Probabilistic reachable and invariant sets for linear systems with correlated disturbance

Mirko Fiacchini [a], Teodoro Alamo [b],

[a] Univ. Grenoble Alpes, CNRS, Grenoble INP, GIPSA-lab, 38000 Grenoble, France

[b] Departamento de Ingeniería de Sistemas y Automática, Universidad de Sevilla, Sevilla 41092, Spain

Abstract

In this paper a constructive method to determine and compute probabilistic reachable and invariant sets for linear discrete-time systems, excited by a stochastic disturbance, is presented. The samples of the disturbance signal are not assumed to be uncorrelated, only bounds on the mean and the covariance matrices are supposed to be known. This allows to consider nonlinear stochastic systems approximations and the effect of nonlinear filters on the disturbance. The correlation bound concept is introduced and employed to determine probabilistic reachable sets and probabilistic invariant sets. Constructive methods for their computation, based on convex optimization, are given.

Key words: Probabilistic sets, Correlated disturbance, Stochastic systems, Predictive control

1 Introduction

Stochastic reachability analysis, aiming at computing or estimating the state evolution of dynamical systems affected by disturbances, gained importance for stochastic control and prediction. A first class of approaches is based on the estimation of the state distribution through polynomial chaos expansions, [1, 3, 20, 24, 30]. The main limitation of these methods is their applicability to low dimensional disturbances, for instance in the case of time-invariant stochastic parametric uncertainties and uncertain initial conditions. A cumulant-based approach is presented in [29] to approximate the state distributions for systems with bounded zonotopic noises. Also methods based on the generalized moment problem [13, 19] have been recently applied to address optimal control problems in presence of stochastic uncertainties in the parameters and the initial conditions, see [22, 23]. Another class of approaches are the sampling-based methods, like Monte Carlo and scenario-based ones, [2, 8, 11, 16, 27], that consist in generating sampled realizations of the possibly correlated uncertainty to infer statistical information of the state evolution or desired property of the trajectories.

⋆ Corresponding author M. Fiacchini. This work has been partially supported by the Agencia Estatal de Investigacin (AEI)-Spain under Grant PID2019-106212RB-C41/AEI/10.13039/501100011033 and by the LabEx PERSYVAL-Lab (ANR-11-LABX-0025-01) funded by the French program Investissement davenir.

Email addresses: mirko.fiacchini@gipsa-lab.fr (Mirko Fiacchini), talamo@us.es (Teodoro Alamo).

The recent interest in the characterization and computation of probabilistic reachable sets and probabilistic invariant sets is also due to the growing popularity of stochastic Model Predictive Control (SMPC), see [21]. Indeed, as in the case of deterministic and robust predictive techniques, several desirable features can be ensured also in the stochastic context by appropriately employing reachable and invariant sets to ensure probabilistic guarantees, for instance, of constraints satisfaction, recursive feasibility and stability properties. The stochastic tube-based approaches, for example, make a wide use of probabilistic invariant or reachable sets to pose deterministic constraints in the nominal prediction such that chance constraints are satisfied, see [9, 15]. Also in [10], probabilistic invariant sets are employed to handle probabilistic state constraints and a method for computing probabilistic invariant ellipsoids is presented.

Concerning the computation of reachable and invariant sets for deterministic systems and for robust control, i.e. in the worst-case disturbance context, several well-established results are present in the literature, for linear [6, 18] and nonlinear systems [12]. In the recent years, some results have been appearing also on probabilistic reachable and invariant sets. The work [17] is completely devoted to the problem of computing probabilistic invariant sets and ultimate bounds for linear systems affected by additive stochastic disturbances. Also the paper [14] presents a characterization of probabilistic sets based on the invariance property in the robust context, whereas [16] employs scenario-based methods to design them.

In most of the works concerning probabilistic reachable and invariant sets computation and SMPC, however, the stochastic disturbance is modelled by an independent sequence of random variables. The assumption of independence in time, and thus uncorrelation, between disturbances, though, is often unrealistic, especially when a linear systems with additive perturbations is employed to model a nonlinear system where the disturbance is modelled by the output of a nonlinear filter. The commonly used approach to get rid of the correlation consists in modelling the correlated disturbance as a white noise filtered by a linear system, i.e. the ARMA model for instance. This approach, though, is not always able to remove the noise correlation, unless the disturbance is effectively given by i.i.d. signals feeding linear filters, which might be not the case in reality. More generally, requiring constant mean and constant covariance matrices of the disturbance, and their exact knowledge, is an often too restrictive assumption, in practice, when dealing with real systems and real data.

In this paper, we consider the problem of characterizing and computing, via convex optimization, outer bounds of probabilistic reachable sets and probabilistic invariant ellipsoids for linear systems excited by disturbances whose realizations are correlated in time. Only bounds on the mean and covariance matrices are required to be known, even stationarity is not necessary. Based on these bounds, the called correlation bound is defined and then employed to determine constructive conditions for computing probabilistic reachable and invariant ellipsoidal sets. The method, resulting in convex optimization problems, is then illustrated through a numerical example, for which the covariance matrices cannot be computed, but bounds exist.

Notation: The set of integers and natural numbers are denoted with $\mathbb{Z}$ and $\mathbb{N}$, respectively. Given $A \in \mathbb{R}^{n\times n}$, $\{\lambda_i(A)\}_{i=1}^n$ denote the n eigenvalues of A; $\rho(A)$ the spectral radius of A; $\sigma_{min}(A) = \min_{i=1,\dots,n} \sqrt{\lambda_i(A^\top A)}$ and $\sigma_{max}(A) = \max_{i=1,\dots,n} \sqrt{\lambda_i(A^\top A)} = \|A\|_2$. The set of symmetric matrices in $\mathbb{R}^{n\times n}$ is denoted $\mathbb{S}^n$. With $\Gamma \succ 0$ ($\Gamma \succeq 0$) it is denoted that Γ is a definite (semi-definite) positive matrix. If $\Gamma \succeq 0$ then $\Gamma^{\frac{1}{2}}$ is the matrix satisfying $\Gamma^{\frac{1}{2}}\Gamma^{\frac{1}{2}} = \Gamma$. For all $\Gamma \succeq 0$ and $r \geq 0$ define $\mathscr{E}(\Gamma, r) = \{x = \Gamma^{1/2}z \in \mathbb{R}^n : z^\top z \leq r\}$; if moreover $\Gamma \succ 0$, then $\mathscr{E}(\Gamma, r) = \{x \in \mathbb{R}^n : x^\top \Gamma^{-1} x \leq r\}$. Given two sets $Y, Z \subseteq \mathbb{R}^n$, their Minkowski set addition is $Y + Z = \{y + z \in \mathbb{R}^n : \; y \in Y, z \in Z\}$, their difference is $Y - Z = \{x \in \mathbb{R}^n : \; x + Z \subseteq Y\}$. The Gaussian (or normal) distribution with mean μ and covariance Σ is denoted $\mathscr{N}(\mu, \Sigma)$, the χ squared cumulative distribution function of order n is denoted $\chi_n^2(x)$. Given a random vector x, $\mathrm{E}\{x\}$ denotes its expected value.

2 Correlation bound

Consider first the nonlinear system $x_{k+1} = f(x_k, d_k)$, where $x_k \in \mathbb{R}^n$ is the state and d_k represents time-varying uncertain parameters and disturbances. A common way of approximating the nonlinear dynamics is by means of a model of the form $x_{k+1} = Ax_k + w_k$ where w_k is an additive terms accounting for the cumulative effects of the modelling errors and the past values of d_k. In this context it is unrealistic to assume that w_k is not correlated with the previous values w_j, with $j \leq k$, especially if j is close to k. Even an assumption on stationarity of w_k is often hardly justifiable since, due to the possibly nonlinear nature of $f(\cdot,\cdot)$, the statistical properties of w_k depend also on the current state x_k and therefore might be time varying. To better deal with these issues, the case of additive uncertainty w_k that is correlated in time and not necessarily stationary is considered.

