POLITECNICO MILANO 1863

SCUOLA DI INGEGNERIA INDUSTRIALE
E DELL'INFORMAZIONE

EXECUTIVE SUMMARY OF THE THESIS

Shaping the stationary state distribution via state-feedback and the scenario approach

TESI MAGISTRALE IN AUTOMATION AND CONTROL ENGINEERING – INGEGNERIA DELL'AUTOMAZIONE

AUTHOR: GIULIO SALIZZONI

ADVISOR: SIMONE GARATTI

ACADEMIC YEAR: 202O-2021

1. Introduction

The subject of study is discrete-time time-invariant linear system which is affected by a stationary disturbance. It is described by the following equation:

$$x_{k+1} = Ax_k + Bu_k + d_k$$

The aim is to design a method to develop a state-feedback controller while optimizing a certain performance, set by the user. The state and the input must satisfy some probabilistic constraints when the system is operating in stationary conditions.

The path followed is similar to the one proposed in [1], but with substantial differences in the assumptions and in the controller.

2. Problem formulation

The problem defined is a chance-constrained optimization program. The controller has a constant term γ and a feedback term, obtained multiplying the state for a matrix K with a suitable dimension. The goal is to define a method to optimize these controller parameters such that the stationary state distribution respects the constraints. Unlike most previous cases where the scenario approach was used, both the cost function and the constraints are not assumed to be convex. The disturbance is a stationary stochastic process with zero mean.

3. Scenario approach

The scenario approach is a methodology, developed in the recent years, to deal with chance-constrained optimization program. This type of problems is generally hard to solve, for both the presence of probabilistic constraints and the application to the stationary process, which prevents the use of analytic methods. Using the scenario approach, however, it is possible to simplify the problem, replacing the probabilistic constraint with the equivalent normal one. As demonstrated by [2], by choosing an appropriate sampling of the constraints it is possible to obtain a standard convex optimization problem. The

solution is then approximately feasible for the original set of constraints. This means that the number of all the original constraints that are violated when applying the solution found goes to zero with the growth of the number of samples used in the program.
One of the strengths of this approach is that it does not require any knowledge on the probability $\mathbb{P}$. The scenarios are sampled from the system, according to the probability but without the necessity to know it. This is particularly useful as not always is possible to determine with precision the probabilistic distribution of certain variables.

4. Truncation

Even after the application of the scenario approach, the problem remains in practice unfeasible, due to the use of the stationary conditions. To calculate them, it is, in fact, necessary to have infinitely long realizations of the disturbance, while in practice there are only finite ones. To solve this problem, an approximation is introduced. Instead of the stationary condition process, a truncated version is used, calculated for the first M terms. To compensate for the approximation error, it is necessary to introduce a tightening δ on the constraints and a bound on the norm of the matrix $(A + BK)$. Thanks to this simplification, together with the scenario approach, the problem is now an optimization program that can be solved through standard methods. An example of resolution is given in the practical case, where the matlab function *fmincon* is used.

5. Norm bound

The difference between the stationary state and the approximated one is:

$$\left\|x_{k,\infty} - x_{k,M}\right\| = \left\|\sum_{s=M}^{\infty}(A + BK)^s d_{k-1-s}\right\|$$

In order to the difference to be negligible, considering the assumption made on the disturbance, the norm of the matrix $(A + BK)^s$ has to go to zero as s goes to infinity. It is possible to apply many conditions on the norm. The simplest one is requiring that $\|A + BK\| < 1$. Then, thanks to the sub-multiplicative property of the norm, all the matrix power will be smaller than one and tend towards zero. This condition is, however, too strict. In fact, it limits the values that the matrix $(A + BK)^s$ also for $s < M$, even if that is not necessary. Between the possibility explored in the thesis, a good one is bounding with an exponential the norm just for power greater than the truncation value. The condition is expressed as follows:

$$\|(A + BK)^t\| \le \lambda^t \ \ for\ t = M, \dots, 2M - 1$$

It is indeed sufficient to express the condition until a certain value, the following one are bounded thanks to the sub-multiplicative property.

6. Violation probability

The violation probability describes the probability that the scenario solutions does not fulfill the constraints for new realizations of the disturbance. In practice, it useful to evaluate the generalization of the scenario decision, so to determine how well it deals with unseen situations. It has been demonstrated that it is possible to bound the violation probability, but to do that it is necessary to introduce the concept of support set.
Often, not all the scenarios used to determine the solution contributed to the process. In these cases, it is possible to remove some of the realizations and the program would provide the same result. A support set is a subsample of the original set that guarantees the same solution. Starting from the cardinality of the support set, i.e., the number of scenarios contained, the cardinality of the original set and a confidence parameter β, a bound for the violation probability can be determined. In case of convex problem, it is known that the cardinality of the support set will be equal, or smaller, to the dimension of the state. Therefore, given the confidence parameter which gives the probability with which the bound is true, it is possible to calculate a priori the number of scenarios needed to obtain a certain maximum violation probability. This is not possible with nonconvex constraints, with which a different approach has to be used. The cardinality of the support set can be determined only a posteriori, therefore the violation probability bound can be used when deciding if the solution found is satisfying or not and whether adopt it or not [3].
Notice that, for our goal, the support set identified does not need to be the smallest one neither an irreducible one (i.e., no scenario can be removed without a change in the solution). However, the smaller the cardinality, the better are the guarantees, so a small set is preferable.

7. Research of the support set

With enough time, it is always possible to find the smallest irreducible support set through brute force, but it is not, in practice, a feasible path. As said, it is not necessary to find the smallest irreducible support set, so it is possible to balance the quality of the result with the time needed. A simple solution, presented in [3] as well in other articles, is a simple greedy algorithm, that removes one scenario at a time from the original set. This algorithm guarantees that the support set found is irreducible, but it needs at least as many steps as the number of scenarios present in the original set. Also, the program has to be repeated many times with a great number of scenarios.

To solve these problems, it is possible to use an opposite approach. The idea is to look for a support set adding a scenario at a time. The critical point is the order used to add the scenarios. A good order, for the problems studied, is following the distance of the state from the constraints, starting from the nearest one to the most distant. Then, when a support set is found, it is possible to apply the simple greedy algorithm to it, to cut off the scenarios not needed and to guarantee the irreducibility. This algorithm showed good performances. It generally uses many less steps than the simple greedy algorithm, it requires to solve simpler problem and has the same guarantees.

8. Complete problem and solution

The compete problem is formulated as follows:

$$\min_{\gamma,K,h} h$$
$$s.t.\, \mathbb{P}_{d_k}\{l(x_{k,\infty},\gamma+Kx_{k,\infty}) \le h \wedge f(x_{k,\infty},\gamma+Kx_{k,\infty}) \le 0\} \ge 1-\varepsilon$$
$$\|(A+BK)^t\| \le \lambda^t \;\; for\; t=M,\dots,2M-1$$

After applying the scenario approach and the truncation, the new version of the problem is:

$$\min_{\gamma,K,h} h$$
$$s.t.\, l(x_{k,M},\gamma+Kx_{k,M}) \le h-\delta$$
$$f(x_{k,M},\gamma+Kx_{k,M}) \le -\delta$$
$$x_{k,M}^{(i)} = (I-A-BK)^{-1}B\gamma + \sum_{s=0}^{M-1}(A+BK)^s d_{k-1-s}$$
$$i=1,\dots,N$$
$$\|(A+BK)^t\| \le \lambda^t \;\; for\; t=M,\dots,2M-1$$

Given a solution (γ^M, K^M, h^M) and considering a cost function dependent only on the controller parameters (this condition is not mandatory, but it is useful to simplify the calculous and the notations), then it is guaranteed that:

$$\mathbb{P}_{d_k}^N\left\{\mathbb{P}_{d_k}\{f(x_{k,\infty},\gamma+Kx_{k,\infty})>0\} \le \varepsilon(s^M)+\frac{\chi_M}{\delta}\right\} \ge 1-\beta$$

s^M is the cardinality of the support set and χ_M is a parameter dependent on the difference between the stationary process and the truncated one, and therefore on the truncation value M and the norm bound exponential λ.

In practice, the probability that the optimal controller parameter found are not suitable for a realization of the disturbance, coherent with the probability distribution, is smaller than a term evaluated. $\varepsilon(s^M)$ depends one the number of scenarios used, the cardinality of the support set found and the confidence parameter β. χ_M/δ depends on the approximation of the stationary process, the bound on the norm and the tightening of the constraints. We can guarantee this with a probability greater than $1-\beta$. Since the dependence of $\varepsilon(s^M)$ on β is logarithmic, it is possible to choose a very low value for the confidence parameter, making the inequality true in almost all the cases, without great repercussion.

9. Choice of the parameters

There are many tunable parameters, giving more freedom to the user but also increasing the choice complexity. In some cases, the number of scenarios N and the truncation value M are limited by the data available. This can simplify a lot the decision process, but that is not always the case. Also the confidence parameter β can be given, but, even if it is not, it is the easiest to tune, since its value can be assigned after having found the solution. The other parameters to assess are λ and δ. If N and M are given, it only necessary to balance the two values. If, for example, δ can not be too big, maybe because the constraints are already very strict, then the value of λ should be quite small, to limit the value of χ_M/δ. If even N and M must be tuned, the tuning becomes more sophisticated. It is necessary to consider the computational load, so there are practical limits on the value of N and M. A general method to decide can be given, the choice has to be done based on the specific case. If the probabilistic distribution of the disturbance is known, and so it is possible to generate the realizations, a solution could be to do many attempts, varying the parameters and looking for the best combination.

10. Simulation example

One of the two examples brought in the thesis is an asymptotically stable system with two states and two control inputs. The state matrix and the control matrix are:

$$A = \begin{bmatrix} 0.9 & -0.1 \\ 0.2 & 0.6 \end{bmatrix}$$
$$B = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

The constraints are given by two parallel straight line, between which the state variables should stay. The norm bound is an exponential with coefficient λ. Its value and the other parameters are reported in the following table.

Parameter	Value
Number of scenarios	N=200
Truncation value	M=20
Norm bound coefficient	$\lambda = 0.72$
Confidence parameter	$\beta = 10^{-5}$

The controller computed, using the algorithm developed, is:

$$K = \begin{bmatrix} -0.4164 & 0.1609 \\ -2.2481 & -0.3233 \end{bmatrix}$$
$$A + BK = \begin{bmatrix} 0.4836 & 0.0609 \\ -2.0481 & 0.2767 \end{bmatrix}$$
$$\gamma = \begin{bmatrix} 0.0752 \\ 0.0394 \end{bmatrix}.$$

In figure 1 it is possible to appreciate the work done by the controller. In orange, there are the state variables after M step if the disturbance realizations are applied without control. In blue, the state variables with the same disturbance realizations but with the developed control. Thanks to the controller, all the scenarios lay between the two constraints. The bound for the violation probability, calculated assuming a value of 0.1 for δ, is 16.52%. A test done using 10000 scenarios (not used in the solution computation), shown in figure 2, has a violation percentage of 2%. In figure 3, it is shown the evolution of $\|(A+BK)^s\|$ compared with the exponential bound applied.

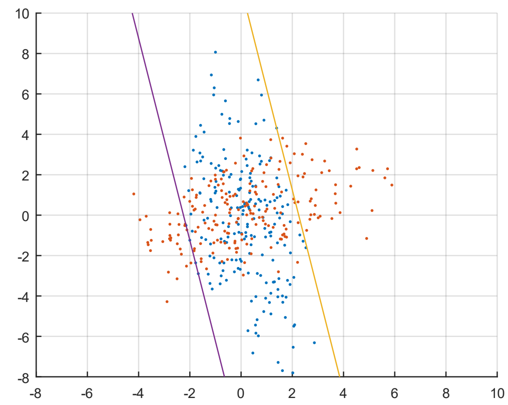


Figure 1: The scenarios without the control action (in orange) and with (in blue). The two straight lines are the two constraints.

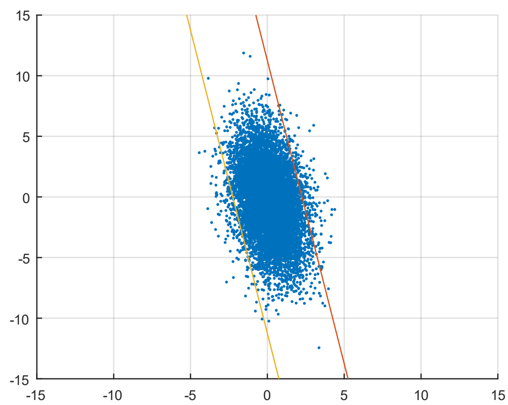


Figure 2: The test performed using several disturbance realizations to assess the controller quality

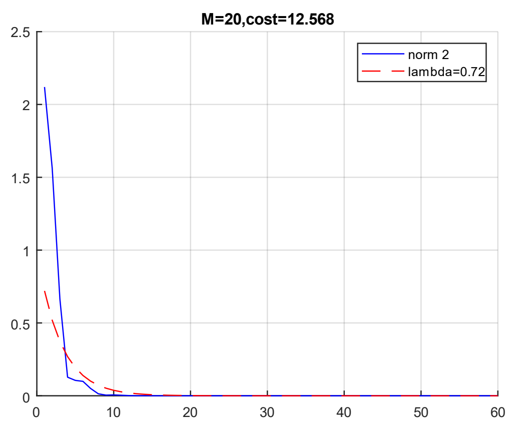


Figure 3: The norm $\|(A+BK)^s\|$ (in blue) and the exponential bound (in red and dashed).

11. Conclusion

The methodology developed allows to deal with chance-constrained optimization problem. The main strength is its generality, requiring very few assumptions. Thanks to the use of the scenario approach, there is no restriction on the disturbance probability distribution neither assumption on its knowledge, and, differently from most of the other cases, convexity of the cost function and the constraints is not required. Also, thanks to the use of a state-feedback controller, there is no assumption on the stability of the state matrix. The user has a lot of personalization possibility, due to the many tunable parameters, allowing a better adaptation to the specific case.

Bibliografia

[1] A. Falsone, L. Deori, D. Ioli, S. Garatti e M. Prandini, «Optimal disturbance compensation for constrained linear systems operating in stationary conditions: A scenario-based approach,» *Automatica,* 2019.

[2] G. C. Calafiore e M. C. Campi, «The scenario approach to robust control design,» *IEEE Transactions on automatic control,* vol. 51, n. 5, pp. 742-753, 2006.

[3] M. C. Campi, S. Garatti e F. A. Ramponi, «A general scenario theory for nonconvex optimization and decision making,» *IEEE Transactions on automatic control,* vol. 63, n. 12, pp. 4067-4078, 2016.

POLITECNICO
MILANO 1863

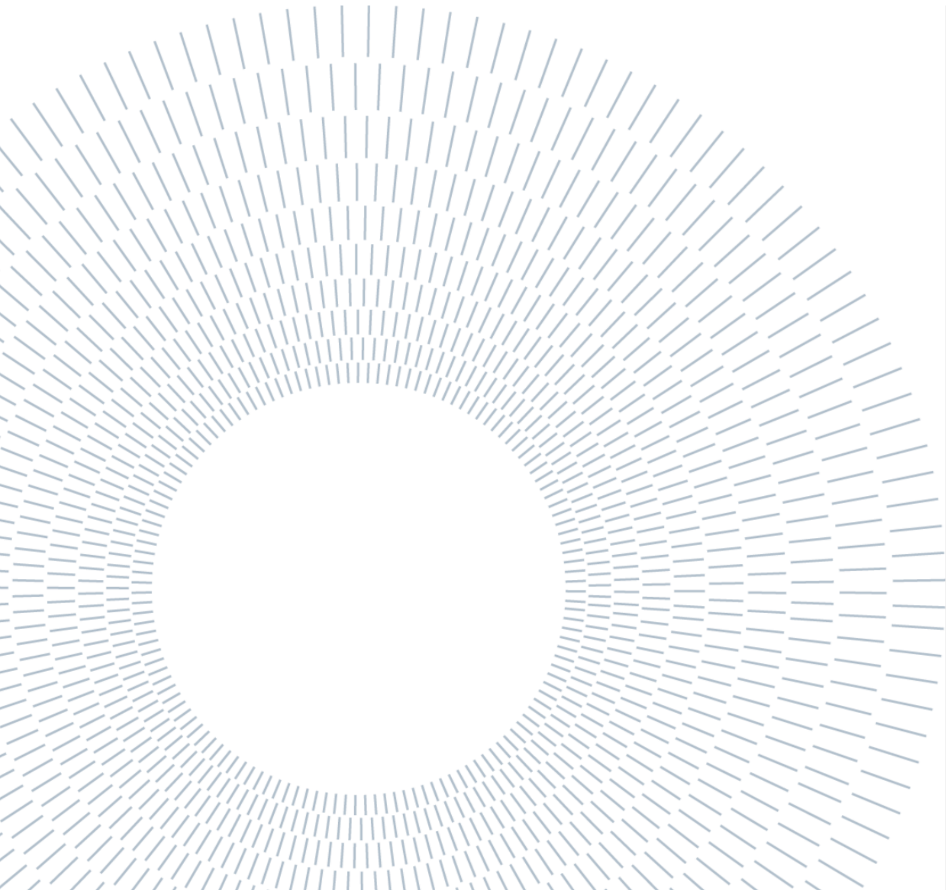
SCUOLA DI INGEGNERIA INDUSTRIALE
E DELL'INFORMAZIONE

Shaping the stationary state distribution via state-feedback and the scenario approach

TESI DI LAUREA MAGISTRALE IN
AUTOMATION AND CONTROL ENGINEERING-
INGEGNERIA DELL'AUTOMAZIONE

Author: **Giulio Salizzoni**

Student ID:	940100
Advisor:	Simone Garatti
Co-advisor:	Alessandro Falsone
Co-advisor:	Maria Prandini
Co-advisor:	Diego Regruto
Academic Year:	2020-21



Abstract

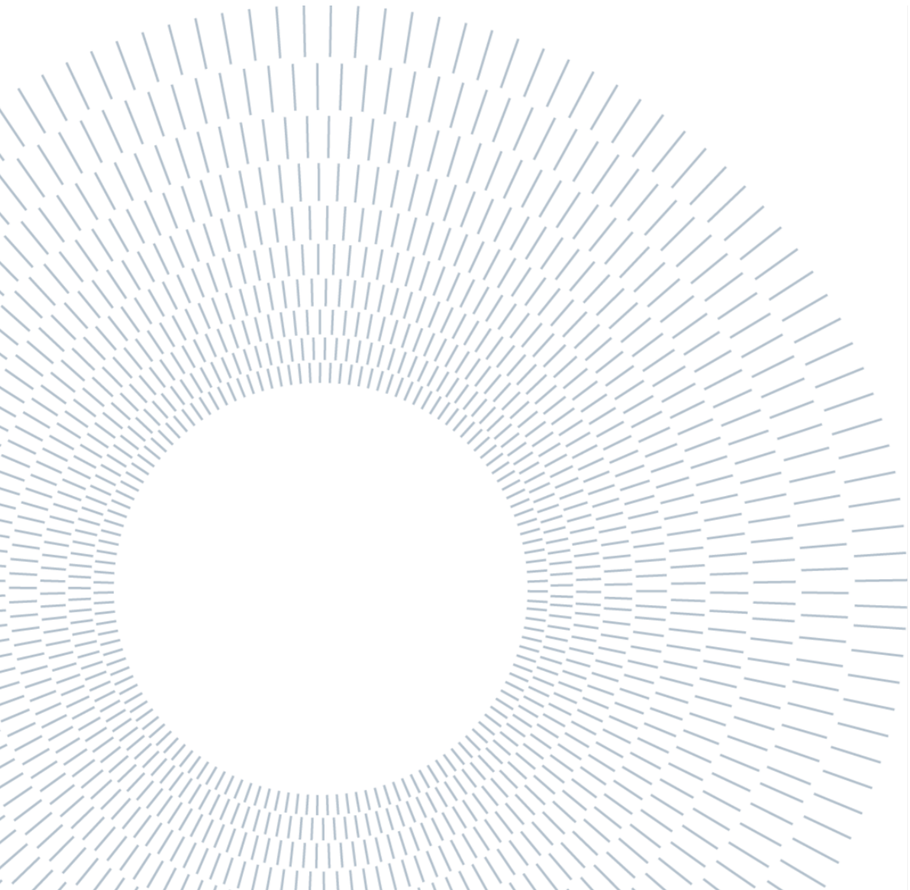
The goal of this work is that of designing a state feedback controller for a discrete-time time-invariant linear system, so as to optimize a given cost function while shaping the stationary state distribution to satisfy a probabilistic constraint. Specifically, we consider a system affected by a disturbance described as a strictly stationary stochastic process, and subject to nonconvex probabilistic constraints to its state variables and control inputs. These requirements are synthesized in a chance-constrained optimization program, where the optimization variables are the controller parameters. In recent years, a new methodology has been developed to solve this type of problems, which are otherwise very hard to assess. It is called the scenario approach and it has the merit of transforming the chance-constrained problem to a standard problem, more precisely the probabilistic constraints are substituted with a finite number of deterministic constraints corresponding to a certain number of randomly selected realizations of the disturbance. In this way, the computational burden is kept moderate. At the same time, however, the approach is complemented by a solid theory that keeps the approximation under control and provides a quantification of the feasibility of the obtained solution for the original probability constraint. In the present setup, however, the scenario approach alone is not sufficient to make the problem solvable, because it would require realizations of infinite length of the disturbance to reconstruct the stationary state. It is necessary thus to apply a truncation, to make the program feasible, and a tightening on the constraints, to compensate for the approximation introduced. The main contribution of this thesis is to account for all these elements and provide a non-trivial extension of the theory of the scenario approach that applies to the present non-standard setup. In addition to the theoretical part, a numerical example, together with the algorithm to solve it, is presented in the final part of the thesis.

Key-words: stochastic linear system, state feedback, nonconvex constraint, scenario approach, optimal chance-constrained control.

