

Towards price-based predictive control of a small scale electricity network

P. Braun^{a,b,*}, L. Grüne^b, C. M. Kellett^a, S. R. Weller^a, K. Worthmann^c

^a*School of Electrical Engineering and Computer Science, University of Newcastle, Callaghan NSW 2308, Australia;* ^b*Mathematical Institute, Universität Bayreuth, 95440 Bayreuth, Germany;* ^c*Institute for Mathematics, Technische Universität Ilmenau, 99693 Ilmenau, Germany*

(Received 00 Month 20XX; accepted 00 Month 20XX)

With the increasing deployment of battery storage devices in residential electricity networks, it is essential that the charging and discharging of these devices be scheduled so as to avoid adverse impacts on the electricity distribution network. In this paper, we propose a non-cooperative, price-based hierarchical distributed optimisation approach that provably recovers the centralised, or cooperative, optimal performance from the point of view of the network operator. The performance of this algorithm is validated on recent measurements taken from an Australian electricity distribution company, Ausgrid.

Keywords: Distributed Optimisation, Gradient Ascent Method, Smart Grid Battery Scheduling, Real-Time Pricing

1. Introduction

As residential battery storage becomes increasingly common, particularly in conjunction with small-scale renewable energy sources such as rooftop solar photovoltaics, it is clear that uncoordinated charging and or discharging of a large number of batteries can lead to undesirable behaviour from the point of view of an electricity distribution company. In particular, if charging occurs across all resources simultaneously an artificially high supply must be maintained. Similarly, if discharging occurs simultaneously the grid must absorb an unnecessarily large amount of energy.

In (Worthmann, Kellett, Braun, Grüne, & Weller, 2015; Braun, Grüne, Kellett, Weller, & Worthmann, 2016) we investigated several model predictive control (MPC) schemes for the control of a small scale electricity network and demonstrated the performance of these schemes using a dataset provided by the company Ausgrid (see (Ratnam, Weller, Kellett, & Murray, 2015) for an overview of the dataset). The control schemes in (Worthmann et al., 2015; Braun, Grüne, et al., 2016) concentrated on reducing fluctuations of the aggregate power demand and on temporary operation of the electricity network in islanded mode (see (Braun, Faulwasser, et al., 2016)). The control algorithms range from centralised approaches to cooperative and non-cooperative distributed algorithms. Whereas for the cooperative distributed approaches convergence to the centralised control setting was shown in (Braun, Grüne, et al., 2016; Braun, Faulwasser, et al., 2016), convergence of the non-cooperative algorithms introduced in (Worthmann et al., 2015) and (Braun, Grüne, Kellett, Weller, & Worthmann, 2015) remains an open question.

In this paper we close this gap by introducing a negotiation algorithm between an energy provider and its customers to compute real-time electricity prices. The algorithm is based on the dual ascent algorithm (see (Bertsekas, 1999), for example) from which convergence can be concluded. Residential customers send their predicted power demand for the next several hours to the energy provider.

*Corresponding author. Email: philipp.braun@newcastle.edu.au

The energy provider communicates real-time prices to its customers which then can change their consumption patterns by using storage devices and update their predicted energy demand. Thus, the negotiation process continues until an equilibrium, which minimises the electricity costs and the fluctuations in the power demand, is found.

Herein we consider the need for prices to correspond to an essentially non-cooperative scheme, whereby customers act purely in order to minimise their costs (or maximise their profits). By contrast, we consider the centralised optimisation problem, and its distributed variant, to be essentially cooperative as customers either blindly implement the centralised control solution or solve an optimisation problem that benefits the network.

An approach for real-time electricity prices with a similar motivation (i.e., peak reduction and cost minimisation) is proposed in (Mohsenian-Rad & Leon-Garcia, 2010). There, prices are set by the energy provider without a negotiation process and the customer minimises its electricity costs by anticipating future electricity prices. In (Samadi, Mohsenian-Rad, Schober, Wong, & Jatskevich, 2010), a similar dual decomposition algorithm is used to create time-varying adaptive prices. However, a different smart grid model is used without local storage devices. Instead, it is assumed that at every time instant $k \in \mathbb{N}$ customers have a lower and an upper bound on their power demand. Then customers maximise their welfare by maximising their power demand (based on a strictly increasing utility function) minus the costs for electricity. In this case, there is no coupling between two consecutive time steps k and $k + 1$ and every time instant can be considered independently. Thus, local storage devices cannot be handled with this approach.

In (Vytelingum, Voice, Ramchurn, Rogers, & Jennings, 2010), a game theoretic approach for real-time pricing is presented. The authors consider micro-storage devices, similar to the batteries used in this paper, for peak-shaving and cost reduction. In contrast to our approach, customers optimise their battery usage based on predicted prices without a negotiation process with the energy provider. The authors in (Yang, Chavali, & Nehorai, 2012) optimise the schedule for controllable loads instead of the optimal usage of batteries. In this reference, a different distributed optimisation algorithm with the same communication structure as the one proposed in our work is used. The approach in (Yang et al., 2012) is not embedded in a receding horizon scheme. In (Floch, Belletti, Saxena, Bayen, & Moura, 2015), a similar distributed dual ascent algorithm in the context for the optimal charging of electric vehicles is considered. The algorithm, however, is not used to model real-time electricity prices and the algorithm is not used in an MPC context. In the context of MPC, dual decomposition algorithms for quadratic separable cost functions are considered in (Doan, Keviczky, Necoara, Diehl, & Schutter, 2009; Doan, Keviczky, & Schutter, 2011; Giselsson, Doan, Keviczky, Schutter, & Rantzer, 2013; Giselsson & Rantzer, 2010), for example.

Recently, in (Gharesifard, Basar, & Dominguez-Garcia, 2016), a continuous-time game theoretic framework is proposed. Similar to our work, (Gharesifard et al., 2016) proposes a central entity (an aggregator) though, with a continuous-time system, there is not a negotiation period, but rather a leader-follower structure where the central entity sets a price curve and the residences employ a distributed gradient-flow algorithm to find the Nash equilibrium. As a further structural difference with the approach proposed herein, in (Gharesifard et al., 2016) residences communicate amongst themselves (not necessarily all-to-all communication), whereas our proposed structure has each residence communicating only with the central entity.

This paper is structured as follows. In Section 2 the model of the small scale electricity network is introduced and the control problem is formulated. In Section 3, the dual ascent algorithm is reviewed and convergence properties of the algorithm are provided. Furthermore, it is shown how the algorithm can be implemented in a distributed way. In Section 4 we show how the the distributed dual ascent algorithm can be applied to the electricity network. In particular we show how the distributed dual ascent algorithm can be used to obtain real-time electricity prices and how convergence of the algorithm can be ensured. Additionally we explain how the algorithm can be embedded in a receding horizon context and discuss its differences to centralised and cooperative distributed control schemes. In Section 5 the open loop and the closed loop performance of the

dual ascent algorithm using the pricing scheme is demonstrated by means of numerical simulations. In particular, we investigate the performance compared to a cooperative control scheme and the impact of the number of storage devices in the electricity network before the paper is concluded in Section 6.

2. Model description and problem fomulation

We consider a dynamical system consisting of \mathcal{I} residential energy systems (RESs), $\mathcal{I} \in \mathbb{N}$, which was first introduced in (Ratnam, Weller, & Kellett, 2013) and extended in (Braun, Faulwasser, et al., 2016). Each RES owns storage devices and has positive and negative loads which result from local generators and residential appliances. The i -th RES, $i \in \mathbb{N}_{\mathcal{I}} = \{1, \dots, \mathcal{I}\}$, is governed by the system dynamics

$$x_i(k+1) = x_i(k) + T (\beta_i u_i^+(k) + u_i^-(k)) \quad (1)$$

$$z_i(k) = w_i(k) + u_i^+(k) + \gamma_i u_i^-(k) \quad (2)$$

with parameters $(\beta_i, \gamma_i) \in (0, 1]^2$ and time index $k \in \mathbb{N}$. The variables $z_i(k)$ [kW], representing the power demand drawn from/supplied to the grid by each RES and given by Equation (2), depend on the net consumptions $w_i(k)$ [kW], i.e., the power generation of solar photovoltaic panels minus the energy demand of the residents, and the battery usage. The dynamics of the battery, Equation (1), are defined through the state of charge $x_i(k)$ in [kWh] and the input variables $u_i^+(k)$ and $u_i^-(k)$ in [kW]. The input $u_i^+(k)$ represents the power drawn from the grid, while $u_i^-(k)$ can be used to discharge the battery and, thus, decrease the power demand. The parameters β_i and γ_i model losses due to energy conversion, and T (in hours [h]) represents the length of the sampling interval.

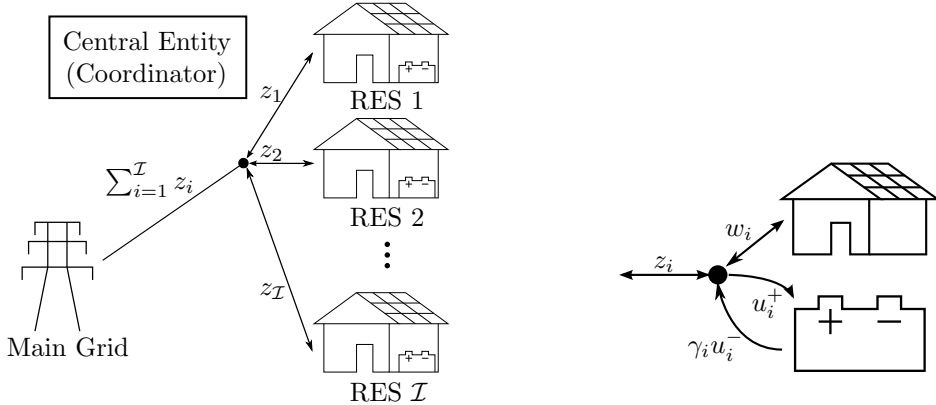


Figure 1. Visualization of the electricity network (left) and visualisation of Equation (2) (right). The power demand z_i depends on the net power consumption w_i and can be manipulated by using the battery.

The storage device of each RES is constrained by

$$0 \leq x_i(k) \leq C_i \quad (3a)$$

$$\underline{u}_i \leq u_i^-(k) \leq 0 \quad (3b)$$

$$0 \leq u_i^+(k) \leq \bar{u}_i \quad (3c)$$

$$0 \leq \frac{u_i^-(k)}{\underline{u}_i} + \frac{u_i^+(k)}{\bar{u}_i} \leq 1. \quad (3d)$$

The constraint (3a) sets an upper and a lower bound on the capacity of the batteries, the constraints (3b) and (3c) define upper and lower bounds on the charging rate of the batteries and (3d) ensures

that the bounds on the charging rate are also respected when the batteries are used for charging and discharging in the time interval from k to $k + 1$.

Each RES is connected with the main grid at the point of common coupling through its individual power demand z_i , $i \in \mathbb{N}_{\mathcal{I}}$. The network wide power demand $\sum_{i=1}^{\mathcal{I}} z_i$ is covered by the main grid. The network is coordinated by the central entity (CE). The CE encompasses the functions of a grid operator (coordinator) and an energy provider. Since the CE is responsible for a stable operation of the electricity grid without outages we are interested in minimising the deviation from the average power demand

$$\bar{\zeta}(k) = \frac{1}{N} \sum_{j=k}^{k+N-1} \frac{1}{\mathcal{I}} \sum_{i=1}^{\mathcal{I}} w_i(j) \quad (4)$$

over a fixed prediction horizon $N \in \mathbb{N}$. In other words, we assume that the CE is interested in the optimal solution of the minimisation problem

$$\begin{aligned} \min_{u_i} \quad & \frac{1}{N} \sum_{j=k}^{k+N-1} \left(\frac{1}{\mathcal{I}} \sum_{i=1}^{\mathcal{I}} z_i(j) - \bar{\zeta}(k) \right)^2 \\ \text{s.t.} \quad & x_i(k) = x_i^0 \text{ and (1), (2), (3)} \\ & \forall i = 1, \dots, \mathcal{I}, \forall j = k, \dots, k + N - 1 \end{aligned} \quad (5)$$

for given states of charge of the battery $x_i^0 \in [0, C_i]$ and fixed predictions of the net consumption $w_i(k), \dots, w_i(k + N - 1)$ for $i = 1, \dots, \mathcal{I}$. Optimal solutions of the optimisation problem (5) are denoted with a \cdot^* ; i.e., $u_i^*(k), \dots, u_i^*(k + N - 1)$ and $z_i^*(k), \dots, z_i^*(k + N - 1)$ for $i = 1, \dots, \mathcal{I}$.

Since the net consumption $w_i(k)$ can only be approximated up to some time into the future, model predictive control (MPC) offers a natural tool to compute a feedback control for the charging and discharging of the batteries. To this end, the finite dimensional optimisation problem (5) is solved at every time step k and the first piece of the optimal solution $u_i^{\text{cl}}(k) = u_i^*(k)$, $i = 1, \dots, \mathcal{I}$, is used as a control value before the time index k is shifted and the process is repeated. Analogously, the power demand of the closed loop solution is denoted by $z_i^{\text{cl}}(k)$ for $i = 1, \dots, \mathcal{I}$ and $k \in \mathbb{N}$. To simplify the notation we use $\bar{\cdot}$ to denote the average with respect to the number of systems, i.e., we define $\bar{z}(k) = \frac{1}{\mathcal{I}} \sum_{i=1}^{\mathcal{I}} z_i(k)$ to denote the average power demand at time k .

To measure the performance of the MPC feedback over a simulation of length $\mathcal{N} \in \mathbb{N}$ we use the following performance metrics:

- The first measure reflects the mean-quadratic-deviation (MQD) from the average:

$$\frac{1}{\mathcal{N}} \sum_{k=0}^{\mathcal{N}-1} \left(\bar{z}^{\text{cl}}(k) - \bar{\zeta}^{\text{cl}} \right)^2, \quad (\text{MQD})$$

$$\frac{1}{N} \sum_{j=0}^{N-1} \left(\bar{z}^*(k+j) - \bar{\zeta}(k) \right)^2, \quad (\text{MQD}_k)$$

where $\bar{\zeta}^{\text{cl}} := \frac{1}{\mathcal{N}} \sum_{k=0}^{\mathcal{N}-1} \bar{z}^{\text{cl}}(k)$ denotes the overall average power demand of the MPC simulation.

