

Comparative Study of Aging-Aware Control Strategies for Grid-Connected Photovoltaic Battery Systems

Huang Zhang, Faisal Altaf, and Torsten Wik

Abstract—Various strategies with different objectives have been proposed to control grid-connected photovoltaic (PV) battery systems where electric vehicle (EV) batteries can be used as stationary energy storage. As the first attempt to enable aging-aware decision-making under various uncertainties, an economically motivated stage cost function is proposed to account for both the grid and the battery degradation cost. Historical operational data and “fixed” forecasted electricity price data are utilized to improve the economic performance of an implicit (or time-varying) optimal policy. Simulation results show that an implicit optimal policy achieved better economic performance (i.e., lowest grid and battery degradation cost) with smaller fluctuation amplitudes than an explicit one. Thus, to improve the aging-aware decision-making under uncertainties for EV batteries further, the implicit optimal policy will be further developed with consideration of other forecasts.

I. INTRODUCTION

To further improve the energy efficiency of transportation systems, reduce reliance on fossil fuels, and decrease carbon emissions, the commercialization of xEVs (i.e., all types of pure electric vehicles (BEVs), hybrid electric vehicles (HEVs), and plug-in hybrid electric vehicles (PHEVs)) continues accelerating in the global market despite supply chain disruptions, geopolitical uncertainty, and high energy prices [1]. An increasing number of xEV batteries can be expected to retire within an 8-10 year span after they reach the end of their first lives in vehicles (typically 70–80% of initial nominal capacity) [2]. At the same time, the necessary and increasing penetration of renewable energy sources into power systems requires urgent solutions to the variable electricity generation that results from the intermittent nature of renewable energy sources [3]. To address this, stationary battery energy storage systems (BESS) have gained a lot of interest due to their advantages of rapid response, and good scalability [3]. However, safe and optimal usage of stationary BESSs based on new and used xEV batteries still faces challenges, particularly concerning decision-making under different types of uncertainty [4] [5]. In residential grid-connected photovoltaic (PV) battery systems, the uncertainty

mainly arises from forecasts of load demand, PV power production, electricity prices, and battery degradation.

In the literature, different strategies for controlling residential PV battery systems have been proposed with different objectives. Two widely used rule-based strategies are the maximizing self-consumption (MSC) strategy and the time-of-use (ToU) strategy in which decisions are based on the PV-load difference and the electricity price [6]. The rule-based strategies are easy to implement but not capable of addressing uncertainties arising from the various aforementioned sources. In contrast, model predictive control (MPC) finds a feedback policy at each time instant by solving an optimal control problem in a receding-horizon fashion, which effectively reduces the impact of uncertainties [7]. For computational reasons, simple models are often preferred in the MPC scheme. As a result, the MPC-based solution is optimal for the given model, but it is often suboptimal for the real system having stochastic dynamics [8]. To address this issue, recent studies attempt to combine reinforcement learning (RL) and MPC in order to achieve economic optimality for the real system while retaining stability guarantees [9] [10] [11]. In practice, the MPC is used as a new type of function approximator within RL to support the approximations of the value function, action-value function, or policy [9]. RL techniques can then be used to tune the parameters of MPC to improve closed-loop performance [12]. However, the battery degradation cost, which is expected to have a significant impact on the economic feasibility of stationary BESSs based on new or used batteries, has not been considered in the aforementioned control strategies.

In this work, a residential grid-connected PV battery system is investigated with an economically motivated stage cost and stochastic dynamics. Specifically, our **key results and contributions** are summarized as follows:

- An economically motivated stage cost function is proposed to account for both the grid and the battery degradation cost. Notably, the battery degradation that accounts for both the calendar aging and cycling aging of a battery is incorporated into aging-aware control strategies for the first time.
- We demonstrate how information from the historical operational data and “fixed” forecasted electricity price data can be utilized to improve the economic performance of learned policies.

II. SYSTEM MODELING

The residential grid-connected photovoltaic (PV) battery system typically consists of a PV generation unit and a bat-

Corresponding author: Huang Zhang is with the Department of Electromobility, Volvo Group Trucks Technology, 405 08 Gothenburg, Sweden (mobile: +46739022691; email: huang.zhang@volvo.com) and with the Department of Electrical Engineering, Chalmers University of Technology, 412 96 Gothenburg, Sweden (email: huangz@chalmers.se)

Faisal Altaf is with the Department of Electromobility, Volvo Group Trucks Technology, 405 08 Gothenburg, Sweden (email: faisal.altaf@volvo.com)

Torsten Wik is with the Department of Electrical Engineering, Chalmers University of Technology, 412 96 Gothenburg, Sweden (email: torsten.wik@chalmers.se)

tery energy storage system (BESS). A general scheme of this residential grid-connected PV battery system is illustrated in Fig. 1. When PV production power is larger than the electric load, residual energy is used to charge the battery or sold to the grid. When PV production power is less than the electric load, the battery discharges, or the electricity is bought from the grid to satisfy the load demand.

