Phase Unwrapping of Signals Propagated Under the Arctic Ice Crust: A Statistical Approach

JOSÉ M. F. MOURA, MEMBER, IEEE, AND ARTHUR B. BAGGEROER, FELLOW, IEEE

Abstract—**The paper studies the phase unwrapping of time signals transmitted under the Arctic ice crust. Unexpected phase discontinuities observed in a previous study [14] of the same data prompted the need for a robust phase unwrapper. The acoustic source generates a narrow-band sequence whose phase experiences random fluctuations. At the receiving hydrophones, the measurements are corrupted by wide-band noise. With this formulation, the reconstruction of the random phase naturally fits the setting of statistical signal processing [16]. We apply discrete optimal nonlinear filtering techniques [5] to design the phase unwrapping algorithm. The scheme presented proves more insensitive to noise than other nonstatistical unwrappers, supporting less stringent prefiltering constraints. Under the conditions of the Arctic experiment, it accommodates processing of each individual sensor's output data, avoiding the need for beam forming. Also, to discriminate between intrinsic phenomena and transients induced by the prefilters, shorter duration impulse responses are desired. Accordingly, the prefilters are designed with large bandwidths, the statistical unwrapper withstanding satisfactorily the lower signal-to-noise ratio environment. To take advantage of this flexibility, the data are processed under alternative conditions which assume different values for the statistical parameters. The similarities of the corresponding unwrapped phase paths help to discriminate those rapid events which are intrinsic from those which are artifacts introduced by the processing. By using real signals that traveled several hundred kilometers through a highly unstable channel, the work shows that nonlinear statistical designs are viable and useful in many practical problems of underwater acoustics.**

I. Introduction

THE present work studies the phase of continuous wave acoustic signals after being transmitted several hundred kilometers under the ice crust in the Arctic Ocean. The actual data collection experiments have been described previously; see [8], [2], [3]. They are part of an ongoing research program started in 1977, which includes onsite data gathering missions in the Arctic, e.g., the Fram 2 (1980) expedition, the Fram 4 (1982), the Marginal Ice Zone Experiment MIZEX 83, MIZEX 84, and MIZEX (1987).

Data collected under Fram 2 were studied in [14]. The processing showed that the phase exhibited unexpected phase discontinuities. These raised puzzling questions. First, they were a rare and fast phenomenon. Out of 300 min of signals analyzed, only two phase jumps were observed, each one lasting for just a few seconds. Second, the phase slips occurred in the same recorded experiment, less than 5 min apart. Finally, and perhaps more important, each slip amounted to strictly less than a cycle (2π). It is well known from communications theory that phase unwrappers are cursed by large sudden phase discontinuities induced by even small noise variations. On a frequency modulation (FM) receiver, they are heard as clicks. The crackling sounds become more frequent at signal-to-noise ratios (SNR) below a threshold level. These phase slips are multiples of full cycles.

Plausible explanations for the jumps detected in [14] include the following: 1) a source anomaly, 2) a channel instability, 3) a recording malfunction, or 4) a processing deficiency. To better understand the two latter ones, a new phase unwrapper has been developed. The new design takes into account the statistical nature of the signals. At low SNR's, the statistical unwrapper is more robust than other phase unwrapping algorithms, exhibiting fewer phase jumps. Without noticeable performance degradation, it accommodates individual processing of the data of each sensor, prior to array beam forming, as well as preprocessing filters with large bandwidths and short impulse response durations. The greater flexibility provided by the statistical algorithm is particularly relevant when attempting to decide if phase discontinuities are intrinsic or are the result of processing artifacts.

In Section II, we give a brief description of the data measuring experiments, summarizing previous processing experience. Phase unwrapping is dealt with in several distinct aspects in Sections III–VI. Section III overviews the problem. Section IV presents the statistical design. Section V details its implementation as a compromise between performance and complexity. This is an adaptation of the algorithm in [10], [11] to our problem. Section VI analyzes the statistical unwrapper behavior under controlled conditions (simulated study in a digital computer). Finally, Section VII applies it to the Arctic data. The preprocessing is examined there, and the robustness to poor

Manuscript received December 1, 1986; revised October 19, 1987. This work was supported by the Office of Naval Research, Arctic Program Office, under Research Contract N00014-86-K-0325. The processing reported here and the text preparation used computational facilities made available by the Laboratory for Information and Decision Systems (M.I.T.), sponsored by the U.S. Army Research Office under Contract DAAG-29-84-K-005.

J. M. F. Moura was with the Department of Electrical Engineering and Computer Science and LIDS, Massachusetts Institute of Technology, Cambridge, MA, on leave from Instituto Superior Técnico, Lisbon, Portugal. He is now with the Department of Electrical and Computer Engineering, Carnegie-Mellon University, Pittsburgh, PA 15213.

A. B. Baggeroer is with the Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, MA 02139.

IEEE Log Number 8719447.

prior statistical knowledge utilized to resolving questions about the nature of fast phase changes.

The paper shows that sophisticated algorithms can be used in the context of many underwater acoustics experiments. Extending the signal-to-noise ratio threshold of good performance, the statistical procedure places less stringent constraints on the preprocessing schemes, reducing the chance of contamination of the raw data by spurious events. The price paid is in terms of sophistication, and hence of computer power needed. However, as argued later on, simple-minded approximations to optimal approaches translate into reasonable compromises that obtain most of the performance gain available at the cost of a fractional increase of the processing complexity. To the authors' knowledge, the present is the first published account of the use of nonlinear stochastic filters in a real-world problem. That the data have propagated and were collected under severe conditions just strengthens the belief that these techniques should find in the future a wider utilization.

II. Data Collection Experiments in the Arctic Ocean

The research program Fram 2 was conducted in the Eurasian Basin of the Arctic Ocean from ice stations on drifting pack ice from March 19, 1980 to May 5, 1980. The scientific program is described in [8]. It involved three ice stations—a main camp, Fram 2, and two remote sites. The camp was approximately 2 × 2 km. It included a 22-channel array of hydrophones, suspended to a depth of 91 m through holes drilled in the ice. The interelement spacing was log periodic, and the overall dimensions were approximately 800 m on each leg of an L-shaped array. Ice breakup forced redeployment into an X shape. Each hydrophone was monitored separately, digitized at a 250 Hz sampling rate with timing provided by a global clock, and recorded on magnetic tape. The signals we are concerned with originated at camp 1, about 343 km to the north of Fram 2. They consisted of coherent acoustic tones centered at 15 Hz. The source was a Hydro-acoustics HLF-3 operated at a depth of 91 m. The peak power was 165–177 dB re 1 μPa at 1 m. The response of the receiving system rolled off above 80 Hz. The disturbances were very low-frequency noise of mechanical origin, wide-band omnidirectional noise in the 1–80 Hz band, and a 60 Hz interference. See [2] for further details.

To establish the stability in time of the acoustic channel for transmission of low frequencies under the ice in the Arctic Ocean, the Fram 2 tonal data have been extensively analyzed in [14]. The processing tasks included beam forming with a standard delay and sum technique, decimation by a factor of 2, quadrature demodulation, and low-pass filtering (LPF). The LPF used a 500-point finite impulse response (FIR) digital filter rolling off from 0 dB at 0.016 Nyquist to −50 dB at 0.024 Nyquist. At the 0.008 s sampling interval, the cutoff frequency is 1 Hz. The quadrature demodulated data were further decimated by a factor of 16, low-pass filtered again to obtain a 128 mHz passband centered on the demodulated tone (cutoff at 64 mHz). The phase sequence was reconstructed by unwrapping the arctan of the quadrature components of the filtered signal; see Section III. Although the received signal is composed of many single paths traveling under the ice crust, the analysis in [14] concludes that the fluctuations of the single-path vectors is negligible compared to the noise. A suitable model assumes that the total envelope amplitude ρ is the sum of a constant signal vector and a noise vector whose quadrature components are Gaussian. The statistical tests carried out substantiated this hypothesis, which again underlies the model to be presented in Section IV. Only two deep fades were identified, both in the same experiment, one larger than 20 dB and the other of 5 dB. The corresponding skips observed in the phase unwrapped path were intriguing in that they did not correspond to full cycle slips. Questions remained about the true nature of these events. The fades duration was short, on the order of the time constants of the filters used in [14]. To determine if the fades were not the result of all channels missing a sample, their duration should be tested using alternative filters with visibly different time constants. However, the phase-demodulated scheme of [14] did not endure shorter FIR's, larger filter bandwidths, or individual processing of single hydrophones. The design of a more reliable phase unwrapper is necessary to understand these anomalies.