- Peak-to-Peak (PTP), i.e., the maximal variation in the demand profile over the *simulation*

interval in closed and open loop:

$$\max_{k=0,\dots,N-1} \{\bar{z}^{\text{cl}}(k)\} - \min_{k=0,\dots,N-1} \{\bar{z}^{\text{cl}}(k)\}, \quad (\text{PTP})$$

$$\max_{j=0,\dots,N-1} \{\bar{z}^*(k+j)\} - \min_{j=0,\dots,N-1} \{\bar{z}^*(k+j)\}. \quad (\text{PTP}_k)$$

To compute the feedback $(u_i^{\text{cl}}(k))_{k \in \mathbb{N}}$, $i \in \mathbb{N}_{\mathcal{I}}$, the optimisation problem (5) must be solved at every time step $k \in \mathbb{N}$. Thus, we introduce the dual ascent algorithm in the next section in order to solve an optimisation problem at a fixed time instant $k \in \mathbb{N}$ on a time window of length $N \in \mathbb{N}$. Afterwards, we discuss the convergence properties of this algorithm. Then, in Section 4, we return to the electricity network application.

The following abbreviations are used in order to simplify the notation:

- $\mathbf{z}_i = (z_i(k), \dots, z_i(k+N-1))^T \in \mathbb{R}^N$ for the *power demand* of the i -th system, $i \in \mathbb{N}_{\mathcal{I}}$, on the considered time window,
- $\bar{\mathbf{z}} = \frac{1}{T} \sum_{i=1}^T \mathbf{z}_i \in \mathbb{R}^N$ for the average *power demand* on the considered time window, and
- $\mathbf{z} = (\mathbf{z}_1^T, \dots, \mathbf{z}_{\mathcal{I}}^T)^T \in \mathbb{R}^{N\mathcal{I}}$.

Similar notation is used for other variables, e.g. \mathbf{x} and \mathbf{w} . Additionally we use the notation $\mathbf{1} = (1, \dots, 1)^T$ for a vector of appropriate size and I is used for the identity matrix of appropriate dimension. As defined at the beginning of this section, $\mathbb{N}_{\mathcal{I}} = \{1, \dots, \mathcal{I}\}$ denotes the RESs from 1 to \mathcal{I} . For a given net consumption \mathbf{w}_i and an initial state of charge x_i^0 , the feasible set of solutions \mathbf{z}_i is

$$\mathbb{D}_i = \left\{ \mathbf{z}_i \in \mathbb{R}^N \left| \begin{array}{l} x_i(k) = x_i^0 \\ \text{System dynamics (1) and (2)} \\ \text{Constraints (3)} \\ \forall j \in \{k, \dots, k+N-1\} \end{array} \right. \right\}. \quad (6)$$

The set \mathbb{D} is defined in the same way for the variable \mathbf{z} (instead of \mathbf{z}_i) with $\mathbb{D}_1 \times \dots \times \mathbb{D}_{\mathcal{I}}$. Note that the feasible sets \mathbb{D}_i , $i \in \mathbb{N}_{\mathcal{I}}$, and \mathbb{D} are compact and convex (polyhedral).

3. The dual ascent method

3.1 The Lagrangian and duality

At a fixed time step $k \in \mathbb{N}$ the optimisation problem (5) can be written in the form

$$\begin{aligned} \min_{\mathbf{y} \in D} \quad & \phi(\mathbf{y}) \\ \text{s.t.} \quad & 0 = \chi(\mathbf{y}) := A\mathbf{y} - b. \end{aligned} \quad (\text{P})$$

with a convex function $\phi: \mathbb{R}^n \rightarrow \mathbb{R}$, a convex set $D \subset \mathbb{R}^n$, a matrix $A \in \mathbb{R}^{m \times n}$ and a vector $b \in \mathbb{R}^m$ defining linear equality constraints $\chi: \mathbb{R}^n \rightarrow \mathbb{R}^m$. The Lagrangian $\mathcal{L}: \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}$ of the primal problem (P) is defined as

$$\mathcal{L}(\mathbf{y}, \lambda) = \phi(\mathbf{y}) + \lambda^T (A\mathbf{y} - b). \quad (7)$$

The variables $\lambda \in \mathbb{R}^m$ are called Lagrange multipliers. Given the primal problem (P) and the Lagrangian (7), we can define the dual function $\psi : \mathbb{R}^m \rightarrow \mathbb{R} \cup \{-\infty\}$,

$$\psi(\lambda) = \inf_{\mathbf{y} \in D} \mathcal{L}(\mathbf{y}, \lambda), \quad (8)$$

and the dual problem

$$\max_{\lambda \in \mathbb{R}^m} \psi(\lambda). \quad (D)$$

Observe that ψ is defined as an extended real-valued function since the infimum does not need to be finite for all $\mathbf{y} \in D$. All definitions and results in the remainder of this section concern the primal (P) and the dual problem (D). Thus we assume that the assumptions and notations used in Section 3.1 hold throughout Section 3.

The primal problem (P), the Lagrangian (7), and the dual problem (D) are related through the following results which allow a characterization of the optimal solution of the primal problem (P) based on the solution of the dual (D), and vice-versa. These results and the corresponding proofs can, for example, be found in (Bertsekas & Tsitsiklis, 1989, Appendix C).

Theorem 1 (Duality theorem): *If the primal problem (P) has an optimal solution \mathbf{y}^* , then the dual problem (D) has an optimal solution λ^* and the optimal values are equal; i.e.,*

$$\phi(\mathbf{y}^*) = \phi^* = \psi^* = \psi(\lambda^*).$$

Theorem 2 (Saddle point theorem): *For the primal problem (P) and the dual problem (D), the following holds: \mathbf{y}^* is optimal for the primal problem (P) and λ^* is optimal for the dual problem (D) if and only if the saddle point condition*

$$\mathcal{L}(\mathbf{y}^*, \lambda) \leq \mathcal{L}(\mathbf{y}^*, \lambda^*) \leq \mathcal{L}(\mathbf{y}, \lambda^*) \quad (9)$$

holds for all $\mathbf{y} \in D$ and for all $\lambda \in \mathbb{R}^m$.

Hence, with these two theorems, if an optimal solution of the dual problem is known, a solution of the primal problem can be computed based on the Lagrangian by solving the minimisation problem

$$\mathbf{y}^* \in \operatorname{argmin}_{\mathbf{y} \in D} \mathcal{L}(\mathbf{y}, \lambda^*). \quad (10)$$

The advantage of the minimisation problem (10) over the original primal problem (P) is that the constraints $\chi(\mathbf{y}) = 0$ are incorporated into the cost function. Likewise, if \mathbf{y}^* is known an optimal solution of the dual problem is obtained by solving the unconstrained problem

$$\lambda^* \in \operatorname{argmax}_{\lambda \in \mathbb{R}^m} \mathcal{L}(\mathbf{y}^*, \lambda). \quad (11)$$

Moreover, if the function ϕ is strictly convex and D is convex and compact, the primal problem has a unique optimal solution \mathbf{y}^* (see (Bertsekas, 1999, Proposition B.10) for example) and thus, the dual problem also has an optimal solution λ^* . We exploit these properties in the dual ascent algorithm introduced next.

3.2 Definition and convergence of the dual ascent method

Instead of solving the primal optimisation problem (P) directly, we search for an optimal solution of the dual problem (D) from which an optimal solution of the primal problem is obtained. If the dual function ψ is known, the dual problem (D) can be solved by a gradient ascent method. The key idea is to find a sequence $(\lambda^\ell)_{\ell \in \mathbb{N}}$ such that $\psi(\lambda^{\ell+1}) > \psi(\lambda^\ell)$ holds for all $\ell \in \mathbb{N}$. Typically, directions $(d^\ell)_{\ell \in \mathbb{N}}$ satisfying the ascending condition

$$\nabla\psi(\lambda^\ell)^T d^\ell > 0 \quad (12)$$

are computed to fulfill this task; e.g., $d^\ell = \nabla\psi(\lambda^\ell)$ is a possible choice. Then, the gradient method is defined by

$$\lambda^{\ell+1} = \lambda^\ell + c^\ell d^\ell \quad (13)$$

using a sequence $(c^\ell)_{\ell \in \mathbb{N}}$ of suitable stepsizes. Here, we distinguish between the following two cases:

- **Constant stepsize** $c > 0$: $c^\ell = c$ for all $\ell \in \mathbb{N}$.
- **Diminishing stepsize**: a sequence $(c^\ell)_{\ell \in \mathbb{N}}$ satisfying

$$c^\ell \geq 0, \quad c^\ell \xrightarrow{\ell \rightarrow \infty} 0 \quad \text{and} \quad \sum_{\ell=0}^{\infty} c^\ell = \infty, \quad (14)$$

e.g. the harmonic numbers $c^\ell = 1/(\ell + 1)$.

Convergence of the gradient method can be proven under Lipschitz continuity of the gradient $\nabla\psi$ (Bertsekas, 1999, Proposition 1.2.3 and Proposition 1.2.4).

Theorem 3 (Convergence of the dual ascent method): *Assume the existence of a Lipschitz constant $L > 0$ such that*

$$\|\nabla\psi(\lambda_1) - \nabla\psi(\lambda_2)\| \leq L\|\lambda_1 - \lambda_2\|$$

holds for all $\lambda_1, \lambda_2 \in \mathbb{R}^m$. Let $(\lambda^\ell)_{\ell \in \mathbb{N}}$ be generated according to (13) with ascent directions $(d^\ell)_{\ell \in \mathbb{N}}$ satisfying (12). In addition, let one of the following two conditions be satisfied.

- **Constant stepsize**: *There exists $\varepsilon > 0$ such that*

$$\varepsilon \leq c^\ell \leq (2 - \varepsilon) \cdot \frac{|\nabla\psi(\lambda^\ell)^T d^\ell|}{L\|d^\ell\|^2}$$

for all $\ell \in \mathbb{N}$.

- **Diminishing stepsize**: *There exist positive scalars c_1, c_2 such that the stepsize $(c^\ell)_{\ell \in \mathbb{N}}$ satisfies (14) and the ascent directions $(d^\ell)_{\ell \in \mathbb{N}}$ satisfy*

$$c_1\|\nabla\psi(\lambda^\ell)\|^2 \leq \nabla\psi(\lambda^\ell)^T d^\ell, \quad \|d^\ell\|^2 \leq c_2\|\nabla\psi(\lambda^\ell)\|^2.$$

Then, either $\psi(\lambda^\ell) \rightarrow -\infty$ for $\ell \rightarrow \infty$ or $(\psi(\lambda^\ell))_{\ell \in \mathbb{N}}$ converges to a finite value and $\nabla\psi(\lambda^\ell) \rightarrow 0$ for $\ell \rightarrow \infty$. Furthermore, every accumulation point of $(\lambda^\ell)_{\ell \in \mathbb{N}}$ is a stationary point of ψ .

3.3 The dual ascent algorithm

Before we present the dual ascent algorithm, we show first how an ascent direction can be computed. Since a closed form expression of the dual function ψ and its derivative $\nabla\psi$ is generally not known, we use the Lagrangian to compute $\nabla\psi(\lambda^\ell)$ for a given Lagrange multiplier $\lambda^\ell \in \mathbb{R}^m$. The corresponding results read as follows (Bertsekas, 1999, Prop. 6.1.1).

Theorem 4: *Let $\phi : \mathbb{R}^n \rightarrow \mathbb{R}$ be strongly convex and suppose that $\mathbf{y}_\lambda \in D$ is the unique minimiser of the Lagrangian $\mathcal{L}(\cdot, \lambda)$ for a given $\lambda \in \mathbb{R}^m$ over the convex and closed set D . Then, the dual function $\psi : \mathbb{R}^m \rightarrow \mathbb{R} \cup \{-\infty\}$ is continuously differentiable and satisfies*

$$\nabla\psi(\lambda) = A\mathbf{y}_\lambda - b.$$

To show the Lipschitz-continuity of $\nabla\psi$ required for the convergence statement in Theorem 3, a few definitions and auxiliary results are needed.

Definition 1: For the dual function $\psi : \mathbb{R}^m \rightarrow \mathbb{R} \cup \{-\infty\}$, we define a restriction D_ψ of the domain of ψ as

$$D_\psi := \{\lambda \in \mathbb{R}^m \mid \psi(\lambda) > -\infty\}.$$

Based on Definition 1, concavity of the dual function can be shown, cf. (Bertsekas, 1999, Proposition 5.1.2).

Proposition 1: *The restricted domain D_ψ of the dual function $\psi : \mathbb{R}^m \rightarrow \mathbb{R} \cup \{-\infty\}$ is convex and ψ is concave on D_ψ .*

Proof. Let $\lambda_1, \lambda_2 \in D_\psi$, and $\beta \in [0, 1]$ be given. Then, for all $\mathbf{y} \in D$, the linearity of the Lagrangian with respect to the Lagrange multiplier λ implies that

$$\mathcal{L}(\mathbf{y}, \beta\lambda_1 + (1 - \beta)\lambda_2) = \beta\mathcal{L}(\mathbf{y}, \lambda_1) + (1 - \beta)\mathcal{L}(\mathbf{y}, \lambda_2).$$

Taking the minimum on both sides leads to the estimate

$$\min_{\mathbf{y} \in D} \mathcal{L}(\mathbf{y}, \beta\lambda_1 + (1 - \beta)\lambda_2) \geq \min_{\mathbf{y} \in D} \beta\mathcal{L}(\mathbf{y}, \lambda_1) + \min_{\mathbf{y} \in D} (1 - \beta)\mathcal{L}(\mathbf{y}, \lambda_2)$$

which by definition is equivalent to

$$\psi(\beta\lambda_1 + (1 - \beta)\lambda_2) \geq \beta\psi(\lambda_1) + (1 - \beta)\psi(\lambda_2);$$

i.e., ψ is concave. Furthermore, for $\lambda_1, \lambda_2 \in D_\psi$, the last inequality implies $\psi(\beta\lambda_1 + (1 - \beta)\lambda_2) > -\infty$, and thus, $\beta\lambda_1 + (1 - \beta)\lambda_2 \in D_\psi$ for all $\beta \in [0, 1]$, which shows convexity of D_ψ and completes the proof. \square

With the help of Theorem 4, an ascent direction can be computed by first minimising the Lagrangian for a fixed λ and then evaluating the function χ .

