IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received November 21, 2020, accepted December 17, 2020, date of publication December 28, 2020,
date of current version January 7, 2021.

Digital Object Identifier 10.1109/ACCESS.2020.3047618

Simultaneous Model and Parameter Estimation for Joint Communication and Positioning

REBECCA CARRIE ADAM[1] AND PETER ADAM HOEHER[2], (Fellow, IEEE)
[1]Centre for Industrial Electronics, University of Southern Denmark, 6400 Sønderborg, Denmark
[2]Chair of Information and Coding Theory, Kiel University, 24143 Kiel, Germany

Corresponding author: Rebecca Carrie Adam (rebeccaadam@sdu.dk)

ABSTRACT Joint communication and positioning based on a unified signal structure yields synergy and can complement and assist system designs enabling higher coverage and quality of service for both communication and positioning. For the time of arrival (TOA) based joint communication and positioning the TOA estimation accuracy is crucial. It is known to translate directly into position estimation accuracy. In the presence of multipath propagation, the estimation accuracy of signal arrival times in return strongly depends on the actual as well as the estimated number of the physical path parameters, the model order. In this work, we assess the performance and the mutual impact of simultaneous model order and parameter estimation for channel-estimation-based joint communication and positioning. Besides introducing a terrestrial channel-estimation-based unified joint communication and positioning system framework, we discuss and numerically compare different methods to sequentially or jointly estimate the parameters and the model order. We show that a TOA error-minimizing model order estimation is preferable over estimating the correct model order. Furthermore, we compare the performance with a proposed focused order-related lower bound. This bound determines the optimal model order for a chosen estimator. It depends on the actual and hypothetical model order and it replaces the here unsuitable Cramer-Rao lower bound. Secondly, the comparison shows that employing the parameter and model order-dependent inverse Fisher information matrix yields a close-to-optimal approach. We numerically show for a realistic channel scenario with many multipath parameters that the method is still accurate.

INDEX TERMS Parameter estimation, model selection, information theoretic criteria, information complexity criterion, Fisher information.

I. INTRODUCTION

Future wireless communication system designs will integrate positioning capabilities as a core system feature from the beginning [1]. When it comes to designing a joint communication and positioning system from scratch, channel estimation-based multipath parameter estimation offers the possibility to obtain positioning-relevant information like the time of arrival (TOA) or the angle of arrival (AOA) and hence can play a central role [2].

Wireless communication standards and novel joint communication and positioning system concepts have adopted the idea to gain synergy by using a unified signal structure for communication and positioning [3]. The authors of [4]–[6] introduce a unified signal structure as an alternative terrestrial approach to the global navigation satellite system (GNSS)-based system. The long term evolution (LTE) signal design considers positioning reference signals from the beginning. More precisely, the authors of [5] investigated the positioning accuracy for such approaches. Furthermore, the authors of [7] investigate a potential receiver design for long term evolution (LTE) signals, whereas the authors of [8] discuss potential mm-wave radar communication system approaches.

For positioning, the multipath delays and the number of multipath components (the number of propagating ray clusters) have to be estimated. The latter determines the model order. We assess a simultaneous model selection and parameter estimation approach and provide numerical results for a joint positioning and communication system. To illustrate the relationship between related contributions and novel contributions, we categorize our overview into open challenges, related contributions, and novel contributions.

The associate editor coordinating the review of this manuscript and approving it for publication was Nan Wu.

A. OPEN CHALLENGES

Although the range of contributions in the area of model selection [9]–[22] as well as in the area of parameter estimation [23]–[30] indicates that both research areas are rather mature, there are still open questions and challenges:

- How to efficiently find the position in the presence of a realistic multipath propagation scenario is still known to be a challenging problem. Such a scenario usually leads to a severe parameter estimation performance degradation [1] compared to a single path scenario. Fading multipath components and non-line-of-sight conditions are difficult to extract from the additive noise [2].
- Furthermore, an increasing number of parameters entails an increasing estimation error due to a higher problem complexity (overestimation). If, on the other hand, the number of parameters is assumed too small (underestimation), a systematic modeling error has to be taken into account. Positioning applications target an especially high estimation accuracy of the positioning-relevant parameters. Therefore, methods yielding the optimal parameter estimation performance for positioning to simultaneously estimate the parameters and model order (number of multipath clusters) can still be improved.
- Especially for estimating the multipath fading channel parameters it is obvious that automatically excluding paths from the model with a power below the noise power is beneficial [28]. Underestimating the model order is beneficial in these cases. For positioning or delay estimation, excluding paths potentially yields a lower delay estimation error [6]. Following this line of thought, the probability of correct detection is not a sufficient performance measure to assess the model order selection performance. Furthermore, the Cramer-Rao Lower Bound (CRLB) [31] requires the correct model order. If the correct model order is too high for the number of observations, the CRLB fails to assess the performance. The formulation of a more practical, order-related lower bound is required. Numerical results for simultaneous model selection and parameter estimation should be compared to this bound.

This contribution addresses all these challenges. We assess a simultaneous model selection and parameter estimation approach for a channel estimation-based joint communication and positioning framework targeting an optimal delay accuracy. Moreover, we propose a novel, practical model order-related lower bound. Similar to a problem stated in radar [32] we assume here that the correct model order does not necessarily yield the optimal performance. Our approach employs the delay's Fisher information matrix in the model selection as a natural solution to exclude undesirable paths in the model order detection. This approach outperforms classical information-theoretic approaches.

B. RELATED CONTRIBUTIONS

In the following, we provide an overview of approaches and contributions related to this work.

1) CONTRIBUTIONS USING REDUCED COMPLEXITY MODELS

Instead of performing model selection and parameter estimation together, [6], [33] propose model simplification to yield an acceptable ranging accuracy to computational cost tradeoff in multipath scenarios. Such approaches have the drawback that they result in a positioning accuracy limited by a modeling error, which can be very high depending on the applied scenario.

2) CONTRIBUTIONS ASSUMING THE CORRECT MODEL ORDER IS OPTIMAL

Many researchers assume that the model order can be estimated correctly and that this is the optimal choice [23], [24], [27], [34], [35]. The authors focused on the parameter estimation results, employed the correct model order, and neglected the possibility that this assumption is not always the optimal choice. Often these contributions focus on proposing a special estimation or detection algorithm. Hence the simulation setups are designed to prove that the algorithm reliably works in a minimalistic setup, like in a two-path scenario. The impact of model selection and realistic channel modeling exceeds the scope of these contributions.

3) CONTRIBUTIONS USING THE MODEL ORDER DETECTION PROBABILITY OF CORRECT DETECTION AS A PERFORMANCE MEASURE

In a range of contributions, researchers investigate and compare different model selection methods by choosing the probability of correct detection as a performance measure for model selection [10], [11], [14], [16]–[19], [22], [36], [37]. These contributions present various strategies to estimate the model order. Nevertheless, they do not consider the impact on the parameter estimation performance or small sample sizes. In [38], [39] the parameter estimation performance is considered. The authors use higher-order arrays combined with the parallel factor (PARAFAC) methods. These methods have the advantage of performing robustly.

4) FREQUENCY-DOMAIN-BASED JOINT MODEL ORDER SELECTION AND PARAMETER ESTIMATION CONTRIBUTIONS

Recent approaches [40], [41] to simultaneously estimate the model order and the parameters rely on a signal model described in the frequency domain. These models come with a systematic error. The parameter estimation results are compared to the CRLB and not to an order specific bound demonstrating the suitability for positioning. These contributions also do not address particular application associated requirements. Consequently, the authors do not consider channel conditions for which an underestimated model yields superior results. In [17], the authors propose

a promising shift-invariance based order selection technique for exponential data modeling. This technique was then further employed in a proposal for joint parameter estimation and model order selection in [40]. The authors could numerically show the impact of model order selection on the parameter estimation. As we will clarify when introducing the system framework used here, formulating the signal model in the frequency domain like in [17], [21], [40]–[44] is impossible without some undesired signal design impairments like oversampling for the here investigated joint communication and positioning framework. Secondly, it is not possible without accepting a limited estimation accuracy. The contributions in [17], [40]–[43] furthermore do not investigate a lower performance bound tailored to simultaneous model selection and parameter estimation for positioning.

5) CONTRIBUTIONS EMPLOYING SEQUENTIAL PROCESSING AND BAYESIAN ESTIMATION

The authors of [45] base their work on a sequential signal model and propose and assess a particle algorithm for sequential Bayesian parameter estimation and model selection. Sequential processing requires initial guesses for the parameter estimation problem. Similar to [45], the authors of [46] employ a Bayesian strategy to jointly estimate the model order and parameter estimation by population Monte Carlo simulation. Bayesian approaches exploit the parameters' a priori distributions. In a practical scenario, the physical channel parameters' a priori distributions will not be known to the receiver side.

6) COMPRESSIVE SENSING JOINT MODEL ORDER SELECTION AND PARAMETER ESTIMATION CONTRIBUTIONS (ON-GRID DELAY MODEL)

Compressive sensing constitutes a valid approach for joint model selection and parameter estimation. In [47], the authors investigated the relation between sparse reconstruction and parameter estimation with model order selection, proposing to use sparsity parameters via compressive sensing. Their approach mimics classic order selection criteria. On the other hand, compressive sensing entails the necessity of defining an inexact ''on-grid'' delay model.

C. NOVEL CONTRIBUTIONS

We summarize the novel contributions in this article as:

- Overcoming the limitations of reduced complexity models by performing model selection and parameter estimation simultaneously or jointly. We propose to determine the delay estimates for different hypothetical model orders. We then propose to choose the estimator and the model order yielding the best delay estimation accuracy.
- Assessing a soft information-based simultaneous model order selection and parameter estimation designed to yield accurate delay estimates for a joint communication and positioning system. We demonstrate that the following technique is particularly suited to perform well in such systems. Here we use a soft information complexity criterion (ICOMP). The ICOMP criterion utilizes the Fisher information matrix as a complexity measure. We can see this as tuning to exclude low energy paths for multipath parameter estimation and model selection.
- Focusing the model order detection to the positioning-relevant parameters, the delays. The results [6] (page 74 and page 123) indicate that simply assuming the correct model order is not the optimal choice for positioning. We formulate a for TOA error performance beneficial automatism to estimate the model order optimally. Our approach takes into account the instantaneous channel conditions.
- Providing a theoretical, more practical, and appropriate focused order-related bound (FORLB) than the CRLB to quantify a lower theoretical bound for the joint model selection and parameter estimation performance in general.
- Applying the simultaneous model selection and parameter estimation to the proposed generalized joint communication and positioning signal model. We show numerical results for a continuous delay model and time-domain processing to circumvent additional systematic error sources as described in the subsubsections I-B4 and I-B6. Furthermore, we use the channel parameters' invariance and apply block-wise estimation. Using global optimization, we overcome the necessity of initial guesses. To target a realistic receiver design, we assume that we do not know the a priori parameter probability distribution.
- Investigating the impact of a realistic fading channel and the availability of a small number of measurements for simultaneous model selection and parameter estimation. We show that our approach yields an accurate parameter estimation performance in realistic channel scenarios.

We organized this article as follows: Firstly, we present a generalized joint communication and positioning system framework based on channel estimation. Then, different model order estimation strategies and parameter estimation strategies are discussed. Afterward, we propose the model order related bound. Finally, we numerically assess the mutual impact of model order estimation and parameter estimation for a typical wireless outdoor-channel model.

D. NOTATION

$\mathrm{Re}(\cdot)$ and $\mathrm{Im}(\cdot)$ stand for the real and the imaginary component of a complex value. We denote vectors by a bold lower case letter notation, whereas we denote matrices by a bold upper case letter notation. Further, $(\cdot)^H$ denotes the Hermitian transpose of a matrix. Hypotheses are represented by $\tilde{(\cdot)}$ and estimates by $\hat{(\cdot)}$. The operator $\mathrm{vec}(\cdot)$ reshapes a matrix into a vector by transferring the matrix column-wise into the vector. The pseudo-inverse of any matrix is denoted by $(\cdot)^\dagger$.

