Network-side Localization via Semi-Supervised Multi-point Channel Charting

Junquan Deng\textsuperscript{1}, Olav Tirkkonen\textsuperscript{2}, Jianzhao Zhang\textsuperscript{1}, Xianlong Jiao\textsuperscript{3} and Christoph Studer\textsuperscript{4}

\textsuperscript{1}Sixty-third Research Institute, National University of Defense Technology, China, Emails: {jqdeng, jianzhao0633}@nudt.edu.cn
\textsuperscript{2}Department of Communications and Networking, Aalto University, Finland, Email: olav.tirkkonen@aalto.fi
\textsuperscript{3}College of Computer Science, Chongqing University, China, Email: xjiao@cqu.edu.cn
\textsuperscript{4}Department of Information Technology and Electrical Engineering, ETH Zurich, Switzerland, Email: studer@ethz.ch

Abstract—We consider the network-side mobile localization problem in future 5G and beyond wireless networks with distributed multi-antenna base stations (BSs). For this application, we propose a semi-supervised multi-point channel charting (SS-MPCC) framework, which consists of (i) collaborative collection of channel state information (CSI) and other side-information by distributed BSs; (ii) local CSI feature extraction and self-learning of a dissimilarity metric, and (iii) global graph construction and constrained manifold learning. We show that side-information from routine network operations, including timestamps, channel qualities, and a small set of labeled samples, can be exploited to construct a consistent global graph. The graph is then mapped to a 2D channel chart using constrained manifold learning for localization purposes. We evaluate the performance of SS-MPCC in a simulated urban outdoor scenario with realistic user motion. Our results show that SS-MPCC achieves a mean localization error of 5.6 m with only 10\% of labeled CSI samples. SS-MPCC does not require accurate synchronization among multiple BSs and is promising for future cellular localization.

Index Terms—MIMO CSI, channel charting, semi-supervised learning, network-side localization

I. INTRODUCTION

Mobile location information acquisition is fundamental in building smart cites and intelligent transportation systems. However, how to realize seamless high-precision localization, especially on the network side, is a challenging problem in GPS-denied environments. Future 5G and beyond radio access networks (RANs) will be densely deployed with massive multiple-input multiple output (mMIMO) Base Stations (BSs) and high-frequency carrier frequencies in order to provide ubiquitous ultra-fast and reliable wireless services. In addition to communication, high-resolution sensing functions are envisioned to be integrated in such RANs, which opens up new opportunities for high-precision positioning and tracking [1]–[4]. Along these lines, 3GPP Rel-16 has specified positioning support for 5G New Radio (NR) [5]. However, current network-side NR positioning techniques are based on triangulation with Angle-of-Arrival (AOA) or trilateration with Time-Difference-of-Arrival (TDOA) measurements at multiple BSs, which require rigorous calibration or accurate synchronization among network elements. Moreover, the performance of such methods degrades in complex multi-path and obstructed environments. Fingerprinting techniques [6]–[8] with machine learning could be applied for cellular positioning in such challenging propagation environments. However, they require a large number of densely-sampled channel measurements, which scales only poorly to large areas and renders automation to dynamic environments challenging. As unlabeled samples are easy to collect, semi-supervised learning (SSL) methods [9] utilizing both labeled and unlabeled data are promising for building a large-scale automatic network-side positioning system. SSL has attracted increasing attention for indoor positioning applications based on Received Signal Strength Indicator fingerprints [10], [11]. However, utilizing mMIMO Channel State Information (CSI) [12], [13] in a networking communication environment for high-precision and seamless positioning and tracking remains a challenging problem.

