Cooperative Localization in Massive Networks

Yifeng Xiong, *Student Member, IEEE*, Nan Wu, *Member, IEEE*, Yuan Shen, *Senior Member, IEEE*,
and Moe Z. Win, *Fellow, IEEE*

Abstract—Network localization is capable of providing accurate and ubiquitous position information for numerous wireless applications. This paper studies the accuracy of cooperative network localization in large-scale wireless networks. Based on a decomposition of the equivalent Fisher information matrix (EFIM), we develop a random-walk-inspired approach for the analysis of EFIM, and propose a position information routing interpretation of cooperative network localization. Using this approach, we show that in large lattice and stochastic geometric networks, when anchors are uniformly distributed, the average localization error of agents grows logarithmically with the reciprocal of anchor density in an asymptotic regime. The results are further illustrated using numerical examples.

Index Terms—Network localization, wireless network, efficiency of cooperation, asymptotic analysis, information inequality.

I. INTRODUCTION

NETWORK LOCALIZATION is a key enabler of various wireless applications requiring location-awareness [1]–[3], including Internet-of-things [4]–[7], autonomous vehicles [8]–[10], and big data [11]–[13]. In most outdoor environments, global navigation satellite system (GNSS) can provide meter-level accuracy for various location-based services, however, their performance are severely degraded in harsh environments such as urban canyons, underground and indoor scenarios. Exploiting the cooperation among nodes, network localization can provide satisfactory positioning performance in these GNSS-denied environments [14]–[16].

Nodes in network localization fall into two categories: agents with unknown positions and anchors with precisely known positions. To elaborate further, anchors are usually used to model base stations in cellular networks, road-side units in vehicular networks, leading nodes in unmanned aerial vehicle swarms, and access points in wireless local area networks. Agents are widely used to model devices requesting positioning services, such as users in cellular networks and vehicles in vehicular networks.

Typically, agents communicate with neighboring nodes, and infer their positions using some signal metrics ex-

tracted from received signals including time-of-arrival (TOA), time-difference-of-arrival (TDOA), and angle-of-arrival (AOA) [17]–[27]. TOA and TDOA can be obtained by measuring the propagation delay of the received signal, which can provide ranging information between nodes. AOA is a metric characterizing the incoming direction of the received signal, which can be obtained using directional antennas or antenna arrays. TOA- and TDOA-based network localization algorithms have been well studied in the literature. Recently, driven by the application of massive multiple-input multiple-output (MIMO) technology in the forthcoming new generation of wireless systems, techniques combining ranging and bearing information have also been studied [28]–[30].

In the classical non-cooperative scheme, network localization is performed using only anchor-agent communication. By contrast, the emerging scheme of cooperative network localization also incorporates agent-agent communications. Intuitively, the inter-agent measurements provide additional position information, and hence has the potential of improving the localization accuracy and preventing the outage of localization services, which has been verified empirically by many existing works [18], [31]–[33]. However, these results are based on the performance of specific localization algorithms, which cannot reflect the essential relation between the localization performance and network parameters.

In order to gain more insights into the cooperative localization scheme, some performance limits have been studied [34]–[42]. These limits have been exploited to address network operation tasks such as power allocation [43], [44]. The tool of squared position error bound (SPEB), which gives a lower bound for the position mean-squared error (MSE) using the inverse of equivalent Fisher information matrix (EFIM) [45]–[47], are used extensively in these works.

Apart from finding applications in network operation tasks, performance limits are also important in their own rights for understanding the efficiency of cooperation among agents. This could help us to determine whether the accuracy improvement introduced by agent cooperation is worth the additional system complexity. Especially, scaling laws of the localization accuracy are particularly desirable for massive networks, since obtaining the performance limits explicitly for these networks would be computationally expansive. However, closed-form expressions of the performance limits are only proposed for some specific network topologies [48], and existing results on the asymptotic localization performance in general large-scale networks are mostly qualitative [49]–[52]. Especially, there is a lack of investigation into the localization accuracy scaling laws in multi-hop networks. For these networks, traditional triangulation techniques fail to provide reliable localization services, whereas cooperative network localization becomes

TABLE I
COMPARISON BETWEEN THE ERROR SCALING LAWS PROPOSED IN THIS PAPER AND THOSE DEVELOPED IN EXISTING CONTRIBUTIONS.

Original Contribution	Scheduling	Measurement Type	Anchor Deployment	Error Scaling (as $N_a \to \infty$)
[45]	Fully cooperated (optimal)	Distance only ($N_t = 1$)	Constant density	Converges to a *constant*
[53]			Located in the innermost circle	*Quadratic* in the number of hops from the anchors
This paper				*Logarithmic* in the radial direction; *Quadratic* in the tangential direction
			Single anchor	*Logarithmic* in the radial direction; *Unbounded* in the tangential direction
		Both distance and angle ($N_t \geqslant 2$)		*Logarithmic* in both directions
		Distance only or both distance and angle ($N_t = 1$ or $N_t \geqslant 2$)	Density $\to 0$	*Logarithmic* in (anchor density)$^{-1}$
[54]–[56]	Sequential			At least proportional to the *square root* of (anchor density)$^{-1}$

particularly useful. The essential difficulty of the asymptotic analysis lies in the information coupling phenomenon [48], because of which the EFIM does not possess a block-diagonal structure, and is intractable to be inverted analytically.

Recently, we have developed a framework for understanding the scaling laws of relative and absolute synchronization errors in large networks [57]. The network synchronization problem share some similarities with the cooperative localization problem. Especially, they both rely on the pairwise measurement between adjacent nodes. However, it is not straightforward to apply the results in [57] to the cooperative localization problem, since the analysis method therein relies on the fact that the local parameter (i.e., the delay) in network synchronization is a scalar, while the location parameters are typically vectors.

In this paper, we generalize the framework proposed in [57] to aid the asymptotic analysis of network localization. In particular, we develop a random walk interpretation of the EFIM for location parameters, by introducing generalized random walks having matrix-valued "pseudo-probabilities". This enables us to depict the asymptotic localization performance quantitatively. For better clarity of illustration, we have summarized the asymptotic error scaling laws proposed in this paper, and contrasted them to the existing ones in Table I.

The contributions of this paper are summarized as follows.

- We propose a decomposition of the EFIM, based on which we propose the concepts of nominal position information (NPI) and efficiency of cooperation (EoC) to characterize the information coupling between an agent and its neighbors.
- We develop a random walk interpretation of the EoC. Inspired by this interpretation, we describe quantitatively an intuition of "position information routing" for network localization.
- We show that in large lattice and stochastic geometric networks[1], the average position MSE of agents grows only

[1]In this paper, we consider the theoretical models of lattice and stochastic geometric networks instead of real-world networks, since massive cooperative location-aware networks typically consist of mobile nodes, and hence do not have fixed topologies. These models are also widely used in other research works that analyze large-scale wireless networks [58]–[68].

logarithmically with the reciprocal of anchor density in an asymptotic regime, if anchors are distributed uniformly in the network.
- We show that for the concentric multi-hop network model considered in [53] where each node is equipped with a single antenna, the directional position error bound (DPEB) [45] grows *logarithmically* on the direction to the centroid of that region, while it increases *quadratically* on the orthogonal direction.

The rest of this paper is organized as follows. Section II formulates the network localization problem, specifies the system model, and describes the general form of EFIM. In Section III we conduct a random-walk-inspired analysis of EFIM, and present the position information routing interpretation. Section IV develops expressions for the asymptotic behavior of EFIM and SPEB in large lattice and stochastic geometric networks. We further discuss the practical implications of our results in Sec. V. The analytical results are illustrated in Section VI, and the conclusions are drawn in Section VII.

Notations: Throughout this paper, a, **a**, **A**, and A represent random variables (scalars), random vectors, random matrices and random sets, respectively; Their realizations, or the corresponding deterministic quantities, are denoted by a, $\boldsymbol{a}$, $\boldsymbol{A}$, and $\mathcal{A}$, respectively. The m-by-n matrix of zeros (resp. ones) is denoted by $\boldsymbol{0}_{m \times n}$ (resp. $\boldsymbol{1}_{m \times n}$). The m-dimensional vector of zeros (resp. ones) is denoted by $\boldsymbol{0}_m$ (resp. $\boldsymbol{1}_m$). The m-by-m identity matrix is denoted by $\boldsymbol{I}_m$. $\boldsymbol{Q}_n \in \mathbb{R}^{2n \times 2}$ denotes a matrix given by $[\boldsymbol{I}_2 \ \boldsymbol{I}_2 \ \dots \ \boldsymbol{I}_2]^{\mathrm{T}}$. These subscripts are omitted if they are clear from the context. $\boldsymbol{E}_{ij}^{M,N}$ denotes a $M \times N$ matrix with all zeros but a 1 on the (i,j)-th entry, which is simplified as $\boldsymbol{E}_{ij}^N$ when $M = N$. The notation $[\cdot]_{i,j}$ denotes the (i,j)-th 2×2 sub-matrix of its argument; $[\cdot]_{r_1:r_2,c_1:c_2}$ denotes a submatrix obtained by extracting the $(2r_1 - 1)$-th to the $(2r_2)$-th row and the $(2c_1 - 1)$-th to the $(2c_2)$-th column of its argument. $[\cdot]^{\dagger}$ denotes the Moore-Penrose pseudo inverse of its argument, while $[\cdot]^{\mathrm{H}}$ denotes the Hermitian transpose of its argument. $\mathrm{tr}\{\cdot\}$ stands for the trace of a square matrix. The Kronecker product between matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ is denoted by $\boldsymbol{A} \otimes \boldsymbol{B}$. $\|\boldsymbol{x}\|_p$ denotes the l_p norm, which represents the l_2

norm by default when the subscript is omitted. $\boldsymbol{u}(\varphi)$ denotes a unit vector $[\cos\varphi \ \sin\varphi]^{\mathrm{T}}$, and $\boldsymbol{J}_{\mathrm{r}}(\varphi) := \boldsymbol{u}(\varphi)\boldsymbol{u}(\varphi)^{\mathrm{T}}$. The function $\mathbb{1}\{\cdot\}$ takes the value 1 if its argument as a logical expression is true, and 0 otherwise. The notation $\min^{(i)}(\mathcal{S})$ denotes the i-th smallest element in the set $\mathcal{S}$.

The notation $\mathbb{E}_{\mathsf{x}}\{\cdot\}$ denotes the expectation with respect to x, and the subscript is removed when there is no confusion. The functions $f_{\mathsf{x}}(\boldsymbol{x})$ and $f_{\mathsf{x}}(\boldsymbol{x}, \boldsymbol{\theta})$ denote the probability density function (PDF) of x and the PDF of x parameterized by $\boldsymbol{\theta}$, respectively. We also define the following function

$$\jmath_{\mathrm{bm}}(\boldsymbol{z}, a(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \boldsymbol{\theta}_3), \boldsymbol{\theta}_1, \boldsymbol{\theta}_2) := \frac{\partial \ln f_{\mathsf{z}}(\boldsymbol{z}; a(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \boldsymbol{\theta}_3))}{\partial \boldsymbol{\theta}_1}$$
$$\cdot \frac{\partial \ln f_{\mathsf{z}}(\boldsymbol{z}; a(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \boldsymbol{\theta}_3))}{\partial \boldsymbol{\theta}_2^{\mathrm{T}}}.$$

Some frequently-used Bachmann-Landau notations are given as follows.

$$\begin{array}{ll} a(n) = O(b(n)) & \limsup_{n\to\infty} a(n)(b(n))^{-1} < \infty \\ a(n) = \Omega(b(n)) & \liminf_{n\to\infty} a(n)(b(n))^{-1} > 0 \\ a(n) = \Theta(b(n)) & a(n) = O(b(n)) \text{ and } a(n) = \Omega(b(n)) \end{array}$$

II. SYSTEM MODEL AND EVALUATION OF EFIM

Consider a network with N_{a} agents and N_{b} anchors. The indices of agents and anchors constitute sets $\mathcal{N}_{\mathrm{a}} = \{1, 2, \ldots, N_{\mathrm{a}}\}$ and $\mathcal{N}_{\mathrm{b}} = \{N_{\mathrm{a}}+1, N_{\mathrm{a}}+2, \ldots, N_{\mathrm{a}}+N_{\mathrm{b}}\}$, respectively. The position of node i is denoted by $\boldsymbol{p}_i = [p_{\mathrm{x}i} \ p_{\mathrm{y}i}]^{\mathrm{T}}$. The vector containing positions of all nodes is denoted by $\boldsymbol{p}_{\mathrm{all}} = [\boldsymbol{p}_1^{\mathrm{T}} \ \boldsymbol{p}_2^{\mathrm{T}} \ \cdots \ \boldsymbol{p}_{N_{\mathrm{a}}+N_{\mathrm{b}}}^{\mathrm{T}}]^{\mathrm{T}}$, and especially for agents we denote $\boldsymbol{p} = [\boldsymbol{p}_1^{\mathrm{T}} \ \boldsymbol{p}_2^{\mathrm{T}} \ \cdots \ \boldsymbol{p}_{N_{\mathrm{a}}}^{\mathrm{T}}]^{\mathrm{T}}$. The angle from node i to j is denoted by φ_{ij}.

In general, all relative position measurements between node i and j are measurements of the displacement vector $\boldsymbol{d}_{ij} := \boldsymbol{p}_i - \boldsymbol{p}_j$. In this paper, we assume in particular that each node i is equipped with an array of N_{t} antennas.[2] The received signals at node j from node i can be modeled as

$$\mathbf{r}_{ij}(t) = N_{\mathrm{t}} \sum_{l=1}^{N_{\mathrm{p}}} \alpha_{ij}^{(l)} \boldsymbol{a}_j \boldsymbol{a}_i^{\mathrm{H}} \boldsymbol{s}_i(t - \tau_{ij}^{(l)}) + \mathbf{z}_{ij}(t) \qquad (1)$$

where N_{p} denotes the number of propagation paths, $\boldsymbol{s}_i(t) = [s_{i1}(t) \ s_{i2}(t) \ \ldots \ s_{iN_{\mathrm{t}}}(t)]^{\mathrm{T}}$ denotes a set of orthonormal transmitting waveforms, $\boldsymbol{a}_i$ is the unit array steering vector of node i, $\alpha_{ij}^{(l)}$ is the amplitude of the received signal of the l-th path, and $\mathbf{z}_{ij}(t)$ denote the noise modeled as white complex Gaussian processes. The propagation delay of the l-th path $\tau_{ij}^{(l)}$ is given by $c^{-1}d_{ij}^{(l)}$ where c is the signal propagation speed and $d_{ij}^{(l)}$ represents the length of the l-th propagation path. We denote by $\mathbf{r}$ the concatenated vector of the Karhunen-Loeve expansions of all $\mathbf{r}_{ij}(t)$. We assume that there is a line-of-sight (LOS) path corresponding to $l = 1$ for each link, meaning that $\tau_{ij}^{(1)} = c^{-1}\|\boldsymbol{p}_i - \boldsymbol{p}_j\|$.

We denote the receiver sensitivity of each node i by $p_{\mathrm{rs},i}$, below which the receiving power is not sufficient for successful signal detection. For simplicity, we say that node i is in the measurement range of node j (and denote this relation as

[2]Our results can also be applied to the case in which nodes are equipped with different number of antennas after minor modifications.

$i \in \mathcal{N}_j$), if we have $\alpha_{ij}^2 \|\boldsymbol{p}_i - \boldsymbol{p}_j\|^{-\gamma} \leqslant p_{\mathrm{rs},j}$ where γ is the path loss exponent and $\alpha_{i,j}$ denotes the signal amplitude of the LOS path; otherwise node j cannot receive measurements from node i. The connectivity between nodes can then be described using a graph $\mathcal{G}_{\mathrm{net}} = \{\mathcal{V}_{\mathrm{net}}, \mathcal{E}_{\mathrm{net}}\}$, where the vertex set $\mathcal{V}_{\mathrm{net}} = \mathcal{N}_{\mathrm{a}} \cup \mathcal{N}_{\mathrm{b}}$, and $\mathcal{E}_{\mathrm{net}}$ is the set of edges given by

$$\mathcal{E}_{\mathrm{net}} := \{(i,j)|i \in \mathcal{V}_{\mathrm{net}}, \ j \in \mathcal{V}_{\mathrm{net}}, \ i \leftrightarrow j\}$$

where $i \leftrightarrow j$ means $i \in \mathcal{N}_j$ (or $j \in \mathcal{N}_i$) holds. Naturally, the connectivity between agents constitutes another graph, denoted by $\mathcal{G}_{\mathrm{agt}}$.

Apart from the agents' positions, there are also some unknown channel parameters denoted by $\boldsymbol{\eta}$. In light of this, we consider $\boldsymbol{\theta} = [\boldsymbol{p}^{\mathrm{T}} \ \boldsymbol{\eta}^{\mathrm{T}}]^{\mathrm{T}}$ as the complete vector of parameters. For any unbiased estimator $\hat{\mathbf{p}}$ of $\boldsymbol{p}$, the MSE matrix of $\hat{\mathbf{p}}$ satisfies

$$\mathbb{E}\{(\boldsymbol{p} - \hat{\mathbf{p}})(\boldsymbol{p} - \hat{\mathbf{p}})^{\mathrm{T}}\} \succcurlyeq \boldsymbol{J}_{\mathrm{e}}^{-1}(\boldsymbol{p})$$

where $\boldsymbol{J}_{\mathrm{e}}(\boldsymbol{p})$ is the EFIM for $\boldsymbol{p}$ with respect to $\boldsymbol{\theta}$. To further characterize the performance limit of a specific agent i, we define the SPEB of agent i as

$$\mathrm{sp}\{\boldsymbol{p}_i\} := \mathrm{tr}\{\boldsymbol{J}_{\mathrm{e}}^{-1}(\boldsymbol{p}_i)\}$$

and the DPEB of agent i along the direction of $\boldsymbol{u}$ as [45]

$$\mathrm{dp}\{\boldsymbol{p}_i; \boldsymbol{u}\} := \boldsymbol{u}^{\mathrm{T}} \boldsymbol{J}_{\mathrm{e}}^{-1}(\boldsymbol{p}_i)\boldsymbol{u}.$$

Using SPEB, we have $\mathbb{E}\{\|\boldsymbol{p}_i - \hat{\mathbf{p}}_i\|^2\} \geqslant \mathrm{sp}\{\boldsymbol{p}_i\}$. Using DPEB, the position error along the direction of $\boldsymbol{u}$ can be bounded by $\mathbb{E}\{\|\boldsymbol{u}^{\mathrm{T}}(\boldsymbol{p}_i - \hat{\mathbf{p}}_i)\|^2\} \geqslant \mathrm{dp}\{\boldsymbol{p}_i; \boldsymbol{u}\}$.

According to [45, Theorem 1], the EFIM for $\boldsymbol{p}$ can be written as

$$\boldsymbol{J}_{\mathrm{e}}(\boldsymbol{p}) = \boldsymbol{D} - \boldsymbol{A} \qquad (2)$$

where $\boldsymbol{D} := \mathrm{diag}(\boldsymbol{D}_1, \ldots, \boldsymbol{D}_{N_{\mathrm{a}}})$ with $\boldsymbol{D}_i$ given by $\boldsymbol{D}_i = \sum_{i \leftrightarrow j} \boldsymbol{J}_{ij}$, and

$$\boldsymbol{A} := - \sum_{i \leftrightarrow j, i \in \mathcal{N}_{\mathrm{a}}, j \in \mathcal{N}_{\mathrm{a}}} \boldsymbol{E}_{ij}^{N_{\mathrm{a}}} \otimes \boldsymbol{J}_{ij}.$$

Matrix $\boldsymbol{J}_{ij}$ denotes the contribution of measurement $\mathbf{r}_{ij}$ to the position information of agent i (as well as j) given by

$$\boldsymbol{J}_{ij} = \mathbb{E}\{\jmath_{\mathrm{bm}}(\mathbf{r}_{ij}, [\boldsymbol{d}_{ij} \ \kappa_{ij}], \boldsymbol{p}_i, \boldsymbol{p}_j)\} \qquad (3)$$

where κ_{ij} denotes nuisance parameters related to channel condition, transmission power, signal waveform, etc.

We further assume that the power transmitted from each antenna is identical in every single node. Under previous assumptions, it is known that in far-field environments, $\boldsymbol{J}_{ij}$ can be approximately decomposed as follows [35, Corollary 1]

$$\boldsymbol{J}_{ij} \approx N_{\mathrm{t}}^2 \left(\widetilde{\xi}_{ij,\mathrm{r}} \boldsymbol{J}_{\mathrm{r}}(\varphi_{ij}) + \widetilde{\xi}_{ij,\mathrm{t}} \boldsymbol{J}_{\mathrm{r}}\left(\varphi_{ij} + \frac{\pi}{2}\right) \right) \qquad (4)$$

where $\widetilde{\xi}_{ij,\mathrm{r}} := p_i \beta^2 \xi_{ij}$, $\widetilde{\xi}_{ij,\mathrm{t}} := p_i f_{\mathrm{c}}^2 \xi_{ij} G_{ij} \|\boldsymbol{d}_{ij}\|^{-2}$, p_i is the transmission power of node i, β is the transmitted signal bandwidth, ξ_{ij} is given by $\xi_{ij} = \zeta_{ij} \|\boldsymbol{d}_{ij}\|^{-\gamma}$ with ζ_{ij} being the ranging coefficient [69] determined by channel and waveform parameters, and G_{ij} denotes the squared array aperture function (SAAF) determined by the shapes of the

arrays equipped on nodes. When $N_\mathrm{t} = 1$, (4) can be simplified as

$$\boldsymbol{J}_{ij} \approx \widetilde{\xi}_{ij,\mathrm{r}} \boldsymbol{J}_\mathrm{r}(\varphi_{ij}). \qquad (5)$$

In general, G_{ij} is determined by the orientations of node i and j. As a special case, G_{ij} is a constant with respect to orientations in the case of uniformly oriented array (UOA) [35, Definition 4]. When $N_\mathrm{t} = 1$, we have $G_{ij} = 0$, and hence (4) reduces to (5).

III. INFORMATION COUPLING AND RANDOM-WALK-INSPIRED ANALYSIS

To obtain the desired MSE lower bound, we have to take the inverse of the EFIM for $\boldsymbol{p}$ derived in Section II, which is a rather difficult task due to the notorious information coupling phenomenon. To alleviate this difficulty, in this section we propose a decomposition of EFIM inspired by the theory of random walk on graphs.

A. Information Coupling and EFIM Decomposition

In a non-cooperative network, it is well-known that the EFIM for agent i can be expressed as [46]

$$\boldsymbol{J}_\mathrm{e}(\boldsymbol{p}_i) = \boldsymbol{D}_i = \sum_{j \in \mathcal{N}_\mathrm{b}} \boldsymbol{J}_{ij} \qquad (6)$$

due to the block-diagonal structure of the EFIM for $\boldsymbol{p}$. Here, the second equality follows from the fact that agents can only communicate with anchors in non-cooperative networks. Unfortunately, this simple formula does not hold for cooperative networks since the corresponding EFIMs are not block-diagonal.

Intuitively, for any agent i in a cooperative network, the neighboring agents can be viewed as "weak anchors" in the sense that they provide less position information, due to their position uncertainty. Hence we may have

$$\boldsymbol{J}_\mathrm{e}(\boldsymbol{p}_i) = \boldsymbol{D}_i \boldsymbol{X}_i \preccurlyeq \boldsymbol{D}_i$$

where the matrix $\boldsymbol{X}_i$ depicts the overall efficiency that agent i utilizes the position information obtained from neighboring nodes. Formally speaking, we have the following result.

Theorem 1: When the inverse of $\boldsymbol{J}_\mathrm{e}(\boldsymbol{p})$ exists, the EFIM of agent i can be decomposed as

$$\boldsymbol{J}_\mathrm{e}(\boldsymbol{p}_i) = \boldsymbol{D}_i(\boldsymbol{I} + \boldsymbol{\Delta}_i)^{-1} \ \forall i \in \mathcal{N}_\mathrm{a} \qquad (7)$$

where $\boldsymbol{\Delta}_i := \sum_{n=1}^{\infty} \boldsymbol{T}_{ii}^{(n)} \succcurlyeq \boldsymbol{0}$ with $\boldsymbol{T}_{ij}^{(n)}$ given by the following recursion

$$\boldsymbol{T}_{ij}^{(n)} = \sum_{k=1}^{N_\mathrm{a}} \boldsymbol{T}_{ik}^{(n-1)} \boldsymbol{T}_{kj}^{(1)} \qquad (8)$$

where

$$\boldsymbol{T}_{ij}^{(1)} = \begin{cases} \boldsymbol{D}_i^{-1} \boldsymbol{J}_{ij}, & i \neq j, \ i \notin \mathcal{N}_\mathrm{b}; \\ \boldsymbol{I}_2, & i = j, \ i \in \mathcal{N}_\mathrm{b}; \\ \boldsymbol{0}_{2 \times 2}, & \text{otherwise.} \end{cases} \qquad (9)$$

Proof: Please refer to Appendix I. □

Considering the similarity between (6) and the definition of $\boldsymbol{D}_i$, matrix $\boldsymbol{D}_i$ can be interpreted as the nominal position information provided by agent i's neighbors when information
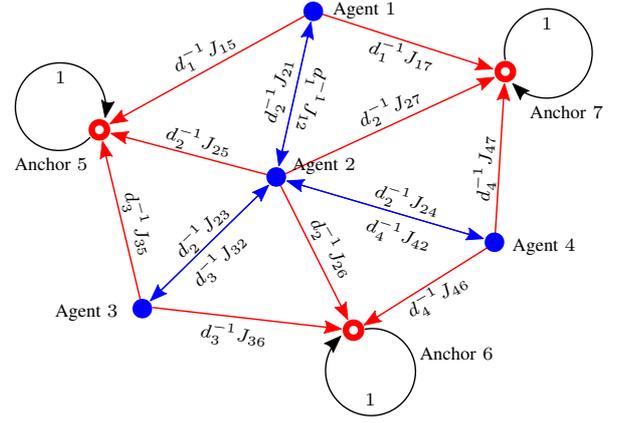

Fig. 1. The transition diagram portraying the random walk characterized by (11b) and (11c). The quantities next to the edges represent the corresponding transition probability.

coupling is ignored, and hence we give the following definition.

Definition 1 (Nominal Position Information): The nominal position information (NPI) of agent i is defined as $\boldsymbol{D}_i$.

