

Blind Radio Tomography

Daniel Romero, *Member, IEEE*, Donghoon Lee, *Student Member, IEEE*,
and Georgios B. Giannakis, *Fellow, IEEE*

Abstract—From the attenuation measurements collected by a network of spatially distributed sensors, radio tomography constructs spatial loss fields (SLFs) that quantify absorption of radiofrequency waves at each location. These SLFs can be used for interference prediction in (possibly cognitive) wireless communication networks, for environmental monitoring or intrusion detection in surveillance applications, for through-the-wall imaging, for survivor localization after earthquakes or fires, and so on. The cornerstone of radio tomography is to model attenuation as the bidimensional integral of the SLF of interest scaled by a weight function. Unfortunately, existing approaches (i) rely on heuristic assumptions to select the weight function; and (ii) are limited to imaging changes in the propagation medium or they require a separate calibration step with measurements in free space. The first major contribution in this paper addresses (i) by means of a blind radio tomographic approach that learns the SLF together with the aforementioned weight function from the attenuation measurements. This challenging problem is tackled by capitalizing on contemporary kernel-based learning tools together with various forms of regularization that leverage prior knowledge. The second contribution addresses (ii) by means of a novel calibration technique capable of imaging static structures without separate calibration steps. Numerical tests with real and synthetic measurements validate the efficacy of the proposed algorithms.

Index Terms—radio tomography, tomographic imaging, channel-gain cartography, kernel-based learning.

I. INTRODUCTION

Tomographic imaging enjoys extensive popularity and widespread usage in natural sciences, notably in medical imaging [1]. The principles underpinning tomographic methods have been carried over to construct *spatial loss fields* (SLFs), which are maps quantifying the attenuation experienced by electromagnetic waves in radio frequency bands at every spatial position [2]. To this end, pairs of collaborating sensors deployed across the area of interest estimate the attenuation introduced by the channel between them. Different from traditional techniques, radio tomography relies on *incoherent* measurements, meaning that no phase information is available. This simplification saves the costs incurred by the accurate synchronization necessary to acquire phase differences among waveforms received at different sensors.

This work was supported by NSF grants 1343248, 1442686, 1508993, 1509040 as well as the FRIPRO TOPPFORSK grant WISECART 250910/F20 from the Research Council of Norway. Parts of this work have been presented at the IEEE Global Conference on Signal and Information Processing, Greater Washington, D.C. (USA), Dec. 7-9, 2016. Code for the simulations is available at the author's websites.

D. Romero was with the Dept. of ECE and Digital Tech. Center, Univ. of Minnesota, USA. He is now with the Dept. of Information and Communication Technology, Univ. of Agder, Norway. E-mail: daniel.romero@uia.no. D. Lee and G. B. Giannakis are with the Dept. of ECE and Digital Tech. Center, Univ. of Minnesota, USA. E-mails: {leex6962,georgios}@umn.edu.

SLFs are instrumental in a number of problems, including radio tomographic imaging [3], channel-gain cartography [4], and device-free passive localization [2], [5], [6]. Specifically, the absorption mapped by SLFs allows one to discern objects in space, thus enabling radio tomographic imaging. The latter is of interest in environmental monitoring for surveillance or intrusion detection [7]. Relative to existing alternatives using camera sensors, radio tomographic imaging features lower hardware costs and benefits from the ability of radio frequency waves to penetrate physical structures such as trees or buildings. The latter characteristic also renders radio tomography appealing for through-the-wall imaging [8], [9], which finds multiple military and civilian applications including security and responding to emergency situations. For example, these techniques may enable the police or emergency services to locate persons in burning buildings, survivors in rescue operations, or kidnappers in hostage situations. Similarly, SLFs are also useful in channel-gain cartography, where the goal is to predict the channel attenuation for links between arbitrary pairs of locations where no sensors are deployed [4]. Channel-gain maps obtained from SLFs solve the classical problem of predicting the interference inflicted to receivers that never transmit, as necessary in cognitive radio and for unlicensed access to television broadcasting systems [10]–[13], where the non-collaborative nature of primary users precludes any direct form of channel estimation between secondary transmitters and primary receivers. Further applications of these channel-gain maps include network planning or interference management in cellular networks.

The fundamental principle underlying radio tomography is that closely located radio links exhibit similar shadowing due to the presence of common obstructions. This correlation is related to the geometry of the propagation environment by the model in [14], [15], which prescribes that the attenuation due to shadowing is proportional to the line integral of a bidimensional SLF. Inspired by this model, [2], [3], [16] proposed various techniques for radio tomographic imaging. Since these techniques avoid calibration issues by estimating the difference between the SLF at consecutive time instants instead of the SLF itself, they reveal the location of changes in the propagation medium but are unable to image static structures. Similarly, [9] builds on the arguments in [17] to replace the SLF with an indicator function of the voxels that contain objects in motion and therefore also suffers from this limitation. In contrast, the scheme in [7] estimates the SLF directly and therefore can image static structures, but involves a separate calibration stage where these structures are absent.

On the other hand, existing radio tomography approaches approximate the aforementioned line integral by a discretized version of the bidimensional integral of the SLF scaled by a

function that quantifies the impact of the absorption at each spatial point on the shadowing attenuation of each link. In [2], [3], [9], [16], such a weight function is the indicator of an ellipse with foci at the transmitter and receiver positions and can be used to account for attenuation introduced by obstructions inside the first Fresnel ellipsoid. Building upon these works, [7] proposes a more sophisticated weight function that assigns a higher weight to obstructions that lie close to the straight path between transmitter and receiver. However, since the choice of these weight functions relies on heuristic arguments, one expects that the performance of radio tomographic methods can be improved upon adopting weight functions that capture the actual propagation phenomena more accurately.

A different body of literature applies radio tomography for channel-gain cartography. Remarkably, [18] applies the radio tomographic model in [15] to motivate a linear time-evolution model for channel-gain cartography. A more explicit application of radio tomography is reported in [19], where the weight function is constant on its elliptic support.

To sum up, all tomography works so far adopt a heuristically selected weight function. In contrast, the main contribution of this paper comprises three *blind* estimators that simultaneously learn the weight function and the SLF from the sensor measurements, therefore suppressing the need for heuristic considerations. The learned weight function reveals insightful information about propagation in the medium of interest and opens the door to validating models such as those in [17] and [7]. As noted in [3], the *non-blind* radio tomography problem, where the weight function is given and only the SLF has to be estimated, is intrinsically ill-posed. Hence, the *blind* problem addressed here is “doubly” challenging, since besides the SLF one has to estimate the weight function. To cope with these difficulties, the weight function is estimated here through kernel-based learning, a framework that is extensively popular due to its simplicity, universality, and because it leads to computationally efficient algorithms [20], [21]. On the other hand, the SLF is estimated using three alternative regularization criteria.

The second contribution is a novel calibration technique that simultaneously estimates the antenna gains and path loss exponent together with the SLF and weight function. Different from existing alternatives, the proposed method can image static structures and does not need a separate calibration stage where obstacles are removed from the propagation medium. Therefore, the present work constitutes a significant step forward in most applications involving tomographic imaging, channel-gain cartography, and device-free localization.

The rest of the paper is organized as follows. Sec. II reviews the radio tomography model, states the problem, and defines the calibrated and uncalibrated scenarios. The novel blind algorithm is then derived for the calibrated scenario in Sec. III, and extended to the uncalibrated scenario in Sec. IV. Numerical tests with synthetic as well as real measurements are provided in Sec. V. Finally, Sec. VI summarizes the main conclusions.

Notation: Sets are represented with calligraphic letters, whereas bold uppercase (lowercase) letters denote matrices (column vectors). \mathbf{I}_T represents the $T \times T$ identity matrix

and $\mathbf{i}_t \in \mathbb{R}^T$ its t -th column. $\mathbf{0}_T$ is the $T \times T$ all-zero matrix. Superscript $(\cdot)^\top$ stands for transposition, \otimes for the Kronecker product, and $\|\cdot\|$ for the Euclidean norm. The vectorization of an $M \times N$ matrix $\mathbf{X} := [\mathbf{x}_1, \dots, \mathbf{x}_N]$ is given by $\text{vec}(\mathbf{X}) := [\mathbf{x}_1^\top, \dots, \mathbf{x}_N^\top]^\top$.

II. BACKGROUND AND PROBLEM STATEMENT

This section introduces the radio tomographic model and formulates both the non-blind and blind radio tomography problems. Although the exposition builds on the framework in [7], the proposed approaches readily carry over to the frameworks in [2], [3], [16] and in [9] just by replacing the SLF here with the difference of consecutive SLFs or with an indicator function of *voxels* containing moving objects.

Consider a bidimensional geographical area indexed by the closed and convex set $\mathcal{A} \subset \mathbb{R}^2$. After averaging out the effects of small-scale fading, the power gain between a transmitter located at $\mathbf{x} \in \mathcal{A}$ and a receiver located at $\mathbf{x}' \in \mathcal{A}$ is given in dB units by

$$g(\mathbf{x}, \mathbf{x}') = g_{\text{TX}}(\mathbf{x}) + g_{\text{RX}}(\mathbf{x}') - \gamma_0 10 \log_{10} \|\mathbf{x} - \mathbf{x}'\|_2 - s(\mathbf{x}, \mathbf{x}') \quad (1)$$

where $g_{\text{TX}}(\mathbf{x})$ (resp. $g_{\text{RX}}(\mathbf{x}')$) is the combined gain of the power amplifier (low-noise amplifier) and transmit (receive) antenna, assumed omnidirectional for simplicity, of the sensor at $\mathbf{x} \in \mathcal{A}$ ($\mathbf{x}' \in \mathcal{A}$); γ_0 is the pathloss exponent, and $s(\mathbf{x}, \mathbf{x}')$ is the attenuation due to shadow fading. All other constant factors have been absorbed into $g_{\text{TX}}(\mathbf{x})$ and $g_{\text{RX}}(\mathbf{x}')$. The radio tomographic model in [7], which generalizes that in [14] and [15], prescribes that

$$s(\mathbf{x}, \mathbf{x}') = \int_{\mathcal{A}} w(\mathbf{x}, \mathbf{x}', \tilde{\mathbf{x}}) f(\tilde{\mathbf{x}}) d\tilde{\mathbf{x}} \quad (2)$$

where $f: \mathcal{A} \rightarrow \mathbb{R}_+$ is the SLF and $w: \mathcal{A} \times \mathcal{A} \times \mathcal{A} \rightarrow \mathbb{R}_+$ is the weight function. Whereas $f(\tilde{\mathbf{x}})$ represents the absorption at location $\tilde{\mathbf{x}} \in \mathcal{A}$, the weight $w(\mathbf{x}, \mathbf{x}', \tilde{\mathbf{x}})$ quantifies the impact of the absorption at $\tilde{\mathbf{x}}$ on the attenuation between \mathbf{x} and \mathbf{x}' . Typically, function w confers a greater weight $w(\mathbf{x}, \mathbf{x}', \tilde{\mathbf{x}})$ to those locations $\tilde{\mathbf{x}}$ lying closer to the line segment between \mathbf{x} and \mathbf{x}' and its selection is described later in this section.

Equation (2) models how the nature and spatial distribution of obstructions in the propagation medium affect the attenuation between each pair of locations. Its relevance is twofold: first, as mentioned in Sec. I, f represents absorption across space and therefore it can be used for imaging; see Sec. V for examples. Second, if both w and f are known, the gain between any two points \mathbf{x} and \mathbf{x}' can be recovered through (1) and (2), which enables channel-gain cartography.

The goal of radio tomography is to estimate f . To this end, N sensors at locations $\{\mathbf{x}_1, \dots, \mathbf{x}_N\} \subset \mathcal{A}$ collaboratively obtain channel-gain measurements. Specifically, at time $t = 1, \dots, T$, sensors $n(t)$ and $n'(t)$ measure $g(\mathbf{x}_{n(t)}, \mathbf{x}_{n'(t)})$ e.g. through pilot sequences, where $n(t), n'(t) \in \{1, \dots, N\} \forall t$. These measurements can be expressed as $\tilde{g}_t = g(\mathbf{x}_{n(t)}, \mathbf{x}_{n'(t)}) + \epsilon_t$, $t = 1, \dots, T$, where ϵ_t stands for measurement error. It is instructive to consider first that $\{g_{\text{TX}}(\mathbf{x}_n)\}_{n=1}^N$, $\{g_{\text{RX}}(\mathbf{x}_n)\}_{n=1}^N$, and γ_0 are known. In such a

calibrated scenario, it follows from (1) that $\{\check{g}_t\}_{t=1}^T$ contain the same information as $\{\check{s}_t\}_{t=1}^T$, where

$$\begin{aligned} \check{s}_t &:= g_{\text{TX}}(\mathbf{x}_{n(t)}) + g_{\text{RX}}(\mathbf{x}_{n'(t)}) \\ &\quad - \gamma_0 10 \log_{10} \|\mathbf{x}_{n(t)} - \mathbf{x}_{n'(t)}\|_2 - \check{g}_t \\ &= s(\mathbf{x}_{n(t)}, \mathbf{x}_{n'(t)}) - \epsilon_t. \end{aligned} \quad (3)$$

Thus, the fusion center may use $\{\check{s}_t\}_{t=1}^T$ rather than $\{\check{g}_t\}_{t=1}^T$. Sec. IV will deal with the *uncalibrated* scenario, where $\{g_{\text{TX}}(\mathbf{x}_n)\}_{n=1}^N$, $\{g_{\text{RX}}(\mathbf{x}_n)\}_{n=1}^N$ or γ_0 are unknown.

So far, works on radio tomography have focused on the non-blind problem, where one estimates f given w as well as the measurements and radio locations $\{(\mathbf{x}_{n(t)}, \mathbf{x}_{n'(t)}, \check{s}_t)\}_{t=1}^T$. The rest of this section describes the selection of w for this problem and formulates its blind counterpart.

The radio tomographic model originally proposed in [14, eq. (4)] and [15, eq. (9)] is expressed in terms of a line integral as

$$s(\mathbf{x}, \mathbf{x}') = \frac{1}{\sqrt{\|\mathbf{x} - \mathbf{x}'\|_2}} \int_{\mathbf{x}}^{\mathbf{x}'} f(\tilde{\mathbf{x}}) d\tilde{\mathbf{x}} \quad (4)$$

and can be viewed as a special case of (2) upon setting

$$w(\mathbf{x}, \mathbf{x}', \tilde{\mathbf{x}}) = \frac{1}{\sqrt{\|\mathbf{x} - \mathbf{x}'\|_2}} \int_{\mathbf{x}}^{\mathbf{x}'} \delta(\|\tilde{\mathbf{x}} - \tilde{\mathbf{x}}\|_2) d\tilde{\mathbf{x}} \quad (5)$$

where δ stands for the Dirac delta. Informally, w in (5) is a function assigning a weight $1/\sqrt{\|\mathbf{x} - \mathbf{x}'\|_2}$ to the point $\tilde{\mathbf{x}}$ if it lies on the line segment between \mathbf{x} and \mathbf{x}' , and zero otherwise. Therefore, this model only accounts for the attenuation introduced by obstacles obstructing the line of sight.

