On the Spatial Predictability of Communication Channels

Mehrzad Malmirchegini, *Student Member, IEEE*, and Yasamin Mostofi, *Member, IEEE*

Abstract—In this paper, we are interested in fundamentally understanding the spatial predictability of wireless channels. We propose a probabilistic channel prediction framework for predicting the spatial variations of a wireless channel, based on a small number of measurements. By using this framework, we then develop a mathematical foundation for understanding the spatial predictability of wireless channels. More specifically, we characterize the impact of different environments, in terms of their underlying parameters, on wireless channel predictability. We furthermore show how sampling positions can be optimized to improve the prediction quality. Finally, we show the performance of the proposed framework in predicting (and justifying the predictability of) the spatial variations of real channels, using several measurements in our building.

Index Terms—Spatial predictability, Wireless channels, Probabilistic modeling and estimation.

I. INTRODUCTION

IN the past few years, the sensor network revolution has created the possibility of exploring and controlling the environment in ways not possible before[2], [3]. The vision of a multi-agent robotic network cooperatively learning and adapting in harsh unknown environments to achieve a common goal is closer than ever. Since each agent has a limited sensing capability, the group relies on networked sensing and decision-making to accomplish the task. Thus, maintaining connectivity becomes considerably important in such networks. In the robotics and control community, considerable progress has been made in the area of networked robotic and control systems [4]. However, ideal or over-simplified models have typically been used to model the communication links among agents. For instance, disk models are commonly used, where the link quality is assumed above an acceptable threshold in a disk around the transmitter, with no connectivity outside of the disk, as shown in Fig. 1 (top-left).

In order to realize the full potentials of these networks, an integrative approach to communication and motion planning issues is essential, i.e., each robot should have an awareness of the impact of its motion decisions on link qualities, when planning its trajectory [5]. This requires each robot to assess the quality of the communication link in the locations that it has not yet visited. As a result, proper prediction of the communication signal strength and *fundamentally understanding the spatial predictability of a wireless channel*, based on
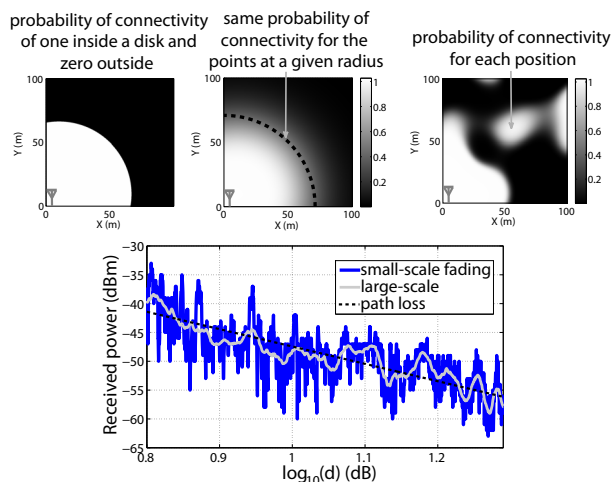
Fig. 1: (top) Different connectivity models for the communication channel to the fixed transmitter at (0,0) coordinate: (top-left) simplified disc model that is commonly used in the robotic-network literature (top-middle) our probabilistic path loss model, and (top-right) our general probabilistic model. (bottom) underlying dynamics of the received signal power across a route in the basement of ECE building.

only a few measurements, become considerably important. In the communications community, rich literature was developed, over the past decades, for the characterization and modeling of wireless channels [6]–[9]. If all the information about object positions, geometry and dielectric properties is available, ray tracing methods could be used to model the spatial variations of the received signal strength in a given area [10]. However, such approaches require knowing the environment, in terms of locations of the objects and their dielectric properties, which is prohibitive for real-time networked robotic applications. Furthermore, such approaches can not provide a fundamental understanding of wireless channel predictability. In the wireless communication literature, it is well established that a communication channel between two nodes can be probabilistically modeled as a multi-scale dynamical system with three major dynamics: small-scale fading (multipath), large-scale fading (shadowing) and path loss [6]–[8]. Fig. 1, for instance, shows the received signal power across a route in the basement of the ECE building at UNM. The three main dynamics are marked on the figure. The measured received signal is the small-scale fading. In order to extract the large-scale component, the received signal should be averaged locally over a distance of 5λ to 40λ (depending on the scenario), where λ is the transmission wavelength [7], [11]. In the example of Fig. 1, for instance, we averaged the channel locally over the length of $5\lambda = 62.5$cm, by using a moving average (frequency of operation is 2.4GHz). Once we have

the large-scale component, the distance-dependent path loss is calculated by finding the best line fit to the log of the received measurements [6], [7], [12].

It is the goal of this paper to utilize such probabilistic link models and fundamentally characterize the spatial predictability of a wireless channel. More specifically, we build on our previously proposed channel prediction work [1], where we developed a probabilistic framework in which each robot can spatially predict the channel, based on a small number of measurements. Fig. 1 (top-middle) and Fig. 1 (top-right), for instance, illustrate how this framework enables a more realistic characterization of wireless channels and their connectivity, as compared to the commonly-used disk model of Fig. 1 (top-left). In this paper, we then mathematically characterize the impact of different environments, in terms of their underlying parameters, on channel spatial predictability. Furthermore, we show the optimum distribution of the sparse sampling positions in order to maximize channel predictability. We emphasize that we are not suggesting that a wireless channel is fully predictable, as it is not. Rather, our goal is to develop a mathematical characterization of how predictable a wireless channel can be and understand the impact of different underlying parameters on its predictability. Thus, we also test our mathematical framework on real channel measurements in Section VI, where we show the impact of different environments on wireless channel predictability. In general, predicting the spatial variations of a random field, based on sparse sampling, has also been of interest in other areas such as meteorology, ecological systems, and acoustic field estimation, just to name a few [13], [14]. However, to the best of authors' knowledge, no framework has yet been developed to mathematically characterize and understand the spatial predictability of a general random field or wireless channels in particular. As such, the contribution of this paper is beyond only understanding the spatial predictability of wireless channels and can possibly benefit other areas that require estimation of a random field, based on sparse measurements. The rest of the paper is organized as follows. In Section II, we describe our proposed probabilistic channel prediction framework. In Section III, we mathematically characterize the impact of different underlying channel parameters on the prediction performance, assuming perfectly-estimated path loss parameters. In Section IV, we mathematically characterize the impact of different environments and sampling positions on the estimation of path loss parameters and show how to optimize the positions of the sparse samples. Then, in Section V, we extend the analysis of Section III to characterize wireless channel predictability in the presence of path loss estimation error. In Section VI, we show the performance of the proposed framework in predicting (and understanding the predictability of) the spatial variations of real channels, using several measurements in our building. We conclude in Section VII.

II. MODEL-BASED ESTIMATION OF CHANNEL SPATIAL VARIATIONS

As mentioned in the previous section, a communication channel between two nodes can be modeled as a multi-scale dynamical system with three major dynamics: small-scale fading (multipath), large-scale fading (shadowing) and path loss. Let $\Upsilon_{\text{RX}}(q)$ denote the received signal strength (power), in the transmission from a fixed transmitter at $q_b \in \mathcal{K}$ to a mobile node at $q \in \mathcal{K}$, where $\mathcal{K} \subset \mathbb{R}^2$ denotes the workspace. Consider the case where the channel to the fixed transmitter is narrowband. Furthermore, assume that the workspace is not changing with time, i.e. the environmental features that impact the wireless transmission in the workspace are time-invariant. Our proposed framework can be extended to time-varying environments, as we briefly discuss later in this section. Then, we have the following at the output of the power detector: $\Upsilon_{\text{RX}}(q) = g(q)P_T + \varrho$, where P_T and $g(q)$ denote the transmitted power and channel gain (square of the amplitude of the baseband equivalent channel), at position q, respectively and ϱ represents the power of the receiver thermal noise [6]. Define $\Upsilon(q) \triangleq \Upsilon_{\text{RX}}(q) - \varrho$. We assume that the receiver can estimate and remove the noise power to obtain $\Upsilon(q)$.[1] $\Upsilon(q)$ is proportional to $g(q)$ and can be modeled as a multiscale dynamical system with three major dynamics: multipath fading, shadowing and path loss. We can then characterize $\Upsilon(q)$ by a 2D non-stationary random field with the following form [6]: $\Upsilon(q) = \Upsilon_{\text{PL}}(q)\Upsilon_{\text{SH}}(q)\Upsilon_{\text{MP}}(q)$, where $\Upsilon_{\text{MP}}(q)$ and $\Upsilon_{\text{SH}}(q)$ are random variables representing the impact of multipath fading and shadowing components respectively and $\Upsilon_{\text{PL}}(q) = \frac{K_{\text{PL}}}{\|q - q_b\|^{n_{\text{PL}}}}$ is the distance-dependent path loss.[2] In this model, the multipath fading coefficient, $\Upsilon_{\text{MP}}(q)$, has a unit average. Let $\Upsilon_{\text{dB}}(q) = 10 \log_{10}\left(\Upsilon(q)\right)$ represent the received signal strength in dB. We have

$$\Upsilon_{\text{dB}}(q) = \underbrace{10 \log_{10}\left(K_{\text{PL}}\right) + \overline{\Upsilon}_{\text{MP, dB}}}_{K_{\text{dB}}} - 10 n_{\text{PL}} \log_{10}\left(\|q - q_b\|\right) + \nu(q) + \omega(q), \qquad (1)$$

where $\overline{\Upsilon}_{\text{MP, dB}} = 10\,\mathbb{E}\left\{\log_{10}\left(\Upsilon_{\text{MP}}(q)\right)\right\}$ is the average of the multipath fading in dB, $\nu(q) = 10 \log_{10}\left(\Upsilon_{\text{SH}}(q)\right)$ is a zero-mean random variable representing the shadowing effect in dB and $\omega(q) = 10 \log_{10}\left(\Upsilon_{\text{MP}}(q)\right) - \overline{\Upsilon}_{\text{MP, dB}}$ is a zeromean random variable, independent of $\nu(q)$, which denotes the impact of multipath fading in dB, after removing its average. In the communication literature, the distributions of $\Upsilon_{\text{MP}}(q)$ and $\Upsilon_{\text{SH}}(q)$ (or equivalently the distributions of $\omega(q)$ and $\nu(q)$) are well established based on empirical data [8]. For instance, Nakagami distribution is shown to be a good match for the distribution of $\Upsilon_{\text{MP}}(q)$ in several environments [6]. In this case, we have the following Nakagami distribution, with parameter m and unit average, for the distribution of $\Upsilon_{\text{MP}}(q)$: $f_{\Upsilon_{\text{MP}}}(x) = \frac{m^m x^{m-1}}{\Gamma(m)} e^{-mx}$, where $\Gamma(.)$ represents the Gamma function. This then results in the following distribution for

---

[1]Most related device drivers provide an estimate on the noise power. MadWiFi, for instance, estimates the noise power by using the often-used formula of $K_{\text{Bol}} \times T_{\text{env}} \times BW$ [15], where K_{Bol} is the Boltzmann's constant, T_{env} is the environment temperature and BW is the utilized bandwidth. Its newer versions can even provide a better online assessment, by using measurements from the silent mode (when no transmission) [16].

[2]In this paper, we follow the convention of [7] and use the term "shadowing" to refer to the large-scale fading after its mean (path loss) is removed in the dB domain. More specifically, Υ_{SH} is the large-scale fading after its average (path loss) is removed in the dB domain. Furthermore, we use the term "multipath fading" to refer to the normalized small-scale fading, i.e. with unit average. Then, Υ_{MP} is the normalized small-scale fading.

$\omega(q)$: $f_\omega(x) = \frac{\ln(10)}{10} 10^{(x+\overline{\Upsilon}_{\text{MP, dB}})/10} f_{\Upsilon_{\text{MP}}}\left(10^{(x+\overline{\Upsilon}_{\text{MP, dB}})/10}\right)$. Some experimental measurements have also suggested Gaussian to be a good enough yet simple fit for the distribution of $\omega(q)$ [17]. We will take advantage of this Gaussian simplification later in our framework. As for the shadowing variable, log-normal is shown to be a good match for the distribution of $\Upsilon_{\text{SH}}(q)$. Then, we have the following zero-mean Gaussian pdf for the distribution of $\nu(q)$: $f_\nu(x) = \frac{1}{\sqrt{2\pi\alpha}} e^{-x^2/2\alpha}$, where α is the variance of the shadowing variations around path loss.

Characterizing the spatial correlation of $\omega(q)$ and $\nu(q)$ is also considerably important for our model-based channel prediction framework. However, we do not attempt to predict the multipath component, $\omega(q)$, due to the fact that it typically decorrelates fast and that the form of its correlation function can change considerably, depending on the angle of arrival and position of the scatterers. Therefore, in our proposed framework we only predict the path loss and shadowing components of the channel. The impact of multipath will then appear in the characterization of the prediction error variance, as we shall see. As for the spatial correlation of shadowing, [18] characterizes an exponentially-decaying spatial correlation function, which is widely used: $\mathbb{E}\left\{\nu(q_1)\nu(q_2)\right\} = \alpha e^{-\|q_1-q_2\|/\beta}$, for $q_1, q_2 \in \mathcal{K}$ where α denotes the *shadowing power* and the *correlation distance*, β, controls the spatial correlation of the channel [18]. For some examples of a time-varying environment, Oestges et al. model the slow temporal-variation of the channel as a zero-mean Gaussian variable with an exponential temporal correlation in the dB domain [27]. Thus, Eq. 1 can be extended to such time-varying cases by adding this additional variable to Eq. 1. In general, however, finding one model for characterizing the time-variations of different features in the environment is a challenging task and a subject of further studies. Such temporal variations can also be treated as disturbance in the prediction process.