Remark 1: Sufficient conditions to ensure uniqueness of the minimiser of the Lagrangian $\mathcal{L}(\cdot, \lambda)$ for a fixed λ are that the function ϕ is strongly convex on a closed convex domain or that the function ϕ is strictly convex on a compact and convex domain, for example. A twice continuously differentiable function $\phi : D \rightarrow \mathbb{R}$ is strongly convex with parameter $\alpha > 0$ if and only if $\nabla^2\phi(\mathbf{y}) - \alpha I$ is positive semidefinite for all $\mathbf{y} \in D$ (Bertsekas, 1999, Proposition B.5).

It remains to show that $\nabla\psi$ is Lipschitz continuous. To this end, the following general property of projections is required (Bertsekas, 1999, Proposition 2.1.3).

Theorem 5 (Projection Theorem): *Let the set $D \subset \mathbb{R}^n$ be non-empty, closed, and convex.*

(a) *For $\mathbf{y} \in \mathbb{R}^n$, there uniquely exists $[\mathbf{y}]^+ \in D$ satisfying*

$$[\mathbf{y}]^+ = \underset{\mathbf{v} \in D}{\operatorname{argmin}} \|\mathbf{v} - \mathbf{y}\|.$$

The vector $[\mathbf{y}]^+$ is called the projection of \mathbf{y} on D .

(b) *The mapping $\eta: \mathbb{R}^n \rightarrow D$ defined by $\eta(\mathbf{y}) = [\mathbf{y}]^+$ is continuous and nonexpansive; i.e.,*

$$\|[\mathbf{y}_1]^+ - [\mathbf{y}_2]^+\| \leq \|\mathbf{y}_1 - \mathbf{y}_2\| \quad \forall \mathbf{y}_1, \mathbf{y}_2 \in \mathbb{R}^n.$$

Theorem 6: *Consider the primal problem (P) with non-empty, closed and convex set D . Assume that ϕ is continuously differentiable and strongly convex with parameter $\alpha > 0$; i.e., for all $\mathbf{y}_1, \mathbf{y}_2 \in D$,*

$$\alpha \|\mathbf{y}_1 - \mathbf{y}_2\|^2 \leq (\nabla\phi(\mathbf{y}_1) - \nabla\phi(\mathbf{y}_2))^T (\mathbf{y}_1 - \mathbf{y}_2). \quad (15)$$

Then, the dual ψ is concave and $\nabla\psi$ is Lipschitz continuous with Lipschitz constant $L := \|A\|^2 / \alpha$.

Proof. The following proof is based on ideas presented in (Cheng, 1987). Let $\mathbf{y}_1, \mathbf{y}_2 \in D$, $\mathbf{y}_1 \neq \mathbf{y}_2$. Since ϕ is assumed to be differentiable and strongly convex, inequality (15) and the Cauchy-Schwartz inequality yield

$$\alpha \|\mathbf{y}_1 - \mathbf{y}_2\|^2 \leq (\nabla\phi(\mathbf{y}_1) - \nabla\phi(\mathbf{y}_2))^T (\mathbf{y}_1 - \mathbf{y}_2) \leq \|\nabla\phi(\mathbf{y}_1) - \nabla\phi(\mathbf{y}_2)\| \cdot \|\mathbf{y}_1 - \mathbf{y}_2\|,$$

or equivalently

$$\|\mathbf{y}_1 - \mathbf{y}_2\| \leq \frac{1}{\alpha} \|\nabla\phi(\mathbf{y}_1) - \nabla\phi(\mathbf{y}_2)\|. \quad (16)$$

For fixed λ , the minimiser $\tilde{\mathbf{y}}_\lambda$ of the Lagrangian $\mathcal{L}(\cdot, \lambda)$ on \mathbb{R}^n , i.e.,

$$\tilde{\mathbf{y}}_\lambda := \underset{\mathbf{y} \in \mathbb{R}^n}{\operatorname{argmin}} \mathcal{L}(\mathbf{y}, \lambda),$$

is given by the solution of

$$\nabla\phi(\tilde{\mathbf{y}}_\lambda) + A^T \lambda^T = 0. \quad (17)$$

Since ϕ is strongly convex, the Lagrangian $\mathcal{L}(\cdot, \lambda)$ is also strongly convex and $\tilde{\mathbf{y}}_\lambda$ is unique. Using (17) to substitute the gradient $\nabla\phi(\cdot)$ in (16) yields

$$\|\tilde{\mathbf{y}}_{\lambda_1} - \tilde{\mathbf{y}}_{\lambda_2}\| \leq \frac{1}{\alpha} \|-A^T \lambda_1 + A^T \lambda_2\| \leq \frac{1}{\alpha} \|A\| \|\lambda_1 - \lambda_2\|.$$

Let $\hat{\mathbf{y}}_{\lambda_1}$ and $\hat{\mathbf{y}}_{\lambda_2}$ denote the unique minimiser of the Lagrangian for fixed Lagrange multiplier λ_1 and λ_2 , respectively, subject to the constraint set D . Then, with Theorem 5, we obtain the estimate

$$\|\hat{\mathbf{y}}_{\lambda_1} - \hat{\mathbf{y}}_{\lambda_2}\| = \|[\tilde{\mathbf{y}}_{\lambda_1}]^+ - [\tilde{\mathbf{y}}_{\lambda_2}]^+\| \leq \|\tilde{\mathbf{y}}_{\lambda_1} - \tilde{\mathbf{y}}_{\lambda_2}\| \leq \frac{\|A\|}{\alpha} \|\lambda_1 - \lambda_2\|. \quad (18)$$

Since the gradient of the dual function can be evaluated using the expression $\nabla\psi(\lambda) = A\widehat{\mathbf{y}}_\lambda - b$ (see also Theorem 4), we obtain the estimate

$$\|\nabla\psi(\lambda_1) - \nabla\psi(\lambda_2)\| = \|(A\widehat{\mathbf{y}}_{\lambda_1} - b) - (A\widehat{\mathbf{y}}_{\lambda_2} - b)\| \leq \|A\|\|\widehat{\mathbf{y}}_{\lambda_1} - \widehat{\mathbf{y}}_{\lambda_2}\|. \quad (19)$$

Combining (18) and (19) provides the Lipschitz continuity estimate

$$\|\nabla\psi(\lambda_1) - \nabla\psi(\lambda_2)\| \leq \frac{\|A\|^2}{\alpha} \|\lambda_1 - \lambda_2\|$$

which shows the assertion. \square

With the results presented in this section the dual ascent algorithm can be stated and is summarised in Algorithm 1.

Algorithm 1 Dual ascent algorithm

Input:

- Strongly convex function $\phi : \mathbb{R}^n \rightarrow \mathbb{R}$,
- Non-empty, closed and convex set $D \subset \mathbb{R}^n$, and
- Linear coupling $\chi : \mathbb{R}^n \rightarrow \mathbb{R}^m$, $\chi(\mathbf{y}) = A\mathbf{y} - b$ with $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$.

Initialization: Set $\ell = 0$ and $\lambda^0 \in \mathbb{R}^m$.

Main loop:

- Solve the minimisation problem

$$\mathbf{y}^\ell := \operatorname{argmin}_{\mathbf{y} \in D} \left(\phi(\mathbf{y}) + \mathbf{y}^T A^T \lambda^\ell \right). \quad (20)$$

- Compute an ascent direction d^ℓ ; e.g., $d^\ell = \chi(\mathbf{y}^\ell)$.
- Choose a stepsize c^ℓ according to Theorem 3 and update the Lagrange multiplier

$$\lambda^{\ell+1} = \lambda^\ell + c^\ell \cdot d^\ell.$$

Increment the iteration counter $\ell = \ell + 1$ and repeat the loop.

3.4 The distributed dual ascent algorithm

The dual ascent Algorithm 1 is an efficient method if the solution of the minimisation problem (20) can be obtained easily or if the objective function ϕ is separable, and hence the minimisation problem (20) can be solved in a distributed way. Here, we will introduce a distributed dual ascent algorithm after we give a formal definition of a separable function.

Definition 2 (Separable function): Let $\phi : \mathbb{R}^n \rightarrow \mathbb{R}$ be a function in the variables

$$\mathbf{y} = (\mathbf{y}_1^T, \mathbf{y}_2^T, \dots, \mathbf{y}_I^T)^T$$

with $\mathbf{y}_i \in \mathbb{R}^{n_i}$, $i \in \mathbb{N}_I$, and $n = \sum_{i=1}^I n_i$. Then, ϕ is called separable if there exist functions

$\phi_i : \mathbb{R}^{n_i} \rightarrow \mathbb{R}$, $i \in \mathbb{N}_{\mathcal{I}}$, such that

$$\phi(\mathbf{y}) = \sum_{i=1}^{\mathcal{I}} \phi_i(\mathbf{y}_i)$$

holds for all $\mathbf{y} \in \mathbb{R}^n$.

We assume that the objective function $\phi : \mathbb{R}^n \rightarrow \mathbb{R}$ can be written in the form

$$\phi(\mathbf{y}) = \bar{\phi} \left(\frac{1}{\mathcal{I}} \sum_{i=1}^{\mathcal{I}} \mathbf{y}_i \right) + \frac{1}{\mathcal{I}} \sum_{i=1}^{\mathcal{I}} \phi_i(\mathbf{y}_i)$$

for strongly convex functions $\bar{\phi} : \mathbb{R}^N \rightarrow \mathbb{R}$ and $\phi_i : \mathbb{R}^{n_i} \rightarrow \mathbb{R}$, $i \in \mathbb{N}_{\mathcal{I}}$, with $N \in \mathbb{N}$ and $N\mathcal{I} = n$. Since the function $\bar{\phi}$ depends on the variables \mathbf{y}_i , $i \in \mathbb{N}_{\mathcal{I}}$, the function ϕ is not separable. Nevertheless, by introducing the variables $\bar{\mathbf{a}} \in \mathbb{R}^N$ and the linear constraints $\chi : \mathbb{R}^{\mathcal{I}N} \times \mathbb{R}^N \rightarrow \mathbb{R}^N$, as

$$\chi(\mathbf{y}, \bar{\mathbf{a}}) = -\frac{1}{\mathcal{I}} \sum_{i=1}^{\mathcal{I}} \mathbf{y}_i + \bar{\mathbf{a}},$$

we observe that the minimisation problem

$$\min_{\mathbf{y} \in D} \phi(\mathbf{y})$$

is equivalent to the minimisation problem

$$\begin{aligned} \min_{\bar{\mathbf{a}} \in \mathbb{R}^N, \mathbf{y}_i \in D_i} \quad & \bar{\phi}(\bar{\mathbf{a}}) + \frac{1}{\mathcal{I}} \sum_{i=1}^{\mathcal{I}} \phi_i(\mathbf{y}_i) \\ \text{s.t.} \quad & -\frac{1}{\mathcal{I}} \sum_{i=1}^{\mathcal{I}} \mathbf{y}_i + \bar{\mathbf{a}} = 0 \end{aligned} \quad (21)$$

for $D_i \subset \mathbb{R}^{n_i}$, $i \in \mathbb{N}_{\mathcal{I}}$, defined such that $D = D_1 \times \dots \times D_{\mathcal{I}}$ holds. The Lagrangian of the minimisation problem (21) is given by

$$\mathcal{L}(\mathbf{y}, \bar{\mathbf{a}}, \lambda) = \bar{\phi}(\bar{\mathbf{a}}) + \bar{\mathbf{a}}^T \lambda + \frac{1}{\mathcal{I}} \left(\sum_{i=1}^{\mathcal{I}} \phi_i(\mathbf{y}_i) - \lambda^T \mathbf{y}_i \right) \quad (22)$$

and, thus, the Lagrangian $\mathcal{L}(\cdot, \cdot, \lambda)$ is separable for fixed Lagrange multiplier $\lambda \in \mathbb{R}^N$. This implies that a solution of the optimisation problem (20) can be obtained by solving the minimisation problems

$$\bar{\mathbf{a}}^{*\lambda} = \operatorname{argmin}_{\bar{\mathbf{a}} \in \mathbb{R}^N} (\bar{\phi}(\bar{\mathbf{a}}) + \bar{\mathbf{a}}^T \lambda)$$

and

$$\mathbf{y}_i^{*\lambda} = \operatorname{argmin}_{\mathbf{y}_i \in D_i} (\phi_i(\mathbf{y}_i) - \mathbf{y}_i^T \lambda)$$

for $i \in \mathbb{N}_{\mathcal{I}}$ in parallel. Since $\bar{\phi}$ and ϕ_i , $i = 1, \dots, \mathcal{I}$, are strongly convex by assumption, also the functions $\bar{\phi}(\bar{\mathbf{a}}) - \bar{\mathbf{a}}^T \lambda$ and $\phi_i(\mathbf{y}_i) + \mathbf{y}_i^T \lambda$, $i = 1, \dots, \mathcal{I}$, are strongly convex for fixed $\lambda \in \mathbb{R}^N$. The corresponding distributed dual ascent algorithm is summarised in Algorithm 2.

Algorithm 2 Distributed dual ascent algorithm

Input:

- Strongly convex functions $\bar{\phi} : \mathbb{R}^N \rightarrow \mathbb{R}$ and $\phi_i : \mathbb{R}^N \rightarrow \mathbb{R}$, $i = 1, \dots, \mathcal{I}$,
- Non-empty, closed and convex sets $D_i \subset \mathbb{R}^N$, $i = 1, \dots, \mathcal{I}$, and
- Linear coupling $\chi : \mathbb{R}^{N\mathcal{I}} \times \mathbb{R}^N \rightarrow \mathbb{R}^N$, $\chi(\mathbf{y}, \bar{\mathbf{a}}) = -\frac{1}{\mathcal{I}} \sum_{i=1}^{\mathcal{I}} \mathbf{y}_i + \bar{\mathbf{a}}$.

