Probabilistic Set Invariance and Ultimate Boundedness

Ernesto Kofman [a], José A. De Doná [b], Maria M. Seron [b]

[a] *Departamento de Control, FCEIA, UNR and CIFASIS - CONICET, Argentina.*

[b] *School of Electrical Engineering and Computer Science and CDSC, The University of Newcastle, Australia.*

Abstract

The notions of invariant sets and ultimate bounds are important concepts in the analysis of dynamical systems and very useful tools for the design of control systems. Several approaches have been reported for the characterisation of these sets, including constructive methods for their computation and procedures to obtain different approximations. However, there are shortcomings in those concepts, in the sense that no general probability distributions can be considered for the disturbances affecting the system (which, for example, precludes the assumption of Gaussian distributions insofar as they are not bounded). Motivated by those shortcomings, we propose in this paper the novel concepts of probabilistic ultimate bounds and probabilistic invariant sets, which extend the notions of invariant sets and ultimate bounds to consider 'containment in probability', and have the important feature of allowing stochastic noises with more general distributions, including the ubiquitous Gaussian distribution, to be considered. We introduce some key definitions for these sets, establish their main properties and develop methods for their computation. A numerical example illustrates the main ideas.

Key words: Invariant sets; Ultimate bounds; Linear systems; Probabilistic methods

1 Introduction

The notions of *invariant sets* and *ultimate bounds* have proven to be important concepts in the analysis of the behaviour of dynamical systems as well as very useful tools for the design of control systems. For some basic definitions of these concepts, including a historic perspective, the reader is referred to the classical references on the topic, for example, the survey paper [2], and the references therein, and the more recent monograph [3]. The characterisation of ultimate bounds and invariant sets, constructive methods for their computation, and procedures to obtain different approximations, can be found in a number of references, including [1,8,9,12].

Paramount to the existence and compactness of these sets, is that the disturbances affecting the system be *bounded*. Take for example a typical linear system model, $x(t+1) = Ax(t) + w(t)$, with matrix A strictly stable. Obviously, no matter how much 'shrinking' effect the term $Ax(t)$ has, if the disturbance factor $w(t)$ is not bounded, then nothing can be said about the boundedness of the state $x(t+1)$ at the next time evolution of the system. Clearly, this constitutes a shortcoming of the notions of set invariance and ultimate bounds, in the sense that no general probability distributions can be considered for the disturbances affecting the system, including Gaussian distributions—insofar as they are not bounded. Motivated by the shortcoming of invariant sets and ultimate bounds mentioned above, in the sense that no general disturbance distributions can be considered, we propose in this paper two novel concepts; namely, *probabilistic ultimate bounds* (PUB) and *probabilistic invariant sets* (PIS) (that is, sets wherein relevant variables can be assured to remain *with a given probability*) and investigate their properties. These sets extend the concepts of invariant sets and ultimate bounds (which are inherently *deterministic* concepts), to consider 'containment in probability', and have the important feature of allowing stochastic noises with more general distributions, including the ubiquitous Gaussian distribution, to be considered.

In this paper we introduce some key definitions, establish the main properties and develop methods for the computation of these sets. A key factor that motivated the developments of this paper was to obtain *explicit characterisations* of the introduced sets that are conceptually simple and computationally efficient. A numerical exam-

⋆ This paper was not presented at any IFAC meeting. Corresponding author E. Kofman. Tel. +54-341-4237248 ext 305.

Email addresses: kofman@fceia.unr.edu.ar (Ernesto Kofman), Jose.Dedona@newcastle.edu.au (José A. De Doná), Maria.Seron@newcastle.edu.au (Maria M. Seron).

ple is provided to illustrate the main ideas. A preliminary version of this work has been reported in the conference paper [7], in which only Gaussian distributions were considered for the system disturbances. The present paper extends that work by generalising the treatment to arbitrary probability distribution functions. For completeness, the Gaussian case—resulting in tighter bounds—is presented here as well, as a particular case of the general result.

It is envisaged that an important number of applications and constructs could be developed utilising these newly introduced concepts, allowing, for example, the extension of existing methods to cases where currently they cannot be applied due to the unboundedness of the disturbances. We discuss briefly below some related literature where the concepts and characterisations introduced in this paper could be useful and/or complementary. The survey paper [4] provides a broad perspective of Probabilistic Robust Control, focusing on control design methods based on the interplay between uncertainty randomisation and convex optimisation. Stochastic developments in Model Predictive Control include, for example, the work in [5], where bounds are imposed on the probability of violation of the state constraints. In [11], the problem of finding the optimal control policy that maximises the probability of the state to belong to a given sequence of sets within a finite time horizon is considered (an important difference with the work reported here is that [11] deals with *feedback policies*, whereas we are concerned with the problem of finding invariant sets for *open-loop* dynamics; in addition, no explicit characterisation of the sets is given in [11] that would allow further comparisons). In the rather different setting of systems described by continuous-time stochastic differential equations, the problem of Stochastic Input to State Stability has been considered in, e.g., [6,10]. Suitable continuous-time extensions of the notions of invariance introduced here could be useful in this context.

