

RESEARCH

Open Access

Joint communication and positioning based on soft channel parameter estimation

Kathrin Schmeink*, Rebecca Adam and Peter Adam Hoehner

Abstract

A joint communication and positioning system based on maximum-likelihood channel parameter estimation is proposed. The parameters of the physical channel, needed for positioning, and the channel coefficients of the equivalent discrete-time channel model, needed for communication, are estimated jointly using a priori information about pulse shaping and receive filtering. The paper focusses on the positioning part of the system. It is investigated how soft information for the parameter estimates can be obtained. On the basis of confidence regions, two methods for obtaining soft information are proposed. The accuracy of these approximative methods depends on the nonlinearity of the parameter estimation problem, which is analyzed by so-called curvature measures. The performance of the two methods is investigated by means of Monte Carlo simulations. The results are compared with the Cramer-Rao lower bound. It is shown that soft information aids the positioning. Negative effects caused by multipath propagation can be mitigated significantly even without oversampling.

1 Introduction

Interest in joint communication and positioning is steadily increasing [1]. Synergetic effects like improved resource allocation and new applications like location-based services or a precise location determination of emergency calls are attractive features of joint communication and positioning. Since the system requirements of communication and positioning are quite different, it is a challenging task to combine them: Communication aims at high data rates with little training overhead. Only the channel coefficients of the equivalent discrete-time channel model, which includes pulse shaping and receive filtering in addition to the physical channel, need to be estimated for data detection. In contrast, positioning aims at precise position estimates. Therefore, parameters of the physical channel like the time of arrival (TOA) or the angle of arrival (AOA) need to be estimated as accurately as possible [2,3]. Significant training is typically spent for this purpose.

In this paper, a joint communication and positioning system based on maximum-likelihood channel parameter estimation is suggested [4]. The estimator exploits the fact that channel and parameter estimation are closely related. The parameters of the physical channel and

the channel coefficients of the equivalent discrete-time channel model are estimated jointly by utilizing a priori information about pulse shaping and receive filtering. Hence, training symbols that are included in the data burst aid both communication *and* positioning.

On the one hand, in [5-7], it is proposed to use a priori information about pulse shaping and receive filtering in order to improve the estimates of the equivalent discrete-time channel model. However, the information about the physical channel is neglected in these publications. On the other hand, channel sounding is performed in order to estimate the parameters of the physical channel [8-10]. But, to the authors best knowledge, the proposed parameter estimation methods are not applied for estimation of the equivalent discrete-time channel model. The estimator proposed in this paper combines both approaches: Channel estimation is mandatory for communication purposes. By exploiting a priori information about pulse shaping and receive filtering, the channel coefficients can be estimated more precisely *and* positioning is enabled. Hence, synergy is created.

This paper focusses on the positioning part of the proposed joint communication and positioning system. Most positioning methods suffer from a bias introduced by multipath propagation. Multipath mitigation is, thus, an important issue. The proposed channel parameter

* Correspondence: kas@tf.uni-kiel.de
Information and Coding Theory Lab Faculty of Engineering, University of Kiel
Kaiserstrasse 2, 24143 Kiel, Germany

estimator performs multipath mitigation in two ways: First, the maximum-likelihood estimator is able to take all relevant multipath components into account in order to minimize the modeling error. Second, soft information can be obtained for the parameter estimates. Soft information corresponds to the variance of an estimate and is a measure of reliability. This information can be exploited by a weighted positioning algorithm in order to improve the accuracy of the position estimate.

On the basis of confidence regions, two different methods for obtaining soft information are proposed: The first method is based on a linearization of the nonlinear parameter estimation problem and the second method is based on the likelihood concept. For linear estimation problems, an exact covariance matrix can be determined in closed form. For nonlinear estimation problems, as it is the case for channel parameter estimation, there are different approximations to the covariance matrix, which are based on a linearization. These approximate covariance matrices are generated by most nonlinear least-squares solvers (e.g., Levenberg-Marquardt method) anyway and can be used after further analysis [11]. Confidence regions based on the likelihood method are more robust than those based on approximate covariance matrices since they do not rely on a linearization, but they are also more complex to calculate. Heuristic optimization methods like genetic algorithms or particle swarm optimization offer a comfortable procedure to determine the likelihood confidence region as demonstrated in [12]. Both methods are only approximate, and their accuracy depends on the nonlinearity of the estimation problem. In [13], Bates and Watts introduce curvature measures that indicate the amount of nonlinearity. These measures can be used to diagnose the accuracy of the proposed methods.

The remainder of this paper is organized as follows: The system and channel model is described in Section 2. The relationship between channel and parameter

estimation is explained and the nonlinear metric of the maximum-likelihood estimator is derived. General aspects concerning nonlinear optimization are discussed. In Section 3, the concept of soft information is introduced. Based on confidence regions, two methods for obtaining soft information concerning the parameter estimates are proposed. In order to further analyze the proposed methods, the curvature measures of Bates and Watts are introduced in Section 4. The curvature measures are calculated for the parameter estimation problem and a first analysis of the problem is given. Afterward, positioning based on the TOA is explained in Section 5, and the performance of the two soft information methods is investigated by means of Monte Carlo simulations. The results are compared with the Cramer-Rao lower bound. Finally, conclusions are drawn in Section 6.

2 System Concept

2.1 System and channel model

Throughout this paper, the discrete-time complex baseband notation is used. Let $x[k]$ denote the k th modulated and coded symbol of a data burst of length K . Some symbols $x[k]$ are known at the receiver side ("training symbols"), whereas others are not known ("data symbols"). It is assumed that data and training symbols can be separated perfectly at the receiver side. The received sample $y[k]$ at time index k can be written as

$$y[k] = \sum_{l=0}^L h_l[k] \cdot x[k-l] + n[k], \quad 0 \leq k \leq K + L - 1, \quad (1)$$

where $h_l[k]$ is the l th channel coefficient of the equivalent discrete-time channel model with effective channel memory length L , and $n[k]$ is a Gaussian noise sample with zero mean and variance σ_n^2 . The noise process is assumed to be white. In Figure 1, the relationship between the physical channel and the equivalent

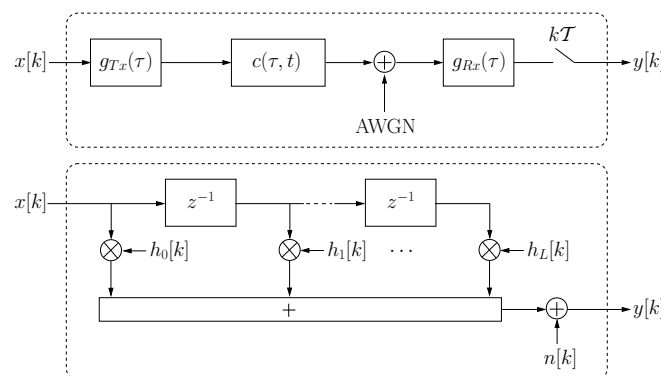


Figure 1 Equivalent discrete-time channel model.

discrete-time channel model is shown. The input/output behavior of the continuous-time channel is *exactly* represented by the equivalent discrete-time channel model, which is described by an FIR filter with coefficients $h_l[k]$. The delay elements z^{-1} correspond to the sampling rate $\frac{1}{T_s}$. In this paper, only symbol-rate sampling $\mathcal{T} = \mathcal{T}_s$ is considered, where T_s is the symbol duration.^a The channel coefficients $h_l[k]$ are samples of the overall impulse response of the continuous-time channel. This impulse response is given by the convolution of the known pulse shaping filter $g(\tau) = g_{Tx}(\tau)$, the unknown physical channel $c(\tau, t)$, and the known receive filter $g_{Rx}(\tau)$. Since the convolution is associative and commutative, pulse shaping and receive filtering can be combined: $g(\tau) = g_{Tx}(\tau) * g_{Rx}(\tau)$, where $*$ denotes the convolution.

The physical channel can be modeled by a weighted sum of delayed Dirac impulses:

$$c(\tau, t) = \sum_{\mu=1}^M f_{\mu}(t) \cdot \delta(\tau - \tau_{\mu}(t)), \quad (2)$$

where M is the number of resolvable propagation paths. The parameters $f_{\mu}(t)$ and $\tau_{\mu}(t)$ denote the complex amplitude and the propagation delay of the μ th path at time t , respectively. Without loss of generality, it is assumed that the multipath components are sorted according to ascending delay: $\tau_1(t) < \tau_2(t) < \dots < \tau_M(t)$. The delay of the first arriving path is called TOA. Positioning is based on the assumption that the TOA corresponds to the distance between transmitter and receiver. This is only true if a line-of-sight (LOS) path exists. In urban or indoor environments, the LOS path is often blocked. In these so-called non-LOS (NLOS) scenarios, the modeling error reduces the positioning accuracy significantly. Additionally, positioning typically suffers from a bias introduced by multipath propagation *even* if a LOS path exists. In order to analyze the multipath mitigation ability of the proposed soft channel parameter estimator, this paper restricts itself to LOS scenarios. However, the influence of NLOS is discussed in Section 5.2.

Given $c(\tau, t)$ and $g(\tau)$, the overall channel impulse response $h(\tau, t)$ can be written as

$$h(\tau, t) = c(\tau, t) * g(\tau) = \sum_{\mu=1}^M f_{\mu}(t) \cdot g(\tau - \tau_{\mu}(t)). \quad (3)$$

After symbol-rate sampling (3) at $t = kT_s$, the channel coefficients can be represented as:

$$h_l[k] = \sum_{\mu=1}^M f_{\mu}[k] \cdot g(lT_s - \tau_{\mu}[k]). \quad (4)$$

In the following, it is assumed that the channel is quasi time-invariant over the training length (block fading). Thus, the time index k in (4) can be omitted.

For simulation of communication systems, it is sufficient to consider excess delays. Without loss of generality, $\tau_1 = 0$ can be assumed then. The effective channel memory length L is, therefore, determined by the excess delay $\tau_M - \tau_1$ plus the effective width T_g of $g(\tau)$.

In case of positioning based on the TOA, however, it is important taking into account that $\tau_1 = \frac{d}{c}$, where d is the distance between transmitter and receiver and c is the speed of light. Denoting the maximum possible delay by τ_M^{\max} , the maximum possible channel memory length can be pre-calculated according to

$$L = \left\lceil \frac{\tau_M^{\max} + T_g}{T_s} \right\rceil. \quad (5)$$

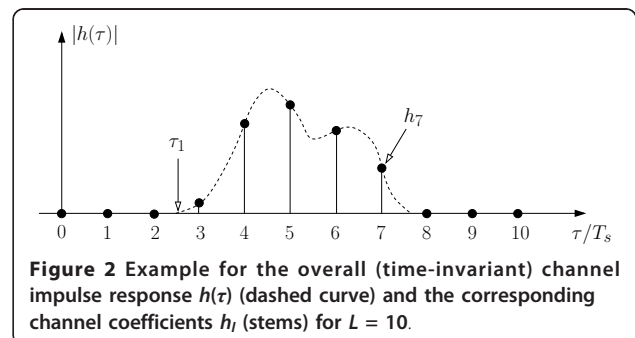
This channel memory length covers all possible propagation scenarios including the worst case. Hence, the channel impulse response is embedded in a sequence of zeros as shown in Figure 2.

