

# Efficient Algorithms for Air-to-ground Channel Reconstruction in UAV-aided Communications

Junting Chen\*, Omid Esrafilian<sup>†</sup>, David Gesbert<sup>†</sup>, and Urbashi Mitra\*

\*Ming Hsieh Department of Electrical Engineering, University of Southern California, USA <sup>†</sup>Communication Systems Department, EURECOM, France

**Abstract**—This paper develops an efficient algorithm to learn and reconstruct from a small measurement samples an air-to-ground radio map with fine-grained propagation details so as to predict the signal strength between a wireless equipped UAV and arbitrary ground users, and ultimately the optimal position of the UAV as a mobile relay. In this paper, a joint data clustering and parameter estimation algorithm is developed to learn an multi-segment propagation model from energy measurements that may contain large observation noise. To reduce the reconstruction complexity, we propose to learn a hidden multi-class virtual obstacle model to help efficiently predict the air-to-ground channel. Numerical results demonstrate that the channel prediction error is significantly reduced, and meanwhile, the radio map reconstruction time is reduced to 1/300.

## I. INTRODUCTION

There is increasing interest exploiting unmanned aerial vehicles (UAVs) as flying relays to assist wireless communications in challenging propagation conditions. In particular, one of the most promising applications is to deploy low altitude UAVs in dense urban areas to improve the link quality of ground users that are in deep shadowing of tall buildings or vegetation. However, such an application is challenging because the UAV-user channel is not known before sending the UAV to the target position. Therefore, it is essential to learn the air-to-ground channel for every possible UAV-user location pair.

Most existing works on UAV-assisted wireless communications model the UAV-user channel gain as a deterministic function of the UAV-user distance [1] or a temporal and spatial correlated process from the prior measurements taken at nearby locations [2], [3]. Such macroscopic models may be applicable to long-range and high altitude UAVs, where the path loss, rather than the terrain shadowing, dominates channel gain. However, in cellular network applications, the presence of the urban architecture or vegetation may critically affect the performance. To address the shadowing issue, the works [4]–[7] established a simplified stochastic model to capture the line-of-sight (LOS) probability of the UAV-user link as a function of the elevation angle at the user and the distribution of the buildings. However, such a stochastic model does not provide specific performance guarantees.

It is not surprising that if the fine-grained air-to-ground propagation including deterministic shadowing effects is known, a better performance may be achieved [8], since the optimal UAV position can be adaptive to the actual shadowing

at the user side rather than merely based on a stochastic shadowing model.

To learn a fine-grained radio map, there are three types of methods considered in the literature: nonparametric methods such as kernel regression [9], semi-parametric methods such as segmented regression [10], [11], and methods based on 3D terrain maps [12], [13]. The kernel methods need to maintain a big table for all the measurement data, since it predicts the channel gain by combining the prior data measured at nearby locations [9]. Such a method is difficult to scale to a large environment, because the channel gain depends on both the UAV and user locations, which form a high-dimensional space. The segmented model [10], [11] exploits some parametric structure based on the log-distance propagation property. Specifically, it classifies the target area into several propagation segments and learns individual propagation laws for each segment. Although it makes better prediction than the kernel method, it stills need to store all the measurement data in order to performance propagation segment classification [11], and hence, it is difficult to scale. Map-based channel prediction approaches [12], [13] exploit sophisticated 3D channel models that take into account the signal diffraction at rooftops and reflection at walls [12], [13]; however, the prediction error can be as large as 10 – 20 dB compared to the measurement data as reported in [13]. In general, propagation is not a simple function of the shapes of the surrounding objects, but it also depends on whether the object is a concrete wall or a tree. Note that in the application of optimizing the UAV position, 10 dB difference in the channel gain can make a a big difference in the optimal UAV position. Moreover, up-to-date 3D city is usually not available. Therefore, it is highly motivated to develop an air-to-ground channel model to achieve better tradeoffs between complexity and precision.

This paper aims at developing efficient algorithms to learn a fine-grained air-to-ground channel for arbitrary UAV and user location pairs. Towards this end, we propose a segmented propagation model based on a hidden multi-class virtual obstacle model. Specifically, we reverse-engineer a multi-class virtual obstacle map based on the multiple propagation segments that are estimated from the sample data. Preliminary results have been developed in [14] for the LOS/NLOS two segment case. In the prediction phase, the air-to-ground channel is reconstructed using a direct path ray-tracing approach based on the estimated multi-class virtual obstacle map. As a result, the proposed model only needs to store the obstacle maps, which are 2D data arrays, together with several

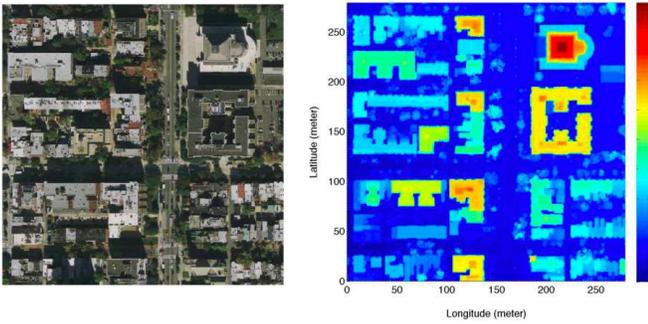


Figure 1. An urban city map captured at center Washington DC in USA, where the left image shows the orthoimagery of the area and the right image shows the building heights obtained from the LIDAR data taken at 2013 available at the USGS database.

propagation parameters. By contrast, the kernel methods and existing segmented approaches [9]–[11] need to store 6D UAV-user channel measurement data. Hence, the proposed method can be understood as projecting the high-dimensional data to dimension-reduced 2D data arrays under a hidden ray-tracing model. Our numerical results demonstrate significantly improved accuracy on radio map reconstruction and a 300-time run-time acceleration.

