

Scenario Reduction with Guarantees for Stochastic Optimal Control of Linear Systems

Cordiano, Francesco; De Schutter, Bart

DOI

[10.23919/ECC64448.2024.10590866](https://doi.org/10.23919/ECC64448.2024.10590866)

Publication date

2024

Document Version

Final published version

Published in

Proceedings of the European Control Conference, ECC 2024

Citation (APA)

Cordiano, F., & De Schutter, B. (2024). Scenario Reduction with Guarantees for Stochastic Optimal Control of Linear Systems. In *Proceedings of the European Control Conference, ECC 2024* (pp. 3502-3508). IEEE. <https://doi.org/10.23919/ECC64448.2024.10590866>

Important note

To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Copyright

Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy

Please contact us and provide details if you believe this document breaches copyrights.
We will remove access to the work immediately and investigate your claim.

Green Open Access added to TU Delft Institutional Repository

'You share, we take care!' - Taverne project

<https://www.openaccess.nl/en/you-share-we-take-care>

Otherwise as indicated in the copyright section: the publisher is the copyright holder of this work and the author uses the Dutch legislation to make this work public.

Scenario Reduction with Guarantees for Stochastic Optimal Control of Linear Systems

Francesco Cordiano, Bart De Schutter

Abstract—Scenario reduction algorithms can be an effective means to provide a tractable description of the uncertainty in optimal control problems. However, they might significantly compromise the performance of the controlled system. In this paper, we propose a method to compensate for the effect of scenario reduction on stochastic optimal control problems for chance-constrained linear systems with additive uncertainty. We consider a setting in which the uncertainty has a discrete distribution, where the number of possible realizations is large. We then propose a reduction algorithm with a problem-dependent loss function, and we define sufficient conditions on the stochastic optimal control problem to ensure out-of-sample guarantees (i.e., against the original distribution of the uncertainty) for the controlled system in terms of performance and chance constraint satisfaction. Finally, we demonstrate the effectiveness of the approach on a numerical example.

I. INTRODUCTION

Dealing with uncertainty is one of the major challenges in optimal control and decision making. Indeed, several real-world applications involve some level of uncertainty, which can negatively affect the system performance in terms of cost, safety, and reliability. Stochastic optimal control [1] has received significant attention in the recent years due to the possibility of explicitly encoding a probabilistic description of the uncertainty in the optimal control problem.

However, one of the main challenges in stochastic optimal control problems is to provide a description of the uncertainty such that: a) it is accurate enough so that performance guarantees can be given for the controlled system; and b) the resulting optimal control problem is numerically tractable and amenable to online optimization. In this regard, a popular technique to represent the uncertainty in a tractable way consists of employing a set of *scenarios*. Scenario-based methods are rooted in multistage stochastic programs [2], where a scenario tree is built over the prediction horizon according to the possible realization of the uncertainty [3]. More generally, scenarios can be conceived as time series with an associated probability, which allows the representation of complex stochastic phenomena, e.g., the price of an asset in trading problems [4], [5], the solar irradiance and the ambient temperature in building heating systems [6], or the degree of deterioration of railways in preventive maintenance [7]. Scenario-based methods rely on an atomic

description of the uncertainty, which is either assumed to be given (e.g., obtained from a forecast [8]), or estimated via scenario generation methods [5], [6], or obtained from a continuous distribution via discretization [9].

However, one of the main shortcomings of scenario-based methods is that a large number of scenarios is required to obtain an accurate representation of the uncertainty, especially in the case of a continuous distribution, which is typically approximated by means of a sample-average approximation [10]. In general, to tackle the issue of the growing complexity, several scenario reduction algorithms have been proposed in the context of stochastic programming, where the aim is to determine a set of scenarios with reduced cardinality that can well approximate some characteristics of the original scenario set. Examples of reduction algorithms can be found in [11], [12], where the accuracy of the approximation is measured in terms of a probability distance (e.g., the Wasserstein distance [13]) between the distributions describing the initial and reduced scenario set, or in [14], [15], where scenarios are grouped according to their similarity using clustering-based algorithms.

Using a limited number of scenarios is justified by quantitative stability results for stochastic programs [16], [17], where it is proven that the optimal value of a stochastic program, or the set of optimal solutions, possesses some type of continuity properties with respect to changes in the probability distribution of the uncertainty. However, when only a limited number of scenarios is used in a stochastic optimal control problem, guarantees hold only for the *in-sample* performance of the system, i.e. the performance against the considered scenario set, whereas the *out-of-sample* (OOS) performance, i.e. the performance against the original distribution of the uncertainty, in terms of incurred cost and actual constraint satisfaction, can be seriously compromised. For example, previous works such as [3], [18], [19], construct a scenario tree by considering only some possible realizations of the uncertainty, but they do not consider chance constraints, which would require the complete tree structure to provide guarantees for the controlled systems. Chance constraints are considered in [20] as a sample average approximation, but guarantees hold only if the number of scenarios approaches infinity. Alternatively, they are considered in [21] via randomized algorithms, where performance guarantees are provided by an appropriate sampling of the constraints. Such data-driven methods are distribution-free (see e.g., [22]) but the sample complexity can be high, which may lead to an increased computational complexity.

This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (Grant agreement No. 101018826 - ERC Advanced Grant CLariNet).

The authors are with the Delft Center for Systems and Control, Delft University of Technology, Delft, The Netherlands, {f.cordiano, b.deschutter}@tudelft.nl

In light of these considerations, in this paper we explore the use of scenario reduction algorithms in stochastic optimal control problems for chance-constrained linear systems with additive uncertainty. We consider a setting where the actual distribution of the uncertainty is known and discrete, and the number of possible realizations (i.e., the scenarios) is finite but huge, which requires a reduction algorithm to allow implementation in an online setting. Our contributions can be summarized in the following points:

- We derive a novel problem-dependent clustering-based scenario reduction algorithm, which is run offline, that, in contrast to more standard algorithms based on k -means [23], [14], or k -medians [24], minimizes a loss function that is shown to have a direct connection with the cost function employed in the optimal control problem;
- In contrast to other scenario-based methods [3], [18], [20], [21], we solve a chance-constrained scenario-based optimal control problem by considering only the scenarios resulting from the proposed reduction algorithm, which leads to computational benefits. We then robustify the optimal control problem by taking into account the effect of the discarded scenarios implicitly via constraint tightening, and by over-approximating the cost function.

The resulting optimal control problem benefits from a reduced computational complexity thanks to the usage of a reduced number of scenarios, while preserving OOS guarantees against the original distribution of the scenarios

The rest of the paper is organized as follows. In Section II, we introduce the scenario-based optimal control problem, and in Section III we discuss a problem-dependent reduction algorithm based on clustering. In Section IV we describe how to modify the constraint set and the cost function in order to preserve OOS guarantees, and in Section V we demonstrate the effectiveness of the proposed approach on a numerical example. Section VI concludes the paper.

