Designing Dual-Tone Radio Interferometric Positioning Systems

Yiyin Wang, *Member, IEEE*, Xiaoli Ma, *Senior Member, IEEE*, Cailian Chen, *Member, IEEE*, and Xinping Guan, *Senior Member, IEEE*

***Abstract*—For many wireless sensor networks, high accuracy and low complexity localization techniques are crucial to their successful deployment. The radio interferometric positioning system (RIPS) has been introduced for accurate localization with low complexity. In this paper, a dual-tone radio interferometric positioning system (DRIPS) is proposed to localize multiple targets simultaneously. In DRIPS, multiple asynchronous targets emit dual-tone signals, and synchronous anchor receivers (nodes with known positions) extract phase information of low-frequency differential tones created by squaring the received dual-tone signals. Multiple time-of-arrivals (TOAs) coupled with unknown offsets due to the asynchronous targets are estimated based on the phase information. Accordingly, a localization algorithm taking the integer ambiguity issue into account is developed. Moreover, considering the case without accurate knowledge of the frequencies of the differential tones, an ESPRIT-type algorithm is proposed to estimate the frequencies. The proposed DRIPS is robust to flat-fading channels, immune to uncertainties of local oscillators, and able to distinguish different targets. In order to show more insights of the performance limit of the DRIPS, Cramér-Rao bounds (CRBs) are derived. Simulation results illustrate the merits of the proposed DRIPS.**

***Index Terms*—Localization, radio interferometry, ranging, sensor networks, synchronization, time-of-arrival.**

I. Introduction

WIRELESS SENSOR NETWORKs (WSNs) have numerous applications, such as industrial process monitoring, disaster rescue, battlefield surveillance, geographical routing, and so on. One key issue for the successful deployment of WSNs is location awareness [1], [2]. The data collected by sensors should be stamped with their corresponding locations in many applications. The stringent power and cost constraints of WSNs challenge the design of low-complexity localization approaches. The high-accuracy requirement further increases the localization difficulties. In general, localization can be accomplished in two steps i) extract metrics bearing location information, and ii) estimate locations based on the metrics obtained from the first step.

According to the metrics employed by localization systems, they can be roughly categorized as systems using coarse or fine metrics. Typical coarse metrics, such as received signal strength (RSS) [3], [4] and wireless connectivity [5], can be achieved in a cost-efficient way. However, the localization accuracy using coarse metrics may not satisfy high requirements of WSNs [6], [7]. Fine metrics, such as time-of-arrival (TOA) [8] and frequency-difference-of-arrival (FDOA) [9], improve localization accuracy, but may increase system cost. For example, time-based localization using ultra-wideband (UWB) impulse radio (IR) [10], [11] can achieve accurate localization. Nevertheless, the prohibitively high Nyquist sampling rate and dedicated hardware of UWB systems impede their feasibility in some applications. Therefore, with the power and cost constraints, it is of great interest to increase the localization accuracy using narrowband signals.

A low-complexity radio interferometric positioning system (RIPS) is proposed in [12] to achieve high localization accuracy. In RIPS, two asynchronous nodes at different locations emit sinusoids, whose frequencies differ slightly from each other. Two synchronous nodes use the RSS indicator (RSSI) to independently extract low-frequency differential tones, whose phase difference is approximately in a linear relationship with the range information among the four nodes. However, the RIPS faces the integer ambiguity issue, requires scheduling for multiple measurements, and cannot handle fading channels. Several modified RIPS-based methods are proposed in [13], [14] to track mobile targets, where Doppler shifts are further explored. Moreover, spinning anchors transmitting radio signals with fixed frequencies (SpinLoc) are employed in [15] to produce specified Doppler signals, and then angle-of-arrivals (AOAs) are estimated as localization metrics. TripLoc in [16] employs AOA estimates based on the RIPS measurements for localization as well. The integer ambiguity issue is resolved in TripLoc thanks to the special design of the antenna arrays. Software defined radios are used to implement interferometric localization with the assistance of AOAs in [17]. The FPGA implementation of the RIPS is investigated in [18] to provide a

Manuscript received July 04, 2014; revised October 21, 2014; accepted December 03, 2014. Date of publication December 30, 2014; date of current version February 09, 2015. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Subhrakanti Dey. A portion of this work was supported by National Basic Research Program of China under the grant no. 2010CB731803, by the National Nature Science Foundation of China under 61301223, 61221003, 61290322, 61174127, and 61273181, Science and Technology Commission of Shanghai Municipality (STCSM), China under 13QA1401900 and 13Z111050008, the Research Found for the Doctoral Program of Higher Education under 20110073120025 and 13Z102090106, and the Georgia Tech Ultra-wideband Center of Excellence (http://www.uwbtech.gatech.edu/).

Y. Wang, C. Chen, and X. Guan are with the Department of Automation, Shanghai Jiao Tong University, and also with the Key Laboratory of System Control and Information Processing, Ministry of Education of China, Shanghai, 200240, China (e-mail: yiyinwang@sjtu.edu.cn; cailianchen@sjtu.edu.cn; xpguan@sjtu.edu.cn).

X. Ma is with the School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA 30332 USA (e-mail: xiaoli@gatech.edu).

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/TSP.2014.2386295

more robust and flexible platform. A stochastic RIPS (SRIPS) is proposed in [19] to make use of radios at 2.4 GHz. Most of these systems inherit the original signal design of the RIPS. On the other hand, fading channels, tedious measurements and the integer ambiguity still impose great challenges on most RIPS-type systems.

Furthermore, an asynchronous RIPS (ARIPS) is proposed in [20] employing dual-tone signals, where two tones are not close to each other and low-frequency differential tones are created by mixing the dual-tone signals from different transmitters. Thus, its dual-tone signal is totally different from the one proposed in this paper. Further, the ARIPS still suffers from fading channels. Recently, a dual-tone radio interferometric positioning system using undersampling techniques (uDRIPS) is proposed in [21]. Although uDRIPS employs the similar dual-tone signal as the one proposed in this paper, its receiver using the undersampling techniques is distinct from the receiver proposed in this paper. In addition, only a single-target scenario is considered in uDRIPS.

In this paper, we propose a dual-tone radio interferometric positioning system (DRIPS) for multi-target localization under flat-fading channels. In DRIPS, multiple asynchronous targets emit dual-tone signals, which are designed to be well separated from each other. The two tones of the dual-tone signal are set apart less than the channel coherence bandwidth, and thus they will be affected by the same fading effects. The flat-fading channel model is well accommodated in the DRIPS. At the anchor receiver, the dual-tone signals are squared similarly as in the RIPS to create low-frequency differential tones, whose phases include TOAs and offsets due to the asynchronous targets. With the help of synchronous anchor receivers, low-complexity localization algorithms are developed using the biased TOAs (TOAs with offsets) and taking a simplified integer ambiguity problem into account. Furthermore, due to various reasons (e.g. randomness of local oscillators (LOs) and unreliable environments), exact knowledge of the frequencies of the differential tones may not be available. An ESPRIT-type frequency estimation algorithm is proposed to deal with this situation. The Cramér-Rao bounds (CRBs) of the range estimates are further derived to show more insights of the performance limit of the DRIPS.

In summary, our proposed DRIPS has the following advantages: i) multiple targets can be localized simultaneously, and no synchronization requirements are imposed on them; ii) the range estimation is robust to flat-fading channels; iii) the synchronization errors from carrier frequency offsets (CFOs) and random initial phases due to unreliable LOs can be eliminated during the phase extraction; iv) the resolvable range restricted by the integer ambiguity issue increases; and v) the estimated TOAs are associated with the targets automatically, as different frequencies are used as physical identities (IDs) for different targets.

The rest of the paper is organized as follows. In Section II, the system model is introduced for a multi-target scenario. The ranging approach in the DRIPS is proposed in Section III. The localization algorithm is accordingly developed in Section IV, where the integer ambiguity issue is addressed as well. Furthermore, the ESPRIT-type algorithm is designed to estimate the frequencies to deal with frequency uncertainties in Section V. The CRBs of the range estimates are derived in Section VI to reveal the relationship between the ranging performance and system parameters. The simulation results are shown in Section VII. The conclusions are drawn at the end of this paper.

Notations: Upper (lower) bold face letters denote matrices (column vectors). The notation $[\mathbf{X}]_{m,n}$ denotes the element on the mth row and nth column of the matrix $\mathbf{X}$, and the notation $[\mathbf{x}]_n$ indicates the nth element of $\mathbf{x}$. The vector $\mathbf{0}_m$ ($\mathbf{1}_m$) and the matrix $\mathbf{I}_m$ are an all-zero (all-one) column vector of length m and an identity matrix of size $m \times m$, respectively. Moreover, the operators $(\cdot)^T$, $(\cdot)^H$, $*$, $|\cdot|$ and $\lfloor\cdot\rfloor$ denote transpose, conjugate transpose, element-wise conjugate, absolute value, and floor function, respectively. All other notations should be self-explanatory.

II. System Model

Let us consider the multi-target scenario where K active targets start to emit dual-tone signals at different time instants, and M anchors (nodes with known positions) receive and process these signals. Without loss of generality, the bandpass complex representation of the dual-tone signal emitted by the kth target can be written as

$$s_k(t) = a_k e^{j\phi_k} e^{j2\pi f_k(t-t_k)} \left(1 + e^{j2\pi g_k(t-t_k)}\right), \quad (1)$$

where f_k is the frequency of the first tone and greater than zero, g_k is the small frequency difference between the two tones and greater than zero as well, ϕ_k denotes the unknown initial phase offset due to the oscillator uncertainty of the target transmitter, a_k is the real-valued amplitude of the dual-tone signal, and t_k is the unknown time instant that the kth target starts to transmit. Note that the unknown t_k introduces an extra phase difference between the two tones. It is due to the lack of time synchronization among the target transmitters and the anchor receivers. Furthermore, the internal delay of the target transmitter may cause an unknown t_k as well. The following assumptions are adopted throughout the paper.

Assumption 1: All M anchor receivers are synchronized. As a result, the start instant t_k of the kth target transmitter is the same with respect to (w.r.t.) all the anchors.

Assumption 2: The frequency difference g_k is smaller than the channel coherence bandwidth.

Although time synchronization is required among anchors as *Assumption 1*, such a requirement is not imposed on the targets. Therefore, asynchronous targets can be localized together in the DRIPS. With *Assumption 2*, the two tones of $s_k(t)$ experience the same channel fading effect [22]. Hence, a flat-fading channel model is adopted here to account for the fading effect.

Via a flat-fading channel, the received signal $r_l(t)$ at the lth anchor is down converted and bandpass filtered. Thus, the resulting signal $x_l(t)$ is modeled as a summation of K components as

$$x_l(t) = \sum_{k=0}^{K-1} x_{l,k}(t), \quad (2)$$

where $x_{l,k}(t)$ is the component from the kth target, and can be represented as

$$\begin{aligned} x_{l,k}(t) =& \beta_{l,k} s_k(t - \tau_{l,k}) e^{-j2\pi f_l t + j\eta_l} && (3)\\ =& \alpha_{l,k}\beta_{l,k} e^{j2\pi((f_k - f_l)t - f_k(t_k + \tau_{l,k}))} \\ & \times \left(1 + e^{j2\pi g_k(t - t_k - \tau_{l,k})}\right), && (4) \end{aligned}$$

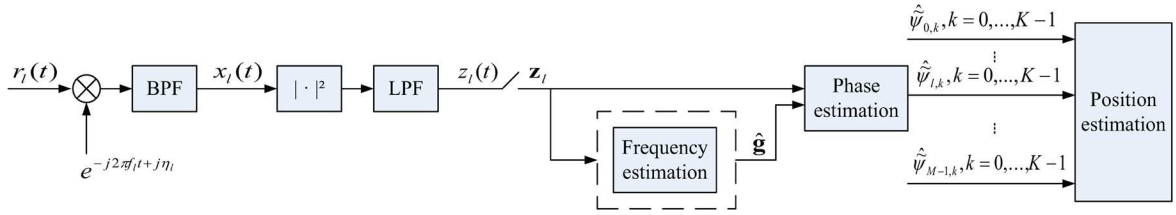


Fig. 1. The receiver structure of the DIRPS.

where (3) indicates the down conversion of the received signal, and (4) is the result of substituting (1) into (3). Note that noise is neglected in (2) for simplicity. The complex coefficient $\beta_{l,k}$ accounts for the flat-fading channel between the lth anchor and the kth target, and can be modeled as a zero-mean complex Gaussian random variable with the variance $\sigma^2_{l,k}$ representing the average power. The complex parameter $\alpha_{l,k}$ absorbs all the random initial phases and the amplitude, and is defined as $\alpha_{l,k} = a_k e^{j(\phi_k+\eta_l)}$, where η_l indicates the unknown random initial phase at the lth anchor. The frequency of the local oscillator (LO) of the lth anchor receiver is denoted by f_l. Moreover, the unknown propagation delay $\tau_{l,k}$ is in a linear relationship w.r.t. the distance $d_{l,k}$ as $d_{l,k} = c\tau_{l,k}$, where $d_{l,k}$ is the unknown distance between the lth anchor and the kth target, and c is the signal propagation speed. In non-line-of-sight scenarios, the relationship $d_{l,k} = c\tau_{l,k}$ may not hold accurately. Calibration methods (e.g. [23]) have to be taken into account. However, this is out of the scope of this paper.

III. Ranging in the DRIPS

In this section, the ranging method based on the phase information is developed. The receiver structure of the DRIPS is illustrated in Fig. 1, where the received signal $r_l(t)$ is first down converted and bandpass filtered. The resulting signal $x_l(t)$ is squared and low-pass filtered to extract multiple low-frequency differential tones, whose phases contain the delay information of interest. Thanks to the low sampling rate for these differential tones, digital signal processing (DSP) blocks are employed sequentially to estimate the frequencies (if accurate knowledge of the tones is not available), the phases and the target positions in a cost-efficient way. We focus on the phase estimation in this section. The position and frequency estimation will be addressed in Section IV and Section V, respectively.

