IoT-Based DC/DC Deep Learning Power Converter Control: Real-Time Implementation

Meysam Gheisarnejad and Mohammad Hassan Khooban, *Senior Member, IEEE*

***Abstract*—Recently, a modularized smart grid (SG) architecture, entitled the Internet of Things (IoT) grid, is developed that accommodates the IoT technology into the dc–dc converters to build a programmable grid with a single voltage bus. This modern architecture can be established with low computing hardware that facilitates the control and management of the IoT-based grids. Due to the uncertainties originated from the integration of the IoT technology and power electronic converters, the deterministic methodologies are unable to precisely model the SG anymore. In response to these challenges, this article addresses a novel adaptive data-driven method based on the active disturbance rejection controller (ADRC) for the voltage regulation of an IoT-based dc–dc buck converter feeding constant power loads. In particular, a deep deterministic policy gradient (DDPG) with the actor-critic architecture is adopted for the online adjusting of the ADRC controller. The established DDPG takes into account the ADRC controller coefficients into the design objective and offers the ADRC controller with the online coefficient setting ability through the neural network learning. The IoT-based system is tested on a real-time testbed with the constrained application protocol protocol and IEEE 802.11 (Wi-Fi) network to assess the applicability of the suggested controller in the presence of network degradations. The impact of both packet loss and interfering traffic on the reduction performance of the DDPG adaptive ADRC controller is investigated, simultaneously. The supremacy of the suggested adaptive data-driven controllers is verified by a comprehensive comparative analysis with the state-of-the-art methodologies.**

***Index Terms*—Constrained application protocol (CoAP), deep deterministic policy gradient (DDPG), Internet of Things (IoT).**

I. Introduction

POWER electronic-based onboard dc microgrids have drawn a lot of attention, mainly because of their benefits in terms of production, size, insolation, flexibility, controllability, etc. [1]. In the context of the microgrid design, dc is preferred to ac due to the majority of the modern electronic equipment, energy storage/generation systems use the dc power, and they do not have some of the problems that the ac microgrids face. As a key technology in the modern power grids [2], [3], the dc microgrids also provide a higher level of security and reliability with lighter congestion over ac. The cascaded dc power converters, however, are prone to the dc bus voltage distraction or even instability caused by the constant power load (CPL) feature of the point-of-load converters [4]–[6].

Up to now, various active nonlinear methodologies have been developed to alleviate the instability effects of the CPLs by improving the control law of injecting, such as backstepping algorithm [7], sliding mode control [8], and nonfragile [9]. The robust control approaches encounter many difficulties due to the unmodeled dynamics and nonlinearity features that make the control design of the dc–dc converters extremely complex or even beyond the capability of nonlinear control techniques. Moreover, the model-based techniques designed based on the linearized model of the dc–dc converters are generally inadequate to compensate for the destructive impedance instabilities imposed by CPLs. To rectify the limitations of the model-based techniques, the data-driven approaches have drawn a lot of attention over the past two decades [10], [11]. The main reason for the popularity of data-driven approaches is that these methods are designed based on the input–output (I/O) data of the controlled system and thus the model identification, unknown dynamics, and mathematical formulation will be disappeared. Among the most common noniterative data-driven control techniques include model-independent approaches, such as the active disturbance rejection control (ADRC) [12], while the other common iterative data-driven approaches are the iterative feedback tuning [13], [14], artificial neural network (NN) [15], and deep reinforcement learning (DRL) [16].

In a scenario with the various types of backup devices and several CPLs, there is a very urgent demand to monitor energy consumption and to control the power generation. However, the conventional dc converters use the point-to-point communication [17] from each CPL to the power supply, which is not economical for such configurations of dc microgrids. This follows a new generation of the dc converters that have the capability of supporting multiple kinds of backup devices with many CPLs, which leads to the main essence of the networked dc converters in the smart grid (SG) area [17], [18]. The Internet of Things (IoT) [19] is regarded as the key enabling technology to transform the existing dc converters to smart dc converters. To enable the remote integration of communication and control for the fast design and deployment of the networked systems, various IoT protocols (e.g., message queuing telemetry transport [20], [21], constrained application protocol (CoAP) [22], [23], extensible messaging and presence protocol [24], etc.) have

Manuscript received February 21, 2020; revised April 26, 2020; accepted May 1, 2020. Date of publication May 10, 2020; date of current version July 31, 2020. Recommended for publication by Associate Editor C. K. Tse. *(Corresponding author: Mohammad Hassan Khooban.)*

Meysam Gheisarnejad is with the Department of Electrical Engineering, Najafabad Branch, Islamic Azad University, Isfahan 1477893855, Iran (e-mail: me.gheisarnejad@gmail.com).

Mohammad Hassan Khooban is with the Department of Engineering, Aarhus University, 8200 Aarhus, Denmark (e-mail: khooban@eng.au.dk).

Color versions of one or more of the figures in this article are available online at https://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TPEL.2020.2993635

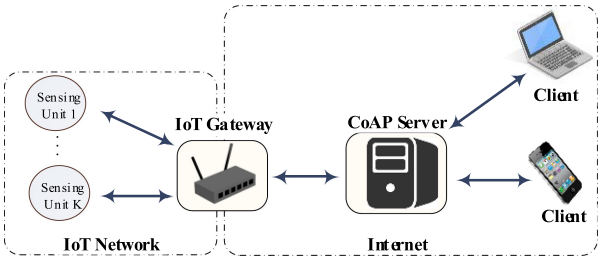

Fig. 1. Overview of a typical IoT architecture.

emerged as the IoT standards (see Fig. 1). For the networking support, Tanyingyong *et al.* [23] implemented the CoAP client as a communication protocol of the networked dc converter, since this protocol is specially designed for the elements with limited resources. Despite the CoAP client used in [23] can transfer the data below the response time required by the SG functions, the processing delays of the IoT elements degrade the IoT-grid outcomes when a sequence of commands execute on multiple IoT elements. In the presence of the processing delay, a methodology based on the sending burst commands and scheduling the outcomes is introduced in [25] for multiple converters. While the IoT-network infrastructure supports the transmission of the large amount of information over the vast geographical regions, it poses new threats to the stability of the power electronic system control. As the degradation factors are unavoidable in the real IoT networks, such as time delay, packet loss, cyber attack, and congestion, these factors affect adversely the stability of the communication-based power systems if these degradations are not handled efficiently [25], [26]. Therefore, it is necessary to provide extra actions for the precise control commands to ensure high stability and robustness against the different types of network degradation processes.

Robustness to the network degradations and dynamic uncertainties makes ADRC a feasible solution for the networked dc–dc converters feeding CPLs. In the ADRC scheme, first, both the network degradations and unknown internal dynamics and even computation inaccuracies are regarded as a total disturbance to be estimated by the extended state observer (ESO). Then, the total disturbance is eliminated by the nonlinear state error feedback (NLSEF) control law while reducing the plant to a set of pure integrators. The application of the original ADRC has been investigated for the cyber-physical systems [27], but ignoring the network degradations in the controller design restricts its applicability. To reduce the insensitivity of the communication-based systems to the network unreliable factors, some modifications have been proposed to ameliorate the ADRC performance. For instance, in [28], Han *et al.* examined the performance of ADRC on a nonlinear model of a 39-bus benchmark over the local area network using the real-time simulation. In [28], the automatic controller is locally installed in each participating production component to prevent the communication delay for the ADRC control signals. To address the effect of the communication constraints, Yu *et al.* [29] proposed a nonlinear sampled-data ESO based ADRC for a communication-based system. However, in all the methods, it is assumed that the shared network is prone to small network degradations. For a higher level of network degradations, some typical predictive approaches, such as smith predictor and predictor observer, can be integrated with the ADRC to estimate the unreliable factors [30]. However, utilizing these strategies lead to increasing the design complexity and difficulties in the practical application.

