Exact Power Constraints in Smart Grid Control

Klaus Trangbaek Mette Petersen Jan Bendtsen Jakob Stoustrup, Senior Member, IEEE

Abstract—This paper deals with hierarchical model predictive control (MPC) of smart grid systems. The objective is to accommodate load variations on the grid, arising from varying consumption and natural variations in the power production e.g. from wind turbines. This balancing between supply and demand is performed by distributing power to consumers in an optimal manner, subject to the requirement that each consumer receives the specific amount of energy the consumer is entitled to within a specific time horizon. However, in order to do so, the high-level controller requires knowledge of how much energy the consumers can receive within a given time horizon. In this paper, we present a method for computing these bounds as convex constraints that can be used directly in the optimisation. The method is illustrated on a simulation example that uses actual wind data as load variation, and fairly realistic consumer models. The example illustrates that the exact bounds computed by the proposed method leads to a better power distribution than a conventional, conservative approach in case of fast changes in the load.

I. INTRODUCTION

One of the greatest challenges in introducing large ratios of renewable energy into existing electric power systems, is the fluctuating and unpredictable nature of power sources that harvest energy from wind, waves and sunlight. One of the main approaches to dealing with this difficulty, is a gradual shift toward so-called "smart grid" infrastructures, where both producers and consumers are equipped with control capabilities that allow them to participate in balancing efforts, etc. [1], and where discrepancies between supply and demand can be evened out via (short-term) storage of energy [2] or by voluntarily displacing consumption in time, so-called demand-side management [3]. One way to achieve this in practise is to exploit large thermal time constants in deep freezers, refrigerators, local heat pumps etc.; extra energy can be stored during off-peak hours, and the accumulated extra cooling can then be used by turning compressors and similar devices on less frequently during peak hours—see e.g., [4] and [5].

Since power systems are multi-variable and subject to constraints, and future reference estimates are often known in advance, e.g., from 24-hour power consumption traces, weather forecasts, etc., a natural choice for the top-level controller will in most case be some sort of model-predictive controller (MPC)—see for instance [6], [7], [8] or [9].

In an earlier paper [10], the authors proposed a hierarchical control architecture for this type of system, which distributes power to consumers in such a way that the consumers participate actively in balancing external load disturbances on the grid while at the same time consuming a pre-determined amount of energy over a given time interval. However, the consumers are obviously subject to both power and energy constraints; hardware such as compressors can only consume certain amounts of power at any given time, and it is only allowed to store or release certain amounts of energy within a given interval, thus avoiding spoiling stored food in deep freezers, ensuring comfortable temperatures in housing, etc. The problem is that these constraints depend in a complicated manner on the current energy stored in the consumers, as well as how the power is distributed between them. Simple constraint estimates will therefore not allow the consumers to be used to their full potential in case of rapid load fluctuations.

In this paper, we present a method for computing non-conservative future constraints for a (possibly large) number of consumers, based on knowledge of their individual power and energy constraints. The method is illustrated on a simulation example that uses actual wind data as load variation, and fairly realistic consumer models. The example illustrates that the exact bounds computed by the proposed method leads to a more efficient power distribution than a conventional, conservative approach in case of fast changes in the load.

We begin by describing the consumers in more detail in Section II, and argue that the power and energy constraints give rise to well-defined convex polytopes. Section III presents the main contribution, a method for computing the optimisation constraints. Section IV then goes into more detail with a hierarchical predictive control scheme for smart grids. Finally, Section V demonstrates the efficiency of the constraint computation.

Our notation is mostly standard. We denote by $(\cdot)^T$ a stacked vector on a finite horizon of length $N_h$, e.g $P^T_{i,k} = [p^T_{i,k} \ p^T_{i,k+1} \ldots \ p^T_{i,k+N_h-1}]^T$.

II. RESOURCE POLYTOPES

We consider a setup of the form shown in Figure 1, where $P_{load}$ is a load imbalance that must be compensated for. The control system is able to do so by either distributing power to consumers IC$_i$, $i = 1, \ldots, N$, or by requesting external power units to produce the required amount $P_{ext}$. In most cases, it is desirable to minimise rapid changes in $P_{ext}$ by distributing these to the ICs instead. Further details about the setup can be found in [10]; for now, we concentrate on the intelligent consumers.