Abstract in lingua italiana

L'obiettivo di questa tesi è realizzare un controllore con feedback di stato per un sistema lineare a tempo discreto e tempo invariante, in modo da ottimizzare una data funzione costo e, allo stesso tempo, rimodellare la distribuzione dello stato stazionario, così da soddisfare un vincolo probabilistico. Più precisamente, consideriamo un sistema affetto da un disturbo definito come un processo stocastico strettamente stazionario, e soggetto a un vincolo probabilistico e non convesso sulle variabili di stato e sull'azione di controllo. Queste richieste sono raccolte in un programma di ottimizzazione soggetto a vincoli probabilistici, in cui le variabili da ottimizzare sono i parametri del controllore. Recentemente è stata sviluppata una nuova metodologia per risolvere questo genere di problemi, altrimenti molto difficili da trattare. Si chiama approccio a scenari e ha il merito di trasformare un problema soggetto a vincoli probabilistici in uno standard. Più precisamente, i vincoli probabilistici sono sostituiti con un numero finito di vincoli deterministici, corrispondenti a un certo numero di realizzazioni del disturbo selezionate casualmente. In questo modo, il carico computazionale è mantenuto limitato. Allo stesso tempo, però, l'approccio è sostenuto da una solida teoria che mantiene l'approssimazione sotto controllo e fornisce una quantificazione, per l'iniziale vincolo probabilistico, dell'attuabilità della soluzione ottenuta. In questa situazione l'approccio a scenari non è sufficiente, da solo, per poter risolvere il problema, in quanto richiederebbe una realizzazione infinitamente lunga del disturbo per ricostruire il processo stazionario. È perciò necessario applicare un troncamento per rendere il programma risolvibile e, per compensare l'approssimazione, deve essere applicato un irrigidimento dei vincoli. Il principale contributo di questa tesi è il tenere in considerazione tutti questi elementi e il fornire un'estensione, non banale, dell'approccio a scenari che si applichi al problema presentato. In aggiunta alla parte teorica, è stato sviluppato un esempio numerico e un algoritmo per risolverlo, presentati nella parte finale della tesi.

Parole chiave: sistema lineare stocastico, feedback di stato; vincoli non convessi, approccio a scenario, controllo ottimale soggetto a vincoli probabilistici



Contents

Introduction

General overview

In this thesis we consider the control of a linear discrete-time stochastic state-space system operating in stationary conditions. It is assumed that the state is measurable and the problem is that of designing a state feedback in order to optimize a given cost function while satisfying some constraints on the state (which can be interpreted as safety operating conditions) and on the control input (to prevent undesired effects due to saturation). Since the system is affected by a disturbance possibly with unbounded support, the evolution of the state is in turn stochastic and with unbounded support too, and for this reason the constraints on the state and on the control action (which is determined by the state since it is obtained by means of a feedback) cannot be deterministically imposed, for each realization of the state, as they would lead to an unfeasible problem. We therefore consider problems in which the constraints are imposed in probability, that is, by requiring that the constraints are satisfied for a fraction of the possible realizations of the state which nevertheless has a fairly large probability of occurring. The problem thus can be interpreted as that of selecting a proper feedback so as to optimally shape the state distribution in order to meet the conditions expressed by the given constraints. Since the system is considered in stationary conditions, the distribution of the state is time invariant and consequently it is sufficient to impose the constraint on a generic instant of time to guarantee its satisfaction for the whole time interval. From the practical side, the state feedback is designed off-line and then applied for every time instants. Given that asymptotic stability must be somehow imposed to accomplish the given goals, the stationary conditions will then be reached in the long run, with an exponential rate of convergence, depending on the maximum absolute value of the close-loop eigenvalues. By the law of large numbers, the state and the input will satisfy the constraints with a frequency that will tend to the chosen level of probability.

The shape of the stationary state distribution for control purpose is a task which is implicitly addressed in a number of control techniques. For example in minimum variance control, [1], generalized minimum variance control, [2], [3], [4], [5], [6] and [7], the shaping of the stationary state distribution is directed towards the minimization of the variance of some suitable output signal. In these approaches, however, it is very difficult to achieve a finer shaping as dictated by the satisfaction of state/input constraints. As a matter of fact, in these approaches it is not possible to include constraints, which are therefore accounted for only indirectly e.g. by introducing a control penalization term to the variance in the cost function. Another approach that instead allows the user to consider constraints that are imposed in probability is Stochastic Model Predictive Control (SMPC, see e.g. the extensive survey [8]), where the feedback control is achieved by means of the so-called receding horizon technique. In this case however on-line computations are needed to design the controller and the designed controller is typically outside the realm of linear state-feedbacks as consider here. Moreover, an exact analysis of the stationary behavior of the control scheme is difficult to achieve and there are no results in the literature on the optimality of the SMPC solution and on the satisfaction of probabilistic state/input constraints in the long run. Eventually, a work this thesis is in the vein of is the recent publication [9], where a similar problem is addressed. [9] however deals with a peculiar situation, complementary to the one studied here, where the state is not accessible, while the disturbance is so. The consider control problem thus is that of designing an open-loop compensator, instead of a more commonly encountered in practice state feedback controller as considered here. In a sense, this thesis extends the methodology of [9] to the important class of state-feedback controllers, but it is important to notice that the extension is highly nontrivial since state-feedback implies a number of extra difficulties. One above all, the dependence of the dynamics on the controller parameters naturally leads to non-convex problems.

The inclusion of probabilistic constraints makes the computation of the controller parameters a chance-constrained optimization program. This type of problem is particularly hard to solve, and generally it is not possible to tackle it with the common design approach. Recently, a new approximate method, called the scenario approach, has been developed, [10] and [11]. The scenario approach allows to solve chance-constrained problems starting from samples of the uncertain affecting the system. The main idea behind it is to solve the problem with deterministic constraints, instead of the probabilistic ones, for a finite number of instances. It has been proved that the solution of this problem can be then generalized to the chance-constrained one, with guarantees on the However, in this case, the complexity introduced by the presence of probabilistic constraints is increased by the fact that they regard the state in stationary conditions. This involves the whole disturbance process, with the consequence that the constraints can not be treated directly applying analytic methods. In this case, the use of the scenario approach alone is not possible, since it would require infinitely-long realizations of the disturbance to compute the optimization. Following a path similar to the one presented in [9], the problem is tackled using the scenario approach in combination with an approximation of the stationary process. In fact, in order to apply the scenario approach on the problem we are analyzing, the stationary state process has to be approximated with a truncated version. To compensate for the error introduced by the approximation, a tightening on the constraints has to be enforced. We also need to introduce a bound on the norm of the controlled system state matrix, to guarantee that the difference between the approximated state and the stationary one is lower than a desired value.

The assumption on the convexity of cost functions and constraints, present in [9], is removed, so to include a wider range of functions. This generalization, however, leads to the impossibility to determine a priori a bound on the satisfaction of the probabilistic constraints. That bound can be still, assessed a posteriori, once the solution has been computed, and it can therefore be used to determine the overall quality of the controller and if it satisfies the user needs.

The parameters computation is done entirely off-line, using data collected from the system, and the optimality of their performance is guaranteed by the method applied. The use of the scenario approach guarantees indeed the feasibility and the satisfaction of the probabilistic constraints when operating in stationary conditions.

Contribution

The problem presented in [9] has been reviewed using a state-feedback controller instead of the disturbance compensator. The presence of a state-feedback controller is generally a more common choice, since it happens with higher frequency to have the state measurements available than the disturbance ones. So, although for specific situations a disturbance compensator is the only choice, the state-feedback controller has a wider range of applications. The extension of the result found in [9] to state-feedback controllers, however, faces some obstacles, such as more intricate dependence of the state from the controller parameters.

The problem has also been extended to include nonconvex constraints, removing the assumption that required the cost function and the constraints to be convex. In this way, the only requirement for the functions is to be Lipschitz continuous, which is not a particularly restrictive condition.

The path followed to solve the problem is similar to the one presented in [9], but revised and expanded to deal with the state-feedback and the nonconvexity of the functions. It is based on the application of the scenario approach in combination with an approximation of the stationary process. In practice, the problem is twice simplified: first it is transformed from a chance-constrained program to a standard one, secondly it is approximated by truncating the stationary state process.

The main contribution is the extension of the guarantees about the solution, provided for the approximated scenario problem, to the original chance-constrained problem. To do that, it has been necessary to blend the scenario approach with an approximation on the stationary state. The combination of the two approaches allows to address the two main difficulties related to the problem, making it solvable by available optimization solvers.

When the approximation was applied, it required a tightening on the constraints and also a limit on the difference between the stationary state process and the truncated one. To maintain under control the value of the difference, a bound on the norm of the controlled system state matrix was introduced. Since this was not necessary in previous works with the scenario approach, we had to study the possible typologies that can be used. Three different versions are presented, with an analytical comparison between the best two.

The methodology developed has also been tested, simulating a practical case. The tests have been conducted under many different conditions, both with stable and unstable systems, to check that the validity of the approach and to show its potentialities.

Structure

In the first chapter, we formulate the problem, expressing mathematically the state equation and the state-feedback. Then, we present the scenario approach, together with the main related concepts, indispensable to understand it, and its results. In the final section, we try to apply the scenario approach to the problem presented. Here emerges the unfeasibility of the program due to the presence of the stationary state process, which relies on infinitely-long realizations.

In practice, it is not possible to collect or generate an infinite number of disturbance measurements, and it wouldn't even be possible to work with that. To address this problem, in the second chapter we introduce an approximation of the stationary state process, based on a finite-long realization of the disturbance. To guarantee that the approximation is negligible, we introduce a condition on the controlled system, so that the difference between the stationary state process and the approximated one is kept under control. We also introduce a tightening on the constraints to compensate for the error caused by the process approximation. We speak about data collection and, in the final section, we introduce the mathematical process to extend to the original problem the guarantees obtained for the approximated version.

In the third chapter, three different bounds for the norm of the controlled system state matrix are introduced. The second and the third are then compared, to evaluated which one, for the same result, is less restrictive.

Finally, in the fourth chapter, we put together all the concepts presented in the previous chapters. We define the final problem and state and prove its solution guarantees, using the theorems and definitions given in the first and second chapters. In the last section, we discuss the choice of the parameters, explaining which considerations must be done when deciding their values.

The fifth chapter presents three algorithm that can be used to determine a support set. The advantages and disadvantages of each are highlighted and used to compare them.

Two simulation examples are presented in the final chapter, to show the main results obtained by this methodology and to give a practical example of the application of the theory exposed.

1. Chance-constrained problem and the scenario approach

1.1 Problem formulation

The thesis addresses the problem of controlling via state function the motion of a discrete-time time-invariant linear system affected by an additive stationary disturbance. The equation describing the evolution of the system is:

$$x_{k+1} = Ax_k + Bu_k + d_k \,. \tag{1-1}$$

The state belongs to a real space with dimension n_x, so that $x_k \in \mathbb{R}^{n_x}$, while the control input has dimension n_u, that is, $u_k \in \mathbb{R}^{n_u}$. The system is also affected by an additive stochastic disturbance for each state variable, $d_k \in \mathbb{R}^{n_x}$. A and B are matrices of appropriate dimensions.

The stochastic process d_k is assumed to be strictly stationary with zero mean and well-defined and known second order moment. The requirement on the zero value of the mean is without any loss of generality. If indeed one is dealing with a stochastic process with a non zero mean, then one can work with an unbiased version of the state. Specifically, introduce the term $\bar{x}_k$ subject to

$$\bar{x}_{k+1} = A\bar{x}_k + W\bar{d} \,, \tag{1-2}$$

where $\bar{d} = \mathbb{E}_d[d_k]$. The problem can be then reformulated as the difference $\Delta x_k = x_k - \bar{x}_k$, whose evolution is described by the equation

$$\Delta x_{k+1} = A\Delta x_k + Bu_k + \Delta d_k \,, \qquad \text{1-3}$$

where Δd_k is the difference between the disturbance d_k and its mean $\bar{d}$ and is therefore a zero mean process. As is clear, controlling Δx_k is completely equivalent to control x_k.

The state is assumed to be observable and the control action is computed as a state feedback, obtained by multiplying the state vector with a matrix gain K, plus a control offset term γ:

$$u_k = \gamma + Kx_k \,. \qquad \text{1-4}$$

The parameters values are taken from the sets $\Gamma \subset \mathbb{R}^{n_u}$ and $\mathrm{K} \subset \mathbb{R}^{n_u \times nx}$. Combining the equation 1-1 with the equation 1-4, we obtain the equation for the controlled system:

$$x_{k+1} = (A + BK)x_k + B\gamma + d_k \,. \qquad \text{1-5}$$

Under the assumption made on the disturbance, for any $k \in \mathbb{Z}$ there exists a measurable function $x_{k,\infty}$ of the process $d_{k-1} = \{\ldots, d_{k-2}, d_{k-1}\}$ such that the process $x_\infty = \{x_{k,\infty}, \mathrm{k} \in \mathbb{Z}\}$ is strictly stationary and with finite first and second order moments. This $x_{k,\infty}$, referred to as the stationary state process, is unique and it is given by the following equation:

$$x_{k,\infty} = (I - A - BK)^{-1}B\gamma + \sum_{s=0}^{\infty}(A + BK)^s d_{k-1-s} \,, \qquad \text{1-6}$$

where convergence of the right-hand side is meant in the mean square sense. The goal is to choose the parameters γ and K in order to optimize a certain cost function while satisfying, at the same time, state and input probabilistic constraints. Specifically, the problem is formulated as a chance-constrained optimization program:

$$\min_{\gamma,K,h} h \qquad \text{1-7}$$
$$s.t.\, \mathbb{P}_{d_k}\{l(x_{k,\infty}, \gamma + Kx_{k,\infty}) \leq h \wedge f(x_{k,\infty}, \gamma + Kx_{k,\infty}) \leq 0\} \geq 1 - \varepsilon.$$

$l(x, u)$ is a function that associate a cost to the state/control input pair or the parameters that determine the control variable. $f(x, u) = \mathbb{R}^{n_x}\mathrm{x}\mathbb{R}^{n_u} \rightarrow \mathbb{R}$ is the constraint applied, which can bound both the state and the control input. Both the cost function and the constraint are evaluated in stationary conditions, when the state x is equal to $x_{k,\infty}$ and the control action u is the state compensator given applying the control equation 1-4 on the stationary state.

Throughout we will work under the following assumption:

Assumption (Lipschitz continuity): the cost function $l(x,u)$ and the constraint function $f(x,u)$ are Lipschitz continuous in $x \in \mathbb{R}^{n_x}$ and in $u \in \mathbb{R}^{n_u}$ with Lipschitz constant, respectively, L_1 and L_2.

The problem presented in 1-7 is particularly hard to solve, due to the presence of probabilistic constraints and the fact that $l(x_{k,\infty}, \gamma + Kx_{k,\infty}) \leq h$ and $f(x_{k,\infty}, \gamma + Kx_{k,\infty}) \leq 0$ are nonconvex. It is therefore necessary to introduce an approximate method to address it.

1.2 The scenario approach

The scenario approach is a general methodology presented for the first time by Calafiore and Campi in [10] and [11]. It was introduced to solve chance-constrained problems where a linear objective is minimized under a probabilistic condition over a convex constraint, parametrized by uncertainty terms. The presence of uncertainty is often encountered in practical problems, making them harder to solve (in many cases, even NP-hard). The word scenario indicates a sampled realization of the uncertainty affecting the system. The idea behind this framework is to rely on a finite number of randomly sampled scenarios and to solve the correspondent convex problem only for these selected cases. In this way, the initial problem is transformed from a chance-constrained control design problem to a standard convex problem, which can be solved in one shot using standard program solvers. One of the main advantages of this method is the possibility to know a-priori the level of probabilistic guarantees of robustness as a function of the number of the scenarios used. It is therefore possible to know the number of scenarios needed to attain the desired level of robustness before computing the optimization.

The general methodology has then been refined by many contributions, exploring possible applications in system and control design, [9], [12], [13] and improving the quality of the results, see e.g. [14], [15], [16], [17], [18]. Nowadays, the theory that certifies the generalization properties has achieved great recognition and the scenario optimization is well understood for convex problems. Recently its validity has been extended also for problems subject to nonconvex constraints [19] [20] [21], which were not included in the previous studies. This requires a paradigm shift, since it is not possible to know a-priori how many constraints will be necessary to guarantee a certain level of probabilistic robustness. However, the theory provides a formidable tool to a-posteriori assess the achieved level of robustness.

In the following subsections, we present an abstract setup to introduce all the necessary elements to understand the scenario approach and then we state the theorem we will use for problem 1-7.

1.2.1 Problem formulation

The starting point is an optimization program, in which the goal is to minimize a linear objective function $c^T\theta$. θ is the optimization variable belonging to the decision space Θ with dimension $d: \theta \in \Theta \subseteq \mathbb{R}^d$, and in the context of control problems it includes the parameters of the controller and possibly other additional parameters. c represents the weights given to the single values among the optimization parameters when computing the cost.

The system our solution is supposed to operate with is affected by uncertainty, which is described through a set Δ and a probability distribution $\mathbb{P}$.

The set Δ is a space containing all the admissible situations. Generally, Δ is a continuous set, so it has an infinite cardinality The probability distribution $\mathbb{P}$ can have different interpretation, based on the problem given. Indeed, it can be the function representing the likelihood of the different situations present in Δ to occur, or it can be a way to attribute different value of importance to the uncertainty instances. The instances δ belongs to the probability space Δ and they can be infinite in number if the set is continuous.

Uncertainty enters the problem because θ, our optimization variable, has to address the satisfaction of the constraints $r(\theta, \delta) \leq 0$, where $r(\theta, \delta): \Theta \times \Delta \rightarrow [-\infty, \infty]$ is a scalar-value function that specifies the constraints we have to enforce (multiple constraints can be reduced to a single scalar-value function by the use of the max operator, see [11]).

A first way to address uncertainty is the worst-case approach where the satisfaction of the constrains is enforced for every $\delta \in \Delta$. This leads to the problem, where we want to minimize a linear cost function $c^T\theta$, where θ is subject to $r(\theta, \delta) \leq 0 \; \forall \delta \in \Delta$. Notice that other studies following [10] and [11] removed the necessity to have a linear function to minimize, using instead a more generic cost function and requiring only it to be convex. Take as examples [9], discussed in the introduction, or [16]. In the first example the function depends on the state, the control input and the disturbance. In the second one, it depends on the optimization variables and the uncertainty parameter.

If the set Δ is continuous, the worst-case problem is a semi-infinite optimization problem, because the number of optimization variables is finite while the number of constraints is infinite. This type of problem is generally difficult to solve and moreover it suffers from conservativism. As a matter of fact, in the worst-case approach, the constraints $r(\theta,\delta) \leq 0$ are enforced for all the possible values $\delta \in \Delta$, with the result that few ill situations can determine a system control with an extremely high cost, as it is shown in [22]. To reduce the conservatism of the worst-case approach, probabilistic robust design has been introduced which is meant to compromise between minimizing the objective function and satisfying most of the constraints. The probabilistic relaxation can be interpreted in this way: instead of satisfying all the constraints, the aim becomes to satisfy all but a small fraction of them, whose probability is no larger than a certain level $\varepsilon \in (0,1)$ (see [11]). Removing some constraints increases the number of possible solutions and therefore the possibility to find solutions with improved cost. It is common experience that even for small values of ε huge improvements for the cost are obtained over the worst-case solution. We now better formalize probabilistic robust design by introducing the relevant notions.

The set of instances $\delta \in \Delta$ for which the constraints are not satisfied is called violation set, leading to the probability of violation, taken from [11].

Definition 1 (Probability of violation): let $\theta \in \Theta$ be given. The probability of violation of θ is defined as:

$$V(\theta) \doteq \mathbb{P}\{\delta \in \varDelta : r(\vartheta,\delta) > 0\}.$$

To better understand the probability of violation, it can be useful to look first at the case where the probability space Δ is not continuous. In this case, $V(\theta)$ quantifies how many instances, on the total, do not satisfy the associated constraints when the optimization variable has value θ. To measure it also in the continuous case, the probability measure $\mathbb{P}$ is used.

From [11] we have also the definition of ε-level feasibility (also called ε-feasibility).

Definition 2 (ε-level): let $\varepsilon \in (0,1)$. We say that $\theta \in \Theta$ is an ε-level robustly feasible (or simply ε-level) solution, if $V(\theta) \leq \varepsilon$.

In probabilistic robust optimization, the goal is to find a solution that minimize the objective function while guaranteeing a certain ε-feasibility. The problem can be then formulated in a general way as follows:

$$\min_{\theta} c^T\theta \qquad \text{1-8}$$
$$subject\ to: \mathbb{P}\{r(\theta,\delta) \leq 0\} \geq 1-\varepsilon.$$

As ε increases in value, progressively more constraints are discarded and, therefore, the feasibility region is enlarged, increasing the number of admissible solutions.

Problem 1-8 is also called chance-constrained problem, where the set of the neglected constraints is chosen in an optimal way, i.e., the constraints not considered are the ones which allows the greatest improvement of the solution in terms of cost. Solving a chance-constrained problem is, however, very hard, given that 1-8 is nonconvex even though $r(\theta,\delta)$ is convex in θ for all δ. For this reason approximation need to be introduced and one effective, which allows the user to keep control on the violation, is the scenario approach.

1.2.2 Scenario approach

The scenario approach, as introduced at the begin of this section, is a method to deal with chance-constrained problems as in 1-8. The fundamental idea behind the scenario approach is to consider just a finite number of instance of δ (scenarios) and to treat them as rigid constraint, without having to deal with the probability distribution of the disturbance and the probabilistic constraints. In fact, one of the strengths of this method is that it does not require, neither assume, any knowledge on $\mathbb{P}$, [17]. The scenarios are, indeed, sampled from the mechanism generating uncertainty, according to the probability, but without the need to know it. If the probability is known, it can be used to draw the scenarios from a model, with potentially great savings in terms of time and resources. In both the cases, all the theoretical results hold independently from $\mathbb{P}$.

Once N realizations of δ are extracted, say $\left(\delta^{(1)},\ldots,\delta^{(N)}\right)$, the scenario approach amounts to solve the following program:

$$\min_{\theta} c^T\theta \qquad \textit{1-9}$$
$$subject\ to\ r\left(\theta,\delta^{(i)}\right) \leq 0 \ \ i = 1,\dots,N$$

whose solution will be denoted by θ_N^*. Notice that *1-9* defines a so called family of decision maps, [17], which are indicated as $M_N: \Delta^N \to \Theta,\ N = 0{,}1{,}2{,}\dots,$ i.e., they are functions which go from the N-dimensional uncertainty domain Δ^N to the decision space Θ. In practice, a decision map indicates that the scenario solution θ_N^* is a decision calculated from a scenario set $\delta^{(1)},\dots,\delta^{(N)}$ according to *1-9*: $\theta_N^* = M_N\left(\delta^{(1)},\dots,\delta^{(N)}\right)$.

It is straightforward to verify that the decision maps M_N satisfy the following property.

