

# Elsevier L<sup>A</sup>T<sub>E</sub>X template<sup>☆</sup>

Elsevier<sup>1</sup>

*Radarweg 29, Amsterdam*

*Elsevier Inc<sup>a,b</sup>, Global Customer Service<sup>b,\*</sup>*

*<sup>a</sup>1600 John F Kennedy Boulevard, Philadelphia*

*<sup>b</sup>360 Park Avenue South, New York*

---

## Abstract

*Keywords:*

---

## 1. Problem Formulation

Given  $Q$  adjacent narrow-band emitters and  $L$  space separated base-stations which intercept the transmitted signals. Every base-station is equipped with an antenna array consisting of  $M$  elements. To avoid the angle ambiguity, the carrier frequency is small than the inverse of the propagation time over the array aperture. Denote the  $q$ th emitters position by the vector of coordinates  $\mathbf{p}_q \in \mathbb{R}^{D \times 1}$  (In general,  $D = 2$  for plane geometry or  $D = 3$  for solid geometry). The complex envelopes of the signals observed by the  $j$ th base-station array are given by

$$\mathbf{y}_j(t) = \sum_{q=1}^Q \alpha_{q,j} \mathbf{a}_j(\mathbf{p}_q) s_q(t - \tau_j(\mathbf{p}_q) - \dot{t}_q) + \mathbf{w}_j(t), \quad 0 \leq t \leq T \quad (1)$$

where  $\mathbf{y}_j(t)$  is the  $j$ th received signal,  $\alpha_{q,j}$  is an unknown complex scalar representing the channel attenuation between the  $q$ th emitter and the  $j$ th base-station,  $\mathbf{a}_j(\mathbf{p}_q)$  is the  $j$ th array response to the signal transmitted from position

---

<sup>☆</sup>Fully documented templates are available in the elsarticle package on CTAN.

<sup>\*</sup>Corresponding author

*Email address:* [support@elsevier.com](mailto:support@elsevier.com) (Global Customer Service)

*URL:* [www.elsevier.com](http://www.elsevier.com) (Elsevier Inc)

<sup>1</sup>Since 1880.

$\mathbf{p}_q$ , and  $s_q(t - \tau_j(\mathbf{p}_q) - \mathring{t}_q)$  is the copy of the  $q$ th transmitted signal waveform, which is transmitted at time  $\mathring{t}_q$  and is received by the  $j$ th base-station after being delayed by  $\tau_j(\mathbf{p}_q)$ . The vector  $\mathbf{w}_j(t)$  represents noise and interference observed by the  $j$ th array. The received signal can be partitioned into  $L$  sections and the length of each section satisfies  $T/L \gg \max \tau_j(\mathbf{p}_q)$  to avoid the delay ambiguity. The emitters are assumed to be stationary during the total observation time  $T$ . Each section can be Fourier transformed and the result of this process is given by the following equation:

$$\mathbf{y}_j(n, l) = \sum_{q=1}^Q \alpha_{q,j} \mathbf{a}_j(\mathbf{p}_q) e^{-j2\pi f_n(\tau_j(\mathbf{p}_q) + \mathring{t}_q)} s_q(n, l) + \mathbf{w}_j(n, l) \quad (2)$$

$$j = 1, \dots, N_r; n = 1, \dots, N; l = 1, \dots, L$$

where  $\mathbf{y}_j(n, l)$  is the Fourier coefficient of the  $l$ th section of the  $j$ th received signal corresponding to frequency  $f_n$ ,  $s_q(n, l)$  is the  $n$ th Fourier coefficient of the  $l$ th section of the  $q$ th transmitted signal waveform, and  $\mathbf{w}_j(n, l)$  represents the  $n$ th Fourier coefficient of the  $l$ th section of the noise waveform.

For compact representation, we define the following vectors:

$$\boldsymbol{\gamma}_{j,q,n,l}(\mathbf{p}_q, \mathring{t}_q | s_q(n, l)) \triangleq \mathbf{a}_j(\mathbf{p}_q) e^{-j2\pi f_n(\tau_j(\mathbf{p}_q) + \mathring{t}_q)} s_q(n, l) \quad (3)$$

Here we assume that the spectrum of transmitted signals, like the synchronization and training sequences in cellular systems, are known to location system but the transmitted time  $\mathring{t}_q$  is unknown (which is assumed known in [1]). We observe that all information about the emitters position is embedded in the vector  $\boldsymbol{\gamma}_{j,q,n,l}$ . This leads to the following representation of (2):

$$\mathbf{y}_j(n, l) = \sum_{q=1}^Q \alpha_{q,j} \boldsymbol{\gamma}_{j,q,n,l} + \mathbf{w}_j(n, l) \quad (4)$$

