Performance Bounds for Look-Ahead Power System Dispatch Using Generalized Multistage Policies

Paul N. Beuchat, Joseph Warrington, *Member, IEEE*, Tyler H. Summers, and Manfred Morari, *Fellow, IEEE*

***Abstract*—We present a combined look-ahead dispatch and reserve optimization formulation, which extends our recent work on time-coupled reserve policies and employs the recent notion of generalized decision rules from the robust optimization literature. This aims to improve the performance of traditional linear decision rules when applied to short-term electrical reserve operation. We derive a primal problem whose solution is a time-coupled rule for up- and down-regulating the power injections of each controllable device, such as a generator or energy storage unit, in response to discovered values of prediction errors. This rule depends on the "bin" into which measured prediction errors fall, so that, for example, up-regulation follows a different rule to down-regulation. We also derive an associated dual problem, whose solution provides a lower-bound on the best possible primal cost. This allows the suboptimality of a candidate solution based on a particular decision rule parameterization to be bounded. The primal and dual solutions are also compared to the so-called prescient case, in which the values of the uncertainty are known in advance. We demonstrate the method using numerical case studies, including the standard IEEE-118 bus network, in which a minimum possible reserve cost is identified using the dual lower-bounding approach.**

***Index Terms*—Automatic generation control, piecewise linear decision rules, power systems, renewables integration, robust optimization.**

I. Introduction

ALL power systems require a reserve system in order to accommodate contingencies and mispredictions of the system's net load (load minus power infeeds from other sources, such as intermittent renewables). In general, the larger the potential misprediction of net load, the larger a reserve margin is required. As a consequence, many power systems are seeing reserve margins, and resulting costs, increase as the share of fluctuating renewable energy sources such as wind and solar power increases [10]. Historically, reserve margins have been specified on a rather static basis to account for the failure of, for example, the largest nuclear plant. However, in power systems featuring a high penetration of intermittent renewables, errors in the prediction of these infeeds are increasingly outweighing the traditional contingency margin [16]. Consequently, short-term reserve needs must be calculated based on estimates of future forecast errors, whose properties are a function of the current and predicted future state of the weather. More precisely, a System Operator performing a look-ahead dispatch will need to model renewable infeeds as a random process whose bounds and moments should be reevaluated each time the dispatch is computed.

Manuscript received July 28, 2014; revised December 07, 2014; accepted January 04, 2015. Paper no. TPWRS-01023-2014.

The authors are with the Automatic Control Laboratory, ETH Zurich, 8092 Zurich, Switzerland (e-mail: beuchatp@control.ee.ethz.ch).

Digital Object Identifier 10.1109/TPWRS.2015.2394354

Renewable energy forecast errors are time-correlated on timescales of a few hours; for example, an under-prediction of wind infeed for one hour ahead makes it more likely that the wind infeed two hours ahead has also been under-predicted. The devices most able to compensate for such forecast errors, such as fast-acting generators and energy storage units (or increasingly, demand response), also have time-coupled constraints and costs on the same timescales. For example, a generator may have a limit on its ramping rate, and a battery must plan how to consume and supply power without exceeding its energy capacity. Consequently, it is attractive to study time-coupled responses to the uncertainty, which in previous work [17] were referred to as *reserve policies*. These can be seen as a logical extension of the increasingly time-coupled reserve products introduced by some System Operators in recent years, for example, the Flexible Ramping product introduced by CAISO [18]. An alternative look-ahead dispatch approach proposed in [6]is based on solving a chance constrained formulation of the OPF problem where the policies computed ensure that the limits and capacities of the system are satisfied with high probability. This contrasts with the robust approach presented in this paper that ensures the constrains are satisfied for all the predicted uncertainties within a given set. In the work of [5], a robust approach is also applied to solving an economic dispatch problem for power systems, but instead they aim of solve the unit-commitment problem with an objective of minimizing the worst case cost.

The reserve polices proposed fall in the class known as linear decision rules (LDRs), where a nominal solution is adapted linearly according to the realized value of the uncertainty. These have gained momentum since the work of Ben-Tal *et al.* [2] showing that optimal LDRs can be computed efficiently under the requirement that inequality constraints should hold for all realisations in the uncertainty set. However, despite their computation attractions, the degree of sub-optimality introduced by LDRs relative to other parameterizations of the adaptation can be quite significant [4], [8].

Georghiou *et al.* [9] presented a *generalized* decision rule method that can reduce suboptimality compared to LDRs. Their approach involves lifting the uncertainty set into a higher dimensional space and then solving for a decision rule as a function mapping the *lifted* uncertainty set to control inputs. This can give decision rules with increased flexibility. The related

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

studies [14] and [1] offer other approaches to increase the flexibility of the decision rules by considering the class of quadratic and polynomial decision rules respectively.

Although still solvable in polynomial time, the increased flexibility allowed by these methods comes with an associated increase in computational cost. Solving the dual of the LDR problem will give a strict lower bound on the optimal cost achievable by any decision rule. In [15] LDRs are applied to the dual problem to efficiently compute a lower bound, and in [9] the lower bound is improved instead by applying generalized decision rules. Hadjiyiannis *et al.* [13]apply the lower-bounding technique to a class of constrained finite-horizon robust control problems.

The piecewise affine policies presented in the present paper offer additional flexibility compared to the affine policies considered in [17]. The numerical example indicates that the cost of accommodating uncertainty, i.e., the reserve cost, can be significantly reduced by this increased flexibility. Additionally, the piecewise affine policies we present are constructed such that the reserve cost cannot be higher (that is, worse) than using simple affine policies.

It is also possible to lower bound the optimal cost of accommodating uncertainty by considering the case where its realized value is known in advance. We refer to this as the *prescient* case. We report that the cost of the optimal prescient dispatch plan may however be a conservative lower bound, in contrast to the dual-based lower bound presented in this work, which accounts for the need to accommodate uncertainty in the dispatch plan.

This paper is organized as follows. Section II outlines the state space approach used to model controllable and uncontrollable grid devices. Section III describes the concept of a generalized decision rule, and Section IV gives a general formulation of the look-ahead optimal dispatch problem. Sections V and VI derive a tractable finite-horizon optimization problem incorporating multistage recourse, providing an upper and lower bound, respectively, for the optimal cost of the general problem. Section VII demonstrates the method via a numerical example, and Section VIII concludes.

Notation: The notation $(.)'$ denotes the vector or matrix transpose; $[.]_k$ represents element k of a vector; $[.]_{:,m}$ column m of a matrix; $\otimes$ the Kronecker product; I_n the identity matrix of size n; $\mathbf{1}_n$ the vector of ones of size n; e_k the unit vector with the kth element equal to 1; $\mathrm{conv}(\cdot)$ denotes the convex hull of a set.

II. Power System Model

The power system model considered in this paper is the same as the model presented in [17]. There are N_p participants connected to a network of N_n nodes. Each participant can be associated with two types of power flows: an *inelastic* power flow, which cannot be controlled and maybe influenced by the modelled uncertainty, and an *elastic* power flow which can be controlled. The convention used is that positive power flows represent injection into the network. The following subsections details the components of the discrete-time model.

A. Inelastic Power Injection

The inelastic power flow from participant i over the future horizon of T time steps is modelled as affine in the uncertainty parameter $r_i + G_i\delta$ and cannot be influenced by any control actions. The vector $r_i \in \mathbb{R}^T$ is the nominal schedule for the power flow and the matrix $G_i \in \mathbb{R}^{T\times N_\delta T}$ maps a random vector $\delta \in \mathbb{R}^{N_\delta T}$ to the power flows. The vector δ represents the system uncertainties, e.g., from errors in the prediction of a load or renewable infeed, and is of the form $[\delta_1' \cdots \delta_T']'$ where each $\delta_k \in \mathbb{R}^{N_\delta}$ is the part of the vector discovered k steps ahead of the current time. We assume that the uncertainty is restricted to a compact set $\Delta := \{\delta \mid S\delta \le h\}$ containing the origin. Additionally, as a minimum, an estimate of its first and second moments $\mathbb{E}[\delta]$ and $\mathbb{E}[\delta\delta']$ are assumed to be available.

B. Elastic Power Flows

The elastic power flow is modelled as the output from a linear time-invariant (LTI) system using a state-space model. The state and input of participant i at time k are denoted $x_k^i \in \mathbb{R}^{n_i}$ and $u_k^i \in \mathbb{R}$, respectively, with

$$x_{k+1}^i = \tilde{A}_i x_k^i + \tilde{B}_i u_k^i \tag{1}$$

where $\tilde{A}_i \in \mathbb{R}^{n_i\times n_i}$ and $\tilde{B}_i \in \mathbb{R}^{n_i}$. The first element of the state $[x_k^i]_1$ represents the power output at time k while the remaining elements are used to model internal dynamics or to include prior states. The input at each time step is a scalar specifying the requested power output from that participant.