Since EFIM characterizes the effective position information acquired by a certain agent, it can be seen from (7) that the term $(\boldsymbol{I} + \boldsymbol{\Delta}_i)^{-1}$ quantifies the efficiency of cooperation between agent i and its neighbors, and can be viewed as a generalized version of the efficiency of cooperation (EoC) defined in [57]. Note that EoC satisfies $\boldsymbol{0} \preccurlyeq (\boldsymbol{I} + \boldsymbol{\Delta}_i)^{-1} \preccurlyeq \boldsymbol{I}$ due to the non-negativity of $\boldsymbol{\Delta}_i$, implying that the information coupling among agents will lead to degeneration of cooperation efficiency. As an extreme case, when all neighboring nodes of agent i are anchors, there is no coupling between agent i and its neighbors, and thus we have $(\boldsymbol{I} + \boldsymbol{\Delta}_i)^{-1} = \boldsymbol{I}$ and $\boldsymbol{J}_\mathrm{e}(\boldsymbol{p}_i) = \boldsymbol{D}_i$.

To simplify further discussions, we will denote

$$\boldsymbol{T} := \boldsymbol{D}^{-1} \boldsymbol{A} \qquad (10)$$

in the rest of this paper.

Remark 1: Most of the results in this subsection have their analogue in [57]. However, it should be noted that the EoC in this paper is a matrix instead of a scalar. Furthermore, the quantity $\boldsymbol{T}_{ij}^{(n)}$ defined in (8) corresponds to the n-step transition probability in [57], but now it can no longer be viewed formally as a probability, since it is a matrix. Fortunately, as we shall discuss in more detail in Sec. III-C, these matrix-valued quantities can still be regarded as probabilities in a certain sense, and some classical results in the random walk theory can still be applied.

B. Random-Walk-Inspired Formalism: The 1-D Toy Example

The EoC characterizes the effect of information coupling. In fact, if we can obtain the EoC of an agent i, we can readily compute its EFIM using (7) without inverting the whole EFIM for $\boldsymbol{p}$. However, (8) and (9) are still rather intricate, and hence it is not clear that how EoC is related to other characteristics of the network. To facilitate understanding, we first consider the toy example of 1-D cooperative localization, in which all nodes

locate along a line. Although this example is much different from real-world scenarios, the technique used in the analysis will be similar in spirit to those applied in the generic case.

In this simplified scenario, quantities such as $\boldsymbol{D}_i$, $\boldsymbol{J}_{ij}$ and $\boldsymbol{\Delta}_i$ are all scalars, and hence we denote them as d_i, J_{ij} and Δ_i, respectively. Using the definition (10), the equations (7), (8) and (9) can be rewritten as

$$\boldsymbol{J}_{\mathrm{e}}(p_i) = \frac{d_i}{1 + \Delta_i} \ \forall i \in \mathcal{N}_{\mathrm{a}} \tag{11a}$$

$$\boldsymbol{T}^n = \boldsymbol{T}^{n-1}\boldsymbol{T} \tag{11b}$$

$$T_{ij} = \begin{cases} d_i^{-1} J_{ij}, & i \neq j, \ i \notin \mathcal{N}_{\mathrm{b}}; \\ 1, & i = j, \ i \in \mathcal{N}_{\mathrm{b}}; \\ 0, & \text{otherwise.} \end{cases} \tag{11c}$$

respectively, where T_{ij} denotes the (i, j)-th entry of matrix $\boldsymbol{T}$. In addition, we will denote the (i, j)-th entry of $\boldsymbol{T}^n$ as $T_{ij}^{(n)}$. Due to the non-negativity of Fisher information, the term T_{ij} is non-negative, and satisfies $\sum_{j \in \mathcal{N}_i} T_{ij} = 1$ according to the definition of $\boldsymbol{D}$. In light of this, we may regard matrix $\boldsymbol{T}$ as the *transition probability matrix* of a random walk, as portrayed in Fig. 1.

From Fig. 1 we can see that anchors correspond to the *absorbing states* of the random walk, since the probability of staying at any state representing an anchor is 1. The quantity Δ_i can now be expressed as

$$\Delta_i = \sum_{n=1}^{\infty} [\boldsymbol{T}^n]_{i,i} \tag{12}$$

which can be interpreted as the sum over all n-step returning probabilities from node i to itself, where $n = 1, 2, \ldots, \infty$. In light of (12), we can rewrite the EoC of any agent i as

$$\frac{1}{1 + \Delta_i} = 1 - F_{ii} \tag{13}$$

where F_{ii} is the *hitting probability* from node i to itself [70, Sec. 1.6]. To elaborate a little further, F_{ii} denotes the probability that the random walk starting from i would ever return to i. Formally, we have

$$F_{ii} := \mathbb{P}\{\mathsf{x}_t = i, \ \exists t \geqslant 0 | \mathsf{x}_0 = i\}$$

where x_t denotes the state of the random walk at time step t. The equation (13) is a classical result in the random walk theory [70, Sec. 1.5].

Another important fact about one-dimensional random walks is that they are always *recurrent* when there is no absorbing states [71, Sec. 28], meaning that we have $F_{ii} = 1$ for all $i \in \mathcal{N}_{\mathrm{a}}$ if there is not any anchor in the network. This implies that when there are some anchors, the hitting probability F_{ii} satisfies

$$F_{ii} + \sum_{j \in \mathcal{N}_{\mathrm{b}}} F_{\mathcal{N}_{\mathrm{b}} \cup \{i\}}(i, j) = 1 \tag{14}$$

where $F_{\mathcal{N}_{\mathrm{b}} \cup \{i\}}(i, j)$ denotes the probability that the random walk starting from i would ever reach the set $\mathcal{N}_{\mathrm{b}} \cup \{i\}$ at least once, and that it is at state j when it reaches $\mathcal{N}_{\mathrm{b}} \cup \{i\}$ for the first time. In other words, (14) can be interpreted as that a random walk starting from i would always reach the set $\mathcal{N}_{\mathrm{b}} \cup \{i\}$ at least once.
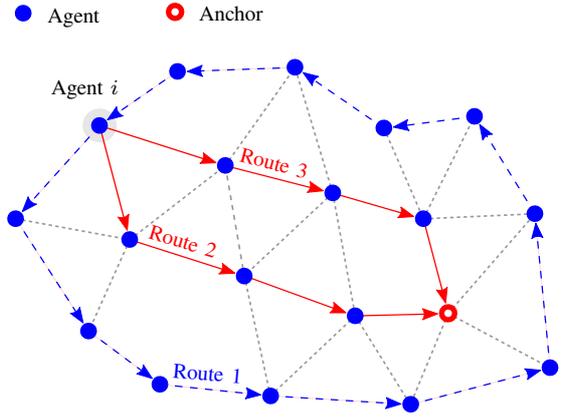

Fig. 2. Graphical illustration of the "position information routing" interpretation. Here the ineffective route (route 1) does not include anchors, and hence does not provide agent i with any position information. However, effective routes (routes 2 and 3) do provide position information.

Combining (13) and (14), we have

$$\frac{1}{1 + \Delta_i} = \sum_{j \in \mathcal{N}_{\mathrm{b}}} F_{\mathcal{N}_{\mathrm{b}} \cup \{i\}}(i, j) \tag{15}$$

and hence

$$d_i = \underbrace{d_i \sum_{j \in \mathcal{N}_{\mathrm{b}}} F_{\mathcal{N}_{\mathrm{b}} \cup \{i\}}(i, j)}_{\boldsymbol{J}_{\mathrm{e}}(\boldsymbol{p}_i)} + d_i F_{ii}. \tag{16}$$

Intuitively, we may now interpret EoC as the "position information routing" efficiency of an agent. To elaborate, (16) shows that the NPI of agent i consists of two parts: effective information (i.e., $\boldsymbol{J}_{\mathrm{e}}(\boldsymbol{p}_i)$) coming from effective routes that start from agent i and arrive at anchors before returning to agent i, and ineffective information coming from ineffective routes that returns to agent i before visiting any anchor, as portrayed in Fig. 2. Since agent i does not have *a priori* position information, it cannot provide itself any information via ineffective routes. By contrast, anchors do provide position information via effective routes.

C. Random-Walk-Inspired Formalism: The 2-D Scenario

In this subsection, we generalize the discussions in Sec. III-B to the more practical 2-D scenario, and develop the "position information routing" interpretation of EoC. To elaborate, the quantities $\boldsymbol{D}_i$, $\boldsymbol{J}_{ij}$, $\boldsymbol{T}_{ij}^{(1)}$ and $\boldsymbol{\Delta}_i$ have a strong resemblance to their 1-D counterparts, except that they are 2×2 matrices. In particular, the term $\boldsymbol{T}_{ij}^{(1)}$ satisfies the (slightly modified) normalization property of $\sum_{j \in \mathcal{N}_i} \boldsymbol{T}_{ij}^{(1)} = \boldsymbol{I}_2$. However, it does not satisfy the non-negativity property. Although $\boldsymbol{T}_{ij}^{(1)}$ does not constitute a properly defined probability measure, in the following discussions, we will show that many results in the classical random walk theory, including (13), can be obtained in a similar form by treating $\boldsymbol{T}_{ij}^{(1)}$ as *pseudo-probability*.

To formalize these ideas, we will use the notations defined in Table II hereafter. Using these notations, matrix $\boldsymbol{T}_{ij}^{(n)}$ can be expressed as

$$\boldsymbol{T}_{ij}^{(n)} = \sum_{\boldsymbol{\omega} \in \Omega_n^{i,j}} \boldsymbol{T}_{\boldsymbol{\omega}}^{(n)}. \tag{17}$$

TABLE II
NOTATIONS RELATED TO THE PSEUDO-PROBABILITIES.

Notation	Definition	Interpretation
$\Omega_n^{i,j}$	$\{\boldsymbol{\omega}\|\boldsymbol{\omega} \in \mathbb{Z}^{n+1}, \omega_1 = i, \omega_{n+1} = j, \omega_k \in \mathcal{N}_a \cup \mathcal{N}_b, \ k = 1, 2, \ldots, n+1\}$	The set of all possible sequences of node indices of length $n+1$ that start with i and end with j.
$\Omega_n^{i,j}(\mathcal{S})$	$\{\boldsymbol{\omega}\|\boldsymbol{\omega} \in \Omega_n^{i,j}, \ \omega_k \notin \mathcal{S} \ \forall 2 \leqslant k \leqslant n\}$	A subset of $\Omega_n^{i,j}$ comprising the sequences that exclude the nodes in the set $\mathcal{S}$ (except i and j).
$\boldsymbol{T}_{\boldsymbol{\omega}}^{(n)}$	$\boldsymbol{T}_{\omega_1\omega_2}^{(1)} \ldots \boldsymbol{T}_{\omega_n\omega_{n+1}}^{(1)}$	The pseudo-probability of a random walk in the *original* network travelling from i to j following path $\boldsymbol{\omega}$ given by $\boldsymbol{\omega} = [\omega_1 \ \omega_2 \ \ldots \ \omega_{n+1}]^{\mathrm{T}}$.
$\widetilde{\boldsymbol{T}}_{\boldsymbol{\omega}}^{(n)}$	$\widetilde{\boldsymbol{T}}_{\omega_1\omega_2}^{(1)} \ldots \widetilde{\boldsymbol{T}}_{\omega_n\omega_{n+1}}^{(1)}$	The pseudo-probability of a random walk in the *auxiliary* network travelling from i to j following the path $\boldsymbol{\omega}$.

Furthermore, we may have the following result by simply imitating (13).

Proposition 1: The EoC of agent i $(\boldsymbol{I} + \boldsymbol{\Delta}_i)^{-1}$ can be rewritten as (as long as it exists)

$$(\boldsymbol{I} + \boldsymbol{\Delta}_i)^{-1} = \boldsymbol{I} - \boldsymbol{F}_{ii} \tag{18}$$

where $\boldsymbol{F}_{ij}$ is defined as $\boldsymbol{F}_{ij} := \sum_{n=1}^{\infty} \boldsymbol{F}_{ij}^{(n)}$ with

$$\boldsymbol{F}_{ij}^{(n)} := \sum_{\boldsymbol{\omega} \in \Omega_n^{i,j}(\{j\})} \boldsymbol{T}_{\boldsymbol{\omega}}^{(n)}. \tag{19}$$

Proof: Please refer to Appendix II-A. □

Apparently, $\boldsymbol{F}_{ij}$ is an analogue of the hitting probability, and hence will be referred to as the *hitting pseudo-probability* hereafter. Following a similar line of reasoning as in the one-dimensional case, next we will show the recurrence of the "pseudo-random walk" in the absence of anchors. To formulate this idea, we define the an auxiliary network of the original network, in which there is no anchor.

Definition 2 (Auxiliary Network): The auxiliary network is obtained by treating all anchors in the original network as agents, for which the EFIM takes the following form

$$\widetilde{\boldsymbol{J}}_e(\boldsymbol{p}_{\mathrm{all}}) = \widetilde{\boldsymbol{D}} - \widetilde{\boldsymbol{A}} \in \mathbb{R}^{2(N_a+N_b) \times 2(N_a+N_b)} \tag{20}$$

where $\widetilde{\boldsymbol{D}} := \mathrm{diag}(\boldsymbol{D}_1, \ldots, \boldsymbol{D}_{N_a+N_b})$ is an analogue of $\boldsymbol{D}$ but includes the NPIs of nodes in $\mathcal{N}_b$ (previously anchors).

Similar to the matrix $\boldsymbol{T}$ in the original network, in the auxiliary network we define

$$\widetilde{\boldsymbol{T}} = \widetilde{\boldsymbol{D}}^{-1}\widetilde{\boldsymbol{A}}. \tag{21}$$

With these definitions we are able to show the recurrence of the random walk over auxiliary networks.

Proposition 2 (Recurrence of Finite Auxiliary Networks): For any network in which there is no anchor (including the auxiliary network) with $N_a + N_b < \infty$, if the graph $\mathcal{G}_{\mathrm{net}}$ is connected, we have $\boldsymbol{F}_{ii} = \boldsymbol{I}_2$ for all $i \in \mathcal{N}_a$.

Proof: Please refer to Appendix II-B. □

Similar to the formula $F_{ii} = 1$ in the one-dimensional case, $\boldsymbol{F}_{ii} = \boldsymbol{I}_2$ can be interpreted as "the random walk starting from i returns almost surely", in the sense of pseudo-probability. Moreover, with the following definition

$$\boldsymbol{F}_{\mathcal{S}}(i,j) = \sum_{n=1}^{\infty} \sum_{\boldsymbol{\omega} \in \Omega_n^{i,j}(\mathcal{S})} \widetilde{\boldsymbol{T}}_{\boldsymbol{\omega}}^{(n)} \tag{22}$$

we may obtain

$$\boldsymbol{F}_{ii} + \sum_{j \in \mathcal{N}_b} \boldsymbol{F}_{\widetilde{\mathcal{R}}_i}(i,j) = \boldsymbol{I}_2, \ \forall i \in \mathcal{N}_a \tag{23}$$

using Proposition 1 and 2, where $\widetilde{\mathcal{R}}_i := \{i\} \cup \mathcal{N}_b$. Thus the EFIM for $\boldsymbol{p}_i$ can be written as

$$\boldsymbol{J}_e(\boldsymbol{p}_i) = \sum_{j \in \mathcal{N}_b} \boldsymbol{D}_i \boldsymbol{F}_{\widetilde{\mathcal{R}}_i}(i,j), \ \forall i \in \mathcal{N}_a. \tag{24}$$

Now we can see that in the 2-D scenario, we may have a similar "position information routing" interpretation of EoC. Indeed, using (22) and the definitions in Table II, we can derive the following result from (24)

$$\boldsymbol{D}_i = \underbrace{\sum_{j \in \mathcal{N}_b} \sum_{n=1}^{\infty} \sum_{\boldsymbol{\omega} \in \Omega_n^{i,j}(\widetilde{\mathcal{R}}_i)} \boldsymbol{D}_i \widetilde{\boldsymbol{T}}_{\boldsymbol{\omega}}^{(n)}}_{\boldsymbol{J}_e(\boldsymbol{p}_i)} + \sum_{n=1}^{\infty} \sum_{\boldsymbol{\omega} \in \Omega_n^{i,i}(\widetilde{\mathcal{R}}_i)} \boldsymbol{D}_i \widetilde{\boldsymbol{T}}_{\boldsymbol{\omega}}^{(n)} \tag{25}$$

which takes a similar form as (16).

IV. LARGE-SCALE NETWORKS

In this section, we develop asymptotic expressions for EoC (hence EFIM) in certain types of large-scale networks, using the method introduced in Sec. III. Especially, we are interested in the asymptotic EoC of agents as their distances to the nearest anchors increase. Intuitively, EoC should be decreasing with the distance to the nearest anchors. This may be interpreted as an *error propagation behavior*, which is a major issue in cooperative networks. To elaborate, when anchors are spatially sparse, the agents being next to anchors are likely to have the best localization accuracy, while those being far from anchors may perform worse since they have to rely entirely on neighboring agents that are themselves subject to position uncertainty.

For simplicity, in this section we consider the case where nodes are equipped with UOAs when $N_t \geqslant 2$, and hence the orientations of nodes do not have an impact on the results.[3]

A. Large Lattice Networks

We first investigate the asymptotic EoC in large lattice networks. The notion "lattice networks" considered here refers to networks in which agents are located in a connected subset of the space $\mathbb{Z}^2$. They may be regarded as highly simplified models of real-world large-scale wireless networks, but one can typically obtain insightful results therefrom due

---

[3]For UOAs, this implies $N_t \geqslant 3$. When nodes are equipped with other types of arrays, if their orientations can be modeled as independently uniform distributed random variables, the results in this section can still serve as looser lower bounds of the expected position MSE due to Jensen's inequality.

to their symmetry. In particular, by saying a lattice network is "large", we mean a network satisfying the following technical conditions.

Condition 1: We consider a finite lattice network within an expanding rotational symmetric region centered at the origin $\mathbf{0}$. The network tends to an infinite lattice network in the sense that $D_{\text{net}} \to \infty$, where $D_{\text{net}} := \max_{\{j,k\} \subseteq \mathcal{N}_a \cup \mathcal{N}_b} \|\boldsymbol{d}_{jk}\|$ is the network diameter. The centroid of the network does not change as it expands. Whenever we say the distance between two nodes $\|\boldsymbol{d}_{ij}\| \to \infty$, its rate of growth is lower than that of D_{net}, i.e., $\|\boldsymbol{d}_{ij}\| = o(D_{\text{net}})$.

To elaborate on our method of analysis, we take a brief look back on the 1-D toy example. For an infinitely large lattice network, it is known that the n-step transition probability $T_{ij}^{(n)}$ tends to be the probability density function of a Gaussian distribution as $n \to \infty$ [71, Sec. 7], taking the following form

$$T_{ij}^{(n)} \sim \frac{1}{\sqrt{2\pi n\sigma^2}} \exp\left\{ -\frac{1}{2n\sigma^2}(p_i - p_j)^2 \right\} \quad (26)$$

where σ^2 is a constant. The desired asymptotic behavior of EoC can then be derived by calculating the asymptotic hitting probabilities with the help of (26).

In its essence, (26) is a result of the celebrated central limit theorem, which may be derived using the technique of characteristic function [71, Sec. 7]. In light of this, we may follow a similar line of reasoning in the analysis of the 2-D scenario, while bearing in mind some technical differences between transition probabilities and pseudo-probabilities. To prevent the paper from being crammed with technical details that are not directly related to the main topic, here we will only present the main results, while the technicalities will be deferred to Appendix III.

We first present the simple case where there is only a single anchor ν located at the origin of the network, namely we have $\boldsymbol{p}_\nu = \mathbf{0}$. Using the notations introduced in Sec. III-C, we have $\mathcal{R}_i = \{\nu, i\}$ for any agent i.

1) Single Anchor, $N_t \geqslant 2$: When $N_t \geqslant 2$, each agent is able to obtain both angular and distance information from its neighbors. In this case, we have the following result on the asymptotic behavior of EFIM.

Proposition 3 (Position Information Path Loss): For lattice networks satisfying Condition 1, when ν is the only anchor, for the $N_t \geqslant 2$ case, the SPEB of agent i scales as

$$\text{sp}\{\boldsymbol{p}_i\} = \Theta(\log \|\boldsymbol{d}_{i\nu}\|) \quad (27)$$

as $\|\boldsymbol{d}_{i\nu}\| \to \infty$, if $\|\boldsymbol{d}_{i\nu}\| = o(D_{\text{net}}/2 - \|\boldsymbol{p}_i\|)$.

Proof: Please refer to Appendix IV-A. □

Proposition 3 implies that the position information an agent acquired decreases logarithmically with the distance to the anchor in an asymptotic regime, as long as agent i is closer to the anchor than the network edge (as implied by the condition $\|\boldsymbol{d}_{i\nu}\| = o(D_{\text{net}}/2 - \|\boldsymbol{p}_i\|)$). This may be viewed as a law of *position information path loss*. An interesting fact is that the path loss of the receiving signal power does not have an effect on the order of position information path loss. We will revisit this fact in Sec. IV-B.

2) Single Anchor, $N_t = 1$: When $N_t = 1$, each agent is only able to receive distance information from its neighbors. Next we show that when there is only a single anchor in the network, this will prevent agents from obtaining the position information on the direction perpendicular to that from them to the anchor.

Proposition 4 (Range-only Position Information Path Loss): For lattice networks satisfying Condition 1, when ν is the only anchor, for the $N_t = 1$ case, the EFIM of agent i is a rank-1 matrix given by

$$\boldsymbol{J}_e(\boldsymbol{p}_i) = \lambda_{i\nu} \boldsymbol{J}_r(\varphi_{i\nu}). \quad (28)$$

The term $\lambda_{i\nu}$ satisfies

$$\lambda_{i\nu} = \Theta\left((\log \|\boldsymbol{d}_{i\nu}\|)^{-1}\right) \quad (29)$$

as long as $\|\boldsymbol{d}_{i\nu}\| = o(D_{\text{net}}/2 - \|\boldsymbol{p}_i\|)$.

Proof: Please refer to Appendix IV-B. □

The fact that $\boldsymbol{J}_e(\boldsymbol{p}_i)$ is rank-1 agrees with our intuition that the network suffers from rotational ambiguity when there is only one anchor. In fact, according to the discussions in Appendix IV-B, (28) holds in any network (i.e., not limited to lattice networks) where there is only a single anchor ν, while the asymptotic behavior of $\lambda_{i\nu}$ may vary.

3) Uniformly Distributed Anchors: Next we investigate the scenario where anchors are uniformly distributed on a lattice with constant density λ_{anc}. Specifically, the positions of anchors take the following form

$$\boldsymbol{p}_i = (\lambda_{\text{anc}})^{-\frac{1}{2}}[k \ l]^{\text{T}}, \ k \in \mathbb{Z}, \ l \in \mathbb{Z}, \ \forall i \in \mathcal{N}_b. \quad (30)$$

Since anchors are spread out over the entire network area, it may not be insightful to investigate the SPEB of a certain agent, and hence we consider the average SPEB of agents.

To relate the previous results in the single anchor case to the multi-anchor scenario considered here, we first conceive a simple EFIM lower bound. Specifically, for any agent i, the projection of its EFIM on an arbitrary direction corresponding to the unit vector $\boldsymbol{u}$, i.e., $\boldsymbol{u}^{\text{T}} \boldsymbol{J}_e(\boldsymbol{p}_i) \boldsymbol{u}$, can be bounded from below by

$$\boldsymbol{u}^{\text{T}} \boldsymbol{J}_e(\boldsymbol{p}_i) \boldsymbol{u} \geqslant \max_{j \in \mathcal{N}_b} \boldsymbol{u}^{\text{T}} \boldsymbol{J}_e^{(j)}(\boldsymbol{p}_i) \boldsymbol{u} \quad (31)$$

where the notation $\boldsymbol{J}_e^{(j)}(\boldsymbol{p}_i)$ denotes the EFIM when anchor j is the only anchor in the network. To elaborate, it is clear that introducing new anchors into the network leads to a non-negative contribution to $\boldsymbol{J}_e(\boldsymbol{p})$, and hence to $\boldsymbol{J}_e(\boldsymbol{p}_i)$, $\forall i \in \mathcal{N}_a$. Therefore, using the definition of positive-semidefiniteness we have

$$\boldsymbol{u}^{\text{T}}\left(\boldsymbol{J}_e(\boldsymbol{p}_i) - \max_{j \in \mathcal{N}_b} \boldsymbol{J}_e^{(j)}(\boldsymbol{p}_i)\right)\boldsymbol{u} \geqslant 0$$

yielding (31). With the help of (31), we generalize the results in the single anchor case as follows.

Proposition 5: For lattice networks satisfying Condition 1, when anchors are distributed as (30), the average SPEB of agents in the "interior area" $\mathcal{I}_\epsilon$ of the network scales as

$$\frac{1}{|\mathcal{I}_\epsilon|} \sum_{\boldsymbol{p}_i \in \mathcal{I}_\epsilon} \text{sp}\{\boldsymbol{p}_i\} = \Theta(\log \lambda_{\text{anc}}^{-1}) \quad (32)$$

as $\lambda_{\mathrm{anc}} \to 0_+$ for any constant $\epsilon > 0$, if $\lambda_{\mathrm{anc}}^{-\frac{1}{2}} = o(D_{\mathrm{net}})$. The area $\mathcal{I}_\epsilon$ is defined as $\mathcal{I}_\epsilon := \{\boldsymbol{p}_i | \|\boldsymbol{p}_i\| < (1-\epsilon)D_{\mathrm{net}}/2, i \in \mathcal{N}_{\mathrm{a}}\}$.

Proof: Please refer to Appendix IV-C. $\qquad\square$

Note that Proposition 5 is valid for both $N_{\mathrm{t}} \geqslant 2$ and $N_{\mathrm{t}} = 1$. Intuitively, though a single anchor can only provide position information on a specific direction in the $N_{\mathrm{t}} = 1$ case, uniformly distributed anchors can provide position information from all directions. Therefore, in light of (31), the asymptotic average SPEB can be bounded as (32).

B. Stochastic Geometric Networks

Stochastic geometric networks are networks in which nodes are uniformly distributed within a specific region $\mathcal{B} \subseteq \mathbb{R}^2$. Nodes in such networks constitute binomial point processes [58]–[62]. Compared to lattice networks, they are better models for practical wireless networks, and have been widely applied to model cellular networks, unmanned aerial vehicle networks, and terrestrial vehicular networks [62]–[65]. In particular, agents in networks employing cooperative localization are typically mobile devices, and hence the stochastic geometric modelling is often preferable to its deterministic counterpart. However, stochastic geometric networks no longer possess the translation-invariant property we relied substantially on for proving the results in Sec. IV.

