Theorem 3: Let $u(s, \epsilon)$ be the solution to the full-order NP problem (1)–(2) with the initial data given by Theorem 2, and let $u^{SF}(s, \epsilon)$ be the transfer function constructed in the aforementioned algorithm. Then for sufficiently small ϵ

1) $u^{SF}(s, \epsilon)$ is a TFS–SBR transfer function, and
2) $\|u(s, \epsilon) - u^{SF}(s, \epsilon)\|_\infty = O(\epsilon)$.

Proof: By Theorem 2, $u(s, \epsilon)$ is a TFS–SBR transfer function. Also, $u^{SF}(s, \epsilon) \in T_\epsilon$ because it has the same slow and fast transfer functions as $u(s, \epsilon)$. Since both of these slow and fast transfer functions are SBR functions, by Lemma 1, $u^{SF}(s, \epsilon)$ is also an SBR transfer function. This proves part 1 of the theorem. Part 2 follows from [9, Theorem 4.1].

Q.E.D.

V. Conclusion

In this note, we have considered a version of the NP interpolation problem. We have obtained conditions which guarantee a TFS solution. We have formulated two smaller NP interpolation problems and have shown that they can be solved in parallel. We have reduced the solvability condition in terms of the solvability conditions for these two smaller problems. The immediate gain here is the reduction in the verification process and hence the computer time. We have used the solutions obtained from these two smaller problems and have constructed a solution which can become arbitrarily close in the H_∞-norm to the solution of the original NP problem. The results of this note should prove useful both from the computational standpoint as well as in the areas such as robust stabilization problem for TFS systems.

References

[1] P. Delsarte, Y. Genin, and Y. Kamp, "On the role of the Nevanlinna–Pick problem in circuit and system theory," *Int. J. Circuit Theory Appl.*, vol. 9, pp. 177–187, 1981.

[2] B. A. Francis and G. Zames, "On the optimal sensitivity theory for SISO feedback systems," *IEEE Trans. Automat. Contr.*, vol. AC-29, pp. 9–16, 1984.

[3] B. C. Chang and J. J. B. Pearson, "Optimal disturbance reduction in linear multivariable systems," *IEEE Trans. Automat. Contr.*, vol. AC-29, pp. 880–887, 1984.

[4] H. Kimura, "Robust stabilizability for a class of transfer functions," *IEEE Trans. Automat. Contr.*, vol. AC-29, no. 9, pp. 788–793, 1984.

[5] P. P. Khargonekar and A. Tannenbaum, "Non-Euclidean metrics and the robust stabilization of systems with parameter uncertainty," *IEEE Trans. Automat. Contr.*, vol. AC-30, no. 10, pp. 1005–1013, 1985.

[6] N. I. Akhiezer, *The Classical Moment Problem.* London, U.K.: Oliver and Boyd, 1965.

[7] J. L. Walsh, *Interpolation and Approximation by Rational Function in the Complex Domain.*, 4th ed., AMS Colloquium Publications, 1965.

[8] J. C. Doyle, B. A. Francis, and A. R. Tanenbaum, *Feedback Control Theory.* New York: Macmillan, 1992.

[9] D. W. Luse and H. K. Khalil, "Frequency domain results for systems with slow and fast dynamics," *IEEE Trans. Automat. Contr.*, vol. AC-30, no. 12, pp. 1171–1179, 1985.

[10] D. W. Luse and J. A. Ball, "Frequency-scale decomposition of H_∞ disk problems," *SIAM J. Contr. Optimiz.*, vol. 27, pp. 814–835,1989.

[11] G. H. Golub and C. F. V. Loan, *Matrix Computations,* 2nd ed. Baltimore, MD, Johns Hopkins Univ., 1990.

Multiscale Smoothing Error Models

Mark R. Luettgen and Alan S. Willsky

***Abstract*— A class of multiscale stochastic models based on scale-recursive dynamics on trees has recently been introduced. These models are interesting because they can be used to represent a broad class of physical phenomena and because they lead to efficient algorithms for estimation and likelihood calculation. In this paper, we provide a complete statistical characterization of the error associated with smoothed estimates of the multiscale stochastic processes described by these models. In particular, we show that the smoothing error is itself a multiscale stochastic process with parameters that can be explicitly calculated.**

I. Introduction

A class of multiscale models describing stochastic processes indexed by the nodes of a tree has recently been introduced in [1], [2]. These models can be used to capture a surprisingly rich class of physical phenomena. For instance, experimental results in [2] illustrate that they can be used to model the statistical self-similarity exhibited by stochastic processes with generalized power spectra of the form $1/f^\beta$, and in [3] we describe how they can be used to represent any 1-D Markov process or 2-D Markov random field. Moreover, this class of models leads to efficient algorithms for estimation and likelihood calculation and as a result provides a useful framework for a variety of signal and image processing problems [1], [2], [4]–[6].