In matrix notation, (4) becomes

$$\mathbf{y}_j(n, l) = \boldsymbol{\Gamma}_{j,n,l}(\mathbf{P}, \mathring{\mathbf{t}}) \boldsymbol{\alpha}_j + \mathbf{w}_j(n, l) \quad (5)$$

where

$$\begin{aligned}\mathbf{\Gamma}_{j,n,l}(\mathbf{P}, \mathbf{\hat{t}}) &\triangleq [\gamma_{j,q,n,l}, \dots, \gamma_{j,q,n,l}] \in \mathbb{C}^{M \times Q} \\ \boldsymbol{\alpha}_j &\triangleq [\alpha_{1,j}, \dots, \alpha_{Q,j}]^T \in \mathbb{C}^{Q \times 1}\end{aligned}\tag{6}$$

$\mathbf{P}, \mathbf{\hat{t}}$  are the sets of emitters' positions and transmitted times respectively. Collecting and catenating all the  $N$  coefficients of the  $l$ th section of the  $j$ th received signal,

$$\mathbf{y}_j(l) = \mathbf{\Gamma}_{j,l}(\mathbf{P}, \mathbf{\hat{t}})\boldsymbol{\alpha}_j + \mathbf{w}_j(l)\tag{7}$$

where

$$\begin{aligned}\mathbf{y}_j(l) &\triangleq [\mathbf{y}_j^T(1, l), \dots, \mathbf{y}_j^T(N, l)]^T \in \mathbb{C}^{NM \times 1} \\ \mathbf{\Gamma}_{j,l}(\mathbf{P}, \mathbf{\hat{t}}) &\triangleq [\mathbf{\Gamma}_{j,1,l}^T(\mathbf{P}, \mathbf{\hat{t}}), \dots, \mathbf{\Gamma}_{j,N,l}^T(\mathbf{P}, \mathbf{\hat{t}})]^T \in \mathbb{C}^{NM \times Q} \\ \mathbf{w}_j(l) &\triangleq [\mathbf{w}_j^T(1, l), \dots, \mathbf{w}_j^T(N, l)]^T \in \mathbb{C}^{NM \times 1}\end{aligned}\tag{8}$$

The problem that we address now is stated as follow: Given the measurements  $\{\mathbf{y}_j(l)\}_{j,l=1}^{N_r, L}$ , how to efficiently and accurately estimate the locations of the emitters  $\mathbf{P}$ .

## 2. the Concentrated Likelihood Function

In this subsection, we will derive the *Concentrated* Likelihood Function (CLF) that depends on the parameters of interest  $\mathbf{P}$  and the auxiliary parameters  $\mathbf{\hat{t}}$ . The transmitted times are conducive to acquiring the location information from the propagation delay  $\tau_j(\mathbf{p})$  since the transmitted times are the same among base-stations. In fact, since  $\mathbf{w}_j(l)$  are independent among different sections of different received signals and  $\mathbf{w}_j(l) \sim N(\mathbf{0}, \sigma^2 \mathbf{I}_{NM})$ , it can be shown that the Log Likelihood Function (LLF) after dropping the constant terms is given by,

$$\mathcal{L}(\mathbf{P}, \mathbf{\hat{t}}, \boldsymbol{\alpha}) \triangleq - \sum_{j=1}^{N_r} \sum_{l=1}^L \|\mathbf{y}_j(l) - \mathbf{\Gamma}_{j,l}(\mathbf{P}, \mathbf{\hat{t}})\boldsymbol{\alpha}_j\|^2\tag{9}$$

where  $\mathbf{P}, \mathbf{t}, \boldsymbol{\alpha}$  stand for any candidate values. Now, define the matrix

$$\mathbf{\Gamma}_j(\mathbf{P}, \mathbf{t}) \triangleq [\mathbf{\Gamma}_{j,1}^T, \dots, \mathbf{\Gamma}_{j,L}^T]^T \in \mathbb{C}^{LNM \times Q} \quad (10)$$

it can be shown that (11) is equivalent to

$$\mathcal{L}(\mathbf{P}, \mathbf{t}, \boldsymbol{\alpha}) = - \sum_{j=1}^{N_r} \|\mathbf{y}_j - \mathbf{\Gamma}_j(\mathbf{P}, \mathbf{t})\boldsymbol{\alpha}_j\|^2 \quad (11)$$

where  $\mathbf{y}_j \triangleq [\mathbf{y}_j^T(1), \dots, \mathbf{y}_j^T(L)]^T$ . Maximizing (11) jointly with respect to all parameters is extremely challenging. Yet, significant computational savings follow from the observation that for any given  $\mathbf{P}, \mathbf{t}$ , the problem of finding the optimal  $\{\boldsymbol{\alpha}_j\}_{j=1}^{N_r}$  becomes a linear least squares (LS) problem [2] whose solution is given by

$$\hat{\boldsymbol{\alpha}}_j = \mathbf{\Gamma}_j^\dagger(\mathbf{P}, \mathbf{t})\mathbf{y}_j \quad (12)$$

where  $\mathbf{\Gamma}_j^\dagger$  is the Moore-Penrose pseudo-inverse of  $\mathbf{\Gamma}_j$  given by  $(\mathbf{\Gamma}_j^H \mathbf{\Gamma}_j)^{-1} \mathbf{\Gamma}_j^H$ . Note here that  $\mathbf{\Gamma}_j$  has full column rank and, therefore,  $(\mathbf{\Gamma}_j^H \mathbf{\Gamma}_j)^{-1}$  always exists. Now, by substituting  $\{\hat{\boldsymbol{\alpha}}_j\}_{j=1}^{N_r}$  for  $\boldsymbol{\alpha}$  back in (11) and resorting to some straightforward algebraic manipulations, we obtain the so-called *concentrated* likelihood function,

$$\mathcal{L}_c(\mathbf{P}, \mathbf{t}) = \sum_{j=1}^{N_r} \mathbf{y}_j^H \mathbf{\Gamma}_j (\mathbf{\Gamma}_j^H \mathbf{\Gamma}_j)^{-1} \mathbf{\Gamma}_j^H \mathbf{y}_j \quad (13)$$