III. Phase Unwrapping of Discrete Time Signals

Phase unwrapping of discrete time signals (DTS) is important in many signal processing applications, from speech to geophysics or underwater acoustics. In the communications field, it is known as phase demodulation. Given a complex-valued time series $x[n]$, $n = 0, 1, \cdots$, the phase is constructed as the continuous sequence

$$\tan\left(\tilde{\phi}[n]\right) = \frac{\text{Im}\left\{x[n]\right\}}{\text{Re}\left\{x[n]\right\}}. \tag{3-1a}$$

The no-memory subroutine computing the inverse tangent function only contributes to the principal value of the phase, i.e., its value modulo 2π. To obtain a sequence with values in $\mathbb{R}^1$ requires the unwrapping of (3-1a). Several authors have studied this problem. We consider only the question of unwrapping the phase of time signals, not that of the Fourier spectrum. A common procedure used in practice, for example in [14], chooses at time n the unwrapped phase $\tilde{\phi}[n]$ as that version of

$$\arctan\left(\frac{\text{Im}\left\{x[n]\right\}}{\text{Re}\left\{x[n]\right\}}\right) + k2\pi \tag{3-1b}$$

where k is an integer that leads to a deviation in absolute value from $\tilde{\phi}[n-1]$ not larger than π. In [13] and [4], the problem is addressed by looking at the shifted version of $\tilde{\phi}[n]$ that suitably accounts for a multiple number of shifts of $\pm\pi$, depending on the zero crossings of Re $\{x[n]\}$. To avoid a global dependence on possible local

errors resulting from incorrect identification of these zeros, an algorithm using Sturm sequences is developed in [13] and applied in [4] to symmetric and antisymmetric sequences.

The models underlying the phase unwrapping algorithms discussed in the literature do not account for two features that are present in many signal processing applications of interest, namely, that the measurement of $x[n]$ are compounded and corrupted by noise, and that the phase itself may be a random sequence. Bringing this to bear has several implications. On the one hand, the problem becomes harder; it lies in the realm of stochastic filtering with nonlinear signals. On the other hand, a more robust phase estimate is to be expected since the noisy nature of the involved signals is taken into consideration. We will see that the phase unwrapper developed under the statistical framework eliminates most of the phase jumps that are apparent in other unwrappers. The algorithm combines, in an optimal statistical sense, the information provided by the data sequence and the *prior* information regarding the constraints the data should satisfy. A reasonable model considers, for example, that the phase is a second-order random sequence; see Section IV. The statistical unwrapper is designed as a discrete-time stochastic nonlinear filter. The theory provides a nonlinear recursive scheme that updates, via Bayes' law, the probability distribution function (pdf) of the phase process given the measurements. If the phase has values in the field of the reals (not digitized), the pdf is defined on the real line. In a finite resource environment, actual computation requires finite representations of this function. Except in trivial applications, no finite representation can capture the complete description of the pdf, approximations being mandatory. The next two sections describe the design theory of optimal discrete nonlinear filtering and its practical implementation in the context of the problem under study.

IV. A Statistical Phase Unwrapping Algorithm

The phase process $\phi[n]$ is

$$\phi[n+N] + a_{N-1}\phi[n+N-1] + \cdots + a_0\phi[n] = b_N u[n+N] + \cdots + b_0 u[n], \tag{4-1a}$$

i.e., it is the output of a linear discrete time invariant system driven by independent, identically distributed (iid) Gaussian random variables $u[n]$ with spectral level q. Some or all of the a_i's, $i = 0, \cdots, N-1$ and some or all but one of the b_j's, $j = 0, \cdots, N$ may be zero. The initial conditions

$$\phi[0], \cdots, \phi[N-1]$$

are jointly Gaussian random variables, independent of the $\{u[n]\}$ for all n. Below, it is useful to have in mind the one-dimensional model

$$\phi[n+1] = a\phi[n] + u[n]. \tag{4-1b}$$

The signal process is

$$\tilde{x}[n] = \sqrt{2P}\exp\{j\phi[n]\} = x_c[n] + jx_s[n]. \tag{4-1c}$$

The measurements are

$$\tilde{z}[n] = \tilde{x}[n] + \tilde{w}[n]$$

with

$$\tilde{w}[n] = W_n \exp\{j\theta_n\} = w_c[n] + jw_s[n]$$

where $\{w_c[n], w_s[n]\}$ are zero-mean iid Gaussian random variables with spectral level r, independent of $\{u[n]\}$ for every n, and independent of the initial phase values or the initial data. We actually work with real signals, and let

$$z[n] = \begin{bmatrix} z_1[n] \\ z_2[n] \end{bmatrix}$$

where $z_1[n] = \tilde{z}_c[n]$, $z_2[n] = \tilde{z}_s[n]$. Similarly for $x[n]$ and $w[n]$. Introduce the normalized noise-to-signal ratio σ in the filter band

$$\sigma = \frac{r \cdot W}{P} \tag{4-2}$$

where W is the filter equivalent (half) bandwidth.

There are two possible phase estimates. One is the (first component of the) mean of the conditional predicted pdf

$$P_n[\Lambda] = \Pr\{\phi[n] \le \Lambda | z[0], \cdots, z[n-1]\}$$

and the other is the mean of the filtered pdf

$$F_n[\Lambda] = \Pr\{\phi[n] \le \Lambda | z[0], \cdots, z[n]\}.$$

To compute the conditional mean phase $\hat{\phi}[n]$, one needs to propagate the pdf's P_n and F_n. This is provided by Bayes' law, e.g., [5], which reduces to two steps.

Prediction Step:

$$P_n = F_{n-1} * Q_n. \tag{4-3}$$

Filtering Step:

$$F_n = (P_n \cdot H_n) \times K_n. \tag{4-4}$$

In (4-3), the predicted pdf P_n at time n is obtained by a convolution $(*)$ operation of the filtered pdf F_{n-1} available at time $n-1$ with a Gaussian kernel Q_n corresponding to the pdf of the iid random variables $\{u_n\}$. Having present (4-1b), (4-3) should be no surprise since the pdf of the sum of two independent random variables is the convolution of the corresponding pdf's.

In (4-4), a pointwise multiplication $(\cdot)$ of P_n with a Gaussian kernel H_n (the pdf of the random variables $\{w_n\}$) followed by a normalization constructs F_n at time n. Actually, (4-4) is accomplished usually in two substeps:

$$F_n^1 = P_n \cdot \hat{H}_n \tag{4-5a}$$

$$F_n = F_n^1 \times K_n \tag{4-5b}$$

where

$$\tilde{H}_n(\Lambda) = \exp\left\{\frac{z_1[n]\cos\Lambda + z_2[n]\sin\Lambda}{rW}\right\} \quad (4\text{-}5c)$$

is a periodic unnormalized kernel. Equation (4-5b) is a normalization procedure guaranteeing that F_n remains a pdf with

$$K_n = \frac{1}{\int F_n^1[d\Lambda]}. \quad (4\text{-}5d)$$

Equation (4-4) is easily justified. By Bayes' law, F_n is

$$\Pr\left\{\phi[n]\,|\,z[0], \cdots, z[n]\right\} = \frac{\Pr\left\{\phi[n], z[0], \cdots, z[n]\right\}}{\Pr\left\{z[0], \cdots, z[n]\right\}}$$

which in turn

$$= \frac{\Pr\left\{\phi[n], z[n]\,|\,z[0], \cdots, z[n-1]\right\}}{\Pr\left\{z[n]\,|\,z[0], \cdots, z[n-1]\right\}}$$

or

$$= \frac{\Pr\left\{\phi[n]\,|\,z[0], \cdots, z[n-1]\right\}\Pr\left\{z[n]\,|\,\phi[n], z[0], \cdots, z[n-1]\right\}}{\Pr\left\{z[n]\,|\,z[0], \cdots, z[n-1]\right\}}.$$

In the last equality, the first factor in the numerator of the right-hand side is P_n of (4-4). Given the independence between the phase sequence $\phi[n]$ and the measurement noise sequence $w[n]$, the second factor in the numerator is the kernel H_n of (4-4). Finally, the denominator is the inverse of the normalizing constant K_n.

The filtered phase estimate sequence is obtained by

$$\hat{\phi}[n] = \int \Lambda F_n[d\Lambda]. \quad (4\text{-}6)$$

The recursion (4-3)–(4-6) provides the basic equations for the statistical phase unwrapper. Because the support of the densities P_n and F_n is the whole real line $\mathbb{R}^1$ and not just a cycle $[-\pi, \pi]$, the estimate provided by (4-6) takes values on the line $\mathbb{R}^1$, just as the phase (4-1) does. The unwrapping of the phase sequence is accomplished *directly* by statistical means. No use is made of the cyclic arctan of (3-1b), so the corresponding decision problems of how to unwrap the phase are avoided. When noise disturbs the measurements, and in particular for high noise-to-signal power ratios, the fact that the noise is accounted for by the model and taken into consideration by the receiver design explains why the algorithm withstands many of the phase jumps experienced by other techniques. This will be elaborated upon later on. The tradeoff may lie in the complexity of (4-3)–(4-6). This is certainly true when they are compared to a straightforward unwrapper (see Section III), being the price paid for better performance. It is questionable, however, if the statistical phase unwrapper is significantly more complex than the unwrappers of [4] or [13]. As discussed in the next section, the implementation of the filter may be quite simple and still achieve most of the performance gain available. Hence, we have the interest of the present algorithm to many underwater signal processing applications.

