Blind High-Resolution Localization and Tracking of Multiple Frequency Hopped Signals

Xiangqian Liu, *Student Member, IEEE*, Nicholas D. Sidiropoulos, *Senior Member, IEEE*, and Ananthram Swami, *Senior Member, IEEE*

***Abstract*—This paper considers the problem of blind localization and tracking of multiple frequency-hopped spread-spectrum signals using a uniform linear antenna array without knowledge of hopping patterns or directions of arrival. As a preprocessing step, we propose to identify a hop-free subset of data by discarding high-entropy spectral slices from the spectrogram. High-resolution localization is then achieved via either quadrilinear regression of four-way data generated by capitalizing on both spatial and temporal shift invariance or a new maximum likelihood (ML)-based two-dimensional (2-D) harmonic retrieval algorithm. The latter option achieves the best-known model identifiability bound while remaining close to the Cramér–Rao bound even at low signal-to-noise ratios (SNRs). Following beamforming using the recovered directions, a dynamic programming approach is developed for joint ML estimation of signal frequencies and hop instants in single-user tracking. The efficacy of the proposed algorithms is illustrated in pertinent simulations.**

***Index Terms*—Array signal processing, direction-of-arrival (DOA) estimation, frequency estimation, frequency hopping, harmonic analysis.**

I. Introduction

FREQUENCY-HOPPED code-division multiple access (FH-CDMA) is an appealing spread spectrum technique in wireless communication because frequency hopping provides resistance to multiple-access interference without requiring stringent power control to alleviate the near-far problem, as required for direct-sequence CDMA [23]. Frequency-hopped spread spectrum (FHSS) has recently been adopted in two commercial standards: IEEE 802.11 (Wireless LAN) and Bluetooth (Wireless PAN). It is also the prevailing spread-spectrum technique in military communications [22], largely due to its robustness to jamming coupled with low probability of intercept/detection (LPI/LPD) and good near–far properties. Blind separation and localization of FHSS signals is a challenging problem that feeds into multiple facets of military communications, from interception of noncooperative communications to jammer localization and mitigation. On the technical side, the problem is challenging not only because hopping patterns and directions of arrival are unknown, but other parameters, such as bin-width, hop-rate, and timing, are at least partially unknown in a realistic scenario. Carrier hopping means that one has to deal with switching exponentials, rather than pure exponentials, and it also induces hopping in the receive antenna array spatial steering vectors, due to wavelength-dependent phase shifting from one array element to another.

Although it is possible to write down a full-blown model for the observed data, which accounts for the known signal structure plus all unknown parameters, and then aim for joint parameter estimation, this approach does not easily lead to conceptual analysis or efficient computation. Mixtures of switching exponentials are far more difficult to deal with than mixtures of pure exponentials and much less studied. The presence of uncertainties and nuisance parameters further aggravates this problem, as the model effectively exhibits less structure. Finally, carrier hopping implies a hybrid continuous-discrete joint estimation and detection problem that is not amenable to a computationally efficient solution.

Blind FHSS methods proposed so far in the literature fall under two broad categories. One general approach consists of employing a coarsely channelized spectrogram-like receiver filterbank, coupled with low-resolution tracking [1], [21]. Due to adjacent-bin leakage, coarse channelization is problematic in the presence of bin uncertainty, especially for co-channel users employing different hopping schemes and/or in the presence of Doppler shift. Furthermore, *multiuser* tracking in the spectrogram domain is computationally very intensive (exponential in the number of users) while only providing low-resolution frequency and hop instant estimates [1].

Nonparametric spectrogram-based approaches to FHSS analysis are useful as a first step toward more refined solutions. More sophisticated model-based approaches have also been developed [11], [26]. These employ a receive antenna array and a parametric data model. Assuming that a (mostly) hop-free snapshot of spatio-temporal data is available, these methods decouple the estimation problem from hop detection. For unstructured receive antenna arrays, the remaining problem is a multichannel version of the classical one-dimensional (1-D) harmonic retrieval problem that is mathematically equivalent to the popular uniform linear array (ULA) sensor array processing model, except that Vandermonde structure is present in the temporal, rather than the spatial dimension. For structured receive antenna arrays, such as ULAs, further parameterization is possible. Under the usual narrowband far-field scenario, a ULA induces Vandermonde structure in the spatial dimension as well,

Manuscript received July 5, 2001; revised December 10, 2001. This work was supported by subcontract participation in DARPA/ATO under Contract MDA 972-01-0056, the ARL Communications and Networks CTA, and the National Science Foundation/Wireless 0096164. The associate editor coordinating the review of this paper and approving it for publication was Dr. Naofal M. W. Al-Dhahir.

X. Liu and N. D. Sidiropoulos are with the Department of Electrical and Computer Engineering, University of Minnesota, Minneapolis, MN 55455 USA (e-mail: xiang@ece.umn.edu; nikos@ece.umn.edu).

A. Swami is with the Army Research Laboratory, Adelphi, MD 20783 USA (e-mail: a.swami@ieee.org).

Publisher Item Identifier S 1053-587X(02)02382-6.

thereby yielding a 2-D harmonic retrieval (2-D HR) problem. Either way, the solution is typically provided by some variation of ESPRIT [11], [26].

Parametric approaches can provide high-resolution localization and frequency estimates. Localization parameters (directions of arrival, or, more generally, steering vectors) are quasi-stationary and can be used for spatial beamforming. Given sufficient spatial degrees of freedom, beamforming can effectively suppress interference, enabling computationally simpler single user tracking. The parametric approaches proposed to date [11], [26] have the following drawbacks.

- The hop detection step has not been given proper attention. Wong [26] proposed the use of rank detection criteria to estimate hop intervals. However, rank detection is difficult, especially at low-to-moderate SNR, and computationally complex because it entails eigenvalue decomposition. The algorithm of [26] requires rough synchronization with the desired signal's hop interval. Lemma *et al.* [11] simply treats hopping as unmodeled dynamics.
- Identifiability is either limited [26] or not fully investigated [11]. For example, the maximum number of source signals that can be resolved by the algorithm proposed in [26] is less than six.
- ESPRIT-based algebraic algorithms are suboptimal, performance-wise.
- Single user tracking has not been addressed.

We propose herein a simple nonparametric hop-detection method using spectrogram entropy analysis so that a hop-free subset of data can be identified for direction-of-arrival (DOA) estimation purposes. Assuming a ULA receiver and a far-field scenario, a hop-free data snapshot can be modeled as a 2-D harmonic mixture. Through appropriate spatial-temporal smoothing, the 2-D harmonic mixture model can be transformed into a quadrilinear model, and quadrilinear alternating least squares (QALS) [14] can be employed to recover the DOAs and frequencies. QALS will be shown to outperform earlier ESPRIT-based methods (exemplified by JAFE [11]) while exhibiting robustness to unmodeled dynamics caused by undetected hops.

After the DOAs have been recovered, single-user tracking amounts to the joint estimation of hop instants, frequencies, and phases of a source signal from a desired DOA over a time interval of interest. Seeking a high-resolution solution, we develop a dynamic programming method following MMSE beamforming of mixture data. In [6] and [16], dynamic programming has been used to construct maximum likelihood sequence estimators for tracking frequencies. However, [6] and [16] assume known frequency grid, hop timing, and hop period. Hence, they are not applicable in the asynchronous user tracking situation considered herein, where hop timing and period are generally unknown, and the frequencies are not necessarily located on a grid, e.g., due to carrier frequency offset or Doppler shift.

In addition to QALS, maximum likelihood (ML)-based 2-D HR is also considered in this paper. Motivated by a recent stochastic identifiability result regarding 2-D HR [7], we develop a novel 2-D HR algorithm, called the multidimensional embedding—alternating least squares (MDE-ALS) algorithm, which is shown to achieve the identifiability bound in [7] and remain close to the Cramér-Rao bound (CRB) even at low SNR. While a variety of techniques have been developed for 2-D HR, e.g., [2], [5], and [13], they do not achieve the identifiability bound given in [7] (see also Section III-B) and can only approach the CRB at relatively high SNR.

The main contributions of this paper can be summarized as follows:

- a simple nonparametric hop-detection method based on spectral entropy;
- a multilinear LS algorithm for localization, which is robust to unmodeled dynamics;
- optimal ML single-user tracking using a dynamic programming approach;
- improved ML-based 2-D HR algorithm with full identifiability, staying close to the CRB for a wide range of SNRs.

We remark that our development is geared toward slow FH (SFH) signals and FSK modulation, which is mostly the case encountered in current systems. Carrier shifts due to hopping or symbol modulation are treated as conceptually equivalent, albeit of different magnitude. This means that certain types of modulation (e.g., Gaussian FSK) that induce continuous (rather than instantaneous) frequency shifts are not directly amenable to our analysis, which requires that the signal between two hops is a pure exponential. Note, however, that certain kinds of CPM (e.g., MSK) can be handled by our approach.[1]

The rest of this paper is organized as follows. The FH signal model is introduced in Section II. Section III contains a derivation of the proposed blind localization algorithm, including the hop instant detection method. The issue of single-user tracking is addressed in Section IV. The novel 2-D harmonic retrieval algorithm is developed in Section V. Section VI presents simulation results. Conclusions are drawn in Section VII.

Some notation conventions that will be used in this paper follow.

$\mathbf{A}^T$	Transpose of $\mathbf{A}$.
$\mathbf{A}^H$	Conjugate transpose of $\mathbf{A}$.
$\mathbf{A}^\dagger$	Pseudo-inverse of $\mathbf{A}$.
$\mathbf{A}(i,f)$	(i,f)th element of $\mathbf{A}$.
$\mathbf{a}_i$	ith column of $\mathbf{A}$.
$\mathbf{A}^{(m)}$	Submatrix of $\mathbf{A}$ formed by its first m rows.
$\mathbf{D}_i(\mathbf{A})$	Diagonal matrix constructed from the ith row of $\mathbf{A}$.
$\mathbf{A} \odot \mathbf{B}$	Khatri-Rao (column-wise Kronecker) product of $\mathbf{A}$ and $\mathbf{B}$.
$\|\cdot\|_F$	Frobenius norm.

II. Data Modeling

A schematic of the FHSS communication scenario under consideration is shown in Fig. 1. A total of d far-field frequency-hopped signals impinge on a ULA of M antennas, each from a nominal DOA with negligible angle spread.

The baseline separation of the ULA is δ. The array steering vector in response to a signal from DOA α can be written as

$$\mathbf{a}(\theta) = [1 \quad \theta \quad \cdots \quad \theta^{M-1}]^T, \quad \theta = e^{j2\pi\Delta\sin(\alpha)}$$

[1]With suitable frequency spacing, FSK modulation can yield continuous phase transitions at the symbol boundaries.