Consider the discrete-time system

$$x_{k+1} = Ax_k + w_k, \tag{1}$$

where $x_k \in \mathbb{R}^n$ is the state and $w_k \in \mathbb{R}^n$ an additive disturbance given by a sequence of random variables that are supposed to be correlated in time.

In this paper, the only assumptions on the disturbance w_k is that its time-dependent mean is bounded, a bound on $\mathrm{E}\{w_k w_k^\top\}$ exists, and the covariance between w_i and w_j exponentially vanishes with $|j-i|$.

Assumption 1 *There exist $m, b, \gamma \in \mathbb{R}$, with $\gamma \in [0,1)$, such that the sequence w_k satisfies:*

$$\mu_k^\top \mu_k \leq m, \quad \forall k \in \mathbb{N}, \tag{2}$$
$$\|\mathrm{cov}(w_i, w_j)\|_2^2 \leq b\gamma^{j-i}, \quad \forall i \leq j, \tag{3}$$

with $\mathrm{E}\{w_k\} = \mu_k$ *and* $\mathrm{cov}(w_i, w_j) = \mathrm{E}\{(w_i - \mu_i)(w_j - \mu_j)^\top\}$.

Note that no assumption on $\{w_k\}_{k\in\mathbb{N}}$ is posed other than the existence of bounds on the mean and the covariance matrices. Neither weak stationarity is required, as both the mean and the covariance matrices are allowed to be functions of time. This aspect might be crucial in practice, as no exact knowledge of the matrices nor guarantee of stationarity are often available.

Proposition 1 *If Assumption 1 is satisfied, then non-negative $\alpha, \beta, \gamma \in \mathbb{R}$ and $\tilde{\Gamma} \in \mathbb{S}^n$ exist, with $\gamma \in [0,1)$ and $\tilde{\Gamma} \succ 0$, such that*

$$\Gamma_{k,k} \preceq \tilde{\Gamma}, \quad \forall k \in \mathbb{N}, \tag{4}$$
$$\Gamma_{i,j}\tilde{\Gamma}^{-1}\Gamma_{i,j}^\top \preceq (\alpha + \beta\gamma^{j-i})\tilde{\Gamma}, \qquad \forall i \leq j, \tag{5}$$

hold, with $\Gamma_{i,j} = \mathrm{E}\{w_i w_j^\top\}$, for all $i, j \in \mathbb{N}$.

Proof: From Assumption 1, it follows that $\mathrm{cov}(w_k, w_k) \preceq \|\mathrm{cov}(w_k, w_k)\|_2 I \preceq \sqrt{b} I$ and $\mu_k \mu_k^\top \preceq mI$, and then

$$\Gamma_{k,k} = \mathrm{E}\{w_k w_k^\top\} = \mathrm{cov}(w_k, w_k) + \mu_k \mu_k^\top \preceq (\sqrt{b} + m)I,$$

which means that (4) holds with $\tilde{\Gamma} = (\sqrt{b} + m)I$. From Assumption 1, and since $ACB^\top + BCA^\top \preceq ACA^\top + BCB^\top$ for every A, B and C of appropriate dimensions and $C \succ 0$, it follows

$$\begin{aligned}
&\Gamma_{i,j}\tilde{\Gamma}^{-1}\Gamma_{i,j}^\top = \mathrm{cov}(w_i,w_j)\tilde{\Gamma}^{-1}\mathrm{cov}(w_i,w_j)^\top + \mu_i\mu_j^\top\tilde{\Gamma}^{-1}\mu_j\mu_i^\top \\
&+ \mathrm{cov}(w_i,w_j)\tilde{\Gamma}^{-1}\mu_j\mu_i^\top + \mu_i\mu_j^\top\tilde{\Gamma}^{-1}\mathrm{cov}(w_i,w_j)^\top \\
&\preceq 2\Big(\mathrm{cov}(w_i,w_j)\tilde{\Gamma}^{-1}\mathrm{cov}(w_i,w_j)^\top + \mu_i\mu_j^\top\tilde{\Gamma}^{-1}\mu_j\mu_i^\top\Big) \\
&\preceq 2\sigma_{max}(\tilde{\Gamma}^{-1})\Big(\mathrm{cov}(w_i,w_j)\mathrm{cov}(w_i,w_j)^\top + \mu_i\mu_j^\top\mu_j\mu_i^\top\Big) \\
&\preceq 2\sigma_{max}(\tilde{\Gamma}^{-1})\Big(m^2 + b\gamma^{j-i}\Big)I \preceq 2\frac{\sigma_{max}(\tilde{\Gamma}^{-1})}{\sigma_{min}(\tilde{\Gamma})}\Big(m^2 + b\gamma^{j-i}\Big)\tilde{\Gamma} \\
&= 2\sigma_{max}^2(\tilde{\Gamma}^{-1})\Big(m^2 + b\gamma^{j-i}\Big)\tilde{\Gamma}
\end{aligned}$$

since $\sigma_{max}(\tilde{\Gamma}^{-1}) = 1/\sigma_{min}(\tilde{\Gamma})$ from $\tilde{\Gamma} \succ 0$, and then (5) holds with $\alpha = 2\sigma_{max}^2(\tilde{\Gamma}^{-1})m^2$ and $\beta = 2\sigma_{max}^2(\tilde{\Gamma}^{-1})b$. ■

Note that, although the existence of bounds (2) and (3) on the mean and covariance matrices is the only posed assumption, it is not necessary to know them. The results of this paper only require, in fact, the knowledge of bounds (4) and (5), that can be estimated from data.

The following definition of correlation bound encloses the key concept that permits to characterize and compute probabilistic reachable and invariant sets for linear systems affected by correlated disturbance.

Definition 1 (Correlation bound) *The random sequence $\{w_k\}_{k\in\mathbb{Z}}$ is said to have a correlation bound Γ_w for matrix A if the recursion $z_{k+1} = Az_k + w_k$ with $z_0 = 0$, satisfies*

$$A\mathrm{E}\{z_k w_k^\top\} + \mathrm{E}\{w_k z_k^\top\}A^\top + \mathrm{E}\{w_k w_k^\top\} \preceq \Gamma_w, \tag{6}$$

or, equivalently

$$\mathrm{E}\{z_{k+1}z_{k+1}^\top\} \preceq A\mathrm{E}\{z_k z_k^\top\}A^\top + \Gamma_w, \tag{7}$$

for all $k \geq 0$.

It will be proved in the next section that, if the matrix A in (1) is Schur, i.e. $\rho(A) < 1$, and Assumption 1 holds, then a correlation bound exists.

2.1 Computation of a correlation bound

As it will be shown in the subsequent sections, a correlation bound permits to determine sequences of probabilistic reachable sets and probabilistic invariant sets. For this, it is necessary to provide a condition and a method to obtain a correlation bound. Such a condition is presented in the following proposition.

Proposition 2 *Given the system (1) with $\rho(A) < 1$, let $\{w_k\}_{k\in\mathbb{Z}} \in \mathbb{R}^n$ be a random sequence such that conditions (4) and (5) hold with $\tilde{\Gamma} \succ 0$, $\alpha \geq 0$, $\beta \geq 0$ and $\gamma \in (0, 1)$. Given $\eta \in [\rho(A)^2, 1)$, consider $\varphi \geq 1$ and $S \in \mathbb{S}^n$ satisfying*

$$S \preceq \tilde{\Gamma} \preceq \varphi S, \qquad ASA^\top \preceq \eta S. \tag{8}$$

Then for every $p \in (\eta, 1)$, the matrix

$$\Gamma_w = \Big(\alpha\varphi\frac{\eta}{p-\eta} + \beta\varphi\frac{\gamma\eta}{p-\gamma\eta} + \frac{p}{1-p} + 1\Big)\tilde{\Gamma} \tag{9}$$

is a correlation bound for the sequence $\{w_k\}_{k\in\mathbb{Z}}$ and matrix A.

