EUROPEAN COMMISSION

HORIZON 2020 PROGRAMME – TOPIC: Hybridisation of battery systems for stationary energy storage

Interoperable, modular and Smart hybrid energy STORage systeM for stationarY applications

GRANT AGREEMENT No. 963527

Deliverable Report

D4.5 –Report on upscaling assessment of the HESS to extra use case applications

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 963527

Deliverable No.	iSTORMY D4.5	
Related WP	WP4	
Deliverable Title	Report on upscaling assessment of the HESS to extra use case applications	
Deliverable Date	29-11-2023	
Deliverable Type	REPORT	
Dissemination level	Public (PU)	
Written By	Denis Vettoretti, Diego Cifelli, Adolfo Anta	
Checked by	Eneko Unamuno (MGEP)	
Reviewed by	Tim Meulenbroeks (TNO)	
Approved by	Project Coordinator (VUB)	
Status	Version 1.1	

Disclaimer/ Acknowledgment

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 963527. The information and views set out in this publication does not necessarily reflect the official opinion of the European Commission. Neither the European Union institutions and bodies nor any person acting on their behalf, may be held responsible for the use which may be made of the information contained therein.

Summary

The iSTORMY project aims at developing an interoperable and modular Hybrid Energy Storage System (HESS) by demonstrating various use cases and seamlessly interfacing the grid to provide multiple services. This report extends the results of the iStormy demonstrator by considering additional grid services and applications of storage systems, and thus assessing the features of HESS with respect to traditional storage systems in terms of grid applications. Standard grid services and typical storage applications were historically envisioned having a single, uniform unit in mind. It is therefore unclear how to operate a hybrid storage system in order to provide these services, namely, how to coordinate the usage of the different modules in an optimal way. In this report we propose methods to coordinate the different modules in a hybrid storage system, thereby reducing the uncertainty behind the operation of a HESS. In particular, three use cases are considered as examples: first, the coordinated provision of frequency and voltage support (that is, the simultaneous provision of active and reactive power), which is expected to be needed in weak grids. Second, peak shaving at an industrial site, and by extension, any kind of service that can be framed as an optimization problem, with a well-defined cost to be minimized and where forecast for prices and/or demand are available. Third, the impact of primary frequency control on the aging of the different modules of a HESS has been evaluated, considering a typical frequency evolution in continental Europe. Two different criteria are used to define the primary frequency service, namely the Nordic criterium and the CE regulation. These methods describe how to leverage the capabilities of HESS to provide typical grid services that were originally thought for single energy storage systems. While for certain services the hybridization can be easily handled using simple extensions of the standard methods for classical uniform storage systems, in other use cases it represents an increased complexity, and appropriate methods need to be developed to optimally operate a HESS while considering the different characteristics of each module.

Contents

1 Table of Figures

2 Tables

Table 1 $Q(U)$ and $P(f)$ parameterization 14
Table 2 - Piecewise coefficients for the HE module 25
Table 3 - Piecewise coefficients for the HP module 25
Table 4 - Optimality gap for different number of regions of the HE module 27
Table 5: Requirements for LER according to Nordic Synchronous Area regulation 30

3 Introduction

The ongoing energy transition requires new paradigms and new assets capable of addressing novel grid needs. Storage units can address several of these system requirements, even more when these units are composed of different sources, such as batteries, supercapacitors, electrolysers, thermal storage or pumped hydro plants. In this spirit, hybrid power plants can leverage heterogeneous capabilities in a collocated manner, in order to identify and exploit synergies between different sources, such as PV and wind. Among all these new system needs, the transition from conventional generation to inverter-based generation is motivating the provision of ancillary services from units that interface the grid via power electronics. This implies that inverters will be required to provide a certain amount of active and reactive power, based on for instance frequency and voltage. Ideally these services are provided in an optimal way, that is, maximizing the efficiency of the units composed of inverters and storage. At the same time, since storage units are becoming larger, they are now typically composed of different units with different characteristics to address different needs, e.g., they can be composed of a high-energy and a high-power module and their correspondent submodules.

Given that typical grid services were designed before hybrid power plants (or hybrid storage systems) were envisaged, there are some uncertainties on how to leverage the capabilities of each module, and how to distribute the setpoints among the different modules present in a hybrid unit. While simple rule-based approaches may be used, they represent very suboptimal strategies that do not maximize the revenue potential of HESS. In particular, for homogeneous storage units it has been observed that the optimal operation is needed to present a profitable business case in certain countries. In this spirit, the goal is to look at some exemplary use cases and analyse how to coordinate the different units in an optimal way. This task attempts at reducing this uncertainty by spelling out possible methods to coordinate the different modules within the unit. To the best of our knowledge, there are very few results in this direction, given that hybrid storage units are not widespread. In all the use cases, it is assumed that there is a low level EMS (as in [1]) that splits setpoints into the different modules, and the focus is on developing high level control solutions to address the requirements of each one of the presented use cases.

4 Coordinated voltage/frequency control for weak distribution grids

The increase of renewable energy sources is radically changing the characteristics of the electrical grid introducing new challenges that need to be addressed in order to ensure grid stability. The transition from conventional generation to inverter-based generation requires to adapt power converters for the provision of ancillary services. Frequency response and voltage support start to become part of the grid code in many countries all around the world and hybrid storage systems, thanks to their nature, can play an important role in the grid of tomorrow.

In this spirit a coordinated voltage/frequency control for a weak distribution grid is implemented in a hybrid storage system, with the aim of guaranteeing voltage and frequency stability.

4.1 Grid model

The topology of the grid used for this use-case is inspired by the IEEE13 Node test feeder. This model is widely used to test common features of distribution grids. The operation voltage is 4.16 kV, and the feeder is characterized by being short and highly loaded that makes it a good representation of a weak grid. The unbalanced loading characteristic of the original model is removed because it is not relevant for this use-case.

Figure 1 shows a one-line diagram representation of the grid used to conduct the analysis. The HESS composed by a high energy (HE) and high power (HP) module is connected to node 675 that is the node further away from the primary substation (650) and therefore, it also represents the weakest point in the grid. At node 675 voltage variations are more severe making it the optimal location to observe how the voltage/frequency controllers react and interact together.

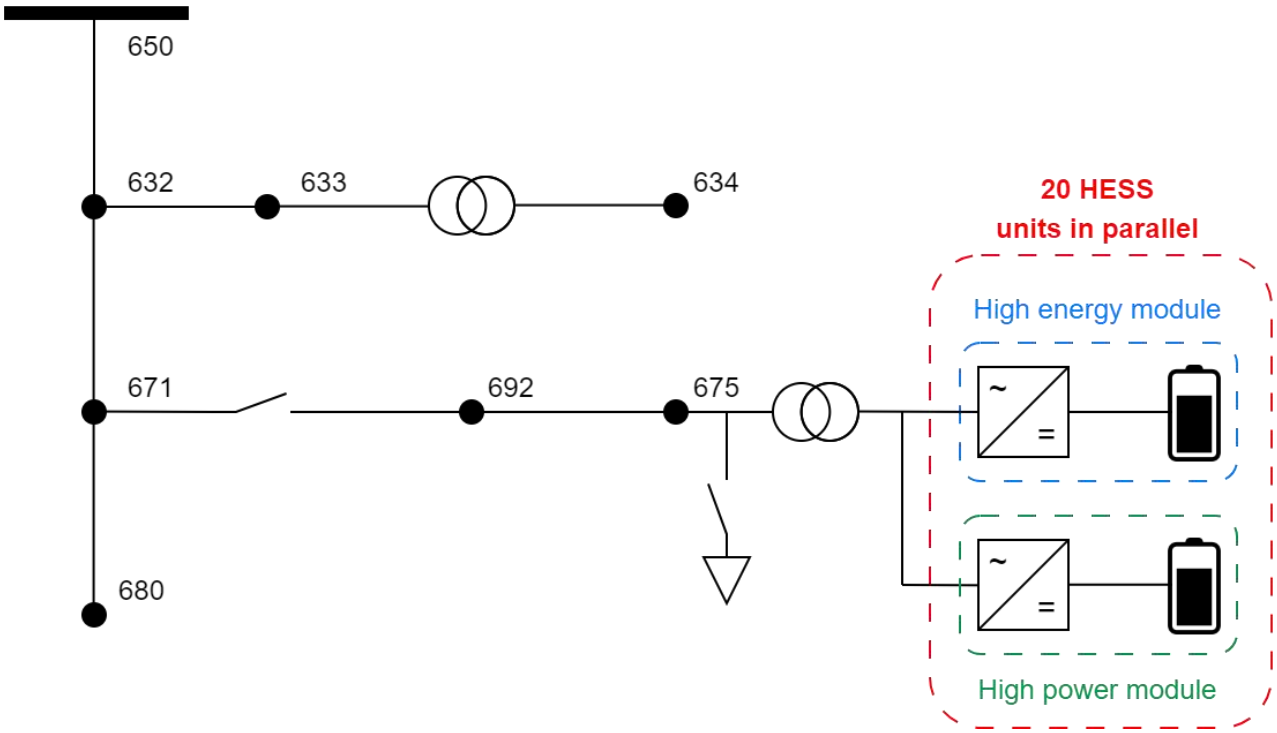

Figure 1 IEEE 13 node test feeder adapted in order to represent a weak grid.

To carry out the analysis some simplifications have been introduced in order to make possible to run the simulations in a reasonable amount of time. The HESS is connected to the feeder with a transformer where the low-voltage side is at 400V. The inverter models have been simplified for both the HP and HE modules. The fast-switching frequency used by the manufactures of the inverters imposes strong requirements on simulation time steps. For the functionality analysed

here these fast dynamics are not relevant and thus the power electronics have been simplified and an average inverter model is used instead of a switching one. Originally the HP module is composed by one inverter with a total nominal power of 92 kW and the HE module by two inverters with a total nominal power of 32 kW.

The total nominal power of one HESS unit is not sufficient to induce visible voltage deviations, therefore 20 HESS units are connected in parallel reaching a nominal power of 2.16 MW. Connecting multiple inverters in parallel drastically increases the simulation time. For this reason, the HP and HE modules are modelled as a reduced order aggregate model inverter. The aggregate inverter model represents an aggregation of 20 units as can be seen in Figure 1. The aggregate model dimensioning follows the procedure suggested in [2].

To properly evaluate the controller algorithms the grid model was developed in MATLAB/Simulink and electromagnetic simulations (EMT) with a time step of 100us are performed. In Figure 2 the MATLAB/Simulink diagram is shown, where the yellow block represents the feeder, the red block a load disturbance, and on the right side of the figure there are the two aggregate inverters modules representing 20 HESS units.

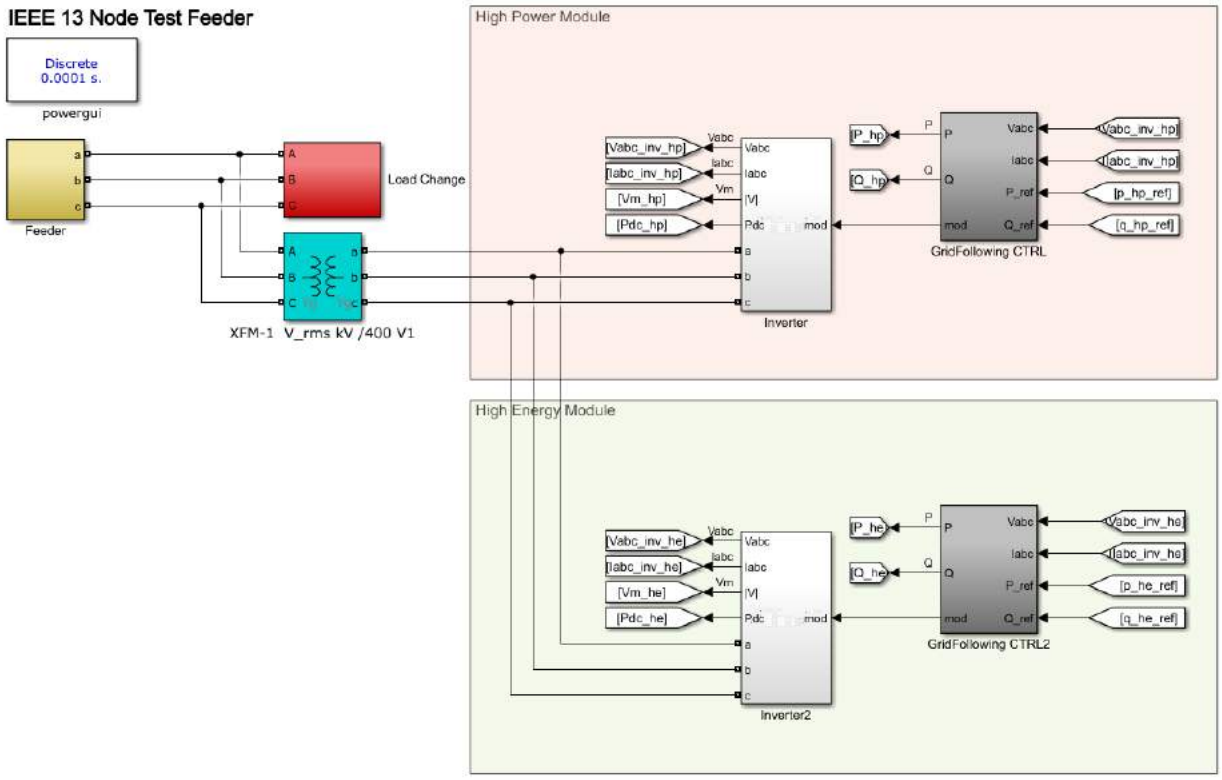

Figure 2 MATLAB/Simulink diagram of IEEE13 test node feeder used for the simulations.

4.2 Controllers

The voltage and frequency controllers developed for this use case are combined with the EMS module presented in the previous tasks within the iStormy project [3]. Starting from the left side of Figure 3 there are two blocks named P(f) and Q(U), that implement the voltage and frequency controllers. A detailed description of them will follow in the next sections. The P(f) and Q(U) controllers receive as input the frequency and voltage measurements from the grid respectively and provide as output active (P_{unlim}) and reactive (Q_{unlim}) power setpoints.

The EMS block in grey colour is a slightly modified replica of the one developed in [3], [1] and it incorporates a battery model in addition to a power splitting algorithm that generates the setpoints for the HE and HP batteries. In the original EMS model, a grid was not considered and the setpoints for the HE and HP where directly used to compute the SoC of the batteries. In this case instead, the active power measurements ($P_{HP\ meas}$, $P_{HE\ meas}$) from the HE and HP inverters are used to compute the SoC of the batteries and fed into the EMS block. The power losses of the inverters are considered

negligible and not relevant for the current use case. The EMS block receives two additional inputs, the grid frequency measured with a phase-locked loop controller (PLL) and the P_{unlim} setpoint from the P(f) controller that are used by the power splitting algorithm to compute the active set points for the battery modules ($P_{HP\ inlim}$, $P_{HE\ unlim}$). The last block in yellow on the right side of Figure 3 represents the active and reactive power prioritization block that will be describe in more details in 4.2.3.