However, objects that do not obstruct the line of sight can still introduce attenuation if they lie close to it. The approach in [3] and [16] captures this effect by assigning the weight $1/\sqrt{\|\mathbf{x} - \mathbf{x}'\|_2}$ to all locations $\tilde{\mathbf{x}}$ lying within an ellipse with foci at \mathbf{x} and \mathbf{x}' through the function

$$w(\mathbf{x}, \mathbf{x}', \tilde{\mathbf{x}}) := \begin{cases} 0 & \text{if } \|\mathbf{x} - \tilde{\mathbf{x}}\|_2 + \|\tilde{\mathbf{x}} - \mathbf{x}'\|_2 \\ & > \|\mathbf{x} - \mathbf{x}'\|_2 + \lambda/2, \\ 1/\sqrt{\|\mathbf{x} - \mathbf{x}'\|_2} & \text{otherwise} \end{cases} \quad (6)$$

where $\lambda > 0$ is selected by the user. The weight function in (6) is referred to as *normalized ellipse function* in [7], where λ is set to the carrier wavelength so that $w(\mathbf{x}, \mathbf{x}', \tilde{\mathbf{x}})$ for fixed \mathbf{x} and \mathbf{x}' becomes the indicator of the first Fresnel zone, defined by $\{\tilde{\mathbf{x}} \in \mathbb{R}^2 : \|\mathbf{x} - \tilde{\mathbf{x}}\|_2 + \|\tilde{\mathbf{x}} - \mathbf{x}'\|_2 \leq \|\mathbf{x} - \mathbf{x}'\|_2 + \lambda/2\}$.

Despite capturing propagation effects more accurately than (5), function (6) still assigns the same weight across all $\tilde{\mathbf{x}}$ within the first Fresnel zone, regardless of the distance from $\tilde{\mathbf{x}}$ to the direct path. Since the attenuation introduced by an obstacle is expected to be a decreasing function of this distance, [7] proposes the following *inverse area elliptical function*¹

$$w(\mathbf{x}, \mathbf{x}', \tilde{\mathbf{x}}) := \begin{cases} 0 & \text{if } \|\mathbf{x} - \tilde{\mathbf{x}}\|_2 + \|\tilde{\mathbf{x}} - \mathbf{x}'\|_2 > \|\mathbf{x} - \mathbf{x}'\|_2 + \lambda/2, \\ \frac{1}{4} & \text{otherwise} \\ \pi \zeta_{\beta}(\mathbf{x}, \mathbf{x}', \tilde{\mathbf{x}}) \sqrt{\|\mathbf{x} - \mathbf{x}'\|_2^2 + \zeta_{\beta}^2(\mathbf{x}, \mathbf{x}', \tilde{\mathbf{x}})} & \end{cases} \quad (7)$$

¹The disagreement between (7) and [7, eq. (19)] owes to typographical errors in [7] and to the fact that the problem [7, eq. (20)] is solved here in closed form.

where $\beta > 0$ is selected by the user and $\zeta_{\beta}(\mathbf{x}, \mathbf{x}', \tilde{\mathbf{x}}) := \max^{1/2}(\beta^2, (\|\mathbf{x} - \tilde{\mathbf{x}}\|_2 + \|\tilde{\mathbf{x}} - \mathbf{x}'\|_2)^2 - \|\mathbf{x} - \mathbf{x}'\|_2^2)$. Similar to (6), if $\tilde{\mathbf{x}}$ is out of the first Fresnel zone, then (7) prescribes a zero weight. Otherwise, if $\tilde{\mathbf{x}}$ lies inside a smaller ellipse with foci at \mathbf{x} and \mathbf{x}' and minor axis length β , then $w(\mathbf{x}, \mathbf{x}', \tilde{\mathbf{x}})$ equals the reciprocal of the area of that ellipse. Finally, if $\tilde{\mathbf{x}}$ lies outside of the smaller ellipse but inside the first Fresnel zone, then $w(\mathbf{x}, \mathbf{x}', \tilde{\mathbf{x}})$ equals the reciprocal of the area of the smallest ellipse containing $\tilde{\mathbf{x}}$ and having \mathbf{x} and \mathbf{x}' as foci. Although w in (7) is intuitively more accurate than its predecessors, the rationale behind its selection is heuristic and may not accurately capture real propagation phenomena. This idea is reinforced by noting that (7), as well as (6), is discontinuous on the boundary of the first Fresnel zone.

To bypass this need for heuristically selecting w , the goal of this paper is to learn w from the data $\{(\mathbf{x}_{n(t)}, \mathbf{x}_{n'(t)}, \check{s}_t)\}_{t=1}^T$. However, since f is generally unknown, the *blind* radio tomography problem involves learning w and f given just the measurements $\{(\mathbf{x}_{n(t)}, \mathbf{x}_{n'(t)}, \check{s}_t)\}_{t=1}^T$.

III. BLIND RADIO TOMOGRAPHY ESTIMATORS

As explained in Sec. II, existing radio tomography schemes estimate f from the measurements $\{(\mathbf{x}_{n(t)}, \mathbf{x}_{n'(t)}, \check{s}_t)\}_{t=1}^T$ after setting w based on heuristic arguments. In contrast, the present section proposes three estimators that obtain both f and w from $\{(\mathbf{x}_{n(t)}, \mathbf{x}_{n'(t)}, \check{s}_t)\}_{t=1}^T$. To this end, Secs. III-A, III-B, and III-C formulate the estimation problem as a generic optimization program that can accommodate different forms of prior information through regularization. Subsequently, Sec. III-D proposes three solvers for different regularizers and Sec. III-E introduces a computationally efficient approximation for large measurement records.

A. Blind radio tomography as a function estimation problem

In the radio tomography literature, the integral in (2) is approximated as

$$s(\mathbf{x}, \mathbf{x}') \simeq c \sum_{l=1}^L w(\mathbf{x}, \mathbf{x}', \tilde{\mathbf{x}}_l) f(\tilde{\mathbf{x}}_l) \quad (8)$$

where $\{\tilde{\mathbf{x}}_l\}_{l=1}^L$ is a grid of points in \mathcal{A} and c is a constant that can be set to unity without loss of generality by absorbing any scaling factor in f . It can be recognized from (8) that the shadowing value $s(\mathbf{x}, \mathbf{x}')$ depends on f only through its values at the grid points. Thus, it suffices to estimate the L entries of the real-valued vector $\mathbf{f} := [f(\tilde{\mathbf{x}}_1), \dots, f(\tilde{\mathbf{x}}_L)]^T$ rather than the function $f(\mathbf{x})$ for all \mathbf{x} . On the other hand, finding w is more challenging since $s(\mathbf{x}, \mathbf{x}')$ may be evaluated at arbitrary real-valued coordinate pairs $(\mathbf{x}, \mathbf{x}')$ and, therefore, $w(\mathbf{x}, \mathbf{x}', \tilde{\mathbf{x}})$ needs to be known for all $\mathbf{x}, \mathbf{x}' \in \mathcal{A}$. Therefore, estimating w does not boil down to a vector estimation problem; instead, w must be estimated as a function on $\mathcal{A}^3 \subset \mathbb{R}^6$.

Intuitively, the number of measurements required to estimate w with a target accuracy depends on the size of its domain, in this case \mathcal{A}^3 . Conversely, for a given number of measurements, one expects that the larger the domain of w is, the lower the quality of its estimate will be. For this reason, the rest of this

section presents two techniques to reduce the aforementioned problem of estimating a function on \mathcal{A}^3 into the problem of estimating a function on a smaller domain by exploiting the known structure of w . It is worth noting that, despite enhancing estimation performance, these techniques are not necessary for the methods in this paper, which can handle in principle any weight function defined on \mathcal{A}^3 .

The first technique relies on the assumption that w depends on $\{\mathbf{x}, \mathbf{x}', \tilde{\mathbf{x}}\}$ only through $\phi_1(\mathbf{x}, \mathbf{x}') := \|\mathbf{x} - \mathbf{x}'\|_2$ and $\phi_2(\mathbf{x}, \mathbf{x}', \tilde{\mathbf{x}}) := \|\mathbf{x} - \tilde{\mathbf{x}}\|_2 + \|\tilde{\mathbf{x}} - \mathbf{x}'\|_2$, which respectively denote the length of the line of sight from \mathbf{x} to \mathbf{x}' and path going through the intermediate point $\tilde{\mathbf{x}}$. In other words, such a weight function can be expressed as² $w(\mathbf{x}, \mathbf{x}', \tilde{\mathbf{x}}) = w(\phi(\mathbf{x}, \mathbf{x}', \tilde{\mathbf{x}}))$, where $\phi(\mathbf{x}, \mathbf{x}', \tilde{\mathbf{x}}) := [\phi_1(\mathbf{x}, \mathbf{x}'), \phi_2(\mathbf{x}, \mathbf{x}', \tilde{\mathbf{x}})]^\top$. Besides being intuitively reasonable, this assumption is satisfied by all weight functions in the literature; cf. Sec. II. For example, the weight function in (6) satisfies $w(\mathbf{x}, \mathbf{x}', \tilde{\mathbf{x}}) = w(\phi(\mathbf{x}, \mathbf{x}', \tilde{\mathbf{x}}))$ for

$$w(\phi) := \begin{cases} 0, & \text{if } \phi_2 > \phi_1 + \lambda/2 \\ 1/\sqrt{\phi_1} & \text{otherwise} \end{cases} \quad (9)$$

with $\phi := [\phi_1, \phi_2]^\top$. Similarly, the weight function in (7) can be expressed as $w(\mathbf{x}, \mathbf{x}', \tilde{\mathbf{x}}) = w(\phi(\mathbf{x}, \mathbf{x}', \tilde{\mathbf{x}}))$ for

$$w(\phi) := \begin{cases} 0, & \text{if } \phi_2 > \phi_1 + \lambda/2 \\ \min \left[\Omega(\phi_1, \phi_2), \Omega \left(\phi_1, \sqrt{\phi_1^2 + \beta^2} \right) \right] & \text{otherwise} \end{cases} \quad (10)$$

where $\Omega(\phi_1, \phi_2) := 4/(\pi\phi_2\sqrt{\phi_2^2 - \phi_1^2})$. Thus, it is reasonable to seek an estimate of w among the class of functions satisfying this assumption, thereby reducing the problem of estimating a function of 6 variables to that of estimating a function of only 2. More formally, the sought $w(\phi)$ will be defined for $\phi \in \mathcal{B}_0 := \{\phi \in \mathbb{R}^2 : 0 \leq \phi_1 \leq D, \phi_1 \leq \phi_2 \leq 2D\} \subset \mathbb{R}^2$, where $D := \sup_{\mathbf{x}, \mathbf{x}' \in \mathcal{A}} \|\mathbf{x} - \mathbf{x}'\|_2$ is the *diameter* of \mathcal{A} and the condition $\phi_1 \leq \phi_2$ follows from the triangle inequality. Since $\mathcal{B}_0 \subset \mathbb{R}^2$, this re-parameterization of w is expected to significantly reduce the number of measurements needed to attain a target estimation accuracy.

The second technique to reduce the size of the domain of w relies on the fact that this function is expected to take significant values only on a small region of the space, the rest being close to zero. For example, one may assume along the lines of (9) and (10) that $w(\mathbf{x}, \mathbf{x}', \tilde{\mathbf{x}})$ only takes non-zero values within a certain ellipsoid with foci at \mathbf{x} and \mathbf{x}' , such as the Fresnel ellipsoid. In terms of the re-parameterization in the previous paragraph, such an ellipsoid can be expressed as the set of $\tilde{\mathbf{x}}$ such that $\phi_2(\mathbf{x}, \mathbf{x}', \tilde{\mathbf{x}}) \leq \phi_1(\mathbf{x}, \mathbf{x}') + \lambda/2$, implying that one may confine the support of w to the reduced set $\mathcal{B} := \{\phi \in \mathbb{R}^2 : 0 \leq \phi_1 \leq D, \phi_1 \leq \phi_2 \leq 2D, \phi_2 \leq \phi_1 + \lambda/2\} \subset \mathcal{B}_0$. Besides this ellipsoid, the user can select further regions \mathcal{B} or ellipsoids. Although w is expected to take small values off the Fresnel ellipsoid, these values may be estimated by selecting a larger domain set, even $\mathcal{B} = \mathcal{B}_0$. However, this operation comes with a caveat: as expected,

²Although the symbol w is used to represent both functions $w(\mathbf{x}, \mathbf{x}', \tilde{\mathbf{x}})$ and $w(\phi)$, there is no ambiguity since the former function takes 6 scalar arguments whereas the latter takes 2.

the larger \mathcal{B} , the larger the number of measurements required to maintain the estimation performance. Therefore, the size of this region must be increased only if a sufficiently large number of measurements is given.

The two techniques introduced in this section are applied next to simplify (8). To this end, apply the re-parameterization prescribed by the first technique to obtain the shadowing attenuation of the t -th measurement from (8), which yields

$$s(\mathbf{x}_{n(t)}, \mathbf{x}_{n'(t)}) \simeq \sum_{l=1}^L w(\phi(\mathbf{x}_{n(t)}, \mathbf{x}_{n'(t)}, \tilde{\mathbf{x}}_l)) f(\tilde{\mathbf{x}}_l) \quad (11)$$

after absorbing c in f . The second technique, which confines the support of w to \mathcal{B} , allows a reduction in the number of summands in (11) by disregarding those with $\phi(\mathbf{x}_{n(t)}, \mathbf{x}_{n'(t)}, \tilde{\mathbf{x}}_l) \notin \mathcal{B}$ since they result in $w(\phi(\mathbf{x}_{n(t)}, \mathbf{x}_{n'(t)}, \tilde{\mathbf{x}}_l)) = 0$. For the t -th measurement define

$$\begin{aligned} \mathcal{L}_t &:= \{l : 1 \leq l \leq L, \phi(\mathbf{x}_{n(t)}, \mathbf{x}_{n'(t)}, \tilde{\mathbf{x}}_l) \in \mathcal{B}\} \\ &:= \{i_{t,1}, \dots, i_{t,L_t}\} \end{aligned} \quad (12)$$

as the set comprising the indices of the L_t grid points $\tilde{\mathbf{x}}_l$ for which $\phi(\mathbf{x}_{n(t)}, \mathbf{x}_{n'(t)}, \tilde{\mathbf{x}}_l)$ is in \mathcal{B} . With this notation, (11) becomes

$$s(\mathbf{x}_{n(t)}, \mathbf{x}_{n'(t)}) \simeq \sum_{l=1}^{L_t} w(\phi_{t,l}) f(\tilde{\mathbf{x}}_{i_{t,l}}) \quad (13)$$

where $\phi_{t,l} := \phi(\mathbf{x}_{n(t)}, \mathbf{x}_{n'(t)}, \tilde{\mathbf{x}}_{i_{t,l}})$, $l = 1, \dots, L_t$.

In short, expression (13), which was obtained through the re-parameterization and support confinement techniques in this section, will prove decisive to lower the computational complexity and improve the estimation performance of the estimators proposed in the rest of the paper.

B. Function estimation via kernel-based learning

Before formulating the blind radio tomography problem, this section reviews kernel-based learning, which is one of the most prominent frameworks for non-linear function estimation due to its simplicity, good performance, low computational complexity, and universality, in the sense that any continuous function vanishing at infinity can be learned with arbitrary accuracy under general conditions; see e.g. [21].