The interaction between the power electronic converters and communication infrastructures brings new uncertainties to the stability of the IoT-based dc–dc converters [18], [23], [25]. These vulnerabilities of the power electronic systems become more severe when the dc–dc converters feed the CPLs. To implement an efficient control strategy to handle the destabilization problem of the networked dc–dc converter, this work proposes a self-adaptive ADRC for the voltage regulation of dc–dc buck converters in the IoT framework. The procedure of tuning the nonlinear feedback coefficients is difficult to acquire the ideal solution and desirable efficiency using the ADRC. To solve the problem, several strategies, such as heuristic techniques and fuzzy logic, have been used to design the coefficients embedded in the nonlinear feedback law [31]. Recently, a deep deterministic policy gradient (DDPG) has become a hotspot in the area of artificial intelligence and machine learning as a powerful tool to solve the high dimensional problems [32]–[34]. The core idea of DDPG is to assess a decision-making specification in a certain kind of the system by giving a performance index to learn a set of agents through lots of plant data produced in the decision-making procedure. The main contributions toward an adaptive data-driven control design for the regulation of the IoT-based system, considering the networked degradations, are summarized as follows.

1) An IoT-based dc–dc power electronic converter with the CoAP protocol is developed, which provides studying the CPL's instability and network degradations, simultaneously.
2) A data-driven control scheme based on ADRC is implemented for the complicated IoT-based system, instead of depending on the more exact knowledge of the controlled plant.
3) The DDPG algorithm, as an integration of the interactive training process of reinforcement learning (RL) and deep learning with the artificial NNs, is adopted for the optimal setting of the ADRC.
4) Simultaneously exploring the effects of the network degradations, such as packet loss and interfering traffic, on the IoT-based system performance can also be regarded as a remarkable outcome of this article.

II. Model Descriptions of the DC/DC Buck Converter

Fig. 2 illustrates the structure of a typical pulsewidth modulation (PWM) type dc–dc buck converter, which contains a dc input voltage E, a filter capacitor, a filter inductor, a PWM gate drive regulated switch, a diode D, and a CPL [1], [7].

Following a PWM pattern generated by a hysteresis controller, the power transistor switches between the cutoff mode and saturation mode. To explore the large-signal feature of the dc–dc converter tackling the CPLs, a dependent current generator as

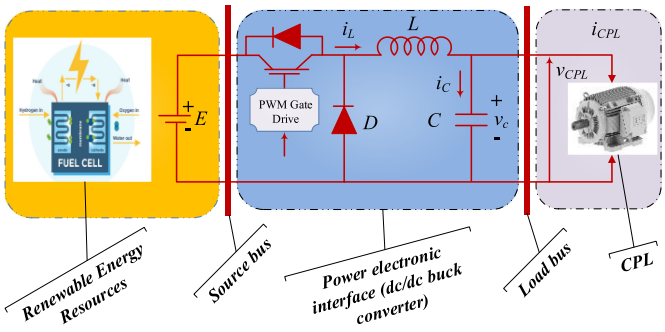

Fig. 2. Simplified circuit diagram of a dc/dc buck converter feeding CPL.

given as follows is often employed to investigate the behavior of the load for the various ranges of voltages and frequencies:

$$i_{\mathrm{CPL}}(t) = \frac{P}{v_{\mathrm{CPL}}(t)} \quad \forall\, v_{\mathrm{CPL}}(t) > \varepsilon \tag{1}$$

where i_{CPL} and v_{CPL} represent the instantaneous current and voltage of the CPL, respectively, P is the rated power, and ε has a small positive value.

Based on the average switching scheme [35], the state-space averaged form of the buck converter is mathematically expressed as

$$\begin{bmatrix} \frac{dx_1}{dt} \\ \frac{dx_2}{dt} \end{bmatrix} = \begin{bmatrix} \frac{1}{L}\left[E \cdot u - x_2\right] \\ \frac{1}{C}\left(x_1 - \frac{x_2}{R} - \frac{P}{x_2}\right) \end{bmatrix} \tag{2}$$

$$y = v_o \tag{3}$$

where the average of the inductor current i_L and voltage capacitor v_C is represented by x_1 and x_2, respectively, the capacitance and inductance of the buck converter are represented by C and L, respectively, and $u \in \{0, 1\}$ is the control input. For simplicity and to express the dynamic model according to the prevalent control theory, it is assumed the parameters x_1 and x_2 be constrained in the modeling of CPL. Likewise, the reference voltage is represented by v_{ref}, and the voltage tracking error is represented by $e = v_o - v_{\mathrm{ref}}$.

The control objective for the dc–dc buck converter is to the capacitor voltage v_C (or output voltage v_o) track the reference voltage v_{ref} in the presence of large disturbances. Since the resistive load has a damping property and CPL reduces the damping, the worst condition scenario, from the stability point of view, is that the total load has a pure CPL nature. This motivates the development of a robust control strategy to stabilize the voltage outcome against the worst-case instability condition caused by the CPL.

III. IoT-Grid Architecture

As a practical use, a programmable and low-cost networked dc–dc buck converter is developed, where the IoT technology is adopted as the networking infrastructure. By adding the network support to the power electronic converters, the remote monitoring/control can be realized for the microgrids with the distributed energy resources [fuel cell, photovoltaic generator and storage units (e.g., batteries and flywheels)] and local loads

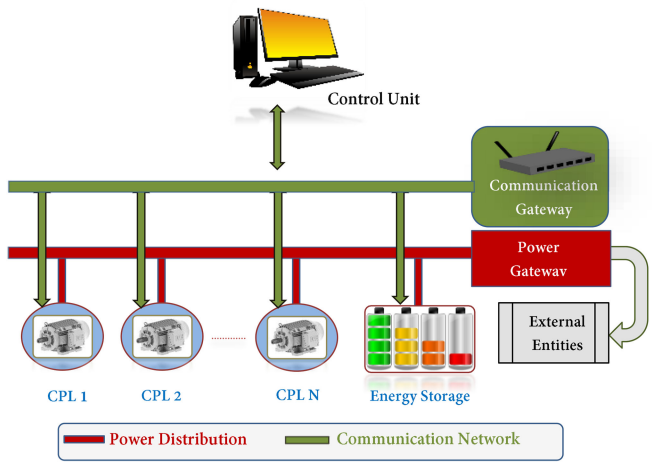

Fig. 3. Illustration of the IoT grid.

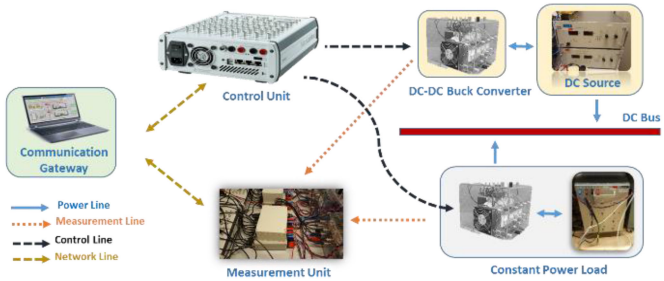

Fig. 4. Experimental testbed of the IoT-based dc–dc buck converter.

[18], [23]. The general architecture of the IoT grid, which is made of two functional networks including the physical energy grid and communication network, is illustrated in Fig. 3 [25]. The physical energy grid is realized by a power distribution network, which is responsible for the power distribution in the microgrids. The IoT grid is connected to other grids by the power gateway, while the communication gateway is utilized to interconnect the IoT grid to external entities. The distributed energy resources and loads are used as the smart devices that can transmit the measured data and control signals to/from the communication gateway. The grid control unit is made by a cluster of the dc–dc converters, which maintains the voltage of the power distribution in a single value while providing various voltages to the CPLs. The main task of the communication network is to offer a remote communication for the various subsystems of the power distribution network, i.e., providing information transfers between the controllable energy resources and control part in the physical layer.

As a proof of the concept, an IoT-based dc–dc buck converter prototype is constructed as illustrated in Fig. 4. In this application, the CoAP protocol over IEEE 802.11 (Wi-Fi) with a data rate of 6 Mb/s is adopted for the data collecting and transmission. A laptop is implemented as the main Internet gateway. The control methodology is implemented in a dSPACE MicroLabBox with DS1202 Power PC dual-core 2 GHz processor board and DS1302 I/O board. The real-time simulation is developed on Windows 10, Core i7, 2.6 GHz, and 8 GB of RAM. The

real-time interface (RTI) based on the MicroLabBox is adopted to implement the software models on the dSPACE platform. A MATLAB C-code generator Simulink Coder is established in this application to automatically run the Simulink models in the real time-tested platform. The connection between the model and dSPACE I/O board can be realized by, directly, dragging the I/O module from the RTI block library onto the model. In this setup, the model code is generated by the Simulink Coder, while the blocks that implement the I/O abilities of the dSPACE platform in the Simulink models are provided by RTI.