Property 1 (consistency): for every non-negative integers N and n, and for every choice $\delta^{(1)},\dots,\delta^{(N)}$ and $\delta^{(N+1)},\dots,\delta^{(N+n)}$, the following three properties hold:

i. If $\delta^{(i_1)},\dots,\delta^{(i_N)}$ is a permutation of $\delta^{(1)},\dots,\delta^{(N)}$, then it holds that $M_N\left(\delta^{(1)},\dots,\delta^{(N)}\right) = M_N\left(\delta^{(i_1)},\dots,\delta^{(i_N)}\right)$;
ii. If $r\left(\theta_N^*,\delta^{(N+i)}\right) \leq 0$ for all *i*=1,…,*n*, then it holds that $\theta_{N+n}^* = M_{N+n}\left(\delta^{(1)},\dots,\delta^{(N+n)}\right) = M_N\left(\delta^{(1)},\dots,\delta^{(N)}\right) = \theta_N^*$;
iii. If $r\left(\theta_N^*,\delta^{(N+i)}\right) > 0$ for one or more *i*=1,…,*n*, then it holds that $\theta_{N+n}^* = M_{N+n}\left(\delta^{(1)},\dots,\delta^{(N+n)}\right) \neq M_N\left(\delta^{(1)},\dots,\delta^{(N)}\right) = \theta_N^*$;

This property is indicating that M_N is permutation-invariant, so it leads to the same solution despite the order in which the scenarios are disposed. The second point demands that, given a set of scenarios and its corresponding solution, if other scenarios are added to a set and the scenario solution is feasible for them, then the solution does not change. Finally, the third point is complementary to the second: if the solution is not feasible for at least one of the added scenarios, then it can not be the solution for the enlarged scenario set. Property 1 is key for the theory of [17] to hold true. This theory will be presented next. Even though our discussion will be limited to program 1-9, notably the theory of [17] applies for the generic decision maps satisfying Property 1. This generality leaves open the possibility of extending the results of this work to algorithms other than 1-9 to approximately solve 1-7.

1.2.3 Support constraints and irreducible support set

The solution θ_N^* to problem 1-9, found using the scenario approach, depends on the extracted scenarios and is therefore a random variable. Consequentially, also the probability of violation $V(\theta_N^*)$, due to its dependence on the decision, is a random variable.

Now we would like to determine what guarantees can be provided on the ε-feasibility of the solution θ_N^*. To determine the violation probability, however, it is first necessary to introduce the concept of support subsample and irreducible support subsample.

Definition 3 (support subsample): given a sample of scenarios $\left(\delta^{(1)},\ldots,\delta^{(N)}\right)$, a support subsample (or support set) S is a k-tuple of elements taken from $\left(\delta^{(1)},\ldots,\delta^{(N)}\right)$, i.e., $S=\left(\delta^{(i_1)},\ldots,\delta^{(i_k)}\right)$ with $i_1<i_2<i_k$, whose solution is the same as the one obtained with the original sample:

$$M_k\left(\delta^{(i_1)},\ldots,\delta^{(i_k)}\right)=M_N\left(\delta^{(1)},\ldots,\delta^{(N)}\right)$$

A support set is said to be irreducible if removing any element from it would modify the solution. It is important to notice that, in general, for the same sample $\delta^{(1)},\ldots,\delta^{(N)}$, it is possible to find more than one irreducible support set. The cardinality s_N^* of a support set is the number of elements and it will be called complexity of the support set. Following the definition given in [17], the complexity of *1-9* is the smallest cardinality of a support subsample and it can be equal to zero in the case the support set is void.

Once the set $\delta^{(1)},\ldots,\delta^{(N)}$ is given, it is possible to determine the complexity according to Definition 3, so s_N^* is an observable quantity. Finding an irreducible support set is a simple task, it is sufficient for example to eliminate one scenario at time and check whether the solution changes or not. If yes the scenario is kept in the support set, otherwise is discarded However, finding the one with minimal cardinality can be a truly formidable task, not always feasible in terms of computational effort.

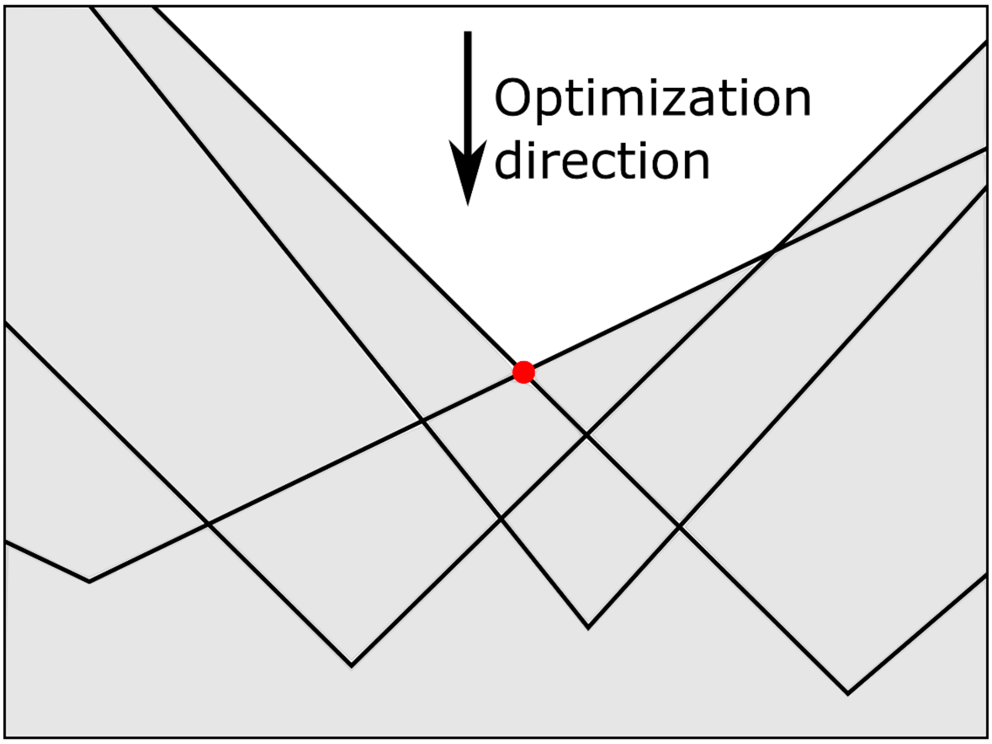


Figure 1-1 The image represents a scenario program characterized by V-shaped constraints. If one of the two support constraints is removed, the solution improves.

When treating convex problems, the maximum value for the complexity is known even before computing the solution the complexity. In fact, thanks to the convexity, the minimum number of support samples necessary to determine the solution is, at most, equal to the number of optimization variables, therefore $s_N^* \leq d$. Looking to figure 1-1, where there are two optimization variables, it is clear that there are only two support constraints, all the other can be removed without affecting the solution. It could be possible to have just one support constraint, if the lowest point for optimization coincides with the angular point of one of the constraints. On the other hand, it is impossible to have an irreducible support set containing three or more constraints.

For nonconvex problem, on the other hand the cardinality of the support set can be computed only after knowing the solution and it is not possible to bound it with the number of optimization variables, as done with convex constraint. From figure 1-2, it is possible to see an example where all the constraints belong to the support set. In fact, the solution is determined by the intersection of just two constraints but removing any other one would lead to a new optimal point.

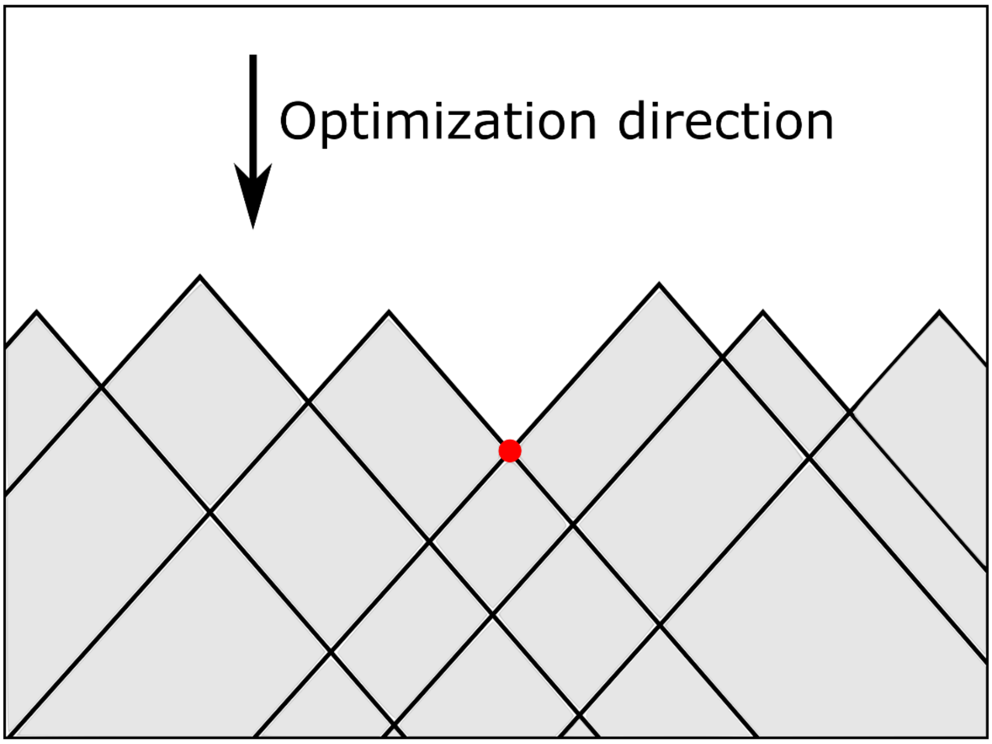


Figure 1-2 This figure represents a nonconvex problem suitable for explanation's sake. This is indeed a particular case in which the support subsample contains all the present constraint.

1.2.4 Results of the scenario approach

The main result of the scenario approach, [15], [17], [18], is that there is a relation between the unknown probability of violation $V(\theta)$ and the complexity s_N^* of 1-9, so that s_N^* can be used to bound the value of $V(\theta)$. The more general result has been proven in [17] which in particular applies for all nonconvex scenario programs irrespective of . The result is as follows:

Theorem 1: Given a confidence parameter $\beta \in (0,1)$, for any $k = 0,1,\dots,N-1$ consider the polynomial equation in the v variable

$$\binom{N}{k}(1-v)^{N-k} - \frac{\beta}{N}\sum_{m=k}^{N-1}\binom{m}{k}(1-v)^{m-k} = 0, \tag{1-10}$$

and let $\varepsilon(k)$ be the unique solution over the interval (0,1). Also define $\varepsilon(N) = 1$. For any $\mathbb{P}$ it holds that

$$\mathbb{P}\{V(\theta_N^*) > \varepsilon(s_N^*)\} \leq \beta. \tag{1-11}$$

Where N is the number of scenarios used, $\theta_N^* = M_N\left(\delta^{(1)},\ldots,\delta^{(N)}\right)$ and s_N^* is the complexity of 1-9, i.e. it is the smallest cardinality of a support set.

What the theorem says is that, if N scenarios $\delta^{(1)},\ldots,\delta^{(N)}$ are randomly extract, the solution θ_N^* is ε-feasible with a high probability $1-\beta$, irrespective of the problem at hand and the probability distribution $\mathbb{P}$.

The role of the confidence parameter β is marginal, since ε depends logarithmically on β, so it is possible to set it equal to a very low value, such as 10^{-7}. In this way, the probability violation is bounded by $\varepsilon(s_N^*)$ with practical certainty and a usable upper estimation to assess $V(\theta_N^*)$ is given.

Thanks to Theorem 1, we can evaluate the ε-feasibility of the solution θ_N^* computed using the scenario approach and hence we can assess the feasibility for the chance-constrained problems of type 1-8. To determine the function $\varepsilon(\cdot)$, it is necessary to solve numerically equation 1-10. In Appendix A, an algorithm based on bisection is provided.

If an explicit result is needed, one can use:

$$\tilde{\varepsilon}(k) := \begin{cases} 1 & \text{if k = N,} \\ 1 - \sqrt[N-k]{\dfrac{\beta}{N\binom{N}{k}}} & \text{otherwise,} \end{cases} \tag{1-12}$$

in place of ε, see While $\tilde{\varepsilon}$ is explicit and ready to use, the provided evolution is loose as compared to that of ε computed by solving equation 1-10. So it is suggested to use the numerical solution of 1-10 whenever is possible. Importantly, since $\varepsilon(k)$ is monotonically increasing with k, it must be noticed that Theorem 1 continues to hold if s_N^* is taken as the cardinality of any support set, not necessarily a minimal one or even an irreducible one. However, finding the smallest support subsample leads to better guarantees on the result reliability, so it is crucial being able to identify small support set. In chapter 5 three searching algorithms are presented to this purpose.

Eventually, notice that for convex problems, since $s_N^* \leq d$, it holds that $\mathbb{P}\{V(\theta_N^*) > \varepsilon(d)\} \leq \beta$, a result that allows one to decide the requirement number of scenarios N to achieve that the violation $V(\theta_N^*)$ is below a given threshold with high confidence. This result however applies for convex problem only and it is therefore of no use in this work.

1.3 Naïve application of the scenario approach to problem 1-7

We can apply the scenario approach to the problem 1-7 described in the first section to deal with the probabilistic constraints. As explained in the subsection 1.2.3, the value of *N*, i.e., the number of scenarios, required to obtain desired ε-level for the violation can not be determined before computing the solution. To deal with this type of problem is therefore necessary to treat ε as a target value, and not as a strict requirement. The choice about the value of *N* will be discussed in chapter 4, for now a generic value is taken. The optimization variables we are considering are the matrix gain K, the control offset γ and the value h (so the solution will be $\theta = (\gamma, K, h)$). The scenarios are unilaterally infinitely long realizations of the disturbance, i.e., $\delta = \{d_k, d_{k-1}, \dots, d_{k-j}, \dots\}$. To use the scenario approach, we assume to extract/collect *N* infinitely long realizations $\left\{d_k^{(i)}, d_{k-1}^{(i)}, \dots, d_{k-j}^{(i)}, \dots\right\}$ $i = 1, \dots, N$.

The scenario version of the problem 1-7 is then formulated as follows:

$$\min_{\gamma, K, h} h \tag{1-13}$$

$$\text{subject to } l\left(x_{k,\infty}^{(i)}, \gamma + K x_{k,\infty}^{(i)}\right) \leq h$$

$$f\left(x_{k,\infty}^{(i)}, \gamma + K x_{k,\infty}^{(i)}\right) \leq 0$$

$$x_{k,\infty}^{(i)} = (I - A - BK)^{-1} B\gamma + \sum_{s=0}^{\infty} (A + BK)^s d_{k-1-s}$$

$$i = 1, \dots, N.$$

Given a solution θ_N^* and given the cardinality s_N^* of a support set, it is possible to state, using theorem 1, that the violation probability for the original problem 1-7 is smaller than $\varepsilon(s_N^*)$ with confidence $1 - \beta$. This could be a satisfactory solution for our problem, but there is a major issue.

The problem expressed in **Errore. L'origine riferimento non è stata trovata.** can not be solved in practice because it requires N infinitely-long realizations of d_k to obtain $x_{k,\infty}^{(i)}$. Indeed, only finite-length realizations of the disturbance are available in practice, so a further step must be done to make the scenario problem applicable in practice. As is clear, the guarantees obtained with the scenario approach are particularly interesting, so we do not want to lose them when modifying the approach. In the next chapters, we will explore the idea to make problem 1-13 solvable by approximating the stationary state process $x_{k,\infty}$ with a truncated version. The approximation obviously introduces an error, so it will be necessary to guarantee that the difference between $x_{k,\infty}$ and the truncated version is negligible and also to apply some restriction to maintain the guarantees obtained through theorem 1. At the end, an approximated version of problem 1-13 will be obtained establishing also a link with the initial problem 1-7 thanks to the guarantees provided by the scenario approach.

2. The proposed scenario optimization

Thanks to the scenario approach, it has been possible to deal with the probabilistic constraints of the original problem formulation 1-7. The proposed method has opened an interesting resolution path, but the obtained scenario problem is still not feasible, since it depends on realizations of the disturbance which are infinitely long. In practice, it is possible to work only with finitely-long realizations, therefore it is necessary to introduce an approximation of the stationary state process. Unfortunately, the error introduce by the approximation prevents one to directly use the existing scenarios theory for the evolution of the violation of the solution. A specific theory will be developed to this purpose, as a major contribution of the thesis.

2.1 The approximated scenario problem

As said, the impossibility to deal with infinite-long realizations of the disturbance as in 1-13 demands the introduction of an approximation of $x_{k,\infty}$. Given an integer $M \in \mathbb{N}$, we will consider a truncated version of $x_{k,\infty}$ as follows

$$x_{k,M} = (I - A - BK)^{-1}B\gamma + \sum_{s=0}^{M-1} (A + BK)^s d_{k-1-s} \,, \qquad \text{2-1}$$

which requires only a finitely-long realization of the disturbance $\{d_{k-1}, \dots, d_{k-M}\}$. The truncation is influenced by the data availability and the complexity that our solver can deal with. The approximation of $x_{k,\infty}$ with $x_{k,M}$ improves as M increases towards infinity. Therefore, to obtain a result as close as possible to the one guaranteed by the scenario approach, M should be chosen as high as possible, considering the available data and the computational load. The truncated version of the scenario program 1-13 requires to generate/collect N infinitely-long realizations of the disturbance $\left\{d_{k-1}^{(i)}, \dots, d_{k-N}^{(i)}\right\}, i = 1, \dots, N$ and then to enforce the constraints for the truncated state $x_{k,M}$ corresponding to these realizations. However, to compensate the error introduced by the truncation, we need to introduce a tightening in the constraints. While the scope of a tightening is not clear here, its role will become obvious later, during the development of the result. The new problem is equal to the scenario problem 1-13 except for the tightening term δ and the use of the truncated version of the state:

$$\min_{\gamma, K, h} h \tag{2-2}$$

$$\text{subject to } l\left(x_{k,M}^{(i)}, \gamma + K x_{k,M}^{(i)}\right) \leq h - \delta$$

$$f\left(x_{k,M}^{(i)}, \gamma + K x_{k,M}^{(i)}\right) \leq -\delta$$

$$x_{k,M}^{(i)} = (I - A - BK)^{-1} B\gamma + \sum_{s=0}^{M-1} (A + BK)^s d_{k-1-s} d_{k-1-s}^{(i)}$$

$$i = 1, \dots, N.$$

The solution will be denoted by γ^*, K^*, h^*. This problem corresponds to a standard optimization program, and it is possible to solve it via available solvers.

The computational effort needed to solve 2-2 depends on both N and M. The greater the value of M, the greater is the problem complexity. In fact, the general cost function and the constraint are applied to $x_{k,M}$, which depends non-linearly on the optimization variables γ and K. In particular, the relationship with K depends on the value of M, which appears at the exponent of a matrix containing K. So, even in the case that the distribution probability $\mathbb{P}$ is known and it is possible to have an unlimited quantity of data, the computational complexity gives a limit to the value of M. Furthermore, as we will see later, the values of M and δ determine, together with other parameters, the guaranteed probability violation, so there is an incentive in trying to maximize M and minimize δ. For the choice of the number of scenarios N, more discussion will be given in chapter 4, where there will be a broader understanding of the problem. We anticipate indeed here that 2-2 is not yet the final version of the problem, because another condition will be introduced.

2.1.1 Types of data collection and possible limitation

The data used for the disturbance realizations can be collected in two ways. In the first case, the probability distribution $\mathbb{P}$ is known and the data required is generated according to it. In this case, it is easy to generate large quantity of data without significative cost. In the second case, the probability distribution is not known, but it is possible to sample the system to collect the data. It can happen, for certain configurations, that is not possible to sample directly the disturbance, with the measurements being available only for the state. However, it is easy to obtain the needed quantity from the measurable ones by using equation 1-1:

$$d_k = x_{k+1} - Ax_k - Bu_k \qquad 2\text{-}3$$

When sampling the disturbance (or the state), it is important to keep the system stable. If the state matrix A is stable or asymptotically stable, there is no need of a control action, so it is possible to just impose $u_k = 0\ \forall k$. If, on the other hand, the state matrix A is unstable, it is necessary to implement a control system. It does not have to be complex or optimal, the only requirement is to counteract the instability of the state. A possible simple solution is using a simple proportional controller $u_k = -Kx_k$. Then, the state equation becomes:

$$x_{k+1} = (A - BK)x_k + d_k \qquad 2\text{-}4$$

The matrix K has to be such that the matrix $A - BK$ is stable or asymptotically stable.

If the data needed have to be sampled from a real system, the quantity that can be collected is limited by many factors. First, the time needed by the system to evolve and the one needed to sample. Second, the cost to run the system, both in terms of resources used and human labor needed. For these reasons, the case in which the probability is known is generally preferred, but the second case is more frequent.

2.2 Violation assessment of the solution of 2-2

A main problem here is to evaluate the feasibility of the solution to 2-2 which respect the original chance-constrained problem 1-7. Given the approximated problem 2-2 and its solution (γ^*, K^*, h^*), we can determine a support set and assess its violation probability, with respect to the constraints over the truncated state. Specifically, the notion of violation is related in the present context to

$$\mathbb{P}\{\delta \in \Delta : l(x,u) \le h \wedge f(x,u) \le 0\}. \qquad 2\text{-}5$$

To simplify the notation, we introduce the function $g(x,u)$ defined as

$$g(x,u) = max\{l(x,u) - h, f(x,u)\}, \qquad 2\text{-}6$$

and then, the following equivalence is true:

$$l(x,u) \le h \wedge f(x,u) \le 0 \Leftrightarrow g(x,u) \le 0. \qquad 2\text{-}7$$

So, considering the solution (γ^*, K^*, h^*), we have that

$$g^*_{k,M} = max\{l(x^*_{k,M}, \gamma^* + K^* x^*_{k,M}) - h^*, f(x^*_{k,M}, \gamma^* + K^* x^*_{k,M})\}, \qquad 2\text{-}8$$

where $x^*_{k,M}$ is the truncated version of the state when the controller γ^*, K^* is used

$$x^*_{k,M} = (I - A - BK^*)^{-1} B\gamma^* + \sum_{s=0}^{M-1} (A + BK^*)^s d_{k-1-s}\,. \qquad 2\text{-}9$$

Then, thanks to properties of the scenario approach, knowing the cardinality s^* of the support set to 2-2 and setting the confidence parameter β, using the theorem 2 we can state that the violation probability for the computed solution is guaranteed to be smaller $\varepsilon(s^*)$ than with a confidence equal to $1 - \beta$:

$$\mathbb{P}^N_{d_k}\{\mathbb{P}_{d_k}\{g^*_{k,M} > -\delta\} \le \varepsilon(s^*)\} \ge 1 - \beta\,. \qquad 2\text{-}10$$

This result found is valid for the constraints evaluated for the approximated state $x^*_{k,M}$. However, to reapproach problem 1-7, we need to obtain a similar statement for the stationary state process $x^*_{k,\infty}$ obtained when the controller is γ^*, K^*, that is

$$x^*_{k,\infty} = (I - A - BK^*)^{-1} B\gamma^* + \sum_{s=0}^{\infty} (A + BK^*)^s d_{k-1-s}\,. \qquad 2\text{-}11$$

Precisely, addressing problem 1-7 requires to find a bound for the violation probability of $g^*_{k,\infty} \le 0$, where $g^*_{k,\infty}$ is the max of the cost function and the constraint function evaluated for the stationary state process $x^*_{k,\infty}$:

$$g^*_{k,\infty} = max\{l(x^*_{k,\infty}, \gamma^* + K^* x^*_{k,\infty}) - h^*, f(x^*_{k,\infty}, \gamma^* + K^* x^*_{k,\infty})\}. \tag{2-12}$$

In order to bound $\mathbb{P}_{d_k}\{g^*_{k,\infty} > 0\}$, we would like to use the guarantees obtained for $g^*_{k,M}$ and expressed in 2-10. For that to be possible, it is intuitive that the difference between $x^*_{k,M}$ and $x^*_{k,\infty}$ has to be kept under control. In the next section, we introduce the mathematical steps to obtain the sought result and we highlight the necessity to introduce another constraint to 2-2, as exhaustively discussed in chapter 3. In chapter 4, the final solution to problem 1-7 will be given together with the complete proof.

