Fig. 9. Estimation error of V_x . I proposed algorithm, II pseudolinear filter ($m=4^\circ$, $\sigma=6^\circ$; Example 3).

Fig. 10. Estimation error of V_y . I proposed algorithm, II pseudolinear filter ($m=4^\circ$, $\sigma=6^\circ$; Example 3).

No attempt has been made to optimize the discretization of the parameter space. This is an area for further investigation.

REFERENCES

- [1] R. S. Bucy and K. D. Senne, "Digital synthesis of non-linear filters," *Automatica*, vol. 7, pp. 287-298, 1971.
- [2] R. S. Bucy, R. A. Geesey, and K. D. Senne, "Passive receiver design via nonlinear filtering theory," in *Proc. 3rd Hawaii Int. Conf. on Syst. Sci.*, vol. 1, 1970, pp. 477-480.
- [3] K. D. Senne and R. S. Bucy, "Digital realization of optimal discrete-time nonlinear estimators," in *Proc. 4th Annu. Princeton Conf. on Syst. Sci.*, 1970, pp. 280-284.
- [4] R. L. Klein and A. H. Wang, "Implementation of nonlinear estimators using spline interpolation and quadrature formulas," in *Proc. 1st Int. Conf. on Inform. Sci. and Syst.*, Patras, Greece, 1977, pp. 197-201.
- [5] A. G. Lindgren and K. F. Gong, "Position and velocity estimation via bearing observations," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-14, pp. 564-577, 1978.
- [6] V. J. Aidala, "Kalman filter behavior in bearings-only tracking applications," *IEEE Aerosp. Electron. Syst.*, vol. AES-15, pp. 29-39, 1979.
- [7] D. J. Murphy, "Noisy bearings-only target motion analysis," Ph.D. dissertation, Dep. Elec. Eng., Northeastern Univ., 1970.
- [8] L. Ljung, "Asymptotic behavior of the extended Kalman filter as a parameter estimator for linear systems," *IEEE Trans. Automat. Contr.*, vol. AC-24, pp. 36-50, Feb. 1979.
- [9] D. G. Lainiotis, "Optimal adaptive estimation: Structure and parameter adaptation," *IEEE Trans. Automat. Contr.*, vol. AC-16, pp. 160-170, 1971.
- [10] F. L. Sims, D. G. Lainiotis, and D. T. Magill, "Recursive algorithm for the calculation of the adaptive Kalman filter weighting coefficients," *IEEE Trans. Automat. Contr.*, vol. AC-14, pp. 215-218, 1969.
- [11] R. W. Bass et al., "ASW target motion and measurement models," Comput. Software Analysis, Inc., Tech. Rep. TR-72 024-01, 1972.
- [12] R. S. Bucy and P. D. Joseph, *Filtering for Stochastic Processes with Application to Guidance*. New York: Wiley, 1968.
- [13] K. J. Åström and P. Eykhoff, "System identification—A survey," *Automatica*, vol. 7, pp. 123-162, 1971.
- [14] D. G. Lainiotis, "Partitioning: A unifying framework for adaptive systems—Part I: Estimation," *Proc. IEEE*, vol. 64, pp. 1126-1143, Aug. 1976; "Partitioning: A unifying framework for adaptive systems: Part II: Control," *Proc. IEEE*, vol. 64, pp. 1182-1198, Aug. 1976.

Edgeworth Expansions in State Perturbation Estimation

WARREN W. WILLMAN, MEMBER, IEEE

Abstract—A formal asymptotic Hermite series expansion is developed for the conditional state density in a scalar estimation problem which differs from the standard Kalman filtering context only by the addition of a small quadratic term in the dynamics. This case is one that might result from a higher order description of perturbations in a more general state estimation situation. The parameters of the expansion here are generated recursively from the measurement data by an initial value system of $3n$ ordinary differential equations for an expansion of n th order asymptotic accuracy in the coefficient of the above quadratic term. To at least third order, this expansion has the special form of the standard Edgeworth expansion of corresponding order for the conditional state density, and its parameters can be expressed in terms of this density's cumulants, which are generated by a simpler set of recursive equations.

INTRODUCTION

There is considerable interest in state estimation problems where a state variable evolves according to linear dynamics with additive white Gaussian noise, and where linear state measurements are available which are also corrupted by the same type of noise. One reason this case is regarded as important is that it is the form assumed by the first-order description of noise-induced state and measurement perturbations from nominal values in a more general class of situations. If the description of such perturbations is carried out to one higher order of accuracy, the effect is typically to introduce quadratic perturbation terms in the dynamics and measurements. Some higher order expansions of this sort are considered in more detail in Gustafson and Speyer [1]. The resulting estimation problem can often be rescaled so that the state and measurement perturbations are of order unity and the coefficients of the quadratic terms become the relatively small quantities.

A formal analysis has been made of estimation problems with this latter type of structure [2]. The results had the formal appearance of giving an approximation to the conditional probability density of the current state, given the currently available measurements, which was accurate to first order in the small coefficients. Similar formal results have also been obtained for the corresponding "smoothing problem" [3], where the density is conditioned on future measurements as well. In each case this first-order approximation to the conditional state density had the interesting property of being, at least in formal appearance, the standard first-order Edgeworth expansion [4] of that density. Also, the parameters of this expansion (first, second, and third central moments) could be generated recursively from the measurement data by a system of initial value equations. This formalism has been verified [5] under fairly general conditions for the special case of a scalar discrete-time problem in which the only nonlinearity is a small product of the state and noise variables in the dynamics, and in a sense which is strong enough to imply that the optimizing control for a quadratic performance criterion which is generated according to this Edgeworth approximation is indeed accurate to first order in the small coefficients.

