Nonlinear Constrained Optimal Control of Wave Energy Converters with Adaptive Dynamic Programming

Jing Na, *Member, IEEE,* Bin Wang, Guang Li, *Member, IEEE,* Siyuan Zhan, and Wei He
Senior Member, IEEE

Abstract—In this paper, we address the energy maximization problem of wave energy converters (WEC) subject to nonlinearities and constraints, and present an efficient online control strategy based on the principle of adaptive dynamic programming (ADP) for solving the associated Hamilton-Jacobi-Bellman (HJB) equation. To solve the derived constrained nonlinear optimal control problem, a critic neural network (NN) is used to approximate the time-dependant optimal cost value and then calculate the practical suboptimal causal control action. The proposed novel WEC control strategy leads to a simplified ADP framework without involving the widely used actor NN. The significantly improved computational efficacy of the proposed control makes it attractive for its practical implementation on a WEC to achieve a reduced unit cost of energy output, which is especially important when the dynamics of a WEC are complicated and need to be described accurately by a high-order model with nonlinearities and constraints. Simulation results are provided to show the efficacy of the proposed control method.

Index Terms—Wave energy converters, adaptive dynamic programming, constrained optimal control.

I. INTRODUCTION

SEA waves provide a promising renewable energy resource for wave energy, which can be potentially harnessed. However, despite decades of research efforts, wave energy is still far from being mature for commercialization due to higher unit cost of generated electricity than other renewable energies, e.g., solar and wind [1]. Reducing the unit cost of energy output of a wave energy converter (WEC) relies on both good device design and an efficient control strategy. It has been recognized that development of advanced control strategies for WECs is one of the most promising cost-reduction pathways. Early conventional WEC control methods,

Jing Na and Bin Wang are with Faculty of Mechanical and Electrical Engineering, Kunming University of Science and Technology, Kunming, 650500, China. (E-mail: najing25@163.com; wangbin2173@126.com).

Guang Li and Siyuan Zhan are with the School of Engineering and Materials Science, Queen Mary University of London, London, UK, E1 4NS, (Email: g.li@qmul.ac.uk, s.zhan@qmul.ac.uk).

Wei He is with the School of Automation and Electrical Engineering, University of Science and Technology Beijing, Beijing, China,(Email:weihe@ieee.org).

e.g., latching control [2] and declutching control, are mainly based on the impedance matching principle, that is, maximum energy can be harvested when the resonance frequency of the WEC matches the dominant frequency of the incoming waves [3]. These control methods can be easily implemented in idealized regular waves but may become complicated for its implementation in real irregular waves.

The control of WEC is essentially an energy maximization problem subject to the wave excitation forces as the persistent disturbances, which is essentially different from traditional optimal control for reference tracking or regulation problems. Hence, a new branch of advanced optimal control strategies need to be developed to resolve the WEC control problem. Specifically, model predictive control (MPC) has been tailored for WEC systems in recent years [4]–[8]. MPC has the ability to explicitly incorporate constraints into the WEC control problem and utilize wave prediction information to satisfy the non-causal optimal control requirement [3]. It is shown that MPC-based control can double the energy output compared with the conventional WEC control strategies [7]. However, one critical problem of implementing the MPC strategy is its heavy online computational burden since a direct maximization control target of the WEC control can result in a non-convex optimization problem. This problem can become numerically intractable if the dynamics of a WEC are complicated and need to be described accurately by a high-order model with nonlinearities to retain the modeling fidelity and constraints for safe operations. To address this issue, some alternative methods have been proposed, such as control based on convex optimization with a modified objective function [9], adaptive control [10], nonlinear MPC with pseudospectral control [11], and nonlinear MPC based on a combination of the pseudospectral method and the differential flatness [12], where the plant model is assumed to be accurate.