2.2 Channel parameter estimation

Channel estimation is mandatory for data detection. Typically, training symbols are inserted in the data burst for estimation of the equivalent discrete-time channel model. If the channel is quasi time-invariant over the training sequence (block fading), least-squares channel estimation (LSCE) can be applied. In this paper, a training preamble of length K_t is assumed. For the interval $L \leq k \leq K_t - 1$, the received samples according to (1) can be expressed in vector/matrix notation as

$$\mathbf{y} = \mathbf{X}\mathbf{h} + \mathbf{n}, \quad (6)$$

where \mathbf{X} is the training matrix with Toeplitz structure, $\mathbf{y} = [y[L], y[L + 1], \dots, y[K_t - 1]]^T$ is the observation vector, $\mathbf{h} = [h_0, h_1, \dots, h_L]^T$ is the channel coefficient vector, and \mathbf{n} is a zero mean Gaussian noise vector with



covariance matrix $C_n = \sigma_n^2 \mathbf{I}$. The least-squares channel estimates are given by

$$\tilde{\mathbf{h}} = (\mathbf{X}^H \mathbf{X})^{-1} \mathbf{X}^H \mathbf{y} = \mathbf{h} + \boldsymbol{\varepsilon}. \quad (7)$$

Using the assumptions above, the estimation error $\boldsymbol{\varepsilon}$ is zero mean and Gaussian with covariance matrix $C_\varepsilon = \sigma_n^2 (\mathbf{X}^H \mathbf{X})^{-1}$ [14]. For a pseudo-random training sequence, the matrix $(\mathbf{X}^H \mathbf{X})$ becomes a scaled identity matrix with scaling factor $K_t - L$, and the covariance matrix of the estimation error reduces to $C_\varepsilon = \frac{\sigma_n^2}{K_t - L} \mathbf{I} = \sigma_\varepsilon^2 \mathbf{I}$.

The main idea of joining communication and positioning is based on the relationship in (4). If the parameters of the physical channel are stacked into a vector $\boldsymbol{\theta} = [\text{Re}\{f_1\}, \text{Im}\{f_1\}, \tau_1, \text{Re}\{f_2\}, \dots, \tau_M]$, (4) can be rewritten as:

$$h_l(\boldsymbol{\theta}) = \sum_{\substack{\mu=1 \\ v=3(\mu-1)}}^M (\theta_{v+1} + j\theta_{v+2}) g(lT_s - \theta_{v+3}). \quad (8)$$

The parameters $\boldsymbol{\theta}$ can be estimated by fitting the model function (8) to the least-squares channel estimates \tilde{h}_l . Hence, the channel estimates are not only used for data detection, but they are also exploited for positioning. Furthermore, refined channel estimates \hat{h}_l are obtained by evaluating (8) for the parameter estimate $\hat{\boldsymbol{\theta}}$ [4].^b On the one hand, positioning is enabled since the TOA τ_1 is estimated. On the other hand, data detection can be improved because refined channel estimates are obtained.

The maximum-likelihood estimate $\hat{\boldsymbol{\theta}}$ is given by the set $\boldsymbol{\theta}$ that maximizes the likelihood function [14]

$$\begin{aligned} p(\tilde{\mathbf{h}}; \boldsymbol{\theta}) &= \prod_{l=0}^L \frac{1}{\pi \sigma_\varepsilon^2} \exp \left\{ -\frac{1}{\sigma_\varepsilon^2} \left| \tilde{h}_l - h_l(\boldsymbol{\theta}) \right|^2 \right\} \\ &= \left(\frac{1}{\pi \sigma_\varepsilon^2} \right)^{L+1} \exp \left\{ -\frac{1}{\sigma_\varepsilon^2} \sum_{l=0}^L \left| \tilde{h}_l - h_l(\boldsymbol{\theta}) \right|^2 \right\}. \end{aligned} \quad (9)$$

For LSCE with pseudo-random training, this is equivalent to maximizing the likelihood function

$$\begin{aligned} p(\mathbf{y}; \boldsymbol{\theta}) &= \prod_{k=L}^{K_t-1} \frac{1}{\pi \sigma_n^2} \exp \left\{ -\frac{1}{\sigma_n^2} \left| y[k] - \sum_{l=0}^L h_l(\boldsymbol{\theta}) x[k-l] \right|^2 \right\} \\ &= \left(\frac{1}{\pi \sigma_n^2} \right)^{K_t-L} \exp \left\{ -\frac{1}{\sigma_n^2} \sum_{k=L}^{K_t-1} \left| y[k] - \sum_{l=0}^L h_l(\boldsymbol{\theta}) x[k-l] \right|^2 \right\}. \end{aligned} \quad (10)$$

with respect to $\boldsymbol{\theta}$. The second approach in (10) may seem more natural to some readers since the parameters are estimated directly from the received samples. But since both approaches are equivalent, as proven in the

“Appendix”, it seems more convenient to the authors to apply the first approach: Channel estimates are usually already available in communication systems and the metric derived from (9) is less complex than the metric derived from (10). Hence, only the first approach is considered in the following.

Since the noise is assumed to be Gaussian, the maximum-likelihood estimator corresponds to the least-squares estimator:

$$\begin{aligned} \hat{\boldsymbol{\theta}} &= \arg \max_{\tilde{\boldsymbol{\theta}}} \left\{ p(\tilde{\mathbf{h}}; \tilde{\boldsymbol{\theta}}) \right\} = \arg \max_{\tilde{\boldsymbol{\theta}}} \left\{ \ln p(\tilde{\mathbf{h}}; \tilde{\boldsymbol{\theta}}) \right\} \\ &= \arg \min_{\tilde{\boldsymbol{\theta}}} \underbrace{\left\{ \sum_{l=0}^L \left| \tilde{h}_l - h_l(\tilde{\boldsymbol{\theta}}) \right|^2 \right\}}_{\Omega(\tilde{\boldsymbol{\theta}})}. \end{aligned} \quad (11)$$

The minimization of the metric $\Omega(\tilde{\boldsymbol{\theta}})$ in (11) cannot be solved in closed form since $\Omega(\tilde{\boldsymbol{\theta}})$ is nonlinear. An optimization method has to be applied. In order to choose a suitable optimization method to find $\hat{\boldsymbol{\theta}}$, different system aspects have to be taken into account, and a tradeoff depending on the requirements has to be found. The goal is to find the *global* minimum of $\Omega(\tilde{\boldsymbol{\theta}})$.

Unfortunately, $\Omega(\tilde{\boldsymbol{\theta}})$ has many *local* minima due to the superposition of random multipath components. Consequently, the optimization method of choice should be either a global optimization method or a local optimization method in combination with a good initial guess, i. e., an initial guess that is sufficiently close to the global optimum. Both choices involve different benefits and drawbacks. To find a good initial guess is difficult and, therefore, may be seen as a drawback itself. But in case a priori knowledge in form of a good initial guess is available, a search in the complete search space would be unnecessary.

For channel parameter estimation, it is suggested to divide the problem into an acquisition and a tracking phase. In the acquisition phase, a global optimization method is applied, and in the tracking phase, the parameter estimate of the last data burst may be used as an initial guess for a local optimization method. This is suitable for channels that do not change too rapidly from data burst to data burst. In this paper, particle swarm optimization (PSO) [15-17] is suggested for the acquisition phase, and the Levenberg-Marquardt method (LMM) [18,19] is proposed for the tracking phase.

PSO is a heuristic optimization method that is able to find the global optimum without an initial guess and without gradient information. PSO is easy to implement because only function evaluations have to be performed. So-called particles move randomly through the search

space and are attracted by good fitness values $\Omega(\tilde{\theta})$ in their past and of their neighbors. In this way, the particles explore the search space and are able to find the global optimum. It is a drawback that PSO does not assure global convergence. There is a certain probability (depending on the signal-to-noise ratio) that PSO converges prematurely to a local optimum (outage). Furthermore, PSO is sometimes criticized because many iterations are performed in comparison to gradient-based optimization algorithms.

The LMM belongs to the standard nonlinear least-squares solvers and relies on a good initial guess. The gradient of the metric has to be supplied by the user. For the LMM, convergence to the optimum in the neighborhood of the initial guess is assured. Second derivative information is used to speed up convergence: The LMM varies smoothly between the inverse-Hessian method and the steepest decent method depending on the topology of the metric [18]. Furthermore, an approximation to the covariance matrix of the parameter estimates is calculated inherently by the LMM. The LMM is designed for small residual problems. For large residual problems (at low signal-to-noise ratio), it may fail (outage).

3 Soft Information

3.1 Definition of soft information

The concept of soft information is already widely applied: In the area of communication, soft information is used for decoding, detection, and equalization. In the field of navigation, soft information is exploited for sensor fusion [20]. This paper aims at obtaining soft information for the parameter estimates in order to improve the positioning accuracy *before* sensor fusion is applied.

Soft information is a measure of reliability of the (hard) estimates. The intention is to determine the a posteriori distribution of the estimates. Hence, the (hard) estimate is the mean of the distribution, and the soft information corresponds to the variance of the distribution. For linear estimation problems with known noise covariance matrix, the a posteriori distribution of the estimates can be determined in closed form [14]. If the noise is Gaussian distributed, the estimator is, furthermore, a minimum variance unbiased estimator (MVU). However, only few problems are linear. A popular estimator for more general problems is the maximum-likelihood estimator as already described in Section 2.2 for channel parameter estimation. The maximum-likelihood estimator is asymptotically (for a large number of observations or at a high signal-to-noise ratio) unbiased and efficient [14]. Furthermore, an asymptotic a posteriori distribution can be determined. For Gaussian noise with covariance matrix $C = \sigma^2 I$, the

asymptotic covariance matrix of the estimates is given by the inverse of the Fisher information matrix evaluated at the true parameters [14]. The parameter estimate $\hat{\theta}$ given by (11) is asymptotically distributed as follows:

$$\hat{\theta} \sim \mathcal{N}(\theta, I(\theta)^{-1}), \quad (12)$$

where $I(\theta)$ is the Fisher information matrix with entries

$$\begin{aligned} [I(\theta)]_{mn} &= -E \left\{ \frac{\delta^2 \ln p(\tilde{h}; \theta)}{\delta \theta_m \delta \theta_n} \right\} \\ &= \frac{2}{\sigma^2} \text{Re} \left\{ \sum_{l=0}^L \frac{\delta h_l^*(\theta)}{\delta \theta_m} \frac{\delta h_l(\theta)}{\delta \theta_n} \right\}, \end{aligned} \quad (13)$$

in which the star \star denotes the conjugate complex. Given the Jacobian matrix of (8),

$$[J(\theta)]_{lm} = \frac{\delta h_l(\theta)}{\delta \theta_m}, \quad (14)$$

the Fisher information matrix can be written as well as

$$I(\theta) = \frac{2}{\sigma^2} \text{Re} \left\{ J(\theta)^H J(\theta) \right\}. \quad (15)$$

The variance of parameter θ_m is given by the m th diagonal entry of the asymptotic covariance matrix:

$$C_{\text{asympt}} = I(\theta)^{-1} = \frac{\sigma^2}{2} \left(\text{Re} \left\{ J(\theta)^H J(\theta) \right\} \right)^{-1}. \quad (16)$$

In general, the true value of the parameters is not known. Therefore, the asymptotic covariance matrix cannot be determined and an approximation has to be found. Different approximate covariance matrices are given in the literature that should be used with caution since the approximation may be very poor [11,21]. In the following section, a short description of confidence regions is included because they are closely related to soft information: Some of the confidence regions rely on the approximate covariance matrices mentioned above.

3.2 Confidence regions

In [11], Donaldson and Schnabel investigate different methods to construct confidence regions and confidence intervals. Confidence regions and intervals are closely related to soft information since they also indicate reliability: The estimated parameters $\hat{\theta}$ do not coincide with the true parameters θ because of the measurement noise. A confidence region indicates the area around the estimated parameters in which the true parameters might be with a specific probability. This probability is

called the confidence level and is often expressed as a percentage. A commonly used confidence level is 95%.