A. Ranging Approach for DRIPS

In order to extract low-frequency differential tones, the preprocessed signal $x_l(t)$ is squared as

$$|x_l(t)|^2 = \sum_{k=0}^{K-1}\sum_{q=0}^{K-1} x_{l,k}(t)x^*_{l,q}(t), \tag{5}$$

where the signal norm $|x_l(t)|^2$ includes not only the autocorrelation terms from the same target, but also the cross-correlation terms from different targets. The autocorrelation terms are of most interest. Plugging (4) into the autocorrelation terms $|x_{l,k}(t)|^2$, where $k \in \mathcal{K}$ and $\mathcal{K} \triangleq \{0, 1, \ldots, K-1\}$, we arrive at

$$|x_{l,k}(t)|^2 = 2|\alpha_{l,k}\beta_{l,k}|^2\left(1+\cos(2\pi g_k t + \psi_{l,k})\right), \quad k \in \mathcal{K}, \tag{6}$$

where the phase of interest is $\psi_{l,k} = -2\pi g_k(t_k + \tau_{l,k})$. According to (6), these low-frequency differential tones contain the delay information of interest in their phases. However, the delay information $\tau_{l,k}$ is coupled with the unknown time instant t_k. Note that f_l and f_k are absent in (6). Although the carrier frequency offsets (CFOs) of the target transmitters and anchor receivers may change f_l and f_k, they do not have any impact on the differential tones. Furthermore, the complex flat-fading channel coefficient $\beta_{l,k}$ and the complex parameter $\alpha_{l,k}$ do not disturb the phases of the differential tones either, since only $|\alpha_{l,k}\beta_{l,k}|^2$ is involved in (6). Therefore, the DRIPS is robust to flat-fading channels and uncertainties of LOs.

Since we are only interested in the autocorrelation terms ($|x_{l,k}(t)|^2, k \in \mathcal{K}$) for the delay ($\tau_{l,k}$) estimation, the cross-correlation terms ($x_{l,k}(t)x^*_{l,q}(t)$, where $k \neq q$ and $k, q \in \mathcal{K}$) have to be separated from the autocorrelation terms in the frequency domain. Furthermore, the autocorrelation terms cannot interfere with each other in order to extract the delay information correctly.

Theorem 1: If the maximum frequency difference ($g_{\max} \triangleq \max\{g_0, g_1, \ldots, g_{K-1}\}$)

$$\begin{aligned} g_{\max} << \min\{&|f_k - f_q|, |f_k - f_q + g_k|, |f_k - f_q - g_q|, \\ &|f_k - f_q + g_k - g_q|\}, \\ &\forall k, \forall q \in \mathcal{K} \text{ and } k \neq q, \end{aligned} \tag{7}$$

then the autocorrelation and the cross-correlation terms can be separated in the frequency domain.

Theorem 2: If the autocorrelation terms do not interfere with each other, then all the frequency differences should be distinct, i.e.,

$$g_k \neq g_q, \quad \forall k, \forall q \in \mathcal{K} \text{ and } k \neq q. \tag{8}$$

Proofs of *Theorem 1* and *Theorem 2* are provided in Appendix A and B, respectively. *Theorem 1* helps to eliminate the nuisance cross-correlation terms by low-pass filtering the squared received signal, and *Theorem 2* prevents the autocorrelation terms from interfering with each other.

As a result, only the DC component and the low-frequency differential tones originated from the autocorrelation terms remain after low-pass filtering, which is denoted by $\mathrm{LPF}\{\cdot\}$. The output of the low pass filter (LPF) can be modeled as

$$\begin{aligned} z_l(t) &= \mathrm{LPF}\left\{|x_l(t)|^2\right\} \\ &= 2\sum_{k=0}^{K-1}|\alpha_{l,k}\beta_{l,k}|^2\left(1+\cos(2\pi g_k t + \psi_{l,k})\right). \end{aligned} \tag{9}$$

The bandwidth of the LPF is $B_L/2$, which satisfies that $B_L > 2g_{\max}$. Therefore, the output $z_l(t)$ can be sampled at a relatively low rate $1/T_s$ ($1/T_s \geq 2g_{\max}$). The subsequent DSP blocks can

be implemented in a cost-efficient way. Collect N samples of $z_l(t)$ into a vector $\mathbf{z}_l$, whose model is given by

$$\mathbf{z}_l = \mathbf{F}_N(\mathbf{g})\mathbf{y}_l + \rho\mathbf{1}_N, \tag{10}$$

where $\mathbf{y}_l = [\boldsymbol{\gamma}_l^T \ \boldsymbol{\gamma}_l^H]^T$ with $\boldsymbol{\gamma}_l = [\gamma_{l,0}, \ldots, \gamma_{l,K-1}]^T$, $\gamma_{l,k} = |\alpha_{l,k}\beta_{l,k}|^2 e^{j\psi_{l,k}}$, $\mathbf{F}_N(\mathbf{g}) = [\boldsymbol{\Phi}_N(\mathbf{g}) \ \ \boldsymbol{\Phi}_N^*(\mathbf{g})]$ with $\boldsymbol{\Phi}_N(\mathbf{g}) = [\boldsymbol{\phi}_N(g_0) \ldots \boldsymbol{\phi}_N(g_{K-1})]$, $\mathbf{g} = [g_0, \ldots, g_{K-1}]^T$, $\boldsymbol{\phi}_N(f) = [1, e^{j2\pi f T_s}, e^{j2\pi 2 f T_s}, \ldots, e^{j2\pi(N-1)fT_s}]^T$, and $\rho = 2\sum_{k=0}^{K-1} |\alpha_{l,k}\beta_{l,k}|^2$.

In the data model (10), parameter ρ is a nuisance parameter. Thus, the orthogonal projection $\mathbf{P}_N$ of $\mathbf{1}_N$ is used to eliminate the term $\rho\mathbf{1}_N$, where $\mathbf{P}_N = \mathbf{I}_N - \frac{1}{N}\mathbf{1}_N\mathbf{1}_N^T$. Premultiplying $\mathbf{P}_N$ on both sides of (10), we arrive at

$$\mathbf{P}_N\mathbf{z}_l = \mathbf{P}_N\mathbf{F}_N(\mathbf{g})\mathbf{y}_l, \tag{11}$$

where $\rho\mathbf{1}_N$ is removed. Assuming accurate knowledge of $\mathbf{g}$, we can estimate $\mathbf{y}_l$ by a least-squares (LS) estimator as

$$\hat{\mathbf{y}}_l = \left(\mathbf{P}_N\mathbf{F}_N(\mathbf{g})\right)^{\dagger}\mathbf{P}_N\mathbf{z}_l, \tag{12}$$

where $\mathbf{F}_N(\mathbf{g})$ of size $N \times 2K$ should be a tall matrix. There are in total $2K+1$ unknown parameters in (10). Thus, the condition that $N > 2K+1$ has to be fulfilled.

Note that by extracting the phase information of $\hat{\mathbf{y}}_l$, we can only obtain the estimate of the phase $\widetilde{\psi}_{l,k}$ instead of $\psi_{l,k}$ as

$$\widehat{\widetilde{\psi}}_{l,k} = \frac{1}{2}\left(\arg\left([\hat{\mathbf{y}}_l]_{k+1}\right) - \arg\left([\hat{\mathbf{y}}_l]_{K+k+1}\right)\right), \tag{13}$$

where $\widetilde{\psi}_{l,k}$ is the fractional part of $\psi_{l,k}$, and given by $\widetilde{\psi}_{l,k} = \psi_{l,k} - 2\pi\lfloor\psi_{l,k}/2\pi\rfloor$ with an unknown integer $\lfloor\psi_{l,k}/2\pi\rfloor$. This is the well-known integer ambiguity problem due to phase wrapping [24], [25]. Recalling that $\psi_{l,k} = -2\pi g_k(t_k + \tau_{l,k})$, the delay of interest $\tau_{l,k}$ is coupled with the unknown time instant t_k. Let us define their corresponding phase parameters as $\varphi_k = -2\pi g_k t_k$, $\widetilde{\varphi}_k = \varphi_k - 2\pi\lfloor\varphi_k/2\pi\rfloor$, $\theta_{l,k} = -2\pi g_k\tau_{l,k}$ and $\widetilde{\theta}_{l,k} = \theta_{l,k} - 2\pi\lfloor\theta_{l,k}/2\pi\rfloor$. Hence, $\psi_{l,k} = \varphi_k + \theta_{l,k}$. We further achieve

$$\widetilde{\psi}_{l,k} = \begin{cases} \widetilde{\varphi}_k + \widetilde{\theta}_{l,k} & \text{if } 0 \le \widetilde{\varphi}_k + \widetilde{\theta}_{l,k} < 2\pi \\ \widetilde{\varphi}_k + \widetilde{\theta}_{l,k} - 2\pi & \text{if } 2\pi \le \widetilde{\varphi}_k + \widetilde{\theta}_{l,k} < 4\pi \end{cases}, \tag{14}$$

and

$$\lfloor\psi_{l,k}/2\pi\rfloor = \begin{cases} \lfloor\varphi_k/2\pi\rfloor + \lfloor\theta_{l,k}/2\pi\rfloor & \text{if } 0 \le \widetilde{\varphi}_k + \widetilde{\theta}_{l,k} < 2\pi \\ \lfloor\varphi_k/2\pi\rfloor + \lfloor\theta_{l,k}/2\pi\rfloor + 1 & \text{if } 2\pi \le \widetilde{\varphi}_k + \widetilde{\theta}_{l,k} < 4\pi. \end{cases} \tag{15}$$

As we observe from (14) and (15), $\widetilde{\varphi}_k$ is coupled with $\widetilde{\theta}_{l,k}$. By making a phase difference as $\widetilde{\psi}_{l,k} - \widetilde{\psi}_{m,k}$, we cannot guarantee to eliminate the effect of $\widetilde{\varphi}_k$, since there are two possible values of $\widetilde{\psi}_{l,k}$. Therefore, we will take this integer ambiguity problem into account for the development of localization algorithms in Section IV. Now let us extract the delay information from the phase estimate. Recalling that $\varphi_k = -2\pi g_k t_k$ and $\theta_{l,k} = -2\pi g_k\tau_{l,k}$, we obtain an aggregate estimate of $t_k + \tau_{l,k}$ as

$$\hat{t}_k + \hat{\tau}_{l,k} = -\left(\widehat{\widetilde{\psi}}_{l,k}/2\pi + \lfloor\psi_{l,k}/2\pi\rfloor\right)/g_k, \tag{16}$$

where the unknown $\lfloor\psi_{l,k}/2\pi\rfloor$ accounts for the integer ambiguity and $\widehat{\widetilde{\psi}}_{l,k}$ is given by (13). We call this aggregate estimate (16) as the biased TOA estimate.

B. Remarks and Discussions

The following remarks are in order.

Remark 1 (Violation of Theorem 1): When *Theorem 1* is violated, the cross-correlation terms $x_{l,k}(t)x_{l,q}^*(t)$ add new columns to $\mathbf{F}_N(\mathbf{g})$. The estimation performance degrades as more parameters have to be estimated. When the number of samples N is less than or equal to the number of columns of $\mathbf{F}_N(\mathbf{g})$, which indicates that the cross-correlation terms add columns no less than $N - 2K - 1$, the matrix $\mathbf{P}_N\mathbf{F}_N(\mathbf{g})$ is rank deficient and (11) is underdetermined. We cannot solve (11) uniquely.

Remark 2 (Violation of Theorem 2: When *Theorem 2* is violated, the number of columns of $\mathbf{F}_N(\mathbf{g})$ is reduced, and we can still estimate $\mathbf{y}_l$ as in (12). However, for the targets with the same g_k, the corresponding delay information cannot be obtained correctly from the signal phase, as they interfere with each other.

Remark 3 (Multiple Asynchronous Targets): In DRIPS, each anchor receiver can estimate all the biased TOAs related to the asynchronous targets together. This is different from most RIPS-type systems, where phase estimation has to be carried out sequentially. Hence, there is no need to schedule the target transmission in DRIPS. This reduces the localization latency significantly. Although the more targets are to be estimated simultaneously, the larger number of samples are required. This increases the localization latency slightly. In general, the localization latency of DRIPS is shorter than that of most RIPS-type systems. Furthermore, a new target can enter the network freely and be localized as long as the frequency assignment for its dual-tone signal is available. Thus, from the network point of view, the target broadcast scheme in the DRIPS is power efficient, and reduces the communication overhead. The computational burden is left to powerful anchor receivers, and the hardware of targets can be simplified.

Remark 4 (DRIPS for Fading Channels): The proposed DRIPS is robust to flat-fading channels. Thanks to small g_k's, which are smaller than the channel coherence bandwidth, the dual tones go through the same fading channel [22, pp. 899]. Moreover, the bandwidth of the tones is much smaller than the channel coherence bandwidth as well. Therefore, the flat-fading channel model fits well for the DRIPS. On the other hand, the performance of the RIPS-type systems degrades under flat-fading channels, because the differential tone is created by mixing the sinusoids from different transmitters at different locations, and its phase is influenced by the unknown channel coefficients from different links. Appendix C shows the detailed reasons about the failure of a tailored RIPS under flat-fading channels.

Remark 5 (CFOs and Random Initial Phases of LOs): Since CFOs and random initial phases of LOs are absent in (6) and (9), they do not disturb the frequencies and phases of the differential tones. No extra calibration is needed. On the other hand, although the random initial phases can be eliminated by making the phase difference in the RIPS [12], the CFOs would cause

frequency uncertainties. Therefore, in the RIPS the frequencies have to be estimated in the first place.

Remark 6 (Integer Ambiguity Issues): Let us recall that $\theta_{l,k} = -2\pi g_k \tau_{l,k}$, $g_k > 0$ and $\tau_{l,k} \geq 0$. If $-2\pi < \theta_{l,k} \leq 0$, we achieve $\widetilde{\theta}_{l,k} = 2\pi + \theta_{l,k}$ and $\lfloor \theta_{l,k}/2\pi \rfloor = -1$. It is then possible to estimate the target position with a simplified integer ambiguity problem. We will further develop the corresponding localization method in Section IV.

Remark 7 (Physical IDs): Based on *Theorem 2*, all g_k's are distinguishable. Hence, they can be used as identities (IDs) for the targets. Making use of these physical IDs, the biased TOA estimates associate with the corresponding targets automatically. Furthermore, under the circumstance that the dual-tone signal experiences deep fading, the frequencies can hop. For example, the frequencies f_k's can be swapped among the targets to explore the frequency diversity.

Remark 8 (Synchronous Anchors): Following *Assumption 1*, the anchor receivers have to be synchronized first. The clock offset (denoted by ε_l) of the lth anchor receiver would add an additional offset to all the phases of interest as $\psi_{l,k} = -2\pi g_k(t_k + \tau_{l,k}) + 2\pi g_k \varepsilon_l, k \in \mathcal{K}$. Hence, the clock offset degrades the localization performance of the DRIPS. The accurate synchronization among anchors is necessary. As the clock offset ε_l is coupled with $\tau_{l,k}$ in the same way as t_k, the synchronization accuracy has to be within nanosecond (ns) if the localization requires a sub-meter accuracy. It can be accomplished by wired (Ethernet over fiber [26], [27]) or wireless approaches (physical layer timestamping [28]). Furthermore, for the anchors with large clock offsets, their range estimates can be filtered out by higher level algorithms as outliers. The impact of synchronization errors on localization performance will be investigated in the simulation section. Moreover, the clock skews of anchor receivers introduce frequency errors. Thus, frequency estimation has to be accomplished before the phase estimation. An ESPRIT-type algorithm will be proposed in Section V for the frequency estimation.

