Spatial Loss Field for Outdoor Micro-cells

Julian D. Villegas G., *Member, IEEE,* and Claude Oestges, *Fellow, IEEE*

Abstract

In propagation analysis it is assumed that existing shadowing losses are due to obstacles through the propagation path between two transceivers. Although classically modelled as a log-normal random variable, special attention is being recently put into models where shadowing is considered as a consequence of an Underlying Spatial Loss Field (USLF) accounting for attenuation, under Radio Channel Tomography (RCT) framework. In this paper we address the task of making shadowing predictions for transceiver locations, without the need of measurements for those locations. A spatial loss field estimate is produced from measurements for a given set of transceiver locations, and used as an input to predict the shadowing at another set of transceiver locations. Moreover, new weighting functions are proposed and developed from channel parameter estimates of the propagation paths.

Index Terms

shadowing, spatial loss field, measurements.

I. INTRODUCTION

Path Loss estimation is indispensable for link budget analysis. Losses are typically modelled as composed by three terms: distance-related losses, short-term fading losses and losses due to obstacles (shadowing). Knowledge about shadowing losses is required for the computation of performance metrics as the Blocking Probability Error (BP) [1]. A traditional approach to model the shadowing consider it as a zero-mean log-normal random variable in time and space, whose time and space correlation is isotropic. In that approach, shadowing losses are fully characterized by the variance of the gaussian random variable (in the dB scale), and the de-correlation distance, a parameter that measures how fast the autocorrelation function decays in time or distance.

Despite such model for the shadowing is validated through extensive measurement campaigns, no insight is gained about shadowing losses dependence over space. Such dependence might be of interested, for example, to identify locations of high shadowing correlation that can be used as prior information for adaptive power control algorithms.

It is usual to consider indoor scenarios to reconstruct losses due to obstacles through the use of imaging reconstruction methods, with the increasing deployment of wireless sensor networks. Applications for tracking and localization of objects have been shown in [12], as well as some of the theoretical background used in the literature of Radio Channel Tomography (RCT). To the knowledge of the authors, this is the first paper to consider to apply these ideas for an outdoor scenario, while providing predictions for the signal of interest (shadowing) at new transceiver locations without the need of making measurements.

The shadowing space dependence is captured by the concept of an Underlying Spatial Loss Field (USLF). It is called underlying, because its purpose is to introduce the notion of “intrinsic” losses, or losses that are specific to the space under consideration, regardless of the transmitter-receiver locations. This notion of “intrinsic” losses is motivated by the idea that the shadowing is produced by obstacles, and so, is related with the geometry of the scenario. Shadowing at given transmitter-receiver locations is linearly related with the USLF by some coefficients. These coefficients -or weights- capture propagation characteristics (i.e. propagation paths), and weight the relevance of the USLF at each space location for the transmitter-receiver locations where the shadowing needs to be computed.

The idea of considering the shadowing or obstruction losses as produced by a USLF was first presented in [5] and [7]. There, is also explained the methodology used to produce the weights for the USLF from the LoS path formed by the transmitter and the receiver. This methodology was later extended in [6] in order to allow more paths than only the LoS. Kernel-based learning of the weights was proposed in [8]. Given the weights and the shadowing, the estimation of the USLF reduces to a least-squares problem with a regularization term. Conditions on the underlying field must be impossed in order to solve the problem. Low rank and sparse solutions for the USLF were developed in [10].

The aim of this paper is to present a simple way to predict shadowing for new transmitter-receiver locations, based on previous measurement campaigns. In spite of very restrictive conditions to estimate the field, as highly-correlated and few transceiver locations, shadowing predictions show acceptable correlations with respect to the ground truth shadowing signal for the transceiver locations in the same scenario.

The methodology used to estimate the underlying field and ultimately make predictions for the shadowing is presented in Fig. 1. First, the Channel Impulse Response (CIR) is measured for pairs of transmitter and receiver coordinates. Details about the measurement campaign are

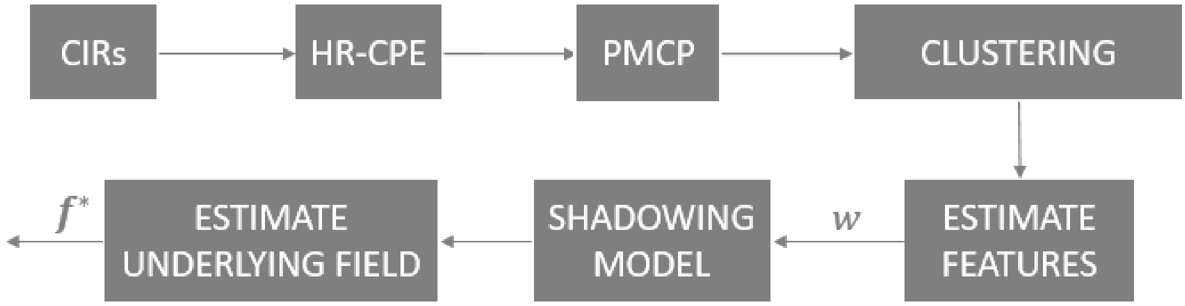

Fig. 1. Methodology used along this article.

presented in section II. In second place, parameters for the model of the CIR are estimated by a High-Resolution Channel Parameter Estimation (HR-CPE) algorithm. A briefly overview of the channel model, its parameters and the HR-CPE algorithm are presented in section III.

Due to the rate at which the channel is sensed and the number of propagation paths (a parameter of the model) per snapshot of measurement, there will be a large number of model parameters estimated. Thus, to reduce this number, parameters are clustered, and the centroids of the clusters are used as the new parameters. This clustering procedure is done in two steps: a Physical Mapping of Channel Parameters (PMCP) step described in section III-A and a Clustering of propagation pahts step described in section IV. The output of the PMCP step are propagation paths for each time snapshot. The output of the Clustering step are the clusters of propagation paths. The clusters of propagation paths are used as the input for the computation of the weights for the USLF. A new methodology to produce such weights, which extends the one proposed in [6] is introduced in section V-B. Estimation of the USLF is accomplished by using the shadowing values at the given transmitter-receiver locations, the computed weights, and the regularization condition explained in section V-A. Finally, under the assumption that the estimated USLF is accurate, the shadowing is predicted at new transmitter-receiver locations that were not used to compute the weights of the USLF. In order to do so, *a priori* information for the propagation paths at the new transmitter-receiver locations is used to produce the weights at those locations. The shadowing prediction results with a discussion about their relevance are presented in section VI. We conclude the paper with section VII, stressing the scope of the paper and its contributions.

The following notation is used throughout the article. Vectors start with lower-case bold letters (**b**). Matrices start with upper-case bold (**M**). In general, sets are represented by calligraphic

TABLE I

MEASUREMENT SETUP

Tx Antenna	UCA - 8 elements
Code Length	5.11 μs
Samples per Code (N_τ)	2044
IRF[a] duration	$\simeq 30\,ms$
Number of snapshots (T)	BS[b]: 6838
Center frequency	3.8 GHz
Bandwidth	200 MHz
Rx Antenna	2 x 4 Planar Array

[a] Impulse Response Function

[b] Base Station

letters ($\mathcal{L}$), except for $\mathcal{N}$ that we use to denote a Gaussian distribution. Special known sets are represented by blackboard-bold letters ($\mathbb{R}$).

II. MEASUREMENT CAMPAING SETUP

A top view of the scenario for the measurement campaign, with the base station location and the routes followed by the mobile transmitter is shown in Fig. 2. Each of the routes was followed by the mobile device, 3 or 4 times. The duration for a complete measurement (a whole ”lap” along one of the red paths) was in average 292 seconds. The channel was sensed at a rate of 30.2 Hz, which means that we get ”snapshots” of the channel each 33 ms (approximately). For each of this snapshots, the 8×8 MIMO impulse response of the channel is estimated. Since the bandwidth was 200 MHz and the samples per code are 2044, the longest path that can be received lasts 5.11 μs. In practice, relevant power delays beyond 500 ns are quite uncommon.

A summary of other settings for the measurement campaign is presented in Table I.

For the base station, we used patch antenna arrays of 8 elements with dual polarization. The mobile device was carrying an Uniform Circular Array (UCA) with 8 dipoles, distributed along its circumference and center.