2.3 Preliminary calculation

As said in the previous section, the result 2-10 is valid for the approximated state, while we would like to have a similar statement for the original problem. With some mathematical steps, we will be able to extend the validity of equation **Errore. L'origine riferimento non è stata trovata.**, provided that a suitable margin is added to $\varepsilon(s^*_N)$. This margin is the price we have to pay for using the truncated scenario approach. The first step is to include $g^*_{k,M}$ in the violation probability expression for $g^*_{k,\infty}$. To do that, we add and subtract $g^*_{k,M}$ and δ in the expression of the violation probability for the constraint evaluated for the stationary state process:

$$\begin{aligned}\mathbb{P}_{d_k}\{g^*_{k,\infty} > 0\} &= \mathbb{P}_{d_k}\{g^*_{k,M} + \delta + g^*_{k,\infty} - g^*_{k,M} - \delta > 0\} \\ &\leq \mathbb{P}_{d_k}\{g^*_{k,M} > \delta \wedge g^*_{k,\infty} - g^*_{k,M} > \delta\} \\ &\leq \mathbb{P}_{d_k}\{g^*_{k,M} > -\delta\} + \mathbb{P}_{d_k}\{g^*_{k,\infty} - g^*_{k,M} > \delta\}.\end{aligned} \tag{2-13}$$

The value of $\mathbb{P}_{d_k}\{g^*_{k,M} > -\delta\}$ is determined, with a certain confidence, by equation 2-10, so we just need to bound the value of $\mathbb{P}_{d_k}\{g^*_{k,\infty} - g^*_{k,M} > \delta\}$. We can rewrite it as:

$$\mathbb{P}_{d_k}\{g^*_{k,\infty} - g^*_{k,M} > \delta\} = \mathbb{P}_{d_k}\{|g^*_{k,\infty} - g^*_{k,M}| > \delta\} \leq \frac{\mathbb{E}[|g^*_{k,\infty} - g^*_{k,M}|]}{\delta}. \tag{2-14}$$

The second inequality is the result of the application of the Chebyshev's inequality (see [23]). In this step emerges the reason behind the tightening δ introduced in problem 2-2. Without it, it would not be possible to bound the probability with the expected value, blocking the following steps.

Thanks to the assumption on the Lipschitz continuity of the functions $l(x,u)$ and $f(x,u)$ we can write the difference between $g^*_{k,\infty}$ and $g^*_{k,M}$ as:

$$\begin{aligned}\mathbb{E}\left[\left|g^*_{k,\infty}-g^*_{k,M}\right|\right] &= \mathbb{E}\left[\left|g^*_{k,\infty}-g\left(x^*_{k,M},\gamma^*+K^*x^*_{k,\infty}\right)+\right.\right.\\ &\quad \left.\left.g\left(x^*_{k,M},\gamma^*+K^*x^*_{k,\infty}\right)-g^*_{k,M}\right|\right]\\ &\le \mathbb{E}\left[\left|g\left(x^*_{k,\infty},\gamma^*+K^*x^*_{k,\infty}\right)-g\left(x^*_{k,M},\gamma^*+K^*x^*_{k,\infty}\right)\right|\right]+\\ &\quad \mathbb{E}\left[\left|g\left(x^*_{k,M},\gamma^*+K^*x^*_{k,\infty}\right)-g\left(x^*_{k,M},\gamma^*+K^*x^*_{k,M}\right)\right|\right]\\ &\le \mathbb{E}\left[L\left\|x^*_{k,\infty}-x^*_{k,M}\right\|\right]+\mathbb{E}\left[L\left\|K^M\left(x^*_{k,\infty}-x^*_{k,M}\right)\right\|\right]\\ &\le L(1+\|K^*\|)\mathbb{E}\left[\left\|x^*_{k,\infty}-x^*_{k,M}\right\|\right].\end{aligned} \quad \text{2-15}$$

L is the maximum between the Lipschitz constant of the cost function and the one of the constraints. The term $\|K^*\|$ is present only when the functions depend not only on the state, but also on the control input. In this case, when computing the solution, a bound on the norm of $\|K^*\|$ has to be added as a constraint in 2-2, so that $L(1+\|K^*\|)$ can be bounded by another constant L'. If, instead, none of the functions depends on the control input, the bound is $\left|g^*_{k,\infty}-g^*_{k,M}\right|\le L\left\|x_{\mathrm{k},\infty}-x_{\mathrm{k},M}\right\|$. From now on we take this second standpoint. However, there are no conceptual differences and it is clear that all results applies when L is replaced by L' as defined above.

Substituting $x^*_{k,M}$ and $x^*_{k,\infty}$ with equations 2-9 and 2-11 in 2-15 the expected value of the difference is finally bounded as follows:

$$\begin{aligned}\mathbb{E}\left[\left|g^*_{k,\infty}-g^*_{k,M}\right|\right] &\le \mathbb{E}\left[L\left\|x^*_{k,\infty}-x^*_{k,M}\right\|\right]=L\mathbb{E}\left[\left\|x^*_{k,\infty}-x^*_{k,M}\right\|\right]\\ \le L\mathbb{E}\left[\left\|\sum_{s=M}^{\infty}(A+BK^*)^s d_{k-1-s}\right\|\right] &\le L\mathbb{E}\left[\sum_{s=M}^{\infty}\|(A+BK^*)^s\|\cdot\|d_{k-1-s}\|\right]\\ &\le L\sum_{s=M}^{\infty}\|(A+BK^*)^s\|\,\mathbb{E}[\|d_{k-1-s}\|].\end{aligned} \quad \text{2-16}$$

The expected values of the disturbance, thanks to the assumption made, is equal to σ^2:

$$\mathbb{E}[\| d_{k-1-s} \|_2]=\sigma^2. \quad \text{2-17}$$

This term is not chosen by the user but depends on the problem at hand.

The only term in inequality 2-16 on which the user has influence is the control gain K^*. Given that the summation contains an infinite number of terms, it is necessary to impose some restriction on the values K can assume. In particular, it is necessary to ensure that the norm of the matrix $(A+BK)^s$ has to become smaller than one at a certain point, for the summation to be finite. Also, the summation value influences the quality of the result, so it will be necessary to find a balance between it and the restrictiveness of the condition imposed to K^*.

There are few different bounds that can be applied on the norm of the matrix $(A + BK)^s$. These will be discussed in the next chapter 3. Later chapter 4 will re the derivation here initiated and the complete result on the bound $\mathbb{P}_{d_k}\{g^*_{k,\infty} > 0\}$ will be provide.

3. Constraints on the norm of $(A + BK)^s$

The need to limit the value of $\mathbb{E}[|g_{k,\infty}^M - g_{k,M}^M|]$ demands us to bound the value of the term $\sum_{s=M}^{\infty}\|(A + BK)^s\|$, which can be accomplished by imposing suitable conditions on K. As a matter of fact, the other two terms determining the value of $\mathbb{E}[|g_{k,\infty}^M - g_{k,M}^M|]$ (which are the Lipschitz value of the constraints function and the expected value of the disturbance) are both given and not modifiable by the user; on the contrary, the matrix $A + BK$ depends on the optimization variables, K, and though an uniform bound to $\sum_{s=M}^{\infty}\|(A + BK)^s\|$ that applies for every K does not exist, this can be achieved by imposing suitable conditions on that limit the values taken by $\|(A + BK)^s\|$. These conditions can be enforced as additional constraints in program x when selecting K. Given that the value of $\mathbb{E}[|g_{k,\infty}^M - g_{k,M}^M|]$ concurs in determining the bound for the violation probability, we are interested on the other hand, in obtaining the smallest possible value for the $\sum_{s=M}^{\infty}\|(A + BK)^s\|$, leading to the strictest constraints. However, feasibility and cost value must be considered too: the size of the feasible set diminishes when the bound strictness increases, potentially causing a large rise in the cost, and if the norm $\|(A + BK)^s\|$ is required to be too small the feasible decision set can even become void, meaning that there is no feasible solution. In this chapter we will analyze three possible constraints, highlighting their differences and the advantages of each one, to find the one with the best compromise between the dimension of the decision set and the bound on the value of $\sum_{s=M}^{\infty}\|(A + BK)^s\|$. We will start from the simplest one and then increasing in complexity.

3.1 Constraint on the spectral radius

One of the simplest constraints that can be applied is requiring that A+BK is contractive according to some sub-multiplicative norm, i.e., that the norm of A+BK is smaller than 1.

$$\|A+BK\| \leq \lambda < 1 \tag{3-1}$$

Thanks to the sub-multiplicative property of the norm, condition implies 3-1 that the norm of the power of the matrix must be bounded by a stricter condition as the exponent increases. Indeed, it holds that

$$\|AB\| \leq \|A\| \cdot \|B\|\,, \tag{3-2}$$

and in the case of matrix power, we have that:

$$\|A^{a+b}\| = \|A^a A^b\| \leq \|A^a\| \cdot \|A^b\|\,. \tag{3-3}$$

Applying this property repeatedly, we can therefore write

$$\|(A+BK)^s\| \leq \|A+BK\| \cdot \|A+BK\| \cdot \ldots \cdot \|A+BK\| (s\ times) \leq \lambda^s\,, \tag{3-4}$$

that is, for any value of s, the norm of the power matrix $\|(A+BK)^s\|$ must stay below an exponential with base λ.

Assuming that the condition 3-1 is satisfied, the bound on the summation can easily be found as follow:

$$\begin{aligned}\sum_{s=M}^{\infty} \|(A+BK)^s\| &\leq \sum_{s=M}^{\infty} \lambda^s = \sum_{s=0}^{\infty} \lambda^s - \sum_{s=0}^{M-1} \lambda^s \\ &= \frac{1}{1-\lambda} - \frac{1-\lambda^M}{1-\lambda} = \frac{\lambda^M}{1-\lambda}\,.\end{aligned} \tag{3-5}$$

We can now substitute it in the equation 2-16, obtaining the value of χ_M:

$$\begin{aligned}\mathbb{E}\left[\left|g_{k,\infty}^M - g_{k,M}^M\right|\right] &\leq L_1 \sum_{s=M}^{\infty} \|(A+BK)^s\|\, \mathbb{E}[\|d_{k-1-s}\|] \\ &\leq L_1 \frac{\lambda^M}{1-\lambda} \sigma^2 = \chi_M\,.\end{aligned} \tag{3-6}$$

The value of χ_M depends on both M and λ. When deciding the values of these parameters, it is important to remember the various need to one has to trade off with. Increasing M leads to a decrease of χ_M, obtaining better guarantees for our solution, but it also means requiring more data and increasing the computational load. For λ the trade-off is between the value of χ_M and the strictness of the constraint. The choice has to be done based on the problem requirements and the user goals.

The main advantage of the solution proposed in the present section is its simplicity. It is possible to include constraint 3-1 in most optimization problems without noticeable increase of the computational load or algorithm complexity. On the other hand, it is unnecessary restrictive, excluding solutions that could potentially be the best choice or, in the worst case, it can make the feasibility domain void. It is possible, as we will see in a further chapter, to bound the norm with an exponential after a transient, just for exponents s equal or greater than M, obtaining the same value for the summation, but with a less rigid constraint. The advantages brought by the solution in the present section thus do not compensate the excessive reduction of the feasibility domain.

3.2 Bounding the tail with an exponential

The condition $\|A + BK\| \leq \lambda < 1$, introduced in the previous section, effectively limits the summation $\sum_{s=M}^{\infty} \|(A + BK)^s\|$, but it is far more restrictive than necessary. In fact, it also constrains the terms $(A + BK)^s$ for $s < M$, which we are not interested about because they do not appear in the sum. Admissible solutions, such as the ones having an initial overshoot, are excluded from the decision set, despite being admissible and guaranteeing the same bound as the summation. Our goal is then to impose to the norm $\|(A + BK)^s\|$ to stay under a given exponential just for certain s onwards. As is clear expressing this condition as an infinite sequence of constraints would be not admissible. Remarkably, it is enough to impose the following finite number of constraints:

$$\|(A + BK)^t\| \leq \lambda^t \; for \; t = T, T+1, \ldots, 2T-1, \qquad \text{3-7}$$

Where $\|\cdot\|$ is a generic sub-multiplicative norm.

Even if the condition 3-7 is imposed till 2T-1, it suffices to guarantee that the norm of $(A + BK)^t$ is dominated by the exponential λ^t for every t greater than T thanks to the sub-multiplicativity property of the norm, we can rewrite $\|(A+BK)^X\|$ as the product of components that can be bounded according to condition 3-7. If X is a multiple of T, it is sufficient to divide it in X/T equal terms:

$$\begin{aligned}\|(A+BK)^X\| &\le \|(A+BK)^T\| \cdot \|(A+BK)^T\| \cdot \ldots \cdot \|(A+BK)^T\| \\ &= \left(\|(A+BK)^T\|\right)^{X/T} \le \left(\lambda^T\right)^{X/T} = \lambda^X.\end{aligned} \tag{3-8}$$

If, instead, X is not a multiple of T, it possible to use a similar division, writing X as $q \cdot T + r$, where $q = \lfloor X/T \rfloor - 1$ and, consequently, $T < r \le 2T - 1$. The bound of the norm is then:

$$\|(A+BK)^X\| \le \left(\|(A+BK)^T\|\right)^q + \|(A+BK)^r\| \le \left(\lambda^T\right)^q + \lambda^r = \lambda^{q \cdot T + r} = \lambda^X . \tag{3-9}$$

Therefore, for every value X greater than T, $\|(A+BK)^X\| \le \lambda^X$, i.e., the norm lays under the exponential with base λ.

Under the condition 3-7 and thanks to the considerations done above, the value of the summation $\sum_{s=M}^{\infty} \|(A+BK)^s\|$ can be bounded as follows:

$$\sum_{s=M}^{\infty} \|(A+BK)^s\| \le \sum_{s=M}^{\infty} \lambda^s = \sum_{s=0}^{\infty} \lambda^s - \sum_{s=0}^{M-1} \lambda^s = \frac{1}{1-\lambda} - \frac{1-\lambda^M}{1-\lambda} = \frac{\lambda^M}{1-\lambda}. \tag{3-10}$$

As anticipated, the result is the same we obtained in the previous section, so the same expression of χ_M is obtained and the same conclusions about the choice of the parameters M and λ can be drawn.

The fact that the bound for $\sum_{s=M}^{\infty} \|(A+BK)^s\|$ found depends only on λ and M is useful for a further consideration. The parameter T, which determines the starting point of the bound and so its strictness, does not appear in the final result, so it can be modified without repercussion on the quality of the guarantee. We would like to have as much freedom as possible in the design of the controller, so a bigger T is preferred. There is no gain in choosing a small T, so we are pushed to select the biggest T possible. The only limit is the value of M. Choosing a T greater than M would make impossible to determine an a priori value that bounds the sum $\sum_{s=M}^{\infty} \|(A+BK)^s\|$ and, consequentially, one for χ_M. Considering these two aspects, the best choice is clearly picking $T = M$, which maximize the value of T while respecting the condition $T \le M$.

The only exception to this rule is because of computational limitations since the bigger T the higher the number of constraints to be imposed according to 3-7 and the harder the resulting non-convex problem.

3.3 Bounding the tail with a recurrent exponential

Starting from the previous constraint, we can try to develop a less strict condition. An idea is to use again an exponential, but shifting it towards left. In this way, the first term bounded is limited by 1 instead of λ^M, the second by λ instead of λ^{M+1} and so on. In mathematical terms, the condition becomes:

$$\|(A+BK)^t\| \leq \lambda^{t-T} \; for \; t = T, T+1, \ldots, 2T. \tag{3-11}$$

It is possible to use slightly different variations of this condition, for example imposing it from T to 2T-1 instead of 2T, or starting the bound from λ and not from 1, but the procedure to reach the result becomes more intricate, so it is preferred to present this version as a first example, which is the simplest one in term of notations.

The first step consists in splitting the summation in two parts (given that M and T are two user-chosen parameters, it will be always true that M is greater than T), in order to obtain a summation that start from T:

$$\sum_{s=M}^{\infty} \|(A+BK)^s\| = \sum_{s=T}^{\infty} \|(A+BK)^s\| - \sum_{s=T}^{M-1} \|(A+BK)^s\| \,. \tag{3-12}$$

To better understand the calculus, we rewrite the first sum showing explicitly its terms:

$$\sum_{s=T}^{\infty} \|(A+BK)^s\| = \|(A+BK)^T\| + \|(A+BK)^{T+1}\| + \ldots + \|(A+BK)^{2T}\| + \ldots \tag{3-13}$$

We collect the components in different summations, each one with 2T-1 elements (the first one is actually composed of 2T elements, but in a further step we will leave out the first term to make it equally long):

$$\sum_{s=T}^{\infty} \|(A+BK)^s\| = \sum_{s=T}^{2T} \|(A+BK)^s\| + \sum_{s=2T+1}^{3T} \|(A+BK)^s\| \\ + \sum_{s=3T+1}^{4T} \|(A+BK)^s\| + \sum_{s=4T+1}^{5T} \|(A+BK)^s\| + \cdots \tag{3-14}$$

Then we use the sub-multiplicative property of the norm to split the terms with too high power in more components. Due to this step, we have the following inequality:

$$\sum_{s=T}^{\infty}\|(A+BK)^s\| \leq \sum_{s=T}^{2T}\|(A+BK)^s\| + \sum_{s=2T+1}^{3T}\|(A+BK)^T\|\|(A+BK)^{s-T}\| + \sum_{s=3T+1}^{4T}\|(A+BK)^{2T}\|\|(A+BK)^{s-2T}\| + \sum_{s=4T+1}^{5T}\|(A+BK)^{3T}\|\|(A+BK)^{s-3T}\| + \cdots \quad \text{3-15}$$

We can now bring outside the summations the terms that do not depend on s and redefine the running index in all the sums so as to obtain

$$\sum_{s=T}^{\infty}\|(A+BK)^s\| \leq 1 + \sum_{s=T+1}^{2T}\|(A+BK)^s\| + \|(A+BK)^T\|\sum_{s=T+1}^{2T}\|(A+BK)^s\| + \|(A+BK)^{2T}\|\sum_{s=T+1}^{2T}\|(A+BK)^s\| + \|(A+BK)^{3T}\|\sum_{s=T+1}^{2T}\|(A+BK)^s\| + \cdots \quad \text{3-16}$$

Using the inequality 3-11, we can now bound the norms of the matrices in the sum with the corresponding exponential, and noticing that

$$\|(A+BK)^{mT}\| = \left\|(A+BK)^{\left\lfloor\frac{m}{2}\right\rfloor 2T}\cdot(A+BK)^{\left(m-2\left\lfloor\frac{m}{2}\right\rfloor\right)T}\right\| \leq \|(A+BK)^{mT}\|^{\left\lfloor\frac{m}{2}\right\rfloor}\cdot\|(A+BK)^T\|^{\left(m-2\left\lfloor\frac{m}{2}\right\rfloor\right)} \leq \lambda^{T\left\lfloor\frac{m}{2}\right\rfloor} \quad \text{3-17}$$

Yields

$$\sum_{s=T}^{\infty}\|(A+BK)^s\| \leq 1 + \sum_{s=T+1}^{2T}\lambda^{s-T} + \quad \text{3-18}$$

$$1 * \sum_{s=T+1}^{2T}\lambda^{s-T} + \lambda^T\sum_{s=T+1}^{2T}\lambda^{s-T} + 1 * \lambda^T\sum_{s=T+1}^{2T}\lambda^{s-T} + \ldots$$

Notice that the above derivations are equivalent to use 3-11 to bound each generic term $\|(A+BK)^t\|$ for $t \geq T$ as follows

$$\|(A+BK)^t\| \leq \lambda^{\left\lfloor\frac{m}{2}\right\rfloor T+r}, \qquad \text{3-19}$$

where m and r, $r = 0,1,\ldots,T-1$, are such that $t = mT + r$. A pictorial representation of the bound 3-19 is given in figure 3.1.

