ARTICLE IN PRESS

Automatica xxx (xxxx) xxx

Contents lists available at ScienceDirect

Automatica

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

Brief paper

On stochastic linear systems with zonotopic support sets☆

Mario E. Villanueva *, Boris Houska

School of Information Science and Technology, ShanghaiTech University, China

ARTICLE INFO

Article history:
Received 8 November 2018
Received in revised form 19 May 2019
Accepted 19 September 2019
Available online xxxx

Keywords:
Lyapunov equations
Robust control
Stochastic processes
Linear systems

ABSTRACT

This paper analyzes stochastic linear discrete-time processes, whose process noise sequence consists of independent and uniformly distributed random variables on given zonotopes. We propose a cumulant-based approach for approximating both the transient and limit distributions of the associated state sequence. The method relies on a novel class of k-symmetric Lyapunov equations, which are used to construct explicit expressions for the cumulants. The state distribution is recovered via a generalized Gram–Charlier expansion with respect to products of a multivariate variant of Wigner's semicircle distribution using Chebyshev polynomials of the second kind. This expansion converges uniformly, under surprisingly mild conditions, to the exact state distribution of the system. A robust feedback control synthesis problem is used to illustrate the proposed approach.

1. Introduction

Stochastic linear discrete-time processes have been of major importance in control theory ever since its early days (Bertsekas & Shreve, 1978; Caines, 1988; Dragan, Morozan, & Stoica, 2010; Kalman, 1960). For instance, stochastic linear system theory is the basis for the classical LQG controller (Stengel, 1994). In this case, the elements of the disturbance sequence of a linear discrete-time system are modeled as random variables with Gaussian probability distributions. This has the advantage that the state has a Gaussian distribution, whose variance can be computed by solving a Lyapunov recursion (Bittanti, Colaneri, & De Nicolao, 1991; Stengel, 1994).

Since real systems are subject to physical limitations, a natural question one may ask is what happens when the process noise has a bounded support. Assuming the system is asymptotically stable, one can assert that in the limit, the state cannot have a Gaussian distribution. This is a consequence of the bounded-input-bounded-output lemma (Blanchini & Miani, 2008). Thus, a Gaussian distribution is, in general, not particularly suited for approximating the distribution of the state of a stochastic system with bounded inputs. In particular, generalizing the *central limit theorem* (Klartag, 2007) for such systems, is impossible.

Numerical methods for computing (or approximating) the probability distribution of the state of a stochastic linear dynamic system with bounded inputs, typically proceed in two steps. In the first step, the reachable set of the system—herein referred to as the support of the probability distribution of the state—is computed or (over-) approximated. Current reachability methods include ellipsoidal calculus (Boyd, El-Ghaoui, Feron, & Balakrishnan, 2004; Kurzhanskiy & Varaiya, 2007), polytopic and zonotopic bounding techniques (Bitsoris, 1988), as well as more general (non-)convex set propagation techniques (Aubin, 1991; Chachuat et al., 2015; Villanueva, Houska, & Chachuat, 2015). A more complete overview of methods for computing reachable sets of linear systems can be found in Blanchini (1999) and Blanchini and Miani (2008).

The second step corresponds to the approximation of the probability distribution of the system's state on its support set. For this task, a variety of methods are available. For example, one could use *sampling-based techniques*. These include Monte-Carlo and quasi Monte-Carlo (Caflisch, 1998); as well as Latin hypercube sampling methods (Loh, 1996; Stein, 1987). However, since these methods are based on exhaustive sampling, they can only reach reasonable accuracies if the problem at hand is of modest dimension (Xiu, 2010).

Another class of methods for approximating a distribution are so called (generalized) *Polynomial chaos expansion* (PCE) methods. These are popular tools for uncertainty quantification (Xiu, 2010; Xiu & Karniadakis, 2002) which must (by now) be considered as a state-of-the-art methodology for analyzing uncertain process control systems (Nagy & Braatz, 2007). PCE is based on expanding the uncertain variables with respect to a finite dimensional basis. As such, these expansions can only be evaluated for dynamic processes with a small to moderate amount of finite-dimensional

☆ This work was supported by the National Natural Science Foundation China (NSFC), Nr. 61473185, as well as Shang-haiTech University, Grant-No. F-0203-14-012. The material in this paper was not presented at any conference. This paper was recommended for publication in revised form by Associate Editor Hyeong Soo Chang under the direction of Editor Ian R. Petersen.

* Corresponding author.

E-mail addresses: meduardov@shanghaitech.edu.cn (M.E. Villanueva), borish@shanghaitech.edu.cn (B. Houska).

https://doi.org/10.1016/j.automatica.2019.108652

uncertain variables. Therefore, in their current state, PCE methods cannot be applied directly to stochastic processes, where the disturbance sequence has infinitely many elements.

The main goal of this paper is the development of numerical algorithms for computing (or approximating) the transient and limit distributions of the state of a stochastic linear system, whose process noise has a bounded zonotopic support. In this paper, we do not follow the sampling or PCE route, since these methods are only efficient in low-dimensional spaces. Instead, this paper proposes a *moment (or cumulant) based approach* (Zhang, 2002) for constructing accurate approximations of probability density functions. Notice that if one is able to compute the cumulants of a distribution, a *generalized Gram–Charlier* (or *Edgeworth*) *series* can be used to reconstruct its underlying probability density function (Berkowitz & Garner, 1970; Cohen, 1998; Kendall & Stuart, 1969). For a history of Gram–Charlier expansions, the reader is referred to Hald (2000, 2002) and for a discussion of their multivariate extensions to Berkowitz and Garner (1970) and Withers and Nadarajah (2014).

Contributions

The two main contributions of this paper can be outlined as follows.

