

Diss. ETH No. 16365

On Graphical Models for Communications and Machine Learning: Algorithms, Bounds, and Analog Implementation

A dissertation submitted to the
Swiss Federal Institute of Technology, Zürich
for the degree of
Doctor of Sciences ETH Zürich

presented by

Justin H. G. Dauwels

ir.

born on November 2, 1977

citizen of Belgium

accepted on the recommendation of
Prof. Dr. Hans-Andrea Loeliger, examiner
Prof. Dr. Shun-ichi Amari, co-examiner
Prof. Dr. Marc Moeneclaey, co-examiner
Dr. Jonathan Yedidia, co-examiner

Hartung-Gorre Verlag, Konstanz, May 2006

Abstract

This dissertation is about a specific problem and about general methods. The specific problem is carrier-phase synchronization, which appears in the context of digital communications. The general methods are message-passing algorithms operating on graphical models, in particular, factor graphs. We consider applications of such algorithms in the context of statistical inference (as in communications, signal processing, and machine learning), statistics, information theory, and the theory of dynamical systems (such as analog electronic circuits).

The primary motivation for this work was (1) to analyze the degradation of digital communications systems due to oscillator non-idealities; (2) the development of synchronization algorithms that minimize this performance degradation.

Clocks are ubiquitous in digital communications systems; real-life clocks are noisy, i.e., their signals are not perfectly periodic, which often leads to a significant degradation of the performance of communications systems.

In the early days of communications, this source of degradation was only of secondary concern. Communications systems used to operate far from the ultimate performance bound, i.e., channel capacity. The main concern was therefore to develop error-correcting techniques that could close the gap between the performance of practical communications systems and channel capacity.

With the recent advent of iterative decoding techniques, communications systems nowadays most often operate close to the ultimate performance limits; issues such as synchronization, which were earlier only of secondary importance, have now become the mayor (remaining) bottlenecks

in the design of communications systems.

In this dissertation, we focus on carrier-phase synchronization, i.e., the alignment of the phase of the local oscillator in the receiver to the phase of the incoming carrier. The questions we address are:

- a) Which physical mechanisms are responsible for phase noise? How can phase noise be modeled?
- b) How can carrier-phase estimation algorithms systematically be derived?
- c) What are the ultimate limits for communication over channels with phase noise? In particular:
 - i) How much does the information rate of a communications channel decrease due to phase noise?
 - ii) How well can the (noisy) carrier phase be estimated?

In contrast to earlier and parallel work, our aim is not the design and optimization of fully operating communications systems. In this thesis, various tools are developed that lead (or may lead) to an answer to the above questions (and many other related questions).

We give a detailed analysis of phase noise in free-running clocks and PLLs (Question 1). We propose a simple intuitive model for phase noise in free-running oscillators. We describe two simple models for passband communications channels. The models take phase offsets into account between the received carrier and the local carrier in the receiver, but disregard timing offsets. In the first model, the phase is constant, in the second, the phase performs a random walk. We investigate under which conditions the two models are valid. Most methods of this thesis will be illustrated by means of both channel models.

Most methods we propose in this dissertation are based on graphical models, more precisely, factor graphs. Factor graphs are used to visualize the structure of the system at hand. They represent the factorization of multivariate functions. Each edge in the graph corresponds to a variable, each node corresponds to a factor. Factor graphs can represent any function, in particular, probabilistic models, error-correcting codes, block diagrams and other common models in communications, signal processing and beyond.

We show how factor graphs can be used as a tool to develop practical estimation and detection algorithms. Our techniques can be applied to model-based signal processing (e.g., phase estimation) and machine learning. In particular, we formulate several standard signal-processing and machine-learning algorithms as message passing on factor graphs, e.g., particle methods, gradient methods, decision-based methods, and expectation maximization. In all those algorithms, local rules are applied at the nodes in a factor graph. In other words, the (global) estimation and detection problem is tackled by a divide-and-conquer strategy: the global computation is carried out by multiple (simple) local computations. The local message-update rules may be used as building blocks for novel estimation and detection algorithms. By listing the possible update rules at each node in the factor graph, one can systematically explore novel algorithms. We demonstrate this idea by deriving phase estimation algorithms for the constant-phase model and the random-walk phase model (Question 2). We also show how the back-propagation algorithm for the training of feed-forward neural networks follows by applying generic message-passing rules. We elaborate on the computation of kernels in the light of message passing on factor graphs.