Next, we describe our proposed model-based channel prediction framework. Consider the case where a wireless channel to a fixed transmitter is sparsely sampled at positions $\mathcal{Q} = \{q_1, q_2, \cdots, q_k\} \subset \mathcal{K}$, in a given environment. These channel measurements can be gathered by one or a number of cooperative homogenous robots, equipped with identical receivers, making measurements along their trajectories. Let a region or an environment refer to an area over which the underlying channel parameters, such as α and β, can be considered constant. The four marked areas of Fig. 10 are examples of such regions. First, consider the case that all the k measurements belong to one region and that we are predicting the channel in the same region. We show how to relax this assumption later in this section. Let $D_\mathcal{Q}$ and $Y_\mathcal{Q} = [y_1, \cdots, y_k]^T \in \mathbb{R}^k$ denote the corresponding distance vector to the transmitter in dB and the vector of all the available channel measurements (in dB) respectively:
$D_\mathcal{Q} = \left[10\log_{10}(\|q_1 - q_b\|), \cdots, 10\log_{10}(\|q_k - q_b\|)\right]^T$ and $Y_\mathcal{Q} = [y_1, \cdots, y_k]^T \in \mathbb{R}^k$. We have,

$$Y_\mathcal{Q} = \underbrace{\begin{bmatrix} 1_k & -D_\mathcal{Q} \end{bmatrix}}_{H_\mathcal{Q}} \theta + \vartheta_\mathcal{Q} + \Omega_\mathcal{Q}, \qquad (2)$$

where 1_k denotes the vector of ones with the length of k, $\theta = $

$[K_{\text{dB}} \ n_{\text{PL}}]^T$ is the vector of the path loss parameters, $\vartheta_\mathcal{Q} = [\nu_1, \cdots, \nu_k]^T$ with $\nu_i = \nu(q_i)$ and $\Omega_\mathcal{Q} = [\omega_1, \cdots, \omega_k]^T$ with $\omega_i = \omega(q_i)$, for $i = 1, \cdots, k$. Based on the log-normal model for shadowing, $\vartheta_\mathcal{Q}$ is a zero-mean Gaussian random vector with the covariance matrix $R_\mathcal{Q} \in \mathbb{R}^{k \times k}$, where $\left[R_\mathcal{Q}\right]_{i,j} = \alpha e^{-\|q_i-q_j\|/\beta}$, for $q_i, q_j \in \mathcal{Q}$. The term $\Omega_\mathcal{Q}$ denotes the impact of multipath fading in dB domain. As mentioned earlier, some empirical data have shown Gaussian to be a good match for the distribution of w_i [17]. For instance, Fig. 2 compares the match of both Nakagami and lognormal to the distribution of multipath fading (Υ_{MP}) for a stationary section of our collected data of Fig. 1. As can be seen, Nakagami provides a considerably good match while lognormal can be acceptable, depending on the required accuracy. Thus, in order to facilitate the mathematical derivations in our prediction framework, we take w_i to have a Gaussian distribution. In addition, multipath fading typically decorrelates considerably fast, making learning of its correlation function, based on sparse possibly non-localized samples, considerably challenging if not infeasible. There is also no one general function that can properly model its correlation in all the environments as its form depends heavily on the angle of arrival and position of the scatterers. While approaches based on the estimation of the power spectrum and linear prediction have been utilized to predict the immediate values of multipath, based on past observations, such approaches require dense sampling in order to capture correlated multipath samples. Finally, even if its correlation function is learned, it typically can not be taken advantage of, in the prediction framework, unless the location of the channel to be predicted is very close to the position of one of the available measurements. Thus, we take $\Omega_\mathcal{Q}$ to be an uncorrelated zero-mean Gaussian vector with the covariance of $\mathbb{E}\left\{\Omega_\mathcal{Q}\Omega_\mathcal{Q}^T\right\} = \sigma^2 I_{k \times k}$, where $I_{k \times k}$ is a $k \times k$ identity matrix and $\sigma^2 = \mathbb{E}\left\{\omega^2(q)\right\} = 100\int_0^\infty \log_{10}^2(x) f_{\Upsilon_{\text{MP}}}(x)dx - 100\left(\int_0^\infty \log_{10}(x) f_{\Upsilon_{\text{MP}}}(x)dx\right)^2$ is the power of multipath fading (in dB domain). In other words, our framework does not attempt to predict the multipath component and assumes the worst case of uncorrelated multipath (worst from a prediction standpoint). The estimated variance of multipath then appears in our assessment of channel prediction error variance, as we shall see. Note, however, that this is only for the purpose of our modeling. When we show the performance of this framework, we use real measurements where the multipath component will have its natural distribution and correlation function. We then define $\Xi_\mathcal{Q} \triangleq \vartheta_\mathcal{Q} + \Omega_\mathcal{Q}$, which is a zero-mean Gaussian vector with the covariance matrix of $R_{\text{tot},\mathcal{Q}} \triangleq R_\mathcal{Q} + \sigma^2 I_{k \times k}$. In our model-based probabilistic framework, we first need to estimate the parameters of the model (θ, α, β and σ^2) and then use these parameters to estimate the channel. Let $f_{Y_\mathcal{Q}}(Y_\mathcal{Q}|\theta, \alpha, \beta, \sigma^2)$ denote the conditional pdf of $Y_\mathcal{Q}$, given the parameters θ, α, β and σ^2. Under the assumption of independent multipath fading variables, Eq. 2 will result in the following:

$$f_{Y_\mathcal{Q}}(Y_\mathcal{Q}|\theta, \alpha, \beta, \sigma^2)$$
$$= \frac{e^{-\frac{1}{2}\left(Y_\mathcal{Q}-H_\mathcal{Q}\theta\right)^T \left(\alpha R_{\text{norm},\mathcal{Q}}(\beta)+\sigma^2 I_{k \times k}\right)^{-1} \left(Y_\mathcal{Q}-H_\mathcal{Q}\theta\right)}}{(2\pi)^{k/2}\left(\det\left[\alpha R_{\text{norm},\mathcal{Q}}(\beta)+\sigma^2 I_{k \times k}\right]\right)^{1/2}}$$
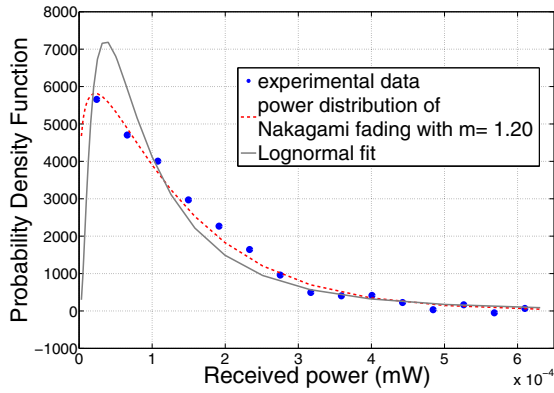
Fig. 2: Comparison of Nakagami and lognormal for the distribution of small-scale fading.

where $R_{\text{norm},\mathcal{Q}} = \frac{1}{\alpha}R_{\mathcal{Q}}$ denotes the normalized version of $R_{\mathcal{Q}}$. Next, we characterize the Maximum Likelihood (ML) estimation of the underlying channel parameters.

$$[\hat{\theta}_{\text{ML}}, \hat{\alpha}_{\text{ML}}, \hat{\beta}_{\text{ML}}, \hat{\sigma}^2_{\text{ML}}]$$
$$= \operatorname{argmax}_{\theta,\alpha,\beta,\sigma^2} \ln\left(f_{Y_{\mathcal{Q}}}(Y_{\mathcal{Q}}|\theta,\alpha,\beta,\sigma^2)\right)$$
$$= \operatorname{argmin}_{\theta,\alpha,\beta,\sigma^2} \left(Y_{\mathcal{Q}} - H_{\mathcal{Q}}\theta\right)^T \left(\alpha R_{\text{norm},\mathcal{Q}}(\beta) + \sigma^2 I_{k\times k}\right)^{-1}$$
$$\times \left(Y_{\mathcal{Q}} - H_{\mathcal{Q}}\theta\right) + \ln\left(\det\left[\alpha R_{\text{norm},\mathcal{Q}}(\beta) + \sigma^2 I_{k\times k}\right]\right),$$

which results in:

$$\hat{\theta}_{\text{ML}} = \left(H_{\mathcal{Q}}^T\left(\hat{\alpha}_{\text{ML}}R_{\text{norm},\mathcal{Q}}(\hat{\beta}_{\text{ML}}) + \hat{\sigma}^2_{\text{ML}}\right)^{-1}H_{\mathcal{Q}}\right)^{-1}$$
$$\times H_{\mathcal{Q}}^T\left(\hat{\alpha}_{\text{ML}}R_{\text{norm},\mathcal{Q}}(\hat{\beta}_{\text{ML}}) + \hat{\sigma}^2_{\text{ML}}\right)^{-1}Y_{\mathcal{Q}}. \quad (3)$$

Finding a closed-form expression for $\hat{\alpha}_{\text{ML}}, \hat{\beta}_{\text{ML}}$ and $\hat{\sigma}^2_{\text{ML}}$, however, is challenging. For the special case where Ω_k is negligible, the ML estimation of channel parameters can be simplified to:

$$\hat{\theta}_{\text{ML},\sigma^2=0} =$$
$$\left(H_{\mathcal{Q}}^T R_{\text{norm},\mathcal{Q}}^{-1}(\hat{\beta}_{\text{ML},\sigma^2=0})H_{\mathcal{Q}}\right)^{-1} H_{\mathcal{Q}}^T R_{\text{norm},\mathcal{Q}}^{-1}(\hat{\beta}_{\text{ML},\sigma^2=0})Y_{\mathcal{Q}},$$

$$\hat{\alpha}_{\text{ML},\sigma^2=0} = \frac{1}{k}\left(Y_{\mathcal{Q}} - H_{\mathcal{Q}}\hat{\theta}_{\text{ML},\sigma^2=0}\right)^T R_{\text{norm},\mathcal{Q}}^{-1}(\hat{\beta}_{\text{ML},\sigma^2=0})$$
$$\times \left(Y_{\mathcal{Q}} - H_{\mathcal{Q}}\hat{\theta}_{\text{ML},\sigma^2=0}\right),$$

$$\hat{\beta}_{\text{ML},\sigma^2=0} = \arg\min_{\beta} \left[Y_{\mathcal{Q}}^T P_{\mathcal{Q},\text{ML}}^T(\beta)R_{\text{norm},\mathcal{Q}}^{-1}(\beta)P_{\mathcal{Q},\text{ML}}(\beta)Y_{\mathcal{Q}}\right]^k$$
$$\times \det\left[R_{\text{norm},\mathcal{Q}}(\beta)\right], \quad (4)$$

where $P_{\mathcal{Q},\text{ML}}(\beta) = I_{k\times k} - H_{\mathcal{Q}}\left(H_{\mathcal{Q}}^T R_{\text{norm},\mathcal{Q}}^{-1}(\beta)H_{\mathcal{Q}}\right)^{-1}H_{\mathcal{Q}}^T R_{\text{norm},\mathcal{Q}}^{-1}(\beta)$. Under the assumption that β is known, it can be shown that $\hat{\theta}_{\text{ML},\sigma^2=0}$ is an unbiased estimator and achieves the Cramer-Rao bound. Furthermore, for large number of sampling points k, we can show that $\hat{\alpha}_{\text{ML},\sigma^2=0}$ is unbiased and achieves the Cramer-Rao bound as well. We skipped the details of the proofs due to the space limitation. The ML estimator will therefore be our benchmark in the estimation of the channel parameters.

As can be seen, in order to estimate θ and α, we first need to estimate β, which is challenging. Furthermore, finding the

ML estimation of the channel parameters for the general case, where $\sigma^2 \neq 0$, is computationally complex. Therefore, we next devise a suboptimum but simpler estimation strategy. Let $\chi = \alpha + \sigma^2$ denote the sum of the shadowing and multipath powers. A Least Square (LS) estimation of θ and χ then results in:

$$\hat{\theta}_{\text{LS}} = \left(H_{\mathcal{Q}}^T H_{\mathcal{Q}}\right)^{-1}H_{\mathcal{Q}}^T Y_{\mathcal{Q}}, \quad (5)$$

$$\hat{\chi}_{\text{LS}|\theta=\hat{\theta}_{\text{LS}}} = \frac{1}{k}Y_{\mathcal{Q}}^T\left(I_{k\times k} - H_{\mathcal{Q}}\left(H_{\mathcal{Q}}^T H_{\mathcal{Q}}\right)^{-1}H_{\mathcal{Q}}^T\right)^2 Y_{\mathcal{Q}}$$
$$= \frac{1}{k}Y_{\mathcal{Q}}^T\left(I_{k\times k} - H_{\mathcal{Q}}\left(H_{\mathcal{Q}}^T H_{\mathcal{Q}}\right)^{-1}H_{\mathcal{Q}}^T\right)Y_{\mathcal{Q}}, \quad (6)$$

where $H_{\mathcal{Q}}$ is full rank, except for the case where the samples are equally-distanced from the transmitter. Since such a special case is very low probable, we assume that $H_{\mathcal{Q}}$ is full rank throughout the paper unless otherwise is stated. We refer to this suboptimal approach as LS throughout the paper. We next discuss a more practical but suboptimum strategy to estimate β. Let $\mathcal{I}_l = \left\{(i,j)|q_i, q_j \in \mathcal{Q} \text{ such that } \|q_i - q_j\| = l\right\}$ denote the pairs of points in $\mathcal{Q}$ which are located at distance l from each other. Let $Y_{\mathcal{Q},\text{cent,LS}} = \left(I_{k\times k} - H_{\mathcal{Q}}\left(H_{\mathcal{Q}}^T H_{\mathcal{Q}}\right)^{-1}H_{\mathcal{Q}}^T\right)Y_{\mathcal{Q}}$ represent the centered version of the measurement vector, when path loss parameters are estimated using the LS estimator of Eq. 5. Define $\hat{r}_{\mathcal{Q}}(l) \triangleq \frac{1}{|\mathcal{I}_l|}\sum_{(i,j)\in\mathcal{I}_l}[Y_{\mathcal{Q},\text{cent,LS}}]_i[Y_{\mathcal{Q},\text{cent,LS}}]_j$ to be the numerical estimate of the spatial correlation function at distance l, where $|.|$ represents the cardinality of the argument set and $[.]_i$ denotes the ith element of the argument vector. We have $[\hat{\alpha}_{\text{LS}}, \hat{\beta}_{\text{LS}}] = \arg\min_{\alpha,\beta}\sum_{l\in\mathcal{L}_{\mathcal{Q}}} w(l)\left[\ln\left(\alpha e^{-l/\beta}\right) - \ln\left(\hat{r}_{\mathcal{Q}}(l)\right)\right]^2$, where $\mathcal{L}_{\mathcal{Q}} = \{l|0 < \hat{r}_{\mathcal{Q}}(l) < \hat{\chi}_{\text{LS}|\theta=\hat{\theta}_{\text{LS}}}\}$ and $w(l)$ can be chosen based on our assessment of the accuracy of the estimation of $\hat{r}_{\mathcal{Q}}(l)$. For instance, if we have very few pairs of measurements at a specific distance, then the weight should be smaller. Let $\mathcal{L}_{\mathcal{Q}} = \{l_1, l_2, \cdots, l_{|\mathcal{L}_{\mathcal{Q}}|}\}$ denote an ordered set of all the possible distances among the measurement points. We have the following Least Square estimator of α and β:
$$\begin{bmatrix}\ln(\hat{\alpha}_{\text{LS}}) \\ \frac{1}{\hat{\beta}_{\text{LS}}}\end{bmatrix} = (M_{\mathcal{L}_{\mathcal{Q}}}^T W_{\mathcal{L}_{\mathcal{Q}}} M_{\mathcal{L}_{\mathcal{Q}}})^{-1}M_{\mathcal{L}_{\mathcal{Q}}}^T W_{\mathcal{L}_{\mathcal{Q}}}b \text{ where } M_{\mathcal{L}_{\mathcal{Q}}} = \begin{bmatrix}1 & -l_1 \\ \vdots & \vdots \\ 1 & -l_{|\mathcal{L}_{\mathcal{Q}}|}\end{bmatrix}, b = \begin{bmatrix}\ln\left(\hat{r}_{\mathcal{Q}}(l_1)\right) \\ \vdots \\ \ln\left(\hat{r}_{\mathcal{Q}}(l_{|\mathcal{L}_{\mathcal{Q}}|})\right)\end{bmatrix}$ and $W_{\mathcal{L}_{\mathcal{Q}}} = \text{diag}[w(l_1), \cdots, w(l_{|\mathcal{L}_{\mathcal{Q}}|})]$. We then have, $\hat{\sigma}^2_{\text{LS}} = \hat{\chi}_{\text{LS}|\theta=\hat{\theta}_{\text{LS}}} - \hat{\alpha}_{\text{LS}}$ for the estimation of the multipath power (in dB domain). Note that the estimated values of the shadowing parameters should satisfy: $0 < \hat{\alpha}_{\text{LS}} \leq \hat{\chi}_{\text{LS}|\theta=\hat{\theta}_{\text{LS}}}$ and $\hat{\beta}_{\text{LS}} > 0$. If due to the lack of enough measurements, any of these are violated, we take $\hat{\alpha}_{\text{LS}}$ and $\hat{\beta}_{\text{LS}}$ to be zero. This means that, in this case, we can not estimate the correlated part of the channel.