We consider a simple discrete time setting, where each IC can be described as an integrator with the energy storage
equation
\[ E_{i,k+1} = E_{i,k} + T_i(P_{i,k} - P_{c,i}) \]  

where \( E_i \) is the energy stored in the \( i \)-th IC, \( P_i \) is the power delivered to the IC, \( P_{c,i} \) is the (constant) power consumed, and \( T_i \) is the size of the time step. For ease of notation, we will assume \( T_i = 1 \) in the following. We further assume that \( P_{c,i} = 0 \). If this is not the case, the average consumption can be considered as part of the external load, and the IC seen as only the controllable consumption around this average.

The consumption rate and energy storage levels are limited by physical constraints:
\[ P_i \leq P_i \leq P_i, \quad E_i \leq E_i \leq E_i. \]  

In order to be of use in predictive control, these constraints must be formulated as constraints on the control variable, \( P_i \), over a finite horizon.

Starting at time step \( k \), the constraints limit the potential consumption of ICs as illustrated in Figure 2. As shown in the figure, over a finite horizon of length \( N_h > 0 \), the constraints on \( P_i \) can also be represented by a polytope in \( \mathbb{R}^{N_h} \). Clearly, the level constraints mean that the constraints at one time step depend on the consumption rate at the previous time step; that is, as long as the constraints \( E_i \leq E_i \leq E_i \) are not active, the polytope is a hypercube, but when the level constraints do become active, they ‘cut off’ convex subsets of the hypercube via intersection by hyperplanes, as indicated in the upper right part of the figure.

III. MINKOWSKI ADDITION OF RESOURCE POLYTOPES

As shown above, the feasible consumption rates are constrained to a convex polytope in \( \mathbb{R}^{N_h} \). Generally, a convex polytope can be represented by a half plane description, \( \Pi = \{ x \in \mathbb{R}^{N_h} | \Phi x \leq \gamma \} \), where \( \Phi \in \mathbb{R}^{M \times N_h} \) is a matrix, \( \gamma \in \mathbb{R}^M \) is a vector, and \( \leq \) is taken element-wise, or by a vertex description, \( \Pi = \{ \sum_{i=1}^{N_v} \alpha_i v_i | 0 \leq \alpha_i, \sum_{i=1}^{N_v} \alpha_i \leq 1 \} \), where \( v_i \) are the \( N_v \) vertex vectors of the convex polytope [11].

Tools for automatic conversion between the two representations exist, for instance MPT [12], but the computations remain challenging, so it is important to choose the most appropriate representation. In this case, the nature of the resource distribution problem leads to convex polytopes with a particular structure, which makes the vertex representation the more attractive choice. Furthermore, conversion to half-plane representation becomes simple and well-conditioned.

The Minkowski sum of a number of convex polytopes, \( \Pi_1, \ldots, \Pi_n \), is defined as the (polytopic) set of all sums of elements from the individual polygons, i.e. \( \Pi_E = \{ \sum_{i=1}^{N} x_i | x_i \in \Pi_i \} \). This is exactly what we need to compute for the ICs, in order to provide the control level with a single constraint set; given consumption capacities for a number of ICs over a horizon, the total capacity will be given by the Minkowski sum of these.

Consider the example convex polytopes in Figure 3. The horizon length is 2, and the axes represent the consumption rates in the first and second time steps. The top plots represent two ICs, the bottom plots the resulting sum. We make the following observations:

- consumption rate and capacity limits cannot simply be added to give the Minkowski sum.
- if we view some of the vertices as multiple, the convex polytopes all have the same basic shape.
- the sum can be considered as a consumer in itself with consumption rate and capacity limits, only now...
the rate limits vary over the horizon, and for \( N_h > 2 \), capacity constraints may be needed for all combinations of samples.

Computing the Minkowski sum of a large number of general polytopes is computationally demanding [13]. Given vertex representations, all possible combinations of vertices can be computed and the Minkowski sum will then be spanned by all these vertex sums. For a high number of consumers, this is not feasible; however, if the vertices can somehow be ordered in the same sequence for each consumer, then it is only necessary to perform the summation vertex by vertex. This is indeed possible for resource polytopes, by permitting some vertices to be identical.

Algorithm 1 is a recursive algorithm that produces a \( N_h \times 12N_h \) matrix \( V \), where each row is a vertex vector representing a possible combination of minimum and maximum consumption over the horizon \( N_h \). The initial resource level is assumed to be zero.

Each call to the algorithm consists of three main operations. First, variables are initialized for storage; next, we iterate through the prediction horizon and compute extreme upper bounds for each sample by assuming that the absolute minimum power has been pulled from the unit at all previous samples. Thirdly, for each sample in turn, we assume that the considered sample is at the extreme upper bound and compute new upper and lower energy bounds according to this assumption. The computation is then repeated with the new set of constraints until the end of the horizon is reach. Finally equivalent computations are made for the lower-bound vertices essentially by exchanging minimisation and maximisation.

