Available online at www.sciencedirect.com

ScienceDirect

IFAC Papers Online CONFERENCE PAPER ARCHIVE

IFAC PapersOnLine 53-2 (2020) 13269–13274

Distributed Control Enforcing Group Sparsity in Smart Grids ⋆

Philipp Sauerteig * **Yuning Jiang** *,** **Boris Houska** **
Karl Worthmann *

* *Institute for Mathematics, Technische Universität Ilmenau, Ilmenau, Germany*
(email: [philipp.sauerteig, karl.worthmann]@tu-ilmenau.de)
** *School of Information Science and Technology, ShanghaiTech University, Shanghai, China*
(email: [jiangyn, borish]@shanghaitech.edu.cn)

Abstract: In modern smart grids, charging of local energy storage devices is coordinated within the distribution grid to compensate the volatile aggregated power demand on the time interval of interest. However, this results in a perpetual usage of all batteries which in return reduces their lifetime. In this paper, we enforce group sparsity by using an $\ell_{p,q}$-regularization on the control to counteract this phenomenon. This leads to a non-smooth convex optimization problem, for which a tailored Alternating Direction Method of Multipliers algorithm is proposed. Furthermore, the algorithm is embedded in a Model Predictive Control framework. Numerical simulations show that the proposed scheme yields sparse control while achieving reasonable overall peak shaving.

Keywords: Distributed control, Predictive control, Structural optimization, Smart power applications

1. INTRODUCTION

The energy transition comes along with a fundamental transition of energy networks from centralized to decentralized power generation. As a result, residential storage devices have been installed in the grid to compensate the volatile generation via renewable energies and the accompanied bidirectional power flow, see e.g. Stecca et al. (2020) for an extensive review of the integration of batteries into the distribution gird. This paradigm shift in energy supply comes along with great optimization potential as discussed in Lezhniuk et al. (2019); Atzeni et al. (2013); Bolognani and Zampieri (2013). For instance, the integration of more and more renewables aggravates the mismatch between generation and demand forcing the grid operator to provide additional control energy, see Morstyn et al. (2018). Hence, one of the main goals in smart grid optimization is to coordinate the local energy storage devices at the household level in such a way that this mismatch is mitigated as illustrated in Worthmann et al. (2015). In Hubert and Grijalva (2011) the authors discuss the importance of optimization algorithms on a residential level.

State-of-the-art method to tackle optimal control problems in a receding horizon fashion is Model Predictive Control (MPC), see e.g. Worthmann et al. (2015) for an MPC approach in smart grids. Efficiently solving the inherent large-scale optimization problem online is crucial for designing a practical MPC scheme. To this end, distributed optimization algorithms exploiting the network structure have been widely applied (Boyd et al., 2011; Braun et al., 2018; Houska et al., 2016). In practice, distributed algorithms can achieve optimality by solving local problems in parallel and only exchanging certain data with a superordinate unit. Thus, distributed optimization builds a bridge between centralized and decentralized optimization (Worthmann et al., 2015). A classical approach is based on dual decomposition. A class of these techniques use first-order methods to solve the corresponding dual problem (Rantzer, 2009; Richter et al., 2011). Alternatively, semi-smooth Newton methods are applied, which require a line search sub-globalization routine (Frasch et al., 2015). A centralized consensus variant of the Alternating Direction Method of Multipliers (ADMM) has been applied to solve a convex control problem in smart grids in Braun et al. (2018). For an introduction to ADMM we refer to Boyd et al. (2011); Hong and Luo (2017). The approach in Braun et al. (2018) and related works, however, enforces all subsystems within the grid to repeatedly charge and discharge their batteries. Studies have shown that the lifespan of a battery decreases with the number of (dis-)charging cycles, see e.g. Xu et al. (2017); Ng et al. (2009) and the references therein. We aim to counteract this so-called cycle aging by enforcing sparsity of the optimal control.

Sparse optimal control has already been proposed to reduce the use of batteries: In Salem et al. (2017) the authors use an ultra sparse matrix rectifier for battery charging while in Jain et al. (2018) an ℓ_1 penalty is applied to enforce sparse communication of the linear quadratic regulator. A Newton-type method proposed in Polyak and Tremba (2019) was designed for solving a general sparse optimal control problem. However, these methods do not

⋆ The first two authors contributed equally.

10.1016/j.ifacol.2020.12.156

consider the scenario of numerous batteries incorporated in a structured network as in our application.

In this paper, we exploit group sparsity by introducing $\ell_{p,1}$ regularization terms in order to extend the lifetime of the batteries. Here, $p \in \{1, 2\}$ denotes the choice of the sparsity pattern. This $\ell_{p,1}$ regularization was introduced in Yuan and Lin (2006) in order to select grouped variables for accurate prediction in statistical learning. The $\ell_{p,1}$ regularization couples the control of each subsystem in time, which differs from the classical additive control regularization. The main advantage of group sparsity is that in each time step some batteries remain unused based on flexible weights. Updating these weights causes all batteries to be used in an alternating fashion while preserving a reasonable performance with respect to peak shaving. Thus, the number of (dis-)charging cycles of each battery is reduced and, hence, its lifespan is extended. For rigorous convergence proofs we refer to literature as such Boyd et al. (2011).