Proof: Note first that φ and S satisfying (8) exist for every $\eta \in [\rho(A)^2, 1)$. From the definition of correlation bound and the equality $z_k = \sum_{i=0}^{k-1} A^{k-1-i}w_i$, matrix Γ_w must satisfy

$$\begin{aligned}
A\mathrm{E}\{(\sum_{i=0}^{k-1} A^{k-1-i}w_i)w_k^\top\} + \mathrm{E}\{w_k(\sum_{i=0}^{k-1} A^{k-1-i}w_i)^\top\}A^\top \\
+\mathrm{E}\{w_k w_k^\top\} \preceq \Gamma_w
\end{aligned}$$

for all $k \in \mathbb{N}$. From condition (5) and

$$\begin{aligned}
0 &\preceq \left(\frac{A^{j-i}\Gamma_{i,j}\tilde{\Gamma}^{-\frac{1}{2}}}{p^{\frac{j-i}{2}}} - p^{\frac{j-i}{2}}\tilde{\Gamma}^{\frac{1}{2}}\right)\left(\frac{A^{j-i}\Gamma_{i,j}\tilde{\Gamma}^{-\frac{1}{2}}}{p^{\frac{j-i}{2}}} - p^{\frac{j-i}{2}}\tilde{\Gamma}^{\frac{1}{2}}\right)^\top \\
&= p^{-(j-i)}A^{j-i}\Gamma_{i,j}\tilde{\Gamma}^{-1}\Gamma_{i,j}^\top(A^{j-i})^\top + p^{j-i}\tilde{\Gamma} - A^{j-i}\Gamma_{i,j} - \Gamma_{i,j}^\top(A^{j-i})^\top
\end{aligned}$$

for every $i, j \in \mathbb{N}$ with $i \leq j$ and $p \neq 0$, it follows that

$$\begin{aligned}
&A^{j-i}\Gamma_{i,j} + \Gamma_{i,j}^\top(A^{j-i})^\top \\
&\preceq (\alpha p^{-(j-i)} + \beta(\gamma p^{-1})^{j-i})A^{j-i}\tilde{\Gamma}(A^{j-i})^\top + p^{j-i}\tilde{\Gamma}.
\end{aligned}$$

Therefore, for every $k \in \mathbb{N}$ it holds

$$\begin{aligned}
&A\mathrm{E}\{(\sum_{i=0}^{k-1} A^{k-1-i}w_i)w_k^\top\} + \mathrm{E}\{w_k(\sum_{i=0}^{k-1} A^{k-1-i}w_i)^\top\}A^\top \\
&+ \mathrm{E}\{w_k w_k^\top\} \preceq \sum_{i=0}^{k-1} A^{k-i}\mathrm{E}\{w_i w_k^\top\} + \sum_{i=0}^{k-1}\mathrm{E}\{w_k w_i^\top\}(A^{k-i})^\top + \tilde{\Gamma} \\
&= \Big(\sum_{i=0}^{k-1} A^{k-i}\Gamma_{i,k} + \Gamma_{i,k}^\top(A^{k-i})^\top\Big) + \tilde{\Gamma} \\
&\preceq \Big(\sum_{i=0}^{k-1}(\alpha p^{-(k-i)} + \beta(\gamma p^{-1})^{k-i})A^{k-i}\tilde{\Gamma}(A^{k-i})^\top + p^{k-i}\tilde{\Gamma}\Big) + \tilde{\Gamma}.
\end{aligned}$$

From (8), it follows that

$$A^j\tilde{\Gamma}(A^j)^\top \preceq \varphi A^j S(A^j)^\top \preceq \varphi\eta^j S \preceq \varphi\eta^j\tilde{\Gamma} \qquad (10)$$

for all $j \in \mathbb{N}$, and then

$$\begin{aligned}
&A\mathrm{E}\{(\sum_{i=0}^{k-1}A^{k-1-i}w_i)w_k^\top\}+\mathrm{E}\{w_k(\sum_{i=0}^{k-1}A^{k-1-i}w_i)^\top\}A^\top+\mathrm{E}\{w_kw_k^\top\}\\
&\preceq \sum_{i=0}^{k-1}\alpha\varphi(\eta p^{-1})^{k-i}\tilde{\Gamma}+\sum_{i=0}^{k-1}\beta\varphi(\gamma\eta p^{-1})^{k-i}\tilde{\Gamma}+\sum_{i=0}^{k-1}p^{k-i}\tilde{\Gamma}+\tilde{\Gamma}\\
&= \Big(\sum_{j=1}^{k}\alpha\varphi(\eta p^{-1})^j+\sum_{j=1}^{k}\beta\varphi(\gamma\eta p^{-1})^j+\sum_{j=1}^{k}p^j\Big)\tilde{\Gamma}+\tilde{\Gamma}\\
&= \Big(\alpha\varphi(\eta p^{-1})\frac{1-(\eta p^{-1})^k}{1-\eta p^{-1}}+\beta\varphi(\gamma\eta p^{-1})\frac{1-(\gamma\eta p^{-1})^k}{1-\gamma\eta p^{-1}}\\
&\quad +p\frac{1-p^k}{1-p}\Big)\tilde{\Gamma}+\tilde{\Gamma}. \qquad (11)
\end{aligned}$$

Two possibilities exist, η can be either positive or zero. If $\eta > 0$ then $0 < \gamma\eta < \eta < p < 1$, and all the terms in the summation in (11) are positive and monotonically increasing with k. If $\eta = 0$ the first two terms in (11) are null and the third one, i.e. $p(1-p^k)/(1-p)$, is positive and monotonically increasing with k, since $0 = \eta < p < 1$. In both cases the supremum is finite and attained for $k \to +\infty$ and then condition (9) implies that Γ_w is a correlation bound for A. ■

Note that, as formally stated in the following corollary, Propositions 1 and 2 imply that, if $\rho(A) < 1$, then Assumption 1 ensures the existence of a correlation bound.

Corollary 1 *If Assumption 1 holds and matrix A in (1) is such that $\rho(A) < 1$, then the random sequence $\{w_k\}_{k\in\mathbb{Z}}$ has a correlation bound for matrix A.*

Proof: The result follows from Propositions 1 and 2. ■

The result of Proposition 2 is used hereafter to design an optimization-based procedure to compute the tightest correlation bound. To obtain the sharpest bound, the parameter multiplying $\tilde{\Gamma}$ in (9) has to be minimized. Note first that such parameter is monotonically increasing with φ and η, for $\varphi \geq 1$ and $\eta \in [\rho(A)^2, 1)$. Nevertheless, the minimizing pair φ and η is not evident, even for a given p, due to the constraint (8). One possibility is to grid the interval $[\rho(A)^2, 1)$ of η and then obtain, for every value of η on the grid, the optimal φ and p. To do so, one should first fix η and then solve the semidefinite programming problem

$$\begin{aligned}
(\varphi^*, S^*) = &\min_{\varphi,S}\ \varphi\\
&\text{s.t.}\quad S \preceq \tilde{\Gamma} \preceq \varphi S\\
&\qquad\ ASA^\top \preceq \eta S.
\end{aligned}$$

Note now that the parameter multiplying $\tilde{\Gamma}$ in (9) is a convex function of p. In fact, $a/(p-a)$ is zero if $a = 0$ and it is finite, convex and decreasing for $p \in (a, +\infty)$ if $a > 0$, whereas $p/(1-p)$ is finite, convex and increasing for $p \in (-\infty, 1)$. Then, the minimum of the function multiplying $\tilde{\Gamma}$ exists and is unique in $(\eta, 1)$. This means that, once φ and η are fixed, the value of p that minimizes the parameter multiplying $\tilde{\Gamma}$ in (9) can be computed by solving the following convex optimization problem in a scalar variable:

$$\begin{aligned}
p^*(\eta,\varphi) = &\min_{p}\ \alpha\varphi\frac{\eta}{p-\eta}+\beta\varphi\frac{\gamma\eta}{p-\gamma\eta}+\frac{p}{1-p}\\
&\text{s.t.}\quad \eta < p < 1.
\end{aligned}$$

Finally, Γ_w can be computed by using in (9) the minimal value of the parameter multiplying $\tilde{\Gamma}$ over the optimal ones obtained for the different η on the grid.

