# Graph Learning and Augmentation Based Interpolation of Signal Strength for Location-Aware Communications

Hong-Ming Chiu
*Dept. of Electronics Engineering*
*National Chiao Tung University*
Hsinchu, Taiwan
hongmingchiu0217@gmail.com

Carrson C. Fung
*Institute of Electronics*
*National Chiao Tung University*
Hsinchu, Taiwan
c.fung@ieee.org

Antonio Ortega
*Signal and Image Processing Institute*
*Department of Electrical and Computer Engineering*
*University of Southern California*
Los Angeles, USA
antonio.ortega@sipi.usc.edu

***Abstract*—A graph learning and augmentation (GLA) technique is proposed herein to solve the received signal power interpolation problem, which is important for preemptive resource allocation in location-aware communications. A graph parameterization results in the proposed GLA interpolator having superior mean-squared error performance and lower computational complexity than the traditional Gaussian process method. Simulation results and analytical complexity analysis are used to prove the efficacy of the GLA interpolator.**

***Index Terms*—Graph learning, graph augmentation, interpolation, location-aware, resource allocation**

## I. Introduction

Achieving high spectral efficiency and low latency are important goals in wireless communication systems such as 5G and beyond. One way of accomplishing these goals is to preemptively allocate communication resources, such as access points (APs), to locations that may require extra power. A data-driven approach can be used to estimate the power at various locations by first collecting a limited number of power signal samples (training samples) in the field, and then interpolating the signal strength at locations where measurements were not taken. Gaussian processes (GP) [1] were proposed in [2]–[4] to estimate received signal power in an entire region that contains a single AP by modeling the signal power as a GP, where its mean and covariance are parameterized by channel model parameters. These parameters are estimated offline using training samples that contain signal power collected at different locations and the locations of which they were taken. The signal power at a new location can then be obtained by cross-correlating the locations of the training samples with the new location, where the correlation process depends on the mean and covariance matrix of the GP estimated in the training step. The correlation will be high provided that the new location is in the proximity of where the training samples were taken. The WiFi-SLAM problem was extended to the case where received signal power has to be estimated from several APs in the field in [5], where the need to know the exact measurement locations was bypassed by modeling the location as a latent variable in the GP model.

Besides spatial estimation of received signal power, [6] also utilized the Gaussian process and kernel "trick" [7] to predict future channel gain. In [6], both the slow- and fast-fading component of the wireless channel are modeled. [8] proposed a Gaussian process technique that can predict received signal power even when there is uncertainty in the location where the training samples are collected. Unfortunately, all these previously proposed techniques rely on accurately estimating the parameters of the channel model, which is unlikely to be possible with sufficient reliability, given that the underlying estimation problem is nonconvex. Although alternating minimization can be used to learn the parameters, it is highly sensitive to initial conditions and is not guaranteed to converge.

A graph learning and augmentation technique is proposed herein to estimate received power signal in a region with a single AP by collecting training samples of power signal at several locations prior to the estimation process. The main novelty of proposed approach is to regard training signals as graph signals [9], corresponding to a graph where each vertex represents a measurement location. The relationships between signals at different nodes are modeled by a generalized graph Laplacian that corresponds to the precision matrix of a Gaussian-Markov random field (GMRF) model [10]. An additional graph augmentation step is required to solve the signal strength interpolation problem. Hence, the proposed interpolator is called graph learning and augmentation (GLA) interpolator. Even though the GP approach relies on estimating the covariance matrix of the GP, which can be regarded as a graph Laplacian, this matrix is populated by hyperparameters that are obtained by solving nonconvex problems, which in turn, affects the quality of the graph matrix. This is not the case for the GLA, as the elements are directly estimated from the data by solving a series of convex problems.

This work has been partially supported by the Ministry of Science and Technology Grant 108-2221-E-009-126 and Ministry of Education project Trusted Intelligent Edge/Fog Computing Technology RSC 108B568.

## II. METHODOLOGY

### A. Notation

Uppercase (lowercase) bold face letters indicate matrices (column vectors). Superscript $(\cdot)^T$ denotes transposition. $[\mathbf{A}]_{ij}$ denotes the $(i,j)^{th}$ element of $\mathbf{A}$. $\mathrm{E}[\cdot]$ stands for statistical expectation of the entity inside the square bracket. $\mathcal{N}(\boldsymbol{\mu}, \mathbf{C})$ denotes Gaussian distribution with mean $\boldsymbol{\mu}$ and covariance matrix $\mathbf{C}$. $\mathcal{U}[a,b]$ denotes Uniform distribution between $a$ and $b$. $\mathbf{A} \succeq \mathbf{0}_N$ designates $\mathbf{A}$ as an $N \times N$ symmetric positive semidefinite matrix. $\mathbf{0}_N$, $\mathbf{1}_N$ and $\mathbf{I}_N$ denote a column vector with $N$ zeros, $N$ ones and $N \times N$ identity matrix, respectively. $\mathrm{tr}(\cdot)$ denotes the trace of the matrix. $(\mathbf{A})_{\text{off}}$ denotes the off-diagonal elements of $\mathbf{A}$. $(\mathbf{A})^{\circ n}$ denotes raising the elements of $\mathbf{A}$ to the power of $n$. $\exp(\mathbf{A})$ denotes element-wise exponential of matrix $\mathbf{A}$. $\mathrm{vech}(\cdot)$ denotes matrix vectorization (i.e. stacking the columns) with only the off-diagonal elements omitted. $\odot$ denotes matrix element-wise multiplication and $|\cdot|$ denotes matrix determinant.