Knowledge of the error statistics of smoothed estimates of such processes is essential for the development of a number of important new applications, including for instance so-called mapping problems [7], the multiscale counterpart to the model validation problem in [8], and certain oceanographic problems [9]. Several such applications have been developed in the context of 1-D Gauss–Markov models by exploiting relatively recent results that show that the smoothing error processes associated with Gauss–Markov models are themselves Gauss–Markov processes [7], [8], [10], [11].[1] In this paper, we derive a dynamic model for the smoothing error process associated with multiscale stochastic models. In particular, we show that the smoothing error is itself a multiscale stochastic process with parameters that can be explicitly computed. These results generalize previous results for Gauss–Markov processes, since these processes correspond to a degenerate form of the multiscale models, and provide the necessary framework for applications such as those mentioned above.

This paper is organized as follows. In Section II we briefly review the class of multiscale stochastic models of interest here and the scale-recursive estimation algorithm associated with them. In Section III we derive a multiscale model for the smoothing error process.

Manuscript received October 19, 1993; revised March 28, 1994. This work was supported by the Air Force Office of Scientific Research under Grant AFOSR-2-J-0002, by the Army Research Office under Grant DAAL03-92-G-0115, by the Office of Naval Research under Grant N00014-91-J-1004, and by the Advanced Research Projects Agency through Air Force Grant F49620-93-1-0604.

M. R. Luettgen is with Alphatech, Inc., Burlington, MA 01803 USA.

A. S. Willsky is with MIT Laboratory for Information and Decision Systems, Cambridge, MA 02139 USA.

IEEE Log Number 9406982.

[1] More generally, Levy *et al.* [12] have recently shown that the smoothing error processes associated with the class of Gaussian reciprocal processes, which contains the class of Gauss–Markov processes, are themselves Gaussian reciprocal. See also [13] for similar results corresponding to 2-D Gauss–Markov random fields.

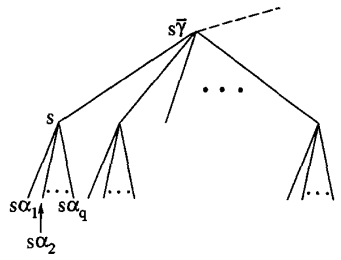

Fig. 1. Multiscale stochastic processes are indexed by a qth-order tree. The parent of a node s on the tree is denoted $s\overline{\gamma}$, and its q offspring are denoted $s\alpha_1, \cdots, s\alpha_q$.

II. Multiscale Stochastic Modeling and Optimal Estimation

The models presented in this section describe multiscale Gaussian stochastic processes indexed by nodes on a tree. A qth order tree is a pyramidal structure of nodes connected such that each node of the tree has q offspring (see Fig. 1). We denote nodes on the tree with an abstract index s, and define an upward (fine-to-coarse) shift operator $\overline{\gamma}$ such that $s\overline{\gamma}$ is the parent of node s. We also define a corresponding set of downward shift operators $\alpha_1, \ldots, \alpha_q$ such that $s\alpha_1, \cdots, s\alpha_q$ are the offspring of node s. In addition, we denote the set of nodes on the tree as $\mathcal{T}$ and the set of nodes that includes node s and all of its descendants as $\mathcal{T}_s$, i.e., $\mathcal{T}_s = \{\sigma | \sigma = s \text{ or } \sigma \text{ is a descendant of } s\}$. Also, the complement of $\mathcal{T}_s$ is denoted $\mathcal{T}_s^c$. The statistical characterization of the model state $x(s) \in \mathcal{R}^n$ is then given by

