Learning Radio Maps for UAV-aided Wireless Networks: A Segmented Regression Approach

Junting Chen*, Uday Yatnalli*†, and David Gesbert*

*Communication Systems Department, EURECOM, Sophia Antipolis, France
†Institute for Communications Engineering, Technical University of Munich, Munich, Germany
Email:{chenju, yatnalli, gesbert}@eurecom.fr

***Abstract*—This paper targets the promising area of unmanned aerial vehicle (UAV)-assisted wireless networking, by which communication-enabled robots operate as flying wireless relays to help fill coverage or capacity gaps in the networks. In order to feed the UAV's autonomous path planning and positioning algorithm, a radio map is exploited, which must be, in practice, reconstructed from UAV-based measurements from a limited subset of locations. Unlike existing methods that ignore the segmented propagation structure of the radio map, this paper proposes a machine learning approach to reconstruct a finely structured map by exploiting both segmentation and signal strength models. A data clustering and parameter estimation problem is formulated using a maximum likelihood approach, and solved by an iterative clustering and regression algorithm. Numerical results demonstrate significant performance advantage in radio map reconstruction as compared to the baseline.**

I. Introduction

The exploitation of aerial robots (UAVs) is a promising technology for future wireless cellular communication networks. In dense urban areas, where wireless signals may be blocked by buildings, or where high user data demand exists, UAVs can serve to carry wireless relays to boost the link quality between the base station (BS) and the users. As such, much of the system gain hinges on the ability to optimally position the UAV as a function of user locations, base station location, and propagation conditions. The latter is known to be determined by a complex arrangement of path loss, shadowing and signal blockage, which ultimately depends on the fine-grained terrain/building topology. In principle, such information can be captured by a *radio map*, which describes the average signal strength for all combinations of BS-UAV-user locations.

There have been interesting recent works on UAV positioning for wireless capacity extension [1]–[7]. Some prior results on UAV position optimization assumed homogeneous path loss models and did not account for the local blockage between the UAV and the user [1]–[3]. However, the presence of the ground obstacles may block the propagation between the UAV and the user, and hence critically affect the performance. To address this issue, the works [4]–[7] established a simplified stochastic model to capture the probability of the line-of-sight (LOS) propagation of the UAV-user channel as a function of the elevation angle at the user and the distribution of the buildings. However, these methods only depend on the macroscopic terrain information, but cannot exploit the fine grained structure of the propagation conditions. Specifically, they cannot capture the fact that a small change of the UAV position may result in a huge difference in the channel gain. For example, the UAV may travel from a LOS location to a NLOS location near a building edge.

To determine the best UAV position, we propose to exploit a finely structured radio map that captures both the signal strength and the propagation conditions. The main difficulty is that the radio map is not easy to obtain, because it may have a complex structure due to the irregular shapes of the terrain and the random locations of the buildings. To circumvent this challenge, we aim at establishing a machine learning approach to reconstruct the radio map from a limited number of signal strength measurements obtained at a subset of locations. In principle, the goal is to train a channel predictor, interpolate the channel gain, and identify the propagation conditions for every UAV-user position.

In related works, a method based on artificial neural network for channel gain prediction was proposed in [8], but it required the input of terrain information such as vegetation type and vegetation density to train the channel predictor. In [9], support vector regression was applied to perform path loss prediction, and in [10], a kernel-based nonparametric algorithm was studied for power map reconstruction from quantized power measurements. However, the main limitation of these works is that they did not exploit the underlying segmented structure in the radio map, which is induced by the terrain or building topology. As a result, they cannot predict the propagation condition. In addition, the work [8] depends on a terrain model, which may be inaccurate or unavailable.

In this paper, we develop a radio map learning and reconstruction approach that both *exploits* and *reveals* the topology-induced structure. To this end, we formulate a joint clustering and regression problem using a maximum likelihood approach. An iterative algorithm is derived, which clusters map points according to propagation model differences and learns the regression model parameters for each cluster. It is demonstrated that the proposed method requires almost as few as $1/10$ of

978-1-4673-8999-0/17/$31.00 ©2017 IEEE

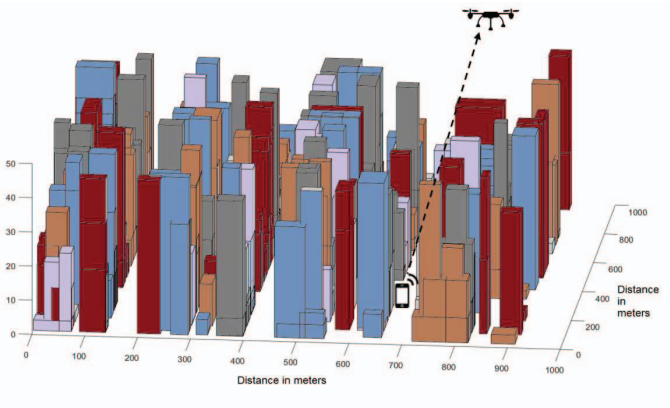

Figure 1. An urban city map where a UAV tries to find a LOS location to the ground user.

training samples for the same reconstruction performance as a k-nearest neighbor (KNN) baseline does.

