

基于多智能体强化学习的储能电能量-调频市场交易策略

张新琪¹, 柳志海², 陈晓晨³, 詹爽³, 程永卓⁴, 何智龙⁵, 杨强²

(1. 大唐海南文昌新能源有限公司, 海南 文昌 571332; 2. 浙江大学 电气工程学院, 浙江 杭州 310027;
3. 大唐(文昌)储能科技有限公司, 海南 文昌 571300; 4. 中国大唐集团海外投资有限公司, 北京 100052;
5. 中国大唐集团科学技术研究总院有限公司华东电力试验研究院, 安徽 合肥 230061)

摘要: 为了提高独立储能在新型电力系统中的市场竞争力, 提出一种基于多智能体强化学习的电能量-调频辅助服务市场联合交易策略。根据南方区域电力市场规则, 构建包含火电、新能源与独立储能的联合出清模型; 在马尔可夫博弈框架下, 将各发电主体建模为智能体, 采用多智能体强化学习算法, 实现非完全信息条件下的策略性报价优化; 为了刻画储能与系统运行之间的跨时段耦合特性, 设计多维状态与动作空间的建模方法, 以实现多时段协同优化。基于IEEE 30节点系统的仿真结果表明, 所提算法可使智能体总收益相较于单一智能体算法提升18.8%。进一步分析结果表明, 随着参与策略性报价的机组数量增加, 各市场主体总体收益呈上升趋势。研究结果验证了所提多智能体强化学习框架在提升储能市场收益与优化系统整体运行效率方面的有效性。

关键词: 独立储能; 电能量市场; 调频辅助服务市场; 联合交易策略; 多智能体; 强化学习

中图分类号: TM73

文献标志码: A

DOI: 10.16081/j.epae.202603013

0 引言

根据国家能源局发布的《加快构建新型电力系统行动方案(2024—2027年)》, 到2025年, 非化石能源消费比重达到20%左右, 风电、光伏等新能源逐步成为电力系统的主体能源^[1]。新能源占比的提升会给电力系统的安全稳定运行带来挑战, 而储能系统被视为应对此类挑战的有效技术手段^[2-4]。与传统发电机组相比, 储能系统具备更快的响应速度、更强的功率调节能力以及能量时空转移特性, 已成为构建新型电力系统的关键技术^[5]。具体而言, 储能系统能够通过提供调频、备用等辅助服务, 增强电网的灵活性与平衡能力^[6]。

为了加快推进储能技术发展并充分发挥其作用, 我国已出台多项支持政策。国家能源局于2024年4月发布了《关于促进新型储能并网和调度运用的通知》, 明确提出以市场化方式推动新型储能的调度运用^[7]。南方电网2025年新修订的“两个细则”中进一步规范了储能并网及参与辅助服务的管理机制, 并同步发布《南方区域调频辅助服务市场交易实施细则》, 明确了独立储能电站等第三方主体的市场准入机制。此外, 针对新能源发展问题, 国家发展改革委、国家能源局发布了《关于深化新能源上网电价市场化改革 促进新能源高质量发展的通知》, 推动新

能源上网电量全面进入电力市场, 通过市场交易形成价格^[8]。因此, 在储能参与电力市场的交易模型中进一步考虑新能源机组的加入, 更加符合当前电力市场政策的发展趋势^[9-10]。

已有较多关于储能参与电力市场的机制设计与调度运行技术方面的研究。在参与现货市场方面, 文献[11]提出了一种考虑多元储能参与现货市场的调度策略; 文献[12]和文献[13]分别建立了共享储能与租赁储能在电力市场中的交易模型。在参与调频市场方面, 文献[14]综述了储能电站参与调频的竞价策略; 文献[15]提出了综合考虑调节精度、响应时间与调节速度的调频性能评价指标体系; 文献[16]引入时间与空间松弛因子, 优化了市场机制; 文献[17]则在山西调频市场背景下研究了独立储能在调频市场中的运营策略问题。随着储能被允许以报量报价方式参与市场, 当前研究逐渐将其建模为“价格制定者”而非“价格接受者”。文献[18]针对储能难以通过电能量市场回收成本的问题, 引入规划收益补偿机制, 构建了报量报价的双层优化模型; 文献[19]提出了用于优化储能在能量与备用市场中运行决策的双层优化模型; 文献[20-21]将构网型储能视为独立市场主体, 建立了其参与辅助服务市场的双层博弈模型; 文献[22]进一步考虑储能作为惯量与调频资源供应商, 提出了其参与电能量-惯量-调频市场的市场出清模型, 并计及了同步机组启停状态的影响。

然而, 上述研究通常假设各机组的市场报价信息为完全已知, 与实际运行过程中每个发电商仅掌握自身数据的情况不符, 限制了发电主体在不完全

收稿日期: 2026-01-09; 修回日期: 2026-03-12

在线出版日期: 2026-03-18

基金项目: 海南省重点研发项目(ZDYF2024GXJS296)

Project supported by Hainan Provincial Key Research and Development Program(ZDYF2024GXJS296)

信息条件下开展随机博弈竞价的能力^[23]。文献[24]采用马尔可夫博弈描述多个发电商在日前市场中的策略性竞价过程,并提出了一种分层纳什分布式强化学习框架,结合分布式算法模拟了市场出清流程。针对马尔可夫博弈建模,现有研究大多采用多智能体强化学习算法进行求解。文献[25-28]围绕火电、虚拟电厂在电力市场中的调度运行问题,构建了多智能体模型以优化投标策略,其中文献[25]采用了多智能体双延迟深度确定性策略梯度(multi-agent twin delayed deep deterministic policy gradient, MATD3)算法,文献[26]则应用多智能体深度确定性策略梯度(multi-agent deep deterministic policy gradient, MADDPG)算法,并提出了基于蒙特卡罗近似的夏普利值方法进行收益分配。上述研究针对火电机组、虚拟电厂等主体在电力市场中的博弈过程,采用强化学习算法进行求解,然而关于储能在不完美信息条件下作为主体参与电力市场交易策略的研究较少。

基于上述背景,本文提出一种基于多智能体强化学习的储能电能-调频市场交易策略,主要贡献包括以下方面。

1) 本文参考南方区域电力市场交易规则,构建了一个同时涵盖火电、新能源与储能的电能-调频辅助服务联合出清框架,能够真实反映新型电力系统中“高比例新能源+灵活储能资源”共存时的市场运行特性。

2) 本文将电力市场视作典型的不完美信息随机博弈环境,基于马尔可夫博弈模型将火电、储能以及新能源机组分别建模为独立智能体。针对储能荷电状态与系统运行存在显著的跨时段耦合特征,采用24维时序动作与观测空间设计,使得智能体能够在考虑全日约束的条件下制定一致性策略,避免分时段出清难以维护储能荷电状态连续性的技术缺陷。在此基础上,采用MADDPG算法实现中心化训练与分散执行,提高策略学习的稳定性和全局最优性。

3) 本文以IEEE 30节点系统为算例进行仿真,量化比较了单智能体强化学习算法与多智能体强化学习算法在收敛速度、总收益等方面的差异。结果表明,MADDPG算法能提升系统总收益、加快收敛速度,证明了多智能体建模在具有策略互动的市场环境下的优势。此外,本文还深入分析了不同的策略性报价机组数量条件下,各主体总收益的变化规律,揭示了在策略性参与者规模扩大时,市场的竞争-协同机制可以提升系统的总体收益。

1 市场交易模型

本文参照南方区域电力市场现货电能交易与调频辅助服务市场的相关实施细则,构建了相应的市场交易模型。在现货电能市场中,火电机组、风

电机组、光伏机组以及独立储能电站均以报量报价方式参与优化出清;在调频辅助服务市场中,火电机组与独立储能电站共同参与调频市场的出清,且独立储能在调频市场中标时段不再参与现货电能市场的出清。本文采用电能-调频市场联合出清机制,其中电能市场以节点边际价格作为结算依据,调频市场则对调频里程予以补偿(对调频容量不给予补偿)。电能市场与调频辅助服务市场的交易框架如图1所示。

