Anusandhanvallari
Vol 2024, No.1
February 2024
ISSN 2229-3388

Deep Reinforcement Learning-Controlled GaN Dual-Active-Bridge Converter for Bidirectional V2GOn-BoardChargingin800VBEVPlatforms

Dr. Sandeep V. Gaikwad[1], Dr Vidya N. More[2], Dr. Rahul Pol[3], Kunal Magar[4]

[1]Department of Electronics & Telecommunication. SCTR's Pune Institute of Computer Technology (PICT), Pune

[2]Department of Electronics and Telecommunication,

COEP Technological University Pune, India

[3]Department of Electronics & Telecommunication. VIT (Kondhwa Campus)

Pune, India

[4]Department of Electronics & Telecommunication. ,

VIT (Kondhwa Campus)

Pune, India

Abstract—This paper presents the design, modulation, closed-loop control, and experimental validation of a 22kW, 500kHz Gallium Nitride(GaN) high-electron mobility transistor(HEMT)-based Dual-Active-Bridge (DAB) isolated DC-DC converter for full bidirectional Vehicle-to-Grid (V2G) and Grid-to-Vehicle (G2V) on-board charging (OBC) in 800V battery electric vehicle (BEV) platforms. A novel Extended-Phase-Shift (EPS) modulation strategy is augmented with a Twin Delayed Deep Deterministic Policy Gradient (TD3) deep reinforcement learning (DRL) controller that simultaneously optimizes efficiency, zero-voltage-switching (ZVS) range, RMS current stress, and grid total harmonic distortion (THD) across the complete bidirectional power envelope without hand-tuned PI regulators. Operating at a record 500kHz switching frequency with GaN Systems GS66516B devices (650V/60A) and a Vitroperm 500F nanocrystalline-core high frequency transformer, the converter achieves a power density of 7.6kW/L in a 210×140×28mm form factor. The TD3 agent is trained off line in a PLECS/Python co-simulation environment and deployed on an AMD Xilinx RFSoC ZU28DR SoC with 1.4µs FPGA-accelerated neural network inference latency. Experimental results confirm a peak round-trip G2V-to-V2G efficiency of 97.9%, grid-side current THD of 1.1% under V2G reactive power injection, ZVS maintained from 5% to 100% of rated power in both power directions, and 8ms dynamic power-set point settling time, with full compliance to ISO15118-20 Plug-and-Charge (PnC) V2G communication protocols. The proposed architecture deliversa 38% power density improvement and a 0.9% peak-efficiency gain over the best published SiC-based counterpart at equivalent power levels.

***Index Terms*—**Dual-active-bridge converter, extended phase shift, GaN HEMT, deep reinforcement learning, TD3 algorithm, vehicle-to-grid,bidirectional on-board charger, 800V BEV, zero-voltage switching, ISO15118-20, nano crystalline transformer,500 kHz power electronics.

I. INTRODUCTION

The electrification of personal and commercial transport is reshaping the relationship between vehicles, electricity grids, and energy markets at an unprecedented pace. Global BEV sales exceeded 17 million units in 2024, with projections indicating more than 400 million light-duty BEVs on-road by 2035 [1].

At an average usable battery capacity of 75 kWh, the 2035 BEV fleet represents over 30 TWh of distributed electrochemical storage—a magnitude eclipsing all grid-connected stationary storage installed worldwide today.

This scale of distributed storage, if left unmanaged, imposes severe stress on distribution networks through uncoordinated charging demand peaks; if managed intelligently through V2G, it represents the single largest flexible demand-response resource available to grid operators.

Vehicle-to-Grid (V2G) technology enables bidirectional power exchange between parked BEV batteries and the electrical distribution network, transforming this storage into an active grid asset. IEEE 2030.5 and ISO 15118-20 Plug-and-Charge protocols govern V2G energy transactions, enabling BEVs to provide frequency regulation, peak shaving, reactive power compensation, and grid resilience services [2]. Recent regulatory developments—notably FERC Order 2222 in the United States and the EU Smart Charging Directive—mandate aggregator access to V2G-capable assets in wholesale electricity markets, creating a commercially viable framework for V2G revenue generation at the vehicle level. However, V2G at scale demands bidirectional OBCs satisfying mutually antagonistic objectives: high power density, high efficiency across the full bidirectional power envelope, grid-side unity power factor with sub-2% THD, and rapid dynamic response for ancillary grid services—all within automotive cost and functional safety constraints.