When completed, the algorithm has generated all potential \( N_h \times 12N_h \) vertices, many of which will be identical, but all of which will be on the boundary of the convex polytope.

For optimisation purposes, a half plane representation is more suitable than the vertex representation resulting from Algorithm 1. As mentioned, the conversion can be performed automatically by existing tools, but for a large number of vertices, it becomes computationally complex. An approximate (over-bounding) but fairly accurate conversion can be performed by considering rate and capacity constraints directly. Assume that a resource polytope is spanned by vertex vectors \( v_i \), \( i = 1..N_h \), and that the \( j \)-th element of \( v_i \), \( v_{i,j} \), is the coordinate for the \( j \)-th time step. Then capacity constraints are simply found as \( E = \max_i \sum_{j=1}^{N_h} v_{i,j} \) and \( E = \min_i \sum_{j=1}^{N_h} v_{i,j} \). Time varying consumption constraints are \( \bar{P}_j = \max_i v_{i,j} \) and \( \underline{P}_j = \min_i v_{i,j} \). We then have the half plane representation

\[
\begin{bmatrix}
I & -I & \bar{w} \\
T & -T & \underline{w} \\
\end{bmatrix}
\]

where \( T \) is a Toeplitz matrix with the \( i \)-th row consisting of \( i \) ones followed by \( N_h - i \) zeros, and \( \bar{T} \) is a column vector consisting of \( N_h \) ones.

### IV. DISTRIBUTION

When controlling a large number of ICs, having a decision variable and a set of constraints for each is not computationally feasible (see [14] for more details).

To alleviate this, we introduce a number of so-called aggregators \( A_j, 1 \leq j \leq N_A \). An aggregator serves as an interface to a subset \( \mathcal{F} \) of ICs, aggregating their capacities into one constraint set. In the following we will present how to aggregate the resource polytope representations for the ICs under an aggregator’s jurisdiction.

The proposed scheme is shown in Figure 4. On both top and aggregator levels, everything is computed over a receding horizon of \( N_h \) time steps.

Starting from the lowest level, the ICs provide the aggregators with their current resource levels, \( \bar{P}_i \). From these, the
The ICs will then absorb their assigned consumption during a time step, after which the entire procedure is repeated for the new resource levels.

A. Computational burden

The main computational burdens are the vertex generation, the top level distribution and the aggregator level distribution, all of which must be performed at each time step. Summation of vertices, the top level optimisation and conversion to half plane representations are not major burdens.

The vertex generation can be performed separately for each IC and is thus easily distributed among aggregators. It could even be performed by the ICs, but then a large amount of data would need to be transferred upwards.

In the current implementation, the computational burden of performing the distribution at the top level or by an aggregator can fairly accurately be approximated by

\[ \text{distribution load: } \beta_d N_h^2 N_d^2 + \beta_h, \]

and the vertex computation at an aggregator by

\[ \text{vertex load: } \beta_v N_h^2 N_d J, \]

where \( N_d \) is the number of associated consumers or aggregators in the layer directly below. Using cpu running times as provided by Matlab simulation studies indicate \( \beta_v \approx 0.125 \beta_0 \) and \( \beta_d \approx 0.015 \beta_0 \) on a standard PC.

B. Communication load

The biggest data flow results from the vertex tables being communicated upward in the hierarchy. Resource flow profiles over the horizon are communicated downwards, but these are quite small in comparison. The vertex tables are of size \( N_h^2 N_0 \times N_j \) and an upper layer must receive one from each of its associated aggregators at each time step. If the aggregators (rather than the consumers) perform the vertex computation, then only the current resource level must be communicated from each consumer, and each consumer need only be given a consumption demand for each time step. There is no need for communication between aggregators on the same level or between consumers.

C. Mid-ranging

At times when the disturbances are steady, the top level should attempt to keep the resources at the consumers at a level that ensures wide manoeuvrability in response to future disturbances, i.e. by keeping the levels away from the capacity limits. Let \( E_{ref,i} \) denote a reference level for consumer \( i \). The desired consumption from the IC point of view is then

\[ P_{i}^* \]

aggregators compute the consumption constraints of each IC over the horizon. Using the vertex representation described in Section III, these constraints are then added to provide constraints provided to the top level.

Once the top level has received the resource polytopes it has full knowledge of the flexibility available in the system. This means that no “iterative price coordination [8]” is needed between the two upper layers of the hierarchy.