Nevertheless, some recent works indicate that large random geometric graphs (RGGs) are in some sense "asymptotically equivalent" to lattice graphs [72]. In light of this, in this subsection, we extend some of the results in Section IV-A to stochastic geometric networks. For technical reasons, we assume that the networks satisfy the following condition.

Condition 2: We consider stochastic geometric networks which remains to be connected as $|\mathcal{R}_{\mathrm{net}}| \to \infty$ with probability $1 - o(N_{\mathrm{a}}^{-1})$ as $|\mathcal{R}_{\mathrm{net}}| \to \infty$.

1) Single anchor: Consider a stochastic geometric network with a single anchor ν. Since the nodes are randomly located within the network area, it may be difficult to depict the asymptotic SPEB of any single agent. Instead, we will characterize the average SPEB of agents within a certain area. Specifically, we consider a ring-shaped region $\mathsf{S}_\nu(R, r)$ centered at $\mathbf{p}_\nu$ given by

$$\mathsf{S}_\nu(R, r) := \left\{ \boldsymbol{x} \Big| R - \frac{r}{2} \leq \|\boldsymbol{x} - \mathbf{p}_\nu\| \leqslant R + \frac{r}{2}, \; \boldsymbol{x} \in \mathbb{R}^2 \right\}$$

where r is a positive constant. Roughly speaking, $\mathsf{S}_\nu(R, r)$ contains all agents having distance approximately R to the anchor.

Theorem 2: For a stochastic geometric network satisfying Condition 2, if $R \to \infty$ as $|\mathcal{R}_{\mathrm{net}}| \to \infty$ but $RN_{\mathrm{a}}^{-1} \to 0$, when $N_{\mathrm{t}} \geqslant 2$, the average SPEB of the agents within $\mathsf{S}_\nu(R, r)$ scales as

$$\frac{1}{|\mathsf{A}_\nu(R, r)|} \sum_{\mathbf{p}_i \in \mathsf{S}_\nu(R, r)} \mathrm{sp}\{\mathbf{p}_i\} = \Theta\big(\log R\big) \tag{33}$$

with probability approaching 1 as $N_{\mathrm{a}} \to \infty$, where $\mathsf{A}_\nu(R, r)$ denote the set of agents located in $\mathsf{S}_\nu(R, r)$. For the $N_{\mathrm{t}} = 1$ case, with probability approaching 1 we also have

$$\frac{1}{|\mathsf{A}_\nu(R, r)|} \sum_{\mathbf{p}_i \in \mathsf{S}_\nu(R, r)} \big(\mathbf{u}(\varphi_{i\nu})^{\mathrm{T}} \mathbf{J}_{\mathrm{e}}(\mathbf{p}_i)\mathbf{u}(\varphi_{i\nu})\big)^{-1} = \Theta\big(\log R\big). \tag{34}$$
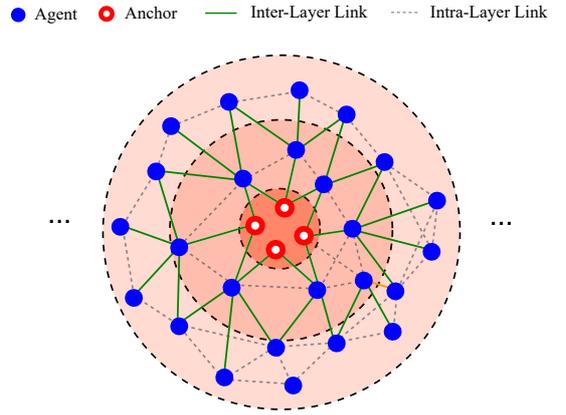

Fig. 3. Illustration of the concentric network model considered in [53]. Agents in the first layer can communicate with anchors directly, while those in the nth layer ($n > 1$) are n hops away from anchors.

Proof: Please refer to Appendix V. $\qquad\square$

In extended networks, it is known that if the agents do not cooperate, the Fisher information an anchor provides for an agent is proportional to the received signal-to-noise ratio (SNR) [46], which typically exhibits a polynomial decay as the distance between these two nodes increases due to the path loss of receiving power. In contrast, Theorem 2 indicates that with cooperation among agents, the rate of position information path loss is only *logarithmic*, showing the great superiority of cooperative localization over its non-cooperative counterpart.

2) Uniformly Distributed Anchors: Now we consider the case that anchors constitute a binomial point process in the network region $\mathcal{R}_{\mathrm{net}}$ with intensity λ_{anc}. As the network region $\mathcal{R}_{\mathrm{net}}$ expands, $\lambda_{\mathrm{anc}} \to 0$ slowly such that $\lambda_{\mathrm{anc}}^{-1} = o(|\mathcal{R}_{\mathrm{net}}|)$, namely there remains at least a constant number of anchors in the network. We assume for convenience that the network region $\mathcal{R}_{\mathrm{net}}$ is a circular area with diameter D_{net}. For such networks, we have the following result.

Proposition 6: The average SPEB of agents in the "interior area" $\mathcal{I}_\epsilon$ of the network scales as

$$\frac{1}{|\mathcal{I}_\epsilon|} \sum_{\boldsymbol{p}_i \in \mathcal{I}_\epsilon} \mathrm{sp}\{\boldsymbol{p}_i\} = O(\log \lambda_{\mathrm{anc}}^{-1}) \tag{35}$$

with probability approaching 1 as $N_{\mathrm{a}} \to \infty$, for any constant $\epsilon > 0$. The area $\mathcal{I}_\epsilon$ is defined as $\mathcal{I}_\epsilon := \{\boldsymbol{p}_i | \|\boldsymbol{p}_i\| < (1 - \epsilon)D_{\mathrm{net}}/2, i \in \mathcal{N}_{\mathrm{a}}\}$.

Proof: Please refer to Appendix VI. $\qquad\square$

For a given anchor intensity λ_{anc}, the distance from a "typical agent" to the nearest anchor would be at the order of $\lambda_{\mathrm{anc}}^{-\frac{1}{2}}$. In light of this, the logarithmic error scaling in Proposition 6 follows intuitively from Theorem 2, since the average SPEB of "typical agents" would increase logarithmically with their distances to the nearest anchors.

V. Discussions

In this section we discuss the practical implications of our results, and relate them to some existing works.

A. Error Propagation in Concentric Multi-hop Wireless Network Model

In the paper [53], the performance limit of cooperative localization in a concentric multi-hop wireless network model (for which $N_{\mathrm{t}} = 1$) is investigated using methods from stochastic geometry. In this network model, anchors locate in the innermost circular area of the network, while agents locate in the outer ring-shaped area, as illustrated in Fig. 3. It is shown in [53] that the expected SPEB of an agent i scales as $\Theta(\|\boldsymbol{d}_{i\bar{\mathrm{A}}}\|^2)$, where $\boldsymbol{d}_{i\bar{\mathrm{A}}}$ is the distance from $\boldsymbol{p}_i$ to the network center.

Using the methods proposed in this paper, we can obtain further insights into the asymptotic localization performance in this network. More precisely, from (31) and Theorem 2, we see that the expected DPEB on the direction from $\boldsymbol{p}_i$ to the network center scales as $\Theta(\log \|\boldsymbol{d}_{i\bar{\mathrm{A}}}\|)$, which is a far lower increasing rate than $\Theta(\|\boldsymbol{d}_{i\bar{\mathrm{A}}}\|^2)$ in [53]. This implies that the DPEB on the orthogonal direction scales as $\Theta(\|\boldsymbol{d}_{i\bar{\mathrm{A}}}\|^2)$. Intuitively, since anchors are confined within a small region, the network is subject to large rotational uncertainty, which leads to the $\Theta(\|\boldsymbol{d}_{i\bar{\mathrm{A}}}\|^2)$ error scaling.

In summary, for concentric networks where $N_{\mathrm{t}} = 1$, when our main interest is to estimate the distance from agent i to the network center, we can expect an error scaling of $\Theta(\log \|\boldsymbol{d}_{i\bar{\mathrm{A}}}\|)$ instead of $\Theta(\|\boldsymbol{d}_{i\bar{\mathrm{A}}}\|^2)$.

B. Trade-off Between Communication Complexity and Localization Accuracy

Typical distributed cooperative localization algorithms include Bayesian message-passing-based methods [18] and distributed convex optimization-based methods [20]. These methods may attain near-optimal localization accuracy, but they generally require iterative information exchange among agents, leading to considerable delay and communication overhead.

By contrast, sequential multi-hop localization methods, as exemplified by DV-hop-based methods [54], work in a layer-by-layer manner and do not require iterations. Basically, DV-hop-based methods consist of a distance estimation stage and a position estimation stage. In the first stage, the distances between each pair of agent and anchor that are n hops from each other are approximated by the sum of all "raw" distance estimates in each hop. The "raw" distance estimates can be obtained by ranging techniques or from range-free metrics such as average hop-distance [55] and number of common neighbors [56]. In the second stage, the positions of all agents are estimated locally based on the anchor-agent distance estimates via standard non-cooperative localization algorithms, e.g., weighted least squares.

Naturally, it is of great interest whether the relative simplicity of DV-hop-based methods comes with the cost of accuracy degradation. For simplicity, we assume that the initial distance estimates for each hop have errors with constant variance, and are mutually independent. Thus for any agent i, the MSEs of the final distance estimates are at least proportional to the number of hops between i and the nearest anchors. This implies that for a given anchor density λ_{anc}, the average localization MSE is at least linear in $\lambda_{\mathrm{anc}}^{-\frac{1}{2}}$. By contrast, from

Proposition 6 we see that for the same λ_{anc}, the optimal localization MSE is at most logarithmic in $\lambda_{\mathrm{anc}}^{-1}$.

In summary, in large networks where anchors are sparse, the localization accuracy of DV-hop-based methods are far from optimal, in the sense that their average localization MSE scales as $\Omega(\lambda_{\mathrm{anc}}^{-\frac{1}{2}})$, as opposed to the optimal scaling $O(\log \lambda_{\mathrm{anc}}^{-1})$. This can be regarded as an accuracy penalty due to their low communication overhead. For a more in-depth analysis on the trade-off between communication complexity and localization accuracy, one may have to resort to rate-distortion methods [73].

C. Reducing the Complexity of Network Operation Techniques in Massive Cooperative Localization Networks

Network operation techniques, including power allocation and node placement methods, have been introduced to cooperative localization networks to improve the localization performance [43], [44]. For massive networks, the pseudo-probability analysis in Appendix III may be used to derive low-complexity optimization algorithms for network operation tasks.

In particular, for a stochastic geometric network, a general network operation problem may be formulated as follows

$$\min_{\boldsymbol{x}} \ C(\boldsymbol{x}) := \mathbb{E}_{\mathbf{p}}\{\mathrm{tr}\left\{\mathbf{J}_{\mathrm{e}}^{-1}(\mathbf{p})\right\}\} \qquad (36)$$

where $\boldsymbol{x}$ denotes the parameters to be optimized, which corresponds to the power allocation coefficients in power allocation problems, or the positions of anchors in anchor placement tasks. In practice, one may approximate the expectation operation in (36) by Monte Carlo sampling, and obtain the following objective function

$$\widetilde{C}(\boldsymbol{x}) = \frac{1}{N_{\mathrm{s}}} \sum_{n=1}^{N_{\mathrm{s}}} \mathrm{tr}\left\{\boldsymbol{J}_{\mathrm{e}}^{-1}\big(\boldsymbol{p}^{(n)}\big)\right\} \qquad (37)$$

where $\boldsymbol{p}^{(n)}$ denotes the n-th sample of the random vector $\mathbf{p}$, and N_{s} represents the number of samples. The minimization of $\widetilde{C}(\boldsymbol{x})$ can then be solved by means of gradient-based methods, or gradient-free methods such as the Nelder-Mead algorithm [74]. In either case, we assume that the optimization algorithm requires K evaluations of the objective function $\widetilde{C}(\boldsymbol{x})$.[4] A naive method is to directly compute the objective function (37), which involves inversion of the $2N_{\mathrm{a}} \times 2N_{\mathrm{a}}$ matrices $\boldsymbol{J}_{\mathrm{e}}^{-1}\big(\boldsymbol{p}^{(n)}\big)$. In this case, the computational complexity of the entire optimization algorithm is thus $O(KN_{\mathrm{s}}N_{\mathrm{a}}^3)$, which could be prohibitive for massive networks having large N_{a}.

Alternatively, we may rearrange the objective function as

$$\begin{aligned}
\widetilde{C}(\boldsymbol{x}) &= \frac{1}{N_{\mathrm{s}}} \sum_{n=1}^{N_{\mathrm{s}}} \sum_{i=1}^{N_{\mathrm{a}}} \mathrm{tr}\left\{\boldsymbol{J}_{\mathrm{e}}^{-1}\big(\boldsymbol{p}_i^{(n)}\big)\right\} \\
&= \frac{1}{N_{\mathrm{s}}} \sum_{n=1}^{N_{\mathrm{s}}} \sum_{i=1}^{N_{\mathrm{a}}} \mathrm{tr}\left\{\boldsymbol{D}_i^{(n)}(\boldsymbol{I} - \boldsymbol{F}_{ii}^{(n)})\right\}
\end{aligned} \qquad (38)$$

where the notations having superscript (n) denote the corresponding quantities in the n-th network sample. In the

---

[4]In gradient-based methods, the gradient can be approximately computed using a fixed number of objective function evaluations.

following discussion, we will omit the superscript when there is no confusion. Under this formulation, the complexity is given by $O(KN_sN_aF)$, where F denotes the complexity of computing the matrix $\boldsymbol{F}_{ii}$. According to Theorem 3 in Appendix III-B, the complexity of computing $\boldsymbol{F}_{ii}$ is dominated by computing the $(2N_b + 1) \times (2N_b + 1)$ matrix $\boldsymbol{P}_{\widetilde{\mathcal{R}}_i}^{-1}$. Note that the matrix $\boldsymbol{P}_{\widetilde{\mathcal{R}}_i}$ can be partitioned as

$$
\boldsymbol{P}_{\widetilde{\mathcal{R}}_i} = \begin{bmatrix} [\boldsymbol{P}_{\widetilde{\mathcal{R}}_i}]_{1,1} & [\boldsymbol{P}_{\widetilde{\mathcal{R}}_i}]_{1,2:N_b+1} \\ [\boldsymbol{P}_{\widetilde{\mathcal{R}}_i}]_{2:N_b+1,1} & \boldsymbol{P}_{\mathcal{N}_b} \end{bmatrix}.
$$

The complexity of computing $\boldsymbol{P}_{\widetilde{\mathcal{R}}_i}^{-1}$ can be reduced by exploiting the partitioned structure. Specifically, the inverse of the matrix $\boldsymbol{P}_{\mathcal{N}_b}$ can be computed in advance, and reused when computing $\boldsymbol{P}_{\widetilde{\mathcal{R}}_i}$ for different i. Finally, for massive networks, the 2×2 blocks in $\boldsymbol{P}_{\widetilde{\mathcal{R}}_i}$ may be approximated by the asymptotic formula in Proposition 9. We now see that the complexity of the entire optimization algorithm is given by $O[KN_s(N_aN_b^2 + N_b^3)]$, which is substantially lower than the $O(KN_sN_a^3)$ complexity of the naive approach when $N_b \ll N_a$.

VI. NUMERICAL RESULTS

In this section we illustrate some results we obtained with numerical examples.

A. Illustration of the Proposed Scaling Laws

In this subsection, we provide numerical results demonstrating the scaling laws of localization error discussed in Sec. IV. Unless stated otherwise, the parameters take the following values by default in all simulations. The path loss exponent $\gamma = 3$, signal bandwidth $\beta = 10\,\mathrm{MHz}$, the ranging coefficient (defined after (4)) $\zeta_{ij} = 1$ for all $i \leftrightarrow j$, the carrier frequency $f_c = 2\,\mathrm{GHz}$, and the SNR measured at $1\mathrm{m}$ from the transmitting antenna is $40\,\mathrm{dB}$. All nodes locate in a disk with diameter D_{net}, and the node density is $2.5 \times 10^{-3}\,\mathrm{m}^{-2}$. For the $N_t \geqslant 2$ case, we consider the scenario where $N_t = 3$ and antennas are arranged in a uniform circular array with diameter $0.3\mathrm{m}$. Receivers on the nodes are able to detect signals when $\mathrm{SNR} \geqslant -15\,\mathrm{dB}$.

First we investigate the law of position information path loss in both lattice and stochastic geometric networks under the $N_t \geqslant 2$ case. Fig. 4 illustrates the SPEB of agents in the single anchor scenario. A single anchor locates at the network centroid $(0\mathrm{m}, 0\mathrm{m})$, and the network diameter D_{net} varies from $500\mathrm{m}$ to $2000\mathrm{m}$. For stochastic geometric networks, the results are averaged over 1000 network snapshots. As implied by Proposition 3 and Theorem 2, in both types of networks, the SPEB of agents grows logarithmically as the distance from the anchor increases, except that there is sudden increment in SPEB at the network edge since the agents at network edge can only communicate with fewer neighbors.

To further illustrate the network boundary effect, we consider the case where there are "holes" in the network region that do not contain any node. This could be related to practical scenarios where the network region contains large obstacles. In particular, we consider a stochastic geometric network with
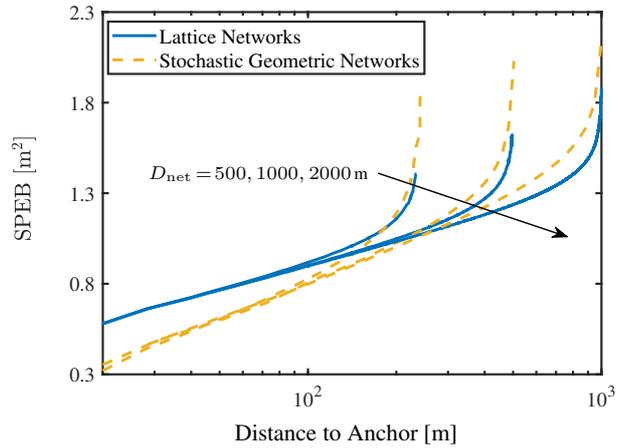

Fig. 4. SPEB as a function of the distance to the single anchor in lattice and stochastic geometric networks under the $N_t \geqslant 2$ case for varying network diameter D_{net}.
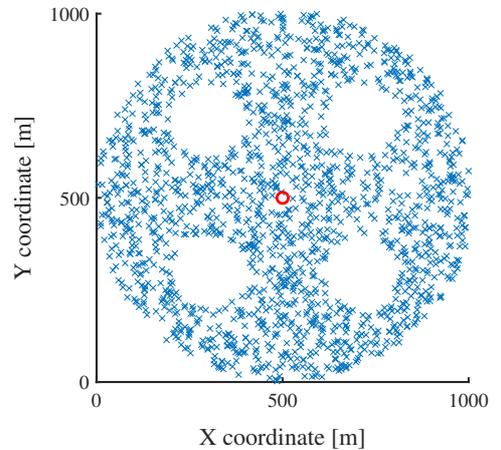
Fig. 5. Illustration of a stochastic geometric network containing four circular "holes", each with radius $r_h = 100\mathrm{m}$. The crosses represent agents, while a single anchor marked by a circle locating at the center of the network region.

$D_{\mathrm{net}} = 1000\mathrm{m}$ containing four circular holes centered at $(300\mathrm{m}, 300\mathrm{m})$, $(300\mathrm{m}, 700\mathrm{m})$, $(700\mathrm{m}, 300\mathrm{m})$, and $(700\mathrm{m}, 700\mathrm{m})$, respectively. Each hole has radius r_h. A single anchor locates at the network centroid $(500\mathrm{m}, 500\mathrm{m})$. An instance of the network is shown in Fig. 5. Under previous assumptions, we portray the relationship between the SPEB of agents and their distances to the anchor is shown in Fig. 6. We observe that when the distance to the anchor is larger than a threshold (around $150\mathrm{m}$), the SPEB of agents in networks containing holes starts to deviate from that of the network without holes. Especially, the deviation is exacerbated when r_h is larger, since the boundary effect caused by the holes becomes stronger as r_h increases.

Next we illustrate the asymptotic behavior of network average SPEB in the uniformly distributed anchor scenario. We consider the $N_t = 1$ case in both lattice and stochastic geometric networks. In lattice networks, anchors are distributed according to (30), while in stochastic geometric networks they form a binomial point process with the same intensity λ_{anc}. For such networks with diameter $D_{\mathrm{net}} = 2000\mathrm{m}$, the average

Fig. 6. SPEB vs. the distance to the single anchor in stochastic geometric networks with $D_{\mathrm{net}} = 1000$m under the $N_{\mathrm{t}} \geqslant 2$ case containing four circular "holes" with various radius r_{h}.

Fig. 8. SPEB as a function of the distance to the single anchor in lattice networks under the $N_{\mathrm{t}} = 2$ case for varying path loss exponent γ.
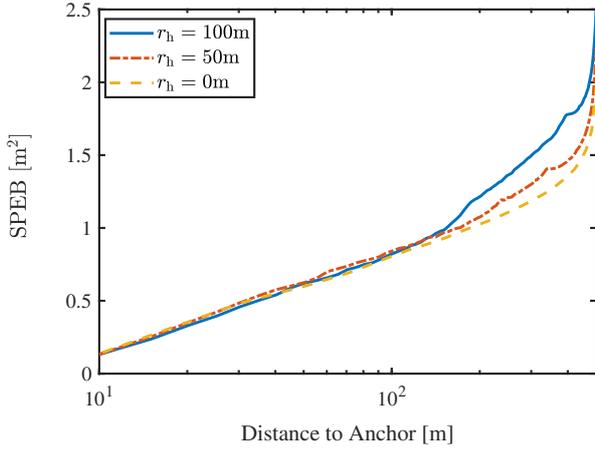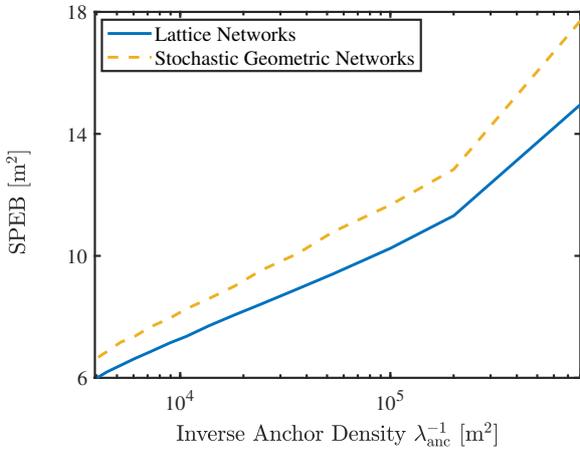

Fig. 7. Network average SPEB as a function of the reciprocal of anchor density $\lambda_{\mathrm{anc}}^{-1}$ in lattice and stochastic geometric networks where anchors are uniformly distributed under the $N_{\mathrm{t}} = 1$ case.
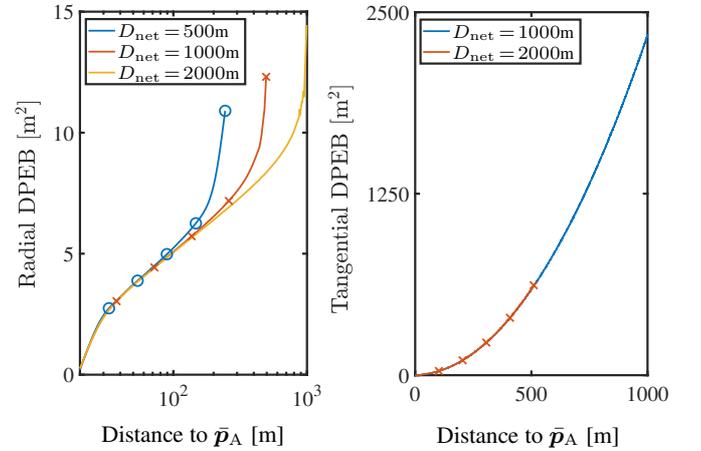
Fig. 9. The radial and tangential DPEB in lattice networks with varying network diameter D_{net} under the $N_{\mathrm{t}} = 1$ case. Left: radial DPEB; right: tangential DPEB.

SPEB of all agents in the networks are shown in Fig. 7. Results for stochastic geometric networks are averaged over 1000 network snapshots. It can be seen that the average SPEB grows logarithmically as the reciprocal of density $\lambda_{\mathrm{anc}}^{-1}$ increases, as indicated by Propositions 5 and 6. The abrupt increment of average SPEB when $\lambda_{\mathrm{anc}}^{-1}$ is large is due to the network edge effect, which is more significant when there are fewer anchors in the network.

We then demonstrate the impact of path loss exponent γ on the SPEB of agents in lattice networks. Here we focus on the single anchor scenario, and two different values of network diameter are considered: $D_{\mathrm{net}} = 1000$m and $D_{\mathrm{net}} = 2000$m. The path loss exponent γ varies from 3 to 3.5, and the SPEB as a function of the distance to the anchor is plotted in Fig. 8. Similar to Fig. 4, all SPEBs grows logarithmically as the distance to network centroid increases, except for those agents located at the network edge. However, the slope of these curves are different. As γ increases, agents receive less information from their neighbors, and hence the SPEB increases faster.

Finally, we illustrate the localization performance under the

$N_{\mathrm{t}} = 1$ case in the concentric network model discussed in Sec. V-A. We consider lattice networks in which around the network centroid, four anchors are located at $(\pm 20\mathrm{m}, 0\mathrm{m})$ and $(0\mathrm{m}, \pm 20\mathrm{m})$, respectively. The DPEBs of agents in networks with D_{net} varying from 500m to 2000m are illustrated in Fig. 9. Here, radial DPEB refers to the DPEB on the direction from an agent to the network centroid, while tangential DPEB refers to the DPEB on the orthogonal direction. It can be seen that the radial DPEB grows in a similar way as the SPEB in the $N_{\mathrm{t}} \geqslant 2$ case illustrated in Fig. 4. In addition, the tangential SPEB increases quadratically as the distance to the network centroid increases. These results are consistent with Proposition 3 and the discussions in Sec. V-A.

