Optimal steady-state disturbance compensation for constrained linear systems: the Gaussian noise case

Alessandro Falsone, *Member, IEEE,* Luca Deori, Daniele Ioli,
Simone Garatti, *Member, IEEE,* Maria Prandini, *Fellow, IEEE*

***Abstract*—We consider the problem of designing a disturbance compensator for a discrete time linear system, so as to optimize a performance index while satisfying probabilistic state and input constraints in steady-state conditions. The problem is formulated as a chance-constrained program that depends on the compensator parameters through the state and input stationary distributions. In this paper, we focus on the Gaussian noise case and provide an analytic expression of the stationary state distribution as a function of the compensator parameters. This expression can be used in the chance-constrained program, which can then be tackled via the scenario approach. Some useful extensions of the set-up are also discussed to further broaden the applicability of the approach. Performance of the proposed design methodology is shown on a building energy management problem where cyclostationary disturbances are compensated, thus providing a stochastic periodic control solution.**

***Index Terms*—Disturbance compensation, optimal constrained control, stochastic linear systems.**

I. Introduction

THIS paper is concerned with the design of a disturbance compensator optimizing performance for a constrained discrete time linear system affected by an additive white Gaussian noise and operating in stationary conditions.

The compensator design entails characterizing and optimally shaping the distribution of the stationary state process, which depends on the compensator parameters. The problem addressed resembles the ones tackled by the minimum variance (MV) ([1]), generalized minimum variance (GMV) ([2], [3], [4], [5], [6], [7]), and H_2 ([8, pag. 273]) control approaches, where the stationary state distribution is shaped so as to minimize the variance of a given output signal. However, in contrast to those design methodologies, we are not limited to the variance as cost function, and we can explicitly enforce (probabilistic) state and input constraints without resorting to the constraint softening solution adopted in the MV and GMV literature, where a penalty term is added to the cost function, see e.g., [9]. In turn, we address the design of a disturbance compensator and not of a state feedback controller.

Our set-up is motivated by problems arising in the energy domain, such as the optimal operation of building Heating Ventilation and Air Conditioning (HVAC) systems and grid-connected solar photovoltaic power plants, where disturbances (specifically, the ambient temperature and the solar energy production) can be easily measured and directly compensated. In some applications also, the state of the system is hardly accessible and feedback control strategies cannot be applied. An example is represented by the thermal control of a building, where the state variables are the temperatures of the slices of the walls modeling the inertia of the building acting as a passive storage, [10], [11], [12].

The optimal design of a steady-state disturbance compensator for a constrained system subject to unbounded stochastic disturbances is naturally formulated as a chance-constrained optimization program, where constraints are imposed in probability. This kind of program is, in general, hard to solve exactly, [13], and analytic ([14], [15], [16]) as well as randomized ([17], [18], [19], [20]) methods have been introduced to solve them by approximating the probabilistic constraints. Unfortunately, these methods are not directly applicable in our setting because of the dependence of the constraints on the compensator parameters through the stationary state distribution.

In this paper, we show that, if the system is affected by white Gaussian noise, then the Gaussian stationary state distribution can be *explicitly characterized through its mean and covariance, which are both computed as analytic functions of the compensator parameters. As a result of this characterization, the so-called scenario approach ([21], [22], [18], [17]) can be directly applied* so as to provide a randomized solution with guarantees of chance-constrained feasibility.

In our recent work [23], [24], we addressed the same problem but without assuming the noise to be either Gaussian or white. though the setup is more general than that of the present paper, the solution in [23], [24] is completely different from the one introduced here because the stationary state distribution is generally not Gaussian and, hence, cannot be characterized only through its first and second order moments. In [23], [24] then, the distribution is approximated by introducing a suitable truncation of the series defining the stationary state process, and one has to resort to a tightening of the constraints to compensate for the introduced error and allow for the guarantees of the scenario approach to still hold, resulting in a possibly conservative solution. In the approach proposed in this paper we avoid any approximation and tightening, while retaining the advantages of the solution in [23], which are: i) *the compensator parameters are computed once and off-line*, and ii) the system with the compensator is *guaranteed to achieve the designed performance and satisfy the probabilistic constraints, in steady-state conditions.* Also, apart from convexity, no specific functional form of cost and constraints is assumed, enhancing the applicability of the

All the authors are with the Dipartimento di Elettronica, Informazione e Bioingegneria, Politecnico di Milano, Via Ponzio 34/5, 20133 Milano, Italy. E-mail: name.surname@polimi.it

approach.

A further contribution of this paper is represented by the extension of the proposed methodology in several directions in order to deal with average cost functions, non-asymptotically stable systems, disturbance compensation in presence of noisy measurements, and cyclostationary disturbances.

Infinite horizon optimal control problems that account for probabilistic constraints are typically tackled within a stochastic model predictive control (SMPC) framework (see [25] for a survey). However, SMPC requires to measure the state of the system and solve on-line an optimization problem at each time instant. Also, we are not aware of any result on the characterization of SMPC performance in the long run, and not even on the satisfaction of the state and input constraints in probability.

The rest of the paper is organized as follows. We first provide some notations. The compensator design problem is formulated in Section II and the proposed resolution strategy is given in Section III. Then we discuss several extensions including periodic stochastic control in Section IV. Finally, a numerical case study related to building energy management is illustrated in Section V and some concluding remarks are drawn in Section VI. The proofs of the main results are deferred to Appendices A-C.

a) Notations: Given a discrete time process $\{v_k,\ k \in \mathbb{Z}\}$, we denote it as $\boldsymbol{v}$ and the probability distribution of $\boldsymbol{v}$ as $\mathbb{P}_{\boldsymbol{v}}$. Correspondingly, the expected value operator with respect to $\mathbb{P}_{\boldsymbol{v}}$ is denoted as $\mathbb{E}_{\boldsymbol{v}}[\,\cdot\,]$. For a vector v, $[v]_i$ denotes its i-th element and for a matrix X, $[X]_{ij}$ denotes the element corresponding to row i and column j. I_n denotes the identity matrix of order n (the subscript is dropped when it is obvious from the context), $J = \mathrm{blkdiag}(J_1, \dots, J_m)$ is the block diagonal matrix built from the square matrices $J_1, \dots, J_m$, $\mathrm{vec}(X)$ is the vectorization operator applied to the matrix X which stacks all columns of X into a single column vector, $\otimes$ denotes the Kronecker product, and X' and $X^\top$ denote the transpose and conjugate transpose of X, respectively. The Euclidean norm of a vector v is denoted as $\|v\|_2$. With $\lceil a \rceil$ we denote the smallest integer greater than or equal to a, $\mathrm{mod}(b, c)$ denotes the remainder of b divided by c, with b and c integers, and $\binom{n}{k}$ denotes the binomial coefficients, which is assumed to be zero if $n < k$, with n and k integers.

II. Problem formulation

Consider a discrete time linear system with state $x_k \in \mathbb{R}^{n_x}$ evolving according to the following recursive equation

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

where $u_k \in \mathbb{R}^{n_u}$ is the control input, $d_k \in \mathbb{R}^{n_d}$ is a stochastic disturbance, and A, B, and W are matrices of appropriate dimensions.

We make the following assumptions.

Assumption 1 (Asymptotic stability)**.** *The spectral radius of matrix A satisfies $\rho_A < 1$.* □

Assumption 2 (Gaussian white noise)**.** *The process $\boldsymbol{d}$ is a sequence of independent and identically distributed Gaussian random vectors with zero mean and covariance matrix $\Sigma_d = S_d S_d^\top$.* □

Note that, in case the state is available for feedback, Assumption 1 can be relaxed to a stabilizability requirement on the couple (A, B), as discussed in Section IV-B. Furthermore, the zero mean assumption is without loss of generality, since if this is not the case, we can introduce $\bar{x}_{k+1} = A\bar{x}_k + W\bar{d}$, where $\bar{d} = \mathbb{E}_{\boldsymbol{d}}[d_k]$, and reformulate the problem in terms of $\Delta x_k = x_k - \bar{x}_k$ which evolves according to $\Delta x_{k+1} = A\Delta x_k + Bu_k + W\Delta d_k$, and is affected by the zero mean process $\Delta d_k = d_k - \bar{d}$.

Let us further assume that the value taken by d_k, at any time $k \in \mathbb{Z}$, is available for compensation purposes.[1] Extension to the case of noisy measurements of d_k will be discussed in Section IV-C.

Our aim is designing a disturbance compensator so as to optimize some performance criterion while satisfying state and input constraints for the controlled system operating in stationary conditions. We consider a compensator of the following form:

$$u_k = \gamma + \vartheta d_k, \tag{2}$$

where the control input u_k is taken to be an affine function of d_k, with the compensator parameters γ and ϑ taking values in the convex and compact sets $\Gamma \subset \mathbb{R}^{n_u}$ and $\Theta \subset \mathbb{R}^{n_u \times n_d}$, respectively.