2 Definitions and Preliminary Background

We consider a discrete time LTI system of the form

$$x(t+1) = Ax(t) + w(t) \tag{1}$$

with $x(t), w(t) \in \mathbb{R}^n$. The disturbance vector $w(t)$ is formed by n independent noise processes, where each of its components, $w_i(t)$, is an i.i.d. zero-mean white random sequence. (Note, in particular, that this implies that the covariance matrix of $w(t)$ is a constant diagonal matrix for all times t.) We assume that the nominal system is asymptotically stable.

2.1 Expected Value and Covariance of $x(t)$

We will review here some basic facts that will be used throughout the paper. We will denote with cov the operation that gives the *covariance matrix* of a random vector, and with var the operation that gives the *variance* of a random scalar variable. Let us define the following covariance matrices:

$$\Sigma_w \triangleq \mathrm{cov}[w(t)] = \mathrm{E}[w(t)w^T(t)]$$

and

$$\Sigma_x(t) \triangleq \mathrm{cov}[x(t)] = \mathrm{E}[(x(t) - E[x(t)])(x(t) - E[x(t)])^T] \tag{2}$$

which are both symmetric positive semidefinite matrices.

Notice that we can compute

$$x(t_0 + N) = A^N x(t_0) + \sum_{j=0}^{N-1} A^j w(t_0 + N - 1 - j) \tag{3}$$

Taking expected values on both sides, we obtain,

$$\begin{aligned} \mathrm{E}[x(t_0 + N)] &= A^N E[x(t_0)] \\ &+ \sum_{j=0}^{N-1} A^j E[w(t_0 + N - 1 - j)] = A^N x(t_0) \end{aligned} \tag{4}$$

since we assume that $x(t_0)$ is given. Notice also that

$$\lim_{N\to\infty} (\mathrm{E}[x(t_0 + N)]) = 0 \tag{5}$$

since A has all its eigenvalues inside the unit circle.

Computing the covariance at both sides of (3), we obtain

$$\begin{aligned} \Sigma_x(t_0 + N) = \mathrm{cov}[x(t_0 + N)] &= \mathrm{cov}[A^N x(t_0)] \\ &+ \sum_{j=0}^{N-1} \mathrm{cov}[A^j w(t_0 + N - 1 - j)] \\ &= \sum_{j=0}^{N-1} A^j \Sigma_w {A^j}^T \end{aligned} \tag{6}$$

which provides an expression to compute the covariance of the state at any instant of time. When matrix A has all its eigenvalues inside the unit circle, it can be easily shown that the latter expression converges as $N \to \infty$. Then, defining

$$\Sigma_x(\infty) = \lim_{N\to\infty} \Sigma_x(t_0 + N)$$

recalling that

$$x(t + N + 1) = Ax(t + N) + w(t + N)$$

and computing the covariance of the above expression with $N \to \infty$, we have that $\Sigma_x \triangleq \Sigma_x(\infty)$ can be obtained from the discrete Lyapunov equation

$$\Sigma_x = A\Sigma_x A^T + \Sigma_w \tag{7}$$

2.2 Definition of PUB and PIS

We will next define the two novel notions we will be concerned with in this paper.

Definition 1 (Probabilistic Ultimate Bounds) *Let $0 < p \leq 1$ and let $S \subset \mathbb{R}^n$. We say that S is a probabilistic ultimate bound (PUB) with probability p for system (1) if for every initial state $x(t_0) = x_0 \in \mathbb{R}^n$ there exist $T = T(x_0) \in \mathbb{N}$ such that the probability $\Pr[x(t) \in S] \geq p$ for each $t \geq t_0 + T$.*

Definition 2 (Probabilistic Invariant Sets) *Let $0 < p \leq 1$ and let $S \subset \mathbb{R}^n$. We say that S is a probabilistic invariant set (PIS) with probability p for system (1) if for any $x(t) \in S$ the probability $\Pr[x(t+k) \in S] \geq p$ for each $k > 0$.*

2.3 Some properties of PUB and PIS

We present here some basic properties derived from the definitions introduced in the previous subsection. The proof of the first result is straightforward and it is thus omitted.

Lemma 3 *If S is a PUB (PIS) with probability p for (1), then it is also a PUB (PIS) with probability $\tilde{p} \geq 0$ for any $\tilde{p} < p$.*

Lemma 4 (PIS⇒PUB) *Let $S_0 \subset \mathbb{R}^n$ be a PIS for (1) with probability p which contains the origin. Given $\varepsilon > 0$ we define $S_\varepsilon = \{x : \operatorname{dist}(x, S_0) \leq \varepsilon\}$. Then, S_ε is a PUB for (1) with probability p.*

PROOF. Let $x_0 = x(t_0) \in \mathbb{R}^n$ be an arbitrary initial state for (1), and let

$$\tilde{x}(t+1) = A\tilde{x}(t) + w(t)$$

with $\tilde{x}(t_0) = 0$. Since $\tilde{x}(t_0) \in S_0$ and S_0 is a probabilistic invariant set with probability p, then $\tilde{x}(t)$ belongs to S_0 with probability greater than or equal to p for all $t > t_0$.