Once the underlying parameters of our model are estimated, channel at position $q \in \mathcal{K}$ can be estimated as follows. We have the following for the probability distribution of $\Upsilon_{\text{dB}}(q)$, conditioned on all the gathered measurements and the underlying parameters: $f(\Upsilon_{\text{dB}}(q)|Y_{\mathcal{Q}},\theta,\alpha,\beta,\sigma^2) \sim \mathcal{N}\left(\tilde{\Upsilon}_{\text{dB},\mathcal{Q}}(q), \sigma^2_{\text{dB},\mathcal{Q}}(q)\right)$ with

$$\tilde{\Upsilon}_{\text{dB},\mathcal{Q}}(q) \triangleq \mathbb{E}\left\{\Upsilon_{\text{dB}}(q) \,\Big|\, Y_{\mathcal{Q}},\theta,\alpha,\beta,\sigma^2\right\}$$

$$= h^T(q)\theta + \phi_{\mathcal{Q}}^T(q)R_{\text{tot},\mathcal{Q}}^{-1}(Y_{\mathcal{Q}} - H_{\mathcal{Q}}\theta) \text{ and}$$

$$\sigma_{\text{dB},\mathcal{Q}}^2(q) \triangleq \mathbb{E}\left\{\left(\Upsilon_{\text{dB}}(q) - \tilde{\Upsilon}_{\text{dB},\mathcal{Q}}(q)\right)^2 \,\middle|\, \theta, \alpha, \beta, \sigma^2\right\}$$

$$= \alpha + \sigma^2 - \phi_{\mathcal{Q}}^T(q)R_{\text{tot},\mathcal{Q}}^{-1}\phi_{\mathcal{Q}}(q), \qquad (7)$$

where $h(q) = \begin{bmatrix} 1 & -D_{\{q\}} \end{bmatrix}^T$, $D_{\{q\}} = 10\log_{10}\left(\|q - q_b\|\right)$ and $\phi_{\mathcal{Q}}(q) = \alpha\left[e^{-\frac{\|q_1 - q\|}{\beta}}, \cdots, e^{-\frac{\|q_k - q\|}{\beta}}\right]^T$ denotes the cross covariance between $\mathcal{Q}$ and q. Therefore, the Minimum Mean Square Error (MMSE) estimation of $\Upsilon_{\text{dB}}(q)$, assuming perfect estimation of the underlying parameters, is given by $\hat{\Upsilon}_{\text{dB},\mathcal{Q}}(q)$. We then have the following by considering the true estimated parameters: $\hat{\Upsilon}_{\text{dB},\mathcal{Q}}(q) = h^T(q)\hat{\theta} + \hat{\phi}_{\mathcal{Q}}^T(q)\hat{R}_{\text{tot},\mathcal{Q}}^{-1}(Y_{\mathcal{Q}} - H_{\mathcal{Q}}\hat{\theta})$, where $\hat{\phi}_{\mathcal{Q}}(q) = \left[\hat{\alpha}\,e^{-\|q - q_1\|/\hat{\beta}}, \cdots, \hat{\alpha}\,e^{-\|q - q_k\|/\hat{\beta}}\right]^T$ and $\hat{R}_{\text{tot},\mathcal{Q}} = \hat{\alpha}R_{\text{norm},\mathcal{Q}}(\hat{\beta}) + \hat{\sigma}^2 I_{k\times k}$.

The prediction quality at position q improves, the more correlated the available channel measurements become with the value of the channel at position q. In order to mathematically assess this, the next lemma characterizes the average number of the available measurements at the β neighborhood of the point to be predicted, for the case of randomly-distributed available channel measurements in 1D. The β-neighborhood of a point q, in the workspace $\mathcal{K}$, is defined as $\{z \in \mathcal{K}|d(z,q) < \beta\}$, where $d(z,q)$ denotes the Euclidian distance between points z and q.

Lemma 1: Consider the case that k channel measurements, at positions $\{q_1, q_2, \cdots, q_k\}$ are available, for predicting the channel at point q. Let $N_\beta(Q,q)$ represent the number of points in $Q = \{q_i\}_{i=1}^k$, which are located in the β neighborhood of q, where q and $\{q_i\}_{i=1}^k$ are i.i.d. random positions, uniformly distributed over the workspace $\mathcal{K} = [0, L]$. We then have, $\overline{N_\beta(Q,q)} = k\left(2\frac{\beta}{L} - \frac{\beta^2}{L^2}\right)$, where $\overline{N_\beta(Q,q)} = \mathbb{E}_{Q,q}\left\{N_\beta(Q,q)\right\}$ and $\mathbb{E}_{Q,q}\{.\}$ represents the expected value w.r.t. Q and q.

Proof: The proof is straightforward. ∎

Special case - probabilistic path loss: If the knowledge of beta is not available or is not used in the prediction (thus beta is assumed zero), then Eq. 7 results in the same probability distribution for all the points that are equally-spaced from the transmitter. An example of this case can be seen in Fig. 1 (top-middle), where we have the same predicted probability of connectivity (probability that the $\Upsilon_{\text{dB}}(q)$ is above a given threshold) for all the points at a given radius from the transmitter. Our more general case of Eq. 7 is then shown in Fig. 1 (top-right), where a probability distribution (and a resulting probability of connectivity) is assigned to each point in the workspace. Both these cases result in a more comprehensive channel prediction than the commonly-used disk model of Fig. 1 (top-left).

Next we show the reconstruction of two real channels, using our proposed method. The performance metric is the Average Normalized Mean Square Error (ANMSE) of the estimated channel, where the following Normalized Mean Square Error, NMSE $= \left[\dfrac{\int_{\mathcal{K}}(\Upsilon_{\text{dB}}(q) - \hat{\Upsilon}_{\text{dB},\mathcal{Q}}(q))^2\,dA}{\int_{\mathcal{K}}\Upsilon_{\text{dB}}^2(q)\,dA}\right]$, is averaged over several different randomly-selected sampling positions, for a given percentage of collected samples. Fig. 3 (top) shows the reconstruction performance for an outdoor channel across
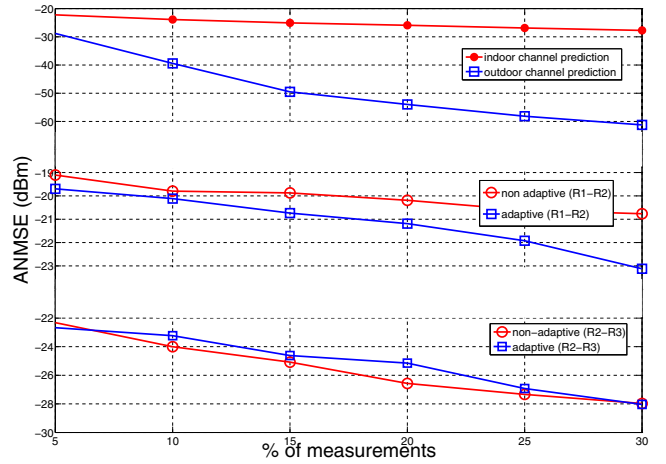

Fig. 3: Impact of different environments on channel prediction performance, using real channel measurements. (top) indoor and outdoor, (middle) main room (R1) and hallway (R2) of Fig. 10 and (bottom) hallways R2 and R3 of Fig. 10.

a street in downtown San Francisco [19] as well as for an indoor channel measurement along a route in the basement of the ECE building at UNM. The indoor experiment uses an 802.11g WLAN card while the outdoor measurement is based on measuring receptions from an AT&T cell tower [19]. For both cases, all the underlying parameters are estimated using the LS approach of this section. Consider the outdoor case, for instance. We have the measurements of the received signal power, every 2mm along a street of length 16m in San Francisco, mounting to 8000 samples. Fig. 3 (top) then shows the prediction performance where only a percentage of the total samples were available to a node. The available measurements are randomly chosen over the street. 5% measurements, for instance, means that a robot has collected 400 samples, randomly over that street, based on which it will predict the channel over the whole street. The prediction error variance is -29dB for the case of 5% measurements. It can be seen that both channels can be reconstructed with a good quality. The outdoor channel, however, can be reconstructed with a considerably better quality. This is expected as the indoor channel suffers from a more severe multipath fading, which makes it less spatially predictable.

A. Space-varying Underlying Parameters and Adaptive Channel Prediction

So far, we considered channel prediction over a small enough space such that the underlying channel parameters can be considered constant over the workspace. However, if the available channel samples belong to a large enough space (such as the entire floor), the underlying parameters can be space-varying. In this part, we show how the previous framework can be extended to an adaptive approach, in order to address the case where the operation, and the corresponding available channel measurements, are over a large space. Basically, a robot can use its localization and mapping information (which it will have for navigation and collision avoidance) to detect when something changes in the structure of its environment. For instance, it can detect when

it moves out of a room to a hallway or when it reaches an intersection. Thus, we assume that the underlying parameters can possibly change when some environmental factors change. From analyzing several real measurements, this is a reasonable assumption. While there could possibly be cases that are not captured by this assumption, i.e. having a drastic change in an underlying parameter without any environmental change, such cases are rare and the robot can not know about it to adapt its strategy anyways.

Let a region denote a place of operation where there is no environmental changes and the underlying parameters can be considered constant (such as a room or a hallway with no intersection that leads to the transmitter). In order to allow the node to give less weight to the available measurements that are collected in different regions and/or are far from the position where the channel needs to be estimated, we introduce a forgetting factor and a distance-dependent weight. This allows the node to adapt the impact of a sample measurement on its prediction framework. The forgetting factor is used to let the node give less impact to a measurement if it belongs to a different region, as compared to the place where the robot needs to predict the channel. On the other hand, the distance-dependent weight allows the robot to give less weight to the farther measurements. Consider the case where the workspace consists of p different regions, i.e. $\mathcal{K} = \bigcup_{i=1}^{p} \mathcal{R}_i$. Let τ_i represent the region, where the ith measurement belongs to, i.e. $q_i \in \mathcal{R}_{\tau_i}$. Define the forgetting matrix F, with the following characteristics: 1) F is symmetric, 2) F is stochastic and 3) $[F]_{i,j} = f_{i,j}$ is proportional to the similarity between regions i and j. The third property implies that, $\max_j f_{j,i} = f_{i,i}$ and $f_{i,j} \geq f_{i,k}$ iff regions i and j have more environmental features in common, as compared to regions i and k. Furthermore, let $\mathcal{G}$ denote the functional space of all non-increasing functions on $\mathbf{R}^+$. For $q \in \mathcal{R}_m \subset \mathcal{K}$, we define the corresponding weight matrix as: $[\Psi_{\mathcal{Q}}(q)]_{i,i} = f_{\tau_i,m} \times g_{\tau_i,m}(\|q - q_i\|)$ and $[\Psi_{\mathcal{Q}}(q)]_{i,j} = 0$ for $i \neq j$, where $g_{\tau_i,m} \in \mathcal{G}$. One candidate for g is an exponential function: $g_{\tau_i,m}(\|q-q_i\|) = e^{-\frac{\|q-q_i\|}{b_{\tau_i,m}}}$. $f_{\tau_i,m}$ and $b_{\tau_i,m}$ are design parameters, which the robot can choose. They impact how conservative the robot will be in taking the measurements of different regions into account. Let $\hat{\theta}_{\text{WLS}}(q) = \min_\theta \left\| \Psi_{\mathcal{Q}}^{\frac{1}{2}}(q)(Y_{\mathcal{Q}} - H_{\mathcal{Q}}\theta) \right\|^2$ denote the weighted LS estimation of the path loss parameters, for prediction at position $q \in \mathcal{R}_m$. We then have, $\hat{\theta}_{\text{WLS}}(q) = \left(H_{\mathcal{Q}}^T \Psi_{\mathcal{Q}}(q) H_{\mathcal{Q}} \right)^{-1} H_{\mathcal{Q}}^T \Psi_{\mathcal{Q}}(q) Y_{\mathcal{Q}}$. The channel and other underlying parameters can be similarly estimated.

Fig. 4 shows the performance of our adaptive approach when a robot moves along a street. The channel measurement is in reception from an AT&T cell tower, in a street in San Francisco [19], which experiences very different path loss exponents due to the presence of an intersection that leads to the transmitter. The robot samples the channel as it moves along the street and estimates the path loss slope, without any a priori information in this environment. The figure compares the performance of the non-adaptive case with that of the adaptive one and shows that we can benefit considerably from the adaptation. Next, Fig. 3 (middle) shows the prediction quality when a number of robots operate in our basement,
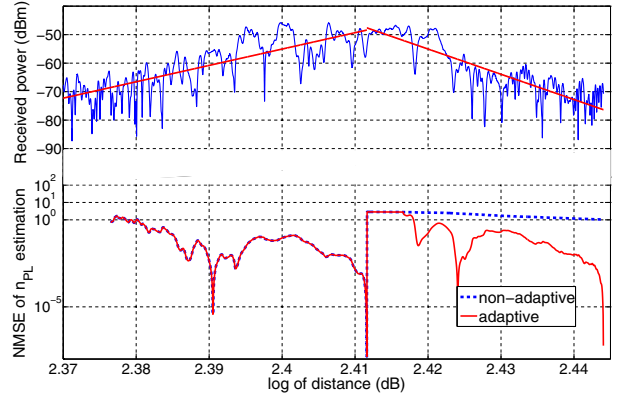

Fig. 4: Performance of our adaptive approach, in estimating the path loss slope, when a robot moves along a street in San Francisco and samples the channel along its trajectory [19]–(top) channel received power across the street along with its best slope fit and (bottom) prediction error variance of the robot, as it moves along the street and measures the channel.

over a large area and cooperate for channel prediction. The regions of operation are R1 and R2, as indicated in Fig. 10. Note that the performance is simulated, in this case, using real channel measurements in this environment. It can be seen that the adaptive approach can improve the performance as compared to the non-adaptive case. In the non-adaptive case, all the gathered and communicated measurements are utilized by each robot for channel prediction, without taking into account that these measurements may belong to different regions. It can be seen that we can benefit a couple of dBs, by using the adaptive approach. In other tests in different environments, we also observed that the adaptation may make a negligible difference if different regions are not that much different, in terms of their underlying parameters, as expected. Fig. 3 (bottom) shows an example of such a case for operation over a different area in our basement. It can be seen that the performance curves are very close.

In this paper, it is our goal to fundamentally understand the impact of different environments (in terms of their underlying parameters) on the proposed channel prediction framework. Consider the four marked regions of Fig. 10 for instance. We want to understand how the channel prediction quality changes (and justify the observed behaviors) when we move from one region to another. Therefore, in the rest of the paper, we consider the non-adaptive channel prediction framework, to predict the channel over a region where the underlying parameters can be considered constant. We then characterize the impact of different environments (in terms of the underlying channel parameters) on the prediction framework.

III. IMPACT OF CHANNEL PARAMETERS ON THE PREDICTION ERROR VARIANCE

In this section, we characterize the impact of the underlying channel parameters on the spatial predictability of a wireless channel. We assume that the underlying parameters are estimated perfectly in this section to avoid error propagation from parameter estimation to channel prediction. In the subsequent sections, we then extend our analysis to take the impact of the estimation error of key underlying parameters into account.

Let $\Upsilon_{\mathrm{dB}}(q) = 10\log_{10}\left(\Upsilon(q)\right)$ represent the received signal strength at position $q \in \mathcal{K}$ in dB. Based on the gathered measurements at $\mathcal{Q} \subset \mathcal{K}$, the goal is to estimate the channel at $q \in \mathcal{K}\backslash\mathcal{Q}$, using the channel predictor, $\tilde{\Upsilon}_{\mathrm{dB},\mathcal{Q}}(q)$ of Eq. 7, with the corresponding error covariance of $\sigma^2_{\mathrm{dB},\mathcal{Q}}(q)$. We next characterize the impact of different channel parameters on this prediction. We first introduce the following lemmas.

Lemma 2: Let $\Psi(t)$ be an invertible matrix for $t \in \mathbb{R}$. We have $\frac{d\Psi^{-n}}{dt} = -\Psi^{-n}\frac{d\Psi^n}{dt}\Psi^{-n}$, where n is a positive integer.

Proof: Taking the derivative from both sides of equation $\Psi^n(t)\Psi^{-n}(t) = I_{k\times k}$, with respect to t, proves the lemma. ∎

Lemma 3: Let J be an n-by-m matrix with the rank of m and Ψ be an n-by-n full rank matrix. If matrix Ψ is positive definite ($\Psi \succ 0$), then $J^T\Psi J$ is positive definite.

Proof: See [20] for a proof. ∎

Theorem 1: The estimation error variance, $\sigma^2_{\mathrm{dB},\mathcal{Q}}$, is an increasing function of α and σ^2 for $\alpha, \sigma^2 \in [0,\infty)$ and an invertible $R_{\mathrm{norm},\mathcal{Q}}$.

