

ORIGINAL RESEARCH

Open access

High-Performance Computing for Scalable Virtual Power Plant Optimization: Parallel Simulated Annealing and MILP Integration

Amira Khalil^{1*}, Sara Fathy¹, Mahmoud Adel², Hany Nabil¹, Rania Mostafa²

Abstract

This research addresses the optimization of scheduling strategies for virtual power plants (VPPs), as demonstrated in Portugal's national New Generation Storage (NGS) initiative, aiming to maximize societal benefits and boost the effectiveness of energy transactions in today's advanced power networks. VPPs play a crucial role in sustainable urban energy ecosystems by combining various distributed energy resources (DERs), such as renewable generation units and battery storage installations. The underlying optimization challenge is expressed as a Mixed-Integer Linear Programming (MILP) model and tackled via a parallelized simulated annealing (SA) heuristic running on high-performance computing (HPC) platforms. The parallel design greatly enhances the speed of searching the solution landscape, thereby allowing the framework to cope with the growing intricacy associated with expanded DER fleets and increasingly complex dispatch requirements. The developed technique aligns closely with smart city aspirations by enabling flexible, high-efficiency energy management, supporting real-time pricing schemes, and extending the durability of essential components, such as energy storage units. Comprehensive simulation experiments reveal the method's strong performance in reducing computation time, maintaining high-quality solutions, and scaling effectively to larger problem sizes, thereby supporting instantaneous operational decisions in competitive energy markets. In addition, the improved synchronization among DERs strengthens overall grid reliability, increases market agility, and fosters the emergence of robust, decarbonized urban energy landscapes. This investigation emphasizes the decisive impact of powerful computing resources in overcoming the complexities of contemporary energy infrastructures, illustrating the potential of cutting-edge algorithmic approaches and HPC systems to deliver scalable, responsive, and environmentally sound energy management solutions tailored to smart urban settings. Ultimately, the outcomes pave the way for energy systems that are both socially inclusive and ecologically sustainable, fully consistent with the essential aims of urban resilience and long-term sustainable progress.

Keywords Virtual power plant, Social welfare optimization, Battery lifespan constraints, Parallel simulated annealing, High-performance computing, Mixed-integer linear programming

*Correspondence:

Amira Khalil
amira.khalil@alexu.edu.eg

¹ Department of Smart Energy Systems, Faculty of Engineering, Alexandria University, Alexandria, Egypt

² Department of Power Grid Optimization, Faculty of Engineering, Cairo University, Cairo, Egypt

Introduction

The growing integration of distributed energy resources (DERs), including renewable energy sources (RESs) such as wind and solar power, offers significant opportunities alongside considerable management challenges for contemporary electricity grids. Virtual power plants (VPPs) have emerged as an effective mechanism for aggregating and orchestrating DERs, thereby improving grid reliability, economic performance, and active participation in electricity markets [1, 2].

VPPs function as unified aggregated units, enabling smooth connectivity with the power grid. They manage a variety of energy assets, including RESs, electric vehicles (EVs), and energy storage systems (ESSs) [3, 4]. Within smart city frameworks, VPPs operate as adaptable enablers of distributed energy governance by consolidating DERs to match generation with consumption, lower greenhouse gas emissions, and increase overall energy utilization efficiency [5]. By incorporating technologies such as Vehicle-to-Grid (V2G), VPPs can deliver grid-supporting services, strengthen system robustness, and support intelligent energy infrastructure designed for city settings [6]. These setups further empower city developers to refine energy distribution patterns by leveraging live data from connected infrastructure elements, such as intelligent public lighting and EV charging points, thereby advancing eco-friendly urban growth [7].

Nevertheless, determining the optimal scheduling and control of VPPs poses major hurdles due to the inherent variability in RES output, electricity consumption patterns, and market price fluctuations. Overcoming these issues requires advanced optimization approaches that harmonize supply–demand interactions while maximizing both financial returns and societal benefits for participants in VPPs [8, 9]. Serving as a key pillar of smart urban strategies, VPPs are required to address the complexities of integrating heterogeneous energy assets while preserving operational flexibility and financial sustainability. The inclusion of EVs as portable DERs introduces additional layers of complexity, as their charging and vehicle-to-grid operations affect not only network balance but also the broader city energy landscape [4, 10]. Recent investigations have tackled these issues through the application of multi-agent frameworks and flexible risk-handling techniques to bolster the durability and security of decentralized energy networks [11]; the use of MILP-driven methods for improving energy exchange, demand scheduling, and asset distribution [12]; and the evaluation of freely available optimization software to advance storage management practices and expand operational capacity in VPPs [13].

Conventional optimization approaches, including MILP, often face limitations when handling the scalability and complexity required by extensive VPP scheduling tasks, particularly when time-sensitive constraints are involved [14, 15]. As VPPs expand in both the number of connected DERs and the complexity of trading frameworks, strong demand exists for reliable optimization solutions that can effectively meet these requirements [16]. As a result, heuristic and metaheuristic techniques—especially simulated annealing (SA)—have attracted considerable interest due to their adaptability and success in exploring intricate search domains commonly found in VPP dispatch problems [17, 18].

Simulated annealing (SA), a probabilistic optimization technique inspired by metallurgical annealing, has found broad use in tackling nonlinear, high-dimensional, and highly constrained optimization problems [16, 19]. Within VPP applications, SA demonstrates utility for coordinating energy allocation, regulating the interplay between energy procurement and storage utilization, and maximizing involvement in energy trading platforms [16, 20]. Contemporary studies indicate that SA delivers strong performance in VPP scheduling scenarios, especially when combined with demand response (DR) initiatives and diverse DER types, such as renewables and storage facilities [9, 18]. Even so, conventional SA implementations may encounter difficulties scaling up in extensive VPP setups that encompass a large number of DERs and elaborate market rules [15, 21].

To overcome these scaling limitations, investigators have adopted high-performance computing (HPC) to support parallelized SA implementations. Advances in parallel processing hardware and software environments have revealed substantial promise for speeding up optimization workflows and managing the vast datasets essential for instantaneous VPP scheduling [15, 22]. HPC-supported parallel SA strategies decompose expansive optimization challenges into smaller, tractable segments, thereby boosting processing performance without compromising solution integrity [23]. By leveraging multi-threading, these parallel SA methods shorten convergence time and facilitate oversight of large VPP configurations, making it practical to optimize energy procurement and delivery with fine temporal granularity [24]. The contribution of HPC to real-time VPP optimization goes well beyond mere increases in processing velocity. Research on parallelized SA illustrates that HPC platforms adeptly handle sophisticated scheduling demands by analyzing granular, time-critical information instantaneously [21, 25]. This functionality is particularly important for VPP managers, who need to adapt promptly to variations in renewable production, consumption levels, and market dynamics [26]. Additionally, parallel SA approaches permit the concurrent assessment of bidding tactics for VPPs, thereby simplifying decision-making in localized energy trading environments [23].

This paper presents an innovative parallelized simulated annealing (SA) algorithm for scheduling virtual power plants (VPPs) that is executed on high-performance computing (HPC) platforms using the OpenMP framework. The proposed model prioritizes maximizing social welfare while accounting for key operational constraints, including battery degradation and the inherent variability of distributed energy resources (DERs). By harnessing parallel processing, this method aims to deliver near-optimal scheduling outcomes while simultaneously boosting computational efficiency and the overall quality of solutions. Extensive simulation studies validate the proposed technique's ability to significantly shorten optimization durations without compromising solution excellence across a diverse array of VPP setups [16, 18].

Related Works

The emergence of virtual power plants (VPPs) has become indispensable for contemporary electricity networks, owing to their ability to aggregate distributed energy resources (DERs), including renewable energy sources (RESs), electric vehicles (EVs), and energy storage systems (ESSs). Robust optimization structures for VPPs are vital to ensure these assets enhance grid reliability, operational efficiency, and participation in energy trading platforms. This section surveys recent techniques and progress in VPP optimization, focusing on strategies for efficient scheduling, energy oversight, and alignment with market dynamics.

Optimization strategies for virtual power plants

A broad spectrum of optimization approaches has been formulated for VPPs, aimed at merging DERs while preserving operational productivity and financial returns. These include heuristic and metaheuristic procedures, game-theoretic frameworks, and models that address multiple objectives. The analysis by Podder *et al.* [1] categorizes 65 distinct VPP optimization methodologies, underscoring those that incorporate RESs alongside energy storage. Their findings indicate that pairing wind energy with demand response mechanisms effectively counters variability issues, thereby rendering VPP scheduling more robust. Moreover, VPPs are recognized as a fundamental approach for embedding distributed energy resources within smart urban power infrastructures. In-depth evaluations by Xie *et al.* [5] emphasize the application of layered control methods and protected communication protocols to effectively govern DERs and storage installations. Such methods support demand-side frequency regulation ancillary services (D-FCASs), bolstering network stability and advancing eco-conscious energy systems [7]. Analogously, examinations of microgrid, smart grid, and VPP innovations by Khan *et al.* [27] emphasize the importance of advanced control systems and engineering solutions for achieving energy sustainability. The fusion of these innovations proves essential for harmonizing functional performance with ecological targets in intelligent urban contexts [28]. Through sophisticated market instruments, VPPs enable participation in electricity markets while balancing power generation and consumption amid fluctuating conditions in the smart grid [10].

Game theory plays a prominent role in VPP optimization by enabling the management of interactions among numerous prosumers within a VPP, facilitating adaptive energy allocation that synchronizes production and consumption with prevailing price signals. For instance, the investigation by Zhang *et al.* [8] illustrates a game-theoretic framework that markedly improves VPP functionality by involving prosumers, yielding equitable benefit allocation for both residential and commercial participants. These methodologies highlight the adaptability VPPs need to accommodate market volatility and DER unpredictability.

Frameworks for multi-level and multi-objective optimization also play a prominent role in VPP scheduling, enabling VPPs to reconcile conflicting objectives such as cost reduction, emission abatement, and user satisfaction. Ali *et al.* [4] introduce a multi-tier, multi-criterion optimization structure that employs cutting-edge algorithms to accomplish proficient load allocation, cost savings in energy, and mitigation of greenhouse gases in VPP operations. This technique, using methods such as AMO-GWO, demonstrates the effectiveness of advanced optimization algorithms in accommodating diverse time-varying constraints while improving VPP performance.

Investigators have developed hierarchical, adaptive optimization structures to meet the demand for instantaneous flexibility in VPPs. Krishna *et al.* [29] examine an energy governance strategy based on the teaching-and-learning-based optimization (TLBO) algorithm and demonstrate its effectiveness in reducing VPP operational expenses. The TLBO-driven model excels at optimizing over a 24-hour timeframe, making it particularly advantageous for scenarios that require short-term energy oversight and rapid decision-making. In a similar vein, Muttaqi *et al.* [30] present an adaptive, forecasting-oriented energy management strategy that integrates model predictive control (MPC) with corrective feedback loops to refine real-time VPP operations, thereby improving financial performance and network reliability by effectively addressing uncertainties.

Robust optimization frameworks are frequently employed in VPP scheduling to address uncertainties arising from renewable energy sources (RESs) and volatile market dynamics. For example, Naughton *et al.* [31] present a robust scheduling framework for VPPs that employs a receding-horizon dispatch strategy. This three-stage optimization process seeks to maximize revenue while effectively addressing uncertainties in both system operations and market conditions. The proposed approach was tested through a practical real-world case study, which confirmed its success in enhancing VPP flexibility and revenue generation under unpredictable operating scenarios.

In addition, multi-scale and multi-time-scale methodologies allow VPPs to align their operations across different temporal resolutions. Yi *et al.* [32] introduce a multi-time-scale scheduling technique that aggregates and disaggregates deferrable loads within a VPP. This enables stronger involvement in both energy and ancillary reserve markets. By tackling the complexities of scheduling deferrable loads, the method achieves greater economic performance while ensuring reliable operation across VPP systems.

Moreover, novel bidding strategies continue to emerge to increase VPP profitability in competitive electricity markets. Tang *et al.* [33], for instance, develop a coordinated bidding and operational model for VPPs, with particular emphasis on optimizing the combined use of energy storage and demand response resources. By incorporating robust optimization principles, their model effectively handles uncertain energy prices and RES intermittency, ultimately delivering higher profits in both day-ahead and real-time market environments.

Taken together, these investigations illustrate the wide range of approaches available to optimize VPPs, each making distinct contributions toward overcoming the challenges of DER management and improving overall VPP performance. Ongoing progress in computational methods and algorithmic design continues to advance the VPP concept, delivering practical solutions for integrating renewable generation and balancing demand in contemporary power systems.

Simulated annealing (SA) is a widely adopted metaheuristic in energy resource management for VPPs, owing to its effectiveness in solving complex, high-dimensional optimization problems. Drawing inspiration from the metallurgical annealing process, SA probabilistically accepts inferior solutions during the initial phases of search. This mechanism helps the algorithm avoid local optima and promotes a more thorough exploration of the global solution landscape [9]. Such behavior makes SA especially appropriate for the non-linear, highly constrained nature of VPP scheduling tasks, particularly in environments marked by substantial uncertainty arising from DER variability and market price fluctuations.

Although SA provides considerable benefits for VPP scheduling optimization, it is worth considering a broader range of metaheuristic algorithms to further improve results in this area. The optimization of VPPs has become a highly active research domain, driven by the growing reliance on renewable energy and the need for effective coordination of decentralized resources. Recent work has explored various metaheuristic approaches—including Ant Colony Optimization (ACO) [34], Genetic Algorithms (GAs) [35], and the Whale Optimization Algorithm (WOA) [36]—to design optimal bidding strategies that aim to maximize earnings in local energy markets (LEMs). These techniques adeptly manage the intricacies of competitive power trading, emphasizing the importance of strategic adaptation and agent-based interactions for better market performance [23]. Angulo *et al.* [23] illustrated the success of such metaheuristics in coordinating diverse energy assets within compact microgrid settings. Similarly, Rädle *et al.* [37] introduced an adaptable control strategy for hierarchical VPPs that employs multiple computational intelligence (CI) metaheuristics, such as SA

and ACO, with objectives of reducing operating reserves and emissions while increasing the operational flexibility of power plants. The integration of hierarchical architectures with CI-based methods exemplifies a continuing effort to refine computational tools for dependable and efficient VPP management, while accommodating the distinct features and limitations of various VPP configurations.