$$x(s) = A(s)x(s\overline{\gamma}) + B(s)w(s) \tag{1}$$

under the assumptions that $x(0) \sim \mathcal{N}(0, P(0)), w(s) \sim \mathcal{N}(0, I), A(s)$ and $B(s)$ are matrices of appropriate size, and $s = 0$ is the root node at the top of the tree. The driving noise $w(s) \in \mathcal{R}^m$ is white, i.e., $w(s)$ and $w(\sigma)$ are independent if $s \neq \sigma$, and independent of the initial condition $x(0)$. The class of models (1) has a statistical structure that can be exploited to develop efficient signal processing algorithms. In particular, note that any given node on the qth-order tree can be viewed as a boundary between $q + 1$ subsets of nodes (q corresponding to paths leading towards offspring and one corresponding to a path leading towards a parent). An important property of the model (1) is that, conditioned on the value of the state at any node, the values of the state corresponding to the $q + 1$ subsets of nodes are independent. This fact is the basis for the development in [1], [2] of an algorithm for computing smoothed estimates of $x(s)$ based on noisy measurements $y(s) \in \mathcal{R}^p$ of the form

$$y(s) = C(s)x(s) + v(s) \tag{2}$$

where $v(s) \sim \mathcal{N}(0, R(s))$, and is independent of both $w(s)$ and $x(0)$. The algorithm for computing the smoothed estimates of $x(s)$ is a generalization to qth-order trees of the well-known Rauch–Tung–Striebel algorithm for smoothing 1-D Gauss–Markov processes. We briefly review the multiscale smoothing algorithm next, and then derive a model for the error associated with the smoothed estimates.

We denote the set of states defined at nodes in $\mathcal{T}_s$ as X_s, i.e., $X_s = \{x(\sigma) | \sigma \in \mathcal{T}_s\}$, and similarly $Y_s = \{y(\sigma) | \sigma \in \mathcal{T}_s\}$. The set of measurements in the subtree strictly below s is denoted Y_s^+, i.e., $Y_s^+ = \{y(\sigma) | \sigma \text{ is a descendant of } s\}$. We also define $\hat{x}(s|Y)$ as the expected value of $x(s)$ given measurements in the set Y and the corresponding error covariance as $P(s|Y)$.

The upward sweep of the smoothing algorithm begins with the initialization of $\hat{x}(s|Y_s^+)$ and $P(s|Y_s^+)$ at the finest level. In particular, for every s at this finest scale we set $\hat{x}(s|Y_s^+)$ to zero and $P(s|Y_s^+)$ to the solution at the finest level of the tree of the Lyapunov equation:

$$P(s) = A(s)P(s\overline{\gamma})A^T(s) + B(s)B^T(s) \tag{3}$$

where $P(s)$ denotes the covariance of the process $x(s)$ at node s. Suppose then that we have $\hat{x}(s|Y_s^+)$ and $P(s|Y_s^+)$ at a given node s. This estimate is *updated* to incorporate the measurement $y(s)$ according to the following:

$$\hat{x}(s|Y_s) = \hat{x}(s|Y_s^+) + K(s)[y(s) - C(s)\hat{x}(s|Y_s^+)] \tag{4}$$

$$P(s|Y_s) = [I - K(s)C(s)]P(s|Y_s^+) \tag{5}$$

where $K(s) = P(s|Y_s^+)C^T(s)[C(s)P(s|Y_s^+)C^T(s) + R(s)]^{-1}$.

Suppose next that we have the updated estimates $\hat{x}(s\alpha_i|Y_{s\alpha_i})$ at all of the immediate descendants of node s. The next step involves the use of these estimates to predict $x(s)$ at the next coarser scale, i.e., to compute $\hat{x}(s|Y_{s\alpha_i})$. Using the following *upward* model for the multiscale process [1], [2]:

$$x(s\overline{\gamma}) = F(s)x(s) + \overline{w}(s) \tag{6}$$

with the measurement equation again given by (2), and where $F(s) = P(s\overline{\gamma})A^T(s)P(s)^{-1}$ and $E[\overline{w}(s)\overline{w}^T(s)] = P(s\overline{\gamma}) - P(s\overline{\gamma})A^T(s)P(s)^{-1}A(s)P(s\overline{\gamma}) \equiv Q(s)$, we compute the fine-to-coarse *predicted* estimates:

$$\hat{x}(s|Y_{s\alpha_i}) = F(s\alpha_i)\hat{x}(s\alpha_i|Y_{s\alpha_i}) \tag{7}$$

$$P(s|Y_{s\alpha_i}) = F(s\alpha_i)P(s\alpha_i|Y_{s\alpha_i})F^T(s\alpha_i) + Q(s\alpha_i). \tag{8}$$