Note that if we apply the compensator (2) to system (1), we get

$$x_{k+1} = Ax_k + B\gamma + (B\vartheta + W)d_k, \tag{3}$$

which originates a well-defined stationary process. Indeed, under Assumptions 1 and 2, by [26, Theorem 1.4, pag. 80], for any $k \in \mathbb{Z}$ there exists a measurable function $x_{k,\infty}$ of the process $\boldsymbol{d_{k-1}} = \{\dots, d_{k-2}, d_{k-1}\}$ such that the process $\boldsymbol{x_\infty} = \{x_{k,\infty},\ k \in \mathbb{Z}\}$ satisfies (3) and is strictly stationary with finite first and second order moments. Furthermore, $x_{k,\infty}$ is unique (see [26, Theorem 3.2, pag. 101]), its probability distribution is Gaussian (see [27, pag. 304]) and is induced from that of $\boldsymbol{d_{k-1}}$. Evidently, the mean and the covariance matrix characterizing the distribution of the Gaussian process $x_{k,\infty}$ depend on the optimization variables γ and ϑ. Since $x_{k,\infty}$ depends on $\boldsymbol{d_{k-1}}$ and process $\boldsymbol{d}$ is white (see Assumption 2), $x_{k,\infty}$ and d_k are independent.

Let $\ell(x, u, d) : \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \times \mathbb{R}^{n_d} \to \mathbb{R}$ be a function associating a cost to the state/control input pair (x, u) when the disturbance value is d. Moreover, suppose that the state x and the input u are subject to a joint constraint defined by imposing that some scalar function $f(x, u) : \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \to \mathbb{R}$ is not positive.[2]

[1] Note that in general is not possible to cancel out the contribution of d_k on the state dynamics (1) by setting u_k so that $Bu_k = -Wd_k$ because B is typically not invertible and u_k may be subject to constraints.

[2] The fact that $f(\cdot)$ is a single constraint function is without loss of generality because if multiple constraint functions $f_1(\cdot), \dots, f_m(\cdot)$ are present, then, we can redefine $f(\cdot)$ as the point-wise maximum of these $f_1(\cdot), \dots, f_m(\cdot)$.

Then, our optimal compensator design problem can be formulated as the following chance-constrained optimization program:

$$\begin{aligned} \min_{\gamma\in\Gamma,\vartheta\in\Theta,h} \quad & h \\ \text{subject to:} \quad & \mathbb{P}_{\boldsymbol{d_k}}\Big\{\ell(x_{k,\infty},\gamma+\vartheta d_k,d_k)\le h, \\ & \qquad \wedge\, f(x_{k,\infty},\gamma+\vartheta d_k)\le 0\Big\}\ge 1-\varepsilon, \end{aligned} \tag{4}$$

where $\varepsilon \in (0,1)$ is a user-chosen probability level and $\mathbb{P}_{\boldsymbol{d_k}}$ is the probability distribution of process $\boldsymbol{d_k} = \{\dots, d_{k-1}, d_k\}$.

The interpretation of the design problem formulation (4) is that we are minimizing the cost and requiring that the constraint is satisfied over all disturbance realizations except for a violation set of probability measure at most ε. This ε acts as a tuning parameter making the solution more conservative as it decreases to 0 (worst case solution).

Notice that functions $\ell(x,u,d)$ and $f(x,u)$ in (4) are evaluated in stationary conditions, i.e., with x set equal to the stationary state $x_{k,\infty}$ and with u given by the disturbance compensator in (2). Solving the optimal compensator design problem then amounts to characterizing and suitably shaping the Gaussian distribution of the stationary process $x_{k,\infty}$. By the joint stationarity of processes $\boldsymbol{x_\infty}$ and $\boldsymbol{d}$, the solution to (4) is optimal and satisfies the probabilistic constraint in (4) for every time instant k. Optimality and feasibility, however, are guaranteed only when the system is operating in stationary conditions. In practice, thanks to Assumption 1, stationarity is always reached in the long run, with a convergence rate that depends on ρ_A. This makes our compensator design particularly appealing because the control law can be computed once (off-line) and then applied at each time step without solving any further on-line optimization problem.

Chance-constrained problems like (4) are generally challenging to solve because of the presence of the probabilistic constraint, which, apart from some notable exceptions (see also Remark 3), is not easy to express analytically as a function of the optimization variables and can be non convex even when both $\ell(\cdot)$ and $f(\cdot)$ are convex functions and the probability distribution is Gaussian, [13], [28].

In the next section, we will show how to approximately solve (4) via the scenario approach, under the following assumption on functions $\ell(x,u,d)$ and $f(x,u)$.

Assumption 3 (Convexity)**.** *The cost function $\ell(x,u,d)$ and the constraint function $f(x,u)$ are convex with respect to $(x,u) \in \mathbb{R}^{n_x} \times \mathbb{R}^{n_u}$.* □

The scenario solution will be feasible for the original chance-constrained problem (4), with a certain confidence level, thus providing a compensator satisfying the joint state-input constraint and with guaranteed performance (the value obtained for the h optimization variable) for all disturbance realizations except for a set of probability measure at most ε.

III. Proposed solution

According to the scenario theory, an approximate solution to (4) can be computed solving a sampled version of (4), where the constraint in probability is replaced with N realization of the inequalities inside the probability. Such realizations are computed evaluating $\ell(\cdot)$ and $f(\cdot)$ at N samples $\{d_k^{(i)}\}_{i=1}^N$ and $\{x_{k,\infty}^{(i)}\}_{i=1}^N$ (which are called scenarios) of the stochastic variables d_k and $x_{k,\infty}$, respectively.

Both d_k and $x_{k,\infty}$ are Gaussian random variables, but whilst one can easily extract random samples of d_k since it has zero mean and known covariance matrix Σ_d, generating samples of $x_{k,\infty}$ is *not* straightforward since its mean and covariance matrix depend on the optimization variables γ and ϑ, which are yet to be determined.

A first and main contribution of this paper is to provide (i) a characterization of the probability distribution of $x_{k,\infty}$ as a function of γ and ϑ and (ii) an expression for the generic i-th sample $x_{k,\infty}^{(i)}$ of $x_{k,\infty}$ whose dependency on γ and ϑ results in $\ell(\cdot)$ and $f(\cdot)$ being convex as a function of γ and ϑ.

Theorem 1 (Characterization of $x_{k,\infty}$)**.** *Under Assumptions 1 and 2, we have that $x_{k,\infty}$ is a stationary Gaussian process with mean $\bar{x} = (I-A)^{-1}B\gamma$ and covariance matrix $\Sigma_x = S_x(\vartheta)S_x(\vartheta)^\top$, with*

$$S_x(\vartheta) = T\bar{\Omega}(\vartheta)(I_{n_d} \otimes S_J), \tag{5}$$

where T is a nonsingular transformation matrix that reduces A to a matrix J in Jordan canonical form, $S_J S_J^\top = \sum_{s=0}^{\infty} \operatorname{vec}(J^s)\operatorname{vec}(J^s)^\top < \infty$, and $\bar{\Omega}(\vartheta) = (\operatorname{vec}(T^{-1}(B\vartheta + W)S_d)' \otimes I_{n_x})$ is linear as a function of ϑ.

Proof. See Appendix B. □

Remark 1 (Computing S_J appearing in (5))**.** *An explicit expression for $\tilde{J} = S_J S_J^\top$ is omitted from the statement of Theorem 1 but is given within its proof (see* (29) *together with* (19)*). In practice, one can compute $\tilde{J}$ according to its definition as a series, by summing $\operatorname{vec}(J^s)\operatorname{vec}(J^s)^\top$ over $s = 0,1,2,\dots$ until the changes in the elements of $\tilde{J}$ are below the machine precision. S_J is then obtained by means of the Cholesky decomposition of $\tilde{J}$.* □

The following result is an immediate consequence of Theorem 1.

Corollary 1 (Sampling from $x_{k,\infty}$)**.** *Under Assumptions 1 and 2, a sample $x_{k,\infty}^{(i)}$ from $x_{k,\infty}$ can be obtained as*

$$x_{k,\infty}^{(i)} = (I-A)^{-1}B\gamma + S_x(\vartheta)e^{(i)}, \tag{6}$$

where $S_x(\vartheta)$ is linear in ϑ and its expression is given by (5)*, and $e^{(i)} \in \mathbb{R}^{n_x^2 n_d}$ is extracted at random according to a $n_x^2 n_d$-variate standard Gaussian probability distribution.* □

Now that we are able to generate samples of the stationary state process *parametrically* in the optimization variables, we can formulate the scenario optimization program

$$\begin{aligned} \min_{\gamma\in\Gamma,\vartheta\in\Theta,h} \quad & h \\ \text{subject to:} \quad & \ell(x_{k,\infty}^{(i)},\gamma+\vartheta d_k^{(i)},d_k^{(i)})\le h \\ & f(x_{k,\infty}^{(i)},\gamma+\vartheta d_k^{(i)})\le 0 \\ & i=1,\dots,N, \end{aligned} \tag{7}$$

where $d_k^{(i)}$ and $x_{k,\infty}^{(i)}$, $i=1,\dots,N$, are samples of independent random variables and each $x_{k,\infty}^{(i)}$ is obtained according to (6) in Corollary 1. Given that $x_{k,\infty}^{(i)}$ depends linearly on γ and ϑ, and since ℓ and f are convex, (7) is a program with N convex constraints that can be solved via standard software.

Denote as $(\gamma^\star,\vartheta^\star,h^\star)$ the solution to the scenario program (7). Under the following technical assumption, we are able to prove the feasibility of $(\gamma^\star,\vartheta^\star,h^\star)$ for the original problem (4) (see Theorem 2), which entails also performance guarantees through the upper bound $h^\star$ on the cost function.