II. THE SYSTEM FRAMEWORK

A. THE SIGNAL MODEL

To calculate the position of a mobile user via the TOA or the AOA without ambiguity, several reference objects with coordinates known to the system are required. Among other network nodes, neighbouring base stations or access points can serve as reference objects. Either the network can determine the position of a device moving in the area spanned by the reference objects, or the device can determine its position via employing signals of multiple base-stations or access points. Since both options are possible and the positioning relevant parameter estimation is carried out for each link separately, in order to achieve a high positioning accuracy, a high single-link parameter estimation accuracy is required. To obtain accurate positioning estimates, we target high-resolution TOA and AOA estimates for each link. Hence, for the sake of clarity, we investigate one single link, keeping in mind that we require at least three links in an actual joint communication and positioning setup. At the transmitter side, a data matrix $\mathbf{X}$ and a known pilot matrix $\mathbf{P}$, which we here assume to choose optimally, are constructed and combined according to rules specified by the actual multiplexing scheme. Together they build the virtual training matrix $\mathbf{V}$. The term virtual training indicates that in addition to the regular pilots the communication data itself can serve as pilot data by iteratively employing the detected data at the receiver side for channel estimation for joint communication and positioning [4]. Different multiplexing schemes will yield a different composition. Independent of the multiplexing scheme the information in $\mathbf{V}$ is supposed to be fully used jointly for communication and positioning in the following manner: We inherently use the pilot symbols in $\mathbf{P}$ as a part of $\mathbf{V}$ for channel estimation for communication at the receiver side. We principally know that the underlying model for the channel depends on the positioning relevant parameters like the multipath delays and that we can use the channel estimates to estimate these parameters. Hence, we propose to doubly exploit the channel estimates for communication, for data detection, and for positioning, for parameter estimation. Note that for the communication detection performance purely training-based estimators are already sufficient, whereas this is not the case for positioning. Positioning algorithms rely on accurate channel estimates. Consequently, we use the total signal energy of the transmitted signal in $\mathbf{V}$ by employing an iterative semi-blind channel estimation strategy [4]. By iteratively including the detected communication data symbols as additional ''virtual'' pilots we can asymptotically approach an optimal channel estimation performance for the positioning side. We illustrate our joint signal design in our framework in Figure 1. Physically, a pulse shaping filter specified by $g_{\mathrm{Tx}}(\tau)$ is applied before transmission over the physical channel and a matched receiver filter $g_{\mathrm{Rx}}(\tau)$ is applied at the receiver side.

Let C denote the number of multipath components, here the model order. The number of receive antennas is denoted by N_r. Further, let the complex path weights of multipath clusters be denoted by $\gamma_{c,v}$ for $v \in \{0 \dots N_r - 1\}$, $c \in \{1, \dots, C\}$ and the cluster delays be denoted by $\boldsymbol{\tau} = [\tau_1, \dots, \tau_C]$. The physical channel $c_v(t, \tau)$ is modeled as the superposition of weighted Diracs delayed by the multipath cluster delays:

$$c_v(t, \tau) = \sum_{c=1}^{C} \gamma_{c,v}(t)\delta(\tau - \tau_c). \tag{1}$$

Thereby, τ_1 is the positioning relevant TOA.

For a moderate mobile velocity the complex path weight $\gamma_{c,v}$ behaves quasi-static in blocks of a certain length. Consequently, we assume block-fading over K times the symbol duration T. That means we assume a quasi-static channel over the time duration KT and that we omit the time-dependency over this block length. From one block to another, however, the complex-valued path amplitudes are assumed to be time-varying. Note that we can assume the delays in $\boldsymbol{\tau}$ to be quasi-static over a certain consecutive block number I. We will use this property to enlarge the number of measurements for the later formulated parameter estimation problem by observing KTI measurements. Further, for the matter of convenience, to formulate an equivalent discrete-time channel model (EDTCM) expression the physical channel is convolved with the convolution of the pulse shaping filter $g_{\mathrm{Tx}}(\tau)$ and the receiver filter $g_{\mathrm{Rx}}(\tau)$, specified by $g(\tau)$

$$g(\tau) = g_{\mathrm{Tx}}(\tau) * g_{\mathrm{Rx}}(\tau). \tag{2}$$

An overall channel impulse response function $h_v(t, \tau)$ with $v \in \{0 \dots, N_r - 1\}$ can be expressed by the convolution of $g(\tau)$ with $c_v(\tau)$

$$h_v(t, \tau) = \sum_{c=1}^{C} \gamma_{c,v}(t)g(\tau - \tau_c). \tag{3}$$

Let the channel memory length be denoted by $L+1$. Sampling at the receiver side with a symbol period T yields the $(L + 1) \times N_r$ EDTCM coefficients $h_{l,v}(i)$ for $l \in \{0, \dots, L\}$, $v \in \{0, \dots, N_r - 1\}$ and $i \in \{0, \dots, I - 1\}$:

$$h_{l,v}(i) = \sum_{c=1}^{C} \gamma_{c,v}(i)g(lT - \tau_c). \tag{4}$$

To obtain a more convenient matrix vector notation let us define the channel matrix $\mathbf{H}$, the so-called spatial signature matrix $\boldsymbol{\Gamma}$ and the delay-dependent pulse matrix $\mathbf{G}(\boldsymbol{\tau})$ entry-wise so that for the single-input multiple-output (SIMO) time series case we have for $i \in \{0, \dots, I - 1\}$

$$[\mathbf{H}]_{l,v}(i) = h_{l,v}(i) \text{ with } \mathbf{H}(i) \in \mathbb{C}^{(L+1\times N_r)} \tag{5}$$

$$[\boldsymbol{\Gamma}]_{c,v}(i) = \gamma_{c,v}(i) \text{ with } \boldsymbol{\Gamma}(i) \in \mathbb{C}^{C\times N_r} \tag{6}$$

$$[\mathbf{G}(\boldsymbol{\tau})]_{l,c} = g(lT - \tau_c) \text{ with } \mathbf{G}(\boldsymbol{\tau}) \in \mathbb{R}^{(L+1)\times C}. \tag{7}$$

Further, a horizontal matrix concatenation for all time indices i yields

$$\mathbf{H} = [\mathbf{H}(0), \dots, \mathbf{H}(I - 1)] \text{ with } \mathbf{H} \in \mathbb{C}^{(L+1\times N_r I)} \tag{8}$$

$$\boldsymbol{\Gamma} = [\boldsymbol{\Gamma}(0), \dots, \boldsymbol{\Gamma}(I - 1)] \text{ with } \boldsymbol{\Gamma} \in \mathbb{C}^{C\times N_r I}. \tag{9}$$

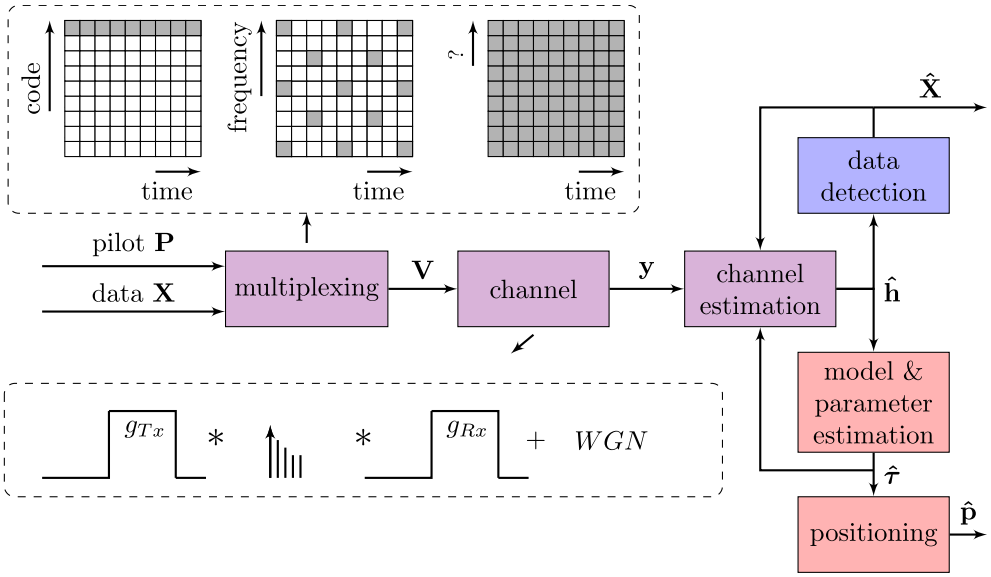

FIGURE 1. The core system framework under investigation yields synergy by relying on a channel estimation concept for communication and positioning.

Then, the matrices $\mathbf{\Gamma}$, $\mathbf{G}(\boldsymbol{\tau})$ and $\mathbf{H}$ are related by

$$\mathbf{H} = \mathbf{G}(\boldsymbol{\tau})\mathbf{\Gamma}. \tag{10}$$

Let $\mathbf{N}(i)$ be a complex Gaussian distributed noise matrix and $\mathbf{Y}(i)$ denote the received signals matrix. The row ranks of both matrices have to match the row rank of $\mathbf{V}$, which depends on the actual multiplexing scheme. Then the system equation can be written as

$$\mathbf{Y}(i) = \mathbf{V}(i)\mathbf{H}(i) + \mathbf{N}(i). \tag{11}$$

As pointed out earlier, the virtual training matrix $\mathbf{V}$ includes a multiplexing dependent combination of the data in $\mathbf{X}$ and the pilots in $\mathbf{P}$. Without loss of generality $\mathbf{V} = \mathbf{X} + \mathbf{P}$ (for TDM and CDM) and $\mathbf{Y}(i), \mathbf{N}(i) \in \mathbb{C}^{(K-L)\times N_r}$.

B. CHANNEL ESTIMATION

Here channel estimation is exploited for communication and for positioning in the receiver as is visualized in Fig. 1. We use a semi-blind channel estimation approach using the total signal energy as training. Why we use this approach, we will illustrate in the following paragraph. For data detection, a purely training-based channel estimation approach yields an acceptable bit-error-rate performance. However, it fails to yield the required accuracy for positioning. In a training-based channel estimation approach, the signal model includes a significant part carrying the communication signal data and a minor part carrying the training symbols. Then the chosen signal design determines how to divide the signal resource block into training and signal parts exactly [34], [59]. Training-based channel estimation utilizes only the received values for the complete signal resource block without exploiting the detected data symbols iteratively. Since the training part is only a small percentage of this utilized resource block, the training-based channel estimation mean squared error is known to approach an error-floor for increasing signal-to-noise ratio [4], [59]. As can be seen from Fig 1 channel estimates are fed to the parameter-estimation algorithm that calculates estimates of the physical path parameters like the delay $\hat{\boldsymbol{\tau}}$. The accuracy of these estimates is proportional to the accuracy of the channel estimates. This dependency becomes clear by assessing the CRLB for the physical path parameters [4]. The parameter estimation accuracy in return impacts the position estimation accuracy. This dependency means that opposed to pure communication systems, joint communication and positioning requires a high channel estimation accuracy exceeding the performance attainable by training-based estimators.

Different approaches can improve the channel estimation performance compared to training-based channel estimation. The straightforward approach to optimize the channel estimation performance targets an optimal training design [60], [61]. Recently, some approaches to improve the pilot-design consider massive MIMO scenarios and the possibility to use parameter estimates gained at the receiver as feedback to the transmitter [62], [63].

Let us assume that the pilot design is optimal. Approaches like successive interference-cancellation as assessed in [59] or signal stripping [57], [58] aim at removing the undesired signal parts from the received signal. They improve the channel estimation performance compared to the training-based estimation by removing the systematic error.

Comparing the mean squared error performance and the CRLBs of the channel estimation approaches training-based, training based with interference-cancellation and semi-blind channel estimation, semi-blind channel estimation outperforms the two competitors [59]. Semi-blind channel estimation utilizes the complete signal plus pilot resource-block energy optimally and shows a superior error performance.