Define $\hat{x}(t) \triangleq x(t) - \tilde{x}(t)$. Then, $\hat{x}(t)$ verifies $\hat{x}(t+1) = A\hat{x}(t)$ with $\hat{x}(t_0) = x_0$. Since A is stable, given $\varepsilon > 0$ there exists $T = T(\varepsilon)$ such that

$$\|\hat{x}(t)\| = \|x(t) - \tilde{x}(t)\| \leq \varepsilon$$

for all $t > t_0 + T$.

Let $t > t_0 + T$. Notice that

$$\tilde{x}(t) \in S_0 \Rightarrow \operatorname{dist}(x(t), S_0) \leq \|x(t) - \tilde{x}(t)\| \leq \varepsilon$$

which implies that $x(t) \in S_\varepsilon$.

Then, for all $t > t_0 + T$, $\Pr[x(t) \in S_\varepsilon] \geq \Pr[\tilde{x}(t) \in S_0] \geq p$ which concludes the proof. □

Lemma 5 (Intersection of PUB) *Let S_1 be a PUB with probability p_1 for system (1) and let S_2 be a PUB with probability p_2 for the same system, with $p_1 + p_2 > 1$. Then, the set $S = S_1 \cap S_2$ is a PUB with probability $p = p_1 + p_2 - 1$.*

PROOF. Given $x(t_0) = x_0$, there exist T_1 and T_2 such that $\Pr[x(t) \in S_1] \geq p_1$ for each $t \geq T_1$ and $\Pr[x(t) \in S_2] \geq p_2$ for each $t \geq T_2$. Let $T \triangleq \max(T_1, T_2)$.

Then, for each $t \geq T$, we know that

$$\Pr[x(t) \notin S_1] \leq 1 - p_1$$

and

$$\Pr[x(t) \notin S_2] \leq 1 - p_2$$

Then,

$$\begin{aligned}\Pr[x(t) \notin S_1 \vee x(t) \notin S_2] &\leq \Pr[x(t) \notin S_1] + \Pr[x(t) \notin S_2] \\ &\leq 2 - p_1 - p_2\end{aligned}$$

Finally,

$$\begin{aligned}\Pr[x(t) \in S] &= \Pr[x(t) \in S_1 \wedge x(t) \in S_2] \\ &= 1 - \Pr[x(t) \notin S_1 \vee x(t) \notin S_2] \\ &\geq p_1 + p_2 - 1\end{aligned}$$

which concludes the proof. □

Lemma 6 (Intersection of PIS) *Let S_1 be a PIS with probability p_1 for system (1) and let S_2 be a PIS with probability p_2 for the same system, with $p_1 + p_2 > 1$. Then, the set $S = S_1 \cap S_2$ is a PIS with probability $p = p_1 + p_2 - 1$.*

PROOF. The proof of this lemma is very similar to that of Lemma 5. □

Corollary 7 (Intersection of several PUB/PIS) *Let $\{S_i\}_{i=1}^r$ be a collection of PUB (PIS) for system (1) with probabilities p_i, $i = 1, \ldots, r$, respectively, with $\sum_{i=1}^r p_i > (r-1)$. Then, the set $S = \cap_{i=1}^r S_i$ is a PUB (PIS) with probability $p = \sum_{i=1}^r p_i - (r-1)$.*

PROOF. The proof follows from a direct application of induction to the result of Lemma 5 (respectively, Lemma 6) for the case of PUB (respectively, PIS). $\square$

Lemma 8 (Union of PUB) *Let S_1 be a PUB with probability p_1 for system* (1) *and let S_2 be a PUB with probability p_2 for the same system, then the set $S_1 \cup S_2$ is a PUB with probability $p = \max\{p_1, p_2\}$.*

PROOF. Given any $x(t_0) \in \mathbb{R}^n$ there exist T_1 and T_2 such that:

$$\Pr[x(t) \in S_1 \cup S_2] \geq \Pr[x(t) \in S_1] \geq p_1, \quad \forall t \geq T_1$$

and

$$\Pr[x(t) \in S_1 \cup S_2] \geq \Pr[x(t) \in S_2] \geq p_2, \quad \forall t \geq T_2$$

These two inequalities imply that

$$\Pr[x(t) \in S_1 \cup S_2] \geq \max\{p_1, p_2\}, \quad \forall t \geq \max\{T_1, T_2\}$$

which concludes the proof. $\square$

Lemma 9 (Union of PIS) *Let S_1 be a PIS with probability p_1 for system* (1) *and let S_2 be a PIS with probability p_2 for the same system, then the set $S_1 \cup S_2$ is a PIS with probability $p = \min\{p_1, p_2\}$.*

PROOF. Given $x(t_0) \in S_1 \cup S_2$, we have that either $x(t_0) \in S_1$ or $x(t_0) \in S_2$.

