This is the author's accepted manuscript of a paper published in the Proceedings of the
64rd IEEE Conference on Decision and Control (CDC 2025).
The final version of record is available at: https://doi.org/10.1109/CDC57313.2025.11313027

A Scenario-Based Approach for Stochastic Economic Model Predictive Control with an Expected Shortfall Constraint

Alireza Arastou, Algo Carè, Ye Wang, Marco Campi, and Erik Weyer

arXiv:2510.26063v1 [eess.SY] 30 Oct 2025

Abstract— **This paper presents a novel approach to stochastic economic model predictive control (SEMPC) that minimizes average economic cost while satisfying an empirical expected shortfall (EES) constraint to manage risk. A new scenario-based problem formulation ensuring controlled risk with high confidence while minimizing the average cost is introduced. The probabilistic guarantees is dependent on the number of support elements over the entire input domain, which is difficult to find for high-dimensional systems. A heuristic algorithm is proposed to find the number of support elements. Finally, an efficient method is presented to reduce the computational complexity of the SEMPC problem with an EES constraint. The approach is validated on a water distribution network, showing its effectiveness in balancing performance and risk.**

I. INTRODUCTION

Model predictive control (MPC) is a powerful control approach for a wide range of systems with constraints. The closed-loop properties of deterministic MPC are well-established [1] making it a natural choice for many applications, such as water distribution networks [2] and traffic systems [3]. In these applications, the control objective can be related to the economic cost incurred during the plant operation. This case is called economic MPC (EMPC) [4].

In practice, knowledge about the system is limited or some parameters are uncertain, introducing randomness in the control problem. Thus, the results from a deterministic MPC might not be reliable [5]. One method to cope with bounded uncertainties is to use a robust MPC strategy [1], [6]. This method can maintain stability and performance under worst-case scenarios. However, it can lead to conservative solutions. This is undesirable in applications where a degree of constraint violation is acceptable or the control objective concerns an economic aspect. Stochastic MPC (SMPC) addresses the above issues and uses a probabilistic objective function, e.g., expected value of the control objective. Moreover, hard constraints are replaced with stochastic chance constraints to allow a degree of constraint violation [5], [7]. Uncertainties and disturbances can be unbounded in this case. An SMPC with chance constraints is usually not computationally tractable. Two methods can be used to solve an SMPC problem: analytical approaches and sample-based (scenario-based) methods [7]. An analytical approach makes use of a priori knowledge about the probability distribution of uncertainties to find a deterministic equivalent of the SMPC problem [5]. A scenario-based method uses samples of the uncertain elements as approximations of their distributions and reformulates the SMPC problem [8]. Bounds are given on the number of samples required such that the solution of a scenario-based MPC meets chance constraints with a desired confidence level for a new scenario [8].

The focus of this paper is on the scenario-based approach. However, different from the standard scenario theory, we will minimize an average economic cost subject to a constraint on the risk of incurring a very large economic cost. Various risk measures have been developed to quantify potential losses under uncertainty [9]. Among these, Value-at-Risk (VaR), Conditional Value-at-Risk (CVaR), and Expected Shortfall (ES) are widely used due to their intuitive interpretation and mathematical tractability [10].

In this paper, a stochastic EMPC (SEMPC) problem with an empirical expected shortfall (EES) constraint is considered. The solution to the SEMPC does not lend itself to be effectively studied by a direct application of the results in [8], [11]–[13]. Moreover, EES is considered as a constraint; thus, the suggested method in [14], [15] cannot be used. The main contributions of the paper are *(i) Probabilistic guarantees on the economic cost for a next unseen scenario, (ii) A heuristic algorithm to find the number of support elements within the feasibility region of the SEMPC problem, (iii) A method to reduce the computational complexity of solving the scenario-based SEMPC.*

The proposed approach is applied to the Richmond water distribution network and the results are discussed. The rest of the paper is organized as follows: the problem statement is given in Section II together with a summary of the scenario approach. Section III gives the probabilistic guarantees. A heuristic algorithm to find the number of support elements within the whole input domain is given in Section IV. An efficient method to reduce computational complexity of solving the obtained control problem is presented in Section V. Simulation results are given in Section VI and conclusions and discussions are given in Section VII.

Algo Carè and Marco Campi were partly supported by the PRIN 2022 project 2022RRNAEX "The Scenario Approach for Control and Non-Convex Design" (CUP: D53D23001440006), funded by the NextGeneration EU program (Mission 4, Component 2, Investment 1.1), and by the PRIN PNRR project P2022NB77E "A data-driven cooperative framework for the management of distributed energy and water resources" (CUP: D53D23016100001), funded by the NextGeneration EU program (Mission 4, Component 2, Investment 1.1).

Alireza Arastou, and Erik Weyer are with the Department of Electrical and Electronic Engineering, The University of Melbourne, Parkville, Melbourne, Victoria, 3010, Australia, (E-mails: aarastou@student.unimelb.edu.au, ewey@unimelb.edu.au)

Algo Carè and Marco Campi are with the Department of Information Engineering, University of Brescia, 25123 Brescia, Italy (E-mails: algo.care@unibs.it; marco.campi@unibs.it).

Ye Wang is with the School of Mathematics and Statistics, The University of Melbourne, Parkville, Melbourne, Victoria, 3010, Australia, (E-mail: ye.wang@unimelb.edu.au)

II. PROBLEM STATEMENT

A stochastic optimization problem is formulated such that an average cost is minimized subject to a constraint on EES. The problem formulation is practically useful in applications such as water distribution networks, where balancing the risk of occasional high costs due to high electricity prices against low average costs is important.