C. Finite-Horizon Stacked Representation

For elastic participants the state vectors and inputs over the time horizon T are stacked together as $\mathbf{x}_i = [x_1^i \cdots x_T^i]$ and $\mathbf{u}_i = [u_0^i \cdots u_{T-1}^i]$, respectively. Hence, the state evolution can be summarized as

$$\begin{aligned}\mathbf{x}_i &= \begin{bmatrix} \tilde{A}_i \\ \tilde{A}_i^2 \\ \vdots \\ \tilde{A}_i^T \end{bmatrix} x_0^i + \begin{bmatrix} \tilde{B}_i & 0 & \cdots & 0 \\ \tilde{A}_i\tilde{B}_i & \tilde{B} & \ddots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ \tilde{A}_i^{T-1}\tilde{B}_i & \cdots & \tilde{A}_i\tilde{B}_i & \tilde{B} \end{bmatrix} \mathbf{u}_i \\ &:= A_i x_0^i + B_i\mathbf{u}_i \end{aligned}\tag{2}$$

where $A_i \in \mathbb{R}^{n_iT\times n_i}$ and $B_i \in \mathbb{R}^{n_iT\times T}$.

The power injection from participant i into the transmission network is hence given by $C_i\mathbf{x}_i$, where $C_i = I_T \otimes e_1'$ selects the first element of the state vector for each time step.

D. Participant Costs and Local Constraints

For elastic participant i, the cost of a state and input trajectory (resp. $\mathbf{x}_i$ and $\mathbf{u}_i$ as defined above) is modelled by the function $J_i : \mathbb{R}^{Tn_i} \times \mathbb{R}^T \to \mathbb{R}$ assumed to be quadratic of the form

$$J_i(\mathbf{x}_i, \mathbf{u}_i) = f_i^{x\prime}\mathbf{x}_i + \frac{1}{2}\mathbf{x}_i' H_i^x \mathbf{x}_i + f_i^{u\prime}\mathbf{u}_i + \frac{1}{2}\mathbf{u}_i' H_i^u \mathbf{u}_i. \tag{3}$$

Additionally, it is assumed that the cost function is convex and that the costs are the same at each stage, which requires that the Hessian matrices take the form $H_i^{(\cdot)} = I_T \otimes \tilde{H}_i^{(\cdot)}$ with each $\tilde{H}_i^x \in \mathbb{R}^{n_i\times n_i}$ symmetric positive semidefinite and $\tilde{H}_i^u \in \mathbb{R}_+$, and the linear coefficients take the form $f_i^{(\cdot)} = 1_T \otimes \tilde{f}_i^{(\cdot)}$.

Time-coupled costs, for example, ramping costs, are readily incorporated into this framework by augmenting the state-vector to include memory of prior states.

The constraints on each participant's input and state sequences are represented by the set $\mathcal{Z}_i$, given by

$$\mathcal{Z}_i := \left\{ \begin{bmatrix} \mathbf{x}_i \\ \mathbf{u}_i \\ \delta \end{bmatrix} \middle| T_i\mathbf{x}_i + U_i\mathbf{u}_i + V_i\delta \le w_i \right\} \tag{4}$$

where the number of constraints used to model participant i is denoted l_i and the parameters are $T_i \in \mathbb{R}^{l_i \times n_i T}$, $U_i \in \mathbb{R}^{l_i \times T}$, $V_i \in \mathbb{R}^{l_i \times N_\delta T}$, and $w_i \in \mathbb{R}^{l_i}$. This allows a wide range of controllable systems to be modelled including those pertinent to conventional and renewable power system participants. For notational simplicity we define $\bar{l} = 1/N_\text{p} \sum_{i=1}^{N_\text{p}} l_i$ as the average number of local constraints per participant.

E. Network Constraints

The network consists of N_n nodes connected by L transmission lines, and is modelled using the same standard linear approximation as in [17], in which lines are lossless, bus voltage magnitudes are constant, and phase angle differences along lines are small. Each participant is connected to one of the nodes, leading to a network model that constrains the combined actions of the participants in two ways. First, at every time step, the net power injection and extraction from inelastic flows $r_i + G_i\delta$ and elastic flows $C_i\mathbf{x}_i$ must balance across all participants

$$\sum_{i=1}^{N_\text{p}} (r_i + G_i\delta + C_i\mathbf{x}_i) = 0. \tag{5}$$

Second, the power flows in the lines, which under the linear network approximation become linear functions of the nodal power injections, cannot exceed their respective ratings, in either direction, at any time. This results in a vector constraint with $2LT$ elements

$$\sum_{i=1}^{N_\text{p}} \Gamma_i(r_i + G_i\delta + C_i\mathbf{x}_i) \le \bar{p} \tag{6}$$

where the block-diagonal matrix $\Gamma_i \in \mathbb{R}^{2LT \times T}$ contains participant i's contributions to power flows in both directions, in each of the L lines, for each of the T time steps, and is determined by which of the N_n nodes the participant is connected. The vector $\bar{p} \in \mathbb{R}^{2LT}$ contains the line ratings. Each block of Γ_i is derived from the procedure in [7], in which the bus phase angles are eliminated.

III. Generalized Decision Rules

In the previous work [17], a tractable reformulation of the primal optimization problem was achieved by restricting the control policies to be affine in the uncertainty δ. Following by the work of [9], this restriction can be partially relaxed by *lifting* the uncertainty parameter into a higher dimensional space of size $N_\xi > N_\delta T$ and allowing a policy that is linear in the lifted uncertainty parameter $\xi \in \mathbb{R}^{N_\xi}$.

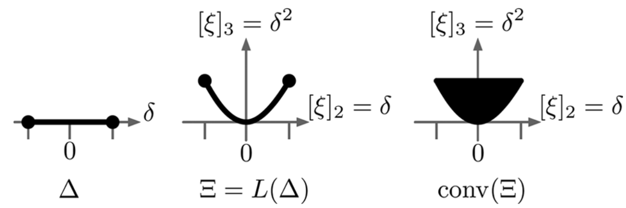

Fig. 1. Original uncertainty set Δ, its lifted counterpart Ξ obtained by applying the lifting $L(\delta) = [1, \delta, \delta^2]' = \xi$ to each element of the original set, and the convex hull of Ξ. The retraction operator for this lifting would ignore the quadratic term. Note that the first element of ξ is not shown because it is always equal to 1.

A. Lifting and Retraction Operator

As per [9], a *lifting operator* is introduced and defined as $L : \mathbb{R}^{N_\delta T} \to \mathbb{R}^{N_\xi}$, $\delta \mapsto \xi$, along with its corresponding *retraction operator* $R : \mathbb{R}^{N_\xi} \to \mathbb{R}^{N_\delta T}$, $\xi \mapsto \delta$, where $\mathbb{R}^{N_\xi}$ referred to as the *lifted space*. The operators L and R are required to satisfy the following properties.

1) L is continuous and satisfies $e_1' L(\delta) = 1$, $\forall \delta \in \Delta$ (i.e., the first component of $\xi = L(\delta)$ is always equal to 1).
2) R is linear.
3) $R \circ L = I_{N_\delta T}$
4) The component mappings of L are linearly independent, meaning that for each $v \in \mathbb{R}^{N_\xi}$

$$L(\delta)' v = 0 \quad \mathbb{P}_\delta\text{-a.s.} \qquad \Rightarrow \qquad v = 0. \tag{7}$$

As an example consider a one-dimensional (1-D) original uncertainty $N_\delta T = 1$and define the lifting operator as $L(\delta) = [1, \delta, \delta^2]' = \xi$, which is continuous as required, and hence the retraction operator can be defined as $R(\xi) = [0, 1, 0]\xi = \delta$, which is linear as required.

The lifted uncertainty set will be denoted $\Xi = L(\Delta)$ and can be computed by applying the lifting to every element of Δ. The lifted uncertainty set will in general be nonconvex and can be expressed as

$$\Xi = L(\Delta) = \left\{ \xi \in \mathbb{R}^{N_\xi} \middle| R(\xi) \in \Delta,\ L \circ R(\xi) = \xi \right\}. \tag{8}$$

Fig. 1 shows the lifted uncertainty set for the 1-D example above, where the original uncertainty set is a bounded line. It is clear that the lifted set is nonconvex, and in the sequel it will be required to have a description of the convex hull of Ξ. For this low-dimensional example, $\text{conv}(\Xi)$ is drawn by inspection, but for general liftings in a higher dimension it is usually not tractable or even possible to obtain a description of the convex hull.

With some abuse of notation, the matrix that describes the linear retraction operator will also be denoted R.

IV. Finite-Horizon Problem Statement

Here, we state the problem of minimizing the expected future costs over control policies, the primal problem, under which participants' planned actions at a given time may be any causal function of the uncertainty information δ available up to that time. Then, a dual formulation is presented, which will be used in subsequent sections to obtain a lower bound on the best achievable cost in the face of uncertainty.

