

Bayesian Cell Filter for Constrained Non-Gaussian Estimation

Sridhar Ungarala and Zhongzhou Chen
Department of Chemical and Biomedical Engineering
Cleveland State University
Cleveland, OH 44115, USA
s.ungarala@csuohio.edu, z.z.chen@csuohio.edu

Abstract—The Bayesian approach provides the most general formulation of the recursive state estimation problem. Except for linear-Gaussian systems, the solution is seldom amenable to implementation. This paper poses the estimation problem in discretized state space. A novel approach is used to model probabilistic dynamics as finite state Markov chains. The Bayesian cell filter can handle nonlinearities, non-Gaussian process and measurement noise and constraints. The filter splits the problem into offline modeling and online estimation tasks. The cell filter is compared with Monte Carlo based particle filter for accuracy and efficiency.

I. INTRODUCTION

Rectification and estimation reduces measurement errors and estimates the true values of states. In general, the errors may be non-additive with non-Gaussian distributions. The system may be nonlinear, accompanied by algebraic constraints. Nonlinearity renders the probability density function (pdf) of the states into time-varying non-Gaussian pdf. Process uncertainties are typically bounded since parameters and disturbances are normally bounded. Algebraic constraints also place bounds on the domains of the pdfs.

The Bayesian approach provides a rigorous solution for estimation by fusing data with *a priori* information in the context of models and constraints. However, generalized solutions for nonlinear/non-Gaussian constrained systems are impossible. Methods such as the extended Kalman filter (EKF) [1], Moving Horizon Estimator (MHE) [2], are special cases of the Bayesian approach, which rely on simplifying assumptions about models, constraints and pdfs.

EKF uses linearization while assuming Gaussian pdfs, and cannot include constraints. MHE reformulates the estimation problem as a quadratic programming problem in a fixed-size window. MHE inherently assumes Gaussian pdfs, which leads to a convenient least squares problem. Although it can include constraints, MHE is computationally intensive and it is not easy to determine an optimal window size.

Monte Carlo based methods such as the particle filter (PF) [3,4] approximate non-Gaussian pdfs with samples. PF follows the evolution of pdfs by simulating particle trajectories in state space. However, the issue of constraints is yet to be explored in PF [5]. Insufficient sampling and degeneracy are of concern for constrained PF since discarding samples violating the constraints can worsen these issues. Furthermore, passing a great number of samples to the nonlinear system at each sampling time entails considerable computational burden for online estimation.

This paper presents a novel approach for estimation in constrained dynamic systems with non-Gaussian pdfs. The fundamental difference between the Bayesian cell filter and existing methods is the separation of the modeling task from the estimation problem. Unlike the Monte Carlo filters, modeling is performed offline thereby significantly reducing the cost of implementing recursive estimation online.

In principle, the evolution of the state pdf is represented by a Foias operator, which is difficult to obtain analytically. This difficulty naturally leads to the popularity of assumptions of linearity and Gaussianity. The proposed approach relaxes these assumptions. It is suboptimal since a continuous pdf is approximated by a discretized probability vector. To this end, the state space is discretized into a finite number of intervals called cells. This is a more realistic quantification of the states, considering the uncertainties in models and limits on sensor resolution. Then, the original deterministic/stochastic system is converted into a finite state Markov chain. A discretized Foias operator is computed using Monte Carlo techniques by Generalized Cell Mapping [6]. The Markov chain can describe the evolution of the state pdf very accurately by following the evolution of the probability vector in cell space [7].

The proposed Bayesian cell filter can readily incorporate algebraic inequality and equality constraints. At the outset, a constrained cell space is created only for the region satisfying all the constraints. The Foias operators are computed only for the constrained cell space. Thus, the estimated states will automatically satisfy all the applicable constraints. Features of non-Gaussian system noise are readily incorporated into the transition matrix of the Markov Chain. Since Monte-Carlo sampling for computing the transition matrix can be done offline, a great number of samples can be used to obtain high accuracy as opposed to the on-line computational burden of PF. The pdf of non-Gaussian measurement noise appears explicitly in the likelihood function so that it can be handled directly. The Bayesian cell filter is general enough to handle any type of nonlinearity, non-Gaussian pdf and algebraic constraint.

