Estimation for Rotational Processes with One Degree of Freedom—Part II: Discrete-Time Processes

ALAN S. WILLSKY, MEMBER, IEEE, AND JAMES TING-HO LO, MEMBER, IEEE

***Abstract*—General error criteria and probability distributions on the circle are studied in connection with estimation by using their Fourier series representations. Conditional probability densities for certain discrete-time folded normal processes, which are analogous to the continuous-time processes associated with the bilinear problems considered in Part I of this series, are computed. An intrinsic physical difference between the discrete-time and continuous-time problems is discussed, and the complexity of the estimation equations in the discrete-time case is analyzed in this setting. Suboptimal sequential filtering schemes are briefly discussed. In addition, Fourier analysis of conditional probability distributions exposes the inherent rich structure in quite general classes of estimation problems on the circle.**

I. INTRODUCTION

IN PART I [1] of this series of papers, we introduced and studied a class of continuous-time estimation problems on the circle S^1. For a certain class of bilinear problems, we were able to derive easily implemented optimal estimation equations. Extensions of these results to arbitrary Abelian Lie groups were presented and were shown to include the vector space Kalman–Bucy filtering results as a special case. As is well known [6]–[8], in the vector space case, the discrete-time optimal linear filter looks much like its continuous-time counterpart. This is due in part to the fact that in the linear–Gaussian case, the relevant conditional densities are Gaussian, *independent* of the continuous or discrete nature of the observations.

It is one of the main points of this paper that the continuous-time results presented in [1] for the circle and arbitrary Abelian Lie groups do *not* extend to the discrete-time case in nearly as nice a manner as in the linear case, and thus we have a striking example of a class of estimation problems for which the continuous-time solution is *much simpler* than its discrete-time analog. However, the reason for the difficulties is easy to interpret physically, and one can use this interpretation to devise suboptimal estimation schemes.

As we shall see, although the discrete-time S^1 estimation problem is considerably more complicated than its continuous-time counterpart, the discrete equations display a great deal of structure that can be utilized to aid in their analysis. This is evident not only in the physically appealing interpretation one has for the discretized bilinear estimation problem, but also in the success one has in utilizing Fourier series techniques to devise estimation methods. Bucy and his associates [9], [10] have studied several specific problems using Fourier series, and the results presented here and in [2]–[4] contain some of their results as special cases (see [3] and [4] for some continuous-time and other discrete-time Fourier series results).

In the next section we investigate the estimation of random variables on the circle with the aid of Fourier series, and in Section III we discuss a discrete-time analog of the bilinear estimation problem discussed in [1]. Section IV deals with the use of Fourier series to solve general discrete-time S^1 estimation problems. We remark that further discussions of the concepts developed in this paper are contained in [2] and [3].

As in [1], we will use several representations of S^1 interchangeably. Referring to the discussion in [1], we will use the $\theta \in [-\pi,\pi)$ representation to define all probability densities on S^1. In addition, the considerations in Section III of this paper are related to another S^1 representation. There exists a natural projection from R^1 to S^1, here identified with $[-\pi,\pi)$:

$$q(x) = x \bmod 2\pi. \tag{1}$$

Two points x_1 and x_2 are projected onto the same point if and only if they differ by an integral multiple of 2π. Also, since $q(x_1+x_2)=[q(x_1)+q(x_2)] \bmod 2\pi$, q is a homomorphism of R^1 into S^1 with kernel

$$\ker q = \{2n\pi \mid n \in Z\}. \tag{2}$$

Clearly, the range of q is all of $[-\pi,\pi)$. Thus, by the First Isomorphism Theorem for groups [11]

Manuscript received July 30, 1973. Paper recommended by H. W. Sorenson, Chairman of the IEEE S-CS Estimation and Identification Committee. The work of A. S. Willsky was performed in part while the author was a Fannie and John Hertz Foundation Fellow in the Department of Aeronautics and Astronautics, Massachusetts Institute of Technology, Cambridge, Mass., and in part at the Decision and Control Sciences Group, M. I. T. Electronics Systems Laboratory, Cambridge, Mass., under partial support from the AFOSR under Grant 72-2273. The work of J. T.-H. Lo was supported in part by the U. S. Office of Naval Research under the Joint Services Electronics Program by Contract N00014-67-A-0298-0006 while the author was with the Division of Engineering and Applied Physics, Harvard University, Cambridge, Mass.

A. S. Willsky is with the Department of Electrical Engineering, Massachusetts Institute of Technology, Cambridge, Mass. 02139.

J. T.-H. Lo is with the Division of Mathematics and Physics, University of Maryland, Baltimore County, Baltimore, Md. 21228.

$$S^1 \cong R^1/\ker q \triangleq R^1/2\pi Z \tag{3}$$

and any point $\theta \in [-\pi, \pi)$ is identified with the equivalence class $\{\theta + 2n\pi | n \in Z\}$ (we note that since $\ker q$ is a closed subgroup of R^1, $R^1/\ker q$ is a Lie group (see [12]), and one can show that the isomorphism (3) is a Lie group isomorphism—i.e., it is smooth). The $\bmod 2\pi$ equivalence of all elements of $\{x + 2n\pi | n \in Z\}$ will be important in Section III.

II. Fourier Analysis of Probability Distributions, Error Criteria, and Optimal Estimation

Let θ be a random variable on S^1 (here we will identify S^1 with $[-\pi,\pi)$) with probability density $p(\theta)$. Given an error function ϕ, the estimation problem is to choose $\tilde{\theta} \in [-\pi, \pi)$ to minimize

$$J(\tilde{\theta}) = \mathcal{E}\big(\phi(\theta - \tilde{\theta})\big) = \int_{-\pi}^{\pi} \phi(\theta - \tilde{\theta}) p(\theta)\, d\theta. \tag{4}$$

We now will write down necessary conditions for $\tilde{\theta}$ to minimize $J(\tilde{\theta})$. We assume that we have $p(\theta)$ and $\phi(\theta)$ in Fourier series form:

$$p(\theta) = \frac{1}{2\pi} + \sum_{n=1}^{\infty} [a_n \sin n\theta + b_n \cos n\theta] \tag{5}$$

$$\phi(\theta) = d_0 + \sum_{n=1}^{\infty} [c_n \sin n\theta + d_n \cos n\theta] \tag{6}$$

where the Fourier coefficients are computed in the usual way [17]. Note that we have the interpretation

$$\pi a_n = \mathcal{E}(\sin n\theta), \qquad \pi b_n = \mathcal{E}(\cos n\theta). \tag{7}$$

We will have more to say about the Fourier series decomposition of p in Section IV.