A. Motivating Example

Consider an EMPC problem for minimization of pumping energy cost in a water distribution network [2]:

$$\min_{u_{0|t},\dots,u_{N-1|t}} \sum_{k=0}^{N-1} \alpha_{k|t}^\top u_{k|t}, + \Delta u_{l|t}^\top R \Delta u_{l|t} + V^f(x_{N|t})$$

$$\text{subject to} \quad \forall k \in \mathbb{I}_{[0:N-1]},$$

$$x_{0|t} = x_0, \tag{1a}$$

$$x_{k+1|t} = A x_{k|t} + B_u u_{k|t} + B_d d_{k|t}, \tag{1b}$$

$$x_{k|t} \in \mathcal{X}, \quad u_{k|t} \in \mathcal{U}, \quad x_{N|t} \in \mathcal{X}^f, \tag{1c}$$

where $x_{k|t} \in \mathbb{R}^{n\times 1}$ are states representing water levels in tanks at time $k+t$ with an initial time t, $u_{k|t} \in \mathbb{R}^{m\times 1}$ are water flows through pumps, $d_{k|t} \in \mathbb{R}^{v\times 1}$ are water demands, and $\alpha_{k|t} \in \mathbb{R}^{m\times 1}$ is the vector of electricity prices. The objective function includes the energy cost $\alpha_{k|t}^\top u_{k|t}$ which is subject to stochastic fluctuations and $\Delta u_{k|t} = u_{k|t} - u_{k-1|t}$ emphasizing input smoothness. $\mathcal{X}^f$ and $V^f(x_{N|t})$ represent terminal constraint and cost. State and input constraints are given by $\mathcal{X}$ and $\mathcal{U}$. $A \in \mathbb{R}^{n\times n}$, $B_u \in \mathbb{R}^{n\times m}$, $B_d \in \mathbb{R}^{n\times v}$ are known matrices.

Denote the cost function in (1a) by $J(\tilde{\alpha}_t, \tilde{u}_t)$, where $\tilde{\alpha}_t = [\alpha_{0|t}^\top, \dots, \alpha_{N-1|t}^\top]^\top$ and $\tilde{u}_t = [u_{0|t}^\top, \dots, u_{N|t}^\top]^\top$. In some cases water companies buy electricity directly from an electricity market and $\tilde{\alpha}_t$ and the cost in (1) are stochastic. The operational objective is to minimize pumping cost in the long run and the criterion

$$\min_{\tilde{u}_t} \mathbb{E}\left\{J(\tilde{\alpha}_t, \tilde{u}_t)\right\} \tag{2}$$

is used. A shortcoming of (2) is that very high energy prices which can happen with a small probability may occasionally lead to unacceptably high operating cost.

B. Risk Measures

Risk measures, such as value at risk (VaR) and expected shortfall (ES), are statistical tools used to evaluate risk [9]. Let $L(u,\delta)$ be a random variable where u is the decision variable and δ is the uncertainty. Let $F_{L,u}$ be cumulative distribution function (CDF) of $L(u,\cdot)$. Given a $\zeta \in (0,1)$, VaR and ES at level ζ are given by

$$\text{VaR}_\zeta(L_u) = \min\{l : F_{L,u}(l) \geq \zeta\}, \tag{3a}$$

$$\text{ES}_\zeta(L_u) = \mathbb{E}\{L_u : L_u \geq \text{VaR}_\zeta(L_u)\}. \tag{3b}$$

The distribution function $F_{L,u}$ is required to compute the risk measures in (3a)-(3b); however, it is not known in many applications. Therefore, the empirical versions of (3a)-(3b) using samples (scenarios) of the random variable is used. In this paper, we focus on EES as suggested in [14], [15]. Given N_s independent realizations of the random variable $(\delta^1, \dots, \delta^{N_s})$, we define $L_i(u) := L(u, \delta^i)$. For a given $u \in \mathcal{U}$, $L_{(i)}(u), i = 1, \dots, N_s$ are the loss functions, L_i, in descending order [15]

$$L_{(1)}(u) \geq L_{(2)}(u) \geq \ldots \geq L_{(N_s)}(u), \tag{4}$$

EES at level $1-\frac{k}{N_s}$ is the average of k-largest realizations,

$$\text{EES}_{1-\frac{k}{N_s}}(L(u)) = \frac{1}{k}\sum_{i=1}^{k} L_{(i)}(u). \tag{5}$$

It is desirable to limit the EES to control the effect of worst case scenarios. SEMPC in (1)-(2) is hence combined with EES in (5) as a constraint in a scenario optimization problem. Let $\tilde{\alpha}_t^i$ be N_s independently drawn scenarios of the vector of electricity prices over the prediction horizon N. The following problem is considered:

$$\min_{u_{0|t},\dots,u_{N-1|t}} \frac{1}{N_s}\sum_{i=1}^{N_s}\Big(\sum_{l=0}^{N-1}(\alpha_{l|t}^i)^\top u_{l|t} \tag{6a}$$

$$+ \Delta u_{l|t}^\top R \Delta u_{l|t} + V^f(x_{N|t})\Big)$$

$$\text{subject to} \quad \forall l \in \mathbb{I}_{[0:N-1]},$$

$$x_{0|t} = x_0, \tag{6b}$$

$$x_{l+1|t} = A x_{l|t} + B_u u_{l|t} + B_d d_{l|t}, \tag{6c}$$

$$x_{l|t} \in \mathcal{X}, \quad u_{l|t} \in \mathcal{U}, \quad x_{N|t} \in \mathcal{X}^f, \tag{6d}$$

$$\frac{1}{k}\sum_{j=0}^{k}\left(\sum_{l=0}^{N-1}(\alpha_{l|t}^{i_j})^\top u_{l|t}\right) \leq M, \tag{6e}$$

$$\text{for any choice of } \{i_1, \dots, i_k\} \subseteq \{1, \dots, N_s\},$$

where (6e) specifies an upper bound of M on EES. $\{i_1, \dots, i_k\}$ in (6e) is any subset of $\{1, \dots, N_s\}$ with cardinality k; however, only the constraint (6e) with indices corresponding to the k-largest scenarios is active.