II. System model

A. Geographical Topology

Consider a (possibly dense) urban area, where a user located at $\mathbf{x}_{\mathrm{U}}$ is surrounded by a number of buildings as illustrated in Fig. 1. Consider a UAV hovering over the city to relay the wireless signal for the user. Denote the position of the UAV as $\mathbf{x}_{\mathrm{D}}$. For exposition purpose, assume that the UAV moves at a fixed height H_{D} above the ground, i.e., $\mathbf{x}_{\mathrm{D}} = (\bar{\mathbf{x}}_{\mathrm{D}}, H_{\mathrm{D}})$, where $\bar{\mathbf{x}}_{\mathrm{D}} \in \mathbb{R}^2$ and H_{D} is larger than the maximum height of the buildings. In addition, we focus on UAV-user link channel reconstruction and assume a fixed user position $\mathbf{x}_{\mathrm{U}}$.[1]

B. Channel Model

Classically, the channel gain in dB can be modeled as

$$\gamma = \beta - 10\alpha \log_{10} d + \xi$$

where α is the path loss exponent, d is the transmitter-to-receiver distance, β can be measured as the average channel gain at the reference point $d = 1$ meter, and ξ is usually modeled as a Gaussian random variable $\mathcal{N}(0, \sigma^2_{\mathrm{SF}})$ to capture the shadowing effect. In addition, the set of parameters α, β, and σ^2_{SF} depend on the propagation scenario, such as LOS or non-line-of-sight (NLOS) propagations.

In this paper, we exploit the spatial structure of the shadowing component ξ, where it may be spatially correlated due to common obstacles and reflectors in a local area. As a result, the random variable ξ may follow different distributions in different areas. Towards this end, consider a general K-segment ray-tracing model for the UAV-user channel. Let $\mathbb{D} \subseteq \mathbb{R}^6$ be the domain of all possible UAV-user position pairs. Consider a partition of $\mathbb{D}$ into K disjoint segments: $\mathbb{D} = \mathcal{D}_1 \cup \mathcal{D}_2 \cup \cdots \cup \mathcal{D}_K$, where $\mathcal{D}_k \cap \mathcal{D}_j = \varnothing$, for $k \neq j$. The UAV-user channel gain in dB is further modeled as

$$\gamma(\mathbf{x}_{\mathrm{D}}) = \sum_{k=1}^{K} \big(\beta_k - 10\alpha_k \log_{10} d(\mathbf{x}_{\mathrm{D}}) + \xi_k\big)\mathbb{I}\{(\mathbf{x}_{\mathrm{D}}, \mathbf{x}_{\mathrm{U}}) \in \mathcal{D}_k\} \tag{1}$$

where $d(\mathbf{x}_{\mathrm{D}}) = \|\mathbf{x}_{\mathrm{D}} - \mathbf{x}_{\mathrm{U}}\|$ is the distance between the UAV and the user, and $\mathbb{I}\{A\}$ is an indicator function taking value 1 if A is true, and 0 otherwise. In addition for $(\mathbf{x}_{\mathrm{D}}, \mathbf{x}_{\mathrm{U}})$ being in the kth propagation segment, α_k represents average path loss exponent, β_k represents the average channel gain at the reference point at $d = 1$ meter, and ξ_k is modeled as $\mathcal{N}(0, \sigma_k^2)$ to capture the local shadowing effect.

Remark 1 (Motivation of the segmented ray-tracing model): Traditional channel models usually classify the propagation into only two scenarios, LOS and NLOS. Yet, there may be a scenario where the direct path penetrates some light obstacles, such as foliage. In this case, the resulting channel gain is stronger than the NLOS case, but weaker than the LOS case. This is characterized as obstructed LOS propagation in [4]. Based on a measurement campaign for the air-to-ground channel in the city center of Bristol in UK reported in [4], it was shown that if the propagation is classified into three segments, LOS, obstructed LOS, and NLOS, the shadowing standard deviation σ_{SF} is roughly 1, 3, and 5 dB, respectively. Such observation gives us the insight that if we can identify more propagation segments, the variance of ξ_k in (1) will decrease, and hence a K-segmented channel model can make prediction of the channel with high precision.

C. Radio Map

Define the *radio map* as

$$\bar{\gamma}(\mathbf{x}_{\mathrm{D}}) = \sum_{k=1}^{K} \big(\beta_k - 10\alpha_k \log_{10} d(\mathbf{x}_{\mathrm{D}})\big)\mathbb{I}\{(\mathbf{x}_{\mathrm{D}}, \mathbf{x}_{\mathrm{U}}) \in \mathcal{D}_k\} \tag{2}$$

where we drop the shadowing component ξ_k in (1) and treat ξ_k as *observation noise* to $\bar{\gamma}(\mathbf{x}_{\mathrm{D}})$. Note that when the variance σ_k^2 of ξ_k is small, $\bar{\gamma}(\mathbf{x}_{\mathrm{D}})$ yields a good approximation of the actual channel $\gamma(\mathbf{x}_{\mathrm{D}})$ in (1).