A simple computation yields

$$J(\tilde{\theta}) = d_0 + \pi \sum_{n=1}^{\infty} \Big[a_n\big(c_n \cos n\tilde{\theta} + d_n \sin n\tilde{\theta}\big) + b_n\big(d_n \cos n\tilde{\theta} - c_n \sin n\tilde{\theta}\big)\Big] \tag{8}$$

and necessary conditions for a local minimum are

$$\frac{d}{d\tilde{\theta}} J(\tilde{\theta}) = 0 \qquad \frac{d^2}{d\tilde{\theta}^2} J(\tilde{\theta}) \geqslant 0. \tag{9}$$

Explicit solution of (9) is possible only for certain error functions. Note that the sums in (8) and (9) are finite if the sum in either (5) or (6) is. The following example indicates how this finiteness can simplify the problem of finding the optimal estimate.

Example 1: Consider the error function

$$\phi(\theta) = 1 - \cos\theta. \tag{10}$$

This criterion has also been considered in [9] and [10]. For this function we have

$$J(\tilde{\theta}) = 1 - \pi\big(a_1 \sin\tilde{\theta} + b_1 \cos\tilde{\theta}\big) \tag{11}$$

and the necessary conditions are

$$a_1 \cos\tilde{\theta} - b_1 \sin\tilde{\theta} = 0, \qquad a_1 \sin\tilde{\theta} + b_1 \cos\tilde{\theta} \geqslant 0. \tag{12}$$

If $a_1 = b_1 = 0$, $J(\tilde{\theta})$ is independent of $\tilde{\theta}$. In any other case, examination of (12) yields the optimal value $\tilde{\theta}_0$ as the unique solution in $[-\pi, \pi)$ of

$$\sin\tilde{\theta}_0 = \frac{a_1}{\sqrt{a_1^2 + b_1^2}} \qquad \cos\tilde{\theta}_0 = \frac{b_1}{\sqrt{a_1^2 + b_1^2}}. \tag{13}$$

Also,

$$J(\tilde{\theta}_0) = 1 - \pi\sqrt{a_1^2 + b_1^2}\,. \tag{14}$$

We note that this particular error function has some very appealing properties. First of all, since the optimal estimate is an explicit function only of the first mode of the probability distribution (i.e., a_1 and b_1), the computational procedure to determine $\tilde{\theta}_0$ is quite simple. In addition, there is strong physical motivation for using this criterion. First note (see [9]) that for small values of θ

$$1 - \cos\theta \cong \tfrac{1}{2}\theta^2. \tag{15}$$

Thus, this is, at least locally, a type of least squares criterion. In fact, suppose that x_1 and x_2 are real-valued random variables such that

$$x_1^2 + x_2^2 = 1. \tag{16}$$

That is, there exists a random variable θ with

$$x_1 = \sin\theta, \qquad x_2 = \cos\theta. \tag{17}$$

Suppose we wish to choose $\tilde{x}_1$ and $\tilde{x}_2$ to minimize

$$J = \tfrac{1}{2}\mathcal{E}\big[(x_1 - \tilde{x}_1)^2 + (x_2 - \tilde{x}_2)^2\big] \tag{18}$$

subject to the constraint

$$\tilde{x}_1^2 + \tilde{x}_2^2 = 1. \tag{19}$$

That is,

$$\tilde{x}_1 = \sin\tilde{\theta}, \qquad \tilde{x}_2 = \cos\tilde{\theta} \tag{20}$$

and, substituting into (18), we have

$$J = \mathcal{E}\big(1 - \cos(\theta - \tilde{\theta})\big). \tag{21}$$

Thus, this error function is a constrained least squares criterion.

We note that although the higher modes do not affect the estimate directly, we shall see in Section IV that these coefficients have an indirect effect on $\tilde{\theta}$. Specifically, we shall find that in dealing with random processes and time-varying densities, the time rates of change of a_1 and b_1 depend, in general, on the other coefficients.

Another possible error function, one that involves the first and second modes of the density, is

$$\phi(\theta)=(1-\cos\theta)^2=\tfrac{3}{2}-2\cos\theta+\tfrac{1}{2}\cos 2\theta. \quad (22)$$

Using the same type of approach as before, one can reduce the problem of finding the optimal estimate to the solution of a quartic polynomial equation and the calculation of several functions—a procedure that can be done easily by computer. However, the complexity, even when we just add in the second mode, is such that no closed form for the optimal error in terms of the Fourier coefficients is available (see [3] for details).

As can be seen, the error analysis becomes increasingly more difficult as the number of nonzero Fourier coefficients increases. For example, direct application of these ideas if $\phi=\rho$ or ρ^2, where ρ is the Riemannian metric (arc length) on S^1 (actually $\rho(\theta)\triangleq\rho(\theta,0)$—see [1]), yields extremely complicated equations. However, the $1/n^2$ behavior of the Fourier coefficients for these two examples suggests truncating the series and applying techniques such as those used in the analysis for $(1-\cos\theta)$ and $(1-\cos\theta)^2$. For some further analysis for the criterion ρ^2, we refer the reader to [22].

In the next section we will encounter densities of the form

$$p(\theta)=\sum_{n=1}^{\infty}c_nF(\theta;\eta_n,\gamma_n), \qquad \sum_{n=1}^{\infty}c_n=1, \quad \gamma_n>0 \quad (23)$$

where F is the folded normal density

$$F(\theta;\eta,\gamma)=\frac{1}{\sqrt{2\pi\gamma}}\sum_{n=-\infty}^{+\infty}\exp\left(-\frac{(\theta+2n\pi-\eta)^2}{2\gamma}\right) \quad (24)$$

(if x is a normal random variable with mean η and variance γ, then $\theta=x \bmod 2\pi$ has the density $F(\theta;\eta,\gamma)$). It should be noted that it can be shown by using a result analogous to that in [13] that the set of densities given by (23) with only finitely many nonzero c_n's is dense in $L^1(-\pi,\pi)$, and this is still true if all the γ_n's are equal to some fixed γ. We do not require that only finitely many c_n's be unequal to zero. The reason for this will be seen in Section III.