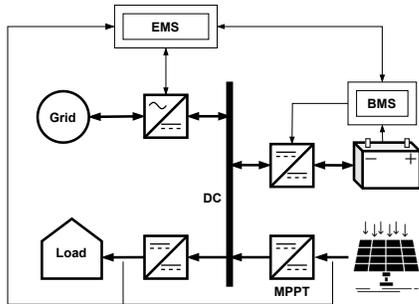


Fig. 1. A general scheme of the residential grid-connected PV battery system.

A. Battery Dynamics

Typically, battery open circuit voltage (OCV) is a nonlinear function of state-of-charge (SoC) and temperature. However, most BESSs are based on lithium iron ferrous phosphate (LFP)/graphit batteries, which have a nearly constant OCV for a wide range of SoC. Therefore, we assume that OCV is constant for the chosen SoC window, then there is a 1:1 relationship between SoC and state-of-energy (SoE). A simple BESS model is given by [4] [5]

$$\frac{d\text{SoC}}{dt} = \frac{1}{E_b} (P_{\text{grid}} + P_{\text{PV}} - P_{\text{load}}) \cdot \begin{cases} \eta, & \text{if } P_{\text{grid}} + P_{\text{PV}} - P_{\text{load}} > 0 \\ \frac{1}{\eta}, & \text{otherwise} \end{cases}, \quad (1)$$

where $\text{SoC} \in [0, 1]$ denotes the battery state of charge, E_b denotes the battery rated capacity in kWh, $P_{\text{grid}} \in [-P_{\text{grid}}^{\text{lim}}, P_{\text{grid}}^{\text{lim}}]$ is the power traded with the grid in kW, P_{PV} is the power generated by the PV unit, P_{load} is the power consumed by the household, and η is the battery round-trip efficiency, which is assumed to be constant.

Discretizing system dynamics using a first-order method with a sampling interval $h = 1$ hour gives the form

$$x(k+1) = x(k) + (u(k) + w(k)) \cdot \begin{cases} \eta, & \text{if } u(k) + w(k) > 0 \\ \frac{1}{\eta}, & \text{otherwise} \end{cases} \quad (2)$$

with

$$\begin{aligned} x &\triangleq \text{SoC}, \quad u \triangleq \frac{h}{E_b} P_{\text{grid}}, \quad u_{\text{lim}} \triangleq \frac{h}{E_b} P_{\text{grid}}^{\text{lim}}, \\ w &\triangleq \frac{h}{E_b} (P_{\text{PV}} - P_{\text{load}}), \end{aligned} \quad (3)$$

where $x \in [0, 1]$ is the SoC of the battery, $u \in [-u_{\text{lim}}, u_{\text{lim}}]$ is the control input that is the energy bought from ($u >$

0) or sold to ($u < 0$) the power grid normalized w.r.t. the battery E_b , and $w \sim \mathcal{N}(\mu, \sigma^2)$ is assumed to be Gaussian disturbance that is the residual energy normalized w.r.t. the battery E_b .

B. Battery Degradation

Battery degradation is a key cost driver during the operation of PV battery systems, and therefore cannot be neglected. Battery degradation consists of two parts, i.e., calendar aging and cycling aging. Calendar aging depends on stress factors, such as storage time, SoC, and temperature, while cycling aging depends on stress factors, such as temperature, current rate (C-rate), depth-of-discharge (DoD), and upper SoC operating limit [13]. The total battery degradation is therefore often expressed by a semi-empirical model that accounts for both calendar and cycling aging using the results of accelerated aging tests.

In this work, we assume a solely time-dependent calendar aging process as well as a charge throughput-dependent cycling aging model [14]. Specifically, the calendar aging at time step k is estimated as,

$$\beta_{\text{cal}}(k) = \frac{h}{L_{\text{cal}}}. \quad (4)$$

where L_{cal} is the number of years until the battery degrades to its EoL at a constant storage temperature without cycling aging. The cycling aging, which is assumed to depend on charge throughput, at time step k is given by,

$$\beta_{\text{cyc}}(k) = 0.5 \cdot \frac{|P_b(k)| \cdot h}{L_{\text{cyc}} \cdot E_b}, \quad (5)$$

where L_{cyc} is the number of equivalent full cycles (EFCs) until the battery reaches its end-of-life (EoL) without calendar aging, and P_b is the battery charge or discharge power. The factor 0.5 is to convert charge throughput to full cycle counting. Then the total battery degradation at time step k , using superposition is given by,

$$\beta_{\text{total}}(k) = \beta_{\text{cal}}(k) + \beta_{\text{cyc}}(k) \quad (6)$$

