Acquisition in Phase Demodulation: Application to Ranging in Radar/Sonar Systems

JOSÉ M. N. LEITÃO
Instituto Sperior Técnico
Portugal

JOSÉ M. F. MOURA, Fellow, IEEE
Carnegie Mellon University

Ranging in radar/sonar systems is cast as a phase/frequency detection problem. In phase/frequency demodulation, there are two stages: acquisition, which resolves the global prior uncertainty before phase/frequency locking; and tracking, which follows the phase/frequency variations. In this work, ranging corresponds to absolute phase acquisition formulated as a global nonlinear filtering problem. The proposed solution propagates the associated multimodal density and generates a process relating global covariance with the single modes common covariance. Acquisition is then defined as the first passage of this process across a given threshold. Statistics such as acquisition time histograms and acquisition performance, characterizing the behavior of the developed estimator/detector, are obtained by Monte Carlo simulations.

Manuscript received August 16, 1992; revised March 3, 1994.

IEEE Log No. T-AES/31/2/09747.

Authors' addresses: J. M. N. Leitão, Instituto de Telecommunicações, and Departmento de Engenharia Electrotécnica e de Computadores, Instituto Superior Técnico, 1096 Lisboa Codex, Portugal; José M. F. Moura, Department of Electrical and Computer Engineering, Carnegie Mellon University, 5000 Forbes Ave., Pittsburgh, PA 15213-3890.

I. INTRODUCTION

The design of localization algorithms for sonar/radar or guidance/navigation systems is addressed here. We are concerned with determining the range and motions of a target. For simplicity, the transmitted signals are narrowband. Using complex exponential notation, let the received signal be

$$\tilde{r}(t) = [\tilde{s}(t) + \tilde{w}(t)]\exp j[2\pi f_0 t] \tag{1}$$

where the narrowband transmitted signal is

$$\tilde{s}(t) = \sqrt{2P}\tilde{g}(t)\exp\left[-j2\pi\frac{R(t)}{\lambda} + \theta\right] \tag{2}$$

with $R(t)$ representing the target-receiver separation (range), λ the wavelength, P the transmitted power, θ an additive phase, f_0 the (assumed known) transmitted frequency, $\tilde{g}(t)$ the narrowband envelope, and $\tilde{w}(t)$ a corrupting noise.

Equation (2) shows that range modulates in phase the transmitted signal. Source localization is cast as the problem of estimating the range $R(t)$, $t \geq 0$, from a record of the observations (1). In this sense, ranging is equivalent to phase demodulation. To motivate the approach we briefly discuss a simple example.

EXAMPLE. For a target following a radial uniform motion with respect to a single omnidirectional sensor (or in the far field of the array),

$$R(t) = R_0 + vt \tag{3}$$

where $R_0 = R(0)$ and v is the radial velocity. Under these circumstances, the phase is given by

$$\psi(t) = 2\pi f_0 t - 2\pi\frac{R(t)}{\lambda} + \theta$$

$$= 2\pi f_0 t - 2\pi\frac{R_0}{\lambda} - 2\pi\frac{vt}{\lambda} + \theta. \tag{4}$$

In active systems (e.g., CW radar), $\theta = 0$, or another known value. The initial range R_0 can be determined from the phase term $\psi_0 = 2\pi R_0/\lambda$, i.e., from the travel time delay. Actually, due to the periodic nature of the transmitted signal, only the residual range ($R_0 \mod \lambda$) can be found.

In passive systems, $\theta \neq 0$ and it is either unknown or random. The residual phase ψ_0 is now indistinguishable from θ. This says that, for a radially moving target, range cannot be determined from the target passive signature. Of course in both cases, the target speed v can be estimated from the Doppler shift that appears as a linear phase modulation.

The preceding example illustrates several points that we now make explicit.

1) Ranging is a global problem where the goal is to overcome global a priori uncertainty, while tracking

can be thought of as a local problem of following the small variations induced by the target motions on the received signal structure.

2) The target motions induce phase modulations on the signal temporal structure. Ranging is equivalent to a phase demodulation problem. In general, in localization, the receiver has an array of sensors. Besides these temporal modulations, the received signals also exhibit spatial wavefront curvature. This spatial information enhances the range determination capability; furthermore it provides bearing discrimination ability to the receiver. We decouple the spatial information from the temporal information and focus on ranging strictly from the phase time modulations (see [1] for an alternative integrated approach to this problem).

3) Ranging raises observability issues. In general, we need target maneuverability, or target acceleration, to be able to recover the range R_0, or $R(t)$ (target in the Fresnel or near field of the array.) On the other hand, estimating v presents in general no difficult observability questions.

4) The preceding comment clarifies why linearizing techniques, based on extended Kalman–Bucy filters (EKBFs), usually work well for target tracking. However EKBFs lack global observability for global ranging. An alternative starting algorithm is needed to achieve this (see [2], where a hybrid strategy along these lines has been discussed for this problem).

In this work, we study ranging as a nonlinear problem and develop an intrinsically nonlinear solution. We apply techniques drawn from stochastic optimal nonlinear filtering (e.g., [3, 4]) to design the ranging algorithm (for a tutorial on stochastic nonlinear filtering see [5]). To apply these tools, we need a suitable model for the target motions, which is assumed as a nominally deterministic trajectory with random perturbations. As referred to above, we decouple the temporal diversity of the problem from the spatial diversity (induced by the array baseline) and consider a pointwise single target being tracked by a single omnidirectional sensor. The simplifying assumption of not considering bearing estimation has two advantages. First, it makes the problem more manageable; performance is assessed via Monte Carlo simulations and these become computationally expensive for high order problems. Second, by decoupling the temporal from the spatial aspects, we focus on ranging issues. This analysis leaves out the coupling between bearing and range estimation but illustrates how global ranging can be accomplished.

There are several classes of phase/frequency problems. Let

$$\psi(t) = 2\pi f_0 t + x_1(t) \tag{5}$$

where $x_1(t)$ is the low pass phase component of the received signal (the transmission wavelength is normalized, i.e., $\lambda = 1$). Decompose $x_1(t)$ as

$$x_1(t) = \tilde{x}_1(t) + 2\pi L(t) \tag{6}$$

where $\tilde{x}_1(t)$ is the cyclic phase, i.e.,

$$\tilde{x}_1(t) = x_1(t) \bmod 2\pi \tag{7}$$

and $L(t)$ is a counting process giving the integer number of 2π intervals necessary to reconstruct the absolute value of the phase.

Depending on the specific goals, phase demodulation problems can be categorized as follows.

1) *Cyclic Phase Demodulation.* Only $\tilde{x}_1(t)$ is of concern. This is the problem in synchronous (or coherent) communication systems, where the initial phase errors lie in $[-\pi, +\pi]$; rapid convergence (acquisition) to a final tracking error (synchronization) is mandatory in digital communications [6].

2) *Absolute Phase Demodulation.* Of interest is $x_1(t)$. What is important is to estimate the global (or absolute) value of the phase (i.e., the distance), [2, 7]. In this context, the following issues can be distinguished:

a) *Acquisition.* The ambiguity associated with the initial phase value overlaps a multiple of 2π cells. The problem is a global one, namely to resolve the initial large indetermination by making the phase estimate converge to the actual phase process.

b) *Tracking.* Once the acquisition has been performed, a tracking period begins. The initial phase estimate is now assumed to be within a small error from the true phase value. The problem is a local one, namely to follow the phase evolution as close as possible.

Range determination corresponds to acquisition in the context of absolute phase demodulation. It requires determination of the global target/receiver separation. In terms of (6), it basically refers to the estimation of $L(t)$. Once acquisition has been achieved, motion following gives rise to a tracking stage.

The classical solution to phase and frequency demodulation is the phase-locked loop (PLL, see e.g. [8, 6]), which is a simple feedback structure performing well in a variety of situations. Its behavior is identical to that of the EKBF applied to the corresponding phase problem.

Optimal nonlinear filtering is essentially a Bayesian approach to the problem of estimating a stochastic process from its noisy observations. Accordingly, all the information about the process to be estimated is contained in its conditional probability density function (filtering density) which, for the problem under study, is in general multimodal. A central issue in implementing a nonlinear filter is to preserve and propagate the relevant shape of the filter density by retaining the essential features of the model nonlinearities; linearization, as done in the EKBF, neglects important information.

Cyclic estimation problems, studied in [11 and 12] are the classical test-bed for optimal nonlinear filtering. These studies show the superior performance of the optimal nonlinear filter over the PLL. Absolute phase tracking has been considered in [13] as a nonlinear filtering problem. In these three references, the optimal filter has been implemented by discretizing the state space and representing the probability functions by the corresponding masses at the grid points (point masses filter). This implementation, being computationally extremely expensive, is however useful as benchmark. The time-delay trackers and detectors developed in [9 and 10] also represent the conditional densities as probability masses on a grid.

The filtering density propagation algorithm presented here is a basic recursive scheme adequate to both absolute and cyclic phase estimation. Reference [14] described this algorithm and used it as an absolute phase tracker applied to real signals propagated under the Arctic ice crust; it proved to be flexible and simple enough for that practical problem, while performing much better than standard procedures. In [15 and 16], the same structure has been adapted to cyclic phase estimation attaining almost the maximal achievable performance gain over the PLL. In recent work (see [19]) we addressed the design of suboptimal phase estimators by adopting minimum Kullback distance criteria (see [17, 18]). Results of this work are used in Section IIIB to assess the quality of the adopted sensor factor representation.

Absolute phase acquisition is inherently a global optimal nonlinear filtering problem, where a local estimator (such as the EKBF) cannot even provide a suboptimal solution. The nonlinear filter (NLF) implementation, developed in Section III, exploiting the periodic nature of the observations model, represents and propagates the involved densities as sums of Gaussian functions with common covariance matrix. This allows for the definition of relative phase variance and the formulation of absolute phase acquisition as a first passage (detection) problem. To our knowledge, there is no other published work along these lines.

The text is organized as follows. Section II presents the class of models considered. Section III develops an efficient phase estimator designed according to the relevant features of the problem under study. Absolute phase acquisition is studied in Section IV. Section IVA, considering a first-order phase dynamics, illustrates the behavior of the NLF, introduces an acquisition mechanism, and studies its performance via Monte Carlo simulation. Incorporating this mechanism, the NLF becomes an estimation/detection structure. The simulations performed include a comparison of this structure against a bank of EKBFs equipped with the same acquisition detection scheme. Subsection IVB extends concepts and methods to a second-order dynamics, suggesting their applicability to higher order models. For reference purposes, the equations of the EKBF corresponding to the addressed problem are derived in the Appendix.

II. MODEL

To design the optimal estimator, suitable models for the dynamics and the observation process are required. Although somewhat restrictive, the models herein considered describe a large class of systems and problems of practical interest. As described in Section I, when the underlying application is a radar/sonar localization system, the phase process and its derivatives represent the relative source/receiver geometry and dynamics.

Consider the received signal described by (1) and (2), which correspond to a phase modulated carrier in an additive white Gaussian noise (AWGN) channel. Let $z_{1,n}$ and $z_{2,n}$ be discrete time versions of the in-phase and quadrature components of the received signal. Assuming constant amplitude (normalized to 1), these observations take the form

$$\begin{bmatrix} z_{1,n} \\ z_{2,n} \end{bmatrix} = \begin{bmatrix} \cos x_{1,n} \\ \sin x_{1,n} \end{bmatrix} + \begin{bmatrix} v_{1,n} \\ v_{2,n} \end{bmatrix} \tag{8}$$

where $\{v_{1,n}\}$ and $\{v_{2,n}\}$ are mutually independent zero mean white Gaussian sequences (WGS) with variance r. The phase process $\{x_{1,n}\}$, discrete version of process $\{x_1(t)\}$ of equation (5), is modeled as the first component of the k-dimensional vector process $\{\mathbf{x}_n\}$ described by the stochastic difference equation

$$\mathbf{x}_{n+1} = \mathbf{A}\mathbf{x}_n + \mathbf{B}\mathbf{u}_n \qquad n = 1,2,\ldots \tag{9}$$

where $\mathbf{A}$ and $\mathbf{B}$ are matrices with appropriate dimensions and $\{\mathbf{u}_n \in \mathbb{R}^m\}$ is a vector zero mean WGS with covariance matrix $\mathbf{Q}$. The initial condition $\mathbf{x}_1$ is a random vector with probability density function $p(\mathbf{x}_1)$.

