ADMM-based Unit and Time Decomposition for Price Arbitrage by Cooperative Price-Maker Electricity Storage Units.

Miguel F. Anjos*[1], James R. Cruise†[2] and Albert Solà Vilalta ‡[1]

[1]Maxwell Institute for Mathematical Sciences, School of Mathematics, University of Edinburgh, Edinburgh, United Kingdom.
[2]Riverlane Research, Cambridge, United Kingdom

October 21, 2021

Abstract

Decarbonization via the integration of renewables poses significant challenges for electric power systems, but also creates new market opportunities. Electric energy storage can take advantage of these opportunities while providing flexibility to power systems that can help address these challenges. We propose a solution method for the optimal control of multiple price-maker electric energy storage units that cooperate to maximize their total profit from price arbitrage. The proposed method can tackle the nonlinearity introduced by the price-maker assumption. The main novelty of the proposed method is the combination of a decomposition by unit and a decomposition in time. The decomposition by unit is based on the Alternating Direction Method of Multipliers and breaks the problem into several one-unit subproblems. Every subproblem is solved using an efficient algorithm for one-unit problems from the literature that exploits an on the fly decomposition in time, and this results in a time decomposition for the whole solution method. Our numerical experiments show very promising performance in terms of accuracy and computational time. In particular, they suggest that computational time scales linearly with the number of storage units.

Key words: control, multiple storage units, price arbitrage, price-maker, energy, cooperative problem, ADMM, horizons.

1 Introduction

Carbon emission reductions are key to limiting the effects of climate change. Under the Paris Agreement (United Nations Framework Convention on Climate Change

*anjos@stanfordalumni.org
†james.cruise@riverlane.io
‡a.sola01@alumni.ub.edu

(2015)), most nations worldwide agreed to keep the temperature rise below 2 degrees Celsius, and preferably below 1.5, compared to pre-industrial levels. This poses unprecedented challenges in many sectors, but there is a general consensus that electrification is crucial to decarbonize our economies (Dennis (2015), International Energy Agency (2021)). We are interested in the role grid-scale electric energy storage can play in this transformation.

In recent years, there has been an increasing interest in electric energy storage. Graves et al. (1999) claim that deregulation of electricity markets creates a more favourable environment for energy storage. Sioshansi et al. (2009) validate this claim by observing an increase in the arbitrage value of energy storage after deregulation in the PJM Interconnection (US) during the 2000s. Moreover, the report by the National Renewable Energy Laboratory (2010) identifies four other reasons for this increasing interest: advances in storage technologies, increase in fossil fuel prices, challenges to siting new transmission and distribution facilities, and opportunities for storage with increasing variable renewable generation.

Electric energy storage can provide many services to the power system, such as price arbitrage, operating reserves, firm capacity, network reinforcement deferral, black-start support, power quality and stability, and aid in the integration of renewables (National Renewable Energy Laboratory (2010)). It is difficult to estimate the value for most of these services. Andrés-Cerezo and Fabra (2020) and Wogrin et al. (2021) present case studies that show how energy storage can aid in the integration of renewables maximizing social welfare, but at the same time the storage owner struggles to recover the investment. A common approach in the literature is to consider price arbitrage only (Sioshansi et al. (2009), Barbry et al. (2019), Cruise et al. (2019)). We will follow this approach, understanding price arbitrage as the ability to make profit by buying electricity when it is cheap, and selling it when it is expensive.

A key feature we want to capture in our model is the possibility to model price-makers, i.e., storage units of sufficient magnitude to affect electricity market-clearing prices. Nowadays, most examples of price-maker storage units are pumped-storage hydro stations, since they can deliver power in the order of GW (Barbour et al. (2016)). Nevertheless, the decreasing prices and rapid advances in storage technologies observed in the reports by the National Renewable Energy Laboratory (2010) and International Energy Agency (2018) imply that, in the near future, storage units using other technologies might become price-makers.

In the present paper, we study how to optimally coordinate a set of price-maker electric energy storage units to maximize profit from price arbitrage. We extend the models by Sioshansi et al. (2009) and Cruise et al. (2019) with one storage unit and the model by Anjos et al. (2020) with two storage units to handle an arbitrary number of units. We present a novel solution method to deal with multiple units that exploits a unit decomposition based on the Alternating Direction Method of Multipliers (ADMM) (Boyd et al. (2011)), and the decomposition in time the algorithm by Cruise et al. (2019) provides to solve the one unit case.

We are interested in the cooperative problem because it provides a benchmark to analyse the benefits of storage units to the power system. Furthermore, in electricity markets with zonal pricing, a utility company might operate several grid-scale storage units in the same zone, which makes this problem relevant for them. Barbour et al. (2016) give examples of these companies in several countries and describe the markets they participate in. For related work on coordination of several storage units or demand response, see Schill and Kemfert (2011), Biegel et al. (2013), Zhang et al. (2013), Del-Rosario-Calaf et al. (2016) and Anjos et al. (2020). Demand response is similar to storage because both shift energy through time, but it presents different practical difficulties.

Following Graves et al. (1999), Sioshansi et al. (2009) and Cruise et al. (2019), we assume an electric energy storage unit is characterized by an energy capacity, power input and output rates, a round-trip efficiency and leakage. The price-maker assumption introduces a nonlinearity in the model, making it challenging to solve. Our model is technology agnostic, and can be used with any storage technology. It can account for both positive and negative electricity prices.

Future electricity prices are uncertain. This uncertainty can be incorporated in the model by introducing some stochasticity (Kazempour et al. (2009), Secomandi (2010), Baslis and Bakirtzis (2011), El-Ghandour and Johnson (2017)). Alternatively, a deterministic model with forecasted prices can be used, reoptimizing once more accurate forecasts become available (Wu et al. (2012), Shafiee et al. (2016), Barbry et al. (2019) and Cruise et al. (2019)). We choose the second option, because Secomandi et al. (2015) argue that assumed probability distributions might be incorrect, and Secomandi (2015) provides evidence of the benefits of a reoptimization approach.

Furthermore, the decomposition in time we do based on Cruise et al. (2019) limits the price forecast information. It combines well with a reoptimization approach, because one can obtain new forecasts and reoptimize the actions of the storage units taking into account this decomposition. We stress that this time decomposition is not a standard rolling horizon approach, where the horizons are set in advance, like the ones described in Sethi and Sorger (1991). Instead, they are obtained on the fly while solving the problem.

In most real-world settings, however, generators are scheduled over fixed-length time periods set in advance (Conejo et al. (2002)). For instance, in day-ahead markets, only next day actions need to be determined. In this setting, our approach gives valuable information on the desirable states of charge of the storage units at the end of the scheduling period by optimizing beyond that point. It also controls how long we should go beyond the end of the scheduling period. This is relevant for storage units that take a relatively long time to fully charge and discharge compared to the length of the scheduling period. Exemples of these units are Cruachan and Foyers pumped-storage stations in Great Britain participating in the day-ahead market, since it takes them more than 20 hours (see Table 2 in Section 6.2) to fully charge or discharge. Our approach allows them to take advantage of both interday and intraday price differences.

Our numerical examples suggest that computational times scale linearly with the number of units and show two orders of magnitude computational time improvements when compared to the solution method by Anjos et al. (2020) for the two storage units case. We are also able to incorporate round-trip efficiencies and leakage, which was not straightforward to do in Anjos et al. (2020). We believe that the main driver for this performance improvement is the combined unit and time decomposition.

This paper is organized in six sections. After this introductory section, Sections 2 and 3 introduce the mathematical formulations of our problem of interest, the N-Units Problem, and a relaxation of it, the Aggregated Unit Problem, respectively. In Section 4 we present the decomposition by unit and in time. In Section 5 we deduce Lagrangian Sufficiency Conditions for optimality, which we will use to analyse our numerical examples in Section 6. Finally, we draw conclusions and suggest future work in Section 7.

2 The Model

Consider a set $\mathcal{S}$ of N electric energy storage units. Following Graves et al. (1999), Sioshansi et al. (2009) and Cruise et al. (2019), we assume that an *electric energy storage unit*, or *unit* for short, is characterized by a 5-tuple $(E, P^i, P^o, \eta, \rho)$, where E denotes the *energy capacity*, P^i the *power input rate*, P^o the *power output rate*, η the *round-trip efficiency* and $1-\rho$ the *leakage*.

The *energy capacity* E is the maximum amount of energy that can be stored in the unit. The *power input* (resp. *output*) *rate* P^i (resp. P^o) is the maximum rate at which energy can be charged (resp. discharged). The *round-trip efficiency* η is the fraction of energy charged available when discharging; it models inefficiencies in the charging and discharging processes. The *leakage* $1-\rho$ is the fraction of energy lost in one time step, so that ρ represents the fraction of energy kept in the unit over one time step. It models inefficiencies while storing energy. We introduce subscripts to identify which unit we refer to, so $(E_j, P^i_j, P^o_j, \eta_j, \rho_j)$ denotes the parameters of unit $j \in \mathcal{S}$.

The ratios E_j/P^i_j and E_j/P^o_j give valuable information of a storage unit. They represent how long it takes unit $j \in \mathcal{S}$ to charge from empty if done at maximum rate P^i_j and to discharge from full if done at maximum rate P^o_j, respectively. For every $j, k \in \mathcal{S}$ such that $1 \le j < k \le N$, we assume

$$\frac{E_j}{P^i_j} \ge \frac{E_k}{P^i_k}, \quad \frac{E_j}{P^o_j} \ge \frac{E_k}{P^o_k}. \tag{1}$$

In other words, if a unit takes longer to fully charge than another unit, it should also take longer to fully discharge. (1) also sets the convention that a *slower* unit has a lower index in $\mathcal{S}$.

We consider a time discretization $\mathcal{T} = \{1, \dots, T\}$. For every $t \in \mathcal{T}$, unit $j \in \mathcal{S}$ can be charged or discharged by a certain amount $x_{j,t} \in [-P^o_j, P^i_j]$. We group all actions in a $N \times T$ matrix $x \in \mathcal{M}_{N\times T}(\mathbb{R})$, where rows $x_j = (x_{j,1}, \dots, x_{j,T}) \in [-P^o_j, P^i_j]^T$ contain the actions of unit $j \in \mathcal{S}$ for all $t \in \mathcal{T}$ and columns $x_t = (x_{1,t}, \dots, x_{N,t}) \in \mathbb{R}^N$ contain the actions of all units $j \in \mathcal{S}$ at $t \in \mathcal{T}$. We set the convention that $x_{j,t} > 0$ represents charging and $x_{j,t} < 0$ represents discharging.

Let $S_{j,t-1}$ be the *State of Charge* (SoC) of unit $j \in \mathcal{S}$ at the end of time step $t-1$ and $x_{j,t}$ be the action taken by unit j during time step $t \in \mathcal{T}$. The SoC $S_{j,t}$ of unit j at the end of time step t is

$$S_{j,t} = \rho_j S_{j,t-1} + x_{j,t}, \tag{2}$$

which incorporates leakage losses. Similarly, given an initial SoC $\bar{S}_{j,0}$ and a set of actions $x_j \in [-P_j^o, P_j^i]^T$ of unit $j \in \mathcal{S}$, the SoC $S_{j,t}(x_j)$ of unit j at the end of time step t is

$$S_{j,t}(x_j) := \rho_j^t \bar{S}_{j,0} + \sum_{l=1}^{t} \rho_j^{t-l} x_{j,l}. \tag{3}$$

Let $I \subset \mathbb{R}$ be the closed interval containing the cumulative actions of the storage units, i.e., $I := [-\sum_j P_j^o, \sum_j P_j^i]$. Every $z \in I$ incurs a cost or generates a revenue, which depends on the time step $t \in \mathcal{T}$. We model both costs and revenues at time $t \in \mathcal{T}$ in a single cost function $C_t : I \longrightarrow \mathbb{R}$, with the convention that $C_t(z) > 0$ represents incurring a cost and $C_t(z) < 0$ generating a revenue. The cooperative aspect of the problem is incorporated here, in making C_t a function of the cumulative actions of the storage units, and not of the action of a specific unit.