III. CHANNEL PARAMETER EXTRACTION

Under the plane wave assumption we can model the signal $\mathbf{y}_t$ at the output of the receiver antenna array as a superposition of L waves [3]. The signal $\mathbf{y}_t$ can be expressed as

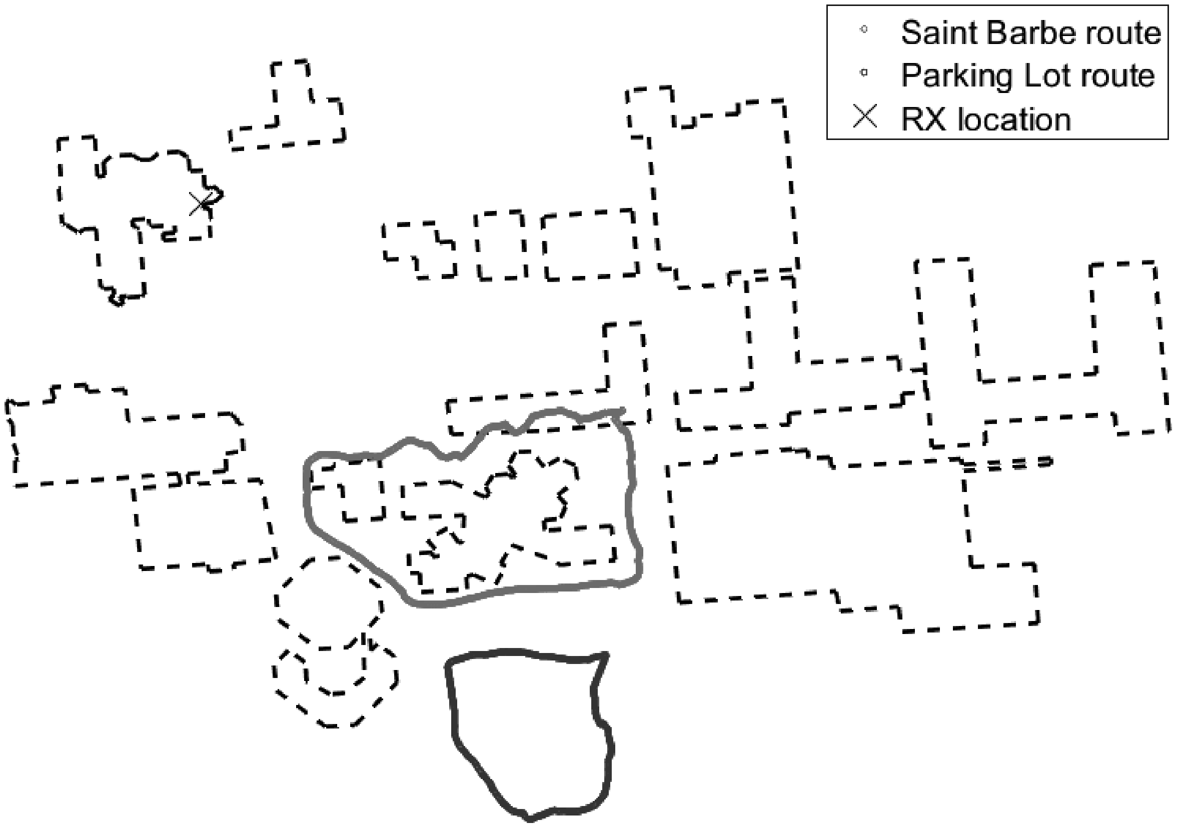

Fig. 2. Top view of the scenario under measurement at Louvain-la-Neuve. Only contours of the buildings considered. The red and blue curves correspond to mobile transmitter routes. The black cross shows the base station location.

$$\mathbf{y}_t = \sum_{l}^{L} \mathbf{s}_{t,l} + \mathbf{v}_t \tag{1}$$

where each of the waves $\mathbf{s}_{t,l}$ is defined as

$$\mathbf{s}_{t,l} = e^{j2\pi\nu_l t} \sum_{p_1=1}^{2} \sum_{p_2=1}^{2} \alpha_{l,p_1,p_2} \mathbf{c}_{2,p_2}(\mathbf{\Omega}_{RX,l}) \mathbf{c}_{1,p_1}(\mathbf{\Omega}_{TX,l})^T \mathbf{u}_{t-\tau_l} \tag{2}$$

and the rest of components of the model are defined as follows:

- α_{l,p_1,p_2} is the complex gain due to path l, polarization p_1 at the transmitter, and polarization p_2 at the receiver.
- ν_l is the Doppler frequency associated to path l.

- $\mathbf{u}_t \in \mathbb{R}^{N_{TX} \times 1}$ is the vector that contains the baseband representation of the signal sent at time t from the transmitter array.
- $\mathbf{c}_{1,p_1}(\boldsymbol{\Omega}_{TX,l}) \in \mathbb{R}^{N_{TX} \times 1}$ is the response of the transmitter antenna array to a wave impinging from direction $\boldsymbol{\Omega}_{Tx,l}$, due to polarization p_1.
- $\mathbf{c}_{2,p_2}(\boldsymbol{\Omega}_{RX,l}) \in \mathbb{R}^{N_{RX} \times 1}$ is the response of the receiver antenna array to a wave impinging from direction $\boldsymbol{\Omega}_{Rx,l}$, due to polarization p_2.
- $\boldsymbol{\Omega}_{TX,l}, \boldsymbol{\Omega}_{RX,l}$ is the Direction of Departure/Arrival of path l; unitary vectors containing the azimuth and elevation angles, or its correspondence in cartesian coordinates.
- $\mathbf{v}_t = \sqrt{\frac{N_0}{2}}\mathbf{W}_t$, where $\mathbf{W}_t \in \mathbb{C}^{N_{RX} \times 1}$ is a vector of independent complex, spatial and temporal white standard Gaussian noise.
- N_{RX} and N_{TX} are the number of RX and TX antennas, respectively (both are equal to 8, as stated in Table I).

For the model presented in (1) and (2) the parameters to be estimated are: the delays τ, the directions of departure $\boldsymbol{\Omega}_{TX}$, the directions of arrival $\boldsymbol{\Omega}_{RX}$, the complex gains α_{p_1,p_2} and the Doppler frequencies ν. Such estimation is done by using the Space Alternating Generalized Expectation-Maximization (SAGE) algorithm [2], [3], [4], [28]. It performs maximum likelihood estimation in two steps: an Expectation and a Maximization step. During the Expectation step, a function for the expectation of the log-likelihood is created with the actual parameter estimates. During the Maximization step, the function created in the Expectation step is used to find the parameters that maximize the function, which become the new estimates. The algorithm ensures that the successive log-likelihood estimates are non-decreasing. At the end of the SAGE algorithm, the set of channel parameters $\{\tau_l, \nu_l, \Omega_{Tx,l}, \Omega_{Rx,l}, \alpha_{l,p_1,p_2}\}_{l=1}^{L}$ is estimated for each time-snapshot t.

The estimation of the number of paths L is not addressed in this paper. The maximum number of paths was chosen to be big enough to account for the most dominant paths in power. Typical values found in the literature range from 20 to 50 [28], [29], [30].

A. Physical Mapping of Interacting Objects

In what follows, we consider only 2D propagation (elevation angles are not taken into account). For the matter of convenience we rename angles. Bi-dimensional angles of departure will be denoted by ϕ. Bi-dimensional angles of arrival will be denoted by θ. For a given time-snapshot

t we have a set of channel parameters $\{\tau_l, \phi_l, \theta_l, \alpha_{l,p_1,p_2}\}_{l=1}^{L}$. Whenever it is clear that we are talking about the channel parameters of a given time snapshot we drop the time index.

For a given snapshot t, each propagation path l is mapped into the scenario by means of the following considerations:

- The reference (0 rad) for angles ϕ_l and θ_l is set to the x-axis (longitude coordinate). For doing so, it is assumed that the maximum power is coming from the virtual line connecting the RX with the TX; putted in other words, the LoS without the presence of the buildings.
- Only the buildings may cause reflection. As shown in Fig. 3, the first bounce is found by extending a line with angle ϕ_l between the TX and the contours of the buildings. The first building the line hits is the first bounce of the path.
- The last bounce is found in a similar way. Extending a line with angle θ_l between the RX an the contours of the buildings. The first building the line hits is the last bounce of the path.

From those considerations, the path l is composed by the three line segments shown in Fig. 3, whose line equations are:

$$\begin{aligned} \mathcal{L}_{l,t}^{1} &= \left\{\mathbf{x}_t + p(\mathbf{p}_{l,t} - \mathbf{x}_t) |\ p \in [0,1]\right\} \\ \mathcal{L}_{l,t}^{2} &= \left\{\mathbf{p}_{l,t} + p(\mathbf{q}_{l,t} - \mathbf{p}_{l,t}) |\ p \in [0,1]\right\} \\ \mathcal{L}_{l,t}^{3} &= \left\{\mathbf{q}_{l,t} + p(\mathbf{r} - \mathbf{q}_{l,t}) |\ p \in [0,1]\right\} \end{aligned} \tag{3}$$

The estimated delay is not taken into account when physically mapping the channel parameters to the scenario. However, it is included in the clustering of the mapped channel parameters, explained in section IV. Its incidence appears specifically in (8). From the previous considerations, we have a basic idea about the detailed propagation paths: there is no knowledge about the real number of bounces and the exact angles at which the waves reflect, transmit or diffract once arriving at the buildings. However, up to this point we are not interested in a more detailed knowledge about the propagation paths. We know that regarding the channel parameters mapping, the only thing that can add complexity to our model is the delay τ_l. We consider that what happens in between the first bounce $\mathbf{p}_l$ and the last one $\mathbf{q}_l$ can be characterized only by an effective path length η_l, that is defined as:

