Super-Resolution Estimation of UWB Channels Including the Dense Component—An SBL-Inspired Approach

Stefan Grebien, Erik Leitinger, *Member, IEEE*, Klaus Witrisal, *Member, IEEE*,
and Bernard H. Fleury, *Senior Member, IEEE*

Abstract— In this paper, we present an iterative algorithm that detects and estimates the specular components (SCs) and estimates the dense component (DC) of single-input—multiple-output (SIMO) ultra-wide-band (UWB) multipath channels. Specifically, the algorithm super-resolves the SCs in the delay–angle-of-arrival domain and estimates the parameters of a parametric model of the delay-angle power spectrum characterizing the DC. Channel noise is also estimated. In essence, the algorithm solves the problem of estimating spectral lines (the SCs) in colored noise (generated by the DC and channel noise). Its design is inspired by the sparse Bayesian learning (SBL) framework. As a result the iteration process contains a threshold condition that determines whether a candidate SC shall be retained or pruned. By relying on results from extreme-value analysis the threshold of this condition is suitably adapted to ensure a prescribed probability of detecting spurious SCs. Studies using synthetic and real channel measurement data demonstrate the virtues of the algorithm: it is able to still detect and accurately estimate SCs, even when their separation in delay and angle is down to half the Rayleigh resolution limit (RRL) of the equipment; it is robust in the sense that it tends to return no more SCs than the actual ones. Finally, the algorithm is demonstrated to outperform a state-of-the-art super-resolution channel estimator in terms of robustness in the estimation of the amplitudes of specular components closely spaced in the dispersion domain.

Index Terms— Parametric channel estimation, multipath propagation, super-resolution line spectral estimation, sparse Bayesian learning (SBL), single-input—multiple-output (SIMO) radio channels, ultra-wide-band (UWB).

I. INTRODUCTION

FUTURE wireless communication technologies will support a variety of services with high quality requirements,

Stefan Grebien is with Joanneum Research Forschungsgesellschaft mbH, 8010 Graz, Austria (e-mail: stefan.grebien@joanneum.at).

Erik Leitinger and Klaus Witrisal are with the Graz University of Technology, 8010 Graz, Austria, and also with the Christian Doppler Laboratory for Location-Aware Electronic Systems, 8010 Graz, Austria (e-mail: erik.leitinger@tugraz.at; witrisal@tugraz.at).

Bernard H. Fleury is with the Institute of Telecommunications, Vienna University of Technology, 1040 Vienna, Austria (e-mail: bernard.fleury@tuwien.ac.at).

addressing performance metrics such as reliability, ultra-low latency, high data rates, and resource-efficient use of the infrastructure [1], [2]. Holistic approaches that combine different functionalities have proven to offer promising solutions to meet these requirements. Illustrative examples are integrated sensing and communications (ISAC) and radio-based simultaneous localization and mapping (SLAM) [2], [3], [4]. These examples emphasize the reliance of these technologies on comprehensive, accurate channel state information. High-performance feasible parametric multi-antenna channel estimators can provide this information.

A. State of the Art

Parametric channel models typically represent multipath propagation as a linear superposition of weighted Dirac delta distributions - or spectral lines - with distinct supports in the underlying dispersion domain (delay, angle of arrival, angle of departure, Doppler frequency, and combinations thereof). Each component in the superposition is meant to represent a specular component (SC). Note that in this paper we shall use the terms SC and spectral line indiscriminately. The finite aperture of the measurement equipment imposes some limitation on the ability to resolve SCs closely spaced in the dispersion domain.

If the number of spectral lines is known, (constrained and unconstrained) maximum-likelihood (ML) methods, see e.g. [5] or subspace-based methods [6], [7] are standard super-resolution[1] tools to estimate their parameters. Expectation-maximization and related algorithms [8], [9] have proven viable approximations of the computationally prohibitive direct implementation of the constrained ML method. These estimators have in common that they do not incorporate the estimation of the number of spectral lines into the estimation problem. Schemes that perform jointly detection of the spectral lines and estimation of their parameters have been designed within a Bayesian framework [10], [11]. Traditional methods combining detection and estimation select among multiple candidate models, each corresponding to a specific hypothesis on the number of spectral lines, the one that optimizes a so-called information criterion, such as the Akaike or Bayesian information criterion, and the minimum description length, see [12] and references therein. Yet, the

---

[1]Super-resolution is the ability of an algorithm to resolve spectral lines even if the separation of their support in the dispersion domain is below the intrinsic resolution of the measurement equipment.

information-based approach suffers from two shortcomings: (a) it is computationally intensive as the adopted information criterion needs to be computed first for each model candidate before a decision can be made; (b) the number of spectral lines of the selected model tends to be positively biased in non-asymptotic regimes of the signal-to-noise-ratio (SNR) and the number of observed samples [13]. Hence, inference schemes designed with this approach are prone to return spurious spectral lines that have no real counterpart. Alternative penalty terms have been proposed that prevent [14] or control [15] this bias.

Model-order selection is inherently realized in sparse signal reconstruction (SSR), see [16] and references therein. SSR aims at recovering a sparse weight vector in an underdetermined linear model with a known and fixed dictionary matrix. To that end it computes an estimate of the weights as the solution to a regularized optimization problem in which the regularization term is selected to promote sparse solutions. A popular instance of SSR is basis pursuit denoising [17], also called LASSO (least absolute shrinkage and selection operator) [18], that uses an ℓ_1-norm regularization. SSR can be formulated within the Bayesian framework as maximum-a-posteriori (MAP) estimation while imposing a sparsity promoting prior on the weight vector. Typically this prior is endowed with a hierarchical structure involving a hyperparameter for each weight. Several hierarchical models have been considered so far: gamma-Gaussian[2] [20], [21], Bernoulli-Gaussian [22], [23], and generalized-gamma–power-exponential [24]. This Bayesian formulation has proven to be a particularly flexible and effective tool for SSR. Since direct implementation of the estimators is typically computationally prohibitive, one has to resort to iterative schemes, often designed using variational inference methods [25], [26].

SSR can be straightforwardly applied in the context of line spectral estimation by discretizing (gridding) the dispersion domain, see e.g. [27], [28], and [29]. The benefit of doing so is that the complex optimization problem needed to estimate the supports of the spectral lines is replaced by a linear programming procedure that returns a sparse estimate of the weight vector. The shortcoming is that gridding of the dispersion domain induces spectral leakage due to the resulting model mismatch. This effect can be mitigated by selecting a denser grid, yet at the cost of increasing the coherence of the dictionary matrix, which impairs the sparse reconstruction capability and increases the computational complexity. Variants of gridding methods that employ some interpolation method [30], [31], [32], [33], [34] or apply a grid refinement technique [29], [35] have been proposed to circumvent the leakage effect.

Atomic noise minimization (ATM) provides an elegant natural means to operate with a continuous, i.e. infinite, dictionary in SSR and thereby to relax the need for discretizing the dispersion domain [36], [37], [38]. However, some specificities of our underlying model — namely a two-dimensional (dim.)

dispersion domain and unknown colored noise — prevent a direct application of the method, see Subsection IV-A and a related discussion in [38]. Moreover, numerical evidence shows that ATM requires the supports of spectral lines to be sufficiently separated in the dispersion domain in order to be able to recover them [37]. In [39] an alternative is proposed that circumvents this shortcoming.

In theory, gridding-based line spectral estimation methods can be straightforwardly extended to account for continuous dispersion parameters by relaxing the discretization constraint and instead including the estimation of the support of the spectral lines in the inference process. Clearly, this approach is an instance of SSR with learning the continuous (vector-valued) parameter of a parameterized dictionary matrix. It has been extensively pursued in connection with the Bayesian formulation of SSR [40], [41], [42], [43], [44], [45], [46]. These algorithms differ in their specific design criteria, such as (i) the chosen sparsity-inducing hierarchical prior model, e.g. gamma-Gaussian [40], [41], [42], [43], [44], Bernoulli-Gaussian [45], [46], (ii) the assumed absence [40], [46] or presence [41], [44] of correlation among the weights of the spectral lines, and (iii) whether point estimates [40], [42], [43], [44], [45] or posterior probability density functions (PDFs) of the dispersion parameters of the SCs are inferred [46]. Experimental evidence shows that the algorithms computing point estimates of the supports of spectral lines show a positive bias in the number of detected spectral lines, i.e. are prone to detect spurious spectral lines. Including inference of the posterior PDF of the supports allows for mitigating this bias, yet at the cost of an increased computational complexity [46]. We remark that (iterative) SSR methods that apply grid refinement techniques such as [29], [35] can be viewed as particular instances of SSR methods with continuous-parameter learning, which adapt their inherent restricted range of the dictionary parameter during the iterations.[3]

The above SSR methods with continuous-parameter dictionary learning include an inherent pruning procedure that determines which ones among the columns of the dictionary matrix are inferred as relevant and switch the others off, see e.g. [47] and [48]. It is shown in [44] that the number of detected spurious spectral lines can be significantly reduced by suitably adapting the threshold of the pruning stage. The analysis provided there relies on some heuristic, yet realistic, assumptions that allow for approximating the probability of detecting a spurious line with the probability that the maximum of a continuous χ^2 random field exceeds the selected threshold [49], [50]. The analysis shows that a prescribed probability of detecting spurious lines can be guaranteed, provided the threshold increases as $C + \log n + \frac{1}{2} \log \log n$ where n is the number of observation samples and C is a constant that depends on that probability [14], [44]. Numerical analyses have shown that using this adapted threshold leads to almost vanishing bias in the number of detected SCs in medium and high SNR regimes with a tendency to underestimate said number in the low SNR regime, see also Section VII.

---

[2]The Bayesian formulation with this choice of hierarchical model is also referred to as sparse Bayesian learning or relevance vector machine to stress its link with automatic relevance determination [19] that uses a similar hierarchical model.

[3]For instance, the "gridless" SBL-based method presented in [29, Sec. IV] is similar to the methods proposed in [42] and [43].

In recent years, an extension of the channel model has been considered, that includes a DC [9], [51], [52]. The DC incorporates diffuse scattering as well as SCs that cannot be resolved with the finite aperture of the measurement equipment. Including the estimation of the DC can improve the accuracy of the estimation of the parameters of resolved SCs [9].

B. Contributions of the Paper

We propose an iterative algorithm that performs combined detection and estimation of SCs and estimation of the DC plus additive white Gaussian noise (WGN) (AWGN) in single-input—multiple-output (SIMO) ultra-wide-band (UWB) multipath channels.[4] The algorithm resolves the SCs in the delay–angle-of-arrival (angle) domain. The contributions of this paper are as follows:

- We model the impact of the DC and AWGN as colored noise, so that the problem becomes that of line spectral estimation [54] in such noise when the relative delays that the (UWB) complex envelope of the sounding wave exhibits when it is sensed by the elements of the antenna array cannot be neglected.
- The design of the algorithm is inspired by the SBL approach [20]. The probabilistic model is extended by assuming that the weights of the spectral lines are independent circularly-symmetric complex Gaussian random variables with unknown variances. In a first stage ML estimation of the variances and all other parameters but the weights is performed after integrating out said weights. These estimates are then used to compute a tractable (Gaussian) approximation of the weights' posterior PDF. The algorithm computes these two stage, the former one in an iterative fashion.
- We suitably modify the threshold inherent to the above ML estimation stage to meet a prescribed probability of detecting spurious spectral lines. To do so we apply results from extreme value analysis [49], [50].
- Using synthetically generated observation data we study in-depth the behavior of the proposed algorithm and especially how the adapted threshold affects its performance.
- We compare the performance of the proposed algorithm with a maximum-likelihood estimation (MLE) algorithm inspired by [9] but with different scheduling and adapted threshold according to [14].
- We apply the algorithm to UWB measurement data collected in an indoor environment. A simple ray-tracing tool is used to identify plausible propagation paths that can be associated to the SCs detected by the algorithm.

The remainder of the paper is organized as follows: In Section II we present the generative signal model for the considered SIMO measurement set-up. Section III describes the probabilistic signal model for inference. We derive the proposed algorithm in Section IV. Section V addresses the analytical correspondence between the probability of detecting spurious spectral lines and the threshold of the ML estimation stage. Section VII reports results from numerical and experimental studies. Concluding remarks are provided in Section VIII.

II. SIGNAL MODEL

A. Continous-Time Signal Model

The experimental measurement setup consists of an UWB transceiver operating in an indoor environment. The transmitter (Tx) is equipped with a single antenna, while an antenna array with colocated elements is emulated at the receiver (Rx) using a single antenna mounted on a positioning table.

Signals are represented by means of their complex envelope with respect to a center frequency f_c. For the sake of simplicity we assume horizontal-only propagation.[5] Under the plane-wave assumption, provided that the location $\boldsymbol{p}$ of the Rx antenna is restricted to a region $\mathcal{R}$ sufficiently confined around a reference point $\bar{\boldsymbol{p}} \in \mathcal{R}$, the antenna output signal can be expressed as

$$r(t;\boldsymbol{p}) = \iint z\big(t;\tau,\varphi,\boldsymbol{p}\big)h(\tau,\varphi)\mathrm{d}\tau\mathrm{d}\varphi + w(t;\boldsymbol{p}). \quad (1)$$

In this expression t is the time variable, while

$$z\big(t;\tau,\varphi,\boldsymbol{p}\big) = \mathrm{e}^{j2\pi f_c g(\varphi,\boldsymbol{p})}s(t - (\tau - g(\varphi,\boldsymbol{p}))) \quad (2)$$

where $s(t)$ is the transmitted signal with bandwidth B and $g\big(\varphi,\boldsymbol{p}\big) = [\cos(\varphi) \sin(\varphi)](\boldsymbol{p} - \bar{\boldsymbol{p}})/c$ with c denoting the speed of light, expresses for a plane wave incident with angle $\varphi \in [-\pi, +\pi)$ the wave's excess (propagation) delay at $\boldsymbol{p}$ relative to its delay at the reference point $\bar{\boldsymbol{p}}$. The function $h(\tau,\varphi) \in \mathbb{C}$ defined on $\mathbb{R} \times [-\pi, +\pi)$ characterizes the spread in (relative) delay τ and angle φ of the signal sensed at $\bar{\boldsymbol{p}}$. Finally, $w(t;\boldsymbol{p})$ is a spatially and temporally white Gaussian noise field with double-sided power spectral density $N_0/2$. Clearly, both $h(\tau,\varphi)$ and $z\big(t;\tau,\varphi,\boldsymbol{p}\big)$ and thereby $r(t;\boldsymbol{p})$ given by (1) also depend on the reference point $\bar{\boldsymbol{p}}$. Since this point is kept fixed, the used notation discards this dependency for the sake of shortness.

Sufficient conditions for the right-hand expression in (1) to accurately represent $r(t;\boldsymbol{p})$ are that (a) the region $\mathcal{R}$ is far away enough from the Tx and the objects in the environment with which the dominant multipath components interact along their path, such as walls, boards, etc., and (b) that the spread function $h(\tau,\varphi)$ stays constant over the bandwidth (frequency aperture) of the sounding signal. The latter assumption implies that the electromagnetic properties of said objects, like reflection and transmission coefficients, are nearly constant over the sounding bandwidth. Note that the directivity of the Rx antenna is absorbed in $h(\tau,\varphi)$.

In this study we assume that the delay-angle spread function $h(\tau,\varphi)$ is the sum of the superposition of a finite number, say K, of spectral lines representing SCs and a (spread) function $\nu(\tau,\varphi)$ describing the DC, i.e.

$$h(\tau,\varphi) = \sum_{k \in \mathcal{K}} \tilde{\alpha}_k \delta(\tau - \tilde{\tau}_k)\delta(\varphi - \tilde{\varphi}_k) + \nu(\tau,\varphi) \quad (3)$$

---

[4]The extension of the algorithm to a multiple-input—multiple-output (MIMO) system is straightforward [53].

[5]An extension to three dim. propagation scenarios including polarization is straightforward, but more involved.
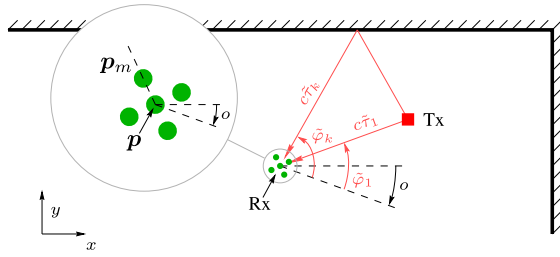
Fig. 1.   Layout of the array with its center of gravity $\bar{p}$, the mth antenna element position p_m and reference orientation o. The first SC originates from propagation along the direct path from the Tx to the Rx with angle $\tilde{\varphi}_1$ and path distance $c\tilde{\tau}_1$. The kth SC is incident with angle $\tilde{\varphi}_k$ and path distance $c\tilde{\tau}_k$.

where $\delta(\cdot)$ denotes the Dirac delta distribution. The kth SC, $k \in \mathcal{K} \triangleq \{1,\dots,K\}$ is characterized by its complex amplitude $\tilde{\alpha}_k \in \mathbb{C}$, its (relative) delay $\tilde{\tau}_k \in \mathbb{R}$ and angle (of arrival) $\tilde{\varphi}_k \in [-\pi,+\pi)$. We model $\nu(\tau,\varphi)$ as a complex circular symmetric (i.e. zero-mean) Gaussian random process [9], [55]. Furthermore, we assume uncorrelated scattering, i.e., $\mathbb{E}[\nu(\tau',\varphi')\nu^*(\tau,\varphi)] = P(\tau,\varphi)\delta(\tau'-\tau)\delta(\varphi'-\varphi)$ [56], where $\mathbb{E}[\cdot]$ denotes expectation and $P(\tau,\varphi)$ is the delay-angle power spectrum (DAPS) of the DC [56]. We make the following additional hypotheses: (a) The spread function $h(\tau,\varphi)$ has bounded support, i.e., without loss of generality, $h(\tau,\varphi) = 0$ whenever $[\tau \ \varphi] \notin [0,T) \times [-\pi,\pi) = \Psi$ with $T > 0$; (b) the equipment is designed in such a way to ensure an aliasing-free estimation of $h(\tau,\varphi)$ over Ψ. Condition (a) implies that $P(\tau,\varphi) = 0$ whenever $[\tau \ \varphi] \notin \Psi$. It also imposes that the dispersion vector $[\tilde{\tau}_k \ \tilde{\varphi}_k]$ of any kth SC, $k \in \mathcal{K}$, belongs to Ψ.