A. General Problem Statement

The objective is to minimize the expected short-term running costs $\sum_{i=1}^{N_\mathrm{p}} \mathbb{E}[J_i(\mathbf{x}_i, \mathbf{u}_i)]$ over a horizon of length T, while fulfilling the local constraints (4) and network constraints (5) and (6) for all possible realizations of δ. This is achieved by application of a sequence of control inputs $\mathbf{u}_i$ which can vary as any general function of the uncertainty parameter δ, called a *policy*. To be practically implemented, the dependence is restricted to be causal, and the most general causal policy is denoted $\pi_i(\delta)$ for participant i. It is assumed that δ_{k+1} is discovered just in time to apply the inputs u_k^i, hence a causal policy is one where $[\pi_i(\delta)]_k$ is only a function of $\delta_1, \ldots, \delta_{k+1}$.

Substituting the input $\mathbf{u}_i = \pi_i(\delta)$ and the state equation into (3), (4), (5), and (6) to eliminate the state, the following *primal* optimization problem is obtained:

$$\begin{aligned}
\min_{\text{Causal } \pi_i} \; & \sum_{i=1}^{N_\mathrm{p}} \mathbb{E}\left[J_i\left(A_i x_0^i + B_i \pi_i(\delta), \pi_i(\delta)\right)\right] \\
\text{s.t.} \; & \sum_{i=1}^{N_\mathrm{p}} [r_i + G_i\delta + C_i(A_i x_0^i + B_i\pi_i(\delta))] = 0, \\
& \qquad\qquad\qquad\qquad \forall \delta \in \Delta \\
& \sum_{i=1}^{N_\mathrm{p}} \Gamma_i [r_i + G_i\delta + C_i(A_i x_0^i + B_i\pi_i(\delta))] \le \bar{p}, \\
& \qquad\qquad\qquad\qquad \forall \delta \in \Delta \\
& \begin{bmatrix} A_i x_0^i + B_i\pi_i(\delta) \\ \pi_i(\delta) \\ \delta \end{bmatrix} \in \mathcal{Z}_i, \quad \forall i, \; \forall \delta \in \Delta
\end{aligned} \tag{$\mathcal{P}$}$$

where problem labels, ($\mathcal{P}$) in this case, will be used throughout the paper to refer to both the specific optimization formulation and the optimal objective value were that problem to be solved.

B. General Lifted Problem Statement

Optimization problem ($\mathcal{P}$) is intractable due to the infinite space of feasible causal policies and the infinite constraints encoded by $\forall \delta \in \Delta$. Before performing a policy restriction similar to that in [17], a problem analogous to ($\mathcal{P}$) is formulated in terms of the lifted uncertainty. Let the policy now be a general function of the lifted uncertainty parameter, denoted $\pi_i^L(\xi)$.

The formal definition of causality for the lifted policy $\pi_i^L(\xi)$ depends on the definition of the lifting operator. In the interests of conciseness we simply note here that in general, an element of the lifted uncertainty can only influence the control decision once all the original uncertainty elements required to compute it become available.

The constraints are now required to hold robustly for all $\xi \in \Xi = L(\Delta)$, and any δ terms in the constraint equations are equivalently replaced by $R\xi$. This leads to the following problem:

$$\begin{aligned}
\min_{\text{Causal } \pi_i^L} \; & \sum_{i=1}^{N_\mathrm{p}} \mathbb{E}\left[J_i\left(A_i x_0^i + B_i \pi_i^L(\xi), \pi_i^L(\xi)\right)\right] \\
\text{s.t.} \; & \sum_{i=1}^{N_\mathrm{p}} [r_i + G_iR\xi + C_i(A_i x_0^i + B_i\pi_i^L(\xi))] = 0, \\
& \qquad\qquad\qquad\qquad \forall \xi \in \Xi \\
& \sum_{i=1}^{N_\mathrm{p}} \Gamma_i [r_i + G_iR\xi + C_i(A_i x_0^i + B_i\pi_i^L(\xi))] \le \bar{p}, \\
& \qquad\qquad\qquad\qquad \forall \xi \in \Xi \\
& \begin{bmatrix} A_i x_0^i + B_i\pi_i^L(\xi) \\ \pi_i^L(\xi) \\ R\xi \end{bmatrix} \in \mathcal{Z}_i, \quad \forall i, \forall \xi \in \Xi.
\end{aligned} \tag{$\mathcal{P}_L$}$$

Due to the lifted optimization problem ($\mathcal{P}$) being over all general causal policies as a function of the lifted uncertainty, a set of causal policies π_i that satisfy the constraints of problem ($\mathcal{P}$) can be mapped to a set of causal lifted policies π_i^L that satisfy the constraints of ($\mathcal{P}$). The reverse is also true that a feasible policy of ($\mathcal{P}$) can be mapped to a feasible policy of ($\mathcal{P}$), hence the problems have the same optimal value [9, Prop. 3.4]. Additional intractability is introduced when Ξ is nonconvex, but, fortunately, replacing the lifted set by its convex hull $\mathrm{conv}(\Xi)$changes neither the objective value nor the optimal policy [3].

In the sequel, the primal problem is solved in polynomial time by restricting the space of policies to finite dimensional decision rules. Although this yields a tractable problem, the result is conservative since the feasible set is restricted. This is less restrictive on the feasible set of ($\mathcal{P}_L$) compared to ($\mathcal{P}$); hence, the motivation for formulating the lifted problem is increased flexibility in the decision rules.

C. General Dual Formulation

To upper bound the conservatism that results from applying decision rules to the primal problem, a lower bound for ($\mathcal{P}$) is required. Formulating the following dual problem is motivated in Section VI because progressively better lower bounds to ($\mathcal{P}$) can be obtained by applying increasingly finely partitioned decision rules to the dual problem.

First, the inequality constraints (the transmission-line constraints and those defining each set $\mathcal{Z}_i$) are converted to equality constraints by introducing a non-negative slack function $s(\delta)$. An inner product with augmented set of $N_c = (T + 2LT + N_\mathrm{p}\bar{l})$ equality constraints and a general dual control policy $\nu(\delta)$ is then appended to the cost function, leading to the following formulation:

$$\begin{aligned}
\min_{s, \text{Causal } \pi_i} \sup_{\nu \in \mathcal{L}_{N_c}} \; & \mathbb{E}\Big[\sum_{i=1}^{N_\mathrm{p}} J_i(\mathbf{x}_i, \pi_i(\delta)) \\
& \qquad + \nu(\delta)' F(\mathbf{x}, \pi(\delta), s(\delta), \delta)\Big] \\
\text{s.t.} \; & s(\delta) \in \mathcal{L}_{2LT + N_\mathrm{p}\bar{l}}, \qquad s(\delta) \ge 0, \quad \forall \delta \in \Delta \\
& \mathbf{x}_i = A_i x_0^i + B_i \pi_i(\delta), \; \forall i
\end{aligned} \tag{$\mathcal{D}$}$$

where the vector $F(\cdot,\cdot,\cdot,\cdot) \in \mathbb{R}^{N_c}$ is the stacked vector of all network and local constraints expressed as equality constraints and is given by

$$F(\mathbf{x}, \pi(\delta), s(\delta), \delta) = \begin{bmatrix} \sum_{i=1}^{N_\mathrm{p}} [r_i + G_i\delta + C_i\mathbf{x}_i] \\ \sum_{i=1}^{N_\mathrm{p}} \Gamma_i[r_i + G_i\delta + C_i\mathbf{x}_i] - \bar{p} + s_N(\delta) \\ T_1\mathbf{x}_i + U_1\pi_1(\delta) + V_1\delta - w_1 + s_1(\delta) \\ \vdots \\ T_{N_\mathrm{p}}\mathbf{x}_i + U_{N_\mathrm{p}}\pi_{N_\mathrm{p}}(\delta) + V_{N_\mathrm{p}}\delta - w_{N_\mathrm{p}} + s_{N_\mathrm{p}}(\delta) \end{bmatrix}. \tag{12}$$

The stacked state vector $\mathbf{x}_i$has been reintroduced to simplify the notation. The notation $\nu \in \mathcal{L}_{N_c}$ and $s(\delta) \in \mathcal{L}_{2LT+N_\mathrm{p}\bar{l}}$ means that the dual control policy and slack function can be any general function of δ, where the subscript indicates the dimension.

In this formulation, the dual control policy $\nu(\delta)$ can be viewed as a penalty function where the supremum over ν will apply a high cost for violation of the constraints encoded in F. Consider a set of causal policies π_i for which at least one constraint is violated. This means that F will have a corresponding nonzero element. Any violation of the constraints encoded in F will lead to an unbounded inner maximization over ν. Thus, problems $(\mathcal{P})$ and $(\mathcal{D})$ effectively have the same feasible region for the outer minimization over π and are therefore equivalent [13].