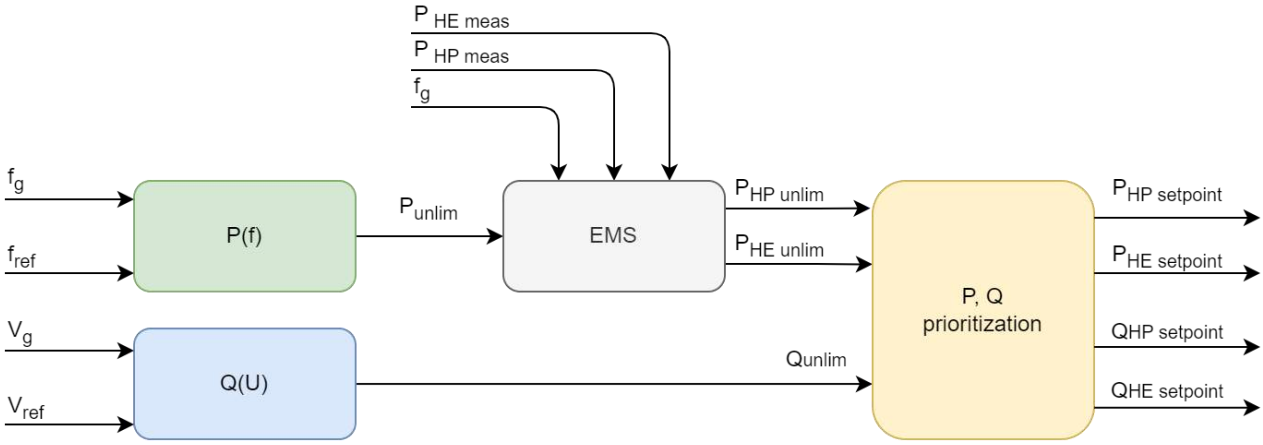

Figure 3 Voltage and frequency controllers combined with the EMS module

4.2.1 Voltage controller: Q(U)

A voltage control reacts to voltage deviations from the nominal value in order to maintain the voltage level within the limits imposed by the grid operators. The grid code varies from country to country and the parameters need to be adapted based on the local regulation. Nevertheless, the actual control law is in all cases represented by a droop curve as shown in Figure 4.

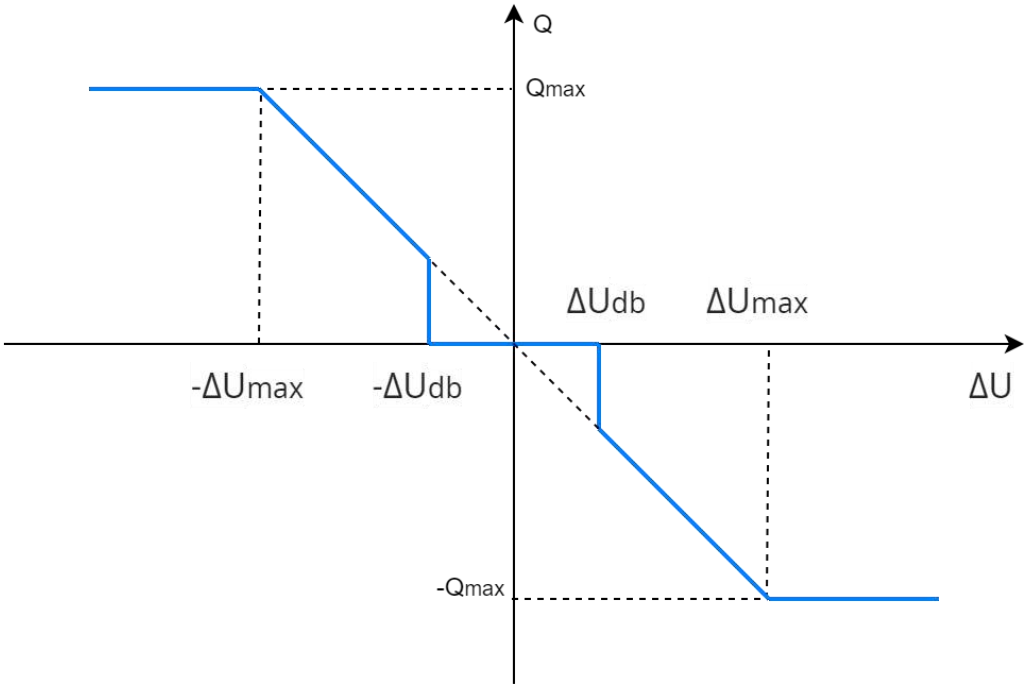

Figure 4 Voltage supporting Q(U) droop controller characteristic.

The Q(U) controller injects reactive power into the grid when the voltage is lower compared to the nominal value and in the opposite way, it absorbs reactive power when the voltage is higher than the nominal value.

The droop curve in Figure 4 can be divided in 5 regions based on the voltage deviations:

- **Dead band,** $\Delta \boldsymbol{U} \in [-\Delta \boldsymbol{U_{db}}, \Delta \boldsymbol{U_{db}}]$**:** when the voltage is inside this interval no action is taken.
- **High voltage band** $\Delta \boldsymbol{U} \in [\Delta \boldsymbol{U_{db}}, \Delta \boldsymbol{U_{max}}]$**:** voltage deviations larger than ΔU_{db} are considered too high and the controller will take a countermeasure to limit the voltage. The device will start to absorb reactive power proportionally to the voltage deviation. The maximum reactive power absorbed by the device can be controlled by -Q_{max}.
- **Critical high voltage band** $\Delta \boldsymbol{U} > \Delta \boldsymbol{U_{max}}$**:** voltage deviations over ΔU_{max} are considered critical and the device will absorb as much reactive power as possible i.e. - Q_{max}.
- **Low voltage band** $\Delta \boldsymbol{U} \in [-\Delta \boldsymbol{U_{max}}, -\Delta \boldsymbol{U_{db}}]$**:** analogous to the high voltage band but in the opposite direction. The device will inject reactive power into the grid proportionally to the voltage deviation.
- **Critically low voltage band** $\Delta \boldsymbol{U} < -\Delta \boldsymbol{U_{max}}$**:** similar to the critical high voltage band but in the opposite direction.

It is possible to shape different droop curve modifying $-\Delta U_{max}, \Delta U_{max}$ $-\Delta U_{db}, \Delta U_{db,}$ $-Q_{max}, -Q_{max}$. However, most of the time these parameters are regulated by the grid operators.

4.2.2 Frequency controller: P(f)

In the power systems domain, the frequency is one of the most import parameters to monitor because it is a key indicator of the power distribution quality. Frequency is also directly related to the rotor speed and synchronization of the conventional generators. Voltage and frequency need to be maintained within a specific range in order to safely operate the grid avoiding damage of the electrical equipment. The frequency measurement is also a key indicator of the balance between load and generation as shown in Figure 5.

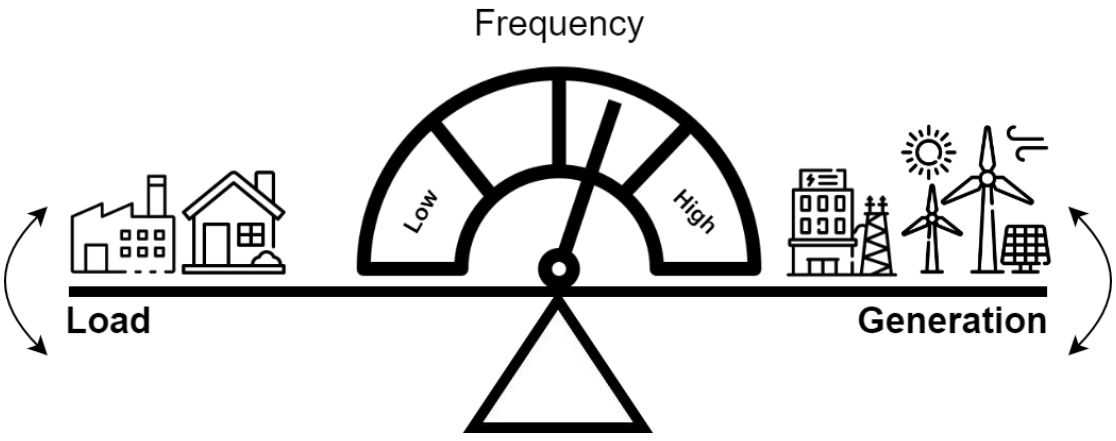

Figure 5 Load and generation balance

If there is a perfect balance between load and generation the frequency is at its nominal voltage. Changing the ratio between load and generation directly affects the frequency. A primary frequency controller P(f) is based on this mechanism, and it will react to a frequency deviation injecting active power into the grid if the frequency is higher than the nominal value and in the opposite direction if the frequency is below the nominal value. A primary frequency controller is characterized by a droop curve as shown in Figure 6. As for the Q(U) controller, the droop curve of the P(f) controller can be divided in 5 regions:

- **Dead band,** $\Delta \boldsymbol{f} \in [-\Delta \boldsymbol{f_{db}}, \Delta \boldsymbol{f_{db}}]$**:** when the frequency is inside this interval no action is taken.

- **High frequency band** $\Delta f \in [\Delta f_{db}, \Delta f_{max}]$**:** frequency deviations larger than ΔU_{db} are considered too high and the controller will take a countermeasure to restore the nominal frequency. The device will start to absorb active power proportionally to the frequency deviation. The maximum active power absorbed by the device can be controlled by $-Pmax$.
- **Critical high frequency band** $\Delta f > \Delta f_{max}$**:** frequency deviations over Δf_{max} are considered critical and the device will absorb as much active power as possible i.e. $-Pmax$.
- **Low frequency band** $\Delta f \in [-\Delta f_{max}, -\Delta f_{db}]$**:** analogous to the high frequency band but in the opposite direction. The device will inject active power into the grid proportionally to the frequency deviation.
- **Critically low voltage band** $\Delta f < -\Delta f_{max}$**:** similar to the critical high frequency band but in the opposite direction.

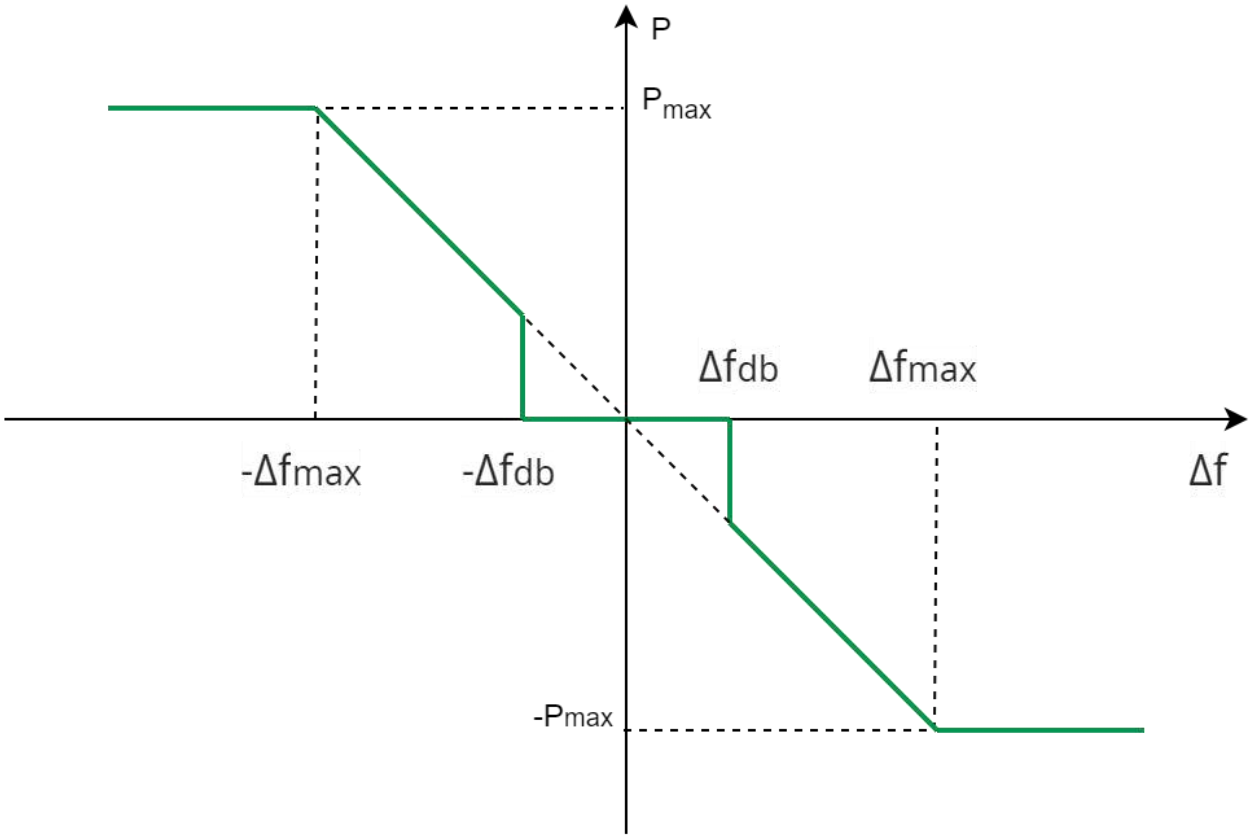

Figure 6 Primary frequency response P(f) droop controller characteristic

Similar to the Q(U) controller, it is possible to shape different droop curves modifying $-\Delta f_{max}$, Δf_{max}, $-\Delta f_{db}$, $\Delta f_{db,}$ $-P_{max}$, $-P_{max}$. However, most of the times these parameters are provided by the grid operators and differ from region to region.

4.2.3 Active and reactive power prioritization

Since both $P(f)$ and $Q(U)$ controllers can be active at the same time it is necessary to prioritize one of them in order to comply with the physical limits of the inverter. For the PQ prioritization algorithm only the power limits of the inverter are considered, that is, $P^2 + Q^2 \leq S_{max}^2$. The user can specify the maximum amount of apparent power that the inverter provides configuring S_{max} for the PQ prioritization module and the prioritization logic i.e., if the active power is prior to the reactive power or vice-versa. The logic behind the PQ prioritization is fairly simple and can be summarized in 2 statements:

Reactive power is set prior to active power

$$\begin{cases} P^2 = S^2 - Q^2 & if\ P^2 + Q^2 > S_{max}^2 \\ no\ limitation & if\ P^2 + Q^2 \leq S_{max}^2 \end{cases}$$

Active power is set prior to reactive power

$$\begin{cases} Q^2 = S^2 - P^2 & if\ P^2 + Q^2 > S_{max}^2 \\ no\ limitation & if\ P^2 + Q^2 \leq S_{max}^2 \end{cases}$$

The sign of the limited P or Q has to be the same as the sign of the unlimited signal.

Other limitation strategies have been considered in the literature, where instead of choosing P or Q to be limited, the prioritization is achieved considering the severity of the voltage or the frequency drop, in a weighted manner. However, in our opinion it does make sense in general to prioritize voltage control as it is a local variable, where other devices cannot really contribute in a meaningful way. However, when it comes to frequency, all inverters in the grid can contribute since the frequency is a global variable.

As shown in Figure 3 the inputs for the PQ prioritization block in yellow are the active power setpoints ($P_{HP\ unlim}, P_{HE\ unlim}$) coming from the EMS block and the reactive power setpoint (Q_{unlim}) coming from the $Q(U)$ controller. The same Q setpoint is used for both the HP and HE module. The PQ prioritization blocks has 4 outputs ($P_{HE}, Q_{HE}, P_{HP}, Q_{HP}$) that are the final active and reactive power setpoints for the HE and HP modules.

4.2.4 Grid following controllers

The $P_{HE}, Q_{HE}, P_{HP}, Q_{HP}$ setpoints generated by the PQ prioritization block are fed to the inverters of the HE and HP modules. The inverters of the HE and HP modules have been equipped with a grid following controller (GFL). This type of controller can be represented by a controlled current source that can be used to achieve the active and reactive power targets regulating the current that is absorbed or injected into the grid. A GFL controller cannot operate in islanded mode. The GLF controller synchronizes to the grid with a PLL that measures the phase angle of the voltage at the point of common coupling (PCC). The active and reactive power regulation are achieved by a cascade of two control loops that form the outer and inner loop of the GFL controller as shown in Figure 7. The outer loop represented in blue is composed by two PI controllers that regulate the active and reactive power injected/absorbed into the grid. Similar to the outer loop, the inner loop represented in red is composed by two PI controllers that are used to regulate the inverter current accordingly to the reference values of the outer loop.