A natural question is whether the obtained solution is reliable given an unseen scenario. Fundamentals of the scenario approach and the challenges of using this method for the problem in (6) are given in the next section.

C. Scenario Theory for General Decision Making [13]

Let Φ_m be a map from a set of scenarios $(\delta^1, \dots, \delta^m)$ to a decision $z \in \mathcal{Z}$, where $\mathcal{Z}$ is a generic decision set. For every δ there is a set $\mathcal{Z}_\delta$ of allowed decisions. Assume that Φ_m has the following properties [13], [16]:

Assumption 1: (Mapping properties)

- **Permutation invariance**: For a permutation of $(\delta^1, \dots, \delta^m)$ denoted by $(\delta^{i_1}, \dots, \delta^{i_m})$, we have $\Phi_m(\delta^1, \dots, \delta^m) = \Phi_m(\delta^{i_1}, \dots, \delta^{i_m})$.
- **Stability in the case of confirmation**: For any integer n, if the set of scenarios given by $(\delta^1, \dots, \delta^m, \delta^{m+1}, \dots, \delta^{m+n})$ is such that

  $$\Phi_m(\delta^1, \dots, \delta^m) \in \mathcal{Z}_{\delta^{m+i}}, \quad \forall i \in \{1, \dots, n\},$$

  then $\Phi_{m+n}(\delta^1, \dots, \delta^{m+n}) = \Phi_m(\delta^1, \dots, \delta^m)$.

- **Responsiveness to contradiction**: Let $(\delta^1, \dots, \delta^m, \delta^{m+1}, \dots, \delta^{m+n})$ be a set of scenarios where n is an integer, such that
$$\exists i \in \{1, \dots, n\} : \Phi_m(\delta^1, \dots, \delta^m) \notin \mathcal{Z}_{\delta^{m+i}},$$
then $\Phi_{m+n}(\delta^1, \dots, \delta^{m+n}) \neq \Phi_m(\delta^1, \dots, \delta^m)$.

Risk is defined as $V(z) = \mathbb{P}\{\delta \in \Delta : z \notin \mathcal{Z}_\delta\}$. Also, $\mathcal{Z}_{\delta^i}$ is called a support element if $\Phi_{m-1}(\delta^1, \dots, \delta^{i-1}, \delta^{i+1}, \dots, \delta^m) \neq \Phi_m(\delta^1, \dots, \delta^m)$ [13]. The following non-degeneracy assumption is in place.

***Assumption** 2:* Almost surely $\Phi_m(\delta^1, \dots, \delta^m)$ coincides with the obtained decision after eliminating all elements that are not support elements.

The following theorem provides a probabilistic certificate for the risk at $z_m^* = \Phi_m(\delta^1, \dots, \delta^m)$.

***Theorem** 1 ([13]):* Assume that Assumptions 1 and 2 hold. Let $\beta \in (0, 1)$ be a confidence parameter. For each $k = 0, 1, \dots, m-1$, consider the polynomial equation

$$\binom{m}{k} t^{m-k} - \frac{\beta}{2m} \sum_{i=k}^{m-1} \binom{i}{k} t^{i-k} - \frac{\beta}{6m} \sum_{i=m+1}^{4m} \binom{i}{k} t^{i-k} = 0, \tag{7}$$

which has exactly two solutions in $[0, +\infty)$, denoted by $\underline{t}(k)$ and $\bar{t}(k)$, with $\underline{t}(k) \leq \bar{t}(k)$. For $k = m$, consider the polynomial equation

$$1 - \frac{\beta}{6m} \sum_{i=m+1}^{4m} \binom{i}{k} t^{i-m} = 0, \tag{8}$$

which has a unique solution $\bar{t}(m)$. Let $\underline{t}(m) = 0$. Set $\underline{\epsilon}(k) := \max\{0, 1 - \bar{t}(k)\}$, and $\bar{\epsilon}(k) := 1 - \underline{t}(k)$, $k = 0, 1, \dots, m$. Let s_m^* be the number of support elements at z_m^*. Under Assumptions 1 and 2, it holds that

$$\mathbb{P}^m \left\{ \underline{\epsilon}(s_m^*) \leq V(z_m^*) \leq \bar{\epsilon}(s_m^*) \right\} \geq 1 - \beta. \tag{9}$$

Proof: The proof is given in [13]. ■

In our case we are interested in a bound on the energy cost $\sum_{l=0}^{N-1} \alpha_{l|t}^\top u_{l|t}$ for a new unseen scenario of the electricity prices (the real cost that will be incurred). In particularly, we would like this cost to be less than M. Bounds of this type was given in [15] for the solution $u_{l|t}^*$ which minimized the EES on the seen N_s scenarios. However, this theory is not applicable to the situation at hand since the solution here is obtained by minimizing a different objective function and the EES is acting as a constraint only.

