Adaptive Deep Reinforcement Learning Control for Distributed Shunt Compensators in Modern Electric Ship Power Systems

Sina Zarrabian[*], Chang Woo Hong[*], Van-Hai Bui[†], Adedoyin Inaolaji[‡]
[*]Department of Electrical and Computer Engineering, United States Naval Academy, Annapolis, MD, USA
[†]Department of Electrical and Computer Engineering, University of Michigan-Dearborn, Dearborn, MI, USA
[‡]Department of Electrical Engineering, University at Buffalo, Buffalo, NY, USA

***Abstract*— Shipboard power systems are increasingly transforming into complex microgrids with high penetration of power electronics, nonlinear loads, and mission-critical supply requirements. These characteristics amplify voltage fluctuations, undesirable harmonics, and instability particularly under contingency conditions and pulsed or rapidly varying loads. Distributed flexible AC transmission systems (D-FACTS) offer effective voltage support and power quality enhancement; however, their performance remains highly dependent on controller tuning, limiting adaptivity under dynamic operational conditions. This paper proposes a hybrid machine learning (ML)-based control framework combining deep learning and reinforcement learning to enhance distributed shunt compensators performance in shipboard power systems. The proposed framework is implemented to a realistic diesel-electric ship power system model with the integration of real-time, hardware-in-the-loop (HIL) simulations. Also, we compare the effectiveness of different shunt controllers with conventional methods. Finally, the results demonstrate a prominent improvement in voltage stability and overall performance across various operating scenarios.**

***Index Terms*—Deep reinforcement learning, D-STATCOM, D-SVC, machine learning, ship power systems.**

I. Introduction

Electrification of naval and commercial vessels is approaching the development of modern shipboard power systems (SBPSs) towards a dominance of inverter-controlled AC-DC microgrids, characterized by low mechanical inertia, limited short-circuit capacity and high nonlinear and pulsed loads, propulsion drives, and other vital systems [1]. These characteristics induce, compared to conventional terrestrial grids, higher requirements for control of the power supply in terms of the stability, dynamics, power quality, and the reliability of the system [2].

In recent years, power electronics-based shunt controllers have shifted the applications of flexible AC transmission systems (FACTS) devices from transmission to distribution, such as distribution-Static Var Compensators (D-SVCs) and Static Synchronous Compensators (D-STATCOMs) [3]. Dynamic reactive support, voltage regulation, and harmonic mitigation are among characteristics that are provided by these devices. For shipboard applications where mass, space, and temperature constraints are crucial, the quick controllability, modularity, and compactness of distribution shunt compensators make them appealing [4].

Ship power systems come with special challenges. In terms of loading, sudden pulsed power loads from propulsion motors to pulsed radars can occur frequently. In the context of equipment arrangement, space onboard is limited, so equipment must stay compact yet lightweight. From a topology point of view, some ships use hybrid AC and DC zones where power flow is bidirectional. From an operational perspective, there are limited short-circuit capability, severe voltage dips, problems like flicker, harmonics, and unbalanced loads. Shunt FACTS devices suit ships well since they react quickly to sudden load variations and contingency conditions such as faults [5]-[6]. Additionally, they provide fast reactive power support to stabilize voltage, improve voltage fluctuations or flicker by injecting current, and mitigate harmonics caused by inverter-based resources (IBRs). In [7], D-STATCOM was used to control voltage in an electric ship power system. An adaptive technique based on the artificial immune system was also used to adjust the STATCOM control parameters. However, a trivial electric ship model was used that lacked today's integration of modern hybrid AC/DC ship energy systems.

Aside from operational characteristics of converter-based shunt compensators, their control has a direct and significant impact on their robustness and effectiveness. Although they are effective in many applications, traditional control strategies (e.g., conventional PID, current control, droop, and virtual impedance schemes, vector controllers) depend on system linearization, tuned controller gains, and physics-based models that might not apply to the broad operating range of a modern ship microgrid [8].