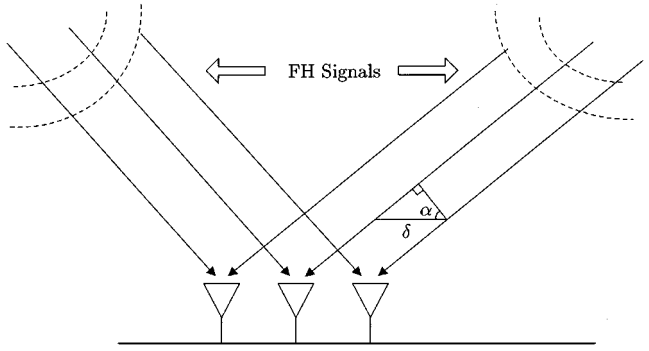


Fig. 1. Multiple far field FH signals impinge on a ULA array.

where $\Delta = \delta/\lambda$, and λ is the wavelength of the incoming signal. The $M \times 1$ signal vector collected from the ULA output at time t can be expressed as

$$\mathbf{x}(t) = \sum_{i=1}^{d} \mathbf{a}(\theta_i)\beta_i s_i(t)$$

where

$$s_i(t) = e^{j(2\pi[f_c + f_i(t)]t + \phi_i(t))}$$

and where β_i is the complex path loss for the ith source that collects the overall attenuation and propagation phase shift and may be frequency dependent. The sources are not necessarily synchronized and may have different symbol rates. f_c is the center frequency, which is assumed common for all sources. The ith signal's instantaneous frequency $f_i(t)$ and phase $\phi_i(t)$ are affected by two factors: frequency hopping and modulation. After down conversion and sampling at a rate of f_s, we collect K samples at each antenna output. For simplicity of exposition, suppose the dth signal hops from frequency $f_{d,x}$ to $f_{d,y}$ between sampling instant $q-1$ and q, whereas the other signal frequencies remain constant within the K sample block. Then, the discrete time baseband-equivalent model can be written as

$$\mathbf{X} = \mathbf{A}\mathbf{\Pi}\mathbf{\Phi}\mathbf{S}^T \tag{1}$$

where

$$\mathbf{A} = [\mathbf{a}(\theta_1) \quad \cdots \quad \mathbf{a}(\theta_{d-1}) \quad \mathbf{a}(\theta_{d,x}) \quad \mathbf{a}(\theta_{d,y})]$$

$$\mathbf{\Pi} = \mathrm{diag}(\beta_1 \quad \cdots \quad \beta_{d-1} \quad \beta_{d,x} \quad \beta_{d,y})$$

$$\mathbf{\Phi} = \mathrm{diag}(e^{j\phi_1} \quad \cdots \quad e^{j\phi_{d-1}} \quad e^{j\phi_{d,x}} \quad e^{j\phi_{d,y}})$$

$$\mathbf{S} = \begin{bmatrix} 1 & \gamma_1 & \cdots & \gamma_1^{q-1} & \gamma_1^{q} & \cdots & \gamma_1^{K-1} \\ & & \vdots & & & \vdots & \\ 1 & \gamma_{d-1} & \cdots & \gamma_{d-1}^{q-1} & \gamma_{d-1}^{q} & \cdots & \gamma_{d-1}^{K-1} \\ 1 & \gamma_{d,x} & \cdots & \gamma_{d,x}^{q-1} & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & \gamma_{d,y}^{q} & \cdots & \gamma_{d,y}^{K-1} \end{bmatrix}^T$$

with $\gamma_i = e^{j2\pi\omega_i}, \omega_i = 2\pi f_i/f_s$. For compactness, absorb $\mathbf{\Pi}$ and $\mathbf{\Phi}$ into $\mathbf{A}$, i.e., let $\mathcal{A} = \mathbf{A}\mathbf{\Pi}\mathbf{\Phi}$ so that (1) becomes

$$\mathbf{X} = \mathcal{A}\mathbf{S}^T. \tag{2}$$

For multiple hops per block, the model generalizes in the obvious fashion; our analysis is applicable to the case of multiple hops per data block. The purpose of blind localization and separation is to recover the DOAs, hop instants, and frequencies for each source, relying on model structure alone, without *a priori* knowledge of the hopping sequences.

III. Blind Localization Algorithm

A. Hop Instant Detection

Consider the spectrogram [15] of the data from the reference antenna output,[2] computed as the squared modulus of the short-term Fourier transform

$$\mathbf{X}_{\mathrm{SPEC}}(k+1, l+1) = \left| \sum_{m=0}^{N_1-1} \mathbf{X}(1, lL+m+1) w(m) e^{-j2\pi km/N_2} \right|^2 \tag{3}$$

for $k = 0, \ldots, N_2 - 1$ and $l = 0, \ldots, (K - N_1)/L$, where $N_2 \geq N_1 \geq L$. Equivalently, (3) means that one splits the signal into $((K - N_1)/L + 1)$ overlapping segments, windows each with a suitably chosen window sequence $w(m)$ defined in the region $0 \leq m \leq N_1 - 1$, and then calculates the DFT for each segment, and the number of frequency samples of the DFT is N_2. Thus, each column of $\mathbf{X}_{\mathrm{SPEC}}$ contains an estimate of the short-term, time-localized frequency content of the signal. Time increases linearly across the columns of $\mathbf{X}_{\mathrm{SPEC}}$, from left to right, whereas frequency increases linearly down the rows, starting at 0.

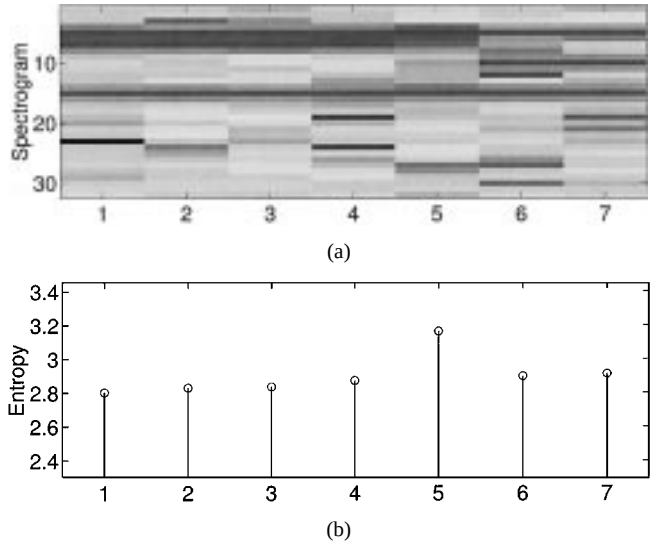
The basic idea for hop instant detection is that a vertical slice (i.e., a column) of the spectrogram corresponding to a given time segment can be viewed as a probability mass function after proper normalization, and slices containing hops exhibit high entropy due to instantaneous frequency spread.

The entropy of a discrete ensemble can be regarded as a measure of uncertainty, and any transfer of probability from one member of the ensemble to another that makes their probabilities more nearly equal will increase the entropy of the ensemble [4]. In a time window that a frequency hops from one bin to another, the signal energy from the frequency-hopped source will be distributed across several frequency bins: Hopping causes spectral spread that manifest itself as higher entropy, and a hop-free subset of data can be obtained by discarding the data corresponding to high-entropy spectral slices. This is illustrated in Fig. 2, where the received signals come from three sources, and only the third one hops during the data block. In the spectrogram plot, the horizontal axis is time, and the vertical axis is the frequency bin. At SNR $= 15$ dB,[3] a length $K = 128$ data sequence is split into seven overlapping segments, each windowed by a Hamming window of length $N_1 = 32$. The number of frequency samples is $N_2 = 32$. From the entropy plot, one can easily infer that the fifth window contains a frequency hop. Hence, the data corresponding to spectrogram windows one to four form a hop-free data subset, i.e., from the first sample to the 80th sample.

Notice that if there are many asynchronous FH signals and the dwell times are relatively short, detecting a hop-free subset may be difficult, and relative long hop-free subsets may not even exist. For long dwell times (for example, slow FH signals), this will not be a problem. In general, the entropy threshold should ideally be determined via decision-theoretic criteria. The

[2]Spectrograms could be averaged across receive antennas.

[3]The nominal model is $\mathbf{Y} = \mathbf{X} + \mathbf{\Xi}$, where $\mathbf{\Xi}(m, n)$ is the additive noise at the mth sensor at the nth sampling instant. Additive noise is assumed to be spatially and temporally white, with variance σ^2. SNR is defined as $10\log_{10}(\|\mathbf{X}\|_F^2/(MK\sigma^2))$; cf. (2).



(a)

(b)

Fig. 2. Hop instant detection via spectrogram entropy analysis. (a) Spectrogram of received signal. (b) Corresponding entropy plot. Parameters of three sources: DOA $= [5^\circ, 10^\circ, 15^\circ]$; frequency $= [3.5, 1.0, 1.5]$ MHz; the third source hops from 1.5 MHz to 2.25 MHz at normalized hoping time instant 0.64, and SNR is 15 dB.

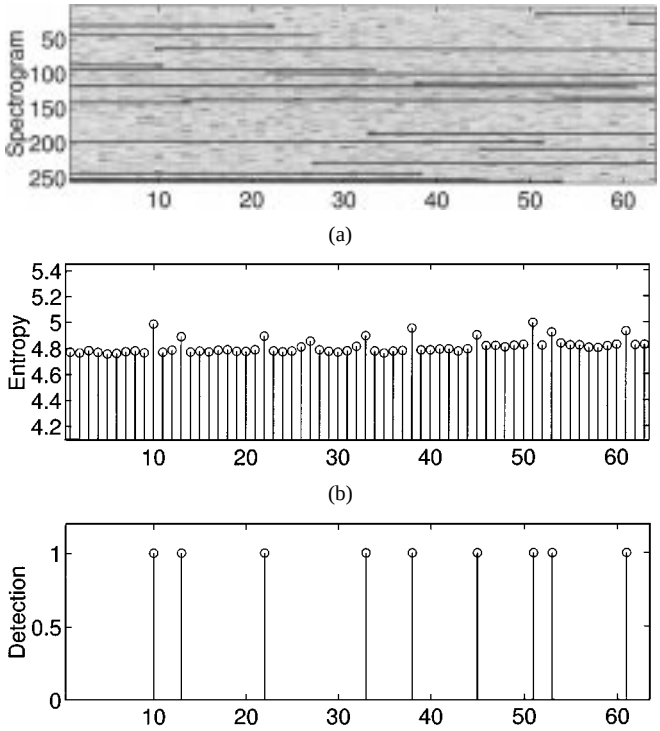
Neyman-Pearson approach seems best suited to the task. However, this requires conditional distributions or estimates thereof, which is quite unrealistic in a blind context. Unsupervised clustering is another possibility, but it requires long observation periods because in most cases, hop events are rare (relative to the null hypothesis). In view of these difficulties, we advocate a simple, pragmatic approach that works reasonably well: Use the average of minimum entropy and maximum entropy as a threshold. The overall strategy for detecting a suitable hop-free subset is to mark spectrogram slices with entropy above the threshold and then select the longest run of unmarked slices. Fig. 3 is an example that illustrates the process of hop detection under a scenario involving ten asynchronous FH sources, each hopping once during the data block. SNR is 15 dB. Using the average of minimum and maximum entropy as a threshold, Fig. 3(c) shows that nine hops are correctly detected. Monte Carlo simulation also shows that for the above experimental setup, using the average of minimum and maximum entropy as threshold yields 0.8 probability of detection at a low false alarm rate of 0.02. Coupled with the fact that, as we will see, localization algorithms are relatively robust to residual isolated hops, this yields a satisfactory overall solution.

From the data model in (2), a hop-free data subset obtained by spectrogram entropy analysis may be written as the $M \times r$ matrix

$$\bar{\mathbf{X}} = \bar{\mathcal{A}}\bar{\mathbf{S}}^T \tag{4}$$

where $\bar{\mathcal{A}}$ and $\bar{\mathbf{S}}$ are submatrices of $\mathcal{A}$ and $\mathbf{S}$ in (2) such that $\bar{\mathcal{A}}$ contains the first d columns of $\mathcal{A}$, and

$$\bar{\mathbf{S}} = \begin{bmatrix} 1 & \gamma_1 & \cdots & \gamma_1^{r-1} \\ & & \vdots & \\ 1 & \gamma_{d-1} & \cdots & \gamma_{d-1}^{r-1} \\ 1 & \gamma_{d,x} & \cdots & \gamma_{d,x}^{r-1} \end{bmatrix}^T$$



(a)

(b)

(c)

Fig. 3. Hop instant detection via spectrogram entropy analysis. Ten asynchronous FH sources; SNR $= 15$ dB. (a) Spectrogram of received signal. (b) Corresponding entropy plot. (c) Hop detection using the average of minimum and maximum entropy as threshold (nine out of ten hops are correctly detected).

where r is an integer that depends on where the last sample ends in the hop-free set. Since $\bar{\mathbf{S}}$ is a Vandermonde matrix, and $\bar{\mathcal{A}}$ is the product of a Vandermonde matrix and a diagonal matrix, the factorization problem in (4) is a 2-D HR problem.