The rationale behind the selection of model (3) is as follows. The SCs originate from electromagnetic interactions with objects in the environment that are essentially non-dispersive, such as line-of-sight (LOS) propagation, specular reflection and transmission, and can be resolved with the used aperture. The DC incorporates the contributions from all other interactions, e.g. diffuse scattering and diffraction. It also includes components from specular interactions that cannot be resolved with the used aperture.

The antenna array is emulated by measuring $r(t;p)$ in (1) at M positions $p_m \in \mathcal{R}$, $m \in \{1,\dots,M\} \triangleq \mathcal{M}$, see Fig. 1. These positions specify the array layout. Without loss of generality the center of gravity of the layout is set to coincide with the reference location, i.e. $\bar{p} = M^{-1}\sum_{m=1}^{M} p_m$. The angle o depicted in the figure determines its orientation.

Inserting the decomposition (3) in (1) yields for the antenna output signal at the mth location

$$
\begin{aligned}
r_m(t) = & \sum_{k\in\mathcal{K}} \tilde{\alpha}_k z\big(t;\tilde{\tau}_k,\tilde{\varphi}_k,p_m\big) \\
& + \iint z\big(t;\tau,\varphi,p_m\big)\nu(\tau,\varphi)\mathrm{d}\tau\mathrm{d}\varphi + w_m(t). \quad (4)
\end{aligned}
$$

We have defined $r_m(t) = r(t;p_m)$ and $w_m(t) = w(t;p_m)$ for short.

B. Discrete-Frequency Signal Model

The signals $r_m(t)$, $m \in \mathcal{M}$ are Nyquist filtered, Fourier transformed, and then synchronously and uniformly sampled with frequency spacing Δ over the bandwidth B to collect for each branch m $N = B/\Delta$ samples that are arranged in a N-dim. vector y_m. These M vectors are then stacked to form the NM-vector $y = [y_1^\mathrm{T} \ \cdots \ y_M^\mathrm{T}]^\mathrm{T}$, which is expressed as

$$
y = Z(\tilde{\psi})\tilde{\alpha} + n \quad \in \mathbb{C}^{NM\times1} \quad (5)
$$

with $\tilde{\alpha} = [\tilde{\alpha}_1 \ \cdots \ \tilde{\alpha}_K]^\mathrm{T} \in \mathbb{C}^{K\times1}$, $\tilde{\psi} = [\tilde{\psi}_1 \ \cdots \ \tilde{\psi}_K] \in \Psi^K$, and $Z(\tilde{\psi}) = [z(\tilde{\psi}_1) \cdots z(\tilde{\psi}_K)] \in \mathbb{C}^{NM\times K}$ with columns given by

$$
z(\tilde{\psi}_k) = [z_1(\tilde{\psi}_k)^\mathrm{T} \ \cdots \ z_M(\tilde{\psi}_k)^\mathrm{T}]^\mathrm{T} \in \mathbb{C}^{NM\times1}, \ k \in \mathcal{K}. \quad (6)
$$

With $Z(f;\tau,\varphi,p_m)$ denoting the Fourier transform of $z(t;\tau,\varphi,p_m)$, i.e.,

$$
Z\big(f;\tau,\varphi,p_m\big) = \mathrm{e}^{j2\pi f_c g(\varphi,p_m)}S(f)\mathrm{e}^{-j2\pi f(\tau - g(\varphi,p_m))} \quad (7)
$$

where $S(f)$ is the Fourier spectrum of $s(t)$, the mth entry in the vector in (6) reads

$$
\begin{aligned}
z_m(\tilde{\psi}_k) \triangleq \big[Z\big(i\Delta;\tilde{\tau}_k,\tilde{\varphi}_k,p_m\big) : i \\
= -(N-1)/2, \ \dots \ ,(N-1)/2 \big]^\mathrm{T} \in \mathbb{C}^{N\times1}, \quad (8)
\end{aligned}
$$

i.e., it contains the Fourier-transformed samples collected at the mth antenna element. The NM-vector $n = v + w$ in (5) aggregates the vectors v and w that collect the samples (arranged in the right order) corresponding to, respectively, the integral term and the noise term in (4) when m ranges in $\mathcal{M}$. From the assumptions on the DC, v is a complex circular symmetric MN-dim. Gaussian random vector with zero mean and $MN \times MN$ covariance matrix $Q = [[Q]_{m,m'}, \ m,m' \in \mathcal{M}]$ with submatrices

$$
[Q]_{m,m'} = \iint P(\tau,\varphi)z_m(\tau,\varphi)z_{m'}(\tau,\varphi)^\mathrm{H}\mathrm{d}\tau\mathrm{d}\varphi \quad (9)
$$

with $(m,m') \in \mathcal{M}^2$. From the assumptions on the noise measurement process w is a complex circular symmetric Gaussian random vector with covariance matrix $\sigma^2 I_{NM}$ where $\sigma^2 = N_0/T_\mathrm{s}$ and $I_{(\cdot)}$ is the identity matrix of dimensions specified by the number given in the subscript. We assume that v and w are uncorrelated. As a result n is a circularly symmetric Gaussian random vector with covariance matrix

$$
R = Q + \sigma^2 I_{NM}. \quad (10)
$$

C. Selected Model for the DC

We impose some structure on the covariance matrix Q in (10) via some assumptions on the behaviour of the DC and simplifying approximations in the derivations of the submatrices in (9). This structure will ensure the feasibility of the estimation algorithm.

a) The DAPS factorizes as $P(\tau,\varphi) = P\,p(\tau)p(\varphi)$. Here, $P = \iint P(\tau,\varphi)\mathrm{d}\tau\mathrm{d}\varphi$ is the power of the DC, and $p(\tau)$ and $p(\varphi)$ are, respectively, the normalized delay power spectrum (DPS) and the normalized angle power spectrum (APS) [57].

b) In the computation of (9) we discard the second occurrence of the term $g(\varphi,p_m)$ in (7), i.e., $Z(f;\tau,\varphi,p_m) =$

$e^{j2\pi f_c g(\varphi, \boldsymbol{p}_m)} S(f) e^{-j2\pi f\tau}$. This step amounts to adopting a narrowband representation that neglects the relative delays across array elements of the modulating signals of incident waves [9], [58]. It follows from Assumptions *a)* and *b)* that the covariance matrix of $\boldsymbol{v}$ factorizes as

$$\boldsymbol{Q} = P\boldsymbol{Q}_{\mathrm{s}} \otimes \boldsymbol{Q}_{\mathrm{f}} \tag{11}$$

with $\otimes$ denoting the Kronecker product [9], [58].[6] The first factor is the spatial correlation matrix

$$\boldsymbol{Q}_{\mathrm{s}} = \int p(\varphi)\boldsymbol{a}_{\mathrm{s}}(\varphi)\boldsymbol{a}_{\mathrm{s}}(\varphi)^{\mathrm{H}} \mathrm{d}\varphi \tag{12}$$

with $\boldsymbol{a}_{\mathrm{s}}(\varphi) = [e^{-j2\pi f_c g(\varphi, \boldsymbol{p}_1)} \ \cdots \ e^{-j2\pi f_c g(\varphi, \boldsymbol{p}_m)}]^{\mathrm{T}} \in \mathbb{C}^{M \times 1}$ being the steering vector of the array, and more generally the spatial steering vector of the sounding equipment. The second factor is the delay correlation matrix

$$\boldsymbol{Q}_{\mathrm{f}} = \int p(\tau)\boldsymbol{a}_{\mathrm{f}}(\tau)\boldsymbol{a}_{\mathrm{f}}(\tau)^{\mathrm{H}} \mathrm{d}\tau \tag{13}$$

with $\boldsymbol{a}_{\mathrm{f}}(\tau) = \left[S(i\Delta)e^{j2\pi i\Delta\tau} : i = -(N-1)/2, \ \ldots \ , (N-1)/2\right]^{\mathrm{T}} \in \mathbb{C}^{N \times 1}$ denoting the frequency steering vector of the sounding equipment.

c) Experimental evidence shows that the DPS typically exhibits an exponentially decaying tail [9], [55], [59] and a smooth onset [55], [60]. This behavior is well represented by a truncated and normalized gamma PDF given by

$$\begin{aligned}p(\tau) &= p(\tau; \boldsymbol{\vartheta}) \\ &= \begin{cases} \dfrac{a}{\theta^\xi \Gamma(\xi)} (\tau - \beta)^{\xi-1} e^{-\frac{\tau-\beta}{\theta}} u(\tau - \beta), & \tau \in [0, T) \\ 0, & \text{elsewhere} \end{cases}\end{aligned} \tag{14}$$

where $u(\tau)$ is the unit step function, $\Gamma(\cdot)$ is the gamma function, and $\boldsymbol{\vartheta} = [\beta \ \theta \ \xi]$ collects the onset ($\beta > 0$), scale ($\theta > 0$), and shape ($\xi > 0$) parameters. The normalization constant $a > 0$ guarantees that $\int p(\tau; \boldsymbol{\vartheta}) \mathrm{d}\tau = 1$. The range of the parameter β is restricted in such a way to ensure that the integral of the truncated tail of the gamma PDF is negligibly small, i.e., $a(\beta) \approx 1$ for any such values of β.

d) We neglect the spatial correlation across antenna elements, i.e., we set $\boldsymbol{Q}_{\mathrm{s}} = \boldsymbol{I}_M$ [9], [61]. This choice provides a good approximation of $\boldsymbol{Q}_{\mathrm{s}}$ under the assumption of uniform APS for the antenna-element spacings used in practice.[7]

By combining Assumptions *a)–d)*, the covariance matrix $\boldsymbol{R}$ takes the following form:

$$\boldsymbol{R} \triangleq \boldsymbol{R}(\boldsymbol{\eta}) = \boldsymbol{I}_M \otimes P\boldsymbol{Q}_{\mathrm{f}} + \sigma^2 \boldsymbol{I}_{MN} \tag{15}$$

where $\boldsymbol{\eta} = [\sigma^2 \ P \ \boldsymbol{\vartheta}]$ with $\boldsymbol{\vartheta}$ defined above and $\boldsymbol{Q}_{\mathrm{f}}$ given in (13) with $p(\tau) = p(\tau; \boldsymbol{\vartheta})$ according to (14). Hence, because

---

[6]In (11) and in the following, the subscripts "s" and "f" of $\boldsymbol{a}_{\mathrm{s}}(\varphi) \in \mathbb{C}^{M \times 1}$ and $\boldsymbol{a}_{\mathrm{f}}(\tau) \in \mathbb{C}^{N \times 1}$ refer to the domain, namely space and frequency, respectively, of these apertures.

[7]This assumption is exact for a uniform linear array with half-a-wavelength element-spacing in the case of 3-dim. propagation with uniform direction dispersion. For horizontal-only propagation with uniform APS, an antenna spacing equal to approximately 40% of the wavelength leads to practically uncorrelated entries in $\boldsymbol{v}$.

of Assumption *c)*, $\boldsymbol{R}(\boldsymbol{\eta})$ is block-diagonal with M identical $N \times N$ diagonal submatrices equal to

$$\tilde{\boldsymbol{R}} \triangleq \tilde{\boldsymbol{R}}(\boldsymbol{\eta}) = P\boldsymbol{Q}_{\mathrm{f}} + \sigma^2 \boldsymbol{I}_N. \tag{16}$$

III. SPARSE BAYESIAN FORMULATION

If the number of components K of the model (5) were known, the vectors of dispersion parameters $\tilde{\boldsymbol{\psi}}$ and complex amplitudes $\tilde{\boldsymbol{\alpha}}$ of the SCs and the parameter vector $\boldsymbol{\eta}$ of colored noise could be inferred using a standard MAP or ML estimation technique. Since we can view the family $\{\boldsymbol{Z}(\tilde{\boldsymbol{\psi}})\}_{\tilde{\boldsymbol{\psi}} \in \Psi}$ as a continuous dictionary, atomic-norm methods seem at first glance to be an inference method particularly tailored to our model. However, as detailed in the discussion of Section IV, some specificities of the model prevent a direct application of these methods.

We propose an approach inspired from SBL [20] for SSR to include the estimation of the unknown K. This approach requires a two-fold modification of the generative signal model that we address below.

A. Discrete-Frequency Signal Model for Inference

In a first step the initial generative signal model (5) is modified as follows. The number of hypothetical SCs is set to a fixed number, say L. Parameter L is a design parameter that is selected large enough so that $K \leq L$. In addition, $L \ll NM$. Actually we only need that $L \leq MN$. The further restriction $\ll$ is for feasibility issues. Similarly as in Subsection II-B, we define the vector $\boldsymbol{\psi} = [\boldsymbol{\psi}_1 \ \cdots \ \boldsymbol{\psi}_L]$ with entries $\boldsymbol{\psi}_l = [\tau_l \ \varphi_l] \in \Psi$, $l \in \{1, \ldots, L\} \triangleq \mathcal{L}$. With these modifications, we arrive at the discrete-frequency signal model given by

$$\boldsymbol{y} = \boldsymbol{Z}(\boldsymbol{\psi})\boldsymbol{\alpha} + \boldsymbol{n} \quad \in \mathbb{C}^{NM \times 1} \tag{17}$$

where $\boldsymbol{\alpha} = [\alpha_1 \ \cdots \ \alpha_L]^{\mathrm{T}} \in \mathbb{C}^{L \times 1}$ and $\boldsymbol{Z}(\boldsymbol{\psi}) = [\boldsymbol{z}(\boldsymbol{\psi}_1) \ \cdots \ \boldsymbol{z}(\boldsymbol{\psi}_L)] \in \mathbb{C}^{NM \times L}$ with $\boldsymbol{z}(\boldsymbol{\psi}_l)$ defined similarly to (6). Under the assumptions made in Subsection II, the likelihood function of this model reads

$$\begin{aligned}f(\boldsymbol{y}|\boldsymbol{\psi}, \boldsymbol{\eta}, \boldsymbol{\alpha}) &= [\pi^{NM} \det(\boldsymbol{R}(\boldsymbol{\eta}))]^{-1} \\ &\quad \times e^{-(\boldsymbol{y} - \boldsymbol{Z}(\boldsymbol{\psi})\boldsymbol{\alpha})^{\mathrm{H}} \boldsymbol{R}(\boldsymbol{\eta})^{-1} (\boldsymbol{y} - \boldsymbol{Z}(\boldsymbol{\psi})\boldsymbol{\alpha})}\end{aligned} \tag{18}$$

with $\det(\cdot)$ denoting the determinant of a matrix. The second step consists in specifying a hierarchical prior for each entry α_l in form of a Gaussian scale mixture. Specifically, we define

$$f(\boldsymbol{\alpha}, \boldsymbol{\gamma}) = f(\boldsymbol{\alpha}|\boldsymbol{\gamma})f(\boldsymbol{\gamma}) = \Pi_{l \in \mathcal{L}} f(\alpha_l|\gamma_l)f(\gamma_l) \tag{19}$$

where $f(\alpha_l|\gamma_l) = \sqrt{\gamma_l/2\pi} \exp\{-\gamma_l|\alpha_l|^2/2\}$, $l \in \mathcal{L}$ and $\boldsymbol{\gamma} = [\gamma_1 \ \cdots \ \gamma_L]^{\mathrm{T}} \in \mathbb{R}_+^{L \times 1}$ with $\mathbb{R}_+ = \{r \in \mathbb{R}; r \geq 0\}$. Note that all entries in $\boldsymbol{\gamma}$ have the same prior with PDF $f(\gamma)$. These entries and their prior are referred to as hyperparameters and hyperprior, respectively. We postulate priors for the parameter vectors $\boldsymbol{\psi}$ and $\boldsymbol{\eta}$ with respective PDFs $f(\boldsymbol{\psi})$ and $f(\boldsymbol{\eta})$. With these specifications, the probabilistic model for inference reads

$$f(\boldsymbol{\psi}, \boldsymbol{\eta}, \boldsymbol{\alpha}, \boldsymbol{y}) = f(\boldsymbol{y}|\boldsymbol{\psi}, \boldsymbol{\eta}, \boldsymbol{\alpha})f(\boldsymbol{\alpha}|\boldsymbol{\gamma})f(\boldsymbol{\gamma})f(\boldsymbol{\psi})f(\boldsymbol{\eta}). \tag{20}$$

B. Inference Method

The proposed method is inspired by SBL [20]. First it computes a MAP estimate of $\boldsymbol{\psi}$, $\boldsymbol{\eta}$, and $\boldsymbol{\gamma}$ from the joint posterior of these random vectors; then it uses these estimates to infer an approximation of the posterior distribution of $\boldsymbol{\alpha}$.