Initialization: Set $\ell = 0$ and $\lambda^0 \in \mathbb{R}^N$.**Main Loop:****Phase 1 (Subsystem i , $i \in \mathbb{N}_{\mathcal{I}}$):** Receive λ^ℓ

- Solve the minimisation problem

$$\mathbf{y}_i^\ell := \operatorname{argmin}_{\mathbf{y}_i \in D_i} \left(\phi_i(\mathbf{y}_i) - \mathbf{y}_i^T \lambda^\ell \right) \quad (23)$$

and send \mathbf{y}_i^ℓ to the CE.**Phase 2 (CE):** Receive \mathbf{y}_i^ℓ for all $i = 1, \dots, \mathcal{I}$

- Solve the minimisation problem

$$\bar{\mathbf{a}}^\ell := \operatorname{argmin}_{\bar{\mathbf{a}} \in \mathbb{R}^N} \left(\bar{\phi}(\bar{\mathbf{a}}) + \bar{\mathbf{a}}^T \lambda^\ell \right). \quad (24)$$

- Compute an ascent direction d^ℓ , e.g. $d^\ell := \chi(\mathbf{y}^\ell, \bar{\mathbf{a}}^\ell)$.
- Choose a stepsize c^ℓ according to Theorem 3 and update the Lagrange multiplier

$$\lambda^{\ell+1} := \lambda^\ell + c^\ell \cdot d^\ell$$

and broadcast $\lambda^{\ell+1}$.Increment the iteration counter $\ell = \ell + 1$ and repeat the loop.

Remark 2: The distributed dual ascent algorithm can be applied to the more general problem

$$\min_{\mathbf{y}_i \in D_i} \sum_{i=1}^{\mathcal{I}+1} \phi_i(\mathbf{y}_i) \quad \text{s.t.} \quad \sum_{i=1}^{\mathcal{I}+1} A_i \mathbf{y}_i - b = 0$$

where $\phi_i : D_i \rightarrow \mathbb{R}$, $D_i \subset \mathbb{R}^{n_i}$, $n_i \in \mathbb{N}$, and $A_i \in \mathbb{R}^{m \times n_i}$ for $i = 1, \dots, \mathcal{I} + 1$, $b \in \mathbb{R}^m$. Since we consider optimisation problems in the unknowns \mathbf{z}_i and $\bar{\mathbf{z}} = \frac{1}{\mathcal{I}} \sum_{i=1}^{\mathcal{I}} \mathbf{z}_i$, $i \in \mathbb{N}_{\mathcal{I}}$, we use the notation used in Algorithm 2 for simplicity of exposition.

4. The dual ascent algorithm for a network of RESs

In the following section, we show how the distributed dual ascent Algorithm 2 can be used to find an approximation of the optimal solution of the minimisation problem

$$\operatorname{argmin}_{\mathbf{z}_i \in D_i} \left\| \frac{1}{\mathcal{I}} \sum_{i=1}^{\mathcal{I}} \mathbf{z}_i - \mathbf{1} \bar{\zeta} \right\|^2 \quad (25)$$

with $\bar{\zeta} = \bar{\zeta}(k)$. The problem (25) is the same as (5) introduced in Section 2, but written in terms of the power demand \mathbf{z} instead of the input \mathbf{u} . Observe that the objective function covers the performance metric (MQD). Since we consider a single optimisation problem corresponding to a fixed time instant $k \in \mathbb{N}$ in this section, we omit the time dependency whenever possible.

4.1 Cooperative application of the dual ascent algorithm

To apply Algorithm 2, we introduce the (auxiliary) variable $\bar{\mathbf{a}} \in \mathbb{R}^N$ and introduce the additional constraint

$$\bar{\mathbf{a}} = \bar{\mathbf{z}} = \frac{1}{\mathcal{I}} \sum_{i=1}^{\mathcal{I}} \mathbf{z}_i.$$

Thus, the optimisation problem (25) can be equivalently written as

$$\begin{aligned} & \underset{\mathbf{z} \in \mathcal{D}, \bar{\mathbf{a}} \in \mathbb{R}^N}{\operatorname{argmin}} && G(\bar{\mathbf{a}}) \\ & \text{s.t.} && \chi(\mathbf{z}, \bar{\mathbf{a}}) = 0 \end{aligned} \quad (26)$$

with $G : \mathbb{R}^N \rightarrow \mathbb{R}$, $G(\bar{\mathbf{a}}) = \frac{\eta}{2} \|\bar{\mathbf{a}} - \mathbf{1}\bar{\zeta}\|^2$ for fixed $\eta > 0$ and $\chi : \mathbb{R}^{\mathcal{I}N} \times \mathbb{R}^N \rightarrow \mathbb{R}^N$, $\chi(\mathbf{z}, \bar{\mathbf{a}}) = \bar{\mathbf{a}} - \frac{1}{\mathcal{I}} \sum_{i=1}^{\mathcal{I}} \mathbf{z}_i$. Note that the scaling factor $\eta > 0$ only changes the minimum G^* but not the (possibly non-unique) minimisers $(\mathbf{z}^*, \bar{\mathbf{a}}^*)$.

Since the function G does not depend on the variables \mathbf{z} , the objective function is not strongly convex and thus, the assumptions of the dual ascent algorithm are not satisfied. Therefore, we define the functions $F_i : \mathbb{R}^N \rightarrow \mathbb{R}$, $F_i(\mathbf{z}_i) = \frac{1}{2} \frac{\delta}{\mathcal{I}} \|\mathbf{z}_i\|^2$ for $i = 1, \dots, \mathcal{I}$ and $\delta > 0$, and define the relaxed optimisation problem

$$\begin{aligned} & \underset{\mathbf{z} \in \mathcal{D}, \bar{\mathbf{a}} \in \mathbb{R}^N}{\operatorname{argmin}} && G(\bar{\mathbf{a}}) + \sum_{i=1}^{\mathcal{I}} F_i(\mathbf{z}_i) \\ & \text{s.t.} && \chi(\mathbf{z}, \bar{\mathbf{a}}) = 0. \end{aligned} \quad (27)$$

Problem (27) is an approximation of the original problem (26) and for $\eta > 0$ and $\delta \rightarrow 0$ the minimisers of the original problem are recovered, as demonstrated in our numerical simulations below. The advantage of using the relaxed problem (27) is that convergence of Algorithm 2 can be shown if the stepsize is chosen appropriately. Before the corresponding result is given we introduce the notation $K : \mathbb{R}^{\mathcal{I}N} \times \mathbb{R}^N \rightarrow \mathbb{R}$,

$$K(\mathbf{z}, \bar{\mathbf{a}}) = G(\bar{\mathbf{a}}) + \sum_{i=1}^{\mathcal{I}} F_i(\mathbf{z}_i),$$

for the overall objective function.

Theorem 7: *Let the functions $G : \mathbb{R}^N \rightarrow \mathbb{R}$, $F_i : \mathbb{R}^N \rightarrow \mathbb{R}$, $i \in \mathbb{N}_{\mathcal{I}}$, and $\chi : \mathbb{R}^{\mathcal{I}N} \times \mathbb{R}^N \rightarrow \mathbb{R}^N$ be defined as*

$$G(\bar{\mathbf{a}}) = \frac{\eta}{2} \|\bar{\mathbf{a}} - \mathbf{1}\bar{\zeta}\|^2, \quad F_i(\mathbf{z}_i) = \frac{\delta \|\mathbf{z}_i\|^2}{2\mathcal{I}}, \quad \text{and} \quad \chi(\mathbf{z}, \bar{\mathbf{a}}) = \bar{\mathbf{a}} - \frac{1}{\mathcal{I}} \sum_{i=1}^{\mathcal{I}} \mathbf{z}_i$$

for a given reference $\bar{\zeta} \in \mathbb{R}^N$ and given parameters $\eta > 0$ and $\delta > 0$. Furthermore, let the ascent direction d^ℓ in Algorithm 2 be defined as $d^\ell = \chi(\mathbf{z}^\ell, \bar{\mathbf{a}}^\ell)$. Additionally, assume that the sequence

$(c^\ell)_{\ell \in \mathbb{N}}$ is defined such that

$$\varepsilon \leq c^\ell \leq (2 - \varepsilon) \frac{\min\{\frac{\delta}{\mathcal{I}}, \eta\}}{\frac{1}{\mathcal{I}} + 1} \quad (28)$$

for a fixed, sufficiently small $\varepsilon > 0$. Then the following holds:

- (i) The sequence $(\mathbf{z}^\ell, \bar{\mathbf{a}}^\ell)_{\ell \in \mathbb{N}}$ converges to the unique optimal solution of the minimisation problem (27), i.e., $(\mathbf{z}^\ell, \bar{\mathbf{a}}^\ell) \rightarrow (\mathbf{z}^*, \bar{\mathbf{a}}^*)$ for $\ell \rightarrow \infty$.
- (ii) The sequence $(\mathbf{z}^\ell, \bar{\mathbf{a}}^\ell)_{\ell \in \mathbb{N}}$ approaches primal feasibility, i.e., $\chi(\mathbf{z}^\ell, \bar{\mathbf{a}}^\ell) \rightarrow 0$ for $\ell \rightarrow \infty$ and every accumulation point λ^* of $(\lambda^\ell)_{\ell \in \mathbb{N}}$ is an optimal solution of the dual problem.

Proof. We show that the overall objective function K is strongly convex and we compute the Lipschitz constant of the gradient of the dual function according to Theorem 6 from which the results follow. The Hessian of the function K is given by

$$\nabla^2 K(\mathbf{z}, \bar{\mathbf{a}}) = \begin{bmatrix} \frac{\delta}{\mathcal{I}} I & & & & \\ & \ddots & & & \\ & & \frac{\delta}{\mathcal{I}} I & & \\ & & & \frac{\delta}{\mathcal{I}} I & \\ & & & & \eta I \end{bmatrix}$$

and hence, K is strongly convex with parameter

$$\alpha = \min\left\{\frac{\delta}{\mathcal{I}}, \eta\right\}. \quad (29)$$

The strong convexity implies the existence of a unique primal solution $(\mathbf{z}^*, \bar{\mathbf{a}}^*)$ (see Remark 1) from which the existence of an optimal dual solution λ^* follows (see Theorem 1). For the constraints χ , we use the notation

$$A := \begin{bmatrix} -\frac{1}{\mathcal{I}} I & \cdots & -\frac{1}{\mathcal{I}} I & I \end{bmatrix}$$

to obtain

$$\chi(\mathbf{z}, \bar{\mathbf{a}}) = A \begin{pmatrix} \mathbf{z}_1 \\ \vdots \\ \mathbf{z}_{\mathcal{I}} \\ \bar{\mathbf{a}} \end{pmatrix}.$$

Since $A \cdot A^T = (1 + \frac{1}{\mathcal{I}}) I$ holds, we obtain the norm $\|A\| = \sqrt{1 + \frac{1}{\mathcal{I}}}$ which provides the Lipschitz constant

$$L = \frac{1 + \frac{1}{\mathcal{I}}}{\min\{\frac{\delta}{\mathcal{I}}, \eta\}} \quad (30)$$

of the dual function ψ according to Theorem 6. The gradient of the dual function is given by $\nabla\psi(\lambda) = \chi(\mathbf{z}_\lambda, \bar{\mathbf{a}}_\lambda)$ where $(\mathbf{z}_\lambda, \bar{\mathbf{a}}_\lambda) \in \mathbb{D} \times \mathbb{R}^N$ denotes the unique minimiser of the optimisation

problem

$$(\mathbf{z}_\lambda, \bar{\mathbf{a}}_\lambda) = \underset{(\mathbf{z}, \bar{\mathbf{a}}) \in \mathcal{D} \times \mathbb{R}^N}{\operatorname{argmin}} \mathcal{L}(\mathbf{z}, \bar{\mathbf{a}}, \lambda)$$

(see Theorem 4). The concavity of the dual function together with the convexity of the domain of ψ implies that every stationary point λ^* (i.e., $\psi(\lambda^*) = \chi(\mathbf{z}_{\lambda^*}, \bar{\mathbf{a}}_{\lambda^*}) = 0$) of the sequence $(\lambda^\ell)_{\ell \in \mathbb{N}}$ is a global maximum of the dual problem. Finally, the saddle point Theorem 2 implies the convergence $(\mathbf{z}^\ell, \bar{\mathbf{a}}^\ell) \rightarrow (\mathbf{z}^*, \bar{\mathbf{a}}^*)$ for $\ell \rightarrow \infty$. This concludes the proof. \square

Remark 3: Note that the functions F_i , $i \in \mathbb{N}_{\mathcal{I}}$, can be defined in a different way. For example the functions

$$F_i(\mathbf{z}_i) = \frac{\delta}{2\mathcal{I}} \|\mathbf{z}_i - \zeta_i \mathbf{1}\|^2 \quad (31)$$

with $\zeta_i \in \mathbb{R}$ penalising the deviation from an individual reference value lead to the same stepsize condition (28). If strongly convex functions different from (31) are used, the condition on the stepsize might change.

Theorem 7 guarantees convergence of the distributed dual ascent Algorithm 2 for every $\eta > 0$, $\delta > 0$ if the stepsize c^ℓ is chosen according to (28). Nevertheless, only for fixed $\eta > 0$ and $\delta \rightarrow 0$ (or equivalently $\eta \rightarrow \infty$ and $\delta > 0$ fixed) is a solution of the original problem (25) recovered. In Section 5.1, the impact of the choice of $\delta > 0$ is investigated numerically.

To obtain a good approximation of the original problem, it is necessary to choose a small δ (if we additionally assume that $\eta = 1$ is fixed). Unfortunately, a small δ leads to a small Lipschitz constant L , and hence a small stepsize and, thus, slow convergence. Additionally, this effect increases with the number of RESs in the network. Even though the theoretical bound on the stepsize becomes quite small for large networks, numerically, larger stepsizes can often be used to obtain good results. If we consider the residual

$$\mathbf{r}^\ell := -\bar{\mathbf{z}}^\ell + \bar{\mathbf{a}}^\ell \quad (32)$$

in the ℓ -th iteration of Algorithm 2, we can define an alternative stepsize rule which guarantees convergence.