### B. Signal and Channel Model

Denote $\mathbf{x}_s, \mathbf{x}_i \in \mathbb{R}^2$ as vectors that contain the $x$ and $y$ coordinates of the source and sink node, respectively. Assume the log-normal shadow fading model [11, p. 102], and without loss of generality, further assume that the transmit signal power equals 0 dBm. The received signal power at $\mathbf{x}_i$ from an AP at $\mathbf{x}_s$ can be expressed in dBm as [12] $p(\mathbf{x}_s, \mathbf{x}_i) = L_0 - 10\eta \log_{10}(\|\mathbf{x}_s - \mathbf{x}_i\|_2) + \psi(\mathbf{x}_s, \mathbf{x}_i)$, where $L_0$ is the reference distance, $\eta$ is the path-loss exponent and $\psi(\mathbf{x}_s, \mathbf{x}_i) \sim \mathcal{N}\left(0, \sigma_\psi^2\right)$ represents the location-dependent shadow fading in dB between $\mathbf{x}_s$ and $\mathbf{x}_i$. $\sigma_\psi^2$ denotes the shadowing variance. It is well known that the spatial correlation of the shadow fading is a decreasing function [13] such that its spatial covariance function can be given as [12] $C(\mathbf{x}_i, \mathbf{x}_j) \triangleq \mathrm{E}[\psi(\mathbf{x}_s, \mathbf{x}_i), \psi(\mathbf{x}_s, \mathbf{x}_j)] = \sigma_\psi^2 \exp\left(-\frac{\|\mathbf{x}_i - \mathbf{x}_j\|_2}{d_c}\right)$, where $d_c$ denotes correlation distance. It is clear that the correlation of the shadow fading between two locations only depends on their relative distance. Notice that the above log-distance path loss model is not the only choice and the subsequent method can be generalized, for example, to wideband PCS microcell model [11, p. 120].

Assume that $N$ sensors are randomly placed at different locations in a fixed-size region to measure signal strengths and that there are $L$ remaining points where the received signal power needs to be estimated. That is, the region is sampled uniformly at $N+L$ locations and $N$ samples are randomly selected to be the measured points. Each sensor is assumed to collect signal for $T$ units of time so that the noisy observed signal strength at time $t \in \{1, \ldots, T\}$ is written as $\mathbf{p}_{st} = \mathbf{J}_s \mathbf{p}_t + \mathbf{w}_{st} \in \mathbb{R}^N$, where $\mathbf{J}_s \in \mathbb{R}^{N \times (N+L)}$ denotes a downsampling matrix that downsamples the signal strength vector $\mathbf{p}_t \in \mathbb{R}^{N+L}$, which contains received signal power at all of the $L+N$ points at time $t$. $\mathbf{w}_{st} \in \mathbb{R}^N$ denotes the zero-mean additive white Gaussian noise (AWGN) vector at time $t$ with $\sigma_w^2$ denoting its variance. Furthermore, define

$$\mathbf{d} \triangleq [-10\log_{10}(\|\mathbf{x}_s - \mathbf{x}_1\|_2) \ldots - 10\log_{10}(\|\mathbf{x}_s - \mathbf{x}_{N+L}\|_2)]^T$$

then

$$\mathbf{p}_t = \mathbf{H}\boldsymbol{\theta} + \boldsymbol{\psi} \text{ and } \mathbf{p}_{st} = \mathbf{J}_s(\mathbf{H}\boldsymbol{\theta} + \boldsymbol{\psi}) + \mathbf{w}_{st}, \quad (1)$$

where $\mathbf{H} = \begin{bmatrix}\mathbf{1}_{N+L} & \mathbf{d}\end{bmatrix} \in \mathbb{R}^{(N+L)\times 2}$, $\boldsymbol{\theta} = \begin{bmatrix}L_0 & \eta\end{bmatrix}^T$, and $\boldsymbol{\psi} \triangleq \begin{bmatrix}\psi(\mathbf{x}_s, \mathbf{x}_1) & \cdots & \psi(\mathbf{x}_s, \mathbf{x}_{N+L})\end{bmatrix}^T$. It is clear that $\mathbf{p}_t$ is also Gaussian distributed $\mathbf{p}_t \sim \mathcal{N}(\boldsymbol{\mu}, \mathbf{C})$ with $[\mathbf{C}]_{ij} = C(\mathbf{x}_i, \mathbf{x}_j)$, for $i, j = \{1, \ldots, N+L\}$ and $\boldsymbol{\mu} = \mathbf{H}\boldsymbol{\theta}$.