For p given by (23) we will examine the optimal estimation problem for the error function $1-\cos\theta$. Note that p in (23) need *not* be unimodal and symmetric about its mode (as it would be if p=folded normal), and thus the results discussed in [1] do not apply.

As discussed in Example 1, in trying to minimize $\mathcal{E}(1-\cos(\theta-\tilde{\theta}))$ with respect to $\tilde{\theta}$, we need only know the lowest mode Fourier coefficients, a_1 and b_1. If p is as in (23),

$$a_1=\frac{1}{\pi}\sum_{n=1}^{\infty}c_ne^{-\gamma_n/2}\sin\eta_n, \qquad b_1=\frac{1}{\pi}\sum_{n=1}^{\infty}c_ne^{-\gamma_n/2}\cos\eta_n \quad (25)$$

and (assuming a_1 and b_1 are not both zero) the optimal estimate $\tilde{\theta}_0$ is either $\tan^{-1}a_1/b_1$ or $\tan^{-1}a_1/b_1+\pi$, depending upon the signs of a_1 and b_1 [9]. In any case, the optimal cost is given by

$$\mathcal{E}\left(1-\cos\left(\theta-\tilde{\theta}_0\right)\right)=1-\left\{\left[\sum_{n=1}^{\infty}c_ne^{-\gamma_n/2}\sin\eta_n\right]^2+\left[\sum_{n=1}^{\infty}c_ne^{-\gamma_n/2}\cos\eta_n\right]^2\right\}^{1/2}. \quad (26)$$

In general, this optimal error is *not* an increasing function of each of the "variances" γ_n individually. However, if all of the variances equal some value γ, it is easy to see that the optimal error *is* an increasing function of γ.

III. Estimation of Discrete-Time Folded Normal Processes

We wish to examine the problem of estimating a discrete-time random process on S^1 given a series of discrete measurements. One possible model for the signal and measurement processes is a discrete approximation to the continuous signal and measurement processes discussed in [1]. We first approximate the measurement equations (see [1] for details)

$$dz(t)=h(x(t),t)\,dt+r^{1/2}(t)\,dw(t) \quad (27)$$

$$z(t)=(Jz)(t)\triangleq\exp(Rz(t)) \quad (28)$$

by the discrete equations

$$\Delta z_k=z_k-z_{k-1}=h_k(x_k)\Delta t+r_k^{1/2}w_k \quad (29)$$

$$Z_k=\exp(z_kR) \quad (30)$$

where

$$R=\begin{bmatrix}0 & 1\\ -1 & 0\end{bmatrix} \quad (31)$$

and where Δt is the intermeasurement time, $x_k=x(k\Delta t)$, $r_k=r(k\Delta t)$, $h_k(\cdot)=h(\cdot,k\Delta t)$, and $w_k=w(k\Delta t)-w((k-1)\Delta t)$.

We can rewrite the z_k equation as

$$Z_k=Z_{k-1}\exp(\Delta z_kR) \quad (32)$$

and we see that, given $Z_1,\cdots,Z_{k-1}$, the new information contained in Z_k is equivalent to the new information in $Z_{k-1}^{-1}Z_k$. In addition, this information is equivalent (see [1]) to knowledge of

$$\tilde{\Delta} z_k = \Delta z_k \bmod 2\pi, \qquad \tilde{\Delta} z_k \in [-\pi,\pi). \tag{33}$$

It is here that we see a marked difference between the discrete- and continuous-time problems. In the continuous-time case [1], the continuity of the random processes involved results in our knowing $dz(t)$, not just $dz(t) \bmod 2\pi$. However, in the discrete-time problem, the ambiguity associated with our lack of knowledge of the number of rotations that occur in the Δt between measurements is reflected in the fact that our information is $\Delta z_k \bmod 2\pi$ and *not* Δz_k. In some sense, this makes the discrete-time problem more interesting, since this "mod 2π" ambiguity (a type of "cycle-slip" phenomenon) directly affects the form of the conditional distributions we will consider (this is *not* the case in the continuous-time problem [1]).

Motivated by this discussion, we can state precisely a discrete-time analog of the continuous-time problem of [1]. Let x_k be a discrete-time real-valued signal process satisfying

$$x_{k+1} = a(x_k,k) + b^{1/2}(x_k,k)w_k, \qquad x_0 = 0 \tag{34}$$

where $\{w_k\}$ is a unit variance white Gaussian sequence. Consider an observation process $\{\tilde{y}_k\}$ defined by

$$y_k = h(x_k,k) + r_k^{1/2} v_k \tag{35}$$

$$\tilde{y}_k = y_k \bmod 2\pi, \qquad \tilde{y}_k \in [-\pi,\pi) \tag{36}$$

where $\{v_k\}$ is a unit variance Gaussian sequence independent of $\{w_k\}$. The problem is to determine the conditional probability distributions $p_x(x,k|\tilde{Y}_k)$ and $p_\theta(\theta,k|\tilde{Y}_k)$ where $\tilde{Y}_k = \{\tilde{y}_1,\cdots,\tilde{y}_k\}$ and

$$\theta_k = x_k \bmod 2\pi, \qquad \theta_k \in [-\pi,\pi) \tag{37}$$

and to determine optimal estimates with respect to some given criteria. We note that one can interpret θ_k as angular orientation and x_k as total angle swept.

Our attack on this problem can be divided into three parts. First, to gain insight into the problem, we consider the general one-stage case (Section III-A). All of the essential features of the multistage problem are contained here. The approach in Section III-A is intuitive and involves the use of delta functions. The reader is referred to [2] and [3] for the rigorous but tedious measure-theoretic details. In Section III-B we specialize to the linear–Gaussian case, and in Section III-C we consider the two-stage and multistage problems for the linear–Gaussian situation. In Section III-C we also briefly discuss some of the computational problems associated with the multistage results. For a limiting argument relating the continuous-time and discrete-time problems, the reader is referred to [2] and [3], and the details of some suboptimal estimation schemes are contained in [3] and [5]. Also, the reader is referred to [3] and [5] for further related results concerning conditional probabilities.

