采用频域约束强化学习的微电网异构混合储能调度方法

崔菲菲[1]，安豆[1]，杨清宇[1,2]，任志刚[1]
（1. 西安交通大学自动化科学与工程学院，710049，西安；
2. 西安交通大学精密微纳制造技术全国重点实验室，710049，西安）

摘要：在微电网源、荷两侧波动加剧的背景下，针对传统强化学习算法隐式引导异构储能调度易导致策略同质化、协同分工失效的问题，提出一种基于频域约束的深度强化学习协同优化调度方法。首先，面向微电网压缩空气储能与电池储能异构混合储能系统，建立马尔可夫决策过程模型以刻画调度系统的状态演变特征与序贯决策特性；其次，基于双延迟深度确定性策略梯度强化学习算法，将谱正则化频域约束惩罚项引入算法策略网络梯度中，以引导电池储能吸收高波动功率，压缩空气储能跟踪低频基波，实现二者差异化分工。仿真结果表明，与主流深度强化学习基准算法相比，所提算法在典型周调度内，可将综合运行成本降低 10.0%，并网功率跟踪误差降低约 68.5%，并网功率越线次数减少约 92.1%，压缩空气储能机械寿命指标与电池储能等效全循环次数分别下降 92% 和 53.8%。该算法可兼顾微电网系统经济性、电网安全稳定性以及底层储能设备的寿命友好性。

关键词：微电网；异构混合储能；协同优化调度；深度强化学习；频域约束
中图分类号：TP273 **文献标志码**：A
DOI：10.7652/xjtuxb202611006 **文章编号**：0253-987X(2026)11-0059-12

Reinforcement Learning-Based Dispatching Method with Frequency-Domain Constraints for Heterogeneous Hybrid Energy Storage Systems in Microgrids

CUI Feifei[1], AN Dou[1], YANG Qingyu[1,2], REN Zhigang[1]
(1. School of Automation Science and Engineering, Xi'an Jiaotong University, Xi'an 710049, China;
2. State Key Laboratory for Manufacturing System Engineering, Xi'an Jiaotong University, Xi'an 710049, China)

Abstract: In the context of intensified fluctuations on both the generation and load sides of microgrids, conventional deep reinforcement learning algorithms often lead to strategy homogenization and failure of synergistic division of labor due to their implicit guidance of heterogeneous energy storage dispatch. To address these drawbacks, a coordinated optimal dispatch method based on frequency-domain-constrained deep reinforcement learning is proposed. First, tailored to the heterogeneous hybrid energy storage system comprising compressed air energy storage and battery energy storage systems in microgrids, a Markov decision process model is established to characterize the state evolution features and sequential decision-making properties of the dispatch system. Next, based on the twin delayed deep deterministic policy gradient (TD3) reinforcement learning algorithm, a spectral-regularized frequency-domain constraint penalty term is incorporated into

收稿日期：2026-03-21。 作者简介：崔菲菲(1999—)，女，博士生；安豆(通信作者)，男，教授，博士生导师。 基金项目：国家重点研发计划资助项目(2024YFB4206500)；国家自然科学基金资助项目(62573344，62373297)。
网络出版时间：2026-08-12 网络出版地址：https://link.cnki.net/urlid/61.1069.T.20260811.1605.004

the gradient of the algorithm's policy network to guide the battery energy storage system to absorb high-fluctuation power and the compressed air energy storage system to track the low-frequency fundamental wave, thereby achieving a differentiated division of labor between the two. Simulation results demonstrate that, compared with mainstream deep reinforcement learning benchmark algorithms, the proposed algorithm can reduce the comprehensive operating cost by 10.0%, decrease the grid-connected power tracking error by approximately 68.5%, and lower the number of grid power limit violations by 92.1% in a typical weekly dispatch scenario. Moreover, the mechanical life index of the compressed air energy storage and the equivalent full cycle number of the battery energy storage are decreased by 92% and 53.8%, respectively. This algorithm can balance the economic efficiency of the microgrid system, the safety and stability of the power grid, and the lifespan friendliness of the underlying energy storage devices.

Keywords: microgrid; heterogeneous hybrid energy storage; coordinated optimal dispatch; deep reinforcement learning; frequency-domain constraint

高比例风电、光伏等可再生能源并网已成为实现能源转型与“双碳”目标的重要途径之一。因此，新型电力系统运行呈现出源侧波动加剧、负荷侧不确定性上升以及柔性资源耦合增强的趋势[1]。为应对上述挑战，源-网-荷-储一体化运行模式通过对分布式电源、电网交互、可调负荷与储能等资源的整合，实现多时间尺度的协同优化调控，从而提升系统运行的可控性与资源利用效率[2-3]。在此架构中，储能系统(ESS)具备双向可控与能量缓冲特性，可用于平抑不平衡功率波动、改善并网功率交换特性、提供时域调节等，是支撑高比例可再生能源友好接入与系统安全稳定运行的关键环节之一[4-5]。

受限于单一储能技术在功率密度、能量密度与寿命成本之间难以同时兼顾，具备互补特性的异构储能协同被认为是提升多时间尺度调节能力的重要方向之一[6-7]。其中，压缩空气储能(CAES)具有高能量密度与长时段能量平移优势，电池储能系统(BESS)具备高功率与承担短时能量平衡优势，两者在频域层面存在协同调度潜力[8]。现有研究中，文献[9]基于能量枢纽架构对 CAES 与 BESS 进行联合建模，并在计及系统不确定性与需求响应的条件下开展多目标鲁棒优化，验证了异构储能在经济与环境双重效益上的协同增益。面向综合能源系统，文献[10]提出了多模式 CAES-BESS 混合架构，旨在刻画不同运行模式下的协同调节规律，提升调度决策的自适应能力。此外，文献[11]利用自适应变分模态分解技术处理光伏波动功率，通过在电池与超级电容间进行频域功率分配，有效缓解了高频分量对设备寿命与运行成本的消极影响。尽管如此，当前异构储能系统的协同调度仍面临跨时间尺度耦合强、决策可行域随状态演化而动态收缩等挑战。

针对微电网协同调度问题，传统研究主要依赖两类优化范式：一是以鲁棒或随机优化为代表的数学规划方法，用于在显式安全约束与不确定性条件下获得可行调度解[12-13]；二是遗传算法等启发式搜索方法，主要用于求解非凸与多目标问题[14]。然而，这些传统方法对模型精度及计算规模较为敏感，面对源-网-荷-储高度耦合且充满不确定性的环境，往往难以兼顾在线调度的时效性与自适应性。为此，深度强化学习(DRL)凭借其在序贯决策问题上的优势，开始被引入微电网的能量管理中。具体地，文献[15-16]通过建立双层深度确定性策略梯度(DDPG)架构响应多时间尺度管理需求，构建了不确定环境下的在线策略求解流程。在异构储能应用方面，文献[17]针对氢储-电池微电网，构建了基于双延迟深度确定性策略梯度(TD3)的能量管理框架，实现了源荷扰动下的多目标动态权衡。文献[18]设计了双层 DRL 策略，实现了超级电容-电池-氢混合储能系统中功率与变流器的协同控制。文献[19]面向风电并网，利用多智能体 DRL 算法完成了混合储能的二次功率分配与波动平抑。文献[20]则针对有源配电网源-荷-储系统，设计了基于 DDPG 的滚动调控方法，建立了不同时间尺度下的决策执行机制。

尽管 DRL 具备处理高维连续动作空间、弱模型依赖以及适应系统不确定性等优势，已逐渐应用于异构储能协同调度问题，现有方法仍面临以下挑战：①常规 DRL 过度依赖复杂的奖励函数来隐式引导异构储能的分工，导致奖励项构造与权重整定繁琐，策略对工况变化的适应性差；②系统具有强耦合特征与复杂物理边界，导致 DRL 决策的可行域大幅收

缩，算法难以稳定兼顾物理约束的满足与策略的最优；③已有研究将频域指标引入奖励中以引导学习，由于其依赖时间窗口统计，会引入严重的马尔可夫历史依赖与回报时延，导致单步动作的信用分配困难，降低算法学习效率。