Accumulating the battery degradation up to a given time step K , the total aging becomes

$$\Psi_{\text{total}}(K) = \sum_{k=1}^K \beta_{\text{total}}(k). \quad (7)$$

In practice, BESS manufacturers often specify a BESS's lifetime in both calendar lifetime in years (L_{cal}) and cycling lifetime in EFCs (L_{cyc}) until the BESS reaches its end-of-life (EoL). The calendar lifetime (L_{cal}) provides a reference value for storage degradation to EoL at a constant storage temperature, while the cycling lifetime (L_{cyc}) provides a reference value for maximum charge throughput until EoL is reached at a controlled DoD. Consequently, $\Psi_{\text{total}}(K) = 0$ corresponds to a fresh battery never cycled, whereas $\Psi_{\text{total}}(K) = 1$ corresponds to a battery at its EoL. Battery usage beyond the EoL threshold (i.e., $\Psi_{\text{total}}(K) > 1$) might be allowed if the replacement threshold is below the EoL threshold, as further discussed in the next section.

Assuming that the EoL threshold is defined as 80% of the battery's rated capacity, which is commonly used for automotive applications [14], the battery state-of-health (SoH) at time step K can be expressed as

$$\text{SoH}(K) = 1 - 0.2 \cdot \Psi_{\text{total}}(K). \quad (8)$$

Hence, when the total battery degradation reaches 1 (i.e., $\Psi_{\text{total}}(K) = 1$), the SoH is equal to 0.8. The battery SoH change at time step k can therefore be expressed as

$$\Delta \text{SoH}(k) = 0.2 \cdot \beta_{\text{total}}(k). \quad (9)$$

III. PROBLEM FORMULATION

The objective for controlling residential grid-connected PV battery systems can be summarized as finding an optimal policy that minimizes the economic cost (i.e., grid cost and battery degradation cost) of operating the system without violating any constraints.

A. Cost Functions

This subsection presents economic stage cost functions, i.e., the grid cost, the battery degradation cost, and the SoC cost.

1) *Grid Cost Function*: The cost function that accounts for selling electricity to or buying electricity from the grid at time step k is given by

$$L_g(x(k), u(k)) = \begin{cases} \sigma_{\text{buy}} u(k), & \text{if } u(k) > 0 \\ \sigma_{\text{sell}} u(k), & \text{otherwise} \end{cases}, \quad (10)$$

where σ_{buy} is the electricity buying price, σ_{sell} is the electricity selling price, and it is assumed that $\sigma_{\text{buy}} \geq \sigma_{\text{sell}}$. Note that the electricity prices are assumed to be fixed in learning explicit optimal policies, but are forecasted in learning implicit optimal policies.

2) *Battery Degradation Cost Function*: The cost function that accounts for the battery degradation at time step k is given by

$$L_b(x(k), u(k)) = \frac{0.2 \cdot \beta_{\text{total}}(k)}{1 - \alpha} \cdot (c_b^i + c_b^r) \cdot E_b, \quad (11)$$

where α is the replacement threshold defined as 60% of the battery rated capacity (i.e., $\alpha = 0.6$), which is often used for less-demanding stationary storage applications. c_b^i and c_b^r denote the battery initial and replacement unit costs, respectively. The maintenance cost within the battery lifetime is negligible and therefore is not considered here [14] [15].

3) *Battery SoC Cost Function*: In order to prolong the battery lifetime as much as possible, the cost function penalizes the battery SoC if it is outside a specified SoC window $[\text{SoC}_{\text{min}}, \text{SoC}_{\text{max}}]$ with a factor c , i.e.,

$$L_s(x(k), u(k)) = c \cdot \max(x(k) - \text{SoC}_{\text{max}}, 0) + c \cdot \max(\text{SoC}_{\text{min}} - x(k), 0) \quad (12)$$

4) *Economic Stage Cost Function*: Given the cost functions with the different objectives stated above, the final economic stage cost can be expressed

$$L(x(k), u(k)) = L_g(x(k), u(k)) + L_b(x(k), u(k)) + L_s(x(k), u(k)). \quad (13)$$

Compared with other cost functions in the literature, our proposed one not only accounts for grid cost and battery degradation cost by penalizing buying electricity, charging and discharging the battery, and violating a specific SoC window, but also is capable of incorporating future predictions, such as "fixed" forecasted electricity price data in the case of learning implicit optimal policies.

B. Markov Decision Processes

In this subsection, we formulate an infinite-horizon Markov Decision Process (MDP) which seeks a deterministic explicit (or time-invariant) optimal policy, and a finite-horizon MDP which seeks an implicit (or time-varying) optimal policy.