In multipath scenarios with negligible angle spread (far field), within a hop-free subset, the different paths for a given source will all give rise to the same 2-D harmonic but with different complex path loss coefficients (for a complex exponential, time shift is equivalent to phase shift plus scaling due to path attenuation). These can all be combined into a single 2-D harmonic, hence reverting back to the coherent case. However, *identifying* a hop-free subset of sufficient length will be more challenging due to additional hops introduced by multipath.

B. Identifiability

In general terms, the 2-D HR problem (including damping factors) can be stated as follows: Given a sum of F 2-D exponentials

$$x_{i,j} = \sum_{f=1}^{F} c_f a_f^{i-1} b_f^{j-1} \tag{5}$$

for $i = 1, \ldots, I$ and $j = 1, \ldots, J$, where $a_f, b_f, c_f \in \mathbb{C}, I$ denotes the number of samples taken along one dimension and J likewise for the other dimension, find the parameter triples (a_f, b_f, c_f) for $f = 1, \ldots, F$. Identifiability of 2-D HR is thoroughly addressed in [7].

Theorem 1—Almost Sure Identifiability of 2-D HR [7]: Given a sum of F 2-D exponentials as in (5) for

$i = 1,\ldots,I \geq 4$ and $j = 1,\ldots,J \geq 4$, the parameter triples $(a_f, b_f, c_f), f = 1,\ldots,F$ are $P_{\mathcal{L}}(\mathbb{C}^{2F})$ almost surely unique, where $P_{\mathcal{L}}(\mathbb{C}^{2F})$ is the distribution used to draw the $2F$ complex exponential parameters $(a_f, b_f), f = 1,\ldots,F$ that are assumed continuous with respect to the Lebesgue measure in $\mathbb{C}^{2F}$, provided that[4]

$$F \leq \left\lfloor \frac{I}{2} \right\rfloor \left\lceil \frac{J}{2} \right\rceil. \tag{6}$$

We stress that collinear sources (same spatial frequencies) or colliding users (same temporal frequencies) need not cause loss of identifiability. For example, the following explicit result is proven in [7].

Proposition 1: Given a sum of F 2-D exponentials

$$x_{i,j} = \sum_{f=1}^{F} c_f a_f^{i-1} b_f^{j-1}$$

for $i = 1,\ldots,I \geq 4$, and $j = 1,\ldots,J \geq 4$, with $b_2 = b_1$, the parameter triples $(a_f, b_f, c_f), f = 1,\cdots,F$ are $P_{\mathcal{L}}(\mathbb{C}^{2F-1})$ a.s. unique, where $P_{\mathcal{L}}(\mathbb{C}^{2F-1})$ is the distribution used to draw the $(2F-1)$ complex exponential parameters $(a_1, a_2, \cdots, a_F, b_1, b_3, \cdots, b_F)$, which are assumed continuous with respect to the Lebesgue measure in $\mathbb{C}^{2F-1}$, provided that

$$F \leq \left\lfloor \frac{I}{2} \right\rfloor \left\lceil \frac{J}{2} \right\rceil.$$

A variety of good algorithms have been developed for 2-D HR, e.g., [2], [5], [13], but much remains to be done.

- Existing algorithms do not come close to supporting as many sources as predicted by Theorem 1. In other words, algorithm identifiability is far from what is theoretically possible (that is, model identifiability).
- Existing algorithms can only approach the CRB at relatively high SNR.

A novel 2-D HR algorithm (the MDE-ALS algorithm) that achieves the identifiability bound predicted by Theorem 1 and stays close to the CRB even at low SNR will be developed in Section V. This algorithm consists of algebraic initialization followed by ML refinement and takes full advantage of 2-D harmonic signal structure. This also suggests that the algorithm might be sensitive to unmodeled dynamics, which are introduced in our present context by imperfect hop detection. We begin with another alternative based on spatial-temporal smoothing and quadrilinear regression. Instead of exploiting the full 2-D harmonic signal structure, this latter algorithm capitalizes on the shift-invariance property of exponentials, much like ESPRIT. The difference is that the end problem is solved by means of QALS regression. This affords a performance advantage relative to earlier ESPRIT-like approaches [11], [26] with improved robustness to unmodeled dynamics and noise color[5] as a side-benefit of LS. These benefits will be illustrated in Section VI. The choice between QALS and MDE-ALS of Section V depends on a number of factors, including number of sources, SNR, hop-detection performance, and complexity considerations. The reason for deferring the presentation of MDE-ALS until Section V is that certain required modules (alternating least squares and Tretter's frequency estimator) are more naturally introduced in the context of QALS and single-user tracking, respectively.

Remark 1: DOA estimation via 1-D HR (ignoring the temporal signal structure but collecting many more snapshots) is inherently bounded in terms of identifiability by the number of elements (I) of the array [25]. The 2-D HR approach, on the other hand, is bounded by the total spatio-temporal sample size (divided by four) $IJ/4$, where J is the number of hop-free temporal samples. Hence, the 2-D HR approach can resolve many more sources or paths and yield better performance due to better exploitation of model structure, as will be shown by simulation.

C. Low-Rank Decomposition of Quadrilinear Arrays

Consider an $I \times J \times K \times L$ *four*-way array $\underline{\mathbf{X}}$ with typical element

$$x_{i,j,k,l} = \sum_{f=1}^{F} a_{i,f} b_{j,f} g_{k,f} h_{l,f} \tag{7}$$

for $i = 1,\ldots,I, j = 1,\ldots,J, k = 1,\ldots,K, l = 1,\ldots,L$, with $a_{i,f}, b_{j,f}, g_{k,f}, h_{l,f} \in \mathbb{C}$. Equation (7) expresses the four-way (quadrilinear) array $\underline{\mathbf{X}}$ as a sum of F rank-one four-way factors. Analogous to the definition of matrix rank, the rank of a four-way array can be defined as the minimum number of rank-one components needed to decompose $\underline{\mathbf{X}}$. Quadrilinear decomposition falls under the umbrella of PARAllel FACtor (PARAFAC) analysis [10], [14], [18].

Define $\mathbf{A} \in \mathbb{C}^{I\times F}$ with $\mathbf{A}(i,f) := a_{i,f}, \mathbf{B} \in \mathbb{C}^{J\times F}$ with $\mathbf{B}(j,f) := b_{j,f}, \mathbf{G} \in \mathbb{C}^{K\times F}$ with $\mathbf{G}(k,f) := g_{k,f}$, and $\mathbf{H} \in \mathbb{C}^{L\times F}$ with $\mathbf{H}(l,f) := h_{l,f}$. Furthermore, define $\mathbf{X}_{k,l} \in \mathbb{C}^{I\times J}, \mathbf{X}_{l,i} \in \mathbb{C}^{J\times K}, \mathbf{X}_{i,j} \in \mathbb{C}^{K\times L}$, and $\mathbf{X}_{j,k} \in \mathbb{C}^{L\times I}$ with corresponding typical elements

$$\mathbf{X}_{k,l}(i,j) = \mathbf{X}_{l,i}(j,k) = \mathbf{X}_{i,j}(k,l) = \mathbf{X}_{j,k}(l,i) = x_{i,j,k,l}.$$

Then, the model in (7) can be written in four different ways in terms of systems of simultaneous matrix equations

$$\mathbf{X}_{k,l} = \mathbf{A}\mathbf{D}_k(\mathbf{G})\mathbf{D}_l(\mathbf{H})\mathbf{B}^T, \quad k = 1,\ldots,K, \quad l = 1,\ldots,L \tag{8}$$

$$\mathbf{X}_{l,i} = \mathbf{B}\mathbf{D}_l(\mathbf{H})\mathbf{D}_i(\mathbf{A})\mathbf{G}^T, \quad l = 1,\ldots,L, \quad i = 1,\ldots,I \tag{9}$$

$$\mathbf{X}_{i,j} = \mathbf{G}\mathbf{D}_i(\mathbf{A})\mathbf{D}_j(\mathbf{B})\mathbf{H}^T, \quad i = 1,\ldots,I, \quad j = 1,\ldots,J \tag{10}$$

$$\mathbf{X}_{j,k} = \mathbf{H}\mathbf{D}_j(\mathbf{B})\mathbf{D}_k(\mathbf{G})\mathbf{A}^T, \quad j = 1,\ldots,J, \quad k = 1,\ldots,K. \tag{11}$$

By stacking the matrices in (8), we can construct a matrix representation of the four-way array $\underline{\mathbf{X}}$:

$$\mathbf{X}^{(IKL\times J)} = \begin{bmatrix} \mathbf{X}_{1,1} \\ \mathbf{X}_{2,1} \\ \vdots \\ \mathbf{X}_{K,L} \end{bmatrix} = (\mathbf{H} \odot \mathbf{G} \odot \mathbf{A})\,\mathbf{B}^T \tag{12}$$

where the superscript $^{(IKL\times J)}$ means that the matrix is of size $IKL \times J$ and that the i-index (I goes first in the product IKL) runs fastest along the columns, whereas the l-index runs slowest. Unlike low-rank matrix decomposition, which

[4]The Theorem holds true if I and J are switched.

[5]Introduced by 2-D smoothing.

is inherently nonunique unless one imposes orthogonality or other strong constraints, low-rank decomposition of multidimensional arrays is essentially unique under a relatively mild rank-like condition [10], [17].

D. Temporal and Spatial Smoothing

Smoothing is a commonly used technique to take advantage of the shift-invariance property of complex exponentials in algebraic frequency estimation methods. We will use smoothing to generate a quadrilinear model from the two-way model of (4). Define m_1 matrices, each of size $M \times l_1$:

$$\mathbf{Y}_u := \bar{\mathbf{X}}(:, u : l_1 + u - 1), \quad u = 1, \ldots, m_1$$

where $\bar{\mathbf{X}}(:, u : l_1 + u - 1)$ stands for columns u to $l_1 + u - 1$ of matrix $\bar{\mathbf{X}}$, and $l_1 = r - m_1 + 1$. m_1 is known as the temporal smoothing factor. Due to the Vandermonde structure of $\bar{\mathbf{S}}$, it holds that

$$\mathbf{Y}_u = \bar{\mathcal{A}}[\bar{\mathbf{S}}(u : l_1 + u - 1, :)]^T = \bar{\mathcal{A}}\mathbf{D}_u(\bar{\mathbf{S}})(\bar{\mathbf{S}}^{(l_1)})^T.$$

By vertically stacking these m_1 submatrices extracted from $\bar{\mathbf{X}}$, we obtain

$$\mathbf{Y} = \begin{bmatrix} \mathbf{Y}_1 \\ \mathbf{Y}_2 \\ \vdots \\ \mathbf{Y}_{m_1} \end{bmatrix} = \bar{\mathbf{S}}^{(m_1)} \odot \bar{\mathcal{A}}(\bar{\mathbf{S}}^{(l_1)})^T. \tag{13}$$

