

Cleveland State University
EngagedScholarship@CSU



Chemical & Biomedical Engineering Faculty
Publications

Chemical & Biomedical Engineering Department

9-2007

Bayesian Estimation Via Sequential Monte Carlo Sampling-Constrained Dynamic Systems

Lixin Lang
Ohio State University

Wen-shiang Chen
Ohio State University

Bhavik R. Bakshi
Cleveland State University

Prem K. Goel
Ohio State University

Follow this and additional works at: https://engagedscholarship.csuohio.edu/encbe_facpub

 Part of the [Process Control and Systems Commons](#)

[How does a university this good benefit you? Let us know!](#)

Publisher's Statement

NOTICE: this is the author's version of a work that was accepted for publication in *Automatica*. Changes resulting from the publishing process, such as peer review, editing, corrections, structural formatting, and other quality control mechanisms may not be reflected in this document. Changes may have been made to this work since it was submitted for publication. A definitive version was subsequently published in *Automatica*, 43, 9, (September 2007) DOI 10.1016/j.automatica.2007.02.012

Original Citation

Lang, L., Chen, W., Bakshi, B. R., Goel, P. K., , & Ungarala, S. (2007). Bayesian estimation via sequential Monte Carlo sampling—Constrained dynamic systems. *Automatica*, 43(9), 1615-1622. doi:10.1016/j.automatica.2007.02.012

Repository Citation

Lang, Lixin; Chen, Wen-shiang; Bakshi, Bhavik R.; Goel, Prem K.; and Ungarala, Sridhar, "Bayesian Estimation Via Sequential Monte Carlo Sampling-Constrained Dynamic Systems" (2007). *Chemical & Biomedical Engineering Faculty Publications*. 65.
https://engagedscholarship.csuohio.edu/encbe_facpub/65

This Article is brought to you for free and open access by the Chemical & Biomedical Engineering Department at EngagedScholarship@CSU. It has been accepted for inclusion in Chemical & Biomedical Engineering Faculty Publications by an authorized administrator of EngagedScholarship@CSU. For more information, please contact library.es@csuohio.edu.

Bayesian estimation via sequential Monte Carlo sampling—Constrained dynamic systems[☆]

Lixin Lang^b, Wen-shiang Chen^a, Bhavik R. Bakshi^{a,*}, Prem K. Goel^b, Sridhar Ungarala^c

^aDepartment of Chemical and Biomolecular Engineering, The Ohio State University, Columbus, OH 43210, USA

^bDepartment of Statistics, The Ohio State University, Columbus, OH 43210, USA

^cDepartment of Chemical Engineering, Cleveland State University, Cleveland, OH 44115, USA

Abstract

Nonlinear and non-Gaussian processes with constraints are commonly encountered in dynamic estimation problems. Methods for solving such problems either ignore the constraints or rely on crude approximations of the model or probability distributions. Such approximations may reduce the accuracy of the estimates since they often fail to capture the variety of probability distributions encountered in constrained linear and nonlinear dynamic systems. This article describes a practical approach that overcomes these shortcomings via a novel extension of sequential Monte Carlo (SMC) sampling or particle filtering. Inequality constraints are imposed by accept/reject steps in the algorithm. The proposed approach provides samples representing the posterior distribution at each time point, and is shown to satisfy the same theoretical properties as unconstrained SMC. Illustrative examples show that results of the proposed approach are at least as accurate as moving horizon estimation, but computationally more efficient and in addition, the approach indicates the uncertainty associated with these estimates.

Keywords: Constrained estimation; Extended Kalman filter; Particle filter; Moving horizon estimation; Data rectification

1. Introduction

Estimation of dynamic processes is generally expressed as follows. Given measurements $y_{1:k} = \{y_1, y_2, \dots, y_k\}$, process models,

$$x_k = f_{k-1}(x_{k-1}, \omega_{k-1}), \quad (1)$$

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

and initial guess $p(x_1)$, estimate the current state, x_k . Here, $x_k \in \mathfrak{R}^{n_x}$ is the state vector and $f_k : \mathfrak{R}^{n_x} \times \mathfrak{R}^{n_\omega} \rightarrow \mathfrak{R}^{n_x}$ is the system equation. Measurements, $y_k \in \mathfrak{R}^{n_y}$ are related to the state vector through the measurement equation $h_k : \mathfrak{R}^{n_x} \times \mathfrak{R}^{n_v} \rightarrow \mathfrak{R}^{n_y}$. System and measurement errors are represented

by ω_k and v_k , respectively, and $p(x)$ denotes the probability density function (pdf) of a random variable x .