The estimates $\hat{x}(s|Y_{s\alpha_i}), i = 1, \cdots, q$ are then *merged* to obtain

$$\hat{x}(s|Y_s^+) = P(s|Y_s^+)\sum_{i=1}^{q} P^{-1}(s|Y_{s\alpha_i})\hat{x}(s|Y_{s\alpha_i}) \tag{9}$$

$$P(s|Y_s^+) = \left[(1-q)P(s)^{-1} + \sum_{i=1}^{q} P^{-1}(s|Y_{s\alpha_i})\right]^{-1}. \tag{10}$$

We assume here that $P(s)$ and $P(s\overline{\gamma}|Y_s)$ are invertible for all s so that the upward model given by (6) and the merge operation given by (9), (10) are well-defined. As discussed at the end of the next section, this restriction can be removed.

The recursion given by the update, predict and merge equations proceeds up the tree until one obtains the smoothed estimate of the root node, $\hat{x}(0|Y_0)$. This estimate initializes a *downward sweep* in which $\hat{x}(s|Y_0)$ is computed according to

$$\hat{x}(s|Y_0) = \hat{x}(s|Y_s) + J(s)[\hat{x}(s\overline{\gamma}|Y_0) - \hat{x}(s\overline{\gamma}|Y_s)] \tag{11}$$

$$P(s|Y_0) = P(s|Y_s) + J(s)[P(s\overline{\gamma}|Y_0) - P(s\overline{\gamma}|Y_s)]J^T(s) \tag{12}$$

$$J(s) = P(s|Y_s)F^T(s)P^{-1}(s\overline{\gamma}|Y_s). \tag{13}$$

Note that (12) characterizes the smoothing error covariance at any given lattice site s, but does not provide information about the correlation structure of the error process. The goal in the next section is to provide a multiscale model for the smoothing error process, i.e., to show that the error satisfies a recursion of the form (1), and to calculate the associated model parameters. This then provides the complete statistical characterization of the smoothing error that we seek.

III. Multiscale Smoothing Error Models

Given two nodes s and $\sigma \in \mathcal{T}_s^c$ on the tree, we can represent $x(\sigma)$ in terms of $x(s\overline{\gamma})$ and an additive noise term $\varphi_{\sigma, s\overline{\gamma}}$:

$$x(\sigma) = \Phi_{\sigma, s\overline{\gamma}}x(s\overline{\gamma}) + \varphi_{\sigma, s\overline{\gamma}} \tag{14}$$

with $\varphi_{\sigma, s\overline{\gamma}}$ independent of the set of states $x(s\overline{\gamma}) \cup X_s$ and the corresponding set of measurements $y(s\overline{\gamma}) \cup Y_s$, by tracing a path

from σ to $s\overline{\gamma}$ and using the upward dynamics (6) and downward dynamics (1) to eliminate state variables along the way. The state transition matrix $\Phi_{\sigma,s\overline{\gamma}}$ in this construction is a function of the upward and downward prediction matrices A and F along the path, whereas $\varphi_{\sigma,s\overline{\gamma}}$ is a linear function of the upward and downward driving noises w and $\overline{w}$. Since $\varphi_{\sigma,s\overline{\gamma}}$ is independent of the set of states $x(s\overline{\gamma}) \cup X_s$ and the measurements $y(s\overline{\gamma}) \cup Y_s$, we have that $\hat{x}(\sigma|Y_s) = \Phi_{\sigma,s\overline{\gamma}}\hat{x}(s\overline{\gamma}|Y_s)$ which, using (14), implies that

$$\tilde{x}(\sigma|Y_s) = \Phi_{\sigma,s\overline{\gamma}}\tilde{x}(s\overline{\gamma}|Y_s) + \varphi_{\sigma,s\overline{\gamma}} \tag{15}$$

where we have defined the error in $\hat{x}(s|Y)$ as $\tilde{x}(s|Y) \equiv x(s) - \hat{x}(s|Y)$. As a result, we see that $\tilde{x}(s|Y_s)$ has the Markov property:

$$\begin{aligned}
&\boldsymbol{E}\{\tilde{x}(s|Y_s)|\tilde{x}(\sigma|Y_s),\ \sigma \in \mathcal{T}_s^c\} \\
&= \boldsymbol{E}\{\tilde{x}(s|Y_s)|\tilde{x}(s\overline{\gamma}|Y_s),\ \{\varphi_{\varsigma,s\overline{\gamma}}|\varsigma \in \mathcal{T}_s^c\}\} \\
&= \boldsymbol{E}\{\tilde{x}(s|Y_s)|\tilde{x}(s\overline{\gamma}|Y_s)\} + \boldsymbol{E}\{\tilde{x}(s|Y_s)|\{\varphi_{\varsigma,s\overline{\gamma}}|\varsigma \in \mathcal{T}_s^c\}\} \\
&= \boldsymbol{E}\{\tilde{x}(s|Y_s)|\tilde{x}(s\overline{\gamma}|Y_s)\}.
\end{aligned} \tag{16}$$