Under positive electricity prices, $C_t(z) < 0$ if $z < 0$, i.e., the storage units generate a revenue $|C_t(z)|$ from discharging an amount $|z|$, and $C_t(z) > 0$ if $z > 0$, i.e., they incur a cost $C_t(z)$ for charging an amount z. This changes with negative prices, since the units incur a cost for discharging and get a revenue from charging, at least up to a certain amount. Negative prices present interesting challenges, but they blur the bigger picture. Therefore, for the remaining of this work, we assume that C_t is monotonically increasing in I for every $t \in \mathcal{T}$, and show in Appendix EC.2 how to relax it. This assumption represents the canonical situation under positive prices, where the units incur a higher cost to charge more and discharging more generates a higher revenue.

We assume that $C_t(0) = 0$ and C_t is convex for every $t \in \mathcal{T}$. The rationale behind the first assumption is that doing nothing does not incur a cost nor generate a revenue. The convexity assumption models the price-maker assumption by making the price per unit of electricity increase (resp. decrease) if we charge (resp. discharge) more. In the literature, price-makers are usually modelled using residual demand curves (de la Torre et al. (2002), Sousa et al. (2014), Shafiee et al. (2016), Barbry et al. (2019)). They represent how the market-clearing price changes by the actions of the price-maker. Our cost function C_t corresponds to the residual demand curve multiplied by the quantity charged or discharged and represents the total cost or revenue of an action made by the units, including cost increases or revenue decreases caused by the price-maker assumption.

Residual demand curves are generally modelled as monotonically increasing stepwise functions (de la Torre et al. (2002), Shafiee et al. (2016), Barbry et al. (2019)), where each jump represents ramping up/down or switching on/off a generator. Following this approach, the cost function obtained by multiplying a monotonically increasing stepwise function by the quantity charged or discharged is a piecewise linear cost function, where the slope of every piece is given by the value of the residual demand curve at those points. This results in a discontinuous, non-convex cost function, but the slopes of its pieces increase as we move to the next piece on the right. Therefore, they can be well approximated by convex functions, which justifies our convexity assumption.

We assume that prices are deterministic and that the storage units know how their actions affect electricity prices. Therefore, at time $t = 0$, the storage units have complete foresight of the cost functions C_t for all $t \in \mathcal{T}$. This assumption is quite strong, but our approach identifies a *forecast horizon* at every time step $t \in \mathcal{T}$, which is a future time beyond which it is not necessary to look to determine the action at time t. They are analogous to the forecast horizons obtained by Cruise et al. (2019), and set a bound on how much future price information is needed.

Price-maker storage units have very small leakage losses. Fuchs et al. (2015) claim that leakage losses in pumped-storage hydro plants are between $0.005 - 0.02\%$ per day. Therefore, we assume that all storage units have the same leakage loss factor ρ, since such small differences should have very limited effects. Without loss of generality, we assume that round-trip efficiencies take effect right before selling electricity. Therefore, all quantities are expressed in terms of energy charged. To model round-trip efficiencies, we introduce functions $h_j : [-P_j^o, P_j^i] \to \mathbb{R}$ defined by

$$h_j(x) := \begin{cases} \eta_j x & \text{if} \quad x < 0 \\ x & \text{if} \quad x \geq 0 \end{cases} \tag{4}$$

for all $j \in \mathcal{S}$. Note that h_j is convex for every $j \in \mathcal{S}$, since $\eta_j \in [0, 1]$. By considering the functions

$$C_t\Big(\sum_{j \in \mathcal{S}} h_j(x_{j,t})\Big), \tag{5}$$

the units pay for the energy lost due to round-trip inefficiencies, but only get rewarded by the energy they output. We note that (5) is convex if C_t is convex and monotonically increasing.

The mathematical formulation of the N-Units Problem is:

Sets:	
$\mathcal{S}$	set of N electric energy storage units
$\mathcal{T}$	set of T time steps
Parameters:	
E_j	energy capacity of unit $j \in \mathcal{S}$
P_j^i	power input rate of unit $j \in \mathcal{S}$
P_j^o	power output rate of unit $j \in \mathcal{S}$
η_j	round-trip efficiency of unit $j \in \mathcal{S}$
ρ	leakage of every unit $j \in \mathcal{S}$
$\bar{S}_{j,0}$	initial SoC of unit $j \in \mathcal{S}$
$\bar{S}_{j,T}$	final SoC of unit $j \in \mathcal{S}$
Decision variables:	
$x_{j,t}$	action taken by unit $j \in \mathcal{S}$ at time $t \in \mathcal{T}$.

$$\underset{x\in\mathcal{M}_{N\times T}(\mathbb{R})}{\text{Minimize}} \quad \sum_{t\in\mathcal{T}} C_t\Big(\sum_{j\in\mathcal{S}} h_j(x_{j,t})\Big) \tag{6}$$

$$\text{subject to} \quad -P_j^o \le x_{j,t} \le P_j^i \qquad \forall j\in\mathcal{S},\ \ \forall t\in\mathcal{T} \tag{7}$$

$$0 \le \rho^t \bar{S}_{j,0} + \sum_{l=1}^{t} \rho^{t-l} x_{j,l} \le E_j \qquad \forall j\in\mathcal{S},\ \ \forall t\in\mathcal{T} \tag{8}$$

$$\rho^T \bar{S}_{j,0} + \sum_{l=1}^{T} \rho^{T-l} x_{j,l} = \bar{S}_{j,T} \qquad \forall j\in\mathcal{S}. \tag{9}$$

Constraints (7) and (8) are the power rate and capacity constraints of the storage units, respectively. We fix the final SoCs with constraints (9), since otherwise the optimization would favour strategies with empty final SoCs. Note that constraints (8) and (9) involve the SoCs $S_{j,t}(x_j)$ of unit $j\in\mathcal{S}$ at time $t\in\mathcal{T}$ defined in (3). We assume that a feasible solution of the N-Units Problem exists. In practice, this can be easily checked by trying to attain $\bar{S}_{j,T}$ from $\bar{S}_{j,0}$ after T time steps, taking into account power rate constraints and leakage losses.

3 The Aggregated Unit Relaxation (AUR)

In this section, we introduce a relaxation of the N-Units Problem, the *Aggregated Unit Relaxation* (AUR). It consists of aggregating the N storage units into a single unit with the sum of energy capacities and power rates. In other words, we replace (7), (8) and (9) with the constraints obtained by adding them over $j\in\mathcal{S}$. These new constraints depend only on the sum of $x_{j,t}$ over $j\in\mathcal{S}$, but not explicitly on $x_{j,t}$. This leads to a relaxation of the N-Units Problem, that we want to interpret as a 1-Unit Problem. Therefore, instead of N actions, only one action can be taken at time $t\in\mathcal{T}$. For every $t\in\mathcal{T}$, we define $z\in\mathbb{R}^T$ by

$$z_t := \sum_{j\in\mathcal{S}} x_{j,t} \tag{10}$$

to represent the action taken by the aggregated unit at $t\in\mathcal{T}$. Furthermore, we define the round-trip efficiency η of the aggregated unit by

$$\eta := \max\{\eta_1,\dots,\eta_N\}. \tag{11}$$

As we did in (4), we introduce a function $h: I\to\mathbb{R}$, defined by

$$h(z) := \begin{cases} \eta z & \text{if} \quad z<0 \\ z & \text{if} \quad z\ge 0 \end{cases}, \tag{12}$$

to model round-trip efficiency. The Aggregated Unit Relaxation (AUR) is:

$$\underset{z\in\mathbb{R}^T}{\text{Minimize}} \quad \sum_{t\in\mathcal{T}} C_t(h(z_t)) \tag{13}$$

$$\text{subject to} \quad -\sum_{j\in\mathcal{S}} P_j^o \le z_t \le \sum_{j\in\mathcal{S}} P_j^i \qquad \forall t\in\mathcal{T} \tag{14}$$

$$0 \leq \rho^t \sum_{j \in \mathcal{S}} \bar{S}_{j,0} + \sum_{l=1}^{t} \rho^{t-l} z_l \leq \sum_{j \in \mathcal{S}} E_j \qquad \forall t \in \mathcal{T} \tag{15}$$

$$\rho^T \sum_{j \in \mathcal{S}} \bar{S}_{j,0} + \sum_{l=1}^{T} \rho^{T-l} z_l = \sum_{j \in \mathcal{S}} \bar{S}_{j,T}. \tag{16}$$

Proposition 1. *The Aggregated Unit Relaxation is a relaxation of the N-Units Problem.*

Proof. Proof. See Appendix EC.1 in the electronic companion document. □

Definition (11) is key to make Proposition 1 work. It is sometimes useful to undo the substitution (10) because the feasible sets of the AUR with the substitution undone and the N-Units Problem become comparable. In particular, we do so in the proof of Proposition 1.

4 The Decomposition by Unit and Time

We propose solving the N-Units Problem using a decomposition by unit and time. The first step is a decomposition by unit, which we present in Section 4.1. It results in $N+1$ independent 1-Unit Subproblems and an iterative method to solve them. See Appendix EC.5 for the explicit formulations of the subproblems. In Section 4.2, we take advantage of the time decomposition from Cruise et al. (2019) to solve 1-Unit Problems, and obtain a time decomposition for the N-Units Problem. A step by step description of our solution method in algorithm format can be found in Appendix EC.3.

4.1 Decomposition by Unit

We begin by decomposing the N-Units Problem by unit using an *Augmented Lagrangian Relaxation* (ALR) combined with the ADMM (Boyd et al. (2011)). The key observation that makes this approach possible is that every constraint of the N-Units Problem in (7), (8) and (9) only involves decision variables $x_{j,t}$ of a single unit $j \in \mathcal{S}$. Note that the objective function (6) is, in general, not separable. This introduces a difficulty to the decomposition process, but we describe below how to overcome it.

As was observed by Anjos et al. (2020), there are two cases in which solving the N-Units Problem reduces to solving 1-Unit Problems. First, if the storage units are price-takers, the objective function (6) becomes piecewise linear and is therefore separable. This reduces the N-Units Problem to N independent 1-Unit Problems. Second, if the storage units satisfy

$$\frac{E_j}{P_j^o} = \frac{E_k}{P_k^o}, \quad \frac{E_j}{P_j^i} = \frac{E_k}{P_k^i}, \tag{17}$$

and $\eta_j = \eta_k$ for every $j, k \in \mathcal{S}$, the N-Units Problem and the AUR are equivalent. Indeed, consider $\beta_1, \dots, \beta_N \in [0, 1]$ defined by the relations

$$E_j = \beta_j \sum_{k \in \mathcal{S}} E_k \tag{18}$$

for every $j \in \mathcal{S}$. For every feasible solution $z \in \mathbb{R}^T$ of AUR, $x \in \mathcal{M}_{N \times T}(\mathbb{R})$ defined by

$$x_{j,t} = \beta_j z_t \tag{19}$$

for every $j \in \mathcal{S}$ and $t \in \mathcal{T}$, is a feasible solution of the N-Units Problem and the values of the corresponding objective functions coincide. Combining this with Proposition 1, we conclude that both problems are equivalent. Therefore, under these assumptions, solving the N-Units Problem reduces to solving the AUR, which is a 1-Unit Problem, and applying (19).

