Probability-Guaranteed Set-Membership State Estimation for Polynomially Uncertain Linear Time-Invariant Systems

Yiming Wan[1], Vicenç Puig[2], Carlos Ocampo-Martinez[2], Ye Wang[2], and Richard D. Braatz[3]

***Abstract*— Conventional deterministic set-membership (SM) estimation is limited to unknown-but-bounded uncertainties. In order to exploit distributional information of probabilistic uncertainties, a probability-guaranteed SM state estimation approach is proposed for uncertain linear time-invariant systems. This approach takes into account polynomial dependence on probabilistic uncertain parameters as well as additive stochastic noises. The purpose is to compute, at each time instant, a bounded set that contains the actual state with a guaranteed probability. The proposed approach relies on the extended form of an observer representation over a sliding window. For the offline observer synthesis, a polynomial-chaos-based method is proposed to minimize the averaged $\mathcal{H}_2$ estimation performance with respect to probabilistic uncertain parameters. It explicitly accounts for the polynomial uncertainty structure, whilst most literature relies on conservative affine or polytopic overbounding. Online state estimation restructures the extended observer form, and constructs a Gaussian mixture model to approximate the state distribution. This enables computationally efficient ellipsoidal calculus to derive SM estimates with a predefined confidence level. The proposed approach preserves time invariance of the uncertain parameters and fully exploits the polynomial uncertainty structure, to achieve tighter SM bounds. This improvement is illustrated by a numerical example with a comparison to a deterministic zonotopic method.**

I. Introduction

Any practically useful state estimation method needs to deal with imprecise system models and inaccurate sensor measurements. Set-membership (SM) state estimation aims at computing a compact set of states that are consistent with the available measurements, the system model, and the unknown-but-bounded uncertainties. Convex sets such as intervals, ellipsoids, and zonotopes have been exploited to represent SM uncertainties and estimates [9], [13]. This conventional deterministic SM approach can be highly conservative, because it takes into account all possible uncertainty realizations even though the worst-case scenario occurs with vanishingly low probability.

Compared to unknown-but-bounded uncertainties, the probabilistic uncertainty description characterizes not only the support but also the likelihood of different uncertainty realizations. This observation motivates combining SM and probabilistic approaches to compute a probability-guaranteed SM state estimate, i.e., a compact set of states that contains the actual state with a guaranteed probability [3], [4], [17]. Such combined approaches in literature assume arbitrarily fast time-varying (TV) parametric uncertainties with an affine or polytopic structure, which could introduce conservative uncertainty overbounding as physical parameters may not vary arbitrarily.

In constrast, this paper proposes a probability-guaranteed SM state estimation approach for uncertain linear *time-invariant* (TI) systems. This approach takes into account polynomial dependence on probabilistic TI uncertain parameters as well as additive stochastic noises. The SM state estimates are constructed by exploiting the extended form of an observer representation over a sliding window. In the offline design phase, a polynomial-chaos-based observer synthesis method is used to minimize the averaged $\mathcal{H}_2$ performance with respect to the probabilistic parametric uncertainties. It explicitly accounts for the polynomial uncertainty structure, whilst most literature relies on conservative affine or polytopic overbounding. Online state estimation relies on restructuring the extended observer form into a TV linear affine transformation of an uncertainty vector. This enables efficient computation in approximating state distributions with Gaussian mixtures (GMs), and deriving probability-guaranteed SM estimates with ellipsoidal calculus.

This paper is organized as follows. The problem statement is presented in Section II. Section III describes the closed-loop representation and its extended form. Section IV proposes the offline observer synthesis, and Section V presents the online SM estimation. A simulation example and some conclusions are reported in Sections VI and VII, respectively.

II. Problem Statement

Consider the uncertain discrete-time linear TI system

$$x_{t+1} = A(\theta)x_t + w_t, \tag{1a}$$

$$y_t = C(\theta)x_t + v_t, \tag{1b}$$

where $t \in \mathbb{Z}_+$, $x \in \mathbb{R}^{n_x}$ is the state, $y \in \mathbb{R}^{n_y}$ is the measured output, $\theta \in \mathbb{R}^{n_\theta}$ is the vector of uncertain TI parameters, and w and v are the stochastic process and measurement noises, respectively.

Assumption 1: The system matrices $A(\theta)$ and $C(\theta)$ in (1) are TI, with known polynomial dependence on θ.

Assumption 2: θ has a bounded support Θ. Its elements are mutually independent random variables with known probability density functions (PDFs).

Assumption 3: w_t and v_t are zero-mean white noises with known PDFs; and θ, w, and v are mutually independent.

Assumptions 1 and 2 are not restrictive, because (i) non-polynomial nonlinear dependence on θ can be accurately

[1]Huazhong University of Science and Technology, 1037 Luoyu Street, Wuhan, 430074, China. E-mail: ywan@hust.edu.cn
[2]Universitat Politècnica de Catalunya, Institut de Robòtica i Informàtica Industrial (CSIC, UPC), C/. Llorens i Artigas 4-6, 08028 Barcelona, Spain. E-mail: {vpuig, cocampo, ywang}@iri.upc.edu
[3]Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139, USA. E-mail: braatz@mit.edu

approximated by polynomials or piecewise polynomials [15]; and (ii) the correlated entries of θ can be handled by transforming θ into a new random vector θ' whose elements are mutually independent. The PDF of θ can be obtained by either offline identification from data [10], or the *a priori* knowledge that specifies the relative importance of different points in the uncertainty region Θ.

Remark 1: The TI random parameter θ is inherently different from TV stochastic parameters that are statistically independent with respect to time. For (1), given the current state x_t, the future state x_{t+i} is correlated with the past state x_{t-j} since both depend on θ. Hence, the Markov property does not hold for uncertainty propagation in (1), whilst it indeed holds from systems with TV stochastic parameters as in [11], [19] and the references therein.

