Pacific Northwest
NATIONAL LABORATORY

PNNL-38288

MoCoDo: Multi-Objective Co-design Optimization for Pareto-Set Identification

Himanshu Sharma
Thiagarajan Ramachandran
Jan Drgona
Wei Wang
Buxin She
Bowen Huang
James Kotary

September 2025

U.S. DEPARTMENT of ENERGY
Prepared for the U.S. Department of Energy
Under contract DE-AC05-76RL01830

DISCLAIMER

PACIFIC NORTHWEST NATIONAL LABORATORY
operated by
BATTELLE
for the
UNITED STATES DEPARTMENT OF ENERGY
under Contract DE-AC05-76RL01830

Printed in the United States of America

Available to DOE and DOE contractors from
the Office of Scientific and Technical Information,
P.O. Box 62, Oak Ridge, TN 37831-0062
www.osti.gov
ph: (865) 576-8401
fox: (865) 576-5728
email: reports@osti.gov

Available to the public from the National Technical Information Service
5301 Shawnee Rd., Alexandria, VA 22312
ph: (800) 553-NTIS (6847)
or (703) 605-6000
email: info@ntis.gov
Online ordering: http://www.ntis.gov

PNNL-38288

MoCoDo: Multi-Objective Co-design Optimization for Pareto-Set Identification

September 2025

Himanshu Sharma
Wei Wang
Bowen Huang
Thiagarajan Ramachandran
Buxin She
James Kotary
Jan Drgona

Prepared for
the U.S. Department of Energy
Under Contract DE-AC05-76RL01830

Pacific Northwest National Laboratory
Richland, Washington 99354

Abstract

This report presents a practical framework to co-design plant-level controls, storage, and market participation demonstrated for integrating large offshore wind farms to the onshore power grid. The approach links long-term sizing and scheduling decisions with fast dynamics so that reliability and economic value are improved together. We formulate multi-scale multi-stage stochastic optimization models to coordinate wind output, battery operation, and reserve policies under uncertainty. To verify dynamic performance, we conduct detailed time-domain simulations on a reduced Western Interconnection model. The study compares two operating modes: conventional maximum power point tracking and a de-loaded strategy that intentionally holds headroom for frequency support. Results show that the co-designed operating point closely matches steady-state targets before the event, maintains stability after the trip, and delivers a higher quality frequency response than purely maximizing energy. The de-loaded policy reduces the lowest frequency dip and moderates power ramp rates while preserving sufficient energy for recovery. On the market side, the framework highlights how coordinated bidding and storage dispatch can enhance revenues and mitigate risks in day-ahead and balancing markets, while accounting for battery wear. Overall, the report demonstrates an end-to-end methodology that connects design, control, markets, and grid code validation. It provides actionable guidance on selecting reserve levels, tuning supervisory controllers, and sizing storage to support dependable, profitable offshore wind integration. The workflow uses open-source modeling tools and scenario-based uncertainty representations, and is adaptable to hybrid plants. Findings inform planners, operators, and regulators seeking to balance performance, compliance, and cost effectively.

Executive Summary

Integrating large amounts of offshore wind into power systems requires solutions that deliver both reliable performance and attractive economics. This report develops and validates a practical approach that joins plant design, supervisory control, and market strategies into one coherent workflow. The central idea is control co-design: decisions about equipment sizing, operating policies, and bidding are made together so that the plant can support the grid during fast events while also capturing value in electricity markets.

The methodology spans multiple time scales. At long horizons, the framework sizes battery storage and sets reserve policies that determine how much headroom the wind farm keeps for grid support. At operational horizons, it schedules power and energy across day-ahead and balancing markets while accounting for uncertainty. At fast time scales, it tunes controllers to shape the plant's response to disturbances. Modern open-source tools are used for optimization, scenario generation, and nonlinear problem solving, enabling transparent, reproducible analyses.

To verify that the resulting designs work under real disturbances, the study performs detailed time-domain simulations on a reduced model of the Western Interconnection. The test includes four offshore wind farms connected to the grid. Two operating modes are examined: conventional maximum power point tracking, which always maximizes wind output, and a de-loaded mode that intentionally holds back power to create headroom for frequency support. A severe contingency is applied by tripping the system's largest generator at 60 seconds.

The simulations show that the co-designed operating point yields steady power injections prior to the event and maintains stable behavior afterwards. Compared with the energy-maximizing mode, the deloaded strategy lowers the lowest frequency dip, reduces power spikes, and leaves sufficient margin to help the system recover. These findings confirm that modest headroom, when coordinated with storage and appropriate controls, can materially improve system resilience without sacrificing overall energy delivery.

The economic analysis complements the dynamic results. By coordinating bidding decisions with storage dispatch, the framework identifies operating strategies that increase revenue and reduce exposure to forecast errors. It also incorporates battery wear considerations, supporting replacement planning and long-term value preservation. In combination, the market and control insights enable the plant to contribute essential grid services while remaining financially sound.

The report offers concrete guidance for developers, operators, and planners. Key recommendations include: adopt supervisory controllers that can switch between energy-maximizing and support providing modes; hold calibrated headroom to meet frequency and ramping needs; co-optimize storage size with reserve policies; and validate designs with time-domain simulations representative of grid codes and regional conditions. Transmission and interconnection assumptions significantly influence the results, so early coordination with grid operators is advised.

In summary, the work delivers an end-to-end process that connects design, control, markets, and high-fidelity validation for offshore wind integration. The approach is adaptable to hybrid plants and regions, and provides a clear pathway to achieving dependable performance and competitive economics as renewable penetration grows. Future work will expand validation across weather regimes, refine market models for evolving products, and explore hydrogen or long-duration storage as complementary flexibility resources for reliability.

Acronyms and Abbreviations

OWF	Offshore Wind Farm
MTDC	Multi Terminal Direct Current
BESS	Battery Energy Storage System
POI	Point of Interconnection
CCD	Control Co-design
RSC	Rotor Stator Control
DA	Day Ahead Market
RT	Real Time Market

Acknowledgments

This research was supported by the Energy System Co-Design with Multiple Objectives and Power Electronics (E-COMP) Initiative, under the Laboratory Directed Research and Development (LDRD) Program at Pacific Northwest National Laboratory (PNNL). PNNL is a multi-program national laboratory operated for the U.S. Department of Energy (DOE) by Battelle Memorial Institute under Contract No. DE-AC05-76RL01830. The authors express their gratitude to Jay Barlow for his invaluable discussions regarding the offshore wind farm use case and his assistance in identifying relevant datasets for the case study.

Contents

Figures

Tables

1.0 Multi-Objective Control Co-design Using Graph-Based Optimization for Variable Generation Grid Integration

1.1 Motivation

Offshore wind farms (OWFs) are gaining increasing attention worldwide for sustainable energy development. In 2022, 8,385 MW of new offshore wind energy projects were commissioned globally [1]. In the U.S., the offshore wind energy production capacity potential reached 52,687 MW in 2023, showing a growth of 15% [1]. With the increasing power extraction from OWFs, it is crucial to develop capabilities for efficient power transmission. Most modern OWFs are developed with multi-terminal DC (MTDC) grids utilizing modular multilevel converters (MMC) due to their advantages over high-voltage alternating current (HVAC) lines. A detailed review of MMC-MTDC grids can be found in [2]. Large renewable energy source integration introduces challenges for AC-grid operators; a comprehensive discussion is presented in [3]. One widely recognized approach to ensure power system stability involves battery energy storage systems (BESS) and their operation and control [4].

BESS sizing on either the AC or DC side is an important decision during OWF interconnection planning. A comprehensive review focusing on determining optimal sizing for wind farm applications is available in [5]. Recent work by Halwany et al. [6] developed a probabilistic approach for onsite energy storage sizing for OWF black start operations. However, this work did not consider onshore BESS sizing and interconnection with an MTDC grid. Santanu et al. [7] also proposed a multi-objective approach for battery sizing in OWFs, considering economic and reliability objectives. This work developed a sequential approach for handling multiple objectives but neglected battery controls and converter dynamics. Moghaddam et al. [8] considered the BESS sizing problem for an onshore wind farm; however, they used a sequential approach where BESS size was chosen first, followed by a control strategy. To the best of the authors' knowledge, no studies have simultaneously accounted for the control operations of offshore wind farms while sizing BESS.

Conventionally, the design problem is solved first, followed by operational control optimizations. However, many studies [9]–[15] have shown that such a sequential approach results in suboptimal system performance. Control co-design (CCD) is a control system design approach that considers the interactions between the control system and the underlying design of the physical system. Comprehensive reviews of CCD and handling uncertainties in its formulation are presented in [16] and [17], respectively. In this report, we aim to develop a CCD approach suitable for BESS design for OWFs. Specifically, we are interested in developing a CCD approach to handle the challenge of large system CCD with tight coupling among subsystems (e.g., MMCs). We aim to develop a generalized framework to pose a co-design optimization problem that can also handle subsystem-level coupling constraints during control and design optimization. Furthermore, many energy system designs require satisfying multiple objectives (e.g., low operation cost, minimal power loss). Few studies on co-design control (CCD) considering multiple objectives have been conducted without explicitly identifying the Pareto front [18]–[20]. Inspired by recent research on optimizing marine energy kites [21], we propose a CCD approach that addresses multiple objectives to find Pareto solutions for integrating offshore wind farms into the grid. Therefore, this report aims to address the aforementioned research objective by developing the CCD approach using a graph-based optimization framework. This framework enables scalability of optimization for large systems and allows for formulating a subsystem-level co-design problem. Additionally, we develop a gradient-based approach for handling multiple objectives to identify the Pareto set.

The remainder of this report is organized as follows: Section 1.2 discusses the proposed

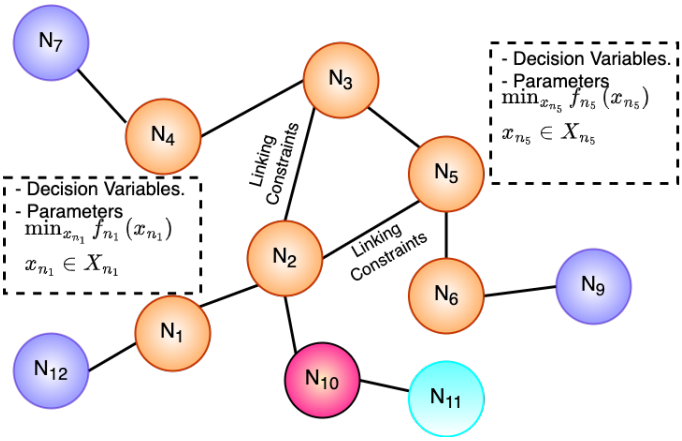


Figure 1: A schematic for the graph-based optimization formulation. Each node defines a subsystem with individual objective functions.

methodology for using graph-based optimization and the gradient-based multi-objective optimization approach for Pareto set identification. Section 1.5 describes the OWF use case for CCD and the developed optimization formulation. In Section 1.6, we discuss the results from the approach, and present our conclusions and future work in Section 1.7.

1.2 Methodology

This section details the proposed methodology developed for performing CCD for OWFs interconnected with an AC-grid through BESS.

1.3 Graph-Based Optimization

Graph-based modeling abstractions have recently been explored in convex optimization [22], infrastructure networks [23], supply chain planning problems [24], and simulation of partial differential equations [25]. The structures of these abstractions are directly tied to the physical topology of the systems. Recent work by [23], [26], [27] has shown that optimization and simulation for complex systems can be represented using a graph-based computational framework. This provides a coherent strategy to capture modeling elements for a system, which are often common in most engineering applications. Figure 1 illustrates the graph-based representation of an optimization problem, consisting of a set of nodes and edges. Each node represents an individual subsystem optimization model (with variables, objectives, constraints, and data), and each edge captures connectivity between node models and coupling constraints. Once constructed, the graph can be communicated to traditional or decomposition optimization solvers (e.g., *Gurobi* or *Ipopt*). In the CCD problem for OWF connecting to an AC grid through MTDC, the system topology of MTDC and AC grids is used to define the nodes and edges. Each node corresponds to a subsystem-level representation. Details about the CCD formulation and how to convert it to a graph-based model will be discussed in Section 1.5. Next, the proposed gradient-based approach for solving multi-objective optimization is described.

1.4 Gradient-Based Multi-Objective Optimization

A gradient-based multi-objective optimization framework is proposed, inspired by the weighted-sum method and the bi-level optimization algorithm in [28]. Given N objective functions that can be split into n nodes, the weighted-sum combination of single objective functions $\{f_1, f_2, \ldots, f_n\}$, $\{g_1, \ldots, g_n\}$, $\{h_1, \ldots, h_n\}$ can be written as:

$$\begin{aligned} \min \quad & w_1(f_1 + f_2 + \ldots + f_n) + w_2(g_1 + g_2 + \ldots + g_n) \\ & + \ldots + w_N(h_1 + h_2 + \ldots + h_n) \end{aligned} \tag{1}$$

where $\{f\}$, $\{g\}$, $\{h\}$ are groups of individual objective functions within different nodes. The proposed multi-objective optimization framework can then be applied to update the weights to find the Pareto frontier, as detailed in Algorithm 1.

Algorithm 1 Gradient-based Approach for Multi-Objective Optimization

Set the step sizes β for updating $\mathbf{w}$ Solve the graph-based problem with initial weights $\mathbf{w}_0 = (w_1, w_2, ..., w_N)$ using *Plasmo.jl* **for** $k = 0$ ***to*** K ***step*** 1 **do**

  Update $\mathbf{w}$ with projected gradient descent $\mathbf{w}_{k+1} = \text{proj}_{\Delta N}(\mathbf{w}_k + \beta h^k_{\mathbf{w}})$ Solve the graph-based problem with updated weights $\mathbf{w}_{k+1} = (w_1, w_2, ..., w_N)$ using *Plasmo.jl*

end

return $\bar{w} = w(\tau)$, where $\tau \sim \mathcal{U}(1, ..., K)$

As shown in Algorithm 1, the inputs of this algorithm include the initial values of weights $\mathbf{w_0}$, step size β, and well-defined multi-objective functions, e.g., $(f_1, f_2, \ldots, f_n)$ and $(g_1, g_2, \ldots, g_n)$ ($N = 2$). The weighted sum of the given objective functions can be written into an n-node graph-based formulation as in (1). During iterations from $k = 1$ to K, the weights $\mathbf{w}$ are updated by the step size β using the per-objective stochastic gradient estimates $h^k_{\mathbf{w}} = [\nabla_{\mathbf{w}_k}\mathbf{f}, \nabla_{\mathbf{w}_k}\mathbf{g}]$. The new weights are then projected onto the N-simplex defined by $\Delta N := \{\mathbf{w} \in \mathbb{R}^N : w_i \geq 0, \forall i \in [N], \sum_{i\in[N]} w_i = 1\}$. For each iteration, the problem is solved as a single objective graph in *Plasmo.jl*. After K iterations, the algorithm's output is uniformly chosen from the generated Pareto frontier weights $(\mathbf{w}_1, \ldots, \mathbf{w}_K)$.

It is imperative to mention that while numerous conventional methods are available for multi-objective optimization problems, the proposed gradient-based algorithm is chosen for three primary reasons. Firstly, computational efficiency is a critical issue when solving large-scale graph-based multi-objective optimization problems, especially when binary variables are present. The graph-based formulation can be easily integrated into our proposed algorithm for computing its projected gradient without losing convexity. Secondly, the proposed algorithm supports vector-valued nonlinear objectives and constraints, which cannot be directly solved in existing *Plasmo.jl* or *JuMP* multi-objective solvers yet. Finally, this approach is more scalable for graph-based formulation when combined with *Plasmo.jl* for application in large-scale graph networks.

1.5 Case Study: Offshore Wind Farm Interconnect

Our methodologies are applied to the well-known WSCC 9-bus system. Two OWFs are connected to buses 4 and 6 through a four-terminal MMC-based MTDC network to transmit wind power [29], [30]. Two BESS are attached to buses 4 and 6 to help reduce total cost and improve system efficiency. The system structure is shown in Figure 2.

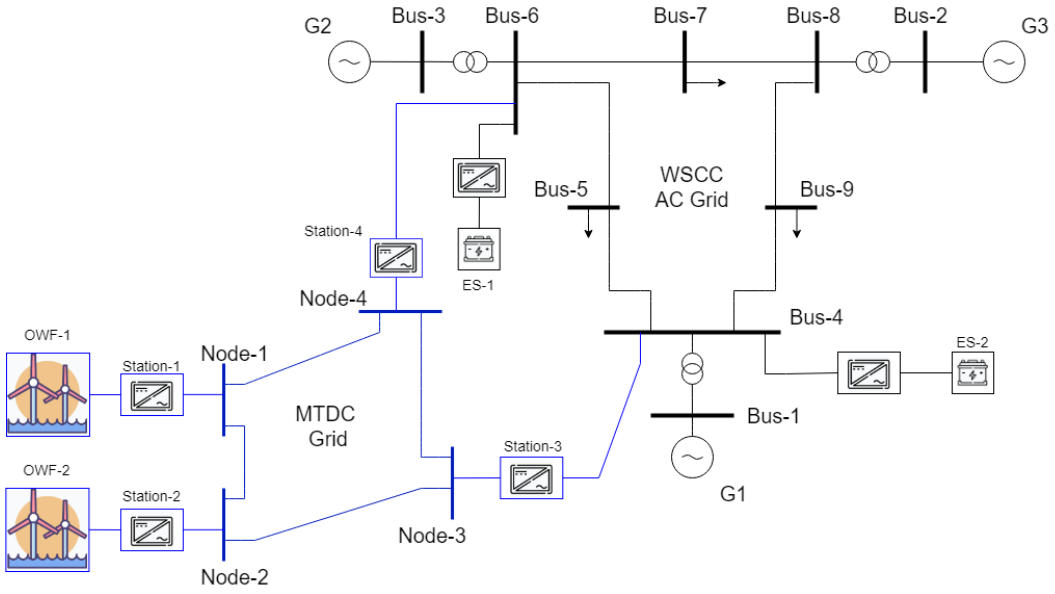


Figure 2: IEEE-9 bus system interconnected with OWFs using MMC-MTDC grid.

OWFs with converters are treated as active power sources, and energy storage devices with converters are modeled as active power loads or sources depending on their charging/discharging operations. MMCs connecting AC and MTDC grids are simply modeled with fixed coefficients [31], [32]. The original nonconvex formulation of our MTDC and AC power flow CCD problem is relaxed to a second-order conic program (SOCP) [33], [34]. This typically yields more accurate solutions than linear models and avoids local optimality issues in solving nonconvex problems. SOCP can be solved by many optimization solvers or calculated through decomposition algorithms.

Control Co-design Optimization Formulation

The following two objectives are considered in our model. This formulation can be easily extended to include additional objectives.

$$\begin{aligned}\min \quad & \sum_{i\in\mathcal{S}} f_i^{EI}(BS_i) + \sum_{i\in\mathcal{B}_{AC}}\sum_{t\in\mathcal{T}} f_i^{G}(P_{i,t}^{G}, Q_{i,t}^{G}) \\ & + \sum_{i\in\mathcal{S}}\sum_{t\in\mathcal{T}} f_i^{EO}(P_{i,t}^{ch}, P_{i,t}^{dis}) \end{aligned} \tag{2}$$

$$\begin{aligned}\min \quad & \sum_{i\in\mathcal{C}}\sum_{t\in\mathcal{T}} P_{i,t}^{Loss} + \sum_{i,j\in\mathcal{L}_{AC}}\sum_{t\in\mathcal{T}} g_{ij}\left(c_{ii,t} + c_{jj,t} - 2c_{ij,t}\right) \\ & + \sum_{ij\in\mathcal{L}_{DC}}\sum_{t\in\mathcal{T}} g_{DC_{ij}}\left(v_{ii,t} + v_{jj,t} - 2v_{ij,t}\right) \end{aligned} \tag{3}$$

The objective function (eq. 2) minimizes total cost, including BESS installation $f^{EI}(\cdot)$, regular generator fuel cost $f^{G}(\cdot)$, and BESS operation cost $f^{EO}(\cdot)$. $\mathcal{S}$, $\mathcal{B}_{AC}$, and $\mathcal{T}$ represent the index sets of batteries, buses in the AC grid, and time intervals, respectively. In the cost functions, BS_i is the size of BESS i; $P_{i,t}^{ch}$ and $P_{i,t}^{dis}$ are their charged and discharged power at time t,

respectively; and $P^G_{i,t}$ and $Q^G_{i,t}$ are active and reactive power output of generator i at time t, respectively.

The objective function (eq. 3) minimizes total power loss, including those in MMCs, AC grid transmission, and DC grid transmission. In this objective function, $\mathcal{C}$, $\mathcal{L}_{AC}$, and $\mathcal{L}_{DC}$ are index sets of MMCs, branches of AC and MTDC grids, respectively. Variable $P^{Loss}_{i,t}$ represents the power loss in converter i at time t. Parameters g_{ij} and $g_{DC_{ij}}$ are conductances of AC and DC branch ij, respectively. Variables $c_{ij,t}$, $v_{ij,t}$, and $s_{ij,t}$ (appearing later) are used for replacing voltages of AC and DC buses to obtain SOCP relaxation. Specifically, if E_i and F_i are the real and imaginary parts of voltage at AC bus i, and V_i is the voltage at DC bus i, then $c_{ij} = E_iE_j + F_iF_j$, $s_{ij} = E_iF_j - E_jF_i$, and $v_{ij} = V_iV_j$. Therefore, for each AC or DC branch, they should satisfy the following relationship:

$$c_{ij,t} = c_{ji,t},\ s_{ij,t} = -s_{ji,t}\ \forall ij \in \mathcal{L}_{AC},\ t \in \mathcal{T} \tag{4}$$

$$c^2_{ij,t} + s^2_{ij,t} \le c_{ii,t}c_{jj,t} \quad \forall ij \in \mathcal{L}_{AC},\ t \in \mathcal{T} \tag{5}$$

$$v_{ij,t} = v_{ji,t},\ v^2_{ij,t} \le v_{ii,t}v_{jj,t} \quad \forall ij \in \mathcal{L}_{DC},\ t \in \mathcal{T} \tag{6}$$

For each bus i at time period t, the following active and reactive power balance constraints should be satisfied:

$$P^G_{i,t} - P^D_{i,t} + P^{Conv}_{i,t} - P^{ch}_{i,t} + P^{dis}_{i,t} = \sum_{j \in \mathcal{B}_{AC}} (G_{ij}c_{ij,t} - B_{ij}s_{ij,t}) \tag{7}$$

$$Q^G_{i,t} - Q^D_{i,t} = -\sum_{j \in \mathcal{B}_{AC}} (G_{ij}s_{ij,t} + B_{ij}c_{ij,t}) \tag{8}$$

Here, parameters $P^D_{i,t}$ and $Q^D_{i,t}$ are active and reactive power demand, and G and B are real and imaginary parts of the AC grid admittance matrix. Variable $P^{Conv}_{i,t}$ is the active power injection from the MMC.

Let $\underline{V}_i$ and $\overline{V}_i$ be the lower and upper bounds of the voltage at bus i. These limits can be imposed as:

$$\underline{V}^2_i \le c_{ii,t} \le \overline{V}^2_i \quad \forall i \in \mathcal{B}_{AC},\ t \in \mathcal{T} \tag{9}$$

The active and reactive power output of generator i at time t have the following ramping and bound limits, where the left and right-hand side values are corresponding parameters:

$$-Ramp^{P-}_i \le P^G_{i,t+1} - P^G_{i,t} \le Ramp^{P+}_i \tag{10}$$

$$-Ramp^{Q-}_i \le Q^G_{i,t+1} - Q^G_{i,t} \le Ramp^{Q+}_i \tag{11}$$

$$P^{\min}_i \le P^G_{i,t} \le P^{\max}_i \tag{12}$$

$$Q^{\min}_i \le Q^G_{i,t} \le Q^{\max}_i \tag{13}$$

Similar to the AC grid, each MTDC bus i at time t has a balance constraint, where $\mathcal{B}_{DC}$ is the index set of buses in MTDC, $P^{WF}_{i,t}$ is offshore wind power generation, G_{DC} is the DC grid admittance matrix, and variable $P^{DC}_{i,t}$ is the power injected into the MMC:

$$P^{WF}_{i,t} - P^{DC}_{i,t} = \sum_{j \in \mathcal{B}_{DC}} v_{ij,t}G_{DC_{ij}} \tag{14}$$

Each MMC has three constraints: power balance, loss estimation, and voltage droop control, presented below. The index i on the left and right-hand sides represents the AC and DC buses it connects, respectively.

$$P_{i,t}^{Conv} + P_{i,t}^{Loss} = P_{i,t}^{DC} \tag{15}$$
$$P_{i,t}^{Loss} = \beta_i \left| P_{i,t}^{DC} \right| \tag{16}$$
$$(k_i P_{i,t}^{Conv} + d_i)^2 \leq v_{ii,t} \tag{17}$$

Here, β is the efficiency coefficient, and k and d are droop control parameters.

Finally, we have the constraints for each BESS i:

$$SC_{i,t} - SC_{i,t-1} = \eta_i^{ch} P_{i,t}^{ch} - \eta_i^{dis} P_{i,t}^{dis} \tag{18}$$
$$0 \leq P_{i,t}^{ch} \leq P_i^{ch\_\max} z_{i,t} \tag{19}$$
$$0 \leq P_{i,t}^{dis} \leq P_i^{dis\_\max} (1 - z_{i,t}) \tag{20}$$
$$0 \leq SC_{i,t} \leq BS_i \tag{21}$$
$$BS_i^{\min} \leq BS_i \leq BS_i^{\max} \tag{22}$$

Constraint (18) is the operation equation, where parameters η^{ch} and η^{dis} are charging and discharging efficiency, respectively. Constraints (19-22) define limits for charging and discharging rates, states of charge, and sizes, respectively.

In this problem, battery size BS is chosen as a design variable, and the remaining operation-related variables are chosen as control variables. These include charged and discharged power of batteries (P^{ch}, P^{dis}), active and reactive power output of generators (P^G, Q^G), AC and MTDC power flow related variables (c, s, v in constraints eq. 4-6, eq. 14), and power through MMC (P^{Conv}, P^{Loss}, P^{DC}).

To convert this problem into a graph-based model, each AC/DC bus, AC/DC branch, converter, and BESS is defined as a graph node. The linking constraints of the edges and objective terms in each node are derived based on the above formulation. An eight-hour time horizon is considered for the problem. The resulting graph is shown in Figure 3, with each color representing an hour. The graph's structure illustrates nodes with associated decision variables connected sequentially in time (sequential hours). Selecting one of the hours on the graph (circled in black) shows the linking constraints among different components, including the same components of the system.

1.6 Results

DC grid branch data is presented in Table 1. Converter droop control and efficiency parameters are $k = 0.02$, $d = 1$, and $\beta = 0.03$. Battery minimum ($BS^{\min}$) and maximum ($BS^{\max}$) capacities are 20 MWh and 120 MWh, respectively, and their initial and minimum states of charge at the last hour are set to half of $BS^{\min}$. Battery charge and discharge efficiencies are 0.8 and 1.1, respectively. Nominal wind farm (WF) outputs are set as 40 MW and 50 MW, respectively. Nominal load and generator cost functions are taken from the 9-bus system data. In each hour, WF output, nominal load, and fuel cost are multiplied by the factors provided in Table 2. The problem is formulated in `Julia` programming using `JuMP` and `Plasmo` packages, and solved by `Ipopt` and `Juniper` on a laptop with an i7-12800H CPU and 16GB RAM.

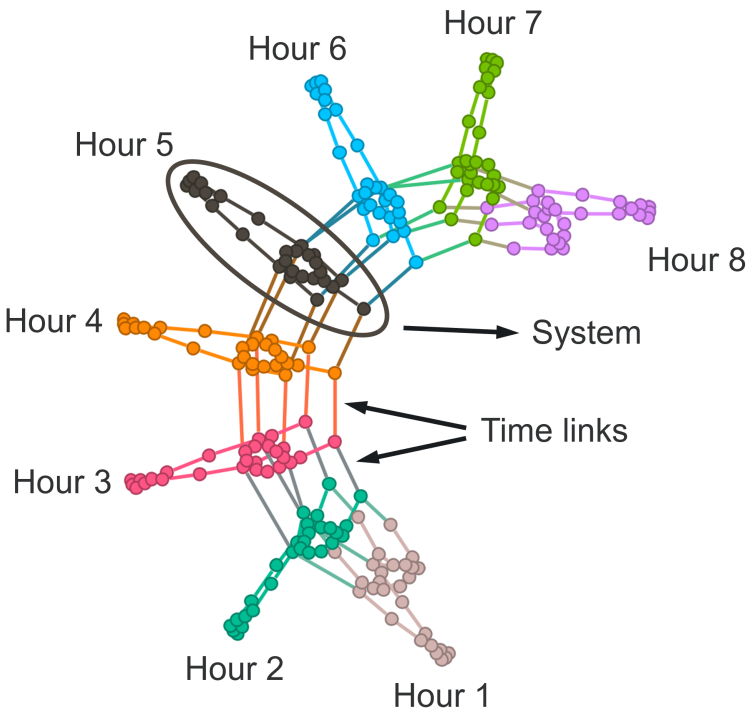


Figure 3: The graph-based co-design optimization problem visualization for the IEEE-9 bus-MTDC use case. Each color indicates a time snapshot of the complete system described as a graph node for 8 hours.

Table 1: DC Grid Branch Data

From	To	R(p.u.)	From	To	R(p.u.)
1	2	0.0016	2	3	0.0048
1	4	0.0048	3	4	0.0042

1.6.1 Single Objective Function

Only the objective function (eq. 2) for minimizing total cost is considered initially. The results, obtained by setting each of the two batteries with fixed sizes ranging from 20 MWh to 120 MWh in 10 MWh increments, are shown in Figure 4.

By setting battery sizes as decision variables, total cost is minimized with an optimal solution of 26.5 MWh and 93.2 MWh. The time for solving this problem is approximately 45 seconds, whereas solving all 11 problems with fixed battery sizes takes almost 15 minutes in total.

To better understand how batteries reduce total cost, their hourly operations are investigated in detail. In Figure 5, the total charging (positive)/discharging (negative) power and state of charge of the batteries are shown by solid lines in the sub-figures, respectively. This includes results with 11 fixed battery sizes and that of CCD by setting battery sizes as decision variables. Dashed lines represent relative load and generation cost levels in Table 2. ”BS” in the legend stands for ”Battery Size”. When load and generation costs are low, such as in the first and last few hours, batteries are charged. This charged power is then used to satisfy demand when load and cost are high. By transferring load from peak to off-peak hours, the total cost is

Table 2: Load, Wind Farm Output, and Cost Levels

Hour	1	2	3	4	5	6	7	8
Load	0.9	1.1	1.25	1.4	1.55	1.3	1.15	1
WF Output	1	0.95	1.05	0.9	0.85	1	1.1	0.95
Fuel Cost	1.1	0.9	1.3	1.5	1.8	1.6	1.4	1.4

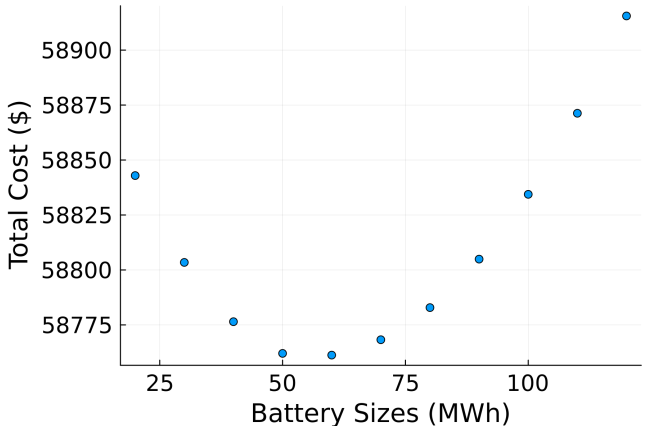


Figure 4: Minimum total system costs with different battery sizes.

reduced. By choosing the best battery sizes based on single-objective control co-design, the model can balance the cost of battery installation, operation, and power generation.

Since batteries transfer load from peak to off-peak time to reduce cost, it is intuitive that different load levels may require different battery sizes. The total cost of the system is calculated based on these 11 fixed battery sizes, considering demand levels between 98% and 104% of the original nominal demand. The results are shown in Figure 6. As the load increases, the optimal battery sizes for achieving minimum total cost also increase. On one hand, higher demand may require larger battery capacities for transferring load from peak to off-peak hours. On the other hand, higher generation costs resulting from larger loads allow for larger budgets for battery installation and operation.

1.6.2 Multiple Objective Functions

In this subsection, both objective functions (eq. 2) and (eq. 3) are minimized together. Initially, a linear scalarization approach is adopted to find the Pareto front by defining individual objective weights. Figure 7(a) compares the Pareto fronts derived from fixed battery sizes with those obtained using the CCD approach. The dashed lines depict the fronts for fixed sizes, intersecting in certain regions, indicating that optimal battery sizes for one objective may be suboptimal for others. In contrast, the solid black line, representing the CCD approach, consistently remains below the dashed lines, signifying a non-dominated solution. This indicates that the CCD approach achieves the lowest total cost for a given level of power loss or the least power loss for a given total cost.

Three points on the Pareto front identified through CCD (refer to Fig. 7(a)) were selected.

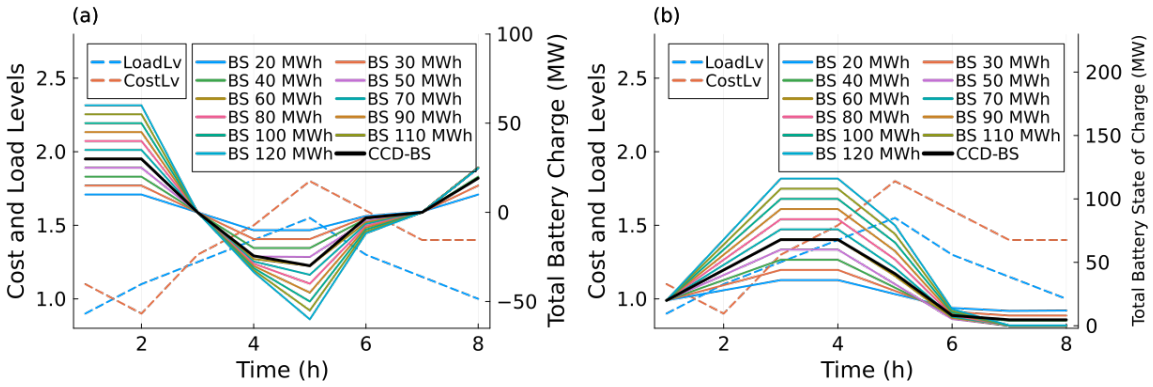


Figure 5: Different battery sizes and relative cost/load levels in each hour (a) Battery charging/discharging operation (b) Battery state of charge.

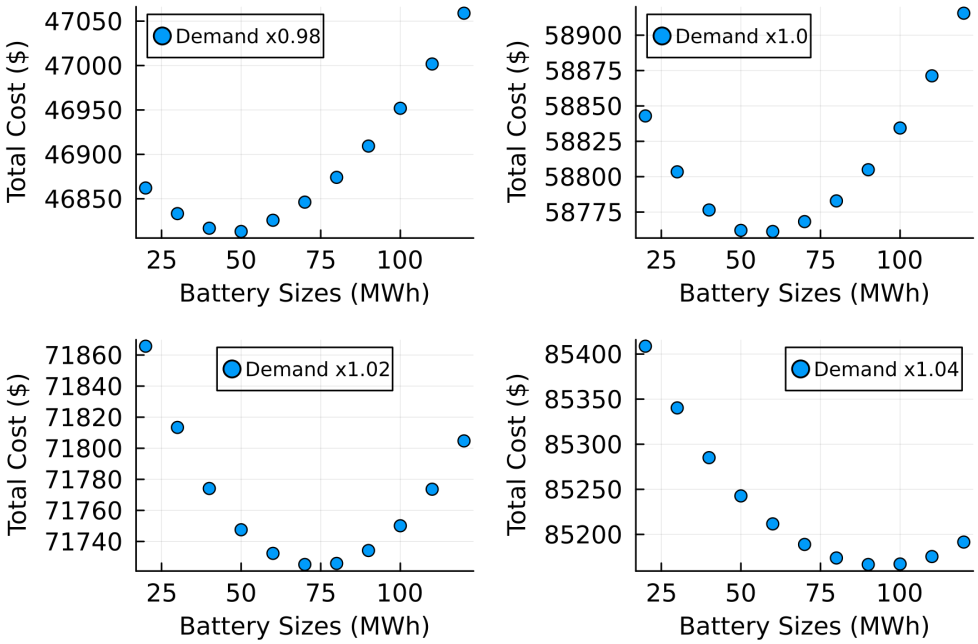


Figure 6: Total system cost for different nominal demand and battery sizes.