To better illustrate the behavior of DPEB when $N_{\mathrm{t}} = 1$, Fig. 10 shows the error ellipses in both lattice and stochastic geometric networks. For better demonstration, only part of the networks is shown in the figure. The error ellipse of agent i is defined as the set of point $\boldsymbol{x}$ satisfying

$$(\boldsymbol{x} - \boldsymbol{p}_i)^{\mathrm{T}} \boldsymbol{J}_{\mathrm{e}}(\boldsymbol{p}_i)(\boldsymbol{x} - \boldsymbol{p}_i) = 1.$$

Here the network diameter $D_{\mathrm{net}} = 500$m. Three anchors

(a) Lattice Network   (b) Stochastic Geometric Network
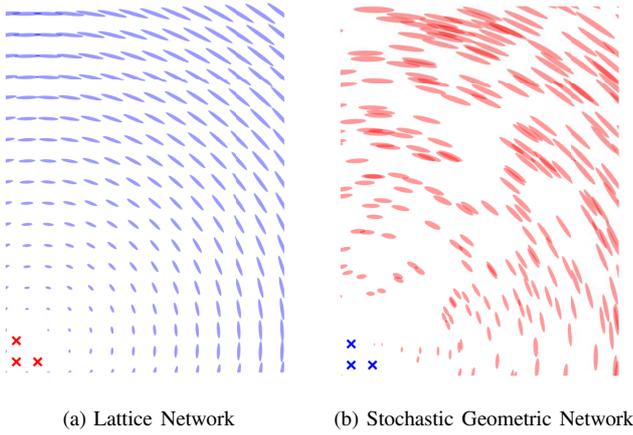
Fig. 10. Error ellipses of agents in lattice and stochastic geometric networks in the constant anchor region scenario when $N_t = 1$. Positions of anchors are marked with "x".

locate at a corner of the scene, marked by "x" in Fig. 10. From the figure it is clear that the tangential DPEB contribute most to the SPEB of agents, in either lattice networks or stochastic geometric networks. Furthermore, the length of major axes of the ellipses increase linearly with the distance to the anchors, as implied by Proposition 3 and the discussions in Sec. V-A.

B. Comparison with Actual Algorithms

In this subsection, we compare the proposed error scaling laws with actual algorithms. In particular, we consider two algorithms, including a centralized algorithm (which is expected to be near-optimal) and a sequential algorithm.

For the simulations presented in this subsection, we consider the practical scenario of a wireless sensor network employing pairwise received signal strength (RSS) and AOA measurements. Specifically, we assume that the network resides in a disk with diameter $D_{\mathrm{net}} = 400$m, in which a single anchor locates at the network center. Agents are randomly distributed in the network region with density 3×10^{-3}m^{-2}. The RSS measurements are obtained according to the channel model in the "park" scenario described in [75], given by

$$y_{ij}^{\mathrm{RSS}} = P_{\mathrm{T}} - L_0 - 10\gamma \log_{10}(\|\boldsymbol{d}_{ij}\|d_0^{-1}) + n_{\mathrm{shadow}} \quad (39)$$

where $\boldsymbol{d}_{ij}$ is the displacement vector between node i and j, $P_{\mathrm{T}} = 5$dBm denotes the transmitting power, $L_0 = 31$dB denotes the path loss at the reference distance $d_0 = 1$m, $\gamma = 3.69$ represents the path loss exponent [75], and n_{shadow} models the shadowing effect, which follows a zero-mean Gaussian distribution with standard deviation $\sigma_{\mathrm{RSS}} = 1.42$dB [75]. The bearing measurement errors are modelled as zero-mean von Mises distributed random variables with standard deviation $\sigma_{\mathrm{AOA}} = 5$ degree [76]. Each node can acquire pairwise measurements with neighbors within range $R_{\mathrm{max}} = 43$m, corresponding to a receiver sensitivity of -100dBm [75].

For the centralized algorithm, we consider an iterative approximate maximum likelihood (AML) estimator, which approximates the likelihood function corresponding to the each pairwise measurement as 2-D Gaussian distributions, with an

iterative refinement on the estimates of measurement variances. Specfically, in the ℓ-th iteration, the likelihood function for a pair of nodes i and j is approximated as follows

$$g_{\boldsymbol{\alpha}_{ij}, r_{ij}}^{(\ell)}(\alpha_{ij}, r_{ij}; \boldsymbol{p}_i, \boldsymbol{p}_j)$$
$$\propto \exp\left\{ -\frac{1}{2}(\boldsymbol{p}_i - \boldsymbol{p}_j - \hat{\boldsymbol{d}}_{ij})^{\mathrm{T}} \boldsymbol{U}_{ij} \boldsymbol{\Lambda}_{ij}^{(\ell)} \boldsymbol{U}_{ij}^{\mathrm{T}}(\boldsymbol{p}_i - \boldsymbol{p}_j - \hat{\boldsymbol{d}}_{ij}) \right\}$$
$$(40)$$

where $\hat{\boldsymbol{d}}_{ij} = r_{ij}[\cos\alpha_{ij} \ \sin\alpha_{ij}]^{\mathrm{T}}$ is an estimate of the displacement vector $\boldsymbol{d}_{ij}$ between node i and j based on the range and bearing measurements, the unitary matrix $\boldsymbol{U}_{ij}$ is given by

$$\boldsymbol{U}_{ij} = \left[\begin{array}{cc} \cos\alpha_{ij} & -\sin\alpha_{ij} \\ \sin\alpha_{ij} & \cos\alpha_{ij} \end{array} \right]$$

and the diagonal matrix $\boldsymbol{\Lambda}_{ij}^{(\ell)}$ denotes the estimate of the measurement variances at the ℓ-th iteration, given by $\mathrm{diag}\big((\hat{\sigma}_{\mathrm{r},ij}^{(\ell)})^{-2}, (\hat{\sigma}_{\mathrm{t},ij}^{(\ell)})^{-2}\big)$. The estimates $\hat{\sigma}_{\mathrm{r},ij}^{(\ell)}$ and $\hat{\sigma}_{\mathrm{t},ij}^{(\ell)}$ are given by

$$\hat{\sigma}_{\mathrm{r},ij}^{(\ell)} = \frac{\ln 10 \sigma_{\mathrm{RSS}}}{10\gamma} \cdot \left\| \hat{\mathbf{p}}_{i,\mathrm{AML}}^{(\ell-1)} - \hat{\mathbf{p}}_{j,\mathrm{AML}}^{(\ell-1)} \right\|$$
$$\hat{\sigma}_{\mathrm{t},ij}^{(\ell)} = \sigma_{\mathrm{AOA}} \cdot \left\| \hat{\mathbf{p}}_{i,\mathrm{AML}}^{(\ell-1)} - \hat{\mathbf{p}}_{j,\mathrm{AML}}^{(\ell-1)} \right\| \quad (41)$$

where $\hat{\mathbf{p}}_{i,\mathrm{AML}}^{(\ell-1)}$ denotes the position estimate of node i in the $(\ell-1)$-th iteration. For $\ell = 1$, the estimates $\hat{\sigma}_{\mathrm{r},ij}^{(1)}$ and $\hat{\sigma}_{\mathrm{t},ij}^{(1)}$ are obtained by replacing $\left\| \hat{\mathbf{p}}_{i,\mathrm{AML}}^{(\ell-1)} - \hat{\mathbf{p}}_{j,\mathrm{AML}}^{(\ell-1)} \right\|$ in (41) with the range estimate r_{ij}. The bearing estimate α_{ij} is obtained directly using the AOA measurement, while the range estimate r_{ij} is computed as

$$r_{ij} = 10^{(P_{\mathrm{T}} - L_0 - y_{ij}^{\mathrm{RSS}})(10\gamma)^{-1}} \cdot d_0.$$

After some manipulations, one could rearrange the likelihood function of $\boldsymbol{p}$ into the following standard form of Gaussian Markov random fields

$$g_{\boldsymbol{\alpha}, \boldsymbol{r}}(\boldsymbol{\alpha}, \boldsymbol{r}; \boldsymbol{p}) \propto \exp\left\{ -\frac{1}{2}\boldsymbol{p}^{\mathrm{T}}\mathbf{P}^{(\ell)}\boldsymbol{p} + (\mathbf{h}^{(\ell)})^{\mathrm{T}}\boldsymbol{p} \right\}$$

where $\mathbf{P}^{(\ell)}$ denotes the precision matrix[5], $\mathbf{h}^{(\ell)}$ represents the linear coefficients, while $\boldsymbol{\alpha}$ and $\boldsymbol{r}$ represent the vectors containing all bearing and range measurements, respectively. The closed-form AML estimate at the ℓ-th iteration can thus be obtained as

$$\hat{\mathbf{p}}_{\mathrm{AML}}^{(\ell)} = (\mathbf{P}^{(\ell)})^{-1}\mathbf{h}^{(\ell)}. \quad (42)$$

A straightforward computation of the matrix $(\mathbf{P}^{(\ell)})^{-1}$ would require centralized operations, which may be unrealistic for large networks. Fortunately, some distributed algorithms such as Gaussian belief propagation are shown to be capable of obtaining the solution in (42) [77], and hence may be viewed as distributed implementations of the AML estimator.

As for the sequential algorithm, we consider a sequential and non-iterative implementation of the aforementioned AML estimator. To elaborate, the agents that can directly obtain measurements with the anchor, which will be referred to as the "1-hop agents", estimate their positions in the first

---

[5]We are slightly abusing the notations here. Please do not confuse the precision matrix with the potential kernel discussed in Appendix III-B.
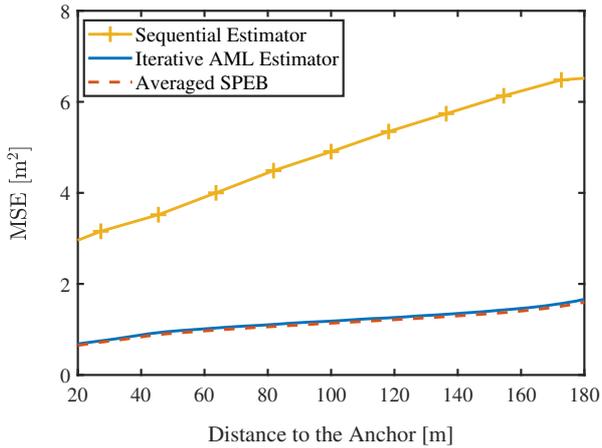
Fig. 11. The MSE of the iterative AML estimator and the sequential estimator, as functions of the distance to the single anchor located at the network centroid. The averaged SPEB is also plotted as a benchmark.

Fig. 12. Comparison between the MSE of the iterative AML estimator and the corresponding averaged SPEB, as functions of the distance to the anchor.

round. In the subsequent rounds of the algorithm, e.g. the n-th round, the n-hop agents estimate their positions based on their measurements with the $(n-1)$-hop agents that are treated as "virtual anchors". More precisely, the sequential estimator for agent i is

$$\hat{\mathbf{p}}_{i,\mathrm{Seq}} = \mathbf{P}_i^{-1} \mathbf{h}_i \qquad (43)$$

where $\mathbf{P}_i \in \mathbb{R}^{2\times 2}$ and $\mathbf{h}_i \in \mathbb{R}^2$ are constructed using the measurements between agent i and its neighboring virtual anchors. To avoid iteration, the estimate of the measurement variances of each pair of nodes i and j is fixed at its initial value $\boldsymbol{\Lambda}_{ij}^{(1)}$.

The localization MSE of an agent i based on the iterative AML estimator, as well as that of the sequential estimator, is illustrated in Fig. 11 as a function of the distance between $\boldsymbol{p}_i$ and the anchor. The results are averaged over 1000 network snapshots, and the averaged SPEB given by

$$\frac{1}{N_{\mathrm{snapshot}} \cdot N_{\mathrm{a}}} \sum_{n=1}^{N_{\mathrm{snapshot}}} \sum_{i=1}^{N_{\mathrm{a}}} \mathrm{sp}\{\boldsymbol{p}_{i,n}\}$$

is also presented as a benchmark. Here, $N_{\mathrm{snapshot}} = 1000$ is the number of snapshots, and $\boldsymbol{p}_{i,n}$ denotes the true position of agent i in the n-th snapshot. We may observe from the figure that the sequential estimator exhibits approximately a linear error scaling with respect to the distance to the network centroid (i.e. the anchor), which is far from the average SPEB. This corroborates our discussion in Sec. V-B.

To observe the performance of the iterative AML estimator more closely, we provide a zoomed-in view of the curves in Fig. 12. As it can be seen from the figure, the AML estimator exhibits a similar asymptotic behavior to the averaged SPEB in the sense that it grows logarithmically as the distance to the network centroid increases. The residual error, which appears to be constant (with respect to the distance to the network centroid), can be interpreted as a performance penalty due to the approximation.

To conclude, these simulation results suggest that the lower bounds of localization error that we proposed are likely to
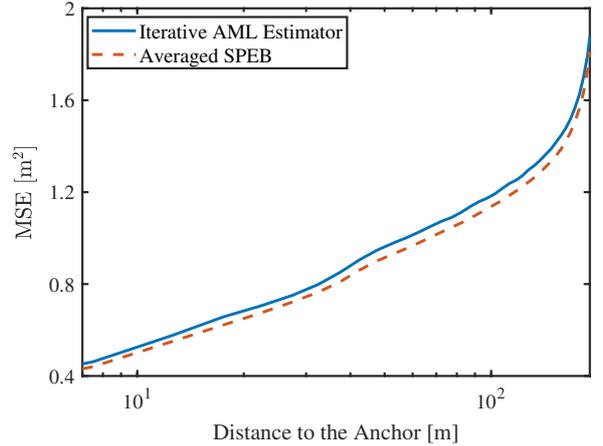
be attainable by actual algorithms. Especially, from the localization performance of the iterative AML estimator we could observe the logarithmic scaling law that we have predicted in the theoretical analysis. The sequential estimator is not capable of attaining the bounds, but it exhibits an error scaling behavior as expected in Sec. V-B.

VII. CONCLUSION

This paper provides a new approach to understanding the effect of cooperation among agents in cooperative network localization. By introducing the NPI-EoC decomposition of EFIM, we develop a random-walk-inspired formalism for the treatment of EoC. We provide an interpretation of this analysis, showing that the EoC among nodes can be regarded as the efficiency of position information routing from anchors to agents. Moreover, following this line of reasoning we show that in large lattice and stochastic geometric networks, the SPEB of an agent scales logarithmically as the reciprocal of anchor density increases, and interpret this result as a consequence of position information path loss. The scaling laws of localization performance proposed in this paper provide a new set of insights into the network localization problem from a network-level perspective, and can serve as guidelines to network deployment and operation tasks in large-scale networks.

APPENDIX I
PROOF OF THEOREM 1

Proof: Denoting $\boldsymbol{T} = \boldsymbol{D}^{-1}\boldsymbol{A}$, from (2) matrix $\boldsymbol{J}_{\mathrm{e}}(\boldsymbol{p})$ can be rewritten as

$$\boldsymbol{J}_{\mathrm{e}}(\boldsymbol{p}) = \boldsymbol{D}(\boldsymbol{I} - \boldsymbol{T}). \qquad (44)$$

Expanding $(\boldsymbol{I} - \boldsymbol{T})^{-1}$ as a matrix power series, the inverse EFIM can be expressed as

$$\boldsymbol{J}_{\mathrm{e}}^{-1}(\boldsymbol{p}) = \left(\boldsymbol{I} + \sum_{n=1}^{\infty} \boldsymbol{T}^n\right) \boldsymbol{D}^{-1}. \qquad (45)$$

The power series converges as long as the inverse of $\boldsymbol{J}_{\mathrm{e}}(\boldsymbol{p}_i)$ exists. Hence the EFIM for agent i takes the following form

$$\boldsymbol{J}_{\mathrm{e}}(\boldsymbol{p}_i) = \boldsymbol{D}_i \left(\boldsymbol{I} + \sum_{n=1}^{\infty} [\boldsymbol{T}^n]_{i,i} \right)^{-1}. \tag{46}$$

Comparing (7) and (46), it now suffices to showing that

$$\sum_{n=1}^{\infty} [\boldsymbol{T}^n]_{i,i} = \sum_{n=1}^{\infty} \boldsymbol{T}_{ii}^{(n)} \tag{47}$$

for all $i, j \in \mathcal{N}_{\mathrm{a}}$ and that $\boldsymbol{\Delta}_i \succcurlyeq \boldsymbol{0}$. Note that the matrix $\boldsymbol{T}$ can be expressed as

$$\boldsymbol{T} = \sum_{i \notin \mathcal{N}_{\mathrm{b}}} \sum_{j \leftrightarrow i} \left(\boldsymbol{E}_{ij}^{N_{\mathrm{a}}} \otimes (\boldsymbol{D}_i^{-1} \boldsymbol{J}_{ij}) \right). \tag{48}$$

We can construct another matrix $\underline{\boldsymbol{T}} \in \mathbb{R}^{2(N_{\mathrm{a}}+N_{\mathrm{b}}) \times 2(N_{\mathrm{a}}+N_{\mathrm{b}})}$ using $\boldsymbol{T}_{ij}^{(1)}$ as

$$\underline{\boldsymbol{T}} = \sum_{i \in \mathcal{N}_{\mathrm{a}} \cup \mathcal{N}_{\mathrm{b}}} \sum_{j \in \mathcal{N}_{\mathrm{a}} \cup \mathcal{N}_{\mathrm{b}}} \boldsymbol{E}_{ij}^{N_{\mathrm{a}}+N_{\mathrm{b}}} \otimes \boldsymbol{T}_{ij}^{(1)} \tag{49}$$

which can be partitioned as

$$\underline{\boldsymbol{T}} = \begin{bmatrix} \boldsymbol{T} & \sum_{i \in \mathcal{N}_{\mathrm{a}}} \sum_{j \in \mathcal{N}_{\mathrm{b}}} \boldsymbol{E}_{ij}^{N_{\mathrm{a}}, N_{\mathrm{b}}} \otimes \boldsymbol{T}_{ij}^{(1)} \\ \boldsymbol{0}_{2N_{\mathrm{b}} \times 2N_{\mathrm{a}}} & \boldsymbol{I}_{2N_{\mathrm{b}}} \end{bmatrix}. \tag{50}$$

By this construction we have $[\underline{\boldsymbol{T}}^n]_{i,j} = \boldsymbol{T}_{ij}^{(n)}$. Furthermore, from (50) we have $[\underline{\boldsymbol{T}}^n]_{1:N_{\mathrm{a}}, 1:N_{\mathrm{a}}} = \boldsymbol{T}^n$, which implies (47).

Now note that $\boldsymbol{J}_{\mathrm{e}}^{-1}(\boldsymbol{p}_i) = (\boldsymbol{I} + \boldsymbol{\Delta}_i) \boldsymbol{D}_i^{-1}$ is a real symmetric matrix, and that $\boldsymbol{J}_{\mathrm{e}}^{-1}(\boldsymbol{p}_i) \succcurlyeq \boldsymbol{D}_i^{-1}$. It then follows that $\boldsymbol{\Delta}_i \succcurlyeq \boldsymbol{0}$, which completes the proof. □

APPENDIX II
PROOF OF RESULTS IN SECTION III-C

A. Proposition 1

Proof: Expressing $\boldsymbol{T}_{ij}^{(n)}$ in terms of $\Omega_n^{i,j}$, we have

$$\begin{aligned} \boldsymbol{T}_{ij}^{(n)} &= \sum_{\boldsymbol{\omega} \in \Omega_n^{i,j}} \boldsymbol{T}_{\boldsymbol{\omega}}^{(n)} \\ &= \sum_{m=1}^{n} \sum_{\boldsymbol{\alpha} \in \Omega_m^{i,j}(\{j\})} \boldsymbol{T}_{\boldsymbol{\alpha}}^{(m)} \sum_{\boldsymbol{\beta} \in \Omega_{n-m}^{j,j}} \boldsymbol{T}_{\boldsymbol{\beta}}^{(n-m)} \\ &= \sum_{m=1}^{n} \boldsymbol{F}_{ij}^{(m)} \boldsymbol{T}_{jj}^{(n-m)}. \end{aligned} \tag{51}$$

Therefore $\boldsymbol{\Delta}_i$ can be written as

$$\boldsymbol{\Delta}_i = \sum_{m=1}^{\infty} \boldsymbol{F}_{ii}^{(m)} \sum_{n=m}^{\infty} \boldsymbol{T}_{ii}^{(n-m)} = \boldsymbol{F}_{ii}(\boldsymbol{I} + \boldsymbol{\Delta}_i). \tag{52}$$

Hence we have $\boldsymbol{I} + \boldsymbol{\Delta}_i = \boldsymbol{I} + \boldsymbol{F}_{ii}(\boldsymbol{I} + \boldsymbol{\Delta}_i)$ implying (18). □

B. Proposition 2

Proof: To prove $\boldsymbol{F}_{ii} = \boldsymbol{I}_2$, it is equivalent to show that both eigenvalues of $\sum_{n=0}^{m} \widetilde{\boldsymbol{T}}_{ii}^{(n)}$ tend to infinity as $m \to \infty$. A sufficient condition is that $\lim_{n \to \infty} \widetilde{\boldsymbol{T}}_{ii}^{(n)}$ is full-rank. From (20) and (21) we see that

$$\boldsymbol{I} - \widetilde{\boldsymbol{T}} = \widetilde{\boldsymbol{D}}^{-1} \widetilde{\boldsymbol{J}}_{\mathrm{e}}(\boldsymbol{p}_{\mathrm{all}}) \tag{53}$$

implying that the eigenspace of $\widetilde{\boldsymbol{T}}$ corresponding to eigenvalue 1 is the null space of $\widetilde{\boldsymbol{D}}^{-1} \widetilde{\boldsymbol{J}}_{\mathrm{e}}(\boldsymbol{p}_{\mathrm{all}})$. Thus, as long as matrix $\widetilde{\boldsymbol{D}}^{-1} \widetilde{\boldsymbol{J}}_{\mathrm{e}}(\boldsymbol{p}_{\mathrm{all}})$ is rank-deficient, the matrix limit $\widetilde{\boldsymbol{T}}^{\infty}$ exists, with its eigenvalues being either 1 or 0. In light of this, the column space of $\widetilde{\boldsymbol{T}}^{\infty}$ is the null space of $\boldsymbol{D}^{-1} \widetilde{\boldsymbol{J}}_{\mathrm{e}}(\boldsymbol{p}_{\mathrm{all}})$. Since there is no anchor in the auxiliary network, matrix $\boldsymbol{D}^{-1} \widetilde{\boldsymbol{J}}_{\mathrm{e}}(\boldsymbol{p}_{\mathrm{all}})$ is indeed rank-deficient with deficiency of at least 2 (corresponding to two-dimensional translation [78]). Therefore, as long as the network is finite, the null space of $\boldsymbol{D}^{-1} \widetilde{\boldsymbol{J}}_{\mathrm{e}}(\boldsymbol{p}_{\mathrm{all}})$ is at least two-dimensional.

More precisely, when only the range information is available (corresponding to $N_{\mathrm{t}} = 1$), the null space of $\widetilde{\boldsymbol{J}}_{\mathrm{e}}(\boldsymbol{p}_{\mathrm{all}})$ is known to be 3-dimensional, spanned by the vectors [78]

$$\begin{aligned} &\boldsymbol{v}_1 \propto [1\ 0\ \dots\ 1\ 0]^{\mathrm{T}}, \quad \boldsymbol{v}_2 \propto [0\ 1\ \dots\ 0\ 1]^{\mathrm{T}}, \\ &\boldsymbol{v}_3 \propto [\boldsymbol{u}(\phi_1 + \pi/2)^{\mathrm{T}}\ \dots\ \boldsymbol{u}(\phi_{N_{\mathrm{a}}+N_{\mathrm{b}}} + \pi/2)^{\mathrm{T}}]^{\mathrm{T}} \end{aligned} \tag{54}$$

where ϕ_i is the angle from node i to the network centroid. Furthermore, the null space of $\widetilde{\boldsymbol{J}}_{\mathrm{e}}(\boldsymbol{p}_{\mathrm{all}})$ is spanned by $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$ when both range and bearing information is available. In either case, $\lim_{n \to \infty} \widetilde{\boldsymbol{T}}_{ii}^{(n)} = \boldsymbol{D}_i^{-1}$ can be seen to be the product of two full-rank matrices hence is full-rank, and the proof is completed. □

APPENDIX III
THE PSEUDO-RANDOM WALK FORMALISM

A. Additional Notations

To facilitate the analysis, we define

$$\boldsymbol{H}_{\mathcal{S}}(i, j) = \begin{cases} \boldsymbol{F}_{\mathcal{S}}(i, j), & i \notin \mathcal{S}; \\ \boldsymbol{I}_2, & i \in \mathcal{S},\ i = j; \\ \boldsymbol{0}_{2 \times 2}, & i \in \mathcal{S},\ i \neq j. \end{cases} \tag{55}$$

In addition, the quantity $\boldsymbol{F}_{\mathcal{S}}(i, j)$ will also be used extensively in the analysis, and hence we reproduce its definition here as

$$\boldsymbol{F}_{\mathcal{S}}(i, j) = \sum_{n=1}^{\infty} \sum_{\boldsymbol{\omega} \in \Omega_n^{i,j}(\mathcal{S})} \widetilde{\boldsymbol{T}}_{\boldsymbol{\omega}}^{(n)}. \tag{56}$$

We may also rearrange $\boldsymbol{H}_{\mathcal{S}}(i, j)$ and $\boldsymbol{F}_{\mathcal{S}}(i, j)$ in a compact matrix form as

$$\begin{aligned} \boldsymbol{H}_{\mathcal{S}} &= \sum_{i \in \mathcal{N}_{\mathrm{a}} \cup \mathcal{N}_{\mathrm{b}}} \sum_{j \in \mathcal{S}} \boldsymbol{E}_{ij}^{N_{\mathrm{a}}+N_{\mathrm{b}}} \otimes \boldsymbol{H}_{\mathcal{S}}(i, j) \\ \boldsymbol{F}_{\mathcal{S}} &= \sum_{i \in \mathcal{N}_{\mathrm{a}} \cup \mathcal{N}_{\mathrm{b}}} \sum_{j \in \mathcal{S}} \boldsymbol{E}_{ij}^{N_{\mathrm{a}}+N_{\mathrm{b}}} \otimes \boldsymbol{F}_{\mathcal{S}}(i, j). \end{aligned} \tag{57}$$

$$\boldsymbol{F}_{\widetilde{\mathcal{R}}_i}(k,l) = \begin{cases} \left[\boldsymbol{P}_{\widetilde{\mathcal{R}}_i}^{-1}\right]_{m,n} - \left(\sum_{j\in\widetilde{\mathcal{R}}_i}\left[\boldsymbol{P}_{\widetilde{\mathcal{R}}_i}^{-1}\right]_{\min^{(j)}(\widetilde{\mathcal{R}}_i),m}\right)\boldsymbol{Z}_{\widetilde{\mathcal{R}}_i}^{-1}\left(\sum_{j\in\widetilde{\mathcal{R}}_i}\boldsymbol{D}_n\left[\boldsymbol{P}_{\widetilde{\mathcal{R}}_i}^{-1}\right]_{n,\min^{(j)}(\widetilde{\mathcal{R}}_i)}\right), & m\neq n; \\[2ex] \boldsymbol{I}_2 + \left[\boldsymbol{P}_{\widetilde{\mathcal{R}}_i}^{-1}\right]_{n,n} - \left(\sum_{j\in\widetilde{\mathcal{R}}_i}\left[\boldsymbol{P}_{\widetilde{\mathcal{R}}_i}^{-1}\right]_{\min^{(j)}(\widetilde{\mathcal{R}}_i),n}\right)\boldsymbol{Z}_{\widetilde{\mathcal{R}}_i}^{-1}\left(\sum_{j\in\widetilde{\mathcal{R}}_i}\boldsymbol{D}_n\left[\boldsymbol{P}_{\widetilde{\mathcal{R}}_i}^{-1}\right]_{n,\min^{(j)}(\widetilde{\mathcal{R}}_i)}\right) & m = n. \end{cases} \quad (65)$$

B. Potential Kernel

For the asymptotic analysis task considered in this paper, the central problem is to characterize the position information an agent i received from an anchor j as a function of their relative displacement $\boldsymbol{d}_{ij}$. Intuitively, the amount of position information should decrease with $\|\boldsymbol{d}_{ij}\|$, exhibiting a "position information path loss" behavior. According to previous analysis, this problem amounts to expressing the pseudo-probabilities $\boldsymbol{F}_{\widetilde{\mathcal{R}}_i}(i,j)$ as functions of $\boldsymbol{d}_{ij}$. In classical random walk theory, such a task relies on the construction of *potential kernels* [79]. In this subsection, we construct analogous quantities.