If $x(t_0) \in S_1$ then, for any $t > t_0$ we have

$$\Pr[x(t) \in S_1 \cup S_2] \geq \Pr[x(t) \in S_1] \geq p_1 \geq \min\{p_1, p_2\}$$

If $x(t_0) \in S_2$ then, for any $t > t_0$ we have

$$\Pr[x(t) \in S_1 \cup S_2] \geq \Pr[x(t) \in S_2] \geq p_2 \geq \min\{p_1, p_2\}$$

In either case, we have $\Pr[x(t) \in S_1 \cup S_2] \geq \min\{p_1, p_2\}$ for all $t > t_0$, which completes the proof. $\square$

Corollary 10 (Union of several PUB/PIS) *Let $\{S_i\}_{i=1}^r$ be a collection of PUB (PIS) for system* (1) *with probabilities p_i, $i = 1, \ldots, r$. Then, the set $S = \cup_{i=1}^r S_i$ is a PUB (PIS) with probability $p = \max\{p_i : i = 1, \ldots, r\}$ ($p = \min\{p_i : i = 1, \ldots, r\}$).*

PROOF. The proof follows from a direct application of induction to the result of Lemma 8 (respectively, Lemma 9) for the case of PUB (respectively, PIS). $\square$

Remark 11 *When $p_i = 1$, $i = 1, \ldots, r$, Corollaries 7 and 10 say that the intersection and the union of* deterministic *invariant sets are* deterministic *invariant sets, which is a well known result.*

3 Computation of Probabilistic Ultimate Bounds

We propose here a method to compute Probabilistic Ultimate Bounds for (1) based on the covariance obtained in (7). We develop first a method based on Chebyshev's inequality which can be used for stochastic processes $w(t)$ with arbitrary distributions. We will then give tighter bounds for the special case of Gaussian noise.

In all the developments that follow, given a parameter (probability) p such that $0 < p < 1$, we will define n parameters $\tilde{p}_i$ such that

$$0 < \tilde{p}_i < 1, \, i = 1, \ldots, n; \qquad \sum_{i=1}^{n} \tilde{p}_i = 1 - p \qquad (8)$$

Also, for a vector x, x_i denotes its ith component, and for a square matrix Σ, the notation $[\Sigma]_{i,i}$ indicates its ith diagonal element.

3.1 General Distribution

Theorem 12 (PUB Computation – General Case) *Consider the system* (1)*. Assume that all the eigenvalues of $A \in \mathbb{R}^{n \times n}$ lie strictly inside the unit circle and suppose that $w(t)$ is zero-mean white noise with covariance matrix Σ_w. Let $0 < p < 1$ and $\tilde{p}_i$, $i = 1, \ldots, n$, be defined as in* (8)*. Then, for any $\varepsilon > 0$, the set $S = \{x : |x_i| \leq b_i + \varepsilon; \; i = 1, \ldots, n\}$ is a probabilistic ultimate bound for the system with probability p, where*

$$b_i \triangleq \sqrt{\frac{[\Sigma_x]_{i,i}}{\tilde{p}_i}}; \quad i = 1, \ldots, n$$

and Σ_x is the solution of the discrete Lyapunov equation (7)*.*

PROOF. Let $\Sigma_x(t)$ be defined as in (2). The term $[\Sigma_x(t)]_{i,i}$ represents the variance of $x_i(t)$. Using Chebyshev's inequality on this component, it results that

$$\Pr[|x_i(t) - \mathrm{E}[x_i(t)]| \geq b_i + \varepsilon/2] \leq \frac{[\Sigma_x(t)]_{i,i}}{(b_i + \varepsilon/2)^2} \qquad (9)$$

Taking into account that $\Sigma_x(t) \to \Sigma_x$ as $t \to \infty$, there exists T_a sufficiently large such that for all $t > T_a$ it results

$$\frac{[\Sigma_x(t)]_{i,i}}{(b_i + \varepsilon/2)^2} \leq \frac{[\Sigma_x]_{i,i}}{b_i^2} = \tilde{p}_i \qquad (10)$$

Also, since $\mathrm{E}[x_i(t)] \to 0$ when $t \to \infty$, there exists T_b sufficiently large such that for all $t > T_b$ it results

$$|\mathrm{E}[x_i(t)]| < \varepsilon/2$$

and then, for all $t > T_b$, it is true that

$$\begin{aligned}\Pr[|x_i(t)| \geq b_i + \varepsilon] &= \Pr[|x_i(t)| - \varepsilon/2 \geq b_i + \varepsilon/2] \\ &\leq \Pr[|x_i(t)| - |\mathrm{E}[x_i(t)]| \geq b_i + \varepsilon/2] \\ &\leq \Pr[|x_i(t) - \mathrm{E}[x_i(t)]| \geq b_i + \varepsilon/2]\end{aligned}$$

Substituting this last inequality into the left hand side of (9) and replacing the right hand side with (10), we obtain

$$\Pr[|x_i(t)| > b_i + \varepsilon] \leq \Pr[|x_i(t)| \geq b_i + \varepsilon] \leq \tilde{p}_i \tag{11}$$

which holds for all $t > T_i \triangleq \max(T_a, T_b)$.

Repeating this analysis for all the components of $x(t)$, and taking $T = \max_i(T_i)$, it results that Inequality (11) holds for all $t > T$ and for $i = 1, \dots, n$.