Given these constraints, the top level can then optimise a performance function using the sum of consumptions \( P_X^* = \sum_{i=1}^{N} P_{i,k}^* \) as a decision variable.

The consumption sum \( P_X^* \) is then distributed between the aggregators, which further distribute it among the ICs. The distribution should be performed in a way that ensures that constraints can be met over the entire horizon.

At the top level, a half plane representation of the constraints \( \Phi_{i,j} P_{i,j}^* \leq \gamma_{i,j} \), \( i = 1, \ldots, N_A \) is computed for each aggregator. The optimisation and distribution is then performed by solving a standard quadratic program at sample \( k \):

\[
\begin{align*}
\min_{P_{i,j}^*} & \quad \sum_{i=1}^{N_A} \left( \sum_{j=1}^{k} P_{i,j}^* - \sum_{j \in J^f} E_{i,j,k} \right)^2 \\
\text{s.t.} & \quad \Phi_{i,j} P_{i,j}^* \leq \gamma_{i,j} \quad i = 1, \ldots, N_A \\
& \quad \sum_{i=1}^{N_A} P_{i,j}^* = P_{X}^* 
\end{align*}
\]

These consumption rates can then be distributed by the aggregators among their associated ICs in a similar manner, i.e., for aggregator \( j \) at sample \( k \) we solve the optimisation problem:

\[
\begin{align*}
\min_{P_{a,j}^*} & \quad \sum_{i=1}^{k} \left( \sum_{j=1}^{i} P_{a,j}^* - E_{i,j,k} \right)^2 \quad (5) \\
\text{s.t.} & \quad \Phi_{i,j} P_{a,j}^* \leq \gamma_{i,j} \quad i = 1, \ldots, N_A \\
& \quad \sum_{i=1}^{k} P_{a,j}^* = P_{X}^* 
\end{align*}
\]

where \( \Phi_{i,j} \gamma_{i,j} \) represent constraints for the individual ICs.

Fig. 4. Control architecture with grid-level control, aggregators and intelligent consumers. At each sample instant, consumer IC \( i \) receives power \( P_i \) and returns information of its own energy storage level \( E_i \).
view is then $\tilde{E}_{i,k} = E_{ref,i} - E_{i,k}$. By adding a term taking this into account, e.g. $\tilde{P} = (\sum_{j=0}^{N_{i}} \tilde{E}_{j,k} - \sum_{i=0}^{N_{i}-1} P_{\Sigma,i,i})^2$, to the performance function at the top level, the optimisation will bring the levels closer to the references when there is no need to use the consumers in the overall balancing.

Note that the aggregators are not required to submit all energy levels of the consumers to the top level. Only the sum associated with each aggregator needs to be communicated in order to compute the total sum.

V. Simulation example

The simulation example considers a power grid consisting of a wind farm, a set of intelligent consumers (ICs), namely heat pumps and refrigeration systems, a set of regular consumers (RCs) and a power plant.

The wind farm and the power plant are the production capacities. The RCs must be supplied a constant power at all times. The average production of the wind farm and power plant thus corresponds to the base consumption of the RCs. The wind farm production however exhibits fast fluctuations, which must be balanced by the ICs and if necessary by the power plant.

The ICs at the lowest level are modeled as simple power and energy constraint units governed by their own energy balance as explained in Section II.

The top level controller has three objectives. As explained earlier the energy supplied to the RCs is the average production of the wind farm and power plant. Fluctuations of the wind power production however must be balanced by the ICs and power plant. The first objective is therefore to keep the energy balance

$$E_{k+1} = E_k + P_{i}(P_{PLANT,k} - P_{WIND,k} - P_{IC,k})$$

close to zero, where $P_{IC,k} = \sum_{i=0}^{N_{i}} P_{i,k}$ is the power absorbed by the ICs, $P_{WIND}$ are fluctuations of the wind farm power production and $P_{PLANT}$ is the power plant deviation from the base line production.

The second objective is to reduce the strain on the power plant by limiting the size of changes in production and the last objective of the top level controller is to bring the ICs close to their reference levels. With these three objectives the optimisation problem at the top level for a prediction horizon $N_h$ is

$$\min_{P_{PLANT,k}, P_{WIND,k}} \sum_{k=1}^{N_h} z_{k+1} = P_{PLANT,k} - P_{WIND,k}$$

where

$$z_{k} = \begin{bmatrix} E_k \\ P_{PLANT,k} - P_{WIND,k} \\ P_{PLANT,k} - P_{IC,k} \end{bmatrix}$$

and $Q_{i} = diag(\tilde{P}_{IC,i}, \tilde{P}_{IC,i}, \tilde{P}_{IC,i})$ for $i = 1, 2, \ldots, N_h - 1$. The terminal weight $Q_{N_h}$ is found by standard dual mode MPC methods [7]. The penalty on $P_{PLANT} - P_{WIND}$ is included to improve closed loop performance.