This nonlinear dynamic estimation problem has received significant attention over the decades and many techniques have been developed including extended Kalman filtering (EKF) (Jazwinski, 1970), unscented Kalman filtering (UKF) (Julier & Uhlmann, 2004), and moving-horizon based least-squares estimation (MHE) (Robertson, Lee, & Rawlings, 1996). EKF linearizes the nonlinear model at each time point and implicitly assumes the noise and variable distributions to be Gaussian. It is computationally efficient, but often fails to provide accurate estimates and may even diverge. UKF avoids the use of linearization and can perform better than EKF, but the deterministic choice of sigma points severely limits flexibility in the shape of the distributions. Also, this approach does not accommodate constraints. MHE formulates a nonlinear programming problem under possible constraints, which it solves in overlapping moving windows, usually of a fixed size. The objective function is usually formulated to find the least-squares solution and the approach has been shown to outperform EKF

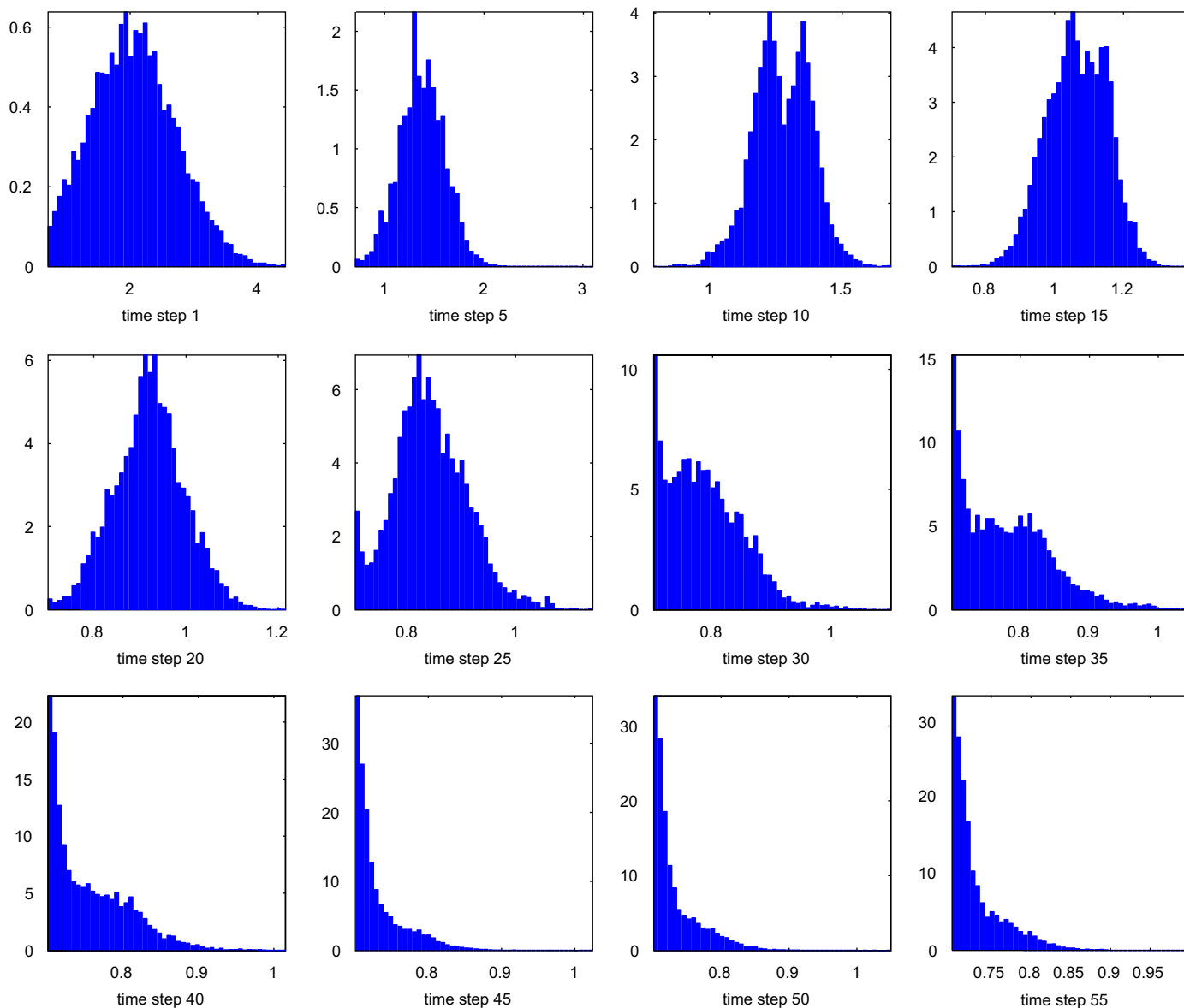


Fig. 1. Evolution of the prior of a McKeithan reaction network. In each subfigure the x -axis represents the variable's range while the y -axis is the density value.

(Haseltine & Rawlings, 2005). However, two significant shortcomings of MHE are that it is not recursive in nature, and it has to rely on multivariate Gaussian or other fixed shape distributions to represent the prior knowledge or arrival cost at the beginning of each moving window. These characteristics may increase the computational cost and reduce estimation accuracy. Fig. 1 shows typical prior distributions encountered in a reaction network example, which is discussed in Section 4.2. These distributions are truncated, unimodal or multimodal, and change shape over time. Any method that requires fitting of such arbitrary shapes is not likely to be practical, and has to rely on approximations such as Gaussian and give up some of its accuracy.