In recent years, SA has been widely used in VPP-related research to address multiple issues in energy dispatch, demand response participation, and grid reliability. One notable application showed that combining SA with a tailored heuristic significantly improved the management of electric vehicles (EVs) functioning as distributed energy resources, enabling optimized charging and discharging patterns that bolster grid stability and increase financial returns for VPP operators [16]. Another SA implementation focused on energy management systems incorporating demand response mechanisms, highlighting the algorithm's ability to accommodate demand variations and substantially lower energy expenses. Simulation results showed that the SA-driven method outperformed particle swarm optimization and several deterministic techniques, underscoring its robustness and flexibility across different network scenarios [17].

Furthermore, investigators have employed simulated annealing (SA) to refine bidding strategies within electricity markets, with research showing that SA-driven algorithms can boost the financial performance of virtual power plants (VPPs) through carefully timed energy trades [18]. By adapting bidding decisions to live market information, SA-based optimization methods have delivered gains in processing speed and revenue generation, reinforcing the importance of SA in dynamic, competitive energy trading environments.

Beyond market engagement, SA has demonstrated strong performance in solving the dynamic economic dispatch challenge for VPPs, particularly when incorporating intermittent renewables such as wind power. By effectively managing the variability of renewable generation, SA-based techniques help lower overall production costs and increase societal benefits, illustrating their value for environmentally responsible energy coordination [19]. Implementations of SA within economic dispatch models have consistently reported reduced generation expenses alongside greater integration of renewable sources, confirming the algorithm's flexibility in adapting to shifts in the energy sector and advancing long-term sustainability objectives.

Overall, integrating SA into energy resource coordination for VPPs underscores its broad applicability and reliability across diverse operational contexts. Its capacity to respond to instantaneous changes and its resilience to uncertainty make SA an essential instrument for achieving efficient, robust, and economically viable VPP performance.

Parallel computing for VPP optimization

The substantial computational demands of VPP optimization have spurred growing interest in high-performance computing (HPC) and parallel processing techniques as practical means to scale up optimization tasks. Conventional optimization approaches frequently prove inadequate when faced with the real-time requirements and expansive scope of VPPs, particularly as these systems grow in terms of connected distributed energy resources (DERs), market involvement, and functional complexity [21]. Parallel computing, when supported by HPC platforms, has shown considerable promise for speeding up optimization workflows and improving the quality of obtained solutions, especially for VPP scenarios that rely on frequent updates driven by live data streams.

Numerous investigations have examined parallel computing as a strategy for tackling the high-dimensional optimization challenges inherent to VPPs. For example, Alvarez *et al.* [24] studied the use of parallel processing in security-constrained optimal power flow (SCOPF) analysis, demonstrating that parallelized implementations could substantially reduce computation time while preserving solution accuracy. Such techniques are relevant to VPP applications, where ensuring security and contingency readiness is critical for dependable grid interconnection.

Studies have further highlighted the role of parallel computing in stochastic and scenario-driven optimization frameworks for addressing uncertainty in VPP operations. A scalable stochastic optimization structure for intricate energy systems,

outlined in Lubin *et al.* [15], utilized parallel interior-point solvers to address extensive dispatch problems that included wind generation, yielding notable advances in scalability and computational performance. This structure highlights the advantages of HPC-supported optimization for VPP scheduling, particularly where the unpredictability of DERs and market conditions limits the effectiveness of deterministic approaches.

Parallel computing also offers valuable potential for demand response programs and market engagement in VPPs. Using exhaustive parallel search techniques, Angulo *et al.* [23] analyzed bidding approaches for local energy markets and found that metaheuristic strategies could deliver near-optimal outcomes within practical time limits. These results indicate that parallel processing can strengthen VPPs' competitiveness in energy markets, enabling more attractive pricing structures and improved resource distribution.

High-performance computing environments, such as the PIPS-IPM solver described in Petra *et al.* [22], have enabled real-time stochastic optimization of power networks with large numbers of variables, thereby confirming the effectiveness of parallel architectures for computationally intensive VPP scheduling problems. In this context, HPC not only accelerates decision-making but also enhances grid reliability by enabling VPPs to adapt swiftly to fluctuations in demand, renewable output, and market pricing signals.

Parallel computing techniques, when integrated with adaptive optimization methods such as simulated annealing (SA), offer additional opportunities to scale VPP optimization efforts. Hybrid frameworks that combine SA with parallel execution have proven successful in addressing the high-dimensional, heavily constrained optimization demands typical of VPPs, as illustrated in Sousa *et al.* [9]. This combined approach supports distributed computation, in which individual nodes within an HPC cluster optimize distinct portions of the VPP's distributed energy resources (DERs). As a result, the process converges more quickly and produces solutions of superior quality compared to traditional sequential methods.

Parallel computing is a groundbreaking strategy for optimizing VPPs, particularly as DER integration continues to grow and the need for instantaneous market responsiveness intensifies. Leveraging multi-threading, distributed processing, and sophisticated parallel algorithms enables VPP operators to effectively coordinate and control extensive DER networks, thereby fostering a more adaptable and agile electricity grid.

Role of High-Performance Computing (HPC) in advanced VPP scheduling

Incorporating high-performance computing (HPC) into virtual power plant (VPP) scheduling represents a major advancement in addressing the challenges posed by expansive energy infrastructure. As distributed energy resources (DERs) become more prevalent and the need for flexible, responsive power systems grows, HPC delivers the processing power needed to support real-time scheduling and oversight of VPPs. Investigations have confirmed that HPC platforms can dramatically improve the scalability of optimization algorithms for energy systems, facilitating the decomposition of comprehensive VPP scheduling problems into smaller, more tractable components [21, 22]. Such scalability is vital to preserving solution accuracy, as VPPs must manage a mix of variable energy sources—including wind, solar, and electric vehicles (EVs)—while complying with market-imposed restrictions.

Parallel computing environments, especially those using HPC clusters, have been deployed for VPP scheduling to accelerate solution times and accommodate increasing complexity. Research indicates that parallelized scheduling procedures markedly shorten computation durations, empowering VPP operators to perform timely adjustments in response to fluctuations in load demand, renewable output predictions, and requirements for grid reliability [24]. These parallel structures support the concurrent optimization of several goals, including cost savings, system dependability, and reduction of carbon emissions—objectives central to contemporary VPP management [15].

Sophisticated stochastic optimization approaches enabled by HPC, originally developed for real-time power network optimization, have been successfully adapted to VPP applications. For example, the PIPS-IPM solver, initially developed for intricate grid systems, has demonstrated effectiveness in solving stochastic dispatch challenges by managing large

datasets and executing computationally intensive operations in real time [22]. This progress highlights HPC's contribution to elevating the accuracy of VPP scheduling, as powerful solvers enable the incorporation of uncertainties in load and generation forecasts, thereby strengthening the resilience of VPP control strategies [23].

Additionally, HPC enables more advanced simulation and optimization, which are especially useful for developing refined bidding strategies in energy markets. By applying exhaustive search procedures and metaheuristic optimization techniques within HPC environments, VPPs can evaluate multiple market situations and identify optimal bidding approaches that maximize earnings amid market unpredictability [9]. Such HPC-supported techniques help ensure that VPPs remain competitive and responsive, adapting to shifting market dynamics while accommodating increased integration of renewable energy [15].

High-performance computing (HPC) has become a vital facilitator in overcoming the escalating complexity and scalability issues encountered in virtual power plant (VPP) optimization. Although prior research has thoroughly examined heuristic and metaheuristic techniques—including genetic algorithms (GA), particle swarm optimization (PSO), and simulated annealing (SA)—for energy scheduling, these methods often face scalability and real-time feasibility limitations when applied to large-scale networks. The present study sets itself apart by using HPC to develop a parallelized variant of the SA algorithm specifically tailored for the VPP scheduling task. By leveraging OpenMP-supported parallelism, the proposed method effectively circumvents the computational limitations commonly associated with SA in high-dimensional, highly constrained optimization problems.

The originality of this research lies in the smooth incorporation of practical operational constraints—such as battery aging effects and the exclusivity of energy acquisition versus dispatch—into the optimization model, thereby ensuring both real-world applicability and flexibility. Moreover, this investigation fills an important gap in the existing literature by demonstrating the scalability of a parallelized SA algorithm for optimizing VPP networks comprising hundreds of prosumers, delivering near-linear speedups across multiple processing cores. Achieving such scalability is essential for contemporary VPPs, as scheduling complexity increases exponentially with the number of distributed energy resources (DERs) and fluctuations in market conditions.

In addition, this research pushes forward the current frontier by placing strong emphasis on the real-time viability of parallelized metaheuristic techniques for VPP optimization. In contrast to earlier works that primarily focus on static or limited-scale problem settings, the proposed method is engineered to adaptively respond to changing market environments, delivering prompt and reliable solutions for real-time energy trading and resource coordination. The HPC-supported structure enables high-frequency decision-making, empowering VPP operators to improve scheduling performance and market engagement even in highly unpredictable energy landscapes.

When compared with previous efforts, which typically assess optimization approaches under idealized or reduced-complexity conditions, this study effectively connects computational progress with tangible deployment. By integrating parallelized metaheuristic optimization with HPC resources, the work presents a scalable, high-efficiency solution tailored to the increasingly complex demands of modern VPPs, thereby making a meaningful contribution to the continued advancement of sustainable energy infrastructure.

Materials and Methods

Problem description

The VPP market optimization service is an essential framework that delivers efficient, real-time energy coordination by managing the interplay between grid requirements and prosumer operations. This service consists of two fundamental elements: forecasting and optimization. The forecasting element processes both live and historical data to anticipate future patterns in energy demand, supply availability, and market behavior, thereby providing the optimization component

with critical predictive information. These predictions allow the optimization service to generate prosumer schedules that maximize overall social welfare by harmonizing energy supply and demand throughout the VPP network.

Central to this optimization structure is the goal of maximizing social welfare via carefully planned scheduling of energy exchanges for every participant in the VPP. Each VPP operates as a prosumer, able to purchase or sell electricity to the grid at each time interval. The optimization service uses both forecast and current data to calculate the optimal buy or sell volumes for each prosumer. Notably, the optimization process does not incorporate prosumer price bids; instead, it coordinates prosumer transactions based on the grid's fixed price signals for each period.

The underlying system relies on an Internet of Things (IoT) data streaming infrastructure that ensures uninterrupted data exchange between prosumer devices and the optimization service. This infrastructure collects and retains information from prosumer assets, including energy production, consumption levels, and battery conditions. It also relays the computed optimal scheduling instructions back to the prosumer devices in real time, allowing them to implement the recommended buy or sell actions. Consequently, this IoT-supported data framework underpins the instantaneous decision-making capacity required for agile VPP market functioning.

The optimization service functions within a well-defined ecosystem comprising three main actors: the grid, the prosumers (VPPs), and the internal VPP market. The grid determines the electricity price for each period, while prosumers are limited to submitting quantity-based offers and adjust their purchasing or selling actions based on these unchanging prices. Inside the VPP market, the optimization service orchestrates these exchanges to maximize total market welfare while respecting all operational limitations. By synchronizing prosumer behaviors with grid-defined parameters, the optimization maintains effective market performance, stability, and operational durability.

The design specifications for this optimization framework are outlined as follows:

- **Rolling-based Social Welfare Maximization:** The core objective centers on enhancing social welfare through the scheduling of energy exchanges in a manner that satisfies the combined requirements of the grid and participating prosumers. This optimization operates on a rolling horizon, enabling the system to adapt promptly to instantaneous shifts in supply and demand and to remain synchronized with prevailing market realities. By recasting the problem as a minimization problem, the framework integrates additional factors, such as battery usage patterns and associated operational expenses, thereby achieving social welfare gains while preserving system efficiency and economic performance.
- **Scalability to Support Growing Market Participation:** The architecture must maintain high computational performance and timely execution even as the number of prosumers increases significantly. This capacity for expansion is vital to sustaining the framework's reliability and effectiveness within a rapidly enlarging virtual power plant (VPP) ecosystem.
- **Battery Longevity Constraints:** To safeguard the durability and sustained functionality of prosumer batteries, the optimization enforces limits on charge and discharge operations. These restrictions keep battery parameters within acceptable ranges, encourage environmentally responsible usage patterns, and extend the service life of energy storage assets.
- **Grid-Fixed Pricing with Quantity-Driven Scheduling:** Under this arrangement, electricity prices are exclusively set by the grid, and prosumers do not submit price proposals. The optimization procedure coordinates each prosumer's transactions to align precisely with these fixed-price indicators, allowing prosumers to specify only the volumes they can produce or use. This structure delivers a clear, price-responsive scheduling mechanism tailored to the VPP market environment.

The VPP market optimization model addresses an intricate, continually evolving challenge that requires seamless integration of predictive analytics, continuous data streaming, and flexible scheduling. To meet the demands of contemporary energy infrastructure, this design must provide strong scalability, instantaneous adaptability, and reliable

data handling. By drawing on forecast information and aligning prosumer behaviors with predefined economic and technical boundaries, the VPP optimization service aims to elevate overall market welfare, promote long-term sustainable energy usage, and enable swift, market-consistent decisions in practical energy management contexts.

Figure 1 illustrates the structure of the VPP management framework operating under a grid-regulated pricing regime. The diagram emphasizes the relationships among prosumers, the VPP market, and the grid. Within this setup, the VPP serves as a central coordinator, overseeing energy trading among prosumers and ensuring compliance with the uniform electricity prices set by the grid.