To handle modeling uncertainties and nonlinearities in the control systems, function approximations (e.g., neural networks (NN) or fuzzy logic systems (FLS)) have been adopted in the adaptive control designs [13]–[15]. However, these adaptive control approaches cannot be directly used for WEC systems since they cannot solve the energy maximization problem. To achieve optimal performance requirement, the idea of reinforcement learning (RL) [16] has been used in the optimal control design, which leads to a new online optimal control design approach, named adaptive dynamic programming (ADP). In this method, NNs are trained to approximate

the solution of the Hamilton-Jacobi-Bellman (HJB) equation and the required control actions [17]. The well-known actor-critic based ADP framework [18] has been initially developed for discrete-time optimal control [19], [20]. Substantial work has also been reported for continuous-time systems [21] with unknown nonlinearities [22], control constraints [23] and tracking command [24], [25]. However, these available ADP methods have been proposed to solve regulation or tracking problems only, and thus they cannot be directly used to solve the energy maximization of WECs [26].

In this paper, we propose a novel fast adaptive suboptimal causal control strategy, which can efficiently tackle the constrained WEC optimal control problem without involving the wave prediction information. Although various WEC devices have been reported, e.g., [27], [28], we will use a benchmark point absorber type of WEC as described in [12] for demonstration and comparison purpose in this paper, where only the vertical heave motion dynamics are considered. To develop an online suboptimal causal control for nonlinear WECs without using offline learning and wave prediction, this paper exploits the applicability of ADP to address the WEC control problem, of which the control objective is to maximize the output energy [8], while constraints imposed on the control input and system states are guaranteed. We firstly reformulate the WEC control as a constrained energy maximisation control problem, where the input constraint is explicitly incorporated into a modified cost function and can be strictly guaranteed. This is a nontrivial advancement compared with the recent work [26], and directly contributes to the reduction of the levelized cost of electricity (LCOE). As a commonly encountered problem in many control applications, actuator saturation is an important issue which needs to be accounted for in power take-off (PTO) design and the hardware selection for feedback control implementation in a WEC control system. The control input constraints are associated with the limits of the components such as the torque limits of hydraulic actuators and hydraulic synchro-generators used in the hydraulic PTO, the torque and excursion limits of the linear generators of a electric PTO and the current/voltage limits of the converters. It is not only critical for safety but also a key factor influencing the trade-off between hardware cost and control performance. A control system fully incorporating the hardware constraints into the controller design can normally outperform its counterpart controlled by an unconstrained controller. Moreover, the proposed control can also tackle the nonlinearities of the WEC dynamics, which cannot be effectively handled by the approach of [26]. The nonlinear effects of the wave-structure interactions can become prominent for large waves. Ignoring these nonlinear effects in the WEC control design can lead to performance degradation. The proposed approach can effectively tackle such nonlinear effects using a lumped term, which helps to reduce the model order and thus the computational burden compared to the approach of [26]. A direct comparison of the proposed constrained control strategy based on ADP with that of [26] using simulations demonstrates the advantages of the proposed one.

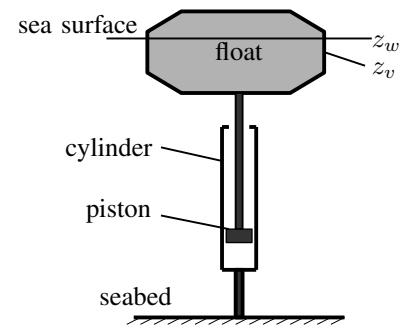The main contributions of this paper can be summarized as follows:

Fig. 1. Structure diagram of the point absorber [12]

1) The cost function adopted in [26] has been further modified to explicitly address the control input constraint. The subsequent ADP synthesis has been also tailored such that the input constraint can be strictly fulfilled in this paper, which helps retain safe operations of WEC and generates more reliable energy output in practice.
2) The nonlinear wave excitation force dynamics are considered directly in the adopted WEC model, such that the order of the WEC model used for the control design and implementation can be reduced compared with [26]. Hence, this ADP based control is more computationally efficient.

The paper is organized as follows. The modeling of a typical WEC and the problem formulation are presented in Section II. The design of online constrained optimal control with ADP is introduced in Section III. Simulation results are given in Section IV. The paper is concluded in Section V.