Section 2 recaps the basic model of residential energy systems. Here, for simplicity we concentrate on the problem of power balance proposed in Worthmann et al. (2014) and extend it in terms of sparsity. In Section 3 we elaborate how to design and implement an ADMM based optimization scheme to solve the underlying optimization problem online in a distributed manner. Our numerical open-loop results in Section 4 show the potential of this approach to prolong the batteries' lifespans while achieving a reasonable overall performance. Moreover, we investigate the closed-loop performance in an MPC scheme numerically.

Notation: Throughout this paper we use the notation $\mathbb{N} = \{1, 2, \ldots\}$ and $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$. Furthermore, we denote $[m : n] := [m, n] \cap \mathbb{N}_0$ for $m, n \in \mathbb{N}_0$ with $m \leq n$ and $\mathbf{1}_\ell = (1, 1, \ldots, 1)^\top \in \mathbb{R}^\ell$ for $\ell \in \mathbb{N}$; $\mathbf{0}_\ell$ is defined analogously. The Kronecker product of two matrices $A \in \mathbb{R}^{k\times l}$ and $B \in \mathbb{R}^{m\times n}$ is given by $A \otimes B = (a_{ij}B)_{i,j} \in \mathbb{R}^{km\times ln}$.

2. PROBLEM FORMULATION

This section recaps the basic model of smart grids consisting of residential energy systems, which are coordinated by a grid operator in order to achieve an overall optimum as described in Worthmann et al. (2014). Furthermore, we introduce a regularization in order to establish group sparsity with respect to the batteries and thus, prolong their lifespans. At the end, the resulting optimal sparse control problem is addressed as a distributed non-smooth convex optimization problem.

2.1 Residential energy systems

Let us consider a smart grid with $\mathcal{I}$, $\mathcal{I} \in \mathbb{N}$, residential energy systems. Each system incorporates its load, power generator, and battery serving as the energy storage device. As shown in Figure 1, the grid operator acts as a Central Entity (CE) compensating the local power demands.

The i-th subsystem, $i \in [1 : \mathcal{I}]$, is described by

$$x_i(n+1) = \alpha_i x_i(n) + T(\beta_i u_i^+(n) + u_i^-(n)) \,, \tag{1a}$$

$$z_i(n) = w_i(n) + u_i^+(n) + \gamma_i u_i^-(n) \,. \tag{1b}$$

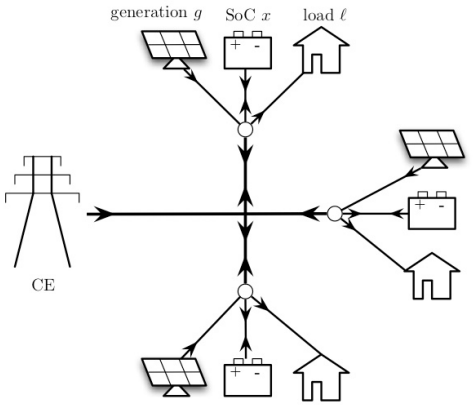


Fig. 1. Network of residential energy systems connected to the grid operator.

State $x_i(n)$ and control inputs $u_i(n) = \left(u_i^+(n)\ u_i^-(n)\right)^\top$ denote the State of Charge (SoC) in kWh, and the charging and discharging rate in kW at time instant $n \in \mathbb{N}_0$, respectively. The output $z_i(n)$ denotes the power demand in kW, and $w_i(n)$ is the net consumption without battery usage, i.e. load minus generation in kW. The parameter $T > 0$ represents the length of the time interval in h, i.e., $T = 0.5$ corresponds to 30 min, while $\alpha_i, \beta_i, \gamma_i \in (0, 1]$ model efficiencies with respect to self-discharge and energy conversion, respectively.

The SoC and the (dis-)charging rate are subject to

$$0 \leq x_i(n) \leq C_i \,, \tag{2a}$$

$$\underline{u}_i \leq u_i^-(n) \leq 0 \,, \tag{2b}$$

$$0 \leq u_i^+(n) \leq \overline{u}_i \,, \tag{2c}$$

$$0 \leq \frac{u_i^-(n)}{\underline{u}_i} + \frac{u_i^+(n)}{\overline{u}_i} \leq 1 \,, \tag{2d}$$

where $C_i \geq 0$ denotes the battery capacity. Note that the model description allows both charging and discharging of a battery during one time step. In order to ensure implementability of both box constraints (2b)–(2c), the stronger condition (2d) is introduced. Furthermore, we allow z_i to be negative, i.e., the subsystems are able to feed superfluous power to the grid. Subsystems without generation or storage device are covered by setting their generation or battery capacity to zero, respectively.