Recently, Bayesian estimation of dynamic processes through sequential Monte Carlo (SMC), also known as particle filtering, has been the subject of extensive research (Arulampalam, Maskell, Gordon, & Clapp, 2002; Doucet, Godsill, & Andrieu,

2000; Gordon, Salmond, & Smith, 1993). Our previous work (Chen, Bakshi, Goel, & Ungarala, 2004) introduced SMC for estimation in dynamic chemical process systems and compared it, for the first time, with MHE for unconstrained estimation. It showed that SMC can not only provide more accurate results than MHE but can also be more computationally efficient. This paper extends our work to deal with constraints, and compares its performance with EKF and MHE via popular case studies from the literature. Using sampling based methods for solving this type of problem does not seem to have received any attention to date. For imposing constraints within SMC, a novel acceptance/rejection algorithm is developed. This approach is equivalent to truncating or modifying the appropriate constrained distribution (prior or likelihood) to satisfy the constraints, which ensures that the posterior also satisfies the constraints. This approach is shown to possess the same theoretical properties as the traditional unconstrained

SMC approach. Furthermore, SMC readily provides insight into the quality of the estimates via regions of highest probability density (uncertainty) at each time point. Case studies demonstrate better accuracy and smaller computational cost of SMC over MHE.

In the following sections, introduction of the generic SMC is first provided. Then the proposed approach for extending sequential Monte Carlo to constrained nonlinear dynamic systems is described. Theoretical analysis of convergence properties of SMC methods is also introduced. Finally, performance of the proposed approach is compared with existing methods via simulation.

2. Sequential Monte Carlo

Using Monte Carlo sampling to solve estimation problems requires an approach for generating samples from the posterior. This may be accomplished via sequential importance sampling, since direct drawing from the posterior distribution is not feasible. For example, one can draw samples $\{x(i), i = 1, 2, \dots, N\}$ from a convenient importance function $\pi(x)$, and the estimate of $E[\varphi(x)]$ regarding a distribution, say, $p(x)$, is given by

$$E[\varphi(x)] = \int \varphi(x) \frac{p(x)}{\pi(x)} \pi(x) dx \\ \approx \frac{1}{N} \sum_{i=1}^N \varphi(x(i)) q^*(i) \quad \text{where } q^*(i) = \frac{p(x(i))}{\pi(x(i))}.$$

In fact, both $p(x)$ and $\pi(x)$ only have to be known up to a constant, in which case $q^*(i)$ are normalized and we have,

$$E[\varphi(x)] \approx \sum_{i=1}^N \varphi(x(i)) q(i) \quad \text{where } q(i) = \frac{q^*(i)}{\sum_{j=1}^N q^*(j)}. \quad (3)$$

Here, information about the relevant distribution is contained in the pairs of samples and weights, $\{x(i), q(i)\}$, known as particles. A basic requirement on the importance function is that its support should include the support of the true distribution (Geweke, 1989).

For dynamic processes, the same operation is used, though a recursive mechanism is needed to update the particles over time. Generally, a prediction step is applied recursively by passing each sample through the state equation (1), to obtain samples corresponding to the prior at time k , $p(x_k | y_{1:k-1})$. This prediction step utilizes information about process dynamics and model accuracy without making any assumptions about the nature of the dynamics and shape or any other characteristic of the distributions. Once the measurement, y_k is available, it can be used to recursively update the previous weights by the following equation (Arulampalam et al., 2002):

$$q_k^*(i) \propto q_{k-1}^*(i) \frac{p(y_k | x_k(i)) p(x_k(i) | x_{k-1}(i))}{\pi(x_k(i) | x_{k-1}(i), y_k)}. \quad (4)$$

This updating step utilizes the measurement model and information about the measurement error. Again, no assumptions about the type of model or distributions are required. The result of these prediction and updating steps is the particles and

weights at time k , $\{x_k(i), q_k(i)\}$. Any posterior moment may then be calculated via Eq. (3). The resulting algorithm is fully recursive and computationally efficient since it avoids integration for obtaining the moments at each time step, nonlinear optimization in a moving window, or restrictive assumptions about the nature of the error or prior distributions and models.

2.1. Convergence properties

Convergence properties of SMC have been studied by Del Moral and Miclo (2000), Crisan (2001), Künsch (2005), and many references cited in these papers. A survey of some convergence results for SMC is provided by Crisan and Doucet (2002). We summarize the state of the art based on the convention of Künsch (2005), followed by a discussion of their validity for the proposed constrained SMC algorithm.