II. WEC MODELING AND CONTROL FORMULATION

The point absorber to be controlled is shown in Fig. 1. For completeness, we first briefly introduce the nonlinear dynamics and modeling of this WEC plant. The point absorber has a float on the sea surface with a variational radius. Note that in the WEC studied in this paper (Fig.1) the heave motion of the float drives a hydraulic system for power generation (see e.g., [29]–[33] for more details about the power generation subsystem with hydraulic generators or linear generators), where a hydraulic cylinder with a piston inside it is vertically installed below the float; the other end of the cylinder is attached to the seabed or anti-heave plate with negligible heave motion relative to the float. The aim of this paper is to propose a new ADP control design methodology to solve the WEC control problem. Hence, we are only concerned with the control of the float vertical motion which can be influenced by manipulated forces acting on the float, while the modeling and control of the PTO mechanisms are not explicitly considered; see [34] and the references therein for more details on the wave-to-wire modeling issues.

We define z_w and z_v as the sea surface level and the height of the float, f_u as the control input (the force acting on the float). It is noted that the motion velocity of the float $v = \dot{z}_v$ can drive the energy generator, such that the extracted energy over a period $[0, T]$ is

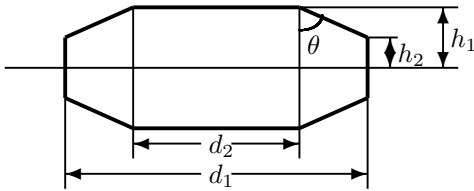$$ E = - \int_0^T f_u v \, dt \qquad (1) $$

Fig. 2. The structure of the float for the calculation of nonlinear hydraulic stiffness [12].

Two constraints have to be considered in the WEC control designs to guarantee safe operation of WECs: first, the relative motion between the float and the sea surface cannot be too large to avoid device damage, and thus the first constraint is given by

$$|z_w - z_v| \leq z_{\max}. \tag{2}$$

for a given constant $z_{\max}$. Second, the required control effort on the device must be within the actuator saturation. Hence, the other constraint is on the required control force f_u as

$$|u| = |f_u| \leq u_{\max}. \tag{3}$$

for a given constant $u_{\max}$. The control objective is to maximize the energy (1) converted from waves subject to the constraints (2) and (3).

The dynamic equation of the float motion is given by

$$m_s \ddot{z}_v = f_s - f_r - f_f + f_u \tag{4}$$

where m_s is the mass of the float, f_s is the buoyancy force, f_r is the radiation force, and f_f is the friction force.

The radiation force is calculated by the Cummins' equation

$$f_r = m_\infty \ddot{z}_v + \int_{-\infty}^{t} h_r(\tau)[\dot{z}_v(t-\tau) - \dot{z}_w(t-\tau)]d\tau \tag{5}$$

where m_∞ is the added mass when the frequency approaches the infinite frequency.

The friction force is proportional to the heave velocity as

$$f_f = D_f \dot{z}_v \tag{6}$$

with D_f as the friction coefficient.

The nonlinearity considered in this WEC model comes from the buoyancy force owing to the nonconstant cross sectional area of the float. The nonlinear function is

$$f_s = \kappa(z) = \begin{cases} \kappa_0 h_2 + \kappa_n(z), & \text{if } h_2 < z \leq h_1; \\ \kappa_0 z, & \text{if } |z| \leq h_2; \\ -\kappa_0 h_2 - \kappa_n(-z), & \text{if } -h_1 \leq z < -h_2. \end{cases} \tag{7}$$

where $z := z_w - z_v$, $\kappa_0 := \pi \rho g d_1^2/4$, with ρ as the density of sea water, g as the standard gravity and

$$\kappa_n(z) = \frac{\pi g \rho}{4}[d_1^2(z - h_2) - (z - h_2)^2 d_1 \xi \\ + (z - h_2)^3 \xi^2/3] \tag{8}$$

with

$$\xi := 2 \tan \theta = \frac{d_1 - d_2}{h_1 - h_2}.$$

The other notations d_1, d_2, h_1, h_2 and θ are shown in Fig. 2.