Channel charting (CC), as put forward in [14], relies on using sporadically collected multi-antenna CSI from unknown locations in order to learn a chart that can be used for applications that do not require absolute position information (e.g., hand-over and user grouping). The principle of CC is to exploit the fact that high-dimensional CSI strongly depends on User Equipment (UE) location, which is low-dimensional. To equip CC with absolute positioning capabilities, a semisupervised autoencoder utilizing a subset CSI of samples with known spatial locations, as well as mobility side information, has been proposed in [15]; this approach does not require dense measurements. More recently, an unified CC architecture based on Sammon’s mapping (SM) and Siamese networks has been presented in [16]; this approach can be used for both supervised and semisupervised positioning, and unsupervised relative localization. In [17], CC is performed for radio signals collected in a real environment from transmissions of a 64-antenna base station. A triplet neural network learns the mapping from CSI samples to CC-coordinates, using time stamp side information to classify pairs of samples to come from nearby or remote spatial locations. Unfortunately, in a semi-supervised setting with challenging non-Line-of-Sight (NLOS) channels and large-scale network coverage areas, the methods of [15], [16] are not capable of exploiting unlabeled data samples to improve
Reference Signals (SRSs), to BSs for channel estimation. An 
etwork operation, the UEs send pilot signals, e.g., Sounding 
moves in the coverage area of multiple mMIMO BSs. During 
A. CSI and Side-information Collection

unit (CU).

manifold learning procedures performed at a centralized 
construction procedures at distributed units (DUs), as well 
procedure, CSI feature extraction and local dissimilarity matrix 

street grid, we show that a small set of CSI measurements 
perform simulations in an urban outdoor multi-cell mmWave 
points in the chart represent the true geographical mobile 
calibrate the multi-point channel chart, so 
chart. Specifically, we use graph-based 
location information of a few labeled samples to increase 
aligned channel chart in which historical and real-time mobile 
distributed multi-antenna BSs collaboratively collect wireless 

Fig. 1. Graph-based semi-supervised multi-point channel charting framework.

location estimates.

Multipoint channel charting (MPCC) [18], [19] is an 
the distortion occurring in single-point channel charting results. In 
this paper, we propose a semi-supervised multi-point 
wireless data from spatio-temporal mobile locations and learn an 
where CSI available at multiple BSs and by 
exploiting redundancy in multi-point CSI to combat the 
channel charting (SS-MPCC) framework for large-scale 
channels can be accurately matched and determined. We 
use timestamp information for the unlabeled samples, and 
location information of a few labeled samples to increase 
method to calibrate the multi-point channel chart, so 
the chart represent the true geographical mobile 
channels can be accurately matched and determined. We 
used graph-based constrained manifold learning methods to 
that points in the chart represent the true geographical mobile 
channel chart. Specifically, we use graph-based constrained manifold learning methods 

to generalize the model to a situation with unlabeled samples 

II. SS-MPCC FRAMEWORK

We now introduce the framework for semi-supervised 
our framework is depicted in 
Figure 1 and consists of a CSI and side-information collection 
procedure, CSI feature extraction and local dissimilarity matrix 
construction procedures at distributed units (DUs), as well 
as global graph construction and semi-supervised constrained 
manifold learning procedures performed at a centralized 
unit (CU).

A. CSI and Side-information Collection

We consider a typical 5G cellular network in which UEs 
moves in the coverage area of multiple mMIMO BSs. During 

channel estimation. An 
estimated mMIMO channel vector at time $t$ on a sub-carrier 
with frequency $f$ is modelled as

$$
\hat{h}_{t,f} = \sum_{l=1}^{P} \alpha_{t,f}^{(l)} s(\phi_{t}^{(l)}) + \mathbf{n},
$$

where $P$ is the number of multi-path components, $\phi_{t}^{(l)}$ is the 
impinging direction-of-arrival (DOA) of the $l$th propagation path, and $\alpha_{t,f}^{(l)}$ is a random complex gain for the $l$th path, which depends on the transmit power, path delays, reflection coefficients of the contributing scatters, and impacts of radio 
channel components. In addition, $\mathbf{n}$ is the estimation error and the vector $s(\phi)$ represents the array steering vector (ASV). For an $M$-antenna uniform linear array with the plane-wave approximation, the ASV is

$$
s(\phi) = \left[1, e^{j\frac{2\pi}{\lambda} s \sin(\phi)}, \ldots, e^{j(M-1)\frac{2\pi}{\lambda} s \sin(\phi)} \right]^T,
$$

where $\lambda$ is the carrier wavelength and $s$ the antenna spacing.

