Optimal Particle Filters for Tracking a Time-Varying Harmonic or Chirp Signal

Efthimios E. Tsakonas, Nicholas D. Sidiropoulos, *Senior Member, IEEE*, and Ananthram Swami, *Fellow, IEEE*

***Abstract*—We consider the problem of tracking the time-varying (TV) parameters of a harmonic or chirp signal using particle filtering (PF) tools. Similar to previous PF approaches to TV spectral analysis, we assume that the model parameters (complex amplitude, frequency, and frequency rate in the chirp case) evolve according to a Gaussian AR(1) model; but we concentrate on the important special case of a single TV harmonic or chirp. We show that the optimal importance function that minimizes the variance of the particle weights can be computed in closed form, and develop procedures to draw samples from it. We further employ Rao–Blackwellization to come up with reduced-complexity versions of the optimal filters. The end result is custom PF solutions that are considerably more efficient than generic ones, and can be used in a broad range of important applications that involve a single TV harmonic or chirp signal, e.g., TV Doppler estimation in communications, and radar.**

***Index Terms*—carrier frequency offset, chirp, Doppler, particle filtering, polynomial phase, radar, time-frequency analysis, time-varying harmonic, tracking.**

I. Introduction and Data Model

SPECTRAL analysis and time-frequency analysis are core tools in signal processing research (e.g., [6] and [17]). Time-varying (TV) spectra arise in a broad range of important applications: from speech, to radar, to wireless communications.

TV spectral analysis tools range from basic nonparametric approaches such as the spectrogram, to the Wigner–Ville and other time-frequency distributions, and on to parametric ones such as polynomial basis expansion models, and TV line spectra mixture models.

Line spectra mixtures (whether stationary or TV) entail a nonlinear observation equation, which complicates parameter estimation. When the evolution of model parameters can be captured in state-space form, particle filtering (PF) tools become particularly appealing for tracking the model parameters, and there have been several contributions in the recent literature dealing with PF approaches to TV spectrum estimation [1], [2], [5], [12], [13], [21].

PF algorithms for tracking time-varying phase and amplitude are considered in [2]. While it is possible to derive instantaneous frequency and frequency rate estimates by taking successive phase differences, such an indirect approach is ad hoc and problematic in practice.

For a multicomponent TV harmonic mixture model, PF approaches have been pursued in [1] and [12]. In [1], the evolution of harmonic parameters (frequencies, complex amplitudes, possibly also decay rates) follows a moving average (MA) model, the measurement follows a Gaussian TV autoregressive (TVAR) model, and an improved auxiliary particle filtering algorithm is applied to track the parameters. In [12], a Gaussian random walk model is employed for the evolution of the parameters, and an unscented PF algorithm is adapted to track them. The use of temporal slices of the spectrogram in the measurement equation of [12] limits the attainable time-frequency resolution. Follow-up work in [13] uses the spectrogram to design the importance distribution for the frequency, the underlying assumption being that frequency is locally constant (see also [5] and [21] for an application of TVAR modeling to the enhancement of speech signals).

Gaussian AR models of the evolution of harmonic mixture parameters are plausible and convenient in many situations—e.g., they can capture smoothness due to inertia or other physical constraints. Following [1] and [12], we also assume that the parameters (complex amplitude, frequency, and frequency rate in the chirp case) evolve according to a Gaussian AR(1) model; but we concentrate on the important special case of a single TV harmonic or chirp signal.

The specific model we use for a TV harmonic is as follows. Let $\mathbf{x}_k := [\omega_k, A_k]^T$ denote the state at time k, where[1] $\omega_k \in \Re$ and $A_k \in \mathbb{C}$ denote instantaneous frequency and complex amplitude. The state is assumed to evolve according to the following AR(1) model:

$$\mathbf{x}_k = \mathbf{H}\mathbf{x}_{k-1} + [\, u_{k-1} \quad w_{k-1} \,]^T$$

where $\mathbf{H}$ is 2×2 diagonal, $\mathbf{H} = \operatorname{diag}\left([b_1, b_2]^T\right)$, with b_ℓ equal to $1 - \epsilon_\ell$ (with $\epsilon_\ell > 0$ typically small, e.g., $\epsilon_\ell = 10^{-3}$). The process noise sequence is independent and identically dis-

Manuscript received April 2, 2007; revised May 7, 2008. First published June 20, 2008; current version published September 17, 2008. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Antonio Napolitano. An earlier version of part of this work appears in conference form in the *Proceedings of the IEEE Nonlinear Statistical Signal Processing Workshop* (NSSPW), Corpus Christi College, Cambridge, U.K., September 13–15, 2006. Supported in part by the Army Research Laboratory (ARL) through participation in the ARL Collaborative Technology Alliance (ARL-CTA) for Communications and Networks under Cooperative Agreement DADD19-01-2-0011, and in part by ARO under ERO Contract N62558-03-C-0012.

E. E. Tsakonas and N. D. Sidiropoulos are with the Department of Electronic and Computer Engineering, Technical University of Crete, 73100 Chania—Crete, Greece (e-mail: etsakwnas@gmail.com; nikos@telecom.tuc.gr).

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

Digital Object Identifier 10.1109/TSP.2008.927462

[1] $\omega_k = \Omega_k T_s$, where Ω_k is the instantaneous frequency of the underlying continuous-time signal at time $t = kT_s$, and T_s is the sampling period. We are interested in estimating ω_k. There is potential for aliasing due to sampling, but we are interested in tracking small offsets and slow drifts.

tributed (i.i.d.). The process noise vector at time k consists of two independent random variables with the following marginal statistics:

$$\left[u_{k-1} \quad w_{k-1}\right]^T \sim \left[\mathcal{N}\left(0,\sigma_\omega^2\right), \mathcal{CN}(0,2\sigma_A^2)\right]^T$$

where $\mathcal{N}$, $\mathcal{CN}$ stand for the (real) normal and circularly symmetric complex normal distribution, respectively. The measurements are related to the state via the measurement equation

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

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

Given a sequence of observations $\{y_k\}_{k=1}^T$, the problem of interest is to estimate the sequence of posterior densities, that is $p\left(\mathbf{x}_k|\{y_l\}_{l=1}^k\right)$, $k\in\{1,\ldots,T\}$. Given $p\left(\mathbf{x}_k|\{y_l\}_{l=1}^k\right)$, one can estimate $\mathbf{x}_k$ via the associated (posterior) mean.

For the above model (and its extension to a TV chirp), we show that the optimal importance function (that minimizes the variance of the particle weights) can be computed in closed form, and develop procedures to draw samples from it. Computing the optimal important function in closed form was not possible for the models in [1], [2], [5], [12], [13], and [21]. We further employ Rao–Blackwellization to come up with reduced-complexity versions of the optimal filters. The resulting filters are considerably more efficient than generic ones, and can be applied in a broad range of applications in digital communications and radar, such as tracking Doppler frequency and frequency rate drift due to irregular motion.

The above model may appear benign in its simplicity, but it is not. First, the measurement nonlinearity is severe. Second, in contrast to a general time-varying phase model, we explicitly model variations in instantaneous frequency. That is, we constrain the phase to be an affine function of time k, but allow time-varying jitter in the slope and the offset. These are precisely the parameters of interest in wireless communications applications. To appreciate the nature of the model, the following illustration is instructive. Fig. 1 depicts a sample path of the evolution of the frequency variable, generated using $b_1 = 0.999$, $\sigma_\omega = 0.001$ and $\omega_0 = 0$. Time variation is—purposefully—extremely slow: the frequency hovers around zero (notice the scaling of the y-axis). Fig. 2 depicts the result of frequency estimation by peak-picking the spectrogram of the noiseless measurements (amplitude fixed to 1 for clarity), using a rectangular window of length 8, maximum overlap, and zero-padding to 256 samples. The result may be surprising at first sight: one would perhaps expect the spectrogram-estimated frequency to hover around zero as well, instead of steadily diverging towards white noise-like behavior. The following simple result, whose proof can be found in the Appendix, sheds light on this "paradox."

1) Claim 1: Consider $e^{j\omega k}$, where k is a constant and ω is a random variable with continuous probability density function (pdf) $f_\omega(\cdot)$. As $k\to\infty$, the pdf of the angle of $e^{j\omega k}$ approaches a uniform pdf over $[0,2\pi)$.

Under our AR(1) model, $e^{j\omega_k k}$ can be written as a function of $e^{j\omega_{k-1}(k-1)}$ times $e^{ju_{k-1}k}$. It follows that $e^{j\omega_k k}$ is asymptotically independent of $e^{j\omega_{k-1}(k-1)}$. In other words, even if we know the frequency at the previous time step (in which case the new frequency is known within small tolerance, due to the driving term), for large k the angle will be uniformly distributed—thus carrying no information about the new frequency. The situation is worse with chirps, due to the presence of the additional quadratic term in the exponent. Clearly, any tracking algorithm (not only the spectrogram or PF) will simply diverge after a certain point in time.[2] The question is, Which approach is best for small to moderate k, and stays on-track longer than others? This is what we explore in the sequel. Our simulations indicate that PF approaches are far better than the spectrogram in this context.

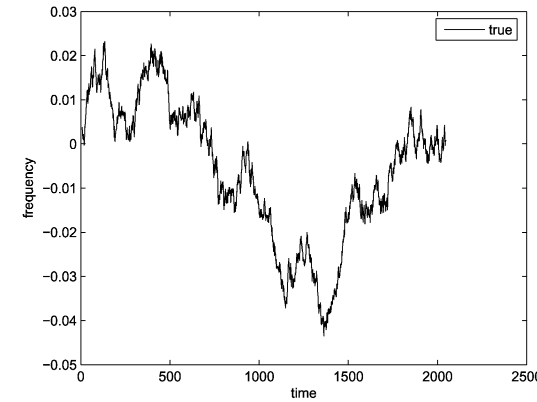

Fig. 1. True frequency hovers around zero (notice scaling of y-axis).

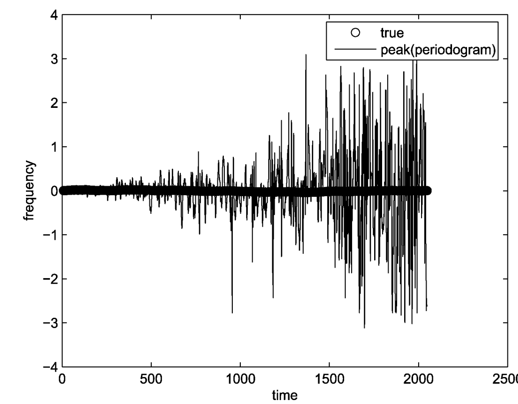

Fig. 2. Peak-picking the spectrogram corresponding to Fig. 1 (fixed complex amplitude $=1$, noiseless measurement, rectangular window of length 8, maximum overlap, zero-padding to 256 samples).