V. Implementation of the Statistical Phase Unwrapper

In a digital environment, recursion (4-3)–(4-6) requires a finite representation of the functions $F_n(\Lambda)$, $P_n(\Lambda)$, $H_n(\Lambda)$, and $Q_n(\Lambda)$. Except in special cases, the finite description is not possible, approximations being required, with inherent resulting errors. The implementation of the algorithm must strike a balance between the complexity of the solution and the errors one is willing to accommodate. The same question arises in many other fields, e.g., integration of partial differential equations (PDE). In a sense, (4-3)–(4-5) are the integral solution of a parabolic PDE. So, a differencing (implicit) grid scheme is one possible representation method, the resulting routine propagating the discretizing grid and the probability mass on the mesh points. The uniform computational flow of this procedure is an advantage that, when tied to machine architectural constraints, leads to speedups of the algorithm. Also, reducing the grid mesh provides a simple way of controlling the corresponding numerical errors. The disadvantage lies on the associated computational effort. For a summary of these issues, see [6], [7].

We take here an alternative route that explores the specifics of the phase unwrapping problem. The algorithm adapts to the present signal processing problem the phase demodulator developed in [10], [11]. Recall that the class of Gaussian functions is closed under convolution and pointwise multiplication. Gaussian functions are completely described by their first- and second-order moments. If we represent F_n, P_n, H_n, and Q_n by finite sum of weighted Gaussian functions, each cycle of the algorithm just propagates the set of means, covariances, and weights associated with each function. The idea of using Gaussian functions for nonlinear filter implementation dates back to [1]. However, the implementation discussed in the remainder of this section is different, particularly simple, and highly tailored to the phase problem.

We first address the general issues associated with the Gaussian representation. Let

$$G(\mu, \sigma, \Lambda) = \exp\left\{-\tfrac{1}{2}(\Lambda - \mu)^T \sigma^{-1}(\Lambda - \mu)\right\} \quad (5\text{-}1)$$

be the (unnormalized) Gaussian function with mean μ and covariance σ. Assume that a representation for F_n by means of a finite sum of weighted Gaussians is available:

$$F_n = \sum_{i=1}^{N_n^F} K_n^F(i)\, G\left(\mu_n^F(i), \sigma_n^F(i), \Lambda\right). \quad (5\text{-}2a)$$

Denote by

$$\mathfrak{R}_n^F = \left\{\mu_n^F(i),\ \sigma_n^F(i),\ K_n^F(i)\right\}_{i \le N_n^F} \quad (5\text{-}2b)$$

the corresponding set of means, covariances, and weights. Each term in (5-2a) will be referred to as a Gaussian mode and abbreviated by $G_n^F(i)$. Indexes and arguments are kept only if necessary. Likewise, define equivalent representations for P_n, H_n, and Q_n. With this setup, we analyze the nonlinear algorithm (4-3)–(4-6). At the prediction step, substitution of Gaussian representations for F_{n-1} and Q_n in (4-3) leads to

$$P_n = \sum_{i,j} K_{n-1}^F(i)\, K_n^Q(j)\, [G_{n-1}^F(i) * G_n^Q(j)] \quad (5\text{-}3a)$$

which, after renumbering, becomes

$$= \sum_{l=1}^{N_n^P} K_n^P(l)\, G(\mu_n^P(l),\ \sigma_n^P(l),\ \Lambda) \quad (5\text{-}3b)$$

with

$$N_n^P = N_{n-1}^F \cdot N_n^Q \quad (5\text{-}3c)$$

prediction Gaussian modes. Similarly, the filtering step (4-4) or (4-5a) leads to

$$F_n = \sum_{i,j} K_n^P(i)\, K_n^H(j)[G_n^P(i) \cdot G_n^H(j)] \quad (5\text{-}4a)$$

$$= \sum_{l=1}^{N_n^F} K_n^F(l)\, G(\mu_n^F(l),\ \sigma_n^F(l),\ \Lambda) \quad (5\text{-}4b)$$

with

$$N_n^F = N_n^P \cdot N_n^H \quad (5\text{-}4c)$$

filtering modes. Using (5-2b), (5-3) and (5-4) can be represented by

$$\mathfrak{R}_{n-1}^F \xrightarrow{(5\text{-}3)} \mathfrak{R}_n^P \quad (5\text{-}4) \quad \mathfrak{R}_n^F.$$

The diagram illustrates that (5-3) and (5-4) require the updating of the sets of means, covariances, and weights associated with each representation. In general, the equations propagating the means and covariances are nothing but the discrete Kalman–Bucy filter equations [9]. The use of generic Gaussian representations will require, besides overhead, as many Kalman–Bucy filters as the number of propagating modes. In turn, the statistical unwrapper substantially reduces this effort.

For the phase model of Section IV, the nonlinearity is in the observations, (4-1c). Accordingly, the kernel Q_n associated with the linear phase process is Gaussian, i.e.,

$$\forall n:\ N_n^Q = 1 \quad (5\text{-}5a)$$

so that by (5-3c)

$$N_n^P = N_{n-1}^F, \quad (5\text{-}5b)$$

there is no mode creation in the prediction step. The observation kernel H_n or $\tilde{H}_n$ given by (4-5c) is not Gaussian. This raises two remarks. One is how to approximate $\tilde{H}_n$ by a train of N_n^H Gaussians. That is a *representation* issue. The other is how to keep the required number of Gaussian modes within a predetermined value given that $N_n^H > 1$ leads, through (5-4c), to an explosive growth of N_n^F. That is a *projection* problem. The simplicity of the approach to be studied and where it departs from the Gaussian sum method of [1] resides exactly on how to resolve satisfactorily these two questions. We address briefly both.

Representation of $\tilde{H}_n$: Rewrite (4-5c) as

$$\tilde{H}_n(\Lambda) = \exp\left[\lambda_n \cos\left(\Lambda - \mu_n^H(0)\right)\right] \quad (5\text{-}6a)$$

where the argument of the maximum

$$\mu_n^H(0) = \arctan\left(\frac{z_1[n]}{z_2[n]}\right) \quad (5\text{-}6b)$$

and the adaptive signal-to-noise ratio (SNR)

$$\lambda_n = \frac{\|z[n]\|}{rW} \quad (5\text{-}6c)$$

with

$$\|z[n]\| = \sqrt{z_1^2[n] + z_2^2[n]}. \quad (5\text{-}6d)$$

Realizing that $\tilde{H}_n$ is periodic, we reexpress it as

$$\tilde{H}_n(\Lambda) = \tilde{H}_n^0(\Lambda) * p(\Lambda) \quad (5\text{-}7)$$

where

$$\tilde{H}_n^0(\Lambda) = \begin{cases} \tilde{H}_n(\Lambda) & \Lambda \in \left[-\pi + \mu_n^H(0),\ \pi + \mu_n^H(0)\right] \\ 0 & \text{elsewhere} \end{cases}$$

is the fundamental period, and the peaked fence

$$p(\Lambda) = \sum_{k=-\infty}^{+\infty} \delta(\Lambda + k2\pi)$$

with $\delta(\cdot)$ the Dirac impulse. The fundamental period $\tilde{H}_n^0(\Lambda)$ is a Tikhonov function, well known in phase-locked loop (PLL) studies; see [18], [17]. It resembles a Gaussian when λ_n is large, and becomes flat as $\lambda_n \to 0$. The idea in [10] is to approximate the basic period $\tilde{H}_n^0(\Lambda)$ by a Gaussian which matches it in three points—the maximum (5-6b) and the points $\pi/2$ apart of the maximum; see Fig. 1 for a typical example.

It is straightforward to conclude that the matching Gaussian $G_n^H(0)$ has a mean $\mu_n^H(0)$ given by (5-6b) and variance

$$\sigma_n^H = \frac{\pi^2\sqrt{2P}}{8\|z[n]\|}\,\sigma. \quad (5\text{-}8a)$$

The representation of the sensor factor becomes, from (5-7),

$$\tilde{H}_n(\Lambda) = G(\mu_n^H(0),\ \sigma_n^H,\ \Lambda) * p(\Lambda) \quad (5\text{-}8b)$$

or

$$\tilde{H}_n(\Lambda) = \sum_{k=-\infty}^{+\infty} G(\mu_n^H(0) + k2\pi,\ \sigma_n^H,\ \Lambda). \quad (5\text{-}8c)$$

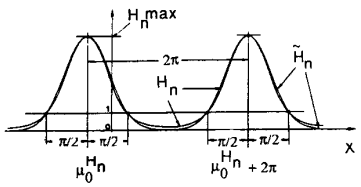


Fig. 1. Gaussian approximation to $\tilde{H}_n^0(\Lambda)$.

Equation (5-8c) represents $\tilde{H}_n$ by Gaussian modes which have the *same* variance and where the means are translates by 2π of $\mu_n^H(0)$. For a linear phase model, substitution in (5-3b) and then in (5-4b) shows that the filtering and prediction modes also have the same variance, i.e.,

$$\sigma_n^P(i) \equiv \sigma_n^P \qquad i = 1, \cdots, N_n^P$$
$$\sigma_n^P(i) \equiv \sigma_n^F \qquad i = 1, \cdots, N_n^F.$$

The end result is that only one Riccati equation is needed for all Kalman–Bucy filters; see the Appendix for a summary of the updating formulas. The overhead in computing (5-6b) and (5-8a) is significantly less than in other techniques that require at each iteration multiple linearizations plus a Riccati equation per Gaussian mode, e.g., [1]. Both facts suggest considerable computational savings. As a final comment, (5-8a) provides an adaptive mechanism of the value of the noise variance used by the filter in its calculations. For weak noise, σ_n^H is a renormalization of σ as it should be. On the other hand, if the noise annihilates one or both of the sensor measurements, σ_n^H is large. Filtering affects the phase estimate update little, the measurements being practically ignored by the filter. We now consider the second issue mentioned before.