Simulation examples are included to demonstrate the proposed approach for non-Gaussian measurement noise, constraints on process noise and constraints on states. Mean squared error of estimation and CPU demands are compared with particle filters.

II. ESTIMATION IN STATE SPACE

Consider the following discrete-time dynamic system and measurement equation,

$$x_k = f(x_{k-1}, w_{k-1}), \quad (1)$$

$$y_k = h(x_k, v_k), \quad (2)$$

where $f : (\mathbb{R}^n \times \mathbb{R}^m) \rightarrow \mathbb{R}^n$ and $h : (\mathbb{R}^n \times \mathbb{R}^p) \rightarrow \mathbb{R}^p$ are nonlinear functions. $w_k \in \mathbb{R}^m$ and $v_k \in \mathbb{R}^p$ are i.i.d. white noise with known pdfs $p_w(w_k)$ and $p_v(v_k)$ respectively. The initial condition x_0 is known via pdf $p(x_0)$. The states are further subject to algebraic constraints,

$$g_1(x_k) = 0, \quad (3)$$

$$g_2(x_k) \leq 0. \quad (4)$$

A. Bayesian Estimation

Bayesian estimation constructs the conditional pdf of the state according to Bayes rule,

$$p(x_k|y_k) \propto p(y_k|x_k)p(x_k|y_{k-1}), \quad (5)$$

where the *a priori* knowledge is represented by $p(x_k|y_{k-1})$, which is modified into the *a posteriori* pdf $p(x_k|y_k)$, in light of the data represented by the likelihood function $p(y_k|x_k)$.

Recursive estimation is performed in three stages at each time instant, (1) a *prediction stage*, where the prior pdf is generated, (2) an *update stage*, where the posterior pdf is computed using prior and likelihood and (3) an *inference stage* where the estimate, \hat{x}_k , is drawn as an inference from the *a posteriori* pdf. Bayesian estimation is intuitively appealing, yet poses formidable difficulties for systems with nonlinearities, non-Gaussian pdfs and constraints.

B. Prediction

The initial pdf, $p(x_0)$, is distorted, translated and spread due to the map f , and the nature of the system noise. The evolution of the state pdf is described by a linear integral operator, P_f , known as the Foias operator [8],

$$p(x_k) = P_f p(x_{k-1}), \quad (6)$$

where the operator is defined as

$$\int P_f p(x_{k-1}) = \int \int_{f^{-1}} \{p_w(w_{k-1}) dw_{k-1}\} p(x_{k-1}) dx_{k-1}. \quad (7)$$

The integration is further constrained by applicable equality and/or inequality constraints placed on the domains of the state pdfs. An extension of (6) may be used for propagating the posterior pdf conditioned on measurements upto time $k-1$ into a conditional prior pdf at time k ,

$$p(x_k|y_{k-1}) = P_f p(x_{k-1}|y_{k-1}). \quad (8)$$

Unfortunately, P_f is difficult to obtain analytically, since the inverse map of f on constraint spaces g_1 and g_2 may have complicated geometry even for relatively simple nonlinearities and constraints.

C. Update

If the measurement map, h , and its inverse map, h^{-1} , are continuously differentiable, the likelihood function is explicitly obtained as follows [1],

$$p(y_k|x_k) = p_v(h^{-1}(x_k, y_k)) \left\| \frac{\partial h^{-1}}{\partial y_k} \right\| = L_h(x_k, y_k). \quad (9)$$

Except for linear or scalar systems, it is generally not possible to compute the likelihood function, $L_h(x_k, y_k)$, analytically since the inverse map h^{-1} may not be well behaved.

The recursive Bayesian estimation solution is the updated *a posteriori* conditional state pdf,

$$p(x_k|y_k) \propto L_h(x_k, y_k) P_f p(x_{k-1}|y_{k-1}). \quad (10)$$

D. Inference

The optimal inference is typically drawn by a conditional expectation on the posterior,

$$\hat{x}_k = \mathbf{E}\{\phi(x_k)\} = \int \phi(x_k) p(x_k|y_k) dx_k, \quad (11)$$

where $\phi(\cdot)$ is any suitable real function. The conditional mean, mode and median are commonly used as optimal inferences according to various optimality criteria.

