$A')^{-1}C'$ gives

$$CP(-SI-A')^{-1}C'+C(SI-A)^{-1}PC'+C(SI-A)^{-1}PC'R^{-1} \cdot CP(-SI-A')^{-1}C'-C(SI-A)^{-1}BQB'(-SI-A')^{-1}C'=0$$

which yields

$$[C(SI-A)^{-1}PC'R^{-1/2}+R^{1/2}][R^{-1/2}CP(-SI-A')^{-1}C'+R^{1/2}] = R+C(SI-A)^{-1}BQB'(-SI-A')^{-1}C'. \quad (10)$$

By virtue of the stabilizability and detectability properties of the system, there is a $\Delta(S)$ such that

$$\Delta(S)\Delta'(-S)=R+C(SI-A)^{-1}BQB'(-SI-A')^{-1}C' \quad (11)$$

where $\Delta(S)$ is square invertible matrix whose inverse is analytic in the right half of the S plane [5]. Therefore,

$$\Delta(S)=C(SI-A)^{-1}PC'R^{-1/2}+R^{1/2} \quad (12)$$

let

$$\hat{B}=PC'R^{-1/2} \quad (13)$$

then from (7)

$$K=\hat{B}R^{-1/2}. \quad (14)$$

Equations (12) and (13) give

$$\Delta(S)=C(SI-A)^{-1}\hat{B}+R^{1/2} \quad (15)$$

and substituting for K from (14) into (9) gives

$$H(S)=(SI-A)^{-1}\hat{B}R^{-1/2}[I+C(SI-A)^{-1}\hat{B}R^{-1/2}]^{-1} = (SI-A)^{-1}\hat{B}[R^{1/2}+C(SI-A)^{-1}\hat{B}]^{-1} \quad (16)$$

from (15)

$$H(S)=(SI-A)^{-1}\hat{B}\Delta^{-1}(S). \quad (17)$$

The numerical procedure of the algorithm is then reduced to solving (11) to obtain the spectral factor $\Delta(S)$, (15) to solve for $\hat{B}$, and (17) to compute the transfer matrix of the minimum variance estimator.

If R is equal to zero, these equations become

$$\Delta(S)\Delta'(-S)=C(SI-A)^{-1}BQB'(-SI-A')^{-1}C', \quad (18)$$

$$C(SI-A)^{-1}\hat{B}=\Delta(S) \quad (19)$$

and

$$H(S)=(SI-A)^{-1}\hat{B}\Delta^{-1}(S). \quad (20)$$

This shows that the perfect measurement system can be treated as a special case of the noisy system. Equations (18) through (20) are exactly the same as those derived in [2].

Conclusions

A unified solution to the transfer function of the minimum variance estimator of the state of a linear continuous-time system has been derived. It is shown that the perfect measurement system is a special case of the generally noise system. The solution in these special cases is identical to that obtained in [2].

Acknowledgment

The author wishes to thank R. S. Bucy for editing the manuscript and suggesting improvements.

References

[1] R. E. Kalman and R. S. Bucy, "New results in linear filtering and prediction theory," *Trans. ASME, Ser. D., J. Basic Eng.*, vol. 83, pp. 95-108, 1961.

[2] U. Shaked and E. Soroka, "A simple solution to the singular linear minimum-variance estimation problem," *IEEE Trans. Automat. Contr.*, vol. AC-32, pp. 81-83, Jan. 1987.

[3] R. S. Bucy and P. Joseph, *Filtering For Stochastic Processes with Application to Guidance*. New York: Interscience, 1968.

[4] A. G. J. MacFarlane, "Return-difference matrix properties for optimal stationary Kalman-Bucy filter," *Proc. IEE*, vol. 118, pp. 373-376, Feb. 1971.

[5] U. Shaked, "A general transfer function approach to linear stationary filtering and steady state optimal control problems," *Int. J. Contr.*, vol. 29, pp. 741-770, Oct. 1976.

[6] R. S. Bucy, "Linear and nonlinear filtering (Invited Paper)," *Proc. IEEE*, vol. 58, pp. 854-864, June 1970.

Robust, Reduced-Order, Nonstrictly Proper State Estimation Via the Optimal Projection Equations with Guaranteed Cost Bounds

WASSIM M. HADDAD AND DENNIS S. BERNSTEIN

Abstract—**A state-estimation design problem involving parametric plant uncertainties is considered. An estimation error bound suggested by multiplicative white noise modeling is utilized for guaranteeing robust estimation over a specified range of parameter uncertainties. Necessary conditions which generalize the optimal projection equations for reduced-order state estimation are used to characterize the estimator which minimizes the error bound. The design equations thus effectively serve as sufficient conditions for synthesizing robust estimators. Additional features include the presence of a static estimation gain in conjunction with the dynamic (Kalman) estimator to obtain a nonstrictly proper estimator.**

I. Introduction

As is well known [1]-[12], the performance of optimal filters based upon nominal parameter values may be severely degraded in the presence of parameter deviations. Thus, it is desirable to obtain *robust* state estimators which provide acceptable performance over the range of parametric uncertainty. The approach of the present paper is related to the guaranteed cost approach developed for control in [13], [14] and applied to estimation in [3]. Specifically, the main idea is to bound the effect of the uncertain parameters on the estimation error over the uncertainty range and then choose estimator gains to minimize the estimation bound. Thus, the *actual* estimation error is guaranteed to lie below the prescribed upper bound.