Higher order Edgeworth expansions of the conditional state density, with parameters generated from the measurement data by a recursive system of equations, have been proposed by Sorenson and Stubberud [6], [7] for approximate state estimation. This idea is pursued here by the development of higher order approximations to the conditional state density in a scalar estimation problem in which the only nonlinearity is a small quadratic term in the noise and state variables of the dynamics, as in the case of [5]. The development here is only formal, however, i.e., it is

Manuscript received December 4, 1979; revised August 25, 1980. Paper recommended by A. Ephremides, Past Chairman of the Estimation Committee.
The author is with the Naval Research Laboratory, Washington, DC 20375.

assumed without rigorous justification that all quantities that *appear* to be of high enough order to be negligible *are* so in some appropriate sense. As a result, this problem is considered in its continuous-time version here, or rather time discretizations thereof, because the formalism is actually simpler in the absence of any precise error analysis. Also, the approximations of the conditional state density are in terms of a more general series of Hermite polynomials, although these can be shown by computation to reduce to the special Edgeworth form out to at least third order. The hope is, of course, that this pattern continues to hold for more general types of higher order approximations, in terms of perturbation variables, of nonlinear state estimation problems.

In other related work, Kuznetsov, Stratonovich, and Tikhonov [8] have developed methods for using the more general type of Hermite series expansions just mentioned to characterize multidimensional distributions of random processes under more general conditions for purposes including nonlinear filtering. They call the parameters of these expansions quasi-moments. In this context the results here consist of explicit recursive equations for determining the quasi-moments of the conditional state density in a more specific case; and also the fact that the corresponding expansion there is formally asymptotic in terms of a certain small parameter. Also, Kushner [9] has developed similar recursive equations for approximating parameters of conditional state distributions, again for a more general case, but these recursions are for the moments rather than the quasi-moments or the cumulants. That work does not deal with any associated approximations for the conditional distribution function itself, however, and the form of the recursions is less explicit than here.

PROBLEM STATEMENT

The particular problem actually treated here is one in which a real scalar state variable x has dynamics

$$x_{i+1} - x_i = \Delta f_i x_i + (1 + \psi_i x_i) w_i; \quad i=0, 1, \dots \quad (1)$$

and of which measurements z are available of the form

$$z_i = x_i + v_i \quad (2)$$

where

- 1) Δ is a time discretization parameter.
- 2) x_0 , w_i , and v_i are mutually independent zero-mean Gaussian random variables such that

$$\text{var}(x_0) = P_0$$

$$\text{var}(w_i) = q_i \Delta$$

$$\text{var}(v_i) = r_i \Delta.$$

- 3) $|\psi_i| \leq h$, where $h > 0$ is regarded as an approximately infinitesimal parameter; f_i , P_0 , $1/P_0$, q_i , r_i , and $1/r_i$ are of order unity.

- 4) For some arbitrary but fixed positive integer n , $0 < \Delta \ll h^{2n}$ (so that $h^k \Delta \gg \Delta^{3/2}$ for $k \leq n$).

- 5) $f_i = f(i\Delta)$, $i=0, 1, \dots$, for some suitably well-behaved time function $f(t)$; likewise for q_i , r_i , and ψ_i .

This problem is meant to be regarded as a formal discretization of the continuous-time situation described by the following Ito differential equations:

$$dx = f(t)x dt + (1 + \psi(t)x) dw \quad (3)$$

$$dy = x dt + dv \quad (4)$$

where $w(t)$ and $v(t)$ are independent Wiener processes with respective variance parameters $q(t)$ and $r(t)$. The measurement process y here is essentially the integral of the z_i measurements in the discretization, i.e.,

$$y(i\Delta) \approx \sum_{j=0}^i z_j \Delta.$$

For any particular choice of $\Delta > 0$, condition 4) can only be satisfied for a restricted range of $h > 0$ and n . For each such choice of these parameters, however, there is a $\Delta(h, n)$ satisfying this condition, so that the discretiza-

tion error in approximating the continuous-time problem of (3) and (4) [and later in the use of (49)–(55)], which is formally of order $\sqrt{\Delta}$, can also be considered small compared to h^n . The important point here is that even though n and h are treated as fixed in the following analysis of the discrete-time problem, if the values of n and $\Delta(n)$ are then regarded as variable for any fixed $h > 0$, these all give discretizations of the same continuous-time problem, which enables the ensuing result to serve formally as an asymptotic expansion in terms of h there.

HERMITE SERIES EXPANSION

To serve basically as an induction hypothesis on the time index, it is assumed temporarily that the conditional probability density of the current state at a generic epoch i , given the measurements $z_0, \dots, z_i$, is of the form

$$p(x) = \left[\frac{e^{-(x-\hat{x})^2/2P}}{\sqrt{2\pi P}} \right] \left[1 + \sum_{k=3}^{3n} a_k H_k \left(\frac{x-\hat{x}}{\sqrt{P}} \right) + \delta \left(\frac{x-\hat{x}}{\sqrt{P}} \right) \right] \quad (5)$$

where the epoch subscript i and the conditioning on $z_0, \dots, z_i$ are now suppressed in the notation, where n is the integer specified in the problem statement, and where

$$H_k(u) = (-1)^k e^{u^2/2} \frac{d^k}{du^k} (e^{-u^2/2})$$

a_{3k} , a_{3k-1} , and a_{3k-2} are of order h^k , $k=1, \dots, n$ and $\delta(u)$ is of order h^{n-1} , except perhaps for values of u so large that

$$\text{Prob} \left\{ x: \left| \frac{x-\hat{x}}{\sqrt{P}} \right| > |u| \right\} \text{ is negligible.}$$

The H_k are modified Hermite polynomials, so that $H_0(u)=1$, $H_1(u)=u$, $H_2(u)=u^2-1$, etc. Also,

$$\frac{d}{du} H_k(u) = k H_{k-1}(u); \quad k=1, 2, \dots \quad (6)$$

$$\int_{-\infty}^{\infty} e^{-u^2/2} H_k(u) H_m(u) du = \begin{cases} 0 & \text{if } m \neq k \\ \sqrt{2\pi} k! & \text{if } m = k \end{cases} \quad (7)$$

and the following product rule holds for $i \leq k$:

$$H_i H_k = \sum_{m=0}^i \frac{i! k!}{m! (i-m)! (k-m)!} H_{k-i-2m}. \quad (8)$$

Hence, the integral, mean and variance of the expansion (5) are unity, $\hat{x}$ and P , except for the contribution of the error term δ . For this assumed form of the conditional density at epoch i , it is shown in the next two sections that except for an additional error which is formally of order $h^{n-1}\Delta$, the conditional density at epoch $i+1$ is also of this form with the same error function δ if the corresponding parameters are certain functions of $\hat{x}$, P , z_{i-1} , and the a_k .