The endpoints of the curve represent optimization for a single objective, whereas the midpoint is the Pareto optimal solution that balances both objectives. Figure 7(b) illustrates the battery operation at these three points. When minimizing power loss, battery usage is highest; conversely, it is lowest when minimizing total cost. The Pareto optimal solution operates the battery at a level that compromises between minimizing power loss and total cost.

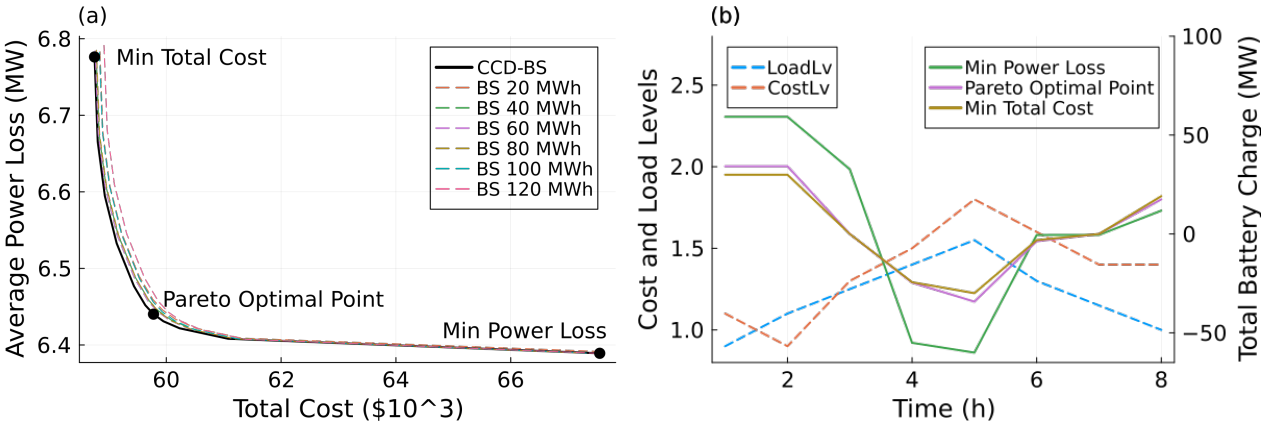


Figure 7: (a) Pareto front of two objective functions with fixed battery sizes and CCD. (b) Battery charging/discharging operations in the three selected solutions on the Pareto front.

Algorithm 1 is applied with different numbers of iterations $K = 10, 30, 100$, to identify weight combinations for the Pareto front, as shown in Fig. 8. The zoomed view shows the algorithm sampling in the region of maximum change.

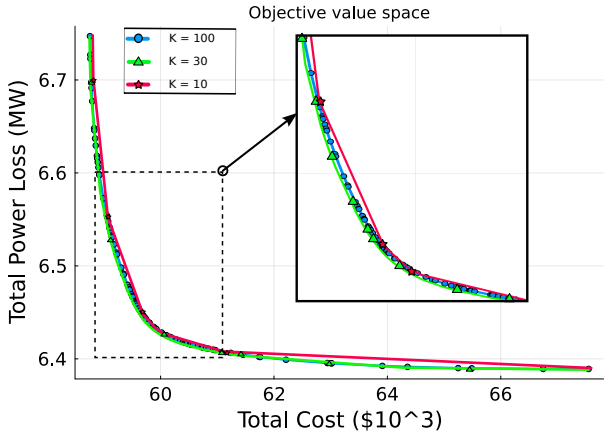


Figure 8: Pareto front of two objective functions with different iterations K.

State-of-the-art multi-objective optimization methods like evolutionary and swarm-based algorithms are effective but become computationally expensive with larger problems. The graph-based optimization outperforms these methods, offering faster computation and fewer iterations. Applied to the IEEE-9-bus system, it computes the Pareto front with impressive efficiency: for a resolution of $K = 100$, it takes only 17 minutes, and for $K = 10$, just 79.9 seconds. This is significantly quicker than traditional solvers like IPOPT, which take over 150 seconds for the same task.

1.7 Conclusions

This study introduces a graph-based, multi-objective CCD method for optimizing storage sizing in an AC-grid linked to OWFs via MTDC, considering power generation and transmission controls. Results indicate that energy storage can shift demand to off-peak hours, reducing costs, with larger storage preferred for higher demand. A gradient-based framework was used to identify the Pareto Front for minimizing cost and power loss, revealing variable optimal energy storage solutions. Future work will apply this CCD to larger systems and incorporate offshore wind variability, demonstrating the advantages of graph-based optimization.

1.7.1 Validation of Control Co-design Offshore Wind Farm Using Simulations (PSCAD)

In Figure 19, the droop control variable k at COTWDPGE for scenario 80 in case CCD18-3 is plotted. The pink band illustrates the distribution of k droop for all scenarios. For most of the time, the ESS variable is at its upper limit. By adjusting the droop status of the wind farm and ESS, the system optimizes its output while ensuring sufficient reserve for frequency support.

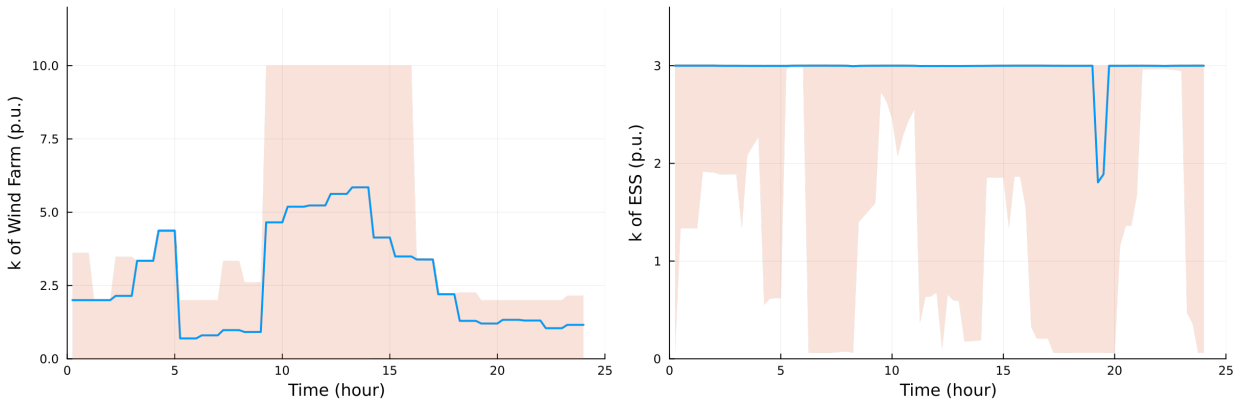


Figure 9: Droop control variable k at COTWDPGE for scenario 80 in case CCD18-3.

To further validate the dynamic performance of the co-design results, a time-domain electromagnetic transient (EMT) simulation was conducted using PSCAD [35]. Both Maximum Power Point Tracking (MPPT) and de-loading control strategies were integrated and tested under the scenario of losing the largest generator at Palo Verde, Arizona, at 60 seconds, which had an output of 2182.3 MW before tripping. Additional details on the PSCAD model can be found in E. Transient trajectories of power and frequency, depicted in Figures 20(a-b) and 20(c), were then analyzed to verify the stability and effectiveness of the co-design results.

Overall, the miniWECC system integrating OWFs remains stable both before and after a generator trip on the main grid. Prior to the 60-second mark, the steady-state output of the four OWFs shows only minor differences compared to the co-design results. After the disturbance, the power output and the increase in power are both smaller than those of the MPPT curves, indicating sufficient headroom. Additionally, the frequency nadir is lower when the OWF operates in MPPT mode without providing frequency support. Consequently, the co-design approach demonstrates superior transient performance, which is crucial given the increasing penetration of wind and other inverter-based resources.

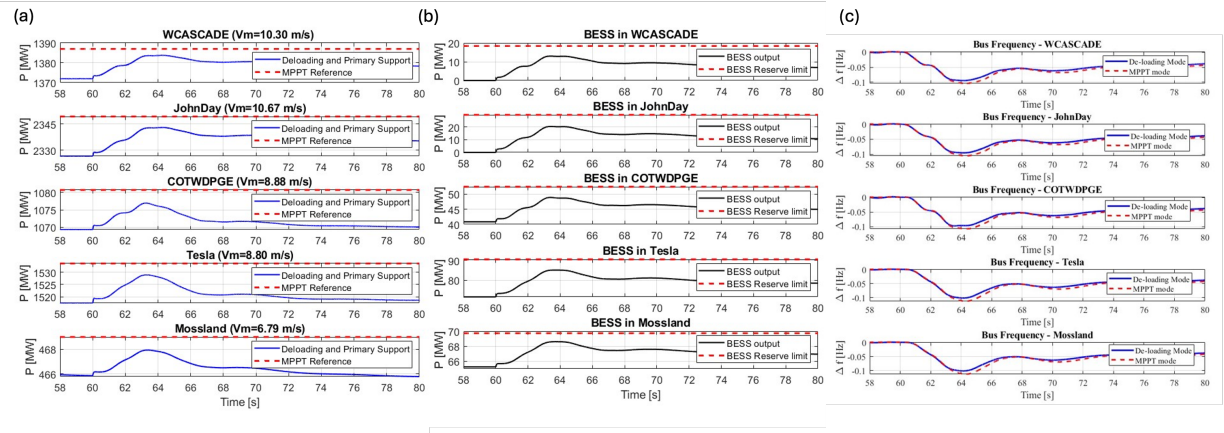


Figure 10: Power curves of 240-bus miniWECC system in PSCAD: (a) OWF; (b): BESS (c) POI frequency curves of 240-bus miniWECC system in PSCAD.

1.7.2 Conclusions and Future Work

Simultaneous consideration of operations, controls, and market participation during the planning of renewable energy systems, such as offshore wind farms, can achieve long-term reliability and optimal performance. Incorporating various uncertainties into the decision-making process enhances overall system resilience. In this report, a multi-timescale, multi-stage stochastic control co-design methodology is developed and presented. This approach models the multi-timescale nature of energy market participation and addresses uncertainties associated with renewable resources and prices at the interconnection point for revenue maximization. The approach is demonstrated using an offshore wind farm grid integration use case. Here, the sizing of HVDC power export cables from offshore to onshore is optimized as part of the process to prevent under-sizing and over-sizing, which could significantly impact project investment costs. Furthermore, energy storage sizing is considered, addressing grid stabilization needs and incorporating market participation. Both offshore wind farm and energy storage participation in ancillary service markets are evaluated for additional revenue opportunities for wind farm owners. The proposed control co-design approach optimizes cable and energy storage sizing, and market participation with droop parameters, showing improved annual revenue for five different capacity wind farms studied. A scenario tree-based stochastic formulation incorporates uncertainty. Additionally, the framework allows for the evaluation of various inflation rates and the implications of energy storage on overall revenue. The study provides validation for droop control parameters through transient PSSE simulations of an offshore wind farm integrated with the miniWECC AC grid model, showing comparable reserve values and performance in meeting the integrated power grid's frequency requirements.

Currently, reliance is placed on historical wind speed and price data, as offshore wind farm development in the U.S. is still maturing. Future efforts could enhance price data estimation using detailed production cost planning with offshore wind farms and consider futuristic climate scenarios for wind speed data. Furthermore, different network topologies (e.g., meshed) could be explored for power wheeling, with incentive designs providing additional revenue streams for offshore wind farm owners. The study can also be expanded by including additional substation-level design and control variables and incorporating reduced-order models of dynamic converter-inverters. As energy market participation becomes increasingly important for future renewable energy developers, the proposed framework could be extended to

next-generation energy market models to estimate revenue and develop participation strategies.

2.0 Control Co-Design for Variable Generation Grid Integration with Energy Storage and Energy Market

2.1 Motivation

Offshore wind farms (OWFs) are pivotal to the global transition towards sustainable energy, harnessing higher wind speeds over open seas. This has driven their rapid international expansion. Over the past two decades, the offshore wind industry has achieved significant technological advancements, such as floating platform turbines and efficient generation systems. These innovations have notably increased wind farm capacity, efficiency, and extended their reach to deeper waters. The U.S. offshore wind energy pipeline grew by 53% in 2024 compared to the previous year, with marked expansions on both the East and West Coasts [36].

As OWF development and grid integration advance, enhancing U.S. grid capacity becomes crucial [37]–[39]. Robust transmission infrastructure, such as High-Voltage Direct Current (HVDC) cables, is essential for efficient long-distance power transfer. HVDC systems offer lower energy losses than Alternating Current (AC) systems, making them ideal for remote offshore wind farms. Voltage Source Converters (VSC) [40] and Modular Multilevel Converters (MMC) [41] are key components for designing multi-terminal HVDC (MTDC) systems.

The design and operation of the MTDC network, which includes offshore substations and electrical cables, are critical for efficient OWF integration and can significantly impact system costs. This study focuses on cable size design to maintain problem tractability. Assessing cable size in conjunction with operational strategy helps prevent under-sizing or over-sizing of transmission cables, directly influencing MTDC network costs.

Beyond MTDC network design, ensuring system stability is vital. Properly sized energy storage systems (ESS) with operational controllers provide flexibility and improved dispatchability [4], [37], [42]. The role of ESS in wind energy is reviewed in [5], with offshore applications benefiting from onshore ESS deployment due to economic and grid service considerations [43], [44].

As the energy mix evolves, OWFs must meet strict grid codes and provide ancillary services beyond capacity markets to ensure grid stability [45]. Market volatility affects renewable resource participation [46], [47]; however, larger projects can achieve enhanced profitability. OWFs can explore new revenue streams by actively participating in energy and ancillary service markets, enabled by technological advancements in power electronics [48], [49]. Farm-level control is crucial for profit maximization [50], and hybrid systems with storage offer significant opportunities in ancillary markets [51], as highlighted by Eguinoa et al. [52].

Related Work Optimization-based models for onshore wind farms with energy storage to participate in multiple energy markets have been extensively studied [53]–[56]. Notably, Hou et al. [57] were among the first to explore OWF market participation with energy storage, though their work focused on operational optimization without considering energy storage sizing. Recently, Bechlenberg et al. [58] addressed energy storage sizing for OWFs optimized for multiple electricity markets. Their work highlights the benefits of integrating OWFs with energy storage, employing a model predictive control approach to manage storage levels and correct prediction uncertainties. However, their ad-hoc approach to electric component sizing, such as cable capacity, could lead to sub-optimal system performance and excessive costs.

Varotto et al. [59] proposed an optimization strategy to evaluate economic benefits through optimal sizing and energy management of Battery Energy Storage Systems (BESS) for hybrid wind-solar offshore farms. A subsequent study [60] framed a multi-objective optimization problem to assess different storage technologies and locations, yet fixed storage sizes and overlooked control tuning for market participation.

Frequency support for the power grid from renewable sources such as OWFs is increasingly vital. This support can be delivered through wind turbines and integrated substations (VSC/MMC-HVDC) with proper control designs [61], [62]. We review control schemes and combinations extensively studied in the literature, particularly critiquing those in [61] for fixed sizing assumptions. Through analysis, we identified optimizing control parameters that could enhance OWFs' energy and frequency support participation alongside energy storage design sizing [63].

Our study focuses on doubly-fed induction generator (DFIG) wind turbines with deloading control, which offers significant kinetic energy storage for frequency support with an appropriate controller design. The deloading controller design aims to reserve a margin for grid support when needed. Rotor-Side Control (RSC) and Pitch Angle Control (PAC) control designs for DFIG turbines are discussed in [64]–[66], with coordinated control strategies detailed in [66], [67]. RSC is preferred for its response speed, particularly with over-speeding for safe operations. Our work examines the primary frequency support of wind farms using an RSC-Droop control strategy.

Contributions: Our literature review identifies two groups of studies: one focusing on design sizing and market participation but neglecting control parameter optimization, and the other emphasizing control design with fixed system design. We advocate for the simultaneous optimization of design sizing and control to improve OWFs' performance as energy market participants. This approach has shown promise in various engineering fields [16], including offshore wind turbines [68], wave energy converters [69], and Energy-Harvesting Ocean Kites [70]. Control co-design offers optimal performance in hybrid renewable systems [71], and new market participation formulations for hybrid systems have been proposed [72], [73]. These works highlight the importance of market participation in comprehensive assessment and planning, as emphasized by Eguinoa et al. [52]. Our work proposes a control co-design approach for optimizing OWF design and control parameters, focusing on market performance.

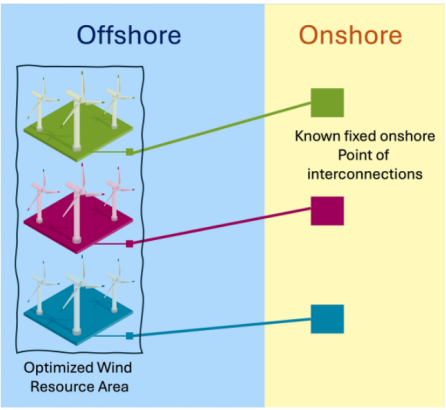
We extend the multi-timescale formulation from earlier studies [73]–[75] to a multi-timescale, multi-stage stochastic control co-design formulation. This accounts for offshore wind and energy price uncertainty. Addressing uncertainties is crucial for variable energy resources like OWFs. We introduce an optimization formulation to evaluate market participation under uncertainty, considering cable size and onshore energy storage capacity as primary design variables. We also optimize the turbine rotor speed droop-gain parameter and use a droop-based controller to integrate energy storage in reserve markets. This study fills a gap by integrating design and operational planning for OWFs with energy storage, assessing developer revenue. Additionally, we include a post-hoc validation step for the optimized droop gain control. The OWF model and AC-Grid Mini WECC models were developed in a transient EMT-based simulation platform, PSSE. A contingency scenario on the AC grid was simulated to compare estimated reserves against PSSE simulations with optimized droop parameters.

The report is organized as follows: Section 2.2 defines the problem statement and details the OWF case study. Section 2.3 outlines the mathematical models and control co-design optimization problem formulation for each step of the multi-scale simulation framework. Section 2.4 presents results, analysis, and validation against the baseline control co-design solutions under different inflation rate considerations, highlighting the role of energy storage sizing and operations. Section 2.4.7 summarizes findings, limitations, and future opportunities.

2.2 Problem Definition

This section examines the control co-design of a proposed OWF on the U.S. West Coast. This OWF is connected to an onshore point of interconnection (POI) via an HVDC cable configured

Schematic of OWFs and POIs



Names and Coordinates of POI and OWFs

POI Names	Latitude & Longitude		Distance	OWF Rated Power (MW)
	OWF Coordinates	POI Coordinates		
WCASCADE,WA	42.781, -124.685	47.347, -122.124	545.060	1500
JOHN DAY, OR	42.836, -124.660	45.678, -120.738	444.423	2350
COTTONWOOD, CA	42.705 -124.609	40.399, -122.265	324.193	1810
TESLA,CA	42.630 -124.558	37.712, -121.565	603.598	2640
MOSSLAND,CA	42.466, -124.510	36.903, -121.807	660.732	1800

Figure 11: Schematic of the optimized wind resource area and radial interconnection shown with tabulated details about individual farm capacity, POI location and distance from onshore.

in a radial topology. The system's topology is illustrated in Figure 11. The choice of the U.S. West Coast use case is inspired by the research conducted by Douville et al. [76]. An ESS is considered for installation onshore to support the OWF's operations and facilitate energy exchange with the grid [77]. The system, comprising the OWF and ESS, participates in the California Independent System Operator (CAISO) electricity markets, including day-ahead (DA) and real-time (RT) energy markets, as well as ancillary service markets by providing upward and downward reserves. For brevity, detailed discussions of these markets are omitted, but further information can be found in Harris et al. [77]. Figure 11 presents the coordinates of the POIs and OWFs, their distances, and the rated power of the OWFs, based on the use cases proposed by Douville et al. [76].

The physical design of the system emphasizes the sizing of the ESS and the HVDC export cables (interties). The primary objective of the control co-design is to minimize installation and material costs while maximizing revenue from energy and ancillary service markets throughout the ESS's operational lifetime. System operations consider power transmission through the HVDC cable and ESS charging/discharging capabilities. Control parameters, specifically the droop control parameters of the OWF and ESS, ensure stable and reliable power and service delivery. In this approach, both the OWF and ESS jointly provide the reserve. Furthermore, it is assumed that when the ESS is utilized for reserve provision, it only needs to do so for brief periods, thus minimally affecting the ESS's state of charge (SoC).

To address the challenge of uncertainty in decision-making, this problem is structured as a multi-stage stochastic optimization aimed at minimizing the system's expected net cost. The scenario tree, based on forecast or historical day-ahead (DA), real-time (RT), and up/down reserve prices, in addition to wind speed data, will be thoroughly discussed in Section 2.3.4.

2.3 Methodology

The schematic of the proposed methodology is shown in Fig 12. The methodology and framework are developed generically such that they can be applied for the simultaneous design and operational control of various energy systems, considering market participation. In this report, the proposed methodology will be described for the control co-design to enable energy market participation of the OWF with ESS. The following subsections present and discuss the details of individual steps outlined in the schematic diagram.

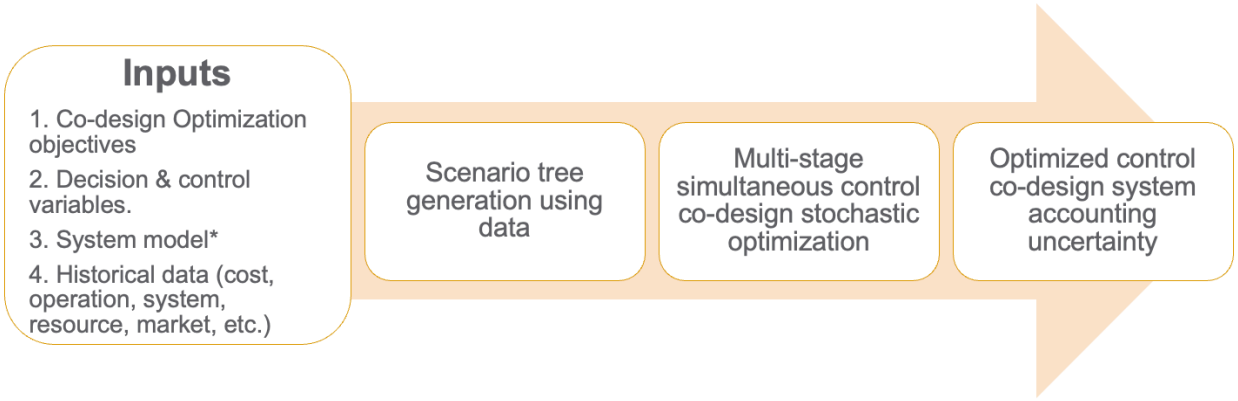


Figure 12: Schematic of the workflow for the proposed multi-time multi-stage stochastic control co-design methodology

2.3.1 Design and Control Decision Variables

Design Variables: The increasing penetration of renewable energy introduces variability and uncertainty, posing stability challenges for power system operations. ESS are effective and widely studied solutions to address these challenges. Proper sizing and optimal operation strategies of ESS integrated with renewable energy systems have shown promising results in enhancing the stability, controllability, and reliability of power systems. Various ESS technologies for wind power integration exist, including pumped hydro storage, compressed air energy storage, flywheel energy storage, and BESS.

In recent years, BESS has experienced significant cost reductions and performance improvements. Due to their rapid response time, high efficiency, substantial power and energy density, and convenient siting capability, BESS are widely used for renewable energy integration. Therefore, this study focuses on the design and sizing of lithium-ion batteries as an onshore integrated energy storage technology for OWFs. The PNNL-curated energy storage database [78] is utilized to provide cost and performance estimates for battery storage technologies.

OWFs are large-capacity energy resources that require efficient power transfer to onshore facilities. Export cable design and sizing are crucial for long-term reliable operations. OWFs, connected to onshore POIs via export cables, can benefit from interconnection or inter-linked export cable topology, as discussed in [38], [76]. These connections enable participation in multiple energy markets or facilitate power wheeling [79] from one region to another to avoid power curtailments. Consequently, sizing export cables solely based on OWF capacity may not be optimal in all scenarios. In this study, cable sizing is considered another design decision variable, accounting for market participation. For simplicity, a radial topology connecting OWFs to onshore through subsea export cables is considered, though the framework can accommodate other connection topologies with minor modifications. The optimization of export cable routing, environmental conditions, or geographic factors are not included in this formulation. Furthermore, dynamic thermal limits, cable aging, and reliability under varying operating conditions in the problem formulation are beyond the scope of the current work. Cable capacity size ranges and associated costs for the optimization problem were identified using data from [38], [76], [80]. This study does not consider design sizing for onshore and offshore converter-inverter substations or any other devices; thus, costs associated with such systems are held constant.

Control Variables: Doubly-fed induction generator (DFIG) turbines with partial-scale power

electronic converters in their rotor circuit are employed, while the stator is directly connected to the power grid. DFIG-based turbines offer various advantages over fixed-generator technologies, including speed control within limited ranges, improved energy conversion efficiency, and active and reactive power capabilities. The control system for DFIG-based offshore wind turbines, depicted in Fig. 13, comprises two principal components: Grid-Side Control (GSC) and Rotor-Side Control (RSC). The GSC is responsible for maintaining the DC voltage across the converters by adjusting the active power dispatched to the grid and regulates AC voltage through reactive power control. Conversely, the RSC maintains rotor-side AC voltage and optimizes power extraction from wind conditions.

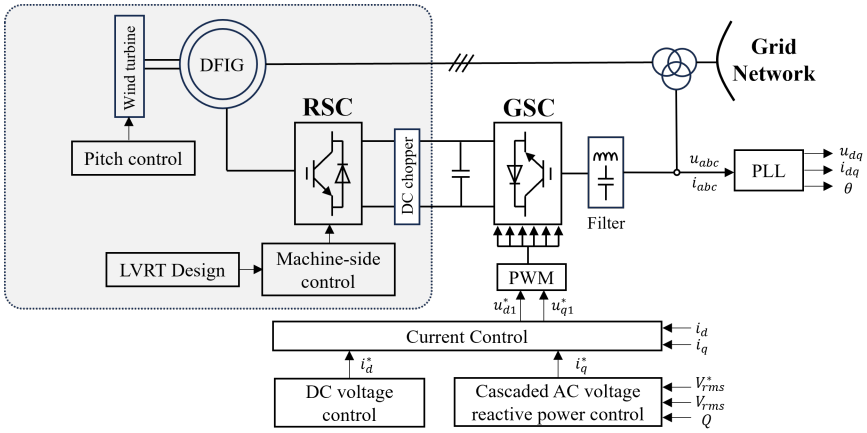


Figure 13: The control diagram of an offshore wind turbine, highlighting the Rotor-Side Control (RSC). The droop gain (k) associated with the RSC controller is optimized to provide the required reserve for primary frequency market participation.

To enhance primary frequency support, the RSC employs a de-loading strategy, establishing a supplementary power reference in response to grid frequency variations. This approach is formalized by the droop relationship outlined in Equation (23), which allocates a power reserve ΔP from the maximum power point to counter frequency deviations on the grid, in accordance with the guidelines specified in [67].

$$\Delta P = k\Delta f = \frac{1}{R}\Delta f \tag{23}$$

where R denotes the speed droop parameter used to share common load changes among multiple generator units.

Despite implementing droop control in the RSC, offshore wind turbines cannot provide frequency support under low or no wind conditions. To overcome this limitation, each OWF is paired with a Grid-Forming Lithium-ion BESS and equipped with a primary droop controller to ensure continuous frequency support capability [81], [82]. The droop controller follows the same frequency support logic as outlined in (23), with its droop gain k jointly optimized alongside the RSC's droop gain.

2.3.2 Energy Market: Day-Ahead, Real-Time, and Reserve

The system participates in Day-Ahead (DA), Real-Time (RT), and ancillary service markets. Before each day begins, operators determine the power allocation for the DA market based on its price, without prior knowledge of RT, up/down reserve prices, or wind speed. Throughout the

day, power and reserve allocations for the remaining markets are adjusted based on observed prices and weather conditions at each time step. In this study, a one-hour interval is assumed for the DA market, and a 15-minute interval for the other markets. The OWF typically allocates most of its power to the DA market, while the RT market is primarily used to mitigate mismatches between DA market participation and actual power generation. The formulation does not explicitly model penalties for under-performance in the day-ahead market, though such penalties are implicitly reflected through changes in revenue.

Predicting market power prices and wind speeds is beyond the scope of this report; thus, historical data is utilized for simulations. CAISO price data for the NP15 region from 2018 and 2022 were retrieved from the CAISO OASIS Portal. The CAISO's OASIS API offers real-time and historical data on Locational Marginal Prices (LMPs) for both DA and RT markets. This data includes various market metrics such as energy prices, congestion costs, and losses, updated frequently, with some available on a sub-hourly basis. The year 2018 represents a typical period before the COVID-19 pandemic, while 2022 is the latest year with both price and wind speed data available. Probability density functions of DA and RT market prices for these two years are illustrated in Figure B.1(a-b), showing a significant energy price increase in 2022 due to rising natural gas costs [83]. The impact of high prices on system design and operation will be discussed in Section 2.4.

2.3.3 Uncertainties Affecting Offshore Wind Farms

The development and operation of OWFs encounter significant uncertainties that affect their safety, security, and overall viability. Wind speed and market price are primary factors among these. These uncertainties directly impact the OWF's output and system revenue, thus influencing the required size of the ESS and its capacity. Additionally, effective management of power injection into the DA market, OWF and ESS droop control parameters, and ESS operation is essential to ensure stable and reliable power supply while maximizing total profits.

Numerous optimization strategies address the challenges posed by these uncertainties in wind farm planning and operation. Techniques such as stochastic optimization [84], robust optimization [85], and distributionally robust optimization [86] have been applied. Considering that the lifespan of a BESS typically exceeds ten years, evaluating the long-term expected revenue of the OWF yields more effective design outcomes. Furthermore, design and market participation decisions are often not made simultaneously. Therefore, a multistage stochastic optimization framework is employed for this problem, which assesses uncertainties at various decision-making stages and optimizes the overall expected performance.

2.3.4 Scenario Generation Scheme

The scenario generation process is essential for modeling and optimizing the performance of offshore wind farms (OWFs), enabling robust decision-making under uncertainty. This process involves creating representative scenarios for key variables like wind speed and energy market prices, capturing their inherent variability to facilitate effective optimization and control. The scenario generation scheme is designed to balance accuracy with computational tractability through a structured approach.

The methodology begins with the collection and normalization of historical data. Appropriate probability distributions, such as the Weibull distribution for wind speed, are then identified, and their goodness-of-fit is validated using statistical methods like the Kolmogorov-Smirnov test. Subsequently, Monte Carlo simulation is employed to generate a large set of scenarios by sampling from the identified distribution. To maintain computational efficiency, this extensive

dataset is reduced using clustering techniques, which group similar scenarios to select a representative subset.

These generated scenarios are then applied to optimize OWF layouts and operations. By simulating various configurations under these conditions, robust designs that maximize energy production and market participation while minimizing risk can be identified. A simplified block diagram of this scheme is presented in Figure 14, with detailed algorithmic descriptions provided in appendix-D as Algorithm 3.

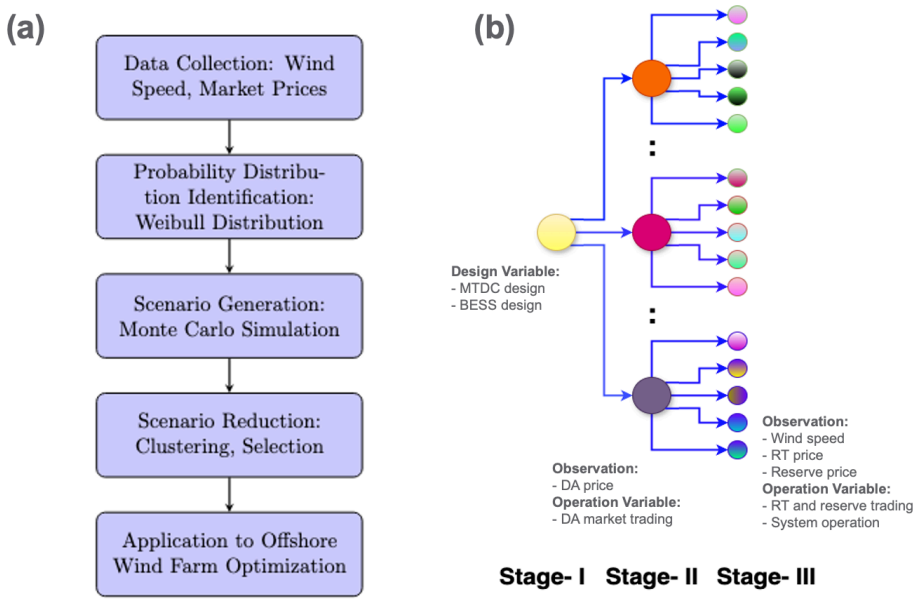


Figure 14: (a) Scenario generation scheme for offshore Wind Farm Optimization. (b) Schematic scenario tree with defining the multi-stage decision variables.

2.3.5 Multi-time, Multi-stage Stochastic Control Co-design Formulation

This section discusses the details of the proposed control co-design multi-time scale, multi-stage stochastic formulation. The development of this methodology has been emphasized earlier in Section 4.1.

2.3.6 Multi-scale Formulation

It is common for different markets to clear at varying frequencies, and wind speed forecast data may not align with these markets. In this report, DA price and wind speed data are considered hourly, while RT and up/down reserve prices are provided every 15 minutes.

Following [74], $\mathcal{T}_1 = \{1, 2, \ldots, 24\}$ and $\mathcal{T}_2 = \{1, 2, \ldots, 96\}$ are used to represent the index sets of each hour and each 15-minute interval in a day, respectively. Additionally, the set $\mathcal{T}_{21}(t) = \{4(t-1) + k \mid k = 1, 2, 3, 4\}$ contains all four 15-minute indices in hour t. This approach can be easily extended to accommodate more time scales and other time resolutions.

2.3.7 Multi-stage Formulation with Random Variables and Uncertain Parameters

There are three stages in this problem: design, DA, and RT. Random events at DA and RT stages are represented as ω^{DA} and ω^{RT}, respectively. When making the design decision, no uncertain parameter is observable. At the DA stage, DA price $\boldsymbol{\lambda}^{\mathsf{DA}}(\omega^{\mathsf{DA}})$ for the next 24 hours is cleared and known. At the RT stage, all other uncertain parameters are revealed, including wind power $\mathbf{P}^{\mathsf{OWF}}(\omega^{\mathsf{DA}}, \omega^{\mathsf{RT}})$ extracted from OWF by maximum power point tracking, RT price $\boldsymbol{\lambda}^{\mathsf{RT}}(\omega^{\mathsf{DA}}, \omega^{\mathsf{RT}})$, up reserve price $\boldsymbol{\lambda}^{\mathsf{ResU}}(\omega^{\mathsf{DA}}, \omega^{\mathsf{RT}})$, and down reserve price $\boldsymbol{\lambda}^{\mathsf{ResD}}(\omega^{\mathsf{DA}}, \omega^{\mathsf{RT}})$.

