Journal of Process Control 20 (2010) 143–157

Contents lists available at ScienceDirect

Journal of Process Control

journal homepage: www.elsevier.com/locate/jprocont

Constrained Bayesian state estimation – A comparative study and a new particle filter based approach

Xinguang Shao, Biao Huang *, Jong Min Lee

Department of Chemical and Materials Engineering, University of Alberta, Edmonton, AB, Canada T6G 2G6

ARTICLE INFO

Article history:
Received 8 April 2009
Received in revised form 2 November 2009
Accepted 2 November 2009

Keywords:
State estimation
Bayesian
Particle Filter
Constraints
Nonlinear
Non-Gaussian

ABSTRACT

This paper investigates constrained Bayesian state estimation problems by using a Particle Filter (PF) approach. Constrained systems with nonlinear model and non-Gaussian uncertainty are commonly encountered in practice. However, most of the existing Bayesian methods are unable to take constraints into account and require some simplifications. In this paper, a novel constrained PF algorithm based on acceptance/rejection and optimization strategies is proposed. The proposed method retains the ability of PF in nonlinear and non-Gaussian state estimation, while take advantage of optimization techniques in constraints handling. The performance of the proposed method is compared with other accepted Bayesian estimators. Extensive simulation results from three examples show the efficacy of the proposed method in constraints handling and its robustness against poor prior information.

1. Introduction

State estimation plays a key role in different applications such as process monitoring, fault detection, process optimization and model based control. Thanks to their distinct advantages in representing uncertainties, Bayesian methods have been extensively studied and used successfully in many areas in the past several decades [12,10,6,30]. Though nonlinear and non-Gaussian processes subject to state constraints are commonly encountered in practical applications, most of the existing Bayesian methods are unable to take constraints into account and require assumptions of linearity or Gaussianity. Therefore, development of Bayesian estimators that can handle nonlinear and non-Gaussian problems with constraints would be useful and has recently become an active research area [27,40,11,39,30,16,17,35].

It is well known that the Bayesian estimation is performed through recursive update of conditional distributions, and the analytical solution is not available in general except for special cases such as Kalman Filter (KF) for unconstrained linear systems with Gaussian noise. For the nonlinear and/or non-Gaussian cases, many suboptimal techniques have been developed, for example, Extended Kalman Filter (EKF) [2], Unscented Kalman Filter (UKF) [15], Moving Horizon Estimator (MHE) [31,27] and sequential Monte Carlo (SMC) filter, also known as Particle Filter (PF) [8,4].

Among these methods, EKF is the most widely used one in applications because of its close resemblance to KF. However, it has been demonstrated that EKF has inherent limitations when applied to systems with severe nonlinearity, and the structure of the algorithm makes EKF unsuitable for constrained systems. UKF avoids the linearization step, and can show a similar performance as EKF. However, the deterministic selection of the "sigma points" limits the flexibility needed to shape a distribution. Variants of KF, EKF and UKF with constraints handling are recently discussed in Vachhani et al. [40,39], Simon [33], Ungarala et al. [37], Julier and LaViola [14], Teixeira et al. [36,35], Kandepu et al. [16], and Kolås et al. [18].

MHE performs state estimation by solving an optimization problem in a receding horizon fashion, which inherently handles constraints for both linear and nonlinear systems [29,28]. However, MHE requires calculations of arrival cost, and the multi-stage nonlinear optimization incurs excessive computational burden for on-line applications [43].

Recently, Monte Carlo sampling based PF has received growing attentions. Since it does not necessitate simplification of nonlinearity or assumption of specific distributions, it is a favorable method for nonlinear and non-Gaussian state estimation problems. Moreover, the recursive update of conditional distributions within a single-horizon window makes PF more computationally attractive than multi-horizon setup of MHE. It has been demonstrated that PF can readily impose constraints onto particles. Preliminary contributions can be found in Lang et al. [23], in which an acceptance/rejection method is used for dealing with inequality

* Corresponding author. Tel.: +1 780 492 9016; fax: +1 780 492 2881.
E-mail address: biao.huang@ualberta.ca (B. Huang).

0959-1524/$ - see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jprocont.2009.11.002

constraint. However, the proposed approach is difficult to handle complicated constraints, such as nonlinear constraints, equality–inequality mixed constraints. Furthermore, the nature of the acceptance/rejection approach is simply removing all the particles outside the constraint region. In some cases (e.g., with a very poor prior estimate), the method could fail due to insufficient number of valid particles. Rajamani and Rawlings [26] also have some preliminary discussions on the combination of PF and MHE; the method is based on optimization technique, which is more sophisticated to handle different types of constraints, but their optimization scheme is applied to the sample mean only, which may not take the full advantage of particle filter, especially when the posterior distribution is non-Gaussian. For instance, in a multimodal case, it is more appropriate to track and constrain individual particles instead of the sample mean only, since the mean could be located between the modes in a feasible region, but as a very poor estimate.

In this paper, a constrained PF algorithm based on hybrid use of acceptance/rejection and optimization strategies is proposed. The proposed method combines the ability of PF to handle nonlinear and non-Gaussian problems and the advantages of optimization techniques in constraints handling. Furthermore, simulation results show that the proposed method enhances the robustness of PF algorithm against poor prior information.

The remainder of this paper is organized as follows: Section 2 introduces the constrained Bayesian state estimation problem. Section 3 briefly reviews several existing Bayesian estimators and their applications for constrained state estimation. Section 4 provides an overview of the generic PF algorithm. In Section 5, two constraint handling strategies are discussed within the generic PF framework, and a novel constrained PF algorithm is proposed. Three examples are illustrated in Section 6. Section 7 gives the conclusions.

2. Problem statement

Consider a discrete time system given by

$$\begin{aligned} x_k &= f_k(x_{k-1}, u_{k-1}) + \omega_{k-1}, \\ y_k &= h_k(x_k) + v_k, \end{aligned} \tag{1}$$

where x_k, u_k, y_k, ω_k and v_k are state, input, output, process noise and measurement noise, respectively; $f_k(\cdot)$, $h_k(\cdot)$ are nonlinear functions; both v_k and ω_k are white noise of possibly non-Gaussian; initial state x_0 may also follow a non-Gaussian distribution $p(x_0)$. System (1) can be alternatively defined in a conditional distribution form by using $p(x_k|x_{k-1})$ and $p(y_k|x_k)$. Note that the variables x_k, y_k, ω_k and v_k are random, while the input term u_k is deterministic. For the simplicity, input term is not considered in the remainder of this paper.

2.1. Bayesian estimation

The objective of Bayesian estimation is to reconstruct the conditional *a posteriori* probability density function (pdf) $p(x_k|Y_k)$, where x_k is the vector of states at time k, and $Y_k = \{y_1, \ldots, y_k\}$ is the vector of noisy measurements up to and including time k. The solution is obtained by recursively solving the following equations based on the Bayes' rule, also known as recursive prediction and update procedures [8]:

$$p(x_k|Y_{k-1}) = \int p(x_k|x_{k-1})p(x_{k-1}|Y_{k-1})dx_{k-1}, \tag{2}$$

$$p(x_k|Y_k) = \frac{p(y_k|x_k)p(x_k|Y_{k-1})}{\int p(y_k|x_k)p(x_k|Y_{k-1})dx_k} = \frac{p(y_k|x_k)p(x_k|Y_{k-1})}{p(y_k|Y_{k-1})}. \tag{3}$$

By definition, it can be seen that the posterior pdf $p(x_k|Y_k)$ is described by three terms: prior pdf $p(x_k|Y_{k-1})$, likelihood pdf $p(y_k|x_k)$ and evidence pdf $p(y_k|Y_{k-1})$. Calculations of these three terms are the main steps of Bayesian estimation.

For unconstrained linear models with Gaussian uncertainties, the optimal solution is given by the closed-form Kalman filter. However, the Gaussianity cannot be guaranteed due to nonlinearity, constraints, or non-Gaussian noise. Therefore, the closed-form solution is generally not possible, and suboptimal methods (e.g., EKF and UKF) have to be resorted to.

Since $p(x_k|Y_k)$ embodies all the statistical information about x_k, the posterior pdf $p(x_k|Y_k)$ is the complete solution of state estimation problem. In many applications, most interests lie in point estimation, such as estimation of mode, mean or median, depending on the criterion for estimation. Fig. 1 demonstrates that these estimates are generally different except for Gaussian distribution.

It should be noted that there are two assumptions in Eqs. (2) and (3): (i) the states follow a first-order Markov process, i.e., $p(x_k|X_{k-1}, Y_{k-1}) = p(x_k|x_{k-1})$, where $X_{k-1} = \{x_0, \ldots, x_{k-1}\}$; (ii) the observations are conditionally independent given the state, i.e., $p(y_k|X_k, Y_{k-1}) = p(y_k|x_k)$.

2.2. Constrained Bayesian estimation

In practical applications, nonlinear and non-Gaussian processes with constraints are commonly encountered. These constraints, which stem from the physical laws or model restrictions, e.g. non-negative mole fractions, limited liquid levels, mass balance, bounded parameters/disturbances, etc., are usually in the form of algebraic equality and inequality relationships, or simply upper and lower bounds. Incorporation of such constraints into estimation will be useful for improving estimation performance.

Take a multimodal posterior pdf (Fig. 2) as an example; for the maximum *a posteriori* (MAP) state estimation where state x_k is concentration,

$$\hat{x}_k := \arg\max_{x_k} p(x_k|Y_k). \tag{4}$$

Mathematically, two optimal solutions can be obtained: one is negative (mode 1) and the other is positive (mode 2). From the knowledge on the constraint (i.e., $x \geqslant 0$), it is easy to find the correct estimate (mode 2).

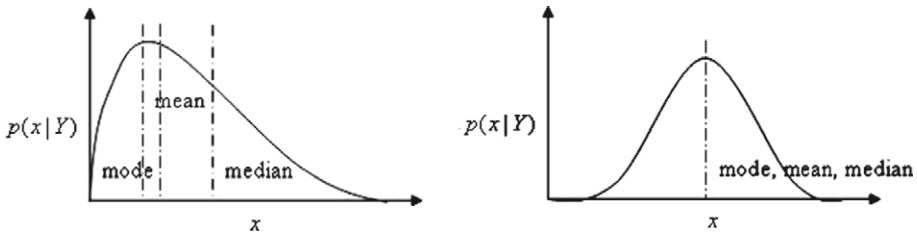
Constraints of stochastic variables affect estimation by reshaping their pdfs in Bayesian framework. For instance, constraints on process noise restrict the transition distribution, $p(x_k|x_{k-1})$; constraints on measurement noise have an influence on the likelihood distribution, $p(y_k|x_k)$; and constraints on states alter the posterior distribution, $p(x_k|y_{1:k})$, as well as the transition and likelihood distributions [38]. Therefore, use of these constraints confines the distributions, leading to improvement of estimation accuracy.

3. Existing Bayesian estimators

This section briefly discusses several existing Bayesian estimators and their applications in constrained estimation to serve as comparisons with the following Monte Carlo sampling based particle filtering approach.

3.1. Kalman filtering based methods

Kalman-filter structure based estimators are widely used for state estimation problems. These methods assume all the system variables follow Gaussian distributions whose statistical information can be fully described by the first and second order moments (i.e., mean and covariance), and the estimate is given by



(a) non-Gaussian pdf (b) Gaussian pdf

Fig. 1. Illustration of point state estimation.

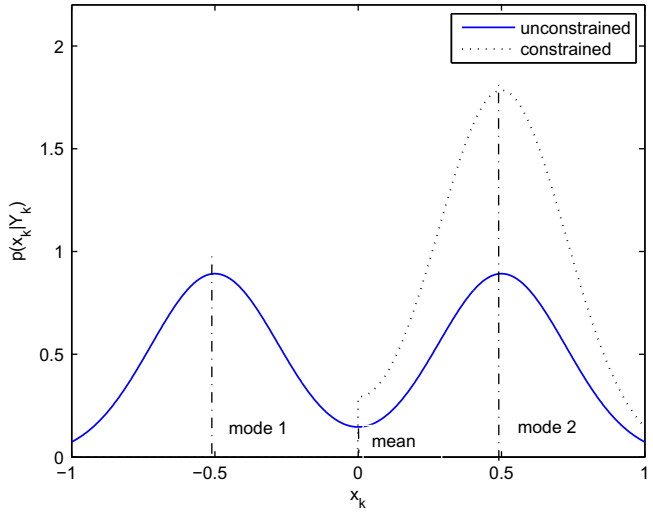


Fig. 2. An example for a multimodal pdf.

$$\hat{x}_k = \hat{x}_k^- + K_k(y_k - \hat{y}_k), \tag{5}$$

where $\hat{x}_k^-$ is the prior estimate based on the information of $p(x_k|Y_{k-1})$; K_k is the Kalman gain at time k. When $f_k(\cdot)$ and $h_k(\cdot)$ in Eq. (1) are both linear functions, Gaussianity is kept all the time, and Kalman filter gives the optimal solution.