II. PRELIMINARIES

A. Notation

The symbol $\mathbf{1}_X$ denotes the indicator function, which takes the value 1 if the logical statement X is true, and 0 otherwise. With $\|\cdot\|_l$ we denote an l -norm, with $l \in \{1, \dots, \infty\}$. The symbols $\oplus$ and $\ominus$ denote, respectively, the Minkowski sum and difference. The symbols $\Pr$ and $\mathbf{E}$ denote, respectively, the probability and expectation operators. Given n matrices $P_1, \dots, P_n$, $\text{diag}(P_1, \dots, P_n)$ is the block-diagonal matrix where the diagonal blocks are the matrices $P_1, \dots, P_n$.

B. Problem statement

We consider a discrete-time linear system affected by stochastic additive uncertainty:

$$x_{t+1} = Ax_t + Bu_t + \eta_t, \quad (1)$$

where $x_t \in \mathbb{R}^n$ and $u_t \in \mathbb{R}^m$ are, respectively, the state and input of the system at time step t , and $\eta_t \in \mathbb{R}^n$ is an

exogenous additive stochastic signal drawn from a distribution $\mathbb{P}$. The goal of the controller is to employ the system dynamics (1) to find a sequence of inputs $u_t^*, \dots, u_{t+N-1}^*$ such that the expected performance of the system over an horizon $N \in \mathbb{Z}_{>0}$

$$\mathbf{E}_{\eta_t, \dots, \eta_{t+N-1}} \left[\sum_{k=1}^N (\|Qx_{t+k}\|_1 + \|Ru_{t+k-1}\|_1) \right] \quad (2)$$

is minimized, while satisfying the following constraints on the state and on the input of the system:

$$\Pr(x_{t+k} \in \mathcal{X}, k = 1, \dots, N) \geq 1 - \varepsilon, \quad (3)$$

$$u_{t+k} \in \mathcal{U}, k = 0, \dots, N-1, \quad (4)$$

where $Q \in \mathbb{R}^{n \times n}$, $R \in \mathbb{R}^{m \times m}$ are matrices with non-negative entries, $\varepsilon \in [0, 1)$ is a risk-tolerance parameter, and $\mathcal{X}, \mathcal{U}$ are polytopic convex sets, i.e.: $\mathcal{X} := \{x \in \mathbb{R}^n : H_x x \leq h_x\}$, and $\mathcal{U} := \{u \in \mathbb{R}^m : H_u u \leq h_u\}$, where H_x, H_u, h_x, h_u are matrices and vectors of appropriate dimensions. For simplicity, in the rest of the paper we assume that Q and R are identity matrices of appropriate dimension, as this does not significantly impact the proposed analysis.

To derive tractable reformulations for (2), (3), a popular choice in the literature is to represent the uncertainty by means of scenarios [4], [6], [10]. In general, the required number of scenarios to obtain a reliable approximation can be large, or even infinite in the case of a continuous distribution [10]. In this paper, for simplicity, we directly consider the case in which $\mathbb{P}$ is discrete, but the number of the possible realization (i.e., the possible scenarios) can be very large. Let the following assumption hold:

Assumption 1: The stochastic signal $\eta_t, \forall t \geq 0$, is drawn from a discrete distribution $\mathbb{P}$, such that $\Pr(\eta_t = \eta_t^{(j)}, \forall t \geq 0) = p^{(j)}, j = 1, \dots, M$, where $M \in \mathbb{Z}_{>0}$ represents the number of possible realizations (or scenarios), and the probabilities $p^{(j)}$ satisfy: $\sum_{j=1}^M p^{(j)} = 1, p^{(j)} \in (0, 1], j = 1, \dots, M$. Both the scenarios $\eta^{(j)}$ and the probabilities $p^{(j)}$ are assumed to be known.

Under Assumption 1, and by resetting, for ease of notation, the closed-loop index t to 0, a stochastic optimal control problem can be formulated as [10]

$$\min_{u_0, \dots, u_{N-1}} \sum_{j=1}^M p^{(j)} \sum_{k=1}^N (\|x_k^{(j)}\|_1 + \|u_{k-1}\|_1) \quad (5a)$$

$$\text{s.t. } x_0^{(j)} = x_0, j = 1, \dots, M, \quad (5b)$$

$$x_{k+1}^{(j)} = Ax_k^{(j)} + Bu_k + \eta_k^{(j)}, \quad (5c)$$

$$\sum_{j=1}^M p^{(j)} \mathbf{1}_{x_k^{(j)} \in \mathcal{X}}, k=1, \dots, N \geq 1 - \varepsilon, \quad (5d)$$

$$u_k \in \mathcal{U}, \quad (5e)$$

$$k = 0, \dots, N-1.$$

In particular, (5a) is the expectation (2) computed over the discrete scenario set, (5c) describes the evolution (1) according to the j -th scenario, and (5d) is an equivalent reformulation of (3) by means of the indicator function [10].

Remark 1: According to Assumption 1, a scenario is conceived as a (multidimensional) time series with an associated probability, as proposed, e.g., in [6], [8]. Hence, the scenario trajectory $x_k^{(j)}, k = 0, \dots, N$, is determined exclusively by scenario $\eta_k^{(j)}, k = 0, \dots, N-1$. This allows to consider time-correlated phenomena (e.g., the solar irradiance in heating control systems [6]), and it can also model the more traditional scenario tree [3], by constructing a profile $\eta^{(j)}$ for each possible sequence of uncertainty realizations.

Remark 2: The usage of the 1-norm, even if less common than the squared 2-norm, is frequently used in optimal control, e.g., for piecewise affine systems, where the corresponding optimization problem can be cast as a mixed-integer linear program (MILP) [25]. Notice that also (5) can be cast as an MILP, as the chance constraint (5d) admits an equivalent reformulation by associating one logical variable to each scenario, whose value equals the indicator function. In addition, the usage of the 1-norm will turn to be useful in the performance analysis in Section IV

