

**Stochastic Modeling and Particle Filtering Algorithms for Tracking a
Frequency-Hopped Signal**

MASTER THESIS

By

Alexandros Valyrakis

Submitted to the Department of Electronic & Computer Engineering in partial fulfillment of
the requirements for the Master Degree.

Technical University of Crete, Chania

Consulting Committee

Professor Sidiropoulos Nicholas (Advisor)

Professor Paterakis Michalis

Associate Professor Liavas Athanasios

November 2007

Acknowledgements

I would like express my sincerest gratitude to Professor Nicholas Sidiropoulos for his restless guidance and support. I am also grateful to Dr. Anathram Swami for the fruitful collaboration the past two years. In addition, I am thankful to all Professors of the Telecommunication Division for offering me valuable experiences throughout the period of my studies. Finally, I would like to thank my officemate Lefteris Karipidis for his assistance whenever needed.

Contents

1	Introduction	1
2	FH Data Model and Problem Statement	4
3	Particle filtering	7
4	Closed-form Optimal Importance Function and Sampling Procedure for Random Frequency Hopping	10
5	Experiments	14
5.1	Simulations	16
5.1.1	Single-channel case	16
5.1.2	Dual-channel case	16
5.1.3	Split-signal case	17
5.2	Measured data	19
6	Conclusions	29

7 Appendix	32
7.1 Derivation of closed form expression for the optimal importance density . . .	32
7.2 Derivation of dominating density	37
7.3 Modified algorithm (M) for general L	39

Abstract

The problem of tracking a frequency-hopped signal without knowledge of its hopping pattern is considered. The problem is of interest in military communications, where, in addition to frequency, hop timing can also be randomly shifted to guard against unauthorized reception and jamming. A conceptually simple stochastic state-space model is proposed to capture the randomness in carrier frequency and hop timing. The model is non-linear and non-Gaussian, but naturally suited for the application of particle filtering tools, and fortuitous as well: it is possible to compute the optimal importance function in closed-form. The optimal importance function is computed for both single- and multi-channel reception, and a rejection procedure is developed to draw samples from the resulting closed form. The performance of several particle filtering algorithms is assessed and compared to the prior art in a range of experiments using both simulated and measured data.

Chapter 1

Introduction

Tracking the parameters (frequency, complex amplitude) of a time-varying complex sinusoid is an important problem that arises in numerous applications. In many cases the parameters can be assumed to vary slowly in time. Slow variations can often be captured using a simple Gaussian auto-regressive model. A related but different problem emerges in the context of frequency hopping (FH) communications, where the carrier frequency is intentionally hopped in a (pseudo-) random and discontinuous fashion. In military communications, FH is used to guard against unauthorized reception and jamming, and hop timing can also be randomized for added protection. In civilian communications (e.g., Bluetooth), FH is used to avoid persistent interference and enable uncoordinated co-existence with other systems.

Several references have considered the problem of tracking a FH signal without knowledge of the hopping pattern [13, 2, 10, 11]. Non-parametric methods based on the spectrogram [2] are simple and useful as exploratory tools, but suffer from limited resolution due to leakage. It

is possible to employ time-frequency distributions that are better-adapted to frequency hopping [3], but the results are still not very satisfactory. Parametric methods for frequency hopping explicitly model the frequency as piecewise-constant, assume a “budget” on the number of hops within a given observation interval, and employ Dynamic Programming (DP) to track the sought frequency and complex amplitude parameters [10, 11]. Other than an upper bound on the number of hops, the methods in [10, 11] do not assume anything else about the frequencies or complex amplitudes, which are treated as deterministic unknowns. The algorithms in [2, 10, 11] are also applicable when hop timing is random.

A different viewpoint is adopted in this thesis. A stochastic non-linear, non-Gaussian state-space formulation is proposed, which captures frequency hopping dynamics in a probabilistic sense. The proposed formulation is naturally well-suited for the application of particle filtering for state estimation. Compared to the prior state-of-art in [10, 11], the new approach has a number of desirable features. For a fixed average hop rate, the complexity of the DP algorithms in [10, 11] is roughly fourth-order polynomial in the number of temporal signal snapshots, T , and DP requires back-tracking - implying that only short data records can be processed, and in batch mode. Particle filtering can be implemented on-line, and its complexity is linear in T . Furthermore, unlike previous methods, the stochastic state-space model can be easily tailored to match a given scenario (e.g., spread bandwidth and modulation).

In closing this section, we remark that particle filtering solutions for tracking *slowly-varying* parameters of a harmonic or chirp signal are discussed in [14, 15]. Interestingly, the case of slowly-varying parameters is much different, and in a sense more difficult than the

one considered here.

Chapter 2

FH Data Model and Problem Statement

For simplicity of exposition the case of two receive antennas is discussed in the sequel, but the derivations in the Appendix cover the case of $L \geq 1$ antennas.

Let $\mathbf{x}_k := [\omega_k, A_k, B_k]^T$, denote the state at time k , where $\omega_k \in [-\pi, \pi)$ denotes instantaneous frequency, while $A_k \in \mathbb{C}$ and $B_k \in \mathbb{C}$ denote the complex amplitude for the first and second antenna, respectively. Let $\mathbf{u}_k := [b_k, \tilde{\omega}_k, \tilde{A}_k, \tilde{B}_k]^T$ denote an auxiliary i.i.d. sequence of vectors with independent components and the following marginal statistics: b_k is a binary random variable with $Pr(b_k = 1) = h$; $\tilde{\omega}_k$ is uniformly distributed over $[-\pi, \pi)$, denoted $\mathcal{U}([-\pi, \pi))$; and \tilde{A}_k and \tilde{B}_k are $\mathcal{CN}(0, 2\sigma_A^2)$, i.e., complex circular Gaussian of variance σ_A^2 for each real dimension. The state is assumed to evolve according to the following stochastic model:

$$\mathbf{x}_k = f(\mathbf{x}_{k-1}, \mathbf{u}_k) = \begin{cases} \mathbf{x}_{k-1} & , \mathbf{u}_k(1) = 0 \\ [\mathbf{u}_k(2), \mathbf{u}_k(3), \mathbf{u}_k(4)]^T & , \mathbf{u}_k(1) = 1 \end{cases}$$

$$= \begin{cases} \mathbf{x}_{k-1} & , w.p. 1 - h \\ [\mathcal{U}([- \pi, \pi)), \mathcal{CN}(0, \sigma_A^2), \mathcal{CN}(0, \sigma_A^2)]^T & , w.p. h \end{cases} ,$$

and the measurements at time k are

$$\mathbf{y}_k(1) = \mathbf{x}_k(2)e^{j\mathbf{x}_k(1)k} + v_k(1),$$

$$\mathbf{y}_k(2) = \mathbf{x}_k(3)e^{j\mathbf{x}_k(1)k} + v_k(2),$$

where $v_k(\cdot)$ denotes i.i.d. $\mathcal{CN}(0, \sigma_n^2)$ measurement noise.

Some comments on our modeling assumptions are useful at this point.

- The above state-space formulation models frequency hopping in a probabilistic fashion. Hops are random, i.i.d., with hop probability h per sample interval. This is different from traditional models of frequency hopping, which assume that the frequency hops periodically. In addition to frequency, hop timing can also be randomly shifted to help guard against interception and repeat-back jamming. In addition, the random hopping model is well-suited for on-line sequential estimation using particle filtering.
- When the (discrete-time, baseband-equivalent) frequency hops, it hops anywhere within $[-\pi, \pi)$ with a uniform density. This is appropriate for FH signals sampled at the Nyquist rate relative to the hopping bandwidth.
- Modulation-induced variations can be neglected when the objective is to estimate carrier frequency, but could also be explicitly modeled using, e.g., a smooth auto-regressive frequency variation model in-between hops, in lieu of the simplified constant model postulated above.

- We assume independent and identically distributed (i.i.d.) Rayleigh fading in space and frequency. This is realistic for frequency-hopped signals in rich multipath environments without a dominant line-of-sight, when the receive antennas are sufficiently well-separated (a few wavelengths apart). Aside from plausibility, this assumption is also convenient for tractability (analytical integration and bounding) considerations. Interestingly, as we will see, it turns out that the resulting filters work well even when the antennas are correlated. This further justifies the spatial independence assumption from a simplicity viewpoint - there is no need to estimate antenna cross-correlation coefficients.

Given a sequence of observations $\{\mathbf{y}_k\}_{k=1}^T$, our *objective* is to estimate the sequence of system states $\{\mathbf{x}_k\}_{k=1}^T$ - that is, the unknown carrier frequencies and complex amplitudes. We will do this in an indirect way - estimating the posterior density $p(\mathbf{x}_k | \{\mathbf{y}_l\}_{l=1}^k)$, from which the state can be estimated using the conditional mean or mode. This leads us to particle filtering.

Chapter 3

Particle filtering

Particle filtering (PF) is an important estimation toolbox that is applicable to general non-linear and/or non-Gaussian state-space models. We refer the reader to [1, 7, 8] for recent tutorial overviews of PF. In PF, a continuous distribution $p(\mathbf{x})$ is approximated by a discrete random measure comprising “particles” (locations) and corresponding weights:

$$p(\mathbf{x}) \approx \sum_{n=1}^N w_n \delta(\mathbf{x} - \mathbf{x}_n),$$

where $\delta(\cdot)$ denotes the Dirac delta functional. If our goal is to estimate the system state at time k from measurements up to and including time k , the key distribution of interest is the posterior density $p(\mathbf{x}_k | \{\mathbf{y}_l\}_{l=1}^k)$. PF starts with a random measure approximation of the initial state distribution, and uses subsequent measurements to estimate $p(\mathbf{x}_k | \{\mathbf{y}_l\}_{l=1}^k)$, $k \in \{1, 2, \dots\}$ in a sequential fashion, i.e., generate a sequence of random measure approximations

$$\hat{p}(\mathbf{x}_k | \{\mathbf{y}_l\}_{l=1}^k) = \sum_{n=1}^N w_{n,k} \delta(\mathbf{x}_k - \mathbf{x}_{n,k})$$

The measure update step is based on the Bayes rule. If we could sample particles from the desired posterior $p(\mathbf{x}_k | \{\mathbf{y}_l\}_{l=1}^k)$, then all particle weights would have been equal. Unfortunately, direct sampling from the desired posterior is not possible in most cases. For this reason, we resort to *importance sampling*: we draw samples from a suitable *importance function* that has the same support as the desired posterior, and we weigh the particles according to the ratio of the two densities at the sample point. The choice of importance function is very important - the more it resembles the desired density the better, but it should also be easy to sample from.