1) *Infinite-Horizon MDPs*: Assuming that the electricity prices are fixed in Eqn (10) and the disturbance (or normalized residual energy) w in Eqn (2) is assumed to have a Gaussian distribution ($\mathbb{P}(w)$) with parameters estimated from historical data. Then the deterministic explicit optimal policy can be defined as

$$\pi^* = \arg \min_{\pi} \mathbb{E} \left[\sum_{k=0}^{\infty} \gamma^k L(x(k), u(k)) | u(k) = \pi(x(k)) \right], \quad (14)$$

where $\gamma \in (0, 1)$ is the discount factor, and the expected value is taken over the Markov Chain distribution of state trajectories under the policy π .

2) *Finite-Horizon MDPs with electricity price forecasts*: Now, assuming that the electricity prices are forecasted for the next 36 hours ($N = 35$) at 13:00 every day, and the disturbance in Eqn (2) is assumed to be a conditional probability distribution ($\mathbb{P}(w|H)$) estimated using historical data. Note that $H \in \{0, 1, \dots, 23\}$ is the hour in a day. Then the deterministic implicit optimal policy can be defined as,

$$\pi^* = \arg \min_{\pi = \{\pi_0, \dots, \pi_N\}} \mathbb{E} \left[T(x(N+1)) + \sum_{k=0}^N \gamma^k L(x(k), u(k), \sigma_e(k)) | u(k) = \pi(x(k), H(k)) \right], \quad (15)$$

where $T(x(N+1)) = \gamma^{N+1} V_{N+1}(x(N+1))$ accounts for long-term electricity wholesale market prediction, and $\sigma_e(k)$ is the forecasted electricity buying or selling price at the time step k . Note that the optimal infinite-horizon state-value function is assigned for $V_{N+1}(x(N+1))$, and the implicit optimal policy is computed at 13:00 every day when the electricity spot prices are released for the next 36 hours.

The explicit optimal policy in Eqn (14) is solved by using value iteration (also called dynamic programming) algorithm, and the implicit optimal policy in Eqn (15) is solved by

using policy iteration. With the discount factor $\gamma < 1$ ($\gamma = 0.99$ in this work), convergence to optimal explicit policy is guaranteed for any finite starting V as stage costs are bounded, but there is no "convergence" for implicit (or time-varying) policy as the state-value function V and implicit policy π remain time-varying. The convergence proof of value iteration and policy iteration can be found in Chapter 4.8 – Appendix: Mathematical Analysis of Reinforcement Learning and Optimal Control by Dimitri Bertsekas [16]. A stability theory for generalized MDPs has been proposed for the undiscounted case in Ref. [17], and for the discounted case in Ref. [18]. To guarantee rigorous stability in practice, an alternative approach that enforces stability constraints in discounted MDPs using MPC was proposed in Ref. [19]. We will enforce specific stability constraints using MPC to our discounted MDPs in the future.

IV. SIMULATION RESULTS

To validate the proposed strategy, approximately two-year measurement data from a residential grid-connected PV battery system for a housing association of 132 households in Gothenburg, Sweden, is used for learning and then the optimal policies are evaluated on another half-year data. This residential PV battery system consists of a stationary BESS that contains 14 lithium-ion battery packs retired from electric buses and a PV generation unit. The specifications of the stationary BESS are given in Table I, and other system parameters are summarized in Table II. Moreover, the economic parameters used in the simulation are given in Table III.

TABLE I
SPECIFICATIONS OF THE LFP/GRAPHITE BATTERY USED IN THE SIMULATION

Parameter	Value	Unit
Battery rated capacity (E_b)	200	kWh
Maximum charge/discharge power (P_b^{lim})	70	kW
Battery round-trip efficiency (η)	96	%
SoC window ($\text{SoC}_{\min} - \text{SoC}_{\max}$)	20-85	%
Calendar life (L_{cal})	13.5*	years
Cycle life (L_{cyc})	6000*	EFC

* The values correspond to 80% of the battery rated capacity, which is a typical EoL criterion for automotive applications.

TABLE II
OTHER SYSTEM PARAMETERS USED IN THE SIMULATION

Parameter	Value	Unit
PV peak power ($P_{\text{PV}}^{\text{max}}$)	170.8	kWp
Grid power limit ($P_{\text{grid}}^{\text{lim}}$)	100	kW
Gaussian disturbance mean (μ)	-0.2074*	-
Gaussian disturbance standard deviation (σ)	0.10*	-

* The values are estimated using the first two-year disturbance data.