In general, one of the main shortcomings in scenario-based formulations of stochastic programs is that the computational complexity of (5) grows rapidly with the number of scenarios, especially due to the non-smooth chance constraint (5d). An effective means to reduce the computational burden of (5) is provided by scenario reduction algorithms [14], where from an initial discrete distribution $\mathbb{P}$, the goal is to find a discrete distribution $\tilde{\mathbb{P}}$ whose domain has a lower cardinality than the domain of $\mathbb{P}$, in a way that $\tilde{\mathbb{P}}$ provides a good description of $\mathbb{P}$ (e.g., in terms of the Wasserstein distance [13]). Hence, let $\tilde{\eta}_t \in \mathbb{R}^n, t \in \mathbb{Z}_{\geq 0}$ be a stochastic signal described by the distribution $\tilde{\mathbb{P}}$, such that: $\Pr(\tilde{\eta}_t = \tilde{\eta}_t^{(j)}, \forall t \geq 0) = \tilde{p}^{(j)}, \sum_{j=1}^{\tilde{M}} \tilde{p}^{(j)} = 1, \tilde{p}^{(j)} \in (0, 1], j = 1, \dots, \tilde{M}$, where $\tilde{M} \in \mathbb{Z}_{>0}, \tilde{M} < M$, represents the cardinality of the reduced set of scenarios. The way in which the scenarios $\tilde{\eta}^{(1)}, \dots, \tilde{\eta}^{(\tilde{M})}$, and the corresponding probabilities $\tilde{p}^{(1)}, \dots, \tilde{p}^{(\tilde{M})}$, are chosen clearly affects the solution of the resulting optimal control problem; hence, in the following, we provide an equivalent reformulation of (5), where the dependence of the cost and the constraint set on the distribution $\mathbb{P}$ is made explicit. Let us introduce the following compact notation for the predicted state and input over the horizon:

$$\mathbf{x}^{(j)} = \begin{bmatrix} x_1^{(j)} \\ \vdots \\ x_N^{(j)} \end{bmatrix}, \mathbf{u} = \begin{bmatrix} u_0 \\ \vdots \\ u_{N-1} \end{bmatrix}, \tilde{\eta}^{(j)} = \begin{bmatrix} \tilde{\eta}_0^{(j)} \\ \vdots \\ \tilde{\eta}_{N-1}^{(j)} \end{bmatrix}, \quad (6)$$

$j = 1, \dots, \tilde{M}$,

for which we have:

$$\mathbf{x}^{(j)} = \mathbf{F}x_0 + \mathbf{G}\mathbf{u} + \mathbf{\Gamma}\tilde{\eta}^{(j)}, \quad (7)$$

where $\mathbf{F}, \mathbf{G}, \mathbf{\Gamma}$ are appropriate matrices derived from the dynamics (1), which can be found e.g. in [21]. Let also $\mathbf{H}_x := \text{diag}(H_x, \dots, H_x)$, $\mathbf{H}_u := \text{diag}(H_u, \dots, H_u)$, $\mathbf{h}_x = [h_x^\top, \dots, h_x^\top]^\top$, $\mathbf{h}_u = [h_u^\top, \dots, h_u^\top]^\top$, and define: $\mathcal{X}_N := \{\mathbf{x} \in \mathbb{R}^{nN} : \mathbf{H}_x \mathbf{x} \leq \mathbf{h}_x\}$, $\mathcal{U}_N := \{\mathbf{u} \in \mathbb{R}^{mN} : \mathbf{H}_u \mathbf{u} \leq \mathbf{h}_u\}$.

Then, the scenario-based optimal control problem that employs the reduced scenario set can be written in function of the resulting probability distribution $\tilde{\mathbb{P}}$ as

$$\min_{\mathbf{u}} \{J(x_0, \tilde{\mathbb{P}}, \mathbf{u}) : \mathbf{u} \in \mathcal{F}(x_0, \tilde{\mathbb{P}})\}, \quad (8)$$

where

$$J(x_0, \mathbf{u}, \tilde{\mathbb{P}}) = \sum_{j=1}^{\tilde{M}} \tilde{p}^{(j)} \left(\|\mathbf{x}^{(j)}\|_1 + \|\mathbf{u}\|_1 \right) \quad (9)$$

is the cost function that we minimize, and

$$\begin{aligned} \mathcal{F}(x_0, \tilde{\mathbb{P}}) &:= \{\mathbf{u} \in \mathbb{R}^{mN} : \sum_{j=1}^{\tilde{M}} \tilde{p}^{(j)} \mathbf{1}_{\mathbf{x}^{(j)} \in \mathcal{X}_N} \geq 1 - \varepsilon, \\ &\mathbf{x}^{(j)} = \mathbf{F}x_0 + \mathbf{G}\mathbf{u} + \mathbf{\Gamma}\tilde{\eta}^{(j)}, \quad (10) \\ &\tilde{\eta}^{(j)} \sim \tilde{\mathbb{P}}, j = 1, \dots, \tilde{M}, \\ &\mathbf{u} \in \mathcal{U}_N\} \end{aligned}$$

denotes the set of feasible inputs. We also observe that the way $\tilde{\mathbb{P}}$ is constructed affects both the cost function (9) and the constraint set (10) of (8), which coincide with (5) for $\tilde{\mathbb{P}} = \mathbb{P}$.

III. SCENARIO REDUCTION

The majority of the reduction algorithms select a set of representative scenarios in a way that some suitable similarity metric between the initial and the resulting distributions is optimized. This can be, e.g., the Wasserstein distance [16], [11], or a discrepancy distance [12], between the probability distributions $\mathbb{P}$ and $\tilde{\mathbb{P}}$. Such reduction algorithms are typically of combinatorial complexity [26]; hence, approximate methods to perform a reduction have been proposed. A popular family of reduction algorithms is based indeed on grouping similar scenarios in a clustering fashion [14], i.e., given the initial set of scenarios $\mathcal{S} := \{\eta^{(1)}, \dots, \eta^{(M)}\}$ and the desired cardinality $\tilde{M}$ for the reduced scenario set, a clustering-based reduction algorithm groups similar scenarios in order to minimize an appropriate cost function, and one representative scenario is chosen for each cluster with probability equal to the sum of the probabilities of the scenarios belonging to the same cluster. In the following, we propose a novel problem-dependent clustering-based reduction algorithm that also includes the more standard k -means [23] and k -medians [24]. By denoting the reduced scenario set as $\tilde{\mathcal{S}} := \{\tilde{\eta}^{(1)}, \dots, \tilde{\eta}^{(\tilde{M})}\}$, the proposed clustering-based scenario reduction algorithm minimizes the following objective function:

$$\mathcal{L}(\tilde{\mathcal{S}}) = \mathbf{E}_{\eta \sim \mathbb{P}} \left[\min_{\tilde{\eta} \in \tilde{\mathcal{S}}} \|\eta - \tilde{\eta}\|_l^l \right] = \sum_{h=1}^M p^{(h)} \min_{\tilde{\eta} \in \tilde{\mathcal{S}}} \|\eta^{(h)} - \tilde{\eta}\|_l^l,$$

where we consider $l \in \{1, 2\}$. Let us now define a cluster as the set of indices corresponding to the closest scenarios to a given center $\tilde{\eta}^{(j)}$, i.e.