First we note that

$$\widetilde{\boldsymbol{J}}_e(\boldsymbol{p}_{\text{all}}) = \widetilde{\boldsymbol{D}}(\boldsymbol{I} - \widetilde{\boldsymbol{T}}) \tag{58}$$

where matrix $\boldsymbol{I} - \widetilde{\boldsymbol{T}}$ is an analogue of the *random walk normalized graph Laplacian* [80]. It is clear that $\boldsymbol{I} - \widetilde{\boldsymbol{T}}$ is non-invertible, but we may calculate its *generalized inverse*.

Lemma 1: We have

$$(\boldsymbol{I} - \widetilde{\boldsymbol{T}}^\infty)\boldsymbol{H}_\mathcal{S} = \lim_{n\to\infty}\{\boldsymbol{P}_n(\boldsymbol{F}_\mathcal{S} - \boldsymbol{I}_\mathcal{S})\} \tag{59}$$

where $\boldsymbol{P}_n = \widetilde{\boldsymbol{G}}_n - \boldsymbol{G}_n$, $\boldsymbol{G}_n = \sum_{k=0}^{n}\widetilde{\boldsymbol{T}}^k$, and

$$\widetilde{\boldsymbol{G}}_n = \sum_{\{i,j\}\subseteq\mathcal{N}_a\cup\mathcal{N}_b}\boldsymbol{E}_{ij}^{N_a+N_b}\otimes\left([\boldsymbol{G}_n]_{i,i}\boldsymbol{D}_i^{-1}\boldsymbol{D}_j\right). \tag{60}$$

Proof: Using the additional notations in Sec. III-A, and after some manipulations, one can obtain

$$(\widetilde{\boldsymbol{T}} - \boldsymbol{I})\boldsymbol{H}_\mathcal{S} = \boldsymbol{F}_\mathcal{S} - \boldsymbol{I}_\mathcal{S} \tag{61}$$

where $\boldsymbol{I}_\mathcal{S} = \sum_{i\in\mathcal{S}}\boldsymbol{E}_{ii}^{N_a+N_b}\otimes\boldsymbol{I}_2$. Multiplying $\boldsymbol{G}_n$ from the left at both sides of (61) and taking limit $n\to\infty$, we have

$$\boldsymbol{M}_\mathcal{S} - \boldsymbol{H}_\mathcal{S} = -\lim_{n\to\infty}\{\boldsymbol{G}_n(\boldsymbol{F}_\mathcal{S} - \boldsymbol{I}_\mathcal{S})\} \tag{62}$$

where $\boldsymbol{M}_\mathcal{S} = \widetilde{\boldsymbol{T}}^\infty\boldsymbol{H}_\mathcal{S}$. Note that (61) implies $\widetilde{\boldsymbol{T}}^\infty(\boldsymbol{F}_\mathcal{S} - \boldsymbol{I}_\mathcal{S}) = \boldsymbol{0}$, and thus the matrix limit in (62) does exist. For the $N_t \geqslant 2$ case, because the relative measurements in the network is invariant under two-dimensional translation, we have

$$\boldsymbol{Q}^T\widetilde{\boldsymbol{J}}_e(\boldsymbol{p}_{\text{all}}) = \boldsymbol{0}, \ \widetilde{\boldsymbol{J}}_e(\boldsymbol{p}_{\text{all}})\boldsymbol{Q} = \boldsymbol{0} \tag{63}$$

where $\widetilde{\boldsymbol{J}}_e(\boldsymbol{p}_{\text{all}})$ satisfies (58). Now, in light of (63), we can construct a matrix $\widetilde{\boldsymbol{G}}_n$ as in (60) satisfying $\widetilde{\boldsymbol{G}}_n(\boldsymbol{F}_\mathcal{S} - \boldsymbol{I}_\mathcal{S}) = \boldsymbol{0}$. Therefore with $\boldsymbol{P}_n := \widetilde{\boldsymbol{G}}_n - \boldsymbol{G}_n$ we have

$$\boldsymbol{H}_\mathcal{S} - \boldsymbol{M}_\mathcal{S} = \lim_{n\to\infty}\{\boldsymbol{P}_n(\boldsymbol{F}_\mathcal{S} - \boldsymbol{I}_\mathcal{S})\}. \tag{64}$$

Thus the proof is completed. □

In Lemma 1, we are essentially trying to construct a *generalized inverse* for $\boldsymbol{I} - \widetilde{\boldsymbol{T}}$. To elaborate, if $\boldsymbol{P} := \lim_{n\to\infty}\boldsymbol{P}_n$ exists, $-\boldsymbol{P}$ can serve as a valid generalized inverse. Then we can solve a system of linear equations for the desired quantity $\boldsymbol{F}_\mathcal{S}$. Indeed, the limit on the right hand side of (59) does exist

as long as $\widetilde{\boldsymbol{T}}^\infty$ exists. However, a subtle fact is that this does not imply the existence of $\lim_{n\to\infty}\boldsymbol{P}_n$. Fortunately, this is true when $N_t \geqslant 2$, for which we have the following result.

Theorem 3: When $N_t \geqslant 2$, for $k = \min^{(m)}(\widetilde{\mathcal{R}}_i)$ and $l = \min^{(n)}(\widetilde{\mathcal{R}}_i)$, the term $\boldsymbol{F}_{\widetilde{\mathcal{R}}_i}(k,l)$ can be expressed as (65), where matrix $\boldsymbol{P}_{\widetilde{\mathcal{R}}_i}$ is given by $\left[\boldsymbol{P}_{\widetilde{\mathcal{R}}_i}\right]_{k,l} = [\boldsymbol{P}]_{m,n}$, $\boldsymbol{Z}_{\widetilde{\mathcal{R}}_i} = \sum_{\{a,b\}\subseteq\widetilde{\mathcal{R}}_i}\boldsymbol{D}_a\left[\boldsymbol{P}_{\widetilde{\mathcal{R}}_i}^{-1}\right]_{\min^{(a)}(\widetilde{\mathcal{R}}_i),\min^{(b)}(\widetilde{\mathcal{R}}_i)}$ is a normalization constant, and

$$\boldsymbol{P} = \sum_{\{a,b\}\subseteq\mathcal{N}_a\cup\mathcal{N}_b}\boldsymbol{E}_{ab}^{N_a+N_b}\otimes\left\{\sum_{p=0}^{\infty}\left(\widetilde{\boldsymbol{T}}_{aa}^{(p)}\boldsymbol{D}_a^{-1}\boldsymbol{D}_b - \widetilde{\boldsymbol{T}}_{ab}^{(p)}\right)\right\}. \tag{66}$$

Proof: We shall first show that the matrix limit $\boldsymbol{P} = \lim_{n\to\infty}\boldsymbol{P}_n$ exists for the case $N_t \geqslant 2$, and hence

$$\boldsymbol{H}_\mathcal{S} - \boldsymbol{M}_\mathcal{S} = \boldsymbol{P}(\boldsymbol{F}_\mathcal{S} - \boldsymbol{I}_\mathcal{S}). \tag{67}$$

First we use the definition of $\boldsymbol{G}_n$ and obtain

$$\left[\boldsymbol{P}\boldsymbol{D}^{-1}\right]_{i,j} = \sum_{n=0}^{\infty}\left(\widetilde{\boldsymbol{T}}_{ii}^{(n)}\boldsymbol{D}_i^{-1} - \widetilde{\boldsymbol{T}}_{ij}^{(n)}\boldsymbol{D}_j^{-1}\right). \tag{68}$$

Matrix $\boldsymbol{P}$ exists if the series in (68) converges for every pair (i,j). It is clear from (63) that

$$\lim_{n\to\infty}\left\|\widetilde{\boldsymbol{T}}_{ii}^{(n)}\boldsymbol{D}_i^{-1} - \widetilde{\boldsymbol{T}}_{ij}^{(n)}\boldsymbol{D}_j^{-1}\right\|_2 = 0$$

holds for all i and j. Moreover, it can be shown that $\left\|\widetilde{\boldsymbol{T}}_{ii}^{(n)}\boldsymbol{D}_i^{-1} - \widetilde{\boldsymbol{T}}_{ij}^{(n)}\boldsymbol{D}_j^{-1}\right\|_2$ exhibits a linear rate of convergence as $n\to\infty$, since $\|\widetilde{\boldsymbol{T}}^n - \widetilde{\boldsymbol{T}}^\infty\|_2$ converges to zero linearly as $n\to\infty$. Therefore, we can conclude that matrix $\boldsymbol{P}$ does exist.

Now, restricting the rows and columns of the matrices in (67) to the elements in the index set $\mathcal{S}$, we can construct $2|\mathcal{S}|\times 2|\mathcal{S}|$ matrices $\boldsymbol{F}_{\mathcal{S},\mathcal{S}}, \boldsymbol{H}_{\mathcal{S},\mathcal{S}}, \boldsymbol{P}_\mathcal{S}, \boldsymbol{M}_{\mathcal{S},\mathcal{S}}$ and $\boldsymbol{D}_\mathcal{S}$ as

$$[\boldsymbol{P}_\mathcal{S}]_{i,j} = [\boldsymbol{P}]_{k,l}, \ \boldsymbol{H}_{\mathcal{S},\mathcal{S}} = \boldsymbol{I}_{2|\mathcal{S}|}, \ [\boldsymbol{M}_{\mathcal{S},\mathcal{S}}]_{i,j} = [\boldsymbol{M}_\mathcal{S}]_{k,l},$$
$$[\boldsymbol{F}_{\mathcal{S},\mathcal{S}}]_{i,j} = [\boldsymbol{F}_\mathcal{S}]_{k,l}, \ [\boldsymbol{D}_\mathcal{S}]_{i,j} = [\boldsymbol{D}]_{k,l} \tag{69}$$

where $k = \min^{(i)}(\mathcal{S})$ and $l = \min^{(j)}(\mathcal{S})$. Thus using (67) we obtain

$$\boldsymbol{F}_{\mathcal{S},\mathcal{S}} - \boldsymbol{I} = \boldsymbol{P}_\mathcal{S}^{-1}(\boldsymbol{I} - \boldsymbol{M}_{\mathcal{S},\mathcal{S}}). \tag{70}$$

From (70) we also have $\boldsymbol{Q}^T\boldsymbol{D}_\mathcal{S}\boldsymbol{P}_\mathcal{S}^{-1}(\boldsymbol{I} - \boldsymbol{M}_{\mathcal{S},\mathcal{S}}) = \boldsymbol{0}$, and hence

$$\boldsymbol{M}_{\mathcal{S},\mathcal{S}} = \boldsymbol{Q}\left(\boldsymbol{Q}^T\boldsymbol{D}_\mathcal{S}\boldsymbol{P}_\mathcal{S}^{-1}\boldsymbol{Q}\right)^{-1}\boldsymbol{Q}^T\boldsymbol{D}_\mathcal{S}\boldsymbol{P}_\mathcal{S}^{-1}. \tag{71}$$

Substituting (71) into (70) we obtain

$$\boldsymbol{F}_{\mathcal{S},\mathcal{S}} = \boldsymbol{I} + \boldsymbol{P}_\mathcal{S}^{-1} - \boldsymbol{P}_\mathcal{S}^{-1}\boldsymbol{Q}\left(\boldsymbol{Q}^T\boldsymbol{D}_\mathcal{S}\boldsymbol{P}_\mathcal{S}^{-1}\boldsymbol{Q}\right)^{-1}\boldsymbol{Q}^T\boldsymbol{D}_\mathcal{S}\boldsymbol{P}_\mathcal{S}^{-1} \tag{72}$$

yielding (65) after some manipulations. □

The matrix $[\boldsymbol{P}_{\widetilde{\mathcal{R}}_i}]_{k,l}$ resembles the potential between m and n in classical random walk theory in the sense that, as will be illustrated in Section IV, it is a function of solely $\boldsymbol{d}_{mn}$ in infinite lattice networks. It is also important in finite networks,

since $\boldsymbol{P}_{\widetilde{\mathcal{R}}_i} \in \mathbb{R}^{2(N_{\mathrm{b}}+1) \times 2(N_{\mathrm{b}}+1)}$ and $\{\boldsymbol{D}_i\}_{i \in N_{\mathrm{a}} \cup N_{\mathrm{b}}}$ retain all the necessary information to describe the information coupling between agent i and other nodes. Typically, the number of anchors N_{b} is far less than the number of agents N_{a}, and thus the dimensionality of $\boldsymbol{P}_{\widetilde{\mathcal{R}}_i}$ is far lower than that of $\boldsymbol{J}_{\mathrm{e}}(\boldsymbol{p})$. Therefore, it is much simpler to derive EoC using $\boldsymbol{P}_{\widetilde{\mathcal{R}}_i}$ instead of manipulating $\boldsymbol{J}_{\mathrm{e}}(\boldsymbol{p})$ directly.

C. Infinite Lattice Networks: Pseudo-Characteristic Function

We now give the asymptotic expressions of EoC in infinitely large lattice networks. Thanks to the symmetry of these networks, we are able to perform harmonic analysis on infinite lattice networks since the underlying graph $\mathcal{G}_{\mathrm{net}}$ is invariant with respect to translation. Since we have assumed that nodes are equipped with UOAs when $N_{\mathrm{t}} \geqslant 2$ which implies that $\boldsymbol{J}_{ij}$ only depends on $\boldsymbol{d}_{ij}$, the singular EFIM $\widetilde{\boldsymbol{J}}_{\mathrm{e}}(\boldsymbol{p}_{\mathrm{all}})$ built upon $\mathcal{G}_{\mathrm{net}}$ should also be translation-invariant. In this regard, we can define *pseudo-characteristic functions* on auxiliary networks.

Definition 3 (Pseudo-Characteristic Function): The pseudo-characteristic function of an auxiliary network associated with an infinite lattice network is defined as

$$\boldsymbol{\Phi}(\boldsymbol{\vartheta}) = \sum_{\boldsymbol{x} \in \mathbb{Z}^2} \widetilde{\boldsymbol{T}}_1(\boldsymbol{0}, \boldsymbol{x}) e^{\jmath \boldsymbol{x}^{\mathrm{T}} \boldsymbol{\vartheta}} \qquad (73)$$

where $\widetilde{\boldsymbol{T}}_k(\boldsymbol{0}, \boldsymbol{x}) = \widetilde{\boldsymbol{T}}_{ij}^{(k)}$ for $\boldsymbol{p}_i = \boldsymbol{0}$ and $\boldsymbol{p}_j = \boldsymbol{x}$, and $\jmath$ is the imaginary unit $\sqrt{-1}$. We define the notation $\widetilde{\boldsymbol{T}}_k(\boldsymbol{0}, \boldsymbol{x})$ to emphasize that it is a function of not only the indices of agents, but also their positions.

To ensure that Theorem 3 is applicable, we first show the recurrence of the random walk on the auxiliary network of infinite lattice networks, using the pseudo-characteristic function.

Proposition 7: For any auxiliary network (or original network without anchor) of infinite lattice networks, we have $\boldsymbol{F}_{ii} = \boldsymbol{I}_2$ for all $i \in \mathcal{N}_{\mathrm{a}}$.

Proof: Using the pseudo-characteristic function we have

$$\sum_{n=0}^{\infty} \widetilde{\boldsymbol{T}}_n(\boldsymbol{0}, \boldsymbol{0}) = \frac{1}{2\pi} \int_{\mathcal{C}} (\boldsymbol{I} - \boldsymbol{\Phi}(\boldsymbol{\vartheta}))^{-1} \mathrm{d}\boldsymbol{\vartheta}$$

We only need to show that the integral does not converge. We consider a decomposition as follows

$$\frac{1}{(2\pi)^2} \int_{\mathcal{C}} (\boldsymbol{I} - \boldsymbol{\Phi}(\boldsymbol{\vartheta}))^{-1} \mathrm{d}\boldsymbol{\vartheta} = \frac{1}{2\pi^2} \int_{\mathcal{C}} \|\boldsymbol{\vartheta}\|^{-2} \boldsymbol{\Sigma}_2^{-1}(\boldsymbol{\vartheta}) \mathrm{d}\boldsymbol{\vartheta}$$
$$+ \frac{1}{(2\pi)^2} \int_{\mathcal{C}} \widetilde{\boldsymbol{\Psi}}(\boldsymbol{\vartheta}) \mathrm{d}\boldsymbol{\vartheta} \qquad (74)$$

where $\widetilde{\boldsymbol{\Psi}} := (\boldsymbol{I} - \boldsymbol{\Phi}(\boldsymbol{\vartheta}))^{-1} - 2\|\boldsymbol{\vartheta}\|^{-2} \boldsymbol{\Sigma}_2^{-1}$. For the first term on the right hand side of (74), note that $c_1 \boldsymbol{I} \preccurlyeq \boldsymbol{\Sigma}_2^{-1}(\boldsymbol{\vartheta}) \preccurlyeq c_2 \boldsymbol{I}$ holds for all $\boldsymbol{\vartheta}$ where $c_1 = (\max_{\boldsymbol{\vartheta}} \lambda_{\max}(\boldsymbol{\Sigma}_2(\boldsymbol{\vartheta})))^{-1}$ and $c_2 = (\min_{\boldsymbol{\vartheta}} \lambda_{\min}(\boldsymbol{\Sigma}_2(\boldsymbol{\vartheta})))^{-1}$ are positive constants, and that

$$\int_{\mathcal{C}} \|\boldsymbol{\vartheta}\|^{-2} \mathrm{d}\boldsymbol{\vartheta} = \int_0^{2\pi} \mathrm{d}\phi \int_0^{\pi} r^{-1} \mathrm{d}r$$

does not converge. Therefore, to prove this proposition, it is now sufficient to show that $\widetilde{\boldsymbol{\Psi}}(\boldsymbol{\vartheta})$ is integrable. Taking the spectral norm of $\widetilde{\boldsymbol{\Psi}}$, we have

$$\|\widetilde{\boldsymbol{\Psi}}(\boldsymbol{\vartheta})\|_2 = \left\| (\boldsymbol{I} - \boldsymbol{\Phi}(\boldsymbol{\vartheta}))^{-1} - \left(\frac{\|\boldsymbol{\vartheta}\|^2}{2} \boldsymbol{\Sigma}_2(\boldsymbol{\vartheta}) \right)^{-1} \right\|_2$$
$$\leqslant \|(\boldsymbol{I} - \boldsymbol{\Phi}(\boldsymbol{\vartheta}))^{-1}\|_2 \cdot \left\| \frac{2}{\|\boldsymbol{\vartheta}\|^2} \boldsymbol{\Sigma}_2^{-1}(\boldsymbol{\vartheta}) \right\|_2 \qquad (75)$$
$$\cdot \left\| (\boldsymbol{I} - \boldsymbol{\Phi}(\boldsymbol{\vartheta})) - \frac{\|\boldsymbol{\vartheta}\|^2}{2} \boldsymbol{\Sigma}_2(\boldsymbol{\vartheta}) \right\|_2.$$

To bound the term $\|(\boldsymbol{I} - \boldsymbol{\Phi}(\boldsymbol{\vartheta}))^{-1}\|_2$, note that $\widetilde{\boldsymbol{T}}_1(\boldsymbol{0}, \boldsymbol{x})$ is Hermitian in infinite lattice networks, and thus using (80) one can verify that $\boldsymbol{\Phi}(\boldsymbol{\vartheta})$ is also Hermitian for any $\boldsymbol{\vartheta}$. Hence $\|(\boldsymbol{I} - \boldsymbol{\Phi}(\boldsymbol{\vartheta}))^{-1}\|_2$ is in fact $[\lambda_{\min}(\boldsymbol{I} - \boldsymbol{\Phi}(\boldsymbol{\vartheta}))]^{-1}$. From the symmetry of $\widetilde{\boldsymbol{T}}_1(\boldsymbol{0}, \boldsymbol{x})$, i.e., $\widetilde{\boldsymbol{T}}_1(\boldsymbol{0}, \boldsymbol{x}) = \widetilde{\boldsymbol{T}}_1(\boldsymbol{0}, -\boldsymbol{x})$, we have

$$\boldsymbol{\Phi}(\boldsymbol{\vartheta}) = \Re\{\boldsymbol{\Phi}(\boldsymbol{\vartheta})\} = \sum_{\boldsymbol{x} \in \mathbb{Z}^2} \widetilde{\boldsymbol{T}}_1(\boldsymbol{0}, \boldsymbol{x}) \cos(\boldsymbol{x}^{\mathrm{T}} \boldsymbol{\vartheta}).$$

Since $\widetilde{\boldsymbol{T}}_1(\boldsymbol{0}, \boldsymbol{x})$ is positive semi-definite for all $\boldsymbol{x}$ and positive definite for some $\boldsymbol{x}$, we can bound matrix $\boldsymbol{I} - \boldsymbol{\Phi}(\boldsymbol{\vartheta})$ from below as

$$\boldsymbol{I} - \boldsymbol{\Phi}(\boldsymbol{\vartheta}) \succcurlyeq 2\pi^{-2} \sum_{\|\boldsymbol{x}\| \leqslant K} (\boldsymbol{x}^{\mathrm{T}} \boldsymbol{\vartheta})^2 \widetilde{\boldsymbol{T}}_1(\boldsymbol{0}, \boldsymbol{x})$$
$$= \|\boldsymbol{\vartheta}\|^2 \boldsymbol{\Phi}_{\mathrm{c}} \qquad (76)$$

for all $\boldsymbol{\vartheta}$ satisfying $\|\boldsymbol{\vartheta}\| \leqslant \pi K^{-1}$, where $\boldsymbol{\Phi}_{\mathrm{c}}$ is a constant positive definite matrix with respect to $\boldsymbol{\vartheta}$. For the case $\|\boldsymbol{\vartheta}\| > \pi K^{-1}$, we also have

$$\boldsymbol{I} - \boldsymbol{\Phi}(\boldsymbol{\vartheta}) \succcurlyeq K^2 \pi^{-2} \|\boldsymbol{\vartheta}\|^2 \boldsymbol{\Phi}_{\min}, \qquad (77)$$

where $\boldsymbol{\Phi}_{\min} := \min_{\boldsymbol{\vartheta} \in \{\boldsymbol{\vartheta} | \|\boldsymbol{\vartheta}\| > \pi K^{-1}\}} \{\boldsymbol{I} - \boldsymbol{\Phi}(\boldsymbol{\vartheta})\} \succ \boldsymbol{0}_{2 \times 2}$. The minimum is in taken in the sense of Löwner partial ordering [81, Sec. 2.4].

Combining (76) and (77), we see that $\|(\boldsymbol{I} - \boldsymbol{\Phi}(\boldsymbol{\vartheta}))^{-1}\|_2 \leqslant c_3 \|\boldsymbol{\vartheta}\|^{-2}$ for some constant $c_3 > 0$, and hence for some constant $c_4 > 0$, (75) can be further simplified as

$$\|\widetilde{\boldsymbol{\Psi}}(\boldsymbol{\vartheta})\|_2 \leqslant c_4 \|\boldsymbol{\vartheta}\|^{-4} \left\| (\boldsymbol{I} - \boldsymbol{\Phi}(\boldsymbol{\vartheta})) - \frac{\|\boldsymbol{\vartheta}\|^2}{2} \boldsymbol{\Sigma}_2(\boldsymbol{\vartheta}) \right\|_2.$$

Denote $\widetilde{\boldsymbol{E}}_{\boldsymbol{x}}\{f(\boldsymbol{x})\} = \sum_{\boldsymbol{x} \in \mathbb{Z}^2} f(\boldsymbol{x}) \widetilde{\boldsymbol{T}}_1(\boldsymbol{0}, \boldsymbol{x})$, and note that $\|\boldsymbol{\vartheta}\| \boldsymbol{\Sigma}_2(\boldsymbol{\vartheta}) = \widetilde{\boldsymbol{E}}_{\boldsymbol{x}}\{(\boldsymbol{x}^{\mathrm{T}} \boldsymbol{\vartheta})^2\}$ and $\widetilde{\boldsymbol{E}}_{\boldsymbol{x}}\{\boldsymbol{x}\} = \boldsymbol{0}_{2 \times 2}$, using Taylor expansion we obtain

$$\left\| (\boldsymbol{I} - \boldsymbol{\Phi}(\boldsymbol{\vartheta})) - \frac{\|\boldsymbol{\vartheta}\|^2}{2} \boldsymbol{\Sigma}_2(\boldsymbol{\vartheta}) \right\|_2$$
$$\leqslant \left\| \widetilde{\boldsymbol{E}}_{\boldsymbol{x}}\left\{ \left| e^{\jmath \boldsymbol{x}^{\mathrm{T}} \boldsymbol{\vartheta}} - 1 - \jmath \boldsymbol{x}^{\mathrm{T}} \boldsymbol{\vartheta} - \frac{1}{2}(\jmath \boldsymbol{x}^{\mathrm{T}} \boldsymbol{\vartheta})^2 \right| \right\} \right\|_2 \qquad (78)$$
$$\leqslant c_5 \|\boldsymbol{\vartheta}\|^{2+\delta} \|\boldsymbol{\Sigma}_{2+\delta}(\boldsymbol{\vartheta})\|_2$$

for some constant $c_5 > 0$ and $\delta > 0$, where $\boldsymbol{\Sigma}_a(\boldsymbol{\vartheta}) := \widetilde{\boldsymbol{E}}_{\boldsymbol{x}}\{|\boldsymbol{x}^{\mathrm{T}}(\boldsymbol{\vartheta}/\|\boldsymbol{\vartheta}\|)|^a\}$. From the system model in Section II we see that $\boldsymbol{\Sigma}_a$ always exists for $a > 0$ and is positive definite with finite norm. Furthermore, it does not depend on the norm $\|\boldsymbol{\vartheta}\|$. Hence the proof is completed. $\square$

By application of the pseudo-characteristic function, we can rewrite Theorem 3 in a simpler form in the case of infinite lattice networks. To achieve this, first note that in infinite

$$\widetilde{\boldsymbol{F}}_{\widetilde{\mathcal{R}}_i}(k,l) = \begin{cases} \left[\boldsymbol{P}_{\widetilde{\mathcal{R}}_i}^{-1}\right]_{m,n} - \left[\boldsymbol{P}_{\widetilde{\mathcal{R}}_i}^{-1}\right]^{(m)} \left(\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{P}_{\widetilde{\mathcal{R}}_i}^{-1}\boldsymbol{Q}\right)^{-1} \left(\left[\boldsymbol{P}_{\widetilde{\mathcal{R}}_i}^{-1}\right]^{(n)}\right)^{\mathrm{T}}, & m \neq n; \\ \boldsymbol{I}_2 + \left[\boldsymbol{P}_{\widetilde{\mathcal{R}}_i}^{-1}\right]_{n,n} - \left[\boldsymbol{P}_{\widetilde{\mathcal{R}}_i}^{-1}\right]^{(n)} \left(\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{P}_{\widetilde{\mathcal{R}}_i}^{-1}\boldsymbol{Q}\right)^{-1} \left(\left[\boldsymbol{P}_{\widetilde{\mathcal{R}}_i}^{-1}\right]^{(n)}\right)^{\mathrm{T}} & m = n. \end{cases} \tag{79}$$

lattice networks the matrix $\boldsymbol{D}$ is proportional to the identity matrix $\boldsymbol{I}$, and thus we have the following corollary.