IV. ADRC Algorithm

The objective of this article is to design an adaptive ADRC control scheme to stabilize the voltage output of the IoT-based dc–dc buck converter in the presence of the CPLs and network degradations, such as packet loss and interfering traffic.

ADRC is straightforward to implement and does not need the complete knowledge of the controlled plant and unmodeled dynamics but the relative gain b_0 of the controlled plant and its order p [28], [29], [36]. Therefore, in the ADRC design, the model of the controlled plant can be formulated as

$$y^{(p)}(t) = b_0 u(t) + f(\cdot) \tag{4}$$

where $f(\cdot)$ denotes the lumped unknown dynamics, resulting from the uncertainties and perturbations.

To ameliorate the tracking performance and prepare smooth signals for the controller, a transition process is generated by a tracking differentiator (TD). Considering the term $y^*(t)$ as the reference signal of the feedback control plant, the conventional form of TD for a second-order system is given as follows:

$$z_{11}(t+1) = z_{11}(t) + h z_{12}(t) \tag{5}$$

$$z_{12}(t+1) = z_{12}(t) + h \cdot fhan\left(z_{11}(t) - y^*(t), z_{12}(t), r, h\right) \tag{6}$$

where $z_{11}(t)$ and $z_{12}(t)$ are the outputs of the TD, h is the sampling period, and r is the variable deciding the track pace. Likewise, $fhan(\cdot)$ is a time-optimal integrated function, which is expressed as

$$\begin{aligned} \xi &= r \cdot h, \quad \xi_0 = \xi \cdot h \\ y &= z_{11} + h z_{12} \\ a_0 &= \sqrt{\xi^2 + 8 \cdot r |y|} \\ a &= \begin{cases} z_{12} + \frac{a_0 - \xi}{2} \text{sign}(y), & |y| > \xi_0 \\ z_{12} + \frac{y}{h}, & |y| \le \xi_0 \end{cases} \\ fhan(\cdot) &= \begin{cases} -r \text{sign}(a), & |a| > \xi \\ -r \frac{a}{\xi}, & |a| \le \xi. \end{cases} \end{aligned} \tag{7}$$

In the ADRC, both the uncertainties and perturbations in the controlled plant are considered as the total disturbance, which is estimated by an ESO. Based on the ESO scheme, a three-order

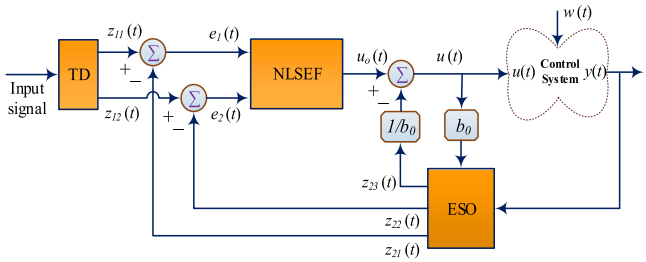

Fig. 5. Illustration of the (a) actor network and (b) critic network.

ESO is described as

$$\begin{cases} e = z_{21} - y \\ \dot{z}_{21} = z_{22} - \beta_{01} e \\ \dot{z}_{22} = z_{23} - \beta_{02} fal\left(e, \frac{1}{2}, \delta\right) + b_0 u \\ \dot{z}_{23} = -\beta_{03} fal\left(\varepsilon, \frac{1}{4}, \delta\right) \end{cases} \tag{8}$$

where z_{21} and z_{22} denote the state outputs, z_{23} denotes the disturbances estimation offered by the ESO, β_{01}, β_{02}, and β_{03} are the adjustable coefficients of the ESO, and δ is a filtering factor. The nonlinear function $fal(\cdot)$ is defined as

$$fal(e, \alpha, \delta) = \begin{cases} |e|^{\alpha} \text{sign}(e), & |e| > \delta \\ e/\delta^{1-\alpha}, & |e| \le \delta. \end{cases} \tag{9}$$

The ADRC also establishes the NLSEF to combine nonlinearly the output signals of TD (z_{11}, z_{12}) and the estimated states (z_{21}, z_{22}) provided by the ESO. The law control of the NLSEF is expressed as follows:

$$\begin{cases} e_1 = z_{11} - z_{21} \\ e_2 = z_{12} - z_{22} \\ u_0 = \beta_1 fal(e_1, \alpha_1, \delta) + \beta_2 fal(e_2, \alpha_2, \delta) \end{cases} \tag{10}$$

where β_1 and β_2 represent the control gains. With the real-time estimation of the unknown disturbance by ESO, the disturbance compensation is obtained by the following control law:

$$u(t) = u_0 - z_{23}/b_0. \tag{11}$$

According to three modules (TD, ESO, and NLSEF), the structure of ADRC is depicted in Fig. 5, where w (t) denotes the external disturbance.

V. Adaptive ADRC Based on DDPG

A. Mathematical Model of DDPG

RL aims to discover an optimal policy to an artificial agent interact with the environment, E, while maximizing the agent's reward during a sequence of time steps. For an observable environment, the interaction process can be expressed by a Markov decision process under the Markovian specification of the underlying environment [33].

1) $s \in \mathbb{S}$: State space, the finite collection of the states that can be considered for the given environment.
2) $a \in \mathbb{A}$: Action space, the finite collection of the available actions that the agent can execute leading to a transition from state s_t at time t to state s_{t+1} at time $t+1$.

3) $r \in \mathbb{R}$: Reward, for an undertaken action at each step time. An immediate feedback is emitted from the environment.
4) $\mathbb{P}(s'|s, a)$: Transition probability distribution, the probability that by performing action a, the state s at time t will move to the state s' at time $t+1$.
5) $\pi(a|s)$: Policy, implies the agent's behavior that maps states s_t to a probability distribution over the control actions a_t.

To formalize the RL problem, the total discounted reward from each state is defined as the return $G_t = \sum_{k=0}^{\infty} \gamma^k r_{t+k}$, where $\gamma \in (0,\ 1)$ denotes the discount factor. The ultimate goal is to maximize the expectation of a discount, that is

$$J = \mathbb{E}_{r_i, s_i \sim E, a_i \sim \pi}\left[G_1\right]. \tag{12}$$

For a specific policy π (where $\pi = a_t$), a state value function V^{π} as given in (13) is defined as a representation of total discounted reward G_t for each $s \in \mathbb{S}$ [considering a deterministic policy $\pi(s_t)$]

$$V^{\pi}(s) = \mathbb{E}_{\pi}\left[G_t | s_t = s\right]. \tag{13}$$

The above-mentioned function can be recursively described with the Bellman equation as follows:

$$V^{\pi}(s) = \mathbb{E}_{\pi}\left[r_t + \gamma V^{\pi}(s_{t+1}) | s_t = s\right]. \tag{14}$$

Similarly, a decomposition of the action-value function can be defined for the state-value function, which starts from the return

$$Q^{\pi}(s, a) = \mathbb{E}_{\pi}\left[G_t | s_t = s, a_t = a\right] \tag{15}$$

to the recursive relationship achieved with the Bellman equation

$$\mathcal{Q}^{\pi}(s, a) = \mathbb{E}_{\pi}\left[r_t + \gamma \mathcal{Q}^{\pi}(s_{t+1}, a_{t+1}) | s_t = s, a_t = a\right]. \tag{16}$$

From the action-value function, the optimal policy $\pi^* = \arg\max_a \mathcal{Q}^*(s, a)$ can be derived, where the Bellman equation employing $\mathcal{Q}^*$can be optimally satisfied.