The simplest and most intuitive choice of importance function is the *prior importance function* $p(\mathbf{x}_k | \mathbf{x}_{n,k-1})$; i.e., the n -th particle is passed through the state-evolution part of the system: $\mathbf{x}_{n,k} = f(\mathbf{x}_{n,k-1}, \mathbf{u}_n)$. A drawback is that particles evolve without taking the latest measurement into account. After all particles have been updated, the weights are adjusted according to $w_{n,k} = w_{n,k-1}p(\mathbf{y}_k | \mathbf{x}_{n,k})$, followed by normalization to $\sum_{n=1}^N w_{n,k} = 1$. Notice that the new measurement affects the weight update.

A common problem in PF is *degeneracy*: the weights of all except a few particles become negligible after several iterations; this can be detected and corrected using *resampling* techniques [1, 7, 8, 4]. The importance function that minimizes the variance of the weights (a precaution against degeneracy) is

$$p(\mathbf{x}_k | \mathbf{x}_{n,k-1}, \mathbf{y}_k) = \frac{p(\mathbf{y}_k | \mathbf{x}_k)p(\mathbf{x}_k | \mathbf{x}_{n,k-1})}{\int_{\mathbf{x}} p(\mathbf{y}_k | \mathbf{x})p(\mathbf{x} | \mathbf{x}_{n,k-1})d\mathbf{x}}.$$

This is often referred to as the *optimal importance function*, and it usually strikes a better estimation performance - complexity trade-off than other alternatives. Notice that the optimal importance function takes into account the latest measurement. Both the prior and the optimal

importance function yield consistent estimates of the desired density when the number of particles goes to infinity.

There are two obstacles to using the optimal importance function: it requires multidimensional integration to compute the normalization factor (which is usually intractable); and sampling from it can be complicated. Thankfully, it turns out that both obstacles can be overcome for our particular model, as shown next.

Chapter 4

Closed-form Optimal Importance

Function and Sampling Procedure for

Random Frequency Hopping

Denote $\mathbf{x}_k := [\omega_k, A_k, B_k]^T$, where $\omega_k \in [-\pi, \pi)$, and $A_k, B_k \in \mathbb{C}$; and likewise $\mathbf{x}_{n,k-1} := [\omega_{n,k-1}, A_{n,k-1}, B_{n,k-1}]^T$. Let $\mathbf{x} := [\omega, A, B]^T$, and

$$D(\mathbf{y}_k, \mathbf{x}_{n,k-1}) := \int_{\mathbf{x}} p(\mathbf{y}_k | \mathbf{x}) p(\mathbf{x} | \mathbf{x}_{n,k-1}) d\mathbf{x}.$$

Then

$$\begin{aligned} D(\mathbf{y}_k, \mathbf{x}_{n,k-1}) &= \int_{\omega \in [-\pi, \pi)} \int_{A \in \mathbb{C}} \int_{B \in \mathbb{C}} \frac{1}{2\pi\sigma_n^2} e^{-\frac{|\mathbf{y}_k(1) - Ae^{j\omega k}|^2}{2\sigma_n^2}} \times \frac{1}{2\pi\sigma_n^2} e^{-\frac{|\mathbf{y}_k(2) - Be^{j\omega k}|^2}{2\sigma_n^2}} \times \\ &\times \left[(1-h)\delta(\omega - \omega_{n,k-1})\delta(A - A_{n,k-1})\delta(B - B_{n,k-1}) + \frac{h}{2\pi} \frac{1}{2\pi\sigma_A^2} \frac{1}{2\pi\sigma_B^2} e^{-\frac{|A|^2}{2\sigma_A^2}} e^{-\frac{|B|^2}{2\sigma_B^2}} \right] dAdBd\omega \end{aligned}$$

This integral can be computed by completing the squares, yielding

$$D(\mathbf{y}_k, \mathbf{x}_{n,k-1}) = \frac{h}{(2\pi)^2(\sigma_n^2 + \sigma_A^2)^2} e^{-\frac{|\mathbf{y}_k(1)|^2 + |\mathbf{y}_k(2)|^2}{2(\sigma_n^2 + \sigma_A^2)}} + \frac{1-h}{(2\pi\sigma_n^2)^2} e^{-\frac{|\mathbf{y}_k(1) - A_{n,k-1} e^{j\omega_{n,k-1}k}|^2}{2\sigma_n^2}} - \frac{|\mathbf{y}_k(2) - B_{n,k-1} e^{j\omega_{n,k-1}k}|^2}{2\sigma_n^2}. \quad (4.1)$$

The weight update for the optimal importance function is

$$w_{n,k} \propto w_{n,k-1} p(\mathbf{y}_k | \mathbf{x}_{n,k-1}) = w_{n,k-1} D(\mathbf{y}_k, \mathbf{x}_{n,k-1}),$$

followed by normalization to 1. We need a way to sample from the optimal importance function. As a first step towards this end, note that $p(\mathbf{x}_k | \mathbf{x}_{n,k-1}, \mathbf{y}_k)$ can be written as a mixture of two pdfs

$$p(\mathbf{x}_k | \mathbf{x}_{n,k-1}, \mathbf{y}_k) = (1 - \tilde{h}) p_0(\mathbf{x}_k | \mathbf{x}_{n,k-1}, \mathbf{y}_k) + \tilde{h} p_1(\mathbf{x}_k | \mathbf{x}_{n,k-1}, \mathbf{y}_k),$$

where

$$p_0(\mathbf{x}_k | \mathbf{x}_{n,k-1}, \mathbf{y}_k) := \delta(\omega_k - \omega_{n,k-1}) \delta(A_k - A_{n,k-1}) \delta(B_k - B_{n,k-1}),$$

$$p_1(\mathbf{x}_k | \mathbf{x}_{n,k-1}, \mathbf{y}_k) := \frac{\frac{1}{(2\pi)^5} \frac{1}{(\sigma_A^2 \sigma_n^2)^2} e^{-\frac{|\mathbf{y}_k(1) - A_k e^{j\omega_k k}|^2}{2\sigma_n^2}} - \frac{|\mathbf{y}_k(2) - B_k e^{j\omega_k k}|^2}{2\sigma_n^2} - \frac{|A_k|^2 - |B_k|^2}{2\sigma_A^2}}{\frac{1}{(2\pi)^2} \frac{h}{(\sigma_n^2 + \sigma_A^2)^2} e^{-\frac{|\mathbf{y}_k(1)|^2 + |\mathbf{y}_k(2)|^2}{2(\sigma_n^2 + \sigma_A^2)}}},$$

and

$$\tilde{h} := h \frac{\frac{1}{(2\pi)^2} \frac{h}{(\sigma_n^2 + \sigma_A^2)^2} e^{-\frac{|\mathbf{y}_k(1)|^2 + |\mathbf{y}_k(2)|^2}{2(\sigma_n^2 + \sigma_A^2)}}}{D(\mathbf{y}_k, \mathbf{x}_{n,k-1})}.$$

It follows that with probability $1 - \tilde{h}$ we simply copy the previous particle, else we draw a particle from $p_1(\mathbf{x}_k | \mathbf{x}_{n,k-1}, \mathbf{y}_k)$. We will use *rejection* for the latter step. Rejection is an exact sampling technique that is well-known and widely-used [6, pp. 40-42]. We wish to draw samples from a density $\phi(\mathbf{x})$, for which there exists a *dominating density* $g(\mathbf{x})$ and a known

constant c such that $\phi(\mathbf{x}) \leq cg(\mathbf{x}), \forall \mathbf{x}$. In practice, we choose $g(\mathbf{x})$ to be easy to sample from, and such that c is as small as possible. We then i) draw a sample \mathbf{x} from $g(\cdot)$ and an independent sample $U \sim \mathcal{U}([0, 1])$; ii) set $\tau := c \frac{g(\mathbf{x})}{\phi(\mathbf{x})}$; iii) if $U\tau \leq 1$ we accept \mathbf{x} else reject it and repeat the process. Rejection requires c trials, on average, to return an accepted sample. The number of trials is geometrically distributed with parameter $1 - \frac{1}{c}$, so the probability of longer trials decays exponentially. In our context, the variance in the number of trials is averaged over many particles, which reduces the variance in execution time per input sample. Still, depending on c , rejection can be computationally very demanding, which is a serious drawback for on-line filtering. On the other hand, rejection is an *exact* sampling method, whereas alternatives would rely on importance sampling.

