UNIVERSITÄT ZU LÜBECK

From the Institute for Electrical Engineering in Medicine
of the University of Lübeck

Director: Prof. Dr. Philipp Rostalski

Model-based Probabilistic Inference for Monitoring Respiratory Effort using the Surface Electromyogram

Dissertation

for Fulfillment of
Requirements
for the Doctoral Degree
of the University of Lübeck

from the Department of Computer Sciences and Technical Engineering

Submitted by

Eike Petersen

from Salzgitter

Lübeck, 2021

Abstract

In assisted mechanical ventilation, patients share the work of breathing with a mechanical ventilator. Medical research has repeatedly emphasized the need to monitor the patient's own respiratory effort under these conditions, but there is currently no noninvasive measurement modality that can achieve this. Surface electromyographic (EMG) measurements of the respiratory muscles have recently been proposed as a potential solution, but several critical signal processing challenges must be overcome to achieve a reliable clinical monitoring methodology. Crosstalk from nearby muscles, including the heart, represents a challenge to measurement fidelity. Moreover, electromyographic measurements by themselves can only indicate the *relative* degree of muscle activation but not the absolute values of pressure generated by the respiratory muscles. In this thesis, several contributions towards the solution of these problems are made.

Firstly, a new, comprehensive *mathematical model* of surface electromyography, muscular force generation, and motor unit pool organization is developed and presented. The model significantly extends previously available models in multiple regards; particular care is taken to model the interrelated distributions of muscle fiber parameters in a way that ensures realistic electro-mechanical macroscopic muscle properties. This model is then used for the performance evaluation of a newly proposed *convolutive blind source separation* method based on the TRINICON framework (TRIple-N Independent component analysis for CONvolutive mixtures), which has been proposed for separating the activity of inspiratory, expiratory, and cardiac activity in multi-channel respiratory surface EMG recordings. (The derivation of the source separation algorithm itself was not part of this thesis.) No ground truth measurements exist for this separation problem, and thus, quantitative evaluation with real measurements is infeasible.

Secondly, the problem of cardiac artifact removal from single-channel surface EMG recordings of the respiratory muscles is addressed comprehensively. Due to the vicinity of the recording electrodes to the heart, these measurements suffer from unusually strong cardiac artifacts, which are nontrivial to remove. Many algorithms of all kinds have been proposed in the literature for this purpose. Here, two contributions are made to the solution of this problem. On the one hand, a novel, model-based, probabilistic method for removing cardiac artifacts from single-channel recordings is proposed, called *probabilistic adaptive template subtraction (PATS)*. The method is based on a parametric model of the beat-to-beat variability of heartbeat morphology while not assuming any restrictive model of the morphology itself. A fully probabilistic inference procedure is proposed, based on maximum likelihood estimation and a bank of Kalman filters and smoothers. On the other hand, a comprehensive *evaluation framework* for algorithms for cardiac artifact removal is developed. Algorithms are evaluated both on real measurements obtained in a dedicated study with healthy participants, conducted solely for this purpose, and on synthetic signals. To comprehensively assess algorithm performance, four different performance measures are considered. The newly proposed algorithm, PATS, performs best among a group of previously proposed algorithms in recovering the raw, uncontaminated EMG signals.

Thirdly, a purely theoretical contribution to the problem of regression under *covariate shift* (i.e., differing training and target distributions), *concept drift* (i.e., a time-varying system), and *model mismatch* is made. This triad of challenges is arguably very common, yet surprisingly little research has been devoted to a principled solution. Here, based on a

probabilistic model of concept drift, a fully probabilistic solution to this problem is presented. As parts of this solution, new, weighted versions of the standard linear Kalman filter and smoother are derived, and an importance-weighted hyperparameter estimation scheme based on prediction error minimization is proposed. Numerical examples demonstrate that the proposed approach can be employed to successfully perform regression in the face of covariate shift, concept drift, and model mismatch.

Finally, utilizing the aforementioned statistical developments, a new method for estimating the pressure P_{mus} generated by the respiratory muscles of patients under assisted mechanical ventilation is presented. The proposed method performs sensor fusion and combines the pneumatic measurements provided by a mechanical ventilator with respiratory surface EMG measurements. To this end, models of respiratory mechanics, of the relationship between respiratory EMG measurements and P_{mus}, and of the dynamic time course of P_{mus} are combined. Unlike many previously proposed approaches, the method does not rely on any special identification maneuvers being performed and can be applied during normal ventilation. In order to improve estimation results, importance weighting and an automatic delay estimation scheme are exploited. Based on an evaluation on eleven recordings from intensive care patients, the proposed algorithm estimates P_{mus} with a mean absolute deviation of just 0.79 mbar, a rank correlation coefficient of 0.74, and 95% limits of agreement of -2.49 mbar and 2.65 mbar. It thus represents an attractive solution for future entirely noninvasive monitoring of respiratory effort in mechanically ventilated patients.

Zusammenfassung

Bei der unterstützten künstlichen Beatmung teilen sich die Patienten die Atemarbeit mit einem Beatmungsgerät. Die medizinische Forschung hat wiederholt die Notwendigkeit betont, die eigene Atemarbeit des Patienten unter diesen Bedingungen zu überwachen, aber es gibt derzeit keine nicht-invasive Messmethode, die dies leisten kann. Kürzlich wurden elektromyografische Oberflächenmessungen (EMG) der Atemmuskulatur als mögliche Lösung vorgeschlagen, doch müssen mehrere kritische Herausforderungen bei der Signalverarbeitung überwunden werden, um eine zuverlässige klinische Überwachungsmethode zu entwickeln. Das Übersprechen von benachbarten Muskeln, einschließlich des Herzens, stellt eine Herausforderung für die Messgenauigkeit dar. Darüber hinaus können elektromyographische Messungen nur den relativen Grad der Muskelaktivierung anzeigen, nicht aber die absoluten Werte des von den Atemmuskeln erzeugten Drucks. In dieser Arbeit werden mehrere Beiträge zur Lösung dieser Probleme geleistet.

Erstens wird ein neues, umfassendes mathematisches Modell der Oberflächenelektromyographie, der Muskelkrafterzeugung und der Organisation von motorischen Einheiten entwickelt und vorgestellt. Das Modell erweitert die bisher verfügbaren Modelle in mehrfacher Hinsicht; besondere Sorgfalt wurde darauf verwendet, die miteinander verknüpften Verteilungen der Muskelfaserparameter so zu modellieren, dass realistische elektromechanische makroskopische Muskeleigenschaften gewährleistet sind. Dieses Modell wird dann für die Leistungsbewertung einer neu vorgeschlagenen *konvolutiven blinden Quellentrennungs-Methode* verwendet, die auf dem TRINICON-Framework (TRIple-N Independent component analysis for CONvolutive mixtures) basiert, das für die Trennung der Aktivität von inspiratorischer, exspiratorischer und kardiogene Aktivität in mehrkanaligen respiratorischen Oberflächen-EMG-Aufzeichnungen vorgeschlagen wurde. (Die Herleitung des Quellentrennungsalgorithmus selbst war nicht Teil dieser Arbeit.) Für dieses Trennungsproblem gibt es keine Referenzmessungen, so dass eine quantitative Auswertung mit realen Messungen nicht möglich ist.

Zweitens wird das Problem der Entfernung kardiogener Artefakte aus einkanaligen Oberflächen-EMG-Aufzeichnungen der Atemmuskulatur umfassend behandelt. Aufgrund der Nähe der Aufzeichnungselektroden zum Herzen leiden diese Messungen unter ungewöhnlich starken kardiogenen Artefakten, deren Entfernung nicht trivial ist. Zu diesem Zweck wurden in der Literatur zahlreiche Algorithmen aller Art vorgeschlagen. Hier werden zwei Beiträge zur Lösung dieses Problems geleistet. Zum einen wird ein neuartiges, modellbasiertes, probabilistisches Verfahren zur Entfernung kardiogener Artefakte aus einkanaligen Aufzeichnungen vorgeschlagen, das als *probabilistische adaptive Muster-Subtraktion* bezeichnet wird. Die Methode basiert auf einem parametrischen Modell der Variabilität der Herzschlagmorphologie zwischen aufeinander folgenden Herzschlägen, ohne dass ein restriktives Modell der Morphologie selbst vorausgesetzt wird. Es wird ein vollständig probabilistisches Inferenzverfahren vorgeschlagen, das auf einer Maximum-Likelihood-Schätzung und einem Ensemble von Kalman-Filtern und -Glättern beruht. Auf der anderen Seite wird ein umfassendes Evaluationsframework für Algorithmen zur Entfernung kardiogener Artefakte entwickelt. Die Algorithmen werden sowohl an realen Messungen, die in einer eigens für diesen Zweck durchgeführten Studie mit gesunden Teilnehmern gewonnen wurden, als auch an synthetischen Signalen bewertet. Um die Leistung der Algorithmen umfassend zu bewerten, werden vier verschiedene Leistungsmaße berücksichtigt. Der neu vorgeschlagene

Algorithmus schneidet bei der Wiederherstellung der unverfälschten EMG-Rohsignale am besten von allen zuvor vorgeschlagenen Algorithmen ab.

Drittens wird ein rein theoretischer Beitrag zum Problem der Regression unter *Kovariatenverschiebung* (d.h. unterschiedliche Trainings- und Zielverteilungen), *Konzeptdrift* (d.h. ein zeitlich veränderliches System) und *Modellfehlern* geleistet. Dieses Dreigespann von Herausforderungen ist praktisch häufig anzutreffen, aber erstaunlich wenig Forschung wurde bisher einer prinzipiellen Lösung gewidmet. Hier wird auf der Grundlage eines probabilistischen Modells des Konzeptdrifts eine vollständig probabilistische Lösung für dieses Problem vorgestellt. Als Teil dieser Lösung werden neue, gewichtete Versionen des standardmäßigen linearen Kalman-Filters und der -Glätters abgeleitet und ein Schema zur Schätzung von Hyperparametern mit Wichtigkeitsgewichtung vorgeschlagen, das auf der Minimierung von Vorhersagefehlern basiert. Numerische Beispiele zeigen, dass der vorgeschlagene Ansatz zur erfolgreichen Durchführung von Regressionen angesichts von Kovariatenverschiebung, Konzeptdrift und Modellfehlern eingesetzt werden kann.

Schließlich wird unter Verwendung der oben erwähnten statistischen Entwicklungen eine neue Methode zur Schätzung des Drucks P_{mus} vorgestellt, der von den Atemmuskeln von Patienten unter unterstützter künstlicher Beatmung erzeugt wird. Die vorgeschlagene Methode führt eine Sensorfusion durch und kombiniert die von einem Beatmungsgerät gelieferten pneumatischen Messungen mit oberflächlichen EMG-Messungen der Atemmuskeln. Zu diesem Zweck werden Modelle der Atemmechanik, der Beziehung zwischen EMG-Messungen der Atemmuskulatur und P_{mus} sowie des dynamischen Zeitverlaufs von P_{mus} kombiniert. Im Gegensatz zu vielen zuvor vorgeschlagenen Ansätzen ist die Methode nicht auf die Durchführung spezieller Identifikationsmanöver angewiesen und kann während der normalen Beatmung angewendet werden. Um die Schätzergebnisse zu verbessern, werden eine Wichtigkeitsgewichtung und eine Methode zur automatischen Identifikation von Signalverzögerungen verwendet. Basierend auf einer Auswertung anhand von elf Aufnahmen von Intensivpatienten schätzt der vorgeschlagene Algorithmus P_{mus} mit einer mittleren absoluten Abweichung von nur 0.79 mbar, einem Rangkorrelationskoeffizienten von 0, 74 und 95%-igen Übereinstimmungsgrenzen von -2.49 mbar und 2.65 mbar. Die vorgeschlagene Methode stellt somit eine attraktive Lösung für die zukünftige nicht-invasive Überwachung der Atemanstrengung bei künstlich beatmeten Patienten dar.

FOREWORD

I N 2014, while working on my master's thesis at the Dräger company, I was working with Marcus Eger on signal processing for electromyographic (EMG) measurements. We wanted to develop algorithms for separating the activity of different respiratory muscles in the measured signals. After a while, I realized that validation was a crucial issue: there is no way of *knowing* which muscles contribute how much activity, so how would I know whether my algorithm worked well? I started working on a simple EMG simulator for the validation. When Marcus asked me how long I believed it would take to finish my work on the simulator, I replied, confidently, "two weeks." About a month later, Marcus asked me again, and I — still only slightly more knowledgeable about EMG modeling than four weeks ago — replied, "Well, it was a bit more complex than I thought, but from here it will probably take me two more weeks." I ended up writing my whole master's thesis about EMG signal modeling (instead of source separation algorithms), and a significant part of my Ph.D. project would still be concerned with this topic. This was my first serious academic adventure, and I thoroughly enjoyed it.

Already while working on my master's thesis, I knew that I wanted to continue doing research afterwards. I had also realized that in the research unit at Dräger in Lübeck, they had all the ingredients of a perfect workplace: very kind and experienced colleagues, thrilling research problems — where thrilling for me meant "technically challenging and societally relevant" — and a beautiful city close to the sea. Unfortunately, no Ph.D. positions were available at the research unit at the time, and so I started looking for other opportunities. It turned out that, at exactly that time, Philipp Rostalski, who later became my Ph.D. supervisor, was founding his new institute at the University of Lübeck and was looking for his first Ph.D. student to work on an EMG-based mechanical ventilation project with the Dräger research unit! We met, quickly agreed that it was the perfect match for both of us, and so I became a Ph.D. student at the university on the same day the institute was founded. I could hardly have been more fortunate.

The six-and-a-half years of which this thesis is a product were, above all, *rich*. Participating in setting up the new institute was both exciting and challenging. Initially, we were just three people: Philipp Rostalski, as my thesis supervisor, Christian Herzog, as our first postdoc, and me. It felt like a startup. During the first week, I brought my coffee machine from home. (At the time of writing, several thousands of cups of coffee later, it's still there and running.) The following years saw the institute quickly grow to over twenty researchers; it became a vibrant space for research at the intersection between theory and practical applications. Besides research, the years filled with conferences, workshops, summer schools, student supervision, teaching duties, administering the institute's IT infrastructure for a while, a long-distance relationship, two moves to different cities, music festivals, board game nights, and a global pandemic. Above all, however, they were filled with wonderful *people*.

First and foremost, I am grateful to Philipp, who succeeded in striking a very fine balance between entrusting me with lots of responsibilities and freedom to explore, while also

CONTENTS

MATHEMATICAL NOTATION

O NE GAINS MANY INSIGHTS during the writing of a technical Ph.D. thesis, one of them being: *notation is a nightmare.* Notation, like so many other things, can only be understood and judged given a *context*, which inevitably varies from section to section, from writer to reader, and — given the interdisciplinary content of this thesis — from reader to reader. I gave my best effort to use clear and consistent notation throughout this thesis, yet I am certain to have fallen short. For any ambiguous and confusing choices, I apologize to the reader in advance.

Matrices, like A and B, will generally be represented by uppercase variables, while vectors, like x and y, will be denoted by lowercase letters. Physical quantities such as electromyographic signals (EMG), lung volume (V), and various pressures (P_{aw}, P_{es}) are an important exception from this rule: they will be denoted as illustrated, but should be understood as (usually discrete-time) time-varying signals, i.e., vectors. Continuously time-varying quantities will be denoted as $x(t)$, whereas individual samples of discrete-time quantities will be denoted as x_k, and the index k will generally be reserved for the sample index. Sometimes, the short-hand notation $x_{3:7}$ will be used to denote x_k for $k = 3, \dots, 7$. Loop variables that change as a function of the iteration will sometimes be denoted as $x^{(\ell)}$ to distinguish this from sample, row, and column indexing. Finally, $\mathcal{N}(\mu_x, \Sigma_x)$ will often be used to denote a normal distribution of a variable x with mean μ_x and covariance matrix Σ_x; sometimes, the alternative notation $\mathcal{N}(x \mid \mu_x, \Sigma_x)$ is used to clarify that the distribution of the variable x is meant.

CHAPTER 1

INTRODUCTION

It makes me so happy. To be at the beginning again, knowing almost nothing.

(Tom Stoppard, Arcadia)

Contents

I N EARLY 2020, the then-emerging COVID-19 epidemic caused a surge
of public interest in a field that is usually much less prominently discussed: *mechanical ventilation.* Under normal circumstances, this branch of medicine receives little public attention, not because it is any less relevant, but rather because it is an old field of medical practice and may be perceived as less exciting than more recently developed fields of inquiry.   However, significant advances to the tremendous benefit of mechanically ventilated patients have been made in recent decades, and yet there is still much room for high-impact clinical improvements, making it an exciting research field. One such avenue for treatment improvement concerns the monitoring of mechanically ventilated patients' *respiratory effort*, a challenge that has been recognized as crucial for providing optimal respiratory support [2, 13].[1] The current clinical gold standard for measuring respiratory effort is to measure esophageal pressure using an esophageal catheter, which is invasive and requires a skilled operator [18, 23]. Esophageal electromyographic (EMG) measurements represent an alternative measure of respiratory effort, which, however, also requires the invasive placement of an esophageal catheter [35, 36].   More recently, *surface* EMG measurements — using electrodes placed on the skin surface above various respiratory muscles — have emerged as an alternative, entirely noninvasive measurement modality [1, 22]. Surface EMG measurements are traditionally heavily used in the prosthetics research field [10, 20] — and indeed, assisted mechanical ventilation can be understood as providing a "respiratory prosthesis" [12]. However, respiratory surface EMG measurements suffer from a comparatively low signal–to–noise ratio due to disturbances such as crosstalk from adjacent muscles, including the heart, rendering sophisticated signal processing algorithms indispensable for obtaining high-fidelity signals. The recording environment in an intensive care unit is also generally challenging, and unlike in traditional prosthetics, no voluntary calibration maneuvers can be performed by patients. Moreover, an intrinsic limitation of EMG measurements is that they can only provide a *relative* measure of respiratory effort since the ratio of the electrical field strength and the force generated by a muscle depend on a multitude of patient-specific (and usually unmeasurable) factors. This is, thus, the core theme of this thesis: to use advanced *mathematical and statistical methods* for modeling electromyography and respiration, and to develop inference algorithms with the ultimate aim of reliable respiratory monitoring using respiratory surface EMG measurements. Along the path, contributions to many different disciplines are made, including the mathematical modeling of electromyography and muscle physiology, blind source separation for respiratory surface EMG measurements, the removal of cardiac artifacts from single-channel respiratory EMG measurements, statistical inference under concept drift, covariate shift, and model mismatch, and, finally, the entirely noninvasive identification of a patient's respiratory effort using surface EMG measurements.

Mechanical Ventilation

Respiratory Effort

Surface EMG

Mathematical and Statistical Methods

1.1 Key contributions and outline of this thesis

To set the stage for the following chapters and introduce the reader to key concepts that may be unfamiliar, chapters 2 and 3 provide background knowledge on statistical inference, respiratory and muscular physiology, mechanical ventilation, and electromyography.

---

[1]While COVID-19 is not a topic of this thesis, it shall be mentioned briefly that the importance of monitoring the respiratory drive of patients with COVID-19 has been emphasized in the emerging literature [4, 9, 11].

As the first major contribution of this thesis, chapter 4 introduces a comprehensive, *mathematical model of surface electromyography*, muscular force generation, and motor unit pool organization. This simulation model is then used (in the same chapter) to evaluate a novel algorithm for *blind source separation* of respiratory surface EMG measurements, which may be used to separate the activities of multiple respiratory muscles from mixture signals, suppress undesired muscle crosstalk, or remove cardiac artifacts.

In the following chapter 5, the related problem of removing *cardiac artifacts* from single-channel respiratory surface EMG measurements is considered. The key contributions of this chapter are twofold. Firstly, a novel, partially model-based, and partially nonparametric algorithm (termed *probabilistic adaptive template subtraction*, PATS) for cardiac artifact removal is proposed. Secondly, a comprehensive evaluation framework for the performance quantification of algorithms proposed for this problem is developed. The framework includes an evaluation in both real measurement signals obtained from a study of healthy subjects and in synthetic signals. Four different performance measures are considered, providing a comprehensive picture of algorithm performance, and ten algorithms are compared, four of which are variants of the newly proposed PATS algorithm. Based on this performance evaluation, the newly proposed algorithm appears to currently be the most accurate available option for recovering the raw EMG signal free of cardiac contamination.

Chapter 6 may appear like an outlier in the context of this thesis, as it treats an abstract statistical inference problem without any apparent connection to respiration or electromyography. However, the methods developed in this chapter will be employed in the following chapter 7 for estimating respiratory effort based on surface EMG measurements. The statistical problem is this: how can one perform regression in the face of *concept drift* (i.e., the system under consideration is time-varying), *covariate shift* (i.e., the distribution of the available training data differs from the target data distribution), and *model mismatch* (i.e., the "true" model is not in the selected model class)? This problem has received surprisingly little consideration in the literature. In this chapter, a fully model-based method for performing regression under these three constraints is developed. Novel, *weighted* formulations of the standard Kalman filter and smoother are derived, combined with an *importance weighting* procedure, and augmented with an importance-weighted hyperparameter estimation procedure based on prediction error minimization. Numerical examples show the validity of the proposed framework for performing time-varying regression under covariate shift and model mismatch.

The following chapter 7 then utilizes the developed statistical framework for *estimating the pressure P_{mus}* generated by the respiratory muscles of a patient under mechanical ventilation. To this end, a model-based sensor fusion method is developed that exploits pneumatic measurements and multiple channels of respiratory surface EMG measurements. All available measurements are combined using models of respiratory mechanics, electromechanical coupling, and the dynamic time course of P_{mus}. It is the first method of this kind described in the literature. The proposed method is evaluated using twelve recordings from nine intensive care patients. The results are very promising and warrant further investigation.

The contributions described in this thesis have been the subject of three peer-reviewed journal publications [26, 32, 33], two peer-reviewed conference publications [25, 27], and three peer-reviewed conference abstracts [29–31] first-authored by the author. In addition,

A Model of Surface Electromyography

Blind Source Separation

Cardiac Artifacts

Concept Drift

Covariate Shift

Model Mismatch

Importance Weighting

Estimating the Pressure P_{mus}

five patent applications [5–8, 21], a survey preprint [34], two peer-reviewed journal publications [17, 19] and six peer-reviewed conference publications [3, 14–16, 24, 28] on related topics have been co-authored by the author. Source code is openly available for the surface EMG simulation model,[2] all cardiac artifact removal algorithms under consideration and the cardiac artifact removal evaluation framework,[3] as well as a the multivariate density estimation scheme developed in chapter 6.[4] Several of the key contributions described in this thesis have not yet been publicized, including the PATS algorithm for cardiac artifact removal, all elements of chapter 6, and the sensor fusion procedure described in chapter 7.

References

[1] G. Bellani *et al.*, "Measurement of diaphragmatic electrical activity by surface electromyography in intubated subjects and its relationship with inspiratory effort," *Respiratory Care*, vol. 63, no. 11, pp. 1341–1349, 2018. DOI: `10.4187/respcare.06176`.

[2] M. Bertoni, S. Spadaro, and E. C. Goligher, "Monitoring patient respiratory effort during mechanical ventilation: Lung and diaphragm-protective ventilation," *Critical Care*, vol. 24, no. 1, 2020. DOI: `10.1186/s13054-020-2777-y`.

[3] H. Buchner, E. Petersen, M. Eger, and P. Rostalski, "Convolutive blind source separation on surface EMG signals for respiratory diagnostics and medical ventilation control," in *38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Florida, USA, 2016. DOI: `10.1109/EMBC.2016.7591513`.

[4] P. Cruces *et al.*, "A physiological approach to understand the role of respiratory effort in the progression of lung injury in SARS-CoV-2 infection," *Critical Care*, vol. 24, no. 1, 2020. DOI: `10.1186/s13054-020-03197-7`.

[5] M. Eger, L. Kahl, P. Rostalski, and E. Petersen, "Apparatus, method, and computer program for influencing the respiration of a person," DE112020000114A5 / WO2020188069A1, 2019.

[6] M. Eger, E. Petersen, J. Graßhoff, and P. Rostalski, "Method and signal processing unit for determining the respiratory activity of a patient," DE112020000233A5 / WO2021052791A1, 2019.

[7] M. Eger, P. Rostalski, E. Petersen, and J. Graßhoff, "Method and signal processing unit for determining the respiratory activity of a patient," DE102019007717B3 / WO2021089215A1, 2019.

[8] ——, "Process and signal processing unit for determining a pneumatic parameter with the use of a lung-mechanical model and of a gradient model," DE102020133460A1 / US20210205561A1, 2020.

[9] P. Esnault *et al.*, "High respiratory drive and excessive respiratory efforts predict relapse of respiratory failure in critically ill patients with COVID-19," *American Journal of Respiratory and Critical Care Medicine*, vol. 202, no. 8, pp. 1173–1178, 2020. DOI: `10.1164/rccm.202005-1582le`.

[10] D. Farina and M. Sartori, "Surface electromyography for man–machine interfacing in rehabilitation technologies," in *Surface Electromyography: Physiology, Engineering, and Applications*, R. Merletti and D. Farina, Eds., John Wiley & Sons, Inc., 2016, pp. 540–560. DOI: `10.1002/9781119082934.ch20`.

---

[2]See `https://github.com/ime-luebeck/semgsim`.

[3]See `https://github.com/ime-luebeck/ecg-removal`.

[4]See `https://github.com/e-pet/mvdensity`.

[11] L. Gattinoni, J. J. Marini, and L. Camporota, "The respiratory drive: An overlooked tile of COVID-19 pathophysiology," *American Journal of Respiratory and Critical Care Medicine*, vol. 202, no. 8, pp. 1079–1080, 2020. DOI: `10.1164/rccm.202008-3142ed`.

[12] R. Giuliani and Y. Ranieri, "Proportional assist ventilation (PAV)," in *Applied Physiology in Respiratory Mechanics*, J. Milic-Emili, Ed. Springer, Milano, 1998. DOI: `10.1007/978-88-470-2928-6_19`.

[13] E. C. Goligher *et al.*, "Clinical strategies for implementing lung and diaphragm-protective ventilation: Avoiding insufficient and excessive effort," *Intensive Care Medicine*, vol. 46, no. 12, pp. 2314–2326, 2020. DOI: `10.1007/s00134-020-06288-9`.

[14] J. Graßhoff, E. Petersen, T. Becher, and P. Rostalski, "Automatic estimation of respiratory effort using esophageal pressure," in *41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Berlin, Germany, 2019. DOI: `10.1109/embc.2019.8856345`.

[15] J. Graßhoff, E. Petersen, and P. Rostalski, "Removing strong ECG interference from EMG measurements," in *Proceedings of the Workshop Biosignals*, Erfurt, Germany, 2018.

[16] J. Graßhoff *et al.*, "A template subtraction method for the removal of cardiogenic oscillations on esophageal pressure signals," in *39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Jeju Island, Korea, 2017. DOI: `10.1109/EMBC.2017.8037299`.

[17] J. Graßhoff *et al.*, "Surface EMG-based quantification of inspiratory effort: A quantitative comparison with Pes," *Critical Care (accepted)*, 2021.

[18] G. Hedenstierna, "Esophageal pressure: Benefit and limitations.," *Minerva anestesiologica*, vol. 78, pp. 959–966, 8 2012, ISSN: 1827-1596.

[19] C. Herzog, E. Petersen, and P. Rostalski, "Iterative approximate nonlinear inference via Gaussian message passing on factor graphs," *IEEE Control Systems Letters*, vol. 3, pp. 978–983, 4 2019. DOI: `10.1109/LCSYS.2019.2919260`.

[20] N. V. Iqbal, K. Subramaniam, and S. Asmi P., "A review on upper-limb myoelectric prosthetic control," *IETE Journal of Research*, vol. 64, no. 6, pp. 740–752, 2017. DOI: `10.1080/03772063.2017.1381047`.

[21] L. Kahl, E. Petersen, J. Graßhoff, and P. Rostalski, "Method and device for determining a respiratory or cardiogenic signal," DE112020000232A5 / WO2021063601A1, 2019.

[22] E. J. W. Maarsingh *et al.*, "Respiratory muscle activity measured with a noninvasive EMG technique: Technical aspects and reproducibility," *Journal of Applied Physiology: Respiratory, Environmental and Exercise Physiology*, vol. 88, no. 6, pp. 1955–1961, 2000. DOI: `10.1152/jappl.2000.88.6.1955`.

[23] F. Mojoli *et al.*, "In vivo calibration of esophageal pressure in the mechanically ventilated patient makes measurements reliable," *Critical Care*, vol. 20, no. 1, 2016. DOI: `10.1186/s13054-016-1278-5`.

[24] M. Olbrich, E. Petersen, C. Hoffmann, and P. Rostalski, "Sparse estimation for the assessment of muscular activity based on sEMG measurements," *IFAC-PapersOnLine*, vol. 51, no. 15, pp. 305–310, 2018. DOI: `10.1016/j.ifacol.2018.09.152`.

[25] E. Petersen, "A mathematical model of surface electromyographic measurements," in *Proceedings of the Workshop Biosignal Processing*, Berlin, 2016.

[26] E. Petersen, H. Buchner, M. Eger, and P. Rostalski, "Convolutive blind source separation of surface EMG measurements of the respiratory muscles," *Biomedical Engineering - Biomedizinische Technik*, vol. 62, no. 2, pp. 171–181, 2017. DOI: `10.1515/bmt-2016-0092`.

[27] E. Petersen, J. Graßhoff, M. Eger, and P. Rostalski, "Surface EMG-based estimation of breathing effort for neurally adjusted ventilation control," in *IFAC World Congress*, ser. IFAC-PapersOnLine, vol. 53, Berlin, Germany: Elsevier BV, 2020, pp. 16 323–16 328. DOI: 10.1016/j.ifacol.2020.12.654.

[28] E. Petersen, C. Hoffmann, and P. Rostalski, "On approximate nonlinear Gaussian message passing on factor graphs," in *IEEE Statistical Signal Processing Workshop*, Freiburg, Germany, 2018. DOI: 10.1109/SSP.2018.8450699.

[29] E. Petersen, L. Kahl, and P. Rostalski, "Electromyography as a tool for personalized rehabilitation," in *Annual Conference of the German Society for Biomedical Engineering (DGBMT)*, Aachen, Germany, 2018.

[30] E. Petersen and P. Rostalski, "Mathematical analysis of a model of intracellular action potential generation," in *22nd Congress of the International Society for Electrophysiology and Kinesiology (ISEK)*, Dublin, Ireland, 2018.

[31] ——, "Static nonlinear transformation of excitation model input as an alternative to feedback control in emg-force models," in *22nd Congress of the International Society for Electrophysiology and Kinesiology (ISEK)*, Dublin, Ireland, 2018.

[32] ——, "A comprehensive mathematical model of motor unit pool organization, surface electromyography, and force generation," *Frontiers in Physiology*, vol. 10, p. 176, 2019, ISSN: 1664-042X. DOI: 10.3389/fphys.2019.00176.

[33] E. Petersen, J. Sauer, J. Graßhoff, and P. Rostalski, "Removing cardiac artifacts from single-channel respiratory electromyograms," *IEEE Access*, vol. 8, pp. 30 905–30 917, 2020. DOI: 10.1109/access.2020.2972731.

[34] E. Petersen *et al.*, "Responsible and regulatory conform machine learning for medicine: A survey of challenges and solutions," *IEEE Access*, vol. 10, pp. 58 375–58 418, 2022. DOI: 10.1109/access.2022.3178382.

[35] C. Sinderby *et al.*, "Neural control of mechanical ventilation in respiratory failure," *Nature Medicine*, vol. 5, no. 12, pp. 1433–1436, 1999. DOI: 10.1038/71012.

[36] C. A. Sinderby, J. C. Beck, L. H. Lindström, and A. E. Grassino, "Enhancement of signal quality in esophageal recordings of diaphragm EMG," *Journal of Applied Physiology: Respiratory, Environmental and Exercise Physiology*, vol. 82, no. 4, pp. 1370–1377, 1997. DOI: 10.1152/jappl.1997.82.4.1370.

CHAPTER 2

PRELIMINARIES A: STATISTICAL INFERENCE

…the result was a proof that, if degrees of plausibility are represented by real numbers, then there is a uniquely determined set of quantitative rules for conducting inference. […] As a result, the imaginary distinction between 'probability theory' and 'statistical inference' disappears […] so it could be used for general problems of scientific inference, almost all of which arise out of incomplete information rather than 'randomness'.

(E. T. Jaynes, Probability theory: the logic of science)

For better or worse, real inference will remain a highly complicated, poorly understood phenomenon.

(E. E. Leamer, Specification searches: ad hoc inference with nonexperimental data)

Contents

THE READER may be surprised by the breadth and depth of the following summary of probability theory and statistical inference, given that this is, after all, a thesis in the field of biomedical engineering, and not in the field of statistics. This depth is prompted by the biomedical engineering problem to be solved. The core scientific questions that arise in the applications briefly described in chapter 1 are these:

- How can different processes be separated from one another when only their superposition is measured?

- How can robust parameter inference be performed in the face of model errors, outliers, and measurement noise?

- How do empirical data distributions influence statistical inference, and how can this information be exploited to achieve more accurate inference?

These are, of course, questions at the heart of the field of statistical inference. To appropriately justify the approach chosen in this thesis, it is, therefore, necessary to set the stage by introducing key concepts and paradigms (section 2.1), as well as a few well-known algorithms for statistical inference in both static (section 2.2) and dynamic systems (sections 2.3 and 2.4).

2.1 Elements of statistical inference

This section introduces some of the core elements of statistical inference theory. These include random variables and probability distributions (section 2.1.1), different paradigms and objective functions for performing statistical inference (section 2.1.2), and some properties of the resulting *estimators* (section 2.1.3). The genesis of statistical *models* and the distinction between a process of interest and statistical *noise* is discussed (section 2.1.4). Finally, brief introductions to regression under model *misspecification* (section 2.1.5) and *covariate shift* (section 2.1.6) are provided.

2.1.1 Random variables and probability distributions

A *random variable* is a quantity whose value is influenced by chance. In many instances, this is a simplifying modeling assumption: while the value of said quantity might be perfectly predictable given a sufficiently complex model of reality and full knowledge of the state of all relevant physical quantities, this is often not practically feasible. In this sense, randomness is often a modeling assumption that compensates for insufficient knowledge [16].

Random
Variable

Formally, a multivariate, real-valued, continuous random variable $\mathcal{X}$ can be defined as a measurable function

$$\mathcal{X} : \Omega \to \mathbb{R}^n$$

that maps from a set of possible outcomes Ω to $\mathbb{R}^n$. Furthermore defining a *probability measure* or *probability distribution*

Probability
Measure

$$P : \mathcal{F} \to [0, 1]$$

Probability

Distribution
Probability

on the $\sigma-$algebra $\mathcal{F}$ on Ω, such that P is countably additive and $P(\Omega) = 1$, we can then define the *probability* of $\mathcal{X}$ taking values in a Lebesgue-measurable set $A \in \mathcal{B}$ as

$$\Pr[\mathcal{X} \in A] = \int_{\mathcal{X}^{-1}(A)} \mathrm{d}P,$$

Probability
Density
Function

where $\mathcal{B}$ is the Borel σ-algebra on $\mathbb{R}^n$. A *probability density function* (PDF) then is any measurable function p with the property that

$$\int_{\mathcal{X}^{-1}(A)} \mathrm{d}P = \int_A p \, \mathrm{d}\mu \quad \forall \, A \in \mathcal{B},$$

where μ is the Lebesgue measure in n dimensions.[1] If A is not only Lebesgue-measurable but also Jordan-measurable, and p is Riemann-integrable on A, then the Lebesgue integral in section 2.1.1 coincides with the Riemann integral [19] and we obtain the formulation

$$\Pr[\mathcal{X} \in A] = \int_A p(x) \, \mathrm{d}x,$$

which is usually given in engineering textbooks [55]. Throughout this thesis, the notation

$$\mathcal{X} \sim P$$

means that the random variable $\mathcal{X}$ has the probability distribution P with probability density function $p(x)$.

Statistical
Moments

Random variables and probability distributions are often characterized by their *statistical moments*. The L-th order moment of a random variable $\mathcal{X}$ is given by

$$M_L = \mathrm{E}[\mathcal{X}^L],$$

Central
Moment

and its L-th order *central moment* by

$$M_L^c = \mathrm{E}[(\mathcal{X} - \mathrm{E}[\mathcal{X}])^L].$$

Expectation

Here, $\mathrm{E}[\cdot]$ denotes the *expectation* operator defined by the Lebesgue integral

$$\mathrm{E}[\mathcal{X}] = \int_\Omega \mathcal{X}(\omega) \, \mathrm{d}P(\omega)$$

which, under the regularity assumptions stated above, is equivalent to the Riemann integral

$$\mathrm{E}[\mathcal{X}] = \int_{\mathbb{R}^n} x \, p(x) \, \mathrm{d}x.$$

When it is not clear from the context, we will sometimes write, e.g.,

$$\mathrm{E}_Q[\mathcal{X}] = \int_{\mathbb{R}^n} x \, q(x) \, \mathrm{d}x$$

---

[1] Equivalently, the probability density function can also be defined as a *Radon-Nikodym derivative* with respect to the Lebesgue measure μ.

to clarify that the expectation is taken with respect to the probability distribution Q. The first-order moment ($L = 1$) is called the *expected value* of the random variable, and the second-order central moment is called its *variance*.

<div style="float:right">Expected Value

Variance</div>

In most practical scenarios, the true probability distribution P is unknown and one only has access to a set of realizations

$$x_k = \mathcal{X}(\omega_k), \quad \omega_k \in \Omega, \quad k = 1, \dots, N^S$$

drawn from this distribution. In this case, a natural definition of the *empirical measure* is

<div style="float:right">Empirical
Measure</div>

$$P_{N^S}(A) = \frac{1}{N^S} \sum_{k=1}^{N^S} \delta_{x_k}(A),$$

where δ_{x_k} denotes the Dirac measure centered on realization x_k. Equivalently, the probability density function of the empirical measure is given by

$$p_{N^S}(x) = \frac{1}{N^S} \sum_{k=1}^{N^S} \delta_{x_k}(x).$$

By the *Glivenko-Cantelli theorem*, P_{N^S} converges uniformly to the correct underlying distribution P almost surely [10]. Moreover, by the *strong law of large numbers*, the expectation of the empirical distribution

<div style="float:right">Glivenko-
Cantelli
Theorem

Strong Law of
Large Numbers</div>

$$\mathrm{E}_{P_{NS}}[\mathcal{X}] = \int_{\mathbb{R}^n} x \, p_{N^S}(x) \, \mathrm{d}x = \frac{1}{N^S} \sum_{k=1}^{N^S} x_k$$

converges almost surely to the expected value $\mathrm{E}_P[\mathcal{X}]$ of the true distribution as $N^S \to \infty$ [10]. The quantity $\mathrm{E}_{P_{NS}}[\mathcal{X}]$ is often called the *sample mean*. Note that the sample mean is itself a random variable since it is influenced by the random sampling of the realizations x_k of $\mathcal{X}$.

<div style="float:right">Sample Mean</div>

2.1.2 Paradigms for statistical inference

Over the last century, many different theoretical frameworks for formulating statistical inference problems have been proposed. As this thesis is ripe with references to many of these different inference paradigms, they deserve a proper introduction, which will be the subject of this section. Particular emphasis is placed on illuminating the relationships and partial equivalences between the different paradigms.

To set the stage for statistical inference, throughout this thesis we assume that the observations y are drawn from a *true* probability distribution, i.e.,

$$y \sim p(y) = p(y \mid x),$$

where $p(y \mid x)$ specifies the relationship between the covariates and the target variable, and the covariates x are assumed to be fixed, i.e., *not* randomly distributed. Moreover, we often assume a probabilistic *regression model* to be given, which is nothing else but a a parameterized family of probability density functions $q(y \mid X, \theta)$. If there is a parameter vector θ^* for which

<div style="float:right">Regression
Model</div>

$$p(y \mid x) \equiv q(y \mid X, \theta^*) \quad \forall \, y, x$$

the regression model is said to be *correctly specified*. Otherwise, *model mismatch* is present,[2] meaning that the model does not described reality correctly. We will return to this issue in detail in section 2.1.5.

Maximum Likelihood Estimation

Arguably, the most "natural" paradigm for statistical inference is *maximum likelihood* (ML) estimation, which aims to find the parameters $\hat{\theta}_{ML}$ that maximize the likelihood of observing the data as they were observed. The ML solution to the parameter inference problem is then given by

$$\hat{\theta}_{ML} = \arg\max_{\theta} q(y \mid X, \theta). \tag{2.1}$$

If individual samples y_k are assumed to be statistically independent, then we have

$$q(y \mid X, \theta) = \prod_{k=1}^{N^S} q_k(y_k \mid x_k, \theta), \tag{2.2}$$

where $pdf[q_k]y_k x_k, \theta$ denotes the likelihood (under the assumed model with parameters θ) of observing y_k given that x_k was observed. By further noting that

$$\hat{\theta}_{ML} = \arg\max_{\theta} q(y \mid X, \theta) = \arg\max_{\theta} \log q(y \mid X, \theta),$$

we obtain the well-known formulation

$$\hat{\theta}_{ML} = \arg\max_{\theta} \log \prod_{k=1}^{N^S} q_k(y_k \mid x_k, \theta) = \arg\max_{\theta} \sum_{k=1}^{N^S} \log q_k(y_k \mid x_k, \theta)$$

$$= \arg\min_{\theta} - \sum_{k=1}^{N^S} \log q_k(y_k \mid x_k, \theta) \tag{2.3}$$

in terms of the sum of the log-likelihood of individual samples.

Maximum A Posteriori Estimation

A very close relative of ML estimation is *maximum a posteriori* (MAP) estimation. Here, the idea is to find the set of parameters $\hat{\theta}_{MAP}$ which maximizes the posterior probability $p(\theta \mid y, x)$, in other words, the most *likely* parameter vector. Using Bayes' rule and (for the last step) assuming statistical independence between individual samples, one obtains

$$\hat{\theta}_{MAP} = \arg\max_{\theta} p(\theta \mid y, X) = \arg\max_{\theta} \frac{p(y \mid X, \theta)\, p(\theta)}{p(y \mid X)}$$

$$= \arg\max_{\theta} q(y \mid X, \theta)\, p(\theta) = \arg\max_{\theta} \log q(y \mid X, \theta)\, p(\theta)$$

$$= \arg\max_{\theta} \log p(\theta) + \sum_{k=1}^{N^S} \log q_k(y_k \mid x_k, \theta), \tag{2.4}$$

which is identical to the ML estimate, except for the regularizing term $\log p(\theta)$, which represents the impact of prior knowledge about the likelihood of different parameter vectors on the resulting parameter estimate. If a uniform distribution is assumed for the prior $p(\theta)$, the MAP solution coincides with the ML solution.

---

[2]This problem has many names. *Model bias*, *model error* and *model misspecification* are a few of them.

So far, we always assumed the regression model to define a family of probability density functions. In other words, we assumed a *probabilistic* regression model. This is not a necessary condition for performing statistical inference, though. Consider a parametric class of deterministic functions $\overline{f}(\theta)$ that simply map *regressors* X to *predictions* $\overline{f}(X, \theta)$, i.e.,

$$X \mapsto \overline{f}(X, \theta) \in \mathbb{R}^{N^S}.$$

Then, inference can also be formulated as wanting to minimize some metric $\overline{\ell}(y, \overline{f}(X, \theta))$ of the discrepancy between the observed outputs y and the model outputs $\overline{f}(X, \theta)$ by choosing the model parameters θ appropriately. In the following, statistical independence between individual samples will always be a model assumption, and we will assume the same model $f(\theta)$ for all samples:

$$\overline{f}(X, \theta) = \begin{bmatrix} f(x_1, \theta) & \cdots & f(x_{N^S}, \theta) \end{bmatrix}^{\mathsf{T}}.$$

Similarly, we will use a per-sample loss function $\ell(y_k, f(x_k, \theta))$ and assume that

$$\overline{\ell}\left(y, \overline{f}(X, \theta)\right) = \sum_{k=1}^{N^S} \ell(y_k, f(x_k, \theta)).$$

Under these assumptions, the overall loss can be rewritten in terms of the empirical expectation of the per-sample loss as

$$\overline{\ell}\left(y, \overline{f}(X, \theta)\right) = \sum_{k=1}^{N^S} \ell(y_k, f(x_k, \theta)) = N^S \cdot \mathrm{E}_{\mathrm{P}_{\mathrm{N}^S}}[\ell(y, f(x, \theta))],$$

thereby providing a statistical interpretation despite the use of a purely deterministic regression model.

Known since the eighteenth century [21, 56], *least-squares* (LS) estimation is possibly the most famous paradigm for parameter inference of all. The approach is simple: minimize the sum — or, equivalently, the average — of the squared distance between model predictions and observations. The least-squares loss function is given by

Least-Squares Estimation

$$\overline{\ell}\left(y, \overline{f}(X, \theta)\right) = \|y - \overline{f}(X, \theta)\|_2^2 = \sum_{k=1}^{N^S} (y_k - f(x_k, \theta))^2 = N^S \cdot \mathrm{E}_{\mathrm{P}_{\mathrm{N}^S}}[(y - f(x, \theta))^2]$$

and the LS estimator is, therefore, given by

$$\hat{\theta}_{\mathrm{LS}} = \arg\min_{\theta} \mathrm{E}_{\mathrm{P}_{\mathrm{N}^S}}[(y - f(x, \theta))^2].$$

Least-squares estimation has proven a tremendously successful paradigm with vast numbers of applications throughout science and engineering [56]. It is identical to ML estimation under the assumption of a deterministic regression function and a Gaussian error distribution, i.e.,

$$q_k(y_k \mid x_k, \theta, \sigma_\varepsilon^2) = \mathcal{N}(f(x_k, \theta), \sigma_\varepsilon^2).$$

In that case,

$$\hat{\theta}_{\mathrm{ML}}, \hat{\sigma}_\varepsilon = \underset{\theta, \sigma_\varepsilon}{\arg\min} - \sum_{k=1}^{N^S} \log q_k(y_k \mid x_k, \theta, \sigma_\varepsilon^2)$$

$$= \underset{\theta, \sigma_\varepsilon}{\arg\min} - \sum_{k=1}^{N^S} \log \frac{1}{\sqrt{2\pi\sigma_\varepsilon^2}} e^{-\frac{1}{2\sigma_\varepsilon^2}(y_k - f(x_k, \theta))^2}$$

$$= \underset{\theta, \sigma_\varepsilon}{\arg\min} \sum_{k=1}^{N^S} \frac{1}{2\sigma_\varepsilon^2}(y_k - f(x_k, \theta))^2 + \frac{N^S}{2} \log 2\pi\sigma_\varepsilon^2,$$

and since the optimization with respect to the regression parameters θ can be carried out independent of the value of σ_ε, it follows that[3]

$$\hat{\theta}_{\mathrm{ML}} = \hat{\theta}_{\mathrm{LS}}$$

in this special case. Similarly, it can be demonstrated that (regularized) LS estimation coincides with other probabilistic estimates such as MAP or Wasserstein distributionally robust estimates [36]. These results, connecting a somewhat arbitrary optimization approach with a probabilistic inference perspective, further legitimize the empirically successful LS paradigm.

Squared distance minimization underlies not only least-squares estimation but also **Minimum Mean Squared Error Estimation** *minimum mean squared error* (MMSE) estimation. The deciding difference between the two is that in MMSE estimation, the expected value of the squared errors in the *parameters* is minimized, whereas an LS estimator minimizes the expected value of the squared errors in the *predictions*. We define

$$\hat{\theta}_{\mathrm{MMSE}} = \underset{\theta}{\arg\min}\, \mathrm{E}[(\theta^* - \theta)^2 \mid y, x] = \underset{\theta}{\arg\min} \int (\theta^* - \theta)^2\, p(\theta \mid y, x)\, \mathrm{d}\theta,$$

and with a short derivation, one obtains [45]

$$\hat{\theta}_{\mathrm{MMSE}} = \mathrm{E}[\theta \mid y, x].$$

Thus, the MMSE estimator is given by the mean of the posterior parameter distribution. Note the difference to the MAP estimator above: while the MMSE estimator is given by the *mean* of the posterior distribution, the MAP estimator is given by its *mode*.

Empirical Risk Minimization *Empirical risk minimization* (ERM), popularized by Vapnik and Chervonenkis [67], is a particularly general statistical inference framework. The goal of *ideal* risk minimization is to minimize the expected value of some per-sample loss function $\ell(y, f(x, \theta))$. This quantity **Risk Functional** is called the *risk functional*

$$\mathcal{R}(\theta) = \mathrm{E}[\ell(y, f(x, \theta))] = \int \ell(y, f(x, \theta))\, p(y, x)\, \mathrm{d}y\, \mathrm{d}x, \qquad (2.5)$$

---

[3]Using the estimate $\hat{\theta}_{\mathrm{ML}} = \hat{\theta}_{\mathrm{LS}}$ for the regression parameters, the optimization with respect to the equation error covariance can then be carried out by calculating the gradient of the resulting expression and setting it to zero. The result is the (asymptotically unbiased) maximum-likelihood error covariance estimate $\hat{\sigma}_\varepsilon^2 = \frac{1}{N^S} \sum_{k=1}^{N^S} (y_k - f(x_k, \hat{\theta}_{\mathrm{LS}}))^2$.

and the associated parameter estimate is given by

$$\hat{\theta}_{\text{RM}} = \arg\min_{\theta} \mathcal{R}(\theta). \tag{2.6}$$

Since the true data-generating PDF $p(y, x) = p(y \mid x)\, p(x)$ is unknown in practice, this minimization problem is infeasible. Therefore, one replaces the expectation with respect to this unknown distribution by the empirical expectation taken over the available data samples. We thus define the *empirical risk functional*

Empirical Risk Functional

$$\mathcal{R}_{\text{emp}}(\theta) = \mathrm{E}_{\mathrm{P}_{\mathrm{N}^S}}[\ell(y, f(x, \theta))] = \frac{1}{N^S} \sum_{k=1}^{N^S} \ell(y_k, f(x_k, \theta)).$$

Statistical inference is then performed by solving the optimization problem

$$\hat{\theta}_{\text{ERM}} = \arg\min_{\theta} \mathcal{R}_{\text{emp}}(\theta),$$

possibly subject to additional constraints on the parameters. Many classical estimation paradigms can be formulated within the ERM framework, including (very obviously) least-squares estimation and maximum likelihood estimation [66, 68]. There is, of course, also a regularized version of ERM, which is called *structural risk minimization* (SRM) [23, 66, 67]. The SRM inference problem is stated as

Structural Risk Minimization

$$\hat{\theta}_{\text{SRM}} = \arg\min_{\theta} \mathcal{R}_{\text{emp}}(\theta) + \lambda\, \ell_{\text{reg}}(\theta)$$

$$= \arg\min_{\theta} \frac{1}{N^S} \sum_{k=1}^{N^S} \ell(y_k, f(x_k, \theta)) + \lambda\, \ell_{\text{reg}}(\theta),$$

where $\ell_{\text{reg}}(\theta)$ is a regularization term that penalizes overly complex models, and λ is a constant that controls the balance between minizing the expected empirical risk and minimizing the complexity of the selected model (more on that in the following section).

Many of the discussed paradigms coincide under specific assumptions on the involved distributions or the regression model. One particularly prominent example of this is the ordinary least squares (OLS) setting, which will be discussed in detail in section 2.2.1, and in which the ML, MMSE, and LS estimates all coincide. Among many others, Luxburg and Schölkopf [39] provide a comprehensive overview of many of the estimation schemes discussed above, taking a statistical learning theory perspective.

As another remark, all of the above estimates are *point estimates* which solve an optimization problem to obtain an in some sense ideal parameter estimate. By contrast, in *fully Bayesian inference* the full posterior distribution $p(\theta \mid y, X)$ is the target of the inference procedure. Fully Bayesian inference provides a much more complete picture of the validity of the resulting estimates, at the cost of increased computational and algorithmic complexity. Markov chain Monte Carlo (MCMC) methods are the classical tool for performing exact fully Bayesian inference [3]; popular approximative algorithms include variational inference [72] and expectation propagation [41]. Of course, if one takes the mode of a posterior distribution resulting from a fully Bayesian inference algorithm, this should be — depending on the accuracy of the employed inference algorithm — again equivalent to the MAP estimator.

Point Estimates

Fully Bayesian Inference

Finally, at peril of stating the obvious, observe that *all* of the inference paradigms introduced above formulate statistical inference as an optimization problem. Every inference algorithm based on these paradigms is, at its heart, an optimization algorithm, whence there is an intimate link between statistical inference and optimization theory.

2.1.3 Properties of statistical estimators and the bias-variance trade-off

Statistical estimators (such as the ones introduced in the previous section) are often discussed and compared in terms of various properties; two of those will be relevant in the context of this thesis: *consistency* and *unbiasedness*. An estimator $\hat{\theta}$ of a quantity θ^* is said to be *consistent* if it converges in probability to the true parameter vector as the number N^S of available samples increases indefinitely, i.e.,

Consistency

$$p\left(|\hat{\theta}_{N^S} - \theta^*| > \delta\right) \to 0 \quad \text{as} \quad N^S \to 0 \quad \forall\, \delta > 0$$

or

$$\text{plim}(\hat{\theta}_{N^S}) = \theta^*.$$

Unbiasedness

As opposed to this asymptotic property, *unbiasedness* is a finite-sample property: An estimator is unbiased if

$$\mathrm{E}[\hat{\theta} - \theta^*] = 0,$$

regardless of the number of available samples. Note that neither of the two properties implies the other in the absence of further assumptions.

At first glance, one might believe unbiased estimators to be uniformly superior to biased estimators, but this is not the case. To understand this, it is helpful to consider the well-known bias-variance decomposition of the mean squared parameter estimation error. In the following, all expectations and the variance are taken over the true, data-generating distribution $p(y, X)$, and $\mu_{\hat{\theta}} := \mathrm{E}[\hat{\theta}]$. We have [42]

$$
\begin{aligned}
\mathrm{E}[(\hat{\theta} - \theta^*)^\mathsf{T}(\hat{\theta} - \theta^*)] &= \mathrm{E}\left[\left((\hat{\theta} - \mu_{\hat{\theta}}) - (\theta^* - \mu_{\hat{\theta}})\right)^\mathsf{T}\left((\hat{\theta} - \mu_{\hat{\theta}}) - (\theta^* - \mu_{\hat{\theta}})\right)\right] \\
&= \mathrm{E}\left[\left(\hat{\theta} - \mu_{\hat{\theta}}\right)^\mathsf{T}\left(\hat{\theta} - \mu_{\hat{\theta}}\right)\right] - 2\,\mathrm{E}\left[\hat{\theta} - \mu_{\hat{\theta}}\right]^\mathsf{T}(\theta^* - \mu_{\hat{\theta}}) + \mathrm{E}\left[\left(\theta^* - \mu_{\hat{\theta}}\right)^\mathsf{T}\left(\theta^* - \mu_{\hat{\theta}}\right)\right] \\
&= \mathrm{E}\left[\left(\hat{\theta} - \mu_{\hat{\theta}}\right)^\mathsf{T}\left(\hat{\theta} - \mu_{\hat{\theta}}\right)\right] + \mathrm{E}\left[\left(\theta^* - \mu_{\hat{\theta}}\right)^\mathsf{T}\left(\theta^* - \mu_{\hat{\theta}}\right)\right] \\
&= \mathrm{Var}[\hat{\theta}] + \mathrm{Bias}^2[\hat{\theta}]. \tag{2.7}
\end{aligned}
$$

Thus, it becomes apparent that a given estimator's mean squared parameter estimation error can always be decomposed into a variance-related term and a bias-related term. As a corollary, a low-variance but biased estimator may, in principle, attain a lower mean squared error than an unbiased but high-variance estimator. The problem of finding a model that strikes the right balance between bias and variance to yield optimal average performance is known as the *bias-variance trade-off*. Two essential means for attaining this optimal balance are the inclusion of prior information and regularization, both of which are discussed in the following section. Both introduce a bias into the estimation procedure while simultaneously reducing the variance of the estimator. (A classic example of a biased estimator that achieves uniformly better performance in terms of the mean squared error than the ML estimator

Bias-variance Trade-off

in the same setting (without the use of prior knowledge) is the *James-Stein estimator* [32].) Finally, without going into much detail, notice that the bias-variance *decomposition* (2.7) does *not* imply the existence of a bias-variance *trade-off*: the decomposition holds *for a given estimator*, but it does not imply that, e.g., modifying that estimator to reduce the bias-related error (say, by increasing the model order) necessarily comes at the cost of increased variance-related error. While bias-variance trade-offs have been demonstrated under specific circumstances, they are no universal property of statistical estimators, and settings have been observed (most notably: deep neural networks) in which both bias and variance can be decreased at the same time by increasing model order [43].

James-Stein Estimator

2.1.4 Model genesis: process and noise, priors and regularization

Until this point, we always assumed that one of the protagonists of statistical inference was simply given to us: the *model*. Given a parametric model, our only task was to estimate parameter values that are — in some sense — optimal. But where does the model come from?

Ideally, a model should represent the synthesis of everything we know about the process under observation *without taking the observations into account that will be used to identify the model's parameters*. This prior knowledge comes in two flavors:

1. *structural knowledge* about the number and kind of different quantities that capture the behavior of the process (internal state variables and observations), and about the ways in which these quantities can and cannot interact, and

2. knowledge about the likely values of the model parameters.

Both types of prior knowledge may be deterministic or probabilistic in nature. As an example, we may *know for a fact* that two quantities A and B do not interact with each other, therefore incorporating this fact into our model, or we may *believe* that they do not interact with 95% certainty, yet still be open to the 5% chance that they, in fact, *do* interact. In the latter case, we might use a Bayesian model averaging technique [15, 29].[4] In the case of parameter knowledge, we might know that a parameter definitely cannot be negative or that it must be between 0 and 1. Alternatively, we might only have a rough idea that it lies somewhere around 35, but could also be significantly smaller or larger. Both kinds of beliefs can be encoded by a prior probability distribution in a Bayesian model, corresponding to the prior $p(\theta)$ in the MAP estimator

$$\hat{\theta}_{\mathrm{MAP}} = \arg\max_{\theta} \log p(\theta) + \sum_{k=1}^{N^S} \log q_k(y_k \mid x_k, \theta),$$

reproduced from eq. (2.4) above.

---

[4]Bayesian model averaging should also be employed in the common case in which there is no a priori knowledge about the class of models that might describe the observed data best. The commonly used method — evaluate model performance empirically, and perform statistical inference with the best-performing model family — incorrectly discards uncertainty about the correct model family.

In section 2.1.2, we already pointed out that the MAP optimization problem is identical to the ML optimization from eq. (2.3),

$$\hat{\theta}_{\text{ML}} = \arg\min_{\theta} -\sum_{k=1}^{N^S} \log q_k(y_k \mid x_k, \theta),$$

Regularization except for the additional, regularizing term $\log p(\theta)$ in the cost. (By *regularization*, one usually means adding an extra term $\ell_{\text{reg}}(\theta)$ to the cost function, which depends on the model parameters θ but *not* on the data.) Adding such a term biases the estimation procedure in some direction that is considered plausible by the modeler and may yield a better bias-variance trade-off (see the previous section). Many different regularization schemes have been proposed in the literature for varying reasons, and — as with inference paradigms — it turns out that many of them coincide under specific assumptions. Returning to the MAP regularization term $\log p(\theta)$ and assuming, for the sake of example, a multivariate Gaussian prior[5]

$$p(\theta) \sim \mathcal{N}(\mu_\theta, \Sigma_\theta)$$

with mean $\mu_\theta = 0$ and covariance matrix Σ_θ, we obtain

$$\log p(\theta) = \log \frac{1}{(2\pi)^{m/2}|\Sigma_\theta|} e^{-\frac{1}{2}\theta^\mathsf{T}\Sigma_\theta^{-1}\theta}$$

$$= -\frac{1}{2}\theta^\mathsf{T}\Sigma_\theta^{-1}\theta + \log \frac{1}{\sqrt{(2\pi)^m|\Sigma_\theta|}}$$

$$= -\frac{1}{2}\|\theta\|^2_{\Sigma_\theta^{-1}} + \text{const}, \tag{2.8}$$

Tikhonov Regularization where the constant does not depend on the model parameters θ and is, therefore, irrelevant for the optimization. Equation (2.8) is known as a *Tikhonov regularization* term, whence we have demonstrated the equivalence of a zero-mean Gaussian prior with covariance matrix Σ_θ on the parameters and Thikonov regularization with kernel matrix Σ_θ^{-1}. If we assume a homoscedastic covariance matrix $\Sigma_\theta = \frac{1}{\lambda}I$ (I being the identity matrix), we obtain

$$\log p(\theta) = -\frac{\lambda}{2}\|\theta\|^2_2 + \text{const},$$

Ridge Regularization a special case which is known as *ridge regularization*. If one uses an empirical Bayes method for estimating the regularization constant λ, the resulting estimate is equivalent to the classical James-Stein estimator [9, 32]. An analogous derivation demonstrates that **LASSO Regularization** $\mathcal{L}_1$-norm regularization, often called *LASSO regularization* (least absolute shrinkage and selection operator) and widely used for its sparsity-inducing properties, is equivalent to assuming a zero-mean Laplacian prior [42]. To summarize, the choice of a regularization scheme, although sometimes an afterthought to model selection, is an integral part of the choice of a model family.

Once again returning to eq. (2.4),

$$\hat{\theta}_{\text{MAP}} = \arg\max_{\theta} \log p(\theta) + \sum_{k=1}^{N^S} \log q_k(y_k \mid x_k, \theta),$$

---

[5]Refer to appendix 2 for some details on the multivariate Gaussian distribution.
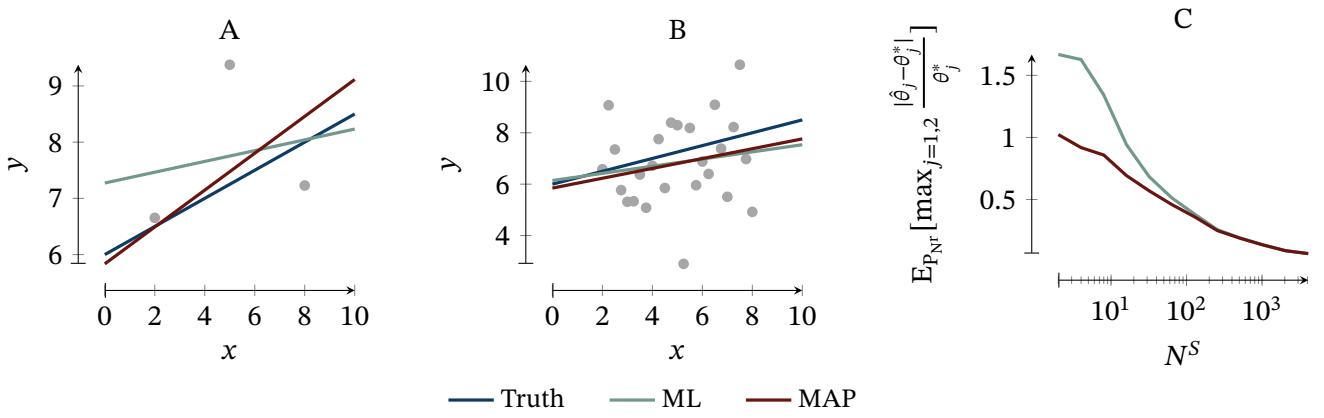
Figure 2.1 – Example illustrating the difference between MAP and ML estimates. The estimation task is to identify θ_1 and θ_2 in $y = \theta_1 x + \theta_2$, given measurements disturbed by Gaussian noise $\varepsilon \sim \mathcal{N}(0, \Sigma_\varepsilon)$. The MAP estimator uses a reasonably accurate guess of the true linear relation. In **panel A**, just three samples are used, and the MAP estimator is significantly closer to the true parameter values than the ML estimator. In **panel B**, 25 samples are used, and the two estimates are already almost indistinguishable. **Panel C** shows the average of the maximum relative parameter estimation error for the two estimators as a function of the sample size N^S, calculated over $N^r = 1000$ realizations.

it is evident that both the prior $p(\theta)$ and the data (x_k, y_k) influence the final parameter estimate $\hat{\theta}_{\mathrm{MAP}}$. If we assume that the collected data are *informative* in the sense that they provide some minimal amount of information about the parameter values of interest, it follows that as the amount of collected data increases, i.e., as N^S grows, the influence of the prior diminishes. This phenomenon is illustrated in fig. 2.1, in which MAP and ML estimates in a simple linear regression problem are compared. It can be observed that the inclusion of (reasonably correct) prior knowledge about the parameters in the MAP estimator leads to better performance in the small sample size regime, whereas for large sample sizes, the MAP and ML estimates converge. This behavior is an instance of a general principle: It has been shown that in many settings, the influence of the prior diminishes as more data become available, and the MAP estimate converges towards the ML estimate [14, 42, 64].[6]

A crucial distinction to be made when constructing a model is between the process of interest and "the rest". Which information contained in the data is considered a reflection of the process of interest, and which information is considered "noise"? Notice that nothing differentiates "noise" from "the process" at the physical level: both are consequences of fundamental physical processes. Yet, for some non-physical reason, one is interested in separating one from the other to obtain a truthful representation of some abstract, high-level "process" one is interested in. A practical definition of "noise" then is this: any deviation from the assumed (deterministic) process model is considered noise.[7] In mathematical terms, we

---

[6]This is a consequence of the Bernstein–von Mises–theorem [64, chapter 10]. Unfortunately, there are counter-examples in which the assumptions of the theorem do not hold — see, e.g., Diaconis and Freedman [14].

[7]This formulation can, of course, be extended to include probabilistic process models as well. For reasons of simplicity, the discussion is limited to deterministic models here, where the complete model only becomes probabilistic by including a probabilistic noise model.

have

$$y_k = f(x_k, \theta) + \varepsilon_k, \tag{2.9}$$

where $\varepsilon \sim p(\varepsilon \mid \theta)$ denotes the (probabilistic) noise component, which may or may not depend on the model parameters θ. To achieve good separation between process and noise based on the available measurements (y_k, x_k) — *the* core challenge of statistical inference — it is therefore not sufficient to create an adequate model $f(X, \theta)$ of the process: an adequate model of the noise distribution $p(\varepsilon \mid \theta)$ is equally essential, and often just as hard or even harder to find.

The ubiquitous noise model in practice is, of course, the normality assumption

$$\varepsilon_k \sim \mathcal{N}(0, \sigma_\varepsilon^2).$$

Central Limit Theorem
The reasons are twofold. Firstly, the *central limit theorem* famously states that a sum of many i.i.d. random variables asymptotically follows a normal distribution, regardless of the distribution of the individual RVs.[8] Accordingly, measurement noise, which is the result of a large number of individual physical processes, is often assumed and experimentally observed to follow a normal distribution. Secondly, and just as importantly, the normality assumption has proved to facilitate the derivation of very efficient inference algorithms, many of which break down once a different distribution is assumed for the noise [56]. Nevertheless, other noise distributions can and are, of course, being assumed by statisticians, informed by domain knowledge about the characteristics of the disturbance.

Model Error
One particular and often–neglected component of the noise signal is the *model error*. If, as defined by eq. (2.9), every deviation from the assumed process model is considered noise, then this must necessarily also include errors resulting from a discrepancy between the process model and reality. Since this particular source of disturbance plays a vital role in this thesis, the following section is devoted to this challenge exclusively.

2.1.5 Model misspecification

A regression model, which as described in the previous section may be composed of a deterministic process model $f(x, \theta)$ and a probabilistic noise model $q(\varepsilon_k \mid \theta)$, is nothing else but a parameterized family of probability distributions $q(y \mid X, \theta)$. Reiterating what was briefly mentioned in section 2.1.2, a regression model is said to be *correctly specified* if there is a parameter vector θ^* such that

Correctly Specified Model

$$q(y \mid X, \theta^*) \equiv p(y \mid X) \quad \forall X,$$

where one assumes that the observations y are drawn from a *true* probability distribution P, i.e.,

$$y \sim p(y) = p(y \mid X). \tag{2.10}$$

Note that, as discussed in the previous section, *both* the process model $f(x, \theta)$ *and* the noise model $q(\varepsilon_k \mid \theta)$ need to be correct in order for the combined model to be correctly specified. Given the obvious violation of the assumption of correct model specification in almost any practical application, what can one say about the properties of the usual estimators in the face of model misspecification? Can any performance bounds or guarantees still be retained?

---

[8]The central limit theorem is another consequence of the *strong law of large numbers* [10].

The answer is, fortunately, yes. It can be shown [2, 76] that under mild regularity conditions on the involved distributions, a maximum likelihood estimator (MLE) — such as OLS — converges almost surely to the parameter vector

$$\hat{\theta}_{\mathrm{KL}} = \arg\min_{\theta} \mathrm{D}_{\mathrm{KL}}(P \parallel Q(\theta))$$

minimizing the *Kullback-Leibler (KL) divergence*

<div style="float:right">Kullback-Leibler Divergence</div>

$$\mathrm{D}_{\mathrm{KL}}(P \parallel Q(\theta)) = \int \log\left(\frac{\mathrm{d}P}{\mathrm{d}Q}\right) \mathrm{d}P \tag{2.11}$$

$$= \int p(y \mid X) \log\left(\frac{p(y \mid X)}{q(y \mid X; \theta)}\right) \mathrm{d}y$$

between the true distribution P and the identified model Q. So, while the true distribution may not be achieved, an ML estimator does minimize the discrepancy between reality and model assumption.[9] Nevertheless, model misspecification represents a significant challenge in many application scenarios. Depending on the (assumed) source of the model mismatch, various strategies have been put forth for solving the challenge.

<div style="float:right">Outliers</div>

Outliers are, by definition, samples that significantly deviate from the assumed model of reality. Given a probabilistic model, these samples are extremely unlikely to occur, yet for some reason, they are observed more frequently than would be expected. They are, thus, a classic example of model misspecification. One way of handling outliers is to adjust the noise model to encompass the outliers' occurrence. The use of a Laplace distribution assumption for the noise model is a common choice that is equivalent to minimizing the mean absolute deviation from the deterministic process model (instead of the mean squared deviation in the case of the normality assumption). Another option specifically tailored to the modeling of outliers is the *normal with unknown variance* (NUV) noise model [38, 71], in which the noise signal is described as a normally distributed signal, the variance of which is occasionally (namely, when an outlier occurs) significantly increased. The occurrence of outliers is determined by means of an *expectation maximization* (EM) algorithm. Both of these options aim at reducing the amount of model misspecification by including outliers in the model itself. An alternative (and closely related) approach is to limit the influence of outliers on the resulting estimates by modifying the cost function. This is the approach taken by classical *robust estimators*, such as M-estimators [30], which abandon the squared cost function in favor of cost functions less sensitive to extreme values of the residuals ε_k. For some of these more outlier-robust cost functions, there is a probabilistic interpretation, e.g., as discussed above, the choice

<div style="float:right">Robust Estimators</div>

$$\ell(y_k, f(x_k, \theta)) = |y_k - f(x_k, \theta)| = |\varepsilon_k|$$

is equivalent to assuming a Laplacian model for the noise distribution $p(\varepsilon_k)$.

Another frequent source of model misspecification is the fact that in the usual statistical models, the covariates X are assumed to be known *exactly*. Since these are often the results

---

[9]Unfortunately, all the usual tests for statistical power and significance, as well as uncertainty measures such as the parameter standard deviation, break down once the model is misspecified. See, e.g., Berk *et al.* [8] and Freedman [18] for a detailed discussion of this topic and alternative procedures for obtaining valid uncertainty estimates. Here, we will ignore this issue and concern ourselves exclusively with the validity of the point estimate itself.

Errors-in-Variables Modeling

of empirical measurements, this assumption is frequently violated. As a remedy, *errors-in-variables* (EIV) versions of many of the classical models have been proposed in the literature [22, 57, 65]. These models assume that

$$x_k = x_k^* + \xi_k \tag{2.12}$$

where x_k^* is the latent (true) value of the covariates for sample k and ξ_k is the measurement error in these signals. Errors in the covariates can lead to significantly biased estimates if they are not considered in the estimation procedure; see, e.g., Wansbeek and Meijer [75] for a derivation of the exact bias when neglecting covariate errors in multivariate linear regression. Correcting this bias is not trivial since the (relative) variances of the measurement and equation errors are often unknown and need to be estimated from the data. For this reason, assuming an EIV model usually renders the inference procedure significantly more complex.

Overfitting

Overfitting is another interesting case of model mismatch. At first glance, one might believe that using a high-dimensional — or "nonparametric" — regression model such as a deep neural network, a Dirichlet process model, or a Gaussian process model would solve the model mismatch problem. This is generally not the case. While the regression model itself might be flexible enough to describe the observed process, what is almost always lacking is an accurate representation of the noise: what are all the sources of disturbances to the process itself and measurements thereof, and what is their relative importance? So long as this crucial model component is not specified correctly, which is virtually never the case, even an infinite-dimensional regression model will not allow perfect identification of the observed process because the line between what is part of the process and what is part of the noise will not be drawn correctly. Despite a highly flexible process model, the *complete* model — process model and noise model in conjunction — is still misspecified.

The overfitting problem is, of course, well-known and routinely addressed in current machine learning applications by means of various regularization schemes. As discussed in the previous section, these are just another way of modifying the regression model. Finding a combination of a high-dimensional regression model and an automatic regularization scheme that leads to regression results that generalize well — i.e., that avoid overfitting — is still a largely unsolved problem. For deep neural networks, in particular, recent applications indicate remarkably good generalization capabilities, but the reason for this unexpectedly good generalization is currently unclear [31, 80]. It has been hypothesized that both the particular structure of these networks as well as the stochastic gradient descent algorithm used for the parameter identification step may contribute to these models' good generalization capabilities [25, 44].

One situation in which model misspecification may lead to seriously impaired estimation performance is when there is *covariate shift*. This situation, and a possible remedy, are the subject of the following section.

2.1.6 Regression under covariate shift

Recall the risk minimization (RM) formulation of eq. (2.5), (2.6):

$$\hat{\theta}_{\mathrm{RM}} = \arg\min_\theta \mathrm{E}_{p_{\mathrm{train}}}[\ell(y, f(x, \theta))]$$
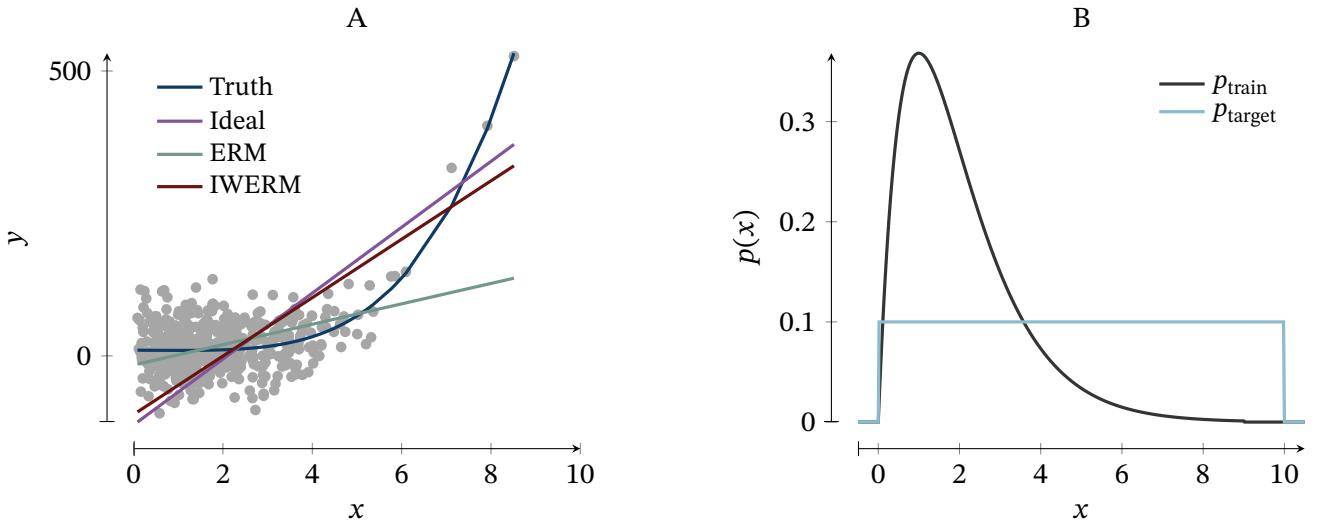
Figure 2.2 – An example illustrating the effect of model mismatch in combination with an imbalanced dataset. The true relationship is given by $y = 0.1x^4 - 0.5x + 10$, and measurements are corrupted by Gaussian noise with constant variance. Samples x_k are drawn from a gamma distribution with shape parameter 2 and scale parameter 1, truncated to $[0, 10]$. The least-squares empirical risk minimization (ERM) solution is well-adapted to the training data distribution p_{train}. If an estimate is desired that minimizes the expected least-squares loss over the *uniform* distribution on $[0, 10]$, then simply weighting each sample by its *importance* $w_k = p_{\text{target}}(x)/p_{\text{train}}(x) \propto 1/p_{\text{train}}(x)$ yields the IWERM estimate, which is close to the optimal solution which is here defined as the solution that minimizes the mean squared error over $p_{\text{target}}(x)$. In the linear regression case, ERM is equivalent to ordinary least squares (OLS), see section 2.2.1, and IWERM can be performed by means of weighted least squares (WLS), see section 2.2.2. The difference to the optimal solution is due to the finite sample number $N^S = 500$; IWERM is a consistent estimator but may be biased for finite sample numbers. For larger values of N^S, the IWERM estimate converges to the ideal solution. Both the ERM and IWERM regression lines shown are averaged over $N^r = 1000$ realizations to illustrate their respective average performance.

An RM estimator minimizes the expected risk over the *training dataset*, i.e., the data used for identifying the parameters. Thus, if the training dataset is dominated by samples from one region of the data space, then the model parameters will be optimized predominantly to fit the data in that region well — at the cost of other, less-frequently observed regions of the data space. Figure 2.2 shows an illustration of this phenomenon for a simple linear regression task. In this example, the best solution to the problem would, of course, be to select a more appropriate process model $f(x, \theta)$. In practice, however, there may be various reasons for using a misspecified model: such a model may be selected for being a well-known standard model, for having computationally attractive properties, having fewer parameters and being identifiable with little data, or, quite simply, because better models are not readily available.

The behavior illustrated in fig. 2.2 is not generally a problem — it is simply the optimal solution to the ERM optimization problem. It is problematic, however, if model performance on the training data distribution

$$p_{\text{train}}(y, x) = p(y \mid x) \, p_{\text{train}}(x)$$

is *not* the ultimate optimality criterion, but model performance should instead be optimized for some *target distribution*

$$p_{\text{target}}(y, x) = p(y \mid x)\, p_{\text{target}}(x)$$

Covariate Shift that differs from $p_{\text{train}}(y, x)$. A discrepancy between $p_{\text{train}}(x)$ and $p_{\text{target}}(x)$ is called *covariate shift* in the literature [47, 52, 59], and if it is present, the RM-optimal parameter estimate changes to

$$\hat{\theta}_{\text{IWRM}} = \arg\min_{\theta} \mathrm{E}_{p_{\text{target}}}[\ell(y, f(x, \theta))], \tag{2.13}$$

where IWRM denotes *importance-weighted risk minimization*, and the choice of the term *importance-weighted* will become apparent in the next paragraph. If one assumes that $p_{\text{target}}(x) \equiv p_{\text{train}}(x)$, then usual RM is recovered.

How can the optimization problem (2.13) be solved in practice? A simple approach has been proposed by Shimodaira [52]. Assuming that $\mathrm{supp}(p_{\text{target}}) \subset \mathrm{supp}(p_{\text{train}})$ and realizing that

$$
\begin{aligned}
\hat{\theta}_{\text{IWRM}} &= \arg\min_{\theta} \mathrm{E}_{p_{\text{target}}}[\ell(y, f(x, \theta))] \\[2mm]
&= \arg\min_{\theta} \iint \ell(y, f(x, \theta))\, p(y \mid x)\, p_{\text{target}}(x)\, \mathrm{d}y\, \mathrm{d}x \\[2mm]
&= \arg\min_{\theta} \int p_{\text{target}}(x) \int \ell(y, f(x, \theta))\, p(y \mid x)\, \mathrm{d}y\, \mathrm{d}x \\[2mm]
&= \arg\min_{\theta} \int \frac{p_{\text{target}}(x)}{p_{\text{train}}(x)} \cdot p_{\text{train}}(x) \int \ell(y, f(x, \theta))\, p(y \mid x)\, \mathrm{d}y\, \mathrm{d}x \\[2mm]
&= \arg\min_{\theta} \mathrm{E}_{p_{\text{train}}}\left[\frac{p_{\text{target}}(x)}{p_{\text{train}}(x)} \cdot \ell(y, f(x, \theta)) \right],
\end{aligned}
$$

the problem can be reformulated as a weighted version of the classical RM problem. Again replacing the unknown exact training distribution $p_{\text{train}}(x)$ with the empirical distribution, we obtain

$$
\begin{aligned}
\hat{\theta}_{\text{IWERM}} &= \arg\min_{\theta} \mathrm{E}_{P_{\text{N}^S}}\left[\frac{p_{\text{target}}(x)}{p_{\text{train}}(x)} \cdot \ell(y, f(x, \theta)) \right] \\[2mm]
&= \arg\min_{\theta} \frac{1}{N^S} \sum_{k=1}^{N^S} \frac{p_{\text{target}}(x_k)}{p_{\text{train}}(x_k)} \cdot \ell(y_k, f(x_k, \theta)).
\end{aligned}
$$

The weights

$$w_k = \frac{p_{\text{target}}(x_k)}{p_{\text{train}}(x_k)}$$

are called the *importance* of sample k, and the weighted estimation procedure is called
Importance-weighted Empirical Risk Minimization *importance-weighted empirical risk minimization* (IWERM).

Shimodaira [52] demonstrates (under mild regularity conditions on the involved distributions and the model $f(\cdot)$) that $\hat{\theta}_{\text{IWERM}}$ is a consistent estimator, i.e.,

$$\mathrm{plim}(\hat{\theta}_{\text{IWERM}}) = \hat{\theta}_{\text{IWRM}} = \arg\min_{\theta} \mathrm{E}_{p_{\text{target}}}[\ell(y, f(x, \theta))].$$

Unfortunately, however, it is not unbiased for finite sample numbers N^S, as is illustrated in fig. 2.2. To achieve an optimal bias-variance trade-off in the face of finite N^S, Sugiyama and Kawanabe [59] propose to either use *adaptive* importance weighting, taking the weights w_k to a power $0 \leq \beta \leq 1$ which is determined by empirical optimization of an information criterion, or to use a classical regularization scheme.

The key to practically implementing IWERM is an efficient method for estimating the importance, which is generally nontrivial. A naive approach would be to first estimate the two involved probability distributions — $p_{\text{train}}(x)$ and $p_{\text{target}}(x)$ — from the respective datasets and then evaluate their ratio. However, density estimation in high-dimensional spaces is known to be a hard problem if no good parametric density models are available [24, 69]. For this reason, efficient algorithms have been developed which do not estimate the two densities, but instead *directly* estimate the importance w_k [35, 60].

The whole discussion in this section was framed in the context of misspecified models. Is importance weighting irrelevant when using a correctly specified model? Indeed, Byrd and Lipton [12] show, as might be expected, that when using over-parameterized deep neural networks for classification tasks on separable datasets, the effect of importance weights on the resulting model diminishes over successive training epochs. The authors thus question the utility of using importance weighting in the over-parameterized learning regime.

Finally, note that there is a close connection between importance-weighted regression and *statistical linearization*. In (multivariate) statistical linearization, the aim is to find

Statistical Linearization

$$\hat{M}, \hat{M}_0 = \underset{M, M_0}{\arg\min} \, \mathrm{E}_{p_{\text{target}}}[(f(x) - Mx - M_0)^2],$$

where $f(x) : \mathbb{R}^n \to \mathbb{R}^1$ is some nonlinear function to be linearized, and $p_{\text{target}}(x)$ is some distribution over which the expected squared linearization error is to be minimized. The difference to the regression setting discussed above is that in statistical linearization, both the nonlinear function $f(x)$ to be approximated and the distribution $p_{\text{target}}(x)$ are known exactly. Statistical linearization has applications in nonlinear filtering and smoothing [28, 63], as will be discussed in some detail in section 2.3.3.

2.2 Inference in static models

Having introduced important basic concepts of statistical inference in the previous section, the present section will briefly summarize a few well-known inference procedures in the context of *static* systems. The classical ordinary least squares (OLS) estimator is introduced in section 2.2.1, including the *recursive* least squares (RLS) algorithm. Section 2.2.2 then discusses an important generalization of OLS, the *weighted* least squares (WLS) algorithm. This algorithm will be of particular importance in the context of this thesis for implementing importance-weighted regression.

2.2.1 Ordinary Least Squares

Consider the static linear model

$$y = x\theta + \varepsilon \tag{2.14}$$

with observations $y \in \mathbb{R}^{N^S}$ and $x \in \mathbb{R}^{N^S \times m}$, parameters $\theta \in \mathbb{R}^m$ to be estimated, and (also unknown) equation error $\varepsilon \in \mathbb{R}^{N^S}$ with covariance matrix $\Sigma_\varepsilon \in \mathbb{R}^{N^S \times N^S}$. We will further

denote the i-th row of x, containing the values of all covariates x^j at sample k, by x_k^T, $x_k \in \mathbb{R}^m$. Furthermore, assume

1. *model correctness*, i.e., there is a parameter vector θ^* for which the model (2.14) correctly describes the stochastic process from which y and x are drawn,

2. *strict exogeneity*, i.e., $\mathrm{E}[\varepsilon \mid x] = 0$, and

3. *linear independence* between all regressors in x.

Ordinary Least Squares

Then, the *ordinary least squares* (OLS) estimator

$$\hat{\theta}_{\mathrm{OLS}} = \arg\min_\theta \sum_{k=1}^{N^S} (y_k - x_k^\mathsf{T}\theta)^2 = \arg\min_\theta \|y - x\theta\|_2^2$$
$$= (x^\mathsf{T} x)^{-1} x^\mathsf{T} y \tag{2.15}$$

is *unbiased* [79], i.e.,

$$\mathrm{E}[\hat{\theta}_{\mathrm{OLS}} \mid X] = \theta^*.$$

The variance-covariance matrix Σ_{OLS} of this parameter estimate can be derived by error propagation from the errors in the observations. To this end, recall that if

$$z = A\xi$$

with a fixed matrix $A \in \mathbb{R}^{k \times k}$ and a stochastic variable $\xi \in \mathbb{R}^k$ with variance-covariance matrix Σ_ξ, then the variance-covariance matrix of the stochastic variable z is given by

$$\Sigma_z = A\Sigma_\xi A^\mathsf{T}. \tag{2.16}$$

Thus, using (2.15) and (2.16), we obtain

$$\Sigma_{\mathrm{OLS}} := \mathrm{Var}[\hat{\theta}_{\mathrm{OLS}} \mid x] = (x^\mathsf{T} x)^{-1} x^\mathsf{T} \Sigma_\varepsilon x (x^\mathsf{T} x)^{-1}. \tag{2.17}$$

The OLS solution (2.15) can be efficiently obtained for large N^S by rearranging eq. (2.15) into the so-called normal form

$$(x^\mathsf{T} x)\hat{\theta}_{\mathrm{OLS}} = x^\mathsf{T} y$$

and using any of the classical numerical methods for solving large, over-determined linear systems, e.g., computing a QR decomposition of the normal matrix $(x^\mathsf{T} x)$.

If one additionally assumes *spherical errors*, i.e.,

$$\mathrm{Var}[\varepsilon \mid x] = \sigma^2 I_{N^S},$$

where I_{N^S} is the identity matrix of dimension N^S, then eq. (2.17) reduces to

$$\Sigma_{\mathrm{OLS}} = \sigma^2 (X^\mathsf{T} X)^{-1}.$$

Under these conditions, the *Gauss-Markov theorem* states that the OLS estimator (2.15) is the *best linear unbiased estimator (BLUE)* of θ, meaning that its variance Σ_{OLS} is optimal in the

class of unbiased linear estimators [79]. Moreover, under the same assumptions, the *reduced chi-squared statistic*

$$\chi_\nu^2 = \frac{\hat{\varepsilon}^\mathsf{T}\hat{\varepsilon}}{\nu} \tag{2.18}$$

is an unbiased estimator of the equation error covariance, i.e.,

$$\mathrm{E}[\chi_\nu^2 \mid x] = \sigma^2.$$

In (2.18),

$$\hat{\varepsilon} = y - x\hat{\theta}_\mathrm{OLS}$$

denotes the estimated equation error, and $\nu = N^S - m$ are the *degrees of freedom* [79].

Finally, if one also assumes $\varepsilon_k \sim \mathcal{N}(0, \sigma^2)$, i.e., the equation errors are normally distributed, it follows that

1. the OLS estimator (2.15) is also normally distributed following $\hat{\theta}_\mathrm{OLS} \sim \mathcal{N}(\theta, \Sigma_\mathrm{OLS})$ [79], and

2. the OLS estimator (2.15) coincides with the maximum likelihood (ML) estimator [26][10]

$$\hat{\theta}_\mathrm{ML} = \arg\max_\theta p(y \mid X, \theta).$$

Correlation between rows of x, as is prevalent, e.g., if the columns of x are time series, is not a problem per se, and all of the above results still hold. Time series can, however, lead to the assumption of strict exogeneity ($\mathrm{E}[\varepsilon \mid x] = 0$) being violated. In this case, it is reassuring to know that under the relaxed assumption of *contemporaneous exogeneity*, i.e. $\mathrm{E}[\varepsilon_k \mid x_k] = 0 \quad \forall\ k$, the OLS estimator (2.15), while no longer necessarily unbiased, is still consistent. As a final remark on OLS, note that at no point we made any assumption about the covariates' origin. In particular, these may be arbitrary nonlinear transformations of other variables or one another, enabling the efficient solution of nonlinear regression problems — as long as they are linear in the regression parameters θ and fulfill the other optimality criteria.