Besides the training symbols, additional data symbols iteratively serve as virtual training [4], [59]. In this contribution, we focus on the simultaneous parameter estimation and model selection for JCAP. Consequently, here we assume that we have an optimal virtual training matrix that can be employed iteratively via semi-blind channel estimation. Since we know that semi-blind channel estimation employs nearly the whole matrix $\mathbf{V}(i)$ as training and for the sake of simplicity, we assume for now that $\mathbf{V}(i)$ is fully constructed of training symbols. Then the channel estimation $\hat{\mathbf{H}}(i)$ is usually carried out according to

$$\hat{\mathbf{H}}(i) = \mathbf{V}^{\dagger}(i)\mathbf{Y}(i) \tag{12}$$

$$\hat{\mathbf{H}} = [\hat{\mathbf{H}}(0), \ldots, \hat{\mathbf{H}}(I-1)] \tag{13}$$

$$\hat{\mathbf{h}} = \mathrm{vec}(\hat{\mathbf{H}}). \tag{14}$$

These channel estimates have twofold purpose: Firstly, we use them for data detection, and secondly, we use them for positioning in the following manner: We interpret the discrete channel estimates in each column of $\hat{\mathbf{H}}$ as snapshot measurements, which we can use to estimate the physical path parameters, by fitting the a priori known model function in (3) to these measurements. Using all columns in $\mathbf{H}$ instead of only a single column (SISO and single-snapshot measurement case) increases the number of observations as well as the number of unknowns.

C. PARAMETER ESTIMATION

Given any reasonable hypothetical parametrization $\tilde{\boldsymbol{\theta}}$ having the delay vector $\boldsymbol{\tau}$ as a subvector and furthermore, given any cost function, $\Omega(\boldsymbol{\theta})$, related to this arbitrary parametrization, the parameter estimation $\hat{\boldsymbol{\theta}}$ in general reads:

$$\hat{\boldsymbol{\theta}} = \arg\min_{\tilde{\boldsymbol{\theta}}}\{\Omega\}. \tag{15}$$

1) DETERMINISTIC MAXIMUM LIKELIHOOD ESTIMATION

Let the stacked parameter vector for a deterministic parametrization $\boldsymbol{\theta}_{\mathrm{dml}}$ be

$$\boldsymbol{\theta}_{\mathrm{dml}} = [\boldsymbol{\tau}, \mathrm{Re}\{\mathrm{vec}(\boldsymbol{\Gamma})\}, \mathrm{Im}\{\mathrm{vec}(\boldsymbol{\Gamma})\}] \in \mathbb{R}^{1\times(C+2N_rIC)}. \tag{16}$$

We have various possibilities to formulate the delay estimator. For the purpose of a convenient notation let us rearrange the values of the received matrix in a long vector instead such that

$$\mathbf{y} = \mathrm{vec}\{\mathbf{Y}\}. \tag{17}$$

Then common way is to formulate the parameter estimation as a received values dependency $\hat{\boldsymbol{\theta}}_{\text{dml-y}}$:

$$\hat{\boldsymbol{\theta}}_{\text{dml-y}} = \arg\min_{\tilde{\theta}}\{(\mathbf{y}-\tilde{\mathbf{y}}(\tilde{\boldsymbol{\theta}}))^H(\mathbf{y}-\tilde{\mathbf{y}}(\tilde{\boldsymbol{\theta}}))\}, \tag{18}$$

$$\tilde{\mathbf{y}} = \mathbf{V}\underbrace{\mathrm{vec}(\mathbf{G}(\tilde{\boldsymbol{\tau}})\tilde{\boldsymbol{\Gamma}})}_{\tilde{\mathbf{h}}}. \tag{19}$$

Assuming that semi-blind channel estimation is inherently performed in a first step, another equivalent and less complex cost function is obtained by formulating the delay estimator $\hat{\boldsymbol{\theta}}_{\text{dml-h}}$ as a dependency of these channel estimates:

$$\hat{\boldsymbol{\theta}}_{\text{dml-h}} = \arg\min_{\tilde{\boldsymbol{\theta}}}\{(\hat{\mathbf{h}}-\tilde{\mathbf{h}}(\tilde{\boldsymbol{\theta}}))^H(\hat{\mathbf{h}}-\tilde{\mathbf{h}}(\tilde{\boldsymbol{\theta}}))\},$$

$$\tilde{\mathbf{h}}(\tilde{\boldsymbol{\theta}}) = \mathrm{vec}(\mathbf{G}(\tilde{\boldsymbol{\tau}})\tilde{\boldsymbol{\Gamma}}). \tag{20}$$

Since the estimation error on the channel estimates and the noise on the received values can be described by a complex Gaussian distribution, this kind of nonlinear least-squares fit in (20) and in (19) is known to be equivalent to the maximum-likelihood (ML) estimator.

Note that, if the delays in $\boldsymbol{\tau}$ were known, the linear parameters $\boldsymbol{\Gamma}$ could be estimated in closed form. Hence, $\tilde{\boldsymbol{\Gamma}}$ in (20) and in (19) can already be substituted beforehand by its estimation result depending on $\boldsymbol{\tau}$:

$$\tilde{\boldsymbol{\Gamma}} = \mathbf{G}^{\dagger}(\tilde{\boldsymbol{\tau}})\hat{\mathbf{H}}. \tag{21}$$

This parameter separability into the nonlinear and linear components and the possibility of resubstituting the closed-form expression applies to both (20) and (19) again and this would leave us with four slightly different estimation options for the delay estimation.

In [48](page 16) we provided proof for the fact that the cost functions (20) and (19) are equivalent and the delay estimates are the same

$$\hat{\boldsymbol{\theta}}_{\text{dml-h}} = \hat{\boldsymbol{\theta}}_{\text{dml-y}} = \hat{\boldsymbol{\theta}}_{\text{dml}}. \tag{22}$$

Consequently, we discard the cost function, which is based on the received values in (19), and will carry on with the cost function based on the channel estimates (20) only. Following the principles of separable nonlinear least-squares estimation [31] (page 255), this substitution yields an elimination of the dependence on the linear parameters

$$\hat{\boldsymbol{\tau}} = \arg\min_{\tilde{\boldsymbol{\tau}}}\{\underbrace{\hat{\mathbf{h}}^H\mathbf{P}_{\mathbf{G}}^{\perp}(\tilde{\boldsymbol{\tau}})\hat{\mathbf{h}}}_{\Omega_{\mathrm{dml}}}\}, \tag{23}$$

with the orthogonal projection matrix

$$\mathbf{P}_{\mathbf{G}}^{\perp}(\tilde{\boldsymbol{\tau}}) = \mathbf{I} - \mathbf{G}(\boldsymbol{\tau})\mathbf{G}^{\dagger}(\boldsymbol{\tau}). \tag{24}$$

Hence, the deterministic ML (DML) estimation task is reduced from a $(2N_rIC+C)$-dimensional problem in (20) to a C-dimensional problem in (23). The estimator (23) is called a deterministic estimator, since the treatment of the parameters is deterministic.

2) STOCHASTIC MAXIMUM LIKELIHOOD ESTIMATION

Exploiting a time series of measurements, that is $I \gg 1$, allows us to treat the problem as either a deterministic or a stochastic estimation problem. More specifically, we can treat the complex path amplitudes as stochastic variables, if they underly the characteristic of a known distribution like the circular complex normal distribution. Then, we can replace the deterministic parameter vector by a set of parameters containing the mean and covariance matrix elements. For instance, assume that the complex path amplitudes are

identically and independently distributed complex normal random variables. Let $\mathbf{C}_{\boldsymbol{\gamma}}$ denote the correlation matrix of the complex path weights. Then, a covariance matrix system representation is

$$\mathbf{C}_{\hat{\mathbf{h}}} = \mathbf{G}(\boldsymbol{\tau})\mathbf{C}_{\boldsymbol{\gamma}}\mathbf{G}^T(\boldsymbol{\tau}) + \mathbf{C}_{\mathbf{w}}, \quad \text{with} \tag{25}$$

$$\mathbf{C}_{\hat{\mathbf{h}}} = \frac{1}{N_r I}\hat{\mathbf{H}}\hat{\mathbf{H}}^H, \tag{26}$$

with the so-called sample correlation matrix $\mathbf{C}_{\hat{\mathbf{h}}}$ and error correlation matrix $\mathbf{C}_{\mathbf{w}}$ ($\sim \ \sigma_{\mathbf{w}}\mathbf{I}$). Let $\boldsymbol{\eta}$ be an auxiliary vector harbouring only the lower diagonal elements of the matrix $\mathbf{C}_{\boldsymbol{\gamma}}$, taken column wise from the matrix. Now the underlying overall parametrization of (26), including both delays and covariance matrix, are given by

$$\boldsymbol{\theta}_{\text{sml}} = [\boldsymbol{\tau}, \text{diag}\{\mathbf{C}_{\boldsymbol{\gamma}}\}, \text{Re}\{\boldsymbol{\eta}\}, \text{Im}\{\boldsymbol{\eta}\}, \sigma_w^2] \ \in \mathbb{R}^{1\times(C^2+C+1)}. \tag{27}$$

Since it is reasonable to assume that $2N_r I > C$ comparing the vector lengths of $\boldsymbol{\theta}_{\text{dml}}$ and $\boldsymbol{\theta}_{\text{sml}}$, it becomes clear that the overall stochastic formulation of the estimation problem yields less overall unknowns than the deterministic problem formulation. Nonetheless, there also are reasons to stick to the more complex deterministic maximum likelihood. Consider that, if we treat the parameters as deterministic it means that we use less a priori information since we exploit no information on any underlying distribution. This proves to be an advantage in cases, where the underlying distribution varies, cannot be specified, or is too complicated. Acknowledging that the number of unknowns is huge, when choosing a deterministic modelling framework at the receiver side, we are on the safe side, in that respect that we can cover the estimation of a broader model range. For this contribution, we are primarily interested in estimating the multipath delays and since we modelled them as static variables over I consecutive measurements, the number of parameters C for the desired vector $\hat{\boldsymbol{\tau}}$ is the same for the stochastic as for the deterministic model as can be seen from (16) and (27).

For the deterministic model the cost function in (23) is specifically determined by the deterministic maximum likelihood (DML) estimator and for the stochastic model it is determined by the concentrated form [49] of the stochastic maximum likelihood (SML) estimator [50]:

$$\Omega_{\text{sml}} = \text{trace}\{\mathbf{P}_{\mathbf{G}}^{\perp}(\tilde{\boldsymbol{\tau}})\mathbf{C}_{\hat{\mathbf{h}}}\}. \tag{28}$$

3) DISCUSSION ON A SUBOPTIMAL FREQUENCY-DOMAIN BASED CLOSED-FORM SOLUTION

Other alternative approaches to solve the underlying estimation problem here can be applied here and therefore we shortly want to discuss them here.

The first other approach to parameter estimation bases on examining the channel coefficients in the frequency domain, yielding two advantages. The frequency domain channel coefficients can simply be divided by the DFT of sampled $g(\tau)$, to obtain a deconvolved time-domain signal. Secondly, the frequency-domain deconvolved signal exhibits an underlying rotational invariance among its subspaces. We can exploit this invariance to find a set of linear equations, which, in return, yield a closed-form search-free solution for the delay estimates, called estimation of signal parameters via rotational invariance (ESPRIT) [24], [25]. Another advantage of the algorithm is that no initial guess for optimization is required to solve the problem. Therefore, it is commonly thought of as a practical method in the initial delay acquisition phase before initiate tracking. Moreover, this method is especially practical for signal models employing orthogonal frequency division multiplexing (OFDM) like in [34], since the signal model already is inherently tailored to frequency domain processing. Hence, it is essential to point out that the ESPRIT algorithm is in principle desirable for the delay estimation in a JCAP framework.