The measurement noise $\{\mathbf{v}_n = [v_{1,n}\ v_{2,n}]^T\}$, the driving input noise $\{\mathbf{u}_n\}$, and the initial condition vector $\mathbf{x}_1$ are all mutually statistically independent.

In the sequel, a first-order example is used to illustrate the behavior of the NLF and introduce the acquisition mechanism. Concepts and methods are then extended to second-order dynamics.

A. First-Order Dynamics

The observations are given by (8). The phase signal is the scalar sequence $\{x_n\}$ modeled by the stochastic difference equation

$$x_{n+1} = ax_n + u_n, \qquad n = 1,2,\ldots \tag{10}$$

where $\{u_n\}$ is a zero mean WGS with variance q, and a is a drift parameter. We can think of (10) as a discretization of

$$\frac{dx(t)}{dt} = (a-1)x(t) + u(t) \tag{11}$$

where the sampling interval has been normalized to 1 ($\Delta = 1$), and the driving term $u(t)$ is white Gaussian noise with spectral density q. For $a = 1$, (10) (or (11)) models a Brownian motion; it describes a fluctuating target located at (normalized) range x_1 (equal to $x(0)$). For $a \neq 1$, it represents the (normalized) range of a target, initially at x_1, and decaying ($a < 1$) or increasing ($a > 1$) with time constant $|a-1|^{-1}$.

B. Second-Order Dynamics

The second-order model considered is representative of a class of positioning and navigation problems [1, 2, 7], like a maneuvering or accelerating target moving in the field of view of a surveying radar/sonar. The phase process is the first component of a two dimensional random sequence $\{\mathbf{x}_n = [x_{1,n} x_{2,n}]^T\}$; the second component is the frequency modulation (Doppler shift) corresponding to moving targets. The sequence $\mathbf{x}_n$ is modeled by

$$\begin{bmatrix} x_{1,n+1} \\ x_{2,n+1} \end{bmatrix} = \begin{bmatrix} 1 & b \\ a & 1 \end{bmatrix} \begin{bmatrix} x_{1,n} \\ x_{2,n} \end{bmatrix} + \begin{bmatrix} 0 \\ u_n \end{bmatrix}, \quad n = 1,2,\ldots \tag{12}$$

where $\{u_n\}$ is a zero mean WGS with variance q. The random variable $\mathbf{x}_1$, independent of $\{u_n\}$, has probability density function $p(\mathbf{x}_1)$. Equation (12) models a target initially at range $x_{1,1}$ which is moving along an essentially circular nominal trajectory with speed $x_{2,1}$. The velocity is perturbed by random accelerations u_n. Again, we can view (12) as arising from the discretization (with sampling interval $\Delta = 1$) of a two-dimensional system

$$\begin{cases} \dfrac{dx_1(t)}{dt} = bx_2(t) \\ \dfrac{dx_2(t)}{dt} = ax_1(t) + u(t) \end{cases} \tag{13}$$

where $u(t)$ is white Gaussian noise. In (13), for $b = 1$, $x_2(t)$ has the interpretation of time derivative of range $x_1(t)$. The second equation couples the range acceleration to the first component. This is similar to the range acceleration equation for a point target moving on the plane along a linear track; the motion equation is (only the range equation is shown)

$$\ddot{R}(t) = \dot{\theta}^2 R(t) \tag{14}$$

where $R(t)$ is the range and $\dot{\theta}$ is the time derivative of the bearing angle. We remind the reader that, according to Section I, we decouple the bearing and range estimation problems, thus assuming $\dot{\theta}$ to be known. This provides motivation for the second order model and an interpretation for a.

Depending on the specific application, the phase models induced by the target/receiver configuration and by the target motions may be different. The first- and second-order models considered allow for the introduction of concepts and illustrate the relevance of the techniques developed herein.

III. NONLINEAR PHASE ESTIMATOR

The nonlinear phase estimator herein presented is tailored to the specifics of the problem at hand, being efficient on two grounds: 1) high performance (recovering most of the available improvement provided by the optimal nonlinear filter), and 2) regular and simple numerical implementation. The next subsection presents the general optimal solution, from which we then derive the NLF implementation.

A. General Optimal Solution

Given the model (8) and (9), consider the problem of estimating $\mathbf{x}_n$, based on the set of present and past observations $\mathbf{Z}_n = \{\mathbf{z}_k,\ 1 \leq k \leq n\}$. Conceived as an NLF, the estimator constructs and propagates the conditional probability density function $F_n = p(\mathbf{x}_n \mid \mathbf{Z}_n)$ herein referred to as the *filtering density*. The solution consists of the recursive application of Bayes' law and Chapman-Kolmogorov equation [3]

$$\text{Prediction:} \quad P_n = S_n * F_{n-1} \tag{15}$$

$$\text{Filtering:} \quad F_n = C_n H_n \bullet P_n \tag{16}$$

where $*$ denotes convolution in the state space, $\bullet$ means pointwise multiplication of the functions defined on the state space, and C_n is a normalizing constant. For a justification of these equations see [11], or more recently [5 or 4]. The *convolution kernel*, $S_n = p(\mathbf{x}_{n+1} \mid \mathbf{x}_n)$, reflects (9) and the assumptions therein; it is Gaussian and given by

$$S_n \propto \mathcal{N}(\mathbf{x}_{n+1} - \mathbf{A}\mathbf{x}_n, \mathbf{B}\mathbf{Q}\mathbf{B}^T) \tag{17}$$

where we used the notation $\mathcal{N}(\mathbf{s}, \mathbf{V}) = \exp\{-1/2\mathbf{s}^T\mathbf{V}^{-1}\mathbf{s}\}$. This kernel acts on the preceding filter density F_{n-1} to give the *prediction density* $P_n = p(\mathbf{x}_n \mid \mathbf{Z}_{n-1})$. This function is updated by the multiplicative effect of the *sensor factor* $H_n = p(\mathbf{z}_n \mid \mathbf{x}_n)$, which, taking into account model (8), is given by

$$H_n \propto \exp\left\{ \frac{z_{1,n} \cos x_{1,n} + z_{2,n} \sin x_{1,n}}{r} \right\}. \tag{18}$$

At each iteration, the optimal estimate $\hat{\mathbf{x}}_n$ minimizes the conditional expectation of a suitable cost function $L(\mathbf{x}_n - \hat{\mathbf{x}}_n)$. For the problem at hand, we look for the minimum mean square error (MMSE) estimate, which is given by the conditional mean

$$\hat{\mathbf{x}}_n = \int \mathbf{x} F_n(\mathbf{x})\, d\mathbf{x}. \tag{19}$$

To implement (15) and (16), approximate finite representations of the operands are required. The particular representation of the sensor factor H_n, presented in the next subsection, is the key to our estimation/detection structure.

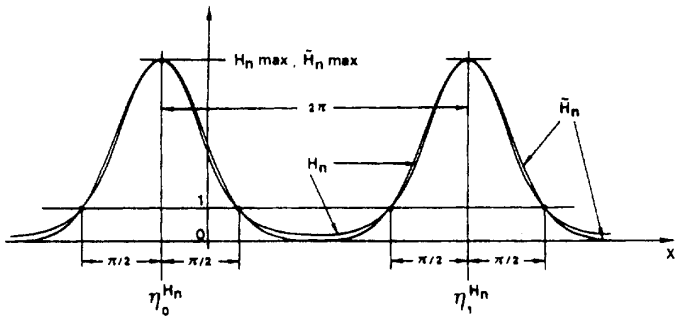

Fig. 1. Two periods of sensor factor H_n given by (20) and its representation $\tilde{H}_n$ given by (23) as train of Gaussian functions. Each basic function $\tilde{H}_n$ matches in three points each period of H_n (points signaled by dots). Resulting common variance σ^{H_n} is given by (24).

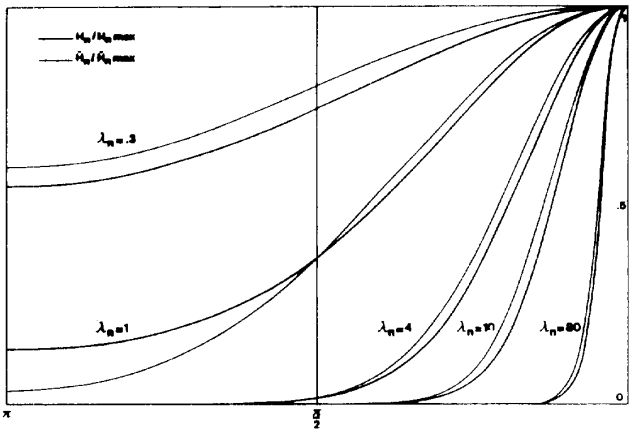

Fig. 2. Shapes of half a period of sensor factor H_n and its representation $\tilde{H}_n$ (both scaled to maximum value of 1), for different values of λ_n. The parameter λ_n is related to the observation data according to (21).

B. Sensor Factor Representation

We rewrite the exponent of the periodic sensor factor (18), so that H_n becomes

$$H_n \propto e^{\lambda_n \cos(x_{1,n} - \eta_0^{H_n})} \tag{20}$$

where

$$\lambda_n = \frac{1}{r}\sqrt{z_{1,n}^2 + z_{2,n}^2} \tag{21}$$

$$\eta_0^{H_n} = \arctan\left(\frac{z_{2,n}}{z_{1,n}}\right). \tag{22}$$

Each normalized period of H_n is a Tikhonov function, well known in PLL studies [8]. It is well approximated by a Gaussian function for large values of λ_n and becomes flat as λ_n tends to zero. This motivates the substitution of the periodic positive function H_n by a train of Gaussian functions, centered on $\eta_i^{H_n} = \eta_0^{H_n} + 2\pi i$, all having a common variance σ^{H_n}:

$$\tilde{H}_n \propto \sum_{i=-\infty}^{+\infty} \mathcal{N}(x_{1,n} - \eta_i^{H_n}, \sigma^{H_n}). \tag{23}$$

This function should reproduce, as much as possible, the shape of H_n, for all values of λ_n.

Fig. 1 shows two periods of H_n and of its representation $\tilde{H}_n$, where the variance σ^{H_n} has been obtained by fitting a Gaussian function to three particular points (signaled by oversized dots) of each of the bell shaped modes of H_n: the maximum $(\eta_i^{H_n}, H_n^{\max})$, which depends on λ_n, and the fixed coordinate points $(\eta_i^{H_n} \pm \pi/2, 1)$. The variance of the resulting $\tilde{H}_n$ is

$$\sigma^{H_n} = \frac{\pi^2}{8\lambda_n}. \tag{24}$$

In the example of Fig. 1, only one Gaussian function contributes significantly to the corresponding periods of $\tilde{H}_n$. As λ_n decreases, lateral effects become more significant, this giving rise to the flatness of $\tilde{H}_n$. Fig. 2 shows, for half a period and various values of λ_n (0.3, 1, 4, 10, and 80), the sensor factor H_n and its representation $\tilde{H}_n$, both scaled to a maximum value of 1 (i.e., plots are for $H_n/H_n^{\max}$ and $\tilde{H}_n/\tilde{H}_n^{\max}$).