The channel vector $\hat{h}_{t,f}$ changes rapidly when the UE 
moves in length scales of the wavelength $\lambda$. This is due 
to changes in the phases of the complex path gains of the 
multipath components. In comparison, the DOAs $\{\phi_{t}^{(l)}\}_{l=1}^{L}$ 
and the powers of $\{\alpha_{t,f}^{(l)} \}_{l=1}^{L}$ change more slowly and smoothly. 
To exploit this property, we use the frequency-domain 
covariance $C = \mathbb{E}_f [\hat{h}_{t,f} \hat{h}_{t,f}^H]$ as the CSI sample for channel 
charting, as it depends on the multipath DOAs and the powers. 
In practice, the covariance can be estimated across multiple 
coherece bandwidths, e.g., different OFDM sub-carriers.

In the considered system, there are two sets of CSI data, 
labeled and unlabeled samples. The first set consists of the 
CSI data with corresponding position coordinates which 
can be gathered in a dedicated site survey, whereas the 
second set is collected from multiple UEs on move. While 
collecting the unlabeled CSI samples, the timestamp and 
UE identifier are also recorded by the BSs, as in [15], [17]. 
We assume there are $B$ BSs. The $L$ different labeled 

B. CSI Feature Extraction

We use the power angular profile (PAP) of multipath 
components (MPCs) as the CSI feature as in [18]. We use 
a Multiple Signal Classification (MUSIC) algorithm [20] to 
extract the PAP feature from a covariance CSI sample $C$. 
The powers of multiple components are proportional to the 
eigenvalues of the covariance matrix. Noticing this, we first 
compute the eigen-decomposition of $C$ and estimate the 
number $\hat{P}$ of prominent MPCs based on the descending
distribution of eigenvalues \( \{\lambda_1, \lambda_2, \ldots, \lambda_M\} \). In practice, \( \hat{P} \) must be smaller than the number \( M \) of BS antennas, and can be estimated as the minimum \( \hat{P} \) satisfying \( \lambda_1 + \cdots + \lambda_{\hat{P}} > \rho \), where \( \rho \) is a target power ratio, which is set as 0.8 in our experiments in Section IV. Then the MUSIC algorithm is applied to find \( \hat{P} \) DOAs \( \{\phi_1, \ldots, \phi_P\} \) in its pseudo-spectrum. Finally, the PAP feature is represented as DOA–power pairs stacked in a vector as \( f = [\lambda_1, \ldots, \lambda_P, \phi_1, \ldots, \phi_P] \).

C. Local CSI Dissimilarity Learning

Given a set of CSI features \( \{f_{n}^{(b)}\}_{b=1}^{L+U} \) extracted from the labeled CSI samples \( L^{(b)} \) and unlabeled \( U^{(b)} \) at a local BS \( b \), we need a method to measure their dissimilarity. As shown in Figure 2, each MPC corresponds to a virtual transmission (Tx) point [21] in the radio propagation environment. Spatially close mobile locations will produce similar virtual Tx points, as they share common scatterers and have similar MPCs. The coordinates of these virtual Tx points can be deduced from the PAP features. For a general multipath PAP feature \( f = [\lambda_1, \ldots, \lambda_P, \phi_1, \ldots, \phi_P] \), its \( \hat{P} \) coordinates with respect to the local BS is given by

\[
\mathcal{F}(f) = [x_1, \ldots, x_n],
\]