The technique used to determine minimizing estimator gains is a generalization of the optimal projection equations for reduced-order state estimation [15]. Thus, the results of the present paper effectively extend the results of [15] to the case of parameter uncertainties. It should be noted that the optimal projection equations, which are necessary conditions for optimality, now serve as *sufficient* conditions for robust estimation by virtue of the fact that a bound on the estimation error is being minimized rather than the estimation error itself. The bound utilized in the present paper was originally suggested by multiplicative white noise

Manuscript received June 8, 1987; revised September 25, 1987. This work was supported in part by the U.S. Air Force Office of Scientific Research under Contract F49620-86-C-0002.

W. M. Haddad is with the Department of Mechanical Engineering, Florida Institute of Technology, Melbourne, FL 32901.

D. S. Bernstein is with the Government Aerospace Systems Division, Harris Corporation, Melbourne, FL 32902.

IEEE Log Number 8718565.

modeling and was used in [16]–[18] for constructing Lyapunov functions for robust fixed-order dynamic compensation. A similar bound was used for full-state feedback in [19].

An additional feature of the present paper is the inclusion of a static feedback gain in conjunction with the dynamic estimator. Thus, the results of the present paper represent a generalization of standard results to the case of nonstrictly proper estimation. Similar treatments in the context of multiplicative noise models were given in [10] and [11] for discrete-time and continuous-time systems, respectively.

II. Notation and Definitions

Note: All matrices have real entries.

$\mathbb{R}$, $\mathbb{R}^{r\times s}$, $\mathbb{R}^r$	Real numbers, $r \times s$ real matrices, $\mathbb{R}^{r\times 1}$.
I_r, $(\ )^T$, $\mathbb{E}$	$r \times r$ identity matrix, transpose, expected value.
$\oplus$, $\otimes$	Kronecker sum, Kronecker product [20].
$\mathbb{S}^r$, $\mathbb{N}^r$, $\mathbb{P}^r$	$r \times r$ symmetric, nonnegative-definite, positive-definite matrices.
$Z_1 \le Z_2$, $Z_1 < Z_2$	$Z_2 - Z_1 \in \mathbb{N}^r$, $Z_2 - Z_1 \in \mathbb{P}^r$, $Z_1, Z_2 \in \mathbb{S}^r$.
$n, l, \hat{l}, n_e, p, q; \tilde{n}$	Positive integers; $n + n_e$.
$x, y, \hat{y}, y_e, x_e, \tilde{x}$	$n, l, \hat{l}, q, n_e, \tilde{n}$-dimensional vectors.
$A, \Delta A; C, \Delta C$	$n \times n$ matrices; $l \times n$ matrices.
$\hat{C}$	$\hat{l} \times n$ matrix.
A_e, B_e, C_e, D_e	$n_e \times n_e$, $n_e \times l$, $q \times n_e$, $q \times \hat{l}$ matrices.
$\tilde{A}, \Delta\tilde{A}$	$\begin{bmatrix} A & 0 \\ B_eC & A_e \end{bmatrix}, \begin{bmatrix} \Delta A & 0 \\ B_e\Delta C & 0 \end{bmatrix}$.
L, R	$q \times n$ matrix, estimation-error weighting in $\mathbb{P}^q$.
$\tilde{R}$	$\begin{bmatrix} L^TRL - L^TRD_e\hat{C} - \hat{C}^TD_e^TRL + \hat{C}^TD_e^TRD_e\hat{C} & -L^TRC_e + C^TD_e^TRC_e \\ -C_e^TRL + C_e^TRD_e\hat{C} & C_e^TRC_e \end{bmatrix}$.
$w_1(\cdot), w_2(\cdot)$	n, l-dimensional white noise.
V_1, V_2	Intensity of $w_1(\cdot)$, $w_2(\cdot)$; $V_1 \in \mathbb{N}^n$, $V_2 \in \mathbb{P}^l$.
V_{12}	Cross intensity of $w_1(\cdot)$, $w_2(\cdot)$.
$\tilde{w}(\cdot), \tilde{V}$	$\begin{bmatrix} w_1(\cdot) \\ B_ew_2(\cdot) \end{bmatrix}, \begin{bmatrix} V_1 & V_{12}B_e^T \\ B_eV_{12}^T & B_eV_2B_e^T \end{bmatrix}$.
α	Positive number.
α_i	Positive number, $i = 1, \cdots, p$.
σ_i	Real number, $i = 1, \cdots, p$.
γ_i	α_i^2/α, $i = 1, \cdots, p$.
$A_\alpha, A_{e\alpha}$	$A + (\alpha/2)I_n$, $A_e + (\alpha/2)I_{n_e}$.
$\tilde{A}_\alpha$	$\begin{bmatrix} A_\alpha & 0 \\ B_eC & A_{e\alpha} \end{bmatrix}$.

III. Robust Estimation Problem

Let $\mathfrak{U} \subset \mathbb{R}^{n\times n} \times \mathbb{R}^{l\times n}$ denote the set of uncertain perturbations $(\Delta A, \Delta C)$ of the nominal plant matrices A and C.

