UNIVERSITÄT DER BUNDESWEHR MÜNCHEN

Blind Estimation of Channel Parameters in CPM Bursts

Andreas Lang

Vollständiger Abdruck der von der Fakultät für Elektrotechnik und Informationstechnik der Universität der Bundeswehr München zur Erlangung des akademischen Grades eines

Doktor-Ingenieurs

genehmigten Dissertation.

Promotionsausschuss

Vorsitzender:	Professor Dr.-Ing. Andreas Knopp
1. Prüfer:	Professor Dr.-Ing. Berthold Lankl
2. Prüfer:	Professor Dr.-Ing. Robert Fischer

Die Dissertation wurde am 08.02.2023 bei der Universität der Bundeswehr München eingereicht und durch die Fakultät für Elektrotechnik und Informationstechnik am 06.06.2023 angenommen. Die mündliche Prüfung fand am 20.07.2023 statt.

Dissertation

Blind Estimation of Channel Parameters in CPM Bursts

Andreas Lang

der Bundeswehr
Universität München

Faculty of Electrical Engineering and Information Technology
Institute of Information Technology
Chair of Communication Systems

Supervisor and first reviewer	Professor Dr.-Ing. Berthold Lankl
Second reviewer	Professor Dr.-Ing. Robert Fischer

July 2023

Abstract

Short burst transmission is of practical relevance in e.g. low power sensor or tactical networks that deploy frequency hopping. The channel conditions can be seen as mutually uncorrelated for each burst due to their spectral and or temporal separation. Because of this time variant nature, a recurring acquisition of the impairment parameters is required for each burst. This thesis proposes a blind joint estimation of several channel parameters in a flat fading environment for continuous phase modulation bursts that is realized by the expectation maximization algorithm. The main contributions are first the determining of the true estimation performance bounds of the considered transmission case, second the formulation of the expectation and maximization steps to enable the joint computation of the maximum likelihood parameter estimates and third the analysis of the parameters' scalar likelihood functions to obtain an optimized initialization grid for the algorithm. It is shown, that the joint estimator produces unbiased estimates in the relevant receive power regions and its performance in terms of the mean squared estimation error achieves the theoretical limits and slightly outperforms a state of the art pilot based estimator. Furthermore, the effective throughput is discussed and bit and frame error rates are compared to each other and to the perfectly synchronized receiver. The proposed method provides a superior performance in these metrics because of the inherently higher spectral efficiency than the pilot based contender. Its computational complexity is quantitatively analyzed and efficient computation steps and further approaches to decrease it are outlined.

This Page Intentionally Left Blank

Kurzfassung

Die Übertragung von kurzen Bursts ist von praktischer Bedeutung in z. B. leistungsbegrenzten Sensor- oder taktischen Netzen, die Frequency Hopping einsetzen. Die Kanalbedingungen können aufgrund ihrer spektralen und/oder zeitlichen Trennung für jeden Burst als unkorreliert angesehen werden. Aufgrund dieser zeitlichen Variabilität ist eine wiederkehrende Erfassung der Kanalparameter für jeden Burst erforderlich. In dieser Arbeit wird eine blinde, gemeinsame Schätzung mehrerer Kanalparameter in einer flachen Fading-Umgebung für Continuous Phase Modulation Bursts vorgeschlagen, die durch den Expectation Maximization-Algorithmus realisiert wird. Die Hauptbeiträge sind erstens die Bestimmung der wahren Grenzen der Schätzgenauigkeit für den betrachteten Übertragungsfall, zweitens die Formulierung der Expecation- und Maximization-Schritte für die gemeinsame Berechnung der Maximum-Likelihood-Parameterschätzungen, und drittens die Analyse der skalaren Likelihood-Funktionen der Parameter, um ein optimiertes Initialisierungsgitter für den Algorithmus zu erhalten. Es wird gezeigt, dass der gemeinsame Schätzer erwartungstreue Schätzungen in den relevanten Empfangsleistungsbereichen erzeugt und seine Leistung in Bezug auf den Schätzfehler die theoretischen Grenzen erreicht und dabei einen pilotbasierenden Schätzer leicht übertrifft. Darüber hinaus wird der effektive Durchsatz diskutiert und die Bit- und Rahmenfehlerraten werden miteinander und mit dem perfekt synchronisierten Empfänger verglichen. Die vorgeschlagene Methode bietet eine überlegene Leistung in diesen Metriken aufgrund der inhärent höheren spektralen Effizienz als der pilotbasierte Konkurrent. Der Rechenaufwand wird quantitativ analysiert und es werden effiziente Berechnungsschritte und weitere Ansätze zur Verringerung der Komplexität aufgezeigt.

Acknowledgments

The last six years have been the most important and the most joyful years of my life because of such important milestones in my life: My first own apartment, moving in together with my now-wife Christina, our wedding and most importantly the births of our two beloved children Paulina and Benedikt.

During all these events I had the luck to work at the Bundeswehr University Munich where I found the perfect conditions for my research of which this thesis is the final product. For this, I want to express my gratitude first and foremost to my thesis advisor Prof. Berthold Lankl. His dedication to supervising his PhD students was remarkable and I will always fondly remember the countless hours of technical and private discussions I had the pleasure to have with him. I want to thank Prof. Robert Fischer for his commitment on serving in my thesis committee. After Prof. Lankl entered his well deserved official retirement, Prof. Andreas Knopp made it possible for me to stay employed at the Uni for which I am very grateful. Thanks to Flo for being a great office mate for many years, you really made coming to work fun. Thanks to Tim for the countless talks we had, especially when we three sat in our office to enjoy a fine drink. I also want to thank Makram and Flo Völk for numerous tips you had for me. Of course a big Thank You goes to all the other people at the institute for making my time there so enjoyable, especially Constanze, Wolfgang and Alejandro (Go Team Kicker!).

My wife and two kids are the pillars in my life and I never want to stop thanking them for being there, supporting me in every way and always bringing a smile to my face. My parents made it possible for me to pursue my studies without ever worrying about finances, thank you both. My father always was and still is my Goto person

for every technical discussion and I don't ever want to miss that. I'm happy to still have my grandmother around me, who always believed in me. I also want to thank my in-laws for supporting our family in many ways.

My oldest friend and best man Mäx ignited the spark of pursuing a PhD by doing so himself. Thinking of all the hard times caused by that decision; still thank you for that... I guess ;) Last, but not least, I'm grateful for all my friends that helped me to balance my life with the right amount of nonsense, cards and beer.

Andreas Lang

Gauting, July 2023

Contents

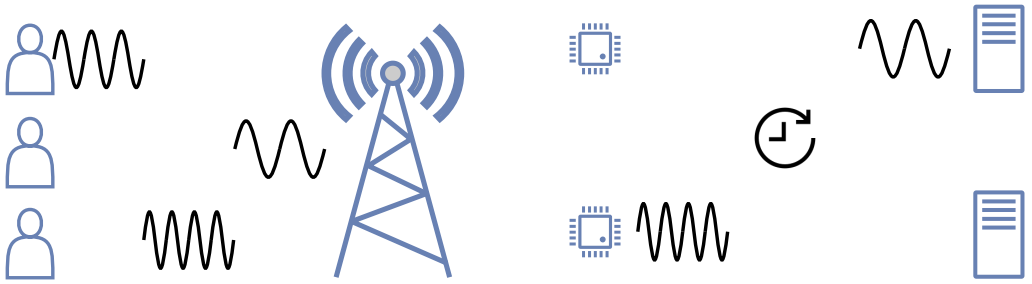
Chapter 1

Introduction

This chapter starts with a brief introduction of two short burst transmissions applications. First, tactical frequency hopping (FH) systems and second, sensor networks of which both widely use continuous phase modulation (CPM), for which important aspects are discussed. The main topic of this work, namely the estimation of channel parameters in short burst CPM transmissions, is then put in focus with a comprehensive literature overview of related works. After the thesis organization is outlined, its contributions are listed before concluding this chapter with common notations used in this work.

1.1 Motivation and Background

The first application of short burst transmission considered are tactical networks, that typically use the very high frequency (VHF) and ultra high frequency (UHF) bandwidths between 30 MHz to 400 MHz. They deploy FH mainly to avoid intentional interference [1] by changing the carrier frequency in a (pseudo) random and discontinuous fashion [2]. For this to succeed, the hop time must not exceed a certain time depending on the jamming device's distance and technology [1]. While a hop time of 1 ms was considered as plenty of protection against jamming devices [3], [4] a few more years back, it is considered as the upper limit of protective hopping times in more recent literature [1].



(a) Frequency hopping network with (potentially more than one) user transmitting data bursts on changing carrier frequencies.

(b) Sensor network sporadically transmitting data bursts on potentially different frequencies.

Figure 1-1: Two burst transmission scenarios that make use of changing carrier frequencies as an electronic protection measure (EPM).

While the data transmission can very well be of a continuous character, the frequency switch constitutes a technical discontinuation and thus each hop is considered as an individual burst. In this case, a carrier frequency switch causes a change in the attenuation of the signal due to the in general frequency selective environment. Furthermore, the transmitter and receiver oscillators' mismatch during one burst is more or less independent from the previous one. If a time division multiple access (TDMA) scheme is used in the network and the transmission pauses for some time, also the start of the next burst is unknown (in a probably well defined range). These conditions make a recurring acquisition of the channel parameters fading factor (FF), carrier frequency offset (CFO), carrier phase offset (CPO) and timing offset (TO) necessary.

In the second considered use case, the data transmission is of burst character not because of the discontinuation in form of a frequency hop, but due to the limited amount of data to be transmitted, e.g. the collecting of sparse sensor data at a server. FH can be used in these networks to gain diversity against bad channel conditions, as it is done in the case of Bluetooth [5]. The pause between two transmission bursts is considered to be long enough, to make the channel parameters mentioned above uncorrelated, which again demands a recurring acquisition of those. Fig. 1-1 visualizes both scenarios.

Both in tactical [6, 7] and sensor [8–10] networks, the use of CPM as waveform is

popular because of its favorable properties in terms of power and bandwidth efficiency. Especially in the context of ultra low power sensors with a planned lifetime of several years, power efficiency is crucial. Moreover, CPM in the form of Gaussian minimum shift keying (GMSK) is deployed in second generation mobile networks [11] and in Bluetooth [5]. Further deployment areas where CPM is considered because of its favorable properties are e.g. millimeter wave [12] and deep space communications [13].

Due to its continuous phase, high side lobes in the signal's spectrum are avoided and hence the average power spectrum is compact. The spectral efficiency can principally be arbitrarily increased by selecting appropriate waveform parameters which generally increase the computational complexity of the optimal detector though. CPM produces a constant amplitude signal, which enables the use of cost and power efficient amplifying hardware near saturation in transmitter and receiver. While the information conveyed by the signal is robust with regard to the non-linear distortions caused by this, the spectral side lobes remain largely suppressed since the phase stays continuous and thus CPM's bandwidth efficiency is retained [14]. The advantages of being able to deploy amplifiers near saturation and thereby saving backoff is e.g. investigated in [15,16] and quantified as several decibels. By ensuring soft phase transitions, dependencies between adjacent symbol intervals arise that cause an inherent trellis coding [14].

For this work, several different CPM waveform configurations are considered and an exemplary baud rate of 42 kBaud is chosen, which is typical for tactical networks [7]. The length of one burst is set to $T_{\mathrm{Burst}} = 1\,\mathrm{ms}$ to fulfill the minimal requirements stated in [1] regarding robustness against intentional interference. In Chapter 6, it is briefly discussed that shorter hop times do not significantly change this thesis' results. The four channel parameters FF, CFO, CPO and TO have to be acquired in every burst and no knowledge about them can be drawn from past bursts. Figure 1-2 shows the burst structure schematically including the data symbol in the middle of the burst and potential pilot and termination sequences in the beginning and end of it.

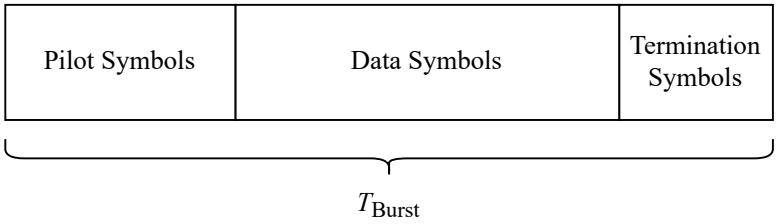


Figure 1-2: General burst structure consisting of data symbols and potential pilot and termination sequences.

1.2 Related Work in Synchronization

This section gives an overview about channel parameter estimation methods structured in three categories: non-data aided (NDA), decision directed (DD) and data aided (DA) estimation, whereas the first two are linked to the term blind estimation, because no known pilot sequence is need to obtain the estimates in contrast to the DA approaches. It is noted, that hybrid techniques exist, usually in the case where a pilot sequence is utilized to obtain an acquisition on the parameters and some kind of feedback loop to track and refine the results. The literature selection concentrates on contributions that lie in the vicinity of channel parameter estimation in CPM burst type transmission. For a comprehensive summary of different estimation methods in CPM, it is referred to [17].

1.2.1 Non-Data Aided

NDA techniques usually utilize the signal statistics and thus are problematic for short bursts, since the stochastic sample size is usually too small. Several NDA estimation techniques for the above channel parameters for CPM transmissions can be found in the literature, though many methods are limited to minimum shift keying (MSK)-type transmissions [18–20] and thus are not applicable in general CPM configurations. There are also some methods that work with arbitrary CPM parameters [21–25], though do not provide very accurate estimates in terms of the mean square estimation error (MSEE) as shown in [26–28]. Furthermore, these methods often only tackle the estimation of a single parameter and not the joint estimation as required in the use

cases described above. This also means in consequence, that the techniques are not robust to deviations of the non-considered channel parameters and their performance deteriorates further under the presence of those.

1.2.2 Decision Directed

Similar to the NDA methods, no pilot sequences are utilized in the parameter estimation task. Instead, initial estimates are improved by taking decision about data symbols into account. Though typical phase locked loop (PLL) methods [29] and Kalman filter approaches [30–33] may use pilots to obtain their initial estimates, which constitutes a degradation of the system's spectral efficiency (SE). Furthermore, after the parameter acquisition, they assume a continuous transmission in order to be able to track the errors over the course of a much larger frame than the considered burst represents. For that reason they are not applicable in burst type transmission. Per survivor processing technique are able to operate without the help pf pilot symbols when only a CPO is estimated [34], but rely on a pilot-based parameter acquisition when the channel becomes more complex [35,36]. In both cases, they consider frames much longer than in this work. Optimal, blind DD estimators were presented by the author in [26–28] based on the expectation maximization (EM) algorithm for CPO, CFO and joint CPO TO estimation. In [37], the single TO estimation was also tackled by means of the EM algorithm.

1.2.3 Data Aided

The most cited work on DA estimation of channel parameters explicitly for burst type CPM transmission is presented in [38]. It uses a specific pilot sequence to obtain optimal estimates CFO, CPO and TO estimates for arbitrary CPM parameter configurations in a feedforward fashion while using a computationally efficient algorithm. Though there are some drawbacks to this method that are inherent to DA techniques. First, the spectral efficiency is degraded by the pilot sequence. Second, (known) pilot sequences can easily be detected and are therefore prone to intentional interference.

Jamming specifically only the pilot sequence for synchronization is a very efficient way to impede the complete transmission. This point is especially relevant in tactical networks, where a focus on robustness in the system design is probably even more present than elsewhere. Another pilot based scalar TO estimation method is presented in [39].

1.3 Thesis Organization

Chapter 2 starts with defining the general CPM signal and introducing the equivalent lowpass representation of burst-type CPM. The CPM trellis is defined and the tilting process is addressed which transforms the possibly periodic trellis into a time invariant one. The characteristics of typical VHF and UHF channels are outlined and based on the coherence times and coherence bandwidths, a static (frequency) flat (SF) channel model is introduced that is comprised by a flat fading component and carrier and timing impairments. An optimal receiver is discussed in detail for this channel model before giving brief overviews over the CPM trellis initialization and the deployment of a channel decoder.

In Chapter 3, principal aspects of estimation theory are covered. The estimation problem at hand is contextually placed in a general framework and classified according to the attributes randomness, dimension and biasedness. A general theoretic performance limit is presented for the estimator, before computing well known explicit bounds, that are generally not as tight as the absolute limit. It was then derived, that for the specific case of the considered channel parameter estimation in burst-type CPM, the computable bounds equal the absolute performance limit and are hence perfectly suited for the practical estimator performance.

The proposed, blind estimator is subsequently discussed in great detail in Chapter 4. The underlying machine learning technique EM is given an overview before putting the considered problem in the general iterative EM framework and mathematically deriving the respective components. Great attention is given to the starting point problem of EM. To ensure the convergence to the global maximum of the four-

dimensional objective function, the scalar log likelihood function (LLF) of the channel parameters are investigated and the principal distance of two adjacent maxima is determined analytically. An algorithmic description of the proposed method is given in pseudo code, whereas for the most demanding computational tasks efficient solutions are suggested.

In Chapter 5, the EM convergence behavior of the estimator is investigated and a suited maximum number of iterations is found. Moreover, the effect of using additional information from a channel decoder is examined. The proposed, blind estimator is then compared to an optimal pilot based method in terms of mean estimation value (MEV), MSEE, effective throughput as well as coded and uncoded error rates in the SF channel and in fast fading, frequency-selective channels. Additionally the proposed estimation method is tested in a cable bound hardware setup. The computational complexity of the estimator is analyzed quantitatively and set in context.

Chapter 6 restates the problem of channel parameter estimation in burst-type CPM transmissions and summarizes the contributions of this thesis. An outlook of worthwhile future research directions is given.

1.4 Contributions

- Formulation of the CPM signal's optimal detection in the deduced SF channel model on base of the Bahl-Cock-Jelinek-Raviv (BCJR) algorithm.

- Computation of the modified Cramér-Rao vector bound (MCRVB)s for the channel parameter estimation in burst-type CPM transmission.

- Derivation of the relations that show the equality of each parameter's MCRVB and the true Cramér-Rao vector bound (CRVB) for the case of known data sequences.

- Complete Derivation of the considered EM auxiliary function and its transformation to enable clear maximization steps.

- Derivation of each parameter's maximization step and demonstration of the practical, mutual uncoupled nature of the parameter's optimizations.
- Analytical investigation of CFO, CPO and TO's likelihood functions and determination of the principal distance of two adjacent maxima in it.
- Design of an optimized grid of EM starting points and presentation of reduction approaches.
- Algorithmic description of the proposed channel parameter estimator and suggestions for efficient implementation of the most computationally complex parts.
- Analysis of the proposed algorithm's convergence behavior and on the effect of the incorporation of channel decoder information.
- Evaluation of the estimator and comparison to a pilot based alternative in terms of MEV, MSEE, effective throughput as well as coded and uncoded error rates in the SF channel and in fast fading, frequency-selective environments.
- Quantitative analysis of the computational complexity of the proposed algorithm's components.

This thesis differentiates itself from available literature foremost in the focus on short burst transmission and in the puristic approach in contrast to hybrid techniques, e.g. for acquisition and tracking, or parameter depending. This becomes especially clear in tackling the starting point problem in an analytical way instead of diluting this problem by using pilot symbols. It is noted, that the EM framework for estimating channel parameters in the context of synchronization of communication systems was first published in [40, 41]. There, only linear modulations were considered and the expectation step was achieved by computing the posterior probabilities through utilizing a Turbo code [42]

The following publications resulted (directly and indirectly) from the work on this thesis.

- Lang, Lankl, 2022: "Blind Vector Parameter Estimation for Burst Type CPM Transmissions in VHF and UHF Channels", *IEEE Open Journal of Vehicular Technology*, 2022, vol. 4, pp. 162-180
- Lang, Lankl, 2019: "Carrier Phase Synchronization in Burst-Type CPM by Means of Expectation Maximization", *2019 IEEE 90th Vehicular Technology Conference (VTC2019-Fall)*, IEEE, 2019. pp. 1-7
- Lang, Lankl, 2020: "Carrier Frequency Offset Estimation in Burst-Type CPM via the EM Algorithm", *2020 IEEE 91st Vehicular Technology Conference (VTC2020-Spring)*, IEEE, 2020. pp. 1-6
- Lang, Lankl, 2021: "Blind Joint Timing and Carrier Phase Estimation in Burst-Type CPM", *2021 IEEE 94th Vehicular Technology Conference (VTC2021-Fall)*, IEEE, 2021. pp. 1-6
- Lang, Lankl, 2018: "Very Short Channel Codes for Burst-Type Transmissions", *MILCOM 2018-2018 IEEE Military Communications Conference (MILCOM)*, IEEE, 2018. pp. 1-6
- Lang, Lankl, 2019: "Tail-Biting in Optimal CPM Detectors ", *IET Electronics Letters*, 2020, vol. 56, No. 3, pp. 141-144
- Lang, Lankl, 2022: "A Comprehensive Study of CPM Trellis Initialization Methods", *MILCOM 2022-2022 IEEE Military Communications Conference (MILCOM)*, IEEE, 2018. pp. 78-83

Parts of Chapters 4 and 5 were already published in the first item of the list, whereas the second to fourth show the research progress to that point. The last three papers cover results on channel codes and (de)-modulation for burst type CPM transmissions that came along the way.

1.5 Notation Conventions

In this section, notations used throughout this work are introduced:

- The imaginary unit is denoted as $j = \sqrt{-1}$. The complex conjugate of a complex number is written as $(a + jb)^* = (a - jb)$.
- The operators $\mathrm{Re}\{\bullet\}$ and $\mathrm{Im}\{\bullet\}$ return the real and imaginary value of a complex argument.
- The modulo operator is denoted as $\bullet \mathrm{mod}\ \bullet$, ceiling, floor and rounding operators as $\lceil\bullet\rceil$, $\lfloor\bullet\rfloor$ and $\lfloor\bullet\rceil$.
- If the sampling theorem is met by a sampling rate of T_0^{-1}, a signal can be represented by the well known interpolation formula [11]

$$x(t) = \sum_{n=-\infty}^{+\infty} x(nT_0) \cdot \mathrm{si}\left(\pi(t/T_0 - n)\right) \ , n \in \mathbb{Z} \tag{1.1}$$

  in which $x(nT_0)$ are the sampling values and $\mathrm{si}(t) = \sin(t)/t$. The energy of the signal is expressed as

$$\begin{aligned} E_x &= \int_{-\infty}^{+\infty} |x(t)|^2 \,\mathrm{d}t && (1.2)\\ &= \int_{-\infty}^{+\infty} \left| \sum_{n=-\infty}^{+\infty} x(nT_0) \cdot \mathrm{si}\left(\pi(t/T_0 - n)\right) \right|^2 \mathrm{d}t && (1.3)\\ &= \sum_{n=-\infty}^{+\infty} |x(nT_0)|^2 \int_{-\infty}^{+\infty} \left|\mathrm{si}\left(\pi(t/T_0 - n)\right)\right|^2 \mathrm{d}t && (1.4)\\ &= T_0 \sum_{n=-\infty}^{+\infty} |x(nT_0)|^2 \ , && (1.5) \end{aligned}$$

  which gives the important relation between energy, sampling interval and sampled values. The sum can be extracted due to the orthogonality of the si functions and the integral can be resolved by Parseval's theorem for example.

- The symmetric rectangular function of width x is denoted as

$$\mathrm{rect}\left(\frac{t}{x}\right) = \begin{cases} 1 & |t| < x/2 \ , \\ 1/2 & |t| = x/2 \ , \\ 0 & |t| > x/2 \ . \end{cases} \tag{1.6}$$

The edges with values 0.5 are given for the sake of mathematical correctness, but need no further consideration in the scope of this work.

- The Dirac delta function is denoted as $\delta(t)$.

- Vectors can be upper and lower case and are written as underlined letters as $\underline{x}$. Their i'th element is referred to as $\underline{x}_i$, whereas a subset of a vector is denoted with its lower (i) and upper (k) index limits as $\underline{x}[i:k]$. Matrices are upper case letters and are denoted by an underset tilde $\underset{\sim}{X}$ with their k'th row and l'th column element are specified by $[\underset{\sim}{X}]_{k,l}$.

- For a probability density function (PDF) and a probability mass function (PMF) the lower case p and upper case P are used, respectively. The stochastic processes are denoted as subscript to the probability symbols p and P. The probabilities' argument is the realization x of the corresponding stochastic process $\boldsymbol{x}$, i.e. $\mathrm{p}_{\boldsymbol{x}}[\boldsymbol{x} = x] \equiv \mathrm{p}_{\boldsymbol{x}}[x]$ and $\mathrm{P}_{\boldsymbol{x}}[\boldsymbol{x} = x] \equiv \mathrm{P}_{\boldsymbol{x}}[x]$. This notation should usually suffice in terms of readability, as p and P are exclusively used for probabilities. A sample function of a stochastic process is written as the process's symbol without the underline. If the PMF's symbol P is underlined, $\underline{\mathrm{P}}_{\boldsymbol{x}}[x]$ shall denote a vector of probabilities for every possible outcome of its argument. The (upright) symbol P should not be able to confuse with the (italic) variable of the modulation index's nominator P, as they are never used in the same context.

- The operators E[•] and Var[•] compute the expected value and the variance, respectively of their arguments.

$$\mathrm{E}[\boldsymbol{x}] = \int_{-\infty}^{+\infty} x \cdot \mathrm{p}_{\boldsymbol{x}}[x]\,\mathrm{d}x \tag{1.7}$$

$$\mathrm{Var}[\boldsymbol{x}] = \int_{-\infty}^{+\infty} (x - \mathrm{E}[\boldsymbol{x}])^2 \cdot \mathrm{p}_{\boldsymbol{x}}[x]\,\mathrm{d}x \ . \tag{1.8}$$

The definition for discrete random variables is analog with sums and PMFs instead of integrals and PDFs.

- The cross correlation function (CCF) of two weakly stationary processes $\boldsymbol{x}$ and

$\boldsymbol{y}$ is defined as

$$R_{\boldsymbol{x},\boldsymbol{y}}(t') = \mathrm{E}\left[x^*(t) \cdot y(t+t')\right] \ , \tag{1.9}$$

whereas the CCF's computation of two sample functions $x(t)$ and $y(t)$ can be calculated as

$$R_{x,y}(t') = \int_{-\infty}^{+\infty} x^*(t) \cdot y(t+t') \, \mathrm{d}t \tag{1.10}$$

or equivalently for sampled signals and switched delay

$$R_{x,y}(t') = T_0 \sum_k x^*(nT_0) \cdot y(nT_0 + t') \ . \tag{1.11}$$

- The transformation between the bandpass and lowpass domain and reverse are defined such that the energies of deterministic and stochastic signals are sustained and that systems do not change energies between in and output [43]

$$x_{\mathrm{LP}}(t) = \frac{1}{\sqrt{2}} \cdot \left[x_{\mathrm{BP}}(t) + j\mathcal{H}\left\{x_{\mathrm{BP}}(t)\right\}\right] \cdot e^{-j2\pi f_{\mathrm{c}} t} \tag{1.12}$$

$$R_{\boldsymbol{x}_{\mathrm{LP}},\boldsymbol{x}_{\mathrm{LP}}}(\tau) = 1 \cdot \left[R_{\boldsymbol{x}_{\mathrm{BP}},\boldsymbol{x}_{\mathrm{BP}}}(\tau) + j\mathcal{H}\left\{R_{\boldsymbol{x}_{\mathrm{BP}},\boldsymbol{x}_{\mathrm{BP}}}(\tau)\right\}\right] \cdot e^{-j2\pi f_{\mathrm{c}} \tau} \tag{1.13}$$

$$h_{\mathrm{LP}}(t) = \frac{1}{2} \cdot \left[h_{\mathrm{BP}}(t) + j\mathcal{H}\left\{h_{\mathrm{BP}}(t)\right\}\right] \cdot e^{-j2\pi f_{\mathrm{c}} t} \tag{1.14}$$

with $\mathcal{H}\{\bullet\}$ denoting the Hilbert transform. The subscripts indicate the domain and f_{c} denoting the respective transformation frequency, which in this work is exclusively the carrier frequency. Derived from (1.13), the bandpass process of a white noise with constant power spectral density $N_0/2$ leads to the equivalent lowpass white noise of constant complex power spectral density N_0 in the relevant bandwidth.

- The log likelihood ratios (LLR) of binary symbols x_n is defined as $\mathrm{L}(x_n) = \log \mathrm{P}_{\boldsymbol{x}_n}(x_n = 0) - \log \mathrm{P}_{\boldsymbol{x}_n}(x_n = 1)$ with log (without subscript) denoting the natural logarithm. If the argument inside the LLR operator is a vector $\underline{x}$, then $\mathrm{L}(\underline{x})$ shall denote the vector of LLRs.

- Several distributions are used throughout this thesis, which are defined as fol-

lows. A random process following a uniform distribution $\boldsymbol{x} \sim \mathcal{U}(a, b)$ between the edge values $a, b \in \mathbb{R}$ with $b > a$ is described by the PDF

$$\mathrm{p}_{\boldsymbol{x}}(x) = \begin{cases} \frac{1}{b-a} & a \leq x \leq b \ , \\ 0 & \text{otherwise} \ . \end{cases} \tag{1.15}$$

The definition of a real normal distributed process $\boldsymbol{x} \sim \mathcal{N}(\mu, \sigma)$ with mean μ and variance σ^2 is given as

$$\mathrm{p}_{\boldsymbol{x}}(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \ . \tag{1.16}$$

The complex variant is defined as two zero mean normal processes as real and imaginary part $\boldsymbol{z} = \boldsymbol{x} + j\boldsymbol{y}$, which gives the complex normal process $\boldsymbol{z} \sim \mathcal{CN}(0, \mathrm{Var}[\boldsymbol{z}] = \mathrm{Var}[\boldsymbol{x}] + \mathrm{Var}[\boldsymbol{y}])$ whose variance is the sum of the real and imaginary parts' variances. A Rayleigh distributed process $\boldsymbol{x} \sim \mathcal{R}(\kappa)$ with the real scaling factor κ has the PDF

$$\mathrm{p}_{\boldsymbol{x}}(x) = \begin{cases} \frac{x}{\kappa^2} e^{-\frac{x^2}{2\kappa^2}} & x > 0 \ , \\ 0 & \text{otherwise} \ . \end{cases} \tag{1.17}$$

This Page Intentionally Left Blank

Chapter 2

System Overview

This paragraph opens up the chapter by giving an overview of the whole transceiver chain that is depicted in Fig. 2-1, which is in itself a graphical glossary. The transmitter structure is typical with binary information symbols $\underline{u}$ that are encoded and mapped to CPM symbols $\underline{a}$ (cf. Section 2.6), which are converted by the CPM modulator into the signal $s(t)$ (cf. Section 2.1). In Section 2.2.1 the phase tilting is explained which is modeled to either take place in the transmitter (on $s(t)$) or the in the receiver (on $r(t)$). While the tilting is in practice usually realized in the receiver, this differentiation will become relevant in Chapter 3. The modeling of the radio channel is discussed in Section 2.3 and Section 2.4 covers the receive filter and the transition from a time-continuous into a time-discrete signal by sampling.

The main part of the receiver comprises three information exchanging blocks:

- The soft-input-soft-output (SISO) CPM detector is handled in Section 2.5.
- The serially concatenated (SC) decoding setup including interleaver and mapper and their according counterparts introduced in Section 2.6. It is noted, that both the serially concatenated (SC) decoding setup and the CPM SISO detector together ware what is usually referred to as the SC CPM setup [44].
- The blind and joint estimation of the channel parameters whose derivation is the main contribution of this work and is explained in detail in Chapter 4.

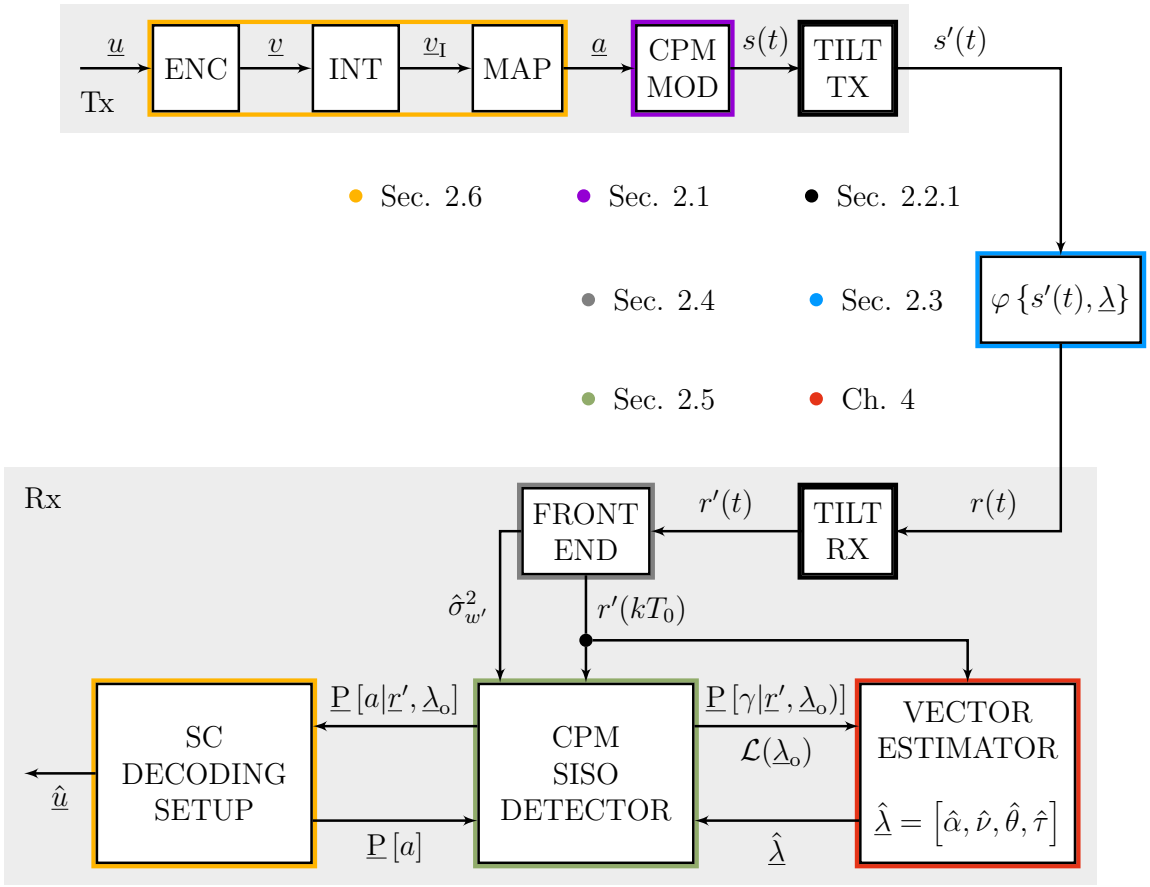


Figure 2-1: Block diagram of the complete transmission system. It furthermore serves as a graphical glossary of the individual parts. Bold letters indicate vectors and hats above letters estimates, respectively. In case of the PMFs, not only a probability vector of one element is passed (as the drawing might imply for the sake of simplicity), but for all symbols. Also for the sake of clarity, the probabilities vectors are not labeled with the corresponding processes. The trial value $\underline{\lambda}_{\mathrm{o}}$ is short for $\underline{\lambda}_{\mathrm{old}}$ and its particular meaning is discussed in Chapter 4. The signal and variable names are introduced in the denoted sections.

The iterative exchange of information by two entities is widely known as the Turbo concept. In this case, the SISO detector forms a Turbo decoder with the channel decoder as well as a Turbo synchronizer with the vector estimator. Hereby the channel code's utilization is optional for the estimation of the channel parameter. Both cases will be treated in the following chapters. This double Turbo setup will show to enable an optimal channel parameter estimation without the help of pilot symbols.

2.1 Continuous Phase Modulation

2.1.1 Bandpass Definition of Continuous CPM Signal

CPM is a format that covers a large range of phase modulations such as MSK or GMSK. A general definition of CPM can be found, among others, in [11]. For this work the CPM signal shall be defined in the style of [14]. Beginning with the formulation of the continuous (i.e. non-burst) passband signal

$$\tilde{s}_{\mathrm{BP}}(t) = \sqrt{\frac{2E_S}{T}} \cos\Big(2\pi f_c t + \tilde{\phi}(t, \underline{\tilde{a}}) + \phi_0\Big) \,, \tag{2.1}$$

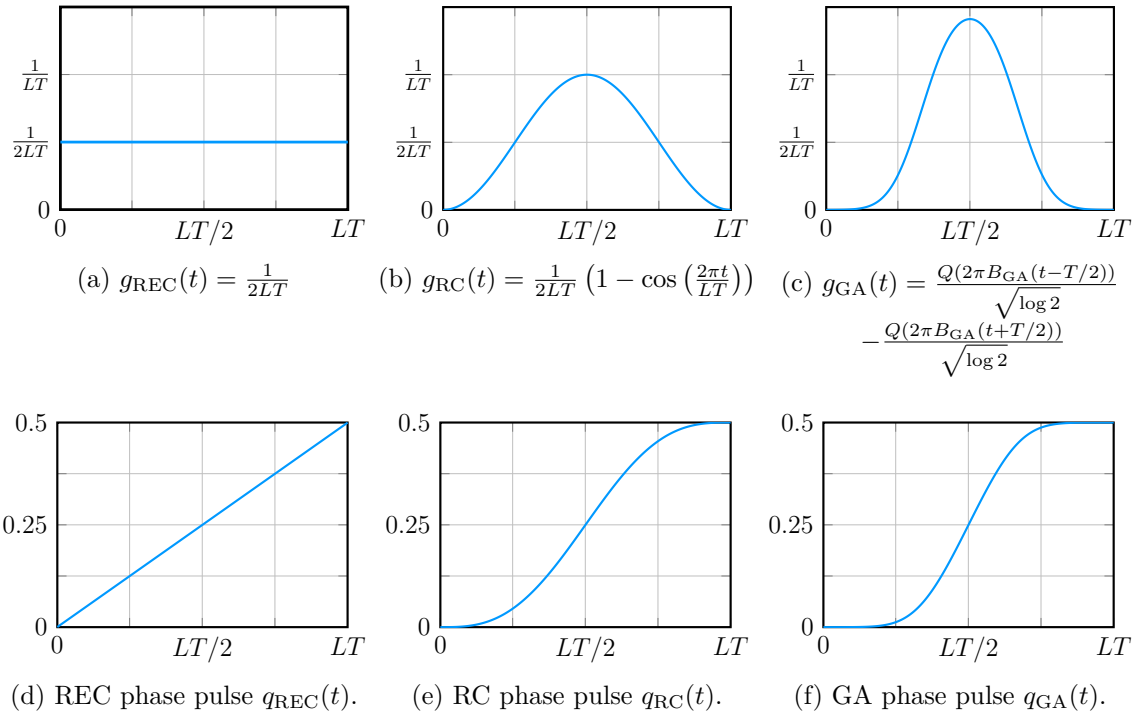
it is clear that CPM has a constant amplitude with power E_S/T with E_S and T depicting the symbol energy and duration, respectively. The carrier frequency f_c and the phase function

$$\tilde{\phi}(t, \underline{\tilde{a}}) = 2\pi h \sum_{n=-\infty}^{+\infty} a_n q(t - nT) \tag{2.2}$$

further define the signal. The choice of an initial phase offset ϕ_0 is arbitrary and set to $\phi_0 = 0$ throughout this work without loss of generality. The modulation index $h = P/Q$ scales the phase pulses $q(t)$ and is typically a fraction of two positive, mutually prime integers P and Q with $P < Q$. The elements of the data symbol vector $\underline{\tilde{a}}$ with a_n being the n-th data symbol modulate the height of $q(t)$. They are independent and identically distributed (i.i.d.) and drawn with equal probability from the symbol alphabet $\mathcal{A}$ whose cardinality $|\mathcal{A}| = M$ is equal to the modulation order M. The alphabet is constructed in the following fashion $\mathcal{A} = \{\pm 1 \pm 3, \ldots, \pm(M-1)\}$, which gives the important cases of binary ($M = 2$) and quaternary ($M = 4$) alphabets of

$$\mathcal{A} = \begin{cases} \{-1, +1\} & M = 2 \,, \\ \{-3, -1, +1, +3\} & M = 4 \,. \end{cases} \tag{2.3}$$

The time derivative of the phase function $\frac{\mathrm{d}\tilde{\phi}}{\mathrm{d}t}$ gives the instantaneous frequency deviation of the CPM signal, whereas the respective derivative of the phase pulse $g(t) = \frac{\mathrm{d}q(t)}{\mathrm{d}t}$ gives a frequency pulse and is typically of a rectangular (REC), raised



(a) $g_{\mathrm{REC}}(t) = \frac{1}{2LT}$ (b) $g_{\mathrm{RC}}(t) = \frac{1}{2LT}\left(1 - \cos\left(\frac{2\pi t}{LT}\right)\right)$ (c) $g_{\mathrm{GA}}(t) = \frac{Q(2\pi B_{\mathrm{GA}}(t-T/2))}{\sqrt{\log 2}} - \frac{Q(2\pi B_{\mathrm{GA}}(t+T/2))}{\sqrt{\log 2}}$

(d) REC phase pulse $q_{\mathrm{REC}}(t)$. (e) RC phase pulse $q_{\mathrm{RC}}(t)$. (f) GA phase pulse $q_{\mathrm{GA}}(t)$.

Figure 2-2: Frequency and phase pulses of REC, RC and GA shape. The mathematical descriptions of the frequency pulses are given in the upper row's captions. The bandwidth time product of the GA pulse in this case is $B_{\mathrm{GA}}T = 1/(L-1)$.

cosine (RC) or Gaussian (GA) form, as displayed in Fig. 2-2.

It is common for the frequency pulses to be confined to a duration of LT with L taking positive integer values and stating the pulse length in symbol durations. Hereby their total area is defined to equal 0.5. This leads to the phase pulse's definition

$$q(t) = \int_0^t g(t')\,\mathrm{d}t' = \begin{cases} 0 & t \leq 0\,, \\ \int_0^t g(t')\,\mathrm{d}t' & 0 < t \leq LT\,, \\ \frac{1}{2} & t > LT\,. \end{cases} \tag{2.4}$$

The most right part of (2.4) shall again emphasize the temporal confinement of the frequency pulse. While full response ($L = 1$) CPM has non overlapping symbols, partial response ($L > 1$) CPM introduces non-linear intersymbol interference (ISI) due to its interfering phase symbols. The pulse choice has direct influence on the

estimation of a time offset through its normalized frequency pulse energy

$$G_2(0)T = T\int_{-\infty}^{+\infty} g^2(t)\,\mathrm{d}t \ . \tag{2.5}$$

The expression $G_2(0)$ stands for the spectral direct component of the squared pulse and reflects the origin of its influence derived in Chapter 3.

GA frequency pulses are defined by the pulse bandwidth time product $B_{\mathrm{GA}}T$ [11] and have to be truncated since they are generally of infinite length. The truncation $L = \left\lceil \frac{1}{B_{\mathrm{GA}}T} + 1 \right\rceil$ is chosen in these cases such that the truncation error is insignificant and (2.4) still holds in very good approximation. As an example, for $B_{\mathrm{GA}}T = 0.5$ and therefore $L = 3$, the approximation error is about $1.122 \cdot 10^{-9}$ with respect to the normalized energy $G_2(0)T$.

2.1.2 Equivalent Lowpass Representation of the Burst-Type CPM Signal

In this section, the CPM signal is transformed into the equivalent lowpass domain and then a description of the burst-type CPM signal bis introduced. As it is notated in Section 1.5 the bandpass lowpass transformation shall preserve the signal's energy and thus the complex envelope with respect to the carrier frequency f_{c} writes as

$$\tilde{s}_{\mathrm{LP}}(t) = \sqrt{\frac{E_S}{T}} \exp\left(j\tilde{\phi}(t, \underline{\tilde{a}})\right) \ . \tag{2.6}$$

The subscript LP indicating the equivalent lowpass domain will be omitted in the rest of this work. The definitions of the phase function $\tilde{\phi}(t, \underline{\tilde{a}})$ and the phase pulse $q(t)$ in Sec. 2.1.1 remain valid in the lowpass domain without adjustments. Figure 2-3 shows the block diagram of the CPM burst modulation.

To define the bursts, windowing is necessary in the partial response case, whereas full response signals are also multiplied with a properly shifted rectangular function for the sake of the definition's consistency. Moreover the information bearing phase function is slightly adjusted. This leads to the following definition of a CPM burst

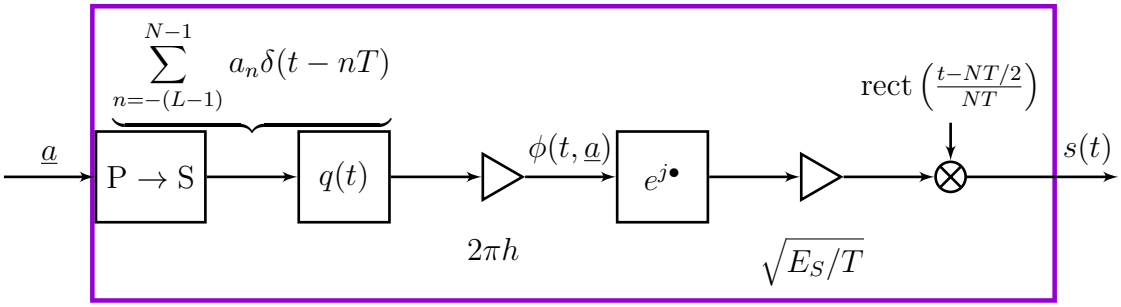


Figure 2-3: Block diagram of the CPM modulator depicted as a violet block in Fig. 2-1. The CPM symbols are transformed from a parallel vector to a serial impulse train (weighted with the values of $\underline{a}$), which is convolved with $q(t)$. The rest depicts (2.7) and (2.8).

consisting of N symbols

$$s(t) = \begin{cases} \sqrt{\frac{E_S}{T}} \exp\left(j\phi(t,\underline{a})\right) & 0 \le t \le NT \ , \\ 0 & \text{otherwise} \end{cases} \tag{2.7}$$

with the adjusted phase function

$$\phi(t,\underline{a}) = 2\pi h \sum_{n=-(L-1)}^{N-1} a_n q(t-nT) \ . \tag{2.8}$$

Hereby, data symbols with negative indices are only used for initialization of the CPM trellis and not transmitted. The phase at the time origin is therefore defined as

$$\phi(t=0, a_{-(L-1)} \dots a_{-1}) = 2\pi h \sum_{n=-(L-1)}^{-1} a_n q(-nT) \ . \tag{2.9}$$

Each respective symbol time slot $nT \le t < (n+1)T$ in (2.8) is described by

$$\phi_n(t,\underline{a}) = 2\pi h \sum_{l=0}^{L-1} a_{n-l} q(t-(n-l)T) + \underbrace{\left[\pi h \sum_{k=-(L-1)}^{n-L} a_k\right] \bmod 2\pi}_{\Psi_n} \ . \tag{2.10}$$

While the first addend in (2.10) defines the phase trajectory of the n'th symbol, the second addend describes a phase offset Ψ_n which is called the phase state and dependent on all completed past symbols. All non-completed past symbols comprise the correlative state vector $\underline{A}_n = \left[a_{n-(L-1)}, \dots, a_{n-1}\right]$, which jointly forms the signal

state $\underline{S}_n = [\Psi_n, \underline{A}_n]$ with the phase state (cf. Section 2.2.2). The symbol vector $\underline{a} = \left[a_{-(L-1)}, \ldots, a_{-1}, a_0, \ldots, a_{N-1}\right]$ is comprised by $L-1$ symbols that define the initial symbol state of the burst and N transmitted data symbols.

2.2 Trellis Representation of the CPM Signal

2.2.1 Phase Tilting of the CPM Signal

As it is thoroughly outlined in [14], CPM can be interpreted as a trellis coded modulation and as such the symbol sequence is a Markov chain. For some parameter combinations the CPM trellis is time variant and its Markov chain is periodic with every two states and e.g. for implementation aspects this is a non-desired property. An elegant solution to generally make the trellis and thus its Markov chain time invariant is to introduce a phase tilt according to [14]. This tilt is usually realized by setting the carrier oscillator frequency in the transmitter or receiver off by

$$\xi = h\frac{M-1}{2T} \; . \tag{2.11}$$

The tilted phase function (cf. (2.10)) is then expressed as

$$\phi'_n(t, \underline{a}) = \phi_n(t, \underline{a}) + 2\pi\xi t \tag{2.12}$$

for $nT \leq t < (n+1)T$. While the tilt is inherently synchronized in time when realizing it in the process of mixing the signal up in the transmitter, it is not in the corresponding down mixing in the receiver. Chapter 3 will shed some more light on the consequences of this matter. To obtain a clean formulation of the tilted signal and to differentiate transmitter and receiver tilting, the boolean variable β is introduced

$$\beta = \begin{cases} 0 & \text{Tilting realized in receiver ,} \\ 1 & \text{Tilting realized in transmitter .} \end{cases} \tag{2.13}$$

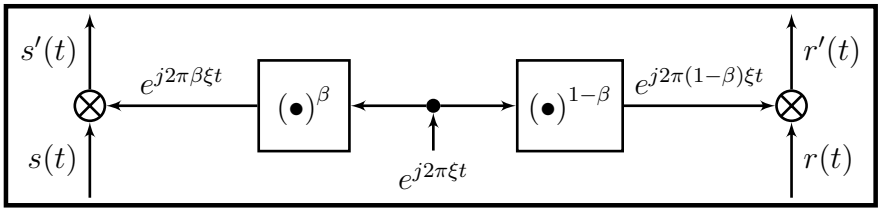


Figure 2-4: Block diagram of the CPM tilting depicted as a black block in Fig. 2-1. The tilting is realized either in the transmitter ($\beta = 1$) or in the receiver ($\beta = 0$).