In the context of the RL algorithms, DDPG is well known as one successful implementation of NNs to the RL paradigm and it has proven to perform well to solve continuous problems. In this algorithm, two deep NNs are designed, i.e., both the functions $\mathcal{Q}(s_t, a_t)$ and $\mu(s_t)$ are approximated by the deep NNs $\mathcal{Q}(s_t, a_t|\theta^Q)$ and $\mu(s_t|\theta^{\mu})$, respectively, where θ^Q and θ^{μ} are the coefficients of the critic and actor networks [32].

The updating of the critic's coefficients is based on the stochastic gradient descent technique to minimize a loss function

$$\mathcal{L}\left(\theta^{\mathcal{Q}}\right) = E_{(s,a)}\left[\left(\mathcal{Q}\left(s_t,\ a_t|\theta^{\mathcal{Q}}\right) - y_t\right)^2\right] \tag{17}$$

where

$$y_t = r_t\left(s_t,\ a_t\right) + \gamma \mathcal{Q}\left(s_{t+1}, \mu(s_t|\theta^{\mu})\ |\theta^{\mathcal{Q}}\right). \tag{18}$$

The actor's coefficients θ^{μ} can be updated according to the following policy gradient:

$$\begin{aligned}\nabla_{\theta^{\mu}} J^{\theta^{\mu}} &\approx \mathbb{E}_{s_t \sim \rho^{\beta}}\left[\nabla_{\theta^{\mu}} \mathcal{Q}\left(s, a|\theta^{\mathcal{Q}}\right)|_{a=\mu(s|\theta^{\mu})} \nabla_{\theta^{\mu}} \mu\left(s|\theta^{\mu}\right)\right] \\ &= \mathbb{E}_{s_t \sim \rho^{\beta}}\left[\nabla_a \mathcal{Q}(s, a|\theta^{\mathcal{Q}})|_{a=\mu^{\theta}(s)} \nabla_{\theta^{\mu}} \mu\left(s|\theta^{\mu}\right)\right]\end{aligned} \tag{19}$$

Algorithm 1: The pseudo-code of the DDPG algorithm.

Initialize critic $\mathcal{Q}(s, a|\theta^Q)$ and actor $\mu(s|\theta^{\mu})$ neural networks with with weights $\theta^{\mathcal{Q}}$ and θ^{μ}
Initialize target network $\mathcal{Q}'$and $\mu'$with weights $\theta^{\mathcal{Q}'} \leftarrow \theta^{\mathcal{Q}}$, $\theta^{\mu'} \leftarrow \theta^{\mu}$
Set up empty replay buffer R
for episode = 1 to M **do**
  Begin with a random noise $\mathcal{N}(x|b) \sim \frac{1}{2b_t}\exp\left(-\frac{x}{b_t}\right)$for exploration
  Receive initial observation state s_1
  for t = 1 to T **do**
    Select action $a_t = \mu(s_t|\theta^{\mu}) + \mathcal{N}$ based on current actor policy and exploration noise
    Apply action $a_t$to environment and observe next state s_{t+1}
    Compute reward r_t according to disparity between values of simulated and observed behavior
    Store following transitions (s_t, a_t, r_t, s_{t+1}) into replay buffer R
    Sample random minibatch of N transitions from R
    Set $y_i = r_i + \gamma \mathcal{Q}'\left(s_{i+1}, \mu'(s_{i+1}|\theta^{\mu'})\ |\theta^{\mathcal{Q}'}\right)$
    Update critic with the loss: $L = \frac{1}{N}\sum_i \left(y_i - \mathcal{Q}(s_i, a_i|\theta^{\mathcal{Q}})\right)^2$
    Update the actor policy using the sampled policy gradient:
    $\nabla_{\theta^{\mu}} J^{\theta^{\mu}} \approx \frac{1}{N}\sum_i \nabla_a \mathcal{Q}(s, a|\theta^{\mathcal{Q}})|_{a=\mu^{\theta}(s)} \nabla_{\theta^{\mu}} \mu(s|\theta^{\mu})$
    Update the target networks:
    $\theta^{\mathcal{Q}'} \leftarrow \tau\theta^{\mathcal{Q}} + (1-\tau)\theta^{\mathcal{Q}'}$
    $\theta^{\mu'} \leftarrow \tau\theta^{\mu} + (1-\tau)\theta^{\mu'}$
  end for
end for

where ρ is the discounted distribution and β is a specific policy to the current policy π.

As first proven by the deep Q network, a reply buffer and the target NNs are also adopted to further enhance the efficiency and robustness during the training. Based on the reply buffer technique, each experience tuple $e = (s_t, a_t, r_t, s_{t+1})$ of each time step is saved in a R-sized experience memory $\mathcal{D} = \{e_1, e_2, \ldots, e_R\}$. In each step of the training process, a minibatch of the tuples are uniformly sampled from the memory R. To avoid the instability of the DDPG learning, two additional NNs $\mathcal{Q}'(s, a|\theta^{\mathcal{Q}'})$ and $\mu'(s|\theta^{\mu'})$, the so-called target networks, are also adopted for the actor and critic NNs. The weight coefficients of the target networks $\theta^{\mathcal{Q}'}$ and $\theta^{\mu'}$ are softly updated by $\theta^{\mathcal{Q}'} \leftarrow \tau\theta^{\mathcal{Q}} + (1-\tau)\theta^{\mathcal{Q}'}$ and $\theta^{\mu'} \leftarrow \tau\theta^{\mu} + (1-\tau)\theta^{\mu'}$ with $\tau \in (0, 1)$. Besides, an exploration noise $\mathcal{N}$ based on Laplace process $\mathcal{N}(x|b) \sim \frac{1}{2b_t}\exp(-\frac{x}{b_t})$ [37] is added to the original actor actions [i.e., $a_t = \mu(s_t|\theta^{\mu}) + \mathcal{N}$] for exploration purposes. The algorithm for the standard DDPG scheme is summarized in Algorithm 1.

B. DDPG Based on the Coefficient Tuning Mechanism of the ADRC

The quality of the ADRC control actions highly depends on the adjustment of its coefficients, which is essentially realized by practical experiments. Numerous works demonstrate that the ADRC scheme can be fully designed with the principle of "separability," namely to design separately the TD, ESO, and the feedback control law and integrate it into the full ADRC.

Among them, α_1, α_2, and δ are the invariable coefficient, the three coefficients of ESO (β_{01}, β_{02}, and β_{03}) can be calculated automatically, only the NLSEF gains (β_1 and β_2) need to be adjusted manually. Due to the presence of the nonlinear function in the NLSEF, manually tuning the size of the NLSEF gains is not in favor of the actual performance and temporary variable

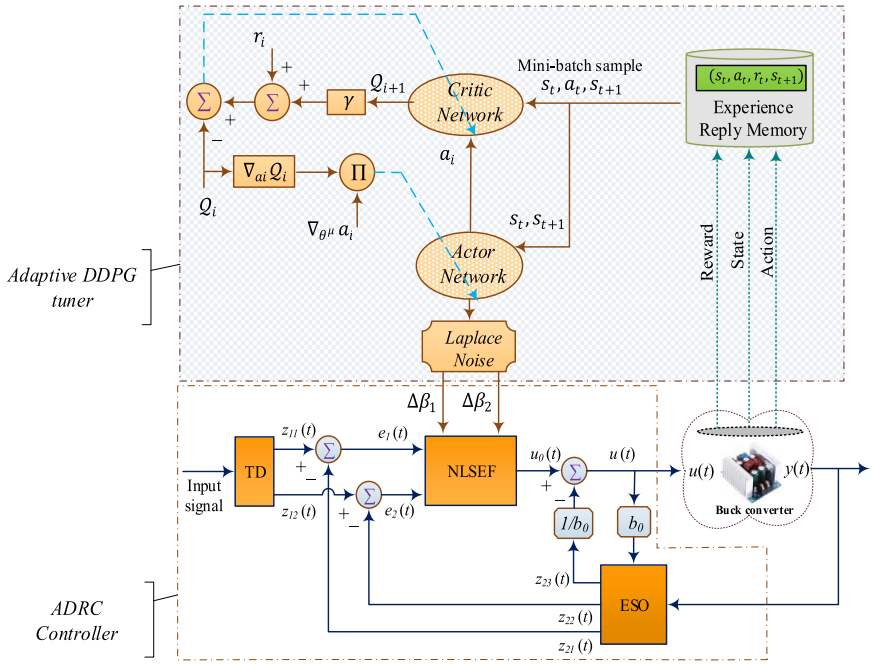

Fig. 6. Illustration of the DDPG adaptive ADRC.

