Received Signal Strength-Based Wireless Localization via Semidefinite Programming: Noncooperative and Cooperative Schemes

Robin Wentao Ouyang, Albert Kai-Sun Wong, and Chin-Tau Lea, *Senior Member, IEEE*

***Abstract*—The received signal strength (RSS)-based approach to wireless localization offers the advantage of low cost and easy implementability. To circumvent the nonconvexity of the conventional maximum likelihood (ML) estimator, in this paper, we propose convex estimators specifically for the RSS-based localization problems. Both noncooperative and cooperative schemes are considered. We start with the noncooperative RSS-based localization problem and derive a nonconvex estimator that approximates the ML estimator but has no logarithm in the residual. Next, we apply the semidefinite relaxation technique to the derived nonconvex estimator and develop a convex estimator. To further improve the estimation performance, we append the ML estimator to the convex estimator with the result by the convex estimator as the initial point. We then extend these techniques to the cooperative localization problem. The corresponding Cramer–Rao lower bounds (CRLB) are derived as performance benchmarks. Our proposed convex estimators comply well with the RSS measurement model, and simulation results clearly demonstrate their superior performance for RSS-based wireless localization.**

***Index Terms*—Cooperative localization, Cramer–Rao lower bound (CRLB), maximum likelihood (ML), received signal strength (RSS), relaxation, semidefinite programming (SDP), wireless localization.**

I. Introduction

Wireless localization has gained considerable attention over the past decade [1]–[4]. The capability of accurately positioning a mobile station in cellular networks enables many innovative applications, for example, emergency services, friend finding, and tracking of the elderly [5]. In addition, wireless localization is an indispensable component of wireless sensor networks since the readings from a large number of sensor nodes are meaningful only when the geolocations of these readings are known. In this paper, we will refer to the base stations in cellular networks and the anchor nodes in wireless sensor networks with known locations as reference nodes (RNs) and the mobile stations and sensor nodes with unknown locations as blind nodes (BNs).

Most of the current localization techniques [6]–[14] for wireless networks are based on the measurement of one or several (hybrid) physical parameters of the radio signal transmitted between the RNs and BNs. These parameters include time of arrival (TOA), time difference of arrival, angle of arrival, and received signal strength (RSS). There exist inherent tradeoffs between the localization accuracy and the implementation complexity of these techniques, and the RSS-based technique provides a low-cost and easy-implementation solution.

Determining the location of a BN given the measurements of one or several aforementioned parameters can be formulated as an estimation problem. The commonly used estimators fall into the following two main categories: 1) the maximum likelihood (ML) estimator [6]–[10] and 2) the linearized least squares (LLS) estimator [11]–[14]. When the statistics of the measurement error are known, the ML estimator is asymptotically optimal. However, due to the nature of the localization problem itself, the formed ML estimator has no closed-form solution, and thus, an iterative solver is required. In addition, the formed ML estimator is nonconvex, and thus, its performance highly depends on the initial point provided for the iterative solver (a local optimization method). A poor initialization often leads to a very bad estimation. Moreover, also due to the nonconvexity, searching for the global minimum of the ML estimator is very difficult.

To overcome this problem, several researchers have proposed the LLS estimator. The idea of the LLS estimator is to reorganize and approximate the original nonlinear equations into a set of linear equations with respect to the BN's location without Taylor-series expansion [15] and then apply the LS technique. Since these equations are linear in the BN's location, a closed-form solution can be obtained. Though the LLS estimator is much easier and has an explicit solution, its accuracy is not as good as the ML estimator, particularly when the variance of the measurement noise is high or the geometric condition is not good. Several other proposed methods can be found in [16]–[18].

Recently, the use of the semidefinite relaxation technique for the wireless localization problem is studied in [19] and [20]. The basic idea is to relax the original nonconvex problem via semidefinite programming (SDP) [21] to produce a convex problem. By taking advantage of the convex optimization technique, the global minimum of a convex problem can be quickly and efficiently found. Biswas *et al.* [19], [20] state that the estimators formed via the semidefinite relaxation technique are highly satisfactory compared with other techniques. However, the SDP estimators in [19] and [20] are proposed for general wireless-localization problems given pairwise distance information. Directly applying these general SDP estimators to the

Manuscript received June 15, 2009; revised October 9, 2009. First published January 12, 2010; current version published March 19, 2010. The work of C.-T. Lea was supported by a China 863 Grant 65306 under Contract 2008AA01A324. The review of this paper was coordinated by Dr. Y. Gao.

The authors are with the Department of Electronic and Computer Engineering, Hong Kong University of Science and Technology, Kowloon, Hong Kong (e-mail: oywtece@ust.hk; eealbert@ust.hk; eelea@ece.ust.hk).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TVT.2010.2040096

RSS-based localization problem may result in large errors due to the extra errors introduced when first estimating pairwise distances from the RSS measurements.

Motivated by the above, in this paper, we design SDP estimators specifically for the RSS-based localization problems. We start with the noncooperative localization problem and derive a nonconvex estimator that approximates the ML estimator but has no logarithm in the residual. Then, we apply the semidefinite relaxation technique to the derived nonconvex estimator to form a convex estimator. To further improve the estimation performance, we append the ML estimator to the convex estimator with the result by the convex estimator as the initial point. These techniques are then extended to the cooperative localization problem. The corresponding Cramer–Rao lower bounds (CRLBs) are derived as performance benchmarks.

Simulation results show that without appending the corresponding ML estimators, our proposed SDP estimators outperform all the other estimators that we compared. Their solutions also serve as better initial points for the corresponding ML estimators to further refine the estimation for RSS-based localization. Further, our proposed SDP estimators perform closely to the corresponding CRLBs, and the performance improvement gained from appending the corresponding ML estimators is minor. That is to say, our proposed SDP estimators can already generate satisfactory results without ML refinement. Although this paper focuses on the 2-D scenario, all the estimators involved can be easily extended to the 3-D scenario.

The remainder of this paper is organized as follows. Section II briefly introduces the RSS measurement model and details the development of the proposed convex estimator for the noncooperative-localization problem. Section III extends these estimators to the cooperative-localization problem. The corresponding CRLBs are derived in Section IV as benchmarks for performance comparison. Simulation results for noncooperative- and cooperative-localization problems are presented in Sections V and VI, respectively. Finally, Section VII draws the conclusions.

II. Received Signal Strength Measurement Model and Convex Estimator Development

A. RSS Measurement Model

For future convenience, we start by introducing the notations used throughout this paper. $\mathbb{R}$, $\mathbb{R}^n$, $\mathbb{R}^{n\times m}$, and $\mathbb{S}^n$ denote the set of real numbers, n-vectors, n by m matrices, and symmetric n by n matrices, respectively. $\mathrm{tr}(A)$ represents the trace of matrix A. $(\cdot)^T$ is the transpose operator. $\mathbb{E}(\cdot)$ is the expectation operator. The identity n by n matrix is noted as I_n. $\|u\|$ denotes the Euclidean norm of vector u. For two symmetric matrices A and B, $A \succeq B$ means that $A - B$ is positive semidefinite. We use $[u]_i$ and $[A]_{i,j}$ to denote the ith element of vector u and the element at the ith row jth column of matrix A, respectively. $[A]_{i:j,k:l}$ denotes the submatrix formed by the intersection of the ith to the jth rows and the kth to the lth columns of matrix A.

We first study the noncooperative-localization problem, in which only one BN is to be localized. We denote the unknown coordinates of the BN as $\theta = [x, y]^T$ $(\theta \in \mathbb{R}^2)$ and the known coordinates of the ith RN as $\phi_i = [a_i, b_i]^T$ $(\phi_i \in \mathbb{R}^2)$, with $i = 1, 2, \ldots, N$ (where N is the total number of RNs that the BN can hear).

The RSS (from the BN and received by the ith RN or *vice versa*), which is denoted as P_i, can be related to the distance between the BN and the ith RN through the path loss model for wireless transmission (in decibels) as [22]

$$L_i = L_0 + 10\gamma \log_{10} \frac{\|\theta - \phi_i\|}{d_0} + m_i, \qquad i = 1, 2, \ldots, N \tag{1}$$

where $L_i = P_T - P_i$ is the path loss (P_T is the transmission power), L_0 denotes the path loss value at the reference distance d_0 $(\|\theta - \phi_i\| \geq d_0)$, γ is the path loss exponent, and m_i is a Gaussian random variable representing the log-normal shadow fading effect in multipath environments. The parametric RSS-based wireless localization problem estimates θ based on (1), given measurements and environmental parameters.

B. Nonconvex Estimator Without Logarithm in the Residual

In (1), m_i's are often modeled as independent and identically distributed (i.i.d.) Gaussian random variables with zero mean and standard deviation σ. The joint conditional pdf of the observation vector $L = [L_1, \ldots, L_N]^T$ $(L \in \mathbb{R}^N)$, given θ, is

$$p(L|\theta) = \prod_{i=1}^{N} \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{ -\frac{\left(L_i - L_0 - 10\gamma \log_{10} \frac{\|\theta - \phi_i\|}{d_0}\right)^2}{2\sigma^2} \right\}. \tag{2}$$

The corresponding ML estimator is therefore

$$\hat{\theta} = \arg\min_{\theta} \sum_{i=1}^{N} \left(10\gamma \log_{10} \frac{\|\theta - \phi_i\|}{d_0} - (L_i - L_0) \right)^2. \tag{3}$$

Define $r_i = 10\gamma \log_{10}(\|\theta - \phi_i\|/d_0) - (L_i - L_0)$ in (3) as the residual, which is the difference between the true value and the measurement. Then, (3) can be considered to minimize a penalty function on the residual vector $r = [r_1, \ldots, r_N]^T$ $(r \in \mathbb{R}^N)$ as

$$\hat{\theta} = \arg\min_{\theta} f(r) \tag{4}$$

where $f(\cdot) = \|\cdot\|^2$.

It is clear that the ML estimator (3) is nonconvex since its domain $\{\theta | \theta \neq \phi_i\}$ is not continuous. In addition, (3) contains $\log_{10}\|\theta - \phi_i\|$, which is neither convex nor concave in the objective function, although we restrict θ in a convex domain (this can be verified by examining the Hessian of $\log_{10}\|\theta - \phi_i\|$, which is neither positive semidefinite nor negative semidefinite). Due to the nonconvexity of (3), finding and confirming the global minimum solution is difficult. A convex estimator would be highly desirable, but the existence of $\log_{10}\|\theta - \phi_i\|$ in r_i encumbers such an idea.