- We provide an explicit formula for computing the cumulants of both the transient and limit distributions of the state sequence of a stochastic discrete-time linear systems, whose process noise is uniformly distributed on a zonotope. This formula is based on a novel class of k-symmetric Lyapunov recursions, as summarized in Theorem 1.
- We develop a generalized Gram–Charlier expansion (GCE) in order to recover accurate numerical approximations of the distribution of the state sequence from its cumulants. This expansion uses Wigner semicircle distributions and their associated Chebyshev polynomials of the second kind in order to construct an expansion that converges under mild regularity assumptions on the reachability properties of the underlying linear system, as summarized in Theorem 2.

Section 2 introduces the problem formulation, while Sections 3 and 4 elaborate on the contributions outlined above. Section 5 illustrates the corresponding methodology by analyzing a case study from the field of robust control. Finally, Section 6 concludes the paper.

Notation and preliminaries

We use the notation $A \otimes B$ to denote the Kronecker product of two matrices A and B. Moreover,

$$A^{\langle k \rangle} = \underbrace{A \otimes A \otimes \cdots \otimes A}_{k \text{ times}}$$

denotes the Kronecker power of A for $k \geq 1$. Here, we define $A^{\langle 0 \rangle} = 1$. The function

$$\mathrm{Sym}_k(A) = \frac{1}{(2k)!} (\nabla \nabla^\intercal)^{\langle k \rangle} \left((x^\intercal)^{\langle k \rangle} A x^{\langle k \rangle} \right)$$

denotes the symmetrizer of order k, as analyzed in Holmquist (1985) and Schott (2003). It is defined for all matrices $A \in \mathbb{R}^{n^k \times n^k}$. Moreover, A is called k-symmetric if $\mathrm{Sym}_k(A) = A$. The symbol ∇ is used to denote the gradient operator and $\nabla\nabla^\intercal F(x)$ denotes the Hessian matrix of a function $F : \mathbb{R}^n \to \mathbb{R}$ at $x \in \mathbb{R}^n$. The trace of a square matrix A is denoted by $\mathrm{Tr}(A)$. The Hadamard (or componentwise) product of two matrices, is denoted by $A \odot B$ for $A, B \in \mathbb{R}^{m \times n}$. The symbol $\mathbb{1}$ denotes a unit matrix of appropriate dimensions.

Throughout this paper $\mathbb{I}^n = [-1, 1]^n$ denotes the closed n-dimensional unit box, $\mathrm{int}(\mathbb{I}^n)$ its interior, and

$$G\mathbb{I}^n + c = \left\{ Gz + c \mid z \in \mathbb{I}^n \right\}$$

a zonotope with center c and shape matrix G. The total variation of an integrable function $f : \mathbb{I}^n \to \mathbb{R}$ is given by

$$\|f\|_{\mathrm{TV}} = \max_{Y' \subseteq \mathbb{I}^n} \left| \int_{Y'} f(y) \mathrm{d}y \right| .$$

The probability of an event is denoted by $\Pr(\cdot)$. For example, if $z \in \mathbb{R}^n$ is a random variable with probability distribution ρ, we use the notation

$$\Pr(z \in Z) = \int_Z \rho(z) \mathrm{d}z$$

to denote the probability that z is in the set $Z \subseteq \mathbb{R}^n$.

2. Problem formulation

This paper is concerned with stochastic linear discrete-time processes of the form

$$x_{k+1} = Ax_k + Bw_k \quad \text{with} \quad x_0 = 0 \, . \tag{1}$$

Here, x_k and w_k denote the state and process noise at time k, respectively. The matrices $A \in \mathbb{R}^{n_x \times n_x}$, and $B \in \mathbb{R}^{n_x \times n_w}$ are assumed to be given.

Assumption 1. The elements, w_k, of the process noise sequence are, for all $k \in \mathbb{N}$, independent and identically distributed (i.i.d.) uniform random variables over $\mathbb{I}^{n_w}$.

The reachable set of the system at time $n \geq 1$,

$$X_n = \left\{ \sum_{k=0}^{n-1} A^k B w_{n-k-1} \;\middle|\; \begin{array}{l} \forall k \in \{0, \ldots, n-1\}, \\ w_k \in \mathbb{I}^{n_w} \end{array} \right\} ,$$

is a zonotope (Blanchini & Miani, 2008; Kolmanovsky & Gilbert, 1998). Notice that X_n can be interpreted as the support of the probability distribution ρ_n, of the state x_n at time n. Moreover, if all eigenvalues of A are in the open unit disk, the limit set

$$X_\infty = \lim_{n \to \infty} X_n$$

exists[1] and it is bounded (Bittanti et al., 1991; Blanchini & Miani, 2008)—although it is in general not a zonotope. Thus, neither the distributions ρ_n nor the limit distribution for $n \to \infty$ (if this limit exists) are Gaussian. Moreover, an explicit characterization of these functions is, in general, impossible. Therefore, the goal of the remainder of this paper is to find computationally tractable approximations of these distributions.

3. Cumulant generating function

This section presents an analysis of the cumulant generating functions

$$\forall y \in \mathbb{R}^{n_x}, \quad C_n(y) = \log \left(\mathbb{E} \left(e^{y^\intercal x_n} \right) \right) \tag{2}$$

that are associated with the distributions ρ_n (Kendall & Stuart, 1969). We analyze these functions on the polar sets

$$X_n^* = \left\{ y \in \mathbb{R}^{n_x} \;\middle|\; \max_{x \in X_n} y^\intercal x \leq 1 \right\}$$

of the reachable sets X_n. Let $\zeta(s) = \sum_{n=1}^\infty \frac{1}{n^s}$ denote Riemann's ζ-function and B_j the jth column of B.

[1] The set X_∞ is a limit of the sequence (X_n), if the Hausdorff distance between X_∞ and X_n converges to 0 for $n \to \infty$.

Please cite this article as: M.E. Villanueva and B. Houska, On stochastic linear systems with zonotopic support sets. Automatica (2019) 108652, https://doi.org/10.1016/j.automatica.2019.108652.