The state of the nonlinear, non-Gaussian and constrained estimator is represented by the entire conditional pdf, hence infinite in size. The computational burden of finding the entire pdf at each sampling time is formidable. The computation of its moments is also a nontrivial task. In view of these limitations, the suboptimal Bayesian cell filter comprising a finite set of quantities is defined [7].

III. ESTIMATION IN CELL SPACE

A. Cell Space

Many estimation problems of practical interest are posed with constraints on states, which lead to finite state space. Consider a finite region $R \subset \mathbb{R}^n$, where the system dynamics, subject to (3) and (4), are likely to be observed. Let R be partitioned into a collection of finite number of connected sets called cells, $\{z^i, i = 1, 2, \dots, N\}$. State space outside the region of interest is a single infinite sized cell called the sink cell z_0 . The continuous state space \mathbb{R}^n is approximated by the discrete cell space $Z = \{z^i\}_{i=0}^N$, and $Z \rightarrow \mathbb{R}^n$ as $N \rightarrow \infty$. State transitions from point to point described by (1) possess analogous cell transitions in cell space. Transitions from cells $\{z^i, i = 1, 2, \dots, N\}$ into the sink cell z_0 are considered terminal. The evolution of the system as a finite state Markov chain over the constrained cell space represents coarse-grained dynamics of the system.

Consider a region of interest $S \subset \mathbb{R}^p$, where measurements of $x \in R$ are likely to be obtained. Let S be discretized into a set of finite measurement cells, $\{d^i, i = 1, 2, \dots, M\}$. A measurement sink cell d_0 represents the infinite space outside S . The collection, $D = \{d_i\}_{i=0}^M$, coarsely represents the set of values likely to be obtained as measurements.

B. Prediction

Let the state pdf, $p(x_k)$, be approximated as a cell probability vector $\mathbf{p}(z_k)$,

$$\mathbf{p}(z_k) = \begin{bmatrix} m_k^0 \\ m_k^1 \\ \vdots \\ m_k^N \end{bmatrix}, \quad (12)$$

where m_k^i is the cell probability mass. Given the current $\mathbf{p}(z_{k-1})$, it is desired to predict the future $\mathbf{p}(z_k)$. The relationship between the initial m_{k-1}^j in cell z^j and the final m_k^i in cell z^i is obtained by a discrete analogue of (6),

$$m_k^i = \sum_{j=0}^N p_{ij} m_{k-1}^j, \quad (13)$$

where p_{ij} is the probability of transition from cell z^j to z^i ,

$$p_{ij} = \int_{z^i} p(x \in z^i | x \in z^j) dx. \quad (14)$$

The evolution of cell probability vectors over cell space is thus, described by a linear transformation,

$$\mathbf{p}(z_k) = \mathbf{P}\mathbf{p}(z_{k-1}), \quad (15)$$

where the transition probability matrix $\mathbf{P} = [p_{ij}]$ is a discretized representation of the linear integral operator P_f in continuous state space occurring in (6). Equation (15) represents a Markov model for the evolution of probability mass in the cell space.

C. Update

The likelihood of obtaining a measurement cell, d , when the state cell is z is given by the cell likelihood matrix, \mathbf{L} , which is a discrete analogue of the likelihood function L_h in (9). Given a measurement cell d_k^i at time k , the cell likelihood vector, $\mathbf{l}(d_k^i|z_k)$, is represented by the appropriate row in the likelihood matrix $\mathbf{L} = [l_{ij}]$. The likelihood mass, l_{ij} , with respect to the measurement cell d^i and the cell z^j is obtained by integrating the likelihood over the cell,

$$l_{ij} = \int_{d^i} p(y \in d^i | x \in z^j) dy. \quad (16)$$

Bayesian estimation in cell space is the problem of recursively constructing the posterior probability vector $\mathbf{p}(z_k|d_k^i)$ given the current likelihood vector $\mathbf{l}(d_k^i|z_k)$ and the previous posterior probability vector $\mathbf{p}(z_{k-1}|d_{k-1}^i)$,

$$\mathbf{p}(z_k|d_k^i) \propto \mathbf{l}(d_k^i|z_k) \otimes \mathbf{P}\mathbf{p}(z_{k-1}|d_{k-1}^i), \quad (17)$$

where \otimes is the Haddamard product.