variations. Aiming at boosting the adaptive ability and satisfying the coefficients setting requirements of the ADRC at different times, we introduce the DRL to adjust β_1 and β_2 embedded in NLSEF online. With the analysis mentioned above, the structure of the DDPG adaptive ADRC for the IoT-based converter is depicted in Fig. 6. In the control structure, the DDPG algorithm is adopted as a parameter tuner to provide the regulatory signals for the adaptive setting of the NLSEF gains. The coefficients of the self-adapting ADRC scheme is as follows:

$$\begin{cases} \beta_1 = \beta'_1 + \Delta\beta_1 \\ \beta_2 = \beta'_2 + \Delta\beta_2 \end{cases} \tag{20}$$

where β'_1 and β'_2 are the pretuning coefficients, and $\Delta\beta_1$ and $\Delta\beta_2$ are the DDPG controller's outputs.

In the controller, the optimal policy of the DDPG algorithm is derived based on solving Bellman's equation in the actor–critic architecture. The critic evaluates the quality of the actor policy control for each immediate reward r_t, and according to the critic data, the actor adjusts the NLSEF gains adaptively. The objective of the control scheme is to regulate the output voltage of the concerned converter v_o such that $v_o = v_{\text{ref}}$ with sensing the state variables (i_L and v_o) and output-voltage error. The actor and critic NNs (and their corresponding target NNs) are feedforward NNs with three hidden layers containing 100, 100, and 20 neurons between the input and output layers, as demonstrated in Fig. 7. The input signals to the actor network are a vector state of the v_o, i_L, and e and their derivative in time step t, i.e., $s_t = \{v_o, i_L, e, (\frac{dv_o}{dt}), (\frac{di_L}{dt}), (\frac{de}{dt})\}$.

The control objective for the dc–dc buck converter is to the capacitor voltage v_C (or output voltage v_o) track the reference voltage v_{ref} in the presence of large disturbances. Thus, to design a DDPG-optimized ADRC controller, the reward function

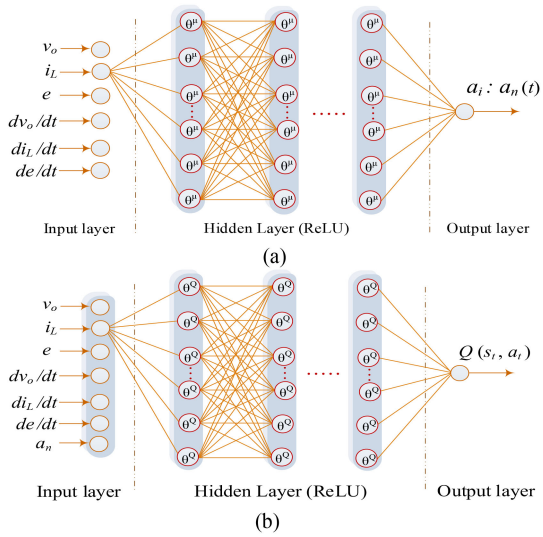

Fig. 7. Illustration of the (a) actor network and (b) critic network.

should be chosen concerning this optimization objective, i.e., the voltage tracking error converges to zero. To achieve this goal, we calculate the reward function r_t based on the sum of the voltage output errors in each step time, as given in

$$r_t = \frac{1}{(v_o - v_{\text{ref}})^2}. \tag{21}$$

Based on the above-mentioned reward signal r_t, the performance of the current DDPG control signals $\{\Delta\beta_1, \Delta\beta_2\}$ is evaluated. When the system disturbances and network degradations happen, r_t will come down with the increase of the voltage output errors, which will result in update the weight

TABLE I
HYPERPARAMETERS SETTINGS OF THE DDPG

Hyper Parameter	Values
Discount factor, γ	0.99
Learning rate, λ	0.0005
Mini-batch size	32
Replay buffer size	20000
Soft target update, τ	0.01

TABLE II
PARAMETERS OF THE BUCK CONVERTER

Parameter	Values
Inductance, L	1 mH
Capacitance, C	1000 μf
Converter input voltage, E	110 V
DC bus voltage reference, V_{ref}	48 V

coefficients of the actor and critic NNs. Therefore, two regulatory signals $\{\Delta\beta_1,\ \Delta\beta_2\}$ are provided to adaptively decrease the effects of both physical and network degradations, i.e., minimize the error between the reference voltage $v_{\rm ref}$ with its actual value v_o. For instance, the actor network takes s_t, then generates two continuous regulatory signals $\{\Delta\beta_1,\ \Delta\beta_2\}$ by the feedforward computation. The critic network receives $\{v_o, i_L,\ e,\ (\frac{dv_o}{dt}),\ (\frac{di_L}{dt}),\ (\frac{de}{dt})\}$ and $\{\Delta\beta_1,\ \Delta\beta_2\}$, and the reward signal r_t is simultaneously derived according to (21). The coefficient weights of the critic network are trained and then the approximate function $\mathcal{Q}(s_t, a_t)$ is outputted. Therefore, the coefficient weights of the DDPG networks are trained for obtaining the regulatory signals to adapt the aforesaid disturbances in the networked dc–dc buck converter. For the implementation of the suggested adaptive tuner, the hyperparameters of the DDPG algorithm are furnished in Table I.

Remark 1: The stability of the suggested data-driven control scheme relies mainly on the original ADRC controller, i.e., without applying the adaptive parameter tuner. Meanwhile, the DDPG tuner provides the regulatory signals such that the DDPG control actions are small, which are not sufficient to lead to the instability.

VI. EXPERIMENTAL RESULTS

This section is devoted to present the different experimental scenarios to validate the data-driven DDPG adaptive ADRC scheme to control the networked dc–dc buck converter. For this purpose, an IoT-based system testbed (as shown in Fig. 4) is built to assess the applicability of the suggested adaptive data-driven scheme from a real-time perspective. The suggested scheme is applied to the IoT-based dc–dc buck converter with the parameters listed in Table II and the experimental outcomes are compared with model predictive control (MPC) [38] and ultralocal model method [39].

To investigate how the suggested scheme deals with the network degradations, we apply the packet loss and interfering traffic to the shared network (IEEE 802.11). In this application, a

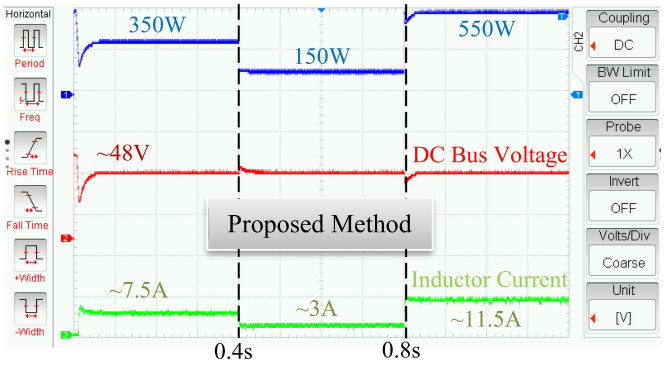

Fig. 8. Transient performance of the DDPG adaptive ADRC controller_ scenario I (the CPL power changes, dc bus voltage, and inductor current are shown with blue, red, and green curves, respectively).

binary two-mode switch is used, which determines the expected packet loss probability. The stochastic factor θ_k of the binary switching is provided by the Bernoulli process [40] with the following definition $\Pr\{\theta_k = 0\} = \alpha_{lp}$, where $\theta_k \epsilon \{0,\ 1\}$, $\theta_k = 0$ means the packet dropout has occurred, $\theta_k = 1$ indicates there is no packet dropout, and $0 < \alpha_{lp} < 1$ is the probability of the packet loss. By applying one interfering node, the random levels of the congestion are imposed on the shared network to study the effect of the network congestion. At each period, the interfering packets are randomly generated and consume $\zeta_{\rm bw}$ percent of the bandwidth.