1.1 报价模型

1) 火电机组模型。

a) 火电机组的发电成本可表示为:

$$C_g(P) = \frac{1}{2} k_g P^2 + h_g P + f_g \quad (1)$$

式中: C_g 为火电机组的发电成本; P 为火电机组的发电功率; k_g 、 h_g 、 f_g 为火电机组的发电成本系数。

b) 火电机组的策略性报价如下:

$$b_{i,t}^C = b_{\min}^C + \alpha_{i,t}^C (b_{\max}^C - b_{\min}^C) \quad (2)$$

$$b_{i,t}^{\text{GPR}} = b_{\min}^{\text{GPR}} + \alpha_{i,t}^{\text{GPR}} (b_{\max}^{\text{GPR}} - b_{\min}^{\text{GPR}}) \quad (3)$$

式中: $b_{i,t}^C$ 、 $b_{i,t}^{\text{GPR}}$ 分别为 t 时段火电机组 i 在电能市场、调频辅助服务市场中的报价函数; $\alpha_{i,t}^C$ 、 $\alpha_{i,t}^{\text{GPR}}$ 分别为 t 时段火电机组 i 在电能市场、调频辅助服务市场中的投标变量,由多智能体强化学习算法给出; b_{\max}^C 、 b_{\min}^C 分别为火电机组在电能市场中报价的最大值、最小值; b_{\max}^{GPR} 、 b_{\min}^{GPR} 分别为火电机组在调频市场中报价的最大值、最小值。

c) 火电机组的利润可表示为:

$$r_i^G = \sum_{t=1}^T [\lambda_{i,t}^E p_{i,t}^C + \lambda_{i,t}^{\text{PR}} p_{i,t}^{\text{GPR}} \rho_{\text{mf},i} - C_{g,t}(p_{i,t}^C + p_{i,t}^{\text{GPR}})] \quad (4)$$

式中: T 为1个周期的时段总数; r_i^G 为1个周期内火电机组 i 的总利润; $\lambda_{i,t}^E$ 为 t 时段电能市场中火电机组 i 的出清价格; $\lambda_{i,t}^{\text{PR}}$ 为 t 时段调频辅助服务市场的调频里程价格; $p_{i,t}^C$ 、 $p_{i,t}^{\text{GPR}}$ 分别为 t 时段火电机组 i 在电能市场中的中标功率、在调频辅助服务市场中的中标容量; $C_{g,t}(p_{i,t}^C + p_{i,t}^{\text{GPR}})$ 为 t 时段火电机组 i 的发电成本; $\rho_{\text{mf},i}$ 为火电机组 i 的调频里程系数。

2) 储能模型。

a) 储能成本可表示为:

$$C_e = c_1 P_k^{\text{ES}} + c_2 \quad (5)$$

式中: C_e 为储能成本; c_1 为储能装置的充放电成本系数; c_2 为储能装置的运行成本系数; P_k^{ES} 为储能装置 k 的充放电功率。

b) 储能的策略性报价为:

$$b_{k,t}^{\text{dis}} = b_{\min}^{\text{dis}} + \alpha_{k,t}^{\text{dis}} (b_{\max}^{\text{dis}} - b_{\min}^{\text{dis}}) \quad (6)$$

$$b_{k,t}^{\text{ch}} = b_{\min}^{\text{ch}} + \alpha_{k,t}^{\text{ch}} (b_{\max}^{\text{ch}} - b_{\min}^{\text{ch}}) \quad (7)$$

$$b_{k,t}^{\text{EPR}} = b_{\min}^{\text{EPR}} + \alpha_{k,t}^{\text{EPR}} (b_{\max}^{\text{EPR}} - b_{\min}^{\text{EPR}}) \quad (8)$$

式中: $b_{k,t}^{\text{dis}}$ 、 $b_{k,t}^{\text{ch}}$ 、 $b_{k,t}^{\text{EPR}}$ 分别为 t 时段储能装置 k 的放电、充电、调频报价函数; b_{\max}^{dis} 、 b_{\min}^{dis} 分别为储能放电报价的最大值、最小值; b_{\max}^{ch} 、 b_{\min}^{ch} 分别为储能充电报价的最

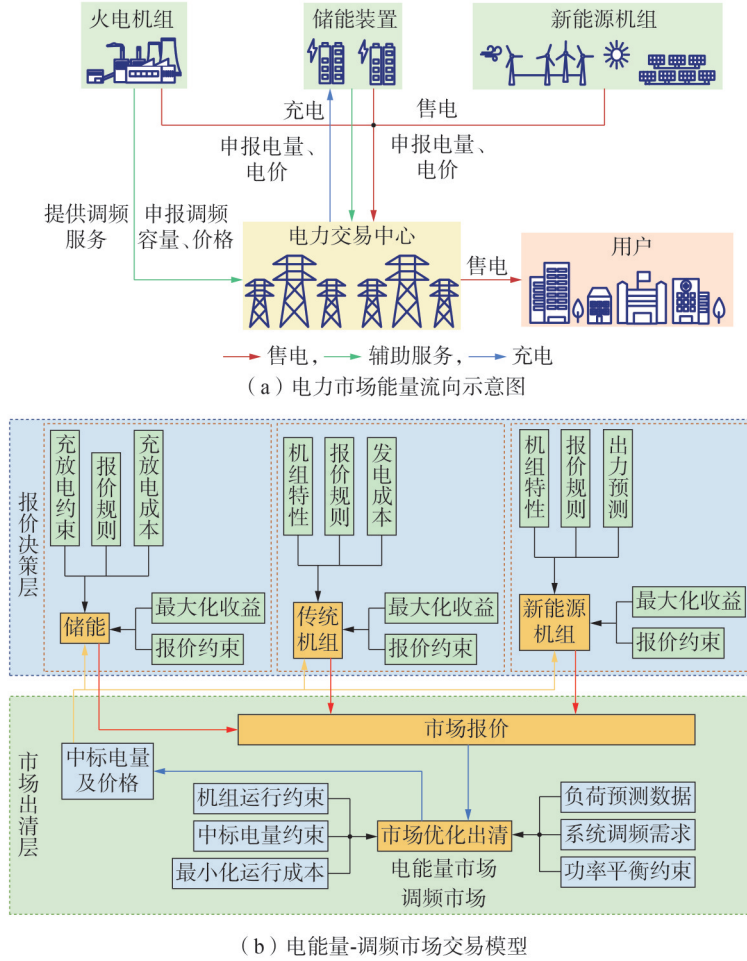


图 1 电能-调频市场交易框架

Fig.1 Trading framework of electric energy and frequency regulation markets

大值、最小值; b_{\max}^{EPR} 、 b_{\min}^{EPR} 分别为储能调频报价的最大值、最小值; $\alpha_{k,t}^{\text{dis}}$ 、 $\alpha_{k,t}^{\text{ch}}$ 、 $\alpha_{k,t}^{\text{EPR}}$ 分别为 t 时段储能装置 k 的放电、充电、调频投标变量。

c) 储能利润可表示为:

$$r_k^{\text{E}} = \sum_{t=1}^T [\lambda_{k,t}^{\text{E}} (p_{k,t}^{\text{dis}} - p_{k,t}^{\text{ch}}) + \lambda_t^{\text{PR}} p_{k,t}^{\text{EPR}} \rho_{\text{mf},k} - C_{e,t} (p_{k,t}^{\text{dis}} + p_{k,t}^{\text{ch}} + p_{k,t}^{\text{EPR}})] \quad (9)$$

式中: r_k^{E} 为 1 个周期内储能装置 k 获得的总利润; $p_{k,t}^{\text{ch}}$ 、 $p_{k,t}^{\text{dis}}$ 分别为 t 时段储能装置 k 在电能市场中的中标充、放电功率; $p_{k,t}^{\text{EPR}}$ 为 t 时段储能装置 k 在调频辅助服务市场中的中标容量; $\lambda_{k,t}^{\text{E}}$ 为 t 时段电能市场中储能装置 k 的出清价格; $\rho_{\text{mf},k}$ 为储能装置 k 的调频里程系数; $C_{e,t} (p_{k,t}^{\text{dis}} + p_{k,t}^{\text{ch}} + p_{k,t}^{\text{EPR}})$ 为 t 时段储能装置 k 的成本。