Kernel-based methods seek function estimates within large classes of functions termed *reproducing kernel Hilbert spaces* (RKHSs) and defined as

$$\mathcal{H} := \left\{ w(\phi) = \sum_{i=1}^{\infty} \alpha_i \kappa(\phi, \phi_i) : \alpha_i \in \mathbb{R}; \phi, \phi_i \in \mathcal{B} \forall i \right\}.$$

In this expression, $\kappa : \mathcal{B} \times \mathcal{B} \rightarrow \mathbb{R}$ denotes a *reproducing* kernel, which is a function satisfying two properties [20]: (i) it is symmetric, meaning that $\kappa(\phi_1, \phi_2) = \kappa(\phi_2, \phi_1)$, $\forall \phi_1, \phi_2 \in \mathcal{B}$; and (ii) it is *positive definite*, meaning that:

$$\sum_{i=1}^I \sum_{i'=1}^I \alpha_i \alpha_{i'} \kappa(\phi_i, \phi_{i'}) \geq 0,$$

$$\forall I > 0, \{\alpha_i\}_{i=1}^I \subset \mathbb{R}, \{\phi_i\}_{i=1}^I \subset \mathcal{B}.$$

A frequent choice of reproducing kernel is the so-called *Gaussian radial basis function*

$$\kappa(\phi, \phi') = \exp\left(-\frac{\|\phi - \phi'\|_2^2}{2\sigma_\kappa^2}\right) \quad (14)$$

where $\sigma_\kappa^2 > 0$ is a user-selected parameter. Being a Hilbert space, \mathcal{H} is endowed with an inner product and, consequently, a norm. Specifically, the norm of a function $w(\phi) = \sum_{i=1}^{\infty} \alpha_i \kappa(\phi, \phi_i) \in \mathcal{H}$ can be obtained through the reproducing kernel as

$$\|w\|_{\mathcal{H}}^2 = \sum_{i=1}^{\infty} \sum_{i'=1}^{\infty} \alpha_i \alpha_{i'} \kappa(\phi_i, \phi_{i'}) \quad (15)$$

and is used in kernel-based learning as a proxy for smoothness of w . Different from other function norms such as the well-known $\|w\|_2^2 := \int |w(\phi)|^2 d\phi$, the RKHS norm does not require (potentially multidimensional) integration. This constitutes a major benefit of adopting the RKHS framework.

Nonparametric kernel-based estimates are commonly sought as the minimizers of judiciously selected regularization criteria. For the present problem, this paper proposes jointly estimating \hat{w} and $\hat{\mathbf{f}}$ as the minimizers of

$$(P1) \quad \min_{w \in \mathcal{H}, \mathbf{f} \in \mathbb{R}^L} \frac{1}{T} \sum_{t=1}^T \left(\check{s}_t - \sum_{l=1}^{L_t} w(\phi_{t,l}) f(\tilde{\mathbf{x}}_{i_{t,l}}) \right)^2 + \mu_w \|w\|_{\mathcal{H}}^2 + \mu_f \rho(\mathbf{f}).$$

Here, the inner summation in the first term is the approximation (13) to $s(\mathbf{x}_{n(t)}, \mathbf{x}_{n'(t)})$. Therefore, the first term in (P1) penalizes estimates w and \mathbf{f} predicting shadowing values that differ from those observed, i.e. $\{\check{s}_t\}_{t=1}^T$. The second term limits overfitting by promoting smooth estimates for w , where the notion of smoothness is captured by the RKHS norm in (15). The convex regularizer $\rho(\mathbf{f})$, for which different choices will be investigated in Sec. III-D, promotes a certain known structure on \mathbf{f} . Finally, the regularization parameters $\mu_w > 0$ and $\mu_f > 0$ balance the trade-off between data fitting, smoothness of w , and compliance of \mathbf{f} with prior knowledge. These parameters can be selected by cross-validation; see e.g. [22, Sec. 1.3]. However, in practice, the extra computation time entailed by this approach is bypassed by fixing these parameters to values that exhibit acceptable performance in a broad collection of typical scenarios.

To sum up, this section formulated the blind radio tomography problem as the function estimation problem in (P1). The rest of the paper will deal with solving (P1).

C. Kernel-based estimate via the representer theorem

A solution to (P1) cannot be found in its present form by numerical means since it involves a search over the infinite dimensional space \mathcal{H} . To circumvent this issue, this section reformulates (P1) as an optimization problem in finitely many scalar variables.

To this end, one can invoke the *representer theorem* [20], [23], which establishes that the minimizer of (P1) with respect

to w admits the expansion

$$\hat{w}(\phi) = \sum_{t=1}^T \sum_{l=1}^{L_t} \alpha_{t,l} \kappa(\phi, \phi_{t,l}) \quad (16)$$

for some $\{\alpha_{t,l}\}_{t,l}$. In other words, although \mathcal{H} contains all functions of the form $w(\phi) = \sum_{i=1}^{\infty} \alpha_i \kappa(\phi, \phi_i)$ for arbitrary $\{\phi_i\}_{i=1}^{\infty} \subset \mathcal{B}$ and $\{\alpha_i\}_{i=1}^{\infty} \subset \mathbb{R}$, one can confine the search for an estimate to those functions of the form (16).

Clearly, after applying the representer theorem, finding the optimum w amounts to finding the optimum $\{\alpha_{t,l}\}_{t,l}$ in (16). As detailed below, these coefficients follow upon substituting (16) into (P1). To this end, let

$$\mathbf{K}_{t,t'} := \begin{bmatrix} \kappa(\phi_{t,1}, \phi_{t',1}) & \cdots & \kappa(\phi_{t,1}, \phi_{t',L_{t'}}) \\ \vdots & \ddots & \vdots \\ \kappa(\phi_{t,L_t}, \phi_{t',1}) & \cdots & \kappa(\phi_{t,L_t}, \phi_{t',L_{t'}}) \end{bmatrix} \quad (17)$$

and let $\mathbf{K}_t := [\mathbf{K}_{t,1}, \dots, \mathbf{K}_{t,T}] \in \mathbb{R}^{L_t \times \tilde{L}}$, where $\tilde{L} := \sum_{t=1}^T L_t$. With this notation and w as in (16), the inner summation in (P1) becomes

$$\begin{aligned} \sum_{l=1}^{L_t} \hat{w}(\phi_{t,l}) f(\tilde{\mathbf{x}}_{i_{t,l}}) &= \sum_{l=1}^{L_t} \sum_{t'=1}^T \sum_{l'=1}^{L_{t'}} f(\tilde{\mathbf{x}}_{i_{t,l}}) \kappa(\phi_{t,l}, \phi_{t',l'}) \alpha_{t',l'} \\ &= \mathbf{f}^\top \mathbf{\Psi}_t^\top \mathbf{K}_t \boldsymbol{\alpha} \end{aligned} \quad (18)$$

where $\boldsymbol{\alpha} := [\alpha_{1,1}, \alpha_{1,2}, \dots, \alpha_{1,L_1}, \alpha_{2,1}, \dots, \alpha_{T,L_T}]^\top \in \mathbb{R}^{\tilde{L}}$ and $\mathbf{\Psi}_t \in \{0, 1\}^{L_t \times \tilde{L}}$ is a matrix whose entries $(l', i_{t,l'})$, $l' = 1, \dots, L_t$, are set to one and the rest are set to zero. Matrix $\mathbf{\Psi}_t$ selects the entries of $\mathbf{f} := [f(\tilde{\mathbf{x}}_1), \dots, f(\tilde{\mathbf{x}}_L)]^\top$ with indices in \mathcal{L}_t , that is $\mathbf{\Psi}_t \mathbf{f} = [f(\tilde{\mathbf{x}}_{i_{t,1}}), \dots, f(\tilde{\mathbf{x}}_{i_{t,L_t}})]^\top$. Likewise, from (15) and (16), the norm in the second term of (P1) equals

$$\|\hat{w}\|_{\mathcal{H}}^2 = \sum_{t,t'=1}^T \sum_{l=1}^{L_t} \sum_{l'=1}^{L_{t'}} \alpha_{t,l} \kappa(\phi_{t,l}, \phi_{t',l'}) \alpha_{t',l'} = \boldsymbol{\alpha}^\top \mathbf{K} \boldsymbol{\alpha} \quad (19)$$

where $\mathbf{K} := [\mathbf{K}_1^\top, \dots, \mathbf{K}_T^\top]^\top \in \mathbb{R}^{\tilde{L} \times \tilde{L}}$.

Therefore, from (18) and (19), (P1) can be rewritten as

(P2)

$$\min_{\boldsymbol{\alpha}, \check{\mathbf{s}}} \frac{1}{T} \|\check{\mathbf{s}} - (\mathbf{I}_T \otimes \mathbf{f}^\top) \mathbf{\Psi}^\top \mathbf{K} \boldsymbol{\alpha}\|_2^2 + \mu_w \boldsymbol{\alpha}^\top \mathbf{K} \boldsymbol{\alpha} + \mu_f \rho(\mathbf{f})$$

where $\check{\mathbf{s}} := [\check{s}_1, \dots, \check{s}_T]^\top$ and $\mathbf{\Psi} := \text{diag}\{\mathbf{\Psi}_1, \dots, \mathbf{\Psi}_T\} \in \{0, 1\}^{\tilde{L} \times LT}$.

To sum up, this section reformulated (P1) as (P2). Upon solving (P2), which will be addressed in the rest of the paper, one directly obtains an estimate of f at the grid points $\{\tilde{\mathbf{x}}_l\}_{l=1}^L$ through \mathbf{f} . To recover the estimate of w , one just needs to substitute the entries of the $\boldsymbol{\alpha}$ solving (P2) into (16).

D. Numerical solvers for three regularizers

This section presents three solvers for (P2) with different choices of $\rho(\mathbf{f})$ that promote certain known structure in \mathbf{f} .

Although (P2) is not jointly convex in $\boldsymbol{\alpha}$ and \mathbf{f} , it is separately convex in each of these vectors. This motivates an

alternating minimization approach where, at step [S1], one minimizes (P2) with respect to α for fixed \mathbf{f} and, at step [S2], one minimizes (P2) with respect to \mathbf{f} for fixed α . More specifically, in [S1], one obtains the $(k+1)$ -st iterate $\alpha[k+1]$, $k = 0, 1, \dots$, as

$$\alpha[k+1] = \arg \min_{\alpha} \frac{1}{T} \|\tilde{\mathbf{s}} - \mathbf{A}[k]\alpha\|^2 + \mu_w \alpha^\top \mathbf{K} \alpha \quad (20)$$

where $\mathbf{A}[k] := (\mathbf{I}_T \otimes \mathbf{f}^\top[k]) \Psi^\top \mathbf{K} \in \mathbb{R}^{T \times \tilde{L}}$ depends on the k -th iterate $\mathbf{f}[k]$. From the first-order optimality conditions of (20), its minimizer can be found in closed form as

$$\alpha[k+1] = (\mathbf{A}^\top[k] \mathbf{A}[k] + \mu_w T \mathbf{K})^{-1} \mathbf{A}^\top[k] \tilde{\mathbf{s}}. \quad (21)$$

To obtain an update equation for [S2], note that

$$\begin{aligned} (\mathbf{I}_T \otimes \mathbf{f}^\top) \Psi^\top \mathbf{K} \alpha &= (\mathbf{I}_T \otimes \mathbf{f}^\top) \left(\sum_{t=1}^T \mathbf{i}_t \otimes \Psi_t^\top \mathbf{K}_t \alpha \right) \\ &= \sum_{t=1}^T (\mathbf{i}_t \otimes \alpha^\top \mathbf{K}_t^\top \Psi_t) \mathbf{f} \end{aligned} \quad (22)$$

where \mathbf{i}_t is the t -th column of \mathbf{I}_T . From (22), one can express the minimizer of (P2) with respect to \mathbf{f} as

$$\mathbf{f}[k+1] = \arg \min_{\mathbf{f}} \frac{1}{T} \|\tilde{\mathbf{s}} - \mathbf{B}[k+1]\mathbf{f}\|^2 + \mu_f \rho(\mathbf{f}) \quad (23)$$

where $\mathbf{B}[k] := \sum_{t=1}^T (\mathbf{i}_t \otimes \alpha^\top[k] \mathbf{K}_t^\top \Psi_t) \in \mathbb{R}^{T \times L}$. Different from (20), no closed-form solution for (23) necessarily exists. Secs. III-D1, III-D2, and III-D3 will present three solvers for (23) with three different choices of $\rho(\mathbf{f})$.

Table I summarizes the alternating minimization solver presented in this section. To minimize memory requirements and execution time, one should avoid explicit construction of Ψ by accounting for its sparsity in the operations involving this matrix. For instance, one can obtain $(\mathbf{I}_T \otimes \mathbf{f}^\top[k]) \Psi^\top$ in [S1] by selecting the appropriate columns of $\mathbf{I}_T \otimes \mathbf{f}^\top[k]$. Regarding computational complexity, [S1] requires $\mathcal{O}(\tilde{L}^3)$ operations whereas the number of operations in [S2] is determined by the selected regularizer.

1) *Tikhonov regularizer*: Following [3], one can adopt the Tikhonov regularizer $\rho(\mathbf{f}) = \mathbf{f}^\top \mathbf{Q} \mathbf{f}$ to promote certain forms of smoothness on \mathbf{f} by suitably selecting \mathbf{Q} . The simplest approach, which can also be adopted in absence of prior information on \mathbf{f} , is to set $\mathbf{Q} = \mathbf{I}_L$. The resulting regularizer $\rho(\mathbf{f}) = \|\mathbf{f}\|_2^2$ promotes smoothness since it heavily penalizes estimates where a small fraction of the entries of \mathbf{f} contain most of the energy $\|\mathbf{f}\|_2^2$. As a more sophisticated alternative, one may set \mathbf{Q} to be the inverse covariance matrix of \mathbf{f} if the latter is known [3], [7]. Such an approach is expected to yield good performance when the SLF does not change drastically over space, e.g. because the wave length is large relative to objects in the area of interest. A further alternative is to set $\rho(\mathbf{f}) = \mathbf{f}^\top \mathbf{Q} \mathbf{f} = \mathbf{f}^\top \mathbf{D}^\top \mathbf{D} \mathbf{f}$, where \mathbf{D} is a matrix approximating a differential operator; e.g. $\mathbf{D} := [\mathbf{D}_x^\top, (\mathbf{D}_y \mathbf{P})^\top]^\top$, with \mathbf{D}_x , \mathbf{D}_y , and \mathbf{P} as in Sec. III-D3; see also [16].

<p>Input: $\{(\mathbf{x}_{n(t)}, \mathbf{x}_{n'(t)}, \tilde{\mathbf{s}}_t)\}_{t=1}^T, \mu_f, \mu_w, \kappa$.</p> <p>1: For every $t = 1, \dots, T$, obtain \mathcal{L}_t in (12). $\phi_{t,l} := \phi(\mathbf{x}_{n(t)}, \mathbf{x}_{n'(t)}, \tilde{\mathbf{x}}_{i_{t,l}}), l = 1, \dots, L_t$.</p> <p>2: Form \mathbf{K} as described in Sec. III-C.</p> <p>3: Initialize $\mathbf{f}[0]$.</p> <p>4: For $k = 0, 1, \dots$ until convergence [S1] α update Compute $\mathbf{A}[k] = (\mathbf{I}_T \otimes \mathbf{f}^\top[k]) \Psi^\top \mathbf{K}$. Obtain $\alpha[k+1]$ via (21). [S2] \mathbf{f} update Compute $\mathbf{B}[k+1] := \sum_{t=1}^T (\mathbf{i}_t \otimes \alpha^\top[k+1] \mathbf{K}_t^\top \Psi_t)$. Obtain $\mathbf{f}[k+1]$ via subroutines in Tables II, III, or IV. 5: Set $\hat{w}(\phi) = \sum_{t=1}^T \sum_{l=1}^{L_t} \alpha_{t,l} [k+1] \kappa(\phi, \phi_{t,l})$. Return function $\hat{w}(\mathbf{x}, \mathbf{x}', \tilde{\mathbf{x}}) := \hat{w}(\phi(\mathbf{x}, \mathbf{x}', \tilde{\mathbf{x}}))$ and vector $\mathbf{f}[k]$.</p>

TABLE I: Proposed blind radio tomography algorithm. Note that Ψ and $\{\Psi_t\}_{t=1}^T$ need not be explicitly constructed.