Corollary 1: *If the stepsize $(c^\ell)_{\ell \in \mathbb{N}}$ in Algorithm 2 is defined as*

$$c^{\ell+1} = \begin{cases} c^\ell & \text{if } \|\mathbf{r}^{\ell+1}\| < \|\mathbf{r}^\ell\| \\ \max \left\{ \frac{c^\ell}{2}, \frac{\min\{\frac{\delta}{2}, \eta\}}{\frac{1}{2}+1} \right\} & \text{if } \|\mathbf{r}^{\ell+1}\| \geq \|\mathbf{r}^\ell\| \end{cases} \quad (33)$$

and $c^0 > 0$, then the convergence properties of Theorem 7 hold.¹

Proof. The result follows immediately from Theorem 7 since the stepsize decreases until the condition (28) is satisfied for all following iterations ℓ . \square

The stepsize rule of Corollary 1 is used in the numerical simulations in Section 5. In our setting,

¹The stepsize rule of Corollary 1 significantly reduces the average number of iterations in our numerical simulations.

the stopping criterion is met in most of the cases with a stepsize

$$c^\ell \geq 2 \frac{\min\{\frac{\delta}{T}, \eta\}}{\frac{1}{T} + 1};$$

i.e., the condition (28) is not satisfied, and fewer iterations are necessary to obtain a solution of the optimisation problem (27).

Remark 4: For the function $G(\bar{\mathbf{a}}) = \frac{\eta}{2} \|\bar{\mathbf{a}} - \bar{\zeta} \mathbf{1}\|^2$, the update $\bar{\mathbf{a}}^{\ell+1}$ of the CE in Algorithm 2 can be computed explicitly as $\bar{\mathbf{a}}^{\ell+1} = \bar{\zeta} \mathbf{1} - \frac{1}{\eta} \lambda^\ell$.

Remark 5: The results in this section focus on the function $G(\bar{\mathbf{a}}) = \|\bar{\mathbf{a}} - \mathbf{1} \bar{\zeta}\|^2$. However, similar results can be derived for arbitrary strongly convex functions G defined on convex and closed sets.

4.2 Price-based implementation of the dual ascent algorithm

In this section, we give a non-cooperative, price-based, implementation of the distributed dual ascent Algorithm 2 in the form of a negotiation process between the CE (energy provider) and the RESs. Whereas in (25) the variables \mathbf{z}_i , $i \in \mathbb{N}_T$, are coupled and the RESs need to cooperate to find an optimal solution, in Algorithm 2, the local optimisation problem (23) of RES i

$$\mathbf{z}_i^{\ell+1} := \operatorname{argmin}_{\mathbf{z}_i \in \mathbb{D}_i} \left(F_i(\mathbf{z}_i) - \lambda^\ell \mathbf{z}_i^T \right)$$

solely depends on the variables \mathbf{z}_i and the dual variables λ defined by the CE. We show how the CE can reduce the fluctuations in the average power demand with real-time electricity price signals. Thus, the dual ascent algorithm corresponds to a negotiation of electricity prices until a price signal is found, which is accepted by the CE and the RESs.

To be able to introduce electricity prices in Algorithm 2, we describe the energy price at a certain time by a function $p : \mathbb{R} \rightarrow \mathbb{R}$ depending on the amount of energy used in the corresponding time interval. In particular, we want the following properties to be satisfied:

- a) The function p is increasing, i.e., a higher energy demand leads to higher costs, and
- b) if the power demand is 0 then also the costs are zero.

The simplest case is to consider linear costs; i.e., at a certain time the energy price can be described by the function

$$p(z; c) = Tcz$$

depending on the sampling time T and a constant $c \in \mathbb{R}_{\geq 0}$ chosen by the energy provider based on the average power demand \bar{z} . With respect to Algorithm 2 these costs are realized by rewriting the original minimisation problem (25) in the form

$$\begin{aligned} \min_{\mathbf{z} \in \mathbb{D}, \bar{\mathbf{a}} \in \mathbb{R}^N} \quad & \|\bar{\mathbf{a}} - \mathbf{1} \bar{\zeta}\|^2 \\ \text{s.t.} \quad & \frac{1}{T} \sum_{i=1}^T \mathbf{z}_i - \bar{\mathbf{a}} = 0. \end{aligned}$$

Indeed, this minimisation problem implies local updates of the form

$$\begin{aligned}\bar{\mathbf{a}}^{*\lambda} &= \operatorname{argmin}_{\bar{\mathbf{a}} \in \mathbb{R}^N} \left(\|\bar{\mathbf{a}} - \mathbf{1}\bar{\zeta}\|^2 - \bar{\mathbf{a}}^T \lambda \right) \\ \mathbf{z}_i^{*\lambda} &= \operatorname{argmin}_{\mathbf{z}_i \in \mathbb{D}_i} \mathbf{z}_i^T \lambda,\end{aligned}$$

(cf. Equations (23) and (24) in Algorithm 2) and we can identify $c = c(k+j) = \lambda_j/T$, $j \in \{0, 1, \dots, N-1\}$, as the parameter of the CE defining the price for electricity at time instant $k+j$. For this setting, the local functions $F_i \equiv 0$, $i \in \mathbb{N}_{\mathcal{I}}$, are convex but not strongly convex. Thus, the assumptions of Algorithm 2 are not satisfied so that convergence cannot be guaranteed. In fact, while the sequence $(\lambda^\ell)_{\ell \in \mathbb{N}}$ seems to converge for a diminishing stepsize, the sequence $(\bar{\mathbf{z}}^\ell)_{\ell \in \mathbb{N}}$ does not need to converge as already observed in the Market Maker approach presented in (Worthmann et al., 2015).

As a remedy, we propose an approach where the electricity prices contain an additional penalty term. More precisely, we propose linear costs with an additional quadratic penalty term

$$p(z; c) := T \cdot a \left(z + b(z - c)^2 - bc^2 \right), \quad (34)$$

with cost coefficients $a, b \in \mathbb{R}_{>0}$ and parameter $c \in \mathbb{R}$. The price still contains the linear term Taz , but in addition, the demand deviating from a given reference c is penalised. Either a power demand above the reference is penalised by additional costs, or the earning is reduced if too much energy is sold to the grid operator. In particular, $p(0; c) = 0$, i.e. Property b) holds. Since, $p(\cdot; c)$ is not monotonically increasing on \mathbb{R} the constants a and b have to be fixed by the CE such that $p(\cdot; c)$ is monotonically increasing on the domain of interest (Property a)). This is always possible since both the net consumption profiles \mathbf{w}_i and the battery capacities C_i , $i \in \mathbb{N}_{\mathcal{I}}$, are bounded.

Minimizing the electricity costs with respect to the function $p(\cdot; c)$ over the prediction horizon $N \in \mathbb{N}$ can be achieved by the minimisation problem

$$\begin{aligned}\min_{\mathbf{z} \in \mathbb{D}, \bar{\mathbf{a}} \in \mathbb{R}^N} \quad & \frac{\eta}{2} \|\bar{\mathbf{a}} - \mathbf{1}\bar{\zeta}\|^2 + \frac{1}{T} \sum_{i=1}^{\mathcal{I}} \left(\rho \mathbf{1}^T \mathbf{z}_i + \frac{\delta}{2} \|\mathbf{z}_i\|^2 \right) \\ \text{s.t.} \quad & -\frac{1}{T} \sum_{i=1}^{\mathcal{I}} \mathbf{z}_i + \bar{\mathbf{a}} = 0\end{aligned} \quad (35)$$

for $\eta, \rho, \delta \in \mathbb{R}_{>0}$. In the context of Algorithm 2, the optimisation problems of the CE and of the RESs read

$$\begin{aligned}\bar{\mathbf{a}}^{*\lambda} &= \operatorname{argmin}_{\bar{\mathbf{a}} \in \mathbb{R}^N} \left(\frac{\eta}{2} \|\bar{\mathbf{a}} - \mathbf{1}\bar{\zeta}\|^2 + \bar{\mathbf{a}}^T \lambda \right), \\ \mathbf{z}_i^{*\lambda} &= \operatorname{argmin}_{\mathbf{z}_i \in \mathbb{D}_i} \left(\rho \mathbf{1}^T \mathbf{z}_i + \frac{\delta}{2} \|\mathbf{z}_i\|^2 - \mathbf{z}_i^T \lambda \right).\end{aligned}$$

The functions $F_i(\mathbf{z}_i) = \frac{1}{T} \left(\rho \mathbf{1}^T \mathbf{z}_i + \frac{\delta}{2} \|\mathbf{z}_i\|^2 \right)$, $i \in \mathbb{N}_{\mathcal{I}}$, are strongly convex due to the quadratic term (with parameter $\alpha = \delta/T$, see Remark 1). Hence, Algorithm 2 is applicable. More importantly, the objective function of RES i , $i \in \mathbb{N}_{\mathcal{I}}$, can be written in the form

$$\rho \mathbf{1}^T \mathbf{z}_i + \frac{\delta}{2} \|\mathbf{z}_i\|^2 - \mathbf{z}_i^T \lambda = \rho \mathbf{1}^T \mathbf{z}_i + \frac{\delta}{2} \left\| \mathbf{z}_i - \frac{\lambda}{\delta} \right\|^2 - \frac{\lambda^T \lambda}{2\delta} \quad (36)$$

yielding the price function $p(\cdot; c)$ by identifying the parameters

$$a = \frac{\rho}{T}, \quad b = \frac{\delta}{2T}, \quad c(k+j) = \frac{1}{\delta} \lambda_j. \quad (37)$$

The constant term $\frac{1}{2\delta} \lambda^T \lambda$ has no impact on the minimiser $\mathbf{z}_i^{\star\lambda}$ and thus, it does not have to be considered in the minimisation problem.

In summary, minimisation of the electricity costs of RES i over the prediction horizon N is equivalent to solving the minimisation problem (23). Moreover, if the minimisation problems (23) are solved iteratively, an optimal solution of (35) is recovered. The price contains the linear component Taz , but is also constructed such that deviations from a given reference value c are penalised.

In Figure 2, the price for electricity is visualised for fixed parameters a and b .² If λ_j and consequently also $c = c(k+j)$ is small, energy is expensive and vice-versa. This figure additionally shows that it is possible to obtain almost linear prices, for which the dual ascent algorithm converges. Note that the parameters ρ and δ can be set individually for every RES; e.g., via specific contracts depending on the respective demand profiles.

Note that the additional linear term $\rho \mathbf{1}^T \mathbf{z}_i$ in (36) does not change the convexity properties of a function. Consequently, Theorem 7 also holds for $F_i(\mathbf{z}_i) = \rho \mathbf{1}^T \mathbf{z}_i + \frac{\delta}{2} \|\mathbf{z}_i\|^2$, $i \in \mathbb{N}_{\mathcal{I}}$.

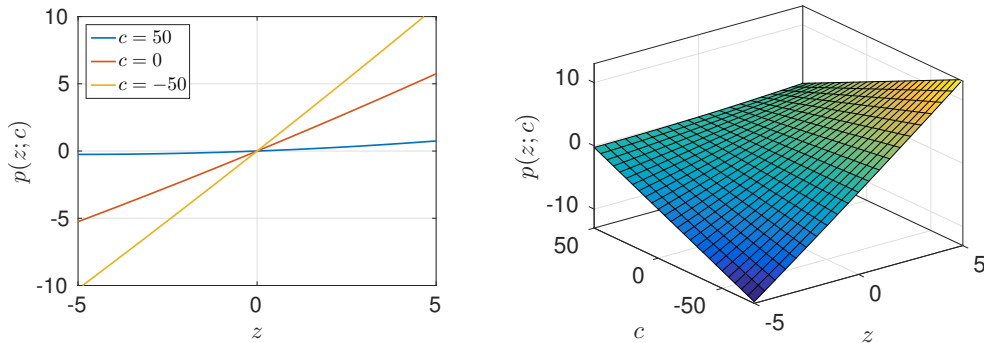


Figure 2. Visualization of the electricity cost function $p(\cdot; c)$ (left) and $p(\cdot; \cdot)$ (right) from (34) for $a = 2.2$ and $b = 0.0091$ (corresponding to $\rho = 1.1$, $\delta = 0.02$ and $T = 0.5$). In the area of interest, the price is almost linear with respect to c and the demand z . The nonlinearity of the cost function can still be seen in $p(z; 50)$ (left).

In Section 5.2, numerical simulations investigating the pricing scheme are provided. In particular, the impact of batteries on the price signals c are investigated.

Remark 6: The objective function $G(\bar{\mathbf{a}}) = \|\bar{\mathbf{a}} - \mathbf{1}\bar{\zeta}\|^2$ can be replaced by any other strongly convex function. Also, constraints on $\bar{\mathbf{a}} \in \mathbb{A}$ for a closed and convex set $\mathbb{A} \subset \mathbb{R}^N$ can be included as long as feasibility of the optimisation problem is guaranteed.

4.3 General properties of the (non-)cooperative control setting

In this section, we highlight properties of the cooperative control setting (i.e., when the optimisation problem (25) is solved directly), and the non-cooperative, price-based control setting. We discuss similarities and differences of both schemes and investigate how the performance (with respect to the performance metrics) deteriorates if a non-cooperative, price-based scheme is used instead of a cooperative scheme.

²The parameters chosen here are just for illustration. The value used for a can be scaled arbitrarily to obtain realistic energy prices.