A probability-guaranteed SM state estimation problem is addressed in this article. Given the system model (1), the observations, and the probabilistic distributional information about θ, w_t, and v_t, the proposed approach aims at constructing, at each time instant t, a set $\mathcal{X}_t$ that contains the actual state with a guaranteed confidence level γ ($0 < \gamma < 1$), i.e.,

$$\Pr\{x_t \in \mathcal{X}_t\} \geq \gamma, \tag{2}$$

with Pr denoting the probability.

III. Closed-loop System Representation

This section constructs a closed-loop representation of the system (1).

By inserting (1b) into (1a), the closed-loop system representation is obtained as

$$\begin{aligned} x_{t+1} &= A(\theta)x_t + w_t + L(y_t - C(\theta)x_t - v_t) \\ &= A_{\text{cl}}(\theta)x_t + Ly_t + w_t - Lv_t, \end{aligned} \tag{3}$$

where L denotes the observer gain to be determined, and

$$A_{\text{cl}}(\theta) = A(\theta) - LC(\theta). \tag{4}$$

Such a closed-loop representation is actually a superposition of the observer

$$\hat{x}_{t+1} = A_{\text{cl}}(\theta)\hat{x}_t + Ly_t \tag{5}$$

and its error dynamics

$$\tilde{x}_{t+1} = A_{\text{cl}}(\theta)\tilde{x}_t + w_t - Lv_t, \tag{6}$$

where $\hat{x}_t$ and $\tilde{x}_t = x_t - \hat{x}_t$ represent the state estimate and its estimation error, respectively. The observer (5) cannot be directly used to compute the estimate $\hat{x}_t$ since it depends on the unknown parameter θ. Instead, the closed-loop representation (3) will be exploited to derive the SM state estimator in the following, and the error dynamics (6) will be used to design the observer gain L in Section IV.

Assume that the observer gain L stabilizes $A_{\text{cl}}(\theta)$ in (4) for all $\theta \in \Theta$. From the closed-loop representation (3), an extended form

$$\begin{aligned} x_t = A_{\text{cl}}^m(\theta)x_{t-m} &+ H_y(\theta)\mathbf{y}_{m,t-1} + H_w(\theta)\mathbf{w}_{m,t-1} \\ &- H_y(\theta)\mathbf{v}_{m,t-1} \end{aligned} \tag{7}$$

can be derived over a sliding time window $[t-m, t-1]$, where

$$\begin{aligned} H_y(\theta) &= \begin{bmatrix} A_{\text{cl}}^{m-1}(\theta)L & A_{\text{cl}}^{m-2}(\theta)L & \cdots & L \end{bmatrix}, \\ H_w(\theta) &= \begin{bmatrix} A_{\text{cl}}^{m-1}(\theta) & A_{\text{cl}}^{m-2}(\theta) & \cdots & I \end{bmatrix}, \\ \mathbf{s}_{m,t-1} &= \begin{bmatrix} s_{t-m}^\top & s_{t-m+1}^\top \cdots & s_{t-1}^\top \end{bmatrix}^\top, \end{aligned} \tag{8}$$

and s represents y, w, and v, respectively. Since the spectral radius of the stabilized $A_{\text{cl}}(\theta)$ is smaller than 1, the initial term $A_{\text{cl}}^m(\theta)x_{t-m}$ is negligible when using a sufficiently large m. Hence (7) can be approximated by

$$x_t \approx H_y(\theta)\mathbf{y}_{m,t-1} + H_w(\theta)\mathbf{w}_{m,t-1} - H_y(\theta)\mathbf{v}_{m,t-1}. \tag{9}$$

The selection of m involves a tradeoff between the error of neglecting $A_{\text{cl}}^m(\theta)x_{t-m}$ and the computational complexity of using (9) to construct SM estimates online.

According to Assumption 1, $H_y(\theta)$ defined in (8) is a polynomial matrix that can be expressed as

$$H_y(\theta) = \hat{H}_{y,0} + \sum_{i=1}^{N} \hat{H}_{y,i}\psi_i(\theta), \tag{10}$$

where $\{\psi_i(\theta), 1 \leq i \leq N\}$ are the N monomial bases included in $H_y(\theta)$, and $\{\hat{H}_{y,i}\}$ are the coefficient matrices of $H_y(\theta)$ with respect to these monomial bases. Accordingly, $H_y(\theta)\mathbf{y}_{m,t-1}$ can be rewritten as

$$\begin{aligned} H_y(\theta)\mathbf{y}_{m,t-1} &= \hat{H}_{y,0}\mathbf{y}_{m,t-1} + \sum_{i=1}^{N} \hat{H}_{y,i}\mathbf{y}_{m,t-1}\psi_i(\theta) \\ &= \hat{H}_{y,0}\mathbf{y}_{m,t-1} + Y_{m,t-1}\Psi_N(\theta), \end{aligned} \tag{11}$$

with

$$Y_{m,t-1} = \begin{bmatrix} \hat{H}_{y,1}\mathbf{y}_{m,t-1} & \cdots & \hat{H}_{y,N}\mathbf{y}_{m,t-1} \end{bmatrix}, \tag{12}$$

$$\Psi_N(\theta) = \begin{bmatrix} \psi_1(\theta) & \psi_2(\theta) & \cdots & \psi_N(\theta) \end{bmatrix}^\top. \tag{13}$$

With the above derivations and the definition

$$\eta(\theta, \mathbf{w}_m, \mathbf{v}_m) := H_w(\theta)\mathbf{w}_{m,t-1} - H_y(\theta)\mathbf{v}_{m,t-1}, \tag{14}$$

(9) can be rewritten in a linear affine form:

$$x_t = \hat{H}_{y,0}\mathbf{y}_{m,t-1} + \underbrace{\begin{bmatrix} Y_{m,t-1} & I \end{bmatrix}}_{Z_{m,t-1}} \underbrace{\begin{bmatrix} \Psi_N(\theta) \\ \eta(\theta, \mathbf{w}_m, \mathbf{v}_m) \end{bmatrix}}_{\boldsymbol{\lambda}(\theta, \mathbf{w}_m, \mathbf{v}_m)}. \tag{15}$$

In the second term of (15), the TV coefficient $Z_{m,t-1}$ is determined by online measurements, whereas the uncertainty vector $\boldsymbol{\lambda}(\theta, \mathbf{w}_m, \mathbf{v}_m)$ has TI polynomial dependence on θ, $\mathbf{w}_m$, and $\mathbf{v}_m$ due to Assumptions 2–3 and the definitions (13)–(15). This structure is critical for the probability-guaranteed SM estimation in Section V.