D. Inference

Each cell is represented by the coordinates of the cell center to which the probability mass is assigned. The cell centers form the elements of the cell vector. An estimate of the state, \hat{x}_k , is obtained by computing the necessary expectations on cell posterior vector using simple dot product with the appropriate function of the cell vector.

IV. CELL-TO-CELL MAPPING

The difficulty in obtaining the Bayesian inference in cell space is not decision theoretic, it is centered on realizing the probability transition matrix, \mathbf{P} , and the likelihood matrix, \mathbf{L} . They can be approximately computed using Monte Carlo integration based on Generalized Cell Mapping (GCM) [6]. A constrained cell space is constructed. A number of initial conditions are uniformly sampled in each cell and the system is simulated to locate the image points (Fig. 1). The transition probability mass, p_{ij} , in the cell transition probability matrix, \mathbf{P} , is computed via this Monte Carlo sampling and simulation as follows,

$$p_{ij} \approx \frac{n_i}{n_j}, \quad (18)$$

where n_j are the number of sampled initial conditions in a cell, z^j , and n_i are the number of mapped images in the image cell, z^i . The approach is completely general for any type of discrete-time nonlinear system, non-Gaussian process noise and algebraic constraints. A similar mapping is performed to compute the likelihood mass, l_{ij} , in the likelihood matrix, \mathbf{L} , for any type of measurement model and non-Gaussian noise.

Small cell sizes are the key to obtaining high resolution information. An accurate mapping of the space inside each cell is of paramount importance for the determination of transition probabilities. If n initial conditions are sampled in each of N cells, GCM requires nN computations of the map f . For instance, 50 samples per cell in a 100×100 cell space will require half a million computations of f . The computational burden grows with the dimensionality of the system. In principle, the sampling and simulation is similar to following the evolution of trajectories in particle filters. Cell mapping is an exercise in large scale simulation to precompute all possible state transitions. The computational cost is one time for any given resolution. The efficiency of the cell filter stems from the fact that the computationally intensive modeling problem is solved off line. The online cost is limited to large matrix multiplications.

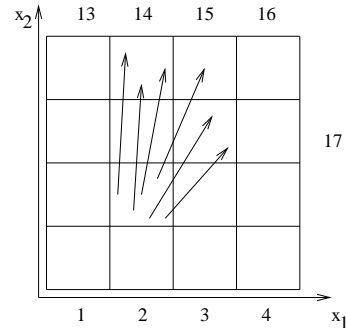


Fig. 1. Generalized cell mapping

V. SIMULATION EXAMPLES

A. Non-Gaussian Measurement Noise

Consider the following one dimensional linear system with zero mean Gaussian process noise, $w_k \sim N(0, 1)$ [9],

$$x_k = x_{k-1} + w_{k-1}, \quad (19)$$

$$y_k = x_k + v_k. \quad (20)$$

The measurements are corrupted with additive non-Gaussian errors. The probability density function of the measurement noise is a sum of two Gaussians,

$$v_k \sim 0.9N(0, 1) + 0.1N(0, 10^2). \quad (21)$$

200 data points are simulated with an initial condition 0.

The state space $x \in [-20, 20]$ is discretized into 200 uniformly defined cells. Generalized Cell Mapping is used with 400 uniformly sampled maps in each cell to compute the cell probability transition matrix, \mathbf{P} . The sparsity pattern of the matrix, \mathbf{P} , is shown in Fig. 2 for a coarser cell space.

Results of the Bayesian cell filter initialized with a Gaussian prior, $x_0 \sim N(0, 1)$, are shown in Fig. 3. The mean squared error of estimation is defined as,

$$\text{MSE} = \frac{1}{Kt} \sum_{k=1}^K (x_k - \hat{x}_k)^T R^{-1} (x_k - \hat{x}_k), \quad (22)$$

where K is the number of data points and t is the length of the state vector. The extended Kalman filter and Moving Horizon Estimator are shown to yield poor results for this system when compared to the particle filter [5]. The average MSE of estimation and the CPU time for 100 realizations on a Pentium 600 MHz machine are shown in Table I. Both cell filter and the particle filter yield about the same MSE, but with widely different computational costs. The cell filter is more efficient since the modeling task with dense sampling is performed offline.