Similar to the formulation of lifted primal problem $(\mathcal{P}_L)$, a lifted dual problem $(\mathcal{D}_L)$ that is equivalent to $(\mathcal{D})$ can be formulated by considering $\pi_i^L(\cdot)$, $\nu^L(\cdot)$, and $s^L(\cdot)$ to be a function of the *lifted* uncertainty ξ and replacing any occurrence of δ by $R\xi$.

V. Tractable Piecewise Affine Decision Rules

We now restrict the set of feasible policies in problem $(\mathcal{P}_\mathcal{L})$ to the set of lifted policies that are linear in the lifted uncertainty

$$\mathbf{u}_i = E_i\xi = \begin{bmatrix} e_i^\xi & D_i^\xi \end{bmatrix} \begin{bmatrix} 1 \\ [\xi]_2 \\ \vdots \\ [\xi]_{N_\xi} \end{bmatrix} \tag{13}$$

where $E \in \mathbb{R}^{T\times N_\xi}$. Remembering that $[\xi]_1 = 1$ by definition of the lifting operator, this means the policy is implicitly affine in the uncertainty with the first column, $e_i^\xi \in \mathbb{R}^T$, describing the nominal schedule and D_i^ξ the causal linear adjustment with respect to the prediction errors measured after the lifting is applied. The affine policies analyzed in [17] are a special case of this, seen by considering a lifting that leaves the uncertainty set unchanged $L(\delta) = [1, \delta']' = \xi$ (i.e., $N_\xi = N_\delta T + 1$), in this case e_i^ξ and D_i^ξ are exactly equivalent to e_i and D_i as defined in [17].

Substituting $\mathbf{u}_i = E_i\xi$ into the cost function $J(\cdot,\cdot)$ leads to the following expression for the cost in terms of the initial condition and the control policy, for participant i:

$$\begin{aligned} \tilde{J}_i(x_0^i, E_i) &= \mathbb{E}\left[J_i(A_ix_0^i + B_iE_i\xi, E_i\xi)\right] \\ &= c_i + f_i^{x\prime}A_ix_0^i + \frac{1}{2}x_0^{i\prime}A_i'H_i^xA_ix_0^i \\ &\quad + (f_i^{x\prime}B_i + f_i^{u\prime} + x_0^{i\prime}A_i'H_i^xB_i)E_i\mathbb{E}[\xi] \\ &\quad + \frac{1}{2}\langle E_i'(B_i'H_i^xB_i + H_i^u)E_i, \mathbb{E}[\xi\xi']\rangle. \end{aligned} \tag{14}$$

This shows that evaluating the cost function for the lifted problem requires knowledge of the first and second moments of the *lifted* uncertainty. For nontrivial liftings, this information cannot be extracted from $\mathbb{E}[\delta]$ and $\mathbb{E}[\delta\delta']$, thus it represents an additional input information requirement which may require additional uncertainty modelling effort to obtain meaningful estimates of $\mathbb{E}[\xi]$ and $\mathbb{E}[\xi\xi']$.

Applying linear policies to lifted problem $(\mathcal{P}_L)$ results in the following finite-dimensional optimization problem:

$$\begin{aligned} \min_{\text{Causal } E_i\in\mathbb{R}^{T\times N_\xi}} \quad & \sum_{i=1}^{N_\mathrm{p}} \tilde{J}_i\left(x_0^i, E_i\right) \\ \text{s.t.} \quad & \sum_{i=1}^{N_\mathrm{p}} [r_i + (G_iR + C_iB_iE_i)\xi + C_iA_ix_0^i] = 0, \\ & \qquad \forall\xi \in \mathrm{conv}(\Xi) \\ & \sum_{i=1}^{N_\mathrm{p}} \Gamma_i[r_i + (G_iR + C_iB_iE_i)\xi + C_iA_ix_0^i] \le \bar{p}, \\ & \qquad \forall\xi \in \mathrm{conv}(\Xi) \\ & \begin{bmatrix} A_ix_0^i + B_iE_i\xi \\ E_i\xi \\ R\xi \end{bmatrix} \in \mathcal{Z}_i, \quad \forall i, \forall\xi \in \mathrm{conv}(\Xi) \end{aligned} \tag{$\mathcal{P}_L^*$}$$

where optimizing over linear policies is a restriction to the feasible space of all general policies, and the optimal objective value will be conservative, $(\mathcal{P}_\mathcal{L}) \le (\mathcal{P})$.

If there exists a tractable half-space representation for $\mathrm{conv}(\Xi)$, then, letting q_ξ denote the number of half-spaces, problem $(\mathcal{P}_L)$ can be readily formulated as a tractable program by introducing a matrix of slack variables of size $(q_\xi \times (2LT + N_\mathrm{p}\bar{l}))$ and following the strong duality argument presented in [12].

If no such representation exists or it cannot be tractably computed, then problem $(\mathcal{P}_L)$ can be upper bounded by following the approach of [9], where the authors introduce a larger uncertainty set $\hat{\Xi} \supseteq \mathrm{conv}(\Xi)$ called the *outer approximation* of Ξ and enforce robustness to this larger set. This $\hat{\Xi}$ set is then chosen such that it has a tractable half-space representation and thus leads to a tractable optimization problem.

A. Piecewise Linear Policies

For the remainder of this section, the lifting operator is specified and an outer approximation $\hat{\Xi}$ is given that is tractable for any original uncertainty set Δ. The lifting operator splits each dimension into pieces and associates each piece with a new dimension of the lifted uncertainty set. Let m_j denote the number of pieces into which dimension $j = 1, \ldots, N_\delta T$ of δ is split, and let the scalar locations of the $(m_j - 1)$ split points be denoted and ordered as follows:

$$z_1^j < z_2^j < \cdots < z_{(m_j-1)}^j \tag{16}$$

where choosing an $m_j = 1$ means that dimension j of the original uncertainty is unaffected by the lifting. This specification means the liftings for each dimension are independent and the lifting operator can be written separately for each dimension as

$$L_j(\delta) = \begin{bmatrix} \min([\delta]_j, z_1^j) \\ \max(\min([\delta]_j, z_2^j) - z_1^j, 0) \\ \vdots \\ \max(\min([\delta]_j, z_{m_j-1}^j) - z_{m_j-2}^j, 0) \\ \max([\delta]_j, z_{m_j-1}^j, 0) \end{bmatrix} \in \mathbb{R}^{m_j} \tag{17}$$

with the full lifting operator defined as

$$\xi = L(\delta) = [\,1, \quad L_1(\delta)', \quad \ldots, \quad L_{N_\delta T}(\delta)'\,]' \tag{18}$$

noting that $L_j(\delta) = [\delta]_j$ if $m_j = 1$. The retraction operator for this piecewise lifting is constructed based on $[\delta]_j = \sum_{k=1}^{m_j} [L_j(\delta)]_k$. The causality constraint on the matrix D_i^ξ as defined by this lifting operator is

$$D_i^\xi = \begin{bmatrix} [D_i^\xi]_{0,0} & 0 & \cdots & 0 \\ [D_i^\xi]_{1,0} & [D_i^\xi]_{1,1} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ [D_i^\xi]_{T-1,0} & \cdots & [D_i^\xi]_{T-1,T-2} & [D_i^\xi]_{T-1,T-1} \end{bmatrix} \tag{19}$$

where $[D_i^\xi]_{l,k} \in \mathbb{R}^{1 \times \left(\sum_{j=kN_\delta+1}^{(k+1)N_\delta} m_j\right)}$ is the response of input u_l^i to the error δ_k after the lifting is applied.

This lifting is chosen based on [9], where the authors give the following tractable analytical expression for $L(\Delta)$ when Δ is a hyper-rectangle:

$$\Delta = \{\delta \in N_\delta T | l_j \le [\delta]_j \le u_j, j = 1, \ldots, N_\delta T\} \tag{20}$$

for which $\mathrm{conv}(\Xi) = \mathrm{conv}(L(\Delta)) = \prod_{j=1}^{N_\delta T} \mathrm{conv}(L_j(\Delta))$, where the polytope for $\mathrm{conv}(L_j(\Delta))$ is defined by

$$S_j L_j(\delta) \ge \left[\frac{-z_1^j}{z_1^j - l_j}, \ \frac{l_j}{z_1^j - l_j}, \ 0, \ \ldots, \ 0\right]'$$

$$S_j = \begin{bmatrix} \frac{-1}{z_1^j - l_j} & 0 & \cdots & & 0 \\ \frac{1}{z_1^j - l_j} & \frac{-1}{z_2^j - z_1^j} & \ddots & & \vdots \\ 0 & \frac{1}{z_2^j - z_1^j} & \ddots & 0 & 0 \\ 0 & 0 & \ddots & \frac{-1}{z_{m_j-1}^j - z_{m_j-2}^j} & 0 \\ \vdots & & \ddots & \frac{1}{z_{m_j-1}^j - z_{m_j-2}^j} & \frac{-1}{u_j^j - z_{m_j-1}^j} \\ 0 & & \cdots & 0 & \frac{1}{u_j^j - z_{m_j-1}^j} \end{bmatrix} \tag{21}$$

which contains $m_j + 1$ hyperplanes. In the case that Δ is not a hyper-rectangle, then under the piecewise lifting defined, an outer approximation $\hat{\Xi}$ can be constructed by first defining $\hat{\Delta}$ as the smallest hyper-rectangle that contains Δ and then taking the following intersection of sets:

$$\hat{\Xi} = \left\{\xi \in \mathbb{R}^{N_\xi} \middle| \xi \in \mathrm{conv}\left(L(\hat{\Delta})\right), SR\xi \le h\right\} \tag{22}$$

where the second term in the definition ensures that all elements $\xi \in \hat{\Xi}$ map back under the retraction to be an element of the original uncertainty set, $R(\xi) \in \Delta$.