Remark 1: One might be tempted to think about periodically resetting the time axis by exploiting the shift property of complex exponentials and absorbing the resulting factor in the phase term. The spectrogram, however, operates on chunks of data without regard to a time reference—effectively resetting the time counter for every new window it processes—yet it suffers from divergence. Furthermore, periodic resetting of the time axis would introduce abrupt periodic changes in the phase, which are inconsistent with phase noise.

Remark 2—Link to Weil's Theorem: Weil's theorem (e.g., see [15]) asserts that the distribution of the fractional part of $\{fk\}_{k\in\mathbb{Z}_+}$, for f irrational (and fixed; $\mathbb{Z}_+$ denotes positive integers) is uniform in $[0,1)$. In the context of Claim 1, let $\omega = 2\pi f$,

[2]In certain applications in digital communications, detecting the onset of divergence could trigger a cold start at the link level to reacquire synchronization using training data.

and $\langle\cdot\rangle$ denote fractional part. Then $e^{j\omega k} = e^{j2\pi fk} = e^{j2\pi\langle fk\rangle}$. The pdf of ω has been assumed continuous, and thus a realization of f will be irrational with probability one. Weil's theorem then shows that the *sample* (empirical) distribution of the angle of $e^{j\omega k}$ for a fixed realization of ω and all k is uniform over $[0, 2\pi)$. In contrast, Claim 1 asserts that the *ensemble* distribution of the angle of $e^{j\omega k}$ is (approximately) uniform over $[0, 2\pi)$ for a fixed large k and ω random with continuous pdf. So, Weil's Theorem applies to sample path averages, whereas Claim 1 to asymptotic ensemble averages. The ensemble distribution converges to the sample path distribution for large $k \in \mathbb{Z}_+$; this is an ergodic property of the random process $e^{j\omega k}$. Interestingly, Claim 1 does not require k to be integer.

II. Particle Filtering

Particle filtering has emerged as an important sequential state estimation method for stochastic nonlinear and/or non-Gaussian state-space models, for which it provides a powerful alternative to the commonly used extended Kalman filter. See [3], [8], and [9] for recent tutorial overviews.

In particle filtering, continuous distributions are approximated by discrete random measures, comprising "particles" and associated weights. That is, a continuous distribution $p(\mathbf{x}_k)$ (k is a time index) is approximated as

$$p(\mathbf{x}_k) \approx \sum_{n=1}^{N} w_{n,k}\delta(\mathbf{x} - \mathbf{x}_{n,k})$$

where $\delta(\cdot)$ denotes the Dirac delta functional, $\mathbf{x}_{n,k}$ is the nth particle (location) for time k and $w_{n,k}$ is the associated weight. A useful simplification stemming from this approximation is that the computation of pertinent expectations and conditional probabilities reduces to summation, as opposed to integration. While this can also be accomplished via direct discretization over a fixed grid, the use of a random measure affords flexibility in adapting the particle locations to better fit the distribution of interest.

If we aim for an online filtering algorithm, in which the state at time k should be estimated from measurements up to and including time k, the key distribution of interest is the posterior density $p\left(\mathbf{x}_k \mid \{y_l\}_{l=1}^{k}\right)$. The basic idea of particle filtering, then, is to begin with a random measure approximation of the initial state distribution, and, as measurements become available, derive updated random measure approximations of $p\left(\mathbf{x}_k \mid \{y_l\}_{l=1}^{k}\right), k \in \{1, 2, \ldots\}$. That is, we seek random measure approximations

$$\hat{p}\left(\mathbf{x}_k \mid \{y_l\}_{l=1}^{k}\right) = \sum_{n=1}^{N} w_{n,k}\delta(\mathbf{x}_k - \mathbf{x}_{n,k})$$

from which the state at time k can be estimated via the associated posterior mean $\hat{\hat{\mathbf{x}}}_k = \sum_{n=1}^{N} w_{n,k}\mathbf{x}_{n,k}$. In particle filtering, the updates—the derivation of $\hat{p}\left(\mathbf{x}_k \mid \{y_l\}_{l=1}^{k}\right)$ from $\hat{p}\left(\mathbf{x}_{k-1} \mid \{y_l\}_{l=1}^{k-1}\right)$—are based on the Bayes rule [3], [8].

A random measure approximation comprises two components: the particles (locations) and the associated weights. If we could sample from the sought posterior $p\left(\mathbf{x}_k \mid \{y_l\}_{l=1}^{k}\right)$, then all particle weights would have been equal. Unfortunately, such direct sampling is not possible in most cases, and thus we resort to sampling from a so-called *importance function* that "resembles" the desired posterior, and from which samples can be drawn with relative ease. The mismatch between the sought density and the importance function is compensated in the calculation of weights, chosen proportional to their ratio evaluated at each particle [3], [8].

Different types of particle filters may be applied to a given state-space model. The various particle filters primarily differ in the choice of importance (or, *proposal*) function. Different importance functions yield different estimation performance—complexity tradeoffs. Perhaps the most intuitive choice of importance function is the *prior importance function* $p(\mathbf{x}_k \mid \mathbf{x}_{n,k-1})$; i.e., the nth particle is updated by propagating it through the state-evolution part of the system. This is a common choice, for simplicity considerations. The drawback is that particles evolve without regard to the latest measurement, which only comes into play in the ensuing weight update. When using the prior importance function, the weight update at time instant k is given by $w_{n,k} = w_{n,k-1}p(y_k \mid \mathbf{x}_{n,k})$, followed by normalization to enforce $\sum_{n=1}^{N} w_{n,k} = 1$.

Regardless of the particular importance function employed, a common problem in particle filtering is *degeneracy*: the weights of all but a few particles tend to become negligible after a few iterations [3], [8]. Degeneracy can be detected via degeneracy measures, and mitigated via *resampling* techniques [3], [8]. Resampling the discrete measure replicates particles with large weights and removes those with negligible weights. All particle weights become equal after resampling. There exist several computationally efficient $[O(N)]$ resampling schemes that can be used to avoid the quadratic cost of brute-force resampling [3], [8].

From the viewpoint of minimizing the variance of the weights, the optimal importance function (OIF) is given by [3], [8]

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

where $\mathbf{x}_{n,k} := [\omega_{n,k}, A_{n,k}]^T$ denotes the nth particle at time k, which is computed by plugging the nth particle at time $k-1$ into the OIF above, and drawing a sample from it. The OIF usually strikes a better performance—complexity tradeoff than other alternatives. There are, however, two difficulties associated with the use of the OIF. First and foremost, it requires integration to compute the normalization factor, which is usually intractable due to nonlinearity. Second, sampling from the optimal importance function is a rather complicated process. Thankfully, for our particular model, it turns out that it is possible to carry out the integration analytically. This is explained next.

III. Optimal Importance Function: TV Harmonic Case

Define a dummy variable $\mathbf{x} := [\omega, A]^T$, and let $D(y_k, \mathbf{x}_{n,k-1}) := \int_{\mathbf{x}} p(y_k|\mathbf{x})p(\mathbf{x}|\mathbf{x}_{n,k-1})d\mathbf{x}$. Then [see the first equation at the bottom of the page]. Letting $m_A := b_2 A_{n,k-1}$, $m_\omega := b_1 \omega_{n,k-1}$, $v := \angle y_k - \angle m_A$, where $\angle(\cdot)$ extracts the angle of its argument, it can be shown[3] that

$$D(y_k, \mathbf{x}_{n,k-1}) = \frac{1}{2\pi\left(\sigma_A^2 + \sigma_n^2\right)} \times e^{-\left(|y_k|^2 + |m_A|^2/2\left(\sigma_A^2+\sigma_n^2\right)\right)} \times \mathcal{B}$$

with the multiplicative factor $\mathcal{B}$ given by

$$\mathbf{I}_0\left(\frac{|m_A||y_k|}{\sigma_A^2+\sigma_n^2}\right) + 2\sum_{\ell=1}^{+\infty} \mathbf{I}_\ell\left(\frac{|m_A||y_k|}{\sigma_A^2+\sigma_n^2}\right) \times e^{-((k\sigma_\omega)^2\ell^2/2)} \cos(\ell k m_\omega - \ell v)$$

where $\mathbf{I}_\ell(\cdot)$ denotes the modified Bessel function of the first kind of order ℓ. The sum term for $\mathcal{B}$ is quite interesting. Due to the negative exponential dependence on k, ℓ, and the properties of modified Bessel functions, it vanishes quickly with k and ℓ. Given y_k, it is easy to come up with a closed-form upper bound on the truncation error, which is, however, overly conservative. Truncation to 20 terms is adequate in all cases considered in our experiments—adding more terms does not affect the results. We used 100 terms as an extra safety margin in our simulations.

We can use rejection [7, pp. 40–42] to generate samples from the optimal importance function [see the last equation at the bottom of the page].

The basic idea of rejection-based sampling can be summarized as follows [7, pp. 40–42]. Suppose 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. The rejection method then works as follows.

Algorithm 1:

1) Draw a sample $\mathbf{x}$ from $g(\cdot)$ and an independent sample U uniformly distributed in $[0, 1]$).
2) Set $\tau := c(g(\mathbf{x})/\phi(\mathbf{x}))$.
3) If $U\tau \leq 1$, then accept and return $\mathbf{x}$; else reject and go to Step 1.

[3]Detailed derivation of this and other results in the paper are available as supplementary material in ieeexplore.

It can be shown that the above rejection method generates samples from the desired density $\phi(\cdot)$, and the mean number of iterations until a sample is accepted is c (thus the desire to keep $c \geq 1$ as small as possible). Furthermore, the distribution of the number of trials is geometric with parameter $1-(1/c)$, which means that the probabilities of longer trials decay exponentially [7, p. 42].

Let

$$\sigma^2 := \frac{\sigma_A^2 \sigma_n^2}{\sigma_A^2 + \sigma_n^2}$$

and

$$\mu := \frac{\sigma_A^2 |y_k| + \sigma_n^2 |m_A|}{\sigma_A^2 + \sigma_n^2}.$$

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

$$g(\mathbf{x}_k \mid \mathbf{x}_{n,k-1}, y_k) = \frac{e^{-\left((\omega_k - m_\omega)^2/2\sigma_\omega^2\right)} e^{-\left((|A_k|-\mu)^2/2\sigma^2\right)}}{(2\pi)^2 \gamma(\mu,\sigma)\, \sigma_\omega \sigma}$$

where

$$c := \frac{e^{-\left(\mu^2/2\sigma^2\right)} + \frac{\sqrt{2\pi}}{\sigma} Q_o\left(-\frac{\mu}{\sigma}\right)}{e^{-\left(|m_A||y_k|/\sigma_A^2+\sigma_n^2\right)}\mathcal{B}}$$

$$\gamma(\mu,\sigma) := Q_o\left(-\frac{\mu}{\sigma}\right) + \frac{\sigma}{\sqrt{2\pi}} e^{-\left(\mu^2/2\sigma^2\right)}$$

$$Q_o\left(-\frac{\mu}{\sigma}\right) := \int_{r=0}^{+\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{-\left((r-\mu)^2/2\sigma^2\right)} dr = \frac{1}{2}\operatorname{erfc}\left(-\frac{\sigma_A^2|y_k| + \sigma_n^2|m_A|}{\sigma_A \sigma_n \sqrt{2\left(\sigma_A^2+\sigma_n^2\right)}}\right).$$

For this particular choice of IF and sampling procedure the weight update step is given by $\mathbf{w}_{n,k} = \mathbf{w}_{n,k-1} D(y_k, \mathbf{x}_{n,k-1})$ and can be carried out before sampling from the optimal importance function (before the particles are propagated to time-step k).