Consider the general state space model described by Eqs. (1) and (2) in the context of a hidden Markov model. The state sequence $\{x_k, k \geq 1\}$ consists of a Markov chain with transition densities a_k such that

$$\Pr(x_k \in dx | x_{1:k-1}) = \Pr(x_k \in dx | x_{k-1}) = a_k(x_{k-1}, dx).$$

Similarly, given x_1, x_2, \dots , measurements y_1, y_2, \dots are conditionally independent and each follows the distribution

$$\Pr(y_k \in dy | x_k) = b_k(x_k, dy).$$

For example, in the widely used state space models with independent additive noises $\omega_k \sim p_\omega$ and $v_k \sim p_v$, the expressions of a_k and b_k are given by

$$a_k(x_{k-1}, x_k) = p_\omega(x_k - f_k(x_{k-1})),$$

$$b_k(x_k, y_k) = p_v(y_k - h_k(x_k)).$$

Given the measurements $y_{1:k}$, the posterior density of x_k is estimated by the empirical distribution based on the N particles and weights at time k , $\{x_k(i), q_k(i)\}$, in Eq. (3),

$$\hat{p}^N(x_k | y_{1:k}) = \sum_{i=1}^N \delta(x_k - x_k(i)) q_k(i).$$

Then Künsch (2005) provides the following theorem.

Theorem 1. *If $x \rightarrow a_k(x, \cdot)$ is continuous, and if for all k , all x and all y ,*

$$0 < b_k(x, y) \leq C(k, y) < \infty,$$

then for all k and all $y_{1:k}$,

$$\|\hat{p}^N(x_k | y_{1:k}) - p(x_k | y_{1:k})\|_1 \rightarrow 0$$

in probability as $N \rightarrow \infty$.

This result states that at each time point, the empirical distribution of the particles converges to the underlying true posterior density when the number of particles is increased.

Table 1
Algorithm for estimation by constrained SMC

-
- FOR times $k = 1, 2, 3, \dots$
 - FOR samples $i = 1, 2, 3, \dots, N$
 - **UNTIL** $(x_k(i), \omega_k(i))$ satisfy constraints,
 - **Draw sample**, $\omega_{k-1}(i)$ from system noise
 - **Calculate** $x_k(i)$ using Eq. (1)
 - **END UNTIL**
 - Assign a weight, $q_k^*(i)$, to $x_k(i)$
 - END FOR
 - Normalize $q_k^*(i)$ to find $q_k(i)$
 - Implement resampling if necessary
 - END FOR
-

The conditions imposed above are quite weak. Furthermore, it is straightforward to verify that these conditions are satisfied for the widely used additive Gaussian noise model.

The convergence of SMC shows that the empirical distribution based on the sampled particles and weights converges to the true distribution. However, the availability of a particle based approximation to the posterior distribution, though critical, does not answer the problem of the approximation error in the estimate of the parameter of interest. The following central limit theorem (Künsch, 2005) shows that under very weak conditions, the SMC approximation of the estimate based on the empirical distribution $\hat{p}^N(x_k|y_{1:k})$ converges to the true estimate at the rate of $1/\sqrt{N}$.

Theorem 2. *Under the conditions in Theorem 1, for each finite k , and all $y_{1:k}$ and functions $\varphi(\cdot)$ that are square integrable with respect to the true posterior distribution,*

$$\sqrt{N} \left(\frac{1}{N} \sum_{i=1}^N \varphi(x_k(i)) q(i) - E(\varphi(x_k)) \right)$$

is asymptotically normal.

This result is a highly simplified version of the one in Künsch (2005). As a practical matter, as time increases, it is important that the asymptotic variance of the Monte Carlo estimate stays bounded.

3. Constrained SMC

3.1. Algorithm

The proposed approach extends existing SMC algorithms to ensure satisfaction of inequality constraints. Equality constraints may be imposed by including them in the state or measurement equations (Ungarala & Bakshi, 2001). This approach, represented by the pseudo-code in Table 1, extends our previous work on unconstrained estimation (Chen et al., 2004). This implementation of SMC uses a convenient and popular choice of importance function as the following equation (Gordon et al., 1993):

$$\pi(x_k|x_{k-1}, y_k) = p(x_k|x_{k-1}).$$

This choice simplifies the recursive weight calculation by Eq. (4) to the following equation:

$$q_k^*(i) = q_{k-1}^*(i) p(y_k|x_k(i)). \quad (5)$$

Updating the prior with the current information then only requires the likelihood value. Other importance functions may be readily used, if necessary.

Enforcement of constraints in SMC is implemented by the extra steps shown in bold in Table 1. These accept/reject steps evaluate the samples $\omega_{k-1}(i)$ and the corresponding $x_k(i)$ generated by the prediction step, via Eq. (1). Only those samples of the generated $\{\omega_{k-1}(i), x_k(i)\}$ that satisfy constraints are accepted. Note that, the noise distributions for $k \geq 2$, or the prior distribution at $k = 1$, itself may be subject to the constraints, and therefore drawing samples from these distributions may require another accept/reject step whenever it is inconvenient to sample directly from the underlying distributions. For example, in Fig. 1, the first subfigure shows samples from a truncated Gaussian distribution. The steps shown in bold face in Table 1 may require a larger number of samples than unconstrained estimation, but as shown by the illustrative examples in Section 4, the computational complexity still remains reasonable and better than MHE. Usually, the number of rejected samples is not large, since most prior samples already satisfy the constraints.