We demonstrate how message-passing algorithms for inference can be implemented as dynamical systems, in particular, as clock-free analog electronic circuits. Those systems operate in continuous time, and do not require a digital clock; therefore, they circumvent the problem of timing synchronization.

We present a numerical algorithm to compute the information rate of continuous channels with memory (Question 3.a). The algorithm is an extension of the methods proposed earlier for discrete channels with memory. Also here, factor graphs and the summary-propagation algorithm are key ingredients. We apply the method to the random-walk phase model. The algorithms we propose for computing Cramér-Rao-type bounds open the door to exciting applications of information geometry, such as (1) natural-gradient-based algorithms; (2) the computation of Fisher kernels.

We propose a numerical algorithm for computing the capacity (or lower bounds on capacity) of continuous memoryless channels (Question 3.a). We present numerical results for the Gaussian channel with average-power and/or peak-power constraints. We outline how the algorithm can be extended to continuous channels with memory (e.g., channels

with phase noise) by means of message-passing techniques.

We propose message-passing algorithms to compute Cramér-Rao-type bounds. Cramér-Rao-type bounds are lower bounds on the minimum mean square estimation error; the bounds may be used to assess the performance of practical (message-passing) estimation algorithms, in particular, our phase-estimation algorithms (Question 3.b). The algorithms we propose for computing Cramér-Rao-type bounds open the door to exciting applications of information geometry, such as (1) natural-gradient-based algorithms; (2) the computation of Fisher kernels.

Keywords: graphical models, summary-propagation, belief propagation, message passing, expectation maximization, EM, steepest descent, particle filter, MCMC, particle methods, Gibbs sampling, importance sampling, decision-based estimation, iterative conditional modes, ICM, carrier phase estimation, phase noise, clock jitter, synchronization, Blahut-Arimoto algorithm, information rate, channel capacity, Cramér-Rao bound, information matrix, kernel methods, Fisher kernel, product kernel, probabilistic kernel, neural networks, back-propagation algorithm, analog electrical circuits, linear feedback shift register, LFSR.

Kurzfassung

Diese Dissertation beschreibt einerseits ein spezifisches Problem und andererseits allgemeine Methoden. Das spezifische Problem ist Trägerphasensynchronisation, welches im Kontext der digitalen Kommunikation auftritt. Die allgemeinen Methoden sind sogenannte “message-passing” Algorithmen, welche auf graphischen Modellen angewandt werden, insbesondere auf Faktorgraphen. Wir betrachten Anwendungen solcher Algorithmen im Kontext von statistischer Inferenz (wie z.B. in der digitalen Kommunikation, in der Signalverarbeitung oder im maschinellen Lernen), Statistik, Informationstheorie und der Theorie dynamischer Systeme (wie z.B. analoge elektrische Schaltungen).

Die primäre Motivation für diese Arbeit war (1) den Leistungsverlust digitaler Kommunikationssysteme infolge nicht-idealen Oszillatoren zu untersuchen; (2) die Entwicklung von Algorithmen, welche diesen Leistungsverlust minimieren.

Oszillatoren sind allgegenwärtig in digitalen Kommunikationssystemen; Signale praktischer Oszillatoren sind verrauscht, d.h., die Oszillator-signale sind nicht exakt periodisch, was oft zu einem bedeutenden Leistungsverlust der Kommunikationssysteme führt.