### C. Power Map Interpolation

The proposed GLA signal strength interpolation algorithm involves three steps. First, using the data collected from $N$ sensors, a generalized graph Laplacian (GGL), $\widehat{\boldsymbol{\Sigma}}_s^{-1}$, is obtained from the algorithm in [10], which measures the relationship between the signal strength at different graph vertices. This is followed by an augmentation step to "expand" the graph model for all $N+L$ points $\widehat{\boldsymbol{\Sigma}}$ so that the final interpolation step can take place. The second step solves for the estimate of $\boldsymbol{\theta}$, which includes the estimate of $L_0$ and $\eta$, by collecting and exploiting $T$ samples of $\mathbf{p}_{st}$ during the estimation process. Finally, an estimate of shadow fading $\widehat{\boldsymbol{\psi}}$ is acquired via MAP estimation by exploiting $\widehat{\boldsymbol{\Sigma}}_s$ and $\widehat{\boldsymbol{\Sigma}}$ obtained in the first step.

### D. Step 1: Graph Learning and Augmentation (GLA)

The GGL $\boldsymbol{\Sigma}_s^{-1} \in \mathbb{R}^{N \times N}$ can be obtained as [10]

$$\widehat{\boldsymbol{\Sigma}}_s^{-1} = \underset{\boldsymbol{\Sigma}_s^{-1} \in \mathcal{L}_g}{\arg\max} \log\left|\boldsymbol{\Sigma}_s^{-1}\right| + \mathrm{tr}\left(\mathbf{S}\boldsymbol{\Sigma}_s^{-1}\right) + \left\|\mathbf{M} \odot \boldsymbol{\Sigma}_s^{-1}\right\|_1, \quad (2)$$

where $\mathbf{M} \succeq \mathbf{0}_N$ denotes a regularization matrix. $\mathbf{S} = \frac{1}{T}\sum_{t=1}^T (\mathbf{p}_{st} - \overline{\mathbf{p}}_s)(\mathbf{p}_{st} - \overline{\mathbf{p}}_s)^T$, with $\overline{\mathbf{p}}_s = \frac{1}{T}\sum_{t=1}^T \mathbf{p}_{st}$ being the sample mean of $\mathbf{p}_{st}$. $\mathcal{L}_g \triangleq \{\mathbf{L} \,|\, \mathbf{L} \succeq \mathbf{0}_N, [\mathbf{L}]_{ij} \leq 0, i \neq j = 1, \ldots, N\}$ defines a set of $N \times N$ GGL matrices. $\mathbf{M} = \mathbf{0}$ in the simulation.

A linear regression approach is used to obtain the complete graph, which contains pairwise weights between all the $N+L$ points. The regression utilizes $\widehat{\boldsymbol{\Sigma}}_s$ to find the channel parameters that will then be used to find the complete graph matrix. Define $\mathbf{D} \in \mathbb{R}^{(N+L)\times(N+L)}$ whose $(i,j)^{th}$ element equals $[\mathbf{D}]_{ij} = \|\mathbf{x}_i - \mathbf{x}_j\|_2$, for $i, j = 1, \ldots, N+L$, and let $\mathbf{D}_s = \mathbf{J}_s \mathbf{D} \mathbf{J}_s^T$ be the matrix that contains only Euclidean distances between measurements. From the channel model and AWGN noise assumptions, the off-diagonal elements of $\widehat{\boldsymbol{\Sigma}}_s$ can be approximated as

$$\left(\widehat{\boldsymbol{\Sigma}}_s\right)_{\text{off}} \approx \left(\mathbf{J}_s \mathbf{C} \mathbf{J}_s^T + \sigma_w^2 \mathbf{I}_N\right)_{\text{off}} = \left(\mathbf{J}_s \sigma_\psi^2 \exp\left(-\frac{\mathbf{D}}{d_c}\right) \mathbf{J}_s^T\right)_{\text{off}}$$
$$= \left(\exp\left(\ln \sigma_\psi^2 - \frac{\mathbf{D}_s}{d_c}\right)\right)_{\text{off}}.$$

Then $(-\ln \widehat{\boldsymbol{\Sigma}}_s)_{\text{off}} = (-\ln \sigma_\psi^2 + \frac{\mathbf{D}_s}{d_c})_{\text{off}}$. The right-hand side can be expanded as an $n$th-order polynomial in $\mathbf{D}_s$ so that