$$\eta_l = c\tau_l - |\mathcal{L}_l^1| - |\mathcal{L}_l^3|, \tag{4}$$

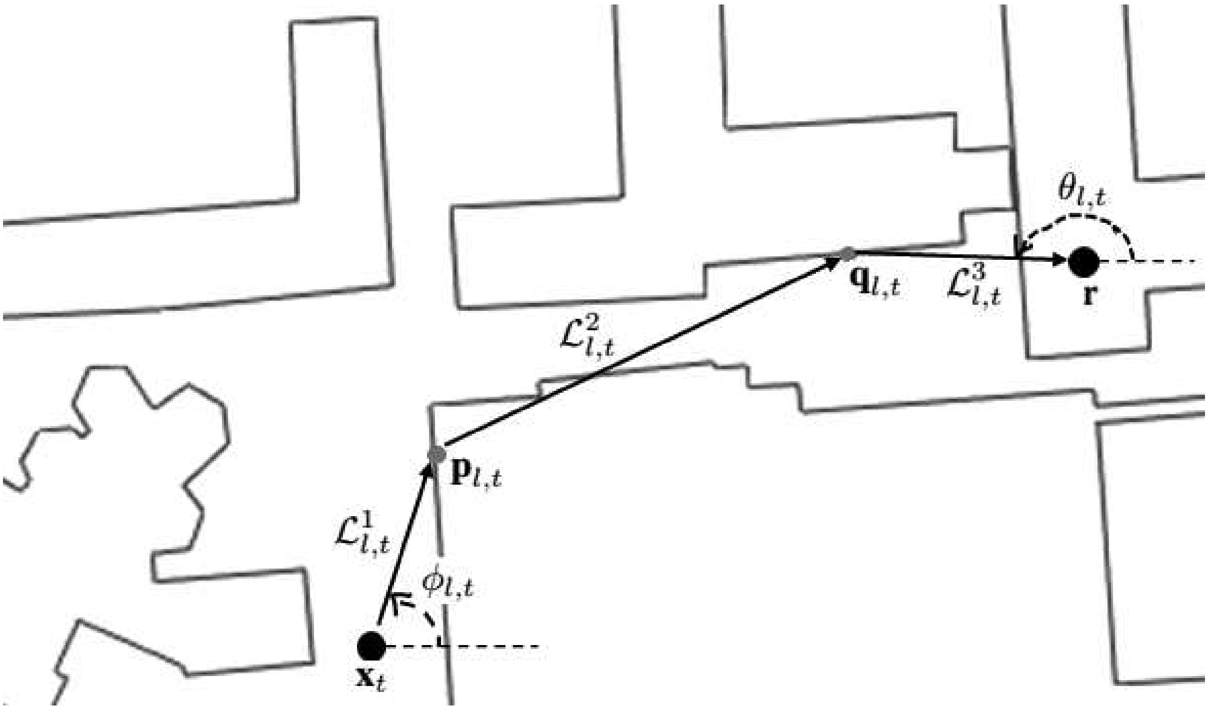

Fig. 3. Physical model used.

where for path l, $|\mathcal{L}_l^1|$ and $|\mathcal{L}_l^3|$ represent the distance between transmitter and first bounce, and the distance between last bounce and receiver, respectively. The line segment $\mathcal{L}^2_{l,t}$ in Fig. 3 is a representation of the "effective" path with propagation of length η_l.

IV. CLUSTERING MULTIPATH COMPONENTS

From the channel parameters mapping, we end up with bounces $\mathbf{p}_{l,t}$ and $\mathbf{q}_{l,t}$ where for each time-snapshot t the number of paths $l = 1, ..., L_t$ can vary. Each bounce is representing a 2D coordinate (latitude and longitude). As for each $\mathbf{p}_{l,t}$ there is a corresponding $\mathbf{q}_{l,t}$ in the same path, we consider them as an unique coordinate $\mathbf{z}_{l,t} \in \mathbb{R}^4 = [\mathbf{p}_{l,t}, \mathbf{q}_{l,t}]$. As shown in Fig. 3, the path l starts at location $\mathbf{x}_t$ of the transmitter (the location of the transmitter at time-snapshot t), then hits $\mathbf{p}_{l,t}$, follows an unknown trajectory whose effective distance is η_l, then hits $\mathbf{q}_{l,t}$, and finally reaches receiver $\mathbf{r}$.

Let $\mathcal{Q}$ be a set comprising all $\mathbf{z}$ to be clustered. A *clusterization* Γ^k is a mapping $\Gamma^k : \mathcal{Q} \to \mathcal{M}$ that assigns to each $\mathbf{z} \in \mathcal{Q}$ a number in the set $\mathcal{M} = \{1, 2, \cdots, k\}$, with k unknown beforehand. In other words, a *clusterization* partitions the original set $\mathcal{Q}$ into *cluster groups* $\mathcal{C}_1, \mathcal{C}_2, ..., \mathcal{C}_k$, such that $\mathcal{C}_i \subset \mathcal{Q}$, $\cup_i \, \mathcal{C}_i = \mathcal{Q}$, and $\cap_i \, \mathcal{C}_i = \varnothing, \;\; i = 1, ..., k$. A *cluster group* $\mathcal{C}_i$ is a set composed by $\mathbf{z}$ with the same assignment $i \in \mathcal{M}$.

Recall that for different clusterizations we end up with different groups $\mathcal{C}_i$. Therefore, it might be more accurate to denote the cluster groups by $\mathcal{C}_i^k$; however, we drop the k index for convenience.

Over each $\mathbf{z}$ a "disk" of radii ϵ can be defined as $Dz(\epsilon) = \{\mathbf{x} \in \mathcal{R}^4 \;\mid\; d(\mathbf{z}, \mathbf{x}) \leq \epsilon\}$, where the function $d(\cdot, \cdot)$ is a distance measure defined in section IV-A. We define as well, the intersection set $\mathcal{I}_i(\epsilon)$ of all the disks $Dz(\epsilon)$ of all the $\mathbf{z}$ in a *cluster group* i: $\mathcal{I}_i(\epsilon) = \cap_{j=1}^{|\mathcal{C}_i|} Dz_j(\epsilon)$, $\forall \mathbf{z}_j \in \mathcal{C}_i$.

Inter-cluster ξ and intra-cluster β distance measures are shown in the following 2 equations, in order to quantify the goodness of the mapping Γ:

$$\beta = \max_{\mathcal{C}_i} \max_{\mathbf{u}, \mathbf{v} \, \in \, \mathcal{C}_i} d(\mathbf{u}, \mathbf{v}) = \max_{\mathcal{C}_i} \beta_i \tag{5}$$

$$\xi = \min_{\mathbf{u} \, \in \, \mathcal{I}_i(\beta_i), \, \mathbf{v} \, \in \, \mathcal{I}_j(\beta_j), \, i \neq j} d(\mathbf{u}, \mathbf{v}) \tag{6}$$

The definition of disks $Dz(\epsilon)$ must be made beforehand. The intersection sets $\mathcal{I}_i(\epsilon)$ might be complicated to compute, or even to define analytically. Therefore, an approximation for (6) is done by considering $\mathcal{I}_i(\beta_i)$ as the centroid $\mathbf{y}_i$ of the cluster $\mathcal{C}_i$: $\mathcal{I}_i(\beta_i) \simeq \dfrac{1}{|\mathcal{C}_i|} \sum_{\mathbf{z} \in \mathcal{C}_i} \mathbf{z} \triangleq \mathbf{y}_i$. Equation (6) is replaced by (7).

$$\xi \simeq \min_{i \neq j} \; d(\mathbf{y}_i, \mathbf{y}_j) \tag{7}$$

Any Γ^k must try to maximize ξ and minimize β simultaneously. Both requirements could be linearly combined in a single restriction. With this in mind, the problem of finding a clusterization for the set of propagation paths of a given time snapshot t could be written as:

$$\underset{\mathcal{C}_1, \ldots, \mathcal{C}_{C_t}}{\operatorname{argmin}} \; \beta^k - \xi^k + \lambda_\tau c \sum_i \sum_{\tau_1, \tau_2 \, \in \, \mathcal{C}_i} |\tau_1 - \tau_2| \tag{8}$$

where c is the speed of light, $\mathcal{C}_1, ..., \mathcal{C}_{C_t}$ are the clusters found, C_t is the number of clusters in the snapshot t, and β^k, ξ^k account for different clusterizations. The third component in (8) are the delays τ. The controllable setting λ is used to weight how much the delays matter into the clusterization. For more on this refer to [13].

A. *Distance measure*

The distance measure used in equations (6) and (5) is defined here. Given two points $\mathbf{u}, \mathbf{v} \in \mathcal{Q}$:

$$\begin{aligned} d(\mathbf{u},\mathbf{v}) = \frac{1}{N}\sum_{i=0}^{N} || \, [\mathbf{u}]_{1:2} + \frac{i}{N}([\mathbf{u}]_{3:4} - [\mathbf{u}]_{1:2}) - \\ ([\mathbf{v}]_{1:2} + \frac{i}{N}([\mathbf{v}]_{3:4} - [\mathbf{v}]_{1:2}))||_2 \end{aligned} \tag{9}$$

where $||\cdot||_2$ is the euclidean norm in $\mathbb{R}^2$, the operator $[\mathbf{x}]_{i:j}$ represents the coordinates i to j of $\mathbf{x}$, and N is the number of divisions in the line segments associated with $\mathbf{u}$ and $\mathbf{v}$. Those line segments are built, by just extending a line from $\mathbf{p}_{l,t}$ to $\mathbf{q}_{l,t}$ for each $\mathbf{z}$. Intuitively, what the distance measure in (9) is doing is to compute the degree of "match" between the line segments associated with $\mathbf{u}$ and $\mathbf{v}$. The defined measure still can be used if instead of using line segments we use piecewise line segments. That would be the case if we have information about the intermediate bounces.

However, the distance defined like in (9) fails to detect observations $\mathbf{z}_1$ and $\mathbf{z}_2$ that might be clustered in the same group. It happens because the distance is not taking into account that the line segments in comparison can cross. For example, the left and middle graphs in Fig. 4 are two pairs of different observations that must give the same distance value (they correspond to equivalent situations according to our criteria for defining a cluster). However, the initial distance h in the left graph will decrease - due to the crossing between line segments -, whereas it will not in the middle graph, giving arise to a different distance computation for each case. To overcome this, we introduce a correction represented in the graph at the right of Fig. 4. There, the new equivalent line segments are composed by the $\mathbf{q}_{l,t}$ and $\mathbf{p}_{l,t}$ of both initial line segments.