To provide a good scenario with a great challenge for the suggested controller, the CPL is assumed to be suddenly varied

$$P = \begin{cases} 350\ \text{W} & \text{for } t\ \in [0,\ 0.4)\,\text{s} \\ 150\ \text{W} & \text{for } t\ \in [0.4,\ 0.8)\ \text{s} \\ 550\ \text{W} & \text{for } t\ \in [0.8,\ 1.2)\ \text{s}. \end{cases} \tag{22}$$

The dc–dc buck converter is regulated by the PWM gate drive where the switch drive signal of PWM is produced by comparing the duty ratio signal with the ramp waveform. For the implementation of the DDPG algorithm, the neurons of the hidden layers, which are placed between the input and output layers, are chosen as 100, 100, 20. The gains of the NLSEF are adaptively adjusted by the DDPG algorithm with the initial values of $\beta_1' = 3$ and $\beta_2' = 7$. Under the CPL's power of (22), the actor and critic networks of the DDPG mechanism with the setting of Table I are trained over 250 episodes to generate the regulatory signals $\{\Delta\beta_1,\ \Delta\beta_2\}$. To assess the improved performance of the DDPG-tuned ADRC controller and conduct a fair comparison with state-of-the-art methods, two prevalent schemes, including MPC [38] and Ultralocal model method [39], are adopted.

Scenario I: In the first step, the probability of the packet loss is assumed as $\alpha_{lp} = 50\%$ while $\zeta_{\rm bw} = 0.2$ of the network bandwidth is occupied. The transient responses under the concerning configuration, including the tracking power of the CPL, the voltage of the dc bus, and the current of the inductor for the suggested method, MPC, and ultralocal model are illustrated in Figs. 8–10, respectively.

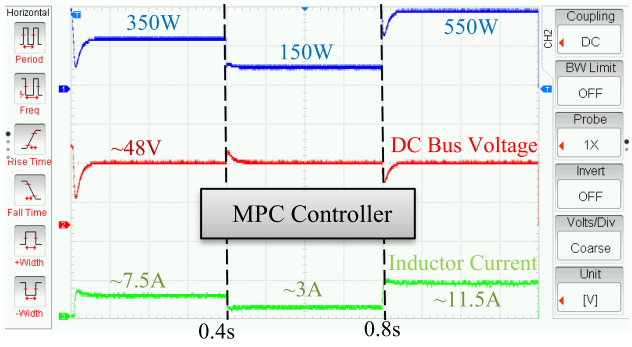

Fig. 9. Transient performance of the MPC controller_ scenario I (the CPL power changes, dc bus voltage, and inductor current are shown with blue, red, and green curves, respectively).

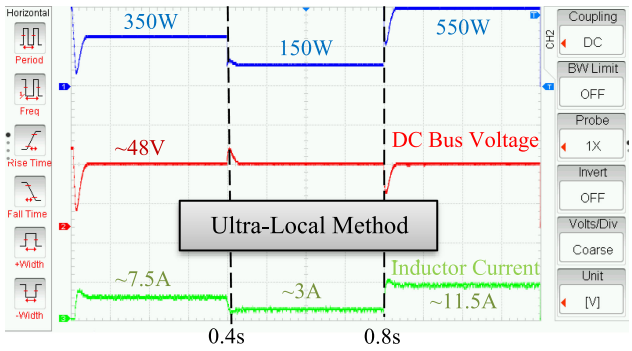

Fig. 10. Transient performance of the ultralocal model controller_ scenario I (the CPL power changes, dc bus voltage, and inductor current are shown with blue, red, and green curves, respectively).

In response to the packet loss and network interfering, one can observe from Fig. 8 that the DDPG adaptive ADRC is robust against the network degradations, as the CPL power, voltage, and inductor current of the converter are properly regulated. In comparison with the MPC and the ultralocal model scheme (see Figs. 9 and 10), the dynamic outcomes of the IoT-based buck converter are regulated faster with lesser deviations, which exhibit a higher level of the reliable operation against the network degradations is yielded by the suggested controller than the conventional strategies. The reason is that the DDPG tuner maintains the design parameters of the ADRC controller to the optimum control point at different times, where the weights in the actor and critic networks are adjusted for the altered operating condition.

Scenario II: In this scenario, a higher level of the packet loss and network interfering are imposed on the communication network to assess the robustness of the designed controllers in a more severe condition of network degradations. To do this, the probability of the packet loss and percentage of the occupied bandwidth is increased to $\alpha_{lp} = 0.8\%$ and $\zeta_{\rm bw} = 0.5$. Figs. 11–13 depict the system outcomes in the scenario, respectively, for the suggested adaptive data-driven controller as well as the MPC and ultralocal model schemes.

Figs. 11–13 reveal how increasing the level of the network degradations adversely affects the system performance, while the IoT-based controllers try to suppress the CPL power, voltage,

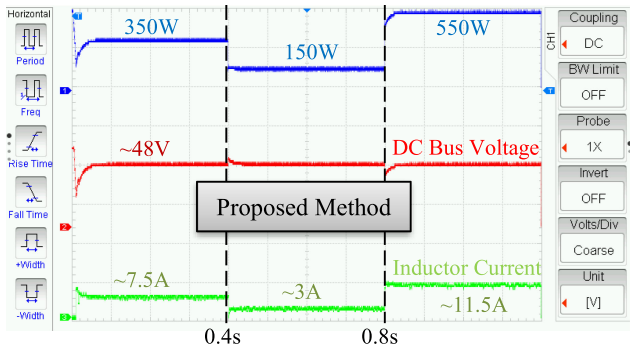

Fig. 11. Transient performance of the DDPG adaptive ADRC controller_ scenario II (the CPL power changes, dc bus voltage, and inductor current are shown with blue, red, and green curves, respectively).

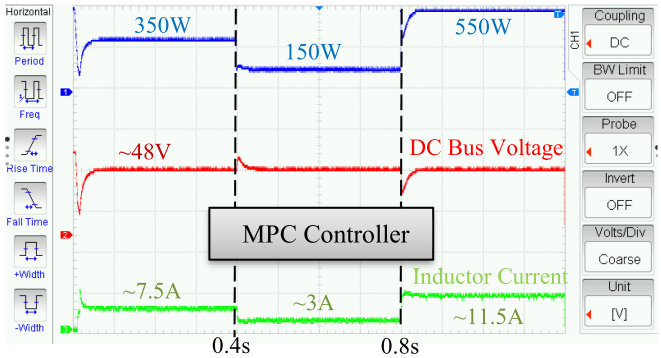

Fig. 12. Transient performance of the MPC controller_ scenario II (the CPL power changes, dc bus voltage, and inductor current are shown with blue, red, and green curves, respectively).

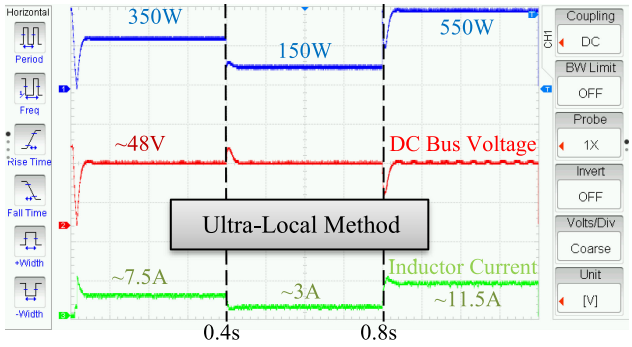

Fig. 13. Transient performance of the ultralocal model controller_ scenario II (the CPL power changes, dc bus voltage, and inductor current are shown with blue, red, and green curves, respectively).

and inductor current deviations. By comparing the outcomes of scenario I with scenario II, it can be observed that the system outputs get deteriorated. It is also observed despite the MPC and ultralocal model schemes have the effect of compensating the steady-state error, their transient dynamic outcomes experience large overshoot, while a higher level of stability is obtained by the suggested controller than the other controllers, i.e., a lower overshoot with the quicker convergence is reached. For completeness, the critical system control specifications, including the maximum output voltage rise/drop of the three control strategies, are given in Table III.