<p>Input: $\mathbf{B}[k+1], \tilde{\mathbf{s}}, \mu_f, \mathbf{Q}$.</p> <p>1: $\mathbf{f}[k+1] =$ $[\mathbf{B}^\top[k+1] \mathbf{B}[k+1] + \mu_f T \mathbf{Q}]^{-1} \mathbf{B}^\top[k+1] \tilde{\mathbf{s}}$.</p> <p>Return $\mathbf{f}[k+1]$.</p>

TABLE II: Subroutine to update \mathbf{f} with Tikhonov regularization.

With the Tikhonov regularizer $\rho(\mathbf{f}) = \mathbf{f}^\top \mathbf{Q} \mathbf{f}$, the \mathbf{f} update in (23) can be expressed in closed form as

$$\begin{aligned} \mathbf{f}[k+1] &= \arg \min_{\mathbf{f}} \frac{1}{T} \|\tilde{\mathbf{s}} - \mathbf{B}[k+1]\mathbf{f}\|^2 + \mu_f \mathbf{f}^\top \mathbf{Q} \mathbf{f} \\ &= [\mathbf{B}^\top[k+1] \mathbf{B}[k+1] + \mu_f T \mathbf{Q}]^{-1} \mathbf{B}^\top[k+1] \tilde{\mathbf{s}}. \end{aligned} \quad (24)$$

Therefore, to obtain an estimate for α and \mathbf{f} , one just has to cyclically apply (21) and (24) until convergence, i.e., one has to execute the algorithm in Table I with the subroutine in Table II at step [S2]. The complexity of [S2] is therefore $\mathcal{O}(T^2 L + L^3)$.

2) *ℓ_1 -norm regularization*: When \mathbf{f} exhibits a sparse pattern, as occurs when the propagation medium comprises a reduced number of relatively small obstructions such as trees, or when trying to detect intruders in areas without obstacles, then one can adopt the sparsity-promoting ℓ_1 -norm regularizer $\rho(\mathbf{f}) = \|\mathbf{f}\|_1$ in (P2) [24]. In this case, (23) becomes

$$\mathbf{f}[k+1] = \arg \min_{\mathbf{f}} \frac{1}{T} \|\tilde{\mathbf{s}} - \mathbf{B}[k+1]\mathbf{f}\|_2^2 + \mu_f \|\mathbf{f}\|_1. \quad (25)$$

Expression (25) is an instance of the LASSO problem [25], which can be efficiently solved e.g. by the fast iterative shrinkage/thresholding algorithm (FISTA) [26].

To keep the computational complexity at a minimum, the proposed algorithm inexactly solves (25) per iteration by applying a single-pass coordinate-descent algorithm over each element of \mathbf{f} [27, Sec. 3.8.6]. To derive the update rule for \mathbf{f} , let $\mathbf{b}_l[k]$ denote the l -th column of $\mathbf{B}[k]$, and let $\mathbf{B}_{-l}[k]$ represent the submatrix of $\mathbf{B}[k]$ resulting from removing the l -th column. Similarly, f_l denotes the l -th entry of \mathbf{f} and \mathbf{f}_{-l} represents a subvector of \mathbf{f} with its l -th entry removed. By defining

<p>Input: $\mathbf{B}[k+1], \tilde{\mathbf{s}}, \mu_f, \mathbf{f}[k]$.</p> <p>1: For $l = 1, 2, \dots, L$</p> <p style="padding-left: 20px;">$\tilde{\mathbf{s}}_l[k] := \tilde{\mathbf{s}} - \mathbf{B}_{-l}[k]\mathbf{f}_{-l}[k]$.</p> <p style="padding-left: 20px;">$f_l[k+1] = \Gamma(\tilde{\mathbf{s}}_l^\top[k]\mathbf{b}_l[k]; \mu_f T/2) / \ \mathbf{b}_l[k]\ _2^2$</p> <p>Return $\mathbf{f}[k+1] = [f_1[k+1], \dots, f_L[k+1]]^\top$.</p>

 TABLE III: Subroutine to update \mathbf{f} with ℓ_1 -norm regularization.

$\tilde{\mathbf{s}}_l[k] := \tilde{\mathbf{s}} - \mathbf{B}_{-l}[k]\mathbf{f}_{-l}[k]$, the minimizer of the objective in (25) at iteration k for fixed \mathbf{f}_{-l} can be written as $f_l[k+1] = \arg \min_{f_l} (1/T)\|\tilde{\mathbf{s}}_l[k] - \mathbf{b}_l[k]f_l\|_2^2 + \mu_f|f_l|$ and can be solved in closed form as $f_l[k+1] = \Gamma(\tilde{\mathbf{s}}_l^\top[k]\mathbf{b}_l[k]; \mu_f T/2) / \|\mathbf{b}_l[k]\|_2^2$, where Γ is the soft-thresholding function defined as $\Gamma(s; \mu) := \text{sign}(s) \max\{0, |s| - \mu\}$. The \mathbf{f} -update for [S2] in the algorithm of Table I is summarized in Table III. The complexity of this subroutine is $\mathcal{O}(TL^2)$.

3) *Total variation regularization:* Following [3], one can adopt a total variation (TV) regularizer, which promotes sharp edges by penalizing non-sparse spatial variations in the estimate of \mathbf{f} [28]. This is useful in presence of solid obstacles with a relatively homogeneous absorption pattern across its volume, e.g., concrete pillars or walls. Denote by $\mathbf{F} \in \mathbb{R}^{L_x \times L_y}$ a matrix such that $\mathbf{f} = \text{vec}(\mathbf{F})$. Its (l, l') -th entry $f_{l, l'}$ corresponds to the value of f at the (l, l') -th point of a bidimensional grid resulting from a spatial arrangement of the points $\{\tilde{\mathbf{x}}_l\}_{l=1}^L$ in L_x rows and L_y columns across \mathcal{A} . The so-called ℓ_1 -based anisotropic TV is defined as

$$\text{TV}(\mathbf{F}) := \sum_{l=1}^{L_x} \sum_{l'=1}^{L_y-1} |f_{l, l'+1} - f_{l, l'}| + \sum_{l'=1}^{L_y} \sum_{l=1}^{L_x-1} |f_{l+1, l'} - f_{l, l'}|. \quad (26)$$

This regularizer adds the absolute differences of function values at grid points located consecutively along a row or a column. Thus, this regularizer promotes sparsity in the local differences and therefore promotes constant regions in the estimate of \mathbf{F} . For $\rho(\mathbf{f}) = \text{TV}(\mathbf{F})$, expression (23) becomes

$$\mathbf{f}[k+1] = \arg \min_{\mathbf{f}} \frac{1}{T} \|\tilde{\mathbf{s}} - \mathbf{B}[k+1]\mathbf{f}\|_2^2 + \mu_f \text{TV}(\mathbf{F}). \quad (27)$$

Efficiently solving (27) is challenging since the TV regularizer is not differentiable. TV problems are generally solved either through algorithms based on the iterative shrinkage-thresholding algorithm (ISTA) [29], or through the alternating direction method of multipliers (ADMM) [30], [31]. Unfortunately, ISTA-based algorithms for TV problems are typically complicated by proximal operations requiring inner loops with additional iterative methods such as the gradient projection (GP) algorithm. To circumvent this challenge, this work pursues an ADMM approach.

To simplify notation, express (26) in terms of \mathbf{f} as $\text{TV}(\mathbf{f}) := \|\mathbf{D}_x \mathbf{f}\|_1 + \|\mathbf{D}_y \mathbf{P} \mathbf{f}\|_1$, where \mathbf{P} is a permutation matrix such that $\mathbf{P} \mathbf{f} = \text{vec}(\mathbf{F}^\top)$, whereas $\mathbf{D}_x := \mathbf{I}_{L_y} \otimes \Delta_{L_x} \in \mathbb{R}^{L_y(L_x-1) \times L}$ and $\mathbf{D}_y := \mathbf{I}_{L_x} \otimes \Delta_{L_y} \in \mathbb{R}^{L_x(L_y-1) \times L}$

<p>Input: $\mathbf{f}[k], \gamma_x[k], \gamma_y[k], \mathbf{d}_x[k], \mathbf{d}_y[k], \mathbf{B}[k+1], \tilde{\mathbf{s}}, \mu_f, \nu, L_x$.</p> <p>[S2-1] Update dual variables:</p> <p>1: $\gamma_x[k+1] = \gamma_x[k] + \nu(\mathbf{D}_x \mathbf{f}[k] - \mathbf{d}_x[k])$</p> <p>2: $\gamma_y[k+1] = \gamma_y[k] + \nu(\mathbf{D}_y \mathbf{P} \mathbf{f}[k] - \mathbf{d}_y[k])$</p> <p>[S2-2] Update \mathbf{d}:</p> <p>3: $\mathbf{d}_x[k+1] = \Gamma(\mathbf{D}_x \mathbf{f}[k] + \gamma_x[k+1]/\nu; \mu_f/\nu)$</p> <p>4: $\mathbf{d}_y[k+1] = \Gamma(\mathbf{D}_y \mathbf{P} \mathbf{f}[k] + \gamma_y[k+1]/\nu; \mu_f/\nu)$</p> <p>[S2-3] Update \mathbf{f}:</p> <p>5: $\mathbf{f}[k+1] =$</p> $\left(\frac{2}{T} \mathbf{B}^\top[k+1] \mathbf{B}[k+1] + \nu \mathbf{D}_x^\top \mathbf{D}_x + \nu \mathbf{D}_y^\top \mathbf{D}_y \right)^{-1}$ $\times \left(\mathbf{D}_x^\top (\nu \mathbf{d}_x[k+1] - \gamma_x[k+1]) \right.$ $\left. + \mathbf{D}_y^\top (\nu \mathbf{d}_y[k+1] - \gamma_y[k+1]) + \frac{2}{T} \mathbf{B}^\top[k+1] \tilde{\mathbf{s}} \right).$ <p>Return $\mathbf{f}[k+1], \gamma_x[k+1], \gamma_y[k+1], \mathbf{d}_x[k+1], \mathbf{d}_y[k+1]$.</p>

 TABLE IV: Subroutine to update \mathbf{f} with TV regularization.

act as discrete gradient operators where Δ_l is an $(l-1) \times l$ matrix whose (i, j) -th entry is 1 if $i = j$; -1 if $j = i+1$; and 0 otherwise. The resulting ADMM algorithm, whose derivation is omitted due to lack of space, is presented in Table IV. This subroutine assumes that the algorithm in Table I initializes $\gamma_x[0], \mathbf{d}_x[0], \gamma_y[0]$, and $\mathbf{d}_y[0]$. One immediate possibility is to set all their entries to zero. Parameter ν is a user-selected step size. The complexity of this subroutine can easily be seen to be $\mathcal{O}(TL^2 + L^3)$. Note that steps 1-4 decouple across entries, which implies that they can be executed in parallel.

E. Approximation for a large number of measurements

The complexity of the algorithm in Table I is dominated by the inversion of the $\tilde{L} \times \tilde{L}$ matrix in (21), which requires $\mathcal{O}(\tilde{L}^3)$ operations. Applications demanding high-resolution estimates of f or w , and hence requiring large T , may therefore incur prohibitive complexity since $\tilde{L} := \sum_{t=1}^T L_t$ increases with T . To bypass such a bottleneck, this section presents a technique to approximate the solution to (P1) for large T at affordable computational complexity.

Observe that the size of the aforementioned $\tilde{L} \times \tilde{L}$ matrix is determined by the number of terms in the sum of (16), which, as dictated by the representer theorem, equals the number of different vectors $\{\phi_{t, l}\}_{l=1, t=1}^{L_t, T}$ where w is evaluated in the objective of (P1). Thus, the size of such a matrix would be reduced if w in (P1) were only evaluated at a reduced set of vectors $\{\bar{\phi}_l\}_{l=1}^{\tilde{L}}$, where $\tilde{L} \ll \tilde{L}$. To this end, approximate $w(\phi_{t, l}) \approx w(\bar{\phi}_{r(t, l)})$ in (P1), where $r(t, l)$ is such that $\bar{\phi}_{r(t, l)}$ is the best approximation of $\phi_{t, l}$ in $\{\bar{\phi}_l\}_{l=1}^{\tilde{L}}$, to obtain

$$\min_{w \in \mathcal{H}, \mathbf{f} \in \mathbb{R}^L} \frac{1}{T} \sum_{t=1}^T \left(\tilde{\mathbf{s}}_t - \sum_{l=1}^{L_t} w(\bar{\phi}_{r(t, l)}) f(\tilde{\mathbf{x}}_{i, l}) \right)^2 + \mu_w \|w\|_{\mathcal{H}}^2 + \mu_f \rho(\mathbf{f}). \quad (\text{P1}')$$

In this case, applying the representer theorem shows that the

minimizer of (P1') can be expressed as

$$\hat{w}(\phi) = \sum_{l=1}^{\bar{L}} \bar{\alpha}_l \kappa(\phi, \bar{\phi}_l) \quad (28)$$

for some $\{\bar{\alpha}_l\}_{l=1}^{\bar{L}}$. The number of summands in (28) is potentially much smaller than that in (16), which reduces the complexity of [S1] and therefore that of the algorithm in Table I.

Before presenting a solver for (P1'), investigating how to approximate the vectors in $\{\phi_{t,l}\}_{l=1,t=1}^{L_t,T}$ with those of $\{\bar{\phi}_l\}_{l=1}^{\bar{L}}$ is in order. To this end, suppose that $\{\bar{\phi}_l\}_{l=1}^{\bar{L}}$ are given. In this case, the best approximation of $\phi_{t,l}$ in the Euclidean distance sense is $\bar{\phi}_{r(t,l)}$, where $r(t,l) := \arg \min_{r \in \{1, \dots, \bar{L}\}} \|\phi_{t,l} - \bar{\phi}_r\|_2$. Thus, for given $\{\bar{\phi}_l\}_{l=1}^{\bar{L}}$, one can naturally quantify the quality of the approximation as the sum of errors $\sum_{t=1}^T \sum_{l=1}^{L_t} \|\phi_{t,l} - \bar{\phi}_{r(t,l)}\|_2^2$. Thus, if $\{\bar{\phi}_l\}_{l=1}^{\bar{L}}$ are not given, it would be prudent to choose the $\{\bar{\phi}_l\}_{l=1}^{\bar{L}}$ minimizing this total error. The resulting minimizers are the \bar{L} centroids of $\{\phi_{t,l}\}_{l=1,t=1}^{L_t,T}$ obtained through K-means with \bar{L} clusters [22, Sec. 9.1]. However, if running K-means on $\{\phi_{t,l}\}_{l=1,t=1}^{L_t,T}$ is too costly for the available computational resources, a fast alternative is to draw the vectors $\{\bar{\phi}_l\}_{l=1}^{\bar{L}}$ uniformly at random and without replacement from $\{\phi_{t,l}\}_{l=1,t=1}^{L_t,T}$.