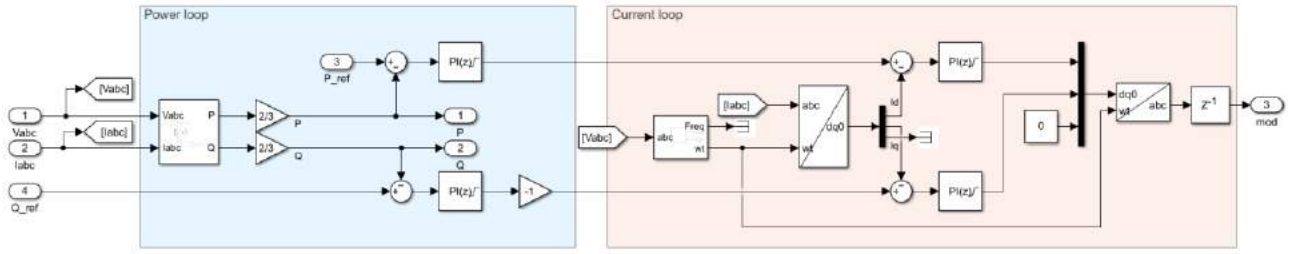

Figure 7 Grid following control diagram

4.3 Grid simulations

Power converters enable a combination of voltage and frequency regulation at the same time. However, in case of extreme events it is necessary to prioritize voltage regulation or frequency regulation to satisfy the physical limits of the converter. In some cases, it is preferred to prioritize frequency regulation, for instance when the power converter is connected to a stiff grid. In these cases, the power converter will contribute balancing the generation and consumption and only secondly regulation the voltage at the PCC.

On the other hand, when the converter is connected to a weak grid it is preferred to prioritize voltage regulation to make sure that the voltage is kept stable at the PCC. Due to the characteristic of the electrical grid, voltage regulation is an action that can be done only locally so it acquires priority over frequency regulation.

In our use case the HESS is connected to a weak point of the grid, therefore voltage regulation has been prioritized.

To properly analysed the behaviour of the HESS with respect to voltage and frequency regulation two use case have been defined:

- **UC1: frequency response**: the behaviour of the $P(f)$ and $Q(U)$ controllers are analysed during a frequency event.
- **UC2: PQ prioritization**: the same conditions of UC1 are used but in addition the grid is perturbed with a load disturbance happening in the proximity of the HESS in order to observe how the PQ prioritization works in case of extreme conditions.

For both use cases the parametrization of the $Q(U)$ and $P(f)$ controllers are reported in Table **1**.

Q(U) configuration			P(f) configuration		
Parameter	**Value**	**Unit**	**Parameter**	**Value**	**Unit**
$-\boldsymbol{\Delta Umax}$	-0.05	p.u.	$-\Delta fmax$	-0.2	p.u.
$\boldsymbol{\Delta Umax}$	0.05	p.u.	$\Delta fmax$	0.2	p.u.
$-\boldsymbol{\Delta Udb}$	-0.01	p.u.	$-\Delta fdb$	-0.01	p.u.
$\boldsymbol{\Delta Udb}$	0.01	p.u.	Δfdb	0.01	p.u.
$\boldsymbol{-Qmax}$	-1	p.u.	$-Pmax$	-1	p.u.
$\boldsymbol{Qmax}$	1	p.u.	$Pmax$	1	p.u.

Table 1 $\boldsymbol{Q(U)}$ and $\boldsymbol{P(f)}$ parameterization

The frequency profile used for the simulations corresponds to the separation of the Continental Europe power system happened on 8 January 2021. The frequency measured with a resolution of 100ms in the Nort-West area is used. This frequency event is representative of a large frequency deviation.

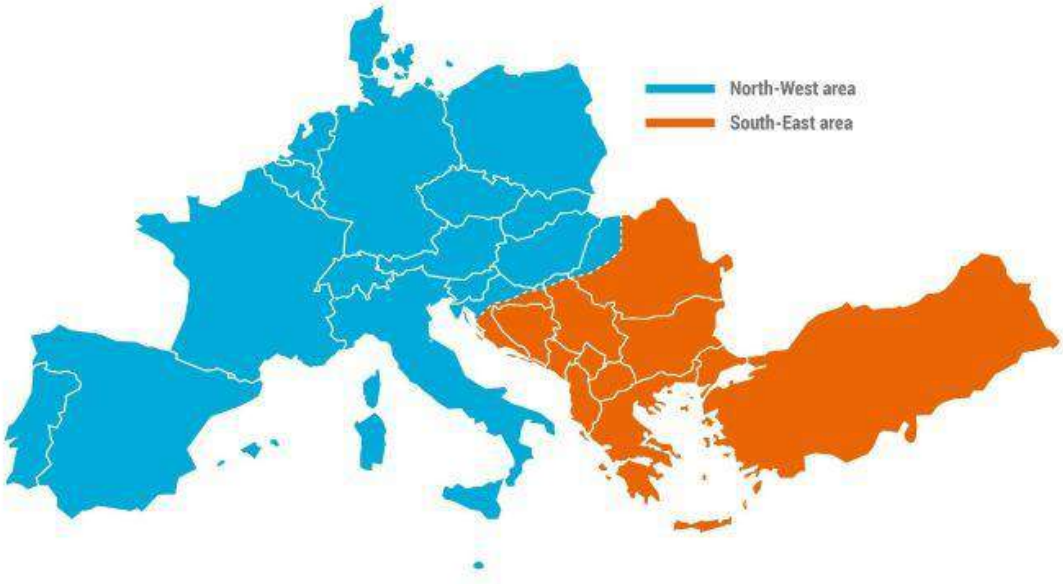

Figure 8 Separation of Continental Europe Synchronous Area on 8 January 2021 [4]

During this event the frequency in the North-West area decreased until $49.74\ Hz$ for a period of 15 seconds and then stabilized at a frequency of $49.84\ Hz$ [4].

4.3.1 UC1 Frequency regulation results

The frequency regulation is achieved thanks to the $P(f)$ controller. For both the HE and HP modules the grid frequency, the voltage deviation and the active and reactive power of the inverters are measured. The results for the HP and HE modules are shown in Figure 9 and Figure 10 respectively. In the plots the deadbands for the $P(f)$ and $Q(U)$ controllers are highlighted by the 2 horizontal black dashed lines. The frequency event starts at $t = 13.5s$.

As soon as the frequency decreases below $-\Delta f_{db}$ the $P(f)$ controller of the HP module starts to inject power into the grid trying to restore the nominal frequency at $50Hz$. When the frequency decreases below $-\Delta f_{max}$ the controller is injecting $Pmax$. The HP module is supposed to provide power only for a short period of time accordingly to the power split algorithm implemented in the EMS module. This can be seen in the bottom plot of Figure 9 where the power in yellow starts to decrease at $t = 25s$ even if the $P(f)$ controller is supposed to provide $Pmax$. As a side effect of the power injection the voltage is slightly increasing, and this is the reason why at $t = 20.5$ the $Q(U)$ controller is providing a small amount of reactive power. Between $t = 20.5s$ and $t = 30s$ the active power is limited because the reactive power is set prior to the active one. It is not possible to see the PQ prioritization in the power plot of Figure 9 because the limitation of the active power is very small. In UC2 the PQ prioritization is analysed more in detail. The power measurements for the HE module in Figure 10 are different compared to the ones of the HP module. When the frequency decreases below $-\Delta f_{db}$ the module starts to inject active power into the grid. Compared to the results of the HP module the active power is ramping up slowly until it reaches $Pmax$ at $t = 25s$. This behaviour is the effect of the EMS power splitting algorithm. The HESS is supposed to utilize the HP module to counteract voltage-frequency deviations that last a short period of time. If the deviation takes longer, the HE module intervenes, and the power output of the HP module is reduced. This can be clearly seen in Figure 11. The reaction time of the HP and HE modules can be set tuning the parameters of the controller implemented in the EMS block.

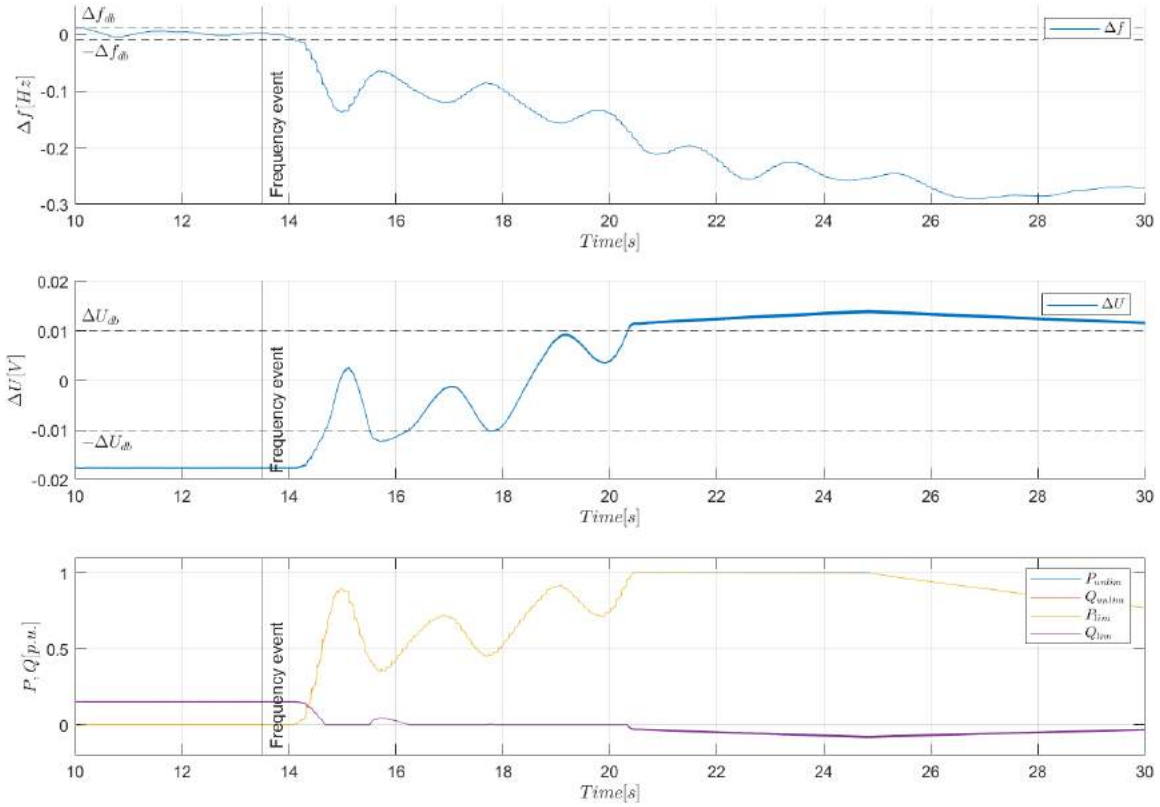

Figure 9 UC1: HP module setpoints

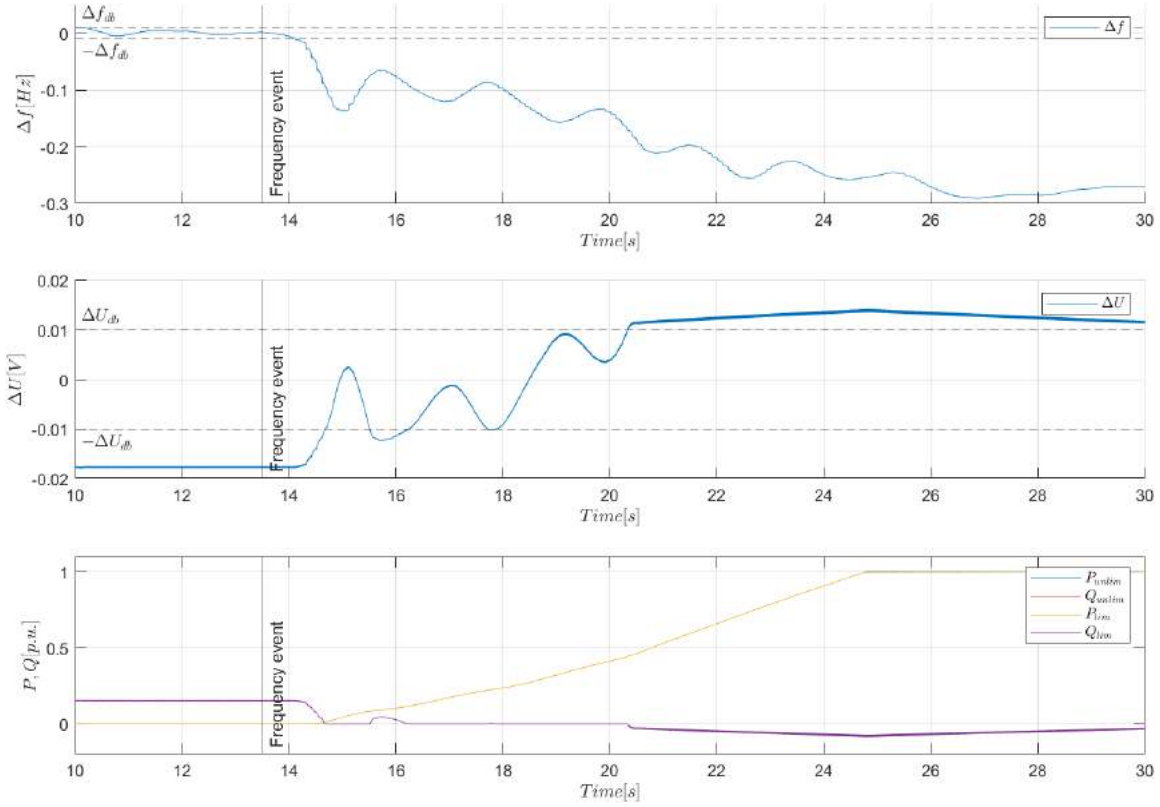

Figure 10 UC1: HE module setpoints

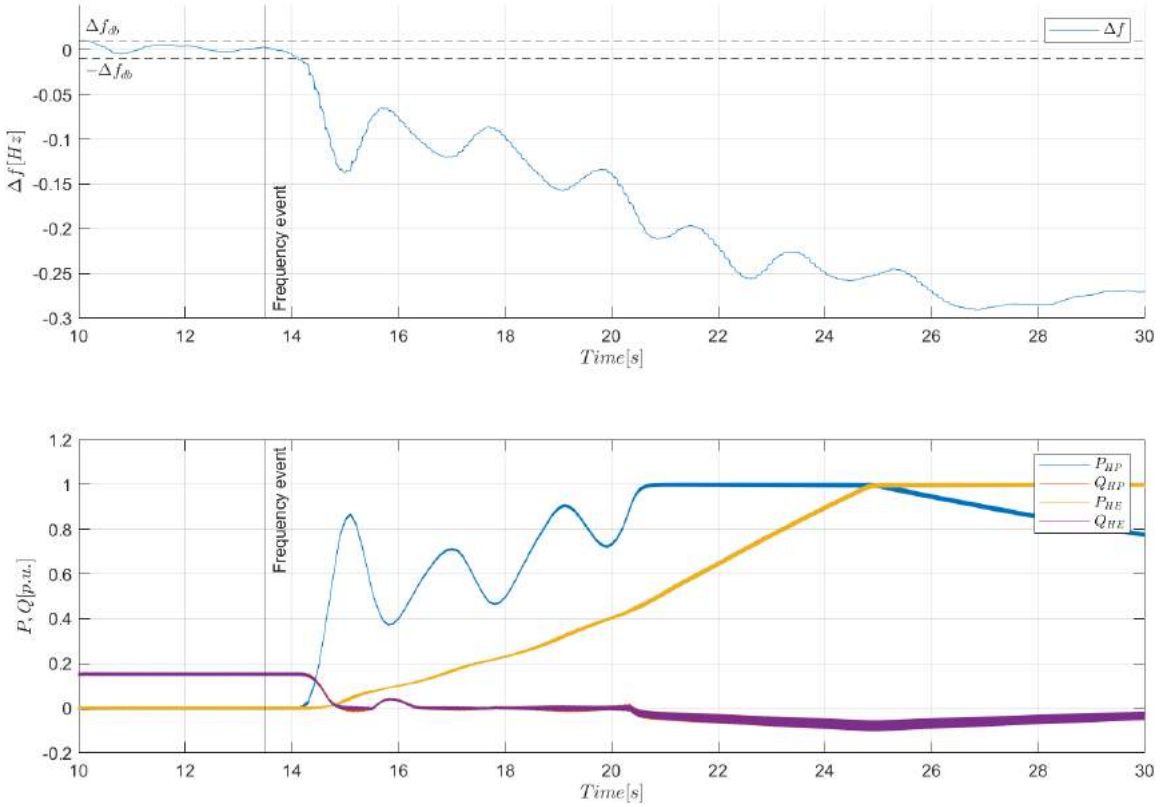

Figure 11 Power output of the HP and HE modules

4.3.2 UC2 PQ prioritization algorithm results

The focus of UC2 is on the PQ prioritization algorithm. The reactive power is set prior to the active power i.e., voltage regulation has priority on frequency regulation. The PQ prioritization is active only when the power limits are reached, this implies that the grid is in such a state for which the $P(f)$ controller is injecting/absorbing a large amount of active power and in the same way the $Q(U)$ controller is demanding reactive power. To achieve these grid conditions the simulation is performed using the frequency profile used in UC1 and additionally a load disturbance is applied at $t = 11s$ as can be seen in Figure 12 and Figure 13. Focusing on the results of the HP module, the load disturbance causes a severe voltage deviation, thus the $Q(U)$ controller as countermeasure injects a large amount of reactive power into the grid in order to stabilize the voltage. At $t = 13.5s$ the frequency event happens, and the $P(f)$ controller starts to inject a large amount of active power. At $t = 14.75s$, $t = 18.68s$ and $t = 20.16s$ the demanded power is then larger than the rated power of the converter and the PQ prioritization block curtails the active power setpoint.