2.3.8 Control Co-design Formulation

This section provides the detailed formulation of the problem. Let $\mathbb{E}_{\omega^{\cdot}}(X)$ denote the expectation of X over all scenarios of $\omega^{\cdot}$. The model is as follows:

$$\min \quad R^{\mathsf{tax}}\left(\lambda^{\mathsf{E}} sz^{\mathsf{E}} + \lambda^{\mathsf{C}} sz^{\mathsf{C}}\right) - R^{\mathsf{Year}} \mathbb{E}_{\omega^{\mathsf{DA}}}\left[\mathcal{F}\left(sz^{\mathsf{E}}, sz^{\mathsf{C}}; \omega^{\mathsf{DA}}\right)\right] \tag{24}$$

$$\text{s.t.} \quad 0 \le sz^{\mathsf{E}} \le 0.05 P^{\mathsf{WFR}}, \; sz^{\mathsf{C}} \ge 0 \tag{25}$$

$$(92-93)$$

In the objective function (73), the first term represents the installation cost of the ESS and HVDC transmission cable, while the second term is the present value of the expected revenue from the OWF over the lifetime of the ESS. Specifically, R^{tax} is a tax credit coefficient for new, qualified clean energy properties. Coefficients λ^{E} and λ^{C} represent the unit costs of ESS and cable, respectively. Variables sz^{E} and sz^{C} denote the rated power of ESS and HVDC cable, respectively. R^{Year} is an annuity factor that converts expected daily revenue to present value over the ESS's lifetime. If the ESS has a lifetime of Y years and the inflation rate is r, then $R^{\mathsf{Year}} = 365\frac{(1+r)^Y - 1}{r(1+r)^Y}$. Constraints (25) limit the ESS to below 5% of OWF rated power P^{WFR} [1]. This limit allows for sufficient battery size exploration while accounting for the cost of larger batteries. Constraints (92-93) pertain to the ESS's expected SoC at the end of the day and its expected operational cycles, which will be discussed later.

For given rated power of ESS and HVDC cable $(sz^{\mathsf{E}}, sz^{\mathsf{C}})$ and a realization of DA scenario ω^{DA}, $\mathcal{F}\left(sz^{\mathsf{E}}, sz^{\mathsf{C}}; \omega^{\mathsf{DA}}\right)$ calculates the daily revenue of the OWF as follows:

$$\begin{aligned}\mathcal{F}\left(sz^{\mathsf{E}}, sz^{\mathsf{C}}; \omega^{\mathsf{DA}}\right) = \max \; & \sum_{t \in \mathcal{T}_1} \lambda_t^{\mathsf{DA}}\left(\omega^{\mathsf{DA}}\right) p_t^{\mathsf{WD}} \\ & + \mathbb{E}_{\omega^{\mathsf{RT}}|\omega^{\mathsf{DA}}}\left[\mathcal{G}\left(sz^{\mathsf{E}}, sz^{\mathsf{C}}, \mathbf{p}^{\mathsf{WD}}; \omega^{\mathsf{DA}}, \omega^{\mathsf{RT}}\right)\right]\end{aligned} \tag{26}$$

The first and second terms represent DA and expected RT stage revenues, respectively. Variable p_t^{WD} represents the power traded to the DA market at time t. For given rated power of ESS and HVDC cable $(sz^{\mathsf{E}}, sz^{\mathsf{C}})$, power traded to DA market $\mathbf{p}^{\mathsf{WD}}$, and realizations of DA scenario ω^{DA} and RT scenario ω^{RT}, $\mathcal{G}(sz^{\mathsf{E}}, sz^{\mathsf{C}}, \mathbf{p}^{\mathsf{WD}}; \omega^{\mathsf{DA}}, \omega^{\mathsf{RT}})$ calculates the RT stage revenue of the OWF. Its objective function is:

$$\begin{aligned}\max \quad & \sum_{t \in \mathcal{T}_2} \lambda_t^{\mathsf{RT}}\left(\omega^{\mathsf{DA}}, \omega^{\mathsf{RT}}\right) p_t^{\mathsf{WR}} + \sum_{t \in \mathcal{T}_2} \lambda_t^{\mathsf{ResU}}\left(\omega^{\mathsf{DA}}, \omega^{\mathsf{RT}}\right)\left(p_t^{\mathsf{ResWU}} + p_t^{\mathsf{ResBU}}\right) \\ & + \sum_{t \in \mathcal{T}_2} \lambda_t^{\mathsf{ResD}}(\omega^{\mathsf{DA}}, \omega^{\mathsf{RT}})(p_t^{\mathsf{ResWD}} + p_t^{\mathsf{ResBD}})\end{aligned} \tag{27}$$

[1] Considering the large capacity of OWF, a higher percentage value would limit practicability and require additional analysis beyond the scope of this work.

These three terms calculate RT, up reserve, and down reserve revenues, respectively. Variable p_t^{WR} represents the power traded to the RT market at time t; p_t^{ResWU} and p_t^{ResBU} denote the up reserve provided by OWF and ESS, respectively; and p_t^{ResWD} and p_t^{ResBD} are the down reserve provided by them.

Problem $\mathcal{G}(sz^{\mathsf{E}}, sz^{\mathsf{C}}, \mathbf{p}^{\mathsf{WD}}; \omega^{\mathsf{DA}}, \omega^{\mathsf{RT}})$ is subject to a series of constraints. The power at the onshore and offshore sides must be balanced and transmitted through the HVDC cable.

$$p_{t_1}^{\mathsf{W}} = V_{1,t_2}(V_{1,t_2} - V_{2,t_2})g \ \forall t_1 \in \mathcal{T}_1,\ t_2 \in \mathcal{T}_{21}(t_1) \tag{28}$$

$$(p_{t_2}^{\mathsf{dis}} - p_{t_2}^{\mathsf{ch}}) - (p_{t_2}^{\mathsf{WR}} + p_{t_1}^{\mathsf{WD}}) = V_{2,t_2}(V_{2,t_2} - V_{1,t_2})g \ \forall t_1 \in \mathcal{T}_1,\ t_2 \in \mathcal{T}_{21}(t_1) \tag{29}$$

$$1.1\ (V_{1,t}(V_{1,t} - V_{2,t})g + p_t^{\mathsf{ResWU}}) \leq sz^{\mathsf{C}} \ \forall t \in \mathcal{T}_2 \tag{30}$$

$$\underline{V}_i \leq V_{i,t} \leq \overline{V}_i \ \forall i \in \{1,2\},\ t \in \mathcal{T}_2 \tag{31}$$

In these constraints, parameter g is the conductance of the transmission cable; $\underline{V}_i$ and $\overline{V}_i$ are the lower and upper voltage limits of bus i, respectively. Here, $i = 1$ refers to the offshore bus and $i = 2$ refers to the onshore bus. At each time t, variable p_t^{W} represents the power actually generated by the OWF; $V_{i,t}$ is the voltage of node i; p_t^{dis} and p_t^{ch} denote the discharged and charged power of the ESS, respectively. Constraints (79) and (80) ensure the power balance on the offshore and onshore sides, respectively. Constraint (81) limits the power transmitted through the HVDC cable, accounting for a 10% additional factor of safety. Voltage limits are imposed by constraint (82).

As stated in Section 2.3.2, the RT market is primarily used to mitigate the difference between DA market participation and actual OWF generation. To prevent excessive power trading in the RT market to exploit price differences between these two markets, the following limits are imposed:

$$p_{t_2}^{\mathsf{WR}} \leq \mathsf{max}\{P_{t_1}^{\mathsf{OWF}}(\omega^{\mathsf{DA}}, \omega^{\mathsf{RT}}) - P_{t_1}^{\mathsf{OWF}}(\omega^{\mathsf{DA}}), 0\} \ \forall t_1 \in \mathcal{T}_1,\ t_2 \in \mathcal{T}_{21}(t_1) \tag{32}$$

$$p_{t_2}^{\mathsf{WR}} \geq \mathsf{min}\{P_{t_1}^{\mathsf{OWF}}(\omega^{\mathsf{DA}}, \omega^{\mathsf{RT}}) - P_{t_1}^{\mathsf{OWF}}(\omega^{\mathsf{DA}}), 0\} \ \forall t_1 \in \mathcal{T}_1,\ t_2 \in \mathcal{T}_{21}(t_1) \tag{33}$$

Here, $P_t^{\mathsf{OWF}}(\omega^{\mathsf{DA}}) = \mathbb{E}_{\omega^{\mathsf{RT}}}[P_t^{\mathsf{OWF}}(\omega^{\mathsf{DA}}, \omega^{\mathsf{RT}})]$. These two constraints ensure that the power traded in the RT market remains within the difference between the actual and expected wind power in each DA scenario.

Considering up and down reserves, the OWF's actual output should stay within the following range:

$$p_{t_1}^{\mathsf{W}} + p_{t_2}^{\mathsf{ResWU}} \leq P_{t_1}^{\mathsf{OWF}}(\omega^{\mathsf{DA}}, \omega^{\mathsf{RT}}) \ \forall t_1 \in \mathcal{T}_1,\ t_2 \in \mathcal{T}_{21}(t_1) \tag{34}$$

$$p_{t_1}^{\mathsf{W}} - p_{t_2}^{\mathsf{ResWD}} \geq 0 \ \forall t_1 \in \mathcal{T}_1,\ t_2 \in \mathcal{T}_{21}(t_1) \tag{35}$$

As discussed in Section 2.3.1, the following droop relationships apply when the OWF and ESS provide fast frequency support:

$$p_t^{\mathsf{ResWU}} = k_t^{\mathsf{ResW}} \Delta f_{\mathsf{max}}^U \ \forall t \in \mathcal{T}_2 \tag{36}$$

$$p_t^{\mathsf{ResWD}} = k_t^{\mathsf{ResW}} \Delta f_{\mathsf{max}}^D \ \forall t \in \mathcal{T}_2 \tag{37}$$

$$p_t^{\mathsf{ResBU}} = k_t^{\mathsf{ResB}} \Delta f_{\mathsf{max}}^U \ \forall t \in \mathcal{T}_2 \tag{38}$$

$$p_t^{\mathsf{ResBD}} = k_t^{\mathsf{ResB}} \Delta f_{\mathsf{max}}^D \ \forall t \in \mathcal{T}_2 \tag{39}$$

$$k_{t_2}^{\mathsf{ResW}} \geq P_{t_1}^{\mathsf{OWF}}(\omega^{\mathsf{DA}}, \omega^{\mathsf{RT}})/R_{max} \ \forall t_1 \in \mathcal{T}_1,\ t_2 \in \mathcal{T}_{21}(t_1) \tag{40}$$

$$k_{t_2}^{\mathsf{ResW}} \leq P_{t_1}^{\mathsf{OWF}}(\omega^{\mathsf{DA}}, \omega^{\mathsf{RT}})/R_{min} \ \forall t_1 \in \mathcal{T}_1,\ t_2 \in \mathcal{T}_{21}(t_1) \tag{41}$$

$$sz^{\mathsf{E}}/R_{max} \leq k_t^{\mathsf{ResB}} \leq sz^{\mathsf{E}}/R_{min} \ \forall t \in \mathcal{T}_2 \tag{42}$$

$$k_{t_2}^{\mathsf{ResW}} + k_{t_2}^{\mathsf{ResB}} \geq P_{t_1}^{\mathsf{OWF}}(\omega^{\mathsf{DA}}, \omega^{\mathsf{RT}})/R_{all} \ \forall t_1 \in \mathcal{T}_1,\ t_2 \in \mathcal{T}_{21}(t_1) \tag{43}$$

Parameters Δf^U_{max} and Δf^D_{max} represent the maximum up and down frequency deviations under planned disturbances, respectively. Variables k_t^{ResW} and k_t^{ResB} are the droop control parameters of the OWF and ESS, respectively. Constraints (36-39) allocate up and down reserves to the OWF and ESS based on droop control parameters and maximum frequency deviations. Constraints (40-43) limit the percentage of OWF and ESS primary frequency support. Additionally, constraint (43) ensures that each OWF, when combined with its ESS, provides joint mandatory primary frequency support to the main grid.

The ESS has the following dynamic and state constraints:

$$SoC_t - SoC_{t-1} = (\eta^{\text{ch}} p_t^{\text{ch}} - \eta^{\text{dis}} p_t^{\text{dis}}) T_2^{\text{length}} \ \forall t \in \mathcal{T}_2 \backslash \{1\} \tag{44}$$

$$SoC_1 - 0.5 sz^{\text{E}} \text{DurH} = (\eta^{\text{ch}} p_1^{\text{ch}} - \eta^{\text{dis}} p_1^{\text{dis}}) T_2^{\text{length}} \tag{45}$$

$$p_t^{\text{ch}} + p_t^{\text{ResBD}} \leq sz^{\text{E}} \quad \forall t \in \mathcal{T}_2 \tag{46}$$

$$p_t^{\text{dis}} + p_t^{\text{ResBU}} \leq sz^{\text{E}} \quad \forall t \in \mathcal{T}_2 \tag{47}$$

$$0 \leq SoC_t \leq sz^{\text{E}} \text{DurH} \quad \forall t \in \mathcal{T}_2 \tag{48}$$

Parameters η^{ch} and η^{dis} represent the charging and discharging efficiency of the ESS, respectively; DurH denotes the duration hour of the ESS; T_2^{length} is the time length of the time indices in set $\mathcal{T}_2$, which is 1/4 hour here. Variable SoC_t is the state of charge of the ESS. The ESS is assumed to be initially half charged. Constraints (85-86) are ESS dynamic equations, and (89-91) impose charging, discharging, and SoC limits, respectively.

Finally, the following two constraints apply. Constraint (92) requires the expected SoC of the ESS to be half of its energy capacity at the end of the day. Constraint (93) ensures its expected total discharge power does not exceed the daily cycle limit cyl^{lim}.

$$\mathbb{E}_{\omega^{\text{RT}}}[SoC_{|\mathcal{T}_2|}] = 0.5 sz^{\text{E}} \text{DurH} \tag{49}$$

$$\mathbb{E}_{\omega^{\text{RT}}}\left[\sum_{t \in \mathcal{T}_2} p_t^{\text{dis}} \eta^{\text{dis}} T_2^{\text{length}}\right] \leq sz^{\text{E}} \text{cyl}^{\text{lim}} \tag{50}$$

2.4 Results

2.4.1 Baseline, CCD, and Comparison Cases Setting

To illustrate the advantages of utilizing the Control Co-Design (CCD) model discussed in the previous section and to explore the benefits of pairing the ESS with the OWF, results for the following cases were computed and compared. The droop parameters are set as: $R_{min} = 0.1$ and $R_{max} = 0.5$ for wind farms; $R_{min} = 0.01$ and $R_{max} = 0.5$ for ESS; and $R_{all} = 0.2$. A total of 100 scenarios were generated, comprising 20 DA stage scenarios and 5 RT stage scenarios for each DA stage scenario.

- Base: ESS rated power sz^{E} is set to 2% of OWF rated power; k^{ResB} and k^{ResW} are fixed at their upper limits for providing maximum frequency support; one or more HVDC cables with a fixed rated power of 2600 MW are used; 3% inflation is applied; and wind and price data are from 2018.

- CCD18-3,18-5, and 18-8: CCD model with 3%, 5%, and 8% inflation rates [2]; wind and price data are from 2018.

[2] Inflation rate choices were based on historical trends identified through tradingeconomics.

- CCD22-3,22-5, and 22-8: CCD model with 3%, 5%, and 8% inflation rates; wind and price data are from 2022.
- No reserve: This case is identical to the CCD18-3 case, but no reserve is provided for frequency support.
- No ESS: This case is identical to the CCD18-3 case, but no ESS is installed.

2.4.2 Design Results and Discussion

The design results for the ESS and the net profits of OWFs are detailed in Fig. 15(a), while the cable design results are shown in Fig. 15(b). The cost for ESS is presented in Fig. 15(c). To ensure consistent comparison, a total cost of $855.4 million for both onshore and offshore converters is included in all design cases.

An examination of the first four cases chosen (CCD18-3, CCD18-5, CCD18-8, and the Base case) in Fig. 15(a) indicates that the Base case, which provides the largest reserve for frequency support, yields the lowest profit. Conversely, the no-reserve case, which allocates all power to market sales, achieves the highest profit compared to the other four cases across all POIs. A general observation is that a higher inflation rate adversely affects net profit for all POIs in both 2018 and 2022 data sets. However, significantly higher profits are observed in cases using 2022 data due to much higher energy prices compared to 2018.

Except for the base case, all other cases utilize the smallest onshore export cable size that satisfies the minimum requirement specified in constraint (81), thereby reducing costs. Figure 16(a) shows the comparison of cable material and installation costs for all other cases versus the base case. The two color bar ranges are used to display material and installation costs, respectively. Additionally, Fig. 16(b) highlights the difference in export cable capacity (MW) between the baseline case and other cases.

2.4.3 ESS Sizing Analysis

Analysis of solution data for all design cases and different offshore wind farm POIs revealed two critical values for ESS size: 3% and the upper limit of 5% of the OWF rated power. In most scenarios, the ESS size is set at one of these two values. The size of the ESS primarily depends on the revenue it can generate during daily operations. On one hand, the system charges the ESS when energy prices are low and discharges it during peak price hours. This revenue strategy relies on the energy price differential across one or more consecutive days. On the other hand, if the ESS provides the reserve for frequency support, the OWF can allocate more power to market sales rather than retaining generation capacity, with this revenue influenced by the absolute energy price.

When energy prices vary within specific ranges, as observed in cases CCD18-3, CCD18-5, and CCD22-8, constraints (40) and (43) are binding, enabling the OWF to send most of its power to the markets. An ESS sized at 3% of the OWF's rated power can adequately meet the reserve requirements specified in constraint (43). Figure 15(b) indicates that if the revenue generated by the ESS does not justify its cost, the system opts for a smaller size, as evidenced by case CCD18-8. Conversely, higher revenue potential supports a larger ESS size, even reaching the upper limit, as demonstrated in cases CCD22-3 and CCD22-5.

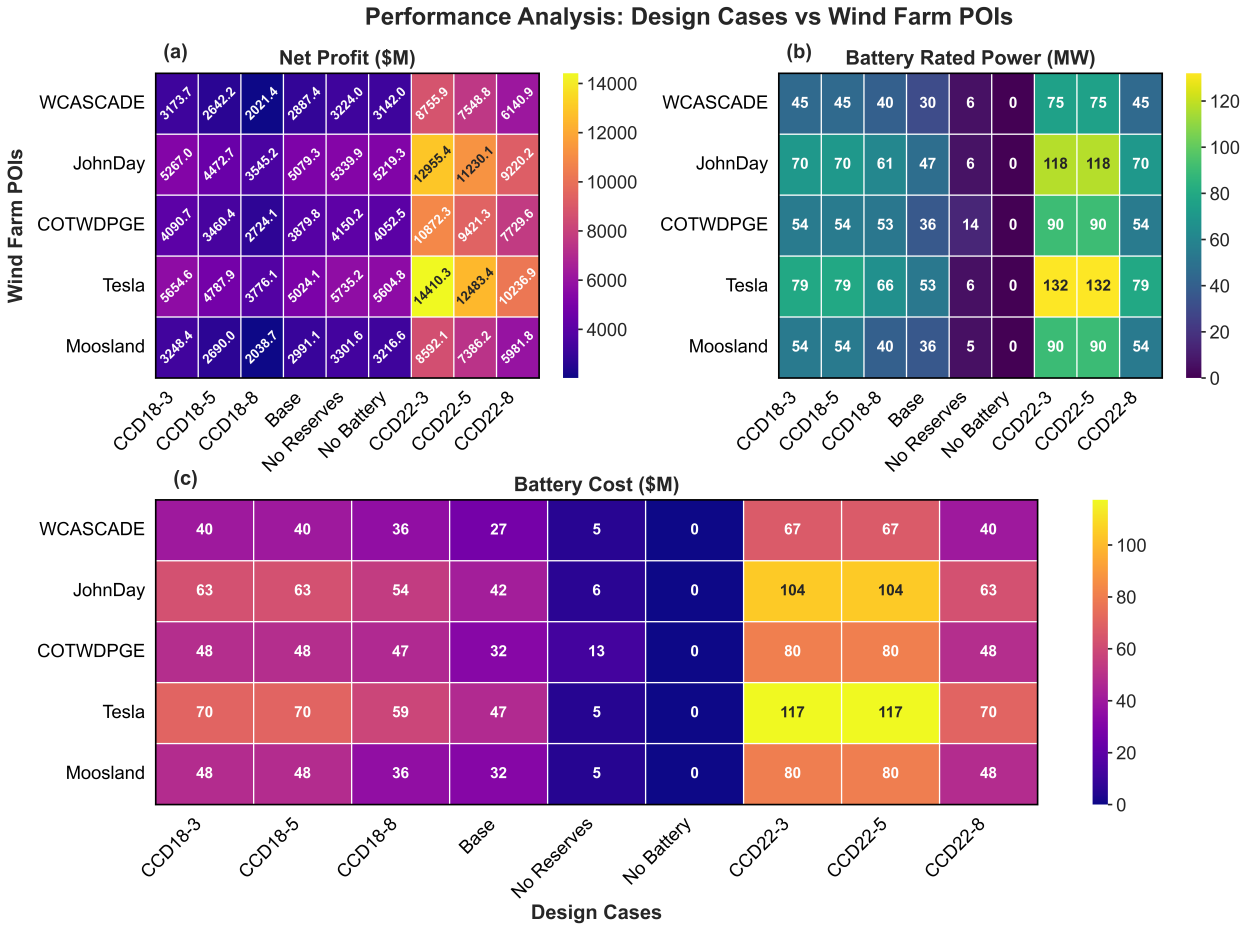


Figure 15: Comparison of (a) Net profit (b) Battery rated power (c) Battery cost for each design case, across different case studies and wind farm POI's.

2.4.4 Market Participation and ESS Operation

To analyze system operations and market participation strategies, detailed information on three typical scenarios—80, 83, and 99—at the POI COTWDPGE in case CCD18-3 is presented in Fig. 17(a-i). Figures 17(a-c) display DA and RT market prices throughout the day. Figures 17(d-f) illustrate the offshore wind farm (OWF) output and power traded in the two markets, with positive values indicating sales to the markets and negative values indicating purchases from the markets. Figures 17(g-i) show the power discharged from the ESS, where a negative value indicates that the ESS is being charged. For Figures 17(d-f), the highlighted background denotes time periods during the day when the RT market price is higher than the DA market price.

In all scenarios, prices exhibit two peaks: a smaller one around 5:00 AM to 7:00 AM and a larger one around 5:00 PM to 7:00 PM. The ESS is typically charged during midday, approximately from 10:00 AM to 3:00 PM, and discharged around the two peak periods, particularly the higher peak. The DA market participation decision considers prices and OWF output across all scenarios within the same DA stage. Participation in the RT market and joint ESS operation helps reconcile discrepancies between actual OWF generation and DA market trades. In scenario 80, the RT price is lower than or close to the DA price, leading to DA market

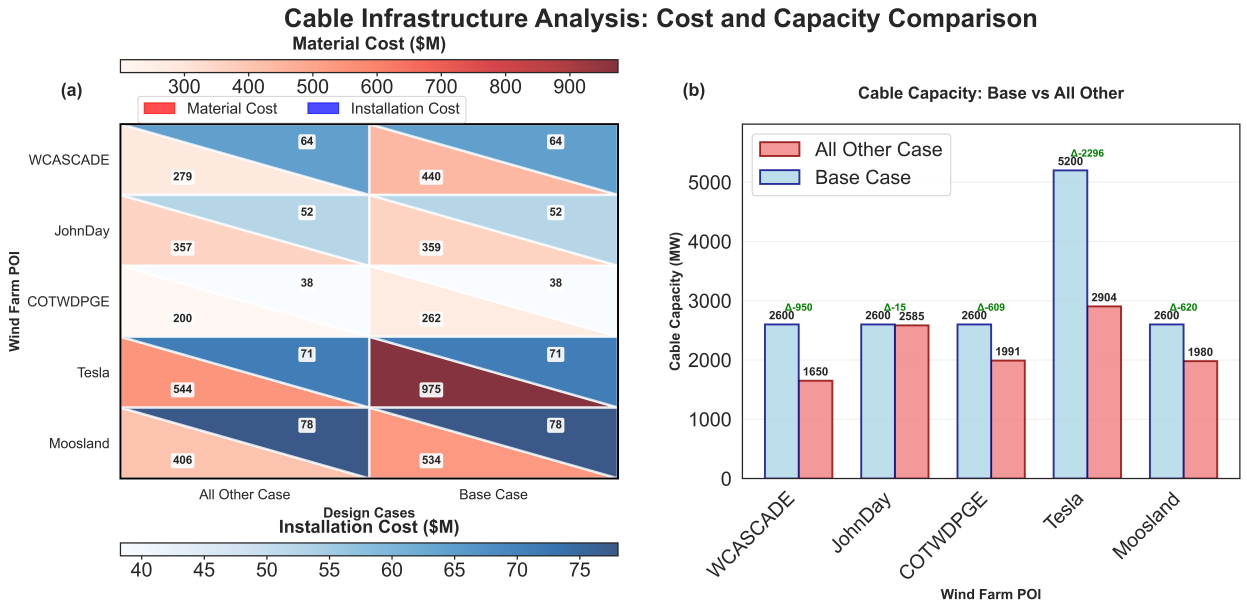


Figure 16: (a) Material and installation cost comparison for all other cases (CCD-18-*, CCD22-*, No reserve, No Battery) relative to the base-case setting. (b) Comparison of export cable capacity (MW) differences between the baseline case and other cases.

trades nearly matching OWF generation and minimal RT market activity. In scenario 83, the RT price exceeds the DA price, with less power allocated to the DA market than OWF generation, allowing surplus for the RT market. Even with reduced generation from 9:00 AM to 4:00 PM, only DA market trades are decreased, while RT trades remain consistent, and the ESS is charged. In scenario 99, mitigation through the RT market is more pronounced. From 12:00 AM to 5:00 AM, without OWF generation, power is purchased from the RT market to satisfy DA commitments. Between 11:00 AM and 1:00 PM, surplus power is sold to the RT market to maintain balance.

2.4.5 Revenue Analysis

Figure 18 displays the expected power traded and revenue from DA and RT markets in a day for the four cases using the same input data and inflation rates. The average unit revenue is derived by dividing total revenue from DA and RT markets by total energy traded. A brief analysis of observations for each design case, compared to the CCD18-3 design case, is provided as follows:

Base Case: This scenario features a smaller ESS size and provides greater reserve support through both the OWF and ESS. With the OWF holding more reserve, total market-traded power and revenue are reduced. However, a similar average unit revenue to CCD18-3 is achieved through ESS operation.

No Reserve Case: Without needing to reserve OWF capacity for frequency support, this scenario trades more power and achieves higher total revenue. As seen from Fig. 18, a relatively small ESS is employed for output shifting, resulting in lower average unit revenue.

No ESS Case: In the absence of ESS, frequency support is solely from the OWF, reducing total market-traded power and revenue. RT market trading, without ESS for output shifting, manages the DA commitment and OWF generation mismatches, resulting in lower average unit

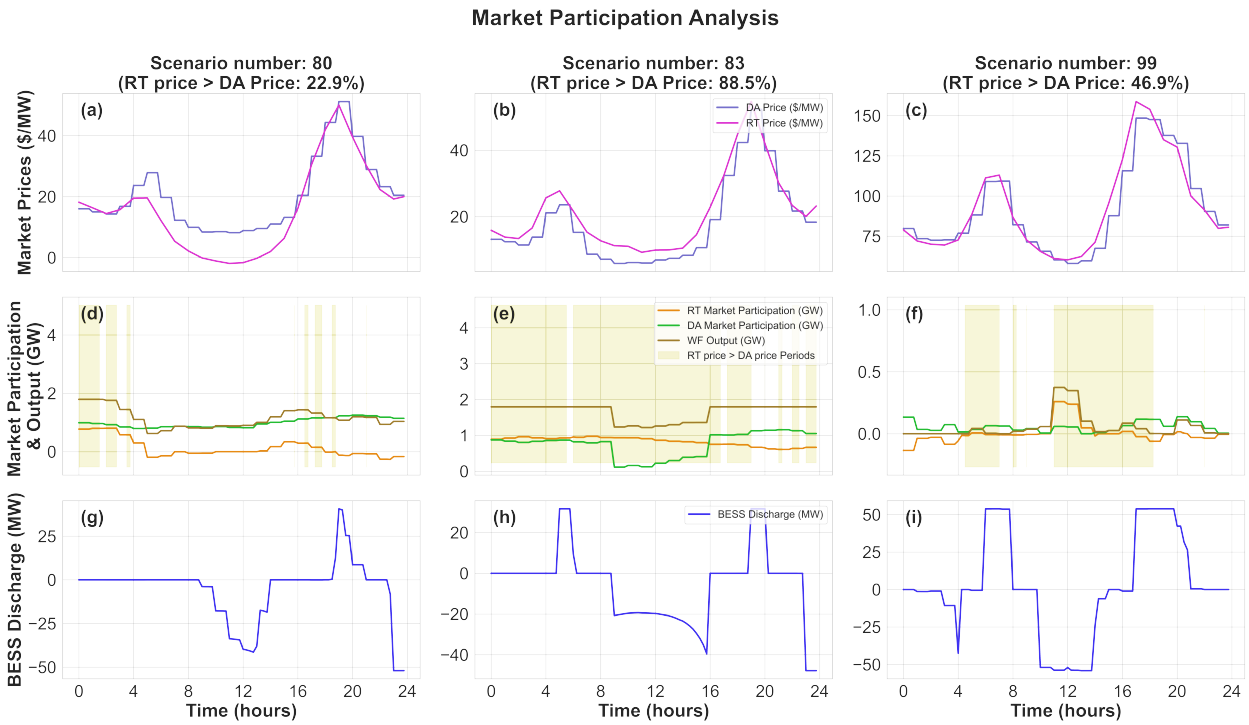


Figure 17: Market price (DA/RT), market participation, and ESS operation of COTWDPGE in scenario 80 of case CCD18-3

revenue than CCD18-3.

2.4.6 Validation of Control Co-design Offshore Wind Farm Using Simulations (PSCAD)

In Figure 19, the droop control variable k at COTWDPGE for scenario 80 in case CCD18-3 is plotted. The pink band illustrates the distribution of k droop for all scenarios. For most of the time, the ESS variable is at its upper limit. By adjusting the droop status of the wind farm and ESS, the system optimizes its output while ensuring sufficient reserve for frequency support.

To further validate the dynamic performance of the co-design results, a time-domain electromagnetic transient (EMT) simulation was conducted using PSCAD [35]. Both Maximum Power Point Tracking (MPPT) and de-loading control strategies were integrated and tested under the scenario of losing the largest generator at Palo Verde, Arizona, at 60 seconds, which had an output of 2182.3 MW before tripping. Additional details on the PSCAD model can be found in E. Transient trajectories of power and frequency, depicted in Figures 20(a-b) and 20(c), were then analyzed to verify the stability and effectiveness of the co-design results.

Overall, the miniWECC system integrating OWFs remains stable both before and after a generator trip on the main grid. Prior to the 60-second mark, the steady-state output of the four OWFs shows only minor differences compared to the co-design results. After the disturbance, the power output and the increase in power are both smaller than those of the MPPT curves, indicating sufficient headroom. Additionally, the frequency nadir is lower when the OWF operates in MPPT mode without providing frequency support. Consequently, the co-design approach demonstrates superior transient performance, which is crucial given the increasing

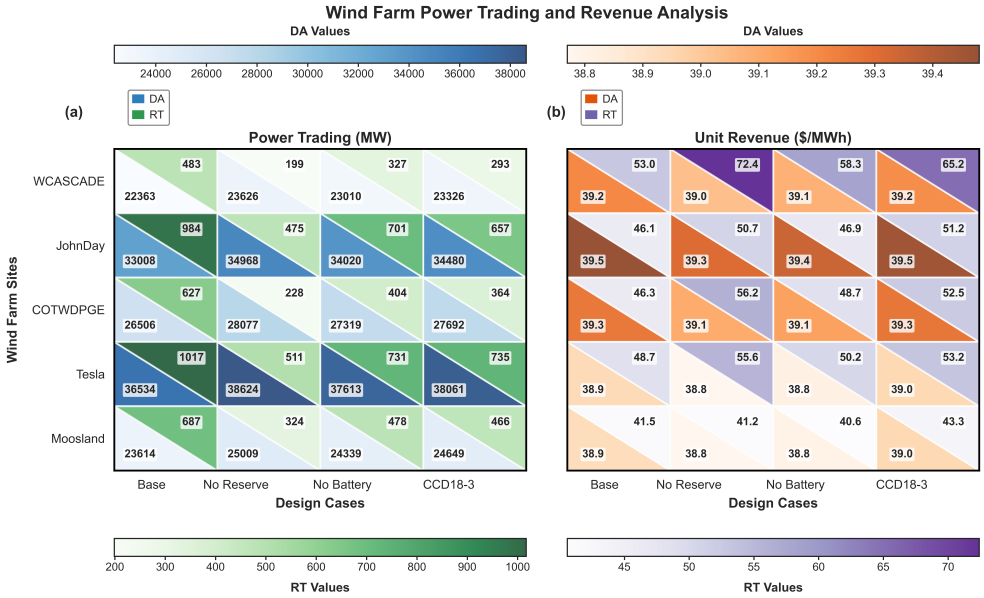


Figure 18: Comparison of selected design cases vs wind farm POI's for (a) Power Traded and (b) Unit revenue of market participation.

penetration of wind and other inverter-based resources.

2.4.7 Conclusions and Future Work

Simultaneous consideration of operations, controls, and market participation during the planning of renewable energy systems, such as offshore wind farms, can achieve long-term reliability and optimal performance. Incorporating various uncertainties into the decision-making process enhances overall system resilience. In this report, a multi-timescale, multi-stage stochastic control co-design methodology is developed and presented. This approach models the multi-timescale nature of energy market participation and addresses uncertainties associated with renewable resources and prices at the interconnection point for revenue maximization. The approach is demonstrated using an offshore wind farm grid integration use case. Here, the sizing of HVDC power export cables from offshore to onshore is optimized as part of the process to prevent under-sizing and over-sizing, which could significantly impact project investment costs. Furthermore, energy storage sizing is considered, addressing grid stabilization needs and incorporating market participation. Both offshore wind farm and energy storage participation in ancillary service markets are evaluated for additional revenue opportunities for wind farm owners. The proposed control co-design approach optimizes cable and energy storage sizing, and market participation with droop parameters, showing improved annual revenue for five different capacity wind farms studied. A scenario tree-based stochastic formulation incorporates uncertainty. Additionally, the framework allows for the evaluation of various inflation rates and the implications of energy storage on overall revenue. The study provides validation for droop control parameters through transient PSSE simulations of an offshore wind farm integrated with the miniWECC AC grid model, showing comparable reserve values and performance in meeting the integrated power grid's frequency requirements.

Currently, reliance is placed on historical wind speed and price data, as offshore wind farm development in the U.S. is still maturing. Future efforts could enhance price data estimation

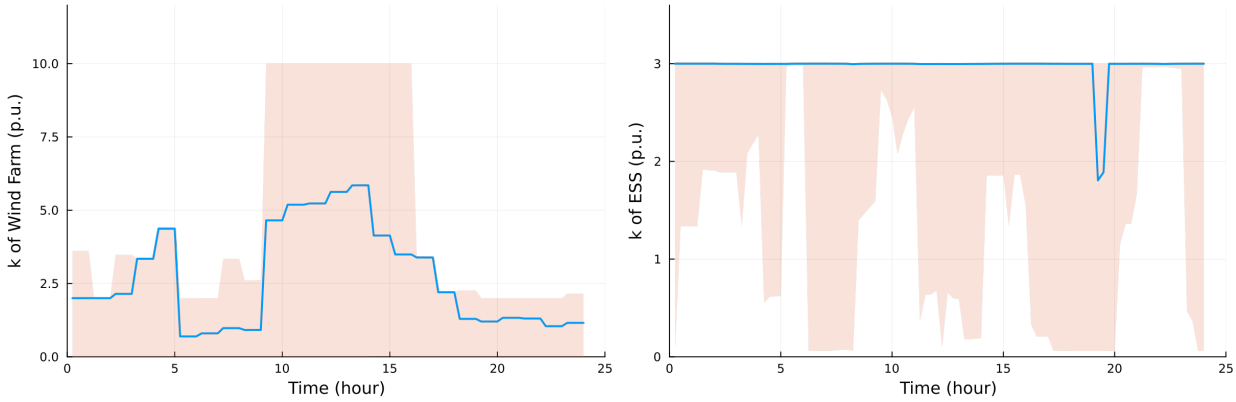


Figure 19: Droop control variable k at COTWDPGE for scenario 80 in case CCD18-3.