We wish to sample from the density $p_1(\mathbf{x}_k | \mathbf{x}_{n,k-1}, \mathbf{y}_k)$. Define

$$\mu_1 := \frac{|\mathbf{y}_k(1)|\sigma_A^2}{\sigma_n^2 + \sigma_A^2}, \quad \mu_2 := \frac{|\mathbf{y}_k(2)|\sigma_A^2}{\sigma_n^2 + \sigma_A^2}, \quad \sigma^2 := \frac{\sigma_n^2\sigma_A^2}{\sigma_n^2 + \sigma_A^2}.$$

Using the triangle inequality, it can be shown that the following is a dominating density:

$$g(\mathbf{x}_k | \mathbf{x}_{n,k-1}, \mathbf{y}_k) = \frac{e^{-\frac{(|A_k| - \mu_1)^2}{2\sigma^2}} e^{-\frac{(|B_k| - \mu_2)^2}{2\sigma^2}}}{(2\pi)^3 \sigma^2 (\sigma e^{-\frac{\mu_1}{2\sigma^2}} + \sqrt{2\pi}Q_1)(\sigma e^{-\frac{\mu_2}{2\sigma^2}} + \sqrt{2\pi}Q_2)},$$

for which it holds that $p_1(\mathbf{x}_k | \mathbf{x}_{n,k-1}, \mathbf{y}_k) \leq cg(\mathbf{x}_k | \mathbf{x}_{n,k-1}, \mathbf{y}_k)$, with

$$c := \frac{1}{\sigma^2} (\sigma e^{-\frac{\mu_1}{2\sigma^2}} + \sqrt{2\pi}Q_1)(\sigma e^{-\frac{\mu_2}{2\sigma^2}} + \sqrt{2\pi}Q_2),$$

$$Q_i := \int_{r=0}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(r-\mu_i)^2}{2\sigma^2}} dr = \frac{1}{2} \operatorname{erfc}\left(-\frac{|\mathbf{y}_k(i)|\sigma_A}{\sigma_n\sqrt{2(\sigma_n^2 + \sigma_A^2)}}\right).$$

Through experimentation, we have found that even better results can be attained using an *outer rejection loop*, which declines candidates $\mathbf{x}_{n,k}$ generated through rejection when the following

metric exceeds a certain small value (set to 3×10^{-3} in our experiments):

$$\tilde{h}(\mathbf{y}_k, \mathbf{x}_{n,k}) := h \frac{1}{(2\pi(\sigma_n^2 + \sigma_A^2))^2} e^{-\frac{|\mathbf{y}_k(1)|^2 + |\mathbf{y}_k(2)|^2}{2(\sigma_n^2 + \sigma_A^2)}},$$

where the function $D(\cdot, \cdot)$ is defined in (4.1). This outer rejection loop selects particles that are consistent with the new measurement (cf. the functional form of the denominator) and, at the same time, have large weight after the associated update. We do not have a full explanation at this point, yet this version of the algorithm appears to yield the best results - in particular, better than the one based on the optimal importance function. Note that the latter is optimal with respect to minimizing the variance of the weights after the update (and typically works better than the one based on the prior importance function), but it is not necessarily optimal in terms of the performance - complexity trade-off.

Chapter 5

Experiments

In our experiments, we compared the following PF algorithms:

- P - using the prior importance function;
- O - using the optimal importance function (both single- and dual-channel versions); and
- M - the modified O algorithm that also uses the outer rejection loop (both single- and dual-channel versions).

In all cases, we used resampling at each time step - in particular, the *residual* resampling implementation of Arnaud Doucet and Nando de Freitas. We also included two baseline algorithms in the comparison: the DP algorithm in [10], and a simple spectrogram estimator (SpE). The DP algorithm in [10] can only be used with short data records due to complexity considerations, and requires an upper bound on the number of hops. The tighter the bound the better its performance, so we gave it the correct number of hops. For each time instant

k , SpE effectively computes the periodogram of $\{y_{k-\ell+1}, \dots, y_k\}$, where ℓ is the window length, and then finds its peak. This can be implemented by peak-picking a full-overlap factor rectangularly-windowed spectrogram. SpE requires minimal parameter tuning (window length ℓ - we used $\ell = 8$ or 16 with zero-padding to 1024 samples in our experiments) and would be the first exploratory method that one would use in practice. For the PF algorithms, the initial state was assumed known. This is reasonable for tracking applications, following initial acquisition. We used Root Mean Square Error (RMSE) of frequency estimation as a performance measure.

We conducted experiments with simulated and measured data. The simulated data were generated using our model, and the PF algorithms utilized the correct model parameters. This is meant to assess the performance potential of the proposed algorithms under controlled conditions. Interestingly, PF algorithms proved to be robust even with higher parameters than the model's. This allows to use PF even if the estimation of the parameters of the model is not accurate. We also conducted experiments using measured data which deviate significantly from our modeling assumptions. The setups and associated results are presented next.

5.1 Simulations

For the simulated data, as model parameters for the initial experiments, we used $\sigma_A^2 = 0.5$, and $\sigma_n^2 = 0.1$ and generated data for variance parameters 1 and 0.2 respectively.

5.1.1 Single-channel case

Fig. 5.2 shows RMSE results for the three single-channel PF algorithms, for $T = 100$, and $h = 0.01$, as a function of the number of particles. The results are averaged over 300 Monte-Carlo (MC) trials. Note that the O algorithm is more accurate than the P algorithm, as expected; and the M algorithm is more accurate than the O algorithm. The advantage of O and M over P is pronounced for a relatively small number of particles.

Fig. 5.3 shows RMSE results for the single-channel PF algorithms vs. DP and SpE for $T = 300$, $h = 0.02$, and 200 MC trials. Notice that PF outperforms DP and SpE when the number of particles is sufficiently large; there is a sharp knee in the RMSE performance of PF at about 1000 particles.

5.1.2 Dual-channel case

Fig. 5.4 shows the RMSE performance of single- vs. dual-channel PF as a function of the number of particles, for $T = 100$, $h = 0.01$, and 400 MC trials. The 'single O alg1' and 'single M alg2' curves are for the correct model parameters. All the other curves use $\hat{\sigma}_A^2 = 1$, and $\hat{\sigma}_n^2 = 0.2$, to study the effect of the parameter mismatch. As it is shown, the approach

is robust even for this kind of model mismatch. So, all the following experiments are conducted for those parameters. Whereas the dual-channel algorithms have been derived under the assumption of uncorrelated complex amplitudes for the different antennas, we also tested the correlated case, and the curves are parameterized by the antenna correlation coefficient. Notice that dual-channel PF continues to work well when the antennas are correlated. Fig. 5.5 shows the corresponding mean execution times. Notice that antenna correlation reduces the mean execution time.

5.1.3 Split-signal case

When the received signal is oversampled by an integer factor, M , ω_k lies in $[-\pi/M, \pi/M)$. In this case it is possible to create a multichannel signal by de-interleaving the baseband sample stream from a single antenna; e.g., splitting the signal into odd and even sample subsequences for $M = 2$. This doubles the frequency and creates a two-channel signal. For a pure (unmodulated) carrier signal, the two sub-channels are perfectly correlated. Carrier hopping, data modulation, and noise partially de-correlate the two channels in practice.

Such de-interleaving has been proposed to enhance the resolution of subspace-based parametric line spectra estimators such as MUSIC and ESPRIT [9]. The basic idea is that multiplying the frequency by a factor of M likewise increases the separation of spectral lines. Interestingly, de-interleaving can also enhance the frequency estimation performance of sub-optimal estimators even for a single harmonic in noise. This suggests that de-interleaving might also be useful in our present context. Thus the idea is to split the received sample

sequence into odd and even subsequences, and apply a dual-channel algorithm.

Fig. 5.6 shows an RMSE comparison of split-signal vs. the original (single-signal stream) O and M PF algorithms for $T = 100$, $h = 0.01$, and 200 MC trials, and Fig. 5.7 shows the corresponding mean execution times. Notice that the split-signal algorithm is more accurate for a low number of particles, but at a significant complexity cost.

5.2 Measured data

We further tested our approach using measured FH data, made available by Telcordia Technologies through the ARL Collaborative Technology Alliance (ARL-CTA) for Communications and Networks under Cooperative Agreement DADD19-01-2-0011. The measurement campaign was conducted in 2004, and comprised a diverse array of scenarios. The particular data that we used corresponds to a line-of-sight (LoS) scenario called *T3-LoS*. Over-the-air testing was performed using a software-defined radio and Agilent 4438 synthesizers; the carrier frequency was 1.875 GHz, with a two-sided bandwidth of 1.25 MHz; the signal bandwidth was 1 MHz. The received signal was sampled at 4 Ms/s with a 12-bit ADC. The (baseband) hopping bandwidth was from -0.5 to 0.5 MHz, divided into 32 equi-spaced frequency bins, and slow FH with binary Gaussian Minimum Shift Keying (GMSK) modulation was used. The maximum modulation-induced frequency deviation was $\frac{1}{128}$ -th of the FH bin width.

The transmitted waveform was known, and included a long synchronization preamble, which also affords accurate channel estimation. The received signal was down-converted and oversampled at the receiver by a factor of 4. This compresses the frequency variation to a quarter of the band, and is (grossly) inconsistent with our model; we therefore sub-sampled the signal prior to processing by a factor of 4. The end result is a full-band SFH GMSK signal with 128 samples per dwell, from which we extracted a segment comprising 2000 samples. The corresponding hopping sequence is plotted in Fig. 5.8.

There are many reasons why testing with measured data is important. The measured data violate several of our assumptions, which is to be expected in practice:

- Due to the presence of a strong LoS component and only minor multipath, the signal's amplitude is approximately constant across dwells. This violates our i.i.d. Rayleigh fading assumption from dwell to dwell.
- Within a dwell, the frequency is not constant; it varies slowly due to GMSK modulation. This induces a time-varying residual phase noise relative to the true carrier. This is illustrated in Fig. 5.9, which shows one dwell of the unmodulated and the received signal after downconversion to the hopping bandwidth, synchronization, amplitude and phase correction. The net effect is that the compound noise term is correlated and non-Gaussian. This is illustrated in Fig. 5.10, which shows a Normal probability plot of the real part of the noise signal. Notice the significant deviation from the Normal distribution in the tails.
- Carrier hopping is not i.i.d. random as postulated in our model; it is periodic (without intentional jitter - see Fig. 5.8).
- The parameters of our model $(h, \sigma_A^2, \sigma_n^2)$ have to be estimated.