We focus on reconstructing $\bar{\gamma}(\mathbf{x}_{\mathrm{D}})$ from a few training samples collected from offline channel measurements at different locations. Existing methods, such as support vector regression, artificial neural networks, and the matrix completion, do not explicitly exploit the segmented structure of $\bar{\gamma}(\mathbf{x}_{\mathrm{D}})$ modeled in (2), and hence are not suitable for radio map reconstruction. Such a reconstruction problem is challenging, because the channel measurement data does not directly contain the segment information. Instead, the system needs a clustering algorithm to partition the training data into K groups according to group-specific models with unknown parameters α_k and β_k to be learned. In the following sections, a segmented regression method is proposed for radio map reconstruction.

[1]The extension to varying height of the UAV and varying position of the user are relatively straight forward. In addition, the method studied in this paper can be applied to BS-UAV link channel reconstruction.

III. Learning the Segmented Model via Segmented Regression

Consider a set of measurement samples $\{(\mathbf{x}^{(l)}, y^{(l)}) : l = 1, 2, \ldots, N\}$, where $\mathbf{x}^{(l)} = \mathbf{x}_{\rm D}^{(l)}$ represents the UAV position and $y^{(l)}$ represents the channel gain measured at position $\mathbf{x}^{(l)}$.[2] According to the segmented propagation model (2), each data sample is to be classified into one of the K segments. Let $\mathbf{z}^{(l)} = (z_1^{(l)}, z_2^{(l)}, \ldots, z_K^{(l)})$ be the label for the data sample $(\mathbf{x}^{(l)}, y^{(l)})$, where $z_k^{(l)} = 1$ means that $(\mathbf{x}^{(l)}, y^{(l)})$ belongs to the kth propagation segment, and $z_k^{(l)} = 0$ otherwise. The labels $\mathbf{z}^{(l)}$ are not observed but to be determined by the training algorithm.

Given that $(\mathbf{x}, y)$ belongs to the kth propagation segment, the observation model can be written as

$$y = \alpha_k g(\mathbf{x}) + \beta_k + \xi_k \tag{3}$$

where $g(\mathbf{x}) = -10\log_{10}(\|\mathbf{x}_{\rm D} - \mathbf{x}_{\rm U}\|_2)$ and $\xi_k \sim \mathcal{N}(0, \sigma_k^2)$ is the observation noise. Note that the variance σ_k^2 is not known prior to the training.

The joint probability density function (PDF) for $(\mathbf{x}, y)$ conditioned on $z_k = 1$ is thus given by

$$p_k(\mathbf{x}, y) = \frac{1}{\sqrt{2\pi}\sigma_k} \exp\left\{-\frac{(y - \alpha_k g(\mathbf{x}) - \beta_k)^2}{2\sigma_k^2}\right\} \tag{4}$$

and the joint PDF of $(\mathbf{x}, y, \mathbf{z})$ can be written as

$$\begin{aligned} p(\mathbf{x}, y, \mathbf{z}) &= \sum_{k=1}^{K} \mathbb{P}\{\mathbf{X} = \mathbf{x}, Y = y \big| Z_k = 1\}\mathbb{P}\{Z_k = 1\} \\ &= \sum_{k=1}^{K} p_k(\mathbf{x}, y)\pi_k \\ &= \prod_{k=1}^{K} p_k(\mathbf{x}, y)^{z_k} \times \prod_{k=1}^{K} \pi_k^{z_k}. \end{aligned} \tag{5}$$

where

$$\pi_k \triangleq \mathbb{P}\{Z_k = 1\}$$

is the marginal probability of a data sample $(\mathbf{x}, y)$ belonging to the kth propagation segment. Note that $\sum_{k=1}^{K} \pi_k = \sum_{k=1}^{K} \mathbb{P}\{Z_k = 1\} = 1$.

A. A Maximum Likelihood Approach

Let $\boldsymbol{\theta} = \{\alpha_k, \beta_k, \sigma_k, \pi_k\}_{k=1}^K$ be the collection of parameters. Based on the measurement samples and the set of labels $\{x^{(l)}, y^{(l)}, z^{(l)}\}$, the likelihood function of the parameter $\boldsymbol{\theta}$ is given by

$$\mathcal{L}(\boldsymbol{\theta}) = \prod_{l=1}^{N} p(\mathbf{x}^{(l}, y^{(l)}, z^{(l)} \big| \boldsymbol{\theta}) \tag{6}$$

[2]As a straight forward extension, one can consider a general case of varying the user position as well: $\mathbf{x}^{(l)} = (\mathbf{x}_{\rm D}^{(l)}, \mathbf{x}_{\rm U}^{(l)})$.

and a maximum likelihood estimation (MLE) of $\boldsymbol{\theta}$ can be obtained as the solution to the following problem

$$\underset{\boldsymbol{\theta}, \{\mathbf{z}^{(l)}\}}{\text{maximize}} \quad \prod_{l=1}^{N} p(\mathbf{x}^{(l}, y^{(l)}, \mathbf{z}^{(l)} \big| \boldsymbol{\theta}) \tag{7}$$

$$\text{subject to} \quad \sum_{k=1}^{K} \pi_k = 1 \tag{8}$$

Maximizing the likelihood function in (6) is equivalent to maximizing the log-likelihood function $\log \mathcal{L}(\boldsymbol{\theta})$. Using the joint PDF $p(\mathbf{x}, y, \mathbf{z})$ in (5), the log-likelihood function can be computed as follows

$$\log \mathcal{L}(\boldsymbol{\theta}) = \sum_{l=1}^{N}\sum_{k=1}^{K} z_k^{(l)} \Big[\log \pi_k + \log p_k(\mathbf{x}^{(l)}, y^{(l)}) \Big]. \tag{9}$$