According to (2.13) the phase tilt is either realized in the transmitter ($\beta = 1$) or in the receiver ($\beta = 0$). Figure 2-4 illustrates the tilting process.

The case of transmitter tilting is unusual in practice and more of theoretical interest in Chapter 3. For that matter, the tilting shall be realized in the receiver, i.e. $s'(t) \equiv s(t)$, unless stated otherwise.

2.2.2 Time Invariant CPM Trellis

In this work only invariant trellises of tilted CPM signals are considered and hence only the tilted phase function from (2.12) is of relevance. With the introduction of the unipolar symbols

$$a'_n = \frac{a_n + M - 1}{2} \in \mathcal{A}' = \{0, 1, \ldots, M-1\} \tag{2.14}$$

(2.12) can be reformulated in the n'th symbol slot as

$$\phi'_n(t, \underline{a}') = 4\pi h \sum_{l=0}^{L-1} a'_{n-l} q(t - (n-l)T) + \underbrace{\left[2\pi h \sum_{k=-L+1}^{n-L} a'_k \right] \bmod 2\pi}_{\Psi'_n} \ . \tag{2.15}$$

As it was already mentioned briefly in Section 2.1.2 the CPM trellis code is described by memory states

$$\underline{S}'_n = [\Psi'_n, \underline{A}'_n] \ , \tag{2.16}$$

The n'th phase state in the tilted case is defined by

$$\Psi'_n = \left[2\pi h \sum_{k=-L+1}^{n-L} a'_k \right] \bmod 2\pi \in \{0, \psi, \ldots, (Q-1)\psi\} \tag{2.17}$$

with $\psi = 2\pi/Q$, whereas

$$\underline{A}'_n = \left[a'_{n-(L-1)}, \ldots, a'_{n-1} \right] \tag{2.18}$$

describes the correlative state with unipolar symbols. The phase state's recursive relation is governed by

$$\Psi'_n = \left[\Psi'_{n-1} + 2\pi h \cdot a'_{n-L} \right] \bmod 2\pi \ . \tag{2.19}$$

In Figure 2-5, an exemplary trellis with its states and further concise properties is shown.

In the tilted CPM trellis, there are $N_{\Psi'} = Q$ phase states and $N_{\underline{A}'} = M^{L-1}$ correlative states, whose product gives the total number of states in each stage of the time invariant trellis

$$N_{\underline{S}'} = N_{\Psi'} \cdot N_{\underline{A}'} = Q \cdot M^{L-1} \ . \tag{2.20}$$

Each state $\underline{S}'_n$ has M output branches, which are determined by the current symbol a'_n. This gives a total number of

$$\Gamma = Q \cdot M^L \tag{2.21}$$

unique branches. A mapping function $m(\Psi'_n, \underline{A}'_n, a'_n)$ is introduced to establish a relation between the integral branch number $\gamma_n \in \{0, \ldots, \Gamma - 1\}$ and the current state $\underline{S}'_n$ with the current data symbol a'_n. Referring to Fig. 2-5, the trellis branches are numbered in ascending order from top to bottom and with the mapping function being defined as

$$\gamma_n = m(\Psi'_n, \underline{A}'_n, a'_n) = \frac{\Psi'_n}{\psi} \cdot M^L + \sum_{k=n-(L-1)}^{n} M^{n-k} \cdot a'_k \ . \tag{2.22}$$

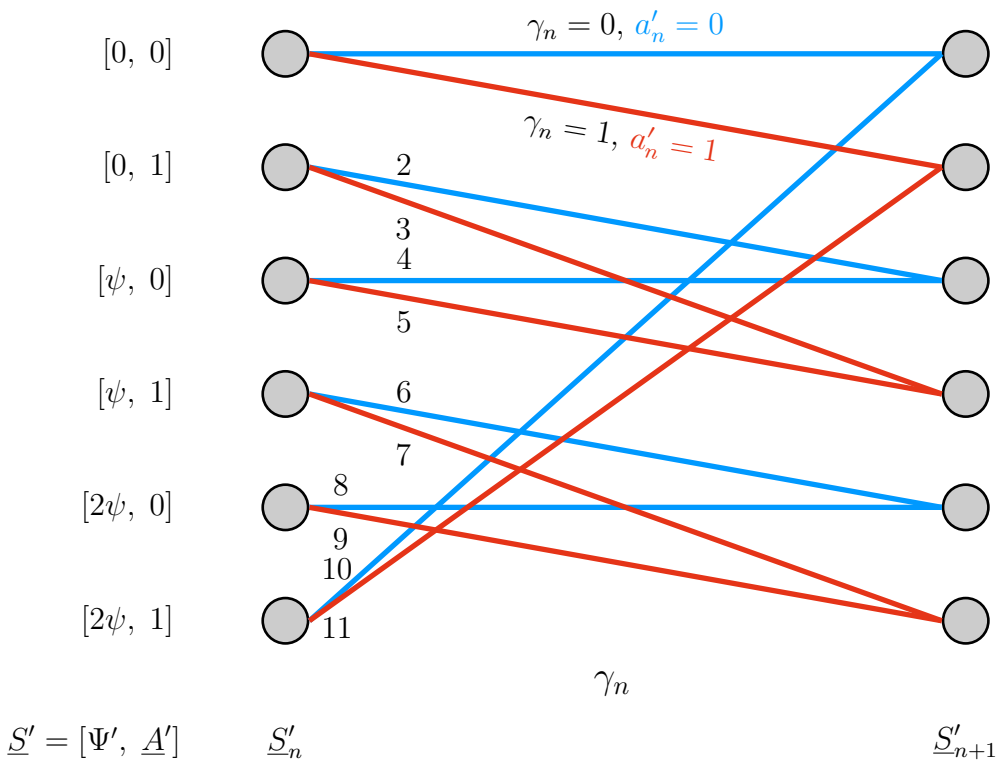


Figure 2-5: Exemplary trellis for a $h = 1/3$, $M = 2$, $L = 2$ waveform. The states are depicted as circles and the branches as lines. While blue lines denote a current $a'_n = 0$ symbol, red lines a 1. In this trellis, relevant variables from the text are: $\psi = 2\pi/3$, $N_{S'} = 6$, $N_{\Psi'} = 3$, $N_{A'} = 2$, $\Gamma = 12$. The ascending numeration of γ_n is given by black numbers labeling the branches. Examples of the successor (cf. (2.26)) and predecessor (cf. (2.27)) sets are $\mathcal{S}([\psi, 1]) = \{[2\psi, 0], [2\psi, 1]\}$ and $\mathcal{P}([\psi, 1]) = \{[0, 1], [\psi, 0]\}$.

Two mapping functions are defined, so that the phase state Ψ'_n and the symbols $\left[a'_{n-(L-1)}, \ldots, a'_n\right]$ that comprise the correlative state and the current symbol can be obtained by the partial inverse functions

$$\Psi'_n = m_{\Psi'_n}(\gamma_n) = \left\lfloor \frac{\gamma_n}{M^L} \right\rfloor \cdot \psi \tag{2.23}$$

$$a'_{n-l} = m_{a'_l}(\gamma_n) = \left\lfloor \frac{\gamma_n}{M^l} \right\rfloor \bmod M \quad , \ l \in \{0, \ldots, L-1\} \ . \tag{2.24}$$

The trellis branches γ_n will mostly be referred to as pseudo symbols (in contrast to the data symbols a'_n) in this work's remainder and its unique signal representation is

given as

$$c_{\gamma_n}(t) = \begin{cases} \exp\left(j\left[4\pi h \sum_{l=0}^{L-1} m_{a'_l}(\gamma_n) q(t+lT) + m_{\Psi'_n}(\gamma_n)\right]\right) & 0 \le t < T\ , \\ 0 & \text{otherwise}\ . \end{cases} \tag{2.25}$$

It is emphasized that the pseudo symbol's signal elements are not shifted by nT in time and tilted throughout this work.

Equations (2.16), (2.18) and (2.19) govern the transition relation between the trellis states and allow for the definition of the set of a state's possible successors and predecessors

$$\mathcal{S}(\underline{s}) \coloneqq \text{The set of all possible successors to state } \underline{s}, \tag{2.26}$$

$$\mathcal{P}(\underline{s}) \coloneqq \text{The set of all possible predecessors to state } \underline{s}. \tag{2.27}$$

and a further mapping function defines the relation between the pseudo symbol γ_n and two consecutive states $\underline{S}'_n$ and $\underline{S}'_{n+1}$

$$\gamma_n = m_{\gamma_n}\left(\underline{S}'_n = \underline{s}, \underline{S}'_{n+1} = \dot{\underline{s}}\right) \tag{2.28}$$

For one burst, there are $N+1$ trellis stages, that are interconnected by the N pseudo symbols. Whereas the initial state $\underline{S}'_0$ is usually defined by convention, a priori information about the final state $\underline{S}'_N$ is not necessarily available, which possibly deteriorates the receiver's performance [45]. Counter actions will be discussed in Sec. 2.5.2.

2.3 Radio Channel for Burst Transmissions in the VHF/UHF Band

The system model considered in this work includes a radio channel over which the transmission takes place. Radio channels have several properties which can impair the

transmission [11] that are discussed and specified in this section. As it was introduced in Section 1.1, the considered transmission frequencies lie in the VHF and UHF band, i.e. the carrier frequency take values between $30\,\mathrm{MHz} < f_\mathrm{c} < 400\,\mathrm{MHz}$.

2.3.1 Contemplations on the Channel Characteristics

The communication scenario involves non-stationary transmitters and receivers, whose movements induce time variance in the channel that is usually expressed by the Doppler spread. Two cases of aerial and ground vehicular network subscribers are considered with exemplary maximum velocities of $v_{\mathrm{max,\,aerial}} = 300\,\mathrm{m\,s^{-1}}$ and $v_{\mathrm{max,\,ground}} = 30\,\mathrm{m\,s^{-1}}$, respectively. The Doppler spread is a (two-sided) bandwidth $B_\mathrm{Dop} = 2f_{\mathrm{Dop,\,max}}$ that is limited by the maximum Doppler shift $f_{\mathrm{D,\,max}}$ to both sides. The Doppler shift f_Dop is maximized when the two transceivers move in exactly opposing directions [43]

$$|f_{\mathrm{Dop,\,max}}| = \frac{v_{\mathrm{max},1} + v_{\mathrm{max},2}}{c_0} \cdot f_\mathrm{c}\ . \tag{2.29}$$

In Table 2.1, the Doppler spreads are given for the two edge values 30 MHz and 400 MHz of the considered transmission band and for the three combinations of both aerial, both ground and one of each transceiver in the connection. The coherence time $T_\mathrm{coh} = B_\mathrm{Dop}^{-1}$ (also given in Table 2.1) gives a measure about the channel's temporal correlation or in other words how fast the channel is changing. In order to assume a quasi static (time invariant) channel over the course of one transmission burst, the coherence time has to be much larger than the burst length, which is $T_\mathrm{Burst} = 1\,\mathrm{ms}$ in the considered system. At least for lower carrier frequencies, this condition is fulfilled.

Depending on the topographical properties of the transmission environment, different power delay profiles can lead to different extents of frequency selective fading [43]. In [46] the results of a measurement campaign from 2012 regarding the channel conditions of ground communications in the range of 34.150 MHz to 363.750 MHz are reported. Four fundamental topographical scenarios were considered (urban, rural, hilly and mountainous) and typical power delay profiles are reported. Moreover, the metric $D_{99\,\%}$ is given for every scenario as the delay spread that is not exceeded

Table 2.1: Doppler spreads and coherence times for several scenarios. $v_{\max} = v_{\max,1} + v_{\max,2}$ is the maximum relative velocity of the respective transceivers.

	Aerial/Aerial		**Ground/Ground**		**Aerial/Ground**	
$\lvert v_{\max}\rvert$	$600\,\mathrm{m\,s^{-1}}$		$60\,\mathrm{m\,s^{-1}}$		$330\,\mathrm{m\,s^{-1}}$	
f_c	30 MHz	400 MHz	30 MHz	400 MHz	30 MHz	400 MHz
$\lvert f_{\mathrm{Dop,\,max}}\rvert$	60 Hz	800 Hz	6 Hz	80 Hz	33 Hz	440 Hz
B_{Dop}	120 Hz	1600 Hz	12 Hz	160 Hz	66 Hz	880 Hz
T_{coh}	8.3 ms	0.63 ms	83.3 ms	6.25 ms	15.2 ms	1.12 ms

Table 2.2: Delay spreads and coherence bandwidths of several environments from [46]. These results were reported for a carrier frequency of $f_C = 74.175\,\mathrm{MHz}$ and the delay spread values specify the upper bound for 99 % of the time.

	Urban	**Rural**	**Hilly**	**Mountainous**
$D_{99\,\%}$	2.2 µs	0.9 µs	4.9 µs	11.8 µs
B_{coh}	455 kHz	1.11 MHz	208 kHz	91 kHz

in 99 % of the measurement time. It is noted, that worst case delay spreads can lie above that value though. The delay spread's inverse is the coherence bandwidth $B_{\mathrm{coh}} = D_{99\,\%}^{-1}$ which gives an indicator of the channel's frequency selectivity. Both parameters are presented in Table 2.2 for the different scenarios.

In case $B_{\mathrm{coh}} \gg 2B_{99\,\%}$, the channel's frequency response can be considered flat, i.e. as a constant factor. Hereby $2B_{99\,\%}$ specifies the two-sided bandwidth which contains 99 % of the CPM signal's power and is calculated in this work according to [43]. In Table 5.1, the bandwidths of the considered waveforms are given and take values between $29.4\,\mathrm{kHz} \leq 2B_{99\,\%} \leq 97.4\,\mathrm{kHz}$.

This work will focus on a ground to ground communication link, for which a realistic channel model is available. In any specific combination of topography (excluding the mountainous environment) and signal waveform, the coherence bandwidth B_{coh} exceeds the signal bandwidth $B_{99\,\%}$ to a smaller or larger extent. For that fact, a

flat channel transfer function (i.e. a frequency non-selective) channel is assumed. Depending on the carrier frequency, the coherence time T_{coh} is larger or much larger than the burst duration T_{Burst} and thus suggests assuming a quasi static (time invariant) channel impulse response over the course of one burst. Furthermore, the mixing oscillator precision in the transceivers shall suffice to assume a constant sampling rate, a constant CFO and a constant CPO (i.e. a lack of phase noise) over the course of one burst. The time invariance of the channel shall be given for each burst, but due to spectral (in FH systems) or temporal (sensor networks) separation of transmission bursts, the channel states of any two bursts are mutually independent and thus no tracking of the channel is possible. To sum up, the channel of each burst will be modeled as an independent SF channel, which will serve as the base of designing the estimator in Chapter 4. The above described simplifications (time invariance and frequency flat fading for each burst) will be evaluated in Section 5.8.1, where the actual power delay profiles of [46] are considered.

2.3.2 System Modeling of the Channel

Based on the assumed SF channel model, only one propagation path with a scalar attenuation is considered and the Doppler spread simplifies to a Doppler shift. Furthermore, a carrier phase mismatch and time delay are considered. All necessary parameters (namely the FF α, CFO ν, CPO θ and TO τ) comprise the parameter vector

$$\underline{\lambda} = [\alpha, \nu, \theta, \tau] \tag{2.30}$$

and define the system operator alongside the additive white Gaussian noise (AWGN) term $w(t)$ with noise power density N_0. The system's output signal is

$$r(t) = \varphi\left\{s'(t), \underline{\lambda}\right\} = \alpha \cdot s'(t-\tau) \cdot e^{j(2\pi\nu(t-t_0)+\theta)} + w(t)\ , \tag{2.31}$$

which is graphically presented in Fig. 2-6. The time reference in the CFO term is relevant in terms of estimation performance and will be treated in detail in Chapter

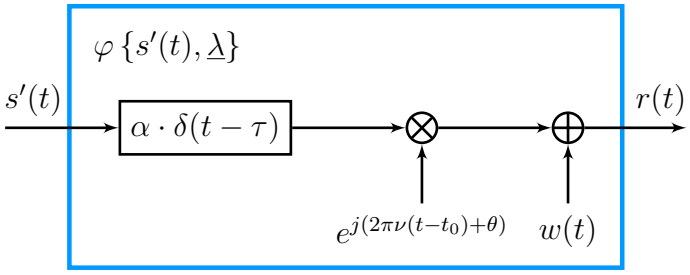


Figure 2-6: Block diagram of the channel's system operator depicted as a blue block in Fig. 2-1. The input of the system is scaled by α, delayed by τ, rotated by θ and shifted in frequency by ν. In the CFO term, the time reference t_0 is incorporated. The complex, additive white noise has a noise power density of N_0.

3. This SF channel model serves as base to designing the synchronization algorithm in Chapter 4.

In the following, the elements of $\underline{\lambda}$ are discussed and defined:

- According to [11], the path gain for a frequency flat channel can be modeled as a Rayleigh distributed random variable

$$\boldsymbol{\alpha} \sim \mathcal{R}(\kappa = 1/\sqrt{2}) \tag{2.32}$$

  in which the scaling factor κ was chosen so that $\mathrm{E}[\boldsymbol{\alpha}^2] = 1$. The phase rotation of the channel is considered below together with the CPO.

- The CFO is the sum of the Doppler shift and the frequency offset between the oscillators in transmitter and receiver. While the first depends on the velocity of the subscribers, the second is determined by the oscillator accuracy and thus both are proportional to the carrier frequency. Since the CFO has two origins with arguably two different statistical properties, it first seems reasonable to consider them separated. While a PDF proportional to a Doppler spectrum would be a sensible choice for the frequency shift in the channel [43], the statistics of the oscillator deviations is hardware dependent. To keep the approach simple, both effects are considered mutually and a uniform distribution is chosen as it is non-informative and thus the most general approach. The oscillator accuracy of typical, commercially available software defined radio

(SDR)s is about 2.5 ppm (parts per million) when not externally synchronized and 0.01 ppm when they are [47]. While hardware with inferior quality may be installed in cheap sensors, tactical radios can be assumed to have high quality oscillators with a self-sustaining quality of 1 ppm. This value is chosen for the further considerations and it is referred to Section 4.4.5 for the consequences of other oscillator choices. In combination with $f_\mathrm{C} = 400\,\mathrm{MHz}$, this leads to a maximum possible offset of

$$|f_{\mathrm{Osc,\,max}}| = 2 \cdot 10^{-6} \cdot 400\,\mathrm{MHz} = 800\,\mathrm{Hz} \tag{2.33}$$

due to each transceiver's hardware inaccuracies in opposing directions. The sum of the Doppler spread (160 Hz, cf. Table 2.1 for ground communications at $f_\mathrm{c} = 400\,\mathrm{Hz}$) and the oscillator effect ($2 \cdot 800\,\mathrm{Hz}$, cf. (2.33)) gives the width of the CFO process's uniform distribution

$$\boldsymbol{\nu} T \sim \mathcal{U}\left(-0.021, +0.021\right) \tag{2.34}$$

in which a symbol rate of $T^{-1} = 42\,\mathrm{kBaud}$ (cf. Table 5.1) was inserted to obtain normalized values for the CFO.

- As it was for the CFO, the origins of the CPO are twofold. First, in a frequency flat channel, the phase rotation of the channel is usually modeled uniformly around the unit circle [11] and second, the generally independent oscillator phases in the transceivers translate to an equally circular uniform distribution of their offset. The sum of both offsets gives the modeled CPO process and its distribution is the convolution of both circular uniform distributions which gives

$$\boldsymbol{\theta} \sim \mathcal{U}\left(-\pi, +\pi\right) \tag{2.35}$$

due to the wrapping at $\pm\pi$.

- In FH systems, a frame synchronization is usually available, which leads a

coarse range of a residual TO to be estimated. The range and the statistics of the values within is naturally dependent on the synchronization technique on which no assumptions shall be made at this point. Therefore the TO is assumed to be uniformly distributed over one symbol duration, i.e.

$$\boldsymbol{\tau}/T \sim \mathcal{U}\left(-0.5, +0.5\right) \ . \tag{2.36}$$

The synchronization algorithm derived in Chapter 4 will not assume any prior knowledge about the parameters, though it will be pointed out how to incorporate such information in the algorithm.

2.4 Front End

As shown in Fig. 2-1, the input signal to the front end is the tilted signal ($\beta = 0$)

$$\begin{aligned} r'(t) &= \varphi\left\{s(t), \underline{\lambda}\right\} \cdot e^{j2\pi\xi t} \\ &= \alpha e^{j(2\pi\nu(t-t_0)+\theta)} \cdot s(t-\tau) \cdot e^{j2\pi\xi t} + w'(t) \ . \end{aligned} \tag{2.37}$$

Due to the tilting happening in the receiver (cf. Section 2.2.1), the additive noise $w'(t)$ is the tilted version of $w(t)$ but with identical statistical properties. At this point, $w'(t)$ is modeled as with the constant complex (lowpass domain) noise power density $N_0 > 0$, is not band limited and thus of infinite power and only has practical relevance by the following band limitation. The receive filter in the lowpass domain is modeled as an optimal rectangular function

$$H_{\mathrm{R}}(f) = \begin{cases} 1 & |f| \leq B_0 \\ 0 & \text{otherwise} \end{cases} \tag{2.38}$$

with B_0 being the one-sided system bandwidth. The system bandwidth is chosen so that $H_{\mathrm{R}}(f)$ shall let the CPM signal pass effectively without distortions (respecting the potential frequency offsets ν and ξ). Therefore, the received signal is still called

r' after the receive filter. The noise power is now limited to $2B_0N_0$. Next, the filtered signal is sampled under (effectively) meeting the Nyquist sampling theorem at a rate of T_0^{-1} and given as

$$r'(kT_0) = r'(t)|_{t=kT_0} \tag{2.39}$$

with k being an integer in the interval $0 \leq k \leq NK-1$ and $K = T/T_0$ representing the oversampling factor. The energy of the signal is computed according to (1.5). The time continuous and discrete representations are exchangeable when the sampling theorem is met [17]. While the continuous representation is convenient to derive the theoretical bounds in Chapter 3, the digital receiver parts in this chapter and in Chapter 3 are described by the time discrete signal. The sampled values of one burst $r'(kt_0)$ can be stored in a vector

$$\underline{r}'\,[0 : NK-1] = [r'(0), r'(T_0), r'(2T_0), \ldots, r'(NT-2T_0), r'(NT-T_0)] \tag{2.40}$$

such that the vector's argument specifies its lower and upper index limit separated by the colon. If, as in the equation above, the complete burst is stored in the vector, the limit specification is usually omitted and only $\underline{r}'$ is written. The vectorized representation is used mainly in the arguments of PDFs and PMFs as well as in the discussion of implementation aspects in Section 4.5.

By choosing the inverse of the (two-sided) system bandwidth (and as such the receive filter width) as the sampling interval $T_0 = 1/(2B_0)$, the sampled noise process $\boldsymbol{w}'(kT_0)$ keeps its whiteness due to the fitting spectral repetition and its presumed stationarity [43]. For that reason, the noise variable is still called $\boldsymbol{w}'$ although it is not the same signal anymore and it is modeled as a complex normal distribution with

the same power as the continuous noise process before the sampling

$$\boldsymbol{w}'(kT_0) \sim \mathcal{CN}\left(0, \sigma^2_{w'} = \frac{N_0}{T_0}\right) , \tag{2.41}$$

$$\mathrm{Re}\left\{\boldsymbol{w}'(kT_0)\right\} \sim \mathcal{N}\left(0, \sigma^2_{\mathrm{Re}\{w'\}} = \frac{N_0}{2T_0}\right) , \tag{2.42}$$

$$\mathrm{Im}\left\{\boldsymbol{w}'(kT_0)\right\} \sim \mathcal{N}\left(0, \sigma^2_{\mathrm{Im}\{w'\}} = \frac{N_0}{2T_0}\right) , \tag{2.43}$$

$$R_{\mathrm{Re}\{\boldsymbol{w}'\},\mathrm{Im}\{\boldsymbol{w}'\}} = 0 . \tag{2.44}$$

The CCF in (2.44) of $\mathrm{Re}\{\boldsymbol{w}'\}$ and $\mathrm{Im}\{\boldsymbol{w}'\}$ states, that the real and imaginary parts of $\boldsymbol{w}'(kT_0)$ are uncorrelated. The system introduced above is rather theoretical and its practical implications shall be discussed briefly. First, the rectangular receive filter is to be replaced by another technically more feasible Nyquist filter. Second, if for whatever reason the filter bandwidth $2B_0$ significantly exceeds the physical signal bandwidth, noise with higher power than necessary and even potential interference by other (communication) signals are passed to the detector succeeding the receive filter. In this constellation, this could happen e.g. when a specific sampling rate T_0^{-1} is desired. If then a filter with a width better suited to the transmit signal is applied, the noise samples will lose their property of being uncorrelated, which is usually dealt with by applying a whitening filter [43].

For soft input detectors such as the below following SISO CPM detector, the knowledge about the noise power or rather variance is crucial. It could e.g. be estimated directly by measuring the received power in an unoccupied time slot in the FH system. Another way in CPM would be to measure the amplitude variance that has to stem from the additive noise in a flat fading channel. The noise variance is estimated in this work by the formula

$$\hat{\sigma}^2_{w'} = \frac{\hat{N}_0}{T_0} = 2\sqrt{\mathrm{Var}\left[\left|\boldsymbol{r}'(kT_0)\right|^2\right]} , \tag{2.45}$$

which is derived in Appendix A.1. The idea behind this is, that in a constant amplitude signal all amplitude variations must be caused by the additive noise. This is of

course only valid in the SF channel model motivated above. The actual estimate is calculated by computing the variance of the realization of the received signal. Large noise samples (in relation to the transmit signal samples) can switch the sign of the superposed received signal. The estimator in (2.45) takes only the absolute value though, so that the noise estimates tend to be too small in these cases.

2.5 Soft-Input-Soft-Output CPM Detector

As the SISO detector acts as the link between the two further central receiver elements (cf. Fig. 2-1), this section introduces it in detail. Recalling (2.37) and (2.39), the detector's input signal is expressed as

$$r'(kT_0) = \alpha e^{j(2\pi\nu(kT_0 - t_0) + \theta)} \cdot s(kT_0 - \tau) \cdot e^{j2\pi\xi kT_0} + w'(kT_0) \tag{2.46}$$

for $0 \leq k < KN$. If the channel TO is assumed to be known (or estimated well enough), the CPM phase is appropriately tilted at this point (cf. (2.12)) and hence the trellis is time invariant. In this work a coherent CPM detector is deployed for which also the CFO and CPO have to be synchronized properly. The absolute knowledge about $\underline{\lambda}$ is in general not available to the receiver, only estimates are. The following derivations are carried out under the assumption, that a reliable estimate $\hat{\underline{\lambda}}$ of the channel parameter vector is known. The noise variance was already estimated in a preceding receiver block and thus $\hat{\sigma}^2_{w'}$ is also available.

2.5.1 MAP Detector for a Trellis Coded CPM Signal

As it was introduced in 2.2 the CPM signal can be interpreted as a trellis coded modulation and thus the utilization of the BCJR algorithm produces a posteriori probabilities of the transmitted data symbols $a'_n \in \{0, \ldots, M-1\}$ [48]. Hence it realizes the maximum a posteriori (MAP) detector for the CPM signal. The symbol

a posteriori probabilities

$$\mathrm{P}_{\boldsymbol{a}'_n|\underline{\boldsymbol{r}}'}\left[a'_n|\underline{r}',\underline{\hat{\lambda}}\right]=\sum_{\Psi'_n}\sum_{A'_n}\mathrm{P}_{\gamma_n|\underline{\boldsymbol{r}}'}\left[\underbrace{m_{\gamma_n}\left(\left[\Psi'_n,\underline{A}'_n\right],a'_n\right)}_{\text{Eq. (2.22)}}|\underline{r}',\underline{\hat{\lambda}}\right] \tag{2.47}$$

for $0\leq n\leq N-1$ are calculated as the sum of probabilities of every trellis branch γ_n that represents the data symbol a'_n and are required as soft inputs for the channel decoder. A posteriori probabilities of the trellis branches themselves

$$\mathrm{P}_{\boldsymbol{\gamma}_n|\underline{\boldsymbol{r}}'}\left[\gamma_n|\underline{r}',\underline{\hat{\lambda}}\right]=\frac{\mathrm{p}_{\boldsymbol{\gamma}_n,\underline{\boldsymbol{r}}'}\left[\gamma_n,\underline{r}'|\underline{\hat{\lambda}}\right]}{\mathrm{p}_{\underline{\boldsymbol{r}}'}\left[\underline{r}'|\underline{\hat{\lambda}}\right]} \tag{2.48}$$

for $\gamma_n\in\{0,\ldots,\Gamma-1\}$ are direct outputs of the BCJR (cf. (2.52)) and are crucial for the parameter estimation in Chapter 4.

The essential components of the BCJR are the forward $f_n^F(\underline{s})$ and backward $f_n^B(\dot{\underline{s}})$ state occupation probability measures as well as the path metric $f_n^M(\underline{s},\dot{\underline{s}})$ and are defined as the following PDFs

$$f_n^F(\underline{s})=\mathrm{p}_{\underline{\boldsymbol{S}}'_n,\underline{\boldsymbol{r}}'}\left[\underline{s},\underline{r}'\left[0:nK-1\right]|\underline{\hat{\lambda}}\right]\ , \tag{2.49}$$

$$f_n^B(\dot{\underline{s}})=\mathrm{p}_{\underline{\boldsymbol{r}}'|\underline{\boldsymbol{S}}'_n}\left[\underline{r}'\left[nK:NK-1\right]|\dot{\underline{s}},\underline{\hat{\lambda}}\right]\ , \tag{2.50}$$

$$f_n^M(\underline{s},\dot{\underline{s}})=\mathrm{p}_{\underline{\boldsymbol{S}}'_{n+1},\underline{\boldsymbol{r}}'|\underline{\boldsymbol{S}}'_n}\left[\dot{\underline{s}},\underline{r}'\left[nK:(n+1)K-1\right]|\underline{s},\underline{\hat{\lambda}}\right]\ . \tag{2.51}$$

It is noted, that the process $\underline{\boldsymbol{r}}'$ is denoted in distributions for the sake of readability although only parts of the received vector $\underline{r}'$ are subject to the distribution. The product of (2.49), (2.50) and (2.51) gives the joint probability that can be inserted in (2.48)

$$\mathrm{p}_{\boldsymbol{\gamma}_n,\underline{\boldsymbol{r}}'}\left[m_{\gamma_n}(\underline{s},\dot{\underline{s}}),\underline{r}'|\underline{\hat{\lambda}}\right]=\mathrm{P}_{\underline{\boldsymbol{S}}'_n,\underline{\boldsymbol{S}}'_{n+1}}\left[\underline{s},\dot{\underline{s}}\right]\cdot f_n^F(\underline{s})\cdot f_n^M(\underline{s},\dot{\underline{s}})\cdot f_{n+1}^B(\dot{\underline{s}}), \tag{2.52}$$

where the mapping function from (2.28) was used. The term $\mathrm{P}_{\underline{\boldsymbol{S}}'_n,\underline{\boldsymbol{S}}'_{n+1}}\left[\underline{s},\dot{\underline{s}}\right]$ denotes a prior information about the state transitions, e.g. from a channel decoder in a SC CPM receiver. In case no prior knowledge about the transition is (yet) available,

the probability equals the constant Γ^{-1}. The forward and backward variables are computed recursively as

$$f_n^F(\underline{s}) = \sum_{\dot{\underline{s}} \in \mathcal{P}(\underline{s})} \mathrm{P}_{\underline{\boldsymbol{S}}'_n, \underline{\boldsymbol{S}}'_{n+1}} [\dot{\underline{s}}, \underline{s}] \cdot f_{n-1}^F(\dot{\underline{s}}) \cdot f_n^M(\dot{\underline{s}}, \underline{s}) \tag{2.53}$$

$$f_n^B(\underline{s}) = \sum_{\dot{\underline{s}} \in \mathcal{S}(\underline{s})} \mathrm{P}_{\underline{\boldsymbol{S}}'_n, \underline{\boldsymbol{S}}'_{n+1}} [\underline{s}, \dot{\underline{s}}] \cdot f_{n+1}^B(\dot{\underline{s}}) \cdot f_n^M(\underline{s}, \dot{\underline{s}}) \ , \tag{2.54}$$

with initial values $f_0^F(\underline{s})$ and $f_N^B(\underline{s})$ discussed in Section 2.5.2. It is straight forward that the sum over all possible forward state occupation probabilities of the last state is equal to the likelihood of the parameter $\underline{\lambda}$

$$\mathcal{L}(\hat{\underline{\lambda}}) = \mathrm{p}_{\underline{\boldsymbol{r}}'} \left[\underline{r}' | \hat{\underline{\lambda}} \right] \tag{2.55}$$

$$= \sum_{\underline{s}} \mathrm{p}_{\underline{\boldsymbol{S}}'_N, \underline{\boldsymbol{r}}'} \left[\underline{s}, \underline{r}' | \hat{\underline{\lambda}} \right] \tag{2.56}$$

$$= \sum_{\underline{s}} f_N^F(\underline{s}) \ . \tag{2.57}$$

This property will find application in the coarse parameter estimation detailed in Section 4.5.1.

The path metric in (2.51) is transformed under the use of Bayes' theorem into

$$f_n^M(\underline{s}, \dot{\underline{s}}) = \mathrm{P}_{\underline{\boldsymbol{S}}'_{n+1} | \underline{\boldsymbol{S}}'_n} \left[\dot{\underline{s}} | \underline{s}, \hat{\underline{\lambda}} \right] \cdot \mathrm{p}_{\underline{\boldsymbol{r}}' | \underline{\boldsymbol{S}}'_n, \underline{\boldsymbol{S}}'_{n+1}} \left[\underline{r}' \left[nK : (n+1)K - 1 \right] | \underline{s}, \dot{\underline{s}}, \hat{\underline{\lambda}} \right] \tag{2.58}$$

$$= 1/M \cdot \mathrm{p}_{\underline{\boldsymbol{r}}' | \underline{\boldsymbol{S}}'_n, \underline{\boldsymbol{S}}'_{n+1}} \left[\underline{r}' \left[nK : (n+1)K - 1 \right] | \underline{s}, \dot{\underline{s}}, \hat{\underline{\lambda}} \right] \tag{2.59}$$

since the event of two successive states $\underline{S}'_n = \underline{s}$ and $\underline{S}'_{n+1} = \dot{\underline{s}}$ in a Markov chain generally does not have any prior information by itself and furthermore is independent on $\underline{\lambda}$. The path metrics as it is expressed in (2.59) are the probability densities of the received pseudo symbols $\gamma_n = m_{\gamma_n}(\underline{s}, \dot{\underline{s}})$ given the trellis transition that define these. The corresponding sampled signal $c_{\gamma_n}(kT_0)$ (cf. (2.25)) of each transition is deterministic and the metrics calculation in an AWGN environment writes as the applied definition of the multivariate $(2K)$ normal distribution. Since all noise samples are

i.i.d., it can be expressed as the product of $2K$ independent normal distributions

$$f_n^M(\underline{s},\underline{\dot{s}}) = \frac{1}{M\left(\pi\hat{\sigma}_{w'}^2\right)^K} \cdot \exp\left(-\frac{1}{\hat{\sigma}_{w'}^2}\sum_{k=0}^{K-1}\left|r'(kT_0+nT) - c_{m_{\gamma_n}(\underline{s},\underline{\dot{s}})}(kT_0,\hat{\underline{\lambda}})\right|^2\right) \quad (2.60)$$

considering $\hat{\sigma}_{w'}^2 = 2\sigma_{\mathrm{Im}\{w'\}}^2 = 2\sigma_{\mathrm{Re}\{w'\}}^2$. The knowledge of $\hat{\underline{\lambda}}$ is used to scale, rotate (with the time origin t_0 being respected in the CFO correction) and delay the pseudo symbols to match the received signal

$$c_{m_{\gamma_n}(\underline{s},\underline{\dot{s}})}(kT_0,\hat{\underline{\lambda}}) = \hat{\alpha}e^{+j\left(2\pi[\hat{\nu}(kT_0+nT-t_0)+\xi\hat{\tau}]+\hat{\theta}\right)}c_{m_{\gamma_n}(\underline{s},\underline{\dot{s}})}(kT_0-\hat{\tau})\ . \quad (2.61)$$

Since the received signal was tilted ahead of a TO correction, the term $+j2\pi\xi\hat{\tau}$ accounts for the thereby induced phase offset. Furthermore the pseudo symbols must experience a CFO according to their position n. The approach of manipulating the pseudo symbols has several advantages over correcting the received signal instead. First, the received signal does not need to be interpolated to obtain correct for the induced channel delay and second, the noise power in (2.60) does not need to be adjusted. By using the equality from Appendix A.2, (2.60) is reformulated as

$$\begin{aligned} f_n^M(\underline{s},\underline{\dot{s}}) = \frac{1}{M}\cdot\frac{1}{\left(\pi\hat{\sigma}_{w'}^2\right)^K}\times & \\ \exp\left(-\frac{1}{\hat{\sigma}_{w'}^2}\sum_{k=0}^{K-1}\left[\left|r'(kT_0+nT)\right|^2 + \left|c_{m_{\gamma_n}(\underline{s},\underline{\dot{s}})}(kT_0,\hat{\underline{\lambda}})\right|^2\right]\right)\times & \quad (2.62)\\ \exp\left(-\frac{1}{\hat{\sigma}_{w'}^2}\sum_{k=0}^{K-1} -2\,\mathrm{Re}\left\{r'(kT_0+nT)\cdot c^*_{m_{\gamma_n}(\underline{s},\underline{\dot{s}})}(kT_0,\hat{\underline{\lambda}})\right\}\right) & \\ \propto \exp\left(-\frac{1}{\hat{N}_0}\hat{\alpha}^2E_\mathrm{S} + \frac{2T_0}{\hat{N}_0}\sum_{k=0}^{K-1}\mathrm{Re}\left\{r'(kT_0+nT)\cdot c^*_{m_{\gamma_n}(\underline{s},\underline{\dot{s}})}(kT_0,\hat{\underline{\lambda}})\right\}\right) & \quad (2.63) \end{aligned}$$

with all constant factors being omitted in (2.63). Whereas the energy of the received symbol $r'(kT_0+nT)$ is obviously constant, all pseudo symbols $c_{m_{\gamma_n}(\underline{s},\underline{\dot{s}})}(kT_0,\hat{\underline{\lambda}})$ have the equal energy $\hat{\alpha}^2E_\mathrm{S}$ due to the constant amplitude property of CPM and thus their squared amplitude is also constant. However, for (2.57) to produce comparable results for different FFs in $\hat{\underline{\lambda}}$, it must be included.

The evaluation of (2.63) can e.g. be realized by convolving $r'(kT_0+nT)$ with a

filter bank with filters $h_\gamma(kT_0) = c_\gamma^*(-kT_0, \hat{\underline{\lambda}})$ and sampling the result in the symbol clock T.

2.5.2 Initialization of the State Occupation Variables

In [45, 49], the negative effect on the symbol error rate is discussed when the start and end states of the CPM trellis are not appropriately initialized. Thus, in order to carry out the recursive calculations of (2.53) and (2.54) effectively, the forward state occupations of the first stage $f_0^F(\underline{s})$ and backward state occupations of the last stage $f_N^B(\dot{\underline{s}})$ need to be initialized. While there are methods for defining the final trellis state, they usually pose spectral (trellis termination [50]) or computational (tail biting [45]) overhead. Especially in burst systems the spectral overhead in terminating may become significant. Several initialization methods are compared in [49]. In this work, the correlative zero state (CZS) from [49] that constitutes a trade off between the spectral efficiency and impact on the symbol error rate is chosen. Figure 2-7 graphically demonstrates an exemplary CZS termination.

By applying the CZS method, the CPM signal is terminated by $N_{\mathrm{CZS}} = L - 1$ zero CPM symbols to reach the correlative state $\underline{A}'_N = \Big[a_{N-(L-1)} = 0, \ldots, a_{N-1} = 0\Big]$. The phase state Ψ'_N is not defined and therefore arbitrary. In contrast, classical termination [50] spends more symbols to reach a predefined phase and correlative state. By setting $\underline{A}'_0 = \underline{A}'_N = 0$, the trellis possesses a defined start and end correlative state. In CPM tail biting the first and last correlative state are also equal but unlike to the CZS approach not known, which entails the computation of $N_{\underline{A}'}$ BCJR instances [45].

In tilted CPM, all signals $r'(t) \cdot e^{jz\psi}$ represent the same digital message as $r'(t)$ [14] with z being an integer and ψ representing the phase state's angle. This property is called rotational invariance and will prove useful in Section 4 since a received signal does not need to be synchronized to the exact CPO but rather only to the remainder $\theta_0 = \theta \bmod \psi$. The reduced range $\theta_0 \in [-\psi/2; +\psi/2]$ makes the estimation easier (cf. Section 4.4) without deteriorating the signal's detection. For this to work, the CPM trellis must allow all states with a CZS $\underline{A}'_0 = \underline{0}$ as a start state opposed to only the all

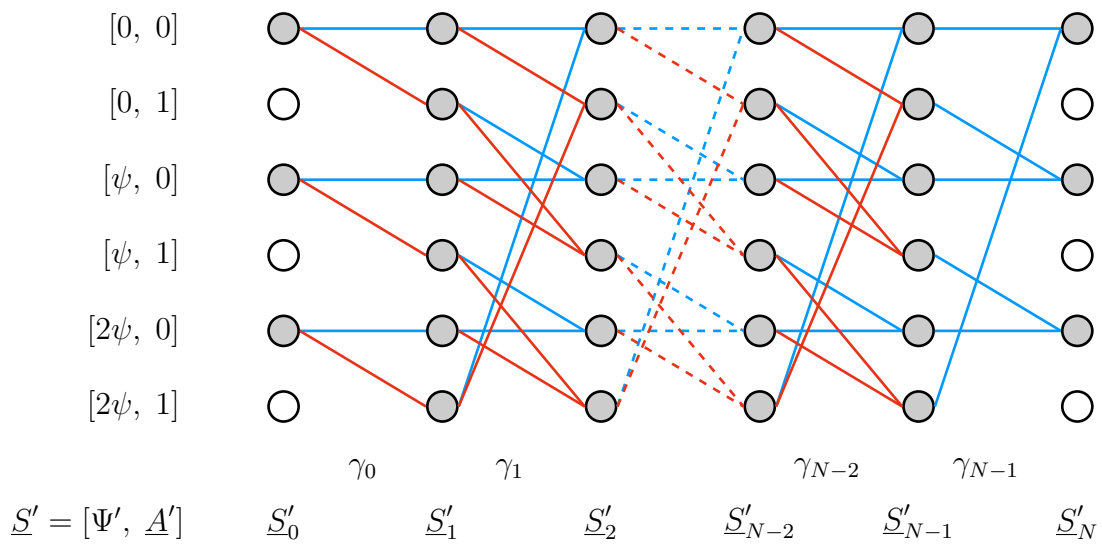


Figure 2-7: Visualization of the trellis (cf. Fig. 2-5) initialization by the CZS termination method. $\underline{A}'_0 = \underline{A}'_{N+1} = \underline{0}$ determines the correlative state of $\underline{S}'_0$ and $\underline{S}'_{N+1}$. While the last phase state Ψ'_{N+1} is generally unknown, Ψ'_0 is left unfixed in case of CPO estimation due to CPM's rotational invariance. White states are unreachable and thus their branches are omitted. The dashed lines indicate the middle part of the trellis.

zero sate $\underline{S}'_0 = [0, \underline{0}]$. Figure 2-7 also accounts for this fact and depicts the start and end trellis states initialization used in this work. In (2.61), $\hat{\theta}$ can be replaced with $\hat{\theta}_0$ without further consequences.

2.6 Channel Coding

Albeit convolutional codes being a common outer channel code for serially concatenated CPM systems with iterative decoding [44], [51], the code of choice in this work is the Polar code [52]. This is due to the need for trellis initialization in convolutional codes for the same reasons as addressed in Sec. 2.5.2. The Polar code as a block code does not have this restriction and can be computationally as well as spectrally more efficient than a comparable convolutional code in the application of burst transmissions [53]. Because of the lower bit error rate at the same code rate [54], a systematic Polar code is used in this work. The decoder is realized according to [55] and called successive cancellation (SCAN) decoder, while variable code rates are achieved by shortening the Polar codes [56].

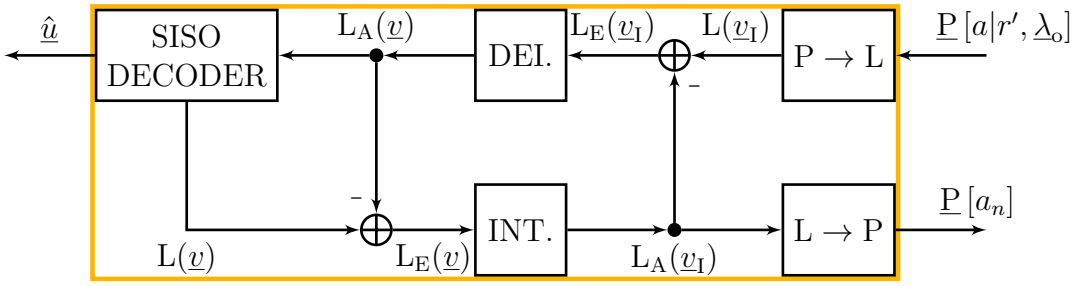


Figure 2-8: Block diagram of the SC CPM channel decoder that depicts the iterative information passing between the decoder and the detector. Bold letters indicate vectors and $\mathrm{L}(\underline{x})$ contains the LLRs of every element of $\underline{x}$. In the case of PMFs, a probability vector of every element $a'_n \in \mathcal{A}',\ 0 \leq n < N$) out of $\underline{a}'$ is passed. For the sake of clarity, the probabilities vectors are not labeled with the corresponding processes.

Figure 2-1 shows the encoding procedure: After systematically encoding the information bits $\underline{u}$, the code bits $\underline{v}$ are interleaved to avoid sequence error propagation. To obtain the CPM data symbols $\underline{a}$, the interleaved bits $\underline{v}_\mathrm{I}$ are converted according to Gray mapping [11]. In Fig. 2-8, the SC counterpart on the receiver side is depicted.

Together with the SISO detector, this block builds the SC CPM scheme with iterative decoding on which [44] and [51] give details. It is shortly explained as follows. The a posteriori probability vectors $\underline{\mathrm{P}}_{\boldsymbol{a}'_n|\underline{r}'}\left[a'_n|\underline{r}', \hat{\underline{\lambda}}\right]$ of the CPM symbols are mapped to code bit probabilities and converted to log likelihood ratios $\mathrm{L}(\underline{v}_\mathrm{I})$ to be passed on the forward branch. After being deinterleaved, they are fed as a priori $\mathrm{L_A}(\underline{v})$ to the channel decoder, which generates a posteriori information $\mathrm{L}(\underline{v})$ that is passed backward. To obtain extrinsic information $\mathrm{L_E}(\underline{v}) = \mathrm{L}(\underline{v}) - \mathrm{L_A}(\underline{v})$, the information of the forward branch is subtracted. To restrict the exchange to extrinsic information is crucial to all kinds of Turbo decoding setups such as e.g. parallel concatenated [42] or block [57] Turbo codes. Afterwards $\mathrm{L_E}(\underline{v})$ is reinterleaved and converted to the probability domain to be passed as a priori probabilities $\underline{\mathrm{P}}_{\boldsymbol{a}'_n}[a_n]$ of the CPM symbols back to the SISO detector. The state transition probabilities $\mathrm{P}_{\underline{\boldsymbol{S}}_n, \underline{\boldsymbol{S}}_{n+1}}[\underline{s}, \dot{\underline{s}}]$ in (2.52), (2.53) and (2.54) are then set to $\underline{\mathrm{P}}_{\boldsymbol{a}'_n}[a_n]$ according to which transition the data symbols a'_n correspond. In case more than one loop is carried out, the extrinsic information in the forward branch must be calculated by $\mathrm{L_E}(\underline{v}_\mathrm{I}) = \mathrm{L}(\underline{v}_\mathrm{I}) - \mathrm{L_A}(\underline{v}_\mathrm{I})$. Investigations about the usefulness of this iterative setup in short burst transmissions were conducted in [53] with the result that none or one iteration provide the best cost

performance trade off. Eventually, estimates $\underline{\hat{u}}$ of the information bits are obtained by the channel decoder.