Lemma 1. *Let Assumption 1 be satisfied. The cumulant generating function C_n satisfies*

$$C_n(y) = \sum_{r=1}^{\infty}\sum_{k=0}^{n-1}\sum_{j=1}^{n_w} \frac{(-1)^{r+1}\zeta(2r)}{r\pi^{2r}} \left(B_j^\intercal (A^\intercal)^k y\right)^{2r} \tag{3}$$

for all $n \in \mathbb{N}$. The infinite sum on the right-hand of the above expression converges uniformly for all $y \in X_n^*$.

Proof. Let $\mathcal{M}_n$ denote the moment generating function,

$$\mathcal{M}_n(y) = \mathbb{E}\left(e^{y^\intercal x_n}\right) = \int_{X_n} e^{y^\intercal x}\rho_n(x)\,\mathrm{d}x\ ,$$

which is closely related to the cumulant generating function, $C_n(y) = \log(\mathcal{M}_n(y))$. By substituting the explicit solution of (1) in the integral above and using Assumption 1 we obtain

$$\mathcal{M}_n(y) = \prod_{k=0}^{n-1}\int_{\mathbb{I}^{n_w}} \frac{\exp\left(y^\intercal A^k B w_{n-k-1}\right)}{2^{n_w}}\mathrm{d}w_{n-k-1}\ .$$

Next, an application of Fubini's theorem and using the definition of the hyperbolic sine function yields

$$\mathcal{M}_n(y) = \prod_{k=0}^{n-1}\prod_{j=1}^{n_w}\frac{1}{y^\intercal A^k B_j}\sinh\left(y^\intercal A^k B_j\right)\ .$$

Since $\xi \mapsto \frac{\sinh(\xi)}{\xi}$ is an entire function with roots at $\xi_\ell = \sqrt{-1}\pi\ell$ for all $\ell \in \mathbb{Z}\setminus\{0\}$, we can apply Weierstrass' factorization theorem (Busam & Freitag, 2009) to write

$$\frac{\sinh(\xi)}{\xi} = \prod_{\ell=1}^{\infty}\left(1+\frac{\xi^2}{\pi^2\ell^2}\right)\ .$$

This representation converges absolutely for any $\xi \in \mathbb{C}$. Thus, we arrive at the convergent product representation

$$\mathcal{M}_n(y) = \prod_{k=0}^{n-1}\prod_{j=1}^{n_w}\prod_{\ell=1}^{\infty}\left(1+\frac{\left(y^\intercal A^k B_j\right)^2}{\pi^2\ell^2}\right). \tag{4}$$

Taking logarithms on both sides yields

$$C_n(y) = \sum_{k=0}^{n-1}\sum_{j=1}^{n_w}\sum_{\ell=1}^{\infty}\log\left(1+\frac{\left(y^\intercal A^k B_j\right)^2}{\pi^2\ell^2}\right)\ .$$

Since we have

$$\log(1+\xi) = \sum_{r=1}^{\infty}\frac{(-1)^{r+1}}{r}\xi^r$$

for all $\xi \in (-1,1)$, we find that the equation

$$C_n(y) = \sum_{r=1}^{\infty}\sum_{k=0}^{n-1}\sum_{j=1}^{n_w}\frac{(-1)^{r+1}\zeta(2r)}{r\pi^{2r}}\left[B_j^\intercal(A^\intercal)^k y\right]^{2r}$$

holds for all y such that $|B_j^\intercal(A^\intercal)^k y| < \pi$, for all $j \in \{1,\ldots,n_w\}$, and all $k \in \{1,\ldots,n\}$. The statement of the lemma follows, as this inequality holds for all $y \in X_n^*$. □

Lemma 1 can be used to construct a computationally tractable explicit expression of the cumulants of ρ_n. Let us introduce the r-symmetric matrix

$$Q_r = \sum_{j=1}^{n_w} B_j^{\langle r\rangle}\left(B_j^{\langle r\rangle}\right)^\intercal \tag{5}$$

and its associated generalized Lyapunov recursion

$$P_{r,k+1} = \mathrm{Sym}_r\left(A^{\langle r\rangle}P_{r,k}\left(A^{\langle r\rangle}\right)^\intercal + Q_r\right) \tag{6}$$

with $P_{r,0} = 0$ for all $k, r \in \mathbb{N}$. For $r = 1$, (6) corresponds to the standard Lyapunov discrete-time recursion.

Theorem 1. *Let Assumption 1 be satisfied and let $P_{r,n}$ denote the solution of (6). If we set*

$$K_{2r,n} = \frac{(-1)^{r+1}(2r)!\zeta(2r)}{r\pi^{2r}}P_{r,n}\ ,$$

then the cumulant generating functions C_n satisfy

$$C_n(y) = \sum_{r=1}^{\infty}\frac{1}{(2r)!}\left(y^{\langle r\rangle}\right)^\intercal K_{2r,n}y^{\langle r\rangle} \tag{7}$$

for all $y \in X_n^*$ and all $n \in \mathbb{N}$.

Proof. An application of the mixed product rule for Kronecker products on (3) yields

$$S_{j,k,r} = \left(B_j^\intercal(A^\intercal)^k y\right)^{2r} = \left(y^\intercal A^k B_j B_j^\intercal (A^\intercal)^k y\right)^{\langle r\rangle}\ .$$

Let us take the sums over k and j on both sides and substitute

$$Q_r = \sum_{j=1}^{n_w} B_j^{\langle r\rangle}\left(B_j^{\langle r\rangle}\right)^\intercal \tag{8}$$

$$P_{r,n} = \sum_{k=0}^{n-1}\mathrm{Sym}_r\left(\left(A^{\langle r\rangle}\right)^k Q_r\left(\left(A^{\langle r\rangle}\right)^\intercal\right)^k\right). \tag{9}$$

This yields the equation

$$\sum_{k=0}^{n-1}\sum_{j=1}^{n_w} S_{j,k,r} = \left(y^{\langle r\rangle}\right)^\intercal P_{r,n}y^{\langle r\rangle}. \tag{10}$$

The relation for $P_{r,n}$ in (9) is the unique solution of (6). Now, Lemma 1 implies that

$$C_n(y) = \sum_{r=1}^{\infty}\frac{(-1)^{r+1}\zeta(2r)}{r\pi^{2r}}\left(y^{\langle r\rangle}\right)^\intercal P_{r,n}y^{\langle r\rangle}\ . \tag{11}$$

The statement of the theorem follows after a re-scaling of the coefficients by means of the expression

$$K_{2r,n} = \frac{(-1)^{r+1}(2r)!\zeta(2r)}{r\pi^{2r}}P_{r,n}\ ,$$

for the even cumulants of ρ_n. □

Notice that the odd cumulants, $K_{2r-1,n} = 0$, vanish due to the symmetry of the probability distribution ρ_n.