Note that the exact expression of log-likelihood function $\log \mathcal{L}(\boldsymbol{\theta})$ is not available, since the segment labels $z_k^{(l)}$ are unknown. However, the statistics of $z_k^{(l)}$ can be computed can be computed in terms of the parameter $\boldsymbol{\theta}$ and the samples $\{\mathbf{x}^{(l)}, y^{(l)}\}$. Specifically, we have

$$\begin{aligned} \bar{z}_k^{(l)}(\boldsymbol{\theta}) &\triangleq \mathbb{E}\{Z_k^{(l)} \big| \mathbf{x}^{(l)}, y^{(l)}, \boldsymbol{\theta}\} \\ &= \mathbb{P}\left\{Z_k^{(l)} = 1 \big| \mathbf{x}^{(l)}, y^{(l)}, \boldsymbol{\theta}\right\} \\ &= \frac{p_k(\mathbf{x}^{(l)}, y^{(l)} \big| \boldsymbol{\theta}) \mathbb{P}\left\{Z_k^{(l)} = 1 \big| \boldsymbol{\theta}\right\}}{\sum_{j=1}^{K} p_j(\mathbf{x}^{(l)}, y^{(l)} \big| \boldsymbol{\theta}) \mathbb{P}\left\{Z_j^{(l)} = 1 \big| \boldsymbol{\theta}\right\}} \\ &= \frac{p_k(\mathbf{x}^{(l)}, y^{(l)} \big| \boldsymbol{\theta})\pi_k}{\sum_{j=1}^{K} p_j(\mathbf{x}^{(l)}, y^{(l)} \big| \boldsymbol{\theta})\pi_j}. \end{aligned} \tag{10}$$

Therefore, we consider to maximize the expected log-likelihood $\mathbb{E}\big\{\log \mathcal{L}(\boldsymbol{\theta})\big\}$ given by

$$\begin{aligned} &\mathbb{E}\big\{\log \mathcal{L}(\boldsymbol{\theta})\big\} \\ &\quad = \sum_{l=1}^{N}\sum_{k=1}^{K} \bar{z}_k^{(l)}(\boldsymbol{\theta}) \Big[\log \pi_k + \log p_k(\mathbf{x}^{(l)}, y^{(l)}) \Big]. \end{aligned} \tag{11}$$

As a result, the MLE problem (7) can be relaxed into

$$\underset{\boldsymbol{\theta}}{\text{maximize}} \quad \mathbb{E}\big\{\log \mathcal{L}(\boldsymbol{\theta})\big\} \tag{12}$$

$$\text{subject to} \quad \sum_{k=1}^{K} \pi_k = 1.$$

B. The Iterative Segmented Regression Algorithm

The expected MLE problem (12) is non-convex, and hence it is difficult to find the optimal solution. Therefore, we focus on an efficient iterative search for a sub-optimal solution to (12). In particular, the following two properties of the objective function (11) can be exploit: first, the $\bar{z}_k^{(l)}(\boldsymbol{\theta})$ is easy to compute using (10), and second, by fixing $\bar{z}_k^{(l)}$, the objective function (11) is concave in π_k, α_k, β_k, and σ_k, respectively.

Let $\boldsymbol{\theta}^{(t)}$ be the collection of parameters at the tth iteration. We have

$$\bar{z}_k^{(l)}(\boldsymbol{\theta}^{(t)}) = \frac{p_k(\mathbf{x}^{(l)}, y^{(l)} \big| \boldsymbol{\theta}^{(t)})\pi_k^{(t)}}{\sum_{j=1}^{K} p_j(\mathbf{x}^{(l)}, y^{(l)} \big| \boldsymbol{\theta}^{(t)})\pi_j^{(t)}}. \tag{13}$$

The objective function (11) then becomes

$$Q(\cdot|\boldsymbol{\theta}^{(t)}) = \sum_{l=1}^{N}\sum_{k=1}^{K} \bar{z}_k^{(l)}(\boldsymbol{\theta}^{(t)})\Big[\log \pi_k + \log\Big(\frac{1}{\sqrt{2\pi}\sigma_k}\exp\Big\{-\frac{(y-\alpha_k g(\mathbf{x})-\beta_k)^2}{2\sigma_k^2}\Big\}\Big)\Big].$$