Assumption 4 (Existence and uniqueness)**.** *For any N, for any sample of the random quantities, the constrained optimization problem* (7) *has a unique solution.* □

Remark 2. *Assumption 4 is quite standard in scenario-based optimization (see, e.g., [18], [19], [29]). The uniqueness part of Assumption 4 can be relaxed by considering a suitable convex tie-break rule to single out a unique solution (see [18]). Also the existence part of Assumption 4 can be relaxed, see e.g. [30].* □

Theorem 2 (Feasibility guarantees)**.** *Choose a confidence parameter $\beta\in(0,1)$ and select N so as to satisfy*

$$\sum_{i=0}^{n}\binom{N}{i}\varepsilon^i(1-\varepsilon)^{N-i}\le\beta, \tag{8}$$

where n is the number of scalar optimization variables in the controller parametrization (γ,ϑ).
Then, if Assumptions 1-4 hold, the solution $(\gamma^\star,\vartheta^\star,h^\star)$ of the scenario program (7) *is feasible for the original chance-constrained problem* (4) *with probability larger than or equal to $1-\beta$, i.e.,*

$$\begin{aligned}\mathbb{P}^N_{\boldsymbol{d_k}}\{\mathbb{P}_{\boldsymbol{d_k}}\{&\ell(x^\star_{k,\infty},\gamma^\star+\vartheta^\star d_k,d_k)\le h^\star\\ &\wedge f(x^\star_{k,\infty},\gamma^\star+\vartheta^\star d_k)\le 0\}\ge 1-\varepsilon\}\ge 1-\beta,\end{aligned} \tag{9}$$

where $x^\star_{k,\infty}$ is the stationary Gaussian process with mean $(I-A)^{-1}B\gamma^\star$ and covariance matrix given by $S_x(\vartheta^\star)S_x(\vartheta^\star)^\top$ defined by (3) *with $\gamma=\gamma^\star$ and $\vartheta=\vartheta^\star$.*

Proof. See Appendix C. □

Remark 3 (Alternative approaches to chance-constraint optimization)**.** *If $\ell(\cdot)$ and $f(\cdot)$ belong to specific classes of functions, for example if they are both linear, then, given the Gaussian nature and complete characterization of the stationary process $x_{k,\infty}$ provided by Theorem 1, one can solve* (4) *by replacing the chance constraint with a tractable analytic approximation as suggested in [15]. Notably, the scenario solution adopted in this paper does not require $\ell(\cdot)$ and $f(\cdot)$ to have a special form but only to be convex.* □

IV. Extensions

The results of this paper encompass also other formulations than the one described in Section II. We discuss them next.

A. Average cost

Suppose that the average cost

$$J(\gamma,\vartheta)=\mathbb{E}_{\boldsymbol{d_k}}[\ell(x_{k,\infty},\gamma+\vartheta d_k,d_k)]$$

can be computed analytically as a function of $(\gamma,\vartheta)\in\Gamma\times\Theta$. This is the case, e.g., when $\ell(\cdot)$ is the classical quadratic cost in the state and control input variables.

Then, one can minimize $J(\gamma,\vartheta)$, instead of the upper bound h on $\ell(x_{k,\infty},\gamma+\vartheta d_k,d_k)$ over a set of probability $1-\varepsilon$ (see problem (4)), thus leading to

$$\begin{aligned}\min_{\gamma\in\Gamma,\vartheta\in\Theta}\quad & J(\gamma,\vartheta)\\ \text{subject to:}\quad & \mathbb{P}_{\boldsymbol{d_k}}\{f(x_{k,\infty},\gamma+\vartheta d_k)\le 0\}\ge 1-\varepsilon.\end{aligned} \tag{10}$$

Under Assumption 3, $J(\cdot)$ is also convex and, *mutatis mutandis*, the theory developed remains valid also in the average cost set-up. Specifically, Theorem 2 still holds with a slightly modified statement, where (4) is replaced with (10), (7) with the sampled counterpart of (10)

$$\begin{aligned}\min_{\gamma\in\Gamma,\vartheta\in\Theta}\quad & J(\gamma,\vartheta)\\ \text{subject to:}\quad & f(x^{(i)}_{k,\infty},\gamma+\vartheta d_k^{(i)})\le 0, i=1,\dots,N,\end{aligned}$$

and $\ell(\cdot)$ is removed from the probability in (9).

To see that $J(\cdot)$ is convex, start noticing that $x_{k,\infty}=\bar{x}(\gamma)+S_x(\vartheta)e_k$ and $\gamma+\vartheta d_k$ are affine maps of γ and ϑ. Then, let $\delta_k=(d_k,e_k)$ and $\tilde{\ell}(\gamma,\vartheta,\delta_k)=\ell(x_{k,\infty},\gamma+\vartheta d_k,d_k)$. Clearly, under Assumption 3, $\tilde{\ell}(\gamma,\vartheta,\delta_k)$ is convex in (γ,ϑ). Then

$$J(\gamma,\vartheta)=\mathbb{E}_{\delta_k}\left[\tilde{\ell}(\gamma,\vartheta,\delta_k)\right]=\int\tilde{\ell}(\gamma,\vartheta,\delta_k)\mathrm{d}\mathbb{P}_{\delta_k}$$

which can be shown to be convex by definition of convexity, as follows. Take $\alpha\in[0,1]$ and $(\gamma_1,\vartheta_1),(\gamma_2,\vartheta_2)\in\Gamma\times\Theta$. Then, we have that

$$\begin{aligned}&J(\alpha\gamma_1+(1-\alpha)\gamma_2,\alpha\vartheta_1+(1-\alpha)\vartheta_2)\\ &\quad=\int\tilde{\ell}(\alpha\gamma_1+(1-\alpha)\gamma_2,\alpha\vartheta_1+(1-\alpha)\vartheta_2,\delta_k)\mathrm{d}\mathbb{P}_{\delta_k}\\ &\quad\le\int\left(\alpha\tilde{\ell}(\gamma_1,\vartheta_1,\delta_k)+(1-\alpha)\tilde{\ell}(\gamma_2,\vartheta_2,\delta_k)\right)\mathrm{d}\mathbb{P}_{\delta_k}\\ &\quad=\alpha\int\tilde{\ell}(\gamma_1,\vartheta_1,\delta_k)\mathrm{d}\mathbb{P}_{\delta_k}+(1-\alpha)\int\tilde{\ell}(\gamma_2,\vartheta_2,\delta_k)\mathrm{d}\mathbb{P}_{\delta_k}\\ &\quad=\alpha J(\gamma_1,\vartheta_1)+(1-\alpha)J(\gamma_2,\vartheta_2),\end{aligned}$$

where the inequality is due to $\tilde{\ell}(\gamma,\vartheta,\delta_k)$ being convex and the monotonicity property of the integral with respect to the measure $\mathbb{P}_{\delta_k}$.

B. Non-asymptotically stable systems

Let us now consider the case in which A in (1) does not satisfy Assumption 1.

Suppose instead that the matrix pair (A,B) is stabilizable and the state is available for feedback. We can then first apply a state feedback stabilizing control law

$$u_k=Kx_k+\tilde{u}_k, \tag{11}$$

that makes the closed-loop system

$$\begin{aligned} x_{k+1} &= (A+BK)x_k + B\tilde{u}_k + Wd_k \\ &= \tilde{A}x_k + B\tilde{u}_k + Wd_k \end{aligned} \tag{12}$$

asymptotically stable. System (12) satisfies Assumption 1 with $\tilde{A}$ in place of A and fits the framework described in Section II, with $\tilde{u}_k$ as control input in place of u_k. We can thus readily apply the proposed methodology to devise an optimal disturbance compensator $\tilde{u}_k = \gamma^\star + \vartheta^\star d_k$ for (12) and then, by (11), feed the original system with the control law $u_k = Kx_k + \gamma^\star + \vartheta^\star d_k$.

The reader should note, however, that matrix K has to be chosen prior to solving (7). Indeed, the joint optimization of K with γ and ϑ would render the expression of $\ell(\cdot)$ and $f(\cdot)$ non-convex as a function of K, thus violating Assumption 3. If minimizing the cost function $\ell(\cdot)$ entails minimizing the variance of the steady-state distribution, then K can be set equal to the linear quadratic regulation (LQR) gain by neglecting the constraints or it can be tuned via the generalized minimum variance approach by accounting for the constraints within the cost function.

C. Compensator using noisy measurements of the disturbance

Suppose that only noisy measurements of the form

$$w_k = Cd_k + n_k$$

of the disturbance d_k are available, where n_k and d_k are jointly Gaussian. We shall show that the proposed approach is still applicable with minor modifications.