The curtailment of the active power can be clearly seen in Figure 12 looking to the P_{unlim} and P_{lim} signal that corresponds to the power setpoint before the PQ prioritization block and the power setpoint after the PQ prioritization block respectively. A certain ripple can be observed in the power flows, which is conjectured to be caused by the interaction of the grid-following controller, the LCL filter and the grid impedance.

With respect to the HE module results in Figure 12 andFigure 13, it is important to highlight how the $Q(U)$ setpoints are identical to the ones of the HP module. The reactive power is not affecting the dc side of the HESS, therefore the power splitting algorithm is not applied to the reactive power. This assumption is valid considering negligible the power electronic losses. In fact, reactive power is a concept primarily associated with alternating current (AC) systems, and it typically does not apply to direct current (DC) systems. Reactive power arises from the phase difference between voltage and current in AC circuits, leading to the oscillating exchange of energy between the source and the load. In a DC system, however, there is a constant and unidirectional flow of electric charge, with no oscillation or phase shift.

On the contrary the power splitting algorithm is applied to the active power since it affects the DC side. The active power as for UC1 is slowing increasing reaching P_{max} at $t = 25s$. The power limit is reached and $t = 23.6s$ and the PQ prioritization algorithm curtails the active power setpoint.

Finally, Figure 14shows a comparison of the HP and HE inverter power measurements. In conclusion, the joint frequency and voltage support can be easily integrated into a HESS, by splitting the active and reactive setpoints among the different modules.

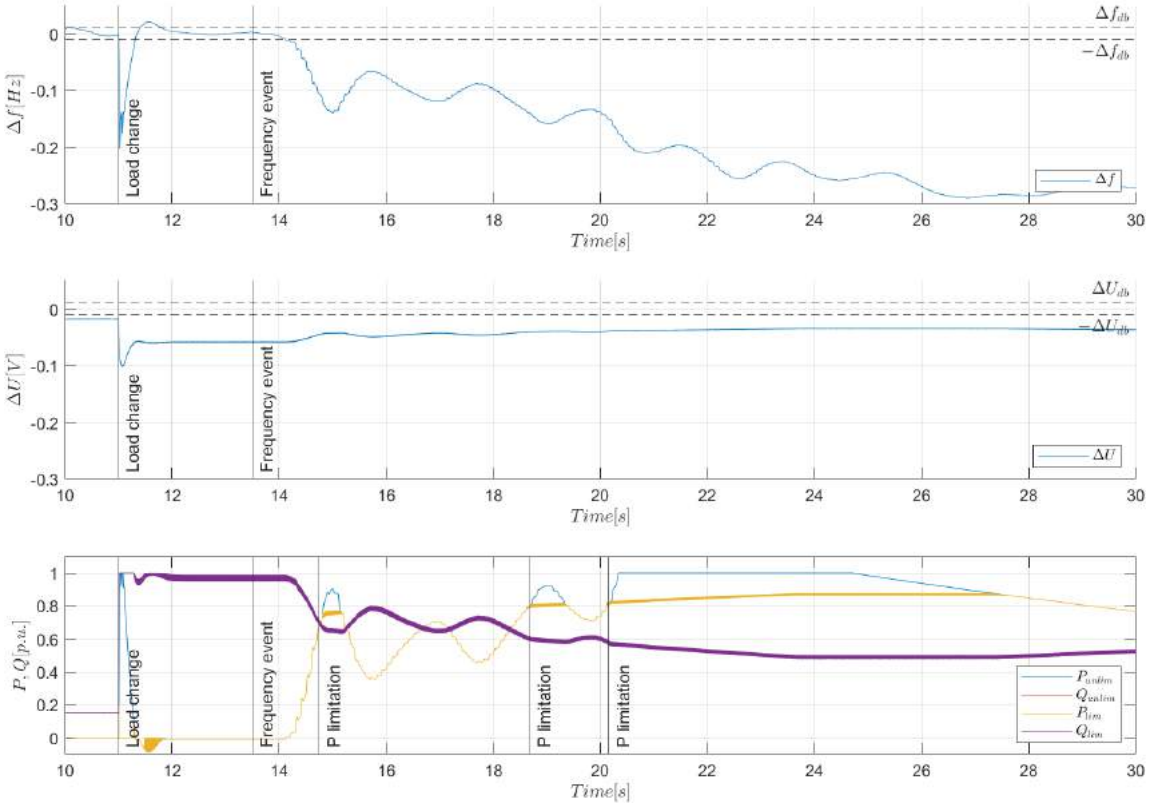

Figure 12 UC2: HP module results

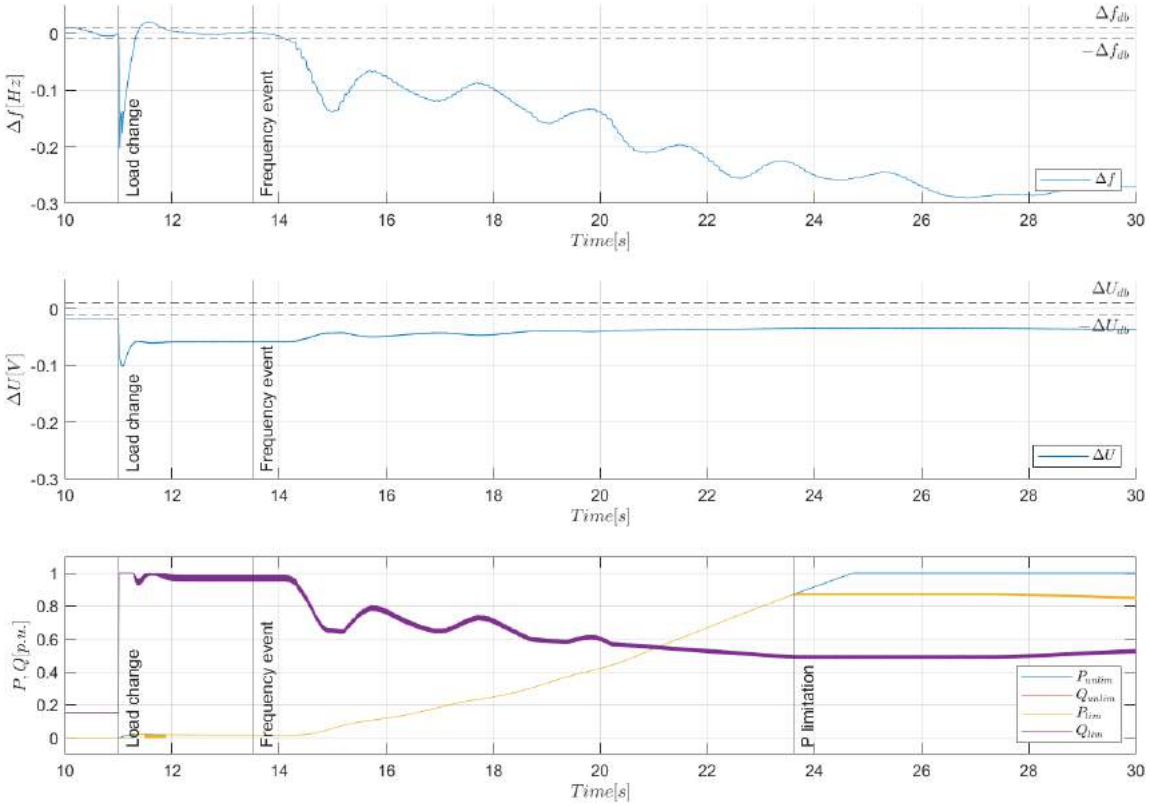

Figure 13 UC2: HE module results

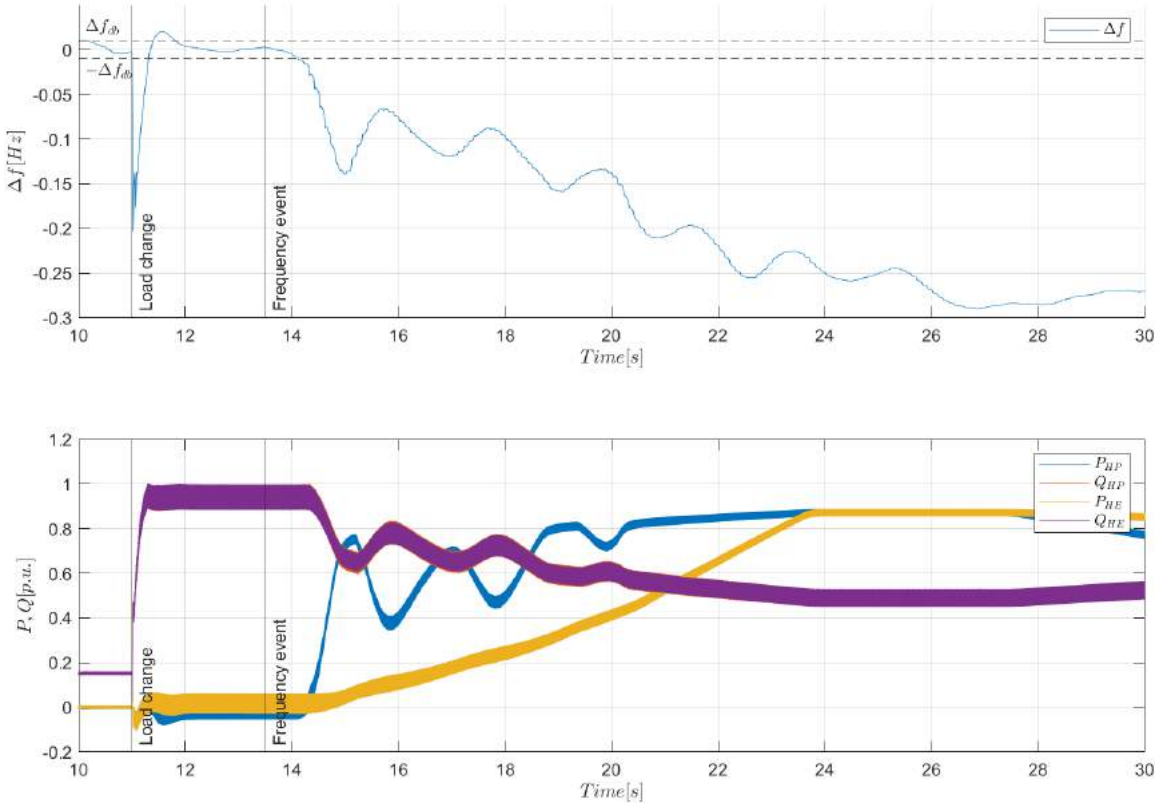

Figure 14 HE and HP modules comparison

5 Peak shaving

Hybrid power plants are emerging as an adequate solution to address emerging power system needs, since they can provide a wide palette of services by leveraging the capabilities of different technologies. At the same time, storage has been recognized as a critical component for the energy transition, in order to deal with the new characteristics of inverter-based generation. In particular, it is becoming clear how a storage system can provide certain grid services, but how to extend these ideas to hybrid storage systems is not fully understood. This chapter extends existing work in this domain to hybrid storage systems and their different dynamics. In particular, we present an optimization-based framework that is able to deal with the typical fast time scales associated with storage, at a reasonable computational cost. The approach relies on piecewise affine linear approximations, that can be integrated into mixed-integer linear programs. An example focused on peak shaving shows the benefits of the proposed approach.

5.1 Implementation of optimization-based grid services in standard battery storage systems

Before tackling the particularities of hybrid storage systems, we review the state-of-the-art strategies for single storage systems. For some of the grid services and applications for storage units, such as market participation or peak-shaving, an optimal solution can be derived, provided that forecast of prices or load demand exist. Therefore, optimization-based algorithms for storage applications have been widely studied in recent years [5], [6], [7]. These techniques typically resort to the use of model predictive control (MPC). Under MPC, an optimizer derives a plan for a certain time horizon, based on some expected forecast for loads, generation and prices. This optimal plan is only used for the next action, and afterwards new measurements are used in order to solve again the optimal problem. Provided that the involved functions are not very complex, the optimization can be solved in a reasonable amount of time. Complexity nonetheless represents a critical point: the considered time step at which the optimization is solved is typically relatively large, sometimes in the order of minutes, in order to minimize the amount of decision variables and thereby reducing the computation time.

While the optimization is typically carried out over a long time horizon and with a large time step, the internal dynamics of a hybrid storage are typically much faster. In standard storage systems, it is typically not necessary to be aware of the internal dynamics, since the energy management system (EMS) does not change setpoints between time steps, given the homogeneity of the components. However, in the case of HESS it is very relevant to keep track of the time-varying setpoints and distinguish between the different modules composing the unit. For instance, it is necessary to model the evolution of the state-of-charge (SoC) of each module, as once a module is depleted it cannot longer be used. Reducing the time step of our optimization to accommodate for the fast dynamics would increase the complexity exponentially, and make the problem not tractable.

Previous works in this area include two-layer setups, where decisions need to be taken at the two time scales. However, in hybrid storage systems, the low level control is typically very fast and therefore not amenable to optimization-based solutions. The work in [8] deals as well with MPC applied to HESS, but no fast dynamics are considered, and the setup is simplified to avoid the inclusion of binary variables. Sequential approaches have also been considered, where first the fast storage unit is controlled, and then the slower unit [9]. It is then treated in a sequential way, rather than simultaneously.

At the same time, most of the existing articles on the matter assume the efficiency of each unit to be constant, unlike in this study. We instead propose a framework to account for the fast internal dynamics of hybrid modules that can be included into an optimization-based solution for storage applications. The proposed solution can deal with nonlinearities, module activation or deactivation, etc. by modelling the inter-step behaviour of each module using piecewise-affine functions, that can be easily integrated into an MPC framework, by means of binary variables.