Similarly, as explained in detail in [27], if we apply ESPRIT for delay estimation to the underlying signal model, we require oversampling. Oversampling yields a better model approximation and therefore can improve the systematic delay estimation errors occurring due to the underlying model mismatch that is present in case the delays are not integer multiples of $1/J$ (which they are naturally not) if we say that J is the oversampling factor. Moreover, the Fourier-transformed data model requires a discrete Fourier transform of the sampled version of $g(\tau)$, which introduces aliasing leading to a delay estimation bias. Contrary to the optimal estimator in (23) the ESPRIT based estimator is a sub-optimal approach. Although the ESPRIT delay estimation method is a practical and in many scenarios especially suitable delay estimation tool for joint communication and positioning, it is not as suitable as the Maximum Likelihood estimator in (23) to assess the in the following sections proposed simultaneous model selection and parameter estimation methods. For the Maximum Likelihood Estimator as can be expected we could show for the proposed signal model in previous works that at least in case the model order is always known the estimator approaches the optimal performance given by the Cramer-Rao Lower Bound (CRLB). The ESPRIT method is already known to yield a suboptimal result and hence cannot approach the CRLB tightly. This error will consequently depend on the chosen parameters. It will depend on the chosen oversampling factor for the signal model. In the numerical results, we will see that the combination of ML-based model selection and parameter estimation as well yields sub-optimal results. Knowingly combining different error sources by employing ESPRIT together with model selection would shift our contributions focus from generally investigating how optimal this combination can practically get to the question of how much the overall performance degradation would be by combining it with different optimal and suboptimal estimation techniques. The second reason for not further investigating the ESPRIT method in this framework in this contribution is the ESPRITS's algorithm's requirement of oversampling in the signal model. For the proposed joint communication and positioning system the effect of oversampling was studied and discussed in depth

in [51], [52] and the authors concluded that the achievable gain for the positioning side of the system is most pronounced for an increase from no oversampling to $J = 1$, to $J = 2$. On the other hand, for the communication side of the system, $J = 1$ already yields sufficient statistics and therefore oversampling for the communication side entails a huge complexity without improving the bit error rate performance. Consequently, if possible, we prefer to choose methods that do not require oversampling in the signal model.

Another similar frequency domain based option to determine the delays is the estimation via multiple signal classification (MUSIC) originally proposed in [29]. Similarly, as for the delay estimation with ESPRIT, the signal model for the associated delay estimation problem relies on a similar signal model and hence also on oversampling as can be understood from studying the data model provided in [30].

4) DISCUSSION ON COMPRESSIVE SENSING APPROACH

To exploit the benefits of compressive sensing to recover the physical path parameters, the authors of [57] state the problem formulation again in the frequency domain. The authors of [57] assess a joint communication and radar sensing framework based on using compressive sensing to determine the physical path parameters. The authors combine the approach with special signal processing techniques. The authors refer to these techniques as signal-stripping and clutter-reduction. With the signal stripping approach, the authors of [57], and [58] yield a problem formulation that circumvents additional communication data based errors. Such techniques aim at targeting a high physical path parameter estimation accuracy. Targeting high precision estimates is a crucial factor for all joint communication and positioning designs. Mainly, the position estimates degrade when inaccurate physical path parameter estimates feed the positioning algorithms. We explained in section II-B that semi-blind channel estimation techniques potentially perform superior due to the higher exploited virtual training percentage of the overall transmit signal energy. The authors of [47], [57] use an on-grid delay model to exploit a one-dimensional compressive sensing strategy. In this contribution, we target an optimal model to circumvent even slight modeling mismatches to simplify the discussion on the numerical results that we provide for simultaneous parameter estimation and model selection. Simultaneous estimation of the model order and the parameters will inherently entail model mismatches unequal to the correct model order. Additional model mismatches would unnecessarily complicate the discussion and interpretation.

5) OPEN QUESTIONS FOR COMBINING MODEL SELECTION AND PARAMETER ESTIMATION

If the delays have been estimated, in a next step the complex path weights can be determined via (21). More importantly, we actually used C in (3) and hence in (23) as if it would be a known number at the receiver. Unfortunately, this usually is not the case. The model order C has to be estimated.

We further investigate flexible methods to determine the optimal model order based on the instantaneous channel conditions and measurements. Some further questions arise, like: Which model order leads to the best parameter estimation performance? Which model order estimation method is most suited for this particular purpose of joint communication and positioning. More specifically, how should model selection be chosen to yield an optimal estimation performance for the positioning relevant parameter, the line-of-sight delay?

For parameter estimation, the optimality criterion is the well known Cramer-Rao Lower Bound (CRLB). Unfortunately, as we will show, for joint model selection and parameter estimation, this bound fails as a practical optimality criterion. Consequently, it is helpful to formulate an alternative lower bound to see, which is the best achievable performance and whether the proposed methods approach this bound and which joint model selection and estimation method is closest to this bound.

Furthermore, it is interesting to see whether the model order should be estimated separately from or together with the parameter estimates.

There is another reason to have a closer look at model order detection for this framework: The most popular and well-known model order selection strategies based on information-theoretic criteria mandatorily require us to fulfil a few conditions regarding the underlying signal. Hence, their applicability to this framework has to be verified.

III. MODEL SELECTION AND PARAMETER ESTIMATION

Now let us assume the model order C is unknown and hence we wish to estimate it. Consequently, we equip (15) with a hypothetical model order $\tilde{C}$ such that $\Omega_{\tilde{C}}$ is the ML cost function for this hypothetical order $\tilde{C}$. Then the delay estimate for $\tilde{C}$ then is $\hat{\boldsymbol{\tau}}_{\tilde{C}}$ with

$$\hat{\boldsymbol{\tau}}_{\tilde{C}} = \arg\min_{\tilde{\boldsymbol{\tau}}_{\tilde{C}}} \{\Omega_{\tilde{C}}\}. \tag{29}$$

whereas $\Omega_{\tilde{C}}$ monotonically decreases, the delay estimation error monotonically increases with an increasing model order $\tilde{C}$. To find the optimal $\hat{\tau}_{\hat{C}}$, we seek a method that smartly balances between minimizing the lack of fit (29) and minimizing the estimation error.

There are threshold-based model section strategies as well as strategies based on information-theoretic criteria. The main drawback of threshold-based strategies is that a tuning parameter has to be introduced and adjusted. On the other hand, the main drawbacks of methods, based on the principles of information-theory often are twofold: Firstly, the underlying signal measurements and the model have to fulfil special requirements, e.g., the measurements should be identically and independently (iid) distributed. Secondly, their derivation inherently requires a large number of available samples they consequently disqualify in the single snapshot measurement scenario. In the following, we first briefly introduce a suitable threshold-based method applicable to this system framework and afterwards we briefly review and discuss the most

popular information-theoretic criteria and their suitability for this system framework. Finally, a method using the Fisher information belonging to the parameter estimates is proposed and assessed for this system.

A. THRESHOLD AND LEAST-SQUARES ERROR DISTRIBUTION BASED METHOD

The method described in the following yields a solution that is applicable if multiple channel estimates are available and it is also applicable if only a single (snapshot) channel estimate is available. Let χ^2 denote the so-called χ^2-distribution [53]. Since the channel estimation error is Gaussian distributed, the least squares error Ω_{dml} is

$$\Omega_{\text{dml}} \sim \frac{\sigma^2}{2}\chi^2. \tag{30}$$

Further, note that the χ^2-distribution depends on the number of degrees of freedom μ. This number μ is equal to the number of all (independent and dependent) measurements minus the number of unknowns, i.e. the parameter estimation problem dimension, which is determined by the number of multipath components. At this point we are considering the deterministic maximum-likelihood scenario. The degrees of freedom μ are calculated as

$$\mu = 2(L+1)N_r I - (2N_r I\tilde{C} + \tilde{C}). \tag{31}$$

By multiplying with the factor two, we have taken into account that the $(L+1)N_r I$ measurements as well as the $N_r I\tilde{C}$ amplitudes of the unknowns in the parameter vector in (16) are complex valued. Note that in the single measurement case, $\mu = 2(L+1) - 3\tilde{C}$. From (23) we see that the estimator for a hypothetical model order $\tilde{C}$ then is

$$\hat{\boldsymbol{\tau}}_{\tilde{C}} = \arg\min_{\tilde{\boldsymbol{\tau}}_{\tilde{C}}} \{\hat{\mathbf{h}}^H \mathbf{P}_{\mathbf{G}}{}^{\perp}(\tilde{\boldsymbol{\tau}}_{\tilde{C}})\hat{\mathbf{h}}\}. \tag{32}$$

Then, the idea is to choose an order-specific threshold $\epsilon_{\tilde{C}}$ based on a tuning parameter, which for this method is a confidence level α, typically chosen close to 1 and which is chosen such that the probability $P\left(\Omega_{LS} \leq \epsilon_{\tilde{C}}\right)$

$$P\left(\Omega_{LS} \leq \epsilon_{\tilde{C}}\right) = \alpha, \tag{33}$$

like in [20]. Let the maximum hypothetical model order be denoted by $C_{\max}$. Then these values of $\epsilon_{\tilde{C}}$ for $\tilde{C} \in \{1, \ldots, C_{\max}\}$ are found by either table-lookup or via calculation by considering the inverse cumulative density functions [53, page 815]. Let Ω_{LS} denote the least squares cost function (here equivalent to the DML cost function). With a once found $\epsilon_{\tilde{C}}$ we can formulate a simultaneous model order and parameter estimator as

$$\hat{\boldsymbol{\tau}}_{\tilde{C}} = \arg\min_{\tilde{\boldsymbol{\tau}}_{\tilde{C}}} \left\{\Omega(\tilde{\boldsymbol{\tau}}_{\tilde{C}})\right\}, \tag{34}$$

$$\hat{C} = \min\left\{\tilde{C} | \frac{2}{\sigma_{\mathbf{w}}^2}\Omega_{\text{LS}}(\tilde{C}, \hat{\boldsymbol{\tau}}_{\tilde{C}}) \leq \epsilon_{\tilde{C}}\right\}. \tag{35}$$

Note that the maximum number of theoretically identifiable paths can be determined by reasonably demanding that μ exceeds zero, that is $\mu > 0$. Then

$$\tilde{C}_{\max} = \begin{cases} \left\lfloor \frac{2}{3}(L+1) \right\rfloor & \text{if } N_r I = 1, \\ \left\lfloor \frac{1}{1+\frac{1}{N_r I}}(L+1) \right\rfloor & \text{else} \end{cases} \tag{36}$$

Considering realistic channel models applicable to this problem, like for instance the WINNER models [54], the actual number of clusters typically ranges from $C \in \{8, \ldots, 20\}$, yielding a total number of $3C \in \{24, 36 \ldots, 80\}$ parameters that have to be estimated given $L+1$ observations. If for instance $L = 9$, then we would have 10 observations. According to (36), the maximum number of theoretically identifiable multipath components would be six. Hence, we would have to take into account a receiver-sided modelling error.