For the above reasons, trying our algorithms on the measured data is a meaningful test of robustness to model mismatch.

The hop period (dwell duration) T_d can be accurately estimated from the spectrogram, or using cyclostationarity. We therefore set $h = 1/T_d = 1/128$. Given T_d , it is possible to segment the signal in fixed-length dwells using a serial acquisition search, and estimate the remaining parameters from the segments. We empirically set $\sigma_A^2 = 0.9$, $\sigma_n^2 = 0.05$. These

values are somewhat lower yet work better than the associated sample averages, which we attribute to model mismatch.

The results of our experiments using measured data are summarized in Fig. 5.11, which shows the RMSE frequency estimation performance of O and M PF algorithms vis-a-vis the SpE. Notice the sharp knee in the performance of the two PF algorithms, which occurs at 500 particles for the M algorithm and between 500 and 1000 particles for the O algorithm. This behavior has been validated by further tests (re-sampling, time-reversal, and adding small amounts of noise to the received waveform).

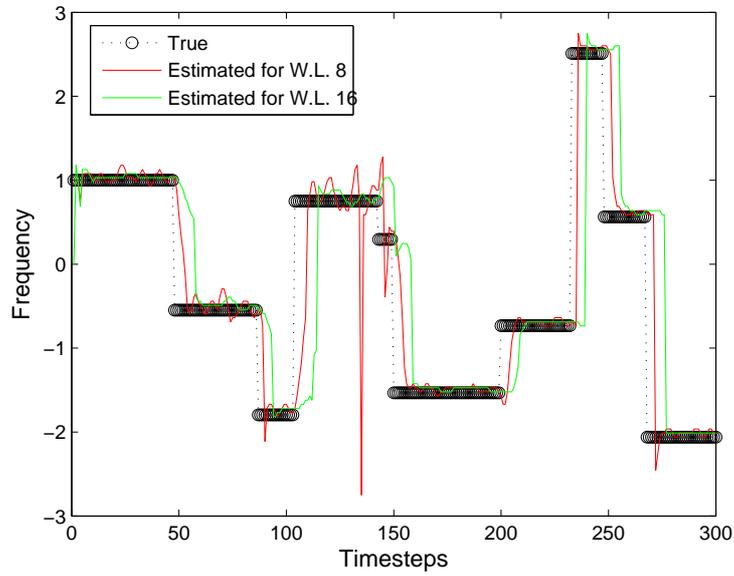


Figure 5.1: Peak-picking the spectrogram: typical results using window length 8 and 16.

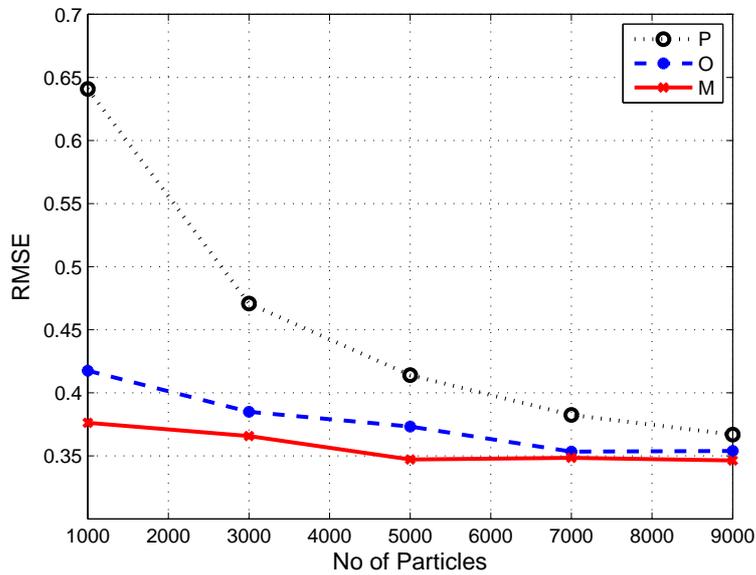


Figure 5.2: RMSE results for the three single-channel PF algorithms (Prior, Optimal, Modified): $T = 100$, $h = 0.01$, 300 MC trials.

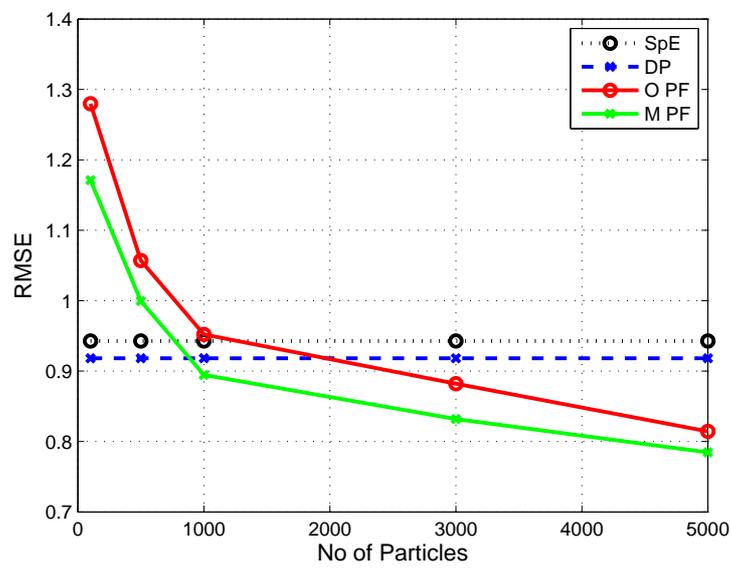


Figure 5.3: RMSE comparison of PF solutions vs. DP and SpE for simulated data: $T = 300$, $h = 0.02$, 200 MC trials.

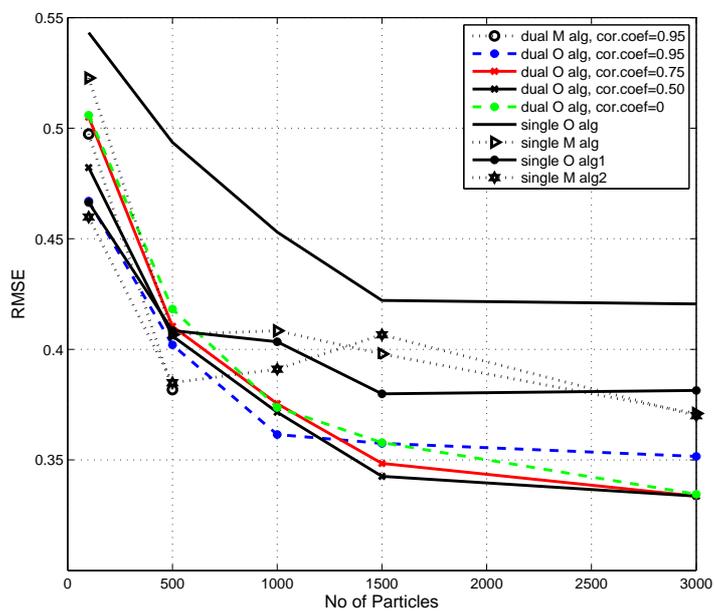


Figure 5.4: RMSE comparison of single- vs. dual-channel PF for $T = 100$, $h = 0.01$, 400 MC trials. For dual-channel PF algorithms, the curves are parameterized by antenna correlation.

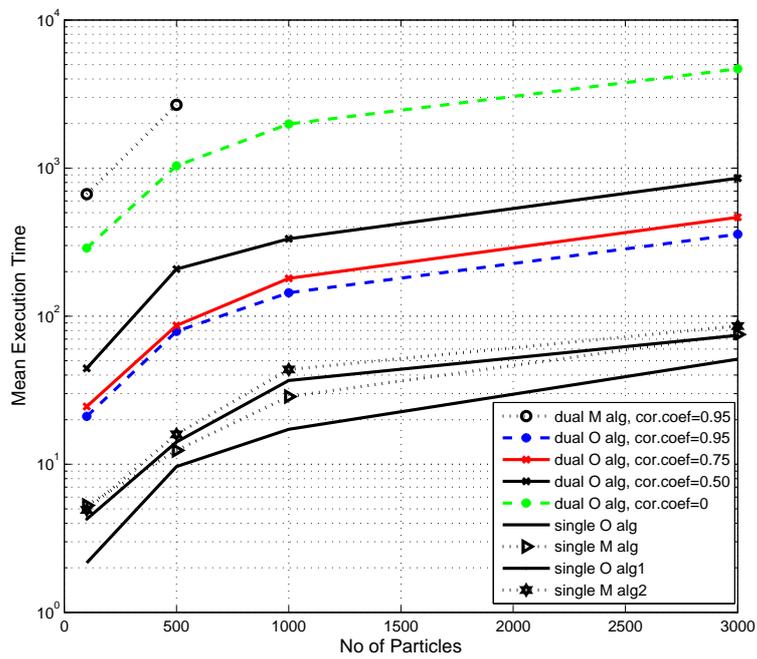


Figure 5.5: Execution time for $T = 100$ samples averaged over 400 MC trials for $h = 0.01$.

For dual-channel PF algorithms, the curves are parameterized by antenna correlation.