2.7 Key Points of the Chapter

- The CPM signal was introduced in the bandpass and lowpass domain and the trellis properties for tilted signals were discussed.
- Channel conditions in the VHF and UHF transmission band were addressed in the ground to ground communication scenario. Hereby the Doppler spread of moving transceivers was defined and delay spreads of typical environments were presented. By comparing the corresponding coherence time and bandwidths with the signal properties, the channel model was reasoned to be simplified to a time invariant and frequency flat model, on which the estimator in Chapter 4 is based on.
- The optimal detection of the CPM signal in the SF channel was treated by means of the BCJR algorithm, which constitutes the MAP detector. Mapping functions between trellis transitions and data symbols have been defined to prepare the notation used in later chapters.

This Page Intentionally Left Blank

Chapter 3

Theoretical Treatment of Maximum Likelihood Parameter Estimation

The general problem of parameter estimation comes in manifold applications and one of them lies in communications systems. The estimation of modulation parameters and of channel parameters are two exemplary cases with this thesis focusing on the latter. While there are several estimation rules in general, the two approaches of maximum likelihood (ML) and MAP are dominant in communications engineering [11]. In the following, the ML approach is followed for the reasons outlined in Section 3.1. The connection to the MAP rule is given when sensible.

According to [58], there are four components that describe the general estimation problem. The specific problem of this thesis under the use of an ML estimation rule fits the framework as follows and as illustrated in Fig. 3-1:

- *Parameter space*: This space is defined in Section 2.3.2 by (2.32), (2.34), (2.35) and (2.36) that govern the single parameter spaces of the vector $\underline{\lambda} = [\alpha, \nu, \theta, \tau]$.
- *Probabilistic Mapping from Parameter Space to Observation Space*: The (blind) estimation problem of channel parameters $\underline{\lambda}$ in communication systems poses mainly two obstacles. First, the parameters are not directly observable, but

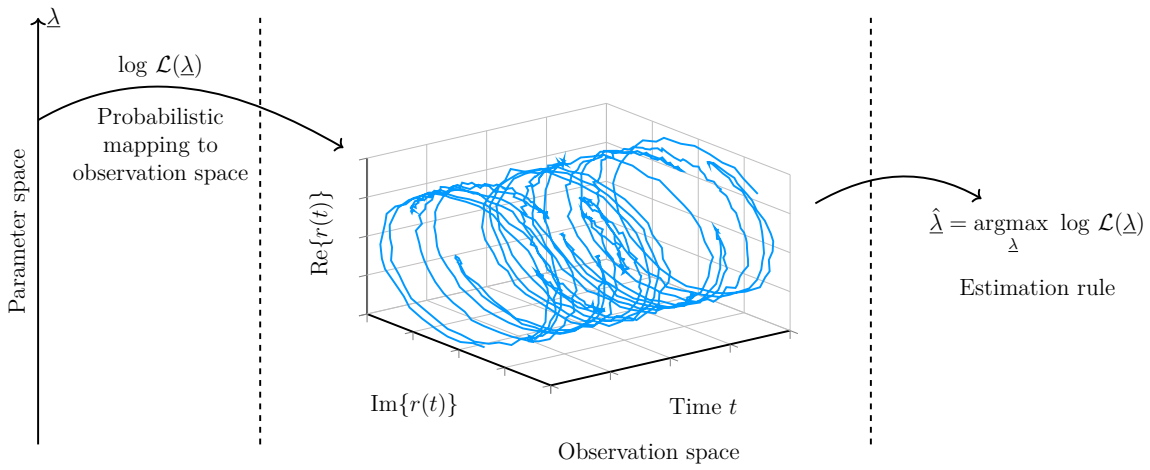


Figure 3-1: ML channel parameter estimation problem set in a general framework in the style of [58]. The observation space consists of a noisy version of the transmit signal.

only through a noisy and deviated version $r(t)$ of the original. Second, as any communication signal, the CPM signal carries unknown data $\underline{a}$, that are a nuisance in terms of parameter estimation. While the first one for itself would not be too hard to overcome (e.g. by the use of pilot sequences), the second one makes the estimation problem challenging. The probabilistic relation between the channel parameters and the observation used in this work is given by the likelihood function $\mathcal{L}(\underline{\lambda})$ or equally by the LLF of $\underline{\lambda}$

$$\log\ \mathcal{L}(\underline{\lambda}) = \log\left(\mathrm{p}_{\underline{\boldsymbol{r}}}\left[\underline{r}|\underline{\lambda}\right]\right) \tag{3.1}$$

$$= \log\left(\sum_{\underline{a}} \mathrm{p}_{\underline{\boldsymbol{r}},\underline{\boldsymbol{a}}}\left[\underline{r},\underline{a}|\underline{\lambda}\right]\right) . \tag{3.2}$$

The difficulty in handling the LLF arises from the fact, that the logarithm's argument in (3.2) is a sum of exponential terms in the case of i.i.d. noise samples (cf. (2.52) and its following equations) and thus cannot be reduced with the logarithm function. In the above equation, $\underline{a}$ is a vector and describes the sequence of data symbols that are in direct relationship to the trellis transitions $\underline{\gamma}$.

- *Observation space*: Usually this is a finite dimensional space, which would require appropriate sampling of the received signal. This was done in (2.39) under

the assumption, that the sampling theorem was met.

- *Estimation rule*: In this work, the ML solution will be sought

$$\hat{\underline{\lambda}} = \underset{\underline{\lambda}}{\operatorname{argmax}} \log \mathcal{L}(\underline{\lambda}) \; . \tag{3.3}$$

As mentioned above, the presence of the unknown data symbols $\underline{a}$ as a latent variable in the LLF (cf. (3.2)) leads to a sum of exponential terms inside the logarithm. Essentially because of this obstacle, the ML solution is hard to be obtained analytically. In Chapter 4 it will be explained in detail, how the EM algorithm can be utilized to overcome this problem by iteratively obtaining an ML estimate.

3.1 Classification of the Estimator

The estimator shall be classified by three features. First, the parameter space can be described by two different types of parameters. On the one hand, a random variable governed by an a priori probability distribution and on the other, an unknown, but deterministic quantity that constitutes the parameter. For the former case, a MAP estimation rule can be reasonable, if the distribution is known a priori. Although a priori PDF's were defined for the four scalar parameters of FF, CFO, CPO and TO in Section 2.3.2, the unknown parameter vector $\underline{\lambda}$ shall be treated as non-random and deterministic. Three of the four parameters were assumed to be uniformly distributed and therefore of limited usefulness in a MAP estimation rule. For this and the sake of generality, all parameters are treated as unknown and deterministic and hence the ML estimation rule shall be applied. Second, the dimension of the parameter $\underline{\lambda}$ is distinguished. If it is a scalar, one speaks of single estimation and in case of a vector of joint estimation. In this work, all four channel parameters from Section 2.3.2 shall be estimated jointly. And third, the unbiasedness of the estimator is considered. If,

on average, correct estimates are produced

$$\mathrm{E}\left[\hat{\underline{\lambda}}_i\right] = \underline{\lambda}_i \tag{3.4}$$

the estimator is called unbiased, which is in general a favorable property. As it will be shown in Chapter 5, the considered methods fulfill (3.4) in general.

The estimator derived in this work is thus classified as an unbiased, joint estimator of a non-random parameter vector and in the following the respective theoretical performance limits shall be discussed.

3.2 The Theoretic Performance Limit of Joint and Unbiased Estimators

A natural question that arises in the field of parameter estimation is the estimator's precision and its gap to the theoretical optimum. The measure that is typically referred to in this context is the Fisher information (FI) [58]. It expresses how much information an observation $r(kT_0)$ carries about an unknown parameter $\underline{\lambda}_i$ by calculating the expected negative curvature of the LLF $\log \mathcal{L}(\underline{\lambda}_k)$. In the case of the above classified estimator, the FI is stored in a symmetric matrix called the Fisher information matrix (FIM) $\underset{\sim}{I}(\underline{\lambda})$, which is expressed as the expectation over the partial derivatives of the LLF [58]

$$\left[\underset{\sim}{I}(\underline{\lambda})\right]_{k,l} = \mathrm{E}_{\underline{\boldsymbol{r}}}\left[\frac{\partial \log\left(\mathrm{p}_{\underline{\boldsymbol{r}}}\left[\underline{r}|\underline{\lambda}\right]\right)}{\partial \underline{\lambda}_k} \cdot \frac{\partial \log\left(\mathrm{p}_{\underline{\boldsymbol{r}}}\left[\underline{r}|\underline{\lambda}\right]\right)}{\partial \underline{\lambda}_l}\right] \tag{3.5}$$

$$= -\mathrm{E}_{\underline{\boldsymbol{r}}}\left[\frac{\partial^2 \log\left(\mathrm{p}_{\underline{\boldsymbol{r}}}\left[\underline{r}|\underline{\lambda}\right]\right)}{\partial \underline{\lambda}_k \partial \underline{\lambda}_l}\right] . \tag{3.6}$$

In the presumed case of properly filtered and sampled signal (cf. Section 2.4), no information gets lost and the noise properties are retained and hence the sampled version can be used without loss of generality. In this case of a joint estimation, the information about each parameter $\underline{\lambda}_i$ is stored in the i'th diagonal element $[\underset{\sim}{I}(\underline{\lambda})]_{i,i}$.

While the FI states directly how much information can be inferred from data obeying the the LLF, a more practical metric to evaluate an estimator's performance is the variance of the estimation error

$$\mathrm{Var}\left[\hat{\underline{\lambda}}_i - \underline{\lambda}_i\right] = \mathrm{E}\left[(\hat{\underline{\lambda}}_i - \underline{\lambda}_i)^2\right] - \mathrm{E}\left[\hat{\underline{\lambda}}_i - \underline{\lambda}_i\right]^2 \tag{3.7}$$

which can be measured directly at the output of an estimator. In the case of unbiasedness, the error variance is equal to the MSEE $\mathrm{E}\left[(\hat{\underline{\lambda}}_i - \underline{\lambda}_i)^2\right]$ because of (3.4). The variance is typically benchmarked against the well known Cramér-Rao bound (CRB) or the CRVB in joint estimation, respectively. The error variance of a jointly estimated $\underline{\lambda}_i$ is hereby lower bounded by the i'th diagonal element of the FIM's inverse

$$\mathrm{CRVB}(\underline{\lambda}_i) = \left[\underset{\sim}{I}^{-1}(\underline{\lambda})\right]_{i,i} \tag{3.8}$$

$$\leq \mathrm{Var}\left[\hat{\underline{\lambda}}_i - \underline{\lambda}_i\right] \ , \tag{3.9}$$

where the estimator is called efficient if $\mathrm{Var}\left[\hat{\underline{\lambda}}_i - \underline{\lambda}_i\right] = \mathrm{CRVB}(\underline{\lambda}_i)$. The non-diagonal elements describe dependencies of the parameters' joint estimation, e.g. $[\underset{\sim}{I}(\underline{\lambda})]_{k,l} \neq 0$ for $k \neq l$ states a coupling in the joint estimation of $\underline{\lambda}_k$ and $\underline{\lambda}_l$ (cf. Section 3.3). So generally non-diagonal elements are desirable to be zero. The scalar counter part to the CRVB, which is called the CRB, will never exceed the vector version

$$\mathrm{CRB}(\underline{\lambda}_i) = \left[\underset{\sim}{I}^{-1}(\underline{\lambda})\right]_{i,i} \tag{3.10}$$

$$\leq \mathrm{CRVB}(\underline{\lambda}_i) \ , \tag{3.11}$$

since the estimation of $\underline{\lambda}_i$ under the knowledge of every other parameter of $\underline{\lambda}$ is at least as accurate as if the other parameters are unknown.

3.3 The Modified Fisher Information Matrix

For the same reasons that prevent (3.2) being analytically tractable, if the LLF is dependent on a latent variable (i.e. the random and unknown data sequence $\underline{a}$), the

true FIM is generally not tractable, either. In [59] and [60] this problem was overcome by defining an analytically tractable information matrix for that case which is called modified Fisher information matrix (MFIM). Here, not the LLF $\log\left(\mathrm{p}_{\underline{\boldsymbol{r}}}\left[\underline{r}|\underline{\lambda}\right]\right)$ must be differentiated, but the simpler structured posterior measure $\log\left(\mathrm{p}_{\underline{\boldsymbol{r}}|\underline{\boldsymbol{a}}}\left[\underline{r}|\underline{a},\underline{\lambda}\right]\right)$, which boils down to an exponential function in the case of an AWGN channel. Furthermore, the expectation is now taken with regard to the simpler structured posterior distribution and thus the MFIM computes as

$$\left[\underset{\sim}{I}_M(\underline{\lambda})\right]_{k,l} = \mathrm{E}_{\underline{\boldsymbol{r}},\underline{\boldsymbol{a}}}\left[\frac{\partial \log\left(\mathrm{p}_{\underline{\boldsymbol{r}}|\underline{\boldsymbol{a}}}\left[\underline{r}|\underline{a},\underline{\lambda}\right]\right)}{\partial \underline{\lambda}_k} \cdot \frac{\partial \log\left(\mathrm{p}_{\underline{\boldsymbol{r}}|\underline{\boldsymbol{a}}}\left[\underline{r}|\underline{a},\underline{\lambda}\right]\right)}{\partial \underline{\lambda}_l}\right] \tag{3.12}$$

$$= -\mathrm{E}_{\underline{\boldsymbol{r}},\underline{\boldsymbol{a}}}\left[\frac{\partial^2 \log\left(\mathrm{p}_{\underline{\boldsymbol{r}}|\underline{\boldsymbol{a}}}\left[\underline{r}|\underline{a},\underline{\lambda}\right]\right)}{\partial \underline{\lambda}_k \partial \underline{\lambda}_l}\right] . \tag{3.13}$$

The derivatives are computed with regard to the true $\underline{\lambda}$. Similar to (3.8), a bound will be derived from the MFIM in the next section that is a lower bound to the true CRVB. In case the received signal is superposed by noise as specified in (2.41) to (2.44), the MFIM from (3.12) can be simplified to

$$\left[\underset{\sim}{I}_M(\underline{\lambda})\right]_{k,l} = \frac{2T_0}{N_0}\sum_{k=0}^{NK-1}\mathrm{E}_{\underline{\boldsymbol{a}}}\left[\mathrm{Re}\left\{\frac{\partial r(kT_0,\underline{a},\underline{\lambda})}{\partial \underline{\lambda}_k} \cdot \frac{\partial r^*(kT_0,\underline{a},\underline{\lambda})}{\partial \underline{\lambda}_l}\right\}\right] \tag{3.14}$$

with

$$\begin{aligned} r(kT_0,\underline{a},\underline{\lambda}) = \Big[&\alpha e^{j(2\pi\nu(kT_0-t_0)+\theta)} \cdot \sqrt{E_\mathrm{S}/T}\, e^{j2\pi h\sum_{n=-(L-1)}^{N-1} a_n q(kT_0-\tau-nT)} \\ &\cdot e^{j2\pi\beta\xi(kT_0-\tau)} + w(t)\Big] \cdot e^{j2\pi(1-\beta)\xi kT_0} , \end{aligned} \tag{3.15}$$

where the definitions of (2.7), (2.8), (2.12), (2.13) and (2.31) are applied. It is noted, that first, the received signal is considered here after the sampling process. Second, the signal is either tilted in the transmitter ($\beta = 1$) when the tilt is subject to the channel delay τ or in the receiver ($\beta = 0$). Third, the time reference in the CFO term is set to half the burst length ($t_0 = NT/2$) in order to obtain a symmetric observation window which will lead to a minimal CPO bound [60]. In a practical view, this is just a little trick in the receiver to set the time reference as preferred. Carrier phase

and frequency are linked by t_0 and the above stated value enables a minimal CPO estimation variance.

The derivation of the MFIM and the MCRVB for CPM were principally carried out in [61] for FF, CFO, CPO and TO. However, these bounds do neither account for a shifted observation window due to the TO, nor for a transmitter tilted CPM signal, which shall be investigated in this work. It is also noted, that the unknown amplitude of the transmitted signal is estimated in [61], whereas in this work the fading factor, i.e. the amplitude of the received signal, is considered. This leads to a difference in the signal-to-noise power ratio (SNR) definition, because of the direct coupling of E_S with the signal's amplitude (cf. (2.7)). The derivation necessary for this work's bounds is carried out in Appendix B.1 with the following MFIM as result

$$\underset{\sim}{I}_M(\underline{\lambda}) = 2\alpha^2 \frac{E_\mathrm{S}}{N_0} N \times \begin{bmatrix} \frac{1}{\alpha^2} & 0 & 0 & 0 \\ 0 & \frac{4}{3}\pi^2\left(\frac{1}{4}N^2T^2 + 3\tau^2\right) & 2\pi\tau & -2\beta\pi^2 h \frac{M-1}{T}\tau \\ 0 & 2\pi\tau & 1 & -\beta\pi h \frac{M-1}{T} \\ 0 & -2\beta\pi^2 h \frac{M-1}{T}\tau & -\beta\pi h \frac{M-1}{T} & \pi^2 h^2\left(4G_2(0)T\frac{M^2-1}{3T^2} + \beta\frac{(M-1)^2}{T^2}\right) \end{bmatrix} \tag{3.16}$$

in which several CPM parameters are used (cf. Table 5.1). The term $G_2(0)T$ describes the normalized frequency pulse energy (cf. (2.5)).

As in the case of the FIM, non-diagonal elements different from zero in the MFIM (cf. (3.16)) stand for a coupling between channel parameters, e.g. $[\underset{\sim}{I}_M(\underline{\lambda})]_{k,l} \neq 0$ indicates, that small deviations of $\hat{\underline{\lambda}}_k$ from its true value lead to a skewed estimation of $\hat{\underline{\lambda}}_l$ and vice versa. Based on this, the following conclusions can be drawn from the MFIM:

- For the MFIM's first element, the FF, no coupling is present, i.e. the first row and the first column consist of zeros exclusively besides the diagonal element. Thus it is assumed, that a reasonably reliable estimate $\hat{\alpha}$ can be obtained without further synchronization of the other parameters.

- As can be seen from the according non-diagonal elements of $\underline{\lambda}_2 = \nu$ and $\underline{\lambda}_3 = \theta$, the coupling depends on the actual value of the TO. For $\tau = 0$, a perfect symmetric observation window is achieved and the coupling disappears, which is in accordance to the observation window contemplations in e.g. [60,62]. If the time reference of the CFO t_0 was set to zero, the coupling would be represented by a much larger term. An intuitive explanation for this behavior is, that in general for a symmetric window the CFO induced phase is known be zero in the middle of the window and have an arbitrary value ϕ_1 in the start and the negative $-\phi_1$ in the end. In contrast of a non-symmetric window only the start phase $\phi_1 = 0$ is known. The consequence of the absence of symmetry in the observation window is a negative impact on the CPO estimation whereas the CFO estimation remains unaffected.

- The TO is coupled to both the CFO and the CPO according to the MFIM for the case of transmitter tilting ($\beta = 1$). Since the tilt (i.e. a known frequency offset) is subject to the unknown TO τ, an unknown phase offset is introduced by this. As in the above case of an asymmetric observation window, this is assumed to have a negative impact on the CPO estimation. Since the tilt introduces a constant, but known frequency offset, the CFO estimation again remains unaffected. It can be observed, that the coupling between τ and ν is described by the value $[\underline{I}_M(\underline{\lambda})]_{2,4} = -2\beta\pi^2 h\frac{M-1}{T}\tau$ and that this value is the product of the CPO's couplings to CFO and TO $[\underline{I}_M(\underline{\lambda})]_{2,3} \cdot [\underline{I}_M(\underline{\lambda})]_{3,4} = [\underline{I}_M(\underline{\lambda})]_{2,4}$. This fact suggests, that the three parameters CFO, CPO and TO are not mutually coupled (as the structure of the MFIM would imply) but only the CPO is coupled to the other two and $[\underline{I}_M(\underline{\lambda})]_{2,4}$ represents the link through the CPO.

- In case of receiver tilting ($\beta = 0$), the basically symmetric observation window ($t_0 = NT/2$, ignoring τ) gives an essentially diagonal matrix, which means that every parameter can principally be estimated independently if the other parameter estimates lie close enough to according their true value. This is in general a favorable property, which is why it was opted to only consider this

specific time reference t_0.

3.4 The Modified Cramér-Rao Vector Bound

The resulting bounds from the MFIM are called MCRVBs and computed the same way as in (3.8) as the diagonal elements of the MFIM's inverse

$$\mathrm{MCRVB}(\underline{\lambda}_i) = \left[\underset{\sim}{I}_M^{-1}(\underline{\lambda})\right]_{i,i} \ . \tag{3.17}$$

Inserting (3.16) into (3.17) eventually gives the following bounds

$$\mathrm{MCRVB}(\alpha) = \frac{1}{2N} \cdot \frac{1}{E_\mathrm{S}/N_0} \ , \tag{3.18}$$

$$\mathrm{MCRVB}(\nu) = \frac{3}{2\pi^2 N^3 T^2} \cdot \frac{1}{\alpha^2 E_\mathrm{S}/N_0} \ , \tag{3.19}$$

$$\mathrm{MCRVB}(\theta) = \left[\frac{1}{2N} + \frac{6\tau^2}{N^3T^2} + \beta\frac{3}{8NG_2(0)T} \cdot \frac{M-1}{M+1}\right] \cdot \frac{1}{\alpha^2 E_\mathrm{S}/N_0} \ , \tag{3.20}$$

$$\mathrm{MCRVB}(\tau) = \frac{3T^2}{8\pi^2 N h^2 \left(M^2-1\right) G_2(0) T} \cdot \frac{1}{\alpha^2 E_\mathrm{S}/N_0} \ . \tag{3.21}$$

While all bounds are indirectly proportional to E_S/N_0, the last three are also indirectly proportional to the squared FF α^2. While the first bound only depends on the noise power, the others rise accordingly as less signal energy is received with decreasing FF. It is noted, that all four bounds are valid only if the data sequence $\underline{a}$ consists of i.i.d. data symbols a_k. This assumption is used throughout this section.

3.4.1 Comparison of the MCRVB with Its Scalar Version

First, the vector bounds are compared to their scalar counter parts called modified Cramér-Rao bound (MCRB), which are computed as the inverse terms of the diagonal elements of the MFIM (cf. (3.17))

$$\mathrm{MCRB}(\underline{\lambda}_i) = \left[\underset{\sim}{I}_M(\underline{\lambda})\right]_{i,i}^{-1} \ . \tag{3.22}$$

The MCRBs lower bound the error variance for single (in contrast to joint) parameter estimation of an unknown parameter under the knowledge of all other channel parameters. The following observations can be made regarding the scalar MCRBs:

- The fading factor bound is equal in the scalar and vector case

$$\mathrm{MCRVB}(\alpha) \equiv \mathrm{MCRB}(\alpha) \,, \tag{3.23}$$

  which reflects the obvious decoupled nature of the FF parameter.

- Although the FF is assumed to be known in the scalar bounds, it still influences the estimation performance and thus is present in the scalar bounds of carrier frequency and phase as well as the timing.

- For the case of CFO ν, it was presumed that the coupling terms in (3.16) only state an indirect coupling to the TO τ through the CPO θ. This statement is directly confirmed by the equivalence of the corresponding scalar and vector bound

$$\mathrm{MCRVB}(\nu) \equiv \mathrm{MCRB}(\nu) \,. \tag{3.24}$$

  Furthermore, the direct coupling to the CPO does not negatively affect the performance limit in the joint estimation.

- The scalar bound of the CPO is given through (3.22) as

$$\mathrm{MCRB}(\theta) = \frac{1}{2N} \cdot \frac{1}{\alpha^2 E_\mathrm{S}/N_0} \,. \tag{3.25}$$

  Clearly it is not affected by the induced phase offsets of the channel delay and the therefore delayed tilting, because the TO responsible for both cases is known. The vector bound can be expressed as

$$\mathrm{MCRVB}(\theta) = \mathrm{MCRB}(\theta) + \mathrm{PEN}(\tau) + \mathrm{PEN}(\beta) \,, \tag{3.26}$$

  where $\mathrm{PEN}(\tau) = \frac{6\tau^2}{N^3T^2} \cdot \frac{1}{\alpha^2 E_\mathrm{S}/N_0}$ and $\mathrm{PEN}(\beta) = \beta \frac{3}{8NG_2(0)T} \cdot \frac{M-1}{M+1} \cdot \frac{1}{\alpha^2 E_\mathrm{S}/N_0}$ denote

penalties on the vector bound in comparison with the scalar bound. They are directly related to $[\underline{\underline{I}}_M(\underline{\lambda})]_{2,3}$ and $[\underline{\underline{I}}_M(\underline{\lambda})]_{3,4}$, respectively. While PEN(β) is a design choice, PEN(τ) is unavoidable in joint estimation. Averaging PEN(τ) with respect to τ (cf. (2.36)), the quotient PEN(τ)/MCRB(θ) $= N^{-2}$ is nearly zero (cf. Appendix B.2) though when the value from Table 5.1 is inserted for N. Thus the relation between the scalar and the vector bound is

$$\mathrm{MCRVB}(\theta) \begin{cases} \approx \mathrm{MCRB}(\theta) & \beta = 0\ , \\ > \mathrm{MCRB}(\theta) & \beta = 1 \end{cases} \tag{3.27}$$

with the first line being valid in very good approximation.

- The scalar TO bound calculates to

$$\mathrm{MCRB}(\tau) = \frac{3T^2}{8\pi^2 N h^2 \left[(M^2-1)\, G_2(0) T + 3\beta(M-1)^2\right]} \cdot \frac{1}{\alpha^2 E_\mathrm{S}/N_0} \tag{3.28}$$

and it can be observed that the transmitter tilting has a positive influence on the scalar estimation performance through the term $3\beta(M-1)^2$. This influence however only holds up if the other channel parameters are known as it vanishes in the according vector bound in (3.21). Therefore, the vector bound is independent from the tilting process whereas the scalar bound is not

$$\mathrm{MCRVB}(\tau) \begin{cases} = \mathrm{MCRB}(\tau) & \beta = 0\ , \\ > \mathrm{MCRB}(\tau) & \beta = 1\ . \end{cases} \tag{3.29}$$

In other words this means, that if no perfect information about the CPO is available, the transmitter tilting brings no advantages (in TO estimation) and additionally causes disadvantages (in CPO estimation). In [63], the principal coupling of CPO and TO through an asymmetric spectrum was described, although the specific effect in case of tilted CPM was not found in the literature.

It is for the last findings, that the tilting is realized exclusively in the receiver (i.e.

$\beta = 0$) throughout the rest of this work. In the last paragraph of Section 2.2.1, this decision was anticipated and all definitions from there on were already adjusted to it.

3.4.2 Comparison of the MCRVB with True Bounds

It is of natural interest how the MCRVB compares to the true CRVB and [60] proved the general validity of

$$\mathrm{MCRVB}(\underline{\lambda}_i) \leq \mathrm{CRVB}(\underline{\lambda}_i) \,, \tag{3.30}$$

which gives an uncertainty whether the estimator is efficient when not reaching the MCRVB. In [62], another scalar bound called the asymptotic Cramér-Rao bound (ACRB) is derived, which approaches the true CRB for high SNR values. The problem stated before, that a sum of exponential terms prevents an easy analytical expression of the true CRB is circumvented by identifying, that this sum has only one dominant addend for high values of E_S/N_0 and consequently only taking this one term into account. It is also shown there, that

$$\mathrm{MCRB}(\underline{\lambda}_i) = \mathrm{ACRB}(\underline{\lambda}_i) \quad \text{, if } \underline{a} \text{ is discrete.} \tag{3.31}$$

In the case of the considered digital communication system, above equation is therefore valid. Furthermore, [62] also shows that in joint estimation the ACRB equals the real vector bound if the FIM $\underset{\sim}{I}$ is not a function of the transmitted sequence $\underline{a}$ and $\underline{a}$ is known to the receiver

$$\mathrm{ACRB}(\underline{\lambda}_i) = \mathrm{CRVB}(\underline{\lambda}_i|\underline{a}) \quad \text{, if } \underset{\sim}{I} \neq f(\underline{a}) \tag{3.32}$$

and is less tight if the matrix depends on $\underline{a}$

$$\mathrm{ACRB}(\underline{\lambda}_i) \leq \mathrm{CRVB}(\underline{\lambda}_i|\underline{a}) \quad \text{, if } \underset{\sim}{I} = f(\underline{a}) \,. \tag{3.33}$$

The receiver's knowledge is thereby indicated by the additional argument in the bound, i.e. $\mathrm{CRVB}(\underline{\lambda}_i|\underline{a})$ denotes the true CRVB in the case of transmitting a known

pilot sequence.

In [64], it was derived under the assumption of an a priori known $\underline{a}$, that data dependence in the FIM $\underset{\sim}{I}$ is only introduced by the timing (τ) and not by frequency (ν) or phase (θ) estimation. This means, that an FIM for the unknown parameters ν and θ fulfills (3.32) (under the knowledge of α and τ). The MCRVBs derived in that case are equal to (3.19) and (3.20), respectively. Combining (3.24), (3.27), (3.31) and (3.32) allows for

$$\mathrm{MCRVB}(\nu) = \mathrm{CRVB}(\nu|\underline{a}) \;, \tag{3.34}$$

$$\mathrm{MCRVB}(\theta) = \mathrm{CRVB}(\theta|\underline{a}) \tag{3.35}$$

to be valid in the joint estimation case considered in this work.

For α, the true CRB (under knowledge of the transmitted sequence $\underline{a}$) can be derived analytically (cf. Appendix B.3) and writes as

$$\mathrm{CRB}(\alpha|\underline{a}) = \frac{1}{2N} \cdot \frac{1}{E_\mathrm{S}/N_0} = \mathrm{MCRVB}(\alpha) \;. \tag{3.36}$$

In the calculation of the MCRVB, no knowledge on the data $\underline{a}$ is assumed, which is why this bound is generally used for blind estimators that include NDA and DD methods. In the case of a DD estimator and the presence of data in the transmission, estimates on that data can be made available in the receiver, e.g. through a feedback loop. Thus the modified bounds with and without i.i.d. data knowledge have the relation [65]

$$\mathrm{MCRVB}(\underline{\lambda}_i|\underline{a}) \leq \mathrm{MCRVB}(\underline{\lambda}_i) \;. \tag{3.37}$$

with the equality being valid at sufficiently high SNR. Collecting the implications of (3.11), (3.30) and (3.37) the derived modified bound from (3.18) equals the true vector bound

$$\mathrm{MCRVB}(\alpha) = \mathrm{CRVB}(\alpha|\underline{a}) \;. \tag{3.38}$$

In case of the remaining parameter TO, the line of argument as for the CFO and CPO is not possible due to the FIM's dependence on the data. However, a formulation

of the true scalar CRB under the knowledge of the random data sequence is given by [63] as

$$\mathrm{CRB}(\tau|\underline{a}) = \frac{\int_{-\infty}^{+\infty}(2\pi f)^2 S_{ss}(f)\,\mathrm{d}f}{\int_{-\infty}^{+\infty} S_{ss}(f)\,\mathrm{d}f} \tag{3.39}$$

in which the bound is only dependent on the characteristics of the power spectral density (PSD) $S_{ss}(f)$ that can be calculated in the CPM case according to [43, 9.3.3]. The denominator holds the power of the received signal, which is known to be $\alpha^2 E_\mathrm{S}/T$ and the nominator is referred to as the normalized moment of inertia in [63]. According to [65], it can be expressed as

$$\begin{aligned}\int_{-\infty}^{+\infty}(2\pi f)^2 S_{ss}(f)\,\mathrm{d}f &= \frac{4\pi^2 h^2}{T}\cdot \mathrm{E}[a_k^2]\cdot\int_{-\infty}^{+\infty}|G(f)|^2\,\mathrm{d}f \\ &= \frac{4\pi^2 h^2}{T}\cdot\frac{M^2-1}{3}G_2(0)\ ,\end{aligned} \tag{3.40}$$

where for the trivial computation of the expected value of the squared CPM symbols, (2.3) is recalled. By inserting the expressions of power and moment of inertia into (3.39), the relation $\mathrm{MCRVB}(\tau) = \mathrm{CRB}(\tau|\underline{a})$ is verified. The same reasoning as in the FF case leads now to the equality

$$\mathrm{MCRVB}(\tau) = \mathrm{CRVB}(\tau|\underline{a})\ . \tag{3.41}$$

Figure 3-2 visualizes the derivations that can be summarized as

$$\mathrm{MCRVB}(\underline{\lambda}_i) = \mathrm{CRVB}(\underline{\lambda}_i|\underline{a})\ . \tag{3.42}$$

Based on the conclusions above, it should be possible for a DD estimator to be MCRVB efficient at SNRs sufficiently high for the detector to principally provide reliable estimates.

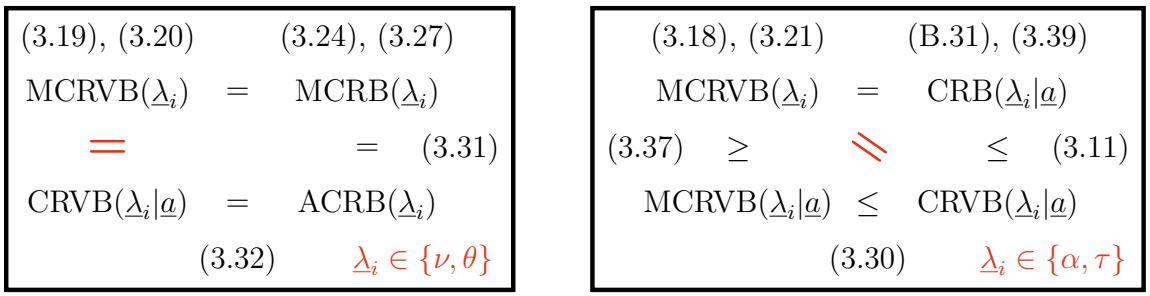


Figure 3-2: Relations between different lower bounds. The left part shows the relations of lower bounds for the CFO ν and the CPO θ, while the right part does it for the FF α and the TO τ. The corresponding equations are given for the bounds and the relations. The red equal sign in bold font is the desired conclusion MCRVB($\underline{\lambda}_i$) = CRVB($\underline{\lambda}_i|\underline{a}$).

3.5 Key Points of the Chapter

- The modified Fisher information matrix for the considered burst type CPM transmission was derived. The coupling terms were discussed and it was concluded, that the matrix is quasi diagonal when the tilting is realized in the receiver.

- Modified Cramér-Rao vector bounds were formulated for the four considered channel parameters and compared to the scalar modified Cramér-Rao bounds. The effect of transmitter tilting was investigated with regard to the bounds. While being advantageous in the case of scalar TO estimation, the advantage vanishes in the joint case with the side effect of a deteriorated CPO estimation.

- The MCRVBs were compared to the asymptotic and true Cramér-Rao bounds and it was shown that for all four cases, that the modified vector bounds are equal to the true vector bounds under the assumption of known data. This result was achieved by gathering relations between several bounds from numerous publications.

- In DD estimators the knowledge on the transmitted sequence becomes available at some point if the SNR is high enough. Thus the true bound in this case is given by CRVB($\underline{\lambda}_i|\underline{a}$), for which a closed form expression could be derived by showing the equality to MCRVB($\underline{\lambda}_i$). This analytically derived true bound will serve as a suited benchmark in Chapter 5.

This Page Intentionally Left Blank

Chapter 4

Channel Parameter Estimation

In the case of CPM signals transmitted in short bursts, several channel parameter must be estimated and corrected to enable a successful detection of the data symbols. The parameters to be estimated in the SE channel were introduced in Section 2.3 as FF α, CFO ν, CPO θ, TO τ. The estimation of e.g. time variant, non-frequency flat channel conditions as well as sample clock rate and phase noise is not subject to this work.

In this chapter, the ML solution of the parameter vector $\underline{\lambda} = [\alpha, \nu, \theta, \tau]$ is sought

$$\hat{\underline{\lambda}} = \underset{\underline{\lambda}}{\operatorname{argmax}} \log \left(\sum_{\underline{a}} \mathrm{p}_{\underline{\boldsymbol{r}},\underline{\boldsymbol{a}}} \left[\underline{r}, \underline{a}' | \underline{\lambda} \right] \right) \tag{4.1}$$

$$= \underset{\underline{\lambda}}{\operatorname{argmax}} \log \left(\sum_{\underline{\gamma}} \mathrm{p}_{\underline{\boldsymbol{r}}',\underline{\boldsymbol{\gamma}}} \left[\underline{r}', \underline{\gamma} | \underline{\lambda} \right] \right) \ , \tag{4.2}$$

where (3.2) was inserted in (3.3) in the first line. In the second line, the data symbols $\underline{a}$ were exchanged with the pseudo symbols $\underline{\gamma}$ and the received signal was exchanged with the tilted version. Both exchanges are deterministic (given proper trellis initialization) and reversible. As it was already outlined in the beginning of Chapter 3, the LLF is not analytically tractable and thus no closed form ML solution can be formulated. An estimator will be derived and thoroughly discussed that obtains the ML solution iteratively by means of the EM algorithm.

The main ideas of Sections 4.2 to 4.5 were published in [66], though are described

in much more detail in this thesis. Numerous subsections were added and several new aspects are highlighted, so that this chapter constitutes a substantial extension over [66].

4.1 Introduction to the Expectation Maximization Algorithm

The EM algorithm is a method to obtain parameter estimates, when some random variables (or an arbitrary subset of these) cannot be observed directly and can be applied to any probabilistic model in general [67,68]. It does so by providing a framework for iteratively approaching the ML solution with an often acceptable computational effort even (and especially) when maximizing the LLF directly is infeasible. A typical use cases is e.g. the estimation of Gaussian mixture models' parameters, where the Baum-Welch algorithm is usually applied, which constitutes the EM solution to this problem [69]. Another EM application is the parameter estimation of hidden Markov model (HMM)s and as it was introduced in Section 2.2, the data transmission with CPM can be interpreted as such. Any HMM is specified by the following parameters [70]:

- The number of states in the model is given as $N_{\underline{S}'} = Q \cdot M^{L-1}$ in Section 2.2.2.
- The observations per state, which is the sampled and tilted received signal $\underline{r}'\left[nK : (n+1)K - 1\right]$ in the according time slot.
- The hidden state transition probabilities $\mathrm{P}_{\underline{\boldsymbol{S}}_n, \underline{\boldsymbol{S}}_{n+1}}\left[\underline{s}, \dot{\underline{s}}\right]$, which are sufficiently defined by the CPM trellis. In case a transition is allowed, the corresponding probability is equal to $1/M$, since the data symbols a'_n are i.i.d., otherwise the transition probability is zero. If a channel code is present in the system, its information can also be incorporated as explained in Section 2.6.
- The observation probability distribution in a given state, which is related to

the path metric in (2.51) by

$$\mathrm{p}_{\underline{\boldsymbol{r}}'|\underline{\boldsymbol{S}}'_n}\left[\underline{r}'\,[nK:(n+1)K-1]\,|\underline{s},\hat{\underline{\lambda}}\right] = \sum_{\dot{\underline{s}}} f_n^M(\underline{s},\dot{\underline{s}})\ . \tag{4.3}$$

It is noted, that the process $\underline{\boldsymbol{r}}'$ is denoted in distributions for the sake of readability although only parts of the received vector $\underline{r}'$ are subject to the distribution.

- The initial state probability distribution, which is given as the forward state occupations of the first stage $f_0^F(\underline{s})$ and the backward state occupations $f_{N+1}^B(\dot{\underline{s}})$ (CZS in the CPM trellis) as specified in Section 2.5.2.

With the underlying model now outlined, the rest of this section gives a general treatment of the EM algorithm (while having the objective in mind and using the according notation). The basic steps of the EM algorithm are given as follows. The bracketed superscript i represents the iteration counter and starts with $i = 1$:

1. An initial parameter vector $\underline{\lambda}_{\mathrm{old}}^{(1)}$ is chosen. The selection of the initial parameter set is discussed in detail in Section 4.4.

2. The expectation step (E-Step) is carried out, which generally comprises the definition of the auxiliary function

$$\mathcal{Q}(\underline{\lambda},\underline{\lambda}_{\mathrm{old}}^{(i)}) = \sum_{\underline{\gamma}} \mathrm{P}_{\underline{\boldsymbol{\gamma}}|\underline{\boldsymbol{r}}'}\left[\underline{\gamma}|\underline{r}',\underline{\lambda}_{\mathrm{old}}^{(i)}\right]\cdot\log\left(\mathrm{p}_{\underline{\boldsymbol{\gamma}},\underline{\boldsymbol{r}}'}\left[\underline{\gamma},\underline{r}'|\underline{\lambda}\right]\right)\ . \tag{4.4}$$

The auxiliary function states the expectation of the log likelihood (LL) of the complete data (observed data $\underline{r}'$ and the latent data $\underline{\gamma}$) with respect to the posterior probabilities $\mathrm{P}_{\underline{\boldsymbol{\gamma}}|\underline{\boldsymbol{r}}'}\left[\underline{\gamma}|\underline{r}',\underline{\lambda}_{\mathrm{old}}^{(i)}\right]$ of the discrete, latent, random variable $\underline{\boldsymbol{\gamma}}$. The logarithm and the density's exponential can be reduced, which makes the auxiliary function in many cases analytically tractable unlike to the LLF itself (cf. (4.2)). In the E-Step, the posterior probabilities must be computed, whereas the explicit computation of the auxiliary function is generally not necessary. The expectation step for the channel parameter estimation is discussed in Section 4.2.

3. In the maximization step (M-Step), the parameter set $\hat{\underline{\lambda}}$ is sought that maximizes the auxiliary function $\mathcal{Q}(\underline{\lambda}, \underline{\lambda}_{\text{old}}^{(i)})$

$$\hat{\underline{\lambda}}_{\text{ML}}^{(i)} = \underset{\underline{\lambda}}{\text{argmax}} \ \mathcal{Q}(\underline{\lambda}, \underline{\lambda}_{\text{old}}^{(i)}) \tag{4.5}$$

while leaving $\underline{\lambda}_{\text{old}}^{(i)}$ fixed to its current value. This step is covered in Section 4.3 regarding the channel parameter estimation. While this M-Step will eventually lead to the ML solution and is used in this thesis, an alternative M-Step shall be given for the sake of completeness

$$\hat{\underline{\lambda}}_{\text{MAP}}^{(i)} = \underset{\underline{\lambda}}{\text{argmax}} \ \mathcal{Q}(\underline{\lambda}, \underline{\lambda}_{\text{old}}^{(i)}) + \log \text{p}_{\underline{\boldsymbol{\lambda}}}(\underline{\lambda}) \ . \tag{4.6}$$

When suitable a priori information about the parameters are available, the MAP estimate can be obtained by this slight modification of the M-Step with no changes necessary in the E-Step [68]. As it was already discussed in the beginning of Chapter 3, this thesis focuses on the ML solution and thus only (4.5) is considered in the remainder with the subscript ML at the estimate $\hat{\underline{\lambda}}^{(i)}$ being dropped.

4. It is checked whether a convergence criterion was reached. If so, the algorithm is finished and the final estimate is given as

$$\hat{\underline{\lambda}} = \hat{\underline{\lambda}}^{(i)} \ . \tag{4.7}$$

Possible convergence criteria are e.g. a sufficiently small difference between a new and old estimate $\left| \hat{\underline{\lambda}}_k^{(i)} - \hat{\underline{\lambda}}_k^{(i-1)} \right|$ or reaching a fixed number of iterations N_{It}. Below in this section, a further (more elegant) criterion is discussed and in Section 5.2, the EM convergence is investigated for the considered transmission system.

5. If no convergence was reached in Step 4, the parameter vector for the next iteration $\underline{\lambda}_{\text{old}}^{(i+1)} = \hat{\underline{\lambda}}^{(i)}$ is updated, i is incremented and the algorithm returns to

step 2.

It is proved e.g. in [69] by applying Jensen's inequality on the concave nature of the logarithmic function, that

$$\log \mathcal{L}(\hat{\underline{\lambda}}^{(i)}) \geq \log \mathcal{L}(\hat{\underline{\lambda}}^{(i)}_{\text{old}}) + \left[\mathcal{Q}\left(\hat{\underline{\lambda}}^{(i)}, \underline{\lambda}^{(i)}_{\text{old}}\right) - \mathcal{Q}\left(\underline{\lambda}^{(i)}_{\text{old}}, \underline{\lambda}^{(i)}_{\text{old}}\right)\right] \tag{4.8}$$

$$\geq \log \mathcal{L}(\underline{\lambda}^{(i)}_{\text{old}}) \; . \tag{4.9}$$

The term in square brackets is non-negative because of (4.5) and thus the second line directly states, that a new estimate $\hat{\underline{\lambda}}$ is a better than (or equal to) $\underline{\lambda}_{\text{old}}$ in the ML sense. The term on the right hand side of (4.8) is called variational lower bound and the difference between this bound and the LLF can be shown to be the Kullback-Leibler divergence of the old and new estimates' posterior PMF of the latent data (cf. Appendix C.1)

$$\text{KL}\left(\text{P}_{\underline{\lambda}^{(i)}_{\text{old}}} \parallel \text{P}_{\hat{\underline{\lambda}}^{(i)}}\right) = \sum_{\underline{\gamma}} \text{P}_{\underline{\gamma}|\underline{r}'}\left[\underline{\gamma}|\underline{r}', \underline{\lambda}_{\text{old}}\right] \log\left(\frac{\text{P}_{\underline{\gamma}|\underline{r}'}\left[\underline{\gamma}|\underline{r}', \underline{\lambda}_{\text{old}}\right]}{\text{P}_{\underline{\gamma}|\underline{r}'}\left[\underline{\gamma}|\underline{r}', \hat{\underline{\lambda}}\right]}\right) \tag{4.10}$$

$$= \left[\log \mathcal{L}(\hat{\underline{\lambda}}^{(i)}) - \log \mathcal{L}(\underline{\lambda}^{(i)}_{\text{old}})\right] - \left[\mathcal{Q}\left(\hat{\underline{\lambda}}^{(i)}, \underline{\lambda}^{(i)}_{\text{old}}\right) - \mathcal{Q}\left(\underline{\lambda}^{(i)}_{\text{old}}, \underline{\lambda}^{(i)}_{\text{old}}\right)\right] \; . \tag{4.11}$$

The probabilities' notations in the divergence are hereby shortened for the sake of brevity. All terms in the second line are computable through (2.57) and (4.27). Since the divergence expresses how similar the posterior distributions of two estimates are, this metric is an elegant way to determine the progress of each EM iteration. In Section 5.2, the divergence between the current and the (converged) ML estimate $\text{KL}\left(\text{P}_{\hat{\underline{\lambda}}^{(i)}} \parallel \text{P}_{\underline{\lambda}_{\text{ML}}}\right)$ is used to evaluate the convergence at different SNRs. In Fig. 4-1, an exemplary, normalized CPO LLF $\mathcal{L}(\theta)$ of a binary, partial response, rectangular CPM scheme is given. Furthermore the auxiliary functions of the first three EM iterations are shown as well as the distance between two LLF peaks $d(\theta)$, which is derived in Section 4.4.1.

Several noteworthy points are displayed in this figure:

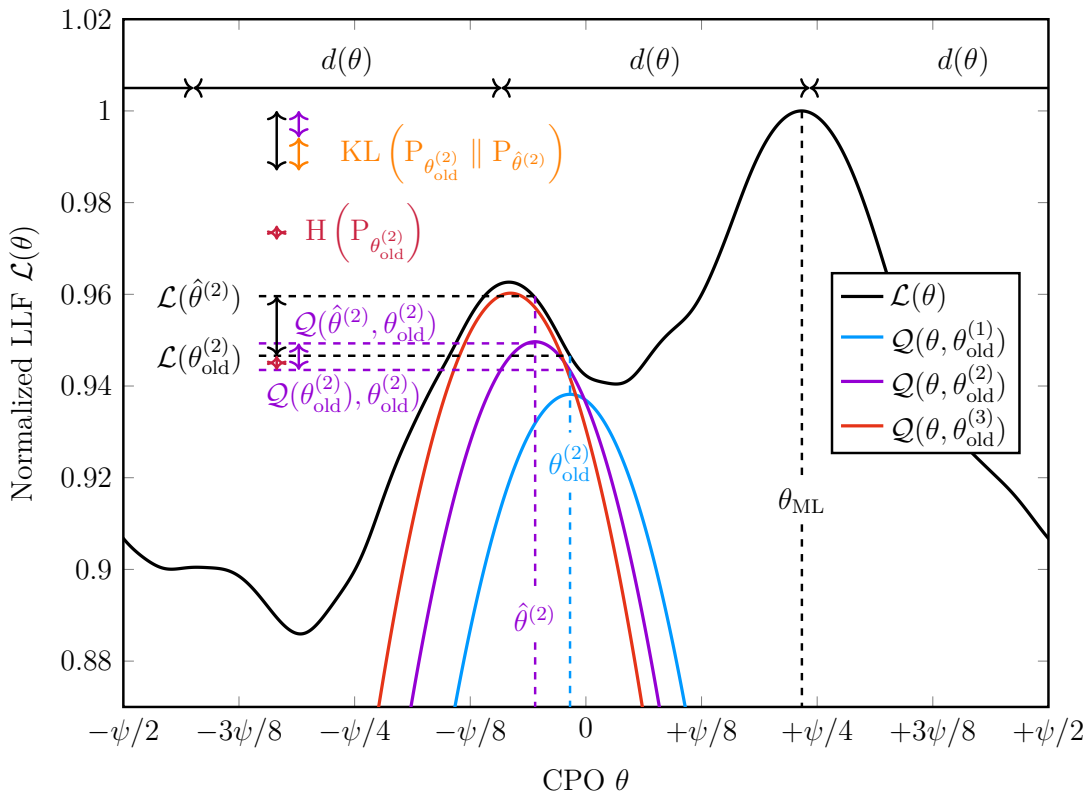


Figure 4-1: LLF $\mathcal{L}(\theta)$ of an $h = 1/3$, $M = 2$, $L = 3$ rectangular CPM scheme. All shown curves are normalized with respect to the maximum of $\mathcal{L}(\theta)$. Three iterations of auxiliary functions $\mathcal{Q}(\theta, \theta^{(i)}_{\mathrm{old}})$ are shown. The start value of the EM algorithm is $\theta^{0}_{\mathrm{old}} = 0$ and the ML solution is $\theta_{\mathrm{ML}} = 0.489$. Relevant variables of the second iteration are indicated by the dashed, horizontal lines. The distance between the LLF's local and global maximum is given by $d(\theta)$ (cf. Section 4.4.1). Furthermore, the values of the Kullback-Leibler divergence (cf. (4.11)) and entropy (cf. (4.12)) are displayed by the double-headed arrows.