The compensator in (2) now takes the form

$$u_k = \gamma + \eta w_k. \tag{13}$$

Substituting (13) in (1) yields

$$\begin{aligned} x_{k+1} &= Ax_k + B\gamma + (B\eta C + W)d_k + B\eta n_k \\ &= Ax_k + B\gamma + \left(B\begin{bmatrix}\eta C & \eta\end{bmatrix} + \begin{bmatrix}W & 0\end{bmatrix}\right)\begin{bmatrix}d_k \\ n_k\end{bmatrix}, \end{aligned}$$

which still fits the structure of (3) setting $\vartheta = [\eta C\ \eta]$ with η being the compensator to design, using $[W\ 0]$ in place of W and $\tilde{d}_k = [d_k^\top\ n_k^\top]^\top$ in place of d_k. Similarly, we need to replace d_k and ϑ with, respectively, $\tilde{d}_k$ and $[\eta C\ \eta]$ in problems (4) and (7). If we then use η as decision variable in place of ϑ, the resulting scenario program (7) is convex in η and, as long as the extended disturbance $\tilde{d}_k$ satisfies Assumption 2, Theorem 2 is still valid. In particular, the joint statistical properties of d_k and n_k will be used in (7) to design the parameters $\gamma^\star$ and $\eta^\star$ of the disturbance compensator $u_k = \gamma^\star + \eta^\star w_k$, adopting the noisy measurements w_k of d_k. Since the same degree of freedom (i.e., the choice of η) is used to simultaneously counteract the effect of d_k and n_k, the effectiveness of the compensator will depend on the relative magnitude between d_k and n_k. An appropriate tuning will be automatically realized when solving program (7).

D. Stochastic periodic control

Let us now consider the case in which the disturbance process $\boldsymbol{d}$ is cyclostationary (rather than stationary), meaning that it has periodic statistical properties. More specifically, according to [31], $\boldsymbol{d}$ is cyclostationary if its mean and autocorrelation function are periodic with some period $T_h > 0$, i.e.

$$\begin{aligned} &\mathbb{E}[d_{s+T_h}] = \mathbb{E}[d_s], \\ &\mathbb{E}[d_{s+T_h} d_{s+T_h+\tau}^\top] = \mathbb{E}[d_s d_{s+\tau}^\top], \end{aligned}$$

for all $s, \tau \in \mathbb{Z}$. Similarly to Section II, we assume that $\boldsymbol{d}$ is a Gaussian process with zero mean, known autocorrelation function, and with independent random variables across different periods. That is, for all $s \in \mathbb{Z}$

$$\mathbb{E}[d_s] = 0, \tag{14}$$

and for all $k \in \mathbb{Z}$ and $t \in \{0, \dots, T_h - 1\}$

$$\mathbb{E}[d_{kT_h+t} d_{kT_h+\tau}^\top] = \begin{cases} \Sigma_d^{t,\tau} & \tau \in \{0, \dots, T_h - 1\} \\ 0 & \text{otherwise.} \end{cases} \tag{15}$$

The condition on the autocorrelation function is the counterpart in the cyclostationary realm of the independence part of Assumption 2. However, differently from Assumption 2, we allow for correlations between d_s and $d_{s'}$ as long as they belong to the same period. By setting $T_h = 1$, (14) and (15) are equivalent to Assumption 2.

Optimal stochastic periodic control of the system

$$x_{s+1} = Ax_s + Bu_s + Wd_s \tag{16}$$

can then be embedded in our framework by adopting the so-called lifting transformation, i.e., by unrolling the original system dynamics over a time window of length equal to the period T_h and referring to the system dynamics from one period k to the next one. To this end, let $X_k = x_{kT_h}$ denote the system state at the beginning of period k, $U_k = [u_{kT_h}^\top \cdots u_{(k+1)T_h-1}^\top]^\top$ the sequence of inputs during period k, and $D_k = [d_{kT_h}^\top \cdots d_{(k+1)T_h-1}^\top]^\top$ the sequence of disturbances affecting the system within period k. Then, iterating (16) starting from $s = kT_h$ up to $(k+1)T_h$, we obtain

$$\begin{aligned} X_{k+1} &= x_{(k+1)T_h} \\ &= A^{T_h} x_{kT_h} + \sum_{s=0}^{T_h-1} A^{T_h-1-s}(Bu_{kT_h+s} + Wd_{kT_h+s}) \\ &= \mathcal{A}X_k + \mathcal{B}U_k + \mathcal{W}D_k, \end{aligned} \tag{17}$$

for appropriately defined matrices $\mathcal{A}$, $\mathcal{B}$, and $\mathcal{W}$. Note that $\mathcal{A} = A^{T_h}$, which inherits the asymptotic stability properties of A. If d_s is Gaussian with zero mean (cf. (14)), also D_k is Gaussian with zero mean. Moreover, by (15), we know that $\mathbb{E}[D_k D_\kappa^\top] = 0$ for all $k, \kappa \in \mathbb{Z}$ with $k \neq \kappa$, meaning that D_k satisfies Assumption 2 with $\Sigma_d = \mathbb{E}[D_k D_k^\top]$, which has $\Sigma_d^{t,\tau}$ in (15) as sub-matrices. System in (17) then fits the structure of (1) and Assumptions 1 and 2 of Section II are satisfied, so that our approach can be readily applied on the lifted system (17).

Note that the compensator parameter matrix ϑ for (17) needs to be restricted to have a lower block-diagonal structure to avoid u_{kT_h+s} being dependent on future values $d_{kT_h+s+1}, \dots, d_{(k+1)T_h-1}$ of the disturbances, for $s = 0, \dots, T_h - 2$. This requirement can be easily accounted for via an additional constraint on the set Θ.

An example of a system affected by a cyclostationary disturbance is given in the following section.

V. Application to building energy management

In this section, we present an energy management application example with the twofold purpose of better motivating the considered disturbance compensation set-up and showing the efficacy of the proposed scenario-based design methodology.

Consider an office building consisting of a single thermally controlled zone. Our goal is to set the zone temperature T_z along a 24-hours time horizon discretized into $T_h = 144$ time slots of $\Delta_T = 10$ minutes each so as to compensate the outdoor temperature T_o, longwave solar radiation LW, and shortwave solar radiation SW, which act as disturbances on the thermal evolution of the building. The temperature T_z is assumed to be regulated by a low-level controller so as to perfectly track in one time slot a desired temperature set-point, which represents the control input in our case study. The cost to be minimized over the one-day time horizon is represented by the electrical energy consumption E_ℓ of the chiller plant that provides the cooling energy request E_c needed to make the zone temperature tracking the desired set-point. Constraints are introduced to represent the chiller actuation limits and to ensure a certain comfort level to the occupants of the building during working hours.

The electric energy consumption E_ℓ depends on the cooling energy request E_c according to a nonlinear convex function $E_\ell = \mathrm{ch}(E_c)$, derived from the Ng-Gordon model of the chiller plant, [32] (see Eq. (21) in [12] with E_c in place of $E_{ch,c}$ and with $T_o = 22°\mathrm{C}$ for simplicity), which can be replaced by a piece-wise affine approximation for computational purposes (cf. [12, Eq. (23)]).

In turn, the cooling energy E_c is the sum of three contributions:

$$E_c = E_w + E_z + E_o, \tag{18}$$

where E_w is the energy exchanged between the zone and the building walls, E_z is associated with the zone thermal inertia, and E_o is related to other thermal phenomena such as occupancy, solar radiation through windows, equipment, lighting, etc.

The first term E_w depends on the thermal behavior of the building described by the temperatures of the layers composing its walls, which together with the zone temperature $T_z(s\Delta_T)$ constitute the state vector x_s. Given the assumption that the actuator is able to perfectly track the reference set-point u, we have $T_z((s+1)\Delta_T) = u_s$, and altogether x_s evolves according to a linear difference equation of the form (16), where d_s is a scalar cyclostationary disturbance with a one-day period that represents the joint contribution of $T_o(s\Delta_T)$, $LW(s\Delta_T)$, and $SW(s\Delta_T)$. The second term E_z in (18) depends affinely on the control input, and the third term E_o is affine in the disturbances.

Interestingly, in our approach we do not need any measurement of the state x_s – that would be hardly available – and exploit instead measurements of d_s for compensation purposes to save electrical energy. Since d_s is cyclostationary with one-day period, we can use the theory developed in Section IV-D, setting $T_h = 144$ and applying the lifting transformation with the following definitions: $X_k = x_{kT_h}$ as the system state at the beginning of the k-th day (time 0 a.m), $U_k = [u_{kT_h} \cdots u_{(k+1)T_h-1}]^\top$ as the sequence of inputs during day k, and $D_k = [d_{kT_h} \cdots d_{(k+1)T_h-1}]^\top$ as the sequence of disturbances affecting the system within day k. By means of principal component analysis of historical data, D_k is described as the weighted sum of a finite number of one-day profiles through coefficients that are modeled as uncorrelated Gaussian variables so that $\mathbb{E}[D_k D_k^\top] = \Sigma_d$ and $\mathbb{E}[D_k D_\kappa^\top] = 0$ for all $k \neq \kappa$. Note that D_k has not zero mean, but our design methodology can still be applied as explained after Assumption 2.

Following [12, Sections 2.1.1-2.1.5], all terms contributing to E_c are affine functions of X_k, U_k, and D_k, thus yielding

$$E_{c,k} = F_c X_k + G_c U_k + H_c D_k + L_c,$$

where $E_{c,k} = [E_{c,kT_h+1} \cdots E_{c,(k+1)T_h}]^\top$ is the cooling energy request within period k. The electrical energy consumption can be thus easily computed as

$$E_{\ell,k} = \mathrm{ch}(E_{c,k}),$$

where the function $\mathrm{ch}(\cdot)$ is meant to be applied component-wise to the $E_{c,k}$ vector.