A recursive algorithm for calculating the OLS estimate $\hat{\theta}_\mathrm{OLS}$, the classical *recursive least squares (RLS)* algorithm, can also easily be derived. The OLS solution based on observations up until index k is given by the solution to the normal equation

$$(x_{1:k}^\mathsf{T} x_{1:k})\hat{\theta}_\mathrm{OLS}^{(k)} = x_{1:k}^\mathsf{T} y_{1:k}$$

or, in scalar form,

$$\left(\sum_{i=1}^{k} x_i x_i^\mathsf{T}\right)\hat{\theta}_\mathrm{OLS}^{(k)} = \sum_{i=1}^{k} y_i x_i.$$

Assuming that $R_k = X_{1:k}^\mathsf{T} X_{1:k}$ is invertible, i.e., the solution is unique, we have

$$R_{k-1} = R_k - x_k x_k^\mathsf{T}.$$

---

[10]We already demonstrated this for the more general, nonlinear case in section 2.1.2.

Algorithm 1: Recursive least squares algorithm (RLS)

---
1 **Function** `RecursiveLeastSquares`

    Input : Mean $\hat{\theta}_{\text{OLS}}^{(0)} = \in \mathbb{R}^m$ and covariance matrix $P_0 \in \mathbb{R}^{m \times m}$ of the
            initial parameter estimate $p(\theta)$, and observations $y_k \in \mathbb{R}$ and
            $x_k \in \mathbb{R}^m$, each for $k = 1, \dots, N^S$.

    Output: Mean $\hat{\theta}_{\text{OLS}}^{(k)}$ and covariance P_k of the posterior distributions
            $p(\theta \mid y_{1:k})$ for $k = 1, \dots, N^S$.

2   **begin**
3   **for** $k = 1$ **to** N^S **do**
4             $e_k = y_k - x_k^\mathsf{T} \hat{\theta}_{\text{OLS}}^{(k-1)}$
5             $P_k = P_{k-1} - \dfrac{P_{k-1} x_k x_k^\mathsf{T} P_{k-1}}{1 + x_k^\mathsf{T} P_{k-1} x_k}$
6             $K_k = P_k x_k$
7             $\hat{\theta}_{\text{OLS}}^{(k)} = \hat{\theta}_{\text{OLS}}^{(k-1)} + K_k e_k$
8   **end**
9   **end**

Thus, it follows that

$$
\begin{aligned}
\hat{\theta}_{\text{OLS}}^{(k)} &= R_k^{-1} \left(\sum_{i=1}^{k-1} y_i x_i + y_k x_k \right) \\
&= R_k^{-1} \left(R_{k-1} \hat{\theta}_{\text{OLS}}^{(k-1)} + y_k x_k \right) \\
&= \hat{\theta}_{\text{OLS}}^{(k-1)} + R_k^{-1} \left(y_k x_k - x_k x_k^\mathsf{T} \hat{\theta}_{\text{OLS}}^{(k-1)} \right) \\
&= \hat{\theta}_{\text{OLS}}^{(k-1)} + R_k^{-1} x_k \left(y_k - x_k^\mathsf{T} \hat{\theta}_{\text{OLS}}^{(k-1)} \right).
\end{aligned}
$$

From the matrix inversion lemma,

$$
R_k^{-1} = R_{k-1}^{-1} - \frac{R_{k-1}^{-1} x_k x_k^\mathsf{T} R_{k-1}^{-1}}{1 + x_k^\mathsf{T} R_{k-1}^{-1} x_k}
$$

and by substituting $P_k := R_k^{-1}$, we obtain algorithm 1, which represents an exact, recursive method for calculating the OLS parameter estimate $\hat{\theta}_{\text{OLS}}$. Note that in general, $e_k \neq \hat{\varepsilon}_k$, because the RLS residual e_k is based on the parameter estimate $\hat{\theta}_{\text{OLS}}^{(k)}$, taking only data up until index k into account. In section 2.3.2, a relationship between the *Kalman filter* algorithm and the RLS algorithm is discussed; this relationship reveals that the RLS estimate can be interpreted as a Gaussian random variable with mean $\hat{\theta}_{\text{OLS}}^{(k)}$ and covariance matrix P_k. Furthermore, it also clarifies the meaning of the initial parameter values $\hat{\theta}_{\text{OLS}}^{(0)}$ and P_0: if these are interpreted as the mean and covariance of a Gaussian prior over the parameters θ, then the resulting estimate is the MAP estimate with respect to that prior. To obtain the classical OLS estimate, one must choose $P_0 = \text{diag}(\infty, \dots, \infty)$.

To conclude this section, the long-known OLS estimator represents the optimal solution to the static linear model identification problem from a large number of different perspectives. The assumptions required for the optimality of the OLS estimator are comparatively weak, and a huge number of conceptual extensions have been developed over the last century for relaxing them even further, such as generalized [1] and total least squares [22, 57]. *Weighted least squares*, which is a special case of generalized least squares, will be introduced in the following section. The most critical assumption in the context of this thesis, however, is the assumption of model correctness. If any of the model assumptions are violated, the remarks given in section 2.1.5 and section 2.1.6 apply. Recall, in particular, fig. 2.2, which provides an example of OLS applied in the face of model misspecification and dataset imbalance.

2.2.2 Weighted least squares

Instead of the ordinary least squares cost function

$$\ell_{\text{OLS}}(\theta) = \|y - x\theta\|_2^2 = \sum_{k=1}^{N^S}(y_k - x_k^{\mathsf{T}}\theta)^2 = \sum_{k=1}^{N^S}|\hat{\varepsilon}_k|^2,$$

one may — for various reasons, which will be explored below — want to minimize the *weighted least squares (WLS)* cost function

$$\ell_{\text{WLS}}(\theta) = \sum_{k=1}^{N^S} w_k|\hat{\varepsilon}_k|^2 = \|W^{1/2}\hat{\varepsilon}\|_2^2, \tag{2.19}$$

where $w_k \geq 0$ is the weight of the kth observation, and $W = \text{diag}(w_1, \dots, w_{N^S})$. Notice that in the case of deterministic regression function and a Gaussian equation error distribution

$$q_k(y_k \mid x_k, \theta, \sigma_\varepsilon^2) = \mathcal{N}(f(x_k, \theta), \sigma_\varepsilon^2),$$

eq. (2.19) is identical to the weighted log-likelihood

$$\sum_{k=1}^{N^S} w_k \log q_k(y_k \mid x_k, \theta, \sigma_\varepsilon^2),$$

see section 2.1.2. Thus, in this case, weighted least squares (WLS) estimation coincides with *weighted maximum likelihood (WML) estimation* [17, 40, 74].

By calculating the gradient of eq. (2.19) and setting it to zero, one obtains the weighted normal equation

Weighted Maximum Likelihood Estimation

$$(x^{\mathsf{T}}Wx)\hat{\theta}_{\text{WLS}} = x^{\mathsf{T}}Wy.$$

The WLS optimal parameter estimate is then given by

$$\hat{\theta}_{\text{WLS}} = (x^{\mathsf{T}}Wx)^{-1}x^{\mathsf{T}}Wy = (\tilde{X}^{\mathsf{T}}\tilde{X})^{-1}\tilde{X}^{\mathsf{T}}\tilde{y},$$

which is the standard OLS solution with

$$\tilde{X} = \sqrt{W} \cdot x = \text{diag}\left(\sqrt{w_1}, \dots, \sqrt{w_{N^S}}\right) \cdot x$$

and

$$\tilde{y} = \sqrt{W} \cdot y.$$

Following the same argument as in the OLS case, the variance-covariance matrix of this estimator is given by

$$\Sigma_{\mathrm{WLS}} := \mathrm{Var}[\hat{\theta}_{\mathrm{WLS}} \mid x] = (x^\mathsf{T} W x)^{-1} x^\mathsf{T} W \Sigma_\varepsilon W^\mathsf{T} x (x^\mathsf{T} W^\mathsf{T} x)^{-1}. \qquad (2.20)$$

Now, a distinction can be made between four cases.

Basic OLS: $W = I$ **and** $\Sigma_\varepsilon = \sigma^2 I$**.** All measurements are weighted equally, and equation errors are assumed to be homoscedastic and uncorrelated. In this case, eq. (2.20) reduces to

$$\Sigma_{\mathrm{WLS}} = \Sigma_{\mathrm{OLS}} = \sigma^2 (x^\mathsf{T} x)^{-1}$$

and the estimator is the BLUE, as described in the previous section.

Classical WLS, Σ_ε **known:** $W = \Sigma_\varepsilon^{-1}$**.** In classical WLS, equation errors are assumed to be *heteroscedastic* (but still uncorrelated, whence W is diagonal), i.e., the equation error covariance varies between samples. It can be shown that in this case, WLS with $W = \Sigma_\varepsilon^{-1}$, i.e., weighting samples by the inverse of their respective equation error covariance, yields the BLUE estimator [1]. Under this assumption,

$$\Sigma_{\mathrm{WLS}} = (X^\mathsf{T} W X)^{-1}.$$

Note that in this setting, the OLS estimator (2.15) is still unbiased, but it is no longer efficient.

Classical WLS, but Σ_ε **unknown.** In practice, of course, the equation error covariance Σ_ε is often unknown and must be estimated as well. In this case, various so-called *feasible generalized least squares (FGLS)* methods can be applied, which estimate both Σ_ε and θ from the data [79].[11]

WLS, but still assuming $\Sigma_\varepsilon = \sigma^2 I$**.** One might also want to weight samples differently *for reasons other than a varying measurement precision*. In this case, still assuming homoscedastic and uncorrelated measurements, we have to settle for eq. (2.20) for the estimator covariance. In this case, σ^2 is usually approximated by the weighted reduced chi-squared statistic

$$\sigma^2 \approx \chi_\nu^2 = \frac{\hat{\varepsilon}^\mathsf{T} W \hat{\varepsilon}}{\nu},$$

where $\nu = n - m$ are again the degrees of freedom. Note that using such an estimator only makes sense if one of the OLS/WLS assumptions is violated, because otherwise $W = \Sigma_\varepsilon^{-1}$ yields the BLUE. One possible reason for such an approach is the implementation of IWERM in the presence of model mismatch and an imbalanced dataset, as discussed in section 2.1.6.

---

[11]Note that *generalized least squares (GLS)* accounts not only for heteroscedasticity, as discussed here, but also for possible correlations in the equation error between samples. It is thus a further generalization of WLS.

2.3 State inference in dynamical models

While the previous section considered estimation procedures for *static* models, both the present and the following section are concerned with *dynamic models*, i.e., systems which describe the evolution of some quantity over time. The present section explores the task of estimating the unknown states in a known dynamical system *given a model of that system* — a task known as *state estimation* — , whereas the following section concerns the identification of the dynamical model itself — a task known as *system identification*. Section 2.3.1 introduces the famous Kalman filter and smoother for linear systems. In section 2.3.2, the application of Kalman filtering and smoothing for time-varying parameter estimation in linear quasi-static systems will be discussed. Finally, state estimation in *nonlinear* dynamic systems requires more sophisticated algorithms and is the subject of section 2.3.3.

Dynamic Models

State Estimation

System Identification

2.3.1 The linear Kalman filter and smoother

The celebrated *Kalman filter*, sometimes also called Kalman-Bucy filter or Stratonovich-Kalman-Bucy filter, has been developed around 1960 by a number of contributors including Kalman [33, 34], Bucy [34], Stratonovich [58] and Swerling [61]. Despite its long history, the Kalman filter is still highly relevant today in almost all branches of science and engineering, and it has even been hypothesized that the human brain performs a variant of Kalman filtering [20, 78]. The modern statistical field of *Bayesian structural time series modeling* makes heavy use of Kalman filters [11], and they are also being used to efficiently implement modern probabilistic non-parametric inference schemes [13, 77].

Kalman Filter

The Kalman filter is the optimal estimator of the states $\theta_k \in \mathbb{R}^{N^\theta}$ in the linear, discrete-time state-space model

$$
\begin{aligned}
\theta_k &= A_k \theta_{k-1} + B_k u_k + \eta_k \\
y_k &= C_k \theta_k + \nu_k,
\end{aligned}
\tag{2.21}
$$

given an initial estimate $\theta_0 \sim \mathcal{N}\left(\hat{\theta}_0^+, P_0^+\right)$ and measurements $y_i \in \mathbb{R}^{N^y}$ up to time index $i = k$. In eq. (2.21), $\nu_k \sim \mathcal{N}\left(0, \Sigma_{\nu,k}\right)$ and $\eta_k \sim \mathcal{N}\left(0, \Sigma_{\eta,k}\right)$ are measurement and process noise, respectively. The matrices $A_k, B_k, C_k, \Sigma_{\nu,k}$ and $\Sigma_{\eta,k}$ are assumed to be known, as well as the input signal $u_k \in \mathbb{R}^{N^u}$.[12] All of the matrices may be time-varying or constant. The Kalman filter represents an efficient recursive algorithm for evaluating the posterior distribution of the states at time k, $p(\theta_k \mid y_1, y_2, \ldots, y_k)$. Given the posterior distribution $p(\theta_{k-1} \mid y_1, y_2, \ldots, y_{k-1})$ at the previous time step, it follows from eq. (2.21) that

$$
p(\theta_k \mid y_{1:k}) = p(\theta_{k-1} \mid y_{1:k-1}) \cdot \delta(\theta_k - A_k \theta_{k-1} - B_k u_k - \eta_k) \cdot \\
\delta(y_k - C_k \theta_k - \nu_k) \cdot p(\eta_{k-1}) \cdot p(\nu_k), \quad (2.22)
$$

where $\delta(\cdot)$ denotes the Dirac delta distribution. If both the previous posterior distribution and the noise distributions are Gaussian, it follows from eq. (2.22) that the next posterior

---

[12]Stochastic input signals can be represented in this model as well, although this requires increasing the dimension of the state space.

Algorithm 2: Linear Kalman filter

---

1 **Function** KF

    Input   :Mean $\hat{\theta}_0^+ \in \mathbb{R}^{N^\theta}$ and covariance $P_0^+ \in \mathbb{R}^{N^\theta \times N^\theta}$ of the initial state estimate $p(\theta_0)$, matrices $A_k, B_k, C_k, \Sigma_{\nu,k}$ and $\Sigma_{\eta,k}$ of suitable dimensions, as well as measurements $y_k \in \mathbb{R}^{N^y}$ and inputs $u_k \in \mathbb{R}^{N^u}$, each for $k = 1, \dots, N^S$.

    Output:Means $\hat{\theta}_k^+$ and covariances P_k^+ of the posterior distributions $p(\theta_k \mid y_{1:k})$ for $k = 1, \dots, N^S$.

2   **begin**

3   **for** $k = 1$ **to** N^S **do**

       // State prediction using eq. (2.21)

4        $\hat{\theta}_k^- = A_k \hat{\theta}_{k-1}^+ + B_k u_k$

5        $P_k^- = A_k P_{k-1}^+ A_k^\mathsf{T} + \Sigma_{\eta,k}$

       // Recursive posterior update using eq. (2.22)

6        $e_k = y_k - C_k \hat{\theta}_k^-$

7        $\Sigma_{e,k} = C_k P_k^- C_k^\mathsf{T} + \Sigma_{\nu,k}$

8        $K_k = P_k^- C_k^\mathsf{T} \Sigma_{e,k}^{-1}$

9        $\hat{\theta}_k^+ = \hat{\theta}_k^- + K_k e_k$

10      $P_k^+ = (I - K_k C_k) P_k^-$

11   **end**

12 **end**

---

will be Gaussian as well:

$$p(\theta_k \mid y_1, y_2, \dots, y_k) = \mathcal{N}\left(\hat{\theta}_k^+, P_k^+\right).$$

Hence, by induction, the posterior distributions at all time steps are Gaussian, and an estimation scheme can be efficiently implemented by operating on the parameters of these Gaussian distributions. If one then chooses the mean $\hat{\theta}_k^+$ of the posterior distribution as the state estimate, the resulting algorithm is the classical Kalman filter algorithm [33, 50, 55], listed in algorithm 2. Moreover, the Kalman filter also yields the covariances P_k^+ of the (Gaussian) posterior distribution.

Given that the Kalman filter simply represents an efficient procedure for calculating the *exact* mean of the posterior state distribution, it is optimal in many ways: if the noise and initial state estimate are Gaussian, it simultaneously maximizes $p\left(\hat{\theta}_k^+ \mid y_{1:k}\right)$ (making it the MAP estimator) and minimizes $\mathrm{E}[(\hat{\theta}_k^+ - \theta_k)^\mathsf{T} \Sigma_{\eta,k}^{-1} (\hat{\theta}_k^+ - \theta_k)]$ (making it the MMSE estimator) [55]. Even if the noise is not assumed to be Gaussian (but still zero-mean and white), the Kalman filter is a *linear* MMSE estimator, i.e., the optimal linear estimator in the MMSE sense [55].

For practical applications, a number of modifications to the original Kalman filter algorithm may be required to obtain a stable and robust algorithm. Some of these modifications, which will be used later in this thesis, are discussed in the following.

Symmetric positive definite state covariance matrices. A frequent problem with the formulation

$$P_k^+ = (I - K_k C_k) P_k^-$$

of the covariance update formula — used in the last step of algorithm 2 — is that the state covariance matrix P_k^+ may loose symmetry or positive definiteness, both of which are necessary for it to be a proper covariance matrix. This problem is easily solved by employing analytical reformulations of the covariance update equation, which are numerically more favorable. Various versions have been proposed; we will employ the *Joseph stabilized covariance update*

$$P_k^+ = (I - K_k C_k) P_k^- (I - K_k C_k)^\mathsf{T} + K_k \Sigma_{\nu,k} K_k^\mathsf{T}. \tag{2.23}$$

Joseph Stabilized Covariance Update

This formulation guarantees that P_k^+ is symmetric positive definite, as long as P_k^- is [55].

Matrix inversions in the update step. Another step of the basic algorithm which frequently causes numerical problems is the matrix inversion that is required to evaluate the expression

$$K_k = P_k^- C_k^\mathsf{T} (C_k P_k^- C_k^\mathsf{T} + \Sigma_{\nu,k})^{-1}$$

for the Kalman gain K_k in algorithm 2. (Note that this is an inversion of an $N^y \times N^y$ matrix, which thus only causes problems when $N^y > 1$, i.e. when more than one measurement signal shall be taken into account.) It turns out that if the measurement noise covariance matrix $\Sigma_{\nu,k}$ is diagonal or static, then, as a consequence of the *matrix inversion lemma*, the matrix inversion can be replaced by a sequence of N^y scalar update steps [55]. This is known as *sequential Kalman filtering* and described in detail in algorithm 3. There, C_{ik} denotes the ith row of C, y_{ik} the ith entry of the measurement vector y_k, and $\Sigma_{\nu,kii}$ the ith diagonal entry of the measurement noise covariance matrix $\Sigma_{\nu,k}$. Simon [55] also discusses a way of applying algorithm 3 to systems with non-diagonal but static measurement noise covariance matrices.

Sequential Kalman Filtering

Missing data. While the two previous modifications merely addressed numerically favorable but analytically equivalent reformulations of the basic algorithm, missing data pose a challenge that is simply not handled by the basic algorithm. What should happen if an entry of the measurement vector y_k is unavailable or an entry of the measurement matrix C_k is unknown? If all of the measurements at time index k are unavailable or the complete measurement matrix is unknown, then the solution is simple. In this case, a correction of the predicted state distribution $p(\theta_k \mid y_{1:k-1})$ based on the measurement y_k is simply not possible, and hence the posterior state distribution at step k is equal to the predicted state distribution: the dynamical model is used to predict the state evolution in open loop, without corrections from measurements.

If only *some* entries are missing, e.g., one of multiple measurement signals is subject to a temporary failure, a slightly more complex approach is required. In this case, a simple approach — assuming a diagonal measurement noise covariance matrix $\Sigma_{\nu,k}$ — is to set the corresponding diagonal entry $\Sigma_{\nu,kii}$ of the measurement noise covariance matrix to ∞. It follows that in algorithm 3, $K_k^{(i)} = 0$, $K_k^{(i)} C_{ki} = 0$ and $K_k^{(i)} \Sigma_{\nu,kii} (K_k^{(i)})^\mathsf{T} =$

Algorithm 3: Sequential state update step for the Kalman filter

---

1 **Function** `SequentialStateUpdate`

    Input : Mean $\hat{\theta}_k^- \in \mathbb{R}^{N^\theta}$ and covariance $P_k^+ \in \mathbb{R}^{N^\theta \times N^\theta}$ of the predicted state estimate $p(\theta_k \mid y_{1:k-1})$, matrices C_k and $\Sigma_{\nu,k}$ of suitable dimensions, as well as measurements $y_k \in \mathbb{R}^{N^y}$. The matrix $\Sigma_{\nu,k}$ *must* be diagonal.

    Output : Mean $\hat{\theta}_k^+$ and covariance P_k^+ of the posterior distribution $p(\theta_k \mid y_1, y_2, \dots, y_k)$.

2   **begin**

3         $\hat{\theta}_k^{+(0)} = \hat{\theta}_k^-$

4         $P_k^{+(0)} = \hat{\theta}_k^-$

        // Iteratively update based on the individual measurements

5   **for** $i = 1$ **to** N^y **do**

            // Scalar update, requiring no matrix inversion

6             $K_k^{(i)} = \dfrac{P_k^{+(i-1)} C_{ki}^\mathsf{T}}{C_{ki} P_k^{+(i-1)} C_{ki}^\mathsf{T} + \Sigma_{\nu,kii}}$

7             $\hat{\theta}_k^{+(i)} = \hat{\theta}_k^{+(i-1)} + K_k^{(i)}(y_{ki} - C_{ki}\hat{\theta}_k^{+(i-1)})$

            // Joseph stabilized covariance update based on eq. (2.23)

8             $P_k^{+(i)} = (I - K_k^{(i)} C_{ki}) P_k^{+(i-1)} (I - K_k^{(i)} C_{ki})^\mathsf{T} + K_k^{(i)} \Sigma_{\nu,kii} (K_k^{(i)})^\mathsf{T}$

9   **end**

10         $\hat{\theta}_k^+ = \hat{\theta}_k^{+(N^y)}$

11         $P_k^+ = P_k^{+(N^y)}$

12   **end**

---

0. This method, of course, generalizes to the simple case discussed above, where all measurement signals are missing.

State constraints For various reasons, it may be desirable to consider a set of inequality constraints

$$D\theta \leq d \tag{2.24}$$

for the possible values of the states. A simple and efficient method to solve the constrained state estimation problem has been proposed by Simon and Chia [54]. At each time step, the unconstrained posterior state estimate $\hat{\theta}_k^+$ is simply projected onto the constraint surface. To this end, one solves the optimization problem

$$\tilde{\theta}_k = \arg\min_x (x - \hat{\theta}_k^+)^\mathsf{T} W(x - \hat{\theta}_k^+), \quad \text{such that} \quad D\theta \leq d, \tag{2.25}$$

where W is any symmetric positive definite weighting matrix. In general, eq. (2.25) represents a quadratic programming problem which can be solved efficiently using any of the standard solvers. The resulting constrained state estimator has various favorable properties and can generally be regarded as uniformly superior to the unconstrained estimator, i.e., the standard linear Kalman filter (if constraints are present, that is) [54]. In the simplest case where only lower and upper bounds $\mathrm{lb} \in \mathbb{R}^{N^\theta}$ and $\mathrm{ub} \in \mathbb{R}^{N^\theta}$ on the states are to be respected, eq. (2.25) simply reduces to

$$\tilde{\theta}_{ki} = \min(\max(\hat{\theta}_{ki}^+, \mathrm{lb}_i), \mathrm{ub}_i) \quad \forall i = 1, \dots, N^\theta, \tag{2.26}$$

regardless of the choice of W. One instance where this is particularly useful is when many data points are missing or faulty: in this case, depending on the properties of the state-space model, an unconstrained state estimate may diverge and lead to bad filter performance or numerical problems. Finally, note that the constrained linear state estimation approach described here is certainly not optimal: to derive a strictly optimal estimator, one needs to depart from the Gaussian framework since constrained state estimates can no longer be Gaussian. Doing so incurs a very significant increase in computational complexity, however. For a comprehensive overview of constrained filtering methods, refer to Simon [53].

Incorporating all of the above improvements into the original algorithm 2, we obtain a practical, robust, and computationally efficient version of the Kalman filter algorithm.

The Kalman filter is a *causal* estimator: it only uses measurements up to time index k for estimating the value of the states at that time. An estimator that estimates the posterior distribution $p(\theta_k \mid y_{1:N^s})$ of the states given *all* measurements is called a (fixed-interval) *Kalman smoother*. Given the same assumptions as for the Kalman filter, the smoothed distribution can also be computed efficiently, usually using a forward-backward recursion. Various analytically equivalent algorithms with differing numerical properties have been proposed in the literature [50, 55, 70]. One classical algorithm is the linear *Rauch-Tung-Striebel smoother* (RTS) [48], which is provided in algorithm 4.

It enjoys the same optimality properties as the Kalman filter: it is both a MAP and MMSE estimator.

Until this point, the Kalman filter and smoother have been discussed from a probabilistic point of view. Both can also be derived as the solutions to an optimization problem, an

Kalman Smoother

Rauch-Tung-Striebel Smoother

37

Algorithm 4: Linear Rauch-Tung-Striebel smoother (RTS)

---

1 **Function** RTS

   Input : Mean $\hat{\theta}_0^+ \in \mathbb{R}^{N^\theta}$ and covariance $P_0^+ \in \mathbb{R}^{N^\theta \times N^\theta}$ of the initial
      state estimate $p(\theta_0)$, matrices $A_k, B_k, C_k, \Sigma_{\nu,k}$ and $\Sigma_{\eta,k}$ of
      suitable dimensions, as well as measurements $y_k \in \mathbb{R}^{N^y}$ and
      inputs $u_k \in \mathbb{R}^{N^u}$, each for $k = 1, \dots, N^S$.

   Output : Means $\hat{\theta}_k^s$ and covariances P_k^s of the posterior
       distributions $p(\theta_k \mid y_{1:N^S})$ for $k = 0, 1, \dots, N^S$.

2   **begin**
    // Forward filtering using the standard Kalman filter
3   Run KF
    // Backward smoothing using the Rauch-Tung-Striebel smoother
4    $\hat{\theta}_{N^S}^s = \hat{\theta}_{N^S}^+$
5   **for** $k = N^S - 1$ **to** 0 **do**
6   Solve $S_k P_{k+1}^- = P_k^+ A_k^\mathsf{T}$ for S_k
7     $P_k^s = P_k^+ + S_k(P_{k+1}^s - P_{k+1}^-)S_k^\mathsf{T}$
8     $\hat{\theta}_k^s = \hat{\theta}_k^+ + S_k(\hat{\theta}_{k+1}^s - \hat{\theta}_{k+1}^-)$
9   **end**
10   **end**

---

alternative perspective that will prove valuable in the following. Recall that we sketched the derivation of the Kalman smoother as the solution to the MAP estimation problem

$$
\begin{aligned}
\hat{\theta}_{0:N^S}^s &= \arg\max_{\theta_{0:N^S}} p(\theta_{0:N^S} \mid y_{1:N^S}) \\
&= \arg\max_{\theta_{0:N^S}} \frac{p(\theta_{0:N^S}, y_{1:N^S})}{p(y_{1:N^S})} \\
&= \arg\max_{\theta_{0:N^S}} p(\theta_{0:N^S}, y_{1:N^S}).
\end{aligned}
\tag{2.27}
$$

Under the Gaussian assumption, the logarithm of the joint likelihood can be expressed analytically as [49]

$$
\log p(\theta_0, \theta_1, \dots, \theta_{N^S}, y_1, \dots, y_{N^S}) = -\frac{1}{2} \sum_{k=1}^{N^S} \left[(y_k - C_k\theta_k)^\mathsf{T} \Sigma_{\nu,k}^{-1}(y_k - C_k\theta_k) + \log|\Sigma_{\nu,k}| \right]
$$

$$
-\frac{1}{2} \sum_{k=0}^{N^S-1} \left[(\theta_{k+1} - A_{k+1}\theta_k - B_{k+1}u_{k+1})^\mathsf{T} \Sigma_{\eta,k+1}^{-1}(\theta_{k+1} - A_{k+1}\theta_k - B_{k+1}u_{k+1}) + \log|\Sigma_{\eta,k+1}| \right]
$$

$$
-\frac{1}{2} \left[(\theta_0 - \hat{\theta}_0^+)^\mathsf{T}(P_0^+)^{-1}(\theta_0 - \hat{\theta}_0^+) + \log|P_0^+| + N^S(N^\theta + N^y)\log 2\pi \right] \tag{2.28}
$$

and thus, the Kalman smoother can also be formulated as an analytical solution to the problem of maximizing eq. (2.28).

2.3.2 Time-varying parameter estimation using a Kalman filter

The Kalman filter can be seen as a generalization of the recursive least squares (RLS) algorithm (discussed in section 2.2.1) to dynamic models. To see this, consider the state-space model

$$\begin{aligned} \theta_k &= \theta_{k-1} + \eta_k \\ y_k &= x_k^\mathsf{T} \theta_k + \nu_k, \end{aligned} \tag{2.29}$$

where θ_k are the states, $y_k \in \mathbb{R}$ one-dimensional measurements, and x_k^T now denotes a time-varying measurement matrix. If one assumes that $\Sigma_\eta = 0$ and $\Sigma_\nu = \Sigma_\varepsilon$, applying the Kalman filter (algorithm 2) to this model yields exactly the recursive least squares algorithm (algorithm 1) for the static regression problem

$$y = x\theta + \varepsilon$$

with $\varepsilon \sim \mathcal{N}(0, \Sigma_\varepsilon)$. The same equivalence can also be demonstrated for the more general case of a multivariate target variable y_k [37, p. 367]. Interpreting the RLS algorithm as a special case of the Kalman filter also justifies the interpretation of the RLS estimate as a Gaussian variable with mean $\hat{\theta}_{\mathrm{OLS}}^{(k)}$ and covariance matrix P_k. The equivalence to the least squares problem can be derived easily by setting $\Sigma_\eta = 0$, $\Sigma_\nu = \Sigma_\varepsilon$ and $P_0^+ = \mathrm{diag}(\infty, \dots, \infty)$ in eq. (2.28), which leaves us with the classical (multiple) least-squares optimization problem

$$\hat{\theta}_{N^S}^+ = \arg\max_{\theta_{N^S}} p(\theta_{N^S} \mid y_{1:N^S})$$

$$= \arg\max_{\theta = \theta_0 = \theta_1 = \dots = \theta_{N^S}} -\frac{1}{2} \sum_{k=1}^{N^S} (y_k - C_k\theta)^\mathsf{T} \Sigma_\varepsilon^{-1} (y_k - C_k\theta).$$

The equivalence stated above holds for the case of *static* parameter values since it is based on the assumption that $\Sigma_\eta = 0$. To obtain an efficient, recursive parameter estimation algorithm for *time-varying* systems, one may simply use the same state-space model but allows for $\Sigma_\eta > 0$. The expected velocity of parameter changes can be incorporated into the model by setting the corresponding diagonal entries of Σ_η appropriately. This also easily enables incorporating different velocities of change for different parameters or expected correlations between changes in different parameters. Of course, smoothers such as the RTS smoother can be employed to further improve the estimates by taking future measurements into account, if that is permitted by the application. This is, thus, a powerful and general method for time-varying multivariate parameter estimation, and it will be further explored in chapter 6.

2.3.3 Kalman filtering and smoothing in nonlinear systems

What happens if the state-space model for which the states are to be estimated is nonlinear? Consider the model[13]

$$\begin{aligned} \theta_k &= f_k(\theta_{k-1}) + \eta_k \\ y_k &= g_k(\theta_k) + \nu_k \end{aligned} \tag{2.30}$$

---

[13]Inputs are omitted here for brevity and because they are not used in this thesis. However, input nonlinearities can be treated equivalently to state transition nonlinearities.

with arbitrary linear or nonlinear (vector-valued) functions $f_k(\cdot)$ and $g_k(\cdot)$. The classical Kalman filter cannot be used to estimate the states in this model since it relies on the linearity and normality assumptions. Note that even if the initial state estimate and the noise distributions are normal, the posterior distributions of the states can be arbitrarily distorted due to being nonlinearly transformed. For general nonlinearities, it is thus no longer possible to use parametric models for the posterior state distributions if *exact* inference is desired. One option is to drop the use of parametric models and use particle-based methods instead, resulting in a family of methods that are usually called **Particle Filter** *particle filters* [6, 50, 55]. These methods, however, are computationally much more demanding and, in particular, suffer heavily from the curse of dimensionality. They are, therefore, often not practical for larger state-space dimensions. Another option is to content oneself with *approximate* inference in exchange for being able to stay within the computationally efficient parametric framework. The most prominent examples of this family of nonlinear filtering methods are the ones that *linearize* any nonlinearities and, thus, stay within the linear-Gaussian framework. The most well-known members of this **Extended and** family are the *extended and unscented Kalman filters* (EKF, UKF) [55, 73].
Unscented To simplify the following discussion, consider the isolated nonlinearity
Kalman Filters

$$y = f(x)$$

where the prior distribution $p(x)$ is known and a Gaussian approximation $q(y)$ to the posterior distribution $p(y)$ is to be computed. Any transformation that maps a Gaussian onto another Gaussian distribution can be represented as a linear transformation, and hence, any approach that results in a Gaussian approximation of the posterior is equivalent to linearizing the nonlinearity:

$$f(x) \approx Mx + M_0.$$

The EKF approach is to simply linearize $f(x)$ around the mean μ_x of the prior distribution $p(x)$. This is, however, not generally the optimal linearization strategy, in the sense that it does not minimize the discrepancy between the true posterior $p(y)$ and its approximation $q(x)$ in any principled way. For example, one might prefer to choose a procedure which minimizes the expected linearization error, i.e.,

$$\hat{M}, \hat{M}_0 = \underset{M,M_0}{\arg\min} \, \mathrm{E}_{p(x)}[(f(x) - Mx - M_0)^2]$$

in the spirit of risk minimization (see sections 2.1.2 and 2.1.6). This optimization problem is
Statistical called *statistical linearization*, and it can be shown [50] that the solution is given by those
Linearization linearization parameters for which the first and second-order moments of the posterior
distribution $p(y)$ are matched exactly. *Moment matching* in this way is also equivalent to
Moment minimizing the Kullback-Leibler divergence between the true posterior and its
Matching approximation [27, 41], i.e.,

$$\underset{M,M_0}{\arg\min} \, \mathrm{E}_{p(x)}[(f(x) - Mx - M_0)^2] = \underset{M,M_0}{\arg\min} \, \mathrm{D}_{\mathrm{KL}}(p(y) \,\|\, q(x)).$$

Thus, one is left with the task of approximating the first two moments of the posterior distribution:

$$\mu_y := \mathrm{E}[y] = \int_{\mathbb{R}^n} f(x) \, p(x) \, \mathrm{d}x \qquad (2.31)$$

and

$$\Sigma_y := \mathrm{Var}[y] = \int_{\mathbb{R}^n} (f(x) - \mu_y)^\mathsf{T}(f(x) - \mu_y)\, p(x)\, \mathrm{d}x. \qquad (2.32)$$

Finding good numerical approximations to equations of the form (2.31), (2.32) is, of course, the sole subject of numerical quadrature, and many different *quadrature rules* of the form

$$\int f(x)\rho(x)\,\mathrm{d}x \approx \sum_{i=1}^{N^\zeta} w_i\, f(\zeta_i)$$

have been developed over the decades. Here, w_x is some known weighting function that is known a priori, the ζ_i are called *quadrature points*, and the w_i are called *quadrature weights*. A method for choosing the quadrature points and weights is called a *quadrature rule*. For the Gaussian weight function $\rho(x) = p(x) \sim \mathcal{N}(0, I)$, the proposed quadrature rules include the *unscented transform* [73], the *cubature transform* [5] and the *Gauß-Hermite integration rule* [4], among many others. One way to quantify the accuracy of different quadrature rules is to compare the polynomial degree up until which functions $f(x)$ can be integrated *exactly*. For example, if a quadrature rule is exact for polynomials up to order three and $f(x)$ is a quadratic polynomial, the mean μ_y is matched exactly, but the covariance Σ_y is *not* — because $(f(x) - \mu_y)^\mathsf{T}(f(x) - \mu_y)$ is a polynomial of order four. Interestingly, it has been noted that empirically, quadrature rules that are exact up to order three are often sufficient, and little performance can be gained by using higher-order quadrature rules [5].

Quadrature Rules

Unscented Transform

Cubature Transform

Gauß-Hermite Integration Rule

Each quadrature rule gives rise to a corresponding approximate nonlinear Kalman filter and smoother. The filtering algorithms are simply obtained by replacing any occurrences of A_k and C_k by the results of the linearization procedure (linearizing about $q\!\left(\hat{\theta}_{k-1}^+\right)$ and $\hat{\theta}_k^-$, respectively), and accounting for the linearization offset terms in lines 4 and 6 of algorithm 2. To obtain the corresponding approximate nonlinear smoothing algorithm, any standard linear smoother implementation can be used, and the linearization used in the forward (filtering) pass is simply reused.

So far, it was assumed that no information or measurement of the output y of the nonlinearity is given. In the state-space model setting, however, past, present, and future measurements provide information about states, and thus, the expected statistical linearization error can be reduced by exploiting this additional information. Various *iterative* approximate nonlinear filtering and smoothing schemes have been proposed that exploit this insight [7, 62]. One very simple method is to perform a full run of a standard approximate nonlinear filter and smoother and then use the resulting, smoothed state estimates to obtain a more accurate linearization of the nonlinearities. This process can be repeated until convergence. The author of this thesis has (co-)authored two publications on the expression of iterative filtering and smoothing schemes as message passing in factor graphs [28, 46], showing that these methods achieve improved accuracy when compared to a standard, non-iterative EKS or UKS. For brevity's sake, these methods will not be further discussed in this thesis.

2.4 Maximum-likelihood parameter inference in dynamical models

System
Identification

While the previous section presented solutions for inferring the hidden states of a dynamical system from measurements, we always assumed the dynamical system itself to be known. This is, of course, rarely the case, which is why successful *system identification* is in almost all cases a necessary prerequisite for performing state estimation. System identification has been the subject of intensive research for decades [37, 51]. Notably, all of the statistical inference paradigms discussed in section 2.1.2 are applicable to the system identification problem as well. While a broad overview of system identification techniques is beyond the scope of this thesis, we will briefly discuss the application of the maximum-likelihood paradigm to system identification. This section is largely based on [50, Chap. 12].

2.4.1 Maximum likelihood estimation in general state-space models

Following the definition given in eq. (2.1), maximum likelihood (ML) parameter estimates are defined by

$$\hat{\psi}_{\mathrm{ML}} = \arg\max_{\psi} q(y \mid \psi), \tag{2.33}$$

where y are the available observations, ψ the model parameters to be estimated, and $q(y \mid \psi)$ denotes the probabilistic model family under consideration.[14] In a dynamic setting, as opposed to the static regression setting considered in section 2.1.2, the individual measurements y_k cannot be considered statistically independent of one another, and hence the simple decomposition of the likelihood used in eq. (2.2) is not feasible. However, the so-called *prediction error decomposition* [50, p. 176]

$$q(y \mid \psi) = q(y_1 \mid \psi) \prod_{k=2}^{N^S} q(y_k \mid y_{1:k-1}, \psi) \tag{2.34}$$

holds nevertheless. The challenge is now to compute $q(y_k \mid y_{1:k-1}, \psi)$.

Assuming that the hidden states θ_k of the system satisfy the Markov property

$$q(\theta_k \mid \theta_{1:k-1}) = q(\theta_k \mid \theta_{k-1}) \tag{2.35}$$

and that each measurement y_k is conditionally independent of all previous measurements given the state θ_{k-1}, i.e.,

$$q(y_k \mid y_{1:k-1}, \theta_k) = q(y_k \mid \theta_k), \tag{2.36}$$

it follows that

$$q(y_k \mid y_{1:k-1}, \psi) = \int_{\mathbb{R}^{N^\theta}} q(y_k \mid \theta_k, \psi)\, q(\theta_k \mid y_{1:k-1}, \psi)\, \mathrm{d}\theta_k. \tag{2.37}$$

---

[14] In the formulation in section 2.1.2, covariates x were also assumed to be given. In the system identification setting, there is only one stream of measurements y available to perform the estimation; hence the slight difference between eq. (2.1) and eq. (2.33).

Note that $q(\theta_k \mid y_{1:k-1}, \theta)$ is precisely the quantity that is computed or approximated by a Bayesian filter such as the Kalman filter presented in the previous section. If we now recall that

$$\arg\max_{\psi} q(y \mid \psi) = \arg\max_{\psi} \log q(y \mid \psi)$$

and

$$\log q(y \mid \psi) = \log \Big[\, q(y_1 \mid \psi) \prod_{k=2}^{N^S} q(y_k \mid y_{1:k-1}, \psi) \Big]$$

$$= \log q(y_1 \mid \psi) + \sum_{k=2}^{N^S} \log q(y_k \mid y_{1:k-1}, \psi),$$

we obtain the simple recursion

$$\log q(y_{1:k} \mid \psi) = \log q(y_{1:k-1} \mid \psi) + \log q(y_k \mid y_{1:k-1}, \psi) \tag{2.38}$$

for the log likelihood of the data up until index k.[15]

Since an analytical solution to eq. (2.33) is usually unavailable, numerical optimization procedures are employed. To implement an efficient optimization scheme, a method to evaluate the marginalization (2.37) is still required. For the general case of nonlinear systems, approximate solutions can be found in the literature [50, Chap. 12]. The following section demonstrates how the integral in eq. (2.37) can be computed analytically in the special case of the linear Gaussian state-space model.

2.4.2 Maximum likelihood estimation in the linear Gaussian state-space model

As a specific instance of the general developments presented so far, assume that $q(y \mid \theta)$ is the PDF corresponding to the linear Gaussian state-space model

$$\begin{aligned} \theta_k &= A_k(\psi)\theta_{k-1} + B_k(\psi)u_k + \eta_k \\ y_k &= C_k(\psi)\theta_k + \nu_k, \end{aligned} \tag{2.39}$$

with $\nu_k \sim \mathcal{N}\big(0, \Sigma_{\nu,k}(\psi)\big)$ and $\eta_k \sim \mathcal{N}\big(0, \Sigma_{\eta,k}(\psi)\big)$, where some or all of the matrices A_k, B_k, C_k, $\Sigma_{\nu,k}$ and $\Sigma_{\eta,k}$ may or may not depend in some way on a parameter vector ψ to be estimated, and may or may not be time-varying. Note that this general formulation incorporates estimating the full matrices, estimating only specific entries of some matrices, and estimating matrices of a particular structure parameterized by ψ. One particular structural assumption that is often useful for estimating the two noise covariance matrices is

$$\Sigma_\eta(\psi) = L_\eta(\psi)^\mathsf{T} L_\eta(\psi),$$

which ensures that the resulting covariance matrix is symmetric and positive semidefinite (as required), regardless of the entries of L_η. This enables unconstrained optimization methods to be used for ψ.

---

[15]The negative log likelihood is sometimes also called the *energy function* [50, Chap. 12].

The model (2.39) clearly satisfies the assumptions eq. (2.35) and eq. (2.36) stated above. The relationship between state θ_k and measurement y_k is given by

$$q(y_k \mid \theta_k) = \mathcal{N}(y_k \mid C_k \theta_k, \Sigma_{\nu,k}).$$

Moreover, from the basic Kalman filter equations (see algorithm 2), we have that

$$q(\theta_k \mid y_{1:k-1}, \psi) = \mathcal{N}\left(\theta_k \mid \hat{\theta}_k^-, P_k^-\right)$$

and it follows that

$$q(y_k \mid y_{1:k-1}, \psi) = \int_{\mathbb{R}^{N^\theta}} q(y_k \mid \theta_k, \psi) \, q(\theta_k \mid y_{1:k-1}, \psi) \, \mathrm{d}\theta_k$$

$$= \int_{\mathbb{R}^{N^\theta}} \mathcal{N}(y_k \mid C_k \theta_k, \Sigma_{\nu,k}) \, \mathcal{N}\left(\theta_k \mid \hat{\theta}_k^-, P_k^-\right) \mathrm{d}\theta_k$$

$$= \mathcal{N}\left(y_k \mid C_k \hat{\theta}_k^-, C_k P_k^- C_k^\top + \Sigma_{\nu,k}\right).$$

This, in combination with the recursive update equation (2.38) for the log likelihood provides us with an efficient way to compute the (logarithm of the) likelihood $q(y \mid \psi)$ during a normal run of a linear Kalman filter, without computing any additional quantities except for the log likelihood itself: we simply obtain the recursive log likelihood update equation [50, Thm. 12.3]

$$\log q(y_{1:k} \mid \psi) = \log q(y_{1:k-1} \mid \psi) + \frac{1}{2}\log|2\pi\Sigma_{e,k}| + \frac{1}{2}e_k^\top \Sigma_{e,k}^{-1} e_k \qquad (2.40)$$

that can be added to the usual Kalman filter recursion, where

$$e_k = y_k - C_k \hat{\theta}_k^-$$

and

$$\Sigma_{e,k} = C_k P_k^- C_k^\top + \Sigma_{\nu,k}$$

as in algorithm 2. Note that e_k and $\Sigma_{e,k}$ both depend on the choice of the parameters ψ, which determine the various matrices used in the filtering procedure.

With this efficient procedure for evaluating the likelihood $q(y \mid \psi)$ at hand, one can now use an arbitrary optimization algorithm for maximizing that quantity. Note that in most instances, the likelihood is a non-convex function of the parameters and may have multiple local minima, indicating that global optimization methods may be preferable. The gradient $\mathrm{d}\log q(y \mid \psi)/\mathrm{d}\psi$ can be calculated as well ([50, Eqs. (A17-A18)] provide expressions for the case of constant matrices A, B, C, Σ_η, Σ_ν), enabling the use of gradient-based optimization methods which are often more efficient. Figure 2.3 shows an example in which ML estimation was used to tune the noise covariances of a one-dimensional Rauch-Tung-Striebel smoother (RTS).
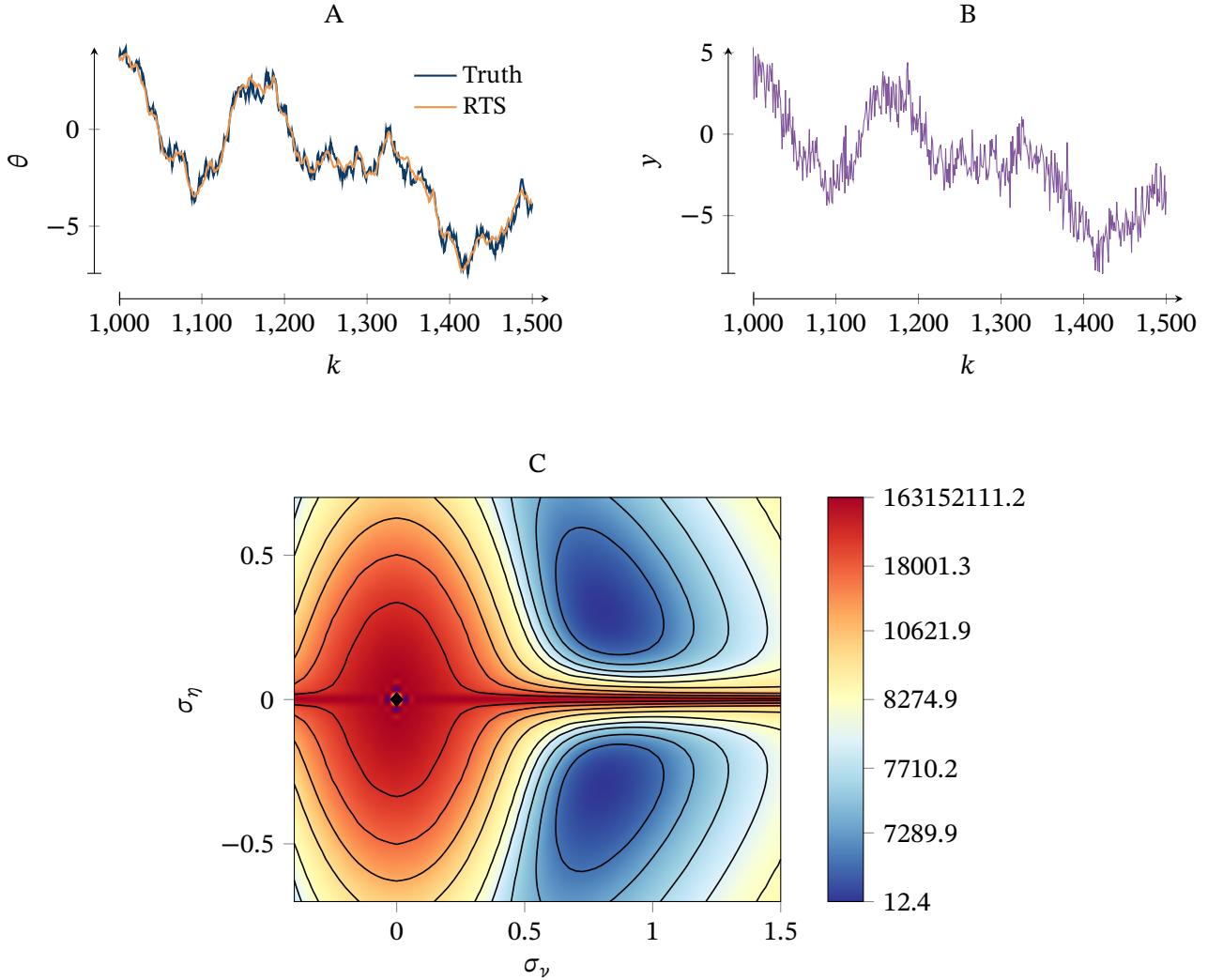
Figure 2.3 – An example illustrating maximum likelihood estimation in a linear state-space model. The state follows the random-walk model $\theta_k = \theta_{k-1} + \eta_k$ with $\sigma_\eta = 0.3$, and measurements are given by $y_k = \theta_k + \nu_k$ with $\sigma_\nu = 0.8$. **Panel A** shows the evolution of the one-dimensional state θ_k (true and estimated by a RTS with optimized noise covariances), and **panel B** the evolution of the measurements y_k. **Panel C** shows the negative log data likelihood $-\log q(y \mid \sigma_\eta, \sigma_\nu)$ as a function of the two noise standard deviations. Aside from identical values with inverted signs, in this case, the negative log likelihood has a single, global minimum at the true values 0.3 and 0.8 for σ_η and σ_ν.

References

[1] A. C. Aitken, "Iv.—on least squares and linear combination of observations," *Proceedings of the Royal Society of Edinburgh*, vol. 55, pp. 42–48, 1936.

[2] H. Akaike, "Information theory and an extension of the maximum likelihood principle," in *Proceeding of the Second International Symposium on Information Theory*, B. N. Petrov and F. Caski, Eds., Akademiai Kiado, Budapest, 1973, pp. 267–281. DOI: 10.1007/978-1-4612-1694-0_15.

[3] C. Andrieu, N. de Freitas, A. Doucet, and M. I. Jordan, "An introduction to MCMC for machine learning," *Machine Learning*, vol. 50, pp. 5–43, 2003. DOI: 10.1023/a:1020281327116.

[4] I. Arasaratnam and S. Haykin, "Discrete-time nonlinear filtering algorithms using Gauss-Hermite quadrature," *Proceedings of the IEEE*, vol. 95, no. 5, pp. 953–977, 2007. DOI: 10.1109/JPROC.2007.894705.

[5] ———, "Cubature Kalman filters," *IEEE Transactions on Automatic Control*, vol. 54, no. 6, pp. 1254–1269, 2009. DOI: 10.1109/TAC.2009.2019800.

[6] M. Arulampalam, S. Maskell, N. Gordon, and T. Clapp, "A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking," *IEEE Transactions on Signal Processing*, vol. 50, no. 2, pp. 174–188, 2002. DOI: 10.1109/78.978374.

[7] B. M. Bell, "The iterated Kalman smoother as a Gauss–Newton method," *SIAM Journal on Optimization*, vol. 4, no. 3, pp. 626–636, 1994. DOI: 10.1137/0804035.

[8] R. Berk *et al.*, "Working with misspecified regression models," *Journal of Quantitative Criminology*, vol. 34, no. 3, pp. 633–655, 2017. DOI: 10.1007/s10940-017-9348-7.

[9] P. J. Bickel *et al.*, "Regularization in statistics," *Test*, vol. 15, no. 2, pp. 271–344, 2006. DOI: 10.1007/bf02607055.

[10] P. Billingsley, *Probability and measure*, 3rd ed. New York, NY, USA: John Wiley & Sons, Inc., 1995.

[11] K. H. Brodersen *et al.*, "Inferring causal impact using Bayesian structural time-series models," *The Annals of Applied Statistics*, vol. 9, no. 1, 2015. DOI: 10.1214/14-aoas788.

[12] J. Byrd and Z. Lipton, "What is the effect of importance weighting in deep learning?" In *Proceedings of the 36th International Conference on Machine Learning*, K. Chaudhuri and R. Salakhutdinov, Eds., vol. 97, PMLR, 2019, pp. 872–881. [Online]. Available: https://proceedings.mlr.press/v97/byrd19a.html.

[13] C. Chen, X. Lin, Y. Huang, and G. Terejanu, "Approximate Bayesian neural network trained with ensemble Kalman filter," in *2019 International Joint Conference on Neural Networks (IJCNN)*, IEEE, 2019. DOI: 10.1109/ijcnn.2019.8851742.

[14] P. Diaconis and D. Freedman, "On the consistency of Bayes estimates," *The Annals of Statistics*, vol. 14, no. 1, pp. 1–26, 1986. [Online]. Available: https://www.jstor.org/stable/2241255.

[15] D. Draper, "Assessment and propagation of model uncertainty," *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 57, no. 1, pp. 45–97, 1995, ISSN: 00359246. [Online]. Available: http://www.jstor.org/stable/2346087.

[16] E. T. Jaynes, *Probability Theory: The Logic of Science*, G. L. Bretthorst, Ed. New York: Cambridge University Press, 2003.

[17] C. Field and B. Smith, "Robust estimation: A weighted maximum likelihood approach," *International Statistical Review / Revue Internationale de Statistique*, vol. 62, no. 3, p. 405, 1994. DOI: 10.2307/1403770.

[18] D. A. Freedman, *Statistical Models*. Cambridge University Press, 2009. DOI: 10.1017/cbo9780511815867.

[19] O. Frink, "Jordan measure and Riemann integration," *The Annals of Mathematics*, vol. 34, no. 3, p. 518, 1933. DOI: 10.2307/1968175.

[20] K. J. Friston, T. Parr, and B. de Vries, "The graphical brain: Belief propagation and active inference," *Network Neuroscience*, vol. 1, no. 4, pp. 381–414, 2017. DOI: 10.1162/netn_a_00018.

[21] C. F. Gauss, *Theory of the motion of the heavenly bodies moving about the sun in conic sections: A Translation of Theoria Motus*. Dover Publications, 1963, vol. 1056.

[22] G. H. Golub and C. Van Loan, "Total least squares," in *Lecture Notes in Mathematics*, Springer Berlin Heidelberg, 1979, pp. 69–76. DOI: 10.1007/bfb0098490.

[23] I. Guyon *et al.*, "Structural risk minimization for character recognition," *Advances in neural information processing systems*, vol. 4, pp. 471–479, 1991.

[24] W. K. Härdle, M. Müller, S. Sperlich, and A. Werwatz, *Nonparametric and semiparametric models*. Springer Science & Business Media, 2012.

[25] M. Hardt, B. Recht, and Y. Singer, "Train faster, generalize better: Stability of stochastic gradient descent," in *Proceedings of The 33rd International Conference on Machine Learning*, M. F. Balcan and K. Q. Weinberger, Eds., vol. 48, New York, New York, USA: PMLR, 2016, pp. 1225–1234. [Online]. Available: http://proceedings.mlr.press/v48/hardt16.html.

[26] F. Hayashi, *Econometrics*. Princeton University Press, 2000, ISBN: 9780691010182.

[27] R. Herbrich, "Minimising the Kullback–Leibler divergence," Microsoft Research, research rep., 2005. [Online]. Available: https://www.microsoft.com/en-us/research/publication/minimising-the-kullback-leibler-divergence/.

[28] C. Herzog, E. Petersen, and P. Rostalski, "Iterative approximate nonlinear inference via Gaussian message passing on factor graphs," *IEEE Control Systems Letters*, vol. 3, pp. 978–983, 4 2019. DOI: 10.1109/LCSYS.2019.2919260.

[29] J. A. Hoeting, D. Madigan, A. E. Raftery, and C. T. Volinsky, "Bayesian model averaging: A tutorial," *Statistical Science*, vol. 14, no. 4, pp. 382–417, 1999. DOI: 10.1214/ss/1009212519.

[30] P. J. Huber, "Robust regression: Asymptotics, conjectures and Monte Carlo," *The Annals of Statistics*, vol. 1, no. 5, pp. 799–821, 1973. [Online]. Available: https://www.jstor.org/stable/2958283.

[31] D. Jakubovitz, R. Giryes, and M. R. D. Rodrigues, "Generalization error in deep learning," in *Applied and Numerical Harmonic Analysis*, Springer International Publishing, 2019, pp. 153–193. DOI: 10.1007/978-3-319-73074-5_5.

[32] W. James and C. Stein, "Estimation with quadratic loss," in *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Contributions to the Theory of Statistics*, Berkeley, Calif.: University of California Press, 1961, pp. 361–379. [Online]. Available: https://projecteuclid.org/euclid.bsmsp/1200512173.

[33] R. E. Kalman, "A new approach to linear filtering and prediction problems," *Journal of Basic Engineering*, vol. 82, no. 1, pp. 35–45, 1960. DOI: 10.1115/1.3662552.

[34] R. E. Kalman and R. S. Bucy, "New results in linear filtering and prediction theory," *Journal of Basic Engineering*, vol. 83, no. 1, pp. 95–108, 1961. DOI: 10.1115/1.3658902.

[35] T. Kanamori, S. Hido, and M. Sugiyama, "A least-squares approach to direct importance estimation," *Journal of Machine Learning Research*, vol. 10, pp. 1391–1445, 2009. [Online]. Available: http://jmlr.org/papers/v10/kanamori09a.html.

[36] D. Kuhn, P. M. Esfahani, V. A. Nguyen, and S. Shafieezadeh-Abadeh, "Wasserstein distributionally robust optimization: Theory and applications in machine learning," in *Operations Research & Management Science in the Age of Analytics*, INFORMS, 2019, pp. 130–166. DOI: 10.1287/educ.2019.0198.

[37] L. Ljung, *System Identification: Theory for the User*, Second Edition. Prentice Hall, 1999.

[38] H.-A. Loeliger *et al.*, "On sparsity by NUV-EM, Gaussian message passing, and Kalman smoothing," in *2016 Information Theory and Applications Workshop (ITA)*, IEEE, 2016. DOI: 10.1109/ita.2016.7888168.

[39] U. von Luxburg and B. Schölkopf, "Statistical learning theory: Models, concepts, and results," in *Handbook of the History of Logic*, Elsevier, 2011, pp. 651–706. DOI: 10.1016/b978-0-444-52936-7.50016-1.

[40] S. Majumder *et al.*, "Statistical inference based on a new weighted likelihood approach," *Metrika*, vol. 84, no. 1, pp. 97–120, 2020. DOI: 10.1007/s00184-020-00778-y.

[41] T. P. Minka, "Expectation propagation for approximate Bayesian inference," *Proceedings of the Seventeenth Conference on Uncertainty in Artificial Intelligence*, 2001. arXiv: 1301.2294v1 [cs.AI]. [Online]. Available: https://dl.acm.org/doi/10.5555/2074022.2074067.

[42] K. P. Murphy, *Machine Learning: A Probabilistic Perspective*. Cambridge, Massachusetts, USA: The MIT Press, 2012.

[43] B. Neal *et al.*, "A modern take on the bias-variance tradeoff in neural networks," in *ICML workshop on Identifying and Understanding Deep Learning Phenomena*, Long Beach, CA, Jun. 15, 2019, 2019. [Online]. Available: https://arxiv.org/abs/1810.08591.

[44] M. Olson, A. Wyner, and R. Berk, "Modern neural networks generalize on small data sets," in *Advances in Neural Information Processing Systems 31*, S. Bengio *et al.*, Eds., Curran Associates, Inc., 2018, pp. 3619–3628. [Online]. Available: http://papers.nips.cc/paper/7620-modern-neural-networks-generalize-on-small-data-sets.pdf.

[45] A. V. Oppenheim and G. C. Verghese, *Signals, systems and inference*. Pearson, 2015.

[46] E. Petersen, C. Hoffmann, and P. Rostalski, "On approximate nonlinear Gaussian message passing on factor graphs," in *IEEE Statistical Signal Processing Workshop*, Freiburg, Germany, 2018. DOI: 10.1109/SSP.2018.8450699.

[47] J. Quiñonero-Candela, M. Sugiyama, N. D. Lawrence, and A. Schwaighofer, *Dataset shift in machine learning*. MIT Press, 2009.

[48] H. E. Rauch, F. Tung, and C. T. Striebel, "Maximum likelihood estimates of linear dynamic systems," *AIAA Journal*, vol. 3, no. 8, pp. 1445–1450, 1965. DOI: 10.2514/3.3166.

[49] S. Roweis and Z. Ghahramani, "A unifying review of linear Gaussian models," *Neural Computation*, vol. 11, no. 2, pp. 305–345, 1999. DOI: 10.1162/089976699300016674.

[50] S. Särkkä, *Bayesian Filtering and Smoothing*. Cambridge University Press, Cambridge, United Kingdom, 2013, ISBN: 978-1-107-03065-7. [Online]. Available: https://users.aalto.fi/~ssarkka/pub/cup_book_online_20131111.pdf.

[51] J. Schoukens and L. Ljung, "Nonlinear system identification: A user-oriented roadmap," *IEEE Control Systems Magazine*, vol. 39, no. 6, pp. 28–99, 2019. DOI: https://doi.org/10.1109/MCS.2019.2938121.

[52] H. Shimodaira, "Improving predictive inference under covariate shift by weighting the log-likelihood function," *Journal of Statistical Planning and Inference*, vol. 90, no. 2, pp. 227–244, 2000. DOI: 10.1016/s0378-3758(00)00115-4.

[53] D. Simon, "Kalman filtering with state constraints: A survey of linear and nonlinear algorithms," *IET Control Theory & Applications*, vol. 4, no. 8, pp. 1303–1318, 2010. DOI: 10.1049/iet-cta.2009.0032.

[54] D. Simon and T. L. Chia, "Kalman filtering with state equality constraints," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 38, no. 1, pp. 128–136, 2002. DOI: 10.1109/7.993234.

[55] D. Simon, *Optimal State Estimation*. Hoboken, New Jersey, USA: John Wiley & Sons, Inc., 2006.

[56] H. W. Sorenson, "Least-squares estimation: From Gauss to Kalman," *IEEE Spectrum*, vol. 7, no. 7, pp. 63–68, 1970. DOI: 10.1109/mspec.1970.5213471.

[57] P. Sprent, *Models in regression and related topics*, ser. Methuen's monographs on applied probability and statistics. Egmont UK Ltd, London, 1969, ISBN: 9780416148305.

[58] R. L. Stratonovich, "Conditional Markov processes," *Theory of Probability and Its Applications*, vol. 5, pp. 156–178, 1960. DOI: 10.1137/1105015.

[59] M. Sugiyama and M. Kawanabe, *Machine Learning in Non-Stationary Environments*. Cambridge, Massachusetts: The MIT Press, 2012.

[60] M. Sugiyama, M. Krauledat, and K.-R. Müller, "Covariate shift adaptation by importance weighted cross validation," *Journal of Machine Learning Research*, vol. 8, pp. 985–1005, 2007. [Online]. Available: https://dl.acm.org/doi/10.5555/1314498.1390324.

[61] P. Swerling, "A proposed stagewise differential correction procedure for satellite tracking and prediction," *Rand Corporation*, 1958.

[62] F. Tronarp, A. F. Garcia-Fernandez, and S. Sarkka, "Iterative filtering and smoothing in nonlinear and non-Gaussian systems using conditional moments," *IEEE Signal Processing Letters*, vol. 25, no. 3, pp. 408–412, 2018. DOI: 10.1109/lsp.2018.2794767.

[63] R. Van der Merwe, "Sigma-point Kalman filters for probabilistic inference in dynamic state-space models," Ph.D. dissertation, 2004.

[64] A. W. Van Der Vaart, *Asymptotic Statistics*. Cambridge University Press, 1998.

[65] S. Van Huffel and P. Lemmerling, Eds., *Total Least Squares and Errors-in-Variables Modeling*. Springer Netherlands, 2002. DOI: 10.1007/978-94-017-3552-0.

[66] V. Vapnik, "An overview of statistical learning theory," *IEEE Transactions on Neural Networks*, vol. 10, no. 5, pp. 988–999, 1999. DOI: 10.1109/72.788640.

[67] V. Vapnik and A. Chervonenkis, *Theory of Pattern Recognition*. Moscow: Nauka, 1974.

[68] V. Vapnik, "Principles of risk minimization for learning theory," in *Advances in neural information processing systems*, 1992, pp. 831–838. [Online]. Available: https://dl.acm.org/doi/10.5555/2986916.2987018.

[69] V. N. Vapnik, *Statistical Learning Theory*. Wiley, 1998, ISBN: 978-0-471-03003-4.

[70] F. Wadehn, L. Bruderer, V. Sahdeva, and H.-A. Loeliger, "New square-root and diagonalized Kalman smoothers," in *2016 54th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, IEEE, 2016. DOI: 10.1109/allerton.2016.7852382.

[71] F. Wadehn *et al.*, "Outlier-insensitive Kalman smoothing and marginal message passing," in *2016 24th European Signal Processing Conference (EUSIPCO)*, IEEE, 2016. DOI: 10.1109/eusipco.2016.7760447.

[72] M. J. Wainwright and M. I. Jordan, "Graphical models, exponential families, and variational inference," *Foundations and Trends® in Machine Learning*, vol. 1, no. 1–2, pp. 1–305, 2007. DOI: 10.1561/2200000001.

[73] E. A. Wan and R. van der Merwe, "The unscented Kalman filter for nonlinear estimation," in *Adaptive Systems for Signal Processing, Communications, and Control Symposium*, 2000, pp. 153–158. DOI: 10.1109/ASSPCC.2000.882463.

[74] X. Wang, C. van Eeden, and J. V. Zidek, "Asymptotic properties of maximum weighted likelihood estimators," *Journal of Statistical Planning and Inference*, vol. 119, no. 1, pp. 37–54, 2004. DOI: 10.1016/s0378-3758(02)00410-x.

[75] T. Wansbeek and E. Meijer, "Measurement error and latent variables," in *A Companion to Theoretical Econometrics*, B. H. Baltagi, Ed. Oxford: Blackwell Publishers Ltd, 2001, pp. 162–179.

[76] H. White, "Maximum likelihood estimation of misspecified models," *Econometrica*, vol. 50, no. 1, p. 1, 1982. DOI: 10.2307/1912526.

[77] W. Wilkinson, P. Chang, M. Andersen, and A. Solin, "State space expectation propagation: Efficient inference schemes for temporal Gaussian processes," in *Proceedings of the 37th International Conference on Machine Learning*, H. Daumé III and A. Singh, Eds., vol. 119, PMLR, 2020, pp. 10 270–10 281. [Online]. Available: http://proceedings.mlr.press/v119/wilkinson20a.html.

[78] R. Wilson and L. Finkel, "A neural implementation of the Kalman filter," in *Advances in Neural Information Processing Systems*, Y. Bengio *et al.*, Eds., vol. 22, Curran Associates, Inc., 2009. [Online]. Available: https://proceedings.neurips.cc/paper/2009/file/6d0f846348a856321729a2f36734d1a7-Paper.pdf.

[79] J. M. Wooldridge, *Introductory Econometrics*. South-Western, Cengage Learning, 2012.

[80] C. Zhang *et al.*, "Understanding deep learning (still) requires rethinking generalization," *Communications of the ACM*, vol. 64, no. 3, pp. 107–115, 2021. DOI: 10.1145/3446776.

Chapter 3

Preliminaries B: physiology, electromyography, and respiratory support

All theory, dear friend, is gray,
And green is the lustrous tree of life.

(Mephistopheles, as quoted in Goethe's Faust*)*

Contents

THE PHYSIOLOGY of human skeletal muscles and the respiratory system, electromyography, and mechanical ventilation are all subjects that fill whole books. Only a brief overview can be provided here; the aim is to provide the background information necessary for the reader to follow along with the following chapters. The discussion is, of course, far from comprehensive, and the interested reader will find pointers to more thorough treatments provided in the corresponding sections. The chapter begins with an introduction to muscular physiology and surface electromyography as a measurement of muscular activity. The second part of the chapter contains an introduction to the physiology of the respiratory system, followed by a brief discussion of mechanical ventilation as a means to support said system. Finally, section 3.3 provides some information on two studies, one in healthy volunteers and one in intensive care patients, data from which will be utilized in later chapters of this thesis.

3.1 Skeletal muscles

This section first provides an introduction to the basic physiology of human skeletal muscles — which include the respiratory muscles — in section 3.1.1, before section 3.1.2 then treats the foundations of surface electromyography.

3.1.1 Physiology

Figure 3.1 illustrates the main anatomical elements involved in movement generation. To initiate muscular contraction, the motor cortex, a part of the central nervous system (CNS) chiefly responsible for controlling body movement, activates a *motor neuron* (or, typically, multiple motor neurons). Each motor neuron *innervates* a number of muscle fibers, together with which it forms a *motor unit (MU)*. Once it is activated by the CNS, the motor neuron propagates an *action potential* along its axon, thereby activating all of the muscle fibers it innervates at once. The depolarization caused by the motor neuron's action potential at the synaptic terminals on the muscle fibers triggers the release of two new action potentials within each muscle fiber, propagating in opposite directions along the muscle fiber. These two action potentials trigger the release of Calcium ions within the fiber and, thus, the shortening of the sarcomeres, ultimately resulting in a shortening of the muscle: a *contraction*. Figure 3.2 shows a detailed schematic representation of this process. There is a brief *electromechanical delay (EMD)* between the action potential and the resulting muscle fiber force twitch and, due to the different time constants of the action potential propagation and force generation processes, the force twitch also lasts much longer than the action potential. Figure 3.3 shows the *excitation-contraction coupling*, i.e., the relationship between electrical and mechanical twitches, of two different types of muscle fibers. Muscles typically consist of a few ten to several hundred motor units (MUs), each of which in turn consists of between less than ten and more than a thousand muscle fibers [57] (in addition to its single motor neuron). The distribution of these MUs' electromechanical properties, as well as the patterns of their activation over time, are intricately organized and have been the subject of much research [57]. Arguably, the two main organizing paradigms of motor units in a muscle are, firstly, the presence of a *common drive* to all MUs and, secondly, the *size principle*.

Motor Neuron

Motor Unit

Action Potential

Electromechanical Delay
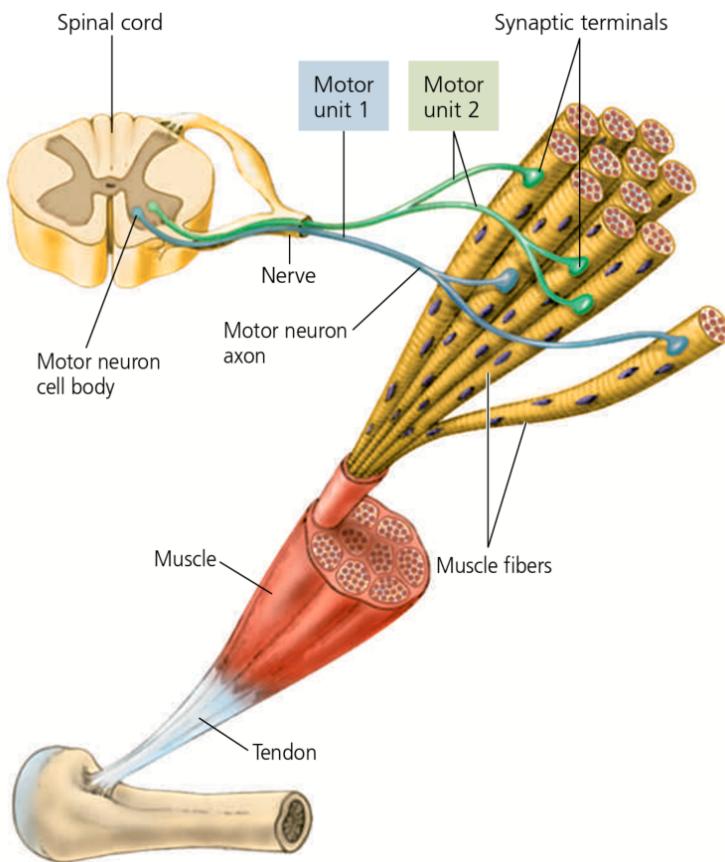
Figure 3.1 – Essential anatomical elements involved in human movement generation. A *motor unit* is defined by the combination of a *motor neuron* and all muscle fibers it innervates. Reproduced with permission from Urry [130].

Common Drive

Onion Skin Phenomenon

Muscle Synergies

In a series of experimental studies, motor units in a muscle — and even in synergistic muscle *pairs* [75] — have been observed to be driven by a shared neural *common drive* signal that governs all of their activity [27, 28, 36, 106].[1] Whether a MU *activates* at a given level of common drive, however, depends on the MU's individual *recruitment threshold*, which differs widely across a muscle's MU pool, along with many other properties. Thus, MUs are recruited in a particular order, which appears to be preserved even during fatiguing contractions [1, 124]. Recruitment and de-recruitment order are inversely related, i.e., those (typically large) motor units that are recruited late — i.e., at high activation levels — are the first to be de-recruited again. In addition, the so-called *onion skin phenomenon* refers to the frequent observation that earlier recruited motor units usually exhibit higher firing rates than later-recruited ones [28]. Note that it is a matter of current scientific debate how broadly this principle really applies [56, 67, 103, 106]. Finally, *muscle synergies* represent an important higher-level extension of the common drive concept. It has long been hypothesized that in many instances, the CNS does not control muscles individually but rather drives the activity

---

[1] Deviations from this pattern have been observed, e.g., for large muscles in which different muscle regions are controlled independently by the CNS [18, 86].
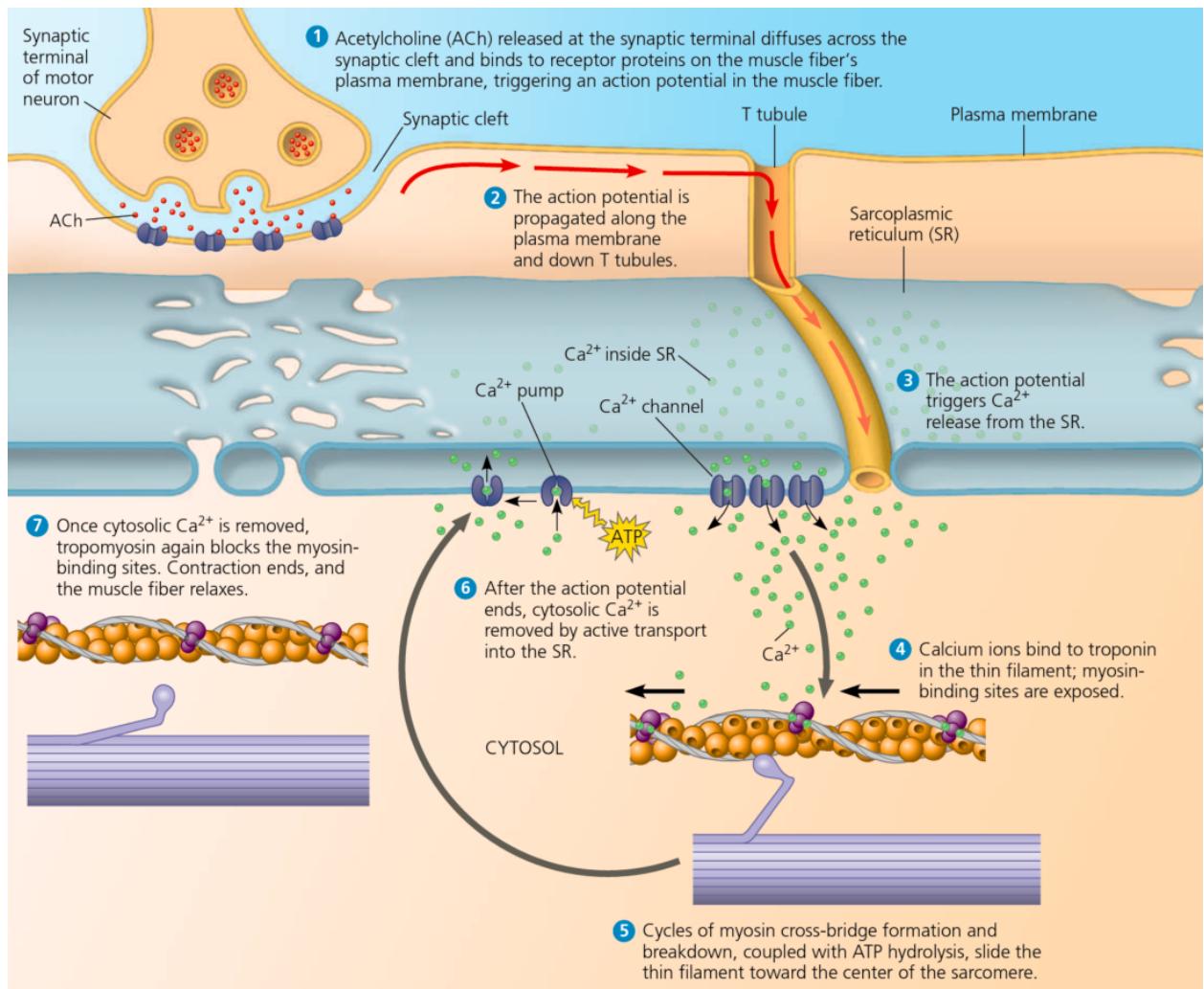
Figure 3.2 – A schematic representation of action potential generation and propagation, force generation, and relaxation in a muscle fiber. Reproduced with permission from Urry [130].
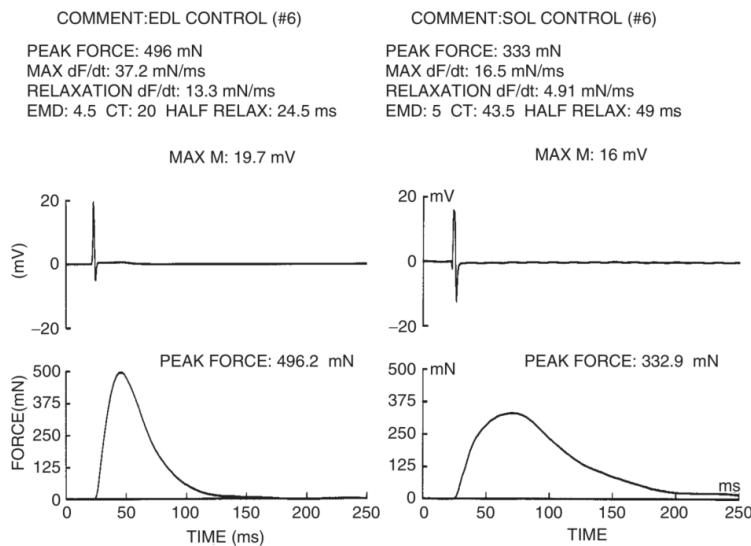
Figure 3.3 – Excitation-contraction coupling of typical fast-twitch and slow-twitch muscle fibers. The top row shows the electrical field fluctuation elicited by a single twitch of the fiber, the bottom row shows the generated force twitch $f(t)$. The left column shows data from a typical fast-twitch fiber (extensor digitorum longus, EDL), the right column data from a typical slow-twitch fiber (soleus, SOL), both from an isolated rat muscle. EMD: electromechanical delay (between EMG and force onset), CT: contraction time (onset to force peak). Reproduced with permission from Moritani *et al.* [97].

of synergistic muscle groups together, using what has been called *muscle synergies*. While some of the neurophysiological details are still a subject of debate, the existence of muscle synergies has been confirmed in a large number of studies [16, 25, 44, 122, 126, 129]. From a control-theoretic point of view, muscle synergies appear to serve as a useful abstraction layer, as evidenced by the successful application of muscle synergy-based proportional prosthesis control [12, 69]. Finally, note that there is, of course, a huge body of literature investigating the general principles using which the CNS and, in particular, the motor cortex perform muscle control [30, 34, 44, 66, 99, 100, 102, 128, 135].

Practically all MU and fiber properties such as EMG twitch amplitude, force twitch amplitude, and recruitment threshold are correlated with MU *size*, where the latter is often defined as the number N_i^{fib} of fibers innervated by MU i [57–59]. In other words: large MUs innervate a large number of muscle fibers, which elicit stronger and faster force and EMG twitch responses, fatigue faster, and are only recruited at higher activation levels than Size Principle smaller MUs. This principle is termed the *size principle*. Empirical studies have shown that the distribution of MU recruitment thresholds — and, thus, motor unit sizes — is highly skewed, with many small MUs being recruited at low activation levels, and a few large MUs being recruited only at high activation levels [57, 112, 132]. Both motor units and muscle fibers have traditionally been grouped into distinct types (fast-twitch, slow-twitch, fatigue-resistant) in the physiological literature [92], with the consequence that various modeling studies have described MUs and fibers as belonging to either of a few distinct classes, with properties drawn from corresponding distributions [4, 24, 113]. It seems, however, that the distributions of MU and muscle fiber properties are actually *continuous*
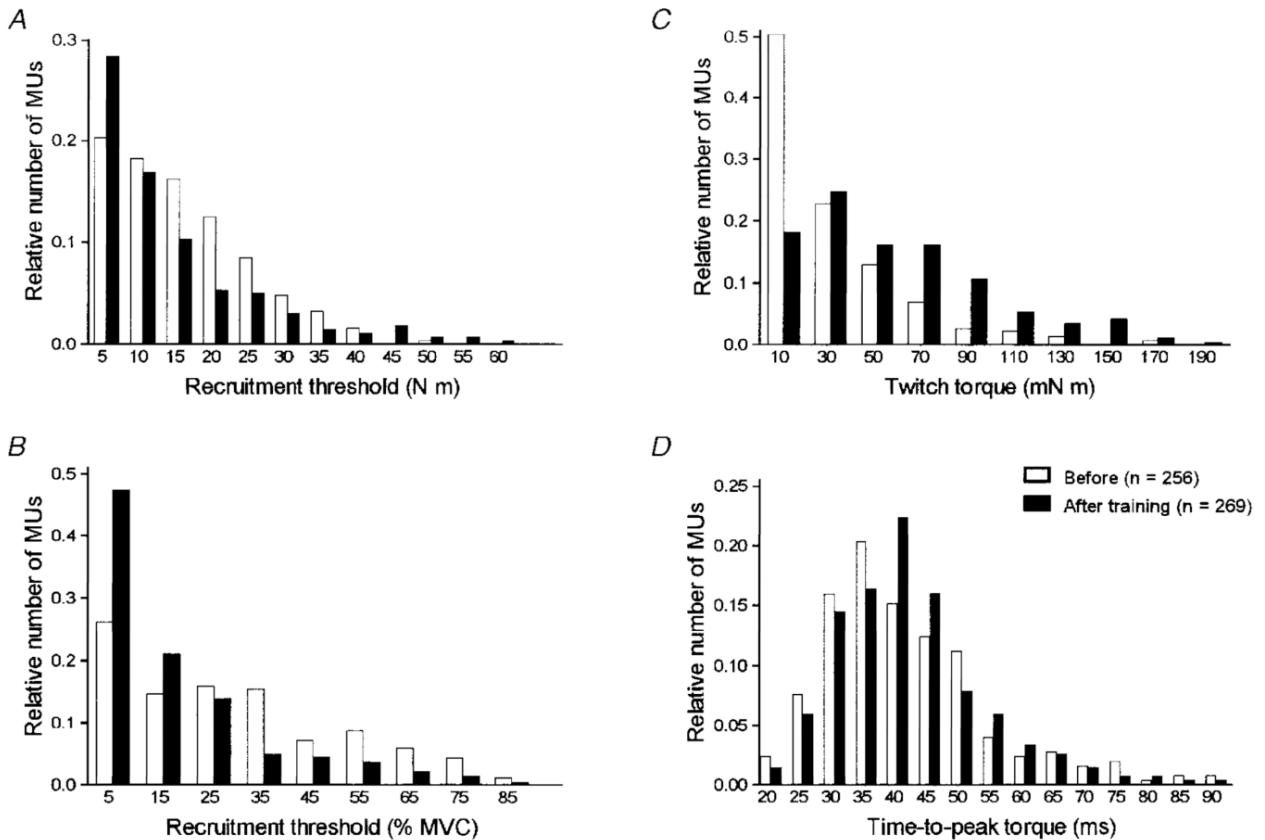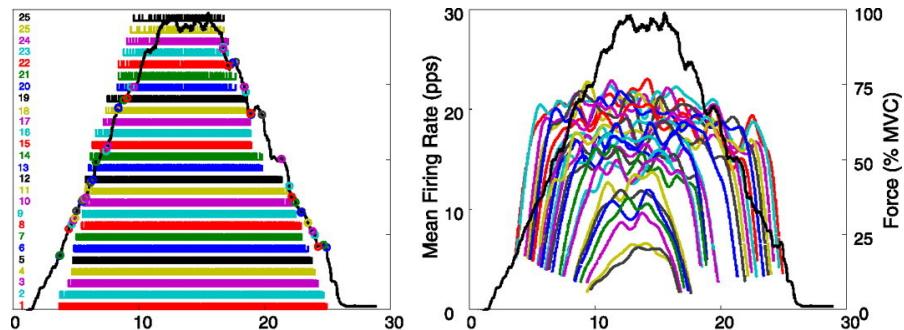
Figure 3.4 – Statistical distributions of motor unit properties in the tibialis anterior in five human subjects, before and after 12 weeks of dynamic training. No distinct clusters of MU properties are visible — in other words, the distributions of these properties are *continuous* and no clearly distinguishable MU types can be identified. Reproduced with permission from Van Cutsem *et al.* [131].

without clear clusters (see fig. 3.4 for an example), thus not warranting a categorical model but rather a model with continuously distributed properties [35, 47, 57, 111]. Figure 3.5 shows an exemplary measurement that is purported to illustrate all of the principles of motor unit pool organization discussed above.

3.1.2 Surface electromyography

Surface electromyography (surface EMG, or sEMG) denotes the measurement of the electrical fields generated by action potentials propagating along muscle fibers during muscle activation. The measurement is performed by placing electrodes on the skin surface, which are connected to an electrical amplifier for further processing. A surface electromyogram provides a noninvasive measure of neural muscle activation and is widely used in fields such as sports science, clinical diagnosis, prosthesis control, ergonomics, and rehabilitation robotics [39, 92]. It has been used for monitoring the activity of respiratory muscles for six decades by means of an esophageal catheter [80, 105], and for more than

A

B
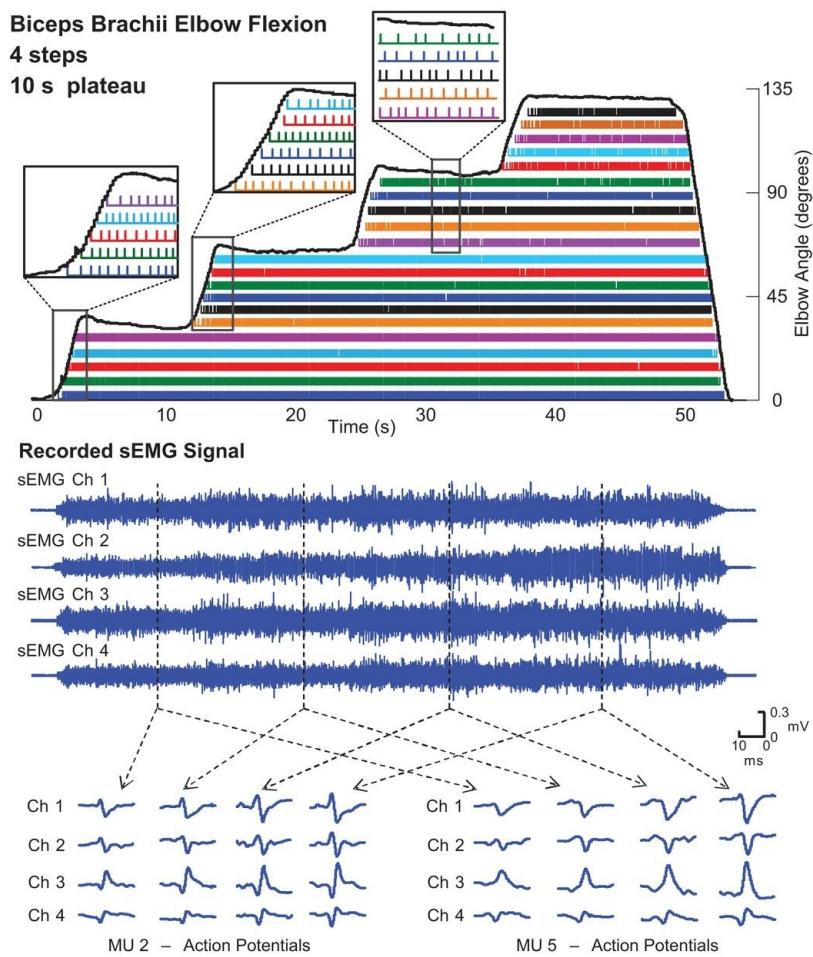
Figure 3.5 – **Panel A:** Force measurements and identified MUAP trains during an isometric constant-force contraction of the vastus lateralis (VL) muscle. Each color (and each row in the left graph) represents one MU's observed firing instants. It can be observed that the MU rate coding conforms with the onion-skin principle, i.e., later-recruited MUs have lower firing rates than MUs recruited earlier. (pps = pulses per second, MVC = maximum voluntary contraction.) Reproduced with permission from De Luca and Contessa [26]. **Panel B:** Biceps brachii surface EMG measurements and analysis of a step-wise elbow flexion movement. The top panel shows the results of an algorithm to identify the discharge times and action potential shapes of different MUs involved in this movement. It can be observed that recruitment order and de-recruitment order are inversely related. Reproduced with permission from De Luca *et al.* [29]. Note that the method used to identify the activity of individual MUs in both studies has been the subject of scientific debate [41].
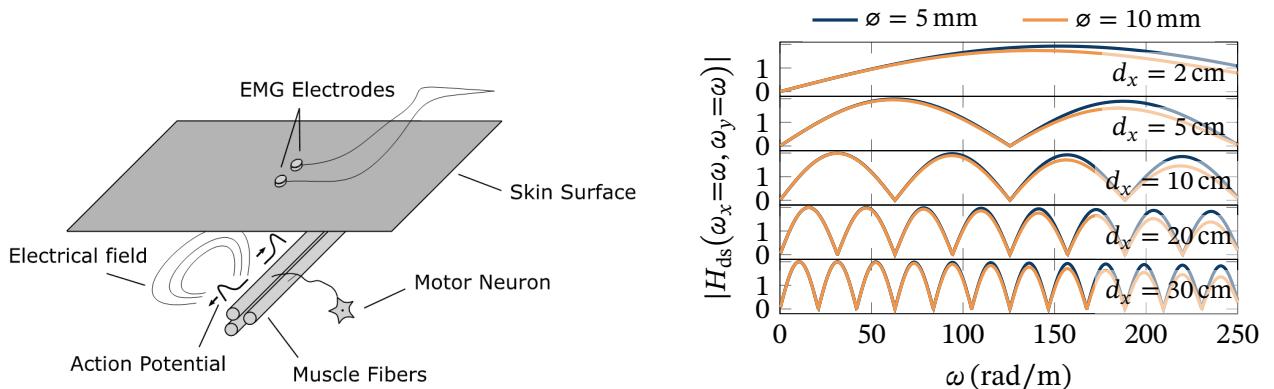
Figure 3.6 – A model of the influence of the electrode detection system. **Left:** a schematic illustration of the assumed geometrical model. At each moment, action potentials propagating along muscle fibers give rise to a three-dimensional potential distribution throughout the surrounding biological tissue. This, in turn, results in a two-dimensional potential distribution on the skin surface (here: the (x, y) plane). A single-differential EMG signal is obtained from the difference between the potentials at two skin electrodes. **Right:** a model of the spatial filtering introduced by the use of non-point electrodes and differential measurements, using eqs. (3.2) to (3.4), for two circular electrodes of varying diameter ø and at varying inter-electrode distances d_x. The graphs show the values of the spatial transfer function of the detection system for $\omega = \omega_x = \omega_y$.

three decades using electrodes placed on the skin surface [33, 98]. The latter is of current research interest [11, 72] and represents the main focus of this thesis. As such, a more detailed overview of existing research on monitoring respiratory effort using surface electrodes will be given in section 7.1.