Robust Estimation Problem: For fixed $n_e \le n$, determine (A_e, B_e, C_e, D_e) such that, for the system consisting of the nth-order disturbed plant

$$\dot{x}(t) = (A + \Delta A)x(t) + w_1(t), \qquad t \in [0, \infty), \tag{3.1}$$

noisy and nonnoisy measurements

$$y(t) = (C + \Delta C)x(t) + w_2(t), \tag{3.2}$$

$$\hat{y}(t) = \hat{C}x(t), \tag{3.3}$$

and n_eth-order nonstrictly proper state estimator

$$\dot{x}_e(t) = A_ex_e(t) + B_ey(t), \tag{3.4}$$

$$y_e(t) = C_ex_e(t) + D_e\hat{y}(t), \tag{3.5}$$

the state-estimation error criterion

$$J(A_e, B_e, C_e, D_e) \triangleq \sup_{(\Delta A, \Delta C)\in\mathfrak{U}} \limsup_{t\to\infty} \mathbb{E}[Lx(t) - y_e(t)]^T R[Lx(t) - y_e(t)] \tag{3.6}$$

is minimized.

For each estimator (A_e, B_e, C_e, D_e) and system variation $(\Delta A, \Delta C) \in \mathfrak{U}$, the disturbed augmented system (3.1)–(3.5) is given by

$$\dot{\tilde{x}}(t) = (\tilde{A} + \Delta\tilde{A})\tilde{x}(t) + \tilde{w}(t), \qquad t \in [0, \infty) \tag{3.7}$$

where $\tilde{x}(t) \triangleq [x^T(t), x_e^T(t)]^T$ and $\tilde{w}(t)$ has intensity $\tilde{V} \in \mathbb{N}^{\tilde{n}}$. The cost can be expressed in terms of the second-moment matrix.

Proposition 3.1: For given (A_e, B_e, C_e, D_e) and $(\Delta A, \Delta C) \in \mathfrak{U}$, the second-moment matrix

$$\tilde{Q}_{\Delta\tilde{A}}(t) \triangleq \mathbb{E}[\tilde{x}(t)\tilde{x}^T(t)], \qquad t \in [0, \infty) \tag{3.8}$$

satisfies

$$\dot{\tilde{Q}}_{\Delta\tilde{A}}(t) = (\tilde{A} + \Delta\tilde{A})\tilde{Q}_{\Delta\tilde{A}}(t) + \tilde{Q}_{\Delta\tilde{A}}(t)(\tilde{A} + \Delta\tilde{A})^T + \tilde{V}, \qquad t \in [0, \infty). \tag{3.9}$$

Furthermore,

$$J(A_e, B_e, C_e, D_e) \triangleq \sup_{(\Delta A, \Delta C)\in\mathfrak{U}} \limsup_{t\to\infty} \operatorname{tr} \tilde{Q}_{\Delta\tilde{A}}(t)\tilde{R}. \tag{3.10}$$

IV. Sufficient Conditions for Robust Performance

The following result is immediate.

Lemma 4.1: Suppose $\tilde{A} + \Delta\tilde{A}$ is stable for all $(\Delta A, \Delta C) \in \mathfrak{U}$. Then

$$J(A_e, B_e, C_e, D_e) = \sup_{(\Delta A, \Delta C)\in\mathfrak{U}} \operatorname{tr} \tilde{Q}_{\Delta\tilde{A}}\tilde{R} \tag{4.1}$$

where $\tilde{Q}_{\Delta\tilde{A}} \in \mathbb{N}^{\tilde{n}}$ is the unique solution to

$$0=(\tilde{A}+\Delta\tilde{A})\tilde{Q}_{\Delta\tilde{A}}+\tilde{Q}_{\Delta\tilde{A}}(\tilde{A}+\Delta\tilde{A})^T+\tilde{V}. \tag{4.2}$$

We seek upper bounds for $J(A_e, B_e, C_e, D_e)$.

Theorem 4.1: Let $\Omega:\mathbb{N}^{\tilde{n}} \times \mathbb{R}^{n_e\times l} \to \mathbb{S}^{\tilde{n}}$ be such that

$$(\Delta A, \Delta C) \in \mathcal{U},$$
$$\Delta\tilde{A}\mathcal{Q}+\mathcal{Q}\Delta\tilde{A}^T\le\Omega(\mathcal{Q}, B_e),\ (\mathcal{Q}, B_e) \in \mathbb{N}^{\tilde{n}}\times\mathbb{R}^{n_e\times l} \tag{4.3}$$

and, for given (A_e, B_e, C_e, D_e), suppose there exists $\mathcal{Q} \in \mathbb{N}^{\tilde{n}}$ satisfying

$$0=\tilde{A}\mathcal{Q}+\mathcal{Q}\tilde{A}^T+\Omega(\mathcal{Q}, B_e)+\tilde{V}, \tag{4.4}$$

and suppose the pair $(\tilde{V}^{1/2}, \tilde{A} + \Delta\tilde{A})$ is stabilizable for all $(\Delta A, \Delta C) \in \mathcal{U}$. Then A_e is asymptotically stable, $A + \Delta A$ is asymptotically stable for all $(\Delta A, \Delta C) \in \mathcal{U}$,

$$\tilde{Q}_{\Delta\tilde{A}}\le\mathcal{Q}, \qquad (\Delta A, \Delta C) \in \mathcal{U} \tag{4.5}$$

where $\tilde{Q}_{\Delta\tilde{A}}$ satisfies (4.2) and

$$J(A_e, B_e, C_e, D_e)\le\text{tr } \mathcal{Q}\tilde{R}. \tag{4.6}$$

Proof: For all $(\Delta A, \Delta C) \in \mathcal{U}$, (4.4) is equivalent to

$$0=(\tilde{A}+\Delta\tilde{A})\mathcal{Q}+\mathcal{Q}(\tilde{A}+\Delta\tilde{A})^T+\Psi(\mathcal{Q}, B_e, \Delta\tilde{A})+\tilde{V} \tag{4.7}$$

where

$$\Psi(\mathcal{Q}, B_e, \Delta\tilde{A}) \triangleq \Omega(\mathcal{Q}, B_e)-(\Delta\tilde{A}\mathcal{Q}+\mathcal{Q}\Delta\tilde{A}^T).$$