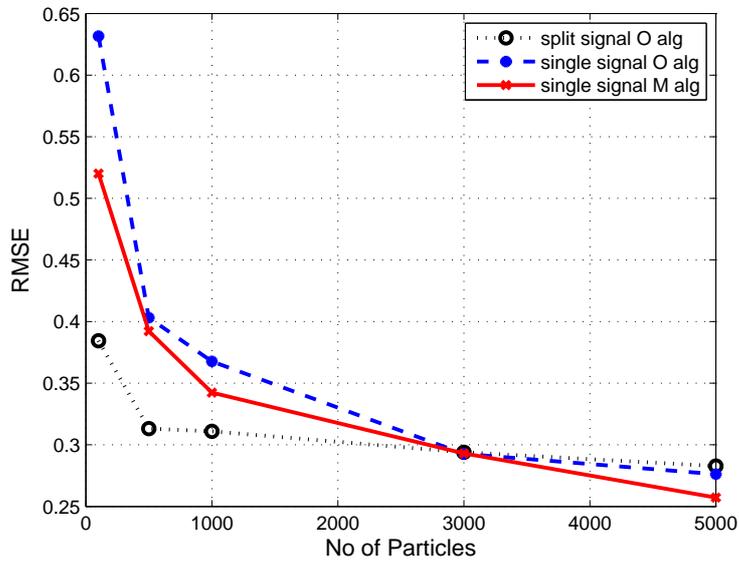


Figure 5.6: RMSE comparison of split-signal vs. original PF algorithms: $T = 100$, $h = 0.01$, 200 MC trials.

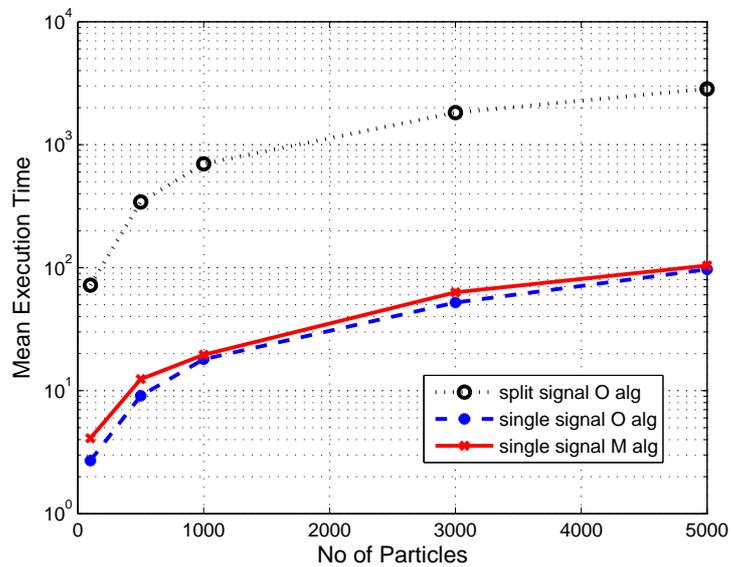


Figure 5.7: Execution time for split-signal and original PF algorithms for $T = 100$ samples averaged over 200 MC trials for $h = 0.01$.

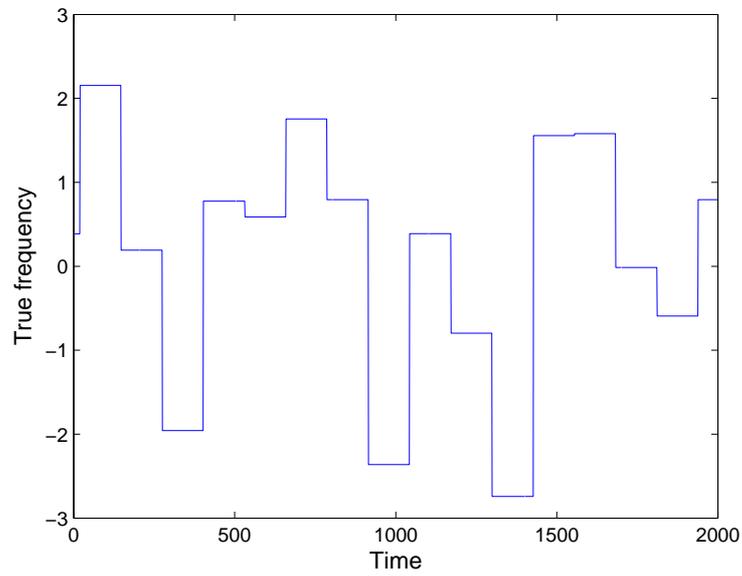


Figure 5.8: Hopping sequence for the measured data ($T = 2000$ temporal samples).

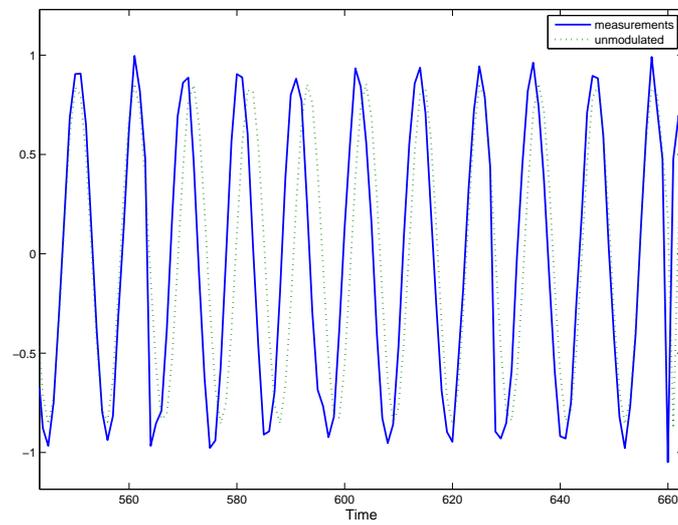


Figure 5.9: One dwell of the carrier and the received signal after synchronization, amplitude and phase correction. Notice the residual phase noise due to modulation.

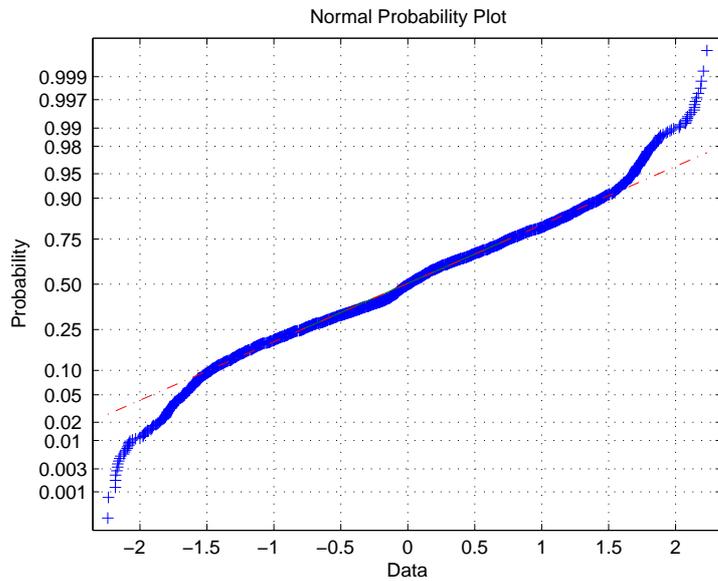


Figure 5.10: Normal probability plot of real part of residual signal for $T = 2000$ time steps. Notice significant deviation from Normal distribution in the tails, primarily due to carrier modulation.

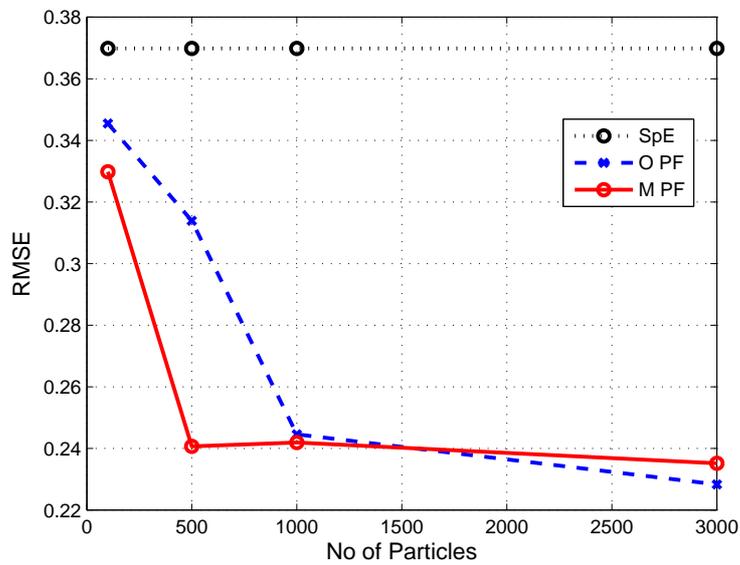


Figure 5.11: RMSE of PF vs. SpE for measured data, $T = 2000$, $h = 1/128$.

Chapter 6

Conclusions

The multichannel filters were developed under the assumption that the channels fade independently, yet simulations (cf. Fig. 5.4 and Fig. 5.5) suggest that the filters remain operational when the channels are (even strongly) correlated; and in fact channel correlation substantially reduces complexity, while RMSE appears insensitive to channel correlation.

Using more channels can reduce RMSE, but complexity grows very quickly with the number of channels, even if the number of particles is kept fixed (cf. Fig. 5.5). This is because every additional channel requires augmenting the state with another complex-valued dimension to model the associated complex amplitude. Dual-channel filters are already much more complex than their single-channel counterparts. As a result, if more than two channels are available, it makes sense to select a pair that exhibits the highest correlation; this will yield good performance at relatively low complexity.

It is also interesting to note that the mean execution time for O and M filters is a slightly

sub-linear¹ function of the number of particles. This can be explained as follows. Recall that previous particles are copied with probability $1 - \tilde{h}$, and new ones are drawn, using rejection, with probability \tilde{h} . When the O and M filters track well, \tilde{h} is small, thus reducing the complexity cost of rejection. Tracking performance is improved with increasing number of particles, thus reducing \tilde{h} , which in turn reduces the per-particle complexity.