V. UNDERLYING SPATIAL FIELD

In [5] the idea of an Underlying Spatial Loss Field (USLF) is used to model the attenuation caused by obstacles. Shadowing values are produced from such field as follows, considering it an isotropical gaussian random field:

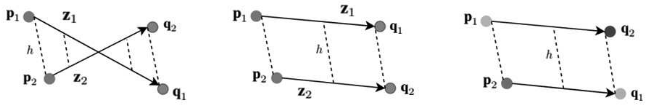

Fig. 4. Left and middle figures represent two situations for which (9) must be equal. The figure in the right represents the corrected version of the figure in the left. Arrows represent the direction followed by the path. z_1, z_2 represent different paths.

$$b(\mathbf{x}, \mathbf{x}') \doteq b_t \in \mathbb{R} = \frac{1}{\sqrt{||\mathbf{x} - \mathbf{x}'||_2}} \int_{\mathbf{x}}^{\mathbf{x}'} f(\hat{\mathbf{x}}) d\hat{\mathbf{x}} \tag{10}$$

where b is the shadowing, t the time snapshot for which $(\mathbf{x}, \mathbf{x}')$ are the locations of the transceivers lying in the scenario $\mathcal{A} \subset \mathbf{R}^2$, and $f(\hat{\mathbf{x}}) : \mathcal{A} \mapsto \mathbb{R}$ is the USLF. The square root factor is used so the variance of the shadowing in link $(\mathbf{x}, \hat{\mathbf{x}})$ is approximately constant.

A generalization of such idea in [6] and rewritten in [8] allows for weighting on $f(\mathbf{x})$

$$b_t = \int_{\mathcal{A}} w(\mathbf{x}, \mathbf{x}', \hat{\mathbf{x}}) f(\hat{\mathbf{x}}) d\hat{\mathbf{x}} \tag{11}$$

Now the weighting function w weights each pixel $\hat{\mathbf{x}}$ depending as well on the transceiver locations $(\mathbf{x}, \mathbf{x}')$ over the whole domain $\mathcal{A}$. In the case of the shadowing defined as in (10), the weights w are given by

$$w(\mathbf{x}, \mathbf{x}', \hat{\mathbf{x}}) = \frac{1}{\sqrt{||\mathbf{x} - \mathbf{x}'||_2}} \int_{\mathbf{x}}^{\mathbf{x}'} \delta(\mathbf{x}, \mathbf{x}') d\hat{\mathbf{x}} \tag{12}$$

where w can be seen as a “selection” function, choosing which infinitesimal pixels are active (non-zero value). A graphical representation of the weights w in (12) is shown in the left image in Fig. 5. More details on alternatives to the weighting function (12) are provided in section V-B.

By considering a discrete grid of pixels, instead of the continuous space $\mathcal{A}$, continuous integration is replaced by summation. If the same principle in (11) is extended to each and all snapshots $t_1, ..., t_T$ we end-up with the following linear system of equations

$$\begin{pmatrix} b_1 \\ b_2 \\ \dots \\ b_T \end{pmatrix} = \begin{pmatrix} w_{1,1} & w_{1,2} & \dots & w_{1,N_xN_y} \\ w_{2,1} & w_{2,2} & \dots & w_{2,N_xN_y} \\ \dots & \dots & \dots & \dots \\ w_{T,1} & w_{T,2} & \dots & w_{T,N_xN_y} \end{pmatrix} \begin{pmatrix} f_1 \\ f_2 \\ \dots \\ f_{N_xN_y} \end{pmatrix} \tag{13}$$

where the set of transceiver pairs is $\{(\mathbf{x}, \mathbf{x}')_i\}_{i=1}^{T}$, N_x are the pixels in the first dimension, N_y the pixels in the second dimension, $f_i = f(\hat{\mathbf{x}}_i)$ is the USLF evaluated at pixel i, and $w_{ij} = w((\mathbf{x}, \mathbf{x}')_i, \hat{\mathbf{x}}_j)$ with $i = 1, ..., T$ and $j = 1, ..., N_xN_y$. This linear system is rewritten in matrix form as $\mathbf{b} = \mathbf{A}\mathbf{f}$, with matrix $\mathbf{A}$ gathering the weights. We call matrix $\mathbf{A}$ the Active Paths Matrix (APM).

As the APM is in general non-square, the system in (13) could have no solution or infinite solutions, depending upon the size and rank of $\mathbf{A}$. In general $\mathbf{A}$ is rank deficient. Moreover, due to its construction from the links, or as will be shown in section V-B, due to its construction from the propagation paths, the APM is very ill-conditioned. However, if $\mathbf{A}$ is full rank the solution to (13) is already available from $\mathbf{f} = \mathbf{A}^{\dagger}\mathbf{b}$, where $\mathbf{A}^{\dagger}$ is either the usual inverse (for the case where APM is square) or the Moore-Penrose pseudo-inverse.

For the non-full rank case, it is needed to impose restrictions on the solution, in order to reduce the infinite space of solutions to one with some desirable properties. The problem of finding $\mathbf{f}$ is now the problem:

$$\mathbf{f}^{*} = \underset{\mathbf{f}}{\operatorname{argmin}} \, ||\mathbf{A}\mathbf{f} - \mathbf{b}||_2^2 + \lambda_R R(\mathbf{f}) \tag{14}$$

That is to say: a minimum least squares problem with some regularization term given by the function $R(\mathbf{s})$. The constant λ_R decides how much the solution follows the condition imposed by the regularization term.

As mentioned in section I, the idea about $\mathbf{f}^{*}$ is that it accounts for the “intrinsic” losses of the scenario. If such assumption is true, then, if also $\mathbf{A}$ is known, shadowing prediction at any transmitter-receiver locations is possible by using: $\mathbf{b} = \mathbf{A}\mathbf{f}^{*}$. The assumption about $\mathbf{f}^{*}$ could be full-filled by fully sensing the space $\mathcal{A}$ as in the left panel of Fig. 6 [9]. In that case, each node is static and works simultaneously as a transmitter and receiver. The channel is sensed over time. However, for outdoor scenarios this measurement setup might not be feasible in practice due to the large number of base stations. For that, a typical measurement setup is shown in the

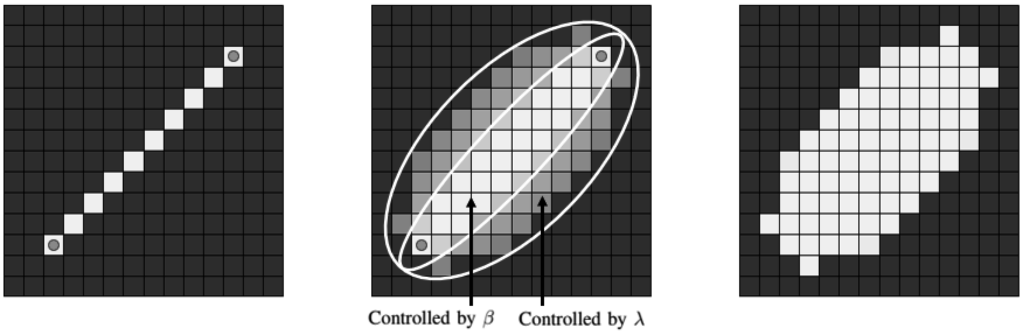

Fig. 5. Different weighting functions. The plots correspond to the weights for the case of 20 × 20 pixels. The dark blue represents zero valued entries. The red circles represent the transceiver locations. Left: NeSh model. Center: IAEM. Right: NEM.

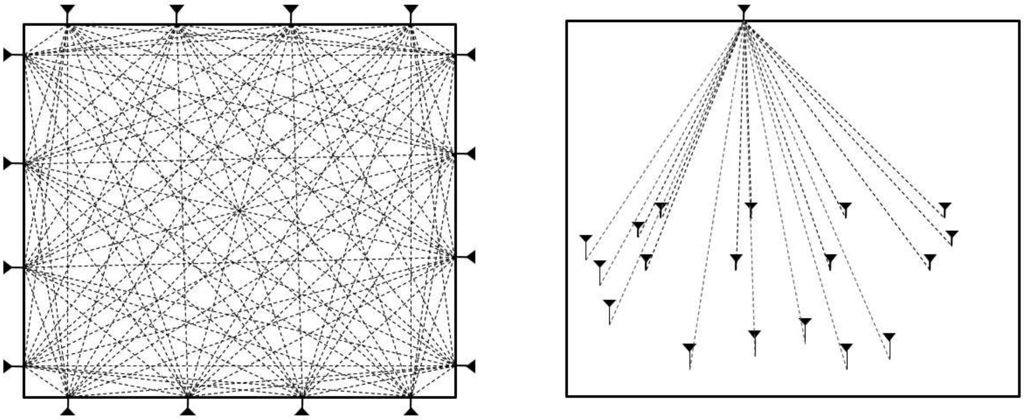

Fig. 6. Measurement setup to sense the space $\mathcal{A}$. Each wireless node (transmitter or receiver) is represented by the stick with the inverted triangle. The black box represents the space $\mathcal{A}$. The dashed lines represent the LoS link between a pair of wireless nodes. Left panel: ideal placement of the wireless nodes in order to have a realistic estimate the USLF. Right panel: realistic placement of the wireless nodes for the estimation of the USLF. The red dashed lines in the right figure represent locations of the wireless nodes that are not used for the estimation of the USLF.

right panel of Fig. 6. For that case, the estimated $\mathbf{f}^*$ will not be an underlying field, since it will not account for all losses correctly. However, as the results show (in section VI), for some new transmitter-receiver locations, the estimated $\mathbf{f}^*$ will be good enough to make good predictions in trend for the shadowing.

