Estimation of Communication Signal Strength in Robotic Networks

Yasamin Mostofi, Mehrzad Malmirchegini and Alireza Ghaffarkhah

Abstract—In this paper we consider estimating the spatial variations of a wireless channel based on a small number of measurements in a robotic network. We use a multi-scale probabilistic model in order to characterize the channel and develop an estimator based on this model. We show that our model-based approach can estimate the channel well for several scenarios, with only a small number of gathered measurements. We furthermore consider a sparsity-based channel estimation approach, in which we utilize the compressibility of the channel in the frequency domain. Our results show that this approach can also be effective in several scenarios. We then discuss the underlying tradeoffs between the two approaches. For the model-based approach, we show the impact of the error in the underlying model as well as the error in the estimation of the parameters of the model on the overall performance. For the sparsity-based approach, we show the impact of channel compressibility on the performance. Overall, the proposed framework can be utilized for communication-aware motion planning in robotic networks, where a prediction of the link qualities is needed.

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. 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. In order to realize this vision, however, an integrative approach to communication and control issues is essential. In the robotics and control community, considerable progress has been made in the area of networked robotic and control systems. However, ideal or over-simplified models have typically been used to model communication links. A mobile cooperative network needs to maintain its connectivity in order to accomplish its task. In order to achieve this, each robot should consider the impact of motion decisions on its link qualities, when planning its trajectory. This requires each robot to assess the quality of the communication links in the locations that it has not yet visited. As a result, proper prediction of the communication signal strength in a given area, based on only a few measurements, becomes considerably important. As the robots move around, they can learn the signal strength at positions along their motion trajectories. However, there is simply not enough time to measure the channel at every location directly. Therefore, the channel should be reconstructed based on a considerably incomplete data set. Mapping the spatial variations of a communication channel, based on a small number of measurements, is an emerging area of research. In [1], we proposed utilizing the sparsity of the communication channel in the frequency domain in order to map the channel with a small number of measurements. In [2], we provided a comprehensive overview of channel characterization and modeling for networked robotic applications, by tapping into the relevant knowledge available in the wireless communication literature. In particular, we showed how certain probabilistic models can characterize the spatial variations of a wireless channel considerably well. In this paper, we show how such models can be used for channel prediction. We furthermore compare the model-based channel estimation approach of this paper with the sparsity-based one of [1]. We discuss the underlying tradeoffs between the two methods and the conditions under which one may perform better than the other.

The paper is organized as follows. In Section II we briefly summarize probabilistic modeling of a wireless channel. In Section III, we show how such models can be used for channel prediction. Section IV briefly summarizes the sparsity-based channel prediction approach, based on using compressive sensing. In Section V, we show the performance of the model-based approach in estimating real channels. We furthermore compare its performance with that of the sparsity-based one and shed light on the underlying tradeoffs between the two. We conclude in Section VI.

II. PROBABILISTIC MODELING OF A WIRELESS CHANNEL

This section provides a brief overview of wireless channel modeling, as needed for the model-based reconstruction of the next section. Readers are referred to [2], [3] for more details. In the wireless communication literature, it is well established that a communication channel between two nodes can be modeled as a multi-scale dynamical system with three major dynamics: multipath fading (small-scale fading), shadowing and path loss. Fig. 1 shows the received signal power across a route in the basement of the ECE building at UNM. The three main dynamics of the received signal power are marked on the figure. As can be seen, the received power can have rapid spatial variations that are referred to as multipath fading. By spatially averaging the received signal locally and over distances that channel can still be considered stationary, a slower dynamic emerges, which is called shadowing. Finally, by averaging over the variations of shadowing, a distance-dependent trend is seen, which is referred to as path loss. Two parameters are important in characterizing the channel: the distribution of a sample of
the channel as well as its spatial correlation. Let \( P_r \) represent the receiver signal power (the solid black curve of Fig. 1). Empirical data has shown Nakagami distribution to be a good match for the distribution of small-scale fading. As for the spatial correlation of small-scale fading, on the other hand, there is no single model that can be a good match for different environments. If the environment is rich in scatterers, for instance, the Fourier transform of the autocorrelation function of small-scale fading will have a form that is referred to as Jakes spectrum [2]. However, there is no general form that can fit most environments.