DENSITY PROPAGATION

In this section the "propagated" density $p[x_{i+1}/z_0, \dots, z_i]$ is computed to an accuracy of order $h^n \Delta$ under the assumption of (5) for epoch i . The result is a series of this same form, together with equations for its parameters in terms of $\hat{x}$, P , and the a_k , such that the residual error function corresponding to δ of (5) is unchanged to this order of accuracy.

By the standard formulas for probability density transformations, if $s = (1 + f\Delta)$, then

$$p(s) = \left[\frac{e^{-(s-\bar{s})^2/2\bar{v}}}{\sqrt{2\pi\bar{v}}} \right] \left[1 + \sum_{k=3}^{3n} a_k H_k \left(\frac{s-\bar{s}}{\sqrt{\bar{v}}} \right) + \delta \left(\frac{s-\bar{s}}{\sqrt{\bar{v}}} \right) \right] \quad (9)$$

where

$$\bar{x} = (1 + f\Delta)\hat{x} \quad (10)$$

and

$$v = (1 + 2f\Delta)P \quad (11)$$

except for terms of formal order Δ^2 . Also, to this same accuracy

$$y = s + \Delta^{1/2}u \quad (12)$$

where y now denotes x_{i+1} and $u = \Delta^{-1/2}(1 + \psi s)w$. Hence,

$$p(u/s) = \frac{e^{-u^2/2q(1+\psi s)^2}}{\sqrt{2\pi q(1+\psi s)^2}}, \quad (13)$$

$$E(y) = \bar{x}, \quad \text{and} \quad (14)$$

$$\text{var}(y) = m = v + q\Delta[(1 + \psi\bar{x})^2 + \psi^2v]. \quad (15)$$

Formally,

$$p(y) = \int_{-\infty}^{\infty} p_s(y - u\Delta^{1/2})p_{u/s}(u, y - u\Delta^{1/2})du \quad (16)$$

where $p_{u/s}(x, y)$ denotes the conditional density of u given s evaluated at $u = x$ for $s = y$, etc. From a formal Taylor series expansion

$$p_{u/s}(u, y - u\Delta^{1/2}) = \frac{e^{-u^2/2b}}{\sqrt{2\pi b}}[1 + \alpha(u)] \quad (17)$$

except for terms of order $\Delta^{3/2}$ (which are negligible to order $h^n\Delta$ by construction), where

$$b = q(1 + \psi y)^2 \quad (18)$$

and

$$\alpha(u) = \frac{\psi\Delta^{1/2}}{1 + \psi y} \left(u - \frac{u^3}{b} \right) + \left(\frac{\psi}{1 + \psi y} \right)^2 \left(u^2 - \frac{5u^4}{2b} + \frac{u^6}{2b^2} \right) \Delta. \quad (19)$$

Substituting (17) into the integral of (16) and completing the square in the resultant exponential factor gives, except for terms of formal order $\Delta^{3/2}$,

$$p(y) = \frac{e^{-(y-\bar{x})^2/2(v+b\Delta)}}{\sqrt{2\pi(v+b\Delta)}} \int_{-\infty}^{\infty} \left[\frac{e^{-(u-\bar{u})^2/2c}}{\sqrt{2\pi c}} \right] [1 + \alpha(u)] \cdot \left[1 + \gamma_n \left(\frac{y - \bar{x} - u\Delta^{1/2}}{\sqrt{v}} \right) + \delta \left(\frac{y - \bar{x} - u\Delta^{1/2}}{\sqrt{v}} \right) \right] du \quad (20)$$

where

$$\gamma_n(x) \text{ denotes } \sum_{k=3}^{2n} a_n H_n(x)$$

and

$$\bar{u} = \frac{b\Delta^{1/2}}{v} (y - \bar{x}) \quad (21)$$

$$c = b - \frac{b^2\Delta}{v} \quad (22)$$

Except for terms which are formally small compared to $h^n\Delta$, the density $p(y)$ can now be obtained by the following steps.

1) Use (6) to develop a Taylor series expansion giving $\gamma_n((y - \bar{x} - u\Delta^{1/2})/\sqrt{v})$ as a series of Hermite polynomials evaluated at $(y - \bar{x})/\sqrt{m}$.

2) Substitute the result into (20) and integrate, adding the contribution of the error term δ separately because if it is not of a "pathological" nature, its contribution to this integral is just $\delta((y - \bar{x})/\sqrt{m})$ except for terms of order $h^{n+1}\Delta$.

3) Use a Taylor series expansion to express the factor outside the integral in (20) as

$$\frac{e^{-(y-\bar{x})^2/2(v+b\Delta)}}{\sqrt{2\pi(v+b\Delta)}} = \left[\frac{e^{-(y-\bar{x})^2/2m}}{\sqrt{2\pi m}} \right] \left[1 + \frac{\psi q\Delta}{\sqrt{m}} (1 + \psi\bar{x})(H_3 + 2H_1) + \psi^2 q\Delta \left(\frac{1}{2}H_4 + 2H_2 + 1 \right) \right] \quad (23)$$

where the H_k are evaluated at $(y - \bar{x})/\sqrt{m}$.

4) Multiply the results of steps 2) and 3) using (8).