Remark 9 (TOA Vs. Q-Range): The Q-range generated by RIPS [12] is a combination of four distances, and involves four nodes. Whenever the measurement of the Q-range involves more than one unknown target, the Q-range will have a complicated relationship w.r.t. the targets' coordinates. Low-complexity localization algorithms using the Q-range are yet developed. On the other hand, rich literature (see [29], [30] and references therein) is available for TOA-based localization. The localization algorithms using the biased TOAs will be developed to tradeoff complexity with performance in Section IV.

Remark 10 (Dual-Tone Vs. Multi-Tone): We only use dual-tone instead of multi-tone signals, since the number of nuisance frequency components increases dramatically with the number of tones. Although using the multi-tone signal can obtain more differential tones bearing delay information, the benefit from extra information is not worthwhile for the loss of the transmission efficiency. Furthermore, as the increase of the number of tones, the bandwidth of the multi-tone signal may exceed the channel coherence bandwidth.

C. Design Parameters for DRIPS

Now let us investigate the system parameter design. Inspired by the orthogonal subcarriers of OFDM systems in [31], we design $f_k \in \mathcal{F}_b$ and $g_k \in \mathcal{G}$, where $\mathcal{F}_b = \{\widetilde{f_0}, \widetilde{f_0} + N_g\Delta g, \ldots, \widetilde{f_0} + (K-1)N_g\Delta g\}$, $\mathcal{G} = \{\Delta g, 2\Delta g, \ldots, K\Delta g\}$, $\widetilde{f_0}$ is the smallest frequency of the first tone, $N_g\Delta g$ (N_g is an integer) is the frequency difference between the first tones of adjacent dual-tone signals, and Δg is the smallest frequency difference between the two tones of the dual-tone signal. An example of the frequency allocation of the DRIPS is shown in Fig. 2, where the preprocessed received signal $x_l(t)$ is composed of multiple dual-tone signals well apart, and the output of the LPF $z_l(t)$ contains a group of differential tones separated equally by Δg. According to *Theorem 1*, we arrive at

$$\Delta f_{\min} = N_g\Delta g - (K-1)\Delta g \quad > \quad K\Delta g, \tag{17}$$

where $\Delta f_{\min}$ is the minimum frequency difference between tones from different targets as shown in Fig. 2. Hence, it is the minimum frequency component created by the cross-correlation terms as well. Moreover, the maximum frequency difference between two tones employed by the same target is $K\Delta g$. Therefore, the cut-off frequency of the LPF f_{cut} should satisfy the condition that $K\Delta g < f_{\text{cut}} < \Delta f_{\min}$ in order to filter out the nuisance cross-correlation terms and retain the autocorrelation terms of interest. Consequently, the frequency difference between the smallest and the largest tone is $\Delta f_{\max}$, which is given by

$$\begin{aligned}\Delta f_{\max} &= (K-1)N_g\Delta g + K\Delta g \\ &= ((K-1)N_g + K)\,\Delta g \leq B,\end{aligned} \tag{18}$$

with B being the bandwidth of $x_l(t)$ designed by the receiver specifications. The larger N_g is, the easier the LPF can be developed. On the other hand, the larger N_g is, the larger $\Delta f_{\max}$ (or B) is. There is a tradeoff to design N_g. Furthermore, recalling that we employ a flat-fading channel model, g_k ($g_k \in \mathcal{G}$) should be smaller than the channel coherence bandwidth $B_{l,k}$, which denotes the channel coherence bandwidth between the lth anchor receiver and the kth target transmitter. As a result, $K\Delta g$ should be smaller than the smallest channel coherence bandwidth, and we arrive at

$$K\Delta g < \min\{B_{l,k}\}, \quad l \in \mathcal{M}, \quad k \in \mathcal{K}, \tag{19}$$

where $\mathcal{M} = \{0, 1, \ldots, M-1\}$ with M being the number of anchors. Moreover, according to (10), the number of samples N should be larger than $2K+1$ ($N > 2K+1$). Thus, given the observation window length of the anchor receiver T_o and the Nyquist sampling rate ($2K\Delta g$), we obtain

$$2KT_o\Delta g > 2K + 1. \tag{20}$$

In summary, the design of Δg is decided by the channel coherence bandwidth and the receiver system specifications, and thus it should fulfill the following condition:

$$\frac{2K+1}{2KT_o} < \Delta g < \min\left\{\frac{B_{l,k}}{K}, \frac{B}{(K-1)N_g + K}\right\}, \tag{21}$$

where $l \in \mathcal{M}$ and $k \in \mathcal{K}$.

Let us show a practical example of parameter design. We assume four active targets ($K = 4$ and $\mathcal{K} = \{0, 1, 2, 3\}$). The frequency allocation $f_k \in \mathcal{F}_b$ and $g_k \in \mathcal{G}, \forall k \in \mathcal{K}$ is

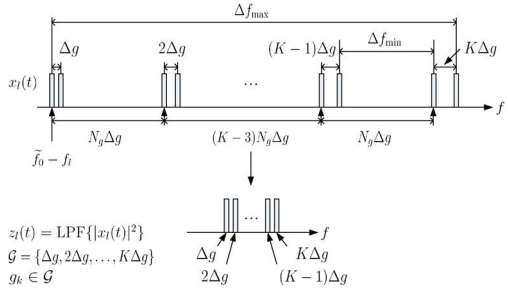


Fig. 2. An example of frequency allocation of DRIPS.

employed. Based on the condition imposed by (17), we have $N_g >> 2K-1=7$. To help the implementation of the LPF, which is used to extract the low-frequency differential tones, N_g can be assigned as 20. Furthermore, $K\Delta g$ should be smaller than the channel coherence bandwidth to validate the flat-fading channel assumption. Considering a typical wireless localization scenario (e.g. the extended typical urban (ETU) channel model in the 3GPP-LTE standard [32]), the channel coherence bandwidth is about 100 kHz. Moreover, the observation window length T_o is set to 1 millisecond (ms). Taking all the constraints into account, we obtain $1.25\ \text{kHz} < \Delta g < 25$ kHz based on (21). Thus, Δg is assigned to 10 kHz.

IV. Localization With an Integer Ambiguity Problem

Let us recall *Remark 6* in Section III-B. If the phase $\theta_{l,k}$ due to the delay $\tau_{l,k}$ is in the range of $(-2\pi, 0]$, we achieve $\widetilde{\theta}_{l,k} = 2\pi + \theta_{l,k}$ and $\lfloor \theta_{l,k}/2\pi \rfloor = -1$. Therefore, we can simplify (14) as

$$\widetilde{\psi}_{l,k} = \begin{cases} \widetilde{\varphi}_k + \theta_{l,k} + 2\pi & \text{if } -2\pi \le \widetilde{\varphi}_k + \theta_{l,k} < 0 \\ \widetilde{\varphi}_k + \theta_{l,k} & \text{if } 0 \le \widetilde{\varphi}_k + \theta_{l,k} < 2\pi \end{cases}, \quad (22)$$

where $\widetilde{\psi}_{l,k}$ only has two possible values. The integer ambiguity problem is simplified. In order to fulfill the condition that $-2\pi < \theta_{l,k} \le 0$, where $\theta_{l,k} = -2\pi g_k d_{l,k}/c$, we need to design g_k according to the range of interest. For example, when the distance of interest is limited to a few hundreds meters or less than 1 km ($0 < d_{l,k} \le 1$ km), we should choose g_k to be smaller than 300 kHz. This adds a further constraint to the choice of g_k. In another point of view, we are able to resolve $d_{l,k}$ in the range of $[0, c/g_k]$. Therefore, we rewrite (22) w.r.t. $d_{l,k}$ and obtain

$$-\frac{c\widetilde{\psi}_{l,k}}{2\pi g_k} = \begin{cases} d_{l,k} + b_k - c/g_k & \text{if } -2\pi \le \widetilde{\varphi}_k + \theta_{l,k} < 0 \\ d_{l,k} + b_k & \text{if } 0 \le \widetilde{\varphi}_k + \theta_{l,k} < 2\pi \end{cases}, \quad (23)$$

where $b_k = -c\widetilde{\varphi}_k/(2\pi g_k)$ caused by the unknown transmission instant t_k.

We collect all the phase estimates $\hat{\widetilde{\psi}}_{l,k}, l \in \mathcal{M}$ from the anchors corresponding to the same kth target into a vector as $\mathbf{v}_k = -c/(2\pi g_k)[\hat{\widetilde{\psi}}_{0,k}, \ldots, \hat{\widetilde{\psi}}_{M-1,k}]^T$. Consequentially, the model of $\mathbf{v}_k$ can be given by

$$\mathbf{v}_k = \mathbf{d}_k + b_k \mathbf{1}_M + (c/g_k)\,\mathbf{u}_k, \quad (24)$$

where $\mathbf{d}_k = [d_{0,k}, d_{1,k}, \ldots, d_{M-1,k}]^T$ with $d_{l,k} = \|\boldsymbol{v}_l - \mathbf{o}_k\|$, $\boldsymbol{v}_l$ and $\mathbf{o}_k$ denoting respectively the coordinates of the lth anchor and the kth target, and $\mathbf{u}_k$ belongs to the set $\{0, -1\}^M$ due to the constraint that $-2\pi < \theta_{l,k} \le 0$. We remark here using the estimated phase difference $\hat{\widetilde{\psi}}_{l,k} - \hat{\widetilde{\psi}}_{m,k}$ to eliminate b_k is equivalent to assuming that $\mathbf{u}_k = \mathbf{0}_M$ or $\mathbf{u}_k = \mathbf{1}_M$. However, $\mathbf{u}_k \in \{0, -1\}^M$. We should take this unknown $\mathbf{u}_k$ into account in the localization algorithm.

Note that (24) is in a nonlinear relationship w.r.t. the unknown $\mathbf{o}_k$ via $\mathbf{d}_k$, and a linear relationship w.r.t. the unknown b_k and the unknown $\mathbf{u}_k$. In general, it is difficult to solve (24) directly. We employ an alternating LS (ALS) solution here [33]. The unknown parameters $\mathbf{o}_k$ and b_k are categorized as a subset, while $\mathbf{u}_k$ is an another subset. Let us define the estimates of the nth iteration as $\hat{\mathbf{o}}_k^{(n)}$, $\hat{b}_k^{(n)}$, and $\hat{\mathbf{u}}_k^{(n)}$. Assuming knowledge of $\hat{\mathbf{u}}_k^{(n-1)}$, we achieve $\hat{\mathbf{o}}_k^{(n)}$ and $\hat{b}_k^{(n)}$ by linearizing (24) and solving an LS problem as in [30]. Moving $(c/g_k)\,\hat{\mathbf{u}}_k^{(n-1)}$ and $b_k\mathbf{1}_M$ to the left side of (24), making element-wise multiplication and rearranging, we arrive at

$$\boldsymbol{\Phi}_a - \widetilde{\mathbf{v}}_k^{(n)} \odot \widetilde{\mathbf{v}}_k^{(n)} = 2\mathbf{A}^T\mathbf{o}_k - 2b_k\widetilde{\mathbf{v}}_k^{(n)} + \left(b_k^2 - \|\mathbf{o}_k\|^2\right)\mathbf{1}_M. \quad (25)$$

where $\odot$ denotes element-wise multiplication, $\widetilde{\mathbf{v}}_k^{(n)} = \mathbf{v}_k - (c/g_k)\hat{\mathbf{u}}_k^{(n-1)}$, $\mathbf{A} = [\boldsymbol{v}_0\ \ldots \boldsymbol{v}_{M-1}]$, and $\boldsymbol{\Phi}_a = [\|\boldsymbol{v}_0\|^2, \|\boldsymbol{v}_1\|^2, \ldots, \|\boldsymbol{v}_{M-1}\|^2]^T$. As a result, we obtain a linear model (25) w.r.t. $\mathbf{o}_k$ and b_k. The estimate of $\mathbf{o}_k^{(n)}$ and $b_k^{(n)}$ can be achieved accordingly based on (25) using the LS estimator. Note that the number of anchors M has to be greater than four to solve (25) in the LS manner. In other words, at least four anchors are needed in the DRIPS. Once we use $\hat{\mathbf{o}}_k^{(n)}$ to calculate $\hat{\mathbf{d}}_k^{(n)}$, it can be used together with $\hat{b}_k^{(n)}$ to update $\hat{\mathbf{u}}_k^{(n)}$ by employing a simple rounding operation as

$$\hat{\mathbf{u}}_k^{(n)} = \begin{cases} 0, & \text{round}\left((g_k/c)\left(\mathbf{v}_k - \hat{\mathbf{d}}_k^{(n)} - \hat{b}_k^{(n)}\mathbf{1}_M\right)\right) \ge 0 \\ -1, & \text{round}\left((g_k/c)\left(\mathbf{v}_k - \hat{\mathbf{d}}_k^{(n)} - \hat{b}_k^{(n)}\mathbf{1}_M\right)\right) < 0 \end{cases} \quad (26)$$

The ALS algorithm is summarized in Algorithm 1.

Algorithm 1: Alternating Least Squares (ALS) Approach to Estimate the Target Position

1) Initial step: $\hat{\mathbf{u}}_k^{(0)} = \mathbf{0}$

2) Update $\hat{\mathbf{o}}_k^{(n)}$ and $\hat{b}_k^{(n)}$ given $\hat{\mathbf{u}}_k^{(n-1)}$ based on the linear model (25)

3) Update $\hat{\mathbf{u}}_k^{(n)}$ given $\hat{\mathbf{o}}_k^{(n)}$ and $\hat{b}_k^{(n)}$ using (26)

4) Go back to step 2), unless $\|\hat{\mathbf{o}}_k^{(n)} - \hat{\mathbf{o}}_k^{(n-1)}\| < \epsilon$ or reach the maximum iteration number

We remark here that when M is small (e.g. $M < 16$), we can apply a brute-force search for $\mathbf{u}_k$ and the LS estimator based on (25) for $\mathbf{o}_k$ as well.