The posterior PDF $f(\boldsymbol{\psi}, \boldsymbol{\eta}, \boldsymbol{\gamma}|\boldsymbol{y})$ is obtained from (20) by marginalizing out the complex amplitude vector $\boldsymbol{\alpha}$, i.e.,

$$f(\boldsymbol{\psi},\boldsymbol{\eta},\boldsymbol{\gamma}|\boldsymbol{y}) \propto \int f(\boldsymbol{y},\boldsymbol{\psi},\boldsymbol{\eta},\boldsymbol{\alpha})\mathrm{d}\boldsymbol{\alpha}$$
$$= f(\boldsymbol{y}|\boldsymbol{\psi},\boldsymbol{\eta},\boldsymbol{\gamma})f(\boldsymbol{\gamma})f(\boldsymbol{\psi})f(\boldsymbol{\eta}) \quad (21)$$

where

$$f(\boldsymbol{y}|\boldsymbol{\psi},\boldsymbol{\eta},\boldsymbol{\gamma}) \propto \det(\boldsymbol{C}(\boldsymbol{\psi},\boldsymbol{\eta},\boldsymbol{\gamma}))^{-1}\mathrm{e}^{-\boldsymbol{y}^{\mathrm{H}}\boldsymbol{C}(\boldsymbol{\psi},\boldsymbol{\eta},\boldsymbol{\gamma})^{-1}\boldsymbol{y}} \quad (22)$$

with

$$\boldsymbol{C}(\boldsymbol{\psi},\boldsymbol{\eta},\boldsymbol{\gamma}) = \boldsymbol{R}(\boldsymbol{\eta}) + \boldsymbol{Z}(\boldsymbol{\psi})\boldsymbol{\Gamma}^{-1}\boldsymbol{Z}(\boldsymbol{\psi})^{\mathrm{H}} \quad (23)$$

and $\boldsymbol{\Gamma} = \mathrm{diag}([\gamma_1 \; \cdots \; \gamma_L])$. The MAP estimates of $\boldsymbol{\psi}$, $\boldsymbol{\eta}$ and $\boldsymbol{\gamma}$ are then computed using (21). From (20) we get

$$f(\boldsymbol{\alpha}|\boldsymbol{y},\boldsymbol{\psi},\boldsymbol{\eta},\boldsymbol{\gamma}) \propto f(\boldsymbol{y}|\boldsymbol{\psi},\boldsymbol{\eta},\boldsymbol{\alpha})f(\boldsymbol{\alpha}|\boldsymbol{\gamma}), \quad (24)$$

which is readily shown to be Gaussian with mean

$$\boldsymbol{\mu} = \boldsymbol{\Sigma}\boldsymbol{Z}(\boldsymbol{\psi})^{\mathrm{H}}\boldsymbol{R}(\boldsymbol{\eta})^{-1}\boldsymbol{y} \quad (25)$$

and covariance matrix

$$\boldsymbol{\Sigma} = \left(\boldsymbol{Z}(\boldsymbol{\psi})^{\mathrm{H}}\boldsymbol{R}(\boldsymbol{\eta})^{-1}\boldsymbol{Z}(\boldsymbol{\psi}) + \boldsymbol{\Gamma}\right)^{-1}. \quad (26)$$

The approximate posterior PDF of $\boldsymbol{\alpha}$ results by plugging the MAP estimates of $\boldsymbol{\psi}$, $\boldsymbol{\eta}$ and $\boldsymbol{\gamma}$ in (24), and thus in (25) and (26).

In our design, we select non-informative improper priors for $\boldsymbol{\psi}$, $\boldsymbol{\gamma}$ and $\boldsymbol{\eta}$: $f(\boldsymbol{\psi}) \propto 1$, $f(\boldsymbol{\gamma}) \propto 1$, $f(\boldsymbol{\eta}) \propto 1$. With this selection, the above MAP estimates coincide with the ML estimates

$$(\hat{\boldsymbol{\psi}}_{\mathrm{ML}}, \hat{\boldsymbol{\eta}}_{\mathrm{ML}}, \hat{\boldsymbol{\gamma}}_{\mathrm{ML}}) = \underset{\boldsymbol{\psi},\boldsymbol{\eta},\boldsymbol{\gamma}}{\arg\max}\, f(\boldsymbol{y}|\boldsymbol{\psi},\boldsymbol{\eta},\boldsymbol{\gamma})$$
$$= \underset{\boldsymbol{\psi},\boldsymbol{\eta},\boldsymbol{\gamma}}{\arg\min}\{\log(\det(\boldsymbol{C}(\boldsymbol{\psi},\boldsymbol{\eta},\boldsymbol{\gamma})))$$
$$+ \boldsymbol{y}^{\mathrm{H}}\boldsymbol{C}(\boldsymbol{\psi},\boldsymbol{\eta},\boldsymbol{\gamma})^{-1}\boldsymbol{y}\} \quad (27)$$

and the posterior PDF of $\boldsymbol{\alpha}$ is inferred using the approximation $f(\boldsymbol{\alpha}|\boldsymbol{y},\hat{\boldsymbol{\psi}}_{\mathrm{ML}},\hat{\boldsymbol{\eta}}_{\mathrm{ML}},\hat{\boldsymbol{\gamma}}_{\mathrm{ML}})$.

IV. ITERATIVE DESIGN OF THE ESTIMATOR

Since the ML estimator in (27) cannot be calculated analytically, even though the likelihood function is given in an analytical form, and a direct numerical solution is computationally prohibitive, we resort to a sequential update of the parameter vectors $\boldsymbol{\psi}$, $\boldsymbol{\eta}$, and $\boldsymbol{\gamma}$ resulting in the estimates $\hat{\boldsymbol{\psi}}$, $\hat{\boldsymbol{\eta}}$, and $\hat{\boldsymbol{\gamma}}$.

1) Estimation of the Supports of the Spectral Lines: Inserting the current estimates $\hat{\boldsymbol{\eta}}$, and $\hat{\boldsymbol{\gamma}}$ in (27) the new estimate of $\boldsymbol{\psi}$ is computed to be

$$\hat{\boldsymbol{\psi}} = \underset{\boldsymbol{\psi}}{\arg\min}\{\log(\det(\boldsymbol{C}(\boldsymbol{\psi},\hat{\boldsymbol{\eta}},\hat{\boldsymbol{\gamma}}))) + \boldsymbol{y}^{\mathrm{H}}\boldsymbol{C}(\boldsymbol{\psi},\hat{\boldsymbol{\eta}},\hat{\boldsymbol{\gamma}})^{-1}\boldsymbol{y}\}.$$
$$(28)$$

2) Estimation of the Parameters of Colored Noise: Similarly, the new estimate of $\boldsymbol{\eta}$ is computed based on the current estimates $\hat{\boldsymbol{\psi}}$, and $\hat{\boldsymbol{\gamma}}$ to be

$$\hat{\boldsymbol{\eta}} = \underset{\boldsymbol{\eta}}{\arg\min}\{\log(\det(\boldsymbol{C}(\hat{\boldsymbol{\psi}},\boldsymbol{\eta},\hat{\boldsymbol{\gamma}}))) + \boldsymbol{y}^{\mathrm{H}}\boldsymbol{C}(\hat{\boldsymbol{\psi}},\boldsymbol{\eta},\hat{\boldsymbol{\gamma}})^{-1}\boldsymbol{y}\}.$$
$$(29)$$

3) Estimation of the Hyperparameters: Finally, given the current estimates $\hat{\boldsymbol{\psi}}$ and $\hat{\boldsymbol{\eta}}$, the new estimate of $\boldsymbol{\gamma}$ is updated according to

$$\hat{\boldsymbol{\gamma}} = \underset{\boldsymbol{\gamma}}{\arg\min}\{\log(\det(\boldsymbol{C}(\hat{\boldsymbol{\psi}},\hat{\boldsymbol{\eta}},\boldsymbol{\gamma}))) + \boldsymbol{y}^{\mathrm{H}}\boldsymbol{C}(\hat{\boldsymbol{\psi}},\hat{\boldsymbol{\eta}},\boldsymbol{\gamma})^{-1}\boldsymbol{y}\}.$$
$$(30)$$

In the sequel we consider instead of (30) a sequential method in which the estimate of each entry in $\boldsymbol{\gamma}$ is updated while the estimate of the other entries are kept fixed [47]:

$$\hat{\gamma}_l = \begin{cases} (|\rho_l|^2 - \zeta_l)^{-1}, & \dfrac{|\rho_l|^2}{\zeta_l} > \kappa \\[2mm] \infty, & \dfrac{|\rho_l|^2}{\zeta_l} \leq \kappa \end{cases} \quad l \in \mathcal{L} \quad (31)$$

with $\kappa = 1$. In this expression

$$\zeta_l = \Big(\boldsymbol{z}(\hat{\psi}_l)^{\mathrm{H}}\boldsymbol{R}(\hat{\boldsymbol{\eta}})^{-1}\boldsymbol{z}(\hat{\psi}_l) - \boldsymbol{z}(\hat{\psi}_l)^{\mathrm{H}}\boldsymbol{R}(\hat{\boldsymbol{\eta}})^{-1}$$
$$\times \boldsymbol{Z}(\hat{\boldsymbol{\psi}}_{\bar{l}})\hat{\boldsymbol{\Sigma}}_{\bar{l}}\boldsymbol{Z}(\hat{\boldsymbol{\psi}}_{\bar{l}})^{\mathrm{H}}\boldsymbol{R}(\hat{\boldsymbol{\eta}})^{-1}\boldsymbol{z}(\hat{\psi}_l)\Big)^{-1} \quad (32)$$
$$\rho_l = \zeta_l \boldsymbol{z}(\hat{\psi}_l)^{\mathrm{H}}\boldsymbol{R}(\hat{\boldsymbol{\eta}})^{-1}\boldsymbol{y}$$
$$- \zeta_l \boldsymbol{z}(\hat{\psi}_l)^{\mathrm{H}}\boldsymbol{R}(\hat{\boldsymbol{\eta}})^{-1}\boldsymbol{Z}(\hat{\boldsymbol{\psi}}_{\bar{l}})\hat{\boldsymbol{\Sigma}}_{\bar{l}}\boldsymbol{Z}(\hat{\boldsymbol{\psi}}_{\bar{l}})^{\mathrm{H}}\boldsymbol{R}(\hat{\boldsymbol{\eta}})^{-1}\boldsymbol{y}$$
$$= \zeta_l \boldsymbol{z}(\hat{\psi}_l)^{\mathrm{H}}\boldsymbol{R}(\hat{\boldsymbol{\eta}})^{-1}\bar{\boldsymbol{y}}_l \quad (33)$$

with

$$\bar{\boldsymbol{y}}_l = \boldsymbol{y} - \boldsymbol{Z}(\hat{\boldsymbol{\psi}}_{\bar{l}})\hat{\boldsymbol{\mu}}_{\bar{l}} \quad (34)$$
$$\hat{\boldsymbol{\mu}}_{\bar{l}} = \hat{\boldsymbol{\Sigma}}_{\bar{l}}\boldsymbol{Z}(\hat{\boldsymbol{\psi}}_{\bar{l}})^{H}\boldsymbol{R}(\hat{\boldsymbol{\eta}})^{-1}\boldsymbol{y} \quad (35)$$
$$\hat{\boldsymbol{\Sigma}}_{\bar{l}} = (\boldsymbol{Z}(\hat{\boldsymbol{\psi}}_{\bar{l}})^{H}\boldsymbol{R}(\hat{\boldsymbol{\eta}})^{-1}\boldsymbol{Z}(\hat{\boldsymbol{\psi}}_{\bar{l}}) + \hat{\boldsymbol{\Gamma}}_{\bar{l}})^{-1} \quad (36)$$
$$\hat{\boldsymbol{\Gamma}}_{\bar{l}} = \mathrm{diag}([\hat{\gamma}_1 \; \cdots \; \hat{\gamma}_{l-1} \; \hat{\gamma}_{l+1} \; \cdots \; \hat{\gamma}_L]) \quad (37)$$

where $\mathrm{diag}(\cdot)$ describes a square diagonal matrix with the elements of the vector given as an argument on the main diagonal and $\hat{\boldsymbol{\psi}}_{\bar{l}} = [\hat{\psi}_1 \; \cdots \; \hat{\psi}_{l-1} \; \hat{\psi}_{l+1} \; \cdots \; \hat{\psi}_L]$.

Note that the computation step of $\hat{\gamma}_l$ (31) contains a condition that determines when the lth spectral line shall be discarded ($\hat{\gamma}_l = \infty$).

4) Estimation of the Weights: Inserting the estimates $\hat{\boldsymbol{\psi}}$, $\hat{\boldsymbol{\eta}}$, and $\hat{\boldsymbol{\gamma}}$ in (24) yields the Gaussian PDF with mean (see (25))

$$\hat{\boldsymbol{\mu}} = \hat{\boldsymbol{\Sigma}}\boldsymbol{Z}(\hat{\boldsymbol{\psi}})^{\mathrm{H}}\boldsymbol{R}(\hat{\boldsymbol{\eta}})^{-1}\boldsymbol{y} \quad (38)$$

and covariance matrix (see (26))

$$\hat{\boldsymbol{\Sigma}} = \left(\boldsymbol{Z}(\hat{\boldsymbol{\psi}})^{\mathrm{H}}\boldsymbol{R}(\hat{\boldsymbol{\eta}})^{-1}\boldsymbol{Z}(\hat{\boldsymbol{\psi}}) + \hat{\boldsymbol{\Gamma}}\right)^{-1} \quad (39)$$

that is used as an approximation of the posterior PDF of $\boldsymbol{\alpha}$. In (39), $\hat{\boldsymbol{\Gamma}} = \mathrm{diag}([\hat{\gamma}_1 \; \cdots \; \hat{\gamma}_L])$.

5) Fitting of the Pruning Threshold κ: Numerical experiments have shown that the iterative algorithm obtained in the above subsections overestimates the number of spectral lines and thereby returns estimates of spurious components.

This bias in the number of detected components increases when either the SNR or the number of samples increases [40, Subsec. V.A], [44]. Following the approach adopted in [14] and [62], we increase the initial threshold $\kappa = 1$ in the pruning condition (31) to $\kappa = \kappa^* > 1$. The value κ^* is set in such a way to reduce the bias. The next section describes in detail this procedure, which yields κ^* given in (48).

A. Discussion

The updating step (28) in its form looks very similar to the classical unconstrained (also called stochastic) ML estimator in sensor array signal processing [5] with the additional assumption that the precision matrix of the weights be diagonal, i.e. equal to $\mathbf{\Gamma}$ as a result of the gamma-Gaussian hierarchical model.[8] Despite the resemblance (28) is not an instance of unconstrained ML estimation. Unconstrained ML estimation requires a scenario where at least as many observations (snapshots, assumed uncorrelated) as the number of sensors are collected,[9] while in our scenario only one observation, i.e. $\mathbf{y}$ in (17), is available. Our estimator also deviates from being an instance of SBL [20] in three respects: (a) the underlying model of SBL is undetermined, which is not the case for our model (17) with $L \ll NM$; (b) the "dictionary matrix", namely $\mathbf{S}(\boldsymbol{\psi})$ in (17), is not fixed but is parameterized by the continuous parameter vector $\boldsymbol{\psi}$ that is estimated; and (c) the inherent threshold of SBL is adapted to control the probability of detecting spurious SCs.[10] Our method belongs to the class of parametric estimators in the nomenclature introduced in [27].

Atomic noise minimization (ATM) provides an elegant, natural means to operate with a continuous, i.e. infinite, dictionary in SSR [36], [37]. At first glance this method looks promising for dealing with the continuous dictionary $\{\mathbf{S}(\boldsymbol{\psi})\}_{\boldsymbol{\psi} \in \Psi}$ in our problem at hand. However, some specificities of the generic model (5) prevent its straightforward application to our scenario. Note that ATM primarily "denoises" the observed signal with the estimation of the spectral lines being subsequently performed based on this denoised signal. While the estimation problem can be solved with an exact semi-definite program when the dispersion domain is one-dim. [36], only an approximate such program could be formulated to date for higher dim. dispersion domains [38].[11] In addition, ATM operates on Nyquist-sampled signals and requires knowledge of the noise characteristics, e.g. its spectral height when noise

is white. These conditions do not hold in our application scenario.

V. COMPUTATION OF THE PRUNING THRESHOLD

To compute the threshold value κ^* we adapt the approach described in [44] to our application scenario; see also [14] for a similar approach applied to constrained ML estimation. To make it tractable the analysis is carried out under the following assumptions.

Assumption 1: The spatial and frequency apertures [64] of the sounding equipment are centro-symmetric[12] [65]. Furthermore, $\mathbf{a}_{\mathrm{f}} = \mathbf{a}_{\mathrm{f}}(0)$, see text below (13), fulfils $\mathbf{J}\mathbf{a}_{\mathrm{f}} = \mathbf{a}_{\mathrm{f}}^*$, where $\mathbf{J}$ is the exchange or reversal matrix [54, Sec. 4.8]. The covariance matrix $\mathbf{R}$ in (15) is known.

It is shown in [53] that as a result of the first part in the assumption the matrix $\mathbf{Q}$ in (11) is centro-hermitian[13] [65] and therefore $\mathbf{R}$ in (15) too. The next assumption reflects an empirical evidence based on extensive simulations of the proposed algorithm.

Assumption 2: Asymptotically as the dimension MN grows large the estimator in Section IV with $\kappa = 1$ exhibits the following behaviour: (a) it resolves all K active SCs and accurately estimates their parameters, i.e. without loss of generality $\hat{\boldsymbol{\psi}}_l \approx \tilde{\boldsymbol{\psi}}_l$ for $l = 1, \ldots, K$; (b) it computes estimates $\hat{\boldsymbol{\psi}}_l$, $l = K+1, \ldots, L$ of $L - K$ (spurious) SC components in such a way that with high probability $\mathbf{z}(\hat{\boldsymbol{\psi}}_l)$ is nearly orthogonal to any columns of $\mathbf{Z}(\hat{\boldsymbol{\psi}}_{\bar{l}})$ for each $l = K+1, \ldots, L$.