The details of this procedure are too complicated to be presented here, but the eventual result after collecting terms and using the fact that a_{3k} , a_{3k-1} , and a_{3k-2} are of order h^k for all $k = 1, \dots, n$, is that (formally)

$$p(y) = \frac{e^{-(y-\bar{x})^2/2m}}{\sqrt{2\pi m}} \left[1 + \sum_{k=3}^{3n} b_k H_k \left(\frac{y-\bar{x}}{\sqrt{m}} \right) + \epsilon \left(\frac{y-\bar{x}}{\sqrt{m}} \right) \right] \quad (23)$$

where

$$\epsilon \left(\frac{y-\bar{x}}{\sqrt{m}} \right) = \delta \left(\frac{y-\bar{x}}{\sqrt{m}} \right) + \text{terms of order } h^{n+1}\Delta,$$

$$\frac{1}{\Delta} (b_3 - a_3) = -\frac{3q}{2v} (1 + \psi\bar{x})^2 a_3 + \frac{q}{\sqrt{v}} (1 + \psi\bar{x}) + \frac{3}{2} q\psi^2 a_3, \quad (24)$$

$$\frac{1}{\Delta} (b_4 - a_4) = -\frac{2q}{v} (1 + \psi\bar{x})^2 a_4 + \frac{3q\psi(1 + \psi\bar{x})}{\sqrt{v}} a_3 + \frac{q\psi^2}{2} (8a_4 + 1), \quad (25)$$

$$\frac{1}{\Delta} (b_5 - a_5) = -\frac{5q}{2v} (1 + \psi\bar{x})^2 a_5 + \frac{4q\psi(1 + \psi\bar{x})}{\sqrt{v}} a_4 + \frac{3q\psi^2}{2} (5a_5 + 2a_3), \quad (26)$$

$$\frac{1}{\Delta} (b_6 - a_6) = -\frac{3q}{v} (1 + \psi\bar{x})^2 a_6 + \frac{q\psi(1 + \psi\bar{x})}{\sqrt{v}} (5a_5 + a_3) + 4q\psi^2 (3a_6 + a_4), \quad (27)$$

and

$$\frac{1}{\Delta} (b_k - a_k) = -\frac{kq}{2v} (1 + \psi\bar{x})^2 a_k + \frac{q\psi(1 + \psi\bar{x})}{\sqrt{v}} [(k-1)a_{k-1} + a_{k-3}] + \frac{q\psi^2}{2} [k(k-2)a_k + 2(k-2)a_{k-2} + a_{k-4}]; \quad k = 7, 8, \dots, 3n. \quad (28)$$

As a consequence, b_{3k} , b_{3k-1} , and b_{3k-2} are also formally of order h^k for all $k = 1, \dots, n$, and differ from a_{3k} , a_{3k-1} , and a_{3k-2} only by order $h^k\Delta$.

UPDATING DENSITY WITH NEW MEASUREMENT

This section consists of a computation of $p[x_{i+1}/z_0, \dots, z_i]$ to order $h^n\Delta$ by using the Bayes rule and the results of the preceding section. The result is again a series of the form of (5), together with equations for its parameters in terms of $\bar{x}$, m , z_{i-1} , and the b_k such that the residual error function is unchanged to order $h^n\Delta$.

Deleting the conditioning on $z_0, \dots, z_i$ and the epoch subscripts of z_{i+1} and r_{i+1} in the notation, and defining

$$v = \Delta^{1/2}n_{i+1} \quad (\text{so } v \text{ has variance } r) \quad (29)$$

gives

$$z = y + \Delta^{-1/2}v \quad (\text{treated here as a term of order } \Delta^{-1/2}) \quad (30)$$

and

$$p(z/y) = \frac{e^{-(z-y)^2\Delta/2r}}{\sqrt{2\pi r/\Delta}}. \quad (31)$$

Since the noise variables and initial state are all statistically independent

$$p(y/z) = gp(y)p(z/y) \quad (32)$$

by the Bayes Rule, where g is a proportionality factor independent of y . The density product in (32) can be expressed as

$$\left[\frac{e^{-(y-\hat{y})^2/2s}}{\sqrt{2\pi s}} \right] \left[\frac{e^{-(y-\bar{x})^2/2(r+m\Delta)}}{\sqrt{2\pi(r+m\Delta)}} \right] \left[1 + \sum_{k=3}^{3n} b_k H_k \left(\frac{y-\bar{x}}{\sqrt{m}} \right) + \epsilon \left(\frac{y-\bar{x}}{\sqrt{m}} \right) \right] \quad (33)$$

where

$$\hat{y} = \bar{x} + \frac{m\Delta}{r+m\Delta}(z-\bar{x}) = \bar{x} + \frac{m\Delta}{r}(z-\bar{x}) + \text{terms of order } \Delta^{3/2}. \quad (34)$$

Now let

$$s = m - \frac{m^2\Delta}{r} + \frac{6\Delta}{r}b_3m^{3/2}(z-\bar{x}). \quad (35)$$

Then the product of the last two factors in expression (33) can be expanded in a formal Taylor series as

$$\left\{ \frac{e^{-(y-\hat{y})^2/2s}}{\sqrt{2\pi s}} \right\} \left\{ 1 + \left[18 \frac{b_3^2 m}{r^2} (z-\bar{x})^2 \Delta^2 - 3 \frac{b_3 \sqrt{m}}{r} (z-\bar{x}) \Delta \right] H_2 \left(\frac{y-\hat{y}}{\sqrt{s}} \right) + \frac{9b_3^2 m}{2r^2} (z-\bar{x})^2 \Delta^2 H_4 \left(\frac{y-\hat{y}}{\sqrt{s}} \right) \right\} \left\{ 1 + \epsilon \left(\frac{y-\bar{x}}{\sqrt{m}} \right) + \sum_{k=3}^{3n} b_k H_k \left[\frac{(y-\hat{y}) + \frac{m\Delta}{r}(z-\bar{x})}{\sqrt{s + \frac{m^2\Delta}{r} - 6 \frac{b_3}{r} m^{3/2}(z-\bar{x})\Delta}} \right] \right\} \quad (36)$$

except for terms that are of order higher than unity in Δ , which by construction are arbitrarily small compared to $h^n \Delta$.