A. Regularization

The regularization term $R(\mathbf{f})$ in (14) accounts for desired properties of the solution $\mathbf{f}^*$. In [10], a sparse and low rank approach is used. The solution vector $\mathbf{f}^*$ is decomposed into two terms, namely $\mathbf{L}$ and $\mathbf{E}$: the former, a low rank matrix and the latter, a sparse matrix. Low-rank condition is achieved by using nuclear norm, and sparsity by using l_1 norm.

$$R(\mathbf{f}) = \lambda_1 ||\mathbf{L}||_* + \lambda_2 ||\mathbf{E}||_1 \tag{15}$$

Such a decomposition is physically justified by thinking about the low-rankness as a consequence of regular placement of building and walls, as in the indoor and urban cases; on the other hand, the sparsity term account for the presence of outliers that don't fit into the low-rank.

Tikhonov regularization (also known as ridge regression) as in [9] and [8], forces the solution to follow the properties in (16) given by equivalent matrices $\mathbf{Q}$ or $\mathbf{G}$, respectively.

$$R_{TIK}(\mathbf{s}) = ||\mathbf{Qs}||_2^2 = (\mathbf{Qs})^T\mathbf{Qs} = \mathbf{s}^T\mathbf{Q}^T\mathbf{Qs} = \mathbf{s}^T\mathbf{Gs} \tag{16}$$

The solution to (14) with (16) as regularizer is found by setting the derivative of $\mathbf{s}$ to zero and solve for $\mathbf{s}$:

$$\begin{aligned} \mathbf{s}_{TIK} &= (\mathbf{A}^T\mathbf{A} + \lambda_R\mathbf{Q}^T\mathbf{Q})^{-1}\mathbf{A}^T\mathbf{b} \\ &= (\mathbf{A}^T\mathbf{A} + \lambda_R\mathbf{G})^{-1}\mathbf{A}^T\mathbf{b} \end{aligned} \tag{17}$$

Common choices start with the minimum norm regularizer $R_{TIK} = ||\mathbf{s}||_2^2$ (by choosing $\mathbf{Q} = \mathbf{I}_{N_x \times N_y}$), which promotes smoothness by forcing the energy to spread over more entries of $\mathbf{s}$ [8]. One can generate $\mathbf{Q}$ assuming, as said in previous section, that $\mathbf{f}$ is an isotropical gaussian random field [5]; in such a case, the covariance matrix $\mathbf{G}^{-1}$ entries are

$$[\mathbf{C}]_{ij} = [\mathbf{G}^{-1}]_{ij} = \sigma_f \exp\left(-\frac{||\mathbf{x}_i - \mathbf{x}_j||_2^p}{\delta_f}\right) \tag{18}$$

with $p = \{1, 2\}$, σ_f is regarded as the scale, and δ_f as the length-scale. The estimated $\mathbf{f}^*$ in (14), has a bayesian interpretion [11]. For prior $\mathbf{f} \sim \mathcal{N}(\mathbf{0}, \mathbf{C})$ and data likelihood $\mathbf{b}|\mathbf{f} \sim \mathcal{N}(\mathbf{Af}, \mathbf{I}_{N_x \times N_y})$, $\mathbf{f}^*$ is the maximum *a posteriori* estimate.

Another regularization choice is the Truncated Singular Value Decomposition (TSVD). It takes the biggest $K < \min(T, N_x N_y)$ eigenvalues from the singular value decomposition of matrix **A**. Given the Singular Value Decomposition (SVD) $\mathbf{A} \in \mathbb{R}^{T \times N_x N_y} = \mathbf{U}\mathbf{D}\mathbf{V}^T$, with

$$\begin{aligned} \mathbf{U} \in \mathbb{R}^{T \times T} &= [\mathbf{u}_1, ..., \mathbf{u}_K, ..., \mathbf{u}_T] \\ \mathbf{V} \in \mathbb{R}^{N_x N_y \times N_x N_y} &= [\mathbf{v}_1, ..., \mathbf{v}_K, ..., \mathbf{v}_{N_x N_y}] \\ \mathbf{D} \in \mathbb{R}^{T \times N_x N_y} &= diag(\sigma_1, ..., \sigma_K, ..., \sigma_{\min(T, N_x N_y)}) \end{aligned} \tag{19}$$

where $\sigma_1 \geq \sigma_2 \geq ... \geq \sigma_K \geq ... \geq \sigma_{\min(T, N_x N_y)}$ and $diag(a, b, c)$ stands for a matrix with non zero entries a, b, c in its principal diagonal. The solution is built from (19) as

$$\mathbf{s}_{TSVD} = \sum_{i=1}^{K} \frac{1}{\sigma_i} \mathbf{u}_i^T \mathbf{b} \mathbf{v}_i \tag{20}$$

B. Weighting function

Besides the model with weights choose as in Equation (12), Fig. 5 shows two alternatives for choosing the weights w. The model for the weights used in (12) is called NeSh [7]. It considers that only the USLF values at space locations on the LoS link between the transmitter and receiver have an impact on the shadowing. Oppose to the LoS assumption in the NeSh, the Normalized Ellipse Model (NEM) and the Inverse Area Ellipse Model (IAEM), assume that there is a beam of rays departuring and a beam of rays arriving, such that USLF values at outer space locations of the LoS are responsible for producing the shadowing. These space locations that are relevant are those located within an ellipse whose foci are the transmitter and the receiver location.

The NEM [6] illustrated in the right of Fig. 5, assigns the same non-zero value to all pixels within an ellipse whose focii are the transceivers. The semi-minor axis of the ellipse is determined by the wavelength, which guarantees that the ellipsoid can be thought as the first Fresnel Zone. The last weighting model is displayed at the center of Fig. 5. IAEM still uses the idea of the ellipse explained previously, but allows pixels to have different values depending upon their distance with respect to the LoS: the farther away from the link, the less value they have.

Both models can be merged into the following equation [8]

$$w(\zeta_1, \zeta_2) = \begin{cases} 0 & \zeta_2 < \zeta_1 + \dfrac{\lambda}{2} \\ \min\left\{\Omega(\zeta_1, \zeta_2), \Omega(\zeta_1, \sqrt{\zeta_1^2 + \beta^2})\right\} & \zeta_2 \geq \zeta_1 + \dfrac{\lambda}{2} \end{cases} \tag{21}$$

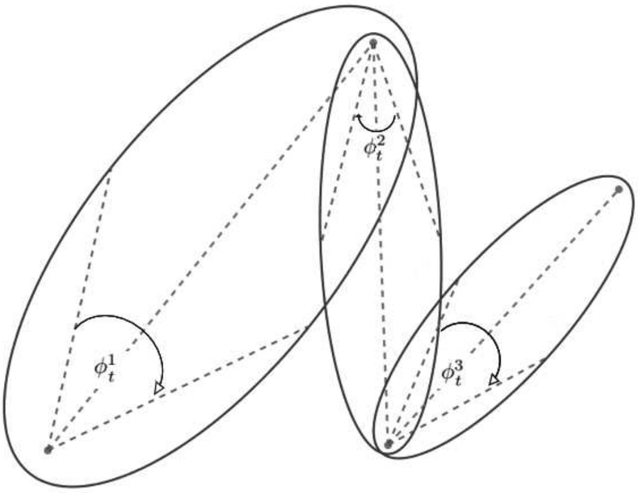

Fig. 7. Proposed weighting model. The red points are the transceivers, first and last bounces of the propagation path. The latter is illustrated by the red dashed line. The blue dashed lines represent the 'beamwidth' of the launched and reflected rays, and control with the semi-minor axis of each ellipse.

with $\Omega(\zeta_1, \zeta_2) = 4/(\pi\zeta_2\sqrt{\zeta_2^2 - \zeta_1^2})$, $\zeta_1 \doteq \zeta_1(\mathbf{x}, \mathbf{x}') = ||\mathbf{x} - \mathbf{x}'||_2$ and $\zeta_2 \doteq \zeta_2(\mathbf{x}, \mathbf{x}', \hat{\mathbf{x}}) = ||\mathbf{x} - \hat{\mathbf{x}}||_2 + ||\mathbf{x}' - \hat{\mathbf{x}}||_2$. For $\beta = 0$ we have the NeSh model, and for $\beta = \sqrt{\zeta_1\lambda}$ we have the NEM. In between values of β define two regions: an outer ellipse controlled by λ and an inner one defined by β. Weights in the inner ellipse have the same value, outside the inner but inside the outer behave like the IAEM, and outside the bigger ellipse have zero value (see Fig. 5).

The three weighting models presented, consider propagation paths as in LoS or Obstructed LoS (OLoS) conditions. Furthermore, as the wavelength can be in the order of centimeters while the distance between the focii of the ellipse is in the order of tens of meters, the semi-minor axis of the ellipse would be very thin. That is to say, that practical beamwidth is very narrow, and NEM as well as IAEM end up being the NeSh model.