$-\ln \widehat{\mathbf{\Sigma}}_s \approx a_0 + a_1\mathbf{D}_s + \cdots + a_n\mathbf{D}_s^{\circ n}$. The parameter $\mathbf{a} = \begin{bmatrix} a_0 & \cdots & a_n \end{bmatrix}^T$ can be easily obtained by solving

$$\begin{aligned}\widehat{\mathbf{a}} &= \underset{\mathbf{a}\in\mathbb{R}^{n+1}}{\arg\min} \left\| \mathrm{vech}\left(-\log\left(\widehat{\mathbf{\Sigma}}_{\boldsymbol{s}}\right)\right) - \mathbf{B}\mathbf{a} \right\|_2^2 \\ &= \left(\mathbf{B}^T\mathbf{B}\right)^{-1}\mathbf{B}^T \mathrm{vech}\left(-\log\left(\widehat{\mathbf{\Sigma}}_{\boldsymbol{s}}\right)\right),\end{aligned}$$

where $\mathbf{B} = \left[\mathbf{1}_{N(N-1)}\ \mathrm{vech}\left(\mathbf{D}_s\right)\ \mathrm{vech}\left(\mathbf{D}_s^{\circ 2}\right) \ldots \mathrm{vech}\left(\mathbf{D}_s^{\circ n}\right)\right]$. The complete graph $\widehat{\mathbf{\Sigma}} \in \mathbb{R}^{(N+L)\times(N+L)}$ is therefore

$$\widehat{\mathbf{\Sigma}} = \exp\left(-\left(\widehat{a}_0 + \widehat{a}_1\mathbf{D} + \widehat{a}_2\mathbf{D}^{\circ 2} + \ldots + \widehat{a}_n\mathbf{D}^{\circ n}\right)\right). \quad (3)$$

### *E. Step 2: Mean Estimation*

With the partial graph $\widehat{\mathbf{\Sigma}}_s$ from the first step, the joint probability density function (pdf) of the random variables $\mathbf{P}_1, \ldots, \mathbf{P}_T$ with unknown parameter $\boldsymbol{\theta}$ can be calculated as

$$f_{\mathbf{P}_1,\ldots,\mathbf{P}_T}(\mathbf{p}_1,\ldots,\mathbf{p}_T;\boldsymbol{\theta}) = \prod_{t=1}^{T} f_{\mathbf{P}_t}(\mathbf{p}_t) \frac{1}{\sqrt{(2\pi)^N |\widehat{\mathbf{\Sigma}}_s|}} \exp\left(-\frac{1}{2}\left(\mathbf{p}_{st} - \mathbf{J}_s\mathbf{H}\boldsymbol{\theta}\right)^T \widehat{\mathbf{\Sigma}}_s^{-1}\left(\mathbf{p}_{st} - \mathbf{J}_s\mathbf{H}\boldsymbol{\theta}\right)\right).$$

Notice that $\mathbf{P}_t$, for $t = 1, \ldots, T$, are assumed to be statistically independent, even though this may not be true in an actual scenario. A maximum likelihood estimator can then be used to find the mean parameter $\boldsymbol{\theta}$ as

$$\begin{aligned}\widehat{\boldsymbol{\theta}} &= \underset{\widehat{\boldsymbol{\theta}}}{\arg\max}\, f_{\mathbf{P}_1,\ldots,\mathbf{P}_T}(\mathbf{p}_1,\ldots,\mathbf{p}_T;\boldsymbol{\theta}) \\ &= \left(\mathbf{H}_s^T\widehat{\mathbf{\Sigma}}_s^{-1}\mathbf{H}_s\right)^{-1}\mathbf{H}_s^T\widehat{\mathbf{\Sigma}}_s^{-1}\overline{\mathbf{p}}_s,\end{aligned}$$

where $\mathbf{H}_s = \mathbf{J}_s\mathbf{H}$. The estimated mean $\widehat{\boldsymbol{\mu}}$ is therefore

$$\widehat{\boldsymbol{\mu}} = \mathbf{H}\left(\mathbf{H}_s^T\widehat{\mathbf{\Sigma}}_s^{-1}\mathbf{H}_s\right)^{-1}\mathbf{H}_s^T\widehat{\mathbf{\Sigma}}_s^{-1}\overline{\mathbf{p}}_s. \quad (4)$$