As a result of Assumption 2, as MN grows large, with high probability (32) and (33) can be approximated for $l = K+1, \ldots, L$ as $\zeta_l \approx \bar{\zeta}(\hat{\boldsymbol{\psi}}_l)$ and $\rho_l \approx \bar{\rho}(\hat{\boldsymbol{\psi}}_l)$, respectively, where we have defined $\bar{\zeta}(\boldsymbol{\psi}_{.}) = (\mathbf{z}(\boldsymbol{\psi}_{.})^{\mathrm{H}}\mathbf{R}^{-1}\mathbf{z}(\boldsymbol{\psi}_{.}))^{-1}$ and $\bar{\rho}(\boldsymbol{\psi}_{.}) = \bar{\zeta}(\boldsymbol{\psi}_{.})\mathbf{z}(\boldsymbol{\psi}_{.})^{\mathrm{H}}\mathbf{R}^{-1}\mathbf{n}$, respectively, with $\boldsymbol{\psi}_{.} = [\tau \ \varphi] \in \Psi$. Therefore, the probability that the algorithm decides that the lth component ($l = K+1, \ldots, L$) is active, i.e. $\hat{\gamma}_l < \infty$ in (31), with threshold set to κ is close to

$$P_{\mathrm{sp}}(\kappa) = \mathbb{P}\left[\sup_{\boldsymbol{\psi}_{.} \in \Psi} |\bar{\rho}(\boldsymbol{\psi}_{.})|^2/\bar{\zeta}(\boldsymbol{\psi}_{.}) \geq \kappa\right] \quad (40)$$

when MN is sufficiently large. Let us consider the circularly-symmetric complex Gaussian random field on Ψ defined as

$$x(\boldsymbol{\psi}_{.}) = \frac{\bar{\rho}(\boldsymbol{\psi}_{.})}{\bar{\zeta}(\boldsymbol{\psi}_{.})^{1/2}} = \frac{\mathbf{z}(\boldsymbol{\psi}_{.})^{\mathrm{H}}\mathbf{R}^{-1}\mathbf{n}}{[\mathbf{z}(\boldsymbol{\psi}_{.})^{\mathrm{H}}\mathbf{R}^{-1}\mathbf{z}(\boldsymbol{\psi}_{.})]^{1/2}}. \quad (41)$$

with $\boldsymbol{\psi}_{.} \in \Psi$. Then, (40) can be recast as

$$P_{\mathrm{sp}}(\kappa) = \mathbb{P}\left[\sup_{\boldsymbol{\psi}_{.} \in \Psi} |x(\boldsymbol{\psi}_{.})|^2 \geq \kappa\right]. \quad (42)$$

Theorem 1: Under Assumption 1 we have the asymptotic equivalence

$$P_{\mathrm{sp}}(\kappa) \sim \left[\frac{1}{\pi}\int_{\Psi}\sqrt{\det(\boldsymbol{\Lambda}(\boldsymbol{\psi}_{.}))}\,\mathrm{d}\boldsymbol{\psi}_{.}\right]\kappa e^{-\kappa} \ , \ \kappa \to \infty. \quad (43)$$

---

[8]Sensor array signal processing considers a signal model similar to (17) where the entries of $\mathbf{y}$ are the outputs of an array of sensors, $\mathbf{Z}(\boldsymbol{\psi})$ is the array response matrix, $\boldsymbol{\psi}$ and $\boldsymbol{\alpha}$ contain respectively the dispersion parameters and the amplitudes of the sources, and $\mathbf{n}$ is the measurement noise vector [5]. The number of sources is assumed to be known and smaller than the number of sensors in order for the model parameters to be identifiable. In practice the number of sources is estimated using an additional model-order selection procedure based on an information theoretic criterion, see Section I.

[9]This condition ensures that the sample array covariance matrix has full rank, which is a mandatory condition in the derivation of the unconstrained ML estimator.

[10]Strictly speaking, SBL is derived under the assumption of AWGN. It can be straightforwardly applied when noise is non-white, by merely applying a whitening filter first.

[11]In [38] the matrix-enhancement-matrix-pencil method [63] is used to compute estimates of the support of spectral lines from the denoised signal.

[12]Specifically, referring to Subsec. II-A for any $m \in \mathcal{M}$, there exists an index $m' \in \mathcal{M}$ such that $\mathbf{p}_{m'} - \mathbf{p} = -(\mathbf{p}_m - \mathbf{p})$. The statement for the vector defining the frequency aperture is similar.

[13]Since these matrices are hermitian, their centro-hermitian property implies per-symmetry.

Furthermore,

$$\frac{1}{\pi} \int_{\Psi} \sqrt{\det(\boldsymbol{\Lambda}(\boldsymbol{\psi}.))} \, \mathrm{d}\boldsymbol{\psi}.$$

$$= 4\pi \int_0^{\Delta^{-1}} \int_0^{2\pi} \Big[\Big(\frac{1}{M} \sum_{m \in \mathcal{M}} d_m^2(\varphi) \Big) b_1(\tau) b_2(\tau) \Big]^{1/2} f_c \, \mathrm{d}\tau \mathrm{d}\varphi. \tag{44}$$

Here, $\boldsymbol{\Lambda}(\boldsymbol{\psi}.)$ is the non-negative definite matrix given in (56) and

$$d_m(\varphi) = \partial g(\varphi, \boldsymbol{p}_m))/\partial \varphi, \quad m \in \mathcal{M} \tag{45}$$

$$b_1(\tau) = \frac{\big(\dot{\boldsymbol{a}}_{\mathrm{f}}(\tau)^H \tilde{\boldsymbol{R}}^{-1} \dot{\boldsymbol{a}}_{\mathrm{f}}(\tau) \big)}{\big(4\pi^2 \boldsymbol{a}_{\mathrm{f}}(\tau)^H \tilde{\boldsymbol{R}}^{-1} \boldsymbol{a}_{\mathrm{f}}(\tau) \big)} \tag{46}$$

$$b_2(\tau) = 1 - \frac{\Re\big\{ \dot{\boldsymbol{a}}_{\mathrm{f}}(\tau)^H \tilde{\boldsymbol{R}}^{-1} \boldsymbol{a}_{\mathrm{f}}(\tau) \big\}^2}{\big(\boldsymbol{a}_{\mathrm{f}}(\tau)^H \tilde{\boldsymbol{R}}^{-1} \boldsymbol{a}_{\mathrm{f}}(\tau) \big) \big(\dot{\boldsymbol{a}}_{\mathrm{f}}(\tau)^H \tilde{\boldsymbol{R}}^{-1} \dot{\boldsymbol{a}}_{\mathrm{f}}(\tau) \big)} \tag{47}$$

with $\tilde{\boldsymbol{R}}$ given in (16) and $\dot{\boldsymbol{a}}_{\mathrm{f}}(\tau) = \partial \boldsymbol{a}_{\mathrm{f}}(\tau)/\partial \tau$. The term $[\frac{1}{M} \sum_m d_m^2(\varphi)]^{1/2}$ incorporates the impact of the array aperture, while $b_1(\tau)$ and $b_2(\tau)$ account for the impact of the frequency aperture (the spectrum $S(f)$) and colored noise.

Proof: See Appendix A.

The function of κ in the asymptotic equivalence (43) provides a tight approximation of $P_{\mathrm{sp}}(\kappa)$ versus κ for κ sufficiently large. Thus, by taking the inverse of that function and evaluating it at a target probability value, say ϵ, we obtain a threshold, say $\kappa^\star = \kappa^\star(\epsilon)$, that yields $P_{\mathrm{sp}}(\kappa^\star)$ close to ϵ. The next lemma essentially gives this inverse function.

Lemma 1: Given $\epsilon \in (0, q/e]$ with e denoting Euler's number and $q = \int_{\Psi} \frac{1}{\pi} \sqrt{\det(\boldsymbol{\Lambda}(\boldsymbol{\psi}.))} \, \mathrm{d}\boldsymbol{\psi}.$ the asymptotic expression in (55), and thereby in (43), is upper-bounded by ϵ provided κ satisfies

$$\kappa \geq \kappa^\star(\epsilon) = -\mathrm{W}_{-1}(-\epsilon/q) \geq 1 \tag{48}$$

where $\mathrm{W}_{-1} : [-e^{-1}, 0) \mapsto \mathbb{R}$ is the second real branch of the Lambert-W function [66].

Proof: See Appendix B.

We conclude from (42), (43) and (48) that $P_{\mathrm{sp}}(\kappa^\star(\epsilon)) \approx \epsilon$ for ϵ sufficiently small and MN sufficiently large. Thus, the function $\kappa^\star(\epsilon)$ provides a means to control the probability of detecting spurious SCs provided MN is sufficiently large. We can use the following asymptotic behavior of the function $\mathrm{W}_{-1}(u)$ as $u \to 0$ to obtain a tight approximation of $\kappa^\star(\epsilon)$ for ϵ small:

$$\mathrm{W}_{-1}(u) = \log(-u) - \log(-\log(-u)) + o(1), u \to 0 \tag{49}$$

where $o(\cdot)$ denotes the little-o notation [66]. Making use of this identity, the equality in (48) can be recast as

$$\kappa^\star(\epsilon) = -\log(\epsilon/q) + \log(-\log(\epsilon/q)) + o(\epsilon), \epsilon \to 0. \tag{50}$$

The right-hand expression with the term $o(\epsilon)$ dropped gives a tight approximation of $\kappa^\star(\epsilon)$ provided ϵ is sufficiently small, and MN is sufficiently large.

A. Examples

We illustrate the right-hand expression in (44) with two examples. We consider a scenario with a sounding signal exhibiting a constant spectrum over its bandwidth, i.e., Assumption 1 is fulfilled, and AWGN only. In this case (44) becomes

$$\frac{1}{\pi} \int_{\Psi} \sqrt{\det(\boldsymbol{\Lambda}(\boldsymbol{\psi}.))} \, \mathrm{d}\boldsymbol{\psi}.$$

$$= 4\pi \sqrt{\frac{N^2 - 1}{12}} \int_0^{2\pi} f_{\mathrm{c}} \sqrt{\tfrac{1}{M} \sum_{m \in \mathcal{M}} d_m^2(\varphi)} \, \mathrm{d}\varphi. \tag{51}$$

Proof: In this case, $b_1(\tau) = 1$ and $b(\tau) = \Delta^2(N^2 - 1)/12 = (B/N)^2(N^2 - 1)/12$. $\square$

If furthermore the array is uniform, square, of dimensions $M' \times M'$, and with equal inter-element spacing $w > 0$, the above expression further simplifies to

$$\frac{1}{\pi} \int_{\Psi} \sqrt{\det(\boldsymbol{\Lambda}(\boldsymbol{\psi}.))} \, \mathrm{d}\boldsymbol{\psi}.$$

$$= 8\pi^2 \sqrt{\frac{N^2 - 1}{12}} \sqrt{f_{\mathrm{c}}^2 \frac{w^2}{c^2} \frac{M - 1}{12}} \quad \text{with} \quad M = M'^2. \tag{52}$$

Proof: The square aperture function of a rectangular uniform array is given as

$$\frac{1}{M} \sum_{m \in \mathcal{M}} d_m^2(\varphi) = M' \frac{w^2}{c^2} \sin^2(\varphi - \psi) \frac{M'(M'^2 - 1)}{12}$$

$$+ M' \frac{w^2}{c^2} \cos^2(\varphi - \psi) \frac{M'(M'^2 - 1)}{12}$$

$$= \frac{w^2}{c^2} \frac{M - 1}{12}. \tag{53}$$

$\square$

In [67] classical detection theory for detecting a single spectral line with unknown frequency embedded in AWGN is applied to obtain a threshold that is derived by analyzing the maximum of the samples of a periodogram taken on the Fourier lattice [68]. In [53] it was shown that such a classical threshold results in a positive bias in the number of detected spurious components that is avoided when the threshold given in (48) is employed. In [69], an improved threshold is proposed that reduces the positive bias in the number of spurious components too.

VI. IMPLEMENTATION

A. Algorithmic Routine

The pseudocode of the proposed algorithm is given in Algorithm 1. It has two main stages: a search and a refine procedure, described in Procedure 1 and Procedure 2, respectively. After initialization the two procedures are executed sequentially in a do-while loop until a stopping criterion is met. Specifically, Algorithm 1 implements a bottom-up strategy: starting with an empty model, i.e. $\hat{L} = 0$, at each iteration of the do-while-loop Procedure 1 searches and adds a candidate SC in the current pool of so far buffered candidate SCs. Procedure 2 estimates and/or re-estimates the parameters of all candidate SCs in the pool, and possibly removes candidate SCs to finally yield an updated pool of $\hat{L}$ candidate SCs. The algorithm terminates once the number $\hat{L}$ of SCs in the pool and
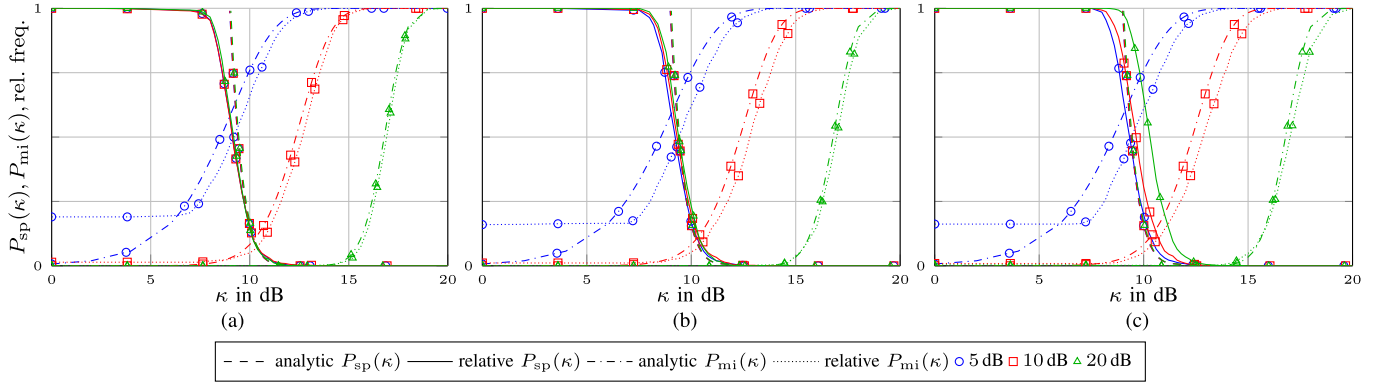
Fig. 2. Comparison of the probabilities $P_{\mathrm{sp}}(\kappa)$ and $P_{\mathrm{mi}}(\kappa)$ with their respective relative frequencies computed from 1000 trials with SNR as a parameter. The three scenarios corresponding each to specific assumptions on $\boldsymbol{Q}$ are described in the text.

Algorithm 1 Main

Input : Measurement vector $\boldsymbol{y}$
Output: $\hat{L}$, $\hat{\boldsymbol{\psi}}$, $\hat{\boldsymbol{\eta}}$, $\hat{\boldsymbol{\mu}}$, and $\hat{\boldsymbol{\Sigma}}$

1   $\hat{\boldsymbol{\psi}}, \hat{\boldsymbol{\gamma}}, \hat{\boldsymbol{\mu}}, \hat{\boldsymbol{\Sigma}} \leftarrow [\,], [\,], [\,], [\,]$
2   $\hat{L} \leftarrow 0$
3   $\hat{\boldsymbol{\eta}} \leftarrow [\hat{\sigma}^2 \;\; \hat{P} \;\; \hat{\boldsymbol{\vartheta}}]$
     with $\hat{\sigma}^2 \leftarrow \frac{\|\boldsymbol{y}\|^2}{NM}$, $\hat{P} \leftarrow 0$, $\hat{\boldsymbol{\vartheta}} \leftarrow \boldsymbol{0}$ (AWGN only), and
     initFlag $\leftarrow$ true
4   **do**
5   search()
6   refine()
7     $\hat{\boldsymbol{\mu}}, \hat{\boldsymbol{\Sigma}} \leftarrow$ (38) and (39), respectively
8   **while** $\hat{L}$ *changes or the maximum number of cycles is reached*

Procedure 2 Refine

1   **Procedure** refine()
2   **do**
3      $\hat{\boldsymbol{\eta}} \leftarrow$ update according to (29)
4      $\hat{\boldsymbol{\psi}} \leftarrow$ update according to (28)
5      $\hat{\kappa}^{\star}(\epsilon) \leftarrow$ (48)
6      $\hat{\boldsymbol{\gamma}} \leftarrow$ update according to (31), $\kappa = \hat{\kappa}^{\star}(\epsilon)$
7   initializeDC()
8   **for** $l \leftarrow 1, \ldots, \hat{L}$ **do**
9   **if** $\hat{\gamma}_l = \infty$ **then**
10   remove lth component from $\hat{\boldsymbol{\psi}}$ and $\hat{\boldsymbol{\gamma}}$
11          $\hat{L} \leftarrow \hat{L} - 1$
12   **end**
13   **end**
14   **while** *not converged*

Procedure 1 Search

1   **Procedure** search()
2   **if** $\hat{L} < L$ **then**
3      $\hat{L} \leftarrow \hat{L} + 1$
4      $l \leftarrow \hat{L}$
5      $\hat{\boldsymbol{\psi}}_l \leftarrow \underset{\boldsymbol{\psi}}{\arg \max} \frac{|\rho_l|^2}{\zeta_l}$ using (32) and (33)
6      $\hat{\gamma}_l \leftarrow (|\rho_l|^2 - \zeta_l)^{-1}$
7   append $\hat{\boldsymbol{\psi}}_l$ to $\hat{\boldsymbol{\psi}}$ and $\hat{\gamma}_l$ to $\hat{\boldsymbol{\gamma}}$
8   **end**

Procedure 3 Initialize DC parameters

1   **Procedure** initializeDC()
2   **if** *(any entry of $\hat{\boldsymbol{\gamma}}$ is ∞ or L reached) and initFlag is true* **then**
3      $\hat{P} \leftarrow \hat{\sigma}^2/(2\Delta)$
4      $\hat{\boldsymbol{\vartheta}} \leftarrow [1/c \;\; 1\mathrm{m}/c \;\; T/2 \;\; 2]$
5      $\hat{\boldsymbol{\eta}} \leftarrow [\hat{\sigma}^2/2 \;\; \hat{P} \;\; \hat{\boldsymbol{\vartheta}}]$
6   initFlag $\leftarrow$ false
7   refine()
8   **end**

---

their parameter estimates as well as the estimated parameters of the DC have converged. It then returns these converged values as the model estimates.

The initial iterations in the do-while-loop of Algorithm 1 are executed while considering measurement noise only, i.e. by using $\boldsymbol{R}$ in (15) with P set to zero whenever $\boldsymbol{R}$ occurs in the update equations of Procedures 1 and 2. This is carried out until a first candidate SC in the pool is pruned in Procedure 2 or $\hat{L}$ reaches a predefined maximum number L, in which case Procedure 3 is executed to initialize the parameters of the DC. Once the initialization is completed, the noise variance estimate $\hat{\sigma}^2$ aggregates a contribution from the DC. The total estimated power over the bandwidth computed from this value is distributed evenly between noise and the DC in Procedure 3. This explains the factor $1/2$ in Lines 3 and 5. From then on

the estimates of the parameters of the DC DPS are updated, i.e. the full covariance matrix $\boldsymbol{R}$ in (15) is accounted for in the update equations of Procedures 1 and 2.

