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

$$x_{l,k}(t) = \beta_{l,k} s_k(t - \tau_{l,k}) e^{-j2\pi f_l t + j\eta_l} \quad (3)$$

$$= \alpha_{l,k}\beta_{l,k} e^{j2\pi((f_k - f_l)t - f_k(t_k + \tau_{l,k}))} \times \left(1 + e^{j2\pi g_k(t - t_k - \tau_{l,k})}\right), \quad (4)$$

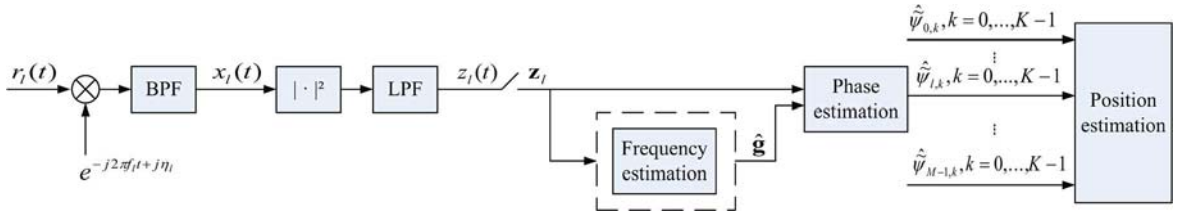

Fig. 1. The receiver structure of the DIRPS.

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

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

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

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

where the unknown $\lfloor\psi_{l,k}/2\pi\rfloor$ accounts for the integer ambiguity and $\hat{\widetilde{\psi}}_{l,k}$ 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

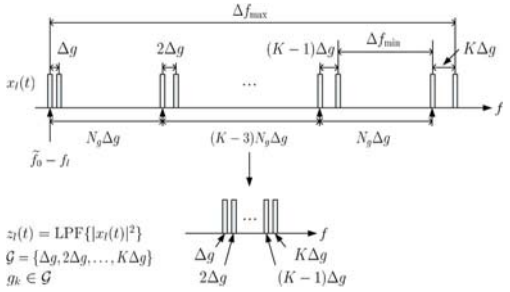

Fig. 2. An example of frequency allocation of DRIPS.

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, $\boldsymbol{\beta} = [\widetilde{\beta}_0, \widetilde{\beta}_1, \ldots, \widetilde{\beta}_{K-1}]^T$ with $\widetilde{\beta}_k = |\beta_k|^2$, and