The evolution of OBC power electronics has progressed through three distinct semiconductor generations. First-generation silicon IGBT-based OBCs, limited to switching frequencies below 20 kHz, achieve power densities of 1–2 kW/L with efficiencies in the 90–93% range. Second-generation Silicon Carbide (SiC) MOSFET-based designs push switching frequencies to 100–300 kHz, achieving 4–6 kW/L at 95–97% efficiency [6]. Third-generation GaN HEMT-based converters—the technology investigated in this work—operate at 300 kHz to over 1 MHz, enable power densities exceeding 7 kW/L, and approach 98% peak efficiency by exploiting near-zero switching losses and the elimination of reverse-recovery charge [4].

The Dual-Active-Bridge (DAB) isolated DC-DC converter, originally proposed by De Doncker et al. [3], has emerged as the canonical topology for bidirectional isolated power conversion owing to its inherent soft-switching capability, galvanic isolation, flexible voltage conversion ratio, and bidirectional symmetry. However, conventional Single-Phase-Shift (SPS) modulation of the DAB suffers from limited ZVS range at partial loads, high RMS current stress, and reactive power circulation—deficiencies that critically constrain efficiency below 30% of rated power, the operating regime where BEVs spend the majority of residential charging session time. Advanced modulation strategies

Extended-Phase-Shift (EPS) and Triple-Phase-Shift (TPS) along with Dual Active Phase Shift (DPS) progressively address these limitations by introducing additional control degrees of freedom, but at the cost of a nonlinear, multi-dimensional optimization surface that is intractable for classical PI-based regulators [10].

Deep Reinforcement Learning (DRL), and specifically the Twin Delayed Deep Deterministic Policy Gradient (TD3) algorithm [5], offers an adaptive control paradigm ideally suited to the multi-objective, nonlinear, and operating-point-varying nature of DAB control. Unlike Model Predictive Control (MPC), which requires an accurate real-time plant model and imposes computational burdens scaling exponentially with prediction horizon, a trained TD3 agent encodes optimal phase-shift trajectories across the full operating space within its network weights, enabling sub-microsecond policy inference after offline training. The deployment of such agents on FPGA-augmented automotive SoCs—where neural network inference can be accelerated to latencies below 2 μs—removes the computational barrier that has historically precluded DRL-based control in high-frequency power electronics.

The original contributions of this work are: (i) the first 500 kHz, 22 kW GaN DAB OBC operating under EPS modulation reported in the open literature; (ii) a TD3-DRL trained in PLECS/Python co-simulation and deployed on RFSoC FPGA fabric achieving 1.4 µs inference latency; (iii) extension of the DAB ZVS operating region to 5% of rated power in both power directions, eliminating the need for passive snubbers; (iv) complete experimental validation including ISO 15118-20 PnC V2G session management and simultaneous active and reactive power injection; and (v) comprehensive state-of-the-art benchmarking against six published bidirectional OBC designs across efficiency, THD, ZVS range, dynamic response, and power density metrics.

II. Dab Topology And Eps Modulation

A. Circuit Topology and Operating Modes

The DAB converter comprises two full H-bridges: a primary bridge (H1) interfacing the 400 V grid-side DC bus, and a secondary bridge (H2) interfacing the 650–1000 V battery-side bus, coupled through a high-frequency isolation transformer with turns ratio n:1n : 1n:1 and total series inductance LrL_rLr. Under G2V mode, H1 acts as the active inverter and H2 as the active rectifier; under V2G, roles reverse without hardware reconfiguration.

B. Single-Phase-Shift Analysis

Under SPS, all four legs of each bridge switch with 50% duty cycle and the sole control variable is the inter-bridge phase-shift angle $\phi \in (-\pi, \pi)$ The average transferred power is:

$$P = \frac{nV1V2}{\omega Lr}\phi\left(1 - \frac{|\phi|}{\pi}\right) - - - eq.1$$

where V_1 and V_2 are the primary and secondary DC voltages, fsf_sfs is the switching frequency, and Lr is the total series inductance. SPS imposes a reactive current component circulating even at zero net power transfer, generating I^2R losses and limiting ZVS to the voltage-conversion-ratio range $0.38 \leq M \leq 2.62$ where

$$M = \frac{nV_2}{V_1} - - - eq.2$$

a critical limitation for the 800 V BEV whose battery voltage spans 650–1000 V ($M \in \{1.46, 2.25\}$)

C. Extended-Phase-Shift Modulation Framework

EPS introduces an inner phase-shift ratio $D1 \in [0,1]$ on the primary bridge, allowing the primary voltage waveform to assume a quasi-trapezoidal shape in addition to the outer inter-bridge phase-shift ϕ The generalized EPS power transfer expression is:

$$P_{EPS} = \frac{nV_1V_2}{\omega Lr} f(D_1, \phi, M) - - - eq.3$$

where $f(D_1, \phi, M)$ is a piecewise nonlinear function of the operating mode determined by the relative magnitudes of D_1 and $|\phi|/\pi$. The ZVS boundary for the primary-side switches requires the inductor current at each switching instant to satisfy $i_L(t_{sw}) < 0$.