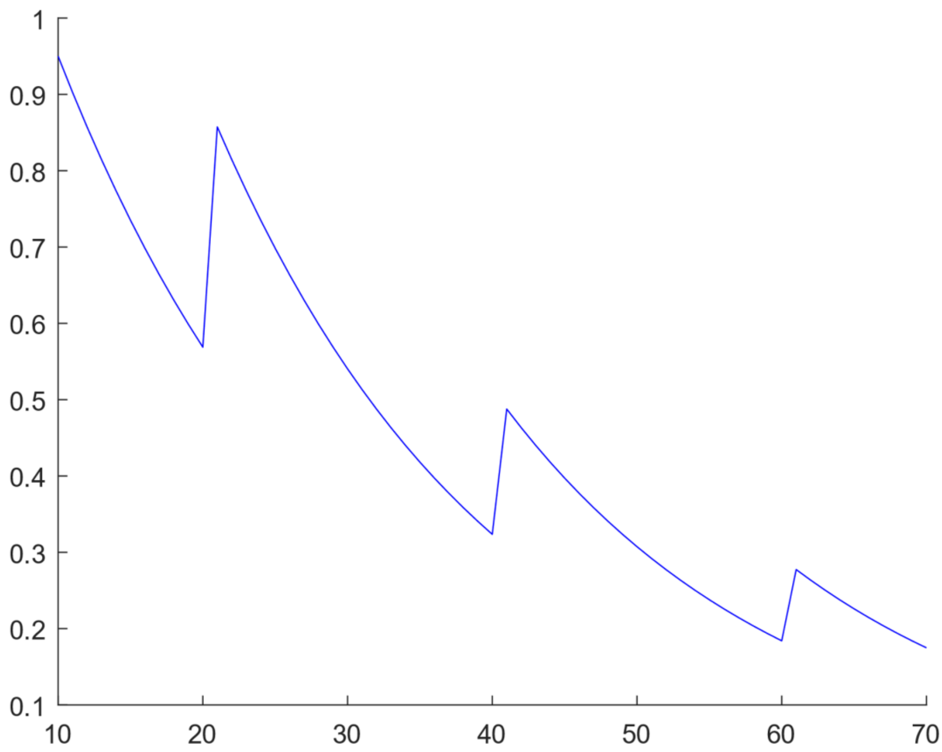


Figure 3-1 Example of the constraint described by equation 3-19. In this case λ is equal to 0.95 and T is equal to 10

To complete the bound to $\sum_{s=T}^{\infty}\|(A+BK)^s\|$, we redefine the summing indexes

$$\sum_{s=T}^{\infty}\|(A+BK)^s\| \leq 1+\sum_{s=1}^{T}\lambda^s+\sum_{s=1}^{T}\lambda^s \qquad \text{3-20}$$

$$+\lambda^T\sum_{s=1}^{T}\lambda^s+\lambda^T\sum_{s=1}^{T}\lambda^s+\lambda^{2T}\sum_{s=1}^{T}\lambda^s+\lambda^{2T}\sum_{s=1}^{T}\lambda^s+\ldots$$

which highlights a clear pattern with which the summations repeat. We can then collect the common terms so obtaining

$$\sum_{s=T}^{\infty} \|(A+BK)^s\| \leq 1 + 2\left(\sum_{s=1}^{T} \lambda^s\right)\left(\sum_{i=0}^{\infty} (\lambda^T)^i\right) \tag{3-21}$$

which, using the properties of geometric series, is equal to

$$\sum_{s=T}^{\infty} \|(A+BK)^s\| \leq 1 + 2\left(\frac{1-\lambda^{T+1}}{1-\lambda} - 1\right)\left(\frac{1}{1-\lambda^T}\right) \tag{3-22}$$

$$= 1 + \frac{2(1-\lambda^{T+1} - 1 + \lambda)}{(1-\lambda)(1-\lambda^T)} = 1 + 2\lambda\frac{1-\lambda^T}{(1-\lambda)(1-\lambda^T)} = \frac{1-\lambda+2\lambda}{1-\lambda} = \frac{1+\lambda}{1-\lambda}.$$

Replacing it in equation 3-12, we obtain:

$$\sum_{s=M}^{\infty} \|(A+BK)^s\| \leq \frac{1+\lambda}{1-\lambda} - \sum_{s=T}^{M-1} \|(A+BK)^s\| \leq \frac{1+\lambda}{1-\lambda}. \tag{3-23}$$

This inequality is valid for every value of M, but it is not the best that can be obtained. In fact, $\sum_{s=T}^{M-1} \|(A+BK)^s\|$ depends on the difference between M and T and it differs from case to case, so there is no generic value to substitute it. The best way to find the limit for the summation $\sum_{s=M}^{\infty} \|(A+BK)^s\|$ is to start from M, and not from T as done. Given n such that $(n-1)T < M \leq nT$ we can split the sum as follows

$$\sum_{s=M}^{\infty} \|(A+BK)^s\| = \sum_{s=M}^{nT} \|(A+BK)^s\| + \sum_{s=nT+1}^{(n+1)T} \|(A+BK)^s\| + \ldots, \tag{3-24}$$

and then proceed in an analogous way as seen before. While the notation becomes complex when trying to do it with general values, once M and T are known the steps are not more complex than what done before. This subdivision leads to the lowest bound for the sum $\sum_{s=M}^{\infty} \|(A+BK)^s\|$, and therefore to the best guarantee.

There are many conditions similar to 3-11 that can be considered as alternatives. Among these, there is one which is still quite simple and, at the cost of a little increase in its restrictiveness, it guarantees a better evolution of $\sum_{s=M}^{\infty} \|(A+BK)^s\|$. The condition is the following:

$$\|(A+BK)^t\| \leq \lambda^{t-T+1}\, for\ t = T, T+1, \dots, 2T\ . \qquad \text{3-25}$$

The difference with the condition 3-11 is the presence of a +1 in the exponent of λ. It is not too much more limiting with respect to 3-11 (in fact, it is possible to make it almost identical to the previous one just incrementing the value of T by 1). On the other side, it guarantees a better bound for the summation:

$$\sum_{s=M}^{\infty} \|(A+BK)^s\| \leq \lambda \frac{1+\lambda^2 - 2\lambda^{T+1}}{(1-\lambda)(1-\lambda^{T+1})} - \sum_{s=T}^{M-1} \|(A+BK)^s\|\ . \qquad \text{3-26}$$

We will skip the derivation, because it is analogous to the one shown above for 3-11. With this condition, the bound diminishes much faster as λ decreases. The improvements in the results can be worth the increased rigidness of the requirements. For this reason, we will use this condition in the practical case.

3.4 Comparison between the exponential and the recurrent exponential

It is useful to understand the relation between conditions 3-7 and 3-11. This comparison is provided in the present section.

At the step T, the constraint with the exponential will be equal to λ^T, while the recurrent one will be equal to λ. It is then clear that we can not compare them using the same λ, because the exponential is much stricter, but it gives a better bound for the summation $\sum_{s=M}^{\infty} \|(A+BK)^s\|$, guaranteeing less conservative evolution of the sum. It can be useful then to plot the values for the summation bounds as function of λ for both the conditions, making easier to find the values of λ leading to the same χ_M for the two conditions.

As it is possible to see in figure 3-2 and in figure 3-3, the difference in value for the same λ is significative, both for high values (over 0,9) and for low ones. The recurrent exponential goes to zero much slower, because its first term is always raised with a lower exponential, due to subtraction of T in the exponent. As T decreases, the shape of the function becomes more similar to the one of the exponential, which on the other side remains unvaried, because it depends only on M (see figure 4-4).

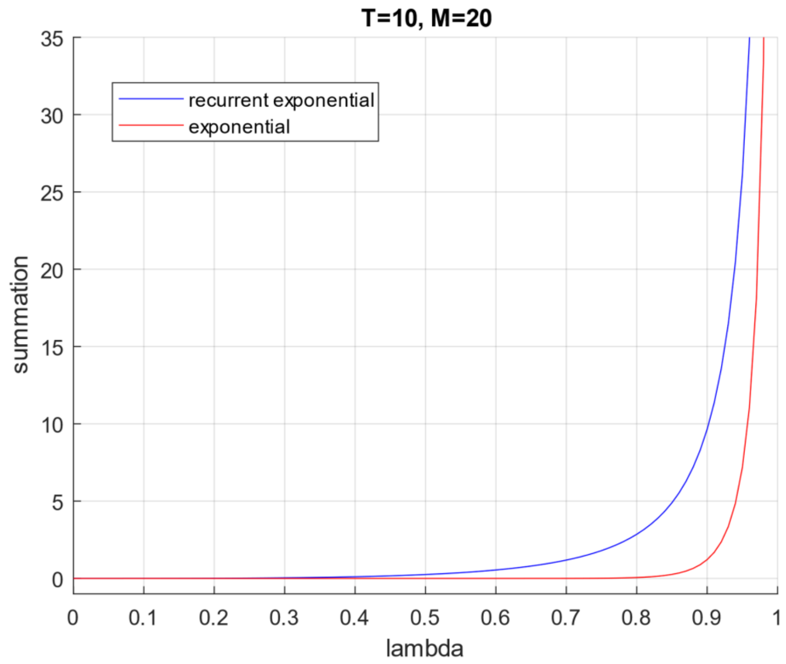


Figure 3-2 In red the value of the summation for the condition 3-7, in blue the one for the condition 3-25. The parameters values are T=10 and M=20

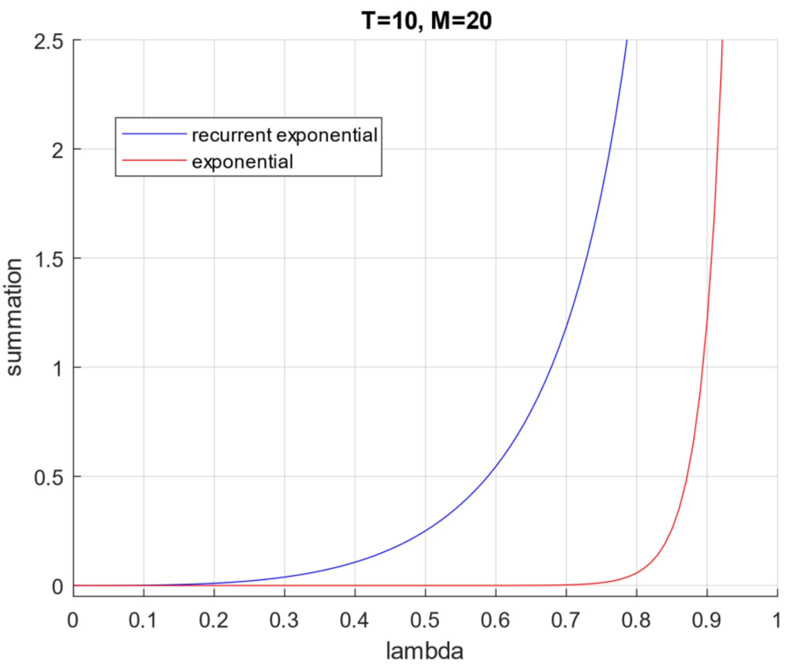


Figure 3-3 The value of the summation tends to zero much faster in the case of the condition 3-7 with respect to the one of the condition 3-25. The parameters values are T=10 and M=20.

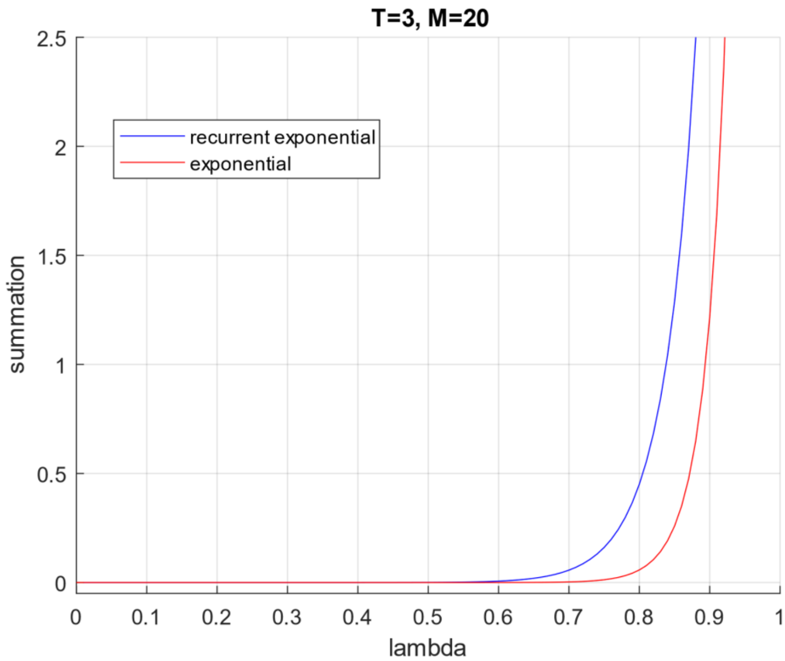


Figure 3-4 In this case T=3 and M=20. It is possible to appreciate how much is changed the summation of the recurrent exponential in comparison to the previous image where T=10. Now, the value of 1 is reached around 0.84 while before was obtained for 0.68 ca

For T=1, the two curves will be identical, but as T increases, the bound to $\sum_{s=M}^{\infty}\|(A+BK)^s\|$ using the recurrent exponential grows much faster, moving the plot towards left in the graph. The difference between the two λs, given the same χ_M for the two approaches, is small for high values of χ_M. When we take the λ for the classic exponential equal to 0.99, the other λ, for the recurrent exponential to give the same χ_M, is 0.98, with a difference of just 0.01. The difference increases for small χ_M, so when the first λ is equal to 0.8, the other must be around 0.45, with a difference of circa 0.35, 35 times higher than the previous case.

Figure 3-5 depicts the relation between the λs for T=10. While the x axis depicts the λ used in the classic exponential bound, the y axis gives the λ used in the bound based on the recurrent exponential.

As we have said before, the recurrent exponential goes to zero much slower. For this reason, very low values for its λ are needed to obtain the value of the summation obtained via the classic exponential when $\lambda = 0.5$.

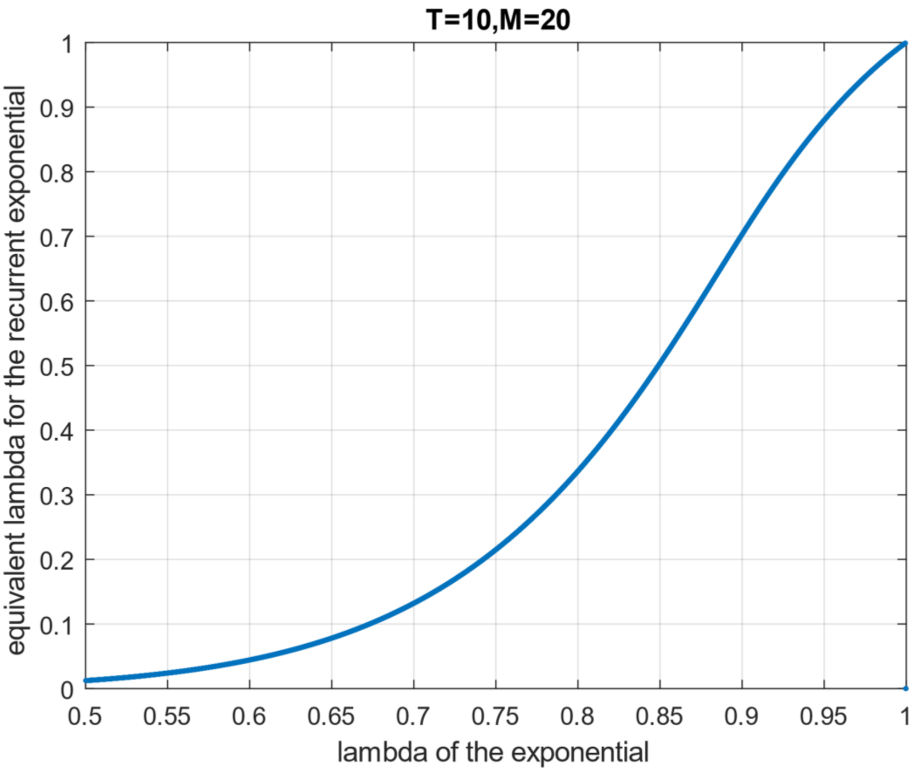


Figure 3-5 Correspondence between the two lambdas for T=10 and M=20

The relation between the two λs depends both on the value of T and the value of M. In figure 3-6 the relation is shown for eight different values of M (maintaining T=10).

Increasing the value of M while keeping constant T corresponds in subtracting more terms in 3-24 and 3-30 equations. But the term subtracted for the 3-30 is greater. For example, when passing from M=20 to M=21, we will subtract $\lambda^M = \lambda^{20}$ to the exponential summation and $\lambda^{M-T+1} = \lambda^9$ to the recurrent exponential one. In this way, we are subtracting a greater value to the condition 3-7, decreasing its summation curve and making it more similar to the other.

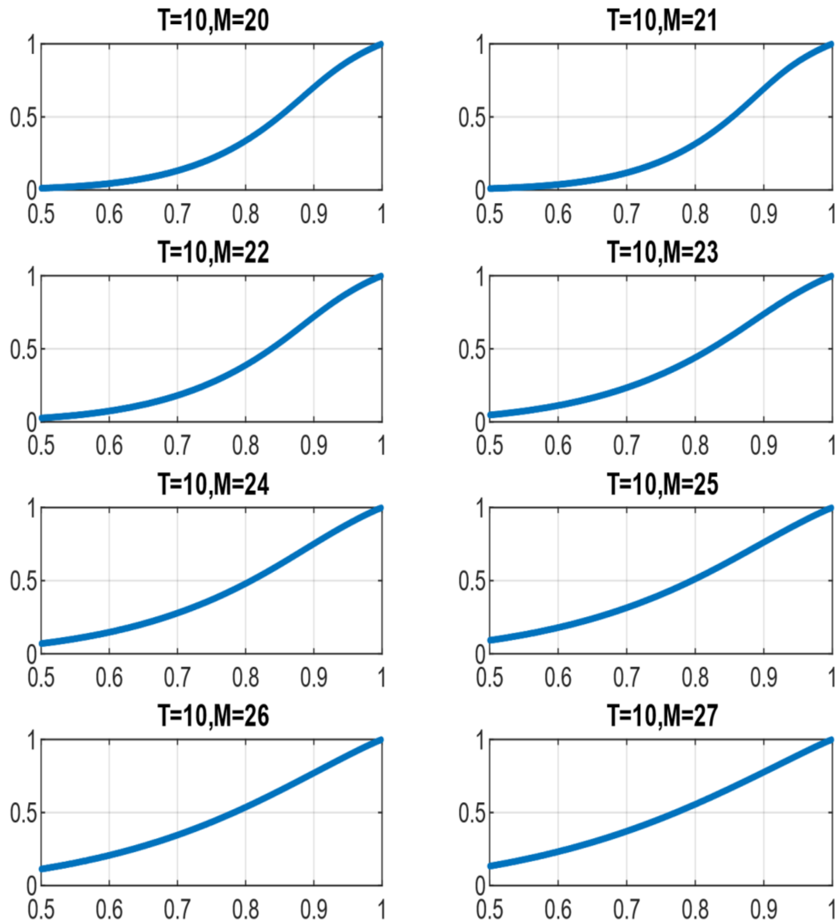


Figure 3-6 Plot of the relationship between the λ of the classic exponential (on the x axis) and the one of the recurrent exponential (on the y axis) for T=10 and for M going from 20 to 27.

Once we have two equivalent λs, leading to the same value of χ_M, it is possible to compare the constraints imposed by 3-7 and 3-11 by plotting the two functions $\|(A + BK)^t\|$ is required to stay below. While comparing them, it is important to remember that the condition 3-7, the one with the recurrent exponential, starts at T, while the condition 3-25 starts at M.

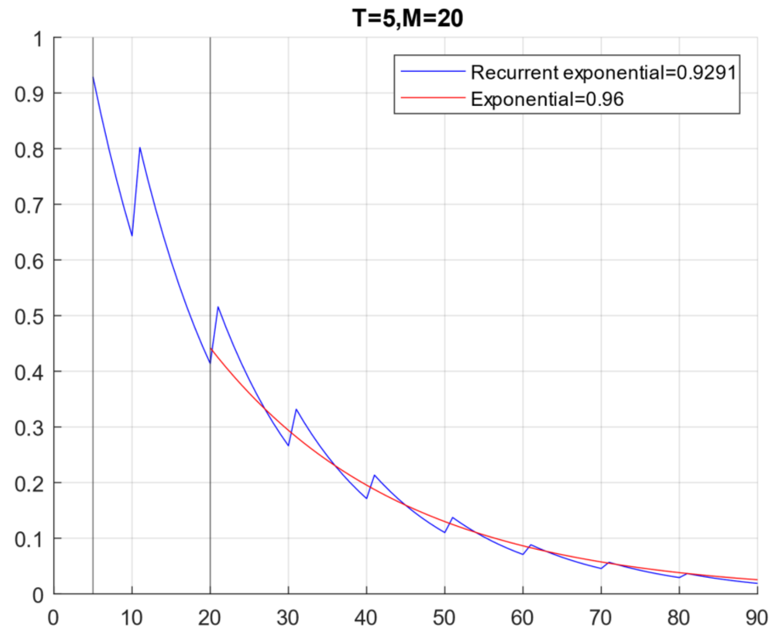


Figure 3-7 Comparison between the two bounds for T=5 and M=20

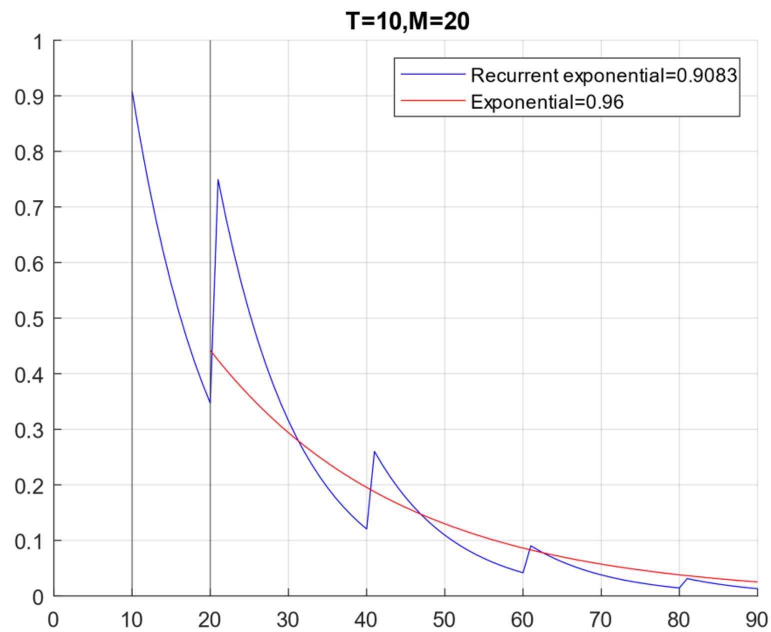


Figure 3-8 Comparison between the two bounds for T=10 and M=20

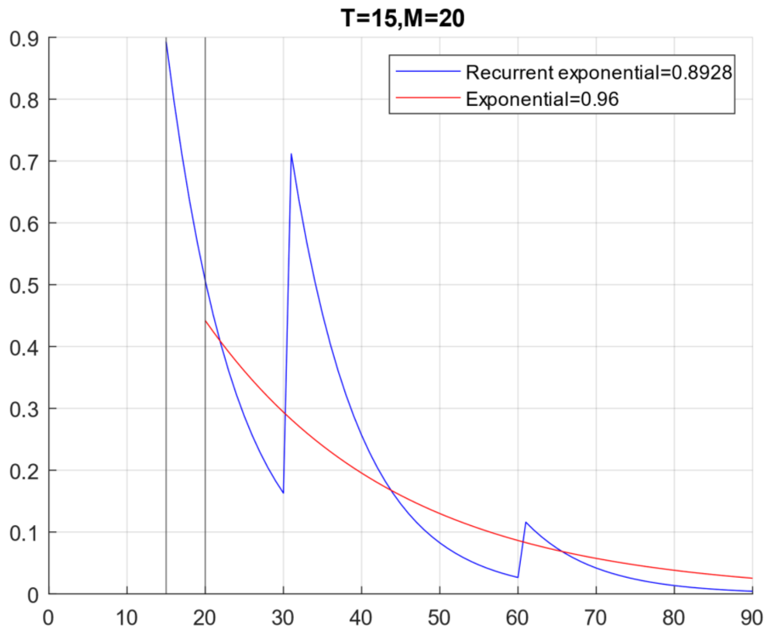


Figure 3-9 Comparison between the two bounds for T=15 and M=20

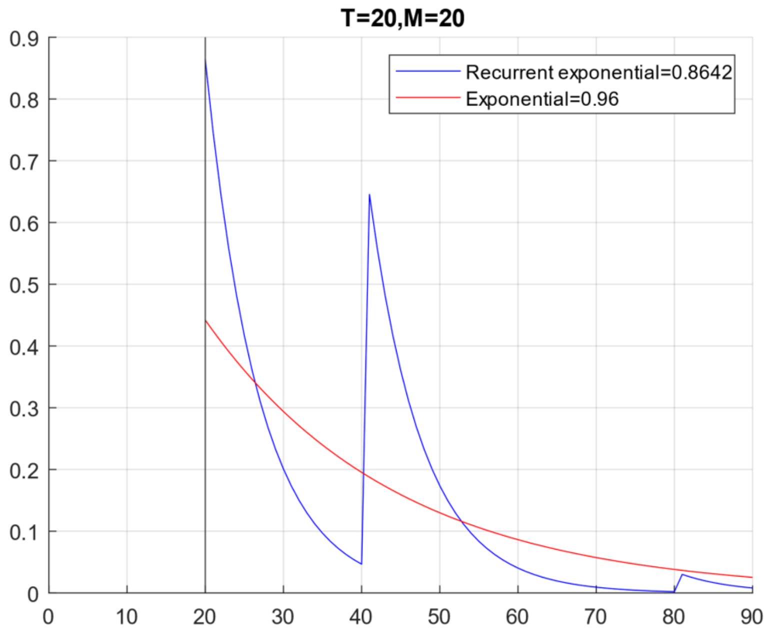


Figure 3-10 Comparison between the two bounds for T=20 and M=20

In the first three cases (T=5, T=10 and T=15), the best condition appears to be 3-7, since the disadvantage given by starting earlier, for the recurrent exponential, is not compensated by a less strict bound to be satisfied by $\|(A + BK)^t\|$. The two functions have indeed similar values for $t \gg M$, so the difference in the starting point is decisive in determining the best one. Condition 3-11 could be better for very specific cases, but in general the condition is definitively worse.