Problem $(\mathcal{P}_L^*)$ can now be solved for any choice of split points per the definitions above; with the retraction operator for the piecewise lifting defined as $R = [0_{N_\delta T \times 1}, R^*] = [0_{N_\delta T \times 1}, \mathrm{diag}(\mathbf{1}_{m_1}', \ldots, \mathbf{1}_{m_{N_\delta T}}')]$. Letting $(\mathcal{P}^*)$ correspond to the problem of finding an optimal policy that is affine in δ (as per [17]) we can state the following lemma.

Lemma 5.1: A feasible policy of $(\mathcal{P}_L^*)$ restricted to be affine in δ $(E_i = [e_i^\xi, D_i R^*], D_i \in \mathbb{R}^{T \times N_\delta T})$ is feasible for $(\mathcal{P}^*)$.

Proof: This proof has been adapted from the proof of [9]Proposition 3.4 to be in notation consistent with this paper. The constraints of $(\mathcal{P}_L^*)$ are all affine in ξ and can be written concisely as $AR\xi + BE_i\xi + C = 0, \forall \xi \in \hat{\Xi}$. Letting the lifted policies be affine in δ the constraint becomes: $AR\xi + B[e_i^\xi, D_i R^*]\xi + C = 0, \forall \xi \in \hat{\Xi}$. From the definition of $\hat{\Xi}$, we have that $R\xi \in \Delta, \forall \xi \in \hat{\Xi}$, therefore the constraint is equivalent to: $A\delta + Be_i^\xi + BD_i\delta + C = 0, \forall \delta \in \Delta$. This is exactly the constraint form for problem $(\mathcal{P}^*)$, hence e_i^ξ and D_i is a feasible policy of $(\mathcal{P}^*)$. ■

The implication of this lemma is that the set of policies considered when solving for the optimal piecewise affine policy, includes the set of all policies that are affine in the original uncertainty dimension δ. When the moments used in the cost function are computed from a set of N uncertainty samples, each denoted $\delta^{(n)}$, then it is also important to check that the optimal cost from applying piecewise affine policies cannot be worse than applying affine policies. Given that the first and second moments are computed from the samples as $\mathbb{E}[\delta\delta'] = 1/N \sum_{n=1}^N \delta^{(n)}\delta^{(n)\prime}$ and $\mathbb{E}[\xi\xi'] = 1/N \sum_{n=1}^N L(\delta^{(n)})L(\delta^{(n)})'$, then we can state the following lemma.

Lemma 5.2: The objective values of $(\mathcal{P}_L^*)$ equals that of $(\mathcal{P}^*)$ when the moments of the uncertainty set are computed from the same data and the lifted policies are restricted to be affine in δ $(E_i = [e_i^\xi, D_i R^*], D_i \in \mathbb{R}^{T \times N_\delta T})$.

Proof: The objective function of $(\mathcal{P}_L^*)$ contains quadratic terms of the form $\mathbb{E}[\xi' E_i' H E_i \xi]$, where H is a symmetric positive semi-definite matrix. Letting the lifted policies be affine in δ the quadratic term becomes $\mathbb{E}\left[\xi'[e_i^{\xi\prime}, \ R^{*\prime}D_i']H[e_i^\xi, \ D_i R^*]\xi\right]$. Given that $[\xi]_1 = 1$, and denoting $\xi = [1, \ \xi^{*\prime}]'$, this becomes: $\mathbb{E}\left[(e_i^{\xi\prime} + \xi^{*\prime}R^{*\prime}D_i')H(e_i^\xi + D_i R^* \xi^*)\right]$. Writing out the expectation as it is computed from data leads to

$$\begin{aligned} &= \frac{1}{N}\sum_{n=1}^N \left[(e_i^{\xi\prime} + \xi^{(n)*\prime}R^{*\prime}D_i')H(e_i^\xi + D_i R^* \xi^{(n)*})\right] \\ &= \frac{1}{N}\sum_{n=1}^N \left[(e_i^{\xi\prime} + \delta^{(n)\prime}D_i' H(e_i^{\xi\prime} + D_i\delta^{(n)})\right] \\ &= \mathbb{E}\left[(e_i^{\xi\prime} + \delta' D_i')H(e_i^{\xi\prime} + D_i\delta)\right] \end{aligned} \tag{23}$$

which equals the quadratic term in the cost function of $(\mathcal{P}^*)$. ■

These lemmas show that, when the moments of the lifted and original uncertainty set are computed from the same data samples, then it is sufficient to compute only the optimal piecewise affine dispatch policy for a chosen lifting. In other words, it is not necessary to compute the optimal affine policy for making a comparison.

VI. Tractable Lower Bound Computation

In a similar fashion to Section V, the supremum over all general dual policies $\nu(\xi)$ is made tractable by restricting $(\mathcal{D}_L)$ to dual policies that are linear in the lifted uncertainty

$$\nu(\xi) = Y\xi, \quad Y \in \mathbb{R}^{N_c \times N_\xi}. \tag{24}$$

The restriction of the dual control policies effectively increases the feasible set of π^L in the infimum and hence is a relaxation of problem $\mathcal{D}_L$, i.e.,

$$\begin{aligned}
&\min_{s,\text{Causal } \pi_i^L} \sup_{\nu \in \mathcal{L}_{N_c}} \mathbb{E}\Bigg[\sum_{i=1}^{N_\mathrm{p}} J_i\left(\mathbf{x}_i, \pi_i^L(\xi)\right) \\
&\qquad\qquad + \nu(\xi)' F(\mathbf{x}, \pi^L(\xi), s(\xi), R\xi)\Bigg] \\
&\geq \min_{\text{Causal } \pi_i^L} \sup_{Y \in \mathbb{R}^{N_c \times N_\xi}} \mathbb{E}\Bigg[\sum_{i=1}^{N_\mathrm{p}} J_i\left(\mathbf{x}_i, \pi_i^L(\xi)\right) \\
&\qquad\qquad + \xi' Y' F(\mathbf{x}, \pi^L(\xi), s(\xi), R\xi)\Bigg].
\end{aligned} \tag{25}$$

As the objective function is linear in Y the supremum can be computed and leads to the following problem:

$$\begin{aligned}
&\min_{s,\text{Causal } \pi_i} \mathbb{E}\left[\sum_{i=1}^{N_\mathrm{p}} J_i\left(\mathbf{x}_i, \pi_i(\xi)\right)\right] \\
&\text{s.t. } s(\xi) \in \mathcal{L}_{2LT+N_\mathrm{p}\bar{l}}, \qquad s(\xi) \geq 0 \quad \forall \xi \in \Xi \\
&\qquad \mathbb{E}[F(\mathbf{x}, \pi(\xi), s(\xi), R\xi)\xi'] = 0 \\
&\qquad \mathbf{x}_i = A_i x_0^i + B_i \pi_i(\xi), \forall i
\end{aligned} \tag{$\mathcal{D}_L^*$}$$

Assuming $\mathrm{conv}(\Xi)$ has a tractable half-space representation denoted by

$$\mathrm{conv}(\Xi) = \{\xi \in \mathbb{R}^{N_\xi} | W_\xi \xi \geq 0, [\xi]_1 = 1\} \tag{27}$$

where $W_\xi \in \mathbb{R}^{q_\xi \times N_\xi}$, and following the argumentation presented in [13], problem $(\mathcal{D}_L^*)$ can be shown to be equivalent to the following tractable problem:

$$\begin{aligned}
&\min_{S_{(\cdot)},\text{Causal } E_i \in \mathbb{R}^{T \times N_\xi}} \sum_{i=1}^{N_\mathrm{p}} \tilde{J}_i\left(x_0^i, E_i\right) \\
&\text{s.t. } S_n \in \mathbb{R}^{2LT \times N_\xi}, \ S_i \in \mathbb{R}^{l_i \times N_\xi} \\
&\sum_{i=1}^{N_\mathrm{p}} [G_i R + C_i B_i E_i + (r_i + C_i A_i x_0^i) e_1'] = 0, \\
&\sum_{i=1}^{N_\mathrm{p}} \Gamma_i [G_i R + C_i B_i E_i + (r_i + C_i A_i x_0^i) e_1'] + S_n = \bar{p} e_1', \\
&T_i B_i E_i + U_i E_i + V_i R + (T_i A_i x_0^i - w_i) e_1' + S_i = 0 \\
&\qquad \forall i = 1, \ldots, N_\mathrm{p} \\
&\begin{bmatrix} W_\xi \\ e_1' \end{bmatrix} \mathbb{E}[\xi\xi'] \begin{bmatrix} S_n', & S_1', & \ldots, & S_{N_\mathrm{p}}' \end{bmatrix} \geq 0.
\end{aligned} \tag{28}$$

The optimal value of (28) equals the optimal value of $(\mathcal{D}_L^*)$ and by (25) is lower bound on the optimal solution of (D), which is in turn equal to that of $(\mathcal{P})$. Thus (28) can be used to assess how close a solution of $(\mathcal{P}_L^*)$ is to obtaining the best achievable cost of accommodating uncertainty.