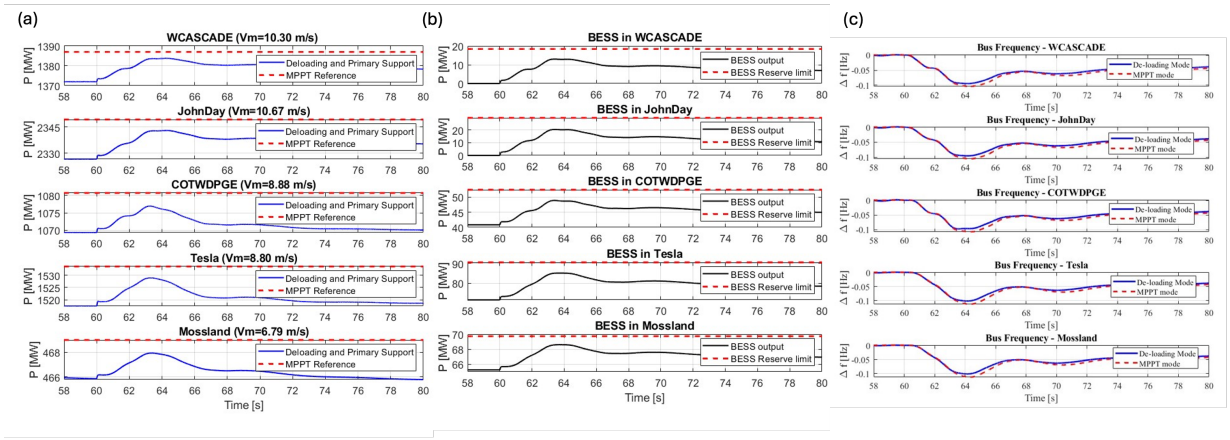


Figure 20: Power curves of 240-bus miniWECC system in PSCAD: (a) OWF; (b): BESS (c) POI frequency curves of 240-bus miniWECC system in PSCAD.

using detailed production cost planning with offshore wind farms and consider futuristic climate scenarios for wind speed data. Furthermore, different network topologies (e.g., meshed) could be explored for power wheeling, with incentive designs providing additional revenue streams for offshore wind farm owners. The study can also be expanded by including additional substation-level design and control variables and incorporating reduced-order models of dynamic converter-inverters. As energy market participation becomes increasingly important for future renewable energy developers, the proposed framework could be extended to next-generation energy market models to estimate revenue and develop participation strategies.

3.0 A Learning-to-Optimize Method for Constrained Parametric Bilevel Problems

3.1 Motivation

Bilevel optimization problems are prevalent in diverse applications, spanning economics and game theory [87], operations management and logistics [88], and engineering system design [89]. Unfortunately, these problems are generally difficult to solve; they are typically NP-hard and may lack efficient frameworks for approximate solutions, especially depending on their specific characteristics [90]. This complexity presents a significant challenge for applications requiring *real-time or repeated solutions*. However, many real-world scenarios demand repeated solutions of related problem instances, such as when economic or engineering design decisions must be made across a variety of scenarios and objectives.

This report introduces a novel framework that applies deep learning to solve a broad class of bilevel problems. These problems include *coupling constraints*, which are constraints that bind both upper- and lower-level decision variables. Our approach employs differentiable optimization solvers at the problem's lower level, enabling gradient-based training of neural networks to approximate the upper-level solution. It also includes internal *correction* routines to enforce coupling constraints between the upper- and lower-level problems. The resulting models are trained to map the parameters of a bilevel problem to its optimal solution. This approach is primarily motivated by optimal control co-design problems, a setting that involves optimizing engineering systems with economic objectives at the upper level, subject to system dynamics determined by optimal control problems at the lower level. We demonstrate how this proposed framework can efficiently generate design solutions in response to varying design objectives and criteria, as encoded by the problem parameters.

Contributions. The main contributions of this report are: **(1)** A novel Learning to Optimize (L2O) method for bilevel optimization problems, which is based on using differentiable optimization to ensure lower-level solution optimality and differentiable projections for upper-level constraint satisfaction. **(2)** We show how these differentiable optimizations can also be used to create internal gradient-based correction mechanisms for satisfying *coupling constraints*. These constraints significantly complicate bilevel problems by linking their upper- and lower-level variables. **(3)** We demonstrate the proposed framework on a collection of challenging bilevel design and control problems. This includes problems with *nonconvex* lower-level components and complex, high-dimensional constraints. In particular, we show its ability to learn high-quality approximators in the optimal control co-design setting, which enables efficient generation of design solutions for varying design criteria.

3.2 Related Work

Bilevel optimization methods. Bilevel optimization problems are inherently challenging, and generally lack efficient solution methods except in specific cases. For sufficiently small problems with a convex structure, common solution methods typically involve single-level reformulations (via KKT conditions or optimal value functions) that can be solved using mixed-integer programming [90]. When a special structure is present, such as linearity at both levels or the absence of coupling constraints, gradient methods, often relying on penalty methods, have been proposed [91]–[93]. Recently, [94], [95] addressed a specific class of bilevel problems using upper-level gradient descent combined with a quadratic programming (QP)-based safety filter to enforce the KKT conditions of the lower-level problem. The complex

nature of more general bilevel problems has led to increased interest in metaheuristics, such as evolutionary algorithms [96], [97] and particle swarm optimization [98], over classical alternatives. A comprehensive review of modern approaches can be found in [99].
Learning to Optimize (L2O). L2O is a subfield of machine learning focused on learning *fast approximations* for challenging optimization problems. One distinct branch aims to learn information that accelerates the convergence of classical solvers [100]. In continuous optimization, this includes predicting active constraints [101], optimization problem parameters [102], stepsizes [103], primal variables [104], [105], and non-Euclidean metrics [106]. For integer variables, L2O involves branching rules [107], cutting planes [108], and variable partitions in large neighborhood search [109]. A separate branch focuses on training ML models to predict optimal solutions directly from a problem representation [110]. A key challenge in this area is maintaining the feasibility of predicted solutions to arbitrary constraints. Proposed solutions include differentiable projections [111], reparameterization tricks [112], [113], dual-variable estimation [114], [115], and gradient-based constraint correction routines [116]. Finally, some recent works have proposed using ML to accelerate bilevel optimization. [117] introduces learning an optimal value reformulation from solved non-parametric examples using a ReLU network embedded in an MIP solver. [118], [119] propose learning solutions of a parametric bilevel program directly, though without any constraints at either level. This report extends the L2O toolkit for directly learning solutions to bilevel optimization problems with continuous variables and a full set of constraints, including coupling constraints.

3.3 Problem Setting

The goal of this report is to learn how to solve parametric bilevel optimization problems represented by (51), where each component may depend on a vector of problem parameters $\boldsymbol{p}$. We use the convention that optimization variables are written as function arguments, while problem parameters are written as subscripts. We consider a pair of problems:

$$\mathcal{B}(\boldsymbol{p}) := \underset{\boldsymbol{y}}{\text{argmin}} \quad \mathcal{L}_{\boldsymbol{p}}(\ \boldsymbol{y}, \boldsymbol{z}\ ) \tag{51a}$$

$$\text{s.t.} \quad \boldsymbol{z} \in \mathcal{O}_{\boldsymbol{p}}(\boldsymbol{y}) \tag{51b}$$

$$\boldsymbol{y} \in \mathcal{C}_{\boldsymbol{p}} \tag{51c}$$

$$\boldsymbol{U}_{\boldsymbol{p}}(\boldsymbol{y}, \boldsymbol{z}) \leq \boldsymbol{0}, \tag{51d}$$

where

$$\mathcal{O}_{\boldsymbol{p}}(\boldsymbol{y}) := \underset{\boldsymbol{z}}{\text{argmin}} \quad l_{\boldsymbol{p}}(\ \boldsymbol{y}, \boldsymbol{z}\ ) \tag{52a}$$

$$\text{s.t.} \quad \boldsymbol{z} \in \mathcal{S}_{\boldsymbol{p}}(\boldsymbol{y}). \tag{52b}$$

The defining feature of problem (51) is its constraint (51b), which requires the variables $\boldsymbol{z}$ to be the solution to another optimization problem (52), which in turn depends on the variables $\boldsymbol{y}$. We refer to problem (51) as the *upper-level problem*, and (52) as the *lower-level problem*, with $\boldsymbol{y}$ and $\boldsymbol{z}$ being the *upper-* and *lower-level* variables, respectively. Our objective is to learn a fast approximator that solves the coupled problems (51,52) across a distribution of problem parameters, denoted as $\mathcal{P}$.

3.3.0.1 Classes of Constraints.

We distinguish three sets of constraints at the upper level. Condition (51c) constrains only the upper-level variables. Constraint (51b) prescribes $\boldsymbol{z}$ as a solution to the lower-level problem (52)

resulting from $\boldsymbol{y}$. Additionally, the *coupling constraints* (51d) significantly complicate the solution of (51). They impose further conditions on the relationship between upper- and lower-level variables in (51), preventing solution concepts based on their separation or decoupling [90]. A large portion of algorithms for bilevel programming cannot accommodate problems with coupling constraints [99]. To the best of our knowledge, no previous work has proposed an L2O framework for parametrically learning to solve problems with coupling constraints.

3.3.0.2 Conditions on the Problem Form.

The proposed framework aims to learn solutions for a broad class of problems in the form of (51,52). However, it relies on a key condition: for all $\boldsymbol{y} \in \mathcal{C}_{\boldsymbol{p}}$ and $\boldsymbol{p} \sim \mathcal{P}$, the solution to the lower-level problem (52) must be *unique* whenever it *exists*. This allows the lower-level problem to define a function. Our proposal further assumes that this function is *differentiable*, a condition that may not always be strictly met in practice. Nevertheless, many well-behaved optimization problems of interest define continuous, almost-everywhere differentiable functions, which are sufficient for effective end-to-end training [111]. Such pseudo-differentiable components are common in deep learning architectures, similar to ReLU activations. While existence of the lower-level solution is not strictly required for all $\boldsymbol{y} \in \mathcal{C}_{\boldsymbol{p}}$ and $\boldsymbol{p} \sim \mathcal{P}$, this condition is assumed to hold for clarity of exposition until it is relaxed using a reformulation trick, introduced in Section 3.5.2. The functions $\mathcal{L}_{\boldsymbol{p}}$ will act as the loss function of a deep network and should therefore be continuous in $\boldsymbol{y}$ and $\boldsymbol{z}$, while $\boldsymbol{U}_{\boldsymbol{p}}$ should be smooth. Following common practice in L2O, these conditions serve as practical guidelines for applying the proposed architecture rather than formal requirements.

3.4 Learning to Solve Bilevel Optimization

This section presents a self-supervised method for learning to solve the parametric bilevel optimization problems described in Section 3.3. Specifically, it trains an ML model to approximate the mapping (51), from problem parameters $\boldsymbol{p}$ to *upper-level* solutions $\boldsymbol{y}^\star = \mathcal{B}(\boldsymbol{p})$. Let $\hat{\mathcal{B}}_\theta$ denote this hypothetical model, with weights θ. We assume a training set of n_p problem instances, each specified by a vector of problem parameters $\{\boldsymbol{p}_{(i)}\}_{i=1}^{n_p}$ drawn from the distribution $\mathcal{P}$. A training procedure for $\hat{\mathcal{B}}_\theta$ should minimize the objective function $\mathcal{L}_{\boldsymbol{p}}(\hat{\boldsymbol{y}})$ attained by its predicted solutions $\hat{\boldsymbol{y}} = \hat{\mathcal{B}}_\theta(\boldsymbol{p})$ in expectation over $\mathcal{P}$, leading to the empirical risk minimization (ERM) goal:

$$\min_{\boldsymbol{\theta}} \quad \mathop{\mathbb{E}}_{\boldsymbol{p}\sim\mathcal{P}} \left[\mathcal{L}_{\boldsymbol{p}}(\hat{\boldsymbol{y}}, \hat{\boldsymbol{z}})\right] \tag{53a}$$

$$\text{s.t.} \quad \hat{\boldsymbol{z}} \in \mathcal{O}_{\boldsymbol{p}}(\boldsymbol{y}) \qquad \forall \boldsymbol{p} \sim \mathcal{P} \tag{53b}$$

$$\hat{\boldsymbol{y}} \in \mathcal{C}_{\boldsymbol{p}} \qquad \forall \boldsymbol{p} \sim \mathcal{P} \tag{53c}$$

$$\boldsymbol{U}_{\boldsymbol{p}}(\hat{\boldsymbol{y}}, \hat{\boldsymbol{z}}) \leq \boldsymbol{0} \qquad \forall \boldsymbol{p} \sim \mathcal{P}, \tag{53d}$$

$$\text{where} \quad \hat{\boldsymbol{y}} \coloneqq \hat{\mathcal{B}}_\theta(\boldsymbol{p}). \tag{53e}$$

Equation (53e) simply defines $\hat{\boldsymbol{y}}$ as the output of $\hat{\mathcal{B}}_\theta$ given $\boldsymbol{p}$. Constraints (53c, 53b, 53d) require that each such output and its resulting pair $(\hat{\boldsymbol{y}}, \hat{\boldsymbol{z}})$ constitutes a feasible solution to the bilevel problem (51,52). Learning solutions subject to such complex constraints is inherently challenging, as any predicted $\hat{\boldsymbol{y}}$ and its corresponding $\hat{\boldsymbol{z}}$ are unlikely to satisfy the coupling constraint (53d) after solving (53b), even if $\hat{\boldsymbol{y}} \in \mathcal{C}_{\boldsymbol{p}}$ as required in (53c).

This section proposes an architecture for the model $\hat{\mathcal{B}}_\theta$ and a method for training it to approximate the ERM goal (53). Its core concept draws on modern techniques for *differentiable optimization*. It adopts a differentiable subroutine that iteratively refines predicted solutions to satisfy coupling constraints while maintaining feasibility to other constraints at each step. Before prescribing the full training method, the main architectural components of $\hat{\mathcal{B}}_\theta$ are introduced.

3.5 Satisfying Constraints with Differentiable Optimization Modules

Recent work in machine learning has focused on developing differentiable optimization solvers, which treat a problem's optimal solution as a differentiable function of its parameters. Their Jacobian matrices are generally modeled by implicit differentiation of an optimality condition at its optimum, such as KKT conditions [120], [121] or a fixed-point condition [122], [123]. The following proposal is agnostic to the specific implementation used. Differentiable optimization is often used to employ optimization problems as modules in trainable machine learning models, and two such modules are used within $\hat{\mathcal{B}}_\theta$.

3.5.0.1 Differentiable Solution of the Lower-Level Problem.

The assumption of uniqueness at the lower level defines an implicit relationship between upper- and lower-level solutions. Thus, the overall bilevel problem is viewed as primarily finding $\boldsymbol{y}^\star := \mathcal{B}(\boldsymbol{p})$. For a given $\boldsymbol{p}$ and any predicted upper-level solution $\hat{\boldsymbol{y}}$, a differentiable solver for problem (52) produces $\hat{\boldsymbol{z}} = \mathcal{O}_{\boldsymbol{p}}(\hat{\boldsymbol{y}})$ along with $\frac{\partial \hat{\boldsymbol{z}}}{\partial \hat{\boldsymbol{y}}}$. This provides a means by which feasible solution pairs can be computed and back-propagated, ensuring that (53b) is always satisfied throughout training.

3.5.0.2 Differentiable Projection at the Upper Level.

In addition to satisfying constraint (53b) with its solution pair, any candidate $\hat{\boldsymbol{y}}$ must also satisfy its own constraint (53c). To achieve this, a projection operator is employed, which solves the following subproblem:

$$\boldsymbol{\Pi}_{\mathcal{C}}(\boldsymbol{y}) = \underset{\boldsymbol{w} \in \mathcal{C}}{\operatorname{argmin}} \, \|\boldsymbol{y} - \boldsymbol{w}\|_2^2. \tag{54}$$

Differentiable projections are essential for constraint satisfaction in L2O [106], [111], [124]. They can be implemented in a differentiable library such as *cvxpylayers* [122] or through automatic differentiation in PyTorch when a closed-form solution is available (e.g., projection onto non-negative constraints using a ReLU function). While other mechanisms could ensure constraint satisfaction [125], [126], the projection operator is chosen for its role in the algorithm presented next, leading to the *projected gradient descent* method, which has well-studied convergence properties [127].

3.5.1 End-to-End Trainable Architecture

This section presents the complete architecture and training routine of a model $\hat{\mathcal{B}}_\theta$ that learns to solve problem (53). Let $\mathcal{N}_\theta$ be a deep neural network with weights θ, which predicts initial estimates $\hat{\boldsymbol{y}} = \mathcal{N}_\theta(\boldsymbol{p})$ of an upper-level solution. Composing $\mathcal{N}_\theta$ with $\boldsymbol{\Pi}_{\mathcal{C}_{\boldsymbol{p}}}$ ensures feasibility to (53c), and further composition with $\mathcal{O}_{\boldsymbol{p}}$ produces a solution pair $(\hat{\boldsymbol{y}}, \mathcal{O}_{\boldsymbol{p}}(\hat{\boldsymbol{y}}))$ that satisfies (53b) and (53c). However, this does not necessarily satisfy the coupling constraint $\boldsymbol{U}_{\boldsymbol{p}}(\hat{\boldsymbol{y}}, \hat{\boldsymbol{z}}) \leq \boldsymbol{0}$ in

(53d). Nonetheless, the differentiability of $\mathcal{O}_{\boldsymbol{p}}$ and $\boldsymbol{\Pi}_{\mathcal{C}_{\boldsymbol{p}}}$ provides a descent direction toward its feasible region. To this end, the *Coupling Constraint Violation* is defined as follows, recognizing $\hat{\boldsymbol{y}}$ as the independent variable and $\hat{\boldsymbol{z}} = \mathcal{O}_{\boldsymbol{p}}(\hat{\boldsymbol{y}})$ as dependent:

$$\nu(\hat{\boldsymbol{y}}) \coloneqq \text{ReLU}\left(\boldsymbol{U}(\ \hat{\boldsymbol{y}}, \mathcal{O}_{\boldsymbol{p}}(\hat{\boldsymbol{y}})\ )\right). \tag{55}$$

The gradient of its Euclidean norm is equal to:

$$\nabla_{\boldsymbol{y}} \|\nu(\boldsymbol{y})\|^2 = 2\nu(\boldsymbol{y}) \frac{d\nu}{d\boldsymbol{y}} = 2\nu(\boldsymbol{y}) \frac{d\nu}{d\boldsymbol{U}} \left[\frac{\partial \boldsymbol{U}}{\partial c} + \frac{\partial \boldsymbol{U}}{\partial \mathcal{O}_{\boldsymbol{p}}} \frac{\partial \mathcal{O}_{\boldsymbol{p}}}{\partial \boldsymbol{y}}\right], \tag{56}$$

whose nontrivial component is the Jacobian $\frac{\partial \mathcal{O}_{\boldsymbol{p}}}{\partial \boldsymbol{y}}$. This information represents backpropagation through the lower-level problem, which can be obtained from one of the differentiable solvers discussed previously. Automatic differentiation in PyTorch [128] is sufficient to complete the remaining chain rule calculations in (56).

A function that reduces $\|\nu(\boldsymbol{y})\|^2$ by performing a gradient descent step of size γ can also be defined:

$$\mathcal{G}(\boldsymbol{y}) \coloneqq \boldsymbol{y} - \gamma \nabla_{\boldsymbol{y}} \|\nu(\boldsymbol{y})\|_2^2. \tag{57}$$

Crucially, this function can also be made differentiable by leveraging functionality for back-propagating gradient calculations (in this case, equation (56)) in automatic differentiation libraries such as PyTorch [129]. The result of function (57) is generally infeasible to (53c); this can be addressed by a (differentiable) projection back onto $\mathcal{C}_{\boldsymbol{p}}$, completing one step of an end-to-end differentiable *Coupling Constraint Correction* routine:

$$\boldsymbol{y}_{k+1} = \boldsymbol{\Pi}_{\mathcal{C}_{\boldsymbol{p}}}\left(\boldsymbol{y}_k - \gamma \nabla \|\nu(\boldsymbol{y}_k)\|_2^2\right). \tag{58}$$

Letting $\boldsymbol{y}_0 = \mathcal{N}_\theta(\boldsymbol{p})$, the architecture for $\hat{\mathcal{B}}_\theta$ follows the prediction of a neural network with m steps of (58). Explicitly,

$$\hat{\mathcal{B}}_\theta(\boldsymbol{p}) = \left[\left(\boldsymbol{\Pi}_{\mathcal{C}_{\boldsymbol{p}}} \circ \mathcal{G}\right)^m \circ \mathcal{N}_\theta\right](\boldsymbol{p}). \tag{59}$$

By construction, this model is end-to-end differentiable and maintains feasibility to (53b) and (53c) while iteratively moving toward satisfaction of (53d). Furthermore, process (58) can be recognized as the classical projected gradient descent method on $\|\nu(\boldsymbol{y})\|_2^2$, which is known to converge to local minima of convex and nonconvex functions under specific conditions [127]. The correction routine (58) is illustrated in Figure 21, where blue arrows represent gradient steps (57), and alternating green arrows represent projections (54) back onto $\mathcal{C}_{\boldsymbol{p}}$. The entire chain of operations is composed with the neural network $\mathcal{N}_\theta$ and unrolled in backpropagation to update its predictions.

3.5.1.1 Convergence of the Correction and Soft Loss.

The proposed architecture builds on the concept of a gradient-based constraint correction mechanism, popularized by the ”DC3” algorithm for L2O [116]. Gradient descent methods are not universally guaranteed to converge, and thus gradient-based constraint corrections are not guaranteed to yield zero violations. However, as noted in [116], when initialized close to an optimum, they are highly effective in practice [130], [131]. Following [116], a ”soft” penalty loss is prescribed:

$$\mathcal{L}^{\text{soft}}_{\boldsymbol{p}_{(i)}}(\hat{\boldsymbol{y}}, \hat{\boldsymbol{z}}) \coloneqq \mathcal{L}_{\boldsymbol{p}_{(i)}}(\hat{\boldsymbol{y}}, \hat{\boldsymbol{z}}) + \lambda \|\nu(\hat{\boldsymbol{y}})\|_2^2. \tag{60}$$

This ensures that upper-level predictions from $\mathcal{N}_\theta$ are initialized close to satisfying the coupling constraint (51d), so that its end-to-end training with the differentiable correction (58) tends to produce feasible solutions to the full problem (51).

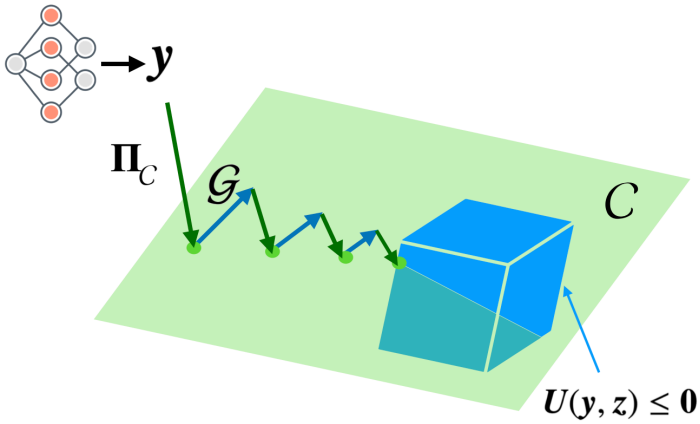


Figure 21: Illustrating the differentiable coupling constraint correction. Gradient steps drive predictions toward $\boldsymbol{U}(\boldsymbol{y}, \boldsymbol{z}) \leq \mathbf{0}$ (blue), while projections maintain them in $\mathcal{C}$ (green).

3.5.1.2 Training Routine.

An overall training scheme is summarized in Algorithm 2, in terms of one epoch of stochastic gradient descent. Each data input sample $\boldsymbol{p}_{(i)}$ represents a distinct instance of problem (51), for which an initial solution estimate $\hat{\boldsymbol{y}}$ is predicted at line 3. The sequence (58) makes up lines 4-8. For each of m *correction steps*, $\hat{\boldsymbol{y}}$ is iteratively refined by taking a step toward feasibility of the coupling constraint (51d). The gradient $\hat{\boldsymbol{g}}$ at line 6 is calculated per equation (56). The loss (60) is then evaluated with respect to the refined estimate $\hat{\boldsymbol{y}}$ and its lower-level pair. If needed, at test time, more than m iterations may be applied in the correction routine. Note that Algorithm 2 describes only the *forward pass* of the training routine. Line 11 encapsulates backpropagation through all components of the model, implemented with a combination of automatic differentiation and implicit differentiation via the differentiable solvers, as described above.

Algorithm 2 Learning Bilevel Optimization with Coupling

Input: parameters $\{\boldsymbol{p}_{(i)}\}_{i=1}^{N}$, weights $\boldsymbol{\theta}$, learning rate α, correction stepsize γ

for $i = 1$ **to** N **do**

  $\hat{\boldsymbol{y}} \leftarrow \mathcal{N}_{\boldsymbol{\theta}}(\boldsymbol{p}_{(i)})$ **for** $k = 1$ **to** m **do**

    $\hat{\boldsymbol{g}} \leftarrow \nabla \left\| \text{ReLU}\left(\boldsymbol{U}_{\boldsymbol{p}_{(i)}}(\hat{\boldsymbol{y}}, \mathcal{O}_{\boldsymbol{p}_{(i)}}(\hat{\boldsymbol{y}}))\right)\right\|^2$ $\hat{\boldsymbol{y}} \leftarrow \hat{\boldsymbol{y}} - \gamma \cdot \hat{\boldsymbol{g}}$ $\hat{\boldsymbol{y}} \leftarrow \square_{\mathcal{C}_{\boldsymbol{p}_{(i)}}}(\hat{\boldsymbol{y}})$

  end

  $\hat{\boldsymbol{z}} \leftarrow \mathcal{O}_{\boldsymbol{p}_{(i)}}(\hat{\boldsymbol{y}})$ $\boldsymbol{g} \leftarrow \nabla_{\theta} \mathcal{L}^{\text{soft}}_{\boldsymbol{p}_{(i)}}(\hat{\boldsymbol{y}}, \hat{\boldsymbol{z}})$ $\boldsymbol{\theta} \leftarrow \boldsymbol{\theta} - \alpha \cdot \boldsymbol{g}$

end

3.5.2 Satisfying Lower-Level Constraints

So far, this section has assumed that the lower-level problem (52) had at least one feasible solution for any $\boldsymbol{p}$ and $\boldsymbol{y} \in \mathcal{C}_{\boldsymbol{p}}$ (see comments on existence and uniqueness in Section 3.3). Otherwise, the optimization problem at line 5 of Algorithm 2 may be infeasible. In such cases, this issue is addressed by reformulating the overall bilevel problem to ensure this property is

BQP	$\left\lvert\frac{\mathcal{L}(\hat{\boldsymbol{y}},\hat{\boldsymbol{z}})-\mathcal{L}(\boldsymbol{y}^\star,\boldsymbol{z}^\star)}{\mathcal{L}(\boldsymbol{y}^\star,\boldsymbol{z}^\star)}\right\rvert$	$\lVert\nu(\boldsymbol{y})\rVert_2$	Time (s)	YALMIP (s)
3×2	$9.2e-4\pm 2e-3$	$5.9e-3\pm 1e-2$	$\mathbf{6.4e-2}$	0.12
6×4	$2.0e-3\pm 5e-3$	$2.8e-4\pm 9e-4$	$\mathbf{6.4e-2}$	1.2
9×6	$1.1e-2\pm 1e-2$	$4.0e-5\pm 7e-4$	$\mathbf{6.7e-2}$	10.3

Table 3: On BQP problems, Test Set Average

Model	$\mathcal{L}(\boldsymbol{y})$		$\lVert\nu(\boldsymbol{y})\rVert_2$		Time (s)	
	TT	HVAC	TT	HVAC	TT	HVAC
Learned (Alg. 2)	$\mathbf{0.122}$	$\mathbf{1.23}$	$\mathbf{1.4e-2}\pm 2e-2$	$3.0e{-2}\pm 5e-2$	$\mathbf{2.7e-2}$	$\mathbf{7.70e-2}$
PSO (baseline)	0.140	1.46	$1.6e-2\pm 2e-2$	$3.0e{-2}\pm 2e-2$	1268.7	1055.9

Table 4: On Control Co-Design, Test Set Average

satisfied. This is done by identifying lower-level constraints that prevent feasibility and promoting them to the upper level, where they become coupling constraints.

Let the lower level's feasible set be partitioned as $\mathcal{S}_{\boldsymbol{p}}(\boldsymbol{y}) = \mathcal{F}_{\boldsymbol{p}}(\boldsymbol{y}) \cup \mathcal{F}^C_{\boldsymbol{p}}(\boldsymbol{y})$, such that $\mathcal{F}_{\boldsymbol{p}}(\boldsymbol{y})$ is nonempty for all $\boldsymbol{p}$ and $\boldsymbol{y} \in \mathcal{C}_{\boldsymbol{p}}$. Problems (51) and (52) are reformulated by replacing $\mathcal{S}_{\boldsymbol{p}}(\boldsymbol{y})$ with $\mathcal{F}_{\boldsymbol{p}}(\boldsymbol{y})$ in problem (52), while $\boldsymbol{z} \in \mathcal{F}^C_{\boldsymbol{p}}(\boldsymbol{y})$ is promoted to the upper-level problem (51). As an upper-level constraint relating $\boldsymbol{y}$ and $\boldsymbol{z}$, it is absorbed into the *coupling constraints* (51d). This technique is applied in the experiments of Section 3.6.2 and detailed in Appendices 3.11.1 and 3.12.

3.6 Experiments

This section evaluates the effectiveness of the proposed methods in learning to solve several parametric bilevel optimization problems. To measure the optimality of learned solutions, the true optimal solutions must be computable. Given the challenges in finding these solutions with existing methods [99], Section 3.6.1 begins by learning solutions to small-scale synthetic problems. For these problems, open-source solvers can provide certified optimal solutions for comparison. Section 3.6.2 then extends the evaluation to more complex bilevel programs in engineering design. These problems are significantly more challenging due to their larger size and complex forms, including *nonconvex optimization* at the lower level. In these cases, learned solutions are compared against those from Particle Swarm Optimization (PSO), a metaheuristic framework commonly used in design optimization [132], [133].

Evaluation Criteria and Conventions. Algorithm 2 is evaluated on its ability to perform the training task specified in Equation 53. Constraints (53c) and (53b) are ensured by the algorithm's construction. Thus, the two main criteria for evaluating the learned solutions $(\hat{\boldsymbol{y}}, \hat{\boldsymbol{z}})$ are their *objective values* (53a) and potential *violations of the coupling constraint* (53d). When true optimal solutions $(\boldsymbol{y}^\star, \boldsymbol{z}^\star)$ are available, the *relative optimality gap* is reported, defined as $\left\lvert\mathcal{L}(\hat{\boldsymbol{y}},\hat{\boldsymbol{z}})-\mathcal{L}(\boldsymbol{y}^\star,\boldsymbol{z}^\star)/\mathcal{L}(\boldsymbol{y}^\star,\boldsymbol{z}^\star)\right\rvert$ and illustrated *in blue*. When true optima are not known, the nominal objective value of the learned solutions, *in green*, is reported alongside the solution produced by a baseline method for comparison. The coupling constraint violation (55) is also reported, and illustrated *in orange*. In general, all metrics are reported on average over the respective test set of problem instances. When a metric should ideally converge to zero, such as the optimality gap and coupling violation, its standard deviation is also reported. Additional implementation details for each experiment, including hyperparameter values, can be found in Appendices 3.10, 3.11.1, 3.12.

3.6.1 Learning Bilevel Quadratic Programming

Experimental evaluation begins with a relatively simpler class of bilevel problems: Bilevel Quadratic Programs (BQP). Both their upper and lower-level problems contain convex

quadratic objective functions and only linear constraints. The upper- and lower-level problems are, respectively:

$$\mathcal{B}(\boldsymbol{c},\boldsymbol{d}) = \underset{\boldsymbol{y}}{\text{argmin}} \quad \frac{1}{2}\boldsymbol{y}^T\boldsymbol{Q}\boldsymbol{y} + \boldsymbol{c}^T\boldsymbol{y} + \boldsymbol{d}^T\boldsymbol{z} + q \tag{61a}$$

$$\textit{s.t.} \quad \boldsymbol{A}\boldsymbol{y} \leq \boldsymbol{b} + \boldsymbol{E}\boldsymbol{z} \tag{61b}$$

$$\boldsymbol{z} \in \mathcal{O}(\boldsymbol{y}), \tag{61c}$$

$$\mathcal{O}(\boldsymbol{y}) = \underset{\boldsymbol{z}}{\text{argmin}} \quad \frac{1}{2}\boldsymbol{z}^T\boldsymbol{H}\boldsymbol{z} + \boldsymbol{e}^T\boldsymbol{z} + \boldsymbol{f}^T\boldsymbol{y} + g \tag{62a}$$

$$\textit{s.t.} \quad \boldsymbol{F}\boldsymbol{z} \leq \boldsymbol{h} + \boldsymbol{G}\boldsymbol{y}. \tag{62b}$$

A neural network is trained to learn solutions as a function of the upper-level linear objective coefficients. That is, in the notation of Section 3.3, $\boldsymbol{p} \coloneqq (\boldsymbol{c},\boldsymbol{d})$.

Experimental Details. For problems where certified optimal solutions can be computed within a reasonable time (i.e., sufficiently small BQP problems), these instances are solved using a mixed-integer programming reformulation. This provides an initial setting for objective comparison against our learned solutions. The open-source YALMIP package is used to solve instances by replacing the lower-level problem with KKT conditions and solving the resulting single-level MIP via a branch-and-bound method [134].

A BQP problem with m and n variables at the upper and lower levels, respectively, is defined as having size $m \times n$. Three sets of BQP problems with sizes 3×2, 6×4, and 9×6 are randomly generated. This is done by drawing elements of matrices $\boldsymbol{A},\boldsymbol{E},\boldsymbol{F},\boldsymbol{G}$ and vectors $\boldsymbol{b},\boldsymbol{e},\boldsymbol{f},\boldsymbol{h}$ from a uniform distribution $U(0,1)$. Positive-semidefinite matrices $\boldsymbol{Q}$ and $\boldsymbol{H}$ are constructed by self-multiplication of such a matrix. Individual problem instances are generated by drawing vectors of linear objective coefficients $(\boldsymbol{c},\boldsymbol{d})$ also from $U(0,1)$. Each set is divided into validation and test portions, each numbering 1000 instances. The prediction model $\hat{\mathcal{B}}_\theta$ is a 5-layer feedforward network followed by 20 steps of Algorithm 2. Algorithm 2 requires differentiable solution of problem (62) at each iteration; for this, the differentiable convex optimization library *cvxpylayers* [122] is employed.