IV. Rao–Blackwellization

For our particular state-space model, it is possible to reduce the dimensionality of the problem via a technique known as Rao–Blackwellization (see [10], [11], [18], and references therein). Conditioned on frequency, our model is AR(1) linear Gaussian on the complex amplitude. The basic idea is to exploit

$$D(y_k, \mathbf{x}_{n,k-1}) = \int_{\omega\in\Re}\int_{A\in\mathbb{C}} \frac{1}{2\pi\sigma_n^2} e^{-(|y_k - Ae^{j\omega k}|^2/2\sigma_n^2)} \left[\frac{1}{\sqrt{2\pi}\sigma_\omega} e^{-((\omega - b_1\omega_{n,k-1})^2/2\sigma_\omega^2)} \frac{1}{2\pi\sigma_A^2} e^{-(|A - b_2 A_{n,k-1}|^2/2\sigma_A^2)}\right] dA d\omega.$$

$$p(\mathbf{x}_k|\mathbf{x}_{n,k-1}, y_k) = \frac{\frac{1}{2\pi\sigma_n^2} e^{-\left(|y_k - A_k e^{j\omega_k k}|^2/2\sigma_n^2\right)} \frac{1}{\sqrt{2\pi}\sigma_\omega} e^{-\left((\omega_k - m_\omega)^2/2\sigma_\omega^2\right)} \frac{1}{2\pi\sigma_A^2} e^{-\left(|A_k - m_A|^2/2\sigma_A^2\right)}}{\frac{1}{2\pi\left(\sigma_A^2+\sigma_n^2\right)} e^{-\left(|y_k|^2 + |m_A|^2/2\left(\sigma_A^2+\sigma_n^2\right)\right)}\mathcal{B}}.$$

this structure to avoid computing everything with plain Monte Carlo sampling. The particle filter is only used to handle the purely nonlinear portion of the state-space.

Reference [18] considers a general nonlinear state-space model that contains a conditionally linear part, and works out the Rao–Blackwellization procedure in detail. Our particular model is a special case of the so-called *Diagonal Model* in [18]; however, we use the OIF to draw samples for the nonlinear part. The choice of importance function is left open in [18] to maintain generality—usually the OIF cannot be computed analytically.

The desired posterior pdf at time k, $p\left(\omega_k, A_k \mid \{y_l\}_{l=1}^k\right)$ can be written as

$$p\Big(\omega_k, A_k \mid \{y_l\}_{l=1}^k\Big) = p\Big(A_k \mid \omega_k, \{y_l\}_{l=1}^k\Big)\, p\Big(\omega_k \mid \{y_l\}_{l=1}^k\Big).$$

This factorization enables us to use particles only to approximate $p\left(\omega_k \mid \{y_l\}_{l=1}^k\right)$, which is a one-dimensional pdf; $p\left(A_k \mid \omega_k, \{y_l\}_{l=1}^k\right)$ can then be analytically computed using the Kalman filter. For state estimation, a Kalman filter is associated to each frequency particle, and the conditional mean filtered estimate of the Kalman filter is used to fill-in the "missing" amplitude dimension.

We use the optimal importance distribution to approximate the marginal posterior density $p\left(\omega_k \mid \{y_l\}_{l=1}^k\right)$. The optimal importance distribution is

$$p\left(\omega_k \mid \omega_{n,k-1}, y_k\right) = \frac{p\left(y_k \mid \omega_k\right) p\left(\omega_k \mid \omega_{n,k-1}\right)}{\int_\omega p\left(y_k \mid \omega\right) p(\omega \mid \omega_{n,k-1}) d\omega}.$$

Letting $\mu_A := b_2^k E\{A_0\}$, $\sigma_{A'}^2 := b_2^{2k} E\left\{|A_0|^2\right\} + \left(1 - b_2^{2k}/1 - b_2^2\right)\left(2\sigma_A^2\right)$, $u := \angle y_k - \angle \mu_A$, it can be shown that [see the equation at the bottom of the page] with

$$\mathcal{B}' = \mathbf{I}_0\left(\frac{|\mu_A||y_k|}{\sigma_{A'}^2 + \sigma_n^2}\right) + 2\left(\sum_{\ell=1}^{+\infty} \mathbf{I}_\ell\left(\frac{|\mu_A||y_k|}{\sigma_{A'}^2 + \sigma_n^2}\right) e^{-\left((k\sigma_\omega)^2/2\right)\ell^2} \cos(\ell k m_\omega - \ell u)\right).$$

The weight update is given by $\mathbf{w}_{n,k} = \mathbf{w}_{n,k-1} D(y_k, \omega_{n,k-1})$, with

$$D(y_k, \omega_{n,k-1}) := \frac{1}{2\pi\left(\sigma_{A'}^2 + \sigma_n^2\right)} \times e^{-\left(|y_k|^2 + |\mu_A|^2/2\left(\sigma_{A'}^2 + \sigma_n^2\right)\right)} \mathcal{B}'.$$

To generate samples distributed according to $p\left(\omega_k \mid \omega_{n,k-1}, y_k\right)$, we could employ the transformation method [7]: this is, after all, a one-dimensional pdf. Still, this requires another integration and some level of approximation (the integral cannot be put in closed form). As an alternative, we found that rejection for this one-dimensional pdf is far more efficient than in the previous case (which involved three real dimensions), and delivers exact samples, which is a definite advantage relative to other sampling methods. A common criticism of rejection for real-time applications is that it takes a random number of draws per particle. With as few as 30 to 50 particles, however, variance is averaged out and the complexity per input measurement is stable enough for our purposes.

Starting from $p\left(\omega_k \mid \omega_{n,k-1}, y_k\right)$ and using the triangle inequality, it is straightforward to show that a suitable dominating density is the transitional prior $p\left(\omega_k | \omega_{n,k-1}\right)$. The constant c associated with the accept-reject algorithm becomes

$$c = \frac{1}{e^{-\left(|\mu_A||y_k|/\sigma_{A'}^2 + \sigma_n^2\right)} \mathcal{B}'}.$$

It is interesting to see that sampling from the optimal importance function can be implemented by rejection over the transitional prior, which is commonly used as importance function *per se*. Pseudo-code for the Rao–Blackwellized optimal filter can be found in Table I.

V. Cramér–Rao Lower Bound

The Cramér–Rao lower bound (CRLB) for our model can be computed using the recursive formula of Tichavsky *et al.* [20] for the calculation of the Fisher information matrix, $\mathbf{J}_k$. The state equation in our particular model is linear, Gaussian; this allows considerable simplification of the general result in [20], thus yielding

$$\mathbf{J}_k = \mathbf{D}_{k-1}^{22} - \mathbf{D}_{k-1}^{21}\left(\mathbf{J}_{k-1} + \mathbf{D}_{k-1}^{11}\right)^{-1} \mathbf{D}_{k-1}^{12}, \quad k \geqslant 0$$

with

$$\mathbf{D}_{k-1}^{11} := -\mathbf{E}\left\{\nabla_{\mathbf{x}_{k-1}}\left[\nabla_{\mathbf{x}_{k-1}} \log \mathbf{p}(\mathbf{x_k} \mid \mathbf{x_{k-1}})\right]^T\right\}$$

$$\mathbf{D}_{k-1}^{12} := \left[\mathbf{D}_{k-1}^{21}\right]^T = -\mathbf{E}\left\{\nabla_{\mathbf{x}_k}\left[\nabla_{\mathbf{x}_{k-1}} \log \mathbf{p}(\mathbf{x_k} \mid \mathbf{x_{k-1}})\right]^T\right\}$$

and

$$\mathbf{D}_{k-1}^{22} := -\mathbf{E}\left\{\nabla_{\mathbf{x}_k}\left[\nabla_{\mathbf{x}_k} \log \mathbf{p}(\mathbf{x}_k \mid \mathbf{x}_{k-1})\right]^T\right\} - \mathbf{E}\left\{\nabla_{\mathbf{x}_k}\left[\nabla_{\mathbf{x}_k} \log \mathbf{p}(\mathbf{y}_k \mid \mathbf{x}_k)\right]^T\right\}.$$

$$p\left(\omega_k \mid \omega_{n,k-1}, y_k\right) = \frac{\frac{1}{2\pi\left(\sigma_{A'}^2 + \sigma_n^2\right)} e^{-\left(|y_k - \mu_A e^{j\omega_k k}|^2/2\pi\left(\sigma_{A'}^2 + \sigma_n^2\right)\right)} \frac{1}{\sqrt{2\pi}\sigma_\omega} e^{-\left((\omega_k - m_\omega)^2/2\sigma_\omega^2\right)}}{\frac{1}{2\pi\left(\sigma_{A'}^2 + \sigma_n^2\right)} e^{-\left(|y_k|^2 + |\mu_A|^2/2\left(\sigma_{A'}^2 + \sigma_n^2\right)\right)} \mathcal{B}'}$$

TABLE I
RBPF USING OIF FOR TRACKING A SINGLE TIME-VARYING HARMONIC (SEE TEXT FOR DEFINITION OF CONSTANTS)

$\left[\{\omega_k^i, m_{A_k}^i, P_{A_k}^i\}_{i=1}^N\right] = RBPF\left[\{\omega_{k-1}^i, m_{A_{k-1}}^i, P_{A_{k-1}}^i\}_{i=1}^N, y_k\right]$

1) Compute normalized importance weights
   - FOR i=1:N,
     $\tilde{\mathbf{w}}_k^i = \frac{1}{2\pi(\sigma_{A'}^2+\sigma_n^2)} e^{-\frac{|y_k|^2+|\mu_A|^2}{2(\sigma_{A'}^2+\sigma_n^2)}} \times \mathcal{B}'$
   - END FOR
   - FOR i=1:N,
     - Normalize: $\mathbf{w}_k^i = \tilde{\mathbf{w}}_k^i / sum\left[\{\tilde{\mathbf{w}}_k^i\}_{i=1}^N\right]$
   - END FOR

2) Resample → equally weighted particles
   $\left[\{\omega_{k-1}^i, m_{A_{k-1}}^i, P_{A_{k-1}}^i\}_{i=1}^N\right] =$
   $RESAMPLE\left[\{\omega_{k-1}^i, m_{A_{k-1}}^i, P_{A_{k-1}}^i, \mathbf{w}_k^i\}_{i=1}^N\right]$