Proof: We first show that the estimation error variance is an increasing function of σ^2. Let $\phi_{\mathrm{norm},\mathcal{Q}}(q) = \frac{1}{\alpha}\phi_{\mathcal{Q}}(q)$ denote the normalized cross covariance between $\mathcal{Q}$ and q. We have $\sigma^2_{\mathrm{dB},\mathcal{Q}}(q) = \alpha + \sigma^2 - \alpha\phi^T_{\mathrm{norm},\mathcal{Q}}(q)\left(R_{\mathrm{norm},\mathcal{Q}} + \frac{\sigma^2}{\alpha}I_{k\times k}\right)^{-1}\phi_{\mathrm{norm},\mathcal{Q}}(q)$. For $\alpha = 0$, we have $\frac{d}{d\sigma^2}\sigma^2_{\mathrm{dB},\mathcal{Q}}(q)\big|_{\alpha=0} = 1 > 0$, $\forall\sigma^2 \in [0,\infty)$. For $\alpha \neq 0$, taking the derivative with respect to σ^2 (using Lemma 2) and then applying Lemma 3 result in: $\frac{d}{d\sigma^2}\sigma^2_{\mathrm{dB},\mathcal{Q}}(q) = 1 + \phi^T_{\mathrm{norm},\mathcal{Q}}(q)\left(R_{\mathrm{norm},\mathcal{Q}} + \frac{\sigma^2}{\alpha}I_{k\times k}\right)^{-2}\phi_{\mathrm{norm},\mathcal{Q}}(q) > 0$, $\forall\sigma^2 \in [0,\infty)$ and for an invertible $R_{\mathrm{norm},\mathcal{Q}}$, which completes the proof. We next prove that $\sigma^2_{\mathrm{dB},\mathcal{Q}}(q)$ is an increasing function of α. First assume that $\sigma^2 \neq 0$. Taking the derivative with respect to α results in: $\frac{d}{d\alpha}\sigma^2_{\mathrm{dB},\mathcal{Q}}(q) = 1 - \phi^T_{\mathrm{norm},\mathcal{Q}}(q)\Big[\left(R_{\mathrm{norm},\mathcal{Q}} + \frac{\sigma^2}{\alpha}I_{k\times k}\right)^{-1} + \frac{\sigma^2}{\alpha}\left(R_{\mathrm{norm},\mathcal{Q}} + \frac{\sigma^2}{\alpha}I_{k\times k}\right)^{-2}\Big]\phi_{\mathrm{norm},\mathcal{Q}}(q)$. Define $f(\alpha) \triangleq \frac{d}{d\alpha}\sigma^2_{\mathrm{dB},\mathcal{Q}}(q)$. f is of class C^∞ on $\mathbb{R}^+$ with the following properties: 1) $f(0) = 1$, 2) $f(\infty) = 1 - \phi^T_{\mathrm{norm},\mathcal{Q}}(q)R^{-1}_{\mathrm{norm},\mathcal{Q}}\phi_{\mathrm{norm},\mathcal{Q}}(q) > 0$ and 3) $\frac{d}{d\alpha}f(\alpha) < 0$.

First property can be easily confirmed. We next prove the second property. Let $R_{\mathrm{norm},\mathcal{Q}\bigcup\{q\}}$ represent the correlation matrix corresponding to $\mathcal{Q}\bigcup\{q\}$. We have $R_{\mathrm{norm},\mathcal{Q}\bigcup\{q\}} = \begin{bmatrix} R_{\mathrm{norm},\mathcal{Q}} & \phi_{\mathrm{norm},\mathcal{Q}}(q) \\ \phi^T_{\mathrm{norm},\mathcal{Q}}(q) & 1 \end{bmatrix}$, where $R_{\mathrm{norm},\mathcal{Q}\bigcup\{q\}}$ is assumed invertible. Thus, under the assumption that $R_{\mathrm{norm},\mathcal{Q}}$ is invertible, the second property can be easily confirmed, using the Schur complement of $R_{\mathrm{norm},\mathcal{Q}}$ block [21]. Next we prove the third property. We have $\frac{d}{d\alpha}f(\alpha) = -\phi^T_{\mathrm{norm},\mathcal{Q}}(q)\frac{d}{d\alpha}\Big[\left(R_{\mathrm{norm},\mathcal{Q}} + \frac{\sigma^2}{\alpha}I_{k\times k}\right)^{-1} + \frac{\sigma^2}{\alpha}\left(R_{\mathrm{norm},\mathcal{Q}} + \frac{\sigma^2}{\alpha}I_{k\times k}\right)^{-2}\Big]\phi_{\mathrm{norm},\mathcal{Q}}(q) = -\phi^T_{\mathrm{norm},\mathcal{Q}}(q)\frac{\sigma^2}{\alpha}\frac{d}{d\alpha}\Big[\left(R_{\mathrm{norm},\mathcal{Q}} + \frac{\sigma^2}{\alpha}I_{k\times k}\right)^{-2}\Big]\phi_{\mathrm{norm},\mathcal{Q}}(q) = -2\sigma^4\phi^T_{\mathrm{norm},\mathcal{Q}}(q)\left(\alpha R_{\mathrm{norm},\mathcal{Q}} + \sigma^2 I_{k\times k}\right)^{-3}\phi_{\mathrm{norm},\mathcal{Q}}(q)$. Since $\left(\alpha R_{\mathrm{norm},\mathcal{Q}} + \sigma^2 I_{k\times k}\right) \succ 0$, we can then easily show that $\frac{d}{d\alpha}f(\alpha) < 0$ using Lemma 3. By using these three properties, we have $f(\alpha) > 0$, which means that the estimation error variance is an increasing function of $\alpha \in [0,\infty)$. Furthermore, if $\sigma^2 = 0$ and $R_{\mathrm{norm},\mathcal{Q}}$ is invertible, then $\frac{d}{d\alpha}\sigma^2_{\mathrm{dB},\mathcal{Q}}(q)\big|_{\sigma^2=0} = f(\infty)$, which is positive as shown for property 2. Therefore, estimation variance is an increasing

function of α in this case too. ∎

We next characterize the impact of β on the prediction quality, using properties of the Euclidean Distance Matrix (EDM) [22]. Given the position set $\mathcal{Q} = \{q_1, q_2, \cdots, q_k\} \subset \mathcal{K}$, the EDM $\Pi = \big[\pi_{i,j}\big] \in \mathbb{R}^{k\times k}$ is defined entry-wise as $\big[\Pi\big]_{i,j} = \pi_{i,j} = \|q_i - q_j\|^2$ for $i, j = 1, 2, \cdots, k$. We have the following properties for the EDM:

1) $\sqrt{\pi_{i,j}} \geq 0$ for $i \neq j$ and $\sqrt{\pi_{i,j}} = 0$ for $i = j$.
2) $\sqrt{\pi_{i,j}} = \sqrt{\pi_{j,i}}$.
3) $\sqrt{\pi_{i,l}} + \sqrt{\pi_{l,j}} \geq \sqrt{\pi_{i,j}}$ for $i \neq j \neq l$.

Theorem 2: Matrix $\Pi = \big[\pi_{i,j}\big] \in \mathbb{R}^{k\times k}$ is EDM if and only if $-V_k^T\Pi V_k \succeq 0$, $\Pi^T = \Pi$ and $\pi_{i,i} = 0$ for $1 \leq i \leq k$, where V_k is the full-rank skinny Schoenberg auxiliary matrix: $V_k \triangleq \frac{1}{\sqrt{2}}\begin{bmatrix} -1_{k-1}^T \\ I_{(k-1)\times(k-1)} \end{bmatrix} \in \mathbb{R}^{k\times k-1}$.

Proof: Readers are referred to [22] for the details of the proof. ∎

Theorem 3: Let $T = \big[t_{i,j}\big] \in \mathbb{R}^{k\times k}$ represent the entry-wise square root of $\Pi = \big[\pi_{i,j}\big] \in \mathbb{R}^{k\times k}$ where $t_{i,j} = \pi_{i,j}^{\frac{1}{2}}$. If Π is EDM, then T is EDM. This case is of interest because it corresponds to the absolute distance matrix.

Proof: Readers are referred to [22]–[24] for the details of the proof. ∎

Lemma 4: The Hadamard product (Schur product) of two positive-definite matrices is positive-definite and the Hadamard product of two positive-semidefinite matrices is positive-semidefinite.

Proof: Readers are referred to Theorem 7.5.3 of [20] for more details. ∎

Theorem 4: The estimation error variance is a decreasing function of $\beta \in (0,\infty)$ for $\sigma^2 \neq 0$ and a non-increasing function of $\beta \in (0,\infty)$ for $\sigma^2 = 0$ and an invertible $R_{\mathrm{norm},\mathcal{Q}}$.

Proof: Case of $\alpha = 0$ is not of interest in this theorem since we are interested in the impact of shadowing. Therefore, in this proof we assume that $\alpha \neq 0$. Let $\delta_{\mathcal{Q}}(q) = \big[\|q_1 - q\|, \|q_2 - q\|, \cdots, \|q_k - q\|\big]^T$ represent the distance vector between the set $\mathcal{Q}$ and position $q \notin \mathcal{Q}$ and $\Delta_{\mathcal{Q}}(q) \triangleq \mathrm{diag}\big[\delta_{\mathcal{Q}}(q)\big]$. Let $\big[T_{\mathcal{Q}}\big]_{i,j} = \| q_i - q_j \|$, $\forall q_i, q_j \in \mathcal{Q}$, denote the absolute distance matrix corresponding to the set $\mathcal{Q}$. First assume that $\sigma^2 \neq 0$. We have Eq. 8, shown at the top of the next page, where ($\bullet$) denotes the Hadamard product. Moreover, it can be confirmed that $R_{\mathrm{norm},\mathcal{Q}}\Delta_{\mathcal{Q}}(q) = R_{\mathrm{norm},\mathcal{Q}} \bullet \left(1_k\delta_{\mathcal{Q}}^T(q)\right)$. Therefore, we have: $\frac{d}{d\beta}\sigma^2_{\mathrm{dB},\mathcal{Q}}(q) = -\frac{1}{\beta^2}\phi_{\mathcal{Q}}^T(q)R^{-1}_{\mathrm{tot},\mathcal{Q}}\Big[\alpha\Big(1_k\delta_{\mathcal{Q}}^T(q) + \delta_{\mathcal{Q}}(q)1_k^T - T_{\mathcal{Q}}\Big) \bullet R_{\mathrm{norm},\mathcal{Q}} + 2\sigma^2\Delta_{\mathcal{Q}}(q)\Big]R^{-1}_{\mathrm{tot},\mathcal{Q}}\phi_{\mathcal{Q}}(q)$. From Lemma 4, we know that the Hadamard product of two positive-semidefinite matrices is positive-semidefinite. Therefore, to prove that $\frac{d}{d\beta}\sigma^2_{\mathrm{dB},\mathcal{Q}}(q)\big|_{\sigma^2\neq 0} < 0$, it suffices to show that $1_k\delta_{\mathcal{Q}}^T(q) + \delta_{\mathcal{Q}}(q)1_k^T - T_{\mathcal{Q}}$ is positive-semidefinite $\big($we know that $\Delta_{\mathcal{Q}}(q) \succ 0\big)$. Let $T_{\{q\}\bigcup\mathcal{Q}} = \begin{bmatrix} 0 & \delta_{\mathcal{Q}}^T(q) \\ \delta_{\mathcal{Q}}(q) & T_{\mathcal{Q}} \end{bmatrix} \in \mathbb{R}^{(k+1)\times(k+1)}$ represent the distance matrix corresponding to $\{q\}\bigcup\mathcal{Q}$. Let e_i denote a unit vector in $\mathbb{R}^{k+1}$, where all the entries are zero except for the ith one. Therefore, the Schoenberg auxiliary matrix can be represented as $V_{k+1} = \frac{1}{\sqrt{2}}\big[e_2 - e_1, \cdots, e_{k+1} - e_1\big]$. We have:

$$\frac{d}{d\beta}\sigma^2_{\text{dB},\mathcal{Q}}(q) = -\frac{1}{\beta^2}\phi^T_{\mathcal{Q}}(q)\Big[\Delta_{\mathcal{Q}}(q)R^{-1}_{\text{tot},\mathcal{Q}} - \alpha R^{-1}_{\text{tot},\mathcal{Q}}\big(T_{\mathcal{Q}} \bullet R_{\text{norm},\mathcal{Q}}\big)R^{-1}_{\text{tot},\mathcal{Q}} + R^{-1}_{\text{tot},\mathcal{Q}}\Delta_{\mathcal{Q}}(q)\Big]\phi_{\mathcal{Q}}(q)$$

$$= -\frac{1}{\beta^2}\phi^T_{\mathcal{Q}}(q)R^{-1}_{\text{tot},\mathcal{Q}}\Big[R_{\text{tot},\mathcal{Q}}\Delta_{\mathcal{Q}}(q) - \alpha T_{\mathcal{Q}} \bullet R_{\text{norm},\mathcal{Q}} + \Delta_{\mathcal{Q}}(q)R_{\text{tot},\mathcal{Q}}\Big]R^{-1}_{\text{tot},\mathcal{Q}}\phi_{\mathcal{Q}}(q)$$

$$= -\frac{1}{\beta^2}\phi^T_{\mathcal{Q}}(q)R^{-1}_{\text{tot},\mathcal{Q}}\Big[\alpha\Big(R_{\text{norm},\mathcal{Q}}\Delta_{\mathcal{Q}}(q) + \Delta_{\mathcal{Q}}(q)R_{\text{norm},\mathcal{Q}} - T_{\mathcal{Q}} \bullet R_{\text{norm},\mathcal{Q}}\Big) + 2\sigma^2\Delta_{\mathcal{Q}}(q)\Big]R^{-1}_{\text{tot},\mathcal{Q}}\phi_{\mathcal{Q}}(q), \quad (8)$$

$$-\big[V^T_{k+1}T_{\{q\}\cup\mathcal{Q}}V_{k+1}\big]_{i,j}$$

$$= -\frac{1}{2}\big(e_{i+1} - e_1\big)^T T_{\{q\}\cup\mathcal{Q}}\big(e_{j+1} - e_1\big)$$

$$= -\frac{1}{2}\Big(e^T_{i+1}T_{\{q\}\cup\mathcal{Q}}e_{j+1} - e^T_1 T_{\{q\}\cup\mathcal{Q}}e_{j+1} - e^T_{i+1}T_{\{q\}\cup\mathcal{Q}}e_1\Big)$$

$$= \frac{1}{2}\Big(\|q_j - q\| + \|q_i - q\| - \|q_i - q_j\|\Big)$$

$$= \frac{1}{2}\Big[1_k\delta^T_{\mathcal{Q}}(q) + \delta_{\mathcal{Q}}(q)1^T_k - T_{\mathcal{Q}}\Big]_{i,j}.$$

Then, matrix $T_{\{q\}\cup\mathcal{Q}}$ is EDM using Theorem 3. Therefore, applying Theorem 2 for EDM $T_{\{q\}\cup\mathcal{Q}}$ results in: $1_k\delta^T_{\mathcal{Q}}(q) + \delta_{\mathcal{Q}}(q)1^T_k - T_{\mathcal{Q}} = -2V^T_{k+1}T_{\{q\}\cup\mathcal{Q}}V_{k+1} \succeq 0$, which completes the proof. Next consider the case where $\sigma^2 = 0$. A similar derivation will result in $\frac{d}{d\beta}\sigma^2_{\text{dB},\mathcal{Q}}(q)\Big|_{\sigma^2=0} \leq 0$, under the assumption that $R_{\text{norm},\mathcal{Q}}$ is invertible. Therefore, the estimation error variance is a non-increasing function of β in this case. ∎

Note that path loss parameters, K_{dB} and n_{PL}, do not affect the estimation error variance in this case. In Section VI, we show the impact of different environments (with different underlying parameters) on channel predictability, using several measurements in our building. We next characterize the impact of the underlying parameters on the estimation of path loss parameters.