IV. A Polynomial Chaos Approach to Observer Gain Design

The closed-loop representation (3) consists of (i) the implicit state estimate $\hat{x}$ from the observer (5); and (ii) the estimation errors attributed to the error dynamics (6). This section is devoted to designing the observer gain L such that the error dynamics (6) is stable and has a minimal averaged

$\mathcal{H}_2$ performance. For this purpose, the conventional robust synthesis approach is not applicable since the error dynamics (6) has polynomial dependence on θ. Therefore, a polynomial chaos (PC) based synthesis approach is proposed next.

A. Preliminaries on polynomial chaos expansion

For a random vector θ, a function $g(\theta) : \mathbb{R}^{n_\theta} \to \mathbb{R}$ with a finite second-order moment admits a polynomial chaos expansion (PCE) [18],[1]

$$g(\theta) = \sum_{i=0}^{\infty} g_i \phi_i(\theta), \tag{16}$$

where $\{g_i\}$ denotes the expansion coefficients and $\{\phi_i(\theta)\}$ denotes the multivariate PC bases. By using the Askey scheme of orthogonal polynomial bases, this expansion exponentially converges in the $\mathcal{L}_2$ sense, which results in accurate approximations even with a relatively small number of terms for mild assumptions on the system dynamics [18]. These basis functions are orthogonal with respect to the PDF $f(\theta)$ of $\theta \in \Theta$, i.e.,

$$\begin{aligned}\langle \phi_i(\theta), \phi_j(\theta) \rangle &= \int_\Theta \phi_i(\theta)\phi_j(\theta) f(\theta)\, \mathrm{d}\theta = \mathbb{E}_\theta\{\phi_i(\theta)\phi_j(\theta)\} \\ &= \begin{cases} 1, & \text{if } i = j \\ 0, & \text{otherwise.} \end{cases}\end{aligned} \tag{17}$$

Throughout this article, $\{\phi_i(\theta)\}$ are normalized as in (17). By exploiting the normalized orthogonality, each PCE coefficient g_i is computed by $g_i = \langle g(\theta), \phi_i(\theta) \rangle$, which can be calculated via numerical integration [18]. In particular, $\phi_0(\theta) = 1$, and the means of $g(\theta)$ and $\phi_i(\theta)$ are

$$\begin{aligned}\mathbb{E}_\theta\{g(\theta)\} &= \langle g(\theta), \phi_0(\theta) \rangle = g_0, \\ \mathbb{E}_\theta\{\phi_i(\theta)\} &= \langle \phi_i(\theta), \phi_0(\theta) \rangle = \begin{cases} 1, & \text{for } i = 0 \\ 0, & \text{for } i > 0. \end{cases}\end{aligned} \tag{18}$$

In practical computations, a PCE (16) with an infinite degree is truncated as $g(\theta) \approx \hat{g}(\theta) = \sum_{i=0}^{N_p} g_i \phi_i(\theta)$ to a finite degree p, whose number of terms is $N_p + 1 = \frac{(n_\theta + p)!}{n_\theta! p!}$.

B. PCE-expanded error dynamics

In the following, the error dynamics (6) is rewritten as

$$\tilde{x}_{t+1}(\theta) = A_{\text{cl}}(\theta)\tilde{x}_t(\theta) + w_t - Lv_t, \tag{19}$$

in order to emphasize the dependence of $\tilde{x}_t$ on θ. Let $\tilde{x}_{i,t}(\theta)$ denote the i^{th} component of the estimation error $\tilde{x}_t(\theta)$. Its truncated PCE with a sufficiently large degree p is expressed as $\tilde{x}_{i,t}(\theta) = \sum_{j=0}^{N_p} \tilde{x}_{i,t}^{(j)} \phi_j(\theta)$, where $\{\tilde{x}_{i,t}^{(j)}\}$ is the set of PCE coefficients. Define

$$\begin{aligned}\tilde{\mathbf{x}}_{i,t}^{\text{PCE}} &= \begin{bmatrix} \tilde{x}_{i,t}^{(0)} & \tilde{x}_{i,t}^{(1)} & \cdots & \tilde{x}_{i,t}^{(N_p)} \end{bmatrix}^\top, \\ \boldsymbol{\phi}(\theta) &= \begin{bmatrix} \phi_0(\theta) & \phi_1(\theta) & \cdots & \phi_{N_p}(\theta) \end{bmatrix}^\top, \\ \tilde{\mathbf{x}}_t^{\text{PCE}} &= \begin{bmatrix} \tilde{\mathbf{x}}_{1,t}^{\text{PCE}} & \cdots & \tilde{\mathbf{x}}_{n_x,t}^{\text{PCE}} \end{bmatrix},\end{aligned} \tag{20}$$

[1] Some publications refer to this expansion as being a *generalized polynomial chaos expansion*, to denote that the PDF of θ is not restricted to be Gaussian.

then the truncated PCE of the estimation error $\tilde{x}_t(\theta)$ is

$$\tilde{x}_t(\theta) = (\tilde{\mathbf{x}}_t^{\text{PCE}})^\top \boldsymbol{\phi}(\theta) = \underbrace{\left(\boldsymbol{\phi}^\top(\theta) \otimes I_{n_x}\right)}_{\Phi_x^\top(\theta)} \underbrace{\text{vec}\left((\tilde{\mathbf{x}}_t^{\text{PCE}})^\top\right)}_{\tilde{\mathbf{X}}_t}, \tag{21}$$

where $\otimes$ and $\text{vec}(\cdot)$ represent the Kronecker product and the vectorization of a matrix, respectively. In the last equation of (21), the property $\text{vec}(EFG) = (G^\top \otimes E)\text{vec}(F)$ is applied.