VII. Numerical Examples

This section presents the results of applying the bounding approach to two case studies. First, a simplified eight-participant system is used to exemplify the benefits of the approach. Second, the IEEE-118 bus network is used to demonstrated that the approach has potential on real-sized systems.

A. Uncertainty Model Description

In order to model uncertainty for the inelastic participants, the power injections from the the wind farm and load were driven with a first order stochastic process. Let $q_k \in \mathbb{R}^{N_\delta}$ denote the state of the uncertainty model at time k. The uncertainty model is driven by the following update equation, with saturation limits on q_k:

$$q_{k+1} = \min\{\max\{q_{\min}, q_k + \beta_k\}, q_{\max}\} \tag{29}$$

where $\beta_k \in \mathbb{R}^{N_\delta}$ is sampled from a multivariate normal distribution with variance $\Sigma \in \mathbb{R}^{N_\delta \times N_\delta}$. The stacked vector $q = [q_1', \ldots, q_T']'$ is then the random future evolution from the initial state q_0. The saturation bounds, $q_{\min}$ and $q_{\max}$, represent the physical limit of the system. The uncertainty model parameters are also shown in Table I.

Monte Carlo sampling of β_k was used to generate $N = 3 \times 10^5$ future evolutions of the first-order stochastic process over the time horizon T, where each evolution is denoted $q^{(n)} \in \mathbb{R}^{N_\delta T}, n = 1, \ldots, N$. The sample mean $\mathbb{E}[q]$ was computed, allowing the nominal prediction r_i for the inelastic participants to be computed as $r_i = G_i \mathbb{E}[q]$. The prediction error for each sample is then computed as $\delta^{(n)} = q^{(n)} - \mathbb{E}[q]$ from which the second moment matrix $\mathbb{E}[\delta\delta']$ can be readily computed and $\mathbb{E}[\delta] = 0$ by definition. As saturation is only applied to the state of the uncertainty model, q, the bounds describing the original uncertainty polytope are

$$\Delta = \left\{ \delta \in \mathbb{R}^{N_\delta T} \,\middle|\, \begin{array}{c} (\mathbf{1}_T \otimes q_{\min}) - \mathbb{E}[q] \leq \delta \\ \delta \leq (\mathbf{1}_T \otimes q_{\max}) - \mathbb{E}[q] \end{array} \right\}. \tag{30}$$

The lifted prediction error is computed for each sample, $\xi^{(n)} = L(\delta^{(n)})$, from which the first and second moments $\mathbb{E}[\xi]$ and $\mathbb{E}[\xi\xi']$ are computed. As Δ is a hyper-rectangle, $\mathrm{conv}(\Xi)$ is constructed as per the description in Section V. These were computed for piecewise lifting functions $L(\cdot)$ with $m_j = \{1,\ 2,\ 4,\ 8\}$ pieces for every dimension of δ.

The number of pieces was varied in the primal $(\mathcal{P}_L^*)$ (respectively dual $(\mathcal{D}_L^*)$) forms of the problem. In order to obtain progressively lower (resp. higher) upper (resp. lower) bounds on the optimal cost of accommodating the uncertainty the split points were spread evenly between the lower and upper bound

TABLE I
SYSTEM PARAMETERS FOR EIGHT-PARTICIPANT SYSTEM

Elastic Particpants - Thermal Generators
Cost function: $\tilde{f}_i^u u_k^i + \tilde{H}_i^u (u_k^i)^2 + \alpha(u_k^i - u_{k-1}^i)^2$
Power output: $0 \leq u_k^i \leq p_{\max}^i$
Ramp contraints: $r_{\min}^i \leq (u_k^i - u_{k-1}^i) \leq r_{\max}^i$
Initial condition: p_0^i

i	$\tilde{f}_i^u$ [\$/MW]	$\tilde{H}_i^u$ [\$/MW^2]	α	$p_{\max}^i$ [MW]	p_0^i	$r_{\min}^i$ [MW/h]	$r_{\max}^i$
1	20	0.02	1.0	1000	500	-80	40
2	20	0.03	0.8	900	400	-100	50
3	20	0.08	0.4	400	150	-350	350
4	20	0.10	0.1	300	100	-300	300

Elastic Particpants - Storage Units
Maximum storage capacity: $s_{\max}^i$
Cost function: $\gamma \left(\frac{s_{\max}^i}{2} - u_k^i \right)^2$
Power output: $p_{\min}^i \leq u_k^i \leq p_{\max}^i$
Ramp contraints: $r_{\min}^i \leq (u_k^i - u_{k-1}^i) \leq r_{\max}^i$
Initial storage level: s_0^i
Initial power injection = 0 [MW]

i	γ [\$/MW^2]	$s_{\max}^i$ [MW]	s_0^i	$p_{\min}^i$ [MW]	$p_{\max}^i$	$r_{\min}^i$ [MW/h]	$r_{\max}^i$
5	0.010	1000	500	-200	200	-100	100
6	0.005	500	250	-200	200	-100	100

Uncertainty Model
r_i, $\mathbb{E}[\delta]$, $\mathbb{E}[\delta\delta']$, $\mathbb{E}[\xi]$, $\mathbb{E}[\xi\xi']$ computed as per the decription in the text using parameters: $N_\delta = 2$, $\Sigma = \text{diag}(150^2, 10^2)$,

$$q_0 = \begin{bmatrix} 350 \\ -1500 \end{bmatrix} \quad q_{\min} = \begin{bmatrix} 0 \\ -1550 \end{bmatrix} \quad q_{\max} = \begin{bmatrix} 750 \\ -1450 \end{bmatrix}$$

Inelastic Participants
$i = 7$, Wind Farm, $\tilde{G}_i = \begin{bmatrix} 1 & 0 \end{bmatrix}$, $\quad i = 8$, Load, $\tilde{G}_i = \begin{bmatrix} 0 & 1 \end{bmatrix}$
$G_i = I_T \otimes \tilde{G}_i$

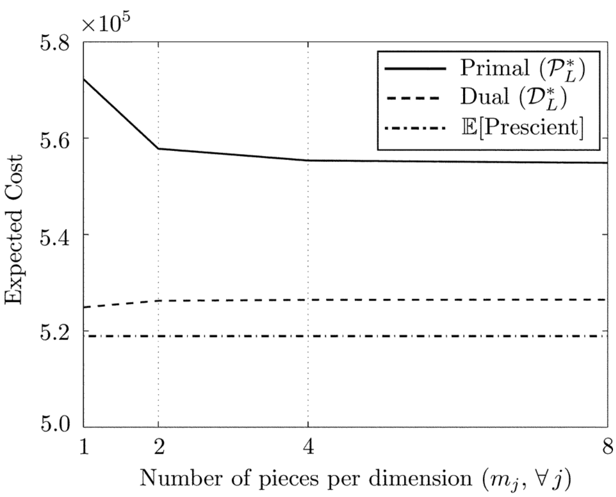

Fig. 2. Variation of upper (solid line) and lower (dashed line) bounds on the optimal expected reserve cost with the number of pieces in the primal or dual policy. The expected prescient case (dashed–dotted line) is also shown as a horizontal line. The largest gap between the dual and prescient costs is a lower bound on the minimum possible cost of accommodating uncertainty in this look-ahead dispatch.

of each dimension. The problems were solved, and the cost results are shown on Fig. 2.

The *prescient* case was computed by measuring the optimal solution to a deterministic version of problem ($\mathcal{P}$) in which the optimal open-loop plans e_i^ξ were found (no reserves, or equivalently no matrices D_i^ξ, are needed in the deterministic case). The optimal cost was averaged over 5000 Monte Carlo realizations of the stochastic process in order to obtain an *expected prescient cost*. This value is plotted alongside the bounds on robust costs on Fig. 2.