$$\mathcal{C}^{(j)} := \{h \in \{1, \dots, M\} : \arg \min_{i \in \{1, \dots, M\}} \|\eta^{(i)} - \tilde{\eta}^{(j)}\|_l = j\}.$$

Hence, the objective function $\mathcal{L}$ can be equivalently written as

$$\mathcal{L}(\tilde{\mathcal{S}}) = \sum_{j=1}^{\tilde{M}} \sum_{h \in \mathcal{C}^{(j)}} p^{(h)} \|\boldsymbol{\eta}^{(h)} - \tilde{\boldsymbol{\eta}}^{(j)}\|_l^l, \quad (11)$$

and the scenario reduction problem becomes

$$\min_{\tilde{\mathcal{S}}} \mathcal{L}(\tilde{\mathcal{S}}). \quad (12)$$

The rationale behind the cost function (11) is that we intend to minimize the expected distance between the scenarios belonging to a given cluster $\mathcal{C}^{(j)}$ and the corresponding center $\tilde{\boldsymbol{\eta}}^{(j)}$, $j = 1, \dots, \tilde{M}$. The reason why in (11) we explicitly consider the expectation over the distribution $\mathbb{P}$ will be clear in the performance analysis in the next section, where we show that there exists a connection between the loss (11) and the impact of the scenario reduction on the system performance. It is well known that (12) is NP-hard [27]; hence, greedy algorithms are frequently used to provide suboptimal solutions to (12). Inspired from [27], we propose the following iterative algorithm:

Algorithm 1 Clustering-based scenario reduction

Require: $\mathcal{S}, p^{(1)}, \dots, p^{(M)}, \tilde{M}$;
Initialize $\tilde{\mathcal{S}} \subseteq \mathcal{S}$ arbitrarily;
while $\mathcal{L}(\tilde{\mathcal{S}})$ decreases **do**
 for $j \in \{1, \dots, \tilde{M}\}$ **do**
 1) $\mathcal{C}^{(j)} \leftarrow \{h \in \{1, \dots, M\} :$
 $\tilde{\boldsymbol{\eta}}^{(j)} = \arg \min_{\tilde{\boldsymbol{\eta}}^{(i)} \in \tilde{\mathcal{S}}} \|\boldsymbol{\eta}^{(h)} - \tilde{\boldsymbol{\eta}}^{(i)}\|_l^l\};$
 2) $\tilde{\boldsymbol{\eta}}^{(j)} \leftarrow \arg \min_{\boldsymbol{\eta} \in \mathbb{R}^{N_n}} \sum_{h \in \mathcal{C}^{(j)}} p^{(h)} \|\boldsymbol{\eta}^{(h)} - \boldsymbol{\eta}\|_l^l;$
 end for
 Assign $\tilde{\mathcal{S}} \leftarrow \{\tilde{\boldsymbol{\eta}}^{(1)}, \dots, \tilde{\boldsymbol{\eta}}^{(\tilde{M})}\};$
end while
Assign $\tilde{p}^{(j)} = \sum_{h \in \mathcal{C}^{(j)}} p^{(h)}, j = 1, \dots, \tilde{M}$;
return $\tilde{\mathcal{S}}, \tilde{p}^{(1)}, \dots, \tilde{p}^{(\tilde{M})}$.

Let us state the following lemma:

Lemma 1: Algorithm 1 converges to a (suboptimal) solution of (12) in a finite number of iterations. In particular, if $p^{(h)} = \frac{1}{M}, h = 1, \dots, M$, the update for $\tilde{\boldsymbol{\eta}}^{(j)}$ in step 2) in Algorithm 1 corresponds to:

- a) $\tilde{\boldsymbol{\eta}}^{(j)} = \sum_{h \in \mathcal{C}^{(j)}} p^{(h)} \boldsymbol{\eta}^{(h)} / |\mathcal{C}^{(j)}|$ if $l = 2$ (i.e., k -means clustering);
- b) $\tilde{\boldsymbol{\eta}}^{(j)}$ such that $\tilde{\boldsymbol{\eta}}_i^{(j)}$ is the median of the i -th entries of each $\boldsymbol{\eta}^{(h)}, h \in \mathcal{C}^{(j)}, i = 1, \dots, nN$, if $l = 1$ (i.e., k -medians clustering).

Proof: The convergence immediately follows by observing that the updates in 1) and 2) lead to a decrease in the cost (11); hence, $\mathcal{L}$ is monotonically decreasing in the iterations of Algorithm 1. Then, a) and b) follow by observing that with $p^{(h)} = \frac{1}{M}$ Algorithm (1) corresponds to k -means (if $l = 2$) and k -medians (if $l = 1$) clustering [27]. ■

Remark 3: Algorithm 1 performs, at each iteration, a greedy selection of the clusters and of the representative scenario for each cluster. This is similar to Lloyd's algorithm [28], [29], which, in contrast to the combinatorial complexity of the original problem (12), has a polynomial complexity $\mathcal{O}(M \cdot \tilde{M})$ (see e.g. [27]). However, it may return a suboptimal solution.

Remark 4: In the case where the probabilities are not all equal, part a) of Lemma 1 corresponds to the weighted average of the scenarios belonging to a given cluster $\mathcal{C}^{(j)}$, where the weights are $p^{(h)}, h \in \mathcal{C}^{(j)}$. Part b) corresponds to the weighted median of the scenarios belonging to a given cluster $\mathcal{C}^{(j)}$, where the weighted median is computed element-wise. For the definition and the computation of the weighted median we refer to [30].

Algorithm 1 can be used in an offline phase to derive a set of scenarios with reduced cardinality, which is then employed as description of the uncertainty for the online optimal control problem. However, using a reduced number of scenarios has an impact on the performance of the system, as shown in the next section.

IV. ENFORCING OUT-OF-SAMPLE GUARANTEES

The distribution $\tilde{\mathbb{P}}$ can indeed be constructed from Algorithm 1, by associating to each scenario $\tilde{\boldsymbol{\eta}}^{(j)}$ the probability $\tilde{p}^{(j)}, j = 1, \dots, \tilde{M}$. The resulting distribution $\tilde{\mathbb{P}}$ will then affect the optimal solution of problem (8), and, moreover, considering a reduced number of scenarios in (8) might compromise the OOS guarantees of the controlled system, considering that it evolves according to the original distribution $\mathbb{P}$. Hence, in this section, we focus on the effect that a scenario reduction can have on the OOS guarantees of the real system, in terms of actual chance constraint satisfaction and performance.