The first two-year data (2021-06-07 16:00:00 - 2023-06-08 15:00) is used for estimating the probability distribution of the disturbance in Eqn (2), $\mathbb{P}(w)$ or $\mathbb{P}(w|H)$, and the

TABLE III
ECONOMIC PARAMETERS USED IN THE SIMULATION

Parameter	Value	Unit
Electricity selling price (σ_{sell})	0.122*	EUR/kWh
Electricity buying price (σ_{buy})	0.178*	EUR/kWh
Battery replacement threshold (α)	60	%
Battery initial unit cost (c_b^i)	463**	EUR/kWh
Battery replacement unit cost (c_b^r)	413**	EUR/kWh
SoC penalty weight (c)	1	EUR

* The electricity prices include spot prices, taxes, and transmission fees.

** The values are taken from Ref. [15].

remaining half-year data (2023-06-08 16:00:00 - 2023-12-31 23:00:00) is used for evaluating the performance of learned explicit and implicit optimal policies.

A. Explicit Optimal Policy

In order to learn the explicit optimal policy, the disturbance in Eqn (2) is assumed to be IDD Gaussian distribution ($\mathbb{P}(w)$) and first estimated using the first two-year data. The resulting parameters are given in Table II.

With the disturbance assumption (i.e., IDD Gaussian distribution) and the electricity price assumption (i.e., fixed), the explicit optimal policy together with its corresponding state-value function are illustrated in Fig. 2. Note that it only takes 8 iterations until the policy converges (termination criterion $\|\pi^* - \pi\|_{\infty} \leq 10^{-3}$). However, additional iterations are needed for the state-value function to converge as well. It can be observed from Fig. 2 that the optimal policy is to sell electricity to the grid when the battery SoC is more than 90%, buy electricity from the grid when the battery SoC is less than 75%, and stop trading with the grid when the battery SoC is 75-90%. Correspondingly, the optimal state-value function decreases monotonically with increasing battery SoC from 0% to about 85%, and then increases with the battery SoC from 85% to 100%. In particular, the maximum value of the optimal state-value function is achieved at $x = 0$, which can be rationalized, since it is required to buy electricity from the grid to charge the battery as its SoC is far below 20%, and also satisfy possible residual load demand. The minimum value of the optimal state-value function is achieved at about $x = 0.85$, which can also be rationalized, i.e., no trading with the grid translates to no grid cost, the battery SoC is within the desired 20-85% window, and the battery can be discharged to satisfy possible residual load demand.

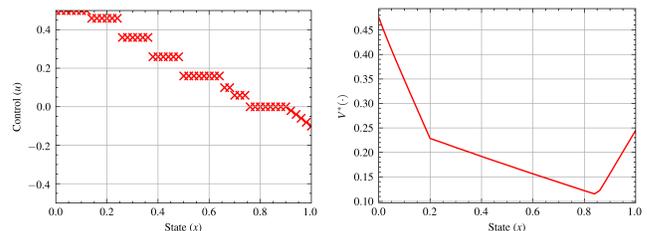


Fig. 2. The explicit optimal policy and state-value function.

The explicit optimal policy is then evaluated on the remaining half-year data, and its resulting battery SoC as a function of time is illustrated in Fig. 3. The initial battery SoC is set to 50%, and it can be observed that the battery SoC first increases to about 70% and then decreases to a stochastic steady-state at about 20%. The explicit optimal policy gradually pushes the battery SoC towards the lower bound of the specified SoC window over the remaining half year to guarantee availability for PV charge. Note that the PV production power gets less and less when the time approaches the end of the remaining half year (2023-12-31 23:00:00) as the sunlight significantly decreases in the winter. However, this long-term increase in demand has to be satisfied by the grid.

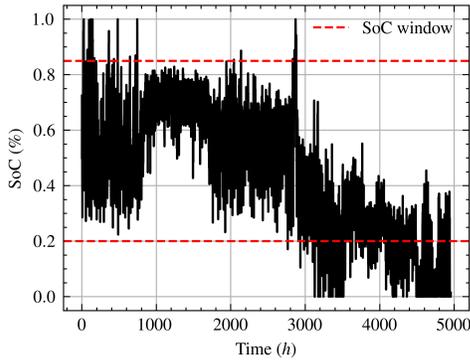


Fig. 3. The battery SoC as a function of time under the explicit optimal policy over the remaining half year.

B. Implicit Optimal Policy

In order to learn the implicit optimal policy, the disturbance in Eqn (2) is assumed to be a conditional probability distribution ($\mathbb{P}(w|H)$) estimated using the first two years of data. The resulting distribution of $\mathbb{P}(w|H)$ is illustrated in Fig. 4. It can be observed that the highest probability of disturbance is between 00:00 - 06:00 during a day over the first two years. It can be rationalized that there is almost zero PV production and low but stable load demand at the same time between 00:00-06:00 during the day. Therefore, the residual energy (or disturbance after normalization) is negative between 00:00 - 06:00. Moreover, it can also be observed that there are positive disturbance values between 07:00 - 19:00 with its maximum value achieved at around 13:00 during the day over the first two years. It can also be rationalized that the PV production peaks while load demand is relatively low (residential case) at around 13:00. Therefore, the residual energy achieves a maximum at around 13:00 during the day.

