

A Monte Carlo Study of Absolute Phase Determination

RICHARD S. BUCY, FELLOW, IEEE, JOSE M. F. MOURA, MEMBER, IEEE, AND
ALBERT J. MALLINCKRODT, SENIOR MEMBER, IEEE

Abstract—The problem of absolute phase tracking and the development of the properties of an optimal estimator of absolute phase are considered. Using Monte Carlo simulation, this estimator's performance is compared with that of the phase-locked loop on the basis of slip distribution growth rate. Further slip prediction is considered and a statistic, based on the entropy of the conditional distribution of the phase given the observations, is shown to be effective.

I. INTRODUCTION

ESTIMATING the phase of a narrow-band signal in the presence of noise is fundamental to the function of many radio-communications, navigation, and radiometric tracking systems. The classical phase-locked-loop (PLL) circuit provides a near-optimal solution to these problems under a wide variety of circumstances, particularly where the phase estimate's output noise-to-signal ratio is small so that the resultant phase-sensor nonlinearity can be well represented in terms of a linear expansion about the operating point.

For a number of reasons, however, receivers wherein accuracy, power efficiency, or range are particularly demanding often are designed to operate at or near the threshold of such nonlinearity. In communications, for example, various bandspreading modulation and coding systems permit a beneficial trade-off between ultimate accuracy and signaling bandwidth up to the onset of some detection threshold, commonly a phase-lock threshold. Similarly in radiometry, measurement accuracy is commonly proportional to the measuring signal's bandwidth, but only up to the point where, as bandwidth is increased, the input noise approaches a detection threshold which is again a phase-lock threshold. Thus it is no accident that the performance of central systems such as deep-space communications and tracking, worldwide navigation, etc., relates directly to phase estimation performance near

threshold, i.e., in just the region for which the PLL no longer provides the optimal phase estimate. Furthermore, in a number of such applications the economic incentive per decibel of threshold extension may well justify a significantly greater effort at improving that threshold. Such incentives, along with the growing availability of ultra-high-speed array processing capability at the minicomputer and potentially microcomputer level have motivated studies of the nonlinear phase estimator which provides the optimal phase estimate in just the critical region where the PLL fails [3].

In the past, the performance of phase or frequency (rate of phase) demodulators has been evaluated by using a cyclic error norm

$$L(\epsilon) = 2(1 - \cos \epsilon), \quad (1)$$

where ϵ is the error process. Equation (1) exhibits the desirable features of being cyclic and of reducing to ϵ^2 for small ϵ . Phase errors larger than 2π (referred to as cycle slips) are of only transient concern. There are practical instances, like in radiometry with DECCA, LORAN, OMEGA, etc., where distance is measured in terms of the transit time phase shift of a radio wave. While some of these systems include means for resolving cyclic ambiguities in the phase, the basic question is that of keeping track of the cumulative phase change (relative to an initial value) with phase defined on the infinite line $(-\infty, \infty)$. For these applications, where the cycle slips have a lasting effect, the minimum mean square error (mmse) or a slight modification thereof is a more appropriate norm. In this paper, the first problem is termed the modulo 2π or cyclic phase demodulation (CPD) problem, while the second one is the absolute phase demodulation problem (APD).

For the modulo 2π problem, nonlinear filters have been built via the direct implementation of the conditional cyclic density. They exhibit, below a threshold level, significant performance improvement over the PLL. Likewise, in principle, the APD is solved by finding the *a posteriori* conditional density $p(x, t|\theta_t)$ of the phase process given the observation field θ_t . However, the problem of effective representation of the conditional density is complicated by the growing support of the conditional density as the slips accumulate. For sufficiently many slips to occur, in a given time interval, the signal-to-noise ratio must be small. In turn, this requires a large grid span for accurate numerical propagation of the conditional density. This remark il-

Manuscript received June 4, 1981; revised April 12, 1982. This research was supported in part by NATO Scientific Affairs Division under Research Grant 1557/78 and in part by the United States Air Force Office of Scientific Research under Grant AFOSR-3100 and Contract F44620-76-C0085. This work was presented in part at the Information Theory Conference, Santa Monica, CA, January 1981.

R. S. Bucy is with the Department of Aerospace Engineering, University of Southern California, Los Angeles, CA 90007.

J. F. Moura is with the Centro de Análise e Processamento de Sinais, Complexo Interdisciplinar, Instituto Superior Técnico, Avenue Rovisco Pais, P-1000 Lisboa, Portugal.

A. J. Mallinckrodt is with the Communications Research Laboratories, Santa Ana, CA 92705.

illustrates the conflict inherent in the simulation of the problem. With time, the problem demands increasingly greater computational resources. The objectives of creating slip statistics of small dispersion, a finite number of mass points in the density representation as the support of the density grows, and of providing estimates which have fewer slips than the PLL, has proven difficult to obtain simultaneously. *Ad hoc* techniques have led only to partial results [1].

By looking at the information carried by the a posteriori probability density function $p(x, t|\theta_t)$, this paper studies the basic properties of the APD, with emphasis on the analysis of the cycle slipping behavior of several estimation structures. A framework for the APD is presented that encompasses the basic model, a direct finite representation for the nonlinear filter, and provides a sufficient statistic for the cycle-slip analysis.

Two important concepts play significant roles: a) the entropy of the conditional distribution, and b) the congruence assumption for the line distribution. As an indicator of disorder, the entropy is a relevant parameter for density coding. Studies of the entropy of the conditional line density and of the entropy of the phase process show a reduction of disorder of one order of magnitude at -3 dB noise-to-signal ratio as a result of observations. The congruence assumption (see Section IV) provides an interesting relation satisfied by the entropy of the congruent conditional density. Using this relation, one measures the (absolute modulus) difference between the entropies of the conditional density on the line and a derived congruent density. This measure is labeled the entropic measure of instability. It provides a means to identify regions of instability in the evolution of the line distribution where the density perceptibly departs from its congruent approximation, lobes being created, significantly reinforced or annihilated. It is in these regions of instability that the estimators can experience cycle slips. From our simulations, it is apparent that good estimators are able to withstand a large number of mass flow situations without loss of a cycle.

In overview, Section II defines the APD model; Section III describes an implementation of the nonlinear filter by a fixed mesh grid with respect to a moving center, discussing its computational advantages, compromises, and possible centering strategies. Section IV presents the congruence assumption and the entropic measure of instability, μ_t . Section V exhibits a sequence of pictures depicting the flow evolution of the line density as related to the time history of μ_t . As predicted by μ_t , undisturbed and turbulent modes are distinguished for the flow of the line density. Section VI shows the experimental evidence related to cycle slips accumulated for several filtering structures as well as histograms of slip behavior. Finally, Section VII advances conclusions and proposes further work to be pursued. This paper contrasts with most previous work on phase modulation in that the latter has not sought to synthesize the optimal demodulator and is instead concerned with analy-

sis or fixes of the phase-locked loop, see [7] and [3] for references.

II. ABSOLUTE PHASE DEMODULATION

For computational feasibility, a scalar phase process Brownian motion x_t is considered.¹ In Ito's differential form,

$$dx_t = d\beta_t \quad (2)$$

with $\{\beta_t, t \geq 0\}$ a standard Brownian motion with $E\beta_t^2 = qt$. The phase demodulation problem constructs an estimate \hat{x}_t of x_t based on the two-dimensional observation vector z_s , $s \leq t$, where

$$dz_t = h(x_t) dt + dv_t \quad (3)$$

with the cyclic sensor

$$h(x_t) = [\cos x_t, \sin x_t]^T \quad (4)$$

and $v_t = [v_t^1, v_t^2]^T$ a vector Brownian motion of covariance matrix $rt I$, independent of β_t .

The standard solution to this problem is given by the phase-locked loop (PLL)

$$d\hat{x}_t = -kh'(\hat{x}_t) dz_t, \quad (5)$$

The loop gain $k = \sqrt{q/2r}$ may be obtained via linearization and minimum mean square error arguments. It is noted that the PLL is the steady state extended Kalman-Bucy filter derived for the problem (2) to (4).