A. Adaption of the constraint set

Since the actual system dynamics evolve according to the distribution $\mathbb{P}$, it might happen that the optimal input obtained by solving (8) (i.e., the problem that employs the reduced scenario set), leads to an actual constraint violation larger than the theoretical value ε . Hence, in this section, we propose a modification to the optimal control problem (8) such that it will guarantee feasibility for the original problem (5).

First, let $\mathcal{E}^{(j)}$ be a set such that $\Gamma(\boldsymbol{\eta}^{(h)} - \tilde{\boldsymbol{\eta}}^{(j)}) \in \mathcal{E}^{(j)}, \forall h \in \mathcal{C}^{(j)}$. That is, for each cluster $\mathcal{C}^{(j)}, j = 1, \dots, \tilde{M}$, the set $\mathcal{E}^{(j)}$ over-approximates the distance between each scenario $\boldsymbol{\eta}^{(h)}, h \in \mathcal{C}^{(j)}$, and $\tilde{\boldsymbol{\eta}}^{(j)}$, i.e., the representative scenario of the cluster. Such set can be easily computed from the knowledge of the initial scenario set and the reduced scenario set (e.g., from element-wise upper and lower bounds on $\eta_k^{(h)} - \tilde{\eta}_k^{(j)}, k = 0, \dots, N - 1$).

The set $\mathcal{E}^{(j)}$ can be employed for a tightening of the original state constraint set, in order to enforce the sought feasibility guarantees. This is formalized in the following theorem:

Theorem 1: Let $\mathcal{X}_N \ominus \mathcal{E}^{(j)}$ be non empty, $\forall j \in \{1, \dots, \tilde{M}\}$, and let us consider the following tightened constraint set:

$$\begin{aligned} \tilde{\mathcal{F}}(x_0, \tilde{\mathbb{P}}) &:= \{\mathbf{u} \in \mathbb{R}^{mN} : \sum_{j=1}^{\tilde{M}} \tilde{p}^{(j)} \mathbf{1}_{\mathbf{x}^{(j)} \in \mathcal{X}_N \ominus \mathcal{E}^{(j)}} \geq 1 - \varepsilon, \\ &\quad \mathbf{x}^{(j)} = \mathbf{F}x_0 + \mathbf{G}\mathbf{u} + \mathbf{\Gamma}\tilde{\boldsymbol{\eta}}^{(j)}, \\ &\quad \tilde{\boldsymbol{\eta}}^{(j)} \sim \tilde{\mathbb{P}}, j = 1, \dots, \tilde{M}, \\ &\quad \mathbf{u} \in \mathcal{U}_N\}. \end{aligned} \quad (13)$$

Then it holds that $\tilde{\mathcal{F}}(x_0, \tilde{\mathbb{P}}) \subseteq \mathcal{F}(x_0, \mathbb{P})$, $\forall x_0 \in \mathcal{X}$.

Proof: Let $\tilde{\boldsymbol{\eta}}^{(j)}$, $j \in \{1, \dots, \tilde{M}\}$, be one of the scenarios belonging to the reduced set, and let $\boldsymbol{\eta}^{(h)}$ be a scenario belonging to the same cluster, i.e., $h \in \mathcal{C}^{(j)}$. Let $\tilde{\mathbf{x}}^{(j)}$ be the scenario trajectory corresponding to scenario $\tilde{\boldsymbol{\eta}}^{(j)}$, and $\mathbf{x}^{(h)}$ the scenario trajectory corresponding to scenario $\boldsymbol{\eta}^{(h)}$. We now show that if $\tilde{\mathbf{x}}^{(j)} \in \mathcal{X}_N \ominus \mathcal{E}^{(j)}$ then $\mathbf{x}^{(h)} \in \mathcal{X}_N$. Indeed, for any input $\mathbf{u} \in \tilde{\mathcal{F}}(x_0, \tilde{\mathbb{P}})$, we have:

$$\begin{aligned} \mathbf{x}^{(h)} &= \mathbf{F}x_0 + \mathbf{G}\mathbf{u} + \mathbf{\Gamma}\boldsymbol{\eta}^{(h)} \\ &= \mathbf{F}x_0 + \mathbf{G}\mathbf{u} + \mathbf{\Gamma}\tilde{\boldsymbol{\eta}}^{(j)} + \mathbf{\Gamma}(\boldsymbol{\eta}^{(h)} - \tilde{\boldsymbol{\eta}}^{(j)}) \\ &\in \mathcal{X}_N \ominus \mathcal{E}^{(j)} \oplus \mathcal{E}^{(j)} \\ &\in \mathcal{X}_N, \end{aligned} \quad (14)$$

where we exploit the known property: $A \ominus B \oplus B \subseteq A$. Let us now prove that the relations in (14) are satisfied with the required probability $1 - \varepsilon$. The fact that $\mathbf{x}_t^{(j)} \in \mathcal{X}_N \ominus \mathcal{E}^{(j)}$ implies $\mathbf{x}^{(h)} \in \mathcal{X}_N$, translates in $\mathbf{1}_{\mathbf{x}^{(j)} \in (\mathcal{X}_N \ominus \mathcal{E}_t^{(j)})} = 1 \Rightarrow \mathbf{1}_{\mathbf{x}^{(h)} \in \mathcal{X}} = 1, \forall h \in \mathcal{C}^{(j)}$, or equivalently: $\mathbf{1}_{\mathbf{x}^{(h)} \in \mathcal{X}} \geq \mathbf{1}_{\mathbf{x}^{(j)} \in (\mathcal{X}_N \ominus \mathcal{E}_t^{(j)})}, \forall h \in \mathcal{C}^{(j)}$. The state constraint for the scenario trajectory $\mathbf{x}_t^{(j)}$ is satisfied with probability $\tilde{p}^{(j)} = \sum_{h \in \mathcal{C}^{(j)}} p^{(h)}$; hence, if $\mathbf{u} \in \tilde{\mathcal{F}}(x_0, \tilde{\mathbb{P}})$, it must hold that

$$\begin{aligned} 1 - \varepsilon &\leq \sum_{j=1}^{\tilde{M}} \tilde{p}^{(j)} \mathbf{1}_{\mathbf{x}^{(j)} \in (\mathcal{X} \ominus \mathcal{E}_t^{(j)})} \\ &\leq \sum_{j=1}^{\tilde{M}} \sum_{h \in \mathcal{C}^{(j)}} p^{(h)} \mathbf{1}_{\mathbf{x}^{(h)} \in \mathcal{X}} \\ &\leq \sum_{j=1}^{\tilde{M}} p^{(j)} \mathbf{1}_{\mathbf{x}^{(j)} \in \mathcal{X}}, \end{aligned}$$

and this proves that $\mathbf{u}$ is feasible for the original set of scenarios, i.e.: $\mathbf{u} \in \mathcal{F}(x_0, \mathbb{P})$. ■

The meaning of Theorem 1 is that the optimal input sequence obtained by solving the scenario-based optimal control problem (8) with the tightened constraint set described by (13) is indeed feasible in terms of chance constraint satisfaction for the actual system, which evolves according to $\mathbb{P}$.