In den frühen Tagen der Kommunikation war diese Art von Leistungsverlust eher von sekundärer Bedeutung. Die Kommunikationssysteme operierten zu diesen Zeiten weit entfernt von der theoretischen Leistungsgrenze, d.h., die Kanalkapazität. Die Hauptsache war deswegen Fehlerkorrigierende Methoden zu entwickeln, welche die Lücke zwischen der Leistung praktischer Kommunikationssysteme einerseits und der Kanalkapazität andererseits schliessen könnten.

Mit der jüngsten Entwicklung iterativer Dekodierungsmethoden arbeiten die heutigen Kommunikationssysteme nahe an der theoretischen Leistungsgrenze; Probleme wie Synchronisation, welche zuvor eher von sekundärer Bedeutung waren, sind heute die wichtigsten (letzten) Engpässe beim Entwurf von Kommunikationssystemen geworden.

In dieser Dissertation konzentrieren wir uns auf Trägerphasensynchronisation. Darunter versteht man das Synchronisieren der Phase des lokalen Empfängersoszillators mit der Phase des empfangenen Signals.

Die Fragen welche wir betrachten sind:

- a) Welche physikalische Mechanismen sind verantwortlich für Phaserauschen? Wie kann Phaserauschen modelliert werden?
- b) Wie können Trägerphasenschätzungsalgorithmen systematisch hergeleitet werden?
- c) Was sind die theoretischen Leistungsgrenzen für Kommunikation über Kanäle mit Phaserauschen, insbesondere:
 - i) Wieviel nehmen die Informationsraten infolge Phaserauschen ab?
 - ii) Wie gut kann die (verrauschte) Trägerphase geschätzt werden?

Im Gegensatz zu früheren und parallelen Arbeiten ist unser Ziel nicht der Entwurf und die Optimierung von vollfunktionierenden Kommunikationssystemen. In dieser Arbeit werden verschiedene Werkzeuge entwickelt, welche zu Antworten auf die obenstehenden Fragen (und vielen anderen verwandten Fragen) führen (oder führen könnten).

Wir beschreiben eine detaillierte Analyse von Phaserauschen in freilaufenden Oszillatoren und PLLs (Frage 1). Wir schlagen ein einfaches intuitives Modell für Phaserauschen in freilaufenden Oszillatoren vor. Wir beschreiben zwei einfache Modelle für Bandpass-Kommunikationskanäle. Die Modelle berücksichtigen Offsets zwischen der Phase des Empfängersoszillators und der Phase des empfangenen Signals; die Modelle vernachlässigen aber Timing-Offsets. Im ersten Modell ist der Phase-Offset konstant, im zweiten Modell wird der Phase-Offset modelliert als

ein Random-Walk Prozess. Wir untersuchen unter welchen Bedingungen beide Modelle gültig sind. Die meisten Methoden dieser Dissertation werden anhand beiden Modellen illustriert.

Viele (wenn nicht alle) Methoden dieser Arbeit sind basieren auf graphischen Modellen, insbesondere Faktorgraphen. Faktorgraphen werden u.a. verwendet um die Struktur des betrachteten Systems zu visualisieren. Sie stellen die Faktorisierung multivariater Funktionen dar. Jede Kante des Graphens entspricht einer Variable, jeder Knoten des Graphens entspricht einem Faktor. Faktorgraphen können beliebige Funktionen darstellen, insbesondere probabilistische Modelle, fehlerkorrigierende Codes, Blockdiagramme und andere Modelle welche oft in der Kommunikation, in der Signalverarbeitung und in anderen Gebieten verwendet werden.