- The LLF is periodic with the phase state width $\psi = 2\pi/Q$, which is in line with the postulation in Section 2.5.2, that the carrier phase must only be locked in any arbitrary phase state due to the rotational invariance. The periodicity is only suggested though by the values $\mathcal{L}(-\psi/2) = \mathcal{L}(+\psi/2)$ in this figure, but would be obvious in a wider CPO range which was discarded for the sake of a clear graph.

- Depending on the starting point (in case of Fig. 4-1, this is $\theta^{(1)}_{\mathrm{old}} = 0$), the EM algorithm does not convergence to the global maximum of the LLF but to a local one. This is a major problem in any kind of gradient ascent technique such as EM and ways to overcome this will be discussed in Section 4.4.

- The auxiliary functions have in general lower values than the LLF [68] and the distance is given by the entropy of the pseudo symbols' posterior distribution (cf. Appendix C.2)

$$\mathrm{H}(\mathrm{P}_{\underline{\lambda}_{\mathrm{old}}^{(i)}}) = \mathcal{L}\left(\underline{\lambda}_{\mathrm{old}}^{(i)}\right) - \mathcal{Q}(\underline{\lambda}_{\mathrm{old}}^{(i)}, \underline{\lambda}_{\mathrm{old}}^{(i)}) \; . \tag{4.12}$$

  The probability's notation in the entropy's argument is hereby shortened for the sake of brevity. If the entropy is zero (i.e. the posterior probabilities are of Dirac delta nature), the auxiliary function of the i'th iteration are tangential to the LLF in $\underline{\lambda}_{\mathrm{old}}^{(i)}$.

- The auxiliary function have a strictly concave shape, which confirms the idea of the EM algorithm, that it makes the maximization easier than directly for the LLF.

- The distance $d(\theta)$ between the LLF's local maxima is no coincidence and depends on the waveform parameters as shown below in Section 4.4.1.

The next two sections cover the derivation of the EM algorithm for the channel parameter estimation using the non-linear CPM. In [40, 41], this was done for the subset CFO, CPO and TO (without providing explicit maximization steps) in general linear modulation using a channel code to formulate the expectation step. Whereas CPM can also be represented as a linear modulation through the pulse amplitude modulation (PAM) decomposition [71, 72], the derivation of the EM estimator is managed in the original form and thus this decomposition finds no application in this thesis. The derivation benefits at some points from the formulation introduced in (2.25), whereas the non-constant amplitude symbols introduced in the decomposition would prove to be obstructive in this regard.

4.2 Expectation Step for Burst Type CPM

For the sake of better readability, the iteration index (i) is dropped in the remainder, unless it is of direct relevance. As the auxiliary function and its maximization is the core of this thesis, the derivation will be carried out in appropriate detail. Beginning with the general formula of the auxiliary function in (4.4), the joint probability is transformed by Bayes' law into

$$\mathrm{p}_{\underline{\gamma},\underline{r}'}\left[\underline{\gamma},\underline{r}'|\underline{\lambda}\right] = \mathrm{p}_{\underline{r}'|\underline{\gamma}}\left[\underline{r}'|\underline{\gamma},\underline{\lambda}\right] \cdot \mathrm{P}_{\underline{\gamma}}\left[\underline{\gamma}|\underline{\lambda}\right] \ . \tag{4.13}$$

The probability density of a received sequence vector $\underline{r}'$ can be split into the product of each pseudo symbol's density because of the independence of observations in Markov chains

$$\mathrm{p}_{\underline{r}'|\underline{\gamma}}\left[\underline{r}'|\underline{\gamma},\underline{\lambda}\right] = \prod_{n=0}^{N-1} \mathrm{p}_{\underline{r}'|\underline{\gamma}}\left[\underline{r}'\left[nK:(n+1)K-1\right]|\gamma_n,\underline{\lambda}\right] \tag{4.14}$$

$$= \prod_{n=0}^{N-1} \frac{1}{\left(\pi\hat{\sigma}_{w'}^2\right)^K} \cdot \exp\left(-\frac{1}{\hat{\sigma}_{w'}^2}\sum_{k=0}^{K-1}\left|r'(kT_0+nT) - c_{\gamma_n}(kT_0,\underline{\lambda})\right|^2\right) \tag{4.15}$$

and the product operator's argument strongly resembles the path metric expression in (2.60). Both (2.60) and (4.15) assume i.i.d. noise samples. It is different in the missing a priori scaling, taking the true value of $\underline{\lambda}$ instead of an estimate and utilizing (2.28) for a lighter notation. The scaled, rotated and delayed pseudo symbols $c_{\gamma_n}(kT_0,\underline{\lambda})$ are defined in (2.61). The sequence probability is independent from the channel parameter vector and can be split up by applying Bayes' law N times and considering, that a state probability in a Markov chain is only dependent on its direct predecessor

$$\mathrm{P}_{\underline{\gamma}}\left[\underline{\gamma}|\underline{\lambda}\right] = \mathrm{P}_{\underline{\gamma}}\left[\underline{\gamma}\right] \tag{4.16}$$

$$= \mathrm{P}_{\gamma_{N-1}|\gamma_{N-2}}\left[\gamma_{N-1}|\gamma_{N-2}\right] \cdot \ \ldots \ \cdot \mathrm{P}_{\gamma_1|\gamma_0}\left[\gamma_1|\gamma_0\right] \cdot \mathrm{P}_{\gamma_0}\left[\gamma_0\right] \tag{4.17}$$

$$= \frac{1}{Q} \cdot \frac{1}{M^{N-L}} \ . \tag{4.18}$$

In the last line, the probability of $\frac{1}{Q \cdot M}$ for the first pseudo symbol γ_0, $1/M$ for every following trellis transition γ_1 to γ_{N-L} and 1 for the terminating symbols γ_{N-L+1} to γ_{N-1} were inserted. The result is basically of combinatoric nature being the inverse of the number of all possible, (a priori) equally probable sequences. Collecting the results from (4.13), (4.14) and (4.18) and inserting them in (4.4) gives

$$\begin{aligned}\mathcal{Q}\left(\underline{\lambda}, \underline{\lambda}_{\text{old}}\right) = C_0 + \sum_{\underline{\gamma}} \mathrm{P}_{\underline{\gamma}|\underline{r}'}\left[\underline{\gamma}|\underline{r}', \underline{\lambda}_{\text{old}}\right] \times \\ \sum_{n=0}^{N-1} \log\left(\mathrm{p}_{\underline{r}'|\gamma_n}\left[\underline{r}'\left[nK:(n+1)K-1\right]|\gamma_n, \underline{\lambda}\right]\right)\end{aligned} \tag{4.19}$$

in which the constant $C_0 = -\log(Q) - (N-L)\log(M)$ was pulled out of the sum. Next, the sum over all pseudo symbol sequences is rewritten as the sums over each individual element and the order of the nested sums is rearranged

$$\begin{aligned}\mathcal{Q}\left(\underline{\lambda}, \underline{\lambda}_{\text{old}}\right) = C_0 + \sum_{n=0}^{N-1}\sum_{\gamma_0=0}^{\Gamma-1}\cdots\sum_{\gamma_{N-1}=0}^{\Gamma-1} \mathrm{P}_{\underline{\gamma}|\underline{r}'}\left[\left[\gamma_0,\ldots,\gamma_{N-1}\right]|\underline{r}', \underline{\lambda}_{\text{old}}\right] \times \\ \log\left(\mathrm{p}_{\underline{r}'|\gamma_n}\left[\underline{r}'\left[nK:(n+1)K-1\right]|\gamma_n, \underline{\lambda}\right]\right)\ .\end{aligned} \tag{4.20}$$

By applying Bayes' law once again and shifting the sum over γ_n and the only dependent term to the front, a part of the above nested sum

$$\begin{aligned}&\sum_{\gamma_0=0}^{\Gamma-1}\cdots\sum_{\gamma_{N-1}=0}^{\Gamma-1} \mathrm{P}_{\underline{\gamma}|\underline{r}'}\left[\left[\gamma_0,\ldots,\gamma_{N-1}\right]|\underline{r}', \underline{\lambda}_{\text{old}}\right] \\ &= \sum_{\gamma_n=0}^{\Gamma-1} \mathrm{P}_{\gamma|\underline{r}'}\left[\gamma_n|\underline{r}', \underline{\lambda}_{\text{old}}\right] \times \\ &\sum_{\gamma_0=0}^{\Gamma-1}\cdots\sum_{\gamma_{n-1}=0}^{\Gamma-1}\sum_{\gamma_{n+1}=0}^{\Gamma-1}\cdots\sum_{\gamma_{N-1}=0}^{\Gamma-1} \mathrm{P}_{\underline{\gamma}|\gamma,\underline{r}'}\left[\left[\gamma_0,\ldots,\gamma_{n-1},\gamma_{n+1},\ldots,\gamma_{N-1}\right]|\gamma_n, \underline{r}', \underline{\lambda}_{\text{old}}\right]\end{aligned} \tag{4.21}$$

reduces greatly, since the sum of Γ^{N-1} terms in the last line equals unity. This step is probably the most crucial one, because besides the complexity reduction, the necessary posterior probabilities $\mathrm{P}_{\gamma|\underline{r}'}\left[\gamma_n|\underline{r}', \underline{\lambda}_{\text{old}}\right]$ are direct results of the CPM

detector (cf. (2.48)). The auxiliary function can now be expressed as

$$\mathcal{Q}\left(\underline{\lambda}, \underline{\lambda}_{\text{old}}\right) = C_0 + C_1 + \sum_{n=0}^{N-1} \sum_{\gamma_n=0}^{\Gamma-1} \mathrm{P}_{\boldsymbol{\gamma}|\underline{\boldsymbol{r}}'}\left[\gamma_n | \underline{r}', \underline{\lambda}_{\text{old}}\right] \times \left(-\frac{1}{\hat{\sigma}_{w'}^2} \sum_{k=0}^{K-1} \left|r'(kT_0 + nT) - c_{\gamma_n}(kT_0, \underline{\lambda})\right|^2\right) , \tag{4.22}$$

where (4.15) was inserted and $C_1 = -NK \log(\pi \hat{\sigma}_{w'}^2)$. Further simplifications follow from (A.2), so that

$$\mathcal{Q}\left(\underline{\lambda}, \underline{\lambda}_{\text{old}}\right) = C_0 + C_1 + C_2 - \frac{\alpha^2 E_{\text{S}} N}{\hat{N}_0} + \sum_{n=0}^{N-1} \sum_{\gamma_n=0}^{\Gamma-1} \mathrm{P}_{\boldsymbol{\gamma}|\underline{\boldsymbol{r}}'}\left[\gamma_n | \underline{r}', \underline{\lambda}_{\text{old}}\right] \cdot 2 \frac{T_0}{\hat{N}_0} \operatorname{Re}\left\{\sum_{k=0}^{K-1} r'(kT_0 + nT) \cdot c_{\gamma_n}^*(kT_0, \underline{\lambda})\right\} \tag{4.23}$$

with $C_2 \cdot \hat{N}_0 = -E_{r'} = -T_0 \sum_{n=0}^{N-1} \sum_{k=0}^{K-1} |r'(kT_0 + nT)|^2$ being the negative value of the received signal's energy. The pseudo symbols' energy $T_0 \sum_{k=0}^{K-1} |c_{\gamma_n}(kT_0, \underline{\lambda})|^2 = \alpha^2 E_{\text{S}}$ is solely influenced by the FF. The estimate of the complex noise power density is given by the relation $\hat{\sigma}_{w'}^2 = \hat{N}_0 T_0^{-1}$ and was previously obtained by (2.45).

The inner sum over γ_n can be drawn inside the real operator, so that an expected, complex conjugated pseudo symbol (cf. (2.61)) can be expressed as

$$\bar{c}_n^*\left(kT_0, \underline{\lambda}\right) = \sum_{\gamma_n=0}^{\Gamma-1} \mathrm{P}_{\boldsymbol{\gamma}|\underline{\boldsymbol{r}}'}\left[\gamma_n | \underline{r}', \underline{\lambda}_{\text{old}}\right] c_{\gamma_n}^*(kT_0, \underline{\lambda}) , \tag{4.24}$$

which in general does not retain the constant amplitude property of the CPM signal. At this point, the scaled real operator's argument $T_0 \sum_{k=0}^{K-1} r'(kT_0 + nT) \cdot \bar{c}_n^*\left(kT_0, \underline{\lambda}\right)$ can be interpreted as a zero lag CCF of the n'th received symbol and its expected version (cf. (1.11)). To extend this interpretation, all expected pseudo symbols are strung together and an expected reference signal based on the CPM detector's soft knowledge on the actual transmitted data symbols is defined

$$\bar{c}(kT_0, \underline{\lambda}) = \sum_{n=0}^{N-1} \bar{c}_n(kT_0 - nT, \underline{\lambda}) \tag{4.25}$$

with the time indicator $0 \leq k \leq NK-1$. The constants are combined in $C_\Sigma = \sum_{i=0}^{2} C_i$ and eventually, the auxiliary function for the case of joint FF, CFO, CPO and TO estimation ($\underline{\lambda} = [\alpha, \nu, \theta, \tau]$) in burst type CPM is expressed as

$$\mathcal{Q}\left(\underline{\lambda}, \underline{\lambda}_{\text{old}}\right) = C_\Sigma - \frac{\alpha^2 E_\text{S} N}{\hat{N}_0} + \frac{2T_0}{\hat{N}_0} \operatorname{Re}\left\{ \sum_{k=0}^{NK-1} r'(kT_0) \cdot \bar{c}^*(kT_0, \underline{\lambda}) \right\} \tag{4.26}$$

$$= C_\Sigma - \frac{\alpha^2 E_\text{S} N}{\hat{N}_0} + \frac{2\alpha}{\hat{N}_0} \operatorname{Re}\left\{ e^{-j\theta} \cdot \chi_{r',\bar{c}}(\nu, \tau) \right\} \tag{4.27}$$

with

$$\chi_{r',\bar{c}}(\nu, \tau) = T_0 \sum_{k=0}^{NK-1} r'(kT_0) \bar{c}^*(kT_0 - \tau) \cdot e^{-j2\pi\nu k T_0} \tag{4.28}$$

stating a zero lag correlation between the received and the expected reference signal. Hereby $\bar{c}(kT_0)$ is the non-adjusted (in contrast to (4.25)), strung together reference signal. In radar terminology, this is called a cross ambiguity function (CAF) and this terminology is picked up. The CAF is dependent on the CFO ν and the TO τ and an exemplary realization is shown in Fig. 4-2. Its maximal value depends on the energies of the two input signals and is

$$\max \chi_{r',\bar{c}}(\nu, \tau) = \sqrt{E_{r'} E_{\bar{c}}} = N E_\text{S} \tag{4.29}$$

in the noise free case, if the posterior probabilities are correct. Because of its concave shape, it can be maximized with appropriate effort with noting that this shape is only present for fixed posterior probabilities as enforced by the EM algorithm.

Having brought the auxiliary function $\mathcal{Q}\left(\underline{\lambda}, \underline{\lambda}_{\text{old}}\right)$ to a manageable form, it is noted, that the expectation step is principally associated with the posterior probabilities' calculation only. In this case however, an altered version of (4.24) (cf. Section 4.3.1) is beneficial to compute in order to enable efficient maximization steps in Section 4.3. Thus the expectation step is conceptually associated with it in addition to the posterior probabilities' computation. The auxiliary function $\mathcal{Q}\left(\underline{\lambda}, \underline{\lambda}_{\text{old}}\right)$ itself is not necessary to be calculated in general and only used to formulate the maximization step. However, it can be explicitly used to determine the Kullback-Leibler divergence

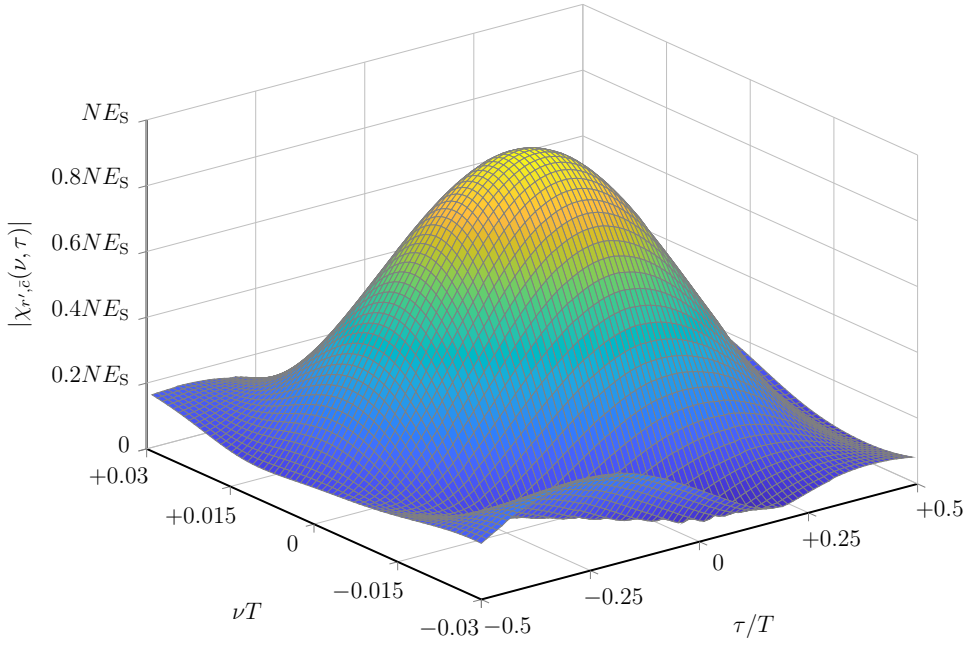


Figure 4-2: Exemplary CAF surface plot of a 4/7-Q2RC signal at $10\log_{10} E_S/N_0 = 10\,\text{dB}$ with the true values of CFO and TO are zero.. The FF $\alpha = 1$ and CPO $\theta = 0$ are fixed in this example.

in (4.11).

4.3 Maximization Step for Burst Type CPM

The M-Step (given in (4.5)) maximizes the auxiliary function $\mathcal{Q}(\underline{\lambda}, \underline{\lambda}_{\text{old}})$ with regard to $\underline{\lambda}$. Because of the auxiliary function's final form in (4.27), this is not analytically possible if the maximization shall be carried out in a joint way. Instead the problem is split up in to three pieces that are mutually decoupled. It is noted, that the constant C_Σ plays no part in the maximization procedure.

4.3.1 Maximization with Regard to CFO and TO

For (4.27) to have the maximal value, the term inside the real operator has to have the largest value possible. Since the complex exponential function can rotate any value of the CAF to be purely real, the CAF can be maximized with respect to its

absolute value $|\chi_{r',\bar{c}}(\nu,\tau)|$. The maximization of the CAF

$$[\hat{\nu},\hat{\tau}] = \underset{\nu,\tau}{\operatorname{argmax}} \; |\chi_{r',\bar{c}}(\nu,\tau)| \tag{4.30}$$

is an important problem in the field of radar signal processing and many publications cover efficient computation methods for specific signal shapes, e.g. [73, 74]. Waveform-independent methods are described in e.g. [75] whose suitability depends on the required estimation range and resolution. The principal idea of maximizing the CPM CAF for the considered CFO and TO estimation is outlined below with a suggestion of an efficient implementation following in Section 4.5.2.

A true joint estimation requires a computation of the CAF's grid points, which shall be avoided due to the undesired computational effort. Instead a subsequent estimation of both parameters is considered and the crucial question at hand is the order in which it shall take place. As introduced in Section 3.3, the CFO and TO entries in the MFIM are not coupled, which means that each parameter can be estimated individually as long as the other is not too far off. The actual modified Fisher information (MFI) values MFI(ν) and MFI(τ) are found in the respective diagonal elements of the MFIM and tell which one's estimation is more robust with regard to the other parameter. The relation (elaborated in Appendix C.3)

$$T^{-2}\mathrm{MFI}(\nu) \gg T^{2}\mathrm{MFI}(\tau) \;, \tag{4.31}$$

of the normalized MFIs suggests estimating the CFO before the TO. While a slight TO mismatch should not influence the joint LLF too much, the exact opposite is expected to happen with a mismatched CFO due to its orders of magnitude larger MFI. Based on (4.31), it can be assumed that the estimation of CFO and TO are practically independent if carried out in the suggested order.

Since (4.28) is (irrespective of the scaling T_0) the definition of the discrete Fourier transform (DFT) of the received and reference signals' product, the cross ambiguity's absolute value can be maximized with regard to the CFO by computing this DFT

and determining the maximizing ν

$$\hat{\nu} = \underset{\nu}{\operatorname{argmax}} \left|\chi_{r',\bar{c}}(\nu, \tau_{\text{old}})\right| \tag{4.32}$$

$$= \underset{\nu}{\operatorname{argmax}} \left| T_0 \sum_{k=0}^{NK-1} r'(kT_0)\bar{c}^*(kT_0 - \tau_{\text{old}}) \cdot e^{-j2\pi\nu kT_0} \right| \tag{4.33}$$

with a possible estimation range of $\hat{\nu} \in \left[-0.5T_0^{-1}; +0.5T_0^{-1}\right)$ and a spectral resolution of $(NT)^{-1}$. To allow for the partial maximization, the TO is set to a fixed value τ_{old} stemming from the last EM iteration or the initialization. The next step of estimating the TO can be realized by the computing the discrete CCF of the frequency corrected receive signal and the reference

$$\hat{\tau} = \underset{\tau}{\operatorname{argmax}} \left|\chi_{r',\bar{c}}(\hat{\nu}, \tau)\right| \tag{4.34}$$

$$= \underset{\tau}{\operatorname{argmax}} \left| T_0 \sum_{k=0}^{NK-1} r'(kT_0)\bar{c}^*(kT_0 - \tau) \cdot e^{-j2\pi\hat{\nu}kT_0} \right| \tag{4.35}$$

whereas the principal TO range is $\hat{\tau} \in (-NT; +NT)$ and the temporal resolution equals T_0. It is noted, that by splitting up the maximization step in (4.30) into two parts and not strictly optimizing it jointly, the presented approach falls into the framework of the expectation conditional maximization (ECM) algorithm [76]. It was reasoned with (4.31), that the estimates are practically uncoupled if obtained in the suggested order and thus the presented estimator will still be called EM further on.

The estimation results' accuracies are limited by the resolution of their corresponding operations. Unless the receiver's system is highly oversampled, the theoretical limits from Section 3.4 cannot be achieved and thus some form of interpolation is required if reaching these limits is desired. This and an efficient implementation suggestion will be covered in Section 4.5.2.

4.3.2 Maximization with Regard to CPO

The maximization of the auxiliary function with respect to the CPO is decoupled from the other parameters and can be realized by the closed form expression

$$\hat{\theta}_0 = \arg\left\{\chi_{r',\bar{c}}(\hat{\nu},\hat{\tau})\right\} - 2\pi\xi\hat{\tau} \tag{4.36}$$

in which the angle of the CAF includes the phase offset due to the delayed tilting. The estimate $\hat{\theta}_0$ exactly lets $\mathrm{Re}\left\{e^{-j\hat{\theta}_0}\cdot\chi_{r',\bar{c}}(\hat{\nu},\hat{\tau})\right\} = |\chi_{r',\bar{c}}(\hat{\nu},\hat{\tau})|$, which is necessary for (4.27) to be maximal. Equation (4.36) computes an estimate of the fractional CPO $\theta_0 = \theta \bmod \psi$, which can be inserted in (2.61) instead of a $\hat{\theta}$ without any consequences on the system performance. The integer part $z\psi$ of the CPO θ is inherently handled by the trellis initialization as explained in Section 2.5.2.

In a practical implementation, the correction term $2\pi\xi\hat{\tau}$ can be omitted in both (4.36) and (2.61) since the CAF angle includes the influence of the delayed tilting. However, this would lead to incorrect CPO estimates, which is avoided if the correction term remains included.

4.3.3 Maximization with Regard to FF

The help function $\mathcal{Q}(\underline{\lambda},\underline{\lambda}_{\mathrm{old}})$ is quadratic with regard to the FF α and due to its negative curvature it is maximized by setting the partial derivation to zero

$$\frac{\partial\mathcal{Q}(\underline{\lambda},\underline{\lambda}_{\mathrm{old}})}{\partial\alpha} = -\frac{2\alpha N E_{\mathrm{S}}}{\hat{N}_0} + \frac{2}{\hat{N}_0}\,\mathrm{Re}\left\{e^{-j\theta}\cdot\chi_{r',\bar{c}}(\nu,\tau)\right\} \stackrel{!}{=} 0 \tag{4.37}$$

and isolating α. The FF estimate is thus obtained through

$$\hat{\alpha} = \frac{|\chi_{r',\bar{c}}(\hat{\nu},\hat{\tau})|}{N E_{\mathrm{S}}}\ , \tag{4.38}$$

where the previously obtained estimates are inserted.

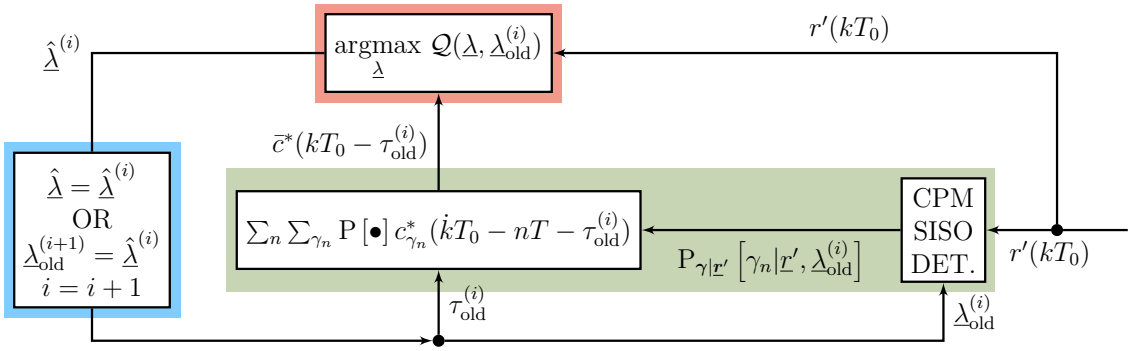


Figure 4-3: Block diagram of the EM algorithm showing the expectation step (highlighted by the green box), the maximization step (highlighted by the red box) and the convergence check (highlighted by the blue box). The E-step consists of the calculation of the posterior trellis branch probabilities $\mathrm{P}_{\boldsymbol{\gamma}|\underline{r}'}\left[\gamma_n|\underline{r}', \underline{\lambda}_{\text{old}}^{(i)}\right]$ (cf. (2.48)) and the creation of the expected reference signal's complex conjugate $\bar{c}^*(kT_0 - \tau_{\text{old}}^{(i)})$. Equations (4.33), (4.35), (4.36), (4.38) and(4.39) comprise the M-step.

Eventually the maximization step according to (4.5) is completed by

$$\hat{\underline{\lambda}} = \left[\hat{\alpha}, \hat{\nu}, \hat{\theta}_0, \hat{\tau}\right] . \tag{4.39}$$

Figure 4-3 visualizes the complete EM algorithm outlined in the beginning of Section 4.1 for the considered channel parameter estimation.

4.3.4 Evaluation of a Maximization with Regard to the Noise Power Density

The auxiliary function was among many other steps simplified by identifying constant terms that thus are of no relevance in maximizing. In the constants C_1 and C_2, the complex noise power density $\hat{N}_0$ is included, which was estimated by (2.45) and was not subject to optimization in the EM algorithm. But since the term appears in $\mathcal{Q}\left(\underline{\lambda}, \underline{\lambda}_{\text{old}}\right)$, the question arises whether it should be included. In order to do so, $\hat{N}_0$ is switched with N_0 and $N_{0,\text{old}}$, respectively in (4.27) and the adjusted auxiliary function is as follows

$$\begin{aligned}\mathcal{Q}\left([\underline{\lambda}, N_0], [\underline{\lambda}_{\text{old}}, N_{0,\text{old}}]\right) = C_0 - NK \log\left(\pi \frac{N_0}{T_0}\right) - \frac{E_{r'} + \alpha^2 E_{\text{S}} N}{N_0} \\ + \frac{2\alpha}{N_0} \operatorname{Re}\left\{e^{-j\theta} \cdot \chi_{r',\bar{c}}(\nu, \tau)\right\} .\end{aligned} \tag{4.40}$$

After determining the corresponding partial derivative as

$$\frac{\partial \mathcal{Q}\left([\underline{\lambda}, N_0], [\underline{\lambda}_{\text{old}}, N_{0,\text{old}}]\right)}{\partial N_0} = -\frac{NK}{N_0} + \frac{E_{r'} + \alpha^2 E_{\text{S}} N}{N_0^2} - 2\frac{\alpha \,\text{Re}\left\{e^{-j\theta} \cdot \chi_{r',\bar{c}}(\nu,\tau)\right\}}{N_0^2}\,, \tag{4.41}$$

it is set to zero and N_0 is isolated to obtain the estimate

$$\hat{N}_0 = \frac{E_{r'} - 2\hat{\alpha}\,|\chi_{r',\bar{c}}(\hat{\nu},\hat{\tau})| + \hat{\alpha}^2 N E_{\text{S}}}{NK} \tag{4.42}$$

with all previous estimates considered. In order to verify this optimum constituting a maximum, it is plugged into the second derivative

$$\left.\frac{\partial^2 \mathcal{Q}\left([\underline{\lambda}, N_0], [\underline{\lambda}_{\text{old}}, N_{0,\text{old}}]\right)}{\partial^2 N_0}\right|_{N_0=\hat{N}_0} = -\frac{(NK)^3}{\left(E_{r'} - 2\hat{\alpha}\,|\chi_{r',\bar{c}}(\hat{\nu},\hat{\tau})| + \hat{\alpha}^2 N E_{\text{S}}\right)^2} \tag{4.43}$$

and confirmed that its sign is negative.

The following case is considered to explain why (4.42) might not be a good method to estimate the noise power density. A noise free ($N_0 = 0$) transmission with an arbitrary channel parameter vector $\underline{\lambda}$ is assumed. As the CAF basically having the structure of a CCF, the maximum value it can take is given by

$$\max\left\{\chi_{r',\bar{c}}(\hat{\nu},\hat{\tau})\right\}|_{N_0=0} = \sqrt{E_{r'} \cdot N E_{\text{S}}} \tag{4.44}$$

$$= \alpha N E_{\text{S}}\,, \tag{4.45}$$

where the second line follows from the noise free received signal having exactly the energy $\alpha^2 N E_{\text{S}}$. For (4.42) to give the estimate $\hat{N}_0 = 0$, two conditions must be met. First, the FF estimate $\hat{\alpha}$ must equal the true value and second, the maximum value of (4.45) must be reached in (4.42). For the second condition to be met in a noise free case, the CFO, CPO and TO estimates must equal the true values. Thus if $\hat{\underline{\lambda}} \neq \underline{\lambda}$, a noise power density greater than zero is estimated. While it might seem tolerable to have deviations in the estimates for the moment being in an iterative procedure, the important difference to consider here is, that the estimate $\hat{N}_0$ has direct (and impactful) deteriorating influence on the path metrics in (2.63). The estimation rule

in (2.45) though is robust with regard to the above problems and delivers estimates of sufficient accuracy.

4.4 Initialization of the Parameters

The EM algorithm falls under the gradient ascent methods and as such it will converge to the next best (possibly local) maximum of its cost function (the LLF in the EM case). It will not find any other (possibly global) maximum and therefore a natural uncertainty about the maximum's quality will remain (cf. Fig. 4-1). To combat this uncertainty, a simple approach would be to let the algorithm start at several parameter points and let each instance converge. This way, the reached maxima can be compared in terms of their cost function and the best candidate can be chosen. If the starting points are dense enough, the chance of reaching the global maximum can be increased. The application of the EM algorithm to the channel parameter estimation suffers from the above elaborated problem and this section covers the selection of a starting point grid such that its size is as small as possible with the constraint, that reaching the global maximum of the LLF shall be guaranteed for at least one of the starting points. For each LLF, the principal minimal distance (PMD) between two adjacent maxima will be determined and the grid is then designed accordingly. The obvious caveat about this approach is the naturally increased computational complexity in contrast to only one starting point, which shall be covered in Section 4.5. In the case of CPO and TO estimation in CPM transmissions, [29] describes the problem of wrong estimates using a PLL and proposes a similar approach of multiple loop initializations to tackle it.

In general, if a transmitted pseudo symbol $c_k(kT_0)$ is received without channel manipulations, the LLF $\mathcal{L}(\underline{\lambda})$ is maximal for $\underline{\lambda} = [\alpha = 1, \nu = 0, \theta = 0, \tau = 0]$. The likelihood function will have further local maxima at some other $\underline{\lambda}$ if this channel parameter vector constitutes a manipulation of $c_k(kT_0)$ such that a different pseudo symbol $c_l(kT_0)$ is mimicked by this. The height difference of maxima in the LLF can still be large and the global maximum is distinguishable, but since EM converges in

any maximum, more than one starting point is required to explore all maxima and find the ML solution.

Since only the phase is modulated in CPM, a manipulation in amplitude will never mimic a different pseudo symbol and consequently no local maxima are expected in the FF dimension of $\mathcal{L}(\underline{\lambda})$. Still, also an FF grid point will be derived in the form of a coarse feed forward estimate. While this is sufficient for the fading factor, coarse estimators of CFO (e.g. [22,24]), CPO (e.g. [25] with the adaption to multilevel CPM from [26]) and TO (e.g. [23]) are not suitable for finding starting points due to their limited performance especially at lower SNRs and for more complex CPM parameters as it was shown in [27, 28], where heuristically derived grids were used for the EM initialization.

In this work, the grid spacing shall be optimized analytically by investigating the scalar LLFs of the channel parameter vector' elements. For the sake of analytical tractability, a noise free transmission of CPM signals with rectangular frequency pulses is assumed. The phase and frequency pulse in this case is given as (cf. Fig 2-2)

$$q_{\mathrm{RE}}(t) = \frac{t}{2LT} \tag{4.46}$$

$$g_{\mathrm{RE}}(t) = \frac{\mathrm{d}q_{\mathrm{RE}}(t)}{\mathrm{d}t} = \frac{1}{2LT} \tag{4.47}$$

respecting the temporal restrictions $0 \leq t < LT$ of (2.4). The results of the following discussion are approximately valid for noisy transmissions and or non-rectangular pulse shapes in the sense, that a pseudo symbol $c_l(kT_0)$ is not exactly but only approximately mimicked by a manipulated $c_k(kT_0)$. The according LLFs still show local maxima at the points that were derived for the noise free, rectangular case. This is indirectly confirmed by the simulations presented in Chapter 5.

Due to the rotational invariance (cf. Section 2.5.2), the phase state Ψ'_n itself does not have any influence in the CPM detection and the CPO LLF is therefore periodic with ψ (cf. Fig. 4-1). Hence the actual phase state does not need to be considered in the following contemplations about how CPO, CFO and TO can change the signal

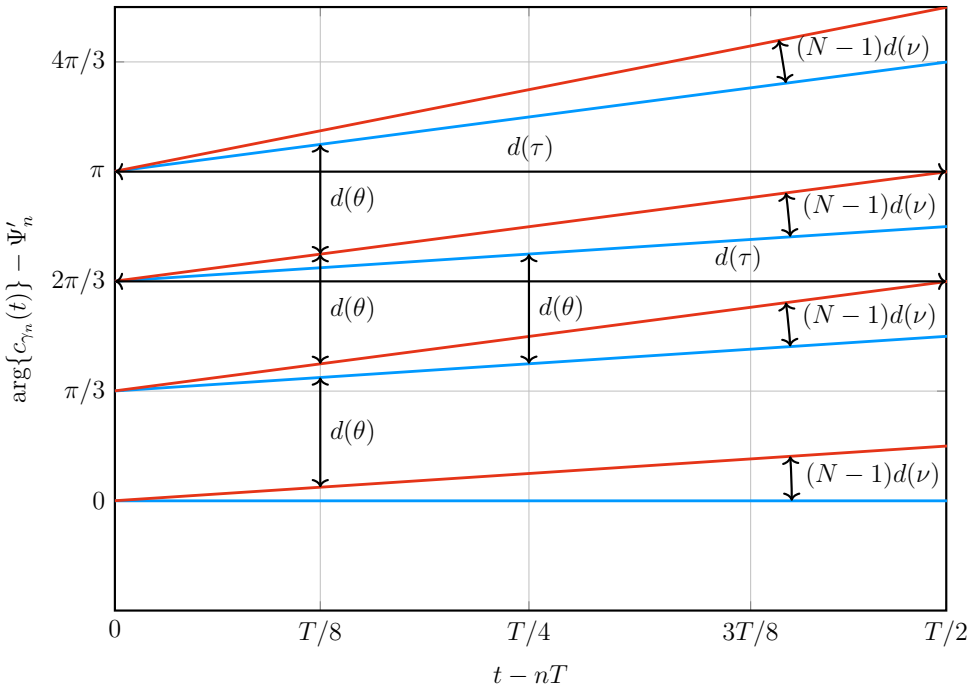


Figure 4-4: Exemplary phase trajectories of eight time invariant trellis branches of an arbitrary phase state for a rectangular waveform with $h = 1/2$, $M = 2$, $L = 3$, which corresponds to the rectangular version of the 1/2-B3GA waveform from Table 5.1. The correlative states are from bottom to top $A'_n = [0, 0]$, $A'_n = [0, 1]$, $A'_n = [1, 0]$ and $A'_n = [1, 1]$. The PMD's are in this case $d(\theta) = \pi/3$, $d(\nu) = (6(N-1)T)^{-1}$ and $d(\tau) = T/2$ with their general formulas being derived in the following subsections.

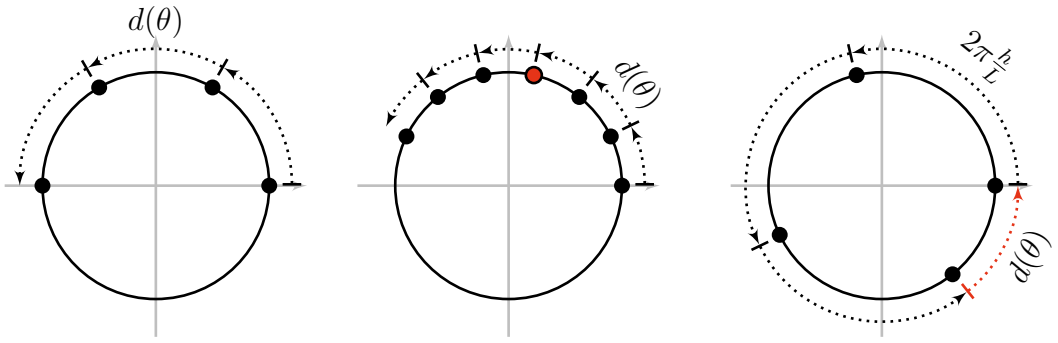
phase in such a way that a further maximum in the LLF arises. Figure 4-4 shows the phase trajectories $\arg\{c_{\gamma_n}(t)\}$ of every pseudo symbol γ_n irrespective of the phase state Ψ'_n of an exemplary CPM scheme.

The starting phase of each trajectory (not to be confused with the phase state Ψ'_n) is given as

$$\phi'_n(t = nT, \underline{a}') = 4\pi h \sum_{l=0}^{L-1} a'_{n-l} q_{\mathrm{RE}}(lT) \tag{4.48}$$

$$= 2\pi \frac{h}{L} \sum_{l=1}^{L-1} a'_{n-l} l \tag{4.49}$$

and depends on the sum of the $L-1$ latest symbols, that are drawn from the discrete symbol alphabet $\mathcal{A}'$ (cf. (2.14)). This constitutes exactly the correlative state, which is why it is called correlative phase value (CPV) henceforth. Their instantaneous frequencies are calculated as the scaled time derivative of (2.15) in the time slot



(a) Rectangular waveform with $h = 1/2$, $M = 2$ and $L = 3$.

(b) Rectangular waveform with $h = 2/7$, $M = 2$ and $L = 4$.

(c) Rectangular waveform with $h = 4/7$, $M = 4$ and $L = 2$.

Figure 4-5: CPVs of different rectangular waveforms. The three cases correspond to the rectangular versions of the three waveforms given in Table 5.1. The red dot in Fig. 4-5b represents a phase occupied by multiple CPVs. The red dashed arrow in Fig. 4-5c represents the second line in (4.57).

$nT \leq t < (n+1)T$

$$f_{\mathrm{i}}(t) = \frac{1}{2\pi} \cdot \frac{\mathrm{d}\phi'_n(t, \underline{a}')}{\mathrm{d}t} \tag{4.50}$$

$$= \frac{h}{LT} \sum_{l=0}^{L-1} a'_{n-l} \,, \tag{4.51}$$

where (4.47) was inserted. The frequency is influenced by the sum of the L latest symbols. Several combinations can lead to the same CPV or instantaneous frequency.

4.4.1 CPO Grid Spacing

Referring to Fig. 4-4, several pairs of pseudo symbols sharing the same $f_{\mathrm{i}}(t)$ can be identified (by their equal slope) and are thus confused when their phase difference is overcome by an appropriate CPO. Due to the sum's structure in (4.49), the minimal distance of two CPV's happens here e.g. when all memory symbols $a'_{n-(L-1)}, \ldots a'_{n-(2)}$ coincide and the symbol a'_{n-1} deviates minimally, i.e. by one and thus takes the value $2\pi\frac{h}{L}$. In Fig. 4-5, three cases (that correspond to the rectangular versions of Table 5.1's waveforms) of the positions of the CPV are shown.

In the cases of Fig. 4-5a and 4-5b, $2\pi\frac{h}{L}$ is also the PMD $d(\theta)$. In Fig. 4-5b, two correlative states $A'_n = [0, 1, 1]$ and $A'_n = [1, 0, 0]$ cause the same initial pseudo

symbol phase for $\gamma_n = 3$ and $\gamma_n = 4$, respectively. This has no influence on the CPO estimation whatsoever, because their differentiation is handled by the CPM trellis as it would be in the perfectly synchronized case. The equivalence of CPVs happens for large enough M and L, if the ratio $\frac{L}{h}$ is an integer.

If this quotient is not an integer as in the case of Fig. 4-5c, PMDs smaller than $2\pi\frac{h}{L}$ can occur if the CPVs exceed one revolution around the unit circle

$$4\pi h \cdot \frac{1}{2}(M-1)\left[\sum_{l=1}^{L-1}\frac{l}{L}\right] > 2\pi - 2\pi\frac{h}{L}\,, \tag{4.52}$$

$$L\left[(M-1)(L-1) - \frac{2}{h}\right] > 2\,, \tag{4.53}$$

where it is checked whether the largest CPV is closer to the zero angle ($\equiv 2\pi$) than $2\pi\frac{h}{L}$. If (4.53) is fulfilled, the PMD is the difference between the zero angle and the CPV that lies closest to it. The corresponding pseudo symbol's number γ_n is given by

$$2\pi\frac{h}{L}\cdot\gamma_n \overset{!}{\approx} 2\pi\,, \tag{4.54}$$

$$\gamma_n \approx \frac{L}{h}\,, \tag{4.55}$$

$$\gamma_n = \left\lfloor\frac{L}{h}\right\rceil \tag{4.56}$$

using the side constraint of γ_n being an integer in the last line. The integer closest to the ratio L/h is logically its rounded value. Gathering the above results gives finally the minimal phase distance at which local maxima in the CPO LLF can occur

$$d(\theta) = \begin{cases} 2\pi\frac{h}{L} & \text{if (4.58) is TRUE}\,, \\ \left|2\pi\left(\frac{h}{L}\cdot\left\lfloor\frac{L}{h}\right\rceil - 1\right)\right| & \text{otherwise}\,, \end{cases} \tag{4.57}$$

with

$$\frac{L}{h} \in \mathbb{Z} \quad \text{OR} \quad L\left[(M-1)(L-1) - \frac{2}{h}\right] \leq 2\,. \tag{4.58}$$

In Fig. 4-1 and 4-4, the PMD $d(\theta)$ is shown for the same exemplary waveform. It

must be noted, that for exotic parameter choices the reduced PMD $\left|2\pi\left(\frac{h}{L}\cdot\left\lfloor\frac{L}{h}\right\rceil-1\right)\right|$ might not guarantee two symbols sharing the same frequency and thus the actual $d(\theta)$ would fall back to the value of $2\pi\frac{h}{L}$. For this to happen, not only (4.58) must be false, but also a very low modulation index must be combined with very long pulses and a high level symbol alphabet (cf. Fig. 4-5 for counter examples, where (4.57) is valid). Such a parameter choice would arguably be rather exotic and at least because of its enormous trellis complexity of little practical interest. Anyway, (4.57) can still be used as lower bound in such a case. It is further noted, that for all waveforms considered in this work (cf. Section 5.1.1), the derived PMD of (4.57) is valid. An exemplary LLF demonstrating the phase PMD was shown in Fig. 4-1.

4.4.2 CFO Grid Spacing

Pseudo symbols sharing the same initial phase only differ in their instantaneous frequency (cf. (4.51)) and thus could be easily confused by an appropriate frequency offset. The necessary CFO for this to happen though is usually even larger than a realistic CFO range (cf. (2.34)). Another effect of a frequency offset is an increasing phase offset over time and the same contemplations as in the CPO case apply. The minimal CFO that can create a side maximum in its LLF induces a phase offset $d(\theta)$ in the last symbol of the burst, so that it would equal the first one besides a very little mismatch in its instantaneous frequency. The PMD therefore calculates as

$$d(\nu)=\frac{d(\theta)}{2\pi(N-1)T}\ . \tag{4.59}$$

Even though all symbols in between have a mismatch in phase and frequency, this still leads to a local maximum due to the falsely matching last symbol. In Fig. 4-4, the PMD is depicted $N-1$-fold first to make it graphically appear and second to set the initial thought in the beginning of this subsection in relation to it.

4.4.3 TO Grid Spacing

For a TO to change one symbol into another to induce a side maximum in its LLF, the two symbols must have the same $f_\mathrm{i}(t)$ as in the CPO case. As can be observed in Fig. 4-4, the necessary TO decreases with increasing $f_\mathrm{i}(t)$. The largest $f_\mathrm{i}(t)$ that two pseudo symbols have in common is given by

$$f_\mathrm{m} = \frac{h}{LT}\left[(M-2) + \sum_{l=1}^{L-1}(M-1)\right] \tag{4.60}$$

$$= \frac{h}{LT}\left[L(M-1)-1\right] \ , \tag{4.61}$$

which occurs when the correlative state contains only $M-1$ but one $M-2$ data symbols. The higher frequency of $\frac{h}{T}(M-1)$ is occupied by only one pseudo symbol ($\gamma_n = \Gamma - 1$) which thus cannot be mistaken by a TO. Two symbols sharing f_m always (but not necessary exclusively) occur at the two greatest CPV (cf. (4.49)). As introduced in Section 4.4.1, these two phases values have the distance $2\pi\frac{h}{L}$. The PMD for a TO to induce a side maximum in its LLF is thus given by the ratio of two adjacent CPV's distance and the maximal mutual frequency

$$d(\tau) = \frac{2\pi h/L}{2\pi f_\mathrm{m}} \tag{4.62}$$

$$= \frac{T}{L(M-1)-1} \ . \tag{4.63}$$

Similar to the CPO case, the validity of the derived PMD must be limited if (4.58) is not valid. For the same kind of exotic parameter choice, it may not be guaranteed, that the highest possible mutual frequency combined with above used phase distance leads to the TO PMD or a lower frequency in combination with a reduced CPO PMD. By the same reasoning as in the CPO subsection, this eventuality is discarded with reference on the hardly existing practical relevance.