Inserting (21) into (19) and left-multiplying by $\Phi_x(\theta)$ leads to

$$\Phi_x(\theta)\Phi_x^\top(\theta)\tilde{\mathbf{X}}_{t+1} = \Phi_x(\theta)A_{\text{cl}}(\theta)\Phi_x^\top(\theta)\tilde{\mathbf{X}}_t + \Phi_x(\theta)B_{\text{cl}}\eta_t, \tag{22}$$

where $B_{\text{cl}} = \begin{bmatrix} I_{n_x} & -L \end{bmatrix}$ and $\eta_t = \begin{bmatrix} w_t^\top & v_t^\top \end{bmatrix}^\top$. Taking the mathematical expectation with respect to θ on both sides of (22) results in the PCE-expanded system

$$\tilde{\mathbf{X}}_{t+1} = \mathcal{A}_{\text{cl}}\tilde{\mathbf{X}}_t + \mathcal{B}_{\text{cl}}\eta_t \tag{23}$$

which describes the dynamics of the PCE coefficient vector $\tilde{\mathbf{X}}_t$, with

$$\begin{aligned}&\mathcal{A}_{\text{cl}} = \mathcal{A} - \mathcal{L}\mathcal{C},\ \mathcal{A} = \mathbb{E}_\theta\{\Phi_x(\theta)A(\theta)\Phi_x^\top(\theta)\}, \\ &\mathcal{L} = I_{N_p+1} \otimes L,\ \mathcal{C} = \mathbb{E}_\theta\{\Phi_y(\theta)C(\theta)\Phi_x^\top(\theta)\}, \\ &\mathcal{B}_{\text{cl}} = \mathbb{E}_\theta\{\Phi_x(\theta)B_{\text{cl}}\} = \begin{bmatrix} B_{\text{cl}}^\top & 0 & \cdots & 0 \end{bmatrix}^\top, \\ &\Phi_y(\theta) = \boldsymbol{\phi}^\top(\theta) \otimes I_{n_y}.\end{aligned} \tag{24}$$

The derivations of $\mathcal{L}$ and $\mathcal{C}$ in (24) exploit the property $\Phi_x(\theta)L = \mathcal{L}\Phi_y(\theta)$, whose proof is referred to Proposition 3 in [1], while $\mathcal{B}_{\text{cl}}$ in (24) is obtained according to (18).

C. PCE-based synthesis

From (21), the averaged $\mathcal{H}_2$ estimation performance with respect to θ for the error dynamics (19) is

$$\begin{aligned}\lim_{t\to\infty} \mathbb{E}_{\theta,\eta}\{\tilde{x}_t^\top(\theta)\tilde{x}_t(\theta)\} &= \lim_{t\to\infty} \mathbb{E}_{\theta,\eta}\{\tilde{\mathbf{X}}_t^\top \Phi_x(\theta)\Phi_x^\top(\theta)\tilde{\mathbf{X}}_t\} \\ &= \lim_{t\to\infty} \mathbb{E}_\eta\{\tilde{\mathbf{X}}_t^\top \tilde{\mathbf{X}}_t\},\end{aligned} \tag{25}$$

since $\mathbb{E}_\theta\{\Phi_x(\theta)\Phi_x^\top(\theta)\} = I_{n_x(N_p+1)}$. Therefore, minimizing the averaged $\mathcal{H}_2$ performance $\lim_{t\to\infty} \mathbb{E}_{\theta,\eta}\{\tilde{x}_t^\top(\theta)\tilde{x}_t(\theta)\}$ is approximated by minimizing $\lim_{t\to\infty} \mathbb{E}_\eta\{\tilde{\mathbf{X}}_t^\top \tilde{\mathbf{X}}_t\}$ for the PCE-expanded system (23).

Proposition 1: The optimization

$$\min_{\mathcal{P},L,S} \ \text{trace}\{\mathcal{P}\} \tag{26a}$$

$$\text{s.t. } \mathcal{P} - \mathcal{A}_{\text{cl}}\mathcal{P}\mathcal{A}_{\text{cl}}^\top - \mathcal{B}_{\text{cl}}Q\mathcal{B}_{\text{cl}}^\top > 0,\ \mathcal{P} > 0, \tag{26b}$$

$$S - A_{\text{cl}}(\theta)SA_{\text{cl}}^\top(\theta) > 0,\ S > 0,\ \forall\theta \in \Theta, \tag{26c}$$

is formulated to minimize $\lim_{t\to\infty} \mathbb{E}_\eta\{\tilde{\mathbf{X}}_t^\top \tilde{\mathbf{X}}_t\}$ while ensuring quadratic stability of the error dynamics (19), where $Q = \text{diag}\{Q_w, Q_v\}$ represents the covariance matrix of $\begin{bmatrix} w_t^\top & v_t^\top \end{bmatrix}^\top$.

With Proposition 1, the $\mathcal{H}_2$ synthesis problem for the error dynamics (19) has been transformed into (26). The cost function (26a) and the first constraint (26b) are standard in $\mathcal{H}_2$ synthesis for the PCE-expanded system (23). Although (26b) ensures stability of (23), it is insufficient to imply stability of the original error dynamics (19), due to the PCE truncation errors introduced in deriving (23) to approximate (19). To deal with this issue, (26c) is additionally imposed such that the quadratic Lyapunov function $V(\tilde{x}_t(\theta)) = \tilde{x}_t^\top(\theta) S \tilde{x}_t(\theta)$ decreases with time, which enforces quadratic stability of the error dynamics (19). Similar strategies have been used in the PCE-based control synthesis [7], [16].