In all scenarios considered, the M filter outperforms the O filter in terms of RMSE but also in terms of execution time for the same number of particles. This suggests that the M filter uses particles more efficiently than the O filter. For both, there is no appreciable improvement in RMSE beyond a certain number of particles, but the breakpoint is smaller for the M filter.

The robustness of O and M filters with respect to model mismatch has been assessed using measured FH data, which violate many of our assumptions. The measured data feature strong LoS (approximately constant signal amplitude), dwell frequency modulation, periodic hop timing, and correlated non-Gaussian noise - thus significantly departing from our working assumptions. Still, the proposed O and M filters work remarkably well, with as few as 500-1000 particles. To appreciate this, consider the following back-of-the-envelope calculation. If a tracking algorithm misses a half-band hop of π rads by a single sample, but otherwise tracks the signal perfectly over a dwell of 128 samples, the associated RMSE will be approximately 0.27. Most hops are less than half-band, yielding lower RMSE in the same scenario; but there will also be errors over the duration of the dwell. M and O filters (with 500 and 1000 particles, respectively) in Fig. 5.11 yield an RMSE of 0.24. This speaks for a key strength of

¹Fig. 5.5 and Fig. 5.7 are in log scale, requiring careful reading to notice this sub-linearity.

PF solutions versus window-based methods (such as SpE): the ability to accurately track hop timing. PF solutions can overcome the need to trade-off bias for variance, at the expense of much higher complexity, which is the main drawback of PF.

At present, O and M filters are too complex for practical on-line implementation, and mainly serve as performance benchmarks. While the particle sampling step can be parallelized and there has been recent progress [4, 5] in parallel and distributed implementation of PF algorithms, much remains to be done to bring on-line implementation within reach. Replacing rejection with a more efficient sampling scheme will certainly help in this direction. On the other hand, rejection generates exact samples, thereby affording the opportunity to assess ultimate performance, without regard to complexity.

Chapter 7

Appendix

7.1 Derivation of closed form expression for the optimal importance density

Assume that L antennas and corresponding receive chains are available, for any $L \geq 1$. Define the state vector The state $\mathbf{x}_k := [\omega_k, A_k^{(1)}, \dots, A_k^{(L)}]^T$ where $\omega_k \in [-\pi, \pi)$ denotes instantaneous frequency and $A_k^{(1)}, \dots, A_k^{(L)} \in \mathbb{C}$ are complex amplitudes . Define the auxiliary i.i.d. sequence of random vectors $\mathbf{u}_k := [b_k, \tilde{\omega}_k, \tilde{A}_k^{(1)}, \dots, \tilde{A}_k^{(L)}]^T$, where b_k is a binary random variable with $Pr(b_k = 1) = h$; $\tilde{\omega}_k$ is $\mathcal{U}([-\pi, \pi))$; and $\tilde{A}_k^{(1)}, \dots, \tilde{A}_k^{(L)}$ are i.i.d. $\mathcal{CN}(0, \sigma_A^2)$.

Then

$$\mathbf{x}_k = f(\mathbf{x}_{k-1}, \mathbf{u}_k) = \begin{cases} \mathbf{x}_{k-1} & , \mathbf{u}_k(1) = 0 \\ [\mathbf{u}_k(2), \dots, \mathbf{u}_k(L+1)]^T & , \mathbf{u}_k(1) = 1 \end{cases}$$

$$= \begin{cases} \mathbf{x}_{k-1} & , w.p. 1 - h \\ [\mathcal{U}([- \pi, \pi]), \mathcal{CN}(0, \sigma_A^2), \dots, \mathcal{CN}(0, \sigma_A^2)]^T & , w.p. h \end{cases} ,$$

$$\mathbf{y}_k(l) = \mathbf{x}_k(l+1)e^{j\mathbf{x}_k(1)k} + v_k(l), \quad l \in \{1, \dots, L\},$$

where $v_k(\cdot)$ denotes $\mathcal{CN}(0, \sigma_n^2)$ measurement noise that is assumed to be i.i.d. in space (l) and time (k).

From the viewpoint of minimizing the variance of the weights, the optimal importance function is given by [1, 7]

$$p(\mathbf{x}_k | \mathbf{x}_{n,k-1}, \mathbf{y}_k) = \frac{p(\mathbf{y}_k | \mathbf{x}_k)p(\mathbf{x}_k | \mathbf{x}_{n,k-1})}{\int_{\mathbf{x}} p(\mathbf{y}_k | \mathbf{x})p(\mathbf{x} | \mathbf{x}_{n,k-1})d\mathbf{x}}.$$

The integral $D(\mathbf{y}_k, \mathbf{x}_{n,k-1}) := \int_{\mathbf{x}} p(\mathbf{y}_k | \mathbf{x})p(\mathbf{x} | \mathbf{x}_{n,k-1})d\mathbf{x}$ can be computed in closed form for general L .

$$D(\mathbf{y}_k, \mathbf{x}_{n,k-1}) = \int_{\omega \in [-\pi, \pi]} \int_{A^{(1)} \in \mathbb{C}} \dots \int_{A^{(L)} \in \mathbb{C}} \prod_{m=1}^L \left[\frac{1}{2\pi\sigma_n^2} e^{-\frac{|\mathbf{y}_k(m) - A^{(m)} e^{j\omega k}|^2}{2\sigma_n^2}} \right]$$

$$\times \left[(1-h)\delta(\omega - \omega_{n,k-1}) \prod_{m=1}^L \delta(A^{(m)} - A_{n,k-1}^{(m)}) + \frac{h}{2\pi} \prod_{m=1}^L \left(\frac{1}{2\pi\sigma_A^2} e^{-\frac{|A^{(m)}|^2}{2\sigma_A^2}} \right) \right] dA^{(1)} \dots dA^{(L)} d\omega = D_1 + D_2,$$

where

$$D_1 = \frac{1-h}{(2\pi\sigma_n^2)^L} \int_{\omega \in [-\pi, \pi]} \int_{A^{(1)} \in \mathbb{C}} \dots \int_{A^{(L)} \in \mathbb{C}} \prod_{m=1}^L \left[e^{-\frac{|\mathbf{y}_k(m) - A^{(m)} e^{j\omega k}|^2}{2\sigma_n^2}} \right]$$

$$\times \left[\delta(\omega - \omega_{n,k-1}) \prod_{m=1}^L \delta(A^{(m)} - A_{n,k-1}^{(m)}) \right] dA^{(1)} \dots dA^{(L)} d\omega =$$

$$\frac{1-h}{(2\pi\sigma_n^2)^L} e^{-\frac{\sum_{m=1}^L |\mathbf{y}_k(m) - A_{n,k-1}^{(m)} e^{j\omega_{n,k-1} k}|^2}{2\sigma_n^2}},$$

and

$$\begin{aligned}
D_2 &= \int_{\omega \in [-\pi, \pi]} \int_{A^{(1)} \in \mathbb{C}} \cdots \int_{A^{(L)} \in \mathbb{C}} \frac{1}{(2\pi\sigma_n^2)^L} \frac{1}{(2\pi\sigma_A^2)^L} \frac{h}{2\pi} e^{-\frac{\sum_{m=1}^L |\mathbf{y}_k(m) - A^{(m)} e^{j\omega k}|^2}{2\sigma_n^2}} e^{-\frac{\sum_{m=1}^L |A^{(m)}|^2}{2\sigma_A^2}} dA^{(1)} \dots dA^{(L)} d\omega \\
&= \frac{h}{(2\pi)^{(2L+1)} \sigma_A^{2L} \sigma_n^{2L}} \int_{\omega \in [-\pi, \pi]} \prod_{m=1}^L \left[\int_{A^{(m)} \in \mathbb{C}} e^{-\frac{|\mathbf{y}_k(m) - A^{(m)} e^{j\omega k}|^2}{2\sigma_n^2}} e^{-\frac{|A^{(m)}|^2}{2\sigma_A^2}} dA^{(m)} \right] d\omega.
\end{aligned}$$

The integral inside the brackets in the above equation can be computed by completing the squares:

$$\mathfrak{J}_{A^{(m)}} = \int_{A^{(m)} \in \mathbb{C}} e^{-\frac{|\mathbf{y}_k(m) - A^{(m)} e^{j\omega k}|^2}{2\sigma_n^2}} e^{-\frac{|A^{(m)}|^2}{2\sigma_A^2}} dA^{(m)} = \int_{A^{(m)} \in \mathbb{C}} e^{-\frac{F}{2\sigma_A^2 \sigma_n^2}} dA^{(m)},$$

where, F is defined as

$$\begin{aligned}
F &:= \sigma_A^2 |\mathbf{y}_k(m) - A^{(m)} e^{j\omega k}|^2 + \sigma_n^2 |A^{(m)}|^2 = \\
&= \sigma_A^2 [|\mathbf{y}_k(m)|^2 - 2\Re(A^{(m)} \mathbf{y}_k^*(m) e^{j\omega k}) + |A^{(m)}|^2] + \sigma_n^2 |A^{(m)}|^2 = \\
&= (\sigma_A^2 + \sigma_n^2) |A^{(m)}|^2 - 2\Re[A^{(m)} (\sigma_A^2 \mathbf{y}_k^*(m) e^{j\omega k})] + \sigma_A^2 |\mathbf{y}_k(m)|^2,
\end{aligned}$$

where $\Re(\cdot)$ ($\Im(\cdot)$) extracts the real (imaginary) part of its argument. Let $W^* := \frac{\sigma_A^2 \mathbf{y}_k(m)^* e^{j\omega k}}{\sigma_A^2 + \sigma_n^2}$;

then

$$\begin{aligned}
F &= (\sigma_A^2 + \sigma_n^2) [|A^{(m)}|^2 - 2\Re(A^{(m)} W^*) + |W|^2] - (\sigma_A^2 + \sigma_n^2) |W|^2 + \sigma_A^2 |\mathbf{y}_k(m)|^2 = \\
&= (\sigma_A^2 + \sigma_n^2) |A^{(m)} - W|^2 + \frac{\sigma_A^2 \sigma_n^2}{\sigma_A^2 + \sigma_n^2} |\mathbf{y}_k(m)|^2.
\end{aligned}$$