B. Small Eight-Participant Example

The properties of the eight-participant system are listed in Table I. For this example, it was assumed that the transmission-line capacities are sufficiently high that the corresponding constraints can be neglected.

Fig. 2 shows that the expected operating costs in the face of uncertainty, (PLa), reduces as the number of pieces in the piecewise affine policy is increased. For the example considered, the two-, four-, and eight-piece policies represent, respectively, a 2.52%, 2.94%, and 3.02% reduction in cost over the one-piece, affine in δ, policies from [17].

In this numerical example, the dual problem gives a better lower bound than the Prescient case. The dual lower bounds ($\mathcal{D}_L^*$) show the expected trend of increasing values as the number of pieces is increased. There is only a 0.30% increase from the one-piece to the eight-piece dual lower bound.

The best achievable operating cost in the face of uncertainty must lie between the Primal and Dual costs. Hence the best possible dual lower bound (that achieved with eight pieces in problem ($\mathcal{D}_L^*$) in the numerical example) can be used to give an upper bound on the sub-optimality of any primal piecewise affine policy. For the numerical example, the one-piece, affine in δ, policy is within 7.9% of the best achievable cost, while the eight-piece policy is within 4.9%. The majority of this reduction in suboptimality over the one-piece policy is achieved by the two-piece policy.

There may be some cases in which the dual control policy restriction to the class of piecewise affine policies leads to lower bounds that are less than the prescient case. Evaluating the prescient case has a small computational load relative to ($\mathcal{P}_L^*$) and ($\mathcal{D}_L^*$), and it is also guaranteed to give a lower bound on the best achievable operating cost in the face of uncertainty. Therefore, it is worthwhile to always compute both and use the higher lower bound for assessing the suboptimality of a primal piecewise affine reserve policy.

C. IEEE 118-Bus Network Case Study

The details of the IEEE 118-Bus Network were extracted from the "Matpower" toolbox [19]. The 118-bus system was then augmented with 18 wind farms that were geographically clustered in three groups of six. The wind farms were specified as inelastic participants driven by a six-dimensional uncertainty model. The bus to which each wind farm is connected and the properties of the uncertainty model are given in Table II. The wind farms were sized so that their maximum cumulative injection was approximately one third of the total load.

Table II also details the ramp constraint imposed on each generator and the transmission line capacities selected, neither of which is specified for the standard 118-bus network. Three sets of transmission lines were chosen that connect major sections of the network graph. In order to specify meaningful transmission

TABLE II
SYSTEM PARAMETERS FOR 118-BUS NETWORK

Elastic Particpants - 54 Thermal Generators
Ramp contraints: $r^i_{\min} \leq (u^i_k - u^i_{k-1}) \leq r^i_{\max}$
For the 35 generators with $p^i_{\max} = 100[MW]$, set:
$-r^i_{\min} = r^i_{\max} = 50[MW]$
For the other generators $p^i_{\max} > 100[MW]$, set:
$-r^i_{\min} = r^i_{\max} = \max\left(20, 0.075\, p^i_{\max}\right)[MW]$
Initial condition p^i_0 spread across all generators in proportion to $p^i_{\max}$.

Uncertainty Model
r_i, $\mathbb{E}[\delta]$, $\mathbb{E}[\delta\delta']$, $\mathbb{E}[\xi]$, $\mathbb{E}[\xi\xi']$ computed using parameters:
$N_\delta = 6$, $q_0 = 50 \otimes \mathbf{1}_6$, $q_{\min} = 0 \otimes \mathbf{1}_6$, $q_{\max} = 100 \otimes \mathbf{1}_6$,

$$\Sigma = \begin{bmatrix} 225 & 25 & 0 & 6 & 0 & 0 \\ 25 & 225 & 4 & 0 & 0 & 0 \\ 0 & 4 & 324 & 12 & 9 & 2 \\ 6 & 0 & 12 & 121 & 0 & 8 \\ 0 & 0 & 9 & 0 & 289 & 24 \\ 0 & 0 & 2 & 8 & 24 & 256 \end{bmatrix}$$

Inelastic Participants - Wind Farms

Group 1		Group 2		Group 3	
Bus	$\tilde{G}_i$	Bus	$\tilde{G}_i$	Bus	$\tilde{G}_i$
1	$[0.75, 0.25]$	53	$[0.90, 0.10]$	105	$[0.70, 0.30]$
2	$[0.80, 0.20]$	54	$[0.85, 0.15]$	106	$[0.80, 0.20]$
3	$[0.85, 0.15]$	55	$[0.15, 0.85]$	108	$[0.60, 0.40]$
4	$[0.25, 0.75]$	56	$[0.15, 0.80]$	109	$[0.50, 0.50]$
5	$[0.20, 0.80]$	57	$[0.35, 0.70]$	110	$[0.40, 0.60]$
6	$[0.15, 0.85]$	58	$[0.35, 0.65]$	111	$[0.30, 0.70]$

G_i is computed differently for each group:
Group 1: $I_T \otimes \left[\tilde{G}_i \quad 0_{1\times 4}\right]$ Group 3: $I_T \otimes \left[0_{1\times 4} \quad \tilde{G}_i\right]$
Group 2: $I_T \otimes \left[0_{1\times 2} \quad \tilde{G}_i \quad 0_{1\times 2}\right]$

Inelastic Participants - Load
The load at each bus was held constant over time, taken directly from the IEEE specifications, summing to a cumulative load of 4242 [MW]

Network Constraints

Line #	30	70	71	75	76	105	106	107	108
$\bar{p}$ [MW]	40	40	50	35	30	55	40	100	95

line capacities, the problem was solved without network constraints, the range of line flows was observed, and the capacities were selected from within this range.

The dispatch problem was formulated for a look-ahead time horizon of $T = 8$ time steps. Fig. 3 shows the expected operating cost in the face of uncertainty for this 118 bus network case study. The primal costs, $(\mathcal{P}^*_L)$, show a similar trend as the smaller example above, with the two- and four-piece policies representing, respectively, a 2.51% and 3.17% reduction in cost over the one-piece policy.

In this case the lower bound provided by the dual problem (DLa)improved negligibly by increasing the number of pieces in the lifting. The expected prescient case was computed from 10 000 Monte Carlo samples and was a tighter lower bound for this case study, although it was only 0.1% better than the dual lower bound. This supports the statement made above for the small numerical example that the prescient case can provide a tighter lower bound in certain cases.

The details of the optimization problem solved to get these results are shown in Table III. As indicated by the number of variables, the size of both the primal and dual optimization problem is proportional to the number of pieces used in the PWA policy.

For the computation times shown in Table III, the optimization problems were solved on a Dual Deca-Core Intel Xeon E5–2690 v2 3.0 GHz using the Gurobi [11] sparse QP solver package. The computation times for both the primal and dual problem increased significantly with the number of pieces. As the case studies have shown minimal improvement in the dual lower bound by increasing the number of pieces, a reasonable approach would be to compute both the dual lower bound using affine policies, and the expected prescient case, and take the higher as the lower bound for assessing the results of the primal problem. For the primal problem, to leverage the cost reductions indicated by the results, a reasonable approach is to use the maximum number of pieces for which the piecewise affine policy can be computed on the resources available in the time frame required.

The breakpoints of the piecewise liftings used in the case studies were chosen to be spread evenly between the lower and upper bound of each dimension. It is possible to formulate a optimization problem to find the best split point locations for a given number of split points per dimension, however, the resulting problem is nonconvex [9] and hence adds significant computational burden. The results of the primal problem in both case studies suggest that even the simplest two-piece lifting that splits each uncertainty dimension in half achieves a significant reduction in cost. Additionally, in both case studies the dual lower bound computation indicate that a two-piecewise affine policy is significantly closer to the best achievable cost when compared to a purely affine policy. Therefore, for a given problem instance, it is reasonable to use evenly spaced splits for defining the piecewise lifting.

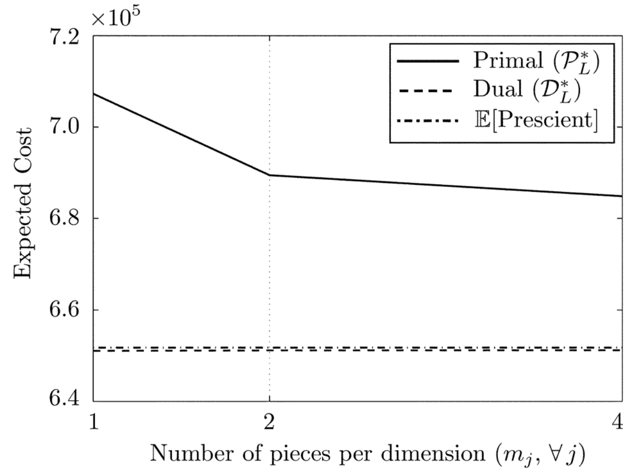

Fig. 3. Variation of upper (solid line) and lower (dashed line) bounds on the optimal expected reserve cost with the number of pieces in the primal or dual policy. The expected prescient case (dashed–dotted line) is also shown as a horizontal line. In this look-ahead dispatch, the prescient case gives the better lower bound.