Projection: The increase in the number of modes results at the filtering step. The projection is then applied at this stage of the algorithm, being accomplished in three phases.

1) Truncation: Each Gaussian mode G^P of P_n is multiplied by only the J most adjacent modes of $\tilde{H}_n$ so that

$$N_n^F = J \times N_n^P.$$

2) Agglutination: Modes $G_n^F(i)$ and $G_n^F(j)$ which are close to each other

$$\left\| G_n^F(i) - G_n^F(j) \right\| < \beta \qquad (5\text{-}9)$$

where β is a preset value are agglutinated. Agglutination of Gaussian functions is a Gaussian function whose mean, variance, and weight are easily evaluated from the corresponding values of the combined modes. For a two-dimensional phase process, (5-9) requires two constants β_1 and β_2.

3) Elimination: Mode $G_n^F(i)$ is eliminated if

$$K_n^F(i) < \epsilon$$

where ϵ is a small positive number. This constraint is equivalent to bounding the maximum number of modes by $1/\epsilon$. Practice shows that usually the average number of modes is well under this limit.

Besides the statistical parameters, constants J, β_i, $i = 1, 2$, and ϵ are needed. By prior experimentation, these are tuned to the particular application under study. If they are chosen so that N_n^F results on the average too small, the filter is on a predictive mode; transients may be missed. If they lead to N_n^F being too large, the algorithm relies heavily on new data, the filtering behavior prevails, the estimates are wigglier, and the long-term (local) mean-square error behavior deteriorates. The details are not further pursued, the interested reader being referred to [12].

The estimate of the phase is

$$\hat{\phi}_n = \frac{\sum_{i=1}^{N_n^F} K_n^F(i)\, \mu_n^F(i)}{\sum_{i=1}^{N_n^F} K_n^F(i)}. \qquad (5\text{-}10)$$

Note that (5-10) is a line estimate, i.e., the unwrapping of the phase has been accomplished without ambiguity.

VI. Behavior Analysis Under Controlled Simulated Conditions

Due to the lack of analytical tools, the performance of nonlinear statistical filters is hard to quantify. In [11], extensive simulations established via Monte Carlo the mean-square error performance of the previous algorithm for several communications problems. Compared to other procedures, like the phase-locked loop (PLL) or the techniques of [1], it performed consistently better and resolved faster larger uncertainties about the phase process (i.e., sudden changes in the input phase signal). Instead of replicating the results of [11], we address here the novel aspects of what happens when there are discrepancies between the model underlying the design and the data collected from the real world. These may arise from errors in the values assumed for modeling parameters or from structural mismatches between the model and the physical system. To understand how the phase unwrapper behaves under different operating conditions, the results of a controlled experiment are reported here. Since our concern is with divergence issues, we focus on the qualitative properties of the procedure. The simulations try to emulate the Arctic data from a noise-to-signal ratio point of view. In this and next sections, the acronyms SPW and APW are used to refer to the statistical phase unwrapper and to the arctan phase unwrapper, respectively.

The simulations generated a chirp signal of the type used to improve resolution in radar systems. This choice is made because of its nonstationary character and its high-frequency contents. Also, it helps understanding the tracker's behavior when the signal moves out of the filter band; see Experiment 3 below. The phase signal is

$$\psi[n] = 2\pi f_0 \Delta n + \phi[n]$$

where

$$\phi[n] = 2\pi \frac{\alpha_0}{2} \Delta^2 (n^2 - n) + \phi[0] + \zeta[n]$$

is the low-pass component, α_0 is the chirp constant, f_0 is the frequency offset, Δ is the sampling interval, and $\zeta[n]$ is a random term. Defining

$$\Lambda_1[n] = \frac{1}{2\pi}\phi[n]$$

$$\Lambda_2[n] = \frac{1}{2\pi}\frac{\phi[n+1] - \phi[n]}{\Delta}$$

and

$$\Lambda[n] = [\Lambda_1[n]\ \Lambda_2[n]]^T$$

a suitable second-order difference equation modeling the low-pass phase process is

$$\Lambda[n+1] = \begin{bmatrix} 1 & \Delta \\ 0 & 1 \end{bmatrix}\Lambda[n] + \begin{bmatrix} 0 \\ 1 \end{bmatrix}\alpha_0\Delta + \begin{bmatrix} 0 \\ 1 \end{bmatrix}\sqrt{\Delta}u[n]$$

$$\Lambda[0] = \begin{bmatrix} \phi[0] \\ 0 \end{bmatrix}. \qquad (6\text{-}1)$$

The process $\Lambda_1[n]$ normalizes the phase to cycles of 2π. In all subsequent plots, phase and phase rate are measured in cycles and cycles per second, respectively. The set $\{u[n]\}$, $n > 0$, corresponds to a sequence of iid Gaussian random variables with zero mean and power level q. The observations are

$$y[n] = \sqrt{2P}\cos\psi[n] + w[n] \qquad (6\text{-}2)$$

where P is the signal power and $\{w[n],\ n \geq 1\}$ are iid Gaussian random variables, with zero-mean (double-sided) spectral level of $r/2$, and independent of the sequence $\{u[u]\}$.

A sequence of experiments was carried out where the response of the SPW is studied. For comparison, the phase path obtained by the APW is also considered. For all figures shown, we have the following.

Signal and Noise:

Δ = sampling interval = 0.008 s.
T = data span = 20 s.
P = signal power = 8×10^{-8} V^2/s.
W = system input bandwidth = 50 Hz.
α_0 = chirp parameter = 0.5/s^2.

Stochastic Phase Unwrapper (SPW):

J = truncation constant = 3.
β_i = agglutination parameters = 0.01, i = 1, 2.
ϵ = elimination parameter = 0.01.

Depending on the experiment, a combination of the following was adopted for the signal and noise:

f_0 = tonal frequency = 1 Hz; 15 Hz.
r = noise variance = 5×10^{-9}; 5×10^{-8} V^2/Hz.
q = phase noise variance = 0.1; 0.5.

The preprocessing consisted of quadrature demodulation with a frequency equal to the nominal tone, followed by low-pass filtering, as illustrated in Fig. 2. The low-pass

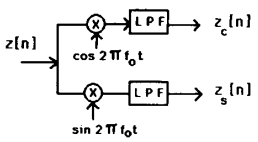


Fig. 2. Preprocessing: low-pass quadrature components.

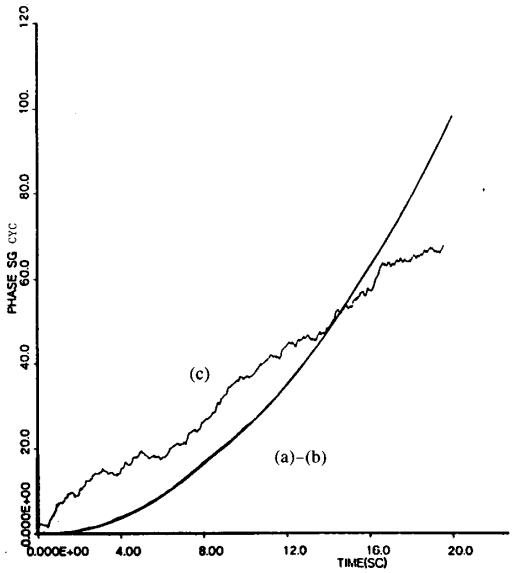


Fig. 3. Experiment 1: low-pass phase. (a) True process. (b) Statistical phase unwrapper estimate. (c) Atan phase unwrapper estimate.

components are then input to SPW and APW. To have a feeling for the robustness of SPW, the algorithm was given in many of the runs erroneous values of the parameters. Details are described next.

Experiment 1: $q = 0.1$; $r = 5 \times 10^{-9}$

Although the chirp signal is not stationary, we need an estimate of its frequency contents. This was obtained by evaluating the Burg spectrum and used to design the low-pass filter (LPF) of Fig. 2. The LPF is an FIR with impulse response duration of 400 ms, i.e., extending over 50 samples, and bandwidth of about 25 Hz, large enough to pass all signal components. The filter is not made narrower so the conditions of next section where the preprocessing is to be as simple as possible are best replicated. Fig. 3 superimposes the phase paths of the true process (a), of SPW (b), and of APW (c). It is clear that APW has lost track, while SPW holds the noisy environment well, its phase estimate being almost undistinguishable from the true process. Also, SPW provides an estimate for the phase rate which, as shown in Fig. 4, is in good agreement with the phase rate of the real process.