Figure 3.6 (left) illustrates the basic setup used for surface EMG measurements, and fig. 3.7 shows both simulated and measured (monopolar) sEMG signals elicited by a single contraction of a single motor unit. Two distinct phases can be observed [91]:

1. The action potential propagates along the fiber, with a propagation velocity of roughly 3 m/s to 5 m/s [87]. This is called the *propagating component* of the signal.

2. The action potential is extinguished once it reaches the fiber end. At this moment, a similar peak can be observed all across the fiber at the same time. This is called the *end-of-fiber component* of the signal.

Considering the frequency characteristics of the two components, the first is rather low-frequent temporally (it is present for the whole duration of the action potential) and rather high-frequent spatially (it is only observable at two particular positions along the fiber at each moment), while the second is rather high-frequent temporally (it is only present for one short moment) and rather low-frequent spatially (it is observed similarly across the whole fiber). Since the biological tissues surrounding the muscle act as a spatial low-pass filter, the propagating component can only be observed directly above the muscle, while the end-of-fiber component is attenuated much less with increasing distance [31, 39]. For this reason, *crosstalk* from muscles located further away from the electrodes consists almost exclusively
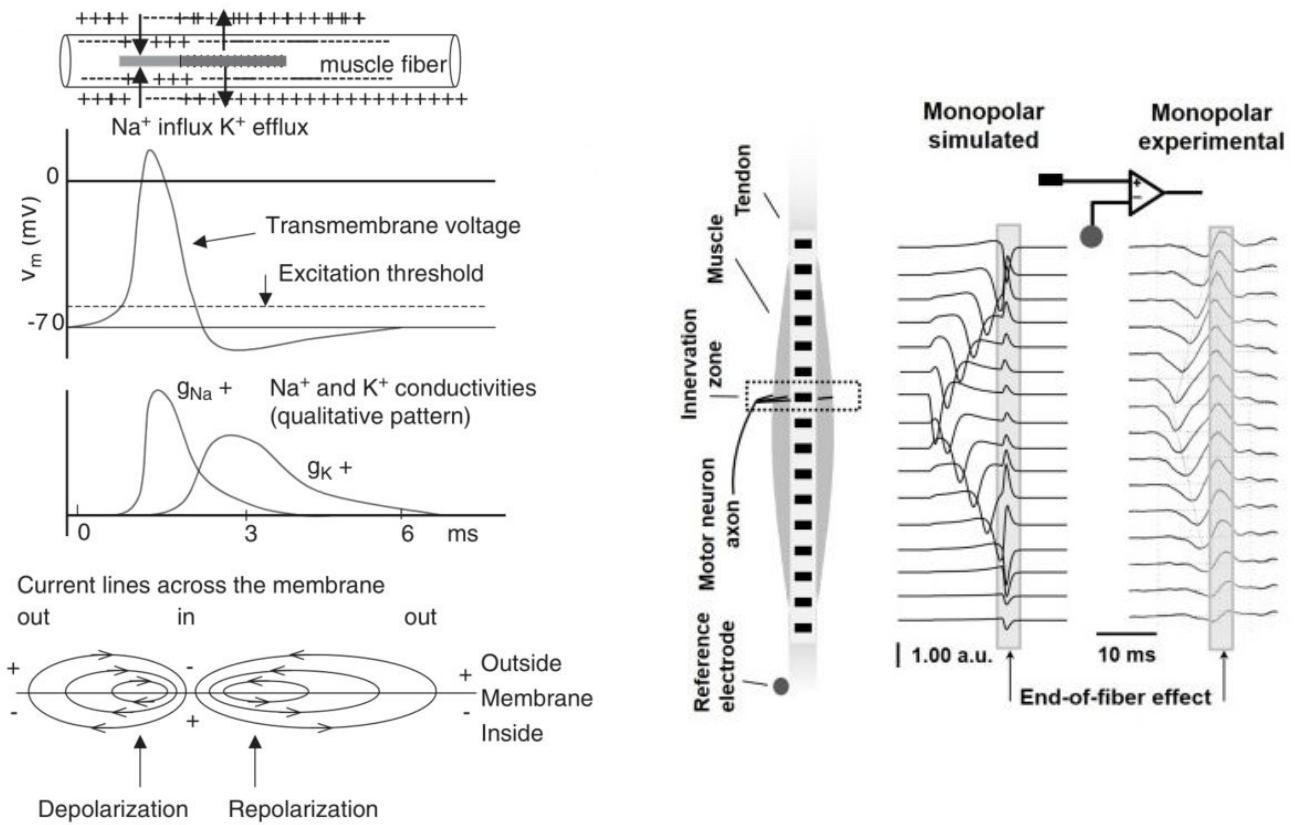
Figure 3.7 – **Left:** a snapshot of action potential propagation along a muscle fiber. The initial depolarization at the neuromuscular junction (NMJ) causes a continuous chemical reaction that propagates the action potential along the fiber, causing ion currents to exit and enter the fiber at different locations. These currents induce an electrical field in the surrounding biological tissue, which acts as a volume conductor. It is this electrical field that is measured by means of (surface) electromyography. Reproduced with permission from Moritani *et al.* [97]. **Right:** simulated and measured monopolar surface potentials elicited by a single motor unit action potential (MUAP), measured by a series of surface electrodes placed at different positions along the muscle. Reproduced with permission from Merletti and Muceli [91]; for details see there.

of end-of-fiber components [31, 39, 91]; an insight which has profound consequences for proper signal processing and analysis [31].

Whereas fig. 3.7 (right) shows a *single* motor unit action potential (MUAP), an EMG measurement usually represents the superposition of hundreds of MUAPs. (Recall from the previous section that each muscle consists of dozens of motor units (MUs) and that each recruited MU activates with a frequency of a few up to several dozen Hz.) Mathematically, the superposition is often [62] modeled linearly as

$$\text{EMG}(t) = \sum_{i=1}^{N^{\text{MU}}} \sum_{j=1}^{N_i^{AP}} \text{MUAP}_i(t - t_{ij}),$$

where t_{ij} denotes the beginning of action potential j of MU i. Similarly, the generated muscle force $F(t)$ is given by the superposition of all the individual force twitches f_i (of which fig. 3.3 shows a single one) generated by each active MU i at each activation j; this is usually similarly modeled as

$$F(t) = \sum_{i=1}^{N^{\text{MU}}} F_i(t) = \sum_{i=1}^{N^{\text{MU}}} \sum_{j=1}^{N_i^{AP}} f_i(t - t_{ij}).$$

Figure 3.5B shows an example of surface EMG measurements from different electrode positions and the corresponding force signal of the biceps brachii muscle.

The relationship between a surface EMG signal and the corresponding generated muscle force is complex and depends on a number of both stationary and non-stationary factors. As illustrated in fig. 3.3 and discussed in the previous section, the amplitude and morphology of the generated EMG and force twitches vary between individual fibers. The occurrence of *amplitude cancellation* in the EMG signal — the effect that the action potential of one MU negates the effect of another MU's simultaneous action potential on the surface potential — influences the EMG-force relationship of different muscles [43]. The efficiency of the force-generating process, also called the *neuromuscular efficiency*, varies over time since it depends on the current length of the muscle fibers [61, 137], the contraction velocity [61, 137], and whether the muscle is *fatigued* due to previous contractions [87]. The EMG-force relationship is further influenced by the *electrical transmission path* between the muscle fibers and the recording electrode, which depends on the location and pinnation angle of the muscle [93], the number and arrangement of the electrodes (as will be discussed below in detail), the type and shape of electrodes, the electrode-skin impedance, the conductivity of the intermediate media such as muscle, fat, and skin tissue. Finally, in almost all cases, it is not the force generated by the muscle that is of interest, but rather some other force that is related to the generated muscle force via bodily biomechanics. As an example, in the case of biceps brachii surface EMG measurements, one is usually interested in the lifting force generated at the wrist, which is related to the force generated by the biceps via the elbow joint and depends on biomechanical factors such as the current elbow angle and the length of the forearm.

Neuromuscular Efficiency

From the previous paragraph, it is apparent that the EMG-force relationship is complex and time-varying. Nevertheless, in many practical applications, the relationship between the two signals can be approximated using a very simple heuristic: an *envelope signal* $\overline{\text{EMG}}(t)$ is calculated from the raw EMG signal, usually by applying some kind of moving average or

Envelope Signal

moving median filter to the *rectified* EMG signal, and then it is assumed that

$$F(t) \propto \overline{\mathrm{EMG}}(t),$$

or, in other words, that there is a simple scaling factor between the two signals. This simple heuristic has been successfully employed and confirmed in a number of studies for a number of different muscles and recording settings [5, 7, 104], although there are muscles for which the relationship is known to be nonlinear [5, 7, 136]. Note that when using this method, the way the envelope signal is calculated implicitly defines an assumption about the dynamic EMG-force transfer function. (Due to the rectifying step, this transfer function will usually be nonlinear.) More complex models of the EMG-force relationship have also been proposed, of course, such as nonlinear relationships between the envelope and force signals [79, 136] and adaptive envelope calculation methods [21, 23]. On the other hand, many applications of surface EMG, e.g., in prosthesis control, do not attempt to estimate the generated muscle force at all. Instead, in these applications, it is usually the aim to identify recurring EMG *patterns*, and then to control the prosthesis based on pattern recognition algorithms [48, 68]. If, however, a prosthesis is supposed to be controllable *continuously* across different force levels, then pattern recognition approaches — which necessarily discretize both patient and prosthesis action — are no longer sufficient, and continuous activation estimation methods are required [46, 70, 138]. Note that if there is a continuous feedback control loop between the user and the assistive device (which, among other things, requires the user to be awake and responsive), the requirements on the precision of the activation estimation procedure may not be high [71].

Electrode Configuration   The choice of an *electrode configuration* is a supremely important step in the design process of any sEMG application, with both the type of electrode to be used as well as the number and geometrical arrangement of electrodes being deciding factors. For the latter, *differential measurements* are a key element to understand. As opposed to *monopolar* measurements, which use only a single measurement electrode in addition to the (always required) reference electrode, single differential measurements employ (at least) a pair of electrodes placed at a distance to one another. The measurement signal is then given by the difference of the signals measured by these two electrodes. The effect of such a configuration is that signal components of low spatial variation are suppressed since these are similarly present in both electrode signals — examples include cardiac and movement artifacts, baseline wander, and crosstalk from other muscles. (This effect will be discussed in more technical detail below.) The same effect is often described as differential measurements having a lower *pickup volume* or *detection volume* than monopolar measurements, meaning that the muscle volume from which activity is picked up by the measurement configuration is smaller [91]. Finally, note that many detection systems more complex than a single differential electrode pair have, of course, been proposed and analyzed [40, 93, 101], with many applications nowadays using high-density electrode grids [2, 63, 73, 76, 83, 93, 123, 133].

Mathematically, the effect of any detection system, including the influence of the number, arrangement, size, shape, and type of the electrodes, can be modeled as a *spatial* transfer function $H_{\mathrm{ds}}(\omega_x, \omega_y)$ which describes how, at each instant in time, the 2-D electrical potential distribution $\phi(x, y, t)$ on the subject's skin is filtered spatially before a point measurement

$$\mathrm{EMG}(t) = \phi(x, y, t) \underset{x,y}{*} h_{\mathrm{ds}}(x, y)\Big|_{x=x_{\mathrm{ea}}, y=y_{\mathrm{ea}}} \tag{3.1}$$

is taken by the measurement setup [37]. In eq. (3.1), the $*_{x,y}$ operator denotes two-dimensional convolution, x_{ec} and y_{ec} specify the position of the electrode arrangement on the skin, and

$$h_{ds}(x,y) = \mathcal{F}_{x,y}^{-1}\{H_{ele}(\omega_x, \omega_y) \cdot H_{ea}(\omega_x, \omega_y)\} <$$

models the two-dimensional impulse response of the detection system, resulting from the combined effects of the electrodes (modeled by $H_{ele}(\omega_x, \omega_y)$) and the electrode arrangement (modeled by $H_{ea}(\omega_x, \omega_y)$). Regarding the latter, consider a regular grid of $N^x \times N^y$ electrodes, aligned with the x and y axes,[2] with interelectrode distances d_x and d_z, respectively, where $N^x = N_a^x + N_b^x + 1$ and $N^y = N_a^y + N_b^y + 1$. The variables subscripted by a and b denote the number of electrodes on the two sides of an arbitrarily chosen reference electrode. Assuming identical electrodes and assigning individual weights $w_{k\ell}$ to each of them, the spatial transfer function of the electrode arrangement is given by [37]

$$H_{ea}(\omega_x, \omega_y) = \sum_{k=-N_a^x}^{N_b^x} \sum_{\ell=-N_a^y}^{N_b^y} w_{k\ell} e^{-i\omega_x k d_x} e^{-i\omega_y \ell d_y}. \tag{3.2}$$

Regarding the modeling of individual electrodes, a large body of research is available [88, 90]. A simple, shape-agnostic electrode model is to assume that the sampled potential measurement $\phi_{ele}(t)$ corresponds to the average potential on the skin surface below the electrode [37], i.e.,

$$\phi_{ele}(t) = \frac{1}{|\Omega_{ele}|} \iint\limits_{\Omega_{ele}} \phi(x,y,t)\, dx, y, \tag{3.3}$$

where Ω_{ele} denotes the region under the electrode, and $|\Omega_{ele}|$ its surface area. For specific electrode shapes, the resulting spatial transfer function can be computed, e.g.,

$$h_{ele}(\omega_x, \omega_y) = \begin{cases} 2\dfrac{J_1\left(r\sqrt{\omega_x^2+\omega_y^2}\right)}{r\sqrt{\omega_x^2+\omega_y^2}} & \text{if} \quad (\omega_x, \omega_y) \neq (0,0) \\ 1 & \text{if} \quad (\omega_x, \omega_y) = (0,0) \end{cases} \tag{3.4}$$

for a circular electrode with radius r [37], where J_1 denotes the first-order Bessel function of the first kind. Note, however, that the very basic electrode model eq. (3.4) neglects many physical effects such as resistive and capacitive properties of electrodes and the fact that the presence of the electrode and the connected amplifier also influences the potential distribution in the biological tissue. Figure 3.6 (right) uses the models discussed above to analyze the spatial filtering effect of several different detection system setups. Two main observations can be made:

1. The differential measurement setup introduces so-called *spectral dips* at spatial frequencies $\omega = 2\pi k/d_x$, $k \in \mathbb{N}$ (in the case of a single differential pair of electrodes at distance d_x). This phenomenon is well-known and studied in the literature [37, 78, 91, 141]. Note that spectral dips may also occur as a result of the spatial filtering characteristics of the biological tissue itself [37]. The usual recommendation is to use

---

[2]Rotated detection systems can, of course, be described very similarly. Refer to Farina and Merletti [37] for details.

a small inter-electrode distance d_x to prevent the occurrence of spectral dips within the relevant frequency range [91].[3]

2. Differential setups with a small inter-electrode distance d_x act as a strong spatial high-pass filter, suppressing spatially low-frequent surface potentials.

The latter observation is the technical justification for the claim made above that differential measurements have a lower *detection volume* and suppress spatially low-frequent artifacts such as cardiac interference, movement artifacts, crosstalk from other muscles, and more.

Signal whitening is a common preprocessing step in surface EMG analysis and demonstrably beneficial in many applications [20, 22, 42]. The underlying idea is that removing temporal correlations from the signal yields a more instantaneous measure of muscular activity. This is usually achieved by calculating a parametric estimate of the power spectral density (PSD) of the signal, and then filtering the signal with the *inverse* of that parametric estimate — yielding a signal with a flat ("white") PSD.[4] Figure 3.8 shows an example of this operation in practice.

Impressive progress has been made over last decades in the recording, processing, and analysis of surface EMG measurements. It has been noted, however, that many of these technological advances have not been translated into clinical products [45]. Recently, it has been hypothesized that this may be primarily due to cultural and organizational barriers [19, 82, 95]. Interestingly, one of the mentioned potential barriers is that the surface EMG is mostly useful in clinical situations which are not life-threatening, thus possibly limiting research funding and practitioner interest [82]. This is, however, certainly not the case for the application discussed in the present thesis.

Much more detailed background information about surface EMG physiology, modeling, and signal processing can be found in the excellent recent tutorial papers by Merletti and Muceli [91] and Merletti and Cerone [89], and in the two seminal textbooks by Merletti and Parker [92] and Merletti and Farina [94]. For more general background on electrophysiology — not limited to surface electromyography —, refer to, e.g., the introductory textbook by Plonsey and Barr [109]. Technical best practices and recommendations regarding surface EMG applications can be found in two consensus publications [14, 60]. Finally, the historically interested reader may enjoy the popular science book by Ashcroft [8] about the history of electrophysiology.

3.2 The respiratory system

The respiratory system's *raison d'être* is to provide the body's cells with oxygen and, simultaneously, remove excess carbon dioxide from the body. In healthy conditions, this goal is achieved — often unnoticed — with extraordinary robustness by a complex, interconnected system of physiological sensing and control mechanisms, which will be the subject of the following section 3.2.1. For various pathological reasons, however, parts of

---

[3]Interestingly, the occurrence of spectral dips can also be exploited to estimate the action potential propagation velocity [38, 77, 78, 93], which is of interest for, e.g., monitoring muscle fatigue [87].

[4]If the resulting amplitude spectrum is flat, where is all the information about the time course of the signal? Why do the whitened signals, such as the one in fig. 3.8, not look like completely random noise? The answer is: this information remains captured in the phase spectrum. Replacing the phase values of a whitened signal by, e.g., some constant value yields a completely random, uninformative signal.
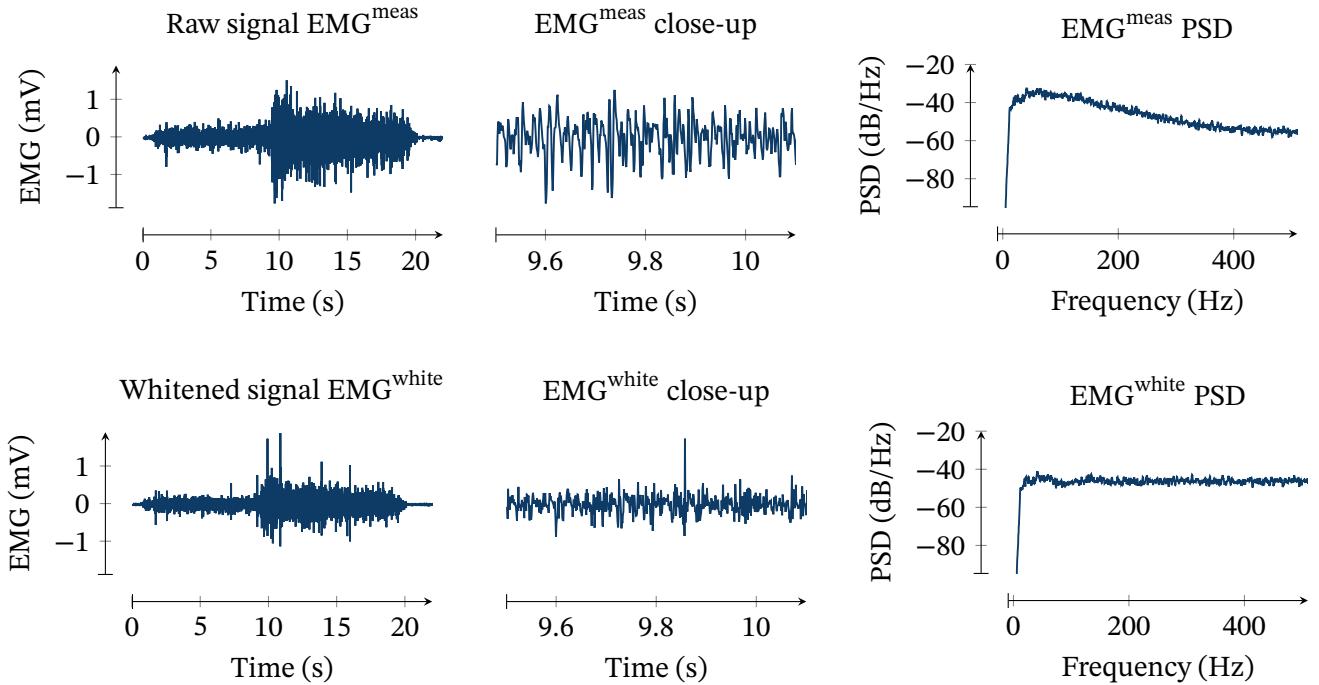
Signal Whitening

Figure 3.8 – An illustrative example of surface EMG signal whitening. The top row shows the original signal EMG^{meas} (preprocessed using only a 10 Hz high-pass Butterworth filter), both in time and frequency space. A sixth-order autoregressive model $H_{burg}(z) = 1/(1 + \sum a_i z^{-i})$ is fit to the power spectral density (PSD) of EMG^{meas} using Burg's method [17, 74]. The original signal EMG^{meas} is then filtered with the *inverse* $H_{burg}(z)^{-1}$ of that model, to obtain the signal EMG^{white} shown in the bottom row.

the respiratory system may cease to fulfill their essential role, making respiratory support necessary. The main themes and technical challenges in providing respiratory support will be discussed in section 3.2.2. Finally, section 3.2.3 presents a few well-known mathematical models of respiratory mechanics.

3.2.1 Physiology and pathophysiology

Figure 3.9 shows a drawing of some of the main components of respiratory anatomy, and fig. 3.10 a simplified view of the respiratory control system in a healthy human. To generate a breath, the respiratory controller, primarily consisting of populations of neurons in the medulla [118] but also various other regions of the brain [64], generates a neural drive signal. This neural drive is propagated towards the motor neurons controlling the muscle fibers that make up the respiratory muscles, causing these muscles to contract. The diaphragm is the main respiratory (inspiratory) muscle, with supporting roles being played by the external intercostal muscles and, particularly in case of high ventilatory demand or diaphragm weakness, the scalenus and sternocleidomastoid neck muscles and the pectoralis chest muscles [118]. As they are contracting, these inspiratory muscles pull the pleural cavity downward and outward, and as a consequence, the pleural cavity and thus
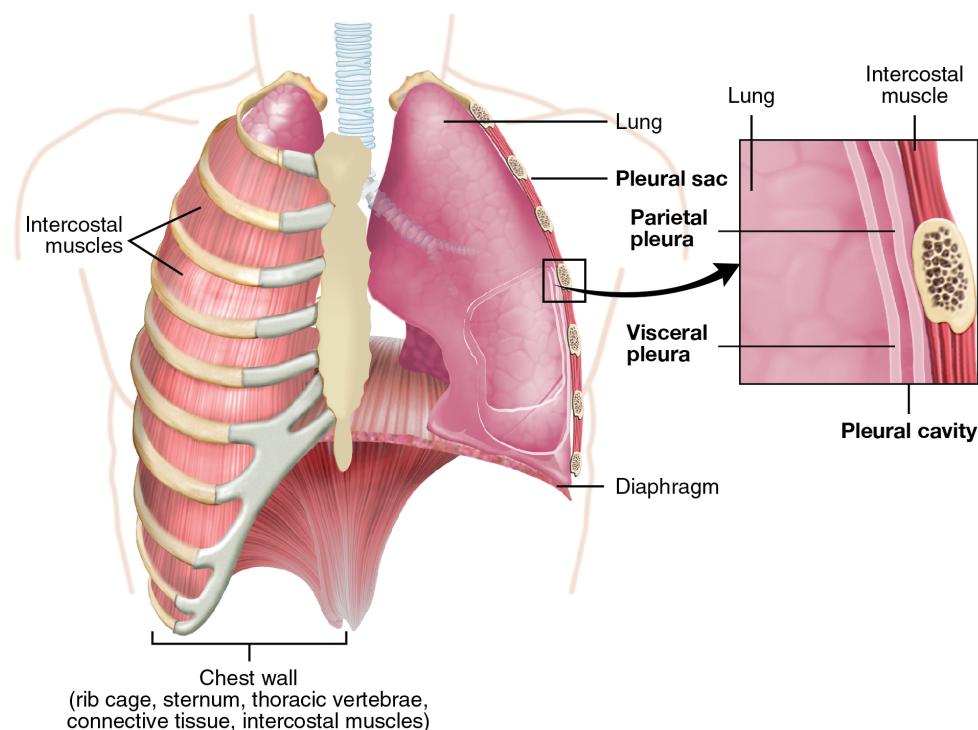
Figure 3.9 – Anatomy of the respiratory system, excluding the respiratory controller, which is located in the brain. Reproduced with permission from Betts *et al.* [15]. Access for free at `https://openstax.org/books/anatomy-and-physiology@22.2/pages/1-introduction`.

the lungs expand. Fresh air streams into the airways to equalize the negative pressure that ensues from this expansion, and oxygen is exchanged against carbon dioxide in the alveoli. Exhalation is usually passive: after some time, the respiratory muscles relax, the pleural cavity and the lungs begin to shrink again, and air is exhaled, equalizing the pressures inside the lungs and outside the airways. Under high ventilatory demand or other special circumstances, forced (or *active*) expiration may also occur, however. In this case, expiratory muscles, including various abdominal muscles[5] and the internal intercostal muscles serve to exert additional pressure and, thereby, force air out of the lungs [118].

Respiration is one of the relatively few physiological processes that can be controlled both consciously or unconsciously by a *central pattern generator* [65, 96, 115, 118]. There are various neural feedback paths that influence the natural control of respiration, including feedback from *stretch receptors* in the lungs to prevent excessive over-inflation of the lungs that may lead to lung damage (the "Hering–Breuer reflex"), and (central and peripheral) *chemoreceptors* that respond to O_2 and CO_2 levels in different regions of the body [110, 118, 125]. Utilizing these feedback paths, the respiratory control center acts to maintain homeostasis in the face of time-varying bodily energy expenditure, ambient temperature, respiratory diseases, and other disturbances.

All components of the respiratory system may cause it to fail in reaching homeostasis.

---

[5]In particular, the rectus abdominis, external and internal oblique, and transversus abdominis muscles [119].
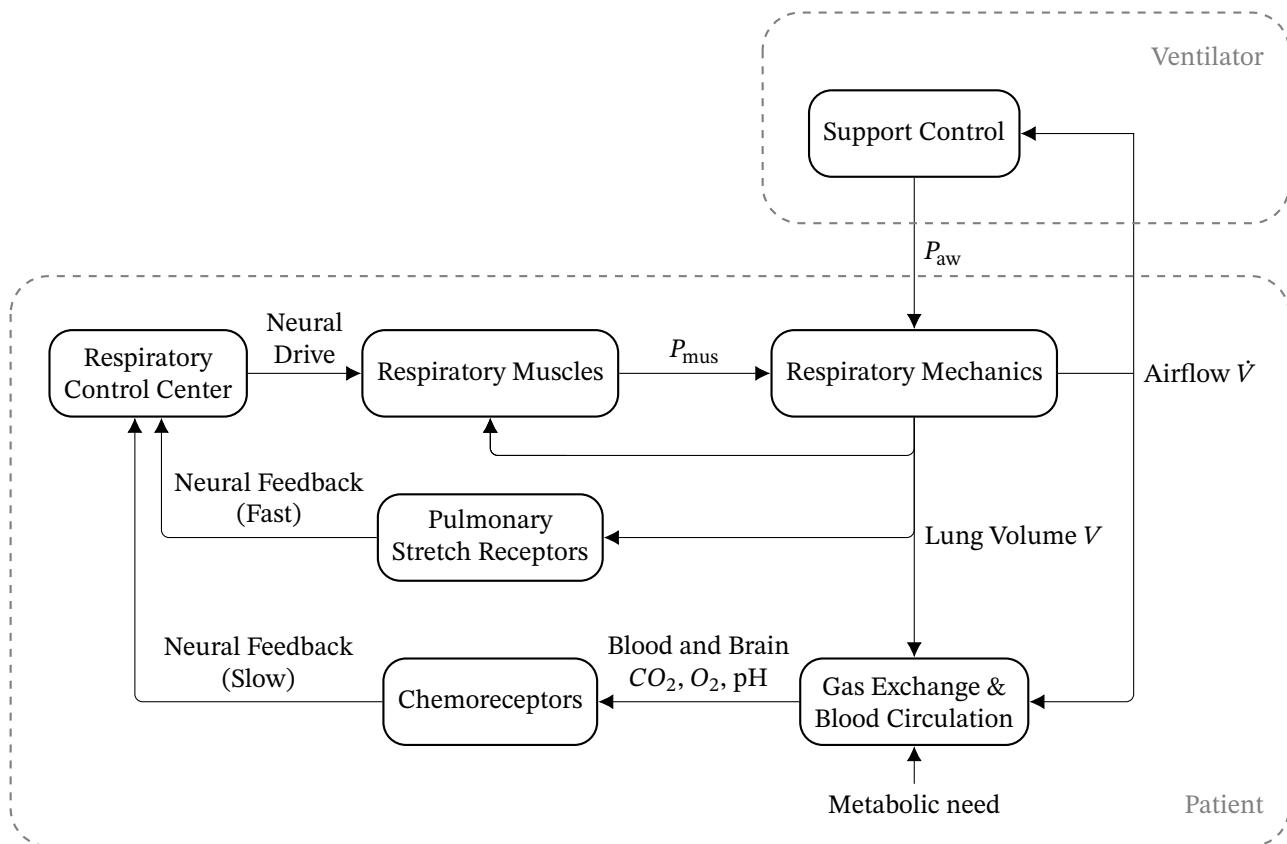
Figure 3.10 – A (simplified) control systems view of the respiratory control system and respiratory support.

The respiratory control center may fail to exert a respiratory drive due to medication or for pathological reasons. The respiratory muscles may be weakened due to continued disuse (e.g., during mechanical ventilation [13, 50]), excessive use (e.g., due to heavily obstructed airways [13, 50]), or infections (such as COVID-19 [120]). They may also be paralyzed during anesthesia. Finally, gas exchange in the lungs may be disturbed for various pathological reasons, such as acute respiratory distress syndrome (ARDS), which may be caused, e.g., by COVID-19 [9, 84]. For all of these reasons (and more that have not been named), patients may require respiratory support by means of a mechanical ventilator.

3.2.2 Respiratory support

If, for any reason, a patient is unable to fulfill her respiratory needs by herself, she may be supported by means of mechanical ventilation. The case in which the patient still has residual respiratory activity, i.e., still exerts some respiratory effort, is called *assisted ventilation* or *respiratory support* (as opposed to *mandatory* ventilation, in which the patient is fully passive, due to, e.g., use of sedatives). Figure 3.10 shows a schematic overview of this setting: there is now a second controller in addition to the patient's own respiratory control center — the mechanical ventilator. The mechanical ventilator typically receives as a

Assisted
Ventilation

measurement the airflow $\dot{V}$ passing through the patient's airways, and it influences the respiration by controlling the external pressure P_{aw} acting on the patient's airways. In the simplest case, the ventilator may just provide a constant positive airway pressure (CPAP), reducing the necessary patient effort to keep the airways and the lungs open. In the intensive care setting, however, a more typical ventilation mode — and the one most

Pressure Support Ventilation

relevant to this thesis — is *pressure support ventilation*. This ventilation mode works by detecting when a patient attempts to breathe (based on simple thresholds on airflow or airway pressure) and then providing a fixed level of pressure support for a pre-set inspiration time.[6] There are many parameters that must be configured by a physician in this ventilation mode, including the baseline level of pressure support, the level of pressure support during inspiration, and the fixed inspiration time.

While it is often a life-saving intervention, mechanical ventilation is also accompanied by many risks, including:

diaphragm atrophy due to disuse if the support level is set too high [13, 49],

diaphragm injury due to excessive strain if the support level is set too low [13, 50, 51],

lung injury due to excessively high driving pressures [13, 49, 139],

and various further complications due to, e.g., intubation, ventilator-associated infections, necessary drug use, and psychological trauma [49, 52]. In order to minimize all of these risks, the aim is generally to transition patients back to fully autonomous respiration as quickly as possible; this process is called "weaning". To address especially the first three

Lung and Diaphragm-protective Ventilation

risks, the paradigm of *lung and diaphragm-protective ventilation* has been developed and widely advocated for in recent years [13, 50, 51]. An essential requirement for following this paradigm is, however, that the pressure P_{mus} generated by the patient's respiratory muscles can be observed. This is currently only feasible with the clinically required reliability using esophageal pressure measurements via an esophageal catheter [3, 85], see fig. 3.11. The following section will briefly discuss how P_{mus} can be estimated from measurements of P_{es}. However, it is not only the overall pressure P_{mus} that is generated by all respiratory muscles together but also the differential activity of the different respiratory muscles that conveys important information about the state of a patient [32, 119]. To the author's knowledge, only two measurement modalities for monitoring the activity of the accessory respiratory muscles (i.e., respiratory muscles other than the diaphragm) are currently known: respiratory surface EMG measurements [55, 134] and mechanomyography [6, 81]. The topic of monitoring a patient's respiratory effort will be discussed in more detail in chapter 7, where a new, noninvasive alternative method (based on respiratory surface EMG measurements) for measuring P_{mus} will be proposed.

One particular complication that may occur during assisted ventilation, and that has

Patient-ventilator Asynchrony

received increased attention in the last two decades, is *patient-ventilator asynchrony*. For a multitude of reasons, the patient and the ventilator may not "breathe in synchrony", which may mean, e.g., that the patient tries to draw a breath but is not supported by the ventilator (because the patient's effort is not detected), or that the patient receives an undesired breath by the ventilator [127]. Patient-ventilator asynchrony is often a result of suboptimal

---

[6]This description is, of course, a simplification. In reality, there are various safety measures and limits built into these modes, such as a maximum delay between subsequent breaths, and implementations differ between manufacturers.
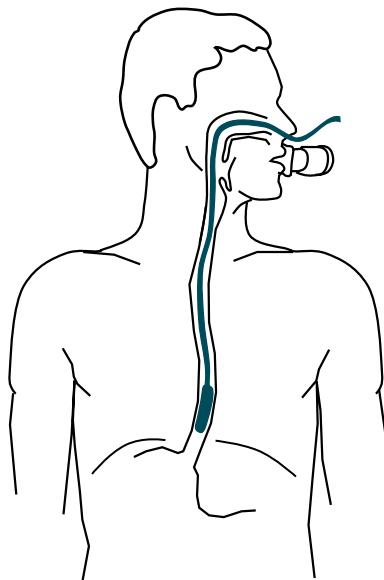
Figure 3.11 – An illustration of esophageal pressure (P_{es}) measurement. An esophageal balloon with a pressure transducer is inserted into the patient's esophagus using a catheter. After correct positioning, the pressure swings measured inside the esophageal balloon closely approximate the swings in the pressure P_{pl} in the pleural cavity. Drawing created and kindly provided by Jan Graßhoff.

ventilator settings, but again — to *detect* patient-ventilator asynchrony, a means to observe the patient's respiratory efforts is required. As a last remark within this very selective overview, researchers have also proposed various *proportional support modes*, which aim to observe a patient's own respiratory effort and provide assistance proportional to that effort [108, 121, 140]. As one might hope, these modes have been found to improve patient-ventilator interaction [107, 117].

Proportional Support Modes

3.2.3 Mathematical models of respiratory mechanics

Figure 3.12 shows two equivalent representations of the standard single-compartment model of respiratory mechanics [10]. Denote by P_{aw} the pressure at the entrance of the airways (i.e., the mouth or the nose), by P_L the pressure within the lungs, by P_{pl} the pressure in the pleural cavity, and by P_{atm} the atmospheric pressure that acts on the body surface. During normal, unassisted breathing, $P_{aw} \equiv P_{atm}$, but during (positive-pressure) mechanical ventilation, we usually have $P_{aw} > P_{atm}$. (Without loss of generality, we will assume $P_{atm} \equiv 0$ in the following.) Furthermore, let V denote the difference between the current lung volume V_L and the lungs' functional residual capacity (FRC) V_0, which is defined as the equilibrium volume level at which $\dot{V} \equiv 0$ for $P_{mus} \equiv 0$. Finally, let $\dot{V}$, its derivative, denote the airflow entering the lungs through the airways. The standard single-compartment model can then be derived from just two assumptions.

Firstly, a pressure drop $\Delta P_R(\dot{V})$ is assumed between P_{aw} and P_L, due to the resistive properties of the airways and the ventilation tube. Usually, this pressure drop is assumed to
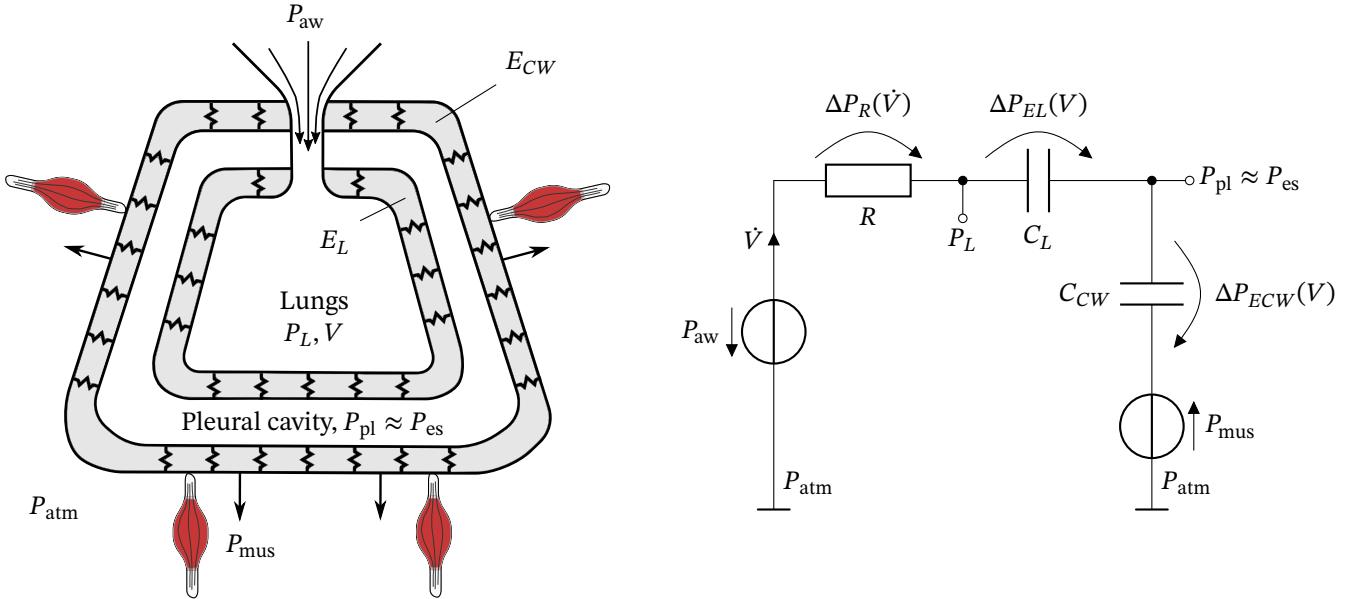
Figure 3.12 – Two equivalent representations of the standard single-compartment model of respiratory mechanics. **Left:** Air must first pass through a narrow, resistive airway opening to reach the lungs, which are represented by an elastic container with volume V_L. This contained is enclosed in a second elastic container, representing the chest wall. The space between the two containers is the pleural cavity. The respiratory muscles exert their forces on said pleural cavity. **Right:** An equivalent electrical circuit. The resistor R models the airway resistance; the two capacitors model the two elastic containers. (C is called the *compliance* and is given by $C = 1/E$.) Voltages represent pressures; currents represent airflows; charges represent volumes. This model is generally less ambiguous than the container model drawn in the left panel.

be either linear or quadratic in the airflow, i.e.,

$$P_{\text{aw}} - P_L = \Delta P_R(\dot{V}) = R\dot{V} \tag{3.5}$$

or (known as *Rohrer's equation* [114])

$$P_{\text{aw}} - P_L = \Delta P_R(\dot{V}) = R_1\dot{V} + R_2\dot{V}|\dot{V}|. \tag{3.6}$$

Secondly, a pressure drop is assumed between the insides and outsides of the two elastic containers "lungs" and "pleural sac", due to their elastic recoil. Usually, this pressure drop is assumed to be either linear or quadratic in the current lung volume, i.e.,

$$P_L - P_{\text{pl}} = \Delta P_{EL}(V) = E_L V, \tag{3.7}$$

Lung Elastance   with E_L denoting the *lung elastance*, or

$$P_L - P_{\text{pl}} = \Delta P_{EL}(V) = E_{L1}V + E_{L2}V^2. \tag{3.8}$$

Equivalently,

$$\Delta P_{ECW}(V) = E_{CW}V, \tag{3.9}$$

with E_{CW} denoting the *chest wall elastance*, or

$$\Delta P_{ECW}(V) = E_{CW1}V + E_{CW2}V^2. \tag{3.10}$$

Finally, the respiratory muscles are assumed to generate a negative pressure $-P_{\mathrm{mus}}$ that acts on the outside of the pleural cavity, and thus we have[7]

$$P_{\mathrm{pl}} + P_{\mathrm{mus}} - P_{\mathrm{atm}} = \Delta P_{ECW}(V). \tag{3.11}$$

Combining eqs. (3.5), (3.7), (3.9) and (3.11) (and assuming $P_{\mathrm{atm}} \equiv 0$), we obtain the *respiratory equation of motion* [10]

$$P_{\mathrm{aw}} = R\dot{V} + (E_L + E_{CW})V - P_{\mathrm{mus}} + P_0. \tag{3.12}$$

(In eq. (7.1), the constant offset P_0 is not strictly necessary but is usually included to identify the offsets, i.e., equilibrium volumes, of the two elastances [10].) From the same equations, we also obtain

$$P_{\mathrm{mus}} = E_{CW}V - P_{\mathrm{pl}}, \tag{3.13}$$

which provides a way to monitor the important quantity P_{mus}, given measurements of P_{pl} and an estimate of E_{CW}. The pleural pressure P_{pl} is difficult (since quite invasive) to measure in practice; fortunately, however, the esophageal pressure P_{es} (following careful and precise placement of the esophageal balloon) has been found to represent an accurate surrogate measurement [10]. Measuring esophageal pressure and accounting for the effect of the chest wall elastance is the current gold standard method for measuring P_{mus} in mechanically ventilated patients [3, 10, 85].[8]

3.3 Study data

A faithful validation of theoretical developments for practical applications always requires representative empirical data. To this end, the developments made in this thesis will be validated using data from two empirical studies. Since the electrode setup used in both studies is similar and differs from standard EMG recording setups, this will be discussed in section 3.3.1. Section 3.3.2 and section 3.3.3 then describe the two studies in further detail.

3.3.1 Measurement setup

In section 3.1.2, the effect of the electrode configuration was assessed at length for general applications, and it was demonstrated that differential measurements with a small inter-electrode distance reduce the so-called *pickup volume* by suppressing spatially low-frequent signal components. Usually, this is a desirable effect. In the particular case of the diaphragm, this is not the case, however: since the diaphragm extends over a large area

---

[7]Throughout this thesis, we define P_{mus} as the negative of the actual pressure generated by the respiratory muscles, i.e., it is positive when a net negative pressure is generated (during inspiration).

[8]The author has co-authored two publications concerning this topic. Graßhoff, Petersen et al. [54] describe a new, automated method for the removal of cardiogenic artifacts from esophageal pressure measurements. Graßhoff, Petersen, et al. [53] describe an automated method for identifying the chest wall elastance E_{CW} and estimating P_{mus} from P_{es}. Both publications will not be further discussed in this thesis.
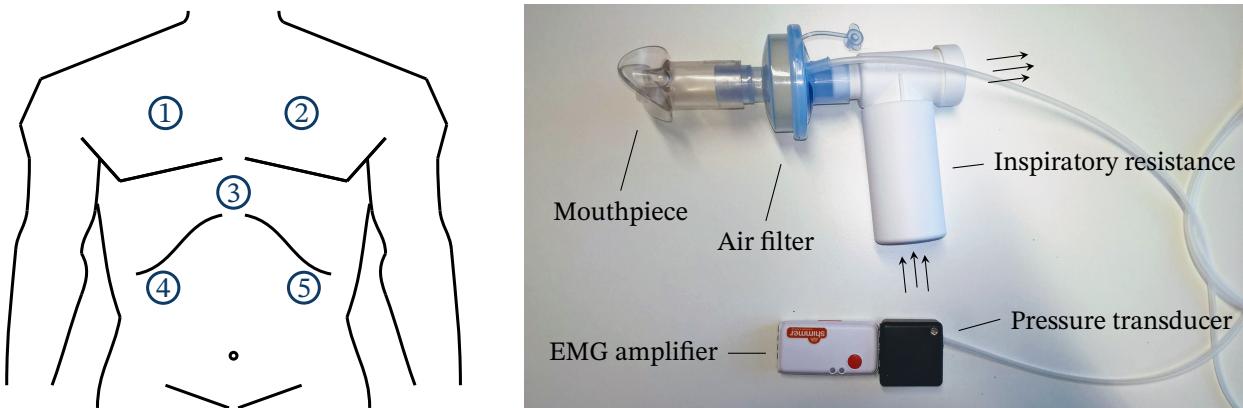
Figure 3.13 – Left: Depiction of the placement of the recording electrodes in study A, described in section 3.3.2. The difference between the signals picked up by electrodes 1 and 2 corresponds to measurement channel 1; the difference difference between the signals picked up by electrodes 4 and 5 corresponds to measurement channel 2. Electrode 3, placed on the sternum, is a reference electrode. Reproduced and modified with permission from Sauer [116]. **Right:** a picture of the study equipment (EMG electrodes + cables not shown).

and extends deeply into the torso (see fig. 3.9), an electrode setup with a large pickup volume is more likely to capture most of the diaphragm's activity. Moreover, differential measurements with a short inter-electrode distance are more prone to being dominated by the activity of superficial abdominal muscles, which lie between the electrodes and the diaphragm. Empirical (unpublished) tests have confirmed these predictions, with monopolar measurements, or differential measurements with a very large inter-electrode distance, yielding an improved signal-to-noise ratio in comparison with differential measurements with a small inter-electrode distance. Moreover, to minimize clinical effort and additional technical processing equipment in the intensive care setting, only a small number of electrodes should be used. For these reasons, the electrode configurations shown in fig. 3.13 and fig. 3.14 were used, which employ long distance differential measurements for each muscle group under consideration to yield a single channel per muscle group (diaphragm, intercostal muscles, rectus abdominis).

3.3.2 Study A: cardiac interference in healthy subjects

For the specific purpose of validating algorithms for removing cardiac interference from trunk EMG recordings — a task which will be the exclusive subject of chapter 5 — a representative study was designed and conducted as part of the Bachelor's thesis of Sauer [116] (under the supervision of the author). A study protocol was written, and ethical approval for the study was obtained from the local ethics committee.

Ten healthy subjects were recruited for the study. Particular care was taken to ensure a diverse and representative study group with respect to age, sex, physical condition, and body mass index (BMI). Participants were asked to lie on a couch and rest quietly for 15 min, while breathing through a pneumatic resistor (PowerBreathe Medic, POWERbreathe International Ltd.) to ensure sufficiently strong respiratory muscle activity. Two channels of surface electromyographic measurements were obtained using a Shimmer3 EMG amplifier (Realtime Technologies Ltd.) at a sampling rate of 1024 Hz, and
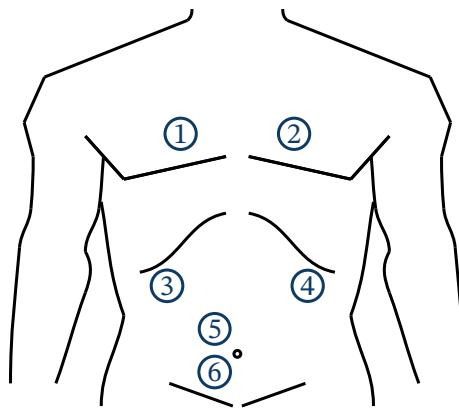
Figure 3.14 – Depiction of the placement of the recording electrodes in study B, described in section 3.3.3. The difference between the signals picked up by electrodes 1 and 2 (second intercostal space, bilaterally on the parasternal line, targeting the parasternal external intercostal muscles) corresponds to the EMG^i measurement channel, the difference difference between the signals picked up by electrodes 3 and 4 (lower costal margin, bilaterally on the midclavicular line, targeting the costal diaphragm) to the EMG^d measurement channel, and the difference between electrodes 5 and 6 (epigastric and hypogastric region, 2cm left or right of median line, targeting the rectus abdominis muscle) to the EMG^r measurement channel.

airway opening pressure was measured using a differential pressure transducer (MP3V5050, Freescale Semiconductors, Inc.) connected to the mouthpiece. Figure 3.13 shows the electrode configuration and the equipment used in the study. Of the total of 20 EMG recordings (ten participants with two channels each), 14 were used for further analyses. Table 3.1 lists the study subjects and some of their attributes, as well as reasons for their exclusion from further analyses, if applicable.

3.3.3 Study B: respiratory effort in mechanically ventilated patients

For the purposes of investigating the estimation of respiratory effort from pneumatic as well as surface electromyographic measurements (the subject of chapter 7), a second dataset is employed. This dataset represents a subset of the data recorded and reported on by Bellani *et al.* [11], and is used here with the kind permission of Dr. Bellani and his colleagues.[9] The study was conducted in an intensive care unit (ICU) at San Gerardo Hospital, Monza, Italy. Adult intubated patients under no or mild sedation were included in the study, and subjects received pressure support ventilation (PSV) using a Servo-i mechanical ventilator (Getinge, Göteborg, Sweden). Esophageal pressure P_{es} was measured using a nasogastric catheter (EDI catheter Maquet, Solna, Sweden) with an esophageal balloon (Adult Esophageal Balloon Catheter 47-9005, Cooper Surgical, Trumbull, Connecticut). Cardiac artifacts in the P_{es} signal were removed using a template subtraction technique proposed by Graßhoff, Petersen, et al. [54], and an estimate $P_{mus\text{-}Pes}$ of P_{mus} was calculated from the P_{es} and V

---

[9]Dr. Bellani and his colleagues recorded more subjects than are reported on in Bellani *et al.* [11]. The subset analyzed here partially differs from the subset analyzed in Bellani *et al.* [11], whence there are differences between the patient group characteristics.

signals using the algorithm described by Graßhoff, Petersen, et al. [53]. (Briefly, E_{CW} was estimated on snippets of 30 s length, and P_{mus} was estimated from eq. (3.13).) Surface EMG signals were collected using three differential pairs of skin electrodes (Kendall 530 Foam Electrodes, Covidien, Dublin, Ireland), which were positioned as depicted in fig. 3.14. The EMG signals were amplified using the Dipha 16 amplifier (Inbiolab, Groningen, Netherlands) and processed using a dedicated signal processing algorithm, which removed cardiac artifacts using a gating procedure, high-pass filtered the raw signals to remove baseline noise, and calculated a moving envelope signal based on a weighted moving average of the absolute EMG signal within a 250 ms window. The envelope calculation was implemented in a *causal* fashion, i.e., only past raw EMG values were taken into account, thus introducing a delay in the EMG envelope signals. The EMG (in µV), ventilator (airway pressure P_{aw} in mbar, airflow $\dot{V}$ in L/s, volume V in L), and P_{es} (in mbar) signals were gathered, synchronized, and recorded using a dedicated acquisition system (Drägerwerk AG & Co. KGaA, Lübeck, Germany). For more details on the study characteristics, refer to Bellani *et al.* [11].

For our analyses comparing EMG-based measures of respiratory effort with P_{es}-based measures (chapter 7), we excluded recordings in which

- either the recording of P_{es} or of the surface EMG signals failed or was subject to strong artifacts,

- the algorithm [53] computing $P_{\mathrm{mus\text{-}Pes}}$ from P_{es} failed (this concerned one dataset which used the NAVA ventilation mode, for which the estimation procedure was not conceived), or

- the patient did not show significant spontaneous breathing activity.

Moreover, only one recording per patient is used, although multiple recordings have been made for most patients. In total, nine subjects were selected for further analyses; their characteristics are listed in table 3.2. The recordings of patients B/08 and B/09 were split on changes in ventilator settings, yielding a total of eleven recordings that are analyzed in chapter 7. As a last remark, the aim of the analyses in chapter 7 was a technical feasibility study and not a formal observational study; thus, subject selection was not performed comprehensively. In other words: it is likely that the data recorded by Dr. Bellani and colleagues contain more suitable subjects than have been analyzed here. (The performance of the algorithm presented in chapter 7 did not influence the selection of the recordings under consideration in any way.)

no.	age	sex	weight (kg)	size (cm)	BMI (kg/m^2)	other features	body fat (%) skin wrinkle (mm)	Reason for exclusion
A/01	21	m	67	182	20	athletic, very lean	7.1 (6, 8, 6, 6, 8, 14, 9)	—
A/02	65	f	71	162	27	just stopped smoking	18.9 (6, 12, 10, 18, 21, 14, 18)	Measurement artifacts in both channels
A/03	22	m	124	187	35	obese	22.4 (18, 36, 15, 26, 28, 26, 20)	Measurement artifacts in channel 2
A/04	27	m	82	186	24	muscular, lean	5.4 (4, 8, 6, 7, 5, 4, 8)	Measurement artifacts in both channels
A/05	21	f	78	169	27	six months pregnant	11.1 (10, 2, 8, 12, 10, 21, 20)	—
A/06	21	m	67	197	17	underweight	3.1 (2, 5, 2, 8, 3, 6, 6)	—
A/07	22	f	56	174	18	lean	3.8 (2, 11, 3, 6, 4, 4, 6)	—
A/08	25	m	74	180	23	lean, athletic to muscular	4.2 (3, 6, 6, 6, 3, 4,8)	—
A/09	50	f	75	168	27	slightly overweight	21.3 (14, 26, 4, 24, 22, 18)	Measurement artifacts in channel 1
A/10	23	f	84	179	26	average	16.5 (6, 18, 11, 17, 15, 32, 21)	—

Table 3.1 – Data on the participants of study A, described in section 3.3.2. The body fat column additionally specifies the measured skin wrinkle size of chest, belly, leg, hip, armpit, triceps, and back. Reproduced and modified with permission from Sauer [116].

no.	age	sex	BMI (kg/m²)	sedation	recording length(s)	occlusions	diagnosis	remarks
B/1	45	m	16	weak	11 min 10 s	10	ARF	—
B/2	60	f	25	weak	11 min 4 s	8	Septic shock	Periodic breathing
B/3	55	m	29	weak	11 min 11 s	0	Pneumonia	Periodic breathing
B/4	76	m	26	none	11 min 30 s	5	COPD	Periodic breathing
B/5	78	m	25	none	8 min 44 s	7	ARF, pneumonia	Active expiration
B/6	59	m	20	N/A	31 min 32 s	4	ARF	—
B/7	78	m	25	N/A	31 min 2 s	2	ARF	—
B/8	N/A	N/A	N/A	N/A	10 min 20 s 11 min 0 s	10/10	N/A	Diagnosis & demographics unknown; split into two parts based on changes in PEEP and PS
B/9	54	m	24	weak	5 min 38 s 6 min 59 s	0/0	COPD, pneumonia	Split into two parts based on a change in PS

Table 3.2 – Data on the selected subjects of study B, described in section 3.3.3. Body-mass index (BMI), acute respiratory failure (ARF), chronic obstructive pulmonary disease (COPD), information unavailable (N/A), positive end-expiratory pressure (PEEP), pressure support level (PS). Data (except for occlusion counts and remarks) kindly provided by Bellani *et al.* [11].

References

[1] A. Adam and C. J. De Luca, "Recruitment order of motor units in human vastus lateralis muscle is maintained during fatiguing contractions," *Journal of Neurophysiology*, vol. 90, no. 5, pp. 2919–2927, 2003. DOI: 10.1152/jn.00179.2003.

[2] B. Afsharipour, S. Soedirdjo, and R. Merletti, "Two-dimensional surface EMG: The effects of electrode size, interelectrode distance and image truncation," *Biomedical Signal Processing and Control*, vol. 49, pp. 298–307, 2019. DOI: 10.1016/j.bspc.2018.12.001.

[3] E. Akoumianaki *et al.*, "The application of esophageal pressure measurement in patients with respiratory failure," vol. 189, no. 5, pp. 520–531, 2014. DOI: 10.1164/rccm.201312-2193ci.

[4] M. Al Harrach *et al.*, "Analysis of the sEMG/force relationship using HD-sEMG technique and data fusion: A simulation study," *Computers in Biology and Medicine*, vol. 83, pp. 34–47, 2017. DOI: 10.1016/j.compbiomed.2017.02.003.

[5] B. A. Alkner, P. A. Tesch, and H. E. Berg, "Quadriceps EMG/force relationship in knee extension and leg press," *Medicine and science in sports and exercise*, vol. 32, no. 2, pp. 459–463, 2000. DOI: 10.1097/00005768-200002000-00030.

[6] J. F. Alonso *et al.*, "Evaluation of respiratory muscles activity by means of cross mutual information function at different levels of ventilatory effort," *IEEE Transactions on Biomedical Engineering*, vol. 54, no. 9, pp. 1573–1582, 2007. DOI: 10.1109/tbme.2007.893494.

[7] C. Anders, G. Brose, G. O. Hofmann, and H.-C. Scholle, "Evaluation of the EMG–force relationship of trunk muscles during whole body tilt," *Journal of Biomechanics*, vol. 41, no. 2, pp. 333–339, 2008. DOI: 10.1016/j.jbiomech.2007.09.008.

[8] F. Ashcroft, *The Spark of Life: Electricity in the Human Body*. W. W. Norton & Co., 2012, 339 pp., ISBN: 0393078035.

[9] S. S. Batah and A. T. Fabro, "Pulmonary pathology of ARDS in COVID-19: A pathological review for clinicians," vol. 176, p. 106 239, 2021. DOI: 10.1016/j.rmed.2020.106239.

[10] J. H. T. Bates, *Lung Mechanics: An Inverse Modeling Approach*. Cambridge University Press, 2009, ISBN: 978-0-521-50960-2.

[11] G. Bellani *et al.*, "Measurement of diaphragmatic electrical activity by surface electromyography in intubated subjects and its relationship with inspiratory effort," *Respiratory Care*, vol. 63, no. 11, pp. 1341–1349, 2018. DOI: 10.4187/respcare.06176.

[12] D. J. Berger and A. d'Avella, "Effective force control by muscle synergies," *Frontiers in Computational Neuroscience*, vol. 8, 2014. DOI: 10.3389/fncom.2014.00046.

[13] M. Bertoni, S. Spadaro, and E. C. Goligher, "Monitoring patient respiratory effort during mechanical ventilation: Lung and diaphragm-protective ventilation," *Critical Care*, vol. 24, no. 1, 2020. DOI: 10.1186/s13054-020-2777-y.

[14] M. Besomi *et al.*, "Consensus for experimental design in electromyography (CEDE) project: Electrode selection matrix," *Journal of Electromyography and Kinesiology*, vol. 48, pp. 128–144, 2019. DOI: 10.1016/j.jelekin.2019.07.008.

[15] J. G. Betts *et al.*, "The respiratory system," in *Anatomy and Physiology*. Houston, Texas: OpenStax, 2013. [Online]. Available: https://openstax.org/books/anatomy-and-physiology@22.2/pages/22-introduction.

[16] E. Bizzi and V. C. K. Cheung, "The neural origin of muscle synergies," *Frontiers in Computational Neuroscience*, vol. 7, 2013. DOI: 10.3389/fncom.2013.00051.

[17] J. P. Burg, "Maximum entropy spectral analysis," *Astronomy and Astrophysics Supplement*, vol. 15, p. 383, 1974.

[18] H. V. Cabral *et al.*, "Is the firing rate of motor units in different vastus medialis regions modulated similarly during isometric contractions?" *Muscle & Nerve*, vol. 57, no. 2, pp. 279–286, 2017. DOI: 10.1002/mus.25688.

[19] I. Campanini, C. Disselhorst-Klug, W. Z. Rymer, and R. Merletti, "Surface EMG in clinical assessment and neurorehabilitation: Barriers limiting its use," *Frontiers in Neurology*, vol. 11, 2020. DOI: 10.3389/fneur.2020.00934.

[20] E. A. Clancy, D. Farina, and G. Filligoi, "Single-channel techniques for information extraction from the surface EMG signal," in *Electromyography: Physiology, Engineering, and Non-Invasive Applications*, John Wiley & Sons, Inc., 2004, pp. 133–168. DOI: 10.1002/0471678384.ch6.

[21] E. Clancy, S. Bouchard, and D. Rancourt, "Estimation and application of EMG amplitude during dynamic contractions," *IEEE Engineering in Medicine and Biology Magazine*, vol. 20, no. 6, pp. 47–54, 2001. DOI: 10.1109/51.982275.

[22] E. Clancy and K. Farry, "Adaptive whitening of the electromyogram to improve amplitude estimation," *IEEE Transactions on Biomedical Engineering*, vol. 47, no. 6, pp. 709–719, 2000. DOI: 10.1109/10.844217.

[23] E. A. Clancy, O. Bida, and D. Rancourt, "Influence of advanced electromyogram (EMG) amplitude processors on EMG-to-torque estimation during constant-posture, force-varying contractions," *Journal of Biomechanics*, vol. 39, no. 14, pp. 2690–2698, 2006. DOI: 10.1016/j.jbiomech.2005.08.007.

[24] P. Contessa and C. J. De Luca, "Neural control of muscle force: Indications from a simulation model," *Journal of Neurophysiology*, vol. 109, pp. 1548–1570, 2013. DOI: 10.1152/jn.00237.2012.

[25] A. d'Avella and E. Bizzi, "Shared and specific muscle synergies in natural motor behaviors," *Proceedings of the National Academy of Sciences*, vol. 102, no. 8, pp. 3076–3081, 2005. DOI: 10.1073/pnas.0500199102.

[26] C. J. De Luca and P. Contessa, "Hierarchical control of motor units in voluntary contractions," *Journal of Neurophysiology*, vol. 107, no. 1, pp. 178–195, 2012. DOI: 10.1152/jn.00961.2010.

[27] C. J. De Luca and B. Mambrito, "Voluntary control of motor units in human antagonist muscles: Coactivation and reciprocal activation," *Journal of Neurophysiology*, vol. 58, no. 3, pp. 525–542, 1987. DOI: 10.1152/jn.1987.58.3.525.

[28] C. J. De Luca and Z. Erim, "Common drive of motor units in regulation of muscle force," *Trends in Neurosciences*, vol. 17, no. 7, pp. 299–305, 1994. DOI: 10.1016/0166-2236(94)90064-7.

[29] C. J. De Luca *et al.*, "Decomposition of surface EMG signals from cyclic dynamic contractions," *Journal of Neurophysiology*, vol. 113, no. 6, pp. 1941–1951, 2015. DOI: 10.1152/jn.00555.2014.

[30] J. L. Dideriksen, D. F. Feeney, A. M. Almuklass, and R. M. Enoka, "Control of force during rapid visuomotor force-matching tasks can be described by discrete time PID control algorithms," *Experimental Brain Research*, vol. 235, no. 8, pp. 2561–2573, 2017. DOI: 10.1007/s00221-017-4995-3.

[31] N. A. Dimitrova, G. V. Dimitrov, and O. A. Nikitin, "Neither high-pass filtering nor mathematical differentiation of the EMG signals can considerably reduce cross-talk," *Journal of Electromyography and Kinesiology*, vol. 12, no. 4, pp. 235–246, 2002. DOI: 10.1016/s1050-6411(02)00008-1.

[32] J. Doorduin, H. W. H. van Hees, J. G. van der Hoeven, and L. M. A. Heunks, "Monitoring of the respiratory muscles in the critically ill," *American Journal of Respiratory and Critical Care Medicine*, vol. 187, no. 1, pp. 20–27, 2013. DOI: 10.1164/rccm.201206-1117cp.

[33] M. L. Duiverman *et al.*, "Reproducibility and responsiveness of a noninvasive EMG technique of the respiratory muscles in COPD patients and in healthy subjects," *Journal of Applied Physiology: Respiratory, Environmental and Exercise Physiology*, vol. 96, no. 5, pp. 1723–1729, 2004. DOI: 10.1152/japplphysiol.00914.2003.

[34] C. L. Ebbesen and M. Brecht, "Motor cortex — to act or not to act?" *Nature Reviews Neuroscience*, vol. 18, no. 11, pp. 694–705, 2017. DOI: 10.1038/nrn.2017.119.

[35] R. M. Enoka and A. J. Fuglevand, "Motor unit physiology: Some unresolved issues," *Muscle and Nerve*, vol. 42, pp. 4–17, 2001. DOI: 10.1002/1097-4598(200101)24:1<4::AID-MUS13>3.0.CO;2-F.

[36] Z. Erim, C. J. De Luca, K. Mineo, and T. Aoki, "Rank-ordered regulation of motor units," *Muscle & Nerve*, vol. 19, no. 5, pp. 563–573, 1996. DOI: https://doi.org/10.1002/(SICI)1097-4598(199605)19:5<563::AID-MUS3>3.0.CO;2-9.

[37] D. Farina and R. Merletti, "A novel approach for precise simulation of the EMG signal detected by surface electrodes," *IEEE Transactions on Biomedical Engineering*, vol. 48, pp. 637–646, 6 2001. DOI: 10.1109/10.923782.

[38] D. Farina and F. Negro, "Estimation of muscle fiber conduction velocity with a spectral multidip approach," *IEEE Transactions on Biomedical Engineering*, vol. 54, no. 9, pp. 1583–1589, 2007. DOI: 10.1109/tbme.2007.892928.

[39] D. Farina, D. F. Stegeman, and R. Merletti, "Biophysics of the generation of EMG signals," in *Surface Electromyography : Physiology, Engineering, and Applications*, John Wiley & Sons, Inc., 2016, pp. 1–24. DOI: 10.1002/9781119082934.ch02.

[40] D. Farina, C. Cescon, and R. Merletti, "Influence of anatomical, physical, and detection-system parameters on surface EMG," *Biological Cybernetics*, vol. 86, no. 6, pp. 445–456, 2002. DOI: 10.1007/s00422-002-0309-2.

[41] D. Farina and A. Holobar, "Characterization of human motor units from surface EMG decomposition," *Proceedings of the IEEE*, vol. 104, no. 2, pp. 353–373, 2016. DOI: 10.1109/jproc.2015.2498665.

[42] D. Farina and R. Merletti, "Comparison of algorithms for estimation of EMG variables during voluntary isometric contractions," *Journal of Electromyography and Kinesiology*, vol. 10, pp. 337–349, 2000. DOI: 10.1016/S1050-6411(00)00025-0.

[43] D. Farina, R. Merletti, and R. M. Enoka, "The extraction of neural strategies from the surface EMG: An update," *Journal of Applied Physiology*, vol. 117, no. 11, pp. 1215–1230, 2014. DOI: 10.1152/japplphysiol.00162.2014.

[44] D. Farina, F. Negro, S. Muceli, and R. M. Enoka, "Principles of motor unit physiology evolve with advances in technology," *Physiology*, vol. 31, no. 2, pp. 83–94, 2016. DOI: 10.1152/physiol.00040.2015.

[45] D. Farina *et al.*, "The extraction of neural information from the surface EMG for the control of upper-limb prostheses: Emerging avenues and challenges," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 22, no. 4, pp. 797–809, 2014. DOI: 10.1109/tnsre.2014.2305111.

[46] A. Fougner *et al.*, "Control of upper limb prostheses: Terminology and proportional myoelectric control—a review," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 20, no. 5, pp. 663–677, 2012. DOI: 10.1109/tnsre.2012.2196711.

[47] A. J. Fuglevand, D. A. Winter, and A. E. Patla, "Models of recruitment and rate coding organization in motor-unit pools," *Journal of Neurophysiology*, vol. 70, pp. 2470–88, 6 1993. DOI: 10.1152/jn.1993.70.6.2470.

[48]   P. Geethanjali, "Myoelectric control of prosthetic hands: State-of-the-art review," *Medical Devices: Evidence and Research*, vol. Volume 9, pp. 247–255, 2016. DOI: 10.2147/mder.s91102.

[49]   E. C. Goligher, N. D. Ferguson, and L. J. Brochard, "Clinical challenges in mechanical ventilation," *The Lancet*, vol. 387, no. 10030, pp. 1856–1866, 2016. DOI: 10.1016/s0140-6736(16)30176-3.

[50]   E. C. Goligher *et al.*, "Clinical strategies for implementing lung and diaphragm-protective ventilation: Avoiding insufficient and excessive effort," *Intensive Care Medicine*, vol. 46, no. 12, pp. 2314–2326, 2020. DOI: 10.1007/s00134-020-06288-9.

[51]   E. C. Goligher *et al.*, "Lung- and diaphragm-protective ventilation," vol. 202, no. 7, pp. 950–961, 2020. DOI: 10.1164/rccm.202003-0655cp.

[52]   M. J. Grap, "Not-so-trivial pursuit: Mechanical ventilation risk reduction," *American Journal of Critical Care*, vol. 18, no. 4, pp. 299–309, 2009. DOI: 10.4037/ajcc2009724.

[53]   J. Graßhoff, E. Petersen, T. Becher, and P. Rostalski, "Automatic estimation of respiratory effort using esophageal pressure," in *41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Berlin, Germany, 2019. DOI: 10.1109/embc.2019.8856345.

[54]   J. Graßhoff *et al.*, "A template subtraction method for the removal of cardiogenic oscillations on esophageal pressure signals," in *39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Jeju Island, Korea, 2017. DOI: 10.1109/EMBC.2017.8037299.

[55]   M. F. Gutiérrez *et al.*, "Does breathing type influence electromyographic activity of obligatory and accessory respiratory muscles?" *Journal of Oral Rehabilitation*, vol. 41, no. 11, pp. 801–808, 2014. DOI: 10.1111/joor.12209.

[56]   A. Gydikov and D. Kosarov, "Some features of different motor units in human biceps brachii," *Pflügers Archiv European Journal of Physiology*, vol. 347, no. 1, pp. 75–88, 1974. DOI: 10.1007/bf00587056.

[57]   C. J. Heckman and R. M. Enoka, "Motor unit," in *Comprehensive Physiology*. American Cancer Society, 2012, pp. 2629–2682, ISBN: 9780470650714. DOI: 10.1002/cphy.c100087.

[58]   E. Henneman, "Relation between size of neurons and their susceptibility to discharge," *Science*, vol. 126, pp. 1345–1347, 3287 1957. DOI: 10.1126/science.126.3287.1345.

[59]   E. Henneman, G. Somjen, and D. O. Carpenter, "Excitability and inhibitibility of motoneurons of different sizes," *Journal of Neurophysiology*, vol. 28, no. 3, pp. 599–620, 1965. DOI: 10.1152/jn.1965.28.3.599.

[60]   H. J. Hermens, B. Freriks, C. Disselhorst-Klug, and G. Rau, "Development of recommendations for SEMG sensors and sensor placement procedures," *Journal of Electromyography and Kinesiology*, vol. 10, no. 5, pp. 361–374, 2000. DOI: 10.1016/s1050-6411(00)00027-4.

[61]   A. V. Hill, "The heat of shortening and the dynamic constants of muscle," *Proceedings of the Royal Society of London. Series B - Biological Sciences*, vol. 126, no. 843, pp. 136–195, 1938. DOI: 10.1098/rspb.1938.0050.

[62]   A. Holobar and D. Farina, "Blind source identification from the multichannel surface electromyogram," *Physiological Measurement*, vol. 35, no. 7, R143–R165, 2014. DOI: 10.1088/0967-3334/35/7/R143.

[63]   A. Holobar *et al.*, "Estimating motor unit discharge patterns from high-density surface electromyogram," *Clinical Neurophysiology*, vol. 120, no. 3, pp. 551–562, 2009. DOI: 10.1016/j.clinph.2008.10.160.

[64] G. Holstege, C. M. Beers, and H. H. Subramanian, Eds., *The Central Nervous System Control of Respiration*. Elsevier Science & Technology, 2014, 440 pp.

[65] S. L. Hooper, "Central pattern generators," *Current Biology*, vol. 10, no. 5, R176–R179, 2000. DOI: 10.1038/npg.els.0000032.

[66] A. J. Ijspeert, "Central pattern generators for locomotion control in animals and robots: A review," *Neural Networks*, vol. 21, no. 4, pp. 642–653, 2008. DOI: 10.1016/j.neunet.2008.03.014.

[67] J. G. Inglis and D. A. Gabriel, "Is the 'reverse onion skin' phenomenon more prevalent than we thought during intramuscular myoelectric recordings from low to maximal force outputs?" *Neuroscience Letters*, vol. 743, p. 135 583, 2021. DOI: 10.1016/j.neulet.2020.135583.

[68] N. V. Iqbal, K. Subramaniam, and S. Asmi P., "A review on upper-limb myoelectric prosthetic control," *IETE Journal of Research*, vol. 64, no. 6, pp. 740–752, 2017. DOI: 10.1080/03772063.2017.1381047.

[69] M. Ison and P. Artemiadis, "Proportional myoelectric control of robots: Muscle synergy development drives performance enhancement, retainment, and generalization," *IEEE Transactions on Robotics*, vol. 31, no. 2, pp. 259–268, 2015. DOI: 10.1109/tro.2015.2395731.

[70] N. Jiang *et al.*, "Intuitive, online, simultaneous, and proportional myoelectric control over two degrees-of-freedom in upper limb amputees," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 22, no. 3, pp. 501–510, 2014. DOI: 10.1109/tnsre.2013.2278411.

[71] N. Jiang *et al.*, "Is accurate mapping of EMG signals on kinematics needed for precise online myoelectric control?" *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 22, no. 3, pp. 549–558, 2014. DOI: 10.1109/tnsre.2013.2287383.

[72] R. M. Kacmarek, L. Berra, and J. Villar, "On the road to surface monitoring of diaphragmatic activity in mechanically ventilated patients," *Respiratory Care*, vol. 63, no. 11, pp. 1457–1458, 2018. DOI: 10.4187/respcare.06685.

[73] T. Kapelner *et al.*, "Predicting wrist kinematics from motor unit discharge timings for the control of active prostheses," *Journal of NeuroEngineering and Rehabilitation*, vol. 16, no. 1, 2019. DOI: 10.1186/s12984-019-0516-x.

[74] S. M. Kay, *Modern spectral estimation: theory and application.* Pearson Education India, 1988.

[75] C. M. Laine *et al.*, "Motor neuron pools of synergistic thigh muscles share most of their synaptic input," *Journal of Neuroscience*, vol. 35, no. 35, pp. 12 207–12 216, 2015. DOI: 10.1523/jneurosci.0240-15.2015.

[76] B. G. Lapatki *et al.*, "A thin, flexible multielectrode grid for high-density surface EMG," *Journal of Applied Physiology*, vol. 96, no. 1, pp. 327–336, 2004. DOI: 10.1152/japplphysiol.00521.2003.

[77] L. Lindstrom, R. Magnusson, and I. Petersén, "Muscular fatigue and action potential conduction velocity changes studied with frequency analysis of EMG signals," *Electromyography*, vol. 10, pp. 341–356, 4 1970, ISSN: 0013-4732.

[78] L. Lindstrom and R. Magnusson, "Interpretation of myoelectric power spectra: A model and its applications," *Proceedings of the IEEE*, vol. 65, no. 5, pp. 653–662, 1977. DOI: 10.1109/proc.1977.10544.

[79] M. M. Liu, W. Herzog, and H. H. Savelberg, "Dynamic muscle force predictions from EMG: An artificial neural network approach," *Journal of Electromyography and Kinesiology*, vol. 9, no. 6, pp. 391–400, 1999. DOI: 10.1016/s1050-6411(99)00014-0.

[80]   R. V. Lourenco and E. P. Mueller, "Quantification of electrical activity in the human diaphragm," *Journal of Applied Physiology: Respiratory, Environmental and Exercise Physiology*, vol. 22, no. 3, pp. 598–600, 1967. DOI: 10.1152/jappl.1967.22.3.598.

[81]   M. Lozano-García *et al.*, "Surface mechanomyography and electromyography provide non-invasive indices of inspiratory muscle force and activation in healthy subjects," *Scientific Reports*, vol. 8, no. 1, 2018. DOI: 10.1038/s41598-018-35024-z.

[82]   B. J. Martin and Y. Acosta-Sojo, "sEMG: A window into muscle work, but not easy to teach and delicate to practice—a perspective on the difficult path to a clinical tool," *Frontiers in Neurology*, vol. 11, 2021. DOI: 10.3389/fneur.2020.588451.

[83]   E. Martinez-Valdes *et al.*, "High-density surface electromyography provides reliable estimates of motor unit behavior," *Clinical Neurophysiology*, vol. 127, no. 6, pp. 2534–2541, 2016. DOI: 10.1016/j.clinph.2015.10.065.

[84]   R. J. Mason, "Thoughts on the alveolar phase of COVID-19," vol. 319, no. 1, pp. L115–L120, 2020. DOI: 10.1152/ajplung.00126.2020.

[85]   T. Mauri *et al.*, "Esophageal and transpulmonary pressure in the clinical setting: Meaning, usefulness and perspectives," *Intensive Care Medicine*, vol. 42, no. 9, pp. 1360–1373, 2016. DOI: 10.1007/s00134-016-4400-x.

[86]   A. McMillan and A. Hannam, "Task-related behaviour of motor units in different regions of the human masseter muscle," *Archives of Oral Biology*, vol. 37, no. 10, pp. 849–857, 1992. DOI: 10.1016/0003-9969(92)90119-s.

[87]   R. Merletti, B. Afsharipour, J. Dideriksen, and D. Farina, "Muscle force and myoelectric manifestations of muscle fatigue in voluntary and electrically elicited contractions," in *Surface Electromyography : Physiology, Engineering, and Applications*, John Wiley & Sons, Inc., 2016, pp. 273–310. DOI: 10.1002/9781119082934.ch10.

[88]   R. Merletti, A. Botter, and U. Barone, "Detection and conditioning of surface EMG signals," in *Surface Electromyography : Physiology, Engineering, and Applications*, John Wiley & Sons, Inc., 2016, pp. 1–37. DOI: 10.1002/9781119082934.ch03.

[89]   R. Merletti and G. Cerone, "Tutorial. surface EMG detection, conditioning and pre-processing: Best practices," *Journal of Electromyography and Kinesiology*, vol. 54, p. 102 440, 2020. DOI: 10.1016/j.jelekin.2020.102440.

[90]   R. Merletti and H. J. Hermens, "Detection and conditioning of the surface EMG signal," in *Electromyography*, John Wiley & Sons, Inc., 2005, pp. 107–131. DOI: 10.1002/0471678384.ch5.

[91]   R. Merletti and S. Muceli, "Tutorial. surface EMG detection in space and time: Best practices," *Journal of Electromyography and Kinesiology*, vol. 49, p. 102 363, 2019. DOI: 10.1016/j.jelekin.2019.102363.

[92]   R. Merletti and P. A. Parker, Eds., *Electromyography, Physiology, Engineering, and noninvasive Applications*. John Wiley & Sons, Inc., 2004.

[93]   R. Merletti, T. M. Vieira, and D. Farina, "Techniques for information extraction from the surface EMG signalhigh-density surface EMG," in *Surface Electromyography : Physiology, Engineering, and Applications*, John Wiley & Sons, Inc., 2016, pp. 126–157. DOI: 10.1002/9781119082934.ch05.

[94]   R. Merletti and D. Farina, Eds., *Surface Electromyography: Physiology, Engineering and Applications*. John Wiley and Sons, Inc., Hoboken, New Jersey, USA., 2016.

[95]   R. Merletti, I. Campanini, W. Z. Rymer, and C. Disselhorst-Klug, "Editorial: Surface electromyography: Barriers limiting widespread use of sEMG in clinical assessment and neurorehabilitation," *Frontiers in Neurology*, vol. 12, 2021. DOI: 10.3389/fneur.2021.642257.

[96] Y. I. Molkov *et al.*, "A closed-loop model of the respiratory system: Focus on hypercapnia and active expiration," *PLoS ONE*, vol. 9, no. 10, M. Koval, Ed., e109894, 2014. DOI: 10.1371/journal.pone.0109894.

[97] T. Moritani, D. Stegeman, and R. Merletti, "Basic physiology and biophysics of EMG signal generation," in *Electromyography: Physiology, Engineering, and Non-Invasive Applications*, John Wiley & Sons, Inc., 2005, pp. 1–25. DOI: 10.1002/0471678384.ch1.

[98] M. J. O'Brien, L. A. van Eykern, and H. A. J. van Vught, "Transcutaneous respiratory electromyographic monitoring," *Critical Care Medicine*, vol. 15, no. 4, pp. 294–299, 1987. DOI: 10.1097/00003246-198704000-00002.

[99] O. S. Oguz, Z. Zhou, S. Glasauer, and D. Wollherr, "An inverse optimal control approach to explain human arm reaching control based on multiple internal models," *Scientific Reports*, vol. 8, no. 1, 2018. DOI: 10.1038/s41598-018-23792-7.

[100] M. Omrani, M. T. Kaufman, N. G. Hatsopoulos, and P. D. Cheney, "Perspectives on classical controversies about the motor cortex," *Journal of Neurophysiology*, vol. 118, no. 3, pp. 1828–1848, 2017. DOI: 10.1152/jn.00795.2016.

[101] N. Östlund, J. Yu, K. Roeleveld, and J. S. Karlsson, "Adaptive spatial filtering of multichannel surface electromyogram signals," *Medical & Biological Engineering & Computing*, vol. 42, no. 6, pp. 825–831, 2004. DOI: 10.1007/bf02345217.

[102] D. J. Ostry and A. G. Feldman, "A critical evaluation of the force control hypothesis in motor control," *Experimental Brain Research*, vol. 153, no. 3, pp. 275–288, 2003. DOI: 10.1007/s00221-003-1624-0.

[103] T. Oya, S. Riek, and A. G. Cresswell, "Recruitment and rate coding organisation for soleus motor units across entire range of voluntary isometric plantar flexions," *The Journal of Physiology*, vol. 587, no. 19, pp. 4737–4748, 2009. DOI: 10.1113/jphysiol.2009.175695.

[104] J. Perry and G. A. Bekey, "EMG-force relationships in skeletal muscle," *Critical reviews in biomedical engineering*, vol. 7, no. 1, pp. 1–22, 1981.

[105] J. M. Petit, G. Milic-Emili, and L. Delhez, "Role of the diaphragm in breathing in conscious normal man: An electromyographic study," *Journal of Applied Physiology*, vol. 15, no. 6, pp. 1101–1106, 1960. DOI: 10.1152/jappl.1960.15.6.1101.

[106] M. Piotrkiewicz and K. S. Türker, "Onion skin or common drive?" *Frontiers in Cellular Neuroscience*, vol. 11, 2017. DOI: 10.3389/fncel.2017.00002.

[107] L. Piquilloud *et al.*, "Neurally adjusted ventilatory assist improves patient–ventilator interaction," *Intensive Care Medicine*, vol. 37, pp. 263–271, 2011. DOI: 10.1007/s00134-010-2052-9.

[108] P. von Platen, A. Pomprapa, B. Lachmann, and S. Leonhardt, "The dawn of physiological closed-loop ventilation—a review," *Critical Care*, vol. 24, no. 1, 2020. DOI: 10.1186/s13054-020-2810-1.

[109] R. Plonsey and R. C. Barr, *Bioelectricity, A Quantitative Approach*, third edition. Springer, 2007.

[110] C.-S. Poon, "Respiratory models and control, Molecular, cellular, and tissue engineering," in *The Biomedical Engineering Handbook*, J. D. Bronzino and D. R. Peterson, Eds., Fourth Edition. CRC Press, 2015, ch. 20, ISBN: 978-1-4398-2531-0.

[111] J. R. Potvin and A. J. Fuglevand, "A motor unit-based model of muscle fatigue," *PLOS Computational Biology*, vol. 13, no. 6, A. M. Haith, Ed., e1005581, 2017. DOI: 10.1371/journal.pcbi.1005581.

[112] R. T. Raikova, P. Krutki, H. Aladjov, and J. Celichowski, "Variability of the twitch parameters of the rat medial gastrocnemius motor units - experimental and modeling study," *Computers in Biology and Medicine*, vol. 37, no. 11, pp. 1572–1581, 2007. DOI: 10.1016/j.compbiomed.2007.02.006.

[113] J. W. Robertson, "Modifying muscle properties in a leading neuromuscular model: The Fuglevand model revisited," Ph.D. dissertation, University of Calgary, 2014.

[114] F. Rohrer, "Der Strömungswiderstand in den menschlichen Atemwegen und der Einfluss der unregelmässigen Verzweigung des Bronchialsystems auf den Atmungsverlauf in verschiedenen Lungenbezirken," *Pflüger's Archiv für die gesamte Physiologie des Menschen und der Tiere*, vol. 162, no. 5, pp. 225–299, 1915. DOI: 10.1007/BF01681259.

[115] J. E. Rubin *et al.*, "Multiple rhythmic states in a model of the respiratory central pattern generator," *Journal of Neurophysiology*, vol. 101, no. 4, pp. 2146–2165, 2009. DOI: 10.1152/jn.90958.2008.

[116] J. Sauer, "Cardiac artefact removal on surface EMG signals of the respiratory muscles," Bachelor's Thesis, Universität zu Lübeck, 2016.

[117] M. Schmidt *et al.*, "Neurally adjusted ventilatory assist and proportional assist ventilation both improve patient-ventilator interaction," *Critical Care*, vol. 19, no. 1, p. 56, 2015. DOI: 10.1186/s13054-015-0763-6.

[118] R. M. Schwartzstein and M. J. Parker, *Respiratory physiology: a clinical approach*. Philadelphia: Lippincott Williams & Wilkins, 2005, ISBN: 9780781757485.

[119] Z.-H. Shi *et al.*, "Expiratory muscle dysfunction in critically ill patients: Towards improved understanding," *Intensive Care Medicine*, vol. 45, no. 8, pp. 1061–1071, 2019. DOI: 10.1007/s00134-019-05664-4.

[120] Z. Shi *et al.*, "Diaphragm pathology in critically ill patients with COVID-19 and postmortem findings from 3 medical centers," *JAMA Internal Medicine*, vol. 181, no. 1, p. 122, 2021. DOI: 10.1001/jamainternmed.2020.6278.

[121] C. Sinderby *et al.*, "Neural control of mechanical ventilation in respiratory failure," *Nature Medicine*, vol. 5, no. 12, pp. 1433–1436, 1999. DOI: 10.1038/71012.

[122] R. E. Singh, K. Iqbal, G. White, and T. E. Hutchinson, "A systematic review on muscle synergies: From building blocks of motor behavior to a neurorehabilitation tool," *Applied Bionics and Biomechanics*, vol. 2018, pp. 1–15, 2018. DOI: 10.1155/2018/3615368.

[123] D. F. Stegeman, B. U. Kleine, B. G. Lapatki, and J. P. Van Dijk, "High-density surface EMG: Techniques and applications at a motor unit level," *Biocybernetics and Biomedical Engineering*, vol. 32, no. 3, pp. 3–27, 2012. DOI: 10.1016/s0208-5216(12)70039-6.

[124] M. S. Stock, T. W. Beck, and J. M. Defreitas, "Effects of fatigue on motor unit firing rate versus recruitment threshold relationships," *Muscle & Nerve*, vol. 45, no. 1, pp. 100–109, 2011. DOI: 10.1002/mus.22266.

[125] C. Straus, I. Arnulf, T. Similowsky, and J.-P. Derenne, "Control of breathing: Neural drive," in *Applied Physiology in Respiratory Mechanics*, Springer Milan, 1998, pp. 1–19. DOI: 10.1007/978-88-470-2928-6_1.

[126] T. Takei *et al.*, "Neural basis for hand muscle synergies in the primate spinal cord," *Proceedings of the National Academy of Sciences*, vol. 114, no. 32, pp. 8643–8648, 2017. DOI: 10.1073/pnas.1704328114.

[127] A. W. Thille *et al.*, "Patient-ventilator asynchrony during assisted mechanical ventilation," *Intensive Care Medicine*, vol. 32, no. 10, pp. 1515–1522, 2006. DOI: 10.1007/s00134-006-0301-8.

[128] E. Todorov and M. I. Jordan, "Optimal feedback control as a theory of motor coordination," *Nature Neuroscience*, vol. 5, no. 11, pp. 1226–1235, 2002. DOI: 10.1038/nn963.

[129] M. C. Tresch and A. Jarc, "The case for and against muscle synergies," *Current Opinion in Neurobiology*, vol. 19, no. 6, pp. 601–607, 2009. DOI: 10.1016/j.conb.2009.09.002.

[130] L. Urry, *Campbell biology in focus*. Boston: Pearson, 2016, ISBN: 9781292109589.

[131] M. Van Cutsem, J. Duchateau, and K. Hainaut, "Changes in single motor unit behaviour contribute to the increase in contraction speed after dynamic training in humans," *The Journal of Physiology*, vol. 513, no. 1, pp. 295–305, 1998. DOI: 10.1111/j.1469-7793.1998.295by.x.