Denote the function $Q(\boldsymbol{\alpha},\boldsymbol{\beta}|\boldsymbol{\theta}^{(t)})$ as $Q(\cdot|\boldsymbol{\theta}^{(t)})$ by fixing all the variables as $\boldsymbol{\theta}^{(t)}$ except for $\{\alpha_k,\beta_k\}$; i.e., $\sigma_k=\sigma_k^{(t)}$ and $\pi_k=\pi_k^{(t)}$. Similar notations apply to $Q(\boldsymbol{\sigma}|\boldsymbol{\theta}^{(t)})$ and $Q(\boldsymbol{\pi}|\boldsymbol{\theta}^{(t)})$. We update the set of parameters $\boldsymbol{\theta}^{(t+1)}$ as the solutions to the following problems

$$\underset{\{\alpha_k,\beta_k\}}{\text{maximize}} \quad Q(\boldsymbol{\alpha},\boldsymbol{\beta}|\boldsymbol{\theta}^{(t)}) \tag{14}$$

$$\underset{\{\sigma_k\}}{\text{maximize}} \quad Q(\boldsymbol{\sigma}|\boldsymbol{\theta}^{(t)}) \tag{15}$$

and

$$\underset{\{\pi_k\}}{\text{maximize}} \quad Q(\boldsymbol{\pi}|\boldsymbol{\theta}^{(t)}) \tag{16}$$
$$\text{subject to} \quad \sum_{k=1}^{K}\pi_k = 1.$$

The optimization problems (14)–(16) are convex. They can be solve by Lagrangian methods and the solutions are given as follows.

Proposition 1 (Update equations): The solutions to the maximization problems (14)–(16) are given by

$$\begin{bmatrix} \alpha_k^{(t+1)} \\ \beta_k^{(t+1)} \end{bmatrix} = \mathbf{A}^{-1}\mathbf{b} \tag{17}$$

$$\sigma_k^{(t+1)} = \sqrt{\frac{\sum_{l=1}^{N}\bar{z}_k^{(l)}(\boldsymbol{\theta}^{(t)})(y^{(l)}-\alpha_k^{(t)}g(\mathbf{x}^{(l)})-\beta_k^{(t)})^2}{\sum_{l=1}^{N}\bar{z}_k^{(l)}(\boldsymbol{\theta}^{(t)})}} \tag{18}$$

and

$$\pi_k^{(t+1)} = \frac{1}{N}\sum_{l=1}^{N}\bar{z}_k^{(l)}(\boldsymbol{\theta}^{(t)}) \tag{19}$$

respectively, where

$$\mathbf{A} = \begin{bmatrix} \sum_{l=1}^{N}\bar{z}_k^{(l)}g(\mathbf{x}^{(l)})^2 & \sum_{l=1}^{N}\bar{z}_k^{(l)}g(\mathbf{x}^{(l)}) \\ \sum_{l=1}^{N}\bar{z}_k^{(l)}g(\mathbf{x}^{(l)}) & \sum_{l=1}^{N}\bar{z}_k^{(l)} \end{bmatrix}$$

and

$$\mathbf{b} = \begin{bmatrix} \sum_{l=1}^{N}\bar{z}_k^{(l)}g(\mathbf{x}^{(l)})y^{(l)} \\ \sum_{l=1}^{N}\bar{z}_k^{(l)}y^{(l)} \end{bmatrix}.$$

Proof: Problems (14)–(15) can be solved by setting the partial derivatives $\frac{\partial}{\partial\alpha_k}Q(\boldsymbol{\alpha},\boldsymbol{\beta}|\boldsymbol{\theta}^{(t)})$, $\frac{\partial}{\partial\beta_k}Q(\boldsymbol{\alpha},\boldsymbol{\beta}|\boldsymbol{\theta}^{(t)})$, and $\frac{\partial}{\partial\sigma_k}Q(\boldsymbol{\sigma}|\boldsymbol{\theta}^{(t)})$ to zero. Problem (16) can be solved by first forming the Lagrangian function

$$\mathfrak{L}(\boldsymbol{\pi},\lambda|\boldsymbol{\theta}^{(t)}) = Q(\boldsymbol{\pi}|\boldsymbol{\theta}^{(t)}) + \lambda\Big(\sum_{k=1}^{K}\pi_k - 1\Big) \tag{20}$$

and computing the solutions to the Karush-Kuhn-Tucker (KKT) conditions $\frac{\partial}{\partial\pi_k}\mathfrak{L}(\boldsymbol{\pi},\lambda|\boldsymbol{\theta}^{(t)})=0$, $\lambda(\sum_{k=1}^{K}\pi_k-1)=0$, and $\lambda\geq 0$. □

As a result, the iterative segmented regression algorithm proceeds as computing update equations (13), (17)–(19) in a recursive way until convergence.

Remark 2 (Initialization): An easy initialization is to first partition the set of measurement samples $\{(\mathbf{x}^{(l)},y^{(l)})\}$ into K groups using K-means algorithm purely based on the channel gain $y^{(l)}$. The initial values $\bar{z}_k^{(l)}(\boldsymbol{\theta}^{(-1)})$ is then assigned as 1, if the lth data sample is clustered into the kth group, and 0 otherwise. The initial values $\boldsymbol{\theta}^{(0)}$ are then computed by (17)–(19).

Remark 3 (Convergence): If the variances σ_k are given, the recursive algorithm specified by steps (13), (17), and (19) is an expectation-maximization (EM) algorithm. This is because with fixed σ_k, the function $Q(\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\pi}|\boldsymbol{\theta}^{(t)})$ is concave jointly in $(\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\pi})$, and hence (17) and (19) give the optimal solution to maximizing $Q(\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\pi}|\boldsymbol{\theta}^{(t)})$. It is known that an EM type algorithm converges to a local maximum of the expected log-likelihood $\mathbb{E}\{\log\mathcal{L}(\boldsymbol{\theta})\}$.