3) Sample from the optimal importance density $p\left(\omega_k|\omega_{k-1}, y_k\right)$:
   - FOR i=1:N,
     - Calculate $c := e^{\frac{|\mu_A||y_k|}{\sigma_{A'}^2+\sigma_n^2}} / \mathcal{B}'$
     - Set $U := 1/eps$ and $\tau := 1/eps$
   - WHILE $(U\tau > 1)$
     - Draw a candidate frequency sample from the dominating density $p\left(\omega_k|\omega_{k-1}^i\right)$:
       $\omega_k^i \sim \mathcal{N}\left(b_1\omega_{k-1}^i, \sigma_\omega^2\right)$
     - Set the acceptance parameter associated with rejection:
       $\tau = c\frac{\text{Dominating}(\omega_k^i)}{\text{Optimal}(\omega_k^i)}$
     - Draw a sample $U \sim \text{Uniform}[0,1]$
   - END WHILE
   - END FOR

4) Use the Kalman Filter relations to obtain analytically the $\{m_{A_k}^i, P_{A_k}^i\}$ associated with each frequency sample:
   - FOR i=1:N,
     $\left[\, m_{A_k}^i, P_{A_k}^i \,\right] = KF\left[\, \omega_k^i, m_{A_{k-1}}^i, P_{A_{k-1}}^i,\; y_k \,\right]$
   - END FOR

At this point, it is convenient to rewrite our model in real-valued form. Upon defining $\mathbf{x}_k' := [\omega_k, \Re(A_k), \Im(A_k)]^T$, where $\Re(\cdot), \Im(\cdot)$ extract the real, respectively imaginary part, we have

$$\begin{aligned}\mathbf{x}_k' &= \mathbf{H}'\mathbf{x}_{k-1}' + \mathbf{u}_{k-1}\\ \mathbf{y}_k &= [\Re\{A_k e^{j\omega_k k}\} \quad \Im\{A_k e^{j\omega_k k}\}]^T + \mathbf{v}_k\end{aligned}$$

where $\mathbf{H}' = \text{diag}\left([b_1, b_2, b_2]^T\right)$, with b_ℓ being $1 - \epsilon_\ell$, $\mathbf{u}_{k-1} \sim \mathcal{N}(\mathbf{0}, \mathbf{Q})$ with $\mathbf{Q} = \text{diag}\left(\left[\sigma_\omega^2, \sigma_A^2, \sigma_A^2\right]^T\right)$, and $\mathbf{v}_k \sim \mathcal{N}(\mathbf{0}, \mathbf{R})$ with $\mathbf{R} = \text{diag}\left(\left[\sigma_n^2, \sigma_n^2\right]^T\right)$. Then

$$\begin{aligned}\mathbf{D}_{k-1}^{11} &= \mathbf{H}'^T\mathbf{Q}^{-1}\mathbf{H}'\\ \mathbf{D}_{k-1}^{12} &= \left[\mathbf{D}_{k-1}^{21}\right]^T := -\mathbf{H}'^T\mathbf{Q}^{-1}\\ \mathbf{D}_{k-1}^{22} &= \mathbf{Q}^{-1} + \mathbf{E}\left\{\tilde{\mathbf{F}_\mathbf{k}}^T\mathbf{R}^{-1}\tilde{\mathbf{F}_\mathbf{k}}\right\}\end{aligned}$$

with $\tilde{\mathbf{F}_\mathbf{k}}$ being the $\mathbf{2}\times\mathbf{3}$ matrix

$$\tilde{\mathbf{F}_\mathbf{k}} = \nabla_{\mathbf{x}_k'}\left[\Re\{A_k e^{j\omega_k k}\} \quad \Im\{A_k e^{j\omega_k k}\}\right]^T.$$

For $\mathbf{D}_{k-1}^{11}$ and $\mathbf{D}_{k-1}^{12}$, note that the expectation operator was dropped because the respective Jacobians are independent of the target state. The expectation operator in the expression for $\mathbf{D}_{k-1}^{22}$ can be easily estimated using MC integration; it can also be calculated analytically, albeit the resulting formula appears cumbersome. Putting terms together yields

$$\begin{aligned}\mathbf{J}_k =& \mathbf{Q}^{-1} + \mathbf{E}\left\{\tilde{\mathbf{F}_\mathbf{k}}^T\mathbf{R}^{-1}\tilde{\mathbf{F}_\mathbf{k}}\right\}\\ &- \mathbf{Q}^{-1}\mathbf{H}'(\mathbf{J}_{k-1}+\mathbf{H}'^T\mathbf{Q}^{-1}\mathbf{H}')^{-1}\mathbf{H}'^T\mathbf{Q}^{-1}, \qquad k \geqslant 0.\end{aligned}$$

The initial density $\mathbf{p}(\mathbf{x}_0)$ is taken to be $\mathcal{N}(\bar{\mathbf{x}}_0, \mathbf{Q}_0)$, in which case $\mathbf{J}_0 = \mathbf{Q}_0^{-1}$.

VI. Numerical Results: TV Harmonic Case

In our simulations, we benchmark the performance of our optimal particle filters against the CRLB and four additional filters: the extended Kalman filter, the SIR PF [14], the auxiliary PF, and a regularized PF. These filters are briefly discussed next.

A. Extended Kalman Filter (EKF)

The EKF equations are well known, but they are rewritten here for convenience. Recall from the previous section the real-valued state-space model. Since the state equation is linear, state prediction is performed using the standard Kalman filter equations

$$\begin{aligned}\widehat{\mathbf{x}}_{k|k-1}' &= \mathbf{H}'\widehat{\mathbf{x}}_{k-1|k-1}'\\ \mathbf{P}_{k|k-1} &= \mathbf{H}'\mathbf{P}_{k-1|k-1}\mathbf{H}'^T + \mathbf{Q}.\end{aligned}$$

Since the measurement equation is nonlinear, the filter update is carried out using

$$\begin{aligned}\widehat{\mathbf{x}}_{k|k}' &= \widehat{\mathbf{x}}_{k|k-1}' + \mathbf{K}_k\left[\mathbf{y}_k - \mathbf{h}_k(\widehat{\mathbf{x}}_{k|k-1}')\right]\\ \mathbf{P}_{k|k} &= \mathbf{P}_{k|k-1} - \mathbf{K}_k\mathbf{S}_k\mathbf{K}_k{}^T\end{aligned}$$

where $\mathbf{S}_k = \widehat{\mathbf{F}}_k\mathbf{P}_{k|k-1}\widehat{\mathbf{F}}_k^T + \mathbf{R}$, $\mathbf{K}_k = \mathbf{P}_{k|k-1}\widehat{\mathbf{F}}_k^T\mathbf{S}_k{}^{-1}$, with $\widehat{\mathbf{F}_k}$ being the $\mathbf{2}\times\mathbf{3}$ Jacobian of the nonlinearity involved in the measurement equation (denoted as $\mathbf{h}_k(\cdot)$), this time evaluated at the filter's estimate $\widehat{\mathbf{x}}_{k|k-1}'$ (see previous section).

B. Regularized PF (RPF)

This algorithm is identical to the sampling importance resampling (SIR) algorithm, which uses the prior importance function, except for a "jittering" of the resampled particles (using a normal distribution kernel) in order to protect the filter from sample impoverishment; see, e.g., [3]. Since the process noise in our model is relatively small, this modification is expected to improve the performance over the standard SIR. However, this filter also has well known disadvantages—the samples are no longer guaranteed to approximate the posterior density asymptotically in the number of particles.

C. Auxiliary SIR (AUX) Filter

The particular algorithm used is the auxiliary SIR filter introduced by Pitt and Shephard (see [16]). This filter tries to explore the state-space in a more sophisticated way than the SIR filter. This is done by resampling at the "previous" time step based on certain point estimates that capture the essential features of the posterior density. This approximation can be inefficient when

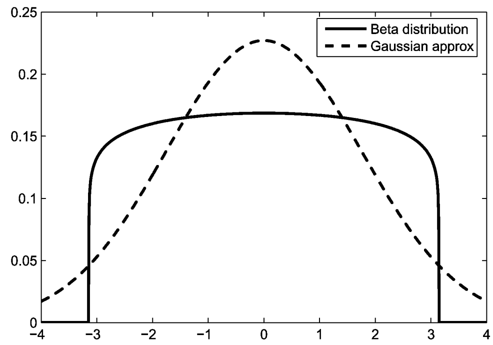

Fig. 3. Probability density of ω_0. Shape parameters: $u_1 = u_2 = 1.1$.

the process noise is large, or when the auxiliary index varies a lot for a fixed prior. When process noise is small enough, though, the AUX filter is reported to improve the performance over the standard SIR.

D. Rao–Blackwellized PF Using OIF (RBPF)

The plain version of PF using the OIF employs rejection for a thee-dimensional distribution, which is not appealing in terms of complexity. The Rao–Blackwellized version performs equally well in terms of tracking performance for the same number of particles, but is much faster—up to 100 times faster in our simulations. Therefore, we only present results for the Rao–Blackwellized version.

E. Initialization Issues

In this section, we investigate the impact of prior knowledge on the CRLB curves. We start by examining the case where almost no prior information about the frequency component of the initial state vector is available. For the initial density of the complex amplitude, we take a narrow Gaussian with mean $E[A_0] = 1 + j$ and standard deviation $\mathrm{std}[A_0] = 0.01$. A beta distribution is used to model the initial density of the frequency component

$$\mathbf{p}(\omega_0) := \frac{\Gamma(u_1+u_2)\left(\frac{\omega_0-\omega_L}{\omega_H-\omega_L}\right)^{u_1-1}\left(1-\frac{\omega_0-\omega_L}{\omega_H-\omega_L}\right)^{u_2-1}}{\Gamma(u_1)\Gamma(u_2)(\omega_H-\omega_L)}$$

for $\omega_o \in [\omega_L, \omega_H]$, where Γ stands for the Gamma function, and u_1, u_2 are the shape parameters. The beta distribution contains the uniform distribution as a special case.

While in simulations we generate ω_0 according to $\mathbf{p}(\omega_0)$, we also need a Gaussian approximation for carrying out CRLB and EKF computations, since both are premised on the assumption that the initial density is Gaussian (note that this is not required for the particle filters). The mean and standard deviation of the best-fitting Gaussian can be found in [4]

$$E[\omega_0] := \omega_L + (\omega_H - \omega_L)\frac{u_1}{u_1+u_2}$$
$$std[\omega_0] := \sqrt{\frac{(\omega_H-\omega_L)^2 u_1 \cdot u_2}{(u_1+u_2)^2(u_1+u_2+1)}}.$$

An illustration of such an approximation is presented in Fig. 3.