4.4.4 FF Initialization

As mentioned above, no local maxima are expected in the FF LLF since a change in the symbol amplitude does not lead to mimicking another valid symbol. In contrast to the other parameters, the CCF in (2.63) is only scaled by the FF and not further affected by it. This is why a different approach is chosen here and the FF grid point is obtained by a coarse estimation in which it is approximately assumed, that the transmit signal and the additive noise are orthogonal to each other. With the knowledge of the noise variance from (2.45), the FF α is estimated by

$$\hat{\alpha} = \sqrt{\frac{\left[\frac{T_0}{NK}\sum_{k=0}^{NK-1}\left|r'(kT_0)\right|^2\right] - T_0\hat{\sigma}_{w'}^2}{E_\mathrm{S}/K}} \tag{4.64}$$

in which the nominator calculates the sample energy of the received signal's transmit component and the denominator expresses the energy of one transmitted sample. The estimate $\hat{\alpha}$ is, under the assumed channel conditions, accurate enough for it to be serving as a single starting value and thus the grid is simply given as

$$\mathcal{I}(\alpha) = \hat{\alpha}\ . \tag{4.65}$$

The consequences of the orthogonality assumption between the signal and the noise, which is in general not true, leads to initial FF estimates that are generally too large. The implications are discussed in Sections 5.4 and 5.5.

4.4.5 Grid For Joint Estimation

The above formulated grid spacings were derived for the case of single parameter estimation. They resemble the heuristically determined spacings in [26, 27], where CPO and CFO, respectively were estimated by means of the EM algorithm. In [28], a joint, EM based estimation of CPO and TO was presented and the (again heuristically derived) initialization grid has roughly half the grid spacings than above derived. This is explained by the fact that both CPO and TO can lead to the same pseudo symbol

confusion (still only strictly valid for rectangular pulse shapes)

$$c_k(t) = e^{jd(\theta)} c_l(t) \tag{4.66}$$

$$= c_l(t + d(\tau)) \ , \ nT - d(\tau) \leq t < (n+1)T - d(\tau) \tag{4.67}$$

$$= e^{j(1-x)d(\theta)} c_l(t + xd(\tau)) \ , \ nT - d(\tau) \leq t < (n+1)T - xd(\tau) \ , \ 0 \leq x \leq 1 \ . \tag{4.68}$$

Referring to Fig. 4-4, both the CPO $d(\theta)$ and the TO $d(\tau)$ or appropriate combinations change $c_{l=5}(t)$ to $c_{k=6}(t)$. It is noted, that the TO produces an exact replica only in the given time slot and a linear extension of the phase trajectory outside of it. Accordingly, if a pseudo symbol is shifted to the right, the time slot is given as $nT + xd(\tau) \leq t < (n+1)T$. In these cases, combinations of CPO and TO lead to similar effects in which the causative offset cannot be reliably determined with the grid spacing derived above. To account for this ambiguity, the CPO and TO grid spacing for joint estimation is halved, which is expected to generate reliable results similar to [28].

In Fig. 4-6, the CPO and TO grid points are added to the LLF contour and also the LLF gradients (i.e. the direction in which the EM algorithm will converge) from these points are shown. The two points marked by dashed lines would be the starting points if the joint estimation would be treated as a scalar one. It can be seen that the EM converges from both of them to local maxima and thus is not able to find the ML solution. The eight points marked by solid lines indicate the adjusted grid points as reasoned above. The ML solution is reached by at least one of them.

The number of grid points in the cases of CFO, CPO, TO is principally determined as the parameter ranges' (cf. (2.34), Section 4.4.1 and (2.36), respectively) quotient with the corresponding minimal distances. The ranges are $\Delta(\underline{\lambda}_2 = \nu) = 0.042T^{-1}$, $\Delta(\underline{\lambda}_3 = \theta) = \psi$, $\Delta(\underline{\lambda}_4 = \tau) = T$ and thus the number of grid points is given as

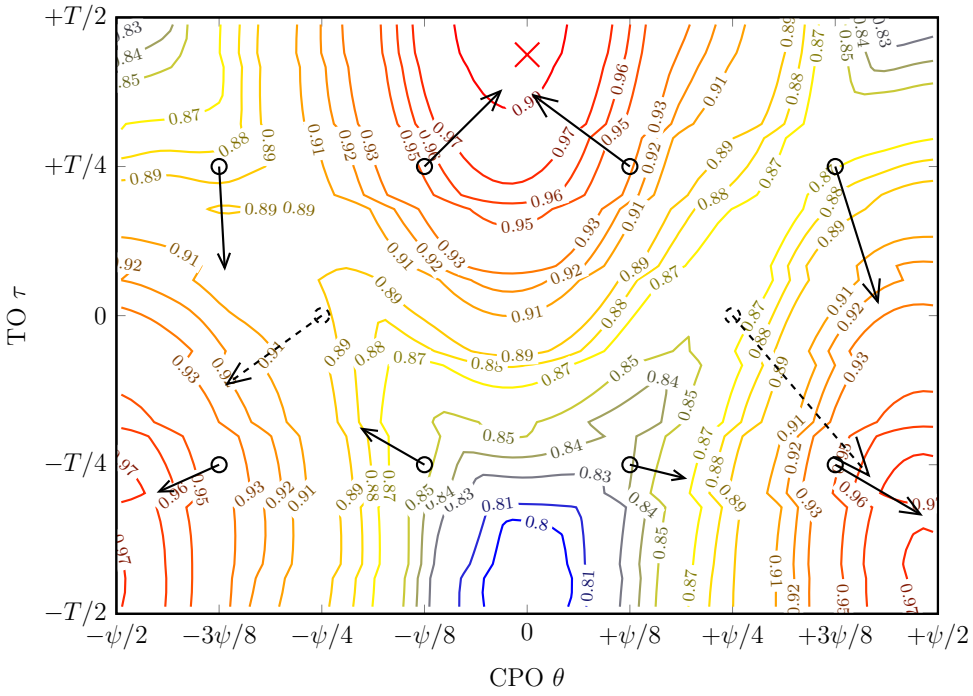


Figure 4-6: Contour plot of an exemplary LLF $\mathcal{L}([\alpha = 1, \nu = 0, \theta, \tau])$ with fixed values for the FF and CFO at $10\log_{10} E_S/N_0 = 10\,\text{dB}$. Its normalized values are written into the contours and the colors indicate the height (ascending from blue to red). The waveform parameters are $h = 1/3$, $M = 2$, $L = 2$ and $g(t)$ has a rectangular form. The true values for the CPO and TO are $\theta = 0$ and $\tau = 7T/16$ with the LLF's global maximum (marked by a red cross) lying in the direct vicinity to that point. Two distinctive local maxima can be found at both bottom corners of the plot at around $(\pm\psi/2, -3T/8)$. The dashed circles indicate the CPO and TO starting point grid strictly according to the LLF analysis in Sections 4.4.1 and 4.4.3. The adjusted grid suggested in the Section 4.4.5 is marked by solid circles. This graph is taken from the author's work in [66].

following

$$N_\text{I}(\alpha) = 1\ , \tag{4.69}$$

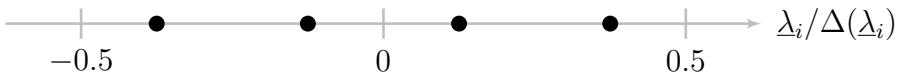
$$N_\text{I}(\nu) = \left\lceil \frac{\Delta(\nu)}{d(\nu)} \right\rceil = \begin{cases} \left\lceil \frac{0.042L(N-1)}{h} \right\rceil & \text{if (4.58) is TRUE}\ , \\ \left\lceil \frac{0.042(N-1)}{\left|\left(\frac{h}{L}\cdot\left\lfloor\frac{L}{h}\right\rceil - 1\right)\right|} \right\rceil & \text{otherwise}\ , \end{cases} \tag{4.70}$$

$$N_\text{I}(\theta) = 2\left\lceil \frac{\Delta(\theta)}{d(\theta)} \right\rceil = \begin{cases} 2\left\lceil \frac{L}{P} \right\rceil & \text{if (4.58) is TRUE}\ , \\ 2\left\lceil \left|\frac{P}{L}\cdot\left\lfloor\frac{L}{h}\right\rceil - Q\right|^{-1} \right\rceil & \text{otherwise}\ , \end{cases} \tag{4.71}$$

$$N_\text{I}(\tau) = 2\left\lceil \frac{\Delta(\tau)}{d(\tau)} \right\rceil = 2\left(L(M-1) - 1\right)\ . \tag{4.72}$$

The grids for CFO, CPO and TO are defined with points distributed uniformly in

(a) Distribution grid for $N_{\mathrm{I}}(\underline{\lambda}_i) = 3$



(b) Distribution grid for $N_{\mathrm{I}}(\underline{\lambda}_i) = 4$

Figure 4-7: Distribution of grid points according to (4.73). One odd and one even number case is displayed.

the respective estimation range $\Delta(\lambda_i), (2 \leq i \leq 4)$

$$\mathcal{I}(\underline{\lambda}_i) = \left\{ \pm \left(\frac{\Delta(\underline{\lambda}_i)}{2} - \frac{\Delta(\underline{\lambda}_i)}{2N_{\mathrm{I}}(\underline{\lambda}_i)} \right), \pm \left(\frac{\Delta(\underline{\lambda}_i)}{2} - \frac{3\Delta(\underline{\lambda}_i)}{2N_{\mathrm{I}}(\lambda_i)} \right), \dots \right\} \tag{4.73}$$

with $\underline{\lambda}_i$ specifying the i'th element of the parameter vector $\underline{\lambda}$. Figure 4-7 displays such a one-dimensional grid. By this choice, the possibility of a non-reachable maximum is erased. The oscillator quality, discussed in Section 2.3.2 has a direct influence on the cardinality of the CFO grid.

Eventually the complete four-dimensional initialization grid with a cardinality of $|\mathcal{I}(\underline{\lambda})| = N_{\mathrm{I}}(\underline{\lambda}) = \prod_{i=1}^{4} N_{\mathrm{I}}(\underline{\lambda}_i)$ is given by the Cartesian product of the individual one-dimensional grids

$$\mathcal{I}(\underline{\lambda}) = \{\mathcal{I}(\alpha) \times \mathcal{I}(\nu) \times \mathcal{I}(\theta) \times \mathcal{I}(\tau)\} \ . \tag{4.74}$$

As emphasized in the beginning of this section, the contemplations above are based on a rectangular pulse shape in a noise free channel. For different choices, the effect of pseudo symbol confusion in the CPM detector is not exactly given, but approximately. No adjustments to the grid were found to be necessary in these cases.

4.4.6 Grid Reduction Approaches

If the EM algorithm is carried out with each and every starting point from $\mathcal{I}(\underline{\lambda})$ until convergence, this will pose a large computational burden on the estimator. Section 4.5.1 will elaborate how this is avoided. But even if not every starting point must be

processed until convergence, reducing the cardinality of $\mathcal{I}(\underline{\lambda})$ will significantly reduce the overall computational complexity of the algorithm (cf. Section 5.9). Generally, this can be done by restricting the parameter ranges. If the external transmission conditions are not to be changed, this could be done by using pilot symbols. In contrast on relying solely on pilot sequences to obtain final estimates, the idea here in general is to obtain a reduced set of starting points that are accurate enough to guarantee the EM algorithm converging to the global maximum of the joint LLF.

For the following contemplations it is assumed, that the pilot symbols are processed by an unbiased and CRVB efficient estimator, that produces normally distributed estimation error (EE)s with zero mean. This distribution's variance is thus given by

$$\sigma^2_{\underline{\lambda}_i} = \mathrm{CRVB}(\underline{\lambda}_i|\underline{a}) \tag{4.75}$$

with the CRVBs all given in Chapter 3. Now the confidence intervals of the distribution above can be used to assess the parameter range that a certain number of pilot symbols is able to provide

$$\Delta_{\mathrm{Pilot}}(\underline{\lambda}_i) = 2n_{\mathrm{CI}}\sqrt{\mathrm{CRVB}(\underline{\lambda}_i|\underline{a})} \tag{4.76}$$

with n_{CI} as the confidence interval level.

Figure 4-8 visualizes how the use of pilot symbols can reduce the parameter range with a certainty of 99.7 %, i.e. $n_{\mathrm{CI}} = 3$ in bursts of length $N = 42$. At the intersection with the corresponding horizontal lines, a gain in form of less necessary grid points is reached. It might seem counterintuitive, that so many pilot symbols are required to reduce especially the CFO range, since a solid CFO estimation should be possible from it. The reason for this is, that rare estimates outside the confidence interval do not heavily influence the variance of such an estimator.

The actual gains from using pilots are clearly dependent on the waveform parameters, since they influence both the number of necessary grid points and the estimation performance. For the considered cases, pilots are most effective in reducing the number of CPO grid points and least so for the CFO. At the point, where all but one

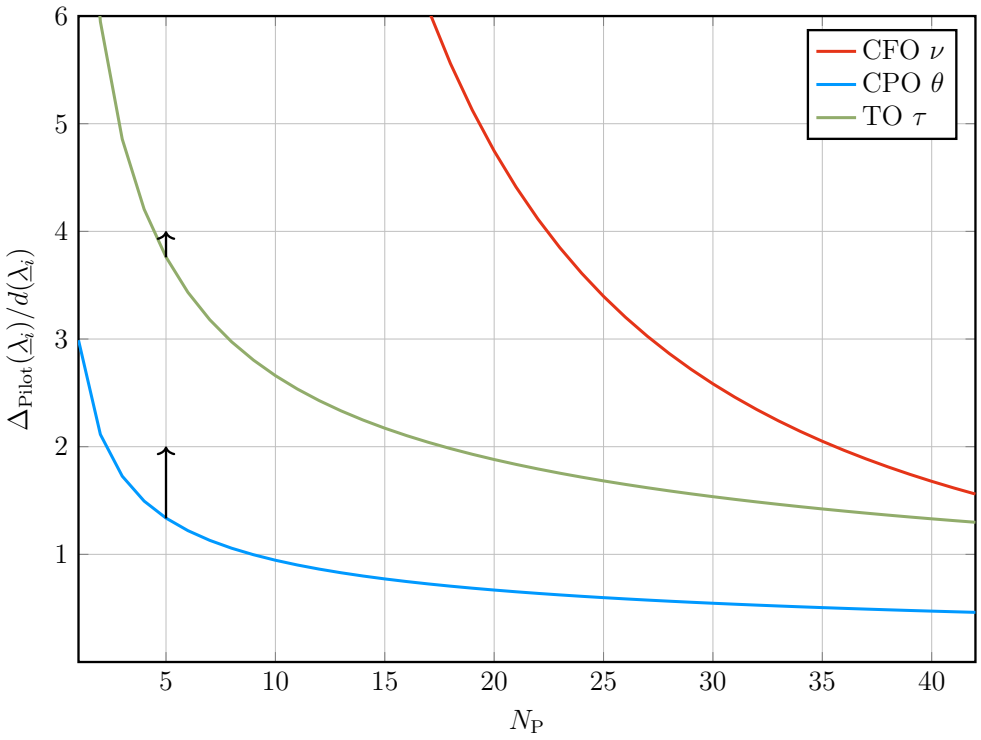


Figure 4-8: Reduction of the parameter range by pilot symbols according to (4.76) for CFO, CPO and TO. The waveform parameters are $h = 2/7$, $M = 2$, $L = 4$ and a pulse of Gaussian shape, which corresponds to one of the simulation waveforms given in 5.1. The curves show the reduction at $\alpha^2 E_{\text{S}}/N_0 = 10$ and $N = 42$ for the 3-$\sigma_{\underline{\lambda}_i}$ confidence interval 99.7 %. The vertical axis shows the ratio of reduced parameter range and the PMD and thus rounding up the curve values to the nearest integer gives the number of necessary grid points (as exemplary indicated by black arrows). Spending five pilot symbols would mean, that two and four grid points are necessary for the CPO and TO estimation in the scalar case and twice as much in the vector case (cf. Section 4.4.5). This setup has a grid cardinality of $|\mathcal{I}(\underline{\lambda})| = 456$ and by spending $N_{\text{P}} = 12$ pilot symbols, the grid size would reduce by a factor of 4.

grid point are eliminated by the pilot, the estimates themselves can probably be used to correct the signal without further refinement of a DD method. Of course the interval's confidence can be chosen less tight and some estimates that are not the ML solution can be accepted. At this point it is also noted, that independent from the use of pilot symbols a less dense grid can be chosen anytime some wrong estimates can be tolerated, which will result in an bit error floor. A possible scenario, where an estimator induces error floor causes not too much of an impact, is when an outer channel code is used that aggregates several bursts or in a fading environment. Furthermore, the pilot symbols cannot only be used to reduce the number of grid points, but can be directly incorporated in the EM algorithm due to their known posterior

probabilities. However, in this work grid reduction measures are not taken. Overall the above provided assessment of the use of pilot symbols can come in handy in system designing.

According to this contemplation and comparing the MSEEs of some NDA estimators investigated in [26–28], it is concluded, that they provide a 3-$\sigma_{\underline{\lambda}_i}$ confidence interval that at best comes close for simple parameter choices (e.g. binary, Q and L not too large) but never undercut the parameter range $\Delta(\underline{\lambda}_i)$. For that reason they are discarded as candidates for providing a reduced set grid points. Furthermore they are in general not robust to the other channel parameters, which will deteriorate their performance compared to the scalar cases investigated in above references. The FF estimator from (4.64) is robust in that regard, since it is energy based and thus not affected by CFO, CPO and TO.

4.5 Implementation Aspects

4.5.1 Algorithmic Synchronization Description

This section will gather all previous results and insights and formulate an algorithm whose goal it is to obtain the ML estimate of the channel parameter vector $\underline{\lambda}$. As a side constraint, the computational complexity shall not be unnecessarily high. In Section 5.9, the complexity will be assessed quantitatively. In Algorithm 1, the procedure of estimating $\hat{\underline{\lambda}}$ is given. In the first loop (lines 1-3), for every grid point the respective LL's are computed by the forward algorithm of the CPM detector (cf. (2.57)). At this point it seems intuitive, that the grid point with the highest LL is chosen as starting point for the EM algorithm. Unfortunately, a grid point nearer to the true $\underline{\lambda}$ can have a lower LL than another point, e.g. when the other point lies in a high local maximum as can be seen in Fig. 4-6. For this reason more than the most likely grid point must be considered in the second loop (lines 5-19) where the EM is carried out.

The selection of starting points for the second loop happens in line 4, where the computation of the LL does not need to be repeated as all necessary values were

Algorithm 1 Calculation of the ML estimate $\hat{\underline{\lambda}}$ by means of the EM algorithm.

Input: $r'(kT_0)$
Output: $\hat{\underline{\lambda}}$

1: **for all** $\underline{\lambda}_i \in \mathcal{I}$ **do**
2: $\quad \mathcal{L}(\underline{\lambda}_i)$ # Eq. (2.57)
3: **end for**
4: $\mathcal{I}_{\text{reduced}} = \bigcup_{\tau_k \in \mathcal{I}_\tau} \underset{\underline{\lambda}}{\text{argmax}}\ \mathcal{L}(\underline{\lambda} = [\bullet, \bullet, \bullet, \tau_k])$ # Select best $\underline{\lambda}$ for every τ_k of TO grid
5: **for all** $\underline{\lambda}_i \in \mathcal{I}_{\text{reduced}}$ **do**
6: $\quad \underline{\lambda}_{\text{old}} = \underline{\lambda}_i$
7: $\quad it = 1$
8: $\quad$ **while** $it \leq N_{\text{It}} + 1$ **do**
9: $\quad\quad$ **if** $it \leq N_{\text{It}}$ **then**
10: $\quad\quad\quad$ E-Step CPM: $\mathrm{P}_{\boldsymbol{\gamma}|\underline{\boldsymbol{r}}'}[\gamma_n|\underline{r}', \underline{\lambda}_{\text{old}}]$ # Eq. (2.48), (4.24)
11: $\quad\quad$ **else if** $it == N_{\text{IT}} + 1$ **then**
12: $\quad\quad\quad$ E-Step SCCPM: $\mathrm{P}_{\boldsymbol{\gamma}|\underline{\boldsymbol{r}}'}[\gamma_n|\underline{r}', \underline{\lambda}_{\text{old}}]$ # Eq. (2.48), (4.24), Sec. 2.6
13: $\quad\quad$ **end if**
14: $\quad\quad$ M-Step: $\hat{\underline{\lambda}}_i = \underset{\underline{\lambda}}{\text{argmax}}\ \mathcal{Q}(\underline{\lambda}, \underline{\lambda}_{\text{old}})$ # Sec. 4.3
15: $\quad\quad \underline{\lambda}_{\text{old}} = \hat{\underline{\lambda}}_i$
16: $\quad\quad it = it + 1$
17: $\quad$ **end while**
18: $\quad l_i = \mathcal{L}(\hat{\underline{\lambda}}_i)$ # Eq. (2.57)
19: **end for**
20: $\hat{\underline{\lambda}} = \underset{\hat{\underline{\lambda}}_i}{\text{argmax}}\ l_i$

already computed. For every TO in $\mathcal{I}_\tau$, the most likely CFO and CPO grid points are chosen. In doing so, the starting points have the best possible TO estimate available so far. As it was deduced in Section 4.3.1, it is crucial, that the auxiliary function is optimized with regard to the CFO before determining the TO. This sequential optimization is supported by the above described choice of starting points, so the CFO estimation can be made based on an adequately estimated TO.

A further way to determine only one appropriate starting point would be to compute the marginal LLFs

$$\mathcal{L}(\underline{\lambda}_k) = \sum_{\underline{\lambda}_l, \underline{\lambda}_m, \underline{\lambda}_n} \mathcal{L}\left(\left[\underline{\lambda}_k, \underline{\lambda}_l, \underline{\lambda}_m, \underline{\lambda}_n\right]\right) \tag{4.77}$$

for which sample values are already available through lines 1-3. The obtained marginal

LLFs are too inaccurate though and significantly more grid points' LL values would have to be computed and thus this approach is discarded.

The number of EM iterations carried out determines the accuracy of the end result as well as the computational complexity, so a criterion to stop at some point must be defined (cf. line 8). This can be e.g. the Kullback-Leibler divergence of the posterior probabilities converges to zero or a fixed number of iteration N_{It} is reached. In general, a fixed number of iterations is advantageous with regard to a guaranteed latency of the estimator. The suited number of $N_{\mathrm{It}} = 2$ will be determined in Section 5.2.

While it increases the computational complexity, the SC CPM setup decreases the number of symbol errors if the detector already provides reasonably reliable symbol estimates. EM as a decision directed estimator naturally gains from this and provides better estimates. As a compromise, the SC CPM setup is only applied in a third and last iteration, cf. lines 9 to 13. The effect of utilizing the channel code will also be investigated in Section 5.2. The maximization step is followed by the preparation of the new iteration. In line 20 the best estimate according to its LL is determined.

If the system's code rate is equal to one, the method of estimating the channel parameter vector will be called vector expectation maxmization (VEM). In this case line 12 achieves the same result as line 10 because the (de)interleaver and the channel code in the SC CPM setup are just shuffling the LLRs, but not changing their values. For code rates smaller than one, the estimation will be referred to as serially concatenated vector expectation maximization (SCVEM).

4.5.2 Efficient Implementation

The most computationally complex parts in Algorithm 1 are the calculation of the path metrics, the expectation and the maximization steps for which some details are elaborated next.

Calculation of Path Metric

The evaluation of (2.63) to compute the path metrics is an intensive task, but can be optimized. For easier comprehension, (2.63) is recalled and slightly reformulated

$$f_n^M(\underline{s}, \dot{\underline{s}}) \propto \exp\left(\frac{2\hat{\alpha}}{\hat{\sigma}_{w'}^2} \operatorname{Re}\left\{e^{-j(\hat{\theta}_0+2\pi\xi\hat{\tau})} \sum_{k=0}^{K-1} e^{-j2\pi\hat{\nu}(kT_0+nT)} r'(kT_0+nT) \cdot c_{\gamma_n}^*(kT_0-\hat{\tau})\right\}\right) \tag{4.78}$$

to assist the following contemplations.

First of all, the delay $\hat{\tau}$ is applying to the reference pseudo symbols, so the resampling operation on the already sampled received signal $r'(kT_0)$ is saved. Furthermore, no potentially manipulated noise statistics due to resampling in the received signal must be taken into account. The shifted references $c_{\gamma_n}^*(kT_0 - \hat{\tau})$ can be computed offline and stored for every TO grid point and generally in the desired TO granularity (cf. Section 4.5.2). The CFO compensation is better suited to be applied to the received signal, where it is a one-time operation on each sample opposed to when applying it to any possible pseudo symbol at any possible position nT. The term respecting the pseudo symbols' energy is dropped since it only must be considered in last line of algorithm and thus is of no significant relevance with regard to complexity contemplations. The result of the sum is just a complex number, so the phase rotation of $\hat{\theta}_0 + 2\pi\xi\hat{\tau}$ is of negligible computational complexity. Every required CPO in the grid can be applied to this particular complex number, i.e. the recomputation of the sum is not necessary in these cases. The same applies for the phase states of the CPM trellis, which are just phase shifts of the zero phase state $\Psi' = 0$. After taking the real part, the scaling due to the noise power and the FF is just one more real multiplication.

MAP Detection and LL Computation

The path metrics are needed for carrying out the MAP detection and thus the inference of the posterior probabilities used in the E-Step (lines 10 and 12). The forward part of it is used to obtain the LLs of each grid point (line 2) and the final decision on

the ML estimate (line 18). Both rely on the trellis representation of the CPM signal and profit from the efficient and accurate implementation presented in [77], where the processing is shifted to the logarithmic domain. For the sake of simplicity, further treatment is omitted here.

Expectation Step

The E-Step (line 10 as well as line 12) consists of two parts. First, the pseudo symbols' posterior probabilities have to calculated, which is done by the BCJR algorithm and second, the weighted reference symbols are calculated in (4.24). The time shifted reference symbols $\bar{c}^*(kT_0 - \tau_{\text{old}})$ can be reused (if not stored offline anyway) in (4.33) since the old estimate τ_{old} from the last iteration is applied here as well as for the path metric calculation. In (4.35), again a realization of delayed reference symbols are required whose preparation counts towards the E-Step for the conceptual reasons mentioned in Section 4.2. For the sake of saving computational power, $\bar{c}^*(kT_0 - \tau_{\text{old}})$ is reused once more. The resulting estimate hence must be treated as the difference $\hat{\tau} - \tau_{\text{old}}$.

Maximization Step

Four maximization operations comprise the M-Step in line 14. While two (CPO and FF) are closed form solutions and need no further consideration in terms of saving computational power, the estimation for the CFO and TO in (4.32) and (4.34) are basically grid searches and are subject to optimization. As mentioned in Section 4.3.1, the necessary operations are a DFT and a CCF, which can be both efficiently implemented by the use of the fast Fourier transform (FFT). The accuracy of both estimates (also given in that section) is lower bounded by the oversampling factor K and in the case of CFO the number of channel symbols N. Both can be improved if the FFT operations are zero padded and thus the result's resolution is improved by ideal interpolation. This approach has the advantage over other techniques (such as e.g. parabolic interpolation) because it does not introduce an interpolation bias [78].

The DFT and CCF are extensive grid searches, i.e. they give results in a much

larger range than the parameter ranges ($\Delta(\nu)T \ll K$ and $\Delta(\tau)/T \ll 2N$ for normalized CFO and TO, respectively). Spectral zoom methods as e.g. [79, 80] can remedy the introduced complexity when extensive grid searches need a high resolution. The number of frequency (or time) bins determines the complexity and depends on the desired estimation range and resolution. As for the range, it is sufficient to search in the width of one LLF maximum, i.e. the PMDs $d(\nu)$ and $d(\tau)$, as other maxima cannot be reached anyways due to getting stuck in local ones. In general, the resolution shall ensure that the introduced quantization error does not have any negative effect on the detector performance. In this work, the resolutions $\varepsilon^{-1}(\nu)$ and $\varepsilon^{-1}(\tau)$ shall be set such that the quantization error is one order of magnitude smaller than the value of the corresponding CRVBs ((3.19) and (3.21))

$$\frac{\varepsilon^2(\bullet)}{12} \stackrel{!}{=} 0.1 \cdot \mathrm{MCRVB}(\bullet) \; . \tag{4.79}$$

The bin spacing is therefore given as

$$\varepsilon(\bullet) = \sqrt{\frac{6}{5} \cdot \mathrm{MCRVB}(\bullet)} \tag{4.80}$$

and the number of required bins expresses as

$$N_{\mathrm{B}}(\bullet) = \left\lceil \frac{d(\bullet)}{\varepsilon(\bullet)} \right\rceil \; . \tag{4.81}$$

The spacing ε is directly proportional to $\sqrt{E_{\mathrm{S}}/N_0}^{-1}$, i.e. in noisy channel conditions, a high resolution (i.e. a low ε) is unnecessary. The number of bins $N_{\mathrm{B}}(\bullet)$ is even at a high SNR only in the double-digit range. The actual set of bins $\mathcal{B}(\bullet)$, that needs to be computed are eventually defined as

$$\mathcal{B}(\nu) = \left\{ \pm \left(\frac{d(\nu)}{2} - \frac{d(\nu)}{2N_{\mathrm{B}}(\nu)} \right), \pm \left(\frac{d(\nu)}{2} - \frac{3d(\nu)}{2N_{\mathrm{B}}(\nu)} \right), \dots \right\} + \nu_{\mathrm{old}} \; , \tag{4.82}$$

$$\mathcal{B}(\tau) = \left\{ \pm \left(\frac{d(\tau)}{2} - \frac{d(\tau)}{2N_{\mathrm{B}}(\tau)} \right), \pm \left(\frac{d(\tau)}{2} - \frac{3d(\tau)}{2N_{\mathrm{B}}(\tau)} \right), \dots \right\} \; . \tag{4.83}$$

As explained in Section 4.5.2, the CFO estimate is searched around its actual value (hence the offset ν_{old}) and for the TO only a delta is determined.

Due to the value of $N_{\text{B}}(\bullet)$, the Goertzel algorithm [80] is a good choice for an efficient implementation of these two maximization steps. In the case of CFO, the Goertzel algorithm calculates the required frequency bins of the DFT in (4.33). For computing the TO, it shall act as a temporal zoom method, as suggested by [28], and for that, the spectral representation of the CCF is computed as

$$\underline{X}_{r',\bar{c}} = \text{FFT}\left\{\underline{r}'\right\}_{\text{ZP}} \odot \text{FFT}\left\{\bar{\underline{c}}\right\}^*_{\text{ZP}} \tag{4.84}$$

with $\odot$ denoting the element wise vector multiplication. It is hereby assumed, that the FFTs are scaled such that the energies of the signal and its spectrum are equal. The subscript ZP to the FFT indicates zero padding to ensure linearity (in contrast to circularity) of the correlation and hence $\underline{X}_{r',\bar{c}}$ is of length N_{ZP}. Computing an inverse discrete Fourier transform (IDFT) of $\underline{X}$ would give the time domain CAF and can be realized by [81]

$$\underline{\chi}_{r',\bar{c}} = j\left\{\frac{1}{\sqrt{N_{\text{ZP}}}}\text{DFT}\left\{j\underline{X}^*_{r',\bar{c}}\right\}\right\}^* . \tag{4.85}$$

The scaling with the inverse square root of the signal's length is to ensure energy equivalence between frequency and time domain. While for the optimization of the auxiliary function with regard to the CFO ν and the TO τ, only the position of the maximum is relevant, the actual value is important for evaluating e.g. the Kullback-Leibler divergence as discussed in Section 4.1. Instead of computing (4.85) completely, once more the Goertzel algorithm can be used to evaluate the DFT at the desired time bins. The operation $j\{\bullet\}^*$ is just a switching of real and imaginary part of the respective complex numbers and thus of negligible computational effort.

4.6 Key Points of the Chapter

- The EM algorithm was introduced and its relation to statistical measures such as the Kullback-Leibler divergence and entropy was elaborated. Principal properties such as the iterative improvement of the estimate in the ML sense and the problem of local maxima in the LLF were described.

- For the considered case of FF, CPO, CFO and TO estimation in CPM bursts, the auxiliary function $\mathcal{Q}\left(\underline{\lambda}, \underline{\lambda}_{\text{old}}\right)$ was derived. The expectation step is comprised by the computation of the posterior probabilities $\mathrm{P}_{\boldsymbol{\gamma}|\underline{\boldsymbol{r}}'}\left[\gamma_n|\underline{r}', \underline{\lambda}_{\text{old}}\right]$ realized by the CPM MAP detector and the expected reference signal $\bar{c}^*(kT_0 - \tau_{\text{old}})$. For the maximization step, efficient expressions were derived. In the coupled case of CFO and TO, the estimations are based on some form of interpolation and a suitable order was proposed, whereas FF and CPO are uncoupled from all other parameters and are obtainable through closed form solutions.

- A closed form estimation expression of the complex noise power density N_0 was derived, but its investigation led to its discard.

- To counter EM's inherent problem of convergence to the nearest maximum in the LLF, which is possibly local, a grid of starting points $\mathcal{I}\left(\underline{\lambda}\right)$ was derived. The grid shall be as sparse as possible but still provide at least one starting point from which the global LLF maximum is reached. For this, the scalar LLFs of CPO, CFO and TO were analyzed and the minimal distance between two peaks in it were determined.

- For the arising computational burden of a large grid of starting points, a rule of thumb based on the CRVB was introduced which describes quantitatively the grid reduction possibilities by the use of pilot symbols.

- An algorithmic description in the form of pseudo code was introduced and possible efficient realization methods for the most computationally complex algorithm parts were suggested.

Chapter 5

Evaluation of the Proposed Estimation Algorithm

In this chapter, many aspects of the estimation approach defined in Algorithm 1 from Section 4.5.1 are examined. The VEM's convergence behavior and the effect of the channel decoder's deployment in the SCVEM setup are investigated in Section 5.2. To recall, the VEM estimator does not take a channel code into account, i.e. the code rate in such a system would be one, while the SCVEM setup utilizes the channel code. Both EM based methods are compared to the pilot based estimator from [38]. Despite its contrary nature (blind versus pilot based estimation), it constitutes the best suited comparison for the methods proposed in this work. Its estimation performance is reported to be optimal, whereas blind methods from the literature fall short of providing estimates of appropriate quality as shown in [26–28]. Features and restrictions of the pilot based method (from now on called vector pilot based (VPB)) are described in Section 5.3. To lay the groundwork for the following section, first the three considered waveforms are introduced in Section 5.1.1. To ensure a fair comparison in Sections 5.4 to 5.8 between a blind and a pilot based estimator the treatment of the efficiency penalty due to the pilot overhead must be resolved. Several approaches are considered depending on the respective evaluation and discussed in Section 5.1.2.

Significant parts of the Sections 5.3 to 5.7, 5.8.1 and 5.9 were published in [66],

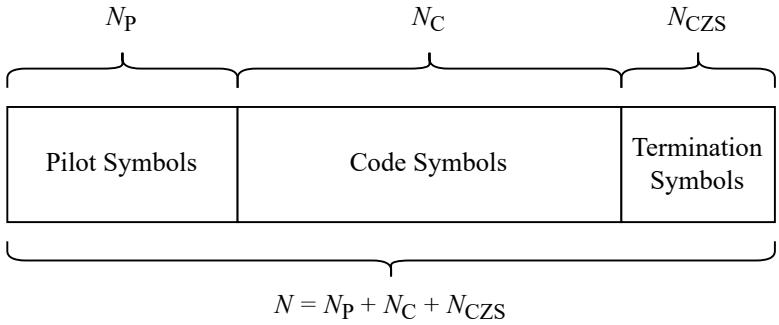


Figure 5-1: General symbol allocation in each burst. The sum of pilot, code and termination symbols equals the burst length in symbols.

largely for different waveform and system parameters. The discussion of the results is generally more detailed in this thesis.

5.1 Simulation Parameters

5.1.1 Parameters of the Considered Waveforms

In order to receive a comprehensive impression about the performance of the channel parameter estimator of Chapter 4, several different CPM waveforms shall be subject to examination. Table 5.1 lists the parameters for three different CPM waveforms.

The first part of the table includes a listing of the CPM specific parameters known from Section 2.1.1: h, M, L and $g(t)$. These are (in that order) the waveform's namesakes. Care has been taken to take a variety of sensible configurations into account, such as different modulation indices, binary and quaternary symbol alphabets, different partial response frequency pulses with two different pulse types. Hereby, 1/2-B3GA constitutes a GMSK-like modulation, 2/7-B4GA represents a waveform with a deeply coupled trellis and 4/7-Q2RC is a very efficient scheme in terms of occupied bandwidth and Euclidean distance of its symbols [14]. Due to this variety, the universal applicability of the presented estimator shall be emphasized. Figure 5-1 visualizes the symbol allocation in a burst.

The number of termination symbols depends on the pulse lengths $N_{CZS} = L - 1$, so it differs for each case. The number of available code symbols $N_C = N - N_{CZS} - N_S$

Table 5.1: Parameters of all simulated CPM waveform schemes. The burst duration and the two-sided signal bandwidth $2B_{99\%}$ are relevant for the channel assumptions in Sec. 2.3.1.

Waveform parameters	**1/2-B3GA**	**2/7-B4GA**	**4/7-Q2RC**
Modulation index h	**1/2**	**2/7**	**4/7**
Modulation order M	2 (**B**inary)	2 (**B**inary)	4 (**Q**uaternary)
Pulse length L	**3** (Partial resp.)	**4** (Partial resp.)	**2** (Partial resp.)
Frequency pulse type $g(t)$	**GA**ussian ($B_{\mathrm{GA}}T = 1/2$)	**GA**ussian ($B_{\mathrm{GA}}T = 1/3$)	**R**aised **C**osine
Termination symbols $N_{\mathrm{CZS}} = L - 1$	2	3	1
Symbol rate T^{-1}		42 kBaud	
Sampling rate T_0^{-1}		336 kHz	
Oversampling factor $K = T/T_0$		8	
Burst duration $T_{\mathrm{Burst}} = NT$		1 ms	
Normalized frequency pulse energy $G_2(0)T$	0.1755	0.1421	0.1875
Normalized one-sided signal bandwidth $B_{99\,\%}T$	0.51	0.35	1.16
Physical signal bandwidth $2B_{99\,\%}$	42.8 kHz	29.4 kHz	97.4 kHz
Modulation efficiency η_{M}	0.98 bit Hz^{-1} s^{-1}	1.43 bit Hz^{-1} s^{-1}	0.86 bit Hz^{-1} s^{-1}

is furthermore dependent on the number of pilot symbols N_P, which is 0 in case of the blind EM based methods. The number of binary code symbol $n_\mathrm{C} = N_\mathrm{C} \log_2(M)$ and binary information symbols k_C relate through the code rate $0 < R = k_\mathrm{C}/n_\mathrm{C} \leq 1$.

Physical properties are specified in the lower part with the symbol and sample rate being fixed throughout every case despite the resulting different bandwidths for the three waveforms. A fair comparison between the waveforms is not a primary goal, but can still be ensured when desired by equalizing the spectral efficiencies, which will be discussed in Section 5.1.2. The sampling interval T_0 is chosen such that the computer simulations meet the requirements of white noise in the receiver as outlined in Section 2.4. The one-sided signal bandwidth $B_{99\,\%}$ depends on the first and second table section, respectively and is together with the burst duration T_Burst essential for the channel characteristic contemplations in Section 2.3.1. The normalized frequency pulse energy $G_2(0)T$ influences the performance in the TO estimation and analytical expressions to calculate them are given in [17] and [65].

5.1.2 Setups for the Different Comparisons

- *Behavior of EM based methods*: In Section 5.2, solely the blind VEM and SCVEM estimators are investigated and no effort to ensure a fair comparison between blind and pilot based approaches is necessary. To keep it simple, a frame length of $N = 32$ is considered with uncoded transmission in the VEM ($R_\mathrm{VEM} = 1$) and coded transmission in the SCVEM case ($R_\mathrm{SCVEM} = 0.5$). It is noted, that since the VEM estimator does not utilize any channel decoder information, the estimation performance is generally independent of the code rate.

- *Equal performance limits*: For Sections 5.4 and 5.5, the number of transmitted symbols per burst is again set to $N = 32$. Since the pilot based (VPB) estimator needs a pilot sequence length dividable by four and [38] proposes a length of at least 32 symbols, this burst length seems to be a natural choice. The VEM setups are again uncoded ($R_\mathrm{VEM} = 1$), whereas the SCVEM setups are coded

($R_{\mathrm{SCVEM}} = 0.5$) with the number of information bits being rounded if necessary. In case of the VPB estimator, the whole burst consists of pilot symbols ($N_{\mathrm{P}} = N$) and it is noted, that its SE is therefore zero. This choice is reasoned with the goal to compare the methods at the same theoretical performance limit for which the same amount of symbols must be available to the respective estimator. The focus in these sections lies solely on the estimation performance, so the SNR is measured with regard to the available symbol energy E_{S} regardless of the information content of these symbols.

- *Equal code rates*: To measure the effective throughput in Section 5.6, both the SCVEM and the VPB system shall benefit from a similar as possible code, so the code rate is set for both cases to $R_{\mathrm{SCVEM}} = R_{\mathrm{VPB}} = 0.5$. Every burst contains a total of $N = 42$ symbols, for which $N_{\mathrm{P}} = 12$ symbols are spent on the pilot in the VPB system. The burst length is now chosen according to the system considerations summarized in Table 5.1 and the VPB pilot length shall be a sensible trade off between estimation capabilities and information content of the burst. The differences in spectral efficiency (the SCVEM system transmits more information than the VPB) do not need to be compensated, since they are integral to the effective throughput.

- *Equal energy per binary information symbol*: In Section 5.7 and 5.8, the number of transmitted symbols per burst $N = 42$ and pilot symbol for the VPB system $N_{\mathrm{P}} = 12$ are again chosen according to the reasoning of the last bullet point. The SNR metric is now set to E_{b}/N_0 instead of E_{S}/N_0 as above, so a comparison is fair in the sense of energy consumption per information. Due to different modulations and and such different symbol alphabets, signal bandwidths and termination constraints, the spectral efficiency will differ. To account for this, first the modulation efficiency defined as

$$\eta_{\mathrm{M}} = \frac{\log_2(M)}{2B_{99\,\%}T} \tag{5.1}$$

Table 5.2: System parameters for every waveform and estimator combination for uncoded transmission in Sections 5.7 and 5.8. The number of code and information are equal due to the fixed code rate $R_{\mathrm{VEM}} = R_{\mathrm{VPB}} = 1$. For the VPB estimator, the number of pilot symbols is set to $N_{\mathrm{P}} = 12$. The resulting SEs are given for each case.

$R = 1$	**1/2-B3GA**		**2/7-B4GA**		**4/7-Q2RC**	
	VEM	VPB	VEM	VPB	VEM	VPB
$k_{\mathrm{C}} = n_{\mathrm{C}}$	40	28	39	27	82	58
η_{S} in $[\mathrm{bit\,s^{-1}\,Hz^{-1}}]$	0.93	0.65	1.33	0.92	0.84	0.59

is introduced and given for each waveform in the last line of Table 5.1. It does not depend on potential coding or termination sequences and gives a baseline of the respective waveform's efficiency. The spectral efficiency in each case calculates as

$$\eta_{\mathrm{S}} = \eta_{\mathrm{M}} \cdot \frac{R \cdot N_{\mathrm{C}}}{N} \tag{5.2}$$

and naturally becomes smaller the more pilot symbols are spent. For uncoded transmission ($R_{\mathrm{VEM}} = R_{\mathrm{VPB}} = 1$), Table 5.2 lists the different SEs. The relation between the energies of one CPM symbol and one binary information symbol is given by

$$E_{\mathrm{S}} = E_{\mathrm{b}} \cdot 2B_{99\,\%}T \cdot \eta_{\mathrm{S}} \tag{5.3}$$

$$= E_{\mathrm{b}} \cdot \log_2(M) \cdot \frac{R \cdot N_{\mathrm{C}}}{N}\,. \tag{5.4}$$

- *Equal spectral efficiencies*: By fixing $\eta_{\mathrm{S}} = \eta_0 = 0.5\,\mathrm{bit}$, a fair comparison is made possible throughout waveforms and estimation methods by transmitting the same amount of information with the same amount of energy per time and occupied bandwidth. The formula for the number of information bits in each case is obtained by reordering (5.2) and respecting, that k_{C} must be an integer

$$k_{\mathrm{C}} = \left\lfloor \frac{\eta_0}{\eta_{\mathrm{M}}} \cdot N \cdot \log_2(M) \right\rfloor \tag{5.5}$$

Table 5.3: System parameters for every waveform and estimator combination for the fixed SE value $\eta_\mathrm{S} = 0.5\,\mathrm{bit}$ in Sections 5.7 and 5.8. The number of code and information bits and their corresponding code rate are given. For the VPB estimator, the number of pilot symbols is set to $N_\mathrm{P} = 12$.

$\eta_\mathrm{S} = 0.5\,\mathrm{bit}$	**1/2-B3GA**		**2/7-B4GA**		**4/7-Q2RC**	
	SCVEM	VPB	SCVEM	VPB	SCVEM	VPB
n_C	40	28	39	27	82	58
k_C	21		15		49	
$R = k_\mathrm{C}/n_\mathrm{C}$	0.53	0.75	0.38	0.56	0.60	0.84

with $\log_2(M) = n_\mathrm{C}/N_\mathrm{C}$. Table 5.3 gives the according code parameters for every waveform and estimator that are used in Sections 5.7 and 5.8.

Since the number of code and information symbols differs in each setup, a flexible channel code is required, which is realized by shortened Polar codes [52, 56]. The code is constructed with the Gaussian approximation method according to [82] and a systematic encoding procedure is chosen [54]. The decoder from [55] is used in which four internal iterations are carried out. To keep a unified block diagram the encoder and decoder block in Fig. 2-1 and 2-8 can be regarded as scramblers that still work according to the systematic Polar encoder rule even if no redundancy bits are added in the uncoded case. In all simulations, that use the SF channel model from Section 2.3.2, an adjusted SNR metric $\alpha^2 E_\mathrm{S}/N_0$ (respectively $\alpha^2 E_\mathrm{b}/N_0$) is used, where the symbol (respectively information bit) energy is scaled by the power decrease caused by the FF α. By this, the fading effect is diminished and a more meaningful assessment of the estimator performances is realized. It is noted, that several SNR definitions are used in the next sections that are apparent in each context. The linguistic term SNR will be used for every definition throughout these sections for the sake of readability.

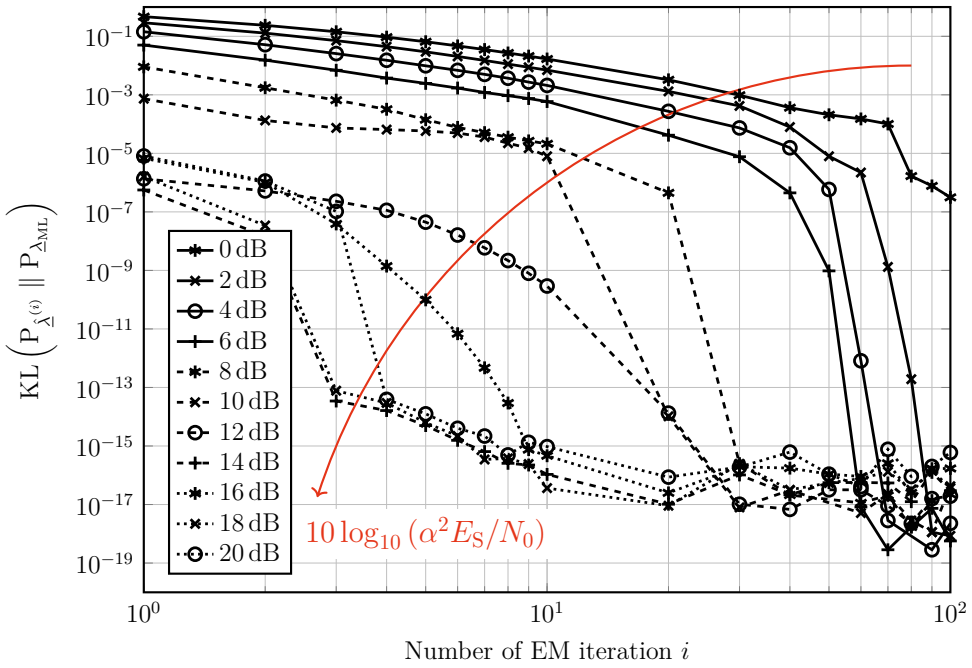


Figure 5-2: Kullback-Leibler divergence of posterior probabilities after the i'th EM iteration and the ML posterior distribution for several SNR points over the number of iterations for the 1/2-B3GA VEM system ($N = 32$). No use of potential channel decoder information was made.

5.2 Convergence Behavior and Influence of Channel Code

Section 4.1 gave an introduction on the progress that is made by each EM iteration in terms of similarity of the CPM symbols' posterior distribution compared to the distribution given the ML solution $\underline{\hat{\lambda}}$ of the channel parameter vector in Section 4.1. It was expressed as the Kullback-Leibler divergence $\mathrm{KL}\left(\mathrm{P}_{\underline{\hat{\lambda}}^{(i)}} \parallel \mathrm{P}_{\underline{\lambda}_{\mathrm{ML}}}\right)$ and it was stated, that the EM algorithm has converged and can be terminated with the optimal result when this divergence approaches zero. Figure 5-2 shows the quantity over the number of iterations for several SNR points when no use is made of the channel code and the corresponding decoder information.