In order to cope for that, we make use of the clusters of propagation paths found through procedures in sections III-A and IV. Since after those procedures we have information about which are the dominant propagation paths, we can use this information to build the weights. The idea of using an ellipse as the region where the USLF produces the shadowing is kept in our model. However, the semi-minor axis of the ellipse, previously linked to the wavelength, have a different interpretation in our model. Fig. 7 shows the concept. For a propagation path at snapshot t, three ellipses are used. The first one, with foci at the transmitter and the first bounce.

The second one, with foci at the first bounce and last bounce; and the third one, with foci at the last bounce and the receiver.

Clusters found in section IV are used. Instead of using a propagation path, we use the centroid of the cluster of propagation paths. In this case, the foci of the second ellipse (the one whose foci are the first and the last bounce), correspond to the centroids of the cluster bounces. For the first ellipse, the angle with respect to the horizontal axis between its foci is the mean angle of departure of the cluster of propagation paths. For the third ellipse, the angle with respect to the horizontal axis between its foci is the mean angle of arrival of the cluster of propagation paths.

Angle ϕ_t^1 in Fig. 7, correspond to the angular spread of the angles of departure of the cluster of propagation paths. Angle ϕ_t^3 in the figure, correspond to the angular spread of the angles of arrival of the cluster of propagation paths. Angle ϕ_t^2 is a parameter of the weighting model proposed. It controls how much the rays spread after the first bounce. For C_t number of clusters in snapshot t, the contribution to the shadowing of each cluster of propagation paths is added, and so (11) is rewritten as

$$b_t \simeq \sum_{c=1}^{C_t} \sum_{\hat{\mathbf{x}}_i \in \mathcal{A}} w(\mathbf{x}_c, \mathbf{x}'_c, \hat{\mathbf{x}}_i) f(\hat{\mathbf{x}}_i) \tag{22}$$

where $w(\mathbf{x}_c, \mathbf{x}'_c, \hat{\mathbf{x}}_i)$ denotes the weight built from the procedure previously explained using the centroid c of a cluster of propagation paths.

VI. RESULTS

The two transmitter routes used to estimate the USLF and predict the shadowing are shown in Fig. 2, as well as the static receiver location. There were used 6838 and 5259 snapshots for the Saint Barbe and Parking Lot routes, respectively. As a convention, all symbols associated with Saint Barbe route are denoted as $\{\}_{ST-BARBE}^{MEAS}$, to stress the fact that the weights built for this transmitter-receiver locations use the propagation paths computed from measurements. Also the APM built from the weights for that route is used to estimate the USLF. On the other hand, all symbols associated with the Parking Lot route are denoted by $\{\}_{P-LOT}^{MDL}$, to stress that for such route the weights are built from *a priori* information about the angles of arrival and departure, instead of the measurements used to estimate the true angles of arrival and departure. Also, the weights for this route were not included in the APM used to estimate the USLF.

Fig. 8. Left panel: Normalized weights for a given snapshot, for the Saint Barbe route. Right panel: Propagation paths and centroids for the same snapshot. Transceivers are represented with black dots. Dashed red lines are propagation paths. Black crosses represent propagation paths first and last bounces. Blue dots are centroid first and last bounces. Thick green line represent the propagation path of the centroid. There are 24 propagation paths and 9 clusters for the snapshot. Six of the clusters are singleton, one has 2 members, another 7 and another 9.

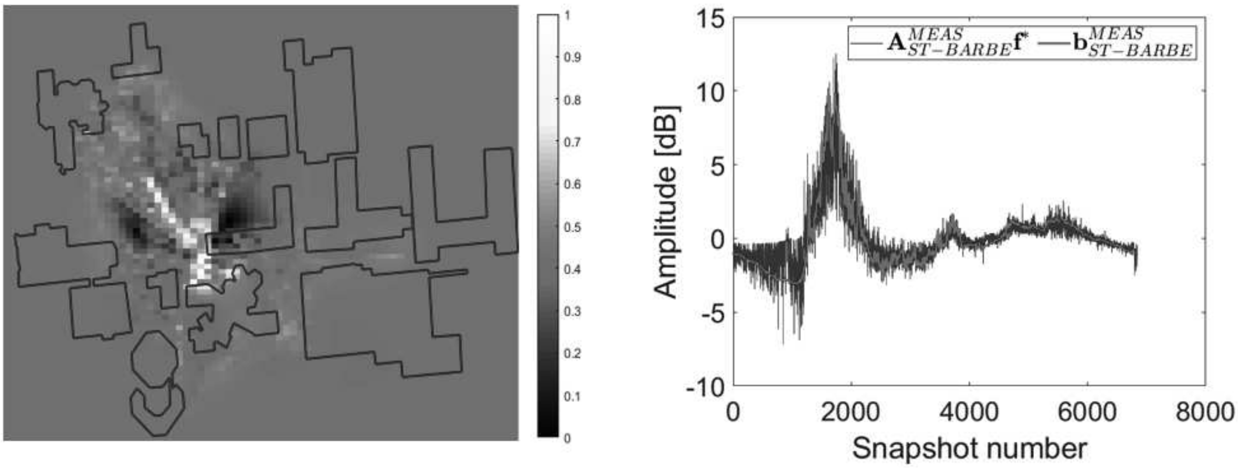

Fig. 9. Left panel: Estimated normalized USLF $\mathbf{f}^{*}$, using Saint Barbe transmitter route. Right panel: Reconstructed $\mathbf{A}\mathbf{f}^{*}$ and measured shadowing from Saint Barbe transmitter route.

A maximum of 27 waves per snapshot is considered. Computed weights for a given snapshot of the Saint Barbe route are shown in Fig. 8, as well as propagation paths and clusters of propagation paths. Hierarchical clustering is used to find out the clusters of propagation paths. A distance threshold of 10, as the maximum distance using (9) between two paths was used to cut the dendogram. The λ_{τ} in (8) is set to zero, which means that no delay information is taken

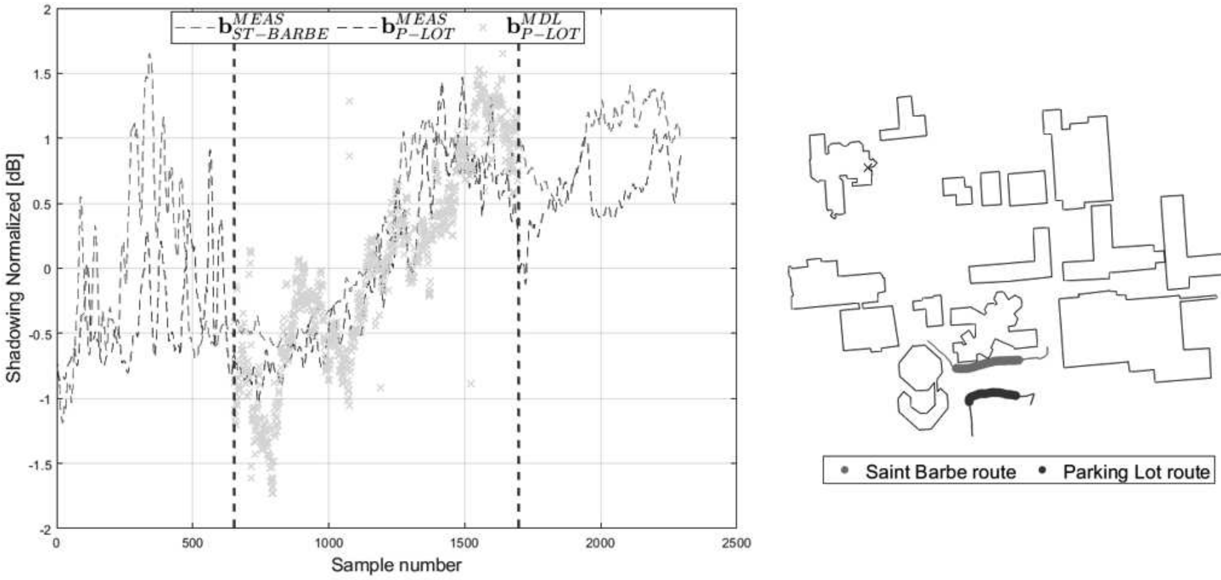

Fig. 10. Left panel: shadowing predicted and from measurements. $\mathbf{b}^{MEAS}_{ST-BARBE}$ is the shadowing from measurements for the Saint Barbe route, $\mathbf{b}^{MEAS}_{P-LOT}$ is the shadowing from measurements for the Parking Lot route, and $\mathbf{b}^{MDL}_{P-LOT}$ is the predicted shadowing for the Parking Lot route. Right panel: transmitter routes considered. The thicker points correspond to the snapshot zone in between the dashed black lines in the left panel.

into consideration.

For all the ellipses associated with the singleton clusters and for the second ellipses of all clusters, the semi-minor axis is set to 40% the distance between one of the foci and the center of the ellipse. It means we assume energy spreads with an angle $\phi_t^2 \simeq 43.6^\circ$ from first bounce towards last bounce, for the case of the second ellipses. This choice is heuristically based.

Tikhonov regularization is used, assuming a stationary gaussian random field with covariance matrix produced by the kernel in (18). Kernel parameters σ_f and δ_f have an impact on the Mean Squared Error (MSE) $= \frac{1}{T}||\mathbf{b}^{MEAS}_{ST-BARBE} - \mathbf{A}^{MEAS}_{ST-BARBE}\mathbf{f}^{*}||^2$ and on the correlation of the predicted shadowing $\mathbf{b}^{MDL}_{P-LOT}$ with the ground truth shadowing $\mathbf{b}^{MEAS}_{P-LOT}$. The better the choice of the parameters to reduce the MSE the worse that choice to lower the correlation.