Note that by (4.3), $\Psi(\mathcal{Q}, B_e, \Delta\tilde{A}) \ge 0$ for all $(\Delta A, \Delta C) \in \mathcal{U}$. Since $(\tilde{V}^{1/2}, \tilde{A} + \Delta\tilde{A})$ is stabilizable for all $(\Delta A, \Delta C) \in \mathcal{U}$, it follows from [21, Theorem 3.6] that $((\tilde{V}+\Psi(\mathcal{Q}, B_e, \Delta\tilde{A}))^{1/2}, \tilde{A}+\Delta\tilde{A})$ is stabilizable for all $(\Delta A, \Delta C) \in \mathcal{U}$. Hence, [21, Lemma 12.2] implies $\tilde{A} + \Delta\tilde{A}$ is asymptotically stable for all $(\Delta A, \Delta C) \in \mathcal{U}$. Since $\tilde{A} + \Delta\tilde{A}$ is lower block triangular, A_e is asymptotically stable and $A + \Delta A$ is asymptotically stable for all $(\Delta A, \Delta C) \in \mathcal{U}$. Next, (4.7) minus (4.2) yields

$$0=(\tilde{A}+\Delta\tilde{A})(\mathcal{Q}-\tilde{Q}_{\Delta\tilde{A}})+(\mathcal{Q}-\tilde{Q}_{\Delta\tilde{A}})(\tilde{A}+\Delta\tilde{A})^T+\Psi(\mathcal{Q}, B_e, \Delta\tilde{A})$$

or, equivalently (since $\tilde{A} + \Delta\tilde{A}$ is asymptotically stable),

$$\mathcal{Q}-\tilde{Q}_{\Delta\tilde{A}}=\int_0^\infty e^{(\tilde{A}+\Delta\tilde{A})t}\Psi(\mathcal{Q}, B_e, \Delta\tilde{A})e^{(\tilde{A}+\Delta\tilde{A})^Tt}\,dt\ge0,$$

which implies (4.5). Finally, (4.5) and (4.1) yield (4.6). □

V. Uncertainty Structure and Guaranteed Cost Bound

The uncertainty set $\mathcal{U}$ is assumed to be of the form

$$\mathcal{U} \triangleq \left\{(\Delta A, \Delta C) \in \mathbb{R}^{n\times n}\times\mathbb{R}^{l\times n} : \Delta A=\sum_{i=1}^{p}\sigma_iA_i,\ \Delta C =\sum_{i=1}^{p} \sigma_iC_i,\ \sum_{i=1}^{p} \sigma_i^2/\alpha_i^2\le1\right\} \tag{5.1}$$

where, for $i = 1, \cdots, p$: $A_i \in \mathbb{R}^{n\times n}$ and $C_i \in \mathbb{R}^{l\times n}$ are fixed matrices denoting the structure of the parametric uncertainty in the dynamics and measurement matrices; α_i is a given positive number; and σ_i is an uncertain real parameter. In practice, the form of ΔA and ΔC permits the modeling of linear parameter uncertainties of arbitrary structure. Note that the uncertain parameters σ_i are assumed to lie in a specified ellipsoidal region in $\mathbb{R}^p$. The augmented system thus has structured uncertainty of the form

$$\Delta\tilde{A}=\sum_{i=1}^{p} \sigma_i\tilde{A}_i \tag{5.2}$$

where

$$\tilde{A}_i \triangleq \begin{bmatrix} A_i & 0 \\ B_eC_i & 0\end{bmatrix}, \qquad i=1, \cdots, p.$$

Remark 5.1: Note that (5.1) allows a particular parameter σ_i to appear in both ΔA and ΔC. Thus, it is possible to consider the case in which the uncertainties ΔA and ΔC are known to be correlated. Of course, for a given i, A_i or C_i can be set to zero so that the similar form of ΔA and ΔC represents no restriction.

We now specify the bounding function Ω satisfying (4.3).

Proposition 5.1: Let α be an arbitrary positive scalar. Then the function

$$\Omega(\mathcal{Q}, B_e) \triangleq \alpha\mathcal{Q}+\alpha^{-1} \sum_{i=1}^{p}\alpha_i^2\tilde{A}_i\mathcal{Q}\tilde{A}_i^T \tag{5.3}$$

satisfies (4.3) with $\mathcal{U}$ given by (5.1).

Proof: Note that

$$0\le\sum_{i=1}^{p} [(\alpha^{1/2}\sigma_i/\alpha_i)I_{\tilde{n}}-(\alpha_i/\alpha^{1/2})\tilde{A}_i] \cdot \mathcal{Q}[(\alpha^{1/2}\sigma_i/\alpha_i)I_{\tilde{n}}-(\alpha_i/\alpha^{1/2})\tilde{A}_i]^T$$
$$=\alpha \sum_{i=1}^{p} (\sigma_i^2/\alpha_i^2)\mathcal{Q}+\alpha^{-1} \sum_{i=1}^{p} \alpha_i^2\tilde{A}_i\mathcal{Q}\tilde{A}_i^T -\sum_{i=1}^{p} \sigma_i(\tilde{A}_i\mathcal{Q}+\mathcal{Q}\tilde{A}_i^T)$$

which, since $\Sigma_{i=1}^p\ \sigma_i^2/\alpha_i^2 \le 1$, implies (4.3). □

Remark 5.2: Note that with (5.3), the modified Lyapunov equation (4.4) becomes

$$0=\tilde{A}_\alpha\mathcal{Q}+\mathcal{Q}\tilde{A}_\alpha^T+\sum_{i=1}^{p} \gamma_i\tilde{A}_i\mathcal{Q}\tilde{A}_i^T+\tilde{V}. \tag{5.4}$$

VI. The Auxiliary Minimization Problem

Our goal is to minimize the error bound (4.6).