Next, to take advantage of the shift-invariance structure of $\bar{\mathcal{A}}$, we define

$$\begin{aligned} \bar{\mathbf{Y}}_v &:= \begin{bmatrix} \mathbf{Y}_1(v : l_2 + v - 1, :) \\ \mathbf{Y}_2(v : l_2 + v - 1, :) \\ \vdots \\ \mathbf{Y}_{m_1}(v : l_2 + v - 1, :) \end{bmatrix} \\ &= \bar{\mathbf{S}}^{(m_1)} \odot \bar{\mathcal{A}}^{(l_2)}\mathbf{D}_v(\bar{\mathbf{A}})(\bar{\mathbf{S}}^{(l_1)})^T \end{aligned}$$

for $v = 1, \ldots, m_2$, where m_2 is the spatial smoothing factor, and $l_2 = M - m_2 + 1$. Recall that $\bar{\mathcal{A}}$ is the first d columns of $\mathcal{A}$, and $\bar{\mathbf{A}}$ is a matrix consisting of the first d columns of $\mathbf{A}$, which does not involve attenuation and phase shift factors; cf. (1) and (2). Similar to (13), stacking the $\bar{\mathbf{Y}}_v$s together and adopting the same notation as in (12), we obtain

$$\begin{aligned} \bar{\mathbf{Y}}^{(l_2 m_1 m_2 \times l_1)} &= \begin{bmatrix} \bar{\mathbf{Y}}_1^T & \bar{\mathbf{Y}}_2^T & \cdots & \bar{\mathbf{Y}}_{m_2}^T \end{bmatrix}^T \\ &= (\bar{\mathbf{A}}^{(m_2)} \odot \bar{\mathbf{S}}^{(m_1)} \odot \bar{\mathcal{A}}^{(l_2)})(\bar{\mathbf{S}}^{(l_1)})^T. \end{aligned} \tag{14}$$

This shows that a two-way model (4) with shift-invariance structure along two modes (i.e., a 2-D harmonic mixture $\mathbf{V}_1\mathbf{V}_2^T$, with $\mathbf{V}_1$ and $\mathbf{V}_2$ both Vandermonde) can be transformed to a quadrilinear model (14) with "residual" Vandermonde structure in all four dimensions.[6] In a nutshell, QALS ignores this residual Vandermonde structure and fits the resulting quadrilinear model in a LS sense. This is explained next.

E. QALS

The principle of alternating least squares (ALS) can be used to fit the quadrilinear model in (14) on the basis of noisy observations [14]. The basic idea behind ALS is to update each time a subset of parameters, using least squares conditioned on interim estimates of the remaining parameters. Because the fit is bounded below and can never increase (each step optimizes the fit conditioned on the remaining parameters), ALS is monotonically convergent in LS fit. In the case of QALS, the parameters are split in four subsets, each corresponding to one of the four parameter matrices.

Least squares model fitting for (14) amounts to minimizing

$$\|\tilde{\mathbf{Y}}^{(l_2 m_1 m_2 \times l_1)} - \bar{\mathbf{A}}^{(m_2)} \odot \bar{\mathbf{S}}^{(m_1)} \odot \bar{\mathcal{A}}^{(l_2)}(\bar{\mathbf{S}}^{(l_1)})^T\|_F^2$$

over $\bar{\mathbf{A}}^{(m_2)}, \bar{\mathcal{A}}^{(l_2)}, \bar{\mathbf{S}}^{(m_1)}$ and $\bar{\mathbf{S}}^{(l_1)}$, where $\tilde{\mathbf{Y}}^{(l_2 m_1 m_2 \times l_1)}$ is the noisy counterpart of $\bar{\mathbf{Y}}^{(l_2 m_1 m_2 \times l_1)}$. The conditional least squares update for $\bar{\mathbf{S}}^{(l_1)}$ is

$$(\hat{\mathbf{S}}^{(l_1)})^T = [\hat{\mathbf{A}}^{(m_2)} \odot \hat{\mathbf{S}}^{(m_1)} \odot \hat{\mathcal{A}}^{(l_2)}]^\dagger \tilde{\mathbf{Y}}^{(l_2 m_1 m_2 \times l_1)}$$

where $\hat{\mathbf{A}}^{(m_2)}, \hat{\mathbf{S}}^{(m_1)}$, and $\hat{\mathcal{A}}^{(l_2)}$ denote previously obtained estimates of $\bar{\mathbf{A}}^{(m_2)}, \bar{\mathbf{S}}^{(m_1)}$, and $\bar{\mathcal{A}}^{(l_2)}$. One may now resort to the complete symmetry of the quadrilinear model [cf. (7)] and data reshaping [cf. (9)–(11)] to figure out corresponding conditional LS updates for $\bar{\mathbf{A}}^{(m_2)}, \bar{\mathbf{S}}^{(m_1)}$ and $\bar{\mathcal{A}}^{(l_2)}$ as follows:

$$\begin{aligned} (\hat{\mathbf{A}}^{(m_2)})^T &= [\hat{\mathbf{S}}^{(l_1)} \odot \hat{\mathcal{A}}^{(l_2)} \odot \hat{\mathbf{S}}^{(m_1)}]^\dagger \tilde{\mathbf{Y}}^{(m_1 l_2 l_1 \times m_2)} \\ (\hat{\mathbf{S}}^{(m_1)})^T &= [\hat{\mathcal{A}}^{(l_2)} \odot \hat{\mathbf{A}}^{(m_2)} \odot \hat{\mathbf{S}}^{(l_1)}]^\dagger \tilde{\mathbf{Y}}^{(l_1 m_2 l_2 \times m_1)} \\ (\hat{\mathcal{A}}^{(l_2)})^T &= [\hat{\mathbf{S}}^{(m_1)} \odot \hat{\mathbf{S}}^{(l_1)} \odot \hat{\mathbf{A}}^{(m_2)}]^\dagger \tilde{\mathbf{Y}}^{(m_2 l_1 m_1 \times l_2)}. \end{aligned}$$

Upon convergence of QALS, $\bar{\mathbf{A}}^{(m_2)}, \bar{\mathbf{S}}^{(m_1)}, \bar{\mathcal{A}}^{(l_2)}$, and $\bar{\mathbf{S}}^{(l_1)}$ will be estimated up to scaling and common permutation of columns [14]. The frequencies and DOAs can then be estimated via simple division or other single 1-D harmonic retrieval techniques (e.g., [24] or periodogram). Since the permutation of columns is common to all four matrices, (f_i, α_i) will be paired up automatically.

IV. Single-User Tracking

After the DOAs have been recovered, single-user tracking amounts to the joint estimation of hop instants, frequencies, and phases of a source signal from a particular DOA over a time interval of interest. In this section, an approach combining linear MMSE beamforming and dynamic programming is developed for this purpose. Using the recovered steering vectors, a linear MMSE beamformer can be applied to obtain signals from a desired DOA while suppressing interference from other directions. Subsequently, dynamic programming (DP) is used for the joint optimal ML estimation of frequencies, phases and hop instants. Fig. 4 outlines the overall procedure.

A. Linear MMSE Beamforming

Suppose we have obtained the antenna steering matrix $\bar{\mathcal{A}}$ by applying QALS to a hop-free data block. Then, a linear MMSE beamformer can be used to separate the source signals, as long as the DOAs do not change or change only slightly. Although $\bar{\mathcal{A}}$ is indeed affected by frequency hopping, the effect can be ignored when the signal's hopping bandwidth is within a few percent of the carrier frequency, which is the case of practical interest. For example, both the IEEE 802.11 frequency-hopping spread spectrum standard and the Bluetooth use 79 distinct frequency channels (23 in Japan) over the 2.4-GHz ISM radio frequency band with 1 MHz channel spacing. The hopping band is approximately 3.3% of the carrier frequency. The situation is

[6]The smoothing parameters m_1, m_2 can be chosen to optimize identifiability [19] or QALS performance. The latter is difficult analytically, but Monte-Carlo simulation is straightforward.

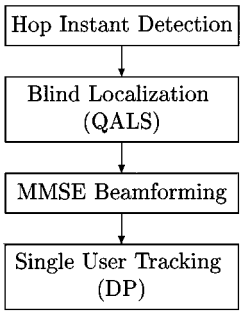


Fig. 4. Procedure of blind localization and tracking.

similar in military communications, where the hopping bandwidth is under 1% of center frequency. Therefore, it is reasonable to assume that $\bar{\mathcal{A}}$ is approximately constant over the time interval of interest. The linear MMSE beamformer is given by

$$\mathbf{W}^H = \frac{1}{\sigma^2}\left(\mathbf{R}_{\mathbf{S}}^{-1} + \frac{1}{\sigma^2}\bar{\mathcal{A}}^H\bar{\mathcal{A}}\right)^{-1}\bar{\mathcal{A}}^H \tag{15}$$

where $\mathbf{R_S}$ is the covariance matrix of the source signal $\mathbf{S}$, and σ^2 is the variance of the additive white Gaussian noise. For $\mathbf{R_S} = \mathbf{I}$ (which is usually the case in practice), (15) simplifies to

$$\mathbf{W}^H = (\sigma^2\mathbf{I} + \bar{\mathcal{A}}^H\bar{\mathcal{A}})^{-1}\bar{\mathcal{A}}^H. \tag{16}$$

MMSE beamforming suffers less from noise gain than does zero-forcing beamforming. Any type of spatial beamforming, including MMSE, combines rows of the observation matrix on a column-by-column basis; this assures that postbeamforming noise is temporally white if the input noise is temporally white. Applying the beamformer to the received $M \times N$ signal matrix $\mathbf{X}$, a typical recovered source signal from a desired DOA will be of the form

$$s(n) = a_i e^{j(\omega_i n + \phi_i)} + \epsilon(n), \quad n_{i-1} \le n < n_i \tag{17}$$

for $i = 1, \ldots, p$, where $0 = n_0 < n_1 < \cdots < n_p = N$ correspond to the $p-1$ unknown switching instants, and $\epsilon(n)$ denotes zero-mean complex white Gaussian noise.[7] The actual number of frequency switches is *a priori* unknown; however, a crude upper bound $(p-1)$ is usually available or can be estimated from the periodogram. If the actual number of hops is less than $p-1$, then two or more contiguous segments in the dynamic programming result will have approximately the same frequency, and thus, one can identify false hops.

B. Dynamic Programming

From (17), the joint maximum likelihood estimation of the frequency vector $\boldsymbol{\omega} = [\omega_1, \ldots, \omega_p]$, phase vector $\boldsymbol{\phi} = [\phi_1, \ldots, \phi_p]$, amplitude vector $\mathbf{a} = [a_1, \ldots, a_p]$, and switching instants $\mathbf{n} = [n_1, \ldots, n_{p-1}]$ amounts to minimizing

$$J(\boldsymbol{\omega}, \boldsymbol{\phi}, \mathbf{a}, \mathbf{n}) = \sum_{i=1}^{p}\sum_{n=n_{i-1}}^{n_i-1} |s(n) - a_i e^{j(\omega_i n + \phi_i)}|^2 \tag{18}$$

over $\boldsymbol{\omega}, \boldsymbol{\phi}, \mathbf{a}, \mathbf{n}$.

For a given time segment defined by hop instants n_{i-1} and n_i, the estimation of ω_i, ϕ_i, and a_i is equivalent to frequency, phase, and amplitude estimation of a complex exponential from a sequence of uniformly spaced samples corrupted by additive white Gaussian noise, and a variety of techniques are available. The optimal maximum likelihood estimator is provided by the peak of the periodogram. Tretter's algorithm [24] offers a particularly attractive solution for our purpose, as it has been shown to essentially attain the CRB for moderate sample sizes and SNR, and has linear complexity [8]. Hence, we adopt [24] for single frequency estimation over a given time segment with the understanding that if SNR is low, then periodogram-based estimation should be used to maintain optimality.