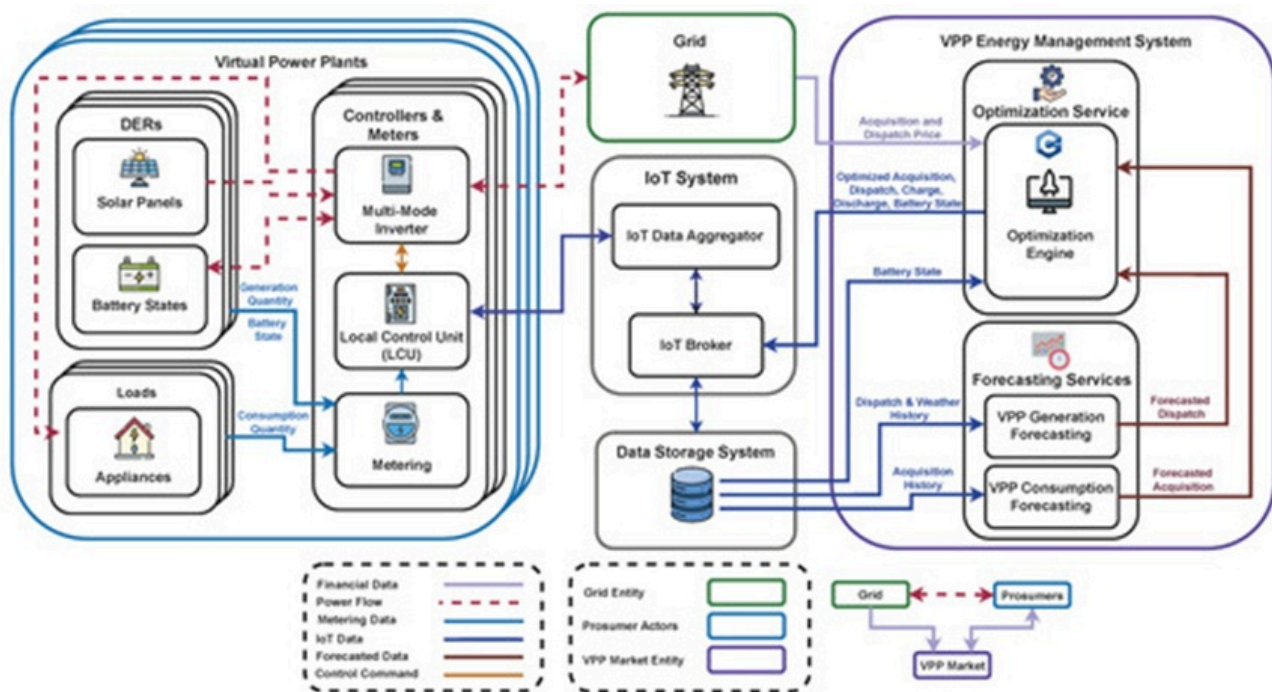


Figure 1. Schematic representation of the VPP management framework, depicting the principal elements and information exchange among prosumers, the VPP market, and the grid. The VPP operates as an intermediary, enabling energy transactions at grid-determined pricing, while a centralized scheduler optimizes resource distribution.

Prosumers regularly send live updates—including current energy production, consumption levels, and battery status—to the VPP market via an Internet of Things (IoT)-supported data-streaming infrastructure. This infrastructure serves as the foundational layer of the entire system, ensuring fluid and reliable data communication.

The VPP's energy management system, serving as a central scheduler, comprises two primary components. The initial component is a forecasting module that processes both real-time and archived data to anticipate future energy generation and load profiles for individual prosumers. These predictive outputs supply indispensable information to the optimization module, which dynamically assigns distributed energy resources (DERs). This module calculates the optimal quantities of energy each prosumer should purchase, sell, or retain in storage at each time interval, while respecting grid-imposed limitations. By coordinating the management of these operations, the VPP framework improves overall system performance, balances energy availability and consumption, and enhances the reliability of the broader energy network.

Problem formulation

The scheduling challenge for the virtual power plant (VPP) is formulated as a Mixed-Integer Linear Programming (MILP) model that maximizes social welfare while respecting constraints on energy procurement, energy delivery, and battery operational dynamics [38].

Parameters and variables

Table 1 presents the essential parameters and decision variables employed in the VPP market optimization model. These encompass critical elements such as electricity prices, battery operational boundaries, and time discretization, as well as decision variables that control energy acquisition, dispatch decisions, and battery charging/discharging. The table organizes each item by category and provides precise notation for both scalar and vector quantities, thereby facilitating a systematic representation of the system's time-varying energy coordination needs.

Table 1. Parameters and decision variables for VPP market optimization.

Category	Symbol	Description	Type	Elements (in Vectors)
Parameters	$\Pi_{t,I}^{acq}$	Acquisition prices for energy per prosumer at time t	Vector	$\Pi_{t,I}^{acq} = [\pi_{t,1}^{acq}, \dots, \pi_{t,N_i}^{acq}]^T$
	$\Pi_{t,I}^{disp}$	Dispatch prices per prosumer at time t	Vector	$\Pi_{t,I}^{disp} = [\pi_{t,1}^{disp}, \dots, \pi_{t,N_i}^{disp}]^T$
	C^B	Battery-related costs	Scalar	-
	C^{fix}	Fixed operational costs	Scalar	-
	Δt	Duration of each time period	Scalar	-
	$\underline{Q}_{t,I}^{acq}, \overline{Q}_{t,I}^{disp}$	Maximum acquisition and dispatch limits per prosumer	Vectors	$[\underline{q}_{t,1}^{acq}, \dots, \underline{q}_{t,N_i}^{acq}]^T, [\overline{q}_{t,1}^{disp}, \dots, \overline{q}_{t,N_i}^{disp}]^T$
	$\underline{Q}_{t,I}^{ch}, \overline{Q}_{t,I}^{dch}$	Maximum charge and discharge limits per prosumer	Vectors	$[\underline{q}_{t,1}^{ch}, \dots, \underline{q}_{t,N_i}^{ch}]^T, [\overline{q}_{t,1}^{dch}, \dots, \overline{q}_{t,N_i}^{dch}]^T$
	$S_{bat,initI}$	Initial battery state per prosumer	Vector	$S_I^{bat,initI} = [s_1^{bat,initI}, \dots, s_{N_i}^{bat,initI}]^T$
	$\underline{S}_{t,I}^{bat}, \overline{S}_{t,I}^{bat}$	Battery capacity limits (min, max) per prosumer	Vectors	$[\underline{s}_{t,1}^{bat}, \dots, \underline{s}_{t,N_i}^{bat}]^T, [\overline{s}_{t,1}^{bat}, \dots, \overline{s}_{t,N_i}^{bat}]^T$
	$Q_{t,I}^{gen}$	Generated energy per prosumer at time t	Vector	$[q_{t,1}^{gen}, \dots, q_{t,N_i}^{gen}]^T$
Continuous decision variables	$Q_{t,I}^{load}$	Energy load consumption per prosumer at time t	Vector	$[q_{t,1}^{load}, \dots, q_{t,N_i}^{load}]^T$
	$Q_{t,I}^{acq}$	Energy acquisition amounts at time t	Vector	$[q_{t,1}^{acq}, \dots, q_{t,N_i}^{acq}]^T$
	$Q_{t,I}^{disp}$	Energy dispatch amounts at time t	Vector	$[q_{t,1}^{disp}, \dots, q_{t,N_i}^{disp}]^T$
	$Q_{t,I}^{ch}, Q_{t,I}^{dch}$	Battery charge and discharge amounts per prosumer at time t	Vectors	$[q_{t,1}^{ch}, \dots, q_{t,N_i}^{ch}]^T, [q_{t,1}^{dch}, \dots, q_{t,N_i}^{dch}]^T$

	$S_{t,I}^{bat}$	Battery state at time t	Vector	$\left[S_{t,1}^{bat}, \dots, S_{t,N_i}^{bat} \right]^T$
	$Q_{t,I}^{disp,paid}, Q_{t,I}^{disp,non-paid}$	Paid and non-paid energy dispatch per prosumer at time t	Vectors	$\begin{bmatrix} q_{t,1}^{disp,non-paid} \\ \vdots \\ q_{t,N_i}^{disp,non-paid} \end{bmatrix}$
Binary decision variables	$\Xi_{t,I}^{acq}$	Acquisition indicator per prosumer at time t	Binary Vector	$\left[\xi_{t,1}^{acq}, \dots, \xi_{t,N_i}^{acq} \right]^T$
	$\Xi_{t,I}^{disp}$	Dispatch indicator per prosumer at time t	Binary Vector	$\left[\xi_{t,1}^{disp}, \dots, \xi_{t,N_i}^{disp} \right]^T$
	$\rho_{t,I}^{ch}$	Charge indicator per prosumer at time t	Binary Vector	$\left[\rho_{t,1}^{ch}, \dots, \rho_{t,N_i}^{ch} \right]^T$
	$\rho_{t,I}^{dch}$	Discharge indicator per prosumer at time t	Binary Vector	$\left[\rho_{t,1}^{dch}, \dots, \rho_{t,N_i}^{dch} \right]^T$

This table compiles all principal parameters and decision variables required for the VPP market-optimization process, where $i \in \{1, \dots, N_i\}$ indexes each prosumer and $t \in \mathcal{T}$ indexes the set of time periods.

Objective function

The central aim of the optimization model is to reduce the VPP's aggregate operational expenditures. These expenditures cover costs for energy procurement and dispatch, battery use (including charging and discharging), and ongoing operational overhead. Minimizing these costs is mathematically equivalent to maximizing social welfare, since it corresponds to the negation of the social welfare expression. By reducing costs, the model effectively increases the overall net benefits for the entire system, thereby promoting economic efficiency across all parties. This approach encapsulates the core social welfare objective by optimally balancing energy provision, consumption patterns, and resource allocation to deliver maximum collective benefit. To ensure realistic, implementable outcomes, the model incorporates a full set of operational constraints, including battery storage capacity limits, charging and discharging rate limits, and energy equilibrium conditions.

The objective function, formulated to capture this cost-minimization goal, serves as the foundational element of the entire optimization procedure. It empowers the VPP to coordinate distributed energy resources (DERs) with high efficiency while successfully managing fluctuating market signals, variable load profiles, and physical limitations. The precise mathematical definition of the objective function is given as

$$\min_{Q_{t,I}^{acq}, Q_{t,I}^{disp}} Z = \sum_{t \in \mathcal{T}} \left((Q_{t,I}^{acq})^T \Pi_{t,I}^{acq} - (Q_{t,I}^{disp})^T \Pi_{t,I}^{disp} \right) \Delta t + (C^B + C^{fix}) \quad (1)$$

Energy balance constraints

The energy balance condition guarantees that, at every time step, the total energy entering the system (from acquisition, local generation, and battery discharge) precisely matches the total energy leaving the system (through dispatch, local consumption, and battery charging). This relationship is mathematically formulated as

$$Q_{t,I}^{acq} + Q_{t,I}^{gen} + Q_{t,I}^{dch} = Q_{t,I}^{disp} + Q_{t,I}^{load} + Q_{t,I}^{ch}, \quad \forall t \in \mathcal{T}, \forall i \in \mathcal{I} \quad (2)$$

Acquisition and dispatch decision constraints

Binary variables govern prosumers' decisions to acquire or dispatch energy. The constraints below enforce valid acquisition and dispatch choices:

$$0 \leq Q_{t,I}^{acq} \leq \overline{Q}_{t,I}^{acq} \cdot \xi_{t,I}^{acq}, \quad 0 \leq Q_{t,I}^{disp} \leq \overline{Q}_{t,I}^{disp} \cdot \xi_{t,I}^{disp} \quad (3)$$

$$0 \leq \xi_{t,I}^{acq}, \xi_{t,I}^{disp} \leq 1, \quad \forall t \in \mathcal{T}, \forall i \in \mathcal{I} \quad (4)$$

Additionally, the next constraint prevents simultaneous acquisition and dispatch operations:

$$\xi_{t,I}^{acq} + \xi_{t,I}^{disp} \leq 1, \quad \forall t \in \mathcal{T}, \forall i \in \mathcal{I} \quad (5)$$

Dispatch breakdown

The total dispatched energy $Q_{t,I}^{disp}$ consists of two distinct components:

$$Q_{t,I}^{disp} = Q_{t,I}^{disp,paid} + Q_{t,I}^{disp,non-paid} \quad (6)$$

Where the following definitions hold:

$Q_{t,I}^{disp,paid}$: the portion of dispatched energy that is compensated at time t .

$Q_{t,I}^{disp,non-paid}$: the portion of dispatched energy that is uncompensated at time t .

The uncompensated dispatch is further restricted by the condition:

$$0 \leq Q_{t,I}^{disp,non-paid} \leq \overline{Q}_{t,I}^{non-paid} \cdot \xi_{t,I}^{disp}, \quad \forall t \in \mathcal{T}, \forall i \in \mathcal{I} \quad (7)$$

The variable $Q_{t,I}^{non-paid}$ is assigned a value of ∞ .

Battery dynamics

The initial battery state of charge at $t=1$ is set as

$$S_{t,I}^{bat} = S_I^{bat,init} + (Q_{t,I}^{ch} - Q_{t,I}^{dch})\Delta t, \quad t = 1, \forall i \in \mathcal{I} \quad (8)$$

For all subsequent time steps, the battery state evolves according to

$$S_{t,I}^{bat} = S_{t-1,I}^{bat} + (Q_{t,I}^{ch} - Q_{t,I}^{dch})\Delta t, \quad \forall t \in \mathcal{T}, \forall i \in \mathcal{I} \quad (9)$$

Charge and discharge constraints

These constraints prevent simultaneous charging and discharging. Binary variables $\rho_{t,I}^{ch}$ and $\rho_{t,I}^{dch}$ regulate charging and discharging activities:

$$0 \leq Q_{t,I}^{ch} \leq \overline{Q}_{t,I}^{ch} \cdot \rho_{t,I}^{ch}, \quad 0 \leq Q_{t,I}^{dch} \leq \overline{Q}_{t,I}^{dch} \cdot \rho_{t,I}^{dch} \quad (10)$$

$$\rho_{t,I}^{ch} + \rho_{t,I}^{dch} \leq 1, \quad \forall t \in \mathcal{T}, \forall i \in \mathcal{I} \quad (11)$$

Battery capacity constraints

The battery state of charge must always remain within its allowable physical limits:

$$\underline{S}_{t,I}^{bat} \leq S_{t,I}^{bat} \leq \overline{S}_{t,I}^{bat}, \quad \forall t \in \mathcal{T}, \forall i \in \mathcal{I} \quad (12)$$

Problem exploration and parallel computation

Within the VPP optimization structure, each participant i at every time step t must maintain energy equilibrium among the grid connection, the battery, and their local generation and load. Choices concerning energy acquisition (purchasing from the grid), dispatch (selling to the grid), battery charging, and battery discharging are mutually exclusive and bounded above by technical limits. This section formally derives the energy balance relationship, incorporating the exclusivity rules and bounding conditions. It further demonstrates how the feasible solution region—restricted to a one-dimensional line rather than a two-dimensional plane—can be explored efficiently and in parallel for each participant.

Energy balance with mutual exclusivity and constraints

For any time step t , the energy balance equation for participant i is expressed as

$$q_{t,I}^{acq} + q_{t,I}^{gen} + q_{t,I}^{dch} = q_{t,I}^{disp} + q_{t,I}^{load} + q_{t,I}^{ch}, \quad \forall t \in \mathcal{T}, \forall i \in \mathcal{I} \quad (13)$$

Where:

$q_{t,i}^{acq}$ represents the energy purchased from the grid,

$q_{t,i}^{disp}$ denotes the energy sold to the grid,

$q_{t,i}^{ch}$ indicates the energy directed to battery charging,

$q_{t,i}^{dch}$ stands for the energy withdrawn from the battery,

$q_{t,i}^{gen}$ corresponds to the locally generated energy (e.g., from photovoltaic panels),

and $q_{load,t,i}$ is the energy demanded by the participant's loads.