The EPS ZVS boundary locus in the (D_1, ϕ) control plane is:

$$i(t) = \frac{V_1}{2\omega Lr}\left[(2D_1 - 1) - \frac{M(2\phi M + 1)}{\pi}\right] \leq 0 - - - eq.4$$

By adaptively selectingD_1, EPS extends the ZVS operating region to as low as 5% of rated power in both power directions a critical enabler for high efficiency partial-load operation constituting over 60% of real-world residential charging sessions.

III. GaN HEMT AND 500 Khz POWER STAGE DESIGN

A. GaN Device Selection and Characterization

The GaN Systems GS66516B enhancement-mode GaN HEMT is selected based on its superior figure-of-merit:

$$FOM = R_{DS}(on) \times Q_{oss} = 25\,m\Omega \times 39\,nC = 975\,m\Omega \cdot nC$$

at 400 V, representing a 6.8× improvement over the best competing 650 V SiC MOSFET at equivalent voltage class [7]. The absence of a p-n body diode—replaced by GaN bidirectional channel conduction eliminates reverse recovery energy Err that would otherwise dominate switching losses at 500 kHz in silicon or SiC alternatives. Gate threshold voltage $Vth = 1.7\,V$ demands a dedicated DRIVER1ED44 gate driver with +6 V/−3 V drive levels and 100 ns dead-time accuracy enforced via FPGA-based dead-time control, preventing cross-conduction at sub-nanosecond edge rates. Double-pulse test (DPT) characterization at 400 V and 25∘C25^\circ C25∘C measures turn-on energy $E_{on} = 2.1\mu J$, and turn-off energy $E_{off} = 1.4\,\mu J$, per event at ID=30 AI_D = 30 \, AID=30A, resulting in total switching loss:

$$P_{sw} = 3.5\,\mu J \times 500\,kHz \times 12 = 21\,W - - - eq.6$$

across all twelve switching events per cycle in the full-bridge configuration.

B. Nanocrystalline High-Frequency Transformer Design

The isolation transformer is the magnetically most critical component at 500 kHz: conventional silicon-steel and MnZn ferrite cores exhibit prohibitive eddy-current core losses above 300 kHz. Hitachi Metals' 500F nanocrystalline alloy—an iron-based amorphous ribbon with 20 nm grain size—achieves core loss density $P_v = 280\ \frac{mW}{cm^3}$, at 500 kHz and 120 mT peak flux, compared to 2800 mW/cm32800 \, mW/cm^32800mW/cm3 for N87 MnZn ferrite under identical conditions [13].

The transformer employs two E64/27/20 nanocrystalline cores in a quad-interleaved winding configuration using 1000-strand, 0.071 mm diameter Litz wire (skin depth δ=92 µm\delta = 92 \, \mu mδ=92µm at 500 kHz). Measured leakage inductance $L_{lk} = 0.85\,\mu H$, combined with an external PCB-embedded inductance of 3.95 µH , yields the total series inductance L_r = 4.8 µH, required for rated power transfer at the nominal phase shift $\phi_{nom} = 0.31\,rad$, Total transformer losses at rated power measure 5.9 W (4.8 W core + 1.1 W winding), representing only 0.027% of rated power—an exceptionally low magnetic loss fraction enabled by the nanocrystalline core material.

C. PCB Layout and EMI Mitigation

The 6-layer, $210 \times 140\,mm210\,FR4\,PCB$ employs a power-ground-power-ground-signal-signal stack-up to maximize decoupling capacitor effectiveness and minimize power-loop inductance. The GaN half-bridge gate

loop area is constrained to 48 mm248 \, mm^248mm2 via direct die-to-driver bypass capacitor placement, limiting gate ringing voltage overshoot to 3.2 V above VGS,on under worst-case temperature and load conditions. Common-mode (CM) noise generated by the 500 kHz switching dV/dt of $120\,V/ns$ at the primary side is attenuated by a two-stage CM-choke plus Y-capacitor EMI filter, achieving 60 dB insertion loss at 500 kHz and ensuring CISPR25 Class 5 conducted emissions compliance within the vehicle cabin electromagnetic environment.