where \( x_i = \rho(\lambda_i) [\lambda_i^{-\frac{1}{2}} \cos \phi_i, \lambda_i^{-\frac{1}{2}} \sin \phi_i]^T \) is the \( i \)th virtual Tx point, and \( \rho(\cdot) \) is a scaling function compensating the scattering loss. We then transform the PAP features \( \{f_{n}^{(b)}\}_{b=1}^{L+U} \) to a point cloud, which consists of all the virtual Tx points with respect to MPCs observed by the BS during CSI sample collection. As depicted in Figure 2, this point cloud exhibits clusters, which reflects distributions of mobile UEs and the radio scatterers in the environments. The dissimilarity between two CSI samples depends on whether their PAP features have virtual points that belong to the same cluster. If two samples share virtual points in a same cluster, one should estimate their distance based on these similar virtual points. For this, take two features \( f_{n}^{(b)} \) and \( f_{m}^{(b)} \) from the feature set, assume that their virtual points are \( [x_{1,m}, \ldots, x_{P,m}] \) and \( [x_{1,n}, \ldots, x_{Q,n}] \) of the form (3). Assume that there are a total of \( C \) clusters in the virtual Tx point cloud. After clustering, a virtual Tx point \( x \) would have a cluster label \( l(x) \in \{1, \ldots, C\} \). Denote the set of clusters by \( C_{m,n} \) where both \( f_{m}^{(b)} \) and \( f_{n}^{(b)} \) have at least one virtual point in each of them. If \( C_{m,n} \) is not empty, then we search a pair of \( \{x_{i,m}, x_{j,n}\} \) from \( [x_{1,m}, \ldots, x_{P,m}] \) and \( [x_{1,n}, \ldots, x_{Q,n}] \) such that \( \max_{i,j} \{|x_{i,m} - x_{j,n}|^2 - |x_{i,n} - x_{j,m}|^2\} \) is minimal and \( l(x_{i,m}) = l(x_{j,n}) \). Then, the distance \( |x_{i,m} - x_{j,n}|^2 \) is used as the dissimilarity metric. If \( C_{m,n} \) is empty, then the distance \( |x_{i,m} - x_{j,n}|^2 \) between the points \( x_{i,m} \) and \( x_{j,n} \), which correspond to the strongest MPCs, is used as the dissimilarity metric.

Using above CSI feature extraction method and the self-learned dissimilarity metric, the local BS can construct a matrix \( D^{(b)} \in \mathbb{R}^{(L+U) \times (L+U)} \) for the CSI samples \( L^{(b)} \) and \( U^{(b)} \), whose element \( D_{i,j}^{(b)} \) captures the dissimilarity between \( C_{i}^{(b)} \) and \( C_{j}^{(b)} \), for \( i, j = 1, \ldots, L + U \).

The CSI feature extraction and dissimilarity learning procedures are performed at the DUs which reside at the network edge close to BSs. Then, the local dissimilarity matrix of all CSI samples and the corresponding timestamp and channel quantities are reported to a logical CU for multi-point channel charting. We thus avoid exchange of the high-dimensional channel covariances between DUs and CU. The CU then fuses the data collected from distributed BSs to form a channel chart via global graph construction and semi-supervised constrained manifold learning, as discussed in Section III.

III. Graph-based SS-MPCC

In [18], manifold-based mapping algorithms, including Sammon’s Mapping (SM), t-Distributed Stochastic Neighbor Embedding (t-SNE) and Laplacian Eigenmaps (LE), were considered for unsupervised MPCC, which can learn the relative position between unlabeled CSI samples. Here, we choose LE the baseline dimensional reduction algorithm, and adapt it for semi-supervised learning. Conventional SM and t-SNE can also be modified to utilize side information. For example, a parametric version of SM using Siamese networks has been considered in [16] for semi-supervised single-point channel charting. However, performing multi-point channel charting with such methods is an open research problem.