B. Constraints on Process Noise

Consider the following linear dynamic system with a non-negative constraint on the process noise [2],

$$x_k = Ax_{k-1} + B|w_{k-1}|, \quad (23)$$

$$y_k = Cx_k + v_k, \quad (24)$$

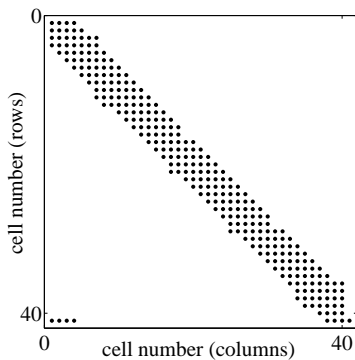


Fig. 2. Sparsity pattern of \mathbf{P} , for example V-A.

TABLE I
AVERAGE MSE AND CPU TIME FOR EXAMPLE V-A.

Method	MSE	CPU sec
Cell Filter, 200 cells	0.2	0.17
Particle Filter, 1000 samples	0.2	71.7

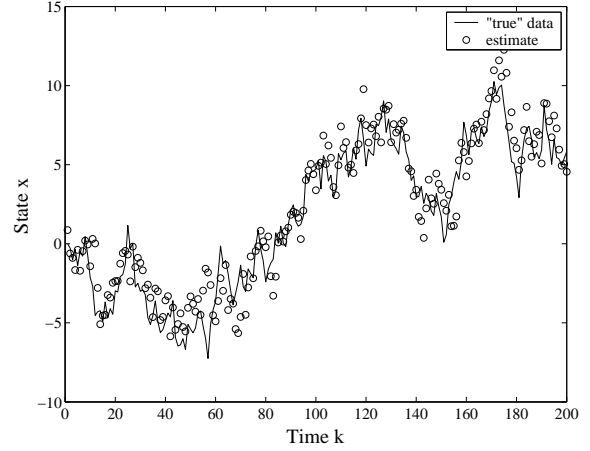


Fig. 3. Bayesian cell filter results for example V-A.

where $A = [0.9962, 0.1949; -0.1949, 0.3815]$, $B = [0.03393; 0.1949]$ and $C = [1 \quad -3]$, $w_k \sim N(0, 1)$ and $v_k \sim N(0, 0.1^2)$. The constrained non-Gaussian pdf of the process noise is shown in Fig. 4, whose mean is $2/\sqrt{2\pi}$ and variance is $(1 - 2\pi)$. Note that only one linear combination of the states is measured. The simulated data consists of 100 data points with the initial condition $[2.5; 0]$.

The region of the state space bound by $x(1) \in [0, 3]$ and $x(2) \in [-1, 1]$ is discretized into 60×40 uniformly defined cell space. The Generalized Cell Mapping (GCM) with 400 uniformly sampled maps per cell is used to compute the transition probability matrix, \mathbf{P} . The maps are simulated with process noise sampled from the constrained pdf. The sparsity pattern of a coarser resolution \mathbf{P} is shown in Fig. 5.

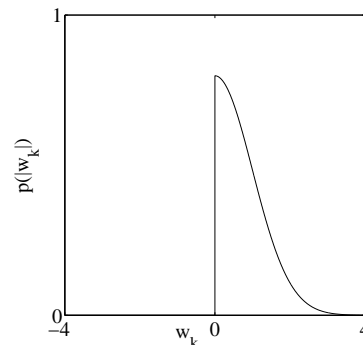


Fig. 4. Non-Gaussian process noise pdf.

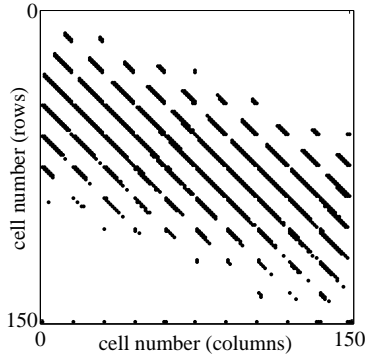


Fig. 5. Sparsity pattern of \mathbf{P} , for example V-B.