A. The One-Stage Problem

Let x and v be independent real-valued random variables with *a priori* densities $p_x(\alpha)$ and $p_v(\nu)$. We wish to compute the conditional density $p_{x|\tilde{y}}(\alpha|\beta)$ where

$$y = h(x) + v, \qquad \tilde{y} = y \bmod 2\pi \tag{38}$$

and $h: R^1 \to R^1$ is a measurable function. By the smoothing property of conditional densities [14],

$$p_{x|\tilde{y}}(\alpha|\beta) = \int_{-\infty}^{+\infty} p_{x|\tilde{y},y}(\alpha|\beta,\xi) p_{y|\tilde{y}}(\xi|\beta)\, d\xi. \tag{39}$$

Now $\tilde{y}$ is a deterministic function of y, so the "density" $p_{y|\tilde{y}}$ is given by

$$p_{y|\tilde{y}}(\beta|\xi) = \frac{p_{\tilde{y}|y}(\beta|\xi)p_y(\xi)}{p_{\tilde{y}}(\beta)} = \frac{\delta(\beta - \xi \bmod 2\pi)p_y(\xi)}{p_{\tilde{y}}(\beta)} \tag{40}$$

where δ is the Dirac delta function. Also, since y determines $\tilde{y}$ uniquely, (39) becomes

$$p_{x|\tilde{y}}(\alpha|\beta) = \frac{1}{p_{\tilde{y}}(\beta)} \int_{-\infty}^{+\infty} p_{x|y}(\alpha|\xi)p_y(\xi)\delta(\beta - \xi \bmod 2\pi)\, d\xi. \tag{41}$$

We now must interpret $\delta(\beta - \xi \bmod 2\pi)$ as a function of ξ. Clearly, what we mean by this is

$$\delta(\beta - \xi \bmod 2\pi) = \sum_{n=-\infty}^{+\infty} \delta(\beta + 2n\pi - \xi). \tag{42}$$

Substituting (42) into (41), we have

$$p_{x|\tilde{y}}(\alpha|\beta) = \sum_{n=-\infty}^{+\infty} \frac{p_y(\beta+2n\pi)}{p_{\tilde{y}}(\beta)} p_{x|y}(\alpha|\beta+2n\pi) \tag{43}$$

$$= \sum_{n=-\infty}^{+\infty} \frac{p_{y|x}(\beta+2n\pi|\alpha)p_x(\alpha)}{p_{\tilde{y}}(\beta)} \tag{44}$$

$$p_{y|x}(\beta+2n\pi|\alpha) = p_v(\beta+2n\pi - h(\alpha)) \tag{45}$$

$$p_y(\beta+2n\pi) = \int_{-\infty}^{+\infty} p_{y|x}(\beta+2n\pi|u)p_x(u)\, du \tag{46}$$

$$p_{\tilde{y}}(\beta) = \sum_{n=-\infty}^{+\infty} p_y(\beta+2n\pi). \tag{47}$$

In these equations the effect of the mod2π measurement ambiguity is quite evident. If we know $y=\beta+2n\pi$ where $\tilde{y}=\beta$, the conditional density for x is $p_{x|y}(\alpha|\beta+2n\pi)$. Thus, for our problem $p_{x|\tilde{y}}(\alpha|\beta)$ is the weighted sum of the conditional densities $p_{x|y}(\alpha|\beta+2n\pi)$ where the weighting constant for the nth term can be shown to be the conditional probability that $y=\beta+2n\pi$ given $\tilde{y}=\beta$ (see [2] and [3]). We note that these results have been derived informally. Rigorous derivations of these and several somewhat more general results are contained in [2] and [3].

B. The Linear–Gaussian Case

In order to interpret these one-stage equations and extend them to the multistage case, we now limit our discussion to the linear–Gaussian case—i.e., where

$$a(x,k)=a_k x \tag{48}$$

$$h(x,k)=h_k x \tag{49}$$

$$b^{1/2}(x,k)=b_k^{1/2} \tag{50}$$

and all *a priori* densities are normal. If $p_x(\alpha)=N(\alpha;\eta,\gamma_1)$, $p_v(\nu)=N(\nu;0,\gamma_2)$, and

$$y=hx+v, \tag{51}$$

then substituting into (43)–(47) and using known results about linear measurements, we have

$$p_{x|\tilde{y}}(\alpha|\beta)=\sum_{n=-\infty}^{+\infty} c_n(\beta)N(\alpha;\eta_n,\gamma_3) \tag{52}$$

where

$$\eta_n=\frac{\gamma_2\eta+\gamma_1 h(\beta+2n\pi)}{h^2\gamma_1+\gamma_2}, \qquad \gamma_3=\frac{\gamma_1\gamma_2}{h^2\gamma_1+\gamma_2} \tag{53}$$

$$c_n(\beta)=\frac{N(\beta+2n\pi;h\eta,h^2\gamma_1+\gamma_2)}{\sum_{k=-\infty}^{+\infty} N(\beta+2k\pi;h\eta,h^2\gamma_1+\gamma_2)}. \tag{54}$$

Thus the nth term is evaluated by an optimal linear estimator which takes as its measurement $\beta+2n\pi$.

C. The Multistage Problem

We now consider the multistage linear problem. It will suffice to consider only two stages, since this will indicate the type of recursions necessary in building the discrete filter. Thus, consider (34)–(37) with a_k, h_k, and $b_k^{1/2}$ given by (48)–(50) and with $p_{v_1}(\nu)=N(\nu;0,r_1)$ and $p_{v_2}=N(\nu;0,r_2)$. Also, let the *a priori* density for x_1 be $N(\alpha;0,b_0)$. The preceding analysis shows that the density $p_{x_1|\tilde{y}_1}(\alpha|\beta_1)$ is given by (52)–(54) with $\eta=0$, $\gamma_1=b_0$, $\gamma_2=r_1$, $\beta=\beta_1$, and $h=h_1$. Then, it is easy to see that the density $p_{x_2|\tilde{y}_1}(\alpha|\beta_1)$ is given by

$$p_{x_2|\tilde{y}_1}(\alpha|\beta_1)=\sum_{n=-\infty}^{+\infty} c_n(\beta_1)N(\alpha;a_1\eta_n,a_1^2\gamma_3+b_1) \tag{55}$$

where the c_n, η_n, and γ_3 are computed from (53) and (54) with the above substitutions.