The wind farm power production, which has to be balanced in the simulations, consist of production data from Horns Rev 2, a 209 MW offshore wind farm in the North Sea owned and operated by DONG Energy. The period covered extends from 15:00 to 22:00 of February 1st 2011 and the average production value has been subtracted, since it is used to supply the RCs.

Two types of ICs are considered, namely a domestic heat pump and a supermarket refrigeration system. The ICs are modeled as simple energy and power constraint units, so the internal dynamic is not modeled in detail. These ICs each have a primary purpose, which must be met, namely to keep the house and frozen goods within a certain acceptable temperature interval. The main mediums, which the heat pump and cooling system must respectively heat and cool, are the concrete floor of a single family home and the freezer content. The thermal energy resource available is thus given by $C \cdot V \cdot \Delta T$, where $V$ and $C$ are the volume and volumetric heat capacity of the main medium and $\Delta T$ is the acceptable temperature interval. The heat pump is assumed to have a coefficient of performance of 3.0, so the electrical energy resource is one third of the thermal resource. Additional parameters are $\tilde{P}_{IC} = 1$, $\tilde{P}_{PLANT} = 2$, $\tilde{P}_{PLANT} = 0.002$, $\tilde{P}_{PLANT} = 0.1$ and $N_h = 4$, corresponding to

$$Q_{N_h} = \begin{bmatrix} 7.4 & -1.7 & 0 \\ -1.7 & 2.1 & 0 \\ 0 & 0 & 10 \end{bmatrix}$$

Four aggregators are included in the simulations and each aggregator handles 400 heat pumps and 400 refrigeration systems. The time step is 15 minutes and $P_{WIND}$ is assumed known over the horizon $N_h$.

The advantage of the resource polytopes is, that the full flexibility of the ICs are communicated to the top level. To illustrate the impact of this, the method is compared to a setup, denoted the cautious method, where the interface between top- and aggregator level does not allow constraints to vary over the horizon $N_h$.

Everything at the top- and aggregator level is computed over a horizon of $N_h$ time steps. Given a horizon of $N_h$ at time step $k$ the cautious aggregators determine $P_{PLANT, \text{Cautious}} = \min(P_{PLANT}, \frac{E_{ref,k}}{N_h})$ and $P_{PLANT, \text{Cautious}} = \max(P_{PLANT}, \frac{E_{ref,k}}{N_h})$. These values are communicated to the top level, which should provide at most $P_{PLANT, \text{Cautious}}$ and at least $P_{PLANT, \text{Cautious}}$ to each IC at each time step over the next prediction horizon. This insures that the energy constraints of the IC’s are not violated.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>10 kW</td>
<td>4.3 kW</td>
<td></td>
</tr>
<tr>
<td>1 K</td>
<td>3 K</td>
<td></td>
</tr>
<tr>
<td>7 kWh</td>
<td>8 kWh</td>
<td></td>
</tr>
<tr>
<td>3 kW</td>
<td>1.3 kW</td>
<td></td>
</tr>
<tr>
<td>-7 kWh</td>
<td>-3 kW</td>
<td></td>
</tr>
</tbody>
</table>

TABLE I
PARAMETERS FOR THE CONSIDERED ICs.
Simulation results are give in Figure 5, which depicts $P_{\text{wind}}$, $P_{\text{plant}}$, $E_{\text{Cautious}}$ and $E_{\text{Resource Polytopes}}$. The two methods have to balance the exact same oscillations with the exact same resource available, but the resource polytope method is noticeably better at maintaining the energy balance. The reason for this can be seen in Figure 6. For the relatively slow heat pump the performance of the two methods is as expected quite similar. For the faster refrigeration system, however, the polytope method is able to get closer $\overline{P}$, when needed, which means a better utilisation of the flexibility.

VI. DISCUSSION

We have presented a novel way to represent resource storage capacity, which has the useful properties that:

- the aggregated constraints of a set of consumers can be computed without approximation by a simple summation,
- conversion to half plane representations, useful for optimisation, can be performed at low cost.

Since the constraint aggregation is exact, the scheme is nearly optimal. With respect to the distribution, it is possible that, in terms of future flexibility, a slightly better distribution could be obtained by a direct distribution at the top level, but for a high number of consumers this does not seem feasible to implement.

REFERENCES