针对上述问题，面向微电网异构混合储能协同运行场景，提出一种基于频域约束的协同调度方法，主要创新工作如下。

(1)建立了微电网异构混合储能的能量-功率互补协同运行机制。该机制通过引导功率型 BESS 吸收高变化率的瞬态波动、引导能量型 CAES 跟踪平缓的基本负荷波，实现异构设备的互补协同。

(2)构建了计及源、荷不确定性与储能设备时序演化特征的马尔可夫决策过程(MDP)模型。该模型将储能出力抽象为连续动作空间，并将多维目标与物理安全约束设计为复合奖励机制，从而把传统单步静态调度问题转化为追求长期期望回报最大化的序贯决策问题。

(3)提出了一种基于频域约束的双延迟深度确定性策略梯度协同优化算法(SpecTD3)，创新性地在策略网络的梯度更新中引入谱正则化频域惩罚项，从底层算法机制上显式限制了 CAES 调度指令中的瞬态跳变分量，引导 BESS 吸收剧烈波动、CAES 跟踪平缓基波，实现异构储能设备的时频解耦出力。

1　异构混合储能协同调度模型

1.1　系统描述

如图 1 所示，含异构混合储能的并网微电网系统由风电、光伏、负荷、CAES、BESS、主电网及能量管理系统(EMS)构成。为提升系统调节能力，配置由 CAES 与 BESS 组成的异构混合储能系统。其中，CAES 内部的热力学结构基于文献[21]设计，配置三级压缩与高低压膨胀机组。太阳能集热回路选取首诺 66(Therminol 66)导热油作为热介质，对膨胀机入口空气进行梯级升温，以提高发电动率。与常规储能不同，CAES 在执行电功率充放电的过程中伴随热能的产生与消耗，副产物热能可通过热网进行购售或供给系统内热用户，从而提升系统能源利用效率。在运行阶段，EMS 负责采集各设备及源、荷信息，对 CAES 与 BESS 的充放电出力进行协调控制，在满足系统及设备安全经济运行约束的同时提升微电网运行的经济性与稳定性。

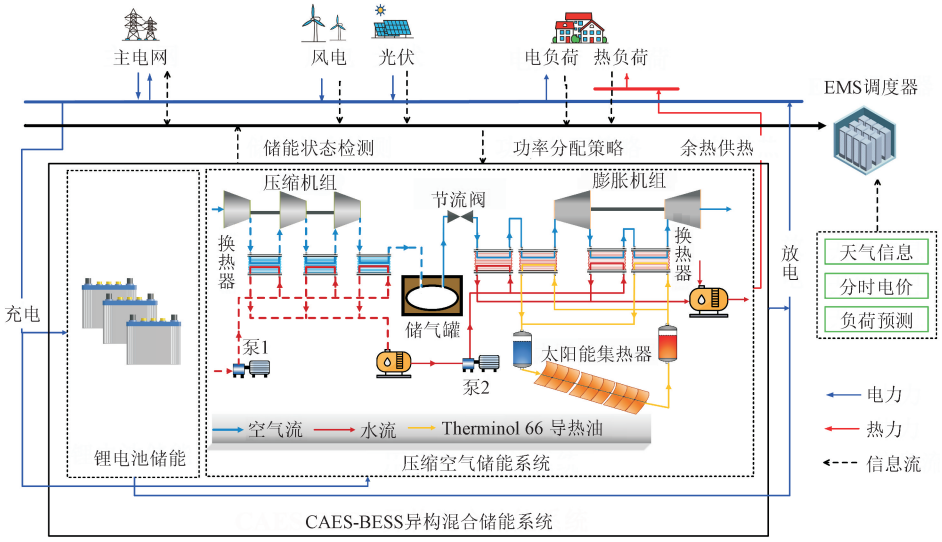


图 1　微电网异构混合储能协同调度系统

Fig. 1　Microgrid system for coordinated dispatch of heterogeneous hybrid energy storage

本文配置的 CAES 与 BESS 在物理特性上具有显著互补性：CAES 作为能量型储能，能量密度高、适合长时段能量平移，但对高频启停与剧烈爬坡敏感；BESS 作为功率型储能，可承担瞬态功率波动平抑，但深度循环会加速电化学衰减。二者通过 CAES 副产热与热网交互形成电-热间接耦合，共同参与微电网的并网功率平滑、峰谷电价套利与可再生能源消纳等辅助服务。

1.2　问题描述

1.2.1　系统模型及约束

在任意调度时段 t 内，微电网内的风光出力、主网交互功率以及异构储能的充放电功率需满足实时

平衡

$$P_{W,t}+P_{P,t}+P_{G,t}-P_{B,t}-P_{C,t}=P_{L,t} \quad (1)$$

式中：$P_{W,t}$、$P_{P,t}$ 分别为风电与光伏的出力功率；$P_{G,t}$ 为与主电网的交互功率(大于 0 表示购电)；$P_{B,t}$ 为 BESS 功率(大于 0 表示充电)；$P_{C,t}$ 为 CAES 充放功率；$P_{L,t}$ 为系统电负荷。

CAES 在运行中产生的副产物热能与热网的交互功率，共同满足系统的热负荷需求。热功率平衡约束表达为

$$H_{C,t}+H_{G,t}=H_{L,t} \quad (2)$$

式中：$H_{C,t}$ 为 CAES 系统产生的可用副产热功率；$H_{G,t}$ 为向城市热网购入与售出的热功率；$H_{L,t}$ 为系统热负荷功率。

为避免对大电网造成冲击，微电网与主电网的交互功率需满足容量上下限约束和爬坡率约束

$$P_{G,\min}\leqslant P_{G,t}\leqslant P_{G,\max} \quad (3)$$

$$|P_{G,t-1}-P_{G,t}|\leqslant P_{G,r} \quad (4)$$

式中：$P_{G,\min}$、$P_{G,\max}$ 分别为微电网向主网购电与售电的最小、最大功率；$P_{G,r}$ 为电网最大允许爬坡量。

1.2.2 CAES-BESS 异构储能系统模型

BESS 在异构储能系统中主要负责平抑短期剧烈功率波动，其荷电状态(SoC)连续时间状态演化过程[22]如下

$$S_{B,t}=\begin{cases}S_{B,t-1}+\dfrac{\eta_{ch}P_{B,t}}{E_B}\Delta t, & P_{B,t}\geqslant 0\\ S_{B,t-1}+\dfrac{P_{B,t}}{\eta_{dis}E_B}\Delta t, & P_{B,t}<0\end{cases} \quad (5)$$

式中：$S_{B,t}$ 为 BESS 在 t 时刻的荷电状态；η_{ch} 为充电效率；η_{dis} 为放电效率；E_B 为 BESS 额定容量。

BESS 安全运行约束如下

$$S_{B,\min}\leqslant S_{B,t}\leqslant S_{B,\max} \quad (6)$$

$$-P_{B,\max}\leqslant P_{B,t}\leqslant P_{B,\max} \quad (7)$$

式中：$S_{B,\min}$、$S_{B,\max}$ 分别为 BESS 最小、最大荷电状态；$P_{B,\max}$ 为 BESS 的最大充放电功率。

为量化频繁充放电对 BESS 寿命的不可逆影响，引入基于循环深度的非线性寿命衰减成本。定义累积循环深度 $\delta_t=\delta_{t-1}+\dfrac{|P_{B,t}|}{E_B}\Delta t$。基于材料疲劳特性与雨流计数等效原则，得到 t 时段的寿命损耗成本 $D_{B,t}$ 计算如下

$$D_{B,t}=a_0bc_b\frac{\delta_t^{b-1}}{E_B}P_{B,t} \quad (8)$$

式中：$a_0=5.24\times10^{-4}$ 和 $b=2.03$ 为循环退化拟合系数[22]；c_b 为电池组单位容量置换成本。

CAES 储能容量大且对高频动作敏感。考虑到在非额定工况下，压缩机与膨胀机的热力学参数具有强非线性特征，基于之前研究结论[10]采用非设计工况下二次多项式拟合结果描述 CAES 的空气质量流量与功率关系，表达式如下

$$\dot{m}_{c,t}=6\times10^{-6}P_{c,t}^2+0.0016P_{c,t}+0.0058 \quad (9)$$

$$\dot{m}_{d,t}=2\times10^{-5}P_{d,t}^2+0.0013P_{d,t}+0.1528 \quad (10)$$

式中：$\dot{m}_{c,t}$ 和 $\dot{m}_{d,t}$ 分别为 t 时刻 CAES 压缩和释放的空气质量流量；$P_{c,t}$ 和 $P_{d,t}$ 分别为压缩机充电和膨胀机放电的功率。