## II. SYSTEM MODEL

### A. Geographical Topology

Consider to learn the air-to-ground channel between a UAV flying above the ground and a user at the street level in an urban environment. An example topology is shown in Fig. 1.<sup>1</sup> The building heights and the bulk of the vegetation are *not* known.

This paper focuses on the long-term channel gain that captures the path loss and the shadowing, which depends on not only the UAV-user distance but also the objects surrounding the user, such as buildings and vegetation.

### B. Segmented Channel Model

Let  $\mathbf{x} = (\mathbf{x}_D, \mathbf{x}_U)$  be a UAV-user location pair, where  $\mathbf{x}_D, \mathbf{x}_U \in \mathbb{R}^3$  denote the location of the UAV and the user, respectively. Let  $d_L(\mathbf{x}) \triangleq 10 \log_{10} d(\mathbf{x})$  be the log-distance, where  $d(\mathbf{x}) = \|\mathbf{x}_D - \mathbf{x}_U\|_2$  is the UAV-user distance. Conventionally, a stochastic model of the channel gain in dB is given as  $\gamma(\mathbf{x}) = \beta + \alpha d_L(\mathbf{x}) + \xi(\mathbf{x})$ , where  $\alpha$  is the path loss exponent,  $\beta$  is the offset that can be measured as at the reference point  $d = 1$  meter (i.e.,  $d_L = 0$ ), and  $\xi(\mathbf{x})$  is modeled as a random variable that captures the global shadowing effect such as signal blockage, diffraction, and reflection.

From the measurement data in [10], [15], the random variable  $\xi$  usually has a high variance (up to 40 dB [10]). On the other hand, it is also observed that when the measurement location is restricted to a small area  $\mathbf{x} \in \mathcal{D}$ , the channel model  $\gamma(\mathbf{x})$  has a small variance for the random variable  $\xi(\mathbf{x})$  [10], [15]. For example, the measurement results in [10] show that, when one focuses on individual streets, the standard

<sup>1</sup>LIDAR (light detection and ranging) is a remote sensing method to measure ranges from an aerial plane to the ground.

deviation of  $\xi$  is roughly 2 dB. Such an observation motivates a segmented channel model as follows.

Let  $\mathbb{D} \subseteq \mathbb{R}^6$  be the domain of all possible UAV-user position pairs of interest. Consider a partition of  $\mathbb{D}$  into  $K + 1$  disjoint segments:  $\mathbb{D} = \mathcal{D}_0 \cup \mathcal{D}_1 \cup \dots \cup \mathcal{D}_K$ , where  $\mathcal{D}_k \cap \mathcal{D}_j = \emptyset$ , for  $k \neq j$ . The air-to-ground propagation is modeled as

$$\gamma(\mathbf{x}) = \sum_{k=1}^K (\beta_k + \alpha_k d_L(\mathbf{x}) + \xi_k) \mathbb{I}\{\mathbf{x} \in \mathcal{D}_k\} + \xi_0 \mathbb{I}\{\mathbf{x} \in \mathcal{D}_0\} \quad (1)$$

where  $\mathbb{I}\{A\}$  is an indicator function taking value 1 if  $A$  is true, and 0 otherwise. The random variables  $\xi_k$ ,  $k = 1, 2, \dots, K$ , are assumed to follow  $\mathcal{N}(0, \sigma_k^2)$  and capture the local shadowing effect. The variable  $\xi_0$  is modeled as  $\mathcal{N}(\beta_0 + \alpha_0 d_L(\mathbf{x}), \sigma_0^2)$  to capture the outliers (i.e., “background noise”) that cannot fit to either one of the propagation segments  $k = 1, 2, \dots, K$ . The motivation of the outlier segment  $\mathcal{D}_0$  is to keep the variance  $\sigma_k^2$  for propagation segments  $k = 1, 2, \dots, K$  small (typically less than 10 dB), while  $\mathcal{D}_0$  is also kept small (the variance  $\sigma_0^2$  is allowed to be large).

We aim at developing efficient algorithms to learn the parameters  $\alpha_k$ ,  $\beta_k$ , and  $\mathcal{D}_k$ , and to predict the channel gain at every possible UAV-user location pair  $\mathbf{x} \in \mathbb{R}^6$  using the following formula

$$\hat{\gamma}(\mathbf{x}) = \sum_{k=1}^K (\beta_k + \alpha_k d_L(\mathbf{x})) \mathbb{I}\{\mathbf{x} \in \mathcal{D}_k\} \quad (2)$$

with rare exceptions where  $\hat{\gamma}(\mathbf{x})$  is not defined in  $\mathbf{x} \in \mathcal{D}_0$ . Note that the smaller variances  $\sigma_k^2$ ,  $k = 1, 2, \dots, K$  are, the better precision  $\hat{\gamma}(\mathbf{x})$  in (2) can achieve. Intuitively, the more segments  $K$  can be characterized, the smaller segment variance  $\sigma_k^2$ , and the higher precision of the prediction model  $\hat{\gamma}(\mathbf{x})$ .

The learning and prediction are based on a set of channel measurements  $y^{(i)} = \gamma(\mathbf{x}^{(i)})$  at different locations  $\{\mathbf{x}^{(i)}, i = 1, 2, \dots, N\}$ .