The optimal nonlinear filter propagates the conditional probability density function $p(x, t|\theta_t)$. This is accomplished either via the Stratonovich-Kushner partial differential equation

$$dp(x, t|\theta_t) = \frac{q}{2} \frac{\partial^2 p(x, t|\theta_t)}{\partial x^2} dt + \left[C - \hat{C}_t \ ; \ S - \hat{S}_t \right] \frac{df_t}{2r} p(x, t|\theta_t) \quad (6)$$

or via Green's function given by the representation theorem [2]. In (6), q and r are the noise variances, C and S represent $\cos x$ and $\sin x$, \hat{C}_t and \hat{S}_t are the components of the conditional mean estimate of the sensor $\hat{h}(x_t)$, and df_t is the innovations vector

$$df_t = dz_t - \hat{h}(x_t) dt \quad (7)$$

Approximating the continuous-time problem of (2) to (4) by the appropriate stochastic difference equations, a discrete-time formulation results. The discrete-time optimal nonlinear filter, straightforwardly derived from Bayes' law, is a recursion that at time n consists of two steps.

Prediction: Computation of the predicted probability function $P_{n|n-1}(x)$, the probability of phase at time $n\Delta$ given the observations up to time $(n-1)\Delta$. This requires a convolution

$$P_{n|n-1}(x) = Q * F_{n-1|n-1}(x), \quad x \in X, \quad (8)$$

¹Because the density is carried on the line a one-state dimensional problem is as computationally demanding as the two-state cyclic density which has the torus for support (see [3]).

with the dynamics represented by the Gaussian kernel Q with the filter probability function $F_{n-1|n-1}$, the probability density of the phase at time $\Delta(n-1)$ given the observations up to time $\Delta(n-1)$. Δ is chosen in a given signal and noise environment so that we obtain ten (10) samples per time constant of the PLL. We have found this to be an effective method for ensuring that the discrete-time problem closely approximates the continuous time one.

2) *Filtering*: Calculation of $F_{n|n}(x)$. This consists of a multiplication

$$F_{n|n}(x) = H_n(x)P_{n|n-1}(x), \quad x \in X, \quad (9)$$

of the sensor function $H_n(x)$ and the predicted density $P_{n|n-1}$. In (8) and (9),

$$Q(x) = \frac{1}{\sqrt{2\pi q_d}} \exp(-x^2/2q_d), \quad (10)$$

$$H_n(x) = k_1 \exp(-\|z_n - h(x)\|^2/2r_d), \quad (11)$$

$$q_d \equiv q\Delta; \quad r_d \equiv \frac{r}{\Delta},$$

X is the probability functions' support, k_1 is a constant which, in actual implementation, is absorbed into numerical normalization procedures, q_d and r_d are the equivalent discrete-time noise variances, $\|\cdot\|$ is the Euclidean norm.

For the modulo 2π or cyclic phase demodulation problem (CPD), the estimator that minimizes the cyclic loss function given by (1) is

$$x_i^* = \tan^{-1}(\hat{S}_i/\hat{C}_i), \quad x_i^* \in (-\pi, \pi). \quad (12)$$

It is constructed from estimates of cyclic quantities. These need not be obtained from $p(x, t|\theta_i)$, but can be obtained from its folded version

$$J(\tilde{x}, t|\theta_i) \triangleq \sum_{k=-\infty}^{\infty} p(2k\pi + \tilde{x}, t|\theta_i), \quad \tilde{x} \in [-\pi, \pi]. \quad (13)$$

The cyclic or folded pdf $J(\tilde{x}, t|\theta_i)$ follows updating schemes parallel to (6) or (8) and (9), referred to in the sequel as cyclic recursion. The CPD has bounded support $X = [-\pi, \pi]$, which leads to a more tractable representation of the density. Alternative implementation techniques and extensive Monte Carlo simulation studies for the discrete-time version of the CPD problem has been reported in [3] and [4].

For the line or absolute phase demodulation problem (APD), one is interested in tracking the process x_t on the real line $(-\infty, \infty)$. As the error process

$$\epsilon_t \triangleq x_t - \hat{x}_t \quad (14)$$

undergoes diffusion on the line, the *a posteriori* distribution spreads out over the real axis. In PLL studies (e.g., [5]), the theoretical performance of the PLL as a line tracker is assessed by decomposing the error process into the sum of two independent components, a cyclic (or modulo 2π) error $\tilde{\epsilon}(t)$ and a counting process $N(t)$:

$$\epsilon_t = \tilde{\epsilon}_t + 2\pi N(t), \quad \tilde{\epsilon}_t \in [-\pi, \pi]. \quad (15)$$

The process $N(t)$ accounts for the phase slips exhibited by the estimate. Under the independence assumption, and assuming $N(t)$ to be Poisson, the error variance is unbounded, and is given by

$$\sigma_{\epsilon_t}^2 \cong \sigma_{\tilde{\epsilon}}^2 + \sigma_N^2(2\pi)^2 t, \quad (16)$$

where σ_N^2 is the so-called diffusion coefficient.

For the nonlinear estimator, experiments have shown that the analogous counting process also has a variance that is linear in t . Later we will compare the variance growth rate of the PLL and the optimal filter. For more general estimator structures (16) may not hold, i.e., it is not obvious that the two error components in (15) are independent. However, the decomposition (15) is an intuitively satisfying tool for analysis, and it will be used to assess the performance of several filters for the APD problem.

We remark that the APD problem is intrinsically a transient problem. The steady-state distribution solution for the nonlinear filtering recursion is the trivial one, i.e., the uniform density on $(-\infty, \infty)$. From a practical point of view, there is a finite time horizon beyond which the accumulated errors are greater than the acceptable practical accuracy requirements, no matter how accurate the noisy measurements are (as long as plant and measurement noises of finite nonzero variance are assumed). In other words, the cyclic measurements provide only local information on the line phase process; how the process departs from a known initial condition (Brownian motion) is the global problem.

This remark motivates our statistical design of the Monte Carlo experiments which we performed to evaluate the distribution of slips and to study the behavior of the conditional density around slip points. We fix the sample path length so that there was small probability that the conditional density was not interior to the finite support chosen to represent the conditional density.

III. ALGORITHM IMPLEMENTATION

As given in Section II, the recursion to be implemented is, in symbolic notation, a convolution

$$P_n = F_{n-1} * Q, \quad (17a)$$

followed by a multiplication

$$F_n = P_n \cdot H. \quad (17b)$$

For the implementation of (17), a differencing scheme was adopted. This grid algorithm provides a controlled simple test on the attainable accuracy, setting a benchmark for other finite representations. The space dimension is discretized using a uniform fixed grid with floating center

$$G_n(l) = C_n + C(l), \quad l = 1, \dots, L, \quad (18a)$$

where the fixed grid is given by

$$C(l) = -(K+1)\pi + (k-1)2\pi + (m-1)\Delta_M, \\ m = 1, \dots, M; \\ k = (l-1) \bmod M + 1 = 1, \dots, K, \quad (18b)$$

with the total number of points being $L = MK$ and the grid mesh $\Delta_M = 2\pi/M$. In (18), K is the number of cycle intervals, M is the number of grid points in each 2π -cycle. In the simulation studies, K is chosen to be odd and M is even. The fixed grid with floating center affords computational savings by requiring a single evaluation of the kernel Q as the grid has *fixed* mesh and by simplifying the computation of the sensor factor H at each iteration.

Three centering strategies were tested. The first trivially sets $C_n = 0$. The second takes C_n as the maximum of the conditional distribution. The third chooses C_n as that grid location which at each iteration minimizes the outward mass flow, i.e., the density loss. The latter scheme leads to the best results and was finally chosen.