Once we average over small-scale variations, another dynamic can be observed, which changes at a slower rate. Let \( \overline{P}_r \) represent the average of the received power. \( \overline{P}_r \) varies over larger distances and is referred to as shadowing. It is the result of the transmitted signal being possibly blocked by a number of obstacles before reaching the receiver. Empirical data has shown lognormal distribution to be a good match for the distribution of shadowing. Let \( \overline{P}_{r,\text{dB}} = 10 \log_{10}(\overline{P}_r) \).

We have the following for the distribution of \( \overline{P}_{r,\text{dB}} \):

\[
p(\overline{P}_{r,\text{dB}}) = \frac{1}{\sqrt{2\pi}\alpha} e^{-\frac{(\overline{P}_{r,\text{dB}}-\mu_{\text{dB}})^2}{2\alpha^2}},
\]

where \( \mu_{\text{dB}} = K_{\text{dB}} - 10\gamma \log d \) and \( \alpha \) is the standard deviation from average. Consider the distance-dependent path loss, \( \mu = K/d^\gamma \), where \( d \) represents the distance between the transmitting and receiving robots and \( \gamma \) denotes the power fall-off rate. Then, it can be seen from Eq. 1 that \( \mu_{\text{dB}} = 10 \log \mu \) represents the average of the large-scale variations. Thus, the distance-dependent path loss characterizes the average of the shadowing variations (which is thus non-stationary due to the varying average), as can be seen from Fig. 1 as well.

Fig. 2 shows the pdf of shadowing using several collected data in the basement of ECE building. It can be seen that the distribution of the log of the shadowing variations (after removing the distance-dependent average) matches a zero-mean normal distribution very well. The standard deviation for this match is \( \alpha = 2.8 \).

As for the spatial correlation of shadowing, there is less mathematical characterizations. Gudmundson [4] characterizes an exponentially-decaying spatial correlation function for shadowing, which is widely used.

### III. Model-based Channel Estimation

As can be seen from the previous section, the received signal power can be modeled probabilistically. Therefore, we can use such models for estimating the strength of a wireless channel in a given area, based on very few measurements. In order to estimate a random field, based on few measurements, we need to have both the spatial correlation and sample distribution of the field. Therefore, since for multipath fading, there is no general model that would fit the correlation, we develop our estimator by only considering the models for shadowing and path loss. As such, the multipath fading component will appear as noise in the estimation. More specifically, we use the fact that Gaussian distribution matches the distribution of the average of the received signal power in dB and that an exponential function can characterize the correlation of shadowing for several scenarios. We, however, note that while such models have shown to match the characterization of a wireless channel in several cases, they clearly can not be a good fit for all scenarios. We show the impact of such modeling errors in Section V.

Let \( q_b \in \mathbb{R}^2 \) denote the position of a fixed transmitter (base station). We are interested in estimating the received signal strength (in the reception from the transmitter) at all positions in a given area, based on only a few direct measurements. The received signal at position \( q \in \mathbb{R}^2 \) can be modeled as follows by only considering shadowing and path loss:

\[
P_{r,\text{dB}} = K_{\text{dB}} - 10\gamma \log(||q-q_b||) + \nu_s, \tag{2}
\]

where \( P_{r,\text{dB}} \) denotes the received signal power in dB and \( \nu_s \) is a zero-mean Gaussian variable representing shadowing impact. Note that there is a modeling error as we are not considering the small-scale fading. In order to estimate the signal at positions that are not measured directly, we need...
to first estimate the underlying parameters of this model, i.e., the parameters of path loss ($K_{dB}$ and $\gamma$) and the correlation parameters of the shadowing term. Next, we will show how to estimate these parameters. Let $y = [y_1, \cdots, y_n]^T \in \mathbb{R}^n$ be a vector of all the available channel measurements in dB, where $n$ denotes the number of available measurements. Based on the models of the previous section, we have