The joint Maximum Likelihood (ML) estimates of  $\mathbf{P}$  and  $\mathbf{t}$  are then obtained as the solution to the following reduced-dimension optimization problem

$$[\hat{\mathbf{P}}, \hat{\mathbf{t}}] = \arg \max_{\mathbf{P}, \mathbf{t}} \mathcal{L}_c(\mathbf{P}, \mathbf{t}) \quad (14)$$

- 10 The reduced-dimension optimization problem (14) still has  $Q(D+1)$  unknown parameters (including the  $D$ -dimensional real-valued location vectors  $\{\mathbf{p}_q\}_{q=1}^Q$  and  $Q$  transmitted times  $\{t_q\}_{q=1}^Q$ ) remain to estimate. The ML estimation will therefore requires a complex multidimensional search over the parameter space.

To solve the multidimensional optimization problem (14), the approach proposed in [1] is to decouple the problem into several  $D$ -dimensional optimization

15

problems under the assumption that the number of sections  $L \rightarrow \infty$  or  $L$  is sufficiently large in practice. Considering the length of section has low limit  $T/L \gg \max \tau_j(\mathbf{p}_q)$ , the number of sections  $L$  is proportional to the total observation time  $T$ . The observation time  $T$  is, however, limited in some practical reasons, like the request for real-time processing or the assumption that emitters are stationary during the observation. It is, therefore, difficult to be satisfied that  $L$  is sufficiently large in some practical cases, which means the decouple approach proposed in [1] may be failure, especially when the emitters are adjacent to one another.

### 3. Global Maximization of The CLF

As done previously within other estimation problems (see [3, 4, 5, 6, 7]), we resort to the optimization framework, which is based on the Pincus' theorem and the powerful Importance Sampling (IS) method, to efficiently and accurately estimate the adjacent emitters' positions using limited observation time. First of all, we introduce the theorem of Pincus [8] as follows.

**Theorem 1.** Let  $F(x_1, \dots, x_n) = F(\mathbf{x})$  be a continuous function on the closure of a bounded domain  $S$  in  $n$ -dimensional Euclidean space of  $\mathbb{R}^n$ . Assume that  $F$  attains a global maximum at exactly on point  $\hat{\mathbf{x}} = [\hat{x}_1, \dots, \hat{x}_n]^T$  of  $S$ . Then for  $i = 1, \dots, n$ ,

$$\hat{x}_i = \lim_{\rho \rightarrow \infty} \frac{\int \dots \int x_i e^{\rho F(\mathbf{x})} dx_1 \dots dx_n}{\int \dots \int e^{\rho F(\mathbf{x})} dx_1 \dots dx_n} \quad (15)$$

Applying this general result to our multidimensional optimization problem (14) with  $\mathbf{x} \triangleq [\mathbf{P}, \mathbf{t}]$  and  $F(\mathbf{x}) \triangleq \mathcal{L}_c(\mathbf{P}, \mathbf{t})$  leads to the following expressions for the required ML estimators,

$$\hat{\mathbf{p}}_q(d) = \int \dots \int \mathbf{p}_q(d) \mathcal{F}(\mathbf{P}, \mathbf{t}) d\mathbf{p}_1(1) \dots d\mathbf{t}_Q \quad (16)$$

$$\hat{t}_q^{(0)} = \int \dots \int t_q \mathcal{F}(\mathbf{P}, \mathbf{t}) d\mathbf{p}_1(1) \dots d\mathbf{t}_Q \quad (17)$$

$$q = 1, \dots, Q; d = 1, \dots, D$$

where

$$\mathcal{F}(\mathbf{P}, \mathring{\mathbf{t}}) \triangleq \lim_{\rho \rightarrow \infty} \frac{e^{\rho \mathcal{L}_c(\mathbf{P}, \mathring{\mathbf{t}})}}{\int \dots \int e^{\rho \mathcal{L}_c(\mathbf{P}, \mathring{\mathbf{t}})} d\mathbf{p}_1(1) \dots d\mathring{t}_Q} \quad (18)$$