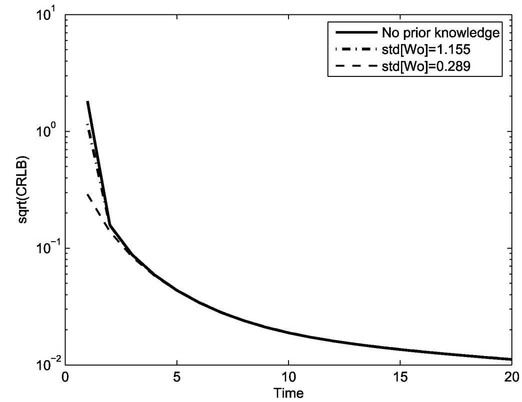

Fig. 4. Comparison of $\sqrt{\mathrm{CRLB}}$ curves (frequency estimation) for inaccurate prior information.

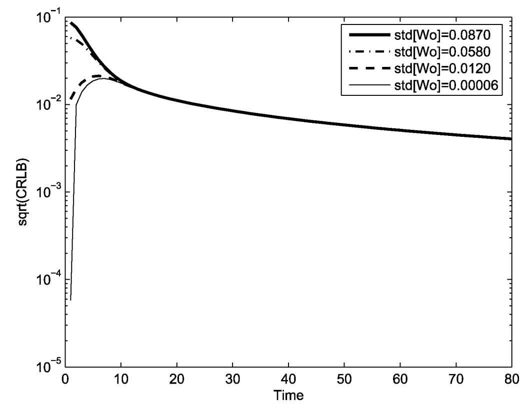

Fig. 5. Comparison of $\sqrt{\mathrm{CRLB}}$ curves (frequency estimation) for accurate prior information.

Figs. 4 and 5 demonstrate the effect of prior information on the CRLB. The following parameters were used: $b_\ell = 0.999$, $\forall \ell$, $\sigma_\omega^2 = 10^{-4}$, $\sigma_A^2 = 10^{-4}$, $\sigma_n^2 = 0.2$, $u_1 = u_2 = 1$—thus the accuracy of prior information $(\mathrm{std}[\omega_0])$ is determined by $\omega_H - \omega_L$. The expectation appearing in the CRLB formulas was approximated using 100 realizations of the state vector. Observe that the CRLB with prior knowledge is initially lower, although the significance of prior information diminishes very quickly over time and the bounds become indistinguishable for $k > 10$. Increasing the value of $\mathrm{std}[\omega_0]$, the CRLB with prior knowledge approaches the one with no prior knowledge.

F. Estimation Performance Results

We now focus on the frequency estimation performance of the five aforementioned filters in a tracking mode, wherein the initial state is assumed to be known exactly—corresponding to a Dirac delta initial distribution. The CRLB and the EKF assume that the initial density is a Gaussian. This mismatch is dealt with by using a very tight density (very small initial variance) to approximate a delta distribution. The expectation appearing in the CRLB was approximated using 100 realizations of the state vector. The error curves corresponding to the five filters were produced by averaging over 200 independent Monte Carlo (MC) runs, each comprising 100 temporal samples. The conditional mean was used to generate point estimates for the particle filters.

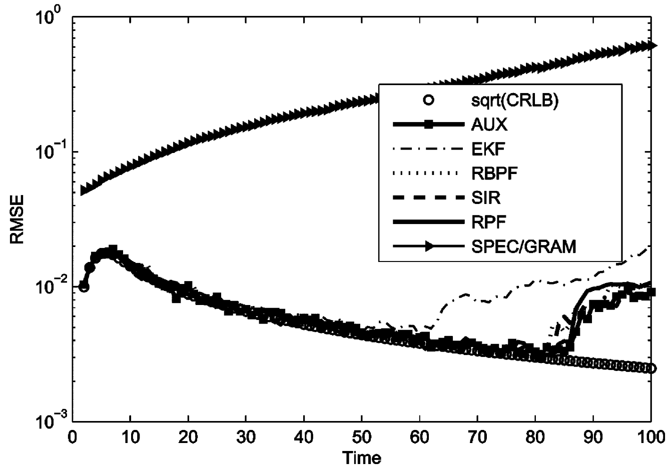

Fig. 6. RMSE (frequency estimation) comparison of the four particle filters, EKF, spectrogram and $\sqrt{\text{CRLB}}$ with accurate prior information. Number of particles: 1000 for SIR, 1000 for RPF, 800 for AUX, 50 for RBPF.

TABLE II
MEAN COMPUTATION TIMES IN SECONDS—(STVH CASE)

EKF	SPEC/GRAM	RBPF	SIR	RPF	AUX
0.00020	0.00015	0.06382	0.07569	0.16431	0.15653

System parameters were set to $b_\ell = 0.999$, $\forall\, \ell$, $\sigma_\omega^2 = 10^{-4}$, $\sigma_A^2 = 10^{-4}$, $\sigma_n^2 = 0.1$, and multinomial resampling was employed.

We compared computational and memory complexities for approximately equal estimation performance. Since accuracy is a major concern, the number of particles for each algorithm was chosen to yield RMSE close to the CRLB. Accordingly, the number of particles, N, was 1000 for SIR, 1000 for RPF, 800 for AUX, and 50 for RBPF.

The results are summarized in Fig. 6, which also includes the spectrogram peak estimator as yet another baseline. A rectangular window comprising eight samples, zero-padding to 128 samples, and maximal overlap factor were used to compute the spectrogram, followed by peak-picking to estimate the instantaneous frequency.

It is satisfying to see that the four particles filters and the EKF operate close to the CRLB, and RBPF in particular performs that well with order-of-magnitude less particles. This being a three-dimensional state-space, such good performance with only 50 particles is not at all obvious. SIR, RPF and AUX filters perform very poorly with less than a few hundred particles in this context. The average computation time per measurement (time-step) for each algorithm is listed in Table II. Observe that RBPF is the fastest among the particle filters, in addition to its far lower memory requirements.

Notice that all filters in Fig. 6 eventually diverge from the CRLB, with EKF being the first to do so. Consistent with our earlier discussion regarding Claim 1 , the spectrogram steadily diverges in this case, and from early on. Interestingly, its performance is order-of-magnitude worse than that of the particle filters.

Remark 3: We note that the particle filters are robust with respect to model parameter mismatch. In particular, RBPF using $\sigma_\omega^2 = 2 \times 10^{-4}$, $\sigma_A^2 = 2 \times 10^{-4}$, $\sigma_n^2 = 0.2$ (i.e., $2\times$ the actual variance parameters used to generate the input data) performs essentially the same as RBPF using the correct variance parameters—the only difference is that the onset of divergence appears slightly earlier (at time index 80 instead of 85).

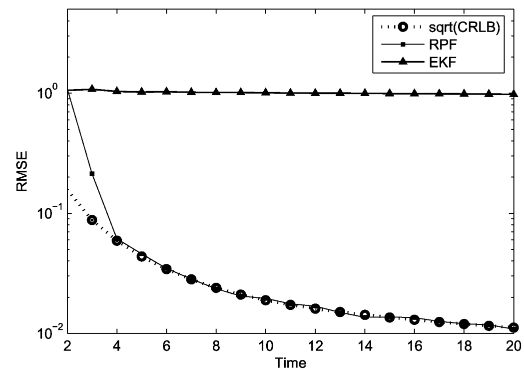

Fig. 7. RMSE (frequency estimation) comparison of RPF, EKF, and $\sqrt{\text{CRLB}}$ with inaccurate prior information.

In Fig. 6, it appears that EKF offers a good performance/complexity tradeoff in the case where the initial information is very accurate; however, its performance is severely degraded when the initial information about the frequency is coarse. In that case, the particle filters can still yield very good performance. To illustrate this, Fig. 7 presents a simple performance comparison between the EKF and RPF (with 1000 particles) when the initial frequency information is inaccurate with $\text{std}[\omega_0] = 1.155$, and otherwise the same system and noise parameters as above. The error curves corresponding to the two filters were produced by averaging over 500 independent MC runs.

VII. EXTENSION TO TV CHIRP SIGNAL

In the following, we extend our results to the case of a TV chirp.

A. TV Chirp Model

Let $\mathbf{x}_k := [r_k, \omega_k, A_k]^T$ denote the state at time k, where $r_k \in \Re$, $\omega_k \in \Re$ and $A_k \in \mathbb{C}$ denote the instantaneous frequency rate, frequency, and complex amplitude respectively. Once again, we shall assume that the state evolves according to the following simple AR(1) model:

$$\mathbf{x}_k = \mathbf{H}\mathbf{x}_{k-1} + \mathbf{v}_{k-1}$$

where $\mathbf{H}$ is 3×3 diagonal, $\mathbf{H} = \text{diag}\left([b_1, b_2, b_3]^T\right)$, with b_ℓ close to 1. The process noise sequence is i.i.d. The process noise vector at time k consists of three independent random variables with the following marginal statistics:

$$\mathbf{v}_{k-1} \sim \left[\mathcal{N}\left(0, \sigma_r^2\right), \mathcal{N}\left(0, \sigma_\omega^2\right), \mathcal{CN}(0, 2\sigma_A^2)\right]^T.$$

The measurement is related to the state via

$$y_k = \mathbf{x}_k(3) e^{j\left(\mathbf{x}_k(1)k^2 + \mathbf{x}_k(2)k\right)} + w_k$$

where w_k denotes i.i.d. $\mathcal{CN}\left(0, 2\sigma_n^2\right)$ measurement noise. Again, the problem of interest is to estimate the sequence of posterior densities, $p\left(\mathbf{x}_k|\, \{y_l\}_{l=1}^k\right)$, $k \in \{1,\ldots,T\}$ given $\{y_k\}_{k=1}^T$.

B. OIF

Let $m_A := b_3 A_{n,k-1}$, $m_\omega := b_2\omega_{n,k-1}$, $m_r := b_1 r_{n,k-1}$. For the TV chirp model, the normalizing factor $D(y_k, \mathbf{x}_{n,k-1}) := \int_{\mathbf{x}} p(y_k|\mathbf{x})p(\mathbf{x}|\mathbf{x}_{n,k-1})d\mathbf{x}$ is given by the multidimensional integral in the equation, shown at the bottom of the page.