3.1.1. Extended Kalman Filter

For nonlinear cases, Gaussianity is no longer guaranteed, and thus approximate solutions are needed. The most popular approximate method based on model linearization is Extended Kalman Filter (EKF), in which the mean and covariance of the posterior distribution approximated as Gaussian are calculated as:

$$\begin{aligned}
&\hat{x}_k^- = f_k(\hat{x}_{k-1}),\\
&F_k = \left.\frac{\partial f_k}{\partial x}\right|_{\hat{x}_{k-1}},\\
&P_k^- = F_k P_{k-1} F_k^T + Q_{k-1},\\
&\hat{y}_k = h_k(\hat{x}_k^-),\\
&H_k = \left.\frac{\partial h_k}{\partial x}\right|_{\hat{x}_k^-},\\
&S_k = H_k P_{k-1} H_k^T + R_k,\\
&K_k = P_k^- H_k^T S_k^{-1},\\
&\hat{x}_k = \hat{x}_k^- + K_k(y_k - \hat{y}_k),\\
&P_k = (I - K_k H_k) P_k^-,
\end{aligned} \tag{6}$$

where Q_{k-1} and R_k are the covariance matrices of the system noise, ω_{k-1}, and measurement noise, v_k, respectively.

The main disadvantages of EKF include: (i) approximated linear model can be inaccurate for highly nonlinear cases, in which estimate may fail to converge to the true state; (ii) update of covariance needs calculation of Jacobian matrices, which can be cumbersome and (iii) constraints are not considered in the algorithm.

For the third issue, a common way to handle constraints in EKF, as well as KF, is "clipping" [11,33], where the estimated state is projected right on the boundary if outside the constraint region. One drawback is that the constraint information is not properly included in the covariance update, which can lead to a poor estimate [16]. Vachhani et al. [40] and Simon [33] propose to combine KF/EKF with optimization formulation to handle equality and inequality constraints. The approaches require higher computational effort but significantly improve estimation accuracy. Gupta and Hauser [9] adopted the optimization strategy to restrict the Kalman gain in KF/EKF to force the estimated state to be in constraint regions.

3.1.2. Unscented Kalman Filter

Instead of approximating nonlinear model, Unscented Kalman Filter (UKF) approximates the posterior distribution by Gaussian distribution directly. It uses a set of deterministically chosen "sigma points" to represent mean and covariance. An estimation procedure for a fully augmented UKF is shown as follows:

$$\begin{aligned}
&\chi_{k-1} = \left[\hat{x}_{k-1}^a \;\; \hat{x}_{k-1}^a + \sqrt{(n^a+\kappa)P_{x_{k-1}}^a} \;\; \hat{x}_{k-1}^a - \sqrt{(n^a+\kappa)P_{x_{k-1}}^a}\right],\\
&\chi_{k,i}^{x,-} = f_k(\chi_{k-1,i}^x) + \chi_{k-1,i}^\omega,\\
&\hat{x}_k^- = \sum_{i=0}^{2n^a} W_i^x \chi_{k,i}^{x,-},\\
&P_k^- = \sum_{i=0}^{2n^a} W_i^c (\chi_{k,i}^{x,-} - \hat{x}_k^-)(\chi_{k,i}^{x,-} - \hat{x}_k^-)^T,\\
&\gamma_{k,i} = h_k(\chi_{k,i}^{x,-}, \chi_{k,i}^v),\\
&\hat{y}_k = \sum_{i=0}^{2n^a} W_i^x \gamma_{k,i},\\
&P_{y_k y_k} = \sum_{i=0}^{2n^a} W_i^c (\gamma_k^i - \hat{y}_k)(\gamma_k^i - \hat{y}_k)^T,\\
&P_{x_k y_k} = \sum_{i=0}^{2n^a} W_i^c (\chi_{k,i}^{x,-} - \hat{x}_k^-)(\gamma_{k,i} - \hat{y}_k)^T,\\
&K_k = P_{x_k y_k} P_{y_k y_k}^{-1},\\
&\hat{x}_k = \hat{x}_k^- + K_k(y_k - \hat{y}_k),\\
&P_k = P_k^- + K_k P_{y_k y_k} K_k^T,
\end{aligned} \tag{7}$$

where $\chi_{k-1} = \left[\chi_{k-1}^{x^T} \;\; \chi_{k-1}^{\omega^T} \;\; \chi_{k-1}^{v^T}\right]^T$ is the "sigma points" of the augmented state, $x_{k-1}^a = [x_{k-1}^T \;\; \omega_{k-1}^T \;\; v_{k-1}^T]^T$, with mean and covariance as

$$\hat{x}^a_{k-1} = \begin{bmatrix} \hat{x}^T_{k-1} & 0 & 0 \end{bmatrix}^T,$$
$$P^a_{x_{k-1}} = \begin{bmatrix} P_{k-1} & 0 & 0 \\ 0 & Q_{k-1} & 0 \\ 0 & 0 & R_{k-1} \end{bmatrix}.$$

$n^a = n^x + n^\omega + n^v$ is the dimension of the augmented state; κ is a tunning parameter; W^x_i and W^c_i are weights for state and covariance. For more details, see van der Merwe [41].

Note that fully augmented UKF is not necessary for all the situations; in order to save computation time for specific problems, readers can refer to Kolås et al. [18] for various versions of UKF algorithms.

Compared to EKF, UKF does not resort to linearization and is free of Jacobian calculation. However, UKF cannot handle constraints either. Kandepu et al. [16] apply "clipping" method in UKF by projecting "sigma points" instead of state, onto the constraint boundary. The advantage is that both the mean and covariance are computed based on the constraint information, making UKF estimate more accurate. Vachhani et al. [39] combine the concept of unscented transformation with optimization formulation for constrained estimation. Teixeira et al. [34,36] propose different UKF approaches based on gain-constrained, truncation and projection strategies to address constraints. Kolås et al. [18] reformulate the UKF update step, and demonstrate that state constraints can be incorporated into different "constraint candidates", i.e., "sigma points" or state in either prediction or update step.

3.2. Moving horizon estimator

An alternative formulation of Eq. (4) is to maximize a conditional *a posteriori* pdf for a sequence of the state trajectory,

$$\{\hat{x}_{k-h}, \ldots, \hat{x}_k\} := \arg\max_{x_{k-h},\ldots,x_k} p(x_{k-h}, \ldots, x_k | Y_k), \tag{8}$$

where $h \in \{0, k\}$ is known as a time horizon parameter.

Using Bayes' rule and Markov assumption, one can have

$$p(x_{k-h}, \ldots, x_k | Y_k) \propto \prod_{j=k-h}^{k} p(y_j|x_j) \prod_{j=k-h}^{k-1} p(x_{j+1}|x_j) p(x_{k-h}|Y_{k-h-1}), \tag{9}$$

where $p(x_{k-h}|Y_{k-h-1})$ is the *a priori* information.

By assuming Gaussian distributions, a quadratic optimization problem can be formulated for solving Eq. (9):

$$\begin{aligned} \min_{x^e_{k-h}, \hat{\omega}_{k-h}, \ldots, \hat{\omega}_{k-1}} \quad & {x^e_{k-h}}^T P^{-1}_{k-h} x^e_{k-h} + \sum_{j=k-h}^{k-1} \hat{\omega}_j^T Q^{-1} \hat{\omega}_j + \sum_{k-h}^{k} \hat{v}_j^T R^{-1} \hat{v}_j \\ \text{s.t.} \quad & \hat{x}_{k-h} = \hat{x}^-_{k-h} + x^e_{k-h}, \\ & \hat{x}_{j+1} = f_k(\hat{x}_j) + \hat{\omega}_j, \quad j = k-h, \ldots, k-1, \\ & y_j = h_j(\hat{x}_j) + \hat{v}_j, \quad j = k-h, \ldots, k, \\ & \hat{x}_j \in \mathbb{X}, \ \hat{\omega}_j \in \mathbb{W}, \ \hat{v}_j \in \mathbb{V}. \end{aligned} \tag{10}$$

Eq. (10) is known as Moving Horizon Estimator (MHE), which can be viewed as a form of iterated EKF [1] for unconstrained system with a horizon size $h = 1$ [27]. The advantage of MHE is that constraints for state or noise can be naturally incorporated into the problem formulation. However, the major problem for MHE is the computational load (see Robertson et al. [31], Rao [27], Rao and Rawlings [28], Rawlings and Bakshi [30], Zavala et al. [43], and reference therein).

4. Particle Filter

Unlike most other Bayesian estimators, PF does not rely on linearization technique or Gaussian assumption. It approximates any *a posteriori* pdf by a set of particles, x^i_k, and their associated weights, $w^i_k \geqslant 0$, in a discrete summation form:

$$\hat{p}(x_k|y_{1:k}) = \sum_{i=1}^{N} w^i_k \delta(x_k - x^i_k), \tag{11}$$

where $\delta(\cdot)$ is the Dirac delta function, and N is the number of particles.

The ideal case for Monte Carlo sampling is to generate particles directly from the true posterior pdf $p(x_k|Y_k)$, which is generally unavailable. Thus an easy-to-implement distribution, the so called *importance sampling distribution* denoted by $q(x_k|X_{k-1}, Y_k)$, is defined before sampling. It is shown that PF is asymptotically unbiased if the support region for $p(x_k|Y_k)$ is a subset of that for $q(x_k|X_{k-1}, Y_k)$ [7]. Liu and Chen [24] show that the optimal importance sampling distribution is:

$$q(x_k|X_{k-1}, Y_k) = p(x_k|X_{k-1}, Y_k). \tag{12}$$

With this choice, one can derive that the importance weights can be computed sequentially as:

$$\tilde{w}^i_k = \tilde{w}^i_{k-1} \frac{p(y_k|x^i_k) p(x^i_k|x^i_{k-1})}{p(x^i_k|X^i_{k-1}, Y_k)}. \tag{13}$$

However, $p(x^i_k|X^i_{k-1}, Y_k)$ is also generally unknown except for linear observation model cases [7]. In practice, the transition kernel is often used for importance sampling, i.e. $q(x_k|X_{k-1}, Y_k) = p(x_k|x_{k-1})$, since the state space model is already known. Eq. (13) then reduces to

$$\tilde{w}^i_k = \tilde{w}^i_{k-1} p(y_k|x^i_k). \tag{14}$$

In order to avoid particles degeneracy problem [42], the importance sampling step is usually followed by a resampling procedure to place the particles in regions with high probability. For different resampling strategies, readers are referred to Boloc et al. [3] and references therein.

Estimation steps of the generic PF (GPF) algorithm is summarized as follows:

Algorithm 1. The GPF algorithm

Step (a) *Initialization*: Generate initial particles $\{x^i_0\}_{i=1}^N$ from *a priori* distribution $p(x_0)$, and set $k = 1$;
Step (b) *Importance sampling*: Generate prior particles, $\{x^{i,-}_k\}_{i=1}^N$, from importance sampling distribution $p(x^{i,-}_k|x^i_{k-1})$;
Step (c) *Weighting*: Evaluate weights of each particle once new measurement is available and normalize the weights as $w^i_k = \tilde{w}^i_k / \sum_{j=1}^N \tilde{w}^j_k$;
Step (d) *Resampling*: Generate posterior particles, $\{x^i_k\}_{i=1}^N$, based on weighting information and resampling strategy, and set $w^i_k = 1/N$;
Step (e) *Output*: Estimate the state by calculating $\hat{x}_k = \sum_{i=1}^N w^i_k \cdot x^i_k$, set $k = k + 1$ and go back to Step (b).