Suppose $\psi(n)(n_{i-1} \le n < n_i)$ are the phase angles obtained by applying a phase unwrapping algorithm (see, e.g., [3] and [20]) to the principal value of the observed signal phase. Then, the estimates of ω_i and ϕ_i are given by [24]

$$\begin{bmatrix}\hat{\omega}_i \\ \hat{\phi}_i\end{bmatrix} = \begin{bmatrix} N_i & -N_i n_{i-1} - U_1 \\ -N_i n_{i-1} - U_1 & N_i n_{i-1}^2 + 2n_{i-1}U_1 + U_2\end{bmatrix} \times \frac{12}{N_i^2(N_i^2-1)}\begin{bmatrix}\sum_{n=n_{i-1}}^{n_i-1} n\psi(n) \\ \sum_{n=n_{i-1}}^{n_i-1}\psi(n)\end{bmatrix} \tag{19}$$

where $N_i = n_i - n_{i-1}$ is the segment sequence length, and

$$U_1 = \sum_{n=0}^{N_i-1} n = (N_i - 1)N_i/2$$

$$U_2 = \sum_{n=0}^{N_i-1} n^2 = (N_i - 1)N_i(2N_i - 1)/6.$$

Given estimates of ω_i and ϕ_i, an estimate of the amplitude a_i can be easily obtained via conditional LS regression.
Writing (17) in vector form as

$$\mathbf{s} = a_i\mathbf{e}(\hat{\omega}_i, \hat{\phi}_i) + \boldsymbol{\epsilon}$$

separating and stacking real and imaginary parts

$$\bar{\mathbf{s}} := \begin{bmatrix}\mathrm{Re}(\mathbf{s}) \\ \mathrm{Im}(\mathbf{s})\end{bmatrix} = a_i\begin{bmatrix}\mathrm{Re}(\mathbf{e}(\hat{\omega}_i, \hat{\phi}_i)) \\ \mathrm{Im}(\mathbf{e}(\hat{\omega}_i, \hat{\phi}_i))\end{bmatrix} + \begin{bmatrix}\mathrm{Re}(\epsilon) \\ \mathrm{Im}(\epsilon)\end{bmatrix} =: a_i\bar{\mathbf{e}} + \bar{\epsilon}$$

from which the conditional LS estimate of the amplitude is easily obtained as

$$\hat{a}_i = \frac{1}{\bar{\mathbf{e}}^T\bar{\mathbf{e}}}\bar{\mathbf{e}}^T\bar{\mathbf{s}}. \tag{20}$$

Similar to [9], we define the cost function for the time segment delimited by $n_{i-1} \le n < n_i$ as

$$\Lambda_i[n_{i-1}, n_i - 1] = \sum_{n=n_{i-1}}^{n_i-1} |s(n) - \hat{a}_i e^{j(\hat{\omega}_i n + \hat{\phi}_i)}|^2.$$

The minimization of (18) is equivalent to the minimization of

$$J(\hat{\boldsymbol{\omega}}, \hat{\boldsymbol{\phi}}, \hat{\mathbf{a}}, \mathbf{n}) = \sum_{i=1}^{p}\Lambda_i[n_{i-1}, n_i - 1]$$

over $\mathbf{n}$, where $\hat{\boldsymbol{\omega}}, \hat{\boldsymbol{\phi}}$, and $\hat{\mathbf{a}}$ are estimates corresponding to $\mathbf{n}$ obtained by using (19) and (20) or periodogram-based estimation if SNR is below the threshold of (19); see [8]. In order to solve this latter minimization problem by dynamic programming, we define

$$\Gamma_h(L) = \min_{\substack{n_1, \ldots, n_{h-1} \\ n_0 = 0, n_h = L+1}} \sum_{i=1}^{h}\Lambda_i[n_{i-1}, n_i - 1] \tag{21}$$

where $0 < n_1 < \cdots < n_{h-1} < L + 1$. Equation (21) can be viewed as the minimization problem of finding the best fit for

[7]Residual interference after MMSE beamforming can be approximated by AWGN since signals from different directions are assumed to be independent.

the length $L+1$ subsequence when a total number of $h-1$ hops is allowed. Hence, $\Gamma_p(N-1)$ is the minimum of $J(\hat{\boldsymbol{\omega}}, \hat{\boldsymbol{\phi}}, \hat{\mathbf{a}}, \mathbf{n})$.

From (21), a recursion for the minimum can be developed as

$$\begin{aligned}\Gamma_h(L) &= \min_{\substack{n_{h-1}\\ n_h=L+1}} \min_{\substack{n_1,\cdots,n_{h-2}\\ n_0=0}} \sum_{i=1}^{h} \Lambda_i[n_{i-1}, n_i - 1] \\ &= \min_{\substack{n_{h-1}\\ n_h=L+1}} \left(\min_{\substack{n_1,\cdots,n_{h-2}\\ n_0=0}} \sum_{i=1}^{h-1} \Lambda_i[n_{i-1}, n_i - 1] \right. \\ &\quad \left. + \Lambda_h[n_{h-1}, n_h - 1] \right) \\ &= \min_{n_{h-1}} (\Gamma_{h-1}[n_{h-1} - 1] + \Lambda_h[n_{h-1}, L])\end{aligned}$$

which simply says that for a signal sequence $[0, L]$, the minimum error for h segments (i.e., $h-1$ hop instants) is the minimum error for the first $h-1$ segments that end at $n = n_{h-1} - 1$ and the error contributed by the last segment from $n = n_{h-1}$ to $n = L$. The solution of the minimization in (18) is obtained for $h = p$ and $L = N - 1$.

Assuming that the minimum length of a segment is two samples (since it is impossible to obtain valid frequency and phase estimates from one sample), the procedure to compute the solution by dynamic programming is as follows.

1) Initialization: For $h = 1$, compute $\Gamma_1(L)$ for $L = 1, \ldots, N - 2p + 1$ as

$$\begin{aligned}\Gamma_1(L) &= \Lambda_1[n_0 = 0, n_1 - 1 = L] \\ &= \sum_{n=0}^{L} |s(n) - \hat{a}_i e^{j(\hat{\omega}_1 n + \hat{\phi}_1)}|^2\end{aligned}$$

where $\hat{\omega}_1, \hat{\phi}_1$, and $\hat{a}_1$ are estimates obtained by applying (19) and (20) to the signal sequence $[0, L]$.

2) Recursion: For $2 \le h \le p - 1$, compute $\Gamma_h(L)$ for $L = 2h - 1, \ldots, N - 2p + 2h - 1$ as

$$\Gamma_h(L) = \min_{2h-2 \le n_{h-1} < L} (\Gamma_{h-1}(n_{h-1} - 1) + \Lambda_h[n_{k-1}, L]).$$

For each L, denote the value of n_{h-1} that minimizes $\Gamma_h(L)$ as $n_{h-1}(L)$, and denote the corresponding $\hat{\omega}_{h-1}, \hat{\phi}_{h-1}, \hat{a}_{h-1}$ as $\hat{\omega}_{h-1}(L), \hat{\phi}_{h-1}(L)$, and $\hat{a}_{h-1}(L)$, respectively. This information will be used for backtracking.

3) Termination: For $h = p$, compute $\Gamma_p(L)$ for $L = N - 1$ as

$$\begin{aligned}\Gamma_p(N-1) = \min_{2p-2 \le n_{p-1} < L} &(\Gamma_{p-1}(n_{p-1} - 1) \\ &+ \Lambda_p[n_{p-1}, N - 1]).\end{aligned}$$

Denote the value of n_{p-1} that minimizes $\Gamma_p(N-1)$ as $n_{p-1}(N-1)$, and denote the corresponding $\hat{\omega}_{p-1}, \hat{\phi}_{p-1}$ and $\hat{a}_{p-1}$ as $\hat{\omega}_{p-1}(N-1), \hat{\phi}_{p-1}(N-1)$, and $\hat{a}_{p-1}(N-1)$, respectively. The minimum of $J(\boldsymbol{\omega}, \boldsymbol{\phi}, \mathbf{a}, \mathbf{n})$ in (18) is given by $\Gamma_p(N-1)$.

4) Backtracking: Finally, the maximum likelihood estimates of hop instants are obtained by using the backward recursion, i.e., $\hat{n}_i = n_i(\hat{n}_{i+1} - 1)$, for $i = p-2, p-3, \ldots, 1$ initialized by $\hat{n}_{p-1} = n_{p-1}(N-1)$. Similarly, the corresponding frequency, phase, and amplitude estimates of each segment can be obtained by their respective backward recursions.

The complexity of the above dynamic programming algorithm is $\mathcal{O}(pN^3)$. In practical FH systems, frequencies hop at a regular rate; therefore, it is enough to estimate two parameters: hop timing and hop period. These can be obtained by applying dynamic programming to a relatively short portion of a long data record, whereas frequency estimation for the remaining data can be accomplished by applying (19) and (20) to each segment of fixed length (hop interval). This will reduce the complexity significantly.

V. Two-Dimensional Harmonic Retrieval

We now return to 2-D HR. The novel MDE-ALS algorithm consists of two steps: algebraic initialization (MDE) and least squares refinement (ALS).

A. Initialization

From (5), we can define $\mathbf{X} \in \mathbb{C}^{I \times J}$ with $\mathbf{X}(i,j) = x_{i,j}$, $\mathbf{A} \in \mathbb{C}^{I \times F}$ with $\mathbf{A}(i,f) = a_f^{i-1}$, $\mathbf{B} \in \mathbb{C}^{J \times F}$ with $\mathbf{B}(j,f) = b_f^{j-1}$, and a diagonal matrix $\mathbf{C} \in \mathbb{C}^{F \times F}$ with $\mathbf{C}(f,f) = c_f$. Then, (5) can be written in matrix form as

$$\mathbf{X} = \mathbf{A}\mathbf{C}\mathbf{B}^T. \tag{22}$$

Define a five-way array $\hat{\mathbf{X}}$ with typical element

$$\hat{x}_{i_1,i_2,i_3,j_1,j_2} := x_{i_1+i_2+i_3-2, j_1+j_2-1} \tag{23}$$

where $i_\alpha = 1, \cdots, I_\alpha \ge 2, j_\beta = 1, \cdots, J_\beta \ge 2$, for $\alpha = 1,2, \beta = 1,2$, and $i_3 = 1,2$. To maximize the number of identifiable harmonics, I_1, I_2, J_1 and J_2 are chosen such that (the maximum number of identifiable harmonics is $\lfloor (I/2) \rfloor \lceil (J/2) \rceil$; see [7] for details)

$$\begin{aligned} I_1 &= \left\lfloor \frac{I}{2} \right\rfloor, \quad J_1 = \left\lceil \frac{J+1}{2} \right\rceil \\ I_2 &= \left\lceil \frac{I}{2} \right\rceil, \quad J_2 = \left\lfloor \frac{J+1}{2} \right\rfloor. \end{aligned} \tag{24}$$