An important issue is how to assess the quality of the above representation. To quantify the discrepancy between H_n and $\tilde{H}_n$, we adopt the concept of Kullback distance [17, 18]. Take $\eta_0^{H_n} = 0$, and consider the normalized central periods of H_n and $\tilde{H}_n$; denote these probability density functions by $h(x,\lambda)$ and $\tilde{h}(x,\sigma)$,

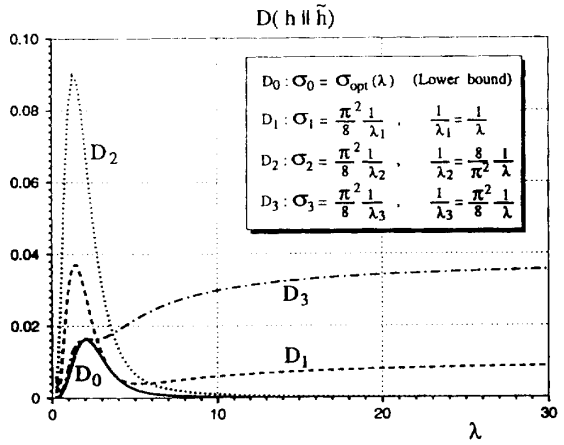

Fig. 3. Kullback distances between the normalized central period of sensor factor h (Tikhonov density [8]) and its representations $\tilde{h}$, as functions of λ. Plot D_0 is the lower bound corresponding to the optimal representation $\sigma_0 = \sigma_{\text{opt}}(\lambda)$. Plots D_1, D_2, and D_3 reflect suboptimal representation defined in inset.

respectively, where we emphasize the dependence of h on x_n and λ_n, and of $\tilde{h}$ on x_n and σ_n, after dropping the temporal index n. The Kullback distance, defined as

$$D(h(x,\lambda)\,||\,\tilde{h}(x,\sigma)) = \int_{-\pi}^{\pi} h(x,\lambda)\ln\frac{h(x,\lambda)}{\tilde{h}(x,\sigma)}dx \quad (25)$$

is a measure of dissimilarity between the probability density functions h and $\tilde{h}$ (the notation has been further simplified by dropping the arguments).

In [19], an optimal representation has been obtained by determining, for any λ, the value of σ that minimizes $D(h||\tilde{h})$ as given by (25). The result is a nonlinear integral equation, whose numerical solution provides the optimal σ as a function of λ, i.e. $\sigma_{\text{opt}}(\lambda)$. The corresponding minimum Kullback distance (lower bound), herein denoted by D_0, is plotted in Fig. 3. As $\lambda \to \infty$, the Tikhonov function h tends to a Gaussian density with variance $\sigma_0 \simeq 1/\lambda$ [8]. This happens when $r \to 0$. In this case $\sqrt{z_1^2 + z_2^2} \to 1$, and, also, from (21), $1/\lambda \simeq r$; this means that, for very small observation noise power, r close to zero, the sensor factor is a train of Gaussian functions with a common variance $\sigma = r$.

If p_1 and p_2 are normal densities with the same mean and different variances, say $p_1 = N(0,\sigma_1)$ and $p_2 = N(0,\sigma_2)$, then (see [18, p. 191])

$$D(p_1\,||\,p_2) = \frac{1}{2}\left[\ln\left(\frac{\sigma_1}{\sigma_2}\right) + \frac{\sigma_2}{\sigma_1} - 1\right]. \quad (26)$$

From Fig. 3, as $\lambda \to \infty$, the parameter $D_0 \to 0$. From (26) this means that the optimal representation $\tilde{h}_{\text{opt}}$ is a perfect replica of h. The maximum of D_0 (worst case) occurs for $\lambda \simeq 2$ where the Tikhonov function h is no longer well approximated by a Gaussian and the contribution of lateral modes are not able to give the representation $\tilde{h}$ the shape of h. For vanishing values of λ, the densities $\tilde{h}$ and h become equally flat and $\tilde{h}$ is again a good approximation to the shape of h.

When σ^{H_n} is given by equation (24), the Kullback distance is given by curve D_1 depicted in Fig. 3. In the limit, as $r \to 0$, $\tilde{h}_1$ assumes a Gaussian shape with variance $\sigma_1 = (\pi^2/8)r$. Computing $D_1(\tilde{h}_1\,||\,h_{\text{opt}})$ according to (26), one gets the horizontal asymptotic value $D_1 = 0.01$ (see Fig. 3). For $3 < \lambda < 4$, the representation is practically optimal ($D_1 \simeq D_0$). The maximum of D_1 (and maximal departure from the optimum D_0) happens for $\lambda \simeq 1.5$. Values of λ close to zero lead again to $D_1 \simeq D_0 \simeq 0$.

To analyze the effect of using expression (24) with mismatches between the assumed noise variance and the true one, we replot the Kullback distance using the values of σ_2 and σ_3 given in the inset of Fig. 3.

Taking σ_2, a variance $(8/\pi^2)r$ (corresponding to an underestimation of about 20%) is presumed. The corresponding plot D_2 coincides with D_0 for large values of λ, but departs significantly from D_0 about its maximum, i.e., for $1 < \lambda < 3$.

Adopting σ_3 is equivalent to presuming a variance $(\pi^2/8)r$, i.e., an overestimation of about 20%. The corresponding plot D_3 is practically optimal, i.e., coincides with D_0, for $\lambda < 2$; however, for large values of λ, D_3 shows a large residual error of 0.0385 (see Fig. 3). This value can be computed from (26).

Plot D_1 is a good compromise between D_2 and D_3. The residual asymptotic error is small (0.01) when compared with the much larger asymptotic value of D_3, and for small values of λ, i.e., $1 < \lambda < 3$, its departure from the optimal D_0 is much smaller than the departure of D_2. Actually, for $2 < \lambda < 4$, D_1 practically coincides with D_0. In view of the above arguments, we adopt in our simulation in part IV the value of σ^{H_n} as computed from (24).

Alternatively, we could have adopted an adaptive technique, leading to a nearly optimal representation; it consists in describing $\sigma_{\text{opt}}(\lambda)$ by three hyperbolic segments: σ_1, for $\lambda < 2$; σ_2, for $2 \le \lambda < 6$; σ_3, for $\lambda > 6$.

In summary we make two comments.

1) To be practical, good representations should preserve the relevant information contained in the represented function and involve the explicit computation of only a small number of parameters. Our description of $\tilde{H}_n$ satisfies these parsimonious requirements. Only two parameters are needed: the location of a maximum ($\eta_0^{H_n}$) of H_n given by (22) and the variance σ^{H_n} calculated according to (24). Both expressions are simple and easy to calculate.

2) As a result of representation (23), and from the fact that S_n as given by (17) is Gaussian, all the densities will be of the Gaussian sum type, provided the initial condition has the same form (see below).

C. Prediction

Assume that F_{n-1} is the sum of $N^{F_{n-1}}$ Gaussian functions, with means $\eta_i^{F_{n-1}} \in R^k$, and common

covariance matrix $\mathbf{V}^{F_{n-1}}$ (dimension $k \times k$):

$$F_{n-1} = \sum_{i=1}^{N^{F_{n-1}}} k_i^{F_{n-1}} \mathcal{N}(\mathbf{x}_{n-1} - \boldsymbol{\eta}_i^{F_{n-1}}, \mathbf{V}^{F_{n-1}}) \quad (27)$$

where the weighting factors $k_i^{F_{n-1}}$ are such that F_{n-1} is a normalized density.

According to (15) the prediction density P_n is the convolution of (27) with the convolution kernel (17). The result is

$$P_n = \sum_{i=1}^{N^{P_n}} k_i^{P_n} \mathcal{N}(\mathbf{x}_n - \boldsymbol{\eta}_i^{P_n}, \mathbf{V}^{P_n}). \quad (28)$$

Since (17) is a single Gaussian, there is no creation of new modes. The number of Gaussian modes of the filtering density, $N^{F_{n-1}}$, is conserved, i.e.,

$$N^{P_n} = N^{F_{n-1}}. \quad (29)$$

The remaining parameters in (28) are

$$\boldsymbol{\eta}_i^{P_n} = \mathbf{A}\boldsymbol{\eta}_i^{F_{n-1}} \quad (30)$$

$$\mathbf{V}^{P_n} = \mathbf{A}\mathbf{V}^{F_{n-1}}\mathbf{A}^T + \mathbf{B}\mathbf{Q}\mathbf{B}^T \quad (31)$$

$$k_i^{P_n} = Ck_i^{F_{n-1}}. \quad (32)$$

Constant C affecting all the terms of P_n can be omitted. Prediction is thus a bank of $N^{F_{n-1}}$ discrete Kalman-Bucy prediction steps. This stems from the linear and Gaussian assumptions on the dynamics (9) and the form (27) of F_{n-1}.

D. Filtering

To implement (16), function H_n is substituted by its representation (23). Since multiplying two Gauss functions yields a Gauss function, the result of multiplying the ith term of P_n by the lth mode of $\tilde{H}_n$ is

$$(P_n \cdot \tilde{H}_n)_{il} \propto k_{il}'^{F_n} \mathcal{N}(\mathbf{x}_n - \boldsymbol{\eta}_{il}^{F_n}, \mathbf{V}^{F_n}) \quad (33)$$

with

$$\boldsymbol{\eta}_{il}^{F_n} = \boldsymbol{\eta}_i^{P_n} - \frac{\eta_{1,i}^{P_n} - \eta_l^{H_n}}{V_{11}^{P_n} + \sigma^{H_n}} \begin{bmatrix} V_{11}^{P_n} \\ \vdots \\ V_{1k}^{P_n} \end{bmatrix} \quad (34)$$

$$\mathbf{V}^{F_n} = \mathbf{V}^{P_n} - \frac{1}{V_{11}^{P_n} + \sigma^{H_n}} \begin{bmatrix} V_{11}^{P_n} & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ V_{1k}^{P_n} & 0 & \cdots & 0 \end{bmatrix} \mathbf{V}^{P_n} \quad (35)$$

$$k_{il}'^{F_n} = k_i'^{P_n} \mathcal{N}(\eta_{1,i}^{P_n} - \eta_l^{H_n}, V_{11}^{P_n} + \sigma^{H_n}). \quad (36)$$

Equations (34) and (35) represent a bank of discrete Kalman-Bucy filtering (updating) steps guided by observations $\eta_l^{H_n}$ with adaptive variance σ^{H_n}. The means $\eta_l^{H_n}$ of each mode of $\tilde{H}_n$ play the role of pseudomeasurements. Very importantly from a computational point of view, note that the covariance matrix $\mathbf{V}^{F_n}$ in (35) is not subscripted by the mode indices (i,l), since all the modes of $\tilde{H}_n$ have the same variance. The bank of filters (34) has several elements, but only one Riccati equation needs to be solved.

E. Dimension Control

At the filtering step the number of terms of F_n increases. To keep the filter dimension N^{F_n} within reasonable limits the following criteria are adopted.

Truncation: Each term of P_n multiplies only the J nearest modes of $\tilde{H}_n$. This number J is called the *multiplication parameter*.

Agglutination: After multiplication, two terms of the product of $P_n \cdot \tilde{H}_n$ are agglutinated in only one Gaussian term if $(\eta_{m,i}^{F_n} - \eta_{m,j}^{F_n})^2 < \beta_m$, where $m = 1,2,\ldots,k$, and $\eta_{m,i}^{F_n}$ denotes the mth component of vector $\boldsymbol{\eta}_i^{F_n}$. The resulting mean, covariance matrix, and weighting factor are easily evaluated from the corresponding values of the combined terms. Vector $\beta \in R^k$ is called the *agglutination parameter* vector.

Elimination: After agglutination and renormalization, mode (il) of $P_n \cdot \tilde{H}_n$ is eliminated if its weight is small, i.e., if

$$k_{il}^{F_n} < \delta.$$

Constant δ is called the *elimination parameter*.

Once the three preceding simplifications have been implemented, the remaining coefficients $k_i^{F_n}$ (after relabeling) are renormalized. The new filtering density becomes

$$F_n = \sum_{i=1}^{N^{F_n}} k_i^{F_n} \mathcal{N}(\mathbf{x}_n - \boldsymbol{\eta}_i^{F_n}, \mathbf{V}^{F_n}) \quad (37)$$

which is a Gaussian sum like F_{n-1} in (27). However, the number of terms is, in general, not the same: an increase ($N^{F_N} > N^{F_{n-1}}$), or a decrease ($N^{F_N} < N^{F_{n-1}}$), may occur. Here resides the ability of F_n to concentrate or to spread out, eliminating or creating new modes adaptively, according to the data. The number of terms N^{F_n} is called in the sequel the *filter dimension*.