3.2. Convergence properties

Remark 1. The constrained SMC algorithm in Table 1 satisfies the theorems given above.

The bold lines in Table 1 ensure that the accepted particles are truly generated from the correct transition densities under the constraints. In other words, the accept/reject operation leads to particles that correctly represent the posterior distribution, which ensures validity of the three theorems for constrained SMC.

4. Case studies

Simulations are run on a 2.0 GHz CPU with 512 MB RAM personal computer using Matlab and GNU/Octave. MHE is implemented by a publicly available package, which uses a specially structured solver for computational efficiency (Tenny, 2002). Allowing for uncertainty, a total of 100 realizations were run for each model with each procedure applied to the data sets generated by that model. The performance is evaluated by the overall mean-squared error (MSE) and, specifically MSE_k^R , the MSE averaged over realizations for each time $k = 1, 2, \dots, N_m$, where N_m is the number of measurements available in each realization. Examining MSE_k^R over time is likely to indicate the long-term behavior of the tested method and provide insight into the distribution of errors over time. They are defined as

$$\text{MSE}_k^R = \frac{1}{100} \sum_{r=1}^{100} (x_{k,r} - \hat{x}_{k,r})^T (x_{k,r} - \hat{x}_{k,r}), \quad (6)$$

$$\text{MSE} = \frac{1}{N_m} \sum_{k=1}^{N_m} \text{MSE}_k^R. \quad (7)$$

In the above equations, $x_{k,r}$ is the true state at time step k of the r th realization, and $\hat{x}_{k,r}$ is the point estimate of $x_{k,r}$. In this paper, the posterior mean is chosen as the point estimate for SMC. More details are available in Lang, Chen, Bakshi, Goel, and Ungarala (2006).

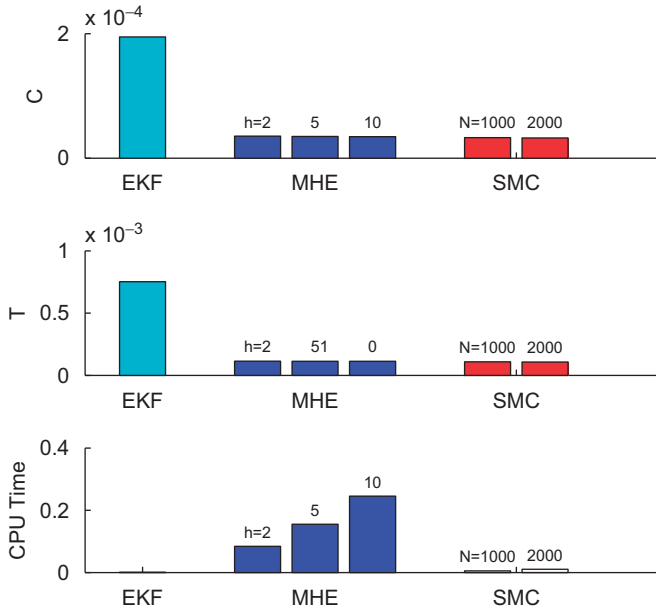


Fig. 2. MSE and CPU time comparison for the constrained CSTR Model.

4.1. Constrained adiabatic CSTR

In this section, a popular process engineering problem, an adiabatic CSTR, is studied (Jang, Joseph, & Mukai, 1986; Henson & Seborg, 1997). Governing equations are as follows:

$$\frac{dC}{dt} = \frac{q}{V}(C_0 - C) - kC e^{-E_A/T},$$

$$\frac{dT}{dt} = \frac{q}{V}(T_0 - T) - \frac{\Delta H}{\rho C_p} kC e^{-E_A/T} - \frac{UA}{\rho C_p V}(T - T_c),$$

where C is the concentration, T is the temperature, q is the flow rate, V is the volume of the reactor, C_0 and T_0 are inflow concentration and temperature, $kC e^{-E_A/T}$ is the reaction rate, ΔH is the heat of reaction, ρ is the density, C_p is the specific heat, T_c is the temperature of the cooling fluid, and U and A are the effective heat transfer coefficient and area. Discretization of continuous differential equations is implemented by finite difference with $\Delta t = 0.005$. The operating conditions are listed in Lang et al. (2006). The system noise at the scale of the normalized state variables is $p(\omega) \sim N(0, \sigma_\omega^2 \cdot I_2)$, where $\sigma_\omega^2 = 2.5 \times 10^{-7}$ and I_m is an $m \times m$ identity matrix, and the measurement noise is $p(v) \sim N(0, \sigma_v^2 \cdot I_2)$, where $\sigma_v^2 = 0.0025$. The initial guess is $p(x_1) \sim N(\mu_{x_1}, \sigma_v^2)$, where $\mu_{x_1} = [0.5; 350]$ is the initial state value. There are 400 measurements in each realization. A non-negative constraint is enforced on the concentration C_k at each time step, k .