The cost function $\ell(\cdot)$ in (4) is thus given by the steady-state daily electrical energy consumption

$$\begin{aligned} \ell(\cdot) &= \mathbb{1}^\top E_{\ell,k,\infty} \\ &= \mathbb{1}^\top \mathrm{ch}(E_{c,k,\infty}) \\ &= \mathbb{1}^\top \mathrm{ch}(F_c X_{k,\infty} + G_c U_k + H_c D_k + L_c), \end{aligned}$$

with $\mathbb{1} = [1 \ \cdots \ 1]^\top$ and $U_k = \gamma + \vartheta D_k$, while the constraint $f(\cdot) \leq 0$ embeds all comfort and actuation constraints evaluated every Δ_T minutes.

In particular, we impose

$$\begin{aligned} &21 \leq [U_{k,\infty}]_i \leq 24 && \forall i \in \{48, \dots, 108\}, \\ &15 \leq [U_{k,\infty}]_i \leq 30 && \forall i \notin \{48, \dots, 108\}, \\ &0 \leq [E_{c,k,\infty}]_i \leq E_c^{\max}, && \forall i \end{aligned}$$

that is, the zone temperature set-point is constrained to be between 21°C and 24°C during the working hours (from 8 a.m. to 6 p.m.), and between 15°C and 30°C otherwise, and the chiller is required not to produce heating energy ($E_c < 0$) but cooling energy only, and not to exceed its maximum capacity $E_c^{\max} = 30$ MJ.

Similarly to [33], in order to reduce the number of parameters of the disturbance compensator $U_k = \gamma + \vartheta D_k$, only a subset of the temperature set-point collected in U_k – those components $[U_k]_{j18}$, $j = 1, \dots, T_h/18$, corresponding to the set-point values sampled every $18\Delta_T$ minutes (i.e.

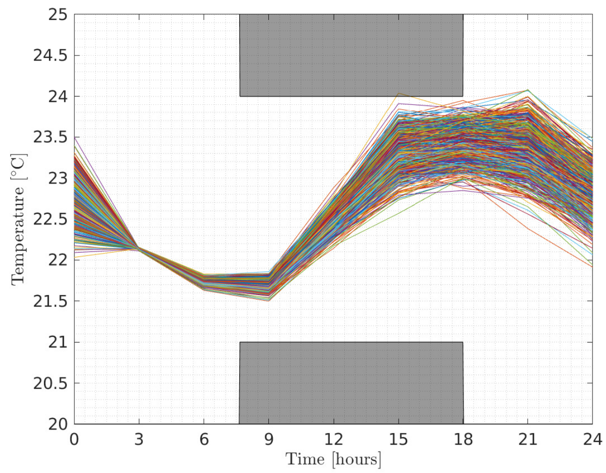


Fig. 1. Temperature set-point obtained with the optimal compensator at day $k = 100$ for 1000 realizations of the disturbance $\boldsymbol{d_k}$. The value at time zero is the set-point at the end of the previous day. Grey areas denote non-comfort zones.

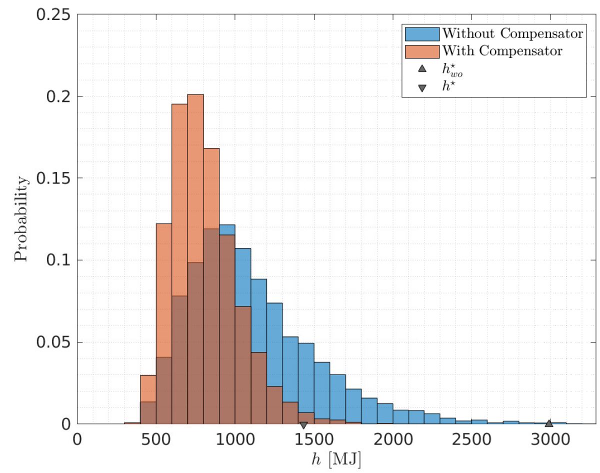


Fig. 2. Histograms of the upper bound h (with compensator) and h_{wo} (without compensator) on the daily electrical energy consumption at day $k = 100$, obtained through N_v realizations of the disturbance $\boldsymbol{d_k}$.

every three hours) – are chosen as a function of D_k, while the intermediate ones are obtained by linear interpolation. Additionally, each value $[U_k]_{j18}$, $j = 1, \dots, T_h/18$ is set as a function of the average of the scalar disturbance d_s computed over the previous time slot of duration $18\Delta_T$. Formally, for all $j = 1, \dots, T_h/18$,

$$[U_k]_{j18} = \varrho_j + \theta_j \frac{1}{18} \sum_{i=18(j-1)}^{18j-1} [D_k]_i$$

$$[U_k]_{j18+i} = \tfrac{18-i}{18}[U_k]_{j18} + \tfrac{i}{18}[U_k]_{(j+1)18} \quad i = 1, \dots, 17.$$

All these requirements on U_k translates into deterministic constraints on γ and ϑ

$$\begin{aligned}
&[\gamma]_{j18} = \varrho_j \\
&[\vartheta]_{j18,s} = \tfrac{1}{18}\theta_j && s = 18(j-1), \dots, 18j-1 \\
&[\vartheta]_{j18,s} = 0 && s \neq 18(j-1), \dots, 18j-1 \\
&[\gamma]_{j18+i} = \tfrac{18-i}{18}[\gamma]_{j18} + \tfrac{i}{18}[\gamma]_{(j+1)18} && i = 1, \dots, 17 \\
&[\vartheta]_{j18+i,s} = \tfrac{18-i}{18}[\vartheta]_{j18,s} + \tfrac{i}{18}[\vartheta]_{(j+1)18,s} && i = 1, \dots, 17, \forall s
\end{aligned}$$

for all $j = 1, \dots, T_h/18$, which parametrize γ and ϑ through the decision variables ϱ_j and θ_j, $j = 1, \dots, T_h/18$.

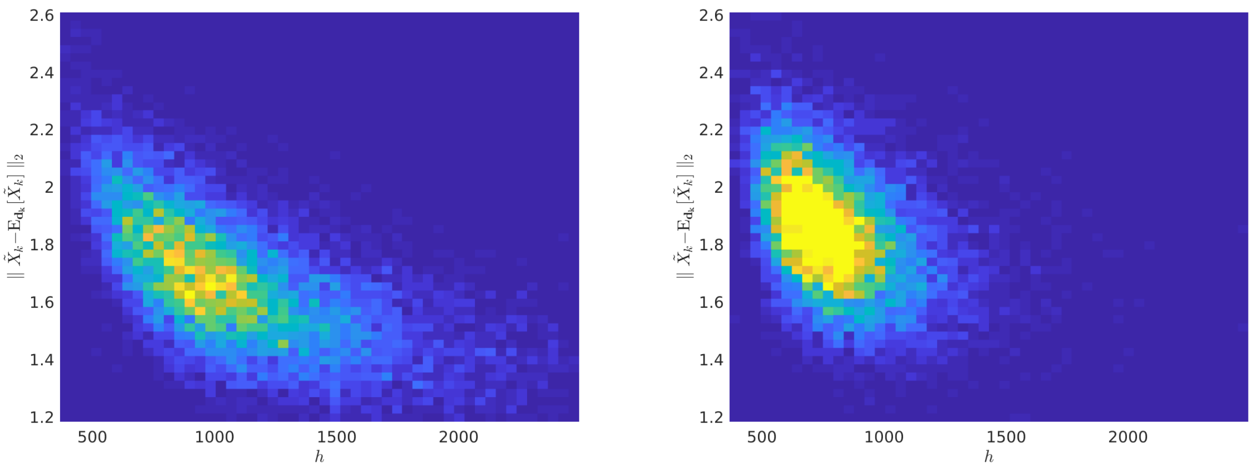
The total number of decision variables is then $n = 2 \cdot T_h/18 = 16$, and, setting a violation $\varepsilon = 0.02$ with a confidence parameter $\beta = 10^{-3}$, we get $N = 1555$ from (8). Solving the scenario program (7), we obtain a disturbance compensator yielding a daily electrical consumption smaller than or equal to $h^\star = 1435$ MJ for all disturbance realizations excepts for a set with probability smaller than or equal to 0.02 with confidence larger than or equal to $1 - 10^{-3}$.

To validate the result, we performed $N_v = 10^4$ simulations of 100 days each, to make the transient behavior vanish and the state reach its stationary distribution. To this purpose, we extracted $100N_v$ one-day realizations of D_k, different from the ones used to solve (7). In each simulation we set the initial thermal state of the building at its expected value in stationary conditions $\bar{X} = (I - A)^{-1}B\gamma^\star$.

Figure 1 shows the zone temperature set-point U_k at day $k = 100$ for 1000 of the N_v simulations. Grey areas denotes non-comfort zones and, as can be seen from the picture, most temperature set-point profiles stay within the comfort limits.