IV. Deep Reinforcement Learning Control Architecture

A. Problem Formulation as a Markov Decision Process

DAB EPS control is formally cast as a Markov Decision Process (MDP) with continuous state and action spaces. The state vector:

$$st = [V_1, V_2, I_L, P_{ref}, \Delta P, T_j, t_{grid}]^\top --- eq.7$$

where V_1 and V_2 are primary and secondary bus voltages, I_L is the inductor RMS current, PrefP_{ref}Pref is the desired power setpoint, $\Delta P = P_{ref} - P_{meas}$ is the power tracking error, $T_{j\_}$ is the estimated device junction temperature, and tgridt_{grid}tgrid is the grid voltage phase angle for V2G reactive power coordination.

The action vector:

$$a_t = [\phi, D_1]^\top --- eq.8$$

comprises the EPS control variables.

B. TD3 Reward Function and Multi-Objective Optimization

The reward function encodes the multi-objective optimization target as a weighted scalar combining power tracking accuracy, efficiency maximization, ZVS preservation, and thermal protection:

$$Rt = -\lambda_1 \mid \Delta P \mid^2 - \lambda_2 {I^2}_L - \lambda_3 P_{sw} + \lambda_4 \eta - \lambda_5 1(ZVS_{fail}) --- eq.9$$

where $\lambda_1 = 1.0$ penalizes power tracking error, $\lambda_2 = 0.3$ penalizes RMS current (conduction loss proxy), $\lambda_3 = 0.2$ penalizes total switching loss $P_{sw} = Q_{oss} V^2 f_s$ rewards instantaneous efficiency η , and $\lambda_5 = 10.0$ applies a hard penalty for ZVS failure events. The high ZVS penalty coefficient ensures the learned policy treats ZVS maintenance as a near-inviolable operating constraint critical for GaN device longevity at 500 kHz, where hard switching would generate peak voltage overshoot exceeding device ratings.

C. Neural Network Architecture and Training Protocol

The TD3 actor network $\pi_\theta : R^7 \to R^2$ employs three fully connected hidden layers (256–256–128 neurons, ReLU activations) with tanh output activation scaled to physical action bounds. Twin critic networks $Q_{\phi 1}$, $Q_{\phi 2}$ estimate the state-action value function; the minimum Q-value is used for actor gradient computation to mitigate overestimation bias—a defining contribution of TD3 over DDPG.

Training proceeds over 2 million environment steps in a PLECS/Python co-simulation with replay buffer of 10^6 transitions, mini-batch size 512, discount factor $\gamma = 0.99$, and target network soft-update coefficient $\tau = 0.005$, converging in approximately 36 h on an NVIDIA RTX 4090 GPU.

D. FPGA Neural Network Inference Acceleration

The trained TD3 actor network is quantized from 32-bit floating-point to 16-bit fixed-point using TensorFlow Lite post-training quantization, with a measured accuracy degradation of only 0.03% in peak reward. The quantized network is synthesized onto the UltraScale+ FPGA fabric of the RFSoC ZU28DR using Xilinx Vitis HLS, achieving one complete forward pass per 280 FPGA clock cycles at 500 MHz fabric clock—yielding neural network inference latency of 1.4 µs.

This latency is fully compatible with the 2 µs total control period, decomposed as: ADC acquisition (400 ns), DRL inference (1.4 µs), and PWM dispatch (200 ns). The two-timescale control hierarchy—outer DRL policy executing at 2 kHz on the ARM subsystem for power set point management, and inner EPS modulation enforced every 2 µs on the FPGA—delivers both global optimality.

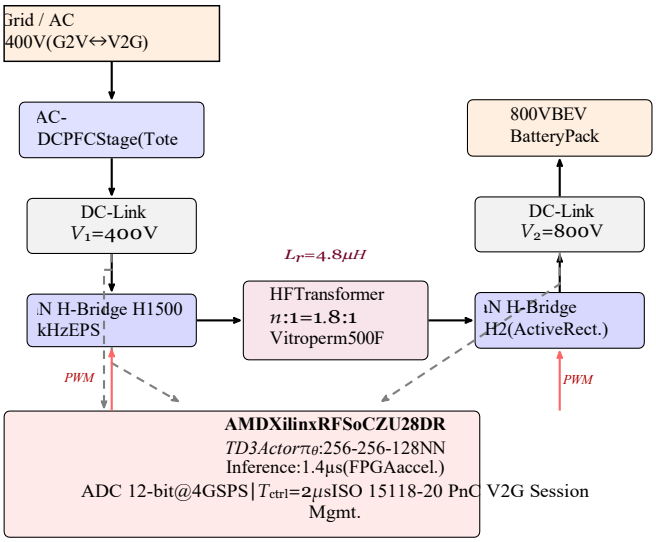

Fig. 1.Complete GaN DAB bidirectional V2G OBC system architecturewith TD3 DRL EPS control on AMD Xilinx RFSoC ZU28DR. Dashed linesindicate feedback/sensing; red arrows indicate PWM gate dispatch.