Mutually exclusive flip-flop mechanism for grid and battery operations

In the VPP model, interactions involving the grid and the battery are governed by a mechanism that enforces mutual exclusivity. This ensures that at each time step t , every participant i is permitted to either purchase energy from the grid or sell energy to the grid, and to either charge the battery or discharge it—but never to perform both actions in the same pair at the same time. This strict exclusivity preserves logical consistency in energy flows and avoids conflicting or impossible simultaneous operations.

The rules enforcing this mutual exclusivity are specified as follows:

$$q_{t,i}^{acq} > 0 \Rightarrow q_{t,i}^{disp} = 0, \quad q_{t,i}^{disp} > 0 \Rightarrow q_{t,i}^{acq} = 0 \quad (14)$$

$$q_{t,i}^{ch} > 0 \Rightarrow q_{t,i}^{dch} = 0, \quad q_{t,i}^{dch} > 0 \Rightarrow q_{t,i}^{ch} = 0 \quad (15)$$

To implement this exclusive behavior, the mechanism operates as a flip-flop. Binary decision variables are employed to toggle between active and inactive states for each type of action, depending on the participant's current needs. Once one operation in a pair is activated (for instance, acquisition or charging), the opposing operation (dispatch or discharging) is automatically disabled.

The binary decision variables that drive this flip-flop logic are defined as

$$\xi_{t,i}^{acq} + \xi_{t,i}^{disp} \leq 1, \quad \rho_{t,i}^{ch} + \rho_{t,i}^{dch} \leq 1 \quad (16)$$

Where the following interpretations hold:

$\xi_{t,i}^{acq} = 1$ activates energy acquisition and forces $\xi_{t,i}^{disp} = 0$.

$\xi_{t,i}^{disp} = 1$ activates energy dispatch and forces $\xi_{t,i}^{acq} = 0$.

$\rho_{t,i}^{ch} = 1$ activates battery charging and forces $\rho_{t,i}^{dch} = 0$.

$\rho_{t,i}^{dch} = 1$ activates battery discharging and forces $\rho_{t,i}^{ch} = 0$.

This two-tiered structure—combining mutual exclusivity with flip-flop control—guarantees that all operations involving the grid and battery remain internally consistent, thereby preserving the overall stability of the VPP optimization framework.

Upper-bound constraints on energy flows

Beyond the mutual exclusivity rules, the decision variables representing energy acquisition, dispatch, charging, and discharging are further limited by upper bounds that reflect the maximum permissible energy transfer rates in each direction:

$$0 \leq \mathbf{q}_{t,I}^{\text{acq}} \leq \overline{\mathbf{q}}_{t,I}^{\text{acq}}, \quad 0 \leq \mathbf{q}_{t,I}^{\text{disp}} \leq \overline{\mathbf{q}}_{t,I}^{\text{disp}} \quad (17)$$

$$0 \leq \mathbf{q}_{t,I}^{\text{ch}} \leq \overline{\mathbf{q}}_{t,I}^{\text{ch}}, \quad 0 \leq \mathbf{q}_{t,I}^{\text{dch}} \leq \overline{\mathbf{q}}_{t,I}^{\text{dch}} \quad (18)$$

These bounds ensure that all chosen actions remain within realistic physical and technical capabilities.

Linearized energy balance equation

To streamline the energy balance relationship, two auxiliary quantities are introduced to represent the net energy exchange with the grid and with the battery, respectively:

$$\mathbf{X}_{t,I} = \mathbf{q}_{t,I}^{\text{acq}} - \mathbf{q}_{t,I}^{\text{disp}} \quad (\text{Net grid interaction}) \quad (19)$$

$$\mathbf{Y}_{t,I} = \mathbf{q}_{t,I}^{\text{ch}} - \mathbf{q}_{t,I}^{\text{dch}} \quad (\text{Net battery activity}) \quad (20)$$

Furthermore, the net difference between the participant's local generation and consumption is expressed as

$$\mathbf{B}_{t,I} = \mathbf{q}_{t,I}^{\text{gen}} - \mathbf{q}_{t,I}^{\text{load}} \quad (\text{Net energy surplus / deficit}) \quad (21)$$

Consequently, the energy balance can be reformulated as a linear expression:

$$\mathbf{Y}_{t,I} = \mathbf{X}_{t,I} + \mathbf{B}_{t,I}, \quad \forall t \in \mathcal{T}, \forall i \in \mathcal{I} \quad (22)$$

The following definitions hold:

- $\mathbf{Y}_t$ denotes the net battery power flow of the player (positive for discharging, negative for charging).
- $\mathbf{X}_t$ denotes the net power exchange with the grid (positive for purchasing, negative for selling).
- $\mathbf{B}_t$ denotes the net power production minus consumption by the player.

This relation expresses the instantaneous energy balance for the player at every time interval, ensuring that battery usage $\mathbf{Y}_t$ exactly matches the difference between grid exchange $\mathbf{X}_t$ and the player's internal net generation/consumption $\mathbf{B}_t$.

Full energy balance including constraints

For each participant, the comprehensive energy balance must respect both mutual-exclusivity rules and capacity limits. This leads to the complete constraint set:

$$\mathbf{Y}_{t,I} = \mathbf{X}_{t,I} + \mathbf{B}_{t,I}, \quad -\overline{\mathbf{q}}_{t,I}^{\text{disp}} \leq \mathbf{X}_{t,I} \leq \overline{\mathbf{q}}_{t,I}^{\text{acq}}, \quad -\overline{\mathbf{q}}_{t,I}^{\text{dch}} \leq \mathbf{Y}_{t,I} \leq \overline{\mathbf{q}}_{t,I}^{\text{ch}} \quad (23)$$

The introduction of binary variables $\dots$ enforces strict mutual exclusivity of decisions at each timestep, thereby eliminating physically impossible states such as simultaneous grid purchase and sale or concurrent battery charge and discharge.

Figure 2 illustrates how the search space is progressively narrowed according to **Eq. 23**. In the beginning, the prosumer's feasible choices occupy a broad two-dimensional plane containing many independent degrees of freedom and numerous potential interaction patterns. The optimization procedure transforms this complex, high-dimensional region into a single constrained line segment delimited by the tight upper and lower limits imposed by **Eq. 23**. Such a contraction guarantees compliance with all mutual-exclusivity and bounding restrictions while dramatically reducing the computational burden. The resulting streamlined decision landscape enables genuine real-time decision-making and allows the overall architecture to scale efficiently to much larger collections of prosumers.

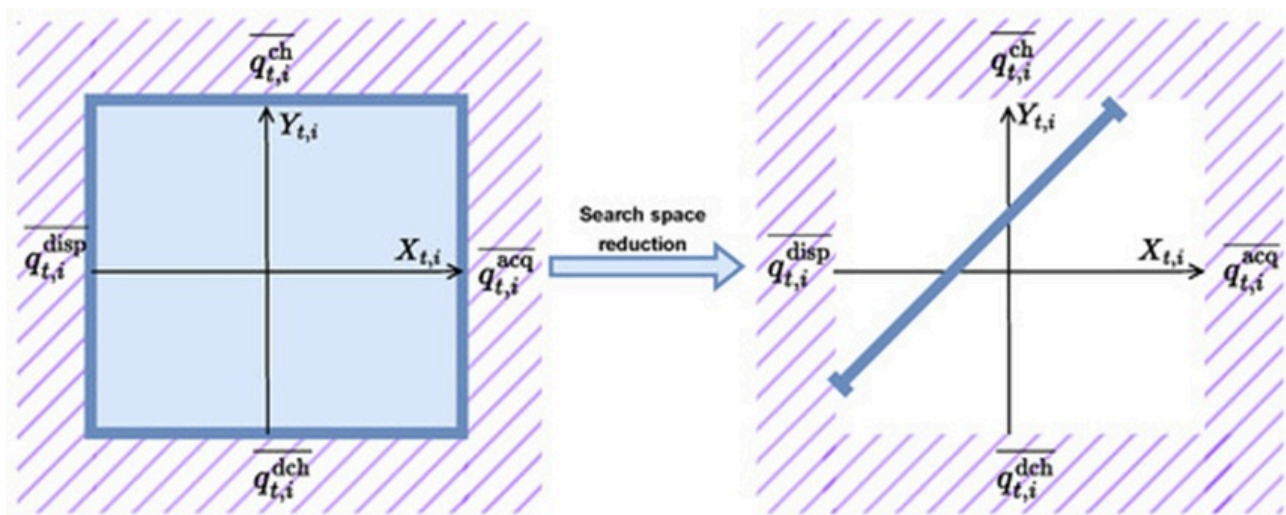


Figure 2. Schematic representation of search-space contraction for an individual prosumer. The starting point is a wide 2D decision plane; optimization collapses it into a narrowly bounded linear trajectory defined by upper and lower thresholds, thereby markedly boosting computational performance and enabling scalable coordination of extensive prosumer populations.

Solver architecture and algorithmic strategy

To tackle the virtual power plant (VPP) scheduling task with both rigorous accuracy and practical speed, two complementary solution strategies were developed: a definitive Mixed-Integer Linear Programming (MILP) solver powered by Gurobi, and a parallelized Simulated Annealing (SA) procedure.

Simulated Annealing was adopted because of its well-documented ability to handle intricate, non-convex, and nonlinear landscapes effectively. Its probabilistic mechanism allows temporary acceptance of inferior moves during early cooling phases, helping the algorithm escape local optima and reducing the risk of premature trapping. The method's conceptual simplicity, combined with its ease of tailoring to domain-specific restrictions, makes it versatile across diverse problem classes [39]. When benchmarked against heavier population-based techniques such as genetic algorithms or particle swarm methods, SA strikes an attractive compromise between execution time and attained solution quality [39]. Furthermore, because it follows a single-solution trajectory rather than maintaining a population, its per-iteration cost remains modest—an advantage that becomes especially valuable when parallelizing the algorithm for demanding applications such as VPP dispatch. In particular, parallelizing SA via OpenMP yields substantial performance gains and renders the approach highly suitable for computationally intensive, large-scale VPP instances [40–43]. Overall, this profile

of strong scalability, reasonable runtime, and dependable solution quality establishes SA as a solid counterpart to the exact MILP method, particularly for scenarios involving very large numbers of controllable assets.

Exact solver

The present study utilizes the Gurobi commercial solver to obtain provably optimal solutions to the VPP scheduling problem cast as an MILP. Gurobi implements a robust branch-and-bound search augmented by powerful presolving, advanced cutting-plane generation, and sophisticated heuristics, all of which reliably find the global optimum within practical runtime budgets. This exact methodology delivers guaranteed best possible schedules.

Although the core contribution of the work centers on the scalable parallel SA technique for realistically sized instances, Gurobi serves as the gold-standard reference against which SA performance is measured. A direct side-by-side comparison of provably optimal MILP results with the near-optimal solutions returned by SA provides a clear, quantitative assessment of the metaheuristic's fidelity. It thereby substantiates its practical value for addressing significantly larger and structurally more challenging VPP coordination problems.

Parallel simulated annealing solver

Although Gurobi, our chosen solver, delivers strong scalability and high performance, it remains a commercial, license-based tool that incurs costs. To increase implementation flexibility and provide a viable option in situations where a free or alternative solver is preferable, we developed a parallelized Simulated Annealing (SA) solver. This implementation aims to deliver high-quality, near-optimal schedules rapidly, even for extensive VPP systems, by leveraging flexible, metaheuristic search strategies tailored to the intricate, time-varying characteristics of VPP optimization.

Simulated Annealing is a stochastic metaheuristic inspired by the metallurgical annealing process, in which a substance is slowly cooled to reach a low-energy configuration. The method excels at addressing challenging optimization tasks characterized by enormous, rugged search landscapes, since it can escape local optima by occasionally accepting worse moves, thereby increasing the probability of finding a globally optimal or very close-to-optimal solution.

In our SA framework, the algorithm repeatedly perturbs the dispatch decisions assigned to each participant at each time interval, aiming to minimize total system operating costs. Scalability is achieved by applying OpenMP to distribute the workload across several threads, assigning each thread to optimize the schedule for a specific player independently. The overall design of this parallel SA method is presented in **Algorithm 1**.

1: Input: Initial state S_0 , Initial temperature T_0 , Cooling rate α , Perturbation scale γ , Number of iterations N_{iter} , Number of players N_i , Number of periods N_t , Battery dynamics, and cost parameters from the input.	
2: Output: Best global state S_{best} , Best global cost Z_{best}	
3: Initialize: Shared memory for global best state S_{best} and global best cost Z_{best} to store optimization results.	
4: Initialize temperature T_0 , objective function (S) for cost calculation.	
5: #pragma omp parallel for shared(S_{best}, Z_{best}) private(S_i, Z_i, T_i) reduction(+: Z_{best})	// Parallelized loop for each player i
6: for $i=1$ to N_i do	// Each player is solved independently in parallel
7: Initialize local state $S_i=S_0$, local cost $Z_i=Z(S_i)$	
8: Set $S_{i,best}=S_i$, $Z_{i,best}=Z_i$, $T_i=T_0$	
9: Initialize period states for player i across all periods.	
10: Calculate initial battery state e_i , based on input parameters and previous period states.	
11: for $k=1$ to N_{iter} do	// Iterative optimization process
12: for $t=1$ to N_t do	// Optimize for each period t for player i
13: Perturb player state S_i^{new} for period t using γ	
14: Compute new balance for the player, update charge/discharge states q_{ch}, q_{dch}	
15: Calculate new battery state e_{bat}^{new} for period t	
16: end for	
17: Compute new cost $Z_i^{new}=Z(S_i^{new})$	// Moved outside of inner loop
18: Calculate acceptance probability $P=\min(1, \exp(\frac{Z_i - Z_i^{new}}{T_i}))$	
19: if $r < P$ then	// Random value $r \in [0, 1]$
20: Update $S_i=S_{newi}$, $Z_i=Z_{newi}$	
21: if $Z_i < Z_{i,best}$ then	
22: Update $S_{i,best}=S_i$, $Z_{i,best}=Z_i$	
23: end if	
24: end if	
25: Apply cooling schedule: $T_i=\alpha T_i$	
26: end for	
27: Update global cost Z_{best} using OpenMP reduction.	
28: end for	
29: Return: Best global state S_{best} , Best global cost Z_{best}	

Algorithm 1. Simulated annealing with parallel optimization using OpenMP

The procedure begins from an initial configuration S_0 at an elevated temperature T_0 . At high temperatures, the algorithm readily accepts inferior moves, guided by a Boltzmann acceptance criterion that promotes broad exploration of the solution space. As cooling progresses according to the parameter α , the likelihood of accepting detrimental changes diminishes, gradually shifting emphasis from exploration toward exploitation of promising regions. This controlled cooling mechanism enables the method to avoid becoming trapped in poor local optima and improves its prospects of converging to a near-optimal schedule for the VPP problem.