For linear problems with Gaussian noise, the confidence regions are elliptical and can be determined exactly by the covariance matrix C_{linear} , which can be computed in closed form [14]. The linear confidence region consists of all parameter vectors $\tilde{\theta}$ that satisfy the following formula:

$$(\tilde{\theta} - \hat{\theta}) C_{\text{linear}}^{-1} (\tilde{\theta} - \hat{\theta}) \leq P \mathcal{F}_{P, N-P}^{1-\alpha} \quad (17)$$

in which $P = 3M$ is the number of parameters, $N = L + 1$ is the number of observations, $1 - \alpha$ is the confidence level, and \mathcal{F} is the Fisher distribution. According to [11], the most common method to determine a confidence region for a nonlinear problem consists of the linearization of the problem in order to obtain an approximate covariance matrix. In this paper, the following approximate covariance matrix is applied:^c

$$C_{\text{approx}} = \frac{s^2}{2} \left(\text{Re} \left\{ J(\hat{\theta})^H J(\hat{\theta}) \right\} \right)^{-1} \quad (18)$$

The only difference between C_{approx} in (18) and C_{asympt} in (16) is that the Jacobian matrix is evaluated at the parameter estimate $\hat{\theta}$ instead of the true parameter θ and that the variance σ^2 is estimated by the residual variance $s^2 = \Omega(\hat{\theta}) / (N - P)$. When C_{linear} in (17) is replaced by C_{approx} in (18), an approximate confidence region for a nonlinear problem is obtained as

$$(\tilde{\theta} - \hat{\theta}) 2 \text{Re} \left\{ J(\hat{\theta})^H J(\hat{\theta}) \right\} (\tilde{\theta} - \hat{\theta}) \leq s^2 P \mathcal{F}_{P, N-P}^{1-\alpha} \quad (19)$$

On the one hand, the computational complexity is quite low and the results are very similar to the well-known linear case. On the other hand, the approximation can be very poor and should be used with caution [11,21]. Another (more complex) way to determine a confidence region is the likelihood method [11]: All parameter vectors $\tilde{\theta}$ that satisfy

$$\Omega(\tilde{\theta}) - \Omega(\hat{\theta}) \leq s^2 P \mathcal{F}_{P, N-P}^{1-\alpha} \quad (20)$$

are included in the likelihood confidence region. This region does not have to be elliptical but can be of any form. The likelihood method is approximate for nonlinear problems as well but more precise and robust than the linearization method since it does not rely on linearization. There is an exact method, which is called lack-of-fit method, that is neglected in this paper due to its high computational complexity and because the likelihood method is already a good approximation according to [11]. The accuracy of the linearization and the

likelihood method strongly depends on the problem and on the parameters. Donaldson and Schnabel [11] suggest to use the curvature measures of Bates and Watts [13], which are introduced in Section 4, as a diagnostic tool. With these measures, it can be evaluated whether the corresponding method is applicable or not.

3.3 Proposed methods to obtain soft information

After this excursion to confidence regions, the way of employing this knowledge for obtaining soft information is now discussed. The first and straightforward idea is to use the variances of the approximate covariance matrix C_{approx} in (18). This method is simple, and many optimization algorithms like the LMM already compute and output C_{approx} or similar versions of it. But without further analysis (see Sections 4 and 5), it is questionable whether this method is precise enough.

The second idea is based on the likelihood confidence regions. Generally, it is quite complex to generate the likelihood confidence region since many function evaluations have to be performed in the surrounding of the parameter estimates $\hat{\theta}$. However, heuristic optimization algorithms like PSO perform many function evaluations in the whole search space anyway, and therefore, they are well suited to determine the likelihood confidence region [12]. A drawback of heuristic algorithms (many function evaluations are required until convergence) is transformed into an advantage with respect to likelihood confidence regions. The procedure proposed in [12] is as follows: In every iteration, each particle determines its fitness $\Omega(\tilde{\theta})$, which is stored with the corresponding parameter set $\tilde{\theta}$ in a table. After the optimum $\hat{\theta}$ with fitness $\Omega(\hat{\theta})$ is found, all parameter sets $\tilde{\theta}$ that fulfill

$$\Omega(\tilde{\theta}) \leq \Omega(\hat{\theta}) \left(1 + \frac{P}{N-P} \mathcal{F}_{P, N-P}^{1-\alpha} \right) \quad (21)$$

are selected from the table and form the likelihood confidence region. It can be observed that the density of points near the parameter estimate $\hat{\theta}$ is higher than at the border of the likelihood confidence region. The reason is that the particles are attracted by good fitness values near the optimum and oscillate in its neighborhood before convergence occurs. Hence, all points $\tilde{\theta}$ form a distribution with mean and variance, where the mean coincides with the parameter estimate $\hat{\theta}$. Therefore, the variance of this distribution can be used as soft information.

In Section 5, the performance of both methods is evaluated and compared. Prior to that the curvature measures of Bates and Watts [13] are introduced for further analysis and understanding.

4 Curvature Measures

4.1 Introduction to curvature measures

In [13], Bates and Watts describe nonlinear least-squares estimation from a geometric point of view and introduce measures of nonlinearity. These measures indicate the applicability of a linearization and its effects on inference. Hence, the accuracy of the confidence regions described in Section 3 can be evaluated using these measures. In the following, the most important aspects of the so-called curvature measures are presented.

First, the nonlinear least-squares problem is reviewed: A set of parameters

$$\boldsymbol{\theta} = [\theta_1, \theta_2, \dots, \theta_P]^T \quad (22)$$

shall be estimated from a set of observations

$$\check{\mathbf{h}} = [\check{h}_0, \check{h}_1, \dots, \check{h}_L]^T \quad (23)$$

with

$$\check{h}_l = h_l(\boldsymbol{\theta}) + \varepsilon_l, \quad (24)$$

where $h_l(\boldsymbol{\theta})$ is a nonlinear function of the parameters $\boldsymbol{\theta}$ and ε_l is additive zero mean measurement noise with variance σ_ε^2 . The least-squares estimate is given by the value $\hat{\boldsymbol{\theta}}$ that minimizes the sum of squares of residuals

$$\Omega(\tilde{\boldsymbol{\theta}}) = \sum_{l=0}^L |\check{h}_l - h_l(\tilde{\boldsymbol{\theta}})|^2, \quad (25)$$

which corresponds to the metric of the maximum-likelihood estimator in the case of Gaussian measurement noise. The sum of squares in (25) can also be written as

$$\Omega(\tilde{\boldsymbol{\theta}}) = \left\| \check{\mathbf{h}} - \mathbf{h}(\tilde{\boldsymbol{\theta}}) \right\|^2. \quad (26)$$

Geometrically, (26) describes the distance between $\check{\mathbf{h}}$ and $\mathbf{h}(\tilde{\boldsymbol{\theta}})$ in the $(L + 1)$ -dimensional *sample space*. If the parameter vector $\tilde{\boldsymbol{\theta}}$ is changed in the P -dimensional *parameter space* (search space), the vector $\mathbf{h}(\tilde{\boldsymbol{\theta}})$ traces a P -dimensional surface in the sample space, which is called *solution locus*. Hence, the function $\mathbf{h}(\tilde{\boldsymbol{\theta}})$ maps all feasible parameters in the P -dimensional parameter space to the P -dimensional solution locus in the $(L+1)$ -dimensional sample space. Because of the measurement noise, the observations do not lie on the solution locus but anywhere in the sample space. The parameter estimate $\hat{\boldsymbol{\theta}}$ corresponds to the point on the solution locus $\mathbf{h}(\hat{\boldsymbol{\theta}})$ with the smallest distance to the point of observations $\check{\mathbf{h}}$.

Since the function $\mathbf{h}(\tilde{\boldsymbol{\theta}})$ is nonlinear, the solution locus will be a curved surface. For inference, the

solution locus is approximated by a tangent plane with an uniform coordinate system. The tangent plane at a specific point $\mathbf{h}(\tilde{\boldsymbol{\theta}}_0)$ can be described by a first-order Taylor series

$$\mathbf{h}(\tilde{\boldsymbol{\theta}}) \cong \mathbf{h}(\tilde{\boldsymbol{\theta}}_0) + \mathbf{J}(\tilde{\boldsymbol{\theta}}_0) (\tilde{\boldsymbol{\theta}} - \tilde{\boldsymbol{\theta}}_0), \quad (27)$$

where $\mathbf{J}(\tilde{\boldsymbol{\theta}}_0)$ is the Jacobian matrix as defined in (14) evaluated at $\tilde{\boldsymbol{\theta}}_0$. The informational value of inference concerning the parameter estimates highly depends on the closeness of the tangent plane to the solution locus. This closeness in turn depends on the curvature of the solution locus. Therefore, the measures of nonlinearity proposed by Bates and Watts indicate the maximum curvature of the solution locus at the specific point $\mathbf{h}(\tilde{\boldsymbol{\theta}}_0)$. It is important to note that there are two different kinds of curvatures since two different assumptions are made concerning the tangent plane. First, it is assumed that the solution locus is planar at $\mathbf{h}(\tilde{\boldsymbol{\theta}}_0)$ and, hence, can be replaced by the tangent plane (*planar assumption*). Second, it is assumed that the coordinate system on the tangent plane is uniform (*uniform coordinate assumption*), i.e., the coordinate grid lines mapped from the parameter space remain equidistant and straight in the sample space. It might happen that the first assumption is fulfilled, but the second assumption is not. Then, the solution locus is planar at the specific point $\mathbf{h}(\tilde{\boldsymbol{\theta}}_0)$, but the coordinate grid lines are curved and not equidistant. If the planar assumption is not fulfilled, the uniform coordinate assumption is not fulfilled either.

In order to determine the curvatures, Bates and Watts introduce so-called *lifted lines*. Similar to the fact that each point $\tilde{\boldsymbol{\theta}}_0$ in the parameter space maps to a point $\mathbf{h}(\tilde{\boldsymbol{\theta}}_0)$ on the solution locus in the sample space, each straight line in the parameter space through $\tilde{\boldsymbol{\theta}}_0$,

$$\tilde{\boldsymbol{\theta}}(m) = \tilde{\boldsymbol{\theta}}_0 + m\mathbf{v}, \quad (28)$$

maps to a lifted line on the solution locus

$$\mathbf{h}_v(m) = \mathbf{h}(\tilde{\boldsymbol{\theta}}_0 + m\mathbf{v}), \quad (29)$$

where \mathbf{v} can be any non-zero vector in the parameter space. The tangent vector of the lifted line for $m = 0$ at $\tilde{\boldsymbol{\theta}}_0$ is given by

$$\dot{\mathbf{h}}_v = \left. \frac{d\mathbf{h}_v(m)}{dm} \right|_0 = \left. \frac{d\mathbf{h}(\tilde{\boldsymbol{\theta}})}{d\tilde{\boldsymbol{\theta}}} \right|_{\tilde{\boldsymbol{\theta}}_0} \left. \frac{d\tilde{\boldsymbol{\theta}}(m)}{dm} \right|_0 = \mathbf{J}(\tilde{\boldsymbol{\theta}}_0)\mathbf{v}. \quad (30)$$