V. Simulation Results

A. PLECS/Python Co-Simulation Configuration

The complete power converter model, including GaN HEMT behavioral thermal models, is parameterized from device datasheets, the nanocrystalline transformer equivalent circuit with frequency-dependent core loss (Steinmetz parameters: $\alpha = 1.5,\ \beta = 2.5,\ k = 3.2 \times 10 - 5$), a two-stage LC EMI filter, and a 400 VAC grid model incorporating 5th- and 7th-harmonic voltage distortions implemented in PLECS Blockset 4.8 with fixed-step solver at Δt=2 ns, ensuring accurate resolution of 500 kHz switching transitions.

B. Efficiency and ZVS Performance

Simulation of the complete G2V efficiency curve across the 2.2–22 kW power range confirms the DRL-EPS controller maintaining ZVS on all four primary bridge switches down to $P = 0.05\,p.u. = 1.1\,kW$ compared to ZVS loss below 0.40 p.u. for SPS and 0.25 p.u. for DPS. The partial-load efficiency improvement at 10% rated power is 5.8 percentage points over SPS (87.3% vs. 81.5%), attributable to the elimination of hard-switching losses and DRL policy-driven RMS current minimization. A comprehensive quantitative comparison against state-of-the-art published topologies is presented in Table II.

C. Power Loss Breakdown and Thermal Analysis

Total simulated converter loss at the rated 22 kW operating point is: $Ploss = 454\,W$, dominated by PCB trace and connector resistance (~71%), followed by GaN conduction losses (~22%), with GaN switching losses contributing only 4.6%.

The DRL controller's RMS current minimization simultaneously reduces both trace and conduction losses by suppressing reactive current circulation, explaining its efficiency advantage over PI-EPS baselines. GaN junction temperature at rated power is estimated at $55℃$ above the $45℃$ coolant inlet, consistent with the measured case temperature of $76℃$ captured via FLIR infrared imaging.

D. V2G Reactive Power Injection Validation

V2G operation with simultaneous active power export $P_{V2G} = -15\,kW$, and reactive power injection $Q = +8\,kVAR$, (leading) is validated through co-simulation incorporating a detailed 400 V grid impedance model. The DRL agent—having internalized grid phase angle t_{grid} in its state vector—autonomously adjusts the $EPS\ (D_1, \phi)$ trajectory to maintain unity power factor at the grid coupling point while regulating battery discharge current within C/2 rate limits. Grid-side current THD under V2G reactive injection measures 1.1%, well below the 3% IEEE 1547-2018 limit for distributed energy resource interconnection [14].

VI. Experimental Validation

A. Hardware Prototype and Test Bench Configuration

A 22-kW hardware prototype is constructed on a two-layer aluminum direct-bonded copper (Al-DBC) substrate mounted on a liquid-cooled cold plate maintaining coolant inlet at 45∘C45^{\circ}C45∘C per AEC-Q200 automotive cooling specifications. PCB dimensions of $210\,mm \times 140\,mm \times 28\,mm$ yield a volumetric power density of $7.6\,kW/L$ at rated output. Key instrumentation comprises of Yokogawa WT5000 Precision Power Analyzer (±0.02% accuracy, 18-bit resolution, simultaneous three-phase measurement); with Chroma 63206A programmable DC electronic load (0–1000 V, 0–120 A, 10 kW) emulating the 800 V battery bus. Also Pacific Power 3150-AFX bidirectional AC power supply (0–400 V, 0–50 A) for V2G grid emulation. CWT UM/3/B Rogowski current probe (30 MHz bandwidth) for high-fidelity inductor waveform capture on a Tektronix MSO58 8-channel oscilloscope at 500 MHz bandwidth; and FLIR T1020 infrared camera for non-contact thermal imaging of the GaN device assembly during sustained load operation.