It remains to include the effect of the additional measurement $\tilde{y}_2$. To do this, we refer to (43)–(47) and assume that $p_v(\nu)=N(\nu;0,\gamma_2)$ and $h(x)=hx$. In this case, the *a priori* density for x_2—the density for x_2 just before we process $\tilde{y}_2$—is $p_{x_2|\tilde{y}_1}$, which is given by (55). This density is of the form

$$p_x(\alpha)=\sum_{k=-\infty}^{+\infty} d_k N(\alpha;\eta_k,\gamma_1) \tag{56}$$

where

$$\sum_{k=-\infty}^{+\infty} d_k=1, \qquad d_k\geqslant 0. \tag{57}$$

In this case (43) becomes

$$p_{x|\tilde{y}}(\alpha|\beta)=\sum_{n=-\infty}^{+\infty}\sum_{k=-\infty}^{+\infty} e_{nk}(\beta)N(\alpha;\mu_{nk},\gamma_3) \tag{58}$$

where γ_3 is given by (53) and

$$\mu_{nk}=\frac{\gamma_2\eta_k+\gamma_1 h(\beta+2n\pi)}{h^2\gamma_1+\gamma_2} \tag{59}$$

$$e_{nk}=\frac{d_k N(\beta+2n\pi;h\eta_k;h^2\gamma_1+\gamma_2)}{\sum_{l=-\infty}^{+\infty}\sum_{m=-\infty}^{+\infty} d_m N(\beta+2l\pi;h\eta_m;h^2\gamma_1+\gamma_2)}. \tag{60}$$

From this computation we can see a pattern. After m measurements, the density $p_x(x,m|\tilde{Y}_m)$ is an m-times infinite sum of normal densities, all with the same variance—the one given by linear theory—and the mean of the $(i_1,i_2,\cdots,i_m)$th term is the linear result if

$$y_1=\tilde{y}_1+2i_1\pi,\cdots,y_m=\tilde{y}_m+2i_m\pi. \tag{61}$$

The coefficient of this term is the conditional probability that (61) holds, and it is a nonlinear function of the measurements [the update equation for these coefficients is (60)].

It is clear that any practical implementation of these results requires some approximate scheme. We will briefly discuss two approaches and refer the reader to [3] and [5] for more detailed remarks. We also note that truncation procedures for sums of normal densities have been considered by Buxbaum and Haddad [15] in relation to a different class of estimation problems, and our first approach is motivated by their work.

Since the coefficients in the series

$$p_{x|\tilde{y}_1,\cdots,\tilde{y}_m}(\alpha|\beta_1,\cdots,\beta_m)$$

$$=\sum_{i_1=-\infty}^{+\infty}\cdots\sum_{i_m=-\infty}^{+\infty} c_{i_1,\cdots,i_m}N(\alpha;\eta_{i_1,\cdots,i_m},\gamma_m) \tag{62}$$

are the conditional probabilities for the various y_i, it would seem reasonable to seek truncation methods based on the properties of these probabilities. Therefore, we consider the following: we are given a positive integer N and wish to recursively choose $2N$ real numbers $\{d_n(m)\}_{n=1}^N$, $\{\mu_n(m)\}_{n=1}^N$ with

$$\sum_{n=1}^{N} d_n(m)=1, \qquad d_n(m)\geq 0 \tag{63}$$

such that the density

$$f(\alpha,m)=\sum_{n=1}^{N} d_n(m)N(\alpha;\mu_n(m),\gamma_m) \tag{64}$$

"approximates" (62). The motivation for this form is the following: we have the interpretation of (61) as being computed by an infinite bank of linear filters, which also compute the appropriate probabilities $c_{i_1,\cdots,i_m}$; therefore, we wish to truncate the bank of filters in order to make the computations feasible. See Fig. 1, which illustrates the basic concept behind such a truncation procedure.

The most obvious method of truncation is to choose the N largest c's at each stage. Call them $c_1(m),\cdots,c_N(m)$, and let

$$d_i(m)=\frac{c_i(m)}{\sum_{j=1}^{N} c_j(m)}, \qquad \mu_i(m)=\eta_i(m). \tag{65}$$

Then we take the density (64) and use it to continue the procedure—i.e., to propagate forward in time [as in (55)] and to process the next measurement [as in (56)–(60)]. Once having applied (56)–(60), we choose the N largest coefficients and repeat the process.

Of course, we cannot directly implement (56)–(60), since these equations involve an infinite number of coefficients. Thus, we must assume (*a priori*) that for any value of $\tilde{y}$, there are only a finite number of values of n such that $P(y=\tilde{y}+2n\pi|\tilde{y})>0$. This approximation is not too severe, since we can show that $P(y=\tilde{y}+2n\pi|\tilde{y}) \sim e^{-n^2}$. Having made this assumption, we compute only a finite number of c's from which we choose the N largest. Finally, we note that Buxbaum and Haddad [15] found this type of approach to be quite effective. Thorough discussions of this and several other truncation methods are contained in [3], [5], and [15].

We remark that [2] and [3] contain a result relating the discrete- and continuous-time problems. This result states that if our discrete-time problem is a discretization of the continuous one, then, as the time Δt between measurements becomes small, the term in the conditional density corresponding to n rotations between measurements ($\Delta y = \Delta\tilde{y}+2n\pi$) goes to zero as $e^{-n^2/\Delta t}$. Thus, for small Δt, a rather crude truncation procedure may provide adequate accuracy. In fact, it may be appropriate to approximate the discrete filter by the continuous filter preceded by a sample and hold (see the continuous filter design in [1]).

