

POLITECNICO
MILANO 1863

SCUOLA DI INGEGNERIA INDUSTRIALE
E DELL'INFORMAZIONE

Literature Review for Explicit Ancillary Service Provision in Multi-Energy Systems: Toward a Multi-Vector Reliability Framework

Tesi di Laurea Magistrale in
Energy Engineering - Ingegneria Energetica

Omer Eltayeb Khogali Mustafa, 250917

Advisor:
Prof. Filipo Bovira

Co-advisors:
Andrea Scrocca

Academic year:
2025-2026

Abstract: This study presents a systematic literature review of explicit ancillary service provision in Multi Energy Systems (MES), motivated by the increasing penetration of inverter-based, stochastic resources and the corresponding decline of rotational inertia in modern power systems. The literature is classified along four dimensions energy carrier (electricity, heat, mobility, and hydrogen), ancillary service type (frequency-related and non-frequency services), optimization technique (mixed-integer linear programming (MILP), two-stage stochastic programming, robust optimization), and market participation context (day-ahead and real-time markets). Twenty two studies are analyzed to identify recurring methodological patterns, including the prevalence of deterministic optimization for infrastructure sizing and the growing adoption of uncertainty aware frameworks for reliability constrained bidding, and to show how flexibility pooling across energy carriers can mitigate the physical limits of single-vector systems. Building on this synthesis, five recurring research gaps are identified: the limited treatment of financial risk in non-compliance scenarios, the persistence of carrier isolated optimization despite the conceptual promise of cross-vector synergy, the reliance on linear approximations for inherently non-linear degradation processes, the neglect of local distribution level constraints in service provision that is conventionally modeled at the transmission level, and the absence of a universal reliability metric capable of comparing the heterogeneous performance of electrical, thermal, and mobility-based assets. In response, the thesis proposes a corresponding research agenda centered on risk aware optimization, deeper cross-vector coordination, high fidelity physical and regulatory modeling, and the standardization of a universal reliability metric, to guide future work toward bankable, reliability aware explicit ancillary service provision in MES.

Key-words: Multi Energy System, Explicit Provision, Implicit Provision, MILP, Uncertainty-Aware Frameworks

1. Introduction

The global energy landscape is undergoing a radical decarbonization process, shifting from a centralized model based on fossil fuel combustion to a decentralized paradigm dominated by Inverter Based Resources (IBRs), such as solar photovoltaics (PV) and wind energy. This transition is driven by the urgent need to mitigate

climate change, yet the integration of high shares of Renewable Energy Sources (RES) introduces fundamental structural changes to the electrical grid's physical layer. Power systems relied on massive synchronous generators that provided inherent rotational inertia, which dampened frequency deviations during disturbances through the kinetic energy stored in their rotating masses.

The displacement of these conventional units by non-synchronous, power electronic interfaces alters the system's physics, by reducing the total system inertia a critical component that has historically provided the grid with a physical buffer against sudden frequency excursions [27]. Consequently, the Rate of Change of Frequency (RoCoF) following a contingency event has increased, leaving Transmission System Operators (TSOs) with a much narrower window for corrective action [34]. As traditional ancillary service providers exit the market, the system experiences a decline in the availability of "must run" units, potentially leading to critical shortages in grid support capability and market liquidity [13].

When the grid switches to mostly RES, it becomes a complicated place to work because weather dependent generation is always changing and planning is always uncertain. Energy markets often clear at a lead time (t^{lead}) that is much longer than the actual delivery time (t^{d}). This means that the difference between the day ahead forecast and the real time realization can make it hard to get the scheduled reserve stag. This uncertainty is especially important when it comes to providing power reserve capacity, where reliability standards are usually higher than (99%). If you do not take these factors into account such as changes in solar irradiance, wind speed errors, and changes in ambient temperature you might think you have more flexibility than you really do. Consequently, the system operator may fail to satisfy mandatory grid reliability standards and remain unable to account for price volatility and cost uncertainty during the market bidding stage[35].

Additionally, new assessment frameworks must take into account the operational limits of distribution networks such as voltage levels and branch flow limits to make sure that flexibility is still possible in all realization scenarios. Because of these limits and the fact that Distributed Energy Resources (DERs) are always changing we need to move from effort flexibility to reliability adjusted flexibility. In this new system, the maximum number of products that can be sold is based on the target failure probability. In modern power networks, the grid can keep supply and demand in balance even when RES are very unstable and go offline without warning. The main technical and financial tools that TSOs use to make sure this security are ancillary services. As traditional providers leave the market to make it more liquid and lower costs through more competition, it becomes technically necessary to connect DERs to ancillary service markets through Virtual Power Plants (VPPs). These services are now the essential pillars that enable the grid to accommodate high penetrations of renewable energy while closely adhering to technical requirements regarding duration, ramp time, and reliability. They are no longer only supplemental. The Feasible Operating Region (FOR) which denotes the set of potential power exchanges between a cluster of resources and the upstream network at the Point of Common Coupling (PCC) is frequently used to quantify these services. The modern VPP manager must effectively characterize this FOR while taking into account the different factors that threaten reserve availability, such as forecast uncertainties and technical DER limitations.

While the services described above remain the primary technical mechanisms for maintaining frequency and voltage stability, relying on electrical assets alone to deliver them is becoming increasingly strained as renewable penetration grows. The IEA projects needs short term flexibility, driven mainly by the rapid expansion of solar PV, are set to roughly triple between 2022 and 2030 in fast growing power systems such as China's [17]. Battery storage has absorbed much of this pressure, with utility scale battery capacity relative to peak demand rising sharply in markets such as California, from under 5% in 2019 to close to 25% in 2024 [18], however batteries remain constrained by limited discharge duration, temperature dependent derating, and the opportunity cost of reserving capacity for ancillary services rather than energy delivery. At the same time on the demand side flexibility despite carrying a technical potential running into the hundreds of gigawatts from sectors such as aluminum production and residential air conditioning alone remains largely untapped with only around 100 GW utilized globally as of 2024 [18].

This growing gap has pushed both researchers and international agencies toward sector coupling the deliberate linking of the power system to the heat, gas, and mobility sectors through technologies such as heat pumps, power to gas, and electric vehicle charging, as a larger source of flexibility than the electricity sector can provide on its own [19]. IRENA's analysis identifies non electrical sectors as carrying substantial untapped flexibility potential precisely because they possess forms of storage, thermal mass in buildings and district heating networks, chemical storage in hydrogen, and the mobile battery capacity of Electric Vehicle (EV) fleets, that are inherently cheaper than dedicated electrochemical storage and do not require new capital intensive electrical infrastructure [20]. Heat pumps and electric boilers, for instance, can absorb renewable surplus that would otherwise be curtailed, while hydrogen electrolysers can do the same for industrial users with few other low

carbon options [20].

This is the structural motivation behind treating MESs, rather than the electrical grid in isolation, as the natural unit of analysis for flexibility going forward. However, coupling sectors together does not by itself guarantee that this flexibility can be relied upon by a system operator, a heat pump shifting its load to avoid a high electricity price is fundamentally different, in terms of certainty and contractual obligation, from a heat pump committed to respond to a system operator's signal within a defined timescale. The distinction between these two modes, flexibility that emerges from cost optimization and flexibility that is contracted and dispatched, is the subject of the following section.

1.1. Ancillary Services: Definition and Operational Role

To keep the electrical grid stable, reliable, and of high quality, TSOs and DSOs depend on ancillary services as their primary technical instruments for grid management. As IBRs have proliferated across modern power systems, these services have transitioned from supplementary support mechanisms to indispensable pillars of system stability. Their fundamental objective is the continuous real time balancing of supply and demand, maintaining grid frequency and voltage within stringent operational limits regardless of the inherent variability of renewable generation. In this context, the controllable flexibility of distributed energy resources represents the critical enabler through which high shares of low carbon generation can be accommodated without compromising the physical security standards historically provided by synchronous generation.

1.1.1 Classification and Time-Scale Differentiation

The operational stability of a modern power system relies on real time balancing between generation and demand. Any mismatch between these two variables results in a deviation of the system frequency (f) from its nominal value 50 Hz. The technical classification of frequency related ancillary services is organized around the physical parameters they regulate and the specific timescales on which they operate, and building a defense strategy for the power system.

Frequency Containment Reserves

Frequency containment reserves (FCR) also referred to as primary regulation, activate automatically within seconds (0–30s) based on local frequency measurements to stabilize the grid by arresting frequency decay. As shown in Table 1, this explicit service requires the system to maintain a specific amount of power in reserve, ready to be deployed according to a pre defined droop characteristic. The provision of FCR follows a continuous, closed loop control logic that unfolds across four sequential stages.

The process begins with frequency sensing, in which high precision sensors at the PCC continuously monitor the grid frequency. Any deviation from the nominal 50 Hz immediately triggers the control response. Once the frequency crosses the deadband threshold, the second stage activates the local controller calculates the required power adjustment using the droop formula expressed in Equation 1, translating the measured frequency deviation directly into a corrective power setpoint.

$$\Delta P = -\frac{1}{s} \cdot (f - f_{nom}) \tag{1}$$

The third stage is the physical power modulation, whose implementation differs depending on the asset type. In Battery Energy Storage System (BESS) and EV systems, the inverter instantly modulates active power injection if the frequency falls below 50 Hz the battery discharges to support the grid, while if it rises above 50 Hz the battery charges to absorb excess energy. In heat pump systems, the controller instead modulates the compressor speed when frequency drops and upward reserve is required, the heat pump reduces its electricity consumption, effectively releasing that power back to the grid without interrupting its thermal function.

The final stage is frequency stabilization at the nadir. The objective at this point is to arrest the frequency decline before it breaches operational safety limits. The FCR collectively provides the power-frequency support required to reach the frequency nadir and stabilize the system within the 0 to 30 second response window, at which point the secondary regulation layer takes over to restore frequency to its nominal value.

Frequency Restoration Reserves

Frequency restoration reserves (aFRR) also referred to as secondary regulation, operate through a centralized control loop that contrasts fundamentally with the local, automatic nature of FCR. Where FCR arrests the frequency deviation, aFRR completes the recovery process by returning the system frequency to its exact nominal value of 50 Hz and replenishing the primary reserves so they are available for the next potential disturbance, as shown in Table 1. The provision of aFRR is an explicit service in which the MES must continuously track a time varying regulation signal dispatched by the TSO, making it technically more demanding than FCR since it requires constant communication infrastructure and high precision setpoint following rather than purely local autonomous response [12].

The process is initiated at the TSO level through centralized error detection. The TSO continuously calculates the Area Control Error (ACE), which accounts simultaneously for the residual frequency deviation and for any imbalances in scheduled power exchange between neighboring control areas [12]. Once the ACE exceeds the activation threshold, the TSO's automatic generation control system transmits regulation signal, typically updated every two to four seconds, directly to the MES aggregator or plant controller, translating the system level error into an asset level power setpoint [34].

Upon receiving this signal, the assets must modulate their output with sufficient precision and duration to track the requested setpoint throughout the activation window. For BESS and EV systems, this requires adjusting active power output in continuous correspondence with the evolving signal since aFRR activation can persist for up to fifteen minutes under standard european balancing rules [12], sufficient energy capacity must be reserved in the state of charge (SOC) to sustain the response without premature depletion. For heat based systems, including CHP units and large scale heat pumps, the thermal inertia of the building or district heating network acts as a natural buffer that decouples the grid obligation from the immediate thermal requirement, while the connected thermal load continues to be satisfied through stored heat, the electrical side of the system modulates its output to follow the TSO's regulation signal [31].

The process concludes with the restoration of primary reserves. As the aFRR response progressively corrects the frequency error and returns the system to its nominal state, the FCR assets that reacted autonomously during the initial 0 to 30 second window are gradually released from their activated position and returned to their neutral standby state [12], restoring the full primary reserve margin in preparation for any subsequent disturbance.