Remark 1 *Note that γ could also be bigger than or equal to 1: this would lead to an (although non realistic) increasingly correlated disturbance. The limit would exist provided that η is smaller than the inverse of γ, for all $p \in (\gamma\eta, 1)$. The case of $\gamma = 1$ is realistic, for instance for the case of constant disturbances, and can modelled by the constant term α.*

The dependence of the bound (9) on the parameter φ can be removed by avoiding using the bound $S \preceq \varphi\tilde{\Gamma}$ as in (10). The corollary below, providing a potentially less conservative correlation bound, follows straightforwardly.

Corollary 2 *Under the hypothesis of Proposition 2, for every $p \in (\eta, 1)$, the matrix*

$$\Gamma_w = \Big(\frac{\alpha\varphi\eta}{p-\eta}+\frac{\beta\varphi\gamma\eta}{p-\gamma\eta}\Big)S+\Big(\frac{p}{1-p}+1\Big)\tilde{\Gamma} \qquad (12)$$

is a correlation bound for matrix A.

Condition (12) provides a further degree of freedom, i.e. the matrix S, that can be used to improve the bound.

3 Probabilistic reachable and invariant sets

Based on the correlation bound, conditions for computing probabilistic reachable and invariant sets are presented. First, two properties are given that are functional to the purpose.

Property 1 *For every $r > 0$ and every $\tilde{\Gamma}, \Sigma \in \mathbb{S}^n$ such that $\tilde{\Gamma} \succeq 0$ and $\Sigma \succ 0$, it holds*

$$\mathscr{E}(A\tilde{\Gamma}A^\top+\Sigma, r) \subseteq A\mathscr{E}(\tilde{\Gamma}, r)+\mathscr{E}(\Sigma, r). \qquad (13)$$

Proof: Notice first that $A\tilde{\Gamma}A^\top+\Sigma \succ 0$ and then

$$\begin{aligned}
\mathscr{E}(A\tilde{\Gamma}A^\top+\Sigma, r) &= \{x \in \mathbb{R}^n :\ x^\top(A\tilde{\Gamma}A^\top+\Sigma)^{-1}x \leq r\}\\
A\mathscr{E}(\tilde{\Gamma}, r)+\mathscr{E}(\Sigma, r) &= \{x = A\tilde{\Gamma}^{1/2}y+\Sigma^{1/2}w \in \mathbb{R}^n :\\
&\qquad y^\top y \leq r,\ w^\top w \leq r\}. \qquad (14)
\end{aligned}$$

For a given $x \in \mathscr{E}(A\tilde{\Gamma}A^\top + \Sigma, r)$, the vectors y and w defined

$$y = \tilde{\Gamma}^{1/2}A^\top(A\tilde{\Gamma}A^\top + \Sigma)^{-1}x, \quad w = \Sigma^{1/2}(A\tilde{\Gamma}A^\top + \Sigma)^{-1}x \tag{15}$$

are such that

$$A\tilde{\Gamma}^{1/2}y + \Sigma^{1/2}w = A\tilde{\Gamma}A^\top(A\tilde{\Gamma}A^\top + \Sigma)^{-1}x + \Sigma(A\tilde{\Gamma}A^\top + \Sigma)^{-1}x = x.$$

Moreover,

$$\begin{aligned} y^\top y &= x^\top(A\tilde{\Gamma}A^\top + \Sigma)^{-1}A\tilde{\Gamma}A^\top(A\tilde{\Gamma}A^\top + \Sigma)^{-1}x \\ &\leq x^\top(A\tilde{\Gamma}A^\top + \Sigma)^{-1}x \leq r \end{aligned}$$

since $A\tilde{\Gamma}A^\top \preceq A\tilde{\Gamma}A^\top + \Sigma$ and $x \in \mathscr{E}(A\tilde{\Gamma}A^\top + \Sigma, r)$. Analogously

$$\begin{aligned} w^\top w &= x^\top(A\tilde{\Gamma}A^\top + \Sigma)^{-1}\Sigma(A\tilde{\Gamma}A^\top + \Sigma)^{-1}x \\ &\leq x^\top(A\tilde{\Gamma}A^\top + \Sigma)^{-1}x \leq r \end{aligned}$$

from $\Sigma \preceq A\tilde{\Gamma}A^\top + \Sigma$. Hence, given $x \in \mathscr{E}(A\tilde{\Gamma}A^\top + \Sigma, r)$, two vectors y and w exist, as defined in (15), such that $x = A\tilde{\Gamma}^{1/2}y + \Sigma^{1/2}w$ and $y^\top y \leq r$ and $w^\top w \leq r$, which means that $x \in A\mathscr{E}(\tilde{\Gamma}, r) + \mathscr{E}(\Sigma, r)$, from (14). Thus (13) is proven. ■

The result in Property 1 is used in the following one to characterize bounds on the covariance matrices and probabilities of the system trajectory.

Property 2 *Suppose that the random sequence $\{w_k\}_{k\in\mathbb{N}}$ has a correlation bound $\Gamma_w \succ 0$ for matrix A with $\rho(A) < 1$. Given $r > 0$, consider the system $z_{k+1} = Az_k + w_k$ with $z_0 = 0$ and the recursion*

$$\Gamma_{k+1} = A\Gamma_k A^\top + \Gamma_w \tag{16}$$

with $\Gamma_0 = 0 \in \mathbb{R}^{n\times n}$. Then,

(i) $\mathrm{E}\{z_k z_k^\top\} \preceq \Gamma_k, \quad \forall k \geq 0,$

(ii) $\Pr\{z_k \in \mathscr{E}(\Gamma_k, r)\} \geq 1 - \frac{n}{r}, \quad \forall k \geq 1,$

(iii) $\mathscr{E}(\Gamma_k, r) \subseteq \mathscr{E}(\Gamma_{k+1}, r) \subseteq A\mathscr{E}(\Gamma_k, r) + \mathscr{E}(\Gamma_w, r), \quad \forall k \geq 1.$

Proof: The claims are proved successively.

(i) Suppose that $\mathrm{E}\{z_k z_k^\top\} \preceq \Gamma_k$ with Γ_k recursively defined through (16). Then

$$\begin{aligned} \mathrm{E}\{z_{k+1}z_{k+1}^\top\} &= \mathrm{E}\{Az_kz_k^\top A^\top + Az_kw_k^\top + w_kz_k^\top A^\top + w_kw_k^\top\} \\ &= A\mathrm{E}\{z_kz_k^\top\}A^\top + A\mathrm{E}\{z_kw_k^\top\} + \mathrm{E}\{w_kz_k^\top\}A^\top + \mathrm{E}\{w_kw_k^\top\} \\ &\preceq A\mathrm{E}\{z_kz_k^\top\}A^\top + \Gamma_w \preceq A\Gamma_kA^\top + \Gamma_w = \Gamma_{k+1}, \end{aligned}$$

where the first inequality follows from the definition of correlation bound.

(ii) This result is based on the Chebyshev inequality, [25,28]. From Markov's inequality, [4, 5], a nonnegative random variable x with expected value μ, satisfies $\Pr\{x > r\} \leq \mu/r$ for all $r > 0$. From $\Gamma_w \succ 0$, it follows that $\Gamma_k \succ 0$ and $\Gamma_k^{-1} \succ 0$ for all $k \geq 1$ and then there exists $D_k \in \mathbb{R}^{n\times n}$ such that $\Gamma_k^{-1} = D_k^\top D_k$ for all $k \geq 1$. Thus

$$\begin{aligned} \mathrm{E}\{z_k^\top\Gamma_k^{-1}z_k\} &= \mathrm{E}\{z_k^\top D_k^\top D_kz_k\} = \mathrm{E}\{\mathrm{tr}\{z_k^\top D_k^\top D_kz_k\}\} \\ &= \mathrm{E}\{\mathrm{tr}\{D_kz_kz_k^\top D_k^\top\}\} = \mathrm{tr}\{D_k\mathrm{E}\{z_kz_k^\top\}D_k^\top\} \\ &\leq \mathrm{tr}\{D_k\Gamma_kD_k^\top\} = \mathrm{tr}\{\Gamma_kD_k^\top D_k\} = \mathrm{tr}\{I\} = n \end{aligned}$$

and then, by applying the Markov's inequality, one gets $\Pr\{z_k^\top\Gamma_k^{-1}z_k > r\} \leq n/r$ and hence $\Pr\{z_k^\top\Gamma_k^{-1}z_k \leq r\} \geq 1 - n/r$, for all $k \geq 1$.