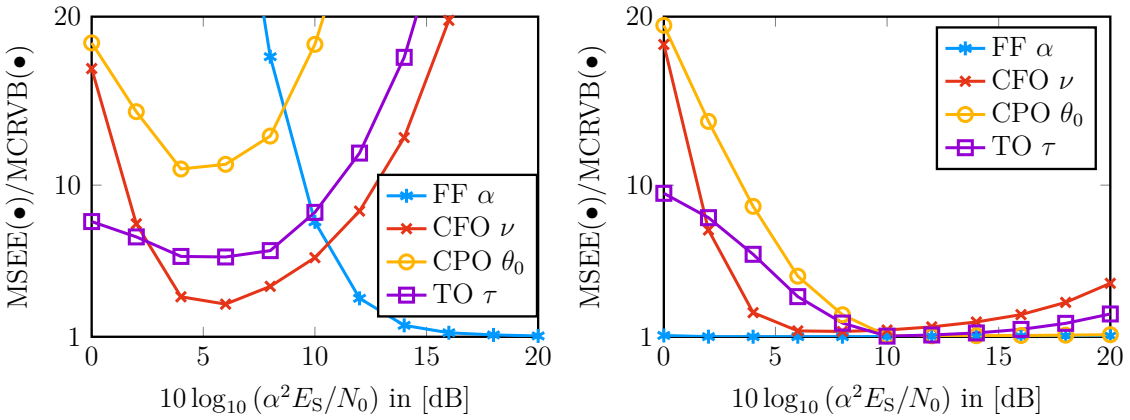
For any SNR the divergence will eventually approach zero when the ML solution is found. In the above case, even the curves corresponding to very low SNR fall to about 10^{-17} (which is basically zero considering the floating point arithmetic and its numerical noise) after several dozen EM iterations. At a higher SNR, the algorithm

converges faster with only about ten iterations.

Although the Kullback-Leibler divergence might be a very elegant metric to describe the progress in each iteration and the EM algorithm's complete convergence, its analysis is not directly helpful for determining a practical convergence threshold for two reasons. First, it is left open whether a divergence of zero is necessary to obtain reasonable estimates in the problem at hand. Second, the number of necessary iterations in order to approach a zero divergence is highly SNR dependent, but a fixed number for any operating point is usually desired in practical systems for a predetermined latency. A more direct approach is displayed in Fig. 5-3 which shows the ratio of the estimator's MSEE to the corresponding theoretical limit over an SNR range of 0 dB to 20 dB. It is recalled from (3.18) to (3.21) that the theoretical limits are indirect proportional to the SNR, so a constant line in the plots represents a decline both in the MSEE and MCRVB.

Figures 5-3a to 5-3d show the quotient before and after the first, after the second and after the tenth EM iteration. Before the first iteration is carried out, the current estimates are obtained by selecting the best grid value (cf. Section 4.4). The accuracy of these estimates does not increase with higher SNR values due to the fixed grid resolution, which is observable in the exponential curves in Fig. 5-3a. The FF estimate is not obtained by a grid evaluation but by the estimation rule (4.64) and its MSEE is able to approach its lower bound at around 17 dB. After one execution of E- and M-Step (Fig. 5-3b), the results already look promising with the FF MSEE being efficient over the complete displayed SNR range and all other estimates touching their respective bound at 10 dB. While the CPO does not suffer from a degradation afterwards, the slightly coupled parameters CFO and TO need one further iteration to hold the bound even at high SNRs (Fig. 5-3c). Further iterations do not bring any considerable improvement as Fig. 5-3d shows. For the results of this section, the total number of EM iterations is set to $N_{\mathrm{It}} = 2$ (cf. Algorithm 1).

As introduced in Section 4.5.1, the SC CPM setup can be used to incorporate the channel decoder information into the EM algorithm. This estimator setup is referred to as SCVEM. In order to evaluate the effect of the channel decoder, the mutual



(a) Normalized MSEEs before the first EM iteration.

(b) Normalized MSEEs after the first EM iteration.

(c) Normalized MSEEs after the second EM iteration.

(d) Normalized MSEEs after the tenth EM iteration.

Figure 5-3: Normalized MSEEs of all channel parameters after the i'th EM iteration for the 1/2-B3GA VEM system ($N = 32$).

information between the transmitted signal and the synchronized signal is measured for the three cases of VEM, SCVEM and the Genie-aided (G-A) parameters estimates. For this, the formula from [83] is used

$$\mathrm{I}\left[\underline{\boldsymbol{L}}_{\mathrm{A}};\underline{\boldsymbol{v}}\right] = \frac{1}{2}\cdot\sum_{v=0}^{1}\int_{-\infty}^{+\infty}\mathrm{p}_{\underline{\boldsymbol{L}}|\underline{\boldsymbol{v}}}(\xi|v)\cdot\log_2\frac{2\cdot\mathrm{p}_{\underline{\boldsymbol{L}}|\underline{\boldsymbol{v}}}(\xi|v)}{\mathrm{p}_{\underline{\boldsymbol{L}}|\underline{\boldsymbol{v}}}(\xi|0)+\mathrm{p}_{\underline{\boldsymbol{L}}|\underline{\boldsymbol{v}}}(\xi|+1)}d\xi\ . \tag{5.6}$$

The mutual information takes values between $0 \leq \mathrm{I}\left[\underline{\boldsymbol{L}}_{\mathrm{A}};\underline{\boldsymbol{v}}\right] \leq 1\,\mathrm{bit}$ and is measured by evaluations of the binary symbols' LLR histograms. The process $\underline{\boldsymbol{L}}_{\mathrm{A}}$ denotes the random variable of the vector $\mathrm{L}_{\mathrm{A}}(\underline{v})$ (LLRs fed into the channel decoder, cf. Section

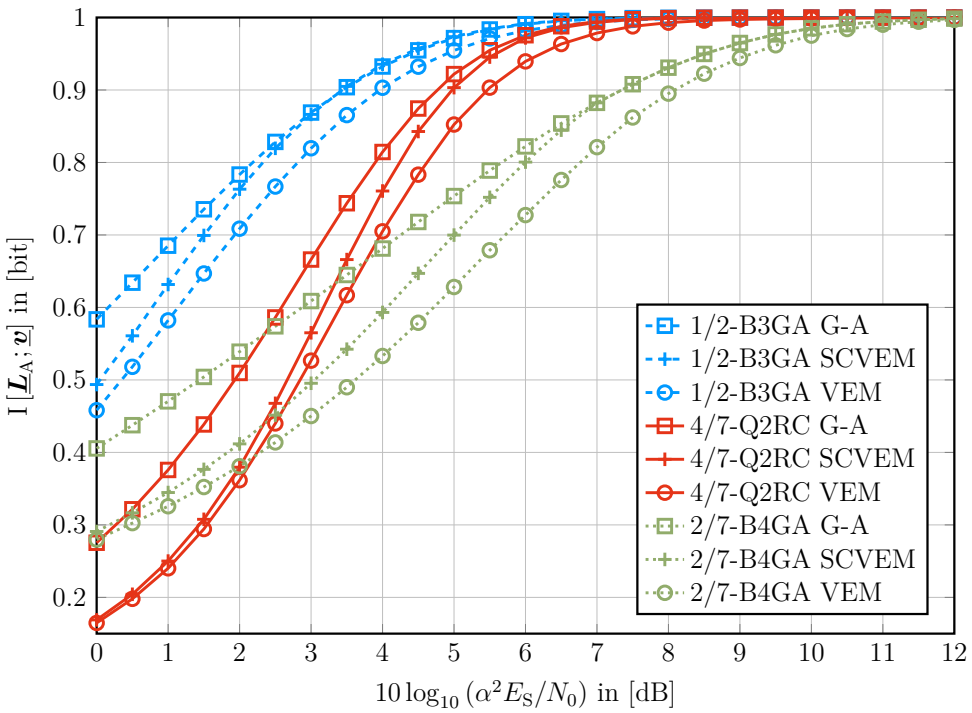


Figure 5-4: Mutual information I [$\underline{\boldsymbol{L}}_A$; $\underline{\boldsymbol{v}}$] of encoded bits and corresponding LLRs. The SCVEM, VEM and G-A ($\hat{\underline{\lambda}} = \underline{\lambda}$, $\hat{\sigma}^2_{w'} = \sigma^2_{w'}$) are compared for each of the three waveforms. The bursts length is $N = 32$.

2.6) and $\underline{\boldsymbol{v}}$ the random variable of encoded bits $\underline{v}$ (cf. Fig. 2-1). Figure 5-4 shows the mutual information for every waveform and the three above described parameter estimation methods. It is noted, that the channel decoder information in the SCVEM setup was only used to obtain better parameter estimates and not to improve the detector performance by providing a priori information to it after the correction of the channel influence. This is done in order to only assess the improvement in estimation performance.

The displayed mutual information serves as a measurement of how reliable the CPM symbols estimates produced by the CPM MAP detector are. For every case it approaches its maximum 1 bit for high SNRs. In general, the information in the G-A case is higher than in the SCVEM setup than in the VEM setup, which is an intuitive result. It is expected, that every parameter estimate has become reliable when the mutual information of the estimator curves approaches 1 bit. This is the case at around 6 dB, 7 dB and 11 dB for the 1/2-B3GA, 4/7-Q2RC and 2/7-B4GA SCVEM

systems, respectively. This statement is also confirmed by the results presented in Section 5.5. It can be observed that each curve approaches the G-A case a few decibels earlier, so that it can be concluded that the SCVEM estimator is operating very near to the optimum. In case of the uncoded VEM systems, the 1 bit mark is approached about 0.5 dB to 1 dB later. At this point, it is mentioned, that a mutual information very close to 1 bit can still entail significant error rates [83]. So contrary to the above drawn conclusion, these systems will generally only be able to provide reliable channel parameter estimates at higher SNRs. The explicit reasons for this are outlined in Section 5.5.

5.3 Notes on the VBP Estimator

For the comparison with the VPB estimator, a few remarks must be made:

- *FF estimation*: Since [38] does not include a method for estimating the FF, the following estimation rule is introduced for the VPB method (which proved to be optimal in the simulations)

$$\hat{\alpha} = \frac{|\Lambda_1(\hat{\nu})| + |\Lambda_2(\hat{\nu})|}{N_\mathrm{P} E_\mathrm{S}} \,. \tag{5.7}$$

  To avoid confusion with this thesis's notation, the functions λ_i from [38] were capitalized to Λ_i. These two functions basically constitute a cross correlation of the received signal and the known pilot sequence of two mutually exclusive signal segments.

- *CFO estimation*: Due to the nature of the pilot in [38] being a preamble, the estimated and corrected CFO introduces a phase offset at the beginning of the data block, if the estimate is not exactly correct. So, depending on the system's noise power, a smaller or larger phase offset will definitely be present at the start of the information block. To combat this effect, [38] suggests the use of a DD phase and timing tracking loop and shows, that the residual CFO estimation error has virtually no effect on error rates with this approach. In order to

estimate the CFO in a symmetric observation window, the time reference in the VPB system is set tot $t_0 = N_\mathrm{P}T/2$ (cf. Section 3.3). It is noted, that [38] is able to estimate a CFO range of $\nu T = [-0.5, +0.5]$ which would induce a high computational complexity in the EM methods due to the need of many CFO grid points. Though a range this large is arguably not of practical relevance (cf. Section 2.3.2).

- *TO estimation*: Contrary to the statement in [38], that a timing offset in the range of $[-0.5T, +0.5T]$ (also cf. (2.36)) can be estimated by a closed form expression, this is not necessarily the case for arbitrary CPM parameters, as it is shown in Appendix D.1. The maximum TO that can theoretically be estimated by the VPB method is

$$\left|\frac{\tau_\mathrm{max}}{T}\right| = \frac{1}{2h(M-1)} \tag{5.8}$$

which can be lower than 0.5 for waveforms with high modulation indices and orders. In practice, the range should be even lower because of noise influences and the range $[-0.75\tau_\mathrm{max}, +0.75\tau_\mathrm{max}]$ is set accordingly including headroom. It is noted, that in [38], the simulated CPM parameters provide enough headroom to be able to estimate the range in (2.36), which is probably the cause why this problem was not addressed in this work. Figure 5-6 visualizes the effect explained above and waiving the TO range reduction would lead to a useless 4/7-Q2RC VPB curve in Fig. 5-11. The VEM method is not limited with regard to the TO range, but profits of the lower range by having potentially less grid points and thus a lower computational complexity.

5.4 Mean Estimation Value

This section analyzes the MEV of the estimators in the SF channel of noteworthy cases. Both VPB and SCVEM follow ML oriented approaches and thus should be unbiased which is the case if the MEV asymptotically equals the true parameter, i.e.

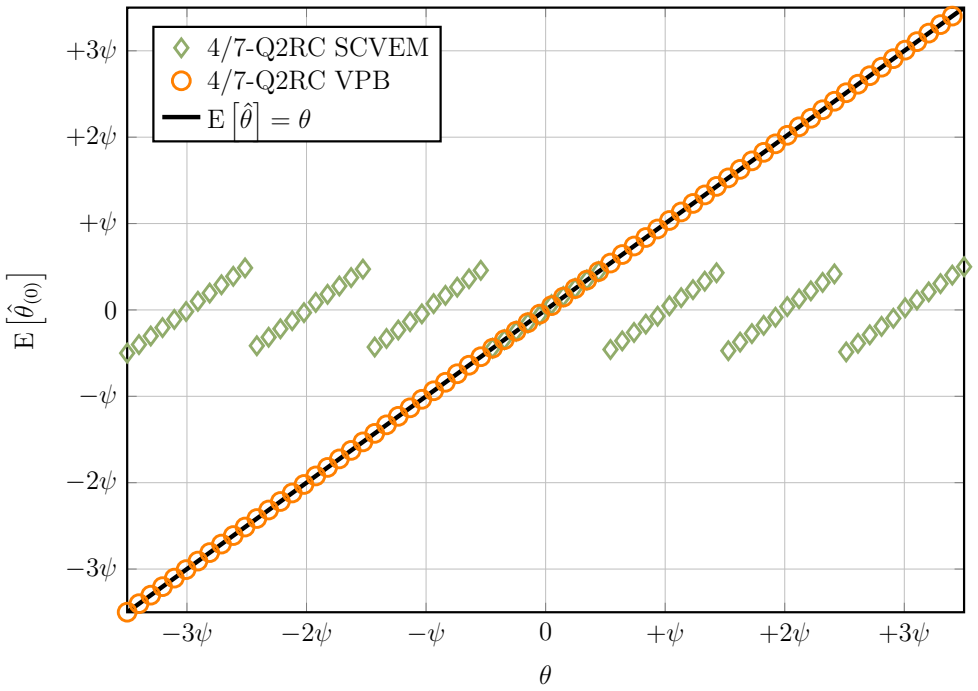


Figure 5-5: CPO MEV E $\left[\hat{\theta}_0\right]$ and E $\left[\hat{\theta}\right]$ of SCVEM and VPB, respectively, in the SF channel for the 4/7-Q2RC waveform ($N = 32$) at $10\log_{10}\left(\alpha^2 E_S/N_0\right) = 15\,\text{dB}$ compared to the bisector line corresponding to an unbiased estimator. This graph is taken from the author's work in [66].

E $\left[\hat{\lambda}_i\right] - \lambda_i = 0$. The cases of the CPO θ and TO τ for the 4/7Q2RC waveform are shown in Fig. 5-5 and 5-6.

The VPB estimator provides unbiased estimates over the whole CPO range $\theta \in [-\pi, +\pi)$ without any signs of unfavorable values. In case of the SCVEM estimator it becomes clear, that the MEV is periodic with the phase state angle ψ. This means that only the fine offset within a phase state is estimated and not the actual CPO that occurred in the channel. For the reason outlined in Section 2.5.2, this is an intended behavior and for that the SCVEM is still considered to provide unbiased CPO estimates. It is noted, that both estimators are prone to handling CPO values in the vicinity of their corresponding range edges in a wrong way, i.e. that $-\pi$ is handled as $+\pi$ or $-\psi/2$ is handled as $+\psi/2$, respectively. This is completely unproblematic in terms of synchronizing because of the equality $e^{-j\pi} = e^{+j\pi}$ and CPM's rotational invariance, respectively. For the displaying of the results in this section, the estimates were mapped in the correct interval to avoid meaningless ripples in the curves. The

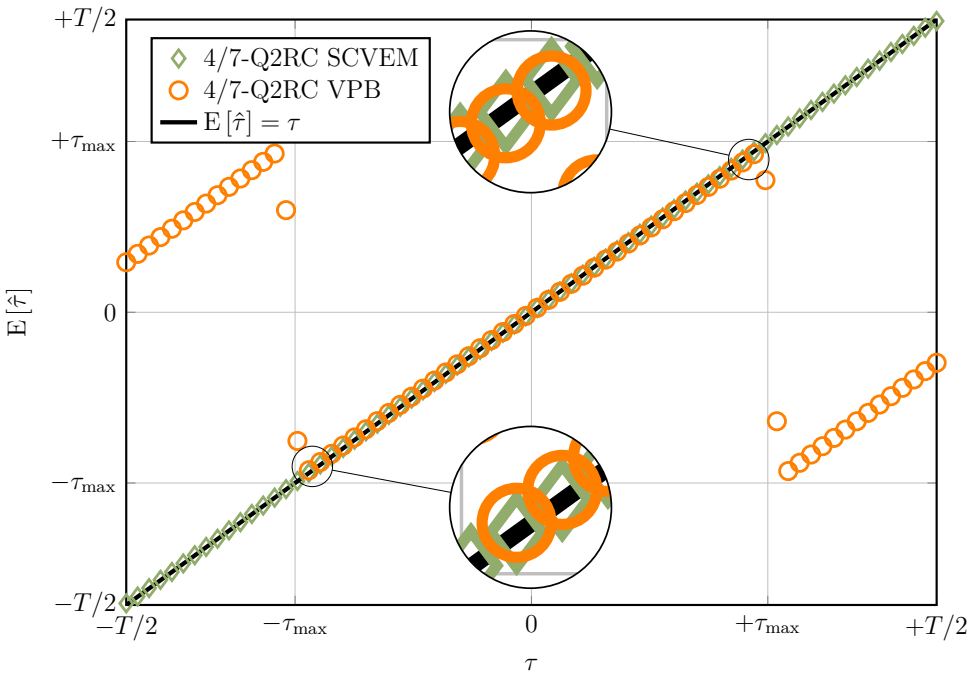


Figure 5-6: TO MEV $\mathrm{E}[\hat{\tau}]$ of SCVEM and VPB in the SF channel for the 4/7-Q2RC waveform ($N = 32$) at $10\log_{10}\left(\alpha^2 E_\mathrm{S}/N_0\right) = 15\,\mathrm{dB}$ compared to the bisector line corresponding to an unbiased estimator. This graph is taken from the author's work in [66].

same applies to Figure 5-10 in Section 5.5.

The TO MEV of the SCVEM estimator matches perfectly with the angle bisector line that indicates unbiasedness. The VPB estimator's curve also gives (almost) unbiased estimates within the range stated in (5.8). Through the magnifications Fig. 5-6, it can be observed, that the estimates are a little skewed towards zero. This effect is very small though and no effects in error rates are expected by this. Outside of that range, the estimates are shifted precisely by $\pm 2\tau_{\max}$ due to the estimator's ambiguity explained in Appendix D.1. It is emphasized that in contrast to the CPO case, this cannot be considered as intentional behavior, since there lies no natural periodicity or the like in time offsets. To avoid the resulting large deteriorations in estimation performance, the TO range in these cases is limited as mentioned in Section 5.3.

An, at the first glance, odd behavior can be observed when comparing the expected FF error of both EM methods. Figure 5-7 shows the errors over the SNR for the 4/7-Q2RC waveform, where the behavior is most pronounced.

The VEM produces positively biased (i.e. estimated too large) FF estimates for

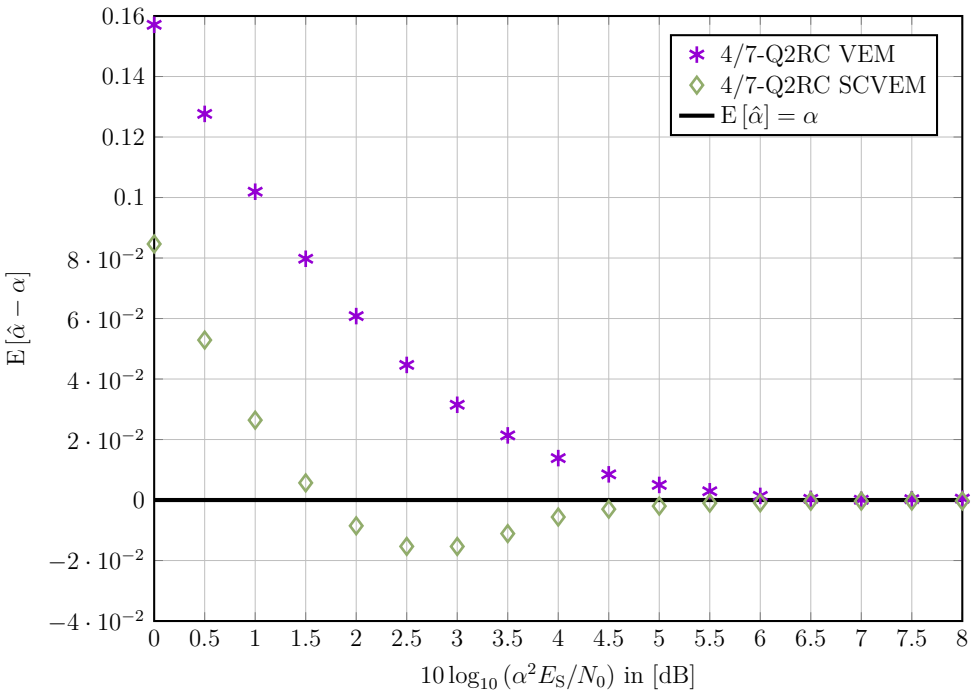


Figure 5-7: Expected FF error $\mathrm{E}[\hat{\alpha} - \alpha]$ of VEM and SCVEM estimators in the SF channel for the 4/7-Q2RC waveform ($N = 32$) over $10 \log_{10}(\alpha^2 E_\mathrm{S}/N_0)$ in [dB].

low SNRs. The reason for this is twofold. First, the coarse estimator (4.64) is recalled, which was reasoned to produce too large estimates. Second, the noise power estimator (2.45) produces too low estimates at low SNRs which are then used in the initial FF estimation in (4.64), which produces even larger estimates as a consequence. The VEM cannot fully correct this, because the symbol estimates from the CPM detector are not reliable enough. At around 6 dB, the estimator becomes unbiased. This behavior is clearly visible in Fig. 5-7.

The SCVEM estimator principally suffers from the same behavior, since it is identical to the VEM up to the second last EM iteration. The unreliable CPM detector estimates are fed to the channel decoder in the last iteration, which produces (even more) unreliable estimates on the binary information symbols (cf. Section 2.6). This detrimental a priori information leads to even less reliable detector estimates which are used to compute the reference symbols in (4.24). The resulting expected reference signal is the input to the CAF in (4.28) which will be of lower power because of the less matching reference. Eventually the CAF is used in (4.38) to obtain the FF

estimate in the E-Step, which will be estimated too low consequently. This behavior is visible in the SCVEM curve being set down from the VEM with estimates even being negatively biased until the SNR becomes high enough to clear up this specific FF estimator characteristic.

5.5 Mean Square Estimation Error

The CRVB discussed in Chapter 3 is arguably the most important benchmark for unbiased estimators, to which the comparison candidates belong in general with the exemptions discussed in last section. This section shows subsequently the MSEE over the SNR of the VEM, SCVEM ($R_{\mathrm{SCVEM}} = 0.5$) and VPB ($N_{\mathrm{P}} = N$) estimators for the four considered channel parameters FF α, CFO ν, CPO θ and TO τ in Figs. 5-8 to 5-11, respectively. By only transmitting pilot symbols in the VPB system, the candidates can be compared to the same theoretical limits. It is noted, that this results in a spectral efficiency of zero for the VPB system though. Splitting the available symbols into pilot and data, this estimator would still perform optimal, but at a higher (i.e. worse) MSEE level. Several blind feedforward estimators were compared to the VEM method in [26–28] and were found to provide an inferior performance. For the sake of the following plots' clarity, these additional curves are left out and it is referred to these references.

All candidates produce efficient FF estimates in the most relevant mid and high SNR regions and thus are all suited for estimation. This is because of the robust (with regard to the CPM detector's symbol estimates) estimation rule in the maximization step. Two observation can be made that are valid for all waveforms but are most prominent in the more complex 4/7-Q2RC case. First, the VPB estimates' MSEE bends up a bit at high SNRs which is due to the simplifying design assumptions in [38] that introduce systematic errors in the estimation of every channel parameter. These are relatively small (so that the effect is only visible in the high SNR region) and are thus not expected to have any considerable impact on bit or frame error rates, which is confirmed by the error rate simulations in Section 5.7. In contrast,

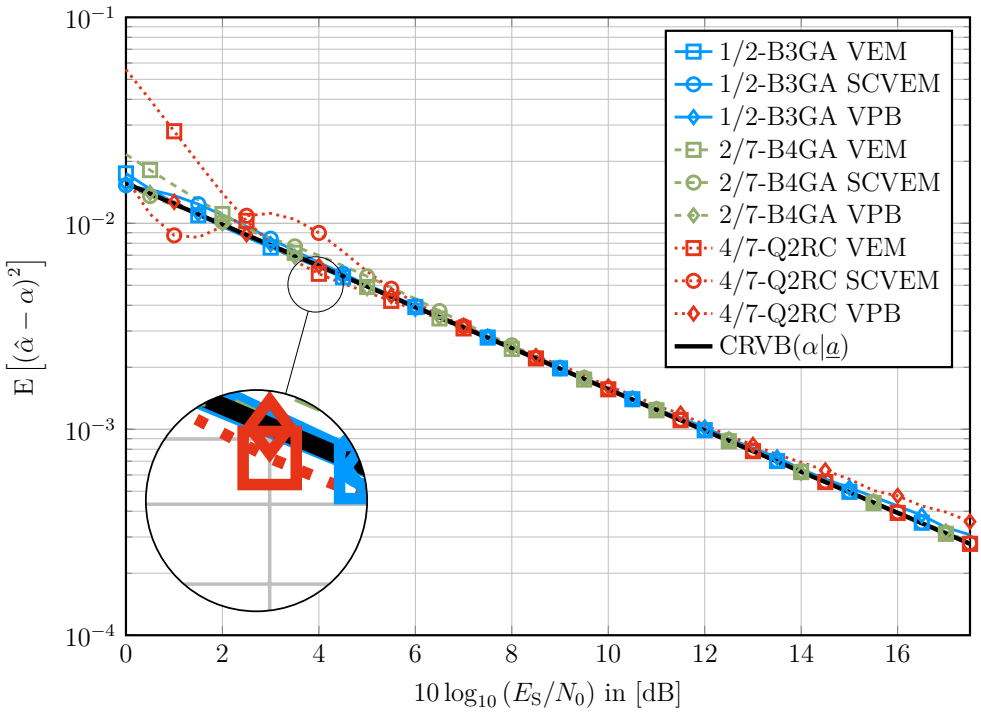


Figure 5-8: Normalized FF MSEEs $\mathrm{E}\left[(\hat{\alpha}-\alpha)^2\right]$ over E_S/N_0 of the VEM, SCVEM and VPB estimators in the SF channel for three waveforms ($N = 32$) compared to the theoretical limit.

this does not happen for the EM based methods, since no simplifications were made in the algorithms derivation and the interpolations were computed optimally (cf. Section 4.5.2). For the second observation, Fig. 5-7 is recalled where the expected estimation errors for VEM and SCVEM are shown. Small positive biases ($\leq 5 \cdot 10^{-2}$) seem to have a positive influence on the MSEE. This case occurs for VEM between 2.5 dB and 6.0 dB and for SCVEM between 0.5 dB and 1.6 dB. These are roughly the ranges, where the estimates' MSEE lies under the CRVB, which is possible for biased estimators. Larger or negative biases causes a clear deterioration in terms of MSEE. Why small positive biases are advantageous is explained with the underlying Rayleigh distribution of the FF, which produces large realizations at a low probability. By shifting the estimates moderately towards this distribution's tail does seemingly decrease the MSEE.

Figure 5-9 shows the normalized frequency MSEE. Even the VPB estimator is not able to produce reliable results at extremely low SNRs, which originates from a natural uncertainty regarding the LLF's global maximum even when a pilot sequence

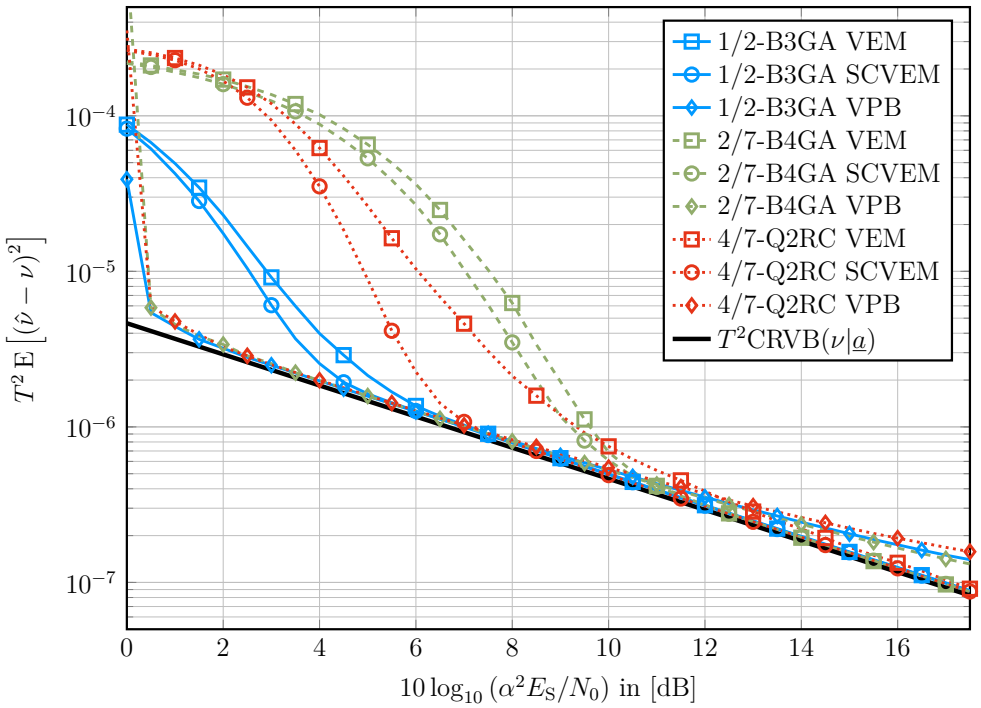


Figure 5-9: Normalized CFO MSEEs $T^2\,\mathrm{E}\left[(\hat{\nu}-\nu)^2\right]$ over $\alpha^2 E_\mathrm{S}/N_0$ of the VEM, SCVEM and VPB estimators in the SF channel for three waveforms ($N = 32$) compared to the theoretical limit.

is used. While the estimates are reliably from fairly low SNRs on, the EM estimates are so only from the point on when the CPM detector produces reliable symbol estimates. This behavior is rather typical for DD methods and is not a problem in general, since accurate estimates become available when the error rate of the system falls and thus its overall performance is of use. This reasoning is also valid for the CPO and TO cases. The gap between the SCVEM and VEM curves is intuitively explained by the better detector performance when a priori information in form of the channel decoder output is fed back to the detector. Especially in the case of 4/7-Q2RC, the estimation performance seems to gain from the decoder deployment. The reason for this will become apparent in a moment.

In the case of CPO, the MSEEs are given with respect to θ for the VPB method and to θ_0 for the EM based estimators. The insights derived from this plot are similar to the CFO case and do not need restating.

In [14], it is reported, that CPM configurations with large index denominator Q are hard to synchronize in phase because of the similar phase states. Is is noted, that

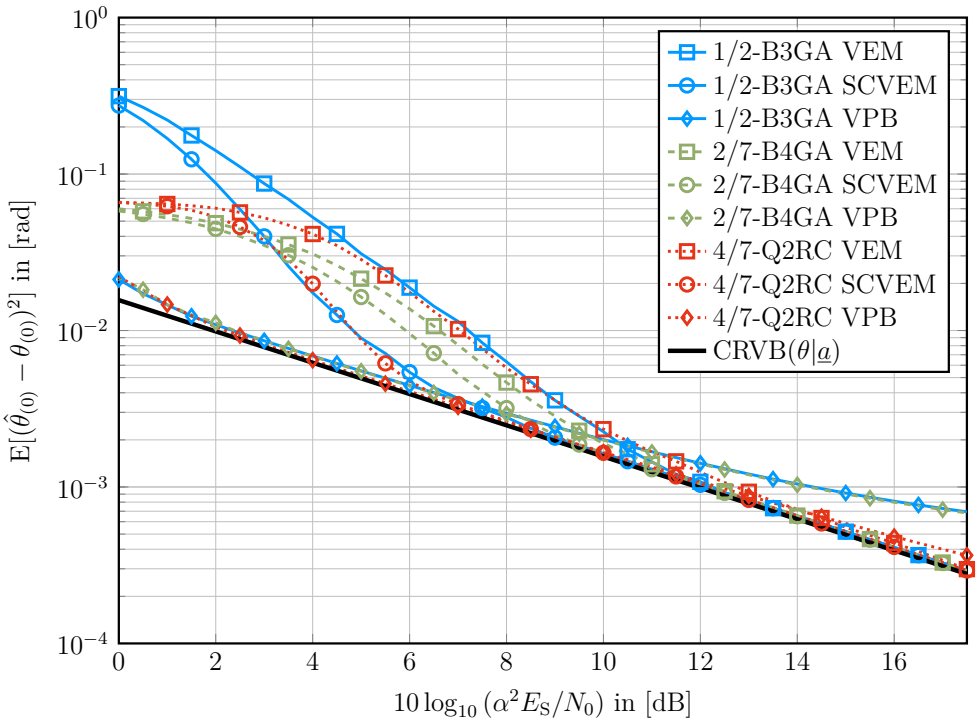


Figure 5-10: Normalized CPO MSEEs $\mathrm{E}\left[\left(\hat{\theta}_0 - \theta_0\right)^2\right]$ (EM based methods) and $\mathrm{E}\left[\left(\hat{\theta} - \theta\right)^2\right]$ (pilot based method) over $\alpha^2 E_\mathrm{S}/N_0$ of the VEM, SCVEM and VPB estimators in the SF channel for three waveforms ($N = 32$) compared to the theoretical limit.

the proposed estimator does not suffer from this as it provides reliable estimates even for $Q = 7$.

Due to the TO CRVB being dependent on the waveform parameters as well as on the actual transmitted symbol sequence, Fig. 5-11 looks more crowded than the plots before. In [64], an optimal pilot sequence $\underline{a}_\mathrm{opt}$ was derived, which is also used here with the respective bounds denoted accordingly. The other parameters' estimation is independent on the transmitted sequence with the bounds of a random sequence $\underline{a}$ and the optimized sequence $\underline{a}_\mathrm{opt}$ being equal. The insights though stay the same with the VPB curves bending up at high SNRs and the VEM and SCVEM curves approaching the CRVB as soon as the CPM detector is able to provide reliable estimates. One exemption constitutes the VEM estimator in case of the 4/7-Q2RC waveform, which suffers from high MSEEs even at high SNRs. This stems from TOs near to $\tau = \pm T/2$ being estimated as $\hat{\tau} = \mp T/2$ occurring with a probability of about 10^{-4}. The signal is basically locked in the correct symbol clock, but with an offset of ± 1 symbol. This

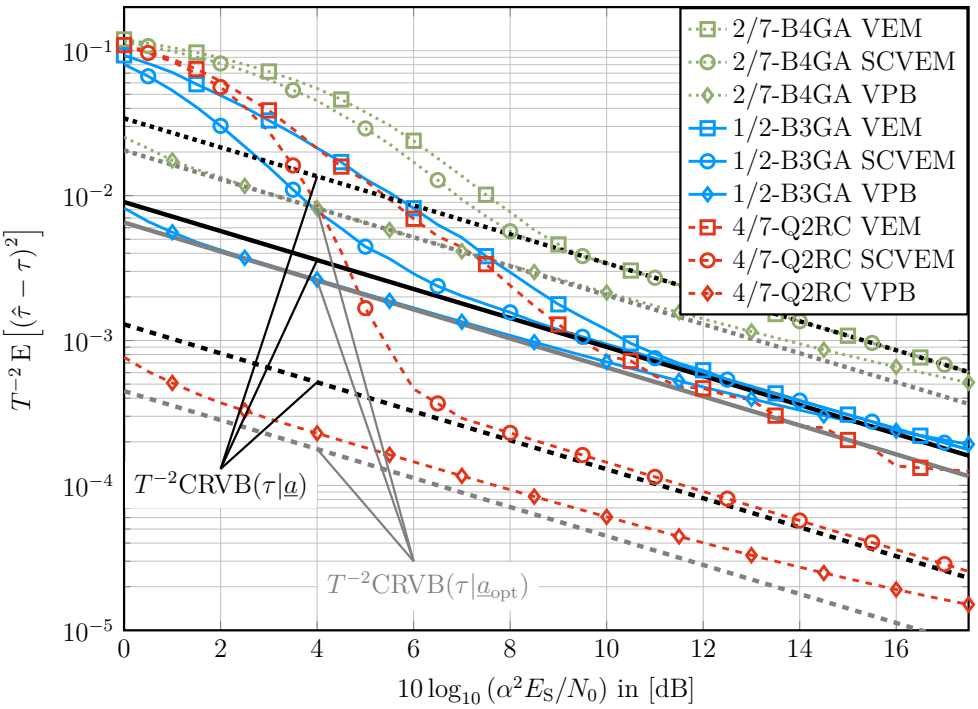


Figure 5-11: Normalized TO MSEEs $T^{-2}\,\mathrm{E}\left[(\hat{\tau}-\tau)^2\right]$ over $\alpha^2 E_{\mathrm{S}}/N_0$ of the VEM, SCVEM and VPB estimators in the SF channel for three waveforms ($N = 32$) compared to the theoretical limits. Each waveform has its own CRVB depending its waveform parameters and furthermore the optimized pilot sequence from [38] $\underline{a}_{\mathrm{opt}}$ gives a lower bound than a random sequence (denoted here by $\underline{a}$).

problem was also reported in [28] and resolved by utilizing the syndrome of the burst's code word. In this work, this is inherently covered by the SCVEM estimator, which utilizes the soft information of the channel decoder. For soft (i.e. large L) phase pulses, the TO estimation is expected to be hard due to similar phase trajectories to other non-delayed symbol sequences. This is reflected in the higher theoretical bound for e.g. 2/7-B4GA in Fig. 5-11, but the proposed estimator shows no general weaknesses with regard to such pulses.

Lastly, it is referred back to Fig. 5-4 in Section 5.2. The information of the VEM synchronized signal has a distance of about 0.5 dB to 1 dB to the SCVEM case. This distance is observable in Figs. 5-9 to 5-11 only for the 2/7-B4GA waveform. In the 1/2-B3GA and especially 4/7-Q2RC case, the TO estimates profit greatly from the decoder information for the reason stated above. As stated in Chapter 3, the joint estimation of all parameters is principally mutually decoupled by the use of a symmetric observation window. This implies, that each parameter can be estimated

correctly even if the others are faulty estimated, though this only true for small estimation errors. In consequence, the VEM system's inadequate TO estimates lead to considerable performance degradations in the CPO and for the case of 4/7-Q2RC even the CFO case.

The results presented in this section shall show what is generally feasible with the considered estimators. If a system's operating point is e.g. at a rather low SNR point, computational effort could potentially be saved, if the CRVB at a high SNR point does not need to be reached.

5.6 Effective Throughput

The effective throughput is defined as the number of binary information symbols that are transmitted in a correctly decoded burst. Moreover, it is normalized by the occupied bandwidth and the burst duration, which gives the formal definition

$$\mathcal{T}_{\text{eff}} = \frac{(1 - \mathcal{E}_{\text{Frame}}) \cdot k_{\text{C}}}{2B_{99\,\%}T_{\text{Burst}}} \tag{5.9}$$

$$= (1 - \mathcal{E}_{\text{Frame}}) \cdot \eta_{\text{S}} \ . \tag{5.10}$$

with the second line following from $T_{\text{Burst}} = NT$, $k_{\text{C}} = \log_2(M) \cdot R \cdot N_{\text{C}}$, (5.1) and (5.2). The probability of the event of a wrongly decoded frame $\mathcal{E}_{\text{Frame}}$ is obtained as the frame error rate (FER) through Monte-Carlo simulations. An upper bound for the throughput can be obtained by applying Shannon's sphere packing bound [84] to the equation above. This bound constitutes the performance in terms of frame errors of an optimal spherical code for a continuous input channel, which itself is a lower bound to the frame error rate (FER). Figure 5-12 shows the normalized effective throughput $\mathcal{T}_{\text{eff}}$ for both waveforms and estimation systems for the fixed code rate of $R_{\text{VEM}} = R_{\text{VPB}} = 0.5$ as stated in Section 5.1.2 and the respective upper bounds obtained through the sphere packing bound.

All throughput curves approach their spectral efficiency according to (5.10) as an asymptote at high SNRs, i.e. when the FER approaches zero. The VPB has

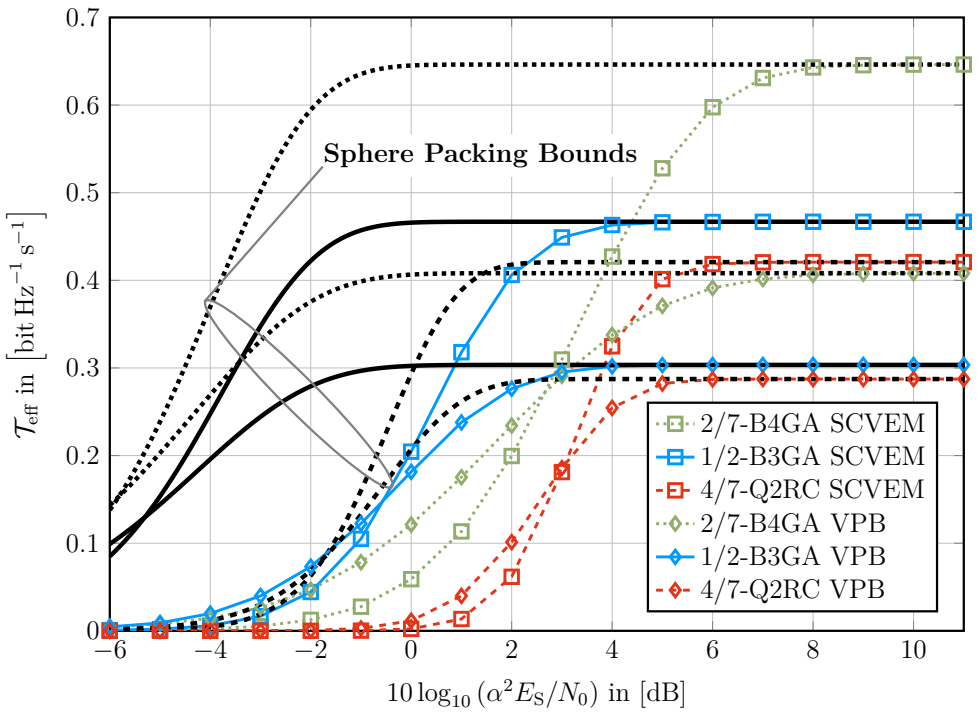


Figure 5-12: Normalized effective Throughput $\mathcal{T}_{\text{eff}}$ with code rates being fixed to $R_{\text{SCVEM}} = R_{\text{VPB}} = 0.5$ over $\alpha^2 E_b/N_0$ in the SF channel for the three waveforms ($N = 42$). The throughput, when the sphere packing bound is applied, upper bounds the performance. The respective corresponding sphere bound curve can be deducted by the asymptotic values.

generally a lower probability of a wrongly decoded frame throughout all SNR regions because of their known pilot structure compared to the EM setups. Though it can only convert this advantage into a slightly higher throughput for very low SNRs and thus its practical relevance is limited. The asymptotes, i.e. the spectral efficiencies of the systems are as expected found to be significantly higher in the SCVEM cases. Comparing the curves to their respective upper bound, it can be observed, that the VPB performs at about 1 dB closer to the theoretic limit for 2/7-B4GA and 4/7-Q2RC whereas both variants show the same gap to the limit in 1/2-B3GA. From this can be derived, that especially for the last waveform, the SCVEM seems to work very well.

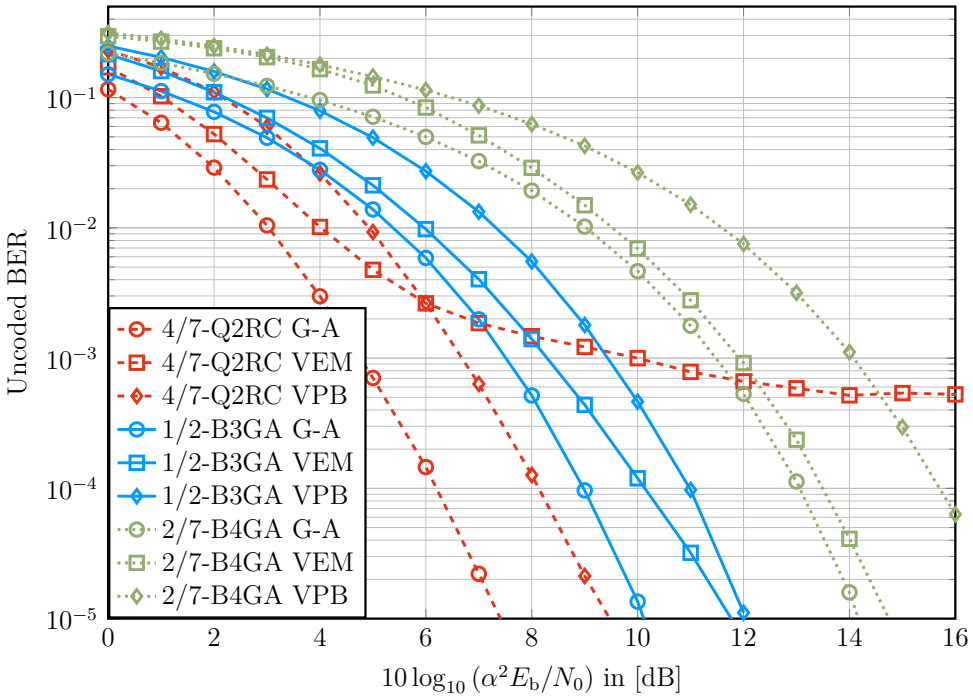


Figure 5-13: Uncoded bit error rate (BER) of the G-A ($\hat{\underline{\lambda}} = \underline{\lambda}$, $\hat{\sigma}^2_{w'} = \sigma^2_{w'}$), VEM and VPB systems over $\alpha^2 E_\mathrm{b}/N_0$ in the SF channel for the three waveforms ($N = 42$).

5.7 Error Rates

This section highlights the error rate performance of the different systems and compares them to the G-A, perfectly synchronized system (basically a transmission over the AWGN channel). The two application cases of short bursts are taken up from the introduction. First, the FH system with short hopping times which is typical for e.g. tactical networks. It is assumed, that such a system will deploy an outer channel code aggregating multiple bursts instead of encoding each burst individually due to the better error correction capabilities of longer codes. For this case, Fig. 5-13 shows the bit error rate of uncoded ($R = 1$) systems. It is noted, that without the CPM trellis termination method introduced in Section 2.5.2, the systems would suffer from a considerable error floor for this metric [49].

For the binary waveforms, the VEM performs near (gap smaller than 1 dB) the G-A system. The BER curve of the quaternary waveform suffers from an error floor at about $5 \cdot 10^{-4}$, which can be traced back to the wrong TO estimates discussed in

Section 5.5. The VPB system naturally suffers from spending $N_\mathrm{P} = 12$ pilot symbols by introducing an SNR gap of $10\log_{10}(N/(N - N_\mathrm{P})) = 1.5\,\mathrm{dB}$. The difference to the total gap to the G-A system of around 2 dB stems from inaccuracies of the channel parameter estimates. Probably for this reason, [38] proposes the use of 32 or even 64 pilot symbols. Due to the total burst length of $N = 42$ symbols, such a high amount of pilot symbols is not justifiable and hence the more balanced number of $N_\mathrm{P} = 12$ was selected.

The second application of short bursts discussed in the introduction is the use in sensor networks. Usually only little data shall be transmitted sporadically, which practically excludes the use of an outer channel code due to latency constraints and thus this application is considered to encode each burst individually. Since corrupt bursts have to be retransmitted in such a scenario and thus the BER is of less relevance, Figure 5-14 shows the coded FER for the G-A, SCVEM and VPB systems. To ensure a fair comparison in this case, each setup shall have the same spectral efficiency $\eta_\mathrm{S} = \eta_0 = 0.5\,\mathrm{bit}$, i.e. convey the same amount of data k_C in the same time and occupied bandwidth and furthermore spend the same energy on the transmission. The according coding parameters can be read from Table 5.3.