With the disturbance assumption (i.e., an empirical probability distribution) and the electricity price assumption (i.e., 36 hours forecasted), the implicit optimal policy together with its corresponding state-value function for the first 24 hours are illustrated in Fig. 5 and Fig. 6, respectively. Note that the implicit policy is computed at 13:00 every day for the next 36 hours. However, only the first 24 policies

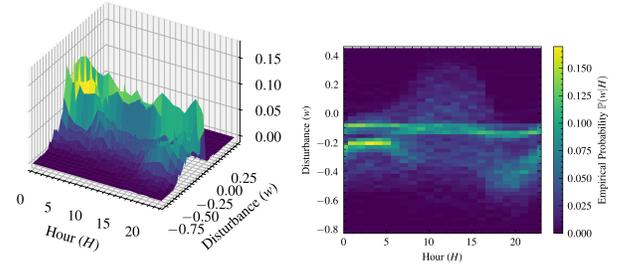


Fig. 4. The probability distribution of disturbance ($\mathbb{P}(w|H)$) estimated using first two-year data in a 3D plot (left) and a top-down 2D plot (right).

are used before it is updated again as the new electricity spot prices are released. It can be observed from Fig. 5 that given an hour in the day, the optimal policy is to buy electricity from the grid when the battery SoC is lower than around 40% and sell electricity to the grid when the battery SoC is higher than around 80% throughout the day. Interestingly, the battery SoC level that separates selling or buying electricity decisions significantly shifts downwards approximately between 09:00 - 17:00 during the day. The significant downward shift is mainly contributed by the peaking of PV production between 09:00 - 17:00 during the day (see Fig. 4). Corresponding, the state-value function in Fig. 6 indicates that given an hour in the day, higher values are achieved when the battery SoC is higher than about 85% or lower than about 20%, while given a battery SoC, lower values are achieved at later hours.

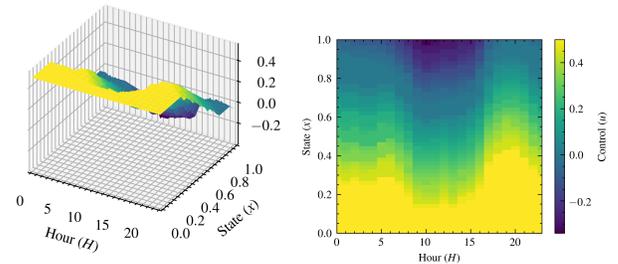


Fig. 5. The implicit optimal policy for the first 24 hours in a 3D plot (left) and a top-down 2D plot (right).

The implicit optimal policy is also evaluated on the remaining half-year data, and its resulting battery SoC as a function of time is illustrated in Fig. 7. The initial battery SoC is set to 50% as well, and it can be observed that the battery SoC first increases to about 70% and then decreases to about 25%. The SoC dynamics under the implicit optimal policy exhibit a similar long-term trend to that under the explicit optimal policy but with smaller fluctuation amplitudes (see Fig. 3) and less operation outside the desired SoC window. In this regard, the implicit optimal policy performs slightly better than the explicit one.

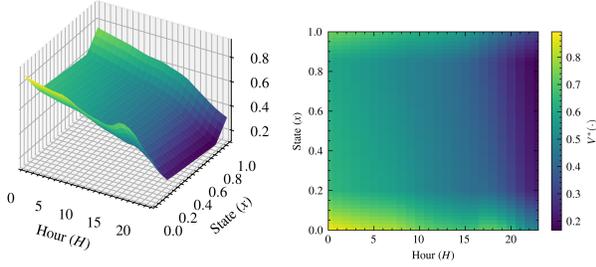


Fig. 6. The state-value function of the implicit optimal policy for the first 24 hours in a 3D plot (left) and a top-down 2D plot (right).

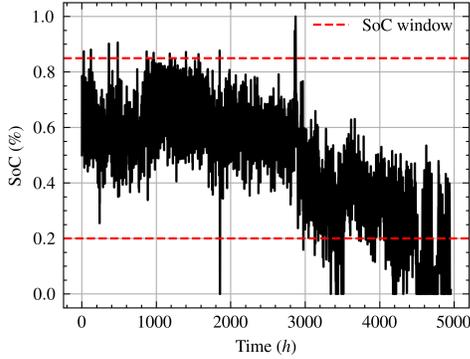


Fig. 7. The battery SoC as a function of time under the implicit optimal policy over the remaining half year.