Then, nest the five-way array $\hat{\mathbf{X}}$ into a three-way array $\mathbf{Z}$ by collapsing two pairs of dimensions as follows (hence, it is called multidimensional embedding):

$$z_{k,l,i_3} := \hat{x}_{\left\lceil \frac{k}{J_1} \right\rceil, \left\lceil \frac{l}{J_2} \right\rceil, i_3, k-(\left\lceil \frac{k}{J_1} \right\rceil - 1)J_1, l-(\left\lceil \frac{l}{J_2} \right\rceil - 1)J_2}$$

for $k = 1, \cdots, I_1 J_1, l = 1, \cdots, I_2 J_2$. Furthermore, define

$$\mathbf{P}_1 = \mathbf{A}^{(I_1)} \odot \mathbf{B}^{(J_1)}, \quad \mathbf{P}_2 = \mathbf{A}^{(I_2)} \odot \mathbf{B}^{(J_2)} \tag{25}$$

and let $\mathbf{Z}_1 = \mathbf{Z}(:,:,1)$ and $\mathbf{Z}_2 = \mathbf{Z}(:,:,2)$. Then, it can be verified that

$$\begin{aligned} \mathbf{Z}_1 &= \mathbf{P}_1 \mathbf{D}_1(\mathbf{A}^{(2)}) \mathbf{C} {\mathbf{P}_2}^T = \mathbf{P}_1 \mathbf{C} {\mathbf{P}_2}^T \\ \mathbf{Z}_2 &= \mathbf{P}_1 \mathbf{D}_2(\mathbf{A}^{(2)}) \mathbf{C} {\mathbf{P}_2}^T = \mathbf{P}_1 \mathbf{C}\mathbf{Q} {\mathbf{P}_2}^T \end{aligned}$$

where $\mathbf{Q} = \mathrm{diag}(a_1, \ldots, a_F)$. It is shown in [7] that $\mathbf{P}_1$ and $\mathbf{P}_2$ are full column rank almost surely[8], and the singular value decomposition of the stacked data yields

$$\begin{bmatrix} \mathbf{Z}_1 \\ \mathbf{Z}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{P}_1 \mathbf{C} \\ \mathbf{P}_1 \mathbf{C}\mathbf{Q} \end{bmatrix} {\mathbf{P}_2}^T = \mathbf{U}_{2I_1J_1 \times F} \boldsymbol{\Sigma}_{F \times F} \mathbf{V}^H_{I_2J_2 \times F} \tag{26}$$

where $\mathbf{U}$ has F columns that together span the column space of $[\mathbf{Z}_1^T \ \mathbf{Z}_2^T]^T$. Since the same space is spanned by the columns of $[(\mathbf{P}_1\mathbf{C})^T \ (\mathbf{P}_1\mathbf{C}\mathbf{Q})^T]^T$, there exists an $F \times F$ nonsingular matrix $\mathbf{T}$ such that

$$\mathbf{U} = \begin{bmatrix} \mathbf{U}_1 \\ \mathbf{U}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{P}_1 \mathbf{C} \\ \mathbf{P}_1 \mathbf{C}\mathbf{Q} \end{bmatrix} \mathbf{T}.$$

[8]See [7] for further results on the structure and rank of Khatri–Rao products of Vandermonde matrices.

TABLE I
MDE-ALS ALGORITHM

1. **Initialization**
   Given $\mathbf{X}$, use the MDE algorithm in Section V-A to obtain initial estimates of $\mathbf{A}$ and $\mathbf{B}$.
2. **Pseudo ALS**
   i) Update $\mathbf{B}$: for $f = 1$ to F,
   - $\mathbf{X}_{(f)} = \mathbf{ACB}^T - c_f\mathbf{a}_f\mathbf{b}_f^T$
   - $\mathbf{b}_f^T = \mathbf{a}_f^\dagger(\mathbf{X} - \mathbf{X}_{(f)})/c_f$
   - Estimate ξ_f from $\mathbf{b}_f$ by (19)
   - Reconstruct $\mathbf{b}_f = [1\ e^{j\xi_f}\ \cdots\ e^{j\xi_f(J-1)}]^T$
   - $c_f = \mathbf{a}_f^\dagger(\mathbf{X} - \mathbf{X}_{(f)})(\mathbf{b}_f^\dagger)^T$

   ii) Update $\mathbf{A}$: as in step i) but exchanging the roles of $\mathbf{A}$ and $\mathbf{B}$, replacing $\mathbf{X}$ with $\mathbf{X}^T$, $\mathbf{X}_{(f)}$ with $\mathbf{X}_{(f)}^T$.
   iii) Go to step i) until convergence, or upper bound on number of iterations is reached.
3. **ALS Refinement**
   Same as step 2, but using periodogram to estimate ζ_f, ξ_f from $\mathbf{a}_f$, $\mathbf{b}_f$.

It then follows that

$$\mathbf{U}_1^\dagger\mathbf{U}_2 = \mathbf{T}^{-1}\mathbf{QT}$$

which is an eigenvalue decomposition problem. The F distinct eigenvalues are the generators of the Vandermonde matrix $\mathbf{A}$. $\mathbf{T}^{-1}$ contains the eigenvectors of $\mathbf{U}_1^\dagger\mathbf{U}_2$ (scaled to unit norm). The remaining parameters can be obtained by

$$\mathbf{P}_1\mathbf{C} = \mathbf{UT}^{-1}, \quad \mathbf{P}_2 = [(\mathbf{P}_1\mathbf{C})^\dagger\mathbf{Z}_1]^T.$$

Notice that the first row of the product $\mathbf{P}_1\mathbf{C}$ is the diagonal of $\mathbf{C}$, i.e., $[c_1, \ldots, c_F]$. Now, the a_fs and b_fs can be readily recovered from $\mathbf{P}_1$ and/or $\mathbf{P}_2$, for example, the second and (J_1+1)th rows of $\mathbf{P}_1$ are $[b_1, \ldots, b_F]$ and $[a_1, \ldots, a_F]$, respectively. Due to the rich structure of the Khatri-Rao product of Vandermonde matrices, better estimates can be obtained via averaging. Note that no pairing issue exists, i.e., (a_f, b_f, c_f) are paired up automatically.

B. LS Refinement

More accurate estimates can be obtained by using pseudo ALS and ALS to refine algebraic estimates. Our discussion below focuses on undamped complex exponentials $a_f = e^{j\zeta_f}$ and $b_f = e^{j\xi_f}$, but ALS can be modified to handle damped complex exponentials. The procedure is summarized in Table I.

In Table I, estimates of ζ_f and ξ_f obtained via (19) are not true least squares solutions; hence, the name *pseudo* ALS. However, pseudo ALS can get close to the optimum quickly due to the low complexity of (19). Final refinement by periodogram-based ALS uses true conditional least squares estimates, which also guarantees eventual monotone convergence[9] of the overall algorithm.

VI. Simulation Results

In the following simulations, we consider an FHSS communication system operating at $f_c = 1$ GHz center frequency. The receiving ULA consists of $M = 6$ equally spaced antennas, whose baseline separation is half a wavelength of f_c. A hopping frequency band with a bandwidth of 8 MHz is occupied by 32 distinct frequency channels with 0.25 MHz channel spacing from 1 to 1.0075 GHz. The minimum hop size[10] between contiguous hops is 0.75 MHz.

A. Blind Localization

In the first experiment, we test the performance of the proposed blind localization algorithm. Three sources, with intermediate frequencies $[3.5\ 1.0\ 1.5]$ MHz modulated at $f_c = 1$ GHz, impinge on the ULA from $\alpha_1 = 5°, \alpha_2 = 10°$, and $\alpha_3 = 15°$. After downconversion, the antenna outputs are sampled at a rate of 8 MHz, and $K = 128$ samples are collected at each antenna. Since we are dealing with slow FH, suppose that during this data block, the third source hops from 1.5 to 2.25 MHz at normalized time instant 0.64, whereas the other two signal frequencies remain constant.

SNR is defined as $10\log_{10}(\|\mathbf{X}\|_F^2/(MK\sigma^2))$; cf. (2). For each SNR value, we first identify a hop-free subset by spectrogram entropy analysis, which we refer to as preprocessing. A data sequence from one antenna is split into seven overlapping segments, each of length 32; hence, the frequency resolution of the spectrogram is 0.25 MHz. The fifth window contains the hop and, hence, should have maximum entropy. The result for SNR $= 15$ dB is shown in Fig. 2, from which we can clearly identify a hop-free data subset corresponding to the first four segments.

Then, QALS is applied to this hop-free data subset for DOA and frequency estimation. Both temporal and spatial smoothing factors are 3; this choice assures that QALS can identify the model parameters, but it does not reflect any attempt to optimize the smoothing factors performance-wise. QALS is randomly initialized, and it typically converges after about 40 iterations. For comparison, we also apply the JAFE algorithm [11] to the same hop-free data subset. JAFE smoothes data along the temporal dimension and then capitalizes on the shift invariance properties of the data to cast the angle and frequency estimation problem into the framework of joint diagonalization of a set of matrices. In our simulations, JAFE was implemented using Jacobi iterations. Estimation error is measured by root mean square error (RMSE) of DOA and frequency estimation. Fig. 5 shows Monte Carlo simulation results comparing QALS to JAFE for DOA and frequency estimation of the three sources. The results demonstrate that QALS offers a 5-dB SNR advantage over JAFE due to the fact that it better exploits (quadrilinear) model structure.

In the second experiment, we test the robustness of QALS. The parameter settings are the same as in the first experiment. This time, however, we apply QALS and JAFE directly to the the raw data, i.e., without first detecting a hop-free data set; hence, the parameters before and after hopping will be estimated at the

[9]Due to the fact that (19) is not strictly LS, the pseudo ALS step does not provide convergence guarantees. For a convergence proof of ALS schemes, see, e.g., [12].

[10]This is similar to IEEE Standard 802.11, which uses 79 frequency channels with 1 MHz channel spacing, and a minimum of 6 MHz hop size is required between contiguous hops.

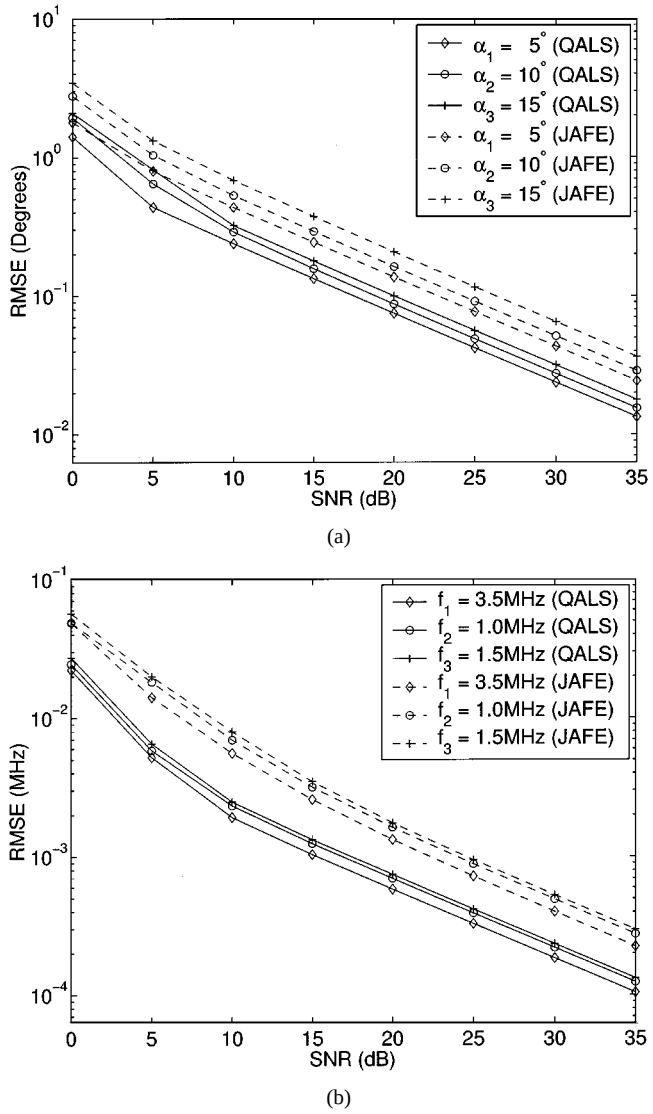


Fig. 5. QALS versus JAFE with preprocessing. (a) DOA estimation. (b) Frequency estimation.