3) 新能源机组模型。

a) 新能源机组的策略性报价为:

$$b_{m,t}^{\text{PV}} = b_{\min}^{\text{PV}} + \alpha_{m,t} (b_{\max}^{\text{PV}} - b_{\min}^{\text{PV}}) \quad (10)$$

$$b_{n,t}^{\text{WT}} = b_{\min}^{\text{WT}} + \alpha_{n,t} (b_{\max}^{\text{WT}} - b_{\min}^{\text{WT}}) \quad (11)$$

式中: $b_{m,t}^{\text{PV}}$ 、 $b_{n,t}^{\text{WT}}$ 分别为 t 时段光伏机组 m 、风电机组 n 在电能市场中的报价函数; b_{\max}^{PV} 、 b_{\min}^{PV} 分别为光伏机组在电能市场中报价的最大值、最小值; b_{\max}^{WT} 、 b_{\min}^{WT}

分别为风电机组在电能市场中报价的最大值、最小值; $\alpha_{m,t}$ 、 $\alpha_{n,t}$ 分别为 t 时段光伏机组 m 、风电机组 n 在电能市场的投标变量。

b) 新能源机组的利润可表示为:

$$r_m^{\text{PV}} = \sum_{t=1}^T \lambda_{m,t}^{\text{E}} p_{m,t}^{\text{PV}} \quad (12)$$

$$r_n^{\text{WT}} = \sum_{t=1}^T \lambda_{n,t}^{\text{E}} p_{n,t}^{\text{WT}} \quad (13)$$

式中: r_m^{PV} 、 r_n^{WT} 分别为 1 个周期内光伏机组 m 、风电机组 n 获得的利润; $p_{m,t}^{\text{PV}}$ 、 $p_{n,t}^{\text{WT}}$ 分别为 t 时段光伏机组 m 、风电机组 n 在电能市场中的中标功率; $\lambda_{m,t}^{\text{E}}$ 、 $\lambda_{n,t}^{\text{E}}$ 分别为 t 时段电能市场中光伏机组 m 、风电机组 n 的出清价格。

1.2 市场出清模型

电能-调频市场联合出清的目标函数可表示为:

$$F_L = \sum_{i=1}^T \left[\sum_{i=1}^{N_G} (b_{i,t}^{\text{G}} p_{i,t}^{\text{G}} + b_{i,t}^{\text{GPR}} p_{i,t}^{\text{GPR}} \rho_{\text{mf},i}) + \sum_{m=1}^{N_{\text{PV}}} b_{m,t}^{\text{PV}} p_{m,t}^{\text{PV}} + \sum_{n=1}^{N_{\text{WT}}} b_{n,t}^{\text{WT}} p_{n,t}^{\text{WT}} + \sum_{k=1}^{N_E} (b_{k,t}^{\text{dis}} p_{k,t}^{\text{dis}} - b_{k,t}^{\text{ch}} p_{k,t}^{\text{ch}} + b_{k,t}^{\text{EPR}} p_{k,t}^{\text{EPR}} \rho_{\text{mf},k}) \right] \quad (14)$$

式中: F_L 为1个周期内系统的总成本; N_G 、 N_{PV} 、 N_{WT} 、 N_E 分别为火电机组、光伏机组、风电机组、储能装置数量。

需要满足的约束条件如下。

1) 网络潮流约束。

各节点的功率平衡约束如式(15)所示, 各支路之间的潮流约束如式(16)和式(17)所示。

$$\sum_{i \in \Phi_g} P_{i,t}^G + \sum_{m \in \Phi_{pv}} P_{m,t}^{PV} + \sum_{n \in \Phi_{wt}} P_{n,t}^{WT} + \sum_{k \in \Phi_{ess}} (p_{k,t}^{dis} - p_{k,t}^{ch}) = \sum_{d \in \Phi_d} P_{d,t}^L + \sum_{(j,q) \in \Phi} P_{j,q,t} \quad (15)$$

$$P_{j,q,t} = \frac{\theta_{j,t} - \theta_{q,t}}{X_{j,q}} \quad (j,q) \in \Phi, \forall t \quad (16)$$

$$-p_{j,q}^{\max} \leq P_{j,q,t} \leq p_{j,q}^{\max} \quad (j,q) \in \Phi, \forall t \quad (17)$$

式中: Φ_g 、 Φ_{pv} 、 Φ_{wt} 、 Φ_{ess} 、 Φ_d 分别为火电机组、光伏机组、风电机组、储能装置、负荷所连节点集合; Φ 为系统支路集合; $p_{d,t}^L$ 为 t 时段负荷 d 的功率; $p_{j,q,t}$ 为 t 时段节点 j 、 q 之间支路的潮流功率; $\theta_{j,t}$ 、 $\theta_{q,t}$ 分别 t 时段节点 j 、 q 的电压相角; $X_{j,q}$ 为节点 j 、 q 之间支路的电抗; $p_{j,q}^{\max}$ 为节点 j 、 q 之间支路潮流功率的最大值。

2) 调频容量约束。

$$\sum_{i=1}^{N_G} P_{i,t}^{GPR} + \sum_{k=1}^{N_E} P_{k,t}^{EPR} = Q_t^D \quad (18)$$

$$\sum_{i=1}^{N_G} P_{i,t}^{GPR} \rho_{mf,i} + \sum_{k=1}^{N_E} P_{k,t}^{EPR} \rho_{mf,k} = M_t^D \quad (19)$$

式中: Q_t^D 为 t 时段系统的调频容量需求; M_t^D 为 t 时段系统的调频里程需求。

3) 储能运行约束。

$$S_{k,t}^{E, \min} \leq S_{k,t}^E \leq S_{k,t}^{E, \max} \quad (20)$$

$$S_{k,t}^E = S_{k,t-1}^E + (\eta_k^{ch} p_{k,t}^{ch} - p_{k,t}^{dis} / \eta_k^{dis}) \Delta t \quad (21)$$

$$S_{k,1}^E = S_{k,24}^E = S_{k, \text{initial}}^E \quad (22)$$

$$\begin{cases} \mu_{k,t}^{ch} + \mu_{k,t}^{dis} + \mu_{k,t}^{pr} \leq 1 \\ 0 \leq p_{k,t}^{dis} \leq \mu_{k,t}^{dis} P_{k, \max}^{dis} \\ 0 \leq p_{k,t}^{ch} \leq \mu_{k,t}^{ch} P_{k, \max}^{ch} \\ 0 \leq p_{k,t}^{EPR} \leq \mu_{k,t}^{pr} P_{k, \max}^{EPR} \end{cases} \quad (23)$$

式中: $S_{k,t}^E$ 为 t 时段储能装置 k 的荷电状态; $S_{k, \max}^E$ 、 $S_{k, \min}^E$ 分别为储能装置 k 荷电状态的最大值、最小值; $S_{k, \text{initial}}^E$ 为储能装置 k 的初始荷电状态, 式(22)限定储能装置 k 在初始时刻的荷电状态 $S_{k,1}^E$ 与在结束时刻的荷电状态 $S_{k,24}^E$ 相等; η_k^{ch} 、 η_k^{dis} 分别为储能装置 k 的充、放电效率; Δt 为单位时段时间间隔; $\mu_{k,t}^{ch}$ 、 $\mu_{k,t}^{dis}$ 、 $\mu_{k,t}^{pr}$ 分别为 t 时段储能装置 k 的充电、放电、调频状态变量, 若处于相应的状态则取值为1, 否则取值为0; $P_{k, \max}^{ch}$ 、 $P_{k, \max}^{dis}$ 分别为储能装置 k 的最大充、放电功率; $P_{k, \max}^{EPR}$ 为储能装置 k 参与调频的最大功率。