The definition of the algorithm requires assigning values to three quantities, M , K , and T . The number M of points in each cycle interval is adjusted to be a reasonable compromise between accuracy and computational effort (computer time and numerical error). The value $M = 17$ was used in most of the numerical studies presented in the sequel. The number K of cycle intervals is related to the time horizon T chosen for the simulation runs duration. In order to obtain accurate statistics for the cycle slip phenomenon, the simulation runs must exhibit a statistically significant number of slips. This may be attained either by choosing a large time horizon T or a small loop signal-to-noise ratio α . The nonlinear theory for the one-dimensional PLL [5] provides an estimate for the mean time to skip a cycle. The simulation parameters T and α are loosely set on the basis of this estimate. For a given T , K is chosen so that the total mass loss experienced in the run is smaller than a given preset value ϵ_p . As a further test on the accuracy of the density representation, the moments of the cyclic density obtained by direct folding of the line density P_n and of the cyclic density computed directly from the cyclic recursion are made to agree within a preset tolerance. Another number of importance in actual simulation concerns the support of the convolution kernel Q . Due to finite computation precision, the kernel need only be computed and stored for a finite number I of points of a fixed grid.

In Fig. 1, the block diagram of our simulation is given. This program flow chart exhibits two loops—a time loop nested in the simulation Monte Carlo loop. The main drive reduces to a set of calls to subroutines accomplishing specific tasks. The nonlinear line filter block includes a convolution loop. This loop is coded as a dot product in the array processor. The array processor also computes and accumulates the data required for entropy evaluation. The program has restartable features, saving the latest seed values for the random number generation. The code per iteration has not been optimized and requires an average of 92.5 ms. The Monte Carlo statistics were collected with a set of 23 experiments of 25×10^3 iterations each (corresponding to 2.5×10^3 PLL time constants) and for a noise-to-signal ratio of -3 dB. The average total mass loss (see eq. (18a)) observed was 3 percent. The statistical design of the Monte Carlo loop is discussed in [3], including the details of the random number generator and an

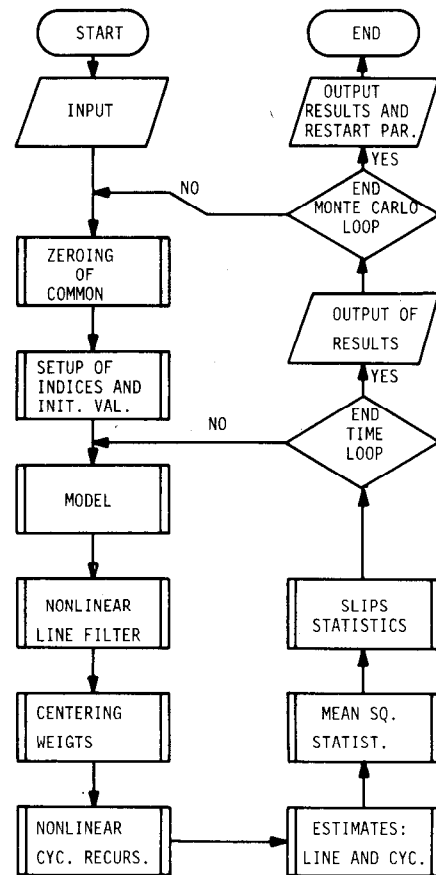


Fig. 1. Program fluxogram.

analysis of the length of the transient phase. The latter analysis justifies our sampling rate of 10 samples per time constant.

IV. CONGRUENCE ASSUMPTION: THE ENTROPIC MEASURE OF INSTABILITY

Intuitively one expects the cyclic density J_n to be symmetric about the cyclic estimate. Also, as the conditional distribution diffuses outward on the line, lobes are created in adjacent cycle intervals. If the slip rate is slow compared to the filter time constant, as occurs at moderately good signal-to-noise ratios, the distributions in the several cycles can be thought of as essentially redundant solutions to the same dynamics problems, namely the cyclic problem, but with weak diffusion coupling between solutions in adjacent cycles. This observation suggests that, for any finite time, the *a posteriori* distribution must be a pseudocyclic function, as illustrated in Fig. 2, whose individual 2π cycles are similar to the cyclic probability J_n , modulated by an overall envelope corresponding to a general outward diffusion of probability with time.

In [1], this pseudoperiodic argument was used to factor $P_n(x)$ as

$$P_n(x) \sim E_n(x)J_n(\tilde{x}), \quad \tilde{x} \in [-\pi, \pi), \quad (19)$$

where J_n is the cyclic density in (13), and $E_n(x)$ is an envelope function on the line. It is stressed that (19) was used as an algorithmic approximation to the *a posteriori* conditional distribution.

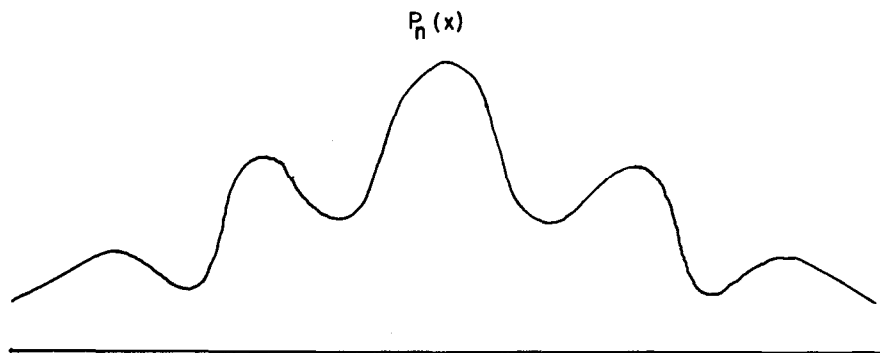


Fig. 2. Pseudoperiodic conditional density.

The simplest envelope assumed is of the histogram type, i.e., the envelope is constant on the K th cycle, having weight

$$W_k^*(n) \triangleq E_n(x), \quad x \in K \text{th cycle.} \quad (20)$$

The weights $\{W_k^*(n)\}$ are equivalently defined as the total mass on the K th cycle. In [1], exact recursive updating formulas for the $\{W_k^*(n)\}$ are presented. Their actual implementation requires approximation, (19) being the one suggested.

Herein (19) is used as a strong signal or “linearized” behavior assumption for the conditional distribution. It is referred to as the congruence assumption.

Congruence Assumption (CA): The *a posteriori* distribution $P_n(x)$ is said to satisfy the congruence assumption (CA) if and only if $P_n(x)$ satisfies (19), where $J_n(\tilde{x})$ is the cyclic conditional density and $E_n(x)$ is a suitably defined envelope on the line.

In the sequel, whenever the congruence assumption is invoked, the histogram envelope of (20) is implied. The congruent line probability is then a train of differently weighted replicas of the cyclic conditional probability $J_n(x)$. In terms of notation, every quantity related to a congruent probability will be tilded, e.g., $\tilde{P}_n(x)$.

As mentioned in the Introduction, a useful functional parameter to study the degree of randomness of the conditional distribution is its (differential) entropy,

$$H_n \triangleq - \sum_{l=1}^L \Delta_M P_n(x_l) \log P_n(x_l), \quad (21)$$

where Δ_M is the space mesh.

The entropy of the message process, being the entropy of a Brownian motion (BM) with variance diffusion coefficient q , is (in continuous time)

$$H_t(\text{BM}) = \frac{1}{2} [1 + \log(2\pi qt)]. \quad (22)$$

Experimental studies show a reduction of at least a factor of 10 for the entropy of the line density, just as a result of observations.

Second, the entropy of any distribution over $K 2\pi$ -cycles is bounded above the the entropy of the uniform distribution (UD)

$$H(\text{UD}) = \log(K2\pi). \quad (23)$$

Whenever the entropy of the line distribution (LD) is greater than (23), i.e.,

$$H_n(\text{LD}) \geq \log(K2\pi), \quad (24)$$

then it is plausible to assert that any reasonable decision scheme will have erred up to time n at least K times when deciding which cycle the process is in, in other words, it will have experienced at least K cycle slips. For example, with the uniform distribution, a randomized strategy will be wrong $(K - 1)/K$ percent of the time. For n sufficiently large,

$$\frac{K - 1}{K} n \geq K. \quad (25)$$

For a congruent distribution (linearized behavior) satisfying (24), the weight distribution may be quite uneven. However, the weights are spread over a larger number of cycles, so that it is still reasonable to take K as a conservative estimate of the number of cycles slipped.