IV. IMPACT OF CHANNEL PARAMETERS ON PATH LOSS ESTIMATION

In this section, we explore the effect of the underlying channel parameters on the estimation of path loss parameters. To provide a benchmark, we first consider the ML estimator of Eq. 3, where we assume that α, β and σ^2 are perfectly known. We then consider the Least Square estimator of Eq. 5 for a more realistic case, where α, β and σ^2 are not known at the time of estimating path loss parameters. Let $\hat{\theta}_{\text{ML}} = [\hat{K}_{\text{dB,ML}} \ \hat{n}_{\text{PL,ML}}]^T$ denote the ML estimation of path loss parameters as denoted by Eq. 3. We have the following error covariance matrix: $C_{\theta,\text{ML}} = \mathbb{E}\Big\{\big(\theta - \hat{\theta}_{\text{ML}}\big)\big(\theta - \hat{\theta}_{\text{ML}}\big)^T\Big\} = \big(H^T_{\mathcal{Q}}R^{-1}_{\text{tot},\mathcal{Q}}H_{\mathcal{Q}}\big)^{-1}$, where $\sigma^2_{\hat{K}_{\text{dB,ML}}} = [C_{\theta,\text{ML}}]_{1,1}$ and $\sigma^2_{\hat{n}_{\text{PL,ML}}} = [C_{\theta,\text{ML}}]_{2,2}$ denote the ML estimation error variance of $\hat{K}_{\text{dB,ML}}$ and $\hat{n}_{\text{PL,ML}}$ respectively. We have the following Theorem.

Theorem 5: Both $\sigma^2_{\hat{K}_{\text{dB,ML}}}$ and $\sigma^2_{\hat{n}_{\text{PL,ML}}}$ are increasing functions of α and σ^2 for $\alpha, \sigma^2 \in [0, \infty)$ and an invertible $R_{\text{norm},\mathcal{Q}}$.

Proof: We have $C_{\theta,\text{ML}} = \Big(H^T_{\mathcal{Q}}\big(\alpha R_{\text{norm},\mathcal{Q}} + \sigma^2 I_{k\times k}\big)^{-1}H_{\mathcal{Q}}\Big)^{-1}$, where $R_{\text{norm},\mathcal{Q}} = \frac{1}{\alpha}R_{\mathcal{Q}}$. Taking the derivative with respect to α results in: $\frac{dC_{\theta,\text{ML}}}{d\alpha} =$

$-C_{\theta,\text{ML}}\frac{d}{d\alpha}\Big(H^T_{\mathcal{Q}}\big(\alpha R_{\text{norm},\mathcal{Q}} + \sigma^2 I_{k\times k}\big)^{-1}H_{\mathcal{Q}}\Big)C_{\theta,\text{ML}} = C_{\theta,\text{ML}}H^T_{\mathcal{Q}}\big(\alpha R_{\text{norm},\mathcal{Q}} + \sigma^2 I_{k\times k}\big)^{-1}R_{\text{norm},\mathcal{Q}}\big(\alpha R_{\text{norm},\mathcal{Q}} + \sigma^2 I_{k\times k}\big)^{-1}H_{\mathcal{Q}}C_{\theta,\text{ML}}$. By using Lemma 3 and the assumption that $R_{\text{norm},\mathcal{Q}} \succ 0$, we can easily see that $\frac{dC_{\theta,\text{ML}}}{d\alpha} \succ 0$. Let $e_1 = [1 \ 0]^T$ and $e_2 = [0 \ 1]^T$ denote unit vectors in $\mathbb{R}^2$. We have: $\frac{d}{d\alpha}\sigma^2_{\hat{K}_{\text{dB,ML}}} = \frac{d}{d\alpha}\big(e^T_1 C_{\theta,\text{ML}}e_1\big) = e^T_1\frac{dC_{\theta,\text{ML}}}{d\alpha}e_1 > 0$ and $\frac{d}{d\alpha}\sigma^2_{\hat{n}_{\text{PL,ML}}} = \frac{d}{d\alpha}\big(e^T_2 C_{\theta,\text{ML}}e_2\big) = e^T_2\frac{dC_{\theta,\text{ML}}}{d\alpha}e_2 > 0$. To show that the estimation error of path loss parameters is an increasing function of σ^2, it suffices to show that $\frac{dC_{\theta,\text{ML}}}{d\sigma^2} \succ 0$. We have, $\frac{dC_{\theta,\text{ML}}}{d\sigma^2} = -C_{\theta,\text{ML}}H^T_{\mathcal{Q}}\frac{d}{d\sigma^2}\big(\alpha R_{\text{norm},\mathcal{Q}} + \sigma^2 I_{k\times k}\big)^{-1}H_{\mathcal{Q}}C_{\theta,\text{ML}} = C_{\theta,\text{ML}}H^T_{\mathcal{Q}}\big(\alpha R_{\text{norm},\mathcal{Q}} + \sigma^2 I_{k\times k}\big)^{-2}H_{\mathcal{Q}}C_{\theta,\text{ML}} \succ 0$, for $\alpha, \sigma^2 \in [0, \infty)$ and an invertible $R_{\text{norm},\mathcal{Q}}$. ∎

In general, the estimation error variance of path loss parameters does not have monotonic behavior as a function of β. To get a better understanding of the impact of correlation distance on the estimation of path loss parameters, we consider two extreme cases of $\beta = 0$ and $\beta = \infty$. More specifically, we characterize the optimum positions of the measurement points at both extremes and find the minimum achievable estimation error variance.

A. Case of $\beta = 0$: In this case, $R_{\text{tot},\mathcal{Q}}\big(\beta = 0\big) = (\alpha + \sigma^2)I_{k\times k}$ and the error covariance matrix of path loss parameters can be characterized as:

$$\lim_{\beta\to 0}C_{\theta,\text{ML}} = (\alpha + \sigma^2)(H^T_{\mathcal{Q}}H_{\mathcal{Q}})^{-1}$$

$$= (\alpha + \sigma^2)\begin{bmatrix} k & -1^T_k D_{\mathcal{Q}} \\ -1^T_k D_{\mathcal{Q}} & D^T_{\mathcal{Q}}D_{\mathcal{Q}} \end{bmatrix}^{-1}$$

$$= \frac{\alpha + \sigma^2}{D^T_{\mathcal{Q}}A_k D_{\mathcal{Q}}}\begin{bmatrix} D^T_{\mathcal{Q}}D_{\mathcal{Q}} & 1^T_k D_{\mathcal{Q}} \\ 1^T_k D_{\mathcal{Q}} & k \end{bmatrix}, \quad (9)$$

where $A_k = kI_{k\times k} - 1_k 1^T_k$. As can be seen, the estimation error variances of both K_{dB} and n_{PL} are functions of sampling positions ($\mathcal{Q}$).

Lemma 5: Matrix $A_k = kI_{k\times k} - 1_k 1^T_k$ has 0 and k as eigenvalues with the multiplicity of 1 and $k - 1$ respectively. Let $v_1 \in \text{span}\{1_k\}$ and $v_2 \in 1^{\perp}_k$, where $1^{\perp}_k = \big\{v\big|v^T 1_k = 0\big\}$. We have $A_k v_1 = 0$ and $A_k v_2 = kv_2$.

Proof: The proof is straightforward and is omitted. ∎

Theorem 6: Let $D^{1_k}_{\mathcal{Q}}$ and $D^{1^{\perp}_k}_{\mathcal{Q}}$ denote the projection of $D_{\mathcal{Q}}$ to $\text{span}\{1_k\}$ and $1^{\perp}_k$ subspaces respectively. The optimum positioning, which minimizes both $\sigma^2_{\hat{K}_{\text{dB,ML}}}$ and $\sigma^2_{\hat{n}_{\text{PL,ML}}}$ for the case of $\beta = 0$, is

$$\mathcal{Q}^{\text{opt}}_{\text{PL},\beta=0} = \arg\max_{\mathcal{Q}}||D_{\mathcal{Q}}||^2_2, \text{ s.t. } \mathcal{Q} \subset \mathcal{K} \text{ and } D^{1_k}_{\mathcal{Q}} = 0. \quad (10)$$

Proof: We have the following optimum positioning in order to minimize the estimation error variance of

K_{dB}, using Rayleigh-Ritz theorem [20]: $\mathcal{Q}^{\text{opt}}_{\sigma^2_{\hat{K}_{\text{dB,ML}},\beta=0}} =$ $\arg\min_{\text{s.t. } \mathcal{Q}\subset\mathcal{K}} \sigma^2_{\hat{K}_{\text{dB,ML}},\beta=0} = \arg\max_{\text{s.t. } \mathcal{Q}\subset\mathcal{K}} \frac{D_{\mathcal{Q}}^T A_k D_{\mathcal{Q}}}{D_{\mathcal{Q}}^T D_{\mathcal{Q}}} = \left\{ Q | Q \subset \mathcal{K} \text{ and } D_{\mathcal{Q}}^{1_k} = 0 \right\}$. This optimization problem can have multiple solutions, depending on the structure of the space, all of which achieve the minimum error variance of $\frac{\alpha+\sigma^2}{k}$. Similarly, we have the following to minimize the estimation error variance of n_{PL}: $\mathcal{Q}^{\text{opt}}_{\sigma^2_{\hat{n}_{\text{PL,ML}},\beta=0}} =$ $\arg\min_{\text{s.t. } \mathcal{Q}\subset\mathcal{K}} \sigma^2_{\hat{n}_{\text{PL,ML}},\beta=0} = \arg\max_{\text{s.t. } \mathcal{Q}\subset\mathcal{K}} D_{\mathcal{Q}}^T A_k D_{\mathcal{Q}} = \arg\max_{\text{s.t. } \mathcal{Q}\subset\mathcal{K} \text{ and } D_{\mathcal{Q}}^{1_k}=0} \|D_{\mathcal{Q}}\|_2^2$. Therefore, Eq. 10 represents the optimum positioning which satisfies both objectives. ∎

Next, we provide an intuitive interpretation. Similar to Eq. 2, the measurement vector can be represented by $Y_{\mathcal{Q}} = (K_{\text{dB}} \times k)u_1 + (-n_{\text{PL}}\|D_{\mathcal{Q}}\|_2)u_2 + \Xi_{\mathcal{Q}}$, where $u_1 = \frac{1_k}{\sqrt{k}}$ and $u_2 = \frac{D_{\mathcal{Q}}}{\|D_{\mathcal{Q}}\|_2}$ are normalized vectors. Then, the problem becomes similar to the decoding problem in CDMA (Code Division Multiple Access) systems. Thus, we have $D_{\mathcal{Q}} \in 1_k^\perp$. Moreover, maximizing k and $\|D_{\mathcal{Q}}\|_2$, which can be interpreted as maximizing the SNR of each term, results in a better estimation of K_{dB} and n_{PL} respectively.

B. Case of $\beta = \infty$: Next we characterize the impact of correlation on the estimation quality of path loss parameters, when β goes to ∞. To simplify the derivations, we define two variables: $\rho = \frac{\alpha}{\sigma^2}$ for $\sigma^2 \neq 0$, which denotes the ratio of the power of shadowing to multipath power (in dB) and $\chi = \alpha + \sigma^2$, which represents the sum of the two powers. The following can be easily confirmed for $\sigma^2 \neq 0$: 1) $\lim_{\beta\to\infty} R_{\text{tot},\mathcal{Q}} = (\rho 1_k 1_k^T + I_{k\times k})\frac{\chi}{1+\rho}$, 2) $\lim_{\beta\to\infty} R_{\text{tot},\mathcal{Q}}^{-1} = \left(I_{k\times k} - \frac{\rho}{1+\rho k}1_k 1_k^T\right)\frac{1+\rho}{\chi}$ (using Matrix Inversion Lemma), 3) $\lim_{\beta\to\infty} 1_k^T R_{\text{tot},\mathcal{Q}}^{-1} 1_k = k\frac{1+\rho}{1+\rho k}\frac{1}{\chi}$, 4) $\lim_{\beta\to\infty} 1_k^T R_{\text{tot},\mathcal{Q}}^{-1} D_{\mathcal{Q}} = \left(1_k^T D_{\mathcal{Q}}\right)\frac{1+\rho}{1+\rho k}\frac{1}{\chi}$, 5) $\lim_{\beta\to\infty} D_{\mathcal{Q}}^T R_{\text{tot},\mathcal{Q}}^{-1} D_{\mathcal{Q}} = \left(D_{\mathcal{Q}}^T D_{\mathcal{Q}} - \frac{\rho}{1+\rho k}\left(1_k^T D_{\mathcal{Q}}\right)^2\right)\frac{1+\rho}{\chi} = D_{\mathcal{Q}}^T\left(\rho A_k + I_{k\times k}\right)D_{\mathcal{Q}}\frac{1+\rho}{1+\rho k}\frac{1}{\chi}$. Using the above equations, we have Eq. 11, shown at the top of next page.

Remark 1: It can be seen from Eq. 11 that Theorem 6 also characterizes the optimum positioning for this case. Moreover, if $\mathcal{Q}^{\text{opt}}_{\text{PL}}$ denotes the solution of Eq. 10, then we have,

$$\sigma^2_{\hat{K}_{\text{dB,ML}}}\Big|_{\beta=0} = \frac{\alpha+\sigma^2}{k}, \ \sigma^2_{\hat{K}_{\text{dB,ML}}}\Big|_{\beta=\infty} = \alpha + \frac{\sigma^2}{k},$$
$$\sigma^2_{\hat{n}_{\text{PL,ML}}}\Big|_{\beta=0} = \frac{\alpha+\sigma^2}{\|D_{\mathcal{Q}^{\text{opt}}_{\text{PL}}}\|^2} \text{ and } \sigma^2_{\hat{n}_{\text{PL,ML}}}\Big|_{\beta=\infty} = \frac{\sigma^2}{\|D_{\mathcal{Q}^{\text{opt}}_{\text{PL}}}\|^2}. \quad (12)$$

As can be seen, the fully correlated case provides a smaller estimation error variance for n_{PL} and larger for K_{dB}. In [5], we showed that the slope of path loss, n_{PL}, has the most impact on the overall channel estimation error variance. Thus, case of $\beta = \infty$ would be more desirable than $\beta = 0$.

Remark 2: Consider the case where multipath effect is negligible, i.e., $\sigma^2 = 0$. We have

$$\lim_{\beta\to\infty} C_{\theta,\text{ML},\sigma^2=0} = \begin{bmatrix} \alpha & 0 \\ 0 & 0 \end{bmatrix}. \quad (13)$$

For this case, the measurement vector becomes $Y_{\mathcal{Q}} = H_{\mathcal{Q}}\theta + \varrho 1_k$, where $\varrho \sim \mathcal{N}(0,\alpha)$ with $\mathcal{N}$ denoting a Gaussian

distribution. Thus, for $k \geq 2$, the slope of path loss, $-n_{\text{PL}}$, can be perfectly estimated. However, the uncertainty of ϱ results in a bias in the estimation of K_{dB}, as can be seen from Eq. 13. It can also be seen that the estimation error covariance is not a function of the sampling positions anymore.

We next characterize the LS estimation of path loss parameters. Let $\hat{\theta}_{\text{LS}}$ denote the LS estimation of path loss parameters as denoted by Eq. 5. We have the following error covariance matrix: $C_{\theta,\text{LS}} = \left(H_{\mathcal{Q}}^T H_{\mathcal{Q}}\right)^{-1} H_{\mathcal{Q}}^T R_{\text{tot},\mathcal{Q}} H_{\mathcal{Q}}\left(H_{\mathcal{Q}}^T H_{\mathcal{Q}}\right)^{-1}$. The following Theorem characterizes some properties of this estimator.

Theorem 7: Let $\hat{\theta}_{\text{LS}}$ and $C_{\theta,\text{LS}}$ represent the Least Square estimator of path loss parameters and the corresponding estimation error covariance matrix respectively. Let $\sigma^2_{\hat{K}_{\text{dB,LS}}}$ and $\sigma^2_{\hat{n}_{\text{PL,LS}}}$ denote the LS error variances of $\hat{K}_{\text{dB,LS}}$ and $\hat{n}_{\text{PL,LS}}$ respectively. We have the following properties:

1) $C_{\theta,\text{LS}} \succeq C_{\theta,\text{ML}}$.
2) $\sigma^2_{\hat{K}_{\text{dB,LS}}}$ and $\sigma^2_{\hat{n}_{\text{PL,LS}}}$ are increasing functions of σ^2 for α, $\sigma^2 \in [0,\infty)$. Moreover, $\sigma^2_{\hat{K}_{\text{dB,LS}}}$ and $\sigma^2_{\hat{n}_{\text{PL,LS}}}$ are increasing functions of α for α, $\sigma^2 \in [0,\infty)$ and an invertible $R_{\text{norm},\mathcal{Q}}$.
3) Both ML and LS estimators provide the same estimation error covariance matrices if $\beta = 0$ or ∞.