4) 火电机组运行约束。

$$P_i^{G, \min} \leq p_{i,t}^G + p_{i,t}^{GPR} \leq P_i^{G, \max} \quad (24)$$

$$-\Delta p_i^{Gd} \leq p_{i,t}^G - p_{i,t-1}^G \leq \Delta p_i^{Gu} \quad (25)$$

式中: $P_i^{G, \max}$ 、 $P_i^{G, \min}$ 分别为火电机组 i 输出功率的最大值、最小值; Δp_i^{Gu} 、 Δp_i^{Gd} 分别为火电机组 i 向上、向下爬坡速率的最大值。

5) 新能源机组相关约束。

$$\beta P_{m,t}^{PV, \text{for}} \leq p_{m,t}^{PV} \leq P_{m,t}^{PV, \text{for}} \quad (26)$$

$$\beta P_{n,t}^{WT, \text{for}} \leq p_{n,t}^{WT} \leq P_{n,t}^{WT, \text{for}} \quad (27)$$

式中: β 为新能源机组的最小出力系数; $P_{m,t}^{PV, \text{for}}$ 、 $P_{n,t}^{WT, \text{for}}$ 分别为 t 时段光伏机组 m 、风电机组 n 的功率预测值。

2 多智能体强化学习模型

2.1 马尔可夫模型

1) 博弈个体。参与博弈的个体包括火电机组、储能装置、新能源机组。

2) 动作空间, 即各博弈个体的报价策略集合。各博弈个体的报价策略为:

$$a_i = \{b_{i,t}^G, b_{i,t}^{GPR}\} \quad (28)$$

$$a_k = \{b_{k,t}^{dis}, b_{k,t}^{ch}, b_{k,t}^{EPR}\} \quad (29)$$

$$a_m = \{b_{m,t}^{PV}\} \quad (30)$$

$$a_n = \{b_{n,t}^{WT}\} \quad (31)$$

式中: a_i 、 a_k 、 a_m 、 a_n 分别为火电机组 i 、储能装置 k 、光伏机组 m 、风电机组 n 的报价策略集合。

3) 状态空间。

$$s_i = \{\lambda_{i,t}^E, \lambda_{i,t}^{PR}, p_{i,t}^G, p_{i,t}^{GPR}\} \quad (32)$$

$$s_k = \{\lambda_{k,t}^E, \lambda_{k,t}^{PR}, p_{k,t}^{dis}, p_{k,t}^{ch}, p_{k,t}^{EPR}\} \quad (33)$$

$$s_m = \{\lambda_{m,t}^E, p_{m,t}^{PV}\} \quad (34)$$

$$s_n = \{\lambda_{n,t}^E, p_{n,t}^{WT}\} \quad (35)$$

式中: s_i 、 s_k 、 s_m 、 s_n 分别为火电机组 i 、储能装置 k 、光伏机组 m 、风电机组 n 的状态空间。

4) 奖励。

奖励为各主体参与电能量和调频辅助服务市场获得的净利润。其中, 本文在新能源机组的奖励函数中设置了弃风弃光惩罚项, 以保证新能源消纳率。

2.2 MADDPG 算法

深度确定性策略梯度 (deep deterministic policy gradient, DDPG) 算法是一种基于确定性策略梯度的无模型执行者-评价者算法, 适用于连续状态与动作空间中的强化学习任务。DDPG 算法的结构包含2个核心神经网络, 分别为用于拟合策略函数的演员网络 (Actor 网络)、逼近动作值函数的评论家网络 (Critic 网络)。Actor 网络负责生成确定性动作并与环境进行交互, Critic 网络则负责评估状态-动作对的价值, 并引导 Actor 网络的学习方向。在 Critic 网络的参数更新过程中, 采用贝尔曼方程计算目标值, 其损失函数通过最小化时序差分误差构建, 如式(36)所示。

$$Q^\mu(s_t, a_t) = E[r_t + \gamma Q^\mu(s_{t+1}, \mu(s_{t+1}))] \quad (36)$$

式中: Q^μ 为在策略 μ 下, 在状态 s_t 执行动作 a_t 后能获

得的累计期望回报; $E[\cdot]$ 为期望函数; r_t 为在状态 s_t 执行动作 a_t 后直接反馈的奖励; γ 为折扣因子; s_{t+1} 为执行动作 a_t 后环境转移到的新状态; $\mu(s_{t+1})$ 为Actor网络根据下一个状态 s_{t+1} 所决定的下一个动作。

网络损失可表示为:

$$L(\theta^Q) = E_{\mu'} \left[\left(Q(s_t, a_t | \theta^Q) - Q_t^{\text{target}} \right)^2 \right] \quad (37)$$

$$Q_t^{\text{target}} = r_t + \gamma Q(s_{t+1}, \mu(s_{t+1}) | \theta^Q) \quad (38)$$

式中: $L(\theta^Q)$ 为损失函数,用于衡量当前Critic网络预测值与目标值之间的差距; μ' 为更新之后的策略集合; θ^Q 为当前Critic网络的参数; $Q(s_t, a_t | \theta^Q)$ 为当前Critic网络的预测值,即对当前状态 s_t 和动作 a_t 做出的 Q 值估计; Q_t^{target} 为目标 Q 值,根据式(38)计算得到。

策略网络梯度的更新公式为:

$$\nabla_{\theta^\mu} J = \nabla_a Q(s, a | \theta^Q) \Big|_{s=s_t, a=\mu(s)} \nabla_{\theta^\mu} \mu(s | \theta^\mu) \Big|_{s=s_t} \quad (39)$$

式中: $\nabla_{\theta^\mu} J$ 为目标梯度; $\nabla_a Q(s, a | \theta^Q) \Big|_{s=s_t, a=\mu(s)}$ 负责评估在状态 s_t 下采取动作 a_t 的好坏; $\nabla_{\theta^\mu} \mu(s | \theta^\mu) \Big|_{s=s_t}$ 负责计算策略网络的 θ^μ 如何影响其输出的动作。

在MADDPG算法中,每个智能体均配置独立的Actor网络与Critic网络。其中,Actor网络负责学习该智能体的策略函数,而Critic网络用于估计其动作值函数^[18]。MADDPG算法采用“集中式训练、分布式执行”的框架:在训练阶段,各智能体的Critic网络能够获取系统中所有智能体的状态与动作信息,以实现更准确的价值评估与策略引导;在实际执行阶段,每个智能体仅能根据自身的局部观测信息,通过其Actor网络独立做出决策。

策略梯度可表示为:

$$\nabla_{\theta_i} J(u_i) = E_{s, a-D} \left[\nabla_{a_i} Q_i^\mu(s, a_1, \dots, a_N) \Big|_{a_i=\mu_i(s)} \nabla_{\theta_i} \mu(a_i | s) \right] \quad (40)$$

式中: $E_{s, a-D}[\cdot]$ 表示从经验回放缓冲区(D)中采样一批经验数据求期望; $\nabla_{a_i} Q_i^\mu(s, a_1, \dots, a_N) \Big|_{a_i=\mu_i(s)}$ 表示 Q 函数对智能体 i 自身动作 a_i 的梯度,表明当前智能体的动作变化如何影响 Q 值; N 为智能体数量; u_i 为智能体 i 的当前Actor网络。

Critic网络的损失函数 $L(\theta_i)$ 可表示为:

$$L(\theta_i) = E_{s, a, r, s'} \left[\left(Q_i^\mu(s, a_1, \dots, a_N) - Q_i^{\text{target}} \right)^2 \right] \quad (41)$$

$$Q_i^{\text{target}} = r_i + \gamma Q_i^\mu(s', a'_1, \dots, a'_N) \Big|_{a'_j=\mu'_j(o_j)} \quad (42)$$

式中: s' 为环境在当前状态与联合动作后转移到的新状态; a'_j 为智能体 j 的下一步动作; μ'_j 为智能体 j 的目标Actor网络; o_j 为智能体 j 的局部观测; r_i 为智能体 i 的奖励。