Let $\text{co}\{(A_i, C_i)\}$ denote the convex hull of the indicated vertices (A_i, C_i) such that $(A(\theta), C(\theta)) \in \text{co}\{(A_i, C_i)\}$ for any $\theta \in \Theta$. Using standard procedures, the synthesis problem (26) is then transformed into

$$\min_{\mathcal{P}, L, S} \ \text{trace}\{\mathcal{P}\} \tag{27a}$$

$$\text{s.t.} \begin{bmatrix} \mathcal{P} & \Pi & \mathcal{B}_{\text{cl}} \\ \Pi^\top & \mathcal{P} & 0 \\ \mathcal{B}_{\text{cl}}^\top & 0 & Q^{-1} \end{bmatrix} > 0, \ \mathcal{P} > 0, \ S > 0, \tag{27b}$$

$$\begin{bmatrix} S & \Gamma_i \\ \Gamma_i^\top & S \end{bmatrix} > 0, \ i = 1, \cdots, q, \tag{27c}$$

where q denotes the number of vertices in the polytopic uncertainty, and Π and Γ_i are defined as

$$\Pi = (\mathcal{A} - \mathcal{L}\mathcal{C})\mathcal{P}, \ \Gamma_i = (A_i - LC_i)S.$$

Equation (27c) is a sufficient condition for ensuring (26c), and Π and Γ_i are bilinear terms with regard to $\mathcal{P}$, S, and L. These bilinear terms cannot be converted into linear terms via conventional change-of-variables due to the block-diagonal structure of $\mathcal{L} = I_{N_p+1} \otimes L$ [5], [14].

V. Probability-Guaranteed Set-Membership State Estimation

In this section, the probability-guaranteed SM state estimate is determined by exploiting the linear affine structure of (15) and the GM distribution.

Since $\boldsymbol{\lambda}(\theta, \mathbf{w}_m, \mathbf{v}_m)$ has TI polynomial dependence on θ, $\mathbf{w}_m$, and $\mathbf{v}_m$ as analyzed in Section III, its PDF is also TI and can be approximated offline by a mixture of Gaussians,

$$p(\boldsymbol{\lambda}) = \sum_{k=1}^{K} \pi_k \mathcal{N}(\boldsymbol{\lambda}; \mu_\lambda^{(k)}, \Sigma_\lambda^{(k)}), \tag{28}$$

where K is the number of Gaussian components, the mixing coefficients $\{\pi_k\}$ satisfy $0 \le \pi_k \le 1$ and $\sum_{k=1}^{K} \pi_k = 1$, and $\mu_\lambda^{(k)}$ and $\Sigma_\lambda^{(k)}$ represent the mean and covariance matrix of each Gaussian component, respectively. The offline procedure for constructing the GM distribution (28) is: (i) generate a sufficient number of samples $\{\theta^{(i)}, \mathbf{w}_m^{(i)}, \mathbf{v}_m^{(i)}\}$, and compute $\boldsymbol{\lambda}^{(i)} = \boldsymbol{\lambda}(\theta^{(i)}, \mathbf{w}_m^{(i)}, \mathbf{v}_m^{(i)})$ according to (13)–(15); and (ii) estimate the GM parameters $\{\pi_k, \mu_\lambda^{(k)}, \Sigma_\lambda^{(k)}\}$ in (28) for the sample distribution of $\{\boldsymbol{\lambda}^{(i)}\}$ using the expectation-maximization (EM) algorithm [2]. Note that K can be determined by comparing multiple models with different K's using Akaike information criterion [2].

Both the EM algorithm and the subsequent analysis rely on formulating Gaussian mixtures in (28) in terms of a discrete latent variable $\mathbf{z}$ [2]. Let $\mathbf{z}$ denote a K-dimensional binary random vector with each element z_k satisfying $z_k \in \{0, 1\}$ and $\sum_{k=1}^{K} z_k = 1$, i.e., a particular element z_k is equal to 1 and all the other elements are null. In order to express the GM distribution $p(\boldsymbol{\lambda})$ as a marginal distribution obtained from the joint distribution $p(\boldsymbol{\lambda}, \mathbf{z})$, define the marginal distribution $p(\mathbf{z})$ and the conditional distribution $p(\boldsymbol{\lambda}|\mathbf{z})$ as $p(z_k = 1) = \pi_k$ and $p(\boldsymbol{\lambda}|z_k = 1) = \mathcal{N}(\boldsymbol{\lambda}; \mu_\lambda^{(k)}, \Sigma_\lambda^{(k)})$, respectively. Therefore, the joint distribution is given by $p(\boldsymbol{\lambda}, \mathbf{z}) = p(\mathbf{z})p(\boldsymbol{\lambda}|\mathbf{z})$, and the GM distribution $p(\boldsymbol{\lambda})$ in (28) is then equivalently expressed as

$$p(\boldsymbol{\lambda}) = \sum_{\mathbf{z}} p(\mathbf{z}) p(\boldsymbol{\lambda}|\mathbf{z}) = \sum_{k=1}^{K} p(z_k = 1) p(\boldsymbol{\lambda}|z_k = 1).$$

This expression explicitly associates every realization of $\boldsymbol{\lambda}$ with a discrete value of $\mathbf{z}$, i.e., a realization of $\boldsymbol{\lambda}$ is generated from a conditional Gaussian distribution $p(\boldsymbol{\lambda}|z_k = 1)$.