Fig. 4 summarizes the propagation algorithm $F_{n-1} \xrightarrow{z_n} F_n$ just described. The design parameters J, β, and δ must be adjusted to the problem under study and tuned by simulation.

F. Initialization

Acquisition problems are concerned with the lack of knowledge about initial conditions. This is captured

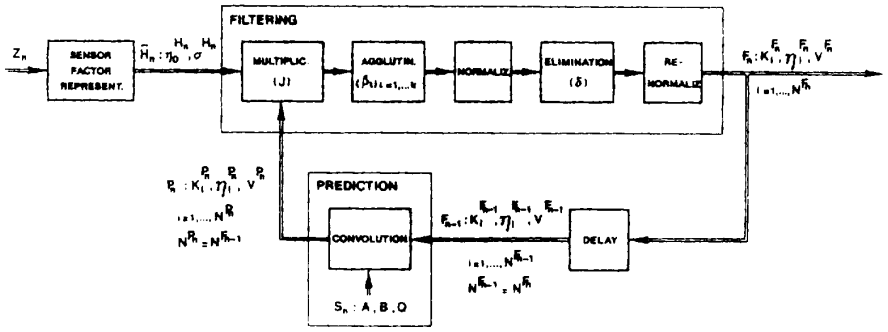

Fig. 4. Block diagram of propagation algorithm $F_{n-1} \xrightarrow{z_n} F_n$ developed in Section III. Updating of filtering density is accomplished in two steps: prediction and filtering, guided by current observations $\mathbf{z}_n$.

by modeling the initial condition as a random variable with appropriate probability density function $p(\mathbf{x}_1)$. This function works as the prediction density in the first iteration: $P_1 = p(\mathbf{x}_1)$.

G. Estimator

At each time step, the optimal estimate is the conditional mean (19) which, for F_n given by (37), takes the form

$$\hat{\mathbf{x}}_n = \frac{\sum_{i=1}^{N^{F_n}} k_i^{F_n} \eta_i^{F_n}}{\sum_{i=1}^{N^{F_n}} k_i^{F_n}}. \tag{38}$$

Note: The use of Gaussian functions in building nonlinear filters is not new (see [20 and 21].) The present work departs, however, from those approaches mainly in the following points:

1) In [20 and 21], the densities are approximated according to a minimal norm criterion (Wiener approximation theorem [20] or related criteria [21]). In [21], the nonlinear observations are linearized around the means of each term of the prediction density, leading to a bank of filtering steps, with no creation of new modes. Additional terms may be created in the prediction step if the individual variances are to be controlled in order to guarantee the nondivergence of each local estimator. The algorithm applied in [22] to cyclic phase and frequency estimation belongs to this class of implementations.

2) Our realization exploits the peculiar features of the periodic sensor factor H_n, preserving its relevant shape (information content); in this sense it is a problem-matched solution. The representation of H_n as a train of Gaussian functions with a common variance leads to a bank of filters with only one Riccati equation. Each filter is driven by pseudomeasurements (which are located at the maxima of H_n) with a common adaptive noise variance (see 24).

IV. ACQUISITION

In this section, we study, in the context of absolute phase acquisition, the behavior of the NLF developed in the previous section. The phase space for these problems is the real line $\mathbb{R}^1$ when the process is first order and the plane $\mathbb{R}^2$ for second-order dynamics. In acquisition, the goal is to overcome a large initial uncertainty regarding the phase (and possibly its derivatives.) Recalling decomposition (6), the major task is to estimate the counting process $L(t)$. This is conceptually a detection problem. As such, it is important that the estimator incorporates a mechanism deciding when acquisition has been accomplished. In radar/sonar applications, this flags the end of the acquisition step and the beginning of the tracking stage.

A. First-Order Dynamics

Consider the scalar phase process modeled by the difference stochastic equation (10), and assume that $P_1 = p(x_1)$ is uniformly distributed on $[x_{\min}, x_{\max}]$; this expresses the total lack of information about the initial phase range within the considered interval.

Since P_1 is not Gaussian, an appropriate initialization procedure is needed before applying the recursive algorithm $F_{n-1} \xrightarrow{z_n} F_n$ to this particular phase dynamics.

Initialization: From the first observation $\mathbf{z}_1$ obtain $\tilde{H}_1$ according to (23), and consider F_1 formed by the minimum number of terms of $\tilde{H}_1$ spanning the support of P_1, i.e., $N^{F_1} = \Delta x_{1,n}/(2\pi) + 2$.

The filtering density F_1 has now the form of (37) (defined on $\mathbb{R}^1$), with $k_i^{F_1} = 1/N^{F_1}$.

The optimal estimate is the conditional mean given by (38).

1) *Simulation Example*: The following example is presented with the goal of gaining insight on the behavior of the NLF.

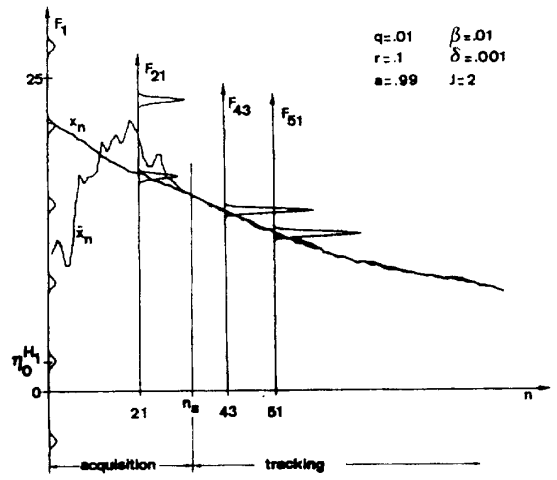

Fig. 5. Typical acquisition behavior of the proposed NLF. Starting with a large indetermination expressed by a multimodal configuration of F_1, the filtering density F_n eventually concentrates around the true phase trajectory. This is accompanied by convergence of the estimate $\hat{x}_n$ to the correct value x_n.

Fig. 5 shows the evolutions of x_n and of the phase estimate $\hat{x}_n$ provided by the NLF, and represents the shape of F_n at four different iterations ($n = 1,21,43,51$). This figure also lists the values of the filter parameters (J,β,δ) and of the model constants (a,q,r). The support of P_1 is the interval $[x_{\min} = 0, x_{\max} = 25]$.

The initializing procedure induces on F_1, besides the modes on the wrong 2π intervals, a mode on the right cell, i.e., where the process x_1 really is. Starting from F_1, the filter proceeds according to the algorithm $F_{n-1} \xrightarrow{\mathbf{z}_n} F_n$ developed in Section III. At each iteration, terms are eliminated when their weights get smaller than δ, while the others are reinforced on account of the normalization procedure. The optimal estimate $\hat{x}_n$ starts roughly from the central region of the interval $[0,25]$, evolves erratically, and finally converges to x_n; when this occurs, F_n is practically concentrated on the correct mode.

2) *Acquisition Mechanism*: As said in the Introduction, absolute phase acquisition (global ranging) raises global observability problems.

Consider the scalar dynamics $x_{n+1} = ax_n$ (limit of (10) as the driving noise variance goes to zero, $q \to 0$) with observations $\mathbf{z}_n = [\cos x_n, \sin x_n]^T$ (limit of (8) when the observation noise variance goes to zero, $r \to 0$). According to the initialization procedure, F_1 is a train of equally weighted Dirac functions (limit of the Gaussian terms of $\tilde{H}_n$ spanning P_1) expressing the exact location of x_1 on each of the possible 2π intervals. If $a \neq 1$, application of the recursive algorithm $F_{n-1} \xrightarrow{\mathbf{z}_n} F_n$ resolves the indetermination in just one step: prediction imposes to each localizing mode of F_1 a different shift; filtering determines the (only) mode of P_2 consistent (coincident) with $\tilde{H}_2$. If $a = 1$, the indetermination will persist ($F_n = F_1$, $n = 1,2,\ldots$). This means that the system is globally observable only if $a \neq 1$.

Consider now the random dynamics (10) with noisy observations (8). If $a = 0$ (and thus the process $\{x_n\}$ is a Brownian motion), the filtering density F_n spreads out by increasing the number of modes N^{F_n}, as newcoming observations $\mathbf{z}_n$ are processed. The initial indetermination grows in this case. Whenever $a \neq 0$, as exemplified in Fig. 5 (where $a = 0.99$), the trend is to concentrate F_n, reducing the number of its modes. The initial multimodal configuration, characterized by a large global variance, becomes practically unimodal and the estimate $\hat{x}_n$ gets close to the true phase value x_n. Once the initial indetermination has been resolved, acquisition is considered accomplished and a tracking stage begins.

To detect the transition from the transient evolution period to the tracking stage, the NLF incorporates an acquisition mechanism based on the relative variance concept introduced below. Let $\sigma_g^{F_n}$ be the global variance of F_n, i.e., $\sigma_g^{F_n} = \int (x_n - \hat{x}_n)^2 F_n \, dx_n$. With $\hat{x}_n$ and F_n given by (38) and (37), respectively, we have

$$\sigma_g^{F_n} = \sigma^{F_n} + \sum_{i=1}^{N^{F_n}} k_i^{F_n} (\eta_i^{F_n})^2 - \left(\sum_{i=1}^{N^{F_n}} k_i^{F_n} \eta_i^{F_n} \right)^2 \tag{39}$$

(σ^{F_n} is, as mentioned before, the common variance of each term of F_n).

The relative variance of F_n is a ratio measuring the concentration of F_n defined as

$$\alpha_n = \frac{\sigma_g^{F_n}}{\sigma^{F_n}} = 1 + \frac{\sum_{i=1}^{N^{F_n}} k_i^{F_n} (\eta_i^{F_n})^2 - \left(\sum_{i=1}^{N^{F_n}} k_i^{F_n} \eta_i^{F_n} \right)^2}{\sigma^{F_n}}. \tag{40}$$

When F_n becomes unimodal, α_n reduces to unity, no matter what the value of σ^{F_n} is; otherwise $\alpha_n > 1$.

Acquisition Time: We now consider that absolute phase acquisition is accomplished when the measure α_n, given by (40), first crosses a given threshold α_a. Acquisition time is a random variable T_a, formalized as the corresponding first passage time. We call the process $\{\alpha_n\}$ and the threshold α_a the *acquisition process* and the *acquisition parameter*, respectively. **Note:** The proposed method is a built-in decision mechanism based on global information available to the filter, i.e., the global shape of its conditional probability density function. Local estimators, e.g., the EKBF (see the Appendix), cannot use an internal criterion like the one just described simply because they lack the information on which to decide.

Acquisition Performance: When a sample function of $\{\alpha_n\}$ crosses the threshold α_a for the first time, two situations can occur:

$$\epsilon_n^2 = (x_n - \hat{x}_n)^2 \begin{cases} < \pi^2 & \text{correct acquisition} \\ \geq \pi^2 & \text{false acquisition} \end{cases} \tag{41}$$

i.e., the errors are either smaller than π, in which case we say that correct acquisition occurred, or larger than π, which leads to a false acquisition. The acquisition performance, denoted by ζ, is defined as the probability of correct acquisition.

The random variable T_a, characterized by the probability density function $p(T_a)$, and the probability of correct acquisition ζ, depend on the filter parameters and on the acquisition threshold α_a. Given the complexity of the problem, it is not possible to get formal expressions for $p(T_a)$ and ζ. Instead, Monte Carlo simulations are performed and the numerical results presented in graphical and table form.

3) *Characterizing the Acquisition Behavior*: To characterize the behavior of the proposed estimator/detector, the following procedure has been adopted: for a given noise condition and a set of filter and model parameters, M runs (of length N) are performed and a number of meaningful statistics are computed.

As usual in phase estimation studies [11–13], the noise condition is expressed in terms of the steady state (tracking) error variance that would be achieved if the observations were linear. This value, denoted by V_∞, is provided by the associated Ricatti equation (see the Appendix). In the PLL literature [8, 6], V_∞ is the inverse of the (carrier) signal-to-noise ratio in the (linearized) loop bandwidth.