As an application of this framework, we analyse the case of peak shaving for large customers, with the objective of reducing the demanded peak power over a certain time horizon. Nonetheless, the framework is amenable to other services. In the case of frequency-based services, it may require to consider a stochastic framework since the evolution of the frequency is not predictable [10].

Modelling and Operation of HESS

Hybrid storage are typically controlled by energy management systems that consider the characteristics of each module such as aging, SoC, dynamic behavior, etc. in order to maximize the overall performance of the unit. For simplicity, in this section we illustrate our ideas using a HESS composed of just two modules, a high energy (HE) and a high power (HP) module. In this case, typically fast power changes are realized using HP modules, while a long withstanding power output can be achieved easier using HE modules. The corresponding EMS can split a given setpoint P_{ref} to limit ramp rates in the HE module, and therefore the setpoints for each module can be defined as:

$$P_{HE} = f(P_{ref}), \quad P_{HP} = P_{ref} - P_{HE}$$

Where P_{HE} and P_{HP} define the setpoints of the high energy and high power modules, and f can represent a slew rate limiter, or a low pass filter, among other possible implementations [1], since aging is clearly influenced by high current-rate and high power-rate. As we will see, our framework is not limited to analytical expressions of f, and even machine learning-based EMS can be considered. Moreover, balancing between the two modules may be needed, either passively where the EMS decides for a certain time window to only use one of the modules, or actively, where typically the HE module restores the SoC of the HP module.

Unlike in most storage applications, in this work we also model the power-dependent efficiency of the system. The losses of a battery depend on the current; assuming that the terminal voltage of the inverter is constant, the losses can be written as a function of the apparent power S and the SoC of the battery [11], [12]. In these types of applications, the focus lies on active power and reactive power provision is not considered in the optimization. Therefore, the losses are assumed to exhibit a quadratic behaviour on P whenever the module is in operation, as reported in the literature [11], [13], [14]:

$$P_{loss_i} \approx b_i \, (c_{0_i} + c_1|P_i| + c_2|P_i|^2 + c_3|P_i|^2 g(SoC_i)$$

with $c_0, c_1, c_2, c_3 \geq 0, g(SoC)$ denoting the SoC-dependent battery losses [12] and where the boolean variable b_1 denotes if the *i*-th module is active or not. Publicly available datasheets confirm the validity of this assumption [15]. Temperature dependences could be included as well into an MPC approach, as explained afterwards. Moreover, modules can be deactivated by the EMS during idle times, in order to reduce the losses of the system.

5.2 Optimization-based services

In order to maximize the revenues generated by a storage system, an optimal operation is sought. Moreover, ideally several compatible services can be addressed by a single optimizer. For these solutions, where typically forecast of generation, load profiles and prices are available, model predictive control is a de facto standard, clearly outperforming rule-based approaches. Under MPC, an optimizer computes a plan, that happens to be optimal provided by that there is no uncertainty, e.g., the load profiles match the expected forecast. Since uncertainties are expected, the optimization problem is recomputed at each time step. As the problem needs to be solved continuously, complexity is an issue, and the optimizer should find a solution (hopefully a global extremum) in one time step. The presence of binary variables, needed for instance to differentiate between a module charging or discharging, can increase the complexity, which is exponential in the number of binary variables in the worst case. Nonetheless, mixed integer linear programming problems (MILPs) can be solved efficiently using commercial solvers, without the need for heuristics. On the other hand, nonlinear constraints and cost functions, especially non-convex, are extremely discouraged, due to their high computational cost. Therefore, ideally we would seek for linear constraints, whenever possible.

As an example of these applications, we first focus on peak shaving, where the maximum power demanded from the grid is to be reduced through the use of batteries. For certain customers, billing occurs not only as a function of energy charges (in €/kWh) but also on peak demand (in €/kW), the latter corresponding to the highest metered consumption peak over a certain time window. By employing a battery, this maximum peak can be reduced by simply discharging the battery at the time instants of high consumption, and charging the battery at moments of low demand. It is important to emphasize that this peak power corresponds to the highest power averaged over a certain interval, given that power meters provide just an *n*-minute resolution (where *n* can be in the order of 15, e.g., in Germany). A simple version of the optimization problem using a slack variable ϵ is typically formulated as:

$$\min_{P_{ref(t)},\epsilon} J(L(t),\epsilon)$$

s.t.

$$Soc(k+1) = SoC(k) + \int_0^h \big(P(t) + P^{loss}(t)\big)dt$$
$$L(t) = l(t) + P(t)$$
$$L(t) \leq L_{peak} + \epsilon$$
$$SoC_i \in [SoC_{min}, SoC_{max}],\ P_i \in [P_{min}, P_{max}]$$

with $P^{loss}(t)$ as in defined before, where $l(t)$ is the load demand, $P(t)$ power from battery, $L(t)$ power demand from the grid and *J* represents the cost to be minimized, that can depend on the load and the maximum peak, for instance. As in standard peak-shaving applications, a variable ϵ is introduced as a slack decision variable, in order to indirectly minimize the maximum peak while avoiding feasibility issues. It represents the maximum power demand from the grid in the optimization horizon above a certain admissible value L_{peak}. Notice that L_{peak} does not need to be constant, and can be adapted to account for time-varying tariffs, or to consider previous peak demands during the billing period. Indeed, if a certain peak demand has been already reached in the previous days, there is no incentive to use the battery to reduce the demand below that peak, as there is no cost savings to be made. Positive values of P correspond to the battery being charged. The cost function may also penalize the usage of the battery, or consider the cost of buying energy from the grid, as we will see later on.

5.3 Model Predictive Control for hybrid storage systems

For single storage systems, the evolution of the SoC is easy to describe as P is constant over the whole time interval of duration h and the integral in the SoC equality in the previous section can be easily replaced by $(P(t) + P^{loss}(t)) \cdot h$, making the problem a standard MILP. Moreover, a single SoC representing all modules (should the design be modular) is enough, since the internal balancing may be ignored for the high-level optimization.

Should a low-level controller modify this reference power during the interval, either to minimize aging or to reduce losses, the expression can become convoluted. Even in the case where the evolution of *P* and the integral can be computed analytically, it may lead to a very convoluted expression that will complicate extremely the optimization problem. In this work we propose a model for the evolution of the SoCs of each module using piecewise affine linear functions, that can be integrated into a MILP formulation.

5.3.1 SoC modelling of HESS fast dynamics

As mentioned before, in HESS a local controller or EMS might vary setpoints and activate or deactivate modules continuously to achieve certain goals. The time scale of this controller is expected to be at least one order of magnitude faster than an optimizer, which time horizon is more in the order of days or weeks. However, for the optimizer it is only relevant how the SoC of each module evolves at each time step k.

This SoC evolution might be highly nonlinear, including saturations, and may cover the deactivation of certain modules, which mathematically corresponds to a sudden change in a variable. In the most general setting, the behaviour can be described by a hybrid dynamic system (see e.g. [16]). In many cases there is no explicit solution for these systems, but we are only interested in the change in SoC after one time step. Hence a model that describes SoC(k+1) as a function of SoC(k), P_{ref} and the characteristics of each module is to be derived, which can be computed numerically. However, in order to integrate it into an MPC setup, an analytical and relatively simple expression is needed. While nonlinear functions are hard to deal with, binary variables and linear constraints are easier to integrate. Therefore, we will resort to piecewise affine (PWA) models, that can approximate nonlinear and/or discontinuous dynamics arbitrarily well. Each linear model describes the behaviour of the storage in a different operating region, which can be all easily computed offline, thereby reducing the online complexity.

For instance, using PWA models the evolution of the SoC of a module *i* at time step *k* using 2 regions can be described by means of two binary variables δ_1 and δ_2:

$$[\delta_1 = 1] \leftrightarrow SoC_i(k+1) = \alpha_1 SoC_i(k) + \beta_1 P_{ref}(k) + \gamma_1 P_{ref}(k+1)$$

$$[\delta_2 = 1] \leftrightarrow SoC_i(k+1) = \alpha_2 SoC_i(k) + \beta_2 P_{ref}(k) + \gamma_2 P_{ref}(k+1)$$

where we have imposed a linear dependence for the SoC on the next and the actual value of P_{ref}. The number of regions (and therefore the number of binary variables δ_i) denote the complexity of our affine model, and α_j, β_j and γ_j are scalars defining the linear model in each region. Increasing the number of regions increases the accuracy of the SoC evolution, at the cost of more binary variables and therefore a more complex optimization problem. The computation of the coefficients α_j, β_j and γ_j can be done using many different methods, analytically in certain cases, or using machine learning tools such as logistic model trees [17], segmented regression or discrete optimization methods [18], among others. Notice that the identification of a linear model in a region is straightforward, but identifying the number and shape of the regions is non-convex and with multiple local minima. The modes of operation δ_1, δ_2 are implicitly defined by hyperplanes of the form $Mx \leq N$, with x being the decision variables of our problem (P_{ref}, SoC, etc.) and where M and N are rectangular matrices of appropriate dimensions. In our case, identifying the switching manifolds or hyperplanes can be guessed, as there are natural selections on how to choose the breaking points, for instance, when the power changes signs or when we observe impulses or nonlinearities (e.g., saturations). An example of this PWA approximation with 3 regions for a typical battery is shown in Figure 15.

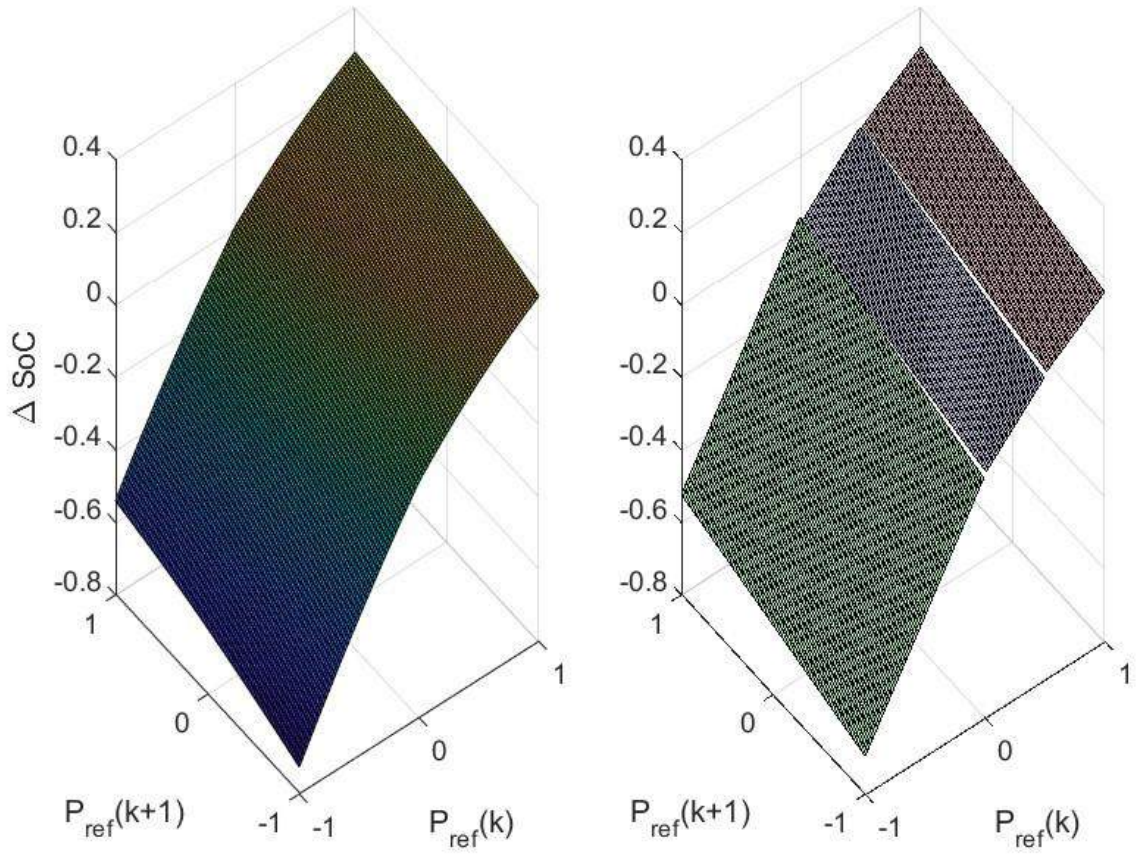

Figure 15 - SoC evolution as a function of the reference power at k and k + 1 (left) and its piecewise linear approximation using regions (right).

Considering extra dependencies, such as the aforementioned SoC dependence for the battery losses, does not directly represent an increase in the computational cost, but it might make the PWA identification more cumbersome, when it comes to identifying the operating regions. Piecewise linear models can be rewritten into mixed logical dynamical systems, which are easy to integrate into a MILP formulation. For the sake of brevity we do not detail the required steps since it has been widely reported in the literature, see [19], [7] for some examples.

Time-step grouping could also be partially used to merge some intermediate time steps and thereby reducing the complexity. Nonetheless, the presented approach, building maps defining the evolution of the SoC at the time steps of interest for the optimization, avoids the high resolution needed for the fast low-level dynamics. Even better, convex piecewise affine cost functions can sometimes be represented without introducing binary variables.

5.3.2 Including balancing

Given the hybrid nature of the storage, it is expected that balancing between the units is needed, given the different ratings and different power setpoints used for each unit. Balancing can happen in different ways, but typically it gets triggered based on the difference of the SoCs. In other words, as soon as $|SoC_{HE} - SoC_{HP}| > Threshold$, then a certain amount of power is transferred between modules. In general, when it comes to balancing, there are two decisions to make: when to balance, and how to do it (that is, how much power from one module to the other module). These two decisions are typically taken based on rules or fixing the amount of power dedicated to balancing. Instead, here we take advantage of the optimization framework to decide when and how balancing should occur.

Interestingly, it is easier to let the optimizer choose at each time step how much balancing is needed, since in this case it can be treated as a continuous variable. On the other hand, if the balancing power is fixed and the optimizer only chooses when to enforce it, a binary decision variable needs to be included. Moreover, in the latter different PWA models would have to be computed to consider 3 independent cases: no balancing, the high-energy module recharging the high-power module, and the opposite case. In the case of treating balancing power as a continuous variable, a single model can be derived, where the evolution of the SoC needs to become now a function of P_{ref} and P_{bal}.

Alternatively, an intermediate solution is also possible: a given rule can be fixed beforehand, while the parameters of said rule (i.e., the variable *Threshold*) are chosen by the optimizer. This can be seen as a similar approach to parametrized MPC [9], and reduces substantially the amount of binary variables since a single parameter is needed for the whole time horizon, at the cost of having less flexibility.