The rest of this section describes how to adapt the algorithm in Table I to obtain a solver for (P1'). The first step is to recognize that all equations after (16) and all algorithms in Secs. III-C and III-D depend on $\{\phi_{t,l}\}_{l=1,t=1}^{L_t,T}$ only through $\mathbf{K}_{t,t'}$ in (17) and its concatenations \mathbf{K}_t and \mathbf{K} . Therefore, the sought solver arises by replacing $\phi_{t,l}$ with $\bar{\phi}_{r(t,l)}$, $l = 1, \dots, L_t$, $t = 1, \dots, T$, in the definitions of these matrices. Specifically, the (l, l') -th entry of $\mathbf{K}_{t,t'}$, which according to (17) is given by $\kappa(\phi_{t,l}, \phi_{t',l'})$, must be replaced with $\kappa(\bar{\phi}_{r(t,l)}, \bar{\phi}_{r(t',l')})$. In matrix form, this is equivalent to replacing $\mathbf{K}_{t,t'}$ with $\mathbf{R}_t \bar{\mathbf{K}} \mathbf{R}_{t'}^\top$, where $\bar{\mathbf{K}}$ is an $\bar{L} \times \bar{L}$ matrix whose (\bar{l}, \bar{l}') element is $\kappa(\bar{\phi}_{\bar{l}}, \bar{\phi}_{\bar{l}'})$ and \mathbf{R}_t is an $L_t \times \bar{L}$ matrix whose l -th row has a one at the $r(t,l)$ -th column and zeros elsewhere. Likewise, \mathbf{K}_t must be replaced with $\mathbf{R}_t \bar{\mathbf{K}} \mathbf{R}^\top$, where $\mathbf{R} := [\mathbf{R}_1^\top, \dots, \mathbf{R}_T^\top]^\top \in \mathbb{R}^{\bar{L} \times \bar{L}}$, and \mathbf{K} with $\mathbf{R} \bar{\mathbf{K}} \mathbf{R}^\top$.

By applying these substitutions and letting $\bar{\alpha} := [\bar{\alpha}_1, \dots, \bar{\alpha}_{\bar{L}}]^\top := \mathbf{R}^\top \alpha$, problem (P2) becomes

$$(P2') \quad \min_{\bar{\alpha}, \mathbf{f}} \frac{1}{T} \|\bar{s} - (\mathbf{I}_T \otimes \mathbf{f}^\top) \Psi^\top \mathbf{R} \bar{\mathbf{K}} \bar{\alpha}\|_2^2 + \mu_w \bar{\alpha}^\top \bar{\mathbf{K}} \bar{\alpha} + \mu_f \rho(\mathbf{f}).$$

Whereas (P2) involves $\bar{L} + L$ variables, (P2') only has $\bar{L} + L$, a potentially much smaller number that confirms the complexity reduction stemming from the approximation in this section.

The alternating minimization algorithm proposed in Sec. III-D to solve (P2) readily carries over to solve (P2'). Specifically, the update in [S1] can be obtained from (20) by replacing $\alpha[k+1]$ with $\bar{\alpha}[k+1]$, α with $\bar{\alpha}$, and $\mathbf{A}[k]$ with $\bar{\mathbf{A}}[k] := (\mathbf{I}_T \otimes \mathbf{f}^\top[k]) \Psi^\top \mathbf{R} \bar{\mathbf{K}}$. In this way, (21) becomes

$$\bar{\alpha}[k+1] = [\bar{\mathbf{A}}^\top[k] \bar{\mathbf{A}}[k] + \mu_w T \bar{\mathbf{K}}]^{-1} \bar{\mathbf{A}}^\top[k] \bar{s}. \quad (29)$$

Similarly, the update in [S2] can be obtained from (23) if $\mathbf{B}[k]$ is replaced with $\bar{\mathbf{B}}[k] := \sum_{t=1}^T (\mathbf{i}_t \otimes \bar{\alpha}^\top[k]) \bar{\mathbf{K}}^\top \mathbf{R}_t^\top \Psi_t$.

Input: $\{(\mathbf{x}_{n(t)}, \mathbf{x}'_{n(t)}, \tilde{s}_t)\}_{t=1}^T, \mu_f, \mu_w, \kappa, \bar{L}$.

- 1: **For every** $t = 1, \dots, T$, **obtain**
 \mathcal{L}_t in (12).
 $\phi_{t,l} := \phi(\mathbf{x}_{n(t)}, \mathbf{x}'_{n(t)}, \tilde{\mathbf{x}}_{i_{t,l}}), l = 1, \dots, L_t$.
- 2: $\{\bar{\phi}_l\}_{l=1}^{\bar{L}} = \text{clustering_algorithm}(\{\phi_{t,l}\}_{t,l}, \bar{L})$.
- 3: **Obtain** $\{r(t,l)\}_{l=1,t=1}^{L_t,T}$ using (III-E).
- 4: **Form** $\bar{\mathbf{K}}, \{\mathbf{R}_t\}_{t=1}^T$, and \mathbf{R} as described in Sec. III-E.
- 5: **Initialize** $\mathbf{f}[0]$.
- 6: **For** $k = 0, 1, \dots$ **until convergence**
[S1] α update
Compute $\bar{\mathbf{A}}[k] := (\mathbf{I}_T \otimes \mathbf{f}^\top[k]) \Psi^\top \mathbf{R} \bar{\mathbf{K}}$.
Obtain $\bar{\alpha}[k+1]$ via (29).
[S2] \mathbf{f} update
Compute $\bar{\mathbf{B}}[k+1] := \sum_{t=1}^T (\mathbf{i}_t \otimes \bar{\alpha}^\top[k+1]) \bar{\mathbf{K}}^\top \mathbf{R}_t^\top \Psi_t$.
Obtain $\mathbf{f}[k+1]$ via subroutines in Tables II, III, or IV.
7: Set $\hat{w}(\phi) = \sum_{l=1}^{\bar{L}} \bar{\alpha}_l[k+1] \kappa(\phi, \bar{\phi}_l)$.
Return function $\hat{w}(\mathbf{x}, \mathbf{x}', \tilde{\mathbf{x}}) := \hat{w}(\phi(\mathbf{x}, \mathbf{x}', \tilde{\mathbf{x}}))$ and vector $\mathbf{f}[k]$.

TABLE V: Proposed blind radio tomography algorithm implementing the complexity reduction approximation in Sec. III-E. Note that the sparse matrices Ψ , $\{\Psi_t\}_{t=1}^T$, $\{\mathbf{R}_t\}_{t=1}^T$, and \mathbf{R} need not be explicitly constructed.

Therefore, the subroutines in Tables II, III, and IV can be invoked with $\bar{\mathbf{B}}[k]$, rather than $\mathbf{B}[k]$, as input argument. After the optimum $\bar{\alpha}$ has been found through this modified iteration, w can be recovered through (28).

Table V summarizes the modified solver. Through the approximation in this section, (21) was replaced with (29). The latter involves inverting an $\bar{L} \times \bar{L}$ matrix, which is considerably smaller than the $\bar{L} \times \bar{L}$ matrix inverted in (21) and no longer increases with T . Thus, the approximation in this section reduced the computational complexity from $\mathcal{O}(\bar{L}^3)$ to $\mathcal{O}(\bar{L})$. Moreover, since the size of $\mathbf{B}[k]$ equals that of $\bar{\mathbf{B}}[k]$, the complexity of [S2] remains the same as in Table I.

IV. RADIO TOMOGRAPHY FROM UNCALIBRATED MEASUREMENTS

To simplify the presentation, Sec. III focused on the scenario with known $\{g_{\text{TX}}(\mathbf{x}_n)\}_{n=1}^N, \{g_{\text{RX}}(\mathbf{x}_n)\}_{n=1}^N$, and γ_0 , where one can obtain the equivalent set of measurements $\{\tilde{s}(\mathbf{x}_{n(t)}, \mathbf{x}'_{n(t)})\}_{t=1}^T$ using (3). Unfortunately, these gains and path loss exponent are difficult to determine accurately in practice. This section extends the method in Sec. III to accommodate the case with unknown $\{g_{\text{TX}}(\mathbf{x}_n)\}_{n=1}^N, \{g_{\text{RX}}(\mathbf{x}_n)\}_{n=1}^N$, and γ_0 , which will be referred to as the *uncalibrated* scenario.

Several approaches are available to handle uncalibrated measurements, including [2], [3], and [16], where gain measurements are replaced with their difference between two time instants. The unknown gains and path loss cancel out since they remain constant over time. Applying such an approach, it becomes possible to reconstruct the difference between the SLF at those time instants, which allows the detection of changes in the propagation environment produced e.g. by intruders, but does not allow for imaging static structures. A static structure is imaged in [7] by gathering two measurement sets. The first set is obtained before placing the structure and is used for estimating $\{g_{\text{TX}}(\mathbf{x}_n)\}_{n=1}^N, \{g_{\text{RX}}(\mathbf{x}_n)\}_{n=1}^N$, and γ_0 .

The second set is obtained after assembling the structure, and is used to estimate f for a postulated w . Unfortunately, in many cases it is not possible to remove a structure, e.g. a building, to calibrate the system. To circumvent the limitations of these approaches, this section proposes estimators for $\{g_{\text{TX}}(\mathbf{x}_n)\}_{n=1}^N$, $\{g_{\text{RX}}(\mathbf{x}_n)\}_{n=1}^N$, and γ_0 , as well as f and w , from the uncalibrated data $\{(\mathbf{x}_{n(t)}, \mathbf{x}_{n'(t)}, \check{y}_t)\}_{t=1}^T$, where $\check{y}_t = g(\mathbf{x}_{n(t)}, \mathbf{x}_{n'(t)}) + \epsilon_t$, $t = 1, \dots, T$; see Sec. II.

Let $\mathcal{G} := \{\gamma_1, \dots, \gamma_C\}$ denote the set of C different gains, either for transmission or reception, so that $g_{\text{TX}}(\mathbf{x}_n) \in \mathcal{G}$ and $g_{\text{RX}}(\mathbf{x}_n) \in \mathcal{G} \forall n$. One can always form this set as $\mathcal{G} = \{g_{\text{TX}}(\mathbf{x}_1), \dots, g_{\text{TX}}(\mathbf{x}_N), g_{\text{RX}}(\mathbf{x}_1), \dots, g_{\text{RX}}(\mathbf{x}_N)\}$ and the proposed algorithm will estimate these $2N$ gains. However, if for example all sensors are of the same manufacturer and model, one may assume that $g_{\text{TX}}(\mathbf{x}_1) = \dots = g_{\text{TX}}(\mathbf{x}_N) = \gamma_1$ and $g_{\text{RX}}(\mathbf{x}_1) = \dots = g_{\text{RX}}(\mathbf{x}_N) = \gamma_2$. In this case, \mathcal{G} contains just two elements, which implies that only two gains have to be estimated and the quality of the estimates will therefore be higher for a given T . Adding ϵ_t to both sides of (1) yields

$$\check{\mathbf{g}} = \mathbf{\Omega}\gamma_1 - \phi_1\gamma_0 - \check{\mathbf{s}} \quad (30)$$

where $\phi_1 := 10 \log_{10}([\phi_1(\mathbf{x}_{n(1)}, \mathbf{x}_{n'(1)}), \dots, \phi_1(\mathbf{x}_{n(T)}, \mathbf{x}_{n'(T)})]^\top)$, $\check{\mathbf{g}} := [\check{g}_1, \dots, \check{g}_T]^\top$, $\gamma_1 := [\gamma_1, \dots, \gamma_C]^\top$, and $\mathbf{\Omega} \in \mathbb{R}^{T \times C}$ is a matrix whose t -th row has ones at columns c_1 and c_2 , where c_1 and c_2 are such that $g_{\text{TX}}(\mathbf{x}_{n(t)}) = \gamma_{c_1}$ and $g_{\text{RX}}(\mathbf{x}_{n'(t)}) = \gamma_{c_2}$. It follows from (30) that $\check{\mathbf{s}} = \mathbf{\Omega}\gamma_1 - \phi_1\gamma_0 - \check{\mathbf{g}} = \tilde{\mathbf{\Omega}}\gamma - \check{\mathbf{g}}$, where $\tilde{\mathbf{\Omega}} := [-\phi_1, \mathbf{\Omega}]$ and $\gamma = [\gamma_0, \gamma_1, \dots, \gamma_C]^\top$. Then, (P2') can be reformulated as

$$(P2'') \quad \min_{\gamma, \check{\alpha}, \mathbf{f}} \frac{1}{T} \|\check{\mathbf{g}} - \tilde{\mathbf{\Omega}}\gamma + (\mathbf{I}_T \otimes \mathbf{f}^\top) \Psi^\top \mathbf{R} \bar{\mathbf{K}} \bar{\alpha}\|_2^2 + \mu_w \bar{\alpha}^\top \bar{\mathbf{K}} \bar{\alpha} + \mu_f \rho(\mathbf{f}).$$

Minimizing (P2'') with respect to γ , one obtains

$$\min_{\check{\alpha}, \mathbf{f}} \frac{1}{T} \|\mathbf{P}_{\tilde{\mathbf{\Omega}}}^\perp (\check{\mathbf{g}} + (\mathbf{I}_T \otimes \mathbf{f}^\top) \Psi^\top \mathbf{R} \bar{\mathbf{K}} \bar{\alpha})\|_2^2 + \mu_w \bar{\alpha}^\top \bar{\mathbf{K}} \bar{\alpha} + \mu_f \rho(\mathbf{f}) \quad (31)$$

where $\mathbf{P}_{\tilde{\mathbf{\Omega}}}^\perp := \mathbf{I}_T - \tilde{\mathbf{\Omega}}(\tilde{\mathbf{\Omega}}^\top \tilde{\mathbf{\Omega}})^{-1} \tilde{\mathbf{\Omega}}^\top$. Comparing (P2') with (31), it follows that the algorithm in Table V can be employed to solve (31) if one replaces $\check{\mathbf{s}}$ with $-\mathbf{P}_{\tilde{\mathbf{\Omega}}}^\perp \check{\mathbf{g}}$, $\mathbf{A}[k]$ with $\bar{\mathbf{A}}_{\tilde{\mathbf{\Omega}}}[k] := \mathbf{P}_{\tilde{\mathbf{\Omega}}}^\perp (\mathbf{I}_T \otimes \mathbf{f}^\top[k]) \Psi^\top \mathbf{R} \bar{\mathbf{K}}$, and $\bar{\mathbf{B}}[k]$ with $\bar{\mathbf{B}}_{\tilde{\mathbf{\Omega}}}[k] := \mathbf{P}_{\tilde{\mathbf{\Omega}}}^\perp \sum_{t=1}^T (\mathbf{i}_t \otimes \bar{\alpha}^\top[k]) \bar{\mathbf{K}}_t^\top \mathbf{R}_t^\top \Psi_t$.

From (P2') and (31), it follows that the price to be paid for working with uncalibrated data is that the information along the column span of $\tilde{\mathbf{\Omega}}$ is neglected. The dimension of this column span therefore embodies the uncertainty in the calibration. As expected, the larger this dimension, the larger the T required to attain a target estimation performance. Since this dimension increases with C , it is important to keep the cardinality of \mathcal{G} as low as possible. To this end, one can assign multiple sensors to the same $\gamma_{c_0} \in \mathcal{G}$. Although some error is incurred if the gains of these sensors are not exactly γ_{c_0} , the overall effect of reducing C may pay off.