For the general case, we introduce variables $z_t \in \mathbb{R}$ defined by the relation

$$h(z_t) = \sum_{j \in \mathcal{S}} h_j(x_{j,t}) \tag{20}$$

for every $t \in \mathcal{T}$, so that the objective function (6) can be replaced by

$$\sum_{t \in \mathcal{T}} C_t(h(z_t)). \tag{21}$$

Note that (21) is separable because it is a function of $z \in \mathbb{R}^T$ only. Further to replacing the objective function (6) by (21) and adding constraints (20) to the N-Units Problem, we add the redundant power and capacity constraints (14), (15) and (16) for the aggregated unit, because they will tighten one of the subproblems. This results in Problem P:

$$\begin{aligned} \underset{x,z}{\text{Minimize}} \quad & \sum_{t \in \mathcal{T}} C_t(h(z_t)) && (22) \\ \text{subject to} \quad & (7), (8), (9), (14), (15), (16), (20). \end{aligned}$$

Note that Problem P is equivalent to the N-Units Problem. Problem P would be decomposable by unit if constraints (20) were not present, which makes them linking constraints. Therefore, we can consider its Augmented Lagrangian Relaxation (ALR):

$$\underset{x,z}{\text{Minimize}} \quad \sum_{t \in \mathcal{T}} \Big[C_t(h(z_t)) + \nu_t \Big(\sum_{j \in \mathcal{S}} h_j(x_{j,t}) - h(z_t) \Big) + \frac{\gamma}{2} \Big(\sum_{j \in \mathcal{S}} h_j(x_{j,t}) - h(z_t) \Big)^2 \Big] \tag{23}$$

$$\text{subject to} \quad (7), (8), (9), (14), (15), (16) \tag{24}$$

for every $\nu = (\nu_1, \dots, \nu_T) \in \mathbb{R}^T$. Note that we introduce a quadratic penalty term in the objective function with a parameter $\gamma > 0$. Thus, ALR is a convex relaxation of Problem P for every $\nu \in \mathbb{R}^T$.

Let g be the dual function of ALR, i.e., given $\nu \in \mathbb{R}^T$, $g(\nu)$ is the optimal value of ALR with ν fixed. Because ALR is a relaxation of Problem P, $g(\nu)$ is a lower bound of the optimal value of Problem P for all $\nu \in \mathbb{R}^T$. We next show that strong duality holds.

Proposition 2. *Maximizing $g(\nu)$ over $\nu \in \mathbb{R}^T$ and the N-Units Problem have equal optimal values.*

Proof. Proof. See Appendix EC.4 in the electronic companion document. □

The equivalence from Proposition 2 gives a max-min problem that can be solved using ADMM. We fix $\nu \in \mathbb{R}^T$ and observe that the problem obtained is the ALR, since for fixed ν there is no maximization over $\nu \in \mathbb{R}^T$ to be done. This comes at the expense of iterating over $\nu \in \mathbb{R}^T$. The ALR is not yet decomposable by unit because the objective function (23) is not separable due to the quadratic terms. To make it separable, we fix all but one of the $N + 1$ vectors of decision variables $z \in \mathbb{R}^T$ or $x_j \in \mathbb{R}^T$ for $j \in \mathcal{S}$. This results in $N+1$ subproblems, one for each vector of decision variables that is not fixed, which are solved iteratively until the linking constraint (20) is violated only up to a certain tolerance. Let Problem Q_j be the problem that has decision variables $x_j \in \mathbb{R}^T$ for $j \in \mathcal{S}$ and Problem Q_0 be the problem that has $z \in \mathbb{R}^T$ as decision variables. See Appendix EC.5 for the mathematical formulation of the subproblems.

Every ADMM iteration consists on solving the $N + 1$ subproblems to update the $N + 1$ vectors of decision variables $z \in \mathbb{R}^T$ or $x_j \in \mathbb{R}^T$ for $j \in \mathcal{S}$, and a ν-update step. We introduce superscripts to indicate which iteration of the ADMM we are in. Let $x_j^{(v)}$, $z^{(v)}$ and $\nu^{(v)}$ denote the values of x_j, z and ν obtained after the ADMM iteration $(v) \in \mathbb{N}$. We initialize all of them to 0.

Subproblems are solved sequentially and not in parallel, to make use of the most updated information. At every ADMM iteration, the vectors of decision variables are updated in the following order:

$$x_1, x_2, \ldots, x_N, z. \tag{25}$$

This is followed by the ν-update using the formula

$$\nu_t^{(v)} = \nu_t^{(v-1)} + \gamma\Big(\sum_{j\in\mathcal{S}} h_j(x_{j,t}^{(v)}) - (z_t^{(v)})\Big) \qquad \forall t \in \mathcal{T}. \tag{26}$$

By the convention introduced in (1), the subproblems of *slower* units are solved first. This choice is made to allow a further decomposition in time, which we describe in Section 4.2 below. In Appendix EC.5, we present the mathematical formulation of the subproblems taking into account the order defined in (25).

4.2 Decomposition in Time

The unit decomposition from Section 4.1 results in $N+1$ 1-Unit Subproblems, which can have a large final time T. Thus, a natural question is whether they need to be solved until T at every ADMM iteration, which can be time consuming. To avoid doing that, we take advantage of the time decomposition introduced in the algorithm by Cruise et al. (2019) to solve 1-Unit Problems.

To solve a 1-Unit Problem, the algorithm of Cruise et al. (2019) first identifies times $1 \leq \tau_1 \leq \bar{\tau}_1 \leq T$ and a segment of optimal actions $(y_1, \ldots, y_{\tau_1})$ of the unit up to time τ_1 such that these actions do not depend on cost functions C_t for $t \geq \bar{\tau}_1$. The times τ_1 and $\bar{\tau}_1$ are defined as the first *decision* and *forecast horizons*. The algorithm then repeats this process, identifying successive decision horizons, forecast horizons and segments of optimal actions, until the final time T. The horizons are not set in advance, but obtained on the fly, and, crucially, this decomposition does not compromise optimality (Cruise et al. (2019)).

The challenge our approach presents is that we obtain $N+1$ different time decompositions, one for every subproblem. We want to have a single time decomposition for the whole solution method. The key idea is that the length of the decision and forecast horizons depend on the E/P^i and E/P^o ratios of the units (Cruise et al. (2019), Anjos et al. (2020)): The larger the ratios are, the longer its decision and forecast horizons are. Therefore, the unit $j \in \mathcal{S}$ with the largest E_j/P_j^i and E_j/P_j^o ratios should determine the decision and forecast horizons of the solution method.

Since (1) holds, such a unit exists, and we made the convention to call it unit 1. This justifies the order we set in (25) to solve the subproblems in every ADMM iteration, since Problem Q_1 should give the largest decision and forecast horizons amongst the subproblems. We solve Problem Q_1 first, and set the forecast horizon of Problem Q_1 as the time until which we solve the remaining subproblems in the current ADMM iteration. Once the ADMM converges, we take the solutions only up to the decision horizon of Problem Q_1. We call these times *ADMM forecast horizon* and *ADMM decision horizon*, respectively. The solution method then restarts from the obtained ADMM decision horizon. Note that price information beyond the ADMM forecast horizon was not needed at this point.

In real-world applications, generators are usually scheduled over fixed-length time periods, which we call scheduling periods. For instance, in day-ahead markets, a scheduling period typically consists of every hour (or half-hour) of the next day. For energy storage, a crucial variable that carries over scheduling periods is the state of charge. Our approach can provide valuable information about the desired state of charge at the end of the scheduling period using the ADMM decision and ADMM forecast horizons. In day-ahead markets, it allows storage units to take advantage of both interday and intraday price differences.

To do so, we apply our solution method until the current ADMM decision horizon is beyond the end of the scheduling period. This determines the states of charge at the end of the scheduling period taking into account the forecasts beyond that time. The optimization then restarts from the end of the scheduling period once we get closer to it and more accurate price forecasts become available. Note that the final time T from the N-Units Problem defined in Section 2 plays a secondary role when solving the problem in practise. It should be set sufficiently into the future so that it does not interfere with the current scheduling period.

5 Lagrangian Sufficiency Conditions for Optimality

In this section, we present the Lagrangian Sufficiency Theorem for the N-Units Problem.

Theorem 1 (Lagrangian Sufficiency Theorem). *If there exist* $x^*, \mu^* \in \mathcal{M}_{N\times T}(\mathbb{R})$ *such that*

(i) x^* *is a feasible solution of the* N*-Units Problem.*

(ii) For every $t \in \mathcal{T}$, $x_t^* = (x_{1,t}^*, \ldots, x_{N,t}^*)$ *minimizes*

$$C_t\Big(\sum_{j\in\mathcal{S}} h_j(x_{j,t})\Big) - \sum_{j\in\mathcal{S}} \mu_{j,t}^* x_{j,t} \tag{27}$$

amongst $x_t = (x_{1,t}, \ldots, x_{N,t}) \in \mathbb{R}^N$ *such that* $x_{j,t} \in [-P_j^o, P_j^i]$ *for every* $j \in \mathcal{S}$.

(iii) For every $j \in \mathcal{S}$, $x_j^* = (x_{j,1}^*, \ldots, x_{j,T}^*) \in \mathbb{R}^T$ *satisfy the complementary slackness conditions*

$$\begin{cases} \rho\mu_{j,t+1}^* = \mu_{j,t}^* & \text{if} \quad 0 < S_{j,t}(x_j^*) < E_j \\ \rho\mu_{j,t+1}^* \leq \mu_{j,t}^* & \text{if} \quad S_{j,t}(x_j^*) = 0 \\ \rho\mu_{j,t+1}^* \geq \mu_{j,t}^* & \text{if} \quad S_{j,t}(x_j^*) = E_j \end{cases} \tag{28}$$

for all $t \in \mathcal{T} \setminus \{T\}$. *Here,* $S_{j,t}(x_j^*)$ *denotes the state of charge of unit* $j \in \mathcal{S}$ *at time* $t \in \mathcal{T}$ *if unit* j *takes actions* $x_j^* \in \mathbb{R}^T$. *It was defined in* (3).

Then, x^* *is an optimal solution of the* N*-Units Problem.*

We omit the proof of this theorem because it is standard and very similar to the proofs of the corresponding theorems for the 1- and 2-Units Problem by Cruise et al. (2019) and Anjos et al. (2020), respectively. Nevertheless, Theorem 1 is useful to analyse the solutions obtained by our method. Observe that $\mu_{j,t}^*$ is a cumulative Lagrange multiplier of capacity constraints and represents a notional reference value per unit volume in unit $j \in \mathcal{S}$ at time t, i.e., the rate at which the residual value of the storage unit, optimally operated up to time T, increases with respect to increasing the state of charge of that unit at time t (Cruise et al. (2019)). Therefore, if the market-clearing price, before the actions of the storage units, is smaller (resp. larger) than $\mu_{j,t}^*$, unit $j \in \mathcal{S}$ should be in the buying (resp. selling) regime at time t.

6 Examples and Computational Performance

In this section, we present and analyse the results obtained from applying our solution method to numerical examples. We begin with a toy example in Section 6.1 and continue with a more realistic example in Section 6.2. Section 6.3 presents a computational performance study. All our numerical experiments were carried out on a personal laptop, an Intel Core i5-8350U CPU at 1.70 GHz and 16GB RAM. The code was written in Microsoft Visual C++ 2019.