It seems that the multidimensional grid search we attempt to avoid is transformed into several multidimensional integrations (16) and (17) bearing the same computational load. By closely analyzing the properties of function (18), however, it turns out that  $\mathcal{F}(\mathbf{P}, \mathring{\mathbf{t}})$  can be seen as a Probability Density Function (pdf) since it is nonnegative and integrates to one. As a pdf, furthermore,  $\mathcal{F}(\mathbf{P}, \mathring{\mathbf{t}})$  indicates the probability of the candidate values being true. The estimators in (16) and (17) can be regarded as statistical expectations, i.e., we have,

$$\hat{\mathbf{p}}_q(d) \triangleq \mathbb{E}_{\mathbf{P}, \mathring{\mathbf{t}}} \{\mathbf{p}_q(d)\} \quad \hat{t}_q \triangleq \mathbb{E}_{\mathbf{P}, \mathring{\mathbf{t}}} \{\mathring{t}_q\} \quad (19)$$

Consequently, the multidimensional integrations are transformed into the statistical expectations, which can be approximated by the sample mean. We will then explain the method to generate  $R$  realizations,  $\{\mathbf{P}^{(r)}\}_{r=1}^R$  and  $\{\mathring{\mathbf{t}}^{(r)}\}_{r=1}^R$ , using the joint pdf  $\mathcal{F}(\mathbf{P}, \mathring{\mathbf{t}})$  and approximate the expectations in (19) using the following sample means,

$$\hat{\mathbf{p}}_q(d) = \frac{1}{R} \sum_{r=1}^R \mathbf{p}_q^{(r)}(d) \quad \hat{t}_q = \frac{1}{R} \sum_{r=1}^R \mathring{t}_q^{(r)} \quad (20)$$

Clearly, as the number of realizations  $R$  increases, the two sample mean in (20) approach the statistical expectations in (19). Unfortunately, the extremely non-linear property of the pdf  $\mathcal{F}(\mathbf{P}, \mathring{\mathbf{t}})$  makes it impractical to generate  $\{\mathbf{P}^{(r)}\}_{r=1}^R$  and  $\{\mathring{\mathbf{t}}^{(r)}\}_{r=1}^R$  using itself. To sidestep this problem, we can resort to the Importance Sampling (IS) method [3, 4] and rewrite (16) and (17) in the following equivalent forms,

$$\hat{\mathbf{p}}_q(d) = \int \dots \int \mathbf{p}_q(d) \frac{\mathcal{F}(\mathbf{P}, \mathring{\mathbf{t}})}{\mathcal{G}(\mathbf{P}, \mathring{\mathbf{t}})} \mathcal{G}(\mathbf{P}, \mathring{\mathbf{t}}) d\mathbf{p}_1(1) \dots d\mathring{t}_Q \quad (21)$$

$$\hat{t}_q = \int \dots \int \mathring{t}_q \frac{\mathcal{F}(\mathbf{P}, \mathring{\mathbf{t}})}{\mathcal{G}(\mathbf{P}, \mathring{\mathbf{t}})} \mathcal{G}(\mathbf{P}, \mathring{\mathbf{t}}) d\mathbf{p}_1(1) \dots d\mathring{t}_Q \quad (22)$$

where  $\mathcal{G}(\mathbf{P}, \mathring{\mathbf{t}})$  is another pdf called *importance function* (IF). By introducing the IF, the multidimensional integrations in (21) and (22) are interpreted as statistical expectations of two transformed random variates (RVs), i.e.,

$$\hat{\mathbf{p}}_q(d) \triangleq \mathbb{E}_{\mathbf{P}, \mathring{\mathbf{t}}} \{ \eta(\mathbf{P}, \mathring{\mathbf{t}}) \mathbf{p}_q(d) \} \quad \hat{t}_q \triangleq \mathbb{E}_{\mathbf{P}, \mathring{\mathbf{t}}} \{ \eta(\mathbf{P}, \mathring{\mathbf{t}}) \hat{t}_q \} \quad (23)$$

where  $\eta(\mathbf{P}, \mathring{\mathbf{t}})$  is called *importance weight* (IW) corresponding to the realizations of  $\mathbf{P}, \mathring{\mathbf{t}}$  and defined as,

$$\eta(\mathbf{P}, \mathring{\mathbf{t}}) \triangleq \frac{\mathcal{F}(\mathbf{P}, \mathring{\mathbf{t}})}{\mathcal{G}(\mathbf{P}, \mathring{\mathbf{t}})} \quad (24)$$

If  $\mathcal{G}(\mathbf{P}, \mathring{\mathbf{t}})$  is carefully designed, the expectations in (23) can be computed at any desired degree of accuracy (by increasing  $R$ ) using the corresponding sample mean estimates. In fact, the choice of  $\mathcal{G}(\mathbf{P}, \mathring{\mathbf{t}})$  is a tradeoff between the ease and efficiency of sampling. The appropriate choice of the IF will be discussed in the following section.

