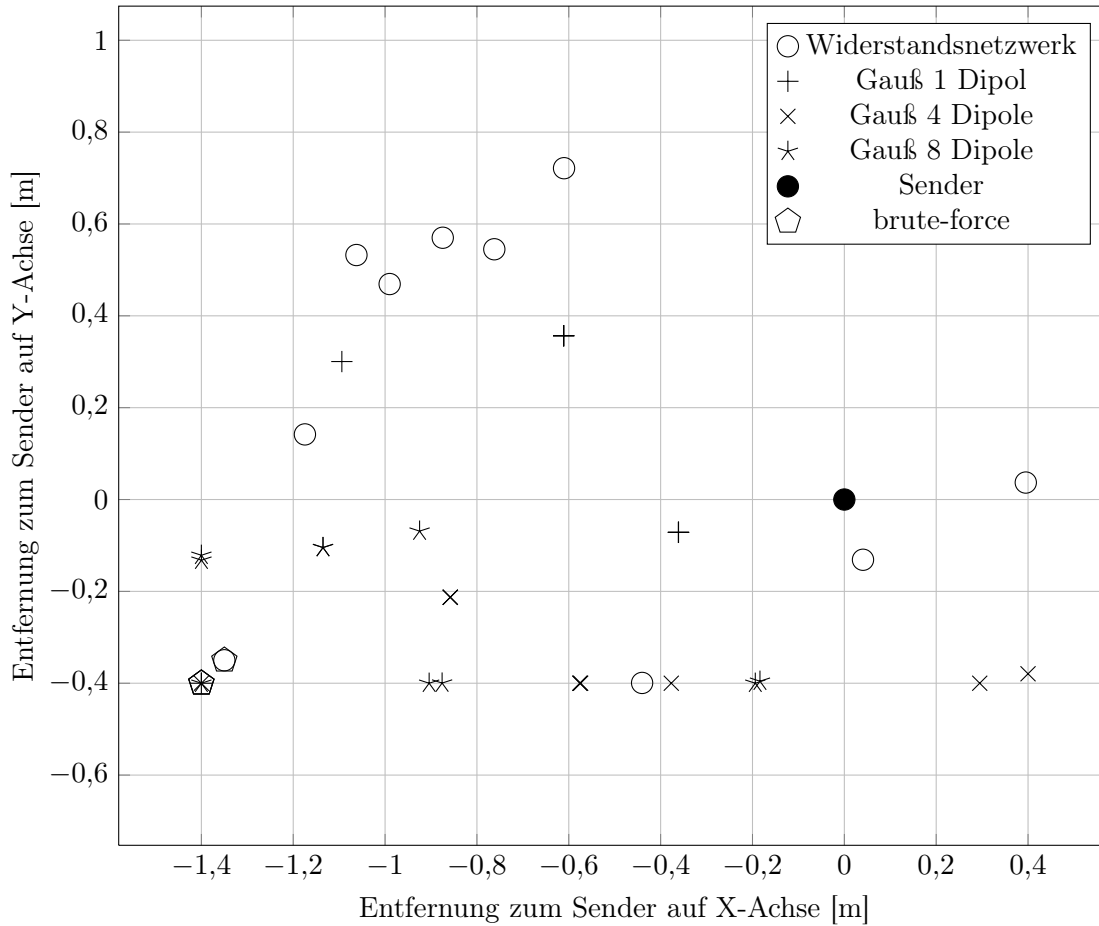


Abbildung B.10.: Messdatensatz 11, berechnete Senderpositionen.

Modell	beste Lokalisierung mit Nelder–Mead				
	Funktionsaufrufe	min. Entf.	ø Entf.	max. Entf.	brute-force-Entf.
WS-Netzwerk	92	0,144	0,838	1,189	1,409
Gauß 1 Dipol	111	0,368	0,608	1,135	1,456
Gauß 4 Dipole	200	0,499	0,709	0,884	1,457
Gauß 8 Dipole	311	0,480	1,006	1,420	1,457

Tabelle B.11.: Messdatensatz Nr. 11: Beste Lokalisierung mit geringster Entfernung zum Sender über alle Startpunkte des Nelder–Mead–Verfahrens pro Modell.