As shown in Fig. 2, SMC consistently requires less computation time and results in slightly better accuracy than MHE, with both methods being much more accurate than EKF. The MSE_k^R values shown in Fig. 3 indicate that EKF performs quite well

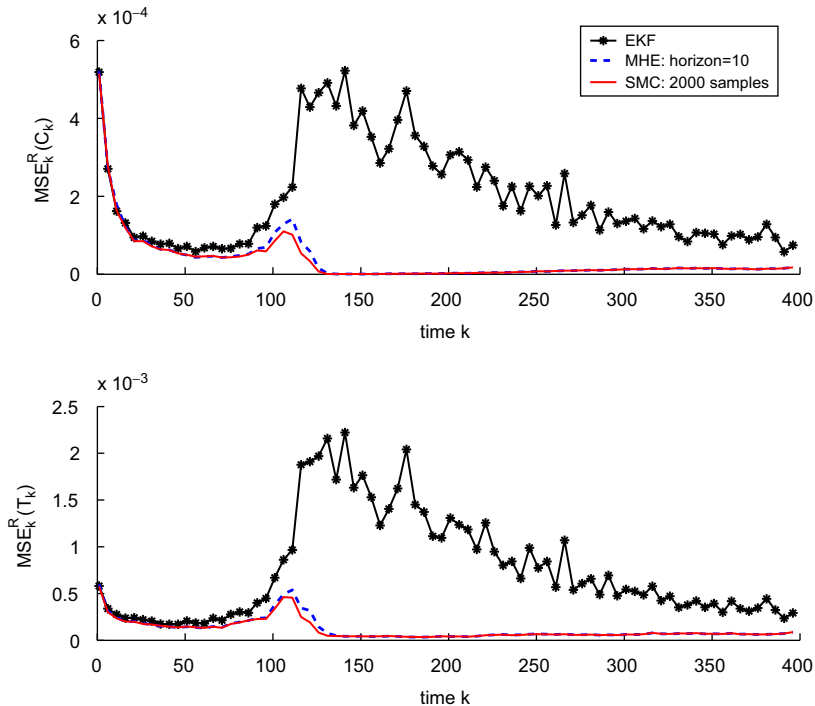


Fig. 3. MSE_k^R comparison for the constrained CSTR model.

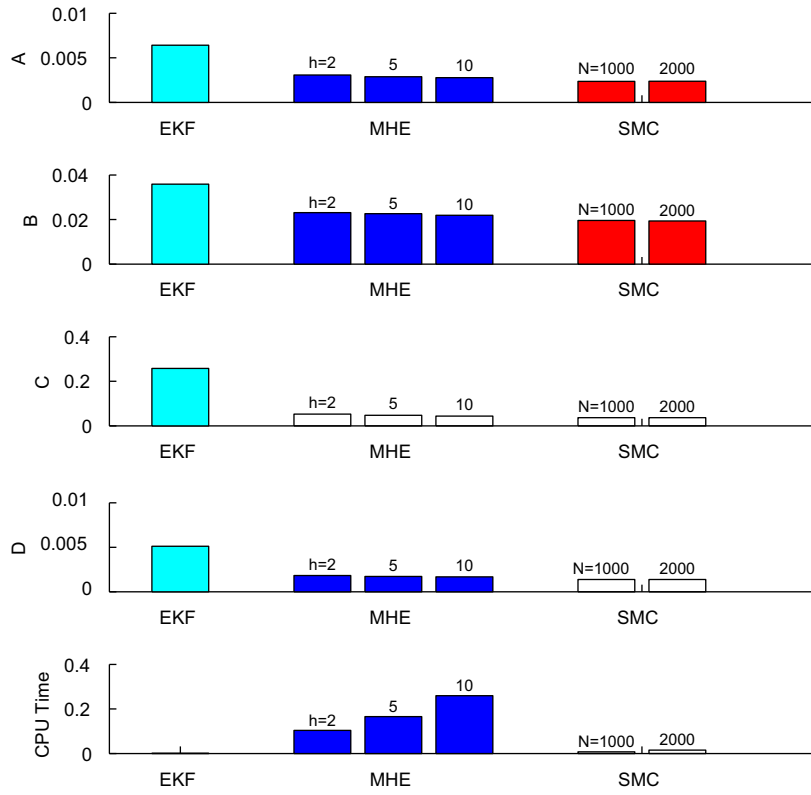


Fig. 4. MSE and CPU time comparison for the constrained McKeithan network.