TABLE III
PERFORMANCE INDICES OF VARIOUS CONTROLLERS

Controllers	Scenario I		Scenario II	
	MOVR	MOVD	MOVR	MOVD
Ultra-Local Method	10.9584	34.8474	11.4371	47.9061
MPC Controller	7.7405	26.9042	8.1513	32.9058
Proposed Method	3.6326	21.4834	6.3175	24.0719

MOVR: Maximum output voltage raise; MOVD: Maximum output voltage drop.

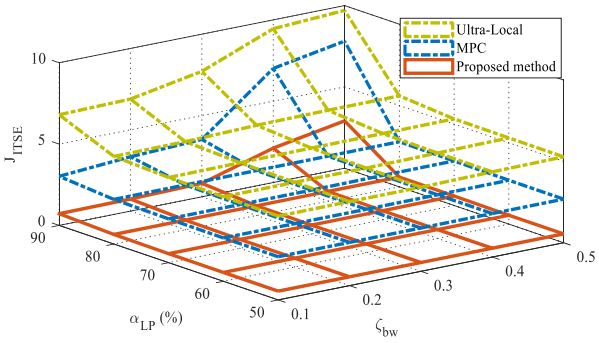

Fig. 14. Robustness analysis against network degradations.

Besides, the robustness of the suggested data-driven method against the various levels of the network degradations are assessed, where α_{lp} is varied within the range of 50%–90% and $\zeta_{\rm bw}$ is varied within the range of 0.1–0.5. The values of the integral time square error corresponding to the various network degradations are presented in Fig. 14. The experimental outcomes of Fig. 14 reveal the less sensitivity of DDPG adaptive ADRC than the state-of-the-art controllers and high efficiency when the suggested data-driven scheme face to the unreliable factors applied to the shared network.

VII. DISCUSSION

The IoT technology offers an open platform that supports the standard protocols and communication interfaces to provide new applications and services. However, the shared networks in the IoT-based systems are subjected to various communication degradations, which can deteriorate the control performance. The stability of the IoT-based dc power electronic converters are intensified when the CPLs are connected to the dc bus. In response to the control challenges of the IoT-based systems, a DDPG adaptive ADRC approach is developed to stabilize the voltage term of an IoT-based dc–dc converter with time-varying CPLs. The main outcomes can be stated as follows.

1) In comparison with the low-order ADRC, which is effective only against small uncertainties, the suggested method can ensure the required control features, i.e. stability, in the presence of the high level of network degradations.
2) In comparison with the high-order ADRC, the suggested method is remarkably less complex, which is very valuable from the practical perspective. The reason is that the high-order ADRC approaches need to compensate for the network degradations using predictor strategies (e.g., smith predictor), thus increasing the design complexity.
3) In comparison with the model-based methodologies (e.g., sliding-mode control and roust nonfragile), the suggested method does not need to the model identification as it is developed based on the I/O measurements.
4) In comparison with the ADRC optimized by the heuristic methodologies (e.g., ant colony optimization), the suggested method can provide optimal performance at whole cycle periods due to its online learning ability.

VIII. CONCLUSION

In this article, the application of the IoT communication is investigated as the key instrument for transferring from the conventional existing grid to the SG. In particular, this article focused on an IoT-based dc–dc buck converter, which allows studying the destructive effect of CPLs in the IoT-grid framework. To promote the quality of the control actions by DRL, this work proposes an adaptive ADRC method for the voltage regulation of networked dc–dc buck converters feedings CPLs. In this way, a DDPG algorithm with the actor–critic framework is employed for the online coefficient setting of the NLSEF at different times. The actor and critic networks are trained using a significant number of the networked converter data and then they can automatically adjust the NLSEF gains for the adaptive compensation control. From the experimental outcomes, it is revealed that as compared with the ultralocal model control and MPC, the suggested method provides a high level of stability for the IoT-based converter when the system is faced with severe packet loss and interfering traffic.

As future work, the effect of both time delay and packet loss on the system can be investigated simultaneously. To ameliorate the operating efficiency, the predictive schemes can be adapted to compensate for the network degradation while the control gains can be optimally designed by DRL.

REFERENCES

[1] M. Wu and D. D.-C. Lu, "A novel stabilization method of LC input filter with constant power loads without load performance compromise in dc microgrids," *IEEE Trans. Ind. Electron.*, vol. 62, no. 7, pp. 4552–4562, Jul. 2015.
[2] L. Herrera, W. Zhang, and J. Wang, "Stability analysis and controller design of dc microgrids with constant power loads," *IEEE Trans. Smart Grid*, vol. 8, no. 2, pp. 881–888, Mar. 2017.
[3] J. Hou, Q. Chen, S.-C. Wong, C. K. Tse, and X. Ruan, "Analysis and control of series/series-parallel compensated resonant converter for contactless power transfer," *IEEE J. Emerg. Sel. Topics Power Electron.*, vol. 3, no. 1, pp. 124–136, Mar. 2015.
[4] S. R. Huddy and J. D. Skufca, "Amplitude death solutions for stabilization of dc microgrids with instantaneous constant-power loads," *IEEE Trans. Power Electron.*, vol. 28, no. 1, pp. 247–253, Jan. 2013.
[5] M. Hajihosseini, M. Andalibi, M. Gheisarnejad, H. Farsizadeh, and M.-H. Khooban, "DC/DC power converter control-based deep machine learning techniques: Real-time implementation," *IEEE Trans. Power Electron.*, 2020.
[6] M. H. Khooban, M. Gheisarnejad, H. Farsizadeh, A. Masoudian, and J. Boudjadar, "A new intelligent hybrid control approach for dc–dc converters in zero-emission ferry ships," *IEEE Trans. Power Electron.*, vol. 35, no. 6, pp. 5832–5841, Jun. 2020.
[7] Q. Xu, C. Zhang, C. Wen, and P. Wang, "A novel composite nonlinear controller for stabilization of constant power load in dc microgrid," *IEEE Trans. Smart Grid*, vol. 10, no. 1, pp. 752–761, Jan. 2019.
[8] S. Singh, D. Fulwani, and V. Kumar, "Robust sliding-mode control of dc/dc boost converter feeding a constant power load," *IET Power Electron.*, vol. 8, pp. 1230–1237, 2015.