Machine learning approaches for control of shunt compensators (especially model-free and data-driven methods such as deep learning, reinforcement learning, and hybrid

[*]DISCLAIMER: The views expressed in this article are those of the author(s) and do not reflect the official policy or position of the U.S. Naval Academy, Department of the Navy, the Department of War, or the U.S. Government.

979-8-3315-8138-1

learning-based supervisory schemes) offer several promising advantages on ship power systems compared to their conventional counterparts. For instance, ML-based controllers facilitate adaptive performance through varying operating modes. They learn control policies without the need for manual retuning controller gains. Moreover, they permit multi-objective optimization, to enhance voltage stability, power quality, and efficiency metrics simultaneously [9]. Besides that, ML makes it possible for the supervisory control to coordinate the most advanced FACTS devices along with the energy storage units.

Recent reviews unveil a very limited comprehensive study of distributed shunt controllers in modern all-electric ship power systems and their impacts in improving voltage quality, reactive power compensation, and stability. This paper tackles the existing gaps by presenting a focused investigation of D-SVC and D-STATCOM performance using a weakly coupled deep Q-network (WCDQN) control algorithm [10]. The study is conducted on a real-world diesel-electric shipboard power system model, incorporating HIL simulations and various operating scenarios. Specifically, the contributions are:

- A dynamic real-world model of shipboard test case with HIL simulations that capture realistic behavior with AC/DC interfaces and pulsed heavy loads.
- A side-by-side performance comparison of D-SVC, and D-STATCOM under various operating scenarios.
- Design and evaluation of an integrated ML-based control architecture that pairs shunt D-FACTS devices with the rest of the system to improve resilience and reduce required device ratings.

II. Modeling of the Shipboard Power System and Shunt Compensators

In this section, the shipboard power system under study is detailed out. Then, the model and control structure of D-SVC and D-STATCOM are presented.

A. *Model of the Ship Power System Under Study*

The ship power system analyzed in this paper is based on the 14 MW national security multi mission vessel (NSMV), which serves as a realistic platform for testing the proposed algorithm [11]. Fig. 1 illustrates the diesel-electric shipboard power system used in the work. For clarity, only one of the two identical engine rooms is shown. The full system includes four diesel-electric engines, each rated at 4.2 MW, providing a combined output of approximately 14.28 MW. Within each engine room, two diesel generators (DGs) supply their respective 6.6 kV busbars. The ship uses two electric motors for propulsion, each rated at 4.5 MW, connected via a single shaft with an efficiency of 94% and operating at a 95% load factor.

The propulsion motors in each engine room are supplied through a variable-speed power converter linked to a three-winding transformer and the bow thruster motor is powered by an autotransformer. The vessel also includes an emergency diesel generator rated at 900 kW. Under normal conditions, the engine rooms serve as the primary source of power for the ship's low-voltage (LV) distribution network.