Auxiliary Minimization Problem: Determine $(\mathcal{Q}, A_e, B_e, C_e, D_e)$ with $\mathcal{Q} \in \mathbb{N}^{\tilde{n}}$ which minimizes

$$\mathcal{J}(\mathcal{Q}, A_e, B_e, C_e, D_e) \triangleq \text{tr } \mathcal{Q}\tilde{R} \tag{6.1}$$

subject to (5.4) and

$$(\tilde{V}^{1/2}, \tilde{A}+\Delta\tilde{A}) \text{ is stabilizable,} \qquad (\Delta A, \Delta C) \in \mathcal{U}. \tag{6.2}$$

Proposition 6.1: If $(\mathcal{Q}, A_e, B_e, C_e, D_e)$ satisfies (5.4) and (6.2) with $\mathcal{Q} \ge 0$, then $\tilde{A} + \Delta\tilde{A}$ is asymptotically stable for all $(\Delta A, \Delta C) \in \mathcal{U}$ and

$$J(A_e, B_e, C_e, D_e)\le\mathcal{J}(\mathcal{Q}, A_e,B_e, C_e, D_e). \tag{6.3}$$

Proof: With Ω given by (5.3), Proposition 5.1 implies that (4.3) is satisfied. Hence, with (6.2), the hypotheses of Theorem 4.1 are satisfied so that the system (3.7) is stable over $\mathcal{U}$ with estimation bound (4.6). Note that (6.3) is merely a restatement of (4.6). □

Remark 6.1: The conservatism of the bound (6.3) is difficult to predict for two reasons. First, the overbounding (4.3) holds with respect to the partial ordering of the nonnegative-definite matrices for which no scalar measure of conservatism is available. And second, the bound (4.3) is required to hold for all nonnegative-definite matrices $\mathcal{Q}$ and estimator gains B_e. The conservatism will thus depend upon the actual values of $\mathcal{Q}$ and B_e determined by solving (5.4).

VII. Necessary Conditions for the Auxiliary Minimization Problem

Rigorous application of the Lagrange multiplier technique requires additional technical assumptions. Specifically, we further restrict $(\mathcal{Q}, A_e,$

B_e, C_e, D_e) to the open set

$$\mathcal{S} \triangleq \{(\mathcal{Q}, A_e, B_e, C_e, D_e) : \mathcal{Q} \in \mathbb{P}^{\tilde{n}},$$
$$\tilde{\mathfrak{A}} \text{ is asymptotically stable,}$$
$$(A_e, B_e, C_e) \text{ is controllable and observable, and}$$
$$\hat{C}(Q_1 - Q_{12}Q_2^{-1}Q_{12}^T)\hat{C}^T > 0\}$$

where

$$\tilde{\mathfrak{A}} \triangleq \tilde{A}_\alpha \oplus \tilde{A}_\alpha + \sum_{i=1}^{p} \gamma_i \tilde{A}_i \otimes \tilde{A}_i$$

and $\mathcal{Q}$ is partitioned as

$$\mathcal{Q} \triangleq \begin{bmatrix} Q_1 & Q_{12} \\ Q_{12}^T & Q_2 \end{bmatrix}$$

where Q_1, Q_{12}, and Q_2 are $n \times n$, $n \times n_e$, and $n_e \times n_e$, respectively. As shown in [11], Q_2 is invertible since (A_e, B_e) is controllable. The definiteness condition holds when $\hat{C}$ has full row rank and $\mathcal{Q}$ is positive definite. As shown in [11], this condition implies the existence of the projection $\hat{\tau}$ defined below.

Remark 7.1: Proposition 6.1 shows that the constraint $(\mathcal{Q}, A_e, B_e, C_e, D_e) \in \mathcal{S}$ is not required for robust estimation. As can be seen from the proof given in [11], the set $\mathcal{S}$ constitutes sufficient conditions under which the Lagrange multiplier technique is applicable to the auxiliary minimization problem. Specifically, $\mathcal{Q} \in \mathbb{P}^{\tilde{n}}$ replaces $\mathcal{Q} \in \mathbb{N}^{\tilde{n}}$ by an open set constraint, while asymptotic stability of $\tilde{\mathfrak{A}}$ serves as a normality condition which further implies that the dual $\mathcal{P}$ of $\mathcal{Q}$ is nonnegative definite. Thus, it is *not* necessary for guaranteed robust estimation that an admissible quadruple obtained by solving the necessary conditions actually be shown to be an element of $\mathcal{S}$.

The following factorization lemma is needed for the statement of the main result. For details, see [15].