(iii) From the definition of Γ_k, it follows $\Gamma_k = \sum_{i=0}^{k-1} A^i\Gamma_w(A^i)^\top$ for $k \geq 1$ and then

$$\Gamma_{k+1} = A^k\Gamma_w(A^k)^\top + \sum_{i=0}^{k-1} A^i\Gamma_w(A^i)^\top = A^k\Gamma_w(A^k)^\top + \Gamma_k \succeq \Gamma_k.$$

This implies $\Gamma_{k+1}^{-1} \preceq \Gamma_k^{-1}$ and hence, $\mathscr{E}(\Gamma_k, r) \subseteq \mathscr{E}(\Gamma_{k+1}, r)$ for all $k \geq 1$. The inclusion $\mathscr{E}(\Gamma_{k+1}, r) \subseteq A\mathscr{E}(\Gamma_k, r) + \mathscr{E}(\Gamma_w, r)$ follows by applying Property 1 with the definition of Γ_{k+1} as in (16). ■

3.1 Probabilistic reachable sets

The simplest confidence regions are ellipsoids, that have been widely used in the context of MPC, see, for example, [9, 15]. The definition of probabilistic reachable sets is recalled.

Definition 2 (Probabilistic reachable set) *It is said that $\Omega_k \subseteq \mathbb{R}^n$ with $k \in \mathbb{N}$ is a sequence of probabilistic reachable sets for system (1), with violation level $\varepsilon \in [0, 1]$, if $x_0 \in \Omega_0$ implies $\Pr\{x_k \in \Omega_k\} \geq 1 - \varepsilon$ for all $k \geq 1$.*

A condition for a sequence of sets to be a probabilistic reachable sets is presented, in terms of correlation bound. The analogous result for uncorrelated disturbance can be found in [14].

Proposition 3 *Suppose that the random sequence $\{w_k\}_{k\in\mathbb{N}}$ has a correlation bound $\Gamma_w \succ 0$ for matrix A with $\rho(A) < 1$. Given $r > 0$, consider the system (1) and the recursion (16) with $x_0 = 0 \in \mathbb{R}^n$, $\Gamma_0 = 0 \in \mathbb{R}^{n\times n}$. Then the sets defined as*

$$\mathscr{R}_{k+1} = A\mathscr{R}_k + \mathscr{E}(\Gamma_w, r), \tag{17}$$

for all $k \in \mathbb{N}$, and $\mathscr{R}_0 = \{0\}$ are probabilistic reachable sets with violation level n/r for every $r > 0$.

Proof: It will be firstly proved that $\mathscr{E}(\Gamma_k, r) \subseteq \mathscr{R}_k$, for all $k \geq 1$. Note first that $\Gamma_1 = A\Gamma_0A^\top + \Gamma_w = \Gamma_w$ and

$\mathscr{R}_1 = A\mathscr{R}_0 + \mathscr{E}(\Gamma_w, r)$. Thus, $\mathscr{E}(\Gamma_1, r) = \mathscr{E}(\Gamma_w, r) = \mathscr{R}_1$ and hence the claim is satisfied for $k = 1$. It suffice now to prove that $\mathscr{E}(\Gamma_k, r) \subseteq \mathscr{R}_k$ implies $\mathscr{E}(\Gamma_{k+1}, r) \subseteq \mathscr{R}_{k+1}$. Supposing $\mathscr{E}(\Gamma_k, r) \subseteq \mathscr{R}_k$ implies

$$\mathscr{E}(\Gamma_{k+1}, r) \subseteq A\mathscr{E}(\Gamma_k, r) + \mathscr{E}(\Gamma_w, r) \subseteq A\mathscr{R}_k + \mathscr{E}(\Gamma_w, r) = \mathscr{R}_{k+1},$$

where the first inclusion follows from (iii) of Property 2. From this and the second claim of Property 2, it follows

$$\Pr\{x_k \in \mathscr{R}_k\} \geq \Pr\{x_k \in \mathscr{E}(\Gamma_k, r)\} \geq 1 - \frac{n}{r}, \quad (18)$$

which implies that $\mathscr{R}_k$ with $k \in \mathbb{N}$ is a sequence of probability reachable sets with violation level n/r. ■

Note that, from (18) it follows that also sets $\mathscr{E}(\Gamma_k, r)$ are probabilistic reachable sets with violation level n/r, which are less conservative than $\mathscr{R}_k$ and simply determined by iteration (16). If, nonetheless, sets $\mathscr{R}_k$ and $\mathscr{E}(\Gamma_k, r)$ require to be computed for every $k \in \mathbb{N}$, a sequence of reachable sets determined by a unique matrix is given below.

Proposition 4 *Suppose that the random sequence $\{w_k\}_{k\in\mathbb{Z}}$ has a correlation bound $\Gamma_w \succ 0$ for matrix A. If $W \in \mathbb{S}^n$ is such that $W \succ 0$ and*

$$AWA^\top \preceq \lambda^2 W, \quad (19)$$

$$\Gamma_w \preceq (1-\lambda)^2 W, \quad (20)$$

with $\lambda \in [0,1)$, then $\Omega_k = \mathscr{E}(W, r(1-\lambda^k)^2)$ is a sequence of probabilistic reachable sets with violation probability n/r. If, moreover, w_k is a Gaussian process with null mean, then $\mathscr{E}(W, r(1-\lambda^k)^2)$ is a reachable set with violation probability $1 - \chi_n^2(r)$.

Proof: It is first proved by induction that

$$\mathrm{E}\{x_k x_k^\top\} \preceq (1-\lambda^k)^2 W \quad (21)$$

for all $k \in \mathbb{N}$, if $x_0 = 0 \in \Omega_0$. The bound holds for $k = 0$ since $x_0 = 0$. Supposing that (21) holds for $k \in \mathbb{N}$, and from Definition 1, it follows

$$\mathrm{E}\{x_{k+1}x_{k+1}^\top\} \overset{(7)}{\preceq} A\mathrm{E}\{x_k x_k^\top\}A^\top + \Gamma_w \overset{(21)}{\preceq} (1-\lambda^k)^2 AWA^\top + \Gamma_w$$
$$\overset{(19),(20)}{\preceq} ((1-\lambda^k)^2\lambda^2 + (1-\lambda)^2)W \preceq \Big((1-\lambda^{k+1})^2$$
$$+2(1-\lambda)\lambda(\lambda^k - 1)\Big)W \preceq (1-\lambda^{k+1})^2 W,$$

since the last inequality holds for all $\lambda \in [0,1)$. Note, as a consequence, that

$$\mathrm{E}\{x_k x_k^\top\} \preceq \lim_{k\to\infty} \mathrm{E}\{x_k x_k^\top\} \preceq W \quad (22)$$

for all $k \in \mathbb{N}$. From the Chebyshev inequality (see, for example, proof of claim (ii) of Property 2) and (21), it follows

$$\Pr\{x_k \in \mathscr{E}(W, r(1-\lambda^k)^2)\} = \Pr\{x_k^\top W^{-1} x_k \leq r(1-\lambda^k)^2\}$$
$$= \Pr\{x_k^\top \big((1-\lambda^k)^2 W\big)^{-1} x_k \leq r\} \geq 1 - \frac{n}{r}. \quad (23)$$