B. Complexity Analysis

The operations that mainly affect the computational complexity of Algorithm 1 are, respectively, the updating step for the dispersion parameter vector (28) and the updating step for the noise parameter vector (29) in Procedure 2. By making use of the matrix inversion and matrix determinant lemmas and exploiting the Kronecker structure of the covariance matrix in (15), the computational complexity of the former updating step is $\mathcal{O}((NM)^2 \hat{L})$, while that of the latter is $\mathcal{O}((NM)^2 \hat{L} + N^3)$ (whether the first or the second term is dominating the other depends on the number of frequency samples N versus

TABLE I
COMMON SYSTEM PARAMETERS FOR ALL
EXPERIMENTS (IF APPLICABLE)

f_{c}	B	$\tilde{B}$	β	θ	ξ	L	inter-element spacing
Hz	GHz	GHz	ns	ns	-	-	cm
6	1.6	1	1/c	5	1.8	50	2

the number of antennas M times the number of estimated components $\hat{L}$). Note that the bottom-up approach steadily increases $\hat{L}$, starting from an empty model. Thus, the approach avoids unnecessary computation. The computational complexity of the modified MLE algorithm used as a benchmark in Section VII-A4 is $\mathcal{O}((N)^2 M \hat{L})$ for updating the dispersion parameter vector and $\mathcal{O}(N^3)$ for updating the noise parameter vector. Note that the RiMAX algorithm in [9] shares the same complexity.

VII. NUMERICAL AND EXPERIMENTAL RESULTS

To validate the proposed algorithm, we first test it in Subsection VII-A with synthetically generated measurements according to the model in (5) with a covariance matrix $\boldsymbol{R}$ given in (9) and (15). Then, in Subsection VII-B we apply the algorithm to measurements collected in an indoor environment.

A. Synthetic Radio Channels

In this study, the signal spectrum $S(f)$ has a root-raised cosine shape with roll-off factor 0.6 and bandwidth $B = 1.6\,\mathrm{GHz}$ centered at $f_{\mathrm{c}} = 6\,\mathrm{GHz}$. The 3-dB bandwidth $\tilde{B}$ is $1\,\mathrm{GHz}$ and yields the RRL $1/\tilde{B} = 1\,\mathrm{ns}$. Each numerical investigation involves 1000 simulation trials. In each trial the Gaussian DC vector $\boldsymbol{v}$ (see text below (8)) is generated using (14) with $\beta = 1\mathrm{m}/c$, $\theta = 5\,\mathrm{ns}$ and $\xi = 1.8$. The power P is specified through the SC-to-DC-ratio $\mathrm{SDR} = 10\log_{10}\left(\frac{1}{M}\|\sum_{k\in\mathcal{K}}\tilde{\alpha}_k \boldsymbol{z}(\tilde{\boldsymbol{\psi}}_k)\|^2/(PB)\right)$. In addition, the Gaussian noise vector $\boldsymbol{w}$ is generated with component variance σ^2 specified through the signal-to-noise ratio $\mathrm{SNR} = 10\log_{10}\left((\frac{1}{M}\|\sum_{k\in\mathcal{K}}\tilde{\alpha}_k \boldsymbol{z}(\tilde{\boldsymbol{\psi}}_k)\|^2 + PB)/\sigma^2\right)$. Parameter settings common to all experiments considered in the paper are reported in Table I.

1) Empirical Substantiation of the "Near-orthogonality" Assumption: This study presents empirical evidence supporting Assumption 2. We consider a synthetic channel with a single SC, i.e., (5) with $K = 1$. The settings specific to the experiment are summarized in Table II. The dispersion parameters of the SC are selected as follows: its delay is fixed to $\tilde{\tau}_1 = \tilde{\tau} = 10\,\mathrm{ns}$ and its angle $\tilde{\varphi}_1 = \tilde{\varphi}$ is drawn uniformly over $[0, 2\pi)$ for each trial and independently across trials. The respective powers of noise, the DC and the SC are set such that $\mathrm{SDR} = -5\,\mathrm{dB}$ and $\mathrm{SNR} = 20\,\mathrm{dB}$ and $\boldsymbol{R}$ is computed using (15).

To substantiate the "near-orthogonality" property claimed in Assumption 2 in each trial the cross-correlation coefficient $\boldsymbol{z}(\hat{\boldsymbol{\psi}}_j)^{\mathrm{H}}\boldsymbol{R}^{-1}\boldsymbol{z}(\hat{\boldsymbol{\psi}}_i)/\left(\boldsymbol{z}(\hat{\boldsymbol{\psi}}_i)^{\mathrm{H}}\boldsymbol{R}^{-1}\boldsymbol{z}(\hat{\boldsymbol{\psi}}_i)\boldsymbol{z}(\hat{\boldsymbol{\psi}}_j)^{\mathrm{H}}\boldsymbol{R}^{-1}\boldsymbol{z}(\hat{\boldsymbol{\psi}}_j)\right)^{\frac{1}{2}}$ is calculated for any pair (i, j) of indices of SCs detected

by the algorithm and the mean and variance of these figures are obtained. The latter quantities are averaged over the 1000 trials to yield 0.0123 and 0.0401, respectively. As a note the average number of detected SCs is 9.52. We also plotted (not reported here due to space constraints) the estimated dispersion vectors of the detected SCs in their domain Ψ. By visual inspection we could qualitatively observe that vectors located outside an elliptically shaped region centered at the dispersion vector $[\tilde{\tau}\ \tilde{\varphi}]$ of the active SC look uniformly distributed. The main axes of the boundary ellipse are set equal to 5 times the root-Cramér-Rao lower bounds (CRLBs) for the estimation of the delay and angle. This choice ensures that estimated dispersion vectors located outside the elliptical region are very unlikely (specifically with probability equal to 0.0001) to be a noisy estimate of $[\tilde{\tau}\ \tilde{\varphi}]$.

2) Detection and Estimation of a Single SC: In this study, we first validate empirically the expression in (43) as an approximation of the probability of detecting spurious SCs as well as an expression approximating the probability of not detecting an active SC that we introduce now. In a single-SC scenario, the probability of missed detection can be approximated in the asymptotic regime $NM \to \infty$ by $\mathbb{P}[|x(\hat{\boldsymbol{\psi}})|^2 < \kappa]$, where $\hat{\boldsymbol{\psi}}$ is the estimated dispersion vector of the detected SC. In this regime the distribution of $2|x(\hat{\boldsymbol{\psi}})|^2$ can be approximated by a non-central χ^2 distribution with 2 degrees of freedom and non-centrality parameter $2\eta = 2\frac{|\tilde{\alpha}|^2}{\sigma^2}\sum_{m\in\mathcal{M}}\boldsymbol{z}_m(\tilde{\boldsymbol{\psi}})\boldsymbol{R}(\boldsymbol{\eta})\boldsymbol{z}_m(\tilde{\boldsymbol{\psi}})$ [44]. Making use of this result, the probability of missed detection in a single-SC scenario is approximated by [44]

$$P_{\mathrm{mi}}(\kappa) = \int_0^\kappa \mathrm{e}^{-(x+\eta)}I_0(2\sqrt{\eta x})\mathrm{d}x. \tag{54}$$

To numerically assess the accuracy of using (42) and (54) as approximations of the probabilities of detecting spurious SCs and missing an active SC, respectively, we modify the settings as summarized in Table II. The threshold κ^* of our algorithm is a varying parameter. Other not explicitly mentioned settings stay as described in Subsection VII-A1. Note that by keeping the delay of the SC fixed the non-centrality parameter stays constant and equal to $\eta = \{8.2, 12.2, 16.7\}\,\mathrm{dB}$ corresponding to the three SNR values.

Fig. 2 shows a comparison of $P_{\mathrm{sp}}(\kappa)$ in (42) (dashed lines) and $P_{\mathrm{mi}}(\kappa)$ in (54) (dash dotted lines) with the relative frequencies of, respectively, detecting a spurious SC (solid lines) and missing the active SC (dotted lines) computed from 1000 trials. To compute the latter quantities we count the occurrence of two events that we now define. First we specify a rectangular region in Ψ centered at the dispersion vector of the active SC and with sides equal to 5 times the square root of the respective CRLBs [70]. The event "false detection" occurs if the estimated dispersion vector of at least one detected SC lies outside the region. The event "missed detection" occurs if no SC is detected or the estimated dispersion parameters of all detected SCs lie outside the region. The study is conducted under two assumptions on the covariance matrix $\boldsymbol{R}$ in (10) used in the generative model: $\boldsymbol{R}$ has the simplified form (15) (Fig. 2a and Fig. 2b) and $\boldsymbol{R}$ has the general

TABLE II
EXPERIMENT SPECIFIC SYSTEM PARAMETERS

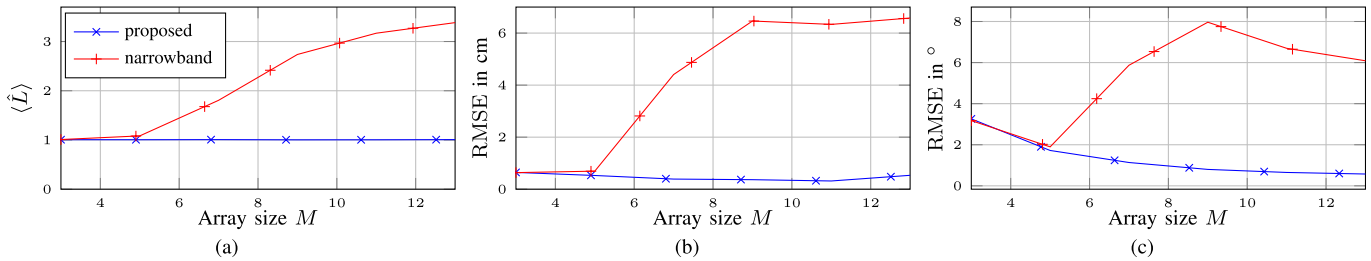|		Exp.1	Exp. 2	Exp. 3	Exp. 4	Exp. 5
		VII-A1	VII-A2	VII-A4	VII-A4	VII-B
SNR	dB	20	$\{5, 10, 20\}$	20	$\{10, 30, 50\}$	40
SDR	dB	-5	-5	∞ ($P = 0$, i.e., WGN only)	6	N/A
N	-	27	54	27	54	235
Δ	MHz	59.26	29.63	59.26	29.63	6.81
K	-	1	1	1	2	N/A
array structure	-	3×3	5×2	$1 \times M$	3×3	3×3
M	-	9	10	1 to 13	9	9
κ^*		4	1 to 1000	$\kappa^\star(\epsilon)$	$\kappa^\star(\epsilon)$	$\kappa^\star(\epsilon)$
ϵ		N/A	N/A	10^{-2}	10^{-3}	10^{-2}

Fig. 3. Wideband versus narrowband detection and estimation of SCs in the AWGN channel: (a) mean number of detected SCs, (b) root mean-square error (RMSE) of the distance estimates of detected SCs, and (c) RMSE of their angle estimates obtained with the proposed algorithm (in blue with crosses) and its narrowband version (in red with pluses) as a function of the array size M.

form (10) (Fig. 2c).[14] Furthermore, under the first assumption we distinguish between the two cases where the matrix $\boldsymbol{R}$ is known (Fig. 2a) and unknown (Fig. 2b) to the algorithm, and thus is estimated in the latter case.

We see in Fig. 2a and Fig. 2b that when the matrix $\boldsymbol{R}$ selected in the generative model matches that used in the design of the algorithm, whether the algorithm knows or does not know said matrix has little impact on its performance. Fig. 2c shows that when there is a mismatch, it only marginally affects the performance of the algorithm. Specifically, a comparison with Fig. 2b shows that at large SNR values the number of spurious SCs is slightly increased due to the mismatch.

3) Wideband Versus Narrowband Detection and Estimation of SCs in the AWGN Channel: State-of-the-art detection and estimation schemes are traditionally designed based on the narrow-band assumption, which neglects the second occurrence of $g(\varphi, \boldsymbol{p}_m)$ in (7) [8], [9]. In this study we show that neglecting this term in the proposed algorithm leads to an increase of the number of detected spurious SCs as the size of the array increases. To quantitatively assess this effect, we modify the simulation scenario as summarized in Table II; σ^2 is assumed known.

In this study and the subsequent ones we adopt the widely used convention in the radar community to convert (propagation) delays in their corresponding equivalent (propagation) distances. Fig. 3 depicts results obtained from 1000 simulation

trials that illustrate the behavior of the proposed algorithm (in blue with crosses) and of a simplified (narrowband) version of it that neglects the second occurrence of $g(\varphi, \boldsymbol{p}_m)$ in (7) (in red with pluses) as a function of the array size M. Fig. 3a, Fig. 3b and 3c report respectively the mean number of detected SCs, the RMSE of the distance estimates, and the RMSE of the angle estimates. When M is increased beyond 5, the narrowband assumption is violated and the narrowband version of the algorithm detects additional spurious SCs with dispersion vectors located in the vicinity of that of the active SC. By contrast, the RMSEs achieved with the proposed algorithm decrease slightly as M is increased.[15]

4) High-Resolution Capability of the Proposed Algorithm: To study the super-resolution capability of the proposed algorithm, we consider a scenario with $K = 2$ SCs with a controlled separation of their respective dispersion vectors. The parameter vector $\tilde{\boldsymbol{\psi}}_1$ of the first SC is drawn uniformly over Ψ. The parameter vector of the second SC is then set to either $\tilde{\boldsymbol{\psi}}_2 = \tilde{\boldsymbol{\psi}}_1 + [\Delta d/c \ 0]$ or $\tilde{\boldsymbol{\psi}}_2 = \tilde{\boldsymbol{\psi}}_1 + [0 \ \Delta\varphi]$. The spacings Δd and $\Delta\varphi$ are fractions of the RRL in, respectively, distance ($c/\tilde{B} \approx 0.3$ m) and angle ($56°$) [71]. The complex amplitudes of both SCs have unit magnitude and their respective phases are drawn uniformly and independently. Other system parameters are summarized in Tables I and II.

In Fig. 4, we compare the performance of the proposed algorithm (circles) with that of a MLE algorithm (trian-

[14]We remind the reader that the algorithm is designed based on the simplified form (15).

[15]To mitigate the impact of spurious SCs caused by large noise deviations, only detected SCs with distance and angle less than respectively 20 cm and $20°$ apart of those of the active SC are considered in the computation of the RMSE values.
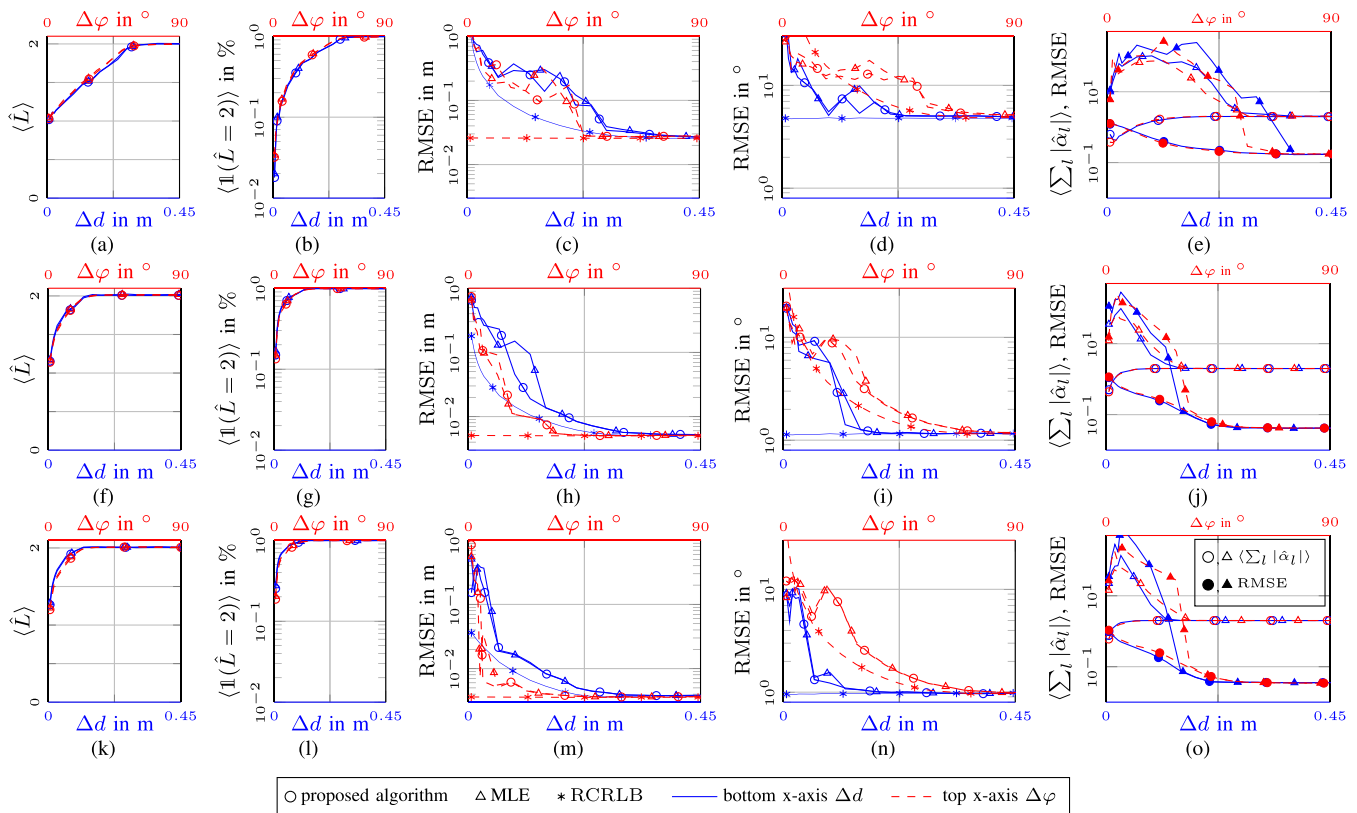
Fig. 4.　Detection and estimation of two closely spaced SCs in the scenario depicted in Subsubsec. VII-A4. Blue solid and red dashed lines refer to the bottom and top horizontal-axis, respectively. The panels in each row depict results obtained for the same SNR value, namely SNR = 10, 30, 50 dB, from the upper to the lower row. Panels in the columns depict from left to right the mean number of detected SCs, the relative frequency that exactly two SCs are detected, the RMSE of the estimated delays, the RMSE of the estimated angles, and the mean (unfilled) and RMSE (filled) of the moduli of the estimated amplitudes.
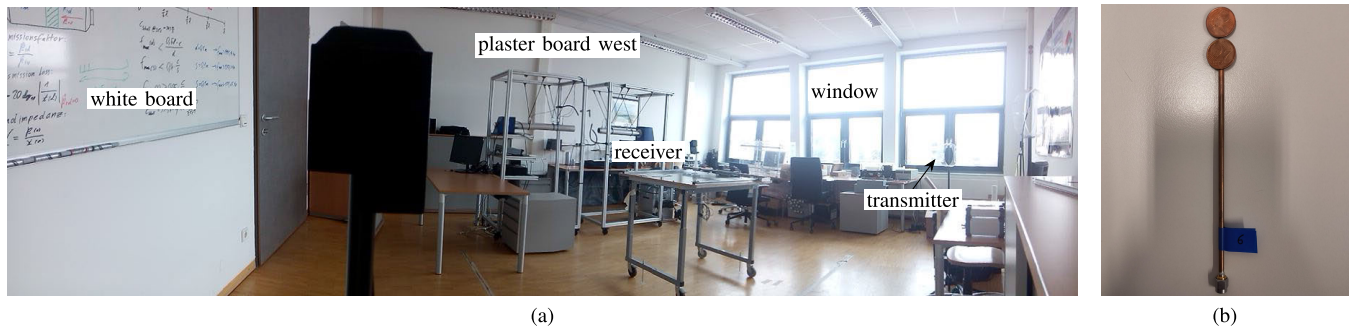