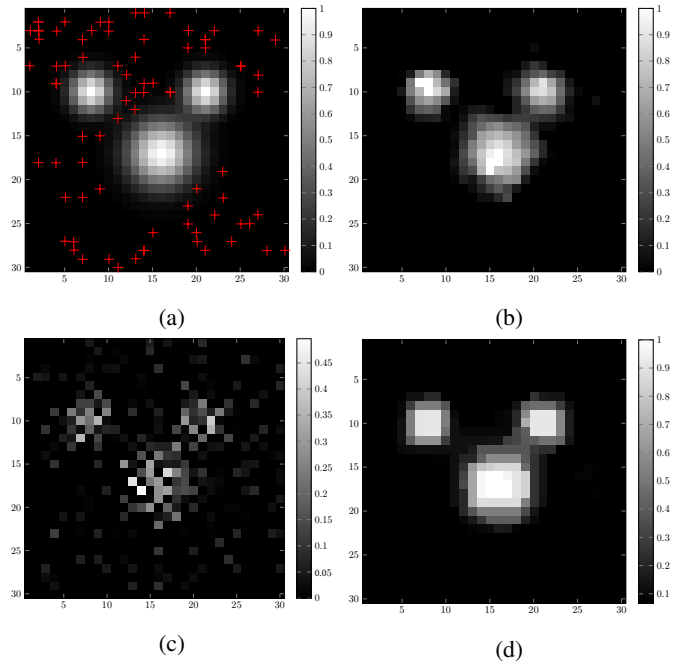


Fig. 1: (a) True SLF \mathbf{F} ; sensor locations are marked with red crosses. (b-d) Estimated $\hat{\mathbf{F}}$ using the algorithm in Table V ($\sigma_1^2 = 5 \times 10^{-2}$, $\sigma_2^2 = 7 \times 10^{-2}$, $\bar{L} = 2,500$, $T = 3,000$) with (b) Tikhonov regularization ($\mathbf{Q} = \mathbf{I}_{900}$, $\mu_f = 10^{-4}$, $\mu_w = 2 \times 10^{-2}$); (c) ℓ_1 -norm regularization ($\mu_f = 2 \times 10^{-4}$, $\mu_w = 9 \times 10^{-5}$); (d) TV regularization ($\mu_f = 4 \times 10^{-5}$, $\mu_w = 3 \times 10^{-4}$, $\rho = 1 \times 10^{-3}$).

V. NUMERICAL TESTS

This section demonstrates the benefits of the proposed algorithms through numerical tests with both synthetic and real measurements.

A. Tests with synthetic measurements

This section illustrates the effectiveness of the proposed estimators in exploiting prior information through the three regularizers in Sec. III-D. The test setup comprises a square area $\mathcal{A} = [0.5, 30.5]^2$ over which the grid $\{\tilde{\mathbf{x}}_l\}_{l=1}^{900} := \{1, \dots, 30\}^2$ of $L = 900$ points is defined. All experiments adopt the weight function in (10) with $\lambda = 0.39$ and $\beta = 1.1$, whereas a certain f will be specified per experiment. A total of $N = 80$ sensors were deployed uniformly at random outside of objects over \mathcal{A} at positions $\{\mathbf{x}_n\}_{n=1}^N$. The t -th measurement is obtained by the $n(t)$ -th and $n'(t)$ -th sensors, where $n(t)$ and $n'(t)$ are drawn per t uniformly at random without replacement from $\{1, \dots, N\}$. To focus on the impact of regularization, the effects of calibration are not accounted for in this section, where $\{g_{\text{TX}}(\mathbf{x}_n)\}_{n=1}^N$, $\{g_{\text{RX}}(\mathbf{x}_n)\}_{n=1}^N$, and γ_0 are assumed known, implying that the fusion center uses the ‘‘shadowing’’ measurements $\{\check{s}_t\}_{t=1}^T$. The latter are generated as $\check{s}_t := s(\mathbf{x}_{n(t)}, \mathbf{x}_{n'(t)}) - \epsilon_t$, where $s(\mathbf{x}_{n(t)}, \mathbf{x}_{n'(t)})$ is obtained through (11) and $\{\epsilon_t\}_{t=1}^T$ are independent zero-mean Gaussian random variables with variance $\sigma_\epsilon^2 = 10^{-2}$ unless otherwise stated.

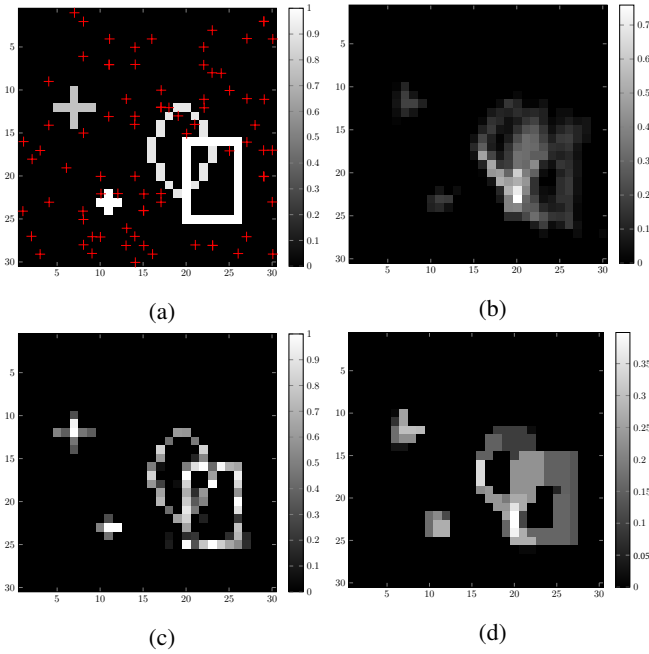


Fig. 2: (a) True SLF F ; sensor locations are marked with red crosses. (b-d) Estimated \hat{F} using the algorithm in Table V ($\sigma_1^2 = 9 \times 10^{-2}$, $\sigma_2^2 = 11 \times 10^{-2}$, $\bar{L} = 2,500$, $T = 3,000$) with (b) Tikhonov regularization ($\mathbf{Q} = \mathbf{I}_{900}$, $\mu_f = 4 \times 10^{-4}$, $\mu_w = 9 \times 10^{-2}$); (c) ℓ_1 -norm regularization ($\mu_f = 5 \times 10^{-5}$, $\mu_w = 1 \times 10^{-3}$); (d) TV regularization ($\mu_f = 1 \times 10^{-4}$, $\mu_w = 3.5 \times 10^{-4}$, $\rho = 5 \times 10^{-4}$).

Algorithm in Table V was tested with a Gaussian kernel $\kappa(\phi, \phi') = \exp(-(1/2)(\phi - \phi')^\top \text{diag}^{-1}\{\sigma_1^2, \sigma_2^2\}(\phi - \phi'))$ and for the three regularizers in Sec. III-D. To cope with the large number of variables, the support confinement technique in Sec. III-A is applied with \mathcal{B} equal to the first Fresnel zone. To simplify computations, the technique in Sec. III-E is applied by drawing the vectors $\{\bar{\phi}_l\}_{l=1}^{\bar{L}}$ uniformly at random and without replacement from $\{\phi_{t,l}\}_{l=1,t=1}^{L_t,T}$.

The rest of the section presents three simulations illustrating the importance of appropriately capturing prior information through the regularizers. Comparisons between the proposed estimator and non-blind alternatives will be provided later.

The first experiment adopts the smooth function $f(\mathbf{x}) = c_f \sum_{j=1}^3 \exp(-\|\mathbf{x} - \bar{\mathbf{x}}_j\|_2^2 / \sigma_{f,j}^2)$, where $\bar{\mathbf{x}}_1 = [15, 13]^\top$, $\bar{\mathbf{x}}_2 = [7, 20]^\top$, $\bar{\mathbf{x}}_3 = [20, 20]^\top$, $\sigma_{f,1}^2 = 13$, $\sigma_{f,2}^2 = \sigma_{f,3}^2 = 5$, and c_f is a constant ensuring that $\max_l f(\bar{\mathbf{x}}_l) = 1$. Fig. 1a depicts the values over the bidimensional grid of this function, whereas Figs. 1b, 1c, and 1d depict its estimates for the three regularizers. As expected, the SLF estimated through Tikhonov regularization is the most satisfactory in this case since it promotes smooth estimates, a property present in f .

In the second experiment, f was set to the sparse function on Fig. 2a. The estimated SLFs for the three regularizers are displayed in Figs. 2b, 2c, and 2d. In this case, the Tikhonov regularizer does not yield a good estimate since f is not smooth. In contrast, the ℓ_1 -norm regularizer leads to the most accurate estimate since it exploits the sparsity of f .

In the third experiment, f was set to the function with flat

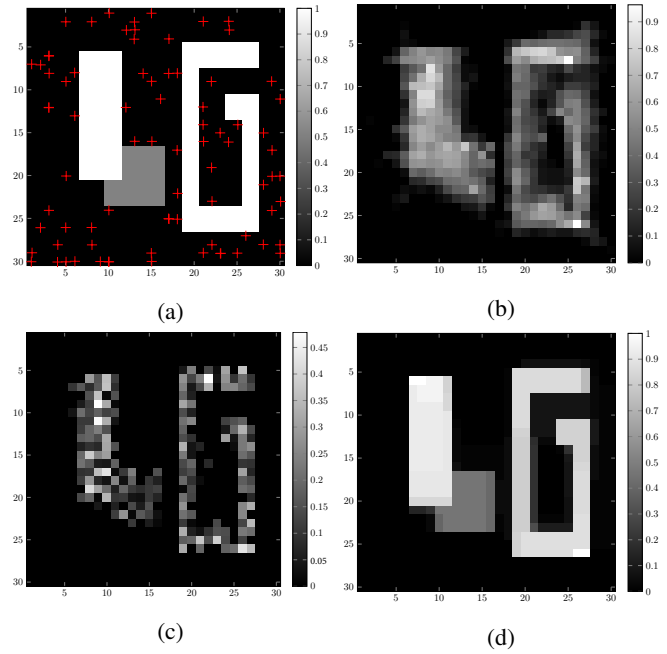


Fig. 3: (a) True SLF F ; sensor locations are marked with red crosses. (b-d) Estimated \hat{F} using the algorithm in Table V ($\sigma_1^2 = 8 \times 10^{-2}$, $\sigma_2^2 = 8.5 \times 10^{-2}$, $\bar{L} = 3,000$, $T = 3,000$) with (b) Tikhonov regularization ($\mathbf{Q} = \mathbf{I}_{900}$, $\mu_f = 4 \times 10^{-4}$, $\mu_w = 9 \times 10^{-2}$); (c) ℓ_1 -norm regularization ($\mu_f = 1 \times 10^{-4}$, $\mu_w = 9 \times 10^{-5}$); (d) TV regularization ($\mu_f = 5 \times 10^{-5}$, $\mu_w = 6 \times 10^{-4}$, $\rho = 5 \times 10^{-4}$).

regions and sharp edges on Fig. 3a. The estimates for the three regularizers are depicted in Figs. 3b, 3c, and 3d. As expected, the TV-regularizer leads to the best estimate since it promotes the kind of structure present in f .

To illustrate the influence of the number of measurements, the estimated SLF in the setup of Fig. 3d is depicted in Fig. 4 for different values of T . As expected, the quality improves for larger T . A visually satisfying estimate is already obtained for $T = 1,800$, which corresponds to $T/N = 22.5$ measurements per sensor on average. However, the needed T will be dictated by the specific application and how quality is quantified there.

To corroborate the ability of the proposed algorithms for learning $w(\phi_1, \phi_2)$, Fig. 5 compares the true $w(\phi_1, \phi_2)$ with $\hat{w}(\phi_1, \phi_2)$ obtained through the TV-regularized estimator in the same setting as Fig. 3. Note that such functions are only defined for $\phi_2 \geq \phi_1$ since the triangle inequality imposes that $\phi_2(\mathbf{x}, \mathbf{x}', \bar{\mathbf{x}}) \geq \phi_1(\mathbf{x}, \mathbf{x}')$ for all $(\mathbf{x}, \mathbf{x}', \bar{\mathbf{x}})$. The fit is satisfactory except in the vicinity of the points where w is discontinuous or non-smooth. These singularities are a consequence of the simplifications adopted in [7] to postulate (10). However, it is reasonable to expect that a function w accurately capturing the actual physics must be continuous and smooth. Hence, these discontinuities are just model artifacts, and the fitting error in those regions is thus not a limitation.

The next experiment investigates the robustness of the proposed algorithms against measurement noise ϵ_t . The normalized error $\|\hat{\mathbf{f}} - \mathbf{f}\|_2 / \|\mathbf{f}\|_2$ averaged over sensor locations and realizations of $\{\epsilon_t\}_{t=1}^T$ will be used to quantify estimation

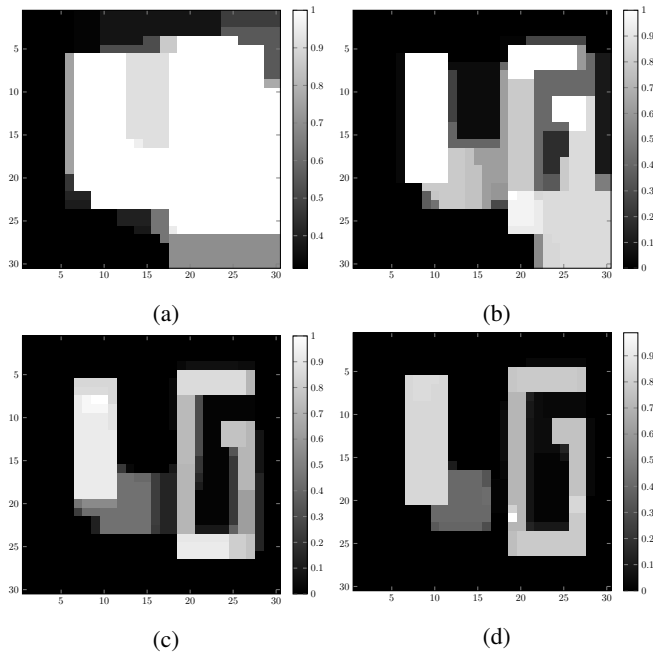


Fig. 4: Estimated $\hat{\mathbf{F}}$ using the algorithm in Table V ($\sigma_1^2 = 8 \times 10^{-2}$, $\sigma_2^2 = 8.5 \times 10^{-2}$, $\bar{L} = 3,000$), with TV regularization ($\mu_f = 1 \times 10^{-4}$, $\mu_w = 3.5 \times 10^{-4}$, $\rho = 5 \times 10^{-4}$). (a) $T = 1,400$; (b) $T = 1,600$; (c) $T = 1,800$; (d) $T = 2,000$.

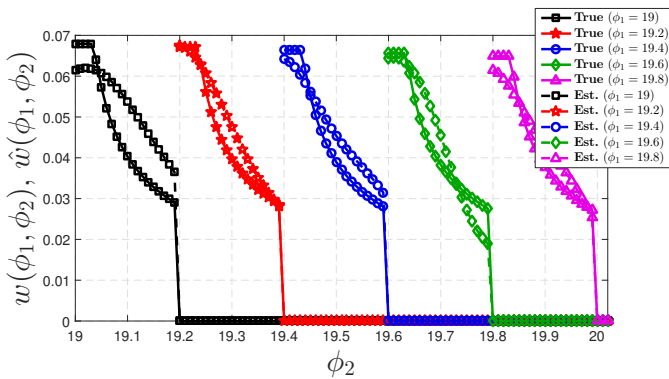


Fig. 5: True and estimated weight functions. Setting of Fig. 3d.

performance, where $\hat{\mathbf{f}}$ is the estimate of the true SLF \mathbf{f} . Fig. 6 depicts this error as a function of the standard deviation of ϵ_t in the setups of Figs. 1b, 2c, and 3d. Note that the latter figures correspond to the right endpoint of the x-axis of Fig. 6. Observe that the estimation performance is not meaningfully sensitive to the standard deviation of the measurement noise so long as the latter is sufficiently small. Moreover, Fig. 6 reveals that the noise power used in Figs. 1b, 2c, and 3d is significantly high; yet the SLF estimates there are of a visually good quality, which suggests that the proposed algorithms are reasonably robust to measurement noise.