Notice that the balancing power needs to be filtered as well, in order to maximize the lifetime of the HE module. Therefore, the original dynamics are modified adding a simple balancing power setpoint:

$$P_{HE} = f\left(P_{ref} + P_{bal}\right), \quad P_{HP} = P_{ref} - P_{HE}$$

In this way, regardless of the balancing power, it is guaranteed by design that the overall unit satisfies the setpoint for active power. We can then compute the PWA models for the SoCs of the HE and the HP module using the knowledge of the power splitting carried out by the EMS and defined by the function *f*. In the case of the HE module, looking at the expression for balancing, we expect the *SoC* at instant *k+1* to be influenced by the sum of P_{ref} and P_{bal}, at time steps *k* and *k+1*. Hence, we fix the expression for SoC_{HE} as:

$$SoC_{HE}(k+1) = SoC_{HE}(k) + (\alpha_{HE}\left(P_{ref}(k) + P_{bal}(k)\right) + \beta_{HE}\left(P_{ref}(k+1) + P_{bal}(k+1)\right) + \gamma_{HE}$$

A more generic expression can also be considered, including dependencies on all possible inputs, at the expense of a more convoluted definition and computation of the ideal splitting manifolds. In the case of the HP module, we expect a similar expression but with an added linear term in $P_{ref}(k+1)$:

$$SoC_{HP}(k+1) = SoC_{HP}(k) + (\alpha_{HP}\left(P_{ref}(k) + P_{bal}(k)\right) + \beta_{HP}\left(P_{ref}(k+1) + P_{bal}(k+1)\right) + \gamma_{HP} + \xi P_{ref}(k+1)$$

Moreover, the optimizer should not include any combination of $P_{ref}(k)$ and $P_{bal}(k)$ that leads to saturation. This saturation could be modelled into the PWA model but at the expense of a more nonlinear map, and therefore more operating modes for our PWA model. Instead, this saturation can easily be included as a linear constraint of the type:

$$1 \geq P_{ref}(k) + P_{bal}(k) \geq -1$$

With these PWA models for the evolution of the SoC and the aforementioned constraint, standard optimization problems as in Section 5.2 can be solved as usual, using a MILP formulation.

5.4 An illustrative example

To illustrate and compile the ideas in this section, we consider the case of a hybrid storage composed of a HE and a HP module, with 50kW/50kWh and 50kW/10kWh as rated power and energy respectively. The inverter characteristics have been taken from publicly available datasheets [15], in particular we consider the losses characteristics from two inverters from Sunways and Satcon. For simplicity, we assume a constant internal resistance for the battery, therefore the SoC dependence in the battery losses is not implemented. The main objective for the HESS is to reduce the peak power demanded by an industrial site, with a certain load profile defined in [20], and scaled to reach a peak power of 250kW. The SoC of the storage units are limited to operate between 0.2 and 0.8, to maximize their lifetime. Moreover, we consider variable costs for the energy absorbed from the grid, with prices taken from the ENTSO-E Transparency platform [21]. In particular, we have selected the prices on 06/03/2023 in Austria. In order to minimize aging from the battery, the energy needed to balance the modules is also penalized (which is already implicitly penalized by means of the losses, but to a lesser degree). The internal EMS implements a ramp rate limitation for the HE module, with the HP module providing the remaining power. For didactic purposes, and to clearly differentiate the responses from the HE and the HP module, we strictly limit the dynamic response of the HE module, so that the setpoint is reached after 3 minutes. Notice that this dynamic response is highly nonlinear and allows us to showcase the potential of the proposed approach. Slow ramp rates are expected from energy storage units relying on fuel cells, compressed air energy storage or pumped hydro as energy source. The coefficients for the PWA approximations are reported in Table 2 and Table 3. We have respectively defined 2 and 3 operating regions for the HE and the HP module, in order to keep complexity low while still preserving accuracy.

Region	α_{HE}	β_{HE}	γ_{HE}	Condition
A	1,48	7,65	-0,07	$P_{ref}(k) + P_{bal}(k) \geq 0$
B	1,59	9,28	-0,07	$P_{ref}(k) + P_{bal}(k) < 0$

Table 2 - Piecewise coefficients for the HE module

Region	α_{HE}	β_{HE}	γ_{HE}	ξ_{HE}	Condition
A	-1,45	-7,94	-0,14	9,39	$P_{ref}(k+1) \in [-0.5,0.5]$
B	-1,63	-9,11	-0,07	10,74	$P_{ref}(k+1) < -0,5$
C	-1,43	-7,79	-0,05	9,22	$P_{ref}(k+1) > 0,5$

Table 3 - Piecewise coefficients for the HP module

The optimization is carried out for a whole day, with a time interval of 15 minutes, which is typically the time window considered for billing purposes in some European countries. The cost for peak shaving is assumed to be

4.411€/kW/month, taken from the tariffs for medium voltage customers in the federal state of Salzburg in 2021. Perfect load forecast has been assumed for simplicity. The considered cost functions for the daily optimization is aiming at minimizing the maximum peak, while penalizing the usage of battery and balancing power:

$$\min_{P_{ref(t)},\epsilon} \sum_{k=0}^{N} a_0\epsilon + a_1|P_{ref}| + a_2|P_{bal}| + a_3(k)L(k)$$

with *N=96* being the number of steps over a day (given the 15-minute time window), a_0 the cost associated to the peak power, a_1, a_2 design parameters in order to penalize the overuse of the battery and a_3 the prices in the intraday market. Quadratic expressions (instead of linear) can be used for each term, which allows us to better define the objectives at the cost of a higher computational complexity. For instance, quadratic expressions for ϵ lead to smaller peak power values. At the same time, given that the cost is only linked to the maximum peak power over a month, it is interesting to adjust L_{peak} based on previous daily peak powers, in order to avoid overusing the battery to reduce irrelevant peaks.

We use CPLEX to solve the MILP problem, and in particular the branch-and-cut algorithm, which leads to a global optimum [22]. We have limited the computation time of the solver to 1 minute, since in our experience in this time it is able to achieve a very low optimality gap, and there is no need to let the solver finish all its computations and reach the global optimum. In order to speed the branch-and-bound, it is possible to use a warm start using a feasible initialization [22]. A final cost is typically needed to avoid stability issues.

Figure 16 shows the evolution of the original demand l, the shaved response L and the power delivered by the battery P_{ref}. During the off-peak times (early in the morning and somewhere in the evening), the battery can be charged and discharged in order to reduce the overall energy costs, according to the intraday prices, while keeping the SoC within the admissible range, with the only constraint of being charged before the maximum peak arrives. Hence the usage of the battery is highly correlated with the energy prices. We would like to emphasize that the shown prices profile is representative and common, hence this type of behaviour is highly expected for the battery usage.

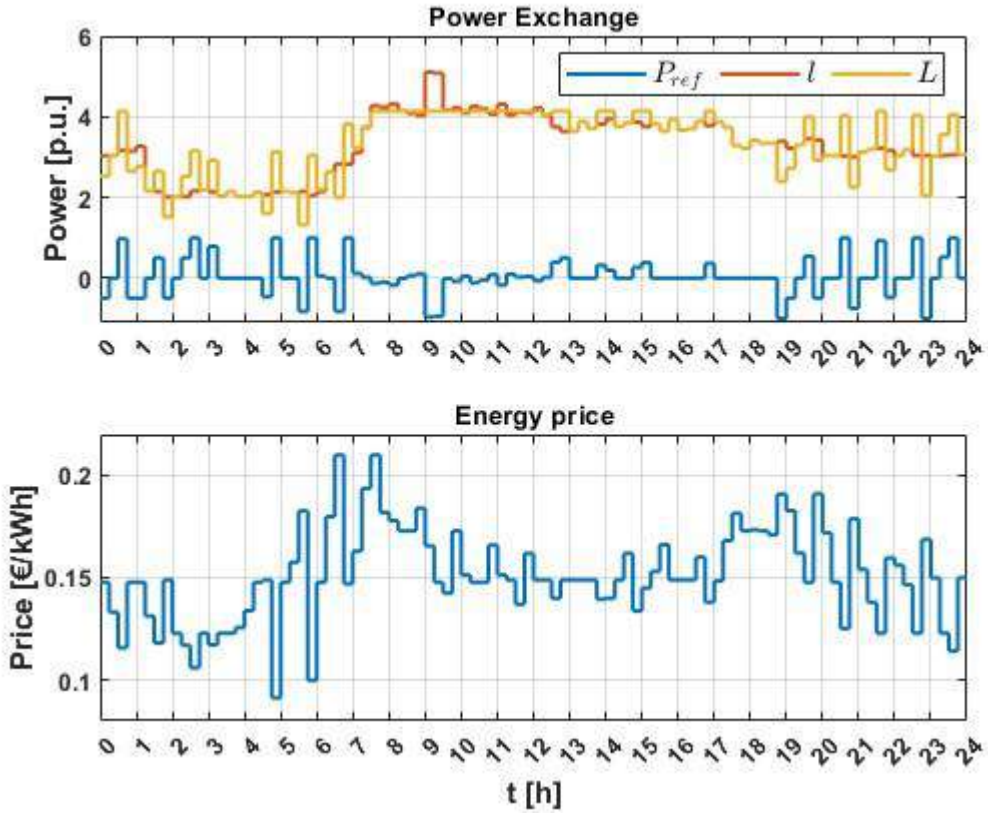

Figure 16 - Evolution of power demand with and without the battery

Figure 17 shows the evolution of the SoC and the corresponding balancing power. Both modules are nearly fully charged before the peak power arrives, including balancing for the HP module. It can also be observed the self-discharging of each battery, which is much more relevant in the case of the HP module. During the off-peak times, the power setpoints for the battery are dictated by the prices, in order to buy whenever energy is cheap and reduce the consumption during the time slots with high prices. A last spike in power is seen at the end of the day in order to satisfy the final constraint on SoC, and caused by the low prices. While not shown in the figure for simplicity, the power for the HE and HP modules are just the filtered version of the sum of the reference and the balancing power.

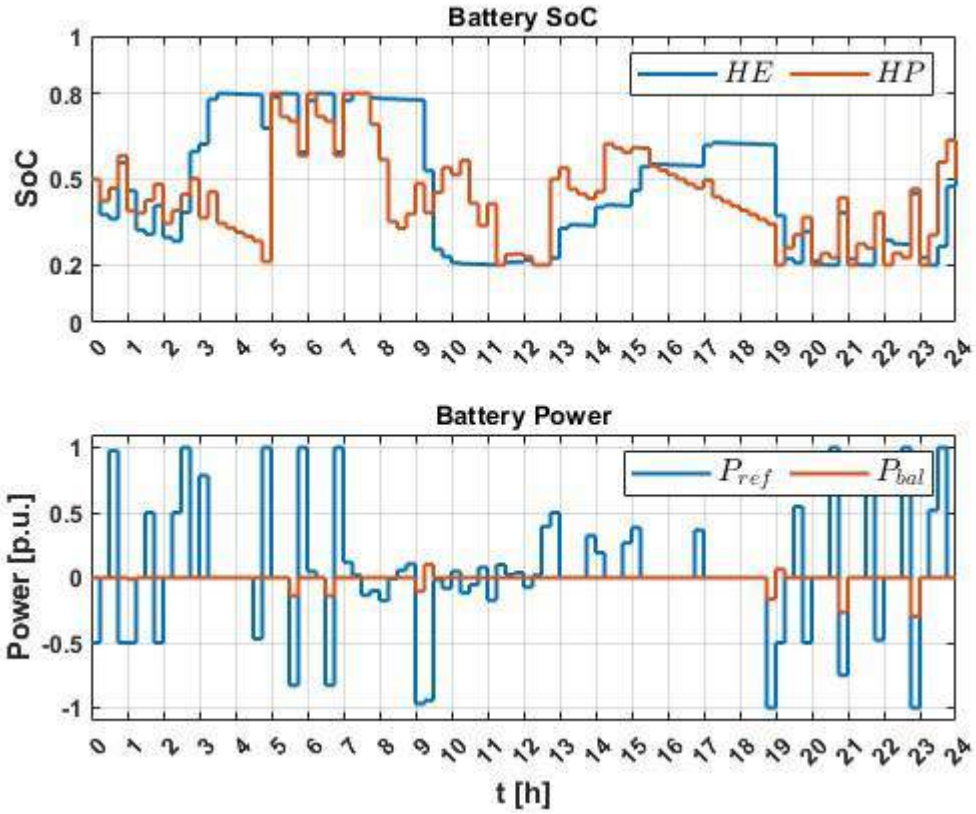

Figure 17 - Evolution of the SoC for both modules and corresponding power reference and balancing power

Finally, we evaluate how complexity scales with respect to the number of regions for the PWA modelling. Table 4 reports the optimality gap after letting the optimizer run for 5 minutes, for 2, 3 and 4 regions for the HE module, where the dependence on binary variables can be clearly identified. Therefore, there is a clear advantage on identifying the minimum number of regions for a certain desired accuracy, as the optimality gap does not increase linearly on the number of regions. On the other hand, given that typical optimization steps are around 15 minutes, there is enough headroom to increase the complexity of the optimization, either by adding more regions, extending the time horizon, or modelling the SoC dependence for the battery losses.

Number of regions	Optimality Gap
2	0,03%
3	0,09%
4	0,24%

Table 4 - Optimality gap for different number of regions of the HE module

In conclusion, the presence of different modules in these type of optimization-based applications raises a hurdle in terms of complexity, compared to standard, uniform storage systems. To address this, we have proposed a strategy for the optimal operation of a HESS that keeps the complexity manageable.

6 Impact of primary frequency control on HESS ageing

Storage systems have been widely used to provide frequency regulation. For a HESS designed to provide primary frequency control, the ageing of the system is fundamental to study the economical profitability. In this third use case we study how the ageing after a year of use and how some control parameters affect ageing.

6.1 Primary frequency control

As already mentioned in Section 4, the grid frequency must be controlled to guarantee the correct operation of the grid. To this regard, primary frequency control regulates the HESS output power to balance the grid and counteract frequency deviations after a disturbance in the grid. This frequency response typically consists of a proportional droop control plus a certain deadband, with the droop curve control in Figure 6. This control gives the power reference to the energy management system (EMS).

The EMS has been developed in deliverable D4.2 of the iStormy project [3]. It relies on power splitting controller, that separates the power reference into the reference to the High-Energy (HE) and High-Power (HP) modules. It is implemented as a high frequency filter in a such a way the fast reference evolution is given to the HP module, meanwhile the slow dynamics is given to the HE module, improving the lifetime of the HESS. Moreover, an SoC balancing regulator is implemented, which balance the SoCs in the two modules to be the same. This is needed to ensure both modules can provide the service without exceeding the SoC constraints. Two options are considered, a filter-based approach and a PI-based approach. The details are reported in deliverable D4.2 [3].

However, a restoration strategy for the SoC of each module has not been considered in the previous tasks, but it is considered in this use case to have the most realistic scenario for the ageing study and reported in next sections.

6.2 Grid codes: CE and Nordic case

Every entity providing primary frequency control must comply with the current grid codes. This is relevant for the adopted SoC restoration strategy. Two EU regulations are considered in the scope of this task, the Austrian and the Nordic one. While the Nordic regulation provides clear requirements how to develop a suitable restoration strategy, the Austrian one gives only some generic constraints, that in our opinion does not uniquely define a valid restoring strategy. Both regulations are explained in the next sections and the corresponding restoration strategy is illustrated.

6.2.1 CE regulation

European Commission Regulation EU 2017/1485 [23] requires that entities participating in the FCR service can provide fully the committed FCR power, continuously for 15 minutes in both directions. Therefore, the SoC of HESS must be regulated inside some bounds so that when the full activation is required, it has enough capacity to absorb/release energy.

The droop curve is the same as reported in Figure 6 with $\Delta f_{db} = 0.01\ Hz$ and $\Delta f_{max} = 0.2\ Hz$.

When the full power is required, the HE module provides its maximum power output, i.e., 32 kW. The power provided by the HP module is scaled on the capacity with respect the HE module, so that after 15 min both modules have the same SoC. Therefore, the maximum FCR power provide by the HP module is 30 kW and the total HESS FCR committed power is $C_{FCR} = 62\ kW$. Considering the SoC operation limit to be 90% and 10% (given that very low and very high values of SoC can really impact the lifetime of the battery), the SoC bounds are computed as follows:

$$\overline{SoC} = 0.9 - \frac{C_{FCR} \cdot 0.25h}{E} - 0.02 = 0.73$$

$$\underline{SoC} = 0.1 + \frac{C_{FCR} \cdot 0.25h}{E} + 0.02 = 0.27$$

A 2% margin is considered in the SoC computation to take in account losses.