Fig. 5.　(a) Picture of the investigated room with the Rx, the Tx and some large-scale items labeled and (b) UWB antenna used at the Rx and the Tx. A floorplan of the room is given in Fig. 7 in Appendix D.

gles) inspired by [9]. The MLE algorithm that we use as a benchmark differs from that in [9] in two respects: the scheduling and the thresholding that both coincide with those implemented in the proposed algorithm. As mentioned in the manuscript, the method used to obtain the threshold is inspired from that in [14].[16] Performance versus spacings Δd and $\Delta \varphi$ are depicted as, respectively, blue solid and red dashed curves.

---

[16]Note that our modifications of the RiMAX algorithm proposed in [9] lead to improved performance. With the threshold suggested in [9], the number of spurious components would be significantly increased. Additionally, with the scheduling and thresholding suggested in [9], close-by components that violate the detection criterion would both be discarded, which would lead to a significant model-mismatch (see [9, Section 5.2.7]).

The first two columns of Fig. 4 present the mean number of detected SCs $\langle \hat{L} \rangle$ and the relative frequency that exactly two SCs are detected $\langle \mathbb{1}(\hat{L} = 2) \rangle$ versus spacing in distance and angle, respectively. Both algorithms are able to reliably find the correct number of SCs provided the respective dispersion vectors of the two SCs are sufficiently apart. At high SNR (see Figs. 4(k)–(o)) the spacing values beyond which this occurs is as low as 0.15 m or 20° for the system setting used in the study. Given that for this setting the RRL in distance is 0.3 m and that in angle is 56°, this result demonstrates the superresolution capability of the proposed algorithm. At lower SNR (see Figs. 4(a)–(e)) these values rise towards the RRLs. Note that these values are not only influenced by AWGN but also by the DC, meaning that at high SNR the resolution
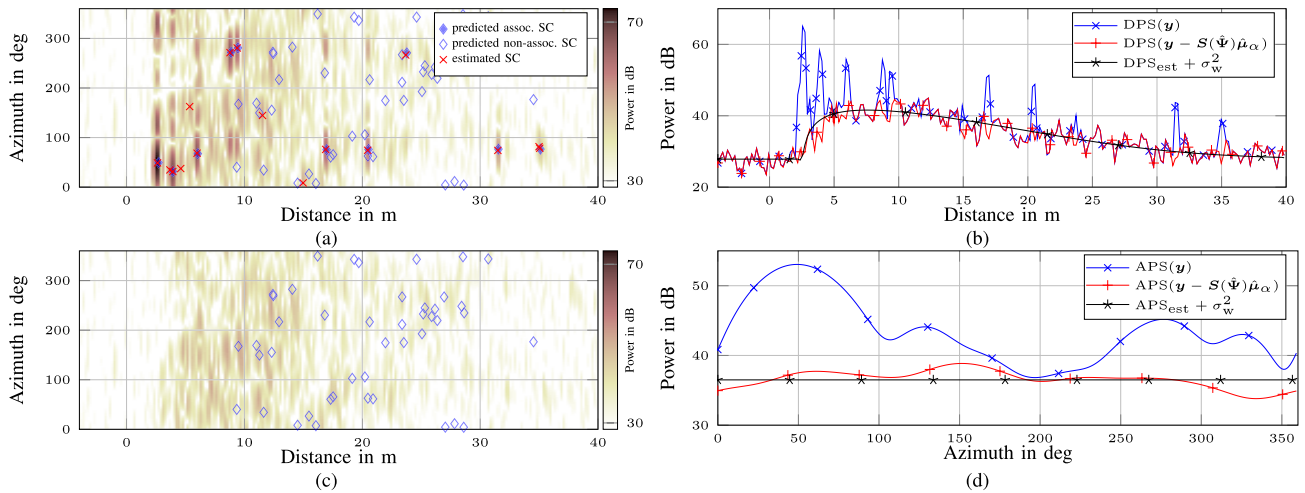
Fig. 6. Selected application of the proposed algorithm to assess the dispersion characteristics (in delay and angle of arrival) of the channel from UWB SIMO measurement data: (a) estimated DAPS and dispersion vectors of detected SCs, (b) various estimated DPS, (c) residual DAPS and dispersion vectors of un-associated SCs, (d) various estimated APS. Detailed descriptions of the measurement setting and the depicted results are given in Subsec. VII-B.

capability of the algorithms is restricted by the SDR. Further worth mentioning is that both algorithms tend to underestimate the number ($K = 2$) of active SCs when the spacing is reduced.

Columns three and four of Fig. 4 depict the RMSE of, respectively, the distance and angle estimates, provided exactly two SCs are detected. We associate the two detected SCs with the true SCs by means of the optimal subpattern assignment (OSPA) metric [72]. To be able to use the metric we normalize the estimated distances and angles with the RRL in distance and in angle, respectively. The root of the sum of the CRLBs (delay and angle) of the two SCs are also depicted (lines with stars). The estimates returned by both algorithms approach their respective root CRLB, provided the spacing of the two SCs in Ψ is large enough.

The last column in Fig. 4 presents the mean of the absolute value of the complex amplitudes $\langle \sum_l |\hat{\alpha}_l| \rangle$ (empty markers) and the RMSE of the absolute value of the complex amplitudes (filled markers) again provided exactly two SCs are detected. The two algorithms perform similarly in estimating the dispersion parameters; however our algorithm outperforms the modified MLE algorithm in estimating the complex amplitudes when the two SCs are closely spaced. This distinct behavior results from the specific structures of the algorithms: the modified MLE algorithm computes a least-squares estimate of the amplitudes, while the proposed algorithm computes a linear minimum mean-square error (MMSE) estimate based on the hyperparameter estimates, see (38) and (39). When the estimated dispersion vectors of the two detected SCs are closely spaced, the least-squares estimator computes the inverse of an ill-conditioned matrix, while the linear MMSE estimator regularizes this matrix.

B. Measured Radio Channels

For the experimental study we used a channel sounding equipment that transmits an m-sequence of $7\,\mathrm{GHz}$ bandwidth at $6.95\,\mathrm{GHz}$ carrier frequency. Details about the equipment can be found in [73]. After applying standard pre-processing steps (subtracting the cross-talk and equalizing the system response), the resulting signal is input to a filter with a root-raised-cosine transfer function with roll-off factor 0.6 and bandwidth $B = 1.6\,\mathrm{GHz}$ centered at $f_\mathrm{c} = 6\,\mathrm{GHz}$. The output signal with reduced bandwidth B is then Fourier transformed and sampled over $[-B/2, +B/2]$ with frequency spacing $\Delta = 6.8085\,\mathrm{MHz}$ to produce a length $N = 235$ vector collecting these samples. A virtual 3×3 antenna array with $2\,\mathrm{cm}$ inter-element spacing is emulated by means of a single antenna mounted on a positioning table. Since the received signals are essentially noise-free, WGN was artificially added with power set such that $\mathrm{SNR} = 40\,\mathrm{dB}$ to emulate the model in (4), see Table II for the parameter settings.

The room where the measurements were performed is depicted in Fig. 5. Based on a layout of it, see Fig. 7 in Appenix D, a classical mirror source method [74] computes the positions of predicted virtual sources associated with rays from the Tx antenna to the center of gravity of the Rx array positions that undergo up to 5 reflections on walls or large objects (windows, boards). To each such predicted source corresponds a predicted SC with dispersion vector computed from the position of the source, see Appendix D.

The following analysis concerns measurements obtained with Rx position p_1 depicted in Fig. 7 in Appendix D. Fig. 6(a) depicts the estimated DAPS computed from the received signal [74]. Note that this power spectrum incorporates the smoothing function of the aperture of the measurement equipment [64]. This will be the case for all power spectra considered in this study. The red crosses and blue diamonds mark the estimated dispersion vectors of the SCs detected by the algorithm and the predicted SCs, respectively. To each detected SC we associate (possibly no, one, or more than one) predicted SC as follows. A predicted SC is associated to a detected SC if their respective distances and angles are no more apart than, respectively, $10\,\mathrm{cm}$ ($1/3$ of the RRL) and $5°$ ($1/10$ of the RRL), see Appendix D for the rationale behind this choice. To most of the detected SCs, a unique predicted SC is associated in this way. The two detected SCs with dispersion parameters $35\,\mathrm{m}$ and $80°$ are associated with the same predicted SC. An association could not be made for four detected SCs. Fig. 6(c) shows the estimated DAPS computed from the residual signal $y - Z(\hat{\Psi})\hat{\mu}$. It also includes

the dispersion vectors of detected SCs (blue diamonds) that could not be associated with any detected SC. Clearly, the strong peaks in the estimated DAPS depicted in Fig. 6(a) have vanished. Fig. 6(b) shows the estimated DPS computed from the original signal (solid blue with crosses) and from the residual signal (solid red with pluses), as well as the theoretical DPS of the DC given in (14) with $\boldsymbol{\eta} = \hat{\boldsymbol{\eta}}$ plus the estimated noise variance $\hat{\sigma}$ (solid black with stars). The first two DPS result from averaging the respective DAPS over the angle domain. The DPS obtained from the residual and reconstructed signals match well. This empirically justifies our choice of model (14). Finally, Fig. 6(d) depicts the estimated APS computed in a similar way as the DPS depicted in Fig. 6(b). The first two estimated APS are obtained by averaging the respective DAPS over the delay domain. It can be seen that the estimated APS of the residual signal is nearly constant over the angle domain.

VIII. CONCLUSION

In this paper, we derive and analyze a super-resolution algorithm for detecting and estimating specular components as well as estimating the power spectrum of the dense component plus noise in an ultra-wide band single-input—multiple-output (SIMO) multipath channel. Estimated parameters are among others the delay, angle-of-arrival, and complex amplitude of the detected specular components as well as the parameters of a parametric model of the delay power spectrum characterizing the dense component. The design of the algorithm is inspired by sparse Bayesian learning. As a result it embodies a pruning condition that determines whether a candidate specular component is considered active or not. The threshold of the pruning condition is adapted to control the probability of detecting spurious specular components.

Numerical studies in a synthetic environment show that the simplifying assumptions underlying the derivation of the algorithm are realistic and that the relative frequencies of detecting spurious specular components and missing active specular components are close to the respective probabilities derived theoretically. These studies also demonstrate several virtues of the algorithm: (a) its ability to still detect and accurately estimate specular components, even when their separation in delay and azimuth is down to half the Rayleigh resolution limit of the equipment; (b) it is robust in the sense that it tends to detect no more specular components than the actual ones. An experimental study illustrates the ability of the proposed algorithm to accurately infer the dispersive characteristics (in delay and angle of arrival) of the UWB SIMO channel. Owing to this high efficiency the proposed algorithm has promising potential applications in all areas of wireless communications that exploit extended channel state information, such as integrated sensing and communications (ISAC) and radio-based localization.

APPENDIX A
PROOF OF THEOREM 1

Proof: As shown in [53] it follows from Assumption 1 that the real and imaginary parts of the Gaussian field $x(\boldsymbol{\psi}.)$ in (41) exhibits the following properties:

1) They have equal constant variance: $\mathbb{E}\big[|\Re\{x(\boldsymbol{\psi}.)\}|^2\big] = \mathbb{E}\big[|\Im\{x(\boldsymbol{\psi}.)\}|^2\big] = 1/2$, $\boldsymbol{\psi}. \in \Psi$.
2) They are independent: $\mathbb{E}[\Re\{x(\boldsymbol{\psi}.)\}\Im\{x(\boldsymbol{\psi}.')\}] = 0$, $\boldsymbol{\psi}., \boldsymbol{\psi}.' \in \Psi$.

Since $\Re\{x(\boldsymbol{\psi}.)\}$ and $\Im\{x(\boldsymbol{\psi}.)\}$ are independent $2|x(\boldsymbol{\psi}.)|^2$ is a random χ^2 field on Ψ with two degrees of freedom [49], [50].[17] Note that unless the DC vanishes, i.e. $P = 0$, see (15), the Gaussian field $x(\boldsymbol{\psi}.)$ is non-stationary and so is $2|x(\boldsymbol{\psi}.)|^2$. The probability that the χ^2 field exceeds a threshold is asymptotically equivalent to the probability of the field's excursion above the threshold when said threshold grows large [49], [50]. Specifically, by applying Weyl's tube formula [50, Theorem 3.3.1] to $2|x(\boldsymbol{\psi}.)|^2$ and making use of [50, Theorem 4.4.1] combined with [50, Section 4.5.2] we obtain

$$\mathbb{P}\Big[\sup_{\boldsymbol{\psi}.} 2|x(\boldsymbol{\psi}.)|^2 \geq 2\kappa\Big]$$
$$\sim \left[\int_{\Psi} \frac{1}{\pi}\sqrt{\det(\boldsymbol{\Lambda}(\boldsymbol{\psi}.))}\,\mathrm{d}\boldsymbol{\psi}.\right]\kappa\mathrm{e}^{-\kappa}, \quad \kappa \to \infty. \tag{55}$$

In this expression $\boldsymbol{\Lambda}(\boldsymbol{\psi}.) \in \mathbb{R}^{2\times2}$ is the covariance matrix

$$\boldsymbol{\Lambda}(\boldsymbol{\psi}.) = \mathbb{E}\left[\left[\frac{\partial x(\boldsymbol{\psi}.)}{\partial \boldsymbol{\psi}.}\right]\left[\frac{\partial x(\boldsymbol{\psi}.)}{\partial \boldsymbol{\psi}.}\right]^{\mathrm{H}}\right]$$
$$= \begin{bmatrix} \mathbb{E}\left[\frac{\partial x(\boldsymbol{\psi}.)\partial x(\boldsymbol{\psi}.)^*}{\partial \tau^2}\right] & \mathbb{E}\left[\frac{\partial x(\boldsymbol{\psi}.)\partial x(\boldsymbol{\psi}.)^*}{\partial \tau \partial \varphi}\right] \\ \mathbb{E}\left[\frac{\partial x(\boldsymbol{\psi}.)\partial x(\boldsymbol{\psi}.)^*}{\partial \varphi \partial \tau}\right] & \mathbb{E}\left[\frac{\partial x(\boldsymbol{\psi}.)\partial x(\boldsymbol{\psi}.)^*}{\partial \varphi^2}\right] \end{bmatrix}. \tag{56}$$

As shown in Appendix C, the entries in (56) read

$$\mathbb{E}\left[\frac{\partial x(\boldsymbol{\psi}.)\partial x(\boldsymbol{\psi}.)^*}{\partial \tau^2}\right] = 4\pi^2 b_1(\tau)b_2(\tau) \tag{57}$$

$$\mathbb{E}\left[\frac{\partial x(\boldsymbol{\psi}.)\partial x(\boldsymbol{\psi}.)^*}{\partial \varphi^2}\right] = \frac{4\pi^2}{M}\sum_{m\in\mathcal{M}} d_m^2(\varphi)f_{\mathrm{c}}^2 \tag{58}$$

$$\mathbb{E}\left[\frac{\partial x(\boldsymbol{\psi}.)\partial x(\boldsymbol{\psi}.)^*}{\partial \tau \partial \varphi}\right] = \mathbb{E}\left[\frac{\partial x(\boldsymbol{\psi}.)\partial x(\boldsymbol{\psi}.)^*}{\partial \varphi \partial \tau}\right] = 0. \tag{59}$$

The right-hand side in (44) follows then from (57)-(59). $\quad\square$

APPENDIX B
PROOF OF LEMMA 2

Proof: With the definition of q the right-hand expression in (55) reads $q\kappa\mathrm{e}^{-\kappa}$. Clearly, this expression determines a non-increasing function of κ defined on $[1,\infty)$ with range $(0, q/e)$. Given $\epsilon \in (0, q/e)$ we seek the minimum value of $\kappa \in [1,\infty)$ such that $q\kappa\mathrm{e}^{-\kappa} \leq \epsilon$ holds. Obviously the sought value solves $q\kappa\mathrm{e}^{-\kappa} = \epsilon$, i.e. equals $\kappa^\star(\epsilon)$ in (48). $\quad\square$

APPENDIX C
COVARIANCE MATRIX OF THE χ^2 RANDOM FIELD $2|x(\boldsymbol{\psi}.)|^2$

The (i,j)-entry of (56) is given in (60), shown at the bottom of the next page, with $\tilde{\boldsymbol{R}}$ given

---

[17] The real and imaginary parts of $\sqrt{2}x(\boldsymbol{\psi}.)$ have unit variance, in accordance with the definition of a χ^2 field.

according to (16). The partial derivatives of the signal $Z(f; \tau, \varphi, \boldsymbol{p}^{(m)}) = \mathrm{e}^{j2\pi f_c g(\varphi, \boldsymbol{p}_m)} S(f) \mathrm{e}^{-j2\pi f\tau}$ w.r.t. τ and φ are, respectively, $\partial \boldsymbol{z}_m(\boldsymbol{\psi})/\partial\tau = -\mathrm{e}^{j2\pi f_c g(\varphi, \boldsymbol{p}_m)}\dot{\boldsymbol{a}}_\mathrm{f}(\tau)$ and $\partial \boldsymbol{z}_m(\boldsymbol{\psi})/\partial\varphi = d_m(\varphi)\exp(j2\pi f_c g(\varphi, \boldsymbol{p}_m))j2\pi f_c \boldsymbol{a}_\mathrm{f}(\tau)$ with $d_m(\varphi) = \partial g(\varphi, \boldsymbol{p}_m))/\partial\varphi$ and $\dot{\boldsymbol{a}}_\mathrm{f}(\tau) = \partial \boldsymbol{a}_\mathrm{f}(\tau)/\partial\tau$. Due to the centro-symmetry of the spatial aperture, see Assumption 1, for any $m \in \mathcal{M}$, there exists an index $m' \in \mathcal{M}$ such that $\boldsymbol{p}_{m'} - \boldsymbol{p} = -(\boldsymbol{p}_m - \boldsymbol{p})$. As a result

$$\sum_{m \in \mathcal{M}} d_m(\varphi) = 0. \tag{61}$$

We now proceed with the computation of the specific entries of (56).