The set of all tangent vectors (for all possible vectors \mathbf{v}) forms the tangent plane. For measuring curvatures,

second-order derivatives are needed additionally. The second-order derivative of the function $\mathbf{h}(\tilde{\boldsymbol{\theta}})$ is the Hessian

$$\left[\mathbf{H}(\tilde{\boldsymbol{\theta}}) \right]_{lij} = \frac{\delta^2 h_l(\tilde{\boldsymbol{\theta}})}{\delta \tilde{\boldsymbol{\theta}}_i \delta \tilde{\boldsymbol{\theta}}_j}, \quad (31)$$

which is a three-dimensional tensor. The l th face of the Hessian is, thus, a $P \times P$ matrix

$$\mathbf{H}_l(\tilde{\boldsymbol{\theta}}) = \begin{bmatrix} \frac{\delta^2 h_l(\tilde{\boldsymbol{\theta}})}{\delta \tilde{\boldsymbol{\theta}}_1 \delta \tilde{\boldsymbol{\theta}}_1} & \dots & \frac{\delta^2 h_l(\tilde{\boldsymbol{\theta}})}{\delta \tilde{\boldsymbol{\theta}}_p \delta \tilde{\boldsymbol{\theta}}_1} \\ \vdots & \ddots & \vdots \\ \frac{\delta^2 h_l(\tilde{\boldsymbol{\theta}})}{\delta \tilde{\boldsymbol{\theta}}_1 \delta \tilde{\boldsymbol{\theta}}_p} & \dots & \frac{\delta^2 h_l(\tilde{\boldsymbol{\theta}})}{\delta \tilde{\boldsymbol{\theta}}_p \delta \tilde{\boldsymbol{\theta}}_p} \end{bmatrix}. \quad (32)$$

The second-order derivative of the lifted line is given by

$$\ddot{\mathbf{h}}_v = \left. \frac{d^2 \mathbf{h}_v(m)}{dm^2} \right|_0 = \mathbf{v}^T \mathbf{H}(\tilde{\boldsymbol{\theta}}_0) \mathbf{v}, \quad (33)$$

in which the tensor product is performed such that

$$\left[\ddot{\mathbf{h}}_v \right]_l = \mathbf{v}^T \mathbf{H}_l(\tilde{\boldsymbol{\theta}}_0) \mathbf{v}. \quad (34)$$

The derivatives of the lifted line $\dot{\mathbf{h}}_v$ and $\ddot{\mathbf{h}}_v$ can be interpreted physically: If a point moves along the lifted line $\mathbf{h}_v(m)$ in the sample space, where m denotes the time, then $\dot{\mathbf{h}}_v$ and $\ddot{\mathbf{h}}_v$ denote the instantaneous velocity and instantaneous acceleration at time $m = 0$, respectively. The acceleration can be decomposed in three parts

$$\ddot{\mathbf{h}}_v = \ddot{\mathbf{h}}_v^P + \ddot{\mathbf{h}}_v^N + \ddot{\mathbf{h}}_v^G \quad (35)$$

as shown in Figure 3. $\ddot{\mathbf{h}}_v^P$ is parallel to the velocity vector $\dot{\mathbf{h}}_v$ and, thus, parallel to the tangent plane. It corresponds to the *change in velocity* of the moving point. $\ddot{\mathbf{h}}_v^N$ is normal to the tangent plane and describes the *change in direction* of the velocity vector $\dot{\mathbf{h}}_v$ normal to the tangent plane. $\ddot{\mathbf{h}}_v^G$ is parallel to the tangent plane and normal to the velocity vector $\dot{\mathbf{h}}_v$. It corresponds to the geodesic acceleration and indicates the *change in direction* of the velocity vector $\dot{\mathbf{h}}_v$ parallel to the tangent plane. Based on these acceleration components, the curvatures of the solution locus at $\tilde{\boldsymbol{\theta}}_0$ can be determined:

$$K_v^N = \frac{\|\ddot{\mathbf{h}}_v^N\|}{\|\dot{\mathbf{h}}_v\|^2} \quad (36)$$

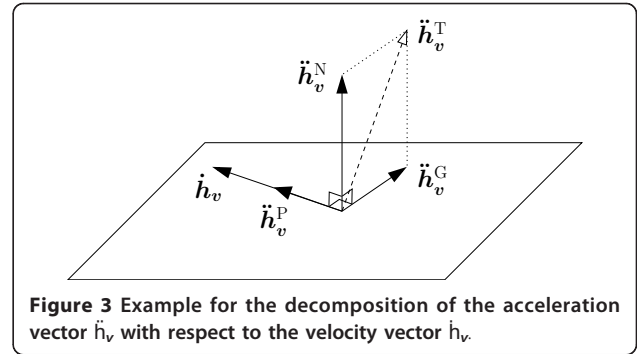


Figure 3 Example for the decomposition of the acceleration vector $\ddot{\mathbf{h}}_v$ with respect to the velocity vector $\dot{\mathbf{h}}_v$.

is the normal curvature in direction of \mathbf{v} and is called *intrinsic curvature* and

$$K_v^T = \frac{\|\ddot{\mathbf{h}}_v^P + \ddot{\mathbf{h}}_v^G\|}{\|\dot{\mathbf{h}}_v\|^2} = \frac{\|\ddot{\mathbf{h}}_v^T\|}{\|\dot{\mathbf{h}}_v\|^2} \quad (37)$$

is the tangential^d curvature in direction of \mathbf{v} and is called *parameter-effects curvature*. The curvatures are divided into normal and tangential components since each component has a different influence on the accuracy of the linear approximation. On the one hand, the intrinsic curvature is an intrinsic property of the solution locus. It only affects the planar assumption. On the other hand, the parameter-effects curvature only influences the uniform coordinate assumption and depends on the specific parameterization of the problem. Hence, a reparameterization may change the parameter-effects curvature but not the intrinsic curvature. In order to assess the effect of the curvatures on inference, they should be normalized. A suitable scaling factor is the so-called *standard radius* $\rho = s\sqrt{P}$ since its square $\rho^2 = s^2P$ appears on the right hand side in (19) and (20), which describe the confidence regions. The *relative curvatures* are given by the curvatures (36) and (37) multiplied with the standard radius:

$$\gamma_v^N = K_v^N \rho, \quad (38)$$

$$\gamma_v^T = K_v^T \rho. \quad (39)$$

If the relative curvatures are small compared with $1/\sqrt{\mathcal{F}_{P,N-P}^{1-\alpha}}$ for all possible directions \mathbf{v} , then the corresponding assumptions are valid. Hence, it is sufficient to determine the *maximum relative curvatures*^e

$$\Gamma^N = \max_v \{\gamma_v^N\}, \quad (40)$$

$$\Gamma^T = \max_v \{\gamma_v^T\} \quad (41)$$

and to compare them to $1/\sqrt{\mathcal{F}_{P,N-P}^{1-\alpha}}$ in order to assess the accuracy of the confidence regions [11]. If the confidence region based on the linearization method (19) with the approximate covariance matrix shall be applied, both the planar assumption and the uniform coordinate assumption have to be fulfilled. That means that the maximum relative curvatures Γ^N and Γ^T have to be small compared with $1/\sqrt{\mathcal{F}_{P,N-P}^{1-\alpha}}$. The confidence region based on the likelihood method (20) is more robust since only the planar assumption needs to be fulfilled and only Γ^N needs to be small compared with $1/\sqrt{\mathcal{F}_{P,N-P}^{1-\alpha}}$.

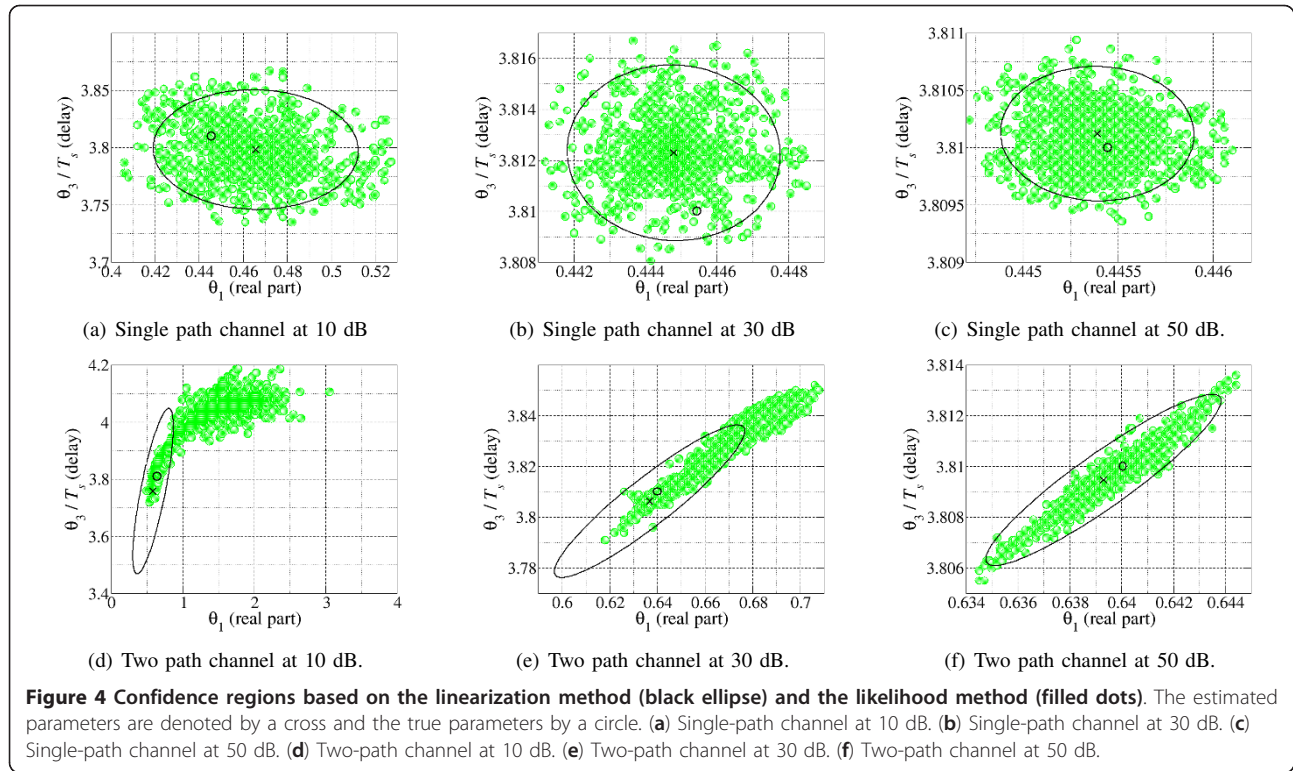
4.2 Analysis of the parameter estimation problem

In the following, the parameter estimation problem is analyzed by calculating the maximum relative curvatures and by plotting the confidence regions (19) and (20) for different signal-to-noise ratios (SNRs). The system setup is as follows: A training preamble of length $K_t = 256$ is assumed that covers 10% of the data burst of length $K = 2,560$. A pseudo-random sequence of BPSK symbols is used as training. Since this paper concentrates on the positioning part of the proposed joint communication and positioning system, it is sufficient to focus on the channel estimation and to neglect the data detection. A Gaussian pulse shape $g(\tau) = gT_x(\tau) * gR_x(\tau) \sim \exp(-(\tau/T_s)^2)$ is assumed. After receive filtering, the noise process is slightly colored, but we have verified that the correlation is negligible with respect to receiver processing. The training sequence is transmitted over the physical channel and at the receiver side channel parameter estimation as suggested in Section 2.2 is performed. For the purpose of curvature analysis, only PSO as described in [16] with $I = 50$ particles and a maximum number of $T = 8,000$ iterations is applied for solving the nonlinear metric $\Omega(\theta)$. PSO delivers the likelihood confidence region automatically as explained in Section 3.3. The approximate covariance matrix is calculated afterward according to (18). A confidence level of 95% is applied ($\alpha = 0.05$). Since the curvature measures depend on the parameter set θ and also on the noise samples, simulations are performed for a fixed channel model at different SNRs. Two different channel models are assumed: A single-path channel ($M = 1$) and a two-path channel ($M = 2$) with a small excess delay ($\Delta\tau_2 = \tau_2 - \tau_1 = 0.81T_s$), both with a memory length $L = 10$. The parameters of the channels are given in Table 1. Furthermore, the maximum relative curvatures Γ^N and Γ^T for different SNRs and the value of $1/\sqrt{\mathcal{F}_{P,N-P}^{0.95}}$ are listed in Table 1. It can be concluded that the planar assumption is always