The results for w_k Gaussian process follow from the definition of the χ squared cumulative distribution, that is $\Pr\{y^\top y \leq r\} = \chi_n^2(r)$ for $y \sim \mathscr{N}(0, I)$ and $r > 0$, see [4,5]. In fact, x_k is a Gaussian random variable, being the finite linear combination of terms of a Gaussian process, that is $x_k \sim \mathscr{N}(0, X_k)$ with $X_k \succeq 0$. Denoting the rank of X_k as q, there exist $M_k \in \mathbb{R}^{n\times q}$ and a random variable $y_k \in \mathbb{R}^q$ such that $M_k y_k = x_k$ and $y_k \sim \mathscr{N}(0, I)$, for all $k \in \mathbb{N}$. From $X_k = \mathrm{E}\{x_k x_k^\top\} = M_k M_k^\top \preceq (1-\lambda^k)^2 W$, that is equivalent to $M_k^\top((1-\lambda^k)^2 W)^{-1} M_k \preceq I$ from Schur complement, it follows that $x_k^\top((1-\lambda^k)^2 W)^{-1} x_k \leq y_k^\top y_k$. Hence

$$\Pr\{x_k \in \mathscr{E}(W, r(1-\lambda^k)^2)\} = \Pr\{x_k^\top \big((1-\lambda^k)^2 W\big)^{-1} x_k \leq r\}$$
$$\geq \Pr\{y_k^\top y_k \leq r\} = \chi_q^2(r) \geq \chi_n^2(r),$$

and then $\Pr\{x_k \in \mathscr{E}(W, r(1-\lambda^k)^2)\} \geq 1 - (1 - \chi_n^2(r))$. ■

Notice that, for every $\lambda \in [\rho(A), 1)$, the convex conditions (19) and (20) admit solutions and, for any matrix W satisfying them, the sets $\mathscr{E}(W, r(1-\lambda^k)^2)$ form a sequence of probabilistic reachable sets with violation probability n/r or $(1 - \chi_n^2(r))$, in the Gaussian process case. Thus, condition (19) and (20) can be used in a convex optimization problem aiming at maximizing or minimizing a measure of the reachable sets, their volume for instance.

3.2 Probabilistic invariant sets

The concept of probabilistic invariant sets, as defined and used in [14,17], is recalled.

Definition 3 (Probabilistic invariant set) *The set $\Omega \subseteq \mathbb{R}^n$ is a probabilistic invariant set for the system (1), with violation level $\varepsilon \in [0,1]$, if $x_0 \in \Omega$ implies $\Pr\{x_k \in \Omega\} \geq 1 - \varepsilon$ for all $k \geq 1$.*

A first condition for a set to be probabilistic invariant, analogous to that proved in [14] for uncorrelated disturbances, is given below.

Property 3 *Suppose that the random sequence $\{w_k\}_{k\in\mathbb{N}}$ has a correlation bound $\Gamma_w \succ 0$ for matrix A. If $W \in \mathbb{S}^n$ and $r > 0$ are such that $W \succ 0$ and*

$$A\mathscr{E}(W, 1) + \mathscr{E}(\Gamma_w, r) \subseteq \mathscr{E}(W, 1), \quad (24)$$

then $\mathscr{E}(W, 1)$ is a probabilistic invariant set with violation probability n/r. If, moreover, w_k is a Gaussian process with null mean, then $\mathscr{E}(W, 1)$ is a probabilistic invariant set with violation probability $1 - \chi_n^2(r)$.

Proof: By definition, it is sufficient to show that $x_0 \in \mathscr{E}(W,1)$ implies $\Pr\{x_k \in \mathscr{E}(W,1)\} \geq 1-n/r$, for all $k \geq 0$. The state x_k can be written as the sum of a nominal term $\bar{x}_k$ and a random vector z_k that depends on the past realizations of the uncertainty. That is, $x_k = \bar{x}_k + z_k$, where $\{\bar{x}_k\}_{k\geq 0}$ and $\{z_k\}_{k\geq 0}$ are given by the recursions

$$\bar{x}_{k+1} = A\bar{x}_k, \qquad z_{k+1} = Az_k + w_k, \tag{25}$$

for all $k \geq 0$, with $\bar{x}_0 = x_0$ and $z_0 = 0$. Below it is first proved that $x_0 \in \mathscr{E}(W,1)$ and (24) imply

$$\bar{x}_k + \mathscr{R}_k \subseteq \mathscr{E}(W,1), \ \forall k \geq 0, \tag{26}$$

with $\mathscr{R}_k$ as in (17). Since $\mathscr{R}_0 = \{0\}$, the inclusion is trivially satisfied for $k=0$. Supposing that $\bar{x}_k + \mathscr{R}_k \subseteq \mathscr{E}(W,1)$ yields

$$\begin{aligned}\bar{x}_{k+1} + \mathscr{R}_{k+1} &= A\bar{x}_k + (A\mathscr{R}_k + \mathscr{E}(\Gamma_w, r))\\ &= A\left(\bar{x}_k + \mathscr{R}_k\right) + \mathscr{E}(\Gamma_w, r)\\ &\subseteq A\mathscr{E}(W,1) + \mathscr{E}(\Gamma_w, r) \subseteq \mathscr{E}(W,1),\end{aligned}$$

and then (26) holds. Condition (26) implies

$$\Pr\{x_k \in \mathscr{E}(W,1)\} = \Pr\{\bar{x}_k + z_k \in \mathscr{E}(W,1)\} \geq \Pr\{z_k \in \mathscr{R}_k\} \geq 1 - \frac{n}{r}$$

for all $k \geq 0$, where the last inequality follows from Proposition 3. The case of Gaussian process follows from the definition of the χ squared cumulative distribution, see also the proof of Proposition 4. ∎

Property 3 implies that the existence of a correlation bound provides a condition for probabilistic invariance that has the same structure as the one corresponding to robust invariance. In the case of ellipsoidal invariant sets, (24) results in a bilinear condition, see [7], that can be solved, for instance, by gridding the space of the Lagrange multiplier and solving an LMI for every value as illustrated below. Nevertheless, as shown afterward, gridding can be avoided by choosing the multiplier in $[\rho(A), 1)$.

Proposition 5 *Suppose that the random sequence $\{w_k\}_{k\in\mathbb{Z}}$ has a correlation bound $\Gamma_w \succ 0$ for matrix A. If $W \in \mathbb{S}^n$ is such that $W \succ 0$ and*

$$\begin{bmatrix} A^\top W^{-1}A - \tau W^{-1} & A^\top W^{-1} \\ W^{-1}A & W^{-1} - (1-\tau)\Gamma_w^{-1}/r \end{bmatrix} \preceq 0 \tag{27}$$

with $\tau \in [0,1)$, then $\mathscr{E}(W,1)$ is a probabilistic invariant set with violation probability n/r. If, moreover, w_k is a Gaussian process with null mean, then $\mathscr{E}(W,1)$ is a probabilistic invariant set with violation probability $1-\chi_n^2(r)$.

Proof: Because of Property 3, it suffices to prove that (27) is equivalent to (24). From Theorem 4.2 in [26], see also [14], condition (24) is equivalent to the existence of non-negative τ_1 and τ_2 such that

$$\begin{bmatrix} A^\top W^{-1}A - \tau_1 W^{-1} & A^\top W^{-1} \\ W^{-1}A & W^{-1} - \tau_2\Gamma_w^{-1}/r \end{bmatrix} \preceq 0 \tag{28}$$

and $1-\tau_1-\tau_2 \geq 0$ hold, the latter implying $\tau_1 \in [0,1]$, $\tau_2 \in [0,1]$ and $\tau_1+\tau_2 \leq 1$. Note first that there is no conservatism in posing $\tau_2 = 1-\tau_1$ in spite of $\tau_2 \leq 1-\tau_1$ since if (28) holds for $\tau_2 < 1-\tau_1$, then it holds also for $\tau_2 = 1-\tau_1$. This implies that posing $\tau_1 = \tau$ and $\tau_2 = 1-\tau$ introduces no conservatism. Moreover, since $W \succ 0$ then τ cannot be 1. Hence (24) is equivalent to (27) with $\tau \in [0,1)$. ∎

Although (27) is a non-convex condition, that can be solved with respect to W^{-1} by gridding τ in $[0,1)$, this can be avoided by choosing $\tau \in [\rho(A), 1)$, as proved below.