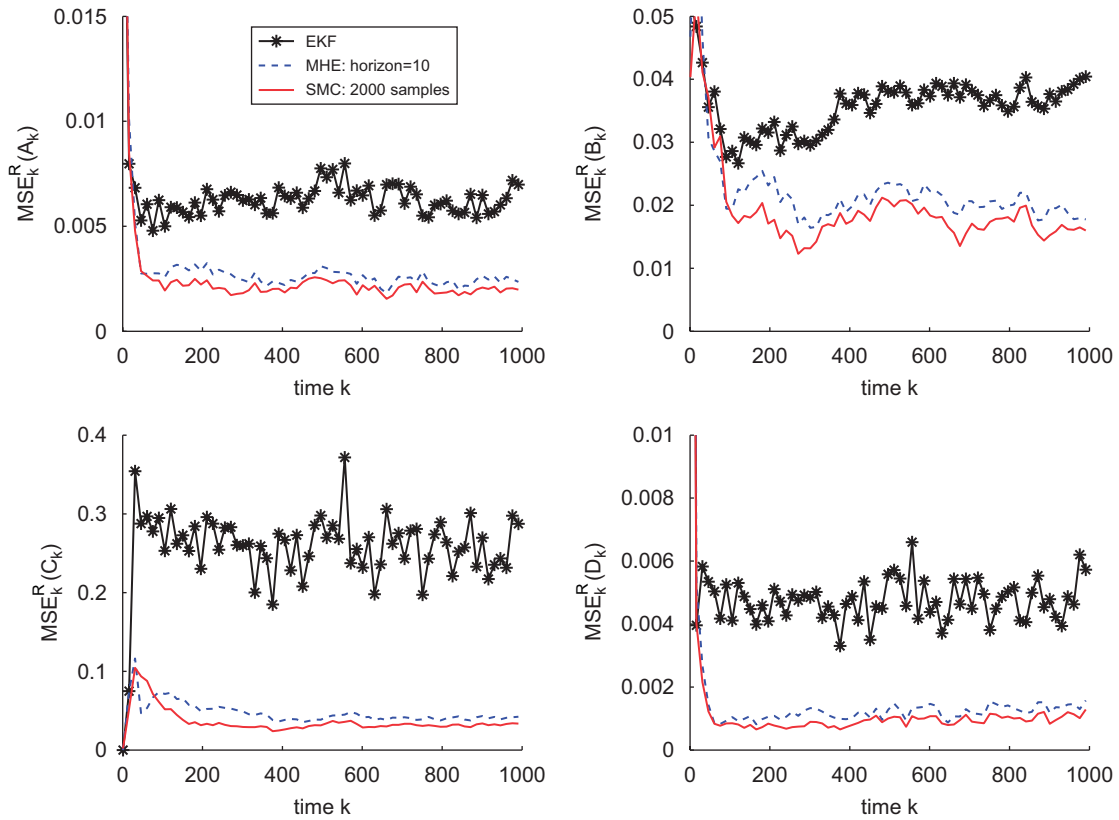


Fig. 5. MSE_k^R comparison for the constrained McKeithan network.

in the beginning but has an abrupt increase in errors after time 100. In fact, this process experiences rapid state changes at that time and the concentration level drops nearly to zero, which is the constraint value. This shows that EKF suffers due to linearization and its inability to enforce constraints. In contrast, MHE uses the smoothing approach for estimating the arrival cost for each window. For this nonlinear problem, the smoother does not reduce to EKF, but utilizes information about the measurements in each window to estimate a more accurate arrival cost than that obtained via filtering (Tenny, 2002). Compared to EKF and MHE, SMC uses neither linearization nor Gaussian approximation. For SMC, the dynamic process is well approximated by the particles that satisfy the constraint.

4.2. Constrained McKeithan network

The McKeithan reaction network represents the kinetics of signaling in a cellular system (Chaves & Sontag, 2002; McKeithan, 1995). Compared to the previous model, it is nonlinear in both the state and measurement equations. In addition, one state component, C , is not present in the measurement equation. The differential form of this model is described by

$$\dot{A} = -k_1 AB + k_3 C + k_4 D,$$

$$\dot{B} = -k_1 AB + k_3 C + k_4 D,$$

$$\dot{C} = k_2 AB - (k_3 + \beta_3) C,$$

$$\dot{D} = \beta_3 C - k_4 D,$$

$$y = [AB^2 \quad AD]^T.$$

Details about model parameters and discretization are in Lang et al. (2006). The initial value is $x_1 = (1, 3, 3, 2)^T$, and the constraint is set to D_k such that $D_k > 0.7$ for all k . The system noise is iid Gaussian with covariance $\Sigma_w = \sigma_w^2 I_4$, where $\sigma_w^2 = 10^{-4}$. The measurement noise is also iid Gaussian with covariance $\Sigma_v = I_2$. The initial prior distribution is set to be $N(x_1, 0.5I_4)$. There are 1000 measurements.

The MSE for different methods is shown in Fig. 4. Like the previous case, EKF is worse than MHE, which is worse than SMC. Furthermore, the MSE_k^R values shown in Fig. 5 for EKF and MHE are worse than SMC for most of the time steps due to the non-Gaussian distributions shown in Fig. 1, which are caused by the complex nonlinearities in the state and measurement equations.