[132] M. Van Cutsem, P. Feiereisen, J. Duchateau, and K. Hainaut, "Mechanical properties and behaviour of motor units in the tibialis anterior during voluntary contractions," *Canadian Journal of Applied Physiology*, vol. 22, no. 6, pp. 585–597, 1997. DOI: 10.1139/h97-038.

[133] A. D. Vecchio *et al.*, "Tutorial: Analysis of motor unit discharge characteristics from high-density surface EMG signals," *Journal of Electromyography and Kinesiology*, vol. 53, p. 102 426, 2020. DOI: 10.1016/j.jelekin.2020.102426.

[134] S. Walterspacher *et al.*, "Activation of respiratory muscles during respiratory muscle training," *Respiratory Physiology & Neurobiology*, vol. 247, pp. 126–132, 2018. DOI: 10.1016/j.resp.2017.10.004.

[135] D. M. Wolpert and Z. Ghahramani, "Computational principles of movement neuroscience," *Nature Neuroscience*, vol. 3, no. Supp, pp. 1212–1217, 2000. DOI: 10.1038/81497.

[136] J. Woods and B. Bigland-Ritchie, "Linear and non-linear surface EMG/force relationships in human muscles. an anatomical/functional argument for the existence of both," *American journal of physical medicine*, vol. 62, no. 6, pp. 287–299, 1983.

[137] G. T. Yamaguchi, *Dynamic Modeling of Musculoskeletal Motion*, 1st edition. Springer Science+Business Media Dordrecht, 2001.

[138] D. Yatsenko, D. McDonnall, and K. S. Guillory, "Simultaneous, proportional, multi-axis prosthesis control using multichannel surface EMG," in *2007 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, IEEE, 2007. DOI: 10.1109/iembs.2007.4353749.

[139] T. Yoshida, Y. Fujino, M. B. P. Amato, and B. P. Kavanagh, "Fifty years of research in ARDS. spontaneous breathing during mechanical ventilation. risks, mechanisms, and management," vol. 195, no. 8, pp. 985–992, 2017. DOI: 10.1164/rccm.201604-0748cp.

[140] M. Younes, "Proportional assist ventilation, a new approach to ventilatory support: Theory," *American Review of Respiratory Disease*, vol. 145, no. 1, pp. 114–120, 1992. DOI: 10.1164/ajrccm/145.1.114.

[141] P. Zipp, "Effect of electrode parameters on the bandwidth of the surface e.m.g. power-density spectrum," *Medical & Biological Engineering & Computing*, vol. 16, no. 5, pp. 537–541, 1978. DOI: 10.1007/bf02457805.

MODELING AND BLIND SOURCE SEPARATION OF RESPIRATORY SURFACE EMG
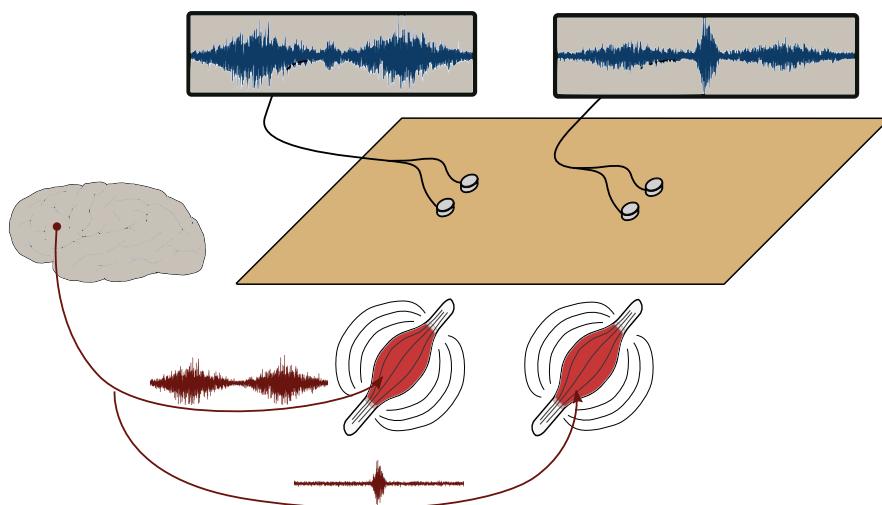
Figure 4.1 – An illustration of electrophysiological *crosstalk*. Two muscles are active at the same time, thus both generating time-varying electrical fields in the surrounding biological tissue. Electrodes — placed, for example, on the skin surface — can only pick up the *superposition* of the two muscles' electrical activity. Separating the effects of multiple sources when only mixture signals are measurable and the properties of the mixing system are unknown is called *blind source separation* (BSS). In this chapter, BSS is used to recover the activity of multiple respiratory muscles from surface EMG measurements at multiple locations, and a comprehensive surface EMG simulation model is developed specifically for the validation of the BSS procedure.

Contents

MUSCLES AND NEURONS generate electrical field fluctuations when they are activated, and these field fluctuations can be picked up by properly placed recording electrodes. All electrophysiological measurements rely on this basic principle, including electromyography (EMG), electrocardiography (ECG), and electroencephalography (EEG), among others. These measurement modalities differ mainly in the type and placement of the electrodes and the subsequent signal processing. As a corollary, which physiological processes an electrophysiological measurement observes — e.g., cardiac or brain activity, respiration, or biceps activation — is determined solely by electrode type and placement as well as subsequent signal processing. If multiple muscles or neurons are close enough to the recording electrode, the recorded signal will inevitably represent a superposition of the activity of these different physiological units. Separating these different contributions from one another in the recorded signal, and thus obtaining a *cleaned* signal exhibiting only the physiological activity of interest, is a ubiquitous task in biomedical signal processing. This is the domain of *source separation algorithms*.

A vast number of algorithms for recovering unknown source signals from measurements of mixtures of these sources have been proposed, based on different assumptions about the source signals, the mixing system, and the available measurements. In this chapter, we will concern ourselves with the development and validation of a novel approach for separating the various sources that are present in respiratory surface EMG measurements: inspiratory and expiratory activity from multiple muscles, crosstalk from other muscles, cardiac activity, and further physiological noise sources.[1] For this purpose, we will employ a particularly general multi-input multi-output (MIMO) blind source separation (BSS) framework, namely the TRIple-N Independent component analysis for CONvolutive mixtures (TRINICON) [9, 10] framework. Besides the design of an effective algorithm, there is a second, crucial challenge when proposing a solution to any physiological source separation: validation. How can the correctness of the estimated sources computed by the algorithm be evaluated? In most physiological settings, reference measurements of the desired sources are inaccessible; this is also the case for the present electrophysiological separation task — it is impossible to gain reference measurements of different physiological processes *measured at the same location and time and with the same equipment as the actual measurements* without also measuring the disturbances resulting from the other source signals. One way to validate a proposed BSS algorithm despite this challenge is to use precise physiological simulation models to generate synthetic mixture signals for which the original, correct sources are known. This is the path we follow here, and for this purpose, a new and comprehensive mathematical model of surface EMG and force signal generation is proposed and implemented. The proposed, TRINICON-based BSS procedure is then validated both on real measurements (qualitatively) and simulated data (quantitatively).

Section 4.1 defines the precise problem under consideration in this chapter and briefly summarizes the state of the art in, firstly, blind source separation for surface EMG signals and, secondly, surface EMG modeling and simulation. Next, section 4.2 proceeds to describe a newly proposed, comprehensive respiratory surface EMG simulation model. In the following section 4.3, the actual source separation algorithm is described, which represents a specific instance of the TRINICON framework proposed by Buchner *et al.* [9]. The algorithm is then validated using simulated measurements in section 4.5. Finally, section 4.6 concludes the

---

[1]The derivation of the algorithm itself has not been performed by the author and is, instead, the work of H. Buchner. It is reproduced here for completeness and not the focus of this chapter.

chapter with a brief discussion and some proposals for further research.[2]

4.1 Problem definition & state of the art

Section 4.1.1 briefly defines the type of separation problem considered in this chapter. In section 4.1.2, a brief overview of previous algorithms for its solution is presented. Since electrophysiological simulation plays an important role in the validation of the BSS algorithm presented in this chapter, and a complex physiological model is developed for precisely that purpose, section 4.1.3 provides an overview of the start of the art in EMG signal modeling.

4.1.1 Problem definition

Given $P = 4$ measurements $\mathrm{EMG}_k^1, \dots, \mathrm{EMG}_k^P$ of respiratory surface EMG measurements, which represent mixtures of multiple physiological phenomena, we aim to isolate Q source signals $s_k^1, \dots, s_k^Q$ representing these individual physiological processes from the mixture measurements. This is a classical (discrete-time) multi-input multi-output (MIMO) separation problem since both multiple sources as well as multiple measurement signals are involved. We will aim for source separation on the *muscle group level*, i.e., the aim is neither to identify the contributions of individual motor units nor those of individual muscles: instead, we want to separate the contributions of inspiratory muscles, expiratory muscles, cardiac activity, and further noise sources. Thus, we assume that $Q = P = 4$. Finally, we do not want to make any specific assumptions about the time course of the individual sources' activation signals,[3] i.e., we want to perform *blind* source separation (BSS).

4.1.2 Source separation algorithms for surface EMG measurements

The literature on MIMO BSS is vast. The proposed approaches can broadly be classified along three dimensions:

The mixing model. Any separation algorithm must make an assumption about the way in which the source signals are *mixed* in the available measurements. The two most popular mixing models are (linear) *instantaneous* mixing and (linear) *convolutive* mixing. *Instantaneous* mixing assumes that the measurement signals $x_k^1 \dots x_k^P$ represent instantaneous linear combinations of the source signals, i.e.,

$$x_k^p = \sum_{q=1}^{Q} h_{qp} s_k^q \quad \forall\, p = 1, \dots, P,$$

---

[2]Parts of this chapter have been the subject of a number of previous journal and conference publications (co-)authored by the author [12, 73, 75, 77, 78]. A preliminary version of the mathematical model of surface electromyography was described by Petersen [75], and a much more complete version, on which the present chapter is built, by Petersen and Rostalski [78]. The model was used by Olbrich *et al.* [73] in a system identification context, an application which is not further discussed here. The BSS algorithm was first described by Buchner *et al.* [12] and comprehensively evaluated using an earlier version of the sEMG model by Petersen *et al.* [77].

[3]We will assume the very general properties of non-Gaussianity, non-whiteness, and non-stationarity; see section 4.3 for details.

where $h_{qp} \in \mathbb{R}$ quantifies the influence of source signal s^q on mixture signal x^p. *Convolutive* mixing, on the other hand, assumes that the source signals are filtered by different linear FIR filters on their way to the different sensors, i.e.

$$x_k^p = \sum_{q=1}^{Q} \underline{h}_{qp}^\mathsf{T} \underline{s}_k^q \quad \forall \; p = 1, \dots, P \tag{4.1}$$

with the FIR coefficients $\underline{h}_{qp} \in \mathbb{R}^L$ and the delayed source signal vector

$$\underline{s}_k^q = \begin{pmatrix} s_k^q & s_{k-1}^q & \cdots & s_{k-L+1}^q \end{pmatrix}^\mathsf{T}.$$

The relationship between P and Q. If $P > Q$ (there are more measurements than sources), the separation problem is said to be *overdetermined*; if $P < Q$ (there are more sources than measurements), the separation problem is said to be *underdetermined*. This has consequences for the type of prior knowledge about the source signals that is required to achieve successful separation.

Assumptions made about the source signals. *Some* assumptions about the characteristics of the source signals must be made to enable recovering the source signals from the mixtures. Typical broad assumptions include statistical independence of the individual sources, sparsity (in the time domain, frequency domain, or a time-frequency representation), as well as non-Gaussianity, non-stationarity and non-whiteness of the source signals. More specific assumptions power various special algorithms.

Possibly the simplest algorithm for performing BSS is a simple principal component analysis (PCA) applied to the signal matrix (or a time-frequency or frequency representation of the signals) [50]. Another closely related [18, 50] and widely popular algorithm is *independent component analysis (ICA)* [16–18, 50, 51], which assumes instantaneous mixing as well as non-Gaussianity and statistical independence of the source signals. In ICA, the mixing coefficients and the source signals are identified by minimizing a suitably chosen cost function such as the *mutual information* between the source signals [50]. *Non-negative matrix factorization (NMF)* [55] represents the basis for another large group of algorithms, based on the assumption of source signal positivity [14]. (As opposed to ICA, NMF does *not* require statistical independence between the source signals.) In a unifying effort, Buchner *et al.* [9] have developed the general TRIple-N Independent component analysis for CONvolutive mixtures (TRINICON) framework for broadband adaptive MIMO signal processing algorithms, which encompasses many of the previously known algorithms as special cases. In this framework, the non-Gaussianity, non-stationarity and non-whiteness of the source signals can be exploited simultaneously to efficiently minimize the mutual information between the source signals. Both instantaneous as well as convolutive mixture models can be treated.

Independent Component Analysis

Non-Negative Matrix Factorization

Shifting from general BSS to the specific application of multi-channel EMG source separation, the literature is expansive as well. In this setting, separation can be performed at a number of different levels [45], see fig. 4.2 for an overview. Firstly, on the highest abstraction level, one can aim to identify the activity of different *muscle groups*, striving to identify one source activity signal — often called a *muscle synergy* — per muscle group. This

will often result in an over-determined separation problem, as there are usually more measurement channels than relevant muscle groups. Muscle-group activity estimation (i.e., *muscle synergy identification*) is often performed using envelope signals and an instantaneous mixing model, using, e.g., NMF [1, 7, 37, 49, 88]. At a more granular level, one may attempt to identify the contributions of *individual muscles*. For this identification problem, one often roughly has $P \approx Q$, except for high-density surface EMG applications, where $P > Q$. Since the solution of this separation task mainly requires the precise identification and elimination of *crosstalk*, many proposed solutions employ a convolutive mixing model in order to capture the effect of the different electrical transmissions paths between the muscles and the different electrodes [46, 48, 68]. Finally, at the lowest level, separation can also be performed at the level of *individual motor units (MUs)*. This is usually only feasible with HD-sEMG, but even then, due to the large number of MUs, the separation task is often underdetermined. At this level, two types of algorithms have been proposed. The *convolution kernel compensation* (CKC) approach proposed by Holobar and Zazula [46] is a variation on classical convolutive blind source separation techniques, whereas the separation algorithm proposed by De Luca *et al.* [24] and Nawab *et al.* [67] relies on template matching for the individual MUAPs. Intense scientific debate has ensued about the validity — and the means for correct validation — of the two approaches [23], with the method proposed by De Luca *et al.* [24] and Nawab *et al.* [67] being commercially available (NeuroMap, Delsys Incorporated, Natick, MA, USA). In particular, there is an open debate about the correct procedure for proper experimental validation of source separation approaches [22, 23, 31, 33], which is one of the reasons why we have selected a simulation-based validation approach here — the other reason being that the discussed validation methods are designed for the validation of MU-level separation algorithms, not for muscle synergy identification.

4.1.3 Surface EMG simulation models

Over the decades, many mathematical models have been proposed for all aspects of the neuromuscular system that contribute to EMG signal generation, ranging from biophysical process models at the sarcomere level [64] over Hodgkin-Huxley type models of action potential propagation [43], models of motor unit rate coding and recruitment [30] and the distributions of electromechanical properties in motor unit pools [38] to purely statistical descriptions of the resulting signals [60]. Since the purpose of the comprehensive model developed in this chapter is the generation of realistic surface EMG signals for the validation of BSS algorithms, certain aspects of physiology can be described in a simplified way, whereas others need to be represented faithfully to obtain realistic simulated sEMG signals. In particular, a precise simulation of processes at the cellular level is *not* required because a simplification of these has rather marginal consequences for mixture signals observed at the skin surface. Three aspects of physiology have a particularly strong effect on the resulting signals:

1. The geometry and conducting properties of the muscle and the surrounding tissue have a huge influence on the recorded signals. This includes the way individual muscle fibers and motor units are distributed throughout the muscle [8, 34, 72].

2. The organization of the motor unit pool similarly has a strong effect on the resulting
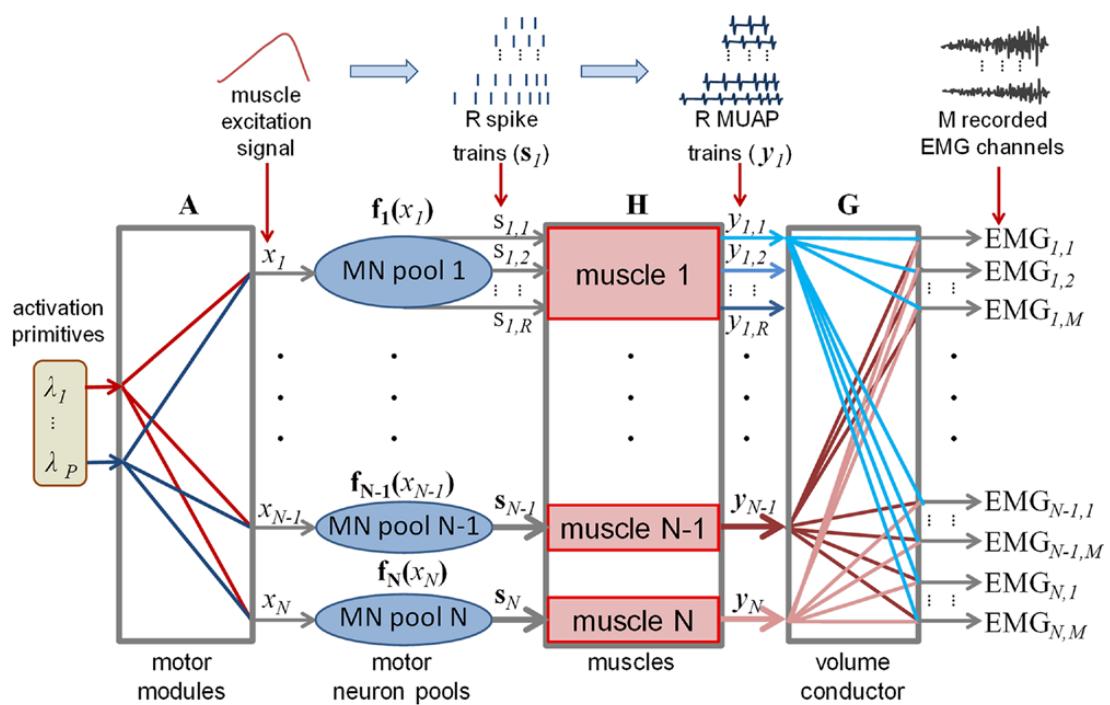
Figure 4.2 – An overview of the different physiological levels of muscle activation, as well as the various transmission and mixing processes that connect activation signals at the different levels. Source identification techniques can be used to recover activation signals at different levels, including the *muscle synergy layer* (recovering *activation primitives*), the *individual muscle layer* (recovering *muscle excitation signals*), and the *individual motor unit layer* (recovering *MU spike trains*). Reproduced with permission from Holobar and Farina [45].

EMG-force characteristics [38]. This includes aspects such as the progressive recruitment of the different motor units, the rate coding of already-recruited MUs, and the stochasticity of the firing process [25].

3. The characteristics of the muscle as a whole, including the EMG-force relationship, are an emerging property that results from the relationship between many individual variables at the motor unit level, such as action potential amplitude, conduction velocity, twitch force, recruitment threshold, and more [38]. The relationship between many of these properties is governed by what is known in the literature as the *size principle* [40–42].

For many years, the model proposed by Fuglevand *et al.* [38] was the state of the art in simulation models of the complete muscle, and it has been reused and extended in a large number of further modeling studies [13, 29, 82, 84, 85]. There have, however, been a number of important advances in different aspects of EMG modeling, warranting an update of this classic model. Farina and Merletti [32], Farina *et al.* [35], and Farina and Rainoldi [36] have proposed analytical models of intracellular action potential propagation and 3-D volume conduction, taking the different conductivities of the (anisotropic) muscle, fat, and skin layers into account, as well as the effect of the electrode recording system. Volume conduction models for more realistic geometries have been proposed using the finite element method (FEM) [8, 57, 58, 63, 92, 96, 97]. The organization of motor unit pools has been the subject of much scientific attention, and as a result, several improved models of rate coding and recruitment have been proposed [20, 21, 30]. The inter-spike interval variability, i.e., the stochastic of the MU firing intervals, has been assessed in detail, models have been proposed, and the effect on the surface EMG and force signals has been analyzed [6, 25, 65]. An integration of many of these advances into a single, comprehensive model was missing, a need that is met by our newly proposed model [78], which will be described in the following section.

4.2 Realistic modeling of respiratory surface EMG measurements

In the context of blind source separation, the purpose of physiological modeling is twofold: firstly, the choice of a proper source separation model should be driven by a deep understanding of physiology. And secondly, realistic simulation models are one of the very few ways in which source separation algorithms can be properly validated, as discussed above. In this section, a comprehensive mathematical model of motor unit pool organization, surface electromyography, and force generation will be described. The model was the subject of a recent journal publication by the author [78] and consolidates various strands of research on improvements in neuromuscular modeling, alluded to in the previous section. Several aspects of the proposed model represent combinations of (improved versions of) previously published model components, while other components are completely new, designed to adhere to recent experimental findings or to provide more modeling flexibility. Two mathematical properties of the employed volume conduction model by Farina and Merletti [32] are proved, both of which enable a clearer interpretation of the model's properties. Finally, a static nonlinear input transformation is proposed that enables easier specification of the simulated force trajectory, as well as a more transparent

analysis of the simulated muscle's properties. In another publication of the author [77], the model was combined with an ECG simulation model and further additive noise to yield a full simulation model of respiratory surface EMG measurements, feasible to serve as a validation tool for BSS algorithms (as was the initial intention in designing the model). The combined model will be employed for exactly this purpose in section 4.5.

The following section 4.2.1 provides a brief overview of the mentioned comprehensive model of surface EMG and force generation. As the model has already been the subject of a very comprehensive publication [78], only its main characteristics are sketched here. The reader is referred to Petersen and Rostalski [78] for a more complete description of the model and a justification of all modeling choices made. Section 4.2.2 then describes the integration of that model with an ECG model to yield a realistic simulation of respiratory surface EMG measurements.

4.2.1 A comprehensive mathematical model of surface EMG and force generation

The model and the simulation process can roughly be separated into three distinct phases:

1. The specification and generation of all simulated muscles, their MU pools, and the properties of these MUs (and the fibers they contain). The modeling components relevant for this phase specify the distributions of and relations between the various properties of the simulated MUs, as well as their geometrical distribution throughout the muscle.

2. The calculation of the motor unit-electrode transfer functions, and the motor unit action potentials (MUAPs) resulting at each simulated electrode from a single activation of each MU. This is where the action potential propagation and volume conduction models play a key role.

3. The simulation of the stochastic firing processes of each individual MU. Each firing of each MU generates a corresponding force twitch and MUAP at each electrode, the superposition of all of which yields the simulated force and surface EMG signals.

Figure 4.3 shows a schematic overview of the path from a muscle activation input to the output force and EMG signals that is followed by the proposed model; it may provide useful orientation to the reader while following along the main text. The following sections will each provide a brief overview of the modeling structure and the main novelties regarding each of these phases.

Motor unit pool organization

First, recall the main features of motor unit pool organization, discussed in more detail in section 3.1.1:

- Muscles are organized into pools of motor units (MUs), each consisting of a motor neuron and the muscle fibers it innervates.

- MUs vary widely in their properties. They are organized according to the *size principle*, which states that most relevant electrical or mechanical properties correlate in some way with MU size, i.e., the number of innervated fibers.
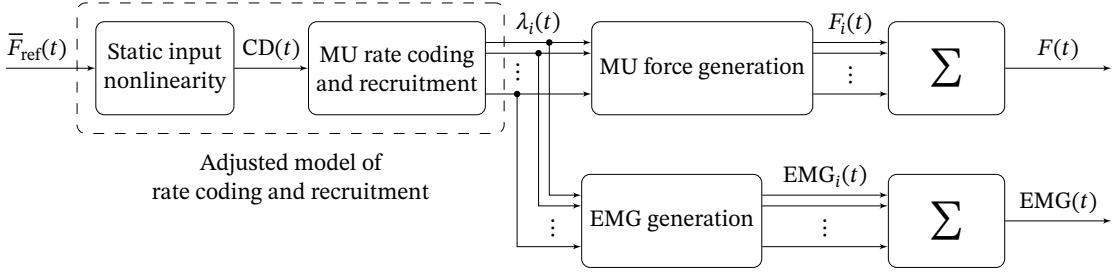
Figure 4.3 – A schematic overview of the proposed comprehensive model of motor unit pool organization, surface electromyography, and force generation. A target muscle force trajectory $\overline{F}_{\text{ref}}(t)$ is defined, from which the common drive level $\text{CD}(t)$ necessary to achieve this force level is calculated by inverting eq. (4.3). Using the rate coding and recruitment model described in eq. (4.2), the instantaneous firing rates $\lambda_i(t)$ of the individual MUs are calculated. From these, firing instants are sampled stochastically, the EMG and force twitches are simulated and superposed, and the total generated muscle force $F(t)$ and the EMG signal $\text{EMG}(t)$ at a particular recording location on the skin surface are obtained. Reproduced and modified with permission from Petersen and Rostalski [78].

- The distribution of MU and muscle fiber properties is *continuous*, i.e., there are no distinct *types* of MUs and muscle fibers.

- MUs in a muscle (and possibly in synergistic muscle groups) receive a *common drive* input signal from the CNS. This common drive controls their recruitment and rate coding.

- Each MU has a distinct recruitment threshold (which is also strongly correlated with MU size) in terms of the common drive to the MU pool. MUs are de-recruited in inverse order of recruitment.

- Under many circumstances, rate coding conforms with the *onion skin phenomenon*, which states that earlier-recruited MUs discharge at a higher rate than later-recruited ones.

The novel mathematical model that we proposed [78] conforms with each of these observations and proposes new model components for most of them.

As opposed to many previous publications [2, 84], we model MU properties as being drawn from a continuous, multivariate distribution (as opposed to distinct types).[4] To begin with the most prominent size principle parameter, two exponential models have been proposed in the literature for the distribution of the *MU recruitment thresholds*. Fuglevand *et al.* [38] have proposed to use the relationship

$$\text{CD}_i^{\text{rec}} = \frac{e^{ai}}{100} \quad \text{with} \quad a = \frac{\log(100 \cdot \text{CD}^{\text{full}})}{N^{\text{MU}}},$$

where CD_i^{rec} denotes the recruitment threshold of MU i in terms of the common drive input signal, and CD^{full} denotes the level of common drive at which the last (and largest) MU is

MU Recruitment Thresholds

---

[4]While Fuglevand *et al.* [38] did model MU properties as following a continuous relationship with MU size, their purely deterministic model ignored the stochasticity of this relationship [40] and did not take into account many of the size principle parameters we consider here.

recruited. Later, De Luca and Contessa [20] have proposed the slightly altered model

$$\mathrm{CD}_i^{\mathrm{rec}} = \frac{bi}{N^{\mathrm{MU}}} \cdot \frac{e^{ai}}{100} \quad \text{with} \quad a = \frac{\log(\frac{100 \cdot \mathrm{CD}^{\mathrm{full}}}{b})}{N^{\mathrm{MU}}}.$$

The decision for either of these models should be made depending on the characteristics of the muscle under consideration. Following the size principle, we then proceed to define all other MU and fiber properties as stochastic functions of the recruitment threshold. These *size principle parameters* in our model include the peak MU twitch force, peak fiber twitch force, peak single-fiber action potential (SFAP) amplitude, action potential conduction velocity, and various model parameters characterizing the speed of contraction. All of these play a role in the characterization of the EMG and force twitches of the individual muscle fibers and MUs. For each of these quantities, a deterministic relationship with the recruitment threshold is defined based on physiological measurements and considerations — for details, refer to Petersen and Rostalski [78]. Most of these physiologically observed relationships have not been taken into account in previous modeling studies. Following Contessa and De Luca [19] (who did this for the few parameters they modeled as size principle parameters), we then draw the actual parameter values from a Weibull distribution with the mean of the distribution specified by the previously described deterministic relationships.

Size Principle Parameters

Besides the *recruitment thresholds* $\mathrm{CD}_i^{\mathrm{rec}}$ of the individual MUs, the second model component determining the discharge behavior of the MUs is the *rate coding model*. Many rate coding models have been previously proposed in the literature, which we all found to suffer from one or multiple drawbacks:

Rate Coding Model

- The classical models [30, 38] have used piece-wise linear models, which represent a stark simplification of reality. Moreover, the non-smoothness of these models likely does not conform with physiology.

- More recently, De Luca and Hostage [21] have proposed a linear-exponential model, which solves some of the above issues while suffering from other drawbacks. The model does not describe the significant increase in the slope of the firing rate characteristics above the point of full recruitment that is necessary to compensate for the fact that new MUs can no longer be recruited from this point on.

- All models proposed so far use very few parameters to describe the firing rate characteristics. This is, of course, an advantage because it reduces model complexity, but on the other, this also limits the flexibility with regards to differing experimental results. As an example, De Luca and Hostage [21] observed a *decline* in the initial firing rates of later-recruited units, whereas Erim *et al.* [30] observed an *increase*. The latter behavior cannot be described using the model of De Luca and Hostage [21]. Similarly, the model of De Luca and Hostage [21] does not allow for an adjustment of the degree of convergence of the firing rates for high activation levels, which has been identified as a crucial feature by Fuglevand *et al.* [38].

For these reasons, we proposed to use the following, new model of rate coding in a motor unit pool:

$$\lambda(\mathrm{CD}(t); \mathrm{CD}_i^{\mathrm{rec}}) = -C_1 \cdot (C_2 - \mathrm{CD}(t)) \cdot \mathrm{CD}_i^{\mathrm{rec}} + C_3 \cdot \mathrm{CD}(t) + C_4 - (C_5 - C_6 \cdot \mathrm{CD}_i^{\mathrm{rec}}) \cdot e^{-\frac{\mathrm{CD}(t) - \mathrm{CD}_i^{\mathrm{rec}}}{C_7}},$$
$$(4.2)$$
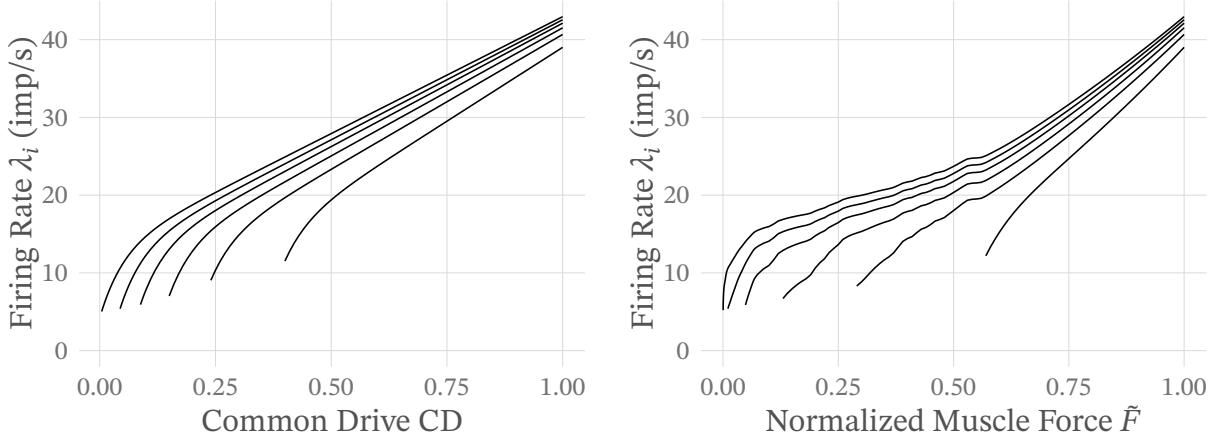
Figure 4.4 – Exemplary firing rate characteristics of every tenth MU generated using the model eq. (4.2). **Panel A** shows the firing rates as a function of the common drive, whereas **panel B** shows them as a function of the (normalized) generated muscle force. Reproduced with permission from Petersen and Rostalski [78].

where $\lambda(\text{CD}(t); \text{CD}_i^{\text{rec}})$ denotes the firing rate of MU i as a function of the common drive $\text{CD}(t)$ and its recruitment threshold CD_i^{rec}, and $C_i, i = 1, \ldots, 7$ are model parameters. The model eq. (4.2) has a number of desirable properties. It models two different phases of the firing rate characteristics: a first phase with a steeper slope of the firing rate characteristics and a second, flatter phase. The transition between the two phases is smooth, and the slope $\mathrm{d}\lambda(\text{CD} = \text{CD}_i^{\text{rec}}; \text{CD}_i^{\text{rec}})/\mathrm{dCD}_i^{\text{rec}}$ of the initial firing rates can be adjusted freely, just like the degree of firing rate convergence for high activation levels. The onion-skin property

$$\frac{\mathrm{d}\lambda(\text{CD}; \text{CD}_i^{\text{rec}})}{\mathrm{dCD}_i^{\text{rec}}} < 0$$

is guaranteed to be fulfilled at all activation levels. Figure 4.4A shows exemplary firing rate characteristics that have been generated using this model.

We have now defined all of the electrical and mechanical properties of the individual MUs, as well as their recruitment and rate coding behavior in terms of the common drive. Although not modeled explicitly, a relationship between the common drive and the resulting muscle force output is thus implicitly defined as an emerging property of the model. Taking the average over individual discharge events, the mean generated muscle force as a (nonlinear) function of the common drive can be computed as

$$\mathrm{E}[F \mid \text{CD}] = \sum_{i=1}^{N^{\text{MU}}} g_i\big(\lambda(\text{CD}; \text{CD}_i^{\text{rec}})\big) \cdot \Omega_i \cdot \lambda(\text{CD}; \text{CD}_i^{\text{rec}}), \tag{4.3}$$

where $g_i(\lambda_i)$ denotes a nonlinear, firing rate–dependent force gain factor[5] and

$$\Omega_i = \int\limits_0^\infty f_i(t)\,\mathrm{d}t = p_i \cdot \kappa^{-m-1} \cdot \Gamma(m+1),$$

denotes the total impulse generated by a single force twitch $f_i(t)$ of MU i, with $\Gamma(x)$ the Gamma function. The remaining quantities are parameters of the force twitch model of Raikova *et al.* [83] which we employed; see there or in Petersen and Rostalski [78] for more details.

The *nonlinear common drive–force relationship* from eq. (4.3) can now be used to further improve the model and our understanding of it. Firstly, it allows us to express the rate coding model (4.2), which is formulated as a function of the common drive, as a function of the generated muscle force instead. This is useful because the common drive in this model is a latent signal that does not correspond to any measurable quantity: in practice, measurements of motor unit recruitment and firing rates are usually given as a function of the generated muscle force. Thus, expressing the rate coding and recruitment model as a function of the generated muscle force is very desirable. Figure 4.4B shows the firing rate characteristics as a function of the generated muscle force, and an important observation can be made: it was mentioned earlier that experimentally observed firing rate characteristics usually exhibit an increase in slope following the point of full recruitment since from there on, an increase in the firing rates of already-recruited MUs is the only way to increase the force output (whereas before that point, firing rate increases and further MU recruitment both act towards this goal). This property now emerges automatically from the model without it being modeled explicitly. Secondly, and at least equally important, we now have a general means to control the simulated muscle in such a way that it deterministically generates the desired muscle force output. To this end, we simply invert the nonlinear relationship between the common drive input and the muscle force output[6] and use this relationship as a *nonlinear input transformation* (see the first block in fig. 4.3), which transforms a desired muscle force output to the common drive input that is required to achieve that level of muscle force. To the author's knowledge, this way of analyzing the nonlinear relationship between the muscle input signal and the generated force output, as well as the use of a linearizing input transformation, are both novel concepts that have not been proposed before. *(margin: Nonlinear Input Transformation)*

(margin: Nonlinear Common Drive–force Relationship)

The last important aspect of a motor unit pool model regards the geometrical arrangement of the individual motor units and the fibers within them. The single-fiber action potential (SFAP) waveform detected by an electrode depends not only on the properties of the action potential and the fiber but also crucially on the distance and the type of tissue separating the fiber and the electrode. Thus, the way in which fibers and motor units are distributed throughout the muscle cross-section has a huge influence on the resulting surface signal. In our model, we propose the following novel method to generate a random *fiber and MU arrangement*: *(margin: Fiber and MU Arrangement)*

1. The muscle is divided into M parts of equal size, each of which will contain N^{MU}/M MUs. This serves to achieve a more uniform distribution of MUs throughout the muscle.

---

[5]We use the expression proposed by Contessa and De Luca [19]; see there or Petersen and Rostalski [78] for details.

[6]The relationship is monotonic, whence inversion is feasible.

2. In each of these muscle parts, MU centers are sampled i.i.d. from a uniform distribution.

3. Each MU is assigned an elliptic region throughout which its fibers are distributed. The area $A_i = N_i^{\text{fib}}/\rho$ of this region is calculated from the number N_i^{fib} of fibers innervated by this MU and the MU fiber density ρ of fibers per unit area.

4. Similarly, fiber locations within the muscle cross-section are sampled i.i.d. from a uniform distribution supported on the MU area. Along the fiber length, the location of the NMJ and the two fiber ends is randomized as well, as described previously [13, 62].[7]

5. Parts of MU territories which exceed the muscle boundaries are cut off, and the number N_i^{fib} of innervated fibers is reduced for these MUs in proportion to the reduction in MU area. To maintain the size principle, the remaining size principle parameters are adjusted to the new MU size as well.

This method achieves a MU and fiber arrangement that attains the following desirable properties:

1. MU regions overlap, which is a well-known feature of muscle physiology.

2. In the probabilistic limit, the overall muscle fiber density remains constant throughout the whole muscle, which is in accordance with various physiological studies [52, 93]. This stands in contrast to various other approaches for handling boundary MUs, as discussed in detail in Petersen and Rostalski [78].

3. All MUs (including the ones on the muscle boundary) conform with the size principle.

A very similar method was used by Fuglevand *et al.* [38]. What differentiates our method from theirs is, firstly, the division of the muscle into muscle regions to reduce fiber density variability, secondly, the special muscle boundary treatment described above, and thirdly, the use of elliptic MU regions as opposed to the less flexible circular regions they used. Carriou *et al.* [13] and Konstantin *et al.* [54] have proposed the use of a deterministic placement algorithm instead of a purely stochastic method as we did. In their method, each new MU is placed such that the distance to the already-placed MUs is maximized. Finally, Navallas *et al.* [66] have followed a slightly different path: they solve the placement problem by explicitly minimizing the variability of the fiber density throughout the muscle. More recently, this method has been extended to also take regionalized MU placement into account [85], which had been observed physiologically [28]. These methods likely produce a more realistic MU and fiber distribution throughout the muscle than our method, but they are also more complex. The choice of the placement algorithm to use will thus depend on the particular target application of the resulting simulation model.

---

[7]This is an important step to prevent unrealistically sharp MUAPs which would result from placing all NMJs and fiber ends at *exactly* the same positions.

Volume conduction and MUAP calculation

Having defined all the electrical and mechanical properties of the individual MUs, as well as the organizing principles of the whole MU pool that makes up a muscle, we now turn to modeling the surface EMG signal: simulating realistic surface EMG (and force) signals was, after all, the initial motivation for creating the present model. For this model component, we reuse the analytic volume conductor model first proposed by Farina and Rainoldi [36] and the intracellular action potential propagation model described by Farina and Merletti [32]. Briefly, each muscle fiber is assumed to lie in an infinite, homogeneous layer of anisotropic *muscle* tissue, above which there are equally infinitely extended layers of isotropic *fat* and *skin* tissue, both of certain thicknesses. Each fiber is represented by a straight line of a certain length, which lies parallel to the skin surface, and with the neuromuscular junction placed at some point along the fiber. The model of Rosenfalck [86] is used for the intracellular action potential (IAP) waveform. The generation (at the NMJ), propagation (along the two fiber halves) and extinction (at the two fiber ends) of the IAP is described as [32]

$$\hat{\imath}(z,t) = \frac{\mathrm{d}}{\mathrm{d}z}\left[\psi(z - z^{\mathrm{NMJ}} - vt)\,p_1(z) - \psi(-z + z^{\mathrm{NMJ}} - vt)\,p_2(z)\right], \qquad (4.4)$$

where z denotes the spatial variable along the muscle fiber, z^{NMJ} the location of the neuromuscular junction (NMJ), and v the IAP's propagation velocity.[8]

$$p_1(z) = H(z - z^{\mathrm{NMJ}}) - H(z - (z^{\mathrm{NMJ}} + L_1)) \quad \text{and} \quad p_2(z) = H(z - (z^{\mathrm{NMJ}} - L_2)) - H(z - z^{\mathrm{NMJ}})$$
$$(4.5)$$

denote the characteristic functions of the two fiber halves, with L_1 and L_2 the distances between the NMJ and the right and left tendon, respectively, and $H(z)$ the Heaviside function. The function

$$\psi(z) = \frac{\mathrm{d}V_m(-z)}{\mathrm{d}z} \qquad (4.6)$$

denotes the voltage gradient across the fiber membrane along the fiber axis, where $V_m(z)$ denotes the trans-fiber membrane voltage wave shape.[9] Here, we use the model

$$V_m(z\,[\mathrm{mm}]) = \begin{cases} 96\,\mathrm{mV}\,\mathrm{mm}^{-3}z^3 e^{-z} - 90\,\mathrm{mV} & \text{if } z > 0, \text{ and} \\ -90\,\mathrm{mV} & \text{otherwise,} \end{cases} \qquad (4.7)$$

originally proposed by Rosenfalck [86], which has also been used in previous studies [13, 32].

Contrary to claims in Farina and Merletti [32] and Petersen and Rostalski [78], $\hat{\imath}(z,t)$ does *not* describe a distributed, time-varying current density source: a fiber-dependent scaling factor and a unit conversion are missing. Assuming a cylindrical homogeneous muscle fiber of radius r, the trans-fiber membrane current per unit length $i_m(z)$ (in A/m) is related to the trans-fiber membrane voltage $V_m(z)$ via [8, 80]

$$i_m(z) = \sigma_{ic}\pi r^2 \frac{\mathrm{d}^2 V_m(z)}{\mathrm{d}z^2},$$

---

[8]We assume a constant IAP propagation velocity here to simplify calculations. It is known, however, that v changes with, e.g., the MU's firing rate [71]. Refer Dideriksen *et al.* [26] for an efficient method to simulate different propagation velocities.

[9]Refer, e.g., to Plonsey and Barr [81] for an in-depth discussion of the relevance of choosing an appropriate model for $V_m(z)$.
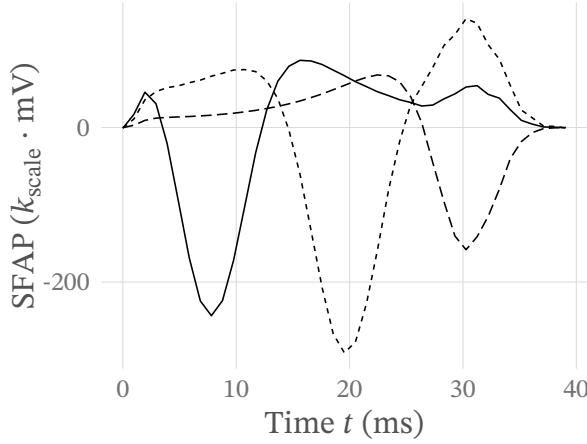
Figure 4.5 – Exemplary simulated single-fiber action potentials (SFAPs). These represent the potential change generated by a single discharge of a single muscle fiber, detected by three different electrodes placed close to the NMJ (solid), in between the NMJ and the fiber end (short dashes) and above the fiber end (long dashes). The constant k_{scale} denotes an arbitrary, static EMG scaling factor. Reproduced with permission from Petersen and Rostalski [78].

where σ_{ic} denotes the intracellular conductivity. Thus, the constant, fiber-dependent scaling factor $\sigma_{ic}\pi r^2$ turns $\hat{\imath}(z, t)$ into a proper current density source, and moreover serves to explicate the (quadratic) dependency of the SFAP amplitude on the fiber diameter.

Combining the IAP model (4.4) to (4.7) with the three-layer volume conductor model of Farina and Rainoldi [36], an analytical expression for the time course of a single action potential measured by a surface electrode at a particular position on the skin surface can be derived [32, 78].[10] The evaluation of said analytical expression requires the evaluation of a double integral [32, 78]. In the author's master's thesis, it was demonstrated that the integration kernel only has removable singularities, and therefore, a numerical integration scheme is guaranteed to converge [74]. The solution of this numerical integration problem is numerically moderately expensive but tractable [13, 78]. Figure 4.5 shows a few single-fiber action potentials (SFAPs) simulated using this model.

While the IAP propagation model (4.4) might appear somewhat *ad hoc*, we could demonstrate in our recent paper [78] that the model can be equivalently formulated as

$$\hat{\imath}(z, t) = \text{GEN}(t)\ \delta(z - z^{\text{NMJ}}) + \psi'(z - z^{\text{NMJ}} - vt)\ p_1(z) + \text{EOF}_1(t)\ \delta(z - z^{\text{NMJ}} - L_1) \quad (4.8)$$
$$+ \psi'(-z + z^{\text{NMJ}} - vt)\ p_2(z) + \text{EOF}_2(t)\ \delta(z - z^{\text{NMJ}} + L_2),$$

with the Dirac distribution δ, the action potential extinction components

$$\text{EOF}_1(t) = -\psi(L_1 - vt) \quad (4.9)$$

and

$$\text{EOF}_2(t) = -\psi(L_2 - vt), \quad (4.10)$$

---

[10]The expression and its derivation are not shown here because they are not new and rather long.

at the two fiber ends, and the potential generation component

$$\text{GEN}(t) = 2\,\psi(-vt) \tag{4.11}$$

at the NMJ. This equivalence is the subject of lemma 1, stated and proved below. While both formulations are mathematically equivalent, the formulation in eq. (4.8) renders, in the author's opinion, the structure of the model and the meaning of the different components more clear: there is a potential generation component at the NMJ, two propagating components moving towards the two fiber ends at constant velocity v, and two corresponding end-of-fiber extinction components. Moreover, we were able to demonstrate that (under mild technical assumptions) it holds for both model formulations that

$$\int_{-\infty}^{\infty} \hat{\imath}(z,t)\,\mathrm{d}z = 0 \quad \forall\, t, \tag{4.12}$$

or in other words: the sum of all incoming and outgoing currents across the fiber is zero at all times. This property is well-justified by the quasi-stationarity of electrophysiological processes [81], and is also in agreement with the predictions of the Hodgkin-Huxley model of action potential propagation [53]. This quasi-stationarity property will be the subject of lemma 2, also stated and proved below. Note that both lemmata and their proofs are reproduced verbatim from Petersen and Rostalski [78].

Lemma 1. *The expressions given in equations (4.4) and (4.8) to (4.11) are equivalent, assuming that $\psi \in C^\infty$.*

Proof. Expanding all terms in equation (4.4) yields

$$\hat{\imath}(z,t) = \frac{\mathrm{d}}{\mathrm{d}z}\big[\,\psi(z - z^{\text{NMJ}} - vt)\,H(z - z^{\text{NMJ}})\big] - \frac{\mathrm{d}}{\mathrm{d}z}\big[\,\psi(z - z^{\text{NMJ}} - vt)\,H(z - z^{\text{NMJ}} - L_1)\big]$$
$$- \frac{\mathrm{d}}{\mathrm{d}z}\big[\,\psi(-z + z^{\text{NMJ}} - vt)\,H(z - z^{\text{NMJ}} + L_2)\big] + \frac{\mathrm{d}}{\mathrm{d}z}\big[\,\psi(-z + z^{\text{NMJ}} - vt)\,H(z - z^{\text{NMJ}})\big].$$

Note that the derivatives can only be understood in the sense of distributions since the derivative of the Heaviside function $H(z)$ can not be defined in the classical sense at $z = 0$. For some basic properties of distributions, refer to appendix 1.

Assuming $\psi \in C^\infty$ and recalling that $\langle(\psi H)', \zeta\rangle = \langle\psi' H + \psi\,\delta, \zeta\rangle$ (refer to appendix 1), the duality $\langle\hat{\imath}(t), \zeta\rangle$ can hence be formulated as

$$\langle\hat{\imath}(t), \zeta\rangle = \langle\psi'_{z^{\text{NMJ}}+vt} H_{z^{\text{NMJ}}}, \zeta\rangle + \langle\psi'_{z^{\text{NMJ}}+vt}\delta_{z^{\text{NMJ}}}, \zeta\rangle - \langle\psi'_{z^{\text{NMJ}}+vt} H_{z^{\text{NMJ}}+L_1}, \zeta\rangle$$
$$- \langle\psi_{z^{\text{NMJ}}+vt}\delta_{z^{\text{NMJ}}+L_1}, \zeta\rangle + \langle\overline{\psi}'_{z^{\text{NMJ}}-vt} H_{z^{\text{NMJ}}-L_2}, \zeta\rangle - \langle\overline{\psi}_{z^{\text{NMJ}}-vt}\delta_{z^{\text{NMJ}}-L_2}, \zeta\rangle$$
$$- \langle\overline{\psi}'_{z^{\text{NMJ}}-vt} H_{z^{\text{NMJ}}}, \zeta\rangle + \langle\overline{\psi}_{z^{\text{NMJ}}-vt}\delta_{z^{\text{NMJ}}}, \zeta\rangle$$
$$= \langle\psi'_{z^{\text{NMJ}}+vt}(H_{z^{\text{NMJ}}} - H_{z^{\text{NMJ}}+L_1}) + \overline{\psi}'_{z^{\text{NMJ}}-vt}(H_{z^{\text{NMJ}}-L_2} - H_{z^{\text{NMJ}}}), \zeta\rangle$$
$$+ \langle(\psi_{z^{\text{NMJ}}+vt} + \overline{\psi}_{z^{\text{NMJ}}-vt})\delta_{z^{\text{NMJ}}}, \zeta\rangle - \langle\psi_{z^{\text{NMJ}}+vt}\delta_{z^{\text{NMJ}}+L_1} + \overline{\psi}_{z^{\text{NMJ}}-vt}\delta_{z^{\text{NMJ}}-L_2}, \zeta\rangle$$

with notations $g' = \frac{\mathrm{d}\,g(z)}{\mathrm{d}z}$, $\overline{g}(x) = g(-x)$, $g_a(x) = g(x - a)$ and $\overline{g}_a(x) = g(-x + a)$. This is exactly equivalent to equations (4.8) to (4.11). $\qquad\square$

Lemma 2. *For compactly supported $\psi(z)$, the intracellular action potential (IAP) model given in equations (4.8) to (4.11) (or equivalently, equation (4.4)) yields a formulation of $\hat{\imath}(z, t)$ that satisfies the condition (4.12). This also holds true if the characteristic functions $p_{1/2}(z)$ are replaced by smooth window functions $w_{1/2}(z)$ which fulfill*

$$w_1(z^{\mathrm{NMJ}} + L_1) = w_1(z^{\mathrm{NMJ}}) = w_2(z^{\mathrm{NMJ}}) = w_2(z^{\mathrm{NMJ}} - L_2) = 0$$

and which yield products $\psi(z)w_{1/2}(z)$ that are differentiable with an integrable derivative. [11]

Proof. First, we consider the case of $\hat{\imath}(z, t)$ as described by equations (4.8) to (4.11), i.e., using the characteristic functions $p_1(z)$, $p_2(z)$. For compactly supported $\psi(z)$,

$$\int_{-\infty}^{\infty} \psi'(z) \, \mathrm{d}z = 0$$

generally holds. Furthermore,

$$\int_{-\infty}^{\infty} \mathrm{GEN}(t) \, \delta(z - z^{\mathrm{NMJ}}) \, \mathrm{d}z = 2\,\psi(-vt) = 2 \int_{-\infty}^{-vt} \psi'(z) \, \mathrm{d}z,$$

$$\int_{-\infty}^{\infty} \mathrm{EOF}_1(t) \, \delta(z - z^{\mathrm{NMJ}} - L_1) \, \mathrm{d}z = -\psi(L_1 - vt) = - \int_{-\infty}^{L_1 - vt} \psi'(z) \, \mathrm{d}z,$$

and equivalently for EOF_2. Combining everything yields

$$\int_{-\infty}^{\infty} \hat{\imath}(z, t) \, \mathrm{d}z = 2 \int_{-\infty}^{-vt} \psi'(z) \, \mathrm{d}z - \int_{-\infty}^{L_1 - vt} \psi'(z) \, \mathrm{d}z - \int_{-\infty}^{L_2 - vt} \psi'(z) \, \mathrm{d}z$$

$$+ \int_{z^{\mathrm{NMJ}}}^{z^{\mathrm{NMJ}} + L_1} \psi'(z - z^{\mathrm{NMJ}} - vt) \, \mathrm{d}z + \int_{z^{\mathrm{NMJ}} - L_2}^{z^{\mathrm{NMJ}}} \psi'(-z + z^{\mathrm{NMJ}} - vt) \, \mathrm{d}z$$

$$= 0.$$

Secondly, considering the case of smooth window functions $w_{1/2}(z)$, we have

$$\int_{-\infty}^{\infty} \hat{\imath}(z, t) \, \mathrm{d}z = \int_{z^{\mathrm{NMJ}}}^{\infty} \frac{\mathrm{d}}{\mathrm{d}z} \left[\psi(z - z^{\mathrm{NMJ}} - vt) \, w_1(z) \right] \mathrm{d}z - \int_{-\infty}^{z^{\mathrm{NMJ}}} \frac{\mathrm{d}}{\mathrm{d}z} \left[\psi(-z + z^{\mathrm{NMJ}} - vt) \, w_2(z) \right] \mathrm{d}z$$

$$= 0$$

by the fundamental theorem of calculus if the conditions on $w_{1/2}(z)$ stated in the lemma hold. $\qquad\square$

---

[11] The Tukey window function employed by Al Harrach *et al.* [2] and Carriou *et al.* [13] falls into the class of window functions supported by this lemma.

Firing train calculation and superposition

Now that the MU pool and all of its properties are set up, and we have a way of calculating the SFAPs as well as the force twitches generated by individual muscle fibers, we can move towards the simulation of complete sEMG and force signals, resulting from many overlapping twitches of many fibers. Assuming linear superposition like most other authors [39, 45], we have

$$\text{EMG}(t) = \sum_{i=1}^{N^{\text{MU}}} \sum_{j=1}^{N_i^{AP}} \text{MUAP}_i(t - t_{ij}), \tag{4.13}$$

where N^{MU} denotes the number of MUs, N_i^{AP} the number of firing events of that MU, t_{ij} the jth firing instant of MU i, and

$$\text{MUAP}_i(t) = \sum_{j=1}^{N_i^{\text{fib}}} \text{SFAP}_{ij}(t) \tag{4.14}$$

the motor unit action potential, obtained by summing over the the SFAPs of all muscle fibers j belonging to MU i. Similarly, for the force signal, we assume that the force generated by the whole muscle is given by the superposition [38]

$$F(t) = \sum_{i=1}^{N^{\text{MU}}} F_i(t) = \sum_{i=1}^{N^{\text{MU}}} \sum_{j=1}^{N_i^{AP}} f_i(t - t_{ij}) \tag{4.15}$$

of the individual force twitches of all MUs, where $F_i(t)$ denotes the force contribution of MU i over time.

The only unknown quantity in section 4.3 and eq. (4.15) are the exact firing instants t_{ij} of all MUs. Their calculation is (slightly) complicated by two facts:

1. The distribution of the inter-spike intervals (ISIs) is known to be *stochastic*, i.e., not exactly determined by the inverse of the firing rate.

2. The common drive, which determines each MU's firing rate, changes over time. Moreover, depending on the exact simulation scenario, these changes may not be known in advance.

We employ the method proposed by Fuglevand *et al.* [38] for the calculation of the ISIs, which conforms with the two above observations. Moreover, to model the decreased ISI variability at increased activation levels, we use the expression proposed by Moritz *et al.* [65] for the coefficient of variation of the ISIs. As opposed to (to the author's knowledge) all previous simulation studies, firing instants in our model are not constrained to sampling instants but may also lie in-between two sampling instants. For more details, refer to Petersen and Rostalski [78].

4.2.2 A simulation model of vector cardiography

To model the effect of cardiac interference on surface EMG measurements, we employ a simple, phenomenological model of vectorcardiography. The model has been proposed by
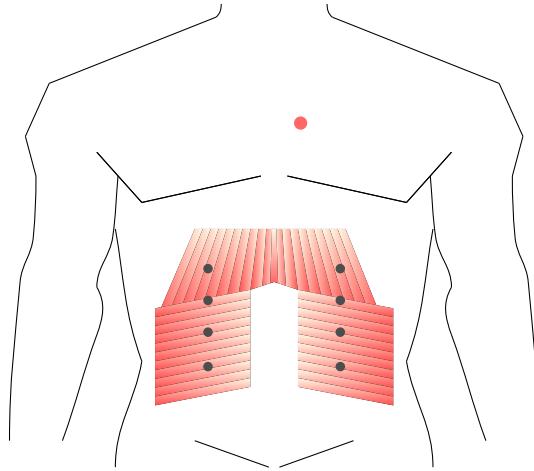
Figure 4.6 – The geometry of the simulation model used for validating source separation on surface respiratory EMG measurements. Three muscles are simulated: the respiratory diaphragm and the left and right transversus abdominis muscles. The red dot on the upper chest marks the position of the heart dipole vector. Reproduced with permission from Petersen *et al.* [77].

Sameni *et al.* [90], with the explicit intent of using it for the validation of a BSS procedure. It describes the heart as a single dipole vector $d(t) \in \mathbb{R}^3$ that rotates with the heart rate within a three-dimensional (infinite) volume conductor, representing the spatial filtering effect of the body tissue. The time-varying measurement signal at an electrode is then given by the projection

$$\phi(t) - \phi_0 = \frac{d(t)^\mathsf{T} r}{4\pi\sigma|r|^3} \tag{4.16}$$

of the heart dipole onto the distance vector $r = x^\text{electrode} - x^\text{heart}$ between the recording electrode and the assumed heart position. In eq. (4.16), $\phi_0 \in \mathbb{R}$ denotes the reference potential, and σ the conductivity of the biological tissues [90]. For the time course of the heart dipole $d(t)$, a Gaussian mixture model inspired by the popular ECG model of McSharry *et al.* [61] is used. For the whole cardiac model component, the implementation provided by the open-source electrophysiological toolbox (OSET) is used [89].

4.2.3 Model parametrization for BSS validation

The proposed model shall now be employed for quantifying the performance of an algorithm for the separation of respiratory surface EMG measurements. The algorithm is discussed in detail in the following section 4.3. To this end, we construct a combined model consisting of the respiratory diaphragm (the main inspiratory muscle), the left and right transversus abdominis muscles (which are important expiratory muscles [94]), and the heart. Four differential measurement channels are simulated; the geometry of the complete simulation model is illustrated in fig. 4.6. The simulated diaphragm consists of 300 MUs, whereas each of the simulated transversus abdominis muscles consists of 100 MUs. The geometry of the simulation model is the same as was used in Petersen *et al.* [77] — however, the simulation was performed again here, since the muscle model has been significantly improved since

that earlier publication (as described in section 4.2.1 and Petersen and Rostalski [78]). The simulated signals are given by

$$\mathrm{EMG}_k^p = \mathrm{EMG}_k^{\mathrm{sim}\,p} + c^{\mathrm{ECG}}\mathrm{ECG}_k^{\mathrm{sim}\,p} + c^\nu \nu_k, \quad p = 1, \ldots, P, \tag{4.17}$$

where $\nu_k \sim \mathcal{N}(0, 1)$, and c^{ECG} and c^ν denote two constant, scalar scaling factors.

4.3 TRINICON-based blind source separation on respiratory surface EMG measurements

This section can be considered an extended description and motivation of the algorithm we proposed recently in Buchner *et al.* [12] and Petersen *et al.* [77]. To briefly recall the problem statement given in section 4.1, we would like to identify $Q = 4$ source signals representing different physiological processes from $P = 4$ measurements $\mathrm{EMG}_k^1, \ldots, \mathrm{EMG}_k^P$ of respiratory surface EMG measurements. This is thus a multi-input multi-output (MIMO) separation problem, and we would like to perform separation at the *muscle group level*. Note that already from this problem specification, it is evident that the aim is *not* to recover some actual physiological signal: there is no physiological signal in the body that could hypothetically be measured, and that captures, e.g., the activity of the inspiratory or expiratory muscles as a group in the desired sense.[12] Moreover, we are not even interested in exactly recovering some specified signal — the aim is simply to obtain separated signals that in some qualitative way capture the activity of the inspiratory muscles, the expiratory muscles, and the heart.

As was also mentioned in section 4.1, the choice of a proper *mixing model* represents an important step in the design of an appropriate source separation procedure. For an EMG signal $\mathrm{EMG}(t)$ measured at a single electrode or a pair of electrodes, we saw in section 4.3 that

$$\mathrm{EMG}(t) = \sum_{i=1}^{N^{\mathrm{MU}}} \sum_{j=1}^{N_i^{AP}} \mathrm{MUAP}_i(t - t_{ij})$$

where $\mathrm{MUAP}_i(t)$ denotes the motor unit action potential (MUAP) measured at the electrode(s) due to a single firing of MU i at time $t = 0$. This can be (and often is [44–46, 68]) rewritten as the convolution

$$\mathrm{EMG}(t) = \sum_{i=1}^{N^{\mathrm{MU}}} \mathrm{MUAP}_i(t) * s^i(t)$$

with the *spike train signal*

$$s^i(t) = \begin{cases} 1 & \text{if } \exists j : t = t_{ij} \\ 0 & \text{else.} \end{cases}$$

---

[12]There is a number of studies claiming to measure "neural respiratory drive". What they usually measure is an invasive EMG signal of the diaphragm obtained from an esophageal catheter. This obviously only measures the activity of one inspiratory muscle, namely, the diaphragm. Moreover, such an invasive EMG signal is, of course, just one of infinitely many conceivable transformations of the neural drive.

Switching to a discrete-time model (and assuming, for simplicity's sake, that firing instants are always sampling instants), one obtains

$$\text{EMG}_k = \sum_{i=1}^{N^{\text{MU}}} \underline{\text{MUAP}}_i^{\mathsf{T}} \underline{s}_k^i \tag{4.18}$$

where

$$\underline{\text{MUAP}}_i = \begin{pmatrix} \text{MUAP}(0) & \text{MUAP}(T_s) & \cdots & \text{MUAP}((L-1) \cdot T_s) \end{pmatrix}^{\mathsf{T}}$$

and

$$\underline{s}_k^i = \begin{pmatrix} s_k^i & s_{k-1}^i & \cdots & s_{k-L+1}^i \end{pmatrix}^{\mathsf{T}},$$

with the sampling time T_s and the MUAP filter length L. If one chooses $Q = N^{\text{MU}}$, eq. (4.18) is exactly identical to the convolutive BSS mixing model introduced in eq. (4.1) . This has been exploited by a number of researchers who used convolutive BSS to decompose multi-channel sEMG signals into the impulse trains of individual MUs [45–47, 68]. This approach, however, requires the use of high-density sEMG measurements to obtain a number of electrodes in the order of the number of relevant MUs,[13] and moreover, it still leaves open the association of the identified MUs with different muscles or muscle groups. For these reasons, this is not a feasible avenue in the setting we're considering.

Instead of using many electrodes and performing separation and identification on the level of single MUs, we thus propose to perform convolutive separation of just $Q = 4$ source signals with $P = 4$ differential sEMG measurements, hypothesizing that this will identify source signals corresponding to the different (inspiratory, expiratory, cardiac) muscles groups that share neural activation signals. This hypothesis corresponds to the assumption that there are inspiratory, expiratory, cardiac, and noise source signals that propagate from four point sources within the body to the four different recording sites, experiencing different linear filtering transformations along the way. This assumption is, of course, highly unrealistic: even the contributions of a *single* muscle (as opposed to a muscle *group*) evoked at different electrodes cannot be described accurately by the same source signal filtered differently with a linear filter for each electrode — due to the differing distances between individual MUs and electrodes. This is, of course, even worse when considering the muscle *group* level, where the inspiratory component of, e.g., an intercostal sEMG measurement is made up of MU firing events completely different from the corresponding diaphragmatic sEMG measurement. The assumed mixing model can thus only hold in a broad, statistical sense: there are stochastic activation signals, the instantaneous power of which roughly captures the neural activation of one of the relevant muscle groups, and the contribution of these activation signals to the different measurements can be identified from their different temporal and frequency signatures. We will assess empirically in section 4.5 whether such a broad, statistical model can be used to reliably separate the activity of the different muscle groups.

For now, the challenge is to estimate $P \cdot Q \cdot L$ filtering coefficients, using only the P available measurement signals $\text{EMG}_k^1, \ldots, \text{EMG}_k^P$. Based on the typical length of a MUAP of a few dozen milliseconds (and a sampling rate of $f_s = 1$ kHz), we choose a demixing filter length of $L = 256$ samples and thus have a total of 4096 coefficients to estimate. Note that we do not

---

[13]In practice, using $Q < N^{\text{MU}}$ electrodes is feasible because relatively few MUs cause most of the potential variation (because they are large or close to the skin surface). For exemplary recent evidence, see, e.g., Botelho *et al.* [8].

attempt to identify the *mixing* filters $\underline{h}_{pq} \in \mathbb{R}^L$ but instead identify *demixing* filters $\underline{w}_{pq} \in \mathbb{R}^L$ such that[14]

$$s_k^q = \sum_{p=1}^{P} \underline{w}_{pq}^{\mathsf{T}} \underline{x}_k^p \quad \forall \, q = 1 \dots Q.$$

As was already mentioned earlier, to solve the resulting convolutive BSS problem, we use the TRIple-N Independent component analysis for CONvolutive mixtures (TRINICON) framework for broadband adaptive MIMO signal processing previously proposed by Buchner *et al.* [9, 10]. This general framework subsumes many previously proposed BSS algorithms, and it incorporates the three complementary assumptions of non-whiteness, non-Gaussianity, and non-stationarity of the source signals.

Algorithm 5 summarizes the algorithm, reproduced and modified (with permission) from Petersen *et al.* [77]. It has been fully derived and implemented by H. Buchner and represents a special case of the general broadband frequency-domain coefficient update equation (10.74) in Buchner *et al.* [9], see section 3.2.2 in that document for more details. First, as a preprocessing step, a spatial prewhitening is performed. To this end, an eigenvalue decomposition of the correlation matrix of the signals measured at the different channels is computed, and the power of the different modes is normalized (see, e.g., Hyvärinen *et al.* [51, p. 159]). This step is crucial in the present application since the level difference between the cardiac component and the EMG and noise components is so large — without the prewhitening step, the numerical BSS problem would often be ill-conditioned. The actual source separation is then performed in the time-frequency domain by minimizing Shannon's mutual information between the identified source signals. Assuming a non-Gaussian joint distribution of the source signals,[15] an expression for the mutual information can be derived analytically [9] that explicitly penalizes stationarity, Gaussianity, and whiteness of the estimated source signals. As a time-frequency distribution, the short-time Fourier transform (STFT) with a Hanning window is chosen, and the resulting numerical minimization problem is solved using a (natural) gradient descent algorithm. The employed broadband separation approach — instead of performing separation purely in the frequency domain — resolves the well-known permutation ambiguity problem that plagues many previously proposed narrowband convolutive BSS methods [10]. To resolve the internal scaling ambiguity — the problem that source signals are only defined up to the application of an arbitrary filtering stage — we employ the *minimum distortion principle* proposed first by Matsuoka *et al.* [59]. While a fully offline implementation is pursued here, a block online implementation of the same algorithm is readily available [9] by simply initializing the optimization problem for each block with the solution of the previous block.

4.4 Simulation-based BSS performance quantification

Even though in a simulation, the true time course of the individual muscles' activity is *known*, it is non-trivial to quantify the performance of a source separation algorithm for

---

[14]The conceptually more difficult identification of the true mixing filters and source signals is known as *multichannel blind deconvolution*, see, e.g., Amari *et al.* [4].

[15]Here, we assume a multivariate Laplacian distribution, because surface EMG signals are known to exhibit spectra that resemble a Laplacian distribution [15, 78].

Algorithm 5: TRINICON-based Convolutive Blind Source Separation

1 **Function** CBSS($x^1, \dots, x^P$)

Input : Mixture measurements $x^1, \dots, x^P \in \mathbb{R}^{N^S}$, demixing filter length L.

Output : Source signals $s^1, \dots, s^Q \in \mathbb{R}^{N^S}$ and demixing filters $\underline{w}_{pq} \in \mathbb{R}^L$.

2 **begin**

1) Spatial prewhitening

Eigenvalue decomposition of the spatial correlation matrix

3 $EDE^{\mathsf{T}} \leftarrow \mathrm{E}[x_k^{1:P}(x_k^{1:P})^{\mathsf{T}}]$ with $x_k^{1:P} = \begin{pmatrix} x_k^1 & \cdots & x_k^P \end{pmatrix}^{\mathsf{T}}$

4 $\begin{pmatrix} \tilde{x}_k^1 & \cdots & \tilde{x}_k^P \end{pmatrix}^{\mathsf{T}} \leftarrow ED^{-1/2}E^{\mathsf{T}}x_k^{1:P}$ $\qquad$ Spatial sphering

2) Main algorithm:

Short-time Fourier transform using a Hanning window (N^B blocks)

5 $X_{m\nu}^P \Leftarrow \tilde{x}_k^P \quad \forall\, p = 1, \dots, P$ $\quad$ Frequency index $\nu = 0, \dots, L-1$, block m

6 $X_{m\nu}^P \leftarrow X_{m\nu}^P - \frac{1}{N}\sum_{m'=0}^{N-1} X_{m'\nu}^P$ $\qquad\qquad$ Center solution

Initialize demixing filter coefficients

7 $\underline{w}_{pq}^0 \leftarrow \begin{pmatrix} 1 & 0 & \cdots & 0 \end{pmatrix}^{\mathsf{T}} \in \mathbb{R}^L \qquad \forall\, p, q$

8 $W_{pq\nu}^0 \Leftarrow \underline{w}_{pqk}^0$ $\qquad\qquad$ Calculate FFT of filter coefficients

9 **for** $\ell = 1$ **to** ℓ_{max} **do**

Circular convolution to obtain (preliminary) source signal estimates

10 $S_{m\nu} \leftarrow W_\nu^{\ell-1}X_{m\nu} \qquad$ with $\qquad$
$X_{m\nu} = \begin{pmatrix} X_{m\nu}^1 & \cdots & X_{m\nu}^P \end{pmatrix}^{\mathsf{T}}$
$S_{m\nu} = \begin{pmatrix} S_{m\nu}^1 & \cdots & S_{m\nu}^Q \end{pmatrix}^{\mathsf{T}}$
$W_\nu^\ell = \begin{pmatrix} W_{11\nu}^\ell & \cdots & W_{P1\nu}^\ell \\ \vdots & \ddots & \vdots \\ W_{1Q\nu}^\ell & \cdots & W_{PQ\nu}^\ell \end{pmatrix}$

11 $b_m^p \leftarrow \sqrt{\frac{1}{M}\sum_{\nu=0}^{M-1}|S_{m\nu}^p|^2}$ $\qquad$ Broadband normalization factors

12 $\Phi_{m\nu} \leftarrow \begin{pmatrix} S_{m\nu}^1/b_m^1 & \cdots & S_{m\nu}^Q/b_m^Q \end{pmatrix}^{\mathsf{T}}$ $\qquad$ Multivariate score function

Perform natural gradient descent coefficient update (step size μ)

13 $W_\nu^\ell \leftarrow W_\nu^{\ell-1} + \mu\left[I - \frac{1}{N^B}\sum_{m=0}^{N^B-1}\Phi_{m\nu}S_{m\nu}^{\mathsf{H}}\right]W_\nu^{\ell-1}$

14 $W_\nu^\ell \leftarrow \mathrm{diag}\left(\left(W_\nu^\ell\right)^{-1}\right)W_\nu^\ell$ $\qquad$ Apply minimum distortion principle

15 **end**

16 $\underline{w}_{pqk}^{\ell_{max}} \Leftarrow W_{pq\nu}^{\ell_{max}}$ $\qquad$ Time domain filter coefficients from inverse FFT

17 $\begin{pmatrix} s_k^1 & \cdots & s_k^Q \end{pmatrix}^{\mathsf{T}} \Leftarrow W_\nu^\ell X_{m\nu}$ $\qquad$ Source signal estimates from inverse STFT

18 **end**

---

recovering this activity. This is due to a number of complicating factors. Firstly, and most trivially to solve, there is, of course, an inherent ambiguity regarding the order of the identified source signals. It is thus a priori unclear which of the identified source signals corresponds to, e.g., inspiratory or expiratory activity, and this must be identified after the separation has been performed. Secondly, as with any convolutive blind source separation procedure [9], each individual identified source signal is only specified up to filtering with an arbitrary FIR filter of the same length L as the demixing filters. Thus, exact point-wise errors between the recovered and the original source signals are not a correct measure of separation performance. Thirdly, and most crucially, in the given setting, the model does not describe the system under consideration correctly[16]: while the mixing model is capable of representing the spatial filtering properties of the biological tissues separating muscles and electrodes, the source model is inaccurate, as discussed in detail above: In reality, there are hundreds of sources (muscle fibers) continuously distributed throughout the upper body, each of which is associated with a different transmission path to the electrodes and, hence, different mixing filters. Our separation model, on the other hand, assumes just four point sources. Thus, one cannot expect the algorithm to correctly recover the exact activity of the different (inspiratory, expiratory, cardiac) muscle groups.

To solve these challenges, we devised the following validation procedure [77]. To obtain an *inspiratory reference (envelope) signal* $\overline{\text{EMG}}^{\text{insp}}$, we consider the raw EMG signal of just the diaphragm muscle to the closest electrode, take the absolute of that signal, and calculate a 101-sample moving average. The same envelope calculation is repeated for the contribution of just the two transversus abdominis muscles to their closest electrodes to obtain a corresponding *expiratory* $\overline{\text{EMG}}^{\text{exp}}$, and for each of the identified source signals s^q to obtain a corresponding envelope signal $\bar{s}^q$. In order to compare the identified source envelope signals with the reference signals while taking the aforementioned filtering ambiguity into account, we minimize the mean absolute error concerning the inspiratory reference signal ($\text{MAE}^{\text{bss,insp}}$) by solving

$$\alpha^*, \beta^*, \kappa^*, q^* = \underset{\alpha \in \mathbb{R}^+, \beta \in \mathbb{R}, \kappa \in \mathbb{Z}, q \in \{1,\dots,Q\}}{\arg\min} \mathrm{E}\left[\left|\overline{\text{EMG}}_k^{\text{insp}} - \alpha \cdot \max\left(0, \bar{s}_{k-\kappa}^q - \beta\right)\right|\right], \qquad (4.19)$$

and equivalently for $\text{MAE}^{\text{bss,exp}}$ using $\overline{\text{EMG}}^{\text{exp}}$. Finally, because we are interested in the *improvement* with respect to the original (simulated) measurement signals, we calculate the ratios

$$\rho_{\text{insp}} = \frac{\text{MAE}^{\text{meas,insp}}}{\text{MAE}^{\text{bss,insp}}} \qquad \text{and} \qquad \rho_{\text{exp}} = \frac{\text{MAE}^{\text{meas,exp}}}{\text{MAE}^{\text{bss,exp}}}$$

of the MAEs optimized according to eq. (4.19), where $\text{MAE}^{\text{meas,insp}}$ and $\text{MAE}^{\text{meas,exp}}$ are calculated using the original EMG measurement signals (instead of the separated source signals). If $\rho_{\text{insp}} > 1$ or $\rho_{\text{exp}} > 1$, an improvement has been achieved by the BSS procedure.

4.5 Results

To assess the effect of different noise levels and realizations on the performance of the BSS procedure, thirteen different noise scales c^ν and four different cardiac component

---

[16]This is, of course, true at least to some degree in any real measurement.
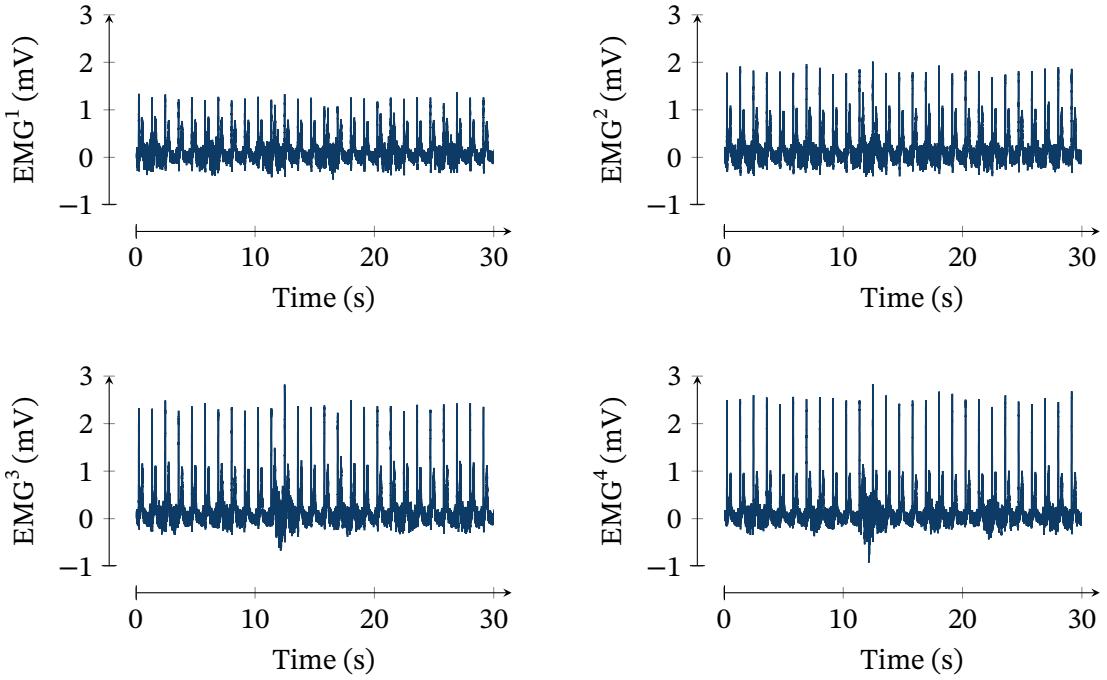
Figure 4.7 – An exemplary simulated multichannel respiratory surface EMG measurement. Each signal consists of three components: a surface EMG component simulated according to the model described in section 4.2.1, a cardiac component simulated according to the model described in section 4.2.2, and a Gaussian white noise component. The first two components are simulated using the parametrization and electrode placement described in section 4.2.3. For the shown example, we have $c^{ECG} = 1$ and a measurement noise SNR of 4 dB.

scales c^{ECG} were tested, and for each combination of scales, the average BSS performance over ten independent realizations was calculated. Thus, 520 test cases were simulated and analyzed in total. Figure 4.7 shows an exemplary simulated dataset, and fig. 4.8 the corresponding, separated signals. Figure 4.9 shows an analysis of the separation performance in the different noise settings, using the performance measures defined in section 4.4. Our BSS algorithm achieves an improvement with respect to the raw measurement signals ($\rho_{insp} > 1$ and $\rho_{exp} > 1$) in all settings and realizations. Especially at $c^{ECG} = 1$ (see fig. 4.7 for an example), very high degrees of improvement (median $\rho_{insp} > 4$, median $\rho_{exp} > 2$) can be observed. There is, however, a high degree of variability between noise realizations regarding the achieved separation success, as can be observed from the large min–max range in fig. 4.9. Moreover, the separation performance declines significantly with increasing cardiac component levels, although these are still within the physiologically observed range.
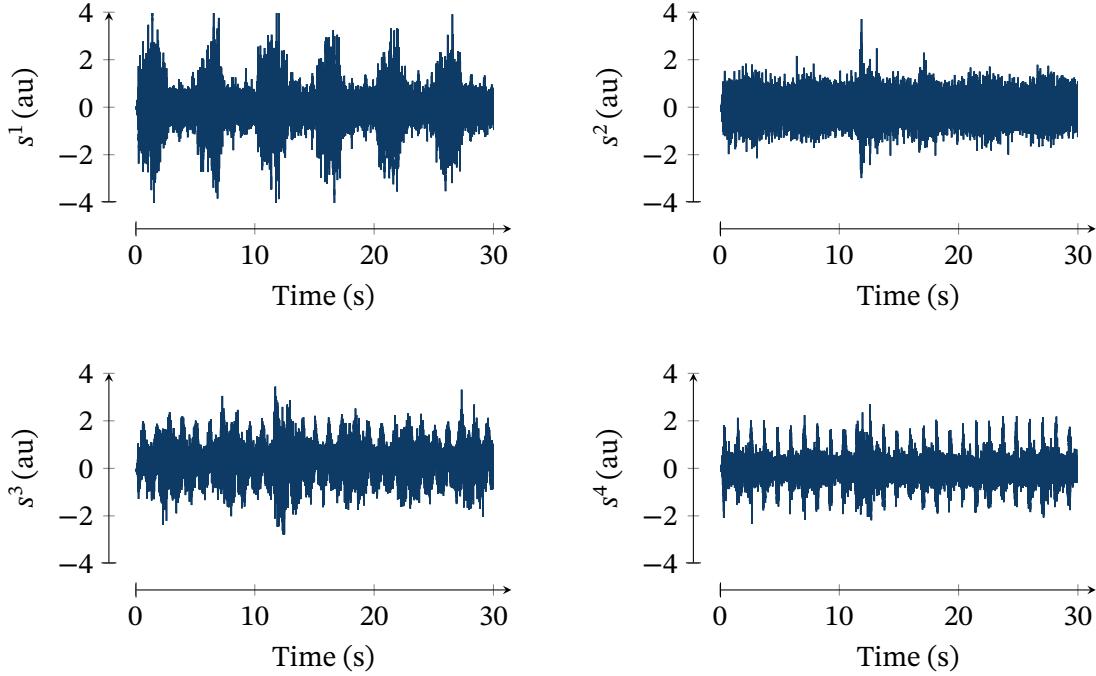
Figure 4.8 – Source signals identified by applying algorithm 5 to the simulated measurement signals shown in fig. 4.7. In this example, source signals s^1 and s^3 captured inspiratory and expiratory activity best, respectively. Separation performance is quantified for this example as $\text{MAE}^{\text{meas, insp}}/\text{MAE}^{\text{bss, insp}} = 4.42$ and $\text{MAE}^{\text{meas, exp}}/\text{MAE}^{\text{bss, exp}} = 1.33$.
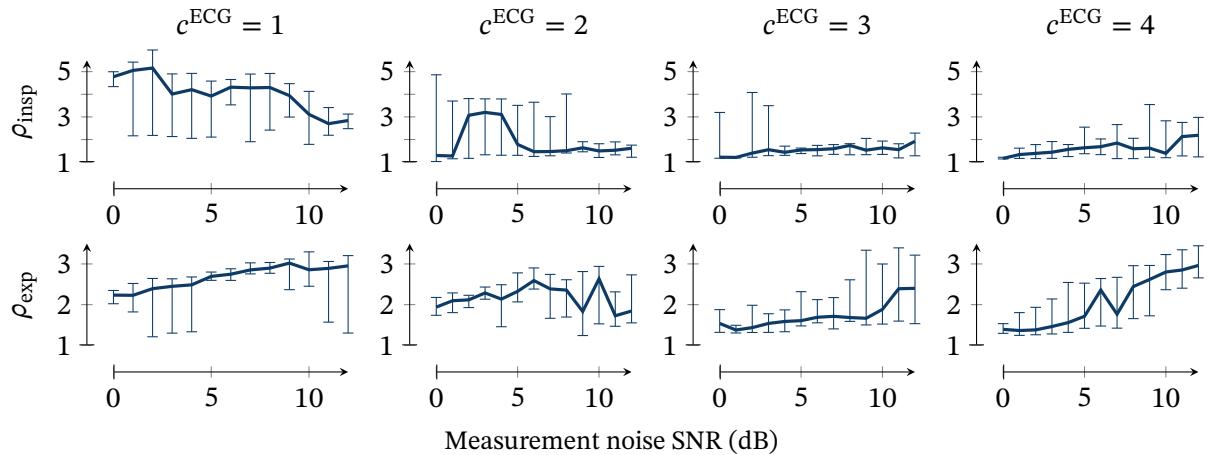
Figure 4.9 – The ratios $\rho_{\text{insp}} = \text{MAE}^{\text{meas, insp}}/\text{MAE}^{\text{bss, insp}}$ (top row) and $\rho_{\text{insp}} = \text{MAE}^{\text{meas, exp}}/\text{MAE}^{\text{bss, exp}}$ (bottom row) for different measurement noise SNRs and cardiac component scales c^{ECG}, averaged over ten noise realizations each. Lines represent median values across noise realizations, error bars represent the min–max range. See section 4.4 for more details on these performance measures.

4.6 Discussion & outlook

To conclude this chapter, we will now proceed to discuss the merits and drawbacks of the selected approaches for surface EMG modeling and simulation and for separating the activity of different muscles in such measurements. Moreover, possible avenues for future improvements will be discussed.

4.6.1 Realistic modeling of respiratory surface EMG measurements

The model proposed by Petersen and Rostalski [78] and described in section 4.2 incorporates much of current physiological knowledge. Previous modeling efforts were often focused on specific components of electrophysiology, whereas an integration of the many different strands of research into a single, comprehensive model was missing. Here, such an integration has been achieved while also proposing new models for many sub-components. Besides the integration of many different model components into a single, comprehensive model, important contributions include

- a novel, flexible model of MU rate coding and recruitment,

- the incorporation of the stochastic dependency of many different MU and fiber properties on MU and fiber size, i.e., a comprehensive and realistic model of the *size principle*,

- the modeling of these relationships as *stochastic* and *continuously distributed* (as opposed to deterministic and clustered into distinct classes),

- deliberate joint modeling of the electrical and mechanical properties of muscle fibers to ensure a realistic EMG-force relationship,

- an analysis of the nonlinear relationship between muscle excitation and the generated muscle force, and the ways in which this influences the modeled rate coding and recruitment behavior, and

- a derivation of an alternative formulation and several important properties of a previously proposed model of intracellular action potential propagation.

The result is a simulation software that enables the highly configurable generation of realistic surface EMG and force measurements. The simulated signals have already been used for the validation of signal processing algorithms and the derivation of new, reduced-order models of physiology [73, 77]. The simulation software is freely and openly accessible at `https://github.com/ime-luebeck/semgsim` [76].

Volume Conduction Model

Although the model already incorporates many relevant aspects of physiology, it could, of course, be further extended. The *volume conduction model* of Farina and Merletti [32] (which is used as part of the proposed model) is quite limited in its capability to represent realistic geometries. This could be replaced by a more complex geometrical model based on the finite element method (FEM), several of which have been proposed in the literature [8, 57, 58, 63, 92, 96, 97]. Regarding the distribution of MUs within the muscle cross-section, strategies accounting for regionalized MU placement have been proposed [85] and could be employed within the present modeling framework. Effects due to *muscle fatigue* are also currently not

modeled, a drawback that could be alleviated by incorporating the model proposed by Potvin and Fuglevand [82]. To accurately model *dynamical contractions*, a number of extensions to the current model would be necessary. In order to correctly model the geometry-dependent MUAP shapes, the volume conduction model must be equipped to represent a time-varying geometry. To this end, an efficient discretization-based approach such as the one proposed by Glaser *et al.* [39] could be pursued. Moreover, as the muscle lengthens and shortens, its force-producing capabilities change. To describe this, a simple Hill-type model of the force-length and force-velocity relationships of the muscle might be employed [101]. Depending on the target application of the model, it might also be desirable to integrate the model within a larger biomechanical model, such as the CEINMS model [79, 91]. As an alternative avenue for further integration, a combination with models of *physiological force control* might be pursued [27, 98]. In the present model, the neural muscle activation signal has to be defined by the user, but this could, of course, also be the result of another physiological model. A realistic model of physiological force control would necessarily introduce a feedback loop in the simulation model because the neural muscle activation depends on the past force generation. Finally, more realistic models of cardiac electrophysiology could, of course, be employed as well, examples of which abound in the literature [56, 69].

<div style="text-align: right;">Dynamical Contractions</div>

<div style="text-align: right;">Physiological Force Control</div>

In an orthogonal line of research, it appears promising to investigate approximation schemes for obtaining similarly realistic simulated signals at significantly *reduced computational effort*. This is not a trivial endeavor, however. The currently significant computation requirements are a consequence of the necessity to compute single-fiber action potentials (SFAPs) for tens of thousands of muscle fibers — each necessitating the approximate evaluation of a double integral, see Eqs. (12) and (23) in Farina and Merletti [32] — and state-dependent impulse train signals for hundreds of MUs. Owing to the complexity of the system under consideration — impulse trains governed by partially stochastic and partially deterministic processes, action potential generation, propagation along the fibers and extinction, as well as diffusion throughout the surrounding tissue layers — the ideal point of attack for an approximation scheme (that does not overly sacrifice model realism) is not obvious. It *is* obvious, however, that simple colored-noise models [95] do not precisely capture many of the biological system's properties. (One obvious limitation of such simple models is that they cannot describe the recurring MUAPs of close-by MUs.) *Mean-field modeling* [3, 99] might represent one possible avenue to address this challenge. These models describe (probabilistically) the stochastic signals resulting from an infinite number of interactive neurons that fire in a partially coupled and partially stochastic fashion. They are popular for the modeling and analysis of EEG signals [3, 99] but have, to the author's knowledge, not been applied to the modeling of EMG signals. Alternatively, *model order reduction* approaches could be investigated. It is well known [8] that a small number of MUs have a disproportionately large effect on the surface EMG signal measured at a particular electrode location, whereas most other MUs are effectively irrelevant or only cause background noise, which could likely be described reasonably well using a colored noise model. Thus, an attractive approach might be to approximate the signal measured at a particular electrode using only a strongly reduced number of MUs close to that electrode, superposed by some kind of colored background noise representing the activity of the remaining, further remote MUs. Many questions remain to be answered, however, such as the optimal number and location of MUs to consider (and computationally efficient ways to determine these — that is, without first

<div style="text-align: right;">Reduced Computational Effort</div>

simulating the complete set-up and then reducing the model afterward) as well as the appropriate type of background noise. Finally, and maybe the least invasively in terms of deviating from the proposed model: whereas in the current model, all single-fiber action potentials (SFAPs) are simulated individually and then superimposed to obtain a MUAP, this superposition could likely be approximated well using just a few SFAPs (and, possibly, some further smoothing operation). Seeing that there are tens of thousands of fibers in a simulated muscle but at most a few hundred MUs, this could already serve to greatly reduce the computational effort required to calculate realistic MUAPs.