Fig. 9 shows the estimated USLF $\mathbf{f}^{*}$ and the reconstructed shadowing $\mathbf{Af}^{*}$, when $\delta_f = 2$, $\sigma_f = 30$ and $\lambda_R = 0.1$ for the Saint Barbe route. The ground truth shadowing shown in red is computed from the measured channel impulse response, by filtering fading with a moving average window of size 10λ. Then, path loss trend with distance is found and removed as well. The higher the regularization constant λ_R the smoother is the USLF and the higher the MSE.

Shadowing prediction was done for the Parking Lot transmitter route. The weights for the new route are produced by assuming some prior information for the angles of arrival and departure, instead of finding them from measurements through SAGE. The purpose is to be able to predict shadowing with no measurement inputs. No information more than the buildings placement is used. We assume that for the transmitter side, angles departure within an interval of $[-\pi, \pi]$ rads, with the LoS as the angular reference. For the receiver side, we assume angles arrive within an interval of $[-70^\circ\ , 70^\circ\ ]$, again, with the LoS as the angular reference. The choices are justified by the beamwidths of the transmitter and receiver antennas. First and last bounces in the buildings are found following the same guidelines in section III-A. This bounces are used to get the weights for the new transmitter locations, using the proposed model in Fig. 7, and explained in section V-B. The predicted shadowing for the Parking Lot route becomes $\mathbf{b}^{MDL}_{P\text{-}LOT} = \mathbf{A}^{MDL}_{P\text{-}LOT}\mathbf{f}^*$, where $\mathbf{f}^*$ is estimated only from the weights associated with the Saint Barbe transmitter route. For computing the new weights, the semi-minor axis of all second ellipses and all ellipses for singleton clusters is set to 70% the distance between one of the foci and the center of the ellipse. The choice of the semi-minor axis for the model and measured cases, is justified as it provides the highest correlation between predicted and measured shadowing. Furthermore, a wider angle ϕ^2_t choice for the prediction is justified as no information of the propagation paths is known beforehand.

Predicted shadowing and ground truth shadowing for transmitter trajectories at Saint Barbe and Parking Lot are shown in Fig. 10. The mean value of all shadowing signals was subtracted to their signals. In addition, the predicted shadowing $\mathbf{b}^{MDL}_{P\text{-}LOT}$ was scaled by $\max(\mathbf{b}^{MEAS}_{ST-BARBE})/\max(\mathbf{b}^{MDL}_{P\text{-}LOT})$. This scaling is justified as the absolute value of the weights is not known, but rather than that, their relative value. This means that the entries of $\mathbf{A}^{MDL}_{P\text{-}LOT}$ can be in a different scale than the entries of $\mathbf{A}^{MDL}_{ST-BARBE}$, and so the same occurs to the entries of $\mathbf{b}^{MDL}_{P\text{-}LOT}$. Both transmitter trajectories shown in the right panel of Fig. 10 are very similar. For instance, there is a high Pearson correlation coefficient ($\rho(\mathbf{b}^{MEAS}_{ST-BARBE}, \mathbf{b}^{MEAS}_{P\text{-}LOT}) = 0.916$) between the shadowing extracted from measurements for both routes. This high correlation is due to the average short distance without obstacles between the routes and their alignment.

Pearson Correlation coefficient between the predicted shadowing at Parking Lot and the shadowing from measurements at Saint Barbe is $\rho(\mathbf{b}^{MEAS}_{M-CURIE}, \mathbf{b}^{MDL}_{P\text{-}LOT}) = 0.834$ (this corresponds to the interval between the black dashed lines in the left panel of Fig.10). This high correlation is showing that the prediction captures the trend of the shadowing, even though no Parking Lot

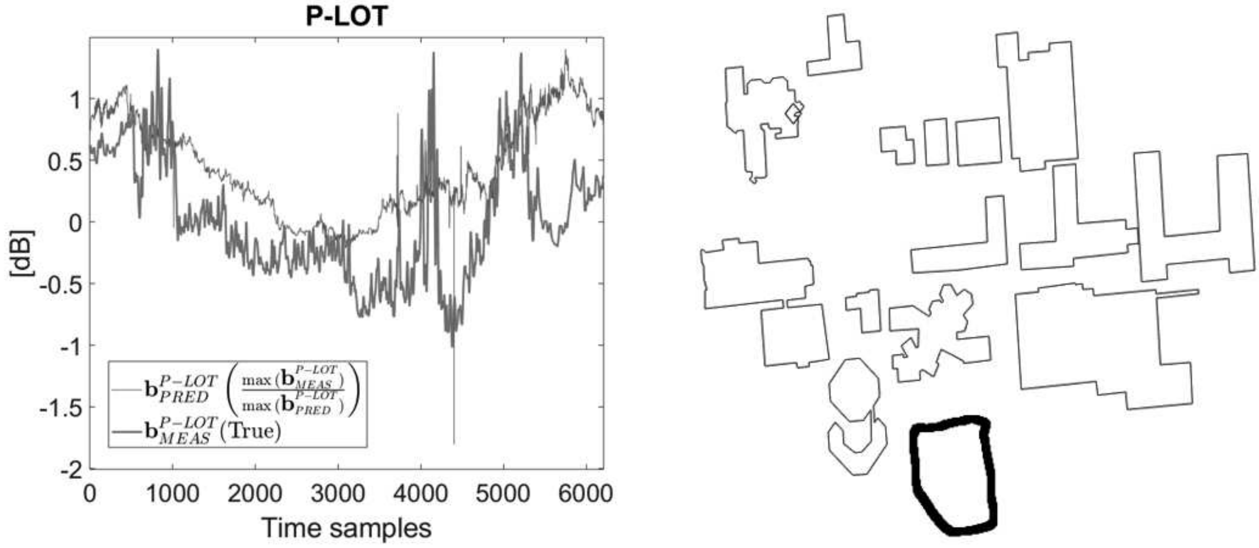

Fig. 11. Shadowing prediction for Parking Lot mobile route. The USLF was estimated from the Saint Barbe route with $N_x = N_y = 70$. Left panel: shadowing prediction and ground-truth value for the Parking lot route shown in the right panel of this figure. Right panel: Parking Lot route. This is the same route presented in Fig. 2; is shown again here for easier comparison with Fig. 12.

measurement was taken into account for the USLF computation.

Moreover, a good prediction in trend for the shadowing is in general conditioned to the estimation of the USLF. Since the locations of the transceivers do not follow the required structure shown in the left panel of Fig. 6, it is expected that predictions for the shadowing at any transmitter-receiver location will not be accurate in trend.

Fig. 11 and 12 show the shadowing predictions for 2 mobile routes. The shadowing predicted for the Parking Lot mobile route has a 0.8 correlation with the ground truth shadowing for the same route. On the other side, the shadowing predicted for the Marie Curie mobile route, does not exhibit a high correlation (the correlation is 0.54). This could be explained by the degree of alignment with respect to the base station, between the route used to estimate the USLF and the route used to predict the shadowing. We mean by the alignment, the angular difference between the line segment going from the center of each route to the base station. The lower the angular difference the higher the alignment, and better the shadowing prediction. The alignment between Saint Barbe and Parking Lot (around 10°) is higher than the one between Saint Barbe and Marie Curie (around 26°).

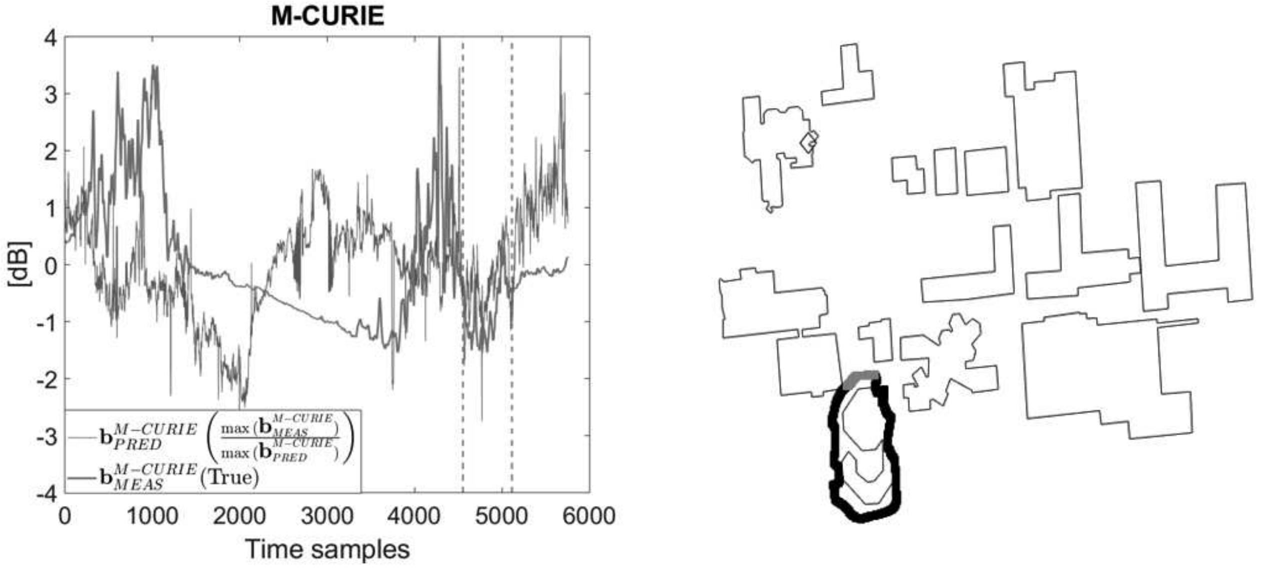

Fig. 12. Shadowing prediction for Marie Curie mobile route. The USLF was estimated from the Saint Barbe route with $N_x = N_y = 70$. Left panel: shadowing prediction and ground-truth value for the route. Right panel: Marie Curie mobile route.