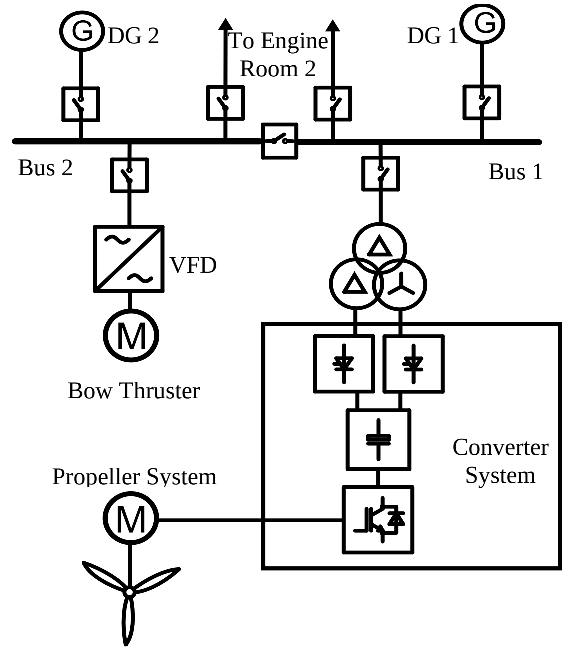

Figure 1. Ship power system under study

B. *Model and Control of D-SVC*

The D-SVC is categorized under the distributed flexible AC transmission system (D-FACTS) devices. It regulates the V-Q profile of the system by adjusting the reactive power to regulate its bus voltage. In this paper, we consider the thyristor switched capacitor-thyristor controlled reactor (TSC-TCR)–type distribution SVC (D-SVC) as shown in Fig. 2. It provides smoother reactive power control, has lower reactance value, and better mitigates harmonics compared to fixed capacitor-type SVC. Control block diagram of the D-SVC is shown in Fig. 2 as well. This block diagram simply shows how D-SVC adjusts its bus voltage. More detail on implementation of the proposed control algorithm is discussed in the next section.

TCR and TSC are connected to the secondary side of a step-down transformer to which shunt filters are connected. This regulates required reactive power by controlling capacitor and inductive banks through thyristor switches. D-SVC dynamic model is represented in (1) as follows [12]:

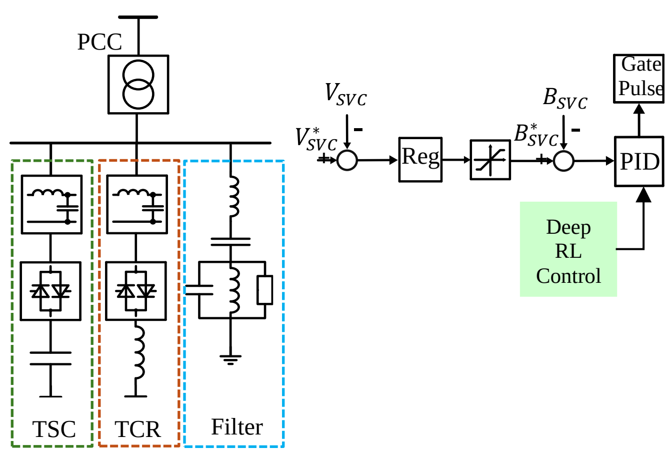

Figure 2. D-SVC schematic and control block diagram

979-8-3315-8138-1

$$\begin{bmatrix}\dot{X}_c\\ \dot{\alpha}\end{bmatrix} = func(x_c,\alpha,V_{SVC},V_{ref})$$

$$\begin{cases} B_{SVC}V_i = I_{SVC} \\ B_{SVC}V_i^2 = Q_{SVC} \\ B_{SVC} = \frac{1}{X_C X_L}\left[X_L - \frac{X_C}{\pi}(\pi - 2\alpha_{SVC} - \sin(2\alpha_{SVC}))\right] \end{cases} \quad (1)$$

where X_c is the control vector, $func$ represents mathematical equations relating the parameters, α is the firing angle of the SCRs, V_{SVC} (V_i) is the voltage of the SVC bus, V_{ref} is the bus reference voltage, and B_{svc} is the susceptance of the SVC.

C. *Model and Control of D-STATCOM*

The D-STATCOM is a shunt var compensator that uses voltage source converters (VSCs) as an interface with the point of connection instead of variable-impedance switched capacitors or inductors. It automatically regulates the voltage by exchanging reactive power with the system, either injecting or absorbing. D-STATCOM can operate in different control modes, such as voltage control, reactive power control, and power factor control. Additionally, it can work in voltage/current mode controls, controlling injected voltage or current, respectively. Figs. 3 and 4 show schematic and control block diagram of D-STATCOM, respectively. D-STATCOM dynamic model can be represented as follows [12]:

$$\frac{d}{dt}\begin{bmatrix} i_{dq(p)} \\ i_{dq(n)} \\ V_{dc}\end{bmatrix} = \begin{bmatrix} \cdots & \cdots & \cdots \\ \vdots & func(R_g, L_g, \omega^*, C) & \vdots \\ \cdots & \cdots & \cdots \end{bmatrix}\begin{bmatrix} i_{dq(p)} \\ i_{dq(n)} \\ V_{dc}\end{bmatrix} + B_{dq}\begin{bmatrix} i_{dq(p)} \\ i_{dq(n)} \\ V_{dc}\end{bmatrix} + D\begin{bmatrix} v_{dq(p)} \\ v_{dq(n)} \\ 0\end{bmatrix} \quad (2)$$

$$B_{dq} = \begin{bmatrix} \cdots & \cdots & \cdots \\ \vdots & func(L_g, \omega^*, C, S_{dq(p)}, S_{dq(n)}, k_m) & \vdots \\ \cdots & \cdots & \cdots \end{bmatrix} \quad (3)$$

where $i_{dq(p)}$, $i_{dq(n)}$, $v_{dq(p)}$, $v_{dq(n)}$, $S_{dq(p)}$ and $S_{dq(n)}$ are positive and negative sequence of the dq-frame current, voltage, and switching functions, respectively. R_g and L_g are the network-side resistance and inductance, respectively. V_{dc} is the dc link voltage, C is the dc link capacitor, ω^* is the angular velocity at the network rated frequency, $D = \omega^*/L_g$, and k_m is the constant coefficient linking AC and DC side voltages.

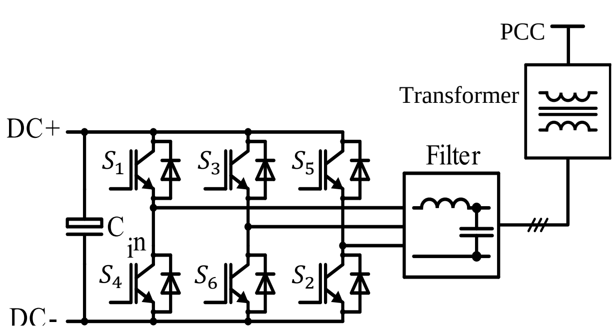

Figure 3. D-STATCOM schematic

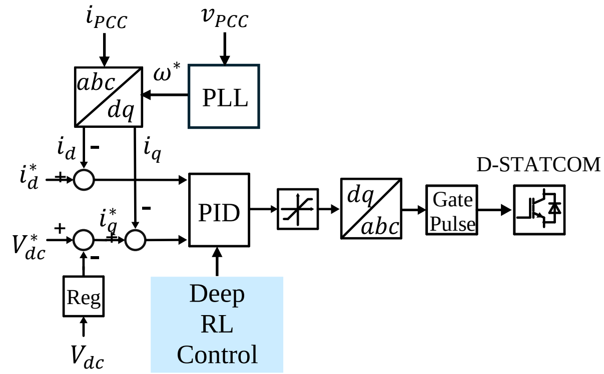

Figure 4. D-STATCOM control block diagram

III. PROPOSED DEEP REINFIRCEMENT LEARNING -BASED ADAPTIVE CONTROL

In this section, we explore how the proposed WCDQN algorithm is applied for adaptive tuning of D-SVC and D-STATCOM PID control parameters.

A. *Introduction to WCDQN Algorithm*

The WCDQN approach [10] is derived from the traditional DQN framework but incorporates Lagrangian relaxation to establish upper and lower bounds. Similar to deep Q-network, it features a primary Q-network, denoted as $Q'(s, a;\, \theta)$, which is trained to approximate the action–value function for the overall decision-making problem. In addition, the method employs several subagent networks $Q_i^\lambda(s_s, a_i;\, \theta_U)$, each responsible for learning the action–value function associated with its corresponding subproblem. Within this formulation, S represents the set of possible states, $\mathcal{A}$ is the discrete action space, and θ denotes the neural network parameters. State transitions occur based on a probability function p, and each state–action pair generates a reward $r(s, a)$. The action-value function in the WCDQN is written as:

$$Q_{i,j+1}^\lambda(s_i, a_i) = Q_{i,j}^\lambda(s_i, a_i) + \rho_n(s_i, a_i)\left[[y_{i,j}^\lambda - Q_{i,j}^\lambda(s_i, a_i)\right] \quad (4)$$

where ρ_n is learning rate and $\sigma_{i,j}^\lambda$ is the target value function defined as:

$$\sigma_{i,j}^\lambda = r_i(s_i, a_i) - \lambda^T \boldsymbol{d}(s_i, a_i) + \gamma\, max_{a_i'}\, Q_{i,j}^\lambda(s_i', a_i') \quad (5)$$