B. Adaptation of the cost function

Even if Theorem 1 ensures $\tilde{\mathcal{F}}(x_0, \tilde{\mathbb{P}}) \subseteq \mathcal{F}(x_0, \mathbb{P})$, it might still happen that the solution of (8) leads to an over-optimistic expected performance of the controlled system, due to the

usage of a limited number of scenarios in the cost function. Hence, by defining

$$J^*(x_0, \mathbb{P}) := \min_{\mathbf{u}} \{J(x_0, \mathbb{P}, \mathbf{u}) : \mathbf{u} \in \mathcal{F}(x_0, \mathbb{P})\}, \quad (15)$$

$$\tilde{J}^*(x_0, \tilde{\mathbb{P}}) := \min_{\mathbf{u}} \{J(x_0, \tilde{\mathbb{P}}, \mathbf{u}) : \mathbf{u} \in \tilde{\mathcal{F}}(x_0, \tilde{\mathbb{P}})\}, \quad (16)$$

we now aim at finding a correction term $c(x_0, \mathbf{u}, \mathbb{P}, \tilde{\mathbb{P}}) \geq 0$ such that

$$\tilde{J}^*(x_0, \tilde{\mathbb{P}}) + c(x_0, \mathbf{u}, \mathbb{P}, \tilde{\mathbb{P}}) \geq J^*(x_0, \mathbb{P}), \quad \forall x_0 \in \mathcal{X}.$$

Theorem 2: Let $\bar{c} \in \mathbb{R}_{>0}$ be a constant such that

$$\bar{c} \geq \sum_{j=1}^{\tilde{M}} \sum_{h \in \mathcal{C}^{(j)}} p^{(h)} \|\mathbf{\Gamma}(\boldsymbol{\eta}^{(h)} - \tilde{\boldsymbol{\eta}}^{(j)})\|_1.$$

Then it holds that

$$\tilde{J}^*(x_0, \tilde{\mathbb{P}}) + \bar{c} \geq J^*(x_0, \mathbb{P}), \quad \forall x_0 \in \mathcal{X}. \quad (17)$$

Proof: Let $\mathbf{u}^*$ and $\tilde{\mathbf{u}}^*$ be the minimizers associated, respectively, to (15) and (16). Then it follows that

$$\begin{aligned} J^*(x_0, \mathbb{P}) - \tilde{J}^*(x_0, \tilde{\mathbb{P}}) &= J(x_0, \mathbf{u}^*, \mathbb{P}) - J(x_0, \tilde{\mathbf{u}}^*, \tilde{\mathbb{P}}) \quad (18) \\ &\leq J(x_0, \tilde{\mathbf{u}}^*, \mathbb{P}) - J(x_0, \tilde{\mathbf{u}}^*, \tilde{\mathbb{P}}) \quad (19) \\ &\leq |J(x_0, \tilde{\mathbf{u}}^*, \mathbb{P}) - J(x_0, \tilde{\mathbf{u}}^*, \tilde{\mathbb{P}})|, \quad (20) \end{aligned}$$

where in (19) we exploit that the input $\tilde{\mathbf{u}}^*$ is feasible for problem (5) thanks to Theorem 1. Then, exploiting the definition of the cost function (9), we have

$$|J(x_0, \tilde{\mathbf{u}}^*, \mathbb{P}) - J(x_0, \tilde{\mathbf{u}}^*, \tilde{\mathbb{P}})| \quad (21)$$

$$= \left| \sum_{j=1}^{\tilde{M}} p^{(j)} \|\mathbf{x}^{(j)}\|_1 - \sum_{j=1}^{\tilde{M}} \tilde{p}^{(j)} \|\tilde{\mathbf{x}}^{(j)}\|_1 \right| \quad (22)$$

$$= \left| \sum_{j=1}^{\tilde{M}} \sum_{h \in \mathcal{C}^{(j)}} p^{(h)} (\|\mathbf{x}^{(h)}\|_1 - \|\tilde{\mathbf{x}}^{(j)}\|_1) \right| \quad (23)$$

$$\leq \sum_{j=1}^{\tilde{M}} \sum_{h \in \mathcal{C}^{(j)}} p^{(h)} \|\mathbf{x}^{(h)} - \tilde{\mathbf{x}}^{(j)}\|_1 \quad (24)$$

$$\leq \sum_{j=1}^{\tilde{M}} \sum_{h \in \mathcal{C}^{(j)}} p^{(h)} \left\| \mathbf{\Gamma}(\boldsymbol{\eta}^{(h)} - \tilde{\boldsymbol{\eta}}^{(j)}) \right\|_1, \quad (25)$$

where in (22) we leverage that the applied input is the same, in (23) we exploit the definition of $\tilde{p}^{(j)}$, $j = 1, \dots, \tilde{M}$, given by Algorithm 1, in (24) we use the triangle inequality, and in (25) we use the dynamics (7). Then (17) follows by selecting a constant $\bar{c}$ greater or equal than the quantity in (25). ■

Notice in particular that the quantity $\bar{c}$ does not depend on the current state x_0 of the system. The essential meaning of Theorem 2 is that it provides an upper bound on the original cost function $J^*(x_0, \mathbb{P})$, for which we prove that this upper bound can be obtained simply by adding a constant to the optimal value of the reduced problem. Similarly as in the constraint tightening (13), the constant $\bar{c}$, can easily be computed from the knowledge of the matrix $\mathbf{\Gamma}$, and by measuring the distance between each scenario and the

representative scenario of the same cluster to which the given scenario belongs.

Remark 5: From the above proof, it follows that

$$J^*(x_0, \mathbb{P}) - \tilde{J}^*(x_0, \tilde{\mathbb{P}}) \quad (26)$$

$$\leq \|\mathbf{\Gamma}\|_1 \sum_{j=1}^{\tilde{M}} \sum_{h \in \mathcal{C}^{(j)}} p^{(h)} \left\| \boldsymbol{\eta}^{(h)} - \tilde{\boldsymbol{\eta}}^{(j)} \right\|_1 \quad (27)$$

$$\leq \|\mathbf{\Gamma}\|_1 \mathcal{L}(\tilde{\mathcal{S}}). \quad (28)$$

Hence, the choice of the objective function in the reduction problem (12) relates to the distance between the optimal values (26) of the optimal control problem (8). In other words, the clustering algorithm in (12) has a problem-dependent nature, i.e., by minimizing the cost function (11) we also reduce the difference in (26).