Lemma 7.1: If $\hat{Q}, \hat{P} \in \mathbb{N}^n$ and rank $\hat{Q}\hat{P} = n_e$, then there exist $n_e \times nG$, Γ, and $n_e \times n_e$ invertible M such that

$$\hat{Q}\hat{P} = G^T M \Gamma \tag{7.1}$$

$$\Gamma G^T = I_{n_e}. \tag{7.2}$$

Recall from [15] that

$$\tau \triangleq \hat{Q}\hat{P}(\hat{Q}\hat{P})^{\#} = G^T\Gamma \tag{7.3}$$

is an oblique projection. Define the complementary projection $\tau_\perp \triangleq I_n - \tau$ and call (G, M, Γ) satisfying (7.1), (7.2) a *projective factorization* of $\hat{Q}\hat{P}$. Furthermore, for arbitrary $Q, \hat{Q} \in \mathbb{R}^{n\times n}$, define the notation

$$V_{2s} \triangleq V_2 + \sum_{i=1}^{p} \gamma_i C_i(Q+\hat{Q})C_i^T.$$

$$Q_s \triangleq QC^T + V_{12} + \sum_{i=1}^{p} \gamma_i A_i(Q+\hat{Q})C_i^T.$$

Theorem 7.1: If $(\mathcal{Q}, A_e, B_e, C_e, D_e) \in \mathcal{S}$ solves the auxiliary minimization problem with $\mathcal{U}$ given by (5.1) and Ω given by (5.3), then there exist $Q, \hat{Q}, \hat{P} \in \mathbb{N}^n$ such that, for some projective factorization (G, M, Γ) of $\hat{Q}\hat{P}$, $(\mathcal{Q}, A_e, B_e, C_e, D_e)$ are given by

$$\mathcal{Q} = \begin{bmatrix} Q+\hat{Q} & \hat{Q}\Gamma^T \\ \Gamma\hat{Q} & \Gamma\hat{Q}\Gamma^T \end{bmatrix}, \tag{7.4}$$

$$A_e = \Gamma(A - Q_s V_{2s}^{-1} C)G^T, \tag{7.5}$$

$$B_e = \Gamma Q_s V_{2s}^{-1}, \tag{7.6}$$

$$C_e = L\hat{\tau}_\perp G^T, \tag{7.7}$$

$$D_e = LQ\hat{C}^T(\hat{C}Q\hat{C}^T)^{-1}, \tag{7.8}$$

and such that Q, $\hat{Q}$, and $\hat{P}$ satisfy

$$0 = A_\alpha Q + QA_\alpha^T + V_1 + \sum_{i=1}^{p} \gamma_i A_i(Q+\hat{Q})A_i^T - Q_s V_{2s}^{-1} Q_s^T + \tau_\perp Q_s V_{2s}^{-1} Q_s^T \tau_\perp^T, \tag{7.9}$$

$$0 = A_\alpha \hat{Q} + \hat{Q}A_\alpha^T + Q_s V_{2s}^{-1} Q_s^T - \tau_\perp Q_s V_{2s}^{-1} Q_s^T \tau_\perp^T, \tag{7.10}$$

$$0 = (A_\alpha - Q_s V_{2s}^{-1} C)^T \hat{P} + \hat{P}(A_\alpha - Q_s V_{2s}^{-1} C) + \hat{\tau}_\perp^T L^T RL\hat{\tau}_\perp - \tau_\perp^T \hat{\tau}_\perp^T L^T RL \hat{\tau}_\perp \tau_\perp, \tag{7.11}$$

$$\text{rank } \hat{Q} = \text{rank } \hat{P} = \text{rank } \hat{Q}\hat{P} = n_e \tag{7.12}$$

where

$$\hat{\tau} \triangleq Q\hat{C}^T(\hat{C}Q\hat{C}^T)^{-1}\hat{C}, \quad \hat{\tau}_\perp \triangleq I_n - \hat{\tau}. \tag{7.13}$$

Furthermore, the auxiliary cost is given by

$$\mathcal{J}(\mathcal{Q}, A_e, B_e, C_e, D_e) = \text{tr } Q\hat{\tau}_\perp^T L^T RL\hat{\tau}_\perp. \tag{7.14}$$

Conversely, if there exist $Q, \hat{Q}, \hat{P} \in \mathbb{N}^n$ satisfying (7.9)–(7.12), then $(\mathcal{Q}, A_e, B_e, C_e, D_e)$ given by (7.4)–(7.8) satisfy (5.4) with $\mathcal{Q} \in \mathbb{N}^{\tilde{n}}$ and with $\mathcal{J}(\mathcal{Q}, A_e, B_e, C_e, D_e)$ given by (7.14).

Proof: The derivation requires only a minor modification of the derivation given in [11]. The only change involves treatment of $\tilde{A}_\alpha$ in place of $\tilde{A}$. □

Remark 7.1: The necessary conditions given in Theorem 7.1 directly generalize the result given in [15]. To recover the result of [15], set $A_i = 0$, $C_i = 0$, $i = 1, \cdots, p$ (to delete the plant uncertainties), and set $\hat{C} = 0$ (to eliminate the static estimation term D_e). It follows from the proof given in [11] that $\hat{C} = 0$ yields $\hat{\tau} = 0$, and thus $\hat{\tau}_\perp = I_n$.

Remark 7.2: Note that $\mathcal{Q}$ given by (7.4) is nonnegative definite.

VIII. Sufficient Conditions for Robust, Reduced-Order Estimation

The main result guaranteeing robust estimation can now be stated.