Manual Frequency Restoration Reserves

Manual frequency restoration reserves (mFRR) also referred to as tertiary regulation, represent the final and slowest layer of the frequency support hierarchy, characterized by manual activation and a primary role in sustained system restoration rather than the rapid corrective action of FCR and aFRR. As shown in Table 1, mFRR is deployed when secondary reserves have been operating at their limits for an extended period, signaling that the power imbalance is too large or too prolonged to be resolved through automated means alone [12]. Its function is to handle sustained power imbalances and to release the aFRR resources that have been depleted,allowing faster assets to return to their optimal operational state while slower, bulk energy carriers

Service Level	Physical Mechanism	Control Loop	Response Time	Goal
FCR	Local Kinetic/Inverter Energy	Decentralized	0-30 S	Stabilize
aFRR	Centralized Setpoint Tracking	Centralized (AGC)	30s – 15m	Restore
mFRR	Sustained Power Shift	Manual Dispatch	15m – Hours	Replace

Table 1: Technical Specifications And Operational Characteristics Of Frequency Support Hierarchy (FCR, aFRR, and mFRR).

absorb the longer-term balancing obligation [34].

The activation process begins with a manual dispatch instruction issued by the TSO. Upon observing that aFRR resources are operating continuously at their activation limits, the system operator issues a manual command to pre-qualified mFRR providers, initiating a controlled power shift that unfolds over a timescale of fifteen minutes to several hours [12]. This extended window, which distinguishes mFRR from the automated services above it in the hierarchy, is precisely what makes it suitable for assets whose physical characteristics favor sustained output over rapid response. The asset modulation stage exploits this extended timescale differently depending on the energy carrier involved. In heat based systems, large scale thermal storage units and gas boilers are particularly well suited to mFRR provision their ability to maintain a steady, controlled increase or decrease in power output for several hours allows CHP units to progressively reset their electrical to thermal dispatch ratios without disrupting thermal supply continuity [31]. In the mobility sector, EV aggregators can leverage the same extended window by shifting the charging schedules of an entire fleet, delaying the charging of a large number of vehicles by one or more hours to provide a sustained reduction in grid demand without requiring any individual vehicle to discharge [35].

The process ends with the systematic replacement of the secondary reserve that mFRR was called to relieve. As the mFRR resource reaches its target output level, the aFRR assets that had been operating at their activation limits are gradually reduced and returned to their standby state [12]. This staged handover ensures that the secondary reserve margin is fully restored before the next potential disturbance, closing the frequency support cycle and returning the system to a condition in which all three reserve layers, FCR, aFRR, and mFRR, are simultaneously available.

1.1.2 Emergency Services and System Restoration

The electricity system relies on a variety of emergency and support services to deal with major problems and help the whole system recover, in addition to regular frequency control such as:

Black start Ccpability is a restoration service where a specific generation unit or a cluster of resources restarts without an external power supply following a total grid shutdown. recent research emphasizes the transition from large centralized plants to dynamic islanding, where a part of the grid disconnects to operate autonomously during a wider fault. where upon detecting a major grid failure, the local controller triggers a separation from the utility grid. and during this islanded state, the system must immediately maintain its own frequency and voltage balance through the coordinated control of dispatchable units and energy storage. In addition to that systems utilizing PV, wind, and BESS must manage the active power mismatch to ensure the local frequency does not collapse before the secondary restoration begins.

Also voltage support which is an operational and emergency service that utilizes reactive power (Q) modulation to maintain nodal voltage magnitudes within safe operational limits, typically regulated at $\pm 5\%$ or $\pm 10\%$ of the nominal value depending on grid code specifications. IBRs such as BESS, and EV aggregators, can inject or absorb reactive power with sub second response times, operating significantly faster than traditional rotating synchronous machinery. Because network voltage is fundamentally a local physical parameter governed by line impedance distributed energy assets must be strategically sited. Proper spatial allocation optimizes local network voltage profiles and prevents localized voltage instability or voltage collapse triggered by the high volatility of intermittent renewable generation. This localized reactive support ultimately prevents the cascading tripping of sensitive power electronic equipment and maximizes the grid's renewable hosting capacity without violating structural network constraints.

Moreover, congestion management and load redispatch in which congestion management involves the strategic redispatching of active power resources and loads to prevent the thermal overloading of transmission and distribution lines. When a specific grid branch approaches its thermal current capacity, the system operator leverages MES flexibility to dynamically unload the congested path. Advanced optimization models incorporate network power flow equations as hard technical constraints within the optimization loop to guarantee that market clearing bids do not translate into physical violations of grid hardware safety thresholds. This active, DER-driven congestion management minimizes the necessity for drastic emergency measures such as rolling blackouts or forced load shedding, where power is directly cut to end consumers to preserve the integrity of the physical infrastructure.

1.1.3 Market-Based Procurement and Reliability Standards

The process of getting extra services has changed from being required for centralized power plants to being open to VPPs and demand side aggregators through competitive, market based systems. Most of the time, Ancillary services products are sold through a tendering process that happens before the delivery time (t^{lead}). Providers are chosen based on a merit order stack of bids tag. There are strict technical requirements for duration, ramp time, and reliability that must be met in order to participate in these markets. Often, reliability targets are higher than 99.9%This high threshold makes it very hard for stochastic resources to estimate extreme quantiles of expected availability. They need to use advanced optimization and uncertainty quantification techniques to do this. Providers in these markets are usually paid for keeping their services available and for the energy they actually deliver during activation. This encourages the creation of MES that are very flexible and reliable.

1.2. Implicit versus Explicit Ancillary Service Provision

The distinction between implicit and explicit ancillary service provision lies in two fundamental dimensions who controls the activation trigger and how the delivered value is recognized and compensated. These two pathways represent different relationships between the energy system and the grid operator, carrying distinct contractual obligations, technical requirements, and revenue mechanisms. Understanding this distinction is central to the study scope, since the transition from implicit to explicit provision is precisely the shift that transforms a MES from a passive participant in energy markets into a dispatchable, accountable grid-support asset.

1.2.1 Implicit Provision As Passive Response to Price Signals

Implicit flexibility often referred to as passive or price based demand response, arises when an energy system adjusts its consumption or generation patterns in response to time varying price signals such as day ahead market (DAM) prices or time of use (ToU) tariffs [1]. The technical objective in this mode is the minimization of the user's internal operational costs. In this configuration, the system does not enter into a formal contract with a system operator to provide a specific grid service. Instead, the flexibility is utilized internally for tasks such as peak shaving or energy arbitrage. For example, a MES might leverage the thermal inertia of a district heating network to shift heat pump operation to hours when electricity prices are low. While this reduces the owner's bill and provides a balancing benefit to the local grid, the response is entirely voluntary. Because the TSO cannot dispatch this resource with 100%İmplicit flexibility is considered a "by-product" of energy optimization rather than a firm grid security asset [42].

1.2.2 Explicit Provision As Committed Grid Support and Dispatchability

In contrast, explicit provision represents a proactive bound commitment where the MES offers its flexibility directly to the TSO or DSO as a discrete ancillary service product [25]. In this mode, the flexibility is no longer a by-product of internal cost saving it is the primary product being traded in specialized balancing markets. The technical process of explicit provision requires the system to reserve a specific amount of power MW or energy MWh capacity that cannot be used for internal optimization. This reserved capacity is placed under the direct or indirect control of the system operator. When a grid contingency occurs such as a frequency drop the MES must respond within the mandatory timescales of FCR, aFRR, or mFRR. The explicit nature is defined by three technical pillars:

- Contractual firmness in which the provider is obligated to deliver the service upon request. Failure to do so results in financial penalties, which is why research focuses so on 99.9% reliability as stated by fabian [37].
- Telemetry and verification where the MES must provide real-time data to the TSO to prove that the resource is available and that it followed the regulation signal accurately [12].
- Remuneration for availability where the system is paid just for "standing by," regardless of whether the grid actually needs to trigger the response. Unlike implicit flexibility, which only saves money on the energy bill [43].

1.2.3 Multi-Energy Systems as Enablers of Ancillary Services

The Concept of Integrated Electricity Heat Gas Mobility Systems

The idea of MES is replacing the old way of managing energy, which involved keeping electricity, thermal, and gas networks separate from each other. An MES is a complete system in which different types of energy, such as electricity, heat, gas, and mobility, work together in the best way possible at different levels, from single buildings to networks that cover whole districts [31]. In these integrated systems, energy is used and it is also changed and stored on different carriers to make the system as a whole work better and support the grid. Hub technologies like Combined Heat and Power (CHP) units, large-scale Heat Pumps (HPs), and Power-to-Gas (P2G) facilities make integration easier by connecting the power grid to other energy infrastructures [14]. The MES can use the built in storage and response times of non electrical sectors to offer services that would be expensive or difficult to do with only electrical assets by treating these carriers as a single, coupled system.

Sector Coupling and the Flexibility Pooling Mechanism

The idea of sector coupling is what makes the MES better. It allows for flexibility pooling, which is the combining of different technical characteristic to make a better flexibility product [4]. It is well known that storing large amounts of electrical energy is difficult and costly. On the other hand, thermal and gas networks have a lot of built in inertia. For instance, district heating networks and building envelopes can store heat energy for several hours with very little loss, making them a low cost, high capacity virtual battery [7]. When these assets are combined, they become more flexible than they were before. BESS or EV fleets can meet short term, high power needs like FFR. Long term energy balancing can be moved to the thermal or gas sectors. This pooling system greatly lowers the cycling stress on costly battery assets and lets the MES take part in several ancillary service markets at the same time. By coordinating these vectors a MES can keep the power profile at the PCC steady even when individual DERs are changing. This makes the resource more reliable and easier for the TSO to use [48].

Strategic Advantages Compared to Single-Energy Systems

Transition from single energy systems to MES offers a sophisticated operational framework characterized by three primary advantages multi vector redundancy, infrastructure based cost efficiency, and cross carrier energy neutrality. These advantages are rooted in the ability to convert energy between different states, allowing the system to bypass the strict conservation laws of a single carrier.

- Multi vector redundancy and system robustnes more specific the technical advantage of MES is the enhancement of system robustness through cross carrier synergy. In a single energy system, such as a standalone BESS or EV fleet, the ability to provide an explicit ancillary service is strictly limited by the real time availability of that specific asset [23]. If an EV fleet is unavailable due to unpredictable mobility patterns, the aggregator faces a service deficit and heavy financial penalties. In an MES, however, the system can utilize functional redundancy, if the electrical storage is constrained, the controller can modulate thermal loads (large scale heat pumps) or adjust the dispatch of a micro CHP unit to meet the committed reserve obligation. This multi vector approach creates a buffer that increases the probability of meeting the $>$99.9% reliability thresholds required in two stage of regulation markets where a single point failure would otherwise result in a breach of contract [16].

- Marginal cost optimization and infrastructure leveraging in which MES provide cost effective pathway to flexibility by leveraging existing non electric infrastructure. Installing dedicated, large scale electrochemical battery systems involves high capital expenditure (CAPEX) and degradation costs associated with frequent cycling [32]. In contrast, MES utilize the inherent storage capacity of thermal networks and gas pipes. Utilizing a 50,000-liter water tank for thermal energy storage (TES) or the linepack of a gas distribution grid for energy buffering is orders of magnitude cheaper per megajoule than lithium-ion technology [6]. By converting surplus electricity into heat or gas during "downward" regulation events, the MES achieves high-fidelity grid support using assets that are already amortized for heating or industrial processes, lowering the marginal cost of providing ancillary services.

- Overcoming the energy neutrality and SoC saturation challenge perhaps the most significant technical advantage is the resolution of the energy neutrality challenge inherent in frequency regulation. Services like aFRR often require the provider to return to a net-zero energy balance over a specific period to ensure the asset is ready for the next event [38]. In single energy systems, a prolonged upward or downward regulation signal often leads to SoC violations the battery becomes either fully empty or fully saturated,

forcing it to stop providing the service. An MES solves this by using energy conversion as a relief valve. For instance, during a sustained downward regulation event, an MES can convert the surplus electricity into heat or hydrogen [15]. Instead of saturating an electrical battery, the energy is moved to a different carrier's storage, preventing SoC saturation in the electrical domain and allowing the system to offer high capacity services for significantly longer durations.

1.3. Scope, Objectives, and Contribution of the Thesis

The aim of this study is to provide a systematic classification of how different energy carriers can explicitly provide ancillary services within MES. This thesis focuses on classifying and comparing the modeling approaches already used in the literature for heat based, electricity based, and mobility based systems, and for the multi-vector architectures that combine them. A central part of this classification is the specific mechanism through which each carrier delivers a service, whether through active and reactive power modulation, thermal load shifting, or capacity reservation, since these mechanisms determine which ancillary service products a given carrier is technically capable of supporting. Building on this classification, the thesis conducts comparative analysis across the diverse system configurations and optimization methodologies identified in the reviewed literature. Rather than evaluating a single technology in isolation, it examines how different carriers perform under comparable grid stress conditions, and extends this comparison to the underlying mathematical paradigms researchers employ, contrasting deterministic linear and mixed-integer programming with uncertainty-aware frameworks such as stochastic and robust optimization. This comparison reveals which modeling choices are most effective for capturing a MES feasible operating region, and exposes a recurring methodological pattern.