Hence, to facilitate the design of a convex estimator, we first derive a nonconvex estimator that approximates the ML estimator (3) but has no logarithm in the residual. To this end, we replace $f(\cdot) = \|\cdot\|^2$ in (4) by another penalty function

$f(\cdot) = \|\cdot\|_\infty = \max_i |[\cdot]_i|$, i.e., the Chebyshev norm, which is also known as the ℓ_∞ norm [21]. Then, the ML estimator (3) is approximated as

$$\hat{\theta} = \arg\min_\theta \max_i \left| 10\gamma \log_{10} \frac{\|\theta - \phi_i\|}{d_0} - (L_i - L_0) \right|. \quad (5)$$

The advantage of such an approximation is that after certain manipulations [21], we can form an equivalent problem without alogarithm, as shown in the following.

Since $\|\theta - \phi_i\| > 0$ and positive scaling of the objective function will not influence the minimizer, (5) is equivalent to

$$\hat{\theta} = \arg\min_\theta \max_i \left| \log_{10} \frac{\|\theta - \phi_i\|^2}{\beta_i^2} \right| \quad (6)$$

where

$$\beta_i^2 = d_0^2 10^{\frac{L_i - L_0}{5\gamma}}. \quad (7)$$

By noting that

$$\begin{aligned} &\left| \log_{10} \frac{\|\theta - \phi_i\|^2}{\beta_i^2} \right| \\ &= \max \left(\log_{10} \frac{\|\theta - \phi_i\|^2}{\beta_i^2}, \log_{10} \frac{\beta_i^2}{\|\theta - \phi_i\|^2} \right) \\ &= \log_{10} \left(\max \left(\frac{\|\theta - \phi_i\|^2}{\beta_i^2}, \frac{\beta_i^2}{\|\theta - \phi_i\|^2} \right) \right) \end{aligned} \quad (8)$$

we can rewrite (6) as

$$\hat{\theta} = \arg\min_\theta \max_i \log_{10} \left(\max \left(\frac{\|\theta - \phi_i\|^2}{\beta_i^2}, \frac{\beta_i^2}{\|\theta - \phi_i\|^2} \right) \right). \quad (9)$$

Since $\log_{10}(x)$ is a strictly monotonically increasing function in its domain $(0, +\infty)$ (there is a one-to-one mapping between $\log_{10}(x)$ and x, and when $\log_{10}(x)$ is maximized, x is also maximized), (9) is equivalent to

$$\hat{\theta} = \arg\min_\theta \max_i \max \left(\frac{\|\theta - \phi_i\|^2}{\beta_i^2}, \frac{\beta_i^2}{\|\theta - \phi_i\|^2} \right). \quad (10)$$

Following is our proposed estimator for the RSS-based noncooperative localization problem, which can simply be written as

$$\hat{\theta} = \arg\min_\theta f(\tilde{r}). \quad (11)$$

where $\tilde{r} = [\tilde{r}_1, \ldots, \tilde{r}_N]^T$ ($\tilde{r} \in \mathbb{R}^N$) is the proposed residual vector with

$$\tilde{r}_i = \max \left(\frac{\|\theta - \phi_i\|^2}{\beta_i^2}, \frac{\beta_i^2}{\|\theta - \phi_i\|^2} \right) \quad (12)$$

and $f(\cdot)$ is an arbitrary convex penalty function (e.g., the ℓ_p norm family [21]). Obviously, $\tilde{r}_i$ contains no logarithm.

The rationality of the proposed estimator (11) lies at least in the following aspects.

1) A close inspection of (1) reveals that it can be expressed as

$$\beta_i^2 = 10^{\frac{m_i}{5}} \|\theta - \phi_i\|^2 \quad (13)$$

where β_i^2 is given by (7). That is to say, the noise in β_i^2 (a function of the measurement L_i) is multiplicative to $\|\theta - \phi_i\|^2$. The residual $\tilde{r}_i$ complies with such a multiplicative noise model since it consists of only the ratio $\|\theta - \phi_i\|^2 / \beta_i^2$ and its inverse.

2) With $f(\cdot)$ in (11) being the Chebyshev norm, (11) is just (10). As has been shown, (10) is equivalent to (5), which is an approximation of the asymptotically optimal ML estimator (3).

Although, compared with the ML estimator (3), (11) has no logarithm in the residual, it is still not convex due to the nonconvexity of the term $\beta_i^2 / \|\theta - \phi_i\|^2$.

C. Convex Estimator Development via Semidefinite Relaxation

In the following, we will develop a convex estimator based on the derived nonconvex estimator (11).

By introducing an auxiliary variable $t = [t_1, \ldots, t_N]^T$ ($t \in \mathbb{R}^N$), (11) can be cast as [21]

$$\begin{aligned} (\hat{\theta}, \hat{t}\,) &= \arg\min_{\theta, t} f(t) \\ \text{s.t.} \ \frac{\|\theta - \phi_i\|^2}{\beta_i^2} &\le t_i, \quad \frac{\beta_i^2}{\|\theta - \phi_i\|^2} \le t_i, \quad i = 1, \ldots, N \end{aligned} \quad (14)$$

where s.t. is short for subject to. Obviously, in the above formulation, $t_i > 0$.

We then rewrite (14) as

$$\begin{aligned} (\hat{\theta}, \hat{t}\,) &= \arg\min_{\theta, t} f(t) \\ \text{s.t.} \ &\|\theta - \phi_i\|^2 \le \beta_i^2 t_i, \\ &\|\theta - \phi_i\|^2 \ge \beta_i^2 t_i^{-1}, \quad i = 1, \ldots, N. \end{aligned} \quad (15)$$

Note that (15) is actually the same as (14) since the constraints in (15) already imply $\|\theta - \phi_i\|^2 \neq 0$ and $t_i > 0$. Therefore, we can derive (14) from (15).

For simplicity, we define $k_i = \|\phi_i\|^2$. By noting that $\|\theta - \phi_i\|^2 = \theta^T\theta - 2\phi_i^T\theta + k_i$, we can rewrite (15) with an auxiliary variable X ($X \in \mathbb{S}^2$) as

$$\begin{aligned} (\hat{\theta}, \hat{X}, \hat{t}\,) &= \arg\min_{\theta, X, t} f(t) \\ \text{s.t.} \ &\text{tr}(X) - 2\phi_i^T\theta + k_i \le \beta_i^2 t_i, \\ &\text{tr}(X) - 2\phi_i^T\theta + k_i \ge \beta_i^2 t_i^{-1}, \quad i = 1, \ldots, N \\ &X = \theta\theta^T. \end{aligned} \quad (16)$$

In the above formulation, $\text{tr}(X) - 2\phi_i^T\theta + k_i \le \beta_i^2 t_i$ are affine constraints, and $\text{tr}(X) - 2\phi_i^T\theta + k_i \ge \beta_i^2 t_i^{-1}$ are convex constraints since $\text{tr}(X)$ is linear in X, $-2\phi_i^T\theta$ is linear in θ, and t_i^{-1} is convex on $t_i > 0$. However, the equality constraint $X = \theta\theta^T$ is not affine; therefore, the above formulation is still not convex.

To obtain a convex estimator, we relax the equality constraint $X = \theta\theta^T$ to an inequality constraint $X \succeq \theta\theta^T$ (semidefinite relaxation) and then express it as a linear matrix inequality (LMI) [23] by using a Schur complement [21], i.e.,

$$\begin{aligned}(\hat{\theta}, \hat{X}, \hat{t}\,) &= \arg\min_{\theta, X, t} f(t)\\ \text{s.t. } & \operatorname{tr}(X) - 2\phi_i^T\theta + k_i \leq \beta_i^2 t_i\\ & \operatorname{tr}(X) - 2\phi_i^T\theta + k_i \geq \beta_i^2 t_i^{-1}, \qquad i = 1, \ldots, N\\ & \begin{bmatrix} X & \theta \\ \theta^T & 1 \end{bmatrix} \succeq 0. \end{aligned} \tag{17}$$

Now, (17) is convex, and the readily developed numerical tools [24], [25] for solving convex optimization problems can be used. An excellent property of a convex optimization problem is that any local minimum is also the global minimum. Therefore, we can guarantee that the global minimum is achieved when a solution is obtained.

In the above formulation, we can also express $\operatorname{tr}(X) - 2\phi_i^T\theta + k_i \geq \beta_i^2 t_i^{-1}$ as LMIs as follows:

$$\begin{bmatrix} \operatorname{tr}(X) - 2\phi_i^T\theta + k_i & \beta_i \\ \beta_i & t_i \end{bmatrix} \succeq 0. \tag{18}$$

An advantage of using LMIs is that t_i^{-1} is circumvented in the expression, and the constraints become linear.

The only difference between (16) and (17) is that we relax the equality constraint in (16) to an inequality constraint in (17). Therefore, if the solution of (17) satisfies $\hat{X} = \hat{\theta}\hat{\theta}^T$, we conclude that $\hat{\theta}$ given by (17) is also the global minimizer of (16) and, thus, the global minimizer of (11) [since (16) is equivalent to (11)]. If not, $\hat{\theta}$ given by (17) is still feasible for (11), except $\hat{\theta} = \phi_i$, since (11) is unconstrained with domain $\{\theta | \theta \neq \phi_i\}$. In addition, $f(\hat{t}\,)$ given by (17) provides a lower bound on the optimal value of (11) since we solve a relaxed problem on a larger set. In the following, we will refer to (17) as an SDP estimator (although it is not strict) to emphasize the utilization of semidefinite relaxation.

D. *Result Analysis*

Compared with (11), besides giving $\hat{\theta}$, the SDP problem (17) also produces $\hat{X}$, which can provide us additional information.