6.1 Toy Example

Consider a set $\mathcal{S}$ of two storage units defined by $(E_1, P_1^i, P_1^o, \eta_1, \rho) = (3, 1, 1, 1, 1)$ and $(E_2, P_2^i, P_2^o, \eta_2, \rho) = (1, 1, 1, 1, 1)$. For simplicity, there is no leakage nor round-trip efficiency losses. Consider $T = 4$ time steps with electricity market-clearing prices given by $p_1 = 20$, $p_2 = 25$, $p_3 = 40$, $p_4 = 45$. Assume that the storage units are price-takers, i.e., for every $t \in \mathcal{T}$, the cost functions C_t are given by

$$C_t(z) = p_t z. \tag{29}$$

Lastly, assume that the initial and final SoCs of the units are $\bar{S}_{1,0} = \bar{S}_{1,4} = 0 = \bar{S}_{2,0} = \bar{S}_{2,4}$.

Since both units are price-takers, the optimal set of actions of unit 1 is

$$x_1 = (+1, +1, -1, -1), \tag{30}$$

and the optimal set of actions of unit 2 is

$$x_2 = (+1, +0, +0, -1). \tag{31}$$

In other words, the units try to charge as much as possible in the first two time steps to discharge in the last two, taking advantage of higher electricity prices in the last two time steps. Since $E_2 = 1$, unit 2 charges at time $t = 1$ and discharges at time $t = 4$ at maximum rates, remaining idle in between, because those times present the largest price difference.

The AUR for this problem is the 1-Unit Problem with $(E, P^i, P^o, \eta, \rho) = (4, 2, 2, 1, 1)$. Since the unit is a price-taker, the optimal solution of AUR is

$$z = (+2, +2, -2, -2). \tag{32}$$

Observe that (32) is not equivalent to the optimal solution of the original 2-Units Problem given by (30) and (31). Indeed, the latter charges less at $t = 2$ and discharges less at $t = 3$ because unit 2 is full at the end of the first time step. This highlights the fact that the AUR is, in general, not a tight relaxation of the N-Units Problem.

We summarize the results obtained by applying our solution method to this problem in Figure 1. On the bottom plot, we can see the stacked SoC of units 1 and 2, and the SoC of the aggregated unit, obtained by solving AUR. In this toy example, we already know that the solution obtained by our method is optimal, because it coincides with (30) and (31). Nevertheless, we can prove optimality using Theorem 1. The challenge is that it is not obvious to identify the cumulative Lagrange multipliers $\mu^*_{j,t}$ required by the theorem.

Since we solve all subproblems using the algorithm by Cruise et al. (2019), we get cumulative Lagrange multipliers $\bar{\mu}_{j,t}$ for every $t \in \mathcal{T}$ from solving Subproblem Q_j for every $j \in \mathcal{S} \cup \{0\}$. We claim that $\mu^*_{j,t}$ is given, for every $j \in \mathcal{S}$ and $t \in \mathcal{T}$, by

$$\bar{\mu}_{0,t} + \bar{\mu}_{j,t}. \tag{33}$$

To prove this claim, it suffices to check that (33) and the SoCs obtained by our solution method satisfy the assumptions of Theorem 1. On the top plot of Figure 1, we show p_t, (33) for every $j \in \mathcal{S}$ and the cumulative Lagrange multipliers μ^{AUR}_t of AUR for every $t \in \mathcal{T}$.

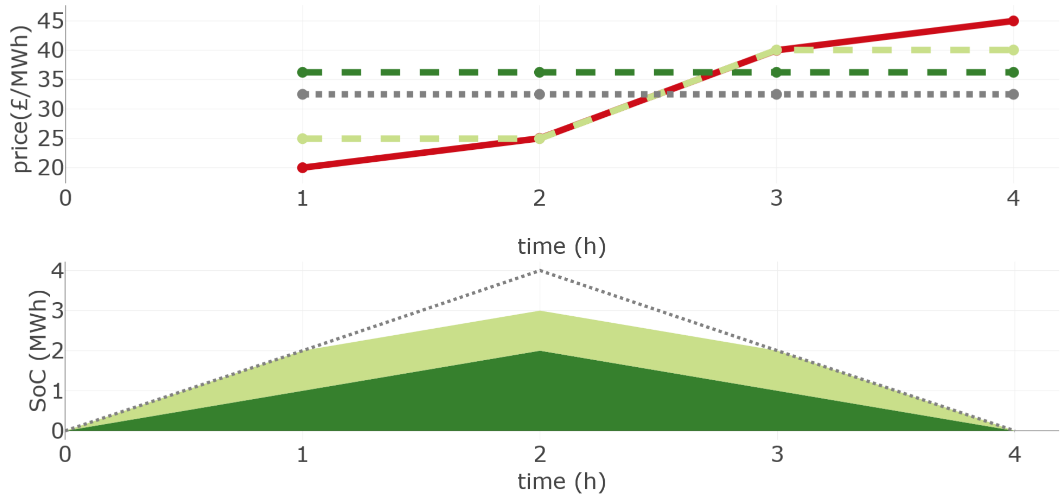

Figure 1: Top: market-clearing prices p_t (£/MWh) [solid red line], $\bar{\mu}_{0,t}+\bar{\mu}_{1,t}$ [dashed dark green line], $\bar{\mu}_{0,t}+\bar{\mu}_{2,t}$ [dashed light green line] and μ_t^{AUR} [dotted grey line]. Bottom: stacked SoC (MWh) of unit 1 [dark green] and unit 2 [light green]. SoC (MWh) of aggregated unit [dotted grey line].

Conditions (i) and (iii) of Theorem 1 can be easily verified by looking at Figure 1. To check condition (ii), observe that $p_1 \le p_2 \le \bar{\mu}_{0,t}+\bar{\mu}_{1,t} \le p_3 \le p_4$ for all $t \in \mathcal{T}$. Therefore, since the units are price-takers, the first component of the minimizers of (27) is P_1^i if $t = 1, 2$ and $-P_1^o$ if $t = 3, 4$. This is precisely (30). A similar argument can be made for unit 2, observing that at times $t = 2$ and $t = 3$, where $x_{2,t} = 0$, we have $\bar{\mu}_{0,t}+\bar{\mu}_{2,t} = p_t$.

Since $\mu_{j,t}^*$ represents a notional value of the energy stored in unit $j \in \mathcal{S}$ at time $t \in \mathcal{T}$ (see Section 5), $\mu_{j,t}^* = \bar{\mu}_{0,t}+\bar{\mu}_{j,t}$ has an interesting interpretation. $\bar{\mu}_{0,t}$ represents the value of the energy stored across the units, and $\bar{\mu}_{j,t}$ represents a unit-dependent adjustment on that value for unit $j \in \mathcal{S}$, highlighting the fact that, sometimes, energy is more valuable stored in one unit than in another. This is supported by the fact that, in this numerical experiment, $\bar{\mu}_{0,t}$ is of the order of the p_t's while $\bar{\mu}_{j,t}$ is an order of magnitude smaller.

6.2 A Realistic Example in GB

In this section we apply our solution method to a realistic representation of the Great Britain (GB) power system. Consider the set $\mathcal{S}$ of 4 electric energy storage units similar to the 4 pumped-storage power stations in GB: Cruachan, Foyers, Ffestiniog and Dinorwig. The energy capacities and power rates of these units are given in the reports by the Renewable Energy Association (2016) and Scottish Renewables (2016). We note that the power input and output rates coincide for every storage unit $j \in \mathcal{S}$, which we denote by P_j. In particular, assumption (1) is automatically satisfied.

Pumped-storage power stations have round-trip efficiencies around 80% (Renewable Energy Association (2016)). Therefore, due to the lack of unit-specific data, we assume that $\eta_j = 0.8$ for every $j \in \mathcal{S}$. Furthermore, since leakage losses are very small, around $0.005 - 0.02\%$ per day (Fuchs et al. (2015)), we assume that there are none, i.e., $\rho = 1$. We summarize the parameters of this problem in Table 2. Notice that we have classified the units by decreasing E/P-ratio, since this value is the main driver of our solution method.

No.	Power Station	E_j (GWh)	P_j (GW)	E_j/P_j (h)	η_j	ρ
1	Cruachan	10	0.44	22.73	0.8	1
2	Foyers	6.3	0.3	21	0.8	1
3	Ffestiniog	2	0.36	5.56	0.8	1
4	Dinorwig	9	1.8	5	0.8	1
0	Aggregated Unit	27.3	2.9	9.41	0.8	1

Table 2: Energy capacities E_j, power rates P_j, E_j/P_j-ratios, round-trip efficiencies η_j and leakage ρ of the selected storage units $j \in \mathcal{S}$.

We consider the set $\mathcal{T}$ of all hours in April 2021. We build cost functions C_t from N2EX day-ahead hourly market-clearing electricity prices p_t (NordPool (2021)) for $t \in \mathcal{T}$ and a market impact factor $\lambda \geq 0$ to be determined later. We assume that the residual demand curve of the storage units at $t \in \mathcal{T}$ is

$$p_t + \lambda p_t z_t, \tag{34}$$

where z_t is the cumulative action of the units at time t. The second term in (34) models the price-maker assumption. Its linear dependence on z_t corresponds to the assumption that the storage units are large enough to be price-makers, but relatively small compared to the size of the system, so that their effect on market-clearing prices is modest. This is consistent with existing energy economics literature (Sioshansi (2010), Sioshansi (2014)). Furthermore, its dependence on p_t reflects the fact that higher market-clearing prices p_t make markets more price-responsive (Barbry et al. (2019)), since the capacity limits of the system are closer. We set the market impact factor to $\lambda = 1/20$, since in GB on-peak prices are generally the double of off-peak prices, which corresponds to a demand difference of approximately 20 GW. This parameter choice corresponds to working in GW, which we do, and is equivalent to the choices made by Cruise et al. (2019) and Anjos et al. (2020).

This results in quadratic cost functions given by

$$C_t(z_t) = (p_t + \lambda p_t z_t) z_t \tag{35}$$

for every $t \in \mathcal{T}$. We apply our solution method for the whole month of April 2021, but focus the analysis in the third week of the month, starting on Monday, 12 April 2021, to have realistic initial and final levels of charge. In Figure 2, we plotted the N2EX day-ahead hourly electricity market-clearing prices p_t during the third week of April 2021. We observe a strong daily cycle, with morning and evening peaks, and two price spikes around $t = 44$ and $t = 68$. Furthermore, the observation motivating our choice of market impact factor λ seems to hold.

In Figure 3, we plotted the stacked SoCs of the 4 storage units obtained by our solution method, and the SoC of the aggregated unit that corresponds to solving

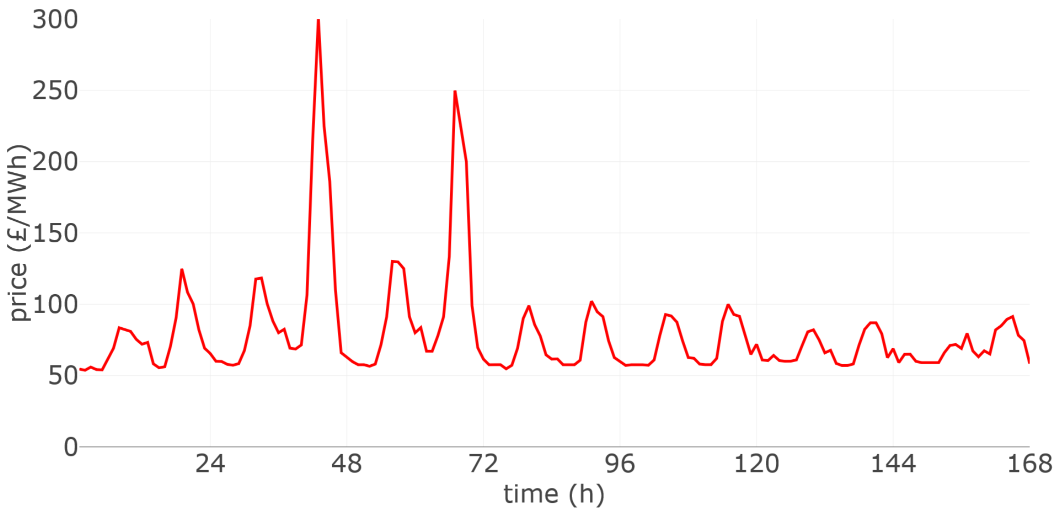

Figure 2: N2EX day-ahead hourly market-clearing prices in the third week of April 2021 (NordPool (2021)).