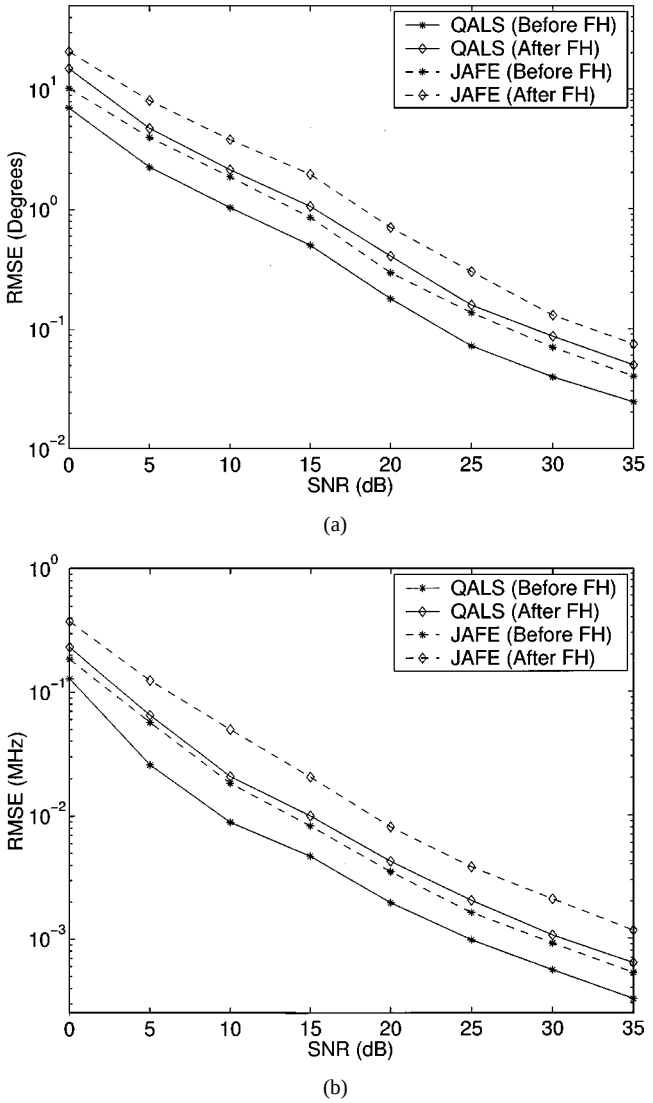


Fig. 6. QALS versus JAFE without preprocessing. (a) DOA estimation. (b) Frequency estimation.

same time. The motivation is that in a heavily loaded asynchronous system, it could be difficult to find a large block of data that is hop free. Reliable detection is also difficult at low SNR. In this case, there are $m_1 - 1$ shift invariance-violating columns for both QALS and JAFE, since spatial smoothing does not introduce additional shift invariance-violating columns. Fig. 6 depicts DOA and frequency estimation results for the hopped source signal both before and after frequency hopping. The figure shows that QALS works reasonably well in the presence of a single undetected hop, and again, it offers a 3–5 dB SNR advantage over JAFE. As expected, the parameter estimation is more accurate before hopping than after hopping since a longer data sample size is available before hopping. In addition, notice that the performance in Fig. 6 is worse and more uneven than that in Fig. 5 which is due to the fact that Fig. 6 involves shift invariance-violating data (unmodeled dynamics).

B. Single-User Tracking

In the third experiment, we test the performance of the proposed single-user tracking method. Suppose we want to track the source signal from DOA $\alpha_3 = 15°$ among the three sources from [5°, 10°, 15°], [11] and we have collected 6×384 data samples (i.e., 3 blocks of data, during each of which every source hops once). Applying the blind localization algorithm (with preprocessing), we obtain an estimate of the steering matrix $\bar{A}$. Then, a linear MMSE beamformer is constructed as in (16) since the hopping bandwidth is less than 1% of the carrier frequency, and thus, the steering vectors can be assumed approximately constant. The signal from $\alpha_3 = 15°$ is separated from the mixture data by applying this beamformer to the whole 6×384 data block. Finally, hop instants and frequencies are jointly estimated by dynamic programming.

An example of single-user tracking is shown in Fig. 7, which is a typical result at SNR $= 15$ dB. The frequency sequence of the original transmitted signal of the desired source is [1.5, 2.25, 6.0, 5.0] MHz, hopped at sampling instants [82, 214, 346]. The figure shows that estimated hop instants and frequencies are very close to the original. If a longer data record needs to be

[11]Recall that with a $M = 6$ element ULA, 5° separation is well under the classical DOA resolution; therefore, the three signals are essentially in the same look direction of the antenna.

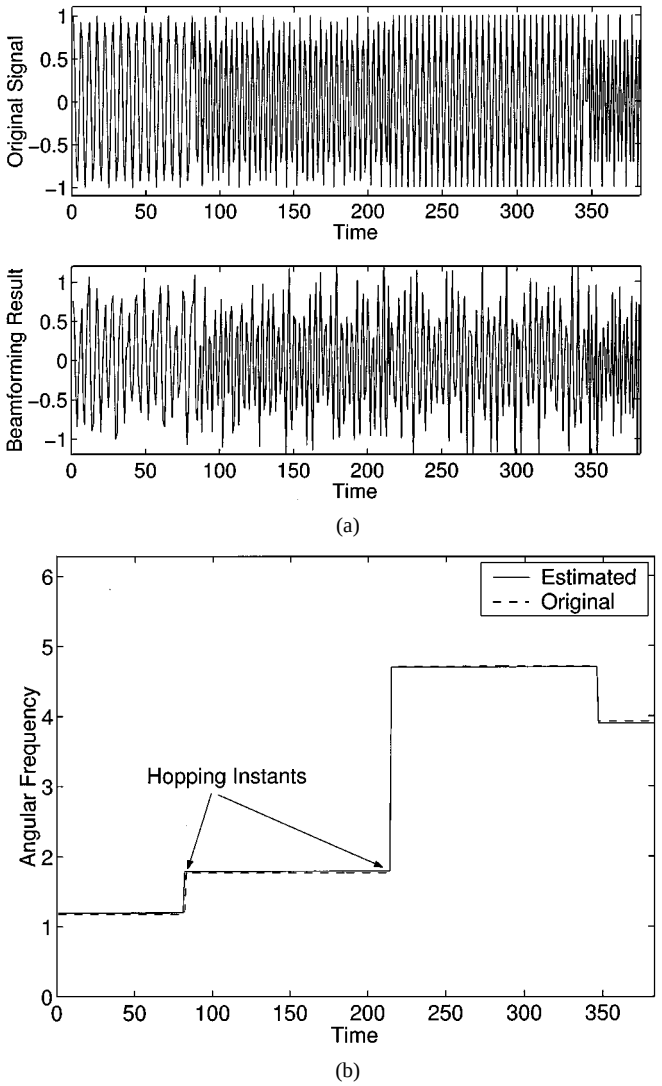


Fig. 7. Example of single-user tracking (SNR $=$ 15 dB). (a) Real part of transmitted signal [top] and beamforming result [bottom]. (b) Dynamic programming joint estimation result.

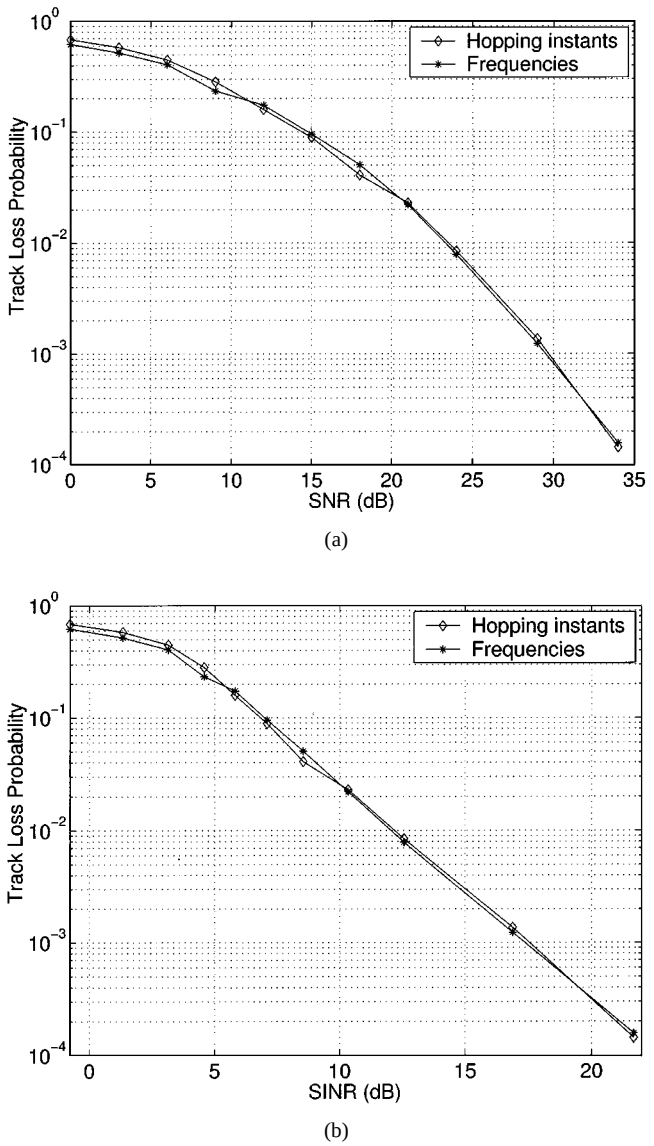


Fig. 8. Track loss probability P_l of single-user tracking. (a) P_l versus SNR. (b) P_l versus SINR.

tracked, we can use (19) and (20) for the remaining data since hop timing and period[12] have been recovered.

In order to evaluate the performance of our single-user tracking algorithm quantitatively, we define *track loss* as a situation wherein an estimated hop instant of the desired source is five samples away from its true value, or an estimated frequency deviates 50 KHz away from its true value.

The track loss probability is estimated by applying the single-user tracking algorithm to 5000 independent realizations of the operating environment; in each realization, the signal from $\alpha_3 = 15^\circ$ contains three hops (four frequencies), with 132 samples between two hops and frequency channel spacing of 0.25 MHz. The track loss probability P_l is depicted in Fig. 8, where (a) is the plot of P_l versus SNR (measured at the antenna array output for mixture data), and (b) is the plot of P_l versus SINR (signal to interference and noise ratio, measured at the beamformer output). The figure shows that the proposed localization and tracking scheme performs quite well.

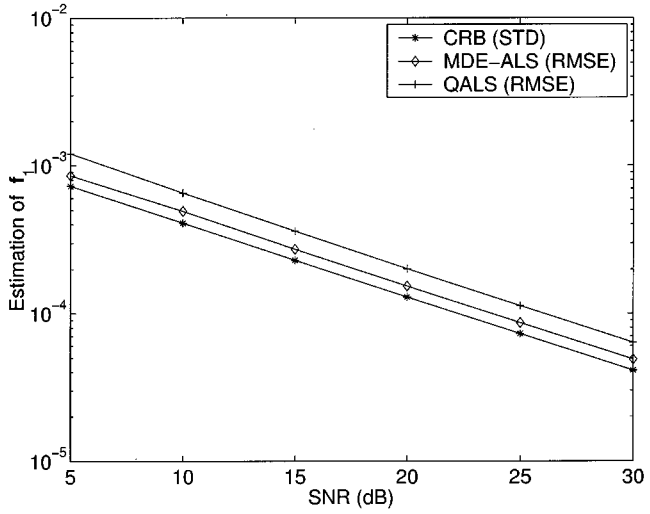
[12]For example, hop period can be obtained by taking the mean of five to ten dwell lengths.