Table 1 Parameters of the investigated channel models and the corresponding maximum rel. curvatures at different SNRs

	M = 1	M = 2	
real part	$\theta_1 = 0.4454$	$\theta_1 = 0.6401$	$\theta_4 = -0.3464$
imaginary part	$\theta_2 = -0.7715$	$\theta_2 = -1.1086$	$\theta_5 = 0.8363$
delay	$\theta_3 = 3.81 T_s$	$\theta_3 = 3.81 T_s$	$\theta_6 = 4.62 T_s$
$1/\sqrt{\mathcal{F}_{P,N-P}^{0.95}}$	0.49591	0.44945	
Γ^N @ 10 dB	0.05429737	0.08281260	
Γ^T @ 10 dB	0.04205545	3.62952012	
Γ^N @ 30 dB	0.00354615	0.01158592	
Γ^T @ 30 dB	0.00272911	0.75507012	
Γ^N @ 50 dB	0.00062543	0.00134538	
Γ^T @ 50 dB	0.00047295	0.08727709	

fulfilled since Γ^N is much smaller than $1/\sqrt{\mathcal{F}_{P,N-P}^{0.95}}$ in all cases. This means the likelihood method is always accurate. For the single-path channel, the uniform coordinate assumption is also fulfilled for all SNRs (see Table 1), i.e., the confidence regions based on the linearization method and the approximate covariance matrices are accurate. This is confirmed by Figure 4a, b, c. In Figure 4, the confidence regions based on the linearization method (black ellipse) and the likelihood method (filled dots) are plotted for the parameter combination of the real part θ_1 and the delay θ_3 of the LOS path normalized with respect to the symbol duration T_s . Both regions are similar for the single-path channel. In case of the two-path channel, a different situation is observed as shown in Figure 4d, e, f. The uniform coordinate assumption is violated at low SNR since Γ^T is *not* much smaller than $1/\sqrt{\mathcal{F}_{P,N-P}^{0.95}}$ (see Table 1). The shape of the likelihood confidence region differs strongly from the ellipse generated by the approximate covariance matrix. Only at high SNR, both shapes coincide. For the two-path channel, the uniform coordinate assumption is valid from approximately 35-40 dB upward. For different channel realizations, different results are obtained. It should be mentioned again that the curvature measures strongly depend on the parameter set θ and on the noise samples. The larger the excess delay $\Delta\tau_2$, the lower is the nonlinearity of the problem, i.e., the uniform coordinate assumption is already valid at lower SNR and vice versa. It can be summarized that the confidence regions based on the linearization method are not accurate at low SNR in a multipath scenario. Hence, the soft information based on the approximate covariance matrix may lead to inaccurate results. The influence of soft information on positioning is investigated in the following section.



5 Positioning

5.1 Positioning based on the time of arrival

There are many different approaches to determine the position, e.g., multiangulation, multilateration, fingerprinting, and motion sensors. This paper focusses on radiolocation based on the TOA, which is also called multilateration. Furthermore, two-dimensional positioning is considered in the following. An extension to three dimensions is straightforward.

The position $\mathbf{p} = [x, y]^T$ of a mobile station (MS) is determined relative to B reference objects (ROs) whose positions $\mathbf{p}_b = [x_b, y_b]^T$ ($1 \leq b \leq B$) are known. For each RO b , the TOA $\hat{\tau}_{1,b}$ is estimated. The TOA corresponds to the distance between this RO and the MS $r_b = \hat{\tau}_{1,b}c$, where c is the speed of light. The estimated distances $\mathbf{r} = [r_1, \dots, r_B]^T$ are called *pseudo-ranges* since they consist of the true distances $\mathbf{d}(\mathbf{p}) = [d_1(\mathbf{p}), \dots, d_B(\mathbf{p})]^T$ and estimation errors $\boldsymbol{\eta} = [\eta_1, \dots, \eta_B]^T$ with covariance matrix $\mathbf{r} = \mathbf{d}(\mathbf{p}) + \boldsymbol{\eta}$:

$$\mathbf{r} = \mathbf{d}(\mathbf{p}) + \boldsymbol{\eta}. \quad (42)$$

The true distance between the b th RO and the MS is a nonlinear function of the position \mathbf{p} given by

$$d_b(\mathbf{p}) = \sqrt{(x - x_b)^2 + (y - y_b)^2}. \quad (43)$$

Thus, positioning is again a nonlinear problem.^f There are alternative ways to solve the set of nonlinear equations described by (42) and (43). In this paper, two different approaches are considered: The iterative Taylor series algorithm (TSA) [22] and the weighted least-squares (WLS) method [23,24].

The TSA is based on a linearization of the nonlinear function (43). Given a starting position $\hat{\mathbf{p}}_0$ (initial guess), the pseudo-ranges can be approximated by a first-order Taylor series

$$\mathbf{r} \cong \mathbf{d}(\hat{\mathbf{p}}_0) + \mathbf{J}(\hat{\mathbf{p}}_0)(\mathbf{p} - \hat{\mathbf{p}}_0) + \boldsymbol{\eta}, \quad (44)$$

in which $\mathbf{J}(\mathbf{p})$ is the Jacobian matrix of (43) with entries

$$[\mathbf{J}(\mathbf{p})]_{b1} = \frac{\delta d_b(\mathbf{p})}{\delta x}, \quad [\mathbf{J}(\mathbf{p})]_{b2} = \frac{\delta d_b(\mathbf{p})}{\delta y}. \quad (45)$$

Defining $\Delta \mathbf{r}_0 = \mathbf{r} - \mathbf{d}(\hat{\mathbf{p}}_0)$ and $\Delta \mathbf{p}_0 = \mathbf{p} - \hat{\mathbf{p}}_0$ results in the following linear relationship

$$\Delta \mathbf{r}_0 \cong \mathbf{J}(\hat{\mathbf{p}}_0)\Delta \mathbf{p}_0 + \boldsymbol{\eta}, \quad (46)$$

that can be solved according to the least-squares approach:

$$\Delta \hat{\mathbf{p}}_0 = \left(\mathbf{J}(\hat{\mathbf{p}}_0)^T \mathbf{W} \mathbf{J}(\hat{\mathbf{p}}_0) \right)^{-1} \mathbf{J}(\hat{\mathbf{p}}_0)^T \mathbf{W} \Delta \mathbf{r}_0. \quad (47)$$

The weighting matrix \mathbf{W} is given by the inverse of the covariance matrix $\mathbf{C}_\eta : \mathbf{W} = \text{diag}\left(\frac{1}{\sigma_{\eta 1}^2}, \dots, \frac{1}{\sigma_{\eta B}^2}\right)$. A new position estimate $\hat{\mathbf{p}}_1$ is obtained by adding the correction factor $\Delta\hat{\mathbf{p}}_0$ to the starting position $\hat{\mathbf{p}}_0$. This procedure is performed iteratively,

$$\hat{\mathbf{p}}_{i+1} = \hat{\mathbf{p}}_i + \Delta\hat{\mathbf{p}}_i, \quad (48)$$

until the correction factor $\Delta\hat{\mathbf{p}}_i$ is smaller than a given threshold. If the initial guess is close to the true position, few iterations are needed. If the starting position is far from the true position, many iterations may be necessary. Additionally, the algorithm may diverge. Hence, finding a good initial guess is a crucial issue. For the numerical results shown in Section 5.2, the position estimate of the WLS method is used as initial guess for the TSA.

The WLS method [23,24] solves the set of nonlinear equations described by (42) and (43) in closed form. Hence, this method is non-iterative and less costly than the TSA. The basic idea is to transform the original set of *nonlinear* equations into a set of *linear* equations. For this purpose, one RO is selected as reference. Without loss of generality, the first RO is chosen here. By subtracting the squared distance of the first RO from the squared distances of the remaining ROs, a linear least-squares problem with solution

$$\hat{\mathbf{p}} = (\mathbf{S}^T \mathbf{W}' \mathbf{S})^{-1} \mathbf{S}^T \mathbf{W}' \mathbf{b} \quad (49)$$

is obtained, in which

$$\mathbf{S} = \begin{bmatrix} x_2 - x_1 & y_2 - y_1 \\ x_3 - x_1 & y_3 - y_1 \\ \vdots & \vdots \\ x_B - x_1 & y_B - y_1 \end{bmatrix} \quad (50)$$

and

$$\mathbf{b} = -\frac{1}{2} \begin{bmatrix} r_2^2 - r_1^2 - R_2^2 + R_1^2 \\ r_3^2 - r_1^2 - R_3^2 + R_1^2 \\ \vdots \\ r_B^2 - r_1^2 - R_B^2 + R_1^2 \end{bmatrix} \quad (51)$$

with $R_b^2 = x_b^2 + y_b^2$. The weighting matrix \mathbf{W}' is given by:

$$\mathbf{W}' = \text{diag}\left(\frac{1}{\sigma_{\eta 2}^4}, \dots, \frac{1}{\sigma_{\eta B}^4}\right).$$

Both, the TSA and the WLS method, apply a weighting matrix that contains the variances of the pseudo-range errors. Reliable pseudo-ranges have higher weights than unreliable ones and, thus, have a stronger influence

on the estimation results. Typically, the true variances are not known. They can only be estimated as described in Section 3: For each link b , the variance of the TOA $\sigma_{\tau_{1,b}}^2$ is determined via the linearization⁸ or the likelihood method. This TOA variance is transformed into a pseudo-range variance $\sigma_{\eta b}^2$ by a multiplication with c^2 . If no information about the estimation error $\boldsymbol{\eta}$ is available, the weighting matrices correspond to the identity matrix \mathbf{I} (no weighting at all).

The Cramer-Rao lower bound (CRLB) provides a benchmark to assess the performance of the estimators [14]:

$$\text{CRLB}(\mathbf{p}) = \sum_{d=1}^2 [\mathbf{I}(\mathbf{p})^{-1}]_{dd}, \quad (52)$$

where

$$\mathbf{I}(\mathbf{p}) = \mathbf{J}(\mathbf{p})^T \mathbf{W} \mathbf{J}(\mathbf{p}) \quad (53)$$

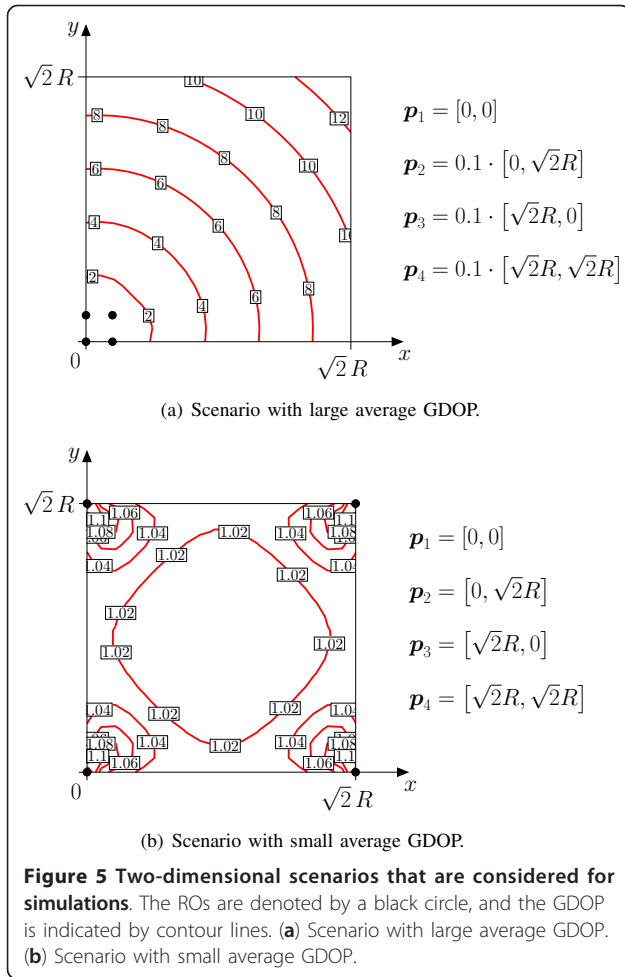
is the Fisher information matrix. If the estimator is unbiased, its mean squared error (MSE) is larger than or equal to the CRLB. If the MSE approaches the CRLB, the estimator is a minimum variance unbiased (MVU) estimator.