Remark 1. Clearly, the generalized Lyapunov equation (12) is time-autonomous. Therefore, it can be "warm-started" using an initial value $P_{r,0} \neq 0$ corresponding to the re-scaled cumulants of an appropriately defined random variable x_0 representing the unknown initial state.

3.1. Limit behavior

In order to analyze the limit distribution of the state sequence induced by (1), we introduce the following technical proposition.

Proposition 1. *The following statements are equivalent.*

(1) The limit set X_∞ exists and it is bounded.
(2) The symmetric discrete-time Lyapunov equation

$$P_{1,\infty} = AP_{1,\infty}A^\intercal + BB^\intercal \tag{12}$$

admits a positive semi-definite solution $P_{1,\infty} \succeq 0$.
(3) The generalized Lyapunov equation

$$P_{r,\infty} = \mathrm{Sym}_r\left(A^{\langle r\rangle}P_{r,\infty}A^{\langle r\rangle} + Q_r\right) \tag{13}$$

admits a positive semi-definite solution $P_{r,\infty} \succeq 0$ for all $r \in \mathbb{N}$.

Proof. The equivalence of the first two statements is well known (Bittanti et al., 1991; Kolmanovsky & Gilbert, 1998). To establish the remaining equivalence, we assume momentarily that the symmetric matrix

$$Y = \sum_{k=0}^{n_x} A^k BB^\intercal (A^k)^\intercal$$

has full rank and such that (12) admits a positive semi-definite solution if and only if all the eigenvalues of A are in the open unit disk. Since the spectrum of $A^{\langle k\rangle}$, is given by

$$\mathrm{spec}\left(A^{\langle k\rangle}\right) = \left\{ \prod_{j=1}^{k} \lambda_j \;\middle|\; \lambda_1, \ldots, \lambda_k \in \mathrm{spec}(A) \right\},$$

Eq. (13) admits a positive semi-definite solution if and only if (12) has a positive semi-definite solution. This shows that the second and third statements of Proposition 1 are equivalent whenever Y has full-rank. Otherwise, if Y does not have full-rank, one can project (12) and (13) onto the subspaces spanned by Y and $Y^{\langle k\rangle}$, respectively. One can then apply an analogous argument in these subspaces to show that the statements of Proposition 1 remain correct without any full-rank assumptions. □

Corollary 1. *Let the generalized Lyapunov equations (13) admit positive semi-definite solutions $P_{r,\infty}$. Then, the limit distribution $\rho_\infty = \lim_{n\to\infty} \rho_n$ exists. Moreover, its even cumulants are given by*

$$K_{2r,\infty} = \frac{(-1)^{r+1}(2r)!\zeta(2r)}{r\pi^{2r}} P_{r,\infty}$$

for all $r \in \mathbb{N}$.

The statement of this corollary follows by combining the results from Proposition 1 and Theorem 1.[2]

Remark 2. Notice that the even moments of ρ_n can be recovered from the cumulants $K_{2k,n}$ using the recursion (Noschese & Ricci, 2003)

$$M_{2r+2,n} = \mathrm{Sym}_{r+1}\left(\sum_{i=0}^{r} \binom{2r+1}{2i+1} M_{2(r-i),n} \otimes K_{2(i+1),n} \right)$$

with $r \in \{0, 1, 2, \ldots\}$. This iteration is started at $M_{0,n} = 1$ and generates even multivariate Bell polynomials (Withers & Nadarajah, 2014).

4. Generalized Gram–Charlier expansions with respect to a multivariate Chebyshev basis

Let ω be the multivariate Wigner semicircle distribution,

$$\forall x \in \mathbb{I}^{n_x}, \qquad \omega(x) = \left(\frac{2}{\pi}\right)^{n_x} \prod_{j=1}^{n_x} \sqrt{1 - x_j^2}$$

and $\mathcal{U}$ the generating function of the multivariate Chebyshev polynomials of the second kind,

$$\mathcal{U}(x, y) = e^{y^\intercal x} \prod_{j=1}^{n_x} \left(\cosh\left(y_j\sqrt{x_j^2 - 1}\right) + \frac{x_j \sinh\left(y_j\sqrt{x_j^2-1}\right)}{\sqrt{x_j^2 - 1}} \right).$$

Our goal is to construct a generalized multivariate GCE of the function ρ_n with respect to ω. Let

$$\forall x \in \mathbb{I}^{n_x},\ \forall r \in \mathbb{N}, \quad \Psi_{2r}(x) = \left(\nabla_y \nabla_y^\intercal\right)^{\langle r\rangle} \mathcal{U}(x, 0)$$

be the even multivariate Chebyshev polynomials (of the second kind) in matrix form.

Theorem 2. *Assume (w.l.o.g.) that $X_n \subseteq \mathrm{int}(\mathbb{I}^{n_x})$ for any given $n \in \mathbb{N}$. Furthermore, let the pair (A, B) be reachable. Then, one can find unique r-symmetric GCE coefficients $\Lambda_{r,n} \in \mathbb{R}^{n_x^r \times n_x^r}$ such that there exists for every $\nu \geq 1$ a constant $V < \infty$ such that the function*

$$\phi_{N,n}(x) = \left(\sum_{r=0}^{N} \mathrm{Tr}\left(\Lambda_{r,n}\Psi_{2r}(x)\right) \right) \omega(x)$$

satisfies

$$\left\| \phi_{N,n} - \rho_n \right\|_{\mathrm{TV}} \leq \frac{V}{\nu(N-\nu)^\nu}$$

for all $n \geq \nu n_x + 1$ and all $N > \nu$. In particular, if $n \geq n_x + 1$, the GCE approximation $\phi_{N,n}$ converges uniformly to ρ_n on $\mathbb{I}^{n_x}$ for $N \to \infty$.