The rest of this section investigates the performance of the proposed blind estimator in channel-gain cartography tasks. To this end, the same setting as in Figs. 3 and 5 is adopted. From the estimates $\hat{\mathbf{f}}$ and \hat{w} obtained through the algorithm in Table V with TV regularization, an estimate of the shadowing attenuation $\hat{s}(\mathbf{x}, \mathbf{x}')$ is obtained through (8) by replacing w and \mathbf{f} with their estimates. For comparison purposes, $s(\mathbf{x}, \mathbf{x}')$

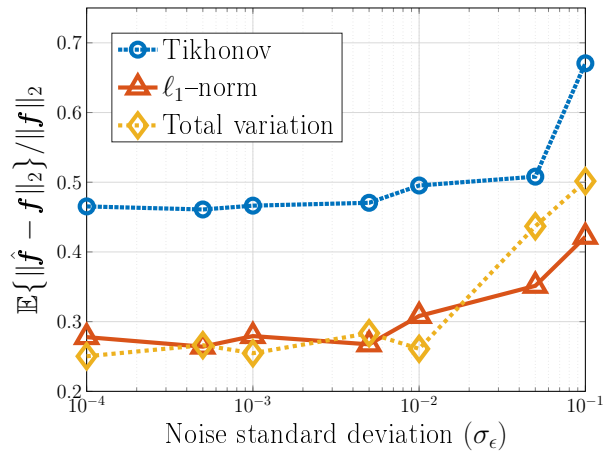


Fig. 6: Normalized error vs. noise standard deviation σ_ϵ obtained by the proposed blind algorithm with the (circle) Tikhonov regularization (setting of Fig. 1b); (triangle) ℓ_1 -norm regularization (setting of Fig. 2c); (diamond) TV regularization (setting of Fig. 3d).

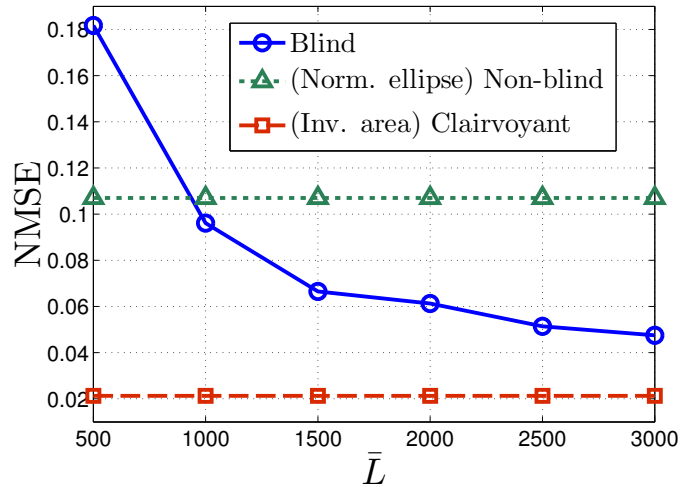


Fig. 7: NMSE vs. number \bar{L} of centroids obtained by (circle) the proposed blind algorithm with the TV regularizer (setting of Fig. 3d with $N = 200$); and (square, triangle) existing non-blind alternatives ($\mu_f = 5 \times 10^{-5}$, $\rho = 5 \times 10^{-4}$). The clairvoyant algorithm (square) adopts (10), which is the weight function used to generate the measurements.

is also estimated from (8), where $w(\phi)$ is pre-specified and \mathbf{f} is estimated through non-blind radio tomography as [3]

$$(P3) \quad \min_{\mathbf{f} \in \mathbb{R}^L} \frac{1}{T} \sum_{t=1}^T \left(\tilde{s}_t - \sum_{l=1}^{L_t} w(\phi_{t,l}) f(\tilde{\mathbf{x}}_{i_t,l}) \right)^2 + \mu_f \text{TV}(\mathbf{f})$$

which can be accomplished through the subroutine in Table IV.

Since $\{g_{\text{TX}}(\mathbf{x}_n)\}_{n=1}^N$, $\{g_{\text{RX}}(\mathbf{x}_n)\}_{n=1}^N$, and γ_0 are known, then knowing $s(\mathbf{x}, \mathbf{x}')$ amounts to knowing $g(\mathbf{x}, \mathbf{x}')$; cf. (1). This suggests adopting a performance metric quantifying error between $s(\mathbf{x}, \mathbf{x}')$ and $\hat{s}(\mathbf{x}, \mathbf{x}')$, such as the following normal-

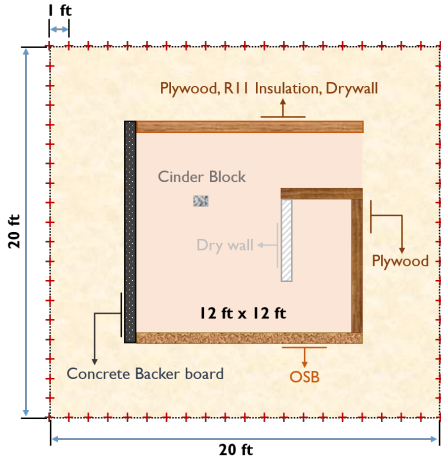


Fig. 8: Testbed configuration.

ized mean-square error (NMSE)

$$\text{NMSE} := \frac{\mathbb{E}\{\int_{\mathcal{A}} [s(\mathbf{x}, \mathbf{x}') - \hat{s}(\mathbf{x}, \mathbf{x}')]^2 d\mathbf{x}d\mathbf{x}'\}}{\mathbb{E}\{\int_{\mathcal{A}} s^2(\mathbf{x}, \mathbf{x}') d\mathbf{x}d\mathbf{x}'\}}$$

where the expectation is over the set $\{\mathbf{x}_n\}_{n=1}^N$ of sensor locations and realizations of $\{\epsilon_t\}_{t=1}^T$. Simulations estimated the expectations by averaging over 20 independent Monte Carlo runs. The integrals are approximated by averaging the integrand over 1,000 pairs $(\mathbf{x}, \mathbf{x}')$ chosen independently and uniformly at random over \mathcal{A} .

Fig. 7 compares the NMSE of the proposed blind algorithm with that of its non-blind counterparts, which assume the weight functions in (9) and (10). Since the measurements were generated using (10), the latter acts as a benchmark. Every point in the horizontal axis corresponds to a different value of the parameter \bar{L} , which in turn corresponds to a certain computational complexity of the blind algorithm. It is observed that the NMSE of the proposed algorithm approaches that of the clairvoyant estimator for sufficiently large \bar{L} . Moreover, assuming the wrong weight function in the non-blind algorithm incurs a five-fold error, thus motivating blind methods.

B. Tests with real measurements

This section validates the proposed estimators using the real data set in [7]. The test setup is depicted in Fig. 8, where $\mathcal{A} = [0.5, 20.5]^2$ is a square with sides of 20 feet (ft), over which a grid $\{\tilde{\mathbf{x}}_i\}_{i=1}^{961} := \{1, \dots, 31\}^2$ of $L = 961$ points is defined. A collection of 20 sensors measure the channel attenuation at 2.425 GHz between pairs of sensor positions, marked with the $N = 80$ crosses. Thus, although the number of actual sensors is 20, the effective number of sensors is $N = 80$. To estimate $\{g_{\text{TX}}(\mathbf{x}_n)\}_{n=1}^N$, $\{g_{\text{RX}}(\mathbf{x}_n)\}_{n=1}^N$, and γ_0 using the approach in [7], a first set of 2,400 measurements was obtained before placing the artificial structure in Fig. 8. Afterwards, the structure comprising one pillar and six walls of different materials is assembled and $T = 2,380$ measurements $\{\check{g}_t\}_{t=1}^T$ are obtained.

The proposed algorithm is tested with the same kernel, support confinement, and approximation technique for large number of measurements as in Sec. V-A. Following [3] and [7],

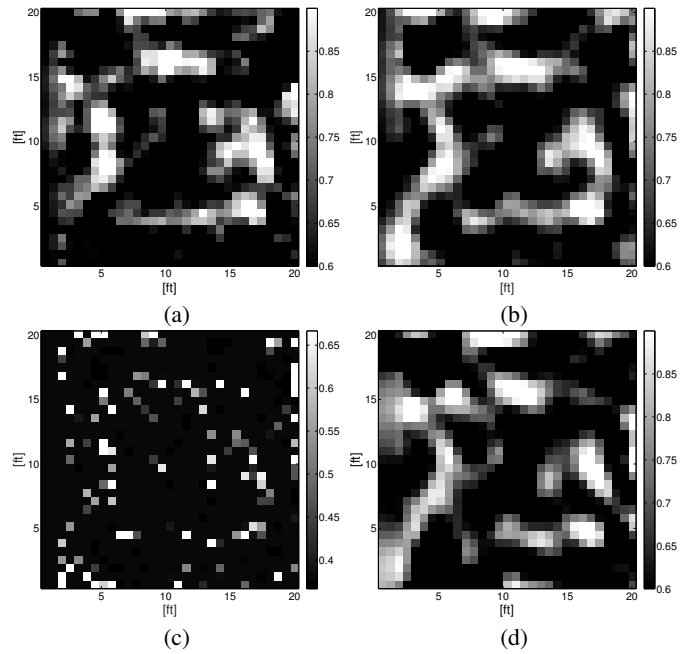


Fig. 9: Estimated $\hat{\mathbf{F}}$ via (a) the non-blind algorithm in [3], [7] and (b-d) the Algorithm in Table V ($\bar{L} = 2,500$) with (b) Tikhonov regularization ($\sigma_1^2 = 1 \times 10^{-1}$, $\sigma_2^2 = 1.5 \times 10^{-1}$, $\mu_f = 6 \times 10^{-2}$, $\mu_w = 3 \times 10^{-2}$); (c) ℓ_1 -norm regularization ($\sigma_1^2 = 8 \times 10^{-2}$, $\sigma_2^2 = 3.5 \times 10^{-1}$, $\mu_f = 6 \times 10^{-2}$, $\mu_w = 1.3 \times 10^{-3}$); (d) TV regularization ($\sigma_1^2 = 1 \times 10^{-1}$, $\sigma_2^2 = 3.8 \times 10^{-1}$, $\mu_f = 4.3$, $\mu_w = 1.3 \times 10^{-3}$, $\rho = 1 \times 10^{-3}$). The first set of measurements was used to estimate path loss and sensor gains.

the (l, l') -th entry of \mathbf{Q}^{-1} in the Tikhonv regularizer was set to $(\sigma_s^2 / \delta_s) \exp(-\|\tilde{\mathbf{x}}_l - \tilde{\mathbf{x}}_{l'}\|_2 / \delta_s)$, which is the covariance of $f(\tilde{\mathbf{x}}_l)$ and $f(\tilde{\mathbf{x}}_{l'})$ predicted by the exponential decay model in [14], for $\sigma_s^2 = 4.76$ and $\delta_s = 1$. An initial $\tilde{\alpha}[0]$ was obtained through kernel ridge regression [20] with the Gaussian kernel by fitting (9), and $\mathbf{f}[0]$ was subsequently obtained from $\tilde{\alpha}[0]$ through the subroutine associated with the selected $\rho(\mathbf{f})$.

The first simulation mimics the setup in [7]. The first set of measurements, i.e., the one obtained without the structure, is used to estimate $\{g_{\text{TX}}(\mathbf{x}_n)\}_{n=1}^N$, $\{g_{\text{RX}}(\mathbf{x}_n)\}_{n=1}^N$, and γ_0 using the approach in [7]. The calibrated measurements $\{\check{s}_t\}_{t=1}^T$ are obtained from $\{\check{g}_t\}_{t=1}^T$ by substituting these estimates into (3).

Fig. 9a displays the non-blind estimate of the SLF in [7], which essentially solves (P3) with $\text{TV}(\mathbf{f})$ replaced with $\mathbf{f}^\top \mathbf{Q} \mathbf{f}$, where \mathbf{Q} is as described earlier. Figs. 9b, 9c, and 9d depict the blind SLF estimates with Tikhonov, ℓ_1 -norm, and TV regularization, respectively. It is observed that both non-blind and the blind algorithms with Tikhonov and TV regularizers successfully recover the form of the artificial structure in the propagation medium, yet some artifacts are introduced possibly due to the presence of the sensors and their tripods. The proposed algorithms attain a reconstruction performance similar to the non-blind algorithm without any need for heuristic assumptions on w . Admittedly, the reconstructed SLF of the blind algorithms is not visually much better than that for the non-blind algorithm due to the low number of measurements. The reason is that the total number

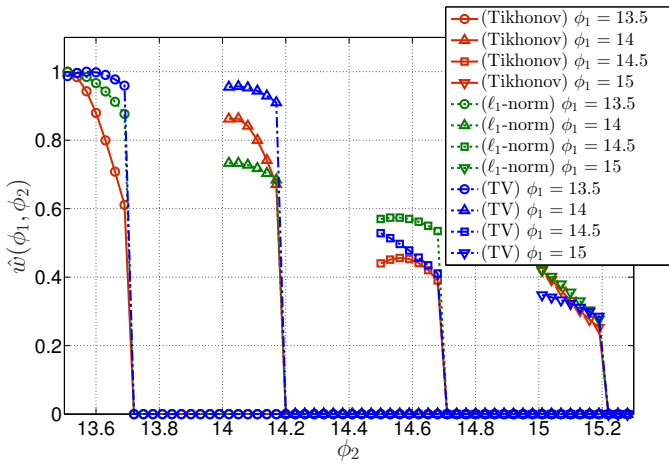


Fig. 10: Estimated weight functions with (red) Tikhonov, (green) ℓ_1 -norm, and (blue) TV regularization.

of unknowns is 2, 982 for the former and just 961 for the latter. On the contrary, for T sufficiently large, it is expected that the blind algorithm achieves a better estimation performance.

In the same setup, Fig. 10 shows the estimate of w obtained with the proposed algorithm. The estimated curves are satisfactory since they are smooth and approximately decreasing within their support. Unfortunately, having limited the support of \hat{w} to the first Fresnel ellipsoid prevents us from estimating the values of w off this ellipsoid. However, this limitation was imposed by the low number of measurements relative to the number of unknowns; one would be able to obtain more satisfactory estimates of w on larger supports if more measurements were available.

The second simulation assesses the performance of the proposed algorithm in the uncalibrated scenario; see Sec. II. In this scenario, the measurements in absence of the artificial structure in Fig. 8 are not used. In contrast, the proposed algorithm adopts the technique in Sec. IV with $g_{\text{TX}}(\mathbf{x}_1) = \dots = g_{\text{TX}}(\mathbf{x}_N) = \gamma_1$ and $g_{\text{RX}}(\mathbf{x}_1) = \dots = g_{\text{RX}}(\mathbf{x}_N) = \gamma_2$. Fig. 11 shows the proposed SLF estimates with the three regularizers. The estimate with TV regularization is visually acceptable, whereas for the Tikhonov regularizer, the result is similar to Figs. 9a and 9b, where twice more measurements were used. Therefore, the technique in Sec. IV suppresses the need for separate calibration stages in which the structure is not present while minimally sacrificing estimation performance.