Then,

$$\Pr[x(t) \notin S] \leq \sum_{i=1}^{n} \Pr[|x_i(t)| > b_i + \varepsilon] \leq \sum_{i=1}^{n} \tilde{p}_i = 1 - p \tag{12}$$

Thus,

$$\Pr[x(t) \in S] \geq p$$

for all $t > T$, which completes the proof. □

Remark 13 *Theorem 12 provides a general but possibly conservative result. It is based on Chebyshev's inequality, which is in general conservative since it holds for arbitrary distributions. Also, it does not exploit the knowledge of the whole covariance matrix of $x(t)$. Thus, another source of conservatism is that, in the estimation of* $\Pr[x(t) \notin S]$ *in* (12)*, this probability is bounded with the sum of the probabilities that the individual components, x_i, become larger than the corresponding bounds, without taking into account the correlation between individual components.*

3.2 Gaussian Distribution

The following theorem, valid for the special case of a Gaussian white noise, provides tighter bounds than those of Theorem 12. The latter is achieved by resorting to specific properties of Gaussian distributions, without the need to employ Chebyshev's inequality.

Theorem 14 (PUB Computation – Gaussian Noise) *Consider the system* (1). *Assume that all the eigenvalues of $A \in \mathbb{R}^{n\times n}$ lie strictly inside the unit circle and suppose that $w(t)$ is zero-mean white Gaussian noise with covariance matrix Σ_w. Let $0 < p < 1$ and $\tilde{p}_i$, $i = 1, \dots, n$, be defined as in* (8). *Then, for any $\varepsilon > 0$, the set $S = \{x : |x_i| \leq b_i + \varepsilon;\ i = 1, \dots, n\}$ is a probabilistic ultimate bound for the system with probability p, where*

$$b_i \triangleq \sqrt{2[\Sigma_x]_{i,i}}\,\mathrm{erf}^{-1}(1 - \tilde{p}_i); \quad i = 1, \dots, n \tag{13}$$

and where Σ_x is the solution of the Lyapunov Equation (7) *and* erf *is the* error function*: $\mathrm{erf}(z) \triangleq \frac{2}{\sqrt{\pi}} \int_0^z e^{-\zeta^2} d\zeta$.*

PROOF. The proof can be found in [7]. □

4 Computation of Probabilistic Invariant Sets

Here, again, we first propose a method to compute probabilistic invariant sets for (1) that can be used for stochastic processes $w(t)$ with arbitrary distributions, and then we provide a method for the particular case of Gaussian noises. In the rest of the developments we will assume the matrix A in (1) to be diagonalisable[1] and we will denote by $\rho(A)$ the *spectral radius* of A (that is, the maximum of the absolute values of the eigenvalues of A). The symbol $\preceq$ will denote the elementwise inequality between two vectors, i.e., for $\alpha, \beta \in \mathbb{R}^n$, $\alpha \preceq \beta$ if and only if $\alpha_i \leq \beta_i$, $i = 1, \dots, n$. For a matrix M with complex entries, M^* will denote the conjugate transpose of M.

4.1 General Distribution

Theorem 15 (PIS Computation – General Case) *Consider the system* (1)*, where matrix A is assumed to be diagonalisable and satisfying $\rho(A) < 1$. Suppose that $w(t)$ is zero-mean white noise with covariance matrix Σ_w. Let $0 < p < 1$ and $\tilde{p}_i$, $i = 1, \dots, n$, be defined as in* (8). *Then, the set $S = \{x : |V^{-1}x| \preceq b\}$ is a probabilistic invariant set for the system with probability p, where V is a similarity transformation such that*

$$\Lambda = \mathrm{diag}(\lambda_1, \dots, \lambda_n) = V^{-1}AV$$

is the Jordan decomposition of matrix A, and the components of $b = [b_1 \ \dots \ b_n]^T$ are computed according to

$$b_i \triangleq \sqrt{\frac{[\Sigma_v]_{i,i}}{\tilde{p}_i}} \frac{1}{(1 - |\lambda_i|)}; \quad i = 1, \dots, n$$

with

$$\Sigma_v = V^{-1}\Sigma_w(V^{-1})^*$$

[1] In an important number of applications, the matrix A in (1) is given by some *closed-loop matrix*, e.g., $\tilde{A} - \tilde{B}K$ or $\tilde{A} - L\tilde{C}$ [where $(\tilde{A}, \tilde{B}, \tilde{C})$ is the open-loop system and K is a feedback gain, L is an observer gain, etc.]. Under standard controllability and observability conditions on $(\tilde{A}, \tilde{B}, \tilde{C})$ the design of K, L, etc., can be readily done by pole placement techniques so that the assumptions on A made here are, without loss of generality, satisfied.

PROOF. With the linear transformation $x(t) = Vz(t)$, system (1) becomes

$$z(t+1) = \Lambda z(t) + V^{-1}w(t) \qquad (14)$$

with $z \in \mathbb{C}^n$, $w(t) \in \mathbb{R}^n$, $V^{-1} \in \mathbb{C}^{n\times n}$, and $\Lambda \in \mathbb{C}^{n\times n}$ being a diagonal matrix.