Proof. The main idea behind the proof for this theorem, is to use the smoothing properties of the convolution operator. Because we assume that (1) is reachable, it is an immediate consequence of the Cayley Hamilton theorem and the pigeonhole principle that x_n is for all $n \geq n_x + 1$ a superposition of at least 2 random variables with bounded (and unimodal) probability distributions. Thus, ρ_n is, for all $n \geq n_x + 1$, a Lipschitz continuous function. Likewise, if $n \geq \nu n_x + 1$, the $(\nu - 1)$th derivative of ρ_n exists and is Lipschitz continuous.

Moreover, the assumption $X_n \subseteq \mathrm{int}(\mathbb{I}^{n_x})$ implies that the auxiliary functions

$$\chi_n(x) = \rho_n(x)\frac{1}{\omega(x)} = \rho_n(x)\left(\frac{\pi}{2}\right)^{n_x} \frac{1}{\prod_{j=1}^{n_x}\sqrt{1 - x_j^2}}$$

have globally Lipschitz continuous $(\nu - 1)$th derivatives on $\mathbb{I}^{n_x}$ for all $n \geq \nu n_x + 1$. The singularities of the derivatives of ω on the boundary of $\mathbb{I}^{n_x}$ cancel out, since $\rho_n(x) = 0$ for all $x \notin X_n$. Thus, the Chebyshev series of

$$\chi_n(x) = \sum_{r=0}^{\infty} \mathrm{Tr}\left(\Lambda_{r,n}\Psi_{2r}(x)\right), \tag{14}$$

is absolutely and uniformly convergent for all $x \in \mathbb{I}^{n_x}$ and all $n \geq \nu n_x + 1$. A proof of this result can be found in Trefethen (2013),[3] where the convergence rate estimate,

$$\left\| \chi_n - \sum_{r=0}^{N} \mathrm{Tr}\left(\Lambda_{r,n}\Psi_{2r}\right) \right\|_{\mathrm{TV}} \leq \mathbf{O}\left(\frac{1}{\nu(N-\nu)^\nu}\right),$$

can also be found. The statement of the theorem follows after multiplying both sides of Eq. (14) with $\omega(x)$. □

[2] Notice that $0 \in X_\infty^*$ if the conditions from Corollary 1 are satisfied, i.e., (3) formally holds for $n = \infty$ in an open neighborhood of 0.

[3] The result in Trefethen (2013) is valid for Chebyshev polynomials of the first kind. However, it can be transferred to the required polynomial basis, using Equation (5.109) in Mason and Handscomb (2002).

4.1. Algorithmic aspects

The coefficients $\Lambda_{r,n}$ of the GCE in Theorem 2 can be computed by comparison of coefficients. Notice that

$$M_{2k,n} = \sum_{r=0}^{k} \left(\int_{\mathbb{I}^{n_x}} (xx^\intercal)^{\langle k \rangle} \mathrm{Tr}\left(\Lambda_{r,n}\Psi_{2r}(x)\right) \omega(x)\,\mathrm{d}x \right) \tag{15}$$

yields a block-triangular linear system for $\Lambda_{r,n}$, with respect to the even moments $M_{2r,n}$ of ρ_n—which are known explicitly. Let $\Theta_{k,r} \in \mathbb{R}^{n_x^k \times n_x^k}$ be the coefficients of the monomials $(xx^\intercal)^{\langle k \rangle}$ in the Chebyshev basis,

$$(xx^\intercal)^{\langle k \rangle} = \sum_{r=0}^{k} \mathrm{Sym}_k \left(\Theta_{k,r} \odot \left(\Psi_{2r}(x) \otimes \mathbb{1}^{\langle k-r \rangle}\right)\right) \ , \tag{16}$$

which can be found by a comparison of coefficients. Moreover, we introduce the constant scaling matrices

$$\Omega_r = \int_{\mathbb{I}^{n_x}} \Psi_{2r}(x) \mathrm{Tr}\left(\mathrm{I}_r \Psi_{2r}(x)\right) \omega(x)\,\mathrm{d}x \ .$$

Here, the symbol $\mathrm{I}_r \in \mathbb{R}^{n^r \times n^r}$ denotes a matrix of ones. The matrix Ξ_r denotes the componentwise inverse of the matrix $\Theta_{r,r} \odot \Omega_r$, such that

$$\Xi_r \odot [\Theta_{r,r} \odot \Omega_r] = \mathrm{I}_r \ .$$

The following theorem establishes a recursion for computing the coefficients $\Lambda_{k,n}$.

Theorem 3. *The Chebyshev coefficients $\Lambda_{k,n}$ of ρ_n (as defined in Theorem 2) can be computed by the recursion*

$$\Lambda_{k,n} = \Xi_k \odot \left(M_{2k} - \sum_{r=0}^{k-1} \mathrm{Sym}_k \left(\Theta_{k,r} \odot \left((\Omega_r \odot \Lambda_{r,n}) \otimes \mathbb{1}^{\langle k-r \rangle}\right)\right) \right)$$

for all $k \in \mathbb{N}$.