At the current time step k, the initial conditions are given by

$$x_i(k) = \hat{x}_i \,,\ i \in [1 : \mathcal{I}], \tag{3}$$

with $\hat{x}_i \in [0, C_i]$. Note that the future net consumption $w_i(n)$, $n \geq k$, is not known in advance. However, we assume that it can be predicted over the subsequent $N \in \mathbb{N}_{\geq 2}$ time steps and call N the prediction horizon. Then, we define the feasible sets

$$\mathbb{U}_i := \left\{ u_i \in \mathbb{R}^{2N} \,\middle|\, \begin{array}{l} \exists\, x_i(k) \,, \ldots \,, x_i(k+N-1) \in \mathbb{R} : \\ \text{initial condition (3) holds,} \\ \text{system dynamics (1a) and} \\ \text{constraints (2) hold for all} \\ \text{time instants } n \in [k : k+N-1] \end{array} \right\}$$

for all $i \in [1 : \mathcal{I}]$ with stacked control inputs

$$u_i = (u_i(k)^\top \ \ldots \ u_i(k+N-1)^\top)^\top \,.$$

Note that the sets $\mathbb{U}_i$ and hence, $\mathbb{U} = \mathbb{U}_1 \times \ldots \times \mathbb{U}_{\mathcal{I}}$ are (convex) polytopes.

2.2 Optimal peak shaving

From a grid operator's point of view it is desirable to provide a constant control energy. Therefore, our goal is to flatten the aggregated power demand profile. To this end, the deviation of the average power demand from a desired reference trajectory $\bar{\zeta} = (\bar{\zeta}(k) \ \ldots \bar{\zeta}(k+N-1))^\top \in \mathbb{R}^N$ is penalized, i.e.,

$$\begin{aligned} &\frac{1}{N} \sum_{n=k}^{k+N-1} \left\| \frac{1}{\mathcal{I}} \sum_{i=1}^{\mathcal{I}} z_i(n) - \bar{\zeta}(n) \right\|_2^2 \\ &\overset{(1b)}{=} \frac{1}{N} \sum_{n=k}^{k+N-1} \left\| \frac{1}{\mathcal{I}} \sum_{i=1}^{\mathcal{I}} \left(u_i^+(n) + \gamma_i u_i^-(n) \right) + \bar{w}(n) - \bar{\zeta}(n) \right\|_2^2 \end{aligned}$$

with averages

$$\bar{z}(n) = \frac{1}{\mathcal{I}} \sum_{i=1}^{\mathcal{I}} z_i(n) \ \text{ and } \ \bar{w}(n) = \frac{1}{\mathcal{I}} \sum_{i=1}^{\mathcal{I}} w_i(n) \ ,$$

$n \in [k : k+N-1]$. Assuming $k \geq N-1$, we choose $\bar{\zeta}$ to be the overall average net consumption

$$\bar{\zeta}(n) = \frac{1}{N} \sum_{j=n-N+1}^{n} \bar{w}(j) \ .$$

By introducing matrices

$$A_i = \frac{1}{\mathcal{I}} \cdot I_N \otimes (1 \ \gamma_i) \in \mathbb{R}^{N \times 2N} \ , \ i \in [1 : \mathcal{I}]$$

and the vector $b = \bar{\zeta} - \bar{w} \in \mathbb{R}^N$ the objective function value can be written as

$$\frac{1}{N} \left\| \sum_{i=1}^{\mathcal{I}} A_i u_i - b \right\|_2^2 \ . \tag{4}$$

In the context of group sparsity each $u_i \in \mathbb{R}^{2N}$ can be considered a group of $u = (u_1^\top \ \ldots \ u_{\mathcal{I}}^\top)^\top \in \mathbb{R}^{2N\mathcal{I}}$. In the following, we will enforce group-sparsity of u by introducing an $\ell_{p,1}$ regularization.

2.3 Group sparse control of batteries

Optimizing (4) with respect to all feasible $u \in \mathbb{U}$ typically results in perpetual charging and discharging of the batteries (Braun et al., 2018). We counteract this phenomenon by establishing group-sparse control, i.e., only a few batteries are active at each time step. To this end we use the weighted mixed $\ell_{p,1}$ norm, i.e.,

$$\begin{aligned} \|u\|_{p,1}^* \ &= \ \sum_{i=1}^{\mathcal{I}} \sigma_i \|u_i\|_p \\ &= \ \sum_{i=1}^{\mathcal{I}} \sigma_i \left(\sum_{n=k}^{k+N-1} \left(\left|u_i^+(n)\right|^p + \left|u_i^-(n)\right|^p \right) \right)^{1/p} \end{aligned}$$

with non-negative weights $\sigma = (\sigma_1 \ \ldots \ \sigma_{\mathcal{I}})^\top \in \mathbb{R}^{\mathcal{I}}$ and $1 \leq p < \infty$ (Yuan and Lin, 2006; Hu et al., 2017). This paper focuses on $p \in \{1, 2\}$. Compared to the $\ell_{2,1}$ case, the $\ell_{1,1}$ norm additionally enforces each non-zero group to be sparse. Hence, the total number of non-zero components is further reduced. In our case, u_i denotes the i-th group of u.