It is straightforward to show for a congruent distribution that

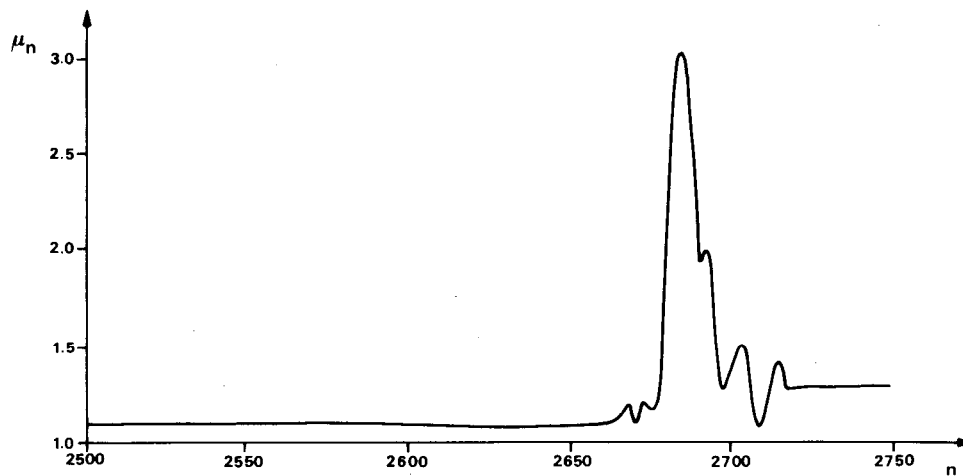
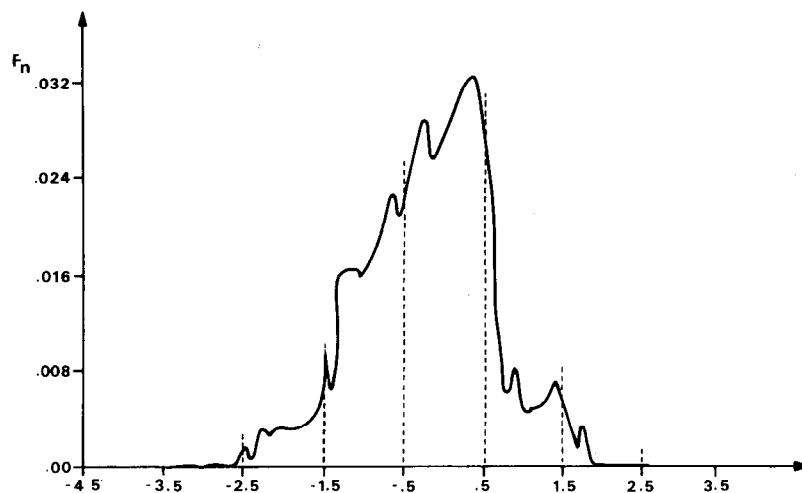
$$\tilde{H}_n = \tilde{H}_n(w) + \tilde{H}_n(J), \quad (26)$$

where $\tilde{H}_n(w)$ is the entropy of the distribution of weights $\{W_k(n)\}$ and $\tilde{H}_n(J)$ is the entropy of the cyclic probability density. As will be seen from experimental studies presented in the following section, the actual form of $P_n(x)$ may depart significantly from the congruent format. To quantify in a contracted way this departure, the following entropic measure of instability is found to be useful:

$$\mu_n \triangleq H_n - \tilde{H}_n, \quad (27)$$

where \tilde{H}_n is the entropy of the congruent distribution $\tilde{P}_n(x)$ corresponding to the conditional PDF $P_n(x)$. This congruent distribution is constructed in the obvious way, by overlapping $P_n(x)$ on a given centered 2π -cycle and replicating this in adjacent intervals, with weights corresponding to the total mass in these intervals. The next section shows how μ_n serves a useful device to predict the smooth or turbulent behavior of the conditional density.

In [8], [9], [10] the joint information was used to *lower bound* the average error performance of systems for given processes. In this paper, in contrast, we are using information concepts applied to the conditional density to signal and detect anomalous behavior along a given sample path, in particular to signal cycle slips.

Fig. 3(a). Q entropic measure history.Fig. 3(b). Density display at $n = 2860$.

V. BEHAVIOR OF THE CONDITIONAL LINE PROBABILITY DISTRIBUTION

The line recursion (17) was implemented as described in Section II. Besides the estimators and statistics presented in Section VI, the output included the history of the entropic measure of instability $\mu_t = H - DH$ introduced in last section and in graphic displays of the density. These were obtained with a Chromatics High Resolution color-graphics output device to the PDP-11-55 minicomputer/API20B array processor configuration.

For a given run, Fig. 3(a), illustrates the trajectory of μ_t from iteration number 2500 to iteration number 2750. One distinguishes two types of behavior: a) where μ_t exhibits small values, remaining in a neighborhood of zero; b) where μ_t takes large values. About iteration number 2680, μ_t exhibits drastic variations. At this iteration, the conditional line density F_n is given by Fig. 3(b). It departs significantly from the congruent assumption. This pattern was found to be present in all the runs simulated. The trajectory of μ_t is a good indication of mass flow between adjacent modes. It will be concluded later on, that it is at these critical points that slips occur. However, the estimators are able to survive without slipping 90–95 percent of the time.

Having the μ_t -history as a guideline, sequences of snapshots of this line density were taken in regions corresponding to the two aforementioned modes of μ_t behavior. They are shown in Figs. 4 to 12. On the basis of these and similar experiments carried out, it is concluded that the evolution of the conditional line distribution exhibits two basic modes:

- 1) A strong signal mode or weak-interaction mode, where the line distribution is very nearly congruent, with similar, weakly interacting lobes (sequence of frames in Figs. 4 to 12). In this mode, the line distribution behaves in a smooth or undisturbed fashion, with weak diffusion coupling between adjacent cycles.
- 2) A weak signal mode or strong-interaction mode, where the line distribution departs significantly from the congruent assumption, with strong coupling between lobes and mass flow, or lobes spanning several adjacent 2π cycles (sequence of pictures in Figs. 4 to 12). In this mode, the line distribution goes through a turbulent state, with the creation, annihilation, or exchange of significant mass between cycle lobes. These pictures confirm that the entropic measure of instability μ_t is a good indicator of the dynamical stability of the conditional distribution in evolution.

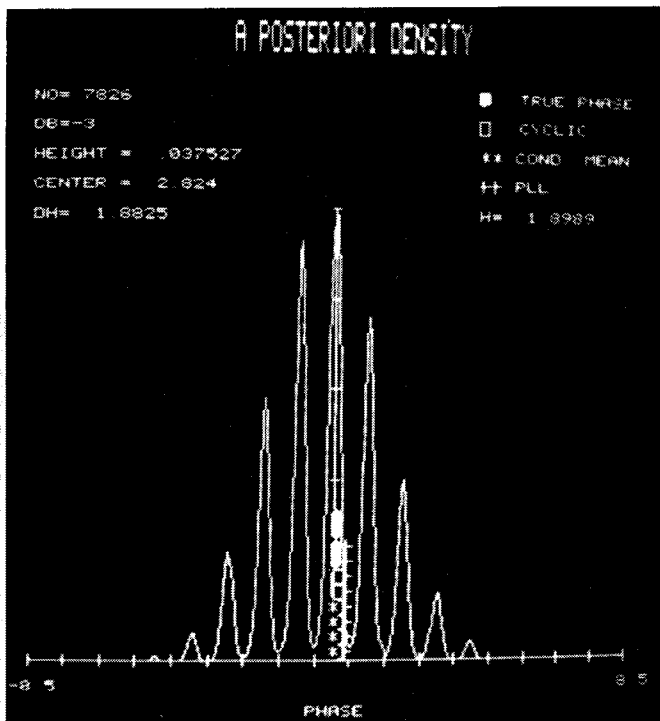


Fig. 4. Phase density 7826.