Parallelization via OpenMP permits independent optimization of each player's subproblem: every thread modifies that player's decisions over the full time horizon while respecting the energy balance equation and battery limitations. This decoupled, thread-based organization substantially shortens overall run times and improves the algorithm's scalability, rendering parallel SA especially well-suited for large-scale, time-critical VPP scheduling applications. By adopting parallel SA, the solver strikes a practical compromise between processing speed and solution accuracy, thereby facilitating flexible, cost-effective management of VPPs in sophisticated electricity markets.

As shown in **Figure 3**, the underlying high-performance computing (HPC) cluster design reinforces this parallelism by dedicating a distinct physical node to each VPP. Each node autonomously performs scheduling optimization for its assigned VPP. Within each node, a soft-computing module runs a parallel SA implementation powered by OpenMP, enabling concurrent processing of multiple players and delivering optimized schedules with high efficiency. This distributed architecture markedly improves both scalability and execution performance, enabling responsive, real-time control of energy systems and ensuring maximum operational agility.

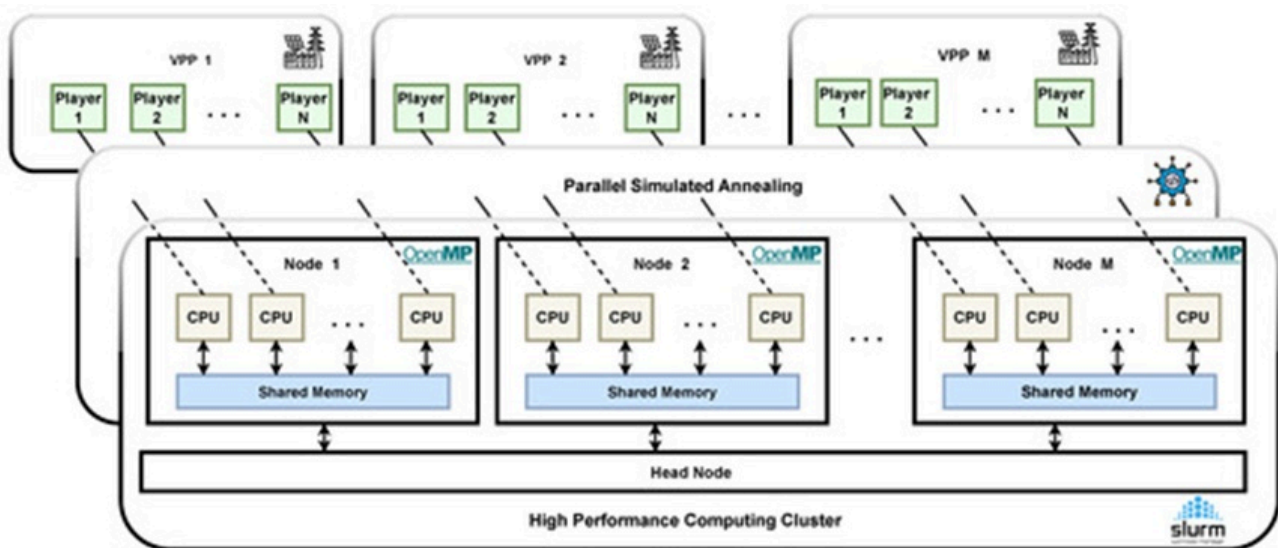


Figure 3. Schematic overview of the high-performance computing cluster setup: The figure depicts the HPC cluster layout, with each physical node responsible for optimizing one independent VPP. Inside each node, a soft-computing component performs parallel Simulated Annealing using OpenMP, enabling effective concurrency across player optimizations and ensuring high-quality scheduling results. This arrangement substantially increases scalability and processing throughput, thereby enabling real-time VPP coordination and enhancing the overall responsiveness of the energy infrastructure.

Experimental design for high-performance computing

The experiments are organized to rigorously test the computational behavior of the parallel SA solver when deployed on HPC resources. The central goal is to quantify the algorithm's scalability and runtime efficiency across different hardware configurations and problem sizes, specifically by varying the number of available CPU cores and the number of participants (prosumers) in the VPP aggregation.

The controlled (independent) variables in the experimental design are:

- Number of cores: Tests were run using between 1 and 32 CPU cores. This spectrum enables measurement of the parallel performance benefits at progressively higher levels of parallel resources.

- Number of players (prosumers): To replicate growing VPP market sizes and evaluate algorithmic scalability, the participant count was varied from 32 to 512. This parameter directly governs problem complexity, since each additional prosumer introduces a separate set of decision variables, state transitions, and constraint evaluations during SA iterations.

The observed (dependent) performance indicators are:

- Averaged execution time (in seconds): This value captures the mean runtime needed to complete the SA procedure for each combination of core count and player count. Tracking how execution time changes with increasing cores and players reveals the algorithm's efficiency and highlights any points at which additional cores yield progressively smaller gains (Amdahl's law effects).
- Speedup ratio: Defined as the single-core execution time divided by the multi-core execution time, this dimensionless metric quantifies the actual acceleration achieved through parallel execution. It serves as a key indicator of parallel efficiency and scalability, helping to pinpoint overheads, load-balancing issues, or saturation points as core utilization rises.

Experimental procedure

For every experimental run, the SA solver was initialized using consistent parameter settings—including the starting temperature, cooling schedule coefficient, and perturbation magnitude—carefully chosen to strike an appropriate balance between thorough exploration of the solution space and convergence toward precise solutions. This ensured uniformity and reproducibility across all trials. The HPC cluster was configured to dedicate isolated computational resources to each experiment, thereby preventing any interference or performance variability arising from concurrent resource contention. To mitigate the effects of the algorithm's inherent randomness, we performed multiple independent executions for each combination of core count and player count, then computed and reported the averaged outcomes.

The experimental protocol was crafted to closely mirror practical VPP scheduling situations in real energy markets, where the number of participating prosumers can fluctuate significantly. By progressively scaling both the problem size (via player count) and the available parallel resources (via core count), the design reveals both absolute runtime reductions and relative efficiency gains attributable to parallel execution. Of particular interest is the speedup ratio, which demonstrates how effectively the OpenMP-parallelized SA implementation maintains strong scaling behavior as the VPP aggregation grows, thereby underscoring the solver's capacity to manage substantially larger market sizes with sustained performance.

Analysis and interpretation

This structured experimental approach facilitates a thorough evaluation of the parallel SA solver's behavior, with primary emphasis on execution time trends and speedup factors across the full range of tested configurations. Analysis of how variations in core availability and prosumer numbers affect these metrics allows precise quantification of parallelization benefits—specifically the degree to which additional cores reduce overall computation time—and helps determine the most advantageous core-to-problem-size ratios. The resulting findings offer concrete guidance for deploying high-performance computing in VPP scheduling applications, clearly illustrating the substantial computational advantages that can be realized when tackling realistic, high-dimensional, large-scale optimization challenges.

Implementation

HPC environment specifications

All VPP scheduling simulations were conducted on the Oblivion HPC cluster, which comprises 88 compute nodes. Each node is equipped with dual Intel Xeon Gold 6154 processors (supplied by Intel Corp.) and 192 GB of RAM, for a total of

36 cores per node. This yields an aggregate of 3168 CPU cores across the system. Such hardware delivers the necessary core density and memory bandwidth to support intensive parallel workloads effectively within our optimization framework.

The Oblivion cluster's architecture features low-latency interconnects and high data throughput, both of which are critical for efficiently parallelizing the VPP scheduling task. By enabling independent solution of each prosumer's subproblem, the multi-core design dramatically accelerates overall execution compared to sequential processing.

To maximize compatibility and extract peak performance from the hardware, we employed GCC version 11.3.0 together with Intel-optimized compilation flags and full OpenMP runtime support, all finely tuned to leverage the specific capabilities of the Xeon Gold processors. This carefully configured environment ensured robust, high-throughput parallel execution, indispensable given the scale and complexity of realistic virtual power plant optimization problems.

Model parameters and outputs

The following subsections specify the key input parameters and resulting output variables of the VPP scheduling optimization model. The inputs establish the boundary conditions, technical limitations, and economic drivers that govern the decision-making process. At the same time, the outputs encapsulate the optimized decisions and performance metrics derived from applying the solver.

Model input parameters

Table 2 lists the principal time-dependent, technical, and financial parameters that shape VPP operations. These values collectively define the operational envelope and constraints within which the optimization must function, allowing the model to respond dynamically to forecasted load profiles, renewable generation availability, and prevailing market prices.

Table 2. Model input parameters for VPP scheduling.

Parameter	Designation	Value range	Unit
$N\mathcal{T}$	Total periods	96	-
$N\mathcal{I}$	Number of players	20–1000	-
Δt	Duration per period	0.25	h
$S_i^{\text{bat,init}}$	Initial battery state	0–1.92	kWh
$\bar{S}_{t,i}^{\text{bat}}$	Minimum battery level	0–1.824	kWh
$\bar{S}_{t,i}^{\text{bat}}$	Maximum battery level	0–9.6	kWh
$\bar{q}_{t,i}^{\text{ch}}$	Max charge rate	0–5	kWh
$\bar{q}_{t,i}^{\text{dch}}$	Max discharge rate	0–5	kWh
$\bar{q}_{t,i}^{\text{acq}}$	Max acquisition amount	4.6–13.8	kWh
$\bar{q}_{t,i}^{\text{disp}}$	Max dispatch amount	4.6–13.8	kWh
C^{fix}	Fixed operational costs	0.2197–0.6249	EUR
$q_{t,i}^{\text{gen}}$	Generated energy forecast	0–8.474	kWh
$q_{t,i}^{\text{load}}$	Consumed energy forecast	0.050–9.822	kWh
$\pi_{t,i}^{\text{acq}}$	Acquisition price	0.1034–0.2314	EUR/kWh

$\pi_{t,i}^{\text{disp}}$	Dispatch price	0.045	EUR/kWh
---------------------------	----------------	-------	---------

Incorporating these parameters ensures that the scheduling framework can accommodate a wide spectrum of realistic operating conditions. It enforces adherence to physical storage limits, grid interaction rules, and cost-related objectives, thereby enabling precise, context-aware optimization of energy flows, battery utilization, procurement, and sales.

Optimization outputs

The principal outcomes generated by the optimization procedure are summarized in [Table 3](#). These variables represent the optimal dispatch plan for the VPP, detailing the recommended energy purchases, sales, charging/discharging actions, and associated cost-minimizing behaviors over the scheduling horizon.

Table 3. Optimization outputs in vector notation.

Parameter	Description	Unit
$q_{t,i}^{\text{acq}}$	Energy amounts to be bought at time t	kWh
$q_{t,i}^{\text{disp}}$	Energy amounts to be sold at time t	kWh
$q_{t,i}^{\text{disp,non-paid}}$	Energy amounts to be discarded (non-paid)	kWh
$q_{t,i}^{\text{ch}}$	Energy is charged to the battery	kWh
$q_{t,i}^{\text{dch}}$	Energy is to be discharged from the battery	kWh
$q_{t,i}^{\text{exp}}$	Total energy amounts to be exported	kWh
$s_{t,i}^{\text{bat}}$	Resulting battery states at time t	kWh

Among the delivered results are the optimized volumes of grid acquisition and injection, battery power flows (charge and discharge), and state-of-charge trajectories. Together, these elements provide a comprehensive, cost-effective roadmap for energy management, enabling the VPP operator to meet load obligations, capitalize on market opportunities, and operate storage assets at the lowest possible cost while preserving system reliability and longevity.

Hyperparameter optimization

Fine-tuning the Simulated Annealing (SA) method for effective VPP scheduling demands meticulous adjustment of its core hyperparameters. In this work, we focused on optimizing four essential parameters—initial temperature, cooling rate, maximum number of iterations, and perturbation scale—to achieve an optimal trade-off between solution quality and computational resources. We began with a broad grid search to map general trends across wide parameter intervals, then proceeded to more precise tuning using sophisticated surrogate-based techniques, specifically Gaussian Process (GP), Random Forest (RF), and Gradient Boosted Regression Tree (GBRT) minimization.

These surrogate modeling approaches were chosen for their proven efficiency in navigating intricate, high-dimensional hyperparameter landscapes. Gaussian Process regression [44-46] excels at representing smooth response surfaces, enabling reliable identification of promising regions with relatively few function evaluations. Random Forest regression [47] was included for its robustness in large-scale or noisy search spaces and its ability to model complex, non-linear parameter interactions, which is especially valuable in the context of SA hyperparameter optimization. Gradient-Boosted Regression Trees [48, 49] were adopted for their iterative refinement and resilience across parameter spaces with varying levels of smoothness or dimensionality. Together, these complementary methods ensured a thorough and resource-

efficient exploration of the hyperparameter space, ultimately yielding a finely tuned SA configuration tailored to the demands of VPP scheduling.

Parameter ranges and search strategy

The investigated ranges for each hyperparameter were informed by established theoretical properties of Simulated Annealing as well as preliminary insights gained from the initial grid search:

- Initial temperature: Spanned from 50.0 to 400.0. This parameter governs the likelihood of accepting inferior moves during the early stages. Values closer to 400.0 promote extensive global exploration by frequently permitting worse solutions, thereby helping escape from poor local optima. As temperature declines, acceptance becomes more selective, progressively favoring improvement of high-quality candidate solutions.
- Cooling rate: Ranged from 0.85 to 0.975. The cooling schedule dictates how quickly the temperature is reduced. Slower cooling (near 0.85) sustains elevated temperatures over more iterations, encouraging prolonged and diverse exploration. Faster cooling (approaching 0.975) hastens the transition to exploitation, concentrating search effort on local refinement.
- Perturbation scale: Covered the interval from 0.025 to 0.2. This factor controls the size of random changes applied to the current solution at each step. Larger step sizes (around 0.2) support aggressive exploration in the initial phase, whereas smaller step sizes (near 0.025) enable precise local adjustments as the algorithm approaches convergence.