This thesis differentiates itself from existing literature in three ways. First, while a substantial share of prior work specializes in a single energy carrier often EV and treats flexibility as an implicit by product of internal cost optimization, this thesis synthesizes findings across electricity, heat, hydrogen, and mobility to show how these vectors function as a coordinated flexibility pool, in which the limitations of one carrier, such as the limited energy capacity of a battery are compensated by the strengths of another such as the thermal inertia of a heating network. Second, rather than treating flexibility as a purely theoretical concept, this review specifically selects and analyzes studies that model ancillary services as formal, contracted commodities subject to rigorous product defining constraints connecting abstract energy management research to real world grid code compliance. Third, the thesis provides a cross comparison of mathematical paradigms across carriers examining why stochastic or robust optimization is favored for some carriers over others and surfacing methodological inconsistencies that a single carrier review would not reveal. Taken together this comparative synthesis forms the basis for the research gaps identified in Section four, which represent the principal scientific contribution of this thesis.

2. Classification Framework for Explicit Ancillary Service Modeling

2.1. Methodological Rationale and the Need for Structure

As traditional ancillary servivce providers exit the market, there is a technical necessity to integrate DERs through VPPs to enhance market liquidity and ensure grid security. Evaluating the potential of these diverse resources ranging from HP to EV requires a well build classification framework to map physical assets to standardized ancillary service products with rigid technical specifications. This structure is critical for addressing the inherent variability of weather dependent generation and the forecasting uncertainties that can put reserve availability at risk . Moreover, an MES specific framework allows for the quantification of the Feasible Operating Region (FOR), representing the set of potential power exchanges between a cluster of resources and the upstream network while adhering to strict reliability thresholds defined by TSO qualification requirements.

2.2. Overview of Comparison Dimensions

The classification framework developed for this thesis evaluates how MES are modeled for explicit ancillary service provision across four interconnected dimensions the energy carrier and physical resource configuration of the system, the spectrum of ancillary service products it targets, the optimization technique employed to manage competing technical and economic objectives, and the market participation context in which the service is procured and delivered. These dimensions are not only organizational labels each one represents a distinct layer of technical and contractual constraints that together determine whether a cluster of resources can reliably

meet grid code requirements. They form the analytical structure through which the twenty two reviewed papers are compared and synthesized throughout this study as showun in figuer (1).

The first dimension concerns the Energy Carrier and System Configuration, which identifies the physical nature of the assets providing the service and distinguishes between large scale centralized power plants and decentralized independent systems. Centralized hybrid plants are typically utility scale configurations in which a BESS is collocated with a large wind or solar farm to mitigate intermittency and provide firm frequency reserves, as documented. Independent systems and microgrids, by contrast, represent localized clusters in which diverse carriers such as heat pumps, photovoltaic arrays, and electric vehicle fleets are coordinated together and are typically modeled as VPP that aggregate small scale flexibility to reach the minimum capacity thresholds required for market entry. Within both configurations, the literature further distinguishes between four primary carrier types, electricity only and hybrid renewable systems, integrated electricity and heat systems, mobility based flexibility through electric vehicle aggregators, and multi-vector systems combining three or more carriers including hydrogen, each of which introduces a distinct set of physical capabilities and constraints that determine which ancillary services it can reliably support.

The second dimension covers the Spectrum of ancillary service Products captured in the reviewed literature, organized into three categories that reflect the physical parameter each service regulates rather than the timescale on which it operates. The first category frequency related services covers frequency containment reserves, aFRR, and mFRR, and remains the primary focus of the reviewed literature due to its high market value and strict technical requirements, these services span a wide range of timescales, from the sub second response of FCR to the sustained power shifts of mFRR, which is itself the mechanism through which longer duration ramping and

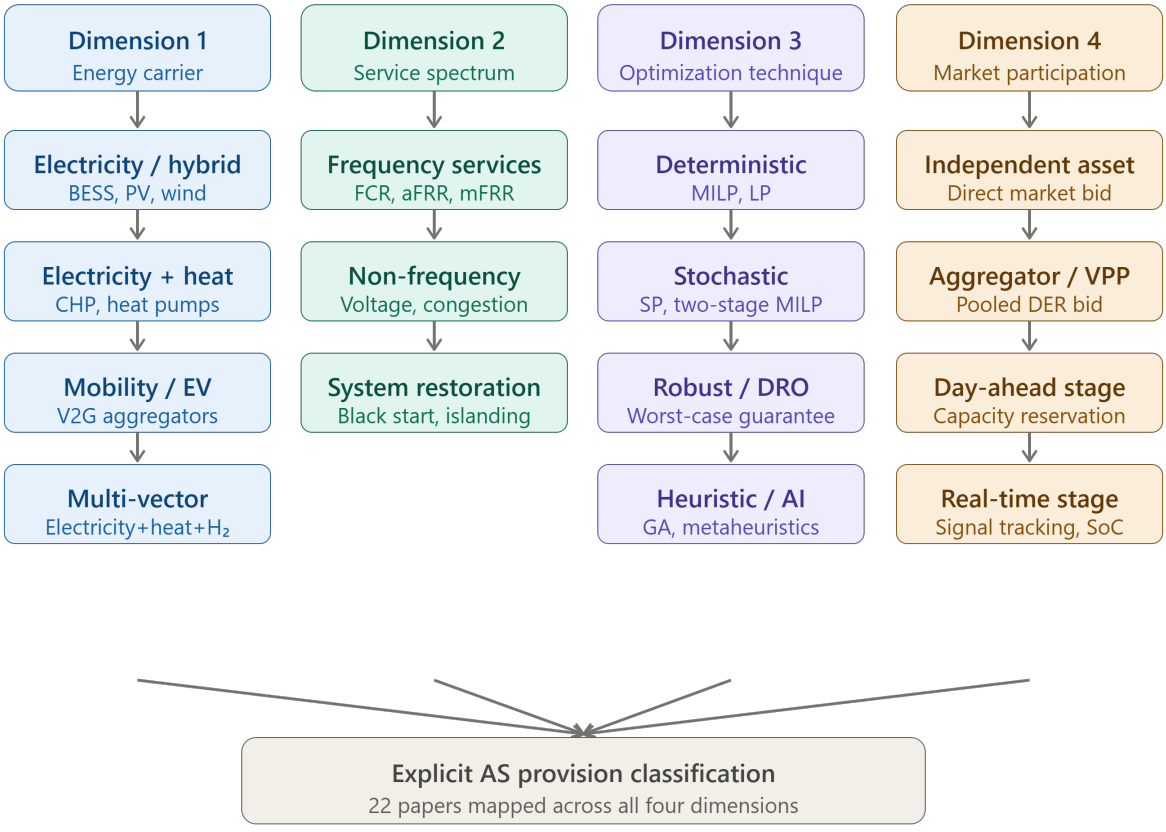


Figure 1: Developed four-dimensional classification framework for mapping multi-energy systems to explicit ancillary service provisions.

balancing needs such as managing steep changes in renewable output over several hours are addressed within this category. The second category, non-frequency services, covers voltage support through reactive power injection and congestion management through strategic redispatch, both of which are increasingly vital in distribution networks where high PV penetration can cause local voltage violations or thermal overloading. The third category,

system restoration, covers black start capability and dynamic islanding, capturing how the reviewed literature treats the recovery of a microgrid or grid segment from a total or partial outage without external power support.

The third dimension addresses the optimization technique employed to model the system's behavior and determine its optimal operating strategy. The reviewed literature organizes into three methodological paradigms. Deterministic optimization, primarily MILP, represents the foundational approach and is most prevalent in studies focused on infrastructure sizing and feasibility assessment, where all input parameters such as weather dependent generation and market prices are assumed known with certainty. Uncertainty-aware frameworks, comprising Stochastic Programming and Robust Optimization, represent a growing share of the literature in response to the high reliability thresholds demanded by transmission system operators stochastic programming optimizes expected performance across a set of probabilistic scenarios, while robust optimization provides a worst case guarantee by optimizing against the most adverse realization within an ambiguity set. Finally, a smaller subset of studies employs heuristic and artificial intelligence-driven methods such as genetic algorithms and metaheuristics to address highly non-linear problems that traditional linear solvers cannot efficiently handle.

The fourth and final dimension classifies the market participation context, addressing both the organizational role of the MES and the temporal horizon across which its commitments are made and delivered. The reviewed literature distinguishes between two organizational models the independent asset or large-scale plant, which bids directly into the market as a single physical entity, and the aggregator or VPP, which coordinates location-dispersed and heterogeneous resources to meet market entry thresholds as a unified bid. Temporally, most explicit service models operate across a two-stage decision process a day ahead scheduling stage in which capacity is reserved and market commitments are made based on probabilistic forecasts, and a real-time activation stage in which the focus shifts to high-fidelity signal tracking and SOC management to ensure the committed service is physically delivered within the mandatory response window.

2.3. Distribution of Energy Carriers and Resource Clusters

The reviewed literature is organized into four clusters based on the physical nature of the energy carriers and the specific assets used to provide the explicit ancillary service. Table (2) demonistrate the The twenty two paper discuseed in the study along side with the four interconnected dimensions used.

Cluster 1: Electricity-Only and Hybrid Renewable Systems

The largest cluster focuses on the coordination of IBR, primarily BESS, PV arrays, and wind turbines, to provide fast frequency and voltage support. Bahloul, Daoud, and Khadem develop a deterministic day-ahead linear programming dispatch strategy for a collocated PV-BESS power plant operating in the Irish DS3 market, demonstrating that multi-service stacking, simultaneously minimizing PV clipping losses while providing frequency and power reserves, creates additional economic value, and that incorporating both calendar and cycling battery ageing models significantly alters the optimal dispatch profile compared to ageing unaware formulations [3]. Santosuosso et al. propose a stochastic economic model predictive control framework in which a wind-BESS hybrid simultaneously trades in the energy market and bids into the frequency containment and frequency restoration reserve markets their key finding is that the SEMPC approach outperforms conventional rule based dispatch by maximizing expected profit while explicitly minimizing storage degradation across probabilistic scenarios of renewable production and ancillary service activation [40]. Wang and Yuan present a multi timescale hierarchical MILP for a smart microgrid that integrates four heterogeneous EV types alongside PV and wind, coordinating day-ahead scheduling with intra day rolling optimization the results show that large capacity EVs provide the most substantial grid services while smaller EVs display distributed charging patterns, with active power deviations remaining consistently below 10% of total load [47]. Kumar et al. use a bi-level optimization framework solved with a genetic algorithm on a real 108-bus Indian distribution system, where level one optimally sites/sizes DERs to minimize annual energy loss and level two manages hourly ancillary services (load deviation, reverse power flow, voltage deviation), their real finding is that including ancillary services provision in the planning stage improves loss reduction, voltage profile, and power factor compared to ignoring it.[26].

In addition, Kassab et al. address the joint problem of microgrid sizing and energy management using two distinct approaches, a cascaded MILP and an embedded accelerated particle swarm optimization-MILP hybrid, applied to a DC-bus campus microgrid with EV flexibility, the Pareto front analysis demonstrates that simultaneous minimization of levelized cost of energy (LCOE) and life cycle emissions is achievable without sacrificing

technical reliability [24]. Chen et al. investigate a VPP comprising both residential loads and industrial cement plants as complementary demand side resources, using a two-stage stochastic MILP, they show that the distinct power and energy characteristics of residential versus industrial microgrids create a synergistic complementarity that improves overall regulation capacity and profitability in the ancillary service market [5]. Iuliano and Vaccaro address the uncertainty quantification problem in microgrid flexibility scheduling by combining affine arithmetic with robust optimization, their formulation computes a family of feasible solutions that includes the robust optimum while overcoming the over-conservatism of conventional range based robust operators, demonstrated on a realistic microgrid case study [21]. Majzoobi and Khodaei develop a MILP-based optimal scheduling model for a microgrid providing ancillary services at three different time resolutions simultaneously, hourly ramping, ten-minute load following, and one-minute frequency regulation their comparative analysis of PJM and European market structures reveals that minimum bid sizes of typically one megawatt effectively exclude small microgrids unless they adopt aggregation frameworks [30]. Finally, Maeyaert, Vandevelde, and Döring develop a two-stage predictive and real-time control framework for BESS operating in a smart distribution grid as a prosumer asset, their key contribution is the demonstration of service stacking, using the same battery to provide frequency containment reserves for the TSO while simultaneously delivering voltage support and congestion relief for the DSO, creating a dual revenue stream [29].