Proof. Using the orthogonality of the Chebyshev polynomials, Eq. (15) can be written as

$$M_{2k,n} = \sum_{r=0}^{k} \left[\int_{\mathbb{I}^{n_x}} \Theta_{k,r} \odot \mathrm{Sym}_k \left(\Psi_{2r}(x) \otimes \mathbb{1}^{k-r}\right) \mathrm{Tr}\left(\Lambda_{r,n}\Psi_{2r}(x)\right) \omega(x)\,\mathrm{d}x \right].$$

Since the Chebyshev polynomials are orthonormal with respect to (scaled) weighting functions ω, we find

$$\int_{\mathbb{I}^{n_x}} \Psi_{2r}(x) \mathrm{Tr}\left(\Lambda_{r,n}\Psi_{2r}(x)\right) \omega(x)\,\mathrm{d}x = \Omega_r \odot \Lambda_{r,n} \ .$$

Thus, the linear equation system for the coefficients $\Theta_{k,r}$ can be further simplified as

$$M_{2k,n} = \sum_{r=0}^{k} \mathrm{Sym}_k \left(\Theta_{k,r} \odot (\Omega_r \odot \Lambda_{r,n}) \otimes \mathbb{1}^{\langle k-r \rangle}\right) \ .$$

Now, the statement of the theorem follows directly by solving the latter equation with respect to $\Lambda_{k,n}$. □

The algorithm in Fig. 1 summarizes the complete procedure for computing the Chebyshev approximation $\phi_{N,n}$ of the state distribution ρ_n. Notice that this approximation is such that the cumulative distribution error,

$$\left| \Pr(x_n \in \mathbb{X}) - \int_{\mathbb{X}} \phi_{N,n}(x)\,\mathrm{d}x \right| \leq \left\| \phi_{N,n} - \rho_n \right\|_{\mathrm{TV}} \ ,$$

is bounded by the term $\left\| \phi_{N,n} - \rho_n \right\|_{\mathrm{TV}}$. Theorem 2 provides a bound on this numerical approximation error.

Input: System matrices A, B, expansion order $N \in \mathbb{N}$.

Algorithm:

1. Solve the generalized Lyapunov recursion (see Section 3)
$$\forall r \in \{1,2,\ldots,N\}, \quad P_{r,k+1} = \mathrm{Sym}_r\left(A^{\langle r \rangle} P_{r,k} \left(A^{\langle r \rangle}\right)^\intercal + Q_r\right)$$
with $Q_r = \sum_{j=1}^{n_w} B_j^{\langle r \rangle} \left(B_j^{\langle r \rangle}\right)^\intercal$ and $P_{r,0} = 0$ for $k \in \{1,2,\ldots,n\}$.
2. Compute the associated cumulants (see Theorem 1),
$$K_{2r,n} = \frac{(-1)^{r+1}(2r)!\zeta(2r)}{r\pi^{2r}} P_{r,n} \ .$$
3. Use the recursion formula (see Remark 2)
$$M_{2r+2,n} = \mathrm{Sym}_{r+1}\left(\sum_{i=0}^{r} \binom{2r+1}{2i+1} M_{2(r-i),n} \otimes K_{2(i+1),n} \right)$$
with $M_{0,n} = 1$ to compute the associated even moments.
4. Compute the coefficients $\Theta_{k,r}$ by comparing the coefficients of the left- and right hand side polynomials in Eq. (16).
5. Compute the GCE coefficients (see Theorem 3)
$$\Lambda_{k,n} = \Xi_k \odot \left(M_{2k} - \sum_{r=0}^{k-1} \mathrm{Sym}_k \left(\Theta_{k,r} \odot \left((\Omega_r \odot \Lambda_{r,n}) \otimes \mathbb{1}^{\langle k-r \rangle}\right)\right) \right).$$

Output: Coefficients $\Lambda_{r,n}$ of the Chebyshev approximation

$$\rho_n(x) \approx \phi_{N,n}(x) = \left(\sum_{r=0}^{N} \mathrm{Tr}\left(\Lambda_{r,n}\Psi_{2r}(x)\right) \right) \omega(x)$$

of the state distribution ρ_n of Eq. (1) (see Theorem 2).

Fig. 1. Algorithm for approximating the state distribution ρ_n.

5. Constrained stochastic linear control systems

In this section we illustrate the applicability of the theory developed so far, through the construction of optimized feedback gains, $\mathcal{K}$, for

$$x_{k+1} = (A_x + A_u\mathcal{K})\,x + Bw_k \quad \text{with} \quad x_0 = 0 \tag{17}$$

with respect to the objective function

$$\lim_{n\to\infty} \mathbb{E}\left(\|x_n\|_2^2 + \|\mathcal{K}x_n\|_2^2\right) = \int_{\mathbb{I}^{n_x}} \left(\|x\|_2^2 + \|\mathcal{K}x\|_2^2\right) \rho_\infty(x,\mathcal{K})\,\mathrm{d}x \ .$$

We enforce the joint-chance state- and control constraint

$$\lim_{n\to\infty} \Pr\left(x_n \notin \mathbb{X} \ \vee \ \mathcal{K}x_n \notin \mathbb{U}\right) \leq \epsilon$$

for given compact state and control constraint sets, $\mathbb{X} \subseteq \mathbb{R}^{n_x}$ and $\mathbb{U} \subset \mathbb{R}^{n_u}$. The parameter $\epsilon > 0$ is a given bound on the constraint violation probability. Let $X_\infty(\mathcal{K})$ be reachable set of (17) for some $\mathcal{K}$ and

$$\mathbb{S}(\mathcal{K}) = \mathbb{X} \cap \left\{ x \in X_\infty(\mathcal{K}) \,\middle|\, \mathcal{K}x \in \mathbb{U} \right\} \ .$$

The joint-chance constraint can then be written as

$$\int_{\mathbb{S}(\mathcal{K})} \rho_\infty(x,\mathcal{K})\,\mathrm{d}x \geq 1 - \epsilon \ .$$