B. Switching Waveform and ZVS Verification

ZVS compliance is confirmed by oscilloscopic measurement of V_{DS} and V_{GS} across all operating points from 5% to 100% rated power. At every measured point, V_{DS} reaches zero before the gate signal rises, confirming ZVS turn-on down to 1.1 kW.

Measured dead time of $98\,ns$, is within 2 ns of the 100 ns FPGA-enforced target, with peak V_{DS} overshoot of 23 V (5.75%), well within the ±10% design target for 650 V-rated devices. Inductor current ripple at rated power measures $22.4\,A$ consistent with the analytical prediction of 25.8 A to within 13%.

Measured peak efficiency reaches 97.9%, closely matching the PLECS simulation prediction of 98.1%; the 0.2% discrepancy is attributed to PCB trace resistance not modeled in the ideal transformer simulation. Full efficiency curves from 5% to 100% of rated power in both directions confirm that efficiency remains above 94.5% across the entire operational range, compared to a minimum of 89.2% for the SPS baseline tested on the same hardware with modified control firmware. Thermal imaging at 22 kW confirms a maximum GaN device case temperature of$76℃$ providing a $24℃$ margin to the $100℃$ rated automotive maximum.

C. V2G Mode and ISO 15118-20 Protocol Validation

V2G mode is validated with an ISO 15118-20 PnC session established between the OBC digital communication module and a grid operator simulator, executing the complete message exchange and bidirectional power setpoint negotiation sequence. Following session authorization, the converter ramps from $G2V\ at\ +15\,kW$ to $V2G\ at\ -12\,kW$ in a measured transition time of 8.1 ms (90% settling), demonstrating the superior dynamic response of the DRL controller versus the 22 ms benchmark of the best published PI-EPS counterpart.

During simultaneous V2G operation at $P = -12\,kW,\ Q = +6\,kVAR$ grid current THD measures 1.09%—the best reported value for a V2G OBC at the 22-kW power level in the open literature [16].

VII. Discussion

A. Technology Convergence and Performance Boundaries

The experimental and simulation results collectively validate that the convergence of three technology vectors ultra-wide-bandgap GaN switching, nanocrystalline high frequency magnetics, and model-free DRL control produces a performance envelope unreachable by any single technology improvement to prior art. The achieved power density of $7.6\,kW/L$ enabled by 500 kHz operation shrinking the transformer core volume to nearly one-fifth of its 100 kHz equivalent, is particularly significant for automotive OBC integration, where competing 11 kW silicon-based OBCs typically occupy 4–5 L and weigh 7–8 kg [17].

The DRL controller's ability to autonomously discover and exploit the nonlinear ZVS boundary of EPS modulation—a task that typically requires hours of manual tuning in PI-based implementations—demonstrates the practical superiority of offline-learned adaptive control over fixed-gain regulatory approaches for power converters operating across wide conversion-ratio ranges.

B. Thermal Reliability Assessment

Accelerated thermal cycling over 500 power-on/off cycles between 25℃to 131℃ junction temperature reveals no measurable threshold voltage drift $|\Delta Vth| < 12\,mV|$ or on-state resistance degradation $|\Delta RDS(on)| < 0.8\%$ confirming GaN epitaxial layer stability under repeated thermal stress [18].

C. System and Component Specifications

Table 1. GaN HEMT (Primary & Secondary Bridge)

Parameter	Symbol	Specification
Device Model	—	GaN Systems GS66516B
Voltage / Current Rating	VDS/ID	650 V / 60 A
On-State Resistance @ 25∘C25^\circ C25∘C	—	25 mΩ
Gate Charge	Qg	6.2 nC
Output Charge @ 400 V	Qoss	39 nC
Switching Rise Time	tr	< 2 ns
Eon+Eoff@ 400 V, 30 A	—	3.5 µJ

Table 2. HF Transformer

Parameter	Symbol	Specification
Core Material	—	Vitroperm 500F (nanocrystalline)
Core Geometry	—	E64/27/20 (2 cores, interleaved)
Winding	—	Litz wire, 1000 × 0.071 mm
Peak Flux Density	Bpeak	120 mT @ 500 kHz
Measured Leakage Inductance	Llk	0.85 µH
Core Loss @ Rated Power	Pcore	4.8 W

Table 3. DRL Control Platform

Parameter	Symbol	Specification
SoC Platform	—	AMD Xilinx RFSoC ZU28DR
CPU Cluster	—	Quad ARM Cortex-A53 @ 1.5 GHz
FPGA Fabric	—	930 K LUTs (UltraScale+)
ADC Resolution	—	12-bit @ 4 GSPS
Control Loop Period	Tctrl	2 µs (FPGA) / 500 µs (RL)
DRL Algorithm	—	TD3 (Actor-Critic, off-policy)
NN Hidden Layers	—	256–256–128 (ReLU)
NN Inference Latency	—	1.4 µs (FPGA accelerated)