Theorem 8.1: Suppose there exist $Q, \hat{Q}, \hat{P} \in \mathbb{N}^n$ satisfying (7.9)–(7.12), let A_e, B_e, C_e, D_e be given by (7.5)–(7.8), and suppose that $(\tilde{V}^{1/2}, \tilde{A} + \Delta\tilde{A})$ is stabilizable for all $(\Delta A, \Delta C) \in \mathcal{U}$ with $\mathcal{U}$ given by (5.1). Then A_e is asymptotically stable, $A + \Delta A$ is asymptotically stable for all $(\Delta A, \Delta C) \in \mathcal{U}$, and the estimation error satisfies the performance bound

$$J(A_e, B_e, C_e, D_e) \le \text{tr } Q\hat{\tau}_\perp^T L^T RL\hat{\tau}_\perp. \tag{8.1}$$

Proof: Theorem 7.1 and Remark 7.2 imply that $\mathcal{Q}$ given by (7.4) is nonnegative definite and satisfies (5.4). With the stabilizability assumption, the result follows from Proposition 6.1. □

Remark 8.1: Suppose $\hat{l} = n$, $\hat{C} = I_n$ (so that perfect measurements of the entire state are available), and Q satisfying (7.9) is positive definite. Then it follows from Theorem 7.1 that $\hat{\tau} = I_n$, $\hat{\tau}_\perp = 0$, $C_e = 0$ (i.e., the dynamic filter is disabled), $D_e = L$, and by (8.1), $J = 0$. This is, of course, the expected result since perfect estimation is achievable in this case.

References

[1] J. A. D'Appolito and C. E. Hutchinson, "Low sensitivity filters for state estimation in the presence of large parameter uncertainties," *IEEE Trans. Automat. Contr.*, vol. AC-14, pp. 310–312, 1969.

[2] P. J. McLane, "Optimal linear filtering for linear systems with state-dependent noise," *Int. J. Contr.*, vol. 10, pp. 41–51, 1969.

[3] B. N. Jain, "Guaranteed error estimation in uncertain systems," *IEEE Trans. Automat. Contr.*, vol. AC-20, pp. 230–232, 1975.

[4] R. C. Chung and P. R. Belanger, "Minimum-sensitivity filter for linear time-invariant stochastic systems with uncertain parameters," *IEEE Trans. Automat. Contr.*, vol. AC-21, pp. 98–100, 1976.
[5] K. Furuta, S. Hara, and S. Mori, "A class of systems with the same observer," *IEEE Trans. Automat. Contr.*, vol. AC-21, pp. 572–576, 1976.
[6] S. P. Bhattacharyya, "The structure of robust observers," *IEEE Trans. Automat. Contr.*, vol. AC-21, pp. 581–588, 1976.
[7] M. Toda and R. V. Patel, "Bounds on estimation errors of discrete-time filters under modeling uncertainty," *IEEE Trans. Automat. Contr.*, vol. AC-25, pp. 1115–1121, 1980.
[8] R. T. Stefani, "Reducing the sensitivity to parameter variations of a minimum-order reduced-order observer," *Int. J. Contr.*, vol. 35, pp. 983–995, 1982.
[9] T. E. Djaferis, "Robust observers for systems with parameters," *Syst. Contr. Lett.*, vol. 7, pp. 385–394, 1986.
[10] W. M. Haddad and D. S. Bernstein, "The optimal projection equations for discrete-time reduced-order state estimation for linear systems with multiplicative white noise," *Syst. Contr. Lett.*, vol. 8, pp. 381–388, 1987.
[11] ——, "The optimal projection equations for reduced-order state estimation: The singular measurement noise case," *IEEE Trans. Automat. Contr.*, vol. AC-32, pp. 1135–1139, 1987.
[12] D. S. Bernstein and D. C. Hyland, "The optimal projection equations for reduced-order modelling, estimation and control of linear systems with multiplicative white noise," *J. Optimiz. Theory Appl.*, to be published.
[13] S. S. L. Chang and T. K. C. Peng, "Adaptive guaranteed cost control of systems with uncertain parameters," *IEEE Trans. Automat. Contr.*, vol. AC-17, pp. 474–483, 1972.
[14] A. Vinkler and L. J. Wood, "Multistep guaranteed cost control of linear systems with uncertain parameters," *J. Guidance Contr.*, vol. 2, pp. 449–456, 1979.
[15] D. S. Bernstein and D. C. Hyland, "The optimal projection equations for reduced-order state estimation," *IEEE Trans. Automat. Contr.*, vol. AC-30, pp. 583–585, 1985.
[16] D. S. Bernstein and S. W. Greeley, "Robust output-feedback stabilization: Deterministic and stochastic perspectives," in *Proc. Amer. Contr. Conf.*, Seattle, WA, June 1986, pp. 1818–1826.
[17] W. M. Haddad, "Robust optimal projection control-system synthesis," Ph.D. dissertation, Dept. Mech. Eng., Florida Inst. Technol., Melbourne, Mar. 1987.
[18] D. S. Bernstein, "Robust static and dynamic output-feedback stabilization: Deterministic and stochastic perspectives," *IEEE Trans. Automat. Contr.*, vol. AC-32, pp. 1076–1084, 1987.
[19] O. I. Kosmidou and P. Bertrand, "Robust-controller design for systems with large parameter variations," *Int. J. Contr.*, vol. 45, pp. 927–938, 1987.
[20] J. W. Brewer, "Kronecker products and matrix calculus in system theory," *IEEE Trans. Circuits Syst.*, vol. CAS-25, pp. 772–781, 1978.
[21] W. M. Wonham, *Linear Multivariable Control: A Geometric Approach.* New York: Springer-Verlag, 1979.