In the fourth case (T=20), it is not anymore obvious which one is the best condition. While the recurrent exponential has a steeper slope, it also starts from a much higher value. The two conditions lead to two different decision sets, but it is not possible to state which one contains the best solution. The choice between the two is left to the user. Even if there is the possibility to choose a value for T smaller than M, it seems that, as for the other exponential, the best decision is to pick it equal to M. The increase in the value of λ allowed by a smaller T is not sufficient to justify the earlier start of the constraint.

4. Complete solution to the given problem

As explained in previous chapters, the methodology proposed to solve the initial problem, the application of the scenario approach followed by an approximation of the stationary process, required to insert a further constraint on the norm of the matrix $(A + BK)^s$. This is not an assumption on the given system, therefore it does not reduce the applicability of the method, rather it is a further requirement on the controller parameters. So, even though this constraint needs to be added to the formulation of the initial problem, it has to be seen as a step of the solution process rather than a part of the problem itself. With this addition, the problem is finally complete, and we have all the ingredients to prove that the violation probability of the solution is bounded by un upper limit with a confidence $1 - \beta$.

In this chapter, the main problem resolution steps are recalled, and the complete problem is introduced for two bounds on the norm: the exponential and the recurrent exponential. The main result, regarding the violation probability of the solution, is stated and proved. In the second part of the chapter there are some considerations and recommendations on the choice of the tunable parameters.

4.1 Complete problem with the exponential bound on the norm

The constraint introduced to bound the value of $\sum_{s=M}^{\infty}\|(A + BK)^s\|$, in order to limit the expected difference $\mathbb{E}[|g_{k,\infty}^M - g_{k,M}^M|]$, need to be added to problem 2-2, together with the other constraints on the state and the control input. If we apply the exponential bound expressed in equation 3-7, the complete version of the problem becomes

$$\min_{\gamma,K,h} h \tag{4-1}$$
$$s.t.\, \mathbb{P}_{d_k}\{l(x_{k,\infty},\gamma + Kx_{k,M}) \leq h \wedge f(x_{k,\infty},\gamma + Kx_{k,\infty}) \leq 0\} \geq 1-\varepsilon$$
$$\| (A+BK)^t \| \leq \lambda^t \; for\; t = T, T+1, \ldots, 2T-1 \,.$$

In words, we are looking for a controller that minimizes a cost function and respects the state-input constraints for the majority of the cases, and, at the same time, guarantees that the norm of the matrix $(A + BK)^t$ lays under an exponential from a certain power over. The problem, as presented here, is hard to solve, mainly due to the presence of probabilistic constraints and the fact that the stationary state $x_{k,\infty}$ depends on an infinite long disturbance realization, which is not feasible in practice.

A solution to address the first problem is the use of the scenario approach, which allows to move from a chance-constrained problem to a problem with a finite number of standard constraints. The scenario approach requires N realizations of the disturbance to work with. In the case of convex problem, it is possible to determine the value of N so as to guarantee the satisfaction of the chance constraint in 4-1 with confidence $1-\beta$ when N is a function of both the confidence required and the value of ε. When dealing instead with nonconvex problems this is not possible anymore, and the wait&judge approach is used, i.e., first the solution is computed and then a bound to the violation probability is calculated. The value of ε is not a strict requirement anymore because it can not be imposed. It is instead determined after that the solution has been calculated.

To avoid the use of infinite long realization it is necessary to introduce an approximation. Instead of using the stationary state $x_{k,\infty}$, the truncated version $x_{k,M}$, obtained from a realization containing M samples, is adopted. To compensate for the approximation, the constraints on the state and the cost function must be tightened by a term δ. This approximation made necessary the introduction of the bound on the norm, which guarantees that the difference $\mathbb{E}[|g_{k,\infty}^M - g_{k,M}^M|]$ remains under a maximum value determined by the user.

Applying the scenario approach, the truncation of $x_{k,\infty}$ and the tightening δ the problem to solve becomes the following approximated scenario program:

$$\min_{\gamma,\mathrm{K},\mathrm{h}} \mathrm{h} \tag{4-2}$$
$$s.t.\, l\left(x_{k,M}^{(i)},\gamma + Kx_{k,M}^{(i)}\right) \leq h-\delta$$

$$f\left(x_{k,M}^{(i)},\gamma+Kx_{k,M}^{(i)}\right)\leq-\delta$$

$$x_{k,M}^{(i)}=(I-A-BK)^{-1}B\gamma+\sum_{s=0}^{M-1}(A+BK)^{s}d_{k-1-s}^{(i)}$$

$$i=1,\dots,N$$

$$\| (A+BK)^{t} \| \leq \lambda^{t}\ for\ t=T,T+1,\dots,2T-1\ ,$$

where the solution is denoted by (γ^{*},K^{*},h^{*}) and s^{*} is the cardinality of the smallest support set identified. N is the number of scenarios used while M is the truncation value. Before stating the theorem on the results of problem 4-2, we retrieve some notations introduced before.

$g_{k,\infty}^{*}$ depends on the state and the control input in stationary conditions and, on the base of the definition of the function $g(x,u)$, it is equal to:

$$g_{k,\infty}^{*}\left(x_{k,\infty}^{*},\gamma^{*}+K^{*}x_{k,\infty}^{*}\right)=max\left(l\left(x_{k,\infty}^{*},\gamma^{*}+K^{*}x_{k,\infty}^{*}\right)-h,f\left(x_{k,\infty}^{*},\gamma^{*}+K^{*}x_{k,\infty}^{*}\right)\right) \qquad 4\text{-}3$$

$\varepsilon(\cdot)$ is the function solution of 1-10, L is the Lipschitz constant and σ^{2} is the expected value of the disturbance.

Theorem 2: given the solution (γ^{*},K^{*},h^{*}) to problem 4-2, fixed a confidence parameter $\beta\in(0,1)$, it holds that

$$\mathbb{P}_{d_k}^{N}\left\{\mathbb{P}_{d_k}\{g_{k,\infty}^{*}>0\}\leq\varepsilon(s^{*})+\frac{\chi_M}{\delta}\right\}\geq1-\beta\ ,$$

where χ_M is equal to

$$\chi_M=L\frac{\lambda^{M}}{1-\lambda}\sigma^{2}\ ,$$

The proof is provided in section 4.3.

Theorem 2 states that the solution to the approximated scenario problem described in 4-2 is valid even for the stationary process, i.e., the original problem 1-7, with a probability violation smaller than $\varepsilon(s^{*})+\chi_M/\delta$. The fraction χ_M/δ represents the loss in the guarantees due to the approximation done.

The controller developed shapes the stationary state distribution, so that the constraints are respected with a certain probability greater than a threshol. If the problem is convex, it is possible for the user to set the threshold by tuning the related parameters. This feature is loss when dealing with nonconvex problem, but it is the price to pay for the great increase in generality of the methodology. In nonconvex cases, however, the threshold for the controller found is known, and it can be used, together with the cost function, to decide if the results obtained are satisfactory or not. Thanks to the generality of this approach (no assumptions are made on the state matrix, the disturbances distribution and its knowledge) and to the fact that it is computed in one shot offline, it can be applied to a very large number of problems.

4.2 Complete problem with the recurrent exponential bound on the norm

For the sake of completeness, the version of the problem with the recurrent exponential bound is also reported, even though the total analogy with the previous section 4.1. Applying the recurrent exponential condition as expressed in equation 3-25 to the problem 1-7 gives

$$\begin{gathered}\min_{\gamma,K,h} h \\ s.t.\, \mathbb{P}_{d_k}\{l(x_{k,\infty},\gamma + Kx_{k,\infty}) \leq h-\delta \wedge f(x_{k,\infty},\gamma + Kx_{k,\infty}) \leq -\delta\} \geq 1-\varepsilon \\ \|(A+BK)^t\| \leq \lambda^{t-T+1}\ for\ t = T, T+1, \ldots, 2T\,.\end{gathered} \tag{4-4}$$

As done in the previous section, the scenario approach is applied, to move from a chance-constrained problem to a standard one, and then the state process truncation is adopted, to remove the dependance on infinitely long realizations. The approximated scenario program with the recurrent exponential bound is

$$\begin{gathered}\min_{\gamma,K,h} h \\ s.t.\, l\left(x_{k,M}^{(i)}, \gamma + Kx_{k,M}^{(i)}, d_k\right) \leq h-\delta \\ f\left(x_{k,M}^{(i)}, \gamma + Kx_{k,M}^{(i)}\right) \leq -\delta \\ x_{k,M}^{(i)} = (I-A-BK)^{-1}B\gamma + \sum_{s=0}^{M-1}(A+BK)^s d_{k-1-s}^{(i)} \\ i = 1, \ldots, N \\ \|(A+BK)^t\| \leq \lambda^{t-T+1}\ for\ t = T, T+1, \ldots, 2T\,.\end{gathered} \tag{4-5}$$

Theorem 3: given the solution (γ^*, K^*, h^*) to problem 4-5, fixed a confidence parameter $\beta \in (0,1)$, it holds that

$$\mathbb{P}_{d_k}^N\{\mathbb{P}_{d_k}\{g_{k,\infty}^M > 0\} \leq \varepsilon(s^M) + \frac{\chi_M}{\delta}\} \geq 1 - \beta \,.$$

where χ_M is equal to

$$\chi_M = L\left(\lambda \frac{1+\lambda^2 - 2\lambda^{T+1}}{(1-\lambda)(1-\lambda^{T+1})} - \sum_{s=T}^{M-1} \|(A+BK)^s\|\right)\sigma^2 \,.$$

The proof is given

The value of $\sum_{s=T}^{M-1}\|(A+BK)^s\|$ is limited by the condition imposed on the norm and it depends on the difference between T and M. Given that their value differs from case to case, it is not possible to give a general threshold always valid, but in every case it can be calculated. An example of a specific case was given in section 2.3.

As said, the only difference between theorem 2 and theorem 3 is in the value of χ_M, so all the comments given in the previous section for the problem with the exponential bound remains valid also in this case.

4.3 Proof of theorems 2 and 3

The introduction of a bound on the norm of the matrix $(A+BK)^s$ was motivated by the necessity of keeping a link between the stationary process $x_{k,\infty}$ to the truncated one $x_{k,M}$. The idea was to limit the difference between the constraints violation in the truncated case and in the stationary case, so that the bound for the violation probability found for the approximated scenario program could be applied to the initial problem. Thanks to the norm bound introduced in the previous chapter, we fill the gap. The proof applies for both theorems 2 and 3, with the only precaution of adjusting the proper value of χ_M in the two cases.

From the first chapter, we know that an upper limit for the violation probability of the solution of the approximated scenario program with confidence $1-\beta$:

$$\mathbb{P}^N_{d_k}\{\mathbb{P}_{d_k}\{g^*_{k,M} > -\delta\} \le \varepsilon(s^M)\} \ge 1-\beta. \tag{4-6}$$

The problem here is that the $g^*_{k,M}$ refers to the constraints in the approximated scenario program referring to the truncated solution, while the result in theorems 2 and 3 is about the satisfaction of $g^*_{k,\infty} \le 0$, i.e., the behavior of the solution to 4-2 in a stationary condition. The next step seen in chapter 2 was to rewrite the violation probability for the non-approximated program, as

$$\begin{aligned}\mathbb{P}_{d_k}\{g^*_{k,\infty} > 0\} &= \mathbb{P}_{d_k}\{g^*_{k,M} + \delta + g^*_{k,\infty} - g^*_{k,M} - \delta > 0\} \\ &\le \mathbb{P}_{d_k}\{g^*_{k,M} > -\delta \vee g^*_{k,\infty} - g^*_{k,M} > \delta\} \\ &\le \mathbb{P}_{d_k}\{g^*_{k,M} > -\delta\} + \mathbb{P}_{d_k}\{g^*_{k,\infty} - g^*_{k,M} > \delta\}\,,\end{aligned} \tag{4-7}$$

where the first term, $\mathbb{P}_{d_k}\{g^*_{k,M} > -\delta\}$, is bounded by equation 4-6, while the second can be bounded using the Chebyshev's inequality:

$$\mathbb{P}_{d_k}\{g^*_{k,\infty} - g^*_{k,M} > \delta\} = \mathbb{P}_{d_k}\{|g^*_{k,\infty} - g^*_{k,M}| > \delta\} \le \frac{\mathbb{E}[|g^*_{k,\infty} - g^*_{k,M}|]}{\delta}\,. \tag{4-8}$$

Thanks to the assumption on the Lipschitz continuity of the cost function and the constraint functions, it is possible to bound the difference between the two functions $g^*_{k,\infty} - g^*_{k,M}$ as a function of the mismatch between $x^*_{k,\infty}$ e $x^*_{k,M}$.

$$\begin{aligned}\mathbb{E}\big[\big|g^*_{k,\infty} - g^*_{k,M}\big|\big] &= \mathbb{E}\big[\big|g^*_{k,\infty} - g\big(x^*_{k,M}, \gamma^* + K^*x^*_{k,\infty}\big) + \\ &\qquad g\big(x^*_{k,M}, \gamma^* + K^*x^*_{k,\infty}\big) - g^*_{k,M}\big|\big] \\ &\le \mathbb{E}\big[\big|g\big(x^*_{k,\infty}, \gamma^* + K^*x^*_{k,\infty}\big) - g\big(x^*_{k,M}, \gamma^* + K^*x^*_{k,\infty}\big)\big|\big] + \\ &\qquad \mathbb{E}\big[\big|g\big(x^*_{k,M}, \gamma^* + K^*x^*_{k,\infty}\big) - g\big(x^*_{k,M}, \gamma^* + K^*x^*_{k,M}\big)\big|\big] \\ &\le \mathbb{E}\big[L\big\|x^*_{k,\infty} - x^*_{k,M}\big\|\big] + \mathbb{E}\big[L\big\|K^M\big(x^*_{k,\infty} - x^*_{k,M}\big)\big\|\big] \\ &\le L(1 + \|K^M\|)\mathbb{E}\big[\big\|x^*_{k,\infty} - x^*_{k,M}\big\|\big]\,.\end{aligned} \tag{4-9}$$

L is the maximum between the Lipschitz value of the cost function and the one of the constraint function. The norm $\|K^*\|$ is limited by the bound imposed on $\|(A + BK)^M\|$. It is not, however, easy to explicate it, so a solution could be to introduce a new requirement directly on $\|K^*\|$. This is necessary only if the cost function or the constraints depend on the input variable. If they depend only on the state, the difference $\big|g^*_{k,\infty} - g^*_{k,M}\big|$ is bounded by just $L\big\|x^*_{k,\infty} - x^*_{k,M}\big\|$, without the need to do any further reasoning about $\|K^*\|$. As done in the previous chapter, for the sake of simplicity, we will indicate $\left|g^*_{k,\infty} - g^*_{k,M}\right| \leq L\left\|x^*_{k,\infty} - x^*_{k,M}\right\|$, knowing that L could be multiplied for $1 + \|K^*\|$ if one of the functions depends on the control input.

The difference between the stationary state process and the approximated version is, thanks to the bound on the norm of $\|(A + BK)^s\|$ finite. We indeed have that

$$\mathbb{E}\left[\left\|x^*_{k,\infty} - x^*_{k,M}\right\|\right] = \tag{4-10}$$

$$\mathbb{E}\left[\left\|(I - A - BK)^{-1}B\gamma + \sum_{s=0}^{\infty}(A + BK)^s d_{k-1-s} - (I - A - BK)^{-1}B\gamma \right.\right.$$
$$\left.\left. + \sum_{s=0}^{M-1}(A + BK)^s d_{k-1-s}\right\|\right]$$
$$= \mathbb{E}\left[\left\|\sum_{s=M}^{\infty}(A + BK)^s d_{k-1-s}\right\|\right] \leq \mathbb{E}\left[\sum_{s=M}^{\infty}\|(A + BK)^s\| \, \|d_{k-1-s}\|\right]$$
$$= \sum_{s=M}^{\infty}\|(A + BK)^s\| \, \mathbb{E}[\|d_{k-1-s}\|] \, .$$

The summation $\sum_{s=M}^{\infty}\|(A + BK)^s\|$ is limited by in view of the results 3-10 and 3-23, where the proper bounds depends on the adopted problem (e.g. $\sum_{s=M}^{\infty}\|(A + BK)^s\| \leq \lambda^M/(1 - \lambda)$ in the case $\|(A + BK)^s\|$ is constrained to stay below to a classic exponential and likewise for the other case). $\mathbb{E}[\|d_{k-1-s}\|]$ is by assumption equal to σ^2. Altogether, we thus have

$$\mathbb{E}[|g^M_{k,\infty} - g^M_{k,M}|] \leq \chi_M \, , \tag{4-11}$$

which used in 4-8 gives

$$\mathbb{P}_{d_k}\{g^M_{k,\infty} - g^M_{k,M} > \delta\} \leq \frac{\mathbb{E}[|g^M_{k,\infty} - g^M_{k,M}|]}{\delta} \leq \frac{\chi_M}{\delta} \, . \tag{4-12}$$

This, together with 4-7, allows us to evaluate the violation probability for the stationary process based on the violation probability with the truncated one.

$$\mathbb{P}_{d_k}\{g^M_{k,\infty} > 0\} \leq \mathbb{P}_{d_k}\left\{g^M_{k,M} > -\delta\right\} + \mathbb{P}_{d_k}\left\{g^M_{k,\infty} - g^M_{k,M} > \delta\right\} \tag{4-13}$$
$$\leq \mathbb{P}_{d_k}\{g^M_{k,M} > -\delta\} + \frac{\chi_M}{\delta} \, .$$

The proof of the theorem is now completed. In indeed holds that

$$\mathbb{P}^N_{d_k}\left\{\mathbb{P}_{d_k}\{g^M_{k,\infty} > 0\} \leq \varepsilon(s^M) + \frac{\chi_M}{\delta}\right\} \geq \tag{4-14}$$

$$\geq \mathbb{P}_{d_k}^N \left\{\mathbb{P}_{d_k}\{g_{k,M}^M > -\delta\} + \frac{\chi_M}{\delta} \leq \varepsilon(s^M) + \frac{\chi_M}{\delta}\right\}$$
$$= \mathbb{P}_{d_k}^N\{\mathbb{P}_{d_k}\{g_{k,M}^M > -\delta\} < \varepsilon(s^M)\} \geq 1-\beta\ ,$$

which is the statement of theorems 2 and 3. This concludes the proof.

4.4 Choice of tunable parameters

Theorems 2 and 3 provide an evolution to the violation, with respect to the original constraints in 1-7, of the scenario solution in the vein of the results of 4-2:

$$V(\gamma^*, K^*, h^*) \leq \varepsilon(s^*) + \frac{\chi_M}{\delta} \qquad \text{4-15}$$

with confidence equal to $1-\beta$

The main differences, with respect to [9] is that ε is no more an user chosen reliability level. In [9] indeed, ε can be given as a requirement which can be satisfied by a priori selection of other user chosen parameters. In our case, ε can not be a hard requirement, because there are no guarantees it is possible to satisfy it. According to the wait&judge perspective, ε is a function of the complexity s^* and the value taken by it can span the whole range [0,1]. 4-15 has indeed to be interpreted thus as solution of the violation of the obtained solution. As compared with the scenario approach, in 4-15 we have to take a margin χ_M/δ over $\varepsilon(s^*)$ which is needed in order to account for the approximation given by the series truncation in the position of the scenario program as well as the tightening of constraints.

The tunable parameters are the number of scenarios N, the truncation value M, the confidence parameter β, the tightening δ, the base of the exponential λ, and, in the case of the recurrent exponential, the bound starting point T. The last two were not present in [9] and increase the possible tuning combination.

Interestingly, we can divide the tunable parameters in two groups, because on the one hand N and β determine, together with the cardinality of the support set, the value of the violation parameter ε, while χ_M/δ is unaffected by the remaining ones. Different considerations can thus be made for these two groups of parameters.

4.4.1 Choice of the parameters affecting χ_M/δ

β is the confidence with which we want to guarantee the result. So when tuning β, it must be considered that increasing its value the bound on the violation probability becomes more strict, but there are less guarantees on the effectiveness of the control system. ε depends logarithmically on β (see [24]), so it is possible to choose very small value without having a significant impact on ε. A difference of one order of magnitude in β corresponds roughly to a difference in one percentage point for the violation probability, see figure 4-1. With β equal to 10^{-6} or 10^{-7}, it is practically certain that $V(\gamma^*, K^*, h^*) \leq \varepsilon(s^*) + \chi_M/\delta$ is valid.