Due to relaxation, $\hat{\theta}$ given by (17) may not be the global minimizer of (11), unless $\hat{X} = \hat{\theta}\hat{\theta}^T$ is satisfied. If we view the unknown global minimizer of (11) as a random variable [19] and denote it as $\tilde{\theta}$, generally, we have

$$\mathbb{E}(\tilde{\theta}) = \hat{\theta} \quad \mathbb{E}(\tilde{\theta}\tilde{\theta}^T) = \hat{X}. \tag{19}$$

Then, the covariance matrix of $\tilde{\theta}$ is given by

$$\operatorname{cov}(\tilde{\theta}) = \mathbb{E}(\tilde{\theta}\tilde{\theta}^T) - \mathbb{E}(\tilde{\theta})\mathbb{E}(\tilde{\theta})^T = \hat{X} - \hat{\theta}\hat{\theta}^T. \tag{20}$$

Since $\hat{X}$ and $\hat{\theta}$ are the solutions of (17), they must satisfy $[\hat{X} \;\; \hat{\theta}; \hat{\theta}^T \;\; 1] \succeq 0$, which is equivalent to $\hat{X} - \hat{\theta}\hat{\theta}^T \succeq 0$. Therefore, we can observe that $\hat{X} - \hat{\theta}\hat{\theta}^T$ can, indeed, serve as a covariance matrix. Such an interpretation coincides with our previous analysis since if $\hat{X} - \hat{\theta}\hat{\theta}^T = 0$, then $\tilde{\theta}$ should be deterministic, and hence, $\tilde{\theta} = \mathbb{E}(\tilde{\theta}) = \hat{\theta}$, i.e., $\hat{\theta}$ given by (17) is also the global minimizer of (11).

Since (17) provides us the information about the covariance matrix of $\tilde{\theta}$, we can utilize it to trust a $\hat{\theta}$, which results in small element variances in $\operatorname{cov}(\tilde{\theta})$, and discard a $\hat{\theta}$, which causes large variances and requires a new estimation.

Such an interpretation also assists in designing approaches to approximately find the global minimizer of (11).

We can assume a certain distribution of $\tilde{\theta}$ that satisfies (19) and (20), and then randomly generate K realizations of $\tilde{\theta}$ (denoted as $\tilde{\theta}_k$, where $k = 1, 2, \ldots, K$) according to the assumed distribution and record the corresponding objective function values $f(\tilde{r}(\tilde{\theta}_k))$ of (11). The approximate global minimizer of (11) is given by $\tilde{\theta}_k$, which results in the smallest $f(\tilde{r}(\tilde{\theta}_k))$ among all the K objective function values. Such an approach is called randomization [26]. It sounds reasonable; however, through simulations, we find that although it can result in a lower objective function value, it does not necessarily produce a better estimation in the Euclidean distance sense. The smallest objective function value is very likely to be caused by a local minimizer (or a point close to a local minimizer) rather than the global minimizer since the solution space is not discrete, and the objective function is nonconvex. A sufficiently large number of randomizations may be required to produce a satisfactory result.

Another approach is through grid search. Since we know that for a Gaussian random variable x, $\mathcal{P}[|x - \mathbb{E}(x)| \leq 2\sigma_x] = 0.97$, where $\mathcal{P}$ denotes probability, and σ_x is the standard deviation of x. Therefore, assuming $\tilde{\theta}$ is Gaussian, it lies in the following space with very high probability: $\mathbb{E}(\tilde{x}) - 2\sigma_{\tilde{x}} \leq \tilde{x} \leq \mathbb{E}(\tilde{x}) + 2\sigma_{\tilde{x}}$, $\mathbb{E}(\tilde{y}) - 2\sigma_{\tilde{y}} \leq \tilde{y} \leq \mathbb{E}(\tilde{y}) + 2\sigma_{\tilde{y}}$, where $\mathbb{E}(\tilde{x}) = [\hat{\theta}]_1$, $\mathbb{E}(\tilde{y}) = [\hat{\theta}]_2$, $\sigma_{\tilde{x}} = \sqrt{[\operatorname{cov}(\tilde{\theta})]_{1,1}}$, and $\sigma_{\tilde{y}} = \sqrt{[\operatorname{cov}(\tilde{\theta})]_{2,2}}$. We can then search through the space with grid steps Δ_x, Δ_y along the respective axis, record the resultant objective function values, and choose the point given the smallest objective function value as the approximate global minimizer of (11). The smaller the step size is, the better the final estimation becomes.

A more convenient and straightforward way is to provide the solution given by the SDP problem (17) as an initial point for the original problem (11) and run a local optimization method [27] to refine the estimation. However, (11) is nondifferentiable, and its equivalent problem (15) is not very smooth. Alternatively, we provide the solution given by (17) as an initial point for the ML estimator (3) due to its smoothness and asymptotical optimality. In the simulation results, we only present the refined solutions by this method since it is more efficient than randomization and grid search, particularly when the dimension of the problem is high (e.g., cooperative localization, which will be discussed in Section III). However, the refined estimation is then an approximate global minimizer of (3) rather than (11).

III. Extension to Cooperative Localization

Assume that there are N RNs and M BNs in a network. Due to limited communication ranges or other physical limitations, probably none of the BNs can directly connect to at least three RNs, and therefore, none of them can localize themselves by traditional noncooperative localization methods. Cooperative localization provides an excellent solution to this problem. It allows BNs to connect with each other and measure the RSS between themselves. Instead of solving for each BN's location independently, all the BNs' locations will be estimated simultaneously after the necessary RSS measurements are collected. For a more detailed description, see [28].

We now describe the cooperative localization problem mathematically. Denote the coordinates of the ith BN (unknown) as θ_i $(i = 1, 2, \ldots, M, \theta_i \in \mathbb{R}^2)$ and the coordinates of the jth RN (known) as ϕ_j $(j = 1, 2, \ldots, N, \phi_j \in \mathbb{R}^2)$. For ease of expression, we further define $\theta = [\theta_1, \theta_2, \ldots, \theta_M]$ $(\theta \in \mathbb{R}^{2\times M})$. The tuple sets $\{(i,j)\}$, $\{(i,k)\}$ indicating the existence of pairwise BN/RN and BN/BN RSS measurements are denoted as $\mathcal{A}$ and $\mathcal{B}$, respectively $(i, k = 1, 2, \ldots, M, i \neq k, j = 1, 2, \ldots, N)$. Further, define a subset of $\mathcal{B}$ with those $\{(i,k)|i < k\}$ as $\mathcal{C}$.

The RSS measurement model for cooperative localization is

$$L_{ij}^{\mathcal{A}} = L_0 + 10\gamma \log_{10} \frac{\|\theta_i - \phi_j\|}{d_0} + m_{ij}, \qquad (i,j) \in \mathcal{A}$$
$$L_{ik}^{\mathcal{B}} = L_0 + 10\gamma \log_{10} \frac{\|\theta_i - \theta_k\|}{d_0} + n_{ik}, \qquad (i,k) \in \mathcal{B} \quad (21)$$

where m_{ij}, n_{ik} are i.i.d. Gaussian random variables with zero mean and standard deviation σ. Assume that $L_{ki}^{\mathcal{B}} = L_{ik}^{\mathcal{B}}$ for $(i,k) \in \mathcal{B}$.

The ML estimator based on the above measurement model is easily formulated as

$$\hat{\theta} = \arg\min_{\theta} \sum_{i,j:(i,j)\in\mathcal{A}} \left(10\gamma \log_{10} \frac{\|\theta_i - \phi_j\|}{d_0} - \left(L_{ij}^{\mathcal{A}} - L_0\right)\right)^2 + \sum_{i,k:(i,k)\in\mathcal{C}} \left(10\gamma \log_{10} \frac{\|\theta_i - \theta_k\|}{d_0} - \left(L_{ik}^{\mathcal{B}} - L_0\right)\right)^2 \quad (22)$$

which is nonconvex.

Similar to the procedures done for noncooperative localization problem, we can also formulate a convex estimator for cooperative localization by applying semidefinite relaxation to our proposed nonconvex estimator.

First, we extend (15) [the equivalent problem of (11)] for cooperative localization as

$$\begin{aligned}(\hat{\theta}, \hat{t}) &= \arg\min_{\theta,t} f(t)\\ \text{s.t. } &\|\theta_i - \phi_j\|^2 \le \beta_{ij}^2 t_{ij}\\ &\|\theta_i - \phi_j\|^2 \ge \beta_{ij}^2 t_{ij}^{-1} \qquad \forall (i,j) \in \mathcal{A}\\ &\|\theta_i - \theta_k\|^2 \le \zeta_{ik}^2 s_{ik}\\ &\|\theta_i - \theta_k\|^2 \ge \zeta_{ik}^2 s_{ik}^{-1} \qquad \forall (i,k) \in \mathcal{C}\end{aligned} \quad (23)$$

where t is a vector constructed by stacking all the t_{ij}'s and s_{ik}'s, $\beta_{ij}^2 = d_0^2 10^{(L_{ij}^{\mathcal{A}} - L_0)/5\gamma}$ for $(i,j) \in \mathcal{A}$ and $\zeta_{ik}^2 = d_0^2 10^{(L_{ik}^{\mathcal{B}} - L_0)/5\gamma}$ for $(i,k) \in \mathcal{C}$.

It is easy to show that [20]

$$\|\theta_i - \phi_j\|^2 = [e_i; -\phi_j]^T [\theta \;\; I_2]^T [\theta \;\; I_2] [e_i; -\phi_j]$$
$$\|\theta_i - \theta_k\|^2 = (e_i - e_k)^T \theta^T \theta (e_i - e_k)$$

where e_i is an M-vector with 1 at the ith entry and 0 elsewhere (semicolon means line break).

Further, define $X = \theta^T\theta$, $g_{ij} = [e_i; -\phi_j]$ for $(i,j) \in \mathcal{A}$ and $h_{ik} = [e_i - e_k; 0_2]$ for $(i,k) \in \mathcal{C}$ $(g_{ij}, h_{ik} \in \mathbb{R}^{M+2})$, and then, (23) can be expressed as

$$\begin{aligned}(\hat{\theta}, \hat{X}, \hat{t}) &= \arg\min_{\theta,X,t} f(t)\\ \text{s.t. } &g_{ij}^T [X \;\; \theta^T; \theta \;\; I_2] g_{ij} \le \beta_{ij}^2 t_{ij}\\ &g_{ij}^T [X \;\; \theta^T; \theta \;\; I_2] g_{ij} \ge \beta_{ij}^2 t_{ij}^{-1} \quad \forall (i,j) \in \mathcal{A}\\ &h_{ik}^T [X \;\; \theta^T; \theta \;\; I_2] h_{ik} \le \zeta_{ik}^2 s_{ik}\\ &h_{ik}^T [X \;\; \theta^T; \theta \;\; I_2] h_{ik} \ge \zeta_{ik}^2 s_{ik}^{-1} \quad \forall (i,k) \in \mathcal{C}\\ &X = \theta^T\theta.\end{aligned} \quad (24)$$

We apply the semidefinite relaxation technique again to relax the constraint $X = \theta^T\theta$ in (24) to $X \succeq \theta^T\theta$, which is equivalent to $[X \;\; \theta^T; \theta \;\; I_2] \succeq 0$.