CAES 的 SoC 转移方程表示为

$$S_{C,t}=S_{C,t-1}+\frac{\Delta p_{C,t}}{p_{\max}-p_{\min}} \quad (11)$$

式中：$S_{C,t}$ 为 CAES 在 t 时刻的荷电状态；$p_{\max}$ 和 $p_{\min}$ 分别为储气室最大和最小允许运行压强；$\Delta p_{C,t}$ 为 t 时段内储气室压强变化量，由充放气质量流量决定

$$\Delta p_{C,t}=\frac{(\dot{m}_{c,t}u_{c,t}-\dot{m}_{d,t}u_{d,t})RT}{V}\Delta t \quad (12)$$

式中：$u_{c,t}$、$u_{d,t}\in\{0,1\}$ 为机组工作状态标识；R 为空气气体常数；T 与 V 分别为储气室运行温度、容积。

机组运行状态及功率边界需满足以下硬约束

$$\begin{cases}S_{C,\min}\leqslant S_{C,t}\leqslant S_{C,\max}\\ 0\leqslant P_{c,t}\leqslant u_{c,t}P_{c,\max}\\ 0\leqslant P_{d,t}\leqslant u_{d,t}P_{d,\max}\\ |P_{c,t}-P_{c,t-1}|\leqslant R_c\\ u_{c,t}+u_{d,t}\leqslant 1\end{cases} \quad (13)$$

式中：$S_{C,\min}$、$S_{C,\max}$ 分别为 CAES 的荷电状态最小、最大值；$P_{c,\max}$、$P_{d,\max}$ 分别为充电状态下压缩机最大功率和放电状态下膨胀机最大功率；R_c 为 CAES 机组的最大爬坡功率限制。

CAES 在膨胀释放阶段产生的副产物热功率为 $H_{C,t}$。热回收水循环的质量流量在此工况下设定为 $\dot{m}_h=3.81$ kg/s，其线性换算模型表示为

$$H_{C,t}=c_p\dot{m}_h(T_h-T_o)u_{d,t} \quad (14)$$

式中：c_p 为水的比定压热容；T_h 与 T_o 分别为换热器的高温端与冷端温度。

CAES 运行损耗主要由机组持续压缩/膨胀带来的机械磨损引起。采用基于经验系数的线性模型对 CAES 运行成本 $D_{C,t}$ 进行量化计算

$$D_{C,t}=c_q|P_{c,t}-P_{d,t}| \quad (15)$$

式中：c_q 为单位功率运行带来的成本系数。

1.2.3 优化目标

调度任务核心是在保障系统物理安全下，实现

经济性最优和并网误差最小化。综合经济成本项设计为

$$C=\sum_{t=1}^{T}(C_{\mathrm{e},t}+C_{\mathrm{h},t}+D_{\mathrm{B},t}+D_{\mathrm{C},t}) \tag{16}$$

式中：$C_{\mathrm{e},t}$ 为主网电能交互成本，$C_{\mathrm{e},t}=\pi_{\mathrm{e},t}P_{\mathrm{G},t}\Delta t$，$\pi_{\mathrm{e},t}$ 为分时电价；$C_{\mathrm{h},t}$ 为热网交互成本，$C_{\mathrm{h},t}=\pi_{\mathrm{h}}H_{\mathrm{G},t}\Delta t$，$\pi_{\mathrm{h}}$ 为购热价格。

为平抑系统内部风光资源及负荷的强随机波动，减小对主电网的冲击，设定并网误差项

$$E_{\mathrm{G}}=\sqrt{\frac{1}{T}\sum_{t=1}^{T}(P_{\mathrm{G},t}-P_{\mathrm{r}})^2} \tag{17}$$

式中：P_{r} 为上级电网发布的基准功率。

2　马尔可夫决策过程建模

微电网异构储能系统的连续协同调度任务本质上是一个典型的序贯决策问题。系统当前 t 时刻的调度动作不仅决定当下的运行成本和功率流动，还会改变储能系统的 SoC，影响未来时刻的调节潜力与决策空间。由于该系统的状态转移具有无后效性，即下一时刻的状态仅由当前状态和当前采取的动作决定，与历史状态无关，因此该调度过程可以被建模为 MDP 模型[23]。标准 MDP 由五元组〈S，A，P，R，γ〉构成：S 为系统状态空间，A 为智能体 Agent 的动作空间，P 表示状态转移概率分布，R 为智能体与环境交互后获得的奖励函数，$\gamma\in[0,1]$ 为折扣因子。在此调度场景中，EMS 即为强化学习智能体，通过与微电网环境的持续交互，学习使长期累积奖励最大化的最优策略。

状态空间 S 需充分涵盖反映微电网运行状况及外部环境的关键特征。设计了 10 维连续状态变量向量 $\boldsymbol{s}_t$，包含可再生资源出力、储能设备状态、需求侧功率、分时电价、上一时刻关键设备运行状态。为加速神经网络收敛，将功率量 P 和 H 分别归一化为 $\widetilde{P}$ 和 $\widetilde{H}$。状态空间统一表达式为

$$\begin{aligned}\boldsymbol{s}_t=[&\widetilde{P}_{\mathrm{W},t},\ \widetilde{P}_{\mathrm{P},t},\ C_{\mathrm{C},t},\ C_{\mathrm{B},t},\ \widetilde{P}_{\mathrm{L},t},\ \widetilde{H}_{\mathrm{L},t},\ \pi_{\mathrm{e},t},\\&\widetilde{P}_{\mathrm{G},t-1},\ \widetilde{P}_{\mathrm{C},t-1},\ \widetilde{P}_{\mathrm{B},t-1}]^{\mathrm{T}}\end{aligned} \tag{18}$$

动作空间 A 由智能体对 CAES 和 BESS 的控制指令构成。采用连续动作空间表示，归一化动作向量 $\boldsymbol{a}_t=[\widetilde{P}_{\mathrm{C},t},\ \widetilde{P}_{\mathrm{B},t}]^{\mathrm{T}}$。考虑到设备的物理容量与当前时刻 SoC 的限制，输出动作受设备物理状态的动态硬截断约束。在每个调度时刻 t，根据 BESS 与 CAES 当前的 SoC、剩余容量及功率爬坡上下限，动态计算各设备允许动作区间；策略网络输出原始动作经 tanh 函数激活后，通过允许动作区间截断得到执行动作。

奖励函数 R 用于评价调度策略的优劣。微电网异构储能协同调度是一个典型的高维、多目标且包含多重软硬性物理约束的优化问题。在 DRL 框架下，传统硬约束难以被直接嵌入求解。因此，将系统协同调度的多维优化目标与物理设备的安全运行边界统一映射至复合奖励函数中。引导智能体在追求系统经济性与功率平滑性的同时，学习如何严格遵循电网及异构储能的物理运行约束规则。复合奖励函数 r_t 表达式为

$$r_t=R_{\mathrm{O},t}+R_{\mathrm{D},t}+R_{\mathrm{G},t}+R_{\mathrm{C},t}+R_{\mathrm{S},t} \tag{19}$$