The SCVEM setup performs very close to the G-A case using the 1/2-B3GA waveform, which was already predicted analyzing the throughput in the last section, while the gap widens for 4/7-Q2RC (1 dB) and 2/7-B4GA (2 dB). In either case, the VPB receiver performs about 1.5 dB worse than the SCVEM contender, which is due to the inferior estimation capabilities of only being able to use 12 pilot symbols and the higher code rate to compensate for the pilot-induced loss of spectral efficiency. For the sake of completeness, it is mentioned, that no kind of synchronization leads to a FER of practically 100 % regardless of the SNR and waveform.

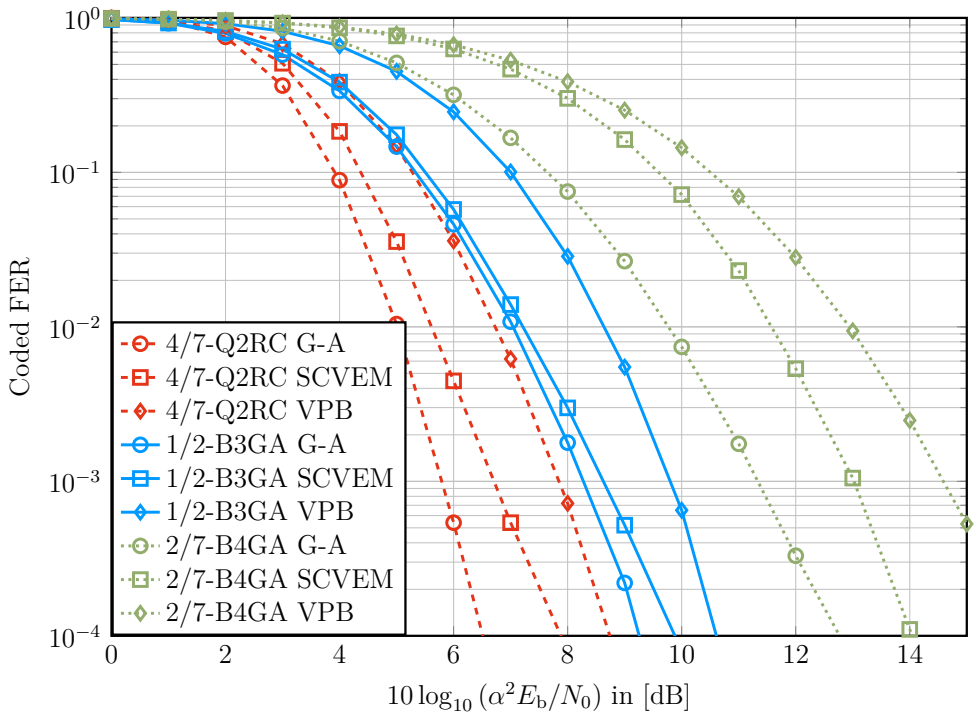


Figure 5-14: Coded FER of the G-A ($\hat{\underline{\lambda}} = \underline{\lambda}$, $\hat{\sigma}^2_{w'} = \sigma^2_{w'}$), VEM and VPB systems over $\alpha^2 E_\mathrm{b}/N_0$ in the SF channel for the three waveforms ($N = 42$). The channel code equalizes each system's SE by applying an according code rate.

5.8 Practical Assessments

5.8.1 Simulation in Fast Fading Channels

This section investigates the application of the proposed channel parameter estimator when not using the simplified, derived SF channel model from Section 2.3.2. Instead, time variant and frequency selective channels from [46] are simulated.

First, a tactical network scenario is considered in the hilly profile [46] where the maximum Doppler spread of $B_\mathrm{Dop} = 160\,\mathrm{Hz}$ is considered, i.e. a normalized maximum Doppler shift of $f_{\mathrm{Dop,\,max}} \cdot T = 0.002$. As introduced in Chapter 1, the transmission in this case is uncoded. Figure 5-15 shows the respective error rates of the unsynchronized (denoted as UNS in the plots) and VEM receivers. The SNR definition in this section does not include the scaling factor α^2 as before to reflect the fading characteristics of the channel. A receiver with G-A perfect channel state information (CSI) using it in a non-adaptive minimum mean square error (MMSE) equalizer [11]

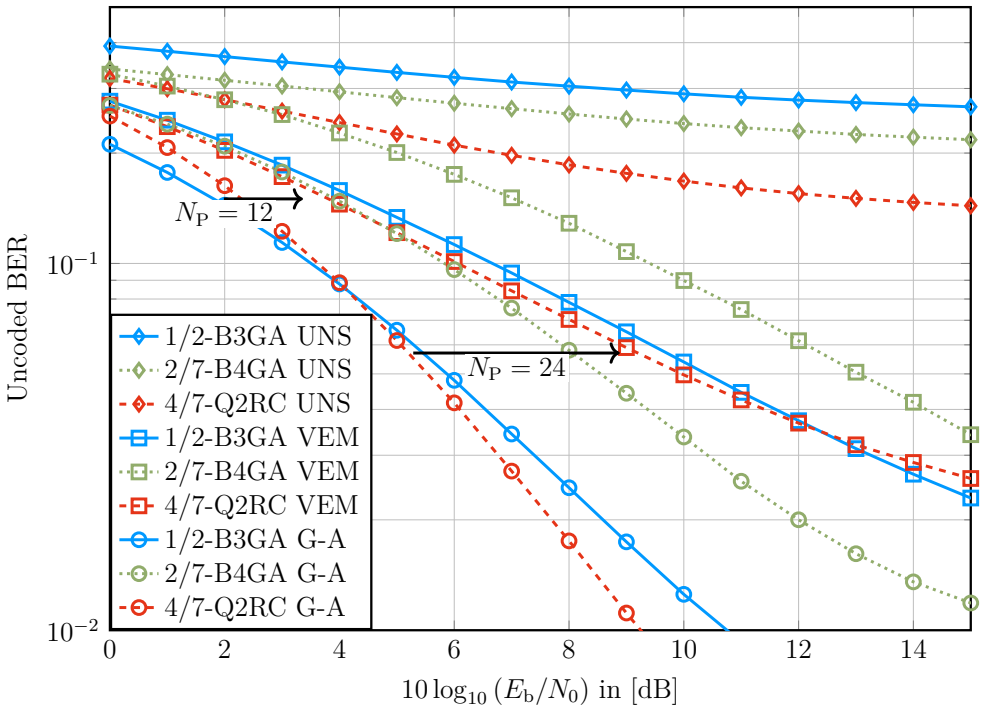


Figure 5-15: Uncoded BER in the Hilly environment [46] with $f_{\mathrm{Dop,\,max}} = 80\,\mathrm{Hz}$. The SCVEM estimator is compared to the unsynchronized and G-A case for all three waveforms and $N = 42$. The arrows imply the effect of an SNR penalty due to the deployment of pilot symbols at the two discussed target error rates.

serves as a comparison candidate. For the sake of simplicity, an adaptive approach was discarded. Though the deployed equalizer serves nonetheless as a suitable bound for the also non-adaptive VEM estimator.

Two operating points are considered in the following that both have application in tactical networks. First, the mixed excitation linear prediction (MELP) vocoder [85] is considered for an audio transmission, for which numerous extension for an increased robustness against transmission errors exist. In [86], such an extension is presented that produces an intelligible speech output even at uncoded BERs of around 15 %. All three unsynchronized systems do not (or do barely) even achieve this target error rate at high SNRs but reaching an error floor before. This fact makes clear the absolute need of some kind of synchronization or respectively, a channel equalizer. The VEM receivers show an SNR gap of around 1.7 dB to 3.0 dB at the target error rate given by [86]. The second operating point stems from an outer channel code's deployment including larger interleavers which creates diversity in the

channel. Assuming an aggregation of several bursts (or parts of, for that matter) to obtain an outer code word of length 511 binary symbols, a ($n_\mathrm{C} = 511, k_\mathrm{C} = 268$) Bose Chaudhuri Hocquenghem (BCH) code can be applied to achieve a SE comparable to the average in Table 5.3. This code thereby has a code rate $R = 0.52$ and an error correction capability of 29 symbols [87]. If the uncoded transmission does not exceed an error rate of $29/511 = 0.057$ on average, an outer error rate of zero is theoretically feasible. It is emphasized, that this is a very rough approach to obtain this target error rate, but the principal performance which the receivers have to provide in this setup should have been made plausible. As can be seen in Fig. 5-15, the VEM systems' gap increases to around 4 dB for all waveforms. At this point it is recalled, that G-A systems do not spend any pilot symbols to obtain their CSI. The use of pilots would lead to an SNR penalty of $10\log_{10}\left(\frac{N-N_\mathrm{CZS}}{N-N_\mathrm{P}-N_\mathrm{CZS}}\right)$ respecting the termination symbols with no information. For the sake of simplicity, $N_\mathrm{CZS} = 2$ symbols are considered for termination, which is the average of the three waveforms and $N_\mathrm{P} = 12$ pilot symbols are deployed as in Section 5.7. The then resulting penalty of 1.55 dB is given in Fig. 5-15 at the first target error rate as an arrow to imply the effect. The gap at the second operating point corresponds to $N_\mathrm{P} = 24$. Since perfect CSI is only obtained by $N_\mathrm{P} \to \infty$, the performance with only a finite number of pilot symbols would surely be worse than the G-A case, too. The decreased effective throughput remains ignored in this consideration.

The second application discussed in Chapter 1 is the burst communication in sensor networks. Here sporadic transmissions of little data are paused for much longer than in the tactical networks considered above. As in Section 5.7, the coded FER is used as a metric for this scenario and shown in Fig. 5-16. The environment is assumed to have a more urban character than for the tactical networks and the respective transceivers to be immobile ($f_{\mathrm{Dop,\,max}} \cdot T = 0$).

The gap between the SCVEM and the G-A receiver widens similarly as in Fig. 5-15 and an error floor is reached surely long before the transmission is arguably considered as reliable in typical communication scenarios. Though sensor networks with sparse data transmission usually do not have hard latency constraints, as they

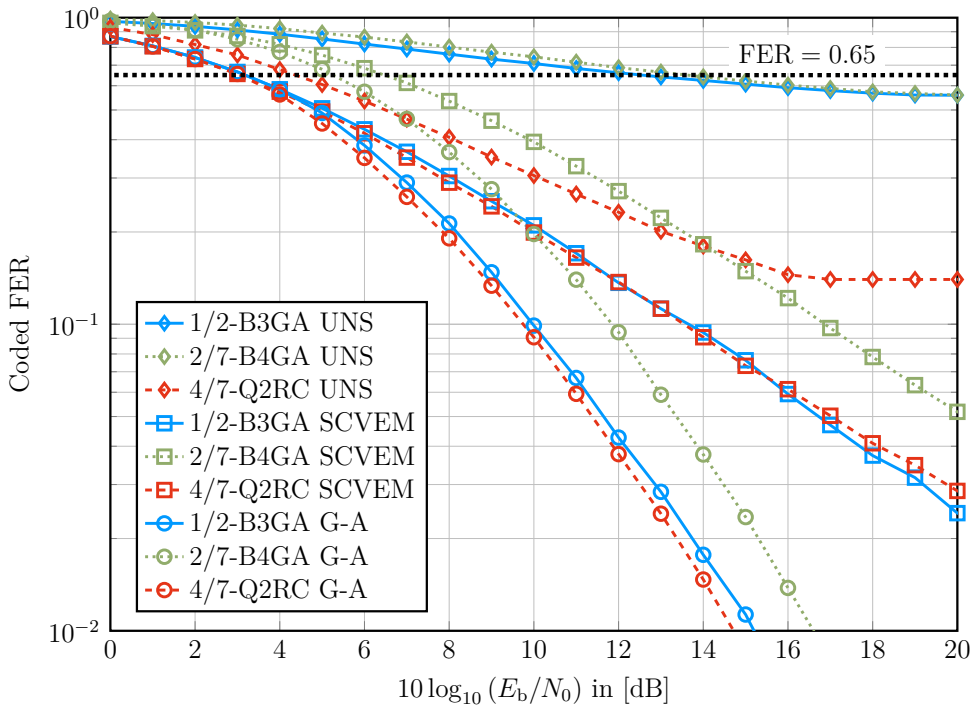


Figure 5-16: Coded FER in the Urban environment [46] with $f_{\text{Dop, max}} = 0\,\text{Hz}$. The SCVEM estimator is compared to the unsynchronized and G-A case for all three waveforms and $N = 42$. The target FER of 0.65 is taken from [88] and is optimal in terms of transmission energy in networks without latency constraints.

can retransmit a faulty frame without causing further delays in the system (given that a faulty frame can be recognized and that information can be made available to the transmitter). This fact can be used to use retransmissions to diversify the channel, as transmitting twice suffering from a FER of 0.1 is equal to one transmission suffering from a FER of 0.01 in terms of reliability. If the second case requires at least 3 dB more power than the first one, a gain is present by this approach. In [88], this topic was investigated and an optimal operating point of an FER of 0.65 was derived for sensor networks in Rayleigh fading channels. Considering this target FER in Fig. 5-16, the SCVEM performs equally well as the G-A for the 1/2-B3GA and 4/7-Q2RC waveforms and 1 dB worse for 2/7-B4GA. While for the quaternary configuration, the unsynchronized receiver can even be regarded as suited, the other two waveforms suffer from gaps up to 10 dB.

The simplifying assumptions that led to the SF channel model in Section 2.3.2 were not entirely justified as this section showed by providing error rates using the

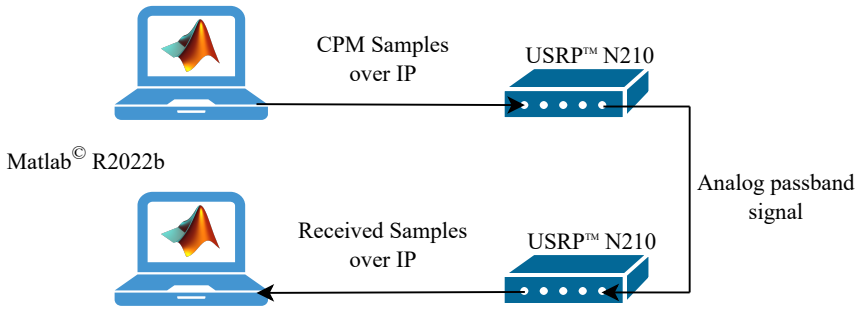


Figure 5-17: Schematic of SDR transmission. The CPM baseband samples were generated on a computer and passed over an Ethernet connection to a USRP™ N210 [47]. The analog signal was transmitted by cable to a second SDR of the same model and the baseband samples were again passed over Ethernet to a computer, where they were processed. The signal processing and hardware connection were realized in Matlab© [89] under the use of the respective hardware support package. The SDR configurations are listed in Table 5.4.

parameters discussed in Section 2.3 directly in an simulation. By providing realistic operating points in terms of target error rates, the EM based systems could be reasoned to be sufficiently well performing compared to a G-A system with (in the tactical network scenario) and even without (sensor network scenario) considering the fact, that pilot symbols are usually deployed in practice to perform channel estimation.

5.8.2 Measurements in Hardware Setup

The objective of this section is to show that no potential hardware impairments were withheld when modeling the channel in Section 2.3.2. This shall be done indirectly by comparing the simulation with a cable bound connection of two SDRs. The SDR setup is sketched in Fig. 5-17.

The baseband signal sample generation is done on a first Matlab© instance and then passed over Ethernet to the first SDR. The analog passband signal is transmitted by a cable connection to a second SDR and after downmixing and sampling, the respective samples are passed back to the same computer and processed in a second Matlab© instance. Both SDRs were connected through an Ethernet switch to the computer. The SDR properties and configurations are given in Table 5.4.

Unfortunately, the available front ends did not both offer a carrier frequency in

Table 5.4: SDR properties and configurations for the transmitter and receiver hardware.

	Transmitter	**Receiver**
Model name	USRP™ N210	USRP™ N210
Front end module	CBX	SBX/CBX
Carrier frequency range	1.18 GHz to 6.02 GHz	380 MHz to 4.42 GHz
Carrier frequency	2.45 GHz	
Baseband sampling rate	337.8 kHz	

the VHF/UHF band and thus it had to opted for a higher value. The selected carrier frequency of 2.45 GHz lies in the industrial, scientific and medical (ISM) band and the results are not expected to have limited meaning because of this choice. Due to restrictions in the SDRs' master clock rate (fixed to 100 MHz) and valid interpolation/decimation factors, the possible baseband sampling rate values are of discrete nature. The nearest possible rate was chosen to 336 kHz (cf. sampling rate in Table 5.1) by setting the interpolation/decimation factors to 296.

Figure 5-18 shows the uncoded BER of the VEM receiver in the cable connection setup compared to the computer simulated performance which serves as reference in this case. Only the waveform configuration of 1/2-B3GA is considered because of the lack of further insights in the other waveforms' cases. The power metric in this figure is chosen as the SNR with the definition

$$\mathrm{SNR} = \frac{P_\mathrm{S}}{P_\mathrm{N}} \tag{5.11}$$

$$= \frac{T_0/(NT)\sum_{k=0}^{NK-1}|s(kT_0)|^2}{T_0/(NT)\sum_{k=0}^{NK-1}|w(kT_0)|^2} \tag{5.12}$$

$$= \sum_{k=0}^{NK-1}\frac{|s(kT_0)|^2}{|w(kT_0)|^2}\,. \tag{5.13}$$

The white noise signal is added in the second Matlab© instance with the desired power. In order to evaluate the hardware effects, the bit error rate is compared to

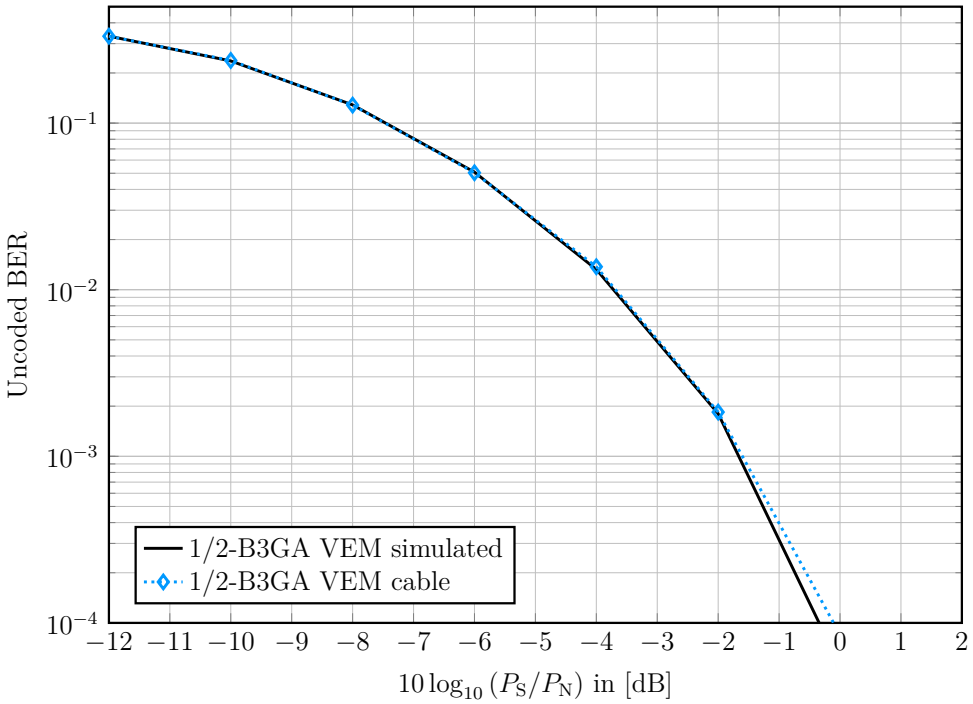


Figure 5-18: Uncoded BER in a cable connection setup with SDRs.

the simulated VEM receiver. The SNR in this case is obtained by the relation

$$\mathrm{SNR} = \frac{E_S T_0}{N_0 T} \,. \tag{5.14}$$

The achieved error rates in both cases are virtually indistinguishable until an SNR of about 2 dB, which indirectly confirms the assumptions of (quasi) constant CFO, CPO and sample clock over the course of $T_{\mathrm{Burst}} = 1\,\mathrm{ms}$. Beyond that SNR, the hardware setup shows a slight degradation compared to the simulation. This is traced back to slight IQ imbalances that were observed in the hardware.

It is noted, that the receiver processing is not done in real time. It is emphasized in the following what this section shall show and what it does not claim. The goal was to verify, that no hardware behavior significantly deteriorates the software receiver's performance and all hardware aspects were correctly assumed in Section 2.3.2. Neither potential real time processing capabilities was to be proved, nor a complete radio channel including transceiver hardware was to be assessed. Both points were beyond the scope of this thesis due to the eventually non-feasible implementation effort in a

limited time frame.

5.9 Analysis of Computational Complexity

An important aspect assessing practical relevance of an estimator is its computational complexity. The setup from Section 5.5 is considered, where both the SCVEM and VPB offer comparable estimation accuracy over a wide SNR range. This section will point out the main drivers of complexity in the SCVEM receiver (and link it to Algorithm 1) as well as compare it with its contender VPB. For the sake of simplicity, only a subset of the complete receiver that is dominant with respect to the complexity is considered and furthermore, each operation (real additions, real multiplications, exponential function etc.) is treated equally. These inaccuracies are accepted to allow for a simple and still informative comparison of synchronization complexity. The absolute number depends on several parameters which were mainly introduced in Sections 4.4.5 and 4.5.2 and are listed in Table 5.5. The SNR operating points are determined as the point where the SCVEM estimator has reached each parameter's CRVB for the respective waveform (cf. Figs. 5-8 to 5-11). In other words, this is the SNR at which the CPM detector is able to produce reliable symbol estimates. For the the basic system parameters, Table 5.1 is recalled.

Calculation of likelihoods in line 2, Algorithm 1: For the calculation of $|\mathcal{I}(\underline{\lambda})|$ likelihoods, the according path metrics of every grid point must be computed (cf. Section 4.5.2 for the necessary steps), which amounts to

$$\begin{aligned}\mathcal{C}_{\mathcal{L},\,1} =& NM^L \left[N_\mathrm{I}(\tau)N_\mathrm{I}(\nu)\cdot(8K-2) + N_\mathrm{I}(\underline{\lambda})\cdot 3 + QN_\mathrm{I}(\underline{\lambda})\cdot 2\right] \\ &+ N_\mathrm{I}(\underline{\lambda})NQM^{L-1}\cdot(2M-1)\,.\end{aligned} \tag{5.15}$$

The first three addends represent the grid points' path metric computation that is comprised by the multiplication of $r'(kT_0+nT)\cdot c^*_{\gamma_n}(kT_0-\hat{\tau})$ and the product's summation, the metrics' rotations to include every phase state and CPO combination, the multiplication with the constant quotient $2\hat{\alpha}/\hat{\sigma}^2_{w'}$ and the exponential in (4.78), respectively. The multiplication with a priori information is ignored since none is

Table 5.5: Relevant parameters for computational complexity. $N_{\mathrm{I}}(\tau)$ for 4/7-Q2RC is determined according to the decreased range explained in (5.8)

	1/2-B3GA	**2/7-B4GA**	**4/7-Q2RC**
$N_{\mathrm{I}}(\nu)$	8	19	5
$N_{\mathrm{I}}(\theta)$	6	4	2
$N_{\mathrm{I}}(\tau)$	4	6	4
$N_{\mathrm{I}}(\underline{\lambda})$	192	456	40
$10\log_{10}(\alpha^2 E_{\mathrm{S}}/N_0)$	9 dB	10 dB	8 dB
$N_{\mathrm{B}}(\nu)$	7	4	10
$N_{\mathrm{B}}(\tau)$	14	6	13
N_{ZP}	512	512	512
n_{MC}	32	32	64

available yet. The execution of BCJR's forward part, which computes the likelihood according to (2.57) is considered in the last term. The complexity $\mathcal{C}_{\mathcal{L},\,2}$ of line 18 is derived similarly.

E-Step in line 10, Algorithm 1: The computation of the posterior probabilities for multiple coarse estimates and iterations is done by the execution of the CPM detector (including the calculation of the path metrics for $\underline{\lambda}_{\mathrm{old}}$) and takes

$$\mathcal{C}_{\mathrm{E},\,1} = N_{\mathrm{I}}(\tau) N_{\mathrm{It}} N \left[Q M^{L-1} \cdot (6M-2) + M^L \left(8K - 2 + 3(Q-1) + 2Q\right)\right] \tag{5.16}$$

operations. The complexity evaluation of (4.24) is ignored, since most of the posterior probabilities are zero or very close to zero and thus its overall complexity is negligible.

E-Step in line 12: To include the channel decoder in the SC CPM setup, two executions of the CPM detector (the path metrics must only be calculated once, because the appliance of the decoder information is just an additional multiplication) and one of the decoder are necessary. Summarized, this line is responsible for

$$\begin{aligned}\mathcal{C}_{\mathrm{E},\,2} = {} & N_{\mathrm{I}}(\tau) N \left[2Q M^{L-1} \cdot (6M-2) + M^L \left(8K - 2 + 3(Q-1) + 3Q\right)\right] \\ & + N_{\mathrm{I}}(\tau) \left(40 \log_2(n_{\mathrm{MC}}) \cdot n_{\mathrm{MC}}\right)\end{aligned} \tag{5.17}$$

operations (cf. [55] for the complexity on the decoder). $n_{\mathrm{MC}} = 2^{\lceil \log_2(n_{\mathrm{C}}) \rceil}$ is the length in bits for the mother Polar code, that is shortened in the process of providing flexible code rates. It is noted, that of course the specific decoder complexity is solely dependent on the respective choice.

M-Step in line 14, Algorithm 1: In the maximization step, only the optimizations with regard to CFO and TO are considered, since the other parameters are estimated with simple closed form expressions. As suggested in Section 4.5.2, the Goertzel algorithm is used to evaluate the CAF at $N_{\mathrm{B}}(\nu)$ CFO and $N_{\mathrm{B}}(\tau)$ TO values (cf. (4.81)). The number of operations for the computation of one bin by this algorithm can be deducted from [80] and amounts to $3NK+4$ for the case at hand. The required resolution is determined by the SNR (through the respective CRVB), which is given in Table 5.5. In the CFO case, the optimization amounts to

$$\mathcal{C}_{\mathrm{M},\,1} = 6NK + N_{\mathrm{B}}(\nu) \cdot (3NK + 4) \tag{5.18}$$

operations with the first term accounting for the multiplication of $r'(kT_0)$ and $\bar{c}^*(kT_0 - \tau_{\mathrm{old}})$ (cf. (4.33)) and the second for the Goertzel executions. It is noted, that the scaling with T_0 in (4.33) does not need to be carried out (and thus $2NK$ multiplications were saved) as long as this factor is instead considered in the FF optimization in (4.38), where it is just one scalar multiplication. To compute a partially interpolated CAF, the method described in Section 4.5.2 is carried out. First, the received and the reference signal are each transformed by means of an FFT. Hereby both signals are zero padded to ensure a linear (in contrast to circular) correlation and to utilize a radix-2 FFT implementation. While [28] computes a spectrally zero padded inverse FFT of this product, this thesis again suggests the Goertzel algorithm to compute the few necessary time bins, which takes in total

$$\mathcal{C}_{\mathrm{M},\,2} = 2\,(4N_{\mathrm{ZP}} \log_2(N_{\mathrm{ZP}}) - 6N_{\mathrm{ZP}} + 8) + 6N_{\mathrm{ZP}} + N_{\mathrm{B}}(\tau) \cdot (3N_{\mathrm{ZP}} + 4) \tag{5.19}$$

operations. N_{ZP} describes the properly zero padded length in samples and is given in Table 5.5. The first addend represents the two necessary complex input FFTs with

the complexity taken from [90], the second the sample wise multiplication of the two spectral signals and the last addend the Goertzel algorithm.

The deployment of the Goertzel algorithm in the CFO estimation is (depending on the waveform) 10.1 to 25.2 times less complex than an FFT that provides the required resolution through zero padding and enables a radix-2 implementation, whereas the cost saving factor in the TO case ranges from 3.8 to 18.6. While there are enough highly efficient parallelized FFT implementations, only the computation of the bins can be parallelized, but not so the Goertzel algorithm itself. The cost savings make its consideration worth though.

VPB method: The main complexity part in the VPB estimation is comprised by two zero padded FFTs of signal products ((22) and (23) in [38]) that takes

$$\mathcal{C}_{\mathrm{VPB}} = 2 \cdot (6N_{\mathrm{P}}KK_{\mathrm{f}} + 4N_{\mathrm{P}}KK_{\mathrm{f}}\log_2(N_{\mathrm{P}}KK_{\mathrm{f}}) - 6N_{\mathrm{P}}KK_{\mathrm{f}} + 8) \tag{5.20}$$

operations. The interpolation factor was chosen as $K_{\mathrm{f}} = 4$ in contrast to [38] suggesting the factor two being enough. It was found, that at higher SNR values (considered in this thesis, though not considered in [38]) this factor was necessary for more satisfying results.

Table 5.6 lists the number of operations per frame (considering the SF model in the MSEE simulation's section) for every crucial part of the VEM estimator and sets that metric in relation to a the G-A synchronized system, i.e. a coherent CPM detector with subsequent channel decoder in an AWGN channel and the VPB estimator. Relevant parameters can be gathered from Tables 5.1 and 5.5.

It is obvious that the gross share of complexity falls in the calculation of likelihoods for the coarse estimation. Any way to shrink the parameter ranges would be helpful to reduce that part, e.g. by inserting few pilot symbols as discussed in Section 4.4.6. Choosing $N_{\mathrm{P}} = 12$, the grid cardinality is reduced to 110 and the total complexity would be 2.7 times less than in Table 5.6 for the 2/7-B4GA case. Investing in better oscillator hardware, e.g. at a quality of 0.1 ppm would be another way to reduce the number of grid points by decreasing the CFO range. Furthermore, typical trellis

Table 5.6: Listing of complexities in terms of number of operations per burst ($T_{\mathrm{Burst}} = 1\,\mathrm{ms}$). $\mathcal{C}_{\mathrm{VEM}}$ constitutes the sum of the parts above it. Several relations in the bottom in order shall give a rough intuition about the practical application.

	1/2-B3GA	**2/7-B4GA**	**4/7-Q2RC**
$\mathcal{C}_{\mathcal{L},\,1}$	$999,424$	$13,541,376$	$1,541,120$
$\mathcal{C}_{\mathrm{E},\,1}$	$161,792$	$792,576$	$542,720$
$\mathcal{C}_{\mathrm{E},\,2}$	$118,784$	$563,712$	$425,984$
$\mathcal{C}_{\mathrm{M},\,1}$	$6,940$	$4,624$	$9,256$
$\mathcal{C}_{\mathrm{M},\,2}$	$55,368$	$43,048$	$53,828$
$\mathcal{C}_{\mathcal{L},\,2}$	$73,728$	$321,024$	$217,600$
$\mathcal{C}_{\mathrm{VEM}}$	$1,416,036$	$15,266,360$	$2,790,508$
$\mathcal{C}_{\mathrm{VEM}}/\mathcal{C}_{\mathrm{VPB}}$	17.3	186	34.1
$\mathcal{C}_{\mathrm{VEM}}/\mathcal{C}_{\mathrm{G\text{-}A}}$	53.2	211	33.5
$\mathcal{C}_{\mathrm{FPGA}}/\mathcal{C}_{\mathrm{VEM}}$	5.2	0.49	2.7

reductions and efficient logarithmic implementations of the BCJR to decrease the number of operations were not considered in this work. Anyway, the coarse estimation and also the $N_{\mathrm{I}}(\tau)$ EM instances are perfectly suited to be parallelized.

Clearly the VPB complexity is magnitudes smaller than SCVEM's. Though in the quantitative comparison, it is ignored to the benefit of the VPB that a phase and timing tracking loop has to be carried out in the following detector, whereas a simple coherent detector can be deployed in the SCVEM case. The fact, that the spectral efficiency of the blind system is always higher than that of the pilot based can also be considered as a soft factor favoring the EM based approach.

The last line in Table 5.6 normalizes the potential computational power of an exemplary, bottom end field programmable array (FPGA) Xilinx Spartan-7 XC7S6 operating at 741 MHz and incorporating ten digital signal processing (DSP) units (i.e. $\mathcal{C}_{\mathrm{FPGA}} = 7.41 \cdot 10^6$ potential operations per T_{Burst}) by the number of operations required by the estimator. Whereas in the 1/2-B3GA and 4/7-Q2RC case, multiple instances of the SCVEM can potentially be processed on the same chip, this is not

possible for the 2/7-B4GA waveform without any grid reduction. It is noted though, that in case of tactical radios, more capable hardware would probably be used.

The actual computational power that needs to be spent on the parameter estimation is always dependent on the desired operating point of the system. A system that shall provide a bit error rate of $6 \cdot 10^{-2}$ (cf. Fig. 5-15) surely poses higher requirements at the estimates' quality than the discussed sensor network that is intended to work at a frame error rate of 0.65 (cf. Fig 5-16) and consequently needs to spend more computational resources on the estimation.

5.10 Key Points of the Chapter

- Multiple simulations have been carried out for this thesis, which were presented in this chapter. The results shall provide a comprehensive analysis of the channel parameter estimator described in the last chapter.

- The convergence behavior of the VEM estimator was investigated by considering the Kullback-Leibler divergence between the posterior distribution of the data symbols given the ML estimate $\hat{\underline{\lambda}}_{\mathrm{ML}}$ and the distribution after each iteration. It was shown, that a complete convergence is only achieved after dozens of iterations and the question arose whether this is necessary or less accurate results will do just as well.

- It was answered by showing the progress of each iteration in terms of the MSEE normalized by the respective CRVB. After two iterations no more considerable gains in accuracy can be expected and hence this number was adopted for Algorithm 1.

- The effect of one additional iteration under the use of the SC CPM setup was investigated by computing the mutual information between the transmitted and the received, synchronized signal and compared to the cases of not using that setup in the additional iteration and the G-A received signal. The information

value of the G-A case was reached considerably earlier by the SCVEM and thus a positive effect in terms of MSEE and error rates could be expected.

- The MEV was examined in some cases to give some insights about the effect of CPM's rotational invariance and a weak point of the pilot-based estimator VPB. Furthermore the expected error in the FF estimation of VEM and SCVEM was investigated and conclusions were drawn originating on the characteristics of the noise power and initial FF estimation as well as on the FF optimization rule in the M-Step.

- The MSEEs of the VEM, SCVEM and the VPB estimator were benchmarked against their respective theoretical limits, namely the CRVBs. The EM based methods approach the CRVB once the SNR is high enough for the (SC) CPM detector to provide reliable symbol estimates as expected for a DD technique, whereas the VPB gives appropriate estimates across the total considered SNR range. The earlier anticipated advantages of the SCVEM over the VEM setup were confirmed.

- The inherent advantage of the SCVEM over the VPB method was emphasized by displaying the effective throughputs. Although VPB performs closer to the theoretical optimum given by the respective sphere packing bounds, only at arguably irrelevant low SNRs could it provide higher throughputs than the SCVEM receiver. By not needing to apply pilot symbols, the throughput tops at a substantially higher level in that case.

- The investigation of error rates confirmed the above results of the EM based estimators being able to provide rates closer to the G-A receiver than the VPB method due to a sufficient estimation performance paired with a higher SE.

- The assumptions made on the channel characteristics made in Section 2.3.2 were evaluated by simulating error rates in an appropriate fast fading multi path environment and were found to be too optimistic. By setting realistic

performance goals, the EM based receivers were shown to provide sufficiently accurate results though.

- By setting up a cable connection between two SDRs, hardware effects were evaluated and the respective assumptions made in Section 2.3.2 were verified.
- The SCVEM computational complexity was evaluated quantitatively and compared to the VPB and G-A cases. Furthermore, a rough assessment on the real time capabilities of the SCVEM receiver was carried out with the bottom line, that though being much more complex than the pilot based approach, it can easily be executed on even very low-end hardware in many cases.

Chapter 6

Conclusion

6.1 Summary

In burst type CPM transmissions, as occurring in sensor networks or frequency hopping systems, the channel parameters need to be acquired over and over gain. Tracking of those is often not feasible in these setups due to a potentially large temporal (sensor networks) or spectral (frequency hopping networks) or both (multi user frequency hopping networks) separation of the transmission bursts. Non-data-aided, feedforward estimation methods from the literature are not suitable to their inferior performance due to the lack of a statistically reliable signal snippet, whereas phase-locked-loop solutions are by design not suited to the recurring parameter acquisition. Pilot-based estimators can work well, but share the disadvantages of a decreased spectral efficiency and furthermore are prone to intentional interference due to their potentially same and known recurring symbol sequence.

The goal of this thesis was to provide a non-pilot estimator for the above scenarios. First, potential channel conditions were investigated and a static (over the course of one burst) and flat fading channel model was motivated. In this model, the channel parameters fading factor, carrier frequency and phase offset as well as the timing offset have to estimated and corrected for that matter. The expectation maximization framework was harnessed to derive an algorithm that fulfills this task in an iterative fashion by converging to the joint ML solution. While the FF and CPO can

be estimated by closed form expressions, the CFO and TO are virtually jointly estimated through two subsequent one-dimensional grid searches, whose optimal order was deduced from the parameters' Fisher information. The typical gradient ascent algorithm's problem of converging into a local maximum of its cost function (in this case local maxima of the LLF) was countered by an initial grid covering the joint LLF, whose most promising nodes served as starting values for the EM algorithm. The grid size was analytically optimized by analyzing the scalar LLF of the respective channel parameters and determining a principal distance between maxima. Efficient implementation suggestions were given for the crucial (in terms of computational complexity) steps were given.

The proposed algorithm was analyzed in its convergence behavior by evaluating metrics such as the Kullback-Leibler divergence and the mutual information to find a suitable EM iteration count to terminate the algorithm. The estimator performance was compared in terms of the MSEE to the theoretical optimum, namely the CRVB. The generally infeasible derivation of this true (i.e. as tight as possible) bound was done by determining the easier to calculate MCRVB and subsequently showing for all four parameters, that these two coincide. It was thus shown, that the proposed estimator SCVEM reaches this bound for every case and is thus an efficient estimator. Examining the metrics effective throughput, bit and frame error rates, it was shown, that the SCVEM performs better than a state of the art pilot based CPM burst estimator across all considered scenarios. This advantage can be attributed to the higher spectral efficiency of the SCVEM systems paired with the appropriate estimation performance. The main drawback of the proposed method is its computational complexity which was analyzed quantitatively and it could be reasoned, that a real time application of such a system is easily feasible with in most cases no limitations whatsoever.

External requirements, such as shorter and flexible hop times [1] are easily implemented with the proposed estimation algorithm and would arguably pose further problems to pilot based approaches. Also time variant channels can be incorporated in the VEM framework by splitting up bursts as far as necessary to again rightfully as-

sume constant nature of the parameters over the course of one subburst. In the latter case, subsequent subburst channel conditions do not need to be assumed independent to each other.

6.2 Possible Future Work

- The proposed algorithm can be extended to also include multi-h CPM configurations, which promise better performance through higher Euclidean symbol distances [91]. A typical point of criticism is the difficult CPO synchronization, for which the proposed SCVEM seems to a promising solution to.

- As was indicated in simulations, the proposed method is competitive in even fast fading and frequency-selective environments, although it was not designed for such scenarios. Further performance improvements can be expected, if multi-tap channels are considered in the algorithm design process. Current research [92] adopts the EM framework to tackle the problem with promising results, though only for less restrictive conditions than considered in simulations mentioned above (time invariant channel, dominant line of sight (LOS) component, only GMSK modulation, longer bursts). For similar, less restrictive conditions, solutions were already proposed, e.g. in [93–99]. However, for the use cases introduced in Chapter 1, the literature leaves a gap for a blind channel estimation method and the EM algorithm seems like a promising approach.

- A substantial problem of the proposed method (and for that matter also for [92]) is the computational burden in the receiver despite the implementation suggestions. Standard approaches to reduce the detector complexity [100, 101] are promising, but their impact on the estimation performance needs to be investigated. Furthermore, machine learning techniques such as the genetic algorithm [102] can be imagined to ease up the evaluation of the complete grid of initialization points.

This Page Intentionally Left Blank

Appendix A

Calculation from Chapter 2

A.1 Derivation of Noise Estimation

The presented noise estimation (cf. (2.45)) in a constant amplitude waveform relies on a flat fading channel and the received signal amplitude's variation is utilized for which additive noise must be the sole cause. Letting a and b represent the real and imaginary part of the received signal $r' = a + jb$, the variance of the squared absolute value writes as

$$\begin{aligned} \mathrm{Var}[|a+jb|^2] &= \mathrm{Var}[a^2+b^2] \\ &= \mathrm{Var}[a^2] + \mathrm{Var}[b^2] + 2\mathrm{Cov}[ab] \\ &= \mathrm{E}[a]^2\,\mathrm{Var}[b] + \mathrm{E}[b]^2\,\mathrm{Var}[a] + \mathrm{Var}[a]\,\mathrm{Var}[b] \\ &= \mathrm{Var}[a]\,\mathrm{Var}[b] = \frac{\sigma_{w'}^4}{4}\,. \end{aligned} \tag{A.1}$$

General variance properties have been used and $\mathrm{Var}\,[a] = \mathrm{Var}\,[b] = \sigma_{w'}^2/2$, $\mathrm{E}[a] = \mathrm{E}[b] = 0$ and $\mathrm{Cov}[ab] = 0$ are given by (2.41)-(2.44). By eventually doubling the end result's square root, the variance of w' is given as in (2.45).

A.2 Transformation of Complex Difference's Absolute Value

Two complex number $z_1 = a + jb$ and $z_2 = c + jd$ are given.

$$\begin{aligned} |z_1 - z_2|^2 &= |a + jb - (c + jd)|^2 \\ &= [(a - c) + j(b - d)] \cdot [(a - c) - j(b - d)] \\ &= a^2 - 2ac + c^2 + b^2 - 2bd + d^2 \\ &= |a + jb|^2 + |c + jd|^2 - 2\,\mathrm{Re}\left\{(a + jb)(c + jd)^*\right\} \ . \end{aligned} \tag{A.2}$$

Appendix B

Calculations from Chapter 3

B.1 Derivation of MFIM for CPM Bursts

The MFIM is calculated based on (3.15), which describes the tilted and sampled received signal. To facilitate the calculations, they are carried our in the continuous time domain, as this is equivalent according to [17, 2.3]. Equation (3.14) is recalled and adjusted to hold the continuous received signal (cf. (1.5))

$$[\underset{\sim}{I}_M(\underline{\lambda})]_{k,l} = \frac{2}{N_0} \int_{\tau}^{NT+\tau} \mathrm{E}_{\underline{a}} \left[\mathrm{Re} \left\{ \frac{\partial r(t, \underline{a}, \underline{\lambda})}{\partial \lambda_k} \cdot \frac{\partial r^*(t, \underline{a}, \underline{\lambda})}{\partial \lambda_l} \right\} \right] \mathrm{d}t \ . \tag{B.1}$$

Because of the MFIM's symmetry ($[\underset{\sim}{I}_M(\underline{\lambda})]_{k,l} = [\underset{\sim}{I}_M(\underline{\lambda})]_{l,k}$), not all entries have be calculated individually. As a first step, the four partial derivatives with regard to channel parameters are calculated for both the received signal and its complex con-

jugate with the time reference set to $t_0 = NT/2$

$$\frac{\partial r(t,\underline{a},\underline{\lambda})}{\partial \alpha} = \frac{1}{\alpha} \cdot r(t,\underline{a},\underline{\lambda}) \tag{B.2}$$

$$\frac{\partial r^*(t,\underline{a},\underline{\lambda})}{\partial \alpha} = \frac{1}{\alpha} \cdot r^*(t,\underline{a},\underline{\lambda}) \tag{B.3}$$

$$\frac{\partial r(t,\underline{a},\underline{\lambda})}{\partial \nu} = j2\pi(t - NT/2) \cdot r(t,\underline{a},\underline{\lambda}) \tag{B.4}$$

$$\frac{\partial r^*(t,\underline{a},\underline{\lambda})}{\partial \nu} = -j2\pi(t - NT/2) \cdot r^*(t,\underline{a},\underline{\lambda}) \tag{B.5}$$

$$\frac{\partial r(t,\underline{a},\underline{\lambda})}{\partial \theta} = j \cdot r(t,\underline{a},\underline{\lambda}) \tag{B.6}$$

$$\frac{\partial r^*(t,\underline{a},\underline{\lambda})}{\partial \theta} = -j \cdot r^*(t,\underline{a},\underline{\lambda}) \tag{B.7}$$

$$\frac{\partial r(t,\underline{a},\underline{\lambda})}{\partial \tau} = \left[-j\left(2\pi h \sum_{n=-(L-1)}^{N-1} a_n g(t-\tau-nT) - \beta\pi h \frac{M-1}{T}\right)\right] \cdot r(t,\underline{a},\underline{\lambda}) \tag{B.8}$$

$$\frac{\partial r^*(t,\underline{a},\underline{\lambda})}{\partial \tau} = \left[+j\left(2\pi h \sum_{n=-(L-1)}^{N-1} a_n g(t-\tau-nT) - \beta\pi h \frac{M-1}{T}\right)\right] \cdot r^*(t,\underline{a},\underline{\lambda})\ . \tag{B.9}$$

The frequency pulse ($g(t) = \frac{\mathrm{d}p(t)}{\mathrm{d}t}$ is the phase pulse derivative. By using $r(t,\underline{a},\underline{\lambda}) \cdot r^*(t,\underline{a},\underline{\lambda}) = \alpha^2 \frac{E_\mathrm{S}}{T}$ and considering the MFIMs symmetry, the expectations in (B.1) of all relevant and non-zero entries are calculated. The short notation $\mathrm{E}^{\lambda_k,\lambda_l} = \mathrm{E}_{\underline{a}}\left[\mathrm{Re}\left\{\frac{\partial r(t,\underline{a},\underline{\lambda})}{\partial \lambda_k} \cdot \frac{\partial r^*(t,\underline{a},\underline{\lambda})}{\partial \lambda_l}\right\}\right]$ is used

$$\mathrm{E}^{\alpha,\alpha} = \frac{E_\mathrm{S}}{T} \tag{B.10}$$

$$\mathrm{E}^{\nu,\nu} = 4\pi^2(t - NT/2)^2 \cdot \alpha^2 \frac{E_\mathrm{S}}{T} \tag{B.11}$$

$$\mathrm{E}^{\nu,\theta} = 2\pi(t - NT/2) \cdot \alpha^2 \frac{E_\mathrm{S}}{T} \tag{B.12}$$

$$\mathrm{E}^{\nu,\tau} = -2\beta\pi^2 h \frac{M-1}{T}(t - NT/2) \cdot \alpha^2 \frac{E_\mathrm{S}}{T} \tag{B.13}$$

$$\mathrm{E}^{\theta,\theta} = \alpha^2 \frac{E_\mathrm{S}}{T} \tag{B.14}$$

$$\mathrm{E}^{\theta,\tau} = -\beta\pi h \frac{M-1}{T} \cdot \alpha^2 \frac{E_\mathrm{S}}{T} \tag{B.15}$$

$$\mathrm{E}^{\tau,\tau} = \mathrm{E}\left[\left(2\pi h \sum_{n=-(L-1)}^{N-1} a_n g(t-\tau-nT) + \beta\pi h \frac{M-1}{T}\right)^2 \cdot \alpha^2 \frac{E_\mathrm{S}}{T}\right] \tag{B.16}$$

In (B.13) and (B.15), the property $\mathrm{E}[a_n] = 0$ was used. The final step is to integrate and scale the expectations

$$\left[\underset{\sim}{I}_M(\underline{\lambda})\right]_{1,1} = \frac{2}{N_0}\int_{\tau}^{NT+\tau} \mathrm{E}^{\alpha,\alpha}\,\mathrm{d}t = 2N \cdot \frac{E_\mathrm{S}}{N_0} \tag{B.17}$$

$$\left[\underset{\sim}{I}_M(\underline{\lambda})\right]_{2,2} = \frac{2}{N_0}\int_{\tau}^{NT+\tau} \mathrm{E}^{\nu,\nu}\,\mathrm{d}t = \frac{8}{3}\pi^2 N \cdot \left(\frac{1}{4}N^2T^2 + 3\tau^2\right) \cdot \frac{\alpha^2 E_\mathrm{S}}{N_0} \tag{B.18}$$

$$\left[\underset{\sim}{I}_M(\underline{\lambda})\right]_{2,3} = \frac{2}{N_0}\int_{\tau}^{NT+\tau} \mathrm{E}^{\nu,\theta}\,\mathrm{d}t = 4\pi N\tau \cdot \frac{\alpha^2 E_\mathrm{S}}{N_0} \tag{B.19}$$

$$\left[\underset{\sim}{I}_M(\underline{\lambda})\right]_{2,4} = \frac{2}{N_0}\int_{\tau}^{NT+\tau} \mathrm{E}^{\nu,\tau}\,\mathrm{d}t = -4\beta\pi^2 hN\frac{M-1}{T}\tau \cdot \frac{\alpha^2 E_\mathrm{S}}{N_0} \tag{B.20}$$

$$\left[\underset{\sim}{I}_M(\underline{\lambda})\right]_{3,3} = \frac{2}{N_0}\int_{\tau}^{NT+\tau} \mathrm{E}^{\theta,\theta}\,\mathrm{d}t = 2N \cdot \frac{\alpha^2 E_\mathrm{S}}{N_0} \tag{B.21}$$

$$\left[\underset{\sim}{I}_M(\underline{\lambda})\right]_{3,4} = \frac{2}{N_0}\int_{\tau}^{NT+\tau} \mathrm{E}^{\theta,\tau}\,\mathrm{d}t = -2\beta\pi hN\frac{M-1}{T} \cdot \frac{\alpha^2 E_\mathrm{S}}{N_0}\ , \tag{B.22}$$

whereas the element $[\underset{\sim}{I}_M(\underline{\lambda})]_{4,4}$ needs to be given special attention

$$\begin{aligned}\frac{2}{N_0}\int_{\tau}^{NT+\tau} \mathrm{E}^{\tau,\tau}\,\mathrm{d}t = \frac{2}{N_0}\int_{\tau}^{NT+\tau} &\mathrm{E}\left[(2\pi h\sum_{n=-(L-1)}^{N-1} a_n g(t-\tau-nT))^2\right] \\ &- 2\mathrm{E}\left[2\pi h\sum_{n=-(L-1)}^{N-1} a_n g(t-\tau-nT)\right] \\ &+ \left(\beta\pi h\frac{M-1}{T}\right)^2 \mathrm{d}t \cdot \frac{\alpha^2 E_\mathrm{S}}{T}\ .\end{aligned} \tag{B.23}$$