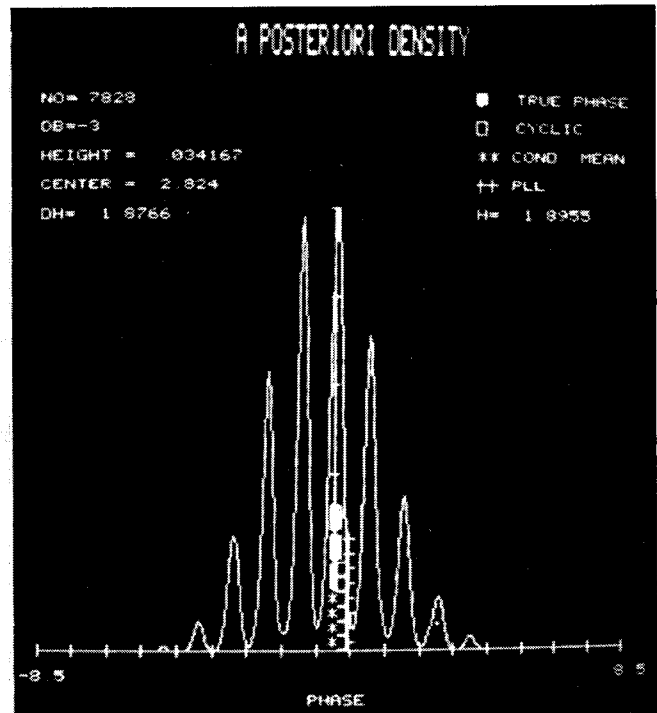


Fig. 5. Phase density 7828.

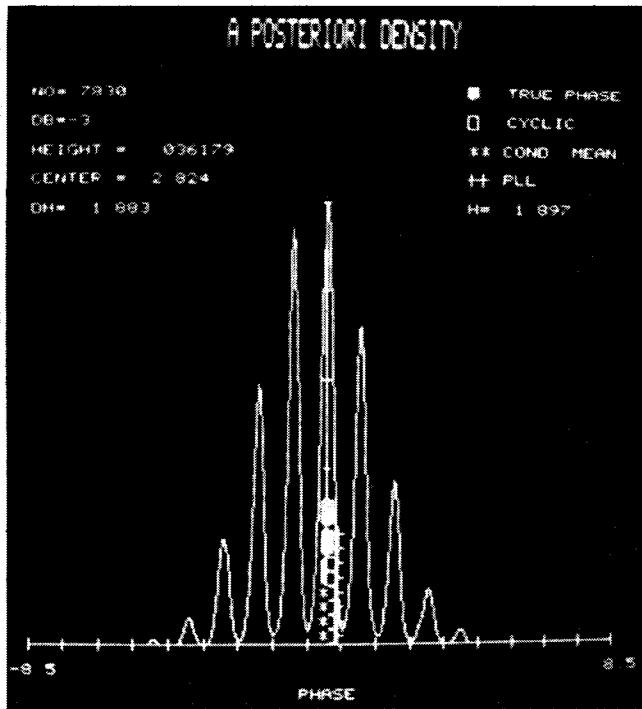


Fig. 6. Phase density 7830.

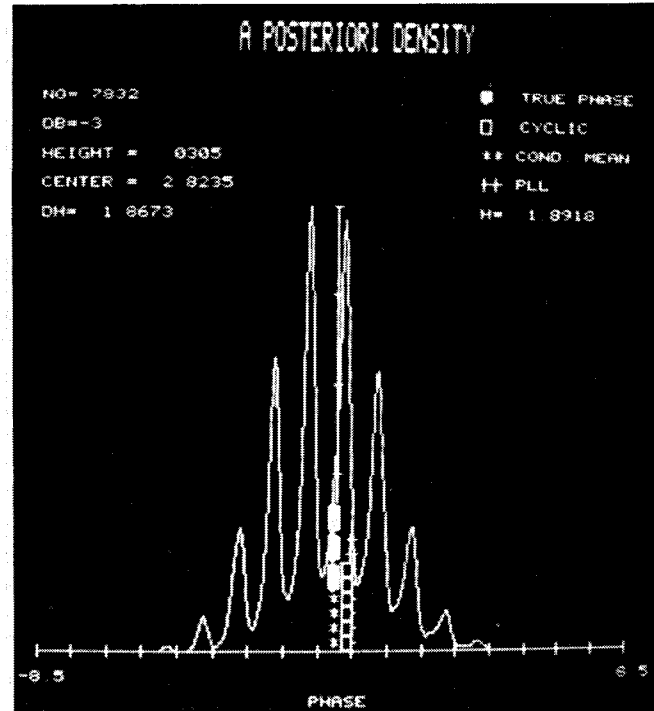


Fig. 7. Phase density 7832.

Note the relation between an information theoretic concept and the behavior of nonlinear filters. The difficulty lies in the fact that computation of μ_i still requires the propagation of the full line density, as the weights W_k^* must be determined.

Figures 4-12 also show, at the arrows, the location of several estimator structures (PLL, line conditional mean, unwrapped cyclic). Although a detailed discussion on these

estimators and in particular of how to decide if a slip occurred or not is postponed to next section, a link is here established between the behavior of the density flow and the cycle slips. It was observed that the cycle slipping is intimately related to the turbulent distribution mode, i.e., where there is oscillation of the entropy difference. After the occurrence of a noisy sequence of data points, when the pattern of density modes changes, the estimators may end up locked at a wrong lobe. It is this phenomenon that

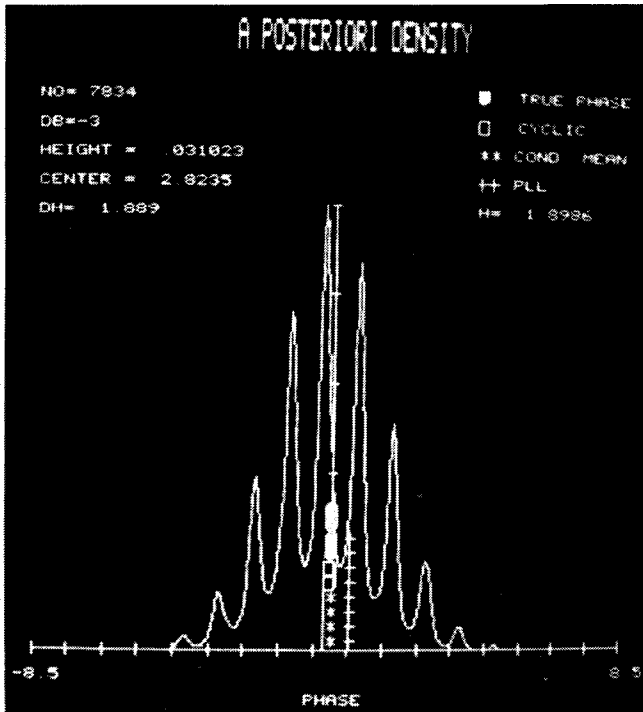


Fig. 8. Phase density 7834.

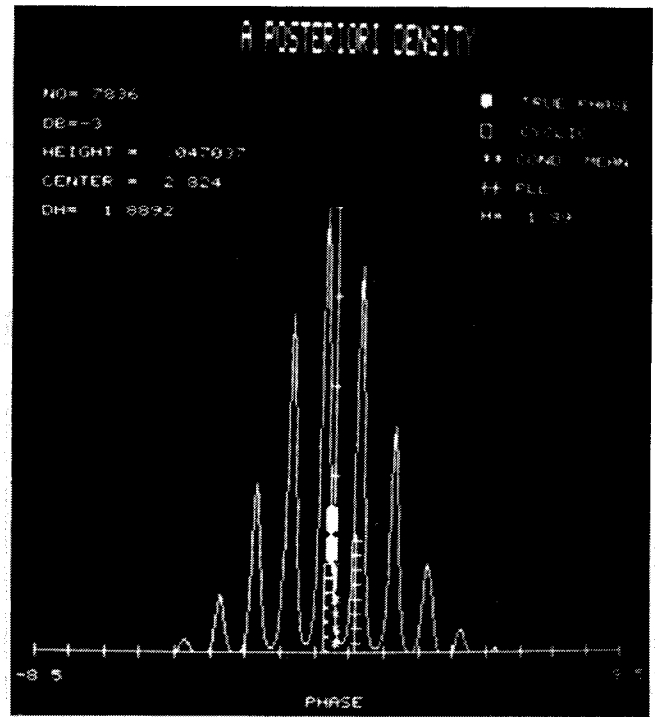


Fig. 9. Phase density 7836.