The main idea of exploiting group sparsity in smart grids is that at a fix time instant $n \in \mathbb{N}_0$ only a few batteries are used to manipulate the power demands. In practice, if the net consumption is below the reference trajectory, i.e., $\bar{w}(n) < \bar{\zeta}(n)$, we create an artificial demand by charging some batteries to compensate the gap, e.g. $u_i^+(n) > 0$ for some i and $u_i^-(n) = 0$ for all i, $i \in [1 : \mathcal{I}]$. Enforcing sparse control using $\ell_{2,1}$ or $\ell_{1,1}$ regularization yields that only the batteries with the most efficient charging rates are used. Analogous argumentation holds true for $\bar{w}(n) > \bar{\zeta}(n)$. To avoid this one-sided usage of the batteries, the weights σ are required to be updated online. In the context of MPC, this yields that the problem setup and thus the resulting group sparsity pattern vary.

2.4 Sparse control problem

Based on the considerations of the previous subsections we study the group sparse control problem

$$\begin{aligned} \min_{u} \ & \frac{1}{N} \left\| \bar{z} - \bar{\zeta} \right\|_2^2 + \kappa \sum_{i=1}^{\mathcal{I}} \left(\sigma_i \|u_i\|_p \right) \\ \text{s.t.} \ & \bar{z} = \sum_{i=1}^{\mathcal{I}} A_i u_i + \bar{w} \\ & u_i \in \mathbb{U}_i \ , \ i \in [1 : \mathcal{I}] \ , \end{aligned} \tag{5}$$

where the parameter $\kappa > 0$ denotes the trade-off between the optimal peak shaving and joint sparse activity of the batteries. Here, the stage costs of each battery based on the ℓ_p norm yield a coupling in $\{u_i(k), \ldots, u_i(k+N-1)\}$ when $p > 1$. This differs from the classical MPC formulation, which takes additive stage costs with respect to time steps into account. Substituting (4) into (5) yields the standard composite optimization form

$$\begin{aligned} \min_{v,u} \ & \frac{1}{N} \left\| \sum_{i=1}^{\mathcal{I}} A_i v_i - b \right\|_2^2 + \sum_{i=1}^{\mathcal{I}} \tilde{\sigma}_i \left\| u_i \right\|_p \\ \text{s.t.} \ & \begin{cases} v_i = u_i \ \ | \ \lambda_i \ , \ i \in [1 : \mathcal{I}] \\ u_i \in \mathbb{U}_i \ , \ i \in [1 : \mathcal{I}] \end{cases} \end{aligned} \tag{6}$$

with $\tilde{\sigma}_i = \kappa \cdot \sigma_i$. Here, we introduce an auxiliary variable $v \in \mathbb{R}^{2N\mathcal{I}}$ and denote by λ_i the Lagrangian multipliers of the constraints $v_i = u_i$. Note that, operator splitting methods have already been developed for solving (6) in a distributed manner; for a detailed theoretical analysis we refer to (Boyd et al., 2011, Sec. 3.2). In the next section, we propose a distributed model predictive control scheme to solve (6) by using the state-of-the-art technique ADMM.

3. DISTRIBUTED SPARSE OPTIMIZATION

In this section we elaborate how to solve (6) in a distributed MPC scheme. First, Algorithm 1 outlines how to solve (6) via ADMM in a distributed manner. The subsystems solve decoupled small-scale problems in parallel and the CE solves an unconstrained Quadratic Programming (QP). Then, the proposed method is embedded into a model predictive control scheme in Algorithm 2.

3.1 Alternating Direction Method of Multipliers

Referring to (Boyd et al., 2011, Sec. 3.1), Algorithm 1 with four main steps outlines to solve (6) by using ADMM.

In the first step the local primal variable u_i and dual variable λ_i are updated in parallel. Then, we solve an unconstrained QP in the consensus step. Note that the solution map is worked out analytically. In order to check the terminal condition, the primal and dual residual r^{pri}, r^{dual} are evaluated in Step 3. Here, in contrast to Boyd and Vandenberghe (2004) we do not use the relative tolerance but a fixed $\varepsilon > 0$. In order to speed up the convergence, we utilize an adaptive strategy to update the dual step size ρ^m. This heuristic increases ρ^m if r^{pri} decreases faster than r^{dual} and vice versa, see e.g. (Boyd et al., 2011, Sec. 3.4.1) for a possible choice of the tuning parameters η and μ. Algorithm 1 requires the grid operator to collect $4N$ floats information from each subsystem and spread $2N+1$ back to the subsystems per iteration. Note that the grid operator does not require to have any information on the local system model.

Algorithm 1 ADMM for solving (6)

Input: initial guesses (u^0, v^0, λ^0) and step size $\rho^0 > 0$, stop tolerance $\varepsilon > 0$, tuning parameter $\eta, \mu > 0$.