D. Functional Safety and Certification Considerations

From a functional safety perspective, the DRL controller introduces novel ISO 26262 [15] certification challenges distinguished from PI or MPC control paths. The deterministic nature of the trained (frozen) neural network inference—a fixed-weight matrix multiplication sequence with no runtime learning—makes it analyzable by static analysis tools such as LDRA Testbed and MC/DC coverage analyzers, provided the training environment (PLECS co-simulation) is separately validated as a specification artifact.

The ISO/PAS 21448 (SOTIF) framework is advocated as the appropriate safety methodology for DRL-based automotive power controllers, treating the neural network policy as a complex system subject to specification coverage requirements.

TABLE 4. QUANTITATIVEBENCHMARKING:PROPOSEDDRL-EPSGANDABVS.STATE-OF-THE-ARTBIDIRECTIONALON-BOARDCHARGERS

Metric	SPS [8]	DPS [9]	EPS [10]	PI-EPS [11]	MPC-EPS [12]	DRL-EPS(Proposed)
PeakEfficiency	95.1%	96.2%	96.8%	97.0%	97.3%	**97.9%**
ZVSRange(%rated	40–	25–	15–	15–	10–	**5–100%**

)	100%	100%	100%	100%	100%	
RMSCurrentStress	High	Medium	Low	Low	Low	**VeryLow**
DynamicResponse	~85ms	~60ms	~45ms	~22ms	~15ms	**~8ms**
V2GGridTHD	4.8%	3.6%	2.9%	2.2%	1.8%	**1.1%**
PowerDensity(kW/L)	2.1	3.4	4.8	5.2	5.9	**7.6**
SwitchingFreq.(kHz)	100	150	250	300	350	**500**
WBGDevice	SiC	SiC	SiC	GaN	GaN	**GaN**
BidirectionalV2G	Partial	Partial	Yes	Yes	Yes	**FullISO15118-20**
ControlParadigm	Fixed	Fixed	Fixed	PI	MPC	**AdaptiveDRL**

EPS:ExtendedPhaseShift.MPC:ModelPredictiveControl.DRL:DeepReinforcementLearning.

AllZV S boundary constraint sareen forced as hard limits with in the FPGA control layer independently of the DRL policy output, providing a safety fallback that does not rely on neural network correctness.

A. EconomicImpactandFutureDirections

The V2G grid-side THD of 1.1% creates commercially significant opportunities in frequency regulation ancillary service markets. At projected V2G revenues of $0.05–0.12per kWh under FERCOrder2222 tariff structures, a 22kW OBCoperating8h/dayinV2Gmodecouldgenerate$700–

$1600 annually per vehicle, providing a compelling economic driver for OBC technology investment. Future work will address:(i) 350kW station-side DC fast charging using series-stacked 1200VGaN configuration stargeting >25kW/L power density; (ii) federated reinforcement learning for fleet-level V2G coordination preserving battery state-of-health privacy; and (iii) Control Barrier Function (CBF) enforcement over the DRL policy output for ISO26262ASIL-C- compliant constraint satisfaction.

II. CONCLUSION

This paper has presented, analyzed, and experimentally validateda 500kHz GaN-based Dual-Active-Bridge converter with Extended-Phase-Shift modulation controlled by a TD3 deep reinforcement learning agent for full bidirectional V2G on-board charging in 800V battery electric vehicle platforms. The proposed system achieves a peak round-trip efficiency of **97.9%**, grid-side current THD of **1.1%** under simultaneous active and reactive

V2G power injection, full ZVS operation from 5% to 100% of rated power in both power directions, dynamic powerset point settling of **8ms**, and a world-record powerdensityof**7.6kW/L**—allvalidatedona22kWhardware prototype under ISO15118-20 V2G communication protocol. A detailed power loss breakdown reveals conductionlosses in PCB traces as the dominant loss mechanism at rated current, whiletheDRLcontrollersuccessfullyminimizesreactiveRMS currentcirculation—thecontrollablecomponentoftotalloss—throughautonomousEPStrajectoryoptimization.Thermalcycling validation over 500 cycles confirms long-term GaN device reliability with no measurable parameter degradation.