代理网络参数的更新公式为:

$$\theta_i^{Q'} \rightarrow \tau \theta_i^{Q'} + (1-\tau) \theta_i^Q \quad (43)$$

$$\theta_i^{\pi'} \rightarrow \tau \theta_i^{\pi'} + (1-\tau) \theta_i^\pi \quad (44)$$

式中: $\theta_i^{Q'}$ 、 θ_i^Q 、 $\theta_i^{\pi'}$ 、 θ_i^π 分别为智能体 i 的目标Critic网络参数、当前Critic网络参数、目标Actor网络参数、当前Actor网络参数; τ 为软更新系数。

2.3 算法求解流程

鉴于策略性报价智能体中包含储能装置,其在市场报价过程中需要满足全天荷电状态约束条件,因此不适宜采用分时段独立出清的方式。为此,本文将所有智能体的动作空间与状态空间设定为涵盖一天24个时段的整体结构,使得每个动作函数 \mathbf{a} 与状态 \mathbf{s} 均为24维向量,因此模型可转化为单步马尔可夫博弈过程。虽然该设置会降低训练效率,但能够有效避免分时段出清对储能荷电状态连续约束条件的破坏,同时,以全天为周期的报价策略也更符合实际市场中集中出清的规则要求。本文所提电力市场交易模型与多智能体强化学习模型算法示意图如图2所示,MADDPG算法的伪代码如附录A表A1所示。

3 算例分析

本文以IEEE 30节点标准测试系统为算例对所提储能参与电力市场的交易模型及求解算法进行仿真分析。系统包含6台火电机组(G_1 — G_6)、2台储能装置(ES_1 、 ES_2)、2台光伏机组(PV_1 、 PV_2)和2台风电机组(WT_1 、 WT_2),各机组的技术参数与报价参数见附录A表A2—A5。其中,5台火电机组、2台储能装置及全部新能源机组(2台光伏机组和2台风电机组)采用策略性报价方式,剩余1台火电机组采用非策略性报价方式。IEEE 30节点标准测试系统的拓扑结构如附录A图A1所示,系统日负荷曲线与新能源典型日出力预测曲线如附录A图A2所示。设置系统的调频总需求为总负荷的8%,历史调频里程与调频容量之比为10。参考文献[19, 23],并结合市场政策进行一定的修改,设定电能市场的报价范围为100~250元/(MW·h),调频辅助服务市场中调频里程的报价范围为25~40元/MW,固定报价机组在2个市场中的报价分别为报价范围的中间值,即175元/(MW·h)和32.5元/MW。MADDPG算法的主要超参数设置如附录A表A6所示。

仿真运行的硬件平台包括AMD Ryzen 7 9700x 8核处理器(5.5 GHz), 48 GB运行内存, NVIDIA GeForce RTX 5080显卡(16 GB显存),软件环境为Python 3.11.14,采用PyTorch框架搭建强化学习算法的神经网络。

3.1 算法收敛性分析

DDPG算法与MADDPG算法在训练过程中各智能体的总奖励曲线如图3所示。

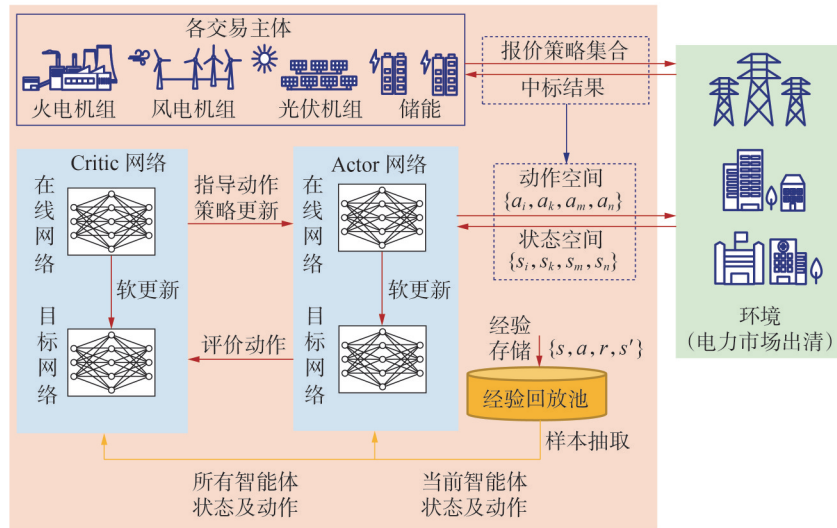


图2 多智能体强化学习模型示意图

Fig.2 Schematic diagram of multi-agent reinforcement learning model

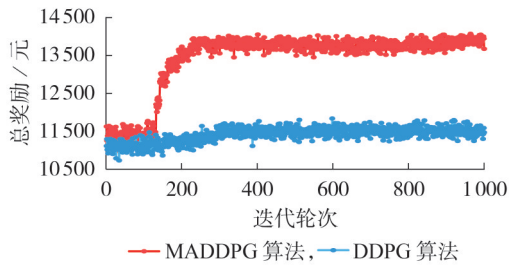


图3 总奖励曲线

Fig.3 Total reward curves

由图3可知, MADDPG算法大约在经过200轮训练后趋于收敛, 而DDPG算法虽然在经过300轮左右训练后也能收敛, 但由于其为单一智能体强化学习架构, 难以在复杂多主体博弈过程中实现全局最优。从收敛后的智能体总奖励(具体计算公式如附录A式(A1)所示)来看, 相较于DDPG算法所得结果11 513.78元, MADDPG算法的结果为13 676.18元, 提升了18.8%; 从市场出清总成本(计算公式见附录A式(A2))来看, DDPG算法所得结果为13 403.34元, MADDPG算法所得结果为15 336.12元。其中, 智能体的总奖励为各策略竞价智能体的净利润之和, 市场出清总成本表示市场的总采购成本, 即机组的净利润与机组运行成本之和。从协同与竞争的角度来看, MADDPG算法的“集中式训练、分布式执行”策略起到了关键作用, 各主体的Critic网络观察各主体的状态, 一直处于竞争状态容易导致总利润降低, 所以主动学习协同策略能够实现整体利益最大化。由于各主体的报价提高, 市场出清总成本随之提升。

3.2 市场出清结果分析

电能量市场出清后的各节点电价如图4所示。由图可知, 位于电网送端的节点的边际电价普遍较低, 而受端节点的边际电价相对较高, 且在部分时段

出清电价超过了机组的报价上限。这一现象主要源于线路传输容量限制引起的阻塞边际成本: 当系统无法按最低成本调度电源以满足全网负荷时, 就会产生额外的阻塞费用。由于电网受端负荷集中且可调用的发电资源有限, 阻塞边际成本显著上升, 从而导致该区域节点的出清电价突破了机组的报价上限。

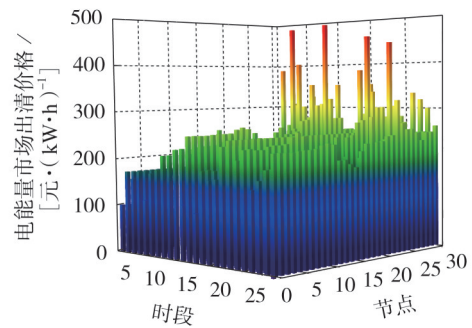


图4 电能量市场出清后的节点电价

Fig.4 Nodal price of electric energy market after clearing

电能量市场中各机组的中标功率如图5(a)所示。可以看出, 系统的大部分出力由火电机组提供, 而储能装置通过在低电价时段充电、高电价时段放电的方式获取电能量市场的收益。值得注意的是, 在少数高电价时段储能装置仍进行充电操作, 这是因为其需要在调频辅助服务市场中预留部分容量; 为了满足后续在调频辅助服务市场的中标条件, 储能装置需要适时在电能量市场中补充电量, 以保障其在调频市场中的可调用能力。调频辅助服务市场的出清价格曲线和各机组的中标功率如图5(b)所示。可以看出, 调频出清价格的变化趋势与系统负荷水平呈现显著正相关性, 具体表现为在负荷较高的时段10—12、17、18, 调频出清价格明显提高。根