V. DRIPS Without Accurate Knowledge of Frequencies

According to (12), the estimate of $\mathbf{y}_l$ depends on accurate knowledge of $\mathbf{g}$. However, the transmitter or receiver may not have accurate knowledge of $\mathbf{g}$ due to various reasons, such as uncertainties of LOs (e.g. clock skews) and unreliable environments (e.g. cheap devices or temperature variations). Therefore,

TABLE I
THE SUMMARY OF THE PROPOSED ALGORITHMS IN DRIPS

	Algorithms
The phase estimator	Proposed by (13) in Section III-A
The position estimator	Proposed by Algorithm 1 in Section IV
The frequency estimator	Proposed by (34) in Section V
Design of f_k and g_k	Follow *Theorem 1* and *Theorem 2* in Section III-A
A special design of Δg	Fulfill (21) in Section III-C

we need to first estimate $\mathbf{g}$ to construct $\mathbf{F}_N(\hat{\mathbf{g}})$ accordingly, and then estimate $\mathbf{y}_l$. The frequency estimation block in Fig. 1 is needed to accomplish the localization task.

The sample vector $\mathbf{z}_l$ is rearranged into a matrix $\mathbf{Z}_l$ of size $p \times q$ with $q = N - p + 1$ as

$$\mathbf{Z}_l = [\,[\mathbf{z}_l]_{1:p} \quad [\mathbf{z}_l]_{2:p+1} \quad [\mathbf{z}_l]_{3:p+2} \quad \cdots, \quad [\mathbf{z}_l]_{q:N}\,], \quad (27)$$

where $[\mathbf{a}]_{m:n}$ represents the vector composed of the mth to the nth elements of the vector $\mathbf{a}$. Consequently, $\mathbf{Z}_l$ can be factorized as

$$\mathbf{Z}_l = \mathbf{F}_p(\mathbf{g})\mathrm{diag}(\mathbf{y}_l)\mathbf{F}_q^T(\mathbf{g}) + \rho\mathbf{1}_p\mathbf{1}_q^T. \quad (28)$$

When $\mathbf{F}_p(\mathbf{g})$ and $\mathbf{F}_q(\mathbf{g})$ are tall matrices, the rank of $\mathbf{Z}_l$ is no greater than $2K+1$. Note that $\mathbf{Z}_l$ reserves the shift-invariant property (see [34], [35]). Hence, we take submatrices of $\mathbf{Z}_l$ as

$$\begin{aligned}\mathbf{Z}_1 &= [\,\mathbf{I}_{p-1} \quad \mathbf{0}_{p-1}\,]\,\mathbf{Z}_l \\ &= \mathbf{F}_{p-1}(\mathbf{g})\mathrm{diag}(\mathbf{y}_l)\mathbf{F}_q^T(\mathbf{g}) + \rho\mathbf{1}_{p-1}\mathbf{1}_q^T, \quad (29)\\ \mathbf{Z}_2 &= [\,\mathbf{0}_{p-1} \quad \mathbf{I}_{p-1}\,]\,\mathbf{Z}_l \\ &= \mathbf{F}_{p-1}(\mathbf{g})\mathbf{\Lambda}\mathrm{diag}(\mathbf{y}_l)\mathbf{F}_q^T(\mathbf{g}) + \rho\mathbf{1}_{p-1}\mathbf{1}_q^T, \quad (30)\end{aligned}$$

where $\mathbf{\Lambda} = \mathrm{diag}([\boldsymbol{\lambda}^T \;\; \boldsymbol{\lambda}^H]^T)$ and $\boldsymbol{\lambda} = [e^{j2\pi g_0 T_s}, e^{j2\pi g_1 T_s}, \ldots, e^{j2\pi g_{K-1}T_s}]^T$. In order to eliminate the nuisance term $\rho\mathbf{1}_{p-1}\mathbf{1}_q^T$, we apply $\mathbf{P}_{p-1}$ and $\mathbf{P}_q$ on both sides of $\mathbf{Z}_1$ and $\mathbf{Z}_2$, where $\mathbf{P}_{p-1}$ and $\mathbf{P}_q$ are the orthogonal projection of $\mathbf{1}_{p-1}$ and $\mathbf{1}_q$, respectively. As a result, we arrive at

$$\widetilde{\mathbf{Z}}_1 \triangleq \mathbf{P}_{p-1}\mathbf{Z}_1\mathbf{P}_q = \mathbf{P}_{p-1}\mathbf{F}_{p-1}(\mathbf{g})\mathbf{T} \quad (31)$$

$$\widetilde{\mathbf{Z}}_2 \triangleq \mathbf{P}_{p-1}\mathbf{Z}_2\mathbf{P}_q = \mathbf{P}_{p-1}\mathbf{F}_{p-1}(\mathbf{g})\mathbf{\Lambda}\mathbf{T}, \quad (32)$$

where $\mathbf{T} = \mathrm{diag}(\mathbf{y}_l)\mathbf{F}_q^T(\mathbf{g})\mathbf{P}_q$. Therefore, applying the ESPRIT algorithm results in

$$\widetilde{\mathbf{Z}}_1^\dagger\widetilde{\mathbf{Z}}_2 = \mathbf{T}^\dagger\mathbf{\Lambda}\mathbf{T}. \quad (33)$$

Let us define the eigenvalue estimates of $\widetilde{\mathbf{Z}}_1^\dagger\widetilde{\mathbf{Z}}_2$ as $\hat{\lambda}_0, \hat{\lambda}_1, \ldots, \hat{\lambda}_{2K-1}$ in decreasing order based on their absolute values. Note that the noise term is omitted in the data model (28). However, it impacts the eigenvalues of $\widetilde{\mathbf{Z}}_1^\dagger\widetilde{\mathbf{Z}}_2$, and its influence on the eigenvalues w.r.t. the same absolute frequency is the same. Therefore, the correct eigenvalue pairs can be found by sorting the eigenvalues based on their absolute values. Accordingly, the estimate of $\mathbf{g}$ can be obtained as

$$\hat{\mathbf{g}} \in \left\{ \frac{\arg\left(\hat{\lambda}_{2k}\hat{\lambda}_{2k+1}^*\right)}{4\pi T_s}, k \in \mathcal{K} \right\}, \quad (34)$$

where the association of the estimated frequencies to the corresponding targets can be accomplished using prior inaccurate knowledge of $\mathbf{g}$. Furthermore, in order to employ the ESPRIT algorithm, the factorization of $\widetilde{\mathbf{Z}}_1$ (31) (or $\widetilde{\mathbf{Z}}_2$ (32)) should be low-rank. Thus, $\mathbf{F}_{p-1}(\mathbf{g})$ and $\mathbf{F}_q(\mathbf{g})$ should be strictly tall. As $\mathbf{P}_{p-1}$ and $\mathbf{P}_q$ are coupled with them, we achieve the following identifiability criteria

$$p \geq 2K + 2, \quad (35)$$

$$q \geq 2K + 1. \quad (36)$$

Plugging $q = N - p + 1$ into (36), we obtain $N - 2K \geq p \geq 2K + 2$ (thus $N \geq 4K + 2$), and choose $p = N/3$, where the number of columns is two times of the number of rows. Furthermore, we notice that the forward-backward averaging technique does not achieve obvious benefits here due to the structures of $\widetilde{\mathbf{Z}}_1$ and $\widetilde{\mathbf{Z}}_2$.

In summary, all the algorithms proposed in the DRIPS are collected in Table I.

VI. CRAMÉR-RAO BOUND

The theoretical performance bounds are of great interest to show the optimal estimation performance, which can be used as the benchmark to evaluate performance of estimators. In this section, we derive the Cramér-Rao bound (CRB) based on the data model (9). Adding the missing noise term, the data model (9) is rewritten as

$$z_l(t) = 2\sum_{k=0}^{K-1} |\alpha_{l,k}\beta_{l,k}|^2 \left(1 + \cos(2\pi g_k t + \psi_{l,k})\right) + \widetilde{m}_l(t), \quad (37)$$

where $\widetilde{m}_l(t)$ is the low-pass filtered $m_l(t)$, which is an aggregate noise term given by

$$m_l(t) = w_l^*(t)\sum_{k=0}^{K-1} x_{l,k}(t) + w_l(t)\sum_{k=0}^{K-1} x_{l,k}^*(t) + |w_l(t)|^2. \quad (38)$$

This aggregate noise term $m_l(t)$ is the result of the squaring operation, and thus includes the signal-noise cross-correlation terms and the noise autocorrelation term. The noise term omitted in the model (2) of the preprocessed signal $x_l(t)$ is denoted by $w_l(t)$, which appears in (38). It can be modeled as a zero-mean complex Gaussian random process with the power spectrum density (PSD) N_0 in the band of interest with B being the bandwidth of the receiver. For the sake of brevity, the subscription l is dropped from now on. The output of the LPF $z(t)$ in (37) is sampled at the Nyquist rate $1/B_L$, where $B_L/2$ is the bandwidth of the LPF. Let us define $\boldsymbol{\psi} = [\psi_0, \psi_1, \ldots, \psi_{K-1}]^T$, $\widetilde{\mathbf{g}} = [\widetilde{g}_0, \widetilde{g}_1, \ldots, \widetilde{g}_{K-1}]^T$ with $\widetilde{g}_k = g_k/B_L$ the normalized frequency, $\widetilde{\boldsymbol{\beta}} = [\widetilde{\beta}_0, \widetilde{\beta}_1, \ldots, \widetilde{\beta}_{K-1}]^T$ with $\widetilde{\beta}_k = |\beta_k|^2$, and

$\boldsymbol{\theta} = [\boldsymbol{\psi}^T \ \widetilde{\mathbf{g}}^T \ \widetilde{\boldsymbol{\beta}}^T]^T$. Making use of these definitions, the sample vector $\mathbf{z}$ can be written as

$$\mathbf{z} = \mathbf{h}(\boldsymbol{\theta}) + \varrho(\widetilde{\boldsymbol{\beta}})\mathbf{1}_{\widetilde{N}} + \widetilde{\mathbf{m}}, \tag{39}$$

where $\varrho(\widetilde{\boldsymbol{\beta}}) = 2\sum_{k=0}^{K-1}|\alpha_k|^2\widetilde{\beta}_k$, $[\mathbf{h}(\boldsymbol{\theta})]_{n+1} = 2\sum_{k=0}^{K-1}|\alpha_k|^2\widetilde{\beta}_k\cos(2\pi\widetilde{g}_k n + \psi_k)$, $[\widetilde{\mathbf{m}}]_{n+1} = \widetilde{m}(n/B_L)$, $n = 0, 1, \ldots, \widetilde{N}-1$, and $\widetilde{N}$ is the total number of samples using the Nyquist rate $1/B_L$.

Both $\boldsymbol{\psi}$ and $\widetilde{\boldsymbol{\beta}}$ are unknown in (39). Moreover, accurate knowledge of $\widetilde{\mathbf{g}}$ may not be available due to uncertainties of LOs. Thus, two cases without and with accurate knowledge of $\widetilde{\mathbf{g}}$ are considered for calculating CRBs. As the derivations are similar, the CRB derivation for the case without accurate knowledge of $\widetilde{\mathbf{g}}$ is exemplified as follows. In this case, $\widetilde{\mathbf{g}}$ has to be estimated as well, and the CRB for $\boldsymbol{\theta}$ is derived. It places lower bounds on the estimation of the amplitudes, phases and frequencies of a group of sinusoids. Since the distribution of $\widetilde{\boldsymbol{\beta}}$ is available, the Bayesian information matrix (BIM) [36] is employed as

$$\mathbf{I}_B(\boldsymbol{\theta}) = E_{\boldsymbol{\theta}}\left[\mathbf{I}_D(\boldsymbol{\theta})\right] + \mathbf{I}_p(\boldsymbol{\theta}), \tag{40}$$

$$\left[\mathbf{I}_D(\boldsymbol{\theta})\right]_{ij} = -E_{\mathbf{z}|\boldsymbol{\theta}}\left[\frac{\partial^2}{\partial\boldsymbol{\theta}_i\partial\boldsymbol{\theta}_j}\ln p(\mathbf{z}|\boldsymbol{\theta})\right], \tag{41}$$

$$\left[\mathbf{I}_p(\boldsymbol{\theta})\right]_{ij} = -E_{\boldsymbol{\theta}}\left[\frac{\partial^2}{\partial\boldsymbol{\theta}_i\partial\boldsymbol{\theta}_j}\ln p(\boldsymbol{\theta})\right], \tag{42}$$

where $\mathbf{I}_D(\boldsymbol{\theta})$ represents information obtained from the data, $\mathbf{I}_P(\boldsymbol{\theta})$ indicates the prior information, and $p(\mathbf{z}|\boldsymbol{\theta})$ is the probability density function (pdf) of $\mathbf{z}$ conditioned on $\boldsymbol{\theta}$.

Since the stochastic properties of $\widetilde{\mathbf{m}}$ are needed to calculate the CRB, they are derived in Appendix D. The main results are summarized as follows:

$$\boldsymbol{\mu}_{\widetilde{\mathbf{m}}} \triangleq E[\widetilde{\mathbf{m}}] = R_B(0)\mathbf{1}_{\widetilde{N}} = N_0 B\mathbf{1}_{\widetilde{N}}, \tag{43}$$

$$\mathbf{C}_{\widetilde{\mathbf{m}}} \triangleq \mathrm{cov}(\widetilde{\mathbf{m}}) = \mathrm{diag}\left(\widetilde{\sigma}_0^2, \widetilde{\sigma}_1^2, \ldots, \widetilde{\sigma}_{\widetilde{N}-1}^2\right), \tag{44}$$

where $\boldsymbol{\mu}_{\widetilde{\mathbf{m}}}$ and $\mathbf{C}_{\widetilde{\mathbf{m}}}$ are the mean and variance of $\widetilde{\mathbf{m}}$, respectively,

$$\widetilde{\sigma}_n^2 = 4\sum_{k=0}^{K-1}|\alpha_k|^2\sigma_k^2\left(R_{B_L}(0) + R_{B_L-g_k}(0)\cos(2\pi\widetilde{g}_k n + \psi_k)\right), \tag{45}$$

and $R_B(\tau) = N_0 B\sin(\pi B\tau)/(\pi B\tau)$. When the difference between $R_{B_L}(\tau)$ and $R_{B_L-g_k}(\tau)$ is negligible, and $\sum_{k=0}^{K-1}|\alpha_k|^2\sigma_k^2\cos(2\pi g_k t + \psi_k) \approx 0$ due to the random unknown ψ_k and a large K, the covariance element $\widetilde{\sigma}_n^2$ can be further approximated as a constant $\bar{\widetilde{\sigma}}^2$ as

$$\bar{\widetilde{\sigma}}^2 = 4R_{B_L}(0)\sum_{k=0}^{K-1}|\alpha_k|^2\sigma_k^2. \tag{46}$$

We further assume $\widetilde{\mathbf{m}}$ is a Gaussian random vector. As a result, $p(\mathbf{z}|\boldsymbol{\theta})$ is the pdf of a Gaussian distribution.