Defining $v(t) \triangleq V^{-1}w(t)$, the covariance of $v(t)$ results $\Sigma_v = V^{-1}\Sigma_w(V^{-1})^*$, and the i–th component of (14) is

$$z_i(t+1) = \lambda_i z_i(t) + v_i(t) \qquad (15)$$

Then,

$$z_i(t_0+N) = \lambda_i^N z_i(t_0) + \sum_{j=0}^{N-1} \lambda_i^j v_i(t_0+N-1-j) \quad (16)$$

The expected value of this last expression is

$$\mathrm{E}[z_i(t_0+N)] = \lambda_i^N z_i(t_0)$$

since we assume that $z(t_0) = V^{-1}x(t_0)$ is given. From (16), the variance of $z_i(t_0+N)$ is

$$\begin{aligned}\mathrm{var}[z_i(t_0+N)] &= \sum_{j=0}^{N-1} |\lambda_i^j|^2 \mathrm{var}[v_i(t_0+N-1-j)] \\ &= \frac{(1-|\lambda_i|^{2N})[\Sigma_v]_{i,i}}{1-|\lambda_i|^2} \qquad (17)\end{aligned}$$

Suppose that $x(t_0) \in S$. Thus, $|z(t_0)| = |V^{-1}x(t_0)| \preceq b$ and $|z_i(t_0)| \le b_i$. Then, it results that

$$|\mathrm{E}[z_i(t_0+N)]| = |\lambda_i^N z_i(t_0)| \le |\lambda_i|^N b_i \qquad (18)$$

From Inequality (18), it follows that

$$\begin{aligned}&\Pr[|z_i(t_0+N)| \ge b_i] \\ &\quad= \Pr[|z_i(t_0+N)| - |\lambda_i|^N b_i \ge b_i(1-|\lambda_i|^N)] \\ &\quad\le \Pr[|z_i(t_0+N)| - |\mathrm{E}[z_i(t_0+N)]| \ge b_i(1-|\lambda_i|^N)] \\ &\quad\le \Pr[|z_i(t_0+N) - \mathrm{E}[z_i(t_0+N)]| \ge b_i(1-|\lambda_i|^N)]\end{aligned}$$

Chebyshev's inequality establishes that

$$\begin{aligned}&\Pr\left[|z_i(t_0+N) - \mathrm{E}[z_i(t_0+N)]| \ge b_i(1-|\lambda_i|^N)\right] \\ &\qquad \le \frac{\mathrm{var}[z_i(t_0+N)]}{b_i^2(1-|\lambda_i|^N)^2} \qquad (19)\end{aligned}$$

and then it results that

$$\begin{aligned}\Pr[|z_i(t_0+N)| \ge b_i] &\le \frac{(1-|\lambda_i|^{2N})[\Sigma_v]_{i,i}}{(1-|\lambda_i|^2)b_i^2(1-|\lambda_i|^N)^2} \\ &= \frac{(1+|\lambda_i|^N)[\Sigma_v]_{i,i}}{(1-|\lambda_i|^2)b_i^2(1-|\lambda_i|^N)} \\ &\le \frac{(1+|\lambda_i|)[\Sigma_v]_{i,i}}{(1-|\lambda_i|^2)b_i^2(1-|\lambda_i|)} \\ &= \frac{[\Sigma_v]_{i,i}}{(1-|\lambda_i|)^2 b_i^2} = \tilde{p}_i\end{aligned}$$

for all $N \ge 1$; where we have used the expression in (17) for $\mathrm{var}[z_i(t_0+N)]$, simple algebraic steps to simplify the above expressions, the definition of b_i given in the statement of the theorem, and in the inequality between the second and third row terms we have used the fact that the term on the second row decreases with N and achieves its maximum at $N = 1$. Then, it results that $\Pr[|z_i(t_0+N)| > b_i] \le \Pr[|z_i(t_0+N)| \ge b_i] \le \tilde{p}_i$. Thus, the probability

$$\begin{aligned}\Pr[|z(t+N)| \not\preceq b] &\le \sum_{i=1}^n \Pr[|z_i(t_0+N)| > b_i] \\ &\le \sum_{i=1}^n \tilde{p}_i = 1-p\end{aligned}$$

for all $N \ge 1$, and then,

$$\begin{aligned}\Pr[|z(t+N)| \preceq b] &= \Pr[|V^{-1}x(t+N)| \preceq b] \\ &= \Pr[x(t+N) \in S] \ge p\end{aligned}$$

which proves that the set S is a probabilistic invariant set with probability p. □

4.2 Gaussian Distribution

Here we again obtain tighter bounds for the case of Gaussian noises by replacing the use of Chebyshev's inequality with specific properties of Gaussian distributions.

Theorem 16 (PIS Computation – Gaussian Noise) *Consider the system (1), where matrix A is assumed to be diagonalisable and satisfying $\rho(A) < 1$. Let*

$$\Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_n) = V^{-1}AV$$

be the Jordan decomposition of matrix A. Suppose that $w(t)$ is zero-mean white Gaussian noise with covariance matrix Σ_w. Let $0 < p < 1$ and $\tilde{p}_i$, $i = 1, \ldots, n$, be defined as in (8) with the restriction that for each pair of complex conjugate eigenvalues λ_i, $\lambda_j = \bar{\lambda}_i$, we take $\tilde{p}_i = \tilde{p}_j$. Then, the set $S = \{x : |V^{-1}x| \preceq b\}$ is a probabilistic invariant set for the system with probability p, where the *components of* $b = [b_1 \ \dots \ b_n]^T$ *are computed according to*

$$b_i \triangleq \frac{\sqrt{2[\Sigma_v]_{i,i}}}{(1-|\lambda_i|)}\mathrm{erf}^{-1}(1-\tilde{p}_i); \quad i = 1, \dots, n$$

with

$$\Sigma_v = V^{-1}\Sigma_w (V^{-1})^*$$

PROOF. The proof can be found in [7]. □

Remark 17 *Definitions 1 and 2 of PUB and PIS are general and can be accomplished by sets of arbitrary shape. However, Theorems 12–16 only provide ways to compute polytopic PUB and PIS.*

Besides the fact that the calculations of polytopic PUB and PIS are simple and lead to explicit formulas, there are several practical applications in which polytopic invariant sets play a key role such as fault tolerant control....