Experiment 2: $f_0 = 15$ Hz; $q = 0.5$; $r = 5 \times 10^{-9}$

In this experiment, besides the higher frequency f_0, the phase noise variance q is boosted up. The Burg spectrum

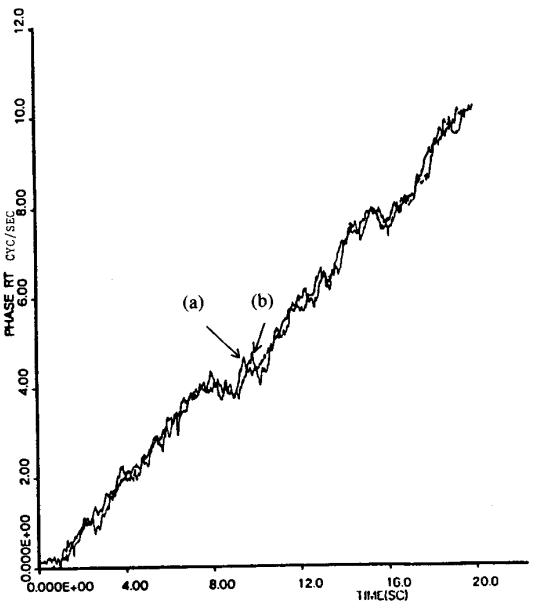


Fig. 4. Experiment 1: low-pass phase rate. (a) True process. (b) Statistical phase unwrapper estimate.

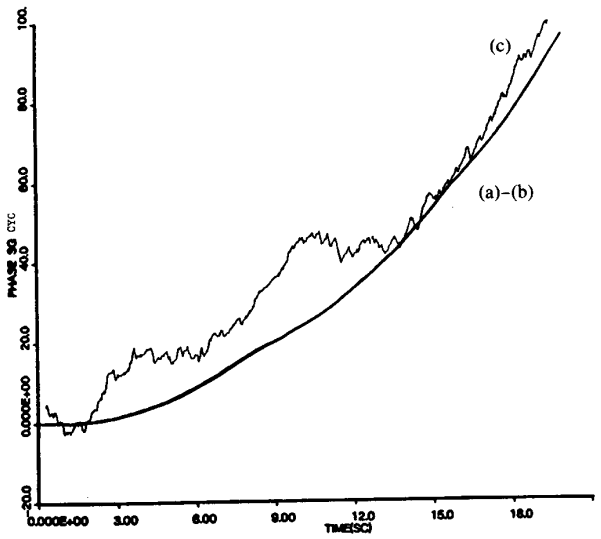


Fig. 5. Experiment 2: low-pass phase. (a) True process. (b) Statistical phase unwrapper estimate. (c) Atan phase unwrapper estimate.

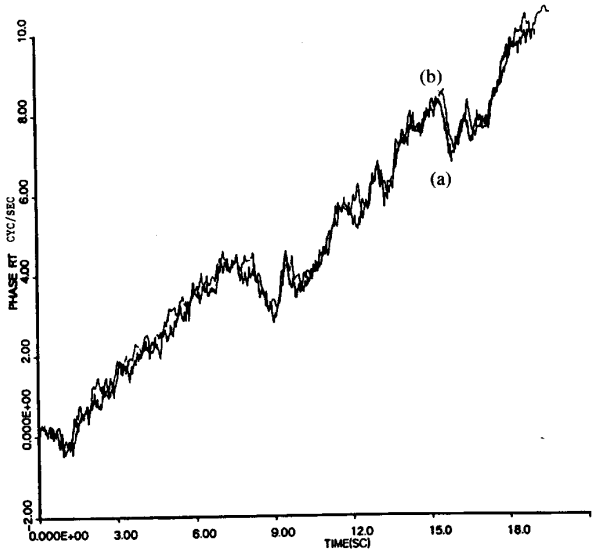


Fig. 6. Experiment 1: low-pass phase rate. (a) True process. (b) Statistical phase unwrapper estimate.

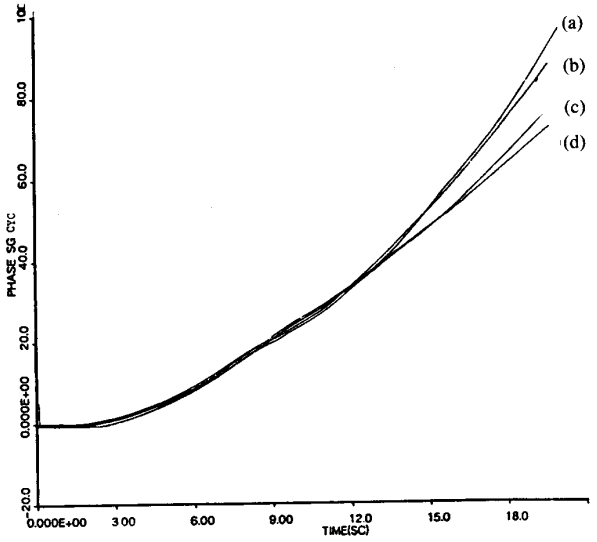


Fig. 7. Experiment 3: low-pass phase. (a) True process ($\alpha_0 = 0.5$, $q = 0.5$). (b) Statistical phase unwrapper estimate ($\alpha_0 = 0.3$, $q = 2$). (c) Statistical phase unwrapper estimate ($\alpha_0 = 0.3$, $q = 0.2$). (d) Statistical phase unwrapper estimate ($\alpha_0 = 0.2$, $q = 0.2$).

estimated the frequency contents of the signal to lie between 15 and 25 Hz. Figs. 5 and 6 repeat for the present experiment Figs. 3 and 4, confirming the observations made above.

Experiment 3: $f_0 = 15$ Hz; $q = 0.5$; $r = 5 \times 10^{-8}$

The signal values are as in the previous example; the noise level is increased by 10 dB. The noise is now of the same strength of the signal. This experiment is targeted to showing how SPW behaves when there are misfits in the model. The LPF used has a cutoff frequency at approximately 5 Hz, well below the bandwidth of the low-pass signal. At midpoint of the data span, the chirp signal falls outside the filter band. In the first half of the observation interval, at the input of SPW, there are signal and noise. In the second half, the signal is missing, resulting in a basic inconsistency between the model and the (simulated) real world. The SPW was run several times over the same data, each run being fed with a different set of assumed values of the chirp parameter α_0 and of the phase noise variance q. Figs. 7 and 8 display the phase and phase rate results. In the first half of the record, there is a reasonable agreement between all estimates and the true process, irrespective of the offsets in the parameters. In the

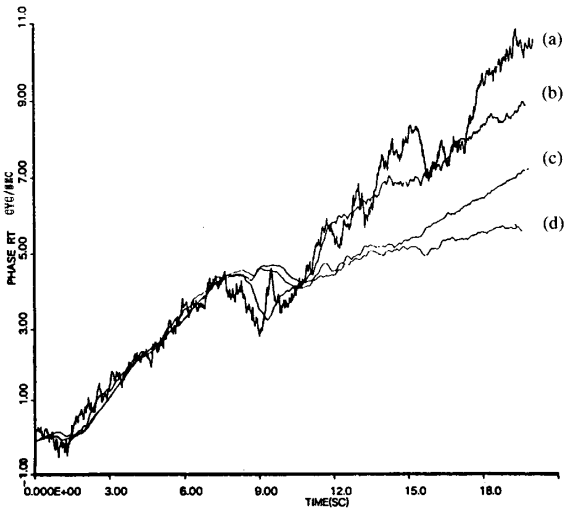


Fig. 8. Experiment 3: low-pass phase rate. (a) True process ($\alpha_0 = 0.5$, $q = 0.5$). (b) Statistical phase unwrapper estimate ($\alpha_0 = 0.3$, $q = 2$). (c) Statistical phase unwrapper estimate ($\alpha_0 = 0.3$, $q = 0.2$). (d) Statistical phase unwrapper estimate ($\alpha_0 = 0.2$, $q = 0.2$).

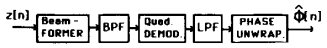


Fig. 9. Arctic data block processing.

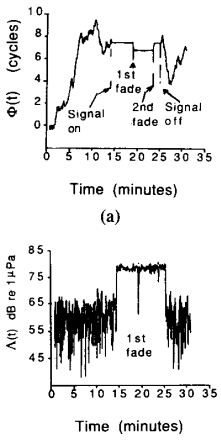


Fig. 10. (a) Phase path in [14, Experiment 4]. (b) Signal amplitude in [14, Experiment 4].

second part, the SPW exhibits a remarkable behavior. It is apparent that each SPW follows a different path. The absence of the signal being unknown to the algorithm, the data being the same, it is concluded that the filter follows in the second half of the observation interval, a path which is principally determined by the prior statistics assumed. Said in other words, when the signal is present, to a large extent, SPW is insensitive to offsets in the parameters' values, filtering being the dominant mode. When the signal is absent, the model assumed in the design of SPW no longer is in accordance with reality, prediction prevailing over the behavior of SPW. In a real-world application, this behavior is useful to test for inconsistencies between the model and the physical system.

VII. Processing of the Arctic Data

Recalling from Section II, the signals to be analyzed were generated by a highly stabilized crystal source vibrating at a nominal 15 Hz frequency. They propagated about 300 km, being received by a 24 channel L-shaped array with leg span of 800 m. The sampling rate of digitization was 250 Hz. The receiving system response rolls off at 80 Hz; this is well above the Nyquist rate. The basic steps of the processing are summarized in Fig. 9.