据市场规则,储能装置不能同时参与电能量市场与调频辅助服务市场,因此调频性能优异的储能并非持续中标,而是需要在大部分时段通过在电能量市场充放电来维持其运行状态,剩余调频需求则主要由火电机组承担。

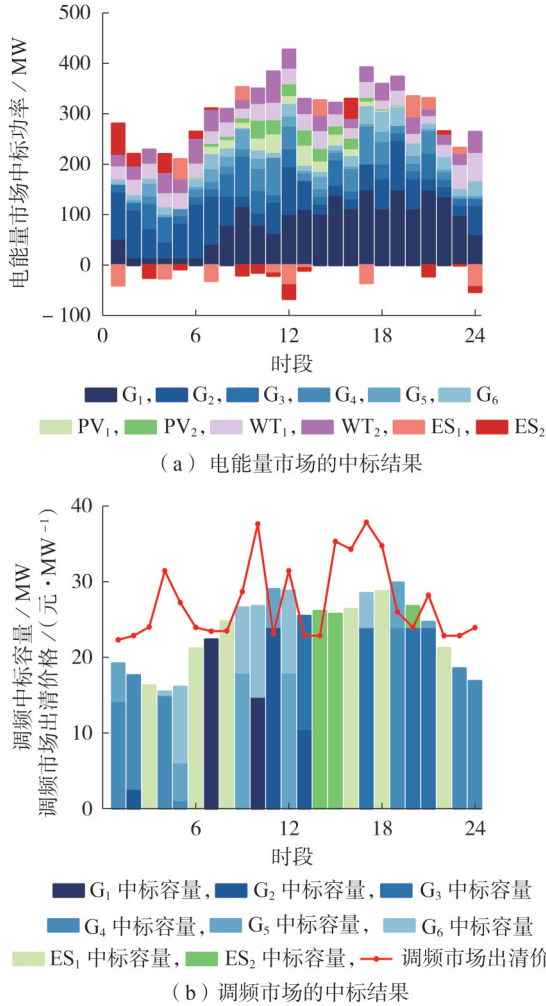


图 5 电能量和调频市场的中标结果

Fig.5 Bidding results of electric energy and frequency regulation markets

3.3 储能收益和报价策略分析

储能装置收益如表 1 所示,火电机组、光伏机组、风电机组的收益见附录 B 表 B1 和表 B2。由表 1 可知,储能装置在调频市场中获得的收益显著高于其在电能量市场中获得的收益。值得注意的是,ES₁、ES₂在电能量市场中产生了净支出(主要用于购电),而其在调频市场中实现了可观的净收益,这表明储能主体通过优先保障调频容量供给、适时在电

表 1 储能装置收益

Table 1 Benefits of energy storage devices

储能装置	电能量收益	调频收益	总收益
ES ₁	-1 725	35 212	33 487
ES ₂	-4 535	21 436	16 901

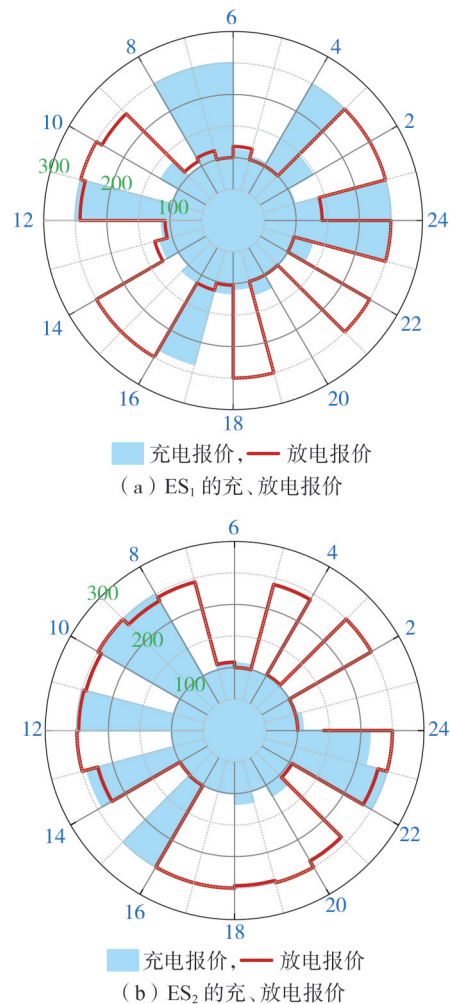
能量市场进行电量补充的策略,实现了整体收益的最大化。

ES₁与 ES₂在电能量市场中的充、放电报价分别如图 6(a)、(b)所示,图中蓝色数值表示时段,绿色数值表示充、放电报价(单位为元/(MW·h))。可以看出,受所在节点位置的影响,2 台储能装置的充、放电报价策略存在显著差异。ES₁在负荷高峰时段未设置过高的放电报价,而是以相对较低的报价确保中标,从而稳定获取电能量市场收益;相较而言,ES₂的充、放电报价策略更为激进,其中标功率总体高于 ES₁,但在部分时段因充电报价过高导致成本上升,反而削弱了其整体收益。

ES₁、ES₂在调频市场中的报价如图 6(c)所示,图中蓝色数值表示时段,绿色数值表示调频报价(单位为元/MW)。由图可知,在调频需求较高的时段(时段 9—13、18—20),储能装置的报价相应提升。由于储能具备更优的调频性能与更高的调频里程系数,系统在出清过程中会优先调用储能资源以满足调频需求。

3.4 策略性报价机组数量的影响分析

为了分析策略性报价机组数量对市场总收益的



(a) ES₁ 的充、放电报价

(b) ES₂ 的充、放电报价

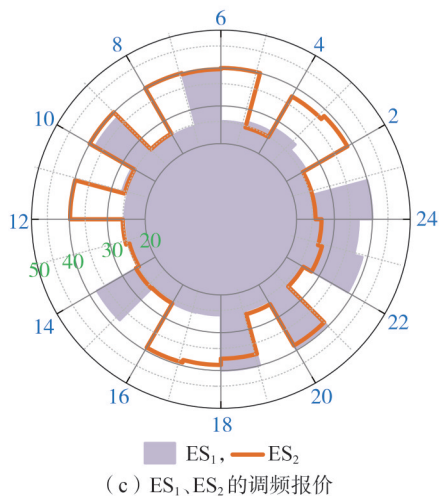


图6 储能装置在电能量和调频市场中的报价
Fig.6 Bidding price of energy storage devices in electric energy and frequency regulation markets

影响,本文分别对系统中设置2—5台策略性报价火电机组的情形进行仿真,各机组总收益如图7所示。由图可知,随着策略性报价火电机组数量的增加,系统内各主体的总收益呈现逐步上升趋势。具体而言,5台策略性报价火电机组情形下的总收益比2台策略性报价火电机组情形下的总收益提升了14.5%,这表明在电力市场交易中充分考虑市场主体间的策略互动行为,有助于促进整体经济效益的提升,进一步验证了本文所提多智能体强化学习框架在模拟实际市场博弈过程中的有效性。

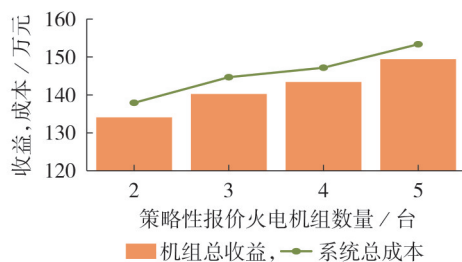


图7 策略性报价机组数量的影响分析
Fig.7 Influence analysis of number of strategic bidding units

3.5 调频需求占比的影响分析

为了分析调频需求占比对储能装置收益的影响,本文对不同的调频需求占比情形进行仿真,所得储能装置在电能量市场和调频市场的收益如图8所示。由图可知:随着调频需求占比的提升,储能在调频市场中的收益显著提高,且增长趋势不断加大;而随着调频需求占比在一定的范围内提升,储能在电能量市场中的购电成本缓慢增大,但达到一定占比后甚至会降低,这是因为随着调频需求增大,电能量市场中的主要参与者——火电机组会在调频市场中采取更为激进的报价策略,在电能量市场中降低报价,