TABLE III
COMPUTATIONAL DETAILS FOR 118-BUS NETWORK

	$(\mathcal{P}^*_L)$			$(\mathcal{D}^*_L)$		
# sides	1	2	4	1	2	4
Time (minutes)	0.9	7.8	95.1	4.4	21.3	175.6
# Variables ($\times 10^6$)	0.37	0.56	0.95	0.20	0.39	0.78
# $\leq$ Constraints ($\times 10^6$)	0.35	0.52	0.87	0.35	0.52	0.87
# = Constraints ($\times 10^6$)	0.18	0.36	0.73	0.19	0.37	0.73

VIII. Conclusion

This paper extended the idea of multi-stage reserve policies to piecewise affine functions of the uncertainty. The number of pieces per dimension and the location of the split points is a design choice for the problem formulation. The numerical example showed that even the most obvious choice of two pieces per dimension (i.e., distinguishing between up- and down-regulating reserves), can give significant cost improvement compared to one-piece, purely affine, policies.

A corresponding dual problem was introduced, whose solution was also parameterized using piecewise affine policies. This resulted in a tractable method for computing a lower bound on dispatch costs that accounts for uncertainty. This dual problem can give a tighter bound on the suboptimality of a particular primal solution, when compared to a lower bound computed based on a deterministic problem (i.e., the prescient case). It may therefore identify a minimum possible cost arising from any *causal* redispatch plan.

In reality, power system look-ahead dispatch decisions are made repeatedly, in response to updating real-time measurements of the grid state. Future work could study how the dual-based lower bound, computed for each instance of a rolling implementation, might be used to estimate the conservativeness of the rolling re-dispatch decision process.

The piecewise affine performance bounds approach was applied to the IEEE 118-bus network to demonstrate its potential viability on large scale problems. For solving the optimization problems repeatedly, the computations times presented for the 118-bus network case study indicate that this is feasible with current solvers. To allow for larger problems to be tackled, future work could also investigate sparsity exploiting and distributed optimization techniques specific to the structure of the power system formulation presented.

References

[1] D. Bampou and D. Kuhn, "Scenario-free stochastic programming with polynomial decision rules," in *Proc. 50th IEEE Conf. Decision Control and Eur. Control Conf.*, Dec. 2011, pp. 7806–7812.

[2] A. Ben-Tal, A. Goryashko, E. Guslitzer, and A. Nemirovski, "Adjustable robust solutions of uncertain linear programs," *Math. Programming*, vol. 99, no. 2, pp. 351–376, 2004.

[3] D. P. Bertsekas, A. Nedic, and A. E. Ozdaglar, *Convex Analysis and Optimization*. Belmont, MA, USA: Athena Scientific, 2003.

[4] D. Bertsimas and V. Goyal, "On the power and limitations of affine policies in two-stage adaptive optimization," *Math. Programming*, vol. 134, no. 2, pp. 491–531, Feb. 2012.

[5] D. Bertsimas, E. Litvinov, X. A. Sun, J. Zhao, and T. Zheng, "Adaptive robust optimization for the security constrained unit commitment problem," *IEEE Trans. Power Syst.*, vol. 28, no. 1, pp. 52–63, Feb. 2013.

[6] D. Bienstock, M. Chertkov, and S. Harnett, "Chance-constrained optimal power flow: Risk-aware network control under uncertainty," *SIAM Rev.*, vol. 56, no. 3, pp. 461–495, 2014.

[7] R. D. Christie, B. F. Wollenberg, and I. Wangensteen, "Transmission management in the deregulated environment," *Proc. IEEE*, vol. 88, no. 2, pp. 170–195, Feb. 2000.

[8] S. J. Garstka and R. J.-B. Wets, "On decision rules in stochastic programming," *Math. Programming*, vol. 7, no. 1, pp. 117–143, Apr. 1974.

[9] A. Georghiou, W. Wiesemann, and D. Kuhn, "Generalized decision rule approximations for stochastic programming via liftings," *Math. Programming*, pp. 1–38, May 2014.

[10] National Grid, "Operating the electricity transmission networks in 2020—update June 2011," 2011.

[11] *"Gurobi Optimizer Reference Manual,"* Gurobi Optimization Inc., 2014.

[12] E. Guslitser, "Uncertainty-immunized solutions in linear programming," Master's thesis, Minerva Optimization Center, Technion—Israel Inst. Technol., Haifa, Israel, 2002.

[13] M. J. Hadjiyiannis, P. J. Goulart, and D. Kuhn, "An efficient method to estimate the suboptimality of affine controllers," *IEEE Trans. Autom. Control*, vol. 56, no. 12, pp. 2841–2853, Dec, 2011.

[14] D. Kuhn and P. Rocha, "A polynomial-time solution scheme for quadratic stochastic programs," *Optimal Theory Appl.*, pp. 576–589, Jan. 2013.

[15] D. Kuhn, W. Wiesemann, and A. Georghiou, "Primal and dual linear decision rules in stochastic and robust optimization," *Math. Programming*, vol. 130, no. 1, pp. 177–209, Dec. 2011.

[16] P. Pinson, "Wind energy: Forecasting challenges for its operational management," *Statistical Sci.*, vol. 28, no. 4, p. 564585, 2013.

[17] J. Warrington, P. J. Goulart, S. Mariethoz, and M. Morari, "Policy-based reserves for power systems," *IEEE Trans. Power Syst.*, vol. 28, no. 4, pp. 4427–4437, Nov. 2013.

[18] L. Xu and D. Tretheway, "Flexible ramping products," California ISO, 2012.

[19] R. D. Zimmerman, C. E. Murillo-Sánchez, and R. J. Thomas, "Matpower: Steady-state operations, planning, and analysis tools for power systems research and education," *IEEE Trans. Power Syst.*, vol. 26, no. 1, pp. 12–19, Feb. 2011.

Paul N. Beuchat received the B.Eng. degree in mechanical engineering and B.Sc. degree in physics from the University of Melbourne, Australia, in 2008, and the M.Sc. degree in robotics, systems, and control from ETH Zürich, Zürich, Switzerland, in 2014, where he is currently working toward the Ph.D. degree at the Automatic Control Laboratory .

From 2009 to 2012, he was a Subsurface Engineer with ExxonMobil, based in Melbourne, Australia. His research interests are control and optimization of large-scale systems, with a focus towards distributed control and application to power systems and coordinate flight.

Joseph Warrington (M'10) received the B.A. and M.Eng. degrees in mechanical engineering from the University of Cambridge, Cambridge, U.K., in 2008, and the Ph.D. degree fromETH Zürich, Zürich, Switzerland, in 2013.

In 2007, he held a SURF Fellowship at the California Institute of Technology. From 2008 to 2009, he was a Control Design Engineer with Wind Technologies Ltd, Cambridge, U.K. Since September 2009 he has been with the Automatic Control Laboratory, ETH Zurich, Zurich, Switzerland. His research interests include distributed and robust control and optimization, with applications in power systems and market mechanism design.

Tyler H. Summers received the B.S. degree in mechanical engineering from Texas Christian University, Fort Worth, TX, USA, in 2004, and the M.S. and Ph.D. degrees in aerospace engineering from the University of Texas in Austin, Austin, TX, USA, in 2007 and 2010, respectively.

He is an ETH Postdoctoral Fellow with the Automatic Control Laboratory, ETH Zürich, Zürich, Switzerland. He was a Fullbright Scholar with the Australian National University, Canberra, Australia, in 2007–2008. His research interests are in feedback control and optimization in complex dynamical networks, with application to electric power networks and robotic networks.

Manfred Morari (F'05) received the Diploma from ETH Zürich, Zürich, Switzerland, and the Ph.D. degree from the University of Minnesota, Minneapolis, MN, USA, both in chemical engineering.

He was appointed head of the Automatic Control Laboratory at ETH Zürich, Zürich, Switzerland, in 1994. Before that, he was the McCollum-Corcoran Professor of Chemical Engineering and Executive Officer for Control and Dynamical Systems at the California Institute of Technology. He has held appointments with Exxon and ICI plc and serves on the technical advisory boards of several major corporations. His interests are in hybrid systems and model predictive control.

Prof. Morari is a Fellow of IFAC and a member of the National Academy of Engineering. He was the recipient of the Hendrik W. Bode Lecture Prize of the IEEE Control Systems Society, the Donald P. Eckman Award and the John R. Ragazzini Award of the Automatic Control Council, the Allan P. Colburn Award and the Professional Progress Award of the AIChE, the Curtis W. McGraw Research Award of the ASEE, and Doctor Honoris Causa from Babes-Bolyai University.