In [14], the structure of Fig. 9 was used, with intermediate stages of decimation applied. The filters were FIR's of 500 sample points duration. After sufficient decimation, the final LPF had a bandpass of 128 mHz centered on the demodulated tone. This processing significantly reduced the noise-induced fluctuations on the phase process, the arctan providing a highly stabilized phase path. There were observed, however, unexplained peaks of very short duration. Identification of these fast events with the well-known cycle slip phenomena associated with phase unwrappers (see [17], [18]) is precluded because they did not correspond to an integer multiple of cycles, but rather to a fraction of a cycle. Because the dynamics of ice flow are much slower than the spikes' duration, it is unlikely that such fast phenomena were due to the temporary annihilation of a strong signal path.

To diagnose if they are or are not artifacts of the processing, it is important to process the data of individual hydrophones and to use shorter FIR's, the latter requiring larger filter bandwidths. In either case, the noise-to-signal power ratio is increased, the arctan phase unwrapper (APW) loses track, and the spikes are masked by the rapid succession of cycle slips. The stochastic phase unwrapper (SPW) of Sections IV and V is here applied. The data analysis pursued corresponded to several versions of the sequence described in Fig. 9. For the sake of brevity, only a representative set is reported here. It conveys and is supportive of the general conclusion provided by all the experiments that were carried out. Both beam-formed and individual sensors' data are studied. The analysis detailed below focuses on an interval of 150 s about the first spike identified in log 4 [14, Fig. 12], here reproduced for illustration purposes in Fig. 10(a). Fig. 10b, also of [14], shows the corresponding amplitude fades. The second spike of the same log was also the subject of our study, with identical conclusions about its nature having been reached. Due to space limitations, we do not present here the corresponding analysis.

To help with the identification of the main features of the received signals, Burg spectral estimates are constructed (see Fig. 11) for the beam-formed Burg spectrum (160 poles). Besides the low-frequency contents deriving from struming and array vibration, one distinguishes in a

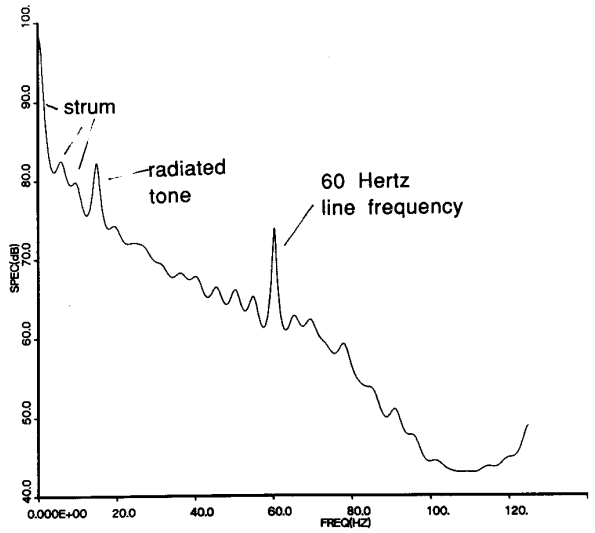


Fig. 11. Burg spectrum (160 poles): beam former.

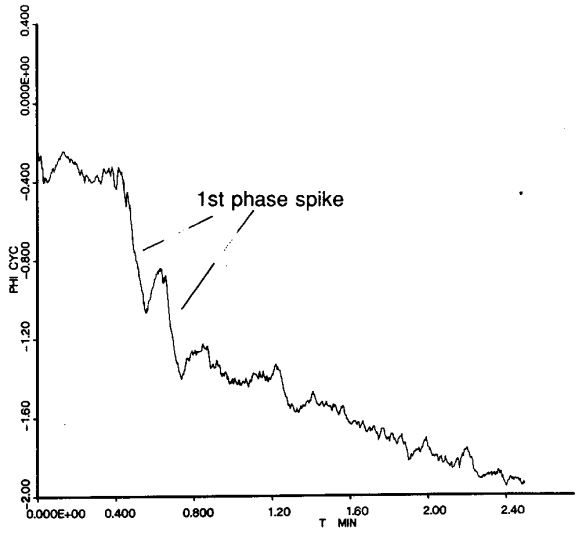


Fig. 12. Statistical unwrapper phase: beam former.

background of wide-band noise the tone of interest at 15 Hz and a vestigial 60 Hz power component. The bandpass filter (BPF) and the low-pass filter (LPF) of the schema in Fig. 14 were designed via the Remez-Exchange algorithm. The BPF is centered at 15 Hz with a passband of 10 Hz. The LPF has a half-bandwidth of 5 Hz. The filters' impulse responses have equivalent durations of 50 points. Compare these values to the 64 mHz cutoff frequency and 500 points duration for the impulse response in [14]. Decimation with a factor of 2 was used only before LPF so that the FIR's of both filters are, respectively, 200 and 400 ms. This is much smaller than the duration of the spike; it is then possible to rule out improbable but possible rare events like loss of a clock pulse in the recording equipment.

Application of SPW requires a suitable statistical model. The phase process of the tone generated by the crystal is taken as a second-order linear process (see Section IV or in Section VI take $\alpha_0 = 0$), the narrow-band signal being assumed a pure sinusoidal carrier whose frequency is chosen at the receiving end as the peak of the Burg spectrum. An unintentional downward trend exhibited by the phase paths below is a result of a slight offset of this assumed frequency with respect to its real value. The statistical parameters, namely q and r (see again Section VI), are not known exactly. A reasonable estimate of the noise-to-signal ratio for the beam-formed data can be obtained from [14]. Alternatively, it can be estimated from the Burg spectral estimates. From Section VI, it is learned that the SPW is robust with respect to perturbations on the exact values of the parameters, and that if filtering is the dominant mode, the global behavior of SPW is not affected by these offsets. If the model misfit is too large, then Experiment 3, Section VI says that the SPW response is markedly different. The processing results are now examined.

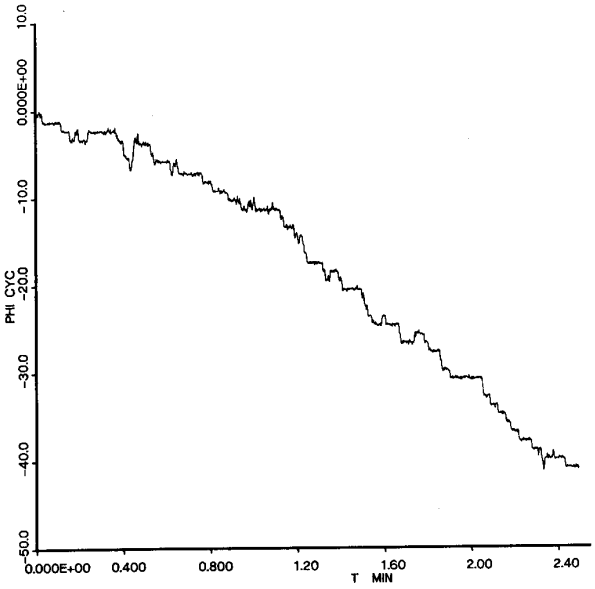


Fig. 13. Atan unwrapper phase: beam former.

The spike of concern is located about 250 s after the signal has been turned on in log 4. The data analyzed correspond to the interval [220 370] s. The origin of the time axes in all figures is relative to the lower end of this interval. Figs. 12–14 concern the beam-formed data. Fig. 12 shows the phase path reconstructed by the SPW, while Fig. 13 displays the corresponding APW curve. SPW shows no cycle slips, while APW extends on the same time interval over about 50 cycles. The latter provides no clues, the spikes masked by the rapid succession of slips. At the SPW output, it is readily apparent that on or after the 240 s time reference (0.4 min in the figures horizontal labelings), the phase slips down with a fast trend to an

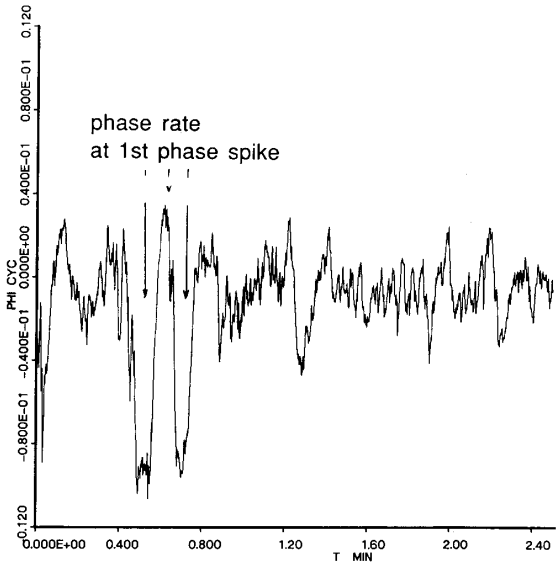


Fig. 14. Statistical unwrapper phase rate: beam former.