LE consists of neighborhood graph construction, weight matrix computation, and low-dimensional mapping procedures. We will show that incorporation of side information and merging of local dissimilarity matrices can be realized in the first two procedures. For application in SS-MPCC, we first construct a graph \( G = (\mathcal{V}, \mathcal{E}) \), where the sets of vertices \( \mathcal{V} \) and edges \( \mathcal{E} \) represent the CSI samples collected from the \( L + U \) locations, and the corresponding neighborhood relations.

The objective of SS-MPCC is to map the \( L + U \) high-dimensional CSI samples into a low-dimensional space with a representational channel chart \( Z = [z_1, \ldots, z_{L+U}] \), so that \( z_n \) should be close to the true mobile location \( p_n \), for all \( n = 1, \ldots, L + U \). It can be addressed using constrained manifold learning methods [22], [23], which is prevalent in machine learning for exploring the underlying low-dimensional structure of the whole dataset with both labeled and unlabeled data. As the first \( L \) locations are known, we set \( \sum_{n=1}^{L} \|z_n - p_n\|^2 \) as a cost function which should be minimized. Furthermore, to preserve the neighborhood relationships of nodes in \( \mathcal{V} \) deduced from CSI dissimilarities, we set \( \sum_{m,n} w_{m,n} \|z_m - z_n\|^2 \) as another cost function, where
$w_{m,n}$ is the edge weight for nodes $m$ and $n$. The overall cost function is therefore given by

$$f(Z) = \sum_{n=1}^{L} \| z_n - p_n \|^2 + \gamma \sum_{m=1, n=1}^{L+U} w_{m,n} \| z_m - z_n \|^2. \quad (4)$$

Here, $\gamma \geq 0$ is a tradeoff coefficient between two objectives. The first term in (4) represents the alignment error for labeled samples, while the second term preserves local smoothness of the channel chart. For $\gamma = 0$, only the anchor points will be used; for $\gamma \to \infty$, only CSI is used and anchors are ignored.

### A. Global Graph Construction

To construct the global graph $\mathcal{G}$ based on the $B$ local dissimilarity matrices $\{ D^{(b)} \}_{b=1,...,B}$ and side information $[t_1, ..., t_U]$ and $P = [p_1, ..., p_L]$, we proceed as:

**Step-1:** Merge the local dissimilarity matrices to form a globally consistent one as in [18]. In practice, the signal from a sampled location $p_n$ may not be received by a BS $b$. In this case, we set $f_n = \emptyset$, and $D^{(b)}_{m,m}, D^{(b)}_{m,n}$ for all $m \neq n$ to be a large value, and set the channel quality $\gamma^{(b)}_n$ be zero. The cell-specific dissimilarities are then fused into a global multi-cell dissimilarity matrix $D$ with elements

$$D_{m,n} = \left( \sum_{b=1}^{B} v_b \right)^{-1} \sum_{b=1}^{B} v_b D_{m,m}^{(b)}, \quad (5)$$

where $v_b$ is a reliability weight. We consider using $v_b = [\min(\gamma^{(m)}_n, \gamma^{(m)}_n)]^\zeta$, where $\{ \gamma^{(b)}_n \}$ are channel qualities between locations $\{ p_j \}$ and BS $b$, estimated from RSRP or other measurements. In addition, The parameter $\zeta$ is a weighting exponent. MPCC is rather robust against changing the weighting exponents, and we set $\zeta = 2$ in this paper.

**Step-2:** Find $k_e$ nearest neighbors based on the global multi-cell dissimilarity matrix $D$ for each node $m$ in $\mathcal{G}$. If node $n$ is in the set of $k_e$ nearest neighbors of node $m$, nodes $m$ and $n$ are connected in the graph.

**Step-3:** If the timestamps $t_m, t_n$ of two nodes $m, n \in \{L + 1, \ldots, L + U\}$ with unlabeled CSI samples satisfy $| t_m - t_n | < T_{th}$, connect nodes $m$ and $n$ in $\mathcal{G}$. Here $T_{th}$ is the threshold for timestamps which is related to the movement speeds of UEs, It will be discussed in Section IV.