式中：$R_{\mathrm{O},t}$ 为综合运行成本惩罚；$R_{\mathrm{D},t}$ 为并网偏差误差惩罚；$R_{\mathrm{G},t}$ 为并网爬坡惩罚；式(16)、(17) 的优化目标通过 $R_{\mathrm{O},t}$、$R_{\mathrm{D},t}$ 惩罚函数实现，式(4)、(13) 爬坡约束分别由 $R_{\mathrm{G},t}$、$R_{\mathrm{C},t}$ 惩罚函数实现。为避免 BESS 的深度循环状态并保持调度裕度，设计 SoC 引导惩罚 $R_{\mathrm{S},t}$，表达式如下

$$\begin{cases}R_{\mathrm{O},t}=-\omega_{\mathrm{O}}(C_{\mathrm{e},t}+C_{\mathrm{h},t}+D_{\mathrm{B},t}+D_{\mathrm{C},t})\\R_{\mathrm{D},t}=-\omega_{\mathrm{D}}\sqrt{(P_{\mathrm{G},t}-P_{\mathrm{r}})^2}\\R_{\mathrm{G},t}=-\omega_{\mathrm{G}}(\Delta P_{\mathrm{G},t})^2\\R_{\mathrm{C},t}=-\omega_{\mathrm{C}}(\Delta P_{\mathrm{C},t})^2\\R_{\mathrm{S},t}=-\omega_{\mathrm{S}}(C_{\mathrm{B},t}-0.5)^2\end{cases} \tag{20}$$

式中：ω_{O}、ω_{D}、ω_{G}、ω_{C} 和 ω_{S} 分别为 5 项惩罚函数的权重。

3　基于 SpecTD3 异构储能调度算法

3.1　算法结构设计

在异构 CAES 和 BESS 混合储能协同连续控制问题中，传统 DRL 依赖复杂的复合奖励函数来隐式引导异构储能的协同分工。该方式不仅导致奖励项构造与权重整定繁琐，在高度耦合与多重物理边界约束下，调度决策的安全可行域被压缩。若仅以运行成本和储能容量作为决策依据，智能体难以挖掘二者的底层动态差异，将其同质化为可高频反复充放的通用能量缓冲池。但 CAES 内部大型旋转机械的安全运行亟需低频次的启停与平缓的功率爬坡，而 BESS 则需避免深度循环状态以延缓电化学衰减。基于上述动机，提出一种考虑频域约束的 SpecTD3 算法，基于主流 TD3 算法，创新性地引入频域约束机制，通过在策略网络的梯度更新中加入针对动作剧烈变化的惩罚引导，从算法底层显式抑制 CAES 调度指令中的瞬态高频跳变分量，以实现 BESS 抑制剧烈波动，CAES 跟踪基波的协同出力。

本文所提 SpecTD3 算法的整体架构如图 2 所示，主要由微电网环境交互模块、经验回放池、双评论家网络(Critics)以及内嵌频域约束模块的策略网络构成。在交互阶段，Agent 首先于每个调度步观测微电网多维状态并输出确定的充放电动作，交互生成的数据元组被存入经验回放池，供后续训练进行小批量随机采样。在网络更新阶段，双 Critic 网络基于采样数据计算时序差分误差并执行梯度下降。如图 2 下方区域所示，策略网络在其梯度更新路径中创新性嵌入了频域约束模块，该模块通过提取动作序列的频域特征并构建损失惩罚，在执行延迟梯度更新时从底层显式抑制了 CAES 动作输出中的高频跳变分量。所提算法无需过度设计复杂奖励函数，基于底层梯度的频域修正，可引导系统训练出电池吸收高波动功率、CAES 跟踪平缓基波的最优协同策略。

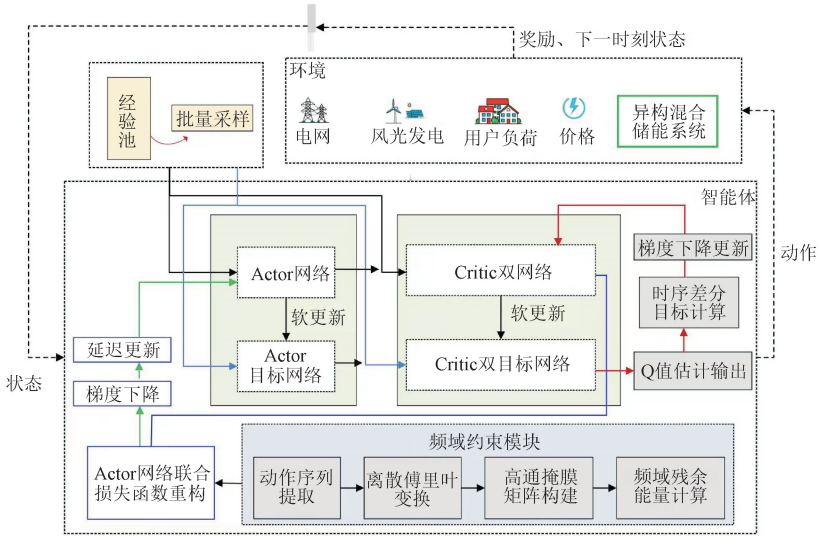


图 2 SpecTD3 异构储能调度算法

Fig. 2 SpecTD3 heterogeneous energy storage dispatch algorithm

综上，本文所提的频域约束是一种通过谱正则化惩罚项嵌入策略网络梯度更新过程的软性约束机制，其约束对象为 CAES 通道的动作输出，可在硬约束所允许的可行域内，进一步引导智能体收敛至协同分工调度策略。

3.2 算法原理

SpecTD3 算法的 Critics 网络继承了 TD3 算法的双网络结构与目标策略平滑机制[24]。两个 Critic 网络结构相同但参数独立，网络记为 Q_{θ_1} 与 Q_{θ_2}，对应参数为 θ_1 与 θ_2。其对应的两个目标 Critic 网络记为 $Q_{\theta'_1}$ 与 $Q_{\theta'_2}$，参数设为 θ'_1 与 θ'_2。为缓解传统 DRL 算法过估计问题，算法在计算目标 Q 值时会对目标动作施加截断的正态分布噪声，以提升价值评估鲁棒性。目标动作 $\tilde{\boldsymbol{a}}$ 表达为

$$\tilde{\boldsymbol{a}} = \mu_{\varphi'}(\boldsymbol{s}') + \boldsymbol{v},$$
$$\boldsymbol{v} \sim \mathrm{clip}(N(\boldsymbol{0}, \sigma^2\boldsymbol{I}), -d, d) \tag{21}$$

式中：$\mu_{\varphi'}$ 代表目标策略网络；$\boldsymbol{s}'$ 为下一时刻的环境状态；$\boldsymbol{v}$ 为噪声向量，服从 $N(\boldsymbol{0}, \sigma^2\boldsymbol{I})$ 正态分布，并由 clip(·) 函数将噪声各维度逐元素截断至区间 $[-d, d]$，d 为噪声绝对值的截断边界。

在获取平滑后的目标动作后，算法选取两个目标 Q 网络输出值中的较小者作为未来价值期望，并结合当前环境反馈的真实奖励 r_i 计算时序差分目标值为

$$y_i = r_i + \gamma(1-d_i)\min_{j=1,2} Q_{\theta'_j}(\boldsymbol{s}'_i, \tilde{\boldsymbol{s}}) \tag{22}$$

式中：d_i 为标识当前回合是否结束的布尔变量。

Critic 通过最小化当前预测值与目标值 y_i 之间的均方误差，沿梯度下降方向以学习率 α_c 更新网络参数 θ_j，损失函数表达为

$$L_{\mathrm{C}}(\theta_j) = \frac{1}{N}\sum_{i=1}^{N}(y_i - Q_{\theta_j}(\boldsymbol{s}_i, \boldsymbol{a}_i))^2 \tag{23}$$

标准策略网络通过沿 Critic 网络输出的 Q 值沿梯度上升方向以学习率 α_a 更新参数 μ_φ。但为避免 CAES 承担剧烈波动，对其输出动作的高频分量施加限制。根据离散傅里叶变换能量守恒原理，控制信号在时域上的剧烈波动等效于频域中高频能量的积累。因此，利用相邻动作节点的一阶时域差分来构建动作序列的频域高频能量模型。