1) Second-Order Partial Derivatives of $x(\boldsymbol{\psi})$ w.r.t. τ: One can easily check that because of (61) the second and third terms of (60) vanish in this case. As a result, we can write

$$\mathbb{E}\left[\frac{\partial x(\boldsymbol{\psi})\partial x(\boldsymbol{\psi})^*}{\partial\tau^2}\right]$$
$$= 2\left(\frac{\sum_{m \in \mathcal{M}}\dot{\boldsymbol{a}}_\mathrm{f}(\tau)^\mathrm{H}\tilde{\boldsymbol{R}}^{-1}\dot{\boldsymbol{a}}_\mathrm{f}(\tau)}{\sum_{m \in \mathcal{M}}\boldsymbol{a}_\mathrm{f}(\tau)^\mathrm{H}\tilde{\boldsymbol{R}}^{-1}\boldsymbol{a}_\mathrm{f}(\tau)} - \frac{\Re\left\{\sum_{m \in \mathcal{M}}\dot{\boldsymbol{a}}_\mathrm{f}(\tau)^\mathrm{H}\tilde{\boldsymbol{R}}^{-1}\boldsymbol{a}_\mathrm{f}(\tau)\right\}^2}{\left|\sum_{m \in \mathcal{M}}\boldsymbol{a}_\mathrm{f}(\tau)^\mathrm{H}\tilde{\boldsymbol{R}}^{-1}\boldsymbol{a}_\mathrm{f}(\tau)\right|^2}\right)$$
$$= 2\frac{\left(\dot{\boldsymbol{a}}_\mathrm{f}(\tau)^\mathrm{H}\tilde{\boldsymbol{R}}^{-1}\dot{\boldsymbol{a}}_\mathrm{f}(\tau)\right)}{\left(\boldsymbol{a}_\mathrm{f}(\tau)^\mathrm{H}\tilde{\boldsymbol{R}}^{-1}\boldsymbol{a}_\mathrm{f}(\tau)\right)}$$
$$\times\left(1 - \frac{\Re\left\{\dot{\boldsymbol{a}}_\mathrm{f}(\tau)^\mathrm{H}\tilde{\boldsymbol{R}}^{-1}\boldsymbol{a}_\mathrm{f}(\tau)\right\}^2}{\left(\boldsymbol{a}_\mathrm{f}(\tau)^\mathrm{H}\tilde{\boldsymbol{R}}^{-1}\boldsymbol{a}_\mathrm{f}(\tau)\right)\left(\dot{\boldsymbol{a}}_\mathrm{f}(\tau)^\mathrm{H}\tilde{\boldsymbol{R}}^{-1}\dot{\boldsymbol{a}}_\mathrm{f}(\tau)\right)}\right)$$
$$= 8\pi^2 b_1(\tau)b_2(\tau) \tag{62}$$

where

$$b_1(\tau) = \left[\left(\dot{\boldsymbol{a}}_\mathrm{f}(\tau)^\mathrm{H}\tilde{\boldsymbol{R}}^{-1}\dot{\boldsymbol{a}}_\mathrm{f}(\tau)\right)/\left(4\pi^2\boldsymbol{a}_\mathrm{f}(\tau)^\mathrm{H}\tilde{\boldsymbol{R}}^{-1}\boldsymbol{a}_\mathrm{f}(\tau)\right)\right] \tag{63}$$

and

$$b_2(\tau) = 1 - \frac{\Re\left\{\dot{\boldsymbol{a}}_\mathrm{f}(\tau)^\mathrm{H}\tilde{\boldsymbol{R}}^{-1}\boldsymbol{a}_\mathrm{f}(\tau)\right\}^2}{\left(\boldsymbol{a}_\mathrm{f}(\tau)^\mathrm{H}\tilde{\boldsymbol{R}}^{-1}\boldsymbol{a}_\mathrm{f}(\tau)\right)\left(\dot{\boldsymbol{a}}_\mathrm{f}(\tau)^\mathrm{H}\tilde{\boldsymbol{R}}^{-1}\dot{\boldsymbol{a}}_\mathrm{f}(\tau)\right)}. \tag{64}$$

The above expression is a delay-dependent loss factor that depends of the structure of the noise vector $\boldsymbol{n}$. Note that if $\boldsymbol{n}$ is white, $b_2(\tau) = 1$, otherwise $b_2(\tau) < 1$, typically.

2) Second-Order Partial Derivatives of $x(\boldsymbol{\psi})$ w.r.t. φ: In this case, the last three terms in (60) vanish, again because of (61). We readily obtain,

$$\mathbb{E}\left[\frac{\partial x(\boldsymbol{\psi})\partial x(\boldsymbol{\psi})^*}{\partial\varphi^2}\right] = \frac{8\pi^2 f_c^2}{M}\sum_{m \in \mathcal{M}} d_m^2(\varphi). \tag{65}$$

3) Second-Order Partial Derivatives of $x(\boldsymbol{\psi})$ w.r.t. τ and φ: In this case (61) make all terms in (60) vanish. Thus,

$$\mathbb{E}\left[\frac{\partial x(\boldsymbol{\psi})\partial x(\boldsymbol{\psi})^*}{\partial\tau\partial\varphi}\right] = 0. \tag{66}$$

APPENDIX D
VALIDATION OF THE SCs DETECTED BY THE ALGORITHM

In this appendix, we provide a qualitative study that attempts to relate the SCs detected by the proposed algorithm to likely propagation mechanisms in the environment where the experimental data were collected. The results of this study supplement those presented in Subsec. VII-B.

A 2-D coordinate system including the layout of the room where the measurements were taken is shown in Fig. 7. Also reported are the two selected positions $\boldsymbol{p}_1$ and $\boldsymbol{p}_2$ of (the center of gravity of) the Rx (virtual) array and the fixed position $\boldsymbol{p}_\mathrm{Tx}$ of the (single) Tx antenna. We recall that the mirror source method [74] computes the positions of predicted virtual sources associated with rays from the Tx antenna to the Rx array positions that undergo up to 5 reflections on walls or large objects (windows, boards). For the sake of conciseness we refer to virtual sources in the sequel as sources. The position, denoted by $\hat{\boldsymbol{p}}_l$, of the predicted source corresponding to the lth SC, $l = 1, \dots \hat{L}$ detected by the algorithm is computed based on the estimated dispersion vector $\hat{\boldsymbol{\psi}}_l$ using the relation $\hat{\boldsymbol{p}}_l = \boldsymbol{p} + c\hat{\tau}_l[\cos(\hat{\varphi}_l)\ \sin(\hat{\varphi}_l)]^\mathrm{T}$, where $\boldsymbol{p}$ either equals $\boldsymbol{p}_1$ or $\boldsymbol{p}_2$. These positions are depicted in Fig. 7 as red crosses and blue pluses for the Rx array positions $\boldsymbol{p}_1$ and $\boldsymbol{p}_2$, respectively.

The procedure described next attempts to associates detected sources and predicted sources. Possibly no, one, or more than one predicted SCs are associated to each detected SC as follows. A predicted SC is associated to a detected SC if their respective distances and angles are no more than, respectively, 10 cm (1/3 of the RRL in distance) and 5° (1/10 of the RRL

$$\mathbb{E}\left[\frac{\partial x(\boldsymbol{\psi})\partial x(\boldsymbol{\psi})^*}{\partial\psi_i\partial\psi_j}\right] = 2\left(\frac{\sum_{m \in \mathcal{M}}\left(\frac{\partial \boldsymbol{z}_m(\boldsymbol{\psi})}{\partial\psi_j}\right)^\mathrm{H}\tilde{\boldsymbol{R}}^{-1}\frac{\partial \boldsymbol{z}_m(\boldsymbol{\psi})}{\partial\psi_i}}{\sum_{m \in \mathcal{M}}\boldsymbol{z}_m(\boldsymbol{\psi})^\mathrm{H}\tilde{\boldsymbol{R}}^{-1}\boldsymbol{z}_m(\boldsymbol{\psi})} - \frac{\sum_{m \in \mathcal{M}}\left(\frac{\partial \boldsymbol{z}_m(\boldsymbol{\psi})}{\partial\psi_i}\right)^\mathrm{H}\tilde{\boldsymbol{R}}^{-1}\boldsymbol{z}_m(\boldsymbol{\psi})\Re\left\{\sum_{m' \in \mathcal{M}}\left(\frac{\partial \boldsymbol{z}_{m'}(\boldsymbol{\psi})}{\partial\psi_j}\right)^\mathrm{H}\tilde{\boldsymbol{R}}^{-1}\boldsymbol{z}_{m'}(\boldsymbol{\psi})\right\}}{\left|\sum_{m \in \mathcal{M}}\boldsymbol{z}_m(\boldsymbol{\psi})^\mathrm{H}\tilde{\boldsymbol{R}}^{-1}\boldsymbol{z}_m(\boldsymbol{\psi})\right|^2}\right.$$
$$- \frac{\sum_{m \in \mathcal{M}}\boldsymbol{z}_m(\boldsymbol{\psi})^\mathrm{H}\tilde{\boldsymbol{R}}^{-1}\frac{\partial \boldsymbol{z}_m(\boldsymbol{\psi})}{\partial\psi_j}\Re\left\{\sum_{m' \in \mathcal{M}}\left(\frac{\partial \boldsymbol{z}_{m'}(\boldsymbol{\psi})}{\partial\psi_i}\right)^\mathrm{H}\tilde{\boldsymbol{R}}^{-1}\boldsymbol{z}_{m'}(\boldsymbol{\psi})\right\}}{\left|\sum_{m \in \mathcal{M}}\boldsymbol{z}_m(\boldsymbol{\psi})^\mathrm{H}\tilde{\boldsymbol{R}}^{-1}\boldsymbol{z}_m(\boldsymbol{\psi})\right|^2}$$
$$\left.+ \frac{\Re\left\{\sum_{m \in \mathcal{M}}\left(\frac{\partial \boldsymbol{z}_m(\boldsymbol{\psi})}{\partial\psi_i}\right)^\mathrm{H}\tilde{\boldsymbol{R}}^{-1}\boldsymbol{z}_m(\boldsymbol{\psi})\right\}\Re\left\{\sum_{m' \in \mathcal{M}}\left(\frac{\partial \boldsymbol{z}_{m'}(\boldsymbol{\psi})}{\partial\psi_i}\right)^\mathrm{H}\tilde{\boldsymbol{R}}^{-1}\boldsymbol{z}_{m'}(\boldsymbol{\psi})\right\}}{\left|\sum_{m \in \mathcal{M}}\boldsymbol{z}_m(\boldsymbol{\psi})^\mathrm{H}\tilde{\boldsymbol{R}}^{-1}\boldsymbol{z}_m(\boldsymbol{\psi})\right|^2}\right) \tag{60}$$
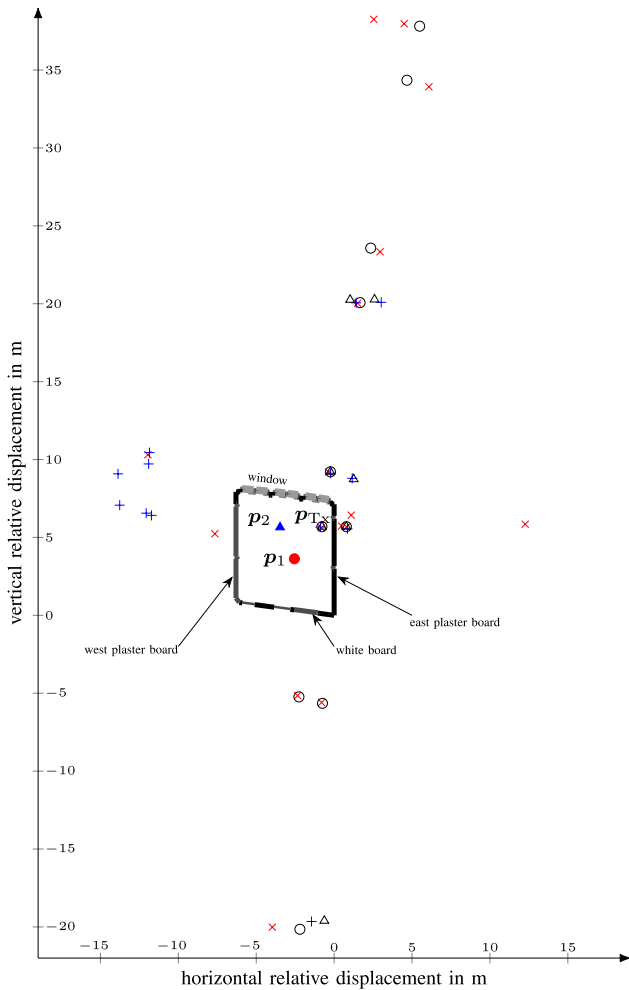
Fig. 7. Floorplan of the investigated environment including the (fixed) Tx position p_{Tx}, the two selected positions p_1 (red filled circle) and p_2 (blue filled triangle) of (the center of gravity of) the Rx array, the locations of estimated (virtual) sources for Rx array position p_1 (red crosses) and Rx array position p_2 (blue pluses) and the positions of associated predicted sources (black circles and triangles, respectively) computed with the mirror source model.

in angle) apart. These selected values are within the same order of magnitude as, respectively, the 5 cm approximate accuracy of the floorplan (measured with a tape measure) and the CRLBs of the estimated distances and angles. The positions of successfully associated predicted sources are depicted in Fig. 7 as black circles and triangles for Rx positions p_1 and p_2, respectively.

The algorithm is able to identify the LOS, most of the predicted first-order reflections and some predicted higher-order reflections for both Rx positions. Worth noting are the rays with reflections up to order five via the white board and the window highlighted in Fig. 7 and Fig. 5 and the second-order rays with reflections via the west plaster board and the east plaster board. The two former items are made of more reflective materials than the two latter. Furthermore, scattering from a metallic frame (see Fig. 5) that was not considered in the mirror source method could explain the detected source located at approximately $[-8\ 6]$ m close to the west plaster board. For Rx position p_2 many detected sources are found in a region around $[-15\ 8]$ m. They are likely to originate from scattering from this metallic frame.