B.2.12. Messdatensatz 12

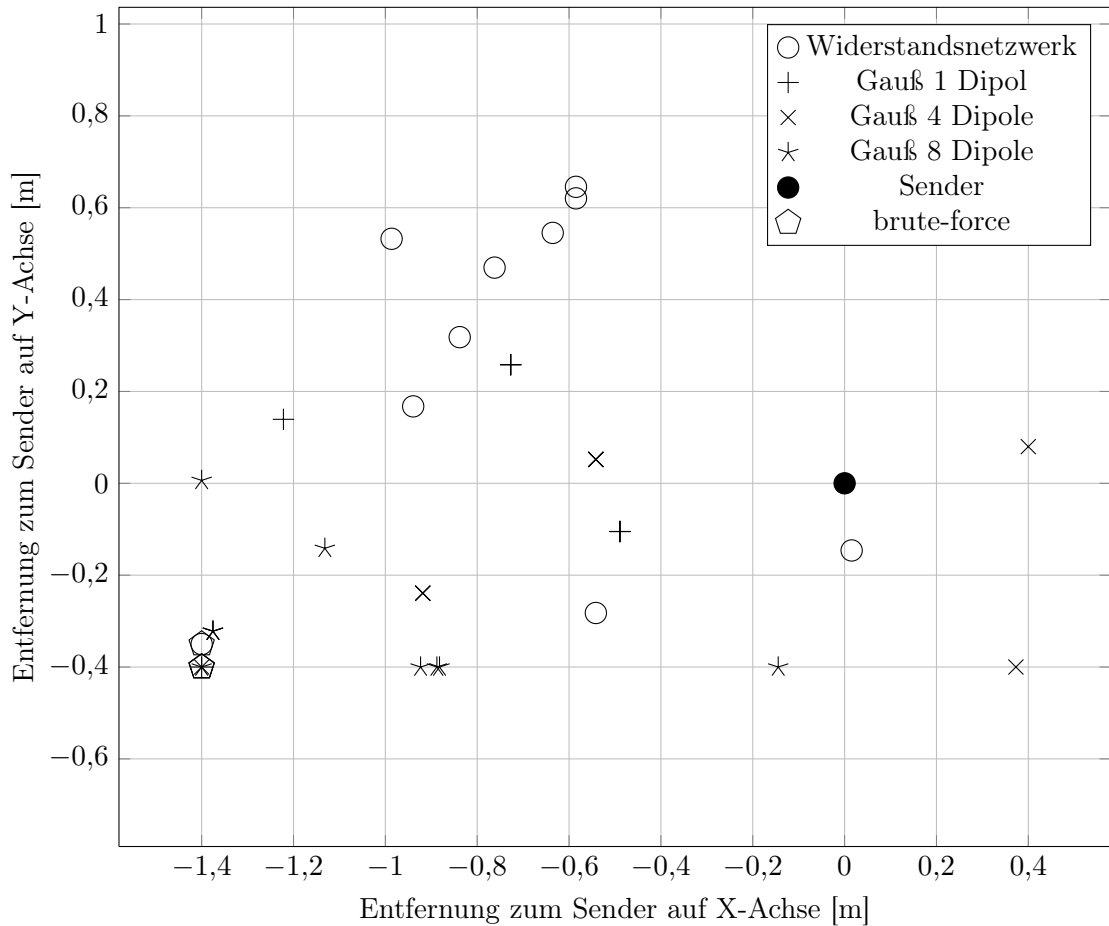


Abbildung B.11.: Messdatensatz 12, berechnete Senderpositionen.

Modell	beste Lokalisierung mit Nelder–Mead				
	Funktionsaufrufe	min. Entf.	ø Entf.	max. Entf.	brute-force-Entf.
WS-Netzwerk	81	0,286	0,830	1,146	1,444
Gauß 1 Dipol	134	0,541	0,723	1,243	1,470
Gauß 4 Dipole	161	0,449	0,683	0,969	1,470
Gauß 8 Dipole	486	0,584	1,160	1,456	1,477

Tabelle B.12.: Messdatensatz Nr. 12: Beste Lokalisierung mit geringster Entfernung zum Sender über alle Startpunkte des Nelder–Mead–Verfahrens pro Modell.