*Remark 1 (Practical Scenarios that Motivate  $\mathcal{D}_0$ ).* Consider that the user locates at a street corner and the UAV locates at a position such that the direct propagation path touches the edge of a building. A slight change of the UAV or the user positions may result in the change of propagation segments according to (2), which leads to a significantly different predicted channel gain  $\hat{\gamma}(\mathbf{x})$ . For example,  $\hat{\gamma}(\mathbf{x})$  may equal to either  $\gamma_{\text{LOS}}(\mathbf{x}) = \beta_{\text{LOS}} + \alpha_{\text{LOS}} d_L(\mathbf{x})$  or  $\gamma_{\text{NLOS}}(\mathbf{x}) = \beta_{\text{NLOS}} + \alpha_{\text{NLOS}} d_L(\mathbf{x})$ . However, in practice, the measured channel gain  $\gamma(\mathbf{x})$  may equal to a value between  $\gamma_{\text{LOS}}$  and  $\gamma_{\text{NLOS}}$  due to energy from scattered paths. Therefore, classifying such a scenario to either pure LOS or pure NLOS case may lead to huge variance  $\sigma_k^2$  and large parameter estimation errors for  $\alpha_k$  and  $\beta_k$ . As a result, introducing an outlier region  $\mathcal{D}_0$  to kick out such rare scenarios may improve the robustness of model (1) in practice.

### C. Multi-class Virtual Obstacle Model

Note that the propagation segments  $\mathcal{D}_k$  are characterized by a set of location points in  $\mathbb{R}^6$ . Nevertheless, there is a hidden sparse structure of  $\mathcal{D}_k$  that can be exploited for reducing

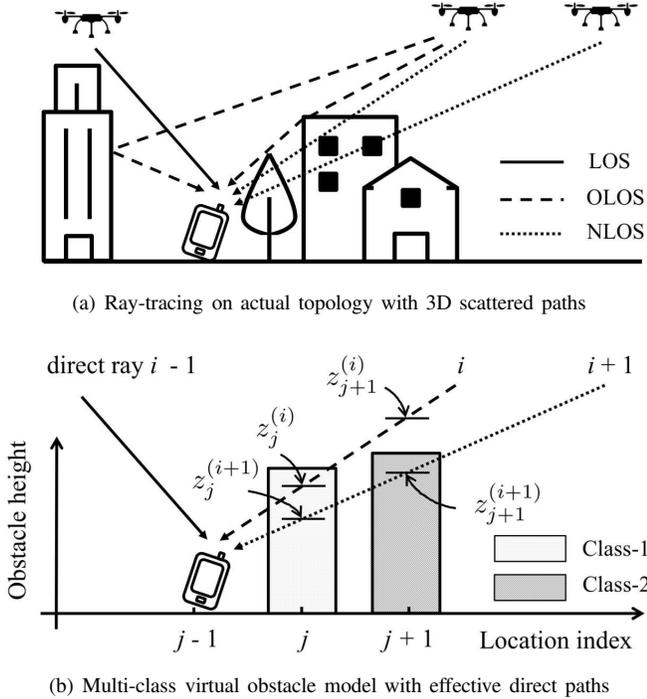


Figure 2. The virtual obstacle model in (b) approximates the propagation over all scattered paths in (a) by an equivalent direct path that passes through *effective* virtual obstacles. For example, the propagation from the second UAV in (a) (including reflection and diffraction) is approximated by an effective direct path which penetrates a light (wooden) wall located at the  $j$ th location in (b).

the model dimension. Specifically, we propose a multi-class virtual obstacle model as follows.

Without loss of generality (w.l.o.g.), assume that  $\mathcal{D}_1$  is the LOS segment, and the propagation in segment  $\mathcal{D}_k$  is less obstructed than that in  $\mathcal{D}_{k+1}$ , for  $k = 1, 2, \dots, K-1$ . Consider to discretize the terrain of interest into  $M$  equally spaced grids, and label the grids with integers  $j = 1, 2, \dots, M$ . For each location  $j$ , there may occupy one virtual obstacle from the  $K-1$  classes, where a class- $k$  obstacle is less obstructed than a class- $(k+1)$  obstacle,  $k = 1, 2, \dots, K-2$ .

Conceptually, the virtual obstacle model tries to approximate the propagation over all scattered paths (Fig. 2 (a)) by an equivalent direct path that passes through *effective* virtual obstacles (Fig. 2 (b)). Such a model makes it possible to reverse engineer a virtual obstacle topology based on the propagation segments  $\mathcal{D}_k$ ,  $k = 1, 2, \dots, K$ . Specifically, if the UAV-user location pair is in the LOS propagation segment,  $\mathbf{x} \in \mathcal{D}_1$ , there must be no obstacle blocking the *direct ray* from the UAV position  $\mathbf{x}_D$  to the user position  $\mathbf{x}_U$  as illustrated by the  $(i-1)$ th direct ray in Fig. 2 (b). Likewise, if  $\mathbf{x}$  belongs to a NLOS segment, there must be a wall blocking the direct ray as shown by the  $(i+1)$ th direct ray in Fig. 2 (b). Moreover, in the case of  $K = 3$  propagation segments, if  $\mathbf{x}$  is in an obstructed LOS (OLOS) segment,  $\mathbf{x} \in \mathcal{D}_2$ , it may suggest that a (virtual) light wall (or vegetation), instead of a concrete wall, may block the direct ray.