假设从小批量经验数据中提取的时间连续动作序列张量为 $\boldsymbol{A} = [\mu_\varphi(\boldsymbol{s}_1), \mu_\varphi(\boldsymbol{s}_2), \cdots, \mu_\varphi(\boldsymbol{s}_N)]^{\mathrm{T}}$。对

该序列执行傅里叶变换，可将其从时域映射至频域 $F(\boldsymbol{A})$。定义截止频率占比 η（$0<\eta<1$）的高通掩膜矩阵 $\boldsymbol{M}$。该掩膜在归一化频率 $f>\eta$ 的频段处值为 1，在低频段处值为 0。基于此，谱正则化频域约束惩罚项 $L_{\mathrm{F}}(\varphi)$ 被定义为动作序列中被滤波器拦截的残余高频能量

$$L_{\mathrm{F}}(\varphi)=\|\boldsymbol{M}\odot F(\boldsymbol{A})\|_2^2 \tag{24}$$

式中：$\odot$ 表示哈达玛积。将该惩罚项引入目标函数后，SpecTD3 的策略网络联合损失函数重构为

$$L_A(\varphi)=-\frac{1}{N}\sum_{i=1}^{N}Q_{\theta_1}(\boldsymbol{s}_i,\mu_\varphi(\boldsymbol{s}_i))+\lambda L_{\mathrm{F}}(\varphi) \tag{25}$$

式中：λ 为频域惩罚系数。

通过上述重构，当策略网络倾向于输出高频动作时，谱正则化模块将受到截止频率 η 的约束，并根据惩罚系数 λ 生成抑制此行为的反向梯度。该梯度在神经网络权重更新层级上将参数拉回平滑决策区域，从而过滤掉短视策略。

4　仿真实验与结果分析

4.1　仿真实验设置

为验证所提出基于 SpecTD3 算法的异构混合储能协同优化调度模型的有效性，采用某工业园区场景算例进行仿真分析，园区具有电、热能源需求特性。仿真以夏季 168 h 典型周为一个完整调度周期，仿真步长设置为 1 h。仿真所采用的太阳能辐照度与风速关键参数，源自美国国家可再生能源实验室官方数据库，如图 3 所示，风光发电模型源于文献[25]。园区电负荷与热负荷的小时级需求剖面参考文献[26]中的典型工业负荷模型，如图 4 所示。系统内各设备的物理约束及配置参数见表 1。购电价格采用峰谷分时电价，见表 2。

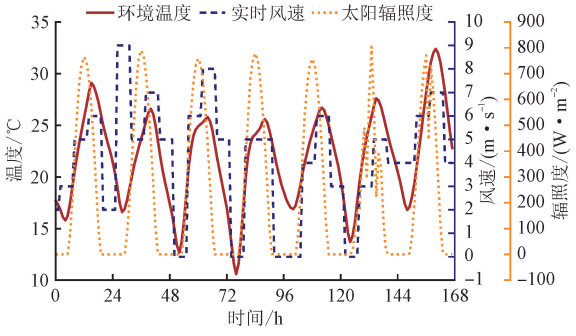


图 3　典型周气象参数变化曲线

Fig. 3　Variation curves of meteorological parameters in a typical week

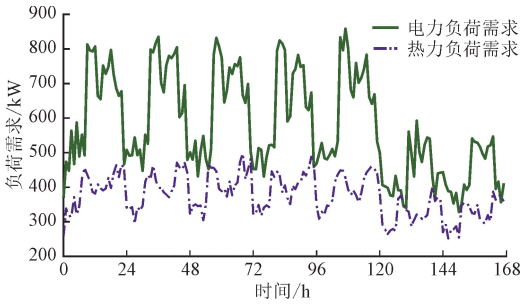


图 4　工业园区负荷需求曲线

Fig. 4　Load demand curves of the industrial park

SpecTD3 算法超参数设置见表 3。为验证所提算法性能，选取 TD3、软演员-评论家算法（SAC）、近端策略优化算法（PPO）3 种主流 DRL 算法作为对比算法。对比算法网络参数和环境模型与 SpecTD3 保持一致。为评估式(19)奖励函数中的 5 项引导效果，设计指标综合运行成本 C、并网功率跟踪误差 E_{G}、电网爬坡越限次数 N_{r}、CAES 机械寿命指标 $M_{\mathrm{C}}=\sum_{t=1}^{T}[(P_{\mathrm{c},t}-P_{\mathrm{c},t-1})/R_{\mathrm{c}}]^2$、BESS 等效全循环次数 $F_{\mathrm{C}}=\sum_{t=1}^{T}|P_{\mathrm{B},t}|/(2E_{\mathrm{B}})$。

表 1　系统核心设备物理与运行参数配置

Table 1　Configuration of physical and operational parameters for core system components

设备分类	参数名称	数值及范围
风电和光伏	额定功率/kW	1 200
	切入风速/(m·s^{-1})	2
	切出风速/(m·s^{-1})	25
	额定风速/(m·s^{-1})	14
	面板面积/m^2	8 000
电网	计划平均功率/kW	200
	最小并网功率/kW	400
	最大并网功率/kW	800
	爬坡率限制/(kW·h^{-1})	80
BESS	额定容量/(kW·h)	2 000
	充电效率/%	92
	放电效率/%	97
	SoC 运行范围/%	5～95
CAES	储气库体积/m^3	2 000
	压力范围/MPa	4.0～6.0
	SoC 运行范围/%	10～95

表 2 分时电价

Table 2 Time-of-use electricity prices

时间段	电价等级	价格/（美元·kW^{-1}·h^{-1}）
00:00—07:00	谷时	0.033 2
08:00—10:00，17:00—21:00	峰时	0.136 4
15:00—16:00，22:00—23:00	平时	0.079 3
11:00—14:00	标准	0.052 0

表 3 SpecTD3 算法超参数

Table 3 Hyperparameter settings for SpecTD3 algorithm

超参数	数值
折扣因子	0.99
软更新系数	0.005
训练批次大小	256
策略噪声标准差	0.2
噪声截断范围	0.5
延迟更新频率	2
学习率	3×10^{-4}
频谱惩罚权重	0.5
截止频率比例	0.2

4.2 算法性能对比分析

图 5 展示了不同 DRL 算法在微电网异构储能协同调度任务中的累计奖励收敛曲线。结果表明，基于 On-policy 的 PPO 算法在训练末期收敛于 −20 440 左右，呈现较大波动，而 SpecTD3、TD3 与 SAC 这 3 种 Off-policy 算法则表现出显著优势。从算法机理层面看，PPO 算法的 On-policy 样本机制丢弃了长时序中稀有的边界试错样本、信赖域裁剪压缩了高维动作空间的策略迁移效率、随机策略采样又加剧了物理硬约束下的执行偏差，三者叠加使其难以处理本文长时序、强约束、高维连续控制的微电网调度场景。相反，具备历史经验反复学习能力的 SAC、TD3 与所提的 SpecTD3 算法均在 100～200 回合区间内实现了快速稳定收敛。所提 SpecTD3 算法最终稳定收敛于最高奖励值 −5 215，相较于基准算法 TD3 和 SAC，分别提升了 17.4% 和 22.1%。综合上述，所提 SpecTD3 算法通过在策略网络梯度更新中引入频域约束模块，算法不仅未破坏基础网络结构的收敛平稳性，且可显式过滤极易触发系统硬约束惩罚的瞬态高频跳变动作，以规避智能体在复杂奖励引导下的无效探索。

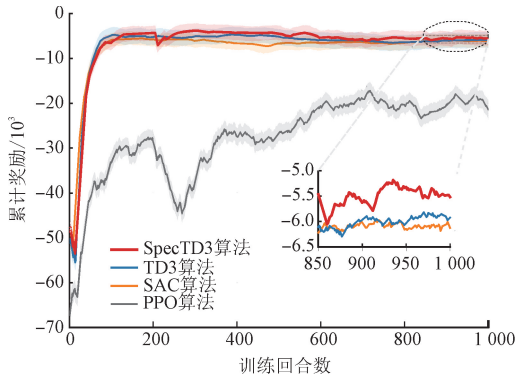
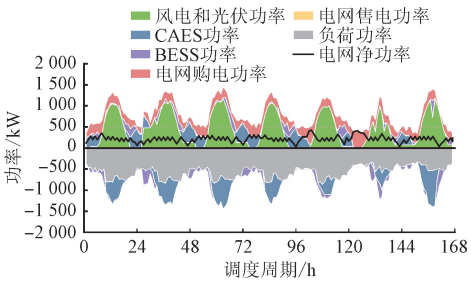