To show the benefits of introducing the disturbance compensator, we also solved the same problem forcing $\vartheta = 0$ in (2), which resulted in an optimal daily electrical consumption of $h^\star_{wo} = 2990$ MJ without (wo) compensation. Note that in this case the total number of decision variable is $n = T_h/18 = 8$ and, to get the same violation of $\varepsilon = 0.02$ with the same confidence parameter $\beta = 10^{-3}$, according to (8), only $N = 975$ scenarios are needed for the solution of (7). Remarkably, by comparing the values of $h^\star$ and $h^\star_{wo}$, it appears that introducing the compensator, we can reduce the daily energy consumption of an amount that is more than a half (precisely, 52%) in the worst case. To strengthen this result we also computed the value of the daily electrical energy consumption h_{wo} at day $k = 100$ for N_v realizations for the disturbances, when input U_k designed without the compensator was applied, and then, we used the same N_v realizations to determine the daily electrical energy consumption h at day $k = 100$ with the compensator. For the case without compensator we set the initial state of all realization equal to $\bar{X}_{wo} = (I - A)^{-1}B\gamma^\star_{wo}$. Figure 2 compares the histograms of the values obtained for h and h_{wo} for the N_v realizations. As can be seen from the picture, most of the mass of the two histograms is on the left of the corresponding optimal values for h (grey triangles), $h^\star$ and $h^\star_{wo}$, respectively, with an empirical violation that is 0.0127 and 0.0029, respectively. In both cases the empirical violation is below our choice of ε, thus validating the guarantees provided by Theorem 2. Furthermore, we can clearly see the benefits of introducing the compensator since the mass of the histogram related to the case with the compensator is shifted towards much smaller values of the energy consumption with respect to the one without compensator.

In Figures 3a and 3b, we also report a bi-variate histogram



(a) without compensator

(b) with compensator

Fig. 3. Histogram of h and $\|\tilde{X}_k - \mathbb{E}_{\boldsymbol{d_k}}[\tilde{X}_k]\|_2$ at day $k = 100$ for N_v realizations of the disturbance $\boldsymbol{d_k}$. Bright colors indicate most probable values, darker colors indicate less likely realizations.

of the two quantities h and $\|\tilde{X}_k - \mathbb{E}_{\boldsymbol{d_k}}[\tilde{X}_k]\|_2$, where the variable $\tilde{X}_k$ is the part of the state vector X_k representing the temperatures of the building walls, without and with the compensator, respectively, obtained based on the N_v realizations. The quantity $\|\tilde{X}_k - \mathbb{E}_{\boldsymbol{d_k}}[\tilde{X}_k]\|_2$ measures the deviation of the temperatures of the building walls with respect to their expected value. As can be seen from the two pictures, the introduction of the compensator reduces the dispersion of the daily electrical energy consumption h (horizontal axis) by moving part of the dispersion induced by the disturbances to the building structure, since the building walls temperatures appear to be more spread (vertical axis).

VI. Conclusions

In this paper we addressed the optimal design of a disturbance compensator for the steady-state operation of a discrete time linear systems affected by white Gaussian noise. The design problem was formalized as a chance-constraint program for the system operating in stationary conditions so that, in the long run, the designed compensator exhibits optimal performance and satisfies the imposed probabilistic constraints. The chance-constraint optimization problem is solved (offline and only once!) by means of an exact characterization of the stationary state distribution as a linear function of the compensator parameters and the use of randomization to deal with the probabilistic constraints. We also discussed the extension of the proposed methodology to the case of average cost function and stochastic periodic compensation for a system affected by cyclostationary Gaussian disturbances, the latter extension being of interest in the energy management application example. The proposed approach is promising and competitive in terms of ease of computation, applicability, and performance guarantees with respect to alternative approaches to optimal constrained control, especially when state measurements are not available for feedback.

Acknowledgment

The authors would like to thank the reviewers for their comments, which helped in the development of the results in Section IV.

Appendix A
A summability result

Lemma 1. *For any $m_1, m_2 \in \mathbb{N}$ and $z_1, z_2 \in \mathbb{C}$ such that $|z_1 z_2| < 1$ we have*

$$\sum_{s=0}^{\infty} \binom{s}{m_1}\binom{s}{m_2} z_1^{s-m_1} z_2^{s-m_2} = \sum_{r=0}^{\underline{m}} \binom{m_1+m_2-r}{r, m_1-r, m_2-r} \frac{z_1^{m_2-r} z_2^{m_1-r}}{(1-z_1 z_2)^{m_1+m_2+1-r}}, \quad (19)$$

where $\underline{m} = \min\{m_1, m_2\}$ and $\binom{n}{k_1,k_2,k_3}$ is the multinomial coefficient.

Proof. Using the fact that

$$\binom{s}{m_1}\binom{s}{m_2} = \sum_{r=0}^{\underline{m}} \binom{m_1+m_2-r}{r, m_1-r, m_2-r}\binom{s}{m_1+m_2-r},$$

we have

$$\begin{aligned}
&\sum_{s=0}^{M} \binom{s}{m_1}\binom{s}{m_2} z_1^{s-m_1} z_2^{s-m_2} \\
&= \sum_{s=0}^{M}\sum_{r=0}^{\underline{m}} \binom{m_1+m_2-r}{r, m_1-r, m_2-r}\binom{s}{m_1+m_2-r} z_1^{s-m_1} z_2^{s-m_2} \\
&= \sum_{r=0}^{\underline{m}} \binom{m_1+m_2-r}{r, m_1-r, m_2-r} z_1^{-m_1} z_2^{-m_2} \\
&\qquad \sum_{s=0}^{M} \binom{s}{m_1+m_2-r} (z_1 z_2)^s, \quad (20)
\end{aligned}$$

where the second equality is obtained by exchanging the two summations first and then collecting all quantities not depending on s outside the inner sum. By taking the limit as $M \to \infty$ on both sides of (20) and using the fact that

$$\sum_{s=0}^{\infty} \binom{s}{p} z^s = \frac{z^p}{(1-z)^{p+1}}$$

for any complex number z with $|z| < 1$ and any $p \in \mathbb{N}$ (see [34, eq. (5.57), pag. 199]), (19) follows, thus concluding the proof. $\square$

Appendix B
Proof of Theorem 1

Given that $\boldsymbol{x_\infty}$ satisfies (3), it is easily seen that $\mathbb{E}_{\boldsymbol{d_k}}[x_{k,\infty}] = \bar{x} = (I - A)^{-1} B\gamma$, while the covariance matrix Σ_x, under Assumption 1, is the unique solution of the Lyapunov equation

$$P = APA^\top + (B\vartheta + W)\Sigma_d(B\vartheta + W)^\top, \tag{21}$$

and is given by

$$\begin{aligned}\Sigma_x &= \mathbb{E}_{\boldsymbol{d_k}}\left[(x_{k,\infty} - \bar{x})(x_{k,\infty} - \bar{x})^\top\right] \\ &= \lim_{M\to\infty} \underbrace{\sum_{s=0}^{M} A^s(B\vartheta + W)S_d S_d^\top (B\vartheta + W)^\top A^{s\top}}_{P_M}. \end{aligned} \tag{22}$$

Let $A = TJT^{-1}$ where J is the Jordan normal form of A and T is a suitable nonsingular transformation matrix. By the Jordan form of A we can compute $A^s = TJ^sT^{-1}$. If we define $\Omega = T^{-1}(B\vartheta + W)S_d$, then, P_M can be compactly rewritten as

$$P_M = T\left(\sum_{s=0}^{M} J^s\Omega\Omega^\top J^{s\top}\right)T^\top. \tag{23}$$

By using the identity

$$J^s\Omega = \underbrace{(\operatorname{vec}(\Omega)' \otimes I_{n_x})}_{\bar{\Omega}}\underbrace{(I_{n_d} \otimes \operatorname{vec}(J^s))}_{\bar{J}_s}, \tag{24}$$

the right hand side of (23) can be rewritten as

$$\begin{aligned} T\left(\sum_{s=0}^{M} J^s\Omega\Omega^\top J^{s\top}\right)T^\top &= T\left(\sum_{s=0}^{M} \bar{\Omega}\bar{J}_s\bar{J}_s^\top\bar{\Omega}^\top\right)T^\top \\ &= T\bar{\Omega}\left(\sum_{s=0}^{M} \bar{J}_s\bar{J}_s^\top\right)\bar{\Omega}^\top T^\top. \end{aligned} \tag{25}$$

Moreover,

$$\begin{aligned} \sum_{s=0}^{M} \bar{J}_s\bar{J}_s^\top &= \sum_{s=0}^{M} (I_{n_d} \otimes \operatorname{vec}(J^s))(I_{n_d} \otimes \operatorname{vec}(J^s))^\top \\ &= \sum_{s=0}^{M} (I_{n_d} \otimes \operatorname{vec}(J^s))(I_{n_d} \otimes \operatorname{vec}(J^s)^\top) \\ &= \sum_{s=0}^{M} (I_{n_d} \otimes \operatorname{vec}(J^s)\operatorname{vec}(J^s)^\top) \\ &= \left(I_{n_d} \otimes \sum_{s=0}^{M} \operatorname{vec}(J^s)\operatorname{vec}(J^s)^\top\right), \end{aligned} \tag{26}$$

where the first equality is due to the definition of $\bar{J}_s$ in (24), the second equality to the distributive property of the conjugate-transpose operator over the Kronecker product, the third equality is given by the mixed product property of the Kronecker product, and the fourth equality is given by the distributive property of addition over the Kronecker product.

Combining (23), (25), and (26), we have

$$P_M = T\bar{\Omega}\left(I_{n_d} \otimes \sum_{s=0}^{M} \operatorname{vec}(J^s)\operatorname{vec}(J^s)^\top\right)\bar{\Omega}^\top T^\top. \tag{27}$$

As indicated by (22), it remains to compute the limit on both sides of (27) for $M \to \infty$ to get the variance of $x_{k,\infty}$.