Proof: The first property says that the ML estimator outperforms the LS one, as expected. We skip the mathematical proof due to space limitations. The second property can be easily confirmed by taking the derivatives with respect to σ^2 and α. We next prove the third property. For $\beta = 0$, we have $R_{\text{tot},\mathcal{Q}} = (\alpha + \sigma^2)I_{k\times k}$, resulting in $\lim_{\beta\to 0} C_{\theta,\text{LS}} = \lim_{\beta\to 0} C_{\theta,\text{ML}} = \frac{\alpha+\sigma^2}{D_{\mathcal{Q}}^T A_k D_{\mathcal{Q}}}\begin{bmatrix} D_{\mathcal{Q}}^T D_{\mathcal{Q}} & 1_k^T D_{\mathcal{Q}} \\ 1_k^T D_{\mathcal{Q}} & k \end{bmatrix}$. For $\beta = \infty$, we have $R_{\text{tot},\mathcal{Q}} = \alpha 1_k 1_k^T + \sigma^2 I_{k\times k}$, $C_{\theta,\text{LS}} = \alpha\left(H_{\mathcal{Q}}^T H_{\mathcal{Q}}\right)^{-1}\left(H_{\mathcal{Q}}^T 1_k 1_k^T H_{\mathcal{Q}}\right)\left(H_{\mathcal{Q}}^T H_{\mathcal{Q}}\right)^{-1} + \sigma^2\left(H_{\mathcal{Q}}^T H_{\mathcal{Q}}\right)^{-1}$ and Eq. 14, shown at the next page. Therefore, we have

$$\lim_{\beta\to\infty} C_{\theta,\text{LS}} = \alpha\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + \frac{\sigma^2}{D_{\mathcal{Q}}^T A_k D_{\mathcal{Q}}}\begin{bmatrix} D_{\mathcal{Q}}^T D_{\mathcal{Q}} & 1_k^T D_{\mathcal{Q}} \\ 1_k^T D_{\mathcal{Q}} & k \end{bmatrix}$$
$$= \begin{bmatrix} \alpha + \frac{D_{\mathcal{Q}}^T D_{\mathcal{Q}}}{D_{\mathcal{Q}}^T A_k D_{\mathcal{Q}}}\sigma^2 & \frac{1_k^T D_{\mathcal{Q}}}{D_{\mathcal{Q}}^T A_k D_{\mathcal{Q}}}\sigma^2 \\ \frac{1_k^T D_{\mathcal{Q}}}{D_{\mathcal{Q}}^T A_k D_{\mathcal{Q}}}\sigma^2 & \frac{k}{D_{\mathcal{Q}}^T A_k D_{\mathcal{Q}}}\sigma^2 \end{bmatrix}.$$

By comparing this equation to Eq. 11, the third property can be verified. ∎

Remark 3: Theorem 7 (3) shows that the optimum positioning of Eq. 10 minimizes the estimation error variance of the LS case too.

We next verify the derived theorems, using a simulated channel. Fig. 5 shows a simulated channel, generated with our probabilistic channel simulator [25], with the following parameters: frequency of operation of 1GHz, $\theta = [-22 \ 3.0]^T$, $\sqrt{\alpha} = \sqrt{8}\text{dB}$ and $\beta = 1\text{m}$. As for multipath fading, this channel experiences a correlated Rician fading, with Jakes power spectrum [8], which results in the multipath fading getting uncorrelated after 0.12m. The pdf of a unit-average Rician distribution, with parameter K_{ric}, is given by [6]: $f_{\Upsilon_{\text{MP}}}(x) = (1+K_{\text{ric}})e^{-K_{\text{ric}}-(1+K_{\text{ric}})x}I_0\left(2\sqrt{xK_{\text{ric}}(K_{\text{ric}}+1)}\right)$, where $I_0(.)$ is the modified zeroth-order Bessel function. Note that $K_{\text{ric}} = 0$ results in an exponential distribution, which

$$\lim_{\beta \to \infty} C_{\theta,\mathrm{ML}} = \lim_{\beta \to \infty} \left(H_{\mathcal{Q}}^T R_{\mathrm{tot},\mathcal{Q}}^{-1} H_{\mathcal{Q}} \right)^{-1} = \lim_{\beta \to \infty} \begin{bmatrix} 1_k^T R_{\mathrm{tot},\mathcal{Q}}^{-1} 1_k & -1_k^T R_{\mathrm{tot},\mathcal{Q}}^{-1} D_{\mathcal{Q}} \\ -1_k^T R_{\mathrm{tot},\mathcal{Q}}^{-1} D_{\mathcal{Q}} & D_{\mathcal{Q}}^T R_{\mathrm{tot},\mathcal{Q}}^{-1} D_{\mathcal{Q}} \end{bmatrix}^{-1}$$

$$= \frac{1 + \rho k}{1 + \rho} \chi \times \frac{1}{(1 + \rho k) D_{\mathcal{Q}}^T A_k D_{\mathcal{Q}}} \times \begin{bmatrix} D_{\mathcal{Q}}^T \left(\rho A_k + I_{k \times k} \right) D_{\mathcal{Q}} & 1_k^T D_{\mathcal{Q}} \\ 1_k^T D_{\mathcal{Q}} & k \end{bmatrix} = \frac{\chi}{1 + \rho} \begin{bmatrix} \rho + \frac{D_{\mathcal{Q}}^T D_{\mathcal{Q}}}{D_{\mathcal{Q}}^T A_k D_{\mathcal{Q}}} & \frac{1_k^T D_{\mathcal{Q}}}{D_{\mathcal{Q}}^T A_k D_{\mathcal{Q}}} \\ \frac{1_k^T D_{\mathcal{Q}}}{D_{\mathcal{Q}}^T A_k D_{\mathcal{Q}}} & \frac{k}{D_{\mathcal{Q}}^T A_k D_{\mathcal{Q}}} \end{bmatrix}$$

$$= \begin{bmatrix} \alpha + \frac{D_{\mathcal{Q}}^T D_{\mathcal{Q}}}{D_{\mathcal{Q}}^T A_k D_{\mathcal{Q}}} \sigma^2 & \frac{1_k^T D_{\mathcal{Q}}}{D_{\mathcal{Q}}^T A_k D_{\mathcal{Q}}} \sigma^2 \\ \frac{1_k^T D_{\mathcal{Q}}}{D_{\mathcal{Q}}^T A_k D_{\mathcal{Q}}} \sigma^2 & \frac{k}{D_{\mathcal{Q}}^T A_k D_{\mathcal{Q}}} \sigma^2 \end{bmatrix}. \tag{11}$$

$$\left(H_{\mathcal{Q}}^T H_{\mathcal{Q}} \right)^{-1} \left(H_{\mathcal{Q}}^T 1_k 1_k^T H_{\mathcal{Q}} \right) \left(H_{\mathcal{Q}}^T H_{\mathcal{Q}} \right)^{-1}$$

$$= \frac{1}{\left(D_{\mathcal{Q}}^T A_k D_{\mathcal{Q}} \right)^2} \begin{bmatrix} D_{\mathcal{Q}}^T D_{\mathcal{Q}} & 1_k^T D_{\mathcal{Q}} \\ 1_k^T D_{\mathcal{Q}} & k \end{bmatrix} \begin{bmatrix} k^2 & -k(1_k^T D_{\mathcal{Q}}) \\ -k(1_k^T D_{\mathcal{Q}}) & (1_k^T D_{\mathcal{Q}})^2 \end{bmatrix} \begin{bmatrix} D_{\mathcal{Q}}^T D_{\mathcal{Q}} & 1_k^T D_{\mathcal{Q}} \\ 1_k^T D_{\mathcal{Q}} & k \end{bmatrix}$$

$$= \frac{1}{\left(D_{\mathcal{Q}}^T A_k D_{\mathcal{Q}} \right)^2} \begin{bmatrix} k^2 (D_{\mathcal{Q}}^T D_{\mathcal{Q}}) - k(1_k^T D_{\mathcal{Q}})^2 & -k(D_{\mathcal{Q}}^T D_{\mathcal{Q}})(1_k^T D_{\mathcal{Q}}) + (1_k^T D_{\mathcal{Q}})^3 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} D_{\mathcal{Q}}^T D_{\mathcal{Q}} & 1_k^T D_{\mathcal{Q}} \\ 1_k^T D_{\mathcal{Q}} & k \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \tag{14}$$
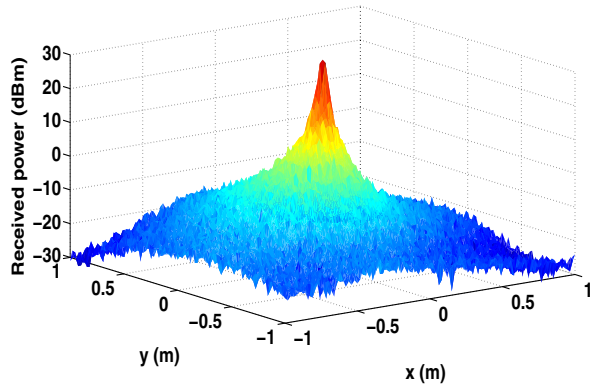

Fig. 5: A 2D simulated channel at 1GHz frequency with the following underlying parameters: $\theta = [-22 \quad 3.0]^T$, $\sqrt{\alpha} = \sqrt{8}$dB, $\beta = 1$m and $\sigma = \sqrt{2}$dB. The transmitter is located at $q_b = [0 \quad 0]^T$.
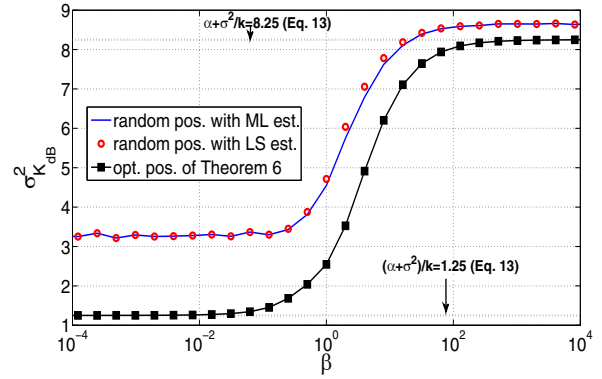

Fig. 6: Impact of β on the estimation of K_{dB} for both optimum positioning of Theorem 6 and random sampling.
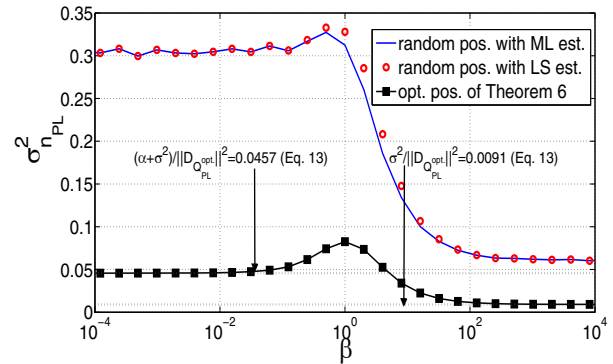

Fig. 7: Impact of β on the estimation of n_{PL} for both optimum positioning of Theorem 6 and random sampling.

experiences a considerable amount of channel variations, while $K_{\mathrm{ric}} = \infty$ results in no fading, i.e., we will have a channel with only path loss and shadowing. Multipath power (in dB), σ^2, is related to K_{ric} as follows: $\sigma^2 = \mathbb{E}\{\omega^2(q)\} = 100 \int_0^\infty \log_{10}^2(x) f_{\Upsilon_{\mathrm{MP}}}(x) dx - 100 \left(\int_0^\infty \log_{10}(x) f_{\Upsilon_{\mathrm{MP}}}(x) dx \right)^2$. For the simulated channel of Fig. 5, $\sigma = \sqrt{2}$ dB, which corresponds to $K_{\mathrm{ric}} = 19$.

Fig. 6 and Fig. 7 show the impact of the correlation distance, β, on the estimation variance of K_{dB} and n_{PL} respectively. In this example, the workspace is a ring with an inner radius of 0.3m and an outer radius of 3.3m, superimposed on the simulated channel of Fig. 5, such that the centers of the rings are positioned at the transmitter. We consider the case where $k = 8$ samples are taken from the workspace. Furthermore, we compare the performance for the case of random uniformly-distributed samples with the case where samples are optimally positioned based on Theorem 6. For this workspace, enforcing $D_{\mathcal{Q}}^{1_k} = 0$ results in $\max \|D_{\mathcal{Q}}\|_2^2 = 100k \log_{10}^2(\frac{10}{3})$, which can

be achieved if and only if half of the samples are distributed on the inner circle while the other half are on the outer one. Therefore, we assume that four samples are equally-spaced on the inner circle while the other four are equally-spaced on the
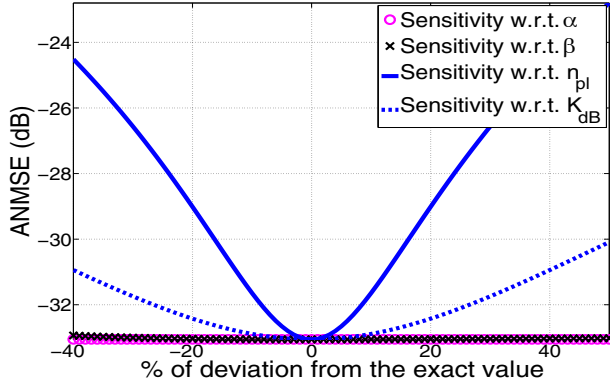
Fig. 8: Average Normalized Mean Square Error (ANMSE), spatially averaged over different channel realization and random sampling positions, as a function of the % of estimation error in $\hat{\alpha}$, $\hat{\beta}$ and $\hat{\theta}$ [5].

outer one.[3] The figures show that the optimum positioning of Theorem 6 can reduce the error, especially for the estimation of n_{PL} (which will have the most impact on the overall channel estimation error [5]). It can also be seen that as β approaches 0 or ∞, both estimators have the same quality as predicted by Theorem 7. Finally, the performances of the two estimators are not that different for other values of β in this example.

V. PERFORMANCE ANALYSIS USING UNBIASED ESTIMATION OF PATH LOSS PARAMETERS

In Section III, we considered the impact of the underlying parameters on the spatial predictability of a wireless channel, where we assumed that the underlying parameters are estimated perfectly. In this section, we extend that analysis to also consider the impact of estimation error in path loss parameters. In [5], authors analyzed the sensitivity of channel predication to the estimation of the underlying parameters. Fig. 8 shows the impact of parameter estimation error on the overall channel prediction performance. For each curve, only one parameter is perturbed while the rest are assumed perfectly estimated. It can be seen that the curves attain their minima when there is no parameter estimation error, as expected. We can furthermore observe that uncertainty in the estimation of different parameters impacts the performance differently. As can be seen, the prediction is more sensitive to path loss parameters (especially path loss exponent n_{PL}). In other words, the effect of an error in the estimation of the shadowing parameters is almost negligible, as compared to the error in path loss estimation. As such, in this section we extend the analysis of Section III to the case where errors in the estimation of path loss parameters are also considered. Consider the case where path loss parameters are estimated using an unbiased estimator.[4] We next characterize the error variance of channel prediction for this case, assuming that the error in the estimation of α, β and σ^2 is negligible. Since we are considering both the ML and LS estimators, we assume

[3]Note that the multipath components of different sampling points become uncorrelated with 100% probability for the case of optimum positioning and 95% probability for the case of random.