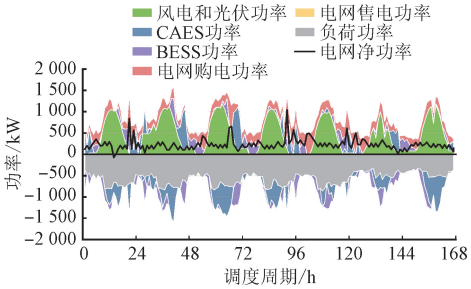
图 5 不同强化学习算法累计奖励训练曲线对比

Fig. 5 Comparison of cumulative reward training curves under different RL algorithms

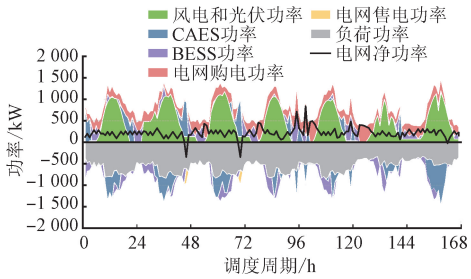
图 6 展示了在典型周内，不同算法控制下微电网系统内部的电力平衡。在该调度场景中，系统面临着风光出力（绿色区域）的随机性与负荷（灰色区域）波动性。结果表明，PPO 算法的电网净功率曲线呈现出剧烈的宽频带振荡情况，未能有效平抑风光波动，且出现电网越限冲击现象。SAC 和 TD3 算法虽然维持了基本的能量平衡，但其净功率曲线在多个时段（如第 36、70、96 h 附近）出现了大幅度的功率突变。相比之下，所提 SpecTD3 算法在整个调度周期内，电网净功率曲线始终保持着高度的平滑稳定，在 200 kW 基准线附近波动。因此，SpecTD3 算法嵌入的频域高频能量惩罚机制，可对异构储能下达的高频跳变指令进行严厉约束，引导智能体学习协同出力策略：主动利用异构储能去吸收源荷功率中的高波动分量，将无冲击的低频能量流分给主电网。



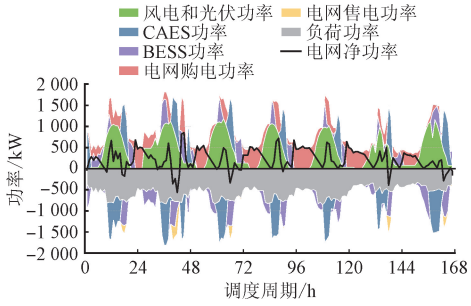
（a）SpecTD3 算法



（b）TD3 算法



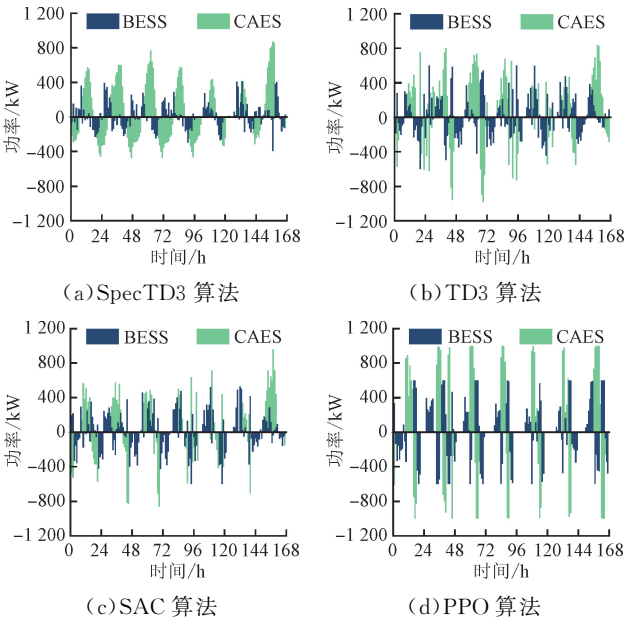
(c)SAC 算法



(d)PPO 算法

图 6　不同算法下电力系统能量平衡演化对比

Fig. 6　Evolutionary comparison of power balance under different algorithms

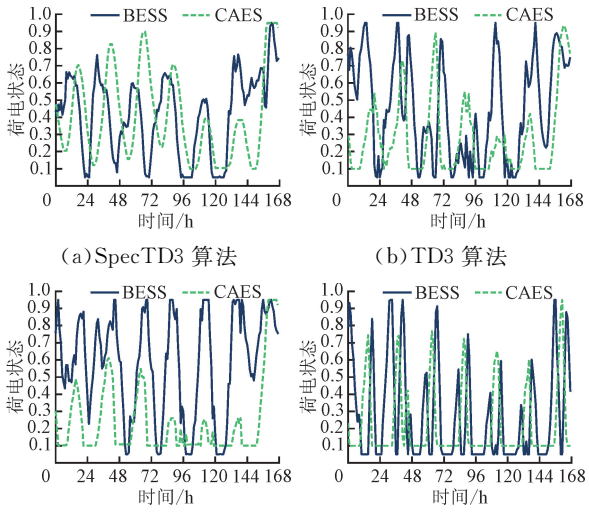


(a)SpecTD3 算法　(b)TD3 算法

(c)SAC 算法　(d)PPO 算法

图 7 不同算法下异构储能出力对比

Fig. 7　Comparison of power output for heterogeneous energy storage under different algorithms



(a)SpecTD3 算法　(b)TD3 算法

(c)SAC 算法　(d)PPO 算法

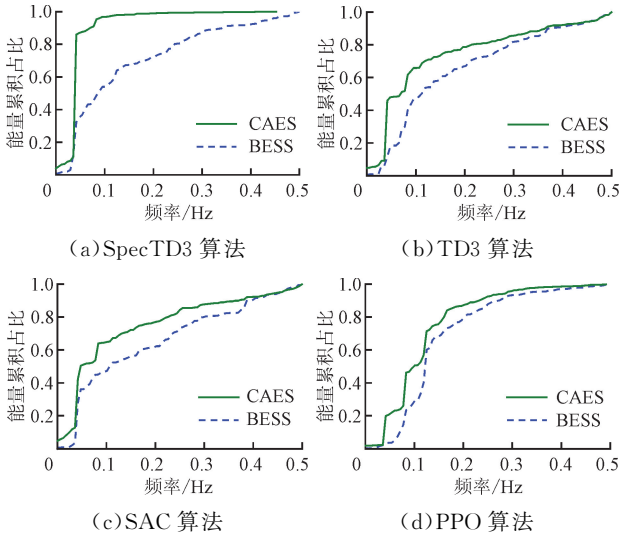
图 8　不同算法下异构储能 SoC 对比

Fig. 8　Comparison of SoC for heterogeneous energy storage under different algorithms

图 7 与图 8 分别展示了不同算法控制下 BESS 与 CAES 的充放电功率分布及对应的 SoC 演化轨迹。结果表明，TD3、SAC 和 PPO 3 种基准算法的出力曲线，未能挖掘出异构储能的底层动态差异，属于同质化调度。具备旋转机械结构的 CAES(浅绿色)频繁出现剧烈高频功率跳变与快速启停。图 8(b)、8(c)、8(d)显示 BESS 的 SoC 轨迹频繁在安全边界(0.05～0.95)之间大幅震荡，经历了多次全深度充放电循环。这种无序的调度策略不仅未能发挥混合储能的互补优势，反而会加速 BESS 的电化学寿命衰减和 CAES 机组的机械磨损与运行风险。相比之下，如图 7(a)与图 8(a)所示，所提 SpecTD3 算法具备时频互补协同调度特性。在图 7(a)中，系统呈现出明确的频域分工：CAES 承担了平滑且连续的能量转移任务，而 BESS 则充当了功率缓冲池，其出力形态呈现出高频、快速、低幅值的特点。图 8(a)表明，CAES 的 SoC 呈现出长周期、大深度平缓充放电循环，而 BESS 的 SoC 则分布在中等 SoC 区间，避免了深度循环疲劳。综上所述，SpecTD3 算法引入的频域约束模块从底层梯度阻断了 CAES 执行高频跳变指令。该机制不仅显式地保障了大型储能设备的安全平稳运行，还隐式地引导 BESS 发挥其快速响应优势，实现了异构储能的互补分工。