#### 4. Appropriate Choice of Importance Function

The appropriate  $\mathcal{G}(\mathbf{P}, \mathring{\mathbf{t}})$  must be designed as close as possible to  $\mathcal{F}(\mathbf{P}, \mathring{\mathbf{t}})$  while allowing at the same time the easy generation of the required parameters realizations  $\{\mathbf{P}^{(r)}\}_{r=1}^R$  and  $\{\mathring{\mathbf{t}}^{(r)}\}_{r=1}^R$ . In order to facilitate the process of generating the required  $R$  parameter realizations, we constrain IF's form separable in terms of the  $Q$  position-transmitted time pairs  $\{(\mathbf{p}_q, \mathring{t}_q)\}_{q=1}^Q$ . In other words, our ultimate goal is to design  $\mathcal{G}(\mathbf{P}, \mathring{\mathbf{t}})$  in a way that allows it to be factorized as follows,

$$\mathcal{G}(\mathbf{P}, \mathring{\mathbf{t}}) \triangleq \prod_{q=1}^Q g_q(\mathbf{p}_q, \mathring{t}_q) \quad (25)$$

This allow us to interpret the pdf of  $Q(D+1)$ -dimensional random vector as the joint pdf of  $Q$  independent  $(D+1)$ -dimensional random vectors. Hence, we can independently generate  $Q$  groups of realizations for  $(D+1)$ -dimensional random vectors  $\{(\mathbf{p}_q^{(r)}, \mathring{t}_q^{(r)})\}_{r,q=1}^{R,Q}$ , instead of generating realizations for  $Q(D+1)$ -dimensional random vector  $(\mathbf{P}, \mathring{\mathbf{t}})$  directly using  $\mathcal{G}(\mathbf{P}, \mathring{\mathbf{t}})$ .

We start with the CLF (13), the separable form (25) requests that the CLF has summation form with respect to  $q$ . Note that the product  $\mathbf{y}_j^H \mathbf{\Gamma}_j$  in CLF is a  $Q$ -dimensional vector. Diagonalizing the inverse matrix  $(\mathbf{\Gamma}_j^H \mathbf{\Gamma}_j)^{-1}$  is, therefore, the key step to the separable form (25). According to the definition of  $\mathbf{\Gamma}_j$  in (10),  $\mathbf{\Gamma}_{j,l}$  in (8),  $\mathbf{\Gamma}_{j,n,l}$  in (6) and  $\gamma_{j,q,n,l}$  in (3), the  $(u, v)$ th element of the matrix product  $\mathbf{\Gamma}_j^H \mathbf{\Gamma}_j$  is given by,

$$\begin{aligned} [\mathbf{\Gamma}_j^H \mathbf{\Gamma}_j]_{u,v} &= \sum_{l=1}^L \sum_{n=1}^N \gamma_{j,q,n,l}^H \gamma_{j,q,n,l} \\ &= \underbrace{\mathbf{a}_j^H(\mathbf{p}_u) \mathbf{a}_j(\mathbf{p}_u)}_{r_{doa,j}^{(u,v)}} \underbrace{\sum_{n=1}^N e^{j2\pi f_n(\tau_j(\mathbf{p}_u) + \hat{t}_u - \tau_j(\mathbf{p}_v) - \hat{t}_v)}}_{r_{toa,j}^{(u,v)}} \underbrace{\sum_{l=1}^L s_u^*(n, l) s_v(n, l)}_{r_s^{(u,v)}} \end{aligned}$$

where  $r_{doa,j}^{(u,v)}, r_{toa,j}^{(u,v)}, r_s^{(u,v)}$  denote the  $(u, v)$ th non-normalized correlation coefficients in aspects of Direction Of Arrival (DOA) and Time Of Arrival (TOA) at the  $j$ th base-station and the transmitted signal, respectively. The matrix  $\mathbf{\Gamma}_j^H \mathbf{\Gamma}_j$  is, therefore, a correlation matrix among all emitters. To avoid additional computational load, we directly assume, instead of strict diagonalizing,  $\mathbf{\Gamma}_j^H \mathbf{\Gamma}_j$ , as diagonal matrix, at the cost of losing the correlation information. The diagonal matrix, for  $j = 1, \dots, N_r$ , is given by,

$$\mathbf{\Gamma}_j^H \mathbf{\Gamma}_j = MNL \cdot \mathbf{I} \quad (26)$$

where  $\mathbf{I}$  is the identity matrix, the coefficient  $MNL$  is the diagonal element of  $\mathbf{\Gamma}_j^H \mathbf{\Gamma}_j$ .

substituting (26) into (13) and ignoring the constant coefficient, we get its separated form as follows,

$$\tilde{\mathcal{L}}_c(\mathbf{P}, \mathbf{t}) = \sum_{j=1}^{N_r} \|\mathbf{\Gamma}_j^H \mathbf{y}_j\|^2 \quad (27)$$

$$= \sum_{q=1}^Q \tilde{\ell}_c(\mathbf{p}_q, \mathbf{t}_q) \quad (28)$$



where

$$\tilde{\ell}_c(\mathbf{p}_q, \hat{t}_q) \triangleq \sum_{j,n,l=1}^{N_r, N, L} |\gamma_{j,q,n,l}^H \mathbf{y}_j(n, l)|^2 \quad (29)$$