By putting Theorems 1 and 2 together, the following optimization problem:

$$\min_{\mathbf{u}} \{J(x_0, \tilde{\mathbb{P}}, \mathbf{u}) + \bar{c} : \mathbf{u} \in \tilde{\mathcal{F}}(x_0, \tilde{\mathbb{P}})\} \quad (29)$$

over-approximates the original one (5). The main advantage in solving (29) is that a reduced scenario set is used, which provides a consistent computational advantage. Nevertheless, the effect of the discarded scenarios is implicitly described in (29) by the constraint tightening and by the over-approximation of the cost function. Thanks to the constraint tightening, any input that is feasible for (29) is also feasible for the original problem (5), and the optimal cost of (29) provides an upper bound for the cost of (5).

V. NUMERICAL EXAMPLE

Let us consider a discrete time dynamical system (1) with

$$A = \begin{bmatrix} 1 & 1 \\ 0 & 0.5 \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad (30)$$

and constraints: $\Pr(x_k \geq -1, k = 1, \dots, N-1) \geq 0.2$, $\|u_k\|_\infty \leq 2$, with initial state $x_0 = [4 \ 3]^\top$. We consider a prediction horizon of $N = 10$, and an initial scenario set of cardinality $M = 200$. Each scenario is a time series, parametrized by a set of parameters that are sampled from a user-defined distribution $\mathbb{P}$, in order to generate different profiles $\boldsymbol{\eta}^{(j)}, j = 1, \dots, M$. For details, the code of this numerical example is available as a Github repository [31].

For simplicity, we assume that all the scenarios are equally likely, i.e.: $p^{(j)} = \frac{1}{M}, j = 1, \dots, M$. Hence, Lemma 1 suggests that the k -medians algorithm is particularly suited, since the chosen norm is the 1-norm. In the following, we provide a comparison between the solution of (8) (i.e., the problem solved considering the reduced scenario set, denoted by P1) and (29) (i.e., the problem that, in addition, considers the sufficient conditions for OOS guarantees, denoted by P2) with (5) (i.e., the original problem that considers the original set with $M = 200$ scenarios, denoted by EXACT). For solving P1 and P2, we run both k -medians (denoted by kMED) and k -means (denoted by kMNS) with $\tilde{M} \in \{5, 25, 50, 75, 100, 125, 150, 175\}$ scenarios. We cast the resulting optimal control problem as an MILP, by associating

a binary variable to each scenario, and by reformulating the chance constraint (5d) accordingly [10]. For all the simulations, we use Gurobi 10.0.2 as a solver on a CPU Intel i7-1185G7 @ 3.00GHz.

In Figure 1, we compare the OOS performance of the system for both P1 and P2. In contrast to P1, it is possible to notice that P2 offers OOS guarantees for the real system. In particular, thanks to the constraint tightening, Figure 1 shows that the actual constraint violation is always lower than the theoretical value ε . In addition, in Figure 1 we note that the optimal cost of P2 provides a more reliable estimate of the cost associated to the initial set of scenarios. In synthesis, P2 provides OOS guarantees thanks to the tightening and the over-approximation of the cost proposed in Section IV, at the cost of a mild conservatism that is gradually reduced by increasing $\tilde{M}$. We observe that both P1 and P2 converge to the same value of actual constraint violation and optimal cost as the number of scenarios is increased. Despite k -medians should be the preferred choice due to the usage of the 1-norm, we note that both k -medians and k -means lead to substantially equivalent results. Finally, in the last plot of Figure 1 we observe that the solver time required to solve P1 and P2 grows exponentially with the number of scenarios due to the mixed-integer reformulation of the chance constraint (5d) (note the logarithmic scale on the y axis in Figure 1). The solver time is comparable for P1 and P2, and P2 offers OOS guarantees even with a small number of scenarios and at a low computational cost with respect to P1.

VI. CONCLUSIONS

In this paper, we have shown that scenario reduction algorithms lead to a significant computational advantage in stochastic optimal control problems, while, however, compromising the OOS performance of the system. To tackle this issue we have proposed a scenario-reduction algorithm tailored for optimal control problems, and sufficient conditions to implicitly keep into account the discarded scenarios in the optimal control problem at a low computational cost. The numerical example demonstrates that the proposed method leads to a consistent computational advantage, since few scenarios can be used without compromising the OOS guarantees, at the cost of a moderate conservatism.

Ongoing work consists of extending the method to the case where the original distribution $\mathbb{P}$ is not known, but described by a finite number of samples, and to provide OOS guarantees with a certain confidence level. Other possible future work consists of extending these results to nonlinear systems and more general sources of uncertainty (e.g., as in the dual adaptive control framework [32]), as well as providing performance bounds for quadratic cost functions.

ACKNOWLEDGMENT

We wish to thank Matin Jafarian for the fruitful discussions and careful reading.

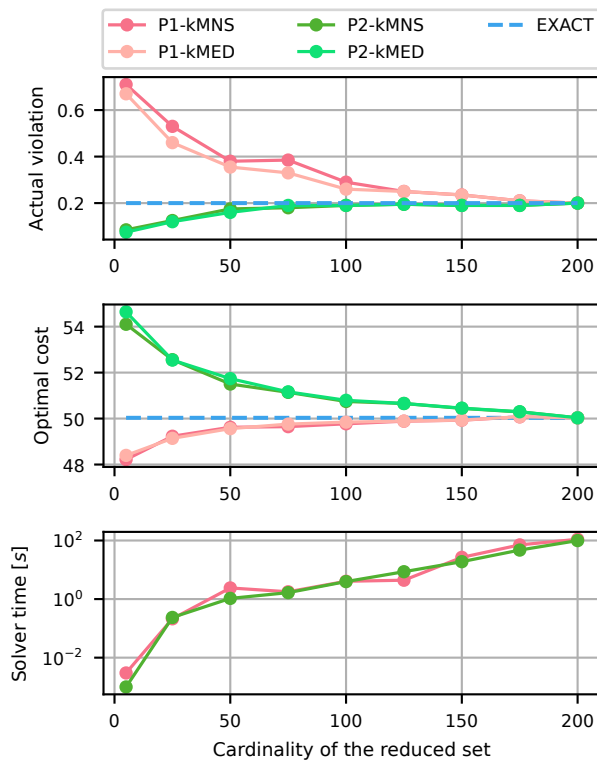


Fig. 1. Performance analysis for the real system. In particular, we notice that P2 over-approximates the original problem (5) both in terms of constraint satisfaction and performance, and the associated computational cost is comparable to P1.