IV. Radio Map Reconstruction

To reconstruct the radio map $\bar{\gamma}(\mathbf{x})$, we need to classify each position $\mathbf{x}$ into one of the K segments based on the parameter $\boldsymbol{\theta}$ learned from the segmented regression algorithm, and apply the segmented model $\{\alpha_k,\beta_k,\sigma_k\}_{k=1}^{K}$ to compute the predicted channel gain $y=\hat{\gamma}(\mathbf{x})$.

There are many methods to build a classifier, such as logistic regression, support vector machines, artificial neural networks, and kernel estimation. Note that the goal of this paper is not to find the best classifier for radio map reconstruction, but to demonstrate the concept of the segmented learning approach in radio map reconstruction. For exposition purpose, we illustrate a simple classification method using KNN approach as follows.

For a given position $\mathbf{x}$, define the index set of the M-nearest neighbors of $\mathbf{x}$ as

$$\mathcal{N}(\mathbf{x}) \triangleq \underset{\mathcal{S}\subseteq\{1,2,\ldots,N\},|\mathcal{S}|=M}{\arg\min}\sum_{m\in\mathcal{S}}\|\mathbf{x}-\mathbf{x}^{(m)}\|. \tag{21}$$

The segment label at position $\mathbf{x}$ is given by

$$\hat{\mathbf{z}}(\mathbf{x}) = \mu\sum_{m\in\mathcal{N}(\mathbf{x})}\mathcal{K}(\mathbf{x},\mathbf{x}^{(m)})\bar{\mathbf{z}}^{(m)} \tag{22}$$

where

$$\mathcal{K}(\mathbf{x},\mathbf{x}^{(m)}) = \exp\Big\{-\frac{\|\mathbf{x}-\mathbf{x}^{(m)}\|^2}{s}\Big\} \tag{23}$$

is a kernel function to weight the data points $(\mathbf{x}^{(m)},y^{(m)},\bar{\mathbf{z}}^{(m)})$ in the set $\mathcal{T}_{\mathrm{L}}$ indexed by the neighbor set $\mathcal{N}(\mathbf{x})$, $s>0$ is the kernel parameter, and $\mu>0$ is to

normalize $\hat{\mathbf{z}}(\mathbf{x})$ such that the components of $\hat{\mathbf{z}}(\mathbf{x})$ sum up to 1; i.e., $\sum_{k=1}^{K} \hat{z}_k(\mathbf{x}) = 1$. The kernel function weights the data samples $m \in \mathcal{N}(\mathbf{x})$ according to their distance to the position $\mathbf{x}$.[3]

Based on the reconstructed label $\hat{\mathbf{z}}(\mathbf{x})$ in (22), two reconstruction methods for the channel gain at position $\mathbf{x}$ are considered.

Soft reconstruction: For UAV position $\mathbf{x}$,

$$\hat{\gamma}_{\mathrm{S}}(\mathbf{x}) = \sum_{k=1}^{K} \big(\beta_k - 10\alpha_k \log_{10} d(\mathbf{x})\big)\hat{z}_k(\mathbf{x}). \tag{24}$$

Note that we have $\hat{\gamma}_{\mathrm{S}}(\mathbf{x}) = \mathbb{E}\{\bar{\gamma}(\mathbf{x})\}$, where the expectation is taken over the estimated distribution of the segment label $\hat{\mathbf{z}}(\mathbf{x})$.

Hard reconstruction: For UAV position $\mathbf{x}$,

$$\hat{\gamma}_{\mathrm{H}}(\mathbf{x}) = \beta_{\hat{k}} - 10\alpha_{\hat{k}} \log_{10} d(\mathbf{x}) \tag{25}$$

where $\hat{k} = \arg\max_{k=1,2,\ldots,K} \hat{z}_k(\mathbf{x})$ is the most likely segment.

As compared to the hard reconstruction, soft reconstruction $\hat{\gamma}_{\mathrm{S}}(\mathbf{x})$ may yield smaller reconstruction errors, since it is a minimum mean-square error (MMSE) estimator of $\bar{\gamma}(\mathbf{x})$ based on the estimated distribution $\hat{\mathbf{z}}(\mathbf{x})$. However, it violates the segmented propagation model of the radio map in (2). By contrast, the hard construction $\hat{\gamma}_{\mathrm{H}}(\mathbf{x})$ first makes a maximum likelihood detection $\hat{k}$ on the propagation segment, and then applies the segmented propagation model (2) to compute the channel gain. As a result, some UAV positioning algorithms can still exploit the segmented propagation property to find the best UAV position [11].

V. NUMERICAL RESULTS

The top view of the dense urban city is shown in Fig. 2, where the buildings, represented by bricks, are densely distributed in a 1000×1000 m Manhattan-like area. The mean height of the building is 22.5 m and the maximum building height is 45 m. The UAV hovers over the city at a height of 50 m. There are four users randomly dropped in the city. We consider two propagation segments for exposition: LOS and NLOS. The model parameters are chosen as $\alpha_1 = 2.27, \beta_1 = -40$, $\alpha_2 = 3.64$, and $\beta_2 = -30$ according to some appropriate scenarios in the WINNER II channel model. The variances of the observation noise are $\sigma_1^2 = 1$ and $\sigma_2^2 = 3$.