For simplicity, we define $Z = [X \; \theta^T; \theta \; I_2]$. Now, the relaxed problem can be written as

$$\begin{aligned}(\hat{Z}, \hat{t}) &= \arg\min_{Z,t} f(t)\\ \text{s.t. } &g_{ij}^T Z g_{ij} \le \beta_{ij}^2 t_{ij}, \quad g_{ij}^T Z g_{ij} \ge \beta_{ij}^2 t_{ij}^{-1} \quad \forall (i,j) \in \mathcal{A}\\ &h_{ik}^T Z h_{ik} \le \zeta_{ik}^2 s_{ik}, \quad h_{ik}^T Z h_{ik} \ge \zeta_{ik}^2 s_{ik}^{-1} \quad \forall (i,k) \in \mathcal{C}\\ &Z \succeq 0, \quad [Z]_{M+1:M+2, M+1:M+2} = I_2\end{aligned} \quad (25)$$

which is convex.

Similarly, $g_{ij}^T Z g_{ij} \ge \beta_{ij}^2 t_{ij}^{-1}$ and $h_{ik}^T Z h_{ik} \ge \zeta_{ik}^2 s_{ik}^{-1}$ can be expressed as LMIs as follows:

$$\begin{bmatrix} g_{ij}^T Z g_{ij} & \beta_{ij} \\ \beta_{ij} & t_{ij} \end{bmatrix} \succeq 0, \qquad \begin{bmatrix} h_{ik}^T Z h_{ik} & \zeta_{ik} \\ \zeta_{ik} & s_{ik} \end{bmatrix} \succeq 0$$

which render these constraints linear.

The estimation of θ is given by $\hat{\theta} = [\hat{Z}]_{M+1:M+2,1:M}$, and the estimation of X is given by $\hat{X} = [\hat{Z}]_{1:M,1:M}$. Furthermore, the estimation of θ_i is given by the ith column of $\hat{\theta}$. Similarly, after we solve (25), we need to check whether $\hat{X} = \hat{\theta}^T\hat{\theta}$. If it does, then the solution given by (25) is also the global minimizer of (24) and, thus, the global minimizer of (23).

A. Result Analysis

Similar to what has been discussed in Section II-D, we view the global minimizer of (23) as random and denote it as $\tilde{\theta}$. We then have

$$\mathbb{E}(\tilde{\theta}) = \hat{\theta}, \quad \mathbb{E}(\tilde{\theta}^T\tilde{\theta}) = \hat{X}.$$

For each $\tilde{\theta}_i$, we define its ensemble variance $\mathrm{var}(\tilde{\theta}_i)$ as the summation of the individual variance of its elements. Then, we have

$$\begin{aligned}\mathrm{var}(\tilde{\theta}_i) &= \mathrm{var}(\tilde{x}_i) + \mathrm{var}(\tilde{y}_i) = \mathrm{tr}\left(\mathbb{E}\left(\tilde{\theta}_i\tilde{\theta}_i^T\right) - \mathbb{E}(\tilde{\theta}_i)\mathbb{E}(\tilde{\theta}_i)^T\right)\\ &= \mathbb{E}\left(\tilde{\theta}_i^T\tilde{\theta}_i\right) - \mathbb{E}(\tilde{\theta}_i)^T\mathbb{E}(\tilde{\theta}_i) = [\hat{X}]_{i,i} - \hat{\theta}_i^T\hat{\theta}_i.\end{aligned}$$

As can be seen, in cooperative localization, the solution of the SDP problem only provides us the ensemble variance of each $\tilde{\theta}_i$ rather than the covariance matrix. For the refinement of the final solution, $\hat{\theta}$, given by (25), is provided as an initial point for the corresponding ML estimator (22).

IV. Cramer–Rao Lower Bound Analysis

The CRLB (denoted as J^{-1}) [29] sets a lower limit on the covariance matrix of any unbiased estimator (unbiased means $\mathbb{E}(\hat{\theta}) = \theta$). That is

$$\mathbb{E}_z \left[(\hat{\theta} - \theta)(\hat{\theta} - \theta)^T \right] \succeq J^{-1} \tag{26}$$

where J is the Fisher information matrix (FIM) [29], with the element $[J]_{i,j}$ defined by

$$[J]_{i,j} = -\mathbb{E}_z \left[\frac{\partial^2 \ln p(z|\theta)}{\partial[\theta]_i \partial[\theta]_j} \right]. \tag{27}$$

z denotes the observation vector, and $p(z|\theta)$ is the joint conditional pdf of the observations given θ.

A. Noncooperative Localization

In noncooperative localization, each BN is localized individually. For a specific BN, $J \in \mathbb{S}^2$. The joint conditional pdf of the observations is given by (2), and therefore, we have

$$[J]_{1,1} = \alpha^2 \sum_{i=1}^{N} \frac{(x-a_i)^2}{\|\theta - \phi_i\|^4} \qquad [J]_{2,2} = \alpha^2 \sum_{i=1}^{N} \frac{(y-b_i)^2}{\|\theta - \phi_i\|^4}$$
$$[J]_{1,2} = [J]_{2,1} = \alpha^2 \sum_{i=1}^{N} \frac{(x-a_i)(y-b_i)}{\|\theta - \phi_i\|^4} \tag{28}$$

where $\alpha = 10\gamma/(\sigma \ln 10)$.

Define the location estimation error as $e = \|\hat{\theta} - \theta\|$. For any unbiased location estimator, its root mean square error (RMSE) $\sqrt{\mathbb{E}(e^2)}$ is lower bounded by

$$\begin{aligned}\sqrt{\mathbb{E}(e^2)} &= \sqrt{\mathbb{E}\left[(\hat{x}-x)^2\right] + \mathbb{E}\left[(\hat{y}-y)^2\right]} \\ &\geq \sqrt{[J^{-1}]_{1,1} + [J^{-1}]_{2,2}} = \sqrt{\mathrm{tr}(J^{-1})}. \end{aligned} \tag{29}$$

Therefore, we define $\sigma_{\mathrm{NC}} \triangleq \sqrt{\mathrm{tr}(J^{-1})}$ as the CRLB on the RMSE of any unbiased noncooperative location estimator.

Relationship Between σ_{NC} and the Environmental Parameters: From (28), it is clear that J can be written as $J = \alpha^2 G(\theta, \phi_i)$, where $G(\theta, \phi_i) \in \mathbb{S}^2$ ($[G]_{i,j} = [J]_{i,j}/\alpha^2$) is a matrix that depends only on the true locations of the BN and the involved RNs but not α. Therefore

$$\sigma_{\mathrm{NC}} = \alpha^{-1}\sqrt{\mathrm{tr}(G^{-1})} \propto \alpha^{-1} = \frac{\ln 10}{10}\frac{\sigma}{\gamma}. \tag{30}$$

It is interesting to note that $\sigma_{\mathrm{NC}} \propto \sigma/\gamma$. That is, σ_{NC} increases as σ increases while it decreases as γ increases. This is because the RSS measurements vary more around the mean power as σ goes larger (causing more uncertainty when relating a measurement to the distance), but they become more sensitive to distance as γ goes larger (causing less uncertainty when relating a measurement to the distance). More interestingly, with the same geometric layout, it is possible to obtain a smaller σ_{NC} under a non-line-of-sight (NLOS) environment than a line-of-sight (LOS) environment since the parameters corresponding to the NLOS environment may result in a smaller σ/γ. For example, according to the IEEE 802.15.4a channel models [30], the residential NLOS environment brings about a smaller σ/γ than the residential LOS environment and, thus, a smaller σ_{NC}.

Relationship Between σ_{NC} and the Size of the Network: Assume that the coordinates of the BN and the RNs are scaled by κ, i.e., $\theta_{\mathrm{new}} = [\kappa x, \kappa y]^T$, $\phi_{i,\mathrm{new}} = [\kappa a_i, \kappa b_i]^T$, which means that the network is scaled without changing the topology (σ and γ are kept the same).

Denote the corresponding J (σ_{NC}) before and after the scaling as J_{old} ($\sigma_{\mathrm{NC,old}}$) and J_{new} ($\sigma_{\mathrm{NC,new}}$), respectively. From (28), we have $J_{\mathrm{new}} = \kappa^{-2} J_{\mathrm{old}}$. Therefore

$$\sigma_{\mathrm{NC,new}} = \sqrt{\mathrm{tr}\left(J_{\mathrm{new}}^{-1}\right)} = \kappa\sqrt{\mathrm{tr}\left(J_{\mathrm{old}}^{-1}\right)} = \kappa\sigma_{\mathrm{NC,old}}. \tag{31}$$

From the above relationship, we can observe that σ_{NC} increases as the network enlarges while keeping the topology unchanged. Specifically, if all the coordinates of the network are scaled by κ, σ_{NC} will also be scaled by κ.

B. Cooperative Localization

In cooperative localization, the involved BNs are localized simultaneously. To facilitate the derivation of the CRLB, we define $\theta_x = [x_1, x_2, \ldots, x_M]^T$ and $\theta_y = [y_1, y_2, \ldots, y_M]^T$ and construct an ensemble (unknown) coordinate vector as $\theta_{\mathrm{en}} = [\theta_x; \theta_y]$. Then, $x_i = [\theta_{\mathrm{en}}]_i$, $y_i = [\theta_{\mathrm{en}}]_{M+i}$ $(i = 1, 2, \ldots, M)$. The expression of the corresponding FIM J (for θ_{en}) is given in the Appendix.

Define the location estimation error for each BN as $e_i = \|\hat{\theta}_i - \theta_i\|$ and further define the normalized RMSE (NRMSE) for the whole network (involving all the located BNs) as $\sqrt{(1/M)\sum_{i=1}^{M}\mathbb{E}(e_i^2)}$. Then, for any unbiased cooperative location estimator, we have the NRMSE lower bounded by

$$\begin{aligned}\sqrt{\frac{1}{M}\sum_{i=1}^{M}\mathbb{E}\left(e_i^2\right)} &= \sqrt{\frac{1}{M}\sum_{i=1}^{M}\left(\mathbb{E}\left[(\hat{x_i}-x_i)^2\right] + \mathbb{E}\left[(\hat{y_i}-y_i)^2\right]\right)} \\ &\geq \sqrt{\frac{1}{M}\sum_{i=1}^{M}\left([J^{-1}]_{i,i} + [J^{-1}]_{M+i,M+i}\right)} \\ &= \sqrt{\frac{1}{M}\mathrm{tr}(J^{-1})}.\end{aligned}$$

Therefore, we define $\sigma_{\mathrm{CO}} \triangleq \sqrt{(1/M)\mathrm{tr}(J^{-1})}$ as the CRLB on the NRMSE of any unbiased cooperative location estimator. Obviously, the two properties (30) and (31) of σ_{NC} also hold for σ_{CO}.