Besides V_∞, the simulation program accepts as input: 1) the model parameters a, q, r, and the initial phase indetermination $[x_{\min}, x_{\max}]$ and 2) the filter parameters J, β, δ, and the acquisition parameter α_a.

The computed statistical data (i.e., the simulation results) are as follows.

a) The acquisition histogram

$$\text{Hist}(n) = \frac{\text{HIST}(n)}{\mu_a}, \qquad n = 1,2,\ldots,N \tag{42}$$

where HIST(n) counts the number of acquisitions that occured at iteration n, and μ_a is the total number of acquisitions in the M Monte Carlo runs. The histogram Hist(n) is thus an estimate of $p(T_a)$. The mean acquisition time $\overline{T}_a$ (computed from Hist(n)) is also provided. The running sum of Hist(n), i.e.,

$$P_r(n) = \sum_{i=1}^{n} \text{Hist}(i) \tag{43}$$

gives an estimate of the probability distribution of T_a. From $P_r(n)$, the program obtains the time T_1 at which $P_r(n = T_1) = 0.95$.

b) The acquisition performance, estimated by the relative frequency

$$\zeta = \frac{\mu_{ca}}{\mu_a} \tag{44}$$

where μ_{ca} is the number of correct acquisitions in the M Monte Carlo runs.

c) The percentage of false acquisitions as a function of n

$$\%f_a(n) = \frac{\text{Hist}_f(n)}{\text{Hist}(n)} \times 100 \tag{45}$$

where $\text{Hist}_f(n)$ is the histogram of false acquisitions.

d) The mean square error at the acquisition instants

$$\overline{\varepsilon_a^2} = \frac{\varepsilon_a^2}{M} \tag{46}$$

where ε_a^2 has been defined in the previous section.

e) The average acquisition process

$$\overline{\alpha}_n = \frac{1}{M}\sum_{i=1}^{M} \alpha_n^{(i)}, \qquad i = 1,2,\ldots,N. \tag{47}$$

The crossing point of $\overline{\alpha}_n$ with the threshold α_a, denoted T_2, can also be used as a figure of merit.

f) The average filter dimension,

$$\overline{N}^{F_n} = \frac{1}{M}\sum_{i=1}^{M} N_{(i)}^{F_n} \tag{48}$$

which is a measure of the computational effort (complexity).

4) *Adjustment of Filter Parameters*: As said in Section IIIE, the filter parameters J, β, and δ, must be adjusted to the problem under study. Here, we adapt them to absolute phase acquisition, which also involves the setting of the acquisition parameter α_a.

Multiplication Parameter J: Returning to Fig. 3, one may ask how many terms are taken in (23) to compute (25). There is no advantage in taking more than two Gaussian functions to describe $\tilde{h}$; increasing this number does not change significantly the Kullback distances $D(h \| \tilde{h})$. From this fact, and from the definition of the multiplication parameter J, we conclude that $J = 2$ is a good choice. We confirmed this by simulation; even for strong noise conditions, $J > 2$ does not induce any noticeable improvements on ζ and $\overline{T}_a$.

Agglutination Parameter β: Ultimately, absolute phase acquisition is a decision based on modes lying about 2π apart. For this reason, the criterion for choosing the value of the agglutination parameter β takes 2π as a reference; namely, we adopt the constraint $\sqrt{\beta} < 0.1 \times (2\pi)$. The value $\beta = 0.25$, verifying this constraint and being sufficiently large to avoid a high filter dimension N^{F_n}, has experimentally proved to be a reasonable option.

Elimination Parameter δ: Fig. 6 illustrates the influence of the elimination parameter δ on the filter behavior. Except for the value of δ, all the conditions, including the parameter values and the noise sequences, are the same in (a) and (b).

1) In Fig. 6(a), at iteration $n = 21$, the filter is propagating only two modes, one of which is the correct one. In subsequent iterations, the mode

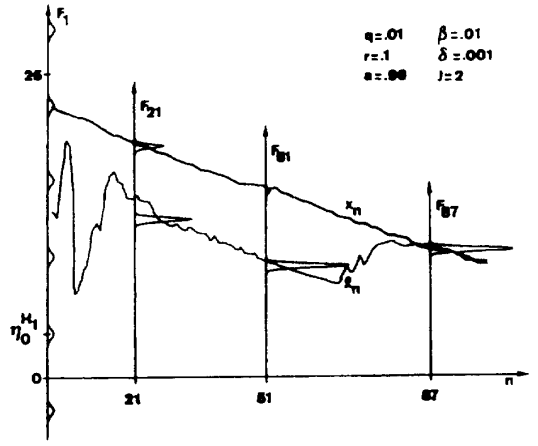

Fig. 6. Influence of elimination parameter δ on behavior of NLF. (a) With $\delta = 0.001$, estimator able to converge to right solution. (b) With $\delta = 0.01$, correct mode around x_n is prematurely eliminated and $\hat{x}_n$ evolves wrongly along an almost parallel trajectory.

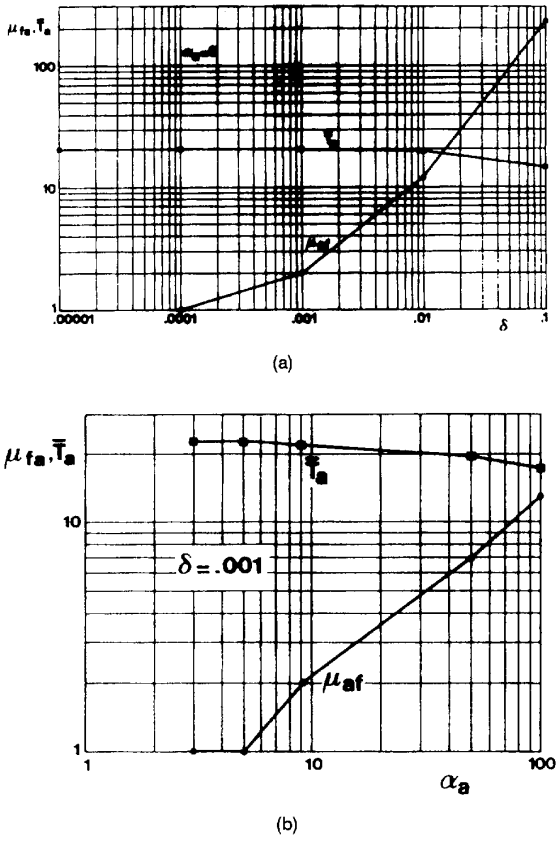

Fig. 7. Experimental results concerning mean acquisition time $\overline{T}_a$ and number of false acquisitions μ_{fa} as functions of: (a) elimination parameter δ (with acquisition parameter $\alpha_a = 9$); (b) acquisition parameter α_a (with $\delta = 0.001$).

corresponding to the false location gets to be much larger during a certain period, as illustrated at iteration $n = 51$. However, the correct mode has not been eliminated, the filter being able to recover and eventually concentrate significant probability mass in this mode (see F_{78}). Before converging to x_n, the estimate $\hat{x}_n$ evolved along a trajectory parallel to that of x_n while the incorrect mode dominated. Finally, the situation was inverted.

2) With a ten times larger value for the elimination parameter δ, the filter is not able to converge to the correct mode, as shown in Fig. 6(b). Here, up to iteration $n = 55$, the evolutions of $\hat{x}_n$ and F_n are similar to the corresponding ones in Fig. 6(a). At this iteration, the excessively large value of δ causes a premature elimination of the correct mode, the filter density reduces to the incorrect mode and the estimate $\hat{x}_n$ is forced to follow a trajectory nearly parallel to $\hat{x}_n$.

Fig. 7 reports simulation results obtained by running 1000 sample functions of length 100. Fig. 7(a) (where $\alpha_a = 9$) shows the influence of the elimination parameter δ on $\overline{T}_a$ and on μ_{fa}. For $\delta < 0.01$, there is no variation in $\overline{T}_a$. The value $\delta = 0.01$ leads to an unacceptably large number of false acquisitions, while $\delta = 0.0001$ induces a high acquisition performance, at the cost of a much larger filter dimension (computational effort). Consequently, the intermediate value $\delta = 0.001$ is adopted as a wise compromise.

Acquisition Parameter α_a: Designing the optimal threshold of a detector is an important issue. In the problem under study, the value of the acquisition parameter α_a is a tradeoff between acquisition performance ζ and acquisition time $\overline{T}_a$.

In Fig. 7(b), the number of false acquisitions μ_{fa} and the mean acquisition time $\overline{T}_a$ are plotted for a set of values of α_a. Increasing the acquisition parameter ($\alpha_a = 50,100$), tends to reduce $\overline{T}_a$ at the expense of

an exceedingly large number of false acquisitions. Our experiments showed that $\alpha_a = 9$ is a reasonable choice.

5) *Evaluation Results*: The usual benchmark in phase tracking experiments is the EKBF. In absolute acquisition, however, this filter alone would follow the corresponding initializing mode, the initial global error persisting as an offset. For comparative studies, an alternative to the NLF under test has been developed by associating an EKBF to each Gaussian mode of F_1. In this way, N^{F_1} identical estimators are run in parallel, guided by the same noise sequence.

The bank propagates the filtering density representation

$$F_n' = \sum_{i=1}^{N^{F_1}} k_i^{F_n'} \mathcal{N}(x_{1,n} - \eta_i^{F_n'}, \sigma^{F_n'}) \tag{49}$$

where $k_i^{F_n'}$, $\eta_i^{F_n'}$, and $\sigma^{F_n'}$ are computed according to the expressions in (59), (60), and (61), respectively. Coupling of elementary modes, and thus the shaping of the multimodal density F_n', is imposed by the normalizing condition $\int F_n'(x)\,dx = 1$.

Starting with N^{F_1} modes, the NLF evolves adaptively by expanding or contracting its dimension N^{F_n} according to the data; the bank of EKBFs (BEKBF) has not this ability, being of constant dimension ($N^{F_n'} = N^{F_1}$).

Given the form of F_n', the measure of relative variance, previously developed for F_n, can be directly incorporated in the BEKBF. The NLF and the BEKBF are thus equipped with the same acquisition detection mechanism.

To evaluate and compare the NLF and the BEKBF, concerning their acquisition ability, a set of simulations have been performed, according to the procedure outlined in IVA3.

Fig. 8 shows, for both filters and for $V_\infty = -20$, -15, -10, and -5 dB, the acquisition histograms Hist(n), the acquisition accumulated probabilities $Pr(n)$, and the percentage of false acquisitions $\%f_a(n)$. Also marked on abscissae are the mean acquisition times $\overline{T}_a$, and the times T_1 such as $Pr(T_1) = 95\%$. The crossing points of $\overline{\alpha}_n$ with α_a, denoted by T_2, can only be seen on the NLF plots. The following parameters have been kept constant: $[x_{1\min} = 0, x_{2\max} = 25]$; $a = -0.99$; $\alpha_a(\text{NLF}) = \alpha_a(\text{BEKBF}) = 9$; $M = 10000$.

The histogram plots provide significant information. We enphasize the two following points.

1) For each value of V_∞ the histograms for NLF and for BEKBF are different . These differences are specially significant for the last two noise conditions ($V_\infty = -10$, -5 dB); namely, the NLF histograms are sharper.

2) From all the plots of the percentage of false acquisition $\%f_a(n)$ we can also conclude that the early acquisitions are false with high probability. They correspond to premature decisions which occur more frequently with the BEKBF.

Fig. 9(a) shows, for both estimators, the dependence of ζ on V_∞. For a given acquisition parameter α_a, the acquisition performance of the NLF is consistently higher than the one of the BEKBF. The gain $\Delta\zeta = \zeta(\text{NLF}) - \zeta(\text{BRKBF})$, negligible for $V_\infty = -20$ dB ($\Delta\zeta = 0.07\%$ with $\alpha_a = 9$), becomes significant for $V_\infty = -5$ dB ($\Delta\zeta = 8.45\%$ with $\alpha_a = 9$). A smaller acquisition threshold, for example $\alpha_a = 5$, leads to better percentages of correct decisions at the cost of extending the acquisition times.