In (5), λ represents the penalty vector applied when a chosen action fails to enhance the system's state, and $r(s, a)$ denotes the reward assigned to taking action a in state s according to the specified reward function. The WCDQN method achieves better performance than standard DQN by refining the loss function so that the error in estimating the Q-values is more accurately captured and minimized. The revised loss function is defined as follows [10]-[11]:

$$\ell(\theta) = \mathbb{E}_{s,a\sim\rho,\lambda\sim\mu}\left[\left(\sigma - Q'(\boldsymbol{s},\boldsymbol{a};\, \theta)\right)^2 + \kappa_U(\, Q'(\boldsymbol{s},\boldsymbol{a};\, \theta) - \sigma_U)_+{}^2\right] \quad (6)$$

where $(x)_+ = \max(x, 0)$, and κ_U is a soft penalty coefficient, and:

979-8-3315-8138-1

$$\sigma = r(s, a) - \gamma\, \mathbb{E}\left[max_{a' \in \mathcal{A}(s')}\, Q'(s', a'; \theta^{-})\right] \quad (7)$$

$$\sigma_U = r(s, a) - \gamma\, \mathbb{E}\left[max_{a' \in \mathcal{A}} Q^{\lambda^*}(s', a'; \theta_U^{-})\right] \quad (8)$$

The WCDQN training procedure largely resembles that of the standard DQN framework but includes several key improvements. At every training step, an action is chosen using an ϵ-greedy strategy from the set of available actions, and the resulting transition is saved in the replay buffer $\boldsymbol{B}(\eth)$ as a tuple $\tau(s, a, r, b, s')$ for later use. The network's value estimates are then updated accordingly. Afterward, each network performs a stochastic gradient descent update using its corresponding loss function, where estimates of the available actions are obtained by sampling mini-batches that contain tuples τ and λ. The penalty factor κ_U may be set to a constant positive value or dynamically adjusted throughout training.

B. WCDQN Control Strategy for Shunt Compensators

The proposed WCDQN is applied to tune D-SVC and D-STATCOM control parameters as follows [13]-[14]:

1) D-SVC:

a) The state space of the system is mapped to $s \in \{V_{SVC(i)}, V_{SVC(i)}^{ref}, B_{SVC}\}$.

b) The action space is defined as $a \in A\{V_{SVC(min)}, V_{SVC(max)}, K_p, K_i, K_d\}$.

c) The reward function is defined as:

$$r(s_i, a_i) = e^{-\left(\frac{\left|\left(\mathrm{V_{SVC(i)}} - \mathrm{V_{SVC(i)}^{ref}}\right)\right|^2}{2\tau^2}\right)} \quad (9)$$

Where τ is a control parameter to guarantee the accuracy.

2) D-STATCOM:

a) The state space of the system is mapped to $s \in \{V_i, V_{i(ref)}, I_i, I_{i(ref)}, \omega_i, V_{dc}, V_{dc(ref)}\}$.

b) The action space is defined as $a \in A\{I_d, I_q, K_p, K_i, K_d\}$.

c) The reward function is defined as:

$$r(s_i, a_i) = e^{-(\varepsilon V_{i,\Delta}^2)} - \left(\xi \left(\frac{dV_i}{dt}\right)^2 + \vartheta T_{s_i}^2\right) \quad (10)$$

where $\frac{dV_i}{dt}$ is measured rate of change of bus voltage, $V_{i,\Delta}$ is the deviation of measured bus voltage from the reference voltage, T_s is the settling time of the system voltage after disturbance, and $\xi, \vartheta, \varepsilon$ are the penalty multipliers.