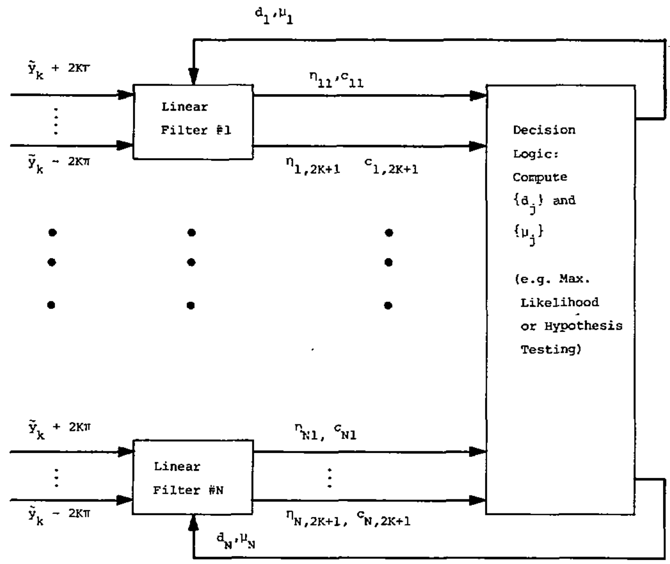

Fig. 1. Conceptual diagram of the truncation method for suboptimal discrete-time filtering.

An additional comment can be made if we are interested in studying $\theta_k = x_k \bmod 2\pi$ as opposed to x_k. In the linear case, the conditional density $p_\theta(\theta,k|\tilde{Y}_k)$ is an infinite sum of folded normal densities. For instance, if we let $\theta = x \bmod 2\pi$ with the density for x given by (52), we have

$$p_{\theta|\tilde{y}}(\alpha|\beta)=\sum_{n=-\infty}^{+\infty} c_n(\beta)F(\alpha;\eta_n,\gamma_3). \tag{66}$$

Since $F(\alpha;\eta,\gamma)=F(\alpha;\eta+2n\pi,\gamma)$, we need only keep those η_n that are inequivalent $\bmod 2\pi$. For example, suppose γ_1 and γ_2 are rationally related—i.e., γ_1/γ_2 is a rational number—and $h=1$. Then let n_0 be the smallest positive integer such that $\gamma_1 n_0$ is an integral multiple of $(\gamma_1+\gamma_2)$. Referring to (53), we have that there are only n_0 distinct folded normal densities in (66) with "means" $\eta_0,\cdots,\eta_{n_0-1}$. In this case, we can write (66) as a sum of a finite number of folded normal densities where the coefficient of each of the terms is computed by summing the $c_n(\beta)$ corresponding to all those η_n that are equivalent $\bmod 2\pi$ to a particular η_j. Thus, if we approximate γ_1 and γ_2 so that n_0 is small, $p_{\theta|\tilde{y}}$ is the sum of only a few terms.

We now comment on the optimal estimation problem. We assume that

$$p_x(\alpha)=\sum_{n=-\infty}^{+\infty} c_n N(\alpha;\eta_n,\gamma). \tag{67}$$

As is well known, the mean $\mathcal{E}(x)$ is the minimum variance estimate of x on R^1, and in this case,

$$\mathcal{E}(x)=\sum_{n=-\infty}^{+\infty} c_n\eta_n. \tag{68}$$

Then in the folded normal case depicted in Fig. 1, the optimal estimate is a linear combination of the outputs of the various linear filters, and the coefficients are the same as the coefficients computed to decide which terms to keep.

If we wish to estimate $\theta = x \bmod 2\pi$, we use the results of Section II. In particular, if we wish to minimize $\mathcal{E}(1-\cos(\theta-\hat{\theta}))$, the optimal estimate $\hat{\theta}_0$ is given by

$$\hat{\theta}_0 = \tan^{-1} \frac{\sum_{n=-\infty}^{+\infty} c_n \sin \eta_n}{\sum_{n=-\infty}^{+\infty} c_n \cos \eta_n}. \tag{69}$$

So, again referring to Fig. 1, the optimal estimate is a nonlinear function of the outputs of the linear filters and the coefficients c_n.

As in the continuous-time case, we can consider the multidimensional (general Abelian Lie group) analog of the results of this section. For instance, let x be an n-dimensional normally distributed random variable, w a k-dimensional normal random variable independent of x, and C a $k \times n$ matrix. Define the measurement $\tilde{y}$:

$$y = Cx + w \tag{70}$$

$$\tilde{y}_i = \begin{cases} y_i, & 1 \leqslant i \leqslant k_1 \\ y_i \bmod 2\pi, & k_1 < i \leqslant k. \end{cases} \tag{71}$$

Then the conditional density $p_{x|\tilde{y}}$ can be written as a $(k-k_1)$-times countably infinite sum of normal distributions, the $(r_1, \cdots, r_{k-k_1})$th of which is the conditional density assuming

$$y_{k_1+i} = \tilde{y}_{k_1+i} + 2r_i\pi, \qquad i = 1, \cdots, k-k_1 \tag{72}$$

and the coefficient of this term is the conditional probability that (72) holds, given $\tilde{y}$.

IV. Fourier Analysis of Conditional Probability Distributions

In Section II we saw how one could express the optimal estimate of a random variable on S^1 in terms of the Fourier coefficients of the probability density of the random variable. In view of this, in considering the conditional probability distribution of a random process on S^1, it is natural to seek methods for tracking the Fourier coefficients. In this section we will consider a general single-stage S^1 estimation problem. Extensions to multistage problems with measurement noise independent from stage to stage is immediate. We note that the results of this section have also been discussed in [2] and [3], and Fourier series techniques have been used in [3], [4], [23], and [10] to aid in the analysis of a variety of discrete- and continuous-time estimation problems.