4.6.2 Source separation for surface respiratory EMG measurements

The source separation problem for surface respiratory EMG measurements has high practical relevance. The increasingly attractive prospects of using respiratory sEMG measurements for clinical monitoring induce demand for methods to robustly clean measurements from crosstalk, cardiac interference, and various other disturbances. For developing such algorithms, proper means of validation are crucial. In this chapter (and in our previous publications on the subject [12, 75, 77, 78] on which this chapter is based), a novel simulation-based method for the numerical quantification of signal separation success in the respiratory setting has been proposed. Moreover, a convolutive BSS algorithm has been derived[17] from the general TRINICON framework, adapted to the specific separation problem at hand. The proposed algorithm is ambitiously designed to solve multiple hard problems at once — suppressing the effect of cardiac interference, removing crosstalk from other muscles, and separating between inspiratory and expiratory muscle activity, all while taking information from only a few available measurement channels into account. Its performance has been quantified using the described validation method.

While the proposed algorithm performs reasonably well in the simulated test cases, it appears likely that even more successful algorithms can be designed. Firstly, and possibly most importantly, notice that algorithm 5 does not exploit *any* application-specific information. The algorithm is completely general, exploiting only the basic statistical properties of non-Gaussianity, non-whiteness, and non-stationarity. It appears probable that incorporating further knowledge about the specific problem at hand might serve to increase algorithm performance. This might include information such as the expected quasi-periodicity of the cardiac component and the envelopes of the respiratory components, the frequency bands in which we expect the different components to be active, or phases during which we expect particular sources to be active or passive. The latter information might be derived from pneumatic measurements such as airflow or current lung volume. Utilizing such additional sources of information for improving separation success represents a shift from *fully* blind source separation to *semi*-blind source separation methods [5]. As an example falling into this class, Roussel *et al.* [87] have proposed a source separation method that exploits the cyclo-stationarity of intramuscular EMG signals. Recently, Buchner *et al.* [11] have presented an extended (Bayesian) version of the already very general TRINICON framework, allowing for the incorporation of both probabilistic and deterministic prior knowledge about the source signals and the mixing system. This extended framework also encompasses the tracking of time-varying systems. (Here, we assumed that the mixing system was static.) Using this extended framework for semi-blind

Semi-Blind Source Separation

---

[17]As was mentioned earlier, the derivation of algorithm 5 is due to H. Buchner, see Buchner *et al.* [12].

source separation to derive a more problem-specific algorithm represents an attractive avenue for future research.

Secondly, as was discussed in more detail in section 4.3, the employed *mixing model* is problematic: while mixing *is* in fact convolutive, there are in reality many more sources (motor units) than sensors (sEMG electrodes) in our setup.[18] This renders the separation problem strongly underdetermined. An obvious first option to solve this problem would be the use of an algorithm for the underdetermined BSS case. This was not further examined here; algorithm 5 is not directly applicable to the underdetermined case and would need adaptation first. Another option to resolve this challenge is to use a dedicated algorithm for removing cardiac interference from all measurement channels (this is the subject of the following chapter). Doing so significantly simplifies the subsequent task of recognizing and removing crosstalk and other disturbances in the cleaned EMG signal. Importantly, once the cardiac component has been removed by means of a preprocessing step, a convolutive source separation model is no longer strictly necessary. One may then use standard envelope-based separation techniques like non-negative matrix factorization (NMF) [55], a standard method for performing activity separation at the muscle group level (except that in other applications, the cardiac interference problem is much less pronounced), as was discussed in section 4.1.2. The following chapter is wholly devoted to the topic of cardiac interference removal.

Mixing Model

Finally, as a completely different alternative, deep learning-based methods such as those recently described by Wen *et al.* [100] could be exploited for the separation task. As these methods are rather new, their prospects for the present application are currently unclear. One significant challenge will likely concern the construction of an informative and sufficiently large and diverse training dataset. To this end, synthetic data such as those generated by the herein proposed simulation model may play a crucial role [70].

References

[1] A. B. Ajiboye and R. F. Weir, "Muscle synergies as a predictive framework for the EMG patterns of new hand postures," *Journal of Neural Engineering*, vol. 6, no. 3, p. 036 004, 2009. DOI: 10.1088/1741-2560/6/3/036004.

[2] M. Al Harrach *et al.*, "Analysis of the sEMG/force relationship using HD-sEMG technique and data fusion: A simulation study," *Computers in Biology and Medicine*, vol. 83, pp. 34–47, 2017. DOI: 10.1016/j.compbiomed.2017.02.003.

[3] S. van Albada and P. Robinson, "Mean-field modeling of the basal ganglia-thalamocortical system. I," *Journal of Theoretical Biology*, vol. 257, no. 4, pp. 642–663, 2009. DOI: 10.1016/j.jtbi.2008.12.018.

[4] S. Amari, S. Douglas, A. Cichocki, and H. Yang, "Multichannel blind deconvolution and equalization using the natural gradient," in *First IEEE Signal Processing Workshop on Signal Processing Advances in Wireless Communications*, IEEE, 1997. DOI: 10.1109/spawc.1997.630083.

[5] M. Babaie-Zadeh and C. Jutten, "Semi-blind approaches for source separation and independent component analysis.," in *European Symposium on Artificial Neural Networks (ESANN)*, 2006, pp. 301–312.

---

[18]This is different in the HD-sEMG case, where the number of sensors is much larger, and researchers have successfully employed convolutive BSS methods similar to ours [46].

[6]   B. K. Barry, M. A. Pascoe, M. Jesunathadas, and R. M. Enoka, "Rate coding is compressed but variability is unaltered for motor units in a hand muscle of old adults," *Journal of Neurophysiology*, vol. 97, no. 5, pp. 3206–3218, 2007. DOI: 10.1152/jn.01280.2006.

[7]   D. J. Berger and A. d'Avella, "Effective force control by muscle synergies," *Frontiers in Computational Neuroscience*, vol. 8, 2014. DOI: 10.3389/fncom.2014.00046.

[8]   D. P. Botelho, K. Curran, and M. M. Lowery, "Anatomically accurate model of EMG during index finger flexion and abduction derived from diffusion tensor imaging," *PLOS Computational Biology*, vol. 15, no. 8, J. A. Papin, Ed., e1007267, 2019. DOI: 10.1371/journal.pcbi.1007267.

[9]   H. Buchner, R. Aichner, and W. Kellermann, "Blind source separation for convolutive mixtures: A unified treatment," in *Audio Signal Processing for Next-Generation Multimedia Communication Systems*, Y. Huang and J. Benesty, Eds., Kluwer Academic Publishers, 2004. DOI: 10.1007/1-4020-7769-6_10.

[10]   ——, "TRINICON-based blind system identification with application to multiple-source localization and separation," in *Blind Speech Separation*, S. Makino, T.-W. Lee, and S. Sawada, Eds., Springer, 2007. DOI: 10.1007/978-1-4020-6479-1_4.

[11]   H. Buchner, K. Helwani, and S. Godsill, "Unsupervised Bayesian estimation and tracking of time-varying convolutive multichannel systems," in *2019 22nd International Conference on Information Fusion (FUSION)*, 2019, pp. 1–8. [Online]. Available: https://ieeexplore.ieee.org/document/9011190.

[12]   H. Buchner, E. Petersen, M. Eger, and P. Rostalski, "Convolutive blind source separation on surface EMG signals for respiratory diagnostics and medical ventilation control," in *38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Florida, USA, 2016. DOI: 10.1109/EMBC.2016.7591513.

[13]   V. Carriou, S. Boudaoud, J. Laforet, and F. S. Ayachi, "Fast generation model of high density surface EMG signals in a cylindrical conductor volume," *Computers in Biology and Medicine*, vol. 74, pp. 54–68, 2016. DOI: 10.1016/j.compbiomed.2016.04.019.

[14]   A. Cichocki, R. Zdunek, and S. Amari, "New algorithms for non-negative matrix factorization in applications to blind source separation," in *IEEE International Conference on Acoustics Speed and Signal Processing Proceedings*, IEEE, 2006. DOI: 10.1109/icassp.2006.1661352.

[15]   E. A. Clancy and N. Hogan, "Probability density of the surface electromyogram and its relation to amplitude detectors," *IEEE Transactions on Biomedical Engineering*, vol. 46, no. 6, pp. 730–739, 1999. DOI: 10.1109/10.764949.

[16]   P. Comon and C. Jutten, *Handbook of Blind Source Separation: Independent component analysis and applications*. Academic press, 2010.

[17]   P. Comon, "Independent component analysis," in *Higher-Order Statistics*, J.-L. Lacoume, Ed., Elsevier, 1992, pp. 29–38. [Online]. Available: https://hal.archives-ouvertes.fr/hal-00346684/.

[18]   ——, "Independent component analysis, a new concept?" *Signal Processing*, vol. 36, no. 3, pp. 287–314, 1994. DOI: 10.1016/0165-1684(94)90029-9.

[19]   P. Contessa and C. J. De Luca, "Neural control of muscle force: Indications from a simulation model," *Journal of Neurophysiology*, vol. 109, pp. 1548–1570, 2013. DOI: 10.1152/jn.00237.2012.

[20]   C. J. De Luca and P. Contessa, "Hierarchical control of motor units in voluntary contractions," *Journal of Neurophysiology*, vol. 107, no. 1, pp. 178–195, 2012. DOI: 10.1152/jn.00961.2010.

[21] C. J. De Luca and E. C. Hostage, "Relationship between firing rate and recruitment threshold of motoneurons in voluntary isometric contractions," *Journal of Neurophysiology*, vol. 104, no. 2, pp. 1034–1046, 2010, Corrigendum. DOI: 10.1152/jn.01018.2009.

[22] C. J. De Luca and S. H. Nawab, "Reply to Farina and Enoka: The reconstruct-and-test approach is the most appropriate validation for surface EMG signal decomposition to date," *Journal of Neurophysiology*, vol. 105, no. 2, pp. 983–984, 2011. DOI: 10.1152/jn.01060.2010.

[23] C. J. De Luca, S. H. Nawab, and J. C. Kline, "Clarification of methods used to validate surface EMG decomposition algorithms as described by Farina et al. (2014)," *Journal of Applied Physiology*, vol. 118, no. 8, pp. 1084–1084, 2015. DOI: 10.1152/japplphysiol.00061.2015.

[24] C. J. De Luca *et al.*, "Decomposition of surface EMG signals," *Journal of Neurophysiology*, vol. 96, no. 3, pp. 1646–1657, 2006. DOI: 10.1152/jn.00009.2006.

[25] J. L. Dideriksen, F. Negro, R. M. Enoka, and D. Farina, "Motor unit recruitment strategies and muscle properties determine the influence of synaptic noise on force steadiness," *Journal of Neurophysiology*, vol. 107, pp. 3357–3369, 2012. DOI: 10.1152/jn.00938.2011.

[26] J. L. Dideriksen, D. Farina, and R. M. Enoka, "Influence of fatigue on the simulated relation between the amplitude of the surface electromyogram and muscle force," *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 2010. DOI: 10.1098/rsta.2010.0094.

[27] J. L. Dideriksen, D. F. Feeney, A. M. Almuklass, and R. M. Enoka, "Control of force during rapid visuomotor force-matching tasks can be described by discrete time PID control algorithms," *Experimental Brain Research*, vol. 235, no. 8, pp. 2561–2573, 2017. DOI: 10.1007/s00221-017-4995-3.

[28] G. C. Elder, K. Bradbury, and R. Roberts, "Variability of fiber type distributions within human muscles," *Journal of Applied Physiology*, vol. 53, no. 6, pp. 1473–1480, 1982. DOI: 10.1152/jappl.1982.53.6.1473.

[29] R. M. Enoka and A. J. Fuglevand, "Motor unit physiology: Some unresolved issues," *Muscle and Nerve*, vol. 42, pp. 4–17, 2001. DOI: 10.1002/1097-4598(200101)24:1<4::AID-MUS13>3.0.CO;2-F.

[30] Z. Erim, C. J. De Luca, K. Mineo, and T. Aoki, "Rank-ordered regulation of motor units," *Muscle & Nerve*, vol. 19, no. 5, pp. 563–573, 1996. DOI: https://doi.org/10.1002/(SICI)1097-4598(199605)19:5<563::AID-MUS3>3.0.CO;2-9.

[31] D. Farina and R. M. Enoka, "Surface EMG decomposition requires an appropriate validation," *Journal of Neurophysiology*, vol. 105, no. 2, pp. 983–4, 2011. DOI: 10.1152/jn.00855.2010.

[32] D. Farina and R. Merletti, "A novel approach for precise simulation of the EMG signal detected by surface electrodes," *IEEE Transactions on Biomedical Engineering*, vol. 48, pp. 637–646, 6 2001. DOI: 10.1109/10.923782.

[33] D. Farina, R. Merletti, and R. M. Enoka, "Reply to De Luca, Nawab, and Kline: The proposed method to validate surface EMG signal decomposition remains problematic," *Journal of Applied Physiology: Respiratory, Environmental and Exercise Physiology*, vol. 118, p. 1085, 8 2015. DOI: 10.1152/japplphysiol.00107.2015.

[34] D. Farina, C. Cescon, and R. Merletti, "Influence of anatomical, physical, and detection-system parameters on surface EMG," *Biological Cybernetics*, vol. 86, no. 6, pp. 445–456, 2002. DOI: 10.1007/s00422-002-0309-2.

[35] D. Farina, L. Mesin, S. Martina, and R. Merletti, "A surface EMG generation model with multilayer cylindrical description of the volume conductor," *IEEE Transactions on Biomedical Engineering*, vol. 51, pp. 415–426, 3 2004. DOI: 10.1109/TBME.2003.820998.

[36] D. Farina and A. Rainoldi, "Compensation of the effect of sub-cutaneous tissue layers on surface EMG: A simulation study," *Medical Engineering & Physics*, vol. 21, pp. 487–497, 6–7 1999. DOI: 10.1016/s1350-4533(99)00075-2.

[37] D. Farina *et al.*, "The extraction of neural information from the surface EMG for the control of upper-limb prostheses: Emerging avenues and challenges," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 22, no. 4, pp. 797–809, 2014. DOI: 10.1109/tnsre.2014.2305111.

[38] A. J. Fuglevand, D. A. Winter, and A. E. Patla, "Models of recruitment and rate coding organization in motor-unit pools," *Journal of Neurophysiology*, vol. 70, pp. 2470–88, 6 1993. DOI: 10.1152/jn.1993.70.6.2470.

[39] V. Glaser, D. Farina, and A. Holobar, "Simulations of high-density surface electromyograms in dynamic muscle contractions," in *2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, IEEE, 2017. DOI: 10.1109/embc.2017.8037599.

[40] C. J. Heckman and R. M. Enoka, "Motor unit," in *Comprehensive Physiology*. American Cancer Society, 2012, pp. 2629–2682, ISBN: 9780470650714. DOI: 10.1002/cphy.c100087.

[41] E. Henneman, "Relation between size of neurons and their susceptibility to discharge," *Science*, vol. 126, pp. 1345–1347, 3287 1957. DOI: 10.1126/science.126.3287.1345.

[42] E. Henneman, G. Somjen, and D. O. Carpenter, "Excitability and inhibitibility of motoneurons of different sizes," *Journal of Neurophysiology*, vol. 28, no. 3, pp. 599–620, 1965. DOI: 10.1152/jn.1965.28.3.599.

[43] A. L. Hodgkin and A. F. Huxley, "A quantitative description of membrane current and its application to conduction and excitation in nerve," *The Journal of Physiology*, vol. 117, no. 4, pp. 500–544, 1952. DOI: 10.1113/jphysiol.1952.sp004764.

[44] A. Holobar, M. A. Minetto, and D. Farina, "Accurate identification of motor unit discharge patterns from high-density surface EMG and validation with a novel signal-based performance metric," *Journal of Neural Engineering*, vol. 11, no. 1, p. 016 008, 2014. DOI: 10.1088/1741-2560/11/1/016008.

[45] A. Holobar and D. Farina, "Blind source identification from the multichannel surface electromyogram," *Physiological Measurement*, vol. 35, no. 7, R143–R165, 2014. DOI: 10.1088/0967-3334/35/7/R143.

[46] A. Holobar and D. Zazula, "Multichannel blind source separation using convolution kernel compensation," *IEEE Transactions on Signal Processing*, vol. 55, no. 9, pp. 4487–4496, 2007. DOI: 10.1109/tsp.2007.896108.

[47] A. Holobar and D. Zazula, "Surface EMG decomposition using a novel approach for blind source separation," in *Informatica Medica Slovenica*, vol. 8, 2003, pp. 2–14.

[48] A. Holobar *et al.*, "Estimating motor unit discharge patterns from high-density surface electromyogram," *Clinical Neurophysiology*, vol. 120, no. 3, pp. 551–562, 2009. DOI: 10.1016/j.clinph.2008.10.160.

[49] F. Hug, N. A. Turpin, A. Guével, and S. Dorel, "Is interindividual variability of EMG patterns in trained cyclists related to different muscle synergies?" *Journal of Applied Physiology*, vol. 108, no. 6, pp. 1727–1736, 2010. DOI: 10.1152/japplphysiol.01305.2009.

[50] A. Hyvärinen, *Survey on independent component analysis*, 1999. [Online]. Available: https://www.cs.helsinki.fi/u/ahyvarin/papers/NCS99.pdf.

[51] A. Hyvärinen, J. Karhunen, and E. Oja, *Independent Component Analysis*. John Wiley & Sons, 2004, 504 pp., ISBN: 9780471464198.

[52] M. A. Johnson, J. Polgar, D. Weightman, and D. Appleton, "Data on the distribution of fibre types in thirty-six human muscles – an autopsy study," *Journal of the Neurological Sciences*, vol. 18, pp. 111–129, 1 1973. DOI: 10.1016/0022-510x(73)90023-3.

[53] P. H. Kleinpenning, T. H. J. M. Gootzen, A. Van Oosterom, and D. F. Stegeman, "The equivalent source description representing the extinction of an action potential at a muscle fiber ending," *Mathematical Biosciences*, vol. 101, pp. 41–61, 1 1990. DOI: 10.1016/0025-5564(90)90101-4.

[54] A. Konstantin *et al.*, "Simulation of motor unit action potential recordings from intramuscular multichannel scanning electrodes," *IEEE Transactions on Biomedical Engineering*, vol. 67, no. 7, pp. 2005–2014, 2020. DOI: 10.1109/tbme.2019.2953680.

[55] D. D. Lee and H. S. Seung, "Algorithms for non-negative matrix factorization," in *Proceedings of the 13th International Conference on Neural Information Processing Systems*, Denver, CO: MIT Press, 2000, pp. 535–541. [Online]. Available: https://proceedings.neurips.cc/paper/2000/file/f9d1152547c0bde01830b7e8bd60024c-Paper.pdf.

[56] A. Loewe *et al.*, "A heterogeneous formulation of the Himeno et al. human ventricular myocyte model for simulation of body surface ECGs," in *2018 Computing in Cardiology Conference (CinC)*, Computing in Cardiology, 2018. DOI: 10.22489/cinc.2018.068.

[57] M. Lowery, N. Stoykov, J. Dewald, and T. Kuiken, "Volume conduction in an anatomically based surface EMG model," *IEEE Transactions on Biomedical Engineering*, vol. 51, no. 12, pp. 2138–2147, 2004. DOI: 10.1109/tbme.2004.836494.

[58] M. Lowery, N. Stoykov, A. Taflove, and T. Kuiken, "A multiple-layer finite-element model of the surface EMG signal," *IEEE Transactions on Biomedical Engineering*, vol. 49, no. 5, pp. 446–454, 2002. DOI: 10.1109/10.995683.

[59] K. Matsuoka, Y. Ohba, Y. Toyota, and S. Nakashima, "Blind separation for convolutive mixture of many voices," in *Proc. IWAENC 2003*, 2003, pp. 279–282.

[60] K. C. McGill, "Surface electromyogram signal modelling," *Medical and Biological Engineering and Computing*, vol. 42, pp. 446–454, 2004. DOI: 10.1007/BF02350985.

[61] P. E. McSharry, G. D. Clifford, L. Tarassenko, and L. A. Smith, "A dynamical model for generating synthetic electrocardiogram signals," *IEEE Transactions on Biomedical Engineering*, 2003. DOI: 10.1109/TBME.2003.808805.

[62] R. Merletti, L. Lo Conte, E. Avignone, and P. Guglielminotti, "Modeling of surface myoelectric signals. I. Model implementation," *IEEE Transactions on Biomedical Engineering*, vol. 46, no. 7, pp. 810–820, 1999. DOI: 10.1109/10.771190.

[63] L. Mesin *et al.*, "A finite element model for describing the effect of muscle shortening on surface EMG," *IEEE Transactions on Biomedical Engineering*, vol. 53, no. 4, pp. 593–600, 2006. DOI: 10.1109/tbme.2006.870256.

[64] M. Mordhorst, T. Heidlauf, and O. Röhrle, "Predicting electromyographic signals under realistic conditions using a multiscale chemo-electro-mechanical finite element model," *Interface Focus*, vol. 5, 2015. DOI: 10.1098/rsfs.2014.0076.

[65] C. T. Moritz, B. K. Barry, M. A. Pascoe, and R. M. Enoka, "Discharge rate variability influences the variation in force fluctuations across the working range of a hand muscle," *Journal of Neurophysiology*, vol. 93, no. 5, pp. 2449–2459, 2005. DOI: 10.1152/jn.01122.2004.

[66] J. Navallas *et al.*, "A muscle architecture model offering control over motor unit fiber density distributions," *Medical and Biological Engineering and Computing*, vol. 48, pp. 875–886, 9 2010. DOI: 10.1007/s11517-010-0642-x.

[67] S. H. Nawab, S. Chang, and C. J. De Luca, "High-yield decomposition of surface EMG signals," *Clinical Neurophysiology*, vol. 121, pp. 1602–1615, 10 2010. DOI: 10.1016/j.clinph.2009.11.092.

[68] F. Negro *et al.*, "Multi-channel intramuscular and surface EMG decomposition by convolutive blind source separation," *Journal of Neural Engineering*, vol. 13, p. 026 027, 2 2016. DOI: 10.1088/1741-2560/13/2/026027.

[69] A. Neic *et al.*, "Efficient computation of electrograms and ECGs in human whole heart simulations using a reaction-eikonal model," *Journal of Computational Physics*, vol. 346, pp. 191–211, 2017. DOI: 10.1016/j.jcp.2017.06.020.

[70] S. I. Nikolenko, *Synthetic Data for Deep Learning.* Springer International Publishing, 2021. DOI: 10.1007/978-3-030-75178-4.

[71] H. Nishizono, H. Kurata, and M. Miyashita, "Muscle fiber conduction velocity related to stimulation rate," *Electroencephalography and Clinical Neurophysiology*, vol. 72, no. 6, pp. 529–534, 1989. DOI: 10.1016/0013-4694(89)90230-7.

[72] C. Nordander *et al.*, "Influence of the subcutaneous fat layer, as measured by ultrasound, skinfold calipers and BMI, on the EMG amplitude," *European Journal of Applied Physiology*, vol. 89, no. 6, pp. 514–519, 2003. DOI: 10.1007/s00421-003-0819-1.

[73] M. Olbrich, E. Petersen, C. Hoffmann, and P. Rostalski, "Sparse estimation for the assessment of muscular activity based on sEMG measurements," *IFAC-PapersOnLine*, vol. 51, no. 15, pp. 305–310, 2018. DOI: 10.1016/j.ifacol.2018.09.152.

[74] E. Petersen, "Integrative mathematical modeling of surface electromyographic modeling," M.S. thesis, University of Hamburg, 2015.

[75] ——, "A mathematical model of surface electromyographic measurements," in *Proceedings of the Workshop Biosignal Processing*, Berlin, 2016.

[76] ——, *sEMGsim*, 2021. [Online]. Available: https://github.com/ime-luebeck/semgsim.

[77] E. Petersen, H. Buchner, M. Eger, and P. Rostalski, "Convolutive blind source separation of surface EMG measurements of the respiratory muscles," *Biomedical Engineering - Biomedizinische Technik*, vol. 62, no. 2, pp. 171–181, 2017. DOI: 10.1515/bmt-2016-0092.

[78] E. Petersen and P. Rostalski, "A comprehensive mathematical model of motor unit pool organization, surface electromyography, and force generation," *Frontiers in Physiology*, vol. 10, p. 176, 2019. ISSN: 1664-042X. DOI: 10.3389/fphys.2019.00176.

[79] C. Pizzolato *et al.*, "CEINMS: A toolbox to investigate the influence of different neural control solutions on the prediction of muscle excitation and joint moments during dynamic motor tasks," *Journal of Biomechanics*, vol. 48, no. 14, pp. 3929–3936, 2015. DOI: 10.1016/j.jbiomech.2015.09.021.

[80] R. Plonsey, "Action potential sources and their volume conductor fields," *Proceedings of the IEEE*, vol. 65, no. 5, pp. 601–611, 1977. DOI: 10.1109/proc.1977.10539.

[81] R. Plonsey and R. C. Barr, *Bioelectricity, A Quantitative Approach*, third edition. Springer, 2007.

[82] J. R. Potvin and A. J. Fuglevand, "A motor unit-based model of muscle fatigue," *PLOS Computational Biology*, vol. 13, no. 6, A. M. Haith, Ed., e1005581, 2017. DOI: 10.1371/journal.pcbi.1005581.

[83] R. T. Raikova, P. Krutki, H. Aladjov, and J. Celichowski, "Variability of the twitch parameters of the rat medial gastrocnemius motor units - experimental and modeling study," *Computers in Biology and Medicine*, vol. 37, no. 11, pp. 1572–1581, 2007. DOI: 10.1016/j.compbiomed.2007.02.006.

[84] J. W. Robertson, "Modifying muscle properties in a leading neuromuscular model: The Fuglevand model revisited," Ph.D. dissertation, University of Calgary, 2014.

[85]   J. W. Robertson and J. A. Johnston, "Modifying motor unit territory placement in the Fuglevand model," *Medical and Biological Engineering and Computing*, vol. 55, pp. 2015–2025, 2017. DOI: 10.1007/s11517-017-1645-7.

[86]   P. Rosenfalck, "Intra and extracellular fields of active nerve and muscle fibers. A physico-mathematical analysis of different models," *Acta Physiologica Scandinavica*, vol. 321, pp. 1–49, 1969.

[87]   J. Roussel *et al.*, "Decomposition of multi-channel intramuscular EMG signals by cyclostationary-based blind source separation," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 25, no. 11, pp. 2035–2045, 2017. DOI: 10.1109/tnsre.2017.2700890.

[88]   A. de Rugy, G. E. Loeb, and T. J. Carroll, "Are muscle synergies useful for neural control?" *Frontiers in Computational Neuroscience*, vol. 7, 2013. DOI: 10.3389/fncom.2013.00019.

[89]   R. Sameni, *OSET: The open-source electrophysiological toolbox*, version 3.14, Released under the GNU General Public License, 2006. [Online]. Available: www.oset.ir.

[90]   R. Sameni, G. Clifford, C. Jutten, and M. Shamsollahi, "Multichannel ECG and noise modeling: Application to maternal and fetal ECG signals," *EURASIP Journal on Advances in Signal Processing*, p. 043 407, 2007. DOI: 10.1155/2007/43407.

[91]   M. Sartori, D. G. Llyod, and D. Farina, "Neural data-driven musculoskeletal modeling for personalized neurorehabilitation technologies," *IEEE Transactions on Biomedical Engineering*, vol. 63, no. 5, pp. 879–893, 2016. DOI: 10.1109/tbme.2016.2538296.

[92]   J. Schneider, J. Silny, and G. Rau, "Influence of tissue inhomogeneities on noninvasive muscle fiber conduction velocity measurements-investigated by physical and numerical modeling," *IEEE Transactions on Biomedical Engineering*, vol. 38, no. 9, pp. 851–860, 1991. DOI: 10.1109/10.83605.

[93]   M.-A. Schnetzer, D. G. Rüegg, R. Baltensperger, and J.-P. Gabriel, "Three-dimensional model of a muscle and simulation of its surface EMG," in *23rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Istanbul, Turkey, 2001. DOI: 10.1109/IEMBS.2001.1020366.

[94]   Z.-H. Shi *et al.*, "Expiratory muscle dysfunction in critically ill patients: Towards improved understanding," *Intensive Care Medicine*, vol. 45, no. 8, pp. 1061–1071, 2019. DOI: 10.1007/s00134-019-05664-4.

[95]   D. F. Stegeman, J. H. Blok, H. J. Hermens, and K. Roeleveld, "Surface EMG models: Properties and applications," *Journal of Electromyography and Kinesiology*, vol. 10, pp. 313–326, 5 2000. DOI: 10.1016/s1050-6411(00)00023-7.

[96]   N. Stoykov, M. Lowery, A. Taflove, and T. Kuiken, "Frequency- and time-domain FEM models of EMG: Capacitive effects and aspects of dispersion," *IEEE Transactions on Biomedical Engineering*, vol. 49, no. 8, pp. 763–772, 2002. DOI: 10.1109/tbme.2002.800754.

[97]   A. Teklemariam *et al.*, "A finite element model approach to determine the influence of electrode design and muscle architecture on myoelectric signal properties," *PLOS ONE*, vol. 11, no. 2, D. Zhang, Ed., e0148275, 2016. DOI: 10.1371/journal.pone.0148275.

[98]   E. Todorov and M. I. Jordan, "Optimal feedback control as a theory of motor coordination," *Nature Neuroscience*, vol. 5, no. 11, pp. 1226–1235, 2002. DOI: 10.1038/nn963.

[99]   J. D. Victor, J. D. Drover, M. M. Conte, and N. D. Schiff, "Mean-field modeling of thalamocortical dynamics and a model-driven approach to EEG analysis," *Proceedings of the National Academy of Sciences*, vol. 108, pp. 15 631–15 638, 2011. DOI: 10.1073/pnas.1012168108.

[100] Y. Wen *et al.*, "A convolutional neural network to identify motor units from high-density surface electromyography signals in real time," *Journal of Neural Engineering*, vol. 18, no. 5, p. 056 003, 2021. DOI: 10.1088/1741-2552/abeead.

[101] G. T. Yamaguchi, *Dynamic Modeling of Musculoskeletal Motion*, 1st edition. Springer Science+Business Media Dordrecht, 2001.

CHAPTER 5

SEPARATING CARDIAC AND RESPIRATORY MUSCLE ACTIVITY
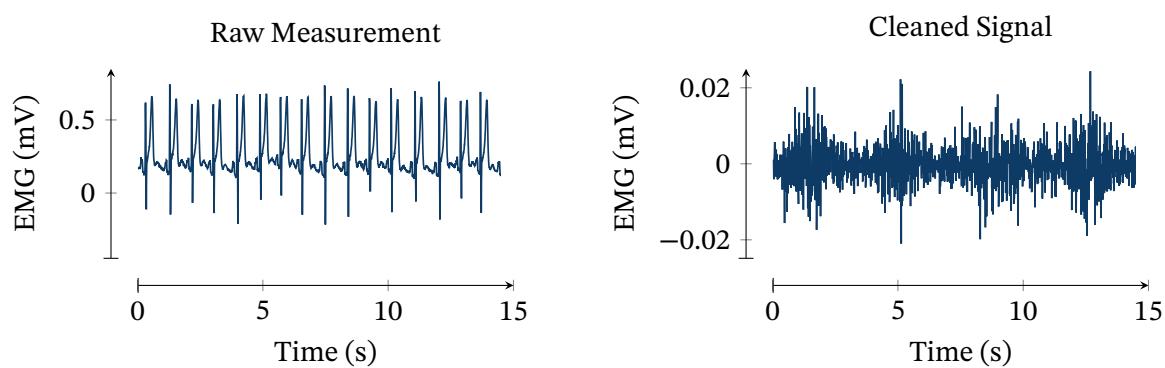
Figure 5.1 – A short excerpt from a respiratory EMG measurement, raw except for powerline interference removal (left) and cleaned from cardiac artifacts using the newly proposed probabilistic adaptive template subtraction (PATS) algorithm (right). The four bursts of activity in the right diagram represent respiratory muscle activity. Due to the large amplitude of the cardiac artifacts, respiratory activity is imperceptible in the raw signal. (Note the different scales of the two diagrams.)

Contents

CARDIAC ARTIFACTS represent a significant source of disturbances in respiratory surface EMG measurements. Due to the strength of the electrical field generated by the cardiac muscle and its close vicinity to the recording electrode positions, the measured signal usually contains cardiac artifacts that are orders of magnitude larger than the respiratory signal component. (Figure 5.1 shows an example.) The exceedingly small signal-to-noise ratio (SNR) distinguishes this separation task from others in, e.g., limb EMG measurements or EEG measurements, where the cardiac component is significantly smaller due to the increased distance from the heart [43]. This low SNR further exacerbates the separation task's difficulty, as even a small relative error in identifying the cardiac artifact results in prohibitively large relative errors in the cleaned respiratory EMG signal. Whereas the previous chapter examined a novel algorithm for identifying multiple physiological source signals from multiple sEMG mixture measurements, the present chapter will treat a single-channel separation problem: the separation of respiratory muscle activity and cardiac activity in single-channel surface electromyographic measurements of the respiratory muscles. Compared to the blind source separation (BSS) problem considered in the previous chapter, the present separation problem is *underdetermined*, since there are more sources to be identified (two) than measurements are available (one) — therefore, the separation procedure can only be successful if it exploits some amount of prior knowledge about the differing properties of cardiac and respiratory muscle activity. Furthermore, algorithm development is impeded by the lack of a reference signal that can be used for validation, just like in the previous chapter: there is no alternative way to measure the target signal *without* cardiac interference, and hence, the *true* result is always unknown.

The chapter begins with a definition of the problem under consideration and a comprehensive overview of the current state of the art (section 5.1), concerning both algorithms for removing cardiac artifacts from single-channel respiratory surface EMG measurement, as well as means for quantifying the quality of the resulting, cleaned signals despite the lack of a reference signal. Section 5.2 then proceeds to introduce a new algorithm based on model-based signal processing techniques, incorporating available domain knowledge while simultaneously retaining enough flexibility to treat real-world signals robustly. The following section 5.3 describes a comprehensive framework for validating cardiac artifact removal algorithms, and section 5.4, using this framework, provides an exhaustive comparative evaluation of many such algorithms (including the newly proposed one). Finally, a brief discussion and outlook (section 5.5) conclude the chapter.[1]

5.1 Problem definition & state of the art

This section commences with a definition of the exact separation problem under consideration (section 5.1.1). We will then proceed to review the current start of the art regarding algorithms for separating cardiac and respiratory muscle activity (section 5.1.2)

---

[1]An earlier version of the algorithm proposed in section 5.2 has been the subject of the Master's thesis of Seemann [67], which was conducted under the supervision of the author. The validation framework discussed in section 5.3 has been the subject of previous journal and conference publications (co-)authored by the author [25, 26, 52], as well as the Bachelor's thesis of Sauer [64], also conducted under the supervision of the author. A patent application concerning an artifact removal method related to the PATS algorithm has been filed by a group of inventors including the author [31].

and methods for quantifying the performance of such algorithms (section 5.1.3).

5.1.1 Problem definition

The specific problem of interest in this chapter is the removal of a cardiac signal component $EMG^{cardiac}$ from a single-channel surface EMG measurement EMG^{meas} of the respiratory muscles while minimizing changes to the characteristics of the respiratory signal component. Examples of possible input and output signals of such a removal algorithm are shown in fig. 5.1. Concerning typical recording sites of interest in this chapter, see fig. 3.13, which shows the electrode positions used in our validation study [52], discussed at more lengths in section 5.3.1.

5.1.2 Algorithms for separating cardiac and respiratory muscle activity

High-Pass
Filtering

Adaptive Noise
Cancellation

Template
Subtraction

The simplest proposed method for separating cardiac and respiratory muscle activity is the use of a simple *high-pass filter* [56]. This method has been used in a large number of applications, and it does serve to reduce the impact of cardiac artifacts significantly. In respiratory EMG measurements, however, there is significant frequency overlap between cardiac and respiratory signal components [7, 52], and hence, a simple high-pass filter on its own cannot completely solve the problem without also significantly distorting the EMG signal [7, 52].[2] Taking one additional piece of information into account to improve separability, several variations of *adaptive noise cancellation* (ANC) algorithms have been proposed [7, 18, 35, 41, 65, 66, 72, 73, 75], which employ a heartbeat impulse signal as a noise reference input. The problem with this class of algorithms is that they cannot cope with varying heartbeat intervals or varying beat morphology; consequently, they have not performed favorably in performance comparisons [60].

Similar to the ANC approaches discussed above, the *template subtraction* algorithm for removing cardiac interference from electrophysiological measurements [11, 12, 38, 46] is also based on the insight that most heartbeats (of the same subject, in the same recording session, given the same recording condition) have very similar waveforms. Cardiac activity can thus be separated from the remaining signal components by simply averaging over all or a moving array of heartbeats, thereby obtaining an average heartbeat *template*, and then subtracting this heartbeat template from each successive detected beat in the signal. Refinements can be achieved by fitting the calculated template in some way to each particular observed beat (this is an important difference to ANC-based approaches). The difference between the original measurement signal and the (fitted) template signal is then supposed to be the EMG component without the cardiac component. Because cardiac activity dominates much of human physiology, template subtraction algorithms have been proposed for removing cardiac components in many physiological measurement methodologies, not just for electromyographic measurements [26, 27]. Methodologically, the template subtraction algorithm exploits the periodicity of cardiac activity while making no assumptions at all about the waveform or frequency spectrum of the remaining signal

---

[2]One might ask why there should be a more significant frequency overlap in respiratory applications than in other EMG applications. The reason lies in the distance of the recording site to the heart: biological tissue that separates the electrode from the field-generating muscle acts as a (spatial) low-pass filter [19, 49, 51, 54], thereby confining the cardiac component to a lower frequency band in, e.g., limb applications.

components (including electromyographic activity), other than that it not be periodic with the heart rate. In comparison to model-based methods (described below), this represents an advantage: arbitrarily rugged and irregular heartbeat waveforms can be represented exactly as long as they stay constant, i.e., as long as the cardiac component is periodic. Thus, despite its rather simplistic, engineered appearance, the template subtraction algorithm could even be called *nonparametric*.

In its most basic and form,[3] the template subtraction algorithm is equivalent to using an adaptive noise cancellation (ANC) algorithm, with the R peak location as the noise input [26]. Another statistical signal processing perspective on the classical template subtraction algorithm has been proposed by Vullings *et al.* [76]. Denoting the signal component corresponding to cardiac beat k by $\theta_k \in \mathbb{R}^\ell$, where ℓ is a standard heart beat length, the assumption of quasi-periodicity can also be represented as

$$\theta_k = \theta_{k-1} + \eta_k, \tag{5.1}$$

where $\eta_k \sim \mathcal{N}(0, \Sigma_\eta)$ represents the (stochastic) changes in beat shape between beats $k-1$ and k. Furthermore assuming noisy measurements

$$y_k = \theta_k + \nu_k \tag{5.2}$$

with $\nu_k \sim \mathcal{N}(0, \Sigma_\nu)$, a simple linear Kalman filter or smoother can be used to estimate the shape of each cardiac beat. If Σ_η and Σ_ν are both diagonal matrices, the ℓ state estimation problems for the components of θ_k are completely decoupled and can be solved very efficiently using ℓ 1-D Kalman filters or smoothers (instead of a single ℓ-D Kalman filter or smoother) — this is a consequence of the model, i.e., template subtraction, not making *any* assumptions about the waveform of a cardiac beat. If a smoother is used and $\Sigma_\eta = 0$, this method coincides exactly with basic (non-moving, non-fitted) template subtraction: it identifies a constant cardiac template that is used for each heartbeat. Note that everything except the cardiac component is supposed to be captured by ν_k, including in particular the electromyographic component of interest and any measurement noise. Equations (5.1) and (5.2) mathematically explicate three strong (and questionable) assumptions the template subtraction algorithm is based on:

1. differences between the waveforms of subsequent heartbeats are completely stochastic,

2. there is no correlation between the changes to subsequent samples within a heartbeat, compared to the previous heartbeat, and

3. all remaining (non-cardiac) components of the signal can be described as Gaussian white noise.

Section 5.2 will examine ways to relax these assumptions, yielding an improved probabilistic adaptive template subtraction (PATS) algorithm.

While algorithms of the template subtraction family estimate cardiac beat morphology in a nonparametric fashion, there is also a large class of (nonlinear) filtering and smoothing algorithms for cardiac artifact removal based on *parametric heartbeat models*. In an Parametric Heartbeat Models

---

[3]In this context, *basic* means that no fitting of the identified template to each individual observed beat is performed.

influential publication, McSharry *et al.* [44] proposed a simple yet flexible dynamical model of cardiac beat morphology, which describes a cardiac signal as a weighted sum of Gaussians. While formulated in continuous-time by McSharry et al., a discrete-time formulation of the same model has been proposed by Sameni *et al.* [60]:

$$\varphi_k = (\varphi_{k-1} + \Delta t \, \omega_k) \mod 2\pi$$

$$z_k = z_{k-1} - \Delta t \sum_{j=1}^{m} \frac{a_j \omega_k}{b_j^2} \overline{\varphi}_j \exp\left(-\frac{\overline{\varphi}_j^2}{2b_j^2}\right) + \eta_k,$$

with z_k denoting the cardiac signal, φ_k the current phase, $\overline{\varphi}_j$ the phase location of cardiac peak j (with $1 \le j \le m$), Δt the sampling time, ω_k the heart rate and $a \in \mathbb{R}^m$ and $b \in \mathbb{R}^m$ the (amplitude and width) parameters of the Gaussian waves representing the different cardiac peaks. For the process noise, $\eta_k \sim \mathcal{N}(0, \Sigma_\eta)$ is assumed. Various nonlinear filtering and smoothing schemes have been proposed that estimate the time-varying model states and parameters $\varphi_k, z_k, \omega_k, a_k, b_k$ and $\overline{\varphi}_k$, and then use z_k as an estimate of the (denoised) cardiac signal component [6, 58–60]. As with all Kalman filtering and smoothing methods, the quality of the resulting estimates depends chiefly on the tuning of the noise covariance matrices. While these algorithms can be tuned to perform very well on individual datasets, in our recent performance comparison, we found it challenging to find an automatic noise covariance tuning method that achieves high performance across all subjects and recordings [52]. There are, of course, many automatic tuning methods known in the general filtering and smoothing literature, but finding a specific method that works well for this problem currently appears to be an open problem.

Signal Decomposition

Another large group of algorithms for separating cardiac and respiratory components in electrophysiological measurements is based on *decomposing* the original mixture signal into different *components* and then solving the separation problem in the resulting, higher-dimensional signal space. The main challenges for this class of algorithms are, firstly, finding an appropriate basis for the decomposition and, secondly, conceiving a robust method to reject cardiac artifacts in the decomposed space. Most of the well-known decomposition methods have been used for this purpose, including the singular value decomposition (SVD) [43, 48], the empirical mode decomposition (EMD) [25, 34], and the wavelet decomposition [2, 3, 21, 25, 34, 52, 71, 74]. For the rejection of cardiac artifacts in the decomposed, higher-dimensional space, simple threshold-based peak rejection methods [25, 52], ICA [3], and simple rejection of the first or last n components [48, 52] have been employed. (Of course, in the multi-channel case, BSS methods can also be employed directly, i.e., without prior signal decomposition [42, 80]. Here, we focus on the single-channel case, however.) Recently, we have proposed two particularly robust and efficient algorithms using the stationary wavelet transform (SWT) and the empirical mode decomposition (EMD) [25, 52].

ECG Denoising

It should be noted that there is, of course, a huge body of literature on *ECG denoising* methods, which includes methods for removing muscle artifacts from ECG recordings. This might, at first glance, appear to be the same problem as removing cardiac inference from electromyographic measurements. There is an important difference between these two settings, however: the accuracy requirements posed on separation algorithms differ wildly due to the inverted SNRs — cardiac and muscular components play inverted roles in these two separation tasks. Nevertheless, algorithms are often applicable to both problems, and

many of the algorithms described above have first been proposed within the context of ECG denoising [30, 60, 76, 79]. The estimation of *ECG-derived respiration* (EDR) [15] is another closely related field that, however, solves a much simpler problem: the identification of *some* kind of respiratory information in an ECG signal. This is a much less challenging task than recovering the exact shape of an interfered respiratory surface EMG signal.

ECG-derived Respiration

Interestingly, despite the recent surge of biomedical applications of machine learning, the author is unaware of any study directly attempting to solve the cardiac artifact removal problem using *machine learning* techniques. There are a few closely related studies, however. Bockelmann *et al.* [13] use deep learning to predict the envelope of an invasively measured diaphragm EMG from raw surface EMG measurements — thereby implicitly solving the cardiac artifact removal problem. Recently, Zhang *et al.* [82] have applied a deep convolutional neural network for removing EMG artifacts from EEG recordings, a closely related problem. Additionally, several groups have proposed the use of current deep learning techniques for ECG signal denoising [9, 10, 28]. The application of machine learning techniques to the cardiac artifact removal problem might represent an attractive avenue for future research, as will be discussed in section 5.5.

Machine Learning

To conclude this section, fig. 5.2 shows a schematic overview of the different families of cardiac artifact removal algorithms discussed above — including the newly proposed PATS algorithm that will be described in section 5.2.

5.1.3 Quantifying separation success

As was already discussed in the previous chapter, quantifying the performance of biological signal separation algorithms represents a difficult challenge. In virtually all practically relevant settings, the signal components to be separated cannot be measured in isolation without also changing their properties, and thus, simple (pointwise) comparison of the separated components with a true reference signal is usually impossible *in vivo*. This is also the case for the present problem of separating cardiac and myographic components in electrophysiological surface measurements.

Three broad avenues for performing validation remain. Firstly, physiological modeling can be used to simulate each of the individual components realistically and then obtain synthetic measurements from their superposition (e.g., [18]). In this setting, the original signal components are known, and the results of a separation procedure can then be compared to them. This is the approach that was pursued in the previous chapter and Petersen *et al.* [50] for the verification of a blind source separation method for respiratory sEMG measurements. Secondly, in a similar yet slightly different approach, synthetic signals can be obtained by superposing physiological measurements that are similar to the signal components of interest but can be measured in isolation. Regarding the particular problem at hand, various groups have superposed different ECG lead measurements (representing the cardiac signal component) with EMG measurements obtained from muscles that are not subject to strong cardiac contamination [1, 20, 52, 60, 64, 80]. As in the previous approach, the true signal components are known, and the separated components can be validated against them. Thirdly and lastly, one can simply content with the fact that pointwise comparisons to a true reference are not feasible and instead consider alternative (application-specific) measures of separation success, such as periodicity with the heart rate (which can be detected easily in the original, unprocessed signal) [52, 59] or the ratio of
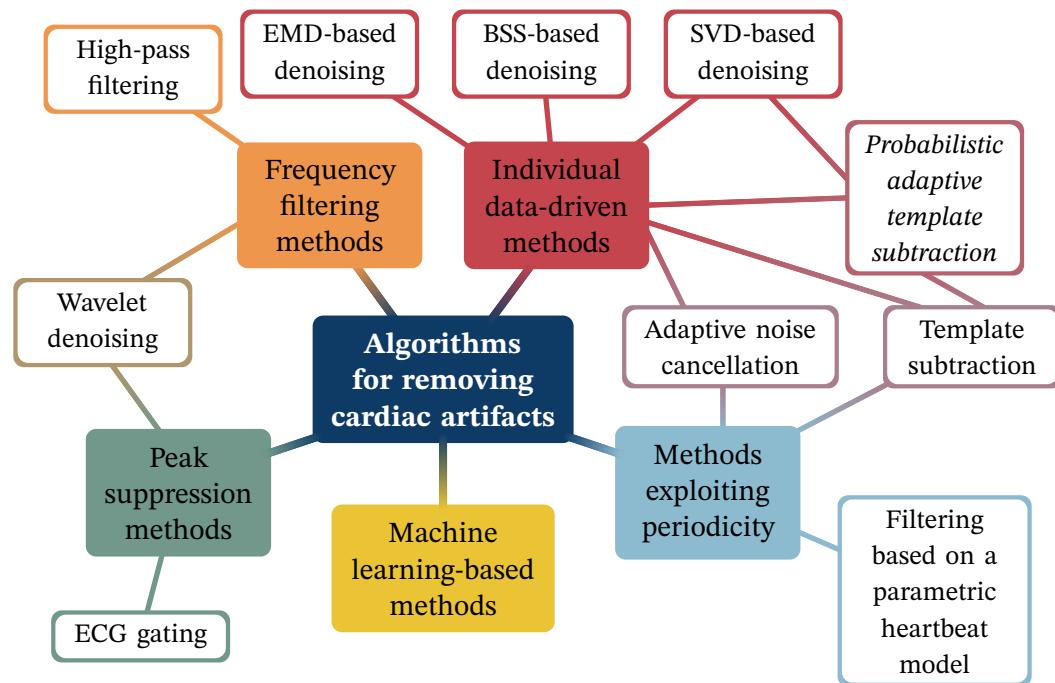
Figure 5.2 – A schematic overview and classification of algorithms for removing cardiac artifacts from single-channel surface EMG measurements. A note on the distinction between *individual data-driven methods* and *machine-learning methods*: the former operate on the data of a single recording session, whereas the latter extract information from a large number of datasets. Probabilistic adaptive template subtraction (PATS) is discussed in section 5.2. Section 5.4 presents a detailed performance comparison of many of these methods. Figure modified and extended with permission from Petersen *et al.* [52].

signal power during cardiac QRS complexes and signal power during the remainder of the signal [52]. In this chapter, the second and third approaches will be pursued. While both approaches individually have their limitations (synthetic signals may always differ in their characteristics from real measurements, and any empirical measure of separation success may fail to detect a certain type of failure [52]), their combination paints a comprehensive picture of the relative performance of different methods under different circumstances.

Previous performance comparisons of algorithms for cardiac artifact removal have been limited to simulated signals [18], data from a single subject [20, 80, 83], or synthetic signals obtained from the superposition of multiple measurement signals (the second approach discussed above) [1, 3, 20, 42, 60, 80, 81]. To the author's knowledge, no previous performance comparison has considered real respiratory sEMG measurements from multiple subjects, except for Slim and Raoof [71], who only considered a single measure of separation success — which, as we will see below, is not sufficient to reliably assess separation performance. Finally, most previous performance analyses compare just a small number of algorithms and do not consider all relevant candidate algorithms. The most comprehensive performance comparison in this regard has recently been provided by Xu *et al.* [81] who compare six different algorithms. However, the authors only consider synthetic data obtained from the superposition of ECG leads and biceps EMG measurements, and

only a single SNR level, which does not capture the wide range of SNRs observed in respiratory sEMG measurements in practice. To summarize, a comprehensive comparative performance evaluation of the most promising proposed algorithms for cardiac artifact removal from respiratory surface EMG measurements has been missing until the author's recent publication Petersen *et al.* [52].

5.2 Probabilistic adaptive template subtraction

In section 5.1.2, the probabilistic formulation of template subtraction given by Vullings *et al.* [76] was described: Denoting the signal component corresponding to cardiac beat k by $\theta_k \in \mathbb{R}^\ell$, where ℓ is a standard heart beat length, Vullings *et al.* [76] propose the random walk model

$$\theta_k = \theta_{k-1} + \eta_k, \tag{5.1}$$

where $\eta_k \sim \mathcal{N}(0, \Sigma_\eta)$ represents the (stochastic) changes in beat shape between beats $k-1$ and k. Furthermore, they assume noisy measurements

$$y_k = \theta_k + \nu_k. \tag{5.2}$$

As discussed in section 5.1.2, this model is based on three simplifying assumptions:

1. differences between the waveforms of subsequent heartbeats are purely random,

2. there is no correlation between the changes (compared to the previous heartbeat) to subsequent samples within a heartbeat, and

3. all remaining (non-cardiac) components of the signal can be described as Gaussian white noise.

None of the three assumptions is typically fulfilled in real measurements. Figure 5.1 (left) shows a clear example of non-stochastic heartbeat waveform evolution: the waveform changes significantly as a function of lung volume due to changes in thoracic geometry [36]. These respiration-related waveform changes, among many other predictable variations such as Mayor waves and other systematic variations [70], clearly violate the first two assumptions stated above. Moreover, surface EMG signals are well known to attain a colored spectrum, as shown in fig. 3.8, which violates the third assumption. For these reasons, an enhanced model and estimation procedure will be proposed in the following, which is derived from the method of Vullings *et al.* [76] while relaxing the first and last of the three assumptions.[4] A proposal for also relaxing the second assumption is made in section 5.2.6.

The newly proposed algorithm can be very roughly summarized as follows: First, all heartbeats are identified and aligned. To identify the statistical properties of the contaminated EMG signal, signal snippets *without* cardiac contamination are analyzed using a simple gating approach, i.e., simply rejecting all samples in the vicinity of a detected heartbeat. An ensemble of independent state-space models is then fit to the time series describing the evolution of individual samples within each heartbeat over the course of the

---

[4]As was mentioned previously, a preliminary version of this algorithm was the subject of the Master's thesis of Seemann [67], which was conducted under the supervision of the author.

full recording, i.e., one model that describes the temporal evolution of the first sample within each heartbeat over time, another model for the evolution of the second sample within each heartbeat over time, and so forth. The identified models are then used for smoothing the time series describing the evolution of the different samples, yielding a denoised version of the cardiac contamination. Finally, the desired cleaned EMG signal is then obtained as the difference between the original signal and this estimate of the cardiac contamination. Algorithm 6 summarizes the new algorithm and may assist the reader in following along the subsequent sections.

5.2.1 Preliminary signal processing

Before discussing the more specific parts of the proposed new method, this section will lay out a number of required preprocessing steps and introduce relevant signals and notation. First, a fourth-order zero-phase 50 Hz bandstop Butterworth filter is used to remove *powerline interference*. Using the filtered signal, R peak locations are then detected using a modified[5] Pan-Tompkins *peak detection algorithm*, yielding a vector $r \in \mathbb{R}^{N^{\text{beats}}}$ of R peak indices. Once R peaks are detected, a vector $c \in \mathbb{R}^{N^{\text{beats}}+1}$ containing the indices of center points between R peaks is also defined, and the beat template length ℓ (in samples) is selected as 1.2 times the average heartbeat length throughout the whole signal. To remove *baseline wander* and slightly simplify the subsequent cardiac artifact removal task [52], the raw signal is then filtered using a 12th-order zero-phase Butterworth high-pass filter with a cut-off frequency of 20 Hz, yielding the signal $\text{EMG}^{\text{preFilt}}$. Moreover, for the purposes of analyzing the characteristics of the EMG component of the measured signal, it will often be useful to consider only those phases of the signal which likely do *not* contain cardiac activity. The QRS segment is usually several dozen ms in length, but this is, of course, not the only phase of cardiac activity. For these reasons, we define a mask "$\neg$ cardiac activity" for a signal EMG as

$$\text{EMG}_{\neg \text{ cardiac activity}} = \big(\text{Filter}(\text{EMG}, H_{>30\text{Hz,ZP}})\big)_{\neg \text{ close to QRS} \wedge \neg \text{ outlier}} \tag{5.3}$$

where "$\neg$ close to QRS" excludes all samples within a $[-150\,\text{ms}, 150\,\text{ms}]$ window around each detected R peak, $\neg$ outlier excludes extreme outliers (defined as having an amplitude above the 0.999 quantile of the signal's amplitude), and $H_{>30\text{Hz,ZP}}$ is a zero-phase high-pass filter with a cut-off frequency of 30 Hz.

Next, a *whitened* representation of the signal is computed, as discussed in section 3.1.2. Our approach differs from the usual whitening approach, however: normally, the whitening filter is determined as the inverse of a spectral estimate of the *whole* signal. Here, however, the whitening filter will be identified using only segments without cardiac activity. We make this choice in order to ensure that everything except for the cardiac signal component can be described as uncorrelated noise, which will become relevant in the following section. A sixth-order autoregressive model H_{burg} is fit to $\text{EMG}^{\text{preFilt}}_{\neg \text{ cardiac activity}}$, and $\text{EMG}^{\text{preFilt}}$ is then filtered with the inverse of that autoregressive model, yielding the whitened signal $\text{EMG}^{\text{white}}$. Spectral estimation on signals with gaps is a current field of

---

[5]After a basic run of the classical Pan-Tompkins algorithm [47], an additional step is performed during which each detected peak is re-aligned by maximizing correlation with the average peak over the whole dataset. We found that this increases robustness in the case of R peaks with multiple local maxima or minima.

Algorithm 6: Probabilistic Adaptive Template Subtraction

---

1 **Function** PATS$\big(\text{EMG}^{meas}, u\big)$

    Input   : A raw, single-channel EMG measurement $\text{EMG}^{\text{meas}} \in \mathbb{R}^{N^S}$ with strong cardiac contamination, and an input signal $u \in \mathbb{R}^{N^{\text{beats}} \times N^u}$ which contains additional information about the shape of each beat.

    Output: A cleaned signal $\text{EMG}^{\text{cleaned}} \in \mathbb{R}^{N^S}$ in which most cardiac interference has been removed, with only EMG activity remaining.

2   **begin**

3         $\text{EMG}^{\text{preFilt}} = \text{Filter}(\text{EMG}^{\text{meas}}, H_{>20\text{Hz}})$

4         $r = \text{FindPeakIdces}(\text{EMG}^{\text{preFilt}}), \quad r \in \mathbb{R}^{N^{\text{beats}}}$

5         $c = \text{FindCenterBetweenPeakIdces}(r), \quad c \in \mathbb{R}^{N^{\text{beats}}+1}$

6         $\ell = \text{RoundOdd}(1.2 \cdot \text{mean}(\text{diff}(r))), \quad \ell \in \mathbb{N}$ Template length

7         $j_{\text{peak}} = \lceil 0.5 \cdot \ell \rceil, \quad j_{\text{peak}} \in \mathbb{N}$ R peak position in templates
Whiten the signal to justify the assumption of uncorrelated measurement noise.

8         $H_{\text{burg}} = \text{PSD Burg}(\text{EMG}^{\text{preFilt}}_{\neg \text{ cardiac activity}})$

9         $\text{EMG}^{\text{white}} = \text{Filter}(\text{EMG}^{\text{preFilt}}, H_{\text{burg}}^{-1})$

10   **for** $j = 1$ **to** ℓ **do**
Identify a maximum-likelihood state-space model for the time course of the jth sample of all beats.

11            $y = \text{EMG}^{\text{white}}_{c+(j-j_{\text{peak}})}$ Vector of all jth samples
Minimize the negative log likelihood of the data. See Eqs. (5.5)–(5.8) for the definitions of the state-space model.

12            $\psi^* = \arg\min_{\psi \in \Psi} -\log p\big(y \mid u, A(\psi), B(\psi), C, \Sigma_\eta(\psi), \mu_\eta(\psi), \Sigma_\nu\big)$
Estimate the denoised time course of the jth sample of all beats. Use a Rauch–Tung–Striebel smoother (RTS), see algorithm 4.

13            $\hat{\theta}^s_{1:N^{\text{beats}}} = \text{RTS}\big(y, u, A(\psi), B(\psi), C, \Sigma_\eta(\psi), \mu_\eta(\psi), \Sigma_\nu\big)$

14            $T_{:,j} = \hat{\theta}^s_{1:N^{\text{beats}}}$ Matrix containing templates of all beats

15   **end**

       (Specialized handling of first and last beats not shown here)

16   **for** $i = 1$ **to** N^{beats} **do**
Row $T_{i,:}$ contains the estimated template of beat i. If an RR intervals is longer or shorter than ℓ, interpolate linearly between adjacent beats.

17            $\text{EMG}^{\text{cardiac}}_{c_i:c_{i+1}} = \text{CreateTemplate}(r_{i-1:i+1}, T_{i-1:i+1,:})$

18   **end**

19        $\text{EMG}^{\text{cleaned, white}} = \text{EMG}^{\text{white}} - \text{EMG}^{\text{cardiac}}$
Reverse whitening to recover original spectral characteristics.

20        $\text{EMG}^{\text{cleaned}} = \text{Filter}(\text{EMG}^{\text{cleaned, white}}, H_{\text{burg}})$

21   **end**

---

research [16], especially in the field of biomedical estimation where missing data are a frequent problem [57]. Our approach here is to treat the different segments of non-cardiac data like individual "experiments" and perform classical multi-experiment autoregressive model parameter identification by joint prediction error minimization across these different experiments [40]. This method is methodologically very similar to the *Burg algorithm for segments* [77] which was recently found to outperform other algorithms in ECG spectrum analysis with gaps [57].

5.2.2 Model proposal

To relax the assumption of purely stochastic heartbeat variation, we propose to replace the whole-beat random walk model (5.1) by a series of ℓ models, one per sample $i = 1, \dots, \ell$ in each heartbeat. In particular, we propose to use third-order autoregressive models with exogenous inputs, i.e.,

$$\text{beat}_{ki} = a_{1i}\text{beat}_{k-1 i} + a_{2i}\text{beat}_{k-2 i} + a_{3i}\text{beat}_{k-3 i} + b_{1:N^u,i}^\mathsf{T} u_k + \eta_{ki}, \tag{5.4}$$

where beat_{ki} denotes sample i of the waveform $\text{beat}_k \in \mathbb{R}^\ell$ of heartbeat k, $u_k \in \mathbb{R}^{N^u}$ denotes sample k of an exogenous input signal $u \in \mathbb{R}^{N^u \times N^{\text{beats}}}$, and $\eta_{ki} \sim \mathcal{N}\left(\mu_{\eta i}, \sigma_{\eta i}^2\right)$ Gaussian noise. We gather all model parameters in a vector

$$\psi_i = \begin{pmatrix} a_{1i} & a_{2i} & a_{3i} & b_{1i} & \cdots & b_{N^u i} & \mu_{\eta i} & \sigma_{\eta i} \end{pmatrix}^\mathsf{T} \quad \in \mathbb{R}^{N^u+5}$$

to be tuned. Furthermore assuming noisy measurements, we obtain a series of ℓ state-space models

$$\begin{aligned} \theta_{ki} &= A(\psi_i)\theta_{k-1 i} + B(\psi_i)u_k + \underline{\eta}_{ki}, \quad \underline{\eta}_{ki} \sim \mathcal{N}\left(\underline{\mu}_{\eta i}, \underline{\Sigma}_{\eta i}\right) \\ \text{EMG}_{ki}^{\text{white}} &= \theta_{ki} + \nu_k, \quad \nu_{ki} \sim \mathcal{N}(0, \Sigma_\nu) \end{aligned} \qquad i = 1, \dots, \ell \tag{5.5}$$

with model matrices

$$A(\psi) = \begin{pmatrix} \psi_1 & \psi_2 & \psi_3 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \quad B(\psi) = \begin{pmatrix} \psi_4 & \cdots & \psi_{N^u+3} \\ 0 & \cdots & 0 \\ 0 & \cdots & 0 \end{pmatrix} \quad C = \begin{pmatrix} 1 & 0 & 0 \end{pmatrix} \tag{5.6}$$

and process noise parameters

$$\underline{\mu}_\eta(\psi) = \begin{pmatrix} \psi_{4+N^u} \\ 0 \\ 0 \end{pmatrix} \quad \underline{\Sigma}_\eta(\psi) = \begin{pmatrix} \psi_{5+N^u}^2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

In eq. (5.5), one time step corresponds to one heartbeat, and $\theta_{ki} \in \mathbb{R}^3$ denotes the state vector corresponding to sample i in beat k. The input signal u_k can be any signal that provides further information about heartbeat k's waveform; several useful choices for u_k will be discussed in section 5.2.5. By $\text{EMG}_{ki}^{\text{white}}$, we denote the sample corresponding to sample i in beat k of the whitened signal $\text{EMG}^{\text{white}}$ (see the previous section). Finally, for the measurement noise covariance, we assume

$$\Sigma_\nu = \text{Var}[\text{EMG}_{\neg\text{ cardiac activity}}^{\text{white}}]. \tag{5.7}$$

where "¬ cardiac activity" is defined as in eq. (5.3). The usage of EMG$^{\text{white}}$ as a measurement signal (instead of EMG$^{\text{preFilt}}$) serves to partially satisfy the third assumption stated in the introduction of this section: while this modification does *not* ensure the Gaussianity of the measurement noise process, it *does* ensure its uncorrelatedness. Furthermore, EMG signals are generally known to attain a distribution that lies between a Gaussian and a Laplacian [17, 51] — thus, the Gaussianity assumption is also not unreasonable.

Notice that, unlike some previously proposed algorithms [60], we do not *stretch* or *shrink* heartbeats to conform with some standardized length. Instead, heartbeats are taken exactly as they are, and simply cut off at a fixed width to the right and to the left of the detected R peak, yielding a set of observed heartbeats of identical lengths. This approach was selected because it was observed empirically that the length and shape of the QRS complex — which represents the most important component of the cardiac contamination — stayed relatively constant, regardless of the total length of the heartbeat. (See fig. 5.3, bottom right panel, for some evidence.) Thus, stretching or compressing the heartbeat to conform with a standardized length would significantly increase the difficulty of the separation problem since QRS complexes would be rendered more different than they really are.

Heartbeats of Different Lengths

5.2.3 Estimation procedure

Given model eqs. (5.5) to (5.7), how can we identify the model parameters ψ_i and the states θ_{ki} for all beat samples $i = 1, \dots, \ell$, finally yielding an estimate of the time course of the cardiac signal component EMG$^{\text{cardiac}}$? An important observation about eqs. (5.5) to (5.7) is that these ℓ models of the time courses of the different beat samples are completely *decoupled*. Instead of a single, $3 \cdot \ell$-dimensional state-space model, one can solve ℓ three-dimensional estimation problems, which tremendously reduces computational effort & memory requirements.

With this in mind, one can solve both estimation problems using the standard tools described in chapter 2: the model parameters ψ_i are identified using *maximum likelihood (ML) estimation* as described in section 2.4, and the subsequent (as well as inner-loop) state estimation problem is solved using a Rauch-Tung-Striebel (RTS) smoother as described in section 2.3.1. To solve the nonlinear likelihood maximization problem (2.33), we employ the energy recursion (2.40). A convex set of feasible values for ψ is defined by

Maximum Likelihood Estimation

$$\Psi = \left\{ \begin{pmatrix} a_1 & a_2 & a_3 & b_1 & \dots & b_{N^u} & \mu_\eta & \sigma_\eta \end{pmatrix}^\mathsf{T} \;\middle|\; \begin{array}{c} -2 \le a_i \le 2 \\ -0.2 \le b_i \le 0.2 \\ -0.5 \le \mu_\eta \le 0.5 \\ 0 \le \sigma_\eta \le 0.2 \end{array} \right\}, \tag{5.8}$$

and we use the interior point algorithm provided by the Matlab `fmincon` command [14].[6] If desired, the computational effort required for solving the parameter optimization problem could be significantly reduced by implementing gradient-based optimization using the gradient expressions provided by Särkkä [61, p. 214], but this was not deemed necessary here.

---

[6]Various solvers were tried and compared heuristically, including the mesh-adaptive direct search algorithm NOMAD [4, 37]. The `fmincon` solver was found to yield equally good results at much lower computation time.

5.2.4 Removal algorithm

Once the whitened signal $\mathrm{EMG}^{\mathrm{white}}$, the model parameters ψ_i and the states θ_{ki} have been obtained for all beat samples i as described in the previous two sections, the first entries of all state estimates θ_{ki} are collated to obtain the final estimate of the cardiac signal component $\mathrm{EMG}^{\mathrm{cardiac}} \in \mathbb{R}^{N^S}$ (see algorithm 6 for details). This cardiac signal estimate is then subtracted from the whitened measurement signal to obtain

$$\mathrm{EMG}^{\mathrm{cleaned,\ white}} = \mathrm{EMG}^{\mathrm{white}} - \mathrm{EMG}^{\mathrm{cardiac}}.$$

Finally, to recover the frequency characteristics of the original signal, we perform the inverse of the whitening operation used initially and obtain the desired, cleaned signal estimate

$$\mathrm{EMG}^{\mathrm{cleaned}} = \mathrm{Filter}(\mathrm{EMG}^{\mathrm{cleaned,\ white}}, H_{\mathrm{burg}}).$$

5.2.5 Choice of shape information signals

So far, the discussion always assumed a suitable input signal $u \in \mathbb{R}^{N^{\mathrm{beats}} \times N^u}$ to be available. This section will discuss several suitable choices for such an input signal. The general requirement for something to be a suitable input signal is simply that, in agreement with the model (5.5), *it should provide some information about the shape of the different beats*. Obviously, there are many choices for u that fulfill this requirement: any signal that correlates with changes in cardiac beat morphology is feasible. Of the many possible choices for input features, two will be considered here:

RR interval length. A first option is to use the current RR interval length $u_{\mathrm{RR},k}$, i.e., the distance between the current and the previous R peak: it is well known that RR interval length is a highly predictive feature regarding ECG beat morphology [5, 33, 39]. Figure 5.3 illustrates the correlation between $u_{\mathrm{RR},k}$ and beat morphology.

SVD-based beat characterization. As a second option for an input signal, a singular value decomposition (SVD)-based characterization of beat morphology will be considered, as has been successfully employed in many previous cardiac signal processing applications [29, 32, 48, 78]. To this end, a matrix M_{QRS} is constructed that contains in each row the samples contained in a $\pm 100\,\mathrm{ms}$ window around each detected R peak. Next, the first N^{PC} principle components $\mathrm{PC}_{\mathrm{QRS},i}$ of M_{QRS} are computed — as is well-known, these represent, in descending order of importance, the most important (in terms of explaining signal variance) directions of QRS complex shape variation. As such, the projection

$$u_{\mathrm{PC},k,i} := M_{\mathrm{QRS},k}\mathrm{PC}_{\mathrm{QRS},i}^{\mathsf{T}}$$

of each QRS complex onto these N^{PC} principal components (the *principal coordinates*) provides a nonparametric and maximally informative characterization $u_{\mathrm{PC},k} \in \mathbb{R}^{N^{\mathrm{PC}}}$ of beat morphology that may be used as part of the input signal u_k. Figure 5.3 illustrates how u_{PC} characterizes beat morphology.

One can, of course, also combine features and use multiple delayed versions of the same features. Note that many other shape-varying factors could also be considered, such as lung volume or other characteristic ECG beat features.
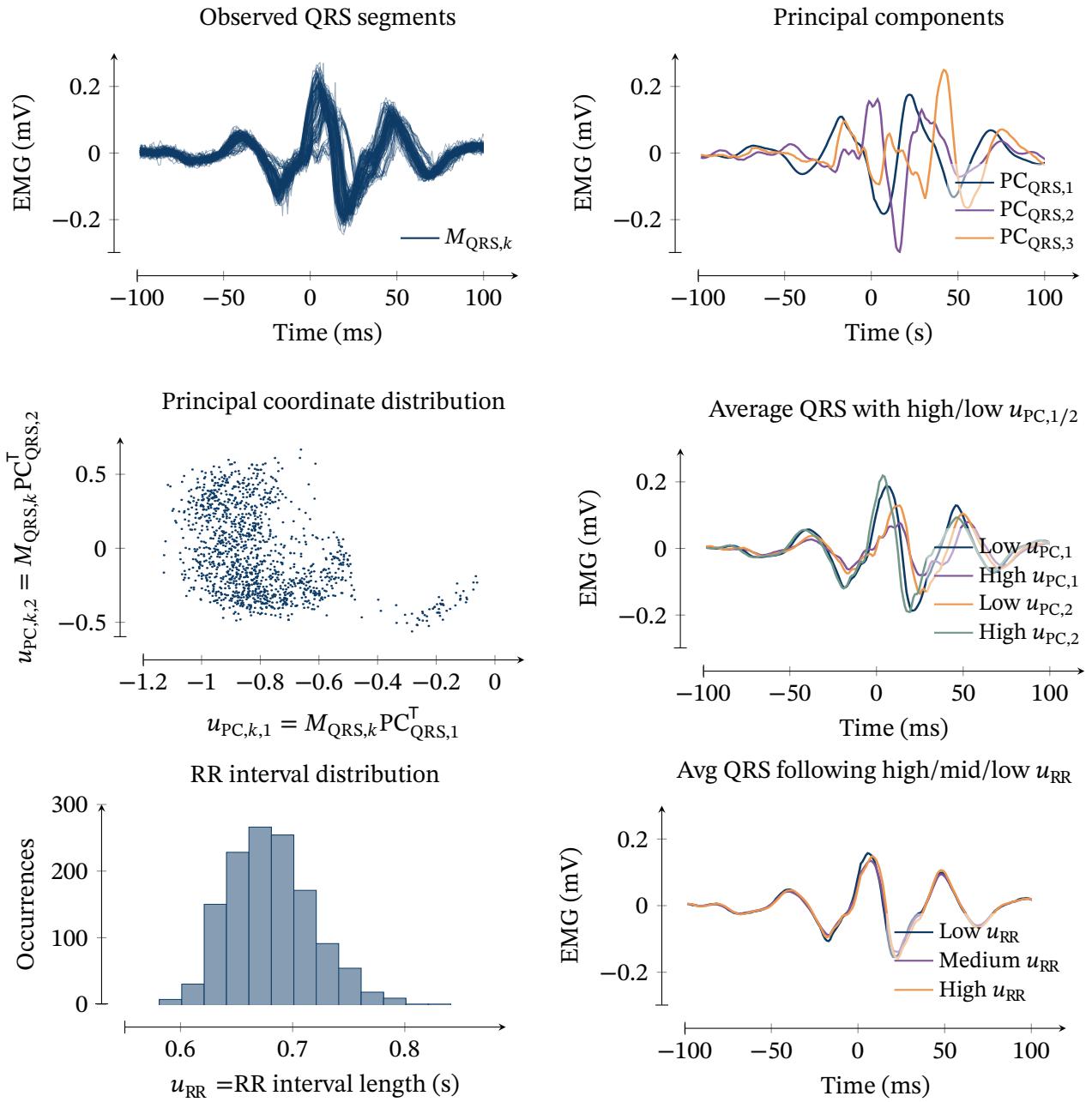
Figure 5.3 – An exploration (based on one sample recording) of two features used for characterizing cardiac beat morphology: principal coordinates and previous RR interval length. *Top row:* observed QRS complexes and their first three principal components. *Middle row:* scatter plot of the first two principal coordinates of each observed heartbeat (left), as well as the averages of all QRS complexes with first/second principal components in the respective bottom and top decile (right). *Bottom row:* observed distribution of RR intervals lengths (left) and averages of all QRS complexes following an RR interval in the 1st, 5th, and 10th decile. A statistical dependency between the QRS segment waveform and both features can be observed (center right panel and bottom right panel), although the relationship with the previous RR interval length is weak in this recording. Thus, both quantities represent potentially useful input features for predicting QRS segment waveform. The data are from recording 2, channel 1, of the study described in section 3.3.2, high-pass filtered with a cut-off frequency of 20 Hz.

5.2.6 Remarks on the proposed solution

To summarize, what are the properties of the proposed model and removal algorithm?

- The approach is adaptive to every single recording and completely unsupervised. All free model parameters are estimated based on the single recording at hand, as opposed to machine-learning approaches based on large labeled training datasets. No calibration procedure of any sort is required.

- The model is nonparametric and completely unconstrained with respect to cardiac beat morphology. Arbitrarily rugged heartbeat shapes can be represented without bias. This is in contrast to methods based on smooth and parametric models of heartbeat morphology.

- Predictable variations of beat morphology due to, e.g., respiratory activity or Mayer waves, are represented by the model. These are incorporated in the model by using the autoregressive heartbeat evolution model (5.5) and input signals such as the RR interval length. To the author's knowledge, none of the known approaches takes these predictable temporal variations into account.

- Unpredictable but systematic variations of beat morphology are *also* controlled for by the model. This is due to the use of the SVD-driven (nonparametric) beat characterization. This allows the algorithm to (continuously) differentiate between different clusters of heartbeats, independent of their time of appearance.

- The noise model is realistic. EMG signals are usually nearly normally distributed, and by means of the specialized whitening step, it is ensured that the noise component is — as assumed by the model (5.5) — uncorrelated. Previously proposed model-based approaches often assumed simple and unrealistic noise models, which, as discussed at length in section 2.1.4, can crucially affect estimation results.

- No measurements or data beyond the single-channel measurement to be denoised are required. Moreover, the measurement need not necessarily be an electrophysiological measurement; the approach is equally well applicable to other physiological measurement modalities affected by cardiac interference, so long as the whitened target signal has an amplitude distribution that is sufficiently close to a Gaussian. (This has not been further tested or verified here, though.)

- Uncertainty quantification is directly available via the RTS smoother's state estimation covariance. This can be used to automatically detect regions where removal performance might be insufficient and to inform further signal processing steps that rely on the cleaned signal.

Of course, the proposed model is one of a myriad of conceivable models. Its features are, however, unique among previously proposed methods, as mentioned above. While model selection is performed with respect to the choice of the input signal in section 5.4, a more extensive model search would likely yield an even better model structure. Note, though, that the class of feasible model structures is limited significantly by the requirement of identifying all model parameters on a single recording. The model proposed here strikes a balance between providing enough flexibility to faithfully represent a large class of possible signal

waveforms on the one hand and incorporating sufficient prior knowledge to enable sample-efficient inference on the other hand.

While the proposed approach solves many problems that other methods exhibit — as confirmed by its good comparative performance, which will be discussed in section 5.4 — one of the unrealistic assumptions noted in the introduction remains: it is still assumed that there is no correlation between the changes to subsequent samples within a heartbeat compared to the previous heartbeat. Mathematically, this is reflected in the assumption that η_{ki} in eq. (5.5) is completely independent of all other process noise terms η_{kj}, which is, of course, not realistic. The assumption can be relaxed by assuming some model of correlation between neighboring process noise realizations η_{ki} and η_{ki+1} — e.g., in the simplest case, a first-order autoregressive model — but this, crucially, introduces *coupling* between all ℓ state and parameter estimation problems that could so far be solved separately from one another. Thus, instead of ℓ separate $(5 + N^u)$-dimensional maximum likelihood parameter estimation problems and ℓ separate 3-dimensional state estimation problems, one then must solve (assuming an otherwise identical model structure) a single $\ell \cdot (5 + N^u)$-dimensional maximum likelihood parameter estimation problem and a single $3 \cdot \ell$-dimensional state estimation problem. Using a memory-efficient implementation of gradient-based ML estimation, this is certainly feasible — it does, however, significantly increase computational and implementation complexity, and was therefore not pursued here. This may represent an attractive opportunity for further research.

5.3 Validation & performance quantification

This section will describe a comprehensive framework for the validation and performance quantification of algorithms for the removal of cardiac interference from single-channel respiratory sEMG measurements.[7] To briefly recapitulate, the key challenge for a meaningful validation regards the lack of a true reference signal. Two different, complementary methods for addressing this challenge are pursued in the following. Firstly, section 5.3.1 describes a study in which real respiratory sEMG measurements were obtained from healthy subjects, and cardiac artifact removal procedures are evaluated based on two empirical measures of separation success. Secondly, section 5.3.2 describes the construction of synthetic data from non-respiratory sEMG measurements (free of cardiographic contamination) and ECG leads (free of myographic contamination), in which case a true reference signal is available for the validation. Section 5.3.3 lists six algorithms that have been proposed previously in the literature and to which the newly proposed PATS algorithm will be compared.

5.3.1 Real study data

In addition to the use of synthetic data for the validation (as described in the following section), the algorithms were also verified using real respiratory measurement data. To this end, study data were collected for this particular purpose in a Bachelor's thesis conducted under the supervision of the author, as discussed in more detail in section 3.3.2. As a brief

---

[7]This section is largely based on Petersen *et al.* [52]. An initial version of the validation framework described in this section has been developed in the Bachelor's thesis of Sauer [64] under the supervision of the author.

summary, two respiratory surface EMG channels (electrode positions as shown in fig. 3.13) were recorded for 15 minutes from ten healthy subjects lying supine on a bench while they breathed through an inspiratory resistor. All signals were assessed for cardiac anomalies by a clinical expert[8]; none were found except for possible sinus arrhythmia in one subject. Of the total of 20 recordings (ten subjects, two channels each), 14 were used for further analyses; reasons for exclusions are noted in table 3.1. All signals were preprocessed using a fourth-order zero-phase Butterworth band-stop filter with a stop frequency of 50 Hz. Furthermore, for all algorithms except the wavelet denoising method, the signals were pre-filtered using a zero-phase high-pass filter with a cut-off frequency of either 10 Hz (for EKS2 and EKS25) or 20 Hz (all remaining algorithms).

As was discussed extensively already, no true reference data are available in the case of real measurements. Hence, a validation scheme must rely on heuristic measures of separation success. Here (as in Petersen *et al.* [52]), two different heuristic performance measures are considered: the improvement in an empirical *signal-to-noise ratio (SNR)* from the raw measurements to the cleaned signals, and the improvement in a *periodicity measure (PM)* that quantifies the periodicity of the signal with the heart rate.

Empirical Signal-to-noise Ratio For estimating the *empirical signal-to-noise ratio* (SNR), we attempt to identify some signal phases during which either only inspiratory muscle activity and no cardiac activity is present, or only cardiac activity and no inspiratory muscle activity. We will then use the first type of sections to estimate the *signal* power, i.e., the EMG power, and the second type of sections to estimate the *noise* power, i.e., the cardiac contaminant power. For the first purpose, we re-use the "$\neg$ cardiac activity" operator defined earlier in eq. (5.3). As a brief reminder, this operator filters the signal with a zero-phase high-pass filter with a cut-off frequency of 30 Hz to remove any remaining baseline activity and then selects those samples of the signal that are *not* within a $[-80\,\text{ms}, 80\,\text{ms}]$ window around each detected R peak. Inspiratory activity is detected by automatic, threshold-based inspection of the airway pressure signal, and we estimate the (myographic) *signal* power as

$$\text{P}\big\{\text{Filter}(\text{EMG}, H_{>30\text{Hz}})_{\text{insp}\wedge\neg\text{ close to QRS}}\big\},$$

where $\text{P}\{X\} = \text{E}[X^2]$ denotes the power of signal X. Similarly, the (cardiac) *noise* power is estimated as

$$\text{P}\big\{\text{EMG}_{\neg\text{ insp}\wedge\text{close to QRS}}\big\}$$

and we finally obtain the empirical SNR (in decibels)

$$\text{SNR}\{\text{EMG}\} = 20 \cdot \log_{10} \frac{\text{P}\big\{\text{Filter}(\text{EMG}, H_{>30\text{Hz}})_{\text{insp}\wedge\neg\text{ close to QRS}}\big\}}{\text{P}\big\{\text{EMG}_{\neg\text{ insp}\wedge\text{close to QRS}}\big\}}. \tag{5.9}$$

Using this definition, the min-max range of the SNR of $\text{EMG}^{\text{preFilt}}$ (which, as a reminder, is the raw signal preprocessed using a high-pass filter with a cut-off frequency of 20 Hz) across all datasets is $-50.82\,\text{dB}$ to $-11.54\,\text{dB}$, with a median of $-36.38\,\text{dB}$. This range will be used below in section 5.3.2 to create a set of synthetic measurements with realistic SNRs. Because

---

[8]The author would like to thank N. Carbon, Department of Anesthesiology, and Operative Intensive Care Medicine, Charité Universitätsmedizin Berlin, Germany, for assessing the study data with respect to cardiac anomalies.

we are interested in the *improvement* in SNR achieved by the different algorithms, we will otherwise mostly concern ourselves with the difference

$$\Delta\text{SNR}\big\{\text{EMG}^{\text{cleaned}}\big\} = \text{SNR}\big\{\text{EMG}^{\text{cleaned}}\big\} - \text{SNR}\big\{\text{EMG}^{\text{preFilt}}\big\} \quad (5.10)$$

between the signal $\text{EMG}^{\text{preFilt}}$ preprocessed by a 20 Hz zero-phase high-pass filter, and the cleaned signal $\text{EMG}^{\text{cleaned}}$.

The second empirical measure of separation success which we will consider in the following is the *periodicity measure* (PM) introduced by Sameni *et al.* [59]. This measure quantifies the quasi-periodicity of a signal with respect to a second, phase-providing signal. In the given setting, it quantifies the correlation between signal samples at identical phase positions in successive cardiac beats: if these are well-correlated, cardiac information is still present in the signal. If they are uncorrelated, we may assume that cardiac interference has been largely removed. To define this measure mathematically, we first denote the temporal distance between sample k and its "successor" sample one heartbeat later by

<div style="text-align: right">Periodicity
Measure</div>

$$\tau_k = \min\{\tau \mid \Phi_{k+\tau} = \Phi_k, \tau > 0\}, \quad (5.11)$$

where Φ_k denotes a phase signal which is obtained by linear interpolation from 0 to 2π between two successive detected R peaks. Then, the PM is defined as the Pearson coefficient of correlation [59]

$$\text{PM} = \frac{|\text{E}[x_k x_{k+\tau_k}]|}{|\text{E}[x_k^2]\,\text{E}[x_{k+\tau_k}^2]|^{\frac{1}{2}}}. \quad (5.12)$$

If PM $= 1$, the signal is perfectly quasi-periodic with the (potentially time-varying) heart rate, whereas if PM $= 0$, there is no remaining periodicity with the heart rate. Analogously to ΔSNR, we will mainly consider the *improvement* (in decibels)

$$\Delta\text{PM}\big\{\text{EMG}^{\text{cleaned}}\big\} = 20 \cdot \log_{10} \frac{\text{PM}\big\{\text{EMG}^{\text{preFilt}}\big\}}{\text{PM}\big\{\text{EMG}^{\text{cleaned}}\big\}} \quad (5.13)$$

between $\text{EMG}^{\text{preFilt}}$ and $\text{EMG}^{\text{cleaned}}$ for assessing and comparing the performance of the different algorithms.

5.3.2 Synthetic data

To create a realistic dataset in which a ground truth for separating cardiac and myographic signal components is known, single-lead ECG signals were superposed with surface EMG measurements above the *musculus gastrocnemius*, which are largely free of cardiac interference. The ECG recordings were obtained from the Physionet PTB diagnostic database [24]. To capture some variability, 15 different leads recorded from two healthy subjects have been used (patient ID131, male, age 26, and patient ID 185, female, age 22). ECG leads are not generally free of myographic noise — which would interfere with the purpose of this whole analysis —, and thus a criterion for the rejection of noisy measurements was developed: using the same window definition as throughout this whole chapter, the power of signal components of $\text{EMG}^{\text{preFilt}}$ inside and outside a $[-80\,\text{ms}, 80\,\text{ms}]$

window around each detected R peak was compared, and signals with a noise level above $-35\,\text{dB}$ were discarded. After this rejection step, 25 ECG signals (of the initial 30) remained. For the surface EMG components, two signals (one with periodic contractions and one with irregular contractions) were recorded from a single, healthy subject (female, age 21) during the Bachelor's thesis of Sauer [64]. All pairs of components were combined at five different SNR levels using

$$\text{EMG}^{\text{syn}} = \text{EMG}^{\text{meas}} + k \cdot \text{ECG}^{\text{meas}},$$

yielding a total of 250 synthetic signals, each with a length of $115\,\text{s}$, ranging from SNR (calculated according to eq. (5.9)) $-53.62\,\text{dB}$ to $-13.91\,\text{dB}$ with a median of $-36.21\,\text{dB}$. These values are very close to those observed in the previous section on real respiratory sEMG measurements.

In addition to the two performance measures ΔSNR and ΔPM defined in the previous section, two additional error measures were computed. Firstly, the normalized norm of the point-wise *raw* error

Raw Signal Error

$$e_{\text{raw}} = \frac{\|\text{EMG}^{\text{meas}} - \text{EMG}^{\text{cleaned}}\|_2}{\|\text{EMG}^{\text{meas}}\|_2} \tag{5.14}$$

was computed to quantify the raw signal recovery error. (In this and the following performance evaluation, extreme outliers for which

$$|\text{EMG}^{\text{cleaned}}| > |\max\{\text{ECG}^{\text{meas}}\}| \tag{5.15}$$

were rejected. These were generally rare and mainly occurred with some of the model-based filtering procedures.) As a secondary performance measure, the normalized norm of the point-wise *envelope* error

Envelope Signal Error

$$e_{\text{env}} = \frac{\|\overline{\text{EMG}}^{\text{meas}} - \beta_1^* \overline{\text{EMG}}^{\text{cleaned}} - \beta_0^*\|_2}{\|\overline{\text{EMG}}^{\text{meas}}\|_2} \tag{5.16}$$

was computed, with $\overline{\text{EMG}}$ denoting the centralized mean average value (MAV) envelope signal calculated with a window length of 128 samples, and

$$\beta_1^*, \beta_0^* = \underset{\beta_1, \beta_2}{\arg\min} \|\overline{\text{EMG}}^{\text{meas}} - \beta_1 \overline{\text{EMG}}^{\text{cleaned}} - \beta_0\|_2. \tag{5.17}$$

This performance measure was selected in order to quantify the error in the *waveform* of the MAV envelope and not so much the error in its offset or amplitude.

5.3.3 Algorithms under comparison

In the following, the performance of a number of algorithms for the removal of cardiac artifacts from respiratory surface EMG measurements will be assessed and compared. The compared algorithms are the following:

PATS The algorithm proposed in the previous section. Four versions of this algorithm will be considered: the basic algorithm without any shape information input signal u (PATS), with the RR interval length $u_{\text{RR},k}$ as an input signal (PATS+RR), with the principal component-based input $u_{\text{PC},k}$ (PATS+PC, see the previous section for details), and with both $u_{\text{RR},k}$ and $u_{\text{PC},k}$ used simultaneously as input signals (PATS+RR+PC).

TS A classical template subtraction algorithm, in the slightly improved version described in Petersen *et al.* [52] as "adaptive template subtraction".

SWT The stationary wavelet transform (SWT)-based denoising method described in detail in Petersen *et al.* [52].

HP200 A simple fourth-order, zero-phase Butterworth high-pass filter with a cut-off frequency of 200 Hz.

EKS2, EKS25 The model-based filtering methods of Sameni [58] and Sameni *et al.* [60] (EKS2) and Akhbari *et al.* [6] (EKS25).

EMD The empirical model decomposition-based removal procedure described in Graßhoff *et al.* [25] and Petersen *et al.* [52].