IV. HIL Setup and Simulation Results

A. HIL Experimental Setup

Fig. 5 shows the HIL setup for this study. The shipboard power system was modeled in real-time using the MATLAB/SIMULINK library. The D-SVC and D-STATCOM hardware developed using LabVolt power modules, including thyristors (SCRs), capacitor banks, inductor banks, smoothing filters, and three-phase IGBT-based voltage source inverter (VSI). The hardware interface was Real-Time Target Machine digital simulator and GPIO unit. The proposed control algorithm is implemented on a real-time platform using MATLAB function running with a 50 µs time step. Gate signals for the VSI module are produced by the Target Machine at a 10 kHz switching frequency. The simulator's GPIO outputs are used to interface with the SCRs and VSI gate inputs.

B. Operating Scenarios and Simulation Results

1) Scenario 1: Sudden Critical Load Change

In this scenario, a sudden critical load change occurs via loss of the bow thruster motor in the engine room #1 at $t = 1\, sec$. The response of the shunt compensators and the proposed control algorithm are examined in this scenario. Voltage and frequency of the system at bus 1 is depicted in Fig. 6. It is observed that the conventional control does not keep up with the changes leading to severe frequency and voltage instability. The D-STATCOM successfully damped voltage and frequency oscillations under 4 secs, with a fast and smooth response while the D-SVC struggled to damp oscillations in a timely manner.

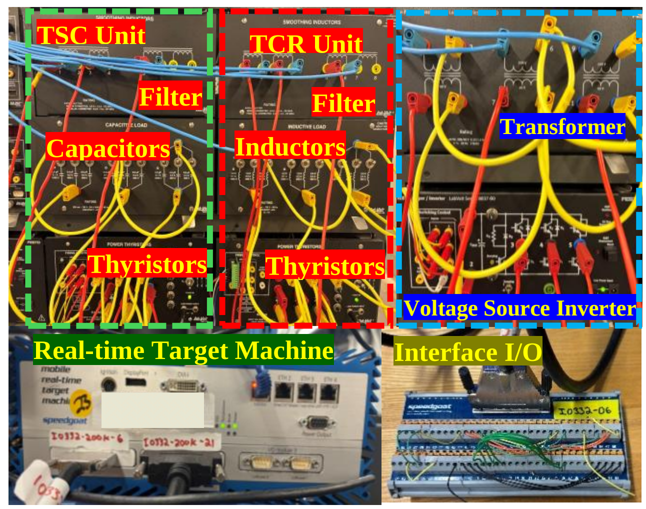

Figure 5. HIL experimental setup

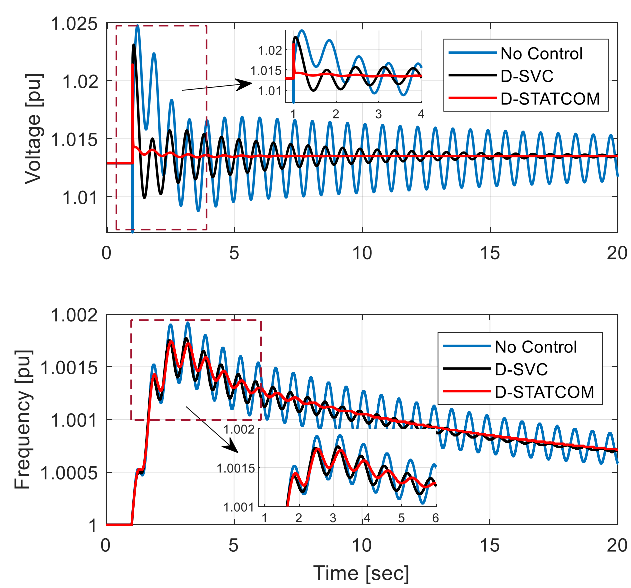

Figure 6. System response to sudden loss of the bow thruster motor: voltage and frequency of Bus 1

979-8-3315-8138-1

1) Scenario 2: Pulsed Power Load

In this scenario a pulsed power load (PPL) is simulated and applied to the system and results are shown in Fig. 7. The D-STATCOM outperforms the D-SVC by providing the least deviation in voltage and frequency profiles at bus 1 during the PPL and can recover the voltage and frequency faster and more effectively. The D-SVC shows higher magnitude deviations in voltage and frequency with a longer damping characteristic. Nevertheless, both shunt compensators bring significant improvement compared to the system's default control.