V. Simulation Results for Noncooperative Localization

This section presents the simulation results for the noncooperative localization problem. In (1), we set $d_0 = 1$ m, $L_0 = 40$ dB, and $\gamma = 3$. We consider the scenario that there are N RNs evenly located on a circle centered at (0, 0) with radius $rad = 20$ (the unit used here is meter and the same below). The location of the ith RN $(i = 1, \ldots, N)$ is given by

$$a_i = rad \cos\frac{2\pi(i-1)}{N}, \qquad b_i = rad \sin\frac{2\pi(i-1)}{N}.$$

In addition to our proposed SDP estimator (17) (we term it $SDP_{\rm RSS}$), the following estimators are chosen for comparison.

One is the general SDP estimator proposed in [20] (we term it SDP) for wireless-localization problems given pairwise distance information, regardless of how such information is obtained (e.g., through TOA or RSS). The general SDP estimator is to relax the following original problem:

$$\hat{\theta} = \arg\min_{\theta} \breve{f}(\breve{r}) \tag{32}$$

where $\breve{r} = [\breve{r}_1, \ldots, \breve{r}_N]^T$ ($\breve{r} \in \mathbb{R}^N$) is the residual vector with

$$\breve{r}_i = \|\theta - \phi_i\|^2 - \hat{d}_i^2 \tag{33}$$

and $\breve{f}(\cdot)$ being a penalty function (the penalty functions considered in [20] are the ℓ_1 norm ($\sum_{i=1}^N |[\cdot]_i|$) and the ℓ_2 norm ($\sqrt{\sum_{i=1}^N [\cdot]_i^2}$)). $\hat{d}_i$ in $\breve{r}_i$ is the measured or estimated pairwise distance between the BN and the ith RN. When the distance is estimated through the RSS measurement by the corresponding ML estimator, $\hat{d}_i$ is given by

$$\hat{d}_i = \arg\min_{d_i} \left(10\gamma \log_{10} \frac{d_i}{d_0} - (L_i - L_0)\right)^2 = d_0 10^{\frac{L_i - L_0}{10\gamma}}. \tag{34}$$

Note that $\hat{d}_i^{\,2} = \beta_i^2$ for RSS-based localization and, thus, (33) is actually

$$\breve{r}_i = \|\theta - \phi_i\|^2 - \beta_i^2. \tag{35}$$

Obviously, $\breve{r}_i$ treats the noise in β_i^2 as additive to $\|\theta - \phi_i\|^2$. However, according to (13), the noise contained in β_i^2 is multiplicative to $\|\theta - \phi_i\|^2$. Therefore, $\breve{r}_i$ is not well tailored to the RSS measurement model and neither is the general SDP estimator, while our proposed SDP estimator complies with the multiplicative noise model.

After a similar procedure done in Section II-C, the general SDP estimator can be expressed as

$$\begin{aligned} (\hat{\theta}, \hat{X}, \hat{t}\,) &= \arg\min_{\theta, X, t} \breve{f}(t) \\ \text{s.t. } & \operatorname{tr}(X) - 2\phi_i^T\theta + k_i - \beta_i^2 \le t_i \\ & \operatorname{tr}(X) - 2\phi_i^T\theta + k_i - \beta_i^2 \ge -t_i, \qquad i = 1, \ldots, N \\ & \begin{bmatrix} X & \theta \\ \theta^T & 1 \end{bmatrix} \succeq 0. \end{aligned} \tag{36}$$

To avoid the impact of different penalty functions on the performance of the estimators, in the simulations, we set both the penalty function $f(\cdot)$ for our proposed SDP estimator (17) and the penalty function $\breve{f}(\cdot)$ for the general SDP estimator (36) as the ℓ_1 norm, which is the most robust among the three commonly used norms—the ℓ_1, ℓ_2, and ℓ_∞ norms. In addition, since in (17) and (36) $t_i \ge 0$, the objective functions of these two problems become $\sum_{i=1}^N t_i$.

Another estimator compared is the LLS estimator proposed in [12] (we term it LLS) with pairwise distance information given by (34).

In addition, the ML estimator (3) is appended to each of these three estimators with the solution produced by the respective estimator as an initial point. We denote the ML estimator with the initial point provided by estimator A as A-ML. The CRLB on the RMSE for any unbiased noncooperative location estimator ($\sigma_{\rm NC}$) is also presented as a benchmark.

The simulations are done via MATLAB. The SDP estimators are solved by `CVX` [24], and the ML estimators are solved by the MATLAB function `lsqnonlin`, which adopts the Levenberg–Marquardt method [27].

A. Effect of the Geometric Layout

It is well known that the geometric layout of the BN and the RNs has a significant impact on the localization accuracy, which is known as the geometric dilution of precision [31]. To investigate its effect, we keep the RN locations fixed ($N = 3$) and choose three different BN locations that represent three typical geometric BN–RN configurations: The first is (0, 2), which is very close to the centroid of the triangle formed by the RNs; the second is (12, −2), which is close to one of the RNs while being far away from the other two; and the third is (22, 8), which is outside the convex hull formed by the RNs.

The RMSEs of different estimators for different BN locations are shown in Fig. 1 (each result is based on 1000 independent runs). We find that under such simulation scenarios, all the ML estimators converge to the same point; therefore, we present their results by only one curve in each subfigure and denote it as ML. Such a phenomenon indicates that the nonconvergence problem of the ML estimator is not severe when N is small. However, this problem becomes evident as N increases, which will be shown later. Through simulations, we find that all these estimators are more or less biased. Therefore, their performance cannot be well lower bounded by the CRLB. However, the CRLB can serve as a benchmark representing the minimum RMSE that any unbiased location estimator can achieve.

From Fig. 1, it is clear that the geometric configurations of the BN and the RNs significantly influence the localization accuracy since the estimators exhibit a very different behavior under different geometric layouts. When the MS is at (0, 2), the performance among all the estimators do not differ as much when σ is small. As σ becomes large, the performance of LLS degrades very quickly. When the MS is at (12, −2), $SDP_{\rm RSS}$ is slightly better than ML, and their RMSEs are even lower than the CRLB, while the other two estimators perform much worse. When the MS is at (22, 8), LLS shows very poor performance, almost exponentially degrading as σ increases. Even though SDP is much better than LLS, its performance is still poor. Nevertheless, $SDP_{\rm RSS}$ and ML exhibit steady and excellent performance, and their RMSEs are even lower than the CRLB for large σ.

As can be seen, our proposed SDP estimator is very robust to bad geometric layout and can generate even better performance than the ML estimator when $N = 3$.

B. Effect of the Standard Deviation

The effect of the standard deviation σ of the log-normal shadow fading variable m_i on the localization accuracy has also been shown in Fig. 1. We can observe that no matter where the BN is, the RMSE of any estimator shows degradation as

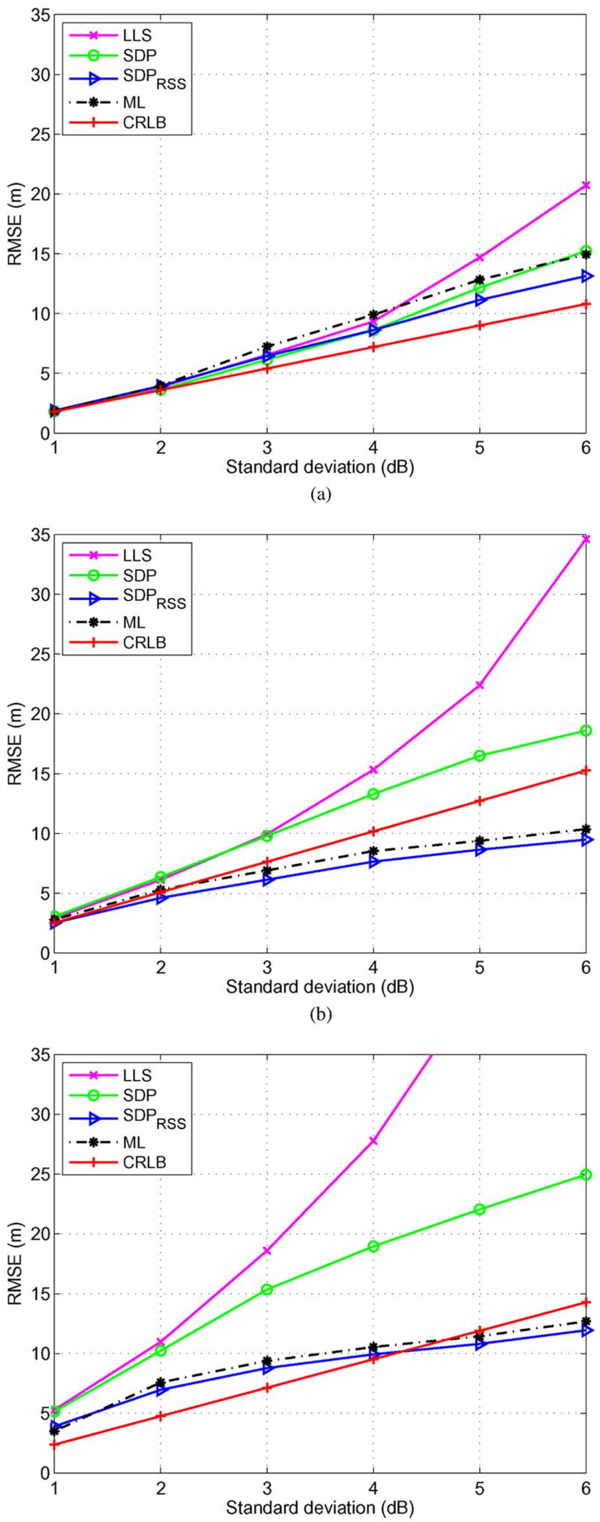


Fig. 1. RMSE versus σ when the BN is at different locations with $N = 3$. (a) BN is at (0, 2). (b) BN is at (12, −2). (c) BN is at (22, 8).

σ increases. Compared with the LLS estimator and the general SDP estimator, the performance degradation of our proposed SDP estimator is much slower and steadier. As has been analyzed, the CRLB on the RMSE for any unbiased noncooperative location estimator σ_{NC} is a linearly increasing function of σ, which coincides with the simulation results.