VII. CONCLUSIONS

In this paper we address the problem of predicting shadowing values for new transmitter locations based on previous measurements. Although prediction is constrained by the degree of correlation among the transmitter trajectory used to estimate the USLF and the trajectory for which prediction is required, it provides an insight into the tendency of shadowing. The prediction must not be regarded as a low-error (in the sense of Mean Squared Error or Mean Absolute Error) prediction, but instead as capable of following the tendency of the real signal. The measure of this tendency is the Pearson Correlation coefficient, which accounts for long linear relationship between two signals. Rapid fluctuations will not be accounted by the Pearson Coefficient, but we are interested in the trend of the signal over a large interval of snapshots.

A situation that remains challenging and that defines the quality of the prediction is the ability to synthesize matrix **A** from a model that does not depend on the measurements, or that captures the structure of many measurements for a given environment type (i.e. semi-urban microcell). Here, by assuming omnidirectional angles of arrival to synthesize **A** at the new locations (Parking Lot route), a high correlation between predicted and ground truth signal is achieved. However, NLoS propagation conditions must probably drop omnidirectional assumption, resulting in low fitness between the real $\mathbf{A}$ and the one synthesize from a model. On the other hand, as no measurement information is assumed available for the prediction stage, the model used to synthesize $\mathbf{A}$ must assume the more general prior information (omnidirectional angles of departure), giving arise to a trade-off between prediction accuracy and prediction need for measurements.

VIII. ACKNOWLEDGMENT

This work was funded by the SWIPT project under UCLouvain ARC program. It was also carried out in the framework of COST Action CA15104 IRACON.

REFERENCES

[1] B. Blaszczyszyn, M. K. Karray, *Quality of Service in Wireless Cellular Networks Subject to Log-Normal Shadowing*, IEEE Transactions on Communications, vol. 61, no. 2, Feb. 2013.

[2] B. H. Fleury, M. Tschudin, R. Heddergott, D. Dahlhaus, K. Ingeman Pedersen, *Channel parameter estimation in mobile radio environments using the SAGE algorithm*, IEEE Journal on Selected Areas in Communications, vol. 17, no. 3, March 1999.

[3] X. Yin, B. H. Fleury, P. Jourdan, A. Stucki, *Polarization estimation of individual propagation paths using the SAGE algorithm*, 14th IEEE Proceedings on Personal, Indoor and Mobile Radio Communications, vol. 2, pp. 1795-1799, 2003.

[4] J. A. Fessler, A. O. Hero, *Space-alternating generalized expectation-maximization algorithm*, IEEE Transactions on Signal Processing, vol. 42, no. 10, Oct. 1994.

[5] P. Agrawal, N. Patwari, *Correlated Link Shadowing Fading in Multi-Hop Wireless Networks*, IEEE Transactions on Wireless Communications, vol. 8, no. 8, Aug. 2009.

[6] B.R. Hamilton, X. Ma, R.J. Baxley, S.M. Matechik, *Propagation Modeling for Radio Frequency Tomography in Wireless Networks*, IEEE Journal of Selected Topics in Signal Processing, vol. 8, pp. 55-65, 2014.

[7] N. Patwari, P. Agrawal, *NeSh: A joint shadowing model for links in a multi-hop network*, IEEE International Conference on Acoustics, Speech and Signal Processing, 2008.

[8] D. Romero, D. Lee, G.B. Giannakis, *Blind Radio Tomography*, IEEE Transactions on Signal Processing, vol. 66, pp. 2055-2069, 2018.

[9] J. Wilson, N. Patwari, F.G. Vasquez, *Regularization Methods for Radio Tomographic Imaging*, Virginia Tech. Wireless Symposium, 2009.

[10] D. Lee, S.J. Kim, G.B. Giannakis, *Channel Gain Cartography for Cognitive Radios Leveraging Low Rank and Sparsity*, IEEE Transactions on Wireless Communications, vol. 16, pp. 5953-5966, 2017.

[11] J. Wilson, N. Patwari, *Radio Tomographic Imaging with Wireless Networks*, IEEE Transactions on Mobile Computing, vol. 9, pp. 621-632, 2010.

[12] J. Wilson, N. Patwari, *See-Through Walls: Motion Tracking Using Variance-Based Radio Tomography Networks*, IEEE Transactions on Mobile Computing, vol. 10, pp. 612-621, 2011.

[13] J.D. Villegas, D. Pimienta del Valle, C. Oestges *Physical Identification of Common Clusters for Outdoor Microcells*, 13th European Conference on Antennas and Propagation (EuCAP), 2019.

[14] M. Gudmundson, *Correlation model for shadow fading in mobile radio systems*, Electronic Letters, vol. 27, pp. 2145-2146, 1991.

[15] P. Hanpinitsak, K. Saito, J. Takada, M. Kim and L. Materum, *Multipath Clustering and Cluster Tracking for Geometry-Based Stochastic Channel Modeling*, IEEE Transactions on Antennas and Propagation, vol. 65, no. 11, pp. 6015-6028, Nov. 2017.

[16] L. Liu, C. Oestges, J. Poutanen, K. Haneda, P. Vainikainen, F. Quitin, F. Tufvesson, P. De Doncker, *The COST 2100 MIMO channel model*, IEEE Wireless Communications, vol. 19, no. 6, pp. 92-99, December 2012.

[17] N. Czink, P. Cera, J. Salo, E. Bonek, J. Nuutinen and J. Ylitalo, *A Framework for Automatic Clustering of Parametric MIMO Channel Data Including Path Powers*, IEEE Vehicular Technology Conference, pp. 1-5, Montreal, 2006.

[18] D. Ganguly, S. Mukherjee, S. Naskar and P. Mukherjee, *A Novel Approach for Determination of Optimal Number of Cluster*, International Conference on Computer and Automation Engineering, pp. 113-117, Bangkok, 2009.

[19] R. He, Q. Li, B. Ai, Y. L. Geng, A. F. Molisch, V. Kristem, Z. Zhong and J. Yu, *A Kernel-Power-Density-Based Algorithm for Channel Multipath Components Clustering*, IEEE Transactions on Wireless Communications, vol. 16, no. 11, pp. 7138-7151, Nov. 2017.

[20] Y. Li, J. Zhang, Z. Ma and Y. Zhang, *Clustering Analysis in the Wireless Propagation Channel with a variational Gaussian Mixture Model*, IEEE Transactions on Big Data, 2018.

[21] D. Shutin, *Clustering wireless channel impulse responses in angular-delay domain*, IEEE 5th Workshop on Signal Processing Advances in Wireless Communications, pp. 253-257, Lisbon, 2004.

[22] R. He, B. Ai, A. F. Molisch, G. L Stber, Q. Li, Z. Zhong and J. Yu, *Clustering Enabled Wireless Channel Modeling Using Big Data Algorithms*, IEEE Communications Magazine, vol. 56, no. 5, pp. 177-183, May 2018.

[23] F. Luan, A. F. Molisch, L. Xiao, F. Tufvesson and S. Zhou, *Geometrical Cluster-Based Scatterer Detection Method with the Movement of Mobile Terminal*, IEEE 81st Vehicular Technology Conference (VTC Spring), pp. 1-6, Glasgow, 2015.

[24] J. Poutanen, *Geometry-based radio channel modeling: Propagation analysis and concept development*,PhD dissertation, Aalto University, 2011.

[25] M. Zhu, *Geometry-based Radio Channel Characterization and Modeling: Parameterization, Implementation and Validation*, PhD dissertation, Lund University, 2014.

[26] N. Czink, X. Yin, H. OZcelik, M. Herdin, E. Bonek and B. H. Fleury, *Cluster Characteristics in a MIMO Indoor Propagation Environment*, IEEE Transactions on Wireless Communications, vol. 6, no. 4, pp. 1465-1475, April 2007.

[27] N. Czink, X. Yin, H. OZcelik, M. Herdin, E. Bonek and B. H. Fleury, *Cluster Characteristics in a MIMO Indoor Propagation Environment*, IEEE Transactions on Wireless Communications, vol. 6, no. 4, pp. 1465-1475, April 2007.

[28] B.H. Fleury, P. Jourdan, A. Stucki, *High-resolution channel parameter estimation for MIMO applications using the SAGE algorithm*, 2002 International Zurich Seminar on Broadband Communications Access - Transmission - Networking.

[29] N. Czink, G. Del Galdo, X. Yin, E. Bonek and J. Ylitalo, *A novel environment characterization metric for clustered MIMO channels*, Wireless Personal Communications, vol. 46, no. 1, pp. 83-98, 2008.

[30] B.H. Fleury, X. Yin, K.G. Rohbrandt, P. Jourdan and A. Stucki, *High-Resolution Bidirection Estimation Based on the SAGE Algorithm: Experience Gathered from Field Experiments*, Proc. XXVIIth General Assmply of the Int. Union of Radio Scientist (URSI), 2002.