The preliminary grid search revealed key behavioral patterns, notably that combinations with higher initial temperatures and slower cooling rates generally enhance early-stage exploration. Building on these observations, we applied the three advanced surrogate optimization techniques—GP, RF, and GBRT—to identify refined hyperparameter sets. Each technique offers distinct advantages for tuning, with RF and GBRT proving especially adept at modeling nonlinear parameter dependencies.

Table 4 presents the final optimized hyperparameter configurations derived for the SA solver in the context of VPP scheduling. The table includes the specific settings recommended by each minimization method along with the corresponding performance scores. Among them, the Random Forest-based approach delivered the best result, achieving a score of 85.83, which highlights its particular effectiveness for this application.

Table 4. Optimized hyperparameter configurations for SA in VPP scheduling. Each method's performance score reflects solution quality under the respective parameter settings.

Optimization method	Optimal parameters	Best score
Gaussian process minimization	[400.0, 0.85, 0.077]	85.77
Random forest minimization	[189.44, 0.895, 0.065]	85.83
Gradient-boosted regression tree minimization	[228.89, 0.959, 0.072]	85.78

Higher scores correspond to superior optimization performance.

Analysis of results

The hyperparameter optimization procedure yielded important observations regarding the impact of each parameter on the Simulated Annealing (SA) solver's effectiveness when applied to VPP scheduling. Elevated initial temperature settings (near 400.0) consistently proved beneficial, as they facilitated wide-ranging exploration at the outset and helped the algorithm avoid suboptimal local optima during the initial stages. In contrast, adopting a slower cooling rate (around 0.85) supported a more measured descent toward convergence, preserving greater solution diversity over many iterations and thereby reducing the risk of early stagnation.

The perturbation scale stood out as a particularly influential factor. Values in the range of approximately 0.065–0.077 struck an effective compromise, delivering sufficient step sizes for meaningful global sampling while still permitting accurate local adjustments as the search matured. Although the maximum iteration limit exerted only modest influence—owing to the relatively constrained interval tested—it nonetheless provided sufficient depth to ensure consistent and stable convergence across the evaluated configurations.

The three surrogate optimization techniques displayed complementary advantages. Random Forest minimization achieved the highest performance score, presumably owing to its proficiency at modeling intricate, nonlinear dependencies among hyperparameters—an aspect especially relevant to SA's adaptive, stochastic nature. Gradient-Boosted Regression Trees likewise produced strong results, yielding a competitive parameter set characterized by a marginally higher cooling rate that promoted quicker exploitation during the later phases of the search.

Convergence phases in simulated annealing

Once the hyperparameters had been optimized within the VPP scheduling framework, the SA algorithm's convergence trajectory can be divided into three well-defined stages: initialization, exploration, and exploitation. These stages capture the methodical progression of the method, beginning with establishing a starting point, followed by broad sampling of the solution landscape, and culminating in focused refinement of high-quality candidates within the allowable bounds. This phased behavior underscores SA's capacity to harmonize expansive global search with precise local improvement, ultimately delivering dependable and high-quality optimization results for VPP dispatch problems.

Figure 4 depicts the typical convergence profile of the SA algorithm in VPP scheduling. The plot traces the algorithm's evolution across its principal phases, illustrating how it systematically traverses the search space and attains robust solutions while respecting the imposed computational limits.

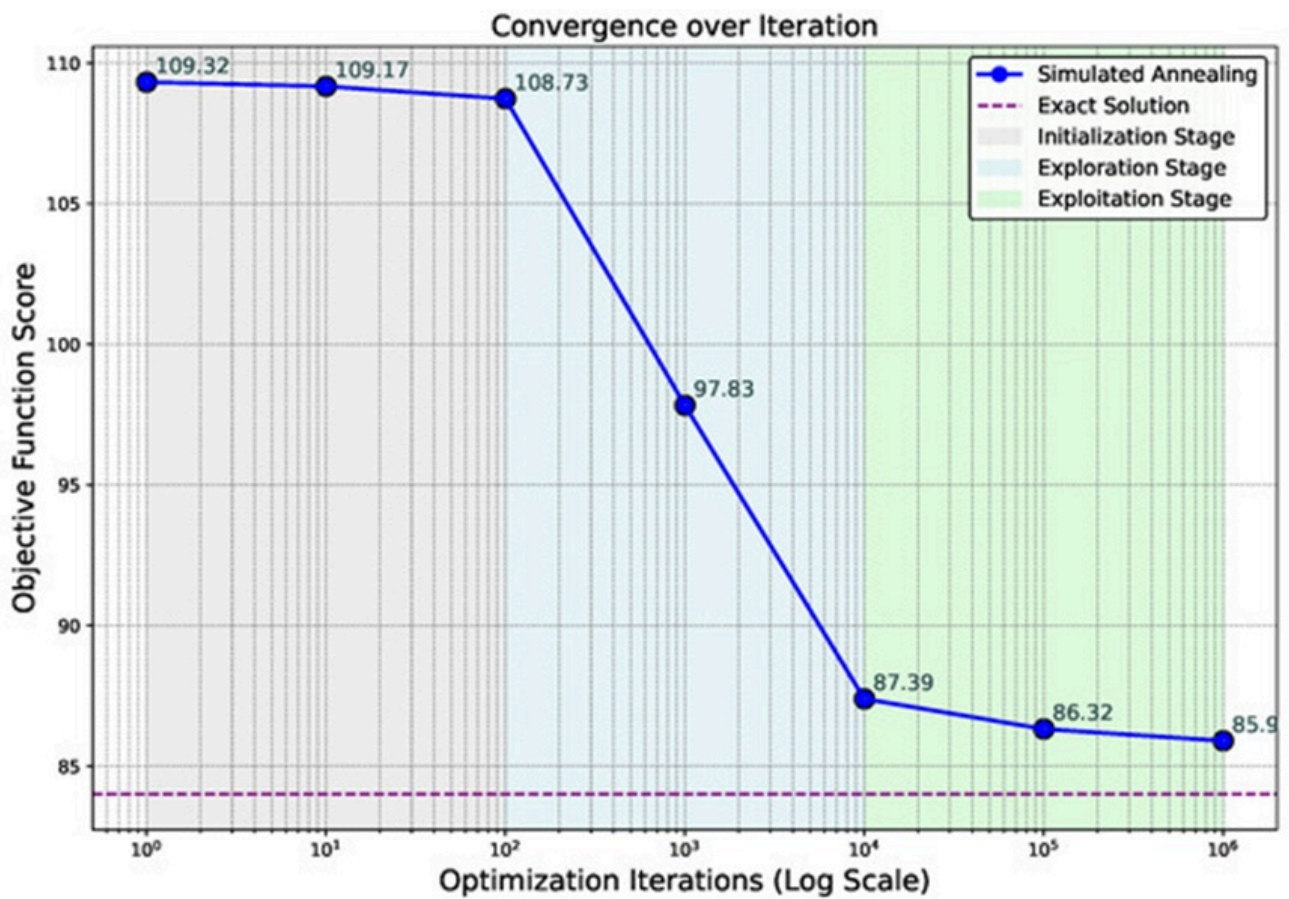


Figure 4. Convergence profile of the SA algorithm for the VPP scheduling task, illustrating the transition through initialization, exploration, and exploitation stages to reach high-quality solutions efficiently.

Parallelization and load balancing techniques

The architecture of the parallel Simulated Annealing solver developed for virtual power plant (VPP) optimization leverages OpenMP to achieve superior computational throughput and ensure equitable workload distribution across available cores. At its core, the parallelization scheme partitions the overall optimization workload into mutually independent subtasks, each dedicated to computing the optimal schedule for a single participant (player) in the VPP. This decomposition strategy minimizes synchronization costs, enhances overall throughput, and enables excellent scalability by allowing simultaneous processing of multiple players across distinct threads.

Thread-based parallelization with OpenMP

In the current implementation, OpenMP constructs are used to orchestrate multiple threads and fully exploit available processing capacity. The key parallel construct is a parallel for loop that uses the `#pragma omp parallel` for directive with static scheduling. This directive evenly allocates the set of players to the available threads at the start of execution. Because every player follows the same optimization procedure over the same number of time periods, the computational effort per task remains highly uniform, resulting in balanced workloads and negligible idle time across threads.

Static scheduling was preferred over dynamic alternatives due to the predictable and consistent nature of the per-player workload, which eliminates the need for runtime task reassignment and thereby reduces associated overhead. The outer parallel loop spans the collection of players, achieving substantial concurrency, while the nested inner loop within each player's optimization routine sequentially manages the temperature cooling schedule, solution perturbation, energy balance checks, constraint enforcement, and objective function evaluation. Since there are no data dependencies between players' optimization processes, inter-thread communication is avoided entirely. This independence maximizes the attainable level of parallelism, and the approaches are particularly well-suited for scaling to VPP instances with a large number of participating prosumers.

Unexplored load balancing strategies

Although the current implementation relies on OpenMP with static scheduling to parallelize the per-player optimization tasks, several alternative load-balancing methods were deliberately left unexplored. In particular, neither dynamic nor guided scheduling—both of which allocate tasks at runtime based on actual thread progress—was employed. Such runtime-adaptive techniques could, in theory, reassign workloads to prevent idle threads in the presence of computational heterogeneity. Nevertheless, they were considered unnecessary here, given the highly uniform nature of the workload: every player undergoes an identical optimization procedure with the same number of time steps and iterations. The simplicity and predictability of static scheduling, which fixes task-to-thread assignments upfront, proved fully adequate while eliminating the additional overhead that dynamic approaches inevitably incur.

Furthermore, more sophisticated load-balancing paradigms, such as hierarchical parallelism or fully task-based execution models, were not investigated. These strategies might offer advantages in scenarios characterized by highly variable task durations, significant inter-task dependencies, or opportunities for much finer-grained concurrency. In the present case, however, the problem exhibits a regular, homogeneous structure in which all players are computationally equivalent and completely independent, rendering such advanced mechanisms superfluous and potentially counterproductive due to added complexity.

Scalability and computational efficiency

The parallel architecture of the SA solver was rigorously evaluated across a spectrum of hardware configurations, specifically from 1 to 32 CPU cores. For each tested combination of core count and player count, we measured both the average execution time and the resulting speedup ratio to quantify scalability and validate the effectiveness of the chosen

parallelization and load distribution approach. The experimental outcomes demonstrated near-ideal linear speedup throughout the entire range up to 32 cores, with only minor departures from the theoretical maximum. This strong scaling behavior verifies that the parallel design is highly effective and delivers consistent performance improvements as the problem size—measured by the number of participating prosumers—increases.

As the number of players increases, the parallelization scheme effectively distributes the computational burden across all available cores, translating additional hardware resources into roughly proportional reductions in total runtime. Such predictable and efficient scaling is particularly valuable in practical VPP scheduling applications, where the system must respond swiftly to fluctuating market signals, renewable generation variability, and shifting load profiles. The close alignment with the ideal speedup, even at larger scales, confirms the robustness of the parallel SA solver, positioning it as a reliable solution for demanding, time-sensitive optimization tasks in expansive real-world energy systems.

Results and Discussion

The parallelized Simulated Annealing solver developed for virtual power plant (VPP) optimization was subjected to a comprehensive performance assessment in a high-performance computing (HPC) environment. The experimental campaign specifically targeted the interplay between the available computational cores and the scale of VPP aggregation (measured as the number of prosumers/players). This section reports key findings on execution time, speedup ratio, and overall scalability, and offers detailed observations on how the solver behaves across diverse combinations of core availability and problem size.

Execution time and speedup analysis

Execution time (averaged over multiple runs) and speedup ratio served as the principal performance indicators for the parallel SA solver. **Figure 5** presents the measured execution time as a function of core count for player aggregations ranging from 32 to 512, with each data point reflecting 1×10^6 iterations of the annealing process. The plots reveal a clear, consistent decline in runtime as more cores are used, underscoring the effectiveness of the OpenMP-based parallelization strategy across all tested scales.

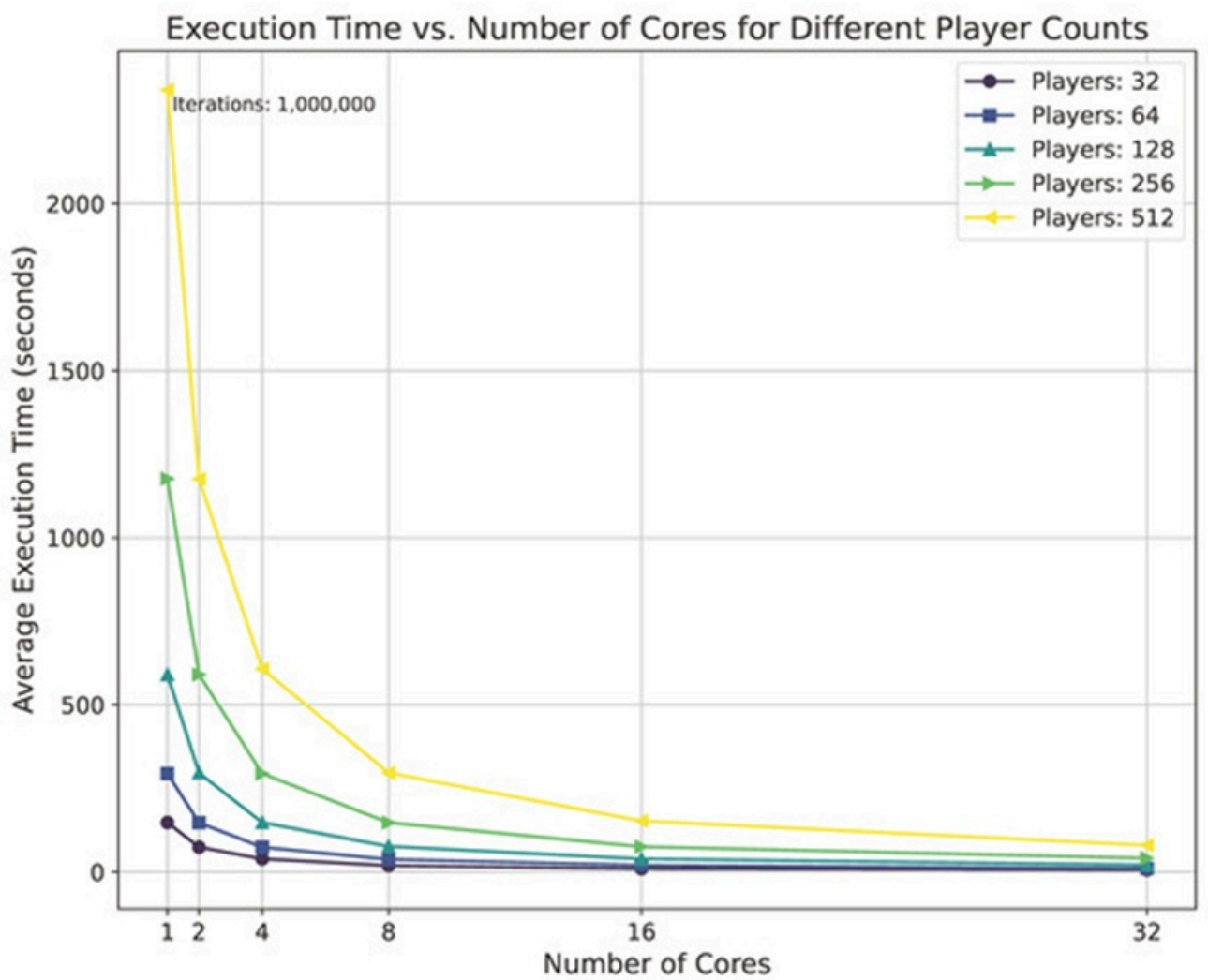


Figure 5. Execution time versus number of cores for varying player counts, measured using 1×10^6 iterations of simulated annealing.