$$y = H\theta + v + w,$$

where $H = \begin{bmatrix} 1 & -10\log(||q_1 - q_0||) \\ \vdots & \vdots \\ 1 & -10\log(||q_n - q_0||) \end{bmatrix}$, (3)

and $\theta = [K_{dB}, \gamma]^T$ contains the path-loss parameters. Here $q_i$ is the position of the $i^{th}$ measurement and $v \in \mathbb{R}^n$ is a zero-mean Gaussian random vector with $R = [r_{i,j}]_{n \times n}$, denoting its covariance matrix. The element $r_{i,j}$ in the covariance matrix characterizes the correlation between the $i^{th}$ and $j^{th}$ samples and is typically modeled by an exponential function as discussed in the previous section: $r_{i,j} = \alpha e^{-||q_i - q_j||/\beta}$. Thus we need to first estimate $K_{dB}, \gamma, \alpha$ and $\beta$ before we can estimate the channel based on Eq. 2. In the following, we first consider a Maximum Likelihood (ML) estimation approach. Let $p(y|\theta, \alpha, \beta)$ denote the conditional pdf of $y$, given the parameters $\theta, \alpha, \beta$:

$$p(y|\theta, \alpha, \beta) = \frac{1}{(2\pi)^{n/2}|R|^{1/2}} \exp\left(-\frac{(y-H\theta)^T R^{-1} (y-H\theta)}{2}\right),$$ (5)

where $|R|$ is the determinant of $R$. We have the following ML estimation of the parameters:

$$\hat{\theta}, \hat{\alpha}, \hat{\beta} = \arg\max_{\theta, \alpha, \beta} \log p(y|\theta, \alpha, \beta) \quad \hat{\alpha} = \frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{\theta})^T \hat{\Sigma}_{\hat{\beta}} (y_i - \hat{\theta}),$$ (6)

where $\hat{\Sigma}_{\hat{\beta}} = \frac{1}{\alpha} \hat{\Sigma}$ is the normalized covariance matrix, which is only a function of the decorrelation distance $\beta$ and the positions of the samples. Then we have:

$$\hat{\theta} = (H^T \hat{\Sigma}_{\hat{\beta}} H)^{-1} H^T \hat{\Sigma}_{\hat{\beta}} (y - H \hat{\theta}),$$

$$\hat{\alpha} = \frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{\theta})^T \hat{\Sigma}_{\hat{\beta}} (y_i - \hat{\theta}).$$ (7)

In order to estimate $\beta$, we insert $\hat{\theta}$ and $\hat{\alpha}$ in Eq. 6 to have the following:

$$\hat{\beta} = \arg\min_{\beta} \beta \left( |R(\beta)| + n \log(\hat{\alpha}) \right)$$

$$= \arg\min_{\beta} \left( y^T M(\beta) \hat{\Sigma}_{\hat{\beta}} M(\beta) y \right)^{1/2},$$

where $M(\beta) = I - H (H^T \hat{\Sigma}_{\hat{\beta}} H)^{-1} H^T \hat{\Sigma}_{\hat{\beta}}$. Note that if $\beta \to 0$, the samples become uncorrelated, which results in

$$\lim_{\beta \to 0} \hat{\beta} = (H^T H)^{-1} H^T y,$$

$$\lim_{\beta \to 0} \hat{\alpha} = \frac{1}{n} \left( y - H \hat{\beta} \right)^T (y - H \hat{\beta}).$$ (9)