图 9 从频域能量累积的视角，直观量化了各算法对异构储能频域分工的控制效果。如图 9(b)、9(c)、9(d)所示，传统基准算法控制下，CAES(绿色实线)与 BESS(蓝色虚线)的能量累积曲线更贴近，表明两者在响应频段上存在交叉耦合，因此偏向无侧重的同质化运行状态。如图 9(a)所示，所提 SpecTD3 算法中，CAES 能量累积曲线在低频段迅速攀升，将 90%以上的能量吞吐集中在长时尺度、极平缓的功率响应区间，而 BESS 的能量累积曲线

则平缓地覆盖了应对瞬态波动的中高频段。



(a)SpecTD3 算法 (b)TD3 算法

(c)SAC 算法 (d)PPO 算法

图 9 不同算法下储能能量分配频域特性对比

Fig. 9 Comparison of energy allocation frequency domain characteristics under different algorithms

表 4 展示了不同算法调度下的策略评估指标对比结果。在系统经济性与并网质量方面，SpecTD3 算法的综合运行成本 C 为 5 241.59 美元，相较于 TD3 基准算法降低了 10.0%。同时，SpecTD3 的并网功率跟踪误差 E_G 为 74.77 kW，较基准算法最高降低了约 68.5%。在 168 h 调度期内，SpecTD3 算法下的电网爬坡越限次数 N_r 被控制在 5 次，相比 SAC 算法降低了 92.1%，大大降低了电网运行风险。结合图 3 与图 4 的源荷数据可见，越限主要出现在风速由高位骤变的时段(如第 48 h 附近)、电负荷夜谷向日峰过渡叠加光伏出力增加的复合突变时刻(第 144 h 附近)，均属源荷净功率瞬态变化幅度逼近储能聚合可调能力上限的极端工况。从设备健康度的量化指标来看，SpecTD3 控制下的 CAES 机械寿命指标 M_C 和 BESS 的等效全循环次数 F_C 各低于基准算法 92% 和 53.8%。上述数据表明，SpecTD3 算法可兼顾系统设备运行寿命、综合经济效益和并网安全。

表 4 不同算法调度下的策略评估指标对比

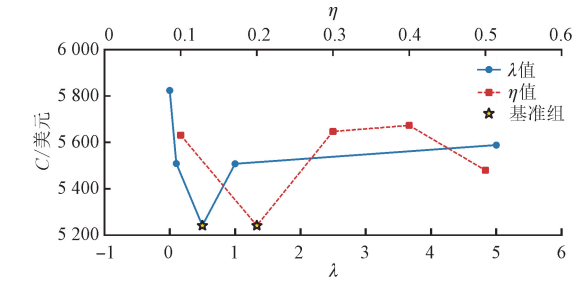
Table 4 Comparison of policy evaluation metrics under different algorithm scheduling

算法	C/美元	E_G/kW	N_r	M_C	F_C
SpecTD3	5 241.59	74.77	5	2.23	4.49
TD3	5 823.56	136.88	29	18.72	7.63
SAC	5 448.25	237.39	63	27.81	9.72
PPO	5 598.68	124.27	30	12.50	6.98

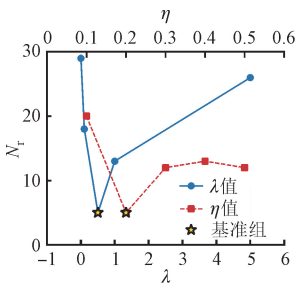
4.3 超参数敏感性分析

所提 SpecTD3 算法中，频域约束模块的核心性能受控于两个关键参数：频域惩罚系数 λ 与截止频率占比 η。前者决定算法对动作指令的抑制权重，后者划定异构储能高低频响应的边界。为此，本节开展了参数敏感性分析，各项核心评估指标随参数变化的趋势如图 10 所示，图例中的“基准组”，即本仿真采用的参数取值($\lambda = 0.5$，$\eta = 0.20$)。

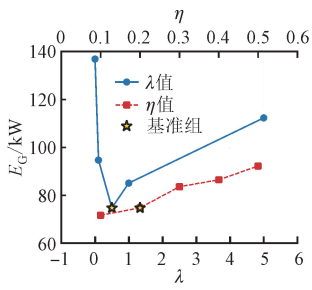
图 10 中蓝色实线表明系统各项指标随 λ 呈典型的 U 型非线性变化特征。当 $\lambda=0$ 时，频域约束模块失效，算法退化为基准 TD3，系统丧失对高频波动的抑制能力。随 λ 增大至 0.5，各项指标收敛至最小值。当 λ 继续增大时，系统指标不降反升。这是因为 DRL 探索过程中存在过度正则化问题。过高的惩罚权重会使智能体为规避单步惩罚，过度限制 CAES 的出力变化，使其丧失正常的功率调节能力。图 10 中红色虚线展示了截止频率占比 η 对调度性能的影响规律。结果表明，当 η 取值较小时，高通滤波器的拦截频带过于严苛，图 10(b)中的爬坡越限次数 N_r 处于畸高水平。这是因为过严的滤波器阻断了 CAES 绝大部分的响应权限，功率调节负担被压在 BESS 上，导致频繁触发并网越限惩罚。随着 η 增大，高通滤波器的物理截止门槛被不断放宽，大量本应由 BESS 处理的中高频源荷波动信号绕过惩罚机制，下发给 CAES 承担，使得 CAES 被强制参与高频调节，频繁的动作指令导致图 10(d)中的机械寿命指标呈现出明显的恶化趋势。



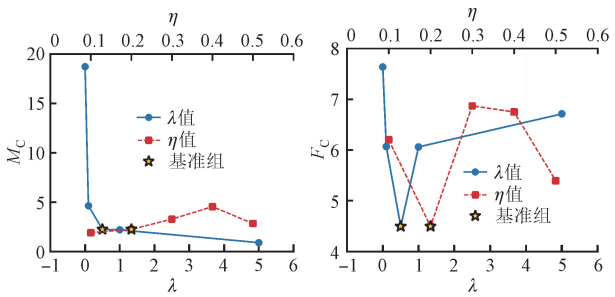
(a)总成本变化图



(b)电网爬坡越线次数



(c)电网功率跟踪误差



(d)CAES 机械寿命指标　　(e)BESS 循环次数

图 10　算法参数敏感性分析结果

Fig. 10　Sensitivity analysis results for algorithms parameters

量化分析结果表明，频域惩罚系数 λ 在[0.5，5.0]的区间内，对微电网综合调度性能具备一定鲁棒性。当 λ 从最优设定点 0.5 增幅达 10 倍时，系统综合运行成本性能恶化率在 6.5%以内。且 CAES 的机械寿命指标始终被有效限制在 2.5 以下的低损耗区间。表明所提协同调度算法对该参数的过估计问题具有较强的容错边界。截止频率占比 η 在 0.1～0.5 区间内取值，系统总成本、跟踪误差值和 CAES 的机械寿命指标相比基准的偏移均可控制在 10%以内，越限次数和 BESS 循环次数指标在最优值附近可维持较低水平，表明算法具有一定鲁棒性。