5. Constrained Particle Filter

As shown in Algorithm 1, the generic PF does not consider constraints. In this section, two methods are introduced to handle constraints in the PF framework, and then a new constrained PF algorithm is proposed. The work discussed here can be applied to variants of PF, such as Auxiliary Particle Filter (APF) [25], Unscented Particle Filter (UPF)[42], and Kernel Particle Filter (KPF) [5].

5.1. Acceptance/rejection

The nature of sample based representation of PF facilitates incorporating constraints into the estimation procedure. Lang et al. [23] and Kyriakides et al. [22] discuss how to accept/reject the particles in the PF algorithm based on constraint knowledge. As a minor modification from their work, a constrained likelihood function is defined as:

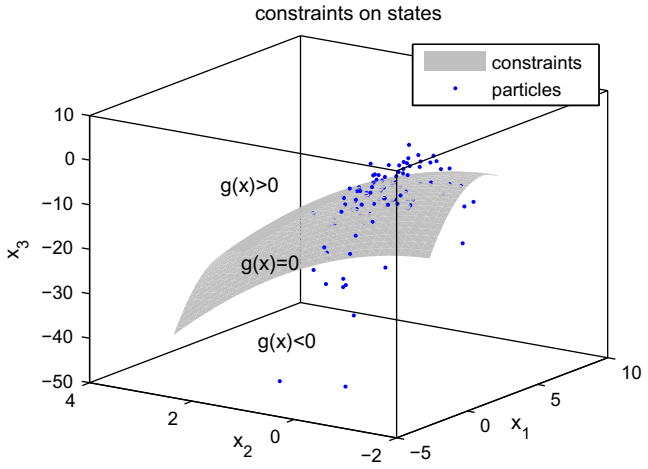


Fig. 3. An example for constraints on states.

$$L_c(x_k^i, y_k^i, \omega_k^i, v_k^i) = \begin{cases} 1, & \text{if } \{x_k^i, y_k^i, \omega_k^i, v_k^i\} \in C_k, \\ 0, & \text{if } \{x_k^i, y_k^i, \omega_k^i, v_k^i\} \notin C_k, \end{cases} \quad i = 1, \ldots, N, \tag{15}$$

where C_k represents a constraint region at time k. Then the weight calculation step, Eq. (13), is modified as

$$\tilde{w}_k^i = w_{k-1}^i \frac{p(y_k|x_k^i) \cdot L_c(x_k^i, y_k^i, \omega_k^i) \cdot p(x_k^i|x_{k-1}^i)}{q(x_k^i|X_{k-1}^i, Y_k)}. \tag{16}$$

This modification enables the algorithm to discard all the particles violating constraints. Fig. 3 depicts an example for constraints on state. Take the equality constraint case, i.e. $g(x) = 0$, as an example, only the particles on the constraint surface will be accepted and reproduced, and all the rest particles will be rejected.

The advantage of acceptance/rejection scheme is twofold. First, it guarantees the particles to stay in constraint region and nearly no extra computation cost is needed. Second, the method retains the Monte Carlo sampling feature of PF which makes it suitable for non-Gaussian problems. However, the disadvantage is that it reduces the number of particles and may yield poor estimation. With poor prior information or complicated constraint conditions (e.g., nonlinear constraints), it is possible that all the particles lie outside the constraint region, which fails the PF algorithm.

5.2. Optimization formulation

A more systematic way to deal with constraints without discarding any particles is to employ optimization technique. In this section, optimization methods used to handle constraints are discussed in PF framework.

5.2.1. Interpretation of Bayesian estimation as an optimization problem

The estimate in Eq. (4) can be further written as

$$\begin{aligned} \hat{x}_k &:= \arg\max_{x_k} p(x_k|Y_k) \propto \arg\max_{x_k} p(y_k|x_k) p(x_k|Y_{k-1}) \\ &= \arg\max_{x_k} (p_{v_k}(y_k - h_k(x_k)) p_{x_k^e}(x_k - x_k^-)). \end{aligned} \tag{17}$$

Note that the measurement noise v_k follows distribution p_{v_k}. Let $x_k = x_k^- + x_k^e$, where x_k^- is the optimal estimate of x_k according to $p(x_k|Y_{k-1})$, and x_k^e is the estimation error which follows distribution $p_{x_k^e}$. Note that exponential and double exponential (Laplacian) distributions are usually used to prescribe pdfs of ω, v and x (Kotz et al. [21, p. 278]; Robertsonb and Lee [32]; Ungarala et al. [38]).

The above equation can be rewritten as the following constrained optimization problem by taking negative logarithm:

$$\begin{aligned} \min_{x_k} \quad & -\log(p_{x_k^e}(x_k - x_k^-)) - \log(p_{v_k}(y_k - h_k(x_k))) \\ \text{s.t.} \quad & x_k = x_k^- + x_k^e, \\ & y_k = h_k(x_k) + v_k, \\ & x_k \in \mathbb{X}_k, \end{aligned} \tag{18}$$

where $\mathbb{X}_k$ denotes a general state constraint region.

If both $p_{x_k^e}$ and p_{v_k} are further assumed as Gaussian distributions, Eq. (18) becomes a constrained nonlinear least square problem:

$$\begin{aligned} \min_{x_k^e} \quad & x_k^{eTP} P_k^{-1} x_k^e + v_k^T R_k^{-1} v_k \\ \text{s.t.} \quad & x_k = x_k^- + x_k^e, \\ & y_k = h_k(x_k) + v_k, \\ & x_k \in \mathbb{X}_k, \end{aligned} \tag{19}$$

where P_k^{-1} and R_k^{-1} may be treated as the weighting matrices, which are quantitative measures of our belief in the prior estimate and the observation model, respectively. Note that for linear models without constraints, solution of Eq. (19) is equivalent to the well known Kalman filter estimate [13, pp. 205–208].

5.2.2. State constraints imposed on particles

According to the steps in Algorithm 1, state constraints in PF can be imposed onto prior particles (particles before resampling procedure), $x_k^{i,-}$, posterior particles (particles after resampling procedure), x_k^i, or estimated mean value, $\hat{x}_k$. The constrained optimization problem presented in Eq. (18) can be adapted as one of the followings:

$$\min_{\tilde{x}_k^{i,-}} -\log(p_{x_k^e}(\tilde{x}_k^{i,-} - x_k^{i,-})), \tag{20}$$

$$\min_{\tilde{x}_k^{i,-}} -\log(p_{x_k^e}(\tilde{x}_k^{i,-} - x_k^{i,-})) - \log(p_{\hat{v}_k}(y_k - h_k(\tilde{x}_k^{i,-}))), \tag{21}$$

$$\min_{\tilde{x}_k^i} -\log(p_{x_k^e}(\tilde{x}_k^i - x_k^i)) - \log(p_{\hat{v}_k}(y_k - h_k(\tilde{x}_k^i))), \tag{22}$$

$$\min_{\tilde{x}_k} -\log(p_{x_k^e}(\tilde{x}_k - \hat{x}_k)) - \log(p_{\hat{v}_k}(y_k - h_k(\tilde{x}_k))), \tag{23}$$

where the diacritic mark "$\sim$" placed above $x_k^{(\cdot)}$ indicates a projected particle/mean. Fig. 4 shows an illustration of particle projection, in which the rectangle represents the state(or output) space where particles (or corresponding outputs) located; the ellipse in state space denotes the state constraint region. Each particle corresponds to one possible state trajectory. If a particle violates the constraint, such a particle will be brought within the constraint region to a most likely location based on Eqs. (20)–(23).

Generally, $p_{x_k^e}$ and $p_{\hat{v}_k}$ can be any distributions. However, for a tractable solution when dealing with constraints, truncated Gaussian, double half Gaussian or Gaussian mixture pdfs are often used to prescribe pdfs of noise and state particles during the implementation [32,27,19,20]. Hence, a quadratic form of objective function can be formed.

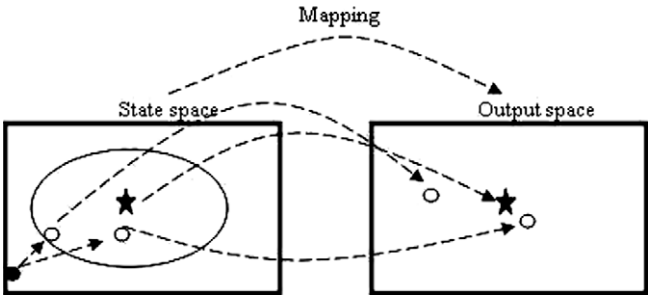


Fig. 4. Illustration of particle projection (○: valid particle, •: violated particle, ★: true state/measurement).

Table 1
Constrained state particles for the generic PF.

Constrained particles	Objective function
Prior particles: $x_k^{i,-}$	Eqs. (20) and (21)
Posterior particles: x_k^i	Eq. (22)
Estimated mean: $\hat{x}_k$	Eq. (23)

The sampling nature of PF has an advantage that covariance of estimated state error can be computed directly from samplers. For the prior particles, covariance can be estimated as:

$$\bar{x}_k^- = \sum_{i=1}^{N} w_k^i x_k^{i,-}, \qquad P_k^- = \frac{\sum_{i=1}^{N} w_k^i (x_k^{i,-} - \bar{x}_k^-)(x_k^{i,-} - \bar{x}_k^-)^T}{1 - \sum_{i=1}^{N} {w_k^i}^2}. \tag{24}$$

For posterior particles, all the weights are set uniform, then the sample covariance can be computed as

$$\bar{x}_k = \frac{1}{N}\sum_{i=1}^{N} x_k^i, \qquad P_k = \frac{1}{N-1}\sum_{i=1}^{N} (x_k^i - \bar{x}_k)(x_k^i - \bar{x}_k)^T. \tag{25}$$

Table 1 lists a summary of imposing constraints onto state particles for the generic PF algorithm.

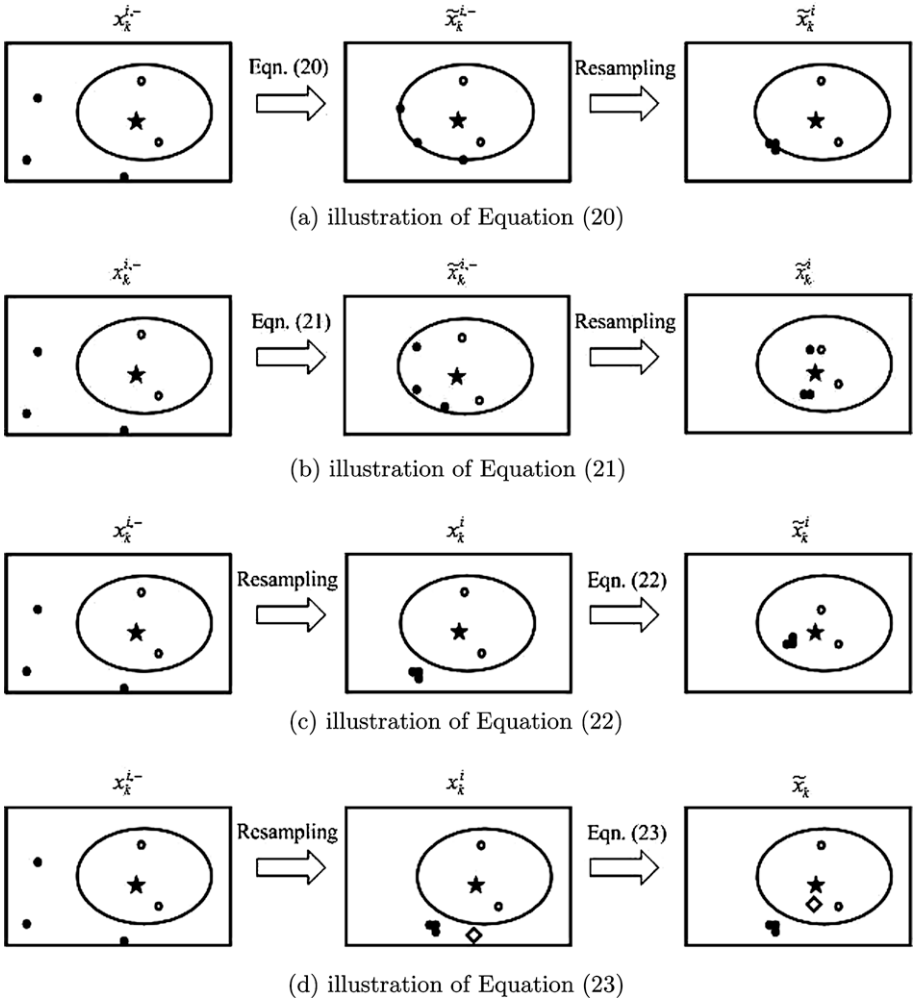
Constraints of other variables could also be imposed onto corresponding particles, such as estimated output, $\hat{y}_k^i$. The choice of which objective function to use, namely which step in Algorithm 1 to implement optimization, depends on the specific system and available computational resources.