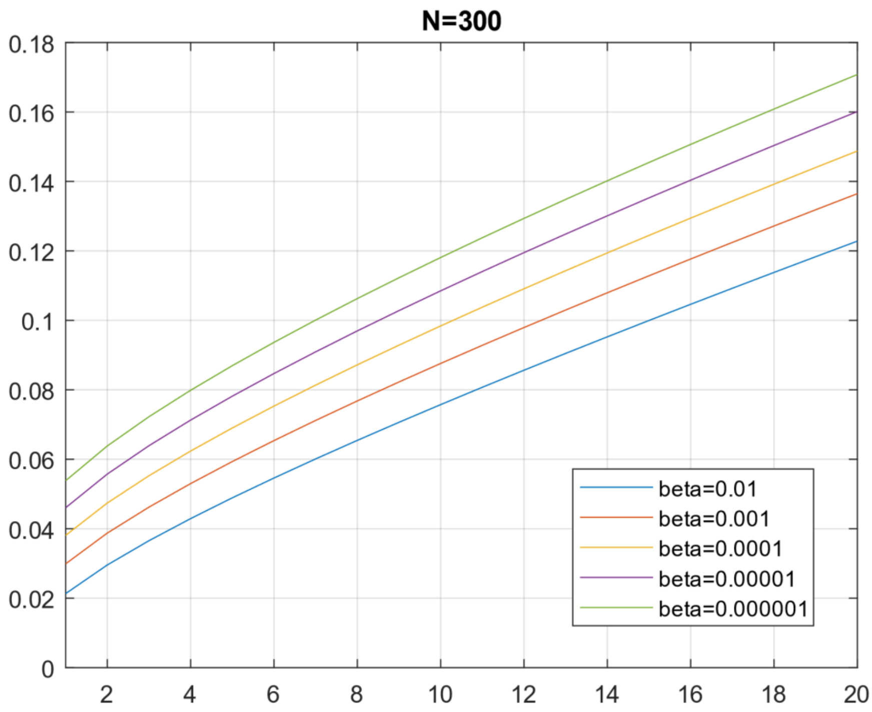


Figure 4-1 Relationship between the support set cardinality and the violation probability with N=300 for different values of β

ε is more influenced by the N, number of scenarios. The bigger its N, the better is the result, as it can be seen in figure 4-2 ($\varepsilon(s_N^*)$ tends to s_N^* when $N \to \infty$, see [24]). However, the number of scenarios is, together with the truncation value M, the biggest factor influencing the computational load. Increasing it too much leads to an extremely complex problem, which would require very long time to solve and without any guarantees on its feasibility. Considering also that N dictates the quantity of data needed, it is crucial, in case the data are obtained by measurements, to have a good value at the first time. If, after a first trial, it is necessary to increase the value of N, it could be indeed a problem to acquire more data. The choice of N has to balance these necessities, considering also the strictness of the other conditions. These choice is particularly hard since is not possible to know the cardinality of the support set, so the choice of N should be done considering the possible range of value ε can assume, instead of thinking to a specific value. In figure 3-1 the function $\varepsilon(\cdot)$ is represented for different value of N.

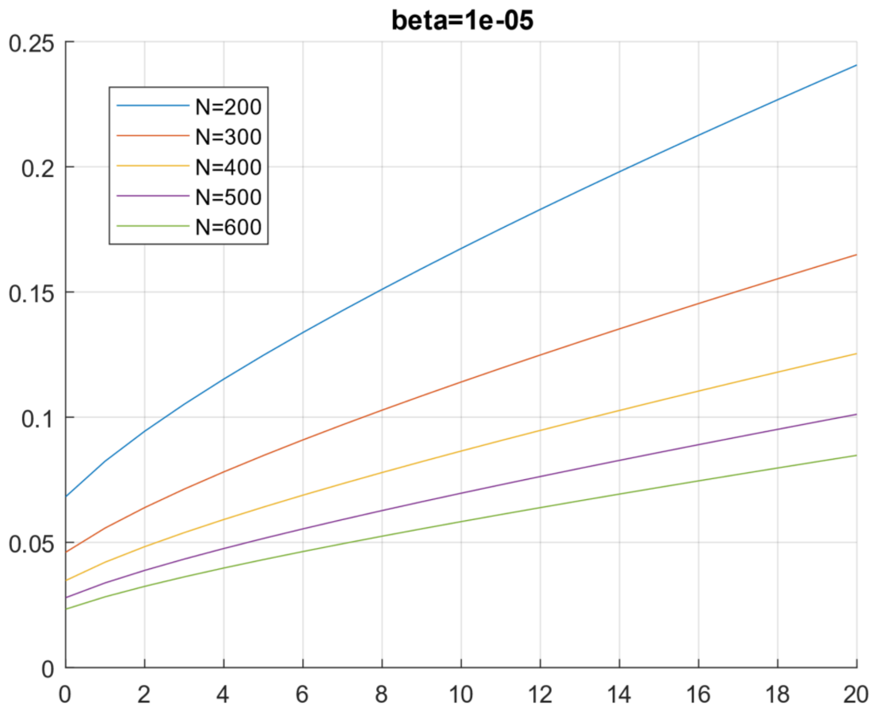


Figure 4-2 Relationship between the support set cardinality (on the x axis) and the violation probability expressed from 0 to 1 with $\beta=10^{-5}$ and for different values of N

4.4.2 Consideration on the probability tightening for the classic exponential condition

The best approach to tune the remaining parameter is starting with the value of χ_M/δ. While the cardinality of the support set, and so the value of ε, depends on the sample used and so is a probabilistic variable (except the case with convex functions), the value of χ_M/δ is decided by the user, through the selection of certain parameters. Its value represents the maximum acceptable margin of guarantees loosening. It should therefore decided first to then tune the other parameters so to obtain the desired value.

δ is the tightening applied to the constraint. Being at the denominator, we would like to set it as big as possible, to minimize the value of the fraction. The limit is given by the cost function and the feasibility. An excessively high value can make the problem unfeasible or can lead to a very big γ, with an incredibly high cost. The idea is therefore to try to maximize δ without getting out of proportion.

The truncation parameter M and the exponential base λ are the two value we can use to tune χ_M. The relation is, in case of exponential,

$$\chi_M = L\frac{\lambda^M}{1-\lambda}\sigma^2 \, . \qquad \text{4-16}$$

Increasing M, considering that $\lambda < 1$, leads to an exponential decrease in the value of χ_M. The trade-off is like the one present for the number of scenarios N. The quality improvements in the result are paid with the need of a greater pool of data and a bigger computational load. Differently from N, the value of M does not influence significatively the feasibility of the problem. The choice of λ regards instead the bound rigidity. A too small λ could lead to a condition too strict, making the problem unfeasible if the dimension of the control input is smaller than the number of states. The value of λ also influences the cost, due to the restriction in the number of possible solutions the program can be chosen between. The consideration here done for the case with an exponential bound on the norm can be replicated, in a similar way, also in the case with a recurrent exponential bound, and, more in general, for every bound on the norm. There will always be, in fact, a trade-off on the bound strictness: its increase it means a reduction of the value of χ_M, but also of the decision set. The compromise between the two must be found case by case and so for the value of M.

With respect to the situation in [9], where there are only M and δ to be tuned, the choice is more complex. While it is possible that M is given by the problem, but λ and

δ must be always picked by the user. The choice is highly related to the problem given. If the norm of the matrix A is already high, it could be better to loosen λ, using the constraints tightening to obtain the value desired. On the other hand, if the constraints are already strict, it is probably a more advisable to tauten λ, leaving more space in the choice of δ.

5. Support set identification

In the first chapter, we were interested in bounding the violation probability $V(\theta)$, which defines, for a certain solution θ, the probability that the constraints are respected when applying a new realization of the disturbance. The violation probability depends on the probability distribution of the disturbance, which is not known, therefore it is not possible to directly compute it. However, it is demonstrated that it is related to the cardinality of a support set as defined in chapter 1, definition 3, and this fact can be used to determine a bound. A support set is a set of scenarios which suffices to obtain the solution to the optimization problems with all the N scenarios in place. If the constraints are convex, the cardinality of the smallest irreducible support set is known to be not greater than the optimization variable dimension, as shown in [15]. In this way, it is always possible to bound the violation probability before even computing the solution. In the case of nonconvex constraints, this is not possible, and a different approach, called wait&judge, is used proposed in [15] and in [19]. The bound on the violation probability is given only after the solution is computed, based on the cardinality of any support set. With the solution at the disposal, it is possible to compute a support set based on its definition, i.e., trying to obtain the same solution using a reduced number of scenarios. There are many different approaches to do it, with huge differences in terms of effectiveness and computational time, so it is useful to spend some time looking for a good algorithm.

5.1 Support set identification methods for nonconvex scenario

The cardinality of the identified support set determines the reliability guarantee. The smaller it is, the better is the guarantee on the results. It is important to remember that, from a theoretical point of view, minimality is not required. However, given the importance it has on the results, it is essential to find an algorithm able to find a support set as small as possible in a limited time. In fact, although it is always possible to find the irreducible support set with minimal cardinality, this would require an exhausting by brute force search over every possible subset of scenarios, and considering the number of scenarios we are dealing with, the task would be overwhelming. Just to give an idea of the entity of the calculation, imagine having a problem with 200 scenarios where the smallest irreducible support set has cardinality 4 (which is not known beforehand). If we try every possible support set, starting from these with cardinality 1, we have to try 200 times using just one scenario, 19'900 times (the number of possible combinations considering that the order does not matter, so it is $(200 * 199)/2$) using two scenarios, 1'313'400 times using three scenarios and then a bigger number of times using four scenarios, but that depends on the specific case. It should now be evident the necessity to have a good algorithm.

So we need an algorithm that can be regarded as a function $F: (\delta^{(1)}, \delta^{(2)}, \dots, \delta^{(N)}) \to \{i_1, i_2, \dots, i_k\}$ where $(\delta^{(i_1)}, \delta^{(i_2)}, \dots, \delta^{(i_k)})$ is a support set. The cardinality of the support set is then:

$$s_N^* := |F(\delta^{(1)}, \delta^{(2)}, \dots, \delta^{(N)})| \,. \qquad 5\text{-}1$$

Among the various possible approaches, I will present three algorithms, the first two are more basic, while the third is a combination of the other two.

5.2 The simple greedy algorithm

The first approach is a simple greedy algorithm presented in [20] which works as follow:

1. Set $L \leftarrow (\delta^{(1)}, \dots, \delta^{(N)})$ and compute $K_N^* \leftarrow$ solution of the related program nonconvex problem;
2. For all $i = 1, .., N$:
    - Let $L' \leftarrow L \backslash \delta^{(i)}$, form the program $NCSP^{(L')}$ with the constraint in L', and let K^*be its solution;
    - If $K^* = K_N^*$, set $L \leftarrow L'$;
3. Output the set $\{i_1, i_2, \dots, i_k\}$ of the indices of the elements in L.

The output of this algorithm can be particularly subject to the order in which the elements $(\delta^{(1)}, \delta^{(2)}, \dots, \delta^{(N)})$ appear, increasing the uncertainty about the result. Another limit of this algorithm is that start using all the scenarios, with two major consequences. First, the number of times that perform the optimization is, at least, N, the number of starting scenarios. Second, given that it removes one scenario at a time, the first optimizations are done with a great number of scenarios, with a longer computational time. To obtain a good reliability guarantee is necessary to have a high N and to find a support set with a small cardinality. It is therefore evident how a proportional dependence from N can drastically affect the time needed to find the support set. The support subsample found is guaranteed to be irreducible, but it can be not the one of minimal length.

5.3 The incremental algorithm

The second algorithm perform on average less cycles than the greedy algorithm. Also, the optimization problem to be solved have an increasing number of scenarios, starting from one, so that the computational time is reduced. The cons is that there isn't a formal guarantee to find an irreducible support set, even though in practice the algorithm proved effective for the problem at hand The idea behind the incremental algorithm is opposite to that of the greedy algorithm. Instead of finding a support set removing one element at time, it creates one adding one scenario at time. In this way, the time needed does not depend anymore on N, but potentially just on the cardinality of the found support set. The main problem to solve is in which order the scenarios have to be added. Choosing the wrong order, a scenario that is essential to construct any support set could potentially put in the last position and the algorithm would find as support set the whole set of scenarios. This example shows the crucial role that the order has. An effective solution has been found by looking at the support sets obtained with the greedy algorithm. In most cases, the scenarios composing the support set were the ones for which $f(x,u)$ was closest to 0 (remember that the constraint requires $f(x,u) \leq 0$). This makes sense intuitively: changing the parameters values of a small amount, is more probable that the scenarios for which $f(x,u)$ is close to 0 will violate the constraint, with respect to the other. An algorithm that orders the scenario based on the value of $f(x,u)$ was created and the support set is obtained by adding one scenario at time from the nearest one to the most distant. The steps are the following:

1. Compute K_N^* from the complete set $(\delta^{(1)}, \dots, \delta^{(N)})$;
2. Using K_N^* and the disturbance realizations $(\delta^{(1)}, \dots, \delta^{(N)})$, calculate the states after M step x_M;
3. Order the realizations based on the value of $f(x,u)$ for $x = x_{K,M}$, obtaining $(\delta^{(i_1)}, \dots, \delta^{(i_N)})$. The case of two or more points are at the same distance is rare and their order is generally not significant. For this reason, no further criterions were evaluated.
4. For all $n = 1, \dots, N$:
   - Let $L' \leftarrow L \cup \delta^{(i_n)}$, form the program $NCSP^{(L')}$ with the constraint in L', and let K^*be its solution;
   - If $K^* = K_N^*$, set $L \leftarrow L'$ and exit the "for" cycle;
5. Output the set $\{i_g, \dots, i_k\}$ of the indices of the elements in L.

This algorithm proved to be much more efficient than the previous one in the case study developed in chapter 6. It generally finds small support set at a much lower computational effort. However, it may be that a scenario that is necessary for the minimal support set is not among the first in the order. In that case, the algorithm finds a support set significatively bigger than the minimal one, and the computational time is affected too. In spite of this drawback, the results are definitively good and it is often the case that the incremental algorithm outperforms the greedy one.

5.4 The incremental greedy algorithm

The third algorithm solves the weak spot highlighted in the incremental algorithm by combining it with the simple greedy one. The idea is to apply completely the second algorithm and then to reduce the support set found using the first algorithm. In this way, we keep the benefit given from the incremental algorithm (to not be dependent on N, far less computational load than the first), solving its problem and further improving the results. In fact, even when the support set is already small, for example with cardinality 4 or 5, applying the simple greedy algorithm often further reduced the support set. With respect to the second algorithm, the computational load and the time needed are increased by roughly a factor 2, but the improvements in the result make it worth using. The support subsample found with this algorithm is guaranteed to be irreducible, but again it could be not the one with the smallest cardinality.

The incremental greedy algorithm is the following:

1. Compute K_N^* from the complete set $(\delta^{(1)},\dots,\delta^{(N)})$;
2. Using K_N^* and the disturbance realizations $(\delta^{(1)},\dots,\delta^{(N)})$, calculate the states after M step x_M;
3. Order the realizations based on the distance between the states after M step and the constraint, obtaining $(\delta^{(i_1)},\dots,\delta^{(i_N)})$. The case of two or more points are at the same distance is rare and their order is generally not significant. For this reason, no further criterions are evaluated;
4. For all $n = 1,..,N$:
   - Let $L' \leftarrow L \cup \delta^{(i_n)}$, form the program $NCSP^{(L')}$ with the constraint in L', and let K^*be its solution;
   - If $K^* = K_N^*$, set $L \leftarrow L'$ and exit the "for" cycle;
5. Considering the set $\{i_1,\dots,i_k\}$ of the indices of the elements in L, $\forall i \in \{i_1,\dots,i_k\}$:
   - Let $L' \leftarrow L \backslash \delta^{(i)}$, form the program $NCSP^{(L')}$ with the constraint in L', and let K^*be its solution;
   - If $K^* = K_N^*$, set $L \leftarrow L'$;
6. Output the set $\{i_1, i_2,\dots,i_j\}$ of the indices of the elements in L.

6. Simulation example

This chapter is meant to illustrate the approach previously presented by means of a simple yet not simplistic example. Many experiments were made, here we report the most significant.

6.1 Problem setup

In both the two simulation examples, we will study states of the second order, in the first case with a single control input, while in the second with two. In the following subsections, the common aspects of the two examples are discussed.

6.1.1 Cost function

The cost function used is a quadratic function evaluating the controller parameters K e γ, expressed as

$$l(K,\gamma) = \sum_{i=1}^{n_u} \sum_{j=1}^{n_x} p_{i,j} \left(K_{i,j}\right)^2 + \sum_{h=1}^{n_u} q_h (\gamma_h)^2. \tag{6-1}$$

This type of cost function allows a great personalization, with the possibility to choose a different weight for each controller parameter. In the two examples proposed, the values taken are $p_{i,j} = 1\ \forall i, \forall j$ and $q_h = 1000\ \forall h$. The bigger weights imposed for γ should lead the optimization program to rely more on K and to use γ only if really needed.

6.1.2 Constraint applied

The constraint function $f(x,u)$ imposes a constraint to the pair (x,u). As for the cost function, there is no assumption on its convexity, on the contrary we will treat them as if it is nonconvex.

In the two simulation examples, a simple linear constraint is used, and which is expressed by the following equation:

$$Fx \leq v \, . \qquad 6\text{-}2$$

In the 2-dimensioanl case, it reduces to ask that all the points stay below or over one or more straight lines. When the disturbance is generated, the evolution of the states without control action is plotted, to confront it with the constraint.

In addition to the constraint on the states value, a constraint on the norm of $A+BK$ is applied. Three different functions, based on what seen in chapter 3, have been developed. The first one just requires that the norm of $A+BK$ is smaller than one. It is the simplest one and it can be useful to compare with the other two. The second includes the condition 3-7, which is:

$$\| (A+BK)^t \| \leq \lambda^t \; for \; t = T, T+1, \ldots, 2T-1 \, . \qquad 6\text{-}3$$

In practice it requires that the norm of the exponential of the matrix lays under an exponential with base λ form a certain T over. The third function includes the condition 3-25:

$$\| (A+BK)^t \| \leq \lambda^{t-T+1} \; for \; t = T, T+1, \ldots, 2T \, . \qquad 6\text{-}4$$

Similar to the previous one, it is less strict, as we have seen in the previous chapters.

In the two simulation examples, the second function is used.

6.1.3 Noise simulation

While in a real practical case the scenarios must be sampled independently one from the other, according to the Δ distribution. In the case developed, the disturbance is generated randomly according to a certain distribution. I used the normal distribution, with the possibility to apply standard deviations σ for every state. For the normal distribution, the general form of the probability density function is

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} . \qquad \text{6-5}$$

However, the distribution type is totally irrelevant in the scenario approach, the only strict requirement is the independence in the sampling. Therefore, it is possible to apply to problems with any type of distribution.

For the two cases presented, the disturbance used has zero mean, as required by the assumptions, and all the standard deviations equal to 1.

6.1.4 fmincon

The function used to find an optimal solution for the problem is fmincon. It is a function available on Matlab that finds the minimum of constrained nonlinear multivariable functions. We pass to it the cost function, the constraint function, and an initial guess from the parameters. It is possible to impose other requirements, such as an upper and lower bound for the parameters, or linear inequalities.

The initial guess can influence the results of the optimization with sometimes significative difference. If it is not possible to have a good initial guess it is recommended to run the program multiple times changing the starting point to have more guarantees on the quality of the result.

6.2 First example

6.2.1 System choice

The system has two states, but only one control input. The state space matrix is

$$A = \begin{bmatrix} 0.95 & 0.2 \\ 0.8 & 0.8 \end{bmatrix}. \tag{6-6}$$

The system is not stable, the two eigenvalues are 1.282 and 0.468, so the controller will have to stabilize it.

The system has just one control input and the control matrix is

$$B = \begin{bmatrix} 1 \\ 0.3 \end{bmatrix}. \tag{6-7}$$

The values of the other main parameters are reported in the following table. The only constraint on the states is to stay below a given straight line, so $v_1 x_1 + v_2 x_2 - v_{max} \leq 0$. The constraint has been chosen so that, when there is no control applied, a high number of realizations violates it. The controller, therefore, has to do a strong reshaping of the state distribution.

Name	**Symbol**	**Value**
Number of scenarios	N	400
Truncation value	M	20
Confidence parameter	β	10^{-5}
Norm bound coefficient	λ	0.95
Constraint coefficient for the first state	V_1	1
Constraint coefficient for the second state	V_2	2
Constraint constant coefficient	V_{max}	2

6.2.2 Results

From image x it is possible to see how would evolve the system after M steps without control (the orange points). In blue, near the center and harder to notice, there are the states after the same number of steps and affected by the same disturbance when the control action developed is applied. In the next image it is possible to better appreciate the job done by the controller.

The result is:

$$K = [-0.1866 \quad -0.1699] \tag{6-8}$$
$$A + BK = \begin{bmatrix} 0.7634 & 0.0301 \\ 0.7440 & 0.7490 \end{bmatrix}$$
$$\gamma = -0.7553\ .$$

It is possible to see that the dependance of the first state from the second is practically canceled. Now, the two eigenvalues are 0.906 and 0.606, so the system is asymptotically stable. γ has a quite high value due to the small distance of the constraint from the origin. The controller needs to move the average position away from the center, to avoid crossing the constraint in case the disturbance as a high value.

In Figure 6-1, it is possible to see in red the states after M step without control and in blue with the developed control. The orange line is the constraint that has to be respected. The state has to stay below the line. Figure 6-2 is a zoom on the center of image x to better highlight the controlled states. As it is possible to see, the points are made more compact and there is a small change in the overall direction. All the points are now below the constraint.

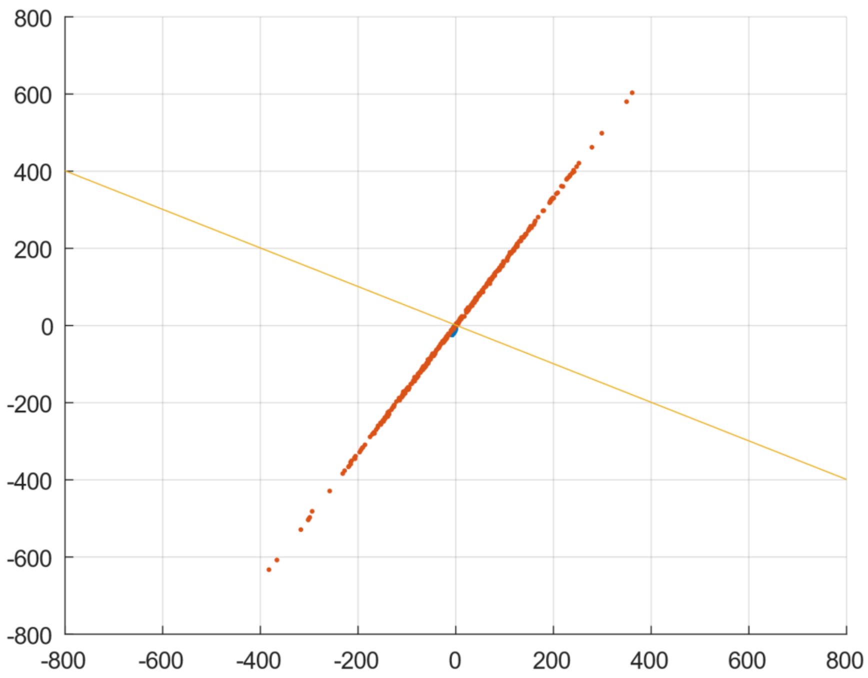


Figure 6-1 The non controlled states are in orange, while the controlled ones in blue. The yellow straight line delimits the area in which the states should stay.