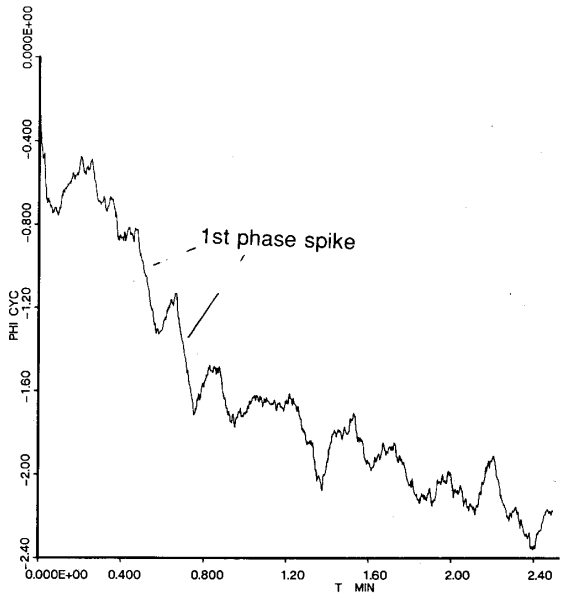


Fig. 15. Statistical unwrapper phase: Channel 2.

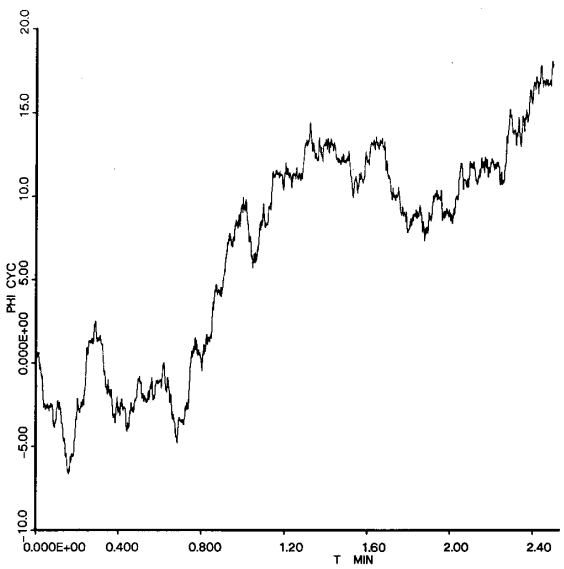


Fig. 16. Atan unwrapper phase: Channel 2.

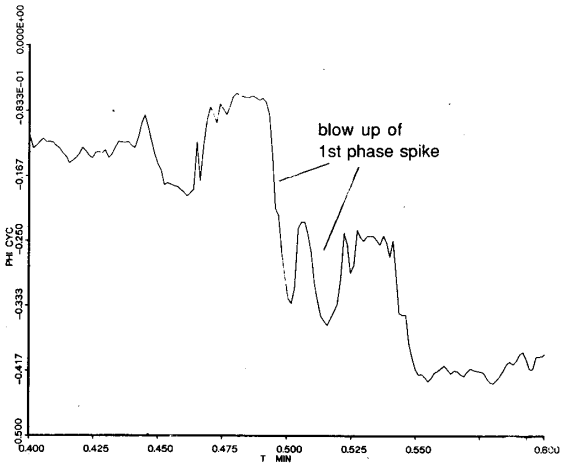


Fig. 17. Statistical unwrapper phase: Channel 10 (blow).

apparent recovery, followed by a fast down trend again, finally locking on the slow global trend. The event lasts for less than 18 s. The burst is also apparent from the frequency rate path provided in Fig. 14. A rough evaluation of the total difference between the upper (left) and lower (right) downward dominant trends, before and after the spike, shows a loss of approximately 0.6–0.8 of a cycle. Being less than 1 cycle, it cannot be attributed to a temporary loss of lock of the type occurring in phase and frequency modulation systems, known as cycle skips.

Figs. 15–20 repeat similar processing curves for the SPW and the APW tracks for individual sensors. Note the different vertical scales. The SPW phase again has a dynamic range of less than two cycles (see Figs. 15 and 19), the down trend being caused by the small offset in the center frequency of the quadrature demodulation block. APW has at least an order of magnitude larger dynamic range (Figs. 17, 19, and 21). In all SPW pictures, at the same time frame (and discounting the relative delays between sensors), the presence of a similar spiky pattern is distinguished in the phase process and its rate. The corresponding APW's curves are inconclusive. From sensor to sensor, the noise environment changes. In particular, in channel 10, the noise is much stronger than in other channels. By retuning the statistical parameters of the

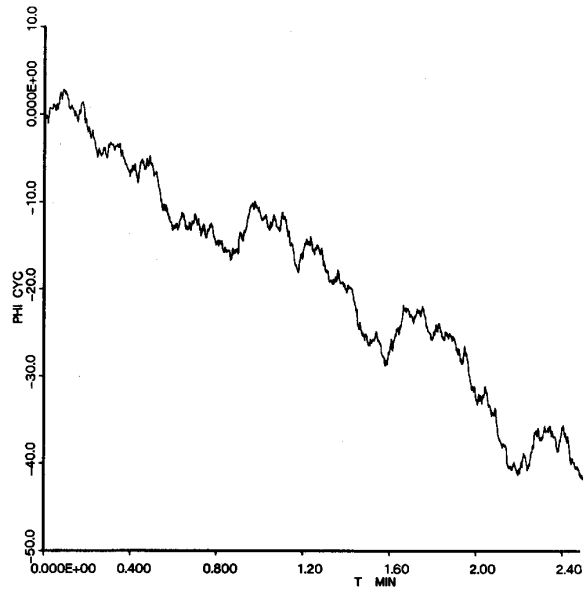


Fig. 18. Atan unwrapper phase: Channel 10.

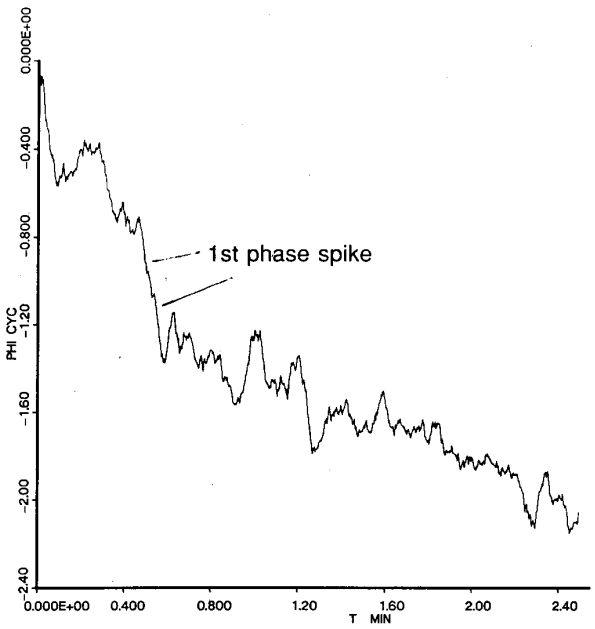


Fig. 19. Statistical unwrapper phase: Channel 20.

SPW, it is still possible to enhance the phase path, clean the noise, and exhibit the pattern of concern (see Fig. 17) where a blow of the SPW phase is presented.

Further experiments were carried out with the beamformed and individual channel's data where different BPF and LPF filters were used and where SPW was provided with alternative values for the model parameters. Consistently, the same spiky behavior was obtained, the curves being very similar to the ones shown. The insight provided by Experiment 3, Section VI justifies then the conclusion that this impulsive event is an intrinsic feature

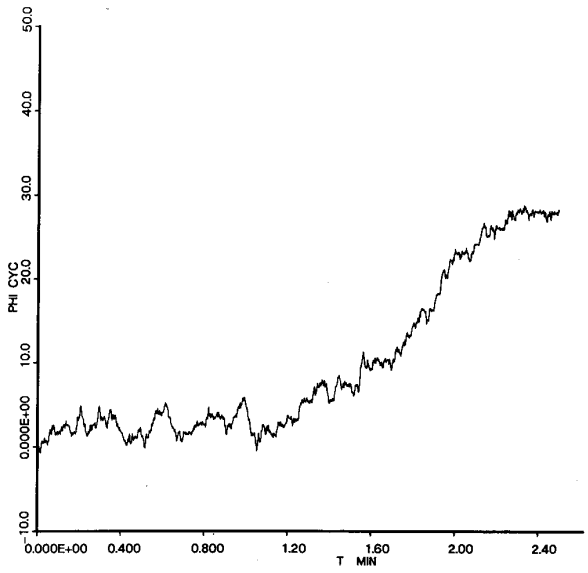


Fig. 20. Atan unwrapper phase: Channel 20.

of the recorded data, and that it is not artificially introduced by the processing. We can only speculate on its real causes: either one path becoming temporarily dominant or a glitch showing in the source.

VIII. Conclusion

The following are relevant issues drawn from the work described.