The ℓ-th element of $\operatorname{vec}(J^s)$ can be expressed as $[\operatorname{vec}(J^s)]_\ell = [J^s]_{rc}$ with $c = \lceil \ell/n_x \rceil$ and $r = \operatorname{mod}(\ell - 1, n_x) + 1$. The Jordan matrix J is known to have a block-diagonal structure $J = \operatorname{blkdiag}(J_1, \ldots, J_{n_b})$, whose b-th block J_b of order m_b is associated to the b-th eigenvalue λ_b of A.

Since $J^s = \operatorname{blkdiag}(J_1^s, \ldots, J_{n_b}^s)$, each element (r, c) of J^s can be defined as follows. Let

$$\begin{aligned} b_r &= \min\left\{\bar{b} \in \mathbb{N} \mid \textstyle\sum_{p=1}^{\bar{b}} m_p \geq r\right\} \\ b_c &= \min\left\{\bar{b} \in \mathbb{N} \mid \textstyle\sum_{p=1}^{\bar{b}} m_p \geq c\right\} \end{aligned}$$

be the row and column index of the block of J^s containing the element (r, c). Due to the block structure of J^s, if $b_r = b_c = b$, then $[J^s]_{rc}$ belongs to the b-th Jordan block, otherwise $[J^s]_{rc}$ is outside the block diagonal and is therefore equal to zero. Accordingly,

$$[J^s]_{rc} = \begin{cases} [J_b^s]_{ij} & b = b_r = b_c \\ 0 & \text{otherwise} \end{cases} \quad \text{with} \quad \begin{cases} i = r - \sum_{p=1}^{b-1} m_p \\ j = c - \sum_{p=1}^{b-1} m_p \end{cases}$$

where i and j denote respectively the row and column index of an element within the b-th Jordan block J_b^s, and can take a value between 1 and m_b. Recall that, for a Jordan block J_b of order m_b, the (i, j)-th element of J_b^s is given by

$$[J_b^s]_{ij} = \begin{cases} \binom{s}{j-i}\lambda_b^{s-(j-i)} & 0 \leq j - i \leq s \\ 0 & \text{otherwise,} \end{cases} \tag{28}$$

with $1 \leq i \leq m_b$ and $1 \leq j \leq m_b$. Thus, we can finally express the (ℓ_1, ℓ_2) element of $\operatorname{vec}(J^s)\operatorname{vec}(J^s)^\top$ as

$$\begin{aligned} &[\operatorname{vec}(J^s)\operatorname{vec}(J^s)^\top]_{\ell_1\ell_2} \\ &\qquad = \begin{cases} [J_{b_1}^s]_{i_1j_1}\overline{[J_{b_2}^s]_{i_2j_2}} & \begin{array}{l} b_1 = b_{r,1} = b_{c,1} \\ b_2 = b_{r,2} = b_{c,2} \end{array} \\ 0 & \text{otherwise,} \end{cases} \end{aligned}$$

where

$$\begin{aligned} b_{r,q} &= \min\left\{\bar{b} \in \mathbb{N} \mid \textstyle\sum_{p=1}^{\bar{b}} m_p \geq \operatorname{mod}(\ell_q - 1, n_x) + 1\right\}, \\ b_{c,q} &= \min\left\{\bar{b} \in \mathbb{N} \mid \textstyle\sum_{p=1}^{\bar{b}} m_p \geq \lceil \ell_q/n_x \rceil\right\}, \\ i_q &= \operatorname{mod}(\ell_q - 1, n_x) + 1 - \textstyle\sum_{p=1}^{b_q-1} m_p, \\ j_q &= \lceil \ell_q/n_x \rceil - \textstyle\sum_{p=1}^{b_q-1} m_p, \end{aligned}$$

$q = 1, 2$ and, according to (28),

$$[J_{b_1}^s]_{i_1 j_1}\overline{[J_{b_2}^s]_{i_2 j_2}} = \begin{cases} \binom{s}{j_1-i_1}\binom{s}{j_2-i_2}\lambda_{b_1}^{s-(j_1-i_1)}\bar{\lambda}_{b_2}^{s-(j_2-i_2)} & \begin{array}{l}0 \le j_1 - i_1 \le s\\ 0 \le j_2 - i_2 \le s\end{array} \\ 0 & \text{otherwise,}\end{cases}$$

$\bar{\lambda}_{b_2}$ being the complex conjugate of λ_{b_2}.

Letting $\tilde{J} = \sum_{s=0}^{\infty} \operatorname{vec}(J^s)\operatorname{vec}(J^s)^\top$ and setting $m_1 = j_1 - i_1$ and $m_2 = j_2 - i_2$, one has

$$\begin{aligned}
[\tilde{J}]_{\ell_1\ell_2} &= \sum_{s=0}^{\infty}[\operatorname{vec}(J^s)\operatorname{vec}(J^s)^\top]_{\ell_1\ell_2} \\
&= \sum_{s=0}^{\infty}\begin{cases}\binom{s}{m_1}\binom{s}{m_2}\lambda_{b_1}^{s-m_1}\bar{\lambda}_{b_2}^{s-m_2} & \begin{array}{l} b_1 = b_{r,1} = b_{c,1}\\ b_2 = b_{r,2} = b_{c,2}\\ 0 \le m_1, m_2 \le s\end{array}\\ 0 & \text{otherwise}\end{cases} \\
&= \begin{cases}\displaystyle\sum_{s=0}^{\infty}\binom{s}{m_1}\binom{s}{m_2}\lambda_{b_1}^{s-m_1}\bar{\lambda}_{b_2}^{s-m_2} & \begin{array}{l} b_1 = b_{r,1} = b_{c,1}\\ b_2 = b_{r,2} = b_{c,2}\\ m_1, m_2 \ge 0\end{array}\\ 0 & \text{otherwise,}\end{cases}
\end{aligned} \tag{29}$$

with m_1 and m_2 depending on ℓ_1 and ℓ_2 through (i_1, j_1) and (i_2, j_2) respectively.

Using Lemma 1 in (29) with $z_1 = \lambda_{b_1}$, $z_2 = \bar{\lambda}_{b_2}$, and $\underline{m} = \min\{m_1, m_2\}$, we obtain

$$[\tilde{J}]_{\ell_1\ell_2} = \sum_{r=0}^{\underline{m}}\binom{m_1+m_2-r}{r, m_1-r, m_2-r}\frac{\lambda_{b_1}^{m_2-r}\bar{\lambda}_{b_2}^{m_1-r}}{(1-\lambda_{b_1}\bar{\lambda}_{b_2})^{m_1+m_2+1-r}}$$

if $b_1 = b_{r,1} = b_{c,1}$, $b_2 = b_{r,2} = b_{c,2}$, and $m_1, m_2 \ge 0$; or $[\tilde{J}]_{\ell_1\ell_2} = 0$ otherwise.

Since the quantities appearing in the expression of $[\tilde{J}]_{\ell_1\ell_2}$ are all finite, we can conclude that $\tilde{J} = \sum_{s=0}^{\infty}\operatorname{vec}(J^s)\operatorname{vec}(J^s)^\top < \infty$ and its (ℓ_1, ℓ_2) element is given by $[\tilde{J}]_{\ell_1\ell_2}$. By taking the limit for $M \to \infty$ on both sides of (27) and recalling (22), we obtain

$$\Sigma_x = T\bar{\Omega}(\vartheta)(I_{n_d}\otimes\tilde{J})\bar{\Omega}(\vartheta)^\top T^\top.$$

We now show that $S_x(\vartheta)S_x(\vartheta)^\top = \Sigma_x$ where $S_x(\vartheta)$ is given by (5):

$$\begin{aligned}
S_x(\vartheta)S_x(\vartheta)^\top &= T\bar{\Omega}(\vartheta)(I_{n_d}\otimes S_J)(T\bar{\Omega}(\vartheta)(I_{n_d}\otimes S_J))^\top \\
&= T\bar{\Omega}(\vartheta)(I_{n_d}\otimes S_J)(I_{n_d}\otimes S_J)^\top\bar{\Omega}(\vartheta)^\top T^\top \\
&= T\bar{\Omega}(\vartheta)(I_{n_d}\otimes S_J)(I_{n_d}\otimes S_J^\top)\bar{\Omega}(\vartheta)^\top T^\top \\
&= T\bar{\Omega}(\vartheta)(I_{n_d}\otimes S_J S_J^\top)\bar{\Omega}(\vartheta)^\top T^\top \\
&= T\bar{\Omega}(\vartheta)(I_{n_d}\otimes\tilde{J})\bar{\Omega}(\vartheta)^\top T^\top \\
&= \Sigma_x.
\end{aligned}$$

This concludes the proof. □

Appendix C
Proof of Theorem 2

Under Assumptions 1 and 2, by Theorem 1 and Corollary 1, $x_{k,\infty}^{(i)}$ in problem (7) is extracted at random according to the distribution of $x_{k,\infty}$, so that (7) is the sampled version of (4). Considering also Assumptions 3 and 4, the result in Theorem 2 then follows immediately from the standard scenario theory, see [18, Theorem 1]. □

References

[1] K. Åström, *Introduction to Stochastic Control Theory*, ser. Mathematics in Science and Engineering. A Series of Monographs and Textbooks. Academic Press, 1970.

[2] D. Clarke and R. Hastings-James, "Design of digital controllers for randomly disturbed systems," *Proc. IEE*, vol. 118, no. 10, pp. 1503–1506, 1971.