Cluster 2: Integrated Electricity and Heat

This cluster exploits the thermal inertia of heating networks as a virtual buffer for the power grid. Zhang et al. propose a two-stage stochastic MILP for a day-ahead integrated electricity and heat system in which large-scale heat pumps explicitly provide both following and regulating reserves alongside wind energy they introduce an analytical inactive transmission constraint identification algorithm that significantly reduces computational burden, and demonstrate that coordinated heat pump reserve provision increases system flexibility potential by up to 25% compared to uncoordinated operation [49]. Ji, Wei, Zhou et al. propose a two-stage distributionally robust optimization model for coordinating energy and reserve dispatch in an integrated electricity-gas-heating system under multiple correlated wind power uncertainties they apply a dirichlet process mixture model to extract uncertainty parameters directly from historical data and construct data-driven ambiguity sets, then solve the decomposed problem using the column and constraint generation algorithm, demonstrating that the DRO approach provides both economic efficiency and a mathematical reliability guarantee superior to standard stochastic programming [22].

Cluster 3: Mobility-Based Flexibility

This cluster addresses the unique stochastic availability of Electric Vehicle fleets and the role of the aggregator in transforming fragmented, temporally uncertain battery capacity into a firm ancillary service product. Alipour et al. develop a stochastic MINLP scheduling model for plug-in EV aggregators participating in both energy and ancillary service markets, characterizing the impact of market price and reserve market uncertainties through a two-stage stochastic programming framework. Their results show that while a deterministic model may suggest full fleet availability, the stochastic noise of vehicle presence typically necessitates a 30% security margin to ensure reliable service delivery [2]. Sarker, Dvorkin, and Ortega-Vazquez formulate a MILP bidding strategy for an EV aggregator in the day-ahead energy and reserve market that explicitly compensates vehicle owners for battery degradation caused by cycling beyond transportation needs their results demonstrate that splitting fleet resources between energy and reserve markets is economically optimal, and that the aggregator's participation generates measurable cost savings for the system operator [41]. El-Hendawi, Wang, and Liu propose both centralized and distributed optimization models for V2G frequency regulation, formulating the centralized problem as a linear program and decomposing it into subproblems using the alternating direction method of multipliers validated against real data from the independent electricity system operator of ontario, the distributed model is shown to overcome the privacy, complexity, and single point of failure limitations of the centralized approach while maintaining equivalent regulation performance [11]. De la Torre, Aguado, and Sauma develop a MILP model for an EV aggregator that jointly incorporates driver patterns, state-of-charge constraints, capacity constraints, and power system security constraints as hard constraints in the optimization by including network power flow equations, the model ensures that ancillary service bids do not create local congestion, and their results show that this integrated approach produces globally optimal allocation between day-ahead energy and regulation capacity [10]. Nespoli et al. address the scalability challenge of national-scale bi-directional EV fleet control by decomposing a mixed-integer quadratic programming problem using ADMM, enabling simultaneous scheduling of thousands of vehicles for ancillary service provision their sensitivity analysis demonstrates that EV clusters located at radial feeder ends provide the highest value for voltage support through reactive power modulation [36]. Ma develops a MILP scheduling model for EV aggregators providing secondary frequency regulation in renewable-powered grids, explicitly modeling EV parking constraints and SOC manage-

Paper Name	System Used	Ancillary Services	Optimization Technique	Type of Power System
Affine arithmetic meets robust optimization in scheduling microgrids-based flexibility services [21]	PV, Wind Energy, Diesel, Gas Turbine	Power Reserve	MILP, Affine arithmetic, Robust optimization	Distribution System.
Application of microgrids in providing ancillary services to the utility grid[30]	PV, Wind Energy, Diesel, Gas Turbine	Frequency Regulation, Power Reserve, Ramping support	MILP	Distribution System / AS Market
Smart microgrid optimization with electric vehicle flexibility for intra-day energy management [47]	PV, Wind Energy, Diesel, BESS, EV	Frequency Regulation, Power Reserve, Voltage support	Multi-timescale hierarchical MILP	Smart Microgrid (Grid-Connected)
Microgrid sizing with EV flexibility: Cascaded MILP and embedded APSO-MILP approaches [24]	PV, Wind energy, BESS, EV	Implicit (Peak shaving, RE integration)	Deterministic MILP, Cascade MILP, APSO-MILP	DC-Bus Campus Microgrid
Optimal microgrid planning for enhancing ancillary service provision [8]	PV, Wind Energy, BESS, Microturbine	Power Reserve and Frequency Regulation	Multi-objective NSGA-II	Medium Voltage (MV) Distribution
A profitability optimization approach of VPPs comprised of residential and industrial microgrids[5]	PV, Energy storage loads	Frequency Regulation	Two-stage stochastic MILP	VPP with Multiple Microgrids
Strategic integration of battery energy storage systems with distributed ancillary services [26]	Wind Turbine, BESS system	Frequency Regulation, Voltage Support	GA, Bi-level MINLP	Active Distribution (High RES)
Optimal dispatch of battery energy storage for multi-service provision considering ageing [3]	PV, BESS Plant	DS3 (Power Reserve), Different Response T	Deterministic Day-ahead MILP	Grid-connected utility-scale
Multi-energy balancing services provision from a hybrid power plant: PV, battery, hydrogen Technologu [9]	BESS, PV Power Plant, Electrolyzer, Fuel Cell, Hydrogen	Power and frequency Reserve (Electricity and Gas)	Two-stage Stochastic MILP	Integrated MES (Electricity-Gas-H2)
Stochastic economic model predictive control for renewable energy trading with storage [40]	Wind Energy, BESS	Frequency and Power Reserve	Two-stage Stochastic MILP	Grid-Connected Hybrid Plant
Coordinated electric vehicles dispatch for multi-service provisions: A comprehensive review [44]	EV and others	Voltage, Frequency and Power	Stochastic Programming, Robust Optimization, RL	Depending on Paper Reviewed
Optimal scheduling of ancillary services provided by an electric vehicle aggregator[10]	EV	Frequency Regulation (Secondary aFRR)	Two-stage Stochastic MILP	Centralized Market-Driven
Battery Storage for Ancillary Services in Smart Distribution Grids[29]	BESS	Voltage, FCR, and Congestion Relief	Two-Stage Predictive and Real-Time Framework	Decentralized Incentive-Driven

Continued on next page...

Paper Name	System Used	Ancillary Services	Optimization Technique	Type of Power System
Bidding strategy design for electric vehicle aggregators considering price volatility[50]	EV	Risk-managed energy procurement	Two-stage Stochastic LP	Centralized Market-Participant
Stochastic scheduling of aggregators of plug-in electric vehicles for AS markets [2]	EV	Power Reserve	MINLP	Distribution Grid
Optimal Participation of an EV Aggregator in Day-Ahead Energy and Reserve Markets[41]	EV, Wind Energy	Power, Frequency regulation	MILP	Distributed, Grid-connected EV
A hierarchical robust scheduling framework for EV aggregators in spot and AS markets[33]	EV	Frequency Regulation	Bi-level Stochastic/robust framework	AS and DER Aggregation
Optimal scheduling of EV aggregators for frequency regulation and cost efficiency[28]	EV, PV, Wind Energy	Frequency Regulation	MILP	Centralized Coordination
National-scale bi-directional EV fleet control for ancillary service provision[36]	EV, PV Energy	Power Reserve	MIQP, ADMM	Distribution-Level AS
Centralized and Distributed Optimization for V2G Applications in Frequency Regulation[11]	EV	Frequency Regulation	LP, ADMM and Lagrange relaxation	Centralized and Distributed
Day-ahead stochastic scheduling of IEH system considering reserve by large-scale heat pumps[49]	Wind, HP, CHP	Freq. and Power Balance	Two-stage Stochastic MILP	Integrated Elec. and Heat (IEHS)
Distributionally robust co-optimization of energy and reserve dispatch for integrated IEGHS [22]	CHP, Boiler, Wind Energy	Power Reserve, Frequency regulation	Two-stage Robust Optimization	Integrated Elec-Gas-Heating

Table 2: Comprehensive Technical Synthesis Of Primary Literature Mapping Energy Carriers, Ancillary Service Products, Optimization Methodologies, and Type of Power System

ment the results show that ignoring parking constraints underestimates charging costs by approximately 20% while the proposed down-regulation strategy reduces total charging costs by around 5% compared to unmanaged charging [28]. Meng et al. propose a hierarchical bi-level robust scheduling framework for EV aggregators for EV aggregators operating simultaneously in the electricity spot market and the frequency regulation ancillary service market using probability-interval hybrid uncertainty modeling and the column and constraint generation algorithm, their distributionally robust optimization model provides a mathematical guarantee of market compliance even under extreme scenarios of EV user behavior and price volatility, outperforming both deterministic and standard stochastic approaches in terms of robustness [33]. Zheng, Wang, and Yang develop a risk-averse stochastic linear programming bidding strategy for EV aggregators in the day-ahead electricity market, using conditional value-at-risk to explicitly manage the financial exposure to real-time price clearing uncertainty their results on the PJM market show that the CVaR-based risk-averse strategy significantly reduces the probability of large financial losses compared to the conventional risk-neutral expected-cost minimization approach [50].

Cluster 4: Multi-Vector Systems and General Reviews

The final cluster represents the most integrated architectures, combining three or more energy carriers to overcome the duration limits of single vector systems and provide services across multiple energy networks simultaneously. Čović, Pavić, and Pandžić propose a two-stage stochastic MILP model for a tri-hybrid PV-BESS-hydrogen power plant that provides balancing services simultaneously in both the power system and the gas network their key contribution is demonstrating that operating power to hydrogen and hydrogen to power units in one system creates cross system imbalances in the other, a problem previously unexplored in the literature, and that this cross system coupling must be explicitly modeled to ensure contractual compliance with both the power TSO and the gas TSO [9]. Su, Wang, and Strbac provide a comprehensive review of coordinated EV dispatch for multi service provisions, classifying ancillary services into energy imbalance, security related, and social welfare categories, and synthesizing control approaches ranging from stochastic programming and robust optimization to reinforcement learning. Their central conclusion is that the future of ancillary services lies in aggregated heterogeneity, where mixing the speed of batteries, the inertia of heat networks, the versatility of EVs, and the energy density of hydrogen into a unified service agnostic VPP produces a portfolio that is mathematically more reliable than any single asset approach [44]. Contreras, Cortes, and Myrzik propose a probabilistic multi objective microgrid planning methodology using the non dominated sorting genetic algorithm to simultaneously minimize power mismatch in islanded mode, maximize ancillary service provision revenue, and minimize investment and operational costs at the medium voltage distribution level, their results demonstrate that proper sizing and location of multi-vector distributed resources, including micro turbines alongside PV and BESS, enables a Net Present Value (NPV) improvement of up to 18% compared to electricity only systems, driven by the cross-carrier energy cascading effect [8].

2.4. Methodological Dimensions Optimization and Uncertainty Management

Having introduced the four classification dimensions in Section 2.2 and established the paper level descriptions across the four carrier clusters in Section 2.3, this section provides a deeper analytical synthesis of dimensions three and four the optimization technique employed and the market participation context in which services are procured and delivered. The objective here is not to re-describe individual papers but to identify the methodological patterns that emerge when the reviewed literature is read as a whole, specifically the correlation between carrier type, service reliability requirement, and modeling choice, and the recurring structural distinction between day-ahead commitment and real-time activation.