To guarantee the SoC to be within the given bounds, a restoration strategy is needed. Since the HE module commits the full power in FCR service, only the HP module can use its available power for restoration purposes. The restoration starts when the HESS reaches one of the bounds and it stops when the nominal SoC (50%) is reached. A time delay of 15 minutes is considered between the activation/deactivation and the flow of power, to take in account the delay to buy/sell energy in the intra-day market.

The power restoration rate is a design choice and different values are compared, evaluating the ability to maintain the SoC in the given bounds and the ageing produced.

6.2.2 Nordic Synchronous Area regulation

In the Nordic Synchronous Area regulation [24], primary frequency control is called Frequency Containment Reserve (FCR) and its divide in three different products:

- FCR-N, in the range 49.9 – 50.1 Hz
- FCR-D upwards, in the range 49.9 – 49.5 Hz
- FCR-D downwards, in the range 50.1 – 50.5 Hz

These are complementary products, one for each different frequency band. While FCR-N is a symmetrical service, and it is designed for small variation around the nominal frequency, FCR-D is asymmetrical and is divided in upwards and downwards response, and it is designed for large frequency deviations in one or the other direction.

For this study, we consider only the provision of FCR-N. The droop curve is the same as reported in Figure 6 with $\Delta f_{db} = 0.01\ Hz$ and $\Delta f_{max} = 0.1\ Hz$.

The regulation defines the so-called Limited Energy Reservoir (LER) entities, which are entities that cannot provide the maximum FCR committed power for more than two hours. This is most of the time the case of BESS and HESS.

The required power and energy of the system are reported in Table 5, where C_{FCR-N} is the committed FCR-N power. For our HESS system, composed by the two modules, C_{FCR-N} is selected to satisfy the requirement and it is chosen as

$$C_{FCR-N} = \min\left(\frac{E_{HP} + E_{HE}}{2 \cdot 1h}, \frac{P_{HP} + P_{HE}}{1.34}\right) = 52.5\ kW$$

Since FCR-N is a symmetric service, the HESS is continuously charging and discharging any time the frequency is outside the deadband. Assuming the frequency deviation distribution has mean zero, the HESS SoC mean will be the nominal value as well. Nevertheless, periods where the frequency deviation is mostly in only one direction (see Figure 19) can happen and then we may hit the SoC operation limits, and the service cannot be provided anymore. Therefore, a restoration strategy is needed and illustrated in the regulations. It defines a Normal State Energy Management (NEM) mode and an Alert State Energy Management (AEM) mode. In Figure 18 the NEM and AEM activation area are shown.

Where NEM is enabled, a restoration process is active and the power setpoint is changed accordingly to:

$$P_{tot} = P_{FCR-N} + P_{NEM} = P_{FCR-N} + 0.34 \cdot C_{FCR-N} \cdot NEM_{current}$$

where P_{FCR-N} is the FCR-N reference, C_{FCR-N} is the maximum committed FCR-N power and $NEM_{current}$ is the rolling mean over the last 5 minutes of $NEM_{Allowed}$, which takes value +1 or -1 if NEM is active in the discharging and charging mode respectively. Therefore, in 5 minutes, the full restoration capability is active. In this way, the rate of which the SoC is moving to its limit is decreased or the direction is reversed bringing the SoC to the nominal value. Therefore, according to the regulation, the HESS needs to reserve 34% of its power for restoration that cannot be used for other purpose.

The NEM may not be enough to guarantee the correct operation of the LER, meaning the SoC can still hit the operation limit. The AEM is designed to avoid this rare but critical situation. When it is active, the reference frequency used to compute, according to the droop curve, the FCR-N power setpoint (namely 50Hz) is changed to be the rolling mean over 5 minutes of the actual measured frequency. Assuming the frequency constant during the five minutes windows, the frequency reference becomes equal to the actual frequency and therefore the power output will be zero, and thereby the SoC remains constant and avoids entering the emergency limits (e.g., SoC below 10% or above 90%). We refer the interested reader to [24] where all details are provided.

	Formula	Value
Required Power Upwards/Downwards [kW]	$\mathbf{1.34 \cdot C_{FCR-N}}$	**70.35**

Required Energy Upwards/Downwards [kWh]	$\mathbf{1h \cdot C_{FCR-N}}$	**52.5**

Table 5: Requirements for LER according to Nordic Synchronous Area regulation

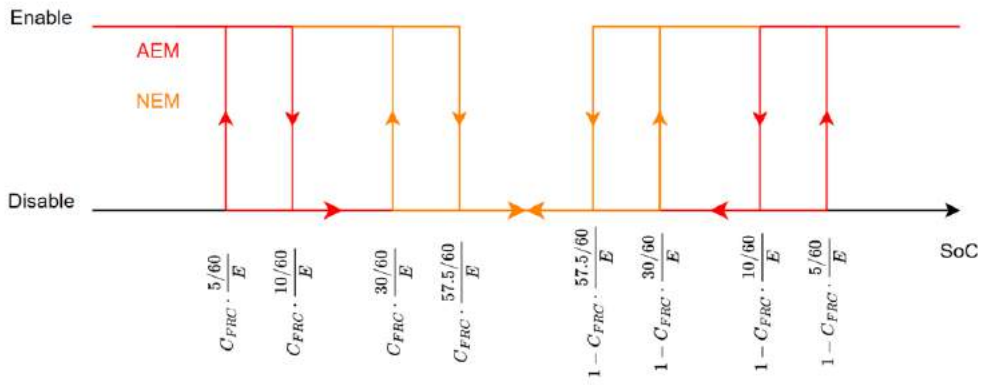

Figure 18: NEM and AEM activation area (E is the total HESS energy)

6.3 Degradation model and implementation

To get a reliable result, the complete battery model developed by one of the iStormy partners (CEA) has been used. However, the ageing model has been built starting from accelerated degradation experiment and using a limited amount of test conditions, therefore the results may be slightly different from a real usage of the HESS. The controller developed in [3] has been added to the model together with the restoration functionally illustrated above for the two cases. In the simulations the initial State-of-Health (SoH) is set to 99% since the degradation model is inaccurate for SoH closed to 100%.

6.3.1 Frequency measurements

To simulate the real operation of the HESS system proving FCR-N, real frequency measurements have been used as input to the system. The time resolution is 0.1s and a year recording is considered, from May 2020 to May 2021. A ten-hour sequence is shown in Figure 19 as an illustrative example. One can also observe, in this specific example, how the frequency is dominated by values above the frequency reference of 50Hz continuously for about ten hours. Therefore, the HESS is charged and without a restoration strategy, it can exceed its capacity limits.

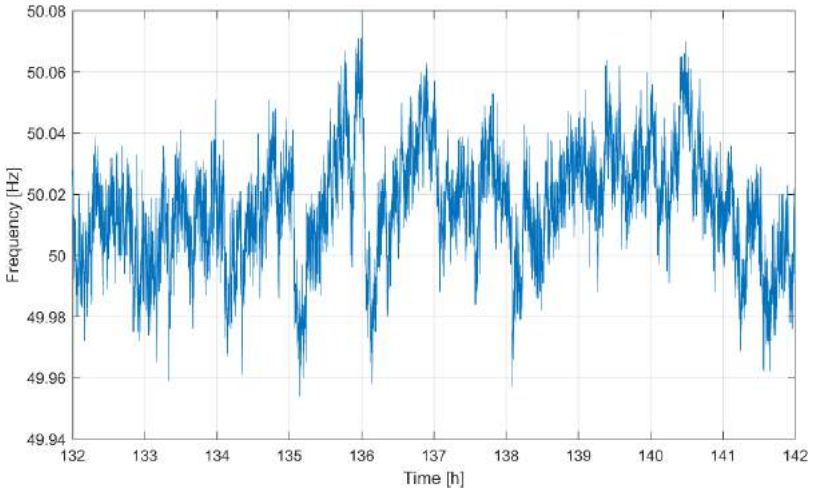

Figure 19: Ten hours frequency measurements sample

6.4 Results

The full system has been simulated for a full year in Simulink and the batteries degradation is computed for different control strategies and compared. We compare the type of balancing controller, filter-based versus PI-based, the cut-off frequency of the power splitting high-pass filter and the restoration power in the CE and Nordic areas. Meanwhile, the PI gains has been tuned in [3] and are fixed.

6.4.1 CE

Considering the CE case, in Figure 20 is shown the SoC behaviour for different values of the restoration power rate. The horizontal black lines represent the restoration activation threshold. For small values (i.e., 2.5 kW and 5.0 kW) there are excursion outside of the safety bounds, meanwhile for larger values they are more contained. From the regulation point of view this may still be acceptable since the service must provide for at least 15 minutes and then, if the operation limits are reached, the service can be stopped. A tighter band would be specified to guarantee staying within bounds, at the expense of more restoration actions.

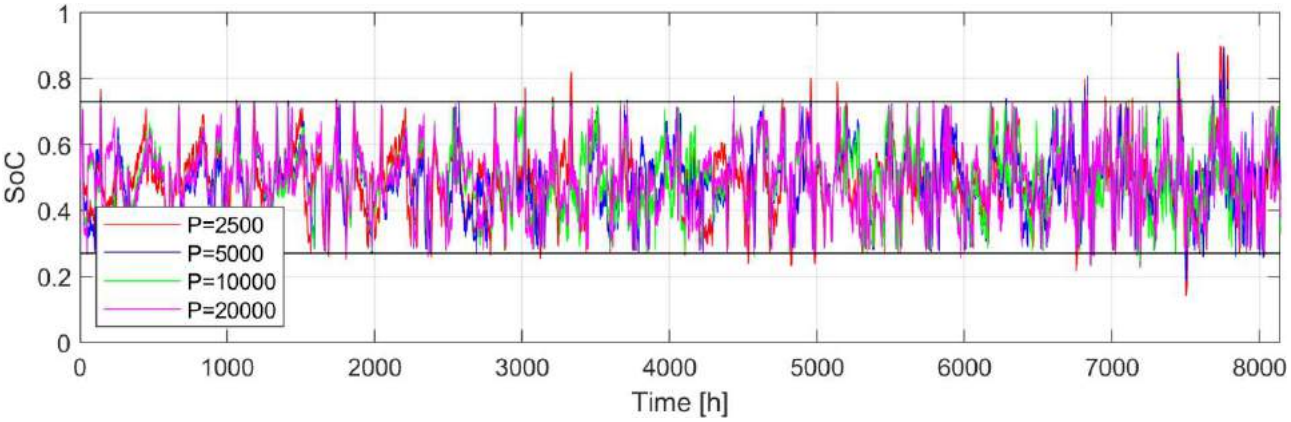

Figure 20: CE case, SoC evolutions with different restoration power rates (in W)

In Figure 21 the SoH is shown for different values of the restoration power rates. There not meaningful differences, but of course high value produces slightly greater degradation.

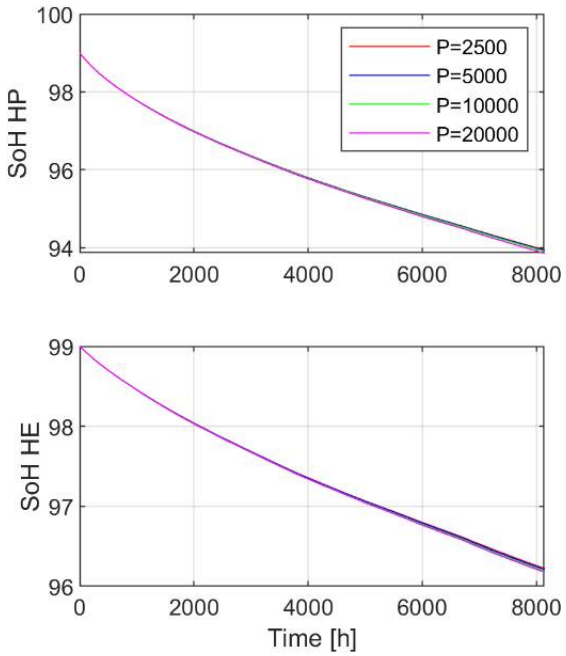

Figure 21: CE case, degradation vs restoration power rate (in W)

The degradation for the two different balancing controller and for different power-splitting filter cut-off frequency is reported in Figure 23 and Figure 22. The control strategy has a minor effect on the ageing, nevertheless the PI-based controller has lower degradation, which is preferrable. As expected, since the cut-off frequency controls how much the HESS uses the HE and HP modules, the HE degradation increases with the increase cut-off frequency, meanwhile, the HP module degradation decreases with the increase of it. Even though after a year the difference is not substantial, there may be significant differences if the entire lifetime of the system is considered. Therefore, this parameter can be chosen to balance the degradation between the modules and optimized to extend the lifetime of the HBESS.

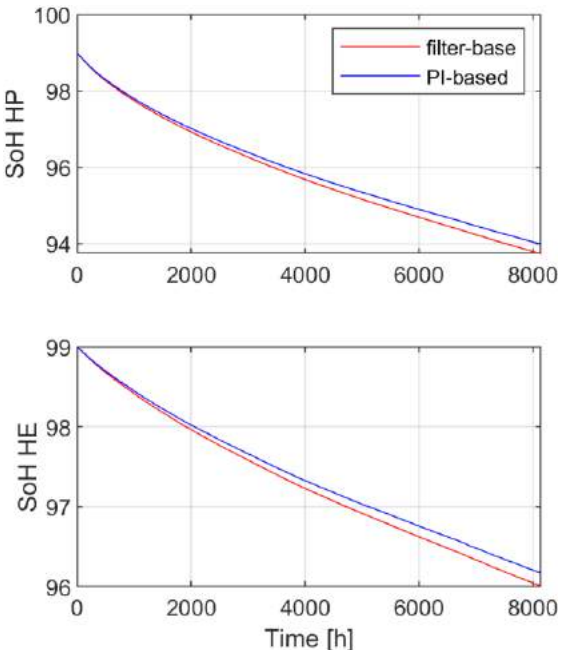

Figure 23: CE case, degradation vs balancing strategy

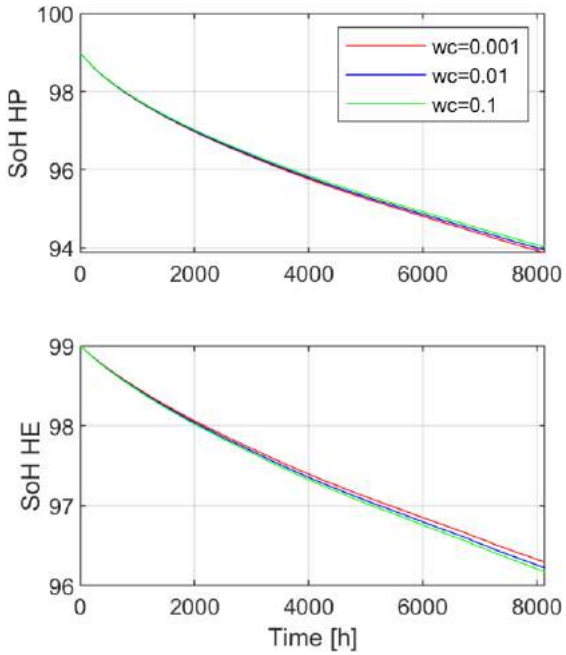

Figure 22: CE case, degradation vs cut-off frequency power splitting filter (in rad/s)

6.4.2 Nordic Synchronous Area

Considering the Nordic regulation, the evolution of the SoC behaviour is shown in Figure 24. In the upper plot, the SoC of the HE module is shown. The HP SoC is not shown as it is not very informative. Indeed, it has the same behaviour since the balancing controller forces them to be extremely similar. In the lower plot the NEM and AEM status are shown. It can be observed that, when the SoC hit the NEM activation threshold, the SoC increase/decrease stops and the discharging/discharging action is activated bringing the SoC to its nominal value. Therefore, the SoC is well constrained in the upper and lower NEM threshold. The AEM is activated only once in the entire year around the 7446th hour. In this case, the NEM is not fast and aggressive enough to stop the SoC increase, and it reaches the AEM threshold. Figure 25 shows how the power reference gets set to zero when the AEM is enabled. The main drawback of this restoration strategy is shown in Figure 26 where the NEM is enabled repeatedly every ~10 hours. This cycling is not healthy for the HESS, but it is necessary to guarantee the continuity of the service.