Therefore, the estimate of $\theta$ and $\alpha$ will not depend on the estimate of $\beta$, as expected. It can be easily confirmed that Eq. 9 is also the result of a Least Squares (LS) estimation of $\theta$ and $\alpha$. The estimation of $\beta$, on the other hand, is typically challenging, as matrix $R$ can become ill-conditioned depending on the sampling pattern. This requires devising suboptimism but robust estimation strategies, as we discuss next. Let $\hat{R}$ represent the estimate of $R$ through numerical averaging. Then, we can find the $\beta$ that results in the best exponential fit to $\hat{R}$ (measured by the Frobenius norm of the difference in dB) as follows:

$$\hat{\beta} = \frac{1}{|D|} \sum_{d_i \in D} \frac{d_i}{\log \frac{\alpha}{\hat{\beta}(d_i)}},$$ (10)

where $D = \{d_i | \frac{\alpha}{\hat{\beta}(d_i)} \geq 1\}$, $|D|$ denotes the size of set $D$ and $\hat{\beta}(d_i)$ is the numerical estimate of $\beta$ at distance $d_i$. This can also be extended to a weighted approach as follows:

$$\hat{\beta} = \frac{1}{\sum_{d_i \in D} w_i} \sum_{d_i \in D} w_i d_i,$$ (11)

where $w_i$ is the corresponding weight for the $i^{th}$ term and can be chosen based on our assessment of the accuracy of the estimation of $\hat{\beta}(d_i)$. For instance, if we have very few measurements at a specific distance between two points, then the weight should be smaller.

In case the location of the transmitting node is not known, then the path loss parameters can be estimated by finding the best line fit to the log of the received measurements (as can be seen from Fig. 1). $\alpha$ can then be estimated by calculating the deviation from this average and $\beta$ can be estimated as explained previously. Alternatively, the position of the base station can also be added to the unknown parameters and jointly estimated.

### A. Channel Estimation

Once the underlying parameters of our model are estimated, channel at any arbitrary position $q \in \mathbb{R}^2$, can be estimated as follows. Let $\mu(q) = 10 \log |P_r(q)|$. We have the following for the probability of $\mu(q)$ conditioned on all the gathered measurements, $y$, and the path-loss and shadowing parameters:

$$p(\mu(q)|y, \theta, \alpha, \beta) \sim \mathcal{N}(\mu_{\text{ave}}(q), \sigma_{\mu}(q)),$$

where $\mathcal{N}(\cdot)$ denotes a normal distribution and

$$h(q) = [1 - 10\log(||q - q_0||)]^T,$$

$$\phi(q) = [e^{-\|q - q_0\|/\beta}, \cdots, e^{-\|q - q_n\|/\beta}]^T.$$ (13)

Then, $\mu_{\text{ave}}(q)$ will be the estimate of the channel (in dB) at position $q$, where we use the estimated $\theta, \alpha, \beta$. 

1948
IV. SPARSITY-BASED CHANNEL ESTIMATION [1]

In this part, we briefly summarize another approach for channel estimation using a small number of measurements. In this approach, the sparsity of the channel in the frequency domain, together with the recent results in the area of compressive sampling theory, are used for channel estimation based on a considerably incomplete data set. Readers are referred to [1], [5] for more details.

A. Compressive sampling theory

A sparse signal is a signal that can be represented with a small number of non-zero coefficients. A compressible signal is a signal that has a transformation where most of its energy is in a very few coefficients, making it possible to approximate the rest with zero. The new theory of compressive sampling [5] shows that, under certain conditions, a compressible signal can be reconstructed using very few observations. Most natural signals are indeed compressible. The best sparse representation of a signal depends on the application and can be inferred from analyzing similar data. Consider a scenario where we are interested in recovering a vector $x \in \mathbb{R}^N$. In our case, $x$ represents the received signal strength over the field of interest. We refer to the domain of $x$ as the primal domain. For 2D signals, vector $x$ can represent the columns of the matrix of interest stacked up to form a vector. Let $z \in \mathbb{R}^n$ where $n \ll N$ represent the incomplete linear measurements of vector $x$ obtained by the sensors. We will have

$$z = \Phi x,$$

where we refer to $\Phi$ as the observation matrix. Clearly, solving for $x$ based on the observation set $z$ is an ill-posed problem as the system is severely under-determined ($n \ll N$). However, suppose that $x$ has a sparse representation in another domain, i.e. it can be represented as a linear combination of a small set of vectors:

$$x = \Gamma X,$$

where $\Gamma$ is an invertible matrix and $X$ is $S$-sparse, i.e. $|\text{supp}(X)| = S \ll N$, where $\text{supp}(X)$ refers to the set of indices of the non-zero elements of $X$ and $|\cdot|$ denotes its cardinality. This means that the number of non-zero elements in $X$ is considerably smaller than $N$. Then we will have

$$z = \Psi X,$$

where $\Psi = \Phi \times \Gamma$. We refer to the domain of $X$ as the sparse domain (or transform domain). If $S \leq n$ and we knew the positions of the non-zero coefficients of $X$, we could solve this problem with traditional techniques like least-squares. In general, however, we do not know anything about the structure of $X$ except for the fact that it is sparse (which we can validate by analyzing similar data). The new theory of compressive sensing allows us to solve this problem.

Theorem 1 (see [5] for details and the proof): If $n \geq 2S$ and under specific conditions, the desired $X$ is the solution to the following optimization problem:

$$\min \|X\|_0, \text{ subject to } z = \Psi X,$$

where $\|X\|_0 = |\text{supp}(X)|$ represents the zero norm of vector $X$.

Theorem 1 states that we only need $2 \times S$ measurements to recover $X$ and therefore $x$ fully. This theorem, however, requires solving a non-convex combinatorial problem, which is not practical.

Instead, consider the following $\ell_1$ relaxation of the aforementioned $\ell_0$ optimization problem:

$$\min \|X\|_1, \text{ subject to } z = \Psi X.$$

(18)

Theorem 2: (see [6], [7], [8], [9] for details) Assume that $X$ is $S$-sparse. The $\ell_1$ relaxation can exactly recover $X$ from measurement $z$ if matrix $\Psi$ satisfies the Restricted Isometry Condition (RIC) [10] for $(2S, \sqrt{2} - 1)$.

Restricted Isometry Condition (RIC) [10]: Matrix $\Psi$ satisfies the RIC with parameters $(Z, \epsilon)$ for $\epsilon \in (0, 1)$ if

$$1 - \epsilon \leq \frac{|\Psi c|_2}{||c||_2} \leq 1 + \epsilon$$

(19)

for all $Z$-sparse vector $c$.

While it is not possible to define all the classes of matrices $\Psi$ that satisfy RIC, it is shown that random partial Fourier matrices [11] satisfy RIC with the probability $1 - O(N^{-M})$ if $n \geq B_M S \times \log^{O(1)} N$, where $B_M$ is a constant, $M$ is an accuracy parameter and $O(\cdot)$ is Big-O notation [5]. This shows that the number of required measurements could be considerably less than $N$. While the recovery of sparse signals is important, in practice signals may rarely be sparse. Most signals, however, will be compressible. In practice, the observation vector $y$ will also be corrupted by noise. The $\ell_1$ relaxation and the corresponding required RIC condition can be easily extended to the case of noisy observations with compressible signals [6]. The $\ell_1$ optimization problem of Eq. 18 can be posed as a linear programming problem [12]. The compressive sensing algorithms that reconstruct the signal based on $\ell_1$ optimization are typically referred to as “Basis Pursuit” [7].

The Restricted Isometry Condition also implies that the columns of matrix $\Psi$ should have a certain near-orthogonality property. Matching Pursuit (MP) approaches, on the other hand, are another class of algorithms that use this property to iteratively reconstruct the signal with less computational complexity. Readers are referred to [10], [13] for more details on this.