Mapping Optimization Techniques to Services and Asset Performance

The relationship between the optimization paradigm chosen by a study and the ancillary service it targets is not arbitrary it reflects a strict physical logic in which the required response timescale and reliability threshold of the service dictate the minimum modeling sophistication necessary to guarantee contractual compliance. For Frequency Containment Reserves and secondary regulation, where market values are highest and TSO reliability thresholds typically exceed 99% the reviewed literature divides into two methodological schools depending on whether the asset is deterministic or stochastic in nature. Studies dealing with infrastructure sizing and economic feasibility assessment, where the goal is to establish whether a system configuration is viable rather than to guarantee a specific delivery probability, predominantly employ deterministic MILP. Bahloul et al. [3] use this approach for their PV-BESS dispatch model, finding that a deterministic framework is sufficient to demonstrate multi-service value but insufficient to capture the tail-risk of frequency signal volatility. Majzoobi and Khodaei [30] similarly use MILP to schedule microgrid ancillary services across three timescales, but their comparative market study reveals the structural barrier deterministic models impose by assuming full asset availability, they systematically overestimate the capacity a system can reliably offer, making the resulting bids non-bankable without an additional reliability correction. Maeyaert et al. [29] adopt a two-stage predictive and real-time MILP to address this, separating the day-ahead sizing decision from the real-time dispatch correction, a structural compromise that retains computational tractability while partially capturing real-time uncertainty.

As the asset base shifts toward inherently stochastic carriers, the literature converges on Stochastic Programming as the standard modeling response. Santosuosso et al. [40] demonstrate this transition explicitly for wind-BESS systems, showing that a stochastic economic model predictive control framework outperforms rule-based dispatch by optimizing across probabilistic scenarios of renewable production and ancillary service activation simultaneously. Wang and Yuan [47] apply a multi timescale hierarchical MILP that coordinates day-ahead

and intra-day decisions for a heterogeneous EV microgrid, finding that the hierarchical structure is necessary to prevent the day ahead schedule from becoming infeasible once real time EV availability deviates from forecast. Chen et al. [5] extend this to the VPP level, demonstrating that the complementarity between residential and industrial load profiles, captured only through stochastic scenario generation, is what makes the combined bid more reliable and profitable than either load type alone.

For the mobility cluster, where human behavior introduces the most severe stochastic availability, the literature moves further along the uncertainty spectrum toward robust and distributionally robust optimization. El-Hendawi, Wang, and Liu [11] show that a distributed LP model decomposed through ADMM provides robust frequency regulation performance against dynamic EV usage patterns without requiring an explicit probability distribution. Meng et al. [33] go further with a bi-level DRO framework that uses probability interval hybrid uncertainty sets to bound both EV behavior and market price fluctuations simultaneously, demonstrating that this approach provides a mathematical compliance guarantee that neither deterministic nor standard stochastic models can match. Zheng, Wang, and Yang [50] address a different dimension of the same problem, using conditional value-at-risk within a stochastic LP to manage financial rather than physical risk, showing that risk neutral expected cost minimization systematically exposes the aggregator to large losses from real-time price clearing that a CVaR-based strategy avoids. Iuliano and Vaccaro [21] contribute distinct approach by combining affine arithmetic with robust optimization, demonstrating that their formulation overcomes the over conservatism inherent in conventional range based robust operators by computing a family of feasible solutions rather than a single worst case optimum.

For tertiary reserves and longer duration balancing, where the sustained power shift requirement exceeds what electrochemical batteries can sustain, the literature shifts toward multi carrier assets and multi objective methods. Čović, Pavić, and Pandžić [9] use a two stage stochastic MILP specifically because the cross system coupling between the power and gas TSOs creates correlated uncertainties that a deterministic model cannot represent. Contreras, Cortes, and Myrzik [8] adopt the non dominated sorting genetic algorithm rather than MILP because their multi objective planning problem, simultaneously minimizing cost, maximizing ancillary service revenue, and reducing emissions across a medium voltage microgrid, involves non convexities that standard linear solvers cannot handle efficiently. Kumar et al. [26] similarly employ a bi-level genetic algorithm combined with MINLP because the non-linear relationship between BESS sizing, network hosting capacity, and ancillary service provision cannot be reduced to a linear approximation without losing the physical accuracy required for investment decisions.

For voltage support and congestion management at the distribution level the reviewed literature relies primarily on network constrained MILP formulations that embed power flow equations as hard constraints. De la Torre, Aguado, and Sauma [10] demonstrate that including network security constraints in the EV aggregator optimization prevents bids that would be physically infeasible at the local feeder level, even when they appear globally optimal. Nespoli et al. [36] extend this to a national scale, using ADMM decomposition to manage the computational burden of simultaneous voltage support across thousands of EV connection points, and showing through sensitivity analysis that EV clusters at radial feeder ends deliver the highest reactive power value per unit of connected capacity. Finally, the comprehensive review by Su, Wang, and Strbac [44] synthesizes these methodological trends into a unified taxonomy, concluding that the future of multi service provision requires service-agnostic controllers that can automatically select which vector and which optimization paradigm to deploy based on the real time marginal cost of each carrier and the specific grid call being answered.

Market Participation and Temporal Context

The market participation dimension reveals a structural divide in how the reviewed literature organizes resources to meet market entry thresholds and how it models the two-stage temporal commitment that defines explicit service provision. On the organizational side, a first group of papers models the independent asset or large scale plant role, in which a single physical entity bids directly into the market without aggregating dispersed resources. Bahloul et al. [3], Čović et al. [9], Santosuosso et al. [40], Kassab et al. [24], Majzoobi and Khodaei [30], Maeyaert et al. [29], Iuliano and Vaccaro [21], Wang and Yuan [47], and Contreras et al. [8] all operate within this model, focusing their optimization on the internal coordination of collocated assets to maximize service reliability at the point of common coupling. The shared challenge across this group is ensuring that the feasible operating region remains contractually valid across all realization scenarios of renewable output and load, which is why this group spans the full range from deterministic MILP to DRO depending on the volatility of the assets involved.

The second group models the aggregator and VPP role, where the coordination challenge shifts from internal asset management to the statistical aggregation of location dispersed, owner operated resources. This role is central to the EV cluster, where Alipour et al. [2], Sarker et al. [41], De la Torre et al. [10], El-Hendawi et al. [11], Nespoli et al. [36], Ma [28], Meng et al. [33], and Zheng et al. [50] all struggle with the same fundamental problem translating a fragmented stochastic fleet of individual vehicle batteries into a single, firm reserve bid that meets TSO minimum capacity thresholds. The aggregation logic is also central to Zhang et al. [49] and Ji et al. [22] in the integrated electricity and heat cluster, where coordinating large numbers of individually small heat pumps and district heating nodes requires the same statistical pooling approach used in EV fleet management. Chen et al. [5] and Su et al. [44] bridge both groups, demonstrating that the complementary characteristics of heterogeneous resource pools, whether across microgrids of different demand types or across carriers of different physical properties, generate reliability improvements that neither homogeneous pool nor single asset model can replicate.

On the temporal side, the reviewed literature converges on a two-stage decision structure that mirrors the physical separation between market clearing and service delivery. The first stage, the day ahead scheduling and bidding horizon, is the focus of Alipour et al. [2], Sarker et al. [41], Zheng et al. [50], De la Torre et al. [10], Santosuosso et al. [40], and Nespoli et al. [36], all of which commit to a reserve quantity based on probabilistic forecasts of availability and price. The second stage, real time activation and signal tracking, is where Bahloul et al. [3], Zhang et al. [49], Ji et al. [22], Wang and Yuan [47], and Maeyaert et al. [29] focus their methodological contribution, prioritizing high-fidelity control of state of charge and power setpoints to ensure that the asset physically follows the TSO regulation signal without violating its operational limits. The studies that perform best in terms of overall market bank ability, notably Santosuosso et al. [40], Meng et al. [33], and Čović et al.[9], are those that explicitly model both stages within a unified optimization framework rather than treating them as separate problems.

The Reliability-Profit Paradox: From Determinism to Uncertainty Awareness

The transition from deterministic to uncertainty-aware modeling that runs through the reviewed literature is not merely a computational upgrade it represents a fundamental shift in how the engineering problem of explicit ancillary services provision is framed. In a deterministic framework such as that employed by Bahloul et al. [3] and Majzoobi and Khodaei [30], the optimization treats asset capability as a static constraint if the battery has a rated capacity of X megawatt hours, the model assumes it will always be available at that capacity. This creates what can be termed a foresight trap the model appears to find a profitable bidding strategy, but that strategy is only valid if the physical world behaves exactly as the deterministic input data assumed. In reality, frequency signal intensity, solar irradiance, wind speed, EV arrival times, and market prices all deviate from their forecasts, and those deviations directly threaten the firmness of the committed service.

Stochastic programming addresses this by optimizing the expected profit across a set of representative scenarios rather than a single forecast. Alipour et al. [2], Sarker et al. [41], Chen et al. [5], and Čović et al. [9] all demonstrate that this approach produces more conservative but more reliable bids, the expected profit is lower than the deterministic optimum, but the probability of non-compliance, and therefore the expected penalty cost, is substantially reduced. However, as Meng et al. [33] argue, stochastic programming's reliance on a fixed probability distribution remains a weakness when the true distribution of EV behavior or market prices is itself uncertain or shifts unexpectedly. This is the motivation for distributionally robust optimization, adopted by Meng et al. [33] and Ji et al. [22], which optimizes against an ambiguity set of plausible distributions rather than a single assumed one, providing a compliance guarantee that holds even when the scenario tree used in stochastic programming proves to be misspecified.

The adoption of these advanced frameworks introduces a well-documented trade-off between model fidelity and computational tractability. A deterministic MILP can schedule a year-long horizon in seconds, while a DRO model must solve a min-max problem identifying the most adversarial scenario within the ambiguity set and optimizing against it. For aggregators managing thousands of EVs or heat pumps, this results in a computational burden that Meng et al. [33] and Ji et al. [22] both address through the column and constraint generation algorithm which decomposes the problem into a tractable sequence of master and subproblems without sacrificing the worst case guarantee.

A further dimension of this paradox is the role of multi-vector synergy as a form of physical hedging that reduces the dependence on advanced mathematical frameworks. As demonstrated by Čović et al. [9] and Contreras et al. [8], an MES that can shift a grid commitment from one carrier to another when the primary asset is

constrained achieves a natural robustness that a single-vector system can only replicate through mathematical conservatism. In this sense physical redundancy across carriers and mathematical robustness through DRO are complementary rather than competing strategies the former reduces the severity of the worst case scenario, while the latter provides a guarantee that even that reduced worst case will not breach the service contract. Together these two mechanisms define what this study terms the conservative feasible operating region the set of power exchanges a MES can commit to with the same confidence that a TSO currently places in traditional synchronous generation.

.

3. Electricity-Only and Hybrid Systems: The Fast Responders

The following sections examine how papers within each cluster relate to, build upon, and at times diverge from one another around shared technical themes, such as the trade-off between profitability and reliability, the role of optimal sizing, and the structural barriers to market entry. The first cluster of the reviewed literature focuses on the most demanding segment of the ancillary services market FCR and secondary regulation aFRR. Figure (2) give a brife inside about the energy duration for different storge capcity Placing a premium on the rapid ramp rates of power electronics. This section combines the findings of ten papers that utilize Battery Energy Storage Systems BESS, Photovoltaic arrays, and Wind turbines as primary assets.