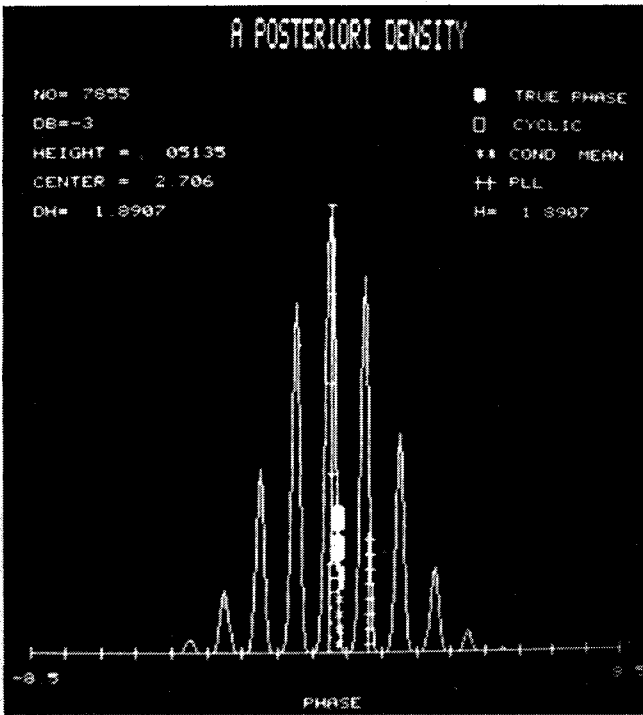


Fig. 10. Phase density 7855.

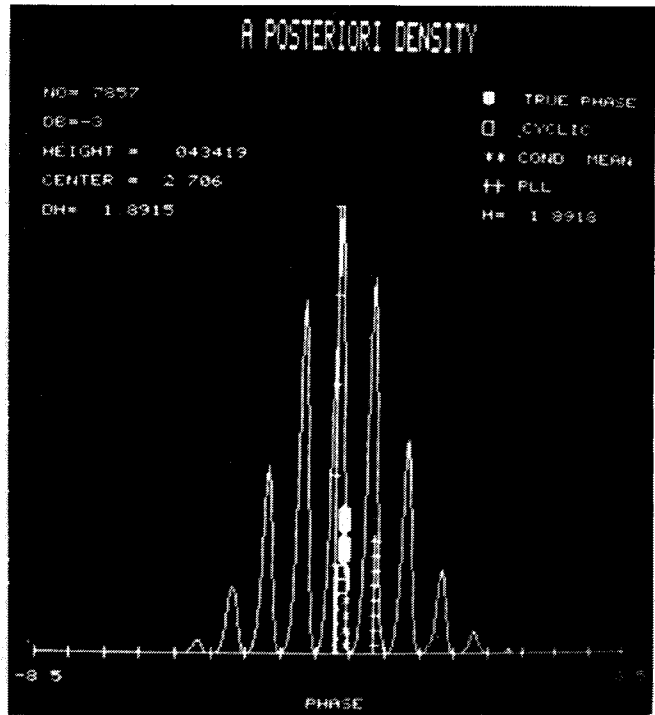


Fig. 11. Phase density 7857.

corresponds to a cycle slip. However, this last phenomenon has small probability and finer resolution between mass flow and slips must be considered.

As long as successive unstable or turbulent modes occur separated in time by several PLL time constants, the PLL slips, under these turbulent modes, constitute events independent of each other; i.e., if the time between sequences of noisy data is more than the loop memory span, the PLL

slips are independent. On the other hand, for the line conditional mean, as for other estimators directly using the line density, one cannot assume this independence. An open question is whether bonding or correlation is sufficiently strong to eventually lead to a corrective movement of the slips. A second related open question is concerned with the quantification of the probability of a slip in terms of μ_r .

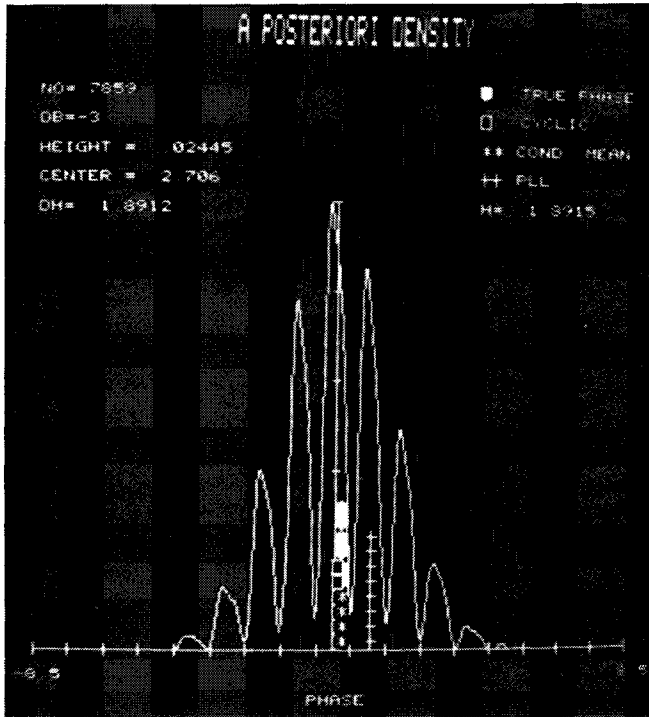


Fig. 12. Phase density 7859.

As an indicator of slips the difference of the entropies of the conditional distribution on the line and the corresponding congruent distribution (i.e., based on the cyclic conditional distribution) gives rise to a large number of false alarms. False alarms are situations where mass flow occurs but the estimate survives the mass flow without slipping either to the right or to the left. An indicator of slips which appears to be as good as can be expected, about a 50 percent false alarm probability, can be constructed by indicating a slip when the difference of the line and congruent entropies exceeds a level while simultaneously the entropy of the conditional density on the line at the present instant exceeds that of the previous instant by a fixed amount. At +3 dB signal-to-noise ratio, $q_D = 0.1$ the slips generated by the first random sequence of the random number generator of [6] with levels of 0.05 predicted slips well for the first 25 000 estimates. The interested reader can reproduce these results by using the generator of [6]. Note that because of our representation of the conditional density on the line viz. 17 2π intervals, our simulation is unable to accurately simulate beyond 25 000 estimates because we lose mass from the finite support.

VI. CYCLE SLIP STATISTICS AND PERFORMANCE BEHAVIOR OF ESTIMATORS

Given a loss function appropriate to the problem at hand, and having the conditional distribution, it is trivial to derive the nonlinear minimum loss estimate \hat{x} . In [1] a hybrid minimum cyclic loss/maximum likelihood criterion was chosen, yielding an estimate $x^* + n2\pi$ where x^* is the minimum cyclic loss estimate and n is the cycle having the largest mass. Here, attention is focussed on three estimates: the PLL, the conditional mean, and the replica of x^* , the

cyclic estimate, unwound about the conditional mean. They are compared on the basis of the average of the accumulated mean square error and of the number of cycle slips observed in a given time span.

A first issue to be resolved concerns counting slips. Decomposing the absolute phase estimation error for estimate \hat{x} as in (15), here repeated

$$\epsilon_t = \tilde{\epsilon}_t + 2\pi N(t), \quad \tilde{\epsilon}_t \in [-\pi, \pi), \quad (28)$$

one might be tempted to count the jumps of $N(t)$ as the cycle slips. However, this is found to be unsatisfactory, as mean time to slip data for the PLL cannot be verified. It is found experimentally that the error process occasionally “loiters” in the vicinity of the unstable equilibrium points $(2n + 1)\pi$ until resolved one way or the other by the dynamics or the data. Fig. 13 illustrates the point. It seems clear in the case illustrated that it is incorrect to count 8 cycle slips due to boundary crossings as shown.

A correct counting of the cycle slips considers the recent past history of the error process. This is achieved by introducing a decision scheme with hysteresis [1], resembling devices used in practice when counting zero crossings of a Brownian motion. This device increases its count only when the error process penetrates sufficiently into one of the adjacent intervals, as illustrated in Fig. 14.