The positioning accuracy depends on the geometry between the ROs and the MS and, thus, varies with the position \mathbf{p} . This effect is called geometric dilution of precision (GDOP) [22,25]. In order to separate the influence of the geometry from the influence of the estimation errors $\boldsymbol{\eta}$ on the positioning accuracy, it is assumed that all pseudo-ranges are affected by the same error variance $\sigma_\eta^2 = 1$, i.e., $\mathbf{W} = \mathbf{I}$. Given this assumption, the GDOP is the square root of the CRLB:

$$\text{GDOP}(\mathbf{p}) = \sqrt{\text{CRLB}(\mathbf{p})|_{\mathbf{W}=\mathbf{I}}}. \quad (54)$$

5.2 Numerical results

In the following, the overall performance of the proposed system concept using soft information is evaluated. For this purpose, two scenarios with different GDOP as shown in Figure 5 are considered. The ROs are denoted by black circles and the GDOP is illustrated by contour lines. For both scenarios, $B = 4$ ROs are located inside a quadratic region with side length $\sqrt{2}R$, where $R = 2T_s c$ is the distance from every RO to the middle point of the region. For the first scenario, the ROs are placed in the lower left part of the region, which results in a large GDOP on average. The second scenario has a small GDOP on average since the ROs are placed in the corners of the region. For the communication links between the MS and the ROs, the same setup as described in Section 4.2 is applied.



Furthermore, power control is assumed, i.e., the SNR for all links is the same. All results reported throughout this paper are for one-shot measurements.

Three different channel models with memory length $L = 10$ are investigated: a single-path channel ($M = 1$), a two-path channel ($M = 2$) with large excess delay ($\Delta\tau_2 \in [T_s, 2T_s]$) and a two-path channel ($M = 2$) with small excess delay ($\Delta\tau_2 \in [\frac{T_s}{10}, T_s]$). For all channel models, the LOS delay $\tau_{1,b}$ for each link b is calculated from the true distance $d_b(\mathbf{p})$. The excess delay of the multipath component $\Delta\tau_2$ for both two-path channels is determined randomly in the corresponding interval. The smaller the excess delay is, the more difficult it is to separate the different propagation paths. The power of the multipath component is half the power of the LOS component. The phase of each component is generated randomly between 0 and 2π . For each link, channel parameter estimation is performed and soft information based on the linearization method and on the likelihood method is obtained. For PSO, $I = 50$ particles and a maximum number of iterations $T = 8,000$ are applied.^h

The estimated LOS delays $\hat{\tau}_{1,b}$ are converted to pseudo-ranges r_b , and the position of the MS is estimated with the TSA and the WLS method applying the different soft information methods. For comparison, positioning without soft information is performed. The position estimate of the WLS method is used as initial guess for the TSA. Furthermore, in the WLS method, the RO with the best weighting factor is chosen as reference.

The performance of the estimators is evaluated by Monte Carlo simulations and the results are compared with the Cramer-Rao lower bound (CRLB). On the one hand, simulations are performed over SNR since the accuracy of the soft information methods depends on the SNR. In each run, a new MS position \mathbf{p} is determined randomly inside the region of Figure 5. On the other hand, simulations are performed over space for a fixed SNR in order to assess the influence of the GDOP. A fixed 4×4 grid of MS positions is applied in this case.

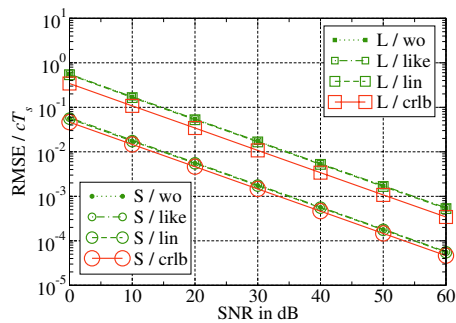
Different channel realizations are generated during the Monte Carlo simulations. Since different channel realizations result in different weighting matrices \mathbf{W} , a mean CRLB is introduced,

$$\overline{\text{CRLB}}(\mathbf{p}) = E \left\{ \sum_{d=1}^2 [\mathbf{I}(\mathbf{p})^{-1}]_{dd} \right\}, \quad (55)$$

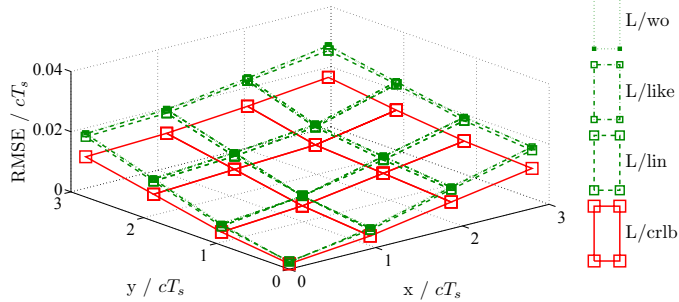
where the expectation is taken with respect to the channel realizations. For the simulations over SNR, the expectation is additionally taken with respect to the random positions \mathbf{p} .

The simulation results are shown in Figure 6. There are eight different graphs (6a, b, c, d, e, f, g, h) arranged in an array with two columns and four rows. In the first column, the results for the simulations over SNR are shown. The second column contains the results for the simulations over space at 30 dB. In each row, the results for a fixed simulation setup are illustrated. All graphs show the root mean squared error (RMSE) of $\hat{\mathbf{p}}$ normalized with respect to $d_s = cT_s$ for positioning without soft information ("wo"), with soft information from the likelihood method ("like"), and with soft information from the linearization method ("lin"). The square root of the mean CRLB (normalized with respect to d_s), which is denoted simply as CRLB in the following, is plotted for comparison ("crlb"). Curves labeled with "L" were obtained for the first scenario with large average GDOP, and curves labeled with "S" were obtained for the second scenario with small average GDOP.

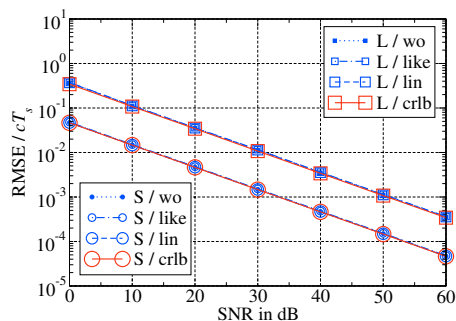
At first, the results for the single-path channel are discussed because this scenario represents an optimal case: Both soft information methods are accurate (see Section 4.2) and due to power control, the pseudo-range errors



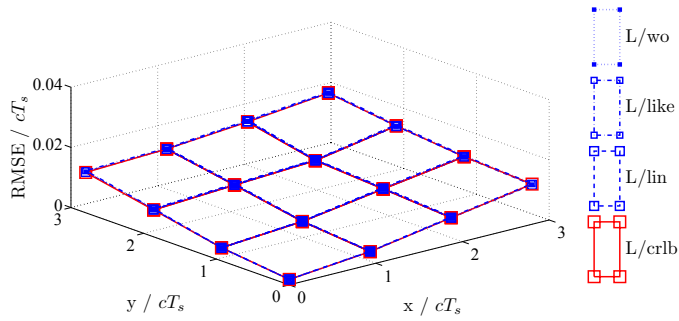
(a) WLS method for a single path channel.



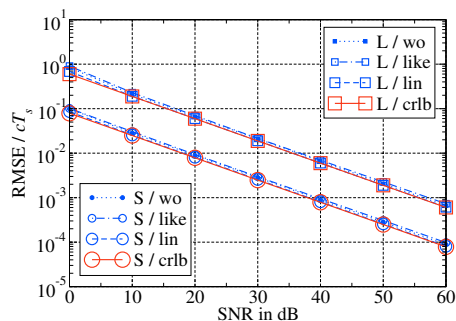
(b) WLS method for a single path channel at 30 dB.



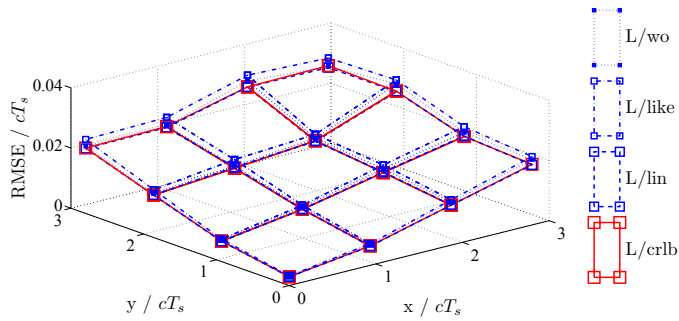
(c) TSA for a single path channel.



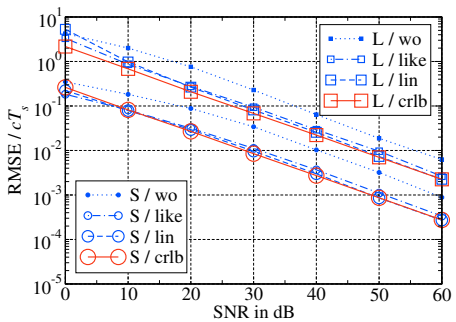
(d) TSA for a single path channel at 30 dB.



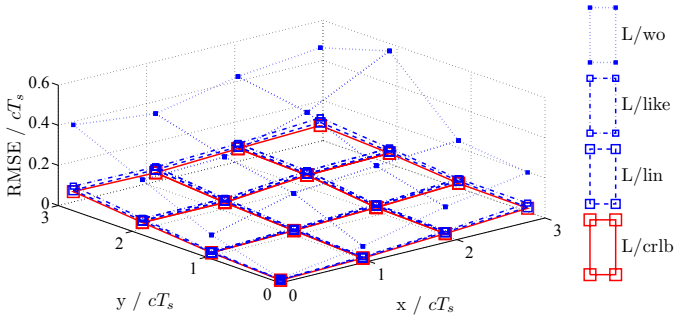
(e) TSA for a two path channel with $\Delta\tau_2 \in [T_s, 2T_s]$.



(f) TSA for a two path channel with $\Delta\tau_2 \in [T_s, 2T_s]$ at 30 dB.



(g) TSA for a two path channel with $\Delta\tau_2 \in [\frac{T_s}{10}, T_s]$.



(h) TSA for a two path channel with $\Delta\tau_2 \in [\frac{T_s}{10}, T_s]$ at 30 dB.