5.2.3. Discussion

In the previous section, several variants of constrained PF algorithms have been presented based on optimization formulations. Illustrations of the differences of these formulations are shown in Fig. 5. In the figures, the rectangle represents the space where particles are generated; the ellipse denotes the state constraint region.

As the figure shows, some of the prior particles are outside the constraint region. By using Eq. (20), as shown in Fig. 5a, violated particles are projected onto the boundary, while particles that are already within the constraints remain unchanged. This equation is equivalent to "clipping", which requires low computational load but probably yields poor performance.

Fig. 5b shows that measurement information is used when imposing constraints onto prior particles. A trade-off between output error and state deviation is made to project the particles into a feasible region before resampling procedure. As in Fig. 5c, constraints are imposed onto particles after resampling procedure. Both Eqs. (21) and (22) reshape posterior distribution by projecting



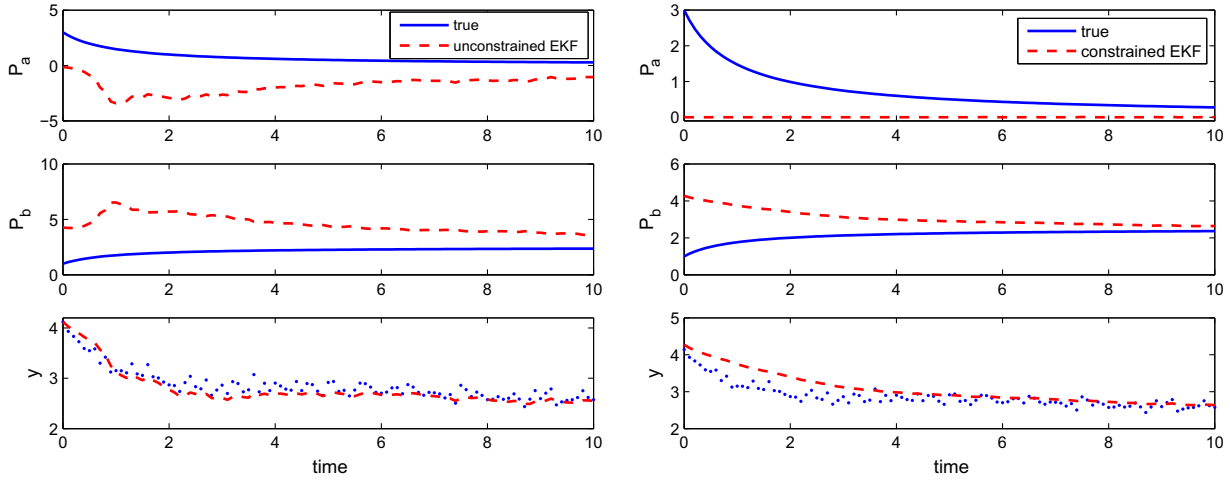
(a) illustration of Equation (20)

(b) illustration of Equation (21)

(c) illustration of Equation (22)

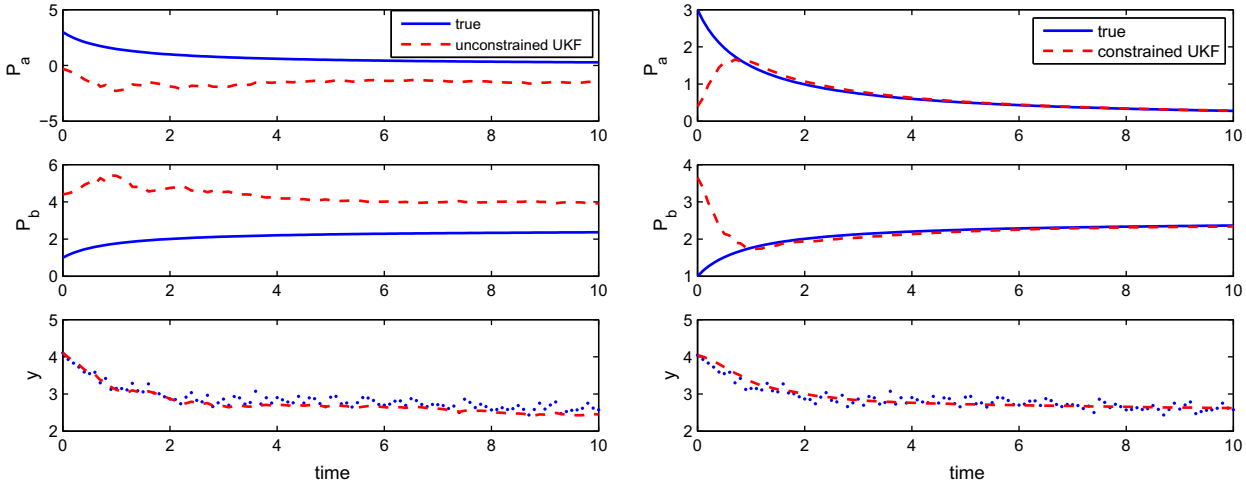
(d) illustration of Equation (23)

Fig. 5. Illustration example of differences among Eqs. (20)–(23) (○: valid particle, •: violated particle, ◇: estimated mean, ★: true state).



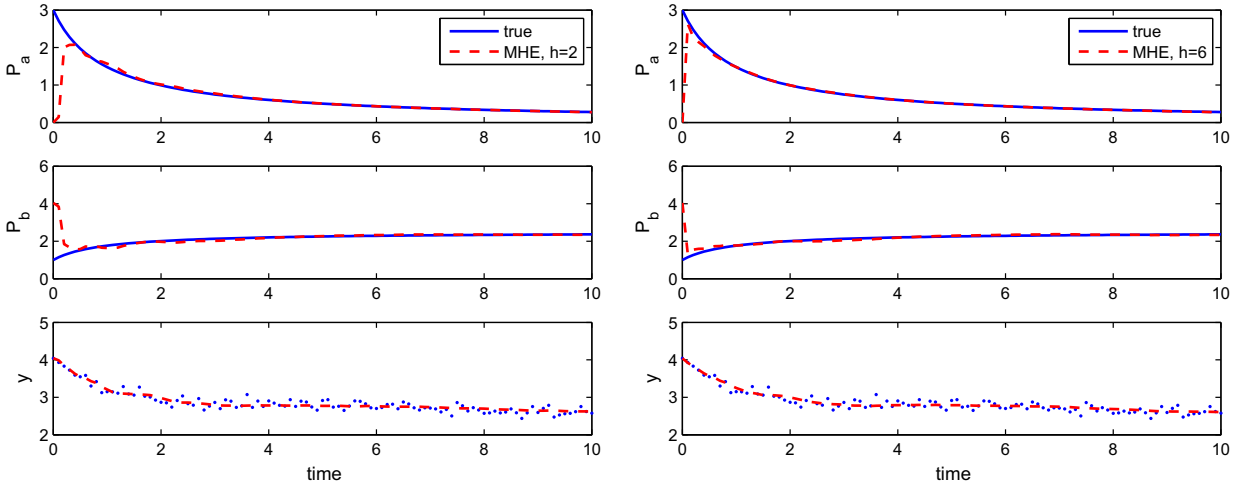
(a) unconstrained EKF (b) constrained EKF

Fig. 6. EKF estimates for example 1.



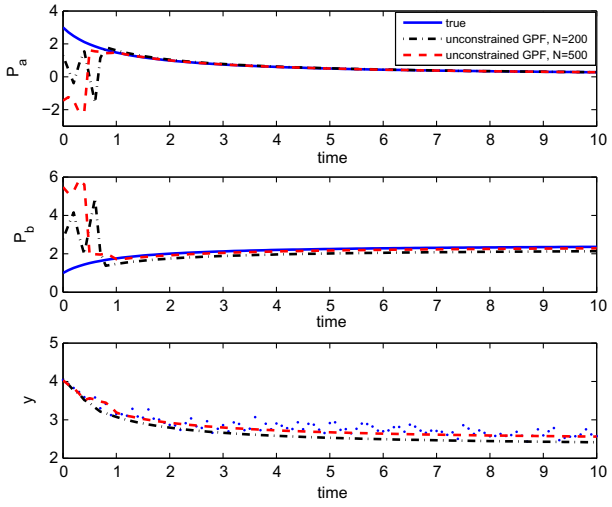
(a) unconstrained UKF (b) constrained UKF

Fig. 7. UKF estimates for example 1.

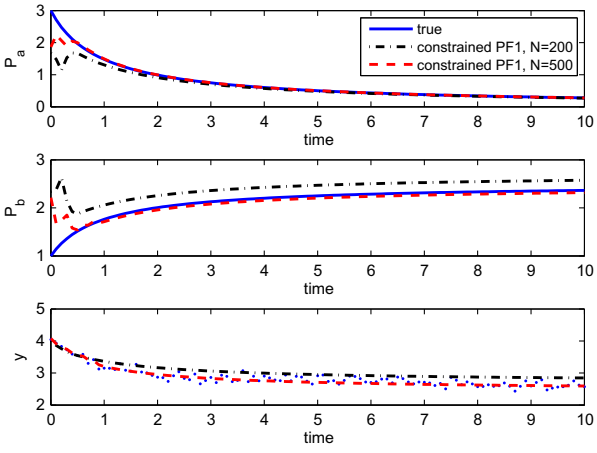


(a) MHE, $h = 2$ (b) MHE, $h = 6$

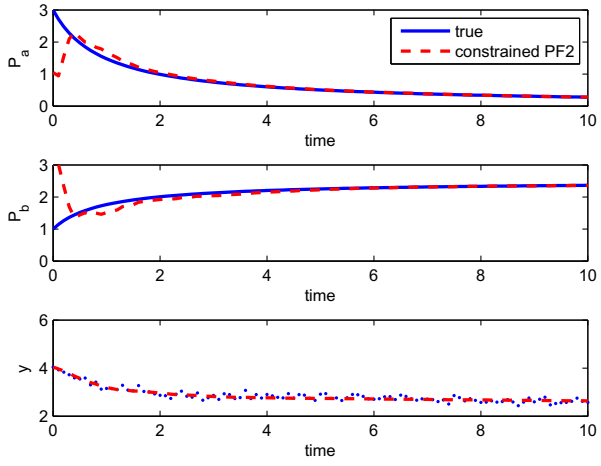
Fig. 8. MHE estimates for example 1.