5　结　论

针对微电网异构混合储能协同调度面临的强耦合、决策域收缩以及传统 DRL 策略同质化问题，建立了面向异构储能协同运行的 MDP 模型，并提出一种基于频域约束的 SpecTD3 优化算法。该方法通过在策略网络更新中引入谱正则化频域惩罚机制，从底层显式抑制了调度指令的瞬态跳变分量，实现了 BESS 与 CAES 的时频解耦与互补协同。多算法对比仿真与敏感性分析结果可得到如下结论。

(1)所提 SpecTD3 算法相较于主流 DRL 基准算法，将系统综合运行成本降低了 10.0%，并网功率跟踪误差下降 68.5%，爬坡越线次数减少 92.1%，同时使 CAES 机械寿命指标与 BESS 等效全循环次数分别下降 92%与 53.8%，兼顾了电网交互质量与底层设备的长期健康。

(2)敏感性分析结果表明，所提算法对核心参数具有较强的鲁棒性：当频域惩罚系数较最优定点大幅增幅 10 倍时，系统综合运行成本恶化率仍被抑制在 6.5%以内；截止频率占比在[0.1，0.5]区间内，系统总成本、跟踪误差及储能寿命等核心指标相比基准值的偏移量可控制在 10% 以内。

参考文献：

[1] Impram S，Varbak Nese S，Oral B. Challenges of renewable energy penetration on power system flexibility：a survey [J]. Energy Strategy Reviews，2020，31：100539.

[2] Zang Tianlei，Wang Shijun，Wang Zian，et al. Integrated planning and operation dispatching of source-grid-load-storage in a new power system：a coupled socio-cyber-physical perspective [J]. Energies，2024，17(12)：3013.

[3] Ahmethodžić L，Musić M，Huseinbegović S. Micro-grid energy management：classification，review and challenges [J]. CSEE Journal of Power and Energy Systems，2023，9(4)：1425-1438.

[4] Rezaeimozafar M，Monaghan R F D，Barrett E，et al. A review of behind-the-meter energy storage systems in smart grids [J]. Renewable and Sustainable Energy Reviews，2022，164：112573.

[5] Rana M M，Uddin M，Sarkar M R，et al. Applications of energy storage systems in power grids with and without renewable energy integration：a comprehensive review [J]. Journal of Energy Storage，2023，68：107811.

[6] Wang Yujie，Wang Li，Li Mince，et al. A review of key issues for control and management in battery and ultra-capacitor hybrid energy storage systems [J]. eTransportation，2020，4：100064.

[7] Fang Ziyu，Shek J K H，Sun Wei. A review of grid-connected hybrid energy storage systems：sizing configurations，control strategies，and future directions [J]. Journal of Energy Storage，2025，118：116226.

[8] Agyekum E B，Abdullah M，Rashid F L，et al. Compressed air energy storage (CAES) systems：technological progress，challenges，and future prospects in renewable energy grids [J]. Energy Conversion and Management：X，2025，28：101402.

[9] Zeynali S，Rostami N，Ahmadian A，et al. Robust multi-objective thermal and electrical energy hub management integrating hybrid battery-compressed air energy storage systems and plug-in-electric-vehicle-based demand response [J]. Journal of Energy Storage，2021，35：102265.

[10] Cui Feifei，An Dou，Xi Huan. Integrated energy hub dispatch with a multi-mode CAES-BESS hybrid system：an option-based hierarchical reinforcement learning approach [J]. Applied Energy，2024，374：123950.

[11] Xiao Gang，Xu Fen，Tong Lianghuai，et al. A hybrid energy storage system based on self-adaptive variational mode decomposition to smooth photovoltaic power

fluctuation [J]. Journal of Energy Storage, 2022, 55: 105509.

[12] Chu Zhongda, Zhang Ning, Teng Fei. Frequency-constrained resilient scheduling of microgrid: a distributionally robust approach [J]. IEEE Transactions on Smart Grid, 2021, 12(6): 4914-4925.

[13] Nakiganda A M, Dehghan S, Markovic U, et al. A stochastic-robust approach for resilient microgrid investment planning under static and transient islanding security constraints [J]. IEEE Transactions on Smart Grid, 2022, 13(3): 1774-1788.

[14] Babu V V, Preetha Roselyn J, Sundaravadivel P. Multi-objective genetic algorithm based energy management system considering optimal utilization of grid and degradation of battery storage in microgrid [J]. Energy Reports, 2023, 9: 5992-6005.

[15] 陈明昊，孙毅，谢志远. 基于双层深度强化学习的园区综合能源系统多时间尺度优化管理 [J]. 电工技术学报，2023，38(7)：1864-1881.
Chen Minghao, Sun Yi, Xie Zhiyuan. The multi-timescale management optimization method for park integrated energy system based on the bi-layer deep reinforcement learning [J]. Transactions of China Electrotechnical Society, 2023, 38(7): 1864-1881.

[16] 熊珞琳，毛帅，唐漾，等. 基于强化学习的综合能源系统管理综述 [J]. 自动化学报，2021，47(10)：2321-2340.
Xiong Luolin, Mao Shuai, Tang Yang, et al. Reinforcement learning based integrated energy system management: a survey [J]. Acta Automatica Sinica, 2021, 47(10): 2321-2340.

[17] Zheng Yi, Jia Jinhua, An Dou. Energy management for microgrids with hybrid hydrogen-battery storage: a reinforcement learning framework integrated multi-objective dynamic regulation [J]. Processes, 2025, 13(8): 2558.

[18] Kumar K, Kwon S, Bae S. Deep reinforcement learning-based control strategy for integration of a hybrid energy storage system in microgrids [J]. Journal of Energy Storage, 2025, 108: 114936.

[19] Wang Xin, Zhou Jianshu, Qin Bin, et al. Coordinated control of wind turbine and hybrid energy storage system based on multi-agent deep reinforcement learning for wind power smoothing [J]. Journal of Energy Storage, 2023, 57: 106297.

[20] 李鹏，钟瀚明，马红伟，等. 基于深度强化学习的有源配电网多时间尺度源荷储协同优化调控 [J]. 电工技术学报，2025，40(5)：1487-1502.
Li Peng, Zhong Hanming, Ma Hongwei, et al. Multi-timescale optimal dispatch of source-load-storage coordination in active distribution network based on deep reinforcement learning [J]. Transactions of China Electrotechnical Society, 2025, 40(5): 1487-1502.

[21] Cui Feifei, An Dou, Teng Shiyang, et al. Cogeneration systems of solar energy integrated with compressed air energy storage systems: a comparative study of various energy recovery strategies [J]. Case Studies in Thermal Engineering, 2023, 51: 103521.

[22] Xu Bolun, Zhao Jinye, Zheng Tongxin, et al. Factoring the cycle aging cost of batteries participating in electricity markets [J]. IEEE Transactions on Power Systems, 2018, 33(2): 2248-2259.

[23] Chen Junyu, Zhao Yue, Wang Minghao, et al. Multi-timescale reward-based DRL energy management for regenerative braking energy storage system [J]. IEEE Transactions on Transportation Electrification, 2025, 11(3): 7488-7500.

[24] Egbomwan O E, Chaoui H, Liu Shichao. A physics-constrained TD3 algorithm for simultaneous virtual inertia and damping control of grid-connected variable speed DFIG wind turbines [J]. IEEE Transactions on Automation Science and Engineering, 2025, 22: 958-969.

[25] 高绪冲，王朝民，张展泳，等. 融合混合储能的微电网多时间尺度优化调度策略 [J/OL]. 南方电网技术，2026-01-05. https://link.cnki.net/urlid/44.1643.tk.20260105.1429.016.
Gao Xuchong, Wang Chaomin, Zhang Zhanyong, et al. Multi-time-scale optimal scheduling strategy for Microgrids integrating hybrid energy storage [J/OL]. Southern Power System Technology, 2026-01-05. https://link.cnki.net/urlid/44.1643.tk.20260105.1429.016.

[26] Li Li, Wang Jing, Zhong Xiaoyi, et al. Combined multi-objective optimization and agent-based modeling for a 100% renewable island energy system considering power-to-gas technology and extreme weather conditions [J]. Applied Energy, 2022, 308: 118376.

（编辑 武红江）