What it does is to adopt a hysteresis threshold definition of “cycle slips.” It is apparent that the tendency of the conditional mean to be conservatively positioned between two main lobes of the conditional distribution reduces its total number of cycle slips.

There is an error of k positive slips at t if and only if

- there exists an s , $s < t$ such that $\epsilon(s) \in \Omega_k = (2\pi k - L, 2\pi k + L)$;
- the first exit time τ of Ω_k is such that $s < \tau < t$ and $\epsilon(\tau) = 2\pi k + L$;
- for every $\mu \in (\tau, t)$, $s(\mu) \in \Omega_{k+1}$.

Similarly, for k negative slips. The above definition only counts a positive slip at the first positive crossing of the level $(2k + 1)\pi + L$ following a positive crossing of the level $(2k + 1)\pi$. The decision threshold L is tuned by adjustment of the PLL experimental and theoretical results.

In order to find slip histograms with sufficiently small confidence intervals, a reasonable number of slips have to be generated in each run. This can be achieved either by lowering the signal-to-noise ratio or by extending the time duration of the run. The simulation results presented correspond to 20 Monte Carlo runs, with the following choice of parameters:

- α = signal-to-noise ratio = +3 dB,
- $k = M = 17$;
- N = number of iterations
= 25×10^3 (2500 PLL time constants),
- $L = \pi$.

For each estimate a running count was kept of the number of slips observed, as well as of the time average accu-

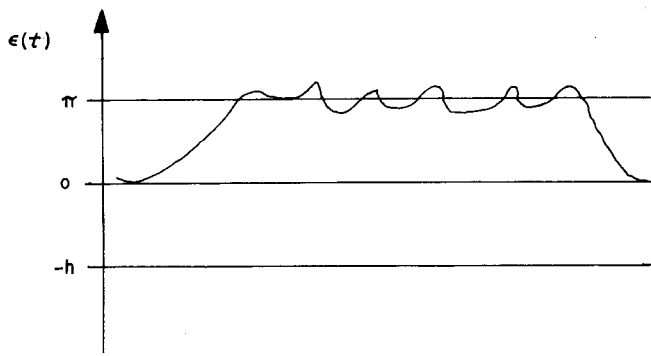


Fig. 13. Error process "loitering" about unstable points.

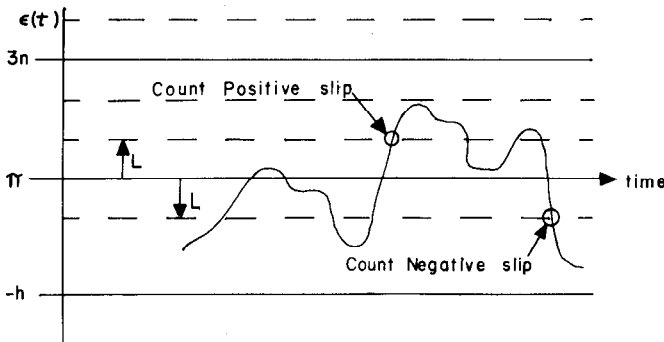


Fig. 14. Counting of cycle slips with hysteresis device.

culated mean-square error

$$\langle \epsilon^2 \rangle = \frac{1}{N} \sum_{i=1}^N (x(i) - \hat{x}(i))^2.$$

The results are reported in Tables I–III.

It is apparent that the tendency of the conditional mean to be conservatively positioned between two main lobes of the conditional PDF penalizes its performance as a (local) cyclic estimator. However, it reduces its total number of cycle slips. The unwound cyclic about the conditional mean, being a noncontinuous estimate, exhibits from time to time additional pseudo slips. The hysteresis device still counts these as slips, but, because they correspond to a transient behavior, they are eventually corrected, having no significant lasting impact on the mean square error performance.

These simulations show that, by carrying forward in time the conditional density, a significant improvement with respect to the PLL can be achieved in terms of reduction of the number of cycle slips observed and of slowing down the growth of the mean square error. Preliminary results, lacking statistical significance, indicate that suboptimal structures with limited memory can achieve most of these improvements.

In Figs. 15 and 16, which represent the total number of times the signal and the estimate differed by $2k\pi$ on 23 sample functions each of duration 25 000 points. The histograms are for the PLL estimate and the cyclic estimate centered at the conditional mean. The results of Figs. 15 and 16 are the results of the same experiment with the two different seeds of the random number generator reported in [6].

TABLE I
SLIP DATA FIRST SAMPLE PATH

Phase Locked				
-728	2741	-7862	-8854	-9641
-10120	-13948	-15140	17898	22517
23199	24171			
Conditional Mean Centered Cyclic Slip Place Data				
2680	3633	-3642	-6177	6346
-6351	6420	-7826	-8824	-9637
12613	-12685	13174	-13222	13237
-13242	13845	-13857	-15093	15097
-15239	15243	-15294	15394	-15344
15370	-15940	15955	-17502	17503
21331	23929	-23940	24072	-24098
24167				

TABLE II
SUMMARY OF AVERAGE STATISTICS FOR PLL AND X MEAN

	PLL	X MEAN	GAIN OVER PLL
$\langle \epsilon^2 \rangle$	2.264	1.796	22%
$\langle \text{slips} \rangle$	11.8	8.5	28%
$\langle \text{iter} \rangle$	1838	2761	50%

TABLE III
AVERAGE NUMBER OF ITERATIONS BETWEEN SLIP $i - 1$ AND i : $\langle \text{iter} \rangle_i$

i	PLL	X MEAN
1	2222	2700
2	2378	2792
3	1969	2495
4	2078	3316
5	1996	3502
6	1848	2220
7	1678	2303
8	2045	—
9	1955	—
10	2016	—

In Table I, we give the original slip place data. An entry $+$ ($-$) k means at time k the estimate slipped one interval positive (negative). There were 240 places on the first sample function of the first Monte Carlo where there was significant mass flow (i.e., the relative minimum of the conditional density in a 4π interval centered at the maximum was greater than 10 percent of the maximum). Notice that in the histogram the PLL estimate accumulates more mass in the tails of the slip distribution. The Q statistic, the rate of growth of slip variance, is consequently larger for the PLL than for the optimal estimate. It is clear from the slip data that the conditional mean centered cyclic estimate is more sensitive and more precise than the phase locked loop estimator.

For the PLL, Lindsey [7] has shown that the slip process $N(t)$ is approximately decomposable as

$$N(t) = N_+(t) - N_-(t),$$

where $N_+(t)$ and $N_-(t)$ are independent Poisson processes, with the same parameter λ in our case so that

$$EN_+(t) = EN_-(t) = \lambda t,$$

$$EN_+^2(t) = EN_-^2(t) = (\lambda t)^2 + \lambda t,$$

and hence

$$E(N_+(t) - N_-(t))^2 = 2(\lambda t),$$

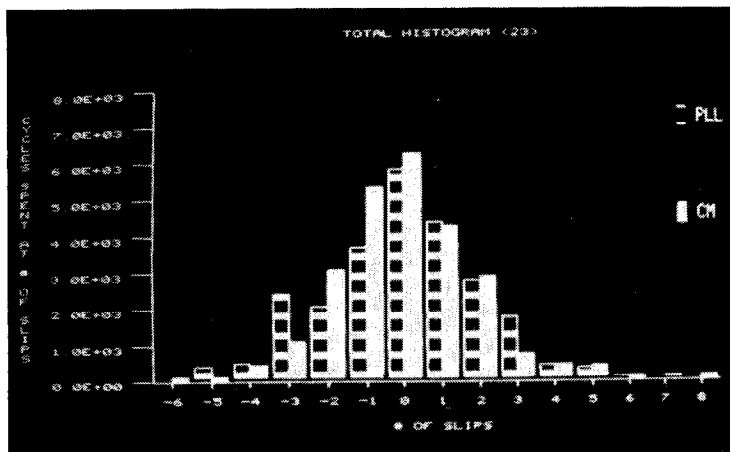


Fig. 15. Slip histogram (seed 1).