5. Conclusions

This paper described a practical approach for extending SMC sampling or particle filtering based Bayesian estimation to constrained dynamic systems. The proposed algorithm is based on previous work on estimation of unconstrained nonlinear dynamic systems (Gordon et al., 1993). It enforces inequality constraints via an acceptance/rejection algorithm that ensures that all samples representing the likelihood, prior or posterior

satisfy relevant constraints. Equality constraints may be readily imposed via the proposed approach by including them in the state or measurement equations. The resulting formulation is general and may be combined with modifications of SMC such as alternate ways of constructing importance functions, dealing with poor initial guess, or enforcing constraints. The proposed algorithm is also shown to possess the theoretical properties of unconstrained SMC. Application of constrained SMC to various case studies indicates that the proposed approach is capable of handling constraints and can provide accurate estimates with efficient computation. Constrained SMC easily outperforms EKF in estimation accuracy and also shows improvement in estimation accuracy over MHE especially when the posterior distributions at each time point cannot be adequately represented by multivariate Gaussian or other fixed-shape distributions. In terms of computational efficiency, SMC consistently requires less CPU time than MHE especially in nonlinear dynamic systems.

Acknowledgments

We thank Prof. Rawlings for the MHE code; Dr. Haseltine for assistance with this program; an anonymous reviewer for his/her insight; and National Science Foundation (CTS-0321911) for financial support.

References

- Arulampalam, M. S., Maskell, S., Gordon, N., & Clapp, T. (2002). A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking. *IEEE Transactions on Signal Processing*, 50(2), 174–188.
- Chaves, M., & Sontag, E. D. (2002). State-estimator for chemical reaction networks of Feinberg–Horn–Jackson zero deficiency type. *European Journal of Control*, 8(4), 343–359.
- Chen, W.-S., Bakshi, B. R., Goel, P. K., & Ungarala, S. (2004). Bayesian estimation of unconstrained nonlinear dynamic systems via sequential Monte Carlo sampling. *Industrial & Engineering Chemistry Research*, 43(14), 4012–4025.
- Crisan, D. (2001). Particle filters—A theoretical perspective. In A. Doucet, N. de Freitas, & N. Gordon (Eds.), *Sequential Monte Carlo methods in practice* (pp. 17–41). New York: Springer, (Chapter 2).
- Crisan, D., & Doucet, A. (2002). A survey of convergence results on particle filtering methods for practitioners. *IEEE Transactions on Signal Processing*, 50(3), 736–746.
- Del Moral, P., & Miclo, L. (2000). Branching and interacting particles systems. Approximations of feynman-kac formulae with applications to non-linear filtering. *Séminaire de Probabilités XXXIV. Lecture notes in mathematics* (pp. 1–145). Berlin: Springer.
- Doucet, A., Godsill, S., & Andrieu, C. (2000). On sequential Monte Carlo sampling methods for Bayesian filtering. *Statistics and Computing*, 10, 197–208.
- Geweke, J. (1989). Bayesian inference in econometric models using Monte Carlo integration. *Econometrica*, 57(6), 1317–1339.
- Gordon, N. J., Salmond, D. J., & Smith, A. F. M. (1993). Novel approach to nonlinear/non-Gaussian Bayesian state estimation. *IEE Proceedings-F*, 140(2), 107–113.
- Haseltine, E. L., & Rawlings, J. B. (2005). Critical evaluation of extended Kalman filtering and moving-horizon estimation. *Industrial & Engineering Chemistry Research*, 44, 2451–2460.
- Henson, M. A., & Seborg, D. E. (1997). *Nonlinear process control*. Upper Saddle River, NJ: Prentice-Hall PTR.

- Jang, S.-S., Joseph, B., & Mukai, H. (1986). Comparison of two approaches to on-line parameter and state estimation of nonlinear systems. *Industrial and Engineering Chemistry Process Design and Development*, 25, 809–814.
- Jazwinski, A. H. (1970). *Stochastic processes and filtering theory*. New York: Academic Press.
- Julier, S. J., & Uhlmann, J. K. (2004). Unscented filtering and nonlinear estimation. *Proceedings of IEEE*, 92(3), 401–422.
- Künsch, H. R. (2005). Recursive Monte Carlo filters: algorithms and theoretical analysis. *Annals of Statistics*, 33(5), 1983–2021.
- Lang, L., Chen, W.-S., Bakshi, B. R., Goel, P. K., & Ungarala, S. (2006). *Bayesian rectification via sequential Monte Carlo sampling of constrained systems*. Technical Report, Ohio State University.
- McKeithan, T. W. (1995). Kinetic proofreading in t-cell receptor signal transduction. *Proceedings of the national academy of sciences*, 92, 5042–5046.
- Robertson, D. G., Lee, J. H., & Rawlings, J. B. (1996). A moving horizon-based approach for least-squares estimation. *AIChE Journal*, 42(8), 2209–2224.
- Tenny, M. J. (2002). *Computational strategies for nonlinear model predictive control*. Ph.D. thesis, University of Wisconsin-Madison.
- Ungarala, S., & Bakshi, B. R. (2001). Multiscale Bayesian estimation and data rectification. In A. A. Petrosian, & F. G. Meyer (Eds.), *Wavelets in signal and image analysis* (pp. 69–110). The Netherlands: Kluwer.