B. CLASSICAL INFORMATION-THEORETIC CRITERIA

The following methods require the availability of multiple channel estimates. Generally information-theoretic criteria consist of two additive terms: The negative log-likelihood $-\ln\left(\mathcal{L}(\hat{\mathbf{h}}|\boldsymbol{\theta}_{\tilde{C}})\right)$ (or a multiple thereof) and a penalty term $\mathcal{P}(\tilde{C})$, which penalizes complexity according to the special information theoretic underlying paradigm. Here, joint parameter estimation and model selection can be performed in the following manner:

$$\mathcal{C}(\hat{\boldsymbol{\theta}}_{\tilde{C}}) = -\ln\left(\mathcal{L}(\hat{\mathbf{h}}|\hat{\boldsymbol{\theta}}_{\tilde{C}})\right) + \mathcal{P}(\tilde{C}), \text{ or} \tag{37}$$

$$\hat{\boldsymbol{\theta}}_{\tilde{C}} = \arg\min_{\tilde{\boldsymbol{\theta}}_{\tilde{C}}} \left\{-\ln\left(\mathcal{L}(\hat{\mathbf{h}}|\tilde{\boldsymbol{\theta}}_{\tilde{C}})\right)\right\}, \tag{38}$$

$$\hat{C} = \arg\min_{\hat{\boldsymbol{\theta}}_{\tilde{C}}} \{\mathcal{C}(\hat{\boldsymbol{\theta}}_{\tilde{C}})\}. \tag{39}$$

First, the optimal parameter vector $\hat{\boldsymbol{\theta}}_{\tilde{C}}$ has to be found for every $\tilde{C}$ and then the optimal model order (depending on $\tilde{C}$) is chosen from the set of optimal parameter vectors $\boldsymbol{\theta}_1, \ldots \boldsymbol{\theta}_{C_{\max}}$. The maximum model order is determined via $C_{\max}$. Due to the opposing monotony of negative log-likelihood and penalty, we can compare for $\tilde{C}$ sequentially knowing that if we have found a minimizing model order, we can stop the comparison. The minimum description length (MDL) criterion [10] is known to be asymptotically consistent. Consequently overfitting, like for the less suited Akeike's information criterion (AIC) [9], [55] is no issue. The MDL criterion origins in the area of coding theory and targets the model providing the shortest description for the measurements and it is known to perform as well as the Bayesian information criterion (BIC) [11]. The MDL penalty depends on the model order k (dimensionality of $\tilde{\boldsymbol{\theta}}_{\tilde{C}}$) and on the number of independent measurements N:

$$\mathcal{C}_{\text{MDL}}(\hat{\boldsymbol{\theta}}_{\tilde{C}}) = -\ln\left(\mathcal{L}(\hat{\mathbf{h}}|\hat{\boldsymbol{\theta}}_{\tilde{C}})\right) + \frac{1}{2}k\ln N. \tag{40}$$

Unfortunately, information-theoretic criteria like the MDL require the specific assumption that the number of parameters

does not grow linearly with the number of measurements [37]. Consequently, this rules out using a deterministic description for the parameter estimation problem. This restriction, and the fact that we need multiple measurements in the first place is a disadvantage compared to the earlier introduced threshold dependent χ^2-based method. If we assume that the complex path amplitudes fade over time in such manner that they are independent and identically and complex Gaussian distributed random variables, the stochastic system description in (26) and the parametrization in (27) apply. In [12] the authors propose a model selection solution that circumvents the effort to perform the multi-dimensional global optimization $C_{\max}$-times, by choosing a different but equivalent problem formulation and parametrization based on the underlying related eigen-decomposition. The approach exploits the fact that the maximum-likelihood estimates of this alternative parametrization can be calculated in closed form. Consequently, it allows to first calculate the model order in closed form. Afterwards, the parameter estimation can be performed for the estimated model order. Let $l_i = \lambda_i,\ \forall i \in \{0,\ldots,L\}$ denote the eigenvalues of $\mathbf{C}_{\hat{\mathbf{h}}}$, sorted in descending order

$$l_0 > l_2 > \ldots > l_L.$$

Following the derivations in [12], the solution for the model order estimation can be formulated for our purpose as

$$\hat{C} = \arg\min_{\tilde{C}} \left\{ -\ln\left(\frac{\prod_{i=\tilde{C}+1}^{L+1} l_i^{1/(L+1-\tilde{C})}}{\frac{1}{L+1-\tilde{C}} \sum_{i=\tilde{C}+1}^{L+1} l_i} \right)^{(L+1-\tilde{C})N_r I} + \frac{1}{2}\tilde{C}(2(L+1)-\tilde{C})\ln(IN_r) \right\}. \tag{41}$$

If we use this method although the complex amplitudes are not independent and identically distributed complex Gaussian distributed random variables, we condone a model mismatch. Then we cannot really say how large the impact of the modeling error will be on the overall performance. Another approach in [13] is based on a subspace decompostion and was designed to cover coherent and incoherent scenarios.

C. THE INFORMATION COMPLEXITY CRITERION

The information complexity criterion (ICOMP) promises to determine the model order more judiciously [14] than the previously discussed criteria, since it uses the inverse estimated Fisher information matrix. It is designed to choose the model order yielding the lowest estimation error and not mandatorily the true model order. For the purpose of joint communication and positioning this behaviour is desirable, since the goal is a low TOA estimation error. The ICOMP can be constructed via (37) too and consists of three additive terms. The first term involves the likelihood and therefore provides a value measuring the lack of fit, whereas the second and the third additive terms constitute penalty terms employing the inverse Fisher information matrix $\mathbf{F}^{-1}_{\hat{\boldsymbol{\theta}}_{\tilde{C}}}$, that is the estimated optimal parameter covariance matrix approximated by

$$\mathbf{F}^{-1}_{\hat{\boldsymbol{\theta}}_{\tilde{C}}} \approx \frac{\sigma_{\mathrm{w}}^2}{2} \mathrm{Re}\left\{ \mathbf{J}^H_{\boldsymbol{\theta}_{\tilde{C}}} \mathbf{J}_{\boldsymbol{\theta}_{\tilde{C}}} \right\}^{-1}. \tag{42}$$

Here, $\mathbf{J}_{\boldsymbol{\theta}_{\tilde{C}}}$ denotes the Jacobian matrix. The information complexity criterion as applied here then is:

$$\mathcal{C}_{\mathrm{ICOMP}}(\hat{\boldsymbol{\theta}}_{\tilde{C}}) = -2\ln\left(\mathcal{L}(\hat{\boldsymbol{h}}|\hat{\theta}_{\tilde{C}})\right) + 2\left(\frac{k}{2}\ln\left(\frac{\mathrm{tr}\left(\mathbf{F}^{-1}_{\hat{\boldsymbol{\theta}}_{\tilde{C}}}\right)}{k} \right) - \frac{1}{2}\ln|\mathbf{F}^{-1}_{\hat{\boldsymbol{\theta}}_{\tilde{C}}}| \right). \tag{43}$$

With $k = 2N_r I\tilde{C} + \tilde{C}$ and the eigenvalues $\lambda_1,\ldots,\lambda_k$ of $\mathbf{F}^{-1}_{\hat{\boldsymbol{\theta}}_{\tilde{C}}}$ (43) can be compactly written as

$$\mathcal{C}_{\mathrm{ICOMP}}(\hat{\boldsymbol{\theta}}_{\tilde{C}}) = -2\mathcal{L}(\hat{\mathbf{h}}|\hat{\boldsymbol{\theta}}_{\tilde{C}}) + k\cdot\ln\left(\frac{\frac{1}{k}\sum_{j=0}^{k}\lambda_j}{\left(\prod_{j=1}^{k}\lambda_j\right)^{1/k}} \right), \tag{44}$$

$$\hat{\boldsymbol{\theta}}_{\tilde{C}} = \arg\min_{\tilde{\boldsymbol{\theta}}_{\tilde{C}}} \left\{ \ln\left(\mathcal{L}(\hat{\mathbf{h}}|\tilde{\boldsymbol{\theta}}_{\tilde{C}})\right) \right\}, \tag{45}$$

$$\hat{C} = \arg\min_{\tilde{C}} \{\mathcal{C}_{\mathrm{ICOMP}}(\hat{\boldsymbol{\theta}}_{\tilde{C}})\}. \tag{46}$$

In (43), the first of the two penalty terms is interpreted as a lack of parsimony term and the second term is interpreted as a profusion of complexity term, which takes into account the parameter estimates interdependencies.

IV. LOWER PERFORMANCE BOUNDS

Usually, when a maximum-likelihood estimator is involved, the well known Cramer-Rao lower bound (CRLB) can be determined as a lower bound for the MSEs. The CRLB is known to be an asymptotical bound that only is valid if three conditions are fulfilled: Firstly, the underlying model has to be specified correctly. Secondly, the SNR has to be high enough. Thirdly, the number of observations has to be large enough. Note that here we cannot even guarantee to fulfil one of these conditions. These conditions indicate that the CRLB alone will not tell us what we are actually interested in, which is to know which model order dependent time of arrival MSE can optimally be attained when employing trustworthy estimates $\hat{\tau}_{1,\tilde{C}}$ for all $\tilde{C} \in \{1,\ldots,C\}$. Before we formulate the optimal attainable model order and the related squared error (SE) of the TOA, we shortly introduce the actual CRLB, since we will additionally use it to explain the numerical results visually. The CRLB matrix belonging to an arbitrary parameter vector $\boldsymbol{\theta}$ is known to be the inverse Fisher information matrix. Obviously, the CRLB depends on the underlying parametrization. Note that the channel

estimation error σ_w^2 can be substituted by either the single entry channel estimation CRLB or by a further noise estimate that extends the parameter vector by another element. For the DML parametrization the CRLB matrix can asymptotically be formulated as a blockmatrix

$$\mathrm{CRLB}(\boldsymbol{\tau}) = \frac{\sigma_w^2}{2}\mathrm{Re}\{\mathbf{J_g}^T(\boldsymbol{\tau})\mathbf{P_g}^{\perp}(\boldsymbol{\tau})\mathbf{J_g}(\boldsymbol{\tau}) \odot (\boldsymbol{\Gamma}^H\boldsymbol{\Gamma})^T\}^{-1}, \tag{47}$$

where

$$[\mathbf{J_g}]_{l,c} = \frac{\partial g(lT - \tau_c)}{\partial \tau_c}. \tag{48}$$

In the single measurement case, the asymptotical result in (47) has to be modified to

$$\mathrm{CRLB}(\boldsymbol{\tau}) = \frac{\sigma_w^2}{2}\mathrm{Re}\{\mathbf{B}^H\mathbf{J_g}^T(\boldsymbol{\tau})\mathbf{P_g}^{\perp}(\boldsymbol{\tau})\mathbf{J_g}(\boldsymbol{\tau})\mathbf{B}\}^{-1}, \tag{49}$$

with

$$\mathbf{B} = \mathrm{diag}\{\boldsymbol{\gamma}\}. \tag{50}$$

Due to space limitations, for the stochastic maximum likelihood CRLB we refer to [56].

From experience we know that if we know the correct model order, at least for low orders and asymptotical conditions (high SNR and many samples), the MSE attains the CRLB. Note that, if we would assume that choosing the correct model order would lead to the best performance in terms of TOA MSE, we would mistakenly determine the CRLB for the correct model order as the best achievable TOA MSE performance, at least if the SNR is high enough. The CRLB calculation requires the correct model and hence model order.

To foresee the best achievable performance, based on the assumption that the estimates in $\hat{\boldsymbol{\tau}}_{\tilde{C}}$ are reliable and close to optimal, we formulate a focused order-related lower bound (FORLB) expression, that depends on the order-dependent focus parameter of interest $\tau_{1,\tilde{C}}$ and on the estimate $\hat{\tau}_{1,\tilde{C}}$:

$$\hat{C}_{\mathrm{opt}} = \arg\min_{\tilde{C}}\{(\tau_1 - \hat{\tau}_{1,\tilde{C}})^2\} \tag{51}$$

$$\mathrm{FORLB}_{\tau_1} = (\tau_1 - \hat{\tau}_{1,\hat{C}_{\mathrm{opt}}})^2. \tag{52}$$

In our application estimating a reduced model order occasionally yields improved parameter estimates. In such cases this focused model order-related lower bound indicates the optimal achievable performance. The FORLB requires parameter estimates that we trust to be optimal or at least close to optimal, for the different hypothetical model orders. Therefore, both bounds can be seen as an aid to optimize estimation and detection algorithms based on specific models and available information. We should not see them as the absolute lowest bound that can be achieved, since changing the estimator or the modelling or the a priori knowledge always changes the bounds as well. Any a priori knowledge would result in a beneficial estimation bias tightening the MSE to CRLB. Here we assume that we have no a priori knowledge apart from the assumption that we can define bounds for the hypothetical delays specified by a coarse estimation step that we can carry out beforehand and that yields accuracy within a sampling period T.