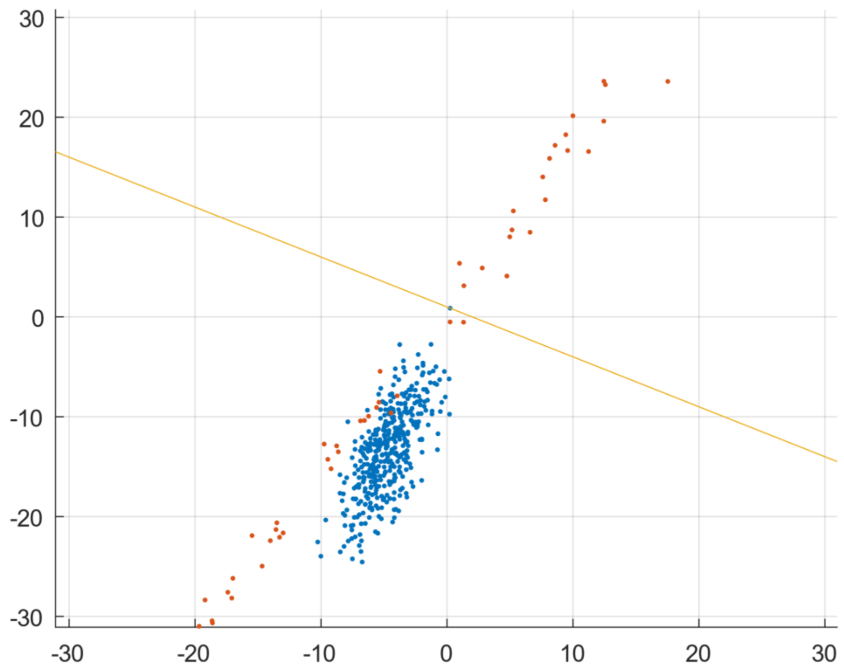


Figure 6-2 Zoom of the image proposed in Figure 6-1. Here it is possible to better appreciate the state distribution when the control is applied.

This case is particularly interesting because the norm of the matrix $(A + BK)^s$ starts over the value of 1, it has an overshoot, but after the 20th step it lays under the bound, as required (see Figure 6-3). The optimization program has used the increased freedom available thanks to the less strict requirement. If the norm bound was required from the first step, the program would have had less option and would have found a solution with a greater cost. This is a limit case, where the program takes all the space it has, waiting until the last step before respecting the bound. But it is not a case so rare, it can happen often, if the initial matrix has the norm bigger than 1.

This case is particular also because the support set is composed of just one scenario. Together with the elevated number of scenarios used, it allows to have a low bound for the violation probability.

An example of a test conducted to evaluate the quality of the resulting controller is shown in Figure 6-4.

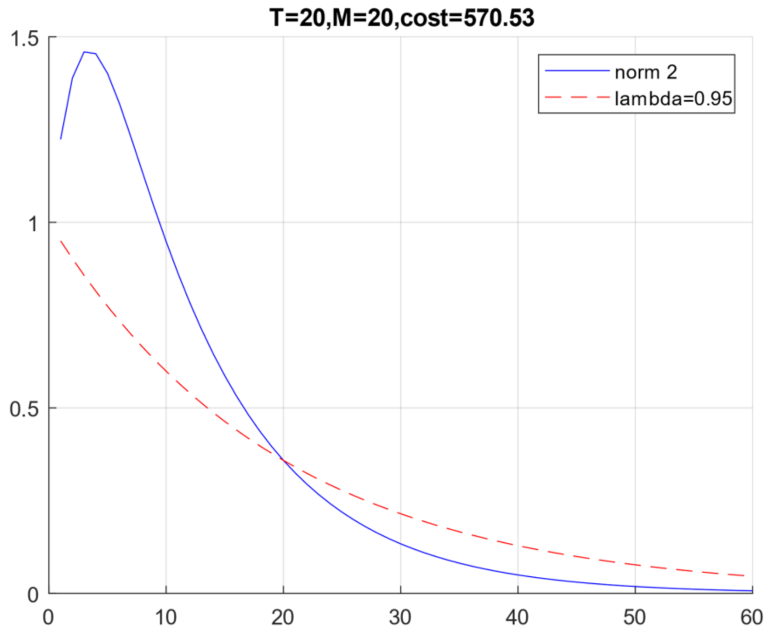


Figure 6-3 Evolution of the norm of the matrix $(A + BK)^s$ as s increases compared to the bound (which is valid for value equal or greater than 20)

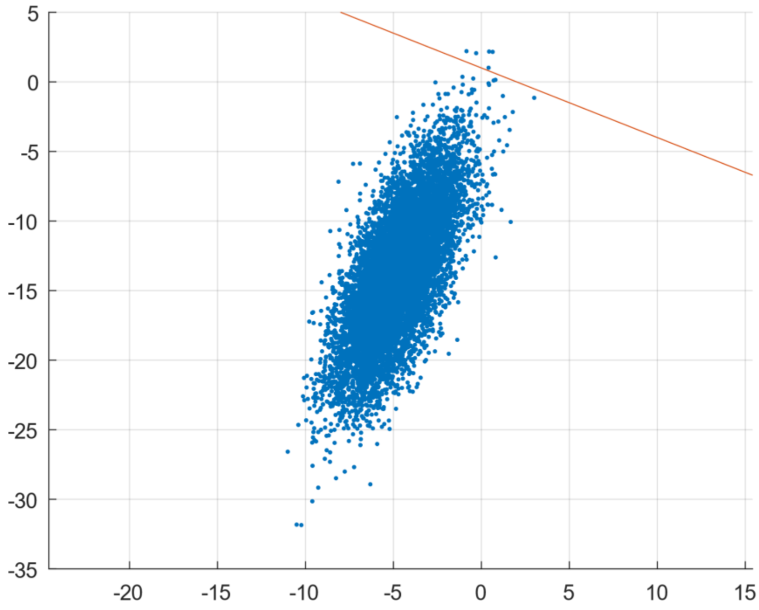


Figure 6-4 Results of a test conducted using the controller developed.

6.3 Second example

6.3.1 System choice

In this case, the system has two states and two inputs, giving more control over the state distribution. The state matrix is

$$A = \begin{bmatrix} 0.9 & -0.1 \\ 0.2 & 0.6 \end{bmatrix}. \tag{6-9}$$

It is asymptotically stable, with the two eigenvalues being 0.7 and 0.8. The control input matrix is

$$B = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}. \qquad \text{6-10}$$

The controller must maintain the states between two parallel straight lines. The first constraint is $v_1 x_1 + v_2 x_2 - v_{max} \le 0$, while the second is $-w_1 x_1 - w_2 x_2 + w_{max} \le 0$.

Name	**Symbol**	**Value**
Number of scenarios	N	200
Truncation value	M	20
Confidence parameter	β	10^{-5}
Norm bound coefficient	λ	0.72
Constraint coefficient for the first state	V_1	1
Constraint coefficient for the second state	V_2	0.2
Constraint constant coefficient	V_{max}	2.25
Constraint coefficient for the first state	W_1	1
Constraint coefficient for the second state	W_2	0.2
Constraint constant coefficient	W_{max}	2.25

6.3.2 Results

The resulting controller is:

$$K = \begin{bmatrix} -0.4164 & 0.1609 \\ -2.2481 & -0.3233 \end{bmatrix} \qquad \text{6-11}$$
$$A + BK = \begin{bmatrix} 0.4836 & 0.0609 \\ -2.0481 & 0.2767 \end{bmatrix}$$
$$\gamma = \begin{bmatrix} 0.0752 \\ 0.0394 \end{bmatrix}.$$

Differently from the previous case, the values for γ are almost negligible. While with only one linear constraint it is useful to move away the points using a constant term in the control input, it is not true in this situation.

The two constraints are facing one another and moving away from one means implicates getting close to the other. Also, thanks to the fact that the number of inputs is equal to the state dimension, there is more control on the values of $A + BK$ and, therefore, the state distribution shape. For these two reasons, and the fact that γ is associated to a high cost, its value is near zero. K alone is, in this case, sufficient to reshape the distribution and to guarantee the constraints observance. In Figure 6-5, it is possible to compare the states without control, in orange, and the controlled one, in blue. The direction among which are directed the noncontrolled points is completely changed by the controller. Thanks to its intervention, the direction is rotated to be parallel to the two straight lines used as constraints.

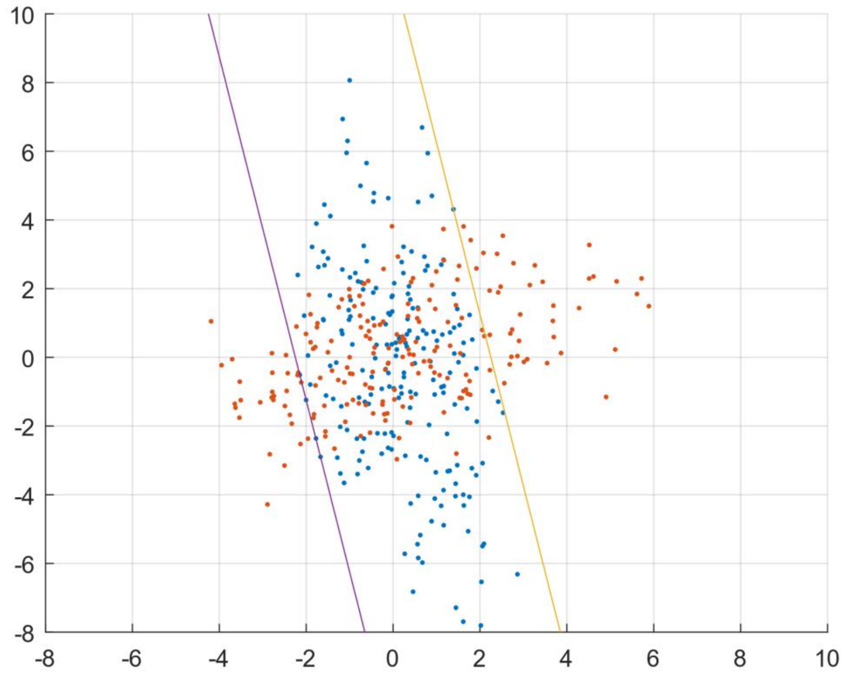


Figure 6-5 Comparison between the non controlled states and the controlled ones.

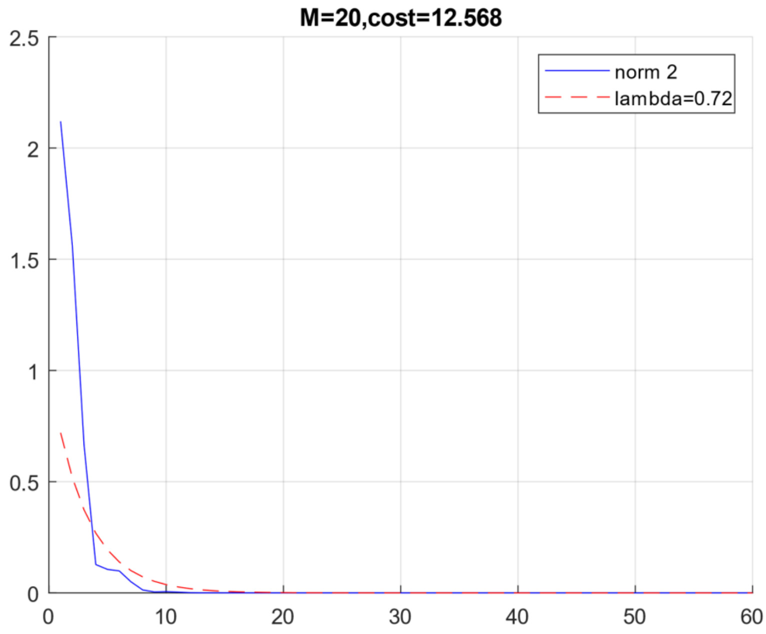


Figure 6-6 Norm of the matrix $(A + BK)^s$ as a function of s (in blue). Dotted in red the bound applied.

The norm of the matrix $A + BK$ starts over 2, but it rapidly decreases to zero, respecting the condition imposed (see Figure 6-6).

The support set found contains four scenarios, so, considering that the number of scenarios used is 200 and β is 10^{-5}, the best epsilon guaranteed is 11.52%. Considering that $\sigma = \sqrt[2]{\mathbb{E}[\| d_{k-1-s} \|_2]} = 1$ and L=1, the value of χ_M is:

$$\chi_M = L\frac{\lambda^M}{1-\lambda}\sigma^2 = \frac{0.72^{20}}{1-0.72} = 0.005\,. \tag{6-12}$$

Being a simulation, we do not have a value, but using 0.1, which corresponds to a reduction of ca 10% of the area between the two straight lines, we would obtain

$$\varepsilon(s^M) + \frac{\chi_M}{\delta} = 0.1152 + \frac{0.005}{0.1} = 0.1652\,. \tag{6-13}$$

Therefore, with a probability greater than 99.99%, $\mathbb{P}_{d_k}\{g_{k,\infty}^M > 0\} \leq 16.52\%$.

In a test done using 10000 realizations, the error percentage was of 2% (see Figure 6-7).

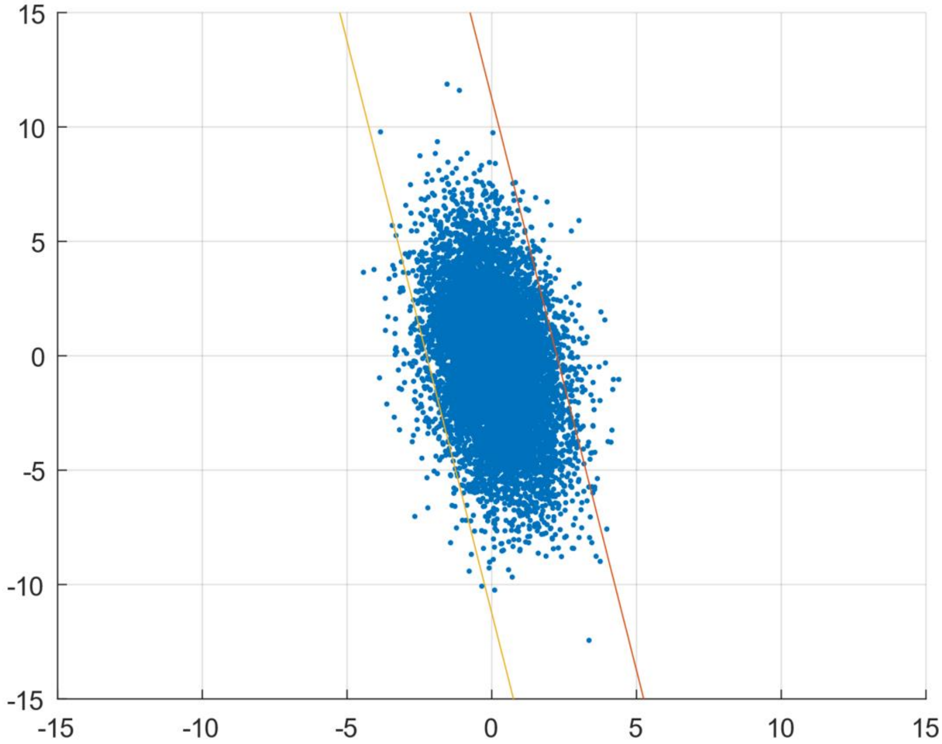


Figure 6-7 Results of a test conducted using the controller developed.

7. Conclusion

The work done in the thesis led to the development of a new method to solve chance-constrained optimization program, which are generally hard to tackle with traditional approach. The main contribution of the thesis is the combination of the scenario approach, effective when working with probabilistic constraints, with the approximation of the stationary state process, necessary to use finitely-long realizations of the disturbance. In this way, the guarantees given for the approximated problem, which were provided by the scenario approach, could be extended to the original problem, thanks to a convenient tightening of the constraints. The use of a state-feedback control and the absence of assumption on the convexity of the constraint function make this approach suitable for a high number of problems. Also, the results obtained testing the method on numerical examples were very satisfactory, with the program being able to find optimal solutions even for strict setups.

Further development of this methodology could include a different approach in the use of the scenarios. In fact, a limit in the scenario approach is that, when the number of scenarios used tends towards infinity, the program is solving all the constraints, and so the solution tends towards the robust one. If, however, some of the sampled scenarios are removed, the solution would better fit the chance-constrained problem. With this approach, if any bad instances is sampled, it would now be discarded and would not, therefore, affect the solution.

8. Bibliografia

[1] K. Åström, Introduction to stochastic control theory, Academic Press, 1970.

[2] D. Clarke e R. Hastings-James, «Design of digital controllers for randomly disturbed systems,» *Proceedings of the Institution of Electrical Engineers,* vol. 118, n. 10, pp. 1503-1506, 1971.

[3] P. J. Gawthrop, «Minimum-variance control,» *Encyclopedia of life support systems (EOLSS),* 2004.

[4] M. Grimble, «Controller performance benchmarking and tuning using generalised minimum variance control,» *Automatica,* vol. 38, n. 12, pp. 2111-2119, 2002.

[5] B. Huang, «Minimum variance control and performance assessment of time-variant processes,» *Journal of Process Control,* vol. 12, n. 6, pp. 707-719, 2002.

[6] V. Peterka, «On steady state minimum variance control strategy,» *Kybernetika,* vol. 8, n. 3, pp. 219-232, 1972.

[7] U. Shaked e P. Kumar, «Minimum variance control of discrete time multivariable armax systems,» *SIAM Journal on Control and Optimization,* vol. 24, n. 3, pp. 396-411, 1986.

[8] A. Mesbah, «Stochastic model predictive control: an overview and perspectives for future research,» *IEEE Control Systems Magazine,,* vol. 36, n. 6, pp. 30-44, 2016.

[9] A. Falsone, L. Deori, D. Ioli, S. Garatti e M. Prandini, «Optimal disturbance compensation for constrained linear systems operating in stationary conditions: A scenario based approach,» *Automatica,* 2019.

[10] G. C. Calafiore e M. C. Campi, «Uncertain convex programs: randomized solutions and confidence levels,» *Mathematical programming,* vol. 102, pp. 25-46, 2005.

[11] G. C. Calafiore e M. C. Campi, «The scenario approach to robust control design,» *IEEE Transactions on Automatic Control,* vol. 51, n. 5, pp. 742-753, 2006.

[12] M. C. Campi, S. Garatti e M. Prandini, «The scenario approach for system and control design,» 2009.

[13] K. You, R. Tempo e X. Pei, «Distributed algorithms for robust convex optimization via the scenario approach,» *IEEE Transactions on automatic control,* vol. 64, n. 3, pp. 880-

894, 2019.

[14] M. C. Campi e S. Garatti, «The exact feasibility of randomized solutions of uncertain convex programs,» *SIAM Journal of optimization,* vol. 19, n. 3, pp. 1211-1230, 2008.

[15] M. C. Campi e S. Garatti, «Wait-and-judge scenario optimization,» *Math. Program.,* vol. 167, n. 1, pp. 155-189, 2018.

[16] A. Carè, S. Garatti e M. C. Campi, «Scenario min-max optimization and the risk of empirical costs,» *SIAM Journal on Optimization,* vol. 25, n. 4, pp. 2061-2080, 2015.

[17] S. Garatti e M. C. Campi, «The risk of making decisions from data through the lens of the scenario approach,» in *19th IFAC Symposium on System Identification*, (Virtual) Padova, Italy, 2021.

[18] S. Garatti e M. C. Campi, «Risk and complexity in scenario optimization,» *Mathematical programming,* 2019.

[19] M. C. Campi, S. Garatti e F. A. Ramponi, «A general scenario theory for nonconvex optimization and decision making,» *IEEE transactions on automatic control,* vol. 63, n. 12, pp. 4067-4078, 2018.

[20] M. C. Campi, S. Garatti e F. A. Ramponi, «Non-convex scenario optimization with application to system identification,» in *IEEE 54th Annual Conference on Decision and Control (CDC)*, Osaka, Japan, 2015.

[21] Y. Yang e C. Sutanto, «Chance-constrained optimization for nonconvex programs using scenario-based methods,» *ISA Transactions,* vol. 90, pp. 157-168, 2019.

[22] M. Vidyasagar, «Randomized algorithms for robust controller synthesis using statistical learning theory,» *Automatica,* vol. 37, n. 10, pp. 1515-1528, 2001.

[23] G. Alsmeyer, «Chebyshev's Inequality,» *International Encyclopedia of Statistical Science,* pp. 239-240, 2011.

[24] S. Garatti e M. C. Campi, «On the consistency of the risk evolution in the scenario approach,» in *60th conference on th Decision and Control*, 2021.

[25] M. C. Campi e S. Garatti, Introduction to the scenario approach, Philadelphia: Mathematical Optimization Society & Society for Industrial and Applied Mathematics, 2018.

9. Appendix

The following MATLAB code returns (k), k = 0, 1,..., d, for user assigned d, N, and β.

```
function out = epsilon(d,N,bet)
out = zeros(d+1,1);
for k = 0:d
  m = [k:1:N];
  aux1 = sum(triu(log(ones(N-k+1,1)*m),1),2);
  aux2 = sum(triu(log(ones(N-k+1,1)*(m-k)),1),2);
  coeffs = aux2-aux1;
  t1 = 0;
  t2 = 1;
  while t2-t1 > 1e-10 t = (t1+t2)/2;
    val = 1 - bet/(N+1)*sum( exp(coeffs-(N-m')*log(t)) );
    if val >= 0
      t2 = t;
    else
      t1 = t;
    end
  end
  out(k+1) = 1-t1;
end
```

Acknowledgement

Countless people supported me during these years at university. I would like to thanks Professors Simone Garatti, Maria Prandini, Alessandro Falsone and Diego Regruto for supporting and supervising my thesis and all the other professors I had, for guiding me in the engineering world. I am grateful to all my friends, the ones I lived with, the ones I met during these years and the ones I already knew, for all the experiences we have done together in these five years. But most of all, I would like to thank my parents and my brother, on whom I can always rely.