III. PROBABILISTIC GUARANTEES

To circumvent the above problem, probabilistic certificates for the EES for the whole input domain are found. If the cost associated with a new scenario is not among the largest k costs for any u in the input domain, EES will not change for any input including the solution of (6). Therefore, (6e) is met given the added scenario. Hence, we are interested in finding guarantees for the situation that the cost associated with a new scenario is not among the largest k costs for any input and hence EES is unchanged. The probability of violation for a new scenario is the probability that the new scenario results in a cost larger than the k-largest cost for at least one input. In the following, we will find a certificate for this probability of violation.

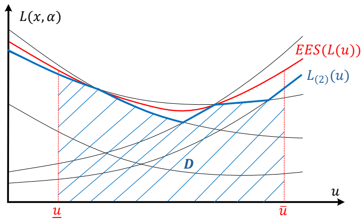


Fig. 1: Region D in (10) with $N_s = 4$, $k = 2$. The blue line is the second-largest cost, while gray lines are $L_i(u)$, $i = 1, 2, 3, 4$.

Denote the region below k-largest cost by D, i.e.,

$$D = \{(\tilde{u}, l) \in \mathcal{U}^N \times \mathbb{R} : \ \tilde{u} \in \mathcal{U}^N, \ 0 \leq l \leq L_{(k)}(\tilde{u})\}. \tag{10}$$

An example for a scalar u with $\mathcal{U} = \{u : \underline{u} \leq u \leq \bar{u}\}$, $N_s = 4$, and $k = 2$ is shown in Fig. 1.

If adding a new scenario to (10) does not change the set D, then the added scenario was not among the largest k costs for any $\tilde{u} \in \mathcal{U}^N$. Therefore, $\mathrm{EES}_{1-\frac{k}{N_s}}(L(\tilde{u}))$ is unchanged for the whole input domain. The probability of violation for the set D is defined as

$$\mathbb{P}\{\tilde{\alpha} \in \Delta : \exists \tilde{u} \in \mathcal{U}^N \text{ such that } (\tilde{u}, L(\tilde{u}, \tilde{\alpha})) \notin D\} \tag{11}$$

$D = \Phi_m(\tilde{\alpha}^1, \dots, \tilde{\alpha}^m)$ can be viewed as a decision in the framework of Section II-C and the corresponding theory can be applied to obtain the probabilistic guarantees. Formally, in the context of EES, we can define the sets

$$\mathcal{Z} = \{\bar{\mathcal{Z}} : \bar{\mathcal{Z}} \subseteq \mathcal{U}^N \times \mathbb{R}^+\}, \tag{12a}$$
$$\mathcal{Z}_{\tilde{\alpha}} := \{\mathcal{S} \in \mathcal{Z} : \forall \tilde{u} \in \Pi_{\mathcal{U}^N}(\mathcal{S}), \ (\tilde{u}, L(\tilde{u}, \tilde{\alpha})) \in \mathcal{S}\} \tag{12b}$$

where $\Pi_{\mathcal{U}^N}(\mathcal{S})$ is the projection of $\mathcal{S}$ on $\mathcal{U}^N$. For the decision D, the probability of violation is

$$V(D) = \mathbb{P}\{\tilde{\alpha} \in \Delta : D \notin \mathcal{Z}_{\tilde{\alpha}}\}. \tag{13}$$

If the corresponding cost of a scenario is above D for at least one $\tilde{u} \in \mathcal{U}^N$, then $D \notin \mathcal{Z}_{\tilde{\alpha}}$. The following theorem provides guarantees for $V(D)$ and consequently for the solution of (6).

***Theorem** 2:* Assume Assumption 2 holds. Given a set of scenarios $(\tilde{\alpha}^1, \dots, \tilde{\alpha}^{N_s})$, let the solution of (10) be given by D and denote the number of support elements within the whole input domain by $s_{N_s}^*$, then, it holds that

$$\mathbb{P}^{N_s} \left\{ \underline{\epsilon}(s_{N_s}^*) \leq V(D) \leq \bar{\epsilon}(s_{N_s}^*) \right\} \geq 1 - \beta, \tag{14}$$

where $\beta \in (0, 1)$ and $\underline{\epsilon}(s_{N_s}^*)$ and $\bar{\epsilon}(s_{N_s}^*)$ are computed from (7)-(8).

Proof: It can be verified that the map in (10) meets the conditions in Assumption 1 . Under Assumption 2, Theorem 1 holds and the certificate in (14) is established. ■

The corollary below shows that (14) can be used to find a certificate for EES at the solution of (6).