[4]The unbiased estimator can be either ML or LS.

that $R_{\text{tot},\mathcal{Q}}$ is invertible in the rest of the section (This is naturally implied if $\sigma^2 \neq 0$). Let $\hat{\theta}_{\text{unb}} = S Y_{\mathcal{Q}}$ denote an unbiased estimator of θ. We have the following for the error covariance matrix $C_{\theta,\,\text{unb}} = S R_{\text{tot},\mathcal{Q}} S^T$, with $S H_{\mathcal{Q}} = I_{k \times k}$. Let $\hat{\Upsilon}_{\text{dB},\mathcal{Q},\hat{\theta}_{\text{unb}}}(q) = \mathbb{E}\left\{ \Upsilon_{\text{dB}}(q) \mid Y_{\mathcal{Q}}, \theta = \hat{\theta}_{\text{unb}}, \alpha, \beta, \sigma^2 \right\}$ denote the estimation of channel at position q, when path loss parameters are estimated using the aforementioned unbiased estimator:

$$\hat{\Upsilon}_{\text{dB},\mathcal{Q},\hat{\theta}_{\text{unb}}}(q) = h^T(q)\hat{\theta}_{\text{unb}} + \phi_{\mathcal{Q}}^T(q) R_{\text{tot},\mathcal{Q}}^{-1}\left(Y_{\mathcal{Q}} - H_{\mathcal{Q}}\hat{\theta}_{\text{unb}}\right)$$
$$= \underbrace{\left[\left(h^T(q) - \phi_{\mathcal{Q}}^T(q) R_{\text{tot},\mathcal{Q}}^{-1} H_{\mathcal{Q}}\right) S + \phi_{\mathcal{Q}}^T(q) R_{\text{tot},\mathcal{Q}}^{-1}\right]}_{G_{\mathcal{Q}}} Y_{\mathcal{Q}}.$$

We have the following characterization for the error variance of channel estimation: $\sigma_{\text{dB},\mathcal{Q},\hat{\theta}_{\text{unb}}}^2(q) \triangleq$ $\mathbb{E}\left\{ \left(\Upsilon_{\text{dB}}(q) - \hat{\Upsilon}_{\text{dB},\mathcal{Q},\hat{\theta}_{\text{unb}}}(q)\right)^2 \mid \alpha,\beta,\sigma^2\right\} =$ $E\left\{ \left(h^T(q)\theta + \Xi_{\{q\}} - G_{\mathcal{Q}} Y_{\mathcal{Q}}\right)\left(h^T(q)\theta + \Xi_{\{q\}} - G_{\mathcal{Q}} Y_{\mathcal{Q}}\right)^T\right\} = \mathbb{E}\left\{ \left(\Xi_{\{q\}} - G_{\mathcal{Q}}\Xi_{\mathcal{Q}}\right)\left(\Xi_{\{q\}} - G_{\mathcal{Q}}\Xi_{\mathcal{Q}}\right)^T\right\} =$ $\alpha + \sigma^2 + G_{\mathcal{Q}} R_{\text{tot},\mathcal{Q}} G_{\mathcal{Q}}^T - G_{\mathcal{Q}}\phi_{\mathcal{Q}}(q) - \phi_{\mathcal{Q}}^T(q) G_{\mathcal{Q}}^T$, where $\Xi_{\{q\}} = \nu(q) + \omega(q)$ denotes the sum of shadowing and multipath power (in dB domain) at position q. It can be easily confirmed that $G_{\mathcal{Q}} R_{\text{tot},\mathcal{Q}} G_{\mathcal{Q}}^T = \left(h^T(q) - \phi_{\mathcal{Q}}^T(q) R_{\text{tot},\mathcal{Q}}^{-1} H_{\mathcal{Q}}\right) S R_{\text{tot},\mathcal{Q}} S^T \left(h^T(q) - \phi_{\mathcal{Q}}^T(q) R_{\text{tot},\mathcal{Q}}^{-1} H_{\mathcal{Q}}\right)^T - \phi_{\mathcal{Q}}^T R_{\text{tot},\mathcal{Q}}^{-1}\phi_{\mathcal{Q}} + G_{\mathcal{Q}}\phi_{\mathcal{Q}}(q) + \phi_{\mathcal{Q}}^T(q) G_{\mathcal{Q}}^T$. Therefore, we have Eq. 15, shown at the top of the next page. The initial uncertainty of channel estimation at q can be represented by $\alpha + \sigma^2 - \phi_{\mathcal{Q}}^T R_{\text{tot},\mathcal{Q}}^{-1}\phi_{\mathcal{Q}}$ if path loss is perfectly removed. Then, the second term of Eq. 15 is an increase in the error due to error propagation in the estimation of θ. As can be seen, $\sigma_{\text{dB},\mathcal{Q},\hat{\theta}_{\text{unb}}}^2(q)$ is not a function of θ since $\hat{\Upsilon}_{\text{dB},\mathcal{Q},\hat{\theta}_{\text{unb}}}(q)$ is an unbiased estimator of channel at position q.

In the previous sections, we showed that as α and σ^2 increase, the estimation of path loss parameters as well as channel prediction quality become worse. Thus, we expect to have the same trend, when considering both path loss estimation error and channel predictability. We next study the effect of correlation distance on the overall performance. Similar to Section IV, we consider two cases of $\beta = 0, \infty$ and characterize the overall channel estimation error variance. Moreover, we propose an optimum positioning scheme that minimizes the overall channel estimation error for these two cases.

Theorem 8: Let $\mathcal{Q}_{\sigma_{\text{dB},\mathcal{Q},\hat{\theta}_{\text{ML/LS}}}^2(q),\beta=0,\infty,\sigma^2\neq0}^{\text{opt}}$[5] denote the optimum positioning which minimizes the overall estimation error variance at q for both $\beta = 0$ and ∞ and $\sigma^2 \neq 0$, considering path loss estimation through either ML or LS. We have $\frac{1}{k}1_k^T D_{\mathcal{Q}_{\sigma_{\text{dB},\mathcal{Q},\hat{\theta}_{\text{ML/LS}}}^2(q),\beta=0,\infty,\sigma^2\neq0}^{\text{opt}}} = D_{\{q\}}$.

Proof: For $\beta = \infty$, we have $\lim_{\beta\to\infty}\phi_{\mathcal{Q}}(q) = \frac{\rho}{1+\rho}\chi 1_k$, $\forall q \in \mathcal{K}$, where $\rho = \frac{\alpha}{\sigma^2}$ and $\chi = \alpha + \sigma^2$. If $\sigma^2 \neq 0$, then using properties 3 and 4 of Section

[5]The notation $\hat{\theta}_{\text{ML/LS}}$ denotes that the estimation of path loss parameters can be either $\hat{\theta}_{\text{ML}|\alpha,\beta,\sigma^2}$ or $\hat{\theta}_{\text{LS}}$.

$$\sigma^2_{\text{dB},\mathcal{Q},\hat{\theta}_{\text{unb}}}(q) = \underbrace{\alpha + \sigma^2 - \phi^T_\mathcal{Q} R^{-1}_{\text{tot},\mathcal{Q}} \phi_\mathcal{Q}}_{\text{initial ch. est. error var. if path loss is perfectly known}} + \underbrace{\left(h^T(q) - \phi^T_\mathcal{Q}(q) R^{-1}_{\text{tot},\mathcal{Q}} H_\mathcal{Q}\right) C_{\theta,\text{ unb}} \left(h^T(q) - \phi^T_\mathcal{Q}(q) R^{-1}_{\text{tot},\mathcal{Q}} H_\mathcal{Q}\right)^T}_{\text{increase in error var. due to error propagation from est. of } \theta}.$$

$$(15)$$

$$\lim_{\beta \to \infty} \left(h^T(q) - \phi^T_\mathcal{Q}(q) R^{-1}_{\text{tot},\mathcal{Q}} H_\mathcal{Q}\right) C_{\theta,\text{ML/LS},\beta=\infty} \left(h^T(q) - \phi^T_\mathcal{Q}(q) R^{-1}_{\text{tot},\mathcal{Q}} H_\mathcal{Q}\right)^T$$

$$= \frac{D^T_\mathcal{Q} D_\mathcal{Q} + \rho(1^T_k D_\mathcal{Q})^2 - 2(1+\rho k)1^T_k D_\mathcal{Q} D_{\{q\}} + k(1+\rho k)D^2_{\{q\}}}{(1+\rho)(1+\rho k)D^T_\mathcal{Q} A_k D_\mathcal{Q}} \chi = \frac{(1^T_k D_\mathcal{Q} - k D_{\{q\}})^2 \rho + D^T_\mathcal{Q} D_\mathcal{Q} - 21^T_k D_\mathcal{Q} D_{\{q\}} + k D^2_{\{q\}}}{(1+\rho)(1+\rho k)D^T_\mathcal{Q} A_k D_\mathcal{Q}} \chi.$$

$$(16)$$

IV-B, we get $\lim_{\beta \to \infty} h^T(q) - \phi^T_\mathcal{Q}(q) R^{-1}_{\text{tot},\mathcal{Q}} H_\mathcal{Q} = \left[1 - \frac{\rho}{1+\rho}\chi \lim_{\beta \to \infty} 1^T_k R^{-1}_{\text{tot},\mathcal{Q}} 1_k \quad - D_{\{q\}} + \frac{\rho}{1+\rho}\chi \lim_{\beta \to \infty} 1^T_k R^{-1}_{\text{tot},\mathcal{Q}} D_\mathcal{Q}\right] = \left[\frac{1}{1+\rho k} \quad \frac{\rho}{1+\rho k} 1^T_k D_\mathcal{Q} - D_{\{q\}}\right]$ and $\lim_{\beta \to \infty} \phi^T_\mathcal{Q} R^{-1}_{\text{tot},\mathcal{Q}} \phi_\mathcal{Q} = \frac{\rho^2 k}{(1+\rho)(1+\rho k)}\chi$. Moreover, from Eq. 11 and Theorem 7, we have $C_{\theta,\text{ML/LS},\beta=\infty} \triangleq \lim_{\beta \to \infty} C_{\theta,\text{ML}} = \lim_{\beta \to \infty} C_{\theta,\text{LS}} = \frac{\chi}{1+\rho} \begin{bmatrix} \rho + \frac{D^T_\mathcal{Q} D_\mathcal{Q}}{D^T_\mathcal{Q} A_k D_\mathcal{Q}} & \frac{1^T_k D_\mathcal{Q}}{D^T_\mathcal{Q} A_k D_\mathcal{Q}} \\ \frac{1^T_k D_\mathcal{Q}}{D^T_\mathcal{Q} A_k D_\mathcal{Q}} & \frac{k}{D^T_\mathcal{Q} A_k D_\mathcal{Q}} \end{bmatrix}$, where $A_k = k I_{k \times k} - 1_k 1^T_k$. After some lines of derivations, Eq. 16, shown at the next page, can be derived. Thus, we have the following for Eq. 15:

$$\lim_{\beta \to \infty} \sigma^2_{\text{dB},\mathcal{Q},\hat{\theta}_{\text{ML/LS}}}(q) = \chi\left[1 - \frac{\rho^2 k}{(1+\rho)(1+\rho k)}\right.$$

$$\left. + \frac{(1^T_k D_\mathcal{Q} - k D_{\{q\}})^2 \rho + D^T_\mathcal{Q} D_\mathcal{Q} - 21^T_k D_\mathcal{Q} D_{\{q\}} + k D^2_{\{q\}}}{(1+\rho)(1+\rho k)D^T_\mathcal{Q} A_k D_\mathcal{Q}}\right]$$

$$= \frac{Q_1 \rho + Q_2}{(1+\rho)(1+\rho k)D^T_\mathcal{Q} A_k D_\mathcal{Q}} \chi, \qquad (17)$$

where $Q_1 = (k+1)D^T_\mathcal{Q} A_k D_\mathcal{Q} + (1^T_k D_\mathcal{Q} - k D_{\{q\}})^2$ and $Q_2 = D^T_\mathcal{Q} A_k D_\mathcal{Q} + D^T_\mathcal{Q} D_\mathcal{Q} - 21^T_k D_\mathcal{Q} D_{\{q\}} + k D^2_{\{q\}}$. It can be easily confirmed that $Q_1 = k Q_2$. Therefore, we have $\lim_{\beta \to \infty} \sigma^2_{\text{dB},\mathcal{Q},\hat{\theta}_{\text{ML/LS}}}(q) = \frac{Q_2}{(1+\rho)D^T_\mathcal{Q} A_k D_\mathcal{Q}}\chi = \frac{Q_2}{D^T_\mathcal{Q} A_k D_\mathcal{Q}}\sigma^2$. Moreover, we have:

$$\lim_{\beta \to 0} \sigma^2_{\text{dB},\mathcal{Q},\hat{\theta}_{\text{ML/LS}}}(q) = \left(1 + h^T(q)(H^T_\mathcal{Q} H_\mathcal{Q})^{-1} h(q)\right)\chi$$

$$= \frac{Q_2}{D^T_\mathcal{Q} A_k D_\mathcal{Q}}\chi. \qquad (18)$$

Thus, the optimum positioning which minimizes channel estimation error variance for both $\beta = 0$ and ∞ is $\mathcal{Q}^{\text{opt}}_{\sigma^2_{\text{dB},\mathcal{Q},\hat{\theta}_{\text{ML/LS}}}(q),\beta=0,\infty,\sigma^2 \neq 0} = \arg\min_\mathcal{Q} \frac{Q_2}{D^T_\mathcal{Q} A_k D_\mathcal{Q}}$. We have, $\frac{Q_2}{D^T_\mathcal{Q} A_k D_\mathcal{Q}} = 1 + \frac{D^T_\mathcal{Q} D_\mathcal{Q} - 21^T_k D_\mathcal{Q} D_{\{q\}} + k D^2_{\{q\}}}{D^T_\mathcal{Q} A_k D_\mathcal{Q}} = 1 + \frac{1}{k}\frac{D^T_\mathcal{Q} A_k D_\mathcal{Q} + (1^T_k D_\mathcal{Q} - k D_{\{q\}})^2}{D^T_\mathcal{Q} A_k D_\mathcal{Q}} = 1 + \frac{1}{k} + \frac{1}{k}\frac{(1^T_k D_\mathcal{Q} - k D_{\{q\}})^2}{D^T_\mathcal{Q} A_k D_\mathcal{Q}}$. It can be easily confirmed that A_k is positive-semidefinite. Thus, under full rank assumption of $H_\mathcal{Q}$ (as discussed in Section II), we have $D^T_\mathcal{Q} A_k D_\mathcal{Q} > 0$. Therefore, to minimize the estimation error variance for both $\beta = 0$ and ∞, we need $1^T_k D_\mathcal{Q} = k D_{\{q\}}$ for $D_\mathcal{Q} \in \mathbb{R}^k \backslash \text{span}\{1_k\}$. ∎
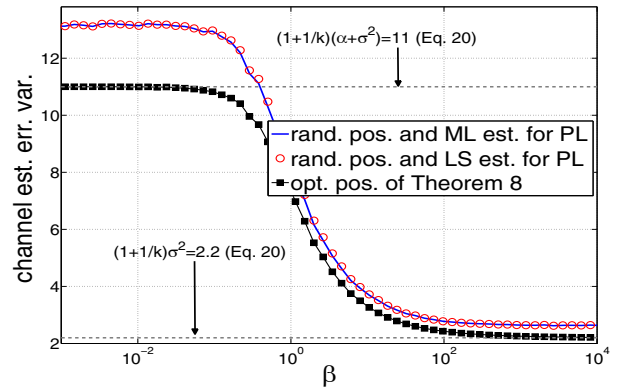

Fig. 9: Impact of β on the channel predictability, when considering path loss estimation error.