From the online measurement $\mathbf{y}_{m,t-1}$, $Y_{m,t-1}$ in (12) can be computed from the coefficient matrices $\{\hat{H}_{y,i}\}$ in (10), and $Z_{m,t-1}$ is constructed from its definition in (15). With $z_k = 1$, the conditional distribution of the state in (15) is then derived as

$$\begin{aligned} &p(x_t|z_k = 1) = \mathcal{N}(x_t; \mu_{x,t}^{(k)}, \Sigma_{x,t}^{(k)}), \\ &\mu_{x,t}^{(k)} = Z_{m,t-1}\mu_\lambda^{(k)}, \ \Sigma_{x,t}^{(k)} = Z_{m,t-1}\Sigma_\lambda^{(k)} Z_{m,t-1}^\top, \end{aligned} \tag{29}$$

by performing a linear transformation $Z_{m,t-1}\boldsymbol{\lambda}$ on the conditional Gaussian component $p(\boldsymbol{\lambda}|z_k = 1)$. Hence, the distribution of the state can be also approximated by a Gaussian mixture

$$p(x_t) = \sum_{k=1}^{K} \pi_k \mathcal{N}(x_t; \mu_{x,t}^{(k)}, \Sigma_{x,t}^{(k)}). \tag{30}$$

Next, the SM estimate $\mathcal{X}_t$ is determined such that (2) holds for the GM distribution of the state in (30).

Theorem 1: Given the GM distribution of the state in (30), $\Pr\{x_t \in \mathcal{X}_t\} \ge \gamma$ holds if the SM estimate $\mathcal{X}_t$ is constructed as

$$\mathcal{X}_t = \bigcup_{k=1}^{K} \mathcal{X}_t^{(k)}, \tag{31}$$

where $\mathcal{X}_t^{(k)}$ is an ellipsoidal confidence set,

$$\mathcal{X}_t^{(k)} = \left\{ x \,\middle|\, (x - \mu_{x,t}^{(k)})^\top (\Sigma_{x,t}^{(k)})^{-1} (x - \mu_{x,t}^{(k)}) < \chi_{n_x}^2(\gamma) \right\}, \tag{32}$$

defined for the k^{th} Gaussian component $p(x_t|z_k = 1)$, and $\chi_{n_x}^2(\gamma) \in \mathbb{R}$ represents the value whose cumulative probability under the chi-square distribution with n_x degrees of freedom is specified by γ.

Proof: With $z_k = 1$, x_t follows the conditional Gaussian distribution $p(x_t|z_k = 1)$, and $\Pr\{x_t \in \mathcal{X}_t|z_k = 1\} \ge \Pr\{x_t \in \mathcal{X}_t^{(k)}|z_k = 1\} \ge \gamma$ holds for $\mathcal{X}_t$ and $\mathcal{X}_t^{(k)}$ defined in (31) and (32). Then, it follows that

$$\begin{aligned} \Pr\{x_t \in \mathcal{X}_t\} &= \sum_{k=1}^{K} \Pr\{x_t \in \mathcal{X}_t|z_k = 1\} \Pr\{z_k = 1\} \\ &\ge \gamma \sum_{k=1}^{K} \Pr\{z_k = 1\} = \gamma. \end{aligned}$$

■

Remark 2: Due to the possible overlaps among the ellipsoid sets $\{\mathcal{X}_t^{(k)}\}$, the actually achieved probability $\Pr\{x_t \in \mathcal{X}_t\}$ can be larger than the predefined confidence level γ.

The proposed approach is summarized in Algorithm 1.

Algorithm 1 Probability-guaranteed SM state estimation

Offline procedures:

1) Select the confidence level γ in (2).
2) Design the observer gain L by solving (27) and select the length m of the time window in (9).
3) Determine the monomial bases $\{\psi_i(\theta), i = 0, \cdots, N-1\}$ and the coefficient matrices $\{\hat{H}_{y,i}\}$ in (10).
4) Compute the GM approximation (28) to the distribution of $\boldsymbol{\lambda}(\theta, \mathbf{w}_m, \mathbf{v}_m)$ using the EM algorithm.

Online procedures at each time instant t:

1) Compute $Y_{m,t-1}$ in (12) from the coefficient matrices $\{\hat{H}_{y,i}\}$ in (10) and the online measurement $\mathbf{y}_{m,t-1}$; and construct $Z_{m,t-1}$ in (15).
2) Determine $\{\mu_{x,t}^{(k)}, \Sigma_{x,t}^{(k)}\}$ in (29) for each Gaussian component of the state distribution.
3) Construct the SM state estimate $\mathcal{X}_t$ in (31) and (32).
4) Compute the interval $[\underline{x}_{i,t}, \bar{x}_{i,t}]$ for each state element $x_{i,t}$ by setting $\underline{x}_{i,t} = \min_k l_i \mu_{x,t}^{(k)} - \sqrt{l_i \Sigma_{x,t}^{(k)} l_i^\top}$ and $\bar{x}_{i,t} = \max_k l_i \mu_{x,t}^{(k)} + \sqrt{l_i \Sigma_{x,t}^{(k)} l_i^\top}$, where l_i is a n_x-dimensional row vector whose i^{th} element is 1 and all the other elements are zero.

VI. Simulation Example

Consider the uncertain system (1) with

$$A(\theta) = \begin{bmatrix} 0.9 + 0.2\theta^3 & -0.4 \\ 0.1 & 0.5 + 0.2\theta^2 \end{bmatrix}, \; C = \begin{bmatrix} 0.2 & 0.1 \end{bmatrix},$$

where the uncertain parameter θ is uniformly distributed over the interval $[-1, 1]$. The white noises w_t and v_t are uniformly distributed over $[-0.2, 0.2]$ and $[-0.06, 0.06]$, respectively.