Using (6) to develop a formal Taylor series expansion to first order in Δ for H_k about $(y-\hat{y})/\sqrt{s}$ in (36), substituting the result in expression (33), and using the product formula (8) and the orders of magnitude (in h) of the b_k gives (32) as a series of the form

$$p(y/z) = g \left[\frac{e^{-(y-\hat{y})^2/2s}}{\sqrt{2\pi s}} \right] \left[1 + \sum_{k=3}^{3n} c_k H_k \left(\frac{y-\hat{y}}{\sqrt{s}} \right) + \epsilon \left(\frac{y-\bar{x}}{\sqrt{m}} \right) \right] \quad (37)$$

except for terms formally of order $h^{n+1}\Delta$ or smaller. From the orthogonality property (7) and the fact that $p(y/z)$ is a probability density

$$1 = \int_{-\infty}^{\infty} p(y/z) dy = g \left[1 + \int_{-\infty}^{\infty} \frac{e^{-(y-\hat{y})^2/2s}}{\sqrt{2\pi s}} \epsilon \left(\frac{y-\bar{x}}{\sqrt{m}} \right) dy \right]. \quad (38)$$

Using the fact that

$$\int_{-\infty}^{\infty} \frac{e^{-(y-\hat{y})^2/2s}}{\sqrt{2\pi s}} \epsilon \left(\frac{y-\hat{y}}{\sqrt{s}} \right) dy = 0$$

by construction and a formal Taylor series expansion of ϵ about this argument shows that

$$g = \frac{1}{1+\alpha} \quad (39)$$

where

$$\alpha = \int_{-\infty}^{\infty} \frac{e^{-(y-\hat{y})^2/2s}}{\sqrt{2\pi s}} \epsilon' \left(\frac{y-\hat{y}}{\sqrt{s}} \right) \left[\frac{y-\bar{x}}{\sqrt{m}} - \frac{y-\hat{y}}{\sqrt{s}} \right] dy. \quad (40)$$

Assuming that ϵ is a "well-behaved" function so that ϵ' is also of order h^{n+1} , it follows from (34) and (35) that

$$\alpha = -\frac{6b_3}{r}(z-\bar{x})\Delta \int_{-\infty}^{\infty} \frac{e^{-(y-\hat{y})^2/2s}}{\sqrt{2\pi s}} (y-\bar{x}) \epsilon' \left(\frac{y-\hat{y}}{\sqrt{s}} \right) dy \quad (41)$$

except for terms which are negligible to order $h^n \Delta$. Since the integral in (41) is of order h^{n+1} and b_3 is of order h ,

$$\alpha = \beta(z-\bar{x})\Delta + \text{negligible terms}$$

where β is of order h^{n+2} . Because the values of $(z-\bar{x})$ are statistically independent innovations for distinct epochs, however, this term too can be ignored, since its cumulative effect in (39) for g over a time interval of order unity (basically a sum of $1/\Delta$ such terms) is a zero-mean random variable with variance of order h^{2n+4} (because $z-\bar{x}$ is of order $\Delta^{-1/2}$), which is negligible to order h^n with arbitrarily high probability by the Chebychev inequality. Hence, g in (37) can be taken as unity here. Deleting other similar zero-mean uncorrelated random terms leads eventually to the following equations determining the c_k of (37):

$$\frac{1}{\Delta}(c_3 - b_3) = (4b_4 - 9b_3^2) \frac{\sqrt{m}}{r} (z-\bar{x}) + \frac{m}{2r} (135b_3^3 - 72b_3b_4 - 3b_3) \quad (42)$$

$$\frac{1}{\Delta}(c_4 - b_4) = (5b_5 - 12b_3b_4) \frac{\sqrt{m}}{r} (z-\bar{x}) + \frac{m}{2r} (216b_4b_3^2 - 9b_3^2 - 120b_3b_5 - 4b_4) \quad (43)$$

$$\frac{1}{\Delta}(c_5 - b_5) = (6b_6 - 15b_3b_5 - 3b_3^2) \frac{\sqrt{m}}{r} (z-\bar{x}) + \frac{m}{2r} (315b_5b_3^2 + 90b_3^3 - 24b_4b_3 - 180b_6b_3 - 5b_5) \quad (44)$$

$$\frac{1}{\Delta}(c_6 - b_6) = (7b_7 - 18b_6b_3 - 3b_4b_3) \frac{\sqrt{m}}{r} (z-\bar{x}) + \frac{m}{r} (216b_6b_3^2 + 54b_4b_3^2 - 15b_5b_3 - 126b_7b_3 - 3b_6) \quad (45)$$

$$\frac{1}{\Delta}(c_k - b_k) = [(k+1)b_{k+1} - 3b_3(kb_k + b_{k-2})] \frac{\sqrt{m}}{r} (z-\bar{x}) + \frac{m}{2r} \{ 9b_3^2 [b_{k-4} + 2kb_{k-2} + k(k+2)b_k] - 6b_3 [k(k+1)b_{k+1} + (k-1)b_{k-1}] - kb_k \}; \quad k=7, \dots, 3n \quad (46)$$

except for terms which are formally negligible compared to $h^n \Delta$. Furthermore, $\epsilon((y-\bar{x})/\sqrt{m})$ is the same as $\epsilon((y-\hat{y})/\sqrt{s})$ except for terms which can be neglected to this order of accuracy. From the form of (42)–(46), c_{3k}, c_{3k-1} , and c_{3k-2} are formally of order h^k , and differ from b_{3k}, b_{3k-1} , and b_{3k-2} by order $h^k \Delta$, for all $k=1, \dots, n$. Also, for $k=3n$, $b_{k+1}=0$ in (46).