### *F. Step 3: MAP Estimator*

Finally, solving the interpolation problem to obtain an estimate of $\mathbf{p}_t$, for $t = T+1, \ldots$ requires finding an estimate for $\boldsymbol{\psi}$. A MAP estimator in this case can exploit the partial and complete graphs, $\widehat{\mathbf{\Sigma}}_s$ and $\widehat{\mathbf{\Sigma}}$, from (2) and (3), respectively. This can be done by first zeroing the mean of $\mathbf{p}_{st}$ in (1) using $\widehat{\boldsymbol{\theta}}$ so that $\mathbf{p}'_{st} \triangleq \mathbf{p}_{st} - \mathbf{H}_s\widehat{\boldsymbol{\theta}} = \mathbf{J}_s\boldsymbol{\psi} + \mathbf{w}_{st}$. Then the posterior pdf of $\boldsymbol{\psi}$ becomes $f_{\mathbf{\Psi}|\mathbf{P}'_{st}}(\boldsymbol{\psi}|\,\mathbf{p}'_{st}) \propto f_{\mathbf{P}'_{st}|\mathbf{\Psi}}(\mathbf{p}'_{st}|\,\boldsymbol{\psi})\, f_{\mathbf{\Psi}}(\boldsymbol{\psi}) = f_{\mathbf{W}_s}(\mathbf{p}'_{st} - \mathbf{J}_s\boldsymbol{\psi})\, f_{\mathbf{\Psi}}(\boldsymbol{\psi})$, where the last equality is true because $\boldsymbol{\psi}$ and $\mathbf{w}_{st}$ are independent. The first term in the posterior pdf can be evaluated as

$$f_{\mathbf{W}_s}(\mathbf{p}'_{st} - \mathbf{J}_s\boldsymbol{\psi}) = \frac{1}{\sqrt{(2\pi\sigma_w^2)^N}} \exp\left(-\frac{1}{2\sigma_w^2}\left\|\mathbf{p}'_{st} - \mathbf{J}_s\boldsymbol{\psi}\right\|_2^2\right)$$

and the prior distribution of $\boldsymbol{\psi}$ equals

$$f_{\mathbf{\Psi}}(\boldsymbol{\psi}) = \frac{1}{\sqrt{(2\pi)^N|\widehat{\mathbf{\Sigma}}|}} \exp\left(-\frac{1}{2}\boldsymbol{\psi}^T\widehat{\mathbf{\Sigma}}^{-1}\boldsymbol{\psi}\right).$$

Then

$$\begin{aligned}\widehat{\boldsymbol{\psi}} &= \underset{\boldsymbol{\psi}\in\mathbb{R}^{N+L}}{\arg\max}\, f_{\mathbf{W}_s}(\mathbf{p}'_{st} - \mathbf{J}_s\boldsymbol{\psi})\, f_{\mathbf{\Psi}}(\boldsymbol{\psi}) \\ &= \underset{\boldsymbol{\psi}\in\mathbb{R}^{N+L}}{\arg\min} \frac{1}{\sigma_w^2}\left(\mathbf{p}'_{st} - \mathbf{J}_s\boldsymbol{\psi}\right)^T\left(\mathbf{p}'_{st} - \mathbf{J}_s\boldsymbol{\psi}\right) + \boldsymbol{\psi}^T\widehat{\mathbf{\Sigma}}^{-1}\boldsymbol{\psi}.\end{aligned}$$

Therefore,

$$\begin{aligned}\widehat{\boldsymbol{\psi}} &= \frac{1}{\sigma_w^2}\left(\widehat{\mathbf{\Sigma}}^{-1} + \frac{1}{\sigma_w^2}\mathbf{J}_s^T\mathbf{J}_s\right)^{-1}\mathbf{J}_s^T\mathbf{p}'_{st} \\ &= \widehat{\mathbf{\Sigma}}\mathbf{J}_s^T\left(\mathbf{J}_s\widehat{\mathbf{\Sigma}}\mathbf{J}_s^T + \sigma_w^2\mathbf{I}_N\right)^{-1}\mathbf{p}'_{st}, \quad (5)\end{aligned}$$

where the last equality is obtained by using the Woodbury identity so as to decrease the computational complexity compared to the prior expression which involves two matrix inverses. Using the result from (4) and (5), the interpolated received signal power $\widehat{\mathbf{p}}_t$ is therefore

$$\widehat{\mathbf{p}}_t = \mathbf{H}\widehat{\boldsymbol{\theta}} + \widehat{\boldsymbol{\psi}} = \widehat{\boldsymbol{\mu}} + \widehat{\mathbf{\Sigma}}\mathbf{J}_s^T\left(\mathbf{J}_s\widehat{\mathbf{\Sigma}}\mathbf{J}_s^T + \sigma_w^2\mathbf{I}_N\right)^{-1}\mathbf{p}'_{st}. \quad (6)$$

## III. Simulation Results

### *A. GLA Interpolation*

The interpolated power map using the proposed GLA interpolation algorithm is shown in Fig. 1. The ground truth is generated using synthetic power map data generated from the log-normal channel model in [11], [12]. The data were generated for a region of size 200m $\times$ 200m with a single AP in the center, with a resolution of 4m. The hyperparameters $[\sigma_\psi, d_c, L_0, \eta] = [9, 70, 10, 3]$. $N = 400$ sensors are randomly placed in the region to collect the signal strength for $T = 100$ units of time. AWGN with variance $\sigma_w^2$ is generated based on the signal-to-noise ratio SNR $= 10\log\left(\frac{\sigma_\psi^2}{\sigma_w^2}\right)$.