The detailed derivations of the BIM and CRBs are shown in Appendix E. The CRBs for the two cases without and with accurate knowledge of $\widetilde{\mathbf{g}}$ are distinguished by the subscript $_{(s)}$. For each case, the CRBs using the accurate covariance matrix ($\mathbf{C}_{\widetilde{\mathbf{m}}}$ in (44)) and the approximated one ($\bar{\widetilde{\sigma}}^2$ in (46)) are differentiated by the subscript $^{(\prime)}$. The CRBs using $\mathbf{C}_{\widetilde{\mathbf{m}}}$ do not have closed forms (see $\mathrm{CRB}(\psi_i)$ in (70) and $\mathrm{CRB}_s(\psi_i)$ in (71) in Appendix E). However, the CRBs employing $\bar{\widetilde{\sigma}}^2$ are calculated in closed form, and show more insight about the performance related to the system parameters. In the case without accurate knowledge of $\widetilde{\mathbf{g}}$, the CRB of ψ_i is given by

$$\mathrm{CRB}'(\psi_i) = \frac{2\widetilde{N}-1}{2\widetilde{N}(\widetilde{N}+1)}\frac{\bar{\widetilde{\sigma}}^2}{|\alpha_i|^4\sigma_i^4}, \tag{47}$$

which indicates that the larger number of samples $\widetilde{N}$, the higher average received signal power $|\alpha_i|^2\sigma_i^2$ and the lower noise variance $\bar{\widetilde{\sigma}}^2$ can contribute to achieve better estimation accuracy. Furthermore, the CRB of ψ_i does not depend on ψ_i itself. In the case that accurate knowledge of $\widetilde{\mathbf{g}}$ is available, the CRB of ψ_i can be simplified as

$$\mathrm{CRB}'_s(\psi_i) = \frac{1}{4\widetilde{N}}\frac{\bar{\widetilde{\sigma}}^2}{|\alpha_i|^4\sigma_i^4}. \tag{48}$$

Assuming the distance is in the resolvable range, the CRBs of the distance estimate can be obtained accordingly as

$$\mathrm{CRB}_{(s)}^{(\prime)}(d_i) = \frac{c^2}{4\pi^2 g_i^2}\mathrm{CRB}_{(s)}^{(\prime)}(\psi_i). \tag{49}$$

According to (49), the accuracy of the distance estimate improves with a larger g_i. Note that $\mathrm{CRB}_s^{(\prime)}(d_i) < \mathrm{CRB}^{(\prime)}(d_i)$, as the estimation performance degrades with the increasing number of unknown parameters.

VII. Simulation Results

The performance of the proposed DRIPS under flat-fading channels is evaluated by simulations in this section. Four active targets are assumed ($K = 4$ and $\mathcal{K} = \{0, 1, 2, 3\}$) similar as the example in Section III-C. Therefore, the parameter design of the example in Section III-C is employed here, and $\Delta g = 10$ kHz. Moreover, the distances between the target-anchor pairs are in the range of $[1, 3]$ km. According to *Remark 6*, in order to avoid the integer ambiguity issue, $K\Delta g$ should be less than 100 kHz. With $K = 4$ and $\Delta g = 10$ kHz, we have that $K\Delta g < 100$ kHz. Thus, $\mathbf{g} = [10, 20, 30, 40]$ kHz. Furthermore, all the targets are assumed to transmit the dual-tone signals with the same amplitude as $a_k = 1, \forall k \in \mathcal{K}$. Note that this does not indicate the received signals have the same amplitude, since the instant RSS is determined by the channel fading effect. Moreover, the average channel power of the flat-fading channel is always assumed to be 1, thus $\sigma_{l,k}^2 = 1, \forall k, \forall l \in \mathcal{M}$. Without loss of generality, the signal-to-noise ratio (SNR) is defined as $1/\sigma^2$, where $\sigma^2 = N_0 B$ the variance of the noise term $w_l(t)$. In the following subsections, the ranging and localization performance of the DRIPS is investigated, respectively.

A. Synchronous Targets to Evaluate Ranging Accuracy

Since the range estimates are based on the estimated TOAs with offsets due to the asynchronous targets, all the targets and the anchors are assumed to be synchronized in order to avoid

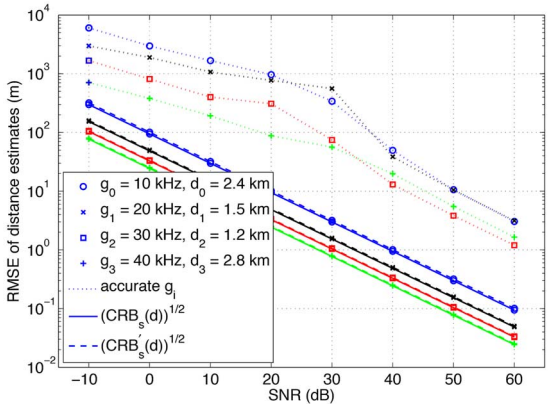


Fig. 3. RMSE of distance estimates versus SNR, $\mathbf{g} = [10, 20, 30, 40]$ kHz and $\mathbf{d} = [2.4, 1.5, 1.2, 2.8]$ km.

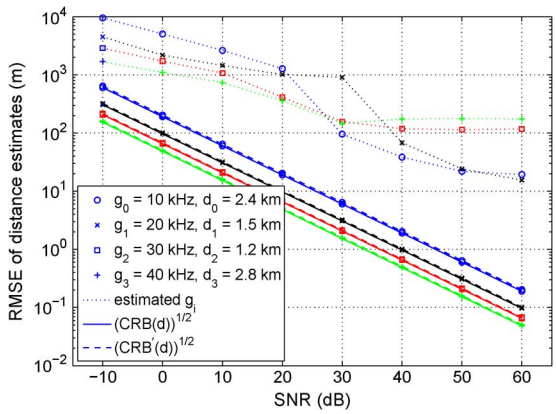


Fig. 4. RMSE of distance estimates versus SNR using the estimated $\mathbf{g}$, where the accurate $\mathbf{g} = [10, 20, 30, 40]$ kHz and $\mathbf{d} = [2.4, 1.5, 1.2, 2.8]$ km.

the offsets and the integer ambiguity issue in this subsection. Thus, $t_k = 0, \forall k \in \mathcal{K}$. In each Monte-Carlo run, a complex coefficient $\beta_{l,k}$ ($\beta_{l,k} \sim \mathcal{CN}(0,1)$) is randomly generated for flat-fading channels. The anchor receives the dual-tone signals from all targets and carries out range estimation.

Example 1: The true distances from the four targets to the lth anchor are fixed to $[2.4, 1.5, 1.2, 2.8]$ km in all the Monte-Carlo runs of this example, and there is no integer ambiguity. Thus, the propagation delays are unknown deterministic parameters, and the channel coefficients are unknown random variables. The root mean square errors (RMSEs) ($\sqrt{\sum_{n=1}^{N}(\hat{d}_{l,k}^{(n)} - d_{l,k})^2/N}$, where $\hat{d}_{l,k}^{(n)}$ is the distance estimate at the nth run) of the distance estimates are evaluated in Figs. 3 and 4. Note that Figs. 3 and 4 correspond to two cases with and without accurate knowledge of $\mathbf{g}$, respectively. Without accurate knowledge of $\mathbf{g}$, the frequencies have to be first estimated by the ESPRIT-type algorithm proposed in Section V, and then the phase information is extracted based on the estimated $\mathbf{g}$.

The square root of the CRBs of the distance estimates are used as benchmarks in both figures. The CRBs using $\mathbf{C}_{\tilde{\mathbf{m}}}$ ($\sqrt{\mathrm{CRB}_{(s)}(d)}$) and the ones using $\tilde{\bar{\sigma}}^2$ ($\sqrt{\mathrm{CRB}'_{(s)}(d)}$) are close to each other. Hence, the approximation is valid. Moreover, the CRBs in Fig. 4 are higher than their counterparts in Fig. 3. Because more parameters are estimated in the case without accurate knowledge of $\mathbf{g}$ for Fig. 4. According to (49), the larger

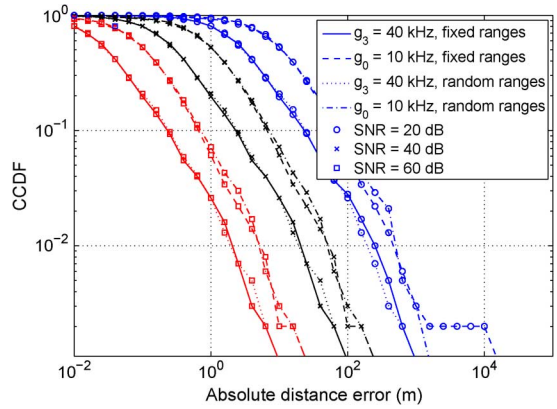


Fig. 5. The CCDF curves of the absolute distance error at $\mathrm{SNR} = \{20\ \mathrm{dB}, 40\ \mathrm{dB}, 60\ \mathrm{dB}\}$ with $g_0 = 10$ kHz and $g_3 = 40$ kHz for fixed and random ranges, respectively.

g_i is, the lower CRB is. Therefore, the lowest CRB is achieved by using g_3 in both figures. Furthermore, there are performance gaps between the CRBs and the RMSEs of the estimator in both figures. The deep fading channel fails the range estimation and causes large errors. The ranging performance further degrades in Fig. 4, as the distances are estimated using the frequency estimates. The estimation method is suboptimal.

Sequentially, the complementary cumulative distribution function (CCDF) curves of the absolute distance error using $g_0 = 10$ kHz (the dashed lines) and $g_3 = 40$ kHz (the solid lines) are illustrated in Fig. 5, respectively. In general, the range estimation that employs larger frequency has less absolute errors. Meanwhile, the higher SNR is, the faster the curve falls. Very few large errors appear due to the deep fading effect.

Example 2: In this example, the propagation delays are randomly generated following a uniform distribution in the range of interest ($[1, 3]$ km) in each Monte-Carlo run. The CCDF curves of the absolute distance errors for random ranges using $g_0 = 10$ kHz (the dashdot lines) and $g_3 = 40$ kHz (the dotted lines) are also illustrated in Fig. 5, respectively. The curves maintain the same tendency as the ones with fixed ranges. Less than 8% absolute errors are larger than 1 m at $\mathrm{SNR} = 60$ dB for both frequencies.

As we observed in the previous simulation results, the deep fading channel may fail the ranging procedure. Under such circumstances, the targets can explore the frequency diversity to switch the frequency allocation in order to avoid the deep fading effect. Therefore, in the following simulations, the median absolute error (MAE) is used as a performance metric, which is robust to performance outliers under deep fading channels. To investigate the influence of the uncertainties of LOs, random frequency discrepancies are added to $\mathbf{g}$ in each run in the following simulations. The frequency discrepancies are uniformly distributed in the range of $[-\zeta\Delta g, \zeta\Delta g]$, where $\zeta \in \{1000\ \mathrm{ppm}, 100\ \mathrm{ppm}, 10\ \mathrm{ppm}\}$ ($1\ \mathrm{ppm} = 10^{-6}$). The clock skew effect due to the target transmitters or anchor receivers can be mimicked by the frequency discrepancies. In this case, it is impossible for anchors to maintain accurate knowledge of $\mathbf{g}$. Therefore, $\mathbf{g}$ is first estimated by the ESPRIT-type algorithm proposed in Section V, and then the range estimation employs the estimated $\mathbf{g}$ similarly as *Example 1*. However,

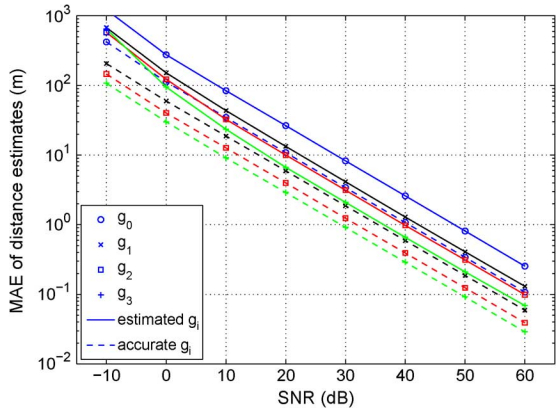


Fig. 6. A comparison of MAEs of distance estimates using the estimated $\mathbf{g}$ and the accurate $\mathbf{g}$ versus SNR with frequency discrepancies $\zeta\Delta g = 10$ Hz.

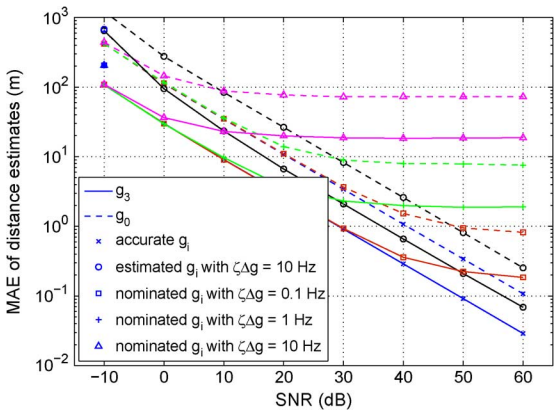


Fig. 7. A comparison of MAEs of distance estimates versus SNR with frequency discrepancies $\zeta\Delta g \in \{0.1\text{ Hz}, 1\text{ Hz}, 10\text{ Hz}\}$, using the estimated, the accurate, and the nominated g_i, respectively.

no frequency discrepancies are employed in *Example 1*. To evaluate the ranging performance, the ranging accuracy using the exact $\mathbf{g}$ (unavailable to receivers) is used as the benchmark in the following comparisons.

The MAE of the distance estimates with $\zeta\Delta g = 10$ Hz is illustrated in Fig. 6. Using the estimated $\mathbf{g}$, the ranging performance can still achieve high accuracy at high SNR. The performance gaps between the benchmarks (the dashed lines) and the ones using the estimated $\mathbf{g}$ (the solid lines) are about 10 dB. The frequency estimation error increases the ranging error.

In Fig. 7, when the frequency discrepancy is neglected, the anchor still applies the nominated $\mathbf{g}$ ($[10, 20, 30, 40]$ kHz) to accomplish ranging, the ranging performance degrades. Furthermore, a larger ζ causes worse ranging performance. With increasing $\zeta\Delta g$, the error floors using the nominated g_i increase. When $\zeta\Delta g = 10$ Hz, the ranging performance using the nominated g_i is unacceptable. When $\zeta\Delta g = 0.1$ Hz, the error floors using the nominated g_i are around 1 m. Thus, the DRIPS is robust to a frequency discrepancy up to 0.1 Hz. Beyond this range, frequency estimation of $\mathbf{g}$ should be carried out. On the other hand, the ranging performance using the estimated g_i still achieves high accuracy even with $\zeta\Delta g = 10$ Hz. The ranging performance gaps between the ones using the estimated and the accurate g_i in Fig. 7 are due to the frequency estimation errors as explained before.