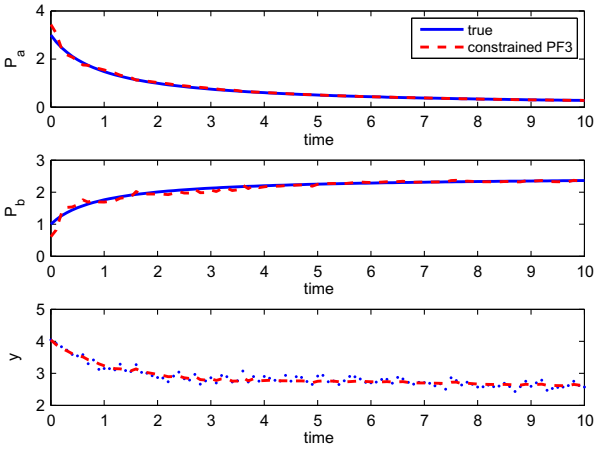
(a) unconstrained GPF



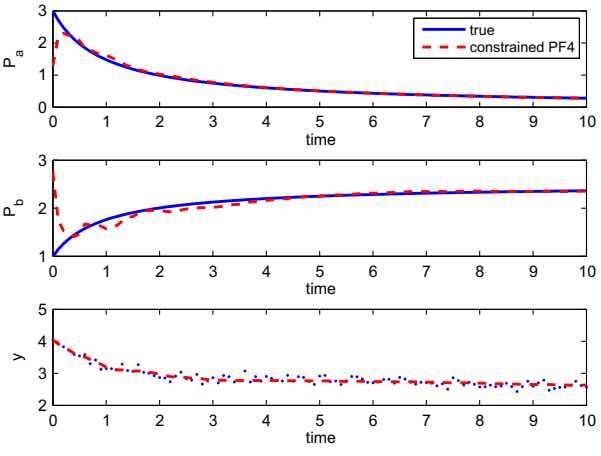
(b) constrained PF1



(c) constrained PF2



(d) constrained PF3



(e) constrained PF4

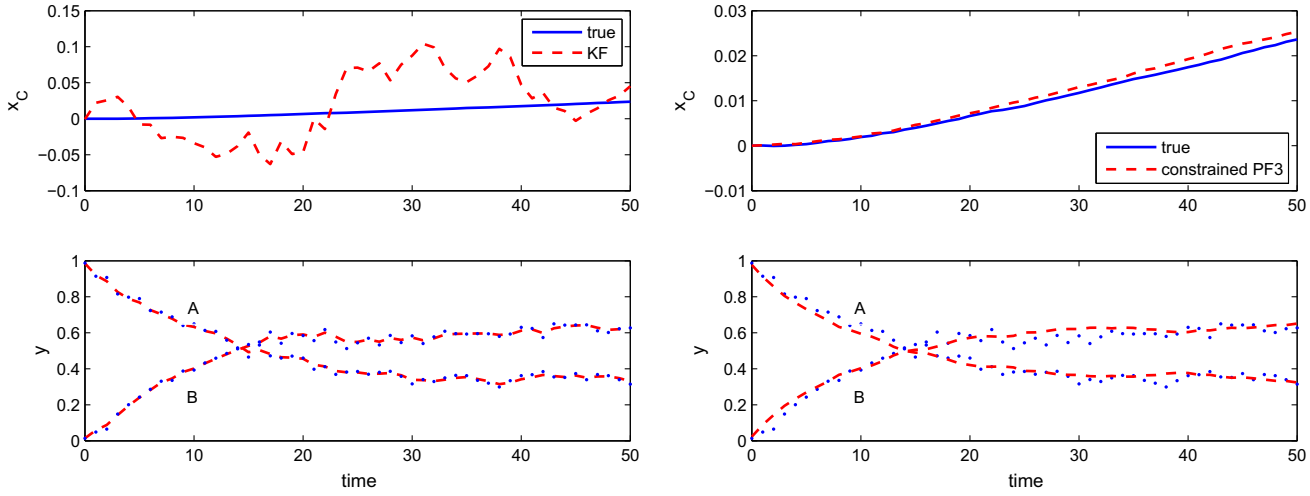
Fig. 9. PF estimates for example 1.

a set of particles, which could provide more accurate estimates when the state distribution is non-Gaussian. However, it requires much higher computational resource.

Eq. (23) imposes constraints onto the estimated mean, as shown in Fig. 5d, without considering the constraints on each particle. Compared to Eqs. (21) and (22), it is computationally less demanding, but it has limitations due to the consideration of mean value only. For instance, the mean could be located between the modes without violating any constraint, but with a very low posterior probability. Note that a fixed-length moving horizon can be added straightforward, leading to a combination of PF and MHE [26].

Table 2
Comparison of estimation performances for example 1.

Estimators		Schemes	MSE P_a	MSE P_b	CPU time (s)
EKF		N/A	6.8743	6.4232	1.547×10^{-4}
Constrained EKF		Clipping	0.8431	1.3061	1.547×10^{-4}
UKF ($\alpha = 1, \beta = 2, \kappa = 1$)		N/A	4.6786	4.3234	3.282×10^{-4}
Constrained UKF		Optimization	0.1532	0.1843	0.1213
MHE	$h = 2$	Optimization	0.1089	0.1186	0.1379
	$h = 6$	Optimization	0.0836	0.0949	0.8446
GPF	$N = 200$	N/A	0.3907	0.4614	0.0192
	$N = 500$	N/A	0.3614	0.3853	0.0493
Constrained PF1	$N = 200$	Accept/reject	0.0578	0.1496	0.0204
	$N = 500$	Accept/reject	0.0183	0.0242	0.0538
Constrained PF2 ($N = 50$)		Hybrid	0.0463	0.0565	0.0398
Constrained PF3 ($N = 50$)		Hybrid	0.0038	0.0055	0.0448
Constrained PF4	$N = 50, h = 2$	Hybrid	0.0147	0.0192	0.0297
	$N = 50, h = 6$	Hybrid	0.0043	0.0011	0.0929



(a) KF (b) constrained PF3

Fig. 10. Simulation results for example 2.

5.3. Constrained PF algorithm

The main concern for the optimization-based PFs is the online computation requirement. In order to reduce the computational cost and make the algorithm robust in the presence of poor prior information, a constrained PF algorithm based on hybrid use of acceptance/rejection and optimization strategies is proposed. The proposed scheme executes optimization only when the estimation performance based on the particles inside constraint region fails a performance test; otherwise, acceptance/rejection method (denoted as constrained PF1) is used.

In order to decide if the performance of the acceptance/rejection-based PF is satisfactory, a chi-square test is used. The rationale is that if the particles inside constraint region provide a good state estimate, the innovation term, $e_k = y_k - \hat{y}_k$, will have mean zero and covariance of Σ. In other words, the squared residual is checked if it follows a central chi-square distribution with p degree of freedom when the measurement error follows Gaussian distribution

$$e_k^T \Sigma^{-1} e_k \sim \chi^2(p), \tag{26}$$

where $e_k \sim N(0, \Sigma)$, and $p = dim(y)$ is the dimension of output.

Given past history data on estimation performance, a sliding time window l can be adopted in Eq. (26):

$$\sum_{j=k-l+1}^{k} e_j^T \Sigma^{-1} e_j \sim \chi^2(l \times p). \tag{27}$$

To reduce the computational cost, the optimization procedure is executed only when Eq. (26) (or Eq. (27)) fails the statistical testing with a given significance level, e.g. $\alpha = 5\%$.

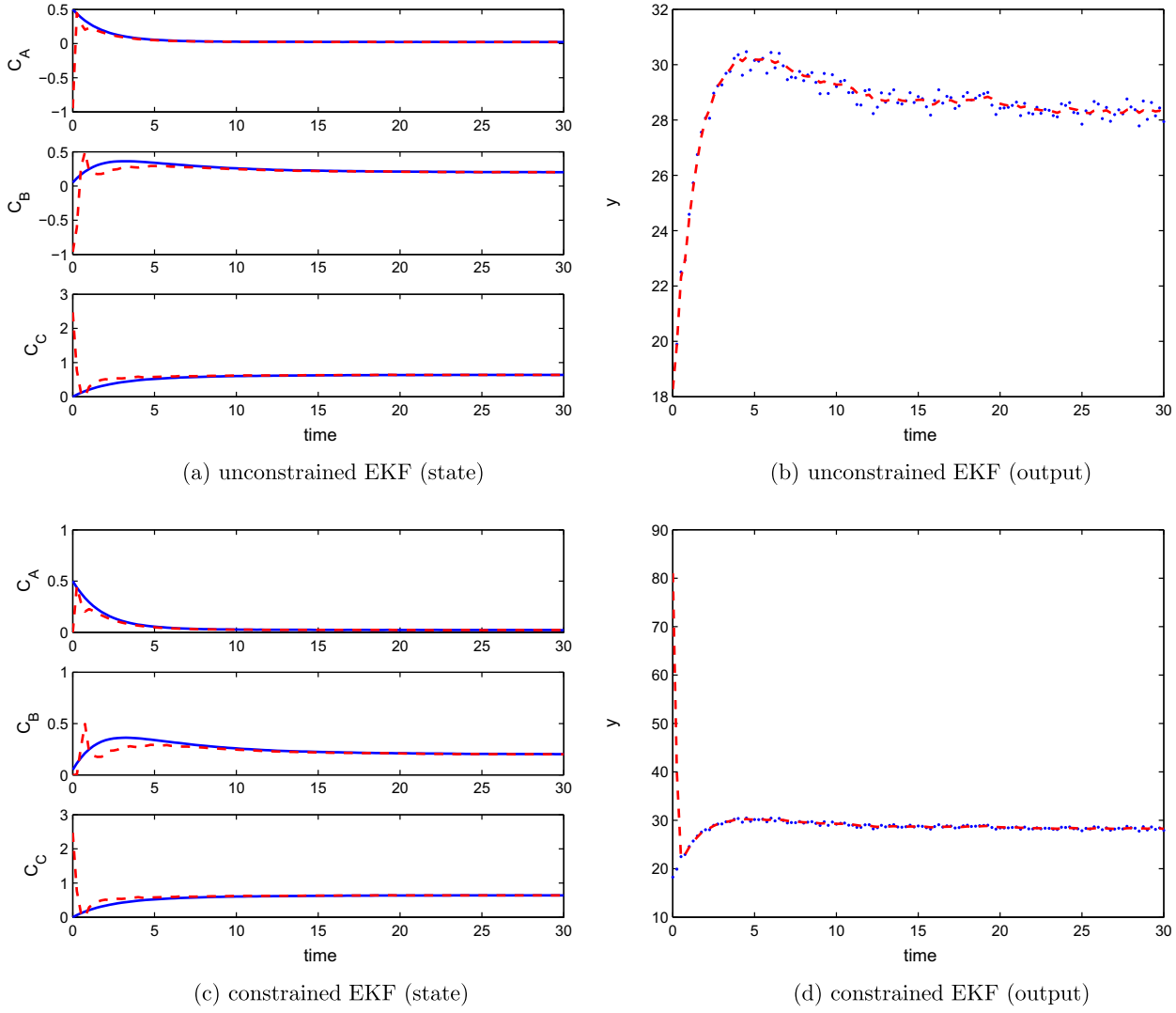
Note that, if the measurement error is assumed as non-Gaussian, the Chi-square test can be simply treated as a quadratic (2nd moment) test of the residual. In that case, a problem-specific threshold would be heuristically chosen instead of using Chi-square table.

The idea of the proposed constrained PF algorithm is summarized in the following, in which Eq. (21) is chosen as an objective function for the optimization procedure (the algorithm is denoted as constrained PF2 in the following section). Similar algorithms based on Eqs. (22) and (23) (denoted as constrained PF3 and constrained PF4, respectively) are provided in Appendix A.

Algorithm 2. A novel constrained PF algorithm based on Eq. (21)

Step (a) *Initialization*: Generate initial particles $\{x_0^i\}_{i=1}^N$ from *a priori* distribution $p(x_0)$, and set $k = 1$;

Step (b) *Importance sampling*: Generate prior particles, $\{x_k^{i,-}\}_{i=1}^N$, from importance sampling distribution $p(x_k^{i,-}|x_{k-1}^i)$;



(a) unconstrained EKF (state)

(b) unconstrained EKF (output)

(c) constrained EKF (state)

(d) constrained EKF (output)

Fig. 11. EKF estimates for example 3.

Step (c) *Weighting*: Calculate constrained likelihood and importance weights according to Eqs. (15) and (16), then normalize the weights as $w_k^i = \tilde{w}_k^i / \sum_{j=1}^{N} \tilde{w}_k^j$;

Step (d) *Chi-square test*: Calculate the weighted sample mean of the valid particles, $\bar{x}_k^- = \sum_{i=1}^{N_1} w_k^i x_k^{i,-}$, where N_1 is the particle number inside the constraint region; and compute the output residual, $e_k = y_k - h(\bar{x}_k^-)$; test the Chi-square criteria with a preset covariance Σ;

Step (e) *Optimization*: Project the violated particles into constraint region by solving Eq. (21) if performance test in step d fails; recalculate importance weights and normalization;

Step (f) *Resampling*: Generate posterior particles, $\{\tilde{x}_k^i\}_{i=1}^N$, based on resampling strategy, and set $w_k^i = 1/N$;

Step (g) *Output*: Estimate the state by calculating $\hat{x}_k = 1/N \cdot \sum_{i=1}^{N} \tilde{x}_k^i$, set $k = k+1$ and go back to Step (b).