The first equality in (16) follows from (15), the second from the orthogonality of $\varphi_{\varsigma,s\overline{\gamma}}$ to $x(s\overline{\gamma})$ and Y_s, and the last from the orthogonality of $\varphi_{\varsigma,s\overline{\gamma}}$ to $x(s)$ and Y_s. Now, using the upward dynamics (6), the upward sweep prediction equation (7), and standard linear least squares formulas we can write

$$\tilde{x}(s|Y_s) = J(s)\tilde{x}(s\overline{\gamma}|Y_s) + \breve{w}(s) \tag{17}$$

where $J(s)$ is given by (13) and where, from (16), $\breve{w}(s)$ is *independent* of $\{\tilde{x}(\sigma|s)|\sigma \in \mathcal{T}_s^c\}$, and has covariance:

$$P(s|Y_s) - P(s|Y_s)F^T(s)P^{-1}(s\overline{\gamma}|Y_s)F(s)P(s|Y_s). \tag{18}$$

Next, note that the independence of $\breve{w}(s)$ and $\{\tilde{x}(\sigma|Y_s)|\sigma \in \mathcal{T}_s^c\}$ implies that $\breve{w}(s)$ is also independent of the *residual* information about $x(s)$ that is contained in the set of all available measurements Y_0, but *not* contained in Y_s. In particular, at each node in $\mathcal{T}_s^c$, a residual component $\nu_s(\sigma)$ that is orthogonal to the measurements in the set Y_s can be defined as

$$\begin{aligned}
\nu_s(\sigma) &= y(\sigma) - \boldsymbol{E}\{y(\sigma)|Y_s\} \\
&= C(\sigma)\tilde{x}(\sigma|Y_s) + v(\sigma).
\end{aligned} \tag{19}$$

Denoting $\nu_s \equiv \{\nu_s(\sigma)|\sigma \in \mathcal{T}_s^c\}$, it is clear that span Y_0 =span $\{Y_s, \nu_s\}$, that $\nu_s \perp Y_s$, and that $\nu_s \perp \breve{w}(s)$. Taking the expected value of both sides of (17) conditioned on ν_s, we obtain

$$\boldsymbol{E}\{\tilde{x}(s|Y_s)|\nu_s\} = J(s)\boldsymbol{E}\{\tilde{x}(s\overline{\gamma}|Y_s)|\nu_s\}. \tag{20}$$

Finally, noting that

$$\hat{x}(s|Y_0) = \hat{x}(s|Y_s) + \boldsymbol{E}\{\tilde{x}(s|Y_s)|\nu_s\} \tag{21}$$

and then subtracting (20) from (17) results in

$$\tilde{x}(s|Y_0) = J(s)\tilde{x}(s\overline{\gamma}|Y_0) + \breve{w}(s) \tag{22}$$

which is a multiscale model for the smoothing error of precisely the same form as (1).

This model is, of course, consistent with the error covariance computation in (12). In particular, using the Lyapunov equation for (22) we obtain

$$\begin{aligned}
P(s|Y_0) &= J(s)P(s\overline{\gamma}|Y_0)J^T(s) + P(s|Y_s) \\
&\quad - P(s|Y_s)F^T(s)P^{-1}(s\overline{\gamma}|Y_s)F(s)P(s|Y_s) \\
&= P(s|Y_s) + J(s)[P(s\overline{\gamma}|Y_0) - P(s\overline{\gamma}|Y_s)]J^T(s).
\end{aligned} \tag{23}$$

In addition, on first-order trees, the model (1) reduces to a standard Gauss–Markov model, and hence (22) generalizes to qth-order trees the corresponding 1-D time-series result. The derivation here is related to, but is in fact substantially simpler than, the derivation based on backwards prediction error models in [8].