Substituting (5), (6) and (7) into (4), it follows

$$m\ddot{z}_v = \kappa(z) - \int_{-\infty}^{t} h_r(\tau)[\dot{z}_v(t-\tau) - \dot{z}_w(t-\tau)]d\tau - D_f \dot{z}_v + f_u \tag{9}$$

where $m = m_s + m_\infty$ is the lumped mass.

The convolution kernel $h_r(\tau)$ can be obtained from mature fluid dynamics software packages, which can be further described by a state-space model

$$\dot{x}_r = A_r x_r + B_r(\dot{z}_v - \dot{z}_w) \tag{10a}$$
$$y_r = C_r x_r \tag{10b}$$

with $y_r := \int_{-\infty}^{t} h_r(\tau)[\dot{z}_v(t - \tau) - \dot{z}_w(t - \tau)]d\tau$.

By choosing the system state vector as $x = [x_1, x_2, x_r]^\top \in \mathbb{R}^n$, the state-space model of the WEC system (4) is given by

$$\dot{x}(t) = Ax(t) + B_u u(t) + B_w w(t) \tag{11a}$$
$$y(t) = C_y x(t) \tag{11b}$$

with $x_1 := z_w - z_v$, $x_2 := \dot{z}_v$, $u := f_u$, $w := \dot{z}_w$, $y := \dot{z}_v$ and

$$A = \begin{bmatrix} 0 & -1 & \mathbf{0}_{1 \times n_r} \\ \kappa(\cdot)/m & -D_f/m & -C_r/m \\ \mathbf{0}_{n_r \times 1} & B_r & A_r \end{bmatrix}, \quad B_u = \begin{bmatrix} 0 \\ 1/m \\ 0 \end{bmatrix},$$

$$B_w = \begin{bmatrix} 1 & 0 & -B_r \end{bmatrix}, \quad C_y = \begin{bmatrix} 0 & 1 & \mathbf{0}_{1 \times n_r} \end{bmatrix}.$$

Note that (11) describes a more realistic nonlinear WEC plant than its simplified, linear counterpart studied in the previous literature [7], [9], which assumes a constant float radius. Moreover, since the nonlinear term $\kappa(z)$ denoting the buoyancy force dynamics is included in the matrix A directly, the WEC model (11) in this paper has lower order than the full-order WEC model studied in [26], where buoyancy force dynamics have to be represented in an extra state-space model and induced into the whole WEC model. This property may make other optimal control designs (e.g., linear MPC) infeasible.

The aim of WEC control design is thus to solve the following constrained optimization problem:

$$\max_{u(t)} \int_0^T -u(t)y(t)dt$$
$$\text{s.t. } \dot{x}(t) = Ax(t) + B_u u(t) + B_w w(t) \tag{12}$$
$$|x_1| \leq z_{\max}, \quad |u(t)| \leq u_{\max}, \quad \forall t \in [t_0, T]$$

In this paper, we aim to resolve the constrained optimal WEC control problem (12) using ADP, where the wave prediction algorithm is not needed. Hence, compared with other control strategies (e.g., MPC), the ADP-based control to be presented provides a suboptimal causal control solution. In some WEC operation scenarios, it is recognized that the loss of energy with a suboptimal causal control can be trivial [35], especially when compared to the obvious benefit of computation load reduction and the avoidance of the cost and maintenance of wave prediction hardware.

III. CONSTRAINED OPTIMAL CONTROL DESIGN VIA ADP

In this section, the optimization problem (12) is first reformulated to explicitly address the constraints on the heave mo-

tion x_1 in (2) and the control input u in (3), and then resolved using the principle of ADP. Then the online implementation will be investigated by developing a recently proposed adaptive law to update the critic NN weight.

A. Optimal control design for WECs with constraints

To guarantee safe operation of a WEC, the constraints imposed on the output x_1 in (2) and the input u in (3) should be fulfilled. Hence, these factors must be considered in the cost function used to synthesis the control. For this purpose, the following modified cost function is introduced:

$$V(x,t) = \int_t^T \left(x_2(\tau)u(\tau) + \frac{\epsilon}{z_{max} - |x_1(\tau)|} + U(u(\tau)) \right) d\tau \tag{13}$$

where $\varepsilon > 0$ is a small constant, z_{max} denotes the constraint on the state x_1. $U(u(\tau))$ is a positive definite function of the control input u, which is set as a nonsquare function as [23]

$$U(u(\tau)) = 2 \int_0^u u_{max} \tanh^{-1} \left(\frac{v}{u_{max}} \right) R dv \tag{14}$$

where $\tanh(\cdot)$ is the hyperbolic function, u_{max} is the saturation bound of u. $R > 0$ is a positive constant to make a trade-off between the control action and energy output (e.g., a large R can reduce the amplitude of control u).