For low level noise ($V_\infty = -20$, -15, -10 dB) the algorithms are equally fast in solving the initial ambiguity (see Fig. 9(b)). For higher noise power ($V_\infty = -5$ dB) the NLF, taking advantage of its adaptive features by increasing or decreasing its dimension according to the data, takes longer to reach a decision than the BEKBF avoiding more often wrong early detections. This justifies the above mentioned better performance as measured by ζ and $\overline{\varepsilon_a^2}$.

Under the experimental conditions considered, the following approximate relation has been observed: $T_1(V_\infty) \simeq T_2(V_\infty) \simeq k\overline{T}_a(V_\infty)$, where k is a constant. Therefore, estimating $\overline{T}_a$ provides additional information about $p(T_a)$.

Fig. 9(c) shows that, for $V_\infty = -5$ dB, the mean square residual error $\overline{\varepsilon_a^2}$, after correct acquisition, is smaller for the NLF, than with the BEKBF. This emphasizes the relative superiority of the first estimator in absolute phase tracking (when acquisition is detected, the estimator enters the tracking stage).

Note: In all the situations discussed, $a = 0.99$ (i.e., the system (10) is stable). Similar results and conclusions can be obtained with other values of a, including $a > 1$ for which system (10) is unstable. For lack of space these results are not reported here.

B. Second-Order Dynamics

In this subsection, we consider the second order phase dynamics modeled by the stochastic difference equation (12), with $b = 0$. The initial indetermination about phase (range) and phase rate (velocity) has probability density function $P_1 = p(\mathbf{x}_1)$, assumed uniformly distributed over the rectangle of $\mathbb{R}^2$, $([x_{1,1\min}, x_{1,1\max}] \times [x_{2,1\min}, x_{2,1\max}])$. The initial range indetermination encompasses a multiple of wavelengths: $\Delta x_{1,1} = x_{1,1\max} - x_{1,1\min} = 2\pi L + \tilde{x}_{1,1}$, $L \in \mathbb{N}$, $\tilde{x}_{1,1} \in [-\pi, +\pi]$.

Like in Section IVA, to apply the propagation algorithm $F_{n-1} \xrightarrow{\mathbf{z}_n} F_n$ an appropriate initialization is needed.

Initialization: With a second order dynamics two observations, $\mathbf{z}_1$ and $\mathbf{z}_2$, are necessary to construct the filter density F_2 as a sum of bidimensional Gaussian functions. The procedure is as follows.

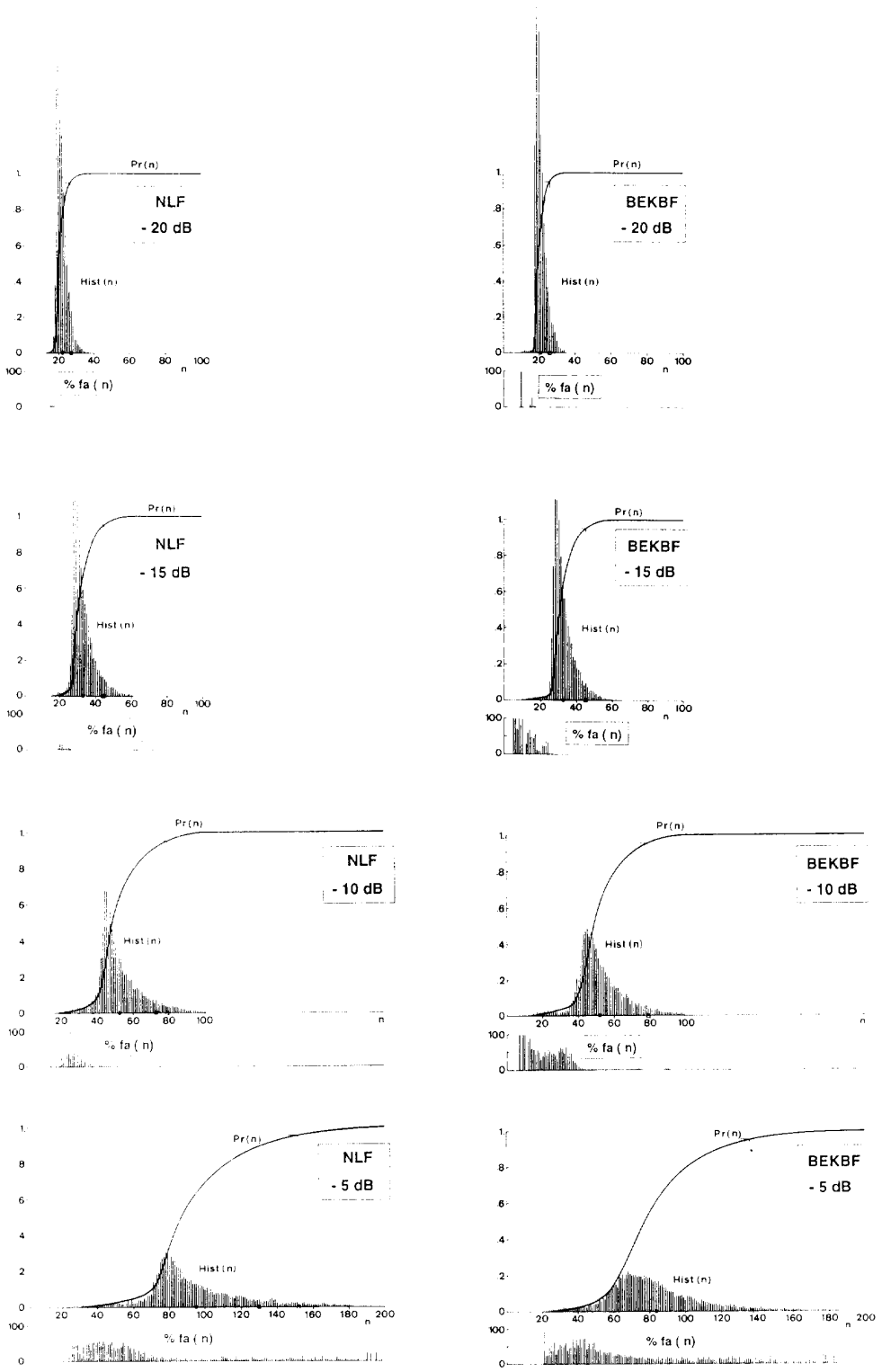

Fig. 8. Acquisition histograms Hist(n), accumulated probabilities $P_r(n)$ and percentage of false acquisitions $\%f_a(n)$ characterizing NLF and the BEKBF; noise conditions are expressed in terms of V_∞ (see Section IVA3).

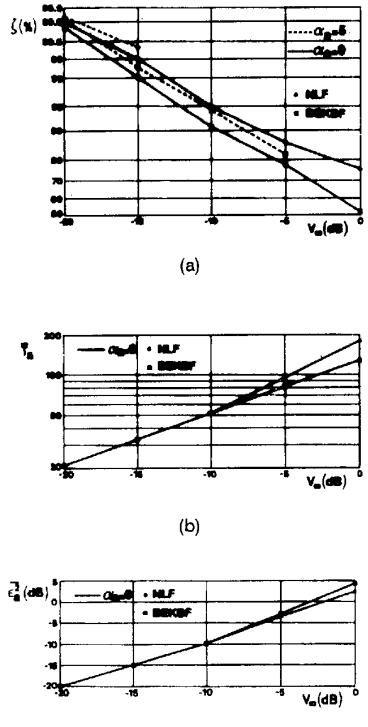

Fig. 9. Simulation results concerning NLF and BEKBF. (a) Acquisition performances ζ versus V_∞, for $\alpha_a = 9$ and $\alpha_a = 5$. (b) Mean acquisition times $\overline{T_a}$ versus V_∞. (c) Mean square error at acquisition instants $\overline{\varepsilon_n^2}$ versus V_∞.

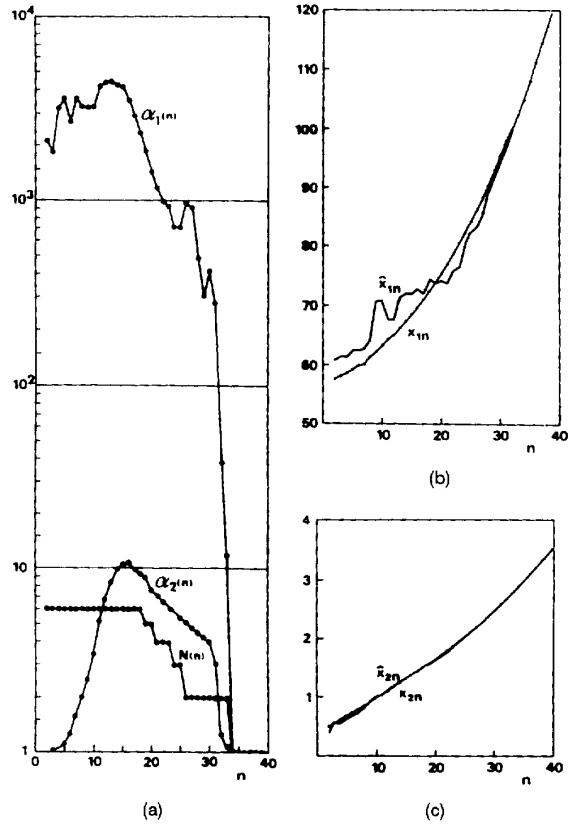

Fig. 10. Simulation example illustrating acquisition behavior for second-order dynamics (12). Plots represent simultaneous evolutions of: (a) acquisition processes $\alpha_1(n)$ and $\alpha_2(n)$ given by (51), and the filter dimensions N^{F_n}; (b) phase $x_{1,n}$ and its estimate $\hat{x}_{1,n}$; (c) phase rate $x_{2,n}$ and its estimate $\hat{x}_{2,n}$.

1) From $\mathbf{z}_1$ obtain $\tilde{H}_1$ and consider F_1 formed by the minimum number of terms of $\tilde{H}_1$ spanning the prior phase range ($N^{F_1} = \Delta x_{1,1}/2\pi + 2$).

2) Predict P_2 according to model (12).

3) From $\mathbf{z}_2$ obtain $\tilde{H}_2$ and consider F_2 formed by the minimum number of terms of $P_2\tilde{H}_2$ whose means subtend the domain where the process $\mathbf{x}_n$ is expected to be found at iteration $n = 2$, i.e., usually the support of P_2.

1) *Acquisition Detection*: The acquisition mechanism of Subection IVA is now extended to the 2-dimensional model (12). The global covariance matrix of F_n (as given by (37) in Section IIIE) is

$$\mathbf{V}_g^{F_n} = \mathbf{V}^{F_n} + \sum_{i=1}^{N^{F_n}} K_i^{F_n} \boldsymbol{\eta}_i^{F_n} \left(\boldsymbol{\eta}_i^{F_n}\right)^T - \boldsymbol{\eta}^{F_n}(\boldsymbol{\eta})^T \quad (50)$$

where $\boldsymbol{\eta}^F = \hat{\mathbf{x}}_n$ is computed according to (38).

The processes $\{\alpha_{1,n}\}_{n\geq 2}$ and $\{\alpha_{2,n}\}_{n\geq 2}$ defined by

$$\alpha_{1,n} = \frac{V_{11g}^{F_n}}{V_{11}^{F_n}}; \qquad \alpha_{2,n} = \frac{V_{22g}^{F_n}}{V_{22}^{F_n}} \quad (51)$$

are measures of the relative dispersion of F_n along the orthogonal components $x_{1,n}$ and $x_{2,n}$, respectively. The decision criterion consists in detecting acquisition when the double condition

$$\begin{cases} \alpha_{1,n} < \alpha_{1,a} \\ \alpha_{2,n} < \alpha_{2,a} \end{cases} \quad (52)$$

is met for the first time. Thresholds $\alpha_{1,a}$ and $\alpha_{2,a}$ are the acquisition parameters; the process $\{\boldsymbol{\alpha}_n = [\alpha_{1,n}; \alpha_{2,n}]^T\}_{n\geq 2}$ is called the acquisition process.