Replacing  $\mathcal{L}_c(\mathbf{P}, \hat{\mathbf{t}})$  and  $\rho$  in (18) with  $\tilde{\mathcal{L}}_c(\mathbf{P}, \hat{\mathbf{t}})$  and an appropriate  $\rho_2$ , we get the separated form of  $\mathcal{G}(\mathbf{P}, \hat{\mathbf{t}})$  (25) and the  $q$ th component is given by,

$$g_q(\mathbf{p}_q, \hat{t}_q) \triangleq \frac{e^{\rho_2 \tilde{\ell}_c(\mathbf{p}_q, \hat{t}_q)}}{\int \int e^{\rho_2 \tilde{\ell}_c(\mathbf{p}_q, \hat{t}_q)} d\mathbf{p}_q d\hat{t}_q} \quad (30)$$

The generation of  $q$ th group of realizations  $\{(\mathbf{p}_q^{(r)}, \hat{t}_q^{(r)})\}_{r=1}^R$  follows, furthermore, the well-known general result from probability theory, which is that the joint pdf,  $g_q(\mathbf{p}_q, \hat{t}_q)$ , can be factorized as the product of marginal and conditional pdfs,

$$g_q(\mathbf{p}_q, \hat{t}_q) = g_q^{\text{mar}}(\mathbf{p}_q) g_q^{\text{con}}(\hat{t}_q | \mathbf{p}_q) \quad (31)$$

The usual method to obtain the marginal pdf  $g_q^{\text{mar}}(\mathbf{p}_q)$  is integrating the joint pdf  $g_q(\mathbf{p}_q, \hat{t}_q)$  with respect to the transmitted time  $\hat{t}_q$ . The analytical integration is, however, difficult since  $g_q(\mathbf{p}_q, \hat{t}_q)$  has very high non-linear about  $\hat{t}_q$ . The discrete approximation is also involving  $D + 1$ -dimensional grid search. We have another alternative method to construct  $g_q^{\text{mar}}(\mathbf{p}_q)$  directly and efficiently,

$$g_q^{\text{mar}}(\mathbf{p}_q) \triangleq \frac{e^{\rho_1 \ell(\mathbf{p}_q)}}{\int e^{\rho_1 \ell(\mathbf{p}_q)} d\mathbf{p}_q} \quad (32)$$

where  $\ell(\mathbf{p}_q)$  is the *single* ( $Q = 1$ ) Likelihood Function (SLF) of the position  $\mathbf{p}_q$ , which is derived from (29) by eliminating the transmitted time  $\hat{t}_q$ . We will, then, describe the derivation in details.

Our derivation starts with a vector definition as follows,

$$\gamma_{q,j,l} \triangleq \Phi_{q,j,l}(\mathbf{p}_q) \mathbf{b}(\hat{t}_q) \quad (33)$$

where

$$\Phi_{q,j,l}(\mathbf{p}_q) \triangleq \Lambda_j(\mathbf{p}_q, s_q(l)) \otimes \mathbf{a}_j(\mathbf{p}_q) \quad (34)$$

$$\mathbf{b}(\mathring{t}_q) \triangleq [e^{-j2\pi f_1 \mathring{t}_q}, \dots, e^{-j2\pi f_N \mathring{t}_q}]^T \quad (35)$$

$$\Lambda_j(\mathbf{p}_q, s_q(l)) \triangleq \begin{bmatrix} s_q(1, l) e^{-j2\pi f_1 \tau_j(\mathbf{p}_q)} & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & s_q(N, l) e^{-j2\pi f_N \tau_j(\mathbf{p}_q)} \end{bmatrix} \quad (36)$$

substituting (33) and (34) into (29), we get

$$\tilde{\ell}_c(\mathbf{p}_q, \mathring{t}_q) = \sum_{j=1}^{N_r} \sum_{l=1}^L \gamma_{j,q,l}^H \mathbf{y}_j(l) \mathbf{y}_j^H(l) \gamma_{j,q,l} \quad (37)$$

$$= \mathbf{b}^H(\mathring{t}_q) \underbrace{\left( \sum_{j=1}^{N_r} \sum_{l=1}^L \Phi_{q,j,l}^H(\mathbf{p}_q) \mathbf{y}_j(l) \mathbf{y}_j^H(l) \Phi_{q,j,l}(\mathbf{p}_q) \right)}_{\triangleq \mathbf{\Pi}(\mathbf{p}_q)} \mathbf{b}(\mathring{t}_q) \quad (38)$$

The right hand of the second equation is a non-normalized Rayleigh Quotient with respect to  $\mathring{t}_q$ . Consequently, maximizing  $\tilde{\ell}_c(\mathbf{p}_q, \mathring{t}_q)$  with respect to  $\mathring{t}_q$ , we get,

$$\ell(\mathbf{p}_q) \triangleq \lambda_{\max}(\mathbf{\Pi}(\mathbf{p}_q)) \quad (39)$$

Substituting (39) into (32), we get the marginal pdf of the  $q$ th emitter's position,

$$g_q^{\text{mar}}(\mathbf{p}_q) = \frac{e^{\rho_1 \lambda_{\max}(\mathbf{\Pi}(\mathbf{p}_q))}}{\int e^{\rho_1 \lambda_{\max}(\mathbf{\Pi}(\mathbf{p}_q))} d\mathbf{p}_q} \quad (40)$$