For $m = 0 : \text{MaxIte}$

1) *Parallel Step:* Compute for all $i \in [1 : \mathcal{I}]$ in parallel

$$u_i^{m+1} = \underset{u_i \in \mathbb{U}_i}{\arg\min}\ \tilde{\sigma}_i \|u_i\|_p + \frac{\rho^m}{2} \left\| u_i - \frac{\lambda_i^m}{\rho^m} - v_i^m \right\|_2^2$$
$$\lambda_i^{m+1} = \lambda_i^m + \rho^m (v_i^m - u_i^{m+1})$$

2) *Consensus Step:* Solve unconstrained QP

$$v^{m+1} = \underset{v}{\arg\min} \frac{1}{N}\|Av - b\|_2^2 + \frac{\rho}{2} \left\| v - u^{m+1} + \frac{\lambda^{m+1}}{\rho} \right\|_2^2$$
$$= \left(\frac{2}{N} A^\top A + \rho^m I \right)^{-1} \left(\frac{2}{N} A^\top b - \lambda^{m+1} + \rho^m u^{m+1} \right)$$

3) *Stop Criterion:* Evaluate

$$r^{\text{pri}} = \rho \|u^{m+1} - v^{m+1}\|_2\ ,\ r^{\text{dual}} = \rho \|(v^{m+1} - v^m)\|_2.$$

If $r^{\text{pri}} \le \varepsilon$ and $r^{\text{dual}} \le \varepsilon$, then terminate.

4) *Adpative Dual Step Size:* Update ρ^{m+1} by

$$\rho^{m+1} \leftarrow \begin{cases} \eta \rho^m & \text{if } r^{\text{pri}} \ge \mu r^{\text{dual}} \\ \rho^m / \eta & \text{if } r^{\text{dual}} \ge \mu r^{\text{primal}} \\ \rho^m & \text{otherwise} \end{cases}$$

End

3.2 Local solver

At the parallel step of Algorithm 1, the u_i update requires to solve a constrained lasso[1] problem. We propose two efficient local solvers to update u_i depending on the case $p \in \{1, 2\}$.

(1) If $p = 1$, the ℓ_1 term in the objective can be reformulated into the constraints Boyd and Vandenberghe (2004) by introducing auxiliary variables $s_i \in \mathbb{R}^{2N}$. This yields the decoupled QP

$$\begin{aligned} \min_{s_i, u_i} \quad & \tilde{\sigma}_i \mathbf{1}^\top s + \frac{\rho^m}{2} \left\| u_i - \frac{\lambda_i^m}{\rho^m} - v_i^m \right\|_2^2 \\ \text{s.t.} \quad & u_i \in \mathbb{U}_i\ ,\ -s \le u \le s\ , \end{aligned} \tag{7}$$

which allows for direct usage of existent QP solvers such as `qpOASES` as in Ferreau et al. (2014).

[1] least absolute shrinkage and selection operator

(2) If $p = 2$, we propose to use a local ADMM solver as follows,

$$\begin{aligned} s_i &= \mathcal{S}_{\tilde{\sigma}_i/\rho^m} \left(v_i^m + u_i^j + \frac{\lambda_i^m - \xi^j}{\rho^m} \right)\ , \\ u_i^{j+1} &= \underset{u_i \in \mathbb{U}_i}{\arg\min}\ \frac{\rho^m}{2} \left\| u_i - s_i - \frac{\xi_i^j}{\rho^m} \right\|_2^2\ , \\ \xi_i^{j+1} &= \xi_i^j + \rho^m (u_i^{j+1} - s_i)\ , \end{aligned} \tag{8}$$

where superscript j represents the iteration of the inner ADMM loop, $\mathcal{S}_a : \mathbb{R}^{2N} \to \mathbb{R}^{2N}$ denotes the *soft thresholding operator* defined by

$$\mathcal{S}_a(x) = \max\{1 - a/\|x\|_2, 0\}\, x\ .$$

Here, the omitted terminal condition is analogous to Step 3) in Algorithm 1 and a fixed dual step size consistent with the current ρ^m is applied.

3.3 Distributed predictive sparse control

The model predictive control scheme requires to solve (6) during each sampling time based on the current measurements. Embedding Algorithm 1, Algorithm 2 outlines an ADMM based distributed predictive sparse control scheme.

Algorithm 2 Distributed predictive control scheme

Offline:

- Initial guess (u^0, λ^0), set $k = 0$, $v^0 = u^0$ and choose weights σ and a tolerance $\varepsilon > 0$.

Online:

1) **Subsystems** measure current SoC $x_i(k)$, predict future net consumption w_i and send it to grid operator.
2) **Grid Operator** computes the reference trajectory $\bar{\zeta}$.
3) Optionally update weights σ. Run Algorithm 1 for solving (6) to obtain u^* and λ^*.
4) **Subsystems** apply $u_i^*(k)$ if $\|u_i(k)\|_2 \ge \varepsilon$ and 0 otherwise.
5) Reinitialize

$$u_i^0 = (u_i^*(k+1)^\top \ \dots\ u_i^*(k+N-1)^\top\ 0_2^\top)^\top$$
$$\lambda_i^0 = (\lambda_i^*(k+1) \ \dots\ \lambda_i^*(k+N-1)\ 0)^\top$$

for all $i \in [1 : \mathcal{I}]$. Then, set $k \leftarrow k+1$ and go to Step 1).