The speedup ratio—computed as single-core execution time divided by multi-core execution time—provides a normalized measure of parallel efficiency. As depicted in [Figure 6](#), the observed speedup closely tracks the ideal linear reference line, particularly for larger configurations (e.g., 256 and 512 players), where the greater volume of independent tasks enables more effective workload distribution among threads. By relying on OpenMP’s default static scheduling, the implementation avoids dynamic overhead while ensuring predictable and balanced thread utilization. These characteristics highlight the strong scalability of the parallel SA framework, making it well-suited to address the growing computational requirements of increasingly complex VPP market environments.

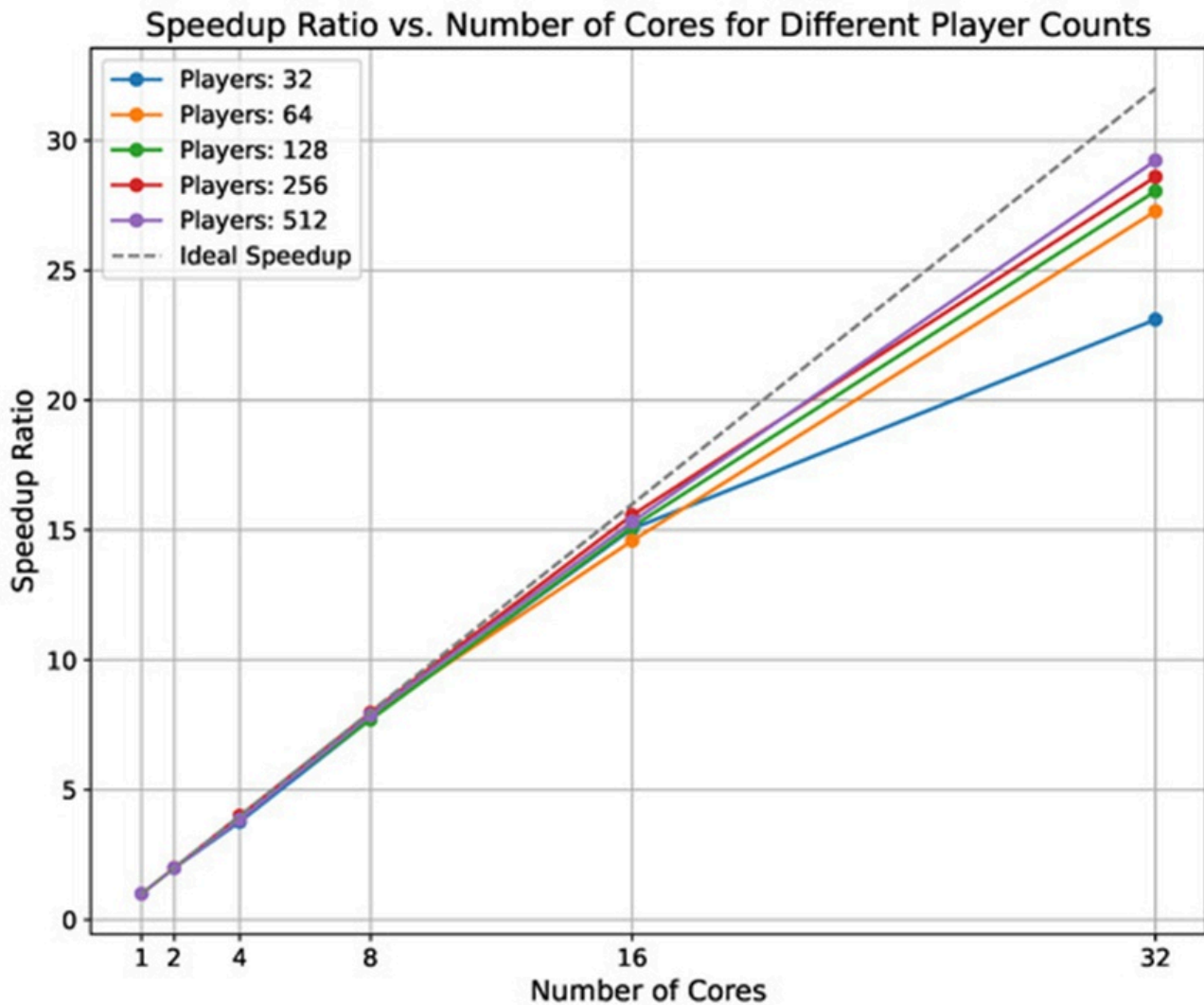


Figure 6. Speedup ratio versus number of cores for different player counts, illustrating near-linear scaling behavior and the benefits of effective parallel task distribution.

These findings highlight the real-world relevance of the solver's performance for extensive VPP scheduling applications. The steady decrease in execution time with increasing core count clearly illustrates the algorithm's ability to meet the stringent real-time processing demands of modern energy systems, where swift decision-making is essential for effective market engagement and optimal resource allocation. The close-to-linear speedup achieved in setups with 256 and 512 players further confirms the method's resilience to increasing problem complexity, positioning it as a strong candidate for volatile, high-volume energy markets.

Moreover, the outcomes underscore the substantial efficiency improvements delivered by the parallelization design. By using OpenMP's static scheduling—which assigns tasks to threads in advance—the solver eliminates much of the overhead typically associated with workload partitioning. This approach optimizes resource usage and mitigates potential bottlenecks, especially in scenarios featuring elevated player numbers. Maintaining such high efficiency is vital for delivering dependable performance across a broad range of VPP operating conditions, regardless of scale.

The speedup ratio's proximity to the theoretical linear benchmark reflects the even workload distribution realized by the parallel SA implementation. This indicates that the algorithm makes excellent use of available processing power, even

within challenging, high-dimensional search spaces. The consistent behavior observed over repeated executions further reinforces the solver's reliability and stability under diverse experimental conditions.

Scalability assessment

Scalability represents a key requirement for the parallelized SA solver, especially when addressing realistic large-scale VPP deployments. Ideal scalability manifests as a near-proportional increase in speedup with additional cores. The experimental data show that the parallel SA solver exhibits excellent scalability up to 32 cores, consistently delivering near-ideal speedups across all tested player configurations. Notably, as the number of players grows, scalability performance strengthens, with speedup profiles for larger aggregations (such as 256 or 512 players) aligning very closely with the ideal linear reference. This pattern confirms that the parallelized SA framework is particularly well-adapted to demanding VPP problems that necessitate significant computational capacity to preserve high efficiency.

The sustained efficiency gains observed as problem size expands demonstrate the robustness of the OpenMP-driven parallelization strategy when managing expansive VPP aggregations. Even in the most computationally intensive case of 512 players, speedup remained stable and close to optimal, illustrating the solver's ability to accommodate substantial numbers of interacting participants without appreciable degradation in parallel effectiveness.

Efficiency and resource utilization

Effective resource utilization constitutes another essential factor in high-performance computing contexts. Throughout the experiments, CPU and memory consumption were monitored to assess distribution across allocated cores. CPU utilization remained consistently high, with very little idle time, confirming that the parallelization scheme effectively balanced the computational load across threads. Memory footprint remained well within acceptable bounds, avoiding contention and ensuring overall system stability.

This proficient management of hardware resources underscores the suitability of the parallel SA solver for real-time VPP optimization tasks. Its ability to scale smoothly across different core counts and player volumes underscores its versatility within HPC infrastructures, establishing it as a powerful and adaptable solution for coordinating large-scale virtual power plants.

Quality of optimization outcomes

Preserving solution quality during parallel execution is a fundamental concern in any distributed optimization framework. To verify that the parallel implementation did not degrade the quality of the results, we systematically compared the solutions produced by the parallel SA solver with those from a purely sequential SA run. The comparison revealed that the parallel version consistently achieved solution qualities equivalent to—or extremely close to—those of the sequential baseline, regardless of the number of cores or players involved. This equivalence substantiates the validity of the parallelization strategy, confirming that it significantly accelerates computation while fully retaining the algorithm's effectiveness in thoroughly exploring the solution space and identifying high-quality (near-)optimal schedules.

Summary of findings

The experimental outcomes confirm that the parallelized Simulated Annealing solver tailored for VPP scheduling delivers substantial improvements in computational performance without any degradation in solution accuracy. The principal conclusions can be summarized as follows:

- Near-ideal speedup ratio: The speedup closely tracked the theoretical linear reference, especially in configurations with larger numbers of players, signifying excellent scalability.

- High efficiency and optimal resource utilization: CPU and memory usage remained consistently high and balanced, ensuring reliable and stable operation across diverse experimental setups.
- Consistent solution quality: The parallel implementation preserved optimization performance, producing results that matched or very closely approximated those obtained from the sequential baseline.

Taken together, these results highlight the parallel SA solver's ability to effectively address the intricate and expansive scheduling requirements of virtual power plant optimization, establishing it as a dependable, high-performing option for demanding HPC-based applications.

Application cases

The developed parallelized SA solver exhibits strong practical relevance for real-world VPP management. Its proven scalability and rapid processing make it especially appropriate for overseeing extensive VPP networks, where numerous prosumers and distributed energy assets must be coordinated under strict time constraints. By efficiently resolving complex, high-dimensional optimization tasks, the solver empowers VPP operators to reduce overall operating expenses, maintain equilibrium between supply and demand, and strengthen their engagement in fast-changing electricity markets. Moreover, its flexibility in responding to price and load variations supports refined resource allocation approaches, encompassing dynamic pricing mechanisms and demand-side management initiatives, all while promoting stable and economical performance. The solver is also capable of managing the complexities of multi-energy systems within VPP frameworks—such as synchronizing power flows, battery storage, and grid-supporting services—thereby boosting system-wide efficiency and environmental performance. These attributes establish the parallelized SA solver as a robust, trustworthy tool for meeting the intensive computational demands of contemporary VPP infrastructures and delivering effective solutions to the challenges of real-time energy coordination.

Practical implications for the VPP market

The present research carries important consequences for grid managers, VPP administrators, and regulatory authorities, extending beyond purely technical deployment scenarios to encompass strategic, operational, and policy-oriented dimensions of implementing the proposed parallelized SA solver in actual energy systems.

For grid operators, the solver's demonstrated scalability and processing efficiency offer a valuable platform for supervising expansive, heterogeneous energy networks involving many distributed participants. Real-time optimization capabilities enable operators to adapt swiftly to shifting market dynamics, seamlessly incorporate variable renewable generation, and uphold system reliability amid fluctuating consumption and production profiles. These features are vital for overcoming operational challenges posed by the growing presence of decentralized energy resources (DERs), including solar photovoltaic systems, wind turbines, and battery storage installations.

For VPP operators, the solver's proficiency in tackling large-scale, multi-variable optimization problems facilitates more effective asset management, thereby improving market participation opportunities. This encompasses implementing sophisticated tariff structures, optimizing the scheduling of energy injections and withdrawals, and effectively executing demand response schemes. In addition, the parallel SA solver enables operators to increase revenue potential in competitive marketplaces while preserving operational robustness. Its suitability for integrated multi-energy environments further promotes coordinated handling of electricity, storage capacity, and ancillary services, contributing to more holistic and sustainable operational strategies.

From a policymaker's perspective, the results underscore the strategic importance of deploying cutting-edge optimization algorithms and high-performance computing resources to advance overarching energy transition objectives. This work can inform advocacy for greater investment in scalable computational platforms that facilitate widespread VPP participation in markets while simultaneously addressing priorities such as grid decarbonization, enhanced system resilience, and equitable energy access. By championing these technological advancements, policymakers can cultivate

an environment that accelerates renewable energy adoption, supports carbon-neutral pathways, and aligns with international sustainability commitments.

Furthermore, the solver's seamless scaling with rising prosumer numbers points toward viable strategies for modernizing electricity markets to accommodate increasing levels of complexity and participation. This progression supports key long-term energy policy aims, including greater operational flexibility, diminished reliance on conventional fossil-based generation, and heightened adaptability to price and supply volatility. As electricity markets continue to transform, the evidence presented here can help shape regulatory frameworks and incentive programs that encourage the widespread adoption of powerful, scalable computational tools that balance economic viability with environmental responsibility.

Conclusion

This research has effectively demonstrated the value of a parallelized Simulated Annealing (SA) approach enabled by OpenMP for enhancing the computational performance and real-time responsiveness of virtual power plant (VPP) scheduling. The incorporation of OpenMP to distribute SA computations across multiple threads resulted in significant reductions in processing time, enabling the solver to manage expansive VPP systems with high efficiency. Experimental findings confirmed that, with judicious hyperparameter tuning and an effective parallelization design, the SA solver exhibited rapid convergence to near-optimal schedules, delivering marked improvements in both execution duration and speedup ratio over a wide range of player counts and core configurations.

Performance assessments revealed that the parallel SA implementation sustained excellent scalability, with speedup ratios remaining very close to the theoretical linear ideal—particularly in scenarios involving larger numbers of players. This outcome highlights the transformative role of parallel computing in improving the scalability, flexibility, and practical viability of VPP operations, marking a meaningful step forward in energy system optimization techniques. The results further confirmed that parallel execution preserved solution quality, as the obtained schedules matched or closely approximated those produced by the sequential SA baseline. Achieving this equilibrium between accelerated computation and reliable optimization outcomes is fundamental to deploying SA successfully in dynamic, large-scale VPP scheduling environments.