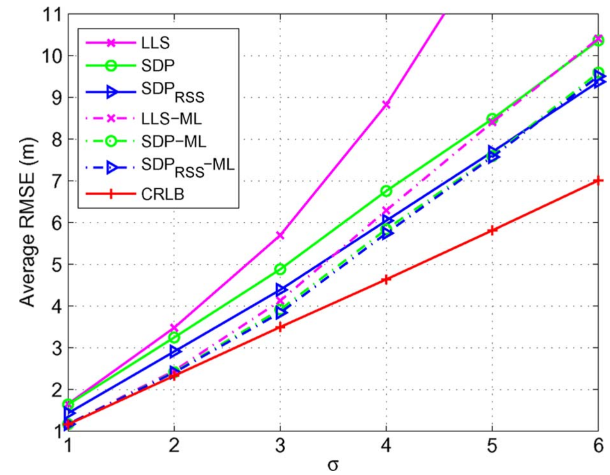


Fig. 2. RMSE versus σ when $N = 6$.

To average out the effect of the geometric layout, we fix the locations of the RNs $(N = 6)$ and uniformly sample 200 random BN locations inside the convex hull formed by these six RNs for each σ. Then, 100 independent localizations are performed to calculate the corresponding RMSE for each sampled location. Finally, the average RMSE is computed by averaging these RMSEs for all the sampled locations. The average RMSEs versus different standard deviations are depicted in Fig. 2. It can be observed that on the average, the performance of SDP_{RSS} is much better than that of SDP. The nonconvergence problem of the ML estimator is exhibited clearly now, as the ML estimators with different initial points produce very different results. When σ is small, the ML estimators can attain the CRLB. However, as σ increases, the ML estimators diverge from the CRLB due to limited data records and large noise variances. When $\sigma = 5$ or 6 dB, SDP_{RSS} exhibits almost the same performance as SDP_{RSS}-ML and SDP-ML. This demonstrates the excellent performance of SDP_{RSS} in harsh environments, where SDP_{RSS} itself already generates good results, and the ML refinement is not necessary.

C. *Effect of the Number of Hearable RNs*

In addition to σ, the number of hearable RNs N also impacts the localization accuracy. In the following simulations, we vary N from 3 to 20 while keeping the BN location inside the square region: $\{(x, y)| -10 \le x \le 10, -10 \le y \le 10\}$. For each N, a similar procedure as has been described in Section V-B is done to calculate the average RMSE.

Fig. 3 shows the average RMSEs versus different numbers of hearable RNs when $\sigma = 4$ dB. It can be observed clearly that without ML estimators, SDP_{RSS} performs best, followed by SDP, and the worst is LLS, no matter how large N is. Although when N is small, the ML estimators does not differ so much, as N increases, SDP_{RSS}-ML outperforms SDP-ML, and SDP-ML performs better than LLS-ML, indicating that, generally, a better initial point leads to a better final result for the nonconvex ML estimator.

In addition, under the simulated scenarios, on the average, SDP_{RSS} is always better than LLS-ML, which illustrates that, generally, LLS cannot serve as a good starting point for the ML

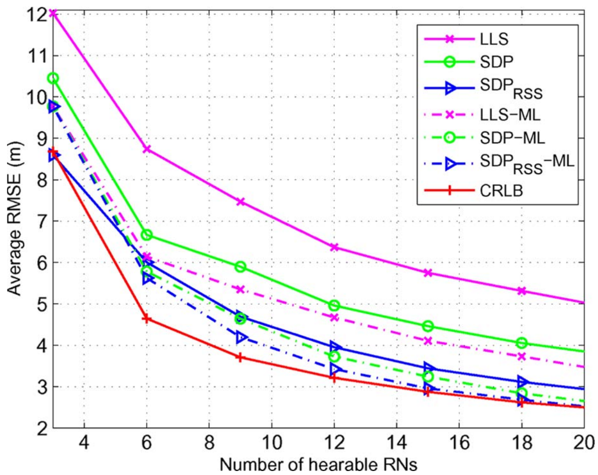


Fig. 3. RMSE versus N when $\sigma = 4$ dB.

estimator. Furthermore, the performance gain by appending the ML estimator to SDP_{RSS} is minor (while the gains for SDP and LLS are much larger), exhibiting the superior performance of our proposed SDP estimator, which can generate excellent results solely without refinement by the ML estimator.

VI. Simulation Results for Cooperative Localization

This section presents the simulation results for the cooperative localization problem. The parameter settings for (21) are the same as those for (1) in Section V. We consider a network with $N = 4$ RNs and $M = 30$ BNs. The four RNs are located at (10, 10), $(-10, 10)$, $(-10, -10)$, and $(10, -10)$, respectively, and the BNs are inside the square region (convex hull) formed by these RNs.

Similar to Section V, our proposed SDP estimator (25) and the general SDP estimator (for cooperative localization) are compared (the penalty functions are set to the ℓ_1 norm). In addition to them, the multidimensional scaling (MDS) estimator [32] (we term it MDS) is chosen for comparison due to its ability to locate several BNs simultaneously. In addition, the ML estimator (22) is appended to each of these compared estimators with the solution produced by the respective estimator as an initial point. The naming for these estimators for cooperative localization is the same as that for noncooperative localization. The CRLB on the NRMSE for any unbiased cooperative location estimator (σ_{CO}) is presented as a benchmark. Without specific indication, each simulation result presented in this section is based on 100 independent runs.

A. Effect of the Network Structure

Through simulations, we find that the structure of the network has a significant influence on the performance of the estimators. Roughly, we classify the network structures as follows: 1) regular and 2) irregular. By regular, we mean that the BNs are connected and spread relatively uniformly throughout the deployment area, and by irregular, we mean that there are obvious large areas without connection and clusters with BNs crowded together. Fig. 4 shows a typical regular network and a typical irregular network with communication range $R = 8$ (we

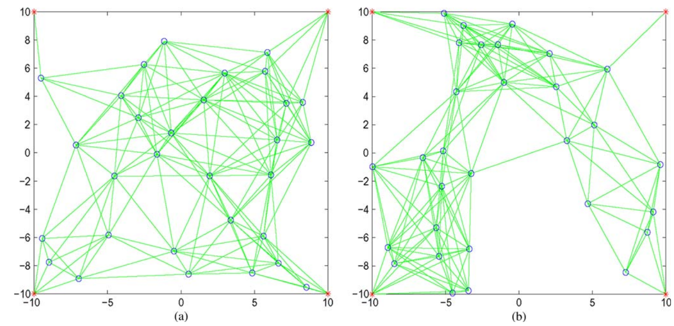


Fig. 4. Network structure. (a) Network I: a typical regular network; BNs are connected and spread relatively uniformly throughout the deployment area. (b) Network II: a typical irregular network with obvious blank areas and clusters.

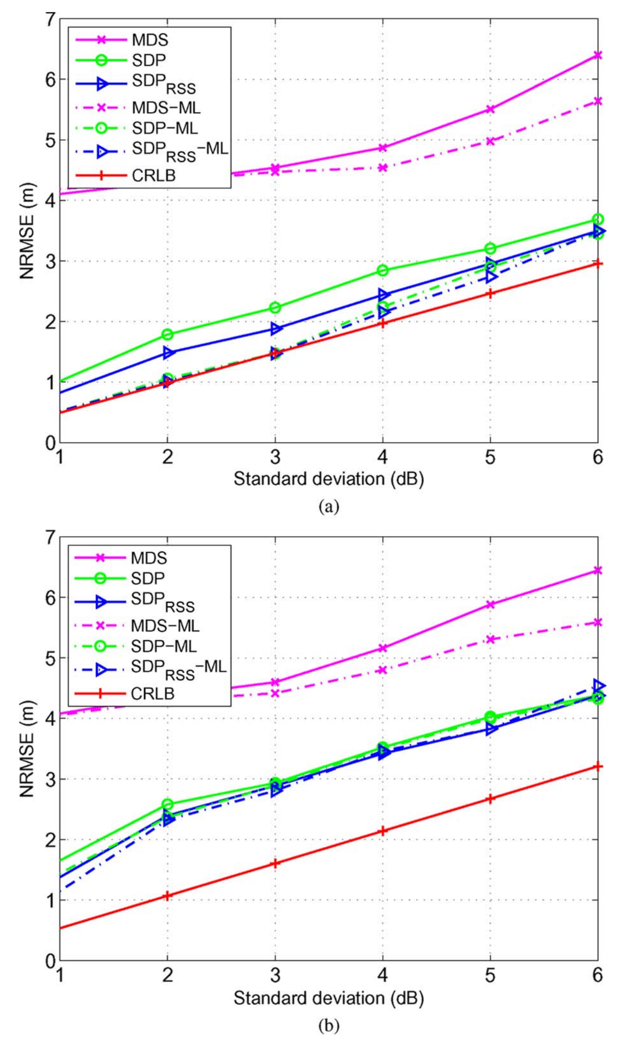


Fig. 5. NRMSE versus σ when $R = 8$ m. (a) Network I. (b) Network II.

assume that the nodes in the network have the same communication range), where the RNs are denoted by red stars, the BNs by blue circles, and the links between the nodes by green lines. A link between two nodes exists if they can communicate with each other.

The NRMSEs of different estimators versus different σ of m_{ij}, n_{ik} in (21) for networks I and II are shown in Fig. 5. As can be observed in Fig. 5(a), for the regular network,

MDS performs very poorly. This is because there are many incomplete data in the required Euclidean distance matrix (due to limited communication range) to recover the configuration of the points. In addition, there are very limited RNs with known locations to accurately determine the required translation, rotation, and scaling. In contrast, the SDP estimators perform much better, and $SDP_{\rm RSS}$ is always better than SDP. As to the ML estimators, $MDS\text{-}ML$ performs poorly since MDS does not perform well. $SDP\text{-}ML$ and $SDP_{\rm RSS}\text{-}ML$ are much better, and they can even attain the CRLB for small σ. As σ becomes large, their performance shows degradation and diverges from the CRLB. When $\sigma = 6$ dB, these two perform the same as $SDP_{\rm RSS}$. This demonstrates again in harsh environments that $SDP_{\rm RSS}$ itself can already generate good results and that the ML refinement almost does not help.

In Fig. 5(b), for the irregular network, MDS and $MDS\text{-}ML$ exhibit similar performance as that for the regular network, while $SDP_{\rm RSS}$ and SDP perform closely, and $SDP_{\rm RSS}$ is slightly better than SDP. All the ML estimators diverge largely from the CRLB in such an irregular network.