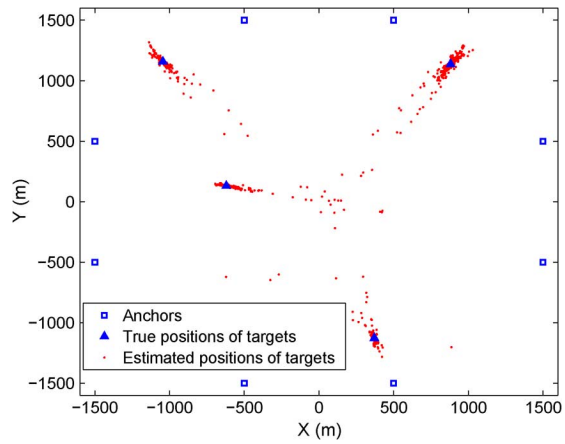


Fig. 8. An example of the estimated target positions using the ALS algorithm at $\mathrm{SNR} = 40$ dB for a fixed geometry.

B. Asynchronous Targets to Evaluate Localization Performance

In this subsection, the localization method proposed in Section IV is evaluated under the scenario, where all the targets are asynchronous, but all the anchors are synchronized. Thus, $t_k \neq 0, k \in \mathcal{K}$. The true transmission time instants $t_k, k \in \mathcal{K}$ are randomly generated in each Monte-Carlo run in a way such that φ_k ($\varphi_k = -2\pi g_k t_k$) is uniformly distributed in the range of $[2\pi/3, 2\pi]$. Note that the integer ambiguity has to be taken into account. Although the true distances are limited in the resolvable range, the integer ambiguity still arises due to the unknown transmission time instants t_k of the asynchronous targets. Flat-fading channels are employed as well. Eight anchors ($M = 8$ and $\mathcal{M} = \{0, 1, \ldots, 7\}$) are used and equally distributed on the edge of a 3 km $\times$ 3 km rectangular as shown in Fig. 8. The four targets are randomly distributed on the 1 m $\times$ 1 m grids inside a smaller rectangular of 2 km $\times$ 2 km.

Fig. 8 illustrates an example of the localization performance using the ALS algorithm at $\mathrm{SNR} = 40$ dB for a fixed geometry over 1,000 Monte Carlo runs. Accurate knowledge of frequencies is assumed here. According to Fig. 8, most position estimates are around the true target positions, while a few have larger estimation errors due to deep fading channels.

The MAEs of location estimates using g_2 and g_3 are shown in Fig. 9, which illustrates the performance comparison among the proposed ALS approach taking the integer ambiguity into account, the subtraction method by making phase difference that does not consider the integer ambiguity, and the localization method of the tailored RIPS (see Appendix C). Moreover, to investigate the influence of the frequency estimation on the localization performance, the maximum frequency discrepancy $\zeta\Delta g$ is assigned as 10 Hz. With accurate knowledge of g_i, only the ALS method (the solid lines) can achieve decent positioning accuracy, and the iteration number for the ALS method is less than 10. The subtraction method (the dotted lines), which eliminates $\widetilde{\varphi}_k$ due to the effect of t_k by making phase differences $\widetilde{\psi}_{l,k} - \widetilde{\psi}_{m,k}$, always reaches a high error floor, as the integer ambiguity is not resolved. Furthermore, the positioning accuracy of the ALS method with the estimated g_i (the dashed lines) degrades compared to the one with the accurate g_i, but still

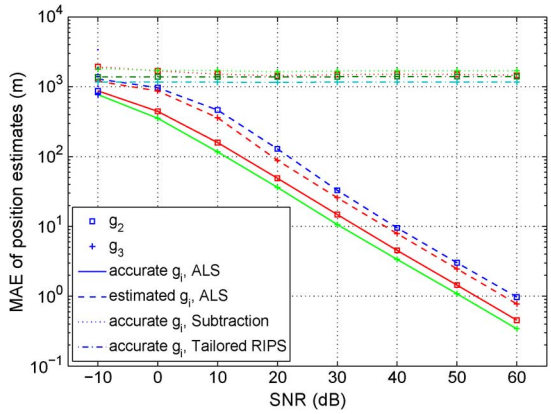


Fig. 9. A comparison of MAEs of the position estimate vs. SNR with frequency discrepancies $\zeta\Delta g = 10$ Hz, using the ALS, the subtraction method, and the tailored RIPS.

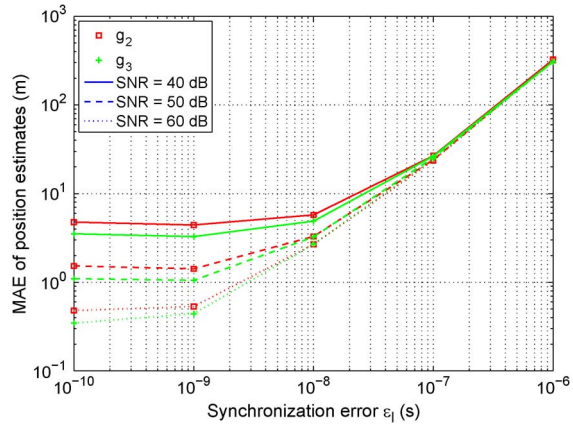


Fig. 10. The impact of anchor synchronization errors on MAEs of the position estimates.

achieves high accuracy at high SNR. Thus, the DRIPS is robust to the flat-fading channel, and can achieve good performance. On the other hands, the localization performance of the tailored RIPS (the dashdot lines) under flat-fading channels is unacceptable as shown in Fig. 9. Besides the eight anchors, a reference anchor placed at the origin is added for the tailored RIPS to transmit a single tone at frequency f_c. In order to avoid the integer ambiguity, f_c is set to 10 kHz. Four targets emit four tones at $f_c + \mathbf{g}$ ($[20, 30, 40, 50]$ kHz), respectively. The conventional LS estimator based on TDOAs [29] is used to estimate the locations in the tailored RIPS. In Appendix C, the reason why the tailored RIPS fails under flat fading channels is explained in detail. The unknown channel coefficients corrupt the range estimation. Therefore, the tailored RIPS also has a high error floor.

Furthermore, we investigate the impact of anchor synchronization errors on localization performance. In each Monte-Carlo run, a synchronization error is randomly generated for each anchor in the range of $[0, \varepsilon_l]$ following a uniform distribution. We assume accurate knowledge of frequencies. The MAE of position estimates using g_2 and g_3 are indicated in Fig. 10. As shown in Fig. 10, the influence of the synchronization errors within a few nanoseconds (ns) can be neglected. However, the localization performance degrades significantly, when the synchronization errors are larger than several tens of ns. The higher the SNR is, the severer the synchronization errors impact. Therefore, for a sub-meter localization accuracy, the synchronization errors should be less than 10 ns. The results shown in Fig. 10 are consistent with the discussions in *Remark 8*.

VIII. Conclusions

This paper has proposed a dual-tone radio interferometric positioning system (DRIPS) for a multi-target scenario. In DRIPS, multiple asynchronous targets transmit dual-tone signals, which are well separated from each other, and whose two tones are very close in frequency. At the anchor receiver, these dual-tone signal can create low-frequency differential tones, whose phases bear the range information. Due to the unknown initial time instants of the targets, the biased time-of-arrival (TOA) estimates are achieved by extracting the phase information of the differential tones. Accordingly, low-complexity localization algorithms are developed with the help of synchronous anchors using the biased TOAs to locate the targets in the resolvable range. The proposed DRIPS is robust to flat-fading channels, and able to localize multiple targets simultaneously. The carrier frequency offsets (CFOs) and random initial phases of the local oscillators are also well accommodated in the DRIPS. Furthermore, when accurate knowledge of the tone frequencies is unavailable, the ESPRIT-type algorithm is proposed to estimate the frequencies. Moreover, the Cramér-Rao bounds (CRBs) are derived to show the performance limit. Simulation results corroborate the efficiency of the proposed DRIPS.

Appendix A
Proof of Theorem 1

Proof: Let us substitute (4) into the cross-correlation terms $x_{l,k}(t)x^*_{l,q}(t)$. Thus, we achieve

$$x_{l,k}(t)x^*_{l,q}(t) = \alpha_{l,k}\alpha^*_{l,q}\beta_{l,k}\beta^*_{l,q}e^{j2\pi(f_k-f_q)t+jf_k\frac{\psi_{l,k}}{g_k}-jf_q\frac{\psi_{l,q}}{g_q}} \times \left(1 + e^{j2\pi(g_k-g_q)t+j\psi_{l,k}-j\psi_{l,q}} + e^{-j2\pi g_q t-j\psi_{l,q}} + e^{j2\pi g_k t+j\psi_{l,k}}\right).$$

Since $\beta_{l,k}$ in general is different from $\beta_{l,q}$, the unknown flat-fading channels may change the phase of $x_{l,k}(t)x^*_{l,q}(t)$. Moreover, $\alpha_{l,k}$ is also different from $\alpha_{l,q}$ due to the randomness of the LOs, and thus the term $\alpha_{l,k}\alpha^*_{l,q}$ impacts on the phase of $x_{l,k}(t)x^*_{l,q}(t)$ as well. According to these unknown parameters, the delay information cannot be extracted correctly from $x_{l,k}(t)x^*_{l,q}(t)$. If $g_{\max} << \min\{|f_k - f_q|, |f_k - f_q + g_k|, |f_k - f_q - g_q|, |f_k - f_q + g_k - g_q|\}, \forall k, \forall q \in \mathcal{K}$ and $k \neq q$, we have $x_{l,k}(t)x^*_{l,q}(t)$ in a different frequency band from $|x_{l,k}(t)|^2$. An LPF can eliminate the nuisance cross terms and only maintain the low-frequency differential tones of $|x_{l,k}(t)|^2$. ■

Appendix B
Proof of Theorem 2

Proof: Suppose that $g_k = g_q$. As a result, the products $|x_{l,k}(t)|^2$ and $|x_{l,q}(t)|^2$ create the low-frequency differential tones with the same frequencies $\pm g_k$. We achieve

$$|x_{l,k}(t)|^2 + |x_{l,q}(t)|^2 = 2\left(|\alpha_{l,k}\beta_{l,k}|^2 + |\alpha_{l,q}\beta_{l,q}|^2\right) + 2\mathrm{Re}\left\{e^{j2\pi g_k t}\left(|\alpha_{l,k}\beta_{l,k}|^2 e^{j\psi_{l,k}} + |\alpha_{l,q}\beta_{l,q}|^2 e^{j\psi_{l,q}}\right)\right\},$$

where the phases are related to the unknown coefficients and unknown delays. Under such a circumstance, the delay information cannot be decoupled from the channel effect and uncertainties of LOs. We cannot extract the delay information for the kth and qth targets as well, since their differential tones interfere with each other. Therefore, $g_k \neq g_q$ in order to avoid interferences between the autocorrelation terms, and derive the delay information of the targets. ■

APPENDIX C
THE TAILORED RIPS

Based on the original RIPS [12], the tailored RIPS involves three anchors and one target in a ranging session. One of the anchors and the target transmit tones at slightly different frequencies. The anchor that transmits is called the reference anchor, and is used for all the other ranging sessions as well. Therefore, for the next ranging session, another two anchors will receive the tones emitted by the target and the reference anchor. Without loss of generality, the kth target and reference anchor transmit tones, and the received composite signal at the lth anchor is given by

$$\widetilde{r}_l(t) = a_k\beta_{l,k}e^{j\phi_k}e^{j2\pi(f_c+g_k)(t-t_k-\tau_{l,k})} + a\beta_l e^{j\phi}e^{j2\pi f_c(t-\widetilde{t}_0-\tau_l)}, \forall l \in \mathcal{M}, \quad (50)$$

where β_l is the complex channel coefficient between the lth anchor and the reference anchor, a is the real-valued amplitude of the tone transmitted by the reference anchor, ϕ is the random initial phase of the reference anchor, $\widetilde{t}_0$ is the known transmission start instant of the reference anchor (as anchors are synchronized), and τ_l is the known propagation delay between the reference and the lth anchor. The rest parameters are just the same as defined before.

The received signal $\widetilde{r}_l(t)$ is processed by a square law device ($|\widetilde{r}_l(t)|^2$), removed the DC component and sampled. The resulted sample vector $\widetilde{\mathbf{r}}_l$ collects the samples of the low-frequency differential tones created by squaring, and can be modeled as

$$\widetilde{\mathbf{r}}_l = [\,\boldsymbol{\phi}_N(g_k) \quad \boldsymbol{\phi}_N^*(g_k)\,]\begin{bmatrix}\widetilde{\gamma}_l \\ \widetilde{\gamma}_l^*\end{bmatrix}, \quad (51)$$

where $\widetilde{\gamma}_l = a_k a\beta_{l,k}\beta_l^* e^{j(\phi_k-\phi)-j2\pi((f_c+g_k)(t_k+\tau_{l,k})-(\widetilde{t}_0+\tau_l)f_c)}$. Thus, an LS estimate of $\widetilde{\gamma}_l$ can be obtained based on (51). The initial phase offsets is canceled out as follows

$$\widetilde{\gamma}_l\widetilde{\gamma}_q^* = a_k^2 a^2\beta_{l,k}\beta_l^*\beta_{q,k}^*\beta_q e^{j2\pi(f_c(\tau_l-\tau_q)+(f_c+g_k)(\tau_{q,k}-\tau_{l,k}))}, \forall l, \forall q \in \mathcal{M}, \quad (52)$$

where τ_l and τ_q are known, and $\tau_{q,k}-\tau_{l,k}$ is the TDOA of interest. However, due to the unknown channel coefficients, it is impossible to correctly recover the TDOA based on (52). Therefore, the tailored RIPS always fails under the fading channel.

APPENDIX D
DERIVATION OF THE STOCHASTIC PROPERTIES OF $\widetilde{\mathbf{m}}$

Note that $\widetilde{\mathbf{m}}$ is the Nyquist sample vector of $\widetilde{m}_l(t)$, where $\widetilde{m}_l(t)$ is the low-pass filtered result of $m_l(t)$. In this appendix, the subscript $_l$ is employed again for the convenience of the derivation. We first derive the stochastic properties of $m_l(t)$ and $\widetilde{m}_l(t)$, respectively, and then the stochastic properties of $\widetilde{\mathbf{m}}$ are calculated accordingly.