Given position realization  $\mathbf{p}_q^{(r)}$  generated by  $g_q^{\text{mar}}(\mathbf{p}_q)$ , the conditional pdf is defined as,

$$g_q^{\text{con}}(\mathring{t}_q | \mathbf{p}_q^{(r)}) = \frac{g_q(\mathbf{p}_q^{(r)}, \mathring{t}_q)}{\int_{\mathring{t}} g_q(\mathbf{p}_q^{(r)}, \mathring{t}_q) d\mathring{t}} \quad (41)$$

Substituting (30) into (41), we have,

$$g_q^{\text{con}}(\mathring{t}_q | \mathbf{p}_q^{(r)}) = \frac{e^{\rho_2 \tilde{\ell}_c(\mathbf{p}_q^{(r)}, \mathring{t}_q)}}{\int_{\mathring{t}} e^{\rho_2 \tilde{\ell}_c(\mathbf{p}_q^{(r)}, \mathring{t}_q)} d\mathring{t}} \quad (42)$$

Note that we use  $\int_{\tilde{t}} e^{\rho_2 \tilde{\ell}_c(\mathbf{p}_q^{(r)}, \tilde{t}_q)} d\tilde{t}$ , instead of  $g_q^{\text{mar}}(\mathbf{p}_q^{(r)})$ , to normalize the conditional pdf (41) correctly.

The process of generating the required realizations  $\{(\mathbf{p}_q^{(r)}, \tilde{t}_q^{(r)})\}_{r,q=1}^{R,Q}$  are as follows: for  $q$ th group of realizations, generate  $\{\mathbf{p}_q^{(r)}\}_{r=1}^R$  using the marginal pdf  $g_q^{\text{mar}}(\mathbf{p}_q)$  (40) and then use the conditional pdf  $g_q^{\text{con}}(\tilde{t}_q|\mathbf{p}_q^{(r)})$  (42) to generate  $\{\tilde{t}_q^{(r)}\}_{r=1}^R$ .

## 5. Generations Of The Required Realizations

We recall the following lemma [9] that will be used to generate the required realizations:

**Lemma 1.** Let  $X \in \mathcal{X}$  be any RV with pdf  $f_X(x)$  and the Cumulative Distribution Function (CDF)  $F_X(x)$  and denote the inverse CDF as  $F_X^{-1}(\cdot) : [0, 1] \rightarrow \mathcal{X}, u \rightarrow x$  s.t.  $F_X(x) = u$ . Then, for any uniform RV,  $U \in [0, 1]$ , the RV  $\tilde{X} = F_X^{-1}(U)$  is distributed according to  $f_X(\cdot)$ .

According to the generating process described before, firstly, we should use the marginal pdf  $g_q^{\text{mar}}(\mathbf{p}_q)$  (40) to generate  $\{\mathbf{p}_q^{(r)}\}_{r=1}^R$ .  $g_q^{\text{mar}}(\mathbf{p}_q)$  is, however, unsuitable to be used along with the result of Lemma 1 since  $\mathbf{p}_q$  is a  $D$ -dimensional vector (defining  $\mathbf{p}_q = (x_q, y_q)$  for  $D = 2$ ) instead of a scalar. To acquire the  $D = 2$  scalar pdfs, consequently,  $g_q^{\text{mar}}(\mathbf{p}_q)$  should be factorized as follows,

$$g_q^{\text{mar}}(x_q, y_q) \triangleq g_x^{\text{mar}}(x_q) g_{y|x}^{\text{con}}(y_q|x_q) \quad (43)$$

where

$$g_x^{\text{mar}}(x_q) \triangleq \int g_q^{\text{mar}}(x_q, y_q) dy_q \quad (44)$$

$$g_{y|x}^{\text{con}}(y_q|x_q) \triangleq \frac{g_q^{\text{mar}}(x_q, y_q)}{g_x^{\text{mar}}(x_q)} \quad (45)$$

In conclusion, the approach of generating the  $q$ th group of realizations  $\{(\mathbf{p}_q^{(r)}, \tilde{t}_q^{(r)})\}_{r=1}^R$  can be stated as follows: generate  $R$  realizations  $\{u_x^{(r)}\}_{r=1}^R \sim U[0, 1]$  and obtain  $x_q^{(r)} = G_x^{-1}(u_x^{(r)})$  where  $G_x(\cdot)$  is the CDF associated to  $g_x^{\text{mar}}(x_q)$  in (44); For  $r = 1, \dots, R$ , then, generate two realizations  $u_y^{(r)}, u_{\tilde{t}}^{(r)} \sim U[0, 1]$  and obtain

65  $y_q^{(r)} = G_{y|x_q}^{-1}(u_y^{(r)})$  and  $\hat{t}_q^{(r)} = G_{\hat{t}|\mathbf{p}_q^{(r)}}^{(-1)}(u_{\hat{t}}^{(r)})$  where  $G_{y|x}(\cdot)$  and  $G_{\hat{t}|\mathbf{p}}(\cdot)$  are the CDF associated to  $g_{y|x}^{\text{con}}(y_q|x_q)$  in (44) and  $g_q^{\text{con}}(\hat{t}_q|\mathbf{p}_q^{(r)})$  in (41) respectively;