All Kalman filtering procedures (PATS, EKS2, EKS25) have been implemented using a sequential filtering scheme to avoid unnecessary matrix inversions [69]. To guarantee the symmetry and positive definiteness of the covariance matrices, the Joseph stabilized covariance update was used [69]. In the EKS25, which employs an extended Kalman filter/smoother, state constraints were implemented using the projection method described by Simon [68]. We found these constraints to lead to significantly improved performance compared to the original implementation of Akhbari *et al.* [6]. For the smoothing step, the Rauch-Tung-Striebel (RTS) smoother [55, 69] was employed. (Sequential filtering, the Joseph stabilized covariance update, state constrained filtering, and the RTS smoother have all been discussed and introduced in section 2.3.1.) A new tuning method was used for the EKS25, which adjusts the EKS25 noise parameters based on the cardiac template amplitude as well as the identified EMG noise covariance. Finally, all Butterworth filters were implemented in second-order sections form for improved numerical stability.

5.4 Results

In this section, the results of the validation procedures described in section 5.3 for all of the algorithms listed in section 5.3.3 will be shown. The full code implementing all algorithms under consideration as well as the validation framework is openly available at `https://github.com/ime-luebeck/ecg-removal`.

5.4.1 Real study data

Figure 5.4 shows exemplary results of all algorithms under consideration (raw and a resulting envelope signal), applied to a snippet of a recording of one study subject. The algorithms remove the cardiac interference to varying degrees of success. Especially in the envelope signals, some degree of residual cardiac contamination (identifiable due to its periodicity with the heartbeat) can be observed in practically all algorithms, with the newly proposed PATS algorithm, SWT denoising, the HP200, and EMD denoising exhibiting the cleanest envelope signals in this particular example. Some of the algorithms markedly suppress the amplitude of the EMG signal, indicating changes in the spectral properties of the signals due to the removal procedure (refer to Petersen *et al.* [52] for an empirical analysis of the spectral effect of all algorithms except the newly proposed PATS algorithm). This effect is

most visible in the results of the HP200 and EMD methods and may be of concern when targeting, e.g., sEMG-based fatigue analyses or other analyses of spectral signal properties. An interesting effect can be observed in the results of the EKS25 procedure on this data snippet, which performs a kind of "gating" by suppressing the whole signal around each QRS complex. Thus, no (or very little) cardiac interference remains, but EMG activity is suppressed during these signal sections as well.

Figure 5.5 shows a box plot of ΔSNR and ΔPM for all algorithms under consideration. With regards to ΔSNR, the best median performance across all datasets is achieved by the EKS25, closely followed by the PATS algorithm. The PATS versions with different shape input signals perform practically identically. The next best algorithm according to median performance in this measure is the SWT, whereas the HP200, EMD, EKS2, and TS performing exhibit rather unfavorable results. Concerning ΔPM, the best results by far are obtained by the HP200 and the EMD, followed by the different versions of the PATS algorithm, and then the SWT. For this measure, differences in performance can be observed between the different PATS versions. While median and worst-case performance are similar across the different input signals, the optimal performance is markedly improved in the versions with shape input signals.

5.4.2 Synthetic data

Figure 5.6 shows the separation performance of the different algorithms assessed on the synthetic datasets as a function of the input signal's SNR. The newly proposed PATS algorithm achieves the best results at most SNR levels concerning all four measurements: e_{raw}, e_{env}, ΔSNR, and ΔPM. The TS and EKS2 algorithms also achieve good results concerning e_{raw}, whereas SWT generally appears superior to these algorithms concerning e_{env} (except at very high input SNR, where TS outperforms SWT). As expected, the EMD and HP200 algorithms are completely unable to recover the original, raw EMG signal, since they strongly filter the signals. To assess the relationship between the exact performance measures e_{raw} and e_{env} (which are only available in synthetic measurements) on the one hand, and the empirical performance measures ΔSNR and ΔPM (which can also be computed for real measuements) on the other hand, the correlation between these measures was analyzed for the synthetic datasets. Figure 5.7 shows scatter diagrams indicating the values these measures take for all 250 synthetic datasets, processed by all ten algorithms. It is apparent that there is a close correlation between ΔSNR and the two exact error measures, whereas the correlation to ΔPM is less pronounced. Spearman's coefficient of correlation was calculated, yielding values of -0.899 between e_{raw} and ΔSNR, -0.905 between e_{env} and ΔSNR, 0.298 between e_{raw} and ΔPM, and 0.4106 between e_{env} and ΔPM. These results confirm that ΔSNR is the far more reliable performance of the two, achieving high correlation with the exact performance measures e_{env} and e_{raw} across a large number of synthetic datasets.

5.5 Discussion & outlook

The contributions of this chapter are twofold. Firstly, a new algorithm for the removal of cardiac artifacts from surface EMG measurements of the respiratory muscles has been
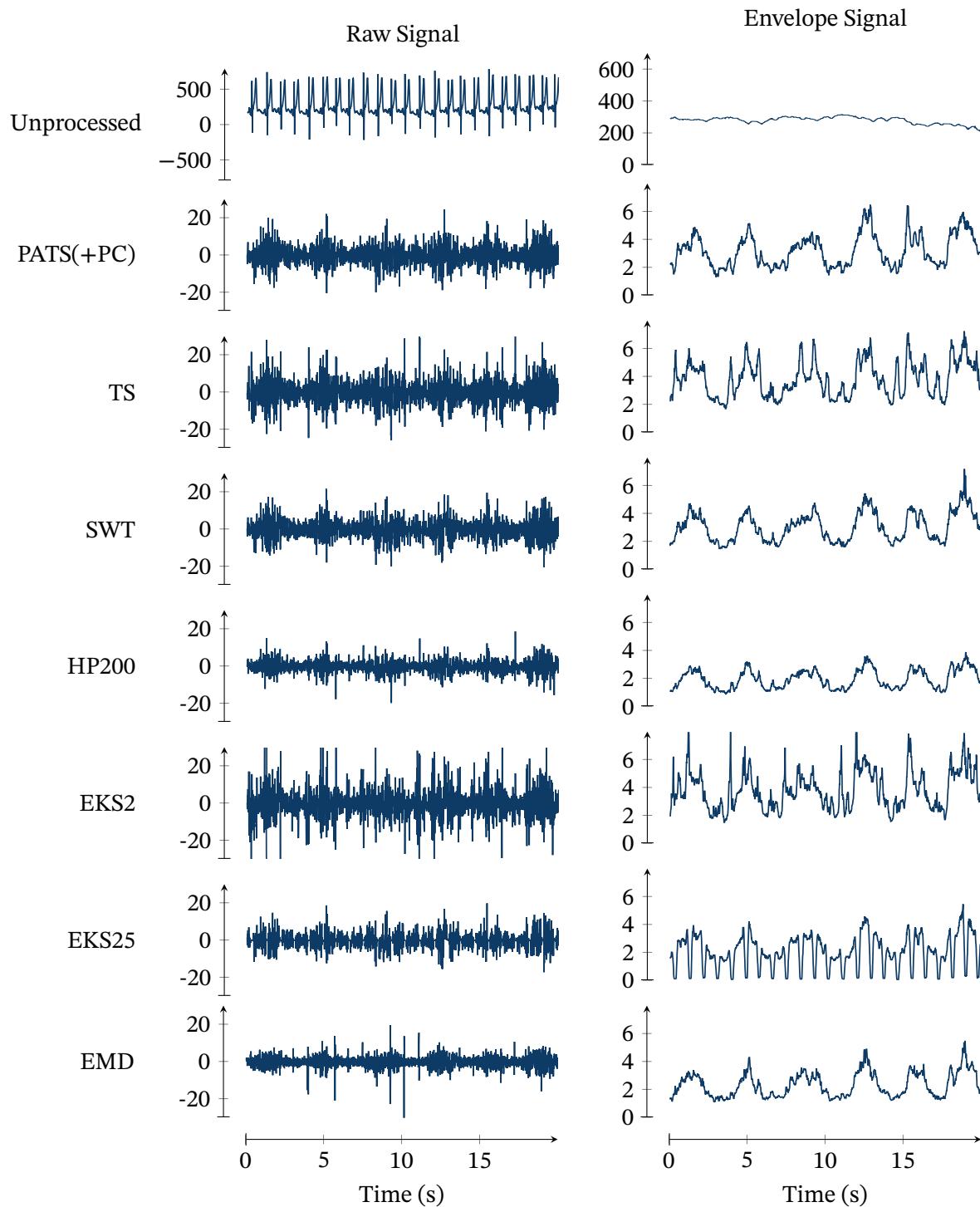
Figure 5.4 – A sample respiratory surface EMG signal snippet, raw except for powerline interference removal (top row) and processed by various cardiac artifact removal algorithms. The left column shows the raw signals, and the right column shows a 195 ms (corresponding to 200 samples) moving median of the absolute of the signal. (This type of envelope was empirically found to be more robust to signal noise than, e.g., a root mean square envelope or a mean absolute value envelope, see Petersen *et al.* [52].) For details on the algorithms refer to section 5.3.3. All signals are given in μV. The shown excerpt is from subject A/10, channel 1, of the study described in section 3.3.2. Algorithm performance varies strongly between measurements and patients.
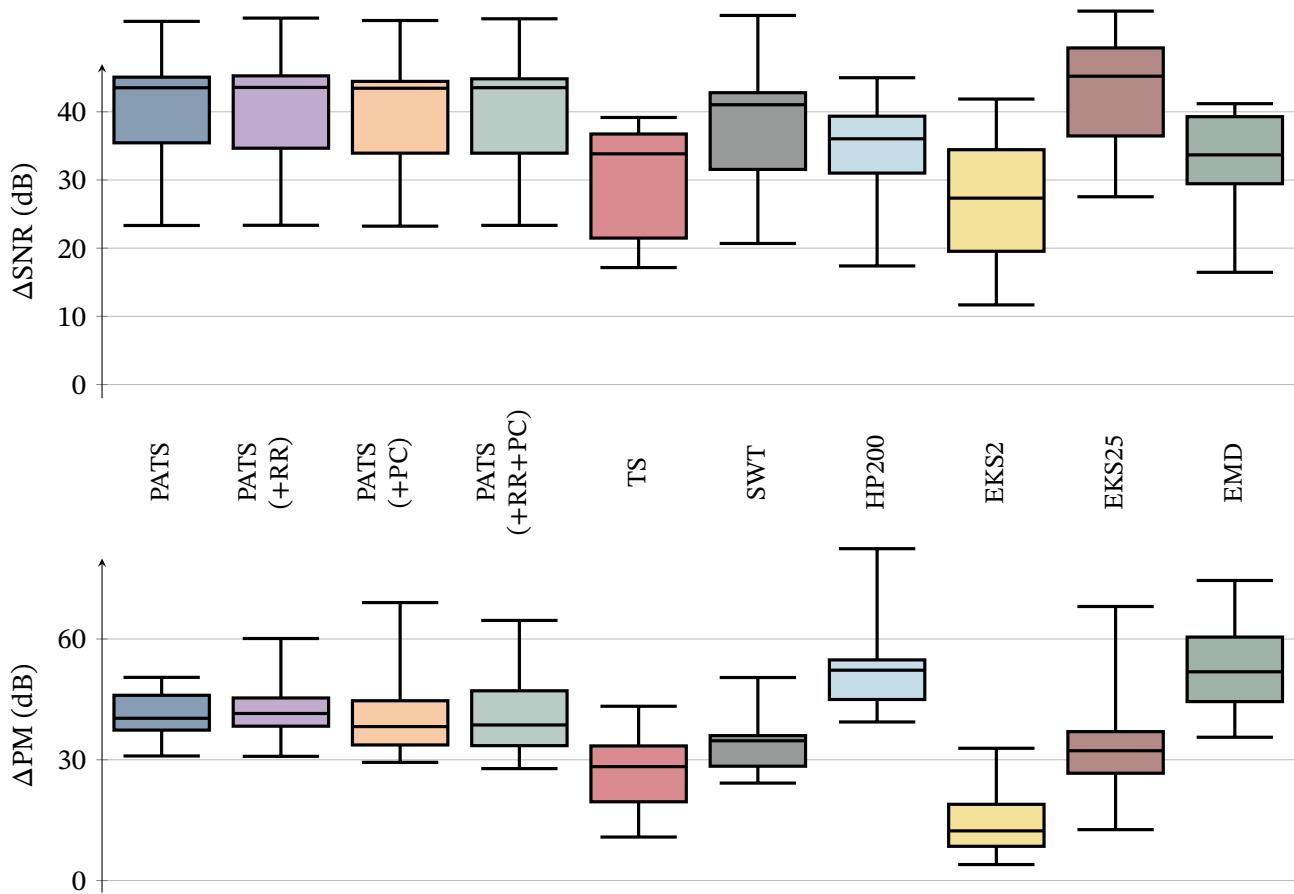
Figure 5.5 – Improvements in SNR and periodicity measure achieved by various algorithms for cardiac artifact removal on real respiratory surface EMG measurements. Standard boxplot: median and interquartile range; whiskers show the min–max range. Refer to section 5.3.3 for a details regarding the various algorithms, and to section 5.3.1 for the definitions of the two performance measures.

proposed. Secondly, a comprehensive framework for the performance evaluation of algorithms for the removal of such artifacts has been described.

PATS

The newly proposed algorithm, called *probabilistic adaptive template subtraction (PATS)* is fully probabilistic, nonparametric, and thus highly adaptive to arbitrary cardiac component waveforms, and explicitly models the temporal variability of subsequent artifacts. It is fully unsupervised and requires no additional tuning or calibration to be applied to a new signal. It does not require any additional measurements beyond the signal to be denoised. Since the method is agnostic to the particular shape of the cardiac component, it could also readily be applied to the removal of other types of periodic artifacts or the denoising of other signals beyond EMG measurements. Probabilistic uncertainty quantification is readily available as part of the method.

The new method was compared to a comprehensive array of previously proposed removal algorithms both in real as well as synthetic measurements. The synthetic measurements were composed of limb surface EMG measurements superposed with single-lead ECG
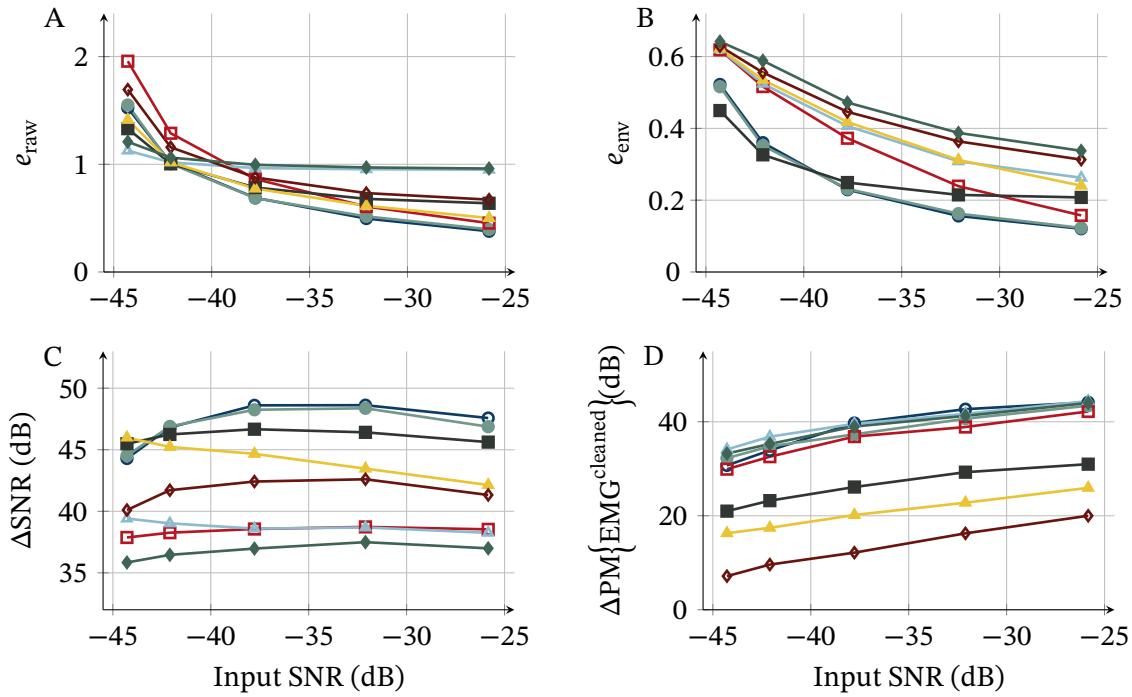
Figure 5.6 – Performance comparison of the different algorithms using synthetic data. (A) Relative pointwise error (see eq. (5.14)), (B) relative error of the EMG envelope (see eq. (5.16)), (C) SNR improvement (see eq. (5.10)) and (D) PM improvement (see eq. (5.13)) achieved by the different algorithms. The x-axis represents the input SNRs of the synthetic signals as defined by eq. (5.9), calculated on the 20 Hz pre-filtered signals. The y-axis represents the median of the respective performance measure over all 50 synthetic datasets at that particular SNR. The different lines show the results obtained by the PATS ($\circ$), PATS+PC ($\bullet$), TS ($\square$), SWT ($\blacksquare$), HP200 ($\triangle$), EKS2 ($\blacktriangle$), EKS25 ($\diamond$), and EMD ($\blacklozenge$) algorithms.

measurements. An interesting complication when creating synthetic measurements in this way is that single-lead cardiac signals are, of course, subject to noise originating (in part) from muscles near the recording sites, which may disturb the subsequent analysis because this constitutes an additional, unknown EMG signal source. For this reason, ECG signals with a high noise level have been rejected. As part of a Bachelor's thesis conducted under the supervision of the author [64], a study was performed in which two channels of respiratory surface EMG were recorded from ten subjects. In real measurements, a key challenge is that no ground truth is available. Thus, two empirical measures were used: an empirical SNR calculation, as well as the periodicity measurement (PM) originally proposed by Sameni *et al.* [59]. An analysis of the degree of correlation of these two empirical performance measures with exact error measures on the synthetic datasets indicates that the proposed empirical SNR measures is the far more reliable of the two performance measures. To the author's knowledge, the performance evaluation framework employed here is currently the most comprehensive framework described in the literature.

The performance results across all these data show that the newly proposed method performs very favorably and seems to be the most accurate method currently available for recovering the raw surface EMG signal from measurements suffering from strong cardiac
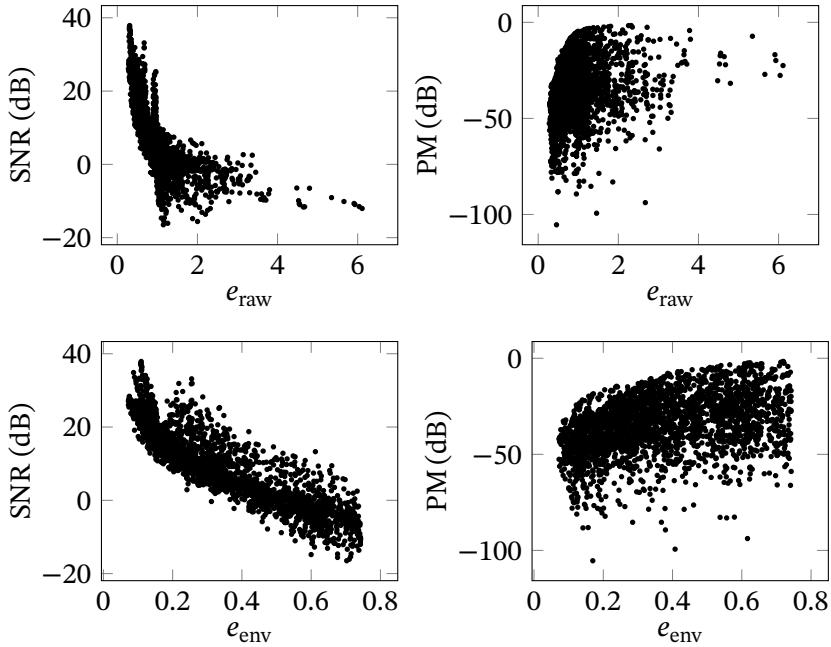
Figure 5.7 – Relation of the four discussed performance measures. Each of the 2500 data points in each graph represents one synthetic dataset (of which there are 250) processed by one of the ten algorithms under consideration.

disturbances. On the other hand, a clear result of the presented analysis is that an artifact removal algorithm must always be selected based on the requirements of the given application: if only a rough characterization of muscular activity is required, a simple high-pass filter with a high cut-off frequency and a subsequent envelope calculation will suffice. If, on the other hand, the precise characteristics of the raw EMG signal are of interest, a more sophisticated method — such as the proposed PATS algorithm — is required. The previously proposed model-based filtering procedures (the EKS2 and the EKS25) performed rather unfavorably in the performance analysis, despite significant effort having been invested into their tuning and improved robustness. The limitation of these techniques lies in the models they assume, which are only very rough approximations of cardiac beat shapes. It was thus, in many cases, impossible to tune the noise covariances such that the result is a clear separation between cardiac and myographic components. The new PATS algorithm, on the other hand, employs no such constrained model of beat morphology and thus does not suffer from this problem. Finally, in another recent performance comparison, Xu *et al.* [81] found TS to outperform many other algorithms (albeit not all of the algorithms considered here were included in their study). They did, however, only consider a single SNR level, whereas our analysis clearly shows that, firstly, SNR differs widely between recordings and, secondly, algorithm performance varies significantly as a function of SNR (see fig. 5.6).

It is, however, crucial to also note the limitations of the employed framework for performance evaluation. Both empirical measures (SNR and PM) have their drawbacks: amplitude suppression (e.g., during the QRS complex) and spectral changes are not

penalized by either of the two performance measures. This may potentially lead to misleadingly good performance measures for algorithms that effectively perform strong "gating" (i.e., suppression of the QRS segment to zero) or strong filtering. It appears likely that in the analysis performed, this affected the results for at least the EKS25, the HP200, as well as the EMD. Finally, the synthetic measurements here were based only on a small number of limb EMG signals, a drawback that could easily be remedied in the future.

The proposed PATS algorithm could be further improved in several ways. Firstly, as was already discussed in detail in section 5.2.6, the algorithm in its current form assumes that the process noise affecting subsequent samples is statistically independent. This assumption could be relaxed, potentially further increasing the algorithm's performance, albeit at the price of significantly increased computational complexity (see section 5.2.6). Secondly, an extension of the algorithm to treat multiple, concurrently occurring sources of disturbances might be of interest, e.g., to be robust to occasional extrasystoles, cardiac pacemaker artifacts, or recurring recording device artifacts. This might be achievable using an approach similar to sparse representation learning-based artifact removal [62, 63]. Thirdly, in many applications, multiple concurrently measured measurements may be available, as opposed to the single-channel approach pursued here. In such cases, it would appear attractive to exploit the additional information that is available from these further measurement channels. Effectively, such a separation problem could be framed as an adaptive blind deterministic multi-channel equalization problem, for which classical algorithms are available in the literature [8, 22, 23]. Such problems are nontrivial to solve, however, and care must be taken to not introduce strong modeling assumptions regarding the relationship between the different channels, as these might otherwise lead to a deterioration of separation performance and not to an improvement. (Recall the difficulties encountered due to the inadequacy of the modeling assumptions of the EKS2 and EKS25 approaches.)

Finally, it was already discussed in section 5.1.2 that the use of data-driven machine learning methods might yield important gains in separation performance in the near future, with Bockelmann *et al.* [13] having first demonstrated promising results in this area. The crucial challenge to solve for this approach concerns the availability of training data: since separation ground truth is unavailable for real measurements, other means of generating labeled training data must be pursued. One avenue might be the use of similar yet different measurement signals, such as performed by Bockelmann *et al.* [13], who used invasive diaphragm EMG measurements as a ground truth. Another possibility is the generation of realistic synthetic signals using comprehensive physiological simulation models such as the one described in the previous chapter. Using such simulated data, the separation ground truth is, of course, readily available and can be used for model training. The use of synthetic data for machine learning is a young field, however, and ripe with challenges [45]. More generally, a reliable machine learning approach for a medical application must overcome many obstacles, including model underspecification, shortcut learning, and inscrutability, to be reliably successful in a practical application [53].

References

[1] S. Abbaspour and A. Fallah, "A combination method for electrocardiogram rejection from surface electromyogram," *The Open Biomedical Engineering Journal*, vol. 8, no. 1, pp. 13–19, 2014. DOI: 10.2174/1874120701408010013.

[2] S. Abbaspour, A. Fallah, M. Lindén, and H. Gholamhosseini, "A novel approach for removing ECG interferences from surface EMG signals using a combined ANFIS and wavelet," *Journal of Electromyography and Kinesiology*, vol. 26, pp. 52–59, 2016. DOI: 10.1016/j.jelekin. 2015.11.003.

[3] S. Abbaspour, M. Lindén, and H. Gholamhosseini, "ECG artifact removal from surface EMG signal using an automated method based on wavelet-ICA," in *pHealth*, 2015. [Online]. Available: https://ebooks.iospress.nl/publication/39241.

[4] M. Abramson *et al.*, *The NOMAD project*. [Online]. Available: https://www.gerad.ca/nomad/.

[5] M. Adnane and A. Belouchrani, "Heartbeats classification using QRS and T waves autoregressive features and RR interval features," *Expert Systems*, vol. 34, no. 6, e12219, 2017. DOI: 10.1111/exsy.12219.

[6] M. Akhbari *et al.*, "ECG denoising and fiducial point extraction using an extended Kalman filtering framework with linear and nonlinear phase observations," *Physiological Measurement*, vol. 37, p. 203, 2 2016. DOI: 10.1088/0967-3334/37/2/203.

[7] P. Akkiraju and D. Reddy, "Adaptive cancellation technique in processing myoelectric activity of respiratory muscles," *IEEE Transactions on Biomedical Engineering*, vol. 39, no. 6, pp. 652–655, 1992. DOI: 10.1109/10.141204.

[8] S. Amari, S. Douglas, A. Cichocki, and H. Yang, "Multichannel blind deconvolution and equalization using the natural gradient," in *First IEEE Signal Processing Workshop on Signal Processing Advances in Wireless Communications*, IEEE, 1997. DOI: 10.1109/spawc.1997. 630083.

[9] K. Antczak, "Deep recurrent neural networks for ECG signal denoising," 2018. arXiv: 1807. 11551 [cs.NE].

[10] C. T. Arsene, R. Hankins, and H. Yin, "Deep learning models for denoising ECG signals," in *2019 27th European Signal Processing Conference (EUSIPCO)*, IEEE, 2019. DOI: 10.23919/eusipco.2019.8902833.

[11] A. Bartolo, C. Roberts, R. R. Dzwonczyk, and E. Goldman, "Analysis of diaphragm EMG signals: Comparison of gating vs. subtraction for removal of ECG contamination," *Journal of Applied Physiology: Respiratory, Environmental and Exercise Physiology*, vol. 80, no. 6, pp. 1898–902, 1996. DOI: 10.1152/jappl.1996.80.6.1898.

[12] R. Bloch, "Subtraction of electrocardiographic signal from respiratory electromyogram," *Journal of Applied Physiology: Respiratory, Environmental and Exercise Physiology*, vol. 55, no. 2, pp. 619–623, 1983. DOI: 10.1152/jappl.1983.55.2.619.

[13] N. Bockelmann *et al.*, "Deep learning for prediction of diaphragm activity from the surface electromyogram," *Current Directions in Biomedical Engineering*, vol. 5, no. 1, pp. 17–20, 2019. DOI: 10.1515/cdbme-2019-0005.

[14] R. H. Byrd, M. E. Hribar, and J. Nocedal, "An interior point algorithm for large-scale nonlinear programming," *SIAM Journal on Optimization*, vol. 9, no. 4, pp. 877–900, 1999. DOI: 10.1137/s1052623497325107.

[15] P. H. Charlton *et al.*, "An assessment of algorithms to estimate respiratory rate from the electrocardiogram and photoplethysmogram," *Physiological Measurement*, vol. 37, no. 4, pp. 610–626, 2016. DOI: 10.1088/0967-3334/37/4/610.

[16] A. D. Chave, "A multitaper spectral estimator for time-series with missing data," *Geophysical Journal International*, vol. 218, no. 3, pp. 2165–2178, 2019. DOI: 10.1093/gji/ggz280.

[17] E. A. Clancy and N. Hogan, "Probability density of the surface electromyogram and its relation to amplitude detectors," *IEEE Transactions on Biomedical Engineering*, vol. 46, no. 6, pp. 730–739, 1999. DOI: 10.1109/10.764949.

[18] Y. Deng, W. Wolf, R. Schnell, and U. Appel, "New aspects to event-synchronous cancellation of ECG interference: An application of the method in diaphragmatic EMG signals," *IEEE Transactions on Biomedical Engineering*, vol. 47, no. 9, pp. 1177–1184, 2000. DOI: 10.1109/10.867924.

[19] N. A. Dimitrova, G. V. Dimitrov, and O. A. Nikitin, "Neither high-pass filtering nor mathematical differentiation of the EMG signals can considerably reduce cross-talk," *Journal of Electromyography and Kinesiology*, vol. 12, no. 4, pp. 235–246, 2002. DOI: 10.1016/s1050-6411(02)00008-1.

[20] J. D. Drake and J. P. Callaghan, "Elimination of electrocardiogram contamination from electromyogram signals: An evaluation of currently used removal techniques," *Journal of Electromyography and Kinesiology*, vol. 16, no. 2, pp. 175–187, 2006. DOI: 10.1016/j.jelekin.2005.07.003.

[21] E.-M. Fadaili *et al.*, "Extration of Diaphragmatic EMG Using DWT," GIPSA-lab, Research Report, 2008, Rapport interne de GIPSA-lab. [Online]. Available: https://hal.archives-ouvertes.fr/hal-00202737.

[22] D. Gesbert, P. Duhame, and S. Mayrargue, "On-line blind multichannel equalization based on mutually referenced filters," *IEEE Transactions on Signal Processing*, vol. 45, no. 9, pp. 2307–2317, 1997. DOI: 10.1109/78.622953.

[23] G. Giannakis and C. Tepedelenlioglu, "Direct blind equalizers of multiple FIR channels: A deterministic approach," *IEEE Transactions on Signal Processing*, vol. 47, no. 1, pp. 62–74, 1999. DOI: 10.1109/78.738240.

[24] A. L. Goldberger *et al.*, "PhysioBank, PhysioToolkit, and PhysioNet," *Circulation*, vol. 101, no. 23, e215–e220, 2000. DOI: 10.1161/01.cir.101.23.e215.

[25] J. Graßhoff, E. Petersen, and P. Rostalski, "Removing strong ECG interference from EMG measurements," in *Proceedings of the Workshop Biosignals*, Erfurt, Germany, 2018.

[26] J. Graßhoff *et al.*, "A template subtraction method for the removal of cardiogenic oscillations on esophageal pressure signals," in *39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Jeju Island, Korea, 2017. DOI: 10.1109/EMBC.2017.8037299.

[27] G. Gratton and P. M. Corballis, "Removing the heart from the brain: Compensation for the pulse artifact in the photon migration signal," *Psychophysiology*, vol. 32, no. 3, pp. 292–299, 1995. DOI: 10.1111/j.1469-8986.1995.tb02958.x.

[28] J. Guan *et al.*, "Automated dynamic electrocardiogram noise reduction using multilayer LSTM network," in *Proceedings of the 15th EAI International Conference on Mobile and Ubiquitous Systems: Computing, Networking and Services*, ACM, 2018. DOI: 10.1145/3286978.3286988.

[29] S. E. Jero, P. Ramu, and S. Ramakrishnan, "Discrete wavelet transform and singular value decomposition based ECG steganography for secured patient information transmission," *Journal of Medical Systems*, vol. 38, no. 10, 2014. DOI: 10.1007/s10916-014-0132-z.

[30] M. A. Kabir and C. Shahnaz, "Denoising of ECG signals based on noise reduction algorithms in EMD and wavelet domains," *Biomedical Signal Processing and Control*, vol. 7, no. 5, pp. 481–489, 2012. DOI: 10.1016/j.bspc.2011.11.003.

[31] L. Kahl, E. Petersen, J. Graßhoff, and P. Rostalski, "Method and device for determining a respiratory or cardiogenic signal," DE112020000232A5 / WO2021063601A1, 2019.

[32] P. Kanjilal, S. Palit, and G. Saha, "Fetal ECG extraction from single-channel maternal ECG using singular value decomposition," *IEEE Transactions on Biomedical Engineering*, vol. 44, no. 1, pp. 51–59, 1997. DOI: 10.1109/10.553712.

[33] M. Korürek and A. Nizam, "A new arrhythmia clustering technique based on ant colony optimization," *Journal of Biomedical Informatics*, vol. 41, no. 6, pp. 874–881, 2008. DOI: 10.1016/j.jbi.2008.01.014.

[34] D. Labate *et al.*, "Empirical mode decomposition vs. wavelet decomposition for the extraction of respiratory signal from single-channel ECG: A comparison," *IEEE Sensors Journal*, vol. 13, no. 7, pp. 2666–2674, 2013. DOI: 10.1109/jsen.2013.2257742.

[35] P. Laguna *et al.*, "Adaptive filter for event-related bioelectric signals using an impulse correlated reference input: Comparison with signal averaging techniques," *IEEE Transactions on Biomedical Engineering*, vol. 39, no. 10, pp. 1032–1044, 1992. DOI: 10.1109/10.161335.

[36] R. Lansing and J. Savelle, "Chest surface recording of diaphragm potentials in man," *Electroencephalography and Clinical Neurophysiology*, vol. 72, no. 1, pp. 59–68, 1989. DOI: 10.1016/0013-4694(89)90031-x.

[37] S. Le Digabel, "Algorithm 909: NOMAD: Nonlinear optimization with the MADS algorithm," *ACM Transactions on Mathematical Software*, vol. 37, no. 4, pp. 1–15, 2011. DOI: 10.1145/1916461.1916468.

[38] S. Levine *et al.*, "Description and validation of an ECG removal procedure for EMGdi power spectrum analysis," *Journal of Applied Physiology: Respiratory, Environmental and Exercise Physiology*, vol. 60, no. 3, pp. 1073–1081, 1986. DOI: 10.1152/jappl.1986.60.3.1073.

[39] C.-C. Lin and C.-M. Yang, "Heartbeat classification using normalized RR intervals and morphological features," *Mathematical Problems in Engineering*, vol. 2014, pp. 1–11, 2014. DOI: 10.1155/2014/712474.

[40] L. Ljung, *System identification toolbox: User's guide*, The MathWorks, Inc. [Online]. Available: https://www.mathworks.com/help/pdf_doc/ident/ident_ug.pdf.

[41] G. Lu *et al.*, "Removing ECG noise from surface EMG signals using adaptive filtering," *Neuroscience Letters*, vol. 462, pp. 14–19, 2009. DOI: 10.1016/j.neulet.2009.06.063.

[42] J. N. Mak, Y. Hu, and K. D. Luk, "An automated ECG-artifact removal method for trunk muscle surface EMG recordings," *Medical Engineering & Physics*, vol. 32, no. 8, pp. 840–848, 2010. DOI: 10.1016/j.medengphy.2010.05.007.

[43] M. Mañanas *et al.*, "Cardiac interference in myographic signals from different respiratory muscles and levels of activity," in *Conference Proceedings of the 23rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, IEEE, 2001. DOI: 10.1109/iembs.2001.1020386.

[44] P. E. McSharry, G. D. Clifford, L. Tarassenko, and L. A. Smith, "A dynamical model for generating synthetic electrocardiogram signals," *IEEE Transactions on Biomedical Engineering*, 2003. DOI: 10.1109/TBME.2003.808805.

[45] S. I. Nikolenko, *Synthetic Data for Deep Learning*. Springer International Publishing, 2021. DOI: 10.1007/978-3-030-75178-4.

[46] O. Pahlm and L. Sörnmo, "Data processing of exercise ECG's," *IEEE Transactions on Biomedical Engineering*, vol. 34, no. 2, pp. 158–165, 1987. DOI: 10.1109/tbme.1987.326040.

[47] J. Pan and W. J. Tompkins, "A real-time QRS detection algorithm," *IEEE Transactions on Biomedical Engineering*, vol. 32, no. 3, pp. 230–236, 1985. DOI: 10.1109/TBME.1985.325532.

[48] E. Peri *et al.*, "Singular value decomposition for removal of cardiac interference from trunk electromyogram," *Sensors*, vol. 21, no. 2, p. 573, 2021. DOI: 10.3390/s21020573.

[49] E. Petersen, "A mathematical model of surface electromyographic measurements," in *Proceedings of the Workshop Biosignal Processing*, Berlin, 2016.

[50] E. Petersen, H. Buchner, M. Eger, and P. Rostalski, "Convolutive blind source separation of surface EMG measurements of the respiratory muscles," *Biomedical Engineering - Biomedizinische Technik*, vol. 62, no. 2, pp. 171–181, 2017. DOI: 10.1515/bmt-2016-0092.

[51] E. Petersen and P. Rostalski, "A comprehensive mathematical model of motor unit pool organization, surface electromyography, and force generation," *Frontiers in Physiology*, vol. 10, p. 176, 2019, ISSN: 1664-042X. DOI: 10.3389/fphys.2019.00176.

[52] E. Petersen, J. Sauer, J. Graßhoff, and P. Rostalski, "Removing cardiac artifacts from single-channel respiratory electromyograms," *IEEE Access*, vol. 8, pp. 30 905–30 917, 2020. DOI: 10.1109/access.2020.2972731.

[53] E. Petersen *et al.*, "Responsible and regulatory conform machine learning for medicine: A survey of challenges and solutions," *IEEE Access*, vol. 10, pp. 58 375–58 418, 2022. DOI: 10.1109/access.2022.3178382.

[54] R. Plonsey and R. C. Barr, *Bioelectricity, A Quantitative Approach*, third edition. Springer, 2007.

[55] H. E. Rauch, F. Tung, and C. T. Striebel, "Maximum likelihood estimates of linear dynamic systems," *AIAA Journal*, vol. 3, no. 8, pp. 1445–1450, 1965. DOI: 10.2514/3.3166.

[56] M. S. Redfern, R. E. Hughes, and D. B. Chaffin, "High-pass filtering to remove electrocardiographic interference from torso EMG recordings," *Clinical Biomechanics*, vol. 8, pp. 44–48, 1993. DOI: 10.1016/S0268-0033(05)80009-9.

[57] L. Rodríguez-Liñares and D. Simpson, "Spectral estimation of HRV in signals with gaps," *Biomedical Signal Processing and Control*, vol. 52, pp. 187–197, 2019. DOI: 10.1016/j.bspc.2019.04.006.

[58] R. Sameni, *OSET: The open-source electrophysiological toolbox*, version 3.14, Released under the GNU General Public License, 2006. [Online]. Available: www.oset.ir.

[59] R. Sameni, M. Shamsollahi, and C. Jutten, "Model-based Bayesian filtering of cardiac contaminants from biomedical recordings," *Physiological Measurement*, vol. 29, no. 5, pp. 595–613, 2008. DOI: 10.1088/0967-3334/29/5/006.

[60] R. Sameni, M. B. Shamsollahi, C. Jutten, and G. D. Clifford, "A nonlinear Bayesian filtering framework for ECG denoising," *IEEE Transactions on Biomedical Engineering*, vol. 54, no. 12, pp. 2172–2185, 2007. DOI: 10.1109/TBME.2007.897817.

[61] S. Särkkä, *Bayesian Filtering and Smoothing*. Cambridge University Press, Cambridge, United Kingdom, 2013, ISBN: 978-1-107-03065-7. [Online]. Available: https://users.aalto.fi/~ssarkka/pub/cup_book_online_20131111.pdf.

[62] U. Satija, B. Ramkumar, and M. S. Manikandan, "A unified sparse signal decomposition and reconstruction framework for elimination of muscle artifacts from ECG signal," in *2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, IEEE, 2016. DOI: 10.1109/icassp.2016.7471781.

[63] ——, "Noise-aware dictionary-learning-based sparse representation framework for detection and removal of single and combined noises from ECG signal," *Healthcare Technology Letters*, vol. 4, no. 1, pp. 2–12, 2017. DOI: 10.1049/htl.2016.0077.

[64]   J. Sauer, "Cardiac artefact removal on surface EMG signals of the respiratory muscles," Bachelor's Thesis, Universität zu Lübeck, 2016.

[65]   T. Schuessler *et al.*, "An adaptive filter for the reduction of cardiogenic oscillations on esophageal pressure signals," in *Proceedings of 17th International Conference of the Engineering in Medicine and Biology Society*, IEEE, 1995. DOI: 10.1109/iembs.1995.579253.

[66]   T. F. Schuessler *et al.*, "An adaptive filter to reduce cardiogenic oscillations on esophageal pressure signals," *Annals of Biomedical Engineering*, vol. 26, no. 2, pp. 260–267, 1998. DOI: 10.1114/1.55.

[67]   V. Seemann, "Adaptive Kalman filter-based cardiogenic artifact removal from surface electromyograms of respiratory muscles," M.S. thesis, Universität zu Lübeck, 2020.

[68]   D. Simon, "Kalman filtering with state constraints: A survey of linear and nonlinear algorithms," *IET Control Theory & Applications*, vol. 4, no. 8, pp. 1303–1318, 2010. DOI: 10.1049/iet-cta.2009.0032.

[69]   D. Simon, *Optimal State Estimation*. Hoboken, New Jersey, USA: John Wiley & Sons, Inc., 2006.

[70]   M. L. Simoons and P. G. Hugenholtz, "Gradual changes of ECG waveform during and after exercise in normal subjects," *Circulation*, vol. 52, pp. 570–577, 1975. DOI: 10.1161/01.CIR.52.4.570.

[71]   Y. Slim and K. Raoof, "Removal of ECG interference from surface respiratory electromyography," *IRBM*, vol. 31, pp. 209–220, 2010. DOI: 10.1016/j.irbm.2010.05.002.

[72]   P. Strobach, K. Abraham-Fuchs, and W. Harer, "Event-synchronous cancellation of the heart interference in biomedical signals," *IEEE Transactions on Biomedical Engineering*, vol. 41, no. 4, pp. 343–350, 1994. DOI: 10.1109/10.284962.

[73]   N. Thakor and Y.-S. Zhu, "Applications of adaptive filtering to ECG analysis: Noise cancellation and arrhythmia detection," *IEEE Transactions on Biomedical Engineering*, vol. 38, no. 8, pp. 785–794, 1991. DOI: 10.1109/10.83591.

[74]   V. von Tscharner, B. Eskofier, and P. Federolf, "Removal of the electrocardiogram signal from surface EMG recordings using non-linearly scaled wavelets," *Journal of Electromyography and Kinesiology*, vol. 21, no. 4, pp. 683–688, 2011. DOI: 10.1016/j.jelekin.2011.03.004.

[75]   M. Ungureanu and W. Wolf, "Basic aspects concerning the event-synchronous interference canceller," *IEEE Transactions on Biomedical Engineering*, vol. 53, no. 11, pp. 2240–2247, 2006. DOI: 10.1109/tbme.2006.877119.

[76]   R. Vullings, B. de Vries, and J. W. M. Bergmans, "An adaptive Kalman filter for ECG signal enhancement," *IEEE Transactions on Biomedical Engineering*, vol. 58, no. 4, pp. 1094–1103, 2011. DOI: 10.1109/tbme.2010.2099229.

[77]   S. de Waele and P. Broersen, "The Burg algorithm for segments," *IEEE Transactions on Signal Processing*, vol. 48, no. 10, pp. 2876–2880, 2000. DOI: 10.1109/78.869039.

[78]   J.-J. Wei, C.-J. Chang, N.-K. Chou, and G.-J. Jan, "ECG data compression using truncated singular value decomposition," *IEEE Transactions on Information Technology in Biomedicine*, vol. 5, no. 4, pp. 290–299, 2001. DOI: 10.1109/4233.966104.

[79]   B. Weng, M. Blanco-Velasco, and K. E. Barner, "ECG denoising based on the empirical mode decomposition," in *2006 International Conference of the IEEE Engineering in Medicine and Biology Society*, IEEE, 2006. DOI: 10.1109/iembs.2006.259340.

[80] N. W. Willigenburg, A. Daffertshofer, I. Kingma, and J. H. van Dieën, "Removing ECG contamination from EMG recordings: A comparison of ICA-based and other filtering procedures," *Journal of Electromyography and Kinesiology*, vol. 22, pp. 485–93, 3 2012. DOI: 10.1016/j.jelekin.2012.01.001.

[81] L. Xu *et al.*, "Comparative review of the algorithms for removal of electrocardiographic interference from trunk electromyography," *Sensors*, vol. 20, no. 17, p. 4890, 2020. DOI: 10.3390/s20174890.

[82] H. Zhang *et al.*, "A novel convolutional neural network model to remove muscle artifacts from EEG," 2020. arXiv: 2010.11709 [eess.SP].

[83] P. Zhou and T. A. Kuiken, "Eliminating cardiac contamination from myoelectric control signals developed by targeted muscle reinnervation," *Physiological Measurement*, vol. 27, pp. 1311–1327, 2006. DOI: 10.1088/0967-3334/27/12/005.

# PARAMETER ESTIMATION UNDER CONCEPT DRIFT, COVARIATE SHIFT, AND MODEL MISMATCH
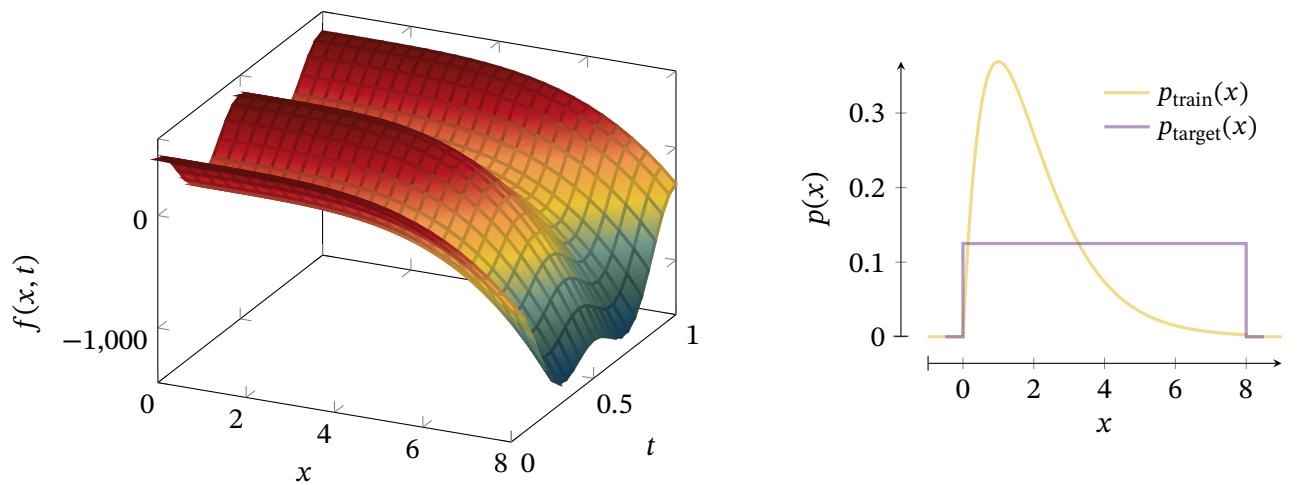
Figure 6.1 – Left: a nonlinear function $f(x, t)$, changing over time and thus exhibiting *concept drift*. **Right:** training data are drawn from $p_{\text{train}}(x)$, but the model should perform well on data drawn from $p_{\text{target}}(x)$: there is *covariate shift*. How can one identify a simple, time-varying model — say, a linear one — that minimizes the expected deviation from the true function under the target distribution $p_{\text{target}}(x)$ at each point in time?

Contents

WHEREAS the previous two chapters have treated two principal challenges in recovering clean surface EMG signals from real-world respiratory measurements, the present chapter and the following chapter 7 shift attention to the identification of respiratory model parameters using such denoised sEMG signals. This practical identification problem poses a number of crucial challenges to an estimation procedure[1]:

Time-varying model parameters — patients' respiratory properties change over time due to disease progression or recovery, change of positioning, and other factors.

Inevitable modeling errors — the respiratory system is highly complex, many of its properties are unobservable using the available measurements, and there are many unpredictable and unknown disturbances.

Covariate shift — respiratory time series recordings are often in large parts highly repetitive due to fixed ventilator control schemes, and only rarely do ventilator maneuvers or other events further explore the full space over which a patient model should be accurate.

The combination of these properties substantially complicates the parameter identification and necessitates the development of new estimation techniques.

The present chapter focuses on the theoretical development of these estimation techniques. Firstly, a novel weighted Kalman filter and smoother are derived, as well as a weighted maximum likelihood method to identify the parameters of the underlying state-space model. Secondly, a simple yet accurate and computationally efficient method for multivariate density estimation is proposed, which is employed to calculate the weights for the importance-weighted parameter estimation procedure. The adequacy of the combination of these two methods to solve importance-weighted, time-varying parameter estimation problems is demonstrated using two synthetic estimation problems. In the following chapter 7, these techniques will be employed to solve the challenges of the respiratory parameter identification problem.[2]

6.1 Problem definition & state of the art

Section 6.1.1 defines the estimation problem under consideration in this chapter; section 6.1.2 briefly surveys the state of the art concerning this estimation problem.

---

[1] For a more detailed discussion of these challenges in the context of sEMG-based respiratory model parameter identification, refer to the following chapter 7.

[2] None of the developments presented in this chapter have been previously published elsewhere.

6.1.1 Problem definition

In section 2.1.6, static weighted likelihood maximization problems of the form

$$\hat{\theta} = \arg\max_{\theta} E_{p_{\text{target}}}[\log q(y \mid x, \theta)]$$

$$= \arg\max_{\theta} E_{p_{\text{train}}}\left[\frac{p_{\text{target}}(x)}{p_{\text{train}}(x)} \log q(y \mid x, \theta)\right]$$

$$\approx \arg\max_{\theta} \frac{1}{N^S} \sum_{k=1}^{N^S} \frac{p_{\text{target}}(x_k)}{p_{\text{train}}(x_k)} \log q(y_k \mid x_k, \theta)$$

were discussed, where $q(y \mid x, \theta)$ is an incorrectly specified model of $p(y \mid x)$ (in other words, model mismatch is present), $p_{\text{target}}(x) \not\equiv p_{\text{train}}(x)$ (there is covariate shift), and x is the target variable. The statistical estimation problem considered in this chapter represents a *dynamic* extension of the above: in addition to model mismatch and covariate shift, we will assume that *concept drift* is present, i.e., the conditional likelihood $p(y \mid x)$ is time-varying. Specifically, we will aim to identify an optimal time-varying parameter trajectory

Concept Drift

$$\hat{\theta}_{0:N^S} = \arg\max_{\theta_{0:N^S}} E_{p_{\text{target}}(x)\times\text{Unif}(k)}[\log q(y_k \mid x_k, \theta_k) + \log q(\theta_k \mid \theta_{k-1})] + \frac{1}{N^S} \log q(\theta_0)$$

$$= \arg\max_{\theta_{0:N^S}} E_{p_{\text{train}}(x)\times\text{Unif}(k)}\left[\underbrace{\frac{p_{\text{target}}(x_k)}{p_{\text{train}}(x_k)}}_{=:w_k} \log q(y_k \mid x_k, \theta_k) + \log q(\theta_k \mid \theta_{k-1})\right] + \frac{1}{N^S} \log q(\theta_0)$$

$$\approx \arg\max_{\theta_{0:N^S}} \underbrace{\frac{1}{N^S} \sum_{k=1}^{N^S} w_k \log q(y_k \mid x_k, \theta_k)}_{\substack{=-\,\mathcal{R}_{\text{emp}}(\theta) \\ \text{Expected log likelihood} \\ \text{of measurements} \\ \text{over target distribution}}} + \underbrace{\frac{1}{N^S} \sum_{k=1}^{N^S} \log q(\theta_k \mid \theta_{k-1}) + \frac{1}{N^S} \log q(\theta_0)}_{\substack{=-\,\ell_{\text{reg}}(\theta) \\ \text{Expected log likelihood of parameter trajectory}}}, \quad (6.1)$$

where $q(\theta_k \mid \theta_{k-1})$ specifies a model of the concept drift, and $q(\theta_0)$ a prior on the initial set of parameter values. The distributions $p_{\text{target}}(x)$ and $p_{\text{train}}(x)$ are *not* assumed to be time-varying here, i.e., the covariate shift is assumed to be static. This assumption is not essential, however, and could be dropped.

Structural Importance-weighted Empirical Risk Minimization

Equation (6.1) represents a *structural importance-weighted empirical risk minimization* problem (see sections 2.1.2 and 2.1.6 for structural risk minimization), where the first part represents the importance-weighted empirical loss $\mathcal{R}_{\text{emp}}(\theta)$, and the second part a regularization term $\ell_{\text{reg}}(\theta)$. Incidentally, if the measurement model $q(y \mid x, \theta)$ is linear in θ, and if the concept drift model can be represented as a linear state-space model, then eq. (6.1) is exactly equivalent to the optimization problem solved by a classical, linear Kalman smoother (see eqs. (2.27) and (2.28)) — except for the sample weights w_k. This near-equivalence will be exploited later on in this chapter. Differently from the existing covariate shift literature, we will *not* treat the distribution $p_{\text{target}}(x)$ as representative of some real, empirical distribution, but rather use it as a *tuning parameter* to control which regions of the data space should be described well by the model, akin to frequency

weighting in system identification [31, 42]. Finally, we will also discuss a principled way to estimate the (hyper-)parameters ψ of the concept drift model $q(\theta_k \mid \theta_{k-1})$.

To summarize, the problem considered in this chapter is to estimate linear model parameters in the face of a time-varying system, model mismatch, and covariate shift.

6.1.2 State of the art

The problem of performing estimation in the face of differing training and target distributions has received much attention in the statistics and machine learning communities. Starting in 1952 with the seminal publication of Horvitz and Thompson [20], statisticians have used inverse probability weighting to account for confounding factors such as sampling biases or missing data in classification [20, 24, 34, 39, 40, 50]. These problems are especially prevalent in medical studies because of low disease prevalence and treatment assignment based on disease status, among other factors [34, 39, 50].

In their highly influential publication, Shimodaira [45] have described a general framework for regression under different training and target distributions $p_{\text{train}}(x)$ and $p_{\text{target}}(x)$ (a situation they call *covariate shift*), inciting a surge of research under the label *covariate shift adaptation* [17, 21, 38, 41, 45, 47, 49, 52]. One of the key observations made by Shimodaira [45] is that

Covariate Shift Adaptation

$$\mathrm{E}_{p_{\text{target}}}[\ell(x;\theta)] = \mathrm{E}_{p_{\text{train}}}\left[\frac{p_{\text{target}}(x)}{p_{\text{train}}(x)}\ell(x;\theta)\right]$$

under the assumption $\mathrm{supp}(p_{\text{target}}) \subset \mathrm{supp}(p_{\text{train}})$, and hence the estimate

$$\hat{\theta} = \arg\min_{\theta} \mathrm{E}_{p_{\text{target}}}[\ell(x;\theta)] = \arg\min_{\theta} \mathrm{E}_{p_{\text{train}}}\left[\frac{p_{\text{target}}(x)}{p_{\text{train}}(x)}\ell(x;\theta)\right]$$

can be determined by means of weighted risk minimization using the original training data, where the weights $w(x)$ are chosen as

$$w(x) = \frac{p_{\text{target}}(x)}{p_{\text{train}}(x)}. \tag{6.2}$$

The density ratio $p_{\text{target}}(x)/p_{\text{train}}(x)$ is often called *importance* [8, 17, 23, 49, 52]. If $\ell(x;\theta)$ is the quadratic loss function, the resulting estimator is the weighted least squares (WLS) estimator discussed in section 2.2.2. The importance weighted estimator is asymptotically bias-free, but in general not efficient [24, 45]): for large sample sizes, the optimal weights are indeed given by equation (6.2) as described above, while for the other extreme case where the model is correctly specified, the optimal weights are given by $w(x) \equiv 1$, i.e., OLS in the case of the quadratic loss function [45]. To facilitate weight selection in practical applications, Shimodaira [45] therefore proposes to consider weights obtained by continuous "interpolation" between the two extreme cases $w(x) \equiv 1$ and $w(x) = p_{\text{target}}(x)/p_{\text{train}}(x)$ by choosing $w^*(x) = w(x)^{\beta}$ with $0 \le \beta \le 1$.

Importance

As a brief remark, notice that the problem of regression or classification under covariate shift is distinct from (though related to) what is called *imbalanced* classification [5, 6, 11, 13]. In the case of covariate shift, the concern is that $p_{\text{train}}(x)$ and $p_{\text{target}}(x)$ differ, whereas

Dataset Imbalance

in the imbalanced classification case, the concern is that an imbalanced *outcome* training distribution

$$p_{\text{train}}(y) = \int p(y \mid x)\, p_{\text{train}}(x)\, \mathrm{d}x$$

may prove problematic in some sense. There does not seem to be a fundamental reason for why such an imbalance in $p_{\text{train}}(y)$ would cause problems, though, as long as covariate shift is not the underlying reason, and as long as a *proper scoring rule* is used [14, 15, 53]. (Most notably, this excludes classification accuracy as a scoring rule.)

Returning to importance-weighted regression, it is a fundamental feature of this method that some samples will be assigned (potentially much) larger weights than others. This can be easily seen to increase *estimator variance* and dependence of the final estimate on just a few influential examples, a problem well-known in the literature [8]. Various remedies for this problem have been proposed, including exponential weight flattening as proposed by Shimodaira [45], weight trimming [28], and stratification of samples into few, discrete categories for which average weights are then computed [7, 24]. Generally, it is well known that the success of such a weighted estimation scheme hinges in large parts on the correctness of the weight specification [8, 24, 28]. Importance weighting is structurally identical to many robust estimation schemes: both solve a weighted likelihood maximization problem; only the weight selection strategies differ [12, 33, 45, 55]. Whereas robust estimators reduce the weights on (outlier) samples that do not conform with the specified model, importance weighting reduces the weight on frequently observed sample types and increases the weight on rarely observed sample types. The two strategies are orthogonal to each other: both frequently occurring and rare samples may conform well or badly with the specified model.

In practice, a crucial question is how to obtain the importance $p_{\text{target}}(x)/\, p_{\text{train}}(x)$, which is assumed to be known by [45], but which is unavailable in many applications. A naive approach would be to use any parametric or non-parametric density estimation method for estimating the two densities $p_{\text{target}}(x)$ and $p_{\text{train}}(x)$. However, density estimation is a hard problem, especially in high-dimensional spaces with a large number of samples, and hence many methods have been proposed for *direct importance estimation* without estimating the two densities individually ([23, 48, 49, 52]). Two popular algorithms are the Kullback-Leibler Importance Estimation Procedure (KLIEP, see Sugiyama *et al.* [48, 49]) and Unconstrained Least-Squares Importance Fitting (uLSIF, see Kanamori *et al.* [23]). As mentioned previously, however, the situation we consider here differs from the standard setting since we assume $p_{\text{target}}(x)$ to be a given (or automatically determined) tuning parameter and not something to be estimated from the data. Thus, the direct importance estimation procedures mentioned above are not applicable to our case, and we will need to resort to *density estimation* procedures for approximating $p_{\text{train}}(x)$. However, as mentioned above, non-parametric multivariate density estimation with a large number of samples is a hard problem [16, 36, 43], and standard software libraries typically do not provide applicable methods even today. The computational effort required for computing a standard kernel density estimate (KDE) is $\mathcal{O}(N^2)$, where N denotes the number of data points taken into consideration, although faster methods based on the non-uniform fast Fourier transform have recently been proposed [35, 36]. Moreover, *evaluating* a standard kernel density estimate based on N data points at M positions requires $\mathcal{O}(MN)$ operations. This quadratic evaluation complexity renders these methods infeasible for applications on long

The left margin contains the following side notes:

Estimator Variance

Direct Importance Estimation

Density Estimation

(potentially streaming) time series, even disregarding for a moment the computational effort required to obtain said kernel density estimate in the first place. To solve this problem, various fast updating mechanisms have been proposed in the literature [27]. In addition to the challenges related to computational complexity, the automatic determination of the kernel parameters in the multivariate case represents another crucial research challenge [3, 36].

Concerning the identification of *time-varying model parameters*, there is, of course, a large class of classical recursive identification methods, including the famous recursive least squares (RLS) algorithm and the Kalman filter and smoother, both of which have been discussed in chapter 2. Specifically, section 2.3.2 discusses how to employ a linear Kalman filter or smoother to perform time-varying parameter estimation in a regression model that is linear in the parameters. Such recursive (or incremental) identification schemes fall under the modern umbrella terms of *learning from streaming data* and *online learning*. In recent years, the topic of time-varying model parameter identification has received renewed interest by the machine learning community under the terms *concept drift* and *concept shift* [19, 25, 29, 56]. Popular methods for identifying time-varying models include adaptive learning methods, windowing techniques, and ensemble methods [19]. Recently, Oh *et al.* [37] have proposed a modified long-short term memory (LSTM) network specifically tailored towards time-varying systems. Mathematically speaking, parameter identification in time-varying systems is, of course, simply a time-varying optimization problem, for which many different approaches are known [46]. Time-varying Model Parameters Online Learning Concept Drift

Surprisingly few researchers have treated the combined problem of identification in the face of both covariate shift *and* concept drift [19, 26, 29, 54]. All prior works known to the author treat (often binary) *classification* problems — the application motivating this research, discussed in the following chapter 7, is a *regression* problem, however. Moreover, many of the proposed approaches represent rather *ad-hoc* modifications of existing learning algorithms to achieve adaptability by, e.g., retraining within a moving window or bagging multiple model instances. In the statistical literature, many authors have treated static weighted likelihood estimation problems as discussed above. However, to the author's knowledge, none have coupled weighted estimation with a model of concept drift as specified in eq. (6.1). Here, we will aim to derive a fully model-based, probabilistic regression procedure accounting for covariate shift and concept drift. For this, we will employ a novel, weighted Kalman filter and smoother, as well as a weighted prediction error identification method to identify the noise covariances. Technically, the new weighted filtering method is similar in spirit to earlier research on *adaptive-gain Kalman filtering*, which also exploits time-varying sample weights (albeit in a different way and towards a different purpose) [4, 44]. Covariate Shift and Concept Drift Adaptive-gain Kalman Filtering

6.2 Importance-weighted state estimation

In this section, a general formulation of a *weighted linear Kalman filter and smoother* will be derived, which can be used to solve a weighted state estimation problem. Given a probabilistic model of the concept drift as well as information about the time-varying importance w_k, this will enable us to perform parameter estimation in the face of concept drift, covariate shift, and model mismatch. Weighted Linear Kalman Filter and Smoother

In section 2.3.2, we saw that the linear Kalman smoother solves the optimization problem

$$
\begin{aligned}
\hat{\theta}^s_{0:N^S} &= \arg\max_{\theta_{0:N^S}} q(\theta_{0:N^S} \mid y_{1:N^S}) \\
&= \arg\max_{\theta_{0:N^S}} \log q(\theta_{0:N^S} \mid y_{1:N^S}) \\
&= \arg\max_{\theta_{0:N^S}} \log q(\theta_{0:N^S}, y_{1:N^S})
\end{aligned}
$$

$$
= \arg\max_{\theta_{0:N^S}} \sum_{k=1}^{N^S} \log q(y_k \mid \theta_k) + \sum_{k=1}^{N^S} \log q(\theta_k \mid \theta_{k-1}) + \log q(\theta_0) \tag{6.3}
$$

$$
= \arg\max_{\theta_{0:N^S}} -\frac{1}{2} \sum_{k=1}^{N^S} \left[(y_k - C_k\theta_k)^{\mathsf{T}} \Sigma_{\nu,k}^{-1} (y_k - C_k\theta_k) + \log|\Sigma_{\nu,k}| \right] \tag{6.4}
$$

$$
-\frac{1}{2} \sum_{k=0}^{N^S-1} \left[(\theta_{k+1} - A_{k+1}\theta_k - B_{k+1}u_{k+1})^{\mathsf{T}} \Sigma_{\eta,k+1}^{-1} \right.
$$

$$
\left. (\theta_{k+1} - A_{k+1}\theta_k - B_{k+1}u_{k+1}) + \log|\Sigma_{\eta,k+1}| \right]
$$

$$
-\frac{1}{2}\left[(\theta_0 - \hat{\theta}_0^+)^{\mathsf{T}}(P_0^+)^{-1}(\theta_0 - \hat{\theta}_0^+) + \log|P_0^+| + N^S(N^\theta + N^y)\log 2\pi \right]
$$

for the hidden state estimates $\hat{\theta}^s_k$, where $q(y_k \mid \theta_k)$ is given by the linear measurement model

$$
y_k = C_k\theta_k + \nu_k, \quad \nu_k \sim \mathcal{N}(0, \Sigma_\nu), \tag{6.5}
$$

and $q(\theta_k \mid \theta_{k-1})$ is given by the linear state progression model

$$
\theta_k = A_k\theta_{k-1} + B_ku_k + \eta_k, \quad \eta_k \sim \mathcal{N}(0, \Sigma_\eta) \tag{6.6}
$$

with a known input signal u_k. Moreover, we noted that solving this optimization problem for the random walk model

$$
\begin{aligned}
\theta_k &= \theta_{k-1} + \eta_k \\
y_k &= x_k^{\mathsf{T}}\theta_k + \nu_k,
\end{aligned} \tag{6.7}
$$

with $\Sigma_\eta = 0$, $\Sigma_\nu = \text{const}$ and $P_0^+ = \text{diag}(\infty, \dots, \infty)$ (and $C_k = x_k^{\mathsf{T}}$) is equivalent to solving the (multiple) multivariate static least-squares optimization problem

$$
\hat{\theta} = \arg\min_{\theta} \sum_{i=1}^{N^S} (y_i - x_i\theta)^{\mathsf{T}} \Sigma_\varepsilon^{-1} (y_i - x_i\theta).
$$

This equivalence can be generalized: allowing for arbitrary matrices A_k and Σ_η (as opposed to $A_k = 1 \,\forall\, k$ and $\Sigma_\eta = 0$ above), the (usual) Kalman smoother can be utilized to perform fully probabilistic regression in the face of concept drift, where $q(\theta_k \mid \theta_{k-1})$ (parameterized by A_k and Σ_η) specifies an assumed model of concept drift and

$$
y_k = x_k^{\mathsf{T}}\theta_k + \nu_k \tag{6.8}
$$

the regression model.[3]

How can this approach be extended to also treat *covariate shift* correctly? In the static regression case, i.e., without concept drift, covariate shift is treated within the maximum-likelihood estimation framework by solving the importance-weighted log likelihood optimization problem

$$\hat{\theta} = \arg\max_{\theta} \sum_{k=1}^{N^S} w_k \log q(y_k \mid \theta) \tag{6.9}$$

instead of its unweighted counterpart, see section 2.1.6. Here, analogously, we will consider the weighted optimization problem

$$\hat{\theta}_{0:N^S}^{ws} = \arg\max_{\theta_{0:N^S}} \sum_{k=1}^{N^S} w_k \log q(y_k \mid \theta_k) + \sum_{k=1}^{N^S} \log q(\theta_k \mid \theta_{k-1}) + \log q(\theta_0) \tag{6.10}$$

$$= \arg\max_{\theta_{0:N^S}} -\frac{1}{2} \sum_{k=1}^{N^S} w_k \left[(y_k - C_k\theta_k)^\mathsf{T} \Sigma_{\nu,k}^{-1} (y_k - C_k\theta_k) + \log|\Sigma_{\nu,k}| \right] \tag{6.11}$$

$$-\frac{1}{2} \sum_{k=0}^{N^S-1} \left[(\theta_{k+1} - A_{k+1}\theta_k - B_{k+1}u_{k+1})^\mathsf{T} \Sigma_{\eta,k+1}^{-1} \right.$$

$$\left. (\theta_{k+1} - A_{k+1}\theta_k - B_{k+1}u_{k+1}) + \log|\Sigma_{\eta,k+1}| \right]$$

$$-\frac{1}{2} \left[(\theta_0 - \hat{\theta}_0^+)^\mathsf{T} (P_0^+)^{-1} (\theta_0 - \hat{\theta}_0^+) + \log|P_0^+| + N^S(N^\theta + N^y) \log 2\pi \right]$$

instead of eq. (6.3). Notice that eq. (6.9) is a special case of eq. (6.10), which is obtained by choosing a flat prior $q(\theta_0)$ and the static state model $q(\theta_k \mid \theta_{k-1}) = \delta(\theta_k - \theta_{k-1})$. Moreover, recall again that eq. (6.10) is a structural importance-weighted empirical risk minimization problem, where the first part represents the empirical risk $\mathcal{R}_{\text{emp}}(\theta)$ and the second part represents a regularization term $\ell_{\text{reg}}(\theta)$. Conveniently, the optimization problem (6.11) can be readily solved using a standard linear Kalman smoother, by simply letting

$$\tilde{\Sigma}_{\nu,k} := \frac{1}{w_k} \Sigma_{\nu,k}$$

and running the standard algorithm with measurement noise covariance matrices $\tilde{\Sigma}_{\nu,k}$ instead of $\Sigma_{\nu,k}$. Thus, we have a simple procedure available for solving general weighted state estimation problems — including, in particular, regression problems in the face of concept drift, covariate shift, and model mismatch.

In the derivation of the standard Kalman filter and smoother, it was assumed that the matrices $\Sigma_{\nu,k}$ correctly denote the measurement noise covariance. Based on this assumption (and others), it was derived that the quantity P_k^+ in the Kalman filter algorithm correctly represents the covariance of the (weighted) state estimate. Since we are now using the weighted matrices $\tilde{\Sigma}_{\nu,k}$ instead, this is no longer the case, and P_k^+ becomes a purely computational quantity that is leveraged for solving eq. (6.10). This is analogous to our discussion of the correct calculation of the WLS estimator covariance in section 2.2.2. The

---

[3]The extension to the fully nonlinear setting is straightforward but not spelled out here for brevity's sake.

correct covariance of the weighted (filtered) state estimate can be derived, however: From eq. (6.6), it follows[4] that

$$\text{Var}[\theta_k] = A_k \text{Var}[\theta_{k-1}] A_k^{\mathsf{T}} + \Sigma_\eta$$

which, assuming $\theta_{k-1} \sim \mathcal{N}\left(\hat{\theta}_{k-1}^{w+}, P_{k-1}^{w+}\right)$, yields

$$P_k^{w-} := \text{Var}[\hat{\theta}_k^{w-}] = A_k P_{k-1}^{w+} A_k^{\mathsf{T}} + \Sigma_\eta.$$

Similarly, from the weighted state update (see algorithm 2 in section 2.3.1, substituting the weighted estimate $\hat{\theta}^{w-}$ for $\hat{\theta}^-$)

$$\hat{\theta}_k^{w+} = \hat{\theta}_k^{w-} + K_k(y_k - C_k \hat{\theta}_k^{w-})$$
$$= (I - K_k C_k)\hat{\theta}_k^{w-} + K_k y_k$$

it follows that

$$
\begin{aligned}
P_k^{w+} &:= \text{Var}[\hat{\theta}_k^{w+}] \\
&= (I - K_k C_k) \text{Var}[\hat{\theta}_k^{w-}](I - K_k C_k)^{\mathsf{T}} + K_k \text{Var}[y_k] K_k^{\mathsf{T}} \\
&= (I - K_k C_k) P_k^{w-} (I - K_k C_k)^{\mathsf{T}} + K_k \Sigma_\nu K_k^{\mathsf{T}}.
\end{aligned}
$$

Weighted Linear Kalman Filter The resulting *weighted linear Kalman filter* algorithm is summarized in algorithm 7. Notice that for $W_k \equiv 1$, the update equations for the covariances P^+ and P^{w+} coincide, as desired. Moreover, for $\Sigma_\eta \equiv 0$, the estimator covariance at the final state is identical to the covariance obtained by performing WLS regression and evaluating eq. (2.20).

The correct estimator covariance of the *smoothed* weighted state estimates can be derived analogously: From the standard RTS smoothing equation (see algorithm 4 in section 2.3.1)

$$\hat{\theta}_k^{ws} = \hat{\theta}_k^{w+} + S_k(\hat{\theta}_{k+1}^{ws} - \hat{\theta}_{k+1}^{w-})$$

it follows that

$$P_k^{ws} = P_k^{w+} + S_k(P_{k+1}^{ws} - P_{k+1}^{w-})S_k^{\mathsf{T}}.$$

Weighted Rauch-Tung-Striebel Smoother Algorithm 8 summarizes the resulting *weighted Rauch-Tung-Striebel smoother* (WRTS).

6.3 Importance-weighted hyperparameter estimation

A crucial question for the practical application of the scheme discussed in the previous section remains: how can one identify the parameters ψ of the state-space model eqs. (6.5) and (6.6), i.e., some or all of the (components of the) matrices A_k, C_k, $\Sigma_{\eta,k}$, and $\Sigma_{\nu,k}$, if they are not known a priori? Many different methods are known in the literature for tuning the parameters of state-space models; here, we will employ a *prediction error identification method (PEM)*. Classical quadratic PEM [31] (which is closely related to ML estimation) chooses the hyperparameters as the solution to the empirical risk minimization problem

Prediction Error Identification

$$
\begin{aligned}
\hat{\psi}_{\text{PEM}} &= \arg\min_\psi \sum_{k=1}^{N^S} \left(y_k - C_k \hat{\theta}_k^-\right)^2 \\
&= \arg\min_\psi \text{E}_{\text{P}_{\text{NS}}(x) \times \text{P}_{\text{NS}}(k)}\left[\left(y_k - C_k \hat{\theta}_k^-\right)^2\right],
\end{aligned}
$$

---

[4] Recall that $\text{Var}[Ax + c] = A \text{Var}[x]A^{\mathsf{T}} + \text{Var}[c]$ (assuming A to be known).

Algorithm 7: Weighted linear Kalman filter

1 **Function** `WKF`

    Input   :Mean $\hat{\theta}_0^+ \in \mathbb{R}^{N^\theta}$ and covariance $P_0^+ = P_0^{w+} \in \mathbb{R}^{N^\theta \times N^\theta}$ of the
                  initial state estimate $p(\theta_0)$, matrices $A_k, B_k, C_k, \Sigma_{\nu,k}$ and $\Sigma_{\eta,k}$
                  of suitable dimensions, measurements $y_k \in \mathbb{R}^{N^y}$ and
                  inputs $u_k \in \mathbb{R}^{N^u}$, as well as a time-varying weighting
                  matrix W_k, each for $k = 1, \dots, N^S$.

    Output:Mean $\hat{\theta}_k^{w+}$ and covariance P_k^{w+} of the optimal solution to the
                  weighted state inference problem (6.11) for $k = 1, \dots, N^S$,
                  each calculated using samples up until index k.

2   **begin**

3   **for** $k = 1$ **to** N^S **do**

            // State prediction using eq. (2.21)

4             $\hat{\theta}_k^{w-} = A_k \hat{\theta}_{k-1}^{w+} + B_k u_k$

5             $P_k^- = A_k P_{k-1}^+ A_k^\mathsf{T} + \Sigma_{\eta,k}$

            // Recursive posterior update using eq. (2.22)

6             $K_k = P_k^- C_k^\mathsf{T} (C_k P_k^- C_k^\mathsf{T} + W_k \Sigma_{\nu,k})^{-1}$

7             $\hat{\theta}_k^{w+} = \hat{\theta}_k^{w-} + K_k (y_k - C_k \hat{\theta}_k^{w-})$

8             $P_k^+ = (I - K_k C_k) P_k^- (I - K_k C_k)^\mathsf{T} + K_k W_k \Sigma_{\nu,k} K_k^\mathsf{T}$

            // Optional: calculate estimator uncertainty

9             $P_k^{w-} = A_k P_{k-1}^{w+} A_k^\mathsf{T} + \Sigma_{\eta,k}$

10            $P_k^{w+} = (I - K_k C_k) P_k^{w-} (I - K_k C_k)^\mathsf{T} + K_k \Sigma_{\nu,k} K_k^\mathsf{T}$

11   **end**

12   **end**

Algorithm 8: Weighted linear Rauch-Tung-Striebel smoother (WRTS)

---

1 **Function** `WRTS`

   Input : Mean $\hat{\theta}_0^{w+} \in \mathbb{R}^{N^\theta}$ and covariance $P_0^+ = P_0^{w+} \in \mathbb{R}^{N^\theta \times N^\theta}$ of the
   initial state estimate $p(\theta_0)$, matrices $A_k, B_k, C_k, \Sigma_{\nu,k}$ and $\Sigma_{\eta,k}$
   of suitable dimensions, measurements $y_k \in \mathbb{R}^{N^y}$ and
   inputs $u_k \in \mathbb{R}^{N^u}$, as well as a time-varying weighting
   matrix W_k, each for $k = 1, \dots, N^S$.

   Output : Mean $\hat{\theta}_k^{ws}$ and covariance P_k^{ws} of the optimal solution to the
   weighted state inference problem (6.11) for $k = 0, 1, \dots, N^S$.

2 **begin**
      // Forward filtering using the standard Kalman filter
3   Run `WKF`
4     $\hat{\theta}_{N^S}^{ws} = \hat{\theta}_{N^S}^{w+}$
      // Backward smoothing using the Rauch-Tung-Striebel smoother
5   **for** $k = N^S - 1$ **to** 0 **do**
6   Solve $S_k P_{k+1}^- = P_k^+ A_k^\mathsf{T}$ for S_k
7         $P_k^s = P_k^+ + S_k(P_{k+1}^s - P_{k+1}^-)S_k^\mathsf{T}$
8         $\hat{\theta}_k^{ws} = \hat{\theta}_k^{w+} + S_k(\hat{\theta}_{k+1}^{ws} - \hat{\theta}_{k+1}^{w-})$
          // Optional: calculate estimator uncertainty
9         $P_k^{ws} = P_k^{w+} + S_k(P_{k+1}^{ws} - P_{k+1}^{w-})S_k^\mathsf{T}$
10   **end**
11 **end**

---

where $\hat{\theta}_k^-$ denotes a forward prediction of the state θ_k, given the measurements $y_{1:k-1}$. Here, however, we are interested in solving time-varying regression problems, i.e., $C_k = x_k^\mathsf{T}$, under *covariate shift*. Thus, the aim is not to minimize the risk over $p_{\text{train}}(x) \approx P_{N^S}$ but instead over $p_{\text{target}}(x)$. Therefore, we will instead solve the importance-weighted empirical risk minimization (IWERM) or importance-weighted prediction error minimization (IWPEM) problem

$$\hat{\psi}_{\text{IWPEM}} = \arg\min_\psi \mathrm{E}_{\mathrm{P}_{N^S}(\mathrm{x}) \times \mathrm{P}_{N^S}(\mathrm{k})} \left[w_k \left(y_k - C_k \hat{\theta}_k^{w-} \right)^2 \right]$$

$$= \arg\min_\psi \sum_{k=1}^{N^S} w_k \left(y_k - C_k \hat{\theta}_k^{w-} \right)^2$$

$$\approx \arg\min_\psi \mathrm{E}_{p_{\text{target}}(\mathrm{x}) \times \mathrm{Unif}(\mathrm{k})} \left[w_k \left(y_k - C_k \hat{\theta}_k^{w-} \right)^2 \right] = \arg\min_\psi \sum_{k=1}^{N^S} w_k \left(y_k - C_k \hat{\theta}_k^{w-} \right)^2,$$

where w_k denotes the importance weight, and $\hat{\theta}_k^{w-}$ the mean of the predicted state distribution computed by the WKF (see algorithm 7). Notice that the expression $y_k - C_k \hat{\theta}_k^{w-}$ is also called the *innovation* of the (weighted) Kalman filter. This quantity can be readily optimized using a numerical optimization scheme, as will be demonstrated in the numerical examples in section 6.5.

6.4 Efficient and accurate multivariate density estimation

The importance-weighted estimation methods described in the previous two sections rely on the availability of accurate importance weights. While $p_{\text{target}}(x)$ will be assumed to be known, $p_{\text{train}}(x)$ must be estimated from the available data. This is a nontrivial endeavor because the setting is multivariate, the number of samples is large, and the resulting density estimate must be evaluated at equally many points (namely, at each measured sample). To the author's knowledge, no efficient method is readily available in standard software packages. Closest to the fulfillment of these requirements may be the fastKDE method [36], which performs highly efficient kernel density estimation (KDE) in the multivariate setting, but which, however, is inefficient if the number M of query points is large. For these reasons, a simple multivariate histogram interpolation method is implemented here, which is described in the following. As density estimation is not the main focus of this chapter and the method is, at its core, far from new, the method is described rather briefly. For the interested reader, the full code is openly available at `https://github.com/e-pet/mvdensity`.

A *multivariate histogram* can be computed very efficiently, and the complexity of the interpolation operation then only depends on the number N^b of histogram bins, *not* the number of measured samples. The resulting interpolated density surface can then be evaluated with complexity $\mathcal{O}(N^b M)$, which is sufficiently cheap. First, artificial boundary histogram bins with zero counts are added to force the smoothed density surface to decline towards zero outside of the histogram, and the geometrical center points of the histogram bins are calculated. For the interpolation step, the *makima* interpolation method is used [22], which represents a modification of the interpolation method proposed by Akima [1]. To prevent potentially negative undershoots during the smoothing step and thus

Multivariate Histogram

guarantee the (semi-)positivity of the resulting PDF estimate, the histogram counts are transformed using the inverse softplus function [9]

$$y = f^{-1}(x) = \log(e^x - 1).$$

Interpolation is then performed on the transformed histogram counts, and the resulting interpolated PDF estimate is transformed back using the softplus function [9]

$$f(y) = \log(1 + e^y).$$

Finally, to obtain a correctly normalized density estimate, the integral of the estimated PDF surface is approximated using a numerical integration scheme, and the estimated surface is divided by this constant.

6.5 Results

The weighted Rauch-Tung-Striebel smoother presented in the previous section will now be applied to two numerical examples to demonstrate its utility for time-varying, importance-weighted parameter estimation. The first is an example of static regression under model mismatch and covariate shift, i.e., the parameters are *not* time-varying. This estimation problem is solved using importance-weighted maximum likelihood estimation. In the second example, model mismatch and covariate shift are also present, and in addition, there is concept drift. A concept drift model is identified using the importance-weighted maximum likelihood approach described in section 6.3, and the time-varying model parameters are identified using the importance-weighted Kalman filter and smoother discussed in section 6.2. First, however, the multivariate density estimation scheme discussed in the previous section will be validated because this method will be employed for the weight estimation in the two importance-weighted estimation problems.

6.5.1 Multivariate density estimation

To validate the accuracy of the proposed density estimation method, three example problems are considered:

- a one-dimensional Gaussian mixture distribution with two components,

- a two-dimensional Gaussian mixture distribution with two components, and

- a five-dimensional Gaussian mixture distribution with three components.

In all three cases, $N^r = 10$ samples of size $N^S = 10^\ell$ are drawn, with $\ell = 1, \dots, 6$. PDF estimation is performed on each of the samples using the method described in the previous section. The differences between the true PDFs and the estimated PDFs are quantified using the *Jensen-Shannon divergence* [30], which is given by

Jensen-
Shannon
Divergence

$$D_{JS}(P \parallel Q) = \frac{1}{2} D_{KL}(P \parallel M) + \frac{1}{2} D_{KL}(Q \parallel M), \quad M = \frac{1}{2}P + \frac{1}{2}Q,$$

where $D_{KL}(P \parallel Q)$ denotes the Kullback-Leibler divergence (defined in eq. (2.11)). As opposed to the Kullback-Leibler divergence, the Jensen-Shannon divergence has the
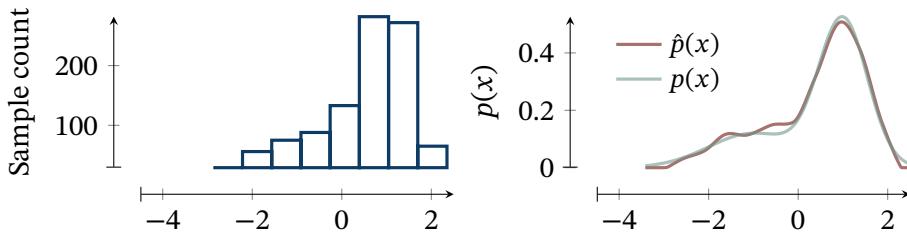
Figure 6.2 – An example of one-dimensional density estimation using the method described in section 6.4. **Left:** Histogram of the observed data. PDF of the true distribution, from which N^S samples have been drawn. **Right:** the true probability distribution $p(x)$, from which the N^S samples yielding the histogram in the left panel have been drawn, and the estimated probability distribution $\hat{p}(x)$. The latter is obtained by simple interpolation of the (inverse softplus-transformed) histogram. In this example, $N^S = 1000$.
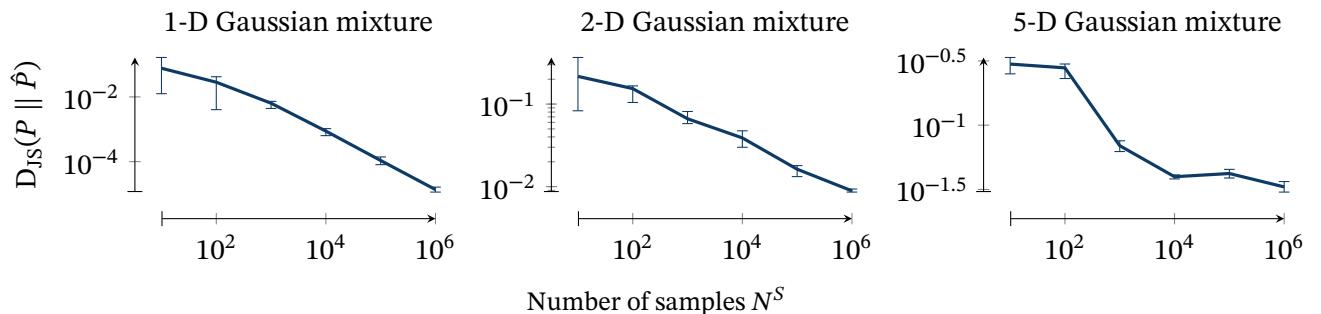
Figure 6.3 – Performance evaluation of the described multivariate density estimation scheme. The graphs show the Jensen-Shannon divergence $D_{JS}(P \parallel \hat{P})$ between the true probability distribution P and the estimation distribution $\hat{P}$ as a function of the sample size N^S, for the three Gaussian mixture (GM) examples described in the text. Values shown are the median over 10 runs, with the error bars indicating the min-max range.

desirable property of being finite for all distributions — more specifically, it is bounded between zero and one, with zero being attained if both distributions are identical.[5] Figure 6.2 shows the result of an exemplary instance of the one-dimensional example, and fig. 6.3 the median, maximum, and minimum Jensen-Shannon divergence for each example as a function of N^S. The method yields accurate results in all three settings, with the accuracy increasing with increasing sample size N^S.

6.5.2 Multivariate static importance-weighted estimation

We will now consider a simple example of multivariate static regression in the face of covariate shift and model misspecification. This example serves to illustrate the general problem of regression under covariate shift; the second example following in section 6.5.3 will illustrate the more complex case with concept drift occurring. We draw $N^S = 5000$ i.i.d. samples $[x_1, x_2]$ from

$$x_1 \sim \Gamma_{[0,10]}(a{=}2, b{=}1), \quad x_2 \sim \mathcal{N}_{[-3,3]}(\mu{=}0, \sigma{=}1),$$

---

[5]Moreover, the Jensen-Shannon divergence induces a proper metric: its square root [10].

where $\Gamma_{[0,10]}(a{=}2, b{=}1)$ denotes the gamma distribution with shape and scale parameters a and b, truncated to $[0, 10]$, and $\mathcal{N}_{[-3,3]}(\mu{=}0, \sigma{=}1)$ analogously denotes a truncated normal distribution. From these covariates, the target variable y is generated following

$$y_k = f(x_k) + \nu_k = f_1(x_{1,k}) + 0.5x_{2,k} + \nu_k$$

with measurement noise $\nu_k \sim \mathcal{N}(\mu{=}0, \sigma{=}40)$ and the nonlinearity

$$f_1(x_1) = 0.1x_1^4 - 0.5x_1 + 10.$$

To the generated data, we will fit the simple linear model

$$\hat{y}_k = \hat{f}(x_k) = \theta_1 x_{1,k} + \theta_2 x_{2,k} + \theta_0.$$

Furthermore, assume that the target distribution is the uniform distribution over the whole data space, i.e.,

$$p_{\text{target}}(x) = \text{Unif}_{[0,10]\times[-3,3]}(x).$$

Notice that (in the case of ML estimation and the assumption of Gaussian measurement noise) this is equivalent to searching for the approximation $\hat{f}(x)$ that minimizes the mean squared error with respect to $f(x)$ over the whole data space $[0, 10] \times [-3, 3]$.