6. Simulation studies

In order to investigate the efficacy of the proposed method, several examples with constraints on state are studied in this section. All the simulations were run on a 2.2 GHz CPU with 1 GB RAM PC using MATLAB 2008a. The mean square error (MSE) and CPU time presented below are based on 100 simulations.

6.1. Two-state batch reaction

6.1.1. Process description

Consider a gas-phase reaction well studied by Vachhani et al. [39], Rawlings and Bakshi [30], Ungarala et al. [37], Kandepu et al. [16] and Kolås [17]:

$$2A \xrightarrow{k} B, \quad k = 0.16,$$

with a stoichiometric matrix

$$v = [-2 \quad 1],$$

and a reaction rate

$$r = kP_a^2.$$

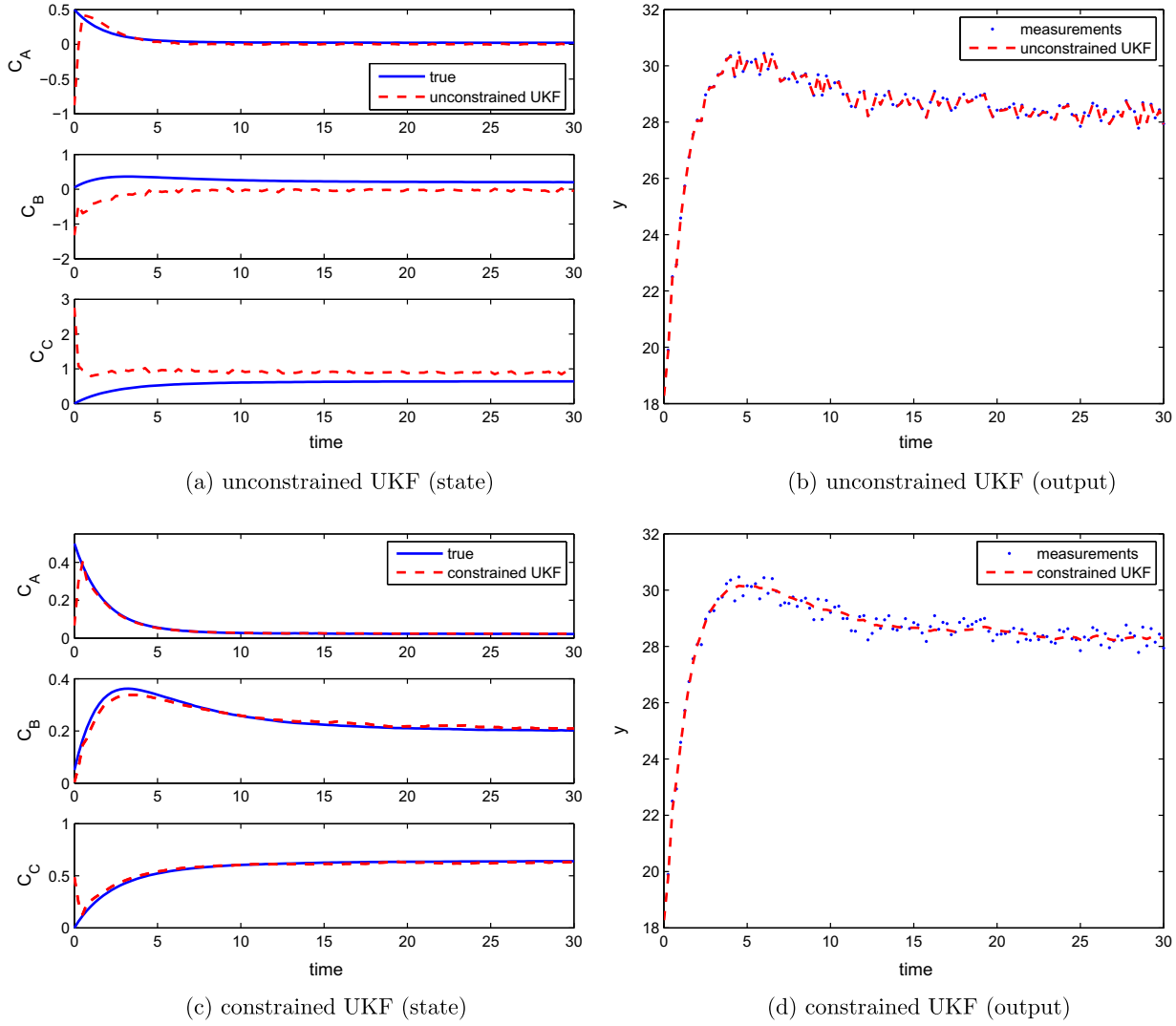
The state and measurement vectors are defined as

$$x = [P_a \quad P_b]^T, \quad y = [1, \quad 1]x,$$

where P_j denotes the *non-negative* partial pressure of species j. It is assumed that the ideal gas law holds and that the reaction occurs in a well-mixed isothermal batch reactor. Then, from first principles, the process model can be written as

$$\dot{x} = f(x) = v^T r. \tag{28}$$

The system is discretized with a sampling interval of $\Delta t = 0.1$ s, and simulated for 100 time steps from the initial condition $x_0 = [3, \quad 1]^T$,



(a) unconstrained UKF (state)

(b) unconstrained UKF (output)

(c) constrained UKF (state)

(d) constrained UKF (output)

Fig. 12. UKF estimates for example 3.

Table 3
Comparison of estimation performances for example 3.

Estimators		Schemes	MSE C_A	MSE C_B	MSE C_C	CPU time (s)
EKF		N/A	0.0176	0.0150	0.0568	2.27×10^{-4}
Constrained EKF		Clipping	0.0024	0.0023	0.0568	2.27×10^{-4}
UKF ($\alpha = 0.001$, $\beta = 2$, $\kappa = 0$)		N/A	0.0184	0.1322	0.1881	0.0012
Constrained UKF		Optimization	0.0017	2.17×10^{-4}	0.0025	0.5330
MHE	$h = 2$	Optimization	0.0037	9.80×10^{-4}	0.0039	0.3436
	$h = 6$	Optimization	0.0021	1.68×10^{-4}	0.0029	2.3896
GPF ($N = 500$)		N/A	Fail			
Constrained PF1 ($N = 500$)		Accept/reject	Fail			
Constrained PF2 ($N = 100$)		Hybrid	0.0014	3.32×10^{-4}	0.0017	0.0379
Constrained PF3 ($N = 100$)		Hybrid	0.0014	2.23×10^{-4}	0.0017	0.0330
Constrained PF4	$N = 100, h = 2$	hybrid	0.0023	2.71×10^{-4}	0.0026	0.0188
	$N = 100, h = 6$	Hybrid	1.82×10^{-4}	1.31×10^{-4}	7.6×10^{-4}	0.1241

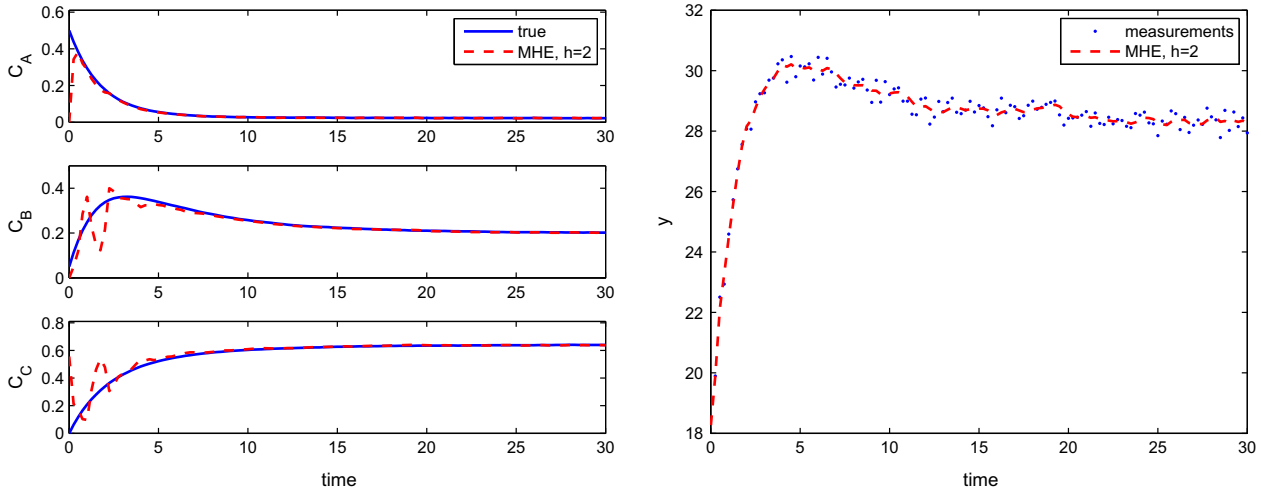
and corrupted by Gaussian noise given by $\omega \sim N\left\{[0,0]^T, 10^{-6} I_2\right\}$, and $v \sim N\{0, 10^{-2}\}$. Estimation starts from a poor initial guess $\bar{x}_0 = [0.1, \quad 4.5]^T$ with a large covariance matrix $P_0 = 6^2 I_2$.

This problem has been popularly studied in the literature because without considering constraints the state estimator can experience a multimodal pdf, which may lead to unphysical estimates.

6.1.2. Simulation results

The proposed constrained PF algorithms are tested on the reactor problem. For fair comparisons, $Q_{\hat{\omega}} = Q_{\omega}, R_{\hat{v}} = R_v$, and the same constraints and noise sequences are used for all the simulations in this example.

Figs. 6a and 7a show that due to the poor initial guess and the multimodal nature, neither unconstrained EKF nor UKF converges



(a) state (b) output

Fig. 13. MHE estimates for example 3, $h = 2$.

to true states within the given simulation time despite good estimates of the output.

Fig. 6b shows the estimate of the constrained EKF using clipping method; it does restrict the state to the constraint region, but the estimation result is still poor. This is because the constraint knowledge was not properly used in updating the covariance matrix.

A quadratic programming (QP) based UKF, proposed by Kolås et al. [18], is used to incorporate state constraints by constraining "sigma points". Fig. 7b shows the estimation performance is improved compared to the unconstrained case. However, a large increase of computation time is observed in solving the optimization problem.

Compared to EKF/UKF based approaches, MHE provides improved estimates in terms of accuracy, see Figs. 8a and 8b, but computation time increases with the horizon size.

Fig. 9a shows the estimation results of unconstrained generic PF with particle size $N = \{200, 500\}$. Compared to its counterparts of unconstrained EKF and unconstrained UKF, the Monte Carlo sampling based PF yields much more accurate estimate in this example; however, it still gives estimates violating physical constraints during initial time points. Compared to MHE, PF shows the advantage in computation time due to its single-horizon formulation.

Results of constrained PFs are shown in Fig. 9b–e, in which constrained PF1 denotes the constrained PF based on acceptance/rejection scheme [23]; constrained PF2 denotes the constrained PF using hybrid scheme with optimization on prior particles (i.e. Eq. (21)); constrained PF3 denotes the constrained PF using hybrid scheme with optimization on posterior particles (i.e. Eq. (22)); and constrained PF4 denotes the constrained PF using hybrid scheme with optimization on estimated mean (i.e. Eq. (23)). The figures show that all of these constrained methods provide physically valid estimates.

Table 2 shows the detailed performance comparisons. It can be seen that optimization-based methods generally yield better estimation, but with much higher computational cost. The table also shows that hybrid use of constraint handling strategies provide the best estimates. In this example, the optimization was only necessary in the first few time steps to compensate the poor initial guess; for most time, acceptance/rejection procedure was used.

It should be also noted that: (i) choice of a particular method should depend on the available computational resource and the accuracy requirement; (ii) both unconstrained GPF and constrained PF1 are sensitive to the poor prior information; therefore, a larger particle size should be chosen. Methods based on the hybrid scheme are more robust; thus a smaller particle size can be used to reduce computational cost; (iii) optimization in constrained PF2 and constrained PF3 is not necessary to be applied to the whole particle set; for instance, optimization in constrained PF2 can be only applied to the prior particles violating constraints; optimization in constrained PF3 can be only applied to the parent particles (i.e. the subset particles selected for resampling) and, (iv) constrained PF4 is actually a combination of constrained PF1 and MHE; thus its estimates will be no poorer than MHE of the same horizon size.