Corollary 1: Given an input $\tilde{u} \in \mathcal{U}^N$, let $\bar{V}(\tilde{u}) = \mathbb{P}\{\tilde{\alpha} \in \Delta : L(\tilde{\alpha}, \tilde{u}) > \text{EES}_{1-\frac{k}{N_s}}(\tilde{u})\}$. Also, let the solution of (6) using a set of scenarios $(\tilde{\alpha}^1, \dots, \tilde{\alpha}^{N_s})$, be given by $\tilde{u}^*$. It holds that

$$\mathbb{P}^{N_s}\left\{\bar{V}(\tilde{u}^*) \le \bar{\epsilon}(s^*_{N_s})\right\} \ge 1 - \beta. \tag{15}$$

Proof: We know that

$$\begin{aligned}&\bar{V}(\tilde{u}^*) \le \mathbb{P}\{\tilde{\alpha} \in \Delta : L(\tilde{\alpha}, \tilde{u}^*) > L_{(k)}(\tilde{u}^*)\}\\&\le \mathbb{P}\{\tilde{\alpha} \in \Delta : \exists \tilde{u} \in \mathcal{U}^N \text{s.t. } L(\tilde{\alpha}, \tilde{u}) > L_{(k)}(\tilde{u})\} = V(D)\end{aligned}$$

It can be concluded that

$$\begin{aligned}&\mathbb{P}^{N_s}\left\{\bar{V}(\tilde{u}^*) \le \bar{\epsilon}(s^*_{N_s})\right\} \ge \mathbb{P}^{N_s}\left\{V(D) \le \bar{\epsilon}(s^*_{N_s})\right\}\\&\ge \mathbb{P}^{N_s}\left\{\underline{\epsilon}(s^*_{N_s}) \le V(D) \le \bar{\epsilon}(s^*_{N_s})\right\} \ge 1 - \beta\end{aligned} \tag{16}$$

■

IV. FINDING THE NUMBER OF SUPPORT ELEMENTS

In (10), the support elements are the scenarios that are among the largest k costs for at least one $\tilde{u} \in \mathcal{U}^N$. Thus, one method to find support elements is to evaluate costs for the whole input domain, order them, and find the largest k costs for every input. However, this method is not computationally feasible for high dimensional systems. In this section, we propose a sample-based method to find the number of support elements for the whole input domain.

The idea is to draw i.i.d samples from the input domain, find the number of support elements using the obtained samples and test the obtained solution on additional samples of the input. If the obtained solution passes the test, the algorithm will output the obtained number of support elements. Otherwise, new scenarios are added to the set of support elements. The proposed method is given in Algorithm 1.

$\hat{p}$ = (number of support elements)$/N_T$ is the empirical probability of finding a new support element when tested against N_T new input samples. Let $\mu - \rho$ be a desired upper bound on the probability of finding a new support element. Thus, it is desirable to set a very small value for $\mu - \rho \in [0, 1]$, so that the probability that we get a new number of support elements for new samples of input be negligible. The break criterion in Algorithm 1 is motivated by the following result, which also provides a guideline for how N_T can be chosen.

Lemma 1: Choose a distribution on $\mathcal{U}^N$, from which $\tilde{u}$ is independently sampled. Let $A \subset \{\tilde{\alpha}_1, \dots, \tilde{\alpha}_{N_s}\}$ be a set of support elements. Let p be the probability of drawing a $\tilde{u}$ such that there is an $\tilde{\alpha}^* \in \{\tilde{\alpha}_1, \dots, \tilde{\alpha}_{N_s}\} \setminus A$ with the property that $(\tilde{\alpha}^*)^\top \tilde{u}$ is among the k-largest values of $(\tilde{\alpha}^i)^\top \tilde{u}$, $i = 1, \dots N_s$. Let $\hat{p}$ be the empirical frequency of this event evaluated on an i.i.d. sample $\tilde{u}_1, \dots, \tilde{u}_{N_T}$. If N_T satisfies

$$\sum_{i=0}^{\lfloor N_T(\mu-\rho) \rfloor} \binom{N_T}{i} \mu^i (1-\mu)^{N_T - i} < \bar{\beta}, \tag{17}$$

it holds that $P^{N_T}\{\hat{p} \le \mu - \rho \text{ and } p > \mu\} < \bar{\beta}$, where P^{N_T} is the probability measure on the input samples.

Proof: The proof is similar to that of Lemma 1 in [17]. ■

The bound in Lemma 1 applies to the procedure run at a single iteration of the while loop for one instance of $\mathcal{I}_k$. However, note that during an execution of Algorithm 1 several versions of $\mathcal{I}_k$ can be tested and the returned $\mathcal{I}_k$ depends on the outcomes of a number of tests (an *a priori* unknown quantity); therefore, the bound in Lemma 1 is not directly applicable to the set of support elements $\mathcal{I}_k$ returned by the algorithm.

The input domain is restricted by the constraints in (6). Thus, if the input samples are drawn from the feasibility region of SEMPC problem, the number of support elements may be reduced in comparison with the candidate number of support elements obtained by sampling from the whole input domain, i.e., $\mathcal{U}^N$ and we can get tighter guarantees from (14).

Let us denote the restricted input domain from (1c) by $\mathcal{P}$. Given an initial state x_0 and a steady-state x^s, assume in (1c) that $\mathcal{X} = \{x : \underline{x} \le x \le \overline{x}\}$, $\mathcal{U} = \{u : \underline{u} \le u \le \overline{u}\}$, and $\mathcal{X}^f = \{x : (x - x^s)^\top \Omega (x - x^s) \le \kappa\}$, where Ω is a positive definite matrix and $\kappa > 0$. $\mathcal{P}$ can be represented by

$$\tilde{u} \in \mathcal{U}^N, \; \bar{B}\tilde{u} \le \bar{A}, \; (\hat{B}\tilde{u} + \gamma)^\top \Omega (\hat{B}\tilde{u} + \gamma) \le \kappa, \tag{18}$$

where $\bar{A}, \bar{B}$, $\hat{B}$, and γ are matrices and vectors with appropriate dimensions that include model dynamics, x_0, and x^s.