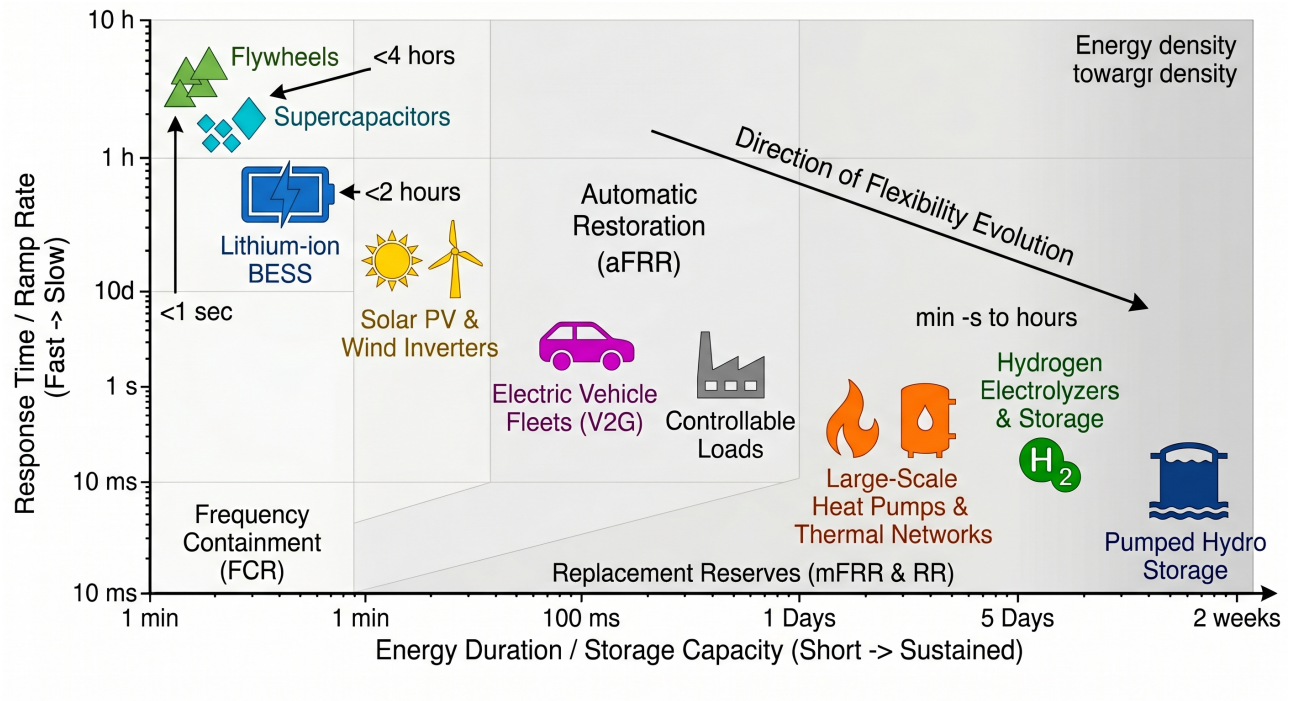


Figure 2: Conceptual Classification Of Flexibility Source Based On Operational Response Dynamic

A close reading of this cluster reveals that the nine reviewed papers divide along a consistent fault line a first group treats the physical asset as fixed and asks how it should be operated, while a second group treats operation as secondary and asks how the asset should be sized or organized in the first place. Bahloul [3], Santosuosso [40], and Maeyaert [29] belong to the first group, and although their methodologies differ considerably, ageing-aware deterministic scheduling, stochastic degradation-minimizing control, and pricing-based service stacking, they converge on the same underlying concern, preserving battery health while maintaining the reliability a TSO demands. None of the three questions whether the battery is the right size the battery's capacity is a given, and the contribution lies entirely in how intelligently that fixed capacity is deployed.

Kumar [26], Kassab [24], and Iuliano and Vaccaro [21] occupy the opposite position where sizing and planning rather than operational control is the central problem. Kumar's finding that ancillary service provision must be considered at the planning stage not bolted on afterward it is reinforced and extended by Kassab, whose comparison of sizing methods shows that a more sophisticated embedded optimization approach outperforms a simpler cascaded one on cost and emissions simultaneously. Iuliano and Vaccaro address a dimension neither paper considers directly, the uncertainty inherent in the renewable forecasts that any sizing decision ultimately

depends on, demonstrating that a more conservative uncertainty-handling method need not sacrifice the computational tractability that makes large-scale planning feasible in the first place.

A further relationship emerges between Majzoobi and Khodaei [30] and the remaining two papers in this cluster. Majzoobi and Khodaei's finding that prevailing market structures impose minimum bid sizes often around 1 MW that exclude smaller individual systems provides an implicit rationale for the aggregation strategies pursued by Chen et al. [5] and Wang and Yuan [47]. Chen et al. respond to this constraint by pooling a residential and an industrial microgrid, whose complementary load profiles allow the combined entity to clear the capacity threshold that neither could reach alone. Wang and Yuan pursue a structurally similar solution through a different mechanism, coordinating a heterogeneous EV fleet alongside PV and wind within a single microgrid, such that large capacity vehicles supply the bulk of the system's grid-service contribution while smaller vehicles contribute more modestly and unpredictably around the margins. Considered independently, these two papers might appear to address unrelated case studies read against Majzoobi and Khodaei's market access finding, they instead appear as two distinct technical responses to the same structural barrier.

Taken as a whole, this cluster suggests that the technical challenge of providing fast, reliable ancillary services from electrical assets is comparatively well understood, while the more consequential and less resolved question is how individually modest resources are sized, planned, or pooled to participate in markets that were not originally designed with their scale in mind. Of the nine papers, Kassab et al. [24] offers the clearest model for addressing this question rigorously.

Integrated Electricity and Heat: The Inertia Providers

This cluster contains only two papers and reading them together reveals less a divide than a progression, Ji et al. [22] picks up the coordination problem Zhang et al. [49] establishes and confronts it under conditions Zhang's setting does not yet have to face. Zhang et al. demonstrate that heat pump thermal inertia can be coordinated with wind generation to provide both following and regulating reserves within a day-ahead stochastic framework, and their inactive transmission constraint identification method keeps this two carrier coordination problem computationally tractable. Ji et al. extend the same underlying logic to a three carrier system, adding gas to the electricity heat coupling, and in doing so confront a harder uncertainty problem correlated wind volatility propagating simultaneously through three linked networks rather than two. It is this added complexity that appears to motivate their methodological choice where Zhang et al.'s stochastic programming approach assumes a fixed probability distribution is adequate, Ji et al. argue, implicitly through their use of distributionally robust optimization with ambiguity sets built directly from historical data, that a fixed distribution assumption becomes a liability once three carriers are coupled rather than two.

Neither paper quantifies a relationship that the broader thesis framework in Section 1.1.1 suggests should exist if a wider thermal comfort tolerance genuinely translates into a larger deliverable reserve, that relationship remains asserted rather than measured anywhere in this cluster. This is best read as an open question this study surfaces, not a finding either paper establishes. Taken as a whole, two papers spanning a clear methodological escalation from two-carrier stochastic coordination to three-carrier distributionally robust coordination, but with no second voice yet confirming, contesting, or extending either result. of the two, Ji et al. [22] stands out as the more methodologically rigorous contribution offers a transferable template for exactly the kind of evidence grounded reliability quantification this thesis argues is missing across the field in section 4.

Mobility-Based Flexibility: The Stochastic Aggregators

Across this cluster, the central observation is that "EV uncertainty" is not one problem but several distinct ones, and the papers divide according to which version of that uncertainty they address. Alipour et al. [2], Sarker et al. [41], and Ma [28] each isolate a different source. Alipour et al. characterize uncertainty in price, vehicle availability, and the probability of being called into the reserve market simultaneously, treating the aggregator's task as balancing low-price charging against reserve participation under all three at once. Sarker et al. narrow the focus to the compensation structure itself, showing that how an aggregator splits resources between energy and reserve markets depends directly on the degradation cost reimbursed to owners, a dimension Alipour et al.'s framework does not price explicitly. Ma isolates a third, more physical source of uncertainty the limited number of vehicles that can charge simultaneously, demonstrating that ignoring this constraint alone, independent of price or availability uncertainty underestimates total charging cost by roughly 20%Ṙead together, these three papers suggest that a fully realistic aggregator model would need to combine all three uncertainty types rather than the single dimension each paper treats in isolation.

A second pair of papers turns specifically to the bidirectional V2G case, where the act of discharging the battery to support the grid introduces a cost the scheduling only papers above do not model. El-Hendawi et al. [11] and Meng et al. [33] both incorporate degradation explicitly into their optimization, but solve fundamentally different problems with it. El-Hendawi et al.'s centralized versus distributed comparison, validated against real Ontario grid data, is primarily a scalability and privacy argument, showing a distributed ADMM-based architecture matches centralized performance without its single point of failure risk. Meng et al., by contrast, use degradation-aware bidding as one input to a bi-level robust framework whose real contribution is coupling the spot and frequency regulation markets under a single worst case guarantee. The two papers are therefore not competing solutions to the same problem, El-Hendawi et al. solve a deployment problem, Meng et al. solve a reliability guarantee problem, and a deployable, contractually reliable aggregator would plausibly need elements of both.

A final group looks past frequency regulation entirely toward constraints the papers above do not consider. De la Torre et al. [10] embed network security constraints directly into the aggregator's bidding problem, preventing bids that would be optimal on paper from creating local congestion in practice, a risk none of the uncertainty focused papers above address. Nespoli et al. [36] extend this network awareness to distribution level voltage support at national scale, and their finding that feeder end clusters provide disproportionate voltage support value gives De la Torre et al.'s congestion avoidance logic a complementary, spatially specific counterpart. Zheng et al. [50], meanwhile, address a risk dimension neither technical paper considers at all, showing that a risk-neutral aggregator minimizing expected cost remains financially exposed to price volatility that a CVaR-based strategy substantially reduces. Technical feasibility and financial viability, in other words are shown here to be genuinely separate constraints rather than two faces of the same uncertainty.

Taken as a whole this cluster suggests that EV aggregation research has matured well beyond simple day-ahead scheduling, but has done so by fragmenting into several largely separate sub problems, behavioral uncertainty, degradation, network impact, and financial risk, that no single paper in this cluster yet unifies. Among these eight contributions, Meng et al. [33] offers the clearest structural model for how that unification might eventually look its bi-level architecture cleanly separates the market coupling layer from the uncertainty handling layer, a design choice that keeps two genuinely different problems, what to bid and how much risk to tolerate, mathematically distinct rather than entangled. A comparable structure, extended to also incorporate the network and financial dimensions raised elsewhere in this cluster, would be a natural direction for the kind of integrated reliability framework this thesis calls for in section 4.

Multi-Vector and Power-to-X Systems: The Synergy Cluster

This is the smallest cluster in the review, three contributions and that thinness is itself part of the finding. Contreras et al. [8] address a planning-stage question the single-vector clusters mostly treat as secondary, how much ancillary service capacity a system planner should sacrifice cost-efficiency to obtain, and their probabilistic multi-objective approach shows the two goals are not as opposed as a purely cost driven planner might assume. It is worth being precise about what this paper contributes to a cluster ostensibly about multi-vector synergy, its system contains no hydrogen, heat, or mobility asset at all, so its relevance here is demonstrating the value of true multi-objective planning, rather than evidentiary for cross carrier coordination specifically.

Čović et al. [9] provide this cluster's only genuine physical multi-carrier evidence, and their central finding carries a caution rather than a celebration unlike a standalone battery, whose operation affects only the electrical system, a power-to-hydrogen and hydrogen-to-power plant's operation affects both the power grid and the gas network simultaneously so an unscheduled balancing action taken for one system's benefit creates an imbalance in the other. Their result that ignoring this cross system effect increases total cost is best read as evidence that multi-vector coordination is harder and more failure prone than the conceptual case for it suggests, not as confirmation that it straightforwardly outperforms single vector alternatives. Su et al. [44] close the cluster at a different altitude entirely, synthesizing the broader coordination literature including the EV cluster discussed above, into a unified taxonomy whose central claim, that heterogeneous resource pooling allows individual carriers' weaknesses to offset one another, is conceptually well aligned with Čović et al.'s finding but is not itself new empirical evidence; it is a review level argument resting on the same thin evidentiary base this cluster as a whole exhibits.

Taken as a whole, this cluster represents the field's most conceptually promising yet least tested direction the case for multi-vector flexibility pooling is clearly and consistently articulated across all three papers, though empirical validation remains concentrated in a single physical case study, with multi-objective planning and

review-level synthesis providing valuable supporting context rather than independent confirmation in their own right. Of the three, Čović et al. [9] is the clearest model for how future work in this space should be structured.

4. Research Gap And Reserach Agenda

The cluster-by-cluster synthesis in Section 3 surfaces five recurring limitations that, considered together, define the gap between the current state of MES ancillary service research and the requirements of bankable, real-world explicit provision. Each subsection below states one limitation and the corresponding direction this thesis proposes for addressing it.