The acquisition time T_a is thus the random variable defined as the first passage time of the vector process $\{\boldsymbol{\alpha}_n\}_{n\geq 2}$ across the double barrier $\boldsymbol{\alpha}_a = [\alpha_{1,a}; \alpha_{2,a}]^T$.

Generalization of the above concepts and criteria to the kth-order dynamics (9) is straightforward.

2) *Simulation Example*: Fig. 10 represents typical evolutions (single run) of: phase and phase rate processes, $\{x_{1,n}\}$ and $\{x_{2,n}\}$, and their respective estimates $\{\hat{x}_{1,n}\}$ and $\{\hat{x}_{2,n}\}$; the two components of the acquisition process $\{\alpha_{1,n}\}$ and $\{\alpha_{2,n}\}$; and the filter dimension N^{F_n}. Simulation parameters are: initial conditions $[x_{11\min} = 70; x_{11\max} = 75]$ and $[x_{21\min} = -1; x_{21\max} = 1]$; noise condition $V_{11\infty} = -20$ dB (see

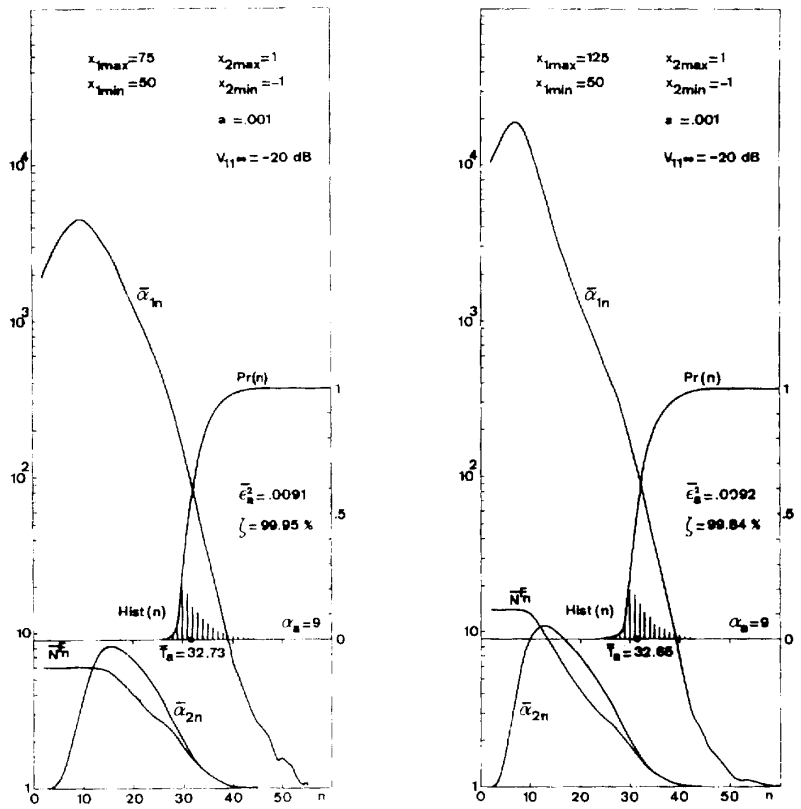

Fig. 11. Simulation outputs showing effects of changing initial indetermination. Simulation parameters of (a) and (b) are the same, with one exception: (a) $\Delta x_{1,1} = 75 - 50 = 25$; (b) $\Delta x_{1,1} = 125 - 50 = 75$. In case (b), although starting with a larger number of modes, the estimator/detector exhibits practically the same acquisition features of case (a).

IVB3); filter and model parameters $J = 2$, $\beta_1 = 0.025$, $\beta_2 = 0.5$, $\delta = 0.001$, $\Delta = 1$, and $a = 0.001$.

Notice that, for $\alpha_{1,n} < 10$, phase acquisition can usually be taken for granted. Since convergence of $\hat{x}_{2,n}$ to $x_{2,n}$ has also been accomplished, with $\alpha_{2,n}$ taking small values. This is the pattern generally observed: phase and phase rate acquisitions are strongly correlated. This is intuitively pleasing. The process $x_{2,n}$ is the discrete version of the derivative of the phase process. It is thus natural that when the phase $x_{1,n}$ has been acquired, acquisition of $x_{2,n}$ has also occured. In general the converse is not true, i.e., acquisition on the derivative $x_{2,n}$ does not imply acquisition on its integral $x_{1,n}$.

Thus, instead of testing all the components of α_n, as suggested in B1, we can take as acquisition process simply its first component $\{\alpha_{1,n}\}$. Accordingly, there is acquisition if

$$\alpha_{1,n} < \alpha_{1,a} \tag{53}$$

for the first time. This decision is considered correct if

$$\varepsilon_n^2 = (x_{1,n} - \hat{x}_{1,n})^2 < \pi^2 \tag{54}$$

otherwise it is labeled as a false acquisition.

The set of simulations reported in the next section is carried out according to the above simplified criterion.

3) *Evaluation Results*: Since the acquisition mechanism is structurally the same for vector and scalar models, the experimental procedure of IVA3 is again used. Where necessary, scalar quantities are substituted by the equivalent vectors.

In the set of simulations presented below, the filter parameters ($J = 2$, $\beta_1 = 0.25$, $\beta_2 = 0.5$, $\delta = 0.001$) and the initial phase rate indetermination $[x_{2,1\min} = -1, x_{2,1\max} = +1]$ are not changed. Noise condition is now expressed in terms of the first diagonal element of the asymptotic steady state covariance matrix, $V_{11\infty}$ (see the Appendix).

By changing input parameters $(a, V_{11\infty}, \Delta x_{1,1}, \alpha_a \equiv \alpha_{1,a})$ and comparing results, the following conclusions are drawn.

1) The acquisition histogram Hist(n), and the detector performance ζ do not depend significantly on the initial phase indetermination $\Delta x_{1,1}$. This is seen by comparing Fig. 11(a) (where $\Delta x_{1,1} = 25$) and Fig. 11(b) (where $\Delta x_{1,1} = 75$).

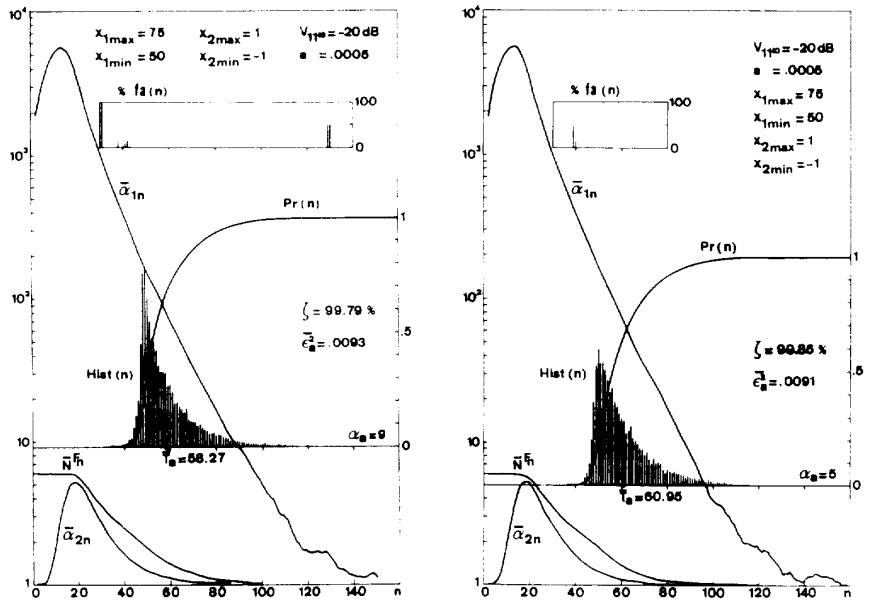

Fig. 12. Simulation outputs showing the effects of changing acquisition parameter α_a. Simulation parameters of (a) and (b) are the same, with one exception: (a) $\alpha_a = 9$; (b) $\alpha_a = 5$. Decreasing α_a yields a larger mean acquisition time $\overline{T}_a$ and a better acquisition performance ζ.

2) Decreasing the acquisition threshold α_a induces a shift of Hist(n) to the right and a better detector performance ζ. These effects, and a shape modification in Hist(n), are apparent from Fig. 12(a) (where $\alpha_a = 9$) and Fig. 12(b) (where $\alpha_a = 5$). They would be much more pronounced under the conditions of Figs. 12 and 13.

3) Smaller values of a, which in the ranging problem correspond to motions closer to being radial, produce longer acquisition times and worse detection performance. This is seen from Fig. 12(a) (where $a = 0.0005$) as compared with Fig. 11(a) (where $a = 0.001$). As for the remaining simulation parameters, both figures are equivalent.

4) The same happens if a is kept constant and the noise condition degrades ($V_{11\infty}$ increases). In Fig. 10, the acquisition histogram grows slowly between $n = 10$ and $n = 60$. During this period most of the decisions have been wrong, as is clear from the plot of the percentage of false acquisitions $\%f_a(n)$. This behavior, observed also under several other conditions, confirms (now with a second order dynamics) that early acquisitions are false with high probability. They correspond to premature decisions.

In summary, the acquisition characteristics found with the scalar dynamics (10) have also been observed with the second order model (12). They are general features for the class of linear models (9) considered in this work.

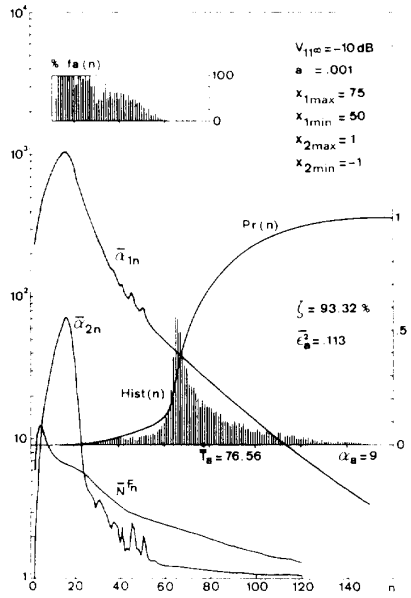

Fig. 13. In terms of simulation parameters, this Figure differs from Fig. 11(a) only in the noise condition: here $V_{11\infty} = -10$ dB, while $V_{11\infty} = -20$ dB in Fig. 11(a). Notice the effect of degrading the noise condition (increasing $V_{11\infty}$): significant shift and spread of the acquisition histogram, most of the early acquisitions being wrong (see plot of $\%f_a(n)$).

C. Acquisition Threshold

In all simulations performed we have considered the acquisition threshold as a constant parameter α_a, not depending on the time index n. Decreasing α_a induces a shift of Hist(n) to the right and reduces the number of false acquisitions. On the other hand, most of these false acquisitions are early events. This suggests that in practice one may want to design a time dependent threshold $\alpha_a(n)$ that avoids premature detections and that shortens the long right tail of Hist(n). A Neyman–Pearson type criterion can be adopted: keep the probability of false acquisition (false alarm) below a preset value and maximize the probability $P_r(T_{\max})$ of detecting within a given maximum delay $T_{\max}$.

V. CONCLUSIONS

This paper is concerned with ranging, the first stage in range/Doppler radar/sonar systems. We cast the problem as the acquisition step in phase/frequency demodulation. Ranging, or acquisition, is a global problem which is intrinsically nonlinear. Linearized solutions as the EKBF or the PLL are not suitable. These perform well as trackers, where one is interested in following the local dynamics, but fail to provide global observability.

An implementation of the optimal NLF was developed by exploiting the relevant characteristics of phase estimation. Under the setup studied here, the filter nonlinearity resides in the filtering step, in which the prediction density is multiplied by the periodic sensor factor. The adopted approach associates a Gaussian function to each period of the sensor factor, as explained and discussed in Subsection IIIB. The resulting filtering density is a weighted sum of Gaussian functions with common covariance matrix.