In order to achieve different sparsity patterns in each time step we optionally update the weights σ in Step 3). In practice, we run Algorithm 1 to a predetermined numerical accuracy such that we choose a tolerance for applying the sparse control at Step 4) in Algorithm 2. Step 5) in Algorithm 2 is a warm-start step, which improves the online convergence performance of Algorithm 1 (Braun et al., 2018). In the following section, we will illustrate the numerical performance of Algorithm 2 by applying it to benchmark problems.

4. NUMERICAL RESULTS

In this section we compare the numerical results of $\ell_{1,1}$ and $\ell_{2,1}$ norm for both open-loop and closed-loop control. To this end we consider heterogeneous systems with randomly generated parameters according to Table 1. Note that

the values were not chosen completed arbitrarily, but inspired by the choice in Braun et al. (2018). However, in our simulations the heterogeneity of the batteries is essential. Furthermore, we set $T = 0.5$ [h], $N = 24$, and

Table 1. Parameters for implementation.

	expected value	standard deviation
C_i	2.0563 [kWh]	0.2431 [kWh]
$\overline{u}_i$	0.5229 [kW]	0.1563 [kW]
$\underline{u}_i$	−0.5105 [kW]	0.1474 [kW]
α_i	0.9913	0.0053
β_i	0.9494	0.0098
γ_i	0.9487	0.0100

$\hat{x}_i = 0.5$ [kWh] for all $i \in [1 : \mathcal{I}]$. The data representing the net consumption is provided by an Australian grid operator (Ratnam et al., 2017).

4.1 Open-loop optimal control

Let us have a closer look at the open-loop performance in this section. Table 2 illustrates the impact of the number of systems $\mathcal{I}$ for the solution sparsity with fix $\kappa = 10^{-3}$. We ran a case study for different weights σ and listed the mean value, the standard deviation, and the median of the percentage of the non-zero components of the optimal control. One can see that the larger the grid the higher the sparsity rate.

Table 2. Impact of the number of subsystems $\mathcal{I}$ on the percentage of non-zero components of the optimal control $u \in \mathbb{R}^{2N\mathcal{I}}$ for $\kappa = 10^{-3}$.

	$\mathcal{I}$	mean	std dev	median
$\ell_{2,1}$	25	30.67	0.85	30.33
	50	24.41	0.46	24.17
	100	18.73	0.26	18.75
$\ell_{1,1}$	25	8.48	0.36	8.38
	50	4.65	0.10	4.65
	100	3.82	0.06	3.81

In the following, we fix $\mathcal{I} = 50$. Therefore, the total number of control variables is 2400. Figure 2 visualizes the impact of the choice of the regularization and the size of the weighting parameter $\kappa \in \{10^{-4}, 10^{-3}\}$ on the open-loop sparsity pattern. Increasing κ the batteries are used less often. For sufficiently large κ some stay even inactive over the whole prediction horizon. Note that the $\ell_{1,1}$ regularization enforces $u_i^+(n) \cdot u_i^-(n) = 0$ for all $i \in [1 : \mathcal{I}]$, $n \in [k : k+N-1]$, i.e., only charging or discharging at one time instant. This is not the case if an $\ell_{2,1}$ regularization is used.

Figure 3 shows the overall performance with respect to (4) depending on $\kappa \in [10^{-5}, 10^{-2}]$. More precisely, the relative deviation from the solution associated with $\kappa = 0$ is depicted. The larger κ, i.e., the less batteries are active, the worse the performance. The $\ell_{1,1}$ regularization with $\kappa = 10^{-4}$ achieves reasonable performance with respect to peak shaving while establishing sparsity.

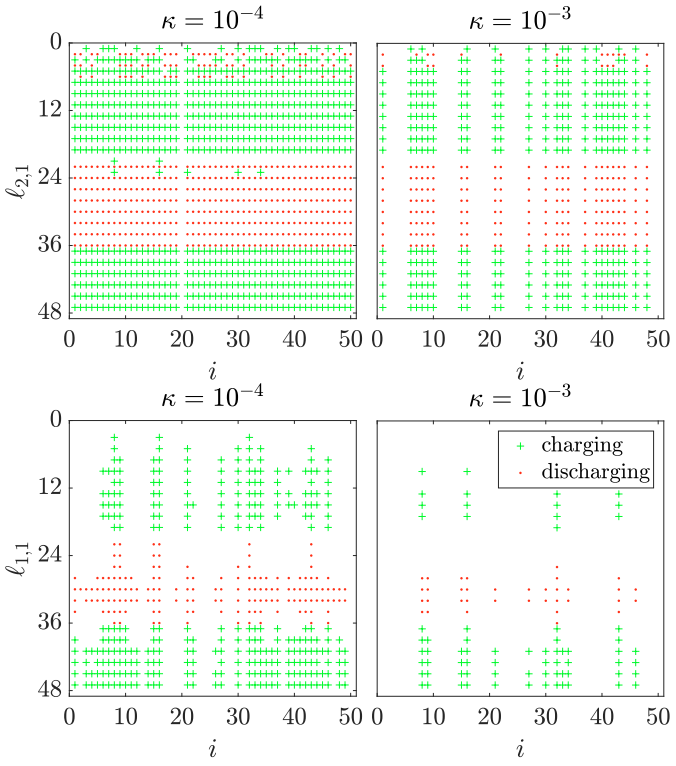
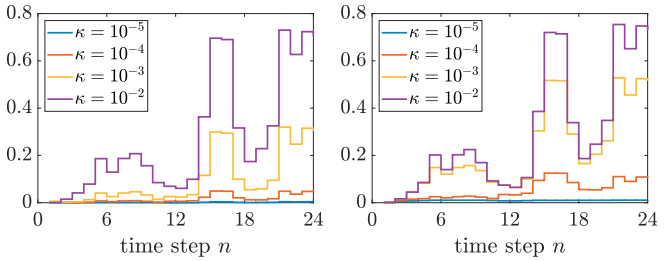