C. Policy Economic Performance Evaluation

Lastly, we evaluate the economic performance of the learned explicit and implicit optimal policies and then compare them with that of the most commonly used strategy, i.e., the rule-based (MSC) strategy. Its basic principle is when PV production is larger than the load demand, the residual energy charges the battery first, and is then sold to the grid if there is any left; when PV production is less than the load demand, the battery discharges first to satisfy the load demand, and then electricity will be bought from the grid if there is still unsatisfied load demand [4]. The resulting economic cost over the remaining half year is summarized in Table IV. The total cost over the remaining half year is divided into the grid cost calculated using Eqn (10) and the battery degradation cost calculated using Eqn (11). The implicit optimal policy achieves the lowest total cost (34779.9 EUR), while the rule-based (MSC) strategy implemented in the residential PV battery system achieves the highest total cost (44772.0 EUR) over the half-year. Moreover, the rule-based (MSC) strategy surprisingly achieves the lowest battery degradation cost, accounting for only 0.86% degradation over the half year, among all the control strategies. To investigate the reason for this, the battery SoC as a function of time under the rule-based (MSC) strategy over the half year is illustrated in Fig. 8. It can be observed that the battery was not used much compared to the explicit and implicit policies over the half-year simulation. Under the rule-based (MSC) strategy,

the battery was quickly discharged first to satisfy the residual demand and then its SoC was kept at the lower bound of the desired SoC window (20-85%) most of the time during the day, until the PV production peaks next time.

To summarize, with the disturbance assumption (i.e., an empirical probability distribution) and the electricity price assumption (i.e., forecasted), the implicit optimal policy achieves the best economic performance (i.e., lowest grid and battery degradation cost) with smaller fluctuation amplitudes than the explicit optimal policy over the half-year simulation. Although the rule-based (MSC) strategy accidentally achieves the lowest battery degradation cost in the residential PV battery system under this study, it was found out that the battery was much less actively used under the rule-based (MSC) strategy, which cannot be generalized as normal operation of other PV battery systems. Moreover, the MSC strategy is solely based on the residual energy without the capability of addressing various uncertainties in grid-connected PV battery systems.

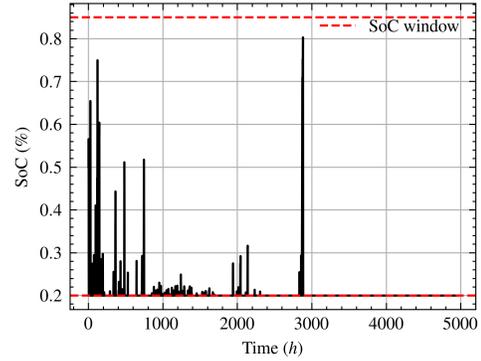


Fig. 8. The battery SoC as a function of time under the rule-based (MSC) strategy over the remaining half year.

V. CONCLUSION

As the first attempt to develop aging-aware control strategies with consideration of various uncertainties in grid-connected photovoltaic battery systems, an economically motivated stage cost function was proposed to account for both the grid and the battery degradation cost. The probability distribution of the disturbance (or normalized residual energy) was first estimated using the first two years of data with different assumptions, i.e., Gaussian distribution and empirical distribution. The implicit optimal policy achieved the best economic performance (i.e., the lowest grid and battery degradation cost) with smaller fluctuations than the explicit one in the remaining half-year simulation, as it considered both historical disturbance information w.r.t. hour of the day and "fixed" forecasted electricity price information.

Although the battery degradation model used in this work includes both calendar and cycling aging, it is still relatively simple and does not consider battery capacity knee occurrence, for example, which is expected to have a significant impact on the economic performance of control strategies. Moreover, only the "fixed" electricity price forecasts are

TABLE IV
ECONOMIC PERFORMANCE EVALUATION OVER THE REMAINING HALF YEAR

Control strategy	Disturbance assumption	Electricity price assumption	Grid cost (EUR)	Battery degradation (%)	Battery degradation cost (EUR)	Total cost (EUR)
Rule-based (MSC)*	-	-	32822.0	2.50	10950.0	44772.0
Rule-based (MSC)	-	-	31516.5	0.86	3748.8	35265.3
Explicit	Gaussian	Fixed	30548.5	1.25	5467.0	36015.4
Implicit	Empirical	Forecasted	29300.4	1.25	5479.4	34779.9

* The realistic maximizing self-consumption (MSC) strategy implemented in the residential PV battery system under study.

incorporated into the Markov Decision Processes (MDPs) formulation. The forecasts for other variables, such as load demand and renewable energy production are expected to improve the economic performance of learned optimal policies. Lastly, we will also enforce constraints in discounted MDPs so that the resulting policy is stabilizing.