Corollary 1: In infinite lattice networks, for $k = \min^{(m)}(\widetilde{\mathcal{R}}_i)$ and $l = \min^{(n)}(\widetilde{\mathcal{R}}_i)$, if $N_{\mathrm{t}} \geqslant 2$, $\boldsymbol{F}_{\widetilde{\mathcal{R}}_i}(k,l)$ can be expressed as (79), where $\boldsymbol{P}_{\widetilde{\mathcal{R}}_i}^{(m)} := \sum_{j=1}^{N_{\mathrm{b}}+1}\left[\boldsymbol{P}_{\widetilde{\mathcal{R}}_i}^{-1}\right]_{i,j}$.

Proof: In infinite lattice networks, for any node i, the neighboring nodes are symmetrically distributed around $\boldsymbol{p}_i$. Using the UOA assumption we have $\boldsymbol{D}_i = \boldsymbol{D}_j \propto \boldsymbol{I}_2$ holds for all nodes i and j, and this corollary follows directly from Theorem 3. □

From Corollary 1 it is clear that $\boldsymbol{F}_{\widetilde{\mathcal{R}}_i}(k,l)$ can be expressed solely in terms of the inverse of the potential kernel, i.e., $\boldsymbol{P}_{\widetilde{\mathcal{R}}_i}^{-1}$, in infinite lattice networks. Matrix $\boldsymbol{P}$ can now be rewritten as

$$\boldsymbol{P} = \lim_{n\to\infty}\left\{\boldsymbol{G}_n(\boldsymbol{0},\boldsymbol{0}) \otimes (\boldsymbol{1}\boldsymbol{1}^{\mathrm{T}}) - \boldsymbol{G}_n\right\}$$

where $\boldsymbol{G}_n(\boldsymbol{x},\boldsymbol{y}) := \sum_{k=0}^{n}\widetilde{\boldsymbol{T}}_k(\boldsymbol{x},\boldsymbol{y})$ and $\boldsymbol{G}_n := \sum_{k=0}^{n}\widetilde{\boldsymbol{T}}_{\boldsymbol{q}}^k$, and each of the 2×2 blocks of $\boldsymbol{P}$ can be expressed in terms of the following matrix limit

$$\boldsymbol{P}(\boldsymbol{x},\boldsymbol{y}) := \lim_{n\to\infty}\sum_{k=0}^{n}\left(\widetilde{\boldsymbol{T}}_k(\boldsymbol{0},\boldsymbol{0}) - \widetilde{\boldsymbol{T}}_k(\boldsymbol{x},\boldsymbol{y})\right). \tag{80}$$

Due to the translation-invariance of infinite lattice networks, $\boldsymbol{P}(\boldsymbol{x},\boldsymbol{y})$ is in fact a function of $\boldsymbol{x} - \boldsymbol{y}$. Hence $[\boldsymbol{P}_{\widetilde{\mathcal{R}}_i}]_{k,l}$ is a function of solely $\boldsymbol{d}_{mn}$, where $k = \min^{(m)}(\widetilde{\mathcal{R}}_i)$ and $l = \min^{(n)}(\widetilde{\mathcal{R}}_i)$. We will show this in the following proposition using the pseudo-characteristic function.

Proposition 8: Matrix limit $\boldsymbol{P}(\boldsymbol{x},\boldsymbol{y})$ can be expressed in terms of the pseudo-characteristic function $\boldsymbol{\Phi}(\boldsymbol{\vartheta})$ as

$$\boldsymbol{P}(\boldsymbol{x},\boldsymbol{y}) = \frac{1}{(2\pi)^2}\int_{\mathcal{C}}(\boldsymbol{I} - \boldsymbol{\Phi}(\boldsymbol{\vartheta}))^{-1}\left(1 - e^{\jmath(\boldsymbol{x}-\boldsymbol{y})^{\mathrm{T}}\boldsymbol{\vartheta}}\right)\mathrm{d}\boldsymbol{\vartheta} \tag{81}$$

where $\mathcal{C} := \{\boldsymbol{\vartheta}|\|\boldsymbol{\vartheta}\| \leqslant \pi\}$.

Proof: Using the definition of characteristic function, for $\boldsymbol{\Phi}^2(\boldsymbol{\vartheta})$ we have

$$\boldsymbol{\Phi}^2(\boldsymbol{\vartheta}) = \sum_{\boldsymbol{x}\in\mathbb{Z}^2}\sum_{\boldsymbol{z}\in\mathbb{Z}^2}\widetilde{\boldsymbol{T}}_1(\boldsymbol{0},\boldsymbol{x})\widetilde{\boldsymbol{T}}_1(\boldsymbol{0},\boldsymbol{z})e^{\jmath(\boldsymbol{x}+\boldsymbol{z})^{\mathrm{T}}\boldsymbol{\vartheta}}.$$

But in infinite lattice networks, we have $\widetilde{\boldsymbol{T}}_1(\boldsymbol{0},\boldsymbol{z}) = \widetilde{\boldsymbol{T}}_1(\boldsymbol{x},\boldsymbol{x}+\boldsymbol{z})$, and thus

$$\boldsymbol{\Phi}^2(\boldsymbol{\vartheta}) = \sum_{\boldsymbol{y}\in\mathbb{Z}^2}\widetilde{\boldsymbol{T}}_2(\boldsymbol{0},\boldsymbol{y})e^{\jmath\boldsymbol{y}^{\mathrm{T}}\boldsymbol{\vartheta}}$$

where $\boldsymbol{y} = \boldsymbol{x} + \boldsymbol{z}$. Repeating this iteration and taking Fourier transform from both sides, we obtain

$$\widetilde{\boldsymbol{T}}_n(\boldsymbol{x},\boldsymbol{y}) = \frac{1}{(2\pi)^2}\int_{\mathcal{C}}\boldsymbol{\Phi}^n(\boldsymbol{\vartheta})e^{\jmath\boldsymbol{\vartheta}^{\mathrm{T}}(\boldsymbol{x}-\boldsymbol{y})}\mathrm{d}\boldsymbol{\vartheta}.$$

Therefore, from the definition of $\boldsymbol{P}(\boldsymbol{x},\boldsymbol{y})$ (80), we see that it can be expressed as

$$\begin{aligned}\boldsymbol{P}(\boldsymbol{x},\boldsymbol{y}) &= \sum_{k=0}^{\infty}\left(\widetilde{\boldsymbol{T}}_k(\boldsymbol{0},\boldsymbol{0}) - \widetilde{\boldsymbol{T}}_k(\boldsymbol{x},\boldsymbol{y})\right) \\ &= \frac{1}{(2\pi)^2}\int_{\mathcal{C}}\left(1 - e^{\jmath\boldsymbol{\vartheta}^{\mathrm{T}}(\boldsymbol{x}-\boldsymbol{y})}\right)\sum_{k=0}^{\infty}\boldsymbol{\Phi}^k(\boldsymbol{\vartheta})\mathrm{d}\boldsymbol{\vartheta} \\ &= \frac{1}{(2\pi)^2}\int_{\mathcal{C}}(\boldsymbol{I} - \boldsymbol{\Phi}(\boldsymbol{\vartheta}))^{-1}\left(1 - e^{\jmath\boldsymbol{\vartheta}^{\mathrm{T}}(\boldsymbol{x}-\boldsymbol{y})}\right)\mathrm{d}\boldsymbol{\vartheta}\end{aligned} \tag{82}$$

hence the proof is completed. □

With Proposition 8, the asymptotic behavior of $\boldsymbol{P}(\boldsymbol{x},\boldsymbol{y})$ can be described as follows.

Proposition 9 (Asymptotic Behavior of $\boldsymbol{P}(\boldsymbol{x},\boldsymbol{y})$): As $\|\boldsymbol{x} - \boldsymbol{y}\| \to \infty$, we have

$$\lambda_m(\boldsymbol{P}(\boldsymbol{x},\boldsymbol{y})) = \Theta\left(\log\frac{\|\boldsymbol{x}-\boldsymbol{y}\|}{\max_{\boldsymbol{\vartheta}}\lambda_{\max}(\boldsymbol{\Sigma}_2(\boldsymbol{\vartheta}))}\right), \; \forall m \in \{1,2\} \tag{83}$$

where $\boldsymbol{\Sigma}_2(\boldsymbol{\vartheta}) := \|\boldsymbol{\vartheta}\|^{-2}\sum_{\boldsymbol{x}\in\mathbb{Z}^2}(\boldsymbol{x}^{\mathrm{T}}\boldsymbol{\vartheta})^2\widetilde{\boldsymbol{T}}_1(\boldsymbol{0},\boldsymbol{x})$. The term $\max_{\boldsymbol{\vartheta}}\lambda_{\max}(\boldsymbol{\Sigma}_2(\boldsymbol{\vartheta}))$ is a constant with respect to $\|\boldsymbol{d}_{i\nu}\|$ but increases with the path loss exponent γ.

Proof: We first decompose $\boldsymbol{P}(\boldsymbol{x},\boldsymbol{y})$ into

$$\begin{aligned}\boldsymbol{P}(\boldsymbol{x},\boldsymbol{y}) &= \frac{1}{(2\pi)^2}\int_{\mathcal{C}}\left(1 - e^{\jmath(\boldsymbol{x}-\boldsymbol{y})^{\mathrm{T}}\boldsymbol{\vartheta}}\right)(\boldsymbol{I} - \boldsymbol{\Phi}(\boldsymbol{\vartheta}))^{-1}\mathrm{d}\boldsymbol{\vartheta} \\ &= \frac{1}{2\pi^2}\int_{\mathcal{C}}\frac{1 - \cos((\boldsymbol{x}-\boldsymbol{y})^{\mathrm{T}}\boldsymbol{\vartheta})}{\|\boldsymbol{\vartheta}\|^2}\boldsymbol{\Sigma}_2^{-1}(\boldsymbol{\vartheta})\mathrm{d}\boldsymbol{\vartheta} \\ &\quad + \frac{1}{(2\pi)^2}\int_{\mathcal{C}}\left(1 - e^{\jmath(\boldsymbol{x}-\boldsymbol{y})^{\mathrm{T}}\boldsymbol{\vartheta}}\right)\widetilde{\boldsymbol{\Psi}}(\boldsymbol{\vartheta})\mathrm{d}\boldsymbol{\vartheta}.\end{aligned} \tag{84}$$

For the first term on the right hand side of (84), we have

$$\begin{aligned}&\int_{\mathcal{C}}\left(1 - \cos((\boldsymbol{x}-\boldsymbol{y})^{\mathrm{T}}\boldsymbol{\vartheta})\right)\|\boldsymbol{\vartheta}\|^{-2}\mathrm{d}\boldsymbol{\vartheta} \\ &\sim \int_0^{2\pi}\mathrm{d}\phi\int_0^{\pi}\left(1 - \cos(\|\boldsymbol{x}-\boldsymbol{y}\|r\sin\phi)\right)r^{-1}\mathrm{d}r \\ &\sim \ln\|\boldsymbol{x}-\boldsymbol{y}\|\end{aligned}$$

where ϕ is the angle between $\boldsymbol{x} - \boldsymbol{y}$ and $\boldsymbol{\vartheta}$. The second term on the right hand side of (84) can be simplified as

$$\frac{1}{(2\pi)^2}\int_{\mathcal{C}}\left(1 - e^{\jmath(\boldsymbol{x}-\boldsymbol{y})^{\mathrm{T}}\boldsymbol{\vartheta}}\right)\widetilde{\boldsymbol{\Psi}}(\boldsymbol{\vartheta})\mathrm{d}\boldsymbol{\vartheta} = \frac{1}{(2\pi)^2}\int_{\mathcal{C}}\widetilde{\boldsymbol{\Psi}}(\boldsymbol{\vartheta})\mathrm{d}\boldsymbol{\vartheta}$$

due to the Riemann-Lebesgue Lemma in Fourier analysis. Therefore, since $\widetilde{\boldsymbol{\Psi}}(\boldsymbol{\vartheta})$ is integrable, the proof is completed. □

In Appendix IV, Proposition 9 will be used to investigate the asymptotic characteristics of EFIM when there is a single anchor in the network.

A. *Proof of Proposition 3*

Proof: We use Theorem 3 to simplify the analysis. Since there is only one anchor, $\boldsymbol{P}_{\widetilde{\mathcal{R}}_i}$ is a 4×4 matrix given by

$$\boldsymbol{P}_{\widetilde{\mathcal{R}}_i} = \begin{bmatrix} \boldsymbol{0}_{2\times 2} & \left[\boldsymbol{P}_{\widetilde{\mathcal{R}}_i}\right]_{1,2} \\ \left[\boldsymbol{P}_{\widetilde{\mathcal{R}}_i}\right]_{2,1} & \boldsymbol{0}_{2\times 2} \end{bmatrix}. \tag{85}$$

Since $\|\boldsymbol{d}_{i\nu}\| = o(D_{\text{net}})$, when the network is sufficiently large, we have $\boldsymbol{D}_i = \boldsymbol{D}_\nu \propto \boldsymbol{I}_2$. Hence from (65) and (85) we have

$$\boldsymbol{F}_{\widetilde{\mathcal{R}}_i}(i,\nu) = \frac{1}{2}\left[\boldsymbol{P}_{\widetilde{\mathcal{R}}_i}\right]_{2,1}^{-1}.$$

According to (24), the EFIM for agent i can be rewritten as

$$\text{sp}\{\boldsymbol{p}_i\} = \text{tr}\left\{2\left(\boldsymbol{D}_i\left[\boldsymbol{P}_{\widetilde{\mathcal{R}}_i}\right]_{2,1}^{-1}\right)^{-1}\right\} \propto \text{tr}\left\{\left[\boldsymbol{P}_{\widetilde{\mathcal{R}}_i}\right]_{2,1}\right\}. \tag{86}$$

From Proposition 9 we see that

$$\text{tr}\left\{\sum_{n=0}^{\infty}\left(\widetilde{\boldsymbol{T}}_{\nu\nu}^{(n),\text{inf}} - \widetilde{\boldsymbol{T}}_{\nu i}^{(n),\text{inf}}\right)\right\} = \Theta(\log\|\boldsymbol{p}_\nu - \boldsymbol{p}_i\|)$$

where $\widetilde{\boldsymbol{T}}_{ij}^{(n),\text{inf}}$ denotes $\widetilde{\boldsymbol{T}}_{ij}^{(n)}$ in an infinite lattice network. Now it suffices to show that

$$\lim_{D_{\text{net}}\to\infty}\left[\boldsymbol{P}_{\widetilde{\mathcal{R}}_i}\right]_{2,1} = \lim_{D_{\text{net}}\to\infty}\sum_{n=0}^{\infty}\left(\widetilde{\boldsymbol{T}}_{\nu\nu}^{(n)} - \widetilde{\boldsymbol{T}}_{\nu i}^{(n)}\right)$$
$$= \sum_{n=0}^{\infty}\left(\widetilde{\boldsymbol{T}}_{\nu\nu}^{(n),\text{inf}} - \widetilde{\boldsymbol{T}}_{\nu i}^{(n),\text{inf}}\right).$$

Note that we have $\widetilde{\boldsymbol{T}}_{ij}^{(n),\text{inf}} = \widetilde{\boldsymbol{T}}_{ij}^{(n)}$ as long as the random walk cannot reach the network edge from node i in n steps. Thus

$$\lim_{D_{\text{net}}\to\infty}\sum_{n=0}^{\infty}\left(\widetilde{\boldsymbol{T}}_{\nu\nu}^{(n)} - \widetilde{\boldsymbol{T}}_{\nu i}^{(n)}\right) - \sum_{n=0}^{\infty}\left(\widetilde{\boldsymbol{T}}_{\nu\nu}^{(n),\text{inf}} - \widetilde{\boldsymbol{T}}_{\nu i}^{(n),\text{inf}}\right)$$
$$= \lim_{m\to\infty}\sum_{n=m}^{\infty}\left(\left(\widetilde{\boldsymbol{T}}_{\nu\nu}^{(n)} - \widetilde{\boldsymbol{T}}_{\nu\nu}^{(n),\text{inf}}\right) - \left(\widetilde{\boldsymbol{T}}_{\nu i}^{(n)} - \widetilde{\boldsymbol{T}}_{\nu i}^{(n),\text{inf}}\right)\right)$$

where $m = \lceil R_{\max}^{-1}(D_{\text{net}}/2 - \max\{\|\boldsymbol{p}_i\|, \|\boldsymbol{p}_\nu\|\})\rceil$ and $R_{\max}$ is the maximum communication range between nodes. Since $\|\boldsymbol{d}_{i\nu}\| = o(D_{\text{net}}/2 - \|\boldsymbol{p}_i\|)$, as the network expands, the distance between i and ν is negligible compared with their distance from the network edge, implying that $m \to \infty$ as $D_{\text{net}} \to \infty$. Thus the proof is completed. □

B. *Proof of Proposition 4*

Proof: When $N_{\text{t}} = 1$, since only information on the relative distances can be obtained, it is known that the null space of $\widetilde{\boldsymbol{J}}_{\text{e}}(\boldsymbol{p}_{\text{all}})$ is 3-dimensional, spanned by [78]

$$\boldsymbol{v}_1 \propto [1\ 0\ \ldots\ 1\ 0]^{\text{T}}, \quad \boldsymbol{v}_2 \propto [0\ 1\ \ldots\ 0\ 1]^{\text{T}},$$
$$\boldsymbol{v}_3 \propto [\boldsymbol{u}(\phi_1 + \pi/2)^{\text{T}}\ \ldots\ \boldsymbol{u}(\phi_{N_{\text{a}}+1} + \pi/2)^{\text{T}}]^{\text{T}} \tag{87}$$

where ϕ_i is the angle from node i to the network centroid. Next we show that these vectors also span the row space of

the matrix $\widetilde{\boldsymbol{T}}^\infty \widetilde{\boldsymbol{D}}^{-1}$. Multiplying $\widetilde{\boldsymbol{T}}^\infty$ from the right at both sides of (58), we have

$$\widetilde{\boldsymbol{J}}_{\text{e}}(\boldsymbol{p}_{\text{all}})\widetilde{\boldsymbol{T}}^\infty = \widetilde{\boldsymbol{D}}(\boldsymbol{I} - \widetilde{\boldsymbol{T}})\widetilde{\boldsymbol{T}}^\infty$$
$$= \boldsymbol{0}. \tag{88}$$

Using the definition of $\widetilde{\boldsymbol{T}}$

$$\widetilde{\boldsymbol{T}} = \widetilde{\boldsymbol{D}}^{-1}\widetilde{\boldsymbol{A}}$$

we obtain

$$\widetilde{\boldsymbol{T}}^{\text{T}} = \boldsymbol{D}\widetilde{\boldsymbol{T}}\widetilde{\boldsymbol{D}}^{-1}. \tag{89}$$

Combining (88) and (89), we have

$$\boldsymbol{D}\widetilde{\boldsymbol{T}}^\infty\widetilde{\boldsymbol{D}}^{-1}\widetilde{\boldsymbol{J}}_{\text{e}}(\boldsymbol{p}_{\text{all}}) = \boldsymbol{0} \tag{90}$$

implying that the rows of $\widetilde{\boldsymbol{T}}^\infty\widetilde{\boldsymbol{D}}^{-1}$ can be written as linear combinations of the vectors given in (87). Therefore, from (87) we see that $\widetilde{\boldsymbol{T}}_{ii}^{(\infty)}\boldsymbol{D}_i^{-1} - \widetilde{\boldsymbol{T}}_{i\nu}^{(\infty)}\boldsymbol{D}_\nu^{-1}$ is rank-1, thus the following quantity

$$\widetilde{\boldsymbol{T}}_{ii}^{(\infty)}\boldsymbol{D}_i^{-1}\boldsymbol{D}_\nu - \widetilde{\boldsymbol{T}}_{i\nu}^{(\infty)} = (\widetilde{\boldsymbol{T}}_{ii}^{(\infty)}\boldsymbol{D}_i^{-1} - \widetilde{\boldsymbol{T}}_{i\nu}^{(\infty)}\boldsymbol{D}_\nu^{-1})\boldsymbol{D}_\nu \tag{91}$$

is also rank-1, and its row space is spanned by $\boldsymbol{u}_{i,\nu}$ which is a unit vector on the direction of $(\boldsymbol{u}(\phi_i+\pi/2) - \boldsymbol{u}(\phi_\nu+\pi/2))\boldsymbol{D}_\nu$.

In the proof of Lemma 1 we have shown that

$$\boldsymbol{H}_{\widetilde{\mathcal{R}}_i} - \boldsymbol{M}_{\widetilde{\mathcal{R}}_i} = \lim_{n\to\infty}\left\{\boldsymbol{P}_n(\boldsymbol{F}_{\widetilde{\mathcal{R}}_i} - \boldsymbol{I}_{\widetilde{\mathcal{R}}_i})\right\}. \tag{92}$$

Under the case where $|\widetilde{\mathcal{R}}_i| = 2$, we have

$$[\boldsymbol{P}_n]_{\widetilde{\mathcal{R}}_i} = \begin{bmatrix} \boldsymbol{0}_{2\times 2} & [\boldsymbol{P}_n]_{i,\nu} \\ [\boldsymbol{P}_n]_{\nu,i} & \boldsymbol{0}_{2\times 2} \end{bmatrix}$$

where $[\boldsymbol{P}_n]_{\widetilde{\mathcal{R}}_i}$ denote the matrix obtained by restricting $\boldsymbol{P}_n$ to the set $\widetilde{\mathcal{R}}_i$, and

$$[\boldsymbol{P}_n]_{i,\nu} = \sum_{k=1}^{n}\left(\widetilde{\boldsymbol{T}}_{ii}^{(k)}\boldsymbol{D}_i^{-1}\boldsymbol{D}_\nu - \widetilde{\boldsymbol{T}}_{i\nu}^{(k)}\right) \tag{93}$$

and vice versa for $[\boldsymbol{P}_n]_{\nu,i}$. Since the left hand side of (92) does exist, from (91), (92) and (93) we have

$$\boldsymbol{X}\boldsymbol{D}_{\widetilde{\mathcal{R}}_i}(\boldsymbol{F}_{\widetilde{\mathcal{R}}_i,\widetilde{\mathcal{R}}_i} - \boldsymbol{I}) = \boldsymbol{0}$$

where

$$\boldsymbol{X} = \begin{bmatrix} \boldsymbol{0} & \widetilde{\boldsymbol{T}}_{ii}^{(\infty)}\boldsymbol{D}_i^{-1} - \widetilde{\boldsymbol{T}}_{i\nu}^{(\infty)}\boldsymbol{D}_\nu^{-1} \\ \widetilde{\boldsymbol{T}}_{\nu\nu}^{(\infty)}\boldsymbol{D}_\nu^{-1} - \widetilde{\boldsymbol{T}}_{\nu i}^{(\infty)}\boldsymbol{D}_i^{-1} & \boldsymbol{0} \end{bmatrix}.$$

Since $\boldsymbol{F}_{\widetilde{\mathcal{R}}_i,\widetilde{\mathcal{R}}_i} - \boldsymbol{I}$ is a 4×4 matrix while $\boldsymbol{X}\boldsymbol{D}_{\widetilde{\mathcal{R}}_i}$ is rank-3, $\boldsymbol{F}_{\widetilde{\mathcal{R}}_i,\widetilde{\mathcal{R}}_i} - \boldsymbol{I}$ has to take the following form

$$\boldsymbol{F}_{\widetilde{\mathcal{R}}_i,\widetilde{\mathcal{R}}_i} - \boldsymbol{I} = -m\boldsymbol{D}_{\widetilde{\mathcal{R}}_i}^{-1}(\boldsymbol{v}_{i,\nu}\otimes[1\ -1]^{\text{T}})(\boldsymbol{v}_{i,\nu}\otimes[1\ -1]^{\text{T}})^{\text{T}} \tag{94}$$

where m is a positive scalar and $\boldsymbol{v}_{i,\nu} := \boldsymbol{u}(\phi_i) - \boldsymbol{u}(\phi_\nu)$. Therefore, the EFIM for agent i is given by

$$\boldsymbol{J}_{\text{e}}(\boldsymbol{p}_i) = \boldsymbol{D}_i\boldsymbol{F}(i,\nu) = m\boldsymbol{J}_{\text{r}}(\varphi_{i\nu}).$$

To investigate the asymptotic behavior of m, note that from (92) we have

$$\boldsymbol{I} - \boldsymbol{M}_{\widetilde{\mathcal{R}}_i,\widetilde{\mathcal{R}}_i} = \boldsymbol{P}_{\widetilde{\mathcal{R}}_i}^{(\text{r})}(\boldsymbol{F}_{\widetilde{\mathcal{R}}_i,\widetilde{\mathcal{R}}_i} - \boldsymbol{I}) \tag{95}$$

where $\boldsymbol{P}_{\widetilde{\mathcal{R}}_i}^{(\mathrm{r})} := \lim_{n \to \infty} [\boldsymbol{P}_n]_{\widetilde{\mathcal{R}}_i} \boldsymbol{U}$, $\boldsymbol{U} = [\boldsymbol{u}_{i,\nu}^{\mathrm{T}} \ \boldsymbol{u}_{i,\nu}^{\mathrm{T}}]^{\mathrm{T}} [\boldsymbol{u}_{i,\nu}^{\mathrm{T}} \ \boldsymbol{u}_{i,\nu}^{\mathrm{T}}]$. As the network expands, $\boldsymbol{P}_{\widetilde{\mathcal{R}}_i}^{(\mathrm{r})}$ will tends to $\boldsymbol{P}_{\widetilde{\mathcal{R}}_i} \boldsymbol{U}$, where $\boldsymbol{P}_{\widetilde{\mathcal{R}}_i}$ is the matrix $\boldsymbol{P}$ in infinite lattice networks restricted to the set $\widetilde{\mathcal{R}}_i$. Hence, these two matrices have the same asymptotic behaviors, implying that m scales as $m = \Theta(\log \|\boldsymbol{d}_{i\nu}\|)$ as long as both i and ν are far from the network edge as discussed in Sec. IV-A. Thus the proof is completed. $\qquad\square$

C. Proof of Proposition 5

Proof: We first prove the $O(\log \lambda_{\mathrm{anc}}^{-1})$ scaling, and then proceed to the $\Omega(\log \lambda_{\mathrm{anc}}^{-1})$ scaling.