In the cooperative setting, an optimal solution of the optimisation problem

$$\min_{\mathbf{z}_i \in \mathbb{D}_i} \left\| \frac{1}{T} \sum_{i=1}^{\mathcal{I}} \mathbf{z}_i - \mathbf{1}\bar{\zeta} \right\|^2 \quad (38)$$

is obtained. With the distributed dual ascent Algorithm 2, we obtain an approximation of the solution of (38) by solving the minimisation problem

$$\begin{aligned} \min_{\mathbf{z}_i \in \mathbb{D}_i, \bar{\mathbf{a}} \in \mathbb{R}^N} \quad & \frac{\eta}{2} \|\bar{\mathbf{a}} - \mathbf{1}\bar{\zeta}\|^2 + \frac{1}{T} \frac{\delta}{2} \sum_{i=1}^{\mathcal{I}} \|\mathbf{z}_i\|^2 \\ \text{s.t.} \quad & -\frac{1}{T} \sum_{i=1}^{\mathcal{I}} \mathbf{z}_i + \bar{\mathbf{a}} = 0 \end{aligned} \quad (39)$$

for fixed $\eta, \delta \in \mathbb{R}_{>0}$. As already pointed out, for fixed η and $\delta \rightarrow 0$, an optimal solution of the original problem is recovered. To obtain a price-based implementation, we introduced the minimisation problem

$$\begin{aligned} \min_{\mathbf{z}_i \in \mathbb{D}_i, \bar{\mathbf{a}} \in \mathbb{R}^N} \quad & \frac{\eta}{2} \|\bar{\mathbf{a}} - \mathbf{1}\bar{\zeta}\|^2 + \frac{1}{T} \sum_{i=1}^{\mathcal{I}} \left(\rho \mathbf{1}^T \mathbf{z}_i + \frac{\delta}{2} \|\mathbf{z}_i\|^2 \right) \\ \text{s.t.} \quad & -\frac{1}{T} \sum_{i=1}^{\mathcal{I}} \mathbf{z}_i + \bar{\mathbf{a}} = 0 \end{aligned} \quad (40)$$

with $\eta, \delta, \rho \in \mathbb{R}_{>0}$. This problem can be solved using a non-cooperative, price-based implementation of Algorithm 2 as indicated in Section 4.2.

In the remainder of this section we concentrate on characteristic properties of the cooperative optimisation problems (38) and (39), and the non-cooperative optimisation problem (40). According to our assumption, the price function $p(\cdot; c)$ is monotonically increasing; i.e., feeding in energy results in a profit in the non-cooperative setting. This implies that the battery of each RES is empty at the end of the prediction horizon³; i.e., $x_i(k+N) = 0$ for all $i \in \mathbb{N}_{\mathcal{I}}$, whereas in the cooperative setting the state of charge of the batteries $x_i(k+N)$ for all $i \in \mathbb{N}_{\mathcal{I}}$ is determined by the optimisation criteria. This is one characteristic which distinguishes the optimal cooperative solution from the optimal non-cooperative solution. In the cooperative control setting, the additional constraints $x_i(k+N) = 0$ can be included by using the constraints

$$\mathbb{D}_i^0 = \left\{ \mathbf{z}_i \in \mathbb{R}^N \left| \begin{array}{l} x_i(k) = x_i^0, x_i(k+N) = 0 \\ \text{System dynamics (1) and (2)} \\ \text{Constraints (3)} \\ \forall j \in \{k, \dots, k+N-1\} \end{array} \right. \right\}$$

instead of \mathbb{D}_i for all $i \in \mathbb{N}_{\mathcal{I}}$ to ensure that the battery is empty at the end of the prediction horizon. If the battery state of charge at the end of the prediction horizon is known, then the energy used in the prediction window is known up to losses in the battery model and can be used to simplify the objective function. To illustrate this fact, we consider the simplified dynamics (1) and (2) without losses; i.e., $\beta_i = \gamma_i = 1$ for all $i \in \mathbb{N}_{\mathcal{I}}$. Here, the empty battery at the end of the prediction horizon implies that the energy demand is constant for all feasible demand profiles $\mathbf{z}_i \in \mathbb{D}_i^0$; i.e.,

$$\mathbf{1}^T \mathbf{z}_i = \frac{x_i(k)}{T} + \sum_{j=k}^{k+N-1} w_i(j).$$

This allows for the following equivalent characterization of the set of minima in the cooperative

³We assume that the prediction horizon N and the maximal discharging rate \underline{u}_i , $i \in \mathbb{N}_{\mathcal{I}}$, are chosen such that the battery can always be discharged within the prediction horizon.

setting:

$$\begin{aligned} \operatorname{argmin}_{\mathbf{z}_i \in \mathbb{D}_i^0} \|\bar{\mathbf{z}} - \bar{\zeta} \mathbf{1}\|^2 &= \operatorname{argmin}_{\mathbf{z}_i \in \mathbb{D}_i^0} \|\bar{\mathbf{z}}\|^2 - 2\bar{\zeta} \mathbf{1}^T \bar{\mathbf{z}} + N\bar{\zeta}^2 \\ &= \operatorname{argmin}_{\mathbf{z}_i \in \mathbb{D}_i^0} \|\bar{\mathbf{z}}\|^2. \end{aligned}$$

Analogously, for the non-cooperative problem one obtains

$$\operatorname{argmin}_{\mathbf{z}_i \in \mathbb{D}_i^0} \frac{\eta}{2} \|\bar{\mathbf{z}}\|^2 + \frac{\delta}{2\mathcal{I}} \sum_{i=1}^{\mathcal{I}} \|\mathbf{z}_i\|^2.$$

This implies that the set of optimal solutions is independent of the reference value $\bar{\zeta}$. In particular, no a priori knowledge of the average demand is needed in the case without losses in the battery model. Furthermore, observe that the constant ρ disappears in the objective function since it only appears in the linear term and does not have an impact on the minimiser. Nevertheless, ρ is necessary to obtain the price-interpretation and monotone electricity costs. We summarise these results in the following corollary.

Corollary 2: *Consider the system dynamics (1) and (2) subject to the constraints (3) and let $\beta_i = \gamma_i = 1$ for all $i \in \mathbb{N}_{\mathcal{I}}$. Additionally, let $\eta, \rho, \delta \in \mathbb{R}_{>0}$ be fixed and replace \mathbb{D}_i by \mathbb{D}_i^0 . Then the minimisers can be characterized as follows*

(i) *Original cooperative problem (38):*

$$\operatorname{argmin}_{\mathbf{z}_i \in \mathbb{D}_i^0} \left\| \frac{1}{\mathcal{I}} \sum_{i=1}^{\mathcal{I}} \mathbf{z}_i \right\|^2.$$

(ii) *Relaxed cooperative problem (39):*

$$\begin{aligned} \operatorname{argmin}_{\mathbf{z}_i \in \mathbb{D}_i^0, \bar{\mathbf{a}} \in \mathbb{R}^N} \quad & \frac{\eta}{2} \|\bar{\mathbf{z}}\|^2 + \frac{\delta}{2\mathcal{I}} \sum_{i=1}^{\mathcal{I}} \|\mathbf{z}_i\|^2 \\ \text{s.t.} \quad & -\frac{1}{\mathcal{I}} \sum_{i=1}^{\mathcal{I}} \mathbf{z}_i + \bar{\mathbf{a}} = \mathbf{0}. \end{aligned}$$

(iii) *Non-cooperative optimisation problem (40):*

$$\begin{aligned} \operatorname{argmin}_{\mathbf{z}_i \in \mathbb{D}_i^0, \bar{\mathbf{a}} \in \mathbb{R}^N} \quad & \frac{\eta}{2} \|\bar{\mathbf{z}}\|^2 + \frac{\delta}{2\mathcal{I}} \sum_{i=1}^{\mathcal{I}} \|\mathbf{z}_i\|^2 \\ \text{s.t.} \quad & -\frac{1}{\mathcal{I}} \sum_{i=1}^{\mathcal{I}} \mathbf{z}_i + \bar{\mathbf{a}} = \mathbf{0}. \end{aligned}$$

In Corollary 2 it is worth emphasising that, when using the constraint set \mathbb{D}_i^0 in place of the constraint set \mathbb{D}_i , the relaxed cooperative optimisation problem (39) is the same as the non-cooperative optimisation problem (40). Specifically, the derivations show that the minimiser of the relaxed problem (39) only differs from the minimiser of the non-cooperative problem (40) because of the constraint $x_i(k+N) = 0$ for all $i \in \mathbb{N}_{\mathcal{I}}$ independent of the parameter ρ (if a model without losses is considered). Moreover, if δ is chosen small, we can expect that a good approximation of a minimiser of the original problem (38) is obtained.

If losses are considered, the term $\mathbf{1}^T \mathbf{z}_i$ is not constant for all $\mathbf{z}_i \in \mathbb{D}_i^0$. Hence, for smaller constants β_i and γ_i , the impact of the linear term in the non-cooperative optimisation problem increases (see Section 5.2).

5. Numerical simulations

In this section, we examine the performance of the distributed dual ascent Algorithm 2 by simulating the algorithm using data obtained from the Australian electricity distribution company Ausgrid (Ratnam et al., 2015). Consistent with the Ausgrid data, we fix $T = 0.5[\text{h}]$ and we use a prediction horizon of $N = 48$, which then corresponds to a prediction of one day. We consider $\mathcal{I} = 100$ RESs defined through the system dynamics (1), (2), and the constraints (3).

5.1 The impact of the relaxation parameter δ

The original optimisation problem (38) cannot be solved using the distributed dual ascent Algorithm 2 since the assumptions on the objective function are not satisfied. Instead the relaxed problem (39) is solved by Algorithm 2 for fixed $\eta, \delta \in \mathbb{R}_{>0}$. In Figure 3, the open loop solution for $\eta = 1$ and varying δ , as well as the solution of the original optimisation problem (38), is shown. For this simulation, every RES is equipped with a battery without losses (i.e., $\beta_i = \gamma_i = 1$ for all $i \in \mathbb{N}_{100}$). For this example batteries of size $C_i = 2[\text{kWh}]$ and $\bar{u}_i = -\bar{d}_i = 0.3[\text{kW}]$ for all $i \in \mathbb{N}_{100}$ are used to avoid that the average demand profile \bar{z} is just a straight line and the constraints on the charging/discharging rates are never active.

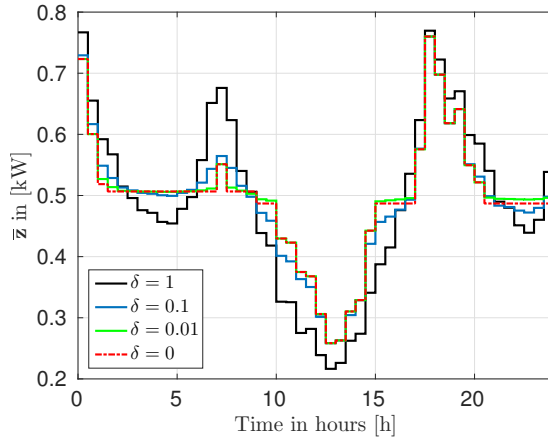


Figure 3. Comparison of the solution of the centralised optimisation problem and the relaxed optimisation problem for different relaxation parameters δ .

For a decreasing parameter δ the difference between the solution \bar{z} of the original optimisation problem and the relaxed optimisation problem shrinks. For $\delta = 0.01$ the difference in the solutions is negligible. The corresponding performance metrics are summarised in Table 1. The same conclusions can be drawn for the setting with losses, i.e., $\beta_i \in (0, 1)$ and $\gamma_i \in (0, 1)$ for $i \in \mathbb{N}_{\mathcal{I}}$ (and $\rho = 0$).

	PTP	MQD
Uncontrolled	1.1016	0.0879
$\delta = 0$	0.5016	0.0100
Distributed dual ascent Algorithm 2		
$\delta = 1$	0.5531	0.0198
$\delta = 0.1$	0.5016	0.0110
$\delta = 0.01$	0.5016	0.0100

Table 1. Performance of the open loop solution of the original problem ($\delta = 0$) compared to the solution using the dual ascent Algorithm 2 for different relaxations δ .

In conclusion, the cooperative application of the distributed dual ascent Algorithm 2 recovers an optimal solution of the original problem if the parameter δ is chosen small enough, justifying the relaxation of the original problem.

5.2 Price-based MPC simulations

In this section we concentrate on the difference of the solutions of the optimisation problems (38) and (40) with $\eta = 1$, $\rho = 1.1$, and $\delta = 0.02$. In particular, we examine the open loop and the closed loop solutions depending on the number of RESs with a battery. From the point of view of the CE (energy provider), we consider the performance metrics, and from the point of view of an individual RES, we consider energy prices and savings.

In the simulations using the minimisation problem (35) (corresponding to the real-time prices (34)), we vary the number of RESs with a battery. For RESs with a battery, we choose the constraints $C_i = 4$, $\bar{u}_i = -\underline{u}_i = 1$ and the initial conditions $x_i(0) = 0$. (For the RESs without a battery we set $C_i = 0$.) According to (Chen et al., 2009) the cycle efficiency of Li-ion batteries is higher than 90%; i.e., the losses in the (dis-)charging process are less than 10%. Hence, we illustrate our findings with the parameters

$$(\beta_i, \gamma_i) \in \{(1, 1), (0.95, 0.95), (0.9, 0.9)\},$$

for all $i \in \mathbb{N}_{\mathcal{I}}$, which corresponds to losses over a charge/discharge cycle of 0% and of about 10% and 19%, respectively.

5.2.1 Open loop (one-shot optimisation) simulations

In Figure 4, we compare the solution of the original cooperative optimisation problem with the solution of the non-cooperative problem for a different number of batteries in the overall system and different constants β_i and γ_i , $i \in \mathbb{N}_{100}$. We vary the number of batteries from 0 to 100 in steps of 10. The red line corresponds to the setting where every RES has a battery.

We observe that for the case without losses, the solution of cooperative control and non-cooperative control only varies slightly due to the penalty term $\frac{\delta}{2} \|\mathbf{z}_i\|^2$ and due to the empty battery at the end of the prediction horizon in the price-based setting. (This effect can be observed in the battery profiles which are excluded for brevity.) As argued in Section 4.3, the additional linear term in the cost functional does not influence the optimal power demand profile. Since $x_i(0) = 0$, the deviation from the average in the first 2 hours cannot be compensated.