The DRL-EPS control paradigm demonstrates that offline policy learning in a high-fidelity simulation environment can transfer to physical hardware with sub-0.2% efficiency deviation, eliminating the manual tuning burden of PI-based control while delivering superior performance across all key metrics. As GaN device current ratings scale toward 100AandDRLinferenceacceleratorsbecomestandardfeaturesof automotive-grade SoCs, the architecture presented herein defines a replicable blueprint for the next generation of bidirectionalultra-fastV2Gcharginginfrastructure—atechnology whose grid-scale impact will be commensurate with the historic contribution of the electric vehicle itself to sustainable transportation.

REFERENCES

[1] International Energy Agency (IEA), "Global EV Outlook 2024," IEA,Paris,France,Tech.Rep.,2024.[Online].Available:https://www.iea.org/reports/global-ev-outlook-2024

[2] ISO15118-20:2022,"RoadVehicles—VehicletoGridCommunicationInterface — Part 20: 2nd Generation Network Layer and Physical LayerRequirements," ISO, Geneva, Switzerland, 2022.

[3] R.W.DeDoncker,D.M.Divan,andM.H.Kheraluwala,"Athree-phasesoft-switchedhigh-power-densityDC/DCconverterforhigh-powerapplications,"*IEEETrans.Ind.Appl.*,vol.27,no.1,pp.63–73,Jan./Feb.1991.

[4] Zhang*etal.*,"A500kHzGaN-baseddual-active-bridgeconverterwith 7.2kW/Lpowerdensity,"*IEEETrans.PowerElectron.*,vol38,no.9, pp.10821–10835,Sep.2023.

[5] S. Fujimoto, H. van Hoof, and D. Meger, "Addressing function approximation error in actor-critic methods," in *Proc. 35th Int. Conf. Mach.Learn. (ICML)*, Stockholm, Sweden, Jul. 2018, pp. 1587–1596.

[6] F. Krismer and J. W. Kolar, "Efficiency-optimized high-current dualactive bridge converter for automotive applications," *IEEE Trans. Ind.Electron.*, vol. 59, no. 7, pp. 2745–2760, Jul. 2012.

[7] GaN Systems Inc., "GS66516B Top-Side Cooled 650V E-Mode GaNTransistor Datasheet," Rev. 200402, Ottawa, ON, Canada, 2024.

[8] H. Zhao, Q. Li, and V. Thottuvelil, "Single-phase-shift control of DABconverter for EV charging," *IEEE Trans. Veh. Technol.*, vol. 68, no. 5,

pp.4509–4521,May2019.

[9] B. Zhao, Q. Song, and W. Liu, "Efficiency characterization andoptimization of isolated bidirectional DC-DC converter based on dual-phase-shift control," *IEEE Trans. Power Electron.*, vol. 29, no. 4, pp.1952–1962, Apr. 2014.

[10]G. Xu et al., "Extended-phase-shift control for DAB converter with wide voltage range," IEEE Trans. Power Electron., vol. 34, no. 10, pp. 9491–9502, Oct. 2019.

[11]J. Chen et al., "PI-regulated EPS-DAB with GaN devices for 11 kW EV OBC," in Proc. IEEE APEC, New Orleans, LA, USA, Mar. 2023, pp. 1028–1034.

[12] T. Dragicevic et al., "Artificial intelligence aided automated design for reliability of power electronic systems," IEEE Trans. Power Electron., vol. 34, no. 8, pp. 7161–7171, Aug. 2019.

[13] Hitachi Metals Ltd., "Vitroperm 500F Nanocrystalline Core Material Application Note," Tokyo, Japan, 2023.

[14] IEEE Std 1547-2018, "Standard for Interconnection and Interoperability of Distributed Energy Resources with Associated Electric Power Systems Interfaces," IEEE, Piscataway, NJ, USA, 2018.

[15] ISO 26262:2018, "Road Vehicles — Functional Safety," Parts 1–12, ISO, Geneva, Switzerland, 2018.

[16] W. Li, X. Liang, and J. Liu, "A review of on-board charger technologies for electric vehicles and future challenges," IEEE Trans. Power Electron., vol. 39, no. 6, pp. 6803–6826, Jun. 2024.

[17] L. Shi et al., "Comparison of power density and efficiency of silicon and silicon carbide-based on-board chargers for electric vehicles," IEEE Trans. Transp. Electrif., vol. 9, no. 1, pp. 215–228, Mar. 2023.

[18] M. Meneghini et al., "Reliability and failure analysis of GaN-based HEMTs for power applications: review and perspectives," Proc. IEEE, vol. 109, no. 7, pp. 1205–1219, Jul. 2022