Remark 1. *In the modified cost function (13) with (14), the first term represents the extracted energy output. It is clear that the maximization of the cost function in (12) is equivalent to the minimization of the first term in the modified cost function (13). The second term is a barrier function of x_1, which is used to address the constraint (2) as [7]. Since $\frac{\varepsilon}{z_{max}-|x_1|} \to \infty$ holds for $|x_1| \to z_{max}$, minimizing this term with a small bounded constant ϵ can prevent the system output x_1 from approaching the boundary defined in (2). The final term is a nonquadratic function of u to constrain the control input [23], which is different to the cost function with a quadratic term $u^\top R u$ adopted in [26]. Hence, as shown in the subsequent developments, the derived control u based on this further modified cost function can strictly guarantee the constraint (3). Thus, the original optimal control problem (12) is converted to the minimization of cost function (13) subject to system (11).*

To obtain the control action, the optimization problem with the cost function (13) and system (11) needs to be solved. This constrained optimal control can be solved by the dynamic programming (DP) algorithm [7], where heavy computational costs may be problematic for practical application, in particular when the WEC is described by a high-order model. In the following, we present an efficient solution using the Hamiltonian method, which needs a Hamiltonian defined as follows [17]

$$H(x,u,V,t) = V_x^\top (Ax + B_w w + B_u u) + x_2 u + \frac{\varepsilon}{z_{max} - |x_1|} + U(u) \tag{15}$$

where $V_x := \partial V(x,t)/\partial x$ is the partial derivative of $V(x,t)$ in (13) with respect to x.

Denote $V^*(x,t)$ as the optimal cost function of the optimal control u^*, which is given by

$$V^*(x,t) = \min_u \int_t^T \left(x_2 u^* + \frac{\epsilon}{z_{max} - |x_1|} + U(u^*) \right) d\tau \tag{16}$$

Then based on the optimal control theory [17], we know that the optimal control u^* satisfies the following HJB equation:

$$\begin{aligned} -V_t^* &= \min_u H(x, u^*, V^*(x,t), t) \\ &= V_x^{*T}(Ax + B_w w + B_u u^*) + x_2 u^* + \frac{\varepsilon}{z_{max} - |x_1|} \\ &\quad + 2 \int_0^{u^*} u_{max} \tanh^{-1} \left(\frac{v}{u_{max}} \right) R dv \end{aligned} \tag{17}$$

where $V_t^* := \partial V^*(x,t)/\partial t$ and $V_x^* := \partial V^*(x,t)/\partial x$.

Then according to the stationary condition [17], we can solve $\partial H(x, u^*, V^*)/\partial u^* = 0$ for the suboptimal causal control action u^* given as

$$\begin{aligned} u^* &= -u_{max} \tanh \left[\frac{1}{2Ru_{max}}(x_2 + B_u^\top V_x^*) \right] \\ &= -u_{max} \tanh (\Psi) \end{aligned} \tag{18}$$

where $\Psi = \frac{1}{2Ru_{max}}(x_2 + B_u^\top V_x^*)$.