First, the proposed approach is applied by implementing Algorithm 1. In the offline procedures, to account for the uniformly distributed parameter θ in the observer synthesis, the Legendre polynomial bases are adopted with a degree 3, and the resulting observer gain is $L = \begin{bmatrix} 2.3828 & 0.3574 \end{bmatrix}^\top$ by solving (27) using the PENBMI solver [6]. In (15), the estimation window length is determined to be $m = 15$ by applying Proposition 1 in [12], such that $\|A_{\text{cl}}^m(\theta)\|_\infty \leq 0.002$. Then, the vector $\Psi_N(\theta)$ of the nominal bases is $\Psi_N(\theta) = \begin{bmatrix} \theta^2 & \theta^3 & \cdots & \theta^{38} & \theta^{39} & \theta^{42} \end{bmatrix}^\top$ for this example. For computing the GM approximation to the distribution of $\boldsymbol{\lambda}(\theta, \mathbf{w}_m, \mathbf{v}_m) \in \mathbb{R}^{41}$, 10^5 random samples are generated, and the EM algorithm is implemented using the MATLAB built-in function `fitgmdist`. The obtained GM distribution of $\boldsymbol{\lambda}$ has 3 Gaussian components. Figure 1 depicts a two-dimensional projection of $\boldsymbol{\lambda}$ onto the subspace of $\begin{bmatrix} \lambda_2 & \lambda_{41} \end{bmatrix}$, where λ_i is the i^{th} element of $\boldsymbol{\lambda}$. The confidence level γ is set to be 95%. The online SM estimates are represented

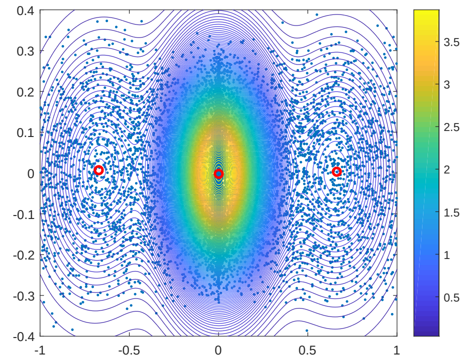

Fig. 1. The GM approximation to the distribution of $[\lambda_2 \; \lambda_{41}]$, where λ_i is the i^{th} element of $\boldsymbol{\lambda} \in \mathbb{R}^{41}$. The scatter plots are random samples of $[\lambda_2 \; \lambda_{41}]$. The contour plot shows the obtained GM distribution of $[\lambda_2 \; \lambda_{41}]$. The red circles represent the means of the 3 Gaussian components.

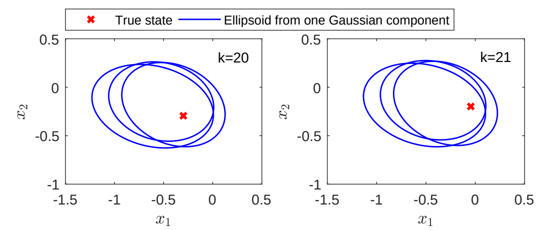

Fig. 2. Proposed SM state estimates at time $k = 20$ and $k = 21$ when the uncertain system (1) is simulated with $\theta = 0.94$.

by the union set of 3 ellipsoids derived from 3 Gaussian components respectively, as illustrated in Figure 2 for the uncertain system simulated with $\theta = 0.94$. From these SM estimates, the intervals of two state elements are determined as shown in Figure 3.

To compare with the proposed method, the zonotopic approach in [9] is also implemented which utilizes only the bounds of θ, w_t, and v_t. To limit the complexity of a state bounding zonotope, zonotope reduction is performed such that the maximal number of zonotope generators is 20, where the definition of a generator is referred to Section 1.1.4 of [8]. For a particular realization $\theta = 0.94$, the state interval bounds derived from the zonotopic method are looser than those of the proposed method, as seen in Figure 3.

To further evaluate the statistical estimation performance, 100 Monte Carlo simulation runs are implemented, with 100 time steps in each run. The tightness of the interval bounds $\{[\underline{x}_{i,t}, \bar{x}_{i,t}], 1 \leq i \leq n_x\}$ for the actual state $\{x_{i,t}, 1 \leq i \leq n_x\}$ is evaluated by the distributions of $\bar{x}_{i,t} - x_{i,t}$ and $x_{i,t} - \underline{x}_{i,t}$, as illustrated in Figure 4. Our proposed method achieves much tighter interval bounds. The more conservative results from the zonotopic method in [9] are due to using a polytopic overbounding for the polynomial uncertainty structure and adopting the TV assumption for the uncertain parameters. An SM estimate is considered valid if the actual state element $x_{i,t}$ belongs to $[\underline{x}_{i,t}, \bar{x}_{i,t}]$. For this simulation example, the rate of valid SM estimates from our proposed method is 99.36%, which is higher than the predefined confidence level 95% due to the reason in Remark

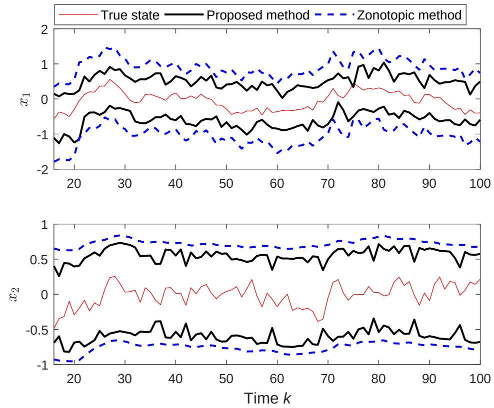

Fig. 3. State interval bounds from our proposed method and the zonotopic method in [9] when the unknown system (1) is simulated with $\theta = 0.94$.