Let us recall (4) and (38) here

$$x_{l,k}(t) = \alpha_{l,k}\beta_{l,k}e^{j2\pi((f_k-f_l)t-f_k(t_k+\tau_{l,k}))} \times \left(1+e^{j2\pi g_k(t-t_k-\tau_{l,k})}\right), \quad (53)$$

$$m_l(t) = w_l^*(t)\sum_{k=0}^{K-1}x_{l,k}(t) + w_l(t)\sum_{k=0}^{K-1}x_{l,k}^*(t) + |w_l(t)|^2, \quad (54)$$

where $\beta_{l,k} \sim \mathcal{CN}(0,\sigma_{l,k}^2)$ and $w_l(t)$ is a zero-mean bandlimited complex Gaussian random process with the autocorrelation function (ACF) $R_B(\tau) = N_0 B\sin(\pi B\tau)/(\pi B\tau)$ with N_0 the PSD and B the bandwidth of the receiver. Moreover, $\beta_{l,k}$ and $w_l(t)$ are independent. Hence, the expectation of $m_l(t)$ is calculated as

$$E\left[m_l(t)\right] = E\left[|w_l(t)|^2\right] = R_B(0) = N_0 B. \quad (55)$$

Furthermore, the ACF $R_{m_l}(t,\tau)$ ($E[m_l(t)m_l(t-\tau)]$) of $m_l(t)$ is computed as

$$R_{m_l}(t,\tau) = R_B^2(0) + 4R_B(\tau)\sum_{k=0}^{K-1}|\alpha_{l,k}|^2\sigma_{l,k}^2\cos\left(2\pi(f_k-f_l+g_k/2)\tau\right) \times\Big(\cos(\pi g_k\tau)+\cos(2\pi g_k t-\pi g_k\tau+\psi_{l,k})\Big). \quad (56)$$

Since $R_{m_l}(t,\tau)$ is a function of t and τ, $m_l(t)$ is nonstationary. However, $m_l(t)$ is cyclostationary (see [37]), as $R_{m_l}(t,\tau) = R_{m_l}(t+nT_p,\tau), n \in \mathcal{Z}$, where $T_p = \mathrm{LCM}\left(1/g_0,\ldots,1/g_{K-1}\right)$ and $\mathrm{LCM}(\cdot)$ is the least common multiple. Consequentially, the spectrum correlation density (SCD) function $S_{m_l}(\alpha,f)$ of $m_l(t)$ is the Fourier transform of $R_{m_l}(t,\tau)$ w.r.t. t and τ.

When $m_l(t)$ is filtered by an ideal LPF, whose frequency response is $H(f) = \mathrm{rect}(f/B_L)$ with $B_L/2$ the bandwidth of the LPF, the output is $\widetilde{m}_l(t)$. Therefore, the SCD function $S_{\widetilde{m}_l}(\alpha,f)$ of $\widetilde{m}_l(t)$ is related to the one of $m_l(t)$ as [37]

$$S_{\widetilde{m}_l}(\alpha,f) = H(f+\alpha/2)H^*(f-\alpha/2)S_{m_l}(\alpha,f). \quad (57)$$

Recall that $B \geq \Delta f_{\max}$ and $B_L/2 > g_{\max}$, where $\Delta f_{\max}$ is defined in (18) and $g_{\max} = \max\{g_0,\ldots,g_{K-1}\}$. Moreover, we always have $B >> B_L$ and assume that $B - \Delta f_{\max} > g_{\max}$. As a result, the ACF $R_{\widetilde{m}_l}(t,\tau)$ of $\widetilde{m}_l(t)$ as the inverse Fourier transform of $S_{\widetilde{m}_l}(\alpha,f)$ w.r.t. α and f, respectively, can be calculated as

$$R_{\widetilde{m}_l}(t,\tau) = R_B^2(0) + 4\sum_{k=0}^{K-1}|\alpha_{l,k}|^2\sigma_{l,k}^2 \times\left(R_{B_L}(\tau)+R_{B_L-g_k}(\tau)\cos(2\pi g_k t+\psi_{l,k})\right). \quad (58)$$

As $R_{\widetilde{m}_l}(t,\tau)$ is still periodic w.r.t. t, the cyclostationarity is reserved in $\widetilde{m}_l(t)$.

Moreover, the expectation of $\widetilde{m}_l(t)$ is

$$E\left[\widetilde{m}_l(t)\right] = E\left[m_l(t)\right]H(0) = R_B(0) = N_0 B, \quad (59)$$

which is a constant. Consequently, the covariance $C_{\widetilde{m}_l}(t,\tau)$ of $\widetilde{m}_l(t)$ is given by

$$\begin{aligned} C_{\widetilde{m}_l}(t,\tau) &= R_{\widetilde{m}_l}(t,\tau) - E\left[\widetilde{m}_l(t)\right] E\left[\widetilde{m}_l(t-\tau)\right] \\ &= 4\sum_{k=0}^{K-1} |\alpha_{l,k}|^2 \sigma_{l,k}^2 \\ &\quad \times \left(R_{B_L}(\tau) + R_{B_L - g_k}(\tau)\cos(2\pi g_k t + \psi_{l,k})\right). \end{aligned} \tag{60}$$

When the difference between $R_{B_L}(\tau)$ and $R_{B_L-g_k}(\tau)$ is negligible, and $\sum_{k=0}^{K-1} |\alpha_{l,k}|^2\sigma_{l,k}^2 \cos(2\pi g_k t + \psi_{l,k}) \approx 0$ due to the random unknown $\psi_{l,k}$ and a large K, we further achieve the approximation of the covariance of $\widetilde{m}_l(t)$ as

$$C'_{\widetilde{m}_l}(\tau) = 4R_{B_L}(\tau)\sum_{k=0}^{K-1} |\alpha_{l,k}|^2\sigma_{l,k}^2, \tag{61}$$

which is only related to τ.

Note that $\widetilde{\mathbf{m}}$ is obtained by sampling $\widetilde{m}_l(t)$ at the Nyquist rate $(1/B_L)$. We assume that $R_{B_L}(\tau) \approx 0$ and $R_{B_L-g_k}(\tau) \approx 0, \tau > 1/B_L$. As a result, the stochastic properties of $\widetilde{\mathbf{m}}$ are summarized in (43) and (44).

Appendix E
Computation of $\mathbf{I}_B(\boldsymbol{\theta})$

Since the Bayesian information matrix (BIM) $\mathbf{I}_B(\boldsymbol{\theta})$ is given by (40) as $\mathbf{I}_B(\boldsymbol{\theta}) = E_{\boldsymbol{\theta}}[\mathbf{I}_D(\boldsymbol{\theta})] + \mathbf{I}_P(\boldsymbol{\theta})$, we first calculate $\mathbf{I}_P(\boldsymbol{\theta})$. Recall that $\widetilde{\beta}_i = |\beta_i|^2$ and $\beta_i \sim \mathcal{CN}(0,\sigma_i^2), i \in \mathcal{K}$. Thus, every $2\widetilde{\beta}_k/\sigma_k^2$ follows an exponential distribution, and they are independent with each other. As a result, the log-likelihood function of $\widetilde{\boldsymbol{\beta}}$ removing constant terms is given by

$$\ln p(\widetilde{\boldsymbol{\beta}}) = -\sum_{k=0}^{K-1} \widetilde{\beta}_k/\sigma_k^2. \tag{62}$$

Taking (62) into (42), we can obtain $\mathbf{I}_P(\boldsymbol{\theta})$ in the following

$$\mathbf{I}_P(\boldsymbol{\theta}) = \begin{bmatrix} \mathbf{0}_{2K\times 2K} & \mathbf{0}_{2K\times K} \\ \mathbf{0}_{K\times 2K} & \operatorname{diag}\left(\boldsymbol{\sigma}_\beta \boldsymbol{\sigma}_\beta^T\right) - \boldsymbol{\sigma}_\beta\boldsymbol{\sigma}_\beta^T \end{bmatrix}, \tag{63}$$

where $\boldsymbol{\sigma}_\beta = \left[1/\sigma_0^2, 1/\sigma_1^2, \ldots, 1/\sigma_{K-1}^2\right]^T$.

Now let us calculate the information matrix of the data $E_{\boldsymbol{\theta}}[\mathbf{I}_D(\boldsymbol{\theta})]$. With the assumption that $\widetilde{\mathbf{m}}$ follows a Gaussian distribution and making use of the stochastic properties of $\widetilde{\mathbf{m}}$, we have

$$\mathbf{z} \sim \mathcal{N}\left(\boldsymbol{\mu}(\boldsymbol{\theta}), \mathbf{C}_{\widetilde{\mathbf{m}}}\right), \tag{64}$$

where $\boldsymbol{\mu}(\boldsymbol{\theta}) = \mathbf{h}(\boldsymbol{\theta}) + \varrho(\widetilde{\boldsymbol{\beta}})\mathbf{1}_{\widetilde{N}} + \boldsymbol{\mu}_{\widetilde{\mathbf{m}}}$, $[\mathbf{h}(\boldsymbol{\theta})]_{n+1} = 2\sum_{k=0}^{K-1} |\alpha_k|^2\widetilde{\beta}_k \cos(2\pi\widetilde{g}_k n + \psi_k)$, and $\varrho(\widetilde{\boldsymbol{\beta}}) = 2\sum_{k=0}^{K-1} |\alpha_k|^2\widetilde{\beta}_k$. Moreover, $\boldsymbol{\mu}_{\widetilde{\mathbf{m}}}$ and $\mathbf{C}_{\widetilde{\mathbf{m}}}$ are respectively the mean and the covariance of $\widetilde{\mathbf{m}}$, which are given by (43) and (44) in Section VI. Applying the CRB for the general Gaussian case [38], we arrive at

$$[\mathbf{I}_D(\boldsymbol{\theta})]_{i,j} = \left[\frac{\partial\boldsymbol{\mu}(\boldsymbol{\theta})}{\partial\theta_i}\right]^T \mathbf{C}_{\widetilde{\mathbf{m}}}^{-1}\left[\frac{\partial\boldsymbol{\mu}(\boldsymbol{\theta})}{\partial\theta_j}\right], \tag{65}$$

where

$$\left[\frac{\partial\boldsymbol{\mu}(\boldsymbol{\theta})}{\partial\widetilde{\beta}_i}\right]_{n+1} = 2|\alpha_i|^2\left(\cos(2\pi\widetilde{g}_i n + \psi_i) + 1\right), \tag{66}$$

$$\left[\frac{\partial\boldsymbol{\mu}(\boldsymbol{\theta})}{\partial\psi_i}\right]_{n+1} = -2|\alpha_i|^2\widetilde{\beta}_i \sin(2\pi\widetilde{g}_i n + \psi_i), \tag{67}$$

$$\left[\frac{\partial\boldsymbol{\mu}(\boldsymbol{\theta})}{\partial\widetilde{g}_i}\right]_{n+1} = -4\pi n|\alpha_i|^2\widetilde{\beta}_i \sin(2\pi\widetilde{g}_i n + \psi_i). \tag{68}$$

Taking all the above ingredients into the BIM $\mathbf{I}_B(\boldsymbol{\theta})$, we arrive at

$$\begin{bmatrix} \mathbf{B}_{\psi\psi} & \mathbf{B}_{\psi\widetilde{g}} & \mathbf{B}_{\psi\widetilde{\beta}} \\ \mathbf{B}_{\psi\widetilde{g}}^T & \mathbf{B}_{\widetilde{g}\widetilde{g}} & \mathbf{B}_{\widetilde{g}\widetilde{\beta}} \\ \mathbf{B}_{\psi\widetilde{\beta}}^T & \mathbf{B}_{\widetilde{g}\widetilde{\beta}}^T & \mathbf{B}_{\widetilde{\beta}\widetilde{\beta}} \end{bmatrix} \triangleq \mathbf{I}_B(\boldsymbol{\theta}), \tag{69}$$

where $\mathbf{B}_{\kappa\varsigma}$ is the submatrix related to κ and ς, with $\kappa,\varsigma \in \{\psi, \widetilde{g}, \widetilde{\beta}\}$. The CRB of $\boldsymbol{\psi}$ for the flat-fading channel is given by

$$\mathrm{CRB}(\psi_i) = \left[\mathbf{I}_B^{-1}(\boldsymbol{\theta})\right]_{i,i}. \tag{70}$$

Note that in (70), we treat $\widetilde{\mathbf{g}}$ as unknowns due to the randomness of LOs. Assuming accurate knowledge of $\widetilde{\mathbf{g}}$, the CRB of $\boldsymbol{\psi}$ can be simplified based on the block matrices inverse as

$$\mathrm{CRB}_s(\psi_i) = \left[\left(\mathbf{B}_{\psi\psi} - \mathbf{B}_{\psi\widetilde{\beta}}\mathbf{B}_{\widetilde{\beta}\widetilde{\beta}}^{-1}\mathbf{B}_{\psi\widetilde{\beta}}^T\right)^{-1}\right]_{i,i}. \tag{71}$$

In order to reduce the computational complexity, we make use of the approximated covariance matrix $\bar{\widetilde{\sigma}}^2\mathbf{I}_{\widetilde{N}}$ to calculate the CRBs, where $\bar{\widetilde{\sigma}}^2$ is given by (46) in Section VI. By the proper design of the bandwidth B_L of the LPF, $\widetilde{g}_k$ can be assumed not near 0 or 1/2. As a result, the BIM can be further simplified as

$$\mathbf{I}'_B(\boldsymbol{\theta}) \triangleq \frac{4\widetilde{N}}{\bar{\widetilde{\sigma}}^2}\begin{bmatrix} \operatorname{diag}(\mathbf{b}_{\psi\psi}) & \operatorname{diag}(\mathbf{b}_{\psi\widetilde{g}}) & \mathbf{0}_{K\times K} \\ \operatorname{diag}(\mathbf{b}_{\psi\widetilde{g}}) & \operatorname{diag}(\mathbf{b}_{\widetilde{g}\widetilde{g}}) & \mathbf{0}_{K\times K} \\ \mathbf{0}_{K\times K} & \mathbf{0}_{K\times K} & \mathbf{B}'_{\widetilde{\beta}\widetilde{\beta}} \end{bmatrix} \tag{72}$$

where $[\mathbf{b}_{\psi\psi}]_{i+1} = |\alpha_i|^4\sigma_i^4$, $[\mathbf{b}_{\psi\widetilde{g}}]_{i+1} = \pi(\widetilde{N}+1)|\alpha_i|^4\sigma_i^4$, $[\mathbf{b}_{\widetilde{g}\widetilde{g}}]_{i+1} = \frac{2}{3}\pi^2(\widetilde{N}+1)(2\widetilde{N}+1)|\alpha_i|^4\sigma_i^4$, and $\mathbf{B}'_{\widetilde{\beta}\widetilde{\beta}} = \boldsymbol{\alpha}\boldsymbol{\alpha}^T - \boldsymbol{\sigma}_\beta\boldsymbol{\sigma}_\beta^T + \operatorname{diag}\left(\frac{1}{2}\boldsymbol{\alpha}\boldsymbol{\alpha}^T + \boldsymbol{\sigma}_\beta\boldsymbol{\sigma}_\beta^T\right)$ with $\boldsymbol{\alpha} = \left[|\alpha_0|^2, |\alpha_1|^2, \ldots, |\alpha_{K-1}|^2\right]^T$. Accordingly, we achieve $\mathrm{CRB}'(\psi_i) = [(\mathbf{I}'_B(\boldsymbol{\theta}))^{-1}]_{i,i}$ as (47) and $\mathrm{CRB}'_s(\psi_i) = [(\operatorname{diag}(\mathbf{b}_{\psi\psi}) - \operatorname{diag}(\mathbf{b}_{\psi\widetilde{g}})\operatorname{diag}(\mathbf{b}_{\widetilde{g}\widetilde{g}})^{-1}\operatorname{diag}(\mathbf{b}_{\psi\widetilde{g}}))^{-1}]_{i,i}$ as (48) in Section VI.