Fig. 2. Impact of the choice of κ and the regularization on the open-loop sparsity pattern for $\mathcal{I} = 50$ systems. The y-axis denotes the components of the control vector $u_i \in \mathbb{R}^{2N}$. Amount of non-zero control values: 51%, 24%, 19%, 5%, respectively.



(a) $\ell_{2,1}$ (b) $\ell_{1,1}$

Fig. 3. Relative deviation $|\bar{z}(n;\kappa) - \bar{z}(n;0)| \,/\, \|\bar{z}(\cdot\,;0)\|_\infty$ of $\bar{z}(\cdot\,;\kappa)$ from $\bar{z}(\cdot\,;0)$ for $\mathcal{I} = 50$ systems.

4.2 MPC closed loop

This section illustrates the closed-loop performance of Algorithm 2 depending on the choice of the penalty term and the weighting parameter. Note that if the weights $\sigma = \sigma(k)$ are changed in every single MPC step k, group sparsity cannot be established, since in each step different devices might be active compared to the previous one. Therefore, in our implementation, we generate varying weights every three hours by using the `MATLAB` command `randn`, which yields normally distributed random numbers, i.e.,

$$\sigma_i(k) \sim \mathcal{N}(0, 1).$$

For the same reason we omit disturbances in the forecasting variables w_i. Note that the choice of the parameters σ as well as the update frequency is not yet optimized. The percentage of non-zero components of the optimal solution is approximately 34% and 8% for the $\ell_{2,1}$ and

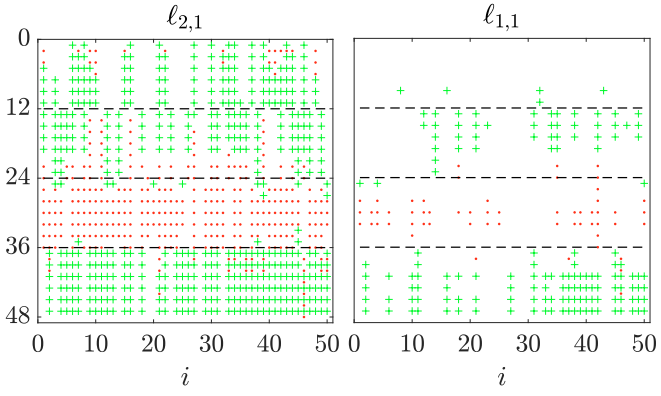


Fig. 4. Closed-loop sparsity pattern for $\mathcal{I} = 50$ systems and $\kappa = 10^{-3}$. The dashed black lines indicate new weights σ.

the $\ell_{1,1}$ case, respectively. Hence, similar to the open-loop simulation, the $\ell_{1,1}$ solution is sparser than the $\ell_{2,1}$ solution. Keep in mind that the open-loop solutions involve several devices to be inactive while others are active the whole time, see Figure 2. Thanks to the updated weights σ this phenomenon does not occur in the closed loop, see Figure 4. After each three hours time interval the sparsity might change. Note that in the plain peak-shaving scenario, i.e. for $\kappa = 0$, more than half of the control values in closed loop (60%) are non-zero indicating the benefit of the proposed method.

5. CONCLUSIONS

In this paper we considered a smart grid optimization problem dealing with optimal control of distributed energy storage devices. We proposed a sparse control problem formulation and designed a distributed optimization scheme to solve it. Our numerical results show that this approach is able to reduce the usage of the batteries in order to prolong their life time in a receding horizon fashion. Future research might consider more complex models, e.g. incorporating the power flow equation. Furthermore, more sophisticated methods to choose the weighting parameters σ might be beneficial.

ACKNOWLEDGEMENTS

Yuning Jiang, Philipp Sauerteig and Karl Worthmann are supported by the German Federal Ministry for Education and Research (BMBF; Grant 05M18SIA). Yuning Jiang and Boris Houska are supported by ShanghaiTech University under Grant No. F-0203-14-012. Karl Worthmann is also indebted to the German Research Foundation (DFG-grant WO 2056/6-1).

REFERENCES

Atzeni, I., Ordóñez, L.G., Scutari, G., Palomar, D.P., and Fonollosa, J.R. (2013). Noncooperative and cooperative optimization of distributed energy generation and storage in the demand-side of the smart grid. *IEEE Transactions on Signal Processing*, 61(10), 2454–2472.