Acquisition is formalized by introducing an internal measure of dispersion of the filtering density, which relates its global covariance matrix with the common covariance matrix of its modes. Acquisition time is then defined as the first passage moment of that measure (acquisition process) across a suited threshold. Incorporating this mechanism, the NLF becomes an estimation/detection structure.

A first-order phase dynamics was used to illustrate the behavior of the NLF and to study its performance via Monte Carlo simulations; these include a comparison of the proposed structure against the BEKBFs equipped with the same detection scheme. To characterize the acquisition mechanism, a set of statistics were obtained, the most important being the acquisition time histogram and the detector performance. The NLF (which can adapt to the data by creating new modes) exhibits better performance.

The acquisition characteristics found for the scalar dynamics were also observed with the second-order model.

APPENDIX. EKBF/PLL

This Appendix derives, for reference purposes, the equations of the discrete EKBF corresponding to models (8) and (9).

The discrete EKBF operates in two steps: prediction and filtering. There is however no representation of the sensor factor H_n as in Section IIIB; instead, the nonlinear functions of the observation model (8) are linearized in each filtering step around the mean of the prediction density, which is assumed Gaussian:

$$P_n' \propto \mathcal{N}(\mathbf{x}_n - \boldsymbol{\eta}^{P_n}, \mathbf{V}^{P_n}). \tag{55}$$

As a consequence of this linearization, the observation factor looses its multimodal structure reducing to the Gaussian form

$$H_n' \propto \mathcal{N}(x_{1,n} - \eta^{H_n'}, r) \tag{56}$$

where

$$\eta^{H_n'} = \eta_1^{P_n'} + z_{2,n}\cos(\eta_1^{P_n'}) - z_{1,n}\sin(\eta_1^{P_n'}). \tag{57}$$

Filtering: Multiplying (55) by (56) leads to

$$F_n' = k^{F_n'}(\mathbf{x}_n - \boldsymbol{\eta}^{F_n}, \mathbf{V}^{F_n}) \tag{58}$$

with

$$k^{F_n'} \propto \mathcal{N}(\boldsymbol{\eta}_1^{P_n'} - \eta^{H_n'}, V_{11}^{P_n'} + r) \tag{59}$$

$$\boldsymbol{\eta}^{F_n'} = \boldsymbol{\eta}^{P_n'} + \frac{z_{2,n}\cos(\eta_1^{P_n'}) - z_{1,n}\sin(\eta_1^{P_n'})}{V_{11}^{P_n'} + r} \times \begin{bmatrix} V_{11}^{P_n'} \\ \vdots \\ V_{1k}^{P_n'} \end{bmatrix} \tag{60}$$

$$\mathbf{V}^{F_n'} = \mathbf{V}^{P_n'} - \frac{1}{V_{11}^{P_n'} + r} \times \begin{bmatrix} V_{11}^{P_n'} & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ V_{1k}^{P_n'} & 0 & \cdots & 0 \end{bmatrix} \mathbf{V}^{P_n'}. \tag{61}$$

When using the EKBF alone, there is no need for computing the factor $k^{F_n'}$. In a structure formed by a parallel of EKBF, as considered in Section IVA5, the weights work as coupling elements, being calculated and normalized in each iteration.

Prediction: From the convolution of (58) with (17) one has again a Gaussian prediction density, P_{n+1},

$$\boldsymbol{\eta}^{P_{n+1}'} = \mathbf{A}\boldsymbol{\eta}^{F_n'} \tag{62}$$

$$\mathbf{V}^{P_{n+1}'} = \mathbf{A}\mathbf{V}^{F_n'}\mathbf{A}^T + \mathbf{B}\mathbf{Q}\mathbf{B}^T \tag{63}$$

The estimate is, at each iteration,

$$\hat{\mathbf{x}}_n = \boldsymbol{\eta}^{F_n'}. \tag{64}$$

Equations (61) and (63) are algebraic Riccati equations describing the deterministic evolutions of $\mathbf{V}^{F_n}$ and $\mathbf{V}^{P_n}$. Starting with $\mathbf{V}^{P_1}$, the covariance matrices eventually reach the steady state characterized by

$$\mathbf{V}^{F_{n+1}} = \mathbf{V}^{F_n} \qquad \text{and} \qquad \mathbf{V}^{P_{n+1}} = \mathbf{V}^{P_n}. \tag{65}$$

Equations (60) and (62) become then a constant gain estimator, formally equivalent to the PLL. In this paper, when referring to the PLL, this asymptotic, constant gain EKBF is meant.

The first diagonal element of the steady state covariance matrix, obeying (65), coincides with the phase variances V_∞ and $V_{11\infty}$ expressing noise conditions in Section IV: $(V^{F_n} = \sigma^{F_n})_{n\to\infty} = V_\infty$, for scalar dynamics; $(V_{11}^{F_n})_{n\to\infty} = V_{11\infty}$ for vector dynamics.

REFERENCES

[1] Moura, J. M. F. (1979)
Passive systems theory with narrow-band and linear constraints: Part III—Spatial/temporal diversity.
IEEE Journal of Oceanic Engineering, **OE-4** (July 1979), 113–119.

[2] Moura, J. M. F. (1981)
The hybrid algorithm. A solution to acquisition and tracking.
Journal of the Acoustical Society of America—JASA, **69** (June 1981), 1663–1672.

[3] Bucy, R. S., and Joseph, P. D. (1968)
Filtering for Stochastic Processes with Applications to Guidance.
New York: Wiley, 1968.

[4] Moura, J. M. F., and Belo, C. A. C. (1989)
Threshold extension by nonlinear techniques.
In Y. T. Chan (Ed.), *Underwater Acoustic Data Processing*.
Boston: Kluwer Academic Publishers, 1989, 433–452.

[5] Moura, J. M. F. (1987)
Stochastic filtering: Linear and nonlinear.
In J. L. Lacoume and R. Stora (Eds.), *Signal Processing—Les Houches XLV*.
Amsterdam: North-Holland Publishing Co., 1987, invited chapter.

[6] Meyr, H., and Ascheid, G. (1990)
Sinchronization in Digital Communications.
New York: Wiley, 1990.

[7] Moura, J. M. F. (1979)
Passive systems theory with narrowband and linear constraints: Part II—Temporal diversity.
IEEE Journal of Oceanic Engineering, **OE-4** (Jan. 1979), 19–30.

[8] Viterbi, A. J. (1971)
Principles of Coherent Communication.
New York: McGraw-Hill, 1971.

[9] Bethel, R. E., and Rahikka, R. G. (1992)
Multisignal time delay detection and tracking.
IEEE Transactions on Aerospace and Electronic Systems, **28** (July 1992), 675–696.

[10] Bethel, R. E., and Rahikka, R. G. (1992)
Multisignal time delay detection and tracking.
IEEE Transactions on Aerospace and Electronic Systems, **28** (July 1992), 675–696.

[11] Bucy, R., Hecht, C., and Senne, K. D. (1972)
An engineer's guide to building nonlinear filters.
Technical report SRL-TR-72-0004 (Vol. 1 and 2), Frank J. Seiler Research Laboratory—USAF Academy, 1972.

[12] Bucy, R. S., and Mallinckrodt, A. J. (1973)
An optimal phase demodulator.
Stochastics, **1** (1973), 3–23.

[13] Bucy, R. S., Moura, J. M. F., and Mallinckrodt, A. J. (1983)
A Monte Carlo study of optimal absolute phase demodulation.
IEEE Transactions on Information Theory, **IT-29** (July 1983).

[14] Moura, J. M. F., and Baggeroer, A. B. (1988)
Phase unwarping of signals propagated under the Arctic ice crust: A statistical approach.
IEEE Transactions on Acoustics Speech and Signal Processing, **36** (May 1988).

[15] Leitão, J. M. N., and Moura, J. M. F. (1981)
Algorithm structures for cyclic phase estimation.
In *GRETSI Symposium on Signal Processing*, Nice, 1981.

[16] Leitão, J. M. N., and Moura, J. M. F. (1981)
Implementation of a 2-dimensional phase estimator.
In R. Boite and P. Dewilde (Eds.), *Circuit Theory and Design*.
North Holland: Delft University Press, 1981, 597–600.

[17] Cover, T., and Thomas, J. (1991)
Elements of Information Theory.
New York: Wiley, 1991.

[18] Kullback, S. (1978)
Information Theory and Statistics.
Peter Smith, 1978.

[19] Leitão, J. M. N., and Moura, J. M. F. (1994)
Nonlinear phase estimators based on the Kullback distance.
In *Proceedings of the 1994 International Conference on Acoustics, Speech, and Signal Processing—ICASSP'94*, Adelaide, Australia, 1994.

[20] Lo, J. T. (1969)
Finite-dimensional sensor orbits and nonlinear filtering.
Technical report AE-114, U.S.C. Aerospace Engineering, Aug. 1969.

[21] Alspach, D. L., and Sorensen, H. W. (1972)
Bayesian estimation using Gaussian sum approximation.
IEEE Transactions on Automatic Control, **17** (Aug. 1972), 439–448.

[22] Tam, P. K. S., and Moore, J. B. (1977)
A Gaussian sum approach to phase and frequency estimation.
IEEE Transactions on Communications, **COM-25** (Sept. 1977), 935–942.

José M. N. Leitão was born in Aldeia Velha–Avis, Portugal, on September 9, 1946. He received the E.E. degree and the Ph.D. degree in electrical engineering, in 1970 and 1983, respectively, from Instituto Superior Técnico (IST), Technical University of Lisbon. He received his "Agregado" degree in electrical and computer engineering, also from IST, in 1992.

He was with the Laboratory of Physiology of the Instituto Gulbenkian de Ciência, in Oeiras, Portugal, from 1970 to 1972. After spending 3 years at the University of Tübingen, Germany, he joined the faculty of IST in 1976, where he is currently Professor Catedrático with the Department of Electrical and Computer Engineering, teaching courses on telecommunications and communication theory. He is also the coordinator of the Communication Theory research group in the Instituto de Telecomunicações, IST. His main research interests are: communication theory, signal and image processing, and pattern recognition.

José M. F. Moura (S'71—M'75—SM'90—F'94) received the engenheiro electrotécnico degree in 1969 from Instituto Superior Técnico (IST), Lisbon, Portugal, and the M.Sc., E.E., and the D.Sc. in Electrical Engineering and Computer Science from the Massachusetts Institute of Technology (M.I.T.), Cambridge, in 1973 and 1975, respectively.

He is currently a Professor of Electrical and Computer Engineering at Carnegie Mellon University (CMU), Pittsburgh, which he joined in 1986. Prior to this, he was on the Faculty of IST where he was an Assistant Professor (1975), Professor Agregado (1978), and Professor Catedrático (1979). He has had visiting appointments at several Institutions, including M.I.T. (Genrad Associate Professor of Electrical Engineering and Computer Science, 1984–1986, also associated with LIDS) and the University of Southern California (research scholar, Department of Aerospace Engineering, Summers 1978–1981). His research interests lie in statistical signal processing (one and two dimensional), array processing, underwater acoustics, and multiresolution techniques. He has organized and codirected two international scientific meetings on signal processing theory and applications.

Dr. Moura has over 140 technical contributions, including invited ones, published in international journals and conference proceedings, and is co-editor of the books *Nonlinear Stochastic Problems* (Reidel, 1983) and *Acoustic Signal Processing for Ocean Exploration* (Kluwer, 1993). He was elected Fellow of the IEEE in November 1993 and corresponding member of the Academy of Sciences of Portugal (Section of Sciences) in July 1992. He is a member of the IEEE Press Board since 1991, a technical Associate Editor for the *IEEE Signal Processing Letters*, and a member of the Underwater Acoustics Technical Committee of The Signal Processing Society. He was an Associate Editor for *The Signal Processing Transactions* (Sonar and Radar) from 1988–1992 and a member of the technical committee of The IEEE International Symposium on Information Theory (ISIT'1993). He is affiliated with several IEEE societies, Sigma Xi, AMS, IMS, and SIAM.