Case of $\sigma^2 = 0$ and $\beta = 0$ can be treated the same as Eq. 18 with $\chi = \alpha$, which results in the same optimum positioning scheme. However, if $\sigma^2 = 0$ and $\beta = \infty$, for $k \geq 2$, the channel variations can be perfectly estimated at each point. Theorem 8 shows that the optimum positioning results in $\|q - q_b\| = \left(\prod_{i=1}^k \|q - q_i\|\right)^{\frac{1}{k}}$. This suggests that the optimum measurement positions should be chosen such that the distance of q to the transmitter be the geometric average of the distances of the measurement points to the transmitter. Let $\mathcal{Q}' = \mathcal{Q}^{\text{opt}}_{\sigma^2_{\text{dB},\mathcal{Q},\hat{\theta}_{\text{ML/LS}}}(q),\beta=0,\infty,\sigma^2 \neq 0}$ denote the optimum positioning. We have,

$$\lim_{\beta \to 0} \sigma^2_{\text{dB},\mathcal{Q}',\hat{\theta}_{\text{ML/LS}}}(q) = \alpha + \sigma^2 + \frac{\alpha + \sigma^2}{k} \text{ and}$$

$$\lim_{\beta \to \infty} \sigma^2_{\text{dB},\mathcal{Q}',\hat{\theta}_{\text{ML/LS}}}(q) = \sigma^2 + \frac{\sigma^2}{k}. \qquad (19)$$

For $\beta = 0$, as k goes to ∞, the estimation error of path loss parameters goes to 0 and the estimation error variance becomes $\alpha + \sigma^2$. This value is an initial uncertainty assuming known path loss parameters. For the case of $\beta = \infty$, on the other hand, the estimation error variance becomes σ^2 as k goes to ∞. Fig. 9 shows the impact of correlation distance β on the estimation performance when path loss parameters are estimated using an ML/LS estimator. The impact of optimum positioning of Theorem 8 can also be seen from the figure.
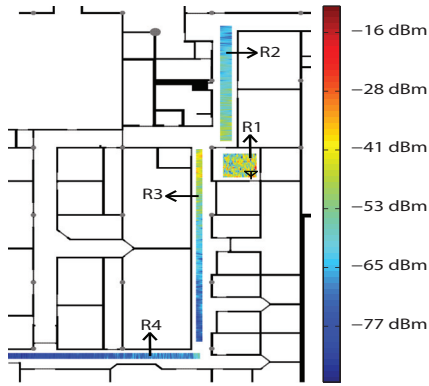
Fig. 10: Blueprint of the portion of the basement of the ECE bldg. at UNM where channel measurements are collected. A colormap of the measured received signal power is superimposed on the map. R1 denotes the main room, where the transmitter is located (as marked on the figure). R2, R3 and R4 correspond to different hallways at the basement. See the PDF file for a colored version of the map.
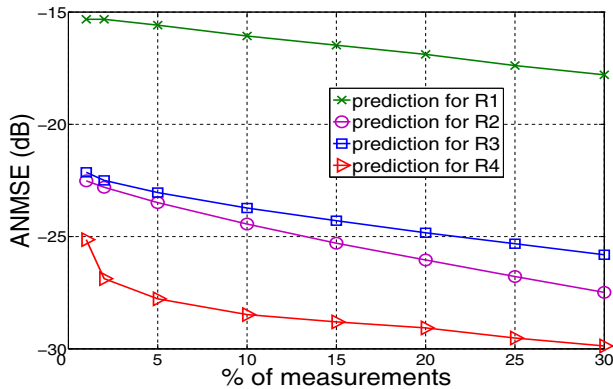

Fig. 11: Comparison of channel predictability for different regions of Fig. 10.

For this example, the workspace is a 2m × 2m square with $\sqrt{\alpha} = \sqrt{8}$dB and $\sigma = \sqrt{2}$dB, where $k = 10$ samples are taken from the workspace (either randomly or optimally). The y axis then represents the estimation performance after averaging over several runs of channel realization and sampling patterns. As can be seen, ML and LS estimators provide very similar performance in this case. Furthermore, optimizing the position of the samples, according to Theorem 8 can improve the performance considerably.

VI. NUMERICAL ANALYSIS ON REAL CHANNEL MEASUREMENTS

In this section, we show the impact of different environments, and their underlying parameters, on channel predictability, using real channel measurements. Fig. 10 shows the blueprint of a portion of the basement of the Electrical and Computer Engineering building at UNM. We used a Pioneer-AT robot to make several measurements along different routes in the basement, in order to map the received signal strength (each route is a straight line). As mentioned earlier, the robot is equipped with an 802.11g wireless card, with transmission at 2.4 GHz. It uses the MadWiFi device driver to measure the re-

ceived signal power [16]. The figure also shows a color-map of our measured received signal power. In order to see the impact of different underlying parameters on channel predictability, the area is divided into four regions of $R1, \cdots, R4$, as can be seen from the figure. Since we are dealing with real data, we can not check the accuracy of the estimation of the underlying parameters. As such, we use all the measurements in each region, to estimate the underlying parameters of that region, which are then used to understand channel predictability of each region. We use the LS estimator of Section II, in order to estimate channel parameters of each region. As can be seen, as the distance to the transmitter increases, n_{PL} (the slope of path loss) increases. This phenomena has previously been reported in the literature as well [26]. Another interesting phenomenon is the shadowing behavior. As can be seen, correlation distance (β) increases as we get farther from the transmitter and move to the hallways. This makes sense as shadowing is the result of the transmitted signal being possibly blocked by a number of obstacles before reaching the receiver. Finally, for region R1 (the main room), multipath fading is the dominant term, as can be seen. This is expected since that room is rich in scatterers and reflectors, with no major obstacle. Next, we consider channel predictability of different regions and relate the observed behaviors to the underlying parameters of Table I.

Fig. 11 compares channel prediction quality of different regions (measured by Average Normalized MSE as defined in Section II), given the parameters of Table I. As can be seen, region R4 has the best performance, as compared to the other regions. From Table I, region R4 has the smallest $\alpha + \sigma^2$ and good amount of correlation, which result in better predictability. On the other hand, region R1 experiences considerable multipath fading and negligible shadowing, which results in the worst predictability. Regions R2 and R3 have similar performances, since one has a higher shadowing correlation while the other experiences lower $\alpha + \sigma^2$. As was shown earlier, path loss parameters do not impact channel predictability. We next study the impact of each individual channel parameter on the estimation performance more closely.

Table I also shows channel parameters corresponding to three pairs of routes in the basement of ECE building at UNM (pairs A, B and C). Each pair is chosen such that only one parameter changes and the rest are almost the same. Fig. 12 (top) shows the impact of the shadowing power on the estimation performance. As can be seen, for A1/A2 pair, the correlation distance and multipath power are almost the same. However, A1 has a smaller shadowing power, which results in a better estimation performance. Fig. 12 (middle) and Fig. 12 (bottom) show the impact of correlation distance and multipath power on the estimation performance respectively. For each case, other channel parameters are almost the same. As can be seen, B2 with its higher correlation distance and C1 with its smaller multipath power provide better predictability. It should be noted that subtracting the estimated thermal noise power (which is provided by the wireless card) from each reception results in a maximum improvement of 0.1dB for all our results.

VII. CONCLUSION

In this paper, we developed a probabilistic channel prediction framework for predicting the spatial variations of a wire-

TABLE I: Channel Parameters for Different Regions and Routes

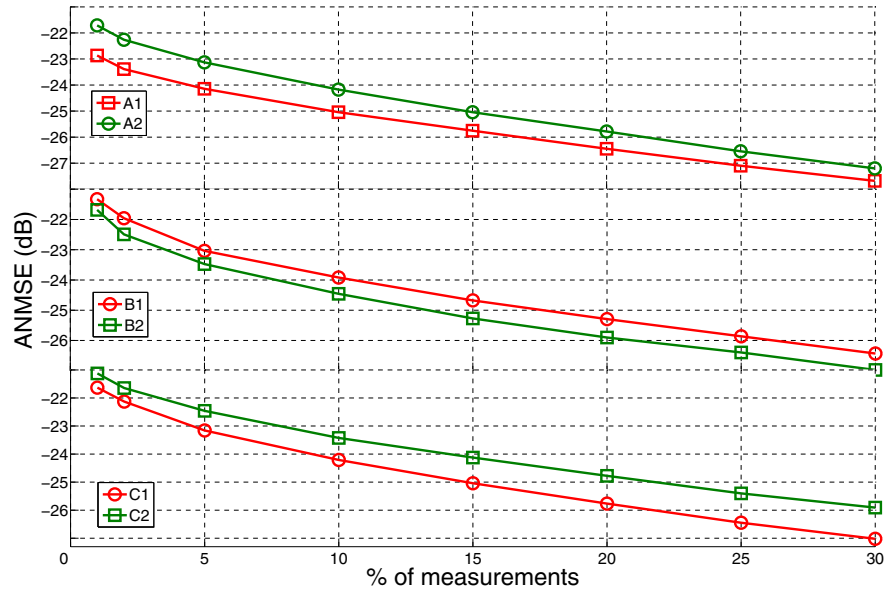| Region	K_{dB}	n_{PL}	α	β	σ^2
R1	-20.8870	1.2272	negligible	negligible	22.1238
R2	-21.4677	2.3878	10.7772	0.0979	2.8862
R3	-17.9694	2.9795	8.6385	0.3231	7.6628
R4	68.7836	9.9392	2.0157	1.4377	7.5687
A1	-	-	8.2164	0.0809	2.9721
A2	-	-	11.6332	0.0860	2.9313
B1	-	-	11.7535	0.2858	6.3979
B2	-	-	11.6029	0.5832	6.1956
C1	-	-	10.4193	0.2258	5.1696
C2	-	-	10.3451	0.2396	7.2873

Fig. 12: Impact of (top) shadowing power, (middle) correlation distance and (bottom) multipath power on channel prediction performance, using real channel measurements of Fig. 10.

less channel, based on a small number of measurements. We then proposed a mathematical foundation for understanding the spatial predictability of wireless channels. More specifically, we characterized the impact of different environments, in terms of their underlying parameters, on wireless channel predictability. We furthermore showed how sampling positions can be optimized to improve the prediction quality. Finally, we showed the performance of the proposed framework in predicting (and justifying the predictability of) the spatial variations of real channels, using several measurements in our building. Overall, the proposed framework can be utilized for communication-aware operation of robotic networks. There are several possible extensions of this work. For instance, we are currently working on characterizing the impact of the underlying parameters on the estimation of shadowing power and correlation distance, which will be the subject of another paper.

REFERENCES

[1] Y. Mostofi, M. Malmirchegini, and A. Ghaffarkhah, "Estimation of communication signal strength in robotic networks," in *Proc. 2010 IEEE International Conf. Robotics Automation*, pp. 1946–1951.

[2] W. Wang, V. Srinivasan, B. Wang, and K. Chua, "Coverage for target localization in wireless sensor networks," *IEEE Trans. Wireless Commun.*, vol. 7, pp. 667–676, Feb. 2008.

[3] J. Lu and T. Suda, "Differentiated surveillance for static and random mobile sensor networks," *IEEE Trans. Wireless Commun.*, vol. 7, pp. 4411–4423, Nov. 2008.

[4] A. Jadbabaie, J. Lin, and A. S. Morse, "Coordination of groups of mobile autonomous agents using nearest neighbor rules," *IEEE Trans. Automatic Control*, vol. 48, pp. 988–1001, 2003.

[5] A. Ghaffarkhah and Y. Mostofi, "Communication-aware motion planning in mobile networks," to appear in *IEEE Trans. Automatic Control*, 2010.

[6] A. Goldsmith, *Wireless Communications*. Cambridge University Press, 2005.

[7] T. S. Rappaport, *Wireless Communications: Principles and Practice*. Prentice-Hall, 2001.

[8] W. C. Jakes, *Microwave Mobile Communications*. Wiley-IEEE Press, 1994.

[9] H. Hashemi, "The indoor radio propagation channel," *Proc. IEEE*, vol. 81, pp. 943–968, July 1993.

[10] K. Remley, H. Anderson, and A. Weisshar, "Improving the accuracy of ray-tracing techniques for indoor propagation modeling," *IEEE Trans. Veh. Technol.*, vol. 49, pp. 2350–2358, Nov. 2000.

[11] K. W. Cheung, "Error in estimating local average power of multipath signals," *Electron. lett.*, vol. 27, 1991.

[12] C. Anderson, T. Rappaport, K. Bae, A. Verstak, N. Ramakrishnan, W. Tranter, C. Shaffer, and L. Watson, "In-building wideband multipath characteristics at 2.5 and 60 GHz," in *Proc. 2002 Veh. Technol. Conf.*, vol. 1.

[13] N. Cressie, "Kriging nonstationary data," *J. American Statistical Association*, vol. 81, no. 395, pp. 625–634, 1986.

[14] N. Cressie, *Statistics for Spatial Data*. Wiley-Interscience, 1993.

[15] J. G. Proakis, *Digital Communications*. McGraw-Hill, 2001.

[16] "Rssi in madwifi." Available: http://madwifi-project.org/wiki/UserDocs/RSSI.

[17] S. Cotton and W. Scanlon, "Higher order statistics for lognormal small-scale fading in mobile radio channels," *IEEE Antennas Wireless Propagation Lett.*, vol. 6, pp. 540–543, 2007.

[18] M. Gudmundson, "Correlation model for shadow fading in mobile radio systems," *Electron. Lett.*, vol. 27, Nov. 1991.

[19] W. M. Smith, "Urban propagation modeling for wireless systems," Ph.D. thesis, Stanford University, 2004.

[20] R. Horn and C. Johnson, *Matrix Analysis*. Cambridge University Press, 1999.

[21] T. Kailath, A. H. Sayed, and B. Hassibi, *Linear Estimation*. Prentice Hall, 2000.

[22] J. Dattorro, *Convex Optimization and Euclidean Distance Geometry*. Meboo, 2005.

[23] I. J. Schoenberg, "Metric spaces and positive definite functions," *Trans. American Mathematical Society*, 1938.

[24] E. Kreyszig, *Introductory Functional Analysis with Applications*. Wiley, 1989.

[25] Y. Mostofi, A. Gonzalez-Ruiz, A. Ghaffarkhah, and D. Li, "Characterization and modeling of wireless channels for networked robotic and control systems—a comprehensive overview," in *Proc. 2009 Int. Conf. Intelligent Robots Syst.*

[26] D. Akerberg, "Properties of a TDMA pico cellular office communication system," in *1988 IEEE Conf. Global Commun.*

[27] C. Oestges, N. Czink, B. Bandemer, P. Castiglione, F. Kaltenberger, and A. J. Paulraj, "Experimental characterization and modeling of outdoor-to-indoor and indoor-to-indoor distributed channels," *IEEE Trans. Veh. Technol.*, vol. 59, no. 5, pp. 2253–2265, June 2010.

Mehrzad Malmirchegini received the B.S. degree in electrical engineering from Iran University of Science and Technology, Tehran, Iran, in 2004 and the M.S. degree from Sharif University of Technology, Tehran, Iran in 2007. He joined the UNM Cooperative Network Lab in 2007, where he is currently a Ph.D. student. From 2004 to 2007, he was a research assistant in the Multi Media Lab at Sharif University. In the summer of 2010, he did summer internship at Motile Robotics Incorporated, Joppa, Maryland. He also did an internship with Samsung Telecommunications America, Dallas, Texas during the summer of 2011. Mehrzad's current research interests include wireless sensor networks, mobile communications and signal processing.

Yasamin Mostofi received the B.S. degree in electrical engineering from Sharif University of Technology, Tehran, Iran, in 1997, and the M.S. and Ph.D. degrees in the area of wireless communication systems from Stanford University, Stanford, CA, in 1999 and 2004, respectively. She is currently an assistant professor in the Department of Electrical and Computer Engineering at the University of New Mexico. Prior to that, she was a postdoctoral scholar in control and dynamical systems at the California Institute of Technology from 2004 to 2006. Dr. Mostofi is the recipient of the Presidential Early Career Award for Scientists and Engineers (PECASE) and the National Science Foundation (NSF) CAREER award. She also received Bellcore fellow-advisor award from Stanford Center for Telecommunications in 1999. She won the 2008-2009 ECE Distinguished Researcher Award at the University of New Mexico. Her current research lies at the intersection of the two areas of communications and control in mobile sensor networks. Current research projects include communication-aware navigation and decision making in robotic networks, compressive sensing and control, obstacle mapping, robotic routers, and cooperative information processing. She serves on the Control Systems Society conference editorial board since 2008. She is a member of the IEEE.