Mathematically, given  $\mathcal{D}_k$ ,  $k = 1, 2, \dots, K$ , the multi-class virtual obstacle model is defined as a set of vectors  $\mathbf{h}^{[k]}$ ,  $k =$

$1, 2, \dots, K-1$ , satisfying the following conditions:

- *Direct rays*: A direct ray  $\mathcal{R}(\mathbf{x})$ , that is a line segment joining a user position  $\mathbf{x}_U$  and a UAV position  $\mathbf{x}_D$  for location pair  $\mathbf{x} = (\mathbf{x}_U, \mathbf{x}_D)$ , belongs to class- $k$  if  $\mathbf{x} \in \mathcal{D}_k$ .
- *Virtual obstacle maps*: A virtual obstacle map of class- $k$  is an  $M$ -array vector  $\mathbf{h}^{[k]}$  with the  $j$ th entry representing the height of a class- $k$  obstacle at the corresponding  $j$ th grid location. Moreover, any class- $k$  direct ray  $\mathcal{R}(\mathbf{x})$ ,  $\mathbf{x} \in \mathcal{D}_k$  and  $k > 1$ , must penetrate at least one class- $(k-1)$  virtual obstacle, but it does not penetrate any virtual obstacles with class  $k' > k-1$ . A class-1 direct ray must not pass any obstacle. Here, a direct ray  $\mathcal{R}(\mathbf{x}^{(i)})$  is said to penetrate a virtual class- $k$  obstacle at location  $j$  if  $z_j^{(i)} < h_j^{[k]}$ , where  $z_j^{(i)}$  is the height above ground when the direct ray  $\mathcal{R}(\mathbf{x}^{(i)})$  passes over grid location  $j$ .

*Remark 2 (Physical Interpretation)*. The objective of the virtual obstacle model is *not* to estimate the actual terrain topology (such as that in Fig. 2 (a)), but to construct a virtual obstacle topology to fit the propagation segments  $\mathcal{D}_k$  under a simplified direct ray model as illustrated in Fig. 2 (b). For example, the diffracted and reflected in Fig. 2 (a) may result in classifying the location pair  $\mathbf{x}^{(i)}$  to an OLOS propagation segment (in  $K = 3$  case). Correspondingly, one may approximate such scenario by putting a wood wall at location  $j$  in Fig. 2 (b) to block the  $i$ th direct ray.

In summary, the goal of this paper is to learn the segment propagation model  $\{\alpha_k, \beta_k, \mathcal{D}_k, k = 1, 2, \dots, K\}$  based on some training measurements  $\{\mathbf{x}^{(i)}, y^{(i)}\}$ , estimate the virtual obstacle map  $\{\mathbf{h}^{[k]}, k = 1, 2, \dots, K-1\}$ , and based on these, reconstruct the air-to-ground channel for every possible UAV-user location  $\mathbf{x}$ .

### III. JOINT CLUSTERING AND REGRESSION WITH THE PRESENCE OF OUTLIERS

In this section, we perform joint cluster and regression to learn the segmented model with the presence of outliers from training energy measurements  $\{\mathbf{x}^{(i)}, y^{(i)} : i = 1, 2, \dots, N\}$ .

#### A. Identify the Outliers

Let  $\mathbf{c}^{(i)} = (c_0^{(i)}, c_2^{(i)}, \dots, c_K^{(i)})$  be the cluster label to estimate, where  $c_k^{(i)} = 1$  means that the data sample  $(\mathbf{x}^{(i)}, y^{(i)})$  is clustered to the  $k$ th segment, and  $c_k^{(i)} = 0$  otherwise. Note that  $c_k^{(i)} \in \{0, 1\}$  and  $\sum_{k=0}^K c_k^{(i)} = 1$ .

Given that  $c_k^{(i)} = 1$ , segmented channel model (1) yields

$$y = \alpha_k d_L(\mathbf{x}) + \beta_k + \xi_k \quad (3)$$

where  $\xi_k \sim \mathcal{N}(0, \sigma_k^2)$  holds for  $k = 0, 1, \dots, K$ . The joint probability density function (PDF)  $p_k(\mathbf{x}, y)$  for  $(\mathbf{x}^{(i)}, y^{(i)})$  conditioned on  $c_k^{(i)} = 1$  is thus given by

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

Note that the energy measurements tend to cluster to one of the  $K$  segments  $k = 1, 2, \dots, K$  with small variance  $\sigma_k^2$ , whereas, the outliers may deviate from any of the cluster

centers. Such property can be exploited to identify the outliers, and a simple algorithm can be designed by setting a threshold on the maximum deviation of the measured energy  $y^{(i)}$  from the segment center  $\alpha_k d_L(\mathbf{x}^{(i)}) + \beta_k$ , *e.g.*,  $|y - \alpha_k d_L(\mathbf{x}) - \beta_k| < 2\sigma_k$ . However, a more systematic way that can benefit from the Bayesian framework is to model the outlier segment using a Gaussian model  $\mathcal{N}(\alpha_0 d_L(\mathbf{x}) + \beta_0, \sigma_0^2)$  with large enough variance  $\sigma_0^2$  (a typical choice can be  $\sigma_0 = 20$ ), where the parameters  $\alpha_0$  and  $\beta_0$  are fit to the entire dataset, *i.e.*, as a least-squared solution to

$$\underset{\alpha_0, \beta_0}{\text{minimize}} \quad \sum_{i=1}^N (y^{(i)} - \alpha_0 d_L(\mathbf{x}^{(i)}) - \beta_0)^2.$$

As a result, an outlier  $(\mathbf{x}^{(i)}, y^{(i)})$  will appear to have small probabilities  $p_k$  for  $k = 1, 2, \dots, K$ , and relatively large probability  $p_0$  (since the propagation segments  $k = 1, 2, \dots, K$  are concentrated).

### B. Maximum Likelihood Estimation

To estimate the remaining parameters, the maximum likelihood approach derived in [11] can be directly applied. The overall solution is summarized as follows:

- Soft clustering: Let  $\theta^{(t)}$  be the parameter obtained from the  $t$ th iteration. The clustering result after the  $t$ th iteration is given by

$$\hat{c}_k^{(i)}(\theta^{(t)}) = \frac{p_k(\mathbf{x}^{(i)}, y^{(i)} | \theta^{(t)}) \pi_k}{\sum_{j=0}^K p_j(\mathbf{x}^{(i)}, y^{(i)} | \theta^{(t)}) \pi_j} \quad (5)$$

and in addition,  $\pi_k^{(t+1)} = \frac{1}{N} \sum_{i=1}^N \hat{c}_k^{(i)}$  for  $k = 0, 1, \dots, K$ .