AUR. We observe that units follow a strong daily cycle, which is expected, since market-clearing prices also exhibit this feature. Note that units 3 and 4 tend to get empty and full on a daily basis, while units 1 and 2 do not, because the latter have E/P-ratios above 20 hours. This does not allow them to fully charge and discharge on a daily basis. Furthermore, all units take advantage of the price spikes, as can be seen in the strong discharge phase around times $t = 44$ and $t = 68$.

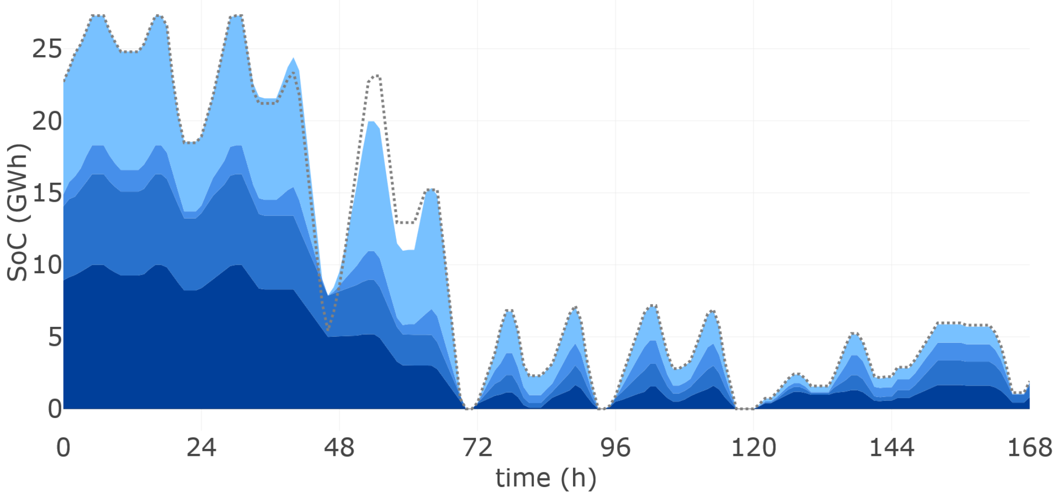

Figure 3: Stacked SoC of unit 1 [very dark blue], unit 2 [dark blue], unit 3 [light blue] and unit 4 [very light blue]. SoC of aggregated unit [dotted grey line].

Figure 3 also supports a comparison between the 4-Units Problem at hand and its relaxation, the AUR. We observe that the cumulative SoC values of the 4-Units Problem mostly coincide or are very close to those of the AUR. There are, however, important differences. Around time $t = 44$, the SoC of the aggregated unit is significantly lower than the cumulative SoC of the 4 units. The reason behind it is that the price spike around that time consists of 6 consecutive hours of prices above 100 £/MWh. Since the aggregated unit has an E/P-ratio greater than 6 hours but units 3 and 4 do not, only the former can take full advantage of this price spike.

A similar behaviour can be observed between the two price spikes around $t = 44$ and $t = 68$. Anticipating the coming price spike at $t = 68$, all units try to charge as much as possible after $t = 48$, when prices are lower. As soon as units 3 and 4 get full at $t = 53$, the four units charge at a smaller cumulative rate than the aggregated unit, which causes the difference in cumulative SoCs after this time. The four units then catch up before the second price spike, at the expense of paying higher prices. Therefore, Figure 3 shows that the AUR can be a relatively tight relaxation of the N-Units Problem. Nevertheless, as observed in this figure and in Section 6.1, there are times in which the relaxation produces infeasible results. By contrast, our solution method is always able to provide a feasible solution for the N-Units Problem.

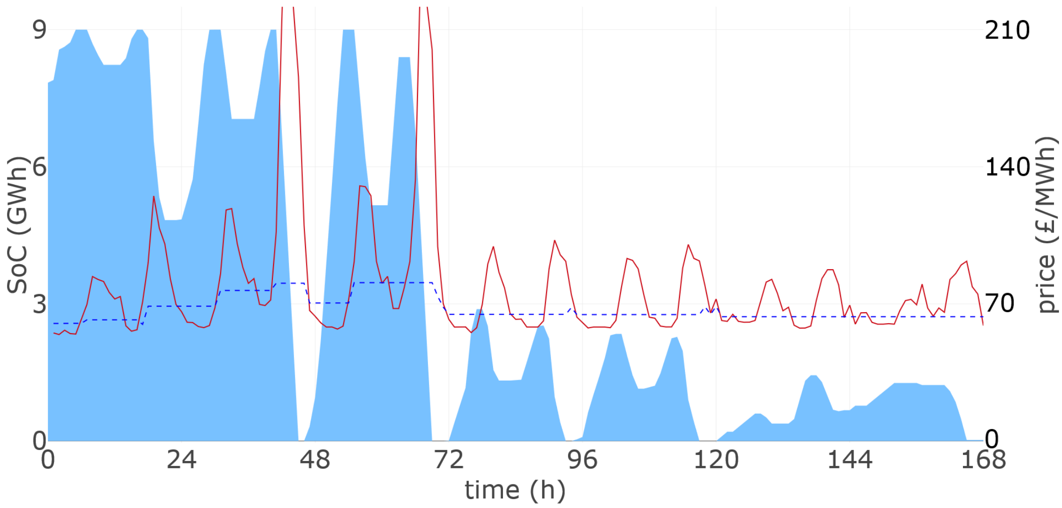

Figure 4: On the left axis: States of charge of unit 4 [very light blue]. On the right axis: electricity market-clearing prices p_t [solid red line] and $\bar{\mu}_{0,t} + \bar{\mu}_{4,t}$ [dashed blue line].

Figure 4 explores whether the solution obtained by our method satisfies the Lagrangian Sufficiency Conditions (Theorem 1). In particular, we want to check if the price signals $\mu^*_{j,t}$ from that theorem are $\bar{\mu}_{0,t} + \bar{\mu}_{j,t}$ for every $t \in \mathcal{T}$ and $j \in \mathcal{S}$, as in the toy example from Section 6.1, see (33). In Figure 4, we plotted the SoC of unit 4, market-clearing prices p_t and the values $\bar{\mu}_{0,t} + \bar{\mu}_{4,t}$ for every $t \in \mathcal{T}$. We focus in unit 4, but we observed very similar behaviours for the other units.

Conditions (i) and (ii) of Theorem 1 can be easily checked. Condition (iii) is satisfied in most time steps, since the values of $\bar{\mu}_{0,t} + \bar{\mu}_{4,t}$ change only if the unit is empty or full, and they mostly change in the right direction: $\bar{\mu}_{0,t} + \bar{\mu}_{4,t}$ increases if the unit is full at time $t \in \mathcal{T}$ and decreases if the unit is empty at that time. There are a few exceptions, around $t = 20$, $t = 94$ and $t = 118$, where $\bar{\mu}_{0,t} + \bar{\mu}_{4,t}$ suffers a small perturbation for one time step. We do not consider them to be significant because they seldom happen and, when they do, the value of $\mu^*_{j,t}$ can be readjusted so that complementary slackness is satisfied. Therefore, in this larger example, we are also able to recover the price signals $\mu^*_{j,t}$ from Theorem 1.

6.3 Computational Performance

In this section we present a study of the computational performance of our solution method. In Figure 5, we report the computational times of an experiment consisting on solving N-Units Problems for $N = 1, \ldots, 5$. We classify the instances in 3 groups, depending on the largest E/P^i and E/P^o ratios of the units. The largest ratios are $5, 10$ and 20, and we pick the remaining units with ratios evenly distributed between 1 and the largest ratio in each group. We set round-trip efficiencies to $\eta_j = 0.8$ for all $j \in \mathcal{S}$ and market impact factor to $\lambda = 1/20$, as we did in Section 6.2. We report the time it took to solve all instances for the month of April 2021.

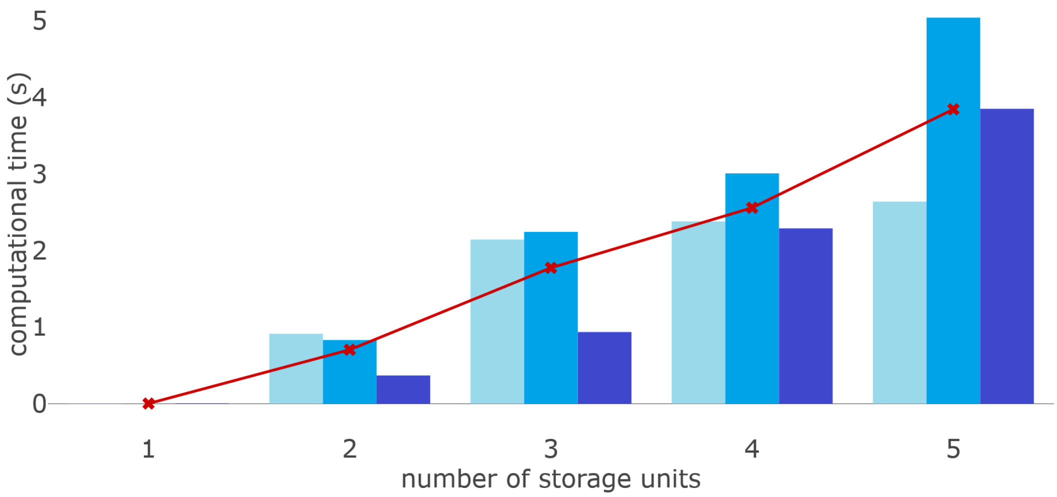

Figure 5: Computational time to solve N-Units Problems with April 2021 data for different number of storage units. Unit with largest E/P^i and E/P^o ratios fixed to 5 [light blue bar], 10 [blue bar] and 20 [dark blue bar]. Average of the previous in red crosses joined by a red line.

The results presented in Figure 5 show that all instances considered could be solved under 6 seconds. Furthermore, there seems to be a mild linear increase in computational time as we increase the number of storage units, which is a very good sign in terms of scalability of our solution method. This is not surprising, since adding one extra unit just means adding one subproblem, and these are generally quick to solve because they are 1-Unit Problems. Moreover, we could solve an instance with 50 storage units with the largest ratios equal to 50 in less than 50s. This gives further evidence of the linear relation between computational time and number of units. We do not observe significant differences in computational time across the three groups considered. Therefore, different energy to power ratios might result in different time decompositions, but this does not result in significant differences in computational time. What is important to obtain this computational performance is having a time decomposition.

Finally, we compare the solution method by Anjos et al. (2020) with our solution method. To do so, we solved the 2-Units Problem considered there with both methods. The data is from January 2020 (NordPool (2021)), and we record the computational time for the whole month. We assume that the storage units are fully efficient and a market impact factor of $\lambda = 1/20$. The solution method by Anjos et al. (2020) took 229.65 seconds solve, whereas ours took 1.68 seconds. This represents a two order of magnitude computational time reduction.