An optimization-based method is proposed in this part to find support elements within $\mathcal{P}$. Let us assume $\mathcal{I}_k$ contains the indices of candidate support elements within the box of input constraint, i.e., $\mathcal{U} = \{u : \underline{u} \le u \le \overline{u}\}$. For an index $j \in \mathcal{I}_k$, if there exists a $\tilde{u} \in \mathcal{P}$, such that $L(\tilde{u}, \tilde{\alpha}^j)$ is among the largest k values, j is a support element in $\mathcal{P}$. Otherwise, we can remove j from $\mathcal{I}_k$. The following program is solved for every $j \in \mathcal{I}_k$

$$\max_{\tilde{u}, \{z_i\}} \sum_{i=1, i \ne j}^{N_s} z_i, \tag{19a}$$

$$\begin{aligned}\text{subject to: } & i = 1, \dots, N_s, \; i \ne j,\\ & \tilde{u} \in \mathcal{P}, \;\; (\tilde{\alpha}^j)^\top \tilde{u} \ge (\tilde{\alpha}^i)^\top \tilde{u} - G(1 - z_i), \quad \text{(19b)}\\ & \sum_{i=1}^{N_s} z_i \ge N - k, , \;\; z_i \in \{0, 1\}, \quad \text{(19c)}\end{aligned}$$

where G is a positive large number. If the above problem is feasible for a $j \in \mathcal{I}_k$, j is a support element in $\mathcal{P}$. If $(\tilde{\alpha}^j)^\top \tilde{u} \ge (\tilde{\alpha}^i)^\top \tilde{u}$, (19b) will be met with either $z_i = 1$ or 0. Since the objective function in (19a) is maximized, then $z_i = 1$ wherever $(\tilde{\alpha}^j)^\top \tilde{u} \ge (\tilde{\alpha}^i)^\top \tilde{u}$. If $(\tilde{\alpha}^j)^\top \tilde{u} < (\tilde{\alpha}^i)^\top \tilde{u}$, $z_i = 0$ in (19b). Therefore, $\sum_{i=1, i \ne j}^{N_s} z_i$ is the number of scenarios that are below j-th scenario and it is required to be greater than $N - k$ in (19c) to make sure the j-th scenario is among top k scenarios. Feasibility of (19) implies that the scenario index by j is of support.

V. COMPUTATIONALLY TRACTABLE SEMPC PROBLEM

The obtained problem in (6) is a convex quadratic program. However, considering every possible choice of k indices from $1, \dots, N_s$ makes the computational burden high.

Algorithm 1 Finding the number of support elements

```
Inputs: $(\tilde{\alpha}^1, \dots, \tilde{\alpha}^{N_s})$, $N_r$, $N_T$, $k$, $\mu$, $\rho$;
Generate $N_r$ i.i.d samples according to a uniform distribution on $\mathcal{U}^N$, i.e., $(\tilde{u}_1, \dots, \tilde{u}_{N_r})$;
Evaluate $L(\tilde{u}_j, \tilde{\alpha}^i)$ for all $i = 1, \dots, N_s$ and $j = 1, \dots, N_r$;
Find the indices of the scenarios that are among the largest $k$ costs and store them in a set $\mathcal{I}_k$;
while $|\mathcal{I}_k| \leq N_s$ do
    Generate $N_T$ new input samples according to a uniform distribution on $\mathcal{U}^N$ for testing $\mathcal{I}_k$, i.e., $(\tilde{u}_1, \dots, \tilde{u}_{N_T})$;
    for each $\tilde{u}_l$, $l = 1, \dots, N_T$ do
        Find the $k$ scenarios with largest values and store them in $\bar{\mathcal{I}}_l$;
        Define $B_l = \begin{cases} 0, & \bar{\mathcal{I}}_l \subset \mathcal{I}_k, \\ 1, & \text{otherwise.} \end{cases}$
        If $B_l = 1$, add the new support elements to $\mathcal{I}_k$;
    end for
    Compute $\hat{p} = \frac{1}{N_T} \sum_{l=1}^{N_T} B_l$;
    if $\hat{p} \leq \mu - \rho$ then
        break;
    end if
end while
Output: A set of support elements $\mathcal{I}_k$.
```

An equivalent computationally suitable formulation can be obtained by using [18]

***Lemma** 2:* Given $\tilde{u} \in \mathcal{U}^N$ and a collection of scenarios $\{L_i(\tilde{u})\}_{i=1}^{N_s}$, the sum of k-largest functions is given by

$$\min_{\bar{t}, \{\lambda_i\}_{i=1}^{N_s}} \; k\bar{t} + \sum_{i=1}^{N_s} \lambda_i \tag{20a}$$

$$\text{subject to}: \lambda_i \geq L_i(\tilde{u}) - \bar{t}, \quad \lambda_i \geq 0, \; i = 1, \dots, N_s. \tag{20b}$$

Proof: See [18]. ■

In the SEMPC problem in (6), (6e) can be replaced by

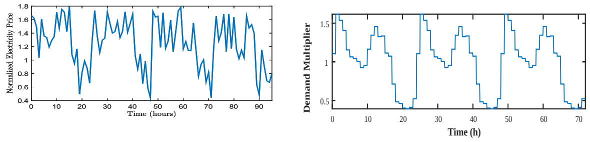
$$\lambda_j \geq \sum_{l=0}^{N-1} (\alpha_{l|t}^j)^\top u_{l|t} - \bar{t}, \; \lambda_j \geq 0, \; , j = 1, \dots, N_s \tag{21a}$$

$$\frac{1}{k}\left(k\bar{t} + \sum_{j=1}^{N_s} \lambda_j\right) \leq M. \tag{21b}$$

(21a)-(21b) reduces the computational burden compared to (6e). t and $\{\lambda_i\}_{i=1}^{N_s}$, which are obtained from (21) are not necessarily the minimizer of (20); thus, (21b) is an upper bound on the sum of the k-largest costs, i.e., EES.