Yet, analogous results to those of Theorems 12–16 may eventually be obtained for PUB and PIS of different shape. For instance, it can be considered the problem of finding the PUB/PIS with minimum volume, or finding the PUB/PIS that turn to be optimal in some other sense. However, those results are beyond the goals of the present work.

5 Example

We consider the system

$$x(t+1) = Ax(t) + Bv(t) = \begin{bmatrix} 0 & 0.6 \\ 0.3 & -0.5 \end{bmatrix} x(t) + \begin{bmatrix} 0 \\ 1 \end{bmatrix} v(t) \tag{20}$$

where $v(t)$ is zero-mean white Gaussian noise with variance $\Sigma_v = 1$. Taking into account that a Gaussian noise is unbounded, under the standard definitions in the literature, the system is not ultimately bounded. We shall then try to compute a probabilistic ultimate bound with probability $p = 0.99$.

Defining $w(t) \triangleq Bv(t)$, the covariance matrix of $w(t)$ can be computed as

$$\Sigma_w = B\sigma_v B^T = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \tag{21}$$

Now, in order to apply Theorem 14, we take

$$\tilde{p}_1 = \tilde{p}_2 = 0.005$$

and then, for any $\varepsilon > 0$, the set $S = \{x : |x_i| \le b_i + \varepsilon;\ i = 1, 2\}$ with $b_1 = 2.1602$, $b_2 = 3.6004$ is a PUB with probability $p = 0.99$ for the system.

In order to verify the correctness and to check for the potential conservatism of the result, we performed several long time simulations of the system (up to time $t = 1 \times 10^6$) and we found in all the cases that between 0.9% and 0.93% of the samples of $x(t)$ were outside the computed set. Theorem 14 states that the probability that a sample be outside the set is less than 1%, which agrees very closely with the simulation results.

We then recomputed the probabilistic ultimate bound sets for different choices of $\tilde{p}_i$ satisfying (8). Figure 1 shows the results for three different sets of parameters.

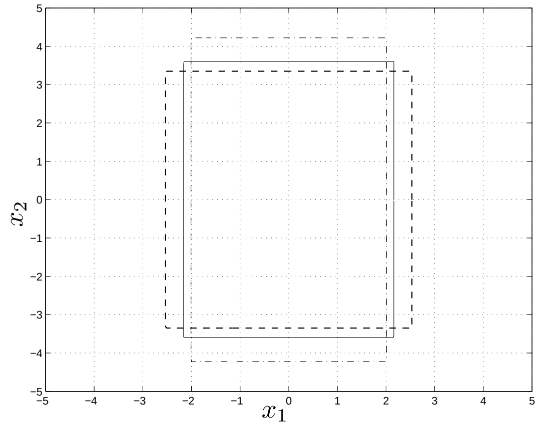

Fig. 1. PUBs for $\tilde{p}_1 = 0.001, \tilde{p}_2 = 0.009$ (dashed), $\tilde{p}_1 = 0.005, \tilde{p}_2 = 0.005$ (solid) and $\tilde{p}_1 = 0.009, \tilde{p}_2 = 0.001$ (dash–dotted)

We note that the intersection of these sets, equal to the intersection of just two of them (the one indicated with dashed lines and the one indicated with dash-dotted lines) is also a PUB with probability 0.98 (see Lemma 5). We remark that the smallest ultimate bounds are obtained in the limiting cases (see (13)) when $\tilde{p}_1 \to 0.01 = 1-p$, $\tilde{p}_2 \to 0$ (smallest bound on x_1, whereas the bound on x_2 tends to infinity), and $\tilde{p}_1 \to 0$, $\tilde{p}_2 \to 0.01 = 1-p$ (smallest bound on x_2, whereas the bound on x_1 tends to infinity), and that the intersection of the sets obtained in these two limiting cases is *visually* indistinguishable from the intersection of the ones indicated in Figure 1 with dashed and dash-dotted lines.

We also computed a PIS with probability $p = 0.99$ for the system. For that goal, we set $\tilde{p}_1 = \tilde{p}_2 = 0.005$ obtaining a set

$$S = \{x : |V^{-1}x| \preceq b\}$$

with

$$V = \begin{bmatrix} 0.9272 & -0.6285 \\ 0.3746 & 0.7778 \end{bmatrix}; \quad b = \begin{bmatrix} 2.4347 \\ 10.5633 \end{bmatrix}$$

Figure 2 plots the PIS and the PUB previously obtained. We note here that from Lemma 4 we have that the PIS