- An alternative to the phase unwrapping of time sequences was presented, designed via the techniques of stochastic nonlinear filtering theory. Taking into consideration the noisy nature of the phase instability and of the measurements disturbances, a reliable unwrapping is achieved.
- The statistical phase unwrapper (SPW) performance depends upon the nature of the misfits between the model underlying the design and the real world data. Errors resulting from offsets of the assumed parameter values affect the local behavior, but do not imply global divergence between the responses of differently tuned SPW's processing the same real data. Errors corresponding to basic misadjustments between the model assumptions and the real world translate into lasting long-term effects in the SPW response, with each differently tuned SPW exhibiting its own distinct phase path (see Section VI). With detuning errors, the SPW reacts dominantly in a filtering mode. With model mismatches, the predictive mode is prevalent in the SPW response.
- The previous remark says that the SPW is robust to parameter errors, and that it can be used to resolve fundamental inconsistencies between the model and the real physical system. This suggests that the nonlinear algorithm has a finite memory span, being able to switch between regions of operation. In one region, greater reliance to new input data is given, while in the other, greater emphasis is placed on the prior knowledge.

- Due to its robustness to imprecise knowledge of the parameters values and to higher noise levels, the SPW supports more flexible preprocessing. The necessity of beam forming may be waived, and the prefilters required to clean the data may be designed with shorter responses and larger bandwidths.
- Application of the SPW to the Arctic data shows that the spiky pattern of concern is present in both beam-formed and single sensors data and when differently tuned SPW's run over the same data. The widely distinct conditions of operation (faster FIR's, beam forming and unbeam forming, largely different levels of noise-to-signal power ratios) leading systematically to the same consistent global behavior support the conclusion that the impulsive behavior is not a processing artifact; rather it is an intrinsic characteristic of the recorded data. The exact origin, e.g., source transient, recording malfunctioning, is not clear. But with high confidence, the SPW helps to rule out the demodulation processing as one of its causes.
- The work has described an application of nonlinear stochastic filtering theory to a real world problem. The flexibility provided by the SPW and its resistance to noise justifies its wider use in environments where the signal-to-noise power ratio is small, as typically happens in underwater acoustics.

Appendix

For simplicity, the Appendix considers the updating equations for the one-dimensional linear phase model

$$\phi[n+1] = a\phi[n] + u[n]$$

(see Section IV for details). They are as follows.

Prediction:

$$\sigma_n^P = a^2\sigma_{n-1}^F + q$$

$$\mu_n^P(i) = a\mu_{n-1}^F(i) \qquad i = 1, \cdots, N_n^P$$

$$K_n^P(i) = K_{n-1}^F(i) \qquad i = 1, \cdots, N_n^P$$

where q is the variance of the noise $u[n]$. Recall from Section V that

$$N_n^P = N_{n-1}^F.$$

Filtering:

$$\frac{1}{\sigma_n^F} = \frac{1}{\sigma_n^P} + \frac{1}{\sigma_n^H}$$

$$\mu_n^F(i,k) = \mu_n^P(i) + \sigma_n^F(\sigma_n^H)^{-1}\left[\mu_n^H(0) - \mu_n^P(i) + k2\pi\right]$$

$$i = 1, \cdots, N_n^P$$

$$K_n^F(i,k) = K_n^P(i)\, G\left(\mu_n^H(0) + k2\pi,\ \sigma_n^P + \sigma_n^H,\ \mu_n^P(i)\right)$$

$$k = J_1 + 1, \cdots, J_1 + J.$$

The resulting number of modes is

$$N_n^F = N_n^P J.$$

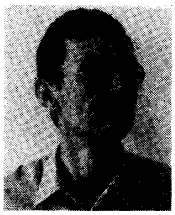
After the filtering recursion is accomplished, projection as described in Section V is carried out, and the modes are renumbered. To obtain a true density, the weights are then normalized.

Acknowledgment

The authors thank Dr. P. Mikhaelevsky for the discussions and help concerning the processing in [14].

References

[1] D. L. Alspach and H. W. Sorensen, "Nonlinear Bayesian estimation using Gaussian sum approximations," *IEEE Trans. Automat. Contr.*, vol. AC-17, pp. 439–448, Aug. 1972.
[2] A. B. Baggeroer and I. Dyer, "Fram 2 in the eastern Arctic," *EOS Trans. Amer. Geophys. Union*, Apr. 6, 1982.
[3] A. B. Baggeroer and G. L. Duckworth, "Seismic exploration in the Arctic Ocean," *Geophysics: The Leading Edge of Exploration*, pp. 22–27, Oct. 1983.
[4] T. E. Bordley, "Phase unwrapping of symmetric, anti-symmetric sequences," Ph.D. dissertation, Dep. Elec. Eng. Comput. Sci., Massachusetts Inst. Technol., Cambridge, 1985.
[5] R. S. Bucy and P. D. Joseph, *Filtering for Stochastic Processes with Applications to Guidance*. New York: Wiley Interscience, 1968; New York: Chelsea, 1987.
[6] R. S. Bucy, F. Ghovanlou, J. M. F. Moura, and K. D. Senne, "Array processing and phase modulation (*Invited Paper*)," *IEEE Comput. Mag.*, Special Issue on Array Processing, pp. 51–61, June 1983.
[7] R. S. Bucy and K. D. Senne, "Nonlinear filtering algorithms for vector processing machines," in *Computers and Mathematics with Applications*, Vol. 6. New York: Pergamon, 1980, pp. 317–338.
[8] I. Dyer and A. B. Baggeroer, "The Fram 2 experiment," *EOS Trans. Amer. Geophys. Union*, vol. 61, no. 6, 1980.
[9] R. E. Kalman and R. S. Bucy, "New results in linear filtering and prediction theory," *Trans. ASME, J. Basic Eng.*, ser. D, vol. 83, pp. 95–108, 1961.
[10] J. M. N. Leitão and J. M. F. Moura, "Nonlinear filtering in phase acquisition," in *Signal Processing Theories and Applications*, M. Kunt and F. de Coulon, Eds. Amsterdam, The Netherlands: Holland, 1980, pp. 437–442.
[11] J. M. N. Leitão, "Estimação não linear óptima de fase. Acquisição e seguimento," Doctoral dissertation, Dep. Elec. Comput. Eng., Instituto Superior Técnico, Lisbon, Portugal, Feb. 1983.
[12] J. M. N. Leitão and J. M. F. Moura, "Phase demodulation: A nonlinear approach," to be submitted.
[13] R. McGowan and R. Kuc, "A direct relation between a signal time series and its unwrapped phase," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-30, pp. 719–725, Oct. 1982.
[14] P. Mikhaelevsky, "Characteristics of CW signals propagated under the ice in the Arctic," *J. Acoust. Soc. Amer.*, vol. 70, no. 6, pp. 1717–1722, 1981.
[15] K. D. Senne, "A machine independent random number generator," *Stochastics*, vol. 1, no. 3, pp. 3–23, 1973.
[16] H. L. Van Trees, *Detection, Estimation and Modulation Theory: Part I*. New York: Wiley, 1968.
[17] —, *Detection, Estimation and Modulation Theory: Part II*. New York: Wiley, 1971.
[18] A. J. Viterbi, *Principles of Coherent Communications*. New York: McGraw-Hill, 1966.



José M. F. Moura (S'71–M'75) received the engenheiro electrotécnico degree in 1969 from Instituto Superior Técnico (IST), Lisbon, Portugal, and the M.Sc., E.E., and the D.Sc. degrees in EECS from the Massachusetts Institute of Technology (M.I.T.), Cambridge, in 1973 and 1975, respectively.

In 1975 he joined the Faculty of IST where he was an Assistant Professor, Professor Agregado (1978), and Professor Catedrático (1979). From 1984 to 1986 he was Genrad Visiting Associate Professor in Electrical Engineering and Computer Science at M.I.T. Since 1986 he has been a Professor of Electrical and Computer Engineering at Carnegie-Mellon University, Pittsburgh, PA. He spent the summers of 1978–1981 as a Visiting Scientist with the Department of Aerospace Engineering, University of Southern California, Los Angeles. His research

interests lie in statistical communications, signal processing, and systems theory. He has over 50 contributions, including invited ones, published in international journals and conference proceedings, and is Co-editor of the book *Nonlinear Stochastic Problems* (Dordrect, The Netherlands: Reidel, 1983).

Dr. Moura is a member of several IEEE societies, of Sigma Xi, AMS, IMS, and Ordem dos Engenheiros.

Arthur B. Baggeroer (S'62–M'68–F'88) received the B.S.E.E. and M.Sc. degrees in electrical engineering from Purdue University, West Lafayette, IN, in 1963 and 1964, respectively, and the D.Sc. degree in electrical engineering from the Massachusetts Institute of Technology (M.I.T.), Cambridge, in 1968.

He joined the Faculty of M.I.T. in 1968, where he is presently a Professor of Electrical and Ocean Engineering and the M.I.T. Director of the M.I.T./WHOI Joint Program in Oceanography and Oceanographic Engineering. He was the Senior Scientist on the Fram II and Fram IV expeditions and took part on Mizex. He has also consulted for the U.S. Navy in array processing studies and for the M.I.T. Lincoln Laboratory, Lexington, on radar signature analysis. His research interests have involved the application of detection and estimation theory methods to radar, sonar, and seismic signal processing in oceanographic systems. He is the author of numerous papers and of the book *State Variables and Communication Theory* (Cambridge, MA: M.I.T. Press, 1970).

Dr. Baggeroer is a member of JASA and Sigma Xi.