V. NUMERICAL RESULTS

The final goal is to assess the performance in realistic multipath scenarios, in the presence of a LOS component. Therefore, in this contribution, we use a clustered delay line model by employing the slightly modified WINNER B1-LOS (outdoor) scenario as provided in [54]. The Rician factor is $R = 3.3$. The main values for the excess delays and cluster powers are tabulated in Table 1 and are only modified in unifying the three parts of the second cluster within a single delay. We know that by utilizing semi-blind channel estimation techniques in combination with a small percentage of pilot versus a large percentage of data, we can at least asymptotically yield a channel estimation MSE performance that converges to the performance that would be reached by transmitting only pilots and no data. Hence, for this contribution, we ignore the performance degradation due to semi-blind channel estimation, and we assume that we have a virtual training matrix, which consists of training data only. Further, here we assume that a length $I = 100$ time series of blocks of length $K = 1000$ BPSK random pilot symbols are transmitted. Each symbol is assigned to a symbol duration of $T = 100$ ns. To build the EDTCM channel taps, we employed T-spaced samples of a raised-cosine function $g(\tau)$ with roll-off equal to 0.3 for $g(\tau)$. We window this function with a half window of size $3T$. We assume, if not said otherwise, $N_r = 1$. Furthermore, we model the channel as a clustered delay line model. We follow [54] to construct the complex path weights. We set the carrier frequency to 2 GHz and we set the receiver velocity to 50 km/h. For the χ^2-based and the ICOMP method we use the DML parametrization and for the MDL results we used the SML Eigen-parametrization.

TABLE 1. WINNER B1 LOS/NLOS clustered delay line model.

	Cluster #	excess delay in [ns]	cluster power in [dB]	AOA in [°]
LOS WINNER-B1 $R = 3.3$, ASA 18°	1	0	0	0
	2 in [54]	30 35 40	-10.5 -12.7 -14.5	45
	2 here	30	-7.5	45
	3	55	-14.8	63
	4	60 65 70	-13.6 -15.8 -17.6	-69
	4	60	-10.6	-69
	5	105	-13.9	61
	6	115	-17.8	-69
	7	250	-19.6	-73
	8	460	-31.4	92

In order to evaluate the performance of joint model selection and parameter estimation we compare the χ^2-based method, the MDL-based method and the ICOMP-based

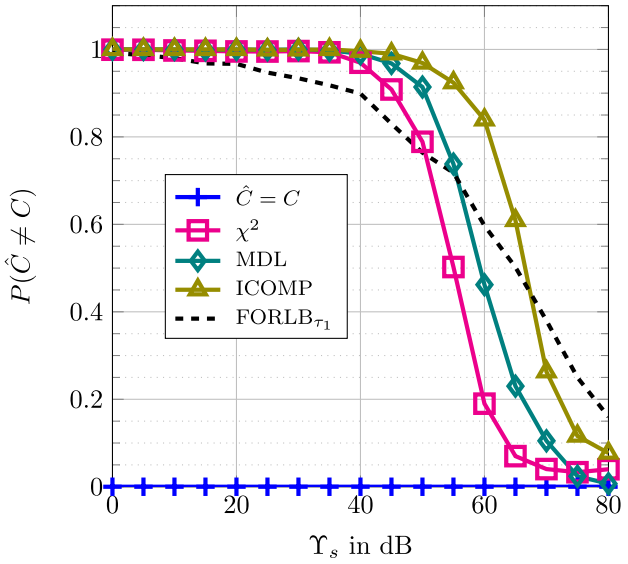

FIGURE 2. Counter intuitively, for the optimal condition, which is the FORLB, the probability of estimating the model order incorrectly, $P(\hat{C} \neq C)$, is not decreasing faster than the estimators for a multipath channel with $C = 4$ in the high signal-to-noise-ratio region.

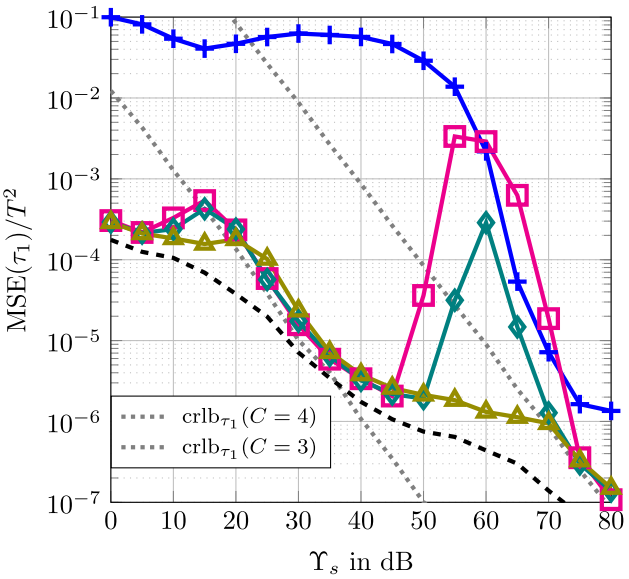

FIGURE 3. The TOA MSE performance for a multipath channel with $C = 4$ shows that the mean optimal FORLB approaches the CRLBs for different correct model orders in different SNR regions. The best performance is obtained by the ICOMP method.

method. The comparison with $\hat{C} = C$ emphasizes that estimating the model order is beneficial. As a lower performance measure, we employ both the order specific CRLBs and the FORLB. Fig. 2 shows the probability of false detection for a channel with $C = 4$ and Fig. 3 shows the MSEs for the TOA for the different algorithms, for $\hat{C} = C$ and the FORLB. Note that we calculated all TOA MSEs by using only the 98 % percentile in order to exclude bad outliers. For the probability of correct detection, two different choices of α show that choosing a value close to 1 yields a similar performance. The probability

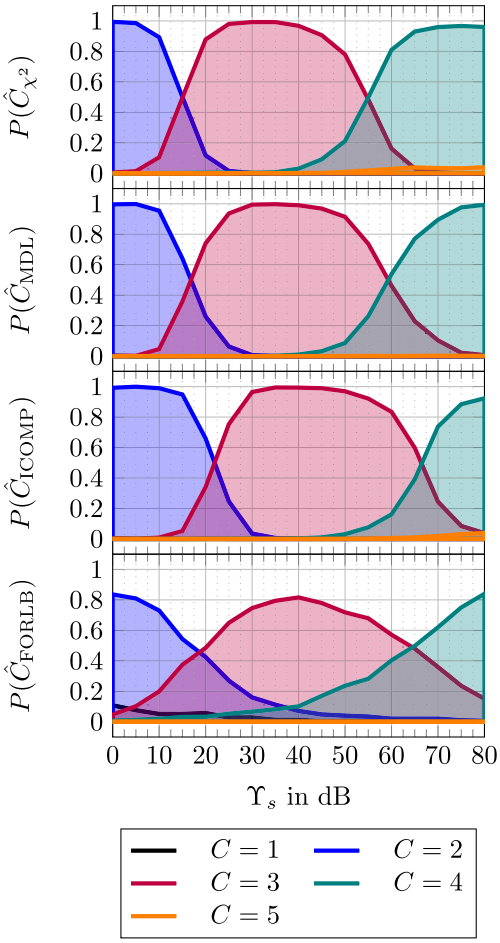

FIGURE 4. The probability distributions for $\hat{C}$ for the different model order selection methods and the FORLB show that optimal distribution for each model order is wider than the distributions obtained via algorithms.

of false detection decreases later, i.e. for higher SNR, for the ICOMP criterion-based method than for the χ^2-based method. Comparing the MSEs of those algorithms, which show the worst behaviour for correct detection, nevertheless, perform better than those, which lead to estimating the correct model order earlier. This behaviour can be understood by studying Fig. 4, which depicts the probability for estimating a hypothetical model order $\tilde{C}$, showing that the methods which lead to a lower TOA MSE shown in Fig. 3, tend to estimate a lower model order. Note that the FORLB indicates that the distribution for the hypothetical model order should generally be wider. Further, we can see that if the probability for estimating a specific hypothetical model order is very high for a specific SNR, the TOA MSE approaches the CRLB$_{\tau_1}$ for that specific model order. Note that the gap between the FORLB$_{\tau_1}$ and the MSEs can be explained by the fact that the FORLB$_{\tau_1}$ is constructed, by laying focus on the TOA, whereas the algorithms are designed to yield a compromise for all components. An optimal solution would target minimizing the squared TOA bias plus the TOA variance. Unfortunately, the TOA bias is unknown, since it depends on the actual τ_1.

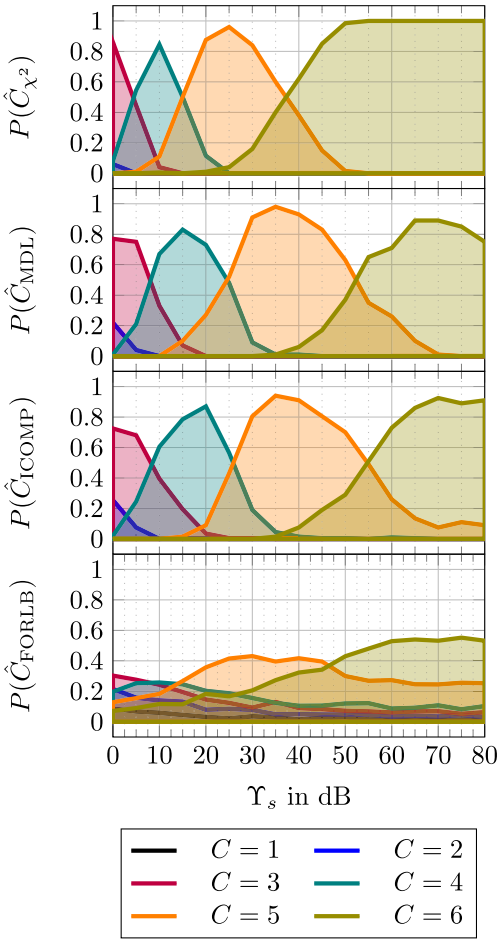

FIGURE 5. The probability distributions for $\hat{C}$ for the different model order selection methods and the FORLB show that optimal distribution for each model order is wider than the distributions obtained via algorithms.

Principally, the same behaviour as in Fig. 4 can also be seen for the realistic scenario $C = 8$ in Fig. 5. We depicted this realistic scenario and the scenario $C = 4$ via squared error distributions instead of MSEs to provide more detailed information. For this realistic scenario, we further see that the TOA SE performance degrades for the very high SNR region Fig. 6 to 11. In the high SNR region, choosing high model orders close to $C_{\max}$ becomes more probable and hence the result more often becomes unreliable. Fig. 3 shows MSEs versus SNRs: The MSEs approach different CRLBS for different SNR regions. Approaching different CRLBs is an atypical trend. The Fig. 6-8 and Fig. 9-11 provide deeper insight into the squared error distribution for the selection and estimation strategies compared to the squared error distribution for the best solution we can choose, the FORLB. The figures show split violin plots.[1] A violin plot shows the actual distribution for each point, which is

[1]A violin plot is similar to the better-known boxplot. The violin plot is more intuitive and informative since it depicts the data distribution. A boxplot, on the other hand, provides insight about only five values that are known to be significant for normally distributed data. These significant values are the median, upper and lower quartile and the user-specified whiskers and sometimes additionally outliers.

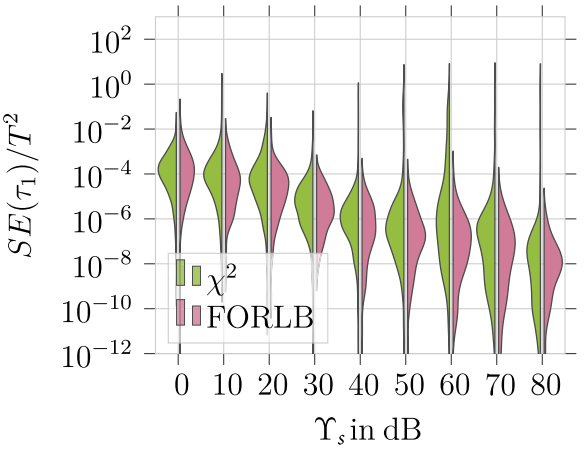

FIGURE 6. Split violin plot comparing the normalized squared error and the FORLB ($C = 4$, χ^2 based).