REFERENCES

[1] A. Conti et al., "Location awareness in beyond 5G networks," *IEEE Commun. Mag.*, vol. 59, no. 11, pp. 22–27, Nov. 2021.
[2] H. Wymeersch et al., "6G radio requirements to support integrated communication, localization, and sensing," in *Proc. Joint Eur. Conf. Netw. Commun. 6G Summit*, Grenoble, France, Jun. 2022, pp. 463–469.
[3] G. Kwon, A. Conti, H. Park, and M. Z. Win, "Joint communication and localization in millimeter wave networks," *IEEE J. Sel. Topics Signal Process.*, vol. 15, no. 6, pp. 1439–1454, Nov. 2021.
[4] E. Björnson, L. Sanguinetti, H. Wymeersch, J. Hoydis, and T. L. Marzetta, "Massive MIMO is a reality—What is next? Five promising research directions for antenna arrays," *Digital Signal Process.*, vol. 94, pp. 3–20, Nov. 2019.
[5] 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*. Cham, Switzerland: Springer, 1993, pp. 99–151.
[6] R. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Trans. Antennas Propag.*, vol. AP-34, no. 3, pp. 276–280, Mar. 1986.
[7] 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.
[8] 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.
[9] A. Richter, "Estimation of radio channel parameters: Models and algorithms," Ph.D. dissertation, Dept. Elect. Eng. Inf. Technol., Ilmenau Univ. Technol., Ilmenau, Germany, 2005.
[10] F. Dublanchet, J. Idier, and P. Duwaut, "Direction-of-arrival and frequency estimation using Poisson–Gaussian modeling," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, Munich, Germany, Apr. 1997, pp. 3501–3504.
[11] C. Andrieu and A. Doucet, "Joint Bayesian model selection and estimation of noisy sinusoids via reversible jump MCMC," *IEEE Trans. Signal Process.*, vol. 47, no. 10, pp. 2667–2676, Oct. 1999.
[12] 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.
[13] 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.
[14] B. Nadler and A. Kontorovich, "Model selection for sinusoids in noise: Statistical analysis and a new penalty term," *IEEE Trans. Signal Process.*, vol. 59, no. 4, pp. 1333–1345, Apr. 2011.
[15] A. Mariani, A. Giorgetti, and M. Chiani, "Model order selection based on information theoretic criteria: Design of the penalty," *IEEE Trans. Signal Process.*, vol. 63, no. 11, pp. 2779–2789, Jun. 2015.
[16] J. A. Tropp and S. J. Wright, "Computational methods for sparse solution of linear inverse problems," *Proc. IEEE*, vol. 98, no. 6, pp. 948–958, Jun. 2010.
[17] S. S. Chen, D. L. Donoho, and M. A. Saunders, "Atomic decomposition by basis pursuit," *SIAM Rev.*, vol. 43, no. 1, pp. 129–159, Jan. 2001.
[18] R. Tibshirani, "Regression shrinkage and selection via the lasso," *J. Roy. Stat. Soc., B, Methodol.*, vol. 58, no. 1, pp. 267–288, Jan. 1996.
[19] D. J. C. MacKay, *Bayesian Methods for Backpropagation Networks*. New York, NY, USA: Springer, 1996, pp. 211–254.
[20] M. E. Tipping, "Sparse Bayesian learning and the relevance vector machine," *J. Mach. Learn. Res.*, vol. 1, pp. 211–244, Jun. 2001.
[21] D. P. Wipf and B. D. Rao, "Sparse Bayesian learning for basis selection," *IEEE Trans. Signal Process.*, vol. 52, no. 8, pp. 2153–2164, Aug. 2004.
[22] F. Champagnat, Y. Goussard, and J. Idier, "Unsupervised deconvolution of sparse spike trains using stochastic approximation," *IEEE Trans. Signal Process.*, vol. 44, no. 12, pp. 2988–2998, Dec. 1996.
[23] C. Soussen, J. Idier, D. Brie, and J. Duan, "From Bernoulli–Gaussian deconvolution to sparse signal restoration," *IEEE Trans. Signal Process.*, vol. 59, no. 10, pp. 4572–4584, Oct. 2011.
[24] R. Giri and B. Rao, "Type I and type II Bayesian methods for sparse signal recovery using scale mixtures," *IEEE Trans. Signal Process.*, vol. 64, no. 13, pp. 3418–3428, Jul. 2016.
[25] C. M. Bishop and M. E. Tipping, "Variational relevance vector machines," in *Proc. Conf. Uncertainty Artif. Intell.* Stanford, CA, USA: Morgan Kaufmann Publishers Inc., Jun. 2000, pp. 46–53.

[26] D. G. Tzikas, A. C. Likas, and N. P. Galatsanos, "The variational approximation for Bayesian inference," *IEEE Signal Process. Mag.*, vol. 25, no. 6, pp. 131–146, Nov. 2008.

[27] P. Stoica, P. Babu, and J. Li, "New method of sparse parameter estimation in separable models and its use for spectral analysis of irregularly sampled data," *IEEE Trans. Signal Process.*, vol. 59, no. 1, pp. 35–47, Jan. 2011.

[28] X. Tan, W. Roberts, J. Li, and P. Stoica, "Sparse learning via iterative minimization with application to MIMO radar imaging," *IEEE Trans. Signal Process.*, vol. 59, no. 3, pp. 1088–1101, Mar. 2011.

[29] R. R. Pote and B. D. Rao, "Maximum likelihood-based gridless DoA estimation using structured covariance matrix recovery and SBL with grid refinement," *IEEE Trans. Signal Process.*, vol. 71, pp. 802–815, 2023.

[30] C. Ekanadham, D. Tranchina, and E. P. Simoncelli, "Sparse decomposition of transformation-invariant signals with continuous basis pursuit," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, May 2011, pp. 4060–4063.

[31] Z. Yang, L. Xie, and C. Zhang, "Off-grid direction of arrival estimation using sparse Bayesian inference," *IEEE Trans. Signal Process.*, vol. 61, no. 1, pp. 38–43, Jan. 2013.

[32] M. Duarte and R. Baraniuk, "Spectral compressive sensing," *Appl. Comput. Harmon. Anal.*, vol. 35, no. 1, pp. 111–129, 2013.

[33] L. Hu, J. Zhou, Z. Shi, and Q. Fu, "A fast and accurate reconstruction algorithm for compressed sensing of complex sinusoids," *IEEE Trans. Signal Process.*, vol. 61, no. 22, pp. 5744–5754, Nov. 2013.

[34] K. Fyhn, M. F. Duarte, and S. H. Jensen, "Compressive parameter estimation for sparse translation-invariant signals using polar interpolation," *IEEE Trans. Signal Process.*, vol. 63, no. 4, pp. 870–881, Feb. 2015.

[35] D. Malioutov, M. Cetin, and A. S. Willsky, "A sparse signal reconstruction perspective for source localization with sensor arrays," *IEEE Trans. Signal Process.*, vol. 53, no. 8, pp. 3010–3022, Aug. 2005.

[36] B. N. Bhaskar, G. Tang, and B. Recht, "Atomic norm denoising with applications to line spectral estimation," *IEEE Trans. Signal Process.*, vol. 61, no. 23, pp. 5987–5999, Dec. 2013.

[37] Z. Yang and L. Xie, "On gridless sparse methods for line spectral estimation from complete and incomplete data," *IEEE Trans. Signal Process.*, vol. 63, no. 12, pp. 3139–3153, Jun. 2015.

[38] Y. Chi and Y. Chen, "Compressive two-dimensional harmonic retrieval via atomic norm minimization," *IEEE Trans. Signal Process.*, vol. 63, no. 4, pp. 1030–1042, Feb. 2015.

[39] Z. Yang and L. Xie, "Enhancing sparsity and resolution via reweighted atomic norm minimization," *IEEE Trans. Signal Process.*, vol. 64, no. 4, pp. 995–1006, Feb. 2016.

[40] D. Shutin and B. H. Fleury, "Sparse variational Bayesian SAGE algorithm with application to the estimation of multipath wireless channels," *IEEE Trans. Signal Process.*, vol. 59, no. 8, pp. 3609–3623, Aug. 2011.

[41] D. Shutin, W. Wang, and T. Jost, "Incremental sparse Bayesian learning for parameter estimation of superimposed signals," in *Proc. SAMPTA*, Sep. 2013, pp. 6–9.

[42] T. L. Hansen, P. B. Jørgensen, N. L. Pedersen, C. N. Manchón, and B. H. Fleury, "Bayesian compressed sensing with unknown measurement noise level," in *Proc. Asilomar Conf. Signals, Syst. Comput.*, Pacific Grove, CA, USA, Nov. 2013, pp. 148–152.

[43] T. L. Hansen, M. A. Badiu, B. H. Fleury, and B. D. Rao, "A sparse Bayesian learning algorithm with dictionary parameter estimation," in *Proc. IEEE 8th Sensor Array Multichannel Signal Process. Workshop (SAM)*, Jun. 2014, pp. 385–388.

[44] D. Shutin and N. Schneckenburger, "Joint detection and super-resolution estimation of multipath signal parameters using incremental automatic relevance determination," 2015, *arXiv:1503.01898*.

[45] T. L. Hansen, B. H. Fleury, and B. D. Rao, "Superfast line spectral estimation," *IEEE Trans. Signal Process.*, vol. 66, no. 10, pp. 2511–2526, May 2018.

[46] M. Badiu, T. L. Hansen, and B. H. Fleury, "Variational Bayesian inference of line spectra," *IEEE Trans. Signal Process.*, vol. 65, no. 9, pp. 2247–2261, May 2017.

[47] M. E. Tipping and A. C. Faul, "Fast marginal likelihood maximisation for sparse Bayesian models," in *Proc. 9th Int. Workshop Artif. Intell. Statist.*, C. M. Bishop and B. J. Frey, Eds. Key West, FL, USA, 2003, pp. 3–6.

[48] D. Shutin, S. R. Kulkarni, and H. V. Poor, "Incremental reformulated automatic relevance determination," *IEEE Trans. Signal Process.*, vol. 60, no. 9, pp. 4977–4981, Sep. 2012.

[49] K. J. Worsley, "Local maxima and the expected Euler characteristic of excursion sets of χ^2, F and t fields," *Adv. Appl. Probab.*, vol. 26, no. 1, pp. 13–42, Mar. 1994.

[50] R. J. Adler, J. E. Taylor, and K. J. Worsley, "Applications of random fields and geometry: Foundations and case studies," 2015. [Online]. Available: https://web.stanford.edu/class/stats317/hrf.pdf

[51] S. Jiang, W. Wang, Y. Miao, W. Fan, and A. F. Molisch, "A survey of dense multipath and its impact on wireless systems," *IEEE Open J. Antennas Propag.*, vol. 3, pp. 435–460, 2022.

[52] S. Semper, M. Döbereiner, C. Steinmetz, M. Landmann, and R. S. Thomä, "High resolution parameter estimation for wideband radio channel sounding," *IEEE Trans. Antennas Propag.*, vol. 71, no. 8, pp. 6728–6743, Jun. 2023.

[53] E. Leitinger, S. Grebien, B. Fleury, and K. Witrisal, "Detection and estimation of a spectral line in MIMO systems," in *Proc. Asilomar Conf. Signals, Syst. Comput.*, Pacific Grove, CA, USA, Nov. 2020, pp. 1090–1095.

[54] P. Stoica and R. Moses, *Spectral Analysis of Signals*. Upper Saddle River, NJ, USA: Prentice-Hall, 2005.

[55] J. Karedal, S. Wyne, P. Almers, F. Tufvesson, and A. F. Molisch, "A measurement-based statistical model for industrial ultra-wideband channels," *IEEE Trans. Wireless Commun.*, vol. 6, no. 8, pp. 3028–3037, Aug. 2007.

[56] B. H. Fleury, "First- and second-order characterization of direction dispersion and space selectivity in the radio channel," *IEEE Trans. Inf. Theory*, vol. 46, no. 6, pp. 2027–2044, Sep. 2000.

[57] R. J. Adler and J. E. Taylor, *Random Fields and Geometry*. New York, NY, USA: Springer, 2007.

[58] W. Weichselberger, "Spatial structure of multiple antenna radio channels: A signal processing viewpoint," Ph.D. dissertation, Inst. Telecommun., Technische Universität Wien, Vienna, Austria, 2003.

[59] J. Andersen, J. Nielsen, G. Pedersen, G. Bauch, and J. M. Herdin, "Room electromagnetics," *IEEE Antennas Propag. Mag.*, vol. 49, no. 2, pp. 27–33, Apr. 2007.

[60] T. Pedersen, "Stochastic multipath model for the in-room radio channel based on room electromagnetics," *IEEE Trans. Antennas Propag.*, vol. 67, no. 4, pp. 2591–2603, Apr. 2019.

[61] J. Salmi, A. Richter, and V. Koivunen, "Detection and tracking of MIMO propagation path parameters using state-space approach," *IEEE Trans. Signal Process.*, vol. 57, no. 4, pp. 1538–1550, Apr. 2009.

[62] F. Talebi and T. Pratt, "Model order selection for complex sinusoids in the presence of unknown correlated Gaussian noise," *IEEE Trans. Signal Process.*, vol. 63, no. 7, pp. 1664–1674, Apr. 2015.

[63] Y. Hua, "Estimating two-dimensional frequencies by matrix enhancement and matrix pencil," *IEEE Trans. Signal Process.*, vol. 40, no. 9, pp. 2267–2280, Sep. 1992.

[64] D. Johnson and D. Dudgeon, *Array Signal Processing: Concepts and Techniques*. Upper Saddle River, NJ, USA: Prentice-Hall, 1993.

[65] I. S. Pressman, "Matrices with multiple symmetry properties: Applications of centrohermitian and perhermitian matrices," *Linear Algebra Appl.*, vol. 284, nos. 1–3, pp. 239–258, Nov. 1998.

[66] R. M. Corless, G. H. Gonnet, D. E. Hare, D. J. Jeffrey, and D. E. Knuth, "On the Lambert-W function," *Adv. Comput. Math.*, vol. 5, no. 1, pp. 329–359, 1996.

[67] B. Mamandipoor, D. Ramasamy, and U. Madhow, "Newtonized orthogonal matching pursuit: Frequency estimation over the continuum," *IEEE Trans. Signal Process.*, vol. 64, no. 19, pp. 5066–5081, Oct. 2016.

[68] S. Kay, *Fundamentals of Statistical Signal Processing: Detection Theory*. Upper Saddle River, NJ, USA: Prentice-Hall, 1998.

[69] M. Xu, J. Zhu, J. Fang, and Z. Xu, "CFAR based NOMP for line spectral estimation and detection," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 59, no. 5, pp. 6971–6990, Jun. 2023.

[70] T. Wilding, S. Grebien, E. Leitinger, U. Muhlmann, and K. Witrisal, "Single-anchor, multipath-assisted indoor positioning with aliased antenna arrays," in *Proc. 52nd Asilomar Conf. Signals, Syst., Comput.*, Pacific Grove, CA, USA, Oct. 2018, pp. 525–531.

[71] H. L. Van Trees, *Optimum Array Processing: Part IV of Detection, Estimation, and Modulation Theory*. Hoboken, NJ, USA: Wiley, 2002.

[72] D. Schuhmacher, B.-T. Vo, and B.-N. Vo, "A consistent metric for performance evaluation of multi-object filters," *IEEE Trans. Signal Process.*, vol. 56, no. 8, pp. 3447–3457, Aug. 2008.

[73] P. Meissner, E. Leitinger, M. Lafer, and K. Witrisal. (2013). *MeasureMINT UWB Database*. [Online]. Available: http://www.spsc.tugraz.at/databases-and-tools/uwb-indoor-channel-experimental-data.html

[74] A. Molisch, *Wireless Communications*. Hoboken, NJ, USA: Wiley, 2005.

Stefan Grebien received the Ph.D. degree (Hons.) in electrical engineering from Graz University of Technology, Austria, in 2020.

From 2015 to 2021, he was with the Signal Processing and Speech Communication Laboratory, Graz University of Technology, and the Christian Doppler Laboratory for Location-Aware Electronic Systems. He is currently a Senior Researcher with the Intelligent Acoustics Solution Group, Joanneum Research Forschungsgesellschaft mbH, Graz. His research interests include estimation theory, Bayesian inference, and channel modeling.

Klaus Witrisal (Member, IEEE) received the Ph.D. degree (cum laude) from Delft University of Technology, Delft, The Netherlands, in 2002, and the Habilitation degree from Graz University of Technology in 2009. He is currently a Full Professor with Graz University of Technology and the Head of the Christian Doppler Laboratory for Location-Aware Electronic Systems, where he participates in a range of national and international research projects with industrial and academic partners. His current research interests include signal processing for wireless communications, propagation channel modeling, and positioning. He served as an Associate Editor for IEEE COMMUNICATIONS LETTERS and KICS/IEEE JOURNAL OF COMMUNICATIONS AND NETWORKS, the Co-Chair for the TWG "Indoor" of the COST Action IC1004 and the EWG "Localization and Tracking" of the COST Action CA15104, the Leading Chair of the IEEE Workshop on Advances in Network Localization and Navigation (ANLN), and the TPC (Co)-Chair of the Workshop on Positioning, Navigation and Communication (WPNC).

Bernard H. Fleury (Senior Member, IEEE) received the dual Diploma degrees in electrical engineering and in mathematics and the Ph.D. degree in electrical engineering from the Swiss Federal Institute of Technology in Zürich (ETHZ), Switzerland, in 1978, 1990, and 1990, respectively.

From 1997 to 2020, he was a Professor in communication theory with the Department of Electronic Systems, Aalborg University, Denmark. From 2000 to 2014, he was Head of Section, first with the Digital Signal Processing Section, and later with the Navigation and Communications Section. From 2006 to 2009, he was partly affiliated as a Key Researcher with the Telecommunications Research Center Vienna (FTW), Austria. From 1978 to 1985 and from 1992 to 1996, he was a Teaching Assistant and a Senior Research Associate, respectively, with the Communication Technology Laboratory, ETHZ. From 1988 to 1992, he was a Research Assistant with the Statistical Seminar, ETHZ. Since October 2020, he is a freelance academic teacher and a researcher and a technical consultant in his scientific fields of expertise. He is currently affiliated with the Institute of Telecommunications, Vienna University of Technology. He is engaged in a close cooperation with the Signal Processing and Speech Communication Laboratory, Graz University of Technology. He has developed, with his staff, a high-resolution method for the estimation of radio channel parameters that has found a wide application and has inspired similar estimation techniques both in academia and in industry. His research interests cover numerous aspects within communication theory, signal processing, and machine learning, mainly for wireless communication systems and networks. His current scientific activities include stochastic modeling and estimation of the radio channel, especially for large systems (operating in large bandwidths and equipped with large antenna arrays) deployed in harsh conditions (highly time-varying environments); iterative message-passing processing (with focus on the design of efficient feasible architectures for wireless receivers); localization techniques in wireless terrestrial systems; and radar signal processing. He has authored and coauthored more than 150 publications and is co-inventor of six filed or published patents in these areas.

Erik Leitinger (Member, IEEE) received the M.Sc. and Ph.D. degrees (Hons.) in electrical engineering from Graz University of Technology, Austria, in 2012 and 2016, respectively.

From 2016 to 2018, he was Post-Doctoral Researcher with the Department of Electrical and Information Technology, Lund University. He is currently a Senior Researcher with Graz University of Technology. His research interests include inference on graphs, statistical signal processing, high-dimensional and nonlinear estimation, localization and navigation, machine learning, stochastic modeling and estimation of radio channels, and estimation/detection theory. He is an Erwin Schrödinger Fellow. He served as the Co-Chair for the Special Session "Positioning Energy Constraint Devices" at the Asilomar Conference on Signals, Systems, and Computers 2020 and the Special Session "Synergistic Radar Signal Processing and Tracking" at the IEEE Radar Conference in 2021. He is an Associate Editor of IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS and *Journal of Advances in Information Fusion*.