6.2. Three-state batch reaction

6.2.1. Process description

Consider a batch reactor system adapted from Ungarala et al. [38]

$$A \underset{k_1}{\overset{k_2}{\rightleftharpoons}} B \overset{k_3}{\rightarrow} C,$$

where $k = [k_1 \quad k_2 \quad k_3] = [0.06 \quad 0.03 \quad 0.001]$. The total number of moles remains constant in the reactor. A set of ODEs is used to describe the process dynamics,

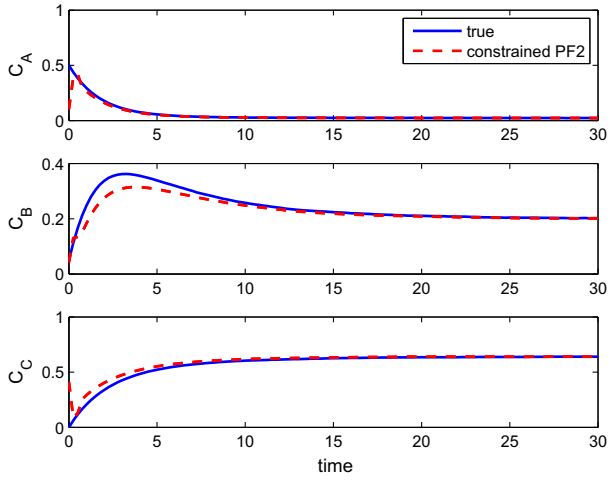
$$\frac{dx}{dt} = \begin{bmatrix} -k_1, & k_2 & 0 \\ k_1 & -k_2 - k_3 & 0 \\ 0 & k_3 & 0 \end{bmatrix} x, \tag{29}$$

where $x = [x_A \quad x_B \quad x_C]^T$ is the vector of model fractions, which must obey the constraints as:

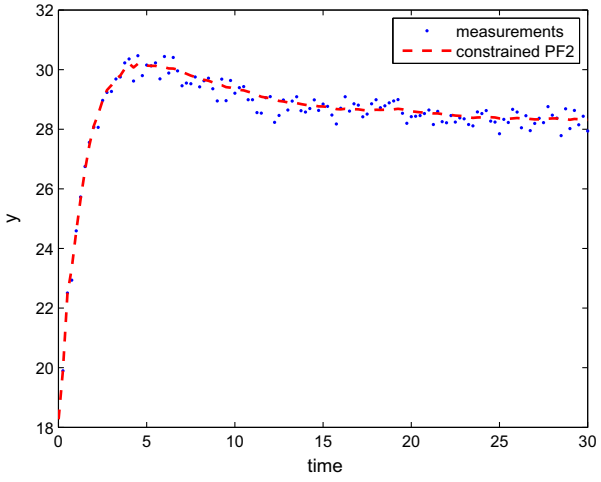
$$\begin{aligned} &0 \leqslant x_i \leqslant 1, \\ &\sum x_i = 1. \end{aligned} \tag{30}$$

The system is discretized with a sampling interval of $\Delta t = 1$, and simulated for 50 time steps from the initial condition $x_0 = [1 \quad 0 \quad 0]^T$. A discretized process function can be obtained as

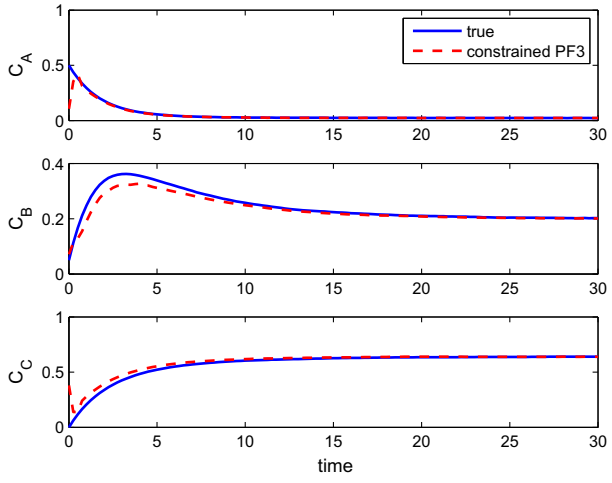
$$x_k = \begin{bmatrix} 1 - k_1, & k_2 & 0 \\ k_1 & 1 - k_2 - k_3 & 0 \\ 0 & k_3 & 1 \end{bmatrix} x_{k-1} + \omega_{k-1}, \tag{31}$$



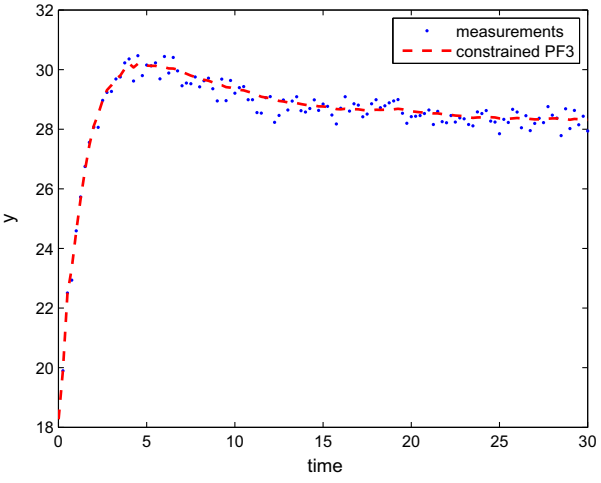
(a) constrained PF2 (state)



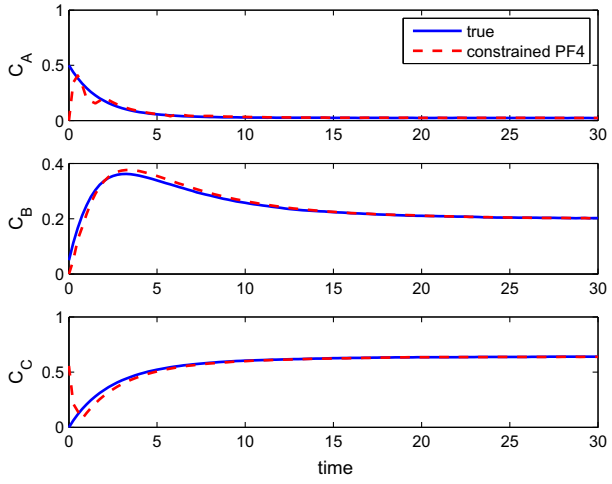
(b) constrained PF2 (output)



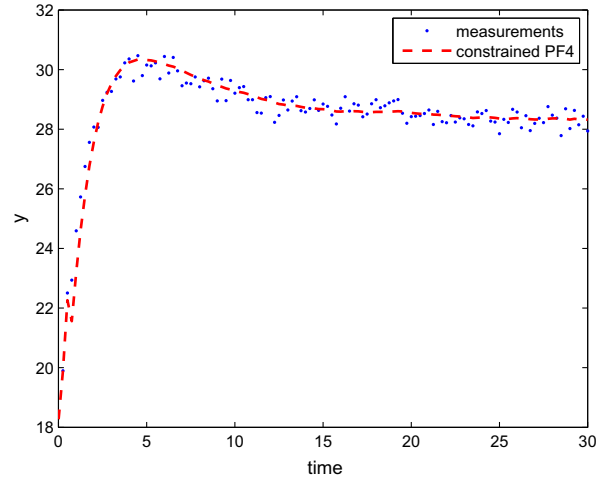
(c) constrained PF3 (state)



(d) constrained PF3 (output)



(e) constrained PF4 (state)



(f) constrained PF4 (output)

Fig. 14. Optimization-based constrained PF estimates for example 3.

where ω is zero mean Gaussian process noise with covariance matrix $Q_\omega = diag([\,0.01^2 \quad 0.01^2 \quad 0.0001^2\,])$. Noisy measurements of mole fractions are only available for species A and B:

$$y_k = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} x_k + v_k, \tag{32}$$

where $v \sim N(0, 0.02^2 I_2)$. The objective is to filter the measurements and estimate the unmeasured state x_C. Estimation starts from a poor initial guess $\bar{x}_0 = [0.8 \quad 0.1 \quad 0.1]^T$ with a covariance matrix $P_0 = diag([1^2 \quad 1^2 \quad 0.01^2])$.

6.2.2. Simulation results

Although it is a linear problem, this system is interesting to study because without considering the equality constraint, the system is not observable since the measurement matrix does not have full column rank. One may estimate x_A and x_B by using Kalman filter with a two-dimension model and then compute x_C from the equality constraint condition. However, the result obtained in Fig. 10a shows that the KF estimate of x_C can easily violate the non-negativity constraint.

Constrained PF based on acceptance/rejection scheme fails due to the poor prior information and the stringent constraint region. The optimization techniques allow for incorporating both equality and inequality constraints in Eq. (30) into the estimation formulation. Constrained PF3 with particle size $N = 100$ is chosen as an estimator for this example. Fig. 10b shows that the estimate accuracy has been significantly improved. The computation time is also reasonable, average CPU time is 0.04 s for each time step in this example.

6.3. Three-state continuous stirred-tank reaction (CSTR)

6.3.1. Process description

Consider a three-state CSTR gas-phase reaction taken from Haseltine and Rawlings [11], Teixeira et al. [36], Kolås et al. [18]

$$A \underset{k_2}{\overset{k_1}{\rightleftharpoons}} B + C,$$

$$2B \underset{k_4}{\overset{k_3}{\rightleftharpoons}} C,$$

$$k = [k_1 \quad k_2 \quad k_3 \quad k_4] = [0.5 \quad 0.05 \quad 0.2 \quad 0.01],$$

with a stoichiometric matrix

$$v = \begin{bmatrix} -1 & 1 & 1 \\ 0 & -2 & 1 \end{bmatrix},$$

and a reactional rate

$$r = \begin{bmatrix} k_1 C_A - k_2 C_B C_C \\ k_3 C_B^2 - k_4 C_c \end{bmatrix}.$$

The states and measurements are defined as to be

$$x = [C_A \quad C_B \quad C_C]^T,$$
$$y = [RT \quad RT \quad RT]x,$$

where C_j denotes the *non-negative* concentration of species j, R is the ideal gas constant, T is the reactor temperature, and $RT = 32.84$. It is assumed that the ideal gas law holds. From first principles, the process model for a well-mixed, isothermal CSTR reactor is

$$\dot{x} = f(x) = \frac{Q_f}{V_R} C_f - \frac{Q_0}{V_R} x + v^T r, \tag{33}$$

where $Q_f = Q_0 = 1, V_R = 100$ and $C_f = [0.5 \quad 0.05 \quad 0]$.

The system is discretized with a sampling interval of $\Delta t = 0.25$, and simulated for 120 time steps from the initial condition $x_0 = [0.5 \quad 0.05 \quad 0]$, and corrupted by Gaussian noise given by $\omega \sim N([0 \quad 0 \quad 0]^T, 10^{-6} I_3)$ and $v \sim N(0, 0.25^2)$. Estimation starts from a poor initial guess $\bar{x}_0 = [0 \quad 0 \quad 3.5]^T$ with a covariance matrix $P_0 = 4^2 I_3$.

6.3.2. Simulation results

With $Q_{\bar{\omega}} = Q_\omega, R_{\bar{v}} = R_v$, and the same constraints and noise sequences, different nonlinear estimators are compared. Fig. 11 shows EKF estimates, where neither unconstrained nor constrained EKF provides satisfactory results. Fig. 12 shows the UKF estimates, where constrained UKF gives good results but with a huge increase of computation time; see Table 3 for detailed comparisons. MHE with a horizon size of $h = 2$ also provides good estimates, see Fig. 13. By increasing the horizon size, MHE estimates can become better, with a further increase of computational cost (see Table 3). Constrained PF1 with $N = 500$ fails in this example, due to the poor initial guess. Fig. 14 shows the estimation results of the proposed constrained PF methods, which provides the best results for this example in terms of computation time and estimation accuracy.

From a large number of simulation runs, it was observed that constrained PF based on acceptance/rejection scheme requires the least amount of computation time, but it easily failed with poor prior information or stringent constraints. Under the same conditions, optimization-based constrained PFs yielded better estimation and showed more robustness; however, they require much higher computation time, which may not be suitable for on-line applications. Hybrid use of the acceptance/rejection and optimization schemes can combine the complementary advantages, and work more efficient in most situations.