Bolognani, S. and Zampieri, S. (2013). A distributed control strategy for reactive power compensation in smart microgrids. *IEEE Transactions on Automatic Control*, 58(11), 2818–2833.

Boyd, S., Parikh, N., Chu, E., Peleato, B., and Eckstein, J. (2011). Distributed Optimization and Statistical Learning via the Alternating Direction Method of Multipliers. *Foundation Trends in Machine Learning*, 3(1), 1–122.

Boyd, S. and Vandenberghe, L. (2004). *Convex optimization*. Cambridge University Press.

Braun, P., Faulwasser, T., Grüne, L., Kellett, C.M., Weller, S.R., and Worthmann, K. (2018). Hierarchical distributed ADMM for predictive control with applications in power networks. *IFAC Journal of Systems and Control*, 3, 10–22.

Ferreau, H.J., Kirches, C., Potschka, A., Bock, H.G., and Diehl, M. (2014). qpOASES: A parametric active-set algorithm for quadratic programming. *Mathematical Programming Computation*, 6(4), 327–363.

Frasch, J.V., Sager, S., and Diehl, M. (2015). A parallel quadratic programming method for dynamic optimization problems. *Mathematical Programming Computation*, 7(3), 289–329.

Hong, M. and Luo, Z.Q. (2017). On the linear convergence of the alternating direction method of multipliers. *Mathematical Programming*, 162(1-2), 165–199.

Houska, B., Frasch, J.V., and Diehl, M. (2016). An augmented Lagrangian based algorithm for distributed nonconvex optimization. *SIAM Journal on Optimization*, 26(2), 1101–1127.

Hu, Y., Li, C., Meng, K., Qin, J., and Yang, X. (2017). Group sparse optimization via $\ell_{p,q}$-regularization. *The Journal of Machine Learning Research*, 18(1), 960–1011.

Hubert, T. and Grijalva, S. (2011). Realizing smart grid benefits requires energy optimization algorithms at residential level. In *ISGT 2011*, 1–8.

Jain, A., Chakrabortty, A., and Biyik, E. (2018). Structurally Constrained ℓ_1-Sparse Control of Power Systems: Online Design and Resiliency Analysis. In *2018 Annual American Control Conference (ACC)*, 4195–4200.

Lezhniuk, P., Kravchuk, S., and Buslavets, O. (2019). Selfoptimization Modes of Electric Grids with Renewable Energy Sources Using the Principle of Least Action. In *2019 IEEE 6th International Conference on Energy Smart Systems (ESS)*, 33–36.

Morstyn, T., Hredzak, B., and Agelidis, V.G. (2018). Control strategies for microgrids with distributed energy storage systems: An overview. *IEEE Transactions on Smart Grid*, 9(4), 3652–3666.

Ng, K., Moo, C.S., Chen, Y.P., and Hsieh, Y.C. (2009). Enhanced Coulomb counting method for estimating state-of-charge and state-of-health of lithium-ion batteries. *Applied Energy*, 86, 1506–1511.

Polyak, B. and Tremba, A. (2019). Sparse solutions of optimal control via Newton method for under-determined systems. *Journal of Global Optimization*, 1–11.

Rantzer, A. (2009). Dynamic dual decomposition for distributed control. In *2009 American Control Conference*, 884–888.

Ratnam, E.L., Weller, S.R., Kellett, C.M., and Murray, A.T. (2017). Residential load and rooftop PV generation: an australian distribution network dataset. *International Journal of Sustainable Energy*, 36(8), 787–806.

Richter, S., Morari, M., and Jones, C.N. (2011). Towards computational complexity certification for constrained MPC based on lagrange relaxation and the fast gradient method. In *50th IEEE Conference on Decision and Control and European Control Conference*, 5223–5229.

Salem, M., Atia, Y., and Mahgoub, O. (2017). Ultra sparse matrix rectifier for battery charging application. In *2017 Intl Conf on Advanced Control Circuits Systems (ACCS) & 2017 Intl Conf on New Paradigms in Electronics & Information Technology (PEIT)*, 305–310.

Stecca, M., R. Elizondo, L., B. Soeiro, T., Bauer, P., and Palensky, P. (2020). A Comprehensive Review of the Integration of Battery Energy Storage Systems into Distribution Networks. *IEEE Open Journal of the Industrial Electronics Society*. DOI 10.1109/OJIES.2020.2981832.

Worthmann, K., Kellett, C.M., Braun, P., Grüne, L., and Weller, S.R. (2015). Distributed and decentralized control of residential energy systems incorporating battery storage. *IEEE Transactions on Smart Grid*, 6(4), 1914–1923.

Worthmann, K., Kellett, C.M., Grüne, L., and Weller, S.R. (2014). Distributed control of residential energy systems using a market maker. *IFAC Proceedings Volumes*, 47(3), 11641–11646.

Xu, B., Zhao, J., Zheng, T., Litvinov, E., and Kirschen, D.S. (2017). Factoring the cycle aging cost of batteries participating in electricity markets. *IEEE Transactions on Power Systems*, 33(2), 2248–2259.

Yuan, M. and Lin, Y. (2006). Model selection and estimation in regression with grouped variables. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 68(1), 49–67.