REFERENCES

- [1] A. Mesbah, "Stochastic model predictive control: An overview and perspectives for future research," *IEEE Control Systems Magazine*, vol. 36, no. 6, pp. 30–44, 2016.
- [2] A. Ruszczyński and A. Shapiro, "Stochastic programming models," *Handbooks in Operations Research and Management Science*, vol. 10, pp. 1–64, 2003.
- [3] D. Bernardini and A. Bemporad, "Stabilizing model predictive control of stochastic constrained linear systems," *IEEE Transactions on Automatic Control*, vol. 57, no. 6, pp. 1468–1480, 2012.
- [4] A. Bemporad, T. Gabriellini, L. Puglia, and L. Bellucci, "Scenario-based stochastic model predictive control for dynamic option hedging," in *49th IEEE Conference on Decision and Control (CDC)*, 2010, pp. 6089–6094.
- [5] K. C. Sharma, R. Bhakar, H. P. Tiwari, and S. Chawda, "Scenario based uncertainty modeling of electricity market prices," in *2017 6th International Conference on Computer Applications In Electrical Engineering-Recent Advances (CERA)*, 2017, pp. 164–168.
- [6] T. Pippia, J. Lago, R. De Coninck, and B. De Schutter, "Scenario-based nonlinear model predictive control for building heating systems," *Energy and Buildings*, vol. 247, p. 111108, 2021.
- [7] Z. Su, A. Jamshidi, A. Núñez, S. Baldi, and B. De Schutter, "Multi-level condition-based maintenance planning for railway infrastructures – a scenario-based chance-constrained approach," *Transportation Research Part C: Emerging Technologies*, vol. 84, pp. 92–123, 2017.
- [8] J. Morales, S. Pineda, A. Conejo, and M. Carrion, "Scenario reduction for futures market trading in electricity markets," *IEEE Transactions on Power Systems*, vol. 24, no. 2, pp. 878–888, 2009.
- [9] G. C. Pflug, "Scenario tree generation for multiperiod financial optimization by optimal discretization," *Mathematical Programming*, vol. 89, no. 2, pp. 251–271, 2001.
- [10] B. K. Pagnoncelli, S. Ahmed, and A. Shapiro, "Sample average approximation method for chance constrained programming: Theory and applications," *Journal of Optimization Theory and Applications*, vol. 142, no. 2, pp. 399–416, 2009.
- [11] J. Dupačová, N. Gröwe-Kuska, and W. Römisch, "Scenario reduction in stochastic programming," *Mathematical Programming*, vol. 95, no. 3, pp. 493–511, 2003.
- [12] R. Henrion, C. Kuchler, and W. Römisch, "Discrepancy distances and scenario reduction in two-stage stochastic integer programming," *Journal of Industrial and Management Optimization*, pp. 363–384, 2008.
- [13] W. Römisch, "Stability of stochastic programming problems," in *Handbooks in Operations Research and Management Science*. Elsevier, 2003, vol. 10, pp. 483–554.
- [14] N. Rujeerapaiboon, K. Schindler, D. Kuhn, and W. Wiesemann, "Scenario reduction revisited: Fundamental limits and guarantees," *Mathematical Programming*, vol. 191, no. 1, pp. 207–242, 2022.
- [15] Y. Wang, Y. Liu, and D. S. Kirschen, "Scenario reduction with sub-modular optimization," *IEEE Transactions on Power Systems*, vol. 32, no. 3, pp. 2479–2480, 2017.
- [16] W. Römisch and R. J.-B. Wets, "Stability of ϵ -approximate solutions to convex stochastic programs," *SIAM Journal on Optimization*, vol. 18, no. 3, pp. 961–979, 2007.
- [17] R. T. Rockafellar and R. J.-B. Wets, *Variational Analysis*, 3rd ed. Springer Science & Business Media, 2009, vol. 317.
- [18] S. Lucia, T. Finkler, D. Basak, and S. Engell, "A new robust NMPC scheme and its application to a semi-batch reactor example," *IFAC Proceedings Volumes*, vol. 45, no. 15, pp. 69–74, 2012.
- [19] F. Holtorf, A. Mitsos, and L. T. Biegler, "Multistage NMPC with online generated scenario trees: Application to a semi-batch polymerization process," *Journal of Process Control*, vol. 80, pp. 167–179, 2019.
- [20] L. Blackmore, M. Ono, A. Bektassov, and B. C. Williams, "A probabilistic particle-control approximation of chance-constrained stochastic predictive control," *IEEE Transactions on Robotics*, vol. 26, no. 3, pp. 502–517, 2010.
- [21] M. Prandini, S. Garatti, and J. Lygeros, "A randomized approach to stochastic model predictive control," in *2012 IEEE 51st IEEE Conference on Decision and Control (CDC)*, 2012, pp. 7315–7320.
- [22] M. C. Campi and S. Garatti, "A sampling-and-discarding approach to chance-constrained optimization: Feasibility and optimality," *Journal of Optimization Theory and Applications*, vol. 148, no. 2, pp. 257–280, 2011.
- [23] J. A. Hartigan and M. A. Wong, "Algorithm as 136: A k -means clustering algorithm," *Applied Statistics*, vol. 28, no. 1, pp. 100–108, 1979.
- [24] C. Whelan, G. Harrell, and J. Wang, "Understanding the k -medians problem," *Proceedings of the International Conference on Scientific Computing (CSC)*, p. 219, 2015.
- [25] A. Richards and J. How, "Mixed-integer programming for control," in *Proceedings of the 2005, American Control Conference, 2005*. IEEE, 2005, pp. 2676–2683.
- [26] H. Heitsch and W. Römisch, "Scenario reduction algorithms in stochastic programming," *Computational Optimization and Applications*, pp. 187–206, 2003.
- [27] S. Dasgupta, "Topics in unsupervised learning," 2008. [Online]. Available: <https://cseweb.ucsd.edu/~dasgupta/291-unsup/>
- [28] S. Lloyd, "Least squares quantization in PCM," *IEEE Transactions on Information Theory*, vol. 28, no. 2, pp. 129–137, 1982.
- [29] Q. Du, V. Faber, and M. Gunzburger, "Centroidal voronoi tessellations: Applications and algorithms," *SIAM Review*, vol. 41, no. 4, pp. 637–676, 1999.
- [30] C. Gurwitz, "Weighted median algorithms for l_1 approximation," *BIT*, vol. 30, no. 2, pp. 301–310, 1990.
- [31] F. Cordiano, "Project code," 2024. [Online]. Available: <https://github.com/fracordi/scen-linear-mpc>
- [32] T. A. N. Heirung, B. E. Ydstie, and B. Foss, "Dual adaptive model predictive control," *Automatica*, vol. 80, pp. 340–348, 2017.