B. Effect of the Standard Deviation

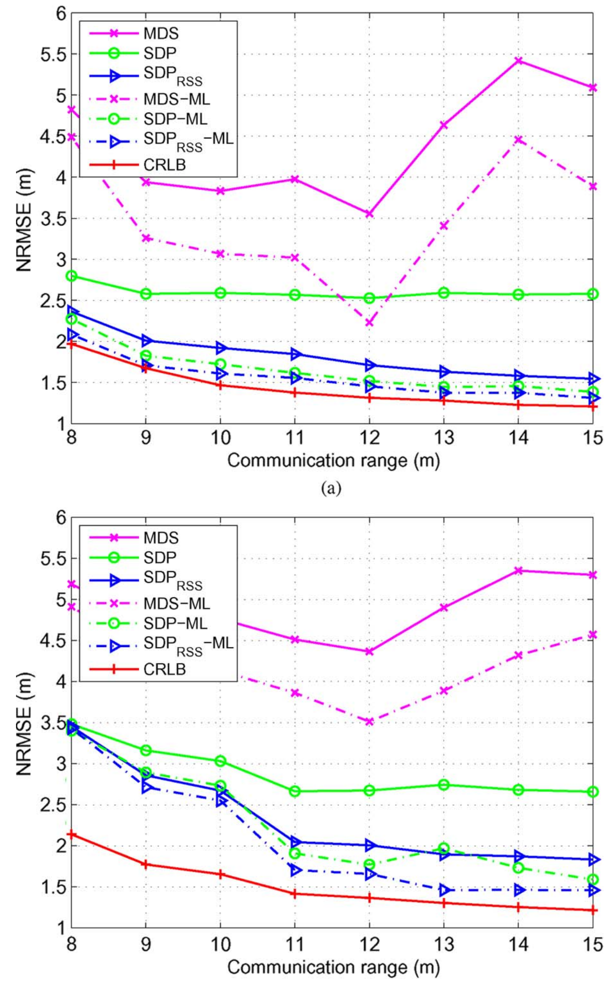
The effect of the standard deviation σ of the log-normal shadow fading variable on the NRMSE has already been shown in Fig. 5. Whether in the regular or the irregular network, all the estimators show performance degradation as the standard deviation σ increases. As has been analyzed, the CRLB on the NRMSE for any unbiased cooperative location estimator $\sigma_{\rm CO}$ is a linearly increasing function of σ, and this coincides with the simulation results.

C. Effect of the Communication Range

We explore the effect of the communication range R on the performance of the estimators in this section. We still use the two networks shown in Fig. 4. The NRMSEs of different estimators versus different R's (σ is set to 4 dB) for networks I and II are shown in Fig. 6.

As can be seen, the CRLB decreases as R increases. This is because when R increases, a node can communicate with more other nodes, thus obtaining more information and facilitating more accurate estimations. For the regular network, excluding the ML estimators, MDS performs worst and is not steady. $SDP_{\rm RSS}$ performs best, and the performance gap between $SDP_{\rm RSS}$ and SDP enlarges as R increases. $SDP_{\rm RSS}$ exhibits steady performance enhancement as R increases, while the NRMSE of SDP is almost unchanged when $R \in [9, 15]$. After the ML estimator refinement, $SDP_{\rm RSS}\text{-}ML$ is the best, and its NRMSE is very close to the CRLB, while $SDP\text{-}ML$ is slightly worse.

For the irregular network, MDS still performs poorly and not steadily. Although when $R = 8$, $SDP_{\rm RSS}$ and SDP perform almost the same, as R increases, the performance gap between them enlarges, and $SDP_{\rm RSS}$ performs better than SDP. In addition, unlike the result for the regular network, $SDP_{\rm RSS}\text{-}ML$ is far above the CRLB for small R for the irregular network. Only when $R \geq 11$ does it perform close to the CRLB. For the irregular network, the performance difference between $SDP_{\rm RSS}\text{-}ML$ and $SDP\text{-}ML$ is more evident than that for the regular network. Sometimes, $SDP\text{-}ML$ is even worse than $SDP_{\rm RSS}$ (e.g., when $R = 10, 13$).



(b)

Fig. 6. NRMSE versus R when $\sigma = 4$ dB. (a) Network I. (b) Network II.

Whether for the regular or the irregular network, as can be seen, appending the ML estimator to refine the result helps a lot for MDS and SDP, while the performance improvement for $SDP_{\rm RSS}$ is minor.

D. Effect of the Number of Hearable RNs

In this section, we investigate the effect of the number of hearable RNs on the performance of the estimators. We set $\sigma = 4$ dB and $R = 8$. First, we use the original four RNs, and then, we increase N to eight and place four additional RNs at (10, 0), (0, 10), (−10, 0), and (0, −10), respectively. The NRMSEs of the estimators versus different N are depicted in Fig. 7.

From the figure, we can observe that the performance of all the estimators improves as N increases from four to eight since more RNs will provide more information and references. When $N = 8$, MDS performs significantly better, and the performance gap between MDS and the SDP estimators becomes much smaller compared with the case when $N = 4$. This is because more RNs will increase the connectivity as well as

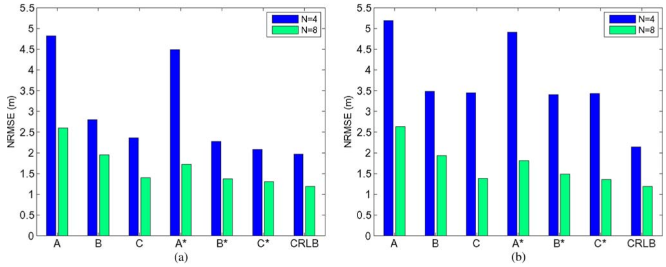


Fig. 7. NRMSE versus N when $\sigma = 4$ dB and $R = 8$ m. (a) Network I. (b) Network II. (A) MDS. (B) SDP. (C) SDP_{RSS}. (A*) MDS-ML. (B*) SDP-ML. (C*) SDP_{RSS}-ML.

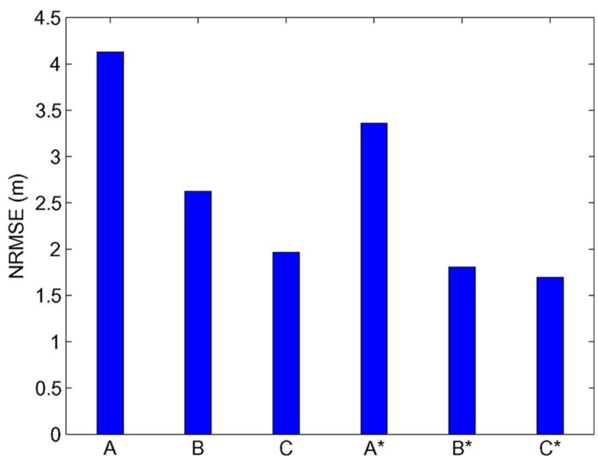


Fig. 8. NRMSEs averaged over 300 randomly generated networks. (A) MDS. (B) SDP. (C) SDP_{RSS}. (A*) MDS-ML. (B*) SDP-ML. (C*) SDP_{RSS}-ML.

facilitate accurately determination of the translation, rotation, and scaling required by MDS.

Another observation is that as more RNs are involved, the impact of the network structure on the performance of the estimators diminishes. We can see that when $N = 8$, the estimators exhibit similar performance for networks I and II. Moreover, SDP_{RSS} performs best, excluding the ML estimators, and SDP_{RSS}-ML is the best among the ML estimators, which is the closest to the CRLB. In addition, the performance improvement from SDP_{RSS} to SDP_{RSS}-ML is minor.

E. Performance Averaged Over Random Network Structures

We present some averaged performance results here, although we consider it to be more appropriate to investigate the performance of the estimators on a network-structure-type basis (e.g., regular or irregular), such as what has been discussed before. We set $N = 4$, $\sigma = 4$ dB, and $R = 10$, generate 300 different networks with $M = 30$ BNs randomly deployed in the convex hull formed by the RNs (the distance between any BNs is kept larger than d_0) and then average over these 300 networks to obtain the NRMSEs. The results are plotted in Fig. 8.

As expected, on the average, SDP_{RSS} is the best, excluding the ML estimators, SDP_{RSS}-ML is the best among all the ML estimators, and the performance enhancement when appending the ML estimator to SDP_{RSS} is very small. That is to say, SDP_{RSS} can generate excellent results solely without refinement by the ML estimator. Moreover, since, for large M, the computational complexity of the ML estimator (22) can exceed that of our proposed SDP estimator (also the general SDP estimator), SDP_{RSS} is the best compromise between localization accuracy and computational efficiency for cooperative localization.

VII. CONCLUSION

To circumvent the nonconvexity of the conventional ML estimator, we have proposed convex SDP estimators specifically for RSS-based wireless localization in this paper. Both noncooperative and cooperative localization problems were investigated. Our proposed SDP estimators complied well with the RSS measurement model. Simulation results showed that without appending ML estimators, our proposed SDP estimators outperformed all the other estimators compared and performed closely to the corresponding CRLBs in RSS-based wireless localization. The solutions of our proposed SDP estimators also served as better initial points for the corresponding ML estimators to further refine the estimations. Moreover, since the performance improvement by appending ML estimators to our proposed SDP estimators is minor while the improvement for other estimators compared is much more significant, our proposed SDP estimators exhibited superior performance for RSS-based wireless localization, which can generate excellent results solely without ML refinement.

APPENDIX
FISHER INFORMATION MATRIX FOR COOPERATIVE LOCALIZATION

It is clear that the corresponding FIM (for θ_{en}) $J \in \mathbb{S}^{2M}$, and it can be partitioned as $J = [J_{xx} \;\; J_{xy}; J_{yx} \;\; J_{yy}]$, where all the block matrices are of dimension M by M, and $J_{yx} = J_{xy}^T$.