The last simulation assesses the performance of the proposed algorithm and competing alternatives for channel-gain cartography. To compare with the algorithm in [7], the same set of shadowing measurements as in the first simulation of this section was used. Data $\{\tilde{s}_t\}_{t=1}^T$ was split into a training set with 80% of the measurements and a test set $\{\tilde{s}_t\}_{t \in \mathcal{E}}$ with the remaining 20%, where the indices in the set \mathcal{E} are drawn uniformly at random without replacement from $\{1, \dots, T\}$. Per Monte Carlo run, the proposed algorithm is executed and an estimate $\hat{s}(\mathbf{x}_n(t), \mathbf{x}_{n'}(t))$ is obtained per $t \in \mathcal{E}$ after substituting \hat{w} and \hat{f} into (13). Afterwards, $\hat{g}(\mathbf{x}_n(t), \mathbf{x}_{n'}(t))$ is obtained by substituting $\hat{s}(\mathbf{x}_n(t), \mathbf{x}_{n'}(t))$ as well as the estimates for $\{g_{\text{TX}}(\mathbf{x}_n)\}_{n=1}^N$, $\{g_{\text{RX}}(\mathbf{x}_n)\}_{n=1}^N$, and γ_0 from the

calibration stage into (1). The performance metric is

$$\text{NMSE} := \frac{\mathbb{E}\{\sum_{t \in \mathcal{E}} (\check{g}_t - \hat{g}(\mathbf{x}_n(t), \mathbf{x}_{n'}(t)))^2\}}{\mathbb{E}\{\sum_{t \in \mathcal{E}} \check{g}_t^2\}}$$

where the expectation is taken over realizations of \mathcal{E} .

Fig. 12 depicts the NMSE of the proposed blind algorithm with Tikhonov regularization along with that of the non-blind algorithm in [7] with w as in (6) and (7). The advantage of the proposed algorithm over its non-blind counterparts is manifest for sufficiently large \bar{L} . As before, greater differences are expected for larger measurement sets.

VI. CONCLUDING SUMMARY

This paper developed blind radio tomographic algorithms that simultaneously estimate the spatial loss field and weight functions of the radio tomographic model, which are of interest in imaging and channel-gain cartography applications. Although the problem is challenging, the usage of kernel-based learning with various regularizers accounting for prior knowledge together with several complexity reduction techniques resulted in an algorithm that needs no heuristic assumption on the weight function, can image static structures, and does not require separate calibration stages. The effectiveness of the novel algorithm was corroborated through synthetic- and real-data experiments. Future research will include online and distributed approaches to blind radio tomography.

REFERENCES

- [1] N. B. Smith and A. Webb, *Introduction to Medical Imaging: Physics, Engineering and Clinical Applications*, Cambridge University Press, 2010.
- [2] N. Patwari and P. Agrawal, "Effects of correlated shadowing: Connectivity, localization, and RF tomography," in *Int. Conf. Info. Process. Sensor Networks*, St. Louis, MO, Apr. 2008, pp. 82–93.
- [3] J. Wilson, N. Patwari, and O. G. Vasquez, "Regularization methods for radio tomographic imaging," in *Virginia Tech Symp. Wireless Personal Commun.*, Blacksburg, VA, Jun. 2009.
- [4] S.-J. Kim, E. Dall'Anese, and G. B. Giannakis, "Cooperative spectrum sensing for cognitive radios using kriged Kalman filtering," *IEEE J. Sel. Topics Sig. Process.*, vol. 5, no. 1, pp. 24–36, Feb. 2011.
- [5] K. Woyach, D. Puccinelli, and M. Haenggi, "Sensorless sensing in wireless networks: Implementation and measurements," in *Int. Symp. Modeling Optimization Mobile, Ad Hoc Wireless Netw.*, Apr. 2006, pp. 1–8.
- [6] N. Patwari and J. Wilson, "RF sensor networks for device-free localization: Measurements, models, and algorithms," *Proc. IEEE*, vol. 98, no. 11, pp. 1961–1973, Nov. 2010.
- [7] B. R. Hamilton, X. Ma, R. J. Baxley, and S. M. Matechik, "Propagation modeling for radio frequency tomography in wireless networks," *IEEE J. Sel. Topics Sig. Process.*, vol. 8, no. 1, pp. 55–65, Feb. 2014.
- [8] E. J. Baranoski, "Through-wall imaging: Historical perspective and future directions," *J. Franklin Institute*, vol. 345, no. 6, pp. 556–569, Sep. 2008.
- [9] J. Wilson and N. Patwari, "See-through walls: Motion tracking using variance-based radio tomography networks," *IEEE Trans. Mobile Comput.*, vol. 10, no. 5, pp. 612–621, May 2011.
- [10] E. Axell, G. Leus, and E. G. Larsson, "Overview of spectrum sensing for cognitive radio," in *Proc. Cognitive Inf. Process.*, 2010, pp. 322–327.
- [11] Q. Zhao and B. M. Sadler, "A survey of dynamic spectrum access," *IEEE Sig. Process. Mag.*, vol. 24, no. 3, pp. 79–89, 2007.
- [12] Federal Communications Commission, "FCC 11-131," *Unlicensed Operation in the TV Broadcast Bands*, 2011.
- [13] S.-J. Kim, E. Dall'Anese, J. A. Bazerque, K. Rajawat, and G. B. Giannakis, "Advances in spectrum sensing and cross-layer design for cognitive radio networks," in *Academic Press Library in Signal Processing: Communications and Radar Signal Processing: 2*, chapter 9, pp. 471–497. Academic Press, 2013.

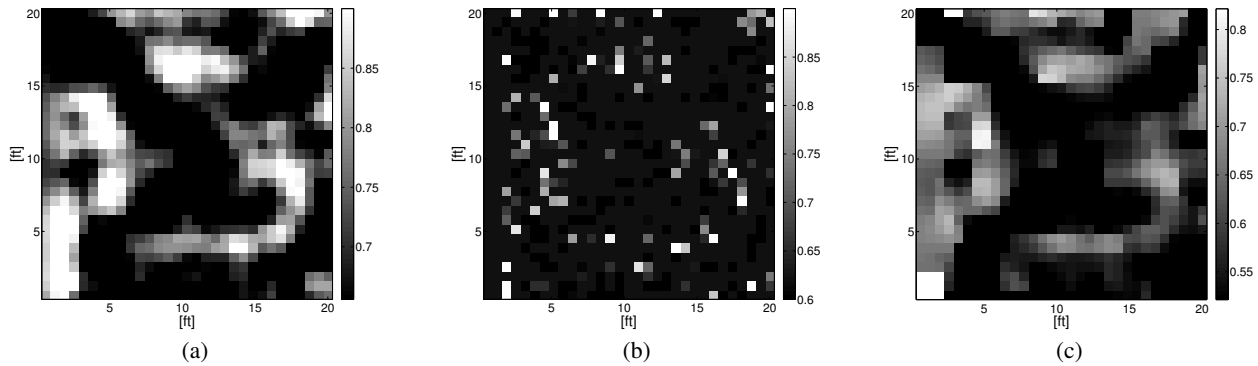


Fig. 11: Estimated $\hat{\mathbf{F}}$ via the algorithm in Sec. IV ($\bar{L} = 2,000$) with (a) Tikhonov regularization ($\sigma_1^2 = 8 \times 10^{-2}$, $\sigma_2^2 = 3.5 \times 10^{-1}$, $\mu_f = 3.2 \times 10^{-1}$, $\mu_w = 5 \times 10^{-3}$); (b) ℓ_1 -norm regularization ($\sigma_1^2 = 8 \times 10^{-2}$, $\sigma_2^2 = 3.5 \times 10^{-1}$, $\mu_f = 3 \times 10^{-3}$, $\mu_w = 1.3 \times 10^{-1}$); and (c) TV regularization ($\sigma_1^2 = 0.1$, $\sigma_2^2 = 3.8 \times 10^{-1}$, $\mu_f = 9.5 \times 10^{-3}$, $\mu_w = 3.3 \times 10^{-3}$, $\rho = 1 \times 10^{-3}$).

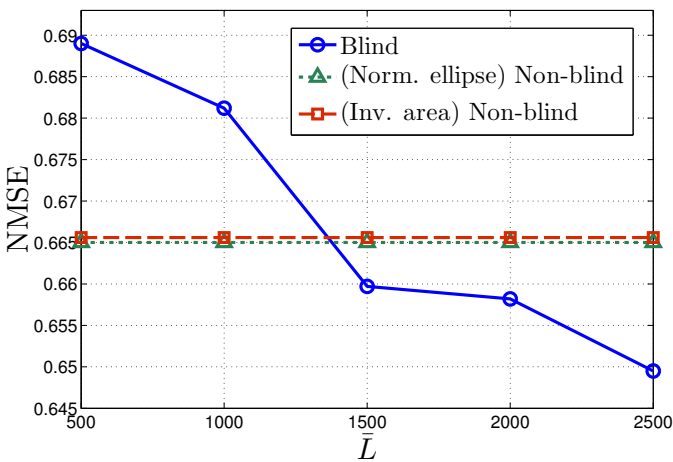


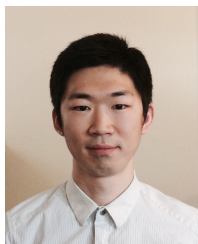
Fig. 12: NMSE vs. number of centroids \bar{L} obtained by (bold) the blind algorithm with the Tikhonov regularizer ($\sigma_1^2 = 9 \times 10^{-2}$, $\sigma_2^2 = 1.5 \times 10^{-1}$, $\mu_f = 3 \times 10^{-2}$, $\mu_w = 2 \times 10^{-2}$); and the non-blind algorithm [3], [7] with (dotted) the normalized ellipse ($\mu_f = 1 \times 10^{-2}$); and (dashed) inverse area elliptical ($\mu_f = 2 \times 10^{-3}$) models.

- [14] N. Patwari and P. Agrawal, "NeSh: A joint shadowing model for links in a multi-hop network," in *Proc. IEEE Int. Conf. Acoust., Speech, Sig. Process.*, Las Vegas, NV, Mar. 2008, pp. 2873–2876.
- [15] P. Agrawal and N. Patwari, "Correlated link shadow fading in multi-hop wireless networks," *IEEE Trans. Wireless Commun.*, vol. 8, no. 9, pp. 4024–4036, Aug. 2009.
- [16] J. Wilson and N. Patwari, "Radio tomographic imaging with wireless networks," *IEEE Trans. Mobile Comput.*, vol. 9, no. 5, pp. 621–632, 2010.
- [17] N. Patwari and J. Wilson, "Spatial models for human motion-induced signal strength variance on static links," *IEEE Trans. Inf. Forensics Security*, vol. 6, no. 3, pp. 791–802, Sep. 2011.
- [18] E. Dall'Anese, S.-J. Kim, and G. B. Giannakis, "Channel gain map tracking via distributed kriging," *IEEE Trans. Veh. Technol.*, vol. 60, no. 3, pp. 1205–1211, 2011.
- [19] D. Lee, S.-J. Kim, and G. B. Giannakis, "Channel gain cartography for cognitive radios leveraging low rank and sparsity," *IEEE Trans. Wireless Commun.*, vol. 16, no. 11, pp. 5953–5966, Nov. 2017.
- [20] B. Schölkopf and A. J. Smola, *Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond*, MIT Press, 2002.
- [21] C. Carmeli, E. De Vito, A. Toigo, and V. Umanita, "Vector valued reproducing kernel Hilbert spaces and universality," *Analysis Appl.*, vol. 8, no. 1, pp. 19–61, 2010.
- [22] C. M. Bishop, *Pattern Recognition and Machine Learning*, Information Science and Statistics. Springer, 2006.

- [23] B. Schölkopf, R. Herbrich, and A. J. Smola, "A generalized representer theorem," in *Computational Learning Theory*. Springer, 2001, pp. 416–426.
- [24] M. A. Kansa and M. G. Rabbat, "Compressed RF tomography for wireless sensor networks: Centralized and decentralized approaches," in *Int. Conf. Distributed Comput. Sensor Syst.*, Marina del Rey, CA, 2009, Springer, pp. 173–186.
- [25] R. Tibshirani, "Regression shrinkage and selection via the lasso," *J. Roy. Statist. Soc. Ser. B*, vol. 58, no. 1, pp. 267–288, 1996.
- [26] A. Beck and M. Teboulle, "A fast iterative shrinkage-thresholding algorithm for linear inverse problems," *SIAM J. Imaging Sci.*, vol. 2, no. 1, pp. 183–202, 2009.
- [27] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning, 2nd ed.*, Springer, New York, 2009.
- [28] L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Physica D*, vol. 60, no. 1–4, pp. 259–268, Nov. 1992.
- [29] A. Beck and M. Teboulle, "Fast gradient-based algorithms for constrained total variation image denoising and deblurring problems," *IEEE Trans. Image Process.*, vol. 18, no. 11, pp. 2419–2434, Nov. 2009.
- [30] G. B. Giannakis, Q. Ling, G. Mateos, I. D. Schizas, and H. Zhu, "Decentralized learning for wireless communications and networking," in *Splitting Methods in Communication, Imaging, Science, and Engineering*, pp. 461–497. Springer, 2016.
- [31] Z. Qin, D. Goldfarb, and S. Ma, "An alternating direction method for total variation denoising," *Optim. Method Softw.*, vol. 30, no. 3, pp. 594–615, 2015.



D. Romero (M'16) received his M.Sc. and Ph.D. degrees in Signal Theory and Communications from the University of Vigo, Spain, in 2011 and 2015, respectively. From Jul. 2015 to Nov. 2016, he was a post-doctoral researcher with the Digital Technology Center and Department of Electrical and Computer Engineering, University of Minnesota, USA. In Dec. 2016, he joined the Department of Information and Communication Technology, University of Agder, Norway, as an associate professor. His main research interests lie in the areas of machine learning, optimization, signal processing, and communications.



Donghoon Lee (S'12) received his B.Eng. and M.Eng. degrees in electrical engineering from Korea University in Seoul, Korea in 2010 and 2013, respectively. Currently, he is working towards the Ph.D. degree with the Department of Electrical and Computer Engineering (ECE) at the University of Minnesota (UMN), Twin Cities. From 2010 to 2011, he worked for the Korea Institute of Science and Technology (KIST) in Seoul, Korea.

His current research interests include statistical signal processing and machine learning, with applications to wireless communications and networking.



Georgios B. Giannakis (Fellow'97) received his Diploma in Electrical Engr. from the Ntl. Tech. Univ. of Athens, Greece, 1981. From 1982 to 1986 he was with the Univ. of Southern California (USC), where he received his MSc. in Electrical Engineering, 1983, MSc. in Mathematics, 1986, and Ph.D. in Electrical Engr., 1986. He was with the University of Virginia from 1987 to 1998, and since 1999 he has been a professor with the Univ. of Minnesota, where he holds an Endowed Chair in Wireless Telecommunications, a University of Minnesota McKnight

Presidential Chair in ECE, and serves as director of the Digital Technology Center.

His general interests span the areas of communications, networking and statistical learning - subjects on which he has published more than 400 journal papers, 700 conference papers, 25 book chapters, two edited books and two research monographs (h-index 128). Current research focuses on learning from Big Data, wireless cognitive radios, and network science with applications to social, brain, and power networks with renewables. He is the (co-) inventor of 32 patents issued, and the (co-) recipient of 9 best journal paper awards from the IEEE Signal Processing (SP) and Communications Societies, including the G. Marconi Prize Paper Award in Wireless Communications. He also received Technical Achievement Awards from the SP Society (2000), from EURASIP (2005), a Young Faculty Teaching Award, the G. W. Taylor Award for Distinguished Research from the University of Minnesota, and the IEEE Fourier Technical Field Award (inaugural recipient in 2015). He is a Fellow of EURASIP, and has served the IEEE in a number of posts, including that of a Distinguished Lecturer for the IEEE-SP Society.