Finally, we note that it is possible to derive a multiscale smoothing error model without assuming invertibility of $P(s)$ and $P(s\overline{\gamma}|Y_s)$. We refer the reader to Appendix D of [14] for a related derivation of a likelihood calculation algorithm for (1), (2) that allows for rank deficient $P(s)$ and $P(s\overline{\gamma}|Y_s)$. A slight variation of the technique used in that derivation can be used to show that a multiscale smoothing error model allowing for rank deficient $P(s)$ and $P(s\overline{\gamma}|Y_s)$ can be written precisely as in (22) but with the gain $J(s)$ given by

$$J(s) = P(s|Y_s)F^T(s)P^{\dagger}(s\overline{\gamma}|Y_s) \tag{24}$$

$$F(s) = P(s\overline{\gamma})A^T(s)P(s)^{\dagger} \tag{25}$$

and the covariance of $\breve{w}(s)$ given by

$$P(s|Y_s) - P(s|Y_s)F^T(s)P^{\dagger}(s\overline{\gamma}|Y_s)F(s)P(s|Y_s) \tag{26}$$

where the superscript $\dagger$ denotes the Moore–Penrose pseudo-inverse [15].

REFERENCES

[1] K. Chou, A. Willsky, A. Benveniste, and M. Basseville, "Recursive and iterative estimation algorithms for multiresolution stochastic processes," in *Proc. 28th IEEE Conf. Decision and Control*, Tampa, FL, Dec. 1989, pp. 1884–1189.

[2] K. Chou, A. Willsky, and A. Benveniste, "Multiscale recursive estimation, data fusion and regularization," Massachusetts Inst. Technol., Lab. Inform. Decision Syst., Tech. Rep. LIDS-P-2085, 1991; see also *IEEE Trans. Automat. Contr.*, Apr. 1994.

[3] M. Luettgen, W. Karl, A. Willsky, and R. Tenney, "Multiscale representations of Markov random fields," *IEEE Trans. Signal Processing*, vol. 41, pp. 3377–3396, 1993.

[4] M. Luettgen, W. Karl, and A. Willsky, "Efficient multiscale regularization with application to the computation of optical flow," *IEEE Trans. Image Processing*, vol. 3, pp. 41–64, 1994.

[5] M. Luettgen and A. Willsky, "Likelihood calculation for a class of multiscale stochastic models, with application to texture discrimination," Lab. Inform. Decision Syst., Massachusetts Inst. Technol., Tech. Rep. LIDS-P-2186, 1993; to appear in *IEEE Trans. Image Processing*.

[6] ——, "Likelihood calculation for a class of multiscale stochastic models," in *Proc. 32nd IEEE Conf. Decision and Control*, San Antonio, TX, Dec. 1993.

[7] M. Bello, A. Willsky, B. Levy, and D. Castañon, "Smoothing error dynamics and their use in the solution of smoothing and mapping problems," *IEEE Trans. Inform. Theory*, vol. IT-32, pp. 483–495, 1986.

[8] M. Bello, A. Willsky, and B. Levy, "Construction and applications of discrete-time smoothing error models," *Int. J. Contr.*, vol. 50, pp. 203–223, 1989.

[9] P. Fieguth, W. Karl, and A. Willsky, "Multiresolution statistical estimation of North Pacific ocean height from Topax/Poseiden satellite altimetry," in *Proc. 1994 SPIE Conf. Neural and Stochastic Methods in Image and Signal Processing*, San Diego, CA, July 1994.

[10] F. Badawi, A. Lindquist, and M. Pavon, "A stochastic realization approach to the smoothing problem," *IEEE Trans. Automat. Contr.*, vol. AC-24, pp. 878–887, 1979.

[11] R. Ackner and T. Kailath, "Discrete-time complementary models and smoothing," *Int. J. Contr.*, vol. 49, pp. 1665–1682, 1989.

[12] B. Levy, R. Frezza, and A. Krener, "Modeling and estimation of discrete time Gaussian reciprocal processes," *IEEE Trans. Automat. Contr.*, vol. 35, pp. 1013–1023, 1990.

[13] B. Levy, "Non-causal estimation for Markov random fields," in *Proc. Int. Symp. MTNS-89, Vol. 1: Realization and Modeling in System Theory*. Basel, Switzerland: Birkhauser-Verlag, 1990.

[14] M. Luettgen, "Image processing with multiscale stochastic models," Ph.D. dissertation, Dep. Elec. Eng. Comput. Sci., Massachusetts Inst. Technol., May 1993.

[15] S. Campbell and C. Meyer, *Generalized Inverses of Linear Transformations*. London, England: Pitman, 1979.