Property 4 *Suppose that the random sequence $\{w_k\}_{k\in\mathbb{N}}$ has a correlation bound $\Gamma_w \succ 0$ for matrix A. Condition (27) admits a solution W for every $\tau \in [\rho(A), 1)$.*

Proof: From $\tau \in [\rho(A), 1)$, there exits $W \in \mathbb{S}^n$ such that $(1-\tau)^{-2}\Gamma_w r \preceq W$ and $AWA^\top \preceq \tau^2 W$ hold, and then:

$$\begin{aligned}(1-\tau)AWA^\top + \tau\Gamma_w r &\preceq \big((1-\tau)\tau^2 + \tau(1-\tau)^2\big)W\\ &= \tau(1-\tau)W.\end{aligned}$$

This implies that $\tau W - AWA^\top \succ 0$ and also

$$\Big(W - AWA^\top/\tau\Big)^{-1} \preceq (1-\tau)\Gamma_w^{-1}/r$$

that is equivalent, from the inversion lemma, to

$$\begin{aligned}&-(1-\tau)\Gamma_w^{-1}/r + \big(W^{-1}\\ &+ (W^{-1}A(\tau W^{-1} - A^\top W^{-1}A)^{-1}A^\top W^{-1}\big) \preceq 0\end{aligned}$$

and then also to (27) from the Schur complement. ∎

From Property 4 it follows that for every $\tau \in [\rho(A), 1)$, the set of matrices W satisfying condition (27), convex in W^{-1}, is non-empty. Moreover, any W in this set provides the probabilistic invariant set $\mathscr{E}(W,1)$ with violation level n/r or $1-\chi_n^2(r)$, in the Gaussian process case. The constraint $\tau \in [\rho(A), 1)$ restricts, though, the set of feasible solutions and then, if one aims at obtaining the minimal probabilistic invariant ellipsoids, gridding τ in $[0, 1)$ might be necessary.

4 Numerical examples

Consider the system (1) with

$$A = \begin{bmatrix} 0.25 & 0 \\ 0.1 & 0.3 \end{bmatrix}.$$

To validate the presented results, it is necessary to generate a random sequence satisfying the bounds (4) and (5). In particular, an example is given for which the value of the covariance matrices cannot be computed, but bounds of the type (4) and (5) can be determined.

Consider the i.i.d. random sequence v_k with Gaussian distribution $\mathcal{N}(0,V)$, for all $k \in \mathbb{N}$, and the switched system with $m \in \mathbb{N}$ modes

$$w_{k+1} = H_{\sigma_k} w_k + F v_k \tag{29}$$

where $\sigma : \mathbb{N} \to \mathbb{N}_m$ is the mode selection signal, assumed arbitrary. Note that $\{w_k\}_{k\in\mathbb{N}}$ is a Gaussian process, since every linear combination of its terms has Gaussian distribution, being a linear combination of elements of v_k, that are i.i.d. with Gaussian distribution. Moreover $\{w_k\}_{k\in\mathbb{N}}$ has null mean since v_k has null mean.

Denote with w_k the state given by (29) with $w_0 = 0$ and switching sequence σ (the dependence of w_k on σ is left implicit); with $\sigma_{[i,j]}$ the subsequence of modes given by the realization of σ from instants i and j with $i < j$, and define $\mathbb{H}_{\sigma_{[i,j]}} = \prod_{k=i}^{j} H_{\sigma_k}$. Suppose there exist $\Gamma \succ 0$ and $\gamma \in [0,1)$ such that

$$H_i \Gamma H_i^\top + FVF^\top \preceq \Gamma, \quad \forall i \in \mathbb{N}_m, \tag{30}$$

$$H_i \Gamma H_i^\top \preceq \gamma \Gamma, \quad \forall i \in \mathbb{N}_m. \tag{31}$$

It can be recursively proved that $\mathrm{E}\{w_k w_k^\top\} \preceq \Gamma$. In fact, the condition holds for $k=0$, from $w_0 = 0$. Suppose that $\mathrm{E}\{w_k w_k^\top\} \preceq \Gamma$ holds for a given $k \in \mathbb{N}$ and since $\mathrm{E}\{v_k w_k^\top\} = 0$, then

$$\mathrm{E}\{w_{k+1} w_{k+1}^\top\} = \mathrm{E}\{H_{\sigma_k} w_k w_k^\top H_{\sigma_k}^\top + F v_k w_k^\top H_{\sigma_k}^\top + H_{\sigma_k} w_k v_k^\top F^\top + F v_k v_k^\top F^\top\} = H_{\sigma_k} \mathrm{E}\{w_k w_k^\top\} H_{\sigma_k}^\top + F \mathrm{E}\{v_k v_k^\top\} F^\top \preceq H_{\sigma_k} \Gamma H_{\sigma_k}^\top + FVF^\top \overset{(30)}{\preceq} \Gamma,$$

for every $\sigma_k \in \mathbb{N}_m$, which means that $\mathrm{E}\{w_{k+1} w_{k+1}^\top\} \preceq \Gamma$. For every $i, j \in \mathbb{N}$ with $i \neq j$, define $\Gamma_{i,j}^{(\sigma)} = \mathrm{E}\{w_i w_j^\top\}$ and note that

$$\Gamma_{k+1,k}^{(\sigma)} = \mathrm{E}\{w_{k+1} w_k^\top\} = \mathrm{E}\{(H_{\sigma_k} w_k + F v_k) w_k^\top\} = \mathrm{E}\{H_{\sigma_k} w_k w_k^\top\} + \mathrm{E}\{F v_k w_k^\top\} = H_{\sigma_k} \mathrm{E}\{w_k w_k^\top\}$$

for all $k \in \mathbb{N}$, and then, from $\mathrm{E}\{w_k w_k^\top\} \preceq \Gamma$, it follows

$$\Gamma_{k+1,k}^{(\sigma)} \Gamma^{-1} \Gamma_{k+1,k}^{(\sigma)\top} = H_{\sigma_k} \mathrm{E}\{w_k w_k^\top\} \Gamma^{-1} \mathrm{E}\{w_k w_k^\top\} H_{\sigma_k}^\top \preceq H_{\sigma_k} \mathrm{E}\{w_k w_k^\top\} H_{\sigma_k}^\top \preceq H_{\sigma_k} \Gamma H_{\sigma_k}^\top \overset{(31)}{\preceq} \gamma \Gamma$$

for every $\sigma_k \in \mathbb{N}_m$. Hence

$$\Gamma_{k+1,k}^{(\sigma_k)} \Gamma^{-1} \Gamma_{k+1,k}^{(\sigma_k)\top} \preceq \gamma \Gamma, \quad \forall \sigma_k \in \mathbb{N}_m.$$

Following analogous considerations it can be proved that

$$\Gamma_{j,i}^{(\sigma)} \Gamma^{-1} \Gamma_{j,i}^{(\sigma)\top} \preceq \mathbb{H}_{\sigma_{[i,j]}} \Gamma \mathbb{H}_{\sigma_{[i,j]}}^\top \preceq \gamma^{j-i} \Gamma, \quad \forall \sigma_{[i,j]} \in \mathbb{N}_m^{j-i}$$

for all $i, j \in \mathbb{N}$ such that $i < j$. Thus conditions (4) and (5) hold with $\tilde{\Gamma} = \Gamma$ and γ solution of (30)-(31), $\Gamma_{j,i} = \Gamma_{j,i}^{(\sigma)}$, $\alpha = 0$ and $\beta = 1$. Note that these bounds hold for every possible realization of the switching sequence σ_k.

An i.i.d. random sequence with distribution $\mathcal{N}(0,V)$, with $V = \mathrm{diag}(1.5, 0.26)$, has been used to feed system (29) with

$$H_1 = \begin{bmatrix} 0.17 & 0.02 \\ 0.07 & 0.14 \end{bmatrix}, H_2 = \begin{bmatrix} 0.15 & 0.025 \\ 0.1 & -0.25 \end{bmatrix}, F = \begin{bmatrix} 0.25 & 0.025 \\ 0.1 & -0.35 \end{bmatrix}$$

and σ_k unknown function of time with value in $\{1,2\}$. The switched system generates a Gaussian process w_k with null mean satisfying the covariance matrix bounds (4) and (5) and the correlation bound (7), with

$$\tilde{\Gamma} = \begin{bmatrix} 0.0098 & 0.0018 \\ 0.0018 & 0.0343 \end{bmatrix}, \quad \Gamma_w = \begin{bmatrix} 0.0113 & 0.0020 \\ 0.0020 & 0.0397 \end{bmatrix}$$

$\alpha = 0$, $\beta = 1$ and $\gamma = 0.0395$, and Γ_w computed using (12).