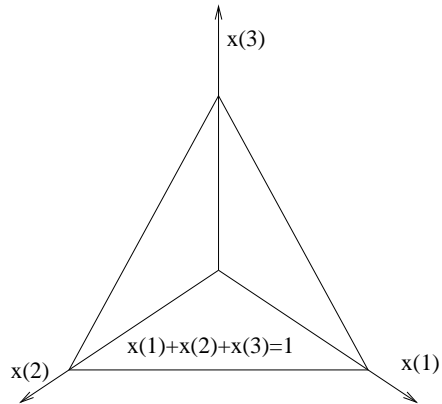


Fig. 7. Constrained region in state space

Estimation is initialized with Gaussian priors $x_0(1) \sim N(2.5, 0.3^2)$ and $x_0(2) \sim N(0, 0.3^2)$. The results are shown in Fig. 6. The cell filter yields better results than the extended Kalman filter and the Moving Horizon Estimators, which show significant biases [2, 5]. The average mean squared error (MSE) and CPU time for 100 realizations are shown in Table II. Even with 1000 samples, the particle filter fails to match the cell filter. Evidently larger number of particles are necessary for online simulation, which deteriorates the computational efficiency of the particle filter.

TABLE II
AVERAGE MSE AND CPU TIME FOR EXAMPLE V-B.

Method	MSE	CPU sec
Cell Filter, 60×40 cells	0.45	2.4
Particle Filter, 1000 samples	0.79	77

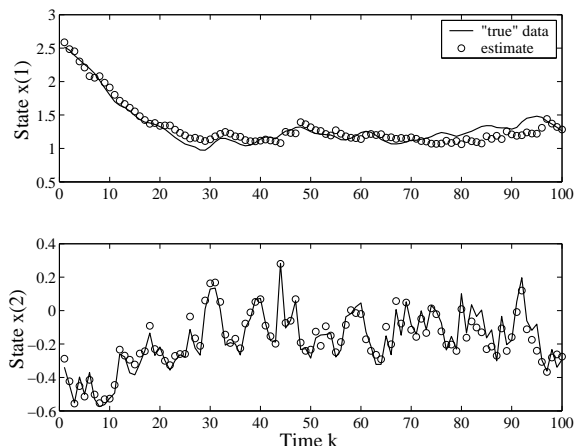


Fig. 6. Bayesian cell filter results for example V-B.

C. Constraints on States

Consider the following three state linear dynamic system with zero mean Gaussian process noise,

$$x_k = Ax_{k-1} + Bw_{k-1}, \quad (25)$$

$$y_k = Cx_k + v_k, \quad (26)$$

where $A = [.9, .2, .01; .05, .7, .1; .05, .1, .89]$, $B = [-0.5; -0.5; 1]$ and $C = I$, $w_k \sim N(0, 0.05^2)$ and all three measurements are corrupted by iid $v_k \sim N(0, 0.1^2)$. The states are further constrained according to,

$$0 \leq x_k \leq 1, \quad (27)$$

$$x_k(1) + x_k(2) + x_k(3) = 1. \quad (28)$$

The simulated data consists of 64 data points with initial condition $[0.25; 0.5; 0.25]$.

The region of state space bound by the unit cube is discretized into $40 \times 40 \times 40$ uniformly defined cell space. Although the system is three dimensional, the constraints limit the states to occupy a triangular plane shown in Fig. 7. The constrained cell space consists of only the cells through which the constraint plane passes, by virtue of which, the size of the cell space is significantly reduced. Using coordinate transformations, the cell space can be reduced to a two dimensional triangular plane. The Generalized Cell Mapping (GCM) with 400 uniformly sampled maps per cell is used to compute the transition probability matrix, \mathbf{P} . Fig. 8 shows the sparsity patterns of a coarse \mathbf{P} .

The estimation is initialized with three Gaussian priors with the respective initial conditions as the mean and a variance of 0.2^2 each. Fig. 9 shows the estimation results for a typical realization. The adherence to constraints is also shown against the failure of the Kalman filter to include the constraints in the estimation procedure. Average MSE and CPU time in Table III illustrate the computational efficiency of the Bayesian cell filter over the particle filter.

VI. CONCLUSIONS

Bayesian estimation of dynamic systems in cell space offers many advantages over traditional suboptimal methods

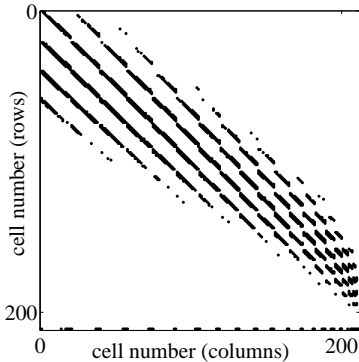


Fig. 8. Sparsity pattern of \mathbf{P} , for example V-C.