7 Conclusions

This paper presents a solution method for the optimal control of multiple price-maker electric energy storage units that cooperate to maximize profit from price arbitrage. It combines a decomposition by unit based on the Alternating Direction Method of Multipliers (ADMM) and an on the fly decomposition in time. Both decompositions improve computational performance. Furthermore, the decomposition in time limits the future price information needed, and, if the storage units need to be scheduled over a fixed-length time period (e.g., day-ahead), it provides valuable information on the desirable states of charge at the end of the scheduling period.

Our numerical experiments suggest impressive performance both in terms of accuracy and computational time. The gap between our problem of interest, the N-Units Problem, and a relaxation of it, the Aggregated Unit Relaxation, is small in all our test cases, and the reasons for the relaxation not being exact are identified. Furthermore, we are able to validate Lagrangian Sufficiency Conditions for optimality in most time steps, identifying different price signals (Lagrange multipliers) for every storage unit, and this determines their actions. Regarding computational performance, our numerical experiments indicate that computational time scales linearly with the number of storage units. Furthermore, it shows two orders of magnitude computational time improvements when compared to the solution method by Anjos et al. (2020).

This work also highlights the importance of the energy to power ratios of storage units. The behaviour of the storage units is driven by the interaction of electricity prices with these ratios. In future work, we would like to study the problem that arises by sending the number of storage units to infinity. We believe it can help analyse the effects of having a large number of small price-taker storage units connected to the power system, which might not affect electricity market-clearing prices individually, but collectively would.

Acknowledgements

The authors thank Fraser Daly, Chris Dent, Jean Lasserre, Seva Shneer and Stan Zachary for very helpful discussions and Jalal Kazempour for delivering a course that inspired part of this research. Albert Solà Vilalta was supported by The Maxwell Institute Graduate School in Analysis and its Applications, a Centre for Doctoral Training funded by the UK Engineering and Physical Sciences Research Council (grant EP/L016508/01), the Scottish Funding Council, Heriot-Watt University and the University of Edinburgh.

References

Andrés-Cerezo D, Fabra N (2020) Storing Power: Market Structure Matters. Working paper, Faculty of Economics, Cambridge.

Anjos MF, Cruise JR, Solà Vilalta A (2020) Control of two energy storage units with market impact: Lagrangian approach and horizons. *Proc. 2020 International Conference on Probabilistic Methods Applied to Power Systems* (PMAPS IEEE, Liege), 1–6.

Aust B, Horsch A (2020) Negative market prices on power exchanges: Evidence and policy implications from Germany. *The Electricity Journal* 33(3):106716.

Bajwa M, Cavicchi J (2017) Growing evidence of increased frequency of negative electricity prices in U.S. wholesale electricity markets. *IAEE Energy Forum Fourth Quarter* .

Barbour E, Wilson IG, Radcliffe J, Ding Y, Li Y (2016) A review of pumped hydro energy storage development in significant international electricity markets. *Renewable and Sustainable Energy Reviews* 61:421–432.

Barbry A, Anjos MF, Delage E, Schell KR (2019) Robust self-scheduling of a price-maker energy storage facility in the New York electricity market. *Energy Economics* 78:629 – 646.

Baslis CG, Bakirtzis AG (2011) Mid-term stochastic scheduling of a price-maker hydro producer with pumped storage. *IEEE Transactions on Power Systems* 26(4):1856–1865.

Biegel B, Andersen P, Pedersen TS, Nielsen KM, Stoustrup J, Hansen LH (2013) Electricity market optimization of heat pump portfolio. *Proc. 2013 IEEE International Conference on Control Applications* (CCA IEEE, Hyderabad), 294–301.

Boyd S, Parikh N, Chu E, Peleato B, Eckstein J (2011) Distributed optimization and statistical learning via the alternating direction method of multipliers. *Found. Trends Mach. Learn.* 3(1):1–122.

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

Conejo A, Arroyo J, Contreras J, Villamor F (2002) Self-scheduling of a hydro producer in a pool-based electricity market. *IEEE Transactions on Power Systems* 17(4):1265–1272.

Cruise JR, Flatley L, Gibbens RJ, Zachary S (2019) Control of energy storage with market impact: Lagrangian approach and horizons. *Operations Research* 67(1):1–9.

de la Torre S, Arroyo JM, Conejo AJ, Contreras J (2002) Price maker self-scheduling in a pool-based electricity market: a mixed-integer LP approach. *IEEE Transactions on Power Systems* 17(4):1037–1042.

De Vos K (2015) Negative wholesale electricity prices in the German, French and Belgian day-ahead, intra-day and real-time markets. *The Electricity Journal* 28(4):36–50.

Del-Rosario-Calaf G, Rocamora S, Corchero C (2016) Design of an algorithm to deliver demand side flexibility from aggregated resources to distribution system operators. *Proc. 2016 IEEE PES Innovative Smart Grid Technologies Conference Europe* (ISGT-Europe IEEE, Ljubljana), 1–6.

Dennis K (2015) Environmentally beneficial electrification: Electricity as the end-use option. *The Electricity Journal* 28(9):100–112.

El-Ghandour L, Johnson TC (2017) A methodology to assess the economic impact of power storage technologies. *Phil. Trans. R. Soc. A* 375(2100):20160303.

Fuchs G, Lunz B, Leuthold M, Sauer DU (2015) Chapter 7 - Overview of Nonelectrochemical Storage Technologies. Moseley PT, Garche J, eds. *Electrochemical Energy Storage for Renewable Sources and Grid Balancing* (Elsevier, Amsterdam), 89–102.

Graves F, Jenkin T, Murphy D (1999) Opportunities for electricity storage in deregulating markets. *The Electricity Journal* 12(8):46 – 56.

International Energy Agency (2018) World energy outlook 2018. Report, International Energy Agency, Paris.

International Energy Agency (2021) Net zero by 2050. A roadmap for the global energy sector. Report, International Energy Agency, Paris.

Kazempour SJ, Hosseinpour M, Moghaddam MP (2009) Self-scheduling of a joint hydro and pumped-storage plants in energy, spinning reserve and regulation markets. *Proc. 2009 IEEE Power & Energy Society General Meeting* (IEEE, Calgary), 1–8.

National Renewable Energy Laboratory (2010) The role of energy storage with renewable electricity generation. Technical report, Office of Energy Efficiency & Renewable Energy, U.S. Department of Energy, Washington, DC.

NordPool (2021) N2EX day ahead auction prices. Accessed May 18, https://www.nordpoolgroup.com/Market-data1/GB/Auction-prices/UK/Hourly/?view=table.

Renewable Energy Association (2016) Energy storage in the UK: An overview. Report, Renewable Energy Association, London.

Schill W, Kemfert C (2011) Modeling strategic electricity storage: The case of pumped hydro storage in Germany. *The Energy Journal* 32(3):59–87.

Scottish Renewables (2016) The benefits of pumped storage hydro to the UK. Report, Pumped Storage Hydro Working Group, Scottish Renewables, Glasgow.

Secomandi N (2010) Optimal commodity trading with a capacitated storage asset. *Management Science* 56(3):449–467.

Secomandi N (2015) Merchant commodity storage practice revisited. *Operations Research* 63(5):1131–1143.

Secomandi N, Lai G, Margot F, Scheller-Wolf A, Seppi DJ (2015) Merchant commodity storage and term-structure model error. *Manufacturing & Service Operations Management* 17(3):302–320.

Sethi S, Sorger G (1991) A theory of rolling horizon decision making. *Annals of Operations Research* 29(1-4):387–415.

Shafiee S, Zamani-Dehkordi P, Zareipour H, Knight AM (2016) Economic assessment of a price-maker energy storage facility in the Alberta electricity market. *Energy* 111:537–547.

Sioshansi R (2010) Welfare impacts of electricity storage and the implications of ownership structure. *The Energy Journal* 31(2):173–198.

Sioshansi R (2014) When energy storage reduces social welfare. *Energy Economics* 41:106–116.

Sioshansi R, Denholm P, Jenkin T, Weiss J (2009) Estimating the value of electricity storage in PJM: Arbitrage and some welfare effects. *Energy Economics* 31(2):269–277.

Sousa JA, Teixeira F, Faias S (2014) Impact of a price-maker pumped storage hydro unit on the integration of wind energy in power systems. *Energy* 69:3–11.

United Nations Framework Convention on Climate Change (2015) Paris agreement. International treaty, UNFCC, Paris.

Wogrin S, Tejada-Arango DA, Delikaraoglou S, Botterud A (2021) Show me the money! Profitability of energy storage systems in low-carbon power systems. *Proc. 2021 IEEE Madrid PowerTech* (IEEE, Madrid), 1–6.

Wu OQ, Wang DD, Qin Z (2012) Seasonal energy storage operations with limited flexibility: The price-adjusted rolling intrinsic policy. *Manufacturing & Service Operations Management* 14(3):455–471.

Zhang W, Lian J, Chang C, Kalsi K (2013) Aggregated modeling and control of air conditioning loads for demand response. *IEEE Transactions on Power Systems* 28(4):4655–4664.

Zhou Y, Scheller-Wolf A, Secomandi N, Smith S (2016) Electricity trading and negative prices: Storage vs. disposal. *Management Science* 62(3):880–898.

Electronic Companion for "ADMM-based Unit and Time Decomposition for Price Arbitrage by Cooperative Price-Maker Electricity Storage Units".

Proofs, Extension and Algorithm

In this electronic companion document we include complementary material. Appendix EC.1 presents the proof of Proposition 1 and Appendix EC.2 discusses how negative electricity prices can be incorporated in the model and solution method. Appendix EC.3 contains a step by step description of our solution method in algorithm format. Appendix EC.4 presents the proof of Proposition 2 and Appendix EC.5 contains the mathematical formulation of the subproblems from the decomposition by unit.

EC.1 Proof of Proposition 1

After undoing substitution (10) in AUR, it becomes clear that (14), (15) and (16) are obtained by adding the N-Unit Problem constraints (7), (8) and (9) over $j \in \mathcal{S}$. Therefore, the feasible set of the N-Units Problem is contained in the feasible set of the AUR. To prove Proposition 1, it suffices to show that (13) is a lower bound of (6) for every feasible solution $x \in \mathcal{M}_{N\times T}(\mathbb{R})$ of the N-Units Problem. We show this by proving that

$$C_t\Big(h\Big(\sum_{j\in\mathcal{S}} x_{j,t}\Big)\Big) \leq C_t\Big(\sum_{j\in\mathcal{S}} h_j(x_{j,t})\Big) \tag{36}$$

for every $t \in \mathcal{T}$. The key ingredient is the following lemma:

Lemma 1. *The inequalities*

$$\sum_{j\in\mathcal{S}} x_{j,t} \leq h\Big(\sum_{j\in\mathcal{S}} x_{j,t}\Big) \leq \sum_{j\in\mathcal{S}} h_j(x_{j,t}) \tag{37}$$

hold for every $t \in \mathcal{T}$ and every feasible solution $x \in \mathcal{M}_{N\times T}(\mathbb{R})$ of the N-Units Problem.

Lemma 1 and the assumption that the cost functions C_t are monotonically increasing in their domain I immediately give (36). This concludes the proof of Proposition 1. We proceed now to prove Lemma 1.

Proof. Proof of Lemma 1. Let $x \in \mathcal{M}_{N\times T}(\mathbb{R})$ be a feasible solution of the N-Units Problem. To show (37), we will do a case distinction depending on whether

$$\sum_{j\in\mathcal{S}} x_{j,t} \geq 0 \tag{38}$$

holds or not.