1) Upper Bound: The average SPEB in the $N_{\mathrm{t}} = 1$ case provides a natural upper bound for that in the $N_{\mathrm{t}} \geqslant 2$ case. Hence we only consider the $N_{\mathrm{t}} = 1$ case here.

Since anchors are distributed according to (30), for any agent i, in each of the following two areas

$$\mathcal{A}_{\mathrm{h}}(i) := \{\boldsymbol{x} | \cos^2 \varphi_{\boldsymbol{x},\boldsymbol{p}_i} \geqslant c_{\mathrm{h}}\}, \ \mathcal{A}_{\mathrm{v}}(i) := \{\boldsymbol{x} | \sin^2 \varphi_{\boldsymbol{x},\boldsymbol{p}_i} \geqslant c_{\mathrm{v}}\}$$

the nearest anchor is within range $O(\lambda_{\mathrm{anc}}^{-\frac{1}{2}})$ of agent i, where c_{h} and c_{v} are constants in $(1/2, 1)$, and $\varphi_{\boldsymbol{x},\boldsymbol{p}_i}$ is the angle from $\boldsymbol{x}$ to $\boldsymbol{p}_i$. Therefore, for any agent i that is far from the network edge, i.e., satisfying $\lambda_{\mathrm{anc}}^{-\frac{1}{2}} = o(D_{\mathrm{net}}/2 - \|\boldsymbol{p}_i\|)$, we have $\mathrm{sp}\{\boldsymbol{p}_i\} = O(\log \lambda_{\mathrm{anc}}^{-1})$ using Propositions 1 and 4.

2) Lower Bound: Similar to the upper bound, it suffices to consider the $N_{\mathrm{t}} \geqslant 2$ case here. Since anchors are distributed according to (30), agents in the following area

$$\mathcal{A}_{\mathrm{sq}} := \Big\{ \boldsymbol{x} = [x_1 \ x_2]^{\mathrm{T}} \Big| x_1 \in \Big(\big(k + \tfrac{1}{4}\big)\lambda_{\mathrm{anc}}^{-\frac{1}{2}}, \big(k + \tfrac{3}{4}\big)\lambda_{\mathrm{anc}}^{-\frac{1}{2}} \Big),$$
$$x_2 \in \Big(\big(l + \tfrac{1}{4}\big)\lambda_{\mathrm{anc}}^{-\frac{1}{2}}, \big(l + \tfrac{3}{4}\big)\lambda_{\mathrm{anc}}^{-\frac{1}{2}} \Big), k \in \mathbb{Z}, l \in \mathbb{Z} \Big\}$$

are at least $1/4 \lambda_{\mathrm{anc}}^{-\frac{1}{2}}$ away from anchors. We can construct a lower bound for the SPEB of any agent i in $\mathcal{A}_{\mathrm{sq}}$, by considering the case where all nodes that are at least $1/4 \lambda_{\mathrm{anc}}^{-\frac{1}{2}}$ away from agent i are anchors. This implies that random walks started from i in the original network do not return if they once arrive a place at least $1/4 \lambda_{\mathrm{anc}}^{-\frac{1}{2}}$ away from $\boldsymbol{p}_i$. Hence we have

$$\boldsymbol{I} - \boldsymbol{F}_{ii} \succcurlyeq \sum_{n=0}^{m} \widetilde{\boldsymbol{T}}_n(\boldsymbol{0}, \boldsymbol{0}) = \frac{1}{(2\pi)^2} \int_{\mathcal{C}} \sum_{n=0}^{m} \boldsymbol{\Phi}^n(\boldsymbol{\vartheta}) \mathrm{d}\boldsymbol{\vartheta} \quad (96\mathrm{a})$$

$$\sim \frac{1}{(2\pi)^2} \int_{\mathcal{C}} \|\boldsymbol{\vartheta}\|^{-2} (\boldsymbol{I} - \boldsymbol{\Phi}^m(\boldsymbol{\vartheta})) \boldsymbol{\Sigma}_2^{-1}(\boldsymbol{\vartheta}) \mathrm{d}\boldsymbol{\vartheta} \quad (96\mathrm{b})$$

where we have applied techniques used in the proof of Proposition 9, and $m = \lfloor (2R_{\max})^{-1} \lambda_{\mathrm{anc}}^{-\frac{1}{2}} \rfloor$. Note that (78) implies that

$$\lim_{m \to \infty} \boldsymbol{\Phi}^m(\boldsymbol{\vartheta} m^{-\frac{1}{2}}) = \lim_{m \to \infty} \Big(\boldsymbol{I} - \frac{\|\boldsymbol{\vartheta}\|^2}{2m} \boldsymbol{\Sigma}_2(\boldsymbol{\vartheta}) \Big)^m$$
$$= \exp \Big\{ -\frac{\|\boldsymbol{\vartheta}\|^2}{2} \boldsymbol{\Sigma}_2(\boldsymbol{\vartheta}) \Big\}.$$

Thus with the substitution $\boldsymbol{\alpha} = \sqrt{m}\boldsymbol{\vartheta}$ we can rewrite the right hand side of (96b) as

$$\frac{1}{(2\pi)^2} \int_{\sqrt{n}\mathcal{C}} \|\boldsymbol{\alpha}\|^{-2} (\boldsymbol{I} - e^{-\frac{\|\boldsymbol{\alpha}\|^2}{2} \boldsymbol{\Sigma}_2(\boldsymbol{\alpha})}) \boldsymbol{\Sigma}_2^{-1}(\boldsymbol{\alpha}) \mathrm{d}\boldsymbol{\alpha}.$$

Using again the arguments in the proof of Proposition 9, we have

$$\Big\| \int_{\mathcal{C}} \|\boldsymbol{\vartheta}\|^{-2} (\boldsymbol{I} - \boldsymbol{\Phi}^m(\boldsymbol{\vartheta})) \boldsymbol{\Sigma}_2^{-1}(\boldsymbol{\vartheta}) \mathrm{d}\boldsymbol{\vartheta} \Big\|_2 \sim \int_0^{\sqrt{m}\pi} \frac{1 - e^{-\frac{r^2}{2}}}{r} \mathrm{d}r \quad (97)$$

as $\lambda_{\mathrm{anc}} \to 0_+$. Note that the right hand side of the last line of (97) scales as $\Theta(\log \lambda_{\mathrm{anc}}^{-1})$. Since the number of agents in the area $\mathcal{A}_{\mathrm{sq}}$ is proportional to the number of total agents, the proof is completed. $\qquad\square$

APPENDIX V
PROOF OF THEOREM 2

Proof: For the single anchor case, we have $\widetilde{\mathcal{R}}_i = \{i, \nu\}$. For any agent i ($i < \nu$ by definition), consider the equation (72) in Sec. III-B

$$\boldsymbol{F}_{\mathcal{S},\mathcal{S}} = \boldsymbol{I} + \boldsymbol{P}_{\mathcal{S}}^{-1} - \boldsymbol{P}_{\mathcal{S}}^{-1} \boldsymbol{Q} (\boldsymbol{Q}^{\mathrm{T}} \boldsymbol{D}_{\mathcal{S}} \boldsymbol{P}_{\mathcal{S}}^{-1} \boldsymbol{Q})^{-1} \boldsymbol{Q}^{\mathrm{T}} \boldsymbol{D}_{\mathcal{S}} \boldsymbol{P}_{\mathcal{S}}^{-1}$$

which can be rewritten as follows for stochastic geometric networks and $\mathcal{S} = \widetilde{\mathcal{R}}_i$

$$\mathbf{F}_{\widetilde{\mathcal{R}}_i, \widetilde{\mathcal{R}}_i} = \boldsymbol{I} + \mathbf{P}_{\widetilde{\mathcal{R}}_i}^{-1} - \mathbf{P}_{\widetilde{\mathcal{R}}_i}^{-1} \boldsymbol{Q} (\boldsymbol{Q}^{\mathrm{T}} \mathbf{D}_{\widetilde{\mathcal{R}}_i} \mathbf{P}_{\widetilde{\mathcal{R}}_i}^{-1} \boldsymbol{Q})^{-1} \boldsymbol{Q}^{\mathrm{T}} \mathbf{D}_{\widetilde{\mathcal{R}}_i} \mathbf{P}_{\widetilde{\mathcal{R}}_i}^{-1}. \quad (98)$$

Note that we have $\mathbf{F}_{\widetilde{\mathcal{R}}_i}(i, \nu) = [\mathbf{F}_{\widetilde{\mathcal{R}}_i, \widetilde{\mathcal{R}}_i}]_{1,2}$, and

$$\mathbf{P}_{\widetilde{\mathcal{R}}_i} = \begin{bmatrix} \mathbf{0}_{2 \times 2} & [\mathbf{P}_{\widetilde{\mathcal{R}}_i}]_{1,2} \\ [\mathbf{P}_{\widetilde{\mathcal{R}}_i}]_{2,1} & \mathbf{0}_{2 \times 2} \end{bmatrix}$$

The submatrices on the diagonal of $\mathbf{P}_{\widetilde{\mathcal{R}}_i}$ are $\mathbf{0}_{2 \times 2}$, as can be seen from the definition of matrix $\boldsymbol{P}$

$$\boldsymbol{P} = \sum_{\{i,j\} \subseteq \mathcal{N}_{\mathrm{a}} \cup \mathcal{N}_{\mathrm{b}}} \boldsymbol{E}_{ij}^{N_{\mathrm{a}} + N_{\mathrm{b}}} \otimes \Big\{ \sum_{n=0}^{\infty} \big(\widetilde{\boldsymbol{T}}_{ii}^{(n)} \boldsymbol{D}_i^{-1} \boldsymbol{D}_j - \widetilde{\boldsymbol{T}}_{ij}^{(n)} \big) \Big\}$$

in which for diagonal blocks (i.e., $i = j$), we have $[\boldsymbol{P}]_{i,i} = \sum_{n=0}^{\infty} \big(\widetilde{\boldsymbol{T}}_{ii}^{(n)} \boldsymbol{D}_i^{-1} \boldsymbol{D}_i - \widetilde{\boldsymbol{T}}_{ii}^{(n)} \big) = \mathbf{0}_{2 \times 2}$. To simplify the notation, we denote $[\mathbf{P}_{\widetilde{\mathcal{R}}_i}^{-1}]_{1,2}$ by $\mathbf{K}_{i,\nu}$ and $[\mathbf{P}_{\widetilde{\mathcal{R}}_i}^{-1}]_{2,1}$ by $\mathbf{K}_{\nu,i}$, whereas denote $[\mathbf{P}_{\widetilde{\mathcal{R}}_i}]_{1,2}$ by $\mathbf{P}_{i,\nu}$ and $[\mathbf{P}_{\widetilde{\mathcal{R}}_i}]_{2,1}$ by $\mathbf{P}_{\nu,i}$. From (66) we have $\mathbf{P}_{i,\nu} = \mathbf{K}_{\nu,i}$. Thus

$$\mathbf{P}_{\widetilde{\mathcal{R}}_i} = \begin{bmatrix} \mathbf{0}_{2 \times 2} & \mathbf{P}_{i,\nu} \\ \mathbf{P}_{\nu,i} & \mathbf{0}_{2 \times 2} \end{bmatrix}, \quad \mathbf{P}_{\widetilde{\mathcal{R}}_i}^{-1} = \begin{bmatrix} \mathbf{0}_{2 \times 2} & \mathbf{K}_{i,\nu} \\ \mathbf{K}_{\nu,i} & \mathbf{0}_{2 \times 2} \end{bmatrix}. \quad (99)$$

The terms $\mathbf{P}_{\widetilde{\mathcal{R}}_i}^{-1} \boldsymbol{Q}$ and $\boldsymbol{Q}^{\mathrm{T}} \mathbf{D}_{\widetilde{\mathcal{R}}_i}$ can be written as

$$\mathbf{P}_{\widetilde{\mathcal{R}}_i}^{-1} \boldsymbol{Q} = \begin{bmatrix} \mathbf{K}_{i,\nu} \\ \mathbf{K}_{\nu,i} \end{bmatrix}, \quad \boldsymbol{Q}^{\mathrm{T}} \mathbf{D}_{\widetilde{\mathcal{R}}_i} = [\mathbf{D}_i \ \mathbf{D}_\nu] \quad (100)$$

Substituting (100) into (98), we have

$$\mathbf{F}_{\widetilde{\mathcal{R}}_i, \widetilde{\mathcal{R}}_i} = \boldsymbol{I} + \begin{bmatrix} \mathbf{0}_{2 \times 2} & \mathbf{K}_{i,\nu} \\ \mathbf{K}_{\nu,i} & \mathbf{0}_{2 \times 2} \end{bmatrix} - \begin{bmatrix} \mathbf{K}_{i,\nu} \\ \mathbf{K}_{\nu,i} \end{bmatrix}$$
$$(\mathbf{D}_i \mathbf{K}_{i,\nu} + \mathbf{D}_\nu \mathbf{K}_{\nu,i})^{-1} [\mathbf{D}_i \ \mathbf{D}_\nu] \mathbf{P}_{\widetilde{\mathcal{R}}_i}^{-1}$$
$$= \begin{bmatrix} \boldsymbol{I}_2 & \mathbf{K}_{i,\nu} \\ \mathbf{K}_{\nu,i} & \boldsymbol{I}_2 \end{bmatrix} - \begin{bmatrix} \mathbf{K}_{i,\nu} \\ \mathbf{K}_{\nu,i} \end{bmatrix}$$
$$(\mathbf{D}_i \mathbf{K}_{i,\nu} + \mathbf{D}_\nu \mathbf{K}_{\nu,i})^{-1} [\mathbf{D}_\nu \mathbf{K}_{\nu,i} \ \mathbf{D}_i \mathbf{K}_{i,\nu}]$$

For the desired term $\mathbf{F}_{\widetilde{\mathcal{R}}_i}(i, \nu)$, we have

$$\mathbf{F}_{\widetilde{\mathcal{R}}_i}(i, \nu) = \mathbf{K}_{i,\nu} \Big(\boldsymbol{I} - (\mathbf{D}_i \mathbf{K}_{i,\nu} + \mathbf{D}_\nu \mathbf{K}_{\nu,i})^{-1} \mathbf{D}_i \mathbf{K}_{i,\nu} \Big)$$
$$= \mathbf{K}_{i,\nu} \Big(\mathbf{D}_i \mathbf{K}_{i,\nu} + \mathbf{D}_\nu \mathbf{K}_{\nu,i} \Big)^{-1} \mathbf{D}_\nu \mathbf{K}_{\nu,i}$$

Note that the expression of $\mathbf{P}_{\widetilde{\mathcal{R}}_i}^{-1}$ above implies $\mathbf{P}_{i,\nu} = \mathbf{K}_{\nu,i}$, the inverse of $\mathbf{F}_{\widetilde{\mathcal{R}}_i}(i,\nu)$ takes the following form

$$
\begin{aligned}
\mathbf{F}_{\widetilde{\mathcal{R}}_i}^{-1}(i,\nu) &= \left(\mathbf{K}_{i,\nu}\left(\mathbf{D}_i\mathbf{K}_{i,\nu} + \mathbf{D}_\nu\mathbf{K}_{\nu,i}\right)^{-1}\mathbf{D}_\nu\mathbf{K}_{\nu,i}\right)^{-1} \\
&= \mathbf{P}_{i,\nu}\mathbf{D}_\nu^{-1}\left(\mathbf{D}_i\mathbf{K}_{i,\nu} + \mathbf{D}_\nu\mathbf{K}_{\nu,i}\right)\mathbf{P}_{\nu,i} \\
&= \left(\boldsymbol{I} + \mathbf{P}_{i,\nu}\mathbf{D}_\nu^{-1}\mathbf{D}_i\mathbf{K}_{i,\nu}\right)\mathbf{P}_{\nu,i} \\
&= \mathbf{P}_{\nu,i}\left(\boldsymbol{I} + \mathbf{K}_{i,\nu}\mathbf{P}_{i,\nu}\mathbf{D}_\nu^{-1}\mathbf{D}_i\right).
\end{aligned}
$$

Now let us consider the convergence from $\mathbf{F}_{\widetilde{\mathcal{R}}_i}^{-1}(i,\nu)$ to $2\mathbf{P}_{\nu,i}$ using the following inequality

$$
\begin{aligned}
&\left\|\mathbf{P}_{\nu,i}(\mathbf{K}_{i,\nu}\mathbf{P}_{i,\nu}\mathbf{D}_\nu^{-1}\mathbf{D}_i - \boldsymbol{I})\right\| \\
&\leqslant \|\mathbf{P}_{\nu,i}\|\left(\|\boldsymbol{I}-\mathbf{D}_\nu^{-1}\mathbf{D}_i\| + \|\mathbf{K}_{i,\nu}\|\|\mathbf{P}_{i,\nu}-\mathbf{P}_{\nu,i}\|\|\mathbf{D}_\nu^{-1}\mathbf{D}_i\|\right)
\end{aligned}
$$
(101)

Here, $\|\cdot\|$ can be any submultiplicative matrix norm. From (101) it is clear that we have to show the convergence from $\mathbf{D}_i$ to $\mathbf{D}_\nu$ and that from $\mathbf{P}_{i,\nu}$ to $\mathbf{P}_{\nu,i}$. Apparently, this requires us to quantify the similarity between RGG and lattice graphs. To this end, we first introduce the concept of *minimax lattice matching error* [82].

Definition 4 (Minimax Lattice Matching Error): Consider a set $\mathcal{S}_{\mathrm{L}}$ of N lattice points arranged as a $\sqrt{N}\times\sqrt{N}$ array within a square in $\mathbb{R}^2$ with area A. For any set $\mathcal{S}_{\mathrm{R}}$ of N uniformly random distributed points, the *minimax lattice matching error* ε is the minimum length such that there exists a one-to-one mapping from $\mathcal{S}_{\mathrm{R}}$ to $\mathcal{S}_{\mathrm{L}}$ in the square, for which the distance between every matched pair does not exceed ε.

Intuitively, if the minimax lattice matching error ε tends to zero, the stochastic geometric network would have an almost identical structure as that of lattice networks, and hence the corresponding average SPEB may also exhibit similar asymptotic behaviors. According to [82], for any stochastic geometric networks containing $N \to \infty$ nodes, if the network keeps connected with probability approaching 1 as its area $|\mathcal{R}_{\mathrm{net}}| \to \infty$, we have $\varepsilon \to 0$ with probability at least $1 - N^{-1}$.

Next, we proceed to analyze the relation between $\widetilde{\mathbf{T}}_{i\nu}^{(1)}$ in the stochastic geometric network and $\widetilde{T}_{j\mu}^{(1)}$ in the minimax matched lattice network, where j and μ correspond to the matched nodes of i and ν, respectively. Note that

$$
\frac{\|\widetilde{\mathbf{T}}_{i\nu}^{(1)} - \widetilde{T}_{j\mu}^{(1)}\|}{\|\widetilde{\mathbf{T}}_{i\nu}^{(1)}\|} \leqslant \|\boldsymbol{I} - \boldsymbol{D}_j^{-1}\boldsymbol{J}_{j\mu}\mathbf{J}_{i\nu}^{-1}\mathbf{D}_i\|
$$

which implies that $\widetilde{\mathbf{T}}_{i\nu}^{(1)}$ converges to $\widetilde{T}_{j\mu}^{(1)}$ if $\boldsymbol{D}_j^{-1}\boldsymbol{J}_{j\mu}\mathbf{J}_{i\nu}^{-1}\mathbf{D}_i$ converges to $\boldsymbol{I}$.

We first show that the eigenvalues of $\boldsymbol{J}_{j\mu}\mathbf{J}_{i\nu}^{-1}$ converge to 1 for all matched pairs (i,j). Since $\boldsymbol{J}_{j\mu}\mathbf{J}_{i\nu}^{-1}$ is a 2×2 matrix, using (4) we can derive its eigenvalues by directly solving the eigenvalues equations, and obtain

$$
\begin{aligned}
\lambda_1 = &\frac{\widetilde{\xi}_{i\nu,\mathrm{r}}}{\widetilde{\xi}_{j\mu,\mathrm{r}}}\cdot\frac{1}{2\mathsf{d}^2}\Big\{(\mathsf{d}^2 d_{\mathrm{L}}^2 + 1)\sin^2\theta + (\mathsf{d}^2+d_{\mathrm{L}}^2)\cos^2\theta \\
&+ \sqrt{\left[(\mathsf{d}^2 d_{\mathrm{L}}^2+1)\sin^2\theta + (\mathsf{d}^2+d_{\mathrm{L}}^2)\cos^2\theta\right]^2 - 4\mathsf{d}^2 d_{\mathrm{L}}^2}\Big\}
\end{aligned}
$$
(102)

$$
\begin{aligned}
\lambda_2 = &\frac{\widetilde{\xi}_{i\nu,\mathrm{r}}}{\widetilde{\xi}_{j\mu,\mathrm{r}}}\cdot\frac{1}{2\mathsf{d}^2}\Big\{(\mathsf{d}^2 d_{\mathrm{L}}^2 + 1)\sin^2\theta + (\mathsf{d}^2+d_{\mathrm{L}}^2)\cos^2\theta \\
&- \sqrt{\left[(\mathsf{d}^2 d_{\mathrm{L}}^2+1)\sin^2\theta + (\mathsf{d}^2+d_{\mathrm{L}}^2)\cos^2\theta\right]^2 - 4\mathsf{d}^2 d_{\mathrm{L}}^2}\Big\}
\end{aligned}
$$
(103)

where d is the distance between node i and ν, d_{L} is the distance between j and μ, and θ is the angle between $\boldsymbol{d}_{i\nu}$ and $\boldsymbol{d}_{j\mu}$. Denoting the minimax lattice matching error by ε, using the triangle inequality we see that $|\mathsf{d} - d_{\mathrm{L}}| \leqslant 2\varepsilon$. Thus for all pairs (i,ν) satisfying $\varepsilon = o(\mathsf{d})$, we have $d_{\mathrm{L}} = \mathsf{d} + o(\mathsf{d})$. From (102) and (103) we can now see that both λ_1 and λ_2 tend to 1 with probability[6] exceeding $1 - 1/N_{\mathrm{a}}$. As for the eigenvectors, we observe that

$$
\boldsymbol{J}_{j\mu} = \boldsymbol{U}_{j\mu}\boldsymbol{\Lambda}_{j\mu}\boldsymbol{U}_{j\mu}^{\mathrm{T}}, \quad \mathbf{J}_{i\nu} = \mathbf{U}_{i\nu}\boldsymbol{\Lambda}_{i\nu}\mathbf{U}_{i\nu}^{\mathrm{T}}
$$
(104)

where $\boldsymbol{\Lambda}_{j\mu}$ and $\boldsymbol{\Lambda}_{i\nu}$ tend to $\boldsymbol{I}_2$ according to previous discussions. Thus it now suffices to show that $\|\boldsymbol{I} - \boldsymbol{U}_{j\mu}^{\mathrm{T}}\mathbf{U}_{i\nu}\| \to 0$. This is equivalent to showing that $\cos(\varphi_{j\mu})\cos(\varphi_{i\nu}) + \sin(\varphi_{j\mu})\sin(\varphi_{i\nu}) \to 1$, which is further equivalent to showing that $\theta \to 0$. The latter is straightforward for all pairs satisfying $\varepsilon = o(\mathsf{d})$, and hence we have $\|\boldsymbol{I} - \boldsymbol{D}_j^{-1}\mathbf{D}_i\| \to 0$.

Since $\mathbf{D}_i = \sum_{k\in\mathrm{N}_i}\mathbf{J}_{\mathbf{d}_{ik}}$, we see that the eigenvalues of $\boldsymbol{D}_j^{-1}\mathbf{D}_i$ also tends to 1 for all matched pairs (i,j) as the maximum communication range satisfies $\varepsilon = o(R_{\max})$. Following a similar line of reasoning as (104), we can also show that the eigenvectors of $\boldsymbol{D}_j^{-1}$ and $\mathbf{D}_i$ tend to be identical. Therefore, for pairs (i,ν) satisfying $\varepsilon = o(\mathsf{d})$, $\widetilde{\mathbf{T}}_{i\nu}^{(1)}$ converges to $\widetilde{T}_{j\mu}^{(1)}$. As the network expands, such pairs will be dominantly many and hence we have the convergence from $\widetilde{\mathbf{T}}_{i\nu}^{(n)}$ to $\widetilde{T}_{j\mu}^{(n)}$ for any finite n. Consequently, we can now conclude that $\|\mathbf{P}_{i,\nu} - \mathbf{P}_{j,\mu}\|$ tends to zero.