Results. Figure 22 illustrates the two main evaluation metrics—relative optimality gap (blue) and coupling violation (orange)—throughout Algorithm 2 for each parametric BQP problem. In each case, the relative optimality gap is reduced by two orders of magnitude over 75 epochs to a value between 10^{-3} and 10^{-2}. The coupling violation (55) is rapidly diminished in the first epoch and generally remains below 10^{-2} within a full standard deviation throughout training. Test set metrics are also reported in Table 3. These results demonstrate Algorithm 2's ability to learn bilevel optimization with negligible error on small-scale BQP problems. Beyond the problem sizes considered here (starting with 12×9), the time taken by YALMIP to fully solve test instances becomes intractable, thus preventing benchmarking against certified optimal solutions on larger instances.

3.6.2 Learning Optimal Control Co-Design

This section introduces more challenging bilevel optimization problems, distinguished by their larger size and more complex forms. Optimal Control Co-Design is a bilevel problem setting where an engineering system is designed to optimize an economic objective at the upper level, subject to its behavior under a known optimal control policy at the lower level. These problems cannot be solved by conventional methods with certificates of optimality. As an

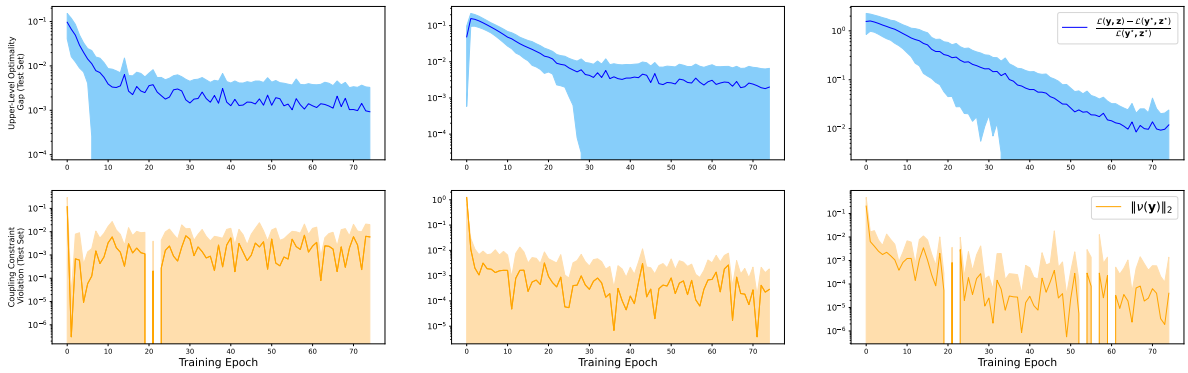


Figure 22: Optimality gap and coupling violation on different-sized BQP problems (left to right: 3×2, 6×4, 9×6). Shows mean and standard deviation over the test set at each training epoch.

alternative, our learned solutions are compared to the results of a Particle Swarm Optimization (PSO)-based metaheuristic method. The PSO framework is commonly applied to problems lacking efficient solution methods, making it a preferred tool in design optimization [132], [133]. Details of the PSO baseline methods, along with illustrations of their results, can be found in Appendix 3.8.

3.6.3 Nonconvex Bilevel Optimization: Control Co-design of a Nonlinear System

This section considers a nonlinear control problem involving two connected tanks controlled by a single pump and a two-way valve. This system serves as a simplified model for pumped-storage hydroelectricity, a form of energy storage used by electric power systems for load balancing. The system dynamics are described by nonlinear ODEs: $\dot{\boldsymbol{x}} = f(\boldsymbol{x}^{(i)}, \boldsymbol{u}^{(i)}, \boldsymbol{y})$:

$$\dot{x_1} = y_1(1 - u_2)u_1 - y_2\sqrt{x_1}, \tag{63a}$$

$$\dot{x_2} = y_1 u_2 u_1 + y_2\sqrt{x_1} - y_2\sqrt{x_2}, \tag{64a}$$

where x_1 and x_2 are the water levels in each tank. Control actions consist of u_1 (pump modulation) and u_2 (valve opening). The nonlinear optimal control problem (66) seeks the control policy that minimizes energy expended to reach a desired terminal state $\boldsymbol{p}$. The function `ODESolve` represents Euler discretization of (63, 64) over $N = 20$ frames to a final time T. This yields new variables $\boldsymbol{x} = [\boldsymbol{x}^{(1)}, \ldots, \boldsymbol{x}^{(N)}] \in \mathbb{R}^{N\times 2}$ and $\boldsymbol{u} = [\boldsymbol{u}^{(1)}, \ldots, \boldsymbol{u}^{(N)}] \in \mathbb{R}^{N\times 2}$, bound by a sequence of nonlinear (i.e., *nonconvex*) equality constraints (66c) for $1 \leq i < N$, with $dt = \frac{T}{N}$. The upper- and lower-level problems are, respectively:

$$\begin{aligned}
\mathcal{B}(\boldsymbol{p}) = \text{argmin}_{\boldsymbol{y}} \quad & \boldsymbol{v}^T\boldsymbol{y} && \text{(65a)}\\
\textit{s.t.} \quad & \boldsymbol{x}, \boldsymbol{u} = \mathcal{O}_{\boldsymbol{p}}(\boldsymbol{y}) && \text{(65b)}\\
& \boldsymbol{y}_{min} \leq \boldsymbol{y} \leq \boldsymbol{y}_{max} && \text{(65c)}\\
& \boldsymbol{x}^{(N)} = \boldsymbol{p}, && \text{(65d)}
\end{aligned}$$

with

$$\mathcal{O}_{\boldsymbol{p}}(\boldsymbol{y}) \coloneqq \underset{0\le \boldsymbol{x},\boldsymbol{u}\le 1}{\operatorname{argmin}} \quad \sum_{k=1}^{N} \|\boldsymbol{u}^{(k)}\|_2^2 \tag{66a}$$

$$\textit{s.t.} \quad \boldsymbol{x}^{(N)} = \boldsymbol{p} \tag{66b}$$

$$\boldsymbol{x}^{(i+1)} = \text{ODESolve}\big(f(\boldsymbol{x}^{(i)}, \boldsymbol{u}^{(i)}, \boldsymbol{y})\big). \tag{66c}$$

The upper-level problem (65) seeks to optimize the design of such a system in terms of its overall cost $\mathcal{L}(\boldsymbol{y}, \boldsymbol{z}) \coloneqq \boldsymbol{v}^T\boldsymbol{y}$, treating the inlet and outlet valve coefficients $\boldsymbol{y} = [y_1, y_2]$ as free design parameters. A feasible design requires satisfying upper and lower bounds on each element of $\boldsymbol{y}$ per (65c). Additionally, the parametric end-state constraint (66b) is duplicated at the upper level in (65d) to emphasize its coupling role. The full reformulation, as per Section 3.5.2, is detailed in Appendix 3.11.1. The initial condition is $\boldsymbol{x}^0 = \boldsymbol{0}$. Taken together, the coupled problems (65, 66) aim to find the parameters $\boldsymbol{y}$ that yield the minimal-cost system design controllable to state $\boldsymbol{p}$ by time T under its control policy (66). In this experiment, the model $\hat{\mathcal{B}}_\theta$ is trained to perform a fast and accurate approximation to this design problem for any scenario specified by a given $\boldsymbol{p} \in [0, 1]$. In the notation of Section 3.3, $\boldsymbol{z} \coloneqq (\boldsymbol{x}, \boldsymbol{u})$.

Experimental Details. This experiment considers $T = 10s$ and $N = 20$, with $\boldsymbol{c}_{min} = 0$ and $\boldsymbol{c}_{max} = \frac{1}{3}$. Problem instances correspond to reference states $\{\boldsymbol{p}\}$, randomly generated from $U(0, 1)$, but with $p_1 < p_2$ to ensure problem (65) feasibility. They are partitioned into training, validation, and test sets of sizes 10000, 1000, and 1000, respectively. Corresponding optimal system designs are learned by $\hat{\mathcal{B}}_\theta$, which consists of an 8-layer feedforward network $\mathcal{N}_\theta$ followed by 20 steps of Algorithm 2. Implementing Algorithm 2 is complex, as it requires a differentiable solution of the nonconvex lower-level programs (66). For this, the differentiable model predictive control solver from [135] is employed, which differentiates problem (66) implicitly via the KKT conditions of the final convex subproblem of a sequential quadratic programming.

Results. Figure 23 (top two plots) illustrates the value of the design objective $\mathcal{L}(\mathbf{y}, \mathbf{z})$, as well as the coupling violation (55), over the test set throughout training. Overall metrics are also found in Table 4 under the header "TT". Despite requiring an average of 1286.7s of solution time per instance, the PSO baseline produces design solutions with 15 percent higher cost than those learned by Algorithm 2, which infers solutions in 0.027s on average. Simultaneously, it achieves nearly identical satisfaction of the coupling constraint. This result is significant, demonstrating the ability to learn nonconvex bilevel optimization with high accuracy.

3.6.4 Control Co-Design of a Building HVAC System

Finally, this report considers the design and control of a building Heating, Ventilation, and Air Conditioning (HVAC) system. Its control policy minimizes energy expenditure while maintaining indoor temperature within prescribed bounds, as described by problem (68). The building consists of two zones, thermally connected to each other and the outside environment by a matrix of conductivity coefficients $\boldsymbol{A}$. State variables $\boldsymbol{x} = [\boldsymbol{x}^{(1)}, \ldots, \boldsymbol{x}^{(N)}] \in \mathbb{R}^{N\times 8}$ consist of the temperatures of each zone's floor, walls, indoor air, and exterior facade at each timestep k. Control actions $\boldsymbol{u} = [\boldsymbol{u}^{(1)}, \ldots, \boldsymbol{u}^{(N)}] \in \mathbb{R}^{N\times 2}$ induce heat flows into each zone, affecting the temperature states $\boldsymbol{x}$ via the actuator design variables $\boldsymbol{Y} \in \mathbb{R}^{8\times 2}$. States are also affected by random disturbances $\boldsymbol{d}$, including heat transfer from occupants and solar irradiation. Thermal constraints (68c) require that the indoor air temperature remains within prescribed time-varying bounds $(\underline{\boldsymbol{p}}, \overline{\boldsymbol{p}})$. This condition couples the upper- and lower-level problems, emphasized by its

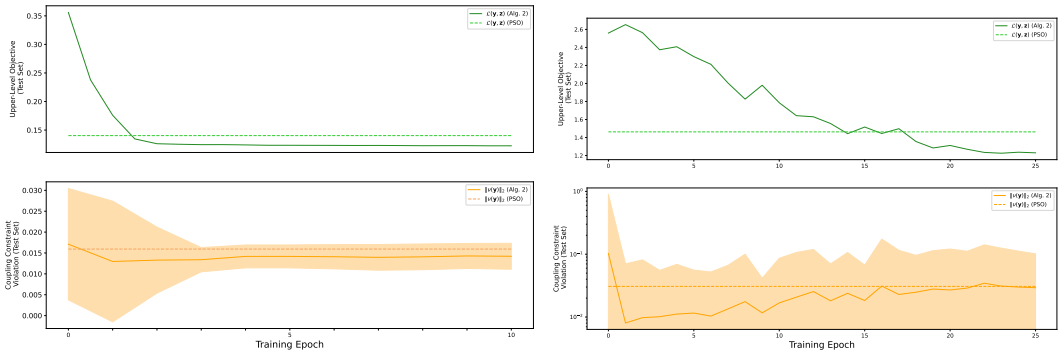


Figure 23: Test set metrics per epoch, on learning control co-design of the two-tank (top) and HVAC (bottom) systems.

duplication at (67d). The design task (67) seeks to learn $\boldsymbol{Y}$ that minimizes a linear cost function $Tr(\boldsymbol{V}^T\boldsymbol{Y})$ while allowing the system to be maintained within thermal bounds using optimal control. In the notation of Section 3.3, $\boldsymbol{p} := (\underline{\boldsymbol{p}}, \overline{\boldsymbol{p}})$, $\boldsymbol{y} := \boldsymbol{Y}$, and $\boldsymbol{z} := (\boldsymbol{x}, \boldsymbol{u}, \boldsymbol{w})$:

$$\begin{aligned} \mathcal{B}(\boldsymbol{p}) = \text{argmin} \quad & Tr(\boldsymbol{V}^T\boldsymbol{Y}) && \text{(67a)} \\ \textit{s.t.} \quad & \boldsymbol{x}, \boldsymbol{u}, \boldsymbol{w} = \mathcal{O}_{\boldsymbol{p}}(\boldsymbol{Y}) && \text{(67b)} \\ & \boldsymbol{Y} \geq \boldsymbol{0} && \text{(67c)} \\ & \underline{\boldsymbol{p}}^{(k)} \leq \boldsymbol{w}^{(k)} \leq \overline{\boldsymbol{p}}^{(k)}, && \text{(67d)} \end{aligned}$$

while

$$\begin{aligned} \mathcal{O}_{\boldsymbol{p}}(\boldsymbol{Y}) = \underset{\boldsymbol{x}, 0 \leq \boldsymbol{u} \leq 1, \boldsymbol{w}}{\text{argmin}} \quad & \sum_{k \in \{1 \dots N\}} \|\boldsymbol{u}^{(k)}\|_2^2 && \text{(68a)} \\ \textit{s.t.} \quad & \boldsymbol{w}^{(k)} = \boldsymbol{C}\boldsymbol{x}^{(k)} && \text{(68b)} \\ & \underline{\boldsymbol{p}}^{(k)} - \underline{\boldsymbol{s}}^{(k)} \leq \boldsymbol{w}^{(k)} \leq \overline{\boldsymbol{p}}^{(k)} + \underline{\boldsymbol{s}}^{(k)} && \text{(68c)} \\ & \boldsymbol{x}^{(k+1)} = \boldsymbol{A}\boldsymbol{x}^{(k)} + \boldsymbol{Y}\boldsymbol{u}^{(k)} + \boldsymbol{E}\boldsymbol{d}^{(k)}. && \text{(68d)} \end{aligned}$$

Experimental Details. This experiment assumes $N = 30$ simulation steps. Problem instances corresponding to the thermal bounds $\underline{\boldsymbol{p}}$ are generated from a β-random walk, with $\overline{\boldsymbol{p}} = \underline{\boldsymbol{p}} + 2.0$, to form training, validation, and test sets of sizes 10000, 1000, and 1000, respectively. A fixed disturbance pattern $\boldsymbol{d}$ is generated from the building control test suite in `NEUROMANCER` [136]. The model $\mathcal{B}_\theta$ is a 6-layer ReLU network followed by 10 steps of Algorithm (2). Differentiable solution of (68) is implemented using `cvxpylayers`.
Results. Figure 23 (bottom two plots) illustrates the value of the design objective $\mathcal{L}(\mathbf{y}, \mathbf{z})$, as well as the coupling violation (55), over the test set throughout training. Overall metrics are also found in Table 4 under "HVAC". While learned designs incur nearly identical coupling constraint violations on average, they are achieved at about 19 percent lower cost and with orders of magnitude lower solving time compared to the baseline.

3.7 Conclusion and Limitations

This report has demonstrated how modern toolsets for differentiable optimization can be used to train machine learning models as fast and accurate approximators of bilevel optimization

with coupling constraints. Experiments on problems in the control co-design domain show that the proposed framework can accurately approximate complex, even nonconvex programs. Two open challenges remain to extend this work. First, it assumes that the lower-level minimizer can be backpropagated, which is not the case in some interesting bilevel problems; future work may focus on extending the framework to set-valued lower solutions. Second, when lower-level objectives are nondifferentiable with respect to the upper variables, extensions based on surrogate gradients or implicit subdifferentials may be investigated. Addressing these points would further broaden the impact and applicability of this work.

3.8 Experimental Details: Particle Swarm Optimization

Particle Swarm Optimization (PSO) is a metaheuristic optimization method that operates with a population (swarm) of candidate solutions (particles). Particles update their solutions using simple rules based on their own best-known position in the search space, as well as the entire swarm's best-known position. PSO is commonly applied to optimization problems with complex objective functions and simple constraints. While simple bounds can be handled naturally in PSO, more complex constraints are often addressed using penalty functions. A survey of constraint-handling techniques in PSO is found in [137].

For both control co-design experiments, PSO baseline methods are implemented using the open-source library *pyswarms* [138]. A penalty-function approach is adopted to handle coupling constraints in our PSO baseline methods. Since the remaining upper-level constraints are simple variable bounds, these are handled natively in the pyswarms PSO algorithm. As in the report's main proposal, the lower-level problem must be feasible relative to the upper-level solutions found at each iteration of PSO. Constraints preventing this condition are treated as coupling constraints and enforced with penalty functions in the lower-level problem. For the control co-design problems, the lower-level problem implementations are identical to (70) and (72). Overall, PSO optimizes a relaxed upper-level objective function:

$$\mathcal{L}_{\boldsymbol{p}}(\boldsymbol{y}) + \kappa\nu(\boldsymbol{y}), \tag{69}$$

subject to $\boldsymbol{y} \in \mathcal{C}_{\boldsymbol{p}}$, which are simple bounds on $\boldsymbol{y}$ in both experimental cases. Evaluation of $\nu(\boldsymbol{y})$ requires optimization at the lower level at each PSO iteration.

In each experiment, the pyswarms solver is configured with default cognitive, social, and inertia parameters ($c_1 = 0.5, c_2 = 0.5, w = 0.9$) and run with 128 particles for 200 iterations. The penalty coefficient κ is chosen such that average coupling constraint violations over the test set are on the order of $1e-2$. This corresponds to $\kappa = 100.0$ in the two-tank experiment and $\kappa = 5.0$ in the HVAC experiment.

3.9 Additional Results: Particle Swarm Optimization

This section illustrates the evolution of the PSO objective throughout its solution of the test set instances. The best objective values, among all particles, are plotted per iteration of PSO. The full PSO objective is illustrated, and is also shown in terms of upper-level design objective and coupling constraint violation penalty. Each metric is reported on average over the test set. Figure 24 corresponds to the two-tank problem, and Figure 25 corresponds to the HVAC problem.

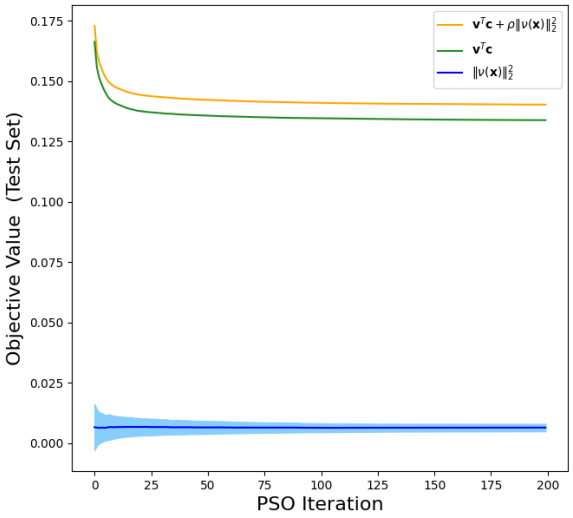


Figure 24: Best Objective Value per PSO Iteration on Two-Tank System Co-Design.

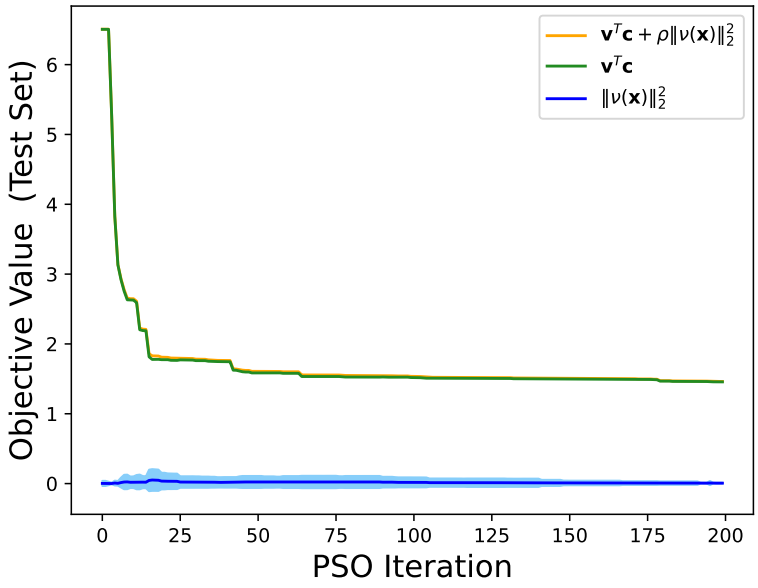


Figure 25: Best Objective Value per PSO Iteration on HVAC System Co-Design.

3.10 Experimental Details: Learning Bilevel Quadratic Programming

This section reports additional details on the experiments presented in Section 3.6.1.

3.11 Hyperparameters and Training

Results in Section 3.6 are derived from the model that achieves the lowest loss among independent training runs, using all combinations of the following hyperparameters:

- Learning rates from among $[10^{-1}, 10^{-2}, 10^{-3}, 10^{-4}, 10^{-5}]$
- Correction stepsizes γ from $[10^{-2}, 10^{-3}, 10^{-4}]$
- $\mathcal{L}^{\texttt{SOFT}}$ penalty weights λ from $[10^2, 10^3]$

The best values are $10^{-3}, 10^{-4}, 10^2$ respectively. All models are trained using the Adam optimizer [139] in PyTorch. In each training run, 10 correction steps are applied in training and 20 are applied at test time.

3.11.1 Experimental Details: Learning Control Co-design of a Two-Tank System

This section reports additional details on the nonlinear system design experiments presented in Section 3.6.2.

3.11.2 Problem Reformulation

The two-tank system design and control problem (65, 66) is bound by the coupling constraint $\boldsymbol{x}^{(N)} = \boldsymbol{p}$, which is redundantly placed at both levels to emphasize its coupling effect. This coupling constraint expresses that a valid system design must be controllable to the end-state $\boldsymbol{p}$. It is recognized that this condition may not be satisfiable for any design variable $\boldsymbol{y}$; for instance, some $\boldsymbol{y}$ may not allow sufficient throughput to fill the tanks from $\mathbf{0}$ to $\boldsymbol{p}$ by time step N.

In practice, the problem is therefore reformulated as follows, as prefaced in Section 3.5.2:

$$\mathcal{B}(\boldsymbol{p}) = \operatorname{argmin}_{\boldsymbol{y}} \quad \boldsymbol{v}^T \boldsymbol{y} \tag{70a}$$

$$\text{s.t.} \quad \boldsymbol{x}, \boldsymbol{u} = \mathcal{O}_{\boldsymbol{p}}(\boldsymbol{y}) \tag{70b}$$

$$\boldsymbol{y}_{min} \leq \boldsymbol{y} \leq \boldsymbol{y}_{max} \tag{70c}$$

$$\boldsymbol{x}^{(N)} = \boldsymbol{p} \tag{70d}$$

$$\text{where} \quad \mathcal{O}_{\boldsymbol{p}}(\boldsymbol{y}) := \underset{0 \leq \boldsymbol{x}, \boldsymbol{u} \leq 1}{\operatorname{argmin}} \quad \sum_{k=1}^{N} \|\boldsymbol{u}^{(k)}\|_2^2 + \rho \|\boldsymbol{x}^{(N)} - \boldsymbol{p}\|^2 \tag{70e}$$

$$\text{s.t.} \quad \boldsymbol{x}^{(i+1)} = \text{ODESolve}\big(f(\boldsymbol{x}^{(i)}, \boldsymbol{u}^{(i)}, \boldsymbol{y})\big) \tag{70f}$$

In this reformulation, the constraint $\boldsymbol{x}^{(N)} = \boldsymbol{p}$ remains at the upper level as a *coupling* constraint, as it binds lower-level variables within the upper-level problem. Thus, it is treated by the coupling constraint correction in Algorithm 2. The constraint is absent from the lower level in this formulation and instead replaced with a penalty on the lower-level objective. $\rho = 100$ is used as the penalty weight in all experimental runs. While equivalent to the original

bilevel problem (65, 66), the above formulation accommodates lower-level feasibility for any candidate design solution at the upper level.

3.11.3 Hyperparameters and Training

Results are derived from the model that achieves the lowest upper-level objective value on average, from among those whose average coupling constraint violation is less than or equal to that of the PSO baseline method after 10 epochs. The results are chosen from among independent training runs using all combinations of the following hyperparameters:

- Learning rates from among $[10^{-1}, 10^{-2}, 10^{-3}, 10^{-4}, 10^{-5}]$
- Correction stepsizes γ from $[10^{-2}, 10^{-3}, 10^{-4}]$
- $\mathcal{L}^{\texttt{SOFT}}$ penalty weights λ from $[10, 10^2, 10^3]$

The best values are $10^{-3}, 10^{-2}, 10$, respectively. All models are trained using the Adam optimizer in PyTorch. In each training run, 5 correction steps are applied in training and 10 are applied at test time. The upper-level objective function has linear coefficients v consisting of all ones in each problem instance, implying that the inlet and outlet valve coefficients should be minimized with equal priority.

3.12 Experimental Details: Learning Control Co-design of a Building HVAC System

This section reports additional details on the HVAC system design experiments presented in Section 3.6.2.

3.13 Problem Reformulation

The building HVAC design and control problem (67,68) is coupled by the thermal constraints $\underline{\boldsymbol{p}}^{(k)} \leq \boldsymbol{w}^{(k)} \leq \overline{\boldsymbol{p}}^{(k)}$, which appear at both levels to emphasize their coupling role. Before Algorithm 2 can be applied, it is recognized that those constraints may not be satisfiable when design variables $\boldsymbol{Y}$ prevent heat flows from converting to temperature changes rapidly enough to stay within those changing bounds.

To arrive at an equivalent problem that ensures feasibility at the lower level for any $\boldsymbol{Y}$, slack variables $\underline{\boldsymbol{s}}^{(k)}$ and $\overline{\boldsymbol{s}}^{(k)}$ are introduced to both sides of (68c), yielding:

$$\underline{\boldsymbol{p}}^{(k)} - \underline{\boldsymbol{s}}^{(k)} \leq \boldsymbol{w}^{(k)} \leq \overline{\boldsymbol{p}}^{(k)} + \underline{\boldsymbol{s}}^{(k)},$$

along with a no-slack condition that maintains equivalence to the original problem:

$$\underline{\boldsymbol{s}}^{(k)} = \overline{\boldsymbol{s}}^{(k)} = \boldsymbol{0} \quad \forall k.$$

This condition is enforced at the upper level and replaced in the lower level by a penalty term:

$$\begin{aligned}
\mathcal{B}(\boldsymbol{p}) = \text{argmin} \quad & \mathit{Tr}(\boldsymbol{V}^T \boldsymbol{Y}) && \text{(71a)}\\
\textit{s.t.} \quad & \boldsymbol{x}, \boldsymbol{u}, \boldsymbol{w} = \mathcal{O}_{\boldsymbol{p}}(\boldsymbol{Y}) && \text{(71b)}\\
& \boldsymbol{Y} \geq \boldsymbol{0} && \text{(71c)}\\
& \underline{\boldsymbol{s}}^{(k)} = \overline{\boldsymbol{s}}^{(k)} = \boldsymbol{0} \quad \forall k && \text{(71d)}
\end{aligned}$$

$$\mathcal{O}_{\boldsymbol{p}}(\boldsymbol{Y}) = \underset{\boldsymbol{x},0\leq\boldsymbol{u}\leq 1,\boldsymbol{w},\boldsymbol{s}}{\text{argmin}} \sum_{k\in\{1\dots N\}} \|\boldsymbol{u}^{(k)}\|_2^2 + \rho\left(\sum_{k\in\{1\dots N\}} \|\underline{\boldsymbol{s}}^{(k)}\|_2^2 + \sum_{k\in\{1\dots N\}} \|\overline{\boldsymbol{s}}^{(k)}\|_2^2\right) \tag{72a}$$

$$\textit{s.t.} \quad \boldsymbol{w}^{(k)} = \boldsymbol{C}\boldsymbol{x}^{(k)} \tag{72b}$$

$$\underline{\boldsymbol{p}}^{(k)} - \underline{\boldsymbol{s}}^{(k)} \leq \boldsymbol{w}^{(k)} \leq \overline{\boldsymbol{p}}^{(k)} + \underline{\boldsymbol{s}}^{(k)} \tag{72c}$$

$$\boldsymbol{x}^{(k+1)} = \boldsymbol{A}\boldsymbol{x}^{(k)} + \boldsymbol{Y}\boldsymbol{u}^{(k)} + \boldsymbol{E}\boldsymbol{d}^{(k)} \tag{72d}$$

In this implementation of Algorithm 2, coupling constraint corrections are applied to (71d). This is the operative coupling constraint in this reformulation, binding the upper-level problem to lower-level variables $\underline{\boldsymbol{s}}^{(k)}$ and $\overline{\boldsymbol{s}}^{(k)}$.

3.13.1 Hyperparameters and Training

Results are presented from the model that achieves the lowest upper-level objective value on average, among those whose average coupling constraint violation is less than or equal to that of the PSO baseline method after 25 epochs. The results are chosen from among independent training runs using all combinations of the following hyperparameters:

- Learning rates from among $[10^{-1}, 10^{-2}, 10^{-3}, 10^{-4}, 10^{-5}]$
- Correction stepsizes γ from $[10^{-4}, 10^{-5}, 10^{-6}]$
- $\mathcal{L}^{\texttt{SOFT}}$ penalty weights λ from $[10^2, 10^3]$

The best values are $10^{-3}, 10^{-4}, 10^2$, respectively. All models are trained using the Adam optimizer in PyTorch. In each training run, 5 correction steps are applied in training and 10 are applied at test time. The upper-level objective function has linear coefficients $\boldsymbol{V}$ consisting of all ones in each problem instance, meaning that all elements of the actuator design variable have equal cost.

4.0 Multi-terminal Topology Co-design for Variable Generation Grid Integration

4.1 Motivation

Offshore wind farms (OWFs) are increasingly important for sustainable energy development. The 2024 Offshore Wind Market Report [36] highlights a 53% growth in the U.S. offshore wind energy pipeline, with significant additions on both coasts. Efficient development of OWFs and their transmission capabilities to the main grid is crucial. The current grid, primarily designed for synchronous generators, faces challenges with integrating large renewable resources, affecting design and stability [140]. Multi-terminal direct current (MTDC) grids, utilizing modular multi-level converters (MMC) and high voltage DC (HVDC) lines, offer enhanced reliability, reduced losses, and improved controllability compared to high voltage alternating current (HVAC) lines. Onshore converters re-convert DC to AC at interconnection points [2].

Planning offshore wind's grid integration requires understanding multi-regional transmission networks. Brinkman et al. [38] analyzed various offshore topologies (radial and networked) for 2030/2050, examining operational, cost, and reliability implications. Douville et al. [76] studied western interconnection with offshore wind, investigating radial and backbone topologies in HVAC/HVDC scenarios for 2030+. Both studies underscored the importance of offshore transmission topology for the value of offshore wind generation and identified Points of Interconnection (POIs) through rigorous modeling. The Atlantic study used Eastern-Interconnect utility data with optimization based on wind resources, POI capacity, and cable distance. Douville et al. [76] identified POIs by assessing onshore transmission capacity and conceptualizing offshore systems for efficient power flows.

OWF power production is weather-dependent, introducing uncertainty in power injection. The point of common coupling at the POI is used to inject power into the onshore grid. Variability in power injection can jeopardize grid stability. Introducing energy storage systems (ESSs) with advanced operational and control strategies is a popular approach for enhancing power system stability [4] and providing economic opportunities. A detailed review of the role of ESS for wind energy resources is provided in [5], with various studies examining the design and control of ESS for wind energy.

Halwany et al. [6] and Moghaddam et al. [8] emphasized the importance of ESS with OWFs. Research by Paul et al. [7] and Sharma et al. [141] developed multi-objective optimization formulations for ESS sizing, while Lin et al. [63] and Zoellick et al. [142] found that ESS enhances OWF revenue through arbitrage and ancillary services. Their study showed that the OWF at Humboldt Bay Substation, despite its small battery size, significantly increased revenue from these markets.

Related Work: Understanding the interaction between energy markets and energy resources is crucial. A comprehensive review on demand-side and market design for renewable systems integration is provided in [143]. Studies by Brunetto and Tina [144] and Di Corato et al. [53] explored the electricity market interaction for wind power with energy storage. Dowling et al. [145] introduced a multi-time scale optimization formulation. Chen et al. [146] and Sorourifar et al. [75] optimized wind farms and battery storage systems using multi-timescale formulations. Gao et al. [73] demonstrated integrated energy systems' participation in electricity markets. Parede et al. [147] examined distributed energy resource interactions in the electricity market. Chen et al. [148] revealed a multi-interval co-optimization approach for battery storage bidding. Bansal et al. [149] introduced a two-stage settlement market mechanism. Ma et al. [150] and Li et al. [151] applied multi-time scale formulations for energy systems. Given the variability in renewable generation and pricing, accounting for

uncertainty is crucial for long-term planning. The importance of stochastic programming for energy systems is discussed in [152], [153]. Several studies [154]–[156] have formulated wind-storage system market participation as a multi-stage stochastic problem.

For long-term OWF planning, developing a methodology to optimize POI choices and ESS siting, sizing, and operation is crucial. Existing literature lacks an integrated approach to deal with these challenges. This report presents a co-design methodology, leveraging multidisciplinary optimization techniques [16], [157], to address these issues comprehensively, demonstrated with a case study. To balance MTDC system and ESS costs, OWF revenue from various electricity markets is considered. We formulate a multiple-timescale optimization model and adopt a multi-stage stochastic optimization approach to handle uncertainties in OWF generation and market prices. Our case study uses a California-based OWF and potential interconnect POIs identified by Douville et al. [76], incorporating historical price data to optimize interconnections between OWF buses.

The remainder of this report is organized as follows: Section **??** discusses the proposed methodology for formulating a multi-stage multi-timescale stochastic optimization model. Section 4.4 describes the OWF use case and data sources. Section **??** discusses the results of the approach. Conclusions and future work are presented in Section 4.6.

4.2 Methodology

This section describes the details of the proposed co-design optimization approach for topology optimization with ESS for OWF integration.

4.2.1 Problem Description

We consider an offshore wind farm (OWF) connected to onshore points of interconnection (POIs) via an MTDC system. Detailed cable routing, including seafloor conditions and conflicts with ocean co-use, is not considered in this analysis. Similarly, the siting of balance-of-plant equipment, such as export cables, substations, and converter stations, is beyond the scope of this analysis. The cost of the MTDC system includes a fixed cost, a rated power-dependent material cost for each installed transmission cable, and fixed costs for onshore and offshore converter stations at their respective locations.

We model the ESS with a fixed duration and optimize its rated power (size) to match the data set used. It can be easily modified to include storage capacity as a decision variable. The cost is calculated solely based on size, without a fixed cost component. We assume that when the ESS reserve is called upon, it only needs to provide service for a short period, minimally affecting its state of charge (SoC). Additionally, the ESS can identify opportunities to maximize profit through market arbitrage.

The OWF participates in day-ahead (DA) and real-time (RT) markets. It trades most of its generated power in the DA market, using the RT market to mitigate mismatches between DA commitment and actual generation. ESSs are installed on the onshore side to facilitate system operation and participate in the ancillary service market by providing up and down reserves.

To optimize MTDC topology and ESS sizes for maximum net profit while considering wind speed and market price uncertainties, a multi-stage stochastic programming model is formulated. In the design stage, MTDC topology, cable size, and ESS locations and sizes are determined based on power generation and market price forecasts. In the day-ahead (DA) stage, given DA market prices and forecasts for power generation, real-time (RT) market, and reserve prices, decisions for DA market trading are made. Finally, in the RT stage, actual

OWF generation, RT market and reserve prices, and system participation and operation in RT and ancillary service markets are determined.

4.3 Multi-stage Multi-timescale Stochastic Optimization Formulation

Problems involving multiple stages and markets may operate on multiple time scales. We assume the DA market and OWF power generation operate on an hourly basis, while the RT and ancillary markets are cleared every 15 minutes. We use $\mathcal{T}_1 = \{1, 2, \ldots, 24\}$ and $\mathcal{T}_2 = \{1, 2, \ldots, 96\}$ to represent the index sets of each hour and each 15-minute interval in a day, respectively. Additionally, the set $\mathcal{T}_{21}(t) = \{4(t-1) + k \mid k = 1, 2, 3, 4\}$ contains all four 15-minute period indices in hour $t \in \mathcal{T}_1$.