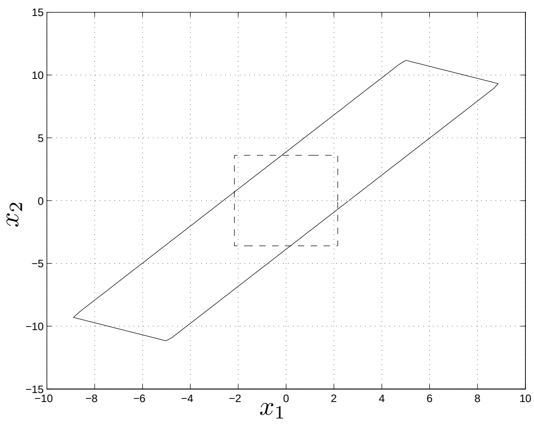

Fig. 2. PIS (solid) and PUB (dashed) for $\tilde{p}_1 = 0.005, \tilde{p}_2 = 0.005$

(enlarged by an $\epsilon > 0$, which can be arbitrarily small) is also a PUB with probability 0.99. Thus, Lemma 5 tells us that the intersection of the two sets of Figure 2 is also a PUB with probability 0.98.

As in the PUB case, we can obtain different PIS changing the values for $\tilde{p}_1, \tilde{p}_2$. Figure 3 shows the PIS obtained

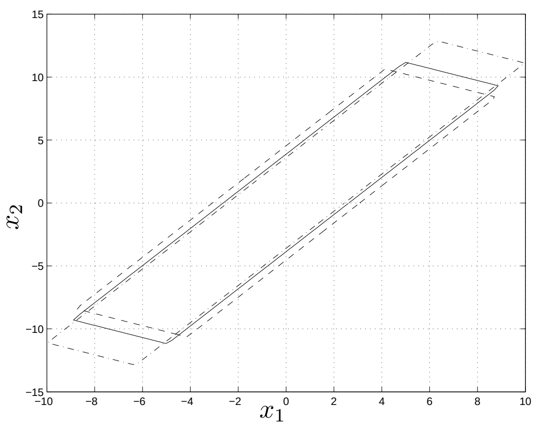

Fig. 3. PIS for $\tilde{p}_1 = 0.005, \tilde{p}_2 = 0.005$ (solid) $\tilde{p}_1 = 0.001, \tilde{p}_2 = 0.009$ (dashed) and $\tilde{p}_1 = 0.009, \tilde{p}_2 = 0.001$ (dash–dotted)

for three different choices of $\tilde{p}_1, \tilde{p}_2$. Here again the intersection of these three sets (equal to the intersection of just two of them) gives a PIS with probability 0.98 (Lemma 6). Similar remarks to the ones made in connection to the PUB sets of Figure 1, regarding the limiting cases for these sets (smallest bounds), apply here to the PIS of Figure 3.

6 Conclusions

We have proposed the novel concepts of probabilistic ultimate bounds and probabilistic invariant sets. These concepts extend the notions of invariant sets and ultimate bounds to consider 'containment in probability', and are intended to alleviate some of the shortcomings of the latter concepts, since they allow stochastic noises with more general distributions, including the ubiquitous Gaussian distribution, to be considered. We have introduced some key definitions for these sets, established their main properties and developed methods for their computation. A numerical example has been included to illustrate the main ideas.

References

[1] Z. Artstein and S. V. Raković. Feedback and invariance under uncertainty via set-iterates. *Automatica*, 44(2):520–525, 2008.

[2] F. Blanchini. Set invariance in control – a survey. *Automatica*, 35(11):1747–1767, 1999.

[3] F. Blanchini and S. Miani. *Set-theoretic methods in control*. Birkhauser, 2007.

[4] G. Calafiore, F. Dabbene, and R. Tempo. Research on probabilistic methods for control system design. *Automatica*, 47:1279–1293, 2011.

[5] M. Cannon, B. Kouvaritakis, S. Raković, and Q. Cheng. Stochastic tubes in model predictive control with probabilistic constraints. *IEEE Transactions on Automatic Control*, 56(1):194–200, 2011.

[6] H. Deng, M. Krstić, and R. Williams. Stabilization of stochastic nonlinear systems driven by noise of unknown covariance. *IEEE Transactions on Automatic Control*, 46(8):1237–1253, 2001.

[7] E. Kofman, J. De Doná, and M. Seron. Probabilistic ultimate bounds and invariant sets for LTI systems with gaussian disturbances. In *Proceedings of 2011 Australian Control Conference*, pages 537–542, Melbourne, Australia, 10–11 November, 2011.

[8] E. Kofman, H. Haimovich, and M. Seron. A systematic method to obtain ultimate bounds for perturbed systems. *International Journal of Control*, 80(2):167–178, 2007.

[9] I. Kolmanovsky and E. G. Gilbert. Theory and computation of disturbance invariant sets for discrete-time linear systems. *Mathematical Problems in Engineering*, 4:317–367, 1998.

[10] S.J. Liu, J.F. Zhang, and Z.P.Jiang. A notion of stochastic input-to-state stability and its application to stability of cascaded stochastic nonlinear systems. *Acta Mathematicae Applicatae Sinica, English Series*, 24(1):141–156, 2008.

[11] G. Pola, J. Lygeros, and M. D. Di Benedetto. Invariance in stochastic dynamical control systems. In *Proceedings of the 17th International symposium on Mathematical Theory of Network and Systems*, Kyoto, Japan, July 2006.

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