The Operational Risk and "Penalty-Free" Simulation Gap

Across all four clusters, the overwhelming majority of reviewed papers optimize for technical feasibility or expected profit without quantifying the financial consequence of failing to deliver a committed service. The reliability focused papers in section 3.1, Bahloul et al. [3] and Santosuosso et al. [40], are characteristic of this pattern both demonstrate how to maintain a reliability buffer or minimize degradation risk, but neither prices what happens if that buffer is breached. Closing this gap does not require inventing a new mathematical tool, since one already exists within the reviewed literature Zheng et al. [50] discussed in Section 3.3, demonstrate that a CVaR-based framework can price and manage an aggregator's financial exposure to market volatility within a tractable stochastic LP. The open task is redirecting this existing methodology, applying a CVaR-style formulation to the specific financial risk of ancillary service non compliance penalties, which no paper in the reviewed corpus currently does shifting the field's modeling paradigm from signal following to genuinely risk-managed bidding. a CVaR-based ancillary service bidding strategy exists in very recent literature (2026), confirming this direction is real but still emerging rather than established [46].

The "Siloed" Optimization and Cross-Vector Interaction Gap

Despite the integrated definition of MES, genuine cross carrier coupling is rare in the reviewed literature, and where it does appear it proves harder to model correctly than the conceptual case for flexibility pooling suggests. As established in section 3.4, only Čović et al. [9] model a system where one carrier's operation physically affects another balancing obligation, and their finding is cautionary, ignoring this cross-system effect increases cost rather than realizing the pooling benefit assumed elsewhere in the literature. Contreras et al. [8], the cluster's other ostensibly multi-vector contribution, was shown to contain no hydrogen, heat or mobility asset at all leaving the field's evidentiary base for cross vector interaction resting on a single case study. The path forward is to extend Čović et al.'s discipline, isolating and quantifying a specific cross-system failure mode, to other carrier combinations, electricity-heat, heat-hydrogen, and so on, rather than continuing to assume that the pooling benefit demonstrated for one plant configuration generalizes automatically across the field.

Linearization Versus Non-Linear Physical Reality

To preserve computational tractability the reviewed literature relies on linear or piecewise approximations of processes that are physically non-linear, most visibly asset degradation. El-Hendawi et al. [11] discussed in Section 3.3, illustrate both the value and the current limit of this approach their model incorporates a degradation cost into the V2G optimization, but through a depth-of-discharge-based formula rather than a full non-linear electrochemical aging model, meaning even one of the cluster's more degradation-conscious papers still simplifies the underlying physics considerably. The same pattern holds for the ageing-aware scheduling in Bahloul et al. [3] where calendar and cycling effects enter as model inputs rather than emergent state dependent processes. Closing this gap requires embedding genuinely non-linear aging and degradation models directly into the optimization loop, without sacrificing the computational tractability that makes large-scale, multi-asset scheduling feasible in the first place, a balance Iuliano and Vaccaro's [21] affine-arithmetic approach in Section 3.1 suggests is achievable for uncertainty modeling and may offer a transferable technique for degradation modeling as well.

Neglect of Local Grid Constraints in Global Service Provision

Ancillary services are conventionally modeled and procured as global TSO products yet delivered through physical assets sitting on local distribution networks and most of the reviewed literature optimizes for the former while abstracting away the latter. Section 3.3 identified the clearest exception to this pattern De la Torre et al. [10] embed network security constraints directly into the EV aggregator's bidding problem specifically to prevent globally optimal bids from creating local congestion, and Nespoli et al. [36] extend this network-awareness to distribution-level voltage support at national scale. That these remain a distinct, smaller sub-group within their own cluster, rather than standard practice followed throughout the literature, is itself evidence of the gap. What is needed is a generalized framework that extends this local constraint aware practice beyond the EV cluster to the other carriers reviewed in this thesis, alongside a formal decision hierarchy specifying how a local constraint, such as a transformer approaching its thermal limit, should qualify or override a global market commitment.

The Absence of a Universal Reliability Metric for Explicit Bidding

This highlighted as the most significant which is that reliability is defined and constrained differently depending on the carrier involved a battery's reliability is SOC limited, as shown throughout section 3.1, a thermal system's reliability is comfort limited, as section 3.2 establishes for Zhang et al. [49] and Ji et al. [22], and an EV fleet's reliability is behavior limited as the uncertainty characterization papers in section 3.3 demonstrate. No paper in the reviewed literature proposes a way to normalize these fundamentally different physical constraints into a single comparable measure that a TSO could use to evaluate a heat pump and a battery in equal terms.

Three methodological elements two from within this review and one from the broader literature point toward how such a metric might eventually be constructed. Within the reviewed corpus Ji et al.'s [22] use of a dirichlet process mixture model to build ambiguity sets directly from historical data rather than assumed distributions, and Meng et al.'s [33] bi-level architecture separating market commitment from uncertainty handling together suggest a feasible structure carrier-specific ambiguity sets feeding into a shared bi-level reliability-scoring layer. This direction is independently supported outside the reviewed corpus a geometric approach to DER flexibility proposes mathematically homogenizing the flexibility domains of fundamentally different DER types into a shared representation [45], while Riaz and Mancarella, the same authors behind the Feasible Operating Region concept introduced in section 1.1, separately propose capacity, ramp, duration, and cost as a consistent set of flexibility metrics applicable across heterogeneous DER types [39]. Neither external paper extends this to a TSO facing reliability score in the specific sense this gap calls for, but both confirm that the underlying normalization problem is recognized and being actively worked on elsewhere in the field, reinforcing that the agenda proposed here is a feasible and timely direction rather than a speculative one.

Digital Twins and Cyber-Physical Resilience

The final direction does not map to a specific gap demonstrated in the reviewed literature. It is included as a forward looking extension of this research agenda rather than a direct response to an established limitation. Real time digital twin technology, capable of running shadow simulations that anticipate a failure point fifteen to thirty minutes before it occurs would extend the predictive capability of every preceding item in this agenda. Similarly as multi-vector aggregation becomes increasingly dependent on continuous communication between dispersed assets and a central controller, how an aggregator remains operationally compliant during a communication lag or cyber incident becomes a meaningful currently unexamined dimension of grid reliability. one this study flags as worth investigating even though no paper in the reviewed corpus addresses it.

Conclusion

This thesis has provided a systematic literature review of explicit ancillary service provision in MES, synthesizing twenty-two studies through a four dimensional classification framework spanning energy carrier, ancillary service type, optimization technique, and market participation context. The review's central technical finding is that the choice of modeling paradigm is not arbitrary but follows a consistent physical logic deterministic optimization remains the standard for infrastructure sizing and feasibility assessment, where computational tractability matters more than capturing tail-risk, while stochastic and robust frameworks become necessary precisely where reliability requirements are strictest and asset behavior is least predictable, most clearly in the mobility cluster, where human behavior introduces the most severe form of uncertainty encountered across the review.

The cluster-by-cluster synthesis in Section 3 further shows that no single energy carrier dominates across all ancillary service types. Electrical assets remain the only resources reviewed that are technically capable of meeting the sub-second response demanded by frequency containment reserves, while integrated electricity-heat systems and the limited multi-vector evidence available point toward genuine, if not yet fully proven, potential for sustained, longer-duration services. This last point deserves particular emphasis while the conceptual case for multi-vector flexibility pooling is compelling and consistently articulated across the reviewed literature, this review found that its empirical validation currently rests on a notably thin evidentiary base, a single physical case study rather than a body of converging results. This is not a weakness of the multi-vector concept itself, but a finding in its own right, and one that meaningfully shapes the research agenda this thesis proposes.

That agenda follows directly from five recurring gaps identified through the cluster synthesis the limited treatment of financial risk in non-compliance scenarios, the persistence of fragmented, carrier-isolated optimization despite the conceptual promise of cross-vector synergy, the reliance on linear approximations for inherently non-linear degradation processes, the neglect of local distribution-level constraints in service provision conventionally modeled at the transmission level, and, most significantly, the absence of a universal reliability metric capable of comparing the fundamentally different physical constraints, SOC limits, thermal comfort bounds, behavioral uncertainty, that govern reliability across electrical, thermal, and mobility-based assets. This last gap is treated as the most consequential of the five, since without a shared basis for comparison, a TSO has no principled way to evaluate a heat pump and a battery on equal terms, which in turn limits how confidently heterogeneous MES portfolios can be treated as bankable reserve products.

The research agenda proposed in response to these gaps is not speculative for each gap, this review identified either a methodological precedent already present within the reviewed corpus or independent, recent confirmation from the broader literature that the same problem is being actively investigated elsewhere, from CVaR-based risk-aware bidding to geometric approaches for normalizing heterogeneous DER flexibility. This convergence suggests the agenda proposed here reflects a genuine and timely direction for the field rather than an isolated assessment.

Beyond its primary academic contribution, this classification and synthesis also offers practical value by mapping which carriers and optimization paradigms are best suited to which service types and reliability requirements, it provides a structured starting point for system operators, aggregators, and investors evaluating Multi-Energy System architectures for ancillary service participation. Ultimately, this thesis argues that the transition toward reliable, bankable explicit ancillary service provision in MES depends less on further isolated technical advances within individual carriers, which this review finds are already comparatively mature, than on the development of the cross-cutting tools, risk-aware bidding frameworks, validated cross-vector coordination, and a universal reliability metric chief among them, needed to make heterogeneous, multi-carrier flexibility trustworthy to the system operators who must ultimately rely on it.

References

[1] Mohammed H Albadi and Ehab F El-Saadany. A summary of demand response in electricity markets. *Electric Power Systems Research*, 78(11):1989–1996, 2008.

[2] Manijeh Alipour, Behnam Mohammadi-Ivatloo, Mohammad Moradi-Dalvand, and Kazem Zare. Stochastic scheduling of aggregators of plug-in electric vehicles for participation in energy and ancillary service markets. *Energy*, 118, 2016.

[3] Mohamed Bahloul, Mohamed Daoud, and Shafi K Khadem. Optimal dispatch of battery energy storage for multi-service provision in a collocated pv power plant considering battery ageing. *Energy*, 293:130744, 2024.

[4] Benjamin Bloess, Wolf-Peter Schill, and Alexander Zeja. Power-to-heat for renewable energy integration: A review of technologies, modeling approaches, and flexibility potentials. *Applied Energy*, 212:1611–1626, 2018.

[5] Yu Chen, Zhen Li, S Yu Samson, Bin Liu, and Xi Chen. A profitability optimization approach of virtual power plants comprised of residential and industrial microgrids for demand-side ancillary services. *Sustainable Energy, Grids and Networks*, 38:101289, 2024.

[6] Simon Clegg and Pierluigi Mancarella. Storing electricity in gas networks: A modeling framework for integrated gas and electricity systems. *IEEE Transactions on Power Systems*, 30(5):2621–2630, 2015.

[7] Simon Clegg and Pierluigi Mancarella. Integrated modeling and assessment of the operational impact of power-to-gas (p2g) on electrical and gas networks. *IEEE Transactions on Sustainable Energy*, 7(1):330–342, 2016.

[8] Sergio F Contreras, Camilo A Cortes, and Johanna MA Myrzik. Optimal microgrid planning for enhancing ancillary service provision. *Journal of Modern Power Systems and Clean Energy*, 7(4):862–875, 2019.

[9] Nikolina Čović, Ivan Pavić, and Hrvoje Pandžić. Multi-energy balancing services provision from a hybrid power plant: Pv, battery, and hydrogen technologies. *Applied energy*, 374:123966, 2024.

[10] S De la Torre, JA Aguado, and E Sauma. Optimal scheduling of ancillary services provided by an electric vehicle aggregator. *Energy*, 265:126147, 2023.

[11] Mohamed El-Hendawi, Zhanle Wang, and Xiaoyue Liu. Centralized and distributed optimization for vehicle-to-grid applications in frequency regulation. *Energies*, 15(12):4446, 2022.

[12] ENTSO-E. Electricity balancing operational guidelines and telemetry rules. Technical report, European Network of Transmission System Operators for Electricity, Brussels, Belgium, 2019.

[13] ENTSO-E. Inertia and its grid operational aspects. Report, European Network of Transmission System Operators for Electricity, October 2020.

[14] Martin Geidl, Göran Koeppel, Patrick Favre-Perrod, B Klockl, Göran Andersson, and K Fröhlich. The energy hub—a powerful concept for future energy systems. *Third International Conference on Critical Infrastructures*, pages 1–7, 2007.

[15] Hans Christian Gils et al. Integrated modeling of variable renewable energy integration through sector coupling and energy storage systems. *Energy Policy*, 102:160–173, 2017.

[16] Nicholas Good and Pierluigi Mancarella. Flexibility metrics and frameworks for multi-energy systems operation and security. *Proceedings of the IEEE*, 107(9):1923–1940, 2019.