Let ω^{DA} and ω^{RT} represent random events at DA and RT stages, respectively. These events, once observed, determine the value of uncertain parameters in each stage. $\mathbb{E}_{\omega^{\cdot}}[X]$ denotes the expectation of X over all scenarios of $\omega^{\cdot}$. The problem is formulated as follows:

$$\min\ R^{\mathsf{T}}\left(\sum_{i\in\mathcal{P}}\left(\lambda^{\mathsf{E}} sz_i^{\mathsf{E}} + \lambda^{\mathsf{On}} y_i\right) + \sum_{l\in\mathcal{L}}(\lambda_l^{\mathsf{C}} sz_l^{\mathsf{C}} + \lambda_l^{\mathsf{F}} x_l) + \sum_{i\in\mathcal{W}}\lambda^{\mathsf{Off}}\right) - R^{\mathsf{Y}}\mathbb{E}_{\omega^{\mathsf{DA}}}\left[\mathcal{F}\left(\mathbf{sz}^{\mathsf{E}}, \mathbf{sz}^{\mathsf{C}}, \mathbf{x}^{\mathsf{C}}; \omega^{\mathsf{DA}}\right)\right] \tag{73}$$

$$\text{s.t. } \underline{P}^{\mathsf{E}} y_i \le sz_i^{\mathsf{E}} \le \overline{P}^{\mathsf{E}} y_i\ \forall i \in \mathcal{P} \tag{74}$$

$$\underline{P}^{\mathsf{C}} x_l \le sz_l^{\mathsf{C}} \le \overline{P}^{\mathsf{C}} x_l\ \forall l \in \mathcal{L} \tag{75}$$

$$\sum_{l\in\mathcal{L}_i^{\mathsf{P}}} x_l \le M y_i\ \forall i \in \mathcal{P} \tag{76}$$

$$(92-93)$$

In the objective function (73), the first term includes the installation and material costs of ESSs, HVDC transmission cables, and converters. Factor R^{T} is a tax credit coefficient for new, qualified clean energy properties. The second term represents the expected revenue from the OWF. Annuity factor R^{Y} converts expected daily revenue to total present value over the lifetime of the ESSs. If their lifetime is Y years, then $R^{\mathsf{Y}} = 365\frac{(1+r)^Y - 1}{r(1+r)^Y}$ with inflation rate r.

$\mathcal{P}$, $\mathcal{W}$, and $\mathcal{L}$ represent the index sets of onshore POIs, offshore wind farm buses, and candidate transmission cables of the MTDC, respectively. For each POI $i \in \mathcal{P}$, $\mathcal{L}_i^{\mathsf{P}}$ denotes the set of its adjacent cables. Binary variable x_l is 1 if candidate cable l is installed, and 0 otherwise. Binary variable y_i is 1 if POI i is connected to the OWF, and 0 otherwise. Continuous variables sz_i^{E} and sz_l^{C} represent the rated power of the ESS at POI i and loading capacity of cable l, respectively. Parameters λ^{On}, λ^{Off}, and λ_l^{F} represent the installation cost of onshore converter, offshore converter, and cable l, respectively; λ^{E} and λ_l^{C} denote the unit material cost of the ESS and cable l, respectively; M is a large number, which can be the total number of buses in this problem.

Constraints (74-75) restrict the rated power of ESSs and transmission cable capacities, with $\underline{P}^{\mathsf{E}}$, $\underline{P}^{\mathsf{C}}$ and $\overline{P}^{\mathsf{E}}$, $\overline{P}^{\mathsf{C}}$ being their lower and upper limits, respectively. If any adjacent cable of POI i is installed, it is connected to the MTDC system, as stated in (76). Constraints (92-93) will be discussed later.

For given rated power of ESSs, MTDC topology, and a realization of DA scenario,

$\mathcal{F}\left(\mathbf{sz}^{\mathsf{E}}, \mathbf{sz}^{\mathsf{C}}, \mathbf{x}^{\mathsf{C}}; \omega^{\mathsf{DA}}\right)$ calculates the daily revenue of the OWF as follows:

$$\mathcal{F}(\mathbf{sz}^{\mathsf{E}}, \mathbf{sz}^{\mathsf{C}}, \mathbf{x}^{\mathsf{C}}; \omega^{\mathsf{DA}}) = \max \sum_{i\in\mathcal{P}} \sum_{t\in\mathcal{T}_1} \lambda_t^{\mathsf{DA}}\left(\omega^{\mathsf{DA}}\right) p_{i,t}^{\mathsf{DA}} + \mathbb{E}_{\omega^{\mathsf{RT}}|\omega^{\mathsf{DA}}}\left[\mathcal{G}\left(\mathbf{sz}^{\mathsf{E}}, \mathbf{sz}^{\mathsf{C}}, \mathbf{x}^{\mathsf{C}}, \mathbf{p}^{\mathsf{DA}}; \omega^{\mathsf{DA}}, \omega^{\mathsf{RT}}\right)\right] \tag{77}$$

The first and second terms are DA and expected RT stage revenues, respectively. Variable $p_{i,t}^{\mathsf{DA}}$ represents the power traded in the DA market at POI i at time t. Parameter $\lambda_t^{\mathsf{DA}}(\omega^{\mathsf{DA}})$ denotes the DA market price at time t in scenario ω^{DA}.

For given rated power of ESSs, MTDC topology, power traded in DA market, and realizations of DA and RT scenarios, $\mathcal{G}\left(\mathbf{sz}^{\mathsf{E}}, \mathbf{sz}^{\mathsf{C}}, \mathbf{x}^{\mathsf{C}}, \mathbf{p}^{\mathsf{DA}}; \omega^{\mathsf{DA}}, \omega^{\mathsf{RT}}\right)$ calculates the RT stage revenue of the OWF. Its objective function is:

$$\begin{aligned} \max \sum_{i\in\mathcal{P}} \sum_{t\in\mathcal{T}_2} \Big[& \lambda_t^{\mathsf{RT}}\left(\omega^{\mathsf{DA}}, \omega^{\mathsf{RT}}\right) p_{i,t}^{\mathsf{RT}} \\ & + \lambda_t^{\mathsf{ReU}}\left(\omega^{\mathsf{DA}}, \omega^{\mathsf{RT}}\right) p_{i,t}^{\mathsf{ReU}} + \lambda_t^{\mathsf{ReD}}\left(\omega^{\mathsf{DA}}, \omega^{\mathsf{RT}}\right) p_{i,t}^{\mathsf{ReD}} \Big] \\ \text{s.t. } & (79-91) \end{aligned} \tag{78}$$

Variables $p_{i,t}^{\mathsf{RT}}$, $p_{i,t}^{\mathsf{ReU}}$, and $p_{i,t}^{\mathsf{ReD}}$ represent the power traded in the RT market, up reserve, and down reserve at POI i at time t, respectively. Parameters $\lambda_t^{\mathsf{RT}}(\omega^{\mathsf{DA}}, \omega^{\mathsf{RT}})$, $\lambda_t^{\mathsf{ReU}}(\omega^{\mathsf{DA}}, \omega^{\mathsf{RT}})$, and $\lambda_t^{\mathsf{ReD}}(\omega^{\mathsf{DA}}, \omega^{\mathsf{RT}})$ denote the RT market, up and down reserve prices at time t in scenario $(\omega^{\mathsf{DA}}, \omega^{\mathsf{RT}})$, respectively.

This problem has a series of constraints. The power at wind farm buses and POIs should be balanced for all $t_1 \in \mathcal{T}_1,\ t_2 \in \mathcal{T}_{21}(t_1)$.

$$p_{i,t_1}^{\mathsf{W}} = \sum_{l\in\mathcal{L}_i^{\mathsf{P}}} V_{i,t_2}(V_{i,t_2} - V_{t_l,t_2}) g_{i,t_l} x_l \quad \forall i \in \mathcal{W} \tag{79}$$

$$\sum_{l\in\mathcal{L}_i^{\mathsf{P}}} V_{i,t_2}(V_{i,t_2} - V_{t_l,t_2}) g_{i,t_l} x_l = (p_{i,t_2}^{\mathsf{dis}} - p_{i,t_2}^{\mathsf{ch}}) - (p_{i,t_2}^{\mathsf{RT}} + p_{i,t_1}^{\mathsf{DA}}) \quad \forall i \in \mathcal{P} \tag{80}$$

Here, o_l, which will appear later, and t_l denote the origin and destination buses of branch l, respectively. Parameter $g_{i,j}$ is the conductance of the branch from bus i to bus j. Variables $p_{i,t}^{\mathsf{W}}$, $p_{i,t}^{\mathsf{dis}}$, $p_{i,t}^{\mathsf{ch}}$, and $V_{i,t}$ represent the actual power generation, discharged power of ESS, charged power of ESS, and voltage at bus i at time t, respectively.

The power through MTDC cables and the voltage at buses should be bounded:

$$1.1|V_{o_l,t}(V_{o_l,t} - V_{t_l,t}) g_{o_l,t_l} x_l| \le sz_l^{\mathsf{C}} \ \forall l \in \mathcal{L},\ t \in \mathcal{T}_2 \tag{81}$$

$$\underline{V}_i \le V_{i,t} \le \overline{V}_i \ \forall i \in \mathcal{P} \cup \mathcal{W},\ t \in \mathcal{T}_2 \tag{82}$$

In constraint (81), a 10% redundancy is imposed for security. In constraint (82), $\underline{V}_i$ and $\overline{V}_i$ denote the lower and upper limits of the voltage at bus i.

Since the RT market is primarily used to mitigate the difference between DA market participation and actual OWF generation, the following limits are imposed:

$$\sum_{i\in\mathcal{P}} p_{i,t_2}^{\mathsf{RT}} \le \max\left\{ \sum_{i\in\mathcal{W}} P_{i,t_1}^{\mathsf{OWF}}(\omega^{\mathsf{DA}}, \omega^{\mathsf{RT}}) - \sum_{i\in\mathcal{W}} P_{i,t_1}^{\mathsf{OWF}}(\omega^{\mathsf{DA}}), 0 \right\} \ \forall t_1 \in \mathcal{T}_1,\ t_2 \in \mathcal{T}_{21}(t_1) \tag{83}$$

$$\sum_{i\in\mathcal{P}} p^{\mathsf{RT}}_{i,t_2} \geq \min\left\{\sum_{i\in\mathcal{W}} P^{\mathsf{OWF}}_{i,t_1}(\omega^{\mathsf{DA}},\omega^{\mathsf{RT}}) - \sum_{i\in\mathcal{W}} P^{\mathsf{OWF}}_{i,t_1}(\omega^{\mathsf{DA}}), 0\right\} \ \forall t_1\in\mathcal{T}_1,\ t_2\in\mathcal{T}_{21}(t_1)$$

Here, $P^{\mathsf{OWF}}_{i,t}(\omega^{\mathsf{DA}},\omega^{\mathsf{RT}})$ represents the maximum wind power extracted at wind farm bus i by maximum power point tracking in scenario $(\omega^{\mathsf{DA}},\omega^{\mathsf{RT}})$ at time t. In addition, $P^{\mathsf{OWF}}_{i,t}(\omega^{\mathsf{DA}}) = \mathbb{E}_{\omega^{\mathsf{RT}}}[P^{\mathsf{OWF}}_{i,t}(\omega^{\mathsf{DA}},\omega^{\mathsf{RT}})]$. These two constraints ensure that the power traded in the RT market remains within the difference between the actual and expected wind power in each DA scenario.

The maximum wind power also bounds the actual power generation from above:

$$0 \leq p^{\mathsf{W}}_{i,t} \leq P^{\mathsf{OWF}}_{i,t}(\omega^{\mathsf{DA}},\omega^{\mathsf{RT}})\ \forall i\in\mathcal{W},\ t\in\mathcal{T}_1 \tag{84}$$

When providing up and down reserve for fast frequency support, ESS has the following operational and dynamic constraints for all $i\in\mathcal{P},\ t\in\mathcal{T}_2$:

$$SoC_{i,t} - SoC_{i,t-1} = (\eta^{\mathsf{ch}} p^{\mathsf{ch}}_{i,t} - \eta^{\mathsf{dis}} p^{\mathsf{dis}}_{i,t}) T^{\mathsf{L}}_2 \tag{85}$$

$$SoC_{i,1} - 0.5 sz^{\mathsf{E}}_i \mathsf{DH} = (\eta^{\mathsf{ch}} p^{\mathsf{ch}}_{i,1} - \eta^{\mathsf{dis}} p^{\mathsf{dis}}_{i,1}) T^{\mathsf{L}}_2 \tag{86}$$

$$\underline{r}^{\mathsf{ReU}} sz^{\mathsf{E}}_i \leq p^{\mathsf{ReU}}_{i,t} \leq \overline{r}^{\mathsf{ReU}} sz^{\mathsf{E}}_i \tag{87}$$

$$\underline{r}^{\mathsf{ReD}} sz^{\mathsf{E}}_i \leq p^{\mathsf{ReD}}_{i,t} \leq \overline{r}^{\mathsf{ReD}} sz^{\mathsf{E}}_i \tag{88}$$

$$p^{\mathsf{ch}}_{i,t} + p^{\mathsf{ReD}}_{i,t} \leq sz^{\mathsf{E}}_i \tag{89}$$

$$p^{\mathsf{dis}}_{i,t} + p^{\mathsf{ReU}}_{i,t} \leq sz^{\mathsf{E}}_i \tag{90}$$

$$0 \leq SoC_{i,t} \leq sz^{\mathsf{E}}_i \mathsf{DH} \tag{91}$$

Parameters $\underline{r}^{\mathsf{ReU}}$, $\underline{r}^{\mathsf{ReD}}$ and $\overline{r}^{\mathsf{ReU}}$, $\overline{r}^{\mathsf{ReD}}$ represent the lower and upper limits of ESS up and down reserves as a percentage of its rated power, respectively; η^{ch} and η^{dis} are ESS charging and discharging efficiency; T^{L}_2 is the time length of the time periods in set $\mathcal{T}_2$, which is 1/4 hour here; DH denotes ESS duration hours. Variable $SoC_{i,t}$ is the state of charge of ESS at POI i. Constraints (85-86) are ESS dynamic equations; and constraints (87-91) impose reserve, ESS charging, discharging, and SoC limits, respectively.

Finally, constraint (92) requires the expected SoC of ESS to be half of its energy capacity at the end of the day, and constraint (93) ensures its expected total discharge power does not exceed daily cycle limit $\mathsf{cyl}^{\mathsf{lim}}$.

$$\mathbb{E}_{\omega^{\mathsf{RT}}}[SoC_{i,|\mathcal{T}_2|}] = 0.5 sz^{\mathsf{E}}_i \mathsf{DH}\ \forall i\in\mathcal{P} \tag{92}$$

$$\mathbb{E}_{\omega^{\mathsf{RT}}}\left[\sum_{t\in\mathcal{T}_2} p^{\mathsf{dis}}_{i,t}\eta^{\mathsf{dis}} T^{\mathsf{L}}_2\right] \leq sz^{\mathsf{E}}_i \mathsf{cyl}^{\mathsf{lim}}\ \forall i\in\mathcal{P} \tag{93}$$

For each scenario, a one-day time horizon is considered to keep the problem tractable while estimating system revenue with acceptable accuracy. If higher accuracy is desired, users of this approach can generate a larger scenario tree and extend the time horizon.

4.4 Case Study

4.4.1 Case Setting

In the case study, the optimized wind farm siting results presented in [76] are used. These sites are primarily located in areas off the West Coast between Coos Bay, Oregon, and

Eureka, California. There are five offshore wind farm buses and five POIs. Nine candidate transmission cables are considered, including five that connect each wind farm bus to a POI and four backbone cables. The locations of these buses, the aggregated power rating of the wind turbines, and candidate cables are illustrated in Fig. 26. The wind farm area is enlarged on the left half of the figure. The goal is to solve for the optimal selection and size of the cables and ESSs such that total system profit is maximized.

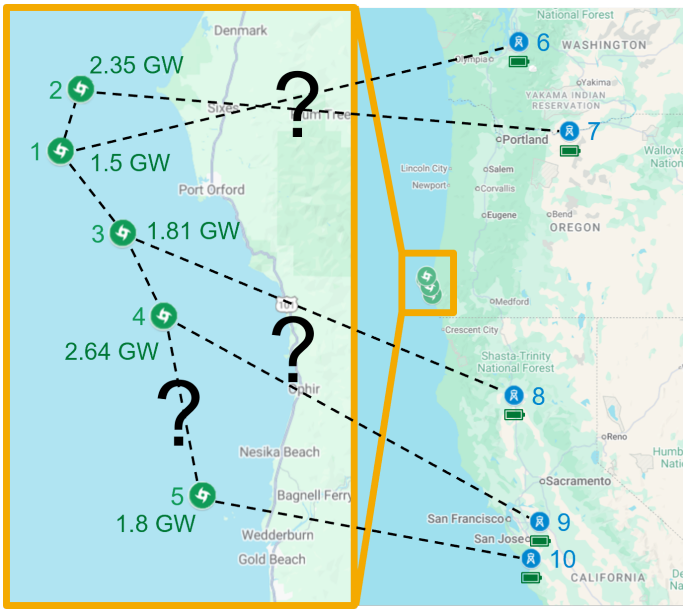


Figure 26: Location of buses, wind farm rated power, and candidate MTDC cables are shown. The wind farm area is enlarged on the left half.

The following five cases are included in this study:

- Radial: optimizes transmission cable sizes of radial topology wind farm buses to POIs and ESSs; 5% inflation rate.
- Mesh5: optimizes the transmission topology, ESS location, and ESS size; 5% inflation rate.
- Mesh5L: same as Mesh5 but each POI has a limited integration capacity of 3.5 GW for wind power.
- Mesh3 and Mesh8: same as Mesh5 but the inflation rates are 3% and 8%, respectively.

The radial topology case serves as a baseline for the study. Inflation rates of 3%, 5%, and 8% have been typical in recent years and are considered in this analysis. Case Mesh5L represents a very common situation where POIs have limited capacity for integrating renewable energy.

The problem is modeled in `Julia` with `JuMP`[158] and solved by `Ipopt` and `Juniper`[159].

4.4.2 Dataset

The data used in this study include the Offshore California Wind Speed Dataset from the NREL WIND Toolkit, along with DA and RT price data for the NP15 region of CAISO, the SCL region in Washington, and the BPAT region in Oregon. Due to limited access to up and

down reserve price data for Washington and Oregon, the same reserve prices from CAISO were applied across all POIs. Historical data from 2022 is extracted from these data sources and utilized to generate a scenario tree by `Julia` package `ScenTrees` [160], comprising 10 DA stage scenarios and 3 RT stage scenarios for each DA scenario.

MTDC cable cost is calculated based on the equations in [161], while converter costs are derived from [38]. ESS cost and efficiency are obtained from the PNNL Energy Storage Cost and Performance Database. The lower and upper limits of the ESS rated power are set as 10 kW and 1 MW, respectively, and 1% to 10% of ESS rated power can be used as reserve for fast frequency response. All ESSs have a 4-hour duration, a lifetime of 16 years, and 2400 life cycles.

4.5 Results

This section discusses the optimization results of designing the MTDC topology and the ESS siting and sizing for the aforementioned cases. Figure 27 illustrates the optimized topology of the MTDC grid.

Table 5 (A) presents the rated power of ESS at POIs; Table 5 (B) provides the size of cables, identified by origin and destination buses, with their lengths in parentheses; Table 5 (C) summarizes the system costs and profits. The MTDC cost includes both cable and converter costs. DA and RT market revenues are calculated over the 16 years of ESS life span and converted to present values.

The cases Radial, Mesh5, and Mesh5L, having the same inflation rate, generate similar revenue. Among these cases, Mesh5 requires the lowest cost for installing and constructing the ESS and MTDC system due to the fewest limitations. Only two cables, between the onshore and offshore sides connecting buses 2–7 and 3–8, are identified through optimization, also highlighting the optimizer's solution to choose their short-length cable path. The Radial case involves building long-distance cables, resulting in much higher costs. Additionally, the five ESSs installed at the POIs further increase the system cost. In case Mesh5L, with limited capacity for integrating wind energy at the POIs, another cable connecting buses 1–6 is constructed to address the limitation. All connected POIs have an ESS installed to facilitate system operation. In cases Mesh3 and Mesh8, different inflation rates result in different total revenues over the ESS lifetime. They share the same MTDC system, which has the most concentrated transmission capacity from offshore to onshore and the lowest cost among all five cases. Overall, the MTDC systems in cases Mesh3, Mesh5, and Mesh8 are similar in terms of construction cost.

The cost of ESS is mainly covered by providing reserve services and arbitraging the market; therefore, ESS size is very sensitive to the inflation rate. With a lower inflation rate, the ESS can generate more revenue, allowing for the adoption of a larger ESS, as seen in the case Mesh3. In contrast, the revenue under a higher inflation rate can only cover the cost of a smaller ESS, as seen in case Mesh8. Comparing with total system revenue, the ESSs in cases Mesh5 and Mesh8 can be viewed as equivalent. It may be desired to expand the role of ESSs, such as using them for output stabilization, to fully utilize the devices.

To investigate system operation and market participation, more detailed information for case Mesh5 in DA stage scenario 7 is plotted in Fig. 28, including market prices, power traded in the two markets, total generation of connected wind farm buses, and ESS SoC of the two connected POIs, respectively. The two POIs have the same RT prices, while the DA price of POI 8 is slightly higher than that of POI 7. They share one RT trading limit as described in constraints (83-84). At POI 7, no power is traded in the DA market; all power is sold to the RT market. At POI 8, the commitment to the DA market is close to the actual

Table 5: Design and Economic Results

(A) ESS Rated Power (kW)					
POI	Radial	Mesh5	Mesh5L	Mesh3	Mesh8
6	11.71	0	11.01	0	0
7	11.71	11.01	10.33	1000	10.36
8	11.71	11.01	10.14	1000	10.36
9	11.71	0	0	0	0
10	11.71	0	0	0	0
(B) Cable Loading (MW)					
1–6	1650	0	3410	0	0
2–7	2585	2585	3850	2398	2398
3–8	1991	8525	3850	8712	8712
4–9	2904	0	0	0	0
5–10	1980	0	0	0	0
1–2	0	0	1265	187	187
1–3	0	1650	3025	1837	1837
3–4	0	4884	4884	4884	4884
4–5	0	1980	1980	1980	1980
(C) Cost and Revenue ($B, except ESS cost is in $k)					
ESS Cost ($k)	52.064	19.584	27.981	1778.128	18.425
MTDC Cost	6.369	4.674	5.342	4.668	4.668
Total Cost	6.369	4.674	5.342	4.670	4.668
DA Revenue	58.703	58.703	58.703	66.744	58.703
RT Revenue	0.427	0.427	0.427	0.485	0.427
Total Revenue	59.130	59.130	59.130	67.229	49.672
Net Profit	52.761	54.456	53.788	62.558	45.003

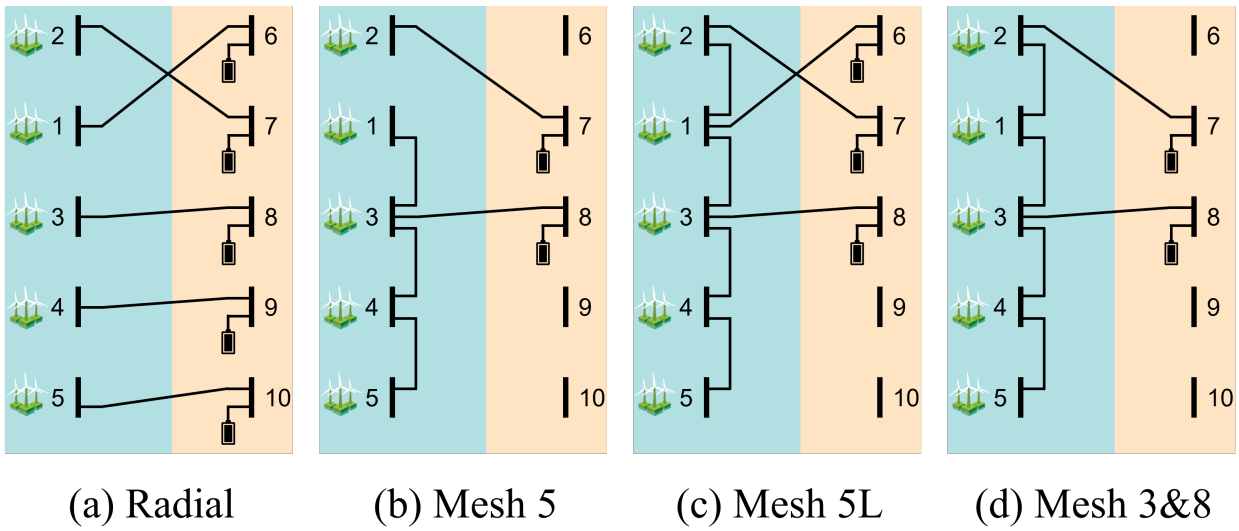


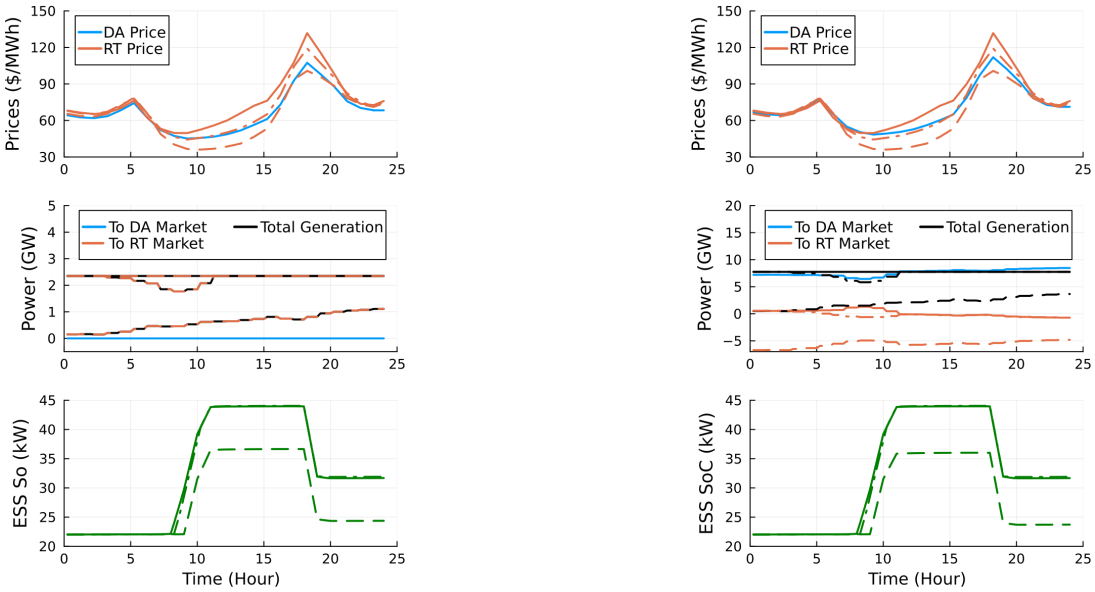
Figure 27: Optimized MTDC topology for each case is shown.

power generation in two of the three RT stage scenarios, and the RT market is used to mitigate the difference. In both POIs, ESSs are charged when market prices are low and discharged when prices are high to generate revenue.

4.6 Conclusion & Future work

In this report, we propose a methodology for simultaneously identifying the optimal MTDC transmission topology of offshore wind farms with the location and size of energy storage while considering energy market participation. The methodology accounts for stochastic variations in wind and energy prices at the wind farm and POIs, respectively. We analyzed and compared the topology design and energy storage location and sizing under different inflation conditions, benchmarking these cases against a radial base case scenario. The results highlight the importance of co-designing the topology compared to the radial case, demonstrating a higher net profit.

This research has the following future directions to explore: (1) designing a more reliable MTDC topology that accounts for contingencies, such as loss of cables; (2) expanding the role of the ESS to include functions like output smoothing and stabilization; (3) enabling wind farms to participate in the ancillary services market.



(a) POI 7

(b) POI 8

Figure 28: System operation of case Mesh5 in DA stage scenario 7. The 3 RT stage scenarios are marked with solid, dash, and dash-dot lines. For the two connected POIs, the first row shows DA and RT market prices; the second row shows power traded in the two markets (positive for selling) and total power generation from connected wind farm buses; the third row shows ESS state of charge.

5.0 Collaboration with various ECOMP initiative project

The project actively engaged with other ECOMP initiative projects to collaborate on various project tasks. Figure 29 shows the connections and collaborative exchanges during fiscal year 2024 (FY24), and Figure 30 shows the collaboration during fiscal year 2025 (FY25). The connected arrows in these figures describe the exchange of information between the MoCoDO project and other ECOMP initiative projects.

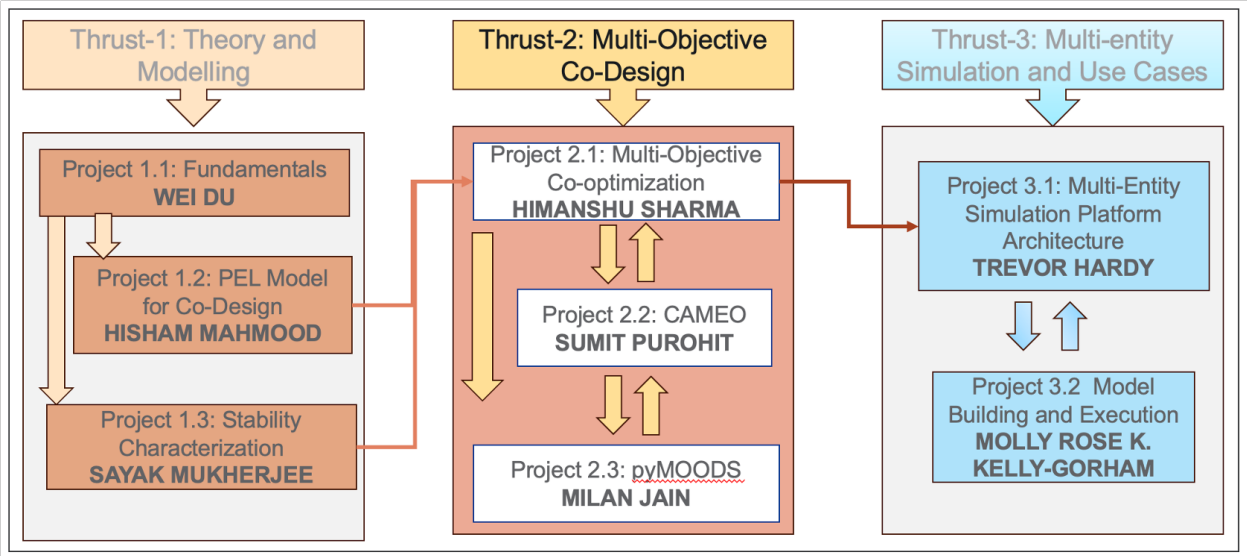


Figure 29: The ECOMP initiative FY24 projects description with arrows defining the interactions of MoCoDO project-2.1 with other initiative projects.

The following sections provide detailed descriptions of the information exchange and collaboration between projects.

5.1 Collaboration with Thrust-1 projects

This section describes the collaborative tasks that the MoCoDO project undertook with Thrust-1 projects.

5.1.1 Validation of Control Co-design Solution

In this task, the project coordinated with ECOMP Thrust-1 project 1.2 in FY24, titled ”Power electronics model for co-design” (led by Hisham Mahmood). The objectives of this collaboration were to validate the control co-design solutions (droop gains) used for providing reserve services to the power grid through faster time-scale simulations (PSSE/PSCAD). The detailed results of this collaboration are described in Section 2.4.6.

5.1.2 Protection System Control Co-design

In this task, the project coordinated with ECOMP Thrust-1 project 1D in FY25, titled ”Real-time test bed for co-design” (led by Quan Nguyen). The objective of this collaboration was to

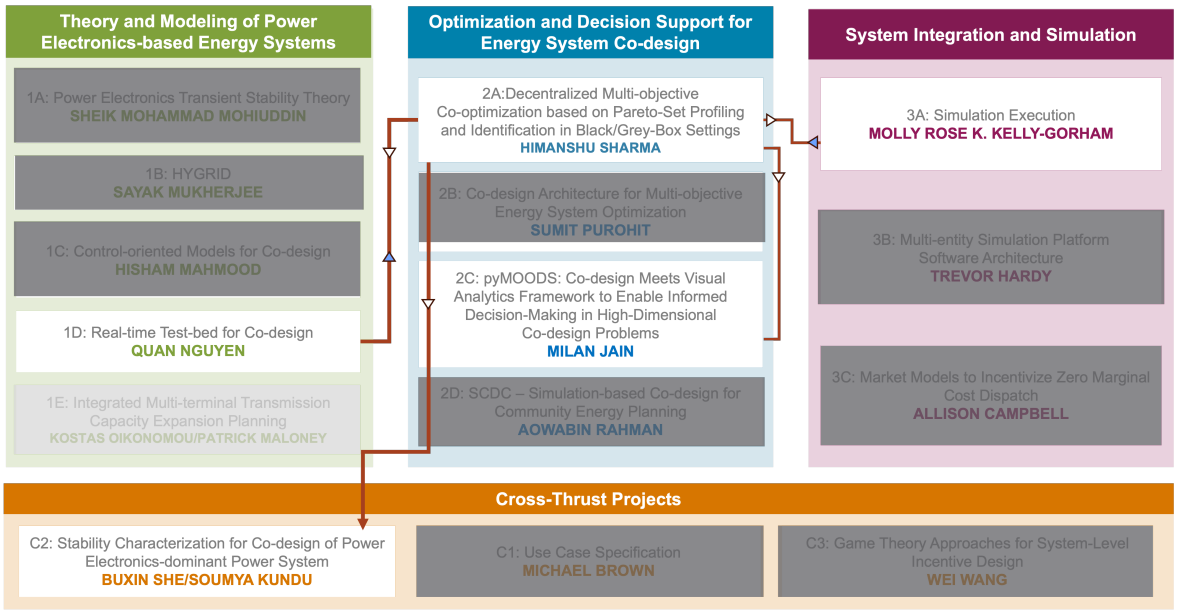


Figure 30: The ECOMP initiative FY25 projects description with arrows defining the interactions of MoCoDO project-2.1 with other initiative projects.

develop a strategy for analyzing the placement of MTDC protection devices on the MTDC network for an MTDC topology identified through the methodology defined in Section-4.2.

Figure 31 describes the workflow developed for this collaborative analysis. In step (a), the co-design framework optimizes the siting and sizing of energy storage with the MTDC topology co-design framework. The identified connections between generation and POI are passed to project-1D for developing the MTDC network protection scheme design in the PSSE/PSCAD simulation environment, shown in step (b). The identified protection scheme with inverter-converted breakers (blue dots) and line breakers (orange dots) can be seen in Figure 31(b), with associated costs provided in this step.

In the next step, fault scenarios are identified, and based on the circuit breaker states, the MTDC co-design energy market optimization formulation is re-run to evaluate the impact on system revenue by including protection devices versus various failure probabilities of fault scenarios, shown in Figure 31(d). The results show that the circuit breaker state (ON/OFF), depending on MTDC faults, affects revenue. The change in revenue is associated with the type of fault observed by the MTDC network, with faster changes observed in cases with larger power loss injections to the onshore POI (6, 7, 8).

5.1.3 Validation of MTDC Topology Co-design solutions for grid stability

In this project task, the team collaborated with ECOMP cross-thrust project-C2 in FY25, titled ”Stability characterization for co-design of power electronics dominant power systems” (led by Buxin She/Soumya Kundu). The objective of this collaboration was to ensure that the MTDC topology design solutions identified by solving the formulation described in Section 4.2 are stable.