VI. Case Study: Richmond Water Network

The Richmond water network is part of the Yorkshire water supply area in U.K. [19]. More details about the model of this network is given in [2], [20]. The control objective is given in (6). The random vectors $\alpha_{l|t}$ are the electricity price, and it is assumed that i.i.d samples of them over the time horizon are available. The samples are drawn from a sum of a deterministic cost with two tariffs and a uniform random variable. A realization of electricity prices is shown in Fig. 2a.



(a) A realization of electricity prices

(b) The demand multiplier

Fig. 2: Electricity price and the demand multiplier

The prediction horizon was $N = 30$ hours, $\mathcal{X} = \{x : \underline{x} \leq x \leq \overline{x}\}$, $\mathcal{U} = \{u : \underline{u} \leq u \leq \overline{u}\}$, $\mathcal{X}^f = \{x : (x - x^s)^\top \Omega (x - x^s) \leq \kappa\}$ with $x^s = 0.5(\underline{x} + \overline{x})$, $\Omega = I_n$, and $\kappa = 0.8$. The water demand at time t is expressed as $d_t = m_t \bar{d}$, where m_t represents the demand multiplier illustrated in Fig. 2b. This multiplier has an average of 1, meaning that the average demand is $\bar{d}$, with different $\bar{d}$ for each individual demand.

A. *Effect of Design Parameters on the Obtained Guarantees*

The probabilistic guarantee is obtained for the whole feasibility region and Corollary 1 ensures that the guarantee can be used for the solution of (21). To this end, the mapping in (10) is considered, and the number of support elements is obtained using Algorithm 1 and (19). As the obtained number of support elements is a lower bound on the actual number; the guarantees are approximate. Note that the scenarios are changed in a moving horizon regime and the same procedure is carried out to find guarantees at each time step.

A comparison of the number of support elements within $\mathcal{P}$ characterized by (1c) and the $\mathcal{U}^N$ box for different design parameters in Algorithm 1 is given in Table I. $k = 2$, $\rho = \frac{\mu}{2}$, $N_r = 3000$, and $\beta = 10^{-6}$ in all of the cases.

TABLE I: Comparison of the number of support elements using various inputs in Algorithm 1. s_{box} and $s_{\mathcal{P}}$ are the number of support elements within the $\mathcal{U}^N$ box and $\mathcal{P}$.

N_s	μ	$\overline{\beta}$	N_T	s_{box}	$s_{\mathcal{P}}$	$\underline{\epsilon}(s_{\mathcal{P}})$	$\bar{\epsilon}(s_{\mathcal{P}})$
2000	10^{-4}	10^{-6}	733984	45	31	0.005	0.037
2000	10^{-3}	10^{-5}	57886	31	23	0.003	0.031
10000	10^{-3}	10^{-5}	57886	39	32	0.001	0.007

The number of input samples in Table I, N_T, is computed from (17). N_T increased considerably, due to the reduced μ to provide a better guarantee for the obtained number of support elements as stated in Lemma 1. Since the testing stage in Algorithm 1 is easy to be carried out, increasing N_T will not affect practicality of the algorithm significantly.

B. *SEMPC with EES constraint*

The proposed control problem in (21) was applied to the Richmond water network with $M = 2150$ and $N_s = 2000$. The water level in tank A (a state) and the flow through pump A (an input) are shown in Fig. 3. To show the effectiveness

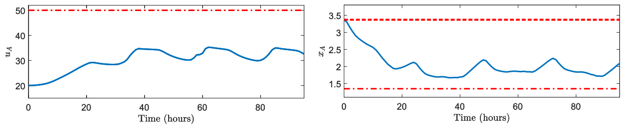


Fig. 3: The simulation results for u_A and x_A (dashed red lines indicate constraints)

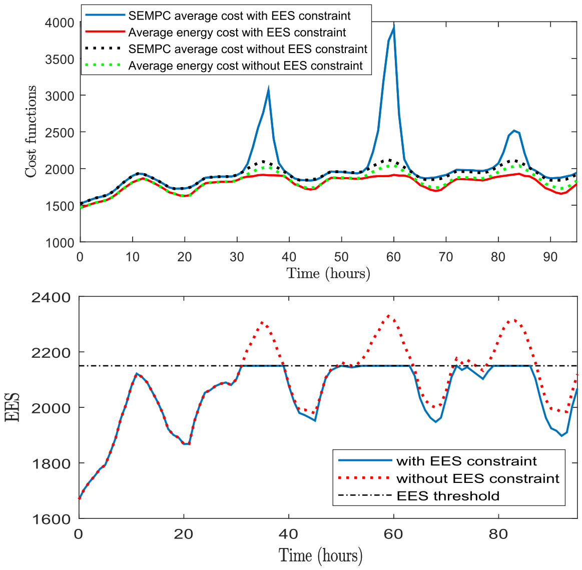


Fig. 4: Average cost and EES with and without the EES constraint

of the proposed method, the SEMPC cost and the EES are shown in Fig. 4 with and without the EES constraint.

It can be seen from these figures that the EES constraint ensures that risk is below the desired value at each time step; however, it resulted in a higher SEMPC cost. The EES constraint resulted in high values for the other terms in the SEMPC cost, such as input smoothness term. Hence, some peaks with large values are visible in Fig. 4.