If (38) holds, we have

$$\sum_{j\in\mathcal{S}} x_{j,t} = h\Big(\sum_{j\in\mathcal{S}} x_{j,t}\Big) = \sum_{\substack{j\in\mathcal{S}\\ x_{j,t}\geq 0}} x_{j,t} + \sum_{\substack{j\in\mathcal{S}\\ x_{j,t}<0}} x_{j,t} \\ \leq \sum_{\substack{j\in\mathcal{S}\\ x_{j,t}\geq 0}} x_{j,t} + \sum_{\substack{j\in\mathcal{S}\\ x_{j,t}<0}} \eta_j x_{j,t} = \sum_{j\in\mathcal{S}} h_j(x_{j,t}), \tag{39}$$

where the first equality follows from the definition of h in (12) and the inequality from the fact that $0 < \eta_j \leq 1$ for every $j \in \mathcal{S}$.

If (38) does not hold, we have

$$\sum_{j\in\mathcal{S}} x_{j,t} \leq \eta \sum_{j\in\mathcal{S}} x_{j,t} = h\Big(\sum_{j\in\mathcal{S}} x_{j,t}\Big) = \sum_{\substack{j\in\mathcal{S}\\ x_{j,t}\geq 0}} \eta x_{j,t} + \sum_{\substack{j\in\mathcal{S}\\ x_{j,t}<0}} \eta x_{j,t} \\ \leq \sum_{\substack{j\in\mathcal{S}\\ x_{j,t}\geq 0}} x_{j,t} + \sum_{\substack{j\in\mathcal{S}\\ x_{j,t}<0}} \eta_j x_{j,t} = \sum_{j\in\mathcal{S}} h_j(x_{j,t}) \tag{40}$$

where the first inequality follows from $0 < \eta \leq 1$ and the second inequality from $0 < \eta_j \leq \eta \leq 1$ for every $j \in \mathcal{S}$ by definition (11). □

EC.2 Negative electricity prices

In this appendix, we describe two challenges that arise in the model introduced in Section 2 and in the results in Section 3 if we do not assume that the cost functions C_t are monotonically increasing in their domain I for every $t \in \mathcal{T}$. We also introduce a mild modification to the objective function (6) and show that it addresses both issues. Therefore, the monotonicity assumption can be safely relaxed, which allows to incorporate negative electricity prices in the model and solution method.

A cost function C_t that is not monotonically increasing in its domain I has a minimum $m_t \in I$ that is not $-\sum_j P^o_j$. Therefore, we are in a situation where taking a cumulative action $z_t < m_t$ results in a lower cumulative state of charge and a higher or equal cost (or a lower or equal revenue) than taking action m_t. It seems unlikely that storage units would take such actions because there exist other actions that result in a higher state of charge with a smaller or equal cost. A prime example of this circumstance are negative electricity prices, where there exists a minimum $m_t \in I$ of C_t satisfying $m_t > 0$ and $C_t(m_t) < 0$. Therefore, convexity of C_t and the assumption that $C_t(0) = 0$ guarantee that $C_t(z) \geq 0$ if $z < 0$ and $C_t(z) < 0$ at least for some $z > 0$. In other words, the storage units incur a cost to discharge and generate a revenue from charging, at least up to a certain amount. One expects that, under negative prices, a storage unit would prefer to charge and get a revenue, instead of discharge and incur a cost.

EC.2.1 Convexity of the Objective Function

The first challenge that arises if there exists $t \in \mathcal{T}$ such that the cost function C_t is not monotonically increasing in its domain I is that the objective function (6) is no longer convex. This is caused by the interaction of round-trip efficiencies with cost functions C_t that are not monotonically increasing in their domain I. In Figure 6, we can observe this effect under negative prices in the 1-Unit Problem context. Convexity is broken by the abrupt change of slope at $z = 0$. The situation only gets worse for multiple units. This lack of convexity introduces a difficulty to approaches that require minimizing the Lagrangian of these functions, as is the case of this work, because the global minimum need not be unique and there might exist local minima.

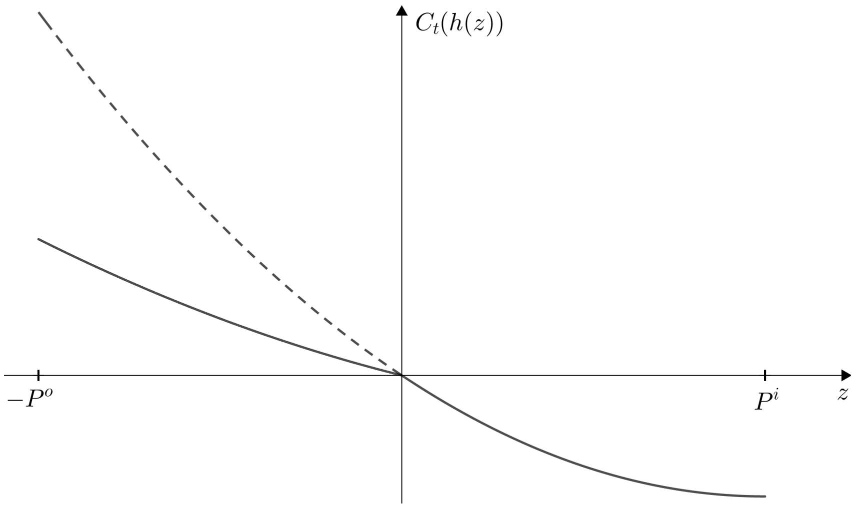

Figure 6: Loss of convexity of cost functions under negative prices due to round-trip efficiencies. $C_t(h(z))$ [solid line] and $C_t(z)$ [dashed line].

This lack of convexity has an interesting interpretation. Under negative prices, the storage units benefit from efficiency losses in the discharging regime because they do not have to pay to dispose of the electricity that is lost. Zhou et al. (2016) provide further insights on optimal storage policies under negative prices and compare storage with disposal, which is similar to the effect introduced by round-trip efficiencies.

To overcome this difficulty, we restore convexity of the objective function (6) by introducing a mild modification. Instead of the objective function (6) of the N-Units Problem, we consider

$$\sum_{t \in \mathcal{T}} C_t^N(x_t) \tag{41}$$

where $C_t^N : \prod_{j \in \mathcal{S}} [-P_j^o, P_j^i] \to \mathbb{R}$ is defined by

$$C_t^N(x_t) := \max\Big\{ C_t\Big(\sum_{j \in \mathcal{S}} h_j(x_{j,t})\Big), C_t\Big(\sum_{j \in \mathcal{S}} x_{j,t}\Big)\Big\} \tag{42}$$

for every $t \in \mathcal{T}$.

Lemma 1 in Appendix EC.1 helps understand the changes introduced by replacing the objective function (6) with (41). If C_t is monotonically increasing in its domain I, Lemma 1 guarantees that we have not introduced any changes at those times. Therefore, this modification is a generalization of the assumptions made in the main body of the paper. Furthermore, although negative electricity prices are becoming more common in day-ahead markets due to renewables and limited flexibility of some generators and of the demand side, they remain rare events (De Vos (2015), Bajwa and Cavicchi (2017), Aust and Horsch (2020)). Thus, we expect C_t to be monotonically increasing for most $t \in \mathcal{T}$, which means that these changes only affect a small proportion of time steps.

Moreover, again by Lemma 1, if the modifications introduced do take effect, they penalize taking actions in the region of the domain of C_t where it is monotonically decreasing, i.e., where reducing further the states of charge would result in a smaller revenue or even incurring a cost. These actions are still allowed, but at a higher cost. As we have argued earlier in this appendix, we expect that taking such actions will rarely be optimal, and therefore there is no harm in adding an extra penalty on them.

In summary, (41) introduces changes in very few time steps, and when it does introduce them, it is most unlikely that the optimal solution will be affected by them. Therefore, we believe that the changes introduced by (41) do not significantly alter the optimal solution, if at all. In Proposition 3, we show that the changes introduced by (41) result in a convex objective function.

Proposition 3. *The function C_t^N defined in (42) is convex for every $t \in \mathcal{T}$.*

Proof. Proof. Let $t \in \mathcal{T}$, $\alpha \in [0,1]$, and $x_t, y_t \in \prod_{j\in\mathcal{S}}[-P_j^o, P_j^i]$ be feasible actions of the N storage units at time t. We want to show that

$$C_t^N(\alpha x_t + (1-\alpha)y_t) \leq \alpha C_t^N(x_t) + (1-\alpha)C_t^N(y_t). \tag{43}$$

Since C_t is convex in the compact set I, there exists a closed interval of minima $[m_t^-, m_t^+] \subseteq I$ of C_t that defines the regions where C_t is monotonically decreasing and increasing. By Lemma 1, we have

$$\alpha \sum_{j\in\mathcal{S}} x_{j,t} + (1-\alpha)\sum_{j\in\mathcal{S}} y_{j,t} \leq \sum_{j\in\mathcal{S}} h_j(\alpha x_{j,t} + (1-\alpha)y_{j,t}). \tag{44}$$

for every $t \in \mathcal{T}$, since $\alpha x_t + (1-\alpha)y_t$ is a feasible action for the N units at time $t \in \mathcal{T}$. We will proceed by case separation based on (44).

<u>Case 1</u>: $\sum_{j\in\mathcal{S}} h_j(\alpha x_{j,t} + (1-\alpha)y_{j,t}) \leq m_t^+$

In this case, we have

$$\begin{aligned} C_t^N(\alpha x_t + (1-\alpha)y_t) &= C_t\Big(\alpha \sum_{j\in\mathcal{S}} x_{j,t} + (1-\alpha)\sum_{j\in\mathcal{S}} y_{j,t}\Big) \\ &\leq \alpha C_t\Big(\sum_{j\in\mathcal{S}} x_{j,t}\Big) + (1-\alpha)C_t\Big(\sum_{j\in\mathcal{S}} y_{j,t}\Big) \\ &\leq \alpha C_t^N(x_t) + (1-\alpha)C_t^N(y_t). \end{aligned} \tag{45}$$

where we used (44) and this case assumption in the first equality, convexity of C_t in the first inequality, and definition (42) in the last inequality.

<u>Case 2</u>: $m_t^- \leq \alpha \sum_{j\in\mathcal{S}} x_{j,t} + (1-\alpha)\sum_{j\in\mathcal{S}} y_{j,t}$

In this case, (44) gives

$$C_t^N(\alpha x_t + (1-\alpha)y_t) = C_t\Big(\sum_{j\in\mathcal{S}} h_j(\alpha x_{j,t} + (1-\alpha)y_{j,t})\Big). \tag{46}$$

Furthermore, since h_j is convex for every $j \in \mathcal{S}$, we have

$$h_j(\alpha x_{j,t} + (1-\alpha)y_{j,t}) \le \alpha h_j(x_{j,t}) + (1-\alpha)h_j(y_{j,t}) \tag{47}$$

for every $j \in \mathcal{S}$ and $t \in \mathcal{T}$. Therefore, in this case, we get

$$m_t^- \le \sum_{j\in\mathcal{S}} h_j(\alpha x_{j,t} + (1-\alpha)y_{j,t}) \le \alpha \sum_{j\in\mathcal{S}} h_j(x_{j,t}) + (1-\alpha)\sum_{j\in\mathcal{S}} h_j(y_{j,t}). \tag{48}$$

Finally, we have

$$\begin{aligned} C_t^N(\alpha x + (1-\alpha)y) &= C_t\Big(\sum_{j\in\mathcal{S}} h_j(\alpha x_{j,t} + (1-\alpha)y_{j,t})\Big) \\ &\le C_t\Big(\alpha\sum_{j\in\mathcal{S}} h_j(x_{j,t}) + (1-\alpha)\sum_{j\in\mathcal{S}} h_j(y_{j,t})\Big) \\ &\le \alpha C_t\Big(\sum_{j\in\mathcal{S}} h_j(x_{j,t})\Big) + (1-\alpha)C_t\Big(\sum_{j\in\mathcal{S}} h_j(y_{j,t})\Big) \\ &\le \alpha C_t^N(x_t) + (1-\alpha)C_t^N(y_t), \end{aligned} \tag{49}$$

where we used (46) in the first equality, (47) and (48) in the first inequality, convexity of C_t in the second inequality and definition (42) in the last inequality.