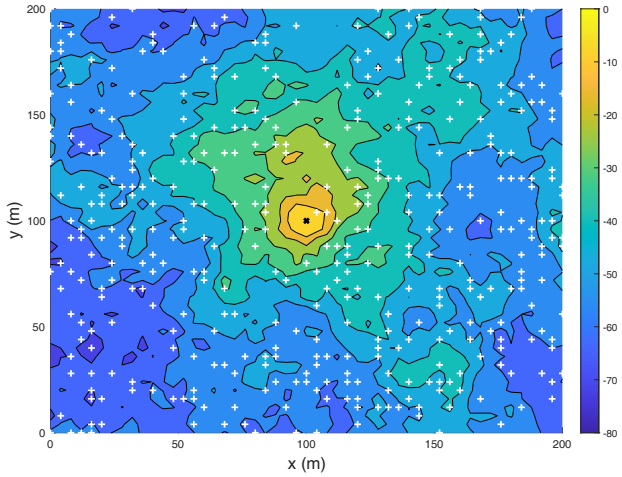
Fig. 1 shows the interpolated result using the proposed GLA algorithm at time instance $t > T$ with SNR = 14dB. Note that the signal strength of the entire field, including the measured points, is interpolated. It can be observed that power map obtained from the proposed method is similar to the ground truth.

### *B. Model Mismatch*

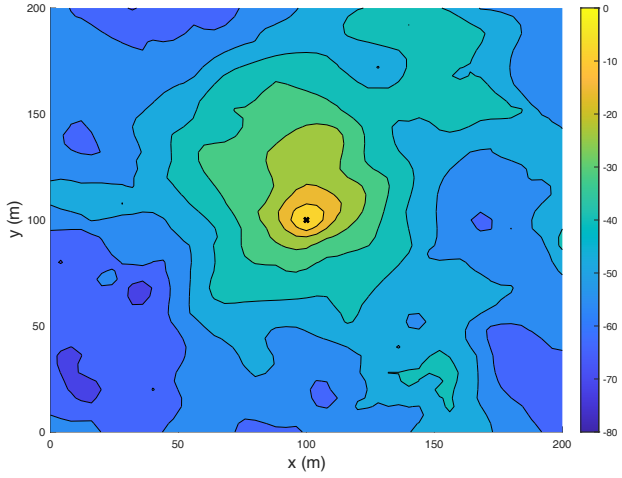
To show the efficacy of the proposed technique, a model mismatch is introduced during the interpolation process. The coordinates of sample points in power map $\mathbf{x}_i$ is assumed to be corrupted by measurement error $\mathbf{w}_i \in \mathbb{R}^2$ up to half of the spatial resolution, such that the actual coordinates are used during the interpolation process is

$$\widetilde{\mathbf{x}}_i = \mathbf{x}_i + \mathbf{w}_i,$$

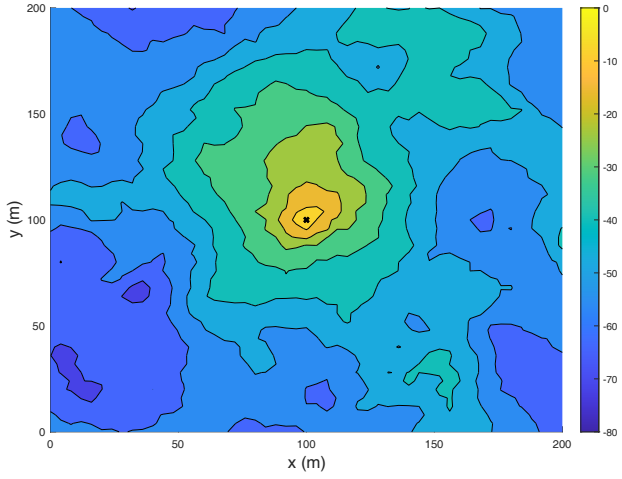
where $i = 1, \ldots, N+L$ and $\mathbf{w}_i \sim \mathcal{U}[-2, 2]$ (in meters). Fig.1c shows the interpolated result with model mismatch and SNR = 14 dB. It can be seen that the result is similar to that of Fig. 1b, where there was no model mismatch.




(a)




(b)




(c)

Fig. 1: Ground truth and interpolated power map in dBm. The white plus sign denotes the position of sensors and the black cross sign denotes the position of AP. SNR = 14 dB. (a) True power map and the position of sensors. (b) Interpolated power map. (c) Interpolated power map with model mismatch.