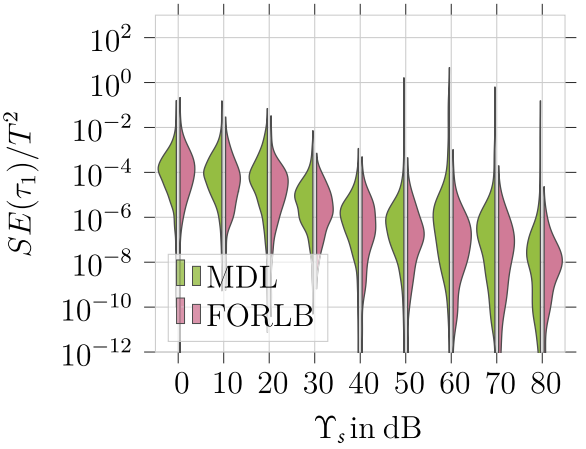

FIGURE 7. Split violin plot comparing the normalized squared error and the FORLB ($C = 4$, *MDL*).

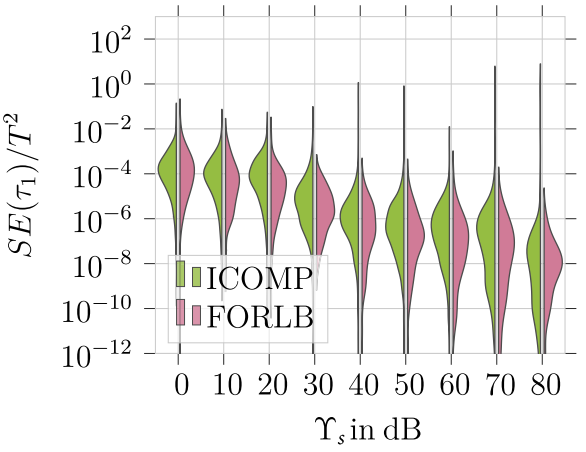

FIGURE 8. Split violin plot comparing the normalized squared error and the FORLB ($C = 4$, *ICOMP*).

reasonable if the data is not normally distributed, as in this case. Hence we can see that all three compared algorithms for $C = 4$ and $C = 8$ provide close to optimal results for a major

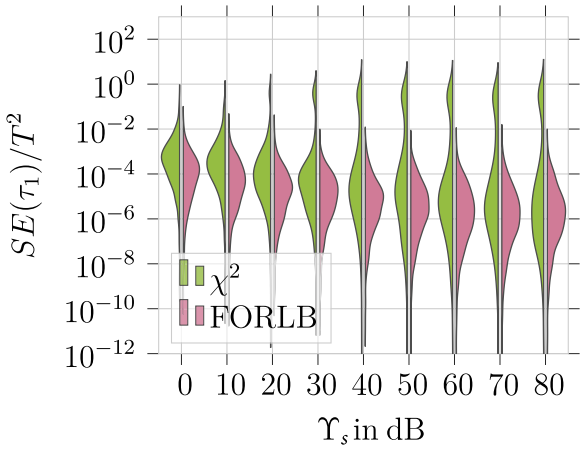

FIGURE 9. Split violin plot comparing the normalized squared error and the FORLB ($C = 8$, χ^2 based).

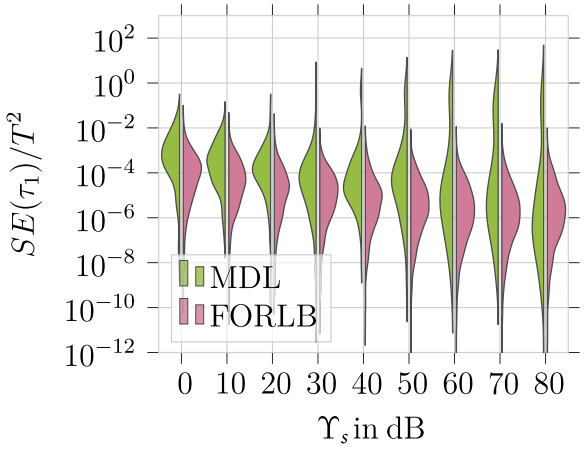

FIGURE 10. Split violin plot comparing the normalized squared error and the FORLB ($C = 8$, *MDL*).

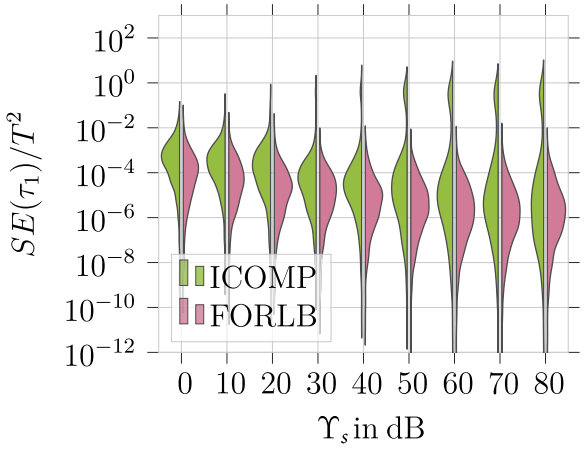

FIGURE 11. Split violin plot comparing the normalized squared error and the FORLB ($C = 8$, *ICOMP*).

part of the estimation approaches. The MSE degradation can be explained by a few non-trustworthy estimates that belong to the minority of the overall data, and it is lowest for the Fisher information based ICOMP detection method. Especially in the very high SNR region, the probability to optimally select a higher model order is higher. We pay this possibility with lower estimation reliability. We can use different mechanisms to either detect an unreliable behaviour or to enforce a lower MSE by defining different additional constraints, like for instance constraining $\hat{\sigma}^2_{\tau_1}$ to be smaller than an upper, user-defined SE limit that we should not exceed.

VI. CONCLUSION

Combining model selection and parameter estimation in a joint communication and positioning system outperforms parameter estimation approaches based on assuming either the correct or a very low or very high model order. Firstly, the calculation and comparison with an assisting focused order-related lower MSE bound, that depends on the correct and the hypothetical model order, can be used to see, which model order we should optimally choose for a particular estimator and SNR. Secondly, the comparison shows that employing the parameter and model order-dependent inverse Fisher information matrix yields a close-to-optimal approach. The approach even works in inherently problematic and ill-conditioned estimation problems, by balancing lack of fit and parsimony in favour of achieving optimal estimation accuracy. Further investigations should find solutions that are even closer to the optimal solution and optimally do not require a successive evaluation of all hypothetical solutions.

REFERENCES

[1] A. Dammann, R. Raulefs, and S. Zhang, "On prospects of positioning in 5G," in *Proc. IEEE Int. Conf. Commun. Workshop (ICCW)*, London, U.K., Jun. 2015, pp. 1207–1213.
[2] H. Dun, C. C. J. M. Tiberius, and G. J. M. Janssen, "Positioning in a multipath channel using OFDM signals with carrier phase tracking," *IEEE Access*, vol. 8, pp. 13011–13028, 2020.
[3] F. Liu, C. Masouros, A. P. Petropulu, H. Griffiths, and L. Hanzo, "Joint radar and communication design: Applications, state-of-the-art, and the road ahead," *IEEE Trans. Commun.*, vol. 68, no. 6, pp. 3834–3862, Jun. 2020.
[4] R. Adam and P. A. Hoeher, "Semi-blind channel estimation for joint communication and positioning," in *Proc. 10th Workshop Positioning, Navigat. Commun. (WPNC)*, Dresden, Germany, Mar. 2013, pp. 1–5.
[5] J. A. del Peral-Rosado, J. A. Lopez-Salcedo, F. Zanier, and G. Seco-Granados, "Position accuracy of joint time-delay and channel estimators in LTE networks," *IEEE Access*, vol. 6, pp. 25185–25199, 2018.
[6] K. Schmeink, "Joint communication and positioning based on interleave-division multiplexing," Ph.D. dissertation, Inst. Elect. Eng. Inf. Eng., Kiel Univ., Kiel, Germany, 2012.
[7] K. Shamaei and Z. M. Kassas, "LTE receiver design and multipath analysis for navigation in urban environments," *Navigation*, vol. 65, no. 4, pp. 655–675, Dec. 2018.
[8] K. V. Mishra, M. R. B. Shankar, V. Koivunen, B. Ottersten, and S. A. Vorobyov, "Toward millimeter-wave joint radar communications: A signal processing perspective," *IEEE Signal Process. Mag.*, vol. 36, no. 5, pp. 100–114, Sep. 2019.
[9] H. Akaike, "A new look at the statistical model identification," *IEEE Trans. Autom. Control*, vol. AC-19, no. 6, pp. 716–723, Dec. 1974.
[10] J. Rissanen, "Modeling by shortest data description," *Automatica*, vol. 14, no. 5, pp. 465–471, Sep. 1978.
[11] G. Schwarz, "Estimating the dimension of a model," *Ann. Statist.*, vol. 6, no. 2, pp. 461–464, Mar. 1978.
[12] M. Wax and T. Kailath, "Detection of signals by information theoretic criteria," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. ASSP-33, no. 2, pp. 387–392, Apr. 1985.