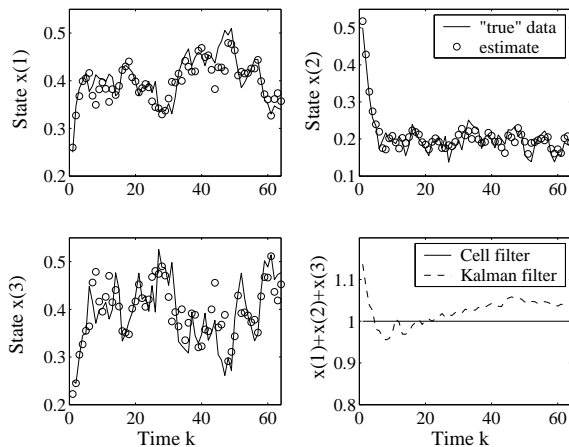


Fig. 9. Bayesian cell filter results for example V-C

such as the extended Kalman filter and the Moving Horizon Estimator. Multivariate relationships typically restrict the applicable mappings in state space. Physical bounds on variables as well as runaway limits on rates of change of variables also place constraints on the transition probabilities in state space. The Bayesian cell filter solves the estimation problem in discretized state space, which includes constraints on variables. No assumptions about the model structure or probability density functions are necessary. As long as the correct prior probability vector

TABLE III
AVERAGE MSE AND CPU TIME FOR EXAMPLE V-C.

Method	MSE	CPU sec
Cell Filter, $40 \times 40 \times 40$ cells	0.11	0.22
Particle Filter, 500 samples	0.11	10.8

and likelihood vector are given to the Bayesian inference problem, it does not matter whether the problem is linear, nonlinear, constrained or non-Gaussian. The system dynamics are modeled as a finite state Markov chain. The Markov models in cell space can be constructed offline, leading to considerable computational savings in online implementation of Monte Carlo based filters. The cell filter is a viable tool for recursive rectification primarily because of the separation of probabilistic modeling task from the rectification task.

VII. ACKNOWLEDGEMENTS

Financial support from EFFRD (SU) and DDREA (ZC) at Cleveland State University is gratefully acknowledged.

REFERENCES

- [1] A. H. Jazwinski, *Stochastic Processes and Filtering Theory*. New York: Academic Press, 1970.
- [2] C. V. Rao and J. B. Rawlings, "Constrained process monitoring: Moving-horizon approach," *AIChE Journal*, vol. 48, no. 1, pp. 97–109, 2002.
- [3] N. Gordon, D. J. Salmond, and A. F. M. Smith, "Novel approach to online/on-Gaussian Bayesian state estimation," *IEE Proc. F*, vol. 140, no. 2, pp. 107–113, 1993.
- [4] W. S. Chen, B. R. Bakshi, P. K. Goel, and S. Ungarala, "Bayesian estimation of unconstrained nonlinear dynamic systems via sequential Monte Carlo sampling," *Ind. Eng. Chem. Res.*, p. In Print, 2004.
- [5] —, "Bayesian estimation of nonlinear systems - dealing with constraints and non-Gaussian errors," in *AIChE 2002 Annual Meeting*, Indianapolis, IN, 2002.
- [6] C. S. Hsu, *Cell-to-Cell Mapping: A Method of Global Analysis for Nonlinear Systems*, ser. Applied Mathematical Sciences. New York: Springer-Verlag, 1987, vol. 64.
- [7] S. Ungarala and Z. Chen, "Bayesian data rectification of nonlinear systems with Markov chains in cell space," in *Proc. American Control Conference*, Denver, CO, 2003, pp. 4857–4862.
- [8] A. Lasota and M. C. Mackey, *Chaos, Fractals, and Noise: Stochastic Aspects of Dynamics*. Springer-Verlag, 1994.
- [9] L. P. M. Johnston and M. A. Kramer, "Estimating state probability distributions from noisy and corrupted data," *AIChE J.*, vol. 44, p. 591, 1998.