Let $v := \angle y_k - \angle m_A - k m_\omega$. It can be shown that

$$D(y_k, \mathbf{x}_{n,k-1}) = \frac{1}{2\pi\left(\sigma_A^2+\sigma_n^2\right)} \times e^{-\left(|y_k|^2+|m_A|^2/2\left(\sigma_A^2+\sigma_n^2\right)\right)} \times \mathcal{R}$$

with the multiplicative factor $\mathcal{R}$ given by

$$\mathcal{R} = \mathbf{I}_0\left(\frac{|m_A||y_k|}{\sigma_A^2+\sigma_n^2}\right) + 2\sum_{\ell=1}^{+\infty}\mathbf{I}_\ell\left(\frac{|m_A||y_k|}{\sigma_A^2+\sigma_n^2}\right) \times e^{-\left(k^2\sigma_\omega^2+k^4\sigma_r^2/2\right)\ell^2}\cos(\ell k^2 m_r - \ell v)$$

where $\mathbf{I}_\ell(\cdot)$ denotes the modified Bessel function of the first kind of order ℓ. Again, the sum can be truncated to a relatively small number of terms (we used 100 terms in our simulations). This is mainly due to the negative exponential dependence on ℓ^2, k^2, k^4 and the decay property of the modified Bessel function with respect to the order ℓ. The OIF can now be written as

$$p\left(\mathbf{x}_k \mid \mathbf{x}_{k-1}, y_k\right) = \frac{\frac{1}{2\pi\sigma_n^2}e^{-\left(\left|y_k - A_k e^{j(r_k k^2+\omega_k k)}\right|^2/2\sigma_n^2\right)}}{\frac{1}{2\pi\left(\sigma_A^2+\sigma_n^2\right)}e^{-\left(|y_k|^2+|m_A|^2/2\left(\sigma_A^2+\sigma_n^2\right)\right)}\times\mathcal{R}} \times \frac{1}{\sqrt{2\pi}\sigma_r}e^{-\left((r_k-m_r)^2/2\sigma_r^2\right)}\frac{1}{\sqrt{2\pi}\sigma_\omega}e^{-\left((\omega_k-m_\omega)^2/2\sigma_\omega^2\right)} \times \frac{1}{2\pi\sigma_A^2}e^{-\left(|A_k-m_A|^2/2\sigma_A^2\right)}.$$

What remains to implement the plain OIF filter for the TV chirp case is to come up with a procedure to draw samples distributed according to the above closed form. We have already described the basic steps of rejection-based sampling. A similar procedure can be applied here. Let again

$$\sigma^2 := \frac{\sigma_A^2\sigma_n^2}{\sigma_A^2+\sigma_n^2}$$

and

$$\mu := \frac{\sigma_A^2|y_k| + \sigma_n^2|m_A|}{\sigma_A^2+\sigma_n^2}.$$

Using the triangle inequality, it can be shown that

$$g(\mathbf{x}_k \mid \mathbf{x}_{n,k-1}, y_k) = \frac{e^{-\left((\omega_k-m_\omega)^2/2\sigma_\omega^2\right)}e^{-\left((r_k-m_r)^2/2\sigma_r^2\right)}e^{-\left((|A_k|-\mu)^2/2\sigma^2\right)}}{(2\pi)^{5/2}\gamma(\mu,\sigma)\,\sigma_\omega\sigma_r\sigma}$$

with

$$\gamma(\mu,\sigma) := Q_o\left(-\frac{\mu}{\sigma}\right) + \frac{\sigma}{\sqrt{2\pi}}e^{-\left(\mu^2/2\sigma^2\right)}$$

$$Q_o\left(-\frac{\mu}{\sigma}\right) := \int_{r=0}^{+\infty}\frac{1}{\sqrt{2\pi}\sigma}e^{-\left((r-\mu)^2/2\sigma^2\right)}dr = \frac{1}{2}\,\mathrm{erfc}\left(-\frac{\sigma_A^2|y_k|+\sigma_n^2|m_A|}{\sigma_A\sigma_n\sqrt{2\left(\sigma_A^2+\sigma_n^2\right)}}\right).$$

For this particular dominating density, it holds $p(\mathbf{x}_k|\mathbf{x}_{n,k-1}, y_k) \leqslant c\cdot g(\mathbf{x}_k|\mathbf{x}_{n,k-1}, y_k)$ with

$$c := \frac{e^{-\left(\mu^2/2\sigma^2\right)} + \frac{\sqrt{2\pi}}{\sigma}Q_o\left(-\frac{\mu}{\sigma}\right)}{e^{-\left(|m_A||y_k|/\sigma_A^2+\sigma_n^2\right)}\mathcal{R}}.$$

Remark 4: In both cases considered (harmonic and chirp) the constant c which determines the complexity of the associated rejection step is dependent on system parameters.

For this particular choice of IF and sampling procedure, the weight update step is given by $\mathbf{w}_{n,k} = \mathbf{w}_{n,k-1}D(y_k, \mathbf{x}_{n,k-1})$ and can be carried out before the particles are propagated to time-step k.

C. Rao–Blackwellization

We can again take advantage of the model structure and partition the state vector into $[r_k, \omega_k]^T$ $(\in \Re^2)$ and A_k $(\in \mathbb{C})$. The sought posterior at time-step k, $p\left(r_k, \omega_k, A_k \mid \{y_l\}_{l=1}^k\right)$ can be factored as

$$p\left(r_k,\omega_k,A_k \mid \{y_l\}_{l=1}^k\right) = p\left(A_k \mid r_k,\omega_k,\{y_l\}_{l=1}^k\right)p\left(r_k,\omega_k \mid \{y_l\}_{l=1}^k\right).$$

$$D(y_k,\mathbf{x}_{n,k-1}) := \int_{r\in\Re}\int_{\omega\in\Re}\int_{A\in\mathbb{C}}\frac{1}{2\pi\sigma_n^2}e^{-\left(|y_k - Ae^{j(rk^2+\omega k)}|^2/2\sigma_n^2\right)} \times\left[\frac{1}{\sqrt{2\pi}\sigma_r}e^{-\left((r-m_r)^2/2\sigma_r^2\right)}\frac{1}{\sqrt{2\pi}\sigma_\omega}e^{-\left((\omega-m_\omega)^2/2\sigma_\omega^2\right)}\frac{1}{2\pi\sigma_A^2}e^{-\left(|A-m_A|^2/2\sigma_A^2\right)}\right] \times dA\,d\omega\,dr.$$

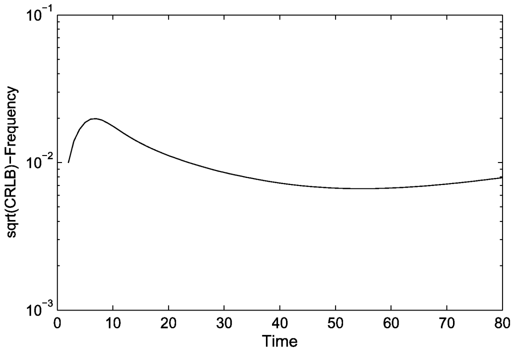

Fig. 8. $\sqrt{\text{CRLB}}$ for the frequency component: very accurate prior information and $k \leq 80$.

Again, $p\left(A_k \mid r_k, \omega_k, \{y_l\}_{l=1}^{k}\right)$ is Gaussian and can be computed using the Kalman Filter. To approximate the marginal posterior $p\left(r_k, \omega_k \mid \{y_l\}_{l=1}^{k}\right)$, we use the optimal importance density

$$p\left(r_k, \omega_k \mid r_{n,k-1}, \omega_{n,k-1}, y_k\right) = \frac{p\left(y_k \mid r_k, \omega_k\right) p\left(r_k, \omega_k \mid r_{n,k-1}, \omega_{n,k-1}\right)}{\int_r \int_\omega p\left(y_k \mid r, \omega\right) p(r, \omega \mid r_{n,k-1}, \omega_{n,k-1}) d\omega dr}$$

which again admits closed form expression. Letting $\mu_A := b_3^k E\{A_0\}$, $\sigma_A'^2 := b_3^{2k} E\left\{|A_0|^2\right\} + \left(1 - b_3^{2k}/1 - b_3^2\right)\left(2\sigma_A^2\right)$, and $\vartheta := \angle y_k - \angle \mu_A - k m_\omega$, it can be shown that

$$p\left(r_k, \omega_k \mid r_{n,k-1}, \omega_{n,k-1}, y_k\right) = \frac{\frac{1}{2\pi\left(\sigma_A'^2+\sigma_n^2\right)} e^{-\left(|y_k - \mu_A e^{j\left(r_k k^2 + \omega_k k\right)}|^2 / 2\pi\left(\sigma_A'^2+\sigma_n^2\right)\right)}}{\frac{1}{2\pi\left(\sigma_A'^2+\sigma_n^2\right)} e^{-\left(|y_k|^2 + |\mu_A|^2 / 2(\sigma_A'^2+\sigma_n^2)\right)} \mathcal{R}'} \times \frac{1}{\sqrt{2\pi}\sigma_\omega} e^{-\left((\omega_k - m_\omega)^2 / 2\sigma_\omega^2\right)} \frac{1}{\sqrt{2\pi}\sigma_r} e^{-\left((r_k - m_r)^2 / 2\sigma_r^2\right)}$$

with

$$\mathcal{R}' := \mathbf{I}_0\left(\frac{|\mu_A||y_k|}{\sigma_A'^2 + \sigma_n^2}\right) + 2\sum_{\ell=1}^{+\infty} \mathbf{I}_\ell\left(\frac{|\mu_A||y_k|}{\sigma_A'^2 + \sigma_n^2}\right) \times e^{-\left(k^2\sigma_\omega^2 + k^4\sigma_r^2/2\right)\ell^2} \cos(\ell k^2 m_r - \ell\vartheta).$$

The weight update is given by $\mathbf{w}_{n,k} = \mathbf{w}_{n,k-1} D(y_k, \omega_{n,k-1}, r_{n,k-1})$, with

$$D(y_k, \omega_{n,k-1}, r_{n,k-1}) := \frac{1}{2\pi\left(\sigma_A'^2 + \sigma_n^2\right)} \times e^{-\left(|y_k|^2 + |\mu_A|^2 / 2\left(\sigma_A'^2+\sigma_n^2\right)\right)} \mathcal{R}'.$$

We shall again employ an accept-reject algorithm to generate samples distributed according to the OIF. Using the triangle inequality and monotonicity of e^{-x}, it is easy to show that $p\left(r_k, \omega_k | r_{n,k-1}, \omega_{n,k-1}\right)$ is a suitable dominating density for which it holds that

$$p\left(r_k, \omega_k \mid r_{n,k-1}, \omega_{n,k-1}, y_k\right) \leq c p\left(r_k, \omega_k \mid r_{n,k-1}, \omega_{n,k-1}\right)$$

with $c = \left(1/e^{-\left(|\mu_A||y_k|/\sigma_A'^2 + \sigma_n^2\right)} \mathcal{R}'\right)$. Again, notice that sampling from the optimal importance function can be implemented by rejection over the transitional prior.