To solve this estimation problem, we employ simple importance-weighted least squares (IWLS) estimation, which, as discussed in section 2.2.2, coincides with the importance-weighted maximum likelihood (IWML) solution (assuming Gaussian measurement noise) and, thus, importance-weighted empirical risk minimization (IWERM). Either WLS (see section 2.2.2) or the WRTS (algorithm 8, with $A = I$ and $\Sigma_\eta = 0$) can be employed; both yield identical results. The data generation process and subsequent estimation are repeated 1000 times in order to be able to quantify the bias and variance of the different estimators. Figure 6.4 shows one realization of the data as well as the average results of the OLS, exact IWLS, and approximate IWLS estimators, compared to the ideal solution. Here, "exact" refers to the use of the exact importance weights

$$w^*(x) = \frac{p_{\text{target}}(x)}{p_{\text{train}}(x)} = \frac{\text{Unif}_{[0,10]\times[-3,3]}(x)}{\Gamma_{[0,10]}(x_1 \mid 2, 1) \cdot \mathcal{N}_{[-3,3]}(x_2 \mid 0, 1)}.$$

This is in contrast to the approximate importance weights

$$w(x) = \frac{p_{\text{target}}(x)}{\hat{p}_{\text{train}}(x)},$$

where $\hat{p}_{\text{train}}(x)$ is an approximation to $p_{\text{train}}(x)$ that is obtained from the observed samples using the method described in section 6.4. As expected, OLS exhibits a small variance but high bias (due to model mismatch and covariate shift), whereas the two IWLS estimators are nearly unbiased but have a much higher variance. Approximate IWLS exhibits a slightly larger bias than exact IWLS but also a significantly smaller variance.
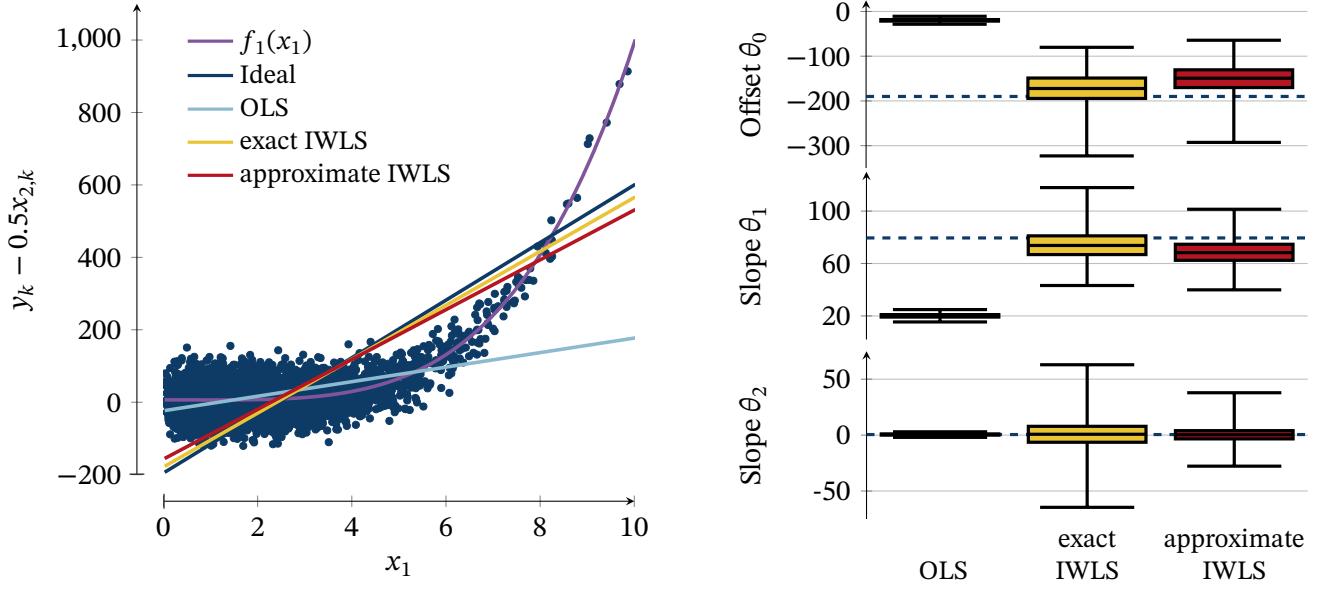
Figure 6.4 – Left: The dots show one realization of the data (with the effect of x_2 canceled out, to illustrate the nonlinearity $f_1(x_1)$). The regression lines represent the average result of the corresponding estimator over 1000 realizations. The "ideal" solution is the one that minimizes the mean squared error between the estimate and the true response surface over the target distribution $p_{\text{target}}(x) = \text{Unif}_{[0,10] \times [-3,3]}$. **Right:** Distribution of the estimated parameter values. The dashed line represents the ideal value, coinciding with the ideal regression line in the left panel. The boxes show the median and interquartile range, and the whiskers the min–max range.

6.5.3 Time-varying importance-weighted estimation

This second example is one-dimensional (mainly to simplify visualization) but features the full combination of covariate shift, model mismatch, and concept drift. We draw $N^S = 50000$ i.i.d. samples from

$$x \sim \Gamma_{[0,8]}(a=2, b=1).$$

From these, measurements are generated following

$$y_k = f(x_k, t_k) + v_k, \quad v_k \sim \mathcal{N}(\mu=0, \sigma=10)$$

with the time-varying nonlinearity

$$f(x, t) = (0.1 \cos(2\pi t) - 0.3)x^4 - 0.5x + 500 - 490 \sin^2(2\pi t),$$

where

$$t_k = \frac{k}{N^S}.$$

Finally, we again assume that the target distribution is the uniform distribution over the whole data space, i.e.,

$$p_{\text{target}}(x) = \text{Unif}_{[0,8]}(x),$$
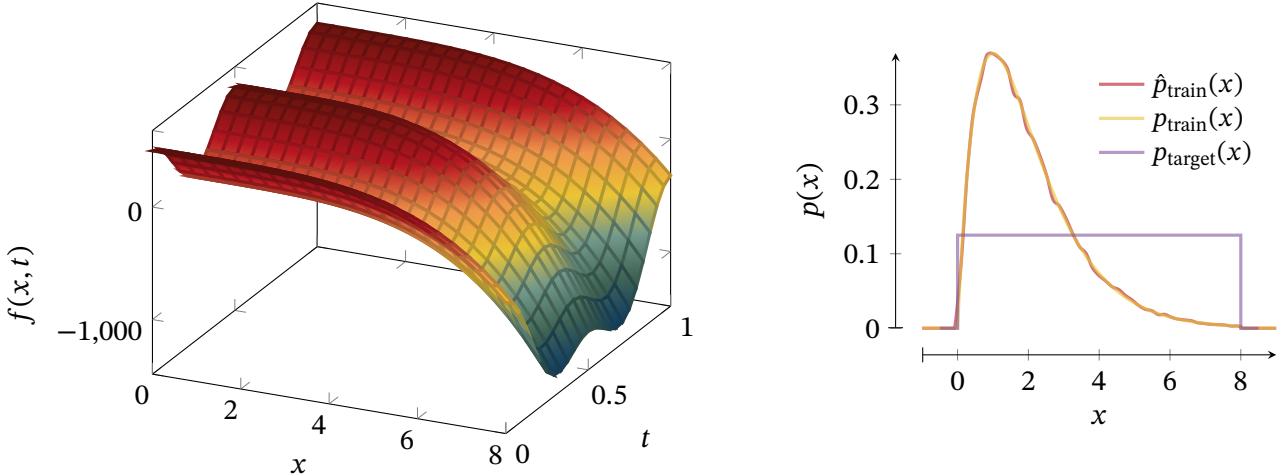
Figure 6.5 – **Left:** the nonlinearity $f(x,t)$, changing over time and thus exhibiting *concept drift*. **Right:** the training and target densities $p_{\text{train}}(x)$ and $p_{\text{target}}(x)$, as well as an approximation $\hat{p}_{\text{train}}(x)$ to $p_{\text{train}}(x)$ identified using the density estimation scheme described in section 6.4.

and we fit a simple, time-varying linear model of the form

$$\hat{f}(x,t) = \theta_1(t)x + \theta_0(t).$$

To approximate the training data distribution (which is assumed unknown), we again employ the method described in section 6.4. Figure 6.5 shows the time-varying nonlinearity $f(x,t)$ as well as the exact and approximated training distribution and the target distribution.

For identifying the time-varying model parameters, the WRTS (algorithm 8) is employed. We assume the time-varying drift model

$$\begin{pmatrix} \theta_{0,k} \\ \delta_{0,k} \\ \theta_{1,k} \\ \delta_{1,k} \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} \theta_{0,k-1} \\ \delta_{0,k-1} \\ \theta_{1,k-1} \\ \delta_{1,k-1} \end{pmatrix} + \eta_k, \quad \eta_k \sim \mathcal{N}\big(\mu=0, \Sigma=\Sigma_\eta\big)$$

for the model parameters, and the measurement model

$$y_k = \begin{pmatrix} 1 & 0 & x_k & 0 \end{pmatrix} \cdot \begin{pmatrix} \theta_{0,k} \\ \delta_{0,k} \\ \theta_{1,k} \\ \delta_{1,k} \end{pmatrix} + \nu_k, \quad \nu_k \sim \mathcal{N}(\mu=0, \Sigma=\Sigma_\nu).$$

We parameterize the noise covariance matrices in terms of a vector ψ of hyperparameters as

$$\Sigma_\eta(\psi) = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & \Sigma_{\eta\delta,11} & 0 & \Sigma_{\eta\delta,12} \\ 0 & 0 & 0 & 0 \\ 0 & \Sigma_{\eta\delta,21} & 0 & \Sigma_{\eta\delta,22} \end{pmatrix}$$

where

$$\Sigma_{\eta\delta}(\psi) = \begin{pmatrix} \Sigma_{\eta\delta,11} & \Sigma_{\eta\delta,12} \\ \Sigma_{\eta\delta,21} & \Sigma_{\eta\delta,22} \end{pmatrix} = L^{\mathsf{T}}(\psi)\,L(\psi)$$

with

$$L(\psi) = \begin{pmatrix} \psi_1 & \psi_2 \\ 0 & \psi_3 \end{pmatrix}$$

and

$$\Sigma_{\nu}(\psi) = \psi_4^2.$$

Finally, we determine the weighted prediction error-optimal set of hyperparameters

$$\hat{\psi}_{\mathrm{IWPEM}} = \arg\min_{\psi} \sum_{k=1}^{N^S} w_k \left(y_k - C_k\,\hat{\theta}_k^{w-}(\psi) \right)^2$$

as per section 6.3. The optimization is solved numerically in two steps. First, a latin hypercube search is performed at 100 points to find a good starting point for the following optimization step. Second, the Matlab `patternsearch` solver is used, which implements a variant of generalized pattern search [2, 51], a derivative-free direct ("black-box") search method. Each function evaluation during the optimization procedure necessitates a full WRTS run. (Notice that the optimization problem exhibits many local minima, whence local solvers converge towards distinctly suboptimal solutions for this problem.)

Figure 6.6 shows the simulated data as well as the estimation results. The IWPEM-optimal IWRTS closely tracks the ideal parameters, with only a small difference between the results obtained using ideal importance weights and those obtained using approximate weights. For comparison, estimation results using the same model but unweighted PEM-optimal hyperparameters and an unweighted RTS are also shown (PEM-RTS), which are starkly disturbed.

6.6 Discussion & outlook

In this chapter, a novel and fully probabilistic method for performing regression in the face of model mismatch, covariate shift, and concept drift has been described. To this end, weighted versions of the linear Kalman filter and Rauch-Tung-Striebel smoother have been derived, and it has been shown that they solve a structural importance-weighted empirical risk minimization problem. The optimal hyperparameters are obtained from solving an importance-weighted maximum likelihood estimation problem, which can similarly be understood as an importance-weighted empirical risk minimization problem. To estimate the importance weights, a simple yet efficient multivariate density estimation scheme has been described. The efficacy of the proposed methodology for obtaining time-varying parameter estimates under model mismatch, covariate shift, and concept drift has been demonstrated in two numerical examples. To the author's knowledge, no method has been described previously that treats this complex setting, although a few methods have been described in the literature that treat *classification* under the same assumptions [19, 26, 29, 54], often in a rather ad-hoc fashion.

As is typically the case for importance-weighted methods [7, 8, 28, 45], the estimators proposed here suffer from increased variance as compared to the classical, uniformly
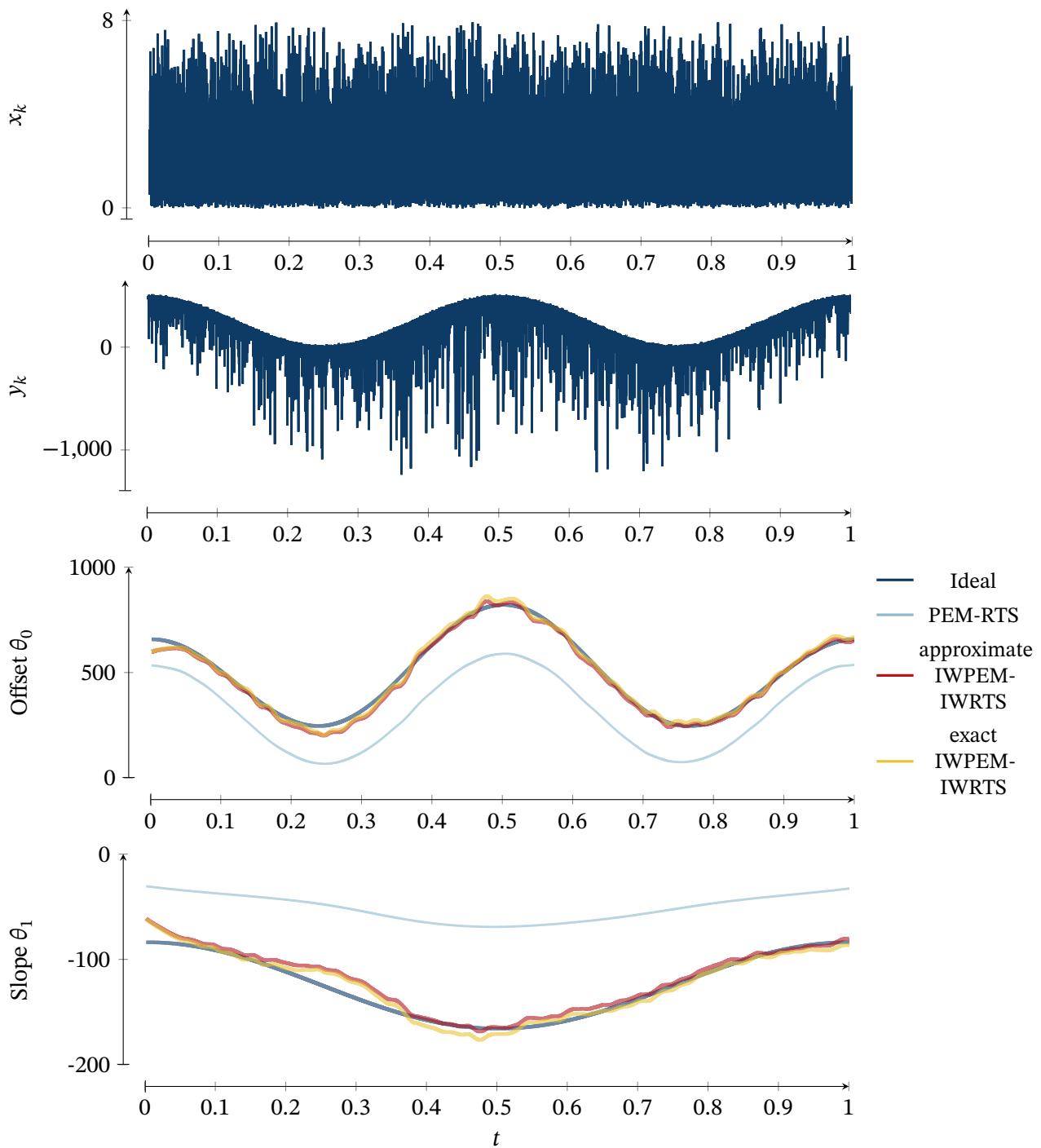
Figure 6.6 – Signals related to the numerical example described in section 6.5.3. (Unweighted) prediction error-optimal Rauch-Tung-Striebel smoother (PEM-RTS), importance-weighted prediction error-optimal weighted Rauch-Tung-Striebel smoother (IWPEM-IWRTS). *Exact* IWPEM-IWRTS uses the exact importance weights $p_{\text{target}}(x)/p_{\text{train}}(x)$, whereas *approximate* IWPEM-IWRTS uses the estimated weights $p_{\text{target}}(x)/\hat{p}_{\text{train}}(x)$.

weighted estimators; this is an artifact of the high weights assigned to individual (rare) samples and is simply the price paid for the reduction in estimator bias. For the same reason, the success of importance-weighted estimation schemes hinges in large parts on the correctness of the weight specification [8, 24, 28]. Here, in contrast to most of the literature on importance weighting, we assumed the target distribution $p_{\text{target}}(x)$ to not be descriptive of some actual observable distribution, but instead to represent a *tuning* parameter that can be used to control how well different regions of the data space should be described by the model. This is akin to frequency weighting in the system identification setting [31, 42], and there are several settings in which this is a natural perspective, one of which will be the subject of the following chapter.

p_{target} as a Tuning Parameter

Various extensions to the proposed method are conceivable. The method is described here for regression models that are linear in the model parameters but could be easily extended to the general *nonlinear setting*, using, e.g., approximative nonlinear filtering algorithms such as the one described by Herzog *et al.* [18]. Other models of concept drift can, of course, be considered, including models of more sudden drifts. The latter could be realized, e.g., using the NUV formalism proposed by Loeliger *et al.* [32]. To choose from multiple competing model proposals, the model selection method of Shimodaira [45] could be applied. The state-space setting is also amenable to the *classification* setting [57]. An extension of the regression framework described here to the solution of *classification* problems under model mismatch, covariate shift, and concept drift thus appears promising and highly relevant to many practical application scenarios. Finally, whereas the estimators described in this chapter have been derived thoroughly, a proper analysis of their statistical properties would be of further interest.

Fully Nonlinear Regression

Classification

References

[1] H. Akima, "A method of bivariate interpolation and smooth surface fitting based on local procedures," *Communications of the ACM*, vol. 17, no. 1, pp. 18–20, 1974. DOI: 10.1145/360767.360779.

[2] C. Audet and J. E. Dennis, "Analysis of generalized pattern searches," *SIAM Journal on Optimization*, vol. 13, no. 3, pp. 889–903, 2002. DOI: 10.1137/s1052623400378742.

[3] A. Bernacchia and S. Pigolotti, "Self-consistent method for density estimation," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, vol. 73, no. 3, pp. 407–422, 2011. DOI: 10.1111/j.1467-9868.2011.00772.x.

[4] N. Boizot, E. Busvelle, and J.-P. Gauthier, "An adaptive high-gain observer for nonlinear systems," *Automatica*, vol. 46, no. 9, pp. 1483–1488, 2010. DOI: 10.1016/j.automatica.2010.06.004.

[5] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: Synthetic minority over-sampling technique," *Journal of Artificial Intelligence Research*, vol. 16, pp. 321–357, 2002. DOI: 10.1613/jair.953.

[6] N. V. Chawla, "Data mining for imbalanced datasets: An overview," in *Data Mining and Knowledge Discovery Handbook*, Springer US, 2009, pp. 875–886. DOI: 10.1007/978-0-387-09823-4_45.

[7]   C. Cortes, Y. Mansour, and M. Mohri, "Learning bounds for importance weighting," in *Advances in Neural Information Processing Systems*, J. Lafferty *et al.*, Eds., vol. 23, Curran Associates, Inc., 2010. [Online]. Available: `https : //proceedings.neurips.cc/paper/2010/file/59c33016884a62116be975a9bb8257e3-Paper.pdf`.

[8]   C. Cortes, M. Mohri, M. Riley, and A. Rostamizadeh, "Sample selection bias correction theory," in *Lecture Notes in Computer Science*, Springer Berlin Heidelberg, 2008, pp. 38–53. DOI: `10.1007/978-3-540-87987-9_8`.

[9]   C. Dugas *et al.*, "Incorporating second-order functional knowledge for better option pricing," in *Advances in Neural Information Processing Systems*, T. Leen, T. Dietterich, and V. Tresp, Eds., vol. 13, MIT Press, 2001. [Online]. Available: `https://proceedings.neurips.cc/paper/2000/file/44968aece94f667e4095002d140b5896-Paper.pdf`.

[10]   D. Endres and J. Schindelin, "A new metric for probability distributions," *IEEE Transactions on Information Theory*, vol. 49, no. 7, pp. 1858–1860, 2003. DOI: `10.1109/tit.2003.813506`.

[11]   A. Fernández *et al.*, *Learning from Imbalanced Data Sets*. Springer International Publishing, 2018. DOI: `10.1007/978-3-319-98074-4`.

[12]   C. Field and B. Smith, "Robust estimation: A weighted maximum likelihood approach," *International Statistical Review / Revue Internationale de Statistique*, vol. 62, no. 3, p. 405, 1994. DOI: `10.2307/1403770`.

[13]   M. Galar *et al.*, "A review on ensembles for the class imbalance problem: Bagging-, boosting-, and hybrid-based approaches," *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, vol. 42, no. 4, pp. 463–484, 2012. DOI: `10.1109/tsmcc.2011.2161285`.

[14]   T. Gneiting and A. E. Raftery, "Strictly proper scoring rules, prediction, and estimation," *Journal of the American Statistical Association*, vol. 102, no. 477, pp. 359–378, 2007. DOI: `10.1198/016214506000001437`.

[15]   R. van den Goorbergh, M. van Smeden, D. Timmerman, and B. V. Calster, "The harm of class imbalance corrections for risk prediction models: Illustration and simulation using logistic regression," *Journal of the American Medical Informatics Association*, vol. 29, no. 9, pp. 1525–1534, 2022. DOI: `10.1093/jamia/ocac093`.

[16]   A. G. Gray and A. W. Moore, "Nonparametric density estimation: Toward computational tractability," in *Proceedings of the 2003 SIAM International Conference on Data Mining*, Society for Industrial and Applied Mathematics, 2003. DOI: `10.1137/1.9781611972733.19`.

[17]   H. Hachiya, M. Sugiyama, and N. Ueda, "Importance-weighted least-squares probabilistic classifier for covariate shift adaptation with application to human activity recognition," *Neurocomputing*, vol. 80, pp. 93–101, 2012. DOI: `10.1016/j.neucom.2011.09.016`.

[18]   C. Herzog, E. Petersen, and P. Rostalski, "Iterative approximate nonlinear inference via Gaussian message passing on factor graphs," *IEEE Control Systems Letters*, vol. 3, pp. 978–983, 4 2019. DOI: `10.1109/LCSYS.2019.2919260`.

[19]   T. R. Hoens, R. Polikar, and N. V. Chawla, "Learning from streaming data with concept drift and imbalance: An overview," *Progress in Artificial Intelligence*, vol. 1, no. 1, pp. 89–101, 2012. DOI: `10.1007/s13748-011-0008-0`.

[20]   D. G. Horvitz and D. J. Thompson, "A generalization of sampling without replacement from a finite universe," *Journal of the American Statistical Association*, vol. 47, no. 260, pp. 663–685, 1952. DOI: `10.1080/01621459.1952.10483446`.

[21] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," in *Proceedings of the 32nd International Conference on Machine Learning*, F. Bach and D. Blei, Eds., vol. 37, Lille, France: PMLR, 2015, pp. 448–456. [Online]. Available: `http://proceedings.mlr.press/v37/ioffe15.html`.

[22] C. Ionita, *Makima piecewise cubic interpolation*, C. Moler, Ed., 2019. [Online]. Available: `https://blogs.mathworks.com/cleve/2019/04/29/makima-piecewise-cubic-interpolation/`.

[23] T. Kanamori, S. Hido, and M. Sugiyama, "A least-squares approach to direct importance estimation," *Journal of Machine Learning Research*, vol. 10, pp. 1391–1445, 2009. [Online]. Available: `http://jmlr.org/papers/v10/kanamori09a.html`.

[24] J. D. Y. Kang and J. L. Schafer, "Demystifying double robustness: A comparison of alternative strategies for estimating a population mean from incomplete data," *Statistical Science*, vol. 22, no. 4, 2007. DOI: `10.1214/07-sts227`.

[25] W. M. Kouw and M. Loog. "An introduction to domain adaptation and transfer learning." arXiv: `1812.11806 [cs.LG]`. (2018).

[26] B. Krawczyk and P. Skryjomski, "Cost-sensitive perceptron decision trees for imbalanced drifting data streams," in *Machine Learning and Knowledge Discovery in Databases*, Springer International Publishing, 2017, pp. 512–527. DOI: `10.1007/978-3-319-71246-8_31`.

[27] N. Langrené and X. Warin, "Fast and stable multivariate kernel density estimation by fast sum updating," *Journal of Computational and Graphical Statistics*, vol. 28, no. 3, pp. 596–608, 2019. DOI: `10.1080/10618600.2018.1549052`.

[28] B. K. Lee, J. Lessler, and E. A. Stuart, "Weight trimming and propensity score weighting," *PLoS ONE*, vol. 6, no. 3, G. Biondi-Zoccai, Ed., e18174, 2011. DOI: `10.1371/journal.pone.0018174`.

[29] Z. Li *et al.*, "Incremental learning imbalanced data streams with concept drift: The dynamic updated ensemble algorithm," *Knowledge-Based Systems*, vol. 195, p. 105 694, 2020. DOI: `10.1016/j.knosys.2020.105694`.

[30] J. Lin, "Divergence measures based on the Shannon entropy," *IEEE Transactions on Information Theory*, vol. 37, no. 1, pp. 145–151, 1991. DOI: `10.1109/18.61115`.

[31] L. Ljung, *System Identification: Theory for the User*, Second Edition. Prentice Hall, 1999.

[32] H.-A. Loeliger *et al.*, "On sparsity by NUV-EM, Gaussian message passing, and Kalman smoothing," in *2016 Information Theory and Applications Workshop (ITA)*, IEEE, 2016. DOI: `10.1109/ita.2016.7888168`.

[33] S. Majumder *et al.*, "Statistical inference based on a new weighted likelihood approach," *Metrika*, vol. 84, no. 1, pp. 97–120, 2020. DOI: `10.1007/s00184-020-00778-y`.

[34] J. D. Mitchell *et al.*, "Inverse probability of treatment weighting (propensity score) using the military health system data repository and national death index," *Journal of Visualized Experiments*, no. 155, 2020. DOI: `10.3791/59825`.

[35] T. A. O'Brien, W. D. Collins, S. A. Rauscher, and T. D. Ringler, "Reducing the computational cost of the ECF using a nuFFT: A fast and objective probability density estimation method," *Computational Statistics & Data Analysis*, vol. 79, pp. 222–234, 2014. DOI: `10.1016/j.csda.2014.06.002`.

[36] T. A. O'Brien *et al.*, "A fast and objective multidimensional kernel density estimation method: fastKDE," *Computational Statistics & Data Analysis*, vol. 101, pp. 148–160, 2016. DOI: `10.1016/j.csda.2016.02.014`.

[37] J. Oh *et al.*, "Relaxed parameter sharing: Effectively modeling time-varying relationships in clinical time-series," in *Proceedings of the 4th Machine Learning for Healthcare Conference*, F. Doshi-Velez *et al.*, Eds., vol. 106, Ann Arbor, Michigan: PMLR, 2019, pp. 27–52. [Online]. Available: http://proceedings.mlr.press/v106/oh19a.html.

[38] J. Quiñonero-Candela, M. Sugiyama, N. D. Lawrence, and A. Schwaighofer, *Dataset shift in machine learning*. MIT Press, 2009.

[39] J. M. Robins, M. Á. Hernán, and B. Brumback, "Marginal structural models and causal inference in epidemiology," *Epidemiology*, vol. 11, no. 5, pp. 550–560, 2000. [Online]. Available: https://www.jstor.org/stable/3703997.

[40] P. R. Rosenbaum and D. B. Rubin, "The central role of the propensity score in observational studies for causal effects," *Biometrika*, vol. 70, no. 1, pp. 41–55, 1983. DOI: 10.1093/biomet/70.1.41.

[41] S. Schneider *et al.*, "Improving robustness against common corruptions by covariate shift adaptation," in *Advances in Neural Information Processing Systems*, H. Larochelle *et al.*, Eds., vol. 33, Curran Associates, Inc., 2020, pp. 11 539–11 551. [Online]. Available: https://proceedings.neurips.cc/paper/2020/file/85690f81aadc1749175c187784afc9ee-Paper.pdf.

[42] J. Schoukens and L. Ljung, "Nonlinear system identification: A user-oriented roadmap," *IEEE Control Systems Magazine*, vol. 39, no. 6, pp. 28–99, 2019. DOI: https://doi.org/10.1109/MCS.2019.2938121.

[43] D. W. Scott, *Multivariate Density Estimation: Theory, Practice, and Visualization*. John Wiley & Sons, Inc., 1992, ISBN: 9780471547709. DOI: 10.1002/9780470316849.

[44] K. D. Sebesta and N. Boizot, "A real-time adaptive high-gain EKF, applied to a quadcopter inertial navigation system," *IEEE Transactions on Industrial Electronics*, vol. 61, no. 1, pp. 495–503, 2014. DOI: 10.1109/tie.2013.2253063.

[45] H. Shimodaira, "Improving predictive inference under covariate shift by weighting the log-likelihood function," *Journal of Statistical Planning and Inference*, vol. 90, no. 2, pp. 227–244, 2000. DOI: 10.1016/s0378-3758(00)00115-4.

[46] A. Simonetto *et al.*, "Time-varying convex optimization: Time-structured algorithms and applications," *Proceedings of the IEEE*, vol. 108, no. 11, pp. 2032–2048, 2020. DOI: 10.1109/jproc.2020.3003156.

[47] M. Sugiyama and M. Kawanabe, *Machine Learning in Non-Stationary Environments*. Cambridge, Massachusetts: The MIT Press, 2012.

[48] M. Sugiyama, M. Krauledat, and K.-R. Müller, "Covariate shift adaptation by importance weighted cross validation," *Journal of Machine Learning Research*, vol. 8, pp. 985–1005, 2007. [Online]. Available: https://dl.acm.org/doi/10.5555/1314498.1390324.

[49] M. Sugiyama *et al.*, "Direct importance estimation for covariate shift adaptation," *Annals of the Institute of Statistical Mathematics*, vol. 60, no. 4, pp. 699–746, 2008. DOI: 10.1007/s10463-008-0197-x.

[50] F. Thoemmes and A. D. Ong, "A primer on inverse probability of treatment weighting and marginal structural models," *Emerging Adulthood*, vol. 4, no. 1, pp. 40–59, 2015. DOI: 10.1177/2167696815621645.

[51] V. Torczon, "On the convergence of pattern search algorithms," *SIAM Journal on Optimization*, vol. 7, no. 1, pp. 1–25, 1997. DOI: 10.1137/s1052623493250780.

[52] Y. Tsuboi *et al.*, "Direct density ratio estimation for large-scale covariate shift adaptation," *Journal of Information Processing*, vol. 17, pp. 138–155, 2009. DOI: 10.2197/ipsjjip.17.138.

[53]   B. C. Wallace and I. J. Dahabreh, "Improving class probability estimates for imbalanced data," vol. 41, no. 1, pp. 33–52, 2013. DOI: 10.1007/s10115-013-0670-6.

[54]   S. Wang, L. L. Minku, and X. Yao, "A systematic study of online class imbalance learning with concept drift," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 29, no. 10, pp. 4802–4821, 2018. DOI: 10.1109/tnnls.2017.2771290.

[55]   X. Wang, C. van Eeden, and J. V. Zidek, "Asymptotic properties of maximum weighted likelihood estimators," *Journal of Statistical Planning and Inference*, vol. 119, no. 1, pp. 37–54, 2004. DOI: 10.1016/s0378-3758(02)00410-x.

[56]   G. Widmer and M. Kubat, "Learning in the presence of concept drift and hidden contexts," *Machine Learning*, vol. 23, no. 1, pp. 69–101, 1996. DOI: 10.1023/a:1018046501280.

[57]   W. Wilkinson, P. Chang, M. Andersen, and A. Solin, "State space expectation propagation: Efficient inference schemes for temporal Gaussian processes," in *Proceedings of the 37th International Conference on Machine Learning*, H. Daumé III and A. Singh, Eds., vol. 119, PMLR, 2020, pp. 10 270–10 281. [Online]. Available: http://proceedings.mlr.press/v119/wilkinson20a.html.

CHAPTER 7

SENSOR FUSION FOR INFERRING P_{MUS}
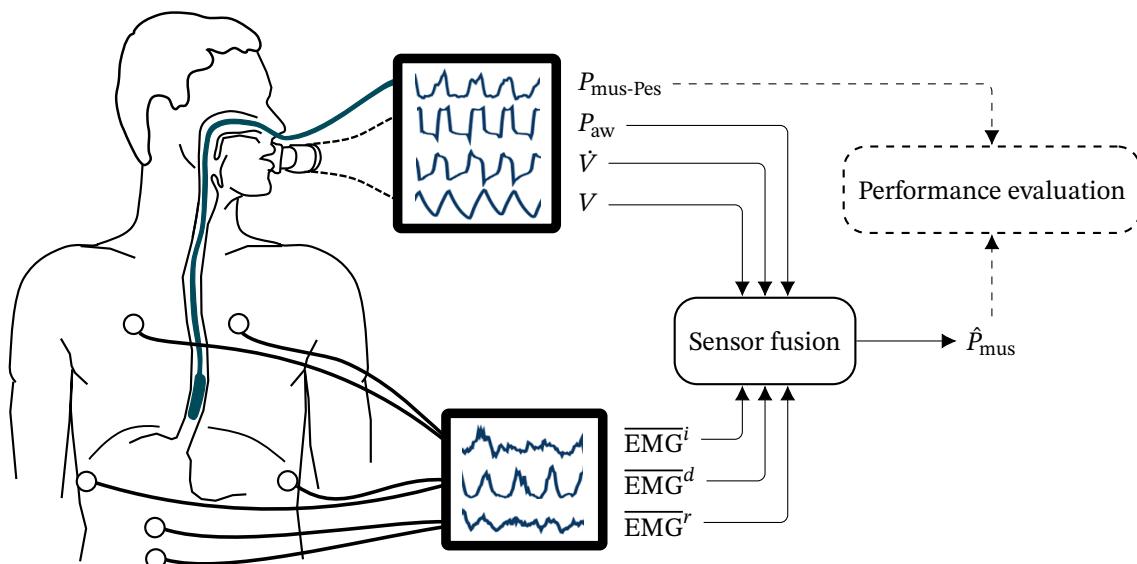
Figure 7.1 – An illustration of the estimation problem to be solved in this chapter. The airway pressure P_{aw} and airflow $\dot{V}$ are standard measurements provided by any mechanical ventilator, and V is obtained by running integration of $\dot{V}$. Three EMG channels are recorded as described in section 3.3.3, cardiac artifacts removed, and their envelope signals $\overline{\text{EMG}}$ calculated. From these measurements, the pressure P_{mus} generated by the respiratory muscles shall be estimated continuously. For validation purposes, esophageal pressure P_{es} was also measured using an esophageal catheter and balloon. From this, the gold standard estimate $P_{\text{mus-Pes}}$ of P_{mus} is calculated, which will be used for performance evaluation. The drawing of the human torso has been created and kindly provided by Jan Graßhoff.

Contents

DURING RESPIRATION, the respiratory muscles generate a pressure (P_{mus}), which acts on the pleural cavity to force air in and out of the lungs. This penultimate chapter of this thesis addresses the problem of estimating that pressure in mechanically ventilated patients based on surface EMG measurements. In recent years, it has become more and more apparent that monitoring P_{mus} in mechanically ventilated patients is crucial to prevent diaphragm atrophy due to misuse, diaphragm injury due to overuse, self-inflicted lung injury due to excessive strain on the lungs, and patient-ventilator asynchrony [13, 34, 78]. The currently available gold standard measurement, esophageal pressure P_{es}, requires the (invasive) insertion of an esophageal catheter, a procedure that is uncomfortable for the patient, error-prone, and expensive. Surface electromyography of the respiratory muscles represents one possible solution for monitoring the patient's respiratory activity noninvasively, but it does not directly yield a measure of the *pressure* generated by the patient due to patient-specific signal scaling factors. The aim of the present chapter is to present an algorithm that estimates the pressure P_{mus} generated by a patient under mechanical ventilation by fusing surface electromyographic measurements with pneumatic measurements from the mechanical ventilator.

This chapter draws heavily on the physiological background, mathematical models, and inference algorithms developed in the previous chapters. A family of probabilistic inference algorithms is derived, which is tailored to the specific characteristics and challenges posed by this particular estimation problem of high clinical importance. Section 7.1 opens the chapter with a precise statement of the problem at hand, and with brief overviews of existing research on the observation of respiratory activity in mechanically ventilated patients, both with and without EMG measurements. The subsequent section 7.2 provides an overview of the approach that is pursued in the following, as well as the challenging properties of the arising inference problem. In section 7.3, physiological models will be developed, identified using real patient data, and their performance compared. These models are then used as part of an algorithm to identify the respiratory pressure generated by a patient under mechanical ventilation, based on both pneumatic and electromyographic data. This algorithm is presented in section 7.4. It is then validated using intensive care unit patient data by comparison with a gold standard reference signal in section 7.5, before section 7.6 concludes the chapter with a discussion and an outlook.[1]

7.1 Problem definition & state of the art

Section 7.1.1 briefly defines the exact estimation problem under consideration in this chapter. The following section 7.1.2 then provides a short overview of existing methods for quantifying the respiratory effort of patients under mechanical ventilation, based on

---

[1]Aspects of this chapter have been the subject of a number of previous journal and conference publications as well as patent applications [26–29] (co-)authored by the author. Graßhoff, Petersen, et al. [36] discuss an algorithm to remove cardiac artifacts from the P_{es} signal, and Graßhoff, Petersen, et al. [35] describe an automated method to obtain $P_{\mathrm{mus\text{-}Pes}}$ from P_{es} measurements. This signal is used as a ground truth validation signal in this chapter. Petersen *et al.* [62] describe, for the first time, a model-based method for estimating P_{mus} from surface EMG measurements. The method presented in this chapter represents a significantly more developed version of the same fundamental approach, which is also validated much more comprehensively. Finally, Graßhoff, Petersen, et al. [37] present an extensive quantitative analysis of the relationship between respiratory surface EMG measurements and P_{mus}. The specific analyses and algorithms presented here have not been described elsewhere.

pneumatic measurements. Since the solution approach pursued in this chapter is based on respiratory surface EMG measurements, section 7.1.3 then specifically examines the existing literature on the use of surface EMG measurements for respiratory monitoring.

7.1.1 Problem definition

The problem under consideration in this chapter is illustrated in fig. 7.1: given both pneumatic as well as electromyographic measurements obtained from a patient under mechanical ventilation, the task is to identify the pressure P_{mus} generated by the patient at each time. As pneumatic measurements, the airway pressure P_{aw}, the airflow $\dot{V}$ in and out of the lungs, as well as a measure V of the volume currently contained in the lung are available. As electromyographic measurements, the envelope signals of differential measurements of intercostal muscle activity ($\overline{\text{EMG}}^i$), diaphragmatic muscle activity ($\overline{\text{EMG}}^d$), and rectus abdominis muscle activity ($\overline{\text{EMG}}^r$) are obtained. The obtained estimate $\hat{P}_{\text{mus}}$ is verified by comparison with a gold standard reference signal $P_{\text{mus-Pes}}$, which is obtained from an esophageal pressure measurement using an esophageal catheter.

7.1.2 Pneumatic measures of respiratory effort

For more than a century [7, 8] it has been known that pressure deflections in the *pleural cavity* reflect, in part, the pressure generated by the respiratory muscles. Since the measurement of that pressure requires a puncture of the pleural cavity, the measurement of *esophageal* pressure P_{es} has been employed as a surrogate measurement in respiratory research for many decades [21, 23], although it has only more recently been translated into clinical practice [5]. To measure esophageal pressure, an esophageal catheter with an attached balloon and pressure transducer is administered to the patient. The balloon must subsequently be precisely positioned such that the pressure inside it approximates the pressure in the pleural cavity well [5, 18]. The resulting esophageal pressure measurement P_{es} reflects not only the pressure generated by the respiratory muscles but also the elastic recoil pressure of the chest wall due to variations in lung volume, as well as cardiac artifacts [5]. Thus, further (model-based) signal processing is required to reliably extract an estimate $P_{\text{mus-Pes}}$ of the pressure generated by the respiratory muscles from P_{es}, see, e.g., Graßhoff, Petersen, et al. [36] and Graßhoff, Petersen, et al. [35]. While esophageal pressure measurement represents the current clinical gold standard for measuring the pressure generated by the respiratory muscles, its drawbacks are readily apparent: the use of an esophageal catheter is invasive and uncomfortable for the patient, and the correct positioning of the esophageal balloon requires a skilled operator [39, 60].

Esophageal Pressure

Owing to the shortcomings of P_{es} measurements, noninvasive alternatives have been sought after for decades, and many potential solutions have been proposed. Virtually all of them are fundamentally based on the *respiratory equation of motion*[2]

Equation of Motion

$$P_{\text{aw},k} + P_{\text{mus},k} = R\dot{V}_k + EV_k + P_0, \tag{7.1}$$

which can be derived from the classical linear single-compartment model of the respiratory system (see section 3.2). The airway pressure P_{aw}, the airflow $\dot{V}$, and the lung volume V are

---

[2]Throughout this thesis, we define P_{mus} as the negative of the pressure generated by the respiratory muscles, i.e., it is positive when a negative pressure is generated (typically during inspiration). This definition enables an intuitive interpretation in terms of respiratory *effort* exerted by the patient.

usually available as measurements from the mechanical ventilator, whereas the pressure P_{mus} generated by the patient, the airway resistance R, the respiratory system elastance E, and the constant offset P_0 are unknown and patient-specific, and thus must be estimated. If estimates for E, R, and P_0 are available (methods to obtain such estimates will be discussed in the next paragraph), eq. (7.1) can be rearranged to yield an estimate of P_{mus} since

$$P_{\text{mus},k} = R\dot{V}_k + EV_k + P_0 - P_{\text{aw},k}. \tag{7.2}$$

In the absence of any additional measurements, assumptions, or constraints, eq. (7.1) is clearly under-determined due to the unknown source term $P_{\text{mus},k}$. Possibly the simplest proposed solution is to estimate E and R during phases in which the patient is known (or expected) to be fully *passive*, i.e., $P_{\text{mus}} \equiv 0$, and then reuse those estimates during phases with patient activity [43, 89]. In the case that $P_{\text{mus}} \equiv 0$, eq. (7.1) reduces to a simple linear regression problem that can be solved for R, E, and P_0 using, e.g., ordinary least squares. Patient passivity is usually ensured either because the patient is sedated or by increasing the pressure support to a high level for a brief period of time [43, 89]. This method, however, has recently been found to yield inaccurate estimates of P_{mus} [61]. It appears likely that at least the following two closely related factors contribute to the observed inaccuracy:

Passive Patient

1. The model (7.1) is subject to significant model errors in describing the respiratory system [9]. Thus, even if some set of "optimal" model parameters were known,[3] the reconstructed P_{mus} signal would still be subject to significant errors.

2. Since the model (7.1) represents a stark simplification of reality, its model parameters must not be understood as a precise quantification of actual physical properties of the respiratory system [9]. Rather, they represent a linearization of a complex, nonlinear, time-varying system and are thus dependent on the patient–ventilator system's current set point. In other words, model parameters determined during phases of patient passivity are likely not directly transferable to other set points, i.e., other ventilator settings or patient activity levels.

An alternative and very popular solution approach that partially alleviates at least the second problem is to use *airway occlusions* [14, 31]. An airway occlusion denotes a brief intervention during which no air can flow to or from the patient, i.e., $\dot{V} = 0$ is enforced. Occlusions can be performed either at the end of an inspiration [31, 61, 70], at the end of an expiration [14], or very briefly at the beginning of a new inspiration ("P0.1") [79]. It has been demonstrated that the airway pressure drop during such occlusions is well-correlated with measures of average respiratory effort [14, 70, 79]. Thus, occlusion maneuvers are well suited to detect excessively high inspiratory effort [14, 61, 70], i.e., classify patients as at-risk or not. At the same time, however, occlusion maneuvers do *not* enable accurate measurement of P_{mus} due to strong measurement bias and variance [14, 61, 70]. One source of errors is expiratory activity, which biases the occlusion pressure measurement, and which cannot be reliably detected during end-inspiratory occlusions [77]. More fundamentally, occlusion-based methods do not enable a *continuous* measurement of P_{mus}, or the identification of a model of the respiratory system. They are limited to occasional measurements at discrete time points and are thus unsuitable for, e.g., proportional support ventilation [46, 65]. Moreover,

Airway
Occlusions

---

[3]A definition of optimality in this context will be the subject of later sections of this chapter.

the (recurring) use of long occlusion maneuvers is, of course, uncomfortable and distressful for the patient.

As an alternative, although also occlusion-based approach, Younes *et al.* [91] proposed to approximate the passive elastance E from the airway pressure measured during brief, end-inspiratory occlusions in the proportional assist ventilation (PAV) mode. In this ventilation mode, the patient is very likely to be passive at the end of inspiration, thus effectively guaranteeing that the effect of E in eq. (7.1) can be observed in isolation. With E known, R can then also be estimated during the same time points, still assuming $P_{\text{mus}} \equiv 0$ at these times [90]. This estimation procedure is commercially available in the *PAV+ ventilation software* (Puritan Bennett, Medtronic); however, it is only applicable in this specific ventilation mode [92]. From E and R, a continuous estimate of P_{mus} can then in principle be calculated using eq. (7.2). This signal has been shown to attain only limited agreement with gold-standard $P_{\text{mus-Pes}}$, however [6]. Similarly to the method of Younes *et al.* [91], Kondili *et al.* [48] proposed a method that automatically identifies expiratory segments during which the patient appears to be passive and estimates R and E during these segments. The pressure P_{mus} can then, again, be estimated continuously using eq. (7.2). This method is not limited to the PAV+ ventilation mode, and its results showed acceptable agreement with the gold standard $P_{\text{mus-Pes}}$, although the method's limits of agreement were still relatively high. While proportional ventilation can benefit the patient even if E and R (and, thus, P_{mus}) cannot be measured accurately [93], this discrepancy becomes problematic once a target level of P_{mus} is used to automatically titrate the support level [6, 20].

PAV+

Over the course of the last 15 years, various researchers have attempted to mitigate the underdetermination of (7.1) following a different approach: by making *assumptions on the time course of P_{mus}*. This can be understood as considering P_{mus} in eq. (7.1) an unknown disturbance and assuming a specific structure of the disturbance's covariance matrix. Redmond *et al.* [68], Saatci and Akan [71], and Vicario *et al.* [84] all assume simple, parametric models for the time course of P_{mus} in individual breaths; Scheel *et al.* [72] assume a simple, second-order dynamic oscillator model for P_{mus}. The parameters of these P_{mus} models, along with R and E, are then fit on each individual breath using least squares estimation or recursive filtering. In a slightly different vein, Vicario *et al.* [83, 84] describe a method that assumes simple monotonicity constraints on the time course of P_{mus}, thus not prescribing a particular (polynomial) time course of P_{mus}. All of these approaches generally result in nonlinear estimation problems. A large number of similar methods has been compared by Rehm *et al.* [69] with respect to their accuracy in estimating the elastance E in different patients and under varying conditions; the authors find the polynomial P_{mus} model proposed by Redmond *et al.* [68] to perform best in their comparison. Importantly, this family of methods is not dependent on any specific types of maneuvers being performed, and is thus (in theory) applicable to ventilation data obtained using any mechanical ventilator in any ventilation mode. To the author's knowledge, none of these methods has been evaluated in a sufficiently large clinical cohort with respect to their accuracy in estimating P_{mus}. Owing to their very nature, all of these methods are vulnerable to deviations of P_{mus} from the stated assumptions, e.g., due to irregular breathing, patient-ventilator asynchrony, or expiratory muscle activity. Another drawback is that they can result in high breath-by-breath parameter estimate variability since they estimate parameters anew on each breath [69].

Assumptions on P_{mus}

7.1.3 Electromyographic measures of respiratory effort

Electromyographic measurements have been employed for respiratory monitoring in humans for at least 70 years. In 1951, Tokizane *et al.* [81] reported on the use of needle EMG measurements of various respiratory muscles for investigating the neural control of respiration. Possibly the first description of esophageal EMG measurements — obtained by inserting a catheter equipped with specialized electrodes into the esophagus — has been provided by Petit *et al.* [63] in 1960. And, finally, Prechtl *et al.* [66] may have contributed the first description of a surface (sometimes also called *transcutaneous*) EMG measurement of the respiratory muscles in 1977. Significant advances in signal processing methodology [25, 53, 58] (also see chapters 4 and 5) and electronics in recent decades have enabled more and more successful translation of EMG-based methods into clinical research and application. The *esophageal* EMG measurement technique was further refined, popularized, and ultimately commercialized by Sinderby *et al.* [75], resulting in the Servo-i ventilator (Getinge) and the neurally adjusted ventilatory assist (NAVA) ventilation mode [74]. Using this ventilation mode, patients receive ventilatory support in proportion to their own breathing efforts (as measured by esophageal EMG). *Surface* EMG measurements — generally preferable due to their noninvasiveness and capability to monitor respiratory muscles other than the diaphragm, but also more challenging to process — have now become a standard tool for monitoring patient-ventilator interactions [19, 49, 85], and respiratory activity of mechanically ventilated patients more generally [2, 47, 51].

Esophageal EMG

Surface EMG

Aside from monitoring the timing and relative levels of a patient's respiratory activity, an increasing number of publications address the problem of *inferring* P_{mus} from EMG measurements. To this end, some kind of patient-specific neuromechanical conversion must be performed, mapping electromyographic activity to mechanical pressure generation. The first steps in this direction were undertaken by Bellani et al. [10, 11], who analyzed the relationship between $P_{\text{mus-Pes}}$ and esophageal EMG measurements. They found the two signals to be highly correlated and related by a simple conversion factor which could be determined during end-expiratory occlusions. During these occlusions (see eq. (7.1)),

Inferring P_{mus} from EMG

$$P_{\text{aw},k} = P_{\text{mus},k} + P_0,$$

and assuming

$$P_{\text{mus},k} = \alpha \overline{\text{EMG}}_k + P_{\text{EMG}}, \tag{7.3}$$

the conversion factor α can be easily determined using, e.g., linear regression, or considering the ratio of the maximum amplitude of the P_{aw} and EMG signals during the occlusion. With α estimated, eq. (7.16) can then be employed to obtain a continuous estimate of P_{mus} throughout normal ventilation. Later, Jansen *et al.* [45] analyzed the same setting — attempting to estimate P_{mus} from esophageal EMG — and found single occlusion maneuvers to yield highly variable estimates of α, i.e., the repeatability was low. Aggregating data from multiple occlusions improved repeatability. Using the same method as in their earlier work on esophageal EMG, Bellani *et al.* [12] analyzed the estimation of $P_{\text{mus-Pes}}$ from *surface* EMG measurements, again based on identifying α during end-expiratory occlusion maneuvers. Their results indicated, for the first time, that estimating $P_{\text{mus-Pes}}$ completely noninvasively in this way is, indeed, feasible.

Promising as these results are, the method of Bellani *et al.* [12] also exhibits a number of opportunities for further enhancement. Most importantly, it only makes use of a tiny fraction of the available data: firstly because parameter estimation is performed only during occlusions, ignoring all other ventilation phases, and secondly because P_{mus} is estimated only from the EMG, ignoring all other available measurement signals ($P_{\text{aw}}, \dot{V}, V$). The low repeatability observed by Jansen *et al.* [45], the need for aggregated analyses as performed by Bellani *et al.* [12], and sometimes rather noisy P_{mus} estimates (since P_{mus} is directly proportional to the — often noisy — EMG measurements) are all consequences of this methodological limitation. It appears plausible that a well-conceived sensor fusion procedure should be capable of combining all available information in an efficient way to obtain a more reliable estimate of P_{mus}, all while not relying on any particular kind of identification maneuver being performed and still entirely noninvasively. In particular, it should be possible to combine the purely pneumatic approaches discussed in the previous section, which are based on the equation of motion (7.1), with the EMG-based approaches discussed in this section. Precisely such a procedure is proposed, evaluated, and discussed in the remainder of this chapter.

Muscle Force Estimation Finally, notice that there is, of course, an expansive and old body of research on *muscle force estimation* from surface EMG measurements in non-respiratory settings [24, 41, 42, 54, 56, 59, 64]. These developments are largely analogous to the present respiratory setting: joint biomechanics take the role of respiratory mechanics, and unknown internal joint forces take the role of the unknown pressure P_{mus}. As this connection appears to be largely unexploited, it may serve as a valuable inspiration. Reliable and continuous estimation of internal forces from surface EMG measurements remains a challenging problem, however, due to the complexity of the involved biomechanics.

7.2 Outline: the estimation problem, its characteristics, and the proposed framework

Briefly recalling section 7.1.1 and fig. 7.1, the problem to be solved is this: given measurement time series of the airway pressure P_{aw}, airflow $\dot{V}$ and its integral, the current lung volume V, and one or multiple EMG channels[4] $\overline{\text{EMG}}$, estimate the time course of the pressure contribution P_{mus} generated by the respiratory muscles of the patient.

7.2.1 Dataset

Throughout this chapter, all analyses are based on study dataset B, described in more detail in section 3.3.3, and based on data recorded and kindly provided by Bellani *et al.* [12]. Briefly, eleven recordings from nine different intensive care patients are used, which comprise standard pneumatic measurements ($P_{\text{aw}}, \dot{V}, V$), P_{es}-based P_{mus} estimates ($P_{\text{mus-Pes}}$) as well as three channels of raw respiratory surface EMG measurements (EMG^i, EMG^d, EMG^r). The EMG signals are preprocessed using a simple gating algorithm for removing cardiac interference, and the weighted moving average absolute value is calculated to yield

---

[4]While in this thesis only surface measurements are considered explicitly, the methodology described should apply equally well to invasive measurements, i.e., measurements of EAdi.

a simple envelope signal. All data are synchronized and sampled at a frequency of $f_s = 100\,\text{Hz}$. For more details, refer to section 3.3.3.

7.2.2 Characteristics of the estimation problem

The stated estimation problem is characterized by several important properties that render it challenging, both practically as well as theoretically:

Unsupervised estimation. There are patient-specific parameters to be inferred anew in each patient. No reference measurements are available for these, i.e., the estimation problem is unsupervised.

Time-varying estimation. The properties of the patient's respiratory system change over time, thus necessitating an adaptive estimation procedure.

High inter-patient variability. Patients present with different medical and physical conditions, treatments, and respiratory behavior, see table 3.2 for an overview of patient characteristics and fig. 7.2 for extracts from two exemplary and quite different recordings.

Outliers and artifacts. Measurements are taken from real human patients under intensive care and are thus subject to a large number of unknown disturbances and measurement artifacts. Any successful estimation procedure must be highly robust to such disturbances.

Model mismatch. The airways, the lungs, the pleural cavity, the rib cage, and the various respiratory muscles form a highly complex nonlinear system with many interacting components and degrees of freedom, very few of which are observable using the available measurements. Thus, any model based on these measurements is highly likely to be subject to model mismatch. Moreover, the presence of strong outliers and various artifacts (see the previous point) renders the formulation of a correct noise model very challenging, thus further contributing to model mismatch.

Covariate shift. The recorded datasets are typically subject to covariate shift in the following sense: Most of the time, the patient breathes normally and repetitively, with little breath-by-breath variation in the recorded waveforms, exploring only a small region of the data space. Occasionally, specific respiratory maneuvers are performed — e.g., occlusions or inspiratory-hold maneuvers — or the patient breathes irregularly in some sense. These sections are rare, but they may hold important (complimentary) information about the patient's respiratory system. Thus, $p_{\text{train}}(x) = p_{\text{train}}\left(\dot{V}_k, V_k, \overline{\text{EMG}}_k^i, \overline{\text{EMG}}_k^d, \overline{\text{EMG}}_k^r\right)$ is highly imbalanced (with an emphasis on frequent, "normal" breaths). On the other hand, $p_{\text{target}}(x)$ might, for instance, be something like a *uniform* distribution over a suitably selected region of the data space: we may want the learned model to describe the behavior of the patient's respiratory system *equally well* in a selected range of conceivable situations, not simply the most frequent breath type. (Otherwise, we risk overfitting to standard breaths and ignoring additional information granted by occasional unusual breaths.)
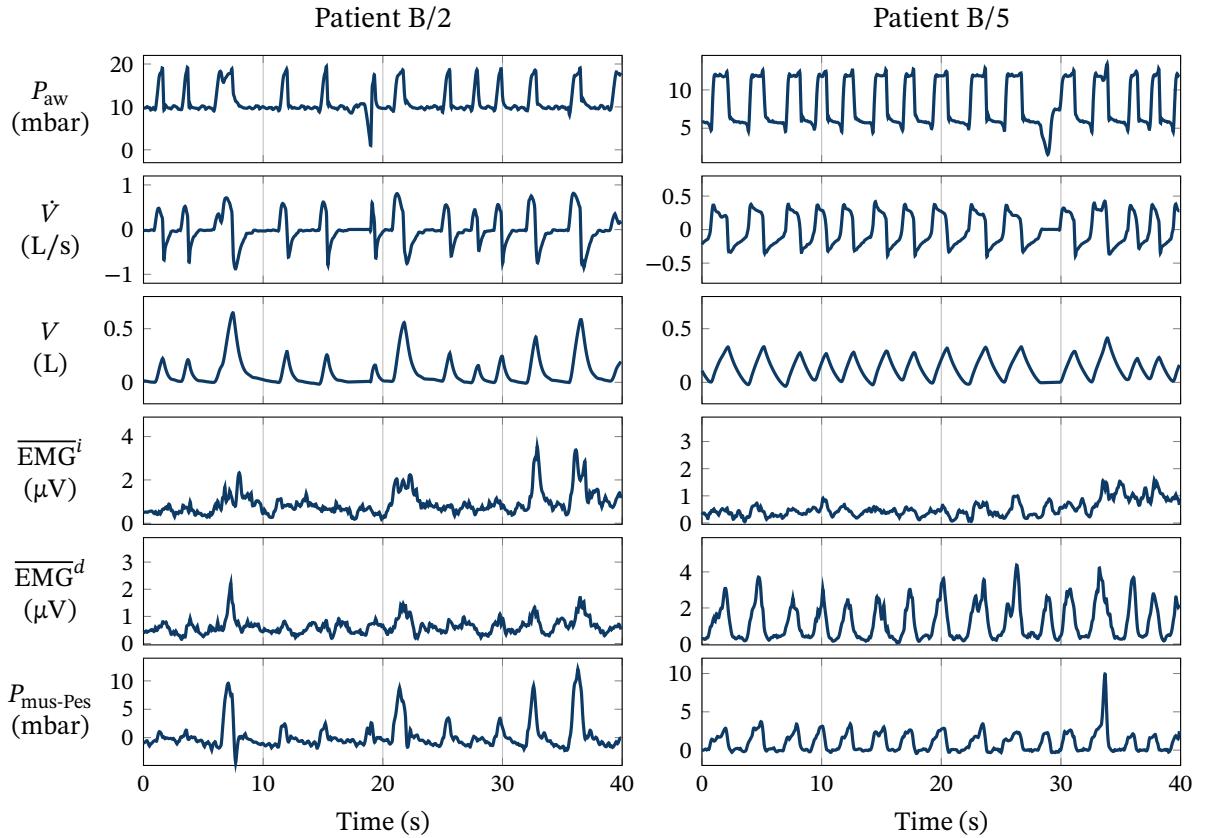
Figure 7.2 – Two exemplary extracts of recordings from two patients in study B. The two patients show widely differing breathing patterns and signal waveforms.

Small evaluation study sample size. Research studies with new measurement methodology in the intensive care setting are notoriously challenging to conduct, thus limiting the typically available study sample size. While the small patient population studied here (nine patients) is certainly not the upper limit, other groups studied populations of 41 [12] and 31 [45] patients, thus still clearly within the small-data regime. This limits the kind of analyses and model selection procedures that can be performed.

Medium-sized subject-level datasets. While certainly not in the realm of „Big Data", the datasets recorded for individual patients consist of some tens of thousands of measurement samples for at least seven measurement signals (P_{aw}, $\dot{V}$, V, $P_{mus\text{-}Pes}$, $\overline{EMG}^d$, $\overline{EMG}^i$, $\overline{EMG}^r$). This yields a dataset size that is large enough to cause computation time and memory efficiency to become relevant concerns, especially when considering prospective real-time application on an embedded medical device.
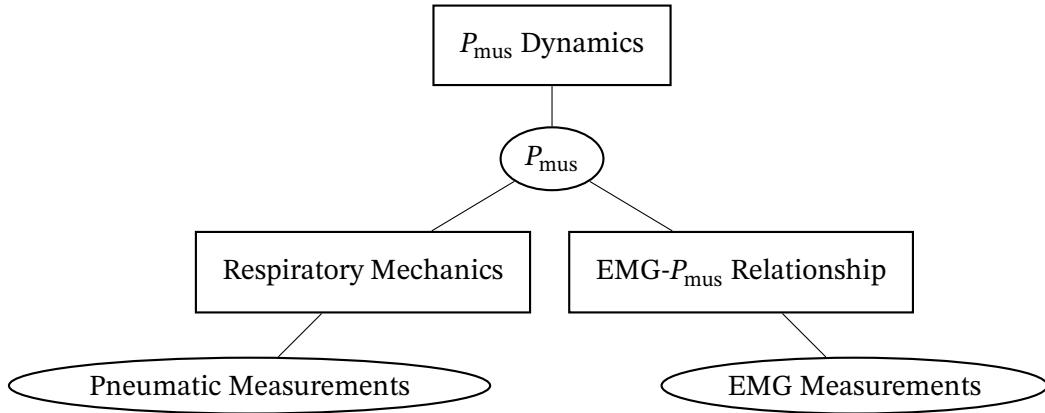
Figure 7.3 – A *factor graph* modeling the general structure of the sensor fusion problem considered in this chapter. Three sources of information about P_{mus} are exploited: (1) the pneumatic measurements P_{aw}, $\dot{V}$ and V, which are related to P_{mus} via a model of respiratory mechanics, (2) surface EMG measurements of the respiratory muscles, which are related to P_{mus} via a model of the EMG-P_{mus} relationship, and (3) a model of the dynamical behavior of the signal $P_{\mathrm{mus}}(t)$ itself. Ellipses denote variables and rectangles denote the constituting factors of the joint probability density function of these variables.

7.2.3 Overview of the proposed solution

As a solution to the described estimation problem, we propose a novel sensor fusion method that combines three sources of information about the $P_{\mathrm{mus}}(t)$ signal:

1. the pneumatic measurements P_{aw}, $\dot{V}$, and V, which are connected to P_{mus} via a model of respiratory mechanics,

2. the electromyographic envelope signals $\overline{\mathrm{EMG}}^i$, $\overline{\mathrm{EMG}}^d$, and $\overline{\mathrm{EMG}}^r$, which are connected to P_{mus} via an electro-mechanical coupling model, and

3. prior knowledge about the dynamical behavior of the signal $P_{\mathrm{mus}}(t)$ itself.

Figure 7.3 shows a *factor graph* [57] that visualizes the connections between these different sources of information. To increase robustness to outliers and artifacts, all signals are subjected to preprocessing, smoothing, and simple (mostly threshold-based) artifact rejection procedures. In order to counter the combination of model mismatch and dataset imbalance, we employ the importance-weighted, time-varying estimation scheme proposed in the previous chapter. Finally, an estimate $\hat{P}_{\mathrm{mus},k}$ is determined by exploiting all available information. Figure 7.4 shows a block diagram illustrating the main steps of the proposed estimation scheme. The following section will discuss the choice of models used for the three components of the sensor fusion procedure in detail before section 7.4 will describe the technical details of the estimation scheme.

7.3 Model selection & identification

We will now proceed to propose models for each of the three sub-components mentioned above and in fig. 7.3, namely, respiratory mechanics, the EMG–P_{mus} relationship, and the
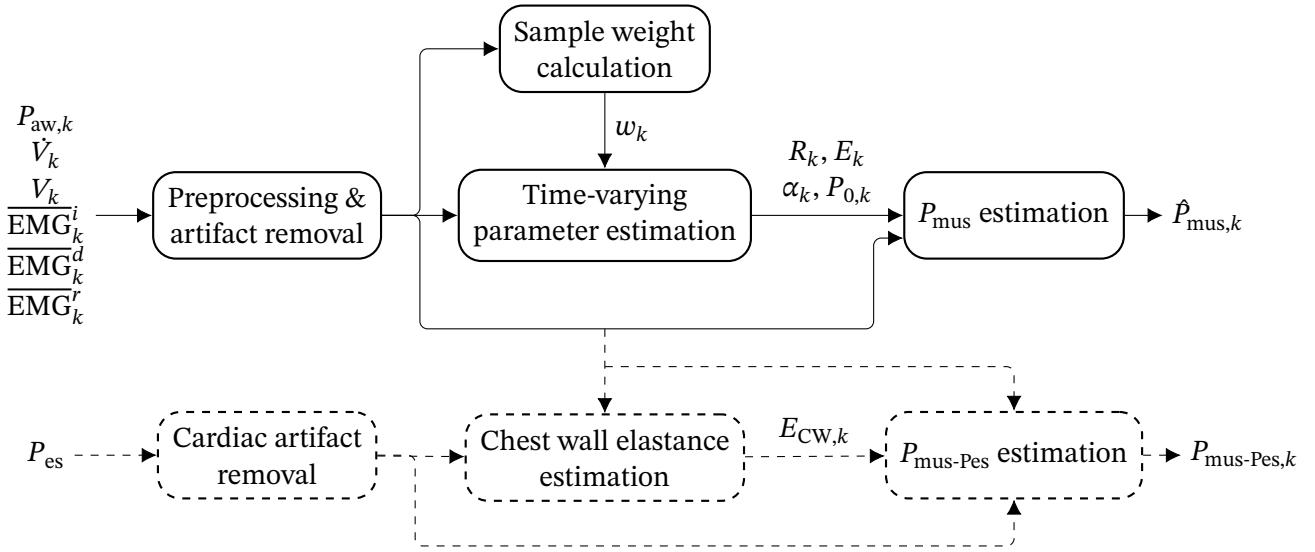
Figure 7.4 – A block diagram illustrating the main processing steps of the proposed, fully noninvasive estimation scheme. The lower, dashed part depicts the validation processing path (requiring invasive measurement of P_{es}).

dynamics of the P_{mus} signal. The capability of each model component to fit real patient measurements will be assessed, and several model (hyper-)parameters will be identified.

7.3.1 Modeling respiratory mechanics

The modeling of respiratory mechanics is, by far, the subcomponent that has received the most comprehensive treatment in the existing literature; a brief overview of the most prominent modeling choices was provided in section 3.2.3. Here, we will employ a simple single-compartment model with a linear-quadratic resistance and a linear elastance, as described in eqs. (3.6), (3.7) and (3.9) in section 3.2.3. The resulting respiratory equation of motion reads

$$P_{\mathrm{aw}} = R_1\dot{V} + R_2|\dot{V}|\dot{V} + EV - P_{\mathrm{mus}} + P_0. \tag{7.4}$$

Many more complex model variants were evaluated during the writing of this thesis but yielded similar or worse estimator performance. These variants are not further discussed here, in the interest of brevity. The only further model variant that will be evaluated in the

Quadratic
Elastance

following is the one with an additional, *quadratic* elastance term, i.e.,

$$P_{\mathrm{aw}} = R_1\dot{V} + R_2|\dot{V}|\dot{V} + E_1V + E_2V^2 - P_{\mathrm{mus}} + P_0. \tag{7.5}$$

To assess the capability of the proposed models of respiratory mechanics to describe real patient data, the model (7.4) was fit to all available study recordings by computing the ordinary least squares (OLS)–optimal parameter values in

$$P_{\mathrm{aw},k} + P_{\mathrm{mus\text{-}Pes},k} = R_1\dot{V}_k + R_2|\dot{V}_k|\dot{V}_k + EV_k + P_0 + \varepsilon_{\mathrm{RM},k}, \tag{7.6}$$

i.e., the values

$$\theta_{\mathrm{RM}}^* = \begin{pmatrix} R_1^* & R_2^* & E^* & P_0^* \end{pmatrix}^{\mathsf{T}}$$
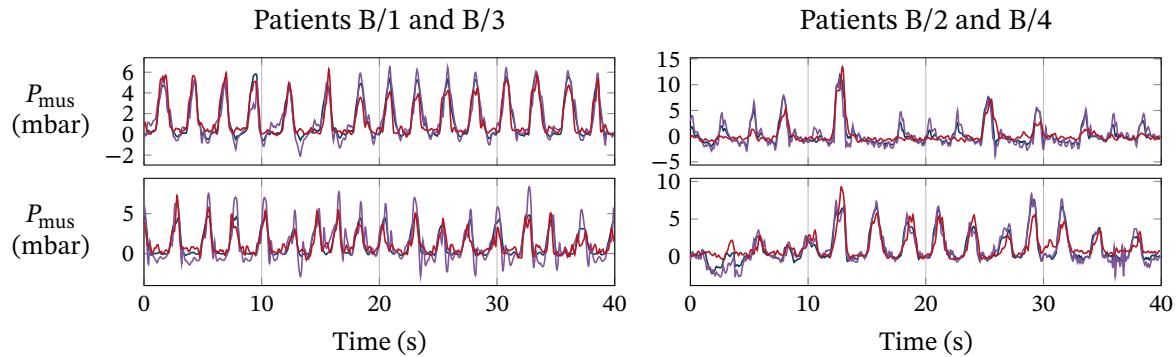
Figure 7.5 – Exemplary fits of the respiratory mechanics model (7.4) (purple) and the electro–mechanical model (7.8) (red) to $P_{\text{mus-Pes}}$ (blue).

that minimize the sum of the squared residuals $\varepsilon^2_{\text{RM},k}$ in eq. (7.6). In order to account for the (potential) temporal variability of the model parameters, fitting was performed on snippets of 30 s length, excluding any snippets that contained less than 20 s of usable data. In addition, to minimize the negative impact of measurement artifacts on the noise estimates, snippets were discarded if either of the following two conditions was met:

- There were extremely large *parameter value difference*s as compared to the previous snippet. More precisely, snippet j is rejected if a) there are more than three pairs of valid adjacent snippets $(\ell, \ell + 1)$ between which parameters can be compared, b) the 0.95–quantile of the absolute values of the parameter differences between the pair of snippets is larger than three times the median of those absolute differences in all parameters, and c) the absolute values of *all* parameter differences in snippet j (as compared to the previous snippet $j - 1$) are larger than the 0.95–quantile of the corresponding absolute parameter difference distribution.

- The *residuals* were exceptionally large. More precisely, snippet j is rejected if a) there are more than three snippets with more than 20 s of data, b) the 0.95–quantile of the mean absolute residual values within each snippet is larger than three times the median of those mean absolute residuals, and c) the mean absolute residual in snippet j is larger than the 0.95–quantile of the mean absolute residuals.

Figure 7.5 shows two exemplary snippets from two recordings and the results of the model fitting procedure. It is apparent that the model (7.4) is capable of achieving a reasonably good fit to these data. Figure 7.6 shows the model residuals $\varepsilon_{\text{RM},k}$ as a function of both $\dot{V}$ and V in all eleven available recordings, indicating that there is no significant structural dependency that is not exploited by the model. This, of course, does *not* mean that the model is identifiable using EMG measurements instead of $P_{\text{mus-Pes}}$, however.

   With the outlier snippets rejected and the model parameters determined, the variance Σ_{RM} of the *residuals* $\varepsilon_{\text{RM},k}$ can then be calculated. Figure 7.7 shows box plots of the residual variances observed in all eleven recordings using both the models (7.4) and (7.5); the median (taken across recordings) standard deviations are 0.7102 mbar and 0.7048 mbar, respectively. Moreover, from the trajectories of the parameter values across multiple 30 s-snippets, an estimate of the variances of the parameter *changes* can also be obtained as follows: assuming

Residual Variance

Parameter

Change
Covariance
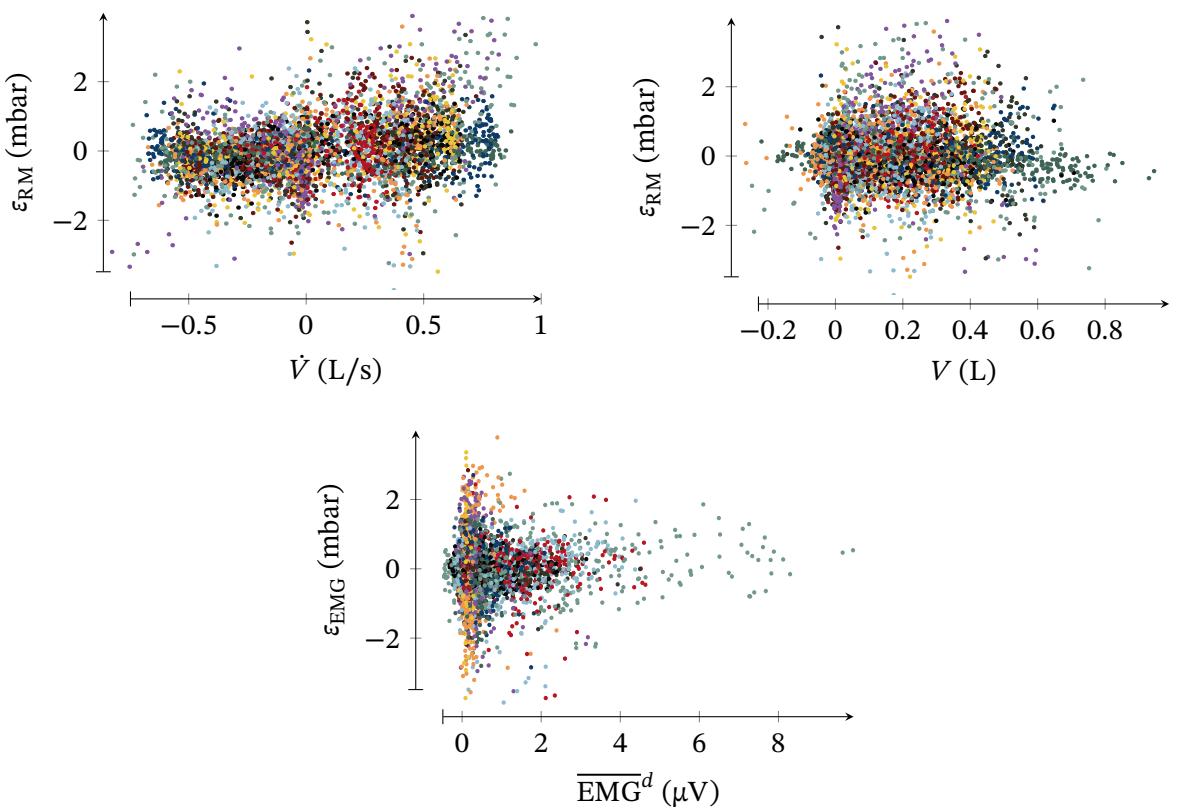
Figure 7.6 – Residual analyses of the model of respiratory mechanics in eq. (7.4) and the EMG–P_{mus} model in eq. (7.8). The two top panels show the residual of the pneumatic model (7.6) as a function of the airflow $\dot{V}$ and lung volume V; the bottom panel shows the residual of the electro-mechanical model (7.8) as a function of $\overline{\text{EMG}}^d$, the main inspiratory EMG channel. Different colors represent the eleven different available recordings.

a random walk model

$$\theta_{\text{RM},k} = \theta_{\text{RM},k-1} + \eta_k, \quad \eta_k \sim \mathcal{N}\left(0, \Sigma_\eta\right)$$

for the model parameters, the total parameter change over a period of 30 s is given by

$$\Delta_{30\,\text{s}}\theta_{\text{RM}} = \theta_{\text{RM},k} - \theta_{\text{RM},k-30f_s} = \sum_{i=k-30f_s+1}^{k} \eta_i \sim \mathcal{N}\left(0, 30 \cdot f_s \cdot \Sigma_\eta\right),$$

where f_s denotes the sampling frequency. Assuming Σ_η to be diagonal, it can thus be estimated as

$$\Sigma_\eta = \frac{1}{30 \cdot f_s} \operatorname{diag}(\operatorname{Var}[\Delta R_1], \operatorname{Var}[\Delta R_2], \operatorname{Var}[\Delta E], \operatorname{Var}[\Delta P_0])$$

where $\operatorname{Var}[\Delta R_1]$ denotes the variance of the observed between-snippet changes in the parameter estimates R_1, and equivalently for the other parameters. Table 7.1 provides the median of the standard deviation of the parameter changes across all recordings.
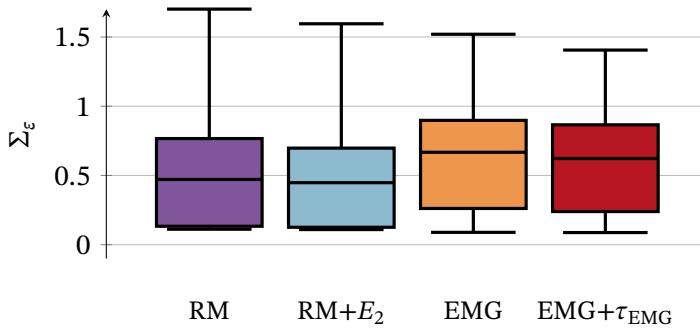
Figure 7.7 – Distribution of the residual variances Σ_{RM} and Σ_{EMG} observed when fitting the respiratory mechanics models (7.4) (RM) and (7.5) (RM+E_2) and the electro-mechanical model (7.8) without (EMG) and with delay estimation (EMG+τ_{EMG}) to the eleven available recordings. The invasively obtained estimate $P_{\text{mus-Pes}}$ is used for the model fits. Boxes show the inter-quartile range and the median value; whiskers indicate the min–max range.

7.3.2 Modeling the EMG–P_{mus} relationship

The EMG-P_{mus} relationship is much less widely explored in the literature than respiratory mechanics. However, several groups of authors have now found very simple, linear models of the shape

$$P_{\text{mus}} \propto \overline{\text{EMG}} \tag{7.7}$$

to describe observed measurements well [12, 37, 62], in line with similar observations in other muscle groups [30]. To fully exploit the information provided by the three available measurement channels, we postulate the combined model

$$P_{\text{mus}} = \alpha^i \overline{\text{EMG}}^i + \alpha^d \overline{\text{EMG}}^d + \alpha^r \overline{\text{EMG}}^r + P_{\text{EMG}} + \varepsilon_{\text{EMG}}, \tag{7.8}$$

where ε_{EMG} denotes the model error. The model (7.8) is consistent with the assumption that

$$P_{\text{mus}} = P_{\text{mus}}^{\text{EMG}i} + P_{\text{mus}}^{\text{EMG}d} + P_{\text{mus}}^{\text{EMG}r}, \tag{7.9}$$

and the linearity assumption (7.7) for the individual muscles. This approach differs from the approach pursued by Bellani *et al.* [12], who assumed

$$P_{\text{mus}} = \alpha(\overline{\text{EMG}}^i + \overline{\text{EMG}}^d + \overline{\text{EMG}}^s), \tag{7.10}$$

i.e., requiring the proportionality factors to be identical for all EMG channels. ($\overline{\text{EMG}}^s$ denotes another respiratory EMG channel not used here.) Due to the wildly varying recording geometry of the different EMG channels, the assumption (7.10) does not appear plausible. It is also different from the approach pursued by Graßhoff *et al.* [37], who select the single most informative channel based on a signal–to–noise ratio calculation and then proceed using only that channel.

Notice that the model (7.8) is robust to a particular type of electromyographic *crosstalk*: Crosstalk assuming that P_{mus} can be correctly represented as

$$P_{\text{mus}} = \begin{pmatrix} \alpha_1 & \alpha_2 & \alpha_3 \end{pmatrix} \begin{pmatrix} s_1 \\ s_2 \\ s_3 \end{pmatrix},$$

Model	R_1	R_2	E_1	E_2	P_0
RM	1.1612	2.1580	0.9914	—	0.3560
RM+E_2	1.0954	2.3169	2.9194	4.6228	0.2739

Model	α^i	α^d	α^r	P_{EMG}
EMG	0.1862	0.3286	1.1831	0.2684
EMG+τ_{EMG}	0.1807	0.2410	1.0078	0.2739

Table 7.1 – Median (across the eleven available recordings) of the observed standard deviation of parameter changes between adjacent 30 s snippets, i.e., $\sqrt{\mathrm{Var}[\Delta_{30\,\mathrm{s}}\theta]}$. The variances are identified by performing regression against $P_{\mathrm{mus\text{-}Pes}}$ in 30 s snippets and analyzing the differences $\Delta_{30\,\mathrm{s}}\theta$ between successive snippets. The analyzed models are the respiratory mechanics model (7.4) (RM), the same model with a quadratic elastance term (RM+E_2), and the electro–mechanical model (7.8), with (EMG+τ_{EMG}) and without (EMG) delay estimation. Units omitted for brevity; refer to section 3.3.3 for the units of all signals.

where the s_i denote muscle activity signals, and assuming further that the three EMG measurements represent mixtures

$$\begin{pmatrix} \overline{\mathrm{EMG}}^i \\ \overline{\mathrm{EMG}}^d \\ \overline{\mathrm{EMG}}^r \end{pmatrix} = M \begin{pmatrix} s_1 \\ s_2 \\ s_3 \end{pmatrix}$$

with an invertible mixing matrix $M \in \mathbb{R}^{3\times3}$, P_{mus} can also be represented as

$$P_{\mathrm{mus}} = \begin{pmatrix} \alpha_1 & \alpha_2 & \alpha_3 \end{pmatrix} M^{-1} M \begin{pmatrix} s_1 \\ s_2 \\ s_3 \end{pmatrix}$$

$$= \begin{pmatrix} \tilde\alpha_1 & \tilde\alpha_2 & \tilde\alpha_3 \end{pmatrix} \begin{pmatrix} \overline{\mathrm{EMG}}^i \\ \overline{\mathrm{EMG}}^d \\ \overline{\mathrm{EMG}}^r \end{pmatrix}.$$

This is of particular interest because often, the $\overline{\mathrm{EMG}}^r$ signal can be used to eliminate abdominal crosstalk from the $\overline{\mathrm{EMG}}^d$ channel.

To assess the ability of the proposed model in describing clinical measurements, the same analyses as described in the previous section were performed: the least-squares-optimal parameters α^i, α^d, α^r and P_{EMG} were determined on 30 s snippets by fitting the observed data to eq. (7.8), using $P_{\mathrm{mus\text{-}Pes}}$ as a surrogate for P_{mus}. Snippet exclusion criteria analogous to those described in the previous section were applied.

To further improve model fit, all EMG envelope signals were delayed by a constant (and
EMG Signal identical across channels) *delay* τ_{EMG}: as is well known (see section 3.1.1 and, in particular,
Delay fig. 3.3), there is a nontrivial dynamic relationship between action potential propagation along a muscle fiber (as picked up by an EMG sensor) and force development (which ultimately results in a change in P_{mus}). For this reason, it is customary in the

(non-respiratory) literature to equip models with an EMG–force delay [40, 56, 64]. Here, we will compare two model variants:

Per-recording EMG signal delay. In this case, the optimal EMG envelope delay τ_{EMG} is determined for each recording individually, by performing OLS regression for all delays in a $[-200\,\mathrm{ms}, 100\,\mathrm{ms}]$ range in $10\,\mathrm{ms}$ increments (resulting in an integer sample shift), and choosing the delay that yields the lowest sum of squared residuals. The median of the identified optimal delays across recordings is $-100\,\mathrm{ms}$. Note the sign: the signals are shifted "to the left".

Fixed EMG signal delay. In this case, all EMG envelope signals in eq. (7.8) are delayed by a fixed delay of $-100\,\mathrm{ms}$ in all recordings.

Notice that a median EMG envelope signal delay of $-100\,\mathrm{ms}$ is not particularly surprising: the causal envelope calculation (cf section 3.3.3) introduced a processing delay of $125\,\mathrm{ms}$ that is mostly compensated for by the identified delay. Moreover, the remaining overall signal shift of $-25\,\mathrm{ms}$ is within the range of experimentally observed EMG-force delays (see fig. 3.3).

Again, fig. 7.5 shows exemplary model fits and fig. 7.6 the model residuals $\varepsilon_{\mathrm{EMG},k}$ as a function of $\overline{\mathrm{EMG}}^d$ in all eleven available recordings. It appears that there is no significant structural dependency that is not exploited by the model. Figure 7.7 shows box plots of the identified equation error variances Σ_{EMG} across all recordings, both with and without delay estimation. The median residual standard deviation is $0.8358\,\mathrm{mbar}$ for both models. Table 7.1 provides the median of the standard deviation of the parameter changes across all recordings. All analyses were performed analogously to those described in the previous section.

7.3.3 Modeling the dynamics of respiratory patient activity

The third and final component of the model, and also the third source of information as described in fig. 7.3, is the model of the temporal evolution of $P_{\mathrm{mus}}(t)$. Notice that the aim here is *not* to develop a realistic simulation model of $P_{\mathrm{mus}}(t)$, but rather to identify a simple forward prediction model that conveys some amount of information about generally plausible $P_{\mathrm{mus}}(t)$ waveforms, and that will be used in the final smoothing step of the estimation procedure.

To obtain this model, we fit a range of model variants to the $P_{\mathrm{mus\text{-}Pes}}$ signals of all datasets:

Autoregressive (AR) models of orders $p = 1, 2, 3, 5,$ and 10. We will denote these by AR_p.

Autoregressive models with exogenous inputs (ARX) and an input–output-delay of 1. For these models, the flow and volume signals are used as exogenous inputs, i.e., the model is

$$P_{\mathrm{mus},k} = \sum_{i=1}^{p} a_i \cdot P_{\mathrm{mus},k-i+1} + \sum_{i=1}^{q} b_{\dot{V},i} \cdot \dot{V}_{k-i} + \sum_{i=1}^{q} b_{V,i} \cdot V_{k-i} + \eta_k,$$

where η_k denotes white noise. We will fix $p = 2$ and consider $q = 1, 2, 3$, and we will denote these models by ARX_{pq}.

State-space (SSM) models of the form

$$\theta_k = A \cdot \theta_{k-1} + b_{\dot{V}} \cdot \dot{V}_{k-1} + b_V \cdot V_{k-1} + \eta_{k-1}, \quad \eta_{k-1} \sim \mathcal{N}\big(0, \Sigma_\eta\big)$$
$$P_{\mathrm{mus},k} = C \cdot \theta_k + \nu_k, \quad \nu_k \sim \mathcal{N}(0, \Sigma_\nu),$$

	AR$_1$	AR$_2$	ARX$_{21}$	ARX$_{22}$	ARX$_{23}$	AR$_3$	AR$_5$	AR$_{10}$
ΔnAIC	3.4939	**0.4391**	0.4372	0.3873	0.3711	0.4237	0.2262	0.1306
n	1	**2**	4	6	8	3	5	10

	SSM$_2^a$	SSM$_2$	SSM$_3^a$	SSM$_3$	SSM$_4^a$	SSM$_4$	SSM$_5^a$	SSM$_5$
ΔnAIC	0.3456	0.2933	0.1738	0.0644	0.1004	0.0035	0.1004	0
n	8	12	15	21	24	32	35	45

Table 7.2 – Difference ΔnAIC of normalized AIC (nAIC) values and the minimum nAIC value observed across all models (achieved by and SSM$_5$), as well as the number of model parameters (n) of multiple dynamic models of P_mus across all available signal snippets. The highlighted AR$_2$ model is the one selected for further use.

where $A, b_{\dot{V}}, b_V, C, \Sigma_\eta$, and Σ_ν are unknown matrices to be identified by the algorithm, as well as autonomous systems of the form

$$\theta_k = A \cdot \theta_{k-1} + \eta_{k-1}, \quad \eta_{k-1} \sim \mathcal{N}(0, \Sigma_\eta)$$
$$P_{\text{mus},k} = C \cdot \theta_k + \nu_k, \quad \nu_k \sim \mathcal{N}(0, \Sigma_\nu).$$

Models of order, i.e., state-space dimension, $p = 2, 3, 4$, and 5 are considered, and we will denote these by SSM$_p$ for the version with inputs, and SSM$_p^a$ for the autonomous version, respectively.

The rationale behind the inclusion of the flow and volume signals as inputs in some of the models is that — for obvious reasons — both signals are correlated in some way with $P_\text{mus}(t)$ and might thus contain information that can be used to better predict $P_\text{mus}(t)$. The AR and ARX models are fit using the standard (unregularized) least-squares approach implemented in the Matlab `arx` function. The SSM models are fit using the subspace identification method implemented in the Matlab `n4sid` function. In both cases, the cost function is given by the one-step-ahead prediction error. For details on both estimation methods, refer to Ljung [55].

Normalized Akaike Information Criterion

For selecting a suitable model out of this group, we compare the values of the *normalized Akaike information criterion (nAIC)* [3, 4, 55] calculated across all available data from all patients. Since the *absolute* value of the (n)AIC conveys no information on model fit — only the *relative* values of different models are of interest for their comparison [17]) — we subtract the minimum value observed across all models (the optimum) from all nAIC values, yielding the performance measure

$$\Delta\text{nAIC}_i = \text{nAIC}_i - \min_j \text{nAIC}_j$$

for model i. Table 7.2 provides the results. The best performance is achieved by the SSM$_5$ model, very closely followed by the SSM$_4$ model. However, except for the AR$_1$ model, all other models attain nearly identical nAIC values. Following a well-known rule of thumb, all models for which ΔnAIC < 2 "have substantial support" [17]; this is clearly the case for all models except AR$_1$. Thus, because it is the simplest of the models with similarly low nAIC
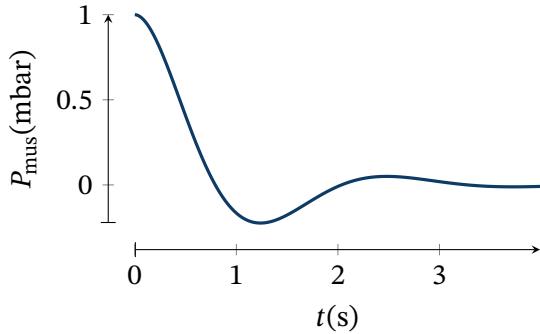
Figure 7.8 – Forward simulation of the model of $P_{\text{mus}}(t)$ described in eq. (7.11) for the initial conditions $P_{\text{mus},k-1} = P_{\text{mus},k-2} = 1$.

values, we will proceed using the AR_2 model in the following. The identified parameters for that model are

$$P_{\text{mus},k} = 1.975 \cdot P_{\text{mus},k-1} - 0.9762 \cdot P_{\text{mus},k-2} + \eta_k, \quad \eta_k \sim \mathcal{N}\left(0, \Sigma_{\eta,=}2.1314 \cdot 10^{-4}\right). \quad (7.11)$$

Figure 7.8 shows a forward simulation of the model for the initial conditions

$$P_{\text{mus},k-1} = P_{\text{mus},k-2} = 1,$$

illustrating that the model represents a damped linear oscillator. The model's frequency response has its maximum absolute value at an oscillation period of 2.83 s.

7.4 Model-based inference for P_{mus} in mechanical ventilation

In this section, the proposed algorithm for estimating P_{mus} from respiratory surface EMG measurements will finally be described in its entirety. To this end, recall from fig. 7.4 that the algorithm proceeds in four main stages: preprocessing and artifact removal, importance weight calculation, time-varying model parameter estimation, and, finally, dynamic P_{mus} signal estimation. The algorithm is entirely independent of any invasively obtained measurements such as P_{es}. All stages are implemented in an offline fashion here, i.e., all stages of the algorithm are applied to the full recording before proceeding to the next algorithmic stage. The different stages of the algorithm will now be detailed in the following sections.

7.4.1 Preprocessing and artifact removal

As they are not the core topic of this chapter, these steps are only described very briefly here. A volume signal is obtained by continuous integration of the $\dot{V}$ measurement signal, with a simple leakage correction algorithm and baseline removal to prevent the volume signal from drifting. Measurement artifacts in all signals are detected based on simple thresholds (too large / too small / too large or small for too long) and rejected.