**Step-4:** Based on the neighborhood graph constructed above, compute the weight matrix $W$ elements $\{ w_{m,n} \}_{L+U}$. Denote $A = \{1, \ldots, L\}$ and $B = \{ L + 1, \ldots, L + U \}$. If the nodes $m$ and $n$ are connected, then the weight $w_{m,n}$ is

$$w_{m,n} = \begin{cases} \frac{-D_{m,n}}{\sigma_m^2} + (1 - \alpha) \frac{-|t_m - t_n|}{\sigma_t^2}, & \text{if } m, n \in A, \\ \frac{-D_{m,n}}{\sigma_m^2} + (1 - \alpha) \frac{-|t_m - t_n|}{\sigma_t^2}, & \text{if } m, n \in B, \\ e^{-\frac{D_{m,n}}{\theta_d}^2}, & \text{otherwise}. \end{cases} \quad (6)$$

Here, $\alpha$ is a tradeoff coefficient for mixed edge weights, and $\theta_s, \theta_t$, and $\theta_d$ are the kernel parameters for CSI dissimilarities, timestamps and location coordinates respectively. The parameters $\alpha, \theta_s, \theta_t$, and $\theta_d$ will be investigated in Section IV. Finally, if the nodes $m$ and $n$ are not connected, set $w_{m,n} = 0$. When $k_e$ is much smaller than the number of nodes in the graph, the resulting graph would typically have a sparse weighted adjacency matrix $W$.

It is straightforward to modify (6) to incorporate only position or timestamp information. For example, if timestamps are absent, then we just need to delete the second line in (6).

### B. Constrained Manifold Learning

Given the graph weight matrix $W$ and the locations $P = [p_1, ..., p_L]$ for the labeled samples, we can find a multi-point channel chart $Z$, by minimizing the cost function in (4). Denote $Q = [P, P] \in \mathbb{R}^{(L+U) \times 2}$, and set the unknown locations $P$ to zeros. Then we rewrite $f(Z)$ as

$$f(Z) = \sum_{n=1}^{L} \| z_n - p_n \|^2 + \gamma \text{Tr} \left( Z^T AZ - Z^T WZ \right) = \text{Tr} \left[ (Z - Q)^T J (Z - Q) + \gamma Z^T L Z \right],$$

where $J = \text{diag} \left\{ 1, \ldots, 1, 0, \ldots, 0 \right\}$, $\Lambda$ is the degree matrix of $\mathcal{G}$ with the $n$th diagonal element $\Lambda_n = \sum_{m=1}^{L+U} w_{n,m}$, and $L = \Lambda - W$ is the graph Laplacian of $\mathcal{G}$.

As in [22], by setting the derivative of objective function to be zero, the aligned channel chart can be estimated as

$$Z = (\gamma L + J)^{-1} J Q. \quad (7)$$

In (7), larger values of $\gamma$ will increase the smoothness of the channel chart $Z$, while $\gamma$ should be close to 0 to map labeled samples to its ground-truth positions. A suitable value of $\gamma$ can be found experimentally as will be shown in IV.

### IV. Simulation Results

We now demonstrate the efficacy of the proposed graph-based SS-MPCC for CSI-based localization with timestamps and a subset of CSI samples with marked positions.

#### A. Simulated Scenario and Evaluation Metrics

We consider an urban outdoor multi-cell mmWave network scenario as depicted in Figure 3. BSs are below rooftop, and signals will be reflected or blocked by the walls. A ray-tracing channel model is used to generate the multi-path channels. There are $B = 6$ mMIMO BSs equipped with ULAs, each has $M = 16$ elements. The antenna arrays of the BSs are oriented perpendicular to the building surfaces where the BSs
are mounted. UEs transmit reference signals with a fixed power of 23 dBm. The carrier frequency is 28 GHz, system bandwidth is 256 MHz with 128 OFDM subcarriers.