B. Sparsity-based channel estimation

Our analysis of several channel measurements has shown the channel to be compressible in the frequency domain for several scenarios. Thus we can also use this framework for channel estimation based on a small number of measurements. In this case, vector $z$ represents all the collected channel measurements. Vector $x$ is then the variable of interest, which denotes the values of the channel over the field of interest. Consequently, $X$ represents the Fourier transform of $x$ (for a 2D channel, this is a vector that is formed by stacking up all the samples of the 2D Fourier). While, in several scenarios, a wireless channel can be considered fairly compressible, there are cases that this may not be true as we show in the next section.
V. CHANNEL PREDICTION AND THE UNDERLYING TRADEOFFS

In this part we show the performance of the aforementioned approaches for channel estimation based on a small number of measurements. As we shall see, each approach has its own strength that can result in a better reconstruction depending on the scenario. Fig. 3 shows channel measurement across a street in San Francisco (data is courtesy of Mark Smith [14]). Fig. 4 measures the sparsity of the channel in the frequency domain. The figure shows $-10\log(\text{NMSE})$, where NMSE denotes the normalized mean square error of the difference between the channel and its sparsified version. In order to generate a sparsified channel, for any point on the x-axis, that percentage of the ordered Fourier coefficients are kept (ordered decreasingly) while the rest are zeroed. Then, the plot characterizes how compressible this channel is. As can be seen, this channel is fairly compressible. Fig. 5 shows the performance of both the sparsity-based and model-based approaches for the reconstruction of this channel, where the x-axis shows the percentage of the measurements gathered (as a % of the whole area of interest). The y-axis shows $-10\log(\text{NMSE})$, where NMSE is the normalized mean square error of the estimation. In this case, the gathered measurements are randomly distributed over the channel. It can be seen that when the number of measurements are small (less than 13.5%), the sparsity-based approach outperforms the model-based one. This makes sense as the model-based approach needs to estimate the underlying parameters. For a very small number of measurements, the error in the estimation of these parameters can be high, resulting in a performance degradation in the overall estimation. As the number of measurements increases, the model-based approach then outperforms the sparsity-based one in this case.

As expected, the model-based approach would be sensitive to the accuracy of the underlying model. In order to see this, Fig. 6 shows another channel measurement in San Francisco [14]. It can be seen that this channel can not be well characterized by only one path loss trend. As a result, we expect that the performance of the model-based approach degrades. Fig. 7 shows the performance of channel reconstruction in this case. It can be seen that the sparsity-based approach outperforms the model-based one in this case. For this case, the channel is considerably compressible in the Fourier domain, which is evident from the good performance of the sparsity-based approach. As discussed previously, the performance of the sparsity-based approach depends on the compressibility of the channel in the frequency domain. There could be cases where the spatial variations of the channel in an area of interest is not that compressible. In order to see this, Fig. 8 shows the reconstruction of the channel in a small 2D area in the basement of the ECE building. The area is 3.28 ft by 49.2125 ft. Fig. 9 shows the sparsity of this channel in the same way that we measured the sparsity for Fig. 4. It can be seen that this channel is not that sparse. As a result, it can be seen from Fig. 8 that the sparsity-based approach does not perform that well and that the model-based approach outperforms the sparsity-based one for most part. It can also be seen that the performance of the model-based approach has degraded considerably as compared to the previous channels, due to the possible mismatch in the underlying model as well as error in the estimation of the parameters.

In general, both approaches can be useful in estimating a wireless channel based on a small number of measurements. We are currently working on a more rigorous characterization of the underlying tradeoffs between the two approaches.
Fig. 6. Another channel measurement across a street in San Francisco [14].

Fig. 7. Estimation Performance for the channel of Fig. 6.

Fig. 8. Performance of the model-based and sparsity-based approaches for a 2D channel in the basement of the ECE building.

Fig. 9. Characterizing the sparsity of the channel of Fig. 8.

VI. CONCLUSIONS

In this paper we considered estimating the spatial variations of a channel based on a small number of gathered measurements. We proposed a model-based estimation approach, in which we used a multi-scale probabilistic model in order to characterize the channel. We furthermore considered a sparsity-based channel estimation approach based on the compressibility of the channel in the frequency domain. Our results showed that both approaches can be effective in estimating the channel with only a few gathered measurements. We then discussed the underlying tradeoffs between the two approaches. Overall, the proposed framework can be useful for communication-aware motion planning in robotic networks, where a prediction of the link qualities is needed.

REFERENCES