We provided the MTDC design topology and export cable characteristics described in Section-4.2 to project-C2. The topologies were then analyzed by project-C2 to understand their stability characteristics, and the results of this analysis can be found in the project-C2

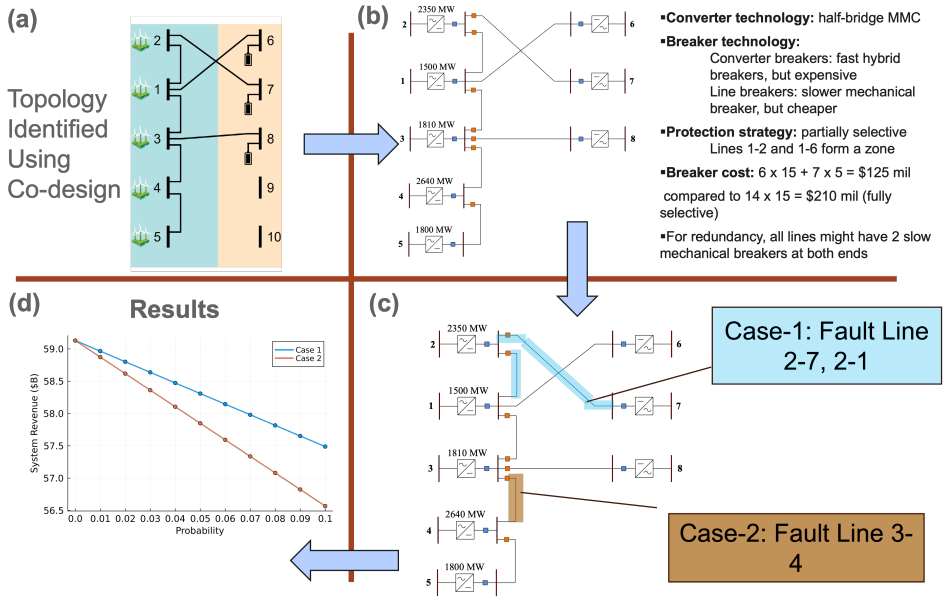


Figure 31: Workflow diagram defining the collaborative exchange of information between projects.

report. As part of future work, the methodology developed in Section-4.2 can be revised to consider feedback from the project-C2 stability analysis to revise or reformulate the MTDC topology optimization problem.

5.2 Collaboration with Thrust-2 projects

This section describes the collaborative tasks that the MoCoDO project undertook with Thrust-2 projects. The collaborative projects shared a common thrust in both FY24 and FY25.

5.2.1 CAMEO: A Co-design Architecture for Multi-objective Energy System Optimization

In this task, the MoCoDO project team collaborated with ECOMP FY24 Thrust-2 project 2.2: CAMEO: A Co-design Architecture for Multi-objective Energy System Optimization (led by Sumit Purohit). The objective of this collaboration was to provide the multi-objective optimization formulation to develop a co-design architecture software library for setting up design problems. Our project provided the multi-objective formulation described in Section 1.7.2-1.2 to develop the backend framework of CAMEO for building design space exploration for defined co-design problems via a modular and automated workflow system, enhancing flexibility and accelerating design and validation cycles. The CAMEO tool was developed to support cloud-scale automation with a user-friendly interface, enabling energy system practitioners to efficiently explore diverse design alternatives. Detailed results from this collaboration can be seen in [162].

5.2.2 PyMOODS: Multi-objective decision support system

In this task, the MoCoDO project team collaborated with ECOMP FY24 project-2.3 and FY25 project 2C, titled PyMOODS: Multi-objective decision support system (led by Milan Jain). The objective of this collaboration was to develop a visualization tool for visualizing multi-objective high-dimensional solutions and providing additional support to stakeholders who need to weigh trade-offs between solutions. The multi-objective Pareto front solutions obtained by solving the co-design problem defined in Section 1.5 were provided to support the development of visuals and a web-user interface. Additionally, the co-design optimization data were used to develop visualization techniques such as generalizers and specializers, tradeoff lattices, and uncertainty-aware scenario characterization to parse and analyze solution data for additional insights. The results of this collaboration can be seen in reports and publications made by project 2C. As part of future work, we will integrate the visualization tool into the optimization library to perform in-situ visualization of solutions as optimization progresses.

5.3 Collaboration with Thrust-3 project

This section describes the collaborative tasks that the MoCoDO project undertook with Thrust-3 projects.

5.3.1 Offshore wind farm co-design and online market bidding strategy

In this task, the MoCoDO project team collaborated with ECOMP FY24 project-3.1 and FY25 project 3A. The objectives of this collaboration were to provide the control co-design solution obtained by solving the formulation presented in Section 1.2 and Section-4.2, shown in Figure 32(a). The next step was to provide and develop the online bidding strategy to participate in the energy market defined in the Multi-entity Simulation Platform (MESP) developed by project 3A. Depending on the interconnection topology type—either radial or meshed (MTDC)—used to integrate offshore wind farms with onshore points of interconnection (POI), the bidding strategy is defined accordingly. Finally, once annual simulations are completed by MESP, we compare the revenue projected in the offline co-design formulation with that in the online setting by interacting through bids in the energy market with additional generators on the AC side.

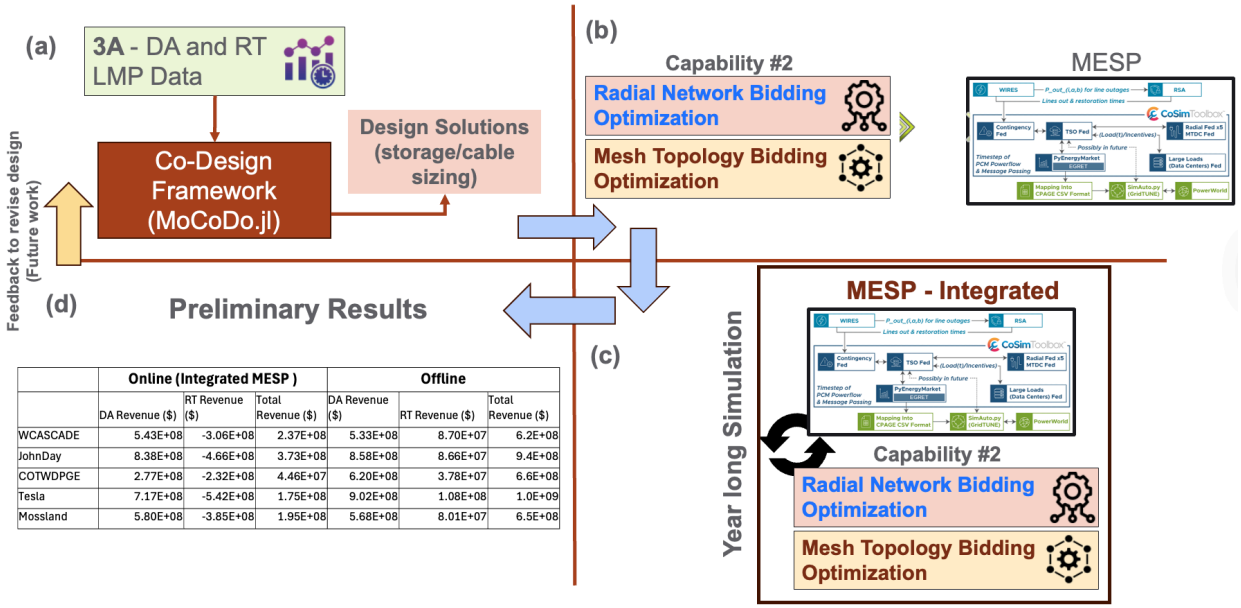
The following sections discuss additional details about the bidding strategy based on the optimization formulation. The nomenclature for mathematical notations is defined in Table 6.

Radial Topology Day Ahead (DA) Market: When making DA bids, hourly wind speed forecasts for the next day are provided as scenarios by MESP (Figure 32(b)). The wind speed in each scenario at each hour is converted to maximum wind power generation $P_{t,s}^{WP}$ based on maximum power point tracking. The optimization objective has two terms: the first term maximizes DA bidding, while the second term minimizes the expected difference between DA bidding and anticipated RT bidding. These two terms are balanced by a weight factor ω. Additional power flow, reserve, and energy storage dynamics are included to solve the problem as described below.

$$\max \quad \sum_{t\in\mathcal{T}} p_t^{DA} - \frac{\omega}{|\mathcal{S}|}\sum_{t\in\mathcal{T}}\sum_{s\in\mathcal{S}} p_{t,s}^{diff} \tag{94}$$

$$\text{s.t.} \quad p_{t,s}^{diff} \geq p_t^{DA} - p_{t,s}^{WR} \quad t\in\mathcal{T}, s\in\mathcal{S} \tag{95}$$

$$p_{t,s}^{diff} \geq p_{t,s}^{WR} - p_t^{DA} \quad t\in\mathcal{T}, s\in\mathcal{S} \tag{96}$$



	Online (Integrated MESP)			Offline		
	DA Revenue ($)	RT Revenue ($)	Total Revenue ($)	DA Revenue ($)	RT Revenue ($)	Total Revenue ($)
WCASCADE	5.43E+08	-3.06E+08	2.37E+08	5.33E+08	8.70E+07	6.2E+08
JohnDay	8.38E+08	-4.66E+08	3.73E+08	8.58E+08	8.66E+07	9.4E+08
COTWDPGE	2.77E+08	-2.32E+08	4.46E+07	6.20E+08	3.78E+07	6.6E+08
Tesla	7.17E+08	-5.42E+08	1.75E+08	9.02E+08	1.08E+08	1.0E+09
Mossland	5.80E+08	-3.85E+08	1.95E+08	5.68E+08	8.01E+07	6.5E+08

Figure 32: Collaborative workflow diagram between MoCoDO and Thrust-3 projects showing the integration of co-design solutions with online market bidding strategies.

$$p_{t,s}^{RU} = p_{t,s}^{ResWU}(1 - P^R/g/V_{\text{max}}^2) + p_{t,s}^{ResBU} \quad t \in \mathcal{T}, s \in \mathcal{S} \tag{97}$$

$$p_{t,s}^{RD} = p_{t,s}^{ResWD} + p_{t,s}^{ResBD} \quad t \in \mathcal{T}, s \in \mathcal{S} \tag{98}$$

$$p_{t,s}^{W} = V_{1,t,s}(V_{1,t,s} - V_{2,t,s})g \quad t \in \mathcal{T}, s \in \mathcal{S} \tag{99}$$

$$(p_{t,s}^{dis} - p_{t,s}^{ch}) - p_{t,s}^{WR} = V_{2,t,s}(V_{2,t,s} - V_{1,t,s})g \quad t \in \mathcal{T}, s \in \mathcal{S} \tag{100}$$

$$V_{\text{min}} \leq V_{1,t,s}, V_{2,t,s} \leq V_{\text{max}} \quad t \in \mathcal{T}, s \in \mathcal{S} \tag{101}$$

$$V_{1,t,s}(V_{1,t,s} - V_{2,t,s})g \leq \text{CabCap} \quad t \in \mathcal{T}, s \in \mathcal{S} \tag{102}$$

$$p_{t,s}^{W} + p_{t,s}^{ResWU} \leq P_{t,s}^{WP} \quad t \in \mathcal{T}, s \in \mathcal{S} \tag{103}$$

$$p_{t,s}^{W} - p_{t,s}^{ResWD} \geq 0 \quad t \in \mathcal{T}, s \in \mathcal{S} \tag{104}$$

$$p_{t,s}^{ResWU} = k_{t,s}^{ResW} \Delta f_{\text{max}}^{U} \quad t \in \mathcal{T}, s \in \mathcal{S} \tag{105}$$

$$p_{t,s}^{ResWD} = k_{t,s}^{ResW} \Delta f_{\text{max}}^{D} \quad t \in \mathcal{T}, s \in \mathcal{S} \tag{106}$$

$$p_{t,s}^{ResBU} = k_{t,s}^{ResB} \Delta f_{\text{max}}^{U} \quad t \in \mathcal{T}, s \in \mathcal{S} \tag{107}$$

$$p_{t,s}^{ResBD} = k_{t,s}^{ResB} \Delta f_{\text{max}}^{D} \quad t \in \mathcal{T}, s \in \mathcal{S} \tag{108}$$

$$P_{t,s}^{WP}/0.5 \leq k_{t,s}^{ResW} \leq P_{t,s}^{WP}/0.1 \quad t \in \mathcal{T}, s \in \mathcal{S} \tag{109}$$

$$\text{BatSize}/0.5 \leq k_{t,s}^{ResB} \leq \text{BatSize}/0.01 \quad t \in \mathcal{T}, s \in \mathcal{S} \tag{110}$$

$$k_{t,s}^{ResW} + k_{t,s}^{ResB} \geq P_{t,s}^{WP}/0.2 \quad t \in \mathcal{T}, s \in \mathcal{S} \tag{111}$$

$$SC_{t,s} - SC_{t-1,s} = \eta^{ch} p_{t,s}^{ch} - \eta^{dis} p_{t,s}^{dis} \quad t \in \mathcal{T}, s \in \mathcal{S} \tag{112}$$

$$p_{t,s}^{ch} + p_{t,s}^{ResBD} \leq \text{BatSize} \quad t \in \mathcal{T}, s \in \mathcal{S} \tag{113}$$

$$p_{t,s}^{dis} + p_{t,s}^{ResBU} \leq \text{BatSize} \quad t \in \mathcal{T}, s \in \mathcal{S} \tag{114}$$

$$0 \leq SC_{t,s} \leq \text{BatSize} * \text{DurH} \quad t \in \mathcal{T}, s \in \mathcal{S} \tag{115}$$

Table 6: Nomenclature: OWF online bidding formulation

Symbol	Definition	Symbol	Definition
$\mathcal{T}$	set of time indices	$p_{t,s}^{ResWU}$	up reserve power from wind farm
$\mathcal{S}$	set of all scenarios	$p_{t,s}^{ResBU}$	up reserve power from battery
t	time-index	$p_{t,s}^{ResWD}$	down reserve power from wind farm
s	individual scenario	$p_{t,s}^{ResBD}$	down reserve power from battery
ω	stochastic scenario	$p_{t,s}^{W}$	wind power generation
g	conductance of export cable	$p_{t,s}^{dis}$	battery discharge power
p_t^{DA}	power traded in DA at time t	$p_{t,s}^{ch}$	battery charge power
$p_{t,s}^{diff}$	difference between DA and RT bidding	p^{chR}	battery charge power by rule
p^{diff}	difference between RT bid and dispatch	p^{disR}	battery discharge power by rule
$p_{t,s}^{WR}$	anticipated RT bidding	P^R	wind farm rated power
$p_{t,s}^{RU}$	up reserve power	$P_{t,s}^{WP}$	wind power by MPPT
$p_{t,s}^{RD}$	down reserve power	$V_{1,t,s}$	bus voltage at wind farm
$V_{2,t,s}$	bus voltage at POI	$\Delta f_{\text{max}}^{U}$	max frequency deviation (up)
V_{min}	minimum voltage limits	$\Delta f_{\text{max}}^{D}$	max frequency deviation (down)
V_{max}	maximum voltage limits	$SC_{t,s}$	state of charge of battery
$k_{t,s}^{ResW}$	droop parameter (wind farm)	SC_{ini}	initial state of charge
$k_{t,s}^{ResB}$	droop parameter (battery)	η^{ch}	charging efficiency
η^{dis}	discharging efficiency	CabCap	export cable capacity
BatSize	battery rated power	DurH	battery storage duration (hours)
TimeLength	time step length in dispatch		

$$\sum_{s\in\mathcal{S}} SC_{t,s} = 0.5 * \text{BatSize} * \text{DurH} * |\mathcal{S}| \quad \text{for } t = \text{T end} \tag{116}$$

Radial Topology Real time (RT) market: Radial Topology Real-time (RT) Market: For RT bidding, battery charging and discharging power are determined by rules as p^{chR} and p^{disR},respectively. Specifically, if the battery's state of charge is less than 10% of its energy capacity, it is charged at 5% of its power rating; if its state of charge is greater than 90% of its energy capacity, it is discharged at 5% of its power rating; otherwise, there is no battery operation. By applying this rule, the battery maintains an appropriate state level to guarantee sufficient flexibility in the dispatch stage.

$$\max \quad p^{WR} \tag{117}$$
$$\text{s.t.} \quad p^{RU} = p^{ResWU}(1 - P^R/g/V_{\text{max}}^2) + p^{ResBU} \tag{118}$$
$$p^{W} = V_1(V_1 - v_2)g \tag{119}$$
$$p^{ch} = p^{chR} \tag{120}$$
$$p^{dis} = p^{disR} \tag{121}$$
$$(p^{dis} - p^{ch}) - p^{WR} = V_2(v_2 - v_1)g \tag{122}$$

$$V_{\min} \leq V_1, V_2 \leq V_{\max} \tag{123}$$
$$V_1(V_1 - V_2)g \leq \text{CabCap} \tag{124}$$
$$p^W + p^{ResWU} \leq P^{WP} \tag{125}$$
$$p^W - p^{ResWD} \geq 0 \tag{126}$$
$$p^{ResWU} = k^{ResW}\Delta f^U_{\max} \tag{127}$$
$$p^{ResWD} = k^{ResW}\Delta f^D_{\max} \tag{128}$$
$$p^{ResBU} = k^{ResB}\Delta f^U_{\max} \tag{129}$$
$$p^{ResBD} = k^{ResB}\Delta f^D_{\max} \tag{130}$$
$$P^{WP}/0.5 \leq k^{ResW} \leq P^{WP}/0.1 \tag{131}$$
$$\text{BatSize}/0.5 \leq k^{ResB} \leq \text{BatSize}/0.01 \tag{132}$$
$$k^{ResW} + k^{ResB} \geq P^{WP}/0.2 \tag{133}$$
$$p^{ch} + p^{ResBD} \leq \text{BatSize} \tag{134}$$
$$p^{dis} + p^{ResBU} \leq \text{BatSize} \tag{135}$$

Radial Topology Dispatch: Radial Topology Dispatch: In the dispatch stage, the p^{chR} and p^{disR} determined during RT bidding are used as minimum battery charging and discharging power, respectively. The details related to dispatch are defined as follows:

$$\min \quad p^{diff} \tag{136}$$
$$\text{s.t.} \quad p^{diff} \geq p^{RT} - p^{WD} \tag{137}$$
$$p^{diff} \geq p^{WD} - p^{RT} \tag{138}$$
$$p^{RU} = p^{ResWU}(1 - P^R/g/V^2_{\max}) + p^{ResBU} \tag{139}$$
$$p^W = V_1(V_1 - V_2)g \tag{140}$$
$$p^{ch} \geq p^{chR} \tag{141}$$
$$p^{dis} \geq p^{disR} \tag{142}$$
$$p^{ch}p^{dis} = 0 \tag{143}$$
$$(p^{dis} - p^{ch}) - p^{WD} = V_2(V_2 - V_1)g \tag{144}$$
$$V_{\min} \leq V_1, V_2 \leq V_{\max} \tag{145}$$
$$V_1(V_1 - V_2)g \leq \text{CabCap} \tag{146}$$
$$p^W + p^{ResWU} \leq P^{WP} \tag{147}$$
$$p^W - p^{ResWD} \geq 0 \tag{148}$$
$$p^{ResWU} = k^{ResW}\Delta f^U_{\max} \tag{149}$$
$$p^{ResWD} = k^{ResW}\Delta f^D_{\max} \tag{150}$$
$$p^{ResBU} = k^{ResB}\Delta f^U_{\max} \tag{151}$$
$$p^{ResBD} = k^{ResB}\Delta f^D_{\max} \tag{152}$$
$$P^{WP}/0.5 \leq k^{ResW} \leq P^{WP}/0.1 \tag{153}$$
$$\text{BatSize}/0.5 \leq k^{ResB} \leq \text{BatSize}/0.01 \tag{154}$$
$$k^{ResW} + k^{ResB} \geq P^{WP}/0.2 \tag{155}$$
$$SC - SC_{ini} = (\eta^{ch}p^{ch} - \eta^{dis}p^{dis}) * \text{TimeLength} \tag{156}$$

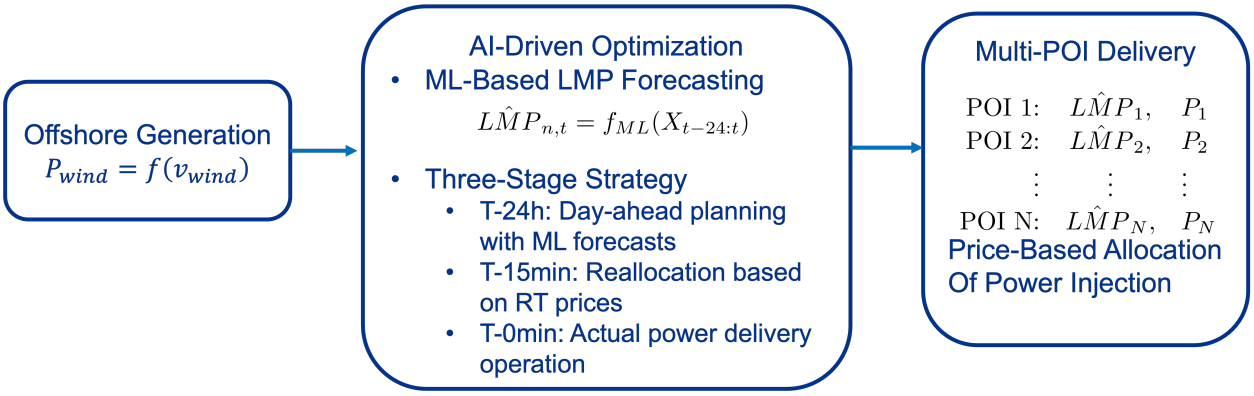


Figure 33: Workflow Overview of the AI-Driven Optimization-Based Energy Market Bidding Strategy for Offshore Wind Farms.

$$p^{ch} + p^{ResBD} \leq \text{BatSize} \tag{157}$$
$$p^{dis} + p^{ResBU} \leq \text{BatSize} \tag{158}$$
$$0 \leq SC \leq \text{BatSize} * \text{DurH} \tag{159}$$

Machine Learning Model Based Energy Market Bidding: We extended our optimization framework to a meshed case by incorporating a predictive model for Locational Marginal Price (LMP) forecasting at the point of interconnection. The objective functions were revised to include LMP predictions as follows:

$$\max \sum_{n \in N} \sum_{t \in T} \widetilde{LMP}_{n,t} \cdot P^{DA}_{n,t} - \lambda \sum_{n,t,s} |P^{actual}_{n,t,s} - P^{DA}_{n,t}| \tag{160}$$

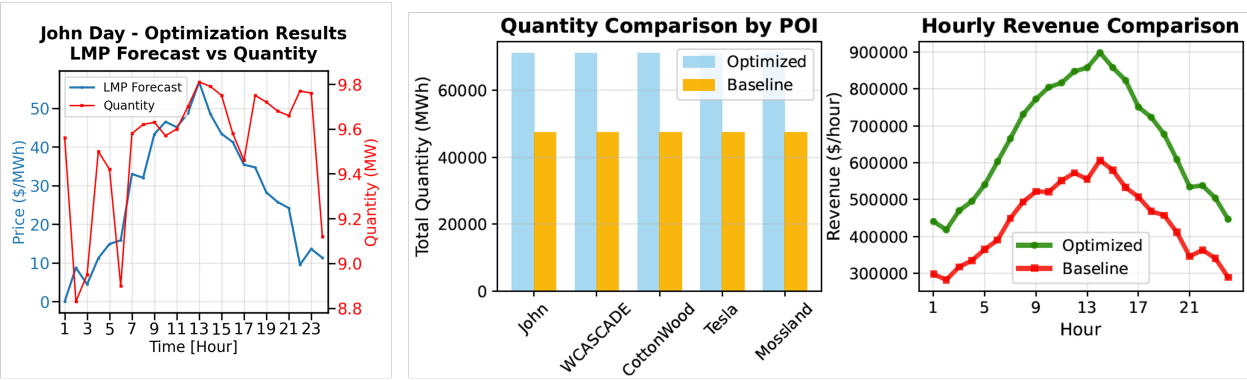


Figure 34: Comparison of LMP Forecasts and Bidding Quantities (JhonDay POI) Analyzing the Impact of LMP Forecasts on Bidding, with Hourly Revenue Comparison Against a Baseline Fixed-Rate LMP Case.

The forecasted day-ahead Locational Marginal Price (LMP) at the offshore wind farm's point of interconnection is denoted as $\widetilde{LMP}$. The original formulation maintains the constraints on power balance and energy storage dynamics. For Multi-Terminal DC (MTDC) networks, LMP predictions were generated for each interconnected Point of Interconnection (POI) to

determine the optimal power quantity based on the forecast. Figure 33 illustrates a high-level workflow overview of this AI-driven, optimization-based energy market bidding strategy.

Offline testing of the predictive bidding strategy, as depicted in Fig. 34, yielded several key findings: bidding quantities closely aligned with LMP forecasts, demonstrating price-responsive bidding; wind power supplied approximately 60% of the total generation, establishing a stable renewable base; a 50.1% increase in revenue, generating an additional $5.36 million compared to the baseline fixed-LMP strategy; and a 50% increase in energy delivery through optimized quantity allocation across POIs. Moreover, battery storage was effectively discharged during periods of high LMP, which provided operational flexibility and enhanced revenue. The optimized dispatch strategy effectively coordinated wind and storage resources to maximize market value.

Future work will extend the approach to energy market participation strategies for radial topologies. Additionally, the current results, generated in an offline setting without considering other AC grid generation participating and modeled via the Multi-Entity Simulation Platform (MESP), will be analyzed in an online setting to assess the overall impact of the proposed strategy.

References

[1] W. Musial, P. Spitsen, P. Beiter, *et al.*, "Offshore wind market report: 2023 edition," 2023. [Online]. Available: https://www.energy.gov/eere/wind/articles/offshore-wind-market-report-2023-edition.

[2] J. A. Ansari, C. Liu, and S. A. Khan, "Mmc based mtdc grids: A detailed review on issues and challenges for operation, control and protection schemes," *IEEE Access*, vol. 8, pp. 168 154–168 165, 2020.

[3] H. Simão, W. Powell, C. Archer, and W. Kempton, "The challenge of integrating offshore wind power in the us electric grid. part ii: Simulation of electricity market operations," *Renewable energy*, vol. 103, pp. 418–431, 2017.

[4] C. Zhao, P. B. Andersen, C. Træholt, and S. Hashemi, "Grid-connected battery energy storage system: A review on application and integration," *Renewable and Sustainable Energy Reviews*, vol. 182, p. 113 400, 2023.

[5] H. Zhao, Q. Wu, S. J. Hu, H.-h. Xu, and C. N. Rasmussen, "Review of energy storage system for wind power integration support," *Applied Energy*, vol. 137, pp. 545–553, 2015. [Online]. Available: https://api.semanticscholar.org/CorpusID:10668291.

[6] N. Halwany, D. Pagnani, M. Ledro, O. E. Idehe, M. Marinelli, and L. Kocewiak, "Optimal sizing of battery energy storage to enable offshore wind farm black start operation," in *21st Wind & Solar Integration Workshop (WIW 2022)*, IET, vol. 2022, 2022, pp. 232–240.

[7] S. Paul, A. P. Nath, and Z. H. Rather, "A multi-objective planning framework for coordinated generation from offshore wind farm and battery energy storage system," *IEEE Transactions on Sustainable Energy*, vol. 11, no. 4, pp. 2087–2097, 2019.

[8] I. N. Moghaddam, B. H. Chowdhury, and S. Mohajeryami, "Predictive operation and optimal sizing of battery energy storage with high wind energy penetration," *IEEE Transactions on Industrial Electronics*, vol. 65, no. 8, pp. 6686–6695, 2017.

[9] J. T. Allison, T. Guo, and Z. Han, "Co-design of an active suspension using simultaneous dynamic optimization," *ASME Journal of Mechanical Design*, vol. 136, no. 8, p. 081 003, 2014.

[10] A. P. "Deshmukh and J. T. Allison, ""multidisciplinary dynamic optimization of horizontal axis wind turbine design".," *Structural and Multidisciplinary Optimization*, vol. 53, no. 1, pp. 15–27, 2016.

[11] H. .-. "Yan and G. .-. Yan, ""integrated control and mechanism design for the variable input-speed servo four-bar linkages"," *"Mechatronics"*, vol. 19, no. 2, pp. 274–285, 2009.

[12] T. Liu, S. Azarm, and N. Chopra, "Decentralized multisubsystem co-design optimization using direct collocation and decomposition-based methods," *Journal of Mechanical Design*, vol. 142, no. 9, 2020.

[13] A. Bhattacharya, S. Vasisht, V. Adetola, S. Huang, H. Sharma, and D. L. Vrabie, "Control co-design of commercial building chiller plant using bayesian optimization.," *Energy and Buildings*, vol. 246, p. 111 077, 2021.

[14] A. L. Nash and N. Jain, “Hierarchical control co-design using a model fidelity-based decomposition framework,” *Journal of Mechanical Design*, vol. 143, no. 1, 2021.

[15] A. L. Nash, H. C. Pangborn, and N. Jain, “Robust control co-design with receding-horizon mpc,” in *2021 American Control Conference (ACC)*, IEEE, 2021, pp. 373–379.

[16] M. Garcia-Sanz, “Control co-design: An engineering game changer,” *Advanced Control for Applications*, vol. 1, no. 1, 2019.

[17] S. Azad and D. R. Herber, “An overview of uncertain control co-design formulations,” *Journal of Mechanical Design*, vol. 145, no. 9, p. 091 709, 2023.

[18] J. T. Allison, D. S. Zalkind, and D. R. Herber, “Open-loop control co-design of semisubmersible floating offshore wind turbines using linear parameter-varying models,” *Journal of Mechanical Design*, vol. 146, pp. 041 704–1, 2024.

[19] A. K. Sundarrajan and D. R. Herber, “Towards a fair comparison between the nested and simultaneous control co-design methods using an active suspension case study,” in *2021 American Control Conference (ACC)*, IEEE, 2021, pp. 358–365.

[20] J. T. Allison, T. Guo, and Z. Han, “Co-design of an active suspension using simultaneous dynamic optimization,” *Journal of Mechanical Design*, vol. 136, no. 8, p. 081 003, 2014.

[21] K. Naik, S. Beknalkar, J. Reed, A. Mazzoleni, H. Fathy, and C. Vermillion, “Pareto optimal and dual-objective geometric and structural design of an underwater kite for closed-loop flight performance,” *Journal of Dynamic Systems, Measurement, and Control*, vol. 145, no. 1, p. 011 005, 2023.

[22] D. Hallac, C. Wong, S. Diamond, *et al.*, “Snapvx: A network-based convex optimization solver,” *The Journal of Machine Learning Research*, vol. 18, no. 1, pp. 110–114, 2017.

[23] J. Jalving, S. Abhyankar, K. Kim, M. Hereld, and V. M. Zavala, “A graph-based computational framework for simulation and optimisation of coupled infrastructure networks,” *IET Generation, Transmission & Distribution*, vol. 11, no. 12, pp. 3163–3176, 2017.

[24] M. Berger, D. Radu, G. Detienne, T. Deschuyteneer, A. Richel, and D. Ernst, “Remote renewable hubs for carbon-neutral synthetic fuel production,” *Frontiers in Energy Research*, p. 200, 2021.

[25] S. Abhyankar, J. Brown, E. M. Constantinescu, D. Ghosh, B. F. Smith, and H. Zhang, “Petsc/ts: A modern scalable ode/dae solver library,” *arXiv preprint arXiv:1806.01437*, 2018.

[26] J. Jalving, Y. Cao, and V. M. Zavala, “Graph-based modeling and simulation of complex systems,” *Computers & Chemical Engineering*, vol. 125, pp. 134–154, 2019.

[27] J. Jalving, S. Shin, and V. M. Zavala, “A graph-based modeling abstraction for optimization: Concepts and implementation in plasmo. jl,” *Mathematical Programming Computation*, vol. 14, no. 4, pp. 699–747, 2022.

[28] A. Gu, S. Lu, P. Ram, and L. Weng, “Min-max bilevel multi-objective optimization with applications in machine learning,” *arXiv preprint arXiv:2203.01924*, 2022.

[29] P. Bresesti, W. L. Kling, R. L. Hendriks, and R. Vailati, “Hvdc connection of offshore wind farms to the transmission system,” *IEEE Transactions on energy conversion*, vol. 22, no. 1, pp. 37–43, 2007.

[30] J. Liang, T. Jing, O. Gomis-Bellmunt, J. Ekanayake, and N. Jenkins, "Operation and control of multiterminal hvdc transmission for offshore wind farms," *IEEE Transactions on Power Delivery*, vol. 26, no. 4, pp. 2596–2604, 2011.

[31] M. Baradar, M. Ghandhari, D. Van Hertem, and A. Kargarian, "Power flow calculation of hybrid ac/dc power systems," in *2012 IEEE Power and Energy Society General Meeting*, IEEE, 2012, pp. 1–6.

[32] Y. Zhou, L. Zhao, W.-J. Lee, Z. Zhang, and P. Wang, "Optimal power flow in hybrid ac and multi-terminal hvdc networks with offshore wind farm integration based on semidefinite programming," in *2019 IEEE Innovative Smart Grid Technologies-Asia (ISGT Asia)*, IEEE, 2019, pp. 207–212.

[33] R. A. Jabr, "Radial distribution load flow using conic programming," *IEEE transactions on power systems*, vol. 21, no. 3, pp. 1458–1459, 2006.

[34] H. Ergun, J. Dave, D. Van Hertem, and F. Geth, "Optimal power flow for ac–dc grids: Formulation, convex relaxation, linear approximation, and implementation," *IEEE transactions on power systems*, vol. 34, no. 4, pp. 2980–2990, 2019.

[35] B. She, H. Mahmood, M. Elizondo, V. Adetola, and Y. Dong, "Configuration and emt simulation of the 240-bus miniwecc system integrating offshore wind farms (OWFs)," *arXiv preprint arXiv:2403.07988*, 2024.

[36] A. McCoy, W. Musial, R. Hammond, D. Mulas Hernando, P. Duffy, P. Beiter, *et al.*, "Offshore wind market report: 2024 edition," National Renewable Energy Laboratory (NREL), Golden, CO (United States), Tech. Rep., 2024. [Online]. Available: https://www.nrel.gov/docs/fy24osti/90525.pdf.

[37] P. Sharma, L. Rese, B. Wang, B. Vyakaranam, and S. Shah, "Grid strength analysis for integrating 30 gw of offshore wind generation by 2030 in the us eastern interconnection," 2023.

[38] G. Brinkman, M. Bannister, S. Bredenkamp, *et al.*, "Atlantic offshore wind transmission study," National Renewable Energy Laboratory (NREL), Golden, CO (United States), Tech. Rep., 2024.

[39] G. E. Barter, A. Robertson, and W. Musial, "A systems engineering vision for floating offshore wind cost optimization," *Renewable Energy Focus*, vol. 34, pp. 1–16, 2020.

[40] N. Flourentzou, V. G. Agelidis, and G. D. Demetriades, "Vsc-based hvdc power transmission systems: An overview," *IEEE Transactions on Power Electronics*, vol. 24, pp. 592–602, 2009. [Online]. Available: https://api.semanticscholar.org/CorpusID:41167598.

[41] L. Zhang, Y. Zou, J. Yu, *et al.*, "Modeling, control, and protection of modular multilevel converter-based multi-terminal hvdc systems: A review," *CSEE Journal of Power and Energy Systems*, vol. 3, pp. 340–352, 2017. [Online]. Available: https://api.semanticscholar.org/CorpusID:69632155.

[42] S. Mao, J. Chen, and M. Liu, "A review of the energy storage system as a part of power system: Modelling, simulation and prospect," *Electric Power Systems Research*, vol. 233, p. 110 448, 2024.

[43] Y. Arellano-Prieto, E. Chavez-Panduro, P. Salvo Rossi, and F. Finotti, "Energy storage solutions for offshore applications," *Energies*, vol. 15, no. 17, p. 6153, 2022.