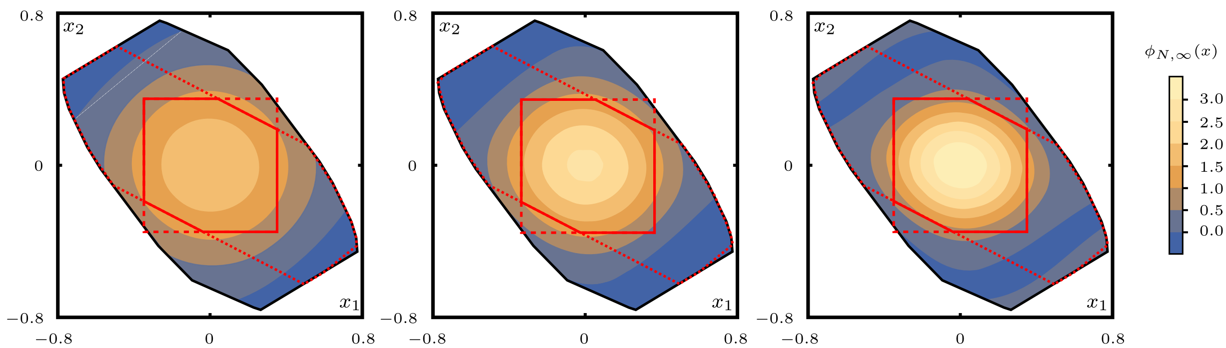

Fig. 2. Density plot of the GCE approximations $\phi_{N,\infty}$ using 4th ($N=2$, left), 8th ($N=4$, center), and 10th ($N=5$, right) moment expansions, respectively. The red dashed line depicts the boundary of the constraint set $\mathbb{X}$. The red dotted line denotes the boundary of the set $\{x \in \overline{X}_\infty(\mathcal{K}^\star) \mid \mathcal{K}^\star x \in \mathbb{U}\}$. The red solid line depicts the boundary of $\mathbb{S}(\mathcal{K}^\star)$.

In summary, the control design problem is given by

$$\begin{aligned} \min_{\mathcal{K}} \ & \int_{\mathbb{I}^{n_x}} \left(\|x\|_2^2 + \|\mathcal{K}x\|_2^2\right) \rho_\infty(x,\mathcal{K})\, \mathrm{d}x \\ \text{s.t.} \ & \int_{\mathbb{S}(\mathcal{K})} \rho_\infty(x,\mathcal{K})\, \mathrm{d}x \geq 1-\epsilon . \end{aligned} \tag{18}$$

Here, one optimizes over stabilizing control gains such that the limit distribution $\rho_\infty(\cdot,\mathcal{K})$ is well defined.

We consider the linear system (17) with

$$A_x = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \quad A_u = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad B = \begin{pmatrix} \frac{3}{20} & \frac{1}{4} \\ -\frac{1}{5} & \frac{3}{20} \end{pmatrix},$$

and $w_k \in [-1,1]^2$ for all $k \in \mathbb{N}_+$. The constraints are

$$\mathbb{X} = [-0.4, 0.4]^2 \quad \text{and} \quad \mathbb{U} = [-0.3, 0.3] .$$

The constraint violation parameter is given by $\epsilon = 0.15$. The feedback gain $\mathcal{K}^\star \approx (-0.42, -0.81)$ was computed by solving a discretization of (18), using the truncated GCE $\phi_{5,\infty}(x,\mathcal{K}^*)$. Fig. 2 shows the contours of the truncated GCEs for $N=2$ (left), $N=4$ (center), and $N=5$ (right). The approximate support $\overline{X}_\infty(\mathcal{K}^\star)$ of $\rho_{N,\infty}$ was computed so as to satisfy

$$X_\infty(\mathcal{K}^\star) \subseteq \overline{X}_\infty(\mathcal{K}^\star) \subseteq X_\infty(\mathcal{K}^\star) \oplus \bar{\epsilon}\, \mathbb{I}^{n_x}$$

with $\bar{\epsilon} = 10^{-4}$ (Rakovic, Kerrigan, Kouramas, & Mayne, 2005).

In order to verify the accuracy of the approximation,

$$\phi_{5,\infty}(x,\mathcal{K}^\star) \approx \rho_\infty(x,\mathcal{K}^\star) ,$$

a Monte-Carlo simulation (2.5×10^6 samples) was implemented. The constraint violation probability computed by Monte-Carlo sampling, $\Pr(x_\infty \notin \mathbb{S}(\mathcal{K}^\star)) \approx 0.14$, must be compared to our approximation

$$= 1 - \int_{\mathbb{S}(\mathcal{K}^\star)} \phi_{5,\infty}(x,\mathcal{K}^\star) = 0.15 .$$

Therefore, for this example, the error associated to the use of a 10th order GCE approximation is approximately 1%. On the other hand, using a second order expansion $\phi_{1,\infty}$ such analysis yields $\Pr(x_\infty \notin \mathbb{S}(\mathcal{K}^\star)) \approx 0.58$ (approximately 44% error). Notice that the assumption of a Gaussian distribution, would yield a similar error in the approximation.

6. Conclusions

This paper proposed a cumulant-based approach used for computing the transient and limit distributions of the state of a stochastic linear system with zonotopic support sets. Explicit expressions for the cumulants of these distributions have been constructed (see Theorem 1) by introducing a novel class of k-symmetric Lyapunov recursions. Moreover, in this paper a generalized Gram–Charlier expansion (GCE) based on Chebyshev polynomials was introduced. This GCE expansion can be used to recover the state distributions from their cumulants under a mild reachability assumption. Theorem 2 provides a uniform convergence result for this GCE approximation. The complete algorithmic procedure has been summarized in Fig. 1, which has been applied to a control synthesis problem for a constrained stochastic linear control system, illustrating the accuracy of the approach.