### C. Performance Assessment

The aggregate mean-squared error (MSE) and regional MSE are adopted as performance metrics, where the aggregated MSE is defined as

$$\frac{1}{R}\sum_{t=T+1}^{T+R}\frac{1}{N+L}\left\|\mathbf{p}_t-\widehat{\mathbf{p}}_t\right\|_F^2,$$

which measures the overall MSE between true and interpolated received power. $R$ is the number of Monte-Carlo simulation.

The regional MSE is defined as

$$\frac{1}{R}\sum_{t=T+1}^{T+R}\left\|p_t\left(\mathbf{x}_s,\mathbf{x}_i\right)-\widehat{p}_t\left(\mathbf{x}_s,\mathbf{x}_i\right)\right\|_2^2,$$

which measures the MSE between true and interpolated received power at location $\mathbf{x}_i$. In addition to the MSE, the run-time performance was also computed on a server with an Intel 8 core i7-2600 3.4GHz CPU.

The following simulations are obtained using the same environment setting and hyperparameters as stated above, and the number of experiments for Monte-Carlo simulation equals $R = 1000$. The aggregate and regional MSE, and run-time of the proposed method and the GP based method in [2] are compared. The block coordinate descent (BCD) method was used to estimate the parameters $[\sigma_\psi, d_c, L_0, \eta]$ in the GP method.

### D. Aggregate MSE

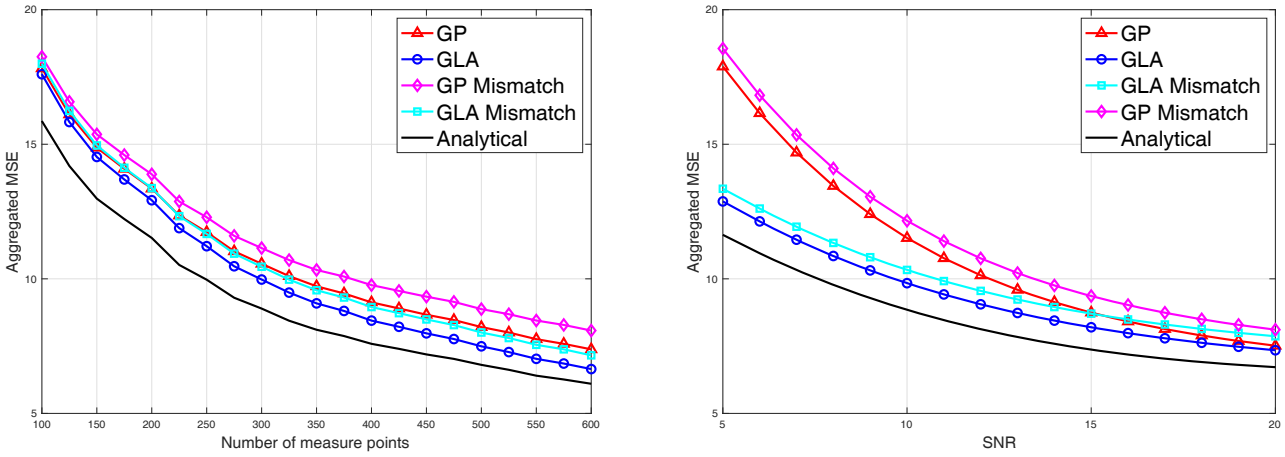
Figs. 2a and 2b show the aggregate MSE for different number of measurements and SNR, respectively. The aggregate MSE with mismatch model are also plotted for comparison and the analytical aggregate MSE [12] is also plotted as a performance benchmark. Observe from Figs. 2a and 2b that the proposed GLA method can still achieve better performance than the GP with model mismatch, especially in low SNR. This is because, unlike the GP approach, GLA does not require solving any nonconvex problems and, thus, does not have any initial condition issues which might cause divergence or inaccurate estimated of parameters.

### E. Regional MSE

Fig. 3 shows the regional MSE for the proposed GLA algorithm. Observe that the regions that are close to measurements have lower MSE compared to the regions that are surrounded by fewer measurements, which agrees with intuition.

### F. Complexity Analysis

In the first step of the proposed GLA algorithm, the BCD algorithm [10] for solving (2) has $O(T_d(N)+N^2)$ complexity for each iteration, where $T_d$ denotes the time complexity required to solve the underlying quadratic programming problem. The graph augmentation step (3) has $O((N+L)^2)$ complexity due to the element-wise squaring of the matrix. The mean and MAP estimator in steps 2 and 3 involve only matrix multiplication and inversion which exhibit worst-case complexity of $O((N+L)^3)$. However, the overall computation




(a) (b)

Fig. 2: Aggregate MSE vs. (a) number of measured points. SNR = 14dB. (b) SNR.