7. Conclusion

Proper use of constraint knowledge is critical for the successful implementation of Bayesian estimators, since it can confine distribution domains of related variables, and make the estimation more accurate. In this paper, two different constraints handling strategies are discussed under the generic PF framework. Several new constrained PF algorithms are implemented based on hybrid use of acceptance/rejection and optimization schemes. Simulation results show that the proposed methods work efficiently for the investigated examples as they combine the advantages of Monte Carlo sampling nature of PF and the benefits of optimization techniques in handling constraints and poor prior information.

It is recommended that different methods should be considered depending on the available computational resource and the accuracy requirements. When one has good initialization knowledge with simple constraints, constraint PF1 [23] should be chosen; when one needs to handle complicated constraints with very limited computational resource, constraint PF4 with single-horizon window should be considered; if computational cost is not the concern and the state distribution is believed as non-Gaussian, then constraint PF2 and PF3 may be selected.

The main contributions of this paper are: (i) different and more efficient ways of incorporating state constraints in PF framework have been discussed and implemented; (ii) variant constrained Bayesian estimators are comparatively studied through several simulation examples. The proposed constrained PFs provide some interesting flexibility for constrained nonlinear/non-Gaussian Bayesian state estimation problems.

Acknowledgements

This work was supported in part by the Natural Sciences and Engineering Research Council of Canada.

Appendix A. Constrained PFs based on Eqs. (22) and (23)

Since Eq. (20) is same with clipping, hereby only the constrained PFs based on Eqs. (22) and (23) are summarized as follows:

Algorithm 3. A novel constrained PF algorithm based on Eq. (22)

Step (a) *Initialization*: Generate initial particles $\{x_0^i\}_{i=1}^N$ from *a priori* distribution $p(x_0)$, and set $k = 1$;
Step (b) *Importance sampling*: Generate prior particles, $\{x_k^{i-}\}_{i=1}^N$, from importance sampling distribution $p(x_k^{i-}|x_{k-1}^i)$;
Step (c) *Weighting*: Calculate constrained likelihood and importance weights according to Eqs. (15) and (16), then normalize the weights as $w_k^i = \tilde{w}_k^i / \sum_{j=1}^N \tilde{w}_k^j$;
Step (d) *Resampling*: Generate posterior particles, $\{x_k^i\}_{i=1}^N$, based on weighting information and resampling strategy;
Step (e) *Chi-square test*: Calculate the sample mean of the posterior particles, $\bar{x}_k = \frac{1}{N}\sum_{i=1}^N x_k^i$, and compute the output residual, $e_k = y_k - h(\bar{x}_k)$; test the Chi-square criteria with a preset Σ;
Step (f) *Optimization*: Project the parent particles (i.e. the subset particles selected for resampling) to new locations by solving Eq. (22) if performance test in step e fails; recalculate the weights and resampling;
Step (g) *Output*: Estimate the state by calculating $\hat{x}_k = 1/N \cdot \sum_{i=1}^N \tilde{x}_k^i$, set $k = k + 1$ and go back to Step (b).

Algorithm 4. A novel constrained PF algorithm based on Eq. (23)

Step (a) Initialization: Generate initial particles $\{x_0^i\}_{i=1}^N$ from *a priori* distribution $p(x_0)$, and set $k = 1$;
Step (b) Importance sampling: Generate prior particles, $\{x_k^{i-}\}_{i=1}^N$, from importance sampling distribution $p(x_k^{i-}|x_{k-1}^i)$;
Step (c) Weighting: Calculate constrained likelihood and importance weights according to Eqs. (15) and (16), then normalize the weights as $w_k^i = \tilde{w}_k^i / \sum_{j=1}^N \tilde{w}_k^j$;
Step (d) Resampling: Generate posterior particles, $\{x_k^i\}_{i=1}^N$, based on weighting information and resampling strategy;
Step (e) Chi-square test: Calculate the sample mean of the posterior particles, $\hat{x}_k = \frac{1}{N}\sum_{i=1}^N x_k^i$, and compute the output residual, $e_k = y_k - h(\hat{x}_k)$; test the Chi-square criteria with a preset Σ;
Step (f) Optimization: Calculate the projected mean, $\tilde{x}_k$, by solving Eq. (23) if performance test in step e fails;
Step (g) Output: Yield the projected mean as PF output; calculate state covariance, $\widetilde{P}_k$, by using EKF method, and regenerate particles from a normal distribution $N(\tilde{x}_k, \widetilde{P}_k)$; set $k = k + 1$ and go back to step b.

References

[1] F. Bell, B. Cathey, The iterated Kalman filter update as a Gauss–Newton method, IEEE Transactions on Automatic Control 38 (2) (1993) 294–297.
[2] J. Bellatoni, K. Dodge, A square root formulation of the Kalman schmidt filter, AIAA Journal 5 (1967) 1309–1314.
[3] M. Boloc, P. Djuric, S. Hong, Resampling algorithms for particle filters: a computational complexity perspective, EURASIP Journal of Applied Signal Processing 15 (2004) 2267–2277.
[4] W. Chen, B. Bakshi, P. Goel, S. Ungarala, Bayesian estimation via sequential Monte Carlo sampling: unconstrained nonlinear dynamic systems, Industrial & Engineering Chemistry Research 43 (2004) 4012–4025.
[5] C. Cheng, R. Ansari, Kernel particle filter: iterative sampling for efficient visual tracking, in: Proceedings of International Conference on Image Processing, 2003.
[6] F. Daum, Nonlinear filters: beyond the Kalman filter, IEEE A&E Systems Magazine 20 (8) (2005) 57–69.
[7] A. Doucet, S. Godsill, On Sequential Simulation-based Methods for Bayesian Filtering, Technical Report, Department of Engineering, Cambridge University, 1998.
[8] N. Gordon, D. Salmond, A. Smith, Novel approach to nonlinear/non-Gaussian Bayesian state estimation, in: IEE Proceedings F Radar and Signal Processing, 1993.
[9] N. Gupta, R. Hauser, Kalman Filtering with Equality and Inequality State Constraints, Technical Report, Oxford University Computing Laboratory, 2007.
[10] J.E. Handschin, D.Q. Mayne, Monte Carlo techniques to estimate the conditional expectation in multistage nonlinear filtering, International Journal of Control 9 (5) (1969) 547–559.
[11] E.L. Haseltine, J.B. Rawlings, Critical evaluation of extended Kalman filtering and moving-horizon estimation, Industrial & Engineering Chemistry Research 44 (2005) 2451–2460.
[12] Y.C. Ho, R.C.K. Lee, A Bayesian approach to problems in stochastic estimation and control, IEEE Transactions on Automatic Control 9 (5) (1964) 333–339.
[13] A. Jazwinski, Stochastic Processes and Filtering Theory, Academic Press, New York, 1970.
[14] S. Julier, J. LaViola, On Kalman filtering with nonlinear equality constraints, IEEE Transactions on Signal Processing 55 (6) (2007) 2774–2784.
[15] S. Julier, J. Uhlmann, Unscented filtering and nonlinear estimation, Proceedings of the IEEE 92 (2004) 401–422.
[16] R. Kandepu, B. Foss, L. Imsland, Applying the unscented Kalman filter for nonlinear state estimation, Journal of Process Control 18 (2008) 753–768.
[17] S. Kolås, Estimation in Nonlinear Constrained Systems with Severe Disturbances, Ph.D. Thesis, Norwegian University of Science and Technology, 2008.
[18] S. Kolås, B. Foss, T. Schei, Constrained nonlinear state estimation based on the UKF approach, Computers & Chemical Engineering 33 (8) (2009) 1386–1401.
[19] J.H. Kotecha, P.M. Djuric, Gaussian particle filtering, IEEE Transactions on Signal Processing 51 (10) (2003) 2592–2601.
[20] J.H. Kotecha, P.M. Djuric, Gaussian sum particle filtering, IEEE Transactions on Signal Processing 51 (10) (2003) 2603–2613.
[21] S. Kotz, T. Kozubowski, K. Podgorski, The Laplace Distribution and Generalizations: A Revisit with Applications to Communications, Economics, Engineering, and Finance, Birkhäuser, Boston, 2001.
[22] I. Kyriakides, D. Morrell, A. Papandreou-Suppappola, A particle filtering approach to constrained motion estimation in tracking multiple targets, in: The 39th Asilomar Conference on Signals, Systems and Computer, 2005.
[23] L. Lang, W. Chen, B. Bakshi, P. Goel, S. Ungarala, Bayesian estimation via sequential Monte Carlo sampling-constrained dynamic systems, Automatica 43 (2007) 1615–1622.
[24] J. Liu, R. Chen, Sequential Monte Carlo methods for dynamical systems, Journal of the American Statistical Association 93 (1998) 1032–1044.
[25] M. Pitt, N. Shephard, Filtering via simulation: auxiliary particle filters, Journal of the American Statistical Association 94 (446) (1999) 590–599.
[26] M. Rajamani, J. Rawlings, Improved state estimation using a combination of moving horizon estimator and particle filters, in: Proceedings of the 2007 American Control Conference, 2007.
[27] C. Rao, Moving Horizon Strategies for the Constrained Monitoring and Control of Nonlinear Discrete-time Systems, Ph.D. Thesis, University of Wisconsin-Madison, 2002.
[28] C. Rao, J. Rawlings, Constrained process monitoring: moving-horizon approach, AIChE Journal 48 (2002) 97–109.
[29] C.V. Rao, J.B. Rawlings, J.H. Lee, Constrained linear state estimation moving horizon approach, Automatica 37 (10) (2001) 1619–1628.
[30] J.B. Rawlings, B.R. Bakshi, Particle filtering and moving horizon estimation, Computers and Chemical Engineering 30 (2006) 1529–1541.
[31] D.G. Robertson, J.H. Lee, J.B. Rawlings, A moving horizon-based approach for least-squares state estimation, AIChE Journal 42 (8) (1996) 2209–2224.
[32] D.G. Robertsonb, J.H. Lee, On the use of constraints in least squares estimation and control, Automatica 38 (7) (2002) 1113–1123.
[33] D. Simon, Optimal State Estimation: Kalman, H_∞ and Nonlinear Approaches, Wiley-Interscience, 2006.
[34] B.O.S. Teixeira, J. Chandrasekar, H. Palanthandalam, L.A.B. Trres, L.A. Aguirre, D.S. Bernstein, Gain-constrained Kalman filtering for linear and nonlinear systems, IEEE Transactions on Signal Processing 56 (9) (2008) 4113–4123.
[35] B.O.S. Teixeira, J. Chandrasekar, L.A.B. Trres, L.A. Aguirre, D.S. Bernstein, State estimation for linear and nonlinear equality-constrained systems, International Journal of Control 82 (5) (2009) 918–936.
[36] B.O.S. Teixeira, L.A.B. Trres, L.A. Aguirre, D.S. Bernstein, Unscented filtering for interval-constrained nonlinear systems, in: Proceedings of the 47th IEEE Conference on Decision and Control, 2008b.
[37] S. Ungarala, E. Dolence, K. Li, Constrained extended Kalman filter for nonlinear state estimation, in: Proceedings of 8th International IFAC Symposium on Dynamics and Control Process Systems, 2007.
[38] S. Ungarala, K. Li, Z. Chen, Constrained Bayesian state estimation using a cell filter, Industrial & Engineering Chemistry Research 47 (2008) 7312–7322.
[39] P. Vachhani, S. Narasimhan, R. Rengaswamy, Robust and reliable estimation via unscented recursive nonlinear dynamic data reconciliation, Journal of Process Control 16 (2006) 1075–1086.
[40] P. Vachhani, R. Rengaswamy, V. Gangwal, S. Narasimhan, Recursive estimation in constrained nonlinear dynamical systems, AIChE Journal 51 (2004) 946–959.
[41] R. vander Merwe, Sigma-point Kalman Filters for Probability Inference in Dynamic State-space Models, Ph.D. Thesis, Oregon Health and Science University, 2004.
[42] R. vander Merwe, A. Doucet, N. deFreitas, E. Wan, The Unscented Particle Filter, Technical Report, Cambridge University Engineering Department, 2000.
[43] V. Zavala, C. Larid, L. Biegler, A fast moving horizon estimation algorithm based on nonlinear programming sensitivity, Journal of Process Control 18 (9) (2008) 876–884.