[13] M. Wax and I. Ziskind, "Detection of the number of coherent signals by the MDL principle," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 37, no. 8, pp. 1190–1196, Aug. 1989.
[14] H. Bozdogan, "Akaike's information criterion and recent developments in information complexity," *J. Math. Psychol.*, vol. 44, no. 1, pp. 62–91, Mar. 2000.
[15] K. P. Burnham and D. R. Anderson, *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach*. New York, NY, USA: Springer, 2002.
[16] J. P. C. L. da Costa, A. Thakre, F. Roemer, and M. Haardt, "Comparison of model selection techniques for high-resolution parameter estimation algorithms," in *Proc. 54th Internationales Wissenschaftliches Kolloquium. Inf. Technol. Electr. Eng.-Devices Syst.*, Illmenau, Germany, Sep. 2009, pp. 1–8.
[17] J.-M. Papy, L. De Lathauwer, and S. Van Huffel, "A shift invariance-based order-selection technique for exponential data modelling," *IEEE Signal Process. Lett.*, vol. 14, no. 7, pp. 473–476, Jul. 2007.
[18] P. M. Djuric, "Asymptotic MAP criteria for model selection," *IEEE Trans. Signal Process.*, vol. 46, no. 10, pp. 2726–2735, Oct. 1998.
[19] P. M. Djuric, "A model selection rule for sinusoids in white Gaussian noise," *IEEE Trans. Signal Process.*, vol. 44, no. 7, pp. 1744–1751, Jul. 1996.
[20] Q. Cheng and Y. Hua, "Detection of cisoids using least square error function," *IEEE Trans. Signal Process.*, vol. 45, no. 6, pp. 1584–1590, Jun. 1997.
[21] E. Radoi and A. Quinquis, "A new method for estimating the number of harmonic components in noise with application in high resolution radar," *EURASIP J. Adv. Signal Process.*, vol. 2004, no. 8, Dec. 2004, Art. no. 615890.
[22] P. Stoica and Y. Selen, "Model-order selection: A review of information criterion rules," *IEEE Signal Process. Mag.*, vol. 21, no. 4, pp. 36–47, Jul. 2004.
[23] B. H. Fleury, M. Tschudin, R. Heddergott, D. Dahlhaus, and K. I. Pedersen, "Channel parameter estimation in mobile radio environments using the SAGE algorithm," *IEEE J. Sel. Areas Commun.*, vol. 17, no. 3, pp. 434–450, Mar. 1999.
[24] R. Roy and T. Kailath, "ESPRIT-estimation of signal parameters via rotational invariance techniques," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 37, no. 7, pp. 984–995, Jul. 1989.
[25] H. Saarnisaari, "TLS-ESPRIT in a time delay estimation," in *Proc. IEEE 47th Veh. Technol. Con. Technol. Motion*, May 1997, pp. 1619–1623.
[26] P. Stoica and R. Moses, *Spectral Analysis of Signals*, 1st ed. Upper Saddle River, NJ, USA: Prentice-Hall, 2005.
[27] A. van der Veen, M. Vanderveen, and A. Paulraj, "Joint angle and delay estimation using shift invariance techniques," *IEEE Trans. Signal Process.*, vol. 46, no. 2, pp. 406–418, Feb. 1998.
[28] A. Richter, "Estimation of radio channel parameters: Models and algorithms," Ph.D. dissertation, Dept. Elect. Eng. Inf. Technol., Ilmenau Univ. Technol., Ilmenau, Germany, 2005.
[29] R. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Trans. Antennas Propag.*, vol. AP-34, no. 3, pp. 276–280, Mar. 1986.
[30] Y.-Y. Wang, J.-T. Chen, and W.-H. Fang, "TST-MUSIC for joint DOA-delay estimation," *IEEE Trans. Signal Process.*, vol. 49, no. 4, pp. 721–729, Apr. 2001.
[31] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Upper Saddle River, NJ, USA: Prentice-Hall, 2010.
[32] Y. Rong, A. A. Samuel, and D. W. Bliss, "Model order identification and parameters estimation using array processings for small sample support," in *Proc. IEEE Radar Conf. (RadarConf)*, Seattle, WA, USA, May 2017, pp. 1165–1169.
[33] J. A. del Peral-Rosado, J. A. López-Salcedo, G. Seco-Granados, F. Zanier, and M. Crisci, "Joint maximum likelihood time-delay estimation for LTE positioning in multipath channels," *EURASIP J. Adv. Signal Process.*, vol. 2014, no. 1, pp. 1–13, Mar. 2014.
[34] B. Yang, K. B. Letaief, R. S. Cheng, and Z. Cao, "Channel estimation for OFDM transmission in multipath fading channels based on parametric channel modeling," *IEEE Trans. Commun.*, vol. 49, no. 3, pp. 467–479, Mar. 2001.
[35] K. Schmeink, R. Block, C. Knievel, and P. A. Hoeher, "Joint channel and parameter estimation for combined communication and navigation using particle swarm optimization," in *Proc. 7th Workshop Positioning, Navigat. Commun.*, Dresden, Germany, Mar. 2010, pp. 4–9.
[36] Z. Dong, J. Hu, B. Du, and Y. He, "Improvement of source number estimation method for single channel signal," *PLoS ONE*, vol. 11, no. 10, Oct. 2016, Art. no. e0164654.
[37] H. Saarnisaari, "Estimation of numbers of multipaths in DS-CDMA by the MDL principle," in *Proc. IEEE 5th Int. Symp. Spread Spectr. Techn. Appl. Spread Technol. Afr.*, Sep. 1998, pp. 258–261.
[38] J. P. C. L. da Costa, F. Roemer, M. Weis, and M. Haardt, "Robust R-D parameter estimation via closed-form PARAFAC," in *Proc. Int. ITG Workshop Smart Antennas (WSA)*, Bremen, Germany, Feb. 2010, pp. 99–106.
[39] J. P. C. L. da Costa, M. Haardt, F. Romer, and G. Del Galdo, "Enhanced model order estimation using higher-order arrays," in *Proc. 41st Asilomar Conf. Signals, Syst. Comput.*, Pacific Grove, CA, USA, Nov. 2007, pp. 412–416.
[40] Z. Zhou, H. C. So, and M. G. Christensen, "Parametric modeling for damped sinusoids from multiple channels," *IEEE Trans. Signal Process.*, vol. 61, no. 15, pp. 3895–3907, Aug. 2013.
[41] Z. Zhou, M. G. Christensen, J. R. Jensen, and S. Zhang, "Parametric modeling for two-dimensional harmonic signals with missing harmonics," *IEEE Access*, vol. 7, pp. 48671–48688, 2019.
[42] J.-K. Hwang and Y.-C. Chen, "A combined detection-estimation algorithm for the harmonic-retrieval problem," *Signal Process.*, vol. 30, no. 2, pp. 177–197, Jan. 1993.
[43] R. Badeau, B. David, and G. Richard, "Selecting the modeling order for the ESPRIT high resolution method: An alternative approach," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, Montreal, QC, Canada, May 2004, p. ii-1025.
[44] C.-H.-J. Ying, A. Sabharwal, and R. L. Moses, "A combined order selection and parameter estimation algorithm for undamped exponentials," *IEEE Trans. Signal Process.*, vol. 48, no. 3, pp. 693–701, Mar. 2000.
[45] D. S. Lee and N. K. K. Chia, "A particle algorithm for sequential Bayesian parameter estimation and model selection," *IEEE Trans. Signal Process.*, vol. 50, no. 2, pp. 326–336, Feb. 2002.
[46] M. Hong, M. F. Bugallo, and P. M. Djuric, "Joint model selection and parameter estimation by population Monte Carlo simulation," *IEEE J. Sel. Topics Signal Process.*, vol. 4, no. 3, pp. 526–539, Jun. 2010.
[47] C. D. Austin, R. L. Moses, J. N. Ash, and E. Ertin, "On the relation between sparse reconstruction and parameter estimation with model order selection," *IEEE J. Sel. Topics Signal Process.*, vol. 4, no. 3, pp. 560–570, Jun. 2010.
[48] K. Schmeink, R. Adam, and P. Hoeher, "Joint communication and positioning based on soft channel parameter estimation," *EURASIP J. Wireless Commun. Netw.*, vol. 2011, 2011, Art. no. 185.
[49] A. G. Jaffer, "Maximum likelihood direction finding of stochastic sources: A separable solution," in *Proc. Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, New York, NY, USA, Apr. 1988, pp. 2893–2894.
[50] B. Ottersten, M. Viberg, P. Stoica, and A. Nehorai, "Exact and large sample maximum likelihood techniques for parameter estimation and detection in array processing," in *Radar Array Processing*, S. Haykin, J. Litva, and T. J. Shephard, Eds. Berlin, Germany: Springer-Verlag, 1993, pp. 99–151.
[51] K. Schmeink, R. Adam, and P. A. Hoeher, "Performance limits of channel parameter estimation for joint communication and positioning," *EURASIP J. Adv. Signal Process.*, vol. 2012, Dec. 2012, Art. no. 178.
[52] K. Schmeink, R. Adam, and P. A. Hoeher, "Influence of oversampling on channel parameter estimation for joint communication and positioning," in *Proc. 8th Workshop Positioning, Navigat. Commun.*, Dresden, Germany, Apr. 2011, pp. 161–165.
[53] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes 3rd Edition: The Art of Scientific Computing*. New York, NY, USA: Cambridge Univ. Press, 2007.
[54] P. Kyösti, J. Meinilä, L. Hentila, X. Zhao, T. Jämsä, C. Schneider, M. Narandzi'c, M. Milojevi'c, A. Hong, J. Ylitalo, V.-M. Holappa, M. Alatossava, R. J. C. Bultitude, Y. Jong, and T. Rautiainen, "IST-4-027756 WINNER II D1.1.2 v1.2 WINNER II channel models," *Inf. Soc. Technol.*, vol. 11, Feb. 2008.
[55] H. Akaike, "Information theory and an extension of the maximum likelihood principle," in *Proc. Int. Symp. Inf. Theory*, New York, NY, USA, 1973, pp. 267–281.
[56] P. Stoica, E. G. Larsson, and A. B. Gershman, "The stochastic CRB for array processing: A textbook derivation," *IEEE Signal Process. Lett.*, vol. 8, no. 5, pp. 148–150, May 2001.
[57] M. L. Rahman, J. A. Zhang, X. Huang, Y. J. Guo, and R. W. Heath, Jr., "Framework for a perceptive mobile network using joint communication and radar sensing," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 56, no. 3, pp. 1926–1941, Jun. 2020.
[58] J. A. Zhang, X. Huang, Y. J. Guo, and M. L. Rahman, "Signal stripping based sensing parameter estimation in perceptive mobile networks," in *Proc. IEEE-APS Topical Conf. Antennas Propag. Wireless Commun. (APWC)*, Verona, Italy, Sep. 2017, pp. 67–70.

[59] H. Schoeneich and P. A. Hoeher, "Iterative pilot-layer aided channel estimation with emphasis on interleave-division multiple access systems," *EURASIP J. Adv. Signal Process.*, vol. 2006, no. 1, pp. 1–15, Dec. 2006.
[60] H. Minn and N. Al-Dhahir, "Optimal training signals for MIMO OFDM channel estimation," *IEEE Trans. Wireless Commun.*, vol. 5, no. 5, pp. 1158–1168, May 2006.
[61] M. Biguesh and A. B. Gershman, "Training-based MIMO channel estimation: A study of estimator tradeoffs and optimal training signals," *IEEE Trans. Signal Process.*, vol. 54, no. 3, pp. 884–893, Mar. 2006.
[62] A. J. Duly, T. Kim, D. J. Love, and J. V. Krogmeier, "Closed-loop beam alignment for massive MIMO channel estimation," *IEEE Commun. Lett.*, vol. 18, no. 8, pp. 1439–1442, Aug. 2014.
[63] Z. Zhou and S. Zhang, "A perturbation approach to the beam tracing in the line-of-sight massive MIMO systems," *IEEE Wireless Commun. Lett.*, vol. 9, no. 9, pp. 1529–1532, Sep. 2020.

REBECCA CARRIE ADAM was born in Hannover, Germany, in 1982. She received the Dipl.Ing. (M.Sc. equivalent) degree in electrical engineering and information technology from Kiel University, in 2009, and the Ph.D. degree from the Chair of Information and Coding Theory, Kiel University, in December 2020.

From 2009 to 2016, she was a Research Assistant with the Chair of Information and Coding Theory, Kiel University, including two years of maternity leave. From 2018 to 2019, she has worked as a Software Developer in industry. Since 2019, she has been working as a Research Assistant with the Centre for Industrial Electronics, University of Southern Denmark, Sønderborg. Her research interests include optimization, estimation theory, and machine learning applied to the areas wireless communication, joint communication and positioning, and electric vehicle fleet charging systems.

Dr. Adam was a recipient of the Best Paper Award at the 10th Workshop on Positioning, Navigation and Communication (WPNC) 2013.

PETER ADAM HOEHER (Fellow, IEEE) received the M.Sc. degree in electrical engineering from RWTH Aachen University, Aachen, Germany, in 1986, and the Ph.D. degree in electrical engineering from the University of Kaiserslautern, Kaiserslautern, Germany, in 1990.

From 1986 to 1998, he was with the German Aerospace Center (DLR), Oberpfaffenhofen, Germany. From 1991 to 1992, he was on leave at AT&T Bell Laboratories, Murray Hill, NJ. In 1998, he joined the University of Kiel, Germany, where he is currently a Full Professor of electrical and information engineering. He has been or is supervising more than 50/200/25 bachelor/master/Ph.D. students and has authored or coauthored some 280 peer-reviewed journal and conference papers being cited more than 12 000 times, wrote two textbooks, and holds about 20 patents. His research interests include communication theory and applied information theory with applications in wireless radio communications, optical wireless communications, molecular communications, underwater communications, and simultaneous wireless information and power transfer.

Dr. Hoeher has been a Fellow of the IEEE for contributions to decoding and detection that include reliability information, since 2014. He received the Hugo-Denkmeier-Award in 1990, the ITG Award in 2007, and several best paper awards. From 1999 to 2006, he has served as an Associate Editor for the IEEE TRANSACTIONS ON COMMUNICATIONS.