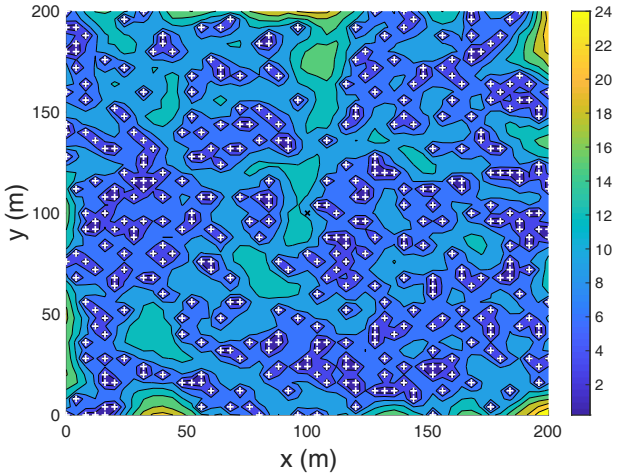



Fig. 3: Regional MSE, where the white plus sign denote sensors' position and the black cross sign denotes AP's position. SNR = 14dB.

complexity in the first step cannot be completely characterized due to the fact that the BCD algorithm requires varying number of iterations to converge. Yet the run-time of the proposed GLA algorithm should exhibit polynomial growth since the BCD is guaranteed to converge due to the convex nature of the underlying problem.

Table I shows the empirical run-time for different ratio of number of measurements to the total number of points $\alpha = \frac{N}{N+L}$ for the GLA and GP approaches. The convergence criterion of the BCD algorithm [10, (59)] for GLA is set to $\epsilon = 10^{-9}$, and the convergence criterion of the BCD algorithm for GP [2, (8)(9)(10)] is set to $\| [\sigma_\psi, d_c]_{pre} - [\sigma_\psi, d_c] \|_2 < 10^{-5}$.

| $\alpha$ | 5% | 10% | 15% | 20% | 25% | 30% |
|---|---|---|---|---|---|---|
| GP | 102 | 1049 | 2803 | 4926 | 6644 | 8560 |
| GLA | 12 | 47 | 157 | 386 | 909 | 1577 |

TABLE I: Empirical run-time (in seconds).

## IV. Conclusion

A graph learning and augmentation signal strength interpolation algorithm is proposed for location-aware communications, which has better MSE and run-time performance than a Gaussian process based method. The GLA approach is better able to capture relationships between measurements of signal strength, which resulted in its superior performance.

## References


[1] C.E. Rasmussen and C.K.I. Williams, *Gaussian Processes for Machine Learning*, MIT Press, 2006.

[2] B. Ferris, D. Hahnel, and D. Fox, "Communication for teams of networked robots," *Proc. of the Robotics: Science and Systems, Philadelphia, PA, US*, 2006.

[3] J. Fink, "Communication for teams of networked robots," *Ph.D. thesis, Electrical and System Engineering, Univ. of Pennsylvania*, 2011.

[4] J. Riihijärvi and P. Mähönen, "Machine learning performance prediction in mobile cellular networks," *IEEE Computational Intelligence Magazine*, vol. 13, no. 1, pp. 51–60, 2018.

[5] B. Ferris, D. Fox, and N. Lawrence, "Wifi-slam using gaussian process latent variable models," *Proc. of the 20th Intl. Joint Conf. on Artificial Intelligence*, 2007.

[6] Q. Liao, S. Valentin, and S. Stancza, "Channel gain prediction in wireless networks based on spatial-temporal correlation," *Proc. of the IEEE 16th Intl. Workshop on Signal Processing Advances in Wireless Communications, Stockholm, Sweden*, 2015.

[7] B. Schölkopf and A.J. Smola, *Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond*, The MIT Press, Cambridge, Massachusetts, 2002.

[8] L.S. Muppirisetty, T. Svensson, and H. Wymeersch, "Spatial wireless channel prediction under location uncertainty," *IEEE Trans. on Wireless Communications*, vol. 15, no. 2, pp. 1031–1044, 2016.

[9] D.I. Shuman et al., "The emerging field of signal processing on graphs," *IEEE Signal Processing Magazine*, vol. 30, no. 3, pp. 83–98, 2013.

[10] H.E. Egilmez, E. Pavez, and A. Ortega, "Graph learning from data under laplacian and structural constraints," *IEEE Journal of Selected Topics in Signal processing*, vol. 11, no. 6, pp. 825–841, 2017.

[11] T.S. Rappaport, *Wireless Communications: Principles and Practice*, 2nd Ed., Prentice Hall, 2001.

[12] R. Di Taranto et al., "Location-aware communications for 5g networks," *IEEE Signal Processing Magazine*, vol. 31, no. 6, pp. 102–112, 2014.

[13] M. Gudmundson, "Correlation model for shadow fading in mobile radio systems," *Electronics Letters*, vol. 27, no. 23, 1991.