First, we compare the ageing effect of the two different power balancing controllers. The parameters of the controllers are set according to the tuning performed on the previous analysis in D4.2 [3]. The cut-off frequency of the splitting filter is set to 0.1 rad/s. In Figure 27 the degradation is shown for the two modules based on the type of controller. The control strategy has a minor effect on the ageing, nevertheless the PI-based controller has lower degradation.

Finally, we compare the ageing effect of the cut-off frequency of the power split filter, considering the PI-based controller as balancing regulator. In Figure 28 the degradation is shown for different cut-off frequency values. As expected, since the cut-off frequency controls how much the HESS uses the HE and HP modules, the HE degradation is proportional to the cut-off frequency; meanwhile, the HP module degradation decreases monotonically with the cut-off frequency. Nevertheless, the difference may not be substantial.

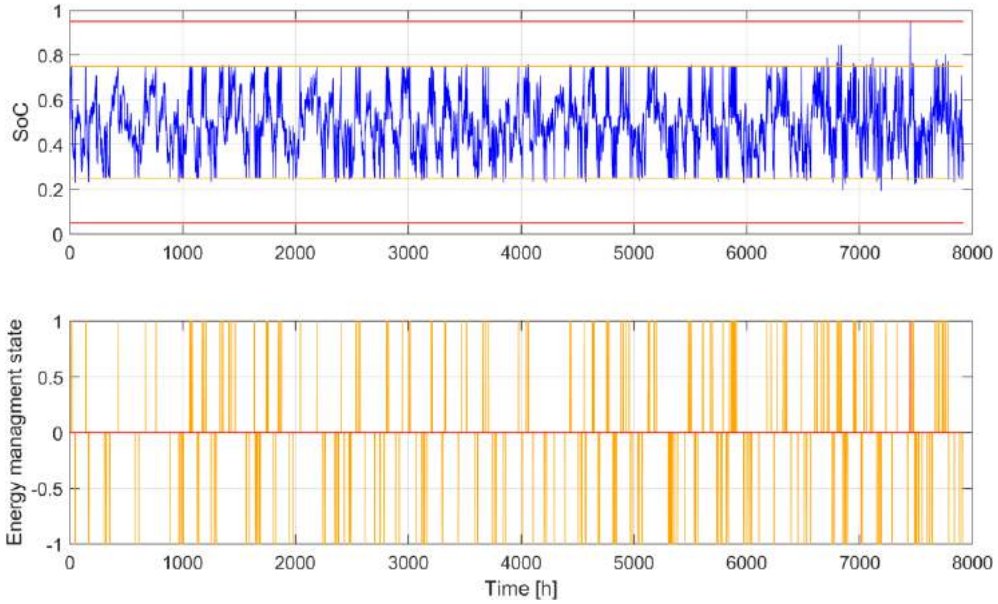

Figure 24: Nordic case: Typical SoC behaviour (up) with NEM (orange) and AEM (red) enable threshold. NEM (orange) and AEM (red) status (down), 1=Enable and discharging, 0=Disable, -1=Enable and charging

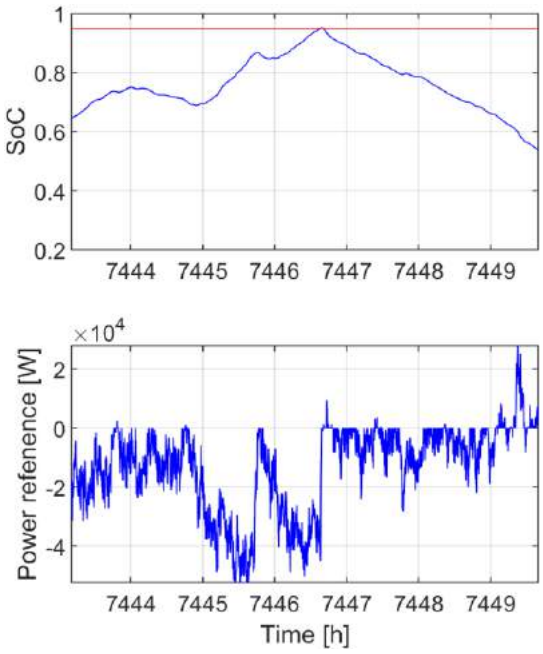

Figure 25: AEM activation

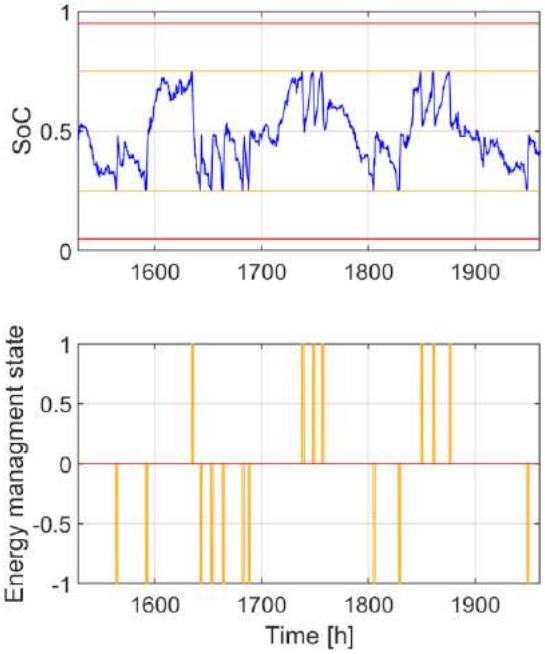

Figure 26: Frequent NEM activation

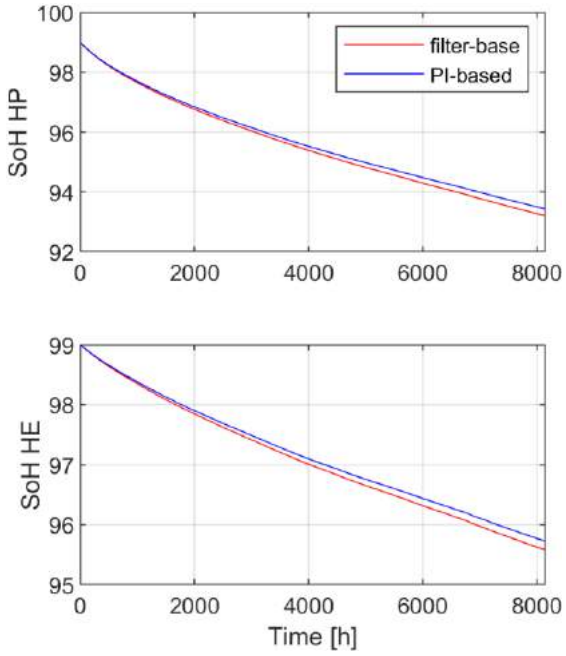

Figure 27: Nordic case, degradation vs balancing strategy

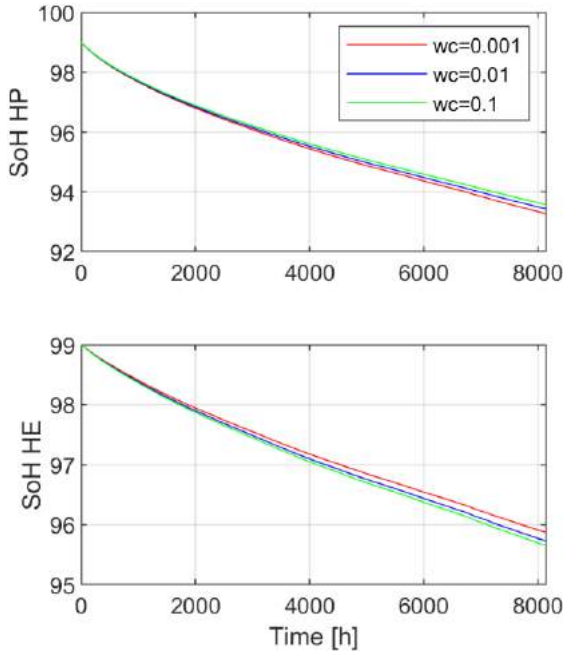

Figure 28: Nordic case, degradation vs cut-off frequency power splitting filter (in rad/s)

7 Conclusions and discussion

This report has explored the capabilities of a HESS to provide grid services of different nature. It was reported how certain services may be straightforward to extend to consider several modules (i.e., coordinated voltage and frequency support); on the other hand, for optimization-based applications, the presence of different modules increases the complexity of the controller in an exponential way, and thus a new method has been proposed. Simpler approaches, that do not consider the particularities of each module, could be designed, but would lead to very suboptimal strategies where one of the modules is often fully charged or discharged, and therefore the advantages of a HESS would disappear. The aging of the different modules while providing primary regulation was also investigated. As expected, the control parameters clearly affect the ageing: in particular, the PID-based balancing strategy is preferred over the filter-based one. Moreover, the cut-off frequency of the splitting filter can affect the long-term use of the system and can be optimized to extend the total lifetime of the HBESS.

The focus of the project (and therefore of this task) was on two modules for simplicity, namely a high-energy and a high-power module, but the extension to a higher number of modules can also be interesting. This is especially the case for the EMS and how to split setpoints among several modules. Likewise, there is no need to consider only battery storage systems, and other hybrid power plants can be included, with additional modelling effects such as ramp rate costs, or logical constraints on the operation of certain modules. Given the massive introduction of power electronics, it is expected that inverter-based generation provides more and more grid services, and even with faster response times, in order to counteract the lack of inertia in the system. In that sense, having a fast response unit that can tolerate high current rates (such a supercapacitor) may become very helpful to extend the lifetime of storage systems.

8 References

[1] E. Unamuno, H. Polat, D. Cabezuelo, J. Galarza, A. Anta, E. Toutain, T. Geury and O. Hegazy, "An Interoperable EMS for the Provision of Grid Services with Hybrid Energy Storage Systems.," in *IECON*, 2022.

[2] V. Purba, B. B. Johnson, S. Jafarpour, F. Bullo and S. V. Dhople, "Dynamic aggregation of grid-tied three-phase inverters.," 2019.

[3] iStormy Project, "Deliverable 4.2 Report on power system level simulation models and standard EMS," 2023.

[4] ENTSO-E, "Continental Europe Synchronous Area Separation on January 2021," in *Interim Report*, 2021.

[5] A. Parisio, E. Rikos and L. Glielmo, "A model predictive control approach to microgrid operation optimization," *IEEE Transactions on Control Systems Technology,* vol. 22, p. 1813–1827, 2014.

[6] M. Carrión and J. M. Arroyo, "A computationally efficient mixed-integer linear formulation for the thermal unit commitment problem," *IEEE Transactions on power systems,* vol. 21, p. 1371–1378, 2006.

[7] M. Falahi, K. Butler-Purry and M. Ehsani, "Dynamic reactive power control of islanded microgrids," *IEEE Transactions on Power Systems,* vol. 28, p. 3649–3657, 2013.

[8] F. R. S. Sevilla, C. Park, V. Knazkins and P. Korba, "Model predictive control of energy systems with hybrid storage," in *2016 IEEE Power and Energy Society General Meeting (PESGM)*, 2016.

[9] T. Pippia, J. Sijs and B. De Schutter, "A parametrized model predictive control approach for microgrids," in *2018 IEEE Conference on Decision and Control (CDC)*, 2018.

[10] J. Engels, B. Claessens and G. Deconinck, "Optimal combination of frequency control and peak shaving with battery storage systems," *IEEE Transactions on Smart Grid,* vol. 11, p. 3270–3279, 2019.

[11] M. Braun, Provision of ancillary services by distributed generators: Technological and economic perspective, Kassel University, 2009.

[12] C. Patsios, B. Wu, E. Chatzinikolaou, D. J. Rogers, N. Wade, N. P. Brandon and P. Taylor, "An integrated approach for the analysis and control of grid connected energy storage systems," *Journal of Energy Storage,* vol. 5, p. 48–61, 2016.

[13] J. W. Kolar, F. Krismer, Y. Lobsiger, J. Muhlethaler, T. Nussbaumer and J. Minibock, "Extreme efficiency power electronics," in *Integr. Power Electron. Syst. (CIPS), 2012 7th Int. Conf.*, Nuremberg, 2012.

[14] R. Grab, F. Hans, M. I. R. Flores, H. Schmidt, S. Rogalla and B. Engel, "Modeling of Photovoltaic Inverter Losses for Reactive Power Provision," *IEEE Access,* 2022.

[15] K.-N. D. Malamaki and C. S. Demoulias, "Minimization of electrical losses in two-axis tracking pv systems," *IEEE transactions on power delivery,* vol. 28, p. 2445–2455, 2013.

[16] R. Goedel, R. G. Sanfelice and A. R. Teel, "Hybrid dynamical systems: modeling stability, and robustness," *Princeton, NJ, USA,* 2012.

[17] N. Landwehr, M. Hall and E. Frank, "Logistic model trees," *Machine learning,* vol. 59, p. 161–205, 2005.

[18] E. Amaldi, S. Coniglio and L. Taccari, "Discrete optimization methods to fit piecewise affine models to data points," *Computers & Operations Research,* vol. 75, p. 214–230, 2016.

[19] A. Bemporad and M. Morari, "Control of systems integrating logic, dynamics, and constraints," *Automatica,* vol. 35, p. 407–427, 1999.

[20] G. Carpinelli, S. Khormali, F. Mottola and D. Proto, "Battery energy storage sizing when time of use pricing is applied," *The Scientific World Journal,* vol. 2014, 2014.

[21] ENTSO-E, "Transparency Platform," *https://transparency.entsoe.eu,* 2023.

[22] A. Richards and J. How, "Mixed-integer programming for control," in *Proceedings of the 2005, American Control Conference, 2005.*, 2005.

[23] “Commission Regulation (EU) 2017/1485 of 2 August 2017 establishing a guideline on electricity transmission system operation,” [Online]. Available: https://eur-lex.europa.eu/eli/reg/2017/1485/2021-03-15.

[24] “Technical Requirements for Frequency Containment Reserve Provision in the Nordic Synchronous Area,” [Online]. Available: https://www.fingrid.fi/globalassets/dokumentit/fi/sahkomarkkinat/reservit/technical-requirements-for-frequency-containment-reserve-provision-in-the-nordic-synchronous-area.pdf.

9 Acknowledgement

The author(s) would like to thank the partners in the project for their valuable comments on previous drafts and for performing the review.

Project partners:

#	Partner short name	Partner Full Name
1	VUB	VRIJE UNIVERSITEIT BRUSSEL
2	PWD	POWERDALE
3	CEG	CEGASA ENERGIA S.L.U.
4	CEA	COMMISSARIAT A L ENERGIE ATOMIQUE ET AUX ENERGIES ALTERNATIVES
5	MGEP	MONDRAGON GOI ESKOLA POLITEKNIKOA JOSE MARIA ARIZMENDIARRIETA S COOP
6	ZIG	ZIGOR RESEARCH & DEVELOPMENT AIE
7	EDF	ELECTRICITE DE FRANCE
8	TNO	NEDERLANDSE ORGANISATIE VOOR TOEGEPAST NATUURWETENSCHAPPELIJK ONDERZOEK TNO
9	PT	PRODRIVE TECHNOLOGIES BV
10	GW	GREENWAY INFRASTRUCTURE SRO
11	AIT	AIT AUSTRIAN INSTITUTE OF TECHNOLOGY GMBH
12	UNR	UNIRESEARCH BV

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 963527