- Parameter regression: For  $k = 1, 2, \dots, K$ ,

$$(\sigma_k^{(t+1)})^2 = \frac{\sum_{i=1}^N \hat{c}_k^{(i)} [y^{(i)} - \alpha_k^{(t+1)} d_L(\mathbf{x}^{(i)}) - \beta_k^{(t+1)}]^2}{\sum_{i=1}^N \hat{c}_k^{(i)}} \quad (6)$$

where

$$\begin{bmatrix} \alpha_k^{(t+1)} \\ \beta_k^{(t+1)} \end{bmatrix} = \mathbf{A}_k^{-1} \begin{bmatrix} \sum_{i=1}^N \hat{c}_k^{(i)} d_L(\mathbf{x}^{(i)}) y^{(i)} \\ \sum_{i=1}^N \hat{c}_k^{(i)} y^{(i)} \end{bmatrix}$$

and

$$\mathbf{A}_k = \begin{bmatrix} \sum_{i=1}^N \hat{c}_k^{(i)} d_L(\mathbf{x}^{(i)})^2 & \sum_{i=1}^N \hat{c}_k^{(i)} d_L(\mathbf{x}^{(i)}) \\ \sum_{i=1}^N \hat{c}_k^{(i)} d_L(\mathbf{x}^{(i)}) & \sum_{i=1}^N \hat{c}_k^{(i)} \end{bmatrix}.$$

## IV. ESTIMATION OF THE HIDDEN MULTI-CLASS VIRTUAL OBSTACLE MAP

To associate the propagation segments  $\mathcal{D}_k$ ,  $k = 1, 2, \dots, K$ , with the hidden virtual obstacle map, one needs to jointly consider the topology of all the buildings and the direct rays from all the UAV-user location pairs in the training dataset. In general, the association is not unique (due to finite

Table I  
RELATION BETWEEN PROPAGATION SEGMENTS AND THE LABEL VARIABLES UNDER THE CASE OF  $K = 3$  SEGMENTS.

Scenario	Segment	Segment label $\mathbf{c}^{(i)}$	Ray label $(L_1^{(i)}, L_2^{(i)})$
LOS	$\mathcal{D}_1$	(1,0,0)	(1,1)
OLOS	$\mathcal{D}_2$	(0,1,0)	(0,1)
NLOS	$\mathcal{D}_3$	(0,0,1)	(0,0)

training dataset) nor is it consistent (due to measurement and propagation segment classification errors). In this section, we first study the geometry characterization of the model, and then we propose a likelihood approach to find an estimate of the virtual obstacle map.

### A. Geometry Characterization

For the direct ray  $\mathcal{R}(\mathbf{x}^{(i)})$  from the  $i$ th UAV-user location pair, let  $L_k^{(i)}$ ,  $k = 1, 2, \dots, K - 1$ , be a binary variable to indicate whether the direct ray is in “virtual” LOS excluding the (light) blockage due to virtual obstacles below class  $k' < k$ . Specifically,

$$L_k^{(i)} = \begin{cases} 1 & \text{no virtual obstacle of class } k' \geq k \\ 0 & \text{some virtual obstacle of } k' \geq k. \end{cases}$$

Equivalently, it follows that

$$L_k^{(i)} = \sum_{m=1}^k c_m^{(i)} \quad (7)$$

which suggests that  $L_k^{(i)} = 1$  if and only if  $\mathbf{x}^{(i)}$  is in propagation segment  $\mathcal{D}_1 \cup \mathcal{D}_2 \cup \dots \cup \mathcal{D}_k$ , and  $L_k^{(i)} = 0$  if and only if  $\mathbf{x}^{(i)} \in \mathcal{D}_{k+1} \cup \mathcal{D}_{k+2} \cup \dots \cup \mathcal{D}_K$ .

For example, in Fig. 2 (b), the  $(i - 1)$ th direct ray has  $L_1^{(i-1)} = 1$ , which means there is not any obstacle blocking the  $(i - 1)$ th direct ray; for the  $i$ th direct ray,  $L_1^{(i)} = 0$ , which means there is some blockage, and  $L_2^{(i)} = 1$ , which means there is no (heavy) blockage from class-2 virtual obstacle or above; for the  $(i + 1)$ th direct ray,  $L_1^{(i+1)} = L_2^{(i+1)} = 0$  means there is blockage by virtual obstacles (as heavy as) class-2. Table I summarizes the relation between the ray labels  $L_k^{(i)}$  and the segment label  $\mathbf{c}^{(i)}$  under  $K = 3$  segment case.

Association can be made between the geometry of the direct rays and the topology of the virtual obstacles as follows. For each UAV-user position pair  $\mathbf{x}^{(i)}$ , denote  $\mathcal{B}^{(i)}$  as the set of grids that the ground projection of the direct ray  $\mathcal{R}(\mathbf{x}^{(i)})$  (see Section II-C) passes through. For each grid location  $j \in \mathcal{B}^{(i)}$ , denote  $z_j^{(i)}$  as the altitude above ground when the direct ray  $\mathcal{R}(\mathbf{x}^{(i)})$  passes over grid location  $j$  (see Fig. 2 (b)). It follows that

$$L_k^{(i)} = \mathbb{I}\{h_j^{[k]} \leq z_j^{(i)}, j \in \mathcal{B}^{(i)}\} \quad (8)$$

which means that if there is no virtual obstacle of class- $k$  blocking the  $i$ th direct ray, then the height of the obstacle must be less than the altitude of the direct ray for *all* corresponding grid locations  $j \in \mathcal{B}^{(i)}$ .