## 6. Implementation Details

In this section, we give all the necessary details for an efficient implementation of our proposed IS-based ML DPD algorithm. To narrow the range of sampling, we would like to generate the realizations around the true parameters. The area of interest, where emitters may be put in, is confined within  $x \in [x_{\min}, x_{\max}]$  and  $y \in [y_{\min}, y_{\max}]$  and the range of transmitted time is  $\hat{t} \in [0, \hat{t}_{\max}]$ , where  $\hat{t}_{\max}$  can be freely chosen as high as desired. The starting point of the algorithm is, therefore, the evaluation of the SLF (39) at multiple grid points  $\{(x^{(i)}, y^{(j)})\}_{i,j=1}^{I,J}$  with relatively large step  $\Delta x$  and  $\Delta y$ . Then, by approximating integrals with discrete sums, we evaluate the  $q$ th marginal pdf with respect to  $\mathbf{p} = (x, y)$  in (32),  $g_q^{\text{mar}}(x, y)$ , at every grid point as follows,

$$g_q^{\text{mar}}(x^{(i)}, y^{(j)}) \simeq \frac{e^{\rho_1 \ell(x^{(i)}, y^{(j)})}}{\sum_{i=1}^I \sum_{j=1}^J e^{\rho_1 \ell(x^{(i)}, y^{(j)})} \Delta x \Delta y} \quad (46)$$

The estimation of  $q$ th emitter position can be initialized as the grid point at which the maximum peak of  $g_q^{\text{mar}}(x^{(i)}, y^{(j)})$  is. The initialization can be written as,

$$[\hat{x}_q, \hat{y}_q] = \arg \max_{(x^{(i)}, y^{(j)})} g_q^{\text{mar}}(x^{(i)}, y^{(j)}), \quad q = 1, \dots, Q \quad (47)$$

The distance between the ML estimator  $\hat{\mathbf{p}}_q$  and the initial estimator  $\mathring{\mathbf{p}}_q$ , however, is non-negligible since the decouple processing in (27) lose the correlation information among emitters. To force  $\{(\mathbf{p}_q^{(r)})\}_{r=1}^R$  to be generated in the vicinity of  $\hat{\mathbf{p}}_q$ , we fix the  $q$ th local area

$$D_{\hat{x}_q} \triangleq [\hat{x}_q - \Delta_x, \hat{x}_q + \Delta_x], \quad D_{\hat{y}_q} \triangleq [\hat{y}_q - \Delta_y, \hat{y}_q + \Delta_y] \quad (48)$$

The distribution of marginal probability  $g_x^{\text{mar}}(x_q)$  in local x-interval  $D_{\hat{x}_q}$  (or conditional probability  $g_{y|x}^{\text{con}}(y_q|x_q)$  in local y-interval  $D_{\hat{y}_q}$ ) is governed by the design parameter  $\rho_1$ . The range parameter  $\Delta_x$  (or  $\Delta_y$ ) should, therefore, match  
70 with  $\rho_1$ , in order to generate the required realizations efficiently.

## References

- [1] A. Weiss, A.J., Direct position determination of multiple radio signals, EURASIP J. Adv. Signal Process 1 (2005) 37–49. doi:10.1155/ASP.2005.37.
- 75 [2] G. H. Golub, V. Pereyra, The differentiation of pseudo-inverses and non-linear least squares problems whose variables separate, Siam Journal on Numerical Analysis 10 (2) (1973) 413–432. doi:10.1137/0710036.
- [3] S. S. . Wang H, Kay S, An importance sampling maximum likelihood direction of arrival estimator, IEEE Transactions on Signal Processing 56 (10)  
80 (2008) 5082–5092. doi:10.1109/TSP.2008.928504.
- [4] S. Kay, S. Saha, Mean likelihood frequency estimation, Signal Processing IEEE Transactions on 48 (7) (2000) 1937–1946. doi:10.1109/78.847780.
- [5] S. Saha, S. M. Kay, Maximum likelihood parameter estimation of superimposed chirps using monte carlo importance sampling, IEEE Transactions on  
85 Signal Processing 50 (2) (2002) 224–230. doi:10.1109/78.978378.
- [6] H. Wang, S. Kay, Maximum likelihood angle-doppler estimator using importance sampling, IEEE Transactions on Aerospace and Electronic Systems 46 (2) (2010) p.610–622. doi:10.1109/TAES.2010.5461644.
- [7] J. Chen, Y. C. Wu, S. C. Chan, T. S. Ng, Joint maximum-likelihood cfo  
90 and channel estimation for ofdma uplink using importance sampling, IEEE Transactions on Vehicular Technology 57 (6) (2008) 3462–3470. doi:10.1109/TVT.2008.920473.
- [8] M. Pincus, A closed form solution of certain programming problems, Operations Research 16 (3) (1968) 690–694. doi:10.1287/opre.16.3.690.
- 95 [9] S. M. Kay, Intuitive Probability and Random Processes Using MATLAB, Springer Science,, 2006. doi:10.1007/0-387-24158-2\_20.