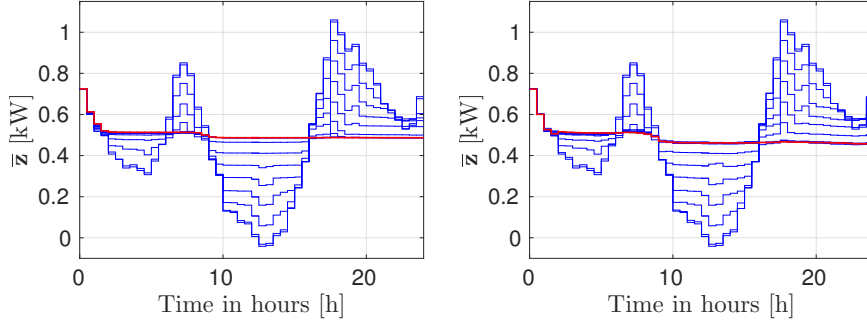
While in the cooperative setting the performance does not deteriorate if losses are considered and if enough storage devices are in the network, in the price-based setting the performance worsens with the loss of energy. If the losses are too big and the gradient in the corresponding electricity costs is too small, it does not pay off to store energy at one time instant to use it at a later time instant.

To make this more precise, we investigate in which case it pays off to use an isolated battery (i.e., we assume that $z(k) = 0$ for all time indices $k \in \mathbb{N}$ where the battery is not used) with efficiency $\beta, \gamma \in (0, 1]$ to buy an amount of energy $z > 0$ at time $j \in \mathbb{N}$ and sell it at time $k \in \mathbb{N}$. To make a profit the inequality

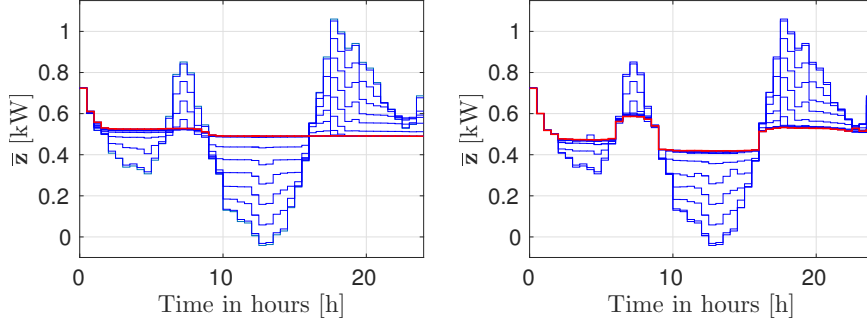
$$p(z, c(j)) + p(-\beta\gamma z, c(k)) < 0$$

has to hold; i.e., the money paid to the central entity is negative. With (34) this inequality can be

Cooperative (left) and price-based (right) optimal power demand without losses



Cooperative (left) and price-based (right) optimal power demand with $\beta_i = \gamma_i = 0.95$



Cooperative (left) and price-based (right) optimal power demand with $\beta_i = \gamma_i = 0.90$

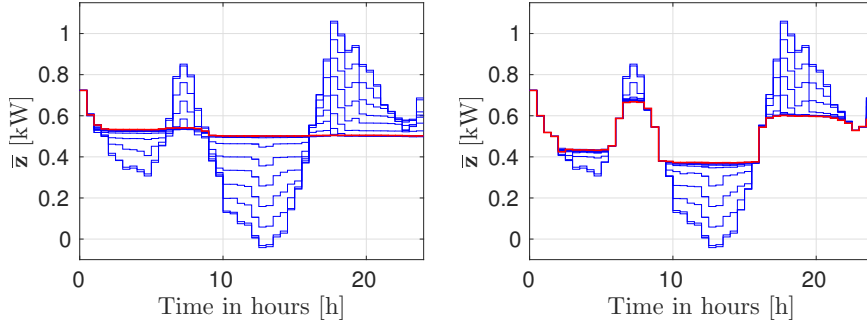


Figure 4. Comparison of the cooperative and the non-cooperative solution of a single optimisation problem with increasing number of batteries in the network (blue lines) and increasing losses (top to bottom). The red line indicates the setting where every RES owns a battery.

rewritten as

$$\begin{aligned} 0 &> (1 - \beta\gamma)z + bz^2 - 2bzc(j) - b\beta^2\gamma^2z^2 + 2b\beta\gamma zc(k) \\ &= (1 - \beta\gamma)z + b(1 + \beta^2\gamma^2)z - 2bz(c(j) + \beta\gamma c(k)). \end{aligned}$$

Reordering the terms leads to

$$0 < z < 2 \frac{c(j) - \beta\gamma c(k)}{1 + \beta^2\gamma^2} - \frac{1}{b} \frac{1 - \beta\gamma}{1 + \beta^2\gamma^2}$$

and finally to the estimate

$$0 < 1 - \beta\gamma < 2b(c(j) - \beta\gamma c(k)) \quad (41)$$

or

$$0 < \beta\gamma(1 - 2bc(k)) + 2bc(j) - 1. \quad (42)$$

Since by assumption, the price function (34) is monotonically increasing on the range of values defined by the minimum and maximum possible loads, the parameter $b > 0$ has to satisfy $b < 1/(2|c(j)|)$ for all $j \in \mathbb{N}$. Thus, $\beta\gamma(1 - 2bc(k))$ is strictly increasing in β, γ and

$$\beta\gamma(1 - 2bc(k)) + 2bc(j) - 1 \leq 2b(c(j) - c(k)) \quad (43)$$

holds for all $\beta, \gamma \in (0, 1]$. To be able to generate a profit, we can conclude that the losses in the battery model have to be small enough such that inequality (42) holds and in the case where no losses in the battery model are considered, i.e., $\beta\gamma = 1$ a profit can always be generated if $c(j) - c(k) > 0$.

Nevertheless, even if $\beta_i = \gamma_i = 0.9$ and hence 19% of the energy is lost, the deviation of the aggregated power demand from the average can be reduced significantly as seen in Figure 4.

The optimal dual variables λ^* corresponding to the non-cooperative settings in Figure 4 (right) are shown in Figure 5.

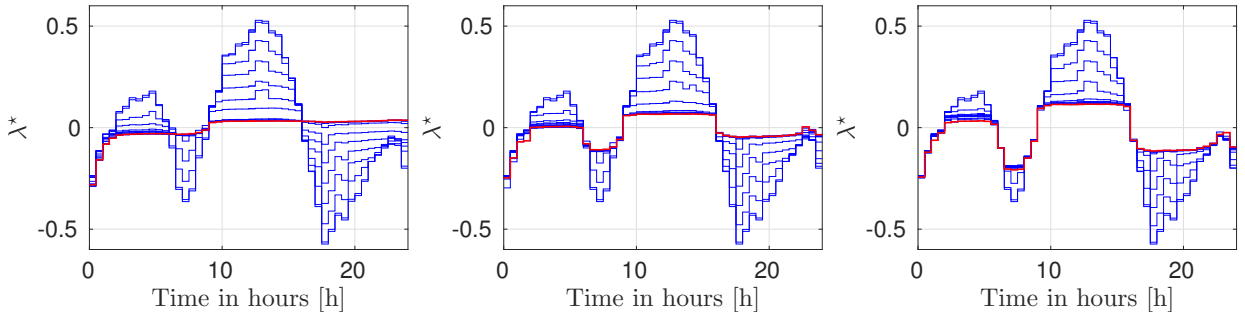


Figure 5. Optimal dual variables λ^* of a single optimisation problem corresponding to the case without losses (left), $\beta_i = \gamma_i = 0.95$ (middle) and $\beta_i = \gamma_i = 0.9$. One can observe that the average demand profiles are mirrored

In Figure 6, the performance metrics of the non-cooperative open loop solutions are illustrated. The usage of storage devices significantly improve the results — even if losses are taken into account. If 60% of the RESs own a battery the results do not improve anymore. In the case where only a small number of RESs have a battery, the losses do not have a significant impact on the performance.

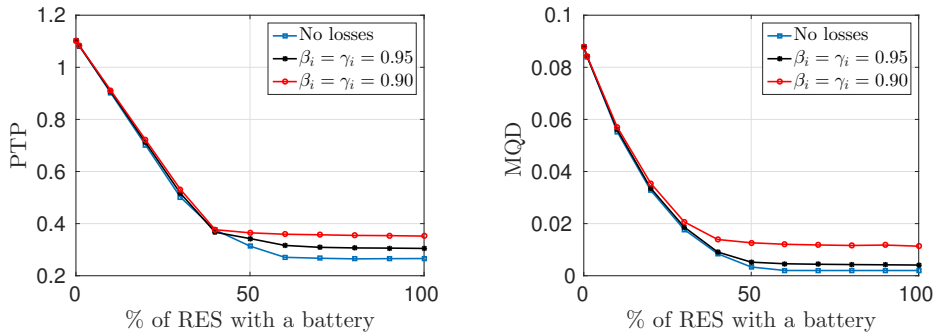


Figure 6. Performance metrics of the non-cooperative open loop solution. Beyond a certain threshold (about 60%), further deployment of battery storage does not yield improved performance.

Not only the energy provider benefits from the smoothed aggregated power demand; individual RESs can benefit as well from lower energy prices if the prices which result from the case without batteries are taken as a reference. In Figure 7 (left), the average electricity costs⁴ for the RESs are shown depending on the number of RESs with a battery. Figure 7 (right) shows that, on average, the costs can be reduced up to 14% in the case without losses and up to 8% if $\gamma_i = \beta_i = 0.9$.

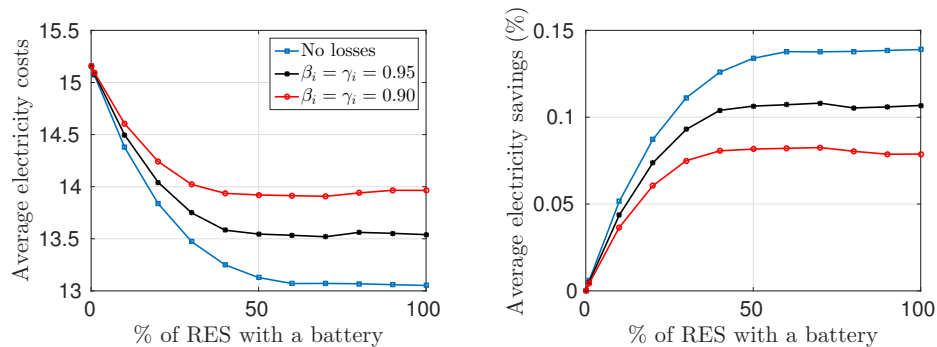


Figure 7. Average open loop electricity costs for every RES and savings with respect to the number of batteries in the network.

Figure 8 shows the savings of the individual systems if 1%, 20%, 50% and 100% of the RESs own a battery. Not only the RESs with a battery reduce their costs, but also most of the RESs without a battery benefit with a significant deployment of battery resources. If many RESs own a battery, the effect of single systems is reduced. This can be observed, for example, looking at RES 1. Moreover, we point out, that even though most of the RESs can reduce their costs, there exist single RESs with increased costs. Figure 8 corresponds to the setting without losses. The cases with loss of energy show a similar behaviour.

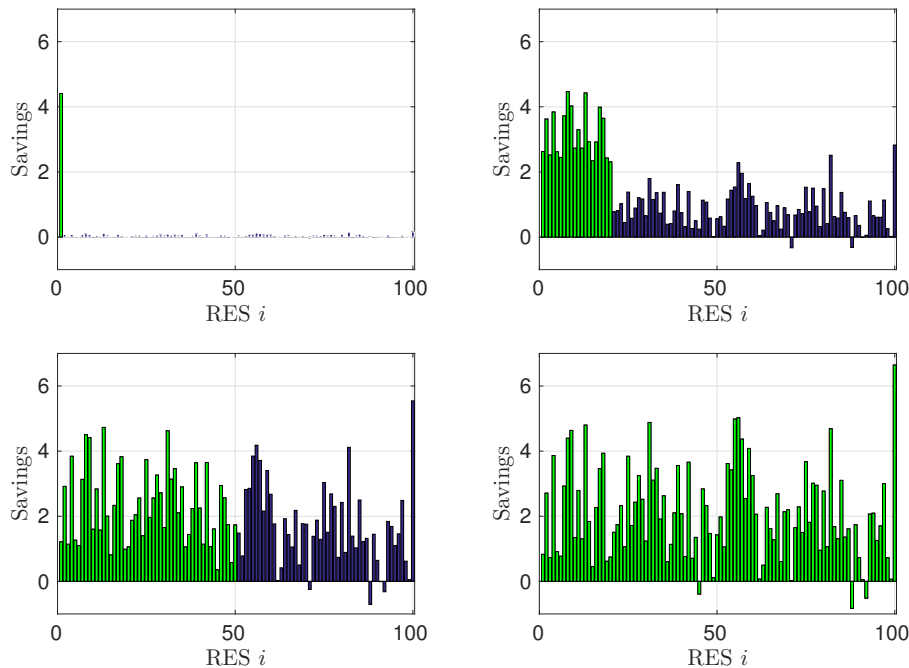


Figure 8. Savings of the individual RESs. Green bars indicate the RESs with batteries and blue bars indicate the RESs without batteries. Observe that not all RESs benefit from batteries. (The results correspond to the setting without losses.)

⁴To obtain realistic energy prices, the numbers can be scaled arbitrarily in Figure 7 (left).

5.2.2 Closed loop (MPC) simulations

In this section we investigate the closed loop performance of Algorithm 2 embedded in a receding horizon scheme. To reduce the number of iterations ℓ in Algorithm 2, warm-start can be used at every time step k . Here, we use the optimal Lagrange variables at time instant k

$$\lambda_k^* = (\lambda_k^*(k), \dots, \lambda_k^*(k + N - 2), \lambda_k^*(k + N - 1))$$

to initialize the Lagrange multipliers at time $k + 1$ as

$$\lambda_{k+1}^0 = (\lambda_k^*(k + 1), \dots, \lambda_k^*(k + N - 1), \lambda_k^*(k + N - 1)).$$

The stepsize at a fixed time instant k is selected according to Corollary 1. Proceeding from time instant k to $k + 1$ we initialize c_{k+1}^0 by $c_{k+1}^0 = 2 \cdot c_k^{\text{end}}$ where c_k^{end} denotes the stepsize of Algorithm 2 at the iteration where the stopping criterion is satisfied. This ensures that the stepsize not only decreases but also increases again at the next time instant.