B.2.13. Messdatensatz 13

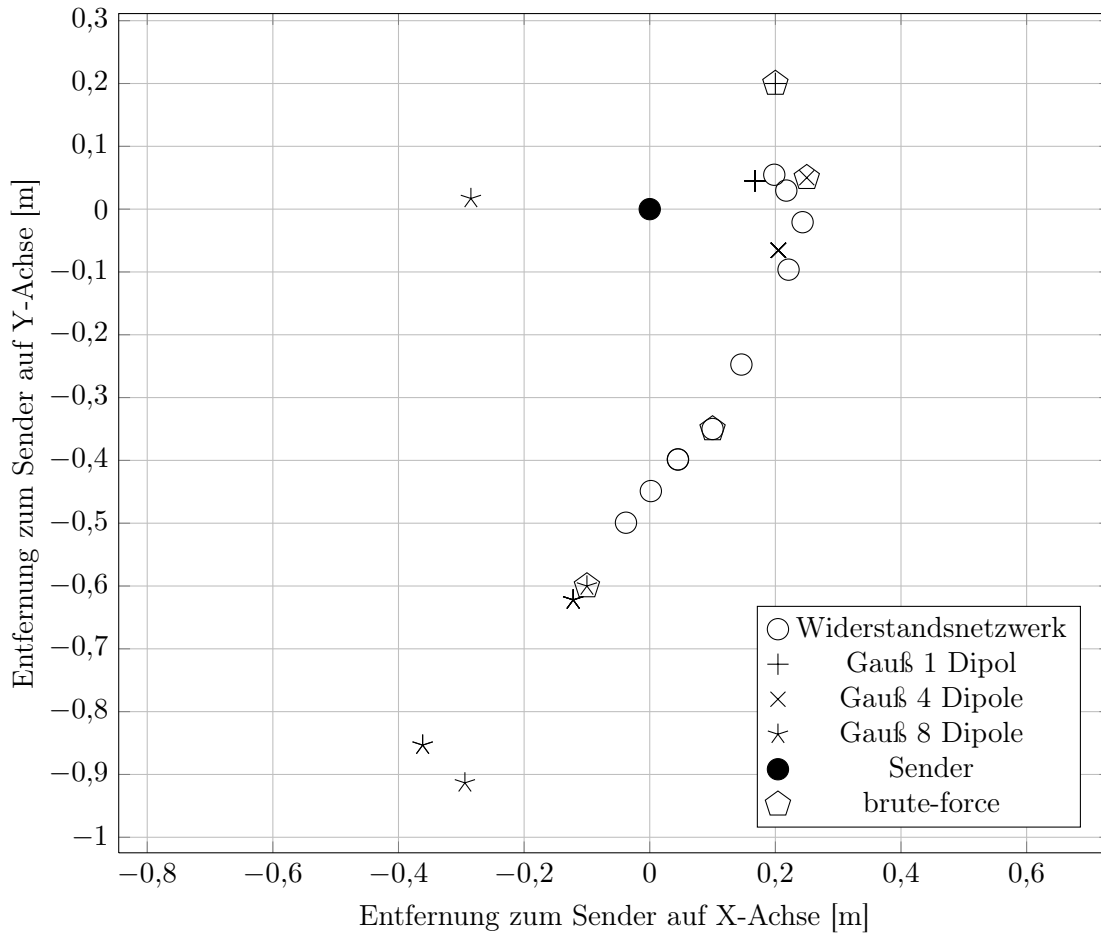


Abbildung B.12.: Messdatensatz 13, berechnete Senderpositionen.

Modell	beste Lokalisierung mit Nelder–Mead				
	Funktionsaufrufe	min. Entf.	ø Entf.	max. Entf.	brute-force-Entf.
WS-Netzwerk	125	0,207	0,328	0,501	0,364
Gauß 1 Dipol	162	0,175	0,175	0,175	0,287
Gauß 4 Dipole	138	0,215	0,215	0,215	0,255
Gauß 8 Dipole	252	0,302	0,715	0,966	0,640

Tabelle B.13.: Messdatensatz Nr. 13: Beste Lokalisierung mit geringster Entfernung zum Sender über alle Startpunkte des Nelder–Mead–Verfahrens pro Modell.