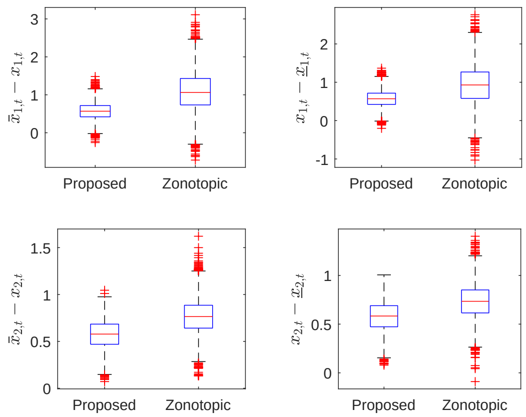

Fig. 4. Distributions of the errors of the state interval bounds $[\underline{x}_{i,t}, \bar{x}_{i,t}]$ from our proposed method and the zonotopic method in [9].

2. Note also that the 3.59% of the obtained zonotopic estimates do not include the actual states. This has a controversy to the expectation that a deterministic zonotopic approach guarantees 100% valid SM estimates by accounting for all possible realizations of uncertainties. The reason is attributed to another limitation of the zonotopic approach in [9]: it is restricted to quadratically stable uncertain systems, whilst the simulated system is unstable at certain values of θ.

In a MATLAB environment (2.5 GHz processor and 6 GB RAM), our proposed method and the zonotopic approach spend 3.5 and 0.7 milliseconds in the worst case, respectively, for each sampling interval. This shows that our proposed method has heavier computational load as it processes a batch of measurements over a sliding window.

VII. Conclusions

A probability-guaranteed set-membership state estimation approach is proposed for uncertain linear time-invariant systems with probabilistic uncertainties. It includes (i) an offline observer synthesis that minimizes the averaged $\mathcal{H}_2$ performance using polynomial chaos, and (ii) an extended observer form with a time-varying linear affine structure derived for online computing the Gaussian mixture distribution of state. The set-membership estimate is then obtained by efficient ellipsoidal calculus. The proposed approach achieves tighter set-membership estimates due to respecting time invariance of uncertain parameters and exploiting the polynomial uncertainty structure.

Acknowledgment

This work is supported by the National Science Foundation of China, Grant No. 61803163, and the Open Project Fund from the State Key Laboratory of Intelligent Control and Decision of Complex Systems (Beijing Institute of Technology).

References

[1] R. Bhattacharya. Robust state feedback control design with probabilistic system parameters. In *Proceedings of 53rd IEEE Conference on Decision and Control*, pages 2828–2833, Los Angeles, 2014.

[2] C. M. Bishop. *Pattern Recognition and Machine Learning*. Springer, Singapore, 2006.

[3] G. Chen, J. Wang, and L. S. Shieh. Interval Kalman filtering. *IEEE Transactions on Aerospace and Electronic Systems*, 33(1):250–259, 1997.

[4] C. Combastel. An extended zonotopic and Gaussian Kalman filter (EZGKF) merging set-membership and stochastic paradigms: Toward non-linear filtering and fault detection. *Annual Reviews in Control*, 42:232–243, 2016.

[5] J. Fisher and R. Bhattacharya. Linear quadratic regulation of systems with stochastic parameter uncertainties. *Automatica*, 45:2831–2841, 2009.

[6] K. Holmström, A. O. Göran, and M. M. Edvall. *User's Guide for TOMLAB/PENOPT*. Tomlab Optimization Inc.

[7] S. Hsu and R. Bhattacharya. Design of linear parameter varying quadratic regulator in polynomial chaos framework. *arXiv preprint arXiv:1711.03582*, 2017.

[8] V. T. H. Le, C. Stoica, T. Alamo, E. F. Camacho, and D. Dumur. *Zonotopes: From Guaranteed State-Estimation to Control*. John Wiley & Sons, New Jersey, USA, 2013.

[9] V. T. H. Le, C. Stoica, T. Alamo, E. F. Camacho, and D. Dumur. Zonotopic guaranteed state estimation for uncertain systems. *Automatica*, 49:3418–3424, 2013.

[10] L. Ljung. *System Identification: Theory for the User*. Prentice Hall, New Jersey, 1987.

[11] Y. Luo, Y. Zhu, X. Shen, and E. Song. Novel data association algorithm based on integrated random coefficient matrices kalman filtering. *IEEE Transactions on Aerospace and Electronic Systems*, 48(1):144–158, 2012.

[12] J. Meseguer, V. Puig, and T. Escobet. Approximating fault detection linear interval observers using λ-order interval predictors. *Int. J. of Adaptive Control & Signal Processing*, 31(7):1040–1060, 2017.

[13] B. T. Polyak, S. A Nazin, C. Durieu, and E. Walter. Ellipsoidal parameter or state estimation under model uncertainty. *Automatica*, 40(7):1171–1179, 2004.

[14] D. Shen, S. Lucia, Y. Wan, R. Findeisen, and R. D. Braatz. Polynomial chaos-based $\mathcal{H}_2$-optimal static output feedback control of systems with probabilistic parameter uncertainties. In *Proceedings of 20th IFAC World Congress*, pages 3595–3600, Toulouse, France, 2017.

[15] E. Süli and D. F. Mayers. *An Introduction to Numerical Analysis*. Cambridge University Press, Cambridge, UK, 2003.

[16] Y. Wan and R. D. Braatz. Mixed polynomial chaos and worst-case synthesis approach to robust observer based linear quadratic regulation. In *Proceedings of the American Control Conference*, pages 6798–6803, Milwaukee, USA, 2018.

[17] G. Wei, S. Liu, Y. Song, and Y. Liu. Probability-guaranteed set-membership filtering for systems with incomplete measurements. *Automatica*, 60:12–16, 2015.

[18] D. Xiu. *Numerical Methods for Stochastic Computations: A Spectral Method Approach*. Princeton University Press, New Jersey, 2010.

[19] C. Yu, N. Xiao, C. Zhang, and L. Xie. An optimal deconvolution smoother for systems with random parametric uncertainty and its application to semi-blind deconvolution. *Signal Processing*, 92(10):2497–2508, 2012.