[44] A. Rabanal, A. M. Smith, C. C. Ahaotu, and E. Tedeschi, “Energy storage systems for services provision in offshore wind farms,” *Renewable and Sustainable Energy Reviews*, vol. 200, p. 114 573, 2024.

[45] F. E. R. Commission *et al.*, “Essential reliability services and the evolving bulk-power system—primary frequency response,” *Order*, no. 842, pp. 2020–06, 2018.

[46] B. C. Eldridge and A. Somani, “Modeling and market design considerations for conventional and decarbonized resources,” Pacific Northwest National Laboratory (PNNL), Richland, WA (United States), Tech. Rep., 2022.

[47] J. Seel, A. D. Mills, and R. H. Wiser, “Impacts of high variable renewable energy futures on wholesale electricity prices, and on electric-sector decision making,” 2018.

[48] C. Loutan, V. Gevorgian, S. Chowdhury, *et al.*, “Avangrid renewables tule wind farm: Demonstration of capability to provide essential grid services,” *CAISO, Folsom, CA*, 2020.

[49] H. Holttinen, J. Kiviluoma, N. Helistö, *et al.*, “Design and operation of energy systems with large amounts of variable generation: Final summary report, iea wind tcp task 25,” 2021.

[50] K. Kölle, T. Göçmen, I. Eguinoa, *et al.*, “Farmconners market showcases results: Wind farm flow control considering electricity prices and revenue,” *Wind Energy Science Discussions*, vol. 2022, pp. 1–29, 2022.

[51] J. H. Kim, F. Kahrl, A. Mills, R. Wiser, C. C. Montañés, and W. Gorman, “Economic evaluation of variable renewable energy participation in us ancillary services markets,” *Utilities Policy*, vol. 82, p. 101 578, 2023.

[52] I. Eguinoa, T. Göçmen, P. B. Garcia-Rosa, *et al.*, “Wind farm flow control oriented to electricity markets and grid integration: Initial perspective analysis,” *Advanced Control for Applications: Engineering and Industrial Systems*, vol. 3, no. 3, e80, 2021.

[53] M. Dicorato, G. Forte, M. Pisani, and M. Trovato, “Planning and operating combined wind-storage system in electricity market,” *IEEE Transactions on Sustainable Energy*, vol. 3, no. 2, pp. 209–217, 2012.

[54] M. Y. Nguyen, D. H. Nguyen, and Y. T. Yoon, “A new battery energy storage charging/discharging scheme for wind power producers in real-time markets,” *Energies*, vol. 5, no. 12, pp. 5439–5452, 2012.

[55] I. N. Moghaddam, B. Chowdhury, and M. Doostan, “Optimal sizing and operation of battery energy storage systems connected to wind farms participating in electricity markets,” *IEEE Transactions on Sustainable Energy*, vol. 10, no. 3, pp. 1184–1193, 2018.

[56] K. Das, A. L. T. P. Grapperon, P. E. Sørensen, and A. D. Hansen, “Optimal battery operation for revenue maximization of wind-storage hybrid power plant,” *Electric Power Systems Research*, vol. 189, p. 106 631, 2020.

[57] P. Hou, G. Yang, P. Enevoldsen, and A. H. Nielsen, “Cooperation of offshore wind farm with battery storage in multiple electricity markets,” in *2018 53rd international universities power engineering conference (UPEC)*, IEEE, 2018, pp. 1–6.

[58] A. Bechlenberg, E. A. Luning, M. B. Saltık, N. B. Szirbik, B. Jayawardhana, and A. I. Vakis, “Renewable energy system sizing with power generation and storage functions accounting for its optimized activity on multiple electricity markets,” *Applied Energy*, vol. 360, p. 122 742, 2024.

[59] S. Varotto, V. Trovato, E. Kazemi-Robati, and B. Silva, “Optimal sizing and energy management of battery energy storage systems for hybrid offshore farms,” in *2024 IEEE 22nd Mediterranean Electrotechnical Conference (MELECON)*, IEEE, 2024, pp. 390–395.

[60] E. Kazemi-Robati, S. Varotto, B. Silva, and I. Temiz, “Multiobjective energy management of multi-source offshore parks assisted with hybrid battery and hydrogen/fuel-cell energy storage systems,” *Applied Energy*, vol. 377, p. 124 529, 2025.

[61] A. Khan, M. Seyedmahmoudian, A. Raza, and A. Stojcevski, “Analytical review on common and state-of-the-art fr strategies for vsc-mtdc integrated offshore wind power plants,” *Renewable and Sustainable Energy Reviews*, vol. 148, p. 111 106, 2021.

[62] C.-H. Lin and Y.-K. Wu, “Overview of frequency-control technologies for a vsc-hvdc-integrated wind farm,” *IEEE Access*, vol. 9, pp. 112 893–112 921, 2021.

[63] C.-H. Lin and Y.-K. Wu, “Coordinated frequency control strategy for vsc-hvdc-connected wind farm and battery energy storage system,” *2022 IEEE Industry Applications Society Annual Meeting (IAS)*, pp. 1–10, 2022. [Online]. Available: https://api.semanticscholar.org/CorpusID:253629766.

[64] G. Ramtharan, N. Jenkins, and J. Ekanayake, “Frequency support from doubly fed induction generator wind turbines,” *IET renewable power generation*, vol. 1, no. 1, pp. 3–9, 2007.

[65] H. Ma and B. Chowdhury, “Working towards frequency regulation with wind plants: Combined control approaches,” *IET Renewable Power Generation*, vol. 4, no. 4, pp. 308–316, 2010.

[66] Z.-S. Zhang, Y.-Z. Sun, J. Lin, and G.-J. Li, “Coordinated frequency regulation by doubly fed induction generator-based wind power plants,” *IET Renewable Power Generation*, vol. 6, no. 1, pp. 38–47, 2012.

[67] K. V. Vidyanandan and N. Senroy, “Primary frequency regulation by deloaded wind turbines using variable droop,” *IEEE transactions on Power Systems*, vol. 28, no. 2, pp. 837–846, 2012.

[68] N. J. Abbas, J. Jasa, D. S. Zalkind, A. Wright, and L. Pao, “Control co-design of a floating offshore wind turbine,” *Applied Energy*, vol. 353, p. 122 036, 2024.

[69] C. A. M. Ströfer, D. T. Gaebele, R. G. Coe, and G. Bacelli, “Control co-design of power take-off systems for wave energy converters using wecopttool,” *IEEE Transactions on Sustainable Energy*, vol. 14, no. 4, pp. 2157–2167, 2023.

[70] K. Naik, S. Beknalkar, A. Mazzoleni, and C. Vermillion, “Fused geometric, structural, and control co-design framework for an energy-harvesting ocean kite,” in *2021 American Control Conference (ACC)*, IEEE, 2021, pp. 3525–3531.

[71] M. Garcia-Sanz, “Hybrid energy systems: Synergy margin and control co-design,” *Advanced Control for Applications: Engineering and Industrial Systems*, e238, 2024.

[72] R. Zhu, K. Das, P. E. Sørensen, and A. D. Hansen, "Enhancing profits of hybrid wind-battery plants in spot and balancing markets using data-driven two-level optimization," *International Journal of Electrical Power & Energy Systems*, vol. 159, p. 110 029, 2024.

[73] X. Gao, B. Knueven, J. D. Siirola, D. C. Miller, and A. W. Dowling, "Multiscale simulation of integrated energy system and electricity market interactions," *Applied Energy*, vol. 316, p. 119 017, 2022.

[74] A. W. Dowling, R. Kumar, and V. M. Zavala, "A multi-scale optimization framework for electricity market participation," *Applied Energy*, vol. 190, pp. 147–164, 2017.

[75] F. Sorourifar, V. M. Zavala, and A. W. Dowling, "Integrated multiscale design, market participation, and replacement strategies for battery energy storage systems," *IEEE Transactions on Sustainable Energy*, vol. 11, no. 1, pp. 84–92, 2018.

[76] T. Douville, M. Severy, S. Datta, *et al.*, "An offshore wind energy development strategy to maximize electrical system benefits in southern oregon and northern california," National Offshore Wind Research and Development Consortium & US Department of the Interior, Bureau of Ocean Energy Management, Tech. Rep. OCS Study BOEM 2023-067, 2023. [Online]. Available: https://www.boem.gov/sites/default/files/documents/BOEM-2023-067.pdf.

[77] A. Harris, I. Guerrero, and M. Severy, "Electricity market options for offshore wind," in *California North Coast Offshore Wind Studies*, M. Severy, Z. Alva, G. Chapman, *et al.*, Eds., Humboldt, CA: Schatz Energy Research Center, 2020.

[78] P. Spitsen and V. Sprenkle, "Energy storage cost and performance database," *Pacific Northwest National Laboratory https://www. pnnl. gov/ESGC-cost-performance*, 2020.

[79] J. Heeter, R. Vora, S. Mathur, P. Madrigal, S. K. Chatterjee, and R. Shah, "Wheeling and banking strategies for optimal renewable energy deployment. international experiences," National Renewable Energy Lab.(NREL), Golden, CO (United States), Tech. Rep., 2016.

[80] R. Fuchs, G. R. Zuckerman, P. Duffy, *et al.*, "The cost of offshore wind energy in the united states from 2025 to 2050," National Renewable Energy Laboratory (NREL), Golden, CO (United States), Tech. Rep., 2024.

[81] T. Zhao, A. Parisio, and J. V. Milanović, "Distributed control of battery energy storage systems for improved frequency regulation," *IEEE Transactions on Power Systems*, vol. 35, no. 5, pp. 3729–3738, 2020.

[82] H. Mahmood and J. Jiang, "Decentralized power management of multiple pv, battery, and droop units in an islanded microgrid," *IEEE Transactions on Smart Grid*, vol. 10, no. 2, pp. 1898–1906, 2017.

[83] "2022 annual report on market issues and performance," California Independent System Operator, Tech. Rep., 2022.

[84] M. Banzo and A. Ramos, "Stochastic optimization model for electric power system planning of offshore wind farms," *IEEE Transactions on Power Systems*, vol. 26, no. 3, pp. 1338–1348, 2010.

[85] R. Chen, H. Sun, Q. Guo, *et al.*, "Reducing generation uncertainty by integrating csp with wind power: An adaptive robust optimization-based analysis," *IEEE Transactions on Sustainable Energy*, vol. 6, no. 2, pp. 583–594, 2015.

[86] F. Alismail, P. Xiong, and C. Singh, “Optimal wind farm allocation in multi-area power systems using distributionally robust optimization approach,” *IEEE Transactions on Power Systems*, vol. 33, no. 1, pp. 536–544, 2017.

[87] K. G. Binmore, *Playing for real: a text on game theory*. Oxford university press, 2007.

[88] A. N. Sadigh, M. Mozafari, and B. Karimi, “Manufacturer–retailer supply chain coordination: A bi-level programming approach,” *Advances in Engineering Software*, vol. 45, no. 1, pp. 144–152, 2012.

[89] F. Bergonti, G. Nava, V. Wüest, *et al.*, “Co-design optimisation of morphing topology and control of winged drones,” in *2024 IEEE International Conference on Robotics and Automation (ICRA)*, IEEE, 2024, pp. 8679–8685.

[90] Y. Beck and M. Schmidt, “A gentle and incomplete introduction to bilevel optimization,” 2021.

[91] M. Cerulli, A. Oustry, C. d’Ambrosio, and L. Liberti, “Solving a class of bilevel programs with quadratic lower level,” 2021.

[92] S. Ghadimi and M. Wang, “Approximation methods for bilevel programming,” *arXiv preprint arXiv:1802.02246*, 2018.

[93] M. Solodov, “An explicit descent method for bilevel convex optimization,” *Journal of Convex Analysis*, vol. 14, no. 2, p. 227, 2007.

[94] S. Sharifi, N. Abolfazli, E. Y. Hamedani, and M. Fazlyab, *Safe gradient flow for bilevel optimization*, 2025. arXiv: 2501.16520 [math.OC]. [Online]. Available: https://arxiv.org/abs/2501.16520.

[95] N. Abolfazli, S. Sharifi, M. Fazlyab, and E. Y. Hamedani, *Perturbed gradient descent via convex quadratic approximation for nonconvex bilevel optimization*, 2025. arXiv: 2504.17215 [math.OC]. [Online]. Available: https://arxiv.org/abs/2504.17215.

[96] A. Sinha, P. Malo, and K. Deb, “Efficient evolutionary algorithm for single-objective bilevel optimization,” *arXiv preprint arXiv:1303.3901*, 2013.

[97] F. Bergonti, G. Nava, V. Wüest, *et al.*, “Co-design optimisation of morphing topology and control of winged drones,” *arXiv preprint arXiv:2309.13948*, 2023.

[98] X. Li, P. Tian, and X. Min, “A hierarchical particle swarm optimization for solving bilevel programming problems,” in *International Conference on Artificial Intelligence and Soft Computing*, Springer, 2006, pp. 1169–1178.

[99] A. Sinha, P. Malo, and K. Deb, “A review on bilevel optimization: From classical to evolutionary approaches and applications,” *IEEE transactions on evolutionary computation*, vol. 22, no. 2, pp. 276–295, 2017.

[100] Y. Bengio, A. Lodi, and A. Prouvost, “Machine learning for combinatorial optimization: A methodological tour d’horizon,” *European Journal of Operational Research*, 2020.

[101] Y. Ng, S. Misra, L. A. Roald, and S. Backhaus, “Statistical learning for dc optimal power flow,” in *2018 Power Systems Computation Conference (PSCC)*, IEEE, 2018, pp. 1–7.

[102] A. Agrawal, S. Barratt, and S. Boyd, “Learning convex optimization models,” *IEEE/CAA Journal of Automatica Sinica*, vol. 8, no. 8, pp. 1355–1364, 2021. doi: 10.1109/JAS.2021.1004075.

[103] B. Amos *et al.*, "Tutorial on amortized optimization," *Foundations and Trends® in Machine Learning*, vol. 16, no. 5, pp. 592–732, 2023.

[104] R. Sambharya, G. Hall, B. Amos, and B. Stellato, "Learning to warm-start fixed-point optimization algorithms," *Journal of Machine Learning Research*, vol. 25, no. 166, pp. 1–46, 2024. [Online]. Available: http://jmlr.org/papers/v25/23-1174.html.

[105] D. Bertsimas and B. Stellato, "Online mixed-integer optimization in milliseconds," *INFORMS Journal on Computing*, vol. 34, no. 4, pp. 2229–2248, 2022. doi: 10.1287/ijoc.2022.1181. [Online]. Available: https://doi.org/10.1287/ijoc.2022.1181.

[106] E. King, J. Kotary, F. Fioretto, and J. Drgona, "Metric learning to accelerate convergence of operator splitting methods for differentiable parametric programming," *arXiv preprint arXiv:2404.00882*, 2024.

[107] E. Khalil, P. Le Bodic, L. Song, G. Nemhauser, and B. Dilkina, "Learning to branch in mixed integer programming," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 30, 2016.

[108] Y. Tang, S. Agrawal, and Y. Faenza, "Reinforcement learning for integer programming: Learning to cut," in *International Conference on Machine Learning*, PMLR, 2020, pp. 9367–9376.

[109] J. Song, r. lanka ravi, Y. Yue, and B. Dilkina, "A general large neighborhood search framework for solving integer linear programs," in *Advances in Neural Information Processing Systems*, H. Larochelle, M. Ranzato, R. Hadsell, M. F. Balcan, and H. Lin, Eds., vol. 33, Curran Associates, Inc., 2020, pp. 20 012–20 023. [Online]. Available: https://proceedings.neurips.cc/paper/2020/file/e769e03a9d329b2e864b4bf4ff54ff39-Paper.pdf.

[110] J. Kotary, F. Fioretto, P. Van Hentenryck, and B. Wilder, "End-to-end constrained optimization learning: A survey," in *Proceedings of the Thirtieth International Joint Conference on Artificial Intelligence, IJCAI-21*, 2021, pp. 4475–4482. doi: 10.24963/ijcai.2021/610. [Online]. Available: https://doi.org/10.24963/ijcai.2021/610.

[111] B. Wilder, B. Dilkina, and M. Tambe, "Melding the data-decisions pipeline: Decision-focused learning for combinatorial optimization," in *Proceedings of the AAAI Conference on Artificial Intelligence (AAAI)*, vol. 33, 2019, pp. 1658–1665.

[112] T. Frerix, M. Nießner, and D. Cremers, "Homogeneous linear inequality constraints for neural network activations," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, 2020, pp. 748–749.

[113] A. V. Konstantinov and L. V. Utkin, "A new computationally simple approach for implementing neural networks with output hard constraints," in *Doklady Mathematics*, Springer, 2024, pp. 1–9.

[114] F. Fioretto, P. V. Hentenryck, T. W. Mak, C. Tran, F. Baldo, and M. Lombardi, "Lagrangian duality for constrained deep learning," in *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*, Springer, 2020, pp. 118–135.

[115] S. Park and P. Van Hentenryck, "Self-supervised primal-dual learning for constrained optimization," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 37, 2023, pp. 4052–4060.

[116] P. L. Donti, D. Rolnick, and J. Z. Kolter, "Dc3: A learning method for optimization with hard constraints," *arXiv preprint arXiv:2104.12225*, 2021.

[117] J. Dumouchelle, E. Julien, J. Kurtz, and E. B. Khalil, "Neur2bilo: Neural bilevel optimization," *arXiv preprint arXiv:2402.02552*, 2024.

[118] J. Shen, X. Chen, H. Heaton, *et al.*, "Learning a minimax optimizer: A pilot study," in *International Conference on Learning Representations*, 2020.

[119] M. Andrychowicz, M. Denil, S. G. Colmenarejo, *et al.*, "Learning to learn by gradient descent by gradient descent," in *Proceedings of the 30th International Conference on Neural Information Processing Systems*, ser. NIPS'16, Barcelona, Spain: Curran Associates Inc., 2016, pp. 3988–3996, isbn: 9781510838819.

[120] B. Amos and J. Z. Kolter, "Optnet: Differentiable optimization as a layer in neural networks," in *International Conference on Machine Learning*, PMLR, 2017, pp. 136–145.

[121] S. Gould, R. Hartley, and D. Campbell, "Deep declarative networks," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 44, no. 8, pp. 3988–4004, 2021.

[122] A. Agrawal, B. Amos, S. Barratt, S. Boyd, S. Diamond, and J. Z. Kolter, "Differentiable convex optimization layers," *Advances in neural information processing systems*, vol. 32, 2019.

[123] A. Agrawal, S. Barratt, S. Boyd, E. Busseti, and W. M. Moursi, "Differentiating through a cone program," *arXiv preprint arXiv:1904.09043*, 2019.

[124] R. Sambharya, G. Hall, B. Amos, and B. Stellato, "End-to-end learning to warm-start for real-time quadratic optimization," in *Learning for Dynamics and Control Conference*, PMLR, 2023, pp. 220–234.

[125] W. Chen, M. Tanneau, and P. Van Hentenryck, "End-to-end feasible optimization proxies for large-scale economic dispatch," *IEEE Transactions on Power Systems*, vol. 39, no. 2, pp. 4723–4734, 2023.

[126] J. Tordesillas, J. P. How, and M. Hutter, "Rayen: Imposition of hard convex constraints on neural networks," *arXiv preprint arXiv:2307.08336*, 2023.

[127] A. Beck, *First-order methods in optimization*. SIAM, 2017.

[128] A. Paszke, S. Gross, S. Chintala, *et al.*, "Automatic differentiation in pytorch," 2017.

[129] A. Paszke, S. Gross, S. Chintala, *et al.*, "Automatic differentiation in pytorch," in *NIPS-W*, 2017.

[130] E. Busseti, W. M. Moursi, and S. Boyd, "Solution refinement at regular points of conic problems," *Computational Optimization and Applications*, vol. 74, pp. 627–643, 2019.

[131] J. D. Lee, I. Panageas, G. Piliouras, M. Simchowitz, M. I. Jordan, and B. Recht, "First-order methods almost always avoid strict saddle points," *Mathematical programming*, vol. 176, pp. 311–337, 2019.

[132] P. Asaah, L. Hao, and J. Ji, "Optimal placement of wind turbines in wind farm layout using particle swarm optimization," *Journal of Modern Power Systems and Clean Energy*, vol. 9, no. 2, pp. 367–375, 2021.

[133] P. Hou, W. Hu, M. Soltani, and Z. Chen, "Optimized placement of wind turbines in large-scale offshore wind farm using particle swarm optimization algorithm," *IEEE Transactions on Sustainable Energy*, vol. 6, no. 4, pp. 1272–1282, 2015.

[134] J. Lofberg, "Yalmip: A toolbox for modeling and optimization in matlab," in *2004 IEEE international conference on robotics and automation (IEEE Cat. No. 04CH37508)*, IEEE, 2004, pp. 284–289.

[135] B. Amos, I. Jimenez, J. Sacks, B. Boots, and J. Z. Kolter, "Differentiable mpc for end-to-end planning and control," *Advances in neural information processing systems*, vol. 31, 2018.

[136] J. Drgona, A. Tuor, J. Koch, M. Shapiro, and D. Vrabie, "NeuroMANCER: Neural Modules with Adaptive Nonlinear Constraints and Efficient Regularizations," 2023. [Online]. Available: https://github.com/pnnl/neuromancer.

[137] M. S. Innocente and J. Sienz, "Constraint-handling techniques for particle swarm optimization algorithms," *arXiv preprint arXiv:2101.10933*, 2021.

[138] L. J. Miranda, "Pyswarms: A research toolkit for particle swarm optimization in python," *Journal of Open Source Software*, vol. 3, no. 21, p. 433, 2018.

[139] Z. Zhang, "Improved adam optimizer for deep neural networks," in *2018 IEEE/ACM 26th international symposium on quality of service (IWQoS)*, Ieee, 2018, pp. 1–2.

[140] D. Wu, G.-S. Seo, L. Xu, C. Su, Ł. Kocewiak, Y. Sun, *et al.*, "Grid integration of offshore wind power: Standards, control, power quality and transmission," *IEEE Open J. of Power Electron.*, 2024.

[141] H. Sharma, W. Wang, B. Huang, T. Ramachandran, and V. Adetola, "Multi-objective control co-design using graph based optimization for offshore wind farm grid integration," in *2024 American Control Conference (ACC)*, 2024, pp. 2580–2585. doi: 10.23919/ACC60939.2024.10644390.

[142] J. Zoellick, G. Adams, A. Mustafa, A. Cooperman, R. Anilkumar, P. Duffy, *et al.*, "Northern california and southern oregon offshore wind transmission study," National Renewable Energy Laboratory (NREL), Golden, CO (United States), Tech. Rep., 2023.

[143] S. Panda, S. Mohanty, P. Rout, B. K. Sahu, S. M. Parida, I. S. Samanta, *et al.*, "A comprehensive review on demand side management and market design for renewable energy support and integration," *Energy Reports*, 2023. doi: 10.1016/J.EGYR.2023.09.049.

[144] C. Brunetto and G. Tina, "Optimal hydrogen storage sizing for wind power plants in day ahead electricity market," *IET Renewable Power Generation*, vol. 1, no. 4, pp. 220–226, 2007.

[145] A. W. Dowling, R. Kumar, and V. M. Zavala, "A multi-scale optimization framework for electricity market participation," *Appl. Energy*, 2017. doi: 10.1016/J.APENERGY.2016.12.081.

[146] X. Chen, X. Gao, D. Guittet, *et al.*, "Beyond price-taker: Multiscale optimization of wind and battery integrated energy systems," *arXiv preprint arXiv:2409.08343*, 2024.

[147] Ã. Parede, J. A. Aguado, C. Essayeh, Y. Xia, I. Savelli, and T. Morstyn, "Stacking revenues from flexible ders in multi-scale markets using tri-level optimization," *IEEE Trans. on Power Syst.*, vol. 39, pp. 3949–3961, 2024.

[148] C. Chen, S. Li, and L. Tong, "Multi-interval energy-reserve co-optimization with soc-dependent bids from battery storage," *arXiv preprint arXiv:2401.15525*, 2024.

[149] R. K. Bansal, E. Mallada, and P. Hidalgo-Gonzalez, "A market mechanism for a two-stage settlement electricity market with energy storage," 2024. [Online]. Available: https://api.semanticscholar.org/CorpusID:268264783.

[150] Q. Ma, W. Wei, and S. Mei, "A two-timescale operation strategy for battery storage in joint frequency and energy markets," *IEEE Trans. on Energy Markets, Policy and Regulation*, vol. 2, pp. 200–213, 2024.

[151] Z. Li, Y. Xia, Y. Bo, and W. Wei, "Optimal planning for electricity-hydrogen integrated energy system considering multiple timescale operations and representative time-period selection," *Appl. Energy*, vol. 362, p. 122 965, 2024.

[152] J. R. Birge and F. Louveaux, *Introduction to stochastic programming*. Springer Science & Business Media, 2011.

[153] W. B. Powell, "A unified framework for stochastic optimization," *European J. of Oper. Res.*, vol. 275, no. 3, pp. 795–821, 2019.

[154] F.-J. Heredia, M. D. Cuadrado, and C. Corchero, "On optimal participation in the electricity markets of wind power plants with battery energy storage systems," *Comput. Oper. Res.*, vol. 96, pp. 316–329, 2018.

[155] Y. Wang, H. Zhao, and P. Li, "Optimal offering and operating strategies for wind-storage system participating in spot electricity markets with progressive stochastic-robust hybrid optimization model series," *Math. Problems in Eng.*, vol. 2019, no. 1, p. 2 142 050, 2019.

[156] M. A. Mohamed, T. Jin, and W. Su, "An effective stochastic framework for smart coordinated operation of wind park and energy storage unit," *Appl. Energy*, vol. 272, p. 115 228, 2020.

[157] J. T. Allison and D. R. Herber, "Special section on multidisciplinary design optimization: Multidisciplinary design optimization of dynamic engineering systems," *AIAA journal*, vol. 52, no. 4, pp. 691–710, 2014.

[158] M. Lubin, O. Dowson, J. Dias Garcia, J. Huchette, B. Legat, and J. P. Vielma, "JuMP 1.0: Recent improvements to a modeling language for mathematical optimization," *Mathematical Programming Computation*, 2023. doi: 10.1007/s12532-023-00239-3.

[159] O. Kröger, C. Coffrin, H. Hijazi, and H. Nagarajan, "Juniper: An open-source nonlinear branch-and-bound solver in julia," in *Integration of Constraint Programming, Artificial Intelligence, and Operations Research*, Springer International Publishing, 2018, pp. 377–386, isbn: 978-3-319-93031-2.

[160] K. Kirui, A. Pichler, and G. C. Pflug, "Scentrees.jl: A Julia package for generating scenario trees and scenario lattices for multistage stochastic programming," *J. of Open Source Softw.*, vol. 5, no. 46, p. 1912, 2020. doi: 10.21105/joss.01912.

[161] A. Y. Nakhai, "Electrical infrastructure cost model for marine energy systems," National Renewable Energy Laboratory (NREL), Golden, CO (United States), Tech. Rep., 2023.

[162] R. Meyur, T. Martin, and S. Purohit, "Cameo: A co-design architecture for multi-objective energy system optimization," *arXiv preprint arXiv:2408.11794*, 2024.

[163] N. W. Miller, W. W. Price, and J. J. Sanchez-Gasca, “Dynamic modeling of ge 1.5 and 3.6 wind turbine-generators,” *GE-Power systems energy consulting*, no. 3.0, 2003.

[164] N. Bodini, M. Optis, Y. Liu, *et al.*, “Causes of and solutions to wind speed bias in nrel’s 2020 offshore wind resource assessment for the california pacific outer continental shelf,” National Renewable Energy Laboratory (NREL), Golden, CO (United States), Tech. Rep., 2024.

Appendix A – Offshore Wind Farm Turbine Technology

We use the GE 1.5 MW wind turbine as a base model [163] to generate sampling points for speed-power fitting. A detailed electromagnetic transient (EMT) model was developed in PSCAD, and simulations were conducted by incrementally increasing the wind speed from the cut-in to the cut-out wind speed. These points are then fitted with a piece-wise quadratic function as presented in Figure B.1(d). The overall fitted curve is aligned with the original GE 1.5 MW wind turbine [163]. We upscale the fitted power curve to the rated power of each OWF to convert wind speed to power output in the control co-design process.

Appendix B – Wind Speed Data

The wind data utilized in the study includes historical wind speed data available from various sources, particularly focusing on offshore wind areas. The dataset includes wind speed measurements at different heights (e.g., 140m) and covers a range of years, notably from 2000 to 2020. The data is essential for generating stochastic wind scenarios and validating wind speed distributions using methods such as the Weibull distribution. Key sources for this wind data include the NREL WIND Toolkit Offshore Summary Dataset, which offers data for specific years such as 2017 and 2020. It is important to note that this dataset exhibits bias relative to floating LiDAR observations, attributable to atmospheric stability effects; see Bodini et.al [164] for a comprehensive analysis . In our study, the dataset is used solely to stochastically generate wind-speed scenarios, so this bias is expected to have limited influence on the reported results. In future work, we will employ additional wind-resource datasets to quantify the sensitivity of co-design solutions to wind-speed data.

The Offshore California Dataset from NREL WIND Toolkit is a 21-year wind resource dataset for offshore California. Produced in 2020, this data set replaces NREL's Wind Integration National Dataset (WIND) Toolkit for offshore California, which was produced and released publicly in 2013 and is currently the principal data set used by stakeholders for wind resource assessment in the continental United States. Both the WIND Toolkit and this new data set are created using the Weather Research and Forecasting (WRF) numerical weather prediction model (NWP)

The Figure B.1(a-b) illustrate wind speed density distributions at various locations (WCASCADE, JohnDay, COTWDPGE, Tesla, and Mossland) for 2018 and 2022. Both years show multi-modal distributions, with peaks around 4-10 m/s and tailing off after 30 m/s. In 2022, the density appears slightly higher around 8-12 m/s across most locations compared to 2018, which might indicate stronger winds or more frequent occurrences of mid-range speeds. This shift can significantly influence wind power generation since wind turbines typically produce maximum power output in this speed range.

Comparing the wind speed distributions to the turbine power curve in Figure B.1(d), the observed wind speeds in 2018 and 2022 indicate that a significant portion of the wind speeds in both years fall within the optimal power generation region (8-12 m/s). The slight increase in frequency in this range during 2022 suggests potentially higher energy output, assuming the wind farms are designed for maximum efficiency in this range.

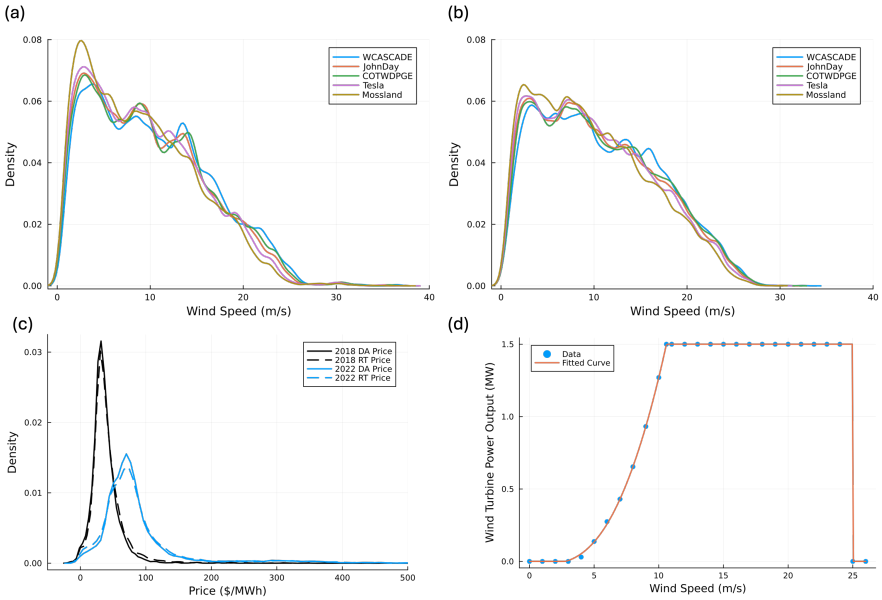


Figure B.1: Distribution of wind-speed for different wind farms locations in (a) 2018 (b) 2022 (c) Day ahead (DA) and Real time (RT) price (2018,2022) (d) Wind turbine power characteristic

Appendix C – Energy Market Data

The energy market data primarily comes from the California Independent System Operator (CAISO). This data is accessed through two sources, the CAISO OASIS API or EIA portal.

CAISO OASIS API includes various market-related metrics such as Locational Marginal Prices (LMPs) for both day-ahead and real-time markets, Ancillary Service Prices (ASP), and system performance data like Regulation Mileage. The CAISO data covers the periods from 2018 onwards, with some limitations in earlier years where LMP data may not be available. This data is crucial for analyzing the market participation of offshore wind farms and optimizing their dispatch in energy markets.

The Wholesale Electricity Market Portal, launched by the U.S. Energy Information Administration (EIA) in March 2024, provides access to electricity market data from seven Regional Transmission Organizations (RTOs) and Independent System Operators (ISOs), including CAISO. The Portal offers datasets such as day-ahead and real-time locational marginal prices (LMPs), load/demand data, generation fuel mix, and city temperatures, gathered from RTO/ISO public data and NOAA. Data availability varies, with updates ranging from hourly to longer intervals. Data can be accessed through dashboards and bulk download flat files via CAISO EIA Portal, though API access is not currently available.

Appendix D – Scenario Generation Algorithm

The scenario generation algorithm is presented in Algorithm-3.

Algorithm 3 Kernel-Based Scenario Tree Generation Algorithm

Input: Time series data for LMP, ASP, Windspeed, etc.
Output: Scenario tree structure
Parameters: SCN1 $= 20$ (second-stage scenarios), SCN2 $= 5$ (third-stage scenarios)
Step 1: Data Preprocessing Load datasets for LMP, ASP, and Windspeed Reshape data into a 3D array (N, T, K), where N is the number of samples, T is time steps, and K is the number of variables
Step 2: Scenario Generation Using Kernel Density Estimation (KDE) **for** *each time step* $t = 1$ **to** T **do**
  Normalize weights $w \leftarrow w / \sum w$ Compute effective sample size N_t and standard deviation σ_t Generate scenario $x[t]$ using KDE from multivariate normal distribution Update weights based on Markovian or Non-Markovian process
end
Return scenario trajectory x
Step 3: Tree Construction Define tree structure as $[1, \text{SCN1}, \text{SCN2}, 1, 1, 1]$ Apply kernel-based sampling to generate $100{,}000$ scenarios Construct the scenario tree using tree approximation
Step 4: Save the Scenario Tree

Appendix E – PSCAD Model and Test Scenarios

The PSCAD model used for validation is modified from the base 240-bus miniWECC system developed in [35]. Grid-following inverter-based resources with dynamic models of REGCA1, REECB1, and REPCA1 are integrated into the system. The REGCA1 model is adjusted so that the *d*- and *q*-axis currents are transformed into a three-phase current using a phase-locked loop. Additionally, frequency and voltage droop deadbands are set at ±0.017 Hz and ±0.01 p.u., respectively.

Five offshore wind farms (OWFs) are connected to WCASCADE, JOHN DAY, COTTONWOOD, TESLA, and MOSSLAND, with rated capacities of 1500 MW, 2350 MW, 1810 MW, 2640 MW, and 1800 MW, respectively. These type 3 turbines include a pitch angle control module, a grid-side control module, a rotor-side control module, and a DC-link chopper, allowing them to switch flexibly between MPPT and de-loading modes. They are assigned wind speeds of 10.68 m/s, 10.68 m/s, 8.86 m/s, 8.86 m/s, and 6.75 m/s, representing typical wind conditions.

Pacific Northwest National Laboratory

902 Battelle Boulevard
P.O. Box 999
Richland, WA 99354
1-888-375-PNNL (7665)

www.pnnl.gov