COMPOSITE RESULT

Since (34), (35), (37), and (42)–(46) describe the situation at epoch $i+1$ corresponding to (5) for epoch i , they, (10), (11), (15), and (24)–(28) form a recursion for generating the conditional state density to a formal accuracy of h^n from the incoming measurements. This follows by what is essentially an induction argument on i . Equation (5) is clearly the form of the Gaussian prior density for x_0 with P_0 as specified, $\hat{x}_0$ and the a_k zero,

and an identically zero error δ . The previous two sections constitute the basis of an induction step showing that the conditional density remains of this form to within an error small compared to h^n , if the parameters are generated according to this recursion and the number of epochs is small compared to $(h\Delta)^{-1}$, since the error accumulated at each step is of order $h^{n+1}\Delta$. In terms of the continuous-time problem being approximated [(3) and (4)] this latter condition corresponds to a time interval small compared to $1/h$.

If the coefficients α_k are defined (for generic epoch i) as

$$\alpha_k = P^{k/2} a_k; \quad k=3, \dots, 3n \quad (47)$$

it is a straightforward but tedious matter of substitution to convert this composite recursion for $\hat{x}$, P , and the a_k to an equivalent one for $\hat{x}$, P , and the α_k , which is of a more convenient form. The quantity $k!\alpha_k$ is the same as the k th (proper) quasi-moment of Kuznetsov, Stratonovich, and Tikhonov [8], [10]. Expressing this result in terms of the formally corresponding continuous-time variables of (3) and (4) for simplicity of notation, and suppressing the "i" variable in the notation, gives

$$p(x/Y) = \left[\frac{e^{-(x-\hat{x})^2/2P}}{\sqrt{2\pi P}} \right] \left[1 + \sum_{k=3}^{3n} \alpha_k P^{-k/2} H_k \left(\frac{x-\hat{x}}{\sqrt{P}} \right) \right] \quad (48)$$

to order h^n , where (since $\sqrt{\Delta} \ll h^n$)

$$\dot{\hat{x}} = f\hat{x} + \frac{P}{r}(z - \hat{x}); \quad \hat{x}(0) = 0 \quad (49)$$

$$\dot{P} = 2fP + q[(1 + \psi\hat{x})^2 + \psi^2 P] - \frac{P^2}{r} + 6\frac{\alpha_3}{r}(z - \hat{x}); \quad P(0) = P_0 \quad (50)$$

$$\begin{aligned} \dot{\alpha}_3 = & 3 \left(f - \frac{P}{r} + q\psi^2 \right) \alpha_3 + Pq\psi(1 + \psi\hat{x}) + \frac{\alpha_3}{2P^2 r} (135\alpha_3^2 - 72P\alpha_4) \\ & + \frac{4\alpha_4}{r}(z - \hat{x}); \quad \alpha_3(0) = 0 \end{aligned} \quad (51)$$

$$\begin{aligned} \dot{\alpha}_4 = & 4 \left(f - \frac{P}{r} + \frac{3}{2}q\psi^2 \right) \alpha_4 + 3\alpha_3 q\psi(1 + \psi\hat{x}) + \frac{1}{2}P^2 q\psi^2 \\ & + \frac{\alpha_3}{2P^2 r} (216\alpha_3\alpha_4 - 9P^2\alpha_3 - 120P\alpha_5) \\ & + \frac{5\alpha_5}{r}(z - \hat{x}); \quad \alpha_4(0) = 0 \end{aligned} \quad (52)$$

$$\begin{aligned} \dot{\alpha}_5 = & 5 \left(f - \frac{P}{r} + 2q\psi^2 \right) \alpha_5 + 4\alpha_4 q\psi(1 + \psi\hat{x}) + 3\alpha_3 Pq\psi^2 \\ & + \frac{\alpha_3}{2P^2 r} [315\alpha_3\alpha_5 - 24P^2\alpha_4 - 90P(2\alpha_6 - \alpha_3^2)] \\ & + \frac{3}{r}(2\alpha_6 - \alpha_3^2)(z - \hat{x}); \quad \alpha_5(0) = 0 \end{aligned} \quad (53)$$

$$\begin{aligned} \dot{\alpha}_6 = & 6 \left(f - \frac{P}{r} + \frac{5}{2}q\psi^2 \right) \alpha_6 + q(5\alpha_5 + P\alpha_3)(1 + \psi\hat{x})\psi \\ & + 4\alpha_4 Pq\psi^2 + \frac{\alpha_3}{2P^2 r} [432\alpha_3\alpha_6 - 30P^2\alpha_5 - 36P(7\alpha_7 - 3\alpha_3\alpha_4)] \\ & + (7\alpha_7 - 3\alpha_3\alpha_4) \frac{(z - \hat{x})}{r}; \quad \alpha_6(0) = 0 \end{aligned} \quad (54)$$

$$\begin{aligned} \dot{\alpha}_k = & k \left(f - \frac{P}{r} + \frac{k-1}{2}q\psi^2 \right) \alpha_k + q\psi(1 + \psi\hat{x})[(k-1)\alpha_{k-1} + P\alpha_{k-3}] \\ & + Pq\psi^2 \left[(k-2)\alpha_{k-2} + \frac{1}{2}P\alpha_{k-4} \right] \\ & + \frac{9\alpha_3^2}{2P^2 r} [k(k+2)\alpha_k + 2kP\alpha_{k-2} + P^2\alpha_{k-4}] \\ & - \frac{3\alpha_3}{Pr} [k(k+1)\alpha_{k+1} + (k-1)P\alpha_{k-1}] \\ & + [(k+1)\alpha_{k+1} - 3\alpha_3\alpha_{k-2}] \frac{(z - \hat{x})}{r}; \end{aligned}$$

$$k \geq 7; \quad \alpha_k(0) = 0 \quad (55)$$

where $Y(t)$ in (48) denotes the function $\{y(s), s: 0 \leq s \leq t\}$, with y now denoting the data of (4) and where z is the formal time-derivative of y , which, for example, could reasonably be implemented here by taking

$$z(t) = \frac{1}{\Delta} \{y(i\Delta) - y[(i-1)\Delta]\}; \quad t \in [i\Delta, (i+1)\Delta). \quad (56)$$

The structure of the equation system (49)–(55) is such that, with the construction of (56), they can be regarded as ordinary differential equations without the need for correction terms of the sort described by Wong and Zakai [11], [12]. Of course, $\alpha_{k+1} = 0$ in (55) for $k = 3n$. Also, many of the terms in this equation become negligible to order h^n as k nears $3n$.