We first visualize the reconstructed radio map by the proposed segmented regression algorithm. Fig. 3 shows the true and reconstructed radio maps for four different user locations indicated in Fig. 2. Measurement samples are collected uniformly from $N = 2500$ UAV positions. The hard reconstruction method in (25) is used to build the reconstructed radio maps. The kernel parameter in (23) is chosen as $s = 300$, based on cross-validations. It is observed that first, the two propagation segments due to terrain topology are clearly identified in the reconstructed radio maps. Note that such map structure is shown to be important in UAV positioning [11]. Second, within each segment, the radio maps are smooth thanks to the log-distance model exploited by the algorithm. Finally, the peaks of the radio maps reveal the user locations as maximum power appears when the UAV is on top of the user.

[3]To determine the choice of kernel function and the parameter s are beyond the scope of this paper. A standard way to choose s through cross-validations.

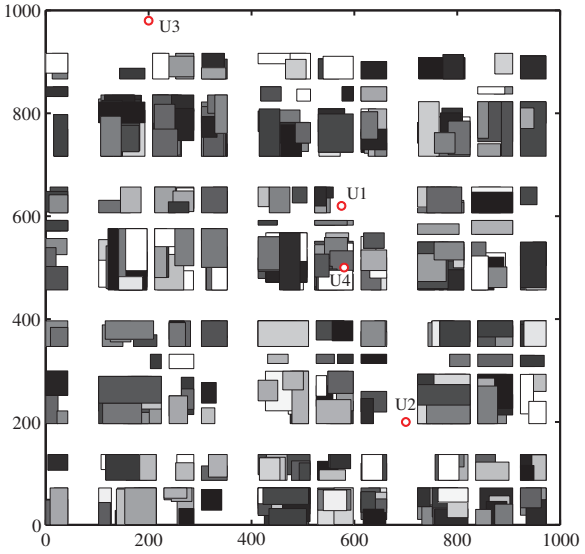

Figure 2. A map of the buildings and the users, where the buildings are represented by bricks, in which, dark bricks represent low buildings and light bricks represent high buildings. User 4 is on the rooftop of the building.

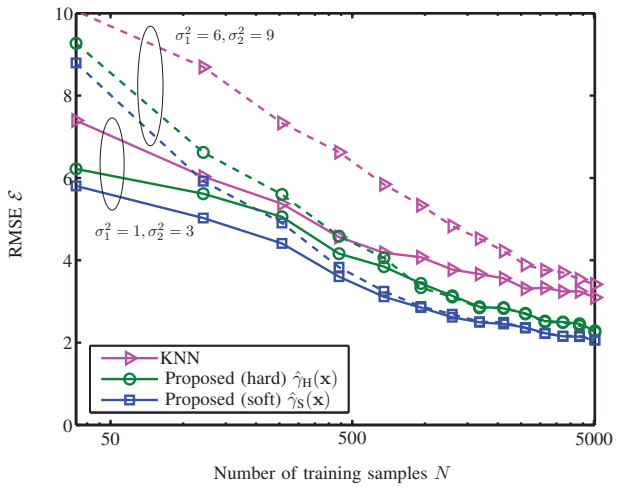

Figure 4. Reconstruction error versus number of training samples N.

We then evaluate the system performance measured by the root mean squared error (RMSE) defined as follows:

$$\mathcal{E} = \sqrt{\mathbb{E}\big\{\big(\hat{\gamma}(\mathbf{x}) - \bar{\gamma}(\mathbf{x})\big)^2\big\}} \tag{26}$$

where the expectation is evaluated empirically over 2.4×10^5 random test positions $\mathbf{x}$. The radio map reconstruction error