B.2.14. Messdatensatz 14

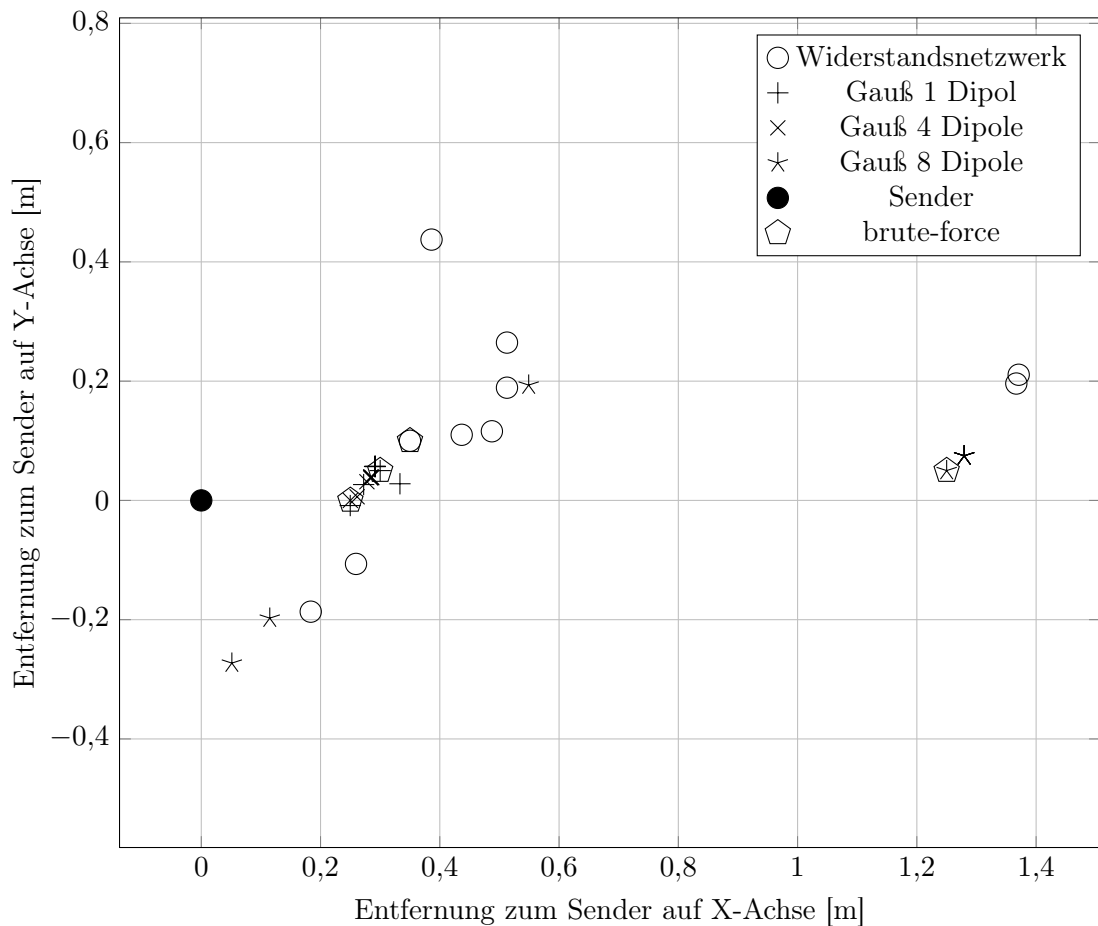


Abbildung B.13.: Messdatensatz 14, berechnete Senderpositionen.

Modell	beste Lokalisierung mit Nelder–Mead				
	Funktionsaufrufe	min. Entf.	ø Entf.	max. Entf.	brute-force-Entf.
WS-Netzwerk	115	0,450	0,785	1,443	0,541
Gauß 1 Dipol	249	0,456	0,494	0,521	0,502
Gauß 4 Dipole	380	0,472	0,589	1,383	0,472
Gauß 8 Dipole	157	0,285	1,002	1,282	1,251

Tabelle B.14.: Messdatensatz Nr. 14: Beste Lokalisierung mit geringster Entfernung zum Sender über alle Startpunkte des Nelder–Mead–Verfahrens pro Modell.