Clearly, the obtained control of (18) is different to that in [26] and can guarantee the satisfaction of the saturation $|u| \le u_{max}$ for all $t > 0$. Considering the property of hyperbolic function and substituting (18) into (14), we have

$$\begin{aligned} U(u^*) &= 2 \int_0^{u^*} u_{max} \tanh^{-1} \left(\frac{v}{u_{max}} \right) R dv \\ &= 2 u_{max} \tanh^{-1}(u^*/u_{max}) R u^* \\ &\quad + u_{max}^2 R \ln \left[1 - (u^*/u_{max})^2 \right] \\ &= 2 u_{max} \Psi R u_{max} \tanh(\Psi) \\ &\quad + u_{max}^2 R \ln \left[1 - \tanh^2(\Psi) \right] \\ &= u_{max}(x_2 + B_u^\top V_x^*) \tanh(\Psi) \\ &\quad + u_{max}^2 R \ln \left[1 - \tanh^2(\Psi) \right]. \end{aligned} \tag{19}$$

We further substitute (19) into HJB equation (17), so that

$$\begin{aligned} -V_t^* &= V_x^{*\top}(Ax + B_w w) - (V_x^{*\top} B_u + x_2) u_{max} \tanh(\Psi) \\ &\quad + \frac{\varepsilon}{z_{max} - |x_1|} + u_{max}(x_2 + B_u^\top V_x^*) \tanh(\Psi) \\ &\quad + u_{max}^2 R \ln \left[1 - \tanh^2(\Psi) \right] \\ &= V_x^{*\top}(Ax + B_w w) + \frac{\varepsilon}{z_{max} - |x_1|} \\ &\quad + u_{max}^2 R \ln \left[1 - \tanh^2(\Psi) \right] \end{aligned} \tag{20}$$

If we solve the HJB equation (20) for the optimal cost value V^*, then the suboptimal causal control given by (18) can be obtained. Since the optimal cost function (13) has a finite-horizon, we know that V^*, V_t^* and V_x^* are dependent on time t [36], [37], and V_t^* appears in the HJB equation (20) though it is not involved in the control (18) explicitly. For this case, the HJB equation (20) is nonlinear and with time-varying nature, and thus it is generally difficult or even not possible to find

its analytical solution. Hence, in the next subsection we will present an alternative solution using the principle of ADP, where a critic NN is online trained to estimate the optimal cost-to-go function by further tailoring the idea presented in [26].

B. Online implementation with ADP

The idea of ADP [17] is to online solve the HJB equation in terms of adaptive methods, where the optimal cost function can be online estimated in terms of a critic NN as [22], [23]. However, as mentioned in the last subsection, the above cost-to-go function $V^*(x, t)$ for WEC is time dependent, which is more difficult to handle than that derived for other optimal control as in [22], [23]. To address this time-dependent issue, several researchers have attempted to further tailor the idea of ADP for optimal regulation problem [36]–[38]. In [36], a critic NN with time-dependent weight was introduced to approximate the time-dependant cost function (13). However, the time-dependent weight were calculated through a backward integration, which is time-consuming and computationally demanding. Alternatively, a critic NN with time-varying activation function (taking time-to-go as its input) and constant weight was suggested in [37], [38]. In this paper, to implement the resulting control algorithm online, we introduce a critic NN with time-varying activation functions as [37], [38].

Without loss of generality, the optimal value function $V^*(x, t)$ is assumed to be a continuous function on a compact set $\Omega \times [0, T]$ [23], and then it can be approximated by the following critic NN with a time-varying regressor function:

$$V^*(x, t) = W^{*\top}\phi(x, T - t) + \varepsilon_n \qquad (21)$$

where $W^* \in \mathbb{R}^l$ is the unknown constant NN weight, $\phi(x, T - t) = [\phi_1(x, T-t), \cdots, \phi_l(x, T-t)]^\top \in \mathbb{R}^l$ is the time-varying NN regressor vector, which is a function of the state x and the time-to-go $T - t$ as [37], [38]; here l denotes the number of neurons, and ε_n defines the residual NN error.

Then its derivatives with respect to x and t are given by

$$\frac{\partial V^*(x, t)}{\partial x} = \nabla\phi^\top(x, T-t)W^* + \nabla\varepsilon_n \qquad (22a)$$

$$\frac{\partial V^*(x, t)}{\partial t} = \nabla\phi_t^\top(x, T-t)W^* + \nabla\varepsilon_{nt} \qquad (22b)$$

where $\nabla\phi(x, T - t) = \partial\phi/\partial x$, $\nabla\phi_t(x, T - t) = \partial\phi/\partial