Given the measurement model in (21), we have

$$[J_{xx}]_{i,j} = \begin{cases} \alpha^2 \left[\sum_{l:(i,l)\in\mathcal{A}} \frac{(x_i-a_l)^2}{\|\theta_i-\phi_l\|^4} + \sum_{k:(i,k)\in\mathcal{B}} \frac{(x_i-x_k)^2}{\|\theta_i-\theta_k\|^4} \right], & i = j \\ -\alpha^2 \delta_{ij} \frac{(x_i-x_j)^2}{\|\theta_i-\theta_j\|^4}, & i \neq j \end{cases}$$

$$[J_{yy}]_{i,j} = \begin{cases} \alpha^2 \left[\sum_{l:(i,l)\in\mathcal{A}} \frac{(y_i-b_l)^2}{\|\theta_i-\phi_l\|^4} + \sum_{k:(i,k)\in\mathcal{B}} \frac{(y_i-y_k)^2}{\|\theta_i-\theta_k\|^4} \right], & i = j \\ -\alpha^2 \delta_{ij} \frac{(y_i-y_j)^2}{\|\theta_i-\theta_j\|^4}, & i \neq j \end{cases}$$

$$[J_{xy}]_{i,j} = \begin{cases} \alpha^2 \left[\sum_{l:(i,l)\in\mathcal{A}} \frac{(x_i-a_l)(y_i-b_l)}{\|\theta_i-\phi_l\|^4} + \sum_{k:(i,k)\in\mathcal{B}} \frac{(x_i-x_k)(y_i-y_k)}{\|\theta_i-\theta_k\|^4} \right], & i = j \\ -\alpha^2 \delta_{ij} \frac{(x_i-x_j)(y_i-y_j)}{\|\theta_i-\theta_j\|^4}, & i \neq j \end{cases}$$

where $\alpha = 10\gamma/(\sigma \ln 10)$, $\delta_{ij} = 1$ for $(i,j) \in \mathcal{B}$, and $\delta_{ij} = 0$ otherwise.

REFERENCES

[1] J. J. Caffery, Jr. and G. L. Stuber, “Subscriber location in CDMA cellular networks,” *IEEE Trans. Veh. Technol.*, vol. 47, no. 2, pp. 406–416, May 1998.

[2] H. Liu, H. Darabi, P. Banerjee, and J. Liu, “Survey of wireless indoor positioning techniques and systems,” *IEEE Trans. Syst., Man, Cybern. C, Appl. Rev.*, vol. 37, no. 6, pp. 1067–1080, Nov. 2007.

[3] Y. Zhao, “Standardization of mobile phone positioning for 3G systems,” *IEEE Commun. Mag.*, vol. 40, no. 7, pp. 108–116, Jul. 2002.

[4] M. Vossiek, L. Wiebking, P. Gulden, J. Wieghardt, C. Hoffmann, and P. Heide, “Wireless local positioning,” *IEEE Microw. Mag.*, vol. 4, no. 4, pp. 77–86, Dec. 2003.

[5] A.-S. Wong, T. K. Woo, A.-L. Lee, X. Xiao, V.-H. Luk, and K. W. Cheng, “An AGPS-based elderly tracking system,” in *Proc. 1st ICUFN*, Jun. 2009, pp. 100–105.

[6] S. Al-Jazzar and M. Ghogho, “A joint TOA/AOA constrained minimization method for locating wireless devices in non-line-of-sight environment,” in *Proc. IEEE 66th VTC—Fall*, 2007, pp. 496–500.

[7] J. Li, J. Conan, and S. Pierre, “Mobile station location estimation for MIMO communication systems,” in *Proc. 3rd ISWCS*, 2006, pp. 561–564.

[8] S. Al-Jazzar and J. J. Caffery, Jr., “ML and Bayesian TOA location estimators for NLOS environments,” in *Proc. IEEE 56th VTC—Fall*, 2002, vol. 2, pp. 1178–1181.

[9] J. Riba and A. Urruela, “A non-line-of-sight mitigation technique based on ML-detection,” in *Proc. IEEE ICASSP*, 2004, vol. 2, pp. ii-153–ii-156.

[10] N. Patwari, A. O. Hero, III, M. Perkins, N. S. Correal, and R. J. O’Dea, “Relative location estimation in wireless sensor networks,” *IEEE Trans. Signal Process.*, vol. 51, no. 8, pp. 2137–2148, Aug. 2003.

[11] J. J. Caffery, Jr., “A new approach to the geometry of TOA location,” in *Proc. IEEE 52nd VTS—Fall VTC*, 2000, vol. 4, pp. 1943–1949.

[12] K. Yu and Y. J. Guo, “NLOS error mitigation for mobile location estimation in wireless networks,” in *Proc. IEEE 65th VTC—Spring*, 2007, pp. 1071–1075.

[13] R. Juang, D.-B. Lin, and H.-P. Lin, “Hybrid SADOA/TDOA mobile positioning for cellular networks,” *IET Commun.*, vol. 1, no. 2, pp. 282–287, Apr. 2007.

[14] S. Venkatraman and J. J. Caffery, Jr., “Hybrid TOA/AOA techniques for mobile location in non-line-of-sight environments,” in *Proc. IEEE WCNC*, 2004, vol. 1, pp. 274–278.

[15] W. Foy, “Position–location solutions by Taylor-series estimation,” *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-12, no. 2, pp. 187–194, Mar. 1976.

[16] A. O. Hero, III and D. Blatt, “Sensor network source localization via projection onto convex sets (POCS),” in *Proc. IEEE ICASSP*, 2005, vol. 3, pp. iii/689–iii/692.

[17] X. Ji and H. Zha, “Sensor positioning in wireless ad-hoc sensor networks using multidimensional scaling,” in *Proc. 23rd IEEE INFOCOM*, 2004, vol. 4, pp. 2652–2661.

[18] L. Doherty, K. S. J. Pister, and L. El Ghaoui, “Convex position estimation in wireless sensor networks,” in *Proc. 20th IEEE INFOCOM*, 2001, vol. 3, pp. 1655–1663.

[19] P. Biswas and Y. Ye, “Semidefinite programming for ad hoc wireless sensor network localization,” in *Proc. 3rd IPSN*, 2004, pp. 46–54.

[20] P. Biswas, T.-C. Liang, K.-C. Toh, Y. Ye, and T.-C. Wang, “Semidefinite programming approaches for sensor network localization with noisy distance measurements,” *IEEE Trans. Autom. Sci. Eng.*, vol. 3, no. 4, pp. 360–371, Oct. 2006.

[21] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.

[22] K. Pahlavan and A. Levesque, *Wireless Information Networks*. Hoboken, NJ: Wiley, 1995.

[23] S. Boyd, *Linear Matrix Inequalities in System and Control Theory*. Philadelphia, PA: Soc. Ind. Appl. Math., 1994.

[24] M. Grant and S. Boyd, CVX: Matlab Software for Disciplined Convex Programming (Web Page and Software), 2009. [Online]. Available: http://stanford.edu/~boyd/cvx

[25] J. Sturm, “Using SeDuMi 1.02, a Matlab toolbox for optimization over symmetric cones,” *Optim. Methods Softw.*, vol. 11, no. 1, pp. 625–653, 1999.

[26] A. d’Aspremont and S. Boyd, Relaxations and Randomized Methods for Nonconvex QCQPs, 2003.

[27] W. Sun and Y.-X. Yuan, *Optimization Theory and Methods*. New York: Springer Science+Business Media, 2006.

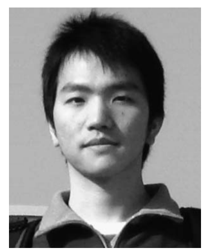
[28] N. Patwari, J. N. Ash, S. Kyperountas, A. O. Hero, III, R. L. Moses, and N. S. Correal, “Locating the nodes: Cooperative localization in wireless sensor networks,” *IEEE Signal Process. Mag.*, vol. 22, no. 4, pp. 54–69, Jul. 2005.

[29] S. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Englewood Cliffs, NJ: Prentice-Hall, 1993.

[30] A. Molisch, K. Balakrishnan, C. Chong, S. Emami, A. Fort, J. Karedal, J. Kunisch, H. Schantz, U. Schuster, and K. Siwiak, IEEE 802.15.4a channel model—Final Rep.

[31] M. A. Spirito, “On the accuracy of cellular mobile station location estimation,” *IEEE Trans. Veh. Technol.*, vol. 50, no. 3, pp. 674–685, May 2001.

[32] X. Li, “Collaborative localization with received-signal strength in wireless sensor networks,” *IEEE Trans. Veh. Technol.*, vol. 56, no. 6, pp. 3807–3817, Nov. 2007.



Robin Wentao Ouyang received the B.E. degree in electronic engineering from the Beijing University of Posts and Telecommunications, Beijing, China, in 2007. He is currently working toward the Ph.D. degree with the Department of Electronic and Computer Engineering, Hong Kong University of Science and Technology, Kowloon, Hong Kong.

His current research interests include wireless localization, tracking, and machine learning.

Albert Kai-Sun Wong received the S.B., S.M., E.E., and Ph.D. degrees in electrical engineering from the Massachusetts Institute of Technology, Cambridge, in 1982, 1984, 1984 and 1988, respectively.

From 1988 to 2000, he was with AT&T and Lucent Technologies/Bell Laboratories, where he held a series of positions as Member of Technical Staff, Distinguished Member of Technical Staff, Technical Manager, Director of Technical Marketing, and Director, Sales and Technical Marketing. From 2000 to 2001, he was the Chief Operating Officer of the Transtech Services Group. From 2002 to 2004, he was a Vice President of the Applied Science and Technology Research Institute of Hong Kong. He also held visiting and adjunct faculty positions at the Chinese University of Hong Kong, Shatin, Hong Kong, Polytechnic University of New York, Brooklyn, and Rutgers University, New Brunswick, NJ. He is currently a Visiting Associate Professor with the Department of Electronic and Computer Engineering, Hong Kong University of Science and Technology, Kowloon, Hong Kong. His current research interests include communication networks, wireless localization and tracking, and photonic and data-switching systems.

Chin-Tau Lea (SM’93) received the B.S and M.S. degrees in electrical engineering from the National Taiwan University, Taipei, Taiwan, in 1976 and 1978, respectively, and the Ph.D. degree in electrical engineering from the University of Washington, Seattle, in 1982.

Since 1996, he has been with the Hong Kong University of Science and Technology (HKUST), Kowloon, Hong Kong, where he is currently a Professor with the Department of Electronic and Computer Engineering. Prior to that, he was with AT&T Bell Laboratories from 1982 to 1985 and with the Georgia Institute of Technology (Georgia Tech), Atlanta, from 1985 to 1995. His current research interests include switching and networking. He is the holder of five U.S. patents.

Dr. Lea is on the Editorial Board of the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS and the Editorial Board of *Computer Networks*. He received the DuPont Young Faculty Award from Georgia Tech in 1987, the IEEE Jack Neubauer Paper Award in 1998, and the School of Engineering Teaching Award from HKUST in 1998.