从而使储能的购电成本降低。

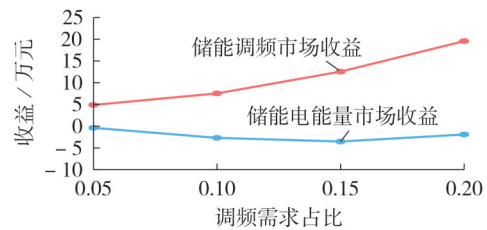


图8 调频需求占比对储能收益的影响
Fig.8 Influence of frequency regulation demand proportion on benefit of energy storage

4 结论和展望

本文针对独立储能在电能量-调频辅助服务市场中的联合交易策略进行研究,提出了一种基于多智能体强化学习的决策方法,综合考虑了各市场主体的策略性报价行为,采用MADDPG算法进行模型求解,并以IEEE 30节点系统为算例验证了所提方法的有效性。所得主要研究结论如下。

1)在算法性能方面,多智能体强化学习算法(MADDPG算法)在收敛速度与收敛效果方面均优于单一智能体算法(DDPG算法)。MADDPG算法大约经过200轮迭代后实现稳定收敛,智能体总收益较DDPG算法提升了18.8%,表明MADDPG算法在处理多主体博弈问题时具有更好的性能表现。

2)在储能参与市场的策略特性方面,受限于不能同时参与电能量与调频市场的规则约束,储能需要合理制定分时段报价与市场选择策略。仿真结果表明,储能倾向于在电能量市场中以较低成本购电,进而通过参与调频辅助服务市场获取更高的收益,体现了其在多市场协同中的策略优势。

3)在策略性主体数量对市场效率的影响方面,随着系统中策略性报价主体数量的增加,各市场参与主体的总收益呈现上升趋势。具体而言,5台策略性火电机组情形下的系统总收益较2台策略性火电机组情形下的系统总收益提升了14.5%,这说明在电力市场设计中充分考虑主体策略互动行为,有助于提升整体市场效率与运行灵活性。

本文为新型电力系统中独立储能参与多市场交易提供了有效的建模与求解思路,所构建的多智能体强化学习博弈框架能够较好地反映实际市场中的不完全信息与策略交互特征,对促进储能参与市场化运营、提升系统整体运行效率具有参考意义。

未来工作将聚焦新能源大量接入的储能电力市场交易策略研究,进一步考虑新能源出力的不确定性和系统负荷的不确定性,采用大语言模型技术辅助竞价智能体进行报价策略的生成。同时,本文没有考虑购电商的竞价策略对出清价格的影响和火电

机组的实际运行特性,后续工作可以考虑建立双边拍卖电力市场模型,研究购电侧的行为对发电侧报价策略的影响,并结合更多实际电力市场的运行数据与生产需求。

附录见本刊网络版(<http://www.epae.cn>)。

参考文献:

- [1] 国家能源局. 关于印发《加快构建新型电力系统行动方案(2024—2027年)》的通知[EB/OL]. (2024-07-25)[2025-10-27]. https://www.gov.cn/zhengce/zhengceku/202408/content_6966863.htm.
- [2] KHAN M K, RAZA M, SHAHBZA M, et al. Recent advancement in energy storage technologies and their applications[J]. *Journal of Energy Storage*, 2024, 92: 112112.
- [3] 王文坚, 陈飞雄, 林世琦. 新型电力系统下调频容量需求计算与储-荷调频方法研究综述[J]. *智慧电力*, 2025, 53(12): 1-12. WANG Wenjian, CHEN Feixiong, LIN Shiqi. A review of frequency regulation capacity demand calculation and energy storage-load frequency regulation methods in new-type power systems[J]. *Smart Power*, 2025, 53(12): 1-12.
- [4] FENG T, CONEJO A J. Electric-vehicle aggregators as cyber-physical energy systems: optimal operation strategy for frequency regulation and energy arbitrage[J]. *Cyber-Physical Energy Systems*, 2025, 1(1): 49-55.
- [5] RANCILIO G, BOVERA F, SPILLER M, et al. BESS and the ancillary services markets: a symbiosis yet? Impact of market design on performance[J]. *Applied Energy*, 2024, 375: 124153.
- [6] 韩泽雷, 鞠平, 秦川, 等. 面向新型电力系统的频率安全研究综述与展望[J]. *电力自动化设备*, 2023, 43(9): 112-124. HAN Zelei, JU Ping, QIN Chuan, et al. Review and prospect of research on frequency security of new power system[J]. *Electric Power Automation Equipment*, 2023, 43(9): 112-124.
- [7] 国家能源局. 国家能源局关于促进新型储能并网和调度运用的通知[EB/OL]. (2024-04-02)[2025-10-27]. http://www.scio.gov.cn/live/2024/34135/wjzc/202406/t20240613_851615.html.
- [8] 国家能源局. 关于深化新能源上网电价市场化改革促进新能源高质量发展的通知[EB/OL]. (2025-02-09)[2025-11-13]. https://www.ndrc.gov.cn/xxgk/zcfb/tz/202502/t20250209_1396066.html.
- [9] 李丹阳, 刘鸿鹏, 韩旭, 等. 考虑新能源市场容量价格的新型电力现货市场交易机制设计[J]. *电力自动化设备*, 2026, 46(2): 194-204. LI Danyang, LIU Hongpeng, HAN Xu, et al. Design of trading mechanism for new electricity spot market considering market capacity price of renewable energy[J]. *Electric Power Automation Equipment*, 2026, 46(2): 194-204.
- [10] 谢开, 刘敦楠, 李竹, 等. 适应新型电力系统的多维协同电力市场体系[J]. *电力系统自动化*, 2024, 48(4): 2-12. XIE Kai, LIU Dunnan, LI Zhu, et al. Multi-dimensional collaborative electricity market system for new power system[J]. *Automation of Electric Power Systems*, 2024, 48(4): 2-12.
- [11] 许敏兴, 苗世洪, 励刚, 等. 考虑多元储能参与现货市场的电力系统日前-日内优化调度[J]. *电力系统自动化*, 2025, 49(24): 51-62. XU Mxing, MIAO Shihong, LI Gang, et al. Day-ahead and intra-day optimal scheduling of power system considering multiple energy storage participating in spot market[J]. *Automation of Electric Power Systems*, 2025, 49(24): 51-62.
- [12] 徐爽, 万靖, 崔世常, 等. 不对称信息下新型储能容量租赁市场的博弈均衡分析[J]. *电力系统保护与控制*, 2024, 52(13): 13-24. XU Shuang, WAN Jing, CUI Shichang, et al. Game equilibrium analysis of energy storage leasing market with asymmetric information[J]. *Power System Protection and Control*, 2024, 52(13): 13-24.
- [13] 王再闯, 郭明阳, 刘斌, 等. 基于策略性双向拍卖机制的共享储能与产销者电力交易模型[J]. *电力自动化设备*, 2026, 46(3): 41-50. WANG Zaichuang, GUO Mingyang, LIU Bin, et al. Electric power trading model between shared energy storage and active energy participant based on strategic double auction mechanism[J]. *Electric Power Automation Equipment*, 2026, 46(3): 41-50.
- [14] 苏烨, 石剑涛, 张江丰, 等. 考虑调频的储能规划与竞价策略综述[J]. *电力自动化设备*, 2021, 41(9): 191-198. SU Ye, SHI Jiantao, ZHANG Jiangfeng, et al. Review on planning and bidding strategy of energy storage considering frequency regulation[J]. *Electric Power Automation Equipment*, 2021, 41(9): 191-198.
- [15] 林阿竹, 柯清辉, 江岳文. 独立储能参与调频辅助服务市场机制设计[J]. *电力自动化设备*, 2022, 42(12): 26-34. LIN Azhu, KE Qinghui, JIANG Yuewen. Market mechanism design of independent energy storage participating in frequency modulation auxiliary service market[J]. *Electric Power Automation Equipment*, 2022, 42(12): 26-34.
- [16] 陆秋瑜, 杨银国, 谢平平, 等. 适应储能参与的调频辅助服务市场机制设计及调度策略[J]. *电网技术*, 2023, 47(12): 4971-4984. LU Qiuyu, YANG Yinguo, XIE Pingping, et al. Market mechanism design and scheduling strategy of auxiliary services for frequency control adapting to energy storage participation[J]. *Power System Technology*, 2023, 47(12): 4971-4984.
- [17] 薛文浩, 吕泉, 齐楚, 等. 现货市场环境独立储能参与一次调频市场运营策略[J]. *电力自动化设备*, 2025, 45(7): 28-34. XUE Wenhao, LÜ Quan, QI Chu, et al. Operation strategy of independent energy storage participating in primary frequency regulation market under spot market environment[J]. *Electric Power Automation Equipment*, 2025, 45(7): 28-34.
- [18] 郝丽丽, 蔡雅琪, 王正风, 等. 计及运行规划收益补偿的电能市场储能报量报价策略[J]. *电力自动化设备*, 2024, 44(5): 36-43, 72. HAO Lili, CAI Yaqi, WANG Zhengfeng, et al. Quantity-price bidding strategy of energy storage in electric energy market considering operation and planning benefit compensation[J]. *Electric Power Automation Equipment*, 2024, 44(5): 36-43, 72.
- [19] NASROLAHPOUR E, KAZEMPOUR J, ZAREIPOUR H, et al. A bilevel model for participation of a storage system in energy and reserve markets[J]. *IEEE Transactions on Sustainable Energy*, 2018, 9(2): 582-598.
- [20] 张硕, 陈媛丽, 李英姿, 等. 计及电力现货机会成本的构网型储能电站调频辅助服务竞价出清双层博弈模型[J]. *中国电机工程学报*, 2024, 44(增刊1): 146-158. ZHANG Shuo, CHEN Yuanli, LI Yingzi, et al. A bi-level game model for frequency regulation ancillary service bidding and clearing of grid-forming energy storage stations considering opportunity cost of electricity spot trading[J]. *Proceedings of the CSEE*, 2024, 44(Supplement 1): 146-158.
- [21] 李国庆, 闫克非, 范高锋, 等. 储能参与现货电能-调频辅助服务市场的交易决策研究[J]. *电力系统保护与控制*, 2022, 50(17): 45-54. LI Guoqing, YAN Kefei, FAN Gaofeng, et al. Transaction deci-