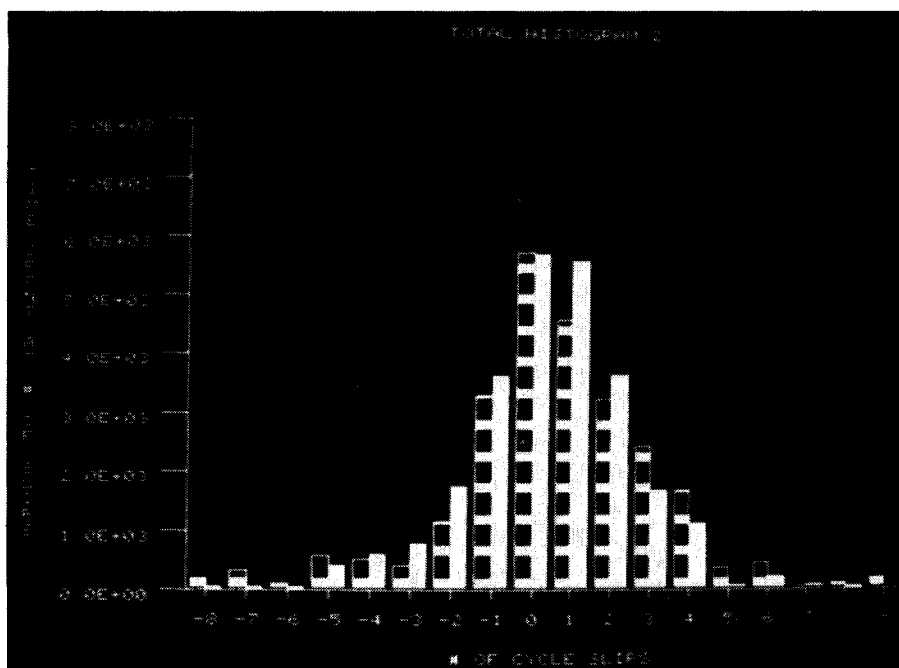


Fig. 16. Slip histogram (seed 2).

where λ is a function of the signal-to-noise ratio. We tested our slip detection process by comparing the simulated average number of slips for the PLL with the above approximation.

This situation leads to a simple idea for a statistic to judge the difference between various estimators of absolute phase, namely λ the variance growth parameter. For example, in the case of the PLL the histogram of the total number of slips determines a variance which grows quadratically with time. From our slip histograms we can compute the histogram variance of the PLL as $EN^2(t/\tau_{PLL}) = 0.005 t/\tau_{PLL}$. It is clear that if $V(t)$ is a discrete random process with variance proportional to t and mean zero, the

$$\frac{1}{n} \sum_{k=1}^n V^2(k\Delta) = \sum_{\alpha} \frac{n_{\alpha}}{n} \alpha^2,$$

where N_{α} = the number of k such that $V(k\Delta) = \alpha$. From the above equality it is clear that the histogram variance

in the case of the PLL measures λ . Our data is consistent with the nonlinear absolute phase estimator generating a slip process $N^*(t)$ analogous to $N(t)$ in distribution and with growth parameter $EN^*(t/\tau_{PLL})^2 = 0.00324 t/\tau_{PLL}$. It is clear that the nonlinear estimator is significantly better than the PLL on the basis of variance growth parameter λ . This is the Q statistic referred to earlier. Analysis of our slip data by computing the variance of the histograms for the PLL and nonlinear estimator by variation of sample path lengths 8000, 16 000 and 24 000 with 23 Monte Carlo runs shows that the variances of the histograms for the nonlinear filter grow linearly with time, at least at -3 dB signal-to-noise ratio.

For each estimate a running count was kept of the slips observed as well as of the time-average accumulated mean square error (normalized by the cycle length of 2π)

$$\langle \epsilon^2 \rangle = \frac{1}{N} \sum_{i=1}^N (x(i) - \hat{x}(i))^2.$$

Table II summarizes the results for the PLL and the line conditional mean X MEAN. In the table

$\langle \text{slips} \rangle =$ average number of slips observed,

$$\langle \text{iter} \rangle = \frac{1}{I} \sum_{i=1}^I \langle \text{iter} \rangle_i,$$

where $\langle \text{iter} \rangle_i =$ average number of iterations between slip $i - 1$ and slip i . The above averages are computed over the 23 Monte Carlo runs. The third column in Table II represents the percentage improvement of X MEAN over the PLL.

Table III gives the values of $\langle \text{iter} \rangle_i$. Notice that, for an equivalent statistical significance, the sequence of available values for X MEAN is smaller than for the PLL.

Table II shows a net improvement of X MEAN over the PLL for the three statistics studied. As a final remark, notice that the value $\langle \text{iter} \rangle = 1838$ in Table II for the PLL is within 11 percent of the theoretical value of 2070, given for $\text{SNR} = -3$ dB by the PLL theory [7]. This serves to adjust the threshold of the hysteresis device used in conjunction with the slip counting.

VII. CONCLUSION

This paper has considered the absolute phase demodulation (APD) problem. It presents experimental evidence based on Monte Carlo simulations run on a digital computer, that links the cycle slipping of several estimates to the turbulent behavior induced by noisy data sequences on the conditional distribution. This turbulent behavior is measured in terms of the distortion of the PDF with respect to its congruent version, being correctly signaled by the entropic measure of instability. This measure presents an important connection between information theory concepts and nonlinear filtering. Statistics on the cycle slips and the absolute mean-square error performance are presented for several estimators, comparing them with respect to the corresponding results for the classical phase locked loop. Even for the one-dimensional phase problem under study, with limited improvement to be expected, significant improvements are reported. Of course, the more realistic model of the phase process where the acceleration is white

and the phase velocity as well as phase must be tracked, merits study. From our discussions of the computational requirements of an accurate numerical and statistical study, it is clear that such a study would be quite demanding computationally. Further studies, presently being carried out, are concerned with the statistical performance of suboptimal finite memory filtering structures.

ACKNOWLEDGMENT

The authors want to thank Faramesh Ghavanlou, Jim Hiroshiga, and Shahen Voskanian, graduate students of Aerospace Engineering, University of Southern California, Los Angeles, CA, for extensive help with graphics and computations associated with the results reported in this paper. We also want to thank Dr. Ken Senne of MIT Lincoln Laboratory, Cambridge, MA, for helpful suggestions and valuable discussions.

REFERENCES

- [1] R. S. Bucy, J. F. Moura, A. J. Mallinckrodt, "Absolute phase demodulation," *Mini and Micro Computers*, vol. 5, no. 2, pp. 116-119; 1980.
- [2] R. S. Bucy and P. D. Joseph, *Filtering for Stochastic Processes with Applications to Guidance*. Tracts in Pure and Applied Math. New York: Interscience, 1968.
- [3] R. S. Bucy, K. D. Senne, and C. Hecht, "New methods for nonlinear filtering," *Francais d'Automatique Informatique, Recherche Operational* Dunod, Paris, pp. 3-54, Feb. 1973.
- [4] R. S. Bucy and K. D. Senne, "Nonlinear filtering algorithms for vector processing machines," *Math. Comput.*, vol. 6, pp. 317-388; 1980.
- [5] F. J. Charles and W. C. Lindsey, "Some analytical and experimental phase lock loop results for low signal-to-noise ratios," in *Proc. IEEE*, vol. 54, pp. 1152-1166; 1966.
- [6] K. D. Senne, "A machine independent random number generator," *Stochastics*, vol. 1, pp. 1-24; 1973.
- [7] W. C. Lindsey, *Synchronization Systems*. Edgewood Cliffs, NJ: Prentice Hall, 1978.
- [8] R. S. Bucy, "Filtering and information," *Inform. Sci.*, vol. 17, pp. 179-187, 1979.
- [9] J. I. Galdos, "A low bound on filtering error with application to phase demodulation," *IEEE Trans. Inform. Theory*, vol. IT-25, no. 4, pp. 452-463, 1979.
- [10] H. L. Weidmann and E. B. Stear, "Entropy analysis of estimating systems," *IEEE Trans. Inform. Theory*, vol. IT-16, no. 3, pp. 264-270, 1970.