We use a simulation of $\mathcal{N} = 387$ time steps⁵ to verify that the performance properties of the open loop simulation carry over to the closed loop. In Figure 9, the closed loop solution for 100 RESs is visualised. In this setting only 50 RESs are equipped with batteries. The average over the state $\bar{\mathbf{x}}^{\text{cl}}$ and the input $\bar{\mathbf{u}}^{\text{cl}} = \bar{\mathbf{u}}^{\text{+cl}} + \bar{\mathbf{u}}^{\text{-cl}}$ is taken only with respect to the RESs with a battery. As one already expects from the open loop control problem, the peaks in the average demand $\bar{\mathbf{z}}^{\text{cl}}$ are decreased significantly. If losses are considered in the battery model, then the battery capacities are not fully used since it does not pay off to buy and sell energy if the gradient in the cost functional is small.

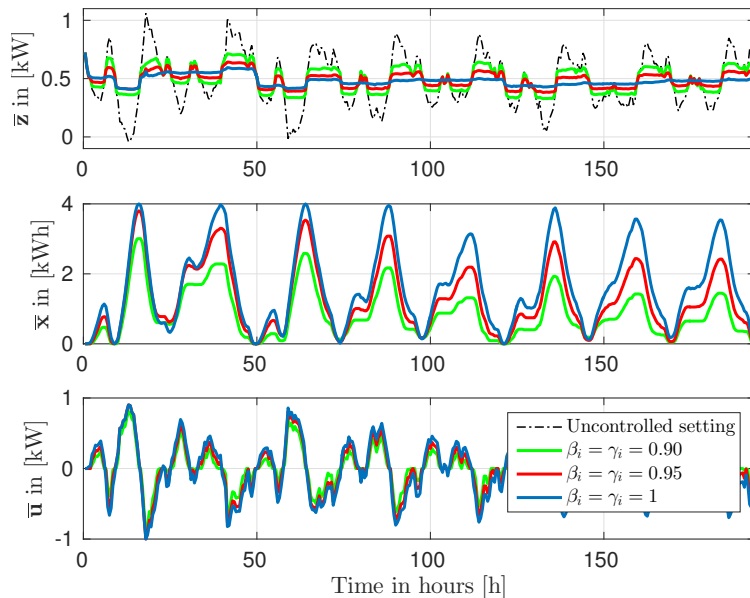


Figure 9. Closed loop trajectories of the non-cooperative control problem for different parameters β_i and γ_i for the case that 50% of the RESs are equipped with a battery.

In Figures 10 and 11, the closed loop counterparts of Figures 6 and 7 are presented. Again, we obtain the same qualitative results as in the open loop setting. Nevertheless, in Figure 11 we observe that the average savings are significantly lower compared to the open loop problem. Note that this occurs due to the high peaks on the first day of the simulation coupled with the assumption

⁵The number of time steps is chosen such that $x_i(0) = x_i(\mathcal{N}) = 0$ for all $i \in \mathbb{N}_{\mathcal{L}}$. This is important to be able to calculate the overall savings of the individual systems. The simulation length $\mathcal{N} = 387$ satisfying $x_i(\mathcal{N}) = 0$ is found by simulating on a longer time interval.

$x_i(0) = 0$ and is not a feature of the closed loop simulation compared to the open loop simulation in general.

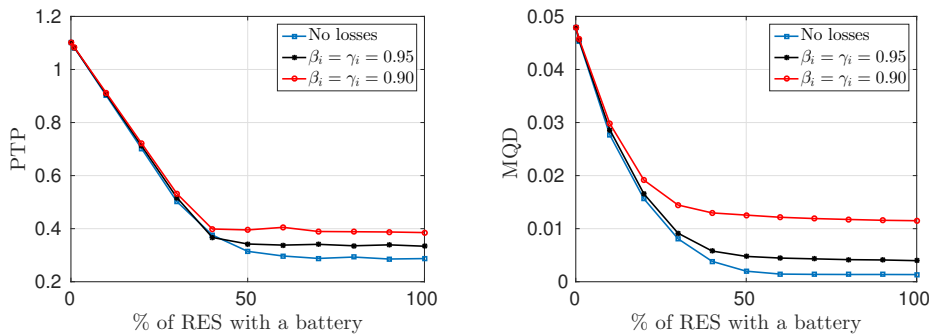


Figure 10. Performance measures of the closed loop control problem for a simulation of length $\mathcal{N} = 387$ with and without losses in the battery model.

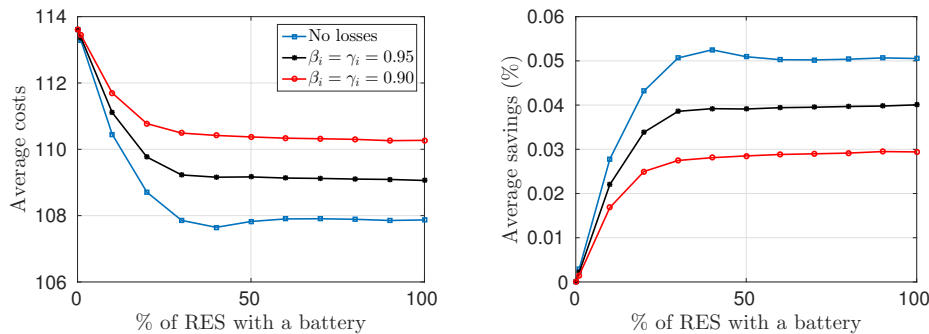


Figure 11. Average closed loop electricity costs for every RES and savings with respect to the number of installed batteries.

The closed loop counterpart of Figure 8 is given in Figure 12. We obtain the same qualitative behaviour as in the open loop case. The results from the point of view of a single RES are summarised additionally in Table 2. If RES 1 is the only RES with a storage device, the expected savings are high. The effect decreases with the increase of RESs equipped with storage devices.

Average savings					Average savings in percent				
RES	1%	20%	50%	100%	RES	1%	20%	50%	100%
1	23.69	11.69	4.74	3.49	1	15.93%	7.87%	3.19%	2.34%
2 – 20	0.10	13.02	7.27	6.26	2 – 20	0.09%	11.67%	6.52%	5.61%
21 – 50	0.12	2.77	6.35	5.25	21 – 50	0.11%	2.59%	5.92%	4.90%
50 – 100	0.08	2.97	4.91	5.89	50 – 100	0.07%	2.52%	4.18%	5.01%

Table 2. Average closed loop savings of the groups of RESs with respect to the percentage of RESs with batteries in the system. (The results correspond to the setting without losses.)

6. Conclusions & Outlook

In this paper we have shown that the distributed dual ascent algorithm converges in a non-cooperative setting to the globally optimal solution. Hence, the distributed dual ascent algorithm can be interpreted as a real-time pricing scheme, which allows an energy provider to control a network of RESs. Moreover, we conducted numerical simulations to investigate the impact of storage

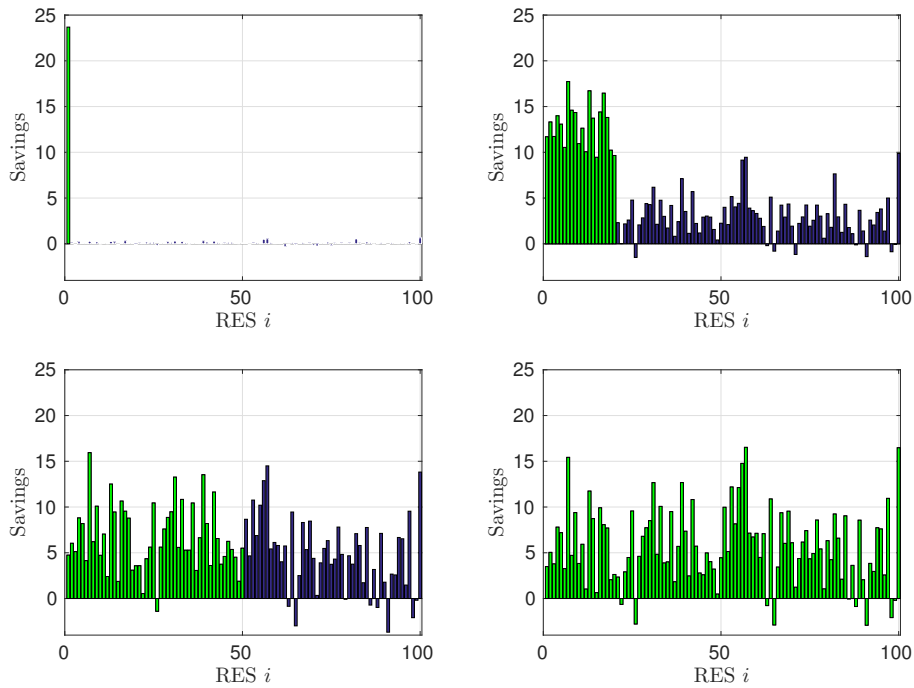


Figure 12. Closed loop savings of the individual RESs. Green bars indicate the RESs with batteries and blue bars indicate the RESs without batteries. Observe that not all RESs benefit from batteries. (The results correspond to the setting without losses.)

devices both on network-wide electricity demand patterns and on the individual electricity bill of each RES.

Whereas in the open loop simulations optimality of the solution of a single optimisation problem can be rigorously proven, performance of the closed loop solutions can only be validated numerically at present. Thus, a rigorous closed-loop performance analysis will be a topic for future research. Another interesting aspect is the tacitly assumed honesty, meaning that each RES reports its unadulterated predicted energy profiles. This, however, may not be the case since electricity prices can be influenced by suitably adapted predictions. Hence, we anticipate an extension of the proposed dual ascent algorithm such that it becomes robust with respect to manipulated information (adulterated predictions).

References

- Bertsekas, D. P. (1999). *Nonlinear programming* (D. P. Bertsekas, Ed.). Athena Scientific.
- Bertsekas, D. P., & Tsitsiklis, J. N. (1989). *Parallel and distributed computation: Numerical methods*. Belmont, MA, USA: Athena Scientific.
- Braun, P., Faulwasser, T., Grüne, L., Kellett, C. M., Weller, S. R., & Worthmann, K. (2016). Maximal islanding time for microgrids via distributed predictive control. In *Proc. 22nd International Symposium on Mathematical Theory of Networks and Systems* (pp. 652–659).
- Braun, P., Grüne, L., Kellett, C. M., Weller, S. R., & Worthmann, K. (2015). A real-time pricing scheme for residential energy systems using a market maker. In *Proc. 5th Australian Control Conference (AuCC)* (pp. 259–262).
- Braun, P., Grüne, L., Kellett, C. M., Weller, S. R., & Worthmann, K. (2016). A distributed optimization algorithm for the predictive control of smart grids. *IEEE Transactions on Automatic Control*. (DOI:10.1109/TAC.2016.2525808)

- Chen, H., Cong, T. N., Yang, W., Tan, C., Li, Y., & Ding, Y. (2009). Progress in electrical energy storage system: A critical review. *Progress in Natural Science*, 19(3), 291 - 312.
- Cheng, Y. C. (1987). Dual gradient method for linearly constrained, strongly convex, separable mathematical programming problems. *Journal of Optimization Theory and Applications*, 53(2), 237–246.
- Doan, M. D., Keviczky, T., Necoara, I., Diehl, M., & Schutter, B. D. (2009). A distributed version of Han’s method for DMPC using local communications only. *Control Engineering and Applied Informatics*, 11(3), 6–15.
- Doan, M. D., Keviczky, T., & Schutter, B. D. (2011). An iterative scheme for distributed model predictive control using Fenchel’s duality. *Journal of Process Control*, 21(5), 746–755.
- Floch, C. L., Belletti, F., Saxena, S., Bayen, A. M., & Moura, S. (2015). Distributed optimal charging of electric vehicles for demand response and load shaping. In *Proc. 54th IEEE Conference on Decision and Control (CDC)* (pp. 6570–6576).
- Ghahsifard, B., Basar, T., & Dominguez-Garcia, A. D. (2016). Price-based coordinated aggregation of networked distributed energy resources. *IEEE Transactions on Automatic Control*, 61(10), 2936–2946.
- Giselsson, P., Doan, M. D., Keviczky, T., Schutter, B. D., & Rantzer, A. (2013). Accelerated gradient methods and dual decomposition in distributed model predictive control. *Automatica*, 49(3), 829–833.
- Giselsson, P., & Rantzer, A. (2010). Distributed model predictive control with suboptimality and stability guarantees. In *Proc. 49th IEEE Conference on Decision and Control (CDC)* (pp. 7272–7277).
- Mohsenian-Rad, A.-H., & Leon-Garcia, A. (2010). Optimal residential load control with price prediction in real-time electricity pricing environments. *IEEE Transactions on Smart Grid*, 1(2), 120–133.
- Ratnam, E. L., Weller, S. R., & Kellett, C. M. (2013). An optimization-based approach for assessing the benefits of residential battery storage in conjunction with solar PV. In *Proc. IREP Symp. Bulk Power System Dynamics and Control–IX* (pp. 1–8).
- Ratnam, E. L., Weller, S. R., Kellett, C. M., & Murray, A. T. (2015). Residential load and rooftop PV generation: An Australian distribution network dataset. *International Journal of Sustainable Energy*, 1–20.
- Samadi, P., Mohsenian-Rad, A.-H., Schober, R., Wong, W. W. S., & Jatskevich, J. (2010). Optimal real-time pricing algorithm based on utility maximization for smart grid. In *Proc. 1st IEEE International Conference on Smart Grid Communications (SmartGridComm)* (pp. 415–420).
- Vytelingum, P., Voice, T. D., Ramchurn, S. D., Rogers, A., & Jennings, N. R. (2010). Agent-based micro-storage management for the smart grid. In *Proc. 9th International Conference on Autonomous Agents and Multiagent Systems (AAMAS)* (pp. 39–46).
- Worthmann, K., Kellett, C. M., Braun, P., Grüne, L., & Weller, S. R. (2015). Distributed and decentralized control of residential energy systems incorporating battery storage. *IEEE Transactions on Smart Grid*, 6(4), 1914–1923.
- Yang, P., Chavali, P., & Nehorai, A. (2012). Parallel autonomous optimization of demand response with renewable distributed generators. In *Proc. 3rd IEEE International Conference on Smart Grid Communications (SmartGridComm)* (pp. 55–60).