Case 3: $\alpha\sum_{j\in\mathcal{S}} x_{j,t} + (1-\alpha)\sum_{j\in\mathcal{S}} y_{j,t} \le m_t^- \le m_t^+ \le \sum_{j\in\mathcal{S}} h_j(\alpha x_{j,t} + (1-\alpha)y_{j,t})$
In this case, we do not know a priori which value $C_t^N(x)$ might take. Nevertheless, we can prove the desired bound for both possible values, which we do by reducing it to Case 1 or Case 2 above.

By convexity of C_t, we have

$$\begin{aligned} C_t\Big(\alpha\sum_{j\in\mathcal{S}} x_{j,t} + (1-\alpha)\sum_{j\in\mathcal{S}} y_{j,t}\Big) &\le \alpha C_t\Big(\sum_{j\in\mathcal{S}} x_{j,t}\Big) + (1-\alpha)C_t\Big(\sum_{j\in\mathcal{S}} y_{j,t}\Big) \\ &\le \alpha C_t^N(x) + (1-\alpha)C_t^N(y). \end{aligned} \tag{50}$$

Note that this argument is analogous to the one in Case 1. By convexity of h_j, we get (47), from which (48) follows in this case as well. One can then show

$$C_t\Big(\sum_{j\in\mathcal{S}} h_j(\alpha x_{j,t} + (1-\alpha)y_{j,t})\Big) \le \alpha C_t^N(x) + (1-\alpha)C_t^N(y) \tag{51}$$

following the same steps as in (49) for Case 2. (50) and (51) conclude the proof. □

EC.2.2 The Aggregated Unit Relaxation as a Relaxation of the N-Units Problem

The second challenge that arises if there exists $t \in \mathcal{T}$ such that the cost function C_t is not monotonically increasing in its domain I is that AUR is no longer a relaxation of the N-Units Problem. Indeed, (36) no longer follows from Lemma

1. The interpretation of this behaviour is analogous to the one in Section EC.2.1. Under negative prices, in the discharging regime, it is beneficial for the storage units to be inefficient. Recall that in (11) we chose the aggregated unit to be as efficient as the most efficient of the N-Units.

To overcome this difficulty, we propose a modification in the objective function (13) of AUR analogous to (41). Instead of the objective function (13), consider

$$\sum_{t\in\mathcal{T}} C_t^A(x), \tag{52}$$

where $C_t^A : \mathcal{M}_{N\times T}(\mathbb{R}) \to \mathbb{R}$ is defined by

$$C_t^A(x) := \max\left\{C_t\Big(h\Big(\sum_{j\in\mathcal{S}} x_{j,t}\Big)\Big), C_t\Big(\sum_{j\in\mathcal{S}} x_{j,t}\Big)\right\}. \tag{53}$$

Observe that this is (41) in the 1-Unit Problem context. Therefore, the results obtained in Section EC.2.1 also hold for (52). In particular, Proposition 3 gives that C_t^A is convex and therefore so is (52). We now want to show that the modified AUR is still a relaxation of the modified N-Units Problem.

Proposition 4. *The Aggregated Unit Relaxation with objective function (52) is a relaxation of the N-Units Problem with objective function (41).*

To show Proposition 4, it only remains to show the bound on the objective function, since we already showed in Section 3 the inclusion of feasible sets. We state the required result as Lemma 2 below.

Lemma 2. *The inequality*

$$C_t^A(x) \le C_t^N(x) \tag{54}$$

holds for every $t \in \mathcal{T}$ and every feasible solution $x \in \mathcal{M}_{N\times T}(\mathbb{R})$ of the N-Units Problem.

Proof. Proof. Let $t \in \mathcal{T}$. Since C_t is convex in the compact set I, there exists a closed interval of minima $[m_t^-, m_t^+] \subseteq I$ of C_t that defines the regions where C_t is monotonically decreasing and increasing. By Lemma 1, we have

$$\sum_{j\in\mathcal{S}} x_{j,t} \le h\Big(\sum_{j\in\mathcal{S}} x_{j,t}\Big) \le \sum_{j\in\mathcal{S}} h_j(x_{j,t}). \tag{55}$$

The proof proceeds by case separation using (55):

<u>Case 1</u>: $m_t^- \le h\Big(\sum_{j\in\mathcal{S}} x_{j,t}\Big)$

This case assumption and (55) give

$$C_t\Big(h\Big(\sum_{j\in\mathcal{S}} x_{j,t}\Big)\Big) \le C_t\Big(\sum_{j\in\mathcal{S}} h_j(x_{j,t})\Big), \tag{56}$$

which implies (54). Note that if C_t is monotonically increasing in its domain I, then we are in this case.

<u>Case 2</u>: $h\Big(\sum_{j\in\mathcal{S}} x_{j,t}\Big) \le m_t^- \le m_t^+ \le \sum_{j\in\mathcal{S}} h_j(x_{j,t})$

Observe that if (56) holds, we can conclude as in Case 1. If it does not hold, then

$$C_t\Big(\sum_{j\in\mathcal{S}} h_j(x_{j,t})\Big) \le C_t\Big(h\Big(\sum_{j\in\mathcal{S}} x_{j,t}\Big)\Big) \le C_t\Big(\sum_{j\in\mathcal{S}} x_{j,t}\Big) \tag{57}$$

where the second inequality follows from this case assumption and (55). Therefore, we have

$$C_t^A(x) = C_t\Big(\sum_{j\in\mathcal{S}} x_{j,t}\Big) = C_t^N(x) \tag{58}$$

from which (54) follows.

Case 3: $\sum_{j\in\mathcal{S}} h_j(x_{j,t}) \le m_t^+$

This case assumption and (55) give (57), and we can conclude as in the second part of Case 2. □

EC.3 The Algorithm

Algorithm 1 is the pseudocode of our solution method:

Algorithm 1: Double Decomposition Algorithm.

Choose $\epsilon > 0$ (convergence tolerance)
Set $\nu = 0$, $z = 0$ and $x_j = 0$ for every $j \in \mathcal{S}$
Set CurrentTime $= 0$
while CurrentTime $< T$ **do**
  repeat
    Solve Problem Q_1 from CurrentTime until next forecast horizon and obtain:
    - ADMM-DecisionHorizon and ADMM-ForecastHorizon
    - $x_{1,t}$ for those times

    for $j = 2, \dots, N$ **do**
      Solve Problem Q_j from CurrentTime until ADMM-ForecastHorizon and update $x_{j,t}$ for those times
    Solve Problem Q_0 from CurrentTime until ADMM-ForecastHorizon and update z_t for those times.
    Update ν_t using (26) from CurrentTime until ADMM-ForecastHorizon
    Update ConstraintViolation as the maximum violation of linking constraints (20) between CurrentTime and ADMM-DecisionHorizon
  until ConstraintViolation $< \epsilon$
  Store values of z_t and $x_{j,t}$ for every $j \in \mathcal{S}$ from CurrentTime until ADMM-DecisionHorizon.
  Update CurrentTime to CurrentTime + ADMM-DecisionHorizon

EC.4 Proof of Proposition 2

Since the N-Units Problem and Problem P are equivalent, it suffices to show it for Problem P. The key observation is that ALR is a partial Lagrangian Relaxation of Problem P with objective function (21) replaced by

$$\sum_{t\in\mathcal{T}}\Big[C_t(h(z_t))+\frac{\gamma}{2}\Big(\sum_{j\in\mathcal{S}}h_j(x_{j,t})-h(z_t)\Big)^2\Big]. \tag{59}$$

This modification of Problem P is still equivalent to Problem P. Indeed, they have the same feasible region and the objective functions are equal in it, since the second term in (59) vanishes in the feasible region of Problem P.

Furthermore, this modification of Problem P is a convex problem with linear constraints. Since we assumed in Section 2 that a feasible solution to the N-Units Problem exists, Slater's conditions (Boyd and Vandenberghe (2004)) hold, and therefore its optimal value is equal to the optimal value of maximizing its partial Lagrangian Relaxation ALR over $\nu\in\mathbb{R}^T$.

EC.5 The Subproblems from the Decomposition by Unit

In this appendix, we present the mathematical formulation of the subproblems that need to be solved to update the decision variables in (25). For every $j\in\mathcal{S}$, $x_j\in\mathbb{R}^T$ is updated by solving Problem Q_j:

$$\underset{x_j\in\mathbb{R}^T}{\text{Minimize}}\quad \sum_{t\in\mathcal{T}}\Big[\nu_t^{(v-1)}h_j(x_{j,t})+\frac{\gamma}{2}\Big(h_j(x_{j,t})+\sum_{\substack{k\in\mathcal{S}\\k<j}}h_k(x_{k,t}^{(v)})+\sum_{\substack{k\in\mathcal{S}\\k>j}}h_k(x_{k,t}^{(v-1)})-h(z_t^{(v-1)})\Big)^2\Big] \tag{60}$$

$$\text{subject to}\quad -P_j\le x_{j,t}\le P_j \qquad \forall t\in\mathcal{T} \tag{61}$$

$$0\le\rho_j^t\bar{S}_{j,0}+\sum_{l=1}^{t}\rho_j^{t-l}x_{j,l}\le E_j \qquad \forall t\in\mathcal{T} \tag{62}$$

$$\rho_j^T\bar{S}_{j,0}+\sum_{l=1}^{T}\rho_j^{T-l}x_{j,l}=\bar{S}_{j,T}. \tag{63}$$

Observe that the objective function (60) uses the values $x_{k,t}^{(v)}$ for $k<j$. The order defined in (25) guarantees that they have already been computed at the time of solving Problem Q_j.

The $z\in\mathbb{R}^T$ update is done by solving Problem Q_0:

$$\underset{z\in\mathbb{R}^T}{\text{Minimize}}\quad \sum_{t\in\mathcal{T}}\Big[C_t(h(z_t))-\nu_t^{(v-1)}h(z_t)+\frac{\gamma}{2}\Big(\sum_{j\in\mathcal{S}}h_j(x_{j,t}^{(v)})-h(z_t)\Big)^2\Big] \tag{64}$$

$$\text{subject to}\quad -\sum_{j\in\mathcal{S}}P_j\le z_t\le\sum_{j\in\mathcal{S}}P_j \qquad \forall t\in\mathcal{T} \tag{65}$$

$$0 \leq \rho^t \sum_{j\in\mathcal{S}} \bar{S}_{j,0} + \sum_{l=1}^{t} \rho^{t-l} z_l \leq \sum_{j\in\mathcal{S}} E_j \qquad \forall t \in \mathcal{T} \tag{66}$$

$$\rho^T \sum_{j\in\mathcal{S}} \bar{S}_{j,0} + \sum_{l=1}^{T} \rho^{T-l} z_t = \sum_{j\in\mathcal{S}} \bar{S}_{j,T}. \tag{67}$$

Observe that Problems Q_0 and Q_j for all $j \in \mathcal{S}$ are 1-Unit Problems with convex cost functions. Therefore, they can be solved with the algorithm of Cruise et al. (2019).