### B. A Likelihood Approach for Virtual Obstacle Estimation

From (5) and (7), an estimate of  $L_k^{(i)} \in \{0, 1\}$  can be given by  $\hat{L}_k^{(i)} = \mathbb{E}\{L_k^{(i)}\} = \mathbb{P}\{L_k^{(i)}\}$ , and

$$\mathbb{P}\{L_k^{(i)}\} = \sum_{m=1}^k \hat{c}_m^{(i)}. \quad (9)$$

Let  $\mathbf{z}^{(i)} = (z_1^{(i)}, z_2^{(i)}, \dots, z_M^{(i)})$  be the altitude vector, where  $z_j^{(i)} = H_{\max}$  for  $j \notin \mathcal{B}^{(i)}$  and  $H_{\max}$  is an upper bound of the virtual obstacle height. The likelihood function of the height vector  $\mathbf{h}^{[k]} = (h_1^{[k]}, h_2^{[k]}, \dots, h_M^{[k]})$  of class- $k$  virtual obstacles can be derived using Bayes' rule as

$$\begin{aligned} \mathbb{P}\{\mathbf{h}^{[k]} = \mathbf{z}\} &= \mathbb{P}\{\mathbf{h}^{[k]} = \mathbf{z} | \mathbf{h}^{[k]} \preceq \mathbf{z}^{(i)}\} \mathbb{P}\{\mathbf{h}^{[k]} \preceq \mathbf{z}^{(i)}\} \\ &+ \mathbb{P}\{\mathbf{h}^{[k]} = \mathbf{z} | \mathbf{h}^{[k]} \not\preceq \mathbf{z}^{(i)}\} \mathbb{P}\{\mathbf{h}^{[k]} \not\preceq \mathbf{z}^{(i)}\} \end{aligned}$$

where  $\mathbf{a} \preceq \mathbf{b}$  means  $a_j \leq b_j$  for all elements  $j$ , and  $\mathbf{a} \not\preceq \mathbf{b}$  means  $a_j > b_j$  for at least one element.

Note that  $\mathbb{P}\{\mathbf{h}^{[k]} = \mathbf{z} | \mathbf{h}^{[k]} \preceq \mathbf{z}^{(i)}\} = 0$  if  $\mathbf{z} \not\preceq \mathbf{z}^{(i)}$ . On the other hand, if  $\mathbf{z} \preceq \mathbf{z}^{(i)}$ ,  $\mathbb{P}\{\mathbf{h}^{[k]} = \mathbf{z} | \mathbf{h}^{[k]} \preceq \mathbf{z}^{(i)}\} \leq 1$ . Therefore,

$$\begin{aligned} \mathbb{P}\{\mathbf{h}^{[k]} = \mathbf{z} | \mathbf{h}^{[k]} \preceq \mathbf{z}^{(i)}\} &\leq \mathbb{I}\{\mathbf{z} \preceq \mathbf{z}^{(i)}\} \\ &= \mathbb{I}\{z_j \leq z_j^{(i)}, j \in \mathcal{B}^{(i)}\} \end{aligned}$$

where the probability function is upper bounded by a trust region specified by the indicator function  $\mathbb{I}\{\cdot\}$ . Similarly,  $\mathbb{P}\{\mathbf{h}^{[k]} = \mathbf{z} | \mathbf{h}^{[k]} \not\preceq \mathbf{z}^{(i)}\} = 0$  if  $\mathbf{z} \preceq \mathbf{z}^{(i)}$ , and it is upper bounded by 1 otherwise. Moreover, since  $\mathbb{P}\{\mathbf{h}^{[k]} \preceq \mathbf{z}^{(i)}\} = \sum_{m=1}^k \hat{c}_m^{(i)}$  due to (8) and (9), we have

$$\begin{aligned} \mathbb{P}\{\mathbf{h}^{[k]} = \mathbf{z}\} &\leq \mathbb{I}\{z_j \leq z_j^{(i)}, j \in \mathcal{B}^{(i)}\} \sum_{m=1}^k \hat{c}_m^{(i)} \\ &+ \left(1 - \mathbb{I}\{z_j \leq z_j^{(i)}, j \in \mathcal{B}^{(i)}\}\right) \left(1 - \sum_{m=1}^k \hat{c}_m^{(i)}\right). \end{aligned} \quad (10)$$

Applying (10) to all location pairs  $\{\mathbf{x}^{(i)}\}$  and clustering labels  $\{\hat{c}_m^{(i)}\}$ , we define an aggregate likelihood function as

$$\begin{aligned} &\frac{1}{N} \sum_{i=1}^N \mathbb{P}\{\mathbf{h}^{[k]} = \mathbf{z}\} \\ &\leq \frac{1}{N} \sum_{i=1}^N \left[ \mathbb{I}\{z_j \leq z_j^{(i)}, j \in \mathcal{B}^{(i)}\} \sum_{m=1}^k \hat{c}_m^{(i)} \right. \\ &\quad \left. + \left(1 - \mathbb{I}\{z_j \leq z_j^{(i)}, j \in \mathcal{B}^{(i)}\}\right) \left(1 - \sum_{m=1}^k \hat{c}_m^{(i)}\right) \right] \\ &\triangleq V_k(\mathbf{z}) \end{aligned}$$

and the estimated heights of the obstacles are given by

$$\hat{\mathbf{h}}^{[k]} = \underset{0 \leq z \leq H_{\max}}{\operatorname{argmax}} V_k(\mathbf{z}), \quad k = 1, 2, \dots, K-1. \quad (11)$$

An alternating algorithm can be applied to solve (11), where the elements  $h_j^{[k]}$  can be updated one by one while fixing all the other elements. As  $V_k(\mathbf{z})$  is a likelihood function,  $V_k(z_j; z_1, z_2, \dots, z_{j-1}, z_{j+1}, \dots, z_M)$  may first increase then

decrease, when one increases  $z_j$  from 0 to 1. Therefore, a bisection search type algorithm can be used to find the best  $z_j$  while fixing all the other elements. The process of alternating update is repeated until convergence. It is observed from our numerical experiment that the algorithm can usually converge after less than  $M$  round updates of  $\mathbf{z}$ .