In the simulated scenario, UEs move along the roads with an average speed of 5 meters per second. We collect $U = 3000$ CSI samples from UE traces and the average distance between neighboring sampled locations is approximately 2 m, which is much larger than the carrier's wavelength. The covariance matrix is estimated over 32 realizations of the channel vector $h$ over multiple adjacent subcarriers. The channel quality is estimated as $E(\|h\|^2)$. For semi-supervised learning, additional $L$ samples are generated randomly on the roads. To measure the performance of CSI-based localization methods, we use the following metrics:

1) **Mean localization error**: For semisupervised localization, a natural way for measuring performance is the mean localization error (MLE) for unlabeled samples:

$$MLE = \frac{1}{U} \sum_{n=1}^{U} \|z_{L+n} - p_{L+n}\|_2.$$  

2) **Kruskal’s stress**: In addition to the MLE, we use Kruskal’s stress (KS) to measure how well the learned channel chart represents the true locations globally:

$$KS = \frac{\sum_{n,m} (\delta_{n,m} - \beta d_{n,m})^2}{\sum_{n,m} \delta_{n,m}^2}.$$  

Here, $\delta_{n,m} = \|p_n - p_m\|$, $d_{n,m} = \|z_n - z_m\|$, and $\beta = \frac{\sum_{n,m} \delta_{n,m} d_{n,m}}{\sum_{n,m} \delta_{n,m}^2}$. KS is in the range $[0, 1]$ and smaller values indicate better preservation of global geometry.

3) **Trustworthiness and continuity**: We use two standard metrics that characterize neighborhood preserving performance: trustworthiness (TW) and continuity (CT) [14]–[16]. The TW measures whether the mapping of high-dimensional samples to the low-dimensional space introduces false neighbors; the CT measures whether similar samples in high-dimensional space remain similar in low-dimensional space. TW and CT have values in $[0, 1]$ and larger values imply better preservation of neighborhood relationship. When evaluating TW and CT, a number $K$ of neighbors should be specified.

**B. Virtual Tx Point Cloud**

Figure 5 depicts the virtual Tx point cloud extracted by BS 1. These points are related to 536 mobile locations inside the coverage area of BS 1, and exhibit clear clusters when compressed to a confined area via a nonlinear scaling function.

**C. Investigations of Algorithm Parameters**

There are seven parameters: the alignment-smoothness tradeoff coefficient $\gamma$, mixed edge weight tradeoff coefficient $\alpha$, number of neighbors $k_e$, timestamp threshold $T_{th}$, and the heat kernel parameters $\theta_s, \theta_d$ and $\theta_t$. We first carried out experiments to study the impact of parameters $k_e$ on the localization performance, as it largely determines the structure of the graph. As shown in Figure 4a, when other parameters are fixed, there is an optimal $k_e$. A very small $k_e$ cannot well characterize the neighborhood relations among nodes. Here we set $k_e = 16$, which leads to the lowest MLE as in Figure 4a. $T_{th}$ plays a similar role as $k_e$, we found that in our simulation scenario, $T_{th} = 2.4 s$ is a good option as depicted in Figure 4d. Small kernel parameters will de-emphasize neighbor connections with larger dissimilarities, while large kernel parameters cannot find the fine-grained local structures. However, when the neighborhood graph is set, their impacts on MLE performance is limited, we set $\theta_s = \theta_d = 5$ and $\theta_t = 2$ in the following simulations. The parameter $\gamma$ has a profound impact on the learned channel chart as can be seen from Figure 4b. Values of $\gamma$ larger than
one leads to a shrunken chart as it helps to minimize the Graph Laplacian. It turns out \( \gamma = \frac{L}{U} \) is a good choice in our simulations; for the mixed edge weight tradeoff coefficient \( \alpha \), we found that \( 1 - L/U \) worked well.

D. Performance Comparison

We now investigate the performance of the proposed SS-MPCC method for network-side localization. We compare SS-MPCC to unsupervised MPCC [18], and kNN method, which is widely used in CSI-based fingerprinting [6]. We will show that SS-MPCC is applicable to a variety of scenarios. It can work with different types and amounts of side-information.