Figure 6 RMSE of \hat{p} normalized with respect to $d_s = cT_s$ for positioning without soft information ("wo"), with soft information from the likelihood method ("like"), and with soft information from the linearization method ("lin"). For comparison $\sqrt{\text{CRLB}(\hat{p})}$ normalized with respect to d_s is plotted ("crlb"). Curves labeled with "L" were obtained for the scenario with large average GDOP, and curves labeled with "S" were obtained for the scenario with small average GDOP. (a) WLS method for a single-path channel. (b) WLS method for a single-path channel at 30 dB. (c) TSA for a single-path channel. (d) TSA for a single-path channel at 30 dB. (e) TSA for a two-path channel with $\Delta\tau_2 \in [T_s, 2T_s]$. (f) TSA for a two-path channel with $\Delta\tau_2 \in [T_s, 2T_s]$ at 30 dB. (g) TSA for a two-path channel with $\Delta\tau_2 \in [\frac{T_s}{10}, T_s]$. (h) TSA for a two-path channel with $\Delta\tau_2 \in [\frac{T_s}{10}, T_s]$ at 30 dB.

for all ROs should be the same. Hence, positioning without and with weighting is supposed to perform equally well. The first row of Figure 6 contains the results for the WLS method, whereas the second row shows the results for the TSA. As supposed previously, the RMSE curves for positioning without soft information and with soft information from the likelihood and the linearization method coincide. The TSA is furthermore a MVU estimator since the RMSE approaches the CRLB for all SNRs and for all positions. The WLS method performs worse: There is a certain gap between the CRLB and the RMSE. In Figure 6b, it can be observed that this gap depends on the position and, thus, on the GDOP: The larger the GDOP is, the larger is the gap. Hence, the gap between RMSE and CRLB in Figure 6a is smaller for the second scenario ("S") since the GDOP is smaller on average. For the two-path channels, a similar behavior of the WLS method was observed. Therefore, only the results for the TSA are considered in the following due to its superior performance.

The third and fourth row of Figure 6 show the simulation results for the two-path channels with large and small excess delay, respectively. It was observed in Section 4.2 that the likelihood method is generally accurate even for multipath channels. In contrast, the accuracy of the linearization method depends on the excess delay and the SNR. The smaller the excess delay, the higher is the nonlinearity of the problem and the less accurate is the linearization method. The accuracy increases with SNR. Hence, it is supposed that the likelihood method outperforms the linearization method. Only at very high SNR, both methods are assumed to perform equally well. Surprisingly, the linearization and the likelihood method show approximately the same performance for all cases. The linearization method performs even slightly better in most cases. Only for very low SNR and a small excess delay the likelihood method outperforms the linearization method. The likelihood method seems to be more susceptible to the GDOP. Hence, the inaccuracy of the covariance matrices at low SNR barely influences the positioning accuracy. Actually, it seems that the absolute value of the weights in the weighting matrices \mathbf{W} and \mathbf{W}' is not crucial. Rather a correct ratio of the weights is relevant. Thus, rough soft information is sufficient as long as the ratio of the pseudo-range variances is accurate. This is fulfilled even for the inaccurate covariance matrices of the linearization method. Hence, it is suggested to apply the linearization method because of its lower computational complexity.

For the two-path channel with large excess delay (Figure 6e, f), the RMSE with or without soft information is almost the same since the multipath components can already be separated by the estimator quite well. For a

small excess delay (Figure 6g, h), the RMSE with soft information is much closer to the CRLB than without soft information. With respect to SNR, a gain of approximately 7-10 dB is achieved (see Figure 6g). Furthermore, positioning with soft information is less susceptible to the GDOP (see Figure 6h). Thus, soft information is well suited to mitigate severe multipath propagation. The smaller the excess delay is, the more important it is to apply soft information for positioning.

The influence of the GDOP can be neglected for the scenario with small average GDOP. The curves labeled with "S" indicate that even for *one-shot estimation without oversampling* a positioning accuracy much smaller than the distance corresponding to the symbol duration, d_s , is achieved for all channel models.

For all simulations, a LOS path has been assumed so far. Hence, the estimated TOA corresponds to distance between transmitter and receiver. However, in urban or indoor environments, the LOS path is often blocked as already mentioned in Section 2.1. Therefore, the influence of NLOS propagation is discussed here. In case of NLOS, a modeling error is introduced that reduces the positioning accuracy significantly. The proposed soft channel parameter estimator does not take a priori information about the physical channel (e.g., probability of NLOS) into account and, hence, is not able to detect such a modeling error. The obtained soft information can only be used to mitigate multipath propagation. In order to mitigate NLOS effects, further processing has to be done (e.g., [24]).

Nevertheless, multipath mitigation is an important issue. The multipath mitigation ability of the proposed soft channel parameter estimation has been presented for $M = 2$ paths due to clarity and simplicity reasons. The influence of the number of multipath components is as follows: The complexity of the soft channel parameter estimator increases with the number of multipath components. Furthermore, the reliability of the estimates decreases with M . Hence, the positioning accuracy deteriorates. If M is large and the scatterers are closely spaced (dense multipath), the estimator becomes biased and the positioning accuracy saturates. In general, it is suggested to consider only the dominant paths if M is large.

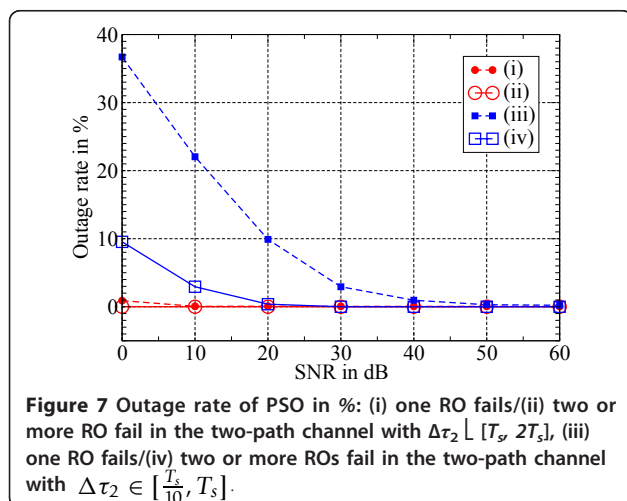
It was mentioned before that the TSA may diverge. Divergence occurred for large GDOP when the initial guess was far from the true position.¹ This happened only rarely. The initial guess is determined by the WLS method which is very susceptible to the GDOP. Hence, the starting position may be far away from the true position for large GDOP.

As mentioned in Section 2.2, PSO does not assure global convergence. For both two-path channels, PSO sometimes converges prematurely. In most of these

cases, it converges to a boundary of the search space, such that the premature convergence can be detected (outage). In Figure 7, the outage rates are shown for both two-path channels: The dashed lines (i) and (iii) denote the probability that the delay estimation fails for one RO and the solid lines (ii) and (iv) denote the probability that two or more ROs fail. If the delay estimation fails for one RO, the position of the MS can be determined nevertheless since only three ROs are necessary for positioning in two dimensions. Only if two or more ROs fail, the position estimation fails, too. By adding more ROs, the outage rate for positioning can be decreased to an arbitrary small amount. The outage rates for the two-channel models differ significantly. For the two-path channel with large excess delay ($\Delta\tau_2 \in [T_s, 2T_s]$), the outage rates (i) and (ii) are negligible. In contrast, the outage rates (iii) and (iv) for the two-path channel with small excess delay ($\Delta\tau_2 \in [\frac{T_s}{10}, T_s]$) are quite high at low SNR but decrease significantly with increasing SNR. The smaller the excess delay is, the higher is the probability that PSO converges prematurely.

6 Conclusions

In this paper, a channel parameter estimator based on the maximum-likelihood approach is proposed for joint communication and positioning. The parameters of the physical channel (e.g., TOA) and the equivalent discrete-time channel model are estimated jointly. In order to mitigate multipath propagation effects and to improve the positioning accuracy, soft information concerning the parameter estimates is used. Two different methods to obtain soft information are proposed: The linearization and the likelihood method.



The accuracy of the methods depends on the nonlinearity of the parameter estimation problem, which is evaluated by the curvature measures of Bates and Watts. It is shown that the likelihood method is always accurate for the parameter estimation problem. The linearization method is only accurate in a single-path channel or at high SNR for a multipath channel. Nevertheless, Monte Carlo simulations for a two-dimensional positioning problem show that this has only very little influence on the positioning. The positioning algorithms that exploit the soft information obtained by the linearization and the likelihood method perform equally well. For severe multipath propagation, the RMSEs for the weighted positioning algorithms are closer to the CRLB than the RMSE of positioning without weighting. A gain of approximately 7-10 dB can be achieved. Hence, multipath propagation effects can be mitigated significantly, even for *one-shot* estimation *without oversampling*. Based on these results, it is suggested to apply the linearization method because of its lower computational complexity.

Endnotes

^aFor oversampling with factor J it follows: $\mathcal{T} = \frac{T_s}{J}$. ^bThe mean squared error of the channel estimates \hat{h}_l is reduced in comparison to the mean squared error of the least-squares channel estimates \check{h}_l , if the number of parameters, $3M$, is less than the number of channel coefficients, $L + 1$, to be estimated. For simulation results please refer to [4]. ^c C_{approx} corresponds to \hat{V}_a in [11] for a complex-valued problem instead of a real-valued problem. ^dThe superscript \tilde{T} , which denotes *tangential*, should not be mistaken for the superscript T , which denotes the *transpose* of a matrix. ^eIn [13] a simplified method to determine the maximum relative curvatures is introduced based on linear transformations of the coordinates in the parameter and the sample space. This method is neglected here because it is out of the scope of this paper. ^fIn a two-dimensional TOA scenario at least three ROs are required. For positioning in three dimensions a fourth RO is needed. ^gFor the linearization method the variance of the TOA corresponds to the 3rd diagonal entry of the approximate covariance matrix C_{approx} . ^hFurthermore, channel parameter estimation was performed for the LMM described in [18] with the true parameters θ as initial guess. Since PSO and the LMM provided approximately the same performance, only PSO is considered here for conciseness. ⁱThe outliers due to divergence were not considered in the calculation of the RMSE.

Appendix

In the following, the equivalence of the maximum-likelihood estimators based on (9) and (10) is shown. First, both metrics are stated in vector/matrix notation. Then, the equivalence of both metrics is proven given the assumptions of Section 2.2. For readability, the terms $\mathbf{h} = \mathbf{h}(\boldsymbol{\theta})$ and $\tilde{\mathbf{h}} = \mathbf{h}(\tilde{\boldsymbol{\theta}})$ are introduced, where $\boldsymbol{\theta}$ denotes the true parameter set and $\tilde{\boldsymbol{\theta}}$ denotes the hypothetical parameter set.