Let θ be a random variable on the circle with *a priori* density

$$p_\theta(\xi) = \frac{1}{2\pi} + \sum_{n=1}^{\infty} [a_n(0)\sin n\xi + b_n(0)\cos n\xi]. \tag{73}$$

Suppose that we take a single (possibly nonlinear) measurement y of θ and that the noise density $p_{y|\theta}(\nu|\xi)$ exists. Since for fixed ν, $p_{y|\theta}(\nu|\xi)$ is periodic in ξ with period 2π, we can write $p_{y|\theta}(\nu|\xi)$ in Fourier series form in ξ:

$$p_{y|\theta}(\nu|\xi) = d_0(\nu) + \sum_{n=1}^{\infty} [c_n(\nu)\sin n\xi + d_n(\nu)\cos n\xi] \tag{74}$$

where the c_n's and d_n's are functions of ν. Applying Bayes' rule, we can compute the Fourier series form for the conditional density $p_{\theta|y}(\xi|\nu)$:

$$p_{\theta|y}(\xi|\nu) = \frac{1}{2\pi} + \sum_{n=1}^{\infty} [a_n(1)\sin n\xi + b_n(1)\cos n\xi] \tag{75}$$

where

$$a_n(1) = \frac{\alpha_n(\nu)}{2\pi c(\nu)}, \qquad b_n(1) = \frac{\beta_n(\nu)}{2\pi c(\nu)} \tag{76}$$

$$c(\nu) = \frac{1}{2\pi} p_y(\nu) = \frac{d_0(\nu)}{2\pi} + \tfrac{1}{2}\sum_{n=1}^{\infty} [a_n(0)c_n(\nu) + b_n(0)d_n(\nu)] \tag{77}$$

$$\begin{aligned}\alpha_k(\nu) = {} & a_k(0)d_0(\nu) + \frac{c_k(\nu)}{2\pi} \\ & + \tfrac{1}{2}\sum_{n=1}^{k-1} [a_n(0)d_{k-n}(\nu) + b_n(0)c_{k-n}(\nu)] \\ & + \tfrac{1}{2}\sum_{n=1}^{\infty} \{[a_{n+k}(0)d_n(\nu) + b_n(0)c_{n+k}(\nu)] \\ & - [a_n(0)d_{n+k}(\nu) + b_{n+k}(0)c_n(\nu)]\}\end{aligned} \tag{78}$$

$$\begin{aligned}\beta_k(\nu) = {} & b_k(0)d_0(\nu) + \frac{d_k(\nu)}{2\pi} \\ & + \tfrac{1}{2}\sum_{n=1}^{k-1} [b_n(0)d_{k-n}(\nu) - a_n(0)c_{k-n}(\nu)] \\ & + \tfrac{1}{2}\sum_{n=1}^{\infty} \{[a_{n+k}(0)c_n(\nu) + b_n(0)d_{n+k}(\nu)] \\ & + [a_n(0)c_{n+k}(\nu) + b_{n+k}(0)d_n(\nu)]\}.\end{aligned} \tag{79}$$

Note that the equations for c, α_k, and β_k are bilinear in the Fourier coefficients of $p_\theta(\xi)$ and $p_{y|\theta}(\nu|\xi)$. Thus, the computation of $p_{\theta|y}$ involves the (in general, nonlinear) computation of the coefficients $\{c_n(\nu)\}$ and $\{d_n(\nu)\}$ (here ν is the actually observed measurement value), the evaluation of the bilinear equations (77)–(79), and the normalization (76). Thus, although the equations (77)–(79) look complicated, they are highly structured [(77)–(79) look like convolutions of infinite series]. We note that this approach is extremely general, in that the only restriction on the form of the measurement is that the conditional density $p_{y|\theta}$ exist. For example, in addition to measurements such as

$$y=(\theta+v)\bmod 2\pi \qquad (80)$$

which are considered in the previous section, using the Fourier series approach we can also consider measurements such as

$$y=\sin\theta+v. \qquad (81)$$

The reader is referred to [3] and [4] for examples of these Fourier techniques and discussions of the extension of these results to multistage problems and the problem of computing conditional distributions for random processes on S^1 given a sequence of observations (as well as the continuous observation case).

We note that Fourier analysis has previously been applied in [23] and [10] to the specific signal form (81), and infinite dimensional optimal estimation equations were derived. These equations can be interpreted as consisting of an infinite bank of phase tracking loops [23], [4], and such an interpretation leads directly to finite dimensional approximations obtained by "truncating" the bank of filters [10], [4] (for some successful simulations, see [4], [24]).

This idea of "truncating" a bank of filters is essentially the same as the "truncation" of the infinite set of Fourier coefficient update equations (75)–(79), a problem that we now briefly consider. First of all, since we can only store a finite number of terms, we *must* truncate the various series in some manner. As discussed in Section II, this is not terribly serious since the coefficients fall off rapidly in size. Also, if we are using an estimation criterion such as $\mathcal{E}(1-\cos(\theta-\hat{\theta}))$ (as one would in phase tracking problems [4], [9]), at any stage we actually use only a_1 and b_1. Of course, (75)–(79) show that, in general, all of the *a priori* coefficients directly influence each of the updated coefficients.

The question of truncating the probability densities is in itself an interesting problem. Suppose we keep the first N modes of p_θ and the first M modes of $p_{y|\theta}$. It is easy to check that, in general, $p_{\theta|y}$ will then have nonzero terms up to the $(N+M)$th mode. Thus, if we are considering a sequence of measurements, we must devise techniques for sequentially truncating the conditional density. As discussed in [3] and [4], if we keep N modes of the *a priori* density p_θ, and if we make some assumption about the shape of p_θ (this is called "assumed density" approximation [19]), we can use $\{a_n,b_n\}_{n=1}^N$ to approximate the higher coefficients, which can then be used in computing the first N "updated" coefficients (of the conditional density $p_{\theta|y}$). Such a procedure may provide better accuracy than a straight truncation of the equations. An example that shows the type of errors that enter when we truncate the various Fourier series is presented in [3].

Once we have approximated the various infinite series and the infinite set of update equations (75)–(79), we are left with the problem of performing the update computations in real time. As noted earlier, the Fourier coefficient update equations possess a convolution-type structure. Therefore, the fast Fourier transform–high-speed convolution techniques of Stockham [21] will be of value in designing efficient methods for performing the necessary computations.

Finally, once an approximation to (75)–(79) is implemented, we can then use the Fourier analysis results of Section II to obtain an optimal estimate *directly as a function of the Fourier coefficients*. It is this feature that, in fact, motivated the development of the Fourier coefficient update coefficients.