D. Cramér–Rao Lower Bound

In this section, we present the CRLB for the TV chirp case. Rewriting the model in real-valued form and using the results in [20], we end up with the desired recursive equation for the calculation of $\mathbf{J}_k$

$$\mathbf{J}_k = \mathbf{Q}^{-1} + \mathbf{E}\left\{\tilde{\mathbf{F}_\mathbf{k}}^T \mathbf{R}^{-1} \tilde{\mathbf{F}_\mathbf{k}}\right\} - \mathbf{Q}^{-1}\mathbf{H}' \times (\mathbf{J}_{k-1} + \mathbf{H}'^T \mathbf{Q}^{-1} \mathbf{H}')^{-1} \mathbf{H}'^T \mathbf{Q}^{-1}, \quad k \geqslant 0$$

where now $\mathbf{H}' = \operatorname{diag}\left([b_1, b_2, b_3, b_3]^T\right)$, with b_ℓ being $1 - \epsilon_\ell$, $\mathbf{Q} = \operatorname{diag}\left(\left[\sigma_r^2, \sigma_\omega^2, \sigma_A^2, \sigma_A^2\right]^T\right)$, $\mathbf{R} = \operatorname{diag}\left(\left[\sigma_n^2, \sigma_n^2\right]^T\right)$, and $\tilde{\mathbf{F}_\mathbf{k}}$ is the $\mathbf{2} \times \mathbf{4}$ matrix defined as

$$\tilde{\mathbf{F}_\mathbf{k}} = \nabla'_{\mathbf{x}_k}\left[\Re\left\{A_k e^{j\left(r_k k^2 + \omega_k k\right)}\right\} \quad \Im\left\{A_k e^{j\left(r_k k^2 + \omega_k k\right)}\right\}\right]^T$$

which is the Jacobian of the nonlinear function involved in the measurement equation, evaluated at the true value of the (real-valued) state vector $\mathbf{x}_k' := [r_k, \omega_k, \Re(A_k), \Im(A_k)]^T$.

The initial information matrix $\mathbf{J}_0$ is calculated from the initial density $\mathbf{p}(\mathbf{x}_0)$, which is assumed Gaussian $\mathcal{N}\left(\bar{\mathbf{x}}_0, \mathbf{Q}_0\right)$. In that case, the recursions may start by choosing $\mathbf{J}_0 = \mathbf{Q}_0^{-1}$.

The best achievable performance concerning the frequency and the frequency rate component of the state vector, in the case of very accurate initial information (an initial pdf with a very small variance), is presented in Figs. 8 and 9, respectively, for $k \leq 80$. The behavior of the bounds as k grows is illustrated in Figs. 10 and 11 respectively. System parameters were set to $b_\ell = 0.999$, $\forall\ \ell$, $\sigma_\omega^2 = 10^{-4}$, $\sigma_A^2 = 10^{-4}$, $\sigma_r^2 = 10^{-10}$, $\sigma_n^2 = 0.2$. The expectation appearing in the CRLB was approximated using 100 realizations of the state vector. Observe from these figures that the bounds are initially growing. This is due to the fact that the initial information was very precise, however, the effect of such an accurate prior knowledge are gradually vanishing over time. Approximately after 600 time steps the CRLB for the frequency rate component of the state vector is starting to decrease. Observe, however, that this is not happening for the frequency component, which exhibits much faster time variation than the frequency rate in this experiment, so the latter is easier to track.

Accurate prior knowledge is not always available. The best achievable error performance in the case of inaccurate prior is illustrated next. A Gaussian initial density $\mathcal{N}\left(m_{r_0}, \sigma_{r_0}^2\right)$ can be used to model the initial density concerning the frequency rate component of the state vector. In the (slowly) TV harmonic case, we have seen that inaccurate initial information only has measurable impact on the initial performance. To illustrate that this is not the case for TV chirp signals, consider a scenario where the initial information about the frequency rate component r_0 is inaccurate, whereas the initial frequency ω_0 is accurately known. For the complex amplitude A_0, we take a narrow Gaussian with mean $E[A_0] = 1 + j$ and standard deviation

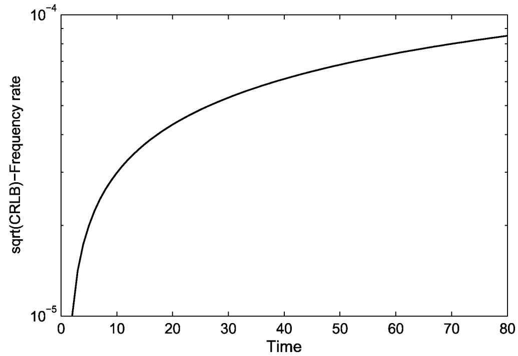

Fig. 9. $\sqrt{\mathrm{CRLB}}$ for frequency rate component: very accurate prior information and $k \leq 80$.

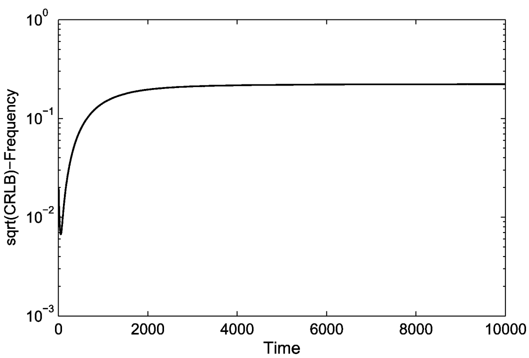

Fig. 10. $\sqrt{\mathrm{CRLB}}$ for frequency component: very accurate prior information and $k \leq 10\,000$.

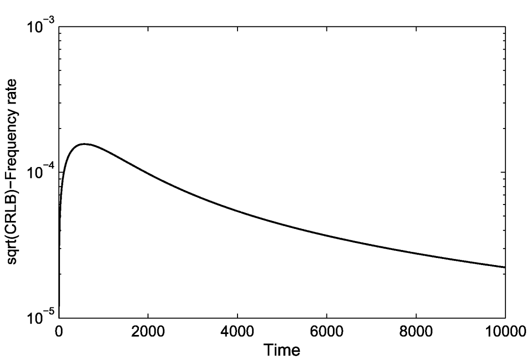

Fig. 11. $\sqrt{\mathrm{CRLB}}$ for frequency rate component: very accurate prior information and $k \leq 10\,000$.

$\mathrm{std}\,[A_0] = 0.01$. The resulting bounds on estimation performance are plotted in Figs. 12 and 13 for the frequency and frequency rate, respectively. Observe that inaccurate initial information concerning r_0 has a deleterious effect on the best achievable error performance for both frequency and frequency rate. Accurate initial information about the frequency rate is critical for acceptable tracking performance in this context.

E. Estimation Performance Results

We now present tracking results for the TV chirp case. We consider two PF algorithms: the SIR filter, which uses the transitional prior as importance distribution, and the Rao–Blackwellized filter which uses the optimal importance density (RBPF). The RMSE results concerning the frequency

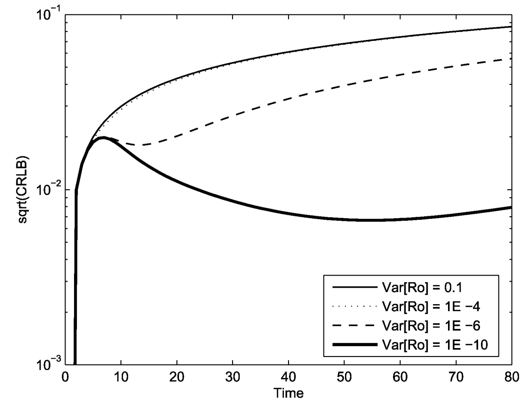

Fig. 12. $\sqrt{\mathrm{CRLB}}$ for frequency component: dependence on the accuracy of prior information.

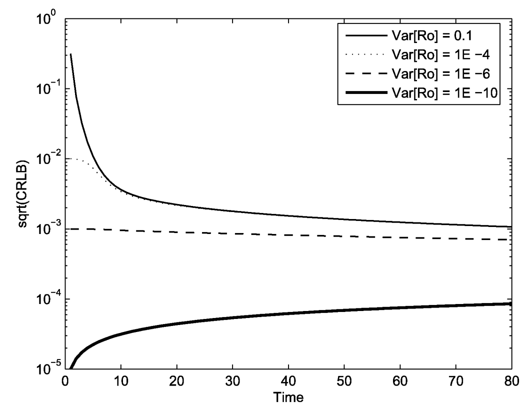

Fig. 13. $\sqrt{\mathrm{CRLB}}$ for frequency rate component: dependence on the accuracy of prior information.

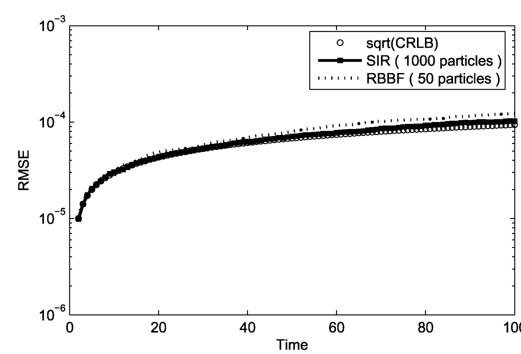

Fig. 14. RMSE performance comparison in TV second-order PPS case: SIR , RBPF, and $\sqrt{\mathrm{CRLB}}$ for the frequency rate parameter. Number of particles: 1000 for SIR, 50 for RBPF.

rate and frequency are presented in Figs. 14 and 15, respectively.

The filters are again considered in a tracking mode—we assume perfect knowledge of the initial state. The expectation appearing in the CRLB was approximated using 100 realizations of the state vector. The error curves corresponding to the two filters were produced by averaging over 500 independent runs, each comprising 100 temporal samples. The conditional mean was used to generate point state estimates. System parameters were set to $b_\ell = 0.999$, $\forall\, \ell$, $\sigma_\omega^2 = 10^{-4}$, $\sigma_A^2 = 10^{-4}$,

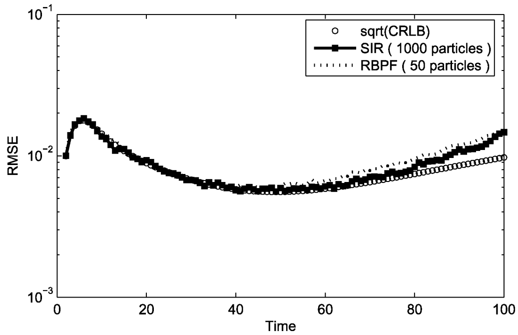

Fig. 15. RMSE performance comparison in TV second-order PPS case: SIR , RBPF, and $\sqrt{\text{CRLB}}$ for the frequency parameter. Number of particles: 1000 for SIR, 50 for RBPF.

$\sigma_r^2 = 10^{-10}$, $\sigma_n^2 = 0.1$, and multinomial resampling was employed at each time step. The number of particles, N, was 1000 for SIR and 50 for RBPF.

Notice from the simulation parameters that we have assigned a very small amount of noise in the frequency rate evolution, thus allowing (capturing) only very small variations in this term. It is however encouraging to observe that although we have used only 50 particles in RBPF's implementation, the two filters yield very similar estimation performance (SIR with 1000 particles seems slightly better), which is also very close to the CRLB. The average computation time per measurement (time-step) was 0.08419 s for SIR and 0.06498 s for RBPF (measured using Matlab tic/toc).