Approximate and Limit Results for Nonlinear Filters with Small Observation Noise: The Linear Sensor and Constant Diffusion Coefficient Case

OFER ZEITOUNI

Abstract—**Recursive approximations for a class of filtering problems are presented. This class is characterized by linear observation sensor, constant diffusion terms, and for the multidimensional problem, potential-like conditions on the drift. For the case of small observation noise, these approximations are used to demonstrate the Gaussian limiting structure of the optimal nonlinear filter.**

I. Introduction

The classical nonlinear filtering problem is of the form

$$dx_t = f(x_t)dt + \sigma(x_t)dw_t,\ x_t \in \mathbb{R}^n, \qquad p(x_0) = p_0(x_0) \tag{1.1}$$

$$dy_t = g(x_t)dt + N_0^{1/2}d\theta_t,\ y_t \in \mathbb{R}^m, \qquad y_0 = 0 \tag{1.2}$$

Manuscript received February 5, 1987; revised September 2, 1987. Paper recommended by Associate Editor, T. J. Tarn. This work was supported in part by a Weizmann Postdoctoral Fellowship.

The author was with the Technion—Israel Institute of Technology, Haifa 32000, Israel. He is now with LIDS, Massachusetts Institute of Technology, Cambridge, MA 02139.

IEEE Log Number 8718566.

where w_t, θ_t are independent Brownian motions and the filtering problem consists of computing statistics of x_t when the observation σ algebra $y_0^t \triangleq \{y_s,\ 0 \leqslant s \leqslant t\}$ is given. By now, it is clear that for all but a few problems, an explicit final dimensional solution does not exist [4]. Therefore, one is led to consider approximations and to consider simplified limiting cases. Especially, the low observation noise case ($N_0 \to 0$) has been considered in the literature [3], [6], [11].

In this paper, we restrict out attention to the special case of linear observations and constant diffusion coefficients, i.e.,

$$dx_t = f(x_t)dt + \sigma dw_t,\ x_t \in \mathbb{R}^n, \qquad p(x_0) = p_0(x_0) \tag{1.1'}$$

$$dy_t = gx_t dt + N_0^{1/2}d\theta_t,\ y_t \in \mathbb{R}^m, \qquad y_0 = 0 \tag{1.2'}$$

where σ, g are matrices of appropriate dimensions. In the multidimensional case, we impose some additional potential-like structural conditions on $f(\cdot)$. For this restricted class of filtering problems, we derive recursive approximations to the conditional density $p_{x_t}(z|y_t^0)$ and to its unnormalized version $\rho_{x_t}(z|y_t^0)$. For the limiting case $N_0 \to 0$, those approximations are used to show that the conditional density, rescaled in a suitable manner, converges to a Gaussian density, with tight estimates on the "tails" of the density. This fact demonstrates the usefulness of Gaussian-based approximations (like the extended Kalman filter or the second-order Gaussian filter).

Related results were obtained by Mayer-Wolf [9] in his dissertation. There, bounds on the filtering error and the Cramer–Rao inequality are used to prove a basic Gaussian limit result, although under different assumptions.

The paper is organized as follows. The one-dimensional problem ($n = m = 1$) is treated in Sections II and III. In Section II, we present our basic approximation theorem, which holds whether N_0 is small or not. We further demonstrate that, if $N_0 \to 0$ the approximations exhibit certain nice limiting behavior, then the rescaled conditional density converges (weakly and pointwise) to a Gaussian density. In Section III, we check out explicitly the limiting behavior of the approximations and derive explicit conditions on $f(\cdot)$ under which the density indeed converges to a Gaussian one. Finally, in Section IV, we extend our results to a class of multidimensional problems.

We make throughout, the following assumption.

(A1) $f(\cdot)$ is continuously differentiable with bounded first partial derivatives.

II. An Approximation Theorem—The One-Dimensional Case

In this section, an approximation theorem for the unnormalized conditional density $\rho(z|y_t^0)$ is presented. Throughout, the one-dimensional case is treated. Multidimensional extensions are postponed to Section IV.

Without loss of generality, we assume $\sigma = 1$ in (1.1′). Recall that under (A-1), a solution to (1.1′) exists and is unique. Moreover, the measure P_1 defined by the pair (1.1′), (1.2′) is absolutely continuous w.r.t. the reference measure P_0 defined by

$$dx_t = \alpha x_t dt + dw_t, \qquad p(x_0) = p_0(x_0) \tag{2.1}$$

$$dy_t = N_0^{1/2}d\theta_t, \qquad y_0 = 0 \tag{2.2}$$

where α is some constant to be defined. The Radon–Nikodym derivative dP_1/dP_0 is [8].

$$\frac{dP_1}{dP_0} = \exp\left(\int_0^t (f(x_s) - \alpha x_s)\,dx_s - 1/2\int_0^t (f^2(x_s) - \alpha^2 x_s^2)\,ds + \int_0^t N_0^{-1} g x_s\,dy_s - 1/2\int_0^t N_0^{-1} g^2 x_s^2\,ds\right). \tag{2.3}$$

As is well known [2], [12], [13], the unnormalized conditional density $\rho(z|y_t^0)$ satisfies

$$\rho(z|y_0^t) \propto E_0\left(\frac{dP_1}{dP_0}\bigg|\, y_0^t,\ x_t = z\right) p_{x_t}(z) \tag{2.4}$$

0018-9286/88/0600-0595$01.00 © 1988 IEEE