[9] N. Vafamand, M. H. Khooban, T. Dragicevic, F. Blaabjerg, and J. Boudjadar, "Robust non-fragile fuzzy control of uncertain DC microgrids feeding constant power loads," *IEEE Trans. Power Electron.*, vol. 34, no. 11, pp. 11300–11308, Nov. 2019.
[10] R.-C. Roman, R.-E. Precup, C.-A. Bojan-Dragos, and A.-I. Szedlak-Stinean, "Combined model-free adaptive control with fuzzy component by virtual reference feedback tuning for tower crane systems," *Procedia Comput. Sci.*, vol. 162, pp. 267–274, 2019.
[11] K. K. Tan, S. Zhao, and J.-X. Xu, "Online automatic tuning of a proportional integral derivative controller based on an iterative learning control approach," *IET Control Theory Appl.*, vol. 1, pp. 90–96, 2007.
[12] J. Sun, J. Yang, S. Li, and W. X. Zheng, "Sampled-data-based event-triggered active disturbance rejection control for disturbed systems in networked environment," *IEEE Trans. Cybern.*, vol. 49, no. 2, pp. 556–566, Feb. 2019.
[13] M.-B. Rădac, R.-E. Precup, E. M. Petriu, and S. Preitl, "Iterative data-driven tuning of controllers for nonlinear systems with constraints," *IEEE Trans. Ind. Electron.*, vol. 61, no. 11, pp. 6360–6368, Nov. 2014.
[14] S. Preitl, R.-E. Precup, Z. Preitl, S. Vaivoda, S. Kilyeni, and J. K. Tar, "Iterative feedback and learning control. Servo systems applications," *IFAC Proc. Vol.*, vol. 40, pp. 16–27, 2007.
[15] A. Luchetta *et al.*, "MLMVNNN for parameter fault detection in PWM dc–dc converters and its applications for buck and boost dc–dc converters," *IEEE Trans. Instrum. Meas.*, vol. 68, no. 2, pp. 439–449, Feb. 2019.
[16] Z. Yan and Y. Xu, "Data-driven load frequency control for stochastic power systems: A deep reinforcement learning method with continuous action search," *IEEE Trans. Power Syst.*, vol. 34, no. 2, pp. 1653–1656, Mar. 2019.
[17] N. Vafamand, M. H. Khooban, T. Dragičević, and F. Blaabjerg, "Networked fuzzy predictive control of power buffers for dynamic stabilization of dc microgrids," *IEEE Trans. Ind. Electron.*, vol. 66, no. 2, pp. 1356–1362, Feb. 2019.
[18] S. Mondal and R. Paily, "Efficient solar power management system for self-powered IoT node," *IEEE Trans. Circuits Syst. I, Regular Papers*, vol. 64, no. 9, pp. 2359–2369, Sep. 2017.
[19] M. Rizzi, P. Ferrari, A. Flammini, and E. Sisinni, "Evaluation of the IoT LoRaWAN solution for distributed measurement applications," *IEEE Trans. Instrum. Meas.*, vol. 66, no. 12, pp. 3340–3349, Dec. 2017.
[20] P. Ferrari, A. Flammini, E. Sisinni, S. Rinaldi, D. Brandão, and M. S. Rocha, "Delay estimation of industrial IoT applications based on messaging protocols," *IEEE Trans. Instrum. Meas.*, vol. 67, no. 9, pp. 2188–2199, Sep. 2018.
[21] V. Tanyingyong, R. Olsson, M. Hidell, P. Sjödin, and B. Ahlgren, "Implementation and deployment of an outdoor IoT-based air quality monitoring testbed," in *Proc. IEEE Global Commun. Conf.*, Abu Dhabi, UAE, 2018, pp. 206–212.
[22] A. Betzler, C. Gomez, I. Demirkol, and J. Paradells, "CoAP congestion control for the internet of things," *IEEE Commun. Mag.*, vol. 54, no. 7, pp. 154–160, Jul. 2016.
[23] V. Tanyingyong, R. Olsson, M. Hidell, P. Sjödin, and B. Pehrson, "Design and implementation of an IoT-controlled dc-dc converter," in *Proc. Sustain. Internet ICT Sustain.*, 2013, pp. 1–4.
[24] S. Bendel, T. Springer, D. Schuster, A. Schill, R. Ackermann, and M. Ameling, "A service infrastructure for the internet of things based on XMPP," in *Proc. IEEE Int. Conf. Pervasive Comput. Commun. Workshops*, 2013, pp. 385–388.
[25] V. Tanyingyong, R. Olsson, J.-W. Cho, M. Hidell, and P. Sjodin, "IoT-grid: IoT communication for smart dc grids," in *Proc. IEEE Global Commun. Conf.*, 2016, pp. 1–7.
[26] R. Sukjaimuk, Q. Nguyen, and T. Sato, "A smart congestion control mechanism for the green IoT sensor-enabled information-centric networking," *Sensors*, vol. 18, 2018, Art. no. 2889.
[27] Y. Huang, J. Wang, D. Shi, J. Wu, and L. Shi, "Event-triggered sampled-data control: An active disturbance rejection approach," *IEEE/ASME Trans. Mechatronics*, vol. 24, no. 5, pp. 2052–2063, Oct. 2019.
[28] W. Han, G. Wang, and A. M. Stankovic, "Active disturbance rejection control in fully distributed automatic generation control with co-simulation of communication delay," *Control Eng. Pract.*, vol. 85, pp. 225–234, 2019.
[29] Y. Yu, Y. Yuan, H. Yang, and H. Liu, "Nonlinear sampled-data ESO-based active disturbance rejection control for networked control systems with actuator saturation," *Nonlinear Dyn.*, vol. 95, pp. 1415–1434, 2019.
[30] S. Chen, W. Xue, Y. Huang, and P. Liu, "On comparison between smith predictor and predictor observer based ADRCs for nonlinear uncertain systems with output delay," in *Proc. Amer. Control Conf.*, 2017, pp. 5083–5088.
[31] Z. Yin, C. Du, J. Liu, X. Sun, and Y. Zhong, "Research on autodisturbance-rejection control of induction motors based on an ant colony optimization algorithm," *IEEE Trans. Ind. Electron.*, vol. 65, no. 4, pp. 3077–3094, Apr. 2018.
[32] Y. Wang, J. Sun, H. He, and C. Sun, "Deterministic policy gradient with integral compensator for robust quadrotor control," *IEEE Trans. Syst., Man, Cybern., Syst.*, 2019.
[33] Y. Wu, H. Tan, J. Peng, H. Zhang, and H. He, "Deep reinforcement learning of energy management with continuous control strategy and traffic information for a series-parallel plug-in hybrid electric bus," *Appl. Energy*, vol. 247, pp. 454–466, 2019.
[34] M. H. Khooban and M. Gheisarnejad, "A novel deep reinforcement learning controller based type-II fuzzy system: Frequency regulation in microgrids," *IEEE Trans. Emerg. Topics Comput. Intell.*, 2020.
[35] J. Wu and Y. Lu, "Feedback linearization adaptive control for a buck converter with constant power loads," in *Proc. IEEE Int. Power Electron. Appl. Conf. Expo.*, 2018, pp. 1–6.
[36] B. Gao, J. Shao, and X. Yang, "A compound control strategy combining velocity compensation with ADRC of electro-hydraulic position servo control system," *ISA Trans.*, vol. 53, pp. 1910–1918, 2014.
[37] M. Zhu, X. Wang, and Y. Wang, "Human-like autonomous car-following model with deep reinforcement learning," *Transp. Res. C, Emerg. Technol.*, vol. 97, pp. 348–368, 2018.
[38] J. Pahasa and I. Ngamroo, "Coordinated control of wind turbine blade pitch angle and PHEVs using MPCs for load frequency control of microgrid," *IEEE Syst. J.*, vol. 10, no. 1, pp. 97–105, Mar. 2016.
[39] H. Ahmed, I. Salgado, and H. Ríos, "Robust synchronization of master-slave chaotic systems using approximate model: An experimental study," *ISA Trans.*, vol. 73, pp. 141–146, 2018.
[40] Q. Shafiee, Č. Stefanović, T. Dragičević, P. Popovski, J. C. Vasquez, and J. M. Guerrero, "Robust networked control scheme for distributed secondary control of islanded microgrids," *IEEE Trans. Ind. Electron.*, vol. 61, no. 10, pp. 5363–5374, Oct. 2014.

Meysam Gheisarnejad received the B.Sc. degree in electronic engineering from the Azad University of Lahijan, Lahijan, Iran, in 2009, and the M.Sc. degree in control engineering from the Azad University of Najafabad, Najafabad, Iran, in 2013.

His research interests include power system dynamics and control, shipboard microgrid, cyber-physical microgrid, power electronics, and renewable energy systems.

Mohammad Hassan Khooban (Senior Member, IEEE) received the Ph.D. degree in power systems and electronics from the Shiraz University of Technology, Shiraz, Iran, in 2017.

From 2016 to 2017, he was a Research Assistant with the University of Aalborg, Aalborg, Denmark, where he conducted research on the advanced control of microgrids and marine power systems. From 2017 to 2018, he was a Postdoctoral Associate with Aalborg University, where he was a Postdoctoral Research Assistant from 2019 to 2020. He is currently an Assistant Professor at Aarhus University. He has authored or coauthored more than 150 publications on journals and international conferences, one book chapter, and holds one patent. His current research interests include control theory and application, power electronics, and its applications in power systems, industrial electronics, and renewable energy systems.

Dr. Khooban is currently a Guest Editor/Associate Editor for the Complexity Journal and the IEEE JOURNAL OF EMERGING AND SELECTED TOPICS IN POWER ELECTRONICS. He also serves extensively as a reviewer for various IEEE/IET transactions and journals on power electronics, circuits, and control engineering. He was nominated by Thomson Reuters to be the World's Top 1% Researchers in Engineering, in 2019.