Our discussion concerning suboptimal phase estimation schemes arising from the Fourier series approach has been quite superficial and is meant only to indicate the important issues. The reader is referred to [3] and [4] for more detailed discussions and for numerical results for a specific continuous-time phase tracking problem.

V. Conclusions

In this paper we have studied several classes of discrete-time estimation problems on the circle. In particular, we have considered the discrete-time analog of the continuous-time bilinear estimation problem examined in [1], and an intrinsic difference between the continuous and discrete problems was discussed. This difference stems from the loss of information between the discrete measurements. It is significant that, unlike the vector space case, this loss of information causes a striking increase in the complexity of the expression for the conditional probability distribution. The problem of suboptimal estimation techniques motivated by the form of the conditional density has been considered briefly.

In addition, Fourier series techniques have been used to investigate a general class of discrete-time S^1 estimation problems. Optimal estimation equations and highly structured sequential equations for conditional densities have been presented. The problem of suboptimal Fourier series estimation has been briefly discussed.

The results of this paper indicate that rather general classes of estimation problems on S^1 possess a great deal of structure, and it is the hope of the authors that this structure can be successfully utilized to develop efficient and accurate estimation techniques. The reader is referred to [3], [4], and [5] in which this question is considered in more detail and in which some further results are reported.

Acknowledgment

The authors wish to thank Prof. R. W. Brockett of Harvard University, Cambridge, Mass., for his support and advice during the preparation of this paper.

References

[1] J. T.-H. Lo and A. S. Willsky, "Estimation for rotational processes with one degree of freedom—Part I: Introduction and continuous-time processes," this issue, pp. 10–21.

[2] ——, "Estimation for rotational processes with one degree of freedom," Div. Eng. Appl. Phys., Harvard Univ., Cambridge, Mass., Tech. Rep. 635, July 1972.

[3] A. S. Willsky, "Dynamical systems defined on groups: Structural properties and estimation," Ph. D. dissertation, Dep. Aeronautics and Astronautics, Massachusetts Inst. Technol., Cambridge, June 1973.

[4] ——, "Fourier series and estimation on the circle with applications to synchronous communication—Part I: Analysis and Part II: Implementation," *IEEE Trans. Inform. Theory*, vol. IT-20, pp. 577–590, Sept. 1974.

[5] ——, "Equivalence class observation problems and the truncation of series of normal densities," to be submitted for publication.

[6] R. E. Kalman, "A new approach to linear filtering and prediction problems," *Trans. ASME*, J. Basic Eng., ser. D, vol. 82, pp. 35–45, 1960.

[7] ——, "New methods in Wiener filtering theory," in *Proc. Symp. Appl. Random Function Theory and Probability*, J. L. Bogdanoff and F. Kozin, Ed. New York: Wiley, 1963.

[8] R. S. Bucy and P. D. Joseph, *Filtering for Stochastic Processes with Applications to Guidance*. New York: Wiley-Interscience, 1968.

[9] R. S. Bucy and A. J. Mallinckrodt, "An optimal phase demodulator," *Stochastics*, vol. 1, pp. 3–23, 1973.

[10] R. S. Bucy, C. Hecht, and K. D. Senne, "New methods for nonlinear filtering," *Rev. Francaise d'Automatique, Informatique et de Recherche Operationelle*, pp. 3–54, Feb. 1973.

[11] J. Rotman, *The Theory of Groups: An Introduction*. Boston, Mass.: Allyn and Bacon, 1965.

[12] C. Chevalley, *Theory of Lie Groups*. Princeton, N. J.: Princeton Univ. Press, 1946.

[13] J. T.-H. Lo, "Finite-dimensional sensor orbits and optimal nonlinear filtering," Dep. Aerosp. Eng., Univ. Southern California, Los Angeles, Tech. Rep. USCAE 114, Aug. 1969; also, *IEEE Trans. Inform. Theory*, vol. IT-18, pp. 583–588, Sept. 1972.

[14] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*. New York: McGraw-Hill, 1965.

[15] P. J. Buxbaum and R. A. Haddad, "Recursive optimal estimation for a class of non-Gaussian processes," in *Proc. Symp. Computer Processing in Commun.*, Polytechnic Inst. Brooklyn, Brooklyn, N. Y., Apr. 1969.

[16] M. Abramowitz and I. A. Stegun, Ed., *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables* (Appl. Math. Series 55). Washington, D.C.: Nat. Bureau Standards, 1964.

[17] F. B. Hildebrand, *Advanced Calculus for Applications*. Englewood Cliffs, N. J.: Prentice-Hall, 1962.

[18] A. S. Willsky, "Estimation on finite cycle groups with applications to coding," to be submitted for publication.

[19] B. W. Licht, "Approximations in optimal nonlinear filtering," Syst. Res. Cen., Case Western Reserve Univ., Cleveland, Ohio, Rep. SRC 70-1, Apr. 1970.

[20] P. J. Buxbaum, "Recursive optimal estimation for a class of non-Gaussian processes," Ph. D. dissertation, Polytechnic Inst. Brooklyn, Brooklyn, N. Y., June 1969.

[21] T. G. Stockham, Jr., "High-speed convolution and correlation," in *1966 Spring Joint Conf., AFIPS Conf. Proc.*, vol. 28. Washington, D.C.: Spartan, 1966, pp. 229–233.

[22] A. S. Willsky and J. T.-H. Lo, "Estimation for rotational processes with one degree of freedom—II: Discrete-time processes," Electron. Syst. Lab., Massachusetts Inst. Technol., Cambridge, Rep. ESL-P-507, July 1973.

[23] A. J. Mallinckrodt, R. S. Bucy, and S. Y. Chang, "Final project report for a design study for an optimal nonlinear received demodulator," Electrac Inc., Rep. to Goddard Space Flight Center, Greenbelt, Md., NASA Contract NAS5-10789, Aug. 31, 1970.

[24] R. S. Bucy and H. Youssef, "Fourier realization of an optimal phase demodulator," in *Proc. 4th Symp. Nonlinear Estimation*, San Diego, Calif., 1973, pp. 34–38.

James Ting-Ho Lo (M'73), for a photograph and biography, see this issue, page 21.

Alan S. Willsky (S'70–M'73), for a photograph and biography, see this issue, page 21.