2) Scenario 3: Balanced Three-Phase Fault

This scenario examines the system performance under a balanced three-phase fault at bus #2. The fault happens at $t = 1\ sec$ and is cleared after 50 ms. Results are presented in Fig. 8. The D-STATCOM offers a better voltage and frequency control within and after the fault, via less voltage dip, faster recovery time at about 3 sec, and significant damping characteristics compared to the D-SVC.

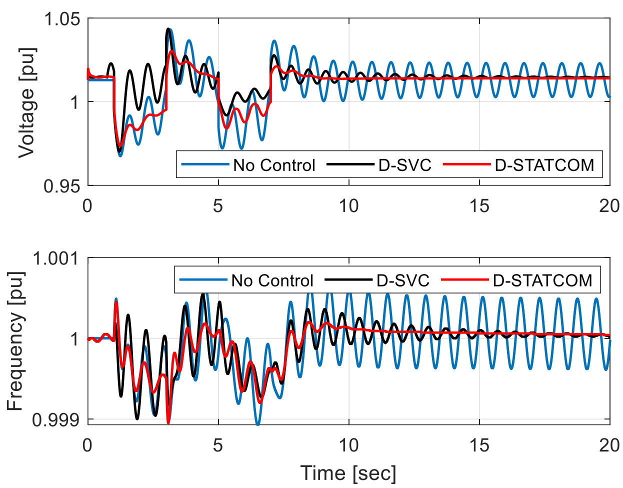

Figure 7. System response to PPL: voltage and frequency of Bus 1

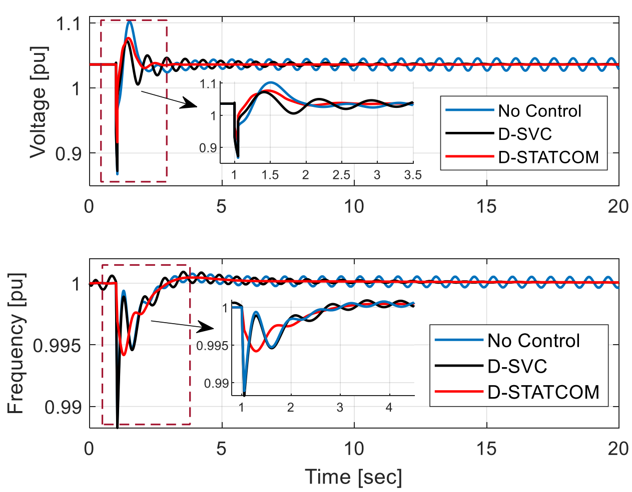

Figure 8. System response to a balanced three-phase fault at bus #2: voltage and frequency of Bus 1

V. CONCLUSION

In summary, this work demonstrates that integration of deep reinforcement learning into the control of distributed shunt compensators can significantly strengthen the resilience and performance of modern shipboard power systems. By applying the proposed WCDQN framework to a realistic diesel-electric vessel and validating it through real-time HIL simulations, we show that intelligent, adaptive control offers clear advantages over conventional methods, especially under rapidly changing or stressed operating conditions. The comparative results highlight notable gains in voltage stability, improved handling of nonlinear and pulsed loads, and a more robust response to system contingencies.

REFERENCES

[1] C. -L. Su, M. -C. Lin and C. -H. Liao, "A Method for Evaluating Energy Efficiency to Justify Power Factor Correction in Ship Power Systems," in IEEE Transactions on Industry Applications, vol. 49, no. 6, pp. 2773-2782, Nov.-Dec. 2013.