Different values of violation probability p_v have been tested, in particular $p_v = 0.1, 0.2, 0.3, 0.4, 0.5$. For every p_v, the values of r such that $\chi_2^2(r) = 1 - p_v$ has been determined and the matrix W solving (27) with minimal trace has been computed to obtain $\mathcal{E}(W,1)$ probabilistic invariant. Then, for every p_v, $N = 1000$ initial states x_0 have been uniformly generated on the boundary of $\mathcal{E}(W,1)$ and assumed independent on w_k. For each x_0, a sequence w_k has been generated through (29) and applied. For every $k = 1, \dots, 100$, the set of states x_k and the number of violation d_k of the constraint $x_k \in \mathcal{E}(W,1)$ have been computed. The frequencies of violation d_k/N, for every p_v and $k = 1, \dots, 100$, are depicted in Fig. 1, that shows that the bound is always satisfied.

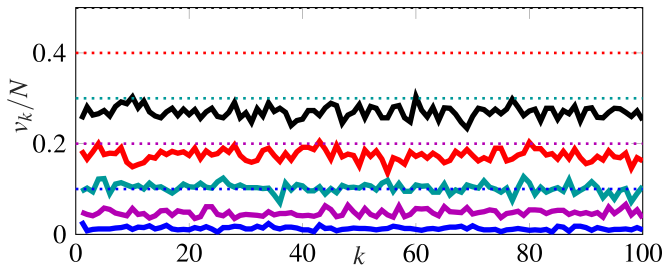

Fig. 1. Frequency of violations d_k/N of $x_k \in \mathcal{E}(W,1)$ for $k = 1, \dots, 100$, with $\alpha = 0$ and $\beta = 1$, obtained for violation probability of: 50% in black; 40% in red; 30% in cyan; 20% in magenta; 10% in blue.

5 Conclusions

This paper presented methods, based on convex optimization, to compute probabilistic reachable and invariant sets for linear systems fed by a stochastic disturbance correlated in time. From the knowledge of bounds on the mean and the covariance matrices, the characterization of the correlation bound is given and then employed for obtaining the reachable and invariant sets.

References

[1] Eva Ahbe, Andrea Iannelli, and Roy S. Smith. Region of attraction analysis of nonlinear stochastic systems using polynomial chaos expansion. Automatica, 122:109187, 2020.

[2] Teodoro Alamo, Roberto Tempo, Amalia Luque, and Daniel R. Ramirez. Randomized methods for design of uncertain systems: Sample complexity and sequential algorithms. Automatica, 52:160–172, 2015.

[3] Lilli Bergner and Christian Kirches. The polynomial chaos approach for reachable set propagation with application to chance-constrained nonlinear optimal control under parametric uncertainties. Optimal Control Applications and Methods, 39(2):471–488, 2018.

[4] Dimitri P. Bertsekas and John N. Tsitsiklis. Introduction to probability. Athena Scientific, 2008.

[5] Patrick Billingsley. Probability and measure. John Wiley & Sons, 2008.

[6] Franco Blanchini and Stefano Miani. Set-theoretic methods in control. Springer, 2008.

[7] Stephen Boyd, Laurent El Ghaoui, Eric Feron, and Venkataramanan Balakrishnan. Linear matrix inequalities in system and control theory, volume 15. Siam, 1994.

[8] Russel E. Caflisch. Monte Carlo and quasi-Monte Carlo methods. Acta numerica, 1998:1–49, 1998.

[9] Mark Cannon, Basil Kouvaritakis, Saša V. Rakovic, and Qifeng Cheng. Stochastic tubes in model predictive control with probabilistic constraints. IEEE Transactions on Automatic Control, 56(1):194–200, 2011.

[10] Mark Cannon, Basil Kouvaritakis, and Xingjian Wu. Probabilistic constrained MPC for multiplicative and additive stochastic uncertainty. IEEE Transactions on Automatic Control, 54(7):1626–1632, 2009.

[11] Alex Devonport and Murat Arcak. Estimating reachable sets with scenario optimization. In Learning for Dynamics and Control (L4DC), 2020.

[12] Mirko Fiacchini, Teodoro Alamo, and Eduardo F. Camacho. On the computation of convex robust control invariant sets for nonlinear systems. Automatica, 46(8):1334–1338, 2010.

[13] Didier Henrion, Milan Korda, and Jean Bernard Lasserre. Moment-SOS Hierarchy, The: Lectures In Probability, Statistics, Computational Geometry, Control And Nonlinear Pdes, volume 4. World Scientific, 2020.

[14] Lukas Hewing, Andrea Carron, Kim P. Wabersich, and Melanie N. Zeilinger. On a correspondence between probabilistic and robust invariant sets for linear systems. In 2018 European Control Conference (ECC), pages 1642–1647. IEEE, 2018.

[15] Lukas Hewing and Melanie N. Zeilinger. Stochastic model predictive control for linear systems using probabilistic reachable sets. In 2018 IEEE Conference on Decision and Control (CDC), pages 5182–5188, 2018.

[16] Lukas Hewing and Melanie N. Zeilinger. Scenario-based probabilistic reachable sets for recursively feasible stochastic model predictive control. IEEE Control Systems Letters, 4(2):450–455, 2019.

[17] Ernesto Kofman, José A. De Doná, and Maria M. Seron. Probabilistic set invariance and ultimate boundedness. Automatica, 48(10):2670–2676, 2012.

[18] Ilya Kolmanovsky and Elmer G. Gilbert. Theory and computation of disturbance invariant sets for discrete-time linear systems. Mathematical problems in engineering, 4(4):317–367, 1998.

[19] Jean-Bernard Lasserre. Moments, positive polynomials and their applications, volume 1. World Scientific, 2010.

[20] Heng Li and Dongxiao Zhang. Probabilistic collocation method for flow in porous media: Comparisons with other stochastic methods. Water Resources Research, 43(9), 2007.

[21] Ali Mesbah. Stochastic model predictive control: An overview and perspectives for future research. IEEE Control Systems Magazine, 36(6):30–44, 2016.

[22] Kaouther Moussa, Mirko Fiacchini, and Mazen Alamir. Robust optimal control-based design of combined chemo-and immunotherapy delivery profiles. In The 8th IFAC Conference on Foundations of Systems Biology in Engineering. IEEE, 2019.

[23] Kaouther Moussa, Mirko Fiacchini, and Mazen Alamir. Robust optimal scheduling of combined chemo-and immunotherapy: Considerations on chemotherapy detrimental effects. In 2020 American Control Conference (ACC), pages 4252–4257. IEEE, 2020.

[24] Zoltan K. Nagy and Richard D. Braatz. Distributional uncertainty analysis using power series and polynomial chaos expansions. Journal of Process Control, 17(3):229–240, 2007.

[25] Jorge Navarro. A very simple proof of the multivariate Chebyshev's inequality. Communications in Statistics-Theory and Methods, 45(12):3458–3463, 2016.

[26] Imre Pólik and Tamás Terlaky. A survey of the S-lemma. SIAM review, 49(3):371–418, 2007.

[27] Hossein Sartipizadeh, Abraham P. Vinod, Behçet Açikmeşe, and Meeko Oishi. Voronoi partition-based scenario reduction for fast sampling-based stochastic reachability computation of linear systems. In 2019 American Control Conference (ACC), pages 37–44. IEEE, 2019.

[28] Bartolomeo Stellato, Bart P. G. Van Parys, and Paul J. Goulart. Multivariate Chebyshev inequality with estimated mean and variance. The American Statistician, 71(2):123–127, 2017.

[29] Mario E. Villanueva and Boris Houska. On stochastic linear systems with zonotopic support sets. Automatica, 111:108652, 2020.

[30] Dongbin Xiu. Fast numerical methods for stochastic computations: a review. Communications in computational physics, 5(2-4):242–272, 2009.