C. Two-Dimensional Harmonic Retrieval

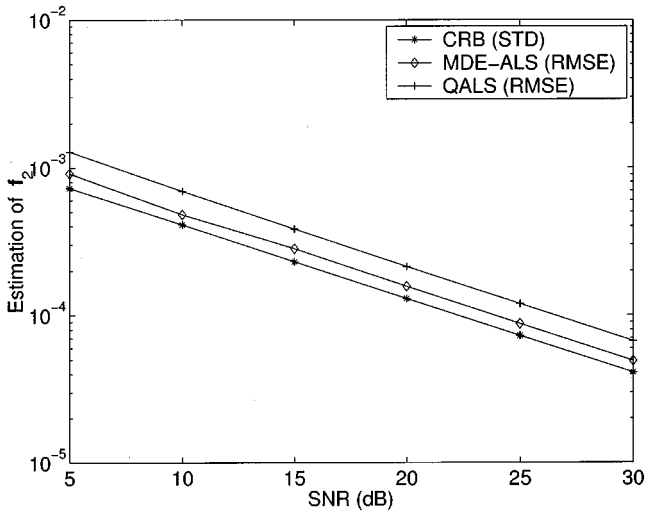
In this experiment, we test the performance of the proposed MDE-ALS algorithm. In [5], Hua derives the CRB for 2-D frequency estimation. We use this CRB to benchmark the performance of our MDE-ALS algorithm. We consider the recovery of three pairs of 2-D harmonics $a_\rho = e^{j2\pi f_{1,\rho}}, b_\rho = e^{j2\pi f_{2,\rho}}$ for $\rho = 1, 2, 3$, where

$$(f_{1,1}, f_{2,1}) = (0.22, 0.15)$$
$$(f_{1,2}, f_{2,2}) = (0.28, 0.21)$$
$$(f_{1,3}, f_{2,3}) = (0.34, 0.30).$$

Sample size along both dimensions is 20. Fig. 9 plots the RMSE achieved by the MDE-ALS and QALS, compared with the corresponding CRB on standard deviation (STD), for the estimation of $\mathbf{f}_1$ and $\mathbf{f}_2$, respectively. It is seen that the MDE-ALS out-



(a)



(b)

Fig. 9. MDE-ALS, QALS, and CRB. (a) Estimation of $\mathbf{f}_1$. (b) Estimation of $\mathbf{f}_2$.

performs QALS (which only takes partial advantage of model structure) and stays close to the CRB for the range of SNR considered.

VII. Conclusion

We have proposed a novel approach for blind localization of frequency hopping signals, without knowledge of their hopping patterns. As a preprocessing step, a hop-free subset may be identified via spectrogram entropy analysis. Following linear MMSE beamforming, a dynamic programming approach was developed for joint optimal ML estimation of hop instants, signal frequencies, and phases in a single-user tracking mode. Simulation results corroborate the high-resolution performance of the localization and tracking algorithms. We also developed a more general 2-D HR algorithm, which was shown to achieve the identifiability bound in [7], while remaining close to the CRB for a wide range of SNRs.

References

[1] L. Aydin and A. Polydoros, "Hop-timing estimation for FH signals using a coarsely channelized receiver," *IEEE Trans. Commun.*, vol. 44, pp. 516–526, Apr. 1996.

[2] M. P. Clark and L. L. Scharf, "Two-dimensional model analysis based on maximum likelihood," *IEEE Trans. Signal Processing*, vol. 42, pp. 1443–1452, June 1994.

[3] I. V. L. Clarkson, "Frequency estimation, phase unwrapping and the nearest lattice point problem," in *Proc. ICASSP*, vol. 3, Phoenix, AZ, Mar. 1999, pp. 1609–1612.

[4] R. G. Gallager, *Information Theory and Reliable Communication*. New York: Wiley, 1968.

[5] Y. Hua, "Estimating two-dimensional frequencies by matrix enhancement and matrix pencil," *IEEE Trans. Signal Processing*, vol. 40, pp. 2267–2280, Sept. 1992.

[6] C. Jauffret and D. Bouchet, "Frequency line tracking on a lofargram: An efficient wedding between probabilistic data association modeling and dynamic programming technique," in *Proc. Asilomar Conf.*, vol. 1, Pacific Grove, CA, Nov. 1996, pp. 486–490.

[7] T. Jiang, N. D. Sidiropoulos, and J. M. F. ten Berge, "Almost-sure identifiability of multidimensional harmonic retrieval," *IEEE Trans. Signal Processing*, vol. 49, pp. 1849–1859, Sept. 2001.

[8] S. Kay, "A fast and accurate single frequency estimator," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 37, pp. 1987–1990, Dec. 1989.

[9] ——, *Fundamentals of Statistical Signal Processing, Volume II: Detection Theory*. Upper Saddle River, NJ: Prentice-Hall, 1998.

[10] J. B. Kruskal, "Three-way arrays: Rank and uniqueness of trilinear decompositions, with application to arithmetic complexity and statistics," *Linear Algebra Appl.*, vol. 18, pp. 95–138, 1977.

[11] A. N. Lemma, A.-J. van der Veen, and E. F. Deprettere, "Joint angle frequency estimation for slow frequency hopping signals," in *Proc. IEEE Workshop Circuits, Syst. Signal Process.*, Mirlo, The Netherlands, Nov. 1998, pp. 363–370.

[12] T. Li and N. D. Sidiropoulos, "Blind digital signal separation using successive interference cancellation iterative least squares," *IEEE Trans. Signal Processing*, vol. 48, pp. 3146–3152, Nov. 2000.

[13] J. Li, P. Stoica, and D. Zheng, "An efficient algorithm for two-dimensional frequency estimation," *Multidimen. Syst. Signal Process.*, vol. 7, no. 2, pp. 151–178, Apr. 1996.

[14] X. Liu and N. D. Sidiropoulos, "Cramér-Rao lower bounds for low-rank decomposition of multidimensional arrays," *IEEE Trans. Signal Processing*, vol. 49, pp. 2074–2086, Sept. 2001.

[15] S. K. Mitra, *Digital Signal Processing: A Computer-Based Approach*. New York: McGraw-Hill, 1998.

[16] L. L. Scharf, "Aspects of dynamic programming in signal and image processing," *IEEE Trans. Automat. Contr.*, vol. AC-26, pp. 1018–1029, Oct. 1981.

[17] N. D. Sidiropoulos and R. Bro, "On the uniqueness of multilinear decomposition of N-way arrays," *J. Chemometr.*, vol. 14, no. 3, pp. 229–239, May 2000.

[18] ——, "PARAFAC techniques for signal separation," in *Signal Processing Advances in Wireless Communications*, P. Stoica, G. Giannakis, Y. Hua, and L. Tong, Eds. Upper Saddle River, NJ: Prentice-Hall, 2000, vol. 2, ch. 4.

[19] N. D. Sidiropoulos and X. Liu, "Identifiability results for blind beamforming in incoherent multipath with small delay spread," *IEEE Trans. Signal Processing*, vol. 49, pp. 228–236, Jan. 2001.

[20] K. Steiglitz and B. Dickinson, "Phase unwrapping by factorization," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-30, pp. 984–991, Dec. 1982.

[21] R. L. Streit and R. F. Barrett, "Frequency line tracking using hidden Markov models," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 38, pp. 586–598, Apr. 1990.

[22] D. J. Torrieri, "Frequency hopping and future army wireless communications," *Proc. IEEE Military Commun. Conf.*, pp. 167–171, 1997.

[23] ——, "Mobile frequency-hopping CDMA systems," *IEEE Trans. Commun.*, vol. 48, pp. 1318–1327, Aug. 2000.

[24] S. A. Tretter, "Estimating the frequency of a noisy sinusoid by linear regression," *IEEE Trans. Inform. Theory*, vol. IT-31, pp. 832–835, Nov. 1985.

[25] M. Wax and I. Ziskind, "On unique localization of multiple sources by passive sensor arrays," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 37, pp. 996–1000, July 1989.

[26] K. T. Wong, "Blind beamforming/geolocation for wideband-FFH's with unknown hop-sequences," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 37, pp. 65–76, Jan. 2001.

Xiangqian Liu (S'99) received the B.S. degree from Beijing Institute of Technology, Beijing, China, in 1997 and the M.S. degree from University of Virginia, Charlottesville, in 1999, both in electrical engineering. He is currently pursuing the Ph.D. degree at the Department of Electrical and Computer Engineering, University of Minnesota, Minneapolis.

His current research interests are in the areas of communications and signal processing, including multiuser detection, blind localization and beamforming, and spread spectrum communication systems.

Nicholas D. Sidiropoulos (M'92–SM'99) received the Diploma in electrical engineering from the Aristotelian University of Thessaloniki, Thessaloniki, Greece, and the M.S. and Ph.D. degrees in electrical engineering from the University of Maryland, College Park (UMCP), in 1988, 1990, and 1992, respectively.

From 1988 to 1992, he was a Fulbright Fellow and a Research Assistant at the Institute for Systems Research (ISR), UMCP. From September 1992 to June 1994, he served his military service as a Lecturer with the Hellenic Air Force Academy. From October 1993 to June 1994, he also was a Member of Technical Staff, Systems Integration Division, G-Systems Ltd., Athens, Greece. He has held Postdoctoral (from 1994 to 1995) and Research Scientist (from 1996 to 1997) positions at ISR-UMCP, before joining the Department of Electrical Engineering, University of Virginia, Charlottesville, in July 1997 as an Assistant Professor. He is currently an Associate Professor with the Department of Electrical and Computer Engineering, University of Minnesota, Minneapolis. His current research interests are primarily in multiway analysis and its applications in signal processing for communications and networking.

Dr. Sidiropoulos is a member of the Signal Processing for Communications Technical Committee (SPCOM-TC) of the IEEE Signal Processing Society and currently serves as Associate Editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING and IEEE SIGNAL LETTERS. He received the NSF/CAREER award (Signal Processing Systems Program) in June 1998.

Ananthram Swami (SM'96) received the B.S. degree from the Indian Institute of Technology, Bombay, the M.S. degree from Rice University, Houston, TX, and the Ph.D. degree from the University of Southern California, Los Angeles, all in electrical engineering.

He has held positions with Unocal Corporation, the University of Southern California, CS-3, and Malgudi Systems. He is currently a Research Scientist with the Communication Networks Branch of the United States Army Research Laboratory, Adelphi, MD, where his work is in the broad area of signal processing for communications. He was a Statistical Consultant to the California Lottery, developed a Matlab-based toolbox for non-Gaussian signal processing, and has held visiting faculty positions at INP, Toulouse, France. He has taught short courses for industry and currently teaches courses on communication theory and signal processing at the University of Maryland.

Dr. Swami a member of the IEEE Signal Processing Society's (SPS) Technical Committee (TC) on Signal Processing for Communications, a member of the IEEE Communication Society's TC on Tactical Communications, and an Associate Editor for the IEEE SIGNAL PROCESSING LETTERS. He was a member of the society's TC on Statistical Signal and Array Processing (from 1993 to 1998); an Associate Editor of the IEEE TRANSACTIONS ON SIGNAL PROCESSING; Vice Chairman of the Orange County Chapter of IEEE-GRS (from 1991 to 1993); and co-organizer and co-chair of the 1993 IEEE SPS Workshop on Higher-Order Statistics, the 1996 IEEE SPS Workshop on Statistical Signal and Array Processing, and the 1999 ASA-IMA Workshop on Heavy-Tailed Phenomena.