[2] A. Vicenzutti, G. Sulligoi, V. Bucci, S. Bertagna, M. Cataneo and P. Borghese, "Naval Smart Grid Preliminary Integration onboard Electric Ships," 2021 IEEE Electric Ship Technologies Symposium (ESTS), Arlington, VA, USA, 2021, pp. 1-7.

[3] J. Andrés Guzmán-Henao, R. Iván Bolaños, L. Fernando Grisales-Noreña, O. D. Montoya and J. C. Hernández, "Hierarchical Energy Planning and Control of DGs, BESS, and D-STATCOMs in Unbalanced Non-Interconnected Distribution Networks," in IEEE Access, vol. 13, pp. 165456-165480, 2025.

[4] A. H. Elmetwaly, A. A. Eldesouky and A. A. Sallam, "An Adaptive D-FACTS for Power Quality Enhancement in an Isolated Microgrid," in IEEE Access, vol. 8, pp. 57923-57942, 2020.

[5] A. Hamidi, S. Golshannavaz and D. Nazarpour, "D-FACTS Cooperation in Renewable Integrated Microgrids: A Linear Multiobjective Approach," in IEEE Transactions on Sustainable Energy, vol. 10, no. 1, pp. 355-363, Jan. 2019.

[6] M. Babaei, R. Jafari-Marandi, S. Abdelwahed and B. Smith, "Application of STATCOM for MVDC shipboard power system," 2017 IEEE Electric Ship Technologies Symposium (ESTS), Arlington, VA, USA, 2017, pp. 142-147.

[7] P. Mitra and G. K. Venayagamoorthy, "An Adaptive Control Strategy for DSTATCOM Applications in an Electric Ship Power System," in IEEE Transactions on Power Electronics, vol. 25, no. 1, pp. 95-104, Jan. 2010.

[8] M. Mahdavi, N. F. Alshammari, A. Awaafo, F. Jurado and P. Gopi, "An Economic Loss Reduction Using Static VAR Compensator and Capacitor Placement in Reconfigurable and Expandable Distribution Grids With a Variable Electric Power Demand," in IEEE Transactions on Industry Applications, vol. 61, no. 4, pp. 5585-5596, July-Aug. 2025.

[9] N. P. Reddy, R. Skjetne, O. S. Os and D. Papageorgiou, "A Comparison of the State-of-the-Art Reinforcement Learning Algorithms for Health-Aware Energy and Emissions Management in Zero-Emission Ships," in IEEE Journal of Emerging and Selected Topics in Industrial Electronics, vol. 5, no. 1, pp. 149-166, Jan. 2024.

[10] El Shar, I. and Jiang, D., "Weakly coupled deep q-networks," *Advances in Neural Information Processing Systems*, vol. 36, 2024.

[11] S. Zarrabian, V. -H. Bui and F. Feng, "Adaptive Deep Reinforcement Learning for Virtual Synchronous Generator Control in Ship Power Systems," 2025 IEEE Electric Ship Technologies Symposium (ESTS), Alexandria, VA, USA, 2025, pp. 143-148.

[12] M. Eremia; C.-C. Liu; A.-A. Edris, "Advanced Solutions in Power Systems: HVDC, FACTS, and Artificial Intelligence," IEEE, 2016.

[13] X. Zhang, Y. Liu, J. Duan, G. Qiu, T. Liu and J. Liu, "DDPG-Based Multi-Agent Framework for SVC Tuning in Urban Power Grid With Renewable Energy Resources," in IEEE Transactions on Power Systems, vol. 36, no. 6, pp. 5465-5475, Nov. 2021.

[14] J. H. Woo, L. Wu, S. M. Lee, J. -B. Park and J. H. Roh, "D-STATCOM d-q Axis Current Reference Control Applying DDPG Algorithm in the Distribution System," in IEEE Access, vol. 9, pp. 145840-145851, 2021.

979-8-3315-8138-1