- sion-making of energy storage stations participating in the spot energy and frequency modulation ancillary service market[J]. Power System Protection and Control, 2022, 50(17): 45-54.
- [22] 刘继春, 刘可欣, 柯贤波, 等. 新型电力系统下储能参与电能-惯量-一次调频多市场交易方法[J]. 电网技术, 2025, 49(3): 1018-1031.
- LIU Jichun, LIU Kexin, KE Xianbo, et al. Participation of energy storage in electric energy-inertia-primary frequency regulation multi-market trading method under a new power system [J]. Power System Technology, 2025, 49(3): 1018-1031.
- [23] 葛晓琳, 凡婉秋, 符杨, 等. 基于改进柔性策略评价的风火储多主体博弈电能-调频市场联合竞价模型[J]. 电网技术, 2023, 47(5): 1920-1930.
- GE Xiaolin, FAN Wanqiu, FU Yang, et al. Joint bidding model of electricity and frequency regulation market with wind fire storage multi-agent games based on improved soft actor-critic[J]. Power System Technology, 2023, 47(5): 1920-1930.
- [24] DU Y, LI F X, ZANDI H L, et al. Approximating Nash equilibrium in day-ahead electricity market bidding with multi-agent deep reinforcement learning[J]. Journal of Modern Power Systems and Clean Energy, 2021, 9(3): 534-544.
- [25] JIANG Y Z, DONG J, HUANG H X. Optimal bidding strategy for the price-maker virtual power plant in the day-ahead market based on multi-agent twin delayed deep deterministic policy gradient algorithm[J]. Energy, 2024, 306: 132388.
- [26] 张继行, 张一, 王旭, 等. 基于多代理强化学习的多新型市场主体虚拟电厂博弈竞价及效益分配策略[J]. 电网技术, 2024, 48(5): 1980-1991.
- ZHANG Jihang, ZHANG Yi, WANG Xu, et al. Game bidding and benefit allocation strategies for virtual power plants with multiple new market entities based on multi-agent reinforcement learning[J]. Power System Technology, 2024, 48(5): 1980-1991.
- [27] LI X Y, LUO F J, LI C J. Multi-agent deep reinforcement learning-based autonomous decision-making framework for community virtual power plants [J]. Applied Energy, 2024, 360: 122813.
- [28] 张兴平, 王腾, 张馨月, 等. 基于多智能体深度确定策略梯度算法的火力发电商竞价策略[J]. 中国电力, 2024, 57(11): 161-172.
- ZHANG Xingping, WANG Teng, ZHANG Xinyue, et al. Bidding strategy for thermal power generation companies based on multi-agent deep deterministic policy gradient algorithm[J]. Electric Power, 2024, 57(11): 161-172.

作者简介:

张新琪(2000—), 男, 工程师, 主要研究方向为新能源发电设备与系统(E-mail: zxq2495@163.com);

柳志海(2002—), 男, 硕士研究生, 主要研究方向为电力市场(E-mail: 22460137@zju.edu.cn);

杨强(1979—), 男, 教授, 博士, 研究方向为电力人工智能理论及应用(E-mail: qyang@zju.edu.cn)。

(编辑 陆丹)

Trading strategy of energy storage in electric energy and frequency regulation markets based on multi-agent reinforcement learning

ZHANG Xinqi¹, LIU Zhihai², CHEN Xiaochen³, ZHAN Shuang³, CHENG Yongzhuo⁴, HE Zhilong⁵, YANG Qiang²

(1. Datang Hainan Wenchang New Energy Co., Ltd., Wenchang 571332, China; 2. College of Electrical Engineering, Zhejiang University, Hangzhou 310027, China; 3. Datang(Wenchang) Energy Storage Technology Co., Ltd., Wenchang 571300, China; 4. China Datang Corporation Overseas Investment Co., Ltd., Beijing 100052, China; 5. East China Electric Power Test & Research Institute, China Datang Corporation Science and Technology General Research Institute Co., Ltd., Hefei 230061, China)

Abstract: To enhance the market competitiveness of independent energy storage in the new power system, a joint trading strategy for participating in electric energy and frequency regulation ancillary service markets based on multi-agent reinforcement learning is proposed. According to the rules of the Southern Regional Electricity Market, a joint clearing model including thermal power, new energy and independent energy storage is constructed. Under a Markov game framework, each power generation entity is modeled as an agent, and a multi-agent reinforcement learning algorithm is adopted to achieve strategic bidding optimization under conditions of incomplete information. To depict the cross-period coupling characteristics between energy storage and system operation, a modeling method of multi-dimensional state and action space is designed to achieve multi-period collaborative optimization. The simulative results of IEEE 30-bus system show that the proposed algorithm can increase the total revenue of the agents by 18.8% compared with the single-agent algorithm. Further analysis results indicate that as the number of units engaging in strategic bidding increases, the overall revenue of each market entity exhibits an upward trend. The research results verify the effectiveness of the proposed multi-agent reinforcement learning framework in enhancing the revenue of energy storage in the markets and optimizing the overall operational efficiency of the system.

Key words: independent energy storage; electric energy market; frequency regulation auxiliary service market; joint trading strategy; multi-agent; reinforcement learning