[17] International Energy Agency. Meeting power system flexibility needs in china by 2030. Technical report, IEA, Paris, France, 2024.

[18] International Energy Agency. Electricity 2026: Flexibility. Technical report, IEA, Paris, France, 2026.

[19] International Renewable Energy Agency. Power system flexibility for the energy transition. Technical report, IRENA, Abu Dhabi, UAE, 2018.

[20] International Renewable Energy Agency. Demand-side flexibility for power sector transformation. Analytical brief, IRENA, Abu Dhabi, UAE, 2019.

[21] Silvia Iuliano and Alfredo Vaccaro. Affine arithmetic meets robust optimization in scheduling microgrids-based flexibility services. *Sustainable Energy, Grids and Networks*, page 101839, 2025.

[22] Daren Ji, Zhinong Wei, Yizhou Zhou, Sheng Research Chen, Guoqiang Sun, and Haixiang Zang. Distributionally robust co-optimization of energy and reserve dispatch for integrated electricity-gas-heating systems. *Sustainable Energy Technologies and Assessments*, 73:104125, 2025.

[23] Liang Ju full et al. Fast frequency response capabilities of multi-energy microgrids considering thermal and electrical storage technologies. *Applied Energy*, 179:1042–1056, 2016.

[24] Fadi Agha Kassab, Berk Celik, Fabrice Locment, Manuela Sechilariu, Sheroze Liaquat, and Timothy M Hansen. Microgrid sizing with ev flexibility: Cascaded milp and embedded apso-milp approaches. *Applied Energy*, 396:126273, 2025.

[25] Daniel S Kirschen and Goran Strbac. *Fundamentals of Power System Economics*. John Wiley & Sons, Chichester, UK, 2nd edition, 2018.

[26] Abhishek Kumar, Nand K Meena, Arvind R Singh, Yan Deng, Xiangning He, RC Bansal, and Praveen Kumar. Strategic integration of battery energy storage systems with the provision of distributed ancillary services in active distribution systems. *Applied Energy*, 253:113503, 2019.

[27] Prabha Kundur. *Power System Stability and Control*. McGraw-Hill, New York, 1994.

[28] Shuqing Ma. Optimal scheduling of electric vehicle aggregators for frequency regulation and cost efficiency in renewable-powered grids. *Computers and Electrical Engineering*, 118:109317, 2024.

[29] Louis Maeyaert, Lieven Vandevelde, and Thomas Döring. Battery storage for ancillary services in smart distribution grids. *Journal of Energy Storage*, 30:101524, 2020.

[30] Alireza Majzoobi and Amin Khodaei. Application of microgrids in providing ancillary services to the utility grid. *Energy*, 123:555–563, 2017.

[31] Pierluigi Mancarella. Mes (multi-energy systems): An overview of concepts, methods, and modeling approaches. *Energy*, 65:1–17, 2014.

[32] Brian V Mathiesen, Henrik Lund, David Connolly, et al. Smart energy systems for coherent 100% renewable energy and transport solutions. *Applied Energy*, 145:139–154, 2015.

[33] Yunfan Meng, Yonghui Sun, Liang Zhao, Chenxu Yin, and Fan Sheng. A hierarchical robust scheduling framework for electric vehicle aggregators in coupled spot and ancillary service markets. *Energy*, 332:137113, 2025.

[34] Federico Milano, Florian Dörfler, Gabriela Hug, David J. Hill, and Gregor Verbič. Foundations and challenges of low-inertia systems. In *2018 Power Systems Computation Conference (PSCC)*, pages 1–25. IEEE, 2018.

[35] Juan Miguel Morales, Antonio J. Conejo, Henrik Madsen, Pierre Pinson, and Marco Zugno. *Integrating Renewables in Electricity Markets: Operational Problems*, volume 205. Springer Science & Business Media, Berlin, Heidelberg, 2013.

[36] Lorenzo Nespoli, Nina Wiedemann, Esra Suel, Yanan Xin, Martin Raubal, and Vasco Medici. National-scale bi-directional ev fleet control for ancillary service provision. *Energy Informatics*, 6(Suppl 1):40, 2023.

[37] Fabian Ocker and Karl-Martin Ehrhart. Participation of demand side flexibility in european balancing markets: An overview of regulatory frameworks. *Energy Policy*, 117:433–446, 2018.

[38] Mark O'Malley et al. Operational boundaries and energy neutrality rules for automatic frequency restoration reserves. *IEEE Transactions on Power Systems*, 35(3):2130–2142, 2020.

[39] S. Riaz and P. Mancarella. Modelling and characterisation of flexibility from distributed energy resources. *arXiv preprint arXiv:2107.05144*, 2021.

[40] Luca Santosuosso, Simon Camal, Francesco Liberati, Alessandro Di Giorgio, Andrea Michiorri, and Georges Kariniotakis. Stochastic economic model predictive control for renewable energy and ancillary services trading with storage. *Sustainable Energy, Grids and Networks*, 38:101373, 2024.

[41] Mushfiqur R Sarker, Yury Dvorkin, and Miguel A Ortega-Vazquez. Optimal participation of an electric vehicle aggregator in day-ahead energy and reserve markets. *IEEE transactions on power systems*, 31(5):3506–3515, 2015.

[42] Pierluigi Siano. Demand response and smart grids—a survey. *Renewable and Sustainable Energy Reviews*, 30:461–478, 2014.

[43] Goran Strbac, Nikos Hatziargyriou, Danny Pudjianto, et al. Value of flexibility in sustainable and low-carbon electricity systems. *Proceedings of the IEEE*, 103(7):1163–1183, 2015.

[44] Chutian Su, Yi Wang, and Goran Strbac. Coordinated electric vehicles dispatch for multi-service provisions: A comprehensive review of modelling and coordination approaches. *Renewable and Sustainable Energy Reviews*, 223:115977, 2025.

[45] To Be Added. Approximating flexibility in distributed energy resources: A geometric approach. *arXiv preprint arXiv:1803.06921*, 2018.

[46] To Be Added. A conditional value-at-risk-based bidding strategy for pvss participation in energy and frequency regulation ancillary markets. *Sustainability*, 18(2):1122, 2026.

[47] PeiHua Wang and Jiaxin Yuan. Smart microgrid optimization with electric vehicle flexibility for intra-day energy management. *International Journal of Electrical Power & Energy Systems*, 172:111190, 2025.

[48] Yi Wang, Jing Qiufull, Goran Strbac, et al. Review of multi-energy system flexibility technologies and optimization methods. *IEEE Transactions on Power Systems*, 34(2):1141–1153, 2019.

[49] Menglin Zhang, Qiuwei Wu, Jinyu Wen, Bo Zhou, Qinyue Guan, Jin Tan, Zhongwei Lin, and Fang Fang. Day-ahead stochastic scheduling of integrated electricity and heat system considering reserve provision by large-scale heat pumps. *Applied Energy*, 307:118143, 2022.

[50] Yanchong Zheng, Yubin Wang, and Qiang Yang. Bidding strategy design for electric vehicle aggregators in the day-ahead electricity market considering price volatility: A risk-averse approach. *Energy*, 283:129138, 2023.

Table 3: Nomenclature and List of Abbreviations

Abbreviation	**Indication / Description**
ACE	Area Control Error
ADMM	Alternating Direction Method of Multipliers
aFRR	Automatic Frequency Restoration Reserves
AGC	Automatic Generation Control
BESS	Battery Energy Storage System
CCG	Column and Constraint Generation
CHP	Combined Heat and Power
DAM	Day-Ahead Market
DER	Distributed Energy Resources
DRO	Distributionally Robust Optimization
DSO	Distribution System Operator
EV	Electric Vehicle
FCR	Frequency Containment Reserves
FFR	Fast Frequency Response
FOR	Feasible Operating Region
G2V	Grid-to-Vehicle
HP	Heat Pump
IBR	Inverter-Based Resources
IEH	Integrated Electricity and Heat
LCOE	Levelized Cost of Electricity
MES	Multi-Energy System
mFRR	Manual Frequency Restoration Reserves
MILP	Mixed-Integer Linear Programming
MPC	Model Predictive Control
NPV	Net Present Value
P2G	Power-to-Gas
P2H	Power-to-Heat
PCC	Point of Common Coupling
PV	Photovoltaic
RES	Renewable Energy Sources
RoCoF	Rate of Change of Frequency
RTM	Real-Time Market
SM / SP	Stochastic Model / Stochastic Programming
SoC	State of Charge
SoCH	State of Charge of Heat
SoH	State of Health
TES	Thermal Energy Storage
TSO	Transmission System Operator
V2G	Vehicle-to-Grid
VPP	Virtual Power Plant

Abstract in lingua italiana

Questo studio presenta una rassegna sistematica della letteratura sulla fornitura esplicita di servizi ausiliari nei sistemi multi-energia (MES), motivata dalla crescente penetrazione di risorse stocastiche basate su inverter e dal corrispondente declino dell'inerzia rotante nei moderni sistemi di potenza. La letteratura viene classificata secondo quattro dimensioni: il vettore energetico (elettricità, calore, mobilità e idrogeno), il tipo di servizio ausiliario (servizi di frequenza e non di frequenza), la tecnica di ottimizzazione (ad esempio, la programmazione lineare intera mista - MILP, la programmazione stocastica a due stadi, l'ottimizzazione robusta) e il contesto di partecipazione al mercato (mercati day-ahead e real-time). Ventidue studi vengono analizzati per identificare modelli metodologici ricorrenti, tra cui la prevalenza dell'ottimizzazione deterministica per il dimensionamento delle infrastrutture e la crescente adozione di framework sensibili all'incertezza (uncertainty-aware) per le offerte vincolate dall'affidabilità, e per dimostrare come il pooling della flessibilità tra diversi vettori energetici possa mitigare i limiti fisici dei sistemi a singolo vettore. Sulla base di questa sintesi, vengono identificati cinque gap di ricerca ricorrenti: la limitata trattazione del rischio finanziario negli scenari di non conformità, la persistenza di ottimizzazioni frammentate e isolate per singolo vettore nonostante la promessa concettuale della sinergia cross-vettore, l'affidamento ad approssimazioni lineari per processi di degradazione intrinsecamente non lineari, la negligenza dei vincoli locali a livello di distribuzione nella fornitura di servizi convenzionalmente modellati a livello di trasmissione, e l'assenza di una metrica di affidabilità universale in grado di confrontare le prestazioni eterogenee di asset elettrici, termici e legati alla mobilità. In risposta a ciò, la tesi propone un corrispondente programma di ricerca incentrato sull'ottimizzazione sensibile al rischio, su un coordinamento cross-vettore più profondo, su una modellazione fisica e regolatoria ad alta fedeltà e sulla standardizzazione di una metrica di affidabilità universale, al fine di guidare i lavori futuri verso una fornitura esplicita di servizi ausiliari che sia bancabile, sicura e consapevole dell'affidabilità nei sistemi multi-energia.

Parole chiave: Sistemi Multi-Energia (MES), Fornitura Esplicita, Fornitura Implicita, MILP, Framework Sensibili all'Incertezza.

Acknowledgements

I would like to express my deepest gratitude to God Almighty for the strength, perseverance, and guidance provided to me throughout this journey and for making the successful completion of this work possible. My thoughts and prayers go out to my fellow Sudanese citizens who are currently enduring the profound hardships and suffering brought about by the ongoing war. I dedicate this milestone to the resilience of my people and pray for peace and restoration in my homeland. I am sincerely grateful to Politecnico di Milano for granting me the opportunity to pursue my Master's degree in such a prestigious and competitive institution. The high quality of the academic material, the rigorous learning environment, and the international atmosphere have been instrumental in my growth as an engineer.

I would like to extend my heartfelt thanks to my advisor, Professor Filippo Bovera, for his invaluable guidance, academic support, and for providing the direction necessary to finalize this research. I also wish to thank my co-advisor, Andrea Scrocca, for his dedicated time and technical insights throughout the development of this thesis. Finally, my deepest love and appreciation go to my family: my father, Eltayeb Khogali, my mother, Egbal Mohammed, and my sister, Waad. Their unwavering belief in me and their sacrifices have been my primary motivation. I am also eternally grateful to all my beloved family members and friends who have offered their support, encouragement, and presence throughout these years. .