The complexity of one round update of  $\mathbf{z}$  is  $\mathcal{O}(M^{1.5}N)$ . Specifically, to update each  $z_j$ , one needs to compute  $N$  summations, in which, there are at most  $|\mathcal{B}^{(i)}| \leq \sqrt{M}$  comparisons. Thus, the complexity to update one  $z_j$  is  $\mathcal{O}(M^{1/2}N)$ , where the constant depends on the implementation of a bisection search for  $z_j$ .

### C. Air-to-ground Channel Reconstruction

For any UAV-user location pair  $\mathbf{x} = (\mathbf{x}_D, \mathbf{x}_U)$ , let  $\mathcal{B}(\mathbf{x})$  be the set of grid locations  $j$  where the direct ray joining  $\mathbf{x}_D$  and  $\mathbf{x}_U$  passes over. Define  $z_j(\mathbf{x})$ ,  $j \in \mathcal{B}(\mathbf{x})$ , as the altitude when the direct ray passes over grid  $j$ . Similar to (8), define  $L_k(\mathbf{x}, \hat{\mathbf{h}}^{[k]}) = \mathbb{I}\{\hat{h}_j^{[k]} \leq z_j^{(i)}(\mathbf{x}), j \in \mathcal{B}(\mathbf{x})\}$  for  $k = 1, 2, \dots, K-1$ . From (7) and since  $c_k \in \{0, 1\}$ , one has

$$\mathbb{I}\{\mathbf{x} \in \mathcal{D}_k\} = c_k(\mathbf{x}) \approx \prod_{m=1}^{k-1} (1 - L_m(\mathbf{x}, \hat{\mathbf{h}}^{[m]})) \prod_{m=k}^{K-1} L_m(\mathbf{x}, \hat{\mathbf{h}}^{[m]}).$$

Using the regression parameters  $\hat{\alpha}_k$  and  $\hat{\beta}_k$  obtained from Section III, the air-to-ground radio map  $\hat{\gamma}(\mathbf{x})$  can be reconstructed as

$$\sum_{k=1}^K (\hat{\beta}_k - \hat{\alpha}_k d_L(\mathbf{x})) \prod_{m=1}^{k-1} (1 - L_m(\mathbf{x}, \hat{\mathbf{h}}^{[m]})) \prod_{m=k}^{K-1} L_m(\mathbf{x}, \hat{\mathbf{h}}^{[m]}). \quad (12)$$

## V. NUMERICAL RESULTS

The urban map for numerical experiment is shown in Fig. 1. The maximum building height in the area is 40 meter. The true radio map is constructed via  $K = 3$  propagation segments based on whether there is a building or vegetation in the way of the direct signal path from the orthoimagery and the building/vegetation heights of the area. The propagation parameters are chosen as  $(\alpha_1, \beta_1, \sigma_1^2) = (-2.2, -28, 1)$ ,  $(\alpha_2, \beta_2, \sigma_2^2) = (-2.8, -24, 2)$ , and  $(\alpha_3, \beta_3, \sigma_3^2) = (-3.6, -22, 3)$ . In the training phase, the UAV is deployed at 400 locations in a uniform pattern with a fixed 50 meter altitude, where at each location, the UAV measures signal strength from 100 uniformly random user locations on the streets.

To reconstruct the air-to-ground radio map at UAV-user location  $\mathbf{x}$ , the KNN baseline computes  $\hat{\gamma}_{\text{KNN}}(\mathbf{x}) = \sum_{m \in \mathcal{N}(\mathbf{x})} \mathcal{K}(\mathbf{x}, \mathbf{x}^{(m)}) y^{(m)}$ , where  $\mathcal{N}(\mathbf{x})$  is the set of 5 nearest points to  $\mathbf{x}$  from the training set  $\{\mathbf{x}^{(i)}\}$ . In addition,  $\mathcal{K}(\mathbf{x}, \mathbf{x}') = \lambda \exp\{-\|\mathbf{x} - \mathbf{x}'\|_2^2 / (2s(\mathbf{x})^2)\}$ , in which  $s(\mathbf{x})^2$  is the variance of  $\|\mathbf{x} - \mathbf{x}'\|_2$  for  $\mathbf{x}' \in \mathcal{N}(\mathbf{x})$  and  $\lambda$  is a normalizing factor such that  $\sum_{m \in \mathcal{N}(\mathbf{x})} \mathcal{K}(\mathbf{x}, \mathbf{x}^{(m)}) = 1$ . The direct reconstruction baseline [11] first determines the propagation segment of  $\mathbf{x}$  using KNN and then applies the corresponding regression function in (2). The proposed reconstruction is obtained from (12).