VIII. Conclusion

We considered the problem of tracking the parameters of a single TV harmonic or chirp signal using particle filtering tools. We showed that the importance function which minimizes the variance of the particle weights can be computed in closed form, and developed suitable rejection-based procedures to sample from the optimal importance function. We further derived efficient versions of the optimal filters based on Rao–Blackwellization. With as few as 50 particles, the optimized particle filters attain estimation performance comparable to that of generic particle filters employing 1000 particles. Using the recursive formula of Tichavsky *et al.* [20], we also computed the pertinent CRLBs and explored their behavior as a function of model parameters and the accuracy of prior information concerning the initial state.

A limitation of all tracking approaches considered is that process noise variance should be small (state evolution should be smooth) for good tracking performance. This is to be expected of course—the models considered are generically unidentifiable and one relies on smoothness to obtain meaningful estimates. Still, many potential applications (e.g., tracking of Doppler shift in mobile terrestrial communications, or residual carrier frequency offset following coarse acquisition) meet this requirement.

There are several extensions that could be pursued: a single higher order TV polynomial phase signal, or multicomponent TV harmonic or chirp signals. Both entail an expansion of the nonlinear part of the state-space and thus hit on the "curse of dimensionality." Custom design of particle filters for these cases hinges on the development of efficient state-space decomposition strategies, which is a matter of engineering art.

Appendix

Proof of Claim 1: The pdf of $x := \omega k$ is given by $f_x(x) = (1/|k|) f_\omega(x/k)$, which is an expanded version of $f_\omega(\cdot)$. Since $e^{jx} = e^{j(x \bmod 2\pi)}$, let us define $\phi := x \bmod 2\pi$. We will prove that, as $k \to \infty$, the pdf of ϕ approaches a uniform pdf over $[0, 2\pi)$.

Split the interval $[0, 2\pi)$ into N equal subintervals of length $\Delta_x = (2\pi/N)$. Take N sufficiently large for $f_x(x)$ and $f_\phi(\phi)$ to be approximately constant over each subinterval. Without loss of generality, choose arbitrary $\xi_i \in [x_{k-1}, x_k] \subseteq [0, 2\pi)$. From the definition of the modulo operation, it follows that

$$\lim_{k\to\infty} f_\phi(\xi_i)\,\Delta_x = \lim_{k\to\infty}\left(\sum_{\mu=-\infty}^{+\infty} f_x(\xi_i + 2\pi\mu)\,\Delta_x\right).$$

We have assumed that f_ω is continuous, and therefore so is f_x; it follows that

$$\begin{aligned}\lim_{k\to\infty} f_\phi(\xi_i)\,\Delta_x &= \sum_{\mu=-\infty}^{+\infty} \lim_{k\to\infty} f_x(\xi_i + 2\pi\mu)\,\Delta_x \\ &= \sum_{\mu=-\infty}^{+\infty} \lim_{k\to\infty} \frac{f_\omega\left(\frac{\xi_i}{k} + \frac{2\pi\mu}{k}\right)}{k}\Delta_x \\ &= \lim_{k\to\infty} \sum_{\mu=-\infty}^{+\infty} \frac{f_\omega\left(\frac{2\pi\mu}{k}\right)}{k}\Delta_x\end{aligned}$$

since ξ_i is bounded. ∎

Acknowledgment

The authors would like to thank the anonymous reviewers, whose comments helped improve the quality of this paper.

References

[1] C. Andrieu, M. Davy, and A. Doucet, "Improved auxiliary particle filtering: Applications to time-varying spectral analysis," presented at the 11th IEEE Workshop on Statistical Signal Processing, Singapore, Aug. 2001.

[2] C. Andrieu, "Optimal estimation of non-stationary phase and amplitude processes," in *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing (ICASSP) 2000*, Jun. 2000, vol. 2, pp. 637–640.

[3] M. S. Arulampalam, S. Maskell, N. Gordon, and T. Clapp, "A tutorial on particle filters for nonlinear/non-Gaussian Bayesian tracking," *IEEE Trans. Signal Process.*, vol. 50, no. 2, pp. 174–188, Feb. 2002.

[4] K. Bury, *Statistical Distributions in Engineering*. Cambridge, U.K.: Cambridge Univ. Press, 1999.

[5] A. T. Cemgil, H. J. Kappen, and D. Barber, "A generative model for music transcription," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 14, no. 2, pp. 679–694, Mar. 2006.

[6] L. Cohen, *Time-Frequency Analysis*. Englewood Cliffs, NJ: Prentice-Hall, 1994.

[7] L. Devroye, *Non-Uniform Random Variate Generation*. New York: Springer-Verlag, 1986.

[8] P. Djuric, J. H. Kotecha, J. Zhang, Y. Huang, T. Ghirmai, M. Bugallo, and J. Miguez, "Particle filtering," *IEEE Signal Process. Mag.*, pp. 19–38, Sep. 2003.

[9] A. Doucet and X. Wang, "Monte Carlo methods for signal processing: A review in the statistical signal processing context," *IEEE Signal Process. Mag.*, vol. 22, no. 6, pp. 152–170, Nov. 2005.

[10] A. Doucet, S. J. Godsill, and C. Andrieu, "On sequential Monte Carlo sampling methods for Bayesian filtering," *Stat. Comput.*, vol. 10, no. 3, pp. 197–208, 2000.

[11] A. Doucet, N. Gordon, and V. Krishnamurthy, "Particle filters for state estimation of jump Markov linear systems," *IEEE Trans. Signal Process.*, vol. 49, no. 3, pp. 613–624, Mar. 2001.

[12] C. Dubois, M. Davy, and J. Idier, "Tracking of time-frequency components using particle filtering," presented at the IEEE Int. Conf. Acoustics, Speech, Signal Processing (ICASSP) 2005, Philadelphia, PA, Mar. 18–23, 2005.

[13] C. Dubois and M. Davy, "Joint detection and tracking of time-varying harmonic components: A flexible Bayesian approach," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 15, no. 4, pp. 1283–1295, May 2007.

[14] N. J. Gordon, D. J. Salmond, and A. F. M. Smith, "Novel approach to nonlinear/non-Gaussian Bayesian state estimation," *Proc. Inst. Electr. Eng.—F*, vol. 140, no. 2, pp. 107–113, 1993.

[15] S. Molchanov and X. Wang, "On the Benford's empirical law," *Random Oper. Stoch. Equ.*, vol. 12, no. 3, pp. 201–210, 2004.

[16] M. Pitt and N. Shephard, "Filtering via simulation: Auxiliary particle filters," *J. Amer. Stat. Assoc.*, vol. 94, no. 446, pp. 590–599, 1999.

[17] P. Stoica and R. L. Moses, *Spectral Analysis of Signals*. Englewood Cliffs, NJ: Prentice-Hall, 2005.

[18] T. Schon, F. Gustafsson, and P.-J. Nordlund, "Marginalized particle filters for mixed linear/nonlinear state-space models," *IEEE Trans. Signal Process.*, vol. 53, no. 7, pp. 2279–2289, Jul. 2005.

[19] E. Tsakonas, N. D. Sidiropoulos, and A. Swami, "Time-frequency analysis using particle filtering: Closed-form optimal importance function and sampling procedure for a single time-varying harmonic," presented at the Nonlinear Statistical Signal Processing Workshop: Classical, Unscented, Particle Filtering Methods, Cambridge, U.K., Sep. 13–15, 2006.

[20] P. Tichavsky, C. H. Muravchik, and A. Nehorai, "Posterior Cramer–Rao bounds for discrete-time dynamical systems," *IEEE Trans. Signal Process.*, vol. 46, no. 5, pp. 1386–1396, May 1998.

[21] J. Vermaak, C. Andrieu, A. Doucet, and S. J. Godsill, "Particle methods for Bayesian modeling and enhancement of speech signals," *IEEE Trans. Speech Audio Process.*, vol. 10, no. 3, pp. 173–185, Mar. 2002.

Efthimios E. Tsakonas is working towards the Diploma degree in the Department of Electronic and Computer Engineering of the Technical University of Crete, Chania-Crete, Greece. His work in this paper was part of his undergraduate Diploma thesis.

He is currently serving his military service, and continues research in between assignments.

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

He has been a Postdoctoral Fellow (1994–1995) and Research Scientist (1996–1997) at the Institute for Systems Research, UMCP, and has held positions as Assistant Professor, Department of Electrical Engineering, University of Virginia-Charlottesville (1997–1999), and Associate Professor, Department of Electrical and Computer Engineering, University of Minnesota—Minneapolis (2000–2002). Since 2002, he has been a Professor in the Department of Electronic and Computer Engineering at the Technical University of Crete, Chania-Crete, Greece, and Adjunct Professor at the University of Minnesota. His current research interests are primarily in signal processing for communications, convex optimization, cross-layer resource allocation for wireless networks, and multiway analysis.

Prof. Sidiropoulos is currently Chair of the Signal Processing for Communications Technical Committee (SPCOM-TC) of the IEEE Signal Processing (SP) Society (2007–2008), where he has served as Member (2000–2005) and Vice-Chair (2005–2006). He is also a member of the Sensor Array and Multichannel processing Technical Committee (SAM-TC) of the IEEE SP Society (2004–2009). He has served as Associate Editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING from 2000 to 2006 and the IEEE SIGNAL PROCESSING LETTERS from 2000 to 2002. He received the U.S. NSF/CAREER award in June 1998 and the IEEE Signal Processing Society Best Paper Award twice (in 2001 and 2007). He is a Distinguished Lecturer of the IEEE SP Society for 2008–2009.

Ananthram Swami (F'08) received the B.Tech. degree from the Indian Institute of Technology (IIT)—Bombay; the M.S. degree from Rice University, Houston, TX, and the Ph.D. degree from the University of Southern California (USC), Los Angeles, all in electrical engineering.

He has held positions with Unocal Corporation, USC, CS-3 and Malgudi Systems. 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 is currently with the U.S. Army Research Laboratory, where his work is in the broad area of signal processing, wireless communications, and sensor and mobile ad hoc networks.

Dr. Swami has served as Associate Editor of the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, the IEEE SIGNAL PROCESSING LETTERS, the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS II, the *IEEE Signal Process. Mag.*, and the IEEE TRANSACTIONS ON SIGNAL PROCESSING. He was Co-Guest Editor of a of the *IEEE Signal Process. Mag.* (SPM) 2004 Special Issue on Signal Processing for Networking, the 2006 SPM Special Issue on Distributed Signal Processing in Sensor Networks, the 2006 *EURASIP Journal on Advances in Signal Processing* Special Issue on Reliable Communications over Rapidly Time-Varying Channels, the 2007 EURASIP *Journal on Wireless Communications and Networking* Special Issue on Wireless Mobile Ad Hoc Networks, and is the Lead Editor for the 2008 IEEE JOURNAL ON SELECTED TOPICS IN SIGNAL PROCESSING Special Issue on Signal Processing and Networking for Dynamic Spectrum Access. He is a Co-Editor, with Q. Zhao, Y.-W. Hong, and L. Tong, of the 2007 Wiley book *Wireless Sensor Networks: Signal Processing & Communications Perspectives*.