We now readily see that $\mathfrak{J}_{A^{(m)}}$ can be written as

$$\mathfrak{J}_{A^{(m)}} = \int_{A^{(m)} \in \mathbb{C}} e^{-\frac{(\sigma_A^2 + \sigma_n^2) |A^{(m)} - W|^2}{2\sigma_A^2 \sigma_n^2}} e^{-\frac{|\mathbf{y}_k(m)|^2}{2(\sigma_A^2 + \sigma_n^2)}} dA^{(m)} =$$

$$\begin{aligned}
&= e^{-\frac{|\mathbf{y}_k(m)|^2}{2(\sigma_A^2 + \sigma_n^2)}} \int_{A^{(m)} \in \mathbb{C}} e^{-\frac{|A^{(m)} - W|^2}{2\frac{\sigma_A^2 \sigma_n^2}{\sigma_A^2 + \sigma_n^2}}} dA^{(m)} = \\
&= e^{-\frac{|\mathbf{y}_k(m)|^2}{2(\sigma_A^2 + \sigma_n^2)}} \times \int_{\Re(A^{(m)})} \int_{\Im(A^{(m)})} e^{-\frac{(\Re(A^{(m)}) - \Re(W))^2}{2\frac{\sigma_A^2 \sigma_n^2}{\sigma_A^2 + \sigma_n^2}} - \frac{(\Im(A^{(m)}) - \Im(W))^2}{2\frac{\sigma_A^2 \sigma_n^2}{\sigma_A^2 + \sigma_n^2}}} d\Re(A^{(m)}) d\Im(A^{(m)}) = \\
&= e^{-\frac{|\mathbf{y}_k(m)|^2}{2(\sigma_A^2 + \sigma_n^2)}} \times \left[\int_{\Re(A^{(m)})} e^{-\frac{(\Re(A^{(m)}) - \Re(W))^2}{2\frac{\sigma_A^2 \sigma_n^2}{\sigma_A^2 + \sigma_n^2}}} d\Re(A^{(m)}) \int_{\Im(A^{(m)})} e^{-\frac{(\Im(A^{(m)}) - \Im(W))^2}{2\frac{\sigma_A^2 \sigma_n^2}{\sigma_A^2 + \sigma_n^2}}} d\Im(A^{(m)}) \right].
\end{aligned}$$

This yields

$$\begin{aligned}
\mathfrak{I}_A^{(m)} &= e^{-\frac{|\mathbf{y}_k(m)|^2}{2(\sigma_A^2 + \sigma_n^2)}} \left[\sqrt{\frac{\pi}{\frac{\sigma_A^2 + \sigma_n^2}{2\sigma_A^2 \sigma_n^2}}} \right] \times \left[\sqrt{\frac{\pi}{\frac{\sigma_A^2 + \sigma_n^2}{2\sigma_A^2 \sigma_n^2}}} \right] = \\
&= 2\pi \frac{\sigma_A^2 \sigma_n^2}{\sigma_A^2 + \sigma_n^2} e^{-\frac{|\mathbf{y}_k(m)|^2}{2(\sigma_A^2 + \sigma_n^2)}}.
\end{aligned}$$

D_2 can now be written as

$$\begin{aligned}
D_2 &= \frac{h}{(2\pi)^{(2L+1)} \sigma_A^{2L} \sigma_n^{2L}} \int_{\omega \in [-\pi, \pi]} \prod_{m=1}^L \left[2\pi \frac{\sigma_A^2 \sigma_n^2}{\sigma_A^2 + \sigma_n^2} e^{-\frac{|\mathbf{y}_k(m)|^2}{2(\sigma_A^2 + \sigma_n^2)}} \right] d\omega \\
&= \frac{h}{(2\pi(\sigma_A^2 + \sigma_n^2))^L} e^{-\frac{\sum_{m=1}^L |\mathbf{y}_k(m)|^2}{2(\sigma_A^2 + \sigma_n^2)}},
\end{aligned}$$

and

$$D(\mathbf{y}_k, \mathbf{x}_{n,k-1}) = \frac{1 - h}{(2\pi\sigma_n^2)^L} e^{-\frac{\sum_{m=1}^L |\mathbf{y}_k(m) - A_{n,k-1}^{(m)} e^{j\omega_{n,k-1} k}|^2}{2\sigma_n^2}} + \frac{h}{(2\pi(\sigma_A^2 + \sigma_n^2))^L} e^{-\frac{\sum_{m=1}^L |\mathbf{y}_k(m)|^2}{2(\sigma_A^2 + \sigma_n^2)}}. \quad (7.1)$$

For the above optimal choice of the importance function, the weight update is given by

$$w_{n,k} \propto w_{n,k-1} p(\mathbf{y}_k | \mathbf{x}_{n,k-1}) = w_{n,k-1} D(\mathbf{y}_k, \mathbf{x}_{n,k-1}),$$

followed by normalization to 1. What is missing is a way to sample from the optimal importance function. As a first step towards this end, note that $p(\mathbf{x}_k | \mathbf{x}_{n,k-1}, \mathbf{y}_k)$ can be written as a

mixture of two pdfs

$$p(\mathbf{x}_k | \mathbf{x}_{n,k-1}, \mathbf{y}_k) = (1 - \tilde{h})p_0(\mathbf{x}_k | \mathbf{x}_{n,k-1}, \mathbf{y}_k) + \tilde{h}p_1(\mathbf{x}_k | \mathbf{x}_{n,k-1}, \mathbf{y}_k),$$

where

$$p_0(\mathbf{x}_k | \mathbf{x}_{n,k-1}, \mathbf{y}_k) := \delta(\omega_k - \omega_{n,k-1}) \prod_{m=1}^L \delta(A_k^{(m)} - A_{n,k-1}^{(m)}),$$

$$p_1(\mathbf{x}_k | \mathbf{x}_{n,k-1}, \mathbf{y}_k) := \frac{\frac{1}{(2\pi\sigma_n^2)^L} \frac{1}{(2\pi\sigma_A^2)^L} \frac{1}{2\pi} e^{-\frac{\sum_{m=1}^L |\mathbf{y}_k^{(m)} - A_k^{(m)}|^2}{2\sigma_n^2}} e^{-\frac{\sum_{m=1}^L |A_k^{(m)}|^2}{2\sigma_A^2}}}{\frac{1}{(2\pi)^L} \frac{1}{(\sigma_n^2 + \sigma_A^2)^L} e^{-\frac{\sum_{m=1}^L |\mathbf{y}_k^{(m)}|^2}{2(\sigma_n^2 + \sigma_A^2)}}},$$

and

$$\tilde{h} := h \frac{\frac{1}{(2\pi)^L} \frac{1}{(\sigma_n^2 + \sigma_A^2)^L} e^{-\frac{\sum_{m=1}^L |\mathbf{y}_k^{(m)}|^2}{2(\sigma_n^2 + \sigma_A^2)}}}{D(\mathbf{y}_k, \mathbf{x}_{n,k-1})}.$$

It follows that with probability $1 - \tilde{h}$ we simply copy the previous particle, else we draw a particle from $p_1(\mathbf{x}_k | \mathbf{x}_{n,k-1}, \mathbf{y}_k)$. We will use rejection sampling techniques for this latter step, as explained next.

7.2 Derivation of dominating density

The density $p_1(\mathbf{x}_k | \mathbf{x}_{n,k-1}, \mathbf{y}_k)$ can be written as

$$p_1(\mathbf{x}_k | \mathbf{x}_{n,k-1}, \mathbf{y}_k) = \frac{1}{(2\pi)^{(L+1)}} \left(\frac{\sigma_n^2 + \sigma_A^2}{\sigma_A^2 \sigma_n^2} \right)^L e^{-\sum_{m=1}^L \left[\frac{|\mathbf{y}_k(m) - A_k^{(m)} e^{j\omega k}|^2}{2\sigma_n^2} + \frac{|A_k^{(m)}|^2}{2\sigma_A^2} \right]} e^{-\frac{\sum_{m=1}^L |\mathbf{y}_k(m)|^2}{2(\sigma_A^2 + \sigma_n^2)}}.$$