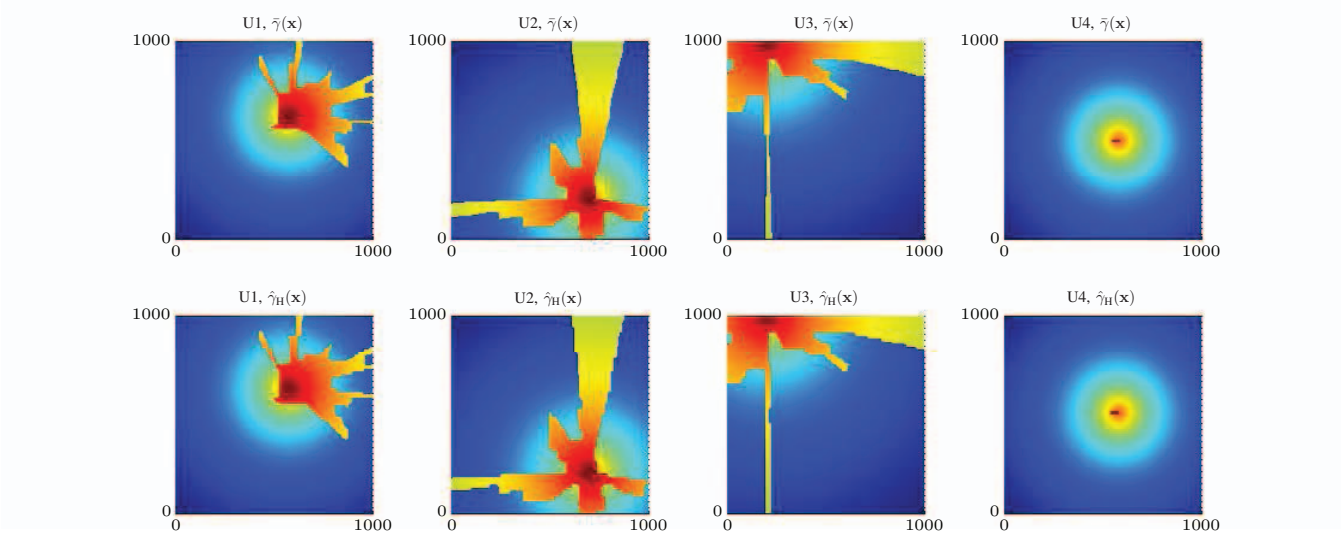

Figure 3. True radio maps $\bar{\gamma}(\mathbf{x})$ and the reconstructed radio maps $\hat{\gamma}_{\mathrm{H}}(\mathbf{x})$ for four different user locations. The propagation structures are clearly reconstructed.

from the proposed segmented regression algorithm is compared over the KNN baseline:

$$\hat{\gamma}_{\mathrm{KNN}}(\mathbf{x}) = \sum_{m \in \mathcal{N}(\mathbf{x})} \mathcal{K}(\mathbf{x}, \mathbf{x}^{(m)}) y^{(m)}$$

where $\mathcal{N}(\mathbf{x})$ defined in (21) is the index set of the M-nearest neighbor data samples of $\mathbf{x}$.

Fig. 4 shows the RMSE versus the number of training samples N over different observation noise variances σ_k in the channel model (1). It is observed that the RMSE decreases as the training size N increases. The proposed schemes achieve the same RMSE using only $1/10$ to $1/2$ training samples as required by the KNN baseline. In addition, the proposed schemes are hardly affected by observation noise, especially under large training sample size N. As a comparison, the performance of the KNN baseline degrades a lot in presence of large noise.

VI. Conclusion

This paper proposed a segmented regression approach to learn the radio map of the air-to-ground channel for applications in UAV-aided wireless communications. Based on the K-segment ray tracing model, a joint clustering and regression problem was formulated using the maximum likelihood approach. An iterative clustering and regression algorithm was developed to learn the segmented model from a few training samples. The radio map was then reconstructed using a kernel-based method. Numerical results demonstrated that the proposed method achieves the same reconstruction error using as few as $1/10$ training samples as required by the KNN baseline, and at the same time, is able to identify the propagation segments, which are useful for optimal UAV positioning.

Acknowledgment

This work was supported by the ERC under the European Union's Horizon 2020 research and innovation program (Agreement no. 670896).

References

[1] Y. Zeng, R. Zhang, and T. J. Lim, "Wireless communications with unmanned aerial vehicles: opportunities and challenges," *IEEE Commun. Mag.*, vol. 54, no. 5, pp. 36–42, 2016.

[2] D. H. Choi, B. H. Jung, and D. K. Sung, "Low-complexity maneuvering control of a UAV-based relay without location information of mobile ground nodes," in *Proc. IEEE Symposium on Computers and Commun.*, 2014, pp. 1–6.

[3] F. Jiang and A. L. Swindlehurst, "Optimization of UAV heading for the ground-to-air uplink," *IEEE J. Sel. Areas Commun.*, vol. 30, no. 5, pp. 993–1005, 2012.

[4] A. Al-Hourani, S. Kandeepan, and A. Jamalipour, "Modeling air-to-ground path loss for low altitude platforms in urban environments," 2014, pp. 2898–2904.

[5] M. Mozaffari, W. Saad, M. Bennis, and M. Debbah, "Drone small cells in the clouds: Design, deployment and performance analysis," in *Proc. IEEE Global Telecomm. Conf.*, 2015, pp. 1–6.

[6] A. Hourani, K. Sithamparanathan, and S. Lardner, "Optimal LAP altitude for maximum coverage," *IEEE Commun. Lett.*, no. 99, pp. 1–4, 2014.

[7] M. Mozaffari, W. Saad, M. Bennis, and M. Debbah, "Optimal transport theory for power-efficient deployment of unmanned aerial vehicles," *arXiv preprint arXiv:1602.01532*, 2016.

[8] E. Ostlin, H.-J. Zepernick, and H. Suzuki, "Macrocell path-loss prediction using artificial neural networks," *IEEE Trans. Veh. Technol.*, vol. 59, no. 6, pp. 2735–2747, 2010.

[9] R. D. Timoteo, D. C. Cunha, and G. D. Cavalcanti, "A proposal for path loss prediction in urban environments using support vector regression," in *Proc. Advanced Int. Conf. Telecommun*, 2014, pp. 1–5.

[10] D. Romero, S.-J. Kim, G. B. Giannakis, and R. Lopez-Valcarce, "Learning power spectrum maps from quantized power measurements," *arXiv preprint arXiv:1606.02679*, 2016.

[11] J. Chen and D. Gesbert, "Optimal positioning of flying relays for wireless networks: A LOS map approach," in *Proc. IEEE Int. Conf. Commun.*, Paris, France, May 2017, to appear.