The expectations $\mathrm{E}[a_k] = 0$ and $\mathrm{E}[a_k a_l] = (M^2-1)/3$ if $k = l$ and zero otherwise are used in the next step as well as $G_2(f) = \int_{-\infty}^{+\infty} g^2(t) \cdot e^{-j2\pi ft}\,\mathrm{d}t$. Furthermore the sum is extended to infinity without affecting the result because of the integral's temporal

limits.

$$\begin{aligned}\frac{2}{N_0}\int_{\tau}^{NT+\tau} \mathrm{E}^{\tau,\tau}\,\mathrm{d}t &= \frac{2}{N_0}\int_{\tau}^{NT+\tau} 4\pi^2h^2\frac{M^2-1}{3}\sum_n g^2(t-\tau-nT) \\ &\qquad + \left(\beta\pi h\frac{M-1}{T}\right)^2 \mathrm{d}t\cdot\frac{\alpha^2E_\mathrm{S}}{T} \qquad \text{(B.24)}\\ &= \frac{2}{N_0}\int_{\tau}^{NT+\tau} 4\pi^2h^2\frac{M^2-1}{3T}\sum_n G_2\left(\frac{n}{T}\right)\cdot e^{j2\pi n(t-\tau)/T} \\ &\qquad + \left(\beta\pi h\frac{M-1}{T}\right)^2 \mathrm{d}t\cdot\frac{\alpha^2E_\mathrm{S}}{T} \qquad \text{(B.25)}\\ &= \frac{2}{N_0}\left(4\pi^2h^2\frac{M^2-1}{3T}\sum_n G_2\left(\frac{n}{T}\right)\cdot\left[\frac{T}{j2\pi n}e^{j2\pi nN}-1\right]\right. \\ &\qquad \left. +NT\left(\beta\pi h\frac{M-1}{T}\right)^2\right)\cdot\frac{\alpha^2E_\mathrm{S}}{T}\,. \qquad \text{(B.26)}\end{aligned}$$

Having in mind, that n is an integer, the term in square brackets is zero for $n \neq 0$ and (by applying the rule of L'Hospital) NT for $n = 0$. After some modifications the MFIM element writes as

$$[\underline{I}_M(\underline{\lambda})]_{4,4} = \frac{2}{N_0}\int_{\tau}^{NT+\tau} \mathrm{E}^{\tau,\tau}\,\mathrm{d}t = 2\alpha^2\frac{E_\mathrm{S}}{N_0}N\pi^2h^2\left(4G_2(0)T\frac{M^2-1}{3T^2}+\beta\frac{(M-1)^2}{T^2}\right) \qquad \text{(B.27)}$$

B.2 Ratio of PEN and MCRB for the CPO

The TO is uniformly distributed in the interval $[-0.5T; 0.5T]$ and thus its variance is $\mathrm{E}[\tau^2] = T^2/12$. The ratio of the averaged TO penalty and scalar bound is then given as

$$\begin{aligned}\frac{\mathrm{PEN}(\theta)}{\mathrm{MCRB}(\theta)} &= \left(\frac{6\,\mathrm{E}[\tau^2]}{N^3T^2}\cdot\frac{1}{\alpha^2E_\mathrm{S}N_0}\right)\Big/\left(\frac{1}{2N}\cdot\frac{1}{\alpha^2E_\mathrm{S}N_0}\right) \\ &= \frac{1}{N^2} \qquad \text{(B.28)}\end{aligned}$$

B.3 Derivation of the true FF CRB

Inserting (3.2) into (3.6) and taking the definition of (3.10) gives the scalar FF bound

$$\mathrm{CRB}(\alpha) = \left(-\mathrm{E}_{\underline{\boldsymbol{a}},\underline{\boldsymbol{r}}} \left[\frac{\partial^2 \log\left(\sum_{\underline{a}} \mathrm{p}_{\underline{\boldsymbol{r}}|\underline{\gamma}}\left[\underline{r}|\alpha,\underline{\gamma}\right] \cdot \mathrm{P}_{\underline{\gamma}}\left[\underline{\gamma}\right] \right)}{\partial^2 \alpha} \right] \right) . \tag{B.29}$$

By assuming a known sequence $\underline{a}$ at the receiver, the bound simplifies to [58, 4.2.3]

$$\mathrm{CRB}(\alpha|\underline{a}) = \frac{N_0}{2} \cdot \left[\int_{\tau}^{NT+\tau} \left| \frac{\partial r(t,\underline{a},\alpha)}{\partial \alpha} \right|^2 \mathrm{d}t \right]^{-1} . \tag{B.30}$$

Inserting (3.15) for $r(t,\underline{a},\alpha)$ gives

$$\begin{aligned} \mathrm{CRB}(\alpha|\underline{a}) &= \frac{N_0}{2} \cdot \left[\int_{\tau}^{NT+\tau} \left| \frac{r(t,\underline{a},\alpha)}{\alpha} \right|^2 \mathrm{d}t \right]^{-1} \\ &= \frac{N_0}{2} \cdot \frac{1}{NTE_{\mathrm{S}}/T} \\ &= \frac{1}{2N} \cdot \frac{1}{E_{\mathrm{S}}/N_0} . \end{aligned} \tag{B.31}$$

This Page Intentionally Left Blank

Appendix C

Calculations from Chapter 4

C.1 Verification of Kullback-Leibler Divergence

The definitions of the LLF and $\mathcal{Q}$ function from (3.1) and (4.4) are inserted into (4.11) and by using logarithm rules, Bayes' law and respecting the independence of the LLFs from the pseudo symbol sequence, the equality to (4.10) is shown.

$$\left[\log \mathcal{L}(\hat{\underline{\lambda}}) - \log \mathcal{L}(\hat{\underline{\lambda}}_{\text{old}})\right] - \left[\mathcal{Q}\left(\hat{\underline{\lambda}}, \underline{\lambda}_{\text{old}}\right) - \mathcal{Q}\left(\underline{\lambda}_{\text{old}}, \underline{\lambda}_{\text{old}}\right)\right] \tag{C.1}$$

$$\begin{aligned}
&= \log\left(\mathrm{p}_{\boldsymbol{r}'}\left[r'(kT_0)|\hat{\underline{\lambda}}\right]\right) - \log\left(\mathrm{p}_{\boldsymbol{r}'}\left[r'(kT_0)|\underline{\lambda}_{\text{old}}\right]\right) \\
&\quad - \Bigg[\sum_{\underline{\gamma}} \mathrm{P}_{\underline{\boldsymbol{\gamma}}|\boldsymbol{r}'}\left[\underline{\gamma}|r'(kT_0), \underline{\lambda}_{\text{old}}\right] \cdot \log\left(\mathrm{p}_{\underline{\boldsymbol{\gamma}},\boldsymbol{r}'}\left[\underline{\gamma}, r'(kT_0)|\hat{\underline{\lambda}}\right]\right) \\
&\qquad - \sum_{\underline{\gamma}} \mathrm{P}_{\underline{\boldsymbol{\gamma}}|\boldsymbol{r}'}\left[\underline{\gamma}|r'(kT_0), \underline{\lambda}_{\text{old}}\right] \cdot \log\left(\mathrm{p}_{\underline{\boldsymbol{\gamma}},\boldsymbol{r}'}\left[\underline{\gamma}, r'(kT_0)|\underline{\lambda}_{\text{old}}\right]\right)\Bigg]
\end{aligned} \tag{C.2}$$

$$\begin{aligned}
&= \log\left(\frac{\mathrm{p}_{\boldsymbol{r}'}\left[r'(kT_0)|\hat{\underline{\lambda}}\right]}{\mathrm{p}_{\boldsymbol{r}'}\left[r'(kT_0)|\underline{\lambda}_{\text{old}}\right]}\right) \\
&\quad - \Bigg(\sum_{\underline{\gamma}} \mathrm{P}_{\underline{\boldsymbol{\gamma}}|\boldsymbol{r}'}\left[\underline{\gamma}|r'(kT_0), \underline{\lambda}_{\text{old}}\right] \times \\
&\qquad \left[\log\left(\frac{\mathrm{P}_{\underline{\boldsymbol{\gamma}}|\boldsymbol{r}'}\left[\underline{\gamma}|r'(kT_0), \hat{\underline{\lambda}}\right]}{\mathrm{P}_{\underline{\boldsymbol{\gamma}}|\boldsymbol{r}'}\left[\underline{\gamma}|r'(kT_0), \underline{\lambda}_{\text{old}}\right]}\right) + \log\left(\frac{\mathrm{p}_{\boldsymbol{r}'}\left[r'(kT_0|\hat{\underline{\lambda}})\right]}{\mathrm{p}_{\boldsymbol{r}'}\left[r'(kT_0|\underline{\lambda}_{\text{old}})\right]}\right)\right]\Bigg)
\end{aligned} \tag{C.3}$$

$$= \sum_{\underline{\gamma}} \mathrm{P}_{\underline{\boldsymbol{\gamma}}|\boldsymbol{r}'}\left[\underline{\gamma}|r'(kT_0), \underline{\lambda}_{\text{old}}\right] \cdot \log\left(\frac{\mathrm{P}_{\underline{\boldsymbol{\gamma}}|\boldsymbol{r}'}\left[\underline{\gamma}|r'(kT_0), \underline{\lambda}_{\text{old}}\right]}{\mathrm{P}_{\underline{\boldsymbol{\gamma}}|\boldsymbol{r}'}\left[\underline{\gamma}|r'(kT_0), \hat{\underline{\lambda}}\right]}\right) \tag{C.4}$$

The iteration index (i) was ignored above.

C.2 Verification of Entropy

Starting from the claim made in (4.12) and ignoring the iteration specifier

$$\mathcal{L}\left(\underline{\lambda}_{\text{old}}\right) - \mathcal{Q}(\underline{\lambda}_{\text{old}}, \underline{\lambda}_{\text{old}}) \tag{C.5}$$

$$= \log\left(\mathrm{p}_{\boldsymbol{r}'}\left[r'(kT_0)|\underline{\lambda}_{\text{old}}\right]\right) - \sum_{\underline{\gamma}} \mathrm{P}_{\underline{\boldsymbol{\gamma}}|\boldsymbol{r}'}\left[\underline{\gamma}|r'(kT_0), \underline{\lambda}_{\text{old}}\right] \cdot \log\left(\mathrm{p}_{\underline{\boldsymbol{\gamma}},\boldsymbol{r}'}\left[\underline{\gamma}, r'(kT_0)|\underline{\lambda}_{\text{old}}\right]\right) \tag{C.6}$$

$$= -\sum_{\underline{\gamma}} \mathrm{P}_{\underline{\boldsymbol{\gamma}}|\boldsymbol{r}'}\left[\underline{\gamma}|r'(kT_0), \underline{\lambda}_{\text{old}}\right] \cdot \log\left(\frac{\mathrm{p}_{\underline{\boldsymbol{\gamma}},\boldsymbol{r}'}\left[\underline{\gamma}, r'(kT_0)|\underline{\lambda}_{\text{old}}\right]}{\mathrm{p}_{\boldsymbol{r}'}\left[r'(kT_0)|\underline{\lambda}_{\text{old}}\right]}\right) \tag{C.7}$$

$$= -\sum_{\underline{\gamma}} \mathrm{P}_{\underline{\boldsymbol{\gamma}}|\boldsymbol{r}'}\left[\underline{\gamma}|r'(kT_0), \underline{\lambda}_{\text{old}}\right] \cdot \log\left(\mathrm{P}_{\underline{\boldsymbol{\gamma}}|\boldsymbol{r}'}\left[\underline{\gamma}|r'(kT_0), \underline{\lambda}_{\text{old}}\right]\right) \tag{C.8}$$

$$= H\left(\mathrm{P}_{\underline{\lambda}_{\text{old}}}\right) \tag{C.9}$$

one arrives at the entropy of the CPM symbols' posterior distribution after a few steps.

C.3 Relation of CFO and TO's MFIs

According to [60] the modified Fisher information (MFI) in this case is computed as follows

$$\left[\underset{\sim}{I}_{\mathrm{M}}(\underline{\lambda})\right]_{ij} = \mathrm{E}_{\underline{a}}\left[\frac{T_{\mathrm{S}}}{\alpha^2 N_0} \operatorname{Re}\left\{\frac{\partial r^*(t, \underline{a}, \underline{\lambda})}{\partial \underline{\lambda}_i} \cdot \frac{\partial r(t, \underline{a}, \underline{\lambda})}{\partial \underline{\lambda}_j}\right\}\right] . \tag{C.10}$$

To obtain the MFI of the CFO and TO, the partial derivatives are chosen to $\underline{\lambda}_i = \underline{\lambda}_j = \nu$ and $\underline{\lambda}_i = \underline{\lambda}_j = \tau$, respectively. Eventually the normalized MFIs are computed to

$$T^{-2}\mathrm{MFI}(\nu) = \frac{2}{3}\pi^2 N \frac{E_{\mathrm{S}}}{N_0} \cdot \left(N^2 + 1\right) , \tag{C.11}$$

$$T^{2}\mathrm{MFI}(\tau) = \frac{2}{3}\pi^2 N \frac{E_{\mathrm{S}}}{N_0} \cdot \left(4h^2 G_2(0) T\right)\left(M^2 - 1\right) \tag{C.12}$$

and by inserting exemplary parameters of the 1/3-B2REC waveform with $G_2(0)T = 0.125$, (4.31) becomes obvious.

This Page Intentionally Left Blank

Appendix D

Calculations from Chapter 5

D.1 TO Restrictions in the VPB Estimator

The formula of calculating the timing offset on a symbol basis in [38] is given by

$$\hat{\varepsilon} = \frac{\arg\left\{\lambda_1(\hat{\nu})\lambda_2^*(\hat{\nu})\right\}}{2(M-1)\pi h} \tag{D.1}$$

with $\varepsilon = \tau/T$ relating to this work's TO definition as stated. Naturally the angle of a complex number is unambiguous only as long as it is in the range $[-\pi, +\pi)$. This gives a maximum unambiguous estimation range of

$$\hat{\varepsilon} \in \left[-\frac{1}{2(M-1)h}, +\frac{1}{2(M-1)h}\right) , \tag{D.2}$$

which is stated in (5.8).

List of Acronyms

ACRB asymptotic Cramér-Rao bound.

AWGN additive white Gaussian noise.

BCH Bose Chaudhuri Hocquenghem.

BCJR Bahl-Cock-Jelinek-Raviv.

BER bit error rate.

CAF cross ambiguity function.

CCF cross correlation function.

CFO carrier frequency offset.

CPM continuous phase modulation.

CPO carrier phase offset.

CPV correlative phase value.

CRB Cramér-Rao bound.

CRVB Cramér-Rao vector bound.

CSI channel state information.

CZS correlative zero state.

DA data aided.

DD decision directed.

DFT discrete Fourier transform.

DSP digital signal processing.

ECM expectation conditional maximization.

EE estimation error.

EM expectation maximization.

EPM electronic protection measure.

FER frame error rate.

FF fading factor.

FFT fast Fourier transform.

FH frequency hopping.

FI Fisher information.

FIM Fisher information matrix.

FPGA field programmable array.

G-A Genie-aided.

GA Gaussian.

GMSK Gaussian minimum shift keying.

HMM hidden Markov model.

i.i.d. independent and identically distributed.

IDFT inverse discrete Fourier transform.

ISI intersymbol interference.

ISM industrial, scientific and medical.

LL log likelihood.

LLF log likelihood function.

LLR log likelihood ratios.

LOS line of sight.

MAP maximum a posteriori.

MCRB modified Cramér-Rao bound.

MCRVB modified Cramér-Rao vector bound.

MELP mixed excitation linear prediction.

MEV mean estimation value.

MFI modified Fisher information.

MFIM modified Fisher information matrix.

ML maximum likelihood.

MMSE minimum mean square error.

MSEE mean square estimation error.

MSK minimum shift keying.

NDA non-data aided.

PAM pulse amplitude modulation.

PDF probability density function.

PLL phase locked loop.

PMD principal minimal distance.

PMF probability mass function.

PSD power spectral density.

RC raised cosine.

REC rectangular.

SC serially concatenated.

SCAN successive cancellation.

SCVEM serially concatenated vector expectation maximization.

SDR software defined radio.

SE spectral efficiency.

SF static (frequency) flat.

SISO soft-input-soft-output.

SNR signal-to-noise power ratio.

TDMA time division multiple access.

TO timing offset.

UHF ultra high frequency.

VEM vector expectation maxmization.

VHF very high frequency.

VPB vector pilot based.

List of Symbols

This section lists the symbols used in this thesis. They are sorted alphabetically with all Greek letter symbols in the back of the list.

$\underline{\tilde{a}}$, $\underline{a}$,$\underline{a}'$	Vector of non-burst, burst and tilted burst data symbols in a CPM burst, respectively.
$\mathcal{A}$, $\mathcal{A}'$	Regular (bipolar) and tilted (unipolar) symbol alphabet, respectively.
A_n, A'_n	n'th correlative state in the regular and tilted CPM trellis, respectively.
B_0	One-sided receive filter bandwidth.
$B_{99\,\%}$	One-sided bandwidth, that contains 99 % of the (one-sided) signal energy.
B_{coh}	Coherence bandwidth of the channel.
B_{Dop}	Doppler spread of the channel.
B_{GA}	Bandwidth of Gaussian (GA) pulse.
$\mathcal{B}(\nu)$, $\mathcal{B}(\tau)$	Sets of desired interpolation points in frequency and timing estimation.
c_0	Speed of light ($299\,792\,458\,\mathrm{m\,s^{-1}}$).
$c_{\gamma_n}(t)$, $c_{\gamma_n}(kT_0)$	Signal corresponding to branch γ_n in the n'th trellis stage and its sampled version.

$c_{\gamma_n}(kT_0, \hat{\underline{\lambda}})$	Sampled, scaled, rotated and delayed signal corresponding to branch γ_n.
$\mathcal{C}$	Number of arithmetic operations.
$d(\nu)$, $d(\theta)$, $d(\tau)$	Principal minimal distances (PMDs) of frequency, phase and timing in the respective log likelihood functions (LLFs).
$D_{99\,\%}$	Delay spread of the channel in 99 % of the measurement time.
E_S, E_b	Energy of one CPM symbol and of one binary information symbol.
$E_{r'}$, $E_{\bar{c}}$	Energy of the received (and tilted) signal and of the expected reference signal.
$\mathcal{E}_\mathrm{Frame}$	Probability of the event of a wrongly decoded frame.
f	Frequency variable.
f_c	Carrier frequency.
$f_{D,\mathrm{max}}$	Maximum Doppler shift.
$f_i(t)$	Instantaneous frequency.
f_m	Largest, mutual instantaneous frequency.
$f_n^F(\underline{s})$, $f_n^B(\dot{\underline{s}})$	Forward and backward state occupations used in the BCJR algorithm.
$f_n^M(\underline{s}, \dot{\underline{s}})$	Path metric used in the BCJR algorithm.
$f_{\mathrm{Osc,\ max}}$	Maximum oscillator offset.
$g(t)$	Frequency pulse function.
$G_2(0)$	Frequency pulse energy.
h	Modulation index.
$H_\mathrm{R}(f)$	Transfer function of receive filter.
$\underset{\sim}{I}(\underline{\lambda})$, $\underset{\sim}{I}_M(\underline{\lambda})$	Fisher information matrix (FIM) and modified Fisher information matrix (MFIM) of the channel parameter vector $\underline{\lambda}$.

$\mathcal{I}(\underline{\lambda}_i)$	Initialization grid for the expectation maximization (EM) algorithm for the channel parameter $\underline{\lambda}_i$.
K	Oversampling factor.
k_{C}	Number of information bits in binary code word.
L	Length of frequency pulse $g(t)$.
$\mathcal{L}(\underline{\lambda})$	Likelihood function of $\underline{\lambda}$.
$m(\Psi'_n, A'_n, a'_n)$	Mapping function to trellis branch number γ_n.
$m_{\gamma_n}\left(\underline{S}'_n, \underline{S}'_{n+1}\right)$	Mapping from two connected states to the trellis branch γ_n.
$m_{a'_l}(\gamma_n)$	Mapping to the unipolar symbols $[a'_{n-(L-1)}, \ldots, a'_n]$.
$m_{\Psi'}(\gamma_n)$	Mapping to phase state Ψ'_n.
M	Modulation order.
n_{C}	Number of code bits in binary code word.
n_{CI}	Level of confidence interval in normal distribution, e.g. n_{CI} specifies the $3-\sigma$ confidence interval.
n_{MC}	Size of mother code.
N	Length of burst in symbols.
N_0, $\hat{N}_0$, $N_{0,\mathrm{old}}$	Complex noise power spectral density, its estimate and a trial value.
$N_{\mathrm{B}}(\nu)$, $N_{\mathrm{B}}(\tau)$	Number of bins in the sets $\mathcal{B}$ of desired frequency and timing interpolation points.
N_{C}	Number of code symbols in one burst.
$N_{\mathrm{I}}(\underline{\lambda}_i)$	Number of grid points in $\mathcal{I}(\underline{\lambda}_i)$.
N_{It}	Number of iterations in the expectation maximization (EM) algorithm.
N_{P}, N_{CZS}	Number of pilot and termination symbols in one burst.

$N_{S'}$, $N_{\Psi'}$, $N_{A'}$	Number of states, phase states and correlative states in the tilted CPM trellis.
N_{ZP}	Number of samples in a properly zero padded signal.
P	Nominator of modulation index.
P_{N}	Noise power.
P_{S}	Signal power.
$\mathcal{P}(\underline{s})$	Set of all predecessors to state $\underline{s}$.
$q(t)$	Phase pulse function.
Q	Denominator of modulation index.
$\mathcal{Q}(\underline{\lambda}, \underline{\lambda}_{\mathrm{old}}^{(i)})$	Auxiliary function that is used in the expectation maximization (EM) algorithm.
$r(t)$, $r'(t)$	Received burst CPM signal in the lowpass domain and its tilted version (depending on β).
$r'(kT_0)$, $\underline{r}'$	Sampled, tilted, received burst CPM signal in the lowpass domain and its vectorization.
R	Code rate $k_{\mathrm{C}}/n_{\mathrm{C}}$.
$s(t)$, $s'(t)$	Transmitted burst CPM signal in the lowpass domain (depending on β).
$\tilde{s}_{\mathrm{BP}}(t)$, $\tilde{s}_{\mathrm{LP}}(t)$	Transmitted continuous non-burst CPM signal in the bandpass and lowpass domain.
S_n, S'_n	n'th state in the trellis and n'th state in the tilted trellis.
$S_{ss}(f)$	Power spectral density (PSD) of the signal $s(t)$.
$\mathcal{S}(\underline{s})$	Set of all successors to state $\underline{s}$.
t	Time variable.

t_0	time reference of the receiver system.
T, T^{-1}	Duration of one CPM symbol and symbol rate.
T_0	Sample clock.
T_{Burst}	Burst duration.
T_{coh}	Coherence time of the channel.
$\mathcal{T}_{\mathrm{eff}}$	Effective Throughput.
$\underline{u}$, $\underline{\hat{u}}$	Vector of information bits and its estimate.
$\underline{v}$, $\underline{v}_{\mathrm{I}}$	Vector of encoded bits and its interleaved version.
$v_{\max}$	Maximum velocity.
$w(t)$, $w'(t)$	Time continuous, white noise and its tilted version.
$w'(kT_0)$	Sampled and tilted white noise.
α, $\hat{\alpha}$	Fading factor (FF) and its estimate.
β	$\beta = 1$ for transmitter tilting and $\beta = 0$ for receiver tilting.
γ_n	CPM trellis edge specifier in the n'th trellis stage.
Γ	Number of edges in the CPM trellis per stage.
$\Delta(\underline{\lambda}_i)$	Parameter ranges, as in which values can a channel parameter take.
$\varepsilon(\nu)$, $\varepsilon(\tau)$	Inverse of the desired resolutions of interpolations in frequency and timing estimation.
η_{M}, η_{S}	Modulation and spectral efficiency.
θ, $\hat{\theta}$, θ_0, $\hat{\theta}_0$	Carrier phase offset (CPO) and its estimate as well as the reduced versions ($\theta_0 = \theta \bmod \psi$).
κ	Scaling factor for Rayleigh distribution.
$\underline{\lambda}$, $\underline{\hat{\lambda}}$, $\underline{\lambda}_{\mathrm{old}}$	Channel parameter vector, its estimate and an old trial value.

ν, $\hat{\nu}$	Carrier frequency offset (CFO) and its estimate.
ξ	Phase tilt to make the CPM trellis time invariant.
$\sigma^2_{w'}$, $\sigma^2_{w'}$	Variance of complex noise process w' and its estimate.
τ, $\hat{\tau}$, τ_{old}, τ_{max}	Timing offset (TO), its estimate and a trial value. τ_{max} is the maximum value that the timing estimator in [38] can handle.
$\tilde{\phi}(t, \underline{\tilde{a}})$, $\phi(t, \underline{a})$, $\phi'(t, \underline{a}')$	Phase function of the continuous (non-burst) CPM signal, phase function of the burst CPM signal and phase function of the tilted burst CPM signal, respectively.
$\phi_{0,1}$	Arbitrary initial phase in the phase function.
$\chi_{r',\bar{c}}(\nu, \tau)$	Cross ambiguity function.
ψ	Radiant value of a phase state.
Ψ_n , Ψ'_n	n'th phase state of the regular and tilted CPM trellis, respectively.

List of Figures

List of Tables

Bibliography

[1] A. Hansson, J. Nilsson, and K. Wiklundh, "Performance Analysis of Frequency-Hopping Ad Hoc Networks With Random Dwell-Time Under Follower Jamming," in *MILCOM 2015-2015 IEEE Military Communications Conference*, pp. 848–853, IEEE, 2015.

[2] A. Valyrakis, E. E. Tsakonas, N. D. Sidiropoulos, and A. Swami, "Stochastic Modeling and Particle Filtering Algorithms for Tracking a Frequency-Hopped Signal," *IEEE Transactions on Signal Processing*, vol. 57, no. 8, pp. 3108–3118, 2009.

[3] C. Brown and P. Vigneron, "Spectrally Efficient CPM Waveforms for Narrowband Tactical Communications in Frequency Hopped Networks," in *MILCOM 2006*, pp. 1–6, IEEE Computer Society, 2006.

[4] D. J. Torrieri, "Fundamental Limitations on Repeater Jamming of Frequency-Hopping Communications," *IEEE Journal on Selected areas in Communications*, vol. 7, no. 4, pp. 569–575, 1989.

[5] B. SIG, "Bluetooth Core Specification Version 5.3, 2021."

[6] NATO, "Ratification Draft STANAGs 5630-5633, Narrowband Waveform," tech. rep., NATO, 2015.

[7] E. Casini, M. Street, P. Vigneron, and R. Barfoot, "SDR-Ready Standardized Waveforms for Tactical VHF and UHF Communications for NATO," tech. rep., NATO C3 AGENCY THE HAGUE (NETHERLANDS), 2010.

[8] C. Enz, N. Scolari, and U. Yodprasit, "Ultra Low-Power Radio Design for Wireless Sensor Networks," in *2005 IEEE International Wkshp on Radio-Frequency Integration Technology: Integrated Circuits for Wideband Comm & Wireless Sensor Networks*, pp. 1–17, IEEE, 2005.

[9] C. Bockelmann, N. Pratas, H. Nikopour, K. Au, T. Svensson, C. Stefanovic, P. Popovski, and A. Dekorsy, "Massive Machine-Type Communications in 5G: Physical and MAC-Layer Solutions," *IEEE Communications Magazine*, vol. 54, no. 9, pp. 59–65, 2016.

[10] A. G. Perotti, M. N. Khormuji, and B. M. Popović, "Simultaneous Wireless Information and Power Transfer by Continuous-Phase Modulation," *IEEE Communications Letters*, vol. 24, no. 6, pp. 1294–1298, 2020.

[11] J. Proakis and M. Salehi, *Digital Communications*. McGraw-Hill., 2008.

[12] R. C. Daniels and R. W. Heath, "60 GHz Wireless Communications: Emerging Requirements and Design Recommendations," *IEEE Vehicular technology magazine*, vol. 2, no. 3, pp. 41–50, 2007.

[13] M. K. Simon, *Bandwidth-Ffficient Digital Modulation With Application to Deep Space Communications*. John Wiley & Sons, 2005.

[14] J. Huber, "Trelliscodierung," Springer, 1992.

[15] T. Svensson and T. Eriksson, "On Power Amplifier Efficiency With Modulated Signals," in *2010 IEEE 71st Vehicular Technology Conference*, pp. 1–5, IEEE, 2010.

[16] F. J. Ortega-Gonzalez, "High Power Wideband Class-E Power Amplifier," *IEEE Microwave and Wireless Components Letters*, vol. 20, no. 10, pp. 569–571, 2010.

[17] U. Mengali, *Synchronization Techniques for Digital Receivers*. Springer Science & Business Media, 2013.

[18] R. Mehlan, Y.-E. Chen, and H. Meyr, "A Fully Digital Feedforward MSK Demodulator with Joint Frequency Offset and Symbol Timing Estimation for Burst Mode Mobile Radio," *IEEE Transactions on Vehicular Technology*, vol. 42, no. 4, pp. 434–443, 1993.

[19] M. Morelli and U. Mengali, "Feedforward Carrier Frequency Estimation with MSK-Type Signals," *IEEE Communications Letters*, vol. 2, no. 8, pp. 235–237, 1998.

[20] M. Morelli and G. M. Vitetta, "Joint Phase and Timing Synchronization Algorithms for MSK-Type Signals," in *1999 IEEE Communications Theory Mini-Conference (Cat. No. 99EX352)*, pp. 146–150, IEEE, 1999.

[21] A. D'andrea, A. Ginesi, and U. Mengali, "Frequency Detectors for CPM Signals," *IEEE Transactions on Communications*, vol. 43, no. 2/3/4, pp. 1828–1837, 1995.

[22] A. D'andrea, A. Ginesi, and U. Mengali, "Digital Carrier Frequency Estimation for Multilevel CPM Signals," in *Proceedings IEEE International Conference on Communications ICC'95*, vol. 2, pp. 1041–1045, IEEE, 1995.

[23] A. N. D'Andrea, U. Mengali, and M. Morelli, "Symbol Timing Estimation with CPM Modulation," *IEEE Transactions on Communications*, vol. 44, no. 10, pp. 1362–1372, 1996.

[24] P. Bianchi, P. Loubaton, and F. Sirven, "On the Blind Estimation of the Parameters of Continuous Phase Modulated Signals," *IEEE Journal on Selected Areas in Communications*, vol. 23, no. 5, pp. 944–962, 2005.

[25] A. D'Amico, A. N. D'Andrea, and U. Mengali, "Feedforward Synchronization Schemes for MSK-Type Signals," *European Transactions on Telecommunications*, vol. 10, no. 6, pp. 597–607, 1999.

[26] A. Lang and B. Lankl, "Carrier Phase Synchronization in Burst-Type CPM by Means of Expectation Maximization," in *2019 IEEE 90th Vehicular Technology Conference (VTC2019-Fall)*, pp. 1–7, IEEE, 2019.

[27] A. Lang and B. Lankl, "Carrier Frequency Offset Estimation in Burst-Type CPM via the EM Algorithm," in *2020 IEEE 91st Vehicular Technology Conference (VTC2020-Spring)*, pp. 1–6, IEEE, 2020.

[28] A. Lang and B. Lankl, "Blind Joint Timing and Carrier Phase Estimation in Burst-Type CPM," in *2021 IEEE 94th Vehicular Technology Conference (VTC2021-Fall)*, pp. 1–6, IEEE, 2021.

[29] J. Huber and W. Liu, "Data-Aided Synchronization of Coherent CPM-Receivers," *IEEE Transactions on Communications*, vol. 40, no. 1, pp. 178–189, 1992.

[30] H. Abeida, J.-M. Brossier, L. Ros, and J. Vila-Valls, "An EM Algorithm for Path Delay and Complex Gain Estimation of Slowly Varying Fading Channel for CPM Signals," in *GLOBECOM 2009-2009 IEEE Global Telecommunications Conference*, pp. 1–6, IEEE, 2009.

[31] Q. Zhao and G. Stuber, "Joint Time and Phase Recovery for CPM and its Asymptotic Behavior," in *GLOBECOM'05 IEEE Global Telecommunications Conference, 2005*, vol. 4, pp. 5–pp, IEEE, 2005.

[32] Q. Zhao and G. L. Stuber, "Robust Time and Phase Synchronization for Continuous Phase Modulation," *IEEE Transactions on Communications*, vol. 54, no. 10, pp. 1857–1869, 2006.

[33] N. Hajiabdolrahim, S. R. Aghdam, and T. Eriksson, "An Extended Kalman Filter Framework for Joint Phase Noise, CFO and Sampling Time Error Estimation," in *2020 IEEE 31st Annual International Symposium on Personal, Indoor and Mobile Radio Communications*, pp. 1–6, IEEE, 2020.

[34] Q. Zhao, H. Kim, and G. Stuber, "Adaptive Iterative Phase Synchronization for Serially Concatenated Continuous Phase Modulation," in *IEEE Military Communications Conference, 2003. MILCOM 2003.*, vol. 1, pp. 78–83, IEEE, 2003.

[35] Q. Zhao, H. Kim, and G. Stuber, "Innovations-Based MAP Estimation with Application to Phase Synchronization for Serially Concatenated CPM," *IEEE Transactions on Wireless Communications*, vol. 5, no. 5, pp. 1033–1043, 2006.

[36] Q. Zhao and G. L. Stuber, "Turbo Synchronization for Serially Concatenated CPM," in *2006 IEEE International Conference on Communications*, vol. 7, pp. 2976–2980, IEEE, 2006.

[37] X. Zhou, R. Xue, D. Zhao, and F. Fang, "Soft Timing Synchronization Algorithm for CPM Signals," in *2013 IEEE International Conference on Green Computing and Communications and IEEE Internet of Things and IEEE Cyber, Physical and Social Computing*, pp. 1381–1385, IEEE, 2013.

[38] E. Hosseini and E. Perrins, "Timing, Carrier, and Frame Synchronization of Burst-Mode CPM," *IEEE Transactions on Communications*, vol. 61, no. 12, pp. 5125–5138, 2013.

[39] C. Brown and P. Vigneron, "Coarse and Fine Timing Sychronisation for Partial Response CPM in a Frequency Hopped Tactical Network," in *MILCOM 2007-IEEE Military Communications Conference*, pp. 1–7, IEEE, 2007.

[40] N. Noels, C. Herzet, A. Dejonghe, V. Lottici, H. Steendam, M. Moeneclaey, M. Luise, and L. Vandendorpe, "Turbo Synchronization: An EM Algorithm Interpretation," in *Communications, 2003. ICC'03. IEEE International Conference on*, vol. 4, pp. 2933–2937, IEEE, 2003.

[41] N. Noels, V. Lottici, A. Dejonghe, H. Steendam, M. Moeneclaey, M. Luise, and L. Vandendorpe, "A Theoretical Framework for Soft-Information-Based Synchronization in Iterative (Turbo) Receivers," *EURASIP Journal on Wireless Communications and Networking*, vol. 2005, no. 2, p. 576206, 2005.

[42] C. Berrou, A. Glavieux, and P. Thitimajshima, "Near Shannon Limit Error-Correcting Coding and Decoding: Turbo-Codes. 1," in *Proceedings of ICC'93-IEEE International Conference on Communications*, vol. 2, pp. 1064–1070, IEEE, 1993.

[43] K.-D. Kammeyer, *Nachrichtenübertragung*. Springer-Verlag, 2013.

[44] P. Moqvist and T. M. Aulin, "Serially Concatenated Continuous Phase Modulation with Iterative Decoding," *IEEE Transactions on Communications*, vol. 49, no. 11, pp. 1901–1915, 2001.

[45] A. Lang and B. Lankl, "Tail-Biting in Optimal CPM Detectors," *Electronics Letters*, vol. 56, no. 3, pp. 141–144, 2019.

[46] J. Fischer, *Physical layer link modeling for mobile ad-hoc networks (MANET)*. Fraunhofer Verlag, 2016.

[47] Ettus Research, *USRP™ N200/N210 NETWORKED SERIES*, 2012.

[48] L. Bahl, J. Cocke, F. Jelinek, and J. Raviv, "Optimal Decoding of Linear Codes for Minimizing Symbol Error Rate (corresp.)," *IEEE Transactions on Information Theory*, vol. 20, no. 2, pp. 284–287, 1974.

[49] A. Lang and B. Lankl, "A Comprehensive Study of CPM Trellis Initialization Methods," in *MILCOM 2022 - 2022 IEEE Military Communications Conference (MILCOM)*, pp. 78–83, IEEE, 2022.

[50] P. Moqvist and T. Aulin, "Trellis Termination in CPM," *Electronics Letters*, vol. 36, no. 23, pp. 1940–1941, 2000.

[51] V. Dantona, C. Hofmann, S. Lattrell, and B. Lankl, "Spectrally Efficient Multilevel CPM Waveforms for VHF Narrowband Communications," in *SCC 2015; 10th International ITG Conference on Systems, Communications and Coding; Proceedings of*, pp. 1–6, VDE, 2015.

[52] E. Arikan, "Channel Polarization: A Method for Constructing Capacity-Achieving Codes for Symmetric Binary-Input Memoryless Channels," *IEEE Transactions on Information Theory*, vol. 55, no. 7, pp. 3051–3073, 2009.

[53] A. Lang and B. Lankl, "Very Short Channel Codes for Burst-Type Transmissions," in *MILCOM 2018-2018 IEEE Military Communications Conference (MILCOM)*, pp. 1–9, IEEE, 2018.

[54] E. Arikan, "Systematic Polar Coding," *IEEE Communications Letters*, vol. 15, no. 8, pp. 860–862, 2011.

[55] U. U. Fayyaz and J. R. Barry, "Low-Complexity Soft-Output Decoding of Polar Codes," *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 5, pp. 958–966, 2014.

[56] V. Bioglio, F. Gabry, and I. Land, "Low-Complexity Puncturing and Shortening of Polar Codes," in *2017 IEEE Wireless Communications and Networking Conference Workshops (WCNCW)*, pp. 1–6, IEEE, 2017.

[57] R. M. Pyndiah, "Near-Optimum Decoding of Product Codes: Block Turbo Codes," *IEEE Transactions on Communications*, vol. 46, no. 8, pp. 1003–1010, 1998.

[58] H. L. Van Trees, "Detection Estimation and Modulation Theory," *Ser. Detection, Estimation, and Modulation Theory*, no. 1, 1968.

[59] A. N. D'Andrea, U. Mengali, and R. Reggiannini, "The Modified Cramer-Rao Bound and Its Application to Synchronization Problems," *IEEE Transactions on Communications*, vol. 42, no. 234, pp. 1391–1399, 1994.

[60] F. Gini, R. Reggiannini, and U. Mengali, "The Modified Cramer-Rao Bound in Vector Parameter Estimation," *IEEE Transactions on Communications*, vol. 46, no. 1, pp. 52–60, 1998.

[61] A. Pollok and R. McKilliam, "Modified Cramér-Rao Bounds for Continuous-Phase Modulated Signals," *IEEE Transactions on Communications*, vol. 62, no. 5, pp. 1681–1690, 2014.

[62] M. Moeneclaey, "On the true and the modified Cramer-Rao bounds for the estimation of a scalar parameter in the presence of nuisance parameters," *IEEE Transactions on Communications*, vol. 46, no. 11, pp. 1536–1544, 1998.

[63] T. Alberty, "Frequency Domain Interpretation of the Cramer-Rao Bound for Carrier and Clock Synchronization," *IEEE Transactions on Communications*, vol. 43, no. 2/3/4, pp. 1185–1191, 1995.

[64] E. Hosseini and E. Perrins, "The Cramer-Rao Bound for Training Sequence Design for Burst-Mode CPM," *IEEE Transactions on Communications*, vol. 61, no. 6, pp. 2396–2407, 2013.

[65] T. Alberty, *Zur Empfängersynchronisation bei der digitalen Nachrichtenübertragung*. N/A, 1996.

[66] A. Lang and B. Lankl, "Blind Vector Parameter Estimation for Burst Type CPM Transmissions," *IEEE Open Journal of Vehicular Technology*, vol. 4, pp. 162–180, 2022.

[67] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum Likelihood from Incomplete Data via the EM Algorithm," *Journal of the Royal Statistical Society: Series B (Methodological)*, vol. 39, no. 1, pp. 1–22, 1977.

[68] C. M. Bishop, *Pattern Recognition and Machine Learning*. Berlin-Heidelberg: Springer New York, 2016.

[69] M. R. Gupta, Y. Chen, *et al.*, "Theory and Use of the EM Algorithm," *Foundations and Trends® in Signal Processing*, vol. 4, no. 3, pp. 223–296, 2011.

[70] L. R. Rabiner, "A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition," *Proceedings of the IEEE*, vol. 77, no. 2, pp. 257–286, 1989.

[71] P. Laurent, "Exact and Approximate Construction of Digital Phase Modulations by Superposition of Amplitude Modulated Pulses (AMP)," *IEEE transactions on communications*, vol. 34, no. 2, pp. 150–160, 1986.

[72] U. Mengali and M. Morelli, "Decomposition of M-ary CPM Signals Into PAM Waveforms," *IEEE Transactions on Information Theory*, vol. 41, no. 5, pp. 1265–1275, 1995.

[73] S. B. Rasool and M. R. Bell, "Efficient Pulse-Doppler Processing and Ambiguity Functions of Nonuniform Coherent Pulse Trains," in *2010 IEEE Radar Conference*, pp. 1150–1155, IEEE, 2010.

[74] Z. Gao, R. Tao, Y. Ma, and T. Shao, "DVB-T Signal Cross-Ambiguity Functions Improvement for Passive Radar," in *2006 CIE International Conference on Radar*, pp. 1–4, IEEE, 2006.

[75] S. Stein, "Algorithms for ambiguity function processing," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 29, no. 3, pp. 588–599, 1981.

[76] X.-L. Meng and D. B. Rubin, "Maximum Likelihood Estimation via the ECM Algorithm: A General Framework," *Biometrika*, vol. 80, no. 2, pp. 267–278, 1993.

[77] P. Robertson, E. Villebrun, and P. Hoeher, "A comparison of optimal and sub-optimal MAP decoding algorithms operating in the log domain," in *Proceedings IEEE International Conference on Communications ICC'95*, vol. 2, pp. 1009–1013, IEEE, 1995.

[78] L. Svilainis, "Review on Time Delay Estimate Subsample Interpolation in Frequency Domain," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency control*, vol. 66, no. 11, pp. 1691–1698, 2019.

[79] G. D. Martin, "Chirp Z-Transform Spectral Zoom Optimization with MATLAB," *Sandia National Laboratories Report SAND2005-7084*, 2005.

[80] G. Goertzel *et al.*, "An algorithm for the evaluation of finite trigonometric series," *The American Mathematical Monthly*, vol. 65, no. 1, pp. 34–35, 1958.

[81] P. Duhamel, B. Piron, and J. M. Etcheto, "On Computing the Inverse DFT," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 36, no. 2, pp. 285–286, 1988.

[82] H. Vangala, E. Viterbo, and Y. Hong, "A Comparative Study of Polar Code Constructions for the AWGN Channel," *arXiv preprint arXiv:1501.02473*, 2015.

[83] S. Ten Brink, "Convergence Behavior of Iteratively Decoded Parallel Concatenated Codes," *IEEE Transactions on Communications*, vol. 49, no. 10, pp. 1727–1737, 2001.

[84] S. Dolinar, D. Divsalar, and F. Pollara, "Code Performance as a Function of Block Size," *TMO progress report*, vol. 42, no. 133, 1998.

[85] L. M. Supplee, R. P. Cohn, J. S. Collura, and A. V. McCree, "MELP: The New Federal Standard at 2400 bps," in *1997 IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 2, pp. 1591–1594, IEEE, 1997.

[86] D. J. Rahikka, J. S. Collura, T. E. Fuja, and T. Fazel, "US Federal Standard MELP Vocoder Tactical Performance Enhancement via MAP Error Correction," in *MILCOM 1999. IEEE Military Communications. Conference Proceedings (Cat. No. 99CH36341)*, vol. 2, pp. 1458–1462, IEEE, 1999.

[87] S. Lin and D. J. Costello, *Error Control Coding*, vol. 2. Prentice Hall New York, 2001.

[88] J. Xu, H. Shi, and J. Wang, "Analysis of Frame Length and Frame Error Rate for the Lowest Energy Dissipation in Wireless Sensor Networks," in *2008 4th International Conference on Wireless Communications, Networking and Mobile Computing*, pp. 1–4, IEEE, 2008.

[89] MATLAB, *version 9.13.0.2105380 (R2022b) Update 2*. Natick, Massachusetts: The MathWorks Inc., 2022.

[90] L. G. Baltar, F. Schaich, M. Renfors, and J. A. Nossek, "Computational Complexity Analysis of Advanced Physical Layers Based on Multicarrier Modulation," in *2011 Future Network & Mobile Summit*, pp. 1–8, IEEE, 2011.

[91] J. Anderson and D. Taylor, "A Bandwidth-Efficient Class of Signal-Space Codes," *IEEE Transactions on Information Theory*, vol. 24, no. 6, pp. 703–712, 1978.

[92] Z. Pan, C. Xie, H. Wang, Y. Wei, and D. Guo, "Blind Turbo Equalization of Short CPM Bursts for UAV-Aided Internet of Things," *Sensors*, vol. 22, no. 17, p. 6508, 2022.

[93] G. K. Kaleh and R. Vallet, "Joint Parameter Estimation and Symbol Detection for Linear or Nonlinear Unknown Channels," *IEEE Transactions on Communications*, vol. 42, no. 7, pp. 2406–2413, 1994.

[94] C. Brutel, J. Boutros, and P. Mege, "Iterative Joint Channel Estimation and Detection of Coded CPM," in *2000 International Zurich Seminar on Broadband Communications. Accessing, Transmission, Networking. Proceedings (Cat. No. 00TH8475)*, pp. 287–292, IEEE, 2000.

[95] K.-D. Kammeyer, V. Kuhn, and T. Petermann, "Blind and Nonblind Turbo Estimation for Fast Fading GSM Channels," *IEEE Journal on Selected Areas in Communications*, vol. 19, no. 9, pp. 1718–1728, 2001.

[96] L. M. Zeger and H. Kobayashi, "A Simplified EM Algorithm for Detection of CPM Signals In A Fading Multipath Channel," *Wireless Networks*, vol. 8, no. 6, pp. 649–658, 2002.

[97] H. Nguyen and B. C. Levy, "Blind ML Detection of CPM Signals via he EMV Algorithm," in *2002 IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 3, pp. III–2457, IEEE, 2002.

[98] H. Nguyen and B. C. Levy, "Blind and semi-Blind Equalization of CPM Signals with the EMV Algorithm," *IEEE Transactions on Signal Processing*, vol. 51, no. 10, pp. 2650–2664, 2003.

[99] M. El Chamaa and B. Lankl, "Noncoherent Symbol Detection of Short CPM Bursts in Frequency-Selective Fading Channels," *IEEE Transactions on Wireless Communications*, vol. 19, no. 2, pp. 771–782, 2019.

[100] G. Colavolpe and R. Raheli, "Reduced-Complexity Detection and Phase Synchronization of CPM Signals," *IEEE Transactions on Communications*, vol. 45, no. 9, pp. 1070–1079, 1997.

[101] A. Barbieri and G. Colavolpe, "Simplified Soft-Output Detection of CPM Signals Over Coherent and Phase Noise Channels," *IEEE Transactions on Wireless Communications*, vol. 6, no. 7, pp. 2486–2496, 2007.

[102] M. Mitchell, *An Introduction to Genetic Algorithms*. MIT press, 1998.