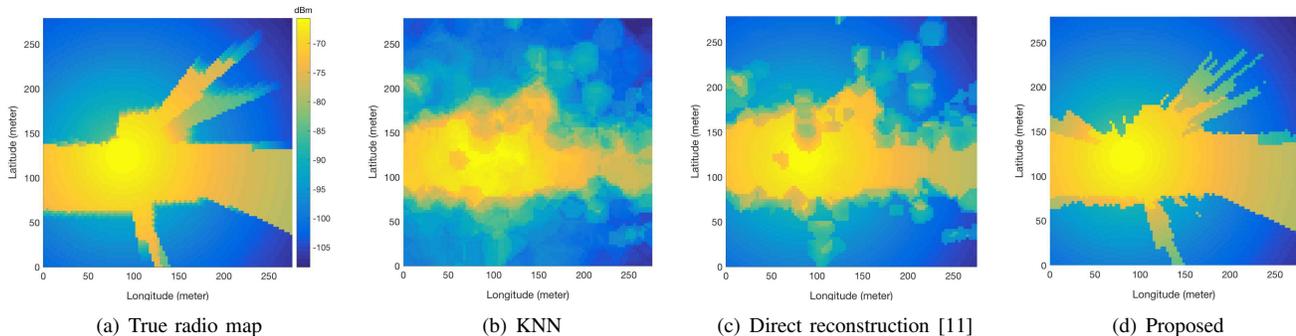


Figure 3. Air-to-ground radio map reconstruction for every UAV location with fixed altitude 50 meters to a user locating at (87, 123, 0) in meters.

Table II  
RUN TIME AND RECONSTRUCTION ERROR PERFORMANCE

	Run-time (sec)	Mean absolute error (dB)	
		$H_{UAV} = 50$ (m)	$H_{UAV} = 70$ (m)
KNN	463	3.6	5.2
Direct [11]	465	3.2	5.0
Proposed	1.50	2.2	2.6

Fig. 3 shows the reconstruction for a user locating at (87, 123, 0) in meters and a UAV at 50 meter altitude. The proposed approach is able to roughly reconstruct the fine-grained propagation structure due to different types of blockage with accurate prediction on the channel gain inside each propagation segment. Table II compares the run-time for constructing a  $93 \times 94$  radio map and the mean absolute error averaging over 100 different user locations. The proposed method reduces the run-time by 300 times and also significantly reduces the mean absolute error of channel prediction.

## VI. CONCLUSION

This paper develops an efficient algorithm to learn and reconstruct the air-to-ground radio channel that captures the fine-grained propagation structure. First, a robust joint clustering and regression algorithm is developed to cluster the training locations to different propagation segments. Then, a hidden multi-class virtual obstacle model, which approximates the air-to-ground propagation by a set of direct rays passing through effective virtual obstacles, is estimated. Finally, the air-to-ground channel is reconstructed for every UAV-user location pair using direct path ray-tracing analysis based on the estimated multi-class virtual obstacle map. Our numerical results demonstrate significantly reduced channel prediction errors and 300X acceleration in radio map reconstruction.

## ACKNOWLEDGMENT

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

## REFERENCES

- [1] 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.
- [2] A. J. Carfang and E. W. Frew, "Real-time estimation of wireless ground-to-air communication parameters," in *Proc. Int. Conf. on Computing, Networking and Commun.*, 2012, pp. 975–979.
- [3] A. Ghaffarkhah and Y. Mostofi, "Channel learning and communication-aware motion planning in mobile networks," in *Proc. American Control Conf.*, 2010, pp. 5413–5420.
- [4] A. Al-Hourani, S. Kandeepan, and A. Jamalipour, "Modeling air-to-ground path loss for low altitude platforms in urban environments," in *Proc. IEEE Global Telecomm. Conf.*, 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. Sithamparamathan, 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," in *Proc. IEEE Int. Conf. Commun.*, Kuala Lumpur, Malaysia, May 2016, pp. 1–6.
- [8] J. Chen and D. Gesbert, "Joint user grouping and beamforming for low complexity massive MIMO systems," in *IEEE Int. Workshop on Signal Process. Advances in Wireless Commun.*, Edinburg, UK, May 2016, pp. 1–6.
- [9] D. Romero, S.-J. Kim, G. B. Giannakis, and R. Lopez-Valcarce, "Learning power spectrum maps from quantized power measurements," *IEEE Trans. Signal Process.*, vol. 65, no. 10, pp. 2547–2560, 2017.
- [10] K. T. Herring, J. W. Holloway, D. H. Staelin, and D. W. Bliss, "Path-loss characteristics of urban wireless channels," *IEEE Trans. on Antennas and Propagation*, vol. 58, no. 1, pp. 171–177, 2010.
- [11] J. Chen, U. Yatnalli, and D. Gesbert, "Learning radio maps for UAV-aided wireless networks: A segmented regression approach," in *Proc. IEEE Int. Conf. Commun.*, Paris, France, May 2017.
- [12] E. Tameh, A. Nix, and M. Beach, "A 3-D integrated macro and microcellular propagation model, based on the use of photogrammetric terrain and building data," in *Proc. IEEE Semiannual Veh. Technol. Conf.*, vol. 3, 1997, pp. 1957–1961.
- [13] J. F. Monserrat, S. Inca, J. Calabuig, and D. Martín-Sacristán, "Map-based channel model for urban macrocell propagation scenarios," *Int. J. of Antennas and Propagation*, vol. 2015, 2015.
- [14] O. Esrafilian and D. Gesbert, "3D city map reconstruction from UAV-based radio measurements," in *Proc. IEEE Global Telecomm. Conf.*, Singapore, Dec. 2017, to appear.
- [15] Q. Feng, J. McGeehan, E. K. Tameh, and A. R. Nix, "Path loss models for air-to-ground radio channels in urban environments," in *Proc. IEEE Semiannual Veh. Technol. Conf.*, vol. 6, 2006, pp. 2901–2905.