Using the triangle inequality, we have

$$\begin{aligned} \frac{|\mathbf{y}_k(m) - A_k^{(m)} e^{j\omega k}|^2}{2\sigma_n^2} + \frac{|A_k^{(m)}|^2}{2\sigma_A^2} &\geq \frac{|\mathbf{y}_k(m)|^2 + |A_k^{(m)}|^2 - 2|\mathbf{y}_k(m)||A_k^{(m)}|}{2\sigma_n^2} + \frac{|A_k^{(m)}|^2}{2\sigma_A^2} = \\ &= \frac{\left(|A_k^{(m)}| - \frac{|\mathbf{y}_k(m)|\sigma_A^2}{\sigma_A^2 + \sigma_n^2} \right)^2}{2\frac{\sigma_A^2 \sigma_n^2}{\sigma_A^2 + \sigma_n^2}} + \frac{|\mathbf{y}_k(m)|^2}{2(\sigma_A^2 + \sigma_n^2)} \end{aligned}$$

and since e^{-x} is monotonically decreasing, for $m = 1, \dots, L$

$$e^{-\frac{|\mathbf{y}_k(m) - A_k^{(m)} e^{j\omega k}|^2}{2\sigma_n^2} - \frac{|A_k^{(m)}|^2}{2\sigma_A^2} + \frac{|\mathbf{y}_k(m)|^2}{2(\sigma_A^2 + \sigma_n^2)}} \leq e^{-\frac{(|A_k^{(m)}| - \mu_m)^2}{2\sigma^2}}$$

where $\sigma^2 := \frac{\sigma_A^2 \sigma_n^2}{\sigma_A^2 + \sigma_n^2}$ and $\mu_m := \frac{|\mathbf{y}_k(m)|\sigma_A^2}{\sigma_A^2 + \sigma_n^2}$ for $m = 1, \dots, L$. Using the above inequalities,

$$\begin{aligned} p_1(\mathbf{x}_k | \mathbf{x}_{n,k-1}, \mathbf{y}_k) &= \frac{1}{(2\pi)^{(L+1)} \sigma^{2L}} e^{-\sum_{m=1}^L \left[-\frac{|\mathbf{y}_k(m) - A_k^{(m)} e^{j\omega k}|^2}{2\sigma_n^2} - \frac{|A_k^{(m)}|^2}{2\sigma_A^2} + \frac{|\mathbf{y}_k(m)|^2}{2(\sigma_A^2 + \sigma_n^2)} \right]} \\ &\leq \frac{1}{(2\pi)^{(L+1)} \sigma^{2L}} e^{-\frac{\sum_{m=1}^L (|A_k^{(m)}| - \mu_m)^2}{2\sigma^2}}. \end{aligned}$$

The only remaining part is to evaluate the normalization factor and the dominating density;

for the former

$$\begin{aligned} c &:= \int_{\omega_k \in [-\pi, \pi]} \int_{A_k^{(1)} \in \mathbb{C}} \dots \int_{A_k^{(L)} \in \mathbb{C}} \frac{1}{(2\pi)^{(L+1)} \sigma^{2L}} e^{-\frac{\sum_{m=1}^L (|A_k^{(m)}| - \mu_m)^2}{2\sigma^2}} dA_k^{(1)} \dots dA_k^{(L)} d\omega_k = \\ &= \frac{1}{(2\pi)^{(L+1)} \sigma^{2L}} \int_{\omega_k \in [-\pi, \pi]} d\omega_k \prod_{m=1}^L \left(\int_{A_k^{(m)} \in \mathbb{C}} e^{-\frac{(|A_k^{(m)}| - \mu_m)^2}{2\sigma^2}} dA_k^{(m)} \right), \end{aligned}$$

with

$$\begin{aligned} \int_{A_k^{(m)} \in \mathbb{C}} e^{-\frac{(|A_k^{(m)}| - \mu_m)^2}{2\sigma^2}} dA_k^{(m)} &= 2\pi \int_{|A_k^{(m)}| \in \mathbb{R}^+} |A_k^{(m)}| e^{-\frac{(|A_k^{(m)}| - \mu_m)^2}{2\sigma^2}} d|A_k^{(m)}| \\ &= (2\pi)(\sigma^2 e^{-\frac{\mu_m}{2\sigma^2}} + \sigma\sqrt{2\pi}Q_m), m = 1, \dots, L, \end{aligned}$$

with

$$Q_m := \int_{r=0}^{+\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(r-\mu_m)^2}{2\sigma^2}} dr = \frac{1}{2} \operatorname{erfc} \left(-\frac{|\mathbf{y}_k(m)|\sigma_A}{\sigma_n\sqrt{2(\sigma_A^2 + \sigma_n^2)}} \right).$$

Hence

$$c = \frac{1}{\sigma^L} \prod_{m=1}^L \left(\sigma e^{-\frac{\mu_m}{2\sigma^2}} + \sqrt{2\pi}Q_m \right)$$

(recall $\sigma^2 = \frac{\sigma_A^2\sigma_n^2}{\sigma_A^2 + \sigma_n^2}$, $\mu_m = \frac{|\mathbf{y}_k(m)|\sigma_A^2}{\sigma_A^2 + \sigma_n^2}$ for $m = 1, \dots, L$) and the dominating density is

$$g(\mathbf{x}_k | \mathbf{x}_{n,k-1}, \mathbf{y}_k) = \frac{1}{(2\pi)^{L+1} \sigma^L} \prod_{m=1}^L \frac{e^{-\frac{(|A_k^{(m)}| - \mu_m)^2}{2\sigma^2}}}{\left(\sigma e^{-\frac{\mu_m}{2\sigma^2}} + \sqrt{2\pi}Q_m \right)}.$$

7.3 Modified algorithm (M) for general L

For general L , the modified PF algorithm (M) uses an outer particle rejection loop with the following metric

$$\tilde{h}(\mathbf{y}_k, \mathbf{x}_{n,k}) := h \frac{\frac{1}{(2\pi(\sigma_n^2 + \sigma_A^2))^L} e^{-\frac{\sum_{m=1}^L |y_k^{(m)}|^2}{2(\sigma_n^2 + \sigma_A^2)}}}{D(\mathbf{y}_k, \mathbf{x}_{n,k})},$$

where $D(\cdot, \cdot)$ is defined in (7.1). This outer loop selects particles that are consistent with the new measurement (cf. the functional form of the denominator) and, at the same time, have large weight after the associated weight update step.

Bibliography

- [1] M.S. Arulampalam, S. Maskell, N. Gordon, T. Clapp, “A Tutorial on Particle Filters for Nonlinear/Non-Gaussian Bayesian Tracking,” *IEEE Trans. Signal Processing*, vol. 50, no. 2, pp. 174–188, Feb. 2002.
- [2] L. Aydin and A. Polydoros, “Hop-timing Estimation for FH Signals Using a Coarsely Channelized Receiver,” *IEEE Trans. Communications*, vol. 44, no. 4, pp. 516–526, Apr. 1996.
- [3] S. Barbarossa and A. Scaglione, “Parameter Estimation of Spread Spectrum Frequency-hopping Signals Using Time-frequency Distributions,” in *Proc. Int. Workshop Signal Proc. Advances in Wireless Communications (SPAWC’97)*, pp. 213–216, Apr. 1997.
- [4] M. Bolic, P. Djuric, and S. Hong, “Resampling Algorithms and Architectures for Distributed Particle Filters,” *IEEE Trans. on Signal Processing*, vol. 53, no. 7, pp. 2442–2450, July 2005.
- [5] O. Brun, V. Teuliere, and J. Garcia, “Parallel Particle Filtering,” *J. Parallel and Distributed Computation*, vol. 62, no. 7, pp. 1186-1202, July 2002.

- [6] L. Devroye, *Non-uniform Random Variate Generation*, Springer-Verlag, New York, 1986.
Available on-line at <http://cgm.cs.mcgill.ca/~luc/rnbookindex.html>
- [7] P. Djuric, J.H. Kotecha, J. Zhang, Y. Huang, T. Ghirmai, M. Bugallo, J. Miguez, “Particle Filtering,” *IEEE Signal Processing Magazine*, pp. 19–38, Sep. 2003.
- [8] A. Doucet, X. Wang, “Monte Carlo Methods for Signal Processing: A Review in the Statistical Signal Processing Context,” *IEEE Signal Processing Magazine*, vol. 22, no. 6, pp. 152-170, Nov. 2005.
- [9] B. Halder, T. Kailath, “Efficient Estimation of Closely Spaced Sinusoidal Frequencies Using Subspace-based Methods,” *IEEE Signal Processing Letters*, vol. 4, no. 2, pp. 49–51, Feb. 1997.
- [10] X. Liu, N. D. Sidiropoulos, and A. Swami, “Blind High Resolution Localization and Tracking of Multiple Frequency Hopped Signals,” *IEEE Trans. Signal Processing*, vol. 50, no. 4, pp. 889–901, Apr. 2002.
- [11] X. Liu, N. D. Sidiropoulos, and A. Swami, “Joint Hop Timing and Frequency Estimation for Collision Resolution in Frequency Hopped Networks,” *IEEE Trans. Wireless Communications*, vol. 4, no. 6, pp. 3063–3074, Nov. 2005.
- [12] N.D. Sidiropoulos, A. Swami, A. Valyrakis, “Tracking a Frequency Hopped Signal using Particle Filtering,” in *Proc. IEEE ICASSP 2006*, May 14-19, 2006, Toulouse, France.

- [13] M. K. Simon, U. Cheng, L. Aydin, A. Polydoros, and B. K. Levitt, "Hop Timing Estimation for Noncoherent Frequency-hopped M-FSK Intercept Receivers," *IEEE Trans. Communications*, vol. 43, no. 2/3/4, pp. 1144–1154, Feb./Mar./Apr. 1995.
- [14] E. Tsakonas, N.D. Sidiropoulos, A. Swami, "Time-Frequency Analysis Using Particle Filtering: Closed-form Optimal Importance Function and Sampling Procedure for a Single Time-varying Harmonic," in *Proc. Nonlinear Statistical Signal Processing Workshop: Classical, Unscented, and Particle Filtering Methods*, Sep. 13-15, 2006, Corpus Christi College, Cambridge, U.K.
- [15] E. Tsakonas, N.D. Sidiropoulos, A. Swami, "Optimal Particle Filters for Tracking Time-Varying Harmonic and Second-Order Polynomial Phase Signals," *IEEE Trans. Signal Processing*, submitted Apr. 2, 2007.
- [16] A. Valyrakis, N.D. Sidiropoulos, A. Swami, "Multichannel Particle Filtering for Tracking a Frequency Hopped Signal," in *Proc. IEEE CAMSAP 2007*, Dec. 12-14, 2007, St. Thomas, U.S. Virgin Islands (submitted).