The metric $\Omega(\tilde{\boldsymbol{\theta}})$ corresponding to (9) was already derived in (11):

$$\begin{aligned} \Omega(\tilde{\boldsymbol{\theta}}) &= \sum_{l=0}^L \left| \tilde{h}_l - h_l(\tilde{\boldsymbol{\theta}}) \right|^2 \\ &= (\tilde{\mathbf{h}} - \tilde{\mathbf{h}})^H (\tilde{\mathbf{h}} - \tilde{\mathbf{h}}). \end{aligned} \quad (56)$$

Equivalently, a metric corresponding to (10) can be derived:

$$\begin{aligned} \Psi(\tilde{\boldsymbol{\theta}}) &= \sum_{k=L}^{K_t-1} \left| y[k] - \sum_{l=0}^L h_l(\tilde{\boldsymbol{\theta}}) x[k-l] \right|^2 \\ &= (\mathbf{y} - \mathbf{X}\tilde{\mathbf{h}})^H (\mathbf{y} - \mathbf{X}\tilde{\mathbf{h}}). \end{aligned} \quad (57)$$

As both metrics have to be minimized, it is sufficient to show that

$$E \left\{ \Omega(\tilde{\boldsymbol{\theta}}) \right\} = c \cdot E \left\{ \Psi(\tilde{\boldsymbol{\theta}}) \right\}, \quad (58)$$

where c is a constant that scales the metric but does not change the location of the minimum. The expectation of the first metric can also be written as

$$\begin{aligned} E \left\{ \Omega(\tilde{\boldsymbol{\theta}}) \right\} &= E \left\{ (\tilde{\mathbf{h}} - \tilde{\mathbf{h}})^H (\tilde{\mathbf{h}} - \tilde{\mathbf{h}}) \right\} \\ &= E \left\{ \tilde{\mathbf{h}}^H \tilde{\mathbf{h}} - 2\tilde{\mathbf{h}}^H \tilde{\mathbf{h}} + \tilde{\mathbf{h}}^H \tilde{\mathbf{h}} \right\} \\ &= E \left\{ (\mathbf{h} + \boldsymbol{\varepsilon})^H (\mathbf{h} + \boldsymbol{\varepsilon}) - 2(\mathbf{h} + \boldsymbol{\varepsilon})^H \tilde{\mathbf{h}} + \tilde{\mathbf{h}}^H \tilde{\mathbf{h}} \right\} \\ &= E \left\{ \mathbf{h}^H \mathbf{h} + 2\mathbf{h}^H \boldsymbol{\varepsilon} + \boldsymbol{\varepsilon}^H \boldsymbol{\varepsilon} - 2\mathbf{h}^H \tilde{\mathbf{h}} - 2\boldsymbol{\varepsilon}^H \tilde{\mathbf{h}} + \tilde{\mathbf{h}}^H \tilde{\mathbf{h}} \right\} \\ &= E \left\{ \mathbf{h}^H \mathbf{h} \right\} + 2E \left\{ \mathbf{h}^H \boldsymbol{\varepsilon} \right\} + E \left\{ \boldsymbol{\varepsilon}^H \boldsymbol{\varepsilon} \right\} \\ &\quad - 2E \left\{ \mathbf{h}^H \tilde{\mathbf{h}} \right\} - 2E \left\{ \boldsymbol{\varepsilon}^H \tilde{\mathbf{h}} \right\} + E \left\{ \tilde{\mathbf{h}}^H \tilde{\mathbf{h}} \right\} \\ &= E \left\{ \mathbf{h}^H \mathbf{h} \right\} + 0 + \sigma_{\varepsilon}^2 \\ &\quad - 2E \left\{ \mathbf{h}^H \tilde{\mathbf{h}} \right\} - 0 + E \left\{ \tilde{\mathbf{h}}^H \tilde{\mathbf{h}} \right\} \\ &= E \left\{ \mathbf{h}^H \mathbf{h} \right\} - 2E \left\{ \mathbf{h}^H \tilde{\mathbf{h}} \right\} + E \left\{ \tilde{\mathbf{h}}^H \tilde{\mathbf{h}} \right\} + \sigma_{\varepsilon}^2 \\ &= E \left\{ (\mathbf{h} - \tilde{\mathbf{h}})^H (\mathbf{h} - \tilde{\mathbf{h}}) \right\} + \frac{\sigma_{\varepsilon}^2}{K_t - L}. \end{aligned} \quad (59)$$

For the second metric follows similarly

$$\begin{aligned} E \left\{ \Psi(\tilde{\boldsymbol{\theta}}) \right\} &= E \left\{ (\mathbf{y} - \mathbf{X}\tilde{\mathbf{h}})^H (\mathbf{y} - \mathbf{X}\tilde{\mathbf{h}}) \right\} \\ &= E \left\{ \mathbf{y}^H \mathbf{y} - 2\mathbf{y}^H \mathbf{X}\tilde{\mathbf{h}} + \tilde{\mathbf{h}}^H \mathbf{X}^H \mathbf{X}\tilde{\mathbf{h}} \right\} \\ &= E \left\{ (\mathbf{X}\mathbf{h} + \mathbf{n})^H (\mathbf{X}\mathbf{h} + \mathbf{n}) \right. \\ &\quad \left. - 2(\mathbf{X}\mathbf{h} + \mathbf{n})^H \mathbf{X}\tilde{\mathbf{h}} + \tilde{\mathbf{h}}^H \mathbf{X}^H \mathbf{X}\tilde{\mathbf{h}} \right\} \\ &= E \left\{ \mathbf{h}^H \mathbf{X}^H \mathbf{X}\mathbf{h} + 2\mathbf{h}^H \mathbf{X}^H \mathbf{n} + \mathbf{n}^H \mathbf{n} \right. \\ &\quad \left. - 2\mathbf{h}^H \mathbf{X}^H \mathbf{X}\tilde{\mathbf{h}} - 2\mathbf{n}^H \mathbf{X}\tilde{\mathbf{h}} + \tilde{\mathbf{h}}^H \mathbf{X}^H \mathbf{X}\tilde{\mathbf{h}} \right\} \\ &= E \left\{ \mathbf{h}^H \mathbf{X}^H \mathbf{X}\mathbf{h} \right\} + 2E \left\{ \mathbf{h}^H \mathbf{X}^H \mathbf{n} \right\} + E \left\{ \mathbf{n}^H \mathbf{n} \right\} \\ &\quad - 2E \left\{ \mathbf{h}^H \mathbf{X}^H \mathbf{X}\tilde{\mathbf{h}} \right\} - 2E \left\{ \mathbf{n}^H \mathbf{X}\tilde{\mathbf{h}} \right\} \\ &\quad + E \left\{ \tilde{\mathbf{h}}^H \mathbf{X}^H \mathbf{X}\tilde{\mathbf{h}} \right\} \\ &= (K_t - L)E \left\{ \mathbf{h}^H \mathbf{h} \right\} + 0 + \sigma_n^2 \\ &\quad - 2(K_t - L)E \left\{ \mathbf{h}^H \tilde{\mathbf{h}} \right\} - 0 + (K_t - L)E \left\{ \tilde{\mathbf{h}}^H \tilde{\mathbf{h}} \right\} \\ &= (K_t - L) \left[E \left\{ \mathbf{h}^H \mathbf{h} \right\} - 2E \left\{ \mathbf{h}^H \tilde{\mathbf{h}} \right\} \right. \\ &\quad \left. + E \left\{ \tilde{\mathbf{h}}^H \tilde{\mathbf{h}} \right\} + \frac{\sigma_n^2}{K_t - L} \right] \\ &= (K_t - L) \left[E \left\{ (\mathbf{h} - \tilde{\mathbf{h}})^H (\mathbf{h} - \tilde{\mathbf{h}}) + \frac{\sigma_n^2}{K_t - L} \right\} \right]. \end{aligned} \quad (60)$$

Comparing (59) and (60) shows that (58) is valid with $c = \frac{1}{K_t - L}$.

Competing interests

The authors declare that they have no competing interests.

Received: 30 November 2010 Accepted: 23 November 2011

Published: 23 November 2011

References

1. R Raulefs, S Plass, C Mensing, The where project: combining wireless communications and navigation, in *Proceedings of 20th Wireless World Research Forum (WWRF)*, Ottawa, Canada (2008)
2. K Pahlavan, AH Levesque, *Wireless information networks*, ch. 13: RF location sensing, Wiley, Hoboken, New Jersey (2005)
3. K Cheung, H So, W-K Ma, Y Chan, A constrained least squares approach to mobile positioning: algorithms and optimality. *EURASIP J Appl Signal Process* 23 (2006). p (2006) Article ID 20858
4. K Schmeink, R Block, PA Hoeher, Joint channel and parameter estimation for combined communication and navigation using particle swarm optimization, in *Proceedings of Workshop on Positioning, Navigation and Communication (WPNC)*, Dresden, Germany (2010)
5. A Khayrallah, R Ramesh, G Bottomley, D Koilpillai, Improved channel estimation with side information, in *Proceedings of IEEE Vehicular Technology Conference (VTC Spring)*, vol. 2. (Phoenix, Arizona, 1997), pp. 1049–1053
6. JW Liang, B Ng, JT Chen, A Paulraj, GMSK linearization and structured channel estimate for GSM signals, in *Proceedings of IEEE Military Communications Conference (MILCOM)*, vol. 2. (Monterey, California, 1997), pp. 817–821
7. H-N Lee, GJ Pottie, Fast adaptive equalization/diversity combining for time-varying dispersive channels. *IEEE Trans Commun.* **46**(9), 1146–1162 (1998). doi:10.1109/26.718557

8. M Feder, E Weinstein, Parameter estimation of superimposed signals using the EM algorithm. *IEEE Trans Acoust Speech Signal Process.* **36**(4), 477–489 (1988). doi:10.1109/29.1552
9. BH Fleury, M Tschudin, R Heddergott, D Dalhaus, KI Pedersen, Channel parameter estimation in mobile radio environments using the SAGE algorithm. *IEEE J Sel Areas Commun.* **17**(3), 434–450 (1999). doi:10.1109/49.753729
10. A Richter, M Landmann, RS Thomä, Maximum likelihood channel parameter estimation from multidimensional channel sounding measurements, in *Proceedings of IEEE Vehicular Technology Conference (VTC Spring)*, Jeju, Korea, pp. 1056–1060 (2003)
11. JR Donaldson, RB Schnabel, Computational experience with confidence regions and confidence intervals for nonlinear least squares, in *Proceedings of 17th Symposium on the Interface of Computer Sciences and Statistics*, Lexington, Kentucky, pp. 83–93 (1985)
12. M Schwaab, EC Biscaia Jr, JL Monteiro, JC Pinto, Nonlinear parameter estimation through particle swarm optimization. *Chem Eng Sci.* **63**(6), 1542–1552 (2008). doi:10.1016/j.ces.2007.11.024
13. DM Bates, DG Watts, Relative curvature measures of nonlinearity. *J R Stat Soc Ser B (Methodological).* **42**(1), 1–25 (1980)
14. SM Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory* (Prentice-Hall, Upper Saddle River, New Jersey, 1993)
15. J Kennedy, R Eberhart, Particle swarm optimization, in *Proceedings of IEEE International Conference on Neural Networks*, vol. 4. Perth, Australia, pp. 1942–1948 (1995)
16. D Bratton, J Kennedy, Defining a standard for particle swarm optimization, in *Proceedings of IEEE Swarm Intelligence Symposium (SIS)*, Honolulu, Hawaii, pp. 120–127 (2007)
17. C Blum, X Li, Swarm Intelligence in Optimization, in *Swarm Intelligence: Introduction and Applications*, ser. *Natural Computing Series*, ed. by C Blum, D Merkle (Springer, 2008), pp. 43–85
18. WH Press, SA Teukolsky, WT Vetterling, BP Flannery, *Numerical Recipes in C++: The Art of Scientific Computing* (Cambridge University Press, Cambridge, 2002)
19. J Nocedal, SJ Wright, *Numerical Optimization* (Springer, New York, 1983)
20. PA Hoehner, P Robertson, E Offer, T Woerz, The soft-output principle: reminiscences and new developments. *Eur Trans Telecommun.* **18**(8), 829–835 (2007). doi:10.1002/ett.1200
21. JE Dennis Jr, RB Schnabel, *Numerical Methods for Unconstrained Optimization and Nonlinear Equations* (Prentice-Hall, Englewood Cliffs, New Jersey, 1983)
22. ED Kaplan, (ed.), *Understanding GPS: Principles and Applications* (Artech House, Boston, 1996)
23. AH Sayed, A Tarighat, N Khajenouri, Network-based wireless location: challenges faced in developing techniques for accurate wireless location information. *IEEE Signal Process Mag.* **22**(4), 24–40 (2005)
24. I Guevenc, CC Chong, F Watanabe, H Inamura, NLOS identification and weighted least-squares localization for UWB systems using multi-path channel statistics. *EURASIP J Appl Signal Process* 1–14 (2008). (2008) Article ID 271984
25. RB Langley, Dilution of precision. *GPS World.* **10**(5), 52–59 (1999)

doi:10.1186/1687-1499-2011-185

Cite this article as: Schmeink et al.: Joint communication and positioning based on soft channel parameter estimation. *EURASIP Journal on Wireless Communications and Networking* 2011 **2011**:185.

Submit your manuscript to a SpringerOpen® journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Immediate publication on acceptance
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► springeropen.com