7.4.2 Calculation of importance weights

Covariate Shift

As was mentioned in section 7.2.2, the estimation problem suffers from *covariate shift* in the following sense: the distribution $p_{\text{train}}\left(\dot{V}, V, \overline{\text{EMG}}^i, \overline{\text{EMG}}^d, \overline{\text{EMG}}^r\right)$ is often highly imbalanced, with a high prevalence of very similar, "normal" breaths. Thus, unweighted (or equally weighted) estimation will identify model parameters that describe these "normal" breaths well, possibly at the cost of reduced performance on more rare, unusual breaths. However, the aim of the parameter identification procedure is not prediction: instead, the parameters are to be used for inferring the hidden quantity P_{mus}. Thus, the aim of inference is to obtain a set of parameters that describes the respiratory system reasonably well under all relevant circumstances. Towards this aim, we propose the assumption that $p_{\text{target}}(x) \approx \text{Unif}(x)$, where the uniform distribution is defined over a suitably restricted space. In particular, we propose four different *importance weighting schemes* that will be evaluated comparatively in section 7.5:

Importance
Weighting
Schemes

Multivariate density estimation, all data (w^{mvd}). To calculate the importance weights, the same procedure as used in the previous chapter is employed: the training data distribution $p_{\text{train}}\left(\dot{V}, V, \overline{\text{EMG}}^d\right)$ is approximated using the multivariate density estimation scheme described in section 6.5.1, and preliminary importance weights are then calculated as

$$\tilde{w}_k^{\text{mvd}} = \frac{p_{\text{target}}\left(\dot{V}_k, V_k, \overline{\text{EMG}}_k^d\right)}{\hat{p}_{\text{train}}\left(\dot{V}_k, V_k, \overline{\text{EMG}}_k^d\right)} \propto \frac{1}{\hat{p}_{\text{train}}\left(\dot{V}_k, V_k, \overline{\text{EMG}}_k^d\right)}.$$

Notice that the two additional EMG signals $\overline{\text{EMG}}^i$ and $\overline{\text{EMG}}^r$ are omitted from the density estimation and weight calculation in order to reduce the dimensionality of the estimation problem and because the diaphragm is the main respiratory muscle. As is usual in the importance-weighting literature (see the discussion in section 6.1.2), we restrict the maximum weight range in order to limit the effect of density estimation uncertainties in regions with very few data points and the variance of the estimator (see fig. 6.4). In particular, we restrict the weight range $\max\left(w^{\text{mvd}}\right) / \min\left(w^{\text{mvd}}\right)$ to 100 by capping any weights larger than $100 \cdot \min\left(\tilde{w}^{\text{mvd}}\right)$. To further smooth out brief outliers, the weighting signal is smoothed using a moving median with a window length of 100 ms and, subsequently, a moving mean with a window length of 500 ms.

Multivariate density estimation, excluding steep slopes ($w^{\text{mvd}-}$). These weights are calculated in the exact same way as the previously described weights w^{mvd}, with one difference: samples with large flow or pressure gradients are excluded from the analyses, and their weights are set to 0 explicitly. The rationale is this: it is known (see, e.g., fig. 3.6 in Bates [9], and also fig. 7.5) that the classical model is inaccurate during phases of rapid pressure and flow changes due to various inertial effects. For the reconstruction of an accurate P_{mus} model, fitting these transient regions is not critical, and doing so might lead to worse model performance in describing P_{mus}. Technically, we reject any samples for which either $d\dot{V}/dt$ or dP_{aw}/dt are greater than their respective 30 s moving 0.95–quantile or smaller than their respective 30 s moving 0.05–quantile.

Simple binning (w^{bin}). With the intent of reducing the dependence on accurate density estimation and at the same time reducing the vulnerability of the weighting scheme to outliers (which would receive a high weight by the two estimation schemes described above), an alternative, simple, and robust weighting scheme has been developed. The aim is to give roughly equal total weight (=number of samples times their weight) to samples during which the resistive, elastic, and electromyographic components of the model are well observable *in isolation*. To this end, first binary significance criteria are defined for $\dot{V}$, V, and $\overline{\text{EMG}}^d$ as $|\dot{V}| > 0.15\,\text{L/s}$, $V > \text{movmedian}_{30\,\text{s}}(V)$ and

$$\overline{\text{EMG}}^d > \frac{2}{3}\,\underset{20\,\text{s}}{\text{movmedian}}\Big(\underset{4\,\text{s}}{\text{movmax}}\Big(\underset{0.2\,\text{s}}{\text{movmedian}}\big(\overline{\text{EMG}}^d\big)\Big)\Big).$$

The weight for samples during which none of the three signals is significant is set to 0, and the weight of all samples during which $\dot{V}$ is significant is set to 1. Next, the weight of all samples during which *only V* is significant is increased such that the total weight on these samples (which is given by the weight times the number of samples) is equal to the number of samples during which $\dot{V}$ is significant (which have received weight 1, see above), with a maximum weight of 100. This step is repeated for the samples during which only $\overline{\text{EMG}}^d$ is significant. Finally, to smooth out brief outliers, the resulting weight signal is smoothed using a moving median with a window length of 0.2 s.

Simple binning, excluding steep slopes ($w^{\text{bin}-}$). Again, these weights are calculated in the exact same way as the previously described weights w^{bin}, with one difference: samples with large flow or pressure gradients are excluded from the analyses, and their weights are set to 0 explicitly. Cf. the description of $w^{\text{mvd}-}$ above for details.

To analyze the influence of the weighting schemes, the results will also be compared to a version using standard uniform weighting ($w^{\mathbb{1}} \equiv 1$).

7.4.3 Estimation of time-varying model parameters

Given the preprocessed signals and the samples weights w_k, we can now proceed to estimate the time-varying model parameters. For this purpose, the weighted Kalman filter and smoother described in the previous chapter are employed. For the model parameters (the dynamical states of the system), we employ the simple *random–walk model*

$$\theta_k = \theta_{k-1} + \eta_k, \quad \eta_k \sim \mathcal{N}\big(0, \Sigma_\eta\big). \tag{7.12}$$

Random–walk Model

For the measurement model, the pneumatic model

$$P_{\text{aw},k} + P_{\text{mus},k} = R_1 \dot{V}_k + R_2 |\dot{V}_k| \dot{V}_k + E V_k + P_0$$

and the EMG–P_{mus} model

$$P_{\text{mus}} = \alpha^i \overline{\text{EMG}}^i + \alpha^d \overline{\text{EMG}}^d + \alpha^r \overline{\text{EMG}}^r + P_{\text{EMG}},$$

are combined to obtain

$$P_{\text{aw},k} = R_{1,k} \dot{V}_k + R_{2,k} |\dot{V}_k| \dot{V}_k + E_k V_k - \alpha_k^i \overline{\text{EMG}}_k^i - \alpha_k^d \overline{\text{EMG}}_k^d - \alpha_k^r \overline{\text{EMG}}_k^r + \underbrace{P_{0,k} - P_{\text{EMG},k}}_{:=\tilde{P}_{0,k}} + \varepsilon_k$$

$$\tag{7.13}$$

with $\varepsilon_k \sim \mathcal{N}(0, \Sigma_\nu)$, which is linear in the model parameter vector

$$\theta_k = \begin{pmatrix} R_{1,k} & R_{2,k} & E_k & \alpha_k^i & \alpha_k^d & \alpha_k^r & \tilde{P}_{0,k} \end{pmatrix}^\mathsf{T}.$$

A measurement noise standard deviation of $\sqrt{\Sigma_\nu} = 1\,\text{mbar}$ is assumed. An equivalent formulation (with the additional state $E_{2,k}$) can be derived analogously for the pneumatic model (7.5) with quadratic elastance. Using the state equation (7.12), the measurement equation (7.13) and the sample weights w_k, the weighted Rauch–Tung–Striebel-Smoother described in algorithm 8 is used to estimate the time-varying model parameters θ_k. Sequential measurement updates, the Joseph stabilized covariance update, missing data handling, and state constraints are all implemented as described in section 2.3.1.[5]

Process Noise Covariance

For the *process noise covariance* Σ_η, three different choices will be compared. Firstly, the process noise can be chosen exactly as determined empirically in section 7.3.1 and section 7.3.2; this will be termed Σ_η^+ because it is the highest level of process noise considered. Secondly, to achieve more stable estimator behavior, a version will be considered in which the process noise is set to $\Sigma_\eta^- = 10^{-4} \cdot \Sigma_\eta^+$. And thirdly, a version without any parameter variability will also be considered, i.e., $\Sigma_\eta^{=0} = 0$. (In this case, the Kalman filter and smoother just reduce to static regression, as discussed before.)

EMG Signal Delay

Moreover, just like in section 7.3.2, two variants will be compared in the following: estimation with a fixed, static *EMG envelope signal delay* of $-100\,\text{ms}$, and with a per-recording identification of the optimal delay τ_{EMG}. Here, however, $P_{\text{mus-Pes}}$ cannot be used, as the algorithm is meant to be fully independent of any invasive measurements. For this reason, we employ the weighted hyperparameter tuning scheme proposed in section 6.3, i.e., we use importance-weighted (empirical) prediction error minimization (IWPEM) and choose the delay as

$$\tau_{\text{EMG}}^* = \underset{\tau_{\text{EMG}} \in \mathbb{Z}}{\arg\min} \sum_{k=2}^{N^S} w_k \big(P_{\text{aw},k} - R_{1,k-1} \dot{V}_k - R_{2,k-1} |\dot{V}_k| \dot{V}_k - E_{k-1} V_k$$
$$+ \alpha_{k-1}^i \overline{\text{EMG}}_{k+\tau_{\text{EMG}}}^i + \alpha_{k-1}^d \overline{\text{EMG}}_{k+\tau_{\text{EMG}}}^d + \alpha_{k-1}^r \overline{\text{EMG}}_{k+\tau_{\text{EMG}}}^r - \tilde{P}_{0,k-1} \big)^2.$$

The NOMAD solver [1, 52], a mesh-adaptive direct search (MADS) algorithm, is used to solve this integer optimization problem efficiently. (Note that every cost function evaluation requires a full Kalman filter and smoother run.)

7.4.4 Estimation of P_{mus}

Given the time-varying estimates θ_k of the pneumatic and electro-mechanical model parameters obtained in the previous section, P_{mus} can now be estimated. To this end, pneumatic and electro-mechanical model components are now, finally, combined with the dynamic model of P_{mus} identified in section 7.3.3. Transforming the autoregressive model from eq. (7.11) into a state-space representation, we obtain

$$\begin{pmatrix} P_{\text{mus},k} \\ P_{\text{mus},k-1} \end{pmatrix} = \begin{pmatrix} 1.975 & -0.9762 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} P_{\text{mus},k-1} \\ P_{\text{mus},k-2} \end{pmatrix} + \begin{pmatrix} \eta_k \\ 0 \end{pmatrix}, \quad \eta_k \sim \mathcal{N}\big(0, \Sigma_\eta = 2.1314 \cdot 10^{-4}\big). \quad (7.14)$$

---

[5]The following, very permissive constraints are implemented: $-100 \le R_1, R_2 \le 100$; $1 \le E \le 100$; $-50 \le \alpha^i \le 50$; $0 \le \alpha^d \le 50$; $-50 \le \alpha^r \le 0$; $-10 \le \tilde{P}_0 \le 30$. (Units omitted here for brevity, see section 3.3.3 for the units of all signals.)

Next, we can compute estimates of P_{mus} from both the respiratory mechanics model (7.4) and the electro-mechanical model (7.8): from the first, we obtain

$$P_{\text{mus},k}^{\text{RM}} = R_{1,k}\dot{V}_k + R_{2,k}|\dot{V}_k|\dot{V}_k + E_k V_k + \tilde{P}_{0,k} - P_{\text{aw},k}, \qquad (7.15)$$

and from the second,

$$P_{\text{mus},k}^{\text{EMG}} = \alpha^i \overline{\text{EMG}}^i + \alpha^d \overline{\text{EMG}}^d + \alpha^r \overline{\text{EMG}}^r. \qquad (7.16)$$

Here, without loss of generality, the constant offset $\tilde{P}_0 = P_0 + P_{\text{EMG}}$ was fully attributed to the pneumatic component. (As a result, the final P_{mus} estimate will have an unknown offset.) Equations (7.15) and (7.16) can now be combined to obtain the measurement equation

$$\begin{pmatrix} P_{\text{mus},k}^{\text{RM}} \\ P_{\text{mus},k}^{\text{EMG}} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} P_{\text{mus},k} \\ P_{\text{mus},k-1} \end{pmatrix} + \nu_k, \quad \nu_k \sim \mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \Sigma_{\text{RM}} & 0 \\ 0 & \Sigma_{\text{EMG}} \end{pmatrix} \right), \qquad (7.17)$$

with Σ_{RM} and Σ_{EMG} being the noise covariances identified empirically in section 7.3.1 and section 7.3.2. Finally, a standard (unweighted) Kalman smoother is applied to eqs. (7.14) and (7.17), yielding the final estimate $\hat{P}_{\text{mus},k}$ (from the first of the two states).

7.5 Results

To evaluate the accuracy and robustness of the proposed algorithm for estimating P_{mus}, the resulting estimate $\hat{P}_{\text{mus},k}$ is compared to the gold-standard reference $P_{\text{mus-Pes}}$, calculated using the methods described by Graßhoff *et al.* [35, 36] (also see section 3.3.3). In order to prevent different signal baselines from influencing the results, a time-varying baseline is subtracted from both signals before comparison. To this end, the moving 0.1-quantile of the respective signal is calculated within a moving 30 s window and subtracted. We compare nine different variants of the proposed estimation algorithm, listed in table 7.3. In order to provide a comprehensive performance assessment, three different *performance measures* are considered:

Performance Measures

1. Spearman's rank correlation (ρ), quantifying the degree of correlation between $P_{\text{mus-Pes}}$ and $\hat{P}_{\text{mus}}$,

2. the mean absolute deviation (MAD) between $P_{\text{mus-Pes}}$ and $\hat{P}_{\text{mus}}$, and

3. the mean absolute *breath amplitude* deviation (MADamp). To calculate this, the moving 0.95-quantile and the moving 0.05-quantile of both $P_{\text{mus-Pes}}$ and $\hat{P}_{\text{mus}}$ are calculated over a moving window of 5 s length. The time-varying amplitudes of both signals are then estimated as the difference between these two quantities, and MADamp is defined as the mean absolute difference between the two amplitude signals.

Figure 7.10 shows box plots of these three quantities for all nine algorithm variants. Concerning worst-case performance, algorithm variant 1 performs best across all three measures. Its median performance is also best or very close to best on all measures: it achieves a median MAD of 0.79 mbar, a median MADamp of 1.78 mbar, and a median ρ of

ID	Weighting scheme	Process noise	EMG delay estimation	Quadratic elastance term E_2
1	$w^{\text{bin}-}$	Σ_η^-	yes	no
2	w^{bin}	Σ_η^-	yes	no
3	w^{mvd}	Σ_η^-	yes	no
4	$w^{\text{mvd}-}$	Σ_η^-	yes	no
5	$w^{\mathbb{1}}$	Σ_η^-	yes	no
6	$w^{\text{bin}-}$	Σ_η^+	yes	no
7	$w^{\text{bin}-}$	$\Sigma_\eta^{=0}$	yes	no
8	$w^{\text{bin}-}$	Σ_η^-	yes	yes
9	$w^{\text{bin}-}$	Σ_η^-	no	no

Table 7.3 – Algorithm variants considered for the evaluation. Variant 1 is the baseline; all other variants vary a single feature with respect to variant 1.

0.74. It can be observed that both the choice of the weighting scheme and the choice of the process noise have a large impact on the estimation performance. Interestingly, choosing the process noise as identified empirically leads to significantly worse performance, whereas performing fully static regression ($\Sigma_\eta^{=0} = 0$) yields almost as good results as using a low level of process (Σ_η^-) on these recordings. Finally, including the delay estimation step does significantly improve worst-case performance, whereas including the quadratic elastance term actually leads to slightly worse performance. Figure 7.11 shows exemplary snippets from eight recordings, illustrating the quality of the obtained estimate $\hat{P}_{\text{mus}}$, and figs. 7.9 and 7.13 shows exemplary snippets from two recordings including the input signals, the estimate $\hat{P}_{\text{mus}}$ obtained with algorithm variant 1, and the trajectories of two of the weighting schemes, $w^{\text{bin}-}$ and w^{mvd}. Moreover, a Bland–Altman analysis has been performed for algorithm variant 1, see fig. 7.12. A variance analysis for repeated measurements [16, 94] yields 95%-limits of agreement of -2.49 mbar and 2.65 mbar (also shown in fig. 7.12). Finally, to assess the potential danger of regressor colinearity, *variance inflation factors (VIFs)* were calculated on each recording between the three main regressors $\dot{V}$, V, and $\overline{\text{EMG}}^d$. The maximum VIF was below 4 in all recordings, indicating that colinearity is not a strong concern [44].

Variance Inflation Factor

7.6 Discussion & outlook

In this chapter, a new, entirely noninvasive method for monitoring the pressure P_{mus} generated by the respiratory muscles of mechanically ventilated patients has been proposed and evaluated. The method performs model-based sensor fusion, exploits both pneumatic as well as surface electromyographic measurements, and uses all available (non-faulty) measurement samples. An importance-weighting procedure is used to ensure that the identified model describes the patient well in all observed regions of the data space, and the model parameters are identified in a time-varying fashion. The resulting estimate $\hat{P}_{\text{mus}}$ shows a median mean absolute deviation of just 0.79 mbar with respect to the gold standard
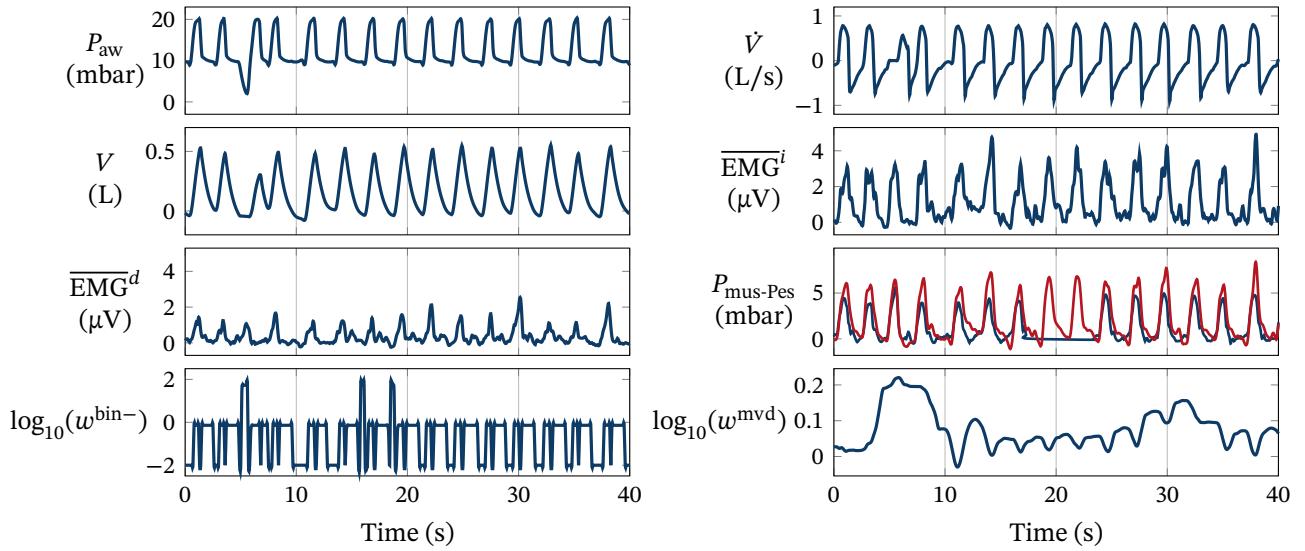
Figure 7.9 – An extract from the recording of patient B/1. The estimate $\hat{P}_{\text{mus}}$ (red) has been obtained using algorithm variant 1. The reference measurement $P_{\text{mus-Pes}}$ showed a recording artifact around 20 s. (The $\overline{\text{EMG}}^r$ channel is not shown here; it has a very small signal amplitude and is uninformative on this snippet.)

reference $P_{\text{mus-Pes}}$ in eleven recordings from nine intensive care patients, a median mean breath amplitude absolute deviation of 1.78 mbar, and a median coefficient of rank correlation of 0.74, all indicating a high estimation accuracy. (In this regard, it is crucial to note that the reference estimate $P_{\text{mus-Pes}}$ is *also* not free of errors, i.e., it is sometimes not clear whether a difference between $P_{\text{mus-Pes}}$ and $\hat{P}_{\text{mus}}$ indicates an error in $P_{\text{mus-Pes}}$, or in $\hat{P}_{\text{mus}}$.) Both importance weighting and (importance-weighted) delay estimation were found to significantly improve algorithm performance, whereas a quadratic elastance term did not improve estimation results.

Compared to purely pneumatic methods for estimating P_{mus}, such as the methods proposed by Kondili *et al.* [48] and Younes *et al.* [91], the proposed method has the benefit of an additional source of information that can be exploited: the EMG measurements. Compared to the EMG-based approach of Bellani *et al.* [12], who used only the EMG and P_{aw} during occlusions for estimating α and then calculated P_{mus} as $\alpha \cdot \overline{\text{EMG}}^d$, both more signals and many more samples are exploited. Moreover, unlike any previous approach except for the method of Kondili *et al.* [48], the proposed method does not rely on *any* special ventilation maneuvers being performed and can (in principle) be applied to measurements obtained during any ventilation mode. The estimation of P_{mus} by incorporating the three sources of available information — EMG signals, pneumatic signals, and knowledge about the dynamics of P_{mus} — means that estimation quality can remain high even if one or multiple of the measurement signals are of low signal quality or even missing. Another advantage in comparison to earlier EMG-based approaches is that due to the exploitation of multiple EMG channels and the proposed EMG–P_{mus} model, muscle crosstalk can be treated correctly in many cases, see the corresponding discussion in section 7.3.2. Even if crosstalk cannot be removed correctly and one channel becomes
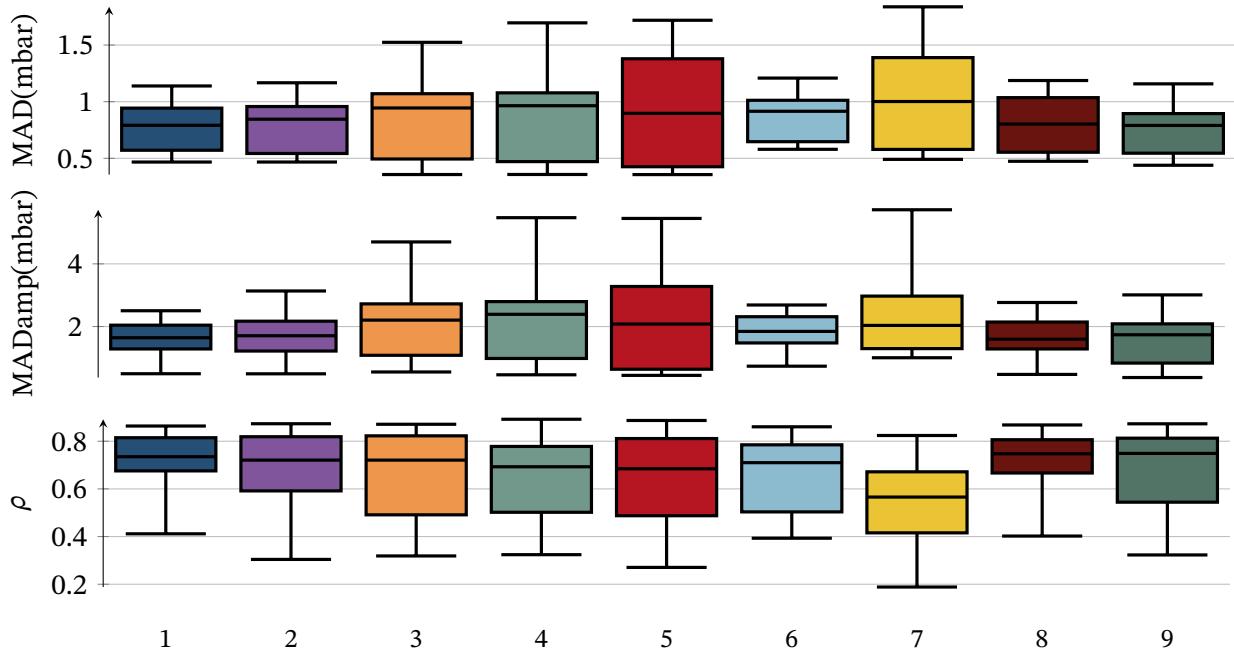
Figure 7.10 – Distribution of three estimation performance measures comparing the gold standard $P_{\text{mus-Pes}}$ and various (baseline-adjusted, see section 7.5) surface EMG-based estimates $\hat{P}_{\text{mus}}$ across all recordings. The performance measures are the mean absolute deviation (MAD), the mean absolute deviation of the *amplitudes* of the two signals (MADamp), and Spearman's rank correlation coefficient ρ. Boxes show the inter-quartile range and the median value, whiskers indicate the min–max range, and number labels indicate the algorithm variant as per table 7.3.

unusable, the estimation can still succeed by just using the two other remaining channels. More generally, the method is, in many cases, able to extract useful P_{mus} estimates despite low EMG signal quality, see fig. 7.13 for an example.

Dynamical Model of P_{mus}

One particularity of the proposed method concerns the utilization of a *dynamical model* of the time course of P_{mus}. A few researchers have previously proposed models for the time course of P_{mus} [68, 71, 72, 84]. None of them, however, have identified their model from actual measurements, as has been done here. The selected autoregressive model of second order may appear simplistic, and more complex models of the respiratory controller can, of course, be conceived. It is crucial, however, to keep in mind the purpose of the model: to predict the likely evolution of P_{mus} a few samples into the future, in order to mediate between the two existing estimates $P_{\text{mus}}^{\text{RM}}$ and $P_{\text{mus}}^{\text{EMG}}$ in a useful way. For this purpose, the identified model is certainly sufficient, if not optimal.

Time-varying Estimation

Performing *time-varying estimation*, as opposed to static regression, did not yield large benefits for the recordings considered here, although slightly better estimates were obtained. However, this, while interesting to note, does not mean that time-varying estimation is not worthwhile. The recordings considered here were rather short (between five and thirty minutes) compared to the time periods over which a respiratory monitoring system would be employed in an intensive care unit. During such long time periods, it appears much more likely that significant parameter changes will occur, if only because the electrode–skin
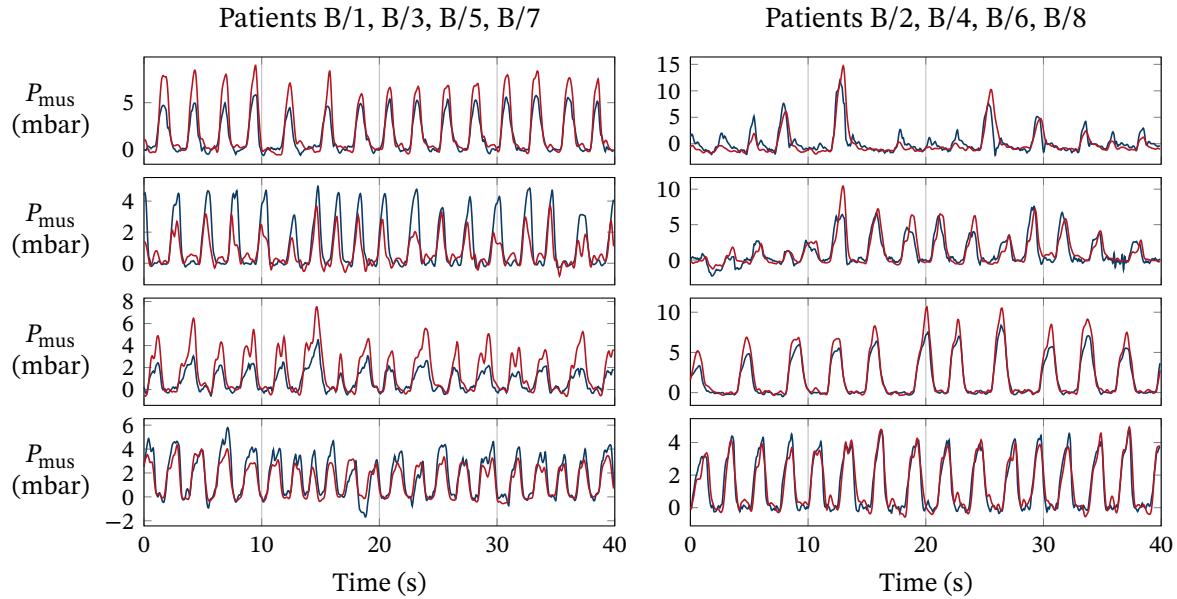
Figure 7.11 – Exemplary results of variant 1 of the proposed, surface EMG-based estimate $\hat{P}_{\text{mus}}$ (red), compared to the gold standard $P_{\text{mus-Pes}}$ (blue).

contact changes or the patient changes position in bed. Of course, it is an equally valid option to repeatedly perform regression over fixed time windows, but for various reasons, it may be considered preferable to use a dedicated time-varying estimation scheme.

There are only a few fundamental limitations to the applicability of the proposed method. Of course, any condition that fully prevents the measurement of respiratory surface EMG measurements means that the proposed method is not applicable. A high degree of *colinearity* (or linear dependence) between the regressors $\dot{V}, V, \overline{\text{EMG}}^i, \overline{\text{EMG}}^d$, and $\overline{\text{EMG}}^r$ might theoretically pose a problem (see section 2.2.1), but this has not been observed here (the maximum variance inflation factors are < 4 in all recordings, see section 7.5) and appears rather unlikely due to the physically constrained nature of the system generating the measurements. Should colinearity become a real concern, it could be alleviated by performing any type of ventilation maneuver that forces a disruption to the usual breathing pattern. (Colinearity between the different EMG channels is, of course, possible and not unlikely, but it does not pose a problem: the particular values of the coefficients α are not of relevance.)

Colinearity

The selection (and weighting) of samples to perform the regression on has a tremendous impact on the quality of the estimate P_{mus}, with the performing of special identification maneuvers only being the most extreme case of *sample selection and weighting*. This is demonstrated by the variability of the estimation results as a function of the employed weighting scheme. In the spirit of the previous chapter 6, to obtain good estimation results, it is crucial to specify the target distribution $p_{\text{target}}\left(\dot{V}, V, \overline{\text{EMG}}^d\right)$ over which the model should perform well, and to weight samples accordingly. To this end, a simple weighting scheme has been proposed that discards regions during which all signals are negligible and regions with large flow or pressure gradients, and that roughly equalizes the total weight on

Sample Selection and Weighting
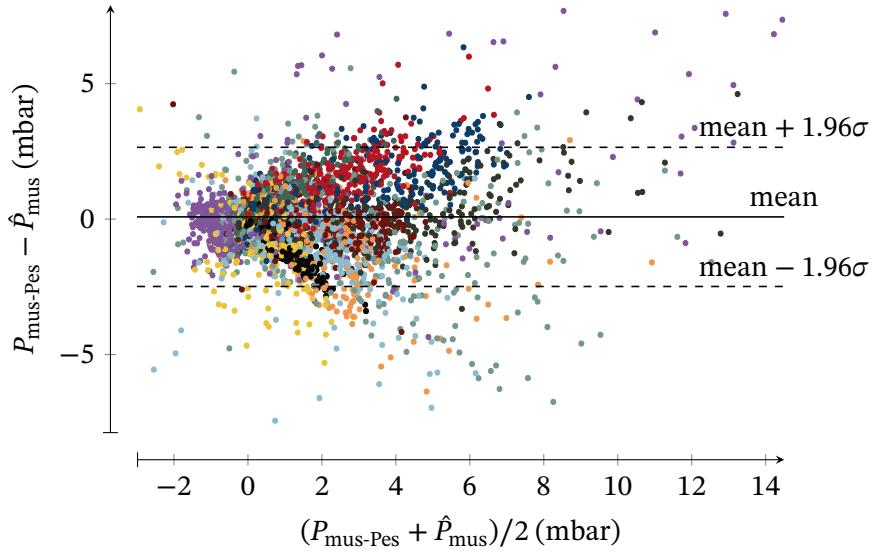
Figure 7.12 – A Bland Altman analysis comparing variant 1 of the proposed surface EMG-based P_{mus} estimator to the gold standard $P_{\text{mus-Pes}}$. The limits of agreement are calculated using the method of Bland and Altman [16]. Different colors indicate samples from different recordings.

samples during which the electro-mechanical scaling factors α, the resistance terms R_1 and R_2, and the elastance E can be observed. Using this simple weighting scheme markedly improves the quality of the estimation results. On the other hand, the influence of the weighting scheme represents clear evidence that the postulated model, consisting of a classical model of respiratory mechanics and a simple electro-mechanical coupling model, suffers from significant model errors. As was discussed at some length in section 2.1.4, this model error can be understood as either a deficiency of the *system* model or as a deficiency of the *noise* model.

System Model Concerning the *system* model, it is well known (and was exploited by the previously described weighting scheme) that the classical model is inaccurate during phases of rapid pressure and flow changes due to inertial effects, among others, and during phases of extreme lung inflation or deflation. This deficiency can also be observed in fig. 7.6, where large residuals are observed close to $\dot{V} = 0$, which occur during either sharply rising flows at the beginning of inspiration or sharply falling flows at the beginning of expiration. The employed physiological model could, of course, be further extended to account for effects like tissue viscoelasticity [9] and intrinsic PEEP [50]. The inclusion of viscoelastic effects will generally lead to models that are nonlinear in the parameters, i.e., a linear Kalman filter and smoother could no longer be used. Instead, one of the nonlinear filter and smoother variants discussed in section 2.3.3 could be employed.

Noise Model Concerning the *noise* model, there is simply a variety of external disturbances and model errors that are unobservable (with current methods). Thus, the Gaussian noise model that has been employed here is very likely not an ideal choice. Instead, a "fat tail" noise model could be employed, resulting in a robust Kalman filter implementation that is more resistant to outliers [80, 86]. In addition, an errors-in-variables formulation could be pursued, in order
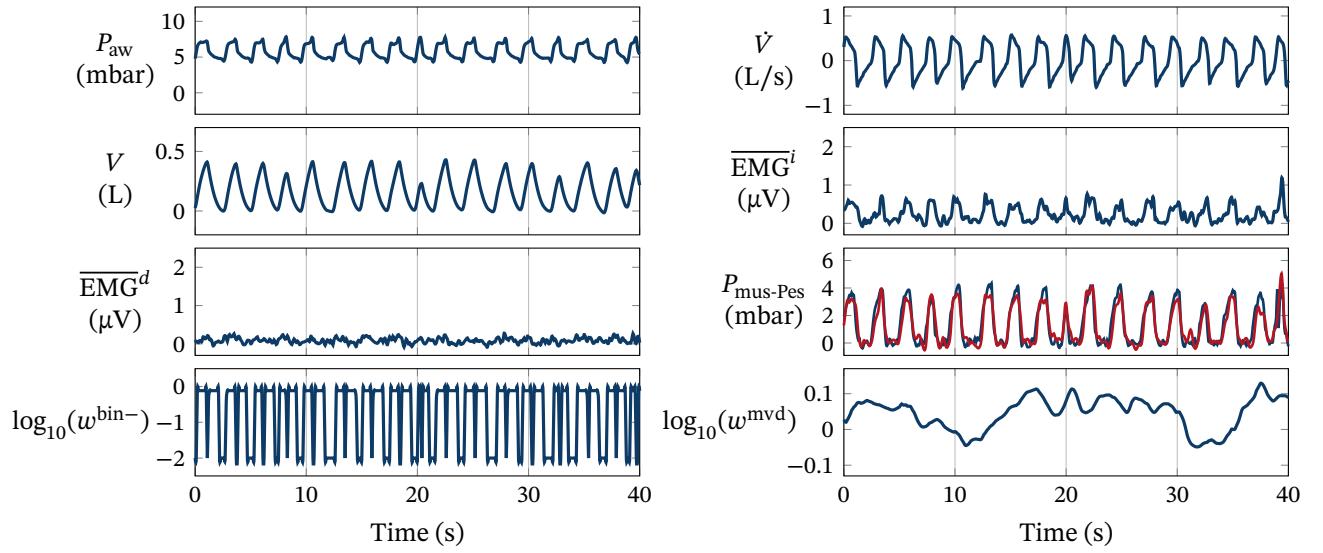
Figure 7.13 – An extract from the second recording of patient B/8, illustrating that the proposed method can yield a high-quality estimate $\hat{P}_{\text{mus}}$ (red) even in the face of low EMG signal quality. The estimate $\hat{P}_{\text{mus}}$ has been obtained using algorithm variant 1. (The $\overline{\text{EMG}}^{r}$ channel is not shown here; it has a very small signal amplitude and is uninformative on this snippet.)

to account for the non-negligible measurement errors in all regressor signals, but particularly in the EMG signals [82]. If a time-varying estimation scheme shall still be used, this would, again, result in a nonlinear filtering and smoothing problem. Moreover, while a constant measurement noise covariance was assumed for the pneumatic P_{mus} estimate $P_{\text{mus}}^{\text{RM}}$ and the electromyographic P_{mus} estimate $P_{\text{mus}}^{\text{EMG}}$ in eq. (7.17), these covariances could also be chosen in a time-varying fashion. This would allow for, e.g., accounting for time-varying signal quality or model credibility in any of the model components. Finally, to further improve the robustness to muscle crosstalk, a dedicated crosstalk detection and removal stage could be introduced into the EMG preprocessing stage, utilizing, for example, the developments presented in chapter 4.

Closely related to the noise model, proper *uncertainty quantification* for the resulting estimate $\hat{P}_{\text{mus}}$ appears highly desirable: (Temporarily) inaccurate estimation results would pose significantly less of a problem if their low accuracy were at least *known* reliably. While the described algorithm is probabilistic in nature and does provide uncertainty estimates in the form of the Kalman filter and smoother state covariances, these covariances neglect a crucial source of uncertainty: *model uncertainty*. As a consequence, when, e.g., a breath is encountered that cannot be described well given the current set of parameter estimates, uncertainty estimates will still *shrink* (except for the influence of the process noise, that is), simply because more samples have been observed. Proper uncertainty quantification is an old yet burgeoning research field [76]; one rather recent approach that might be worthwhile investigating for the present problem is the identification of a Gaussian process model of equation errors in eq. (7.13) [15, 88].

An alternative venue for further improving the reliability and quality of the estimation

Uncertainty Quantification

Identification
Maneuvers

might be to perform dedicated *identification maneuvers*, as many other researchers have done [11, 12, 79, 93]. In this regard, it is noteworthy that since the proposed method does not demand any particular type of maneuvers, less intrusive maneuvers than inspiratory or expiratory hold maneuvers could be conceived that would still improve observability. For example, inspiratory or expiratory flow limitations or temporarily varied levels of pressure support might similarly assist identification while not being as uncomfortable for patients as occlusions. There are large bodies of research on *active inference* [32, 33], dual control [87] and *optimal design of experiments* [67] that could be leveraged for the automated design of ideal ventilatory maneuvers. This remains to be explored in future studies.

Given the recognized importance of monitoring P_{mus} during assisted mechanical ventilation [13, 34, 78], the proposed method may — potentially with some further improvements and moreover assuming similarly positive results in future, more comprehensive performance evaluations — be attractive for future clinical use. One decisive advantage of an sEMG-based monitoring solution is that it is the only currently available means for measuring the activity of respiratory muscles *other* than the diaphragm: the activity of any accessory muscles can be monitored, and their contribution to P_{mus} quantified. In principle, the individual terms $\alpha^i\overline{\mathrm{EMG}}^i$, $\alpha^d\overline{\mathrm{EMG}}^d$, and $\alpha^r\overline{\mathrm{EMG}}^r$ quantify the pressure contribution of the muscles measured by the different EMG channels, as long as the signals are clean and there is no crosstalk. The validity of the estimated contributions is, however, hard to quantify since there is no reference measurement available to compare to. Nevertheless, this could represent an attractive avenue for future research since the monitoring of respiratory muscle activity has been recognized as clinically important [22, 73], Finally, besides monitoring, the proposed method could also be used to implement a

Proportional
Support

proportional support mode [46, 65], see fig. 7.14. Proportional support modes have been shown to improve patient-ventilator interaction [38], even if P_{mus} is not known with a high degree of accuracy [93].
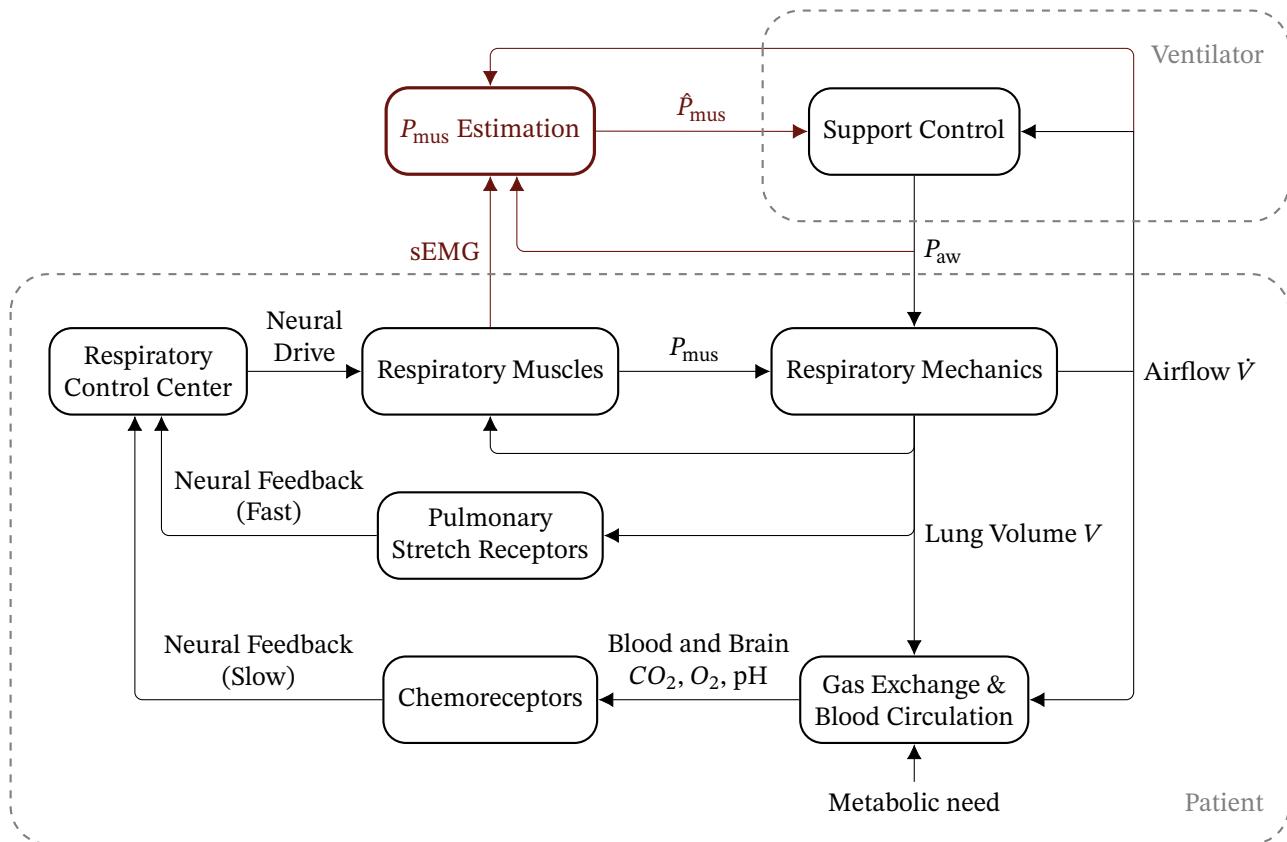
Figure 7.14 – A (simplified) control systems view of sEMG-based proportional support. The patient receives positive pressure support proportional to her own respiratory effort P_{mus}, which is estimated using an sEMG-based method such as the one proposed in this chapter.

References

[1] M. Abramson *et al.*, *The NOMAD project*. [Online]. Available: `https://www.gerad.ca/nomad/`.

[2] H. AbuNurah, D. Russell, and J. Lowman, "The validity of surface EMG of extra-diaphragmatic muscles in assessing respiratory responses during mechanical ventilation: A systematic review," *Pulmonology*, vol. 26, no. 6, pp. 378–385, 2020. DOI: `10.1016/j.pulmoe.2020.02.008`.

[3] H. Akaike, "A new look at the statistical model identification," *IEEE Transactions on Automatic Control*, vol. 19, no. 6, pp. 716–723, 1974. DOI: `10.1109/tac.1974.1100705`.

[4] H. Akaike, "Information theory and an extension of the maximum likelihood principle," in *Proceeding of the Second International Symposium on Information Theory*, B. N. Petrov and F. Caski, Eds., Akademiai Kiado, Budapest, 1973, pp. 267–281. DOI: `10.1007/978-1-4612-1694-0_15`.

[5] E. Akoumianaki *et al.*, "The application of esophageal pressure measurement in patients with respiratory failure," vol. 189, no. 5, pp. 520–531, 2014. DOI: `10.1164/rccm.201312-2193ci`.

[6]   V. Amargiannitakis *et al.*, "Validation of a proposed algorithm for assistance titration during proportional assist ventilation with load-adjustable gain factors," vol. 65, no. 1, pp. 36–44, 2019. DOI: 10.4187/respcare.06988.

[7]   E. Aron, "Der intrapleurale Druck beim lebenden, gesunden Menschen," *Archiv für Pathologische Anatomie und Physiologie und für Klinische Medicin*, vol. 160, no. 1, pp. 226–234, 1900. DOI: 10.1007/bf01939079.

[8]   ——, "Zur Ursache der Einwirkung verdichteter und verdünnter Luft auf den Thierkörper," *Archiv für Pathologische Anatomie und Physiologie und für Klinische Medicin*, vol. 170, no. 2, pp. 264–284, 1902. DOI: 10.1007/bf01929373.

[9]   J. H. T. Bates, *Lung Mechanics: An Inverse Modeling Approach.* Cambridge University Press, 2009, ISBN: 978-0-521-50960-2.

[10]   G. Bellani *et al.*, "The ratio of inspiratory pressure over electrical activity of the diaphragm remains stable during ICU stay and is not related to clinical outcome," *Respiratory Care*, vol. 61, no. 4, pp. 495–501, 2016. DOI: 10.4187/respcare.04400.

[11]   G. Bellani *et al.*, "Estimation of patient's inspiratory effort from the electrical activity of the diaphragm," *Critical Care Medicine*, vol. 41, no. 6, pp. 1483–1491, 2013. DOI: 10.1097/ccm.0b013e31827caba0.

[12]   G. Bellani *et al.*, "Measurement of diaphragmatic electrical activity by surface electromyography in intubated subjects and its relationship with inspiratory effort," *Respiratory Care*, vol. 63, no. 11, pp. 1341–1349, 2018. DOI: 10.4187/respcare.06176.

[13]   M. Bertoni, S. Spadaro, and E. C. Goligher, "Monitoring patient respiratory effort during mechanical ventilation: Lung and diaphragm-protective ventilation," *Critical Care*, vol. 24, no. 1, 2020. DOI: 10.1186/s13054-020-2777-y.

[14]   M. Bertoni *et al.*, "A novel non-invasive method to detect excessively high respiratory effort and dynamic transpulmonary driving pressure during mechanical ventilation," vol. 23, no. 1, 2019. DOI: 10.1186/s13054-019-2617-0.

[15]   I. Bilionis, N. Zabaras, B. A. Konomi, and G. Lin, "Multi-output separable Gaussian process: Towards an efficient, fully Bayesian paradigm for uncertainty quantification," *Journal of Computational Physics*, vol. 241, pp. 212–239, 2013. DOI: 10.1016/j.jcp.2013.01.011.

[16]   J. M. Bland and D. G. Altman, "Agreement between methods of measurement with multiple observations per individual," *Journal of Biopharmaceutical Statistics*, vol. 17, no. 4, pp. 571–582, 2007. DOI: 10.1080/10543400701329422.

[17]   K. P. Burnham and D. R. Anderson, "Multimodel inference," *Sociological Methods & Research*, vol. 33, no. 2, pp. 261–304, 2004. DOI: 10.1177/0049124104268644.

[18]   G. Cammarota *et al.*, "Esophageal balloon calibration during sigh: A physiologic, randomized, cross-over study," vol. 61, pp. 125–132, 2021. DOI: 10.1016/j.jcrc.2020.10.021.

[19]   G. Carteaux *et al.*, "Patient-ventilator asynchrony during noninvasive ventilation," vol. 142, no. 2, pp. 367–376, 2012. DOI: 10.1378/chest.11-2279.

[20]   G. Carteaux *et al.*, "Bedside adjustment of proportional assist ventilation to target a predefined range of respiratory effort," vol. 41, no. 9, pp. 2125–2132, 2013. DOI: 10.1097/ccm.0b013e31828a42e5.

[21]   R. M. Cherniack, L. E. Farhi, B. W. Armstrong, and D. F. Proctor, "A comparison of esophageal and intrapleural pressure in man," *Journal of Applied Physiology*, vol. 8, no. 2, pp. 203–211, 1955. DOI: 10.1152/jappl.1955.8.2.203.

[22]   J. Doorduin, H. W. H. van Hees, J. G. van der Hoeven, and L. M. A. Heunks, "Monitoring of the respiratory muscles in the critically ill," *American Journal of Respiratory and Critical Care Medicine*, vol. 187, no. 1, pp. 20–27, 2013. DOI: 10.1164/rccm.201206-1117cp.

[23] A. Dornhorst and G. Leathart, "A method of assessing the mechanical properties of lungs and air-pssages," *The Lancet*, vol. 260, no. 6725, pp. 109–111, 1952. DOI: 10.1016/s0140-6736(52)92151-x.

[24] J. J. Dowling, "The use of electromyography for the noninvasive prediction of muscle forces," *Sports Medicine*, vol. 24, no. 2, pp. 82–96, 1997. DOI: 10.2165/00007256-199724020-00002.

[25] M. L. Duiverman *et al.*, "Reproducibility and responsiveness of a noninvasive EMG technique of the respiratory muscles in COPD patients and in healthy subjects," *Journal of Applied Physiology: Respiratory, Environmental and Exercise Physiology*, vol. 96, no. 5, pp. 1723–1729, 2004. DOI: 10.1152/japplphysiol.00914.2003.

[26] M. Eger, L. Kahl, P. Rostalski, and E. Petersen, "Apparatus, method, and computer program for influencing the respiration of a person," DE112020000114A5 / WO2020188069A1, 2019.

[27] M. Eger, E. Petersen, J. Graßhoff, and P. Rostalski, "Method and signal processing unit for determining the respiratory activity of a patient," DE112020000233A5 / WO2021052791A1, 2019.

[28] M. Eger, P. Rostalski, E. Petersen, and J. Graßhoff, "Method and signal processing unit for determining the respiratory activity of a patient," DE102019007717B3 / WO2021089215A1, 2019.

[29] ——, "Process and signal processing unit for determining a pneumatic parameter with the use of a lung-mechanical model and of a gradient model," DE102020133460A1 / US20210205561A1, 2020.

[30] D. Farina, R. Merletti, and D. F. Stegeman, "Biophysics of the generation of EMG signals," in *Electromyography. Physiology, Engineering, and noninvasive Applications*, R. Merletti and P. Parker, Eds. Wiley, 2004. DOI: 10.1002/0471678384.ch4.

[31] G. Foti *et al.*, "End-inspiratory airway occlusion: A method to assess the pressure developed by inspiratory muscles in patients with acute lung injury undergoing pressure support," *American Journal of Respiratory and Critical Care Medicine*, vol. 156, no. 4, pp. 1210–1216, 1997. DOI: 10.1164/ajrccm.156.4.96-02031.

[32] K. Friston *et al.*, "Active inference and epistemic value," *Cognitive Neuroscience*, vol. 6, no. 4, pp. 187–214, 2015. DOI: 10.1080/17588928.2015.1020053.

[33] K. J. Friston, T. Parr, and B. de Vries, "The graphical brain: Belief propagation and active inference," *Network Neuroscience*, vol. 1, no. 4, pp. 381–414, 2017. DOI: 10.1162/netn_a_00018.

[34] E. C. Goligher *et al.*, "Lung- and diaphragm-protective ventilation," vol. 202, no. 7, pp. 950–961, 2020. DOI: 10.1164/rccm.202003-0655cp.

[35] J. Graßhoff, E. Petersen, T. Becher, and P. Rostalski, "Automatic estimation of respiratory effort using esophageal pressure," in *41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Berlin, Germany, 2019. DOI: 10.1109/embc.2019.8856345.

[36] J. Graßhoff *et al.*, "A template subtraction method for the removal of cardiogenic oscillations on esophageal pressure signals," in *39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Jeju Island, Korea, 2017. DOI: 10.1109/EMBC.2017.8037299.

[37] J. Graßhoff *et al.*, "Surface EMG-based quantification of inspiratory effort: A quantitative comparison with Pes," *Critical Care (accepted)*, 2021.

[38] A.-F. Haudebourg *et al.*, "Refractory ineffective triggering during pressure support ventilation: Effect of proportional assist ventilation with load-adjustable gain factors," vol. 11, no. 1, 2021. DOI: 10.1186/s13613-021-00935-0.

[39] G. Hedenstierna, "Esophageal pressure: Benefit and limitations.," *Minerva anestesiologica*, vol. 78, pp. 959–966, 8 2012, ISSN: 1827-1596.

[40] W. Herzog, J. Sokolosky, Y. Zhang, and A. Guimarães, "EMG-force relation in dynamically contracting cat plantaris muscle," *Journal of Electromyography and Kinesiology*, vol. 8, no. 3, pp. 147–155, 1998. DOI: 10.1016/s1050-6411(97)00015-1.

[41] A. Hof, "EMG and muscle force: An introduction," *Human Movement Science*, vol. 3, no. 1-2, pp. 119–153, 1984. DOI: 10.1016/0167-9457(84)90008-3.

[42] V. T. Inman *et al.*, "Relation of human electromyogram to muscular tension," *Electroencephalography and Clinical Neurophysiology*, vol. 4, no. 2, pp. 187–194, 1952. DOI: 10.1016/0013-4694(52)90008-4.

[43] G. A. Iotti *et al.*, "Respiratory mechanics by least squares fitting in mechanically ventilated patients: Applications during paralysis and during pressure support ventilation," vol. 21, no. 5, pp. 406–413, 1995. DOI: 10.1007/bf01707409.

[44] G. James, D. Witten, T. Hastie, and R. Tibshirani, "Linear regression," in *An Introduction to Statistical Learning*, Springer US, 2021, pp. 59–128. DOI: 10.1007/978-1-0716-1418-1_3.

[45] D. Jansen *et al.*, "Estimation of the diaphragm neuromuscular efficiency index in mechanically ventilated critically ill patients," *Critical Care*, vol. 22, no. 1, 2018. DOI: 10.1186/s13054-018-2172-0.

[46] A. H. Jonkman *et al.*, "Proportional modes of ventilation: Technology to assist physiology," *Intensive Care Medicine*, vol. 46, no. 12, pp. 2301–2313, 2020. DOI: 10.1007/s00134-020-06206-z.

[47] R. M. Kacmarek, L. Berra, and J. Villar, "On the road to surface monitoring of diaphragmatic activity in mechanically ventilated patients," *Respiratory Care*, vol. 63, no. 11, pp. 1457–1458, 2018. DOI: 10.4187/respcare.06685.

[48] E. Kondili *et al.*, "Estimation of inspiratory muscle pressure in critically ill patients," *Intensive Care Medicine*, vol. 36, no. 4, pp. 648–655, 2010. DOI: 10.1007/s00134-010-1753-4.

[49] A. A. Koopman *et al.*, "Transcutaneous electromyographic respiratory muscle recordings to quantify patient–ventilator interaction in mechanically ventilated children," *Annals of Intensive Care*, vol. 8, no. 1, 2018. DOI: 10.1186/s13613-018-0359-9.

[50] A. Koutsoukou *et al.*, "Expiratory flow limitation and intrinsic positive end-expiratory pressure at zero positive end-expiratory pressure in patients with adult respiratory distress syndrome," *American Journal of Respiratory and Critical Care Medicine*, vol. 161, no. 5, pp. 1590–1596, 2000. DOI: 10.1164/ajrccm.161.5.9904109.

[51] J. V. Kraaijenga, G. J. Hutten, F. H. de Jongh, and A. H. van Kaam, "Transcutaneous electromyography of the diaphragm: A cardio-respiratory monitor for preterm infants," *Pediatric Pulmonology*, vol. 50, no. 9, pp. 889–895, 2014. DOI: 10.1002/ppul.23116.

[52] S. Le Digabel, "Algorithm 909: NOMAD: Nonlinear optimization with the MADS algorithm," *ACM Transactions on Mathematical Software*, vol. 37, no. 4, pp. 1–15, 2011. DOI: 10.1145/1916461.1916468.

[53] R. van Leuteren *et al.*, "Processing transcutaneous electromyography measurements of respiratory muscles, a review of analysis techniques," *Journal of Electromyography and Kinesiology*, vol. 48, pp. 176–186, 2019. DOI: 10.1016/j.jelekin.2019.07.014.

[54] M. M. Liu, W. Herzog, and H. H. Savelberg, "Dynamic muscle force predictions from EMG: An artificial neural network approach," *Journal of Electromyography and Kinesiology*, vol. 9, no. 6, pp. 391–400, 1999. DOI: 10.1016/s1050-6411(99)00014-0.

[55] L. Ljung, *System Identification: Theory for the User*, Second Edition. Prentice Hall, 1999.

[56] D. G. Lloyd and T. F. Besier, "An EMG-driven musculoskeletal model to estimate muscle forces and knee joint moments in vivo," *Journal of Biomechanics*, vol. 36, no. 6, pp. 765–776, 2003. DOI: 10.1016/s0021-9290(03)00010-1.

[57] H.-A. Loeliger *et al.*, "The factor graph approach to model-based signal processing," *Proceedings of the IEEE*, vol. 95, no. 6, pp. 1295–1322, 2007. DOI: 10.1109/jproc.2007.896497.

[58] E. J. W. Maarsingh *et al.*, "Respiratory muscle activity measured with a noninvasive EMG technique: Technical aspects and reproducibility," *Journal of Applied Physiology: Respiratory, Environmental and Exercise Physiology*, vol. 88, no. 6, pp. 1955–1961, 2000. DOI: 10.1152/jappl.2000.88.6.1955.

[59] H. S. Milner-Brown and R. B. Stein, "The relation between the surface electromyogram and muscular force.," *The Journal of Physiology*, vol. 246, no. 3, pp. 549–569, 1975. DOI: 10.1113/jphysiol.1975.sp010904.

[60] F. Mojoli *et al.*, "In vivo calibration of esophageal pressure in the mechanically ventilated patient makes measurements reliable," *Critical Care*, vol. 20, no. 1, 2016. DOI: 10.1186/s13054-016-1278-5.

[61] G. Natalini *et al.*, "Non-invasive assessment of respiratory muscle activity during pressure support ventilation: Accuracy of end-inspiration occlusion and least square fitting methods," vol. 35, no. 4, pp. 913–921, 2020. DOI: 10.1007/s10877-020-00552-5.

[62] E. Petersen, J. Graßhoff, M. Eger, and P. Rostalski, "Surface EMG-based estimation of breathing effort for neurally adjusted ventilation control," in *IFAC World Congress*, ser. IFAC-PapersOnLine, vol. 53, Berlin, Germany: Elsevier BV, 2020, pp. 16 323–16 328. DOI: 10.1016/j.ifacol.2020.12.654.

[63] J. M. Petit, G. Milic-Emili, and L. Delhez, "Role of the diaphragm in breathing in conscious normal man: An electromyographic study," *Journal of Applied Physiology*, vol. 15, no. 6, pp. 1101–1106, 1960. DOI: 10.1152/jappl.1960.15.6.1101.

[64] C. Pizzolato *et al.*, "CEINMS: A toolbox to investigate the influence of different neural control solutions on the prediction of muscle excitation and joint moments during dynamic motor tasks," *Journal of Biomechanics*, vol. 48, no. 14, pp. 3929–3936, 2015. DOI: 10.1016/j.jbiomech.2015.09.021.

[65] P. von Platen, A. Pomprapa, B. Lachmann, and S. Leonhardt, "The dawn of physiological closed-loop ventilation—a review," *Critical Care*, vol. 24, no. 1, 2020. DOI: 10.1186/s13054-020-2810-1.

[66] H. Prechtl, L. V. Eykern, and M. O'Brien, "Respiratory muscle EMG in newborns: A non-intrusive method," *Early Human Development*, vol. 1, no. 3, pp. 265–283, 1977. DOI: 10.1016/0378-3782(77)90040-8.

[67] F. Pukelsheim, *Optimal design of experiments*. Philadelphia: Society for Industrial and Applied Mathematics (SIAM), 2006, ISBN: 0898716047.

[68] D. P. Redmond, P. D. Docherty, Y. S. Chiew, and J. G. Chase, "A polynomial model of patient-specific breathing effort during controlled mechanical ventilation," IEEE, 2015. DOI: 10.1109/embc.2015.7319402.

[69] G. Rehm *et al.*, "Clinical validation of single-chamber model-based algorithms used to estimate respiratory compliance," 2021. arXiv: 2109.10224 [q-bio.QM].

[70] L. Roesthuis, M. van den Berg, and H. van der Hoeven, "Non-invasive method to detect high respiratory effort and transpulmonary driving pressures in COVID-19 patients during mechanical ventilation," vol. 11, no. 1, 2021. DOI: 10.1186/s13613-021-00821-9.

[71] E. Saatci and A. Akan, "Lung model parameter estimation by unscented Kalman filter," IEEE, 2007. DOI: 10.1109/iembs.2007.4352850.

[72] M. Scheel, A. Berndt, and O. Simanski, "Application of Kalman filter for breathing effort reconstruction for OSAS patients in breathing therapy," vol. 66, no. 12, pp. 1064–1071, 2018. DOI: 10.1515/auto-2018-0067.

[73] Z.-H. Shi *et al.*, "Expiratory muscle dysfunction in critically ill patients: Towards improved understanding," *Intensive Care Medicine*, vol. 45, no. 8, pp. 1061–1071, 2019. DOI: 10.1007/s00134-019-05664-4.

[74] C. Sinderby *et al.*, "Neural control of mechanical ventilation in respiratory failure," *Nature Medicine*, vol. 5, no. 12, pp. 1433–1436, 1999. DOI: 10.1038/71012.

[75] C. A. Sinderby, J. C. Beck, L. H. Lindström, and A. E. Grassino, "Enhancement of signal quality in esophageal recordings of diaphragm EMG," *Journal of Applied Physiology: Respiratory, Environmental and Exercise Physiology*, vol. 82, no. 4, pp. 1370–1377, 1997. DOI: 10.1152/jappl.1997.82.4.1370.

[76] R. Smith, *Uncertainty Quantification: Theory, Implementation, and Applications*. Society for industrial and applied mathematics (SIAM), Philadelphia, USA, 2014, 400 pp., ISBN: 161197321X.

[77] S. Soundoulounaki *et al.*, "Airway pressure morphology and respiratory muscle activity during end-inspiratory occlusions in pressure support ventilation," vol. 24, no. 1, 2020. DOI: 10.1186/s13054-020-03169-x.

[78] E. Spinelli *et al.*, "Respiratory drive in the acute respiratory distress syndrome: Pathophysiology, monitoring, and therapeutic interventions," *Intensive Care Medicine*, 2020. DOI: 10.1007/s00134-020-05942-6.

[79] I. Telias, F. Damiani, and L. Brochard, "The airway occlusion pressure (P0.1) to monitor respiratory drive during mechanical ventilation: Increasing awareness of a not-so-new problem," vol. 44, no. 9, pp. 1532–1535, 2018. DOI: 10.1007/s00134-018-5045-8.

[80] J.-A. Ting, E. Theodorou, and S. Schaal, "A Kalman filter for robust outlier detection," in *2007 IEEE/RSJ International Conference on Intelligent Robots and Systems*, IEEE, 2007. DOI: 10.1109/iros.2007.4399158.

[81] T. Tokizane, K. Kawamata, and H. Tokizane, "Electromyographic studies on the human respiratory muscles," *The Japanese Journal of Physiology*, vol. 2, pp. 232–247, 1951. DOI: 10.2170/jjphysiol.2.232.

[82] S. Van Huffel and P. Lemmerling, Eds., *Total Least Squares and Errors-in-Variables Modeling*. Springer Netherlands, 2002. DOI: 10.1007/978-94-017-3552-0.

[83] F. Vicario *et al.*, "Noninvasive estimation of respiratory mechanics in spontaneously breathing ventilated patients: A constrained optimization approach.," *IEEE Transactions on Biomedical Engineering*, vol. 63, pp. 775–787, 4 2015. DOI: 10.1109/tbme.2015.2470641.

[84] F. Vicario *et al.*, "Simultaneous parameter and input estimation of a respiratory mechanics model," in *Modeling, Simulation and Optimization of Complex Processes HPSC 2015*, Springer International Publishing, 2017, pp. 235–247. DOI: 10.1007/978-3-319-67168-0_19.

[85] L. Vignaux *et al.*, "Patient–ventilator asynchrony during non-invasive ventilation for acute respiratory failure: A multicenter study," *Intensive Care Medicine*, vol. 35, pp. 840–846, 2009. DOI: 10.1007/s00134-009-1416-5.

[86] F. Wadehn *et al.*, "Outlier-insensitive Kalman smoothing and marginal message passing," in *2016 24th European Signal Processing Conference (EUSIPCO)*, IEEE, 2016. DOI: 10.1109/eusipco.2016.7760447.

[87] B. Wittenmark, "Adaptive dual control methods: An overview," in *Adaptive Systems in Control and Signal Processing*, Elsevier, 1995, pp. 67–72. DOI: 10.1016/b978-0-08-042375-3.50010-x.

[88] X. Wu, T. Kozlowski, H. Meidani, and K. Shirvan, "Inverse uncertainty quantification using the modular Bayesian approach based on Gaussian process, part 1: Theory," *Nuclear Engineering and Design*, vol. 335, pp. 339–355, 2018. DOI: 10.1016/j.nucengdes.2018.06.004.

[89] Y. Yamada, M. Shigeta, K. Suwa, and K. Hanaoka, "Respiratory muscle pressure analysis in pressure-support ventilation," *Journal of Applied Physiology*, vol. 77, no. 5, pp. 2237–2243, 1994. DOI: 10.1152/jappl.1994.77.5.2237.

[90] M. Younes, W. Riddle, and J. Polacheck, "A model for the relation between respiratory neural and mechanical outputs. III. Validation," *Journal of Applied Physiology*, vol. 51, no. 4, pp. 990–1001, 1981, PMID: 7298442. DOI: 10.1152/jappl.1981.51.4.990.

[91] M. Younes *et al.*, "A method for measuring passive elastance during proportional assist ventilation," *American Journal of Respiratory and Critical Care Medicine*, vol. 164, no. 1, pp. 50–60, 2001. DOI: 10.1164/ajrccm.164.1.2010068.

[92] M. Younes, "Proportional-assist ventilation," in *Principles and Practice of Mechanical Ventilation*, Third edition. New York, NY: McGraw-Hill, 2013, ch. 12. [Online]. Available: accessmedicine.mhmedical.com/content.aspx?aid=57064841.

[93] M. Younes *et al.*, "A method for monitoring and improving patient: Ventilator interaction," *Intensive Care Medicine*, vol. 33, no. 8, pp. 1337–1346, 2007. DOI: 10.1007/s00134-007-0681-4.

[94] G. Zou, "Confidence interval estimation for the bland–altman limits of agreement with multiple observations per individual," *Statistical Methods in Medical Research*, vol. 22, no. 6, pp. 630–642, 2011. DOI: 10.1177/0962280211402548.

CONCLUSION & OUTLOOK

A great discovery solves a great problem but there is a grain of discovery in the solution of any problem. Your problem may be modest; but if it challenges your curiosity and brings into play your inventive faculties, and if you solve it by your own means, you may experience the tension and enjoy the triumph of discovery.

(George Pólya, How to solve it)

Contents

A S THE EVIDENCE for the clinical benefits of monitoring the respiratory drive of critically ill patients has become more and more compelling [3, 8], interest in using respiratory surface EMG measurements for this purpose has risen. There are, however, important signal processing challenges to overcome in order to obtain a reliable monitoring tool that is suitable for the application in respiratory care. In this thesis, several contributions towards this specific purpose have been described, along with contributions to the broader fields of electromyography and statistical inference. Throughout all application challenges considered, model-based, probabilistic inference has been employed.

8.1 Summary

In chapter 4, a comprehensive mathematical model of surface electromyography, muscular force generation, and motor unit pool organization has been described. The model integrates many different aspects of muscular physiology and incorporates many new models for individual subcomponents. The distribution of motor unit and muscle fiber properties was modeled *stochastically* and *continuously* (as opposed to assuming a fixed number of MU and fiber types). The electrical and mechanical properties of muscle fibers were combined deliberately such that a whole simulated muscle attains a realistic EMG-force relationship. Moreover, an alternative reformulation of a previously proposed mathematical model of action potential propagation was described, and several physiologically meaningful mathematical properties of the model were proved. The model was then used to validate a newly proposed blind source separation (BSS) algorithm for the separation of the activity of multiple respiratory muscles in multi-channel respiratory surface EMG measurements. (The derivation of said algorithm was not part of this thesis.) The algorithm is based on a convolutive mixing model and exploits very little prior knowledge about the nature of the sources and the mixing system. Given the difficulty of the separation task, the algorithm yielded good separation results, removing significant amounts of cardiac interference and mostly recovering the original source activity. However, it appears likely that even better-performing algorithms could be designed by exploiting more information about the system at hand. One way of simplifying the separation task is to separate the problems of cardiac artifact removal and muscle crosstalk separation.

Chapter 5 then was wholly dedicated to the task of removing cardiac artifacts from single-channel respiratory surface EMG measurements. To this end, a novel artifact removal algorithm was presented that represents a hybrid between probabilistic, model-based, and nonparametric approaches. A fully nonparametric characterization of the individual cardiac beat shape was combined with a probabilistic model for the temporal evolution of the cardiac beat shape between successive beats and a principled probabilistic inference procedure for learning all relevant model and noise parameters from a single recording, using an array of separate linear Kalman filters and smoothers. The algorithm was then compared to a set of previously proposed algorithms within a comprehensive performance evaluation framework. All algorithms were compared based on both real measurement signals obtained in a study with healthy volunteers, as well as on synthetic signals obtained from the superposition of leg EMG measurements and singe-lead ECG signals. A total of four different evaluation measures was considered, with the newly proposed algorithm performing best for recovering the raw, uncontaminated EMG signals. A secondary finding was that if only a clean EMG *envelope* signal is of interest (as opposed

to the raw signal), a simple high-pass filter with a high cut-off frequency combined with a subsequent moving median-based envelope calculation is already sufficient, and there is no need for more complex algorithms.

Motivated by the estimation problem to be pursued in the subsequent chapter 7, chapter 6 treated a theoretical statistical question: how can one perform parameter inference in the face of *covariate shift*, *concept drift* (i.e., time-varying model parameters), and *model mismatch*? This setting turned out to be surprisingly unexplored in the literature. A novel, principled solution approach for this challenging estimation setting was proposed, based on a combination of importance weighting, Kalman filtering and smoothing, and a probabilistic description of regression in the face of concept drift in a state-space model. Weighted variants of the classical Kalman filter and smoother were derived, which are simple to implement and of equal computational complexity as the standard algorithms. Moreover, a simple and efficient multivariate density estimation scheme and a simple, importance-weighted hyperparameter tuning scheme were proposed, the latter based on weighted prediction error minimization. Numerical examples support the claim that the proposed methodology can be used to successfully perform inference in the face of covariate shift, concept drift, and model mismatch.

Finally, chapter 7 addressed the problem of inferring the pressure P_{mus}, generated by the respiratory muscles of patients under assisted mechanical ventilation, based on both surface EMG measurements of the respiratory muscles and standard pneumatic measurements provided by the mechanical ventilator. To this end, three model components were employed: a classical model of respiratory mechanics, a new (very simple) model relating multiple respiratory EMG measurements and P_{mus}, and a new (also very simple) model of the dynamic time course of P_{mus}. Evidence of the adequacy of the three selected model components was provided based on model analyses in eleven recordings from intensive care patients. Based on these three model components, a probabilistic sensor fusion method was described that uses all available measurement modalities and (almost) all available measurement samples. As the inference problem suffers from covariate shift, concept drift, and model mismatch, the importance-weighted estimation scheme described in the previous chapter was exploited here. An extensive evaluation of different variants of the proposed method on intensive care data indicated the high estimation quality achieved by the proposed algorithm compared to (invasive) gold standard estimates based on esophageal pressure P_{es}. Importance weighting and delay estimation were found to significantly improve the quality of the resulting estimates $\hat{P}_{\mathrm{mus}}$.

8.2 Discussion & Outlook

Many potentially beneficial detail improvements to the algorithms, models, and evaluation strategies described throughout this thesis have already been discussed in the corresponding chapters; these shall not be reiterated here. Instead, this final section shall provide a broader perspective on what has been achieved, as well as potential future uses, improvements, and alternatives to the methods presented in this thesis.

Inter-
disciplinarity
The project pursued in this thesis is highly *interdisciplinary* at its core: it lies at the intersection of the research fields investigating respiratory support, electromyography, and statistical inference. Successful implementation of a surface-EMG-based respiratory monitoring methodology will require deep knowledge in all three fields and will benefit

tremendously from contributions by experts from all three fields.[1] From the perspective of clinical respiratory support research, the realization of the application investigated in this thesis is highly desirable to improve patient treatment outcomes. Thus, practical solutions to implementing a highly robust and practically useful monitoring method are sought after. From the perspective of electromyographic research, the respiratory application may be of particular interest not only because of its intrinsic value but also because it is quite different from existing applications, which are usually in ergonomics, sports medicine, or prosthetics research. On the one hand, the application is analogous to many classical biomechanical problems in many ways if one understands the mechanical ventilator as a "respiratory prosthesis" and respiratory mechanics as an analogon to a biomechanical joint model. On the other hand, besides the differences in muscle properties, geometry, and biomechanics, the application on intensive care patients comes with many challenges and characteristics that differ from those encountered in typical EMG applications. Incidentally, such an application in intensive care may help to accelerate the translation of EMG research into clinical practice: it has been hypothesized that the fact that EMG is usually not used to treat life-threatening problems may have historically posed a barrier to its widespread clinical translation [12]; this would certainly be alleviated by the application discussed in chapter 7. Finally, from the perspective of statistical inference, the application under consideration comes with a convergence of many theoretically interesting challenges: relatively small study population sizes, significant model mismatch and various sources of outliers and disturbances, concept drift and covariate shift, high inter-patient variability, unsupervised estimation (even if a reference measurement is theoretically available, it will ultimately not be available in the patients the method is applied to), and a lack or low quality of reference measurements even in study populations. (For blind source separation and cardiac artifact removal, no reference measurements are available; for P_{mus} estimation, the reference measurement is also subject to significant sources of disturbances.) As has historically often been the case, these practical challenges may spur theoretical innovation, such as the modest developments presented in chapter 6.

The encouraging results presented particularly in chapter 7 support hopes that surface EMG-based respiratory monitoring may soon find its way into the clinic to patients' great benefit. There are, however, technical challenges remaining to be addressed before these hopes can be realized (besides the obvious need for a more comprehensive evaluation in clinical trials). Most importantly, in this thesis, all algorithms were developed in an offline fashion, i.e., they utilize information from the whole recording to which they are applied. Depending on the specific application — retrospective monitoring, real-time monitoring, or proportional support — , algorithms with varying degrees of *real-time* capabilities would need to be developed. This, while certainly possible, is nontrivial in all three application problems considered in this thesis: source separation for crosstalk detection and removal, cardiac artifact removal, and P_{mus} estimation. Concerning source separation, various real-time capable blind source separation frameworks have been proposed in the literature, including variants of the TRINICON framework that has been employed here [1, 4]. In the EMG-based prosthetics control community, multiple groups have addressed the application of non-negative matrix factorization (NMF)-based algorithms for real-time prosthesis control [5, 10, 11]. Regarding cardiac artifact removal, the challenge posed by real-time

Real-Time Application

---

[1] In fact, this thesis has benefited vastly from the author attending conferences and workshops from all three fields and receiving generous advice from experts in all three.

applications depends mainly on the employed algorithm and the exact requirements. While requirements that allow for a delay of up to one heartbeat should be comparatively simple to satisfy, any requirements below that threshold will pose significant challenges to any of the assessed methods and require significant further research efforts. However, it is noteworthy in the context of this thesis that for controlling a mechanical ventilator, a combination of simple gating and signal prediction during the 'gaps' appears likely to perform sufficiently well (due to the typical smoothness and low rate of change of P_{mus}). Finally, the P_{mus} estimation algorithm presented in chapter 7 can, in principle, be applied in real-time — up to the EMG signal processing delay. Incidentally, as the delay analysis in chapter 7 indicated (and as is known from basic physiology, see fig. 3.3), there is a significant although small delay between the raw EMG signal and P_{mus}, indicating that if an appropriate EMG signal processing chain with a low enough delay could be identified, the resulting EMG-based estimate $\hat{P}_{\text{mus}}$ could be available almost synchronously to $P_{\text{mus-Pes}}$.

Integration of Preprocessing and Estimation

The challenges of source separation, cardiac artifact removal, and estimation of P_{mus}, although closely related, have been mainly treated in isolation in this thesis. It appears likely that further gains in estimation performance concerning $\hat{P}_{\text{mus}}$ could be achieved through a tight *integration* between these different processing steps. As discussed in chapter 4 already, the source separation task could be significantly simplified if a dedicated cardiac artifact removal procedure were to be applied first. Then, instead of the original EMG envelope signals, one could use the source signals identified by a source separation procedure as the input signals for estimating P_{mus}. Moreover, uncertainty estimates of both preprocessing steps and general signal quality indices could be incorporated into the estimation procedure described in chapter 7 by assigning time-varying and patient-dependent measurement noise covariances based on the results of these preprocessing steps.

Machine Learning

Method-wise, all three application chapters (chapters 4, 5 and 7) have shown the importance of selecting suitable system and noise models for achieving accurate results, and in all three chapters, further model refinements represent one possible avenue for improvement — human physiology is complex and only rarely fully abides by simple, parametric models. In present days, a standard response to this problem is to use very flexible high-dimensional (*nonparametric* or *machine learning*) models that are — in principle — capable of accurately representing arbitrarily complex relationships. There are, however, serious challenges to overcome in applying such high-dimensional models to the problems discussed in this thesis, including low study population size, high inter-patient variability, patient variability over time, generally low signal-to-noise ratio as is typical for biomedical recordings, and a lack of gold-standard reference signals. Model underspecification is known to cause serious challenges even when huge datasets are available [7] and has demonstrably led to shortcut learning in many medical applications [2, 17]. Linked to the challenges of proper evaluation [7, 14], various studies have reported highly exaggerated performance claims for machine learning methods in medical settings [6, 9, 18] — and all of these problems were reported in applications based on *much* larger datasets than the ones available for time series problems in the intensive care setting. There are, thus, reasonable arguments to be made for the use of simple, parametric models based on prior knowledge for the applications discussed in this thesis. (The author has recently published a survey of challenges and potential solutions to the responsible and regulatory conform application of machine learning methods in medicine [13], touching

upon these and other issues.)

Importance weighting, as discussed introduced in section 2.1.6 and further exploited in chapter 6, represents an alternative avenue for improving inferences based on imperfect models. The fundamental idea is, simply, to weight samples based on how important it is for the model to fit this particular sample well. While traditionally, this approach has been applied with a known (or estimated) target distribution $p_{\text{target}}(x)$ to obtain a model that performs optimally on data drawn from that distribution [15], the target distribution $p_{\text{target}}(x)$ can also be understood as a *design variable*: in which regions of the data space would we like the identified model to perform well? Which regions do we consider more important than others? In chapter 7, it was the application of a weighting scheme based on such considerations that led to a marked improvement in estimation performance. The methods derived in chapter 6, however, are applicable to the fully general case of regression under covariate shift, concept drift, and model misspecification and could thus find applications far beyond this thesis. Extending these methods to the case of *classification* under the same assumptions (covariate shift, concept drift, model misspecification), using, for instance, the framework of Wilkinson *et al.* [16], represents an attractive avenue for future research.

References

[1] R. Aichner, H. Buchner, F. Yan, and W. Kellermann, "A real-time blind source separation scheme and its application to reverberant and noisy acoustic environments," *Signal Processing*, vol. 86, no. 6, pp. 1260–1277, 2006. DOI: 10.1016/j.sigpro.2005.06.022.

[2] M. A. Badgeley *et al.*, "Deep learning predicts hip fracture using confounding patient and healthcare variables," *npj Digital Medicine*, vol. 2, no. 1, 2019. DOI: 10.1038/s41746-019-0105-1.

[3] M. Bertoni, S. Spadaro, and E. C. Goligher, "Monitoring patient respiratory effort during mechanical ventilation: Lung and diaphragm-protective ventilation," *Critical Care*, vol. 24, no. 1, 2020. DOI: 10.1186/s13054-020-2777-y.

[4] H. Buchner, R. Aichner, and W. Kellermann, "TRINICON-based blind system identification with application to multiple-source localization and separation," in *Blind Speech Separation*, S. Makino, T.-W. Lee, and S. Sawada, Eds., Springer, 2007. DOI: 10.1007/978-1-4020-6479-1_4.

[5] L. Cerina, P. Cancian, G. Franco, and M. D. Santambrogio, "A hardware acceleration for surface EMG non-negative matrix factorization," in *2017 IEEE International Parallel and Distributed Processing Symposium Workshops (IPDPSW)*, IEEE, 2017. DOI: 10.1109/ipdpsw.2017.66.

[6] E. Christodoulou *et al.*, "A systematic review shows no performance benefit of machine learning over logistic regression for clinical prediction models," *Journal of Clinical Epidemiology*, vol. 110, pp. 12–22, 2019. DOI: 10.1016/j.jclinepi.2019.02.004.

[7] A. D'Amour *et al.*, "Underspecification presents challenges for credibility in modern machine learning," 2020. arXiv: 2011.03395 [cs.LG].

[8] E. C. Goligher *et al.*, "Clinical strategies for implementing lung and diaphragm-protective ventilation: Avoiding insufficient and excessive effort," *Intensive Care Medicine*, vol. 46, no. 12, pp. 2314–2326, 2020. DOI: 10.1007/s00134-020-06288-9.

[9] R. Jacobucci *et al.*, "Evidence of inflated prediction performance: A commentary on machine learning and suicide research," *Clinical Psychological Science*, vol. 9, no. 1, pp. 129–134, 2021. DOI: 10.1177/2167702620954216.

[10] P. Kim, K.-S. Kim, and S. Kim, "Modified nonnegative matrix factorization using the hadamard product to estimate real-time continuous finger-motion intentions," *IEEE Transactions on Human-Machine Systems*, vol. 47, no. 6, pp. 1089–1099, 2017. DOI: 10.1109/thms.2017.2751549.

[11] C. Lin, B. Wang, N. Jiang, and D. Farina, "Robust extraction of basis functions for simultaneous and proportional myoelectric control via sparse non-negative matrix factorization," *Journal of Neural Engineering*, vol. 15, no. 2, p. 026 017, 2018. DOI: 10.1088/1741-2552/aa9666.

[12] B. J. Martin and Y. Acosta-Sojo, "sEMG: A window into muscle work, but not easy to teach and delicate to practice—a perspective on the difficult path to a clinical tool," *Frontiers in Neurology*, vol. 11, 2021. DOI: 10.3389/fneur.2020.588451.

[13] E. Petersen *et al.*, "Responsible and regulatory conform machine learning for medicine: A survey of challenges and solutions," *IEEE Access*, vol. 10, pp. 58 375–58 418, 2022. DOI: 10.1109/access.2022.3178382.

[14] E. W. Steyerberg and F. E. Harrell, "Prediction models need appropriate internal, internal–external, and external validation," *Journal of Clinical Epidemiology*, vol. 69, pp. 245–247, 2016. DOI: 10.1016/j.jclinepi.2015.04.005.

[15] M. Sugiyama *et al.*, "Direct importance estimation for covariate shift adaptation," *Annals of the Institute of Statistical Mathematics*, vol. 60, no. 4, pp. 699–746, 2008. DOI: 10.1007/s10463-008-0197-x.

[16] W. Wilkinson, P. Chang, M. Andersen, and A. Solin, "State space expectation propagation: Efficient inference schemes for temporal Gaussian processes," in *Proceedings of the 37th International Conference on Machine Learning*, H. Daumé III and A. Singh, Eds., vol. 119, PMLR, 2020, pp. 10 270–10 281. [Online]. Available: http://proceedings.mlr.press/v119/wilkinson20a.html.

[17] J. K. Winkler *et al.*, "Association between surgical skin markings in dermoscopic images and diagnostic performance of a deep learning convolutional neural network for melanoma recognition," *JAMA Dermatology*, vol. 155, no. 10, p. 1135, 2019. DOI: 10.1001/jamadermatol.2019.1735.

[18] L. Wynants *et al.*, "Prediction models for diagnosis and prognosis of covid-19: Systematic review and critical appraisal," *BMJ*, vol. 369, p. m1328, 2020. DOI: 10.1136/bmj.m1328.

APPENDIX

Contents

1 Distributions

The following minimal overview of the theory of continuous linear functionals is reproduced verbatim from and with permission of Petersen and Rostalski [2] for the reader's convenience.

A distribution T is a continuous linear mapping $T : D \to \mathbb{R}$, where D is a given set of so-called test functions. In the following, the value of a distribution T acting on a test function ζ shall be denoted by the duality $\langle T, \zeta \rangle$, with test functions chosen from the set D of compactly supported, smooth functions. Furthermore, two basic properties of distributions are

$$\langle \eta T, \zeta \rangle = \langle T, \eta \zeta \rangle \qquad \forall \eta \in C^\infty, \zeta \in D$$

and

$$\langle T', \zeta \rangle = -\langle T, \zeta' \rangle \qquad \forall \zeta \in D.$$

Moreover (following from the above), considering $\eta \in C^\infty$,

$$\langle (\eta H)', \zeta \rangle = -\langle \eta H, \zeta' \rangle = - \int_{-\infty}^{\infty} x \, \mathrm{d}\,(\eta H)(x)\, \zeta'(x)$$

$$= -[\eta \zeta]_0^\infty + \int_0^\infty x \, \mathrm{d}\eta' \zeta = \langle \delta, \eta \zeta \rangle + \langle \eta' H, \zeta \rangle$$

$$= \langle \eta' H + \eta\, \delta, \zeta \rangle.$$

For more background on distributions (i.e., continuous linear functionals), refer to, e.g., Clarke [1].

2 Multivariate Gaussian calculus

The multivariate Gaussian (or „normal") distribution generalizes the one-dimensional Gaussian (or normal) distribution. A random vector $\mathcal{X} : \Omega \to \mathbb{R}^n$ is normally distributed, i.e.,

$$\mathcal{X} \sim \mathcal{N}(\mu, \Sigma),$$

if every linear combination of its components follows a univariate normal distribution. If the covariance matrix Σ is positive definite, the distribution is said to be „non-degenerate", and in this case, the probability density function (PDF) of the multivariate Gaussian distribution is given by

$$p(x) = \frac{1}{\sqrt{(2\pi)^n |\Sigma|}} e^{-\frac{1}{2}(x-\mu)^{\mathsf{T}}\Sigma^{-1}(x-\mu)}. \tag{1}$$

For $n = 1$, eq. (1) reduces to the standard expression for the univariate case,

$$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}},$$

where σ^2 denotes the variance. Throughout this thesis, sometimes the alternative notation $\mathcal{N}(x \mid \mu_x, \Sigma_x)$ is used to clarify that the distribution of the variable x is meant.

References

[1]   F. Clarke, *Functional Analysis, Calculus of Variations and Optimal Control*, S. Axler and K. Ribet, Eds., ser. Graduate Texts in Mathematics. Springer, 2013.

[2]   E. Petersen and P. Rostalski, "A comprehensive mathematical model of motor unit pool organization, surface electromyography, and force generation," *Frontiers in Physiology*, vol. 10, p. 176, 2019, ISSN: 1664-042X. DOI: 10.3389/fphys.2019.00176.

List of Figures

LIST OF ALGORITHMS

Curriculum vitae

Eike Petersen
* October 3, 1990 in Salzgitter, Germany

2021–today	**Research associate, DTU Compute.** Technical University of Denmark, Kgs. Lyngby, Denmark. *Research focus*: fair machine learning for medicine.
2015–2021	**Research associate, Universität zu Lübeck.** Institute for Electrical Engineering in Medicine, Lübeck, Germany. *Research foci*: mathematical modeling, signal processing, probabilistic inference, bioelectricity, respiration.
2020–2021	**Associated researcher, KI-SIGS** (AI Space for Intelligent Health Systems). Large, Northern German research consortium, funded by the German BMWi. *Research focus*: safe, responsible, and regulatory conform machine learning for medical applications.
September 2019	**Visiting scholar, IEEE-SPS Summer School** on signal processing for autonomous systems, Arenzano, Italy.
October 2018	**Visiting scholar, Lund Center for Control of Complex Engineering Systems**, Lund University, Sweden. Focus period on learning and adaptation for sensorimotor control.
2012–2015	**Master of Science in Industrial Mathematics, University of Hamburg.** *Focus areas*: Modeling, analysis and control of nonlinear dynamical systems. *Thesis*: Integrative mathematical modelling and simulation of surface electromyographic measurements.
2013–2014	**Erasmus exchange**, INSA Toulouse, France. *Focus area*: Numerical methods for PDE systems.
2009–2012	**Bachelor of Science in Computer Science and Engineering, Technical Technical University of Hamburg,** Germany. *Focus areas*: mathematics, computer science, signal processing. *Thesis*: Patient airway pressure determination in medical ventilators using transmission line theory.
2011	**Summer school in Applied Mathematics in Engineering, University of Sarajevo**, Bosnia & Herzegowina. Scholarship funded by the DAAD.
2009–2015	**Student trainee, Dräger Medical**, Lübeck, Germany. Scientific and technical projects related to the control of medical ventilators.
2011	**Research assistant, Technical University of Hamburg,** Institute for Control Systems, Germany.
2011	**Teaching assistant, Technical University of Hamburg,** Institute of Embedded Systems, Germany.
2009	**Abitur (German high-school diploma)**, Johannes-Brahms-Gymnasium, Pinneberg, Germany.