It has also been verified by direct substitution out to third order in h (i.e., to $n=3$ in the notation here) that the expansion of (48) has the special form of a standard n th order Edgeworth expansion of the conditional density $p(x/Y)$. In the third-order case, for example, the result can be expressed as

$$\begin{aligned} p(x/Y) = & \left[\frac{e^{-(x-\hat{x})^2/2P}}{\sqrt{2\pi P}} \right] \left\{ 1 + \frac{\kappa_3}{6P^{3/2}} H_3 + \left[\frac{\kappa_4}{24P^2} H_4 + \frac{\kappa_3^2}{72P^3} H_6 \right] \right. \\ & \left. + \left[\frac{\kappa_5}{120P^{5/2}} H_5 + \frac{\kappa_3\kappa_4}{144P^{7/2}} H_7 + \frac{\kappa_3^3}{1296P^{9/2}} H_9 \right] \right\} \end{aligned} \quad (57)$$

where the H_k s are evaluated at $(x-\hat{x})/\sqrt{P}$ and where κ_3 – κ_5 are the standard cumulants of $p(x/Y)$. These particular cumulants are defined as

$$\left. \begin{aligned} \kappa_3 &= \mu_3 \\ \kappa_4 &= \mu_4 - 3\mu_2^2 \\ \kappa_5 &= \mu_5 - 10\mu_2\mu_3 \end{aligned} \right\} \begin{aligned} &\mu_i = i\text{th central moment of} \\ &p(x/Y) \text{ (so } \mu_2 = P) \end{aligned}$$

and can be generated from the data here to a formal accuracy of order h^3 by the equations

$$\dot{\hat{x}} = f\hat{x} + \frac{P}{r}(z - \hat{x}); \quad \hat{x}(0) = 0, \quad (58)$$

$$\dot{P} = 2fP + q[(1 + \psi\hat{x})^2 + \psi^2 P] - \frac{P^2}{r} + \frac{\kappa_3}{r}(z - \hat{x}); \quad P(0) = P_0, \quad (59)$$

$$\begin{aligned} \dot{\kappa}_3 = & 3 \left(f - \frac{P}{r} + \psi^2 q - \frac{\kappa_4}{2Pr} \right) \kappa_3 + 6Pq\psi(1 + \psi\hat{x}) \\ & + \frac{15\kappa_3^3}{8P^2 r} + \frac{\kappa_4}{r}(z - \hat{x}); \quad \kappa_3(0) = 0, \end{aligned} \quad (60)$$

$$\begin{aligned} \dot{\kappa}_4 = & 4 \left(f - \frac{P}{r} \right) \kappa_4 + 12q\psi[(1 + \psi\hat{x})\kappa_3 + \psi P^2] \\ & - \frac{3\kappa_3^2}{r} + \frac{\kappa_5}{r}(z - \hat{x}); \quad \kappa_4(0) = 0, \end{aligned} \quad (61)$$

and

$$\dot{\kappa}_5 = 5 \left(f - \frac{P}{r} \right) \kappa_5 + 20q\psi(\kappa_4 + 3\psi P\kappa_3) - 10 \frac{\kappa_3\kappa_4}{r}; \quad \kappa_5(0) = 0. \quad (62)$$

The terms in (57) are grouped by orders of magnitude, since it is apparent from (58)–(62) that κ_3 , κ_4 , and κ_5 are of orders h , h^2 , and h^3 , respectively.

Unfortunately, not even a formal demonstration has been found to show for general n that the n th order approximation to $p(x/Y)$ is its n th order Edgeworth expansion. Computationally, the special Edgeworth form has the advantage of reducing the number of independent coefficients by a factor of almost three—from $3n+2$ to $n+2$. Conceptually, this form is interesting because it suggests some kind of connection with the Central Limit Theorem of statistics, from which Edgeworth expansions are originally derived. Finally, it should be noted that, even formally, the results here only constitute an asymptotic expansion for $p(x/Y)$ in powers of the parameter h for some arbitrary but fixed n , not a series converging to this

density for fixed h and increasing n . In the context of refinements of the Central Limit Theorem, the associated Edgeworth expansions sometimes converge in this latter sense also, but often do not for interesting classes of probability densities [4].

REFERENCES

- [1] D. E. Gustafson and J. L. Speyer, "Design of linear regulators for nonlinear stochastic systems," *J. Spacecr. Rockets*, vol. 12, pp. 351-358, June 1975.
- [2] W. W. Willman, "Some formal effects of nonlinearities in optimal perturbation control," *J. Guidance Contr.*, vol. 2, pp. 99-100, Mar. 1979.
- [3] —, "A nonlinear filtering technique for digitized images degraded by film-grain noise," U.S. Naval Res. Lab., Washington, DC, Rep. 8225, Sept. 1980.
- [4] H. Cramer, *Mathematical Methods of Statistics*. Princeton, NJ: Princeton University Press, 1946.
- [5] W. W. Willman, "Precise analysis of an optimal perturbation estimation and control problem," U.S. Naval Res. Lab., Washington, DC, Rep. 8406, Aug. 1980.
- [6] H. W. Sorenson and A. R. Stubberud, "Recursive filtering for systems with small but non-negligible non-linearities," *Int. J. Contr.*, vol. 7, no. 3, pp. 271-280, 1968.
- [7] —, "Non-linear filtering by approximation of the a-posteriori density," *Int. J. Contr.*, vol. 8, pp. 33-51, July 1968.
- [8] P. I. Kuznetsov, P. L. Stratonovich, and V. I. Tikhonov, "Some problems with conditional probability and quasi-moment functions," *Theory Prob. Appl.*, vol. 6, no. 4, pp. 422-427, 1961.
- [9] H. J. Kushner, "Approximations to Optimal Nonlinear Filters," *IEEE Trans. Automat. Contr.*, vol. AC-12, pp. 546-556, Oct. 1967.
- [10] P. I. Kuznetsov, P. L. Stratonovich, and V. I. Tikhonov, "Quasi-moment functions in the theory of random processes," *Theory Prob. Appl.*, vol. 5, no. 1, pp. 80-97, 1960.
- [11] E. Wong and M. Zakai, "On the relation between ordinary and stochastic differential equations," *Int. J. Eng. Sci.*, vol. 3, pp. 213-229, July 1965.
- [12] E. Wong, *Stochastic Processes in Information and Dynamical Systems*. New York: McGraw-Hill, 1971.