ACKNOWLEDGMENT

This work was supported by the Swedish Energy Agency (Grant number P2024-00998). In particular, the authors would like to thank Sébastien Gros from Norwegian University of Science and Technology, Ylva Olofsson from Volvo Energy, Pierre Hult from Riksbyggen, Göteborg Energi, and Johanneberg Science Park for their additional support to this work.

REFERENCES

- [1] O. Alsauskas, E. Connelly, A. Daou, A. Gouy, M. Huismans, H. Kim, J.-B. Le Marois, S. McDonagh, A. Petropoulos, and J. Teter, "Global ev outlook 2023: Catching up with climate ambitions," 2023.
- [2] B. Faessler, "Stationary, second use battery energy storage systems and their applications: a research review," *Energies*, vol. 14, no. 8, p. 2335, 2021.
- [3] A. A. Kebede, T. Kalogiannis, J. Van Mierlo, and M. Bercibar, "A comprehensive review of stationary energy storage devices for large scale renewable energy sources grid integration," *Renewable and Sustainable Energy Reviews*, vol. 159, p. 112213, 2022.
- [4] A. Groß, A. Lenders, T. Zech, C. Wittwer, and M. Diehl, "Using probabilistic forecasts in stochastic optimization," in *2020 International Conference on Probabilistic Methods Applied to Power Systems (PMAPS)*. IEEE, 2020, pp. 1–6.
- [5] A. Groß, C. Wittwer, and M. Diehl, "Stochastic model predictive control of photovoltaic battery systems using a probabilistic forecast model," *European Journal of Control*, vol. 56, pp. 254–264, 2020.
- [6] B. Zou, J. Peng, S. Li, Y. Li, J. Yan, and H. Yang, "Comparative study of the dynamic programming-based and rule-based operation strategies for grid-connected pv-battery systems of office buildings," *Applied energy*, vol. 305, p. 117875, 2022.
- [7] A. Mahfuz-Ur-Rahman, M. R. Islam, K. M. Muttaqi, and D. Sutanto, "An effective energy management with advanced converter and control for a pv-battery storage based microgrid to improve energy resiliency," *IEEE Transactions on Industry Applications*, vol. 57, no. 6, pp. 6659–6668, 2021.
- [8] M. Zanon, T. Faulwasser, and S. Gros, "Practical economic mpc," *PAMM*, vol. 20, no. 1, p. e202000216, 2021.
- [9] S. Gros and M. Zanon, "Data-driven economic nmpc using reinforcement learning," *IEEE Transactions on Automatic Control*, vol. 65, no. 2, pp. 636–648, 2019.
- [10] M. Zanon, S. Gros, and A. Bemporad, "Practical reinforcement learning of stabilizing economic mpc," in *2019 18th European Control Conference (ECC)*. IEEE, 2019, pp. 2258–2263.
- [11] S. Gros and M. Zanon, "Reinforcement learning for mixed-integer problems based on mpc," *IFAC-PapersOnLine*, vol. 53, no. 2, pp. 5219–5224, 2020.
- [12] A. B. Kordabad, W. Cai, and S. Gros, "Mpc-based reinforcement learning for economic problems with application to battery storage," in *2021 European Control Conference (ECC)*. IEEE, 2021, pp. 2573–2578.
- [13] J. Olmos, I. Gandiaga, A. Saez-de Ibarra, X. Larrea, T. Nieva, and I. Aizpuru, "Modelling the cycling degradation of li-ion batteries: Chemistry influenced stress factors," *Journal of Energy Storage*, vol. 40, p. 102765, 2021.
- [14] H. C. Hesse, R. Martins, P. Musilek, M. Naumann, C. N. Truong, and A. Jossen, "Economic optimization of component sizing for residential battery storage systems," *Energies*, vol. 10, no. 7, 2017. [Online]. Available: <https://www.mdpi.com/1996-1073/10/7/835>
- [15] A. A. Kebede, T. Coosemans, M. Messagie, T. Jemal, H. A. Be-habtu, J. Van Mierlo, and M. Bercibar, "Techno-economic analysis of lithium-ion and lead-acid batteries in stationary energy storage application," *Journal of Energy Storage*, vol. 40, p. 102748, 2021.
- [16] D. Bertsekas, *Reinforcement learning and optimal control*. Athena Scientific, 2019, vol. 1.
- [17] S. Gros and M. Zanon, "Economic mpc of markov decision processes: Dissipativity in undiscounted infinite-horizon optimal control," *Automatica*, vol. 146, p. 110602, 2022.
- [18] M. Zanon and S. Gros, "A new dissipativity condition for asymptotic stability of discounted economic mpc," *Automatica*, vol. 141, p. 110287, 2022.
- [19] M. Zanon, S. Gros, and M. Palladino, "Stability-constrained markov decision processes using mpc," *Automatica*, vol. 143, p. 110399, 2022.