Using the number of support scenarios found by Algorithm 1, $\bar{\epsilon}(s_{\mathcal{P}})$ at the SEMPC solution was between 0.025 and 0.045 at each time where the values $k = 2$, $\rho = \frac{\mu}{2}$, $N_r = 3000$, $\mu = 10^{-3}$, $\bar{\beta} = 10^{-5}$, and $\beta = 10^{-6}$ were used. The guarantee at each time step was obtained from Corollary 1.

VII. CONCLUSION

The paper introduces an SEMPC strategy that integrates an EES constraint to effectively manage risk while minimizing average cost. A probabilistic certificate is obtained for the solution of the SEMPC strategy using the number of support elements for the whole feasibility region of the control problem. The support elements are the scenarios that are among the largest k costs for some input. Evaluating the cost function and ordering them for all combinations of inputs to find the number of support elements for a high dimensional system is cumbersome. Thus, a heuristic algorithm is proposed to address this challenge effectively. Moreover, the EES constraint increases the computational complexity of solving the control problem considerably. A reformulation is proposed in this paper to reduce the computational complexity.

REFERENCES

[1] J. B. Rawlings, D. Q. Mayne, and M. Diehl, *Model predictive control: theory, computation, and design*, vol. 2. Nob Hill Publishing Madison, WI, 2017.

[2] Y. Wang, K. Too Yok, W. Wu, A. R. Simpson, E. Weyer, and C. Manzie, “Minimizing pumping energy cost in real-time operations of water distribution systems using economic model predictive control,” *Journal of Water Resources Planning and Management*, vol. 147, no. 7, 2021.

[3] K.-D. Kim and P. R. Kumar, “An MPC-based approach to provable system-wide safety and liveness of autonomous ground traffic,” *IEEE Transactions on Automatic Control*, vol. 59, no. 12, pp. 3341–3356, 2014.

[4] D. Angeli, R. Amrit, and J. B. Rawlings, “On average performance and stability of economic model predictive control,” *IEEE Transactions on Automatic Control*, vol. 57, no. 7, pp. 1615–1626, 2012.

[5] T. A. N. Heirung, J. A. Paulson, J. O’Leary, and A. Mesbah, “Stochastic model predictive control — how does it work?,” *Computers and Chemical Engineering*, vol. 114, pp. 158–170, 2018. FOCAPO/CPC 2017.

[6] A. Bemporad and M. Morari, “Robust model predictive control: A survey,” in *Robustness in identification and control*, pp. 207–226, Springer, 2007.

[7] M. Farina, L. Giulioni, and R. Scattolini, “Stochastic linear model predictive control with chance constraints–a review,” *Journal of Process Control*, vol. 44, pp. 53–67, 2016.

[8] M. C. Campi, S. Garatti, and M. Prandini, “Scenario optimization for MPC,” *Handbook of Model Predictive Control*, pp. 445–463, 2019.

[9] P. Artzner, F. Delbaen, J.-M. Eber, and D. Heath, “Coherent measures of risk,” *Mathematical finance*, vol. 9, no. 3, pp. 203–228, 1999.

[10] R. T. Rockafellar, S. Uryasev, *et al.*, “Optimization of conditional value-at-risk,” *Journal of risk*, vol. 2, pp. 21–42, 2000.

[11] M. C. Campi and S. Garatti, “Wait-and-judge scenario optimization,” *Mathematical Programming*, vol. 167, pp. 155–189, 2018.

[12] M. C. Campi and S. Garatti, “The exact feasibility of randomized solutions of uncertain convex programs,” *SIAM Journal on Optimization*, vol. 19, no. 3, pp. 1211–1230, 2008.

[13] S. Garatti and M. C. Campi, “Risk and complexity in scenario optimization,” *Mathematical Programming*, vol. 191, no. 1, pp. 243–279, 2022.

[14] G. Arici, M. C. Campi, A. Carè, M. Dalai, and F. A. Ramponi, “A theory of the risk for empirical CVaR with application to portfolio selection,” *Journal of Systems Science and Complexity*, vol. 34, pp. 1879–1894, 2021.

[15] F. A. Ramponi and M. C. Campi, “Expected shortfall: Heuristics and certificates,” *European Journal of Operational Research*, vol. 267, no. 3, pp. 1003–1013, 2018.

[16] M. C. Campi, A. Carè, and S. Garatti, “The scenario approach: A tool at the service of data-driven decision making,” *Annual Reviews in Control*, vol. 52, pp. 1–17, 2021.

[17] H. A. Nasir, A. Carè, and E. Weyer, “A scenario-based stochastic MPC approach for problems with normal and rare operations with an application to rivers,” *IEEE Transactions on Control Systems Technology*, vol. 27, no. 4, pp. 1397–1410, 2018.

[18] W. Ogryczak and A. Tamir, “Minimizing the sum of the k largest functions in linear time,” *Information Processing Letters*, vol. 85, no. 3, pp. 117–122, 2003.

[19] U. of Exeter, “Richmond network.” `https://engineering.exeter.ac.uk/research/cws/resources/benchmarks/`, 2000. Accessed February 21, 2023.

[20] A. Arastou, Y. Wang, and E. Weyer, “An optimization-based network partitioning method considering local controllability: Application to water distribution networks,” in *2023 62nd IEEE Conference on Decision and Control (CDC)*, pp. 6370–6375, 2023.