References

[1] 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.

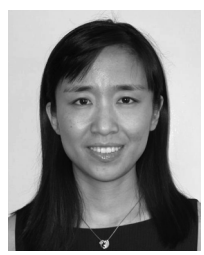
[2] G. Mao, B. Fidan, and B. Anderson, "Wireless sensor network localization techniques," *Comput. Netw.*, vol. 51, no. 10, pp. 2529–2553, Jul. 2007.

[3] P. Bahl and V. Padmanabhan, "RADAR: An in-building RF-based user location and tracking system," in *Proc. IEEE INFOCOM*, Tel Aviv, Israel, Mar. 2000, vol. 2, pp. 775–784.

[4] V. Seshadri, G. V. Zaruba, and M. Huber, "A Bayesian sampling approach to in-door localization of wireless devices using received signal strength indication," in *Proc. IEEE PerCom*, Kauai Island, HI, USA, Mar. 2005, pp. 75–84.
[5] Z. Zhong and T. He, "Achieving range-free localization beyond connectivity," in *Proc. ACM SenSys*, Berkeley, CA, USA, Nov. 2009, pp. 281–294.
[6] E. Elnahrawy, X. Li, and R. P. Martin, "The limits of localization using signal strength: a comparative study," in *Proc. IEEE SECON*, Santa Clara, CA, USA, Oct. 2004, pp. 406–414.
[7] K. Whitehouse, C. Karlof, and D. Culler, "A practical evaluation of radio signal strength for ranging-based localization," *ACM SIGMOBILE Mobile Comput. Commun. Rev.*, vol. 11, no. 1, pp. 41–52, Jan. 2007.
[8] I. Guvenc and C.-C. Chong, "A survey on TOA based wireless localization and NLOS mitigation techniques," *IEEE Commun. Surv. Tutorials*, vol. 11, no. 3, pp. 107–124, Aug. 2009.
[9] H. Wei, R. Peng, Q. Wan, Z. Chen, and S. Ye, "Multidimensional scaling analysis for passive moving target localization with TDOA and FDOA measurements," *IEEE Trans. Signal Process.*, vol. 58, no. 3, pp. 1677–1688, 2010.
[10] S. Gezici, Z. Tian, G. B. Giannakis, H. Kobayashi, A. F. Molisch, H. V. Poor, and Z. Sahinoglu, "Localization via ultra-wideband radios: A look at positioning aspects for future sensor networks," *IEEE Signal Process. Mag.*, vol. 22, pp. 70–84, Jul. 2005.
[11] S. Gezici and H. V. Poor, "Position estimation via ultra-wide-band signals," in *Proc. IEEE*, Feb. 2009, vol. 97, no. 2, pp. 386–403.
[12] M. Maróti, B. Kusy, G. Balogh, P. Völgyesi, A. Nádas, K. Molnár, S. Dóra, and Á. Lédeczi, "Radio interferometric geolocation," in *Proc. ACM SenSys*, San Diego, CA, USA, Nov. 2005, pp. 1–12.
[13] B. Kusy, J. Sallai, G. Balogh, A. Ledeczi, V. Protopopescu, J. Tolliver, F. DeNap, and M. Parang, "Radio interferometric tracking of mobile wireless nodes," in *Proc. ACM MobiSys*, San Juan, Puerto Rico, Jun. 2007, pp. 139–151.
[14] B. Kusy, A. Ledeczi, and X. Koutsoukos, "Tracking mobile nodes using RF Doppler shifts," in *Proc. ACM SenSys*, Sydney, Australia, Nov. 2007, pp. 29–42.
[15] H. Chang, J. Tian, T. Lai, H. Chu, and P. Huang, "Spinning beacons for precise indoor localization," in *Proc. ACM SenSys*, Raleigh, NC, USA, Nov. 2008, pp. 127–140.
[16] I. Amundson, J. Sallai, X. Koutsoukos, A. Lédeczi, and M. Maròti, "RF angle of arrival-based node localisation," *Int. J. Sens. Netw.*, vol. 9, no. 3, pp. 209–224, May 2011.
[17] J. Friedman, A. Davitian, D. Torres, D. Cabric, and M. Srivastava, "Angle-of-arrival-assisted relative interferometric localization using software defined radios," in *Proc. IEEE MILCOM*, Boston, MA, USA, Oct. 2009, pp. 1–8.
[18] S. Szilvasi, J. Sallai, I. Amundson, P. Volgyesi, and A. Ledeczi, "Configurable hardware-based radio interferometric node localization," in *Proc. IEEE Aerosp. Conf.*, Big Sky, MT, USA, Mar. 2010, pp. 1–10.
[19] B. J. Dil and P. J. M. Havinga, "Stochastic radio interferometric positioning in the 2.4 GHz range," in *Proc. ACM SenSys*, Seattle, WA, USA, Nov. 2011, pp. 108–120.
[20] Y. Wang, M. Shinotsuka, X. Ma, and M. Tao, "Design an asynchronous radio interferometric positioning system using dual-tone signaling," in *Proc. IEEE WCNC*, Shanghai, China, Apr. 2013, pp. 2294–2298.
[21] Y. Wang, L. Li, X. Ma, M. Shinotsuka, C. Chen, and X. Guan, "Dual-tone radio interferometric positioning systems using undersampling techniques," *IEEE Signal Process. Lett.*, vol. 21, no. 11, pp. 1311–1315, Nov. 2014.
[22] J. G. Proakis and M. Salehi, *Digital Communications*. New York, NY, USA: McGraw-Hill, 2008, vol. 3.
[23] C. K. Seow and S. Y. Tan, "Non-line-of-sight localization in multipath environments," *IEEE Trans. Mobile Comput.*, vol. 7, no. 5, pp. 647–660, May 2008.
[24] P. J. G. Teunissen, "The least-squares ambiguity decorrelation adjustment: a method for fast GPS integer ambiguity estimation," *J. Geodesy*, vol. 70, no. 1, pp. 65–82, 1995.
[25] C. Wang, Q. Yin, and Q. Wang, "An efficient ranging method for wireless sensor networks," in *Proc. IEEE ICASSP*, Dallas, TX, USA, Mar. 2010, pp. 2846–2849.
[26] P. Moreira, J. Serrano, T. Wlostowski, P. Loschmidt, and G. Gaderer, "White rabbit: Sub-nanosecond timing distribution over ethernet," in *Proc. IEEE ISPCS*, Brescia, Italy, Oct. 2009, pp. 1–5.
[27] White Rabbit Users [Online]. Available: http://www.ohwr.org/projects/white-rabbit
[28] R. Exel, "Clock synchronization in IEEE 802.11 wireless LANs using physical layer timestamps," in *Proc. IEEE ISPCS*, San Francisco, CA, USA, Sep. 2012, pp. 24–28.
[29] P. Stoica and J. Li, "Lecture notes-source localization from range-difference measurements," *IEEE Signal Process. Mag.*, vol. 23, no. 6, pp. 63–66, Nov. 2006.
[30] Y. Wang and G. Leus, "Reference-free time-based localization for an asynchronous target," *EURASIP J. Adv. Signal Process.*, vol. 19, pp. 1–21, Jan. 2012.
[31] R. v. Nee and R. Prasad, *OFDM for Wireless Multimedia Communications*, 1st ed. Norwood, MA, USA: Artech House, 2000.
[32] *LTE; Evolved Universal Terrestrial Radio Access (E-UTRA); User Equipment (UE) Radio Transmission And Reception*, 3GPP TS 36.101, 2011.
[33] Y. Takane, F. W. Young, and J. De Leeuw, "Nonmetric individual differences multidimensional scaling: An alternating least squares method with optimal scaling features," *Psychometrika*, vol. 42, no. 1, pp. 7–67, 1977.
[34] R. Roy and T. Kailath, "ESPRIT-estimation of signal parameters via rotational invariance techniques," *IEEE Trans. Acoust., Speech Signal Process.*, vol. 37, no. 7, pp. 984–995, Jul. 1989.
[35] A. J. Van der Veen, M. C. Vanderveen, and A. Paulraj, "Joint angle and delay estimation using shift-invariance techniques," *IEEE Trans. Signal Process.*, vol. 46, no. 2, pp. 405–418, 1998.
[36] H. L. Van Trees, *Detection, Estimation, Modulation Theory*. New York, NY, USA: Wiley, 1968.
[37] W. A. Gardner, "Exploitation of spectral redundancy in cyclostationary signals," *IEEE Signal Process. Mag.*, vol. 8, no. 2, pp. 14–36, 1991.
[38] S. M. Kay, *Fundamentals of Statistical Signal Processing: Vol. I—Estimation Theory*. Englewood Cliffs, NJ, USA: Prentice-Hall, 1993.

Yiyin Wang (M'11) received the B.S. degree in electrical engineering from Fudan University, Shanghai, China, in 2002, the M.S. degree (*cum laude*) in microelectronics from Delft University of Technology (TU Delft), The Netherlands, and Fudan University, China, in 2005, respectively, and the Ph.D. degree in electrical engineering from TU Delft, in 2011.

She is currently an Associate Professor with the Department of Automation, Shanghai Jiao Tong University, China. Prior to that, she was a Research Assistant at the Circuits and Systems group (CAS), TU Delft, from 2006 to 2007. From February 2010 until July 2010, she visited Georgia Institute of Technology (Georgia Tech), Atlanta. She was a Postdoctoral Fellow at TU Delft and then with Georgia Tech from 2011 to 2013. Her research interests lie in the general area of signal processing for communications and networking. She currently focuses on tracking, localization, and synchronization for wireless sensor networks.



Xiaoli Ma (SM'09) received the B.S. degree in automatic control from Tsinghua University, Beijing, China, in 1998, the M.S. degree in electrical engineering from the University of Virginia in 2000, and the Ph.D. degree in electrical engineering from the University of Minnesota in 2003.

After receiving the Ph.D. degree, she joined the Department of Electrical and Computer Engineering, Auburn University, where she served as an Assistant Professor until December 2005. Since January 2006, she has been with the School of Electrical and Computer Engineering, Georgia Institute of Technology (Georgia Tech), Atlanta, where she is currently a professor. Her research focuses on networking and communications, including network performance analysis, transceiver designs for wireless time- and frequency-selective channels, channel estimation and equalization algorithms, carrier frequency synchronization for OFDM systems, performance analysis, and cooperative designs for wireless networks.

Dr. Ma was awarded the Lockheed Martin Aeronautics Company Dean's Award for Teaching Excellence by the College of Engineering in 2009, and Outstanding Junior Faculty Award by the School of Electrical and Computer Engineering in 2010, at Georgia Tech. She serves as a Senior Area Editor for IEEE SIGNAL PROCESSING LETTERS since 2014 and ELSEVIER DIGITAL SIGNAL PROCESSING since June 2012, and has been an Associate Editor for the IEEE SIGNAL PROCESSING LETTERS (2007–2009) and the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS (2008–2013).

Cailian Chen (S'03–M'06) received the B.Eng. and M.Eng. degrees in automatic control from Yanshan University, China, in 2000 and 2002, respectively, and the Ph.D. degree in control and systems from City University of Hong Kong, Hong Kong SAR, in 2006.

She joined the Department of Automation, Shanghai Jiao Tong University, in 2008 as an Associate Professor. She is now a Full Professor. Her research interests include distributed estimation and control of network systems, wireless sensor and actuator networks, multiagent systems, and intelligent control systems. She has authored and/or coauthored two research monographs and more than 80 referred international journal and conference papers. She is the inventor of 20 patents.

Dr. Chen received the IEEE Transactions on Fuzzy Systems Outstanding Paper Award in 2008. She was one of the First Prize Winners of Natural Science Award from The Ministry of Education of China in 2007. She was honored as New Century Excellent Talents in University of China and Shanghai Rising Star in 2013, Shanghai Pujiang Scholar, Chenguang Scholar, and SMC Outstanding Young Staff of Shanghai Jiao Tong University in 2009. She serves as an Associate Editor of *Peer-to-peer Networking and Applications* (Springer), ISRN *Sensor Networks*, and *Scientific World Journal* (Computer Science), and TPC member of several flagship conferences including IEEE Globecom, IEEE ICC, and IEEE WCCI.

Xinping Guan (M'02–SM'04) received the Ph.D. degree in control and systems from Harbin Institute of Technology, Harbin, China, in 1999.

In 2007, he joined the Department of Automation, Shanghai Jiao Tong University, Shanghai, China, where he is currently a Distinguished University Professor, the Vice Director of University Research Management Office, and the Director of the Key Laboratory of Systems Control and Information Processing, Ministry of Education of China. Before that, he was the Professor and Dean of Electrical Engineering, Yanshan University, China. His current research interests include cyberphysical systems, multiagent systems, wireless networking and applications in smart city and smart factory, and underwater sensor networks. He has authored and/or coauthored four research monographs.

Dr. Guan has published more than 160 papers in IEEE TRANSACTIONS and other peer-reviewed journals, and numerous conference papers. As a Principal Investigator, he has finished/been working on many national key projects. He is the leader of the prestigious Innovative Research Team of the National Natural Science Foundation of China (NSFC). He is an Executive Committee Member of Chinese Automation Association Council and the Chinese Artificial Intelligence Association Council. He is on the editorial board of the IEEE TRANSACTIONS ON SYSTEMS, MAN AND CYBERNETICS—PART C and several Chinese journals. He received the First Prize of Natural Science Award from the Ministry of Education of China in 2006 and the Second Prize of the National Natural Science Award of China in 2008. He was a recipient of the IEEE Transactions on Fuzzy Systems Outstanding Paper Award in 2008. He is a "National Outstanding Youth" honored by NSF of China, "Changjiang Scholar" by the Ministry of Education of China and "State-level Scholar" of "New Century Bai Qianwan Talent Program" of China.