In addition, the study demonstrates how the developed optimization framework supports overarching goals in sustainable urban energy systems. By tackling critical issues such as real-time decision-making, battery degradation constraints, and time-varying electricity tariffs, the work strengthens VPPs' position as core elements of smart city ecosystems. These contributions promote distributed energy governance, bolster grid reliability, and facilitate the smooth integration of renewable generation and electric vehicle fleets—factors indispensable to achieving the energy efficiency and environmental sustainability objectives of future urban environments.

Future research will focus on further optimizing the parallelization strategy to minimize remaining overhead and extend scalability. Potential directions include leveraging more sophisticated OpenMP constructs, experimenting with dynamic and guided scheduling policies, and possibly adopting complementary parallel programming paradigms to achieve superior performance in extremely large VPP aggregations. Another key focus will be expanding the parallelized SA solver to encompass more sophisticated and realistic VPP formulations that account for live market signals, unpredictable demand variations, and interactions across multiple energy carriers. Developing advanced simulation platforms capable of faithfully reproducing evolving smart city conditions will also be prioritized, ensuring that proposed methods remain relevant and robust in rapidly changing urban energy contexts. Moreover, rigorous stress tests involving substantially greater numbers of players will be conducted to probe the limits of scalability and structural resilience under extreme operational loads.

A comprehensive benchmarking effort is planned to compare the performance of the VPP optimization problem when addressed by different metaheuristic techniques, including parallelized versions of Genetic Algorithms (GA) and Particle Swarm Optimization (PSO). These investigations will leverage both CPU- and GPU-based computing environments to

assess the advantages of cutting-edge parallel and distributed architectures. By juxtaposing the proposed parallel metaheuristic solvers with established commercial exact solvers—including Gurobi, CPLEX, and MOSEK—the analysis will provide a thorough appraisal of scalability, computational throughput, and attained solution accuracy. This comparative study will illuminate the distinctive advantages, limitations, and performance trade-offs inherent in each approach, while also evaluating their suitability and real-world potential for tackling extensive, practical energy system challenges.

Acknowledgements

The authors gratefully acknowledge the support of the high-performance computing infrastructure provided through the project “High-Performance Computing for Enhanced Virtual Power Plant Efficiency” (reference 2024.00036.CPCA.A1). This infrastructure has been instrumental in facilitating the computational tasks required for this research. The project has been assigned the DOI <https://doi.org/10.54499/2024.00036.CPCA.A1>, accessed on 27 February 2025.

Conflict of interest

None

Financial support

The European Union funded this research through the Recovery and Resilience Plan (PRR) as part of the Innovation Pact “NGS—New Generation Storage” (reference 58). This initiative is co-financed by NextGenerationEU under the Incentive System “Agendas para a Inovação Empresarial” (“Agendas for Business Innovation”).

Ethics statement

None

Received: 18 Sep 2023

Revised: 25 Oct 2023

Accepted: 02 Dec 2023

Published online: 20 March 2024

Rights and permissions

Open Access The author(s) retain copyright. This article is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License. It may be shared and adapted for non-commercial purposes with appropriate attribution, an indication of changes, and distribution of adaptations under the same license. Third-party material may be subject to separate terms identified in its credit line. View the license at <https://creativecommons.org/licenses/by-nc-sa/4.0/>.

References

Podder AK, Islam S, Kumar NM, Chand AA, Rao PN, Prasad KA, et al. Systematic categorization of optimization strategies for virtual power plants. *Energies*. 2020;13:6251.

Nafkha-Tayari W, Ben Elghali S, Heydarian-Forushani E, Benbouzid M. Virtual power plants optimization issue: A comprehensive review on methods, solutions, and prospects. *Energies*. 2022;15:3607.

Venegas-Zarama JF, Muñoz-Hernandez JI, Baringo L, Diaz-Cachinero P, De Domingo-Mondejar I. A review of the evolution and main roles of virtual power plants as key stakeholders in power systems. *IEEE Access*. 2022;10:47937–64.

Ali JUABW, Kazmi SAA, Altamimi A, Khan ZA, Alrumayh O, Malik MM. Smart energy management in virtual power plant paradigm with a new improved multilevel optimization based approach. *IEEE Access*. 2022;10:50062–77.

Xie Y, Zhang Y, Lee WJ, Lin Z, Shamash YA. Virtual power plants for grid resilience: a concise overview of research and applications. *IEEE/CAA J Autom Sin*. 2024;11:329–43.

Roy S, Das DC, Sinha N. Optimizing smart city virtual power plants with V2G integration for improved grid resilience. In: *Proceedings of the 2024 IEEE International Conference on Interdisciplinary Approaches in Technology and Management for Social Innovation (IATMSI)*, Gwalior, India; 2024.

Ou M, Huang G, Chen S. Design of a smart light device in virtual power plants for new energy regulation. In: *Proceedings of the 2023 4th International Conference on Advanced Electrical and Energy Systems (AEES)*, Shanghai, China; 2023.

Zhang J, Seet B-C, Lie TT. Dynamic energy scheduling for virtual power plant with prosumer resources using game theory. In: *Proceedings of the 2019 IEEE Power and Energy Society General Meeting (PESGM)*, Atlanta, GA, USA; 2019. p. 1–5.

Sousa T, Pinto T, Morais H, Vale Z. Intelligent electric vehicle heuristic for energy resource management using simulated annealing. In: *Proceedings of the 2012 3rd IEEE PES Innovative Smart Grid Technologies Europe (ISGT Europe)*, Berlin, Germany; 2012. p. 1–8.

Venegas-Zarama JF, Muñoz-Hernandez JI. A general description of virtual power plants as smart manager in power systems. In: *Proceedings of the 2022 IEEE European Technology and Engineering Management Summit (E-TEMS)*, Bilbao, Spain; 2022. p. 83–7.

Pereira A, Rodrigues N, Barbosa J, Leitão P. Trust and risk management towards resilient large-scale cyber-physical systems. In: *Proceedings of the 2013 IEEE International Symposium on Industrial Electronics (ISIE)*, Taipei, Taiwan; 2013. p. 1–6.

Filipe Alves, Rui Ribeiro, Maria Petiz, Ali Abbasi, Pedro Carvalho, Ricardo Rodrigues, et al. Virtual power plant optimization service—benchmark of solvers. In: *Proceedings of the 19th Conference on Computer Science and Intelligence Systems (FedCSIS)*, Belgrade, Serbia; 2024. p. 279–87.

Alves F, Petiz M, Ribeiro R, Abbasi A, Carvalho PN, Rodrigues R. An N-Queens benchmark using MILP solvers: Comparison between open-source optimization tools. In: *Optimization, Learning Algorithms and Applications (OL2A 2024)*, Communications in Computer and Information Science. Springer, Cham; 2024;2281.

Aoun A, Adda M, Ilinca A, Ghandour M, Ibrahim H. Optimizing virtual power plant management: a novel MILP algorithm to minimize levelized cost of energy, technical losses, and greenhouse gas emissions. *Energies*. 2024;17:4075.

Lubin M, Petra CG, Anitescu M, Zavala V. Scalable stochastic optimization of complex energy systems. In: *Proceedings of the 2011 International Conference for High Performance Computing, Networking, Storage and Analysis (SC)*, Seattle, WA, USA; 2011. p. 1–64.

Sousa T, Morais H, Vale Z, Faria P, Soares J. Intelligent energy resource management considering vehicle-to-grid: a simulated annealing approach. *IEEE Trans Smart Grid*. 2012;3:535–42.

Sousa T, Vale Z, Morais H. Simulated annealing approach applied to the energy resource management considering demand response for electric vehicles. In: *Proceedings of the 17th International Conference on Intelligent System Applications to Power Systems (ISAP)*, Tokyo, Japan; 2013. p. 1–6.

- Faia R, Pinto T, Vale Z. Optimization of electricity markets participation with simulated annealing. In: Proceedings of the International Conference on Practical Applications of Agents and Multi-Agent Systems (PAAMS), Sevilla, Spain; Springer, Cham;2016. p. 27–39.
- Bouddou R, Benhamida F, Haba M, Belgacem M, Meziane MA. Simulated annealing algorithm for dynamic economic dispatch problem in the electricity market incorporating wind energy. *Ing Syst Inf*. 2020;25:719–27.
- Hannan MA, Abdolrasol MGM, Faisal M, Ker PJ, Begum RA, Hussain A, et al. Binary particle swarm optimization for scheduling MG integrated virtual power plant toward energy saving. *IEEE Access*. 2019;7:107937–51.
- Al-Shafei A, Zareipour H, Cao Y. A review of high-performance computing and parallel techniques applied to power systems optimization. *arXiv*. 2022;arXiv:2207.02388.
- Petra CG, Schenk O, Anitescu M. Real-time stochastic optimization of complex energy systems on high-performance computers. *Comput Sci Eng*. 2014;16:32–42.
- Angulo A, Rodríguez D, Garzón W, Gómez DF, Al Sumaiti A, Rivera S. Algorithms for bidding strategies in local energy markets: Exhaustive search through parallel computing and metaheuristic optimization. *Algorithms*. 2021;14:269.
- Alvarez D, Rodriguez D, Rivera S. Parallel computing for reducing time in security constrained optimal power flow analysis. *Rev Int Metod Numer Calc Diseño Ing*. 2023;39:1–5.
- Epperly T, Edmunds T, Lamont A, Meyers C, Smith S, Yao Y, et al. High-performance computing for electric grid planning and operations. In: Proceedings of the 2012 IEEE Power and Energy Society General Meeting, San Diego, CA, USA; 2012.
- Caramanis M, Ntakou E, Hogan WW, Chakraborty A, Schoene J. Co-optimization of power and reserves in dynamic T&D power markets with nondispatchable renewable generation and distributed energy resources. *Proc IEEE*. 2016;104:807–36.
- Khan R, Islam N, Das SK, Muyeen SM, Muyeen SI, Ali MF, et al. Energy sustainability—survey on technology and control of microgrid, smart grid, and virtual power plant. *IEEE Access*. 2021;9:104663–94.
- Sarmiento-Vintimilla JC, Torres E, Larruskain DM, Pérez-Molina MJ. Applications, operational architectures and development of virtual power plants as a strategy to facilitate the integration of distributed energy resources. *Energies*. 2022;15:775.
- Krishna R, Hemamalini S. Optimal energy management of virtual power plants with storage devices using teaching-and-learning-based optimization algorithm. *Int Trans Electr Energy Syst*. 2022;2022:1727524.
- Muttaqi KM, Sutanto D. Adaptive and predictive energy management strategy for real-time optimal power dispatch from VPPs integrated with renewable energy and energy storage. *IEEE Trans Ind Appl*. 2021;57:1958–72.
- Naughton J, Wang H, Cantoni M, Mancarella P. Co-optimizing virtual power plant services under uncertainty: a robust scheduling and receding horizon dispatch approach. *IEEE Trans Power Syst*. 2021;36:3960–72.
- Yi Z, Xu Y, Gu W, Wu W. A multi-time-scale economic scheduling strategy for virtual power plant based on deferrable loads aggregation and disaggregation. *IEEE Trans Sustain Energy*. 2019;11:1332–46.
- Tang W, Yang H-T. Optimal operation and bidding strategy of a virtual power plant integrated with energy storage systems and elasticity demand response. *IEEE Access*. 2019;7:79798–809.
- Lezama F, Faia R, Soares J, Faria P, Vale Z. Learning bidding strategies in local electricity markets using ant colony optimization. In: Proceedings of the 2020 IEEE Congress on Evolutionary Computation (CEC), Glasgow, UK; 2020. p. 1–8.
- Mousavi SH, Nazemi A, Hafezalkotob A. Using and comparing metaheuristic algorithms for optimizing bidding strategy viewpoint of profit maximization of generators. *J Ind Eng Int*. 2015;11:59–72.

Jain P, Saxena A. Profit maximization bidding strategy for a GENCO using whale optimization algorithm. In: Proceedings of the 2019 Women Institute of Technology Conference on Electrical and Computer Engineering (WITCON ECE), Dehradun, India; 2019. p. 100–3.

Rädle S, Mast J, Gerlach J, Bringmann O. Computational intelligence based optimization of hierarchical virtual power plants. *Energy Syst.* 2021;12:517–44.

Peik-Herfeh M, Seifi H, Sheikh-El-Eslami MK. Decision making of a virtual power plant under uncertainties for bidding in a day-ahead market using point estimate method. *Int J Electr Power Energy Syst.* 2013;44:88–98.

Henderson D, Jacobson SH, Johnson AW. The theory and practice of simulated annealing. In: Glover F, Kochenberger GA, editors. *Handbook of Metaheuristics*. International Series in Operations Research and Management Science. Boston, MA: Springer; 2003; 57:287–319.

Janaki Ram D, Sreenivas TH, Ganapathy Subramaniam K. Parallel simulated annealing algorithms. *J Parallel Distrib Comput.* 1996;37:207–12.

Ferreiro AM, García JA, López-Salas JG, Vázquez C. An efficient implementation of parallel simulated annealing algorithm in GPUs. *J Glob Optim.* 2013;57:863–90.

Chen D-J, Lee C-Y, Park C-H, Mendes P. Parallelizing simulated annealing algorithms based on high-performance computer. *J Glob Optim.* 2007;39:261–89.

Greening DR. Parallel simulated annealing techniques. *Comput Oper Res.* 1990;18:41–9.

Osborne MA, Garnett R, Roberts SJ. Gaussian processes for global optimization. Technical report. University of Oxford, Oxford, UK; 2009.

Lamparth M, Bestehorn M, Märkisch B. Gaussian Processes and Bayesian Optimization for High Precision Experiments. *arXiv* 2022, *arXiv:2205.07625*.

Xu Z, Wang H, Phillips JM, Zhe S. Standard Gaussian Process Can Be Excellent for High-Dimensional Bayesian Optimization. *arXiv* 2024, *arXiv:2402.02746*.

Hutter F, Hoos HH, Leyton-Brown K. Sequential Model-Based Optimization for General Algorithm Configuration. In *Learning and Intelligent Optimization: 5th International Conference, LION 5, Rome, Italy; 2011. Selected Papers 5*; Springer: Berlin/Heidelberg, Germany; 2011. p. 507–23.

Bentéjac C, Csörgő A, Martínez-Muñoz G. A comparative analysis of gradient boosting algorithms. *Artif Intell Rev.* 2021;54:1937–67.

Jiang J, Cui B, Zhang C, Fu F. DimBoost: boosting gradient boosting decision tree to higher dimensions. In: Proceedings of the 2018 International Conference on Management of Data (SIGMOD '18), Houston, TX, USA; 2018. p. 1363–76.