Wir zeigen wie Faktorgraphen als ein Werkzeug verwendet werden können, um praktische Schätz- und Detektionsalgorithmen zu entwickeln. Unsere Techniken können auf modellbasierte Signalverarbeitung (z.B. Phaseschätzung) und auf maschinelles Lernen angewandt werden. Wir formulieren verschiedene Standard-algorithmen der Signalverarbeitung und des maschinellen Lernens als message passing auf Faktorgraphen, z.B. Partikelmethode, Gradientenverfahren, entscheidungsbasierte Methoden, und expectation maximization (EM). In diesen Algorithmen werden lokale Regeln an den Knoten des Faktorgraphen angewandt. Mit anderen Worten, die globale Schätzaufgabe (oder Detektionsaufgabe) wird anhand einer teile-und-herrsche Strategie angepackt: die globale Berechnung wird durch viele (einfache) lokale Berechnungen ersetzt. Die lokale Regeln können als Bausteine für neue Schätz- und Detektionsalgorithmen verwendet werden. Durch dem Auflisten von den möglichen Aufdatierungsregeln an jedem Knoten des Faktorgraphen kann man systematisch neue Algorithmen entwickeln. Wir demonstrieren diese Idee für das konstante-Phasemodell und das Random-Walk Phasemodell (Frage 2). Wir zeigen auch wie der back-propagation Algorithmus für das Trainieren von feed-forward neuronalen Netzwerken als das Anwenden von generischen Aufdatierungsregeln auf einem geeigneten Faktorgraphen aufgefasst werden kann. Wir beschreiben Kernels im Kontext von message passing auf Faktorgraphen.

Wir zeigen wie message-passing Algorithmen für Inferenz als dynamische Systeme implementiert werden können, insbesondere als Uhr-freie analoge elektrische Schaltungen. Diese Systeme arbeiten zeitkontinuierlich und brauchen deswegen keine digitale Uhr; deshalb vermeiden sie das

Problem von Zeitsynchronisation.

Wir schlagen eine numerische Methode vor, um Informationsraten von kontinuierlichen Kanälen zu berechnen (Frage 3.a.). Der Algorithmus ist eine Erweiterung einer Methode welche zuvor für diskrete Kanäle vorgeschlagen wurde. Auch hier sind Faktorgraphen zusammen mit dem summary-propagation Algorithmus ein wichtiger Bestandteil. Wir wenden die Methode auf das random-walk Phasemodell an.

Eine numerische Methode für die Berechnung von Kapazitäten (oder den unteren Grenzen für die Kapazität) von kontinuierlichen Kanälen ohne Gedächtnis (Frage 3.a) wird vorgeschlagen. Wir bieten numerische Ergebnisse für den Gauss'schen Kanal mit mittleren- und maximalen Leistungsbedingungen an. Wir skizzieren wie der Algorithmus mit Hilfe von message-passing Methoden auf Kanäle mit Gedächtnis (z.B. Kanäle mit Phaserauschen) erweitert werden kann.

Wir schlagen message-passing Algorithmen für die Berechnung von Cramér-Rao-Typ Grenzen vor. Cramér-Rao-Typ Grenzen sind untere Grenzen für den minimalen mittleren quadratischen Schätzfehler; diese Grenzen können verwendet werden um praktische (message-passing) Schätzalgorithmen zu bewerten, insbesondere unsere Phaseschätzer (Frage 3.b). Die Algorithmen für die Berechnung von Cramér-Rao-Typ Grenzen welche wir vorschlagen könnten zu neuen interessanten Anwendungen der Informationsgeometrie führen, wie z.B. (1) natürlicher-gradientenbasierte Algorithmen; (2) die Berechnung von Fisher-Kernels.

Stichworte: Graphische Modelle, summary-propagation, belief propagation, message passing, expectation maximization, EM, steepest descent, Partikelfilter, MCMC, Partikelmethode, Gibbs sampling, importance sampling, entscheidungsbasiertes Schätzen, iterative conditional modes, ICM, Trägerphasenschätzung, Phaserauschen, clock jitter, Synchronisation, Blahut-Arimoto Algorithmus, Informationsrate, Kanalkapazität, Cramér-Rao Grenze, Informationsmatrize, Kernel-Methoden, Fisher-Kernel, Produkt-Kernel, probabilistischer Kernel, neuronale Netzwerke, back-propagation Algorithmus, analoge elektrische Schaltungen, LFSR.