References

Aubin, J.-P. (1991). *Viability theory*. Birkhäuser Boston.
Berkowitz, S., & Garner, F. (1970). The calculation of multidimensional Hermite polynomials and Gram-Charlier coefficients. *Mathematics of Computation*, *24*(11), 537–545.
Bertsekas, D. P., & Shreve, S. (1978). *Stochastic optimal control: the discrete-time case*. Academic Press.
Bitsoris, G. (1988). On the positive invariance of polyhedral sets for discrete-time systems. *Systems & Control Letters*, *11*(3), 243–248.
Bittanti, S., Colaneri, P., & De Nicolao, G. (1991). The periodic Riccati equation. In S. Bittani, A. J. Laub, & J. C. Willems (Eds.), *The Riccati equation* (pp. 127–162). Springer Verlag.
Blanchini, F. (1999). Set invariance in control. *Automatica*, *35*(11), 1747–1767.
Blanchini, F., & Miani, S. (2008). *Set-theoretic methods in control*. Springer, Birkhäuser.
Boyd, S., El-Ghaoui, L., Feron, E., & Balakrishnan, V. (2004). *Linear matrix inequalities in system and control theory (vol. 15)*. SIAM.
Busam, R., & Freitag, E. (2009). *Complex analysis*. Springer.
Caflisch, R. E. (1998). Monte Carlo and quasi-Monte Carlo methods. *Acta Numerica*, *7*, 1–49.
Caines, P. E. (1988). *Linear stochastic systems (vol. 77)*. John Wiley NYC.
Chachuat, B., Houska, B., Paulen, R., Perić, N., Rajyaguru, J., & Villanueva, M. E. (2015). Set-theoretic approaches in analysis, estimation and control of nonlinear systems. *IFAC-PapersOnLine*, *48*(8), 981–995.
Cohen, L. (1998). Generalization of the Gram-Charlier/Edgeworth series and application to time-frequency analysis. *Multidimensional Systems and Signal Processing*, *9*(4), 363–372.
Dragan, V., Morozan, T., & Stoica, A.-M. (2010). *Mathematical methods in robust control of discrete-time linear stochastic systems*. Springer.
Hald, A. (2000). The early history of the cumulants and the Gram-Charlier series. *International Statistical Review*, *68*(2), 137–153.
Hald, A. (2002). *On the history of series expansions of frequency functions and sampling distributions (vol. 49)* (pp. 1873–1944). Det Kongelige Danske Videnskabernes Selskab.
Holmquist, B. (1985). The direct product permuting matrices. *Linear and Multilinear Algebra*, *17*(2), 117–141.
Kalman, R. E. (1960). A new approach to linear filtering and prediction problems. *Transactions of the ASME-Journal of Basic Engineering*, *82*(1), 35–45.
Kendall, M. G., & Stuart, A. (1969). *The advanced theory of statistics (vol. 3)*. Charles Griffin.
Klartag, B. (2007). A central limit theorem for convex sets. *Inventiones Mathematicae*, *168*(1), 91–131.

Please cite this article as: M.E. Villanueva and B. Houska, On stochastic linear systems with zonotopic support sets. Automatica (2019) 108652, https://doi.org/10.1016/j.automatica.2019.108652.

Kolmanovsky, I., & Gilbert, E. G. (1998). Theory and computation of disturbance invariant sets for discrete-time linear systems. *Mathematical Problems in Engineering: Theory, Method and Applications*, *4*(4), 317–367.

Kurzhanskiy, A. A., & Varaiya, P. (2007). Ellipsoidal techniques for reachability analysis of discrete-time linear systems. *IEEE Transactions on Automatic Control*, *52*(1), 26–38.

Loh, W.-L. (1996). On latin hypercube sampling. *The Annals of Statistics*, *24*(5), 2058–2080.

Mason, J. C., & Handscomb, D. C. (2002). *Chebyshev polynomials*. Chapman & Hall/CRC.

Nagy, Z., & Braatz, R. D. (2007). Distributional uncertainty analysis using power series and polynomial chaos expansions. *Journal of Process Control*, *17*(3), 229–240.

Noschese, S., & Ricci, P. E. (2003). Differentiation of multivariable composite functions and bell polynomials. *Journal of Computational Analysis and Applications*, *5*(3), 333–340.

Rakovic, S. V., Kerrigan, E. C., Kouramas, K. I., & Mayne, D. Q. (2005). Invariant approximations of the minimal robust positively invariant set. *IEEE Transactions on Automatic Control*, *50*(3), 406–410.

Schott, J. R. (2003). Kronecker product permutation matrices and their application to moment matrices of the normal distribution. *Journal of Multivariate Analysis*, *87*(1), 177–190.

Stein, M. (1987). Large sample properties of simulations using latin hypercube sampling. *Technometrics*, *29*(2), 143–151.

Stengel, R. F. (1994). *Optimal control and estimation*. Dover Publications.

Trefethen, L. N. (2013). *Approximation theory and approximation practice (vol. 128)*. SIAM.

Villanueva, M. E., Houska, B., & Chachuat, B. (2015). Unified framework for the propagation of continuous-time enclosures for parametric nonlinear ODEs. *Journal of Global Optimization*, *62*(3), 575–613.

Withers, C. S., & Nadarajah, S. (2014). The dual multivariate Charlier and Edgeworth expansions. *Statistics & Probability Letters*, *87*, 76–85.

Xiu, D. (2010). *Numerical methods for stochastic computations: a spectral method approach*. Princeton University Press.

Xiu, D., & Karniadakis, G. E. (2002). The Wiener–Askey polynomial chaos for stochastic differential equations. *SIAM Journal on Scientific Computing*, *24*(2), 619–644.

Zhang, D. (2002). *Stochastic methods for flow in porous media: coping with uncertainties*. Academic Press.

Mario E. Villanueva is a postdoctoral researcher at the School of Information Science and Technology at ShanghaiTech University. He received a master and Ph.D. in chemical engineering from Imperial College London in 2011 and 2016, respectively. He was a postdoctoral researcher at Texas A&M University in 2016. Mario Villanueva is the recipient of the 2016 Dudley Newitt Price and the 2018 SIST Excellent Postdoc Award. His research interests include set based computing, robust control, and global optimization.

Boris Houska is an assistant professor at the School of Information Science and Technology at ShanghaiTech University. He received a diploma in mathematics from the University of Heidelberg in 2007, and a Ph.D. in Electrical Engineering from KU Leuven in 2011. From 2012 to 2013 he was a postdoctoral researcher at the Centre for Process Systems Engineering at Imperial College London. Boris Houska's research interests include numerical optimization and optimal control, robust and global optimization, as well as fast model predictive control algorithms.

Please cite this article as: M.E. Villanueva and B. Houska, On stochastic linear systems with zonotopic support sets. Automatica (2019) 108652, https://doi.org/10.1016/j.automatica.2019.108652.