First, as demonstrated by Figure 6, the proposed graph-based SS-MPCC method is capable of unveiling mobile devices’ locations in the considered grid-like urban environment. With only \( 50 \) labeled CSI samples, the network topology is revealed, and the UEs can be located with a MLE as \( 12.6 \) m. When this ratio increases to \( 10\% \), a MLE of \( 5.6 \) m is achieved. Note that such a positioning accuracy does not assume tight synchronization and/or high-precision calibration among multiple BSs, what they need to know is that a CSI sample is transmitted from a UE during the same time interval. This result can be understood by the fact that the timestamps and multi-point data fusion in SS-MPCC help in preserving the neighborhood relationships of the CSI data on each user trace and/or across multiple user traces, which is beneficial in learning the manifold of unlabeled CSI data. As a reward, a small set of labeled data is adequate for network-side positioning and tracking.

To gain more understanding on the efficacy and working mechanisms of SS-MPCC, we investigate its performance with only timestamps, and with only marked positions. For the sake of completeness, results of conventional unsupervised MPCC are also given, as illustrated in Figure 7 and summarized in Table I. The channel chart in Figure 7a generated by unsupervised MPCC is capable of preserving most of the local manifold structures, but with a twisted global structure, and interruptions at some points. By utilizing the timestamps, the chart in Figure 7b is smoother and the traces of UEs are using proper parameters, an aligned channel chart is shown as in Figure 7c. Compared to Figure 7b, the TW and CT scores are smaller, but absolute positioning with a MLE = \( 16.7 \) m is enabled with \( 10\% \) labeled CSI samples.

Finally, we compare SS-MPCC to kNN. kNN is only

<table>
<thead>
<tr>
<th>TABLE I</th>
</tr>
</thead>
<tbody>
<tr>
<td>Performance Comparison</td>
</tr>
<tr>
<td>**</td>
</tr>
<tr>
<td>Timestamps</td>
</tr>
<tr>
<td>L</td>
</tr>
<tr>
<td>MLE [m]</td>
</tr>
<tr>
<td>KS</td>
</tr>
<tr>
<td>TW</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>CT</td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>
applicable when there are labeled samples, it estimates the position of an unlabeled sample by computing the barycenter of $k$ most similar labeled ones, and neglects the timestamp information. Here we set $k = 5$ for kNN, which yields optimal MLE performance. As shown in Table I, SS-MPCC performs much better than kNN for all metrics, especially when timestamps of unlabeled samples are exploited. From the CDFs of MLEs in Figure 8, we see that a portion of locations have large localization errors for kNN, which correspond to the positions without nearby labeled anchor points, especially at the network edges. Furthermore, as is shown in Figure 8, when the number of labeled samples drops from 300 to 150 by half, the performance of MPCC is only slightly affected, with mean MLE increased from 5.6 m to 6.8 m, while the MLE of kNN is increased from 18.8 m to 25.1 m.

V. CONCLUSIONS

We have shown that side-information, such as timestamps and a set of labeled samples, can be incorporated into a semi-supervised multi-point channel charting framework to support large-scale network-side localization and tracking in a distributed manner. Moreover, we have demonstrated that the multi-BS CSI feature fusion and incorporation of timestamps to increase manifold smoothness can be implemented in global graph construction based on graph Laplacian. Finally, through experiments in a simulated urban outdoor scenario, we have shown that SS-MPCC is able to perform large-scale network-side positioning for scenarios with realistic UE motion, even with a very small portion of labeled data. The proposed framework is scalable and automatic in the sense that it could be implemented for varying number of BSs, with spatially sparse labeled samples, and does not require accurate network synchronization. There are many opportunities for extensions of this work, including out-of-sample mapping to locate new samples on the channel chart for real-time positioning applications, and filtering algorithms for multi-target tracking.

REFERENCES