An important corollary of the previous analysis is that both $\|\mathbf{P}_{i,\nu} - \mathbf{P}_{\nu,i}\|$ and $\|\boldsymbol{I} - \mathbf{D}_i\mathbf{D}_\nu^{-1}\|$ tend to zero. In light of this, from (101) we can now work on the simpler term of $2\mathbf{P}_{i,\nu}$. To finish the proof, it now suffices to show that

$$
\frac{\mathrm{tr}\{\sum_{\mathbf{p}_i\in\mathsf{S}_\nu}\left(\mathbf{P}_{i,\nu}\mathbf{D}_i^{-1} - P_{j,\mu}\boldsymbol{D}_j^{-1}\right)\}}{\mathrm{tr}\{\sum_{\mathbf{p}_i\in\mathsf{S}_\nu}\mathbf{P}_{i,\nu}\mathbf{D}_i^{-1}\}} \to 0.
$$

For each 4-tuple (i,j,ν,μ), we have

$$
\begin{aligned}
&\|\mathbf{P}_{i,\nu}\mathbf{D}_i^{-1} - P_{j,\mu}\boldsymbol{D}_j^{-1}\|\cdot\|\mathbf{P}_{i,\nu}\mathbf{D}_i^{-1}\|^{-1} \\
&\leqslant \|\boldsymbol{I} - \mathbf{D}_i\boldsymbol{D}_j^{-1}\| + \|\mathbf{D}_i\|\|\boldsymbol{D}_j^{-1}\|\|\mathbf{K}_{\nu,i}\|\|\mathbf{P}_{i,\nu} - P_{j,\mu}\|
\end{aligned}
$$

which indeed tends to zero since $\|\mathbf{K}_{\nu,i}\| \to 0$.

Similar arguments can also be applied to the $N_{\mathrm{t}} = 1$ case, except that one has to work with the Moore-Penrose pseudo-inverse of $\mathbf{P}_{\widetilde{\mathcal{R}}_i}^{(\mathrm{r})}$. With similar manipulations we can obtain (34), thus the proof is completed. $\square$

APPENDIX VI
PROOF OF PROPOSITION 6

Proof: For any agent i in the interior area of the network, we consider the areas $\mathsf{A}_{\mathrm{h}}(i)$ and $\mathsf{A}_{\mathrm{v}}(i)$ defined as (we use upright letters since they are now random sets)

$$
\begin{aligned}
\mathcal{A}_{\mathrm{h}}(i) &:= \{\boldsymbol{x}\,|\cos^2\varphi_{\boldsymbol{x},\boldsymbol{p}_i} \geqslant c_{\mathrm{h}}\} \\
\mathcal{A}_{\mathrm{v}}(i) &:= \{\boldsymbol{x}\,|\sin^2\varphi_{\boldsymbol{x},\boldsymbol{p}_i} \geqslant c_{\mathrm{v}}\}.
\end{aligned}
$$
(105)

[6]For simplicity, we shall omit the statement of "with probability $1-1/N_{\mathrm{a}}$" in the rest of this proof.

Since anchors form a binomial point process, the probability that there is no anchor within range R from $\mathbf{p}_i$ in $A_h(i)$ is given by

$$p_{0,h}(R) = (|\mathcal{R}_{net}| - 2\cos^{-1}(\sqrt{c_h})\pi R^2)^{N_b}|\mathcal{R}_{net}|^{-N_b}. \quad (106)$$

Choosing $R = k\lambda_{anc}^{-\frac{1}{2}}$, we have

$$p_{0,h}(k\lambda_{anc}^{-\frac{1}{2}}) = \left(1 - \frac{2k^2(\pi\cos^{-1}(\sqrt{c_h}))}{|\mathcal{R}_{net}|\lambda_{anc}}\right)^{|\mathcal{R}_{net}|\lambda_{anc}} \quad (107)$$

and $\lim_{|\mathcal{R}_{net}|\to\infty} p_{0,h}(k\lambda_{anc}^{\frac{1}{2}}) = e^{-2k\pi\cos^{-1}(\sqrt{c_h})}$. Therefore, for $R = \Omega(\lambda_{anc}^{-\frac{1}{2}+\epsilon})$ where $\epsilon > 0$ can be arbitrarily small, we have $p_{0,h}(R) \to 0$ as the network expands. Similar arguments also applies to $A_v(i)$. By application of union bound, we see that with probability approaching 1, there are anchors in both $A_h(i)$ and $A_v(i)$ which are within range $O(\lambda_{anc}^{-1})$ of $\mathbf{p}_i$. Thus using Theorem 2, we can conclude that the average SPEB of agents in the interior area of the network scales as $O(\log \lambda_{anc}^{-1})$. $\qquad\square$

REFERENCES

[1] M. Z. Win, Y. Shen, and W. Dai, "A theoretical foundation of network localization and navigation," *Proc. IEEE*, vol. 106, no. 7, pp. 1136–1165, Jul. 2018.

[2] S. Gezici *et al.*, "Localization via ultra-wideband radios: A look at positioning aspects for future sensor networks," *IEEE Signal Process. Mag.*, vol. 22, no. 4, pp. 70–84, Jul. 2005.

[3] Y. Shen, S. Mazuelas, and M. Z. Win, "Network navigation: Theory and interpretation," *IEEE J. Sel. Areas Commun.*, vol. 30, no. 9, pp. 1823–1834, Oct. 2012.

[4] L. Chen *et al.*, "Robustness, security and privacy in location-based services for future IoT: A survey," *IEEE Access*, vol. 5, pp. 8956–8977, Apr. 2017.

[5] S. D'Oro, L. Galluccio, G. Morabito, and S. Palazzo, "Exploiting object group localization in the internet of things: Performance analysis," *IEEE Trans. Veh. Technol.*, vol. 64, no. 8, pp. 3645–3656, Aug. 2015.

[6] C. L. Nguyen, O. Georgiou, Y. Yonezawa, and Y. Doi, "The wireless localization matching problem," *IEEE Internet Things J.*, vol. 4, no. 5, pp. 1312–1326, Oct. 2017.

[7] M. Z. Win *et al.*, "Efficient multisensor localization for the Internet of things: Exploring a new class of scalable localization algorithms," *IEEE Signal Process. Mag.*, vol. 35, no. 5, pp. 153–167, Sep. 2018.

[8] F. Meyer *et al.*, "Message passing algorithms for scalable multitarget tracking," *Proc. IEEE*, vol. 106, no. 2, pp. 221–259, Feb. 2018.

[9] S. Kuutti, S. Fallah, K. Katsaros, M. Dianati, F. Mccullough, and A. Mouzakitis, "A survey of the state-of-the-art localisation techniques and their potentials for autonomous vehicle applications," *IEEE Internet Things J.*, vol. PP, no. 99, pp. 1–1, 2018.

[10] S. Güler, B. Fidan, S. Dasgupta, B. D. O. Anderson, and I. Shames, "Adaptive source localization based station keeping of autonomous vehicles," *IEEE Trans. Autom. Control*, vol. 62, no. 7, pp. 3122–3135, Jul. 2017.

[11] F. Zabini and A. Conti, "Inhomogeneous Poisson sampling of finite-energy signals with uncertainties in $\mathbb{R}^d$," *IEEE Trans. Signal Process.*, vol. 64, no. 18, pp. 4679–4694, Sep. 2016.

[12] M. Elhamshary, A. Basalmah, and M. Youssef, "A fine-grained indoor location-based social network," *IEEE Trans. Mobile Comput.*, vol. 16, no. 5, pp. 1203–1217, May 2017.

[13] A. Sandryhaila and J. M. F. Moura, "Big data analysis with signal processing on graphs: Representation and processing of massive data sets with irregular structure," *IEEE Signal Process. Mag.*, vol. 31, no. 5, pp. 80–90, Sep. 2014.

[14] T. Wang, Y. Shen, A. Conti, and M. Z. Win, "Network navigation with scheduling: Error evolution," *IEEE Trans. Inf. Theory*, vol. 63, no. 11, pp. 7509–7534, Nov. 2017.

[15] M. Z. Win, W. Dai, Y. Shen, G. Chrisikos, and H. V. Poor, "Network operation strategies for efficient localization and navigation," *Proc. IEEE*, vol. 106, no. 7, pp. 1224–1254, Jul. 2018.

[16] Z. Liu, W. Dai, and M. Z. Win, "Mercury: An infrastructure-free system for network localization and navigation," *IEEE Trans. Mobile Comput.*, vol. 17, no. 5, pp. 1119–1133, May 2018.

[17] M. Sun and K. C. Ho, "Successive and asymptotically efficient localization of sensor nodes in closed-form," *IEEE Trans. Signal Process.*, vol. 57, no. 11, pp. 4522–4537, Nov. 2009.

[18] H. Wymeersch, J. Lien, and M. Z. Win, "Cooperative localization in wireless networks," *Proc. IEEE*, vol. 97, no. 2, pp. 427–450, Feb. 2009.

[19] U. A. Khan, S. Kar, and J. M. F. Moura, "Linear theory for self-localization: Convexity, barycentric coordinates, and Cayley-Menger determinants," *IEEE Access*, vol. 3, pp. 1326–1339, Aug. 2015.

[20] A. Simonetto and G. Leus, "Distributed maximum likelihood sensor network localization," *IEEE Trans. Signal Process.*, vol. 62, no. 6, pp. 1424–1437, Mar. 2014.

[21] W. Yuan, N. W. an Bernhard Etzlinger, H. Wang, and J. Kuang, "Cooperative joint localization and clock synchronization based on Gaussian message passing in asynchronous wireless networks," *IEEE Trans. Veh. Technol.*, vol. 65, no. 9, pp. 7258–7273, Sep. 2016.

[22] S. Mazuelas, A. Conti, J. C. Allen, and M. Z. Win, "Soft range information for network localization," *IEEE Trans. Signal Process.*, vol. 66, no. 12, pp. 3155–3168, Jun. 2018.

[23] A. Conti, S. Mazuelas, S. Bartoletti, W. C. Lindsey, and M. Z. Win, "Soft information for localization-of-things," *Proc. IEEE*, vol. 107, no. 11, pp. 2240–2264, Nov. 2019.

[24] C. Meng, Z. Ding, and S. Dasgupta, "A semidefinite programming approach to source localization in wireless sensor networks," *IEEE Signal Process. Lett.*, vol. 15, pp. 253–256, Feb. 2008.

[25] H. Shen, Z. Ding, S. Dasgupta, and C. Zhao, "Multiple source localization in wireless sensor networks based on time of arrival measurement," *IEEE Trans. Signal Process.*, vol. 62, no. 8, pp. 1938–1949, Apr. 2014.

[26] E. Xu, Z. Ding, and S. Dasgupta, "Source localization in wireless sensor networks from signal time-of-arrival measurements," *IEEE Trans. Signal Process.*, vol. 59, no. 6, pp. 2887–2897, Jun. 2011.

[27] S. Bartoletti, W. Dai, A. Conti, and M. Z. Win, "A mathematical model for wideband ranging," *IEEE J. Sel. Topics Signal Process.*, vol. 9, no. 2, pp. 216–228, Mar. 2015.

[28] R. D. Taranto, S. Muppirisetty, R. Raulefs, D. Slock, T. Svensson, and H. Wymeersch, "Location-aware communications for 5G networks: How location information can improve scalability, latency, and robustness of 5G," *IEEE Signal Process. Mag.*, vol. 31, no. 6, pp. 102–112, Nov. 2014.

[29] N. Garcia, H. Wymeersch, E. G. Larsson, A. M. Haimovich, and M. Coulon, "Direct localization for massive MIMO," *IEEE Trans. Signal Process.*, vol. 65, no. 10, pp. 2475–2487, May 2017.

[30] K. Witrisal *et al.*, "High-accuracy localization for assisted living: 5G systems will turn multipath channels from foe to friend," *IEEE Signal Process. Mag.*, vol. 33, no. 2, pp. 59–70, Mar. 2016.

[31] F. Zhou and Y. Shen, "On the outage probability of localization in randomly deployed wireless networks," *IEEE Commun. Lett.*, vol. 21, no. 4, pp. 901–904, Apr. 2017.

[32] R. M. Vaghefi and R. M. Buehrer, "Cooperative source node tracking in non-line-of-sight environments," *IEEE Trans. Mobile Comput.*, vol. 16, no. 5, pp. 1287–1299, May 2017.

[33] Y. Xiong, N. Wu, H. Wang, and J. Kuang, "Cooperative detection-assisted localization in wireless networks in the presence of ranging outliers," *IEEE Trans. Commun.*, vol. 65, no. 12, pp. 5165–5179, Dec. 2017.

[34] M. Z. Win *et al.*, "Network localization and navigation via cooperation," *IEEE Commun. Mag.*, vol. 49, no. 5, pp. 56–62, May 2011.

[35] Y. Han, Y. Shen, X. P. Zhang, M. Z. Win, and H. Meng, "Performance limits and geometric properties of array localization," *IEEE Trans. Inf. Theory*, vol. 62, no. 2, pp. 1054–1075, Feb. 2016.

[36] J. T. Isaacs, D. J. Klein, and J. P. Hespanha, "Optimal sensor placement for time difference of arrival localization," in *Proc. 48th IEEE Conf. CDC/CCC 2009*. Shanghai, China: IEEE, 2009, pp. 7878–7884.

[37] A. N. Bishop, B. Fidan, B. D. O. Anderson, K. Dogançay, and P. N. Pathirana, "Optimality analysis of sensor-target localization geometries," *Automatica*, vol. 46, no. 3, pp. 479–492, Mar. 2010.

[38] H. K. Achanta, W. Xu, and S. Dasgupta, "Matrix design for optimal sensing," in *Proc. 2013 IEEE Int. Conf. Acoustics, Speech and Signal Process.* Vancouver, Canada: IEEE, 2013, pp. 4221–4225.

[39] S. Dasgupta, S.-R. C. Ibeawuchi, and Z. Ding, "Optimum sensor placement for localization under log-normal shadowing," in *Proc. 2010 Int. Symp. Commun. Inform. Technol. (ISCIT)*. Tokyo, Japan: IEEE, 2010, pp. 204–208.

[40] H. K. Achanta, S. Dasgupta, and Z. Ding, "Optimum sensor placement for localization in three dimensional under log normal shadowing,"

in *Proc. 2012 5th Int. Congress Image and Signal Process. (CISP).* Chongqing, China: IEEE, 2012, pp. 1898–1901.

[41] S. Dasgupta, S.-R. C. Ibeawuchi, and Z. Ding, "Optimum sensor placement for source monitoring under log-normal shadowing in three dimensions," in *Proc. 2009 Int. Symp. Commun. Inform. Technol. (ISCIT).* Piscataway, NJ, USA: IEEE, 2009, pp. 376–381.

[42] Z. Liu, F. Meyer, and M. Z. Win, "On the accuracy of network localization and synchronization," in *IEEE 10th Latin-American Conf. Commun.*, Guadalajara, Mexico, Nov. 2018, pp. 1–6.

[43] W. Dai, Y. Shen, and M. Z. Win, "Energy-efficient network navigation algorithms," *IEEE J. Sel. Areas Commun.*, vol. 33, no. 7, pp. 1418–1430, Jul. 2015.

[44] W. Dai, Y. Shen, and M. Z. Win, "A computational geometry framework for efficient network localization," *IEEE Trans. Inf. Theory*, vol. 64, no. 2, pp. 1317–1339, Feb. 2018.

[45] Y. Shen, H. Wymeersch, and M. Z. Win, "Fundamental limits of wideband localization – Part II: Cooperative networks," *IEEE Trans. Inf. Theory*, vol. 56, no. 10, pp. 4981–5000, Oct. 2010.

[46] Y. Shen and M. Z. Win, "Fundamental limits of wideband localization – Part I: A general framework," *IEEE Trans. Inf. Theory*, vol. 56, no. 10, pp. 4956–4980, Oct. 2010.

[47] Y. Xiong, N. Wu, Y. Shen, and M. Z. Win, "Cooperative network synchronization: Asymptotic analysis," *IEEE Trans. Signal Process.*, vol. 66, no. 3, pp. 757–772, Feb. 2018.

[48] S. Mazuelas, Y. Shen, and M. Z. Win, "Spatiotemporal information coupling in network navigation," *IEEE Trans. Inf. Theory*, vol. 64, no. 12, pp. 7759–7779, Dec. 2018.

[49] J. Schloemann and R. M. Buehrer, "On the value of collaboration in location estimation," *IEEE Trans. Veh. Technol.*, vol. 65, no. 5, pp. 3585–3596, May 2016.

[50] J. Schloemann, H. S. Dhillon, and R. M. Buehrer, "Toward a tractable analysis of localization fundamentals in cellular networks," *IEEE Trans. Wireless Commun.*, vol. 15, no. 3, pp. 1768–1782, Mar. 2016.

[51] Y. Xiong, N. Wu, and H. Wang, "On the performance limits of cooperative localization in wireless sensor networks with strong sensor position uncertainty," *IEEE Commun. Lett.*, vol. 21, no. 7, pp. 1613–1616, Jul. 2017.

[52] Y. Xiong, J. Kuang, Y. Feng, H. Wang, and N. Wu, "Understanding the efficiency of cooperation in location-aware wireless networks," in *Proc. 2017 IEEE Int.l Conf. Commun. Workshops*, Paris, France, May 2017, pp. 828–833.

[53] B. Huang, C. Yu, and B. D. O. Anderson, "Understanding error propagation in multihop sensor network localization," *IEEE Trans. Ind. Electron.*, vol. 60, no. 12, pp. 5811–5819, Dec. 2013.

[54] D. Niculescu and B. Nath, "Ad hoc positioning system (APS)," in *Proc. IEEE Global Commun. Conf.*, San Antonio, TX, Nov. 2001, pp. 2926–2931.

[55] B. Huang, C. Yu, B. D. O. Anderson, and G. Mao, "Estimating distances via connectivity in wireless sensor networks," *Wireless Commun. Mobile Comput.*, vol. 14, no. 5, pp. 541–556, May 2014.

[56] S. Zaidi, A. E. Assaf, S. Affes, and N. Kandil, "Accurate range-free localization in multi-hop wireless sensor networks," *IEEE Trans. Commun.*, vol. 64, no. 9, pp. 3886–3900, Sep. 2016.

[57] Y. Xiong, N. Wu, Y. Shen, and M. Z. Win, "Cooperative network synchronization: Asymptotic analysis," *IEEE Trans. Signal Process.*, vol. 66, no. 3, pp. 757–772, Feb. 2018.

[58] M. Haenggi, J. G. Andrews, F. Baccelli, O. Dousse, and M. Franceschetti, "Stochastic geometry and random graphs for the analysis and design of wireless network," *IEEE J. Sel. Areas Commun.*, vol. 27, no. 7, pp. 1029–1046, Sep. 2009.

[59] D. Dardari, A. Conti, C. Buratti, and R. Verdone, "Mathematical evaluation of environmental monitoring estimation error through energy-efficient wireless sensor networks," *IEEE Trans. Mobile Comput.*, vol. 6, no. 7, pp. 790–802, Jul. 2007.

[60] F. Baccelli and A. Giovanidis, "A stochastic geometry framework for analyzing pairwise-cooperative cellular networks," *IEEE Trans. Wireless Commun.*, vol. 14, no. 2, pp. 794–808, Feb. 2015.

[61] M. Z. Win, P. C. Pinto, and L. A. Shepp, "A mathematical theory of network interference and its applications," *Proc. IEEE*, vol. 97, no. 2, pp. 205–230, Feb. 2009, special issue on *Ultra-Wide Bandwidth (UWB) Technology & Emerging Applications.*

[62] H. ElSawy, A. Sultan-Salem, M. Alouini, and M. Z. Win, "Modeling and analysis of cellular networks using stochastic geometry: A tutorial," *IEEE Commun. Surveys Tuts.*, vol. 19, no. 1, pp. 167–203, 1st Quart. 2017.

[63] T. Hou, Y. Liu, Z. Song, X. Sun, and Y. Chen, "Multiple antenna aided NOMA in UAV networks: A stochastic geometry approach," *IEEE Trans. Commun.*, vol. 67, no. 2, pp. 1031–1044, Feb. 2019.

[64] Z. Tong, H. Lu, M. Haenggi, and C. Poellabauer, "A stochastic geometry approach to the modeling of DSRC for vehicular safety communication," *IEEE Trans. Intell. Transp. Syst.*, vol. 17, no. 5, pp. 1448–1458, May 2016.

[65] M. J. Farooq, H. ElSawy, and M.-S. Alouini, "A stochastic geometry model for multi-hop highway vehicular communication," *IEEE Trans. Wireless Commun.*, vol. 15, no. 3, pp. 2276–2291, Mar. 2016.

[66] H. ElSawy, E. Hossain, and M. Haenggi, "Stochastic geometry for modeling, analysis, and design of multi-tier and cognitive cellular wireless networks: A survey," *IEEE Commun. Surv. Tuts.*, vol. 15, no. 3, pp. 996–1019, Mar. 2013.

[67] M. Mozaffari, W. Saad, M. Bennis, Y.-H. Nam, and M. Debbah, "A tutorial on UAVs for wireless networks: Applications, challenges, and open problems," *IEEE Commun. Surv. Tuts.*, vol. 21, no. 3, pp. 2334–2360, Mar. 2019.

[68] R. W. Heath, M. Kountouris, and T. Bai, "Modeling heterogeneous network interference using Poisson point processes," *IEEE Trans. Signal Process.*, vol. 61, no. 16, pp. 4114–4126, Aug. 2013.

[69] Y. Shen, W. Dai, and M. Z. Win, "Power optimization for network localization," *IEEE/ACM Trans. Netw.*, vol. 22, no. 4, pp. 1337–1350, Aug. 2014.

[70] J.-F. Collet, *Discrete stochastic processes and applications.* Cham, Switzerland: Springer, 2018.

[71] F. Spitzer, *Principles of random walk*, 2nd ed. New York: Springer-Verlag, 1976, vol. 34.

[72] S. Rai, "The spectrum of a random geometric graph is concentrated," *J. Theoretical Prob.*, vol. 20, no. 2, pp. 119–132, 2007.

[73] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. John Wiley & Sons, 2006.

[74] J. A. Nelder and R. Mead, "A simplex method for function minimization," *J. Comput.*, vol. 7, no. 4, pp. 308–313, Jan. 1965.

[75] S. Kurt and B. Tavli, "Path-loss modeling for wireless sensor networks: A review of models and comparative evaluations." *IEEE Antennas Propag. Mag.*, vol. 59, no. 1, pp. 18–37, 1st Quart. 2017.

[76] N. Patwari, J. Ash, S. Kyperountas, A. Hero, R. Moses, and N. Correal, "Locating the nodes: Cooperative localization in wireless sensor networks," *IEEE Signal Process. Mag.*, vol. 22, no. 4, pp. 54–69, Jun. 2005.

[77] Q. Su and Y.-C. Wu, "On convergence conditions of Gaussian belief propagation," *IEEE Trans. Signal Process.*, vol. 63, no. 5, pp. 1144–1155, Mar. 2015.

[78] J. N. Ash and R. L. Moses, "On the relative and absolute positioning errors in self-localization systems," *IEEE Trans. Signal Process.*, vol. 56, no. 11, pp. 5668–5679, Nov. 2008.

[79] F. Spitzer, "Hitting probabilities," *J. Math. and Mech.*, vol. 11, no. 4, pp. 593–614, Jul. 1962.

[80] F. R. Chung, *Spectral graph theory.* American Mathematical Soc., 1997, no. 92.

[81] S. Boyd and L. Vandenberghe, *Convex Optimization*, 1st ed. Cambridge University Press, 2004.

[82] P. W. Shor and J. E. Yukich, "Minimax grid matching and empirical measures," *The Annals of Probability*, pp. 1338–1348, 1991.

Yifeng Xiong (S'15) received his B.S. degree in information engineering, and the M.S. degree in information and communication engineering from Beijing Institute of Technology (BIT), Beijing, China, in 2015 and 2018, respectively. He is currently pursuing the PhD degree with Next-Generation Wireless within the School of Electronics and Computer Science, University of Southampton. His research interests include quantum computation, quantum information theory, graph signal processing, and statistical inference over networks.

Nan Wu (M'11) received his B.S., M.S. and Ph.D. degrees from Beijing Institute of Technology (BIT), Beijing, China in 2003, 2005 and 2011, respectively. From 2008 to 2009, he was a Visiting Ph.D. Student with the Department of Electrical Engineering, Pennsylvania State University, USA. He is currently a Professor with the School of Information and Electronics, BIT. His research interests include signal processing in wireless communication networks. He was a recipient of the National Excellent Doctoral Dissertation Award by MOE of China in 2013. He serves as an Editorial Board Member of the IEEE WIRELESS COMMUNICATIONS LETTERS, the IEEE ACCESS, the International Journal of Electronics and Communications, and the KSII Transactions on Internet and Information Systems.

Yuan Shen (Senior Member, IEEE) received the B.E. degree (Hons.) in electronic engineering from Tsinghua University, Beijing, China, in 2005, and the S.M. degree in electrical engineering and computer science and the Ph.D. degree from MIT, Cambridge, MA, USA, in 2008 and 2014, respectively.

He is currently an Associate Professor with the Department of Electronic Engineering, Tsinghua University. His research interests include network localization and navigation, inference techniques, resource allocation, and cooperative networks. He has served as the TPC Symposium Co-Chair for the IEEE ICC and Globecom for several times. He was the Elected Chair for the IEEE ComSoc Radio Communications Committee for the period of 2019-2020. He is also an Editor of the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, IEEE WIRELESS COMMUNICATIONS LETTERS, and CHINA COMMUNICATIONS.

Moe Z. Win (S'85–M'87–SM'97–F'04) is a professor at the Massachusetts Institute of Technology (MIT) and the founding director of the Wireless Information and Network Sciences Laboratory. Prior to joining MIT, he was with AT&T Research Laboratories and NASA Jet Propulsion Laboratory.

His research encompasses fundamental theories, algorithm design, and network experimentation for a broad range of real-world problems. His current research topics include ultra-wideband systems, network localization and navigation, network interference exploitation, and quantum information science. He has served the IEEE Communications Society as an elected Member-at-Large on the Board of Governors, as elected Chair of the Radio Communications Committee, and as an IEEE Distinguished Lecturer. Over the last two decades, he held various editorial posts for IEEE journals and organized numerous international conferences. Recently, he has served on the SIAM Diversity Advisory Committee.

Dr. Win is an elected Fellow of the AAAS, the EURASIP, the IEEE, and the IET. He was honored with two IEEE Technical Field Awards: the IEEE Kiyo Tomiyasu Award (2011) and the IEEE Eric E. Sumner Award (2006, jointly with R. A. Scholtz). His publications, co-authored with students and colleagues, have received several awards. Other recognitions include the IEEE Communications Society Edwin H. Armstrong Achievement Award (2016), the Cristoforo Colombo International Prize for Communications (2013), the Copernicus Fellowship (2011) and the *Laurea Honoris Causa* (2008) from the Università degli Studi di Ferrara, and the U.S. Presidential Early Career Award for Scientists and Engineers (2004). He is an ISI Highly Cited Researcher.