The Asymptotic Minimum Variance Estimate of Stationary Linear Single Output Processes

U. SHAKED AND B. BOBROVSKY

Abstract—The problem of minimum error variance estimation of single output linear stationary processes in the presence of weak measurement noise is considered. By applying s domain analysis to the case of single input systems and white observation noise, explicit and simple expressions are obtained for the error covariance matrix of estimate and the optimal Kalman gains both for minimum- and nonminimum-phase systems. It is found that as the noise intensity approaches zero, the error covariance matrix of estimating the output and its derivatives becomes insensitive to uncertainty in the system parameters. This matrix depends only on the shape of the high frequency tail of the power-density spectrum of the observation, and thus it can be easily determined from the system transfer function. The theory developed is extended to deal with white measurement noise in multiinput systems where an analogy to the single input nonminimum-phase case is established. The results are also applied to colored observation noise problems and a simple method to derive the minimum error covariance matrices and the optimal filter transfer functions is introduced.

I. INTRODUCTION

The general problem of optimal estimation of linear processes in the presence of measurement noise has been vastly dealt with in the past. Special attention has been devoted to the stationary case with measurement noise of very small intensity [1]. The importance of this case stems from the following reasons

1) In many practical problems in communication the processes are stationary and the signal-to-noise ratio (SNR) is high [2]. In FM communication, for example, the threshold occurs at $\text{SNR} > 100$ [14].

2) The solution for this case, which is relatively easy to obtain, may indicate the maximal accuracy achievable in the estimation process [3].

3) It is shown later that the theory which is developed for cases of weak white measurement noise can be readily extended to deal with colored noise problems. The solution obtained by this approach is shown to be identical to results found by the commonly used algorithm of [4], but significantly easier to derive.

4) This problem is dual to the steady-state "cheap-control" problem of time invariant regulators that has been intensively investigated in the literature [3], [5], [6].

Significant contributions to the solution of this case have been made in the past [2], [5]–[8]. A comprehensive treatment of the single output process case is introduced in the present paper. The asymptotic properties of the closed-loop poles of the optimal filter [9], [10] and the s domain equivalent of the algebraic Riccati equation [11] are first used to derive explicit expressions for k , the optimal Kalman gain, in the case of single stochastic driving input and white measurement noise of small intensity ρ . Choosing a specific canonical state space representation for the asymptotically stable linear process, it is found that these gains are insensitive to the location of the open-loop poles, assuming, of course, that all these poles stay in the left half plane. By denoting the system order by n and the excess of poles over zeros in the system open-loop transfer function by l and $k' \triangleq (k_1, k_2, \dots, k_n)$, it is found that, in the limit as ρ tends to zero, $k_i = \alpha_i \rho^{-\min(i, l)/2}$, $0 \leq i \leq n$ where α_i is related to the input mapping matrix of the minimum-phase image of the system for $i \geq l$ and to the coefficient of the n th order Butterworth polynomial for $i < l$.

The expression for k is substituted in the algebraic Riccati equation to provide a Lyapunov type equation for the MECM (minimum error covariance matrix of estimate). From this equation the different powers of ρ according to which the estimation error covariance of the system output and its first $l-1$ derivatives tend to zero are found explicitly. The theory developed is then extended to also deal with systems of several inputs, and the obtained results are used to derive the solution of the colored measurements noise problem.

The paper is organized as follows. In Section II the explicit solution for the single input case is given. Minimum and nonminimum-phase systems are considered. In the minimum-phase case the pattern according to which the elements of the MECM approach asymptotically to zero is found. In the nonminimum-phase case the submatrix of the MECM whose elements always approach zero in the limit as the noise intensity tends to zero is located. The part of the MECM that does not tend to zero, in this limit, is also determined and it is related to the inherent structural difference that exists between the system and its minimum-phase image.

In Section III the theory is extended to cope with multiinput cases. The different asymptotic behavior of the elements of the MECM in this case is discussed and the similarity to the result found in the single-input nonminimum-phase case is demonstrated. The approach of this section is applied in Section IV to derive solutions to the singular colored measurement noise filtering problems. Two problems are treated. In the first, the intensity of the measurement noise is assumed to approach zero and results that are identical to the solution of the white noise problem are obtained. The second problem concerns the case where the colored measurement noise contains a white noise component of intensity that tends to zero. The asymptotic behavior of the MECM is discussed and the optimal filter, which applies finite gains only, is derived.

An example that illuminates different aspects of the proposed way of solution is given in Section II.

II. THE CASE OF SINGLE INPUT AND WHITE MEASUREMENT NOISE

Given the linear time-invariant system

$$\dot{x} = Ax + bw \quad (1a)$$

$$y = c'x \quad (1b)$$

with the observation process

Manuscript received November 20, 1978; revised December 7, 1979 and August 25, 1980. Paper recommended by A. Ephremides, Past Chairman of the Estimation Committee. The authors are with the Department of Electronic, Communications, Control and Computer Science, Tel-Aviv University, Tel-Aviv, Israel.