[3] V. Peterka, "On steady state minimum variance control strategy," *Kybernetika*, vol. 8, no. 3, pp. 219–232, 1972.

[4] U. Shaked and P. Kumar, "Minimum variance control of discrete time multivariable armax systems," *SIAM Journal on Control and Optimization*, vol. 24, no. 3, pp. 396–411, 1986.

[5] M. Grimble, "Generalized minimum variance control law revisited," *Optimal Control Applications and Methods*, vol. 9, no. 1, pp. 63–77, 1988.

[6] B. Huang, "Minimum variance control and performance assessment of time-variant processes," *Journal of Process Control*, vol. 12, no. 6, pp. 707–719, 2002.

[7] P. J. Gawthrop, "Minimum-variance control," in *Encyclopedia of Life Support Systems (EOLSS)*. UNESCO, 2004.

[8] A. Sinha, *Linear systems: optimal and robust control*. CRC press, 2007.

[9] M. Grimble, "Controller performance benchmarking and tuning using generalised minimum variance control," *Automatica*, vol. 38, no. 12, pp. 2111 – 2119, 2002.

[10] D. Ioli, A. Falsone, and M. Prandini, "Optimal energy management of a building cooling system with thermal storage: a convex formulation," in *9th IFAC Symposium on Advanced Control of Chemical Processes (ADCHEM 2015), Whistler, British Columbia, Canada*, 2015.

[11] D. Ioli, A. Falsone, S. Schuler, and M. Prandini, "A compositional framework for energy management of a smart grid: a scalable stochastic hybrid model for cooling of a district network," in *12th IEEE International Conference on Control and Automation (ICCA 2016), Kathmandu, Nepal*, 2016.

[12] D. Ioli, A. Falsone, A. Papadopoulos, and M. Prandini, "A compositional modeling framework for the optimal energy management of a district network," *Journal of Process Control*, vol. 74, pp. 160 – 176, 2019.

[13] A. Prèkopa, *Stochastic programming*. Kluwer, 1995.

[14] Z. Zhou and R. Cogill, "Reliable approximations of probability-constrained stochastic linear-quadratic control," *Automatica*, vol. 49, no. 8, pp. 2435 – 2439, 2013.

[15] E. Cinquemani, M. Agarwal, D. Chatterjee, and J. Lygeros, "Convexity and convex approximations of discrete-time stochastic control problems with constraints," *Automatica*, vol. 47, no. 9, pp. 2082–2087, 2011.

[16] D. Bertsimas and D. Brown, "Constrained stochastic LQC: A tractable approach," *IEEE Transactions on Automatic Control*, vol. 52, no. 10, pp. 1826–1841, 2007.

[17] M. Campi, S. Garatti, and M. Prandini, "The scenario approach for systems and control design," *Annual Reviews in Control*, vol. 33, no. 2, pp. 149 – 157, 2009.

[18] M. 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.

[19] 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.

[20] T. Alamo, R. Tempo, A. Luque, and D. R. Ramirez, "Randomized methods for design of uncertain systems: Sample complexity and sequential algorithms," *Automatica*, vol. 52, pp. 160 – 172, 2015.

[21] G. Calafiore and M. Campi, "Uncertain convex programs: randomized solutions and confidence levels," *Mathematical Programming*, vol. 102, no. 1, pp. 25–46, 2005.

[22] ——, "The scenario approach to robust control design," *IEEE Transactions on Automatic Control*, vol. 51, no. 5, pp. 742–753, 2006.

[23] A. Falsone, L. Deori, D. Ioli, S. Garatti, and M. Prandini, "Optimal disturbance compensation for constrained linear systems operating in stationary conditions: a scenario-based approach," *Automatica*, vol. 110, 2019.

[24] ——, "Optimally shaping the stationary distribution of a constrained discrete time stochastic linear system via disturbance compensation," in *56th IEEE Conference on Decision and Control (CDC)*, Dec 2017, pp. 629–634.

[25] 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.

[26] P. E. Caines, *Linear stochastic systems*. Wiley New York, 1988.

[27] A. N. Shiryaev, *Probability*. Springer, 1996.

[28] D. Dentcheva, "Optimization models with probabilistic constraints," in *Probabilistic and Randomized Methods for Design under Uncertainty*, G. Calafiore and F. Dabbene, Eds. London: Springer-Verlag, 2006.

[29] A. Caré, S. Garatti, and M. Campi, "Scenario min-max optimization and the risk of empirical costs," *SIAM Journal on Optimization*, vol. 25, no. 4, pp. 2061–2080, 2015.

[30] G. C. Calafiore, "Random convex programs: Dealing with the non-existing solution nuisance," in *49th IEEE Conference on Decision and Control (CDC)*, 2010, pp. 1939–1944.

[31] W. A. Gardner, A. Napolitano, and L. Paura, "Cyclostationarity: Half a century of research," *Signal processing*, vol. 86, no. 4, pp. 639–697, 2006.

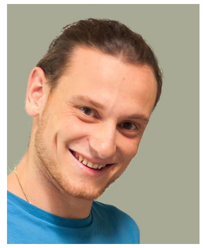
[32] J. Gordon and K. Ng, *Cool Thermodynamics: The Engineering and Physics of Predictive, Diagnostic and Optimization Methods for Cooling Systems*. Cambridge International Science Publishing, 2000.

[33] D. Ioli, A. Falsone, and M. Prandini, "Energy management of a building cooling system with thermal storage: a randomized solution with feedforward disturbance compensation," in *2016 American Control Conference (ACC 2016), Boston, Massachusetts, USA*, July 2016, pp. 2346–2351.

[34] R. L. Graham, D. E. Knuth, and O. Patashnik, *Concrete mathematics*. Adison Wesley, 1989.

Alessandro Falsone received the Bachelor degree in 2011 and the Master degree cum laude in 2013, both in Automation and Control Engineering from Politecnico di Milano. In 2018 he obtained the PhD degree in Information Engineering, System and Control division from Politecnico di Milano. During his PhD studies he also spent three months in the University of Oxford as a visiting researcher. Since 2018 he is a junior assistant professor at the Dipartimento di Elettronica, Informazione e Bioingegneria at Politecnico di Milano. His current research interests include distributed optimization and control, optimal control of stochastic hybrid systems, randomized algorithms, and nonlinear model identification. In 2018 he was the recipient of the Dimitris N. Chorafas Prize. In 2019 he received the IEEE CSS Italy Chapter Best Young Author Journal Paper Award.

Luca Deori received the laurea triennale cum laude in Control Engineering (2010), the laurea specialistica cum laude in Control Engineering (2012), and the Ph.D. in Information Technology Engineering (2015) from Politecnico di Milano, Italy. In 2016, he was a post-doctoral researcher with the Dipartimento di Elettronica, Informazione e Bioingegneria of the Politecnico di Milano, and also held an academic visitor position at the University of Oxford, UK. His research interests include optimization and control of complex uncertain systems, with application to air traffic control, electric vehicle charging control, optimal energy management in buildings, and design of sensing devices.



Daniele Ioli received the Bachelor degree in 2012 and the Master degree in 2014, both in Automation and Control Engineering from Politecnico di Milano. From 2014 to 2018 he was a graduate research assistant at the Dipartimento di Elettronica, Informazione e Bioingegneria at Politecnico di Milano. In 2017 he also spent five months as a visiting researcher at the General Electric Research Center in Munich. His main research interest is modeling and optimal control of smart grids.

Simone Garatti is Associate Professor at the Dipartimento di Elettronica ed Informazione of the Politecnico di Milano, Milan, Italy. He received the Laurea degree and the Ph.D. in Information Technology Engineering in 2000 and 2004, respectively, both from the Politecnico di Milano. In 2003, he held a visiting position at the Lund University of Technology, in 2006 at the University of California San Diego (UCSD), in 2007 at the Massachusetts Institute of Technology (MIT), and in 2019 at the University of Oxford. From 2013 to 2019 he was member of the EUCA Conference Editorial Board, while he is currently associate editor of the International Journal of Adaptive Control and Signal Processing and member of the IEEE-CSS Conference Editorial Board. He is also member of the IEEE Technical Committees on Computational Aspects of Control System Design and on System Identification and Adaptive Control, and of the IFAC Technical Committee on Modeling, Identification and Signal Processing. His research interests include data-driven optimization and decision-making, stochastic optimization for problems in systems and control, system identification, model quality assessment, and uncertainty quantification.

Maria Prandini received her Ph.D. degree in Information Technology in 1998. She was a postdoctoral researcher at UC Berkeley (1998-2000). She also held visiting positions at Delft University of Technology (1998), Cambridge University (2000), UC Berkeley (2005), and ETH Zurich (2006). In 2002, she became an assistant professor of automatic control at Politecnico di Milano, where she is currently a full professor. She was editor for the IEEE Control Systems Society (CSS) Electronic Publications (2013-15), elected member of the IEEE CSS Board of Governors (2015-17), and IEEE CSS Vice-President for Conference Activities (2016-17). She is currently IFAC Vice-President Conferences for the triennium 2020-23. She is program chair of IEEE Conference on Decision and Control 2021, and an associate editor of the IEEE Transactions on Network Systems and Automatica. In 2018, she received the IEEE CSS Distinguished Member Award. She was elevated to IEEE Fellow in 2020. Her research interests include stochastic hybrid systems, randomized algorithms, distributed and data-based optimization, multi-agent systems, and the application of control theory to transportation and energy systems.