B.2.15. Messdatensatz 15

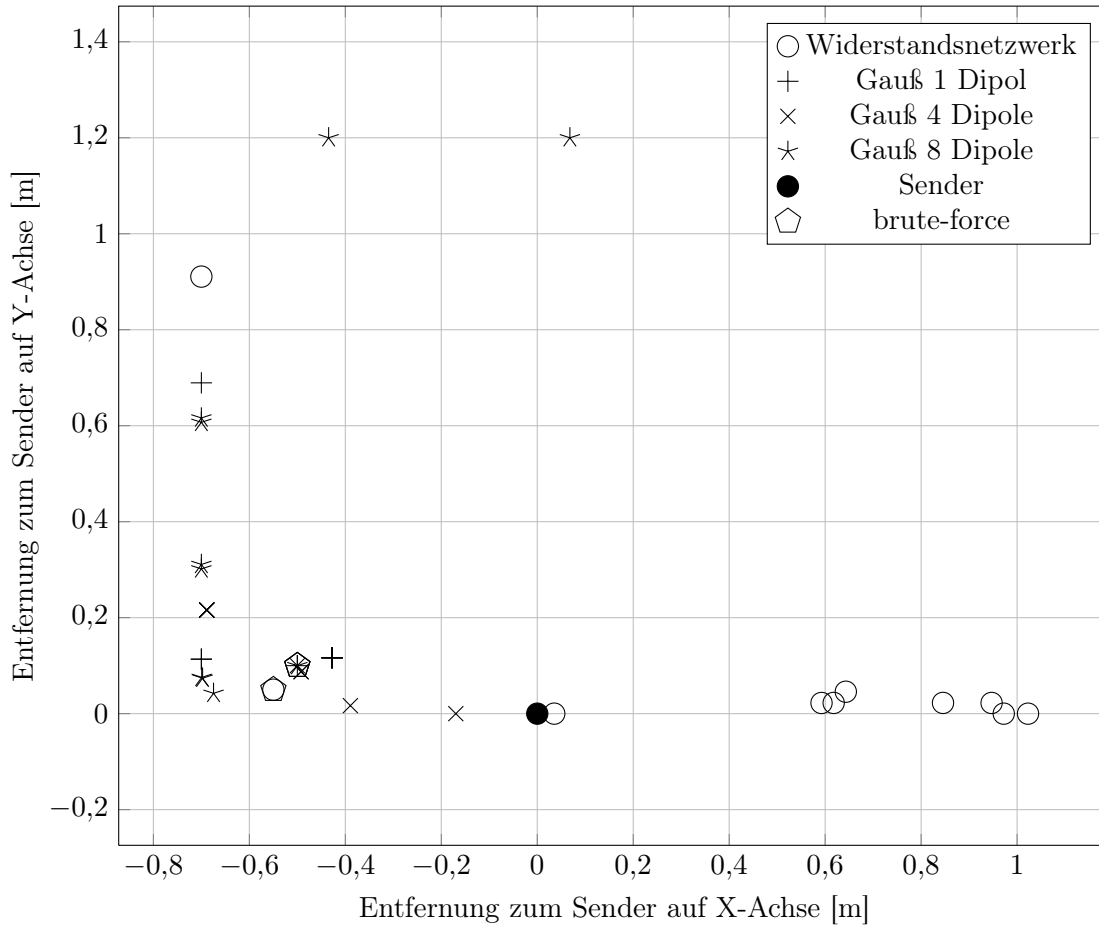


Abbildung B.14.: Messdatensatz 15, berechnete Senderpositionen.

Modell	beste Lokalisierung mit Nelder–Mead				
	Funktionsaufrufe	min. Entf.	ø Entf.	max. Entf.	brute-force-Entf.
WS-Netzwerk	114	0,037	0,761	1,164	0,555
Gauß 1 Dipol	317	0,443	0,565	0,993	0,510
Gauß 4 Dipole	144	0,173	0,579	0,727	0,510
Gauß 8 Dipole	269	0,706	0,891	1,280	0,512

Tabelle B.15.: Messdatensatz Nr. 15: Beste Lokalisierung mit geringster Entfernung zum Sender über alle Startpunkte des Nelder–Mead–Verfahrens pro Modell.

Glossar

Modell	Zusammenhang zwischen Modellparametern und gewonnenen Messdaten.	24, 97
Polytop	Polytop bezeichnet in der Geometrie ein verallgemeinertes Polygon in beliebiger Dimension. Mit k -Polytopen werden Polytope mit k Dimensionen bezeichnet. Ein 0-Polytop ist eine einzelne Ecke (ein Punkt); ein 1-Polytop besteht aus zwei Ecken, die durch eine Kante verbunden sind; ein 2-Polytop besteht aus mehreren, jeweils an einer Ecke verbundenen, einen Zyklus bildenden 1-Polytopen und stellt somit ein Polygon dar; ein 3-Polytop besteht wiederum aus mehreren an den Kanten verbundenen 2-Polytopen und stellt somit ein Polyeder dar; usw.	97
S_{ab}	Ein Sensor, der die Spannung zwischen den Elektroden a und b misst.	34, 35
Simplex	Das (oder manchmal auch der) Simplex oder n -Simplex (Mehrzahl: Simplexe, Simplizia, Simplices oder Simplizes) ist ein Begriff aus der Geometrie und beschreibt ein n -dimensionales Polytop.	38, 39
SPI	Ein vollduplexfähiger synchroner serieller Datenbus mit Master-Slave-Kommunikation.	52
Vorwärtsoperator	Mathematische Formulierung des Modells.	24

Nomenklatur Mathematik

$[0, 1]$	Das Intervall $[0, 1] \in \mathbb{R}$	17
\mathcal{A}	Ein Ereignis im Wahrscheinlichkeitsraum.	v, 17–21, 99, 101
\mathcal{B}	Ein Ereignis im Wahrscheinlichkeitsraum.	v, 17, 21, 99
da	Der Flächeninhalt der Oberfläche in m^2 .	6, 7
\mathcal{D}	Ein Punkt im Datenraum \mathfrak{D} , $\mathcal{D} \in \mathfrak{D}$.	16, 24, 99
\mathfrak{D}	Datenraum.	v, vi, 15, 16, 22–25, 27–29, 34, 99
\oint_S	Das Integral über der geschlossenen Oberfläche S.	6, 7
k_1	$k_1 = \frac{1}{4\pi\epsilon_0\epsilon_r} \frac{Q}{r}$	11, 12, 34–36, 99
$M_{\mathcal{A}}$	Wahrscheinlichkeitsereignis $M_{\mathcal{A}}$: Eine Wahrscheinlichkeitsverteilung, die dem Ereignis \mathcal{A} zugeordnet ist.	20, 21, 99
$M_{\mathcal{B}}$	Wahrscheinlichkeitsereignis $M_{\mathcal{B}}$: Eine Wahrscheinlichkeitsverteilung, die dem Ereignis \mathcal{B} zugeordnet ist.	21, 99
\mathcal{M}	Modell, $\mathcal{M} \in \mathfrak{M}$.	16, 24, 99
\mathfrak{M}	Modellraum.	v, vi, 15, 16, 18, 22–25, 27–29, 31, 99
\cdot	Einfache Multiplikation.	8
\hat{n}	Der Normalenvektor, der senkrecht auf der Fläche S steht.	6, 7

P	$P : \mathcal{A} \rightarrow [0, 1]$. Das Wahrscheinlichkeitsmaß P ist eine Abbildung, die einer Teilmenge \mathcal{A} einer Ereignismenge \mathfrak{X} einen Wert im Intervall $[0, 1]$ zuordnet. Im Regelfall wird dabei mit 0 das unmögliche Ereignis und mit 1 das sichere Ereignis bezeichnet.	17
π	Die Kreiszahl $\pi = 3,1415926\dots$	7, 11, 12, 34, 99, 103
P_{Q_1}	Position der Ladung Q_1 : $P_{Q_1} = (X_1, Y_1, Z_1)$	34, 101
P_{Q_2}	Position der Ladung Q_2 : $P_{Q_2} = (X_2, Y_2, Z_2)$	34, 101
o	Dieses Produkt bedeutet, dass der zu \hat{n} parallele Teil des elektrischen Feldes betrachtet wird.	6, 7
Q_1	Die Ladung an der Position $P_{Q_1} = (X_1, Y_1, Z_1)$ und einer Ladungsstärke von q_1 in Coulomb.	34, 101
Q_2	Die Ladung an der Position $P_{Q_2} = (X_2, Y_2, Z_2)$ und einer Ladungsstärke von q_2 in Coulomb.	34, 101
\mathbb{R}	Die Menge der reellen Zahlen \mathbb{R}	99, 101
\hat{r}	Ein Vektor, der einen Punkt auf der Kugeloberfläche zeigt und parallel zum elektrischen Feld \vec{E} ist.	11, 12, 34
\times	Skalarprodukt	22, 24
\mathfrak{X}	Eine endlichdimensionale Mannigfaltigkeit.	v, vi, 16–22, 24, 25, 101

Nomenklatur Physik

A	Ampere, SI-Einheit für den elektrischen Strom.	103
C	Die SI-Einheit Coulomb für die elektrische Kapazität.	16, 103
\vec{E}	Das gerichtete elektrische Feld.	6, 7, 11, 12, 34
ϵ_0	Die elektrische Feldkonstante (= elektrische Permittivität im Vakuum). $\epsilon_0 = \frac{1}{\mu_0 c^2} = \frac{10^7}{4\pi \cdot 299792458^2} \frac{A^2 s^4}{kgm^3}$ $\approx 8,8541878176 \dots \cdot 10^{-12} \frac{As}{Vm}$ [5, S. 18], [7, S. 322]	6, 7, 11, 12, 34, 99
ϵ_r	Die relative elektrische Permittivität von Wasser. ϵ_r ist abhängig von Druck, Temperatur und der Frequenz des elektrischen Feldes. Unter normalen Bedingungen, also Normaldruck und einer Temperatur von 20 °C, sowie einer Frequenz von bis zu 1 kHz gilt $\epsilon_r \approx 80$, [5, S. 19], [6, S. 293, 313]	11, 12, 34, 99, 103
F	Die SI-Einheit Farad für die elektrische Kapazität. $1F = 1 \frac{C}{V} = 1 \frac{A \cdot s}{V}$	16
m	Die SI-Einheit Meter.	16, 52
Q	Die Ladung Q in <i>Coulomb</i> .	11, 12, 34, 99
q_{enc}	Die eingeschlossene (daher <i>enc</i>) Ladung q in <i>Coulomb</i> .	6, 7
s	Die SI-Einheit Sekunde.	103
V	Volt, SI-Einheit für die elektrische Spannung.	16, 52, 103, 105

Tabellenverzeichnis

3.1.	Verwendete Modellparameter und deren Parametrisierung. Verwendete Einheiten: F: Farad, m: Meter, C: Coulomb.	16
3.2.	Verwendete Daten und deren Parametrisierung. Einheiten: V: Volt, mbar: Millibar	17
6.2.	Allgemeine Parameter	46
6.1.	UML-Klassendiagramm: World, Driver, SimObject, Electrode, Robot .	47
6.3.	Parameter für den Nelder–Mead–Algorithmus	48
6.4.	Parameter für das physikalische Modell	48
6.5.	Aufbau der vom Mikrocontroller an den PC übertragenen Datensätze	51
7.1.	Positionen der Messsonde	54
7.2.	Bestes Modell pro Messdatensatz nach geringster berechneter Entfernung zum Sender. b–f: brute–force–Verfahren nach bestem Modell, N–M: Nelder–Mead–Verfahren	67
7.3.	Bestes Modell pro Messdatensatz nach geringster durchschnittlicher Entfernung zum Sender über alle Startpunkte des Nelder–Mead–Verfahrens. b–f: brute–force–Verfahren nach bestem Modell, N–M: Nelder–Mead–Verfahren	69
B.1.	Verwendete Startpunkte für Nelder–Mead	79
B.2.	Messdatensatz Nr. 2: Beste lokalisierte Positionen	80
B.3.	Messdatensatz Nr. 3: Beste lokalisierte Positionen	81
B.4.	Messdatensatz Nr. 4: Beste lokalisierte Positionen	82
B.5.	Messdatensatz Nr. 5: Beste lokalisierte Positionen	83
B.6.	Messdatensatz Nr. 6: Beste lokalisierte Positionen	84
B.7.	Messdatensatz Nr. 7: Beste lokalisierte Positionen	85
B.8.	Messdatensatz Nr. 8: Beste lokalisierte Positionen	86
B.9.	Messdatensatz Nr. 9: Beste lokalisierte Positionen	87
B.10.	Messdatensatz Nr. 10: Beste lokalisierte Positionen	88
B.11.	Messdatensatz Nr. 11: Beste lokalisierte Positionen	89
B.12.	Messdatensatz Nr. 12: Beste lokalisierte Positionen	90
B.13.	Messdatensatz Nr. 13: Beste lokalisierte Positionen	91
B.14.	Messdatensatz Nr. 14: Beste lokalisierte Positionen	92

B.15. Messdatensatz Nr. 15: Beste lokalisierte Positionen 93

Abbildungsverzeichnis

2.1.	Aufeinander rechtwinklig stehende Vektoren als Basis für eine Sensorkonfiguration	9
2.2.	Feldlinien eines Dipols	11
5.1.	$\frac{\text{Volumen Hyperkugel}}{\text{Volumen Hyperwürfel}}$	38
5.2.	Nelder–Mead–Algorithmus: Reflexion	41
5.3.	Nelder–Mead–Algorithmus: Expansion	42
5.4.	Nelder–Mead–Algorithmus: Innere Kontraktion	42
5.5.	Nelder–Mead–Algorithmus: Äußere Kontraktion	43
5.6.	Nelder–Mead–Algorithmus: Schrumpfen	43
5.7.	Flussdiagramm des Nelder–Mead–Algorithmus	44
7.1.	Vom Sender ohne Last erzeugtes Signal	53
7.2.	Messsonde	54
7.3.	Versuchsaufbau	55
7.4.	Messdaten: Gutes Verhältnis von Nutzsignal und Rauschen	56
7.5.	Messdaten: Übersteuerter Sensor	57
7.6.	Messdaten: Verrauschtes Nutzsignal	58
7.7.	Messdaten: Extraktion der zur Berechnung verwendeten Spannungsdifferenz	59
7.8.	Vergleich Spannungsverteilung: Messwerte und Simulation	60
7.9.	1. Vergleich von $v_{1_{diff}}$ -Werten mit Simulationsmodellen	61
7.10.	2. Vergleich von $v_{1_{diff}}$ -Werten mit Simulationsmodellen	62
7.11.	Ladung vor isolierender Wand mit zugehöriger Bildladung	65
7.12.	Standardabweichung simulierter Spannungen jedes Modells für jeden Datensatz	68
B.1.	Messdatensatz 2, berechnete Senderpositionen.	80
B.2.	Messdatensatz 3, berechnete Senderpositionen.	81
B.3.	Messdatensatz 4, berechnete Senderpositionen.	82
B.4.	Messdatensatz 5, berechnete Senderpositionen.	83
B.5.	Messdatensatz 6, berechnete Senderpositionen.	84
B.6.	Messdatensatz 7, berechnete Senderpositionen.	85
B.7.	Messdatensatz 8, berechnete Senderpositionen.	86
B.8.	Messdatensatz 9, berechnete Senderpositionen.	87
B.9.	Messdatensatz 10, berechnete Senderpositionen.	88
B.10.	Messdatensatz 11, berechnete Senderpositionen.	89

B.11. Messdatensatz 12, berechnete Senderpositionen.	90
B.12. Messdatensatz 13, berechnete Senderpositionen.	91
B.13. Messdatensatz 14, berechnete Senderpositionen.	92
B.14. Messdatensatz 15, berechnete Senderpositionen.	93

Literaturverzeichnis

- [1] BERBERIAN, Sterling K.: *Measure and Integration*. American Mathematical Soc., 2011
- [2] BROOKS, Christopher H.: *An Introduction to Amoeba*
<http://www.cs.usfca.edu/~brooks/papers/amoeba.pdf>
- [3] BULLOCK, Theodore H. ; HOPKINS, Carl D. ; POPPER, Arthur N. (Hrsg.) ; FAY, Richard R. (Hrsg.): *Electroreception*. Springer Science + Business Media Inc., 2005
- [4] DEBARRA, G.: *Measure Theory and Integration*. New Age International, 1981
- [5] FLEISCH, Daniel A. ; FLEISCH, Daniel A. (Hrsg.): *A Student's guide to Maxwell's equations*. Cambridge University Press, Cambridge, 2008
- [6] FLIESSBACH, Torsten ; FLIESSBACH, Torsten (Hrsg.): *Elektrodynamik, 5. Auflage*. Spektrum Akademischer Verlag, Heidelberg, 2008
- [7] HERING, Ekbert ; MARTIN, Rolf ; STOHRER, Martin: *Taschenbuch der Mathematik und Physik, 5. Auflage*. Springer, 2009
- [8] HOPKINS, Carl D.: Design Features for Electric Communication. In: *The Journal of Experimental Biology* (1999), Nr. 202, S. 1217–1228
- [9] KALMIJN, A. J.: Detection of Weak Electric Fields. In: *Sensory Biology of Aquatic Animals* (1988), S. 151–186
- [10] KERNBACH, S. ; DIPPER, T. ; SUTANYO, D.: Multi-Modal Local Sensing and Communication for Collective Underwater systems. In: *Proc. Robotica 2011* (2011)
- [11] KOCH, Karl-Rudolf: *Einführung in die Bayes-Statistik*. Springer, 2000
- [12] LAGARIAS, Jeffrey C. ; REEDS, James A. ; WRIGHT, Margaret H. ; WRIGHT, Paul E.: Convergence Properties of the Nelder-Mead Simplex Method in Low Dimensions. In: *SIAM Journal of Optimization* 9 (1998), Nr. 1, S. 112–147
- [13] MENKE, William ; MENKE, William (Hrsg.): *Geophysical data analysis: discrete inverse theory*. Academic Press, Inc., Orlando, Florida, 1984
- [14] NELDER, John A. ; MEAD, Roger: A Simplex Method for Function Minimization. In: *Computer Journal* (1965), Nr. 7, S. 308–313

- [15] RANJAN, T.N.: Navigation of autonomous underwater vehicle using extended kalman filter. In: *Trends in Intelligent Robotics: 15th Robot World Cup and Congress, FIRA 2010, Bangalore, India, September 15-19, 2010, Proceedings* (2010), S. 1–9
- [16] SCHWARTZ, Melvin ; SCHWARTZ, Melvin (Hrsg.): *Principles of Electrodynamics*. Dover Publications, Inc., New York, 1987
- [17] STRANG, Gilbert ; STRANG, Gilbert (Hrsg.): *Computational Science and Engineering*. Wellesley-Cambridge Press, 2007
- [18] TARANTOLA, Albert: *Probability Density versus Volumetric Probability*
<http://www.cmmacs.ernet.in/inverse/Diverse/DensityVersusVolume/DensityVersusVolume.pdf>. Version: 2004
- [19] TARANTOLA, Albert ; TARANTOLA, Albert (Hrsg.): *Inverse Problem Theory and Methods for Model Parameter Estimation*. Institut de Physique du Globe de Paris, Paris, France, 2005
- [20] VATTA, Frederica ; MENEGHIHI, Fabio ; ESPOSITO, Fabrizio ; MININEL, Stefano ; DI SALLE, Francesco: Realistic and Spherical Head Modeling for EEG Forward Problem Solution: A Comparative Cortex-Based Analysis. In: *Computational Intelligence and Neuroscience 2010* (2010)

Erklärung

Hiermit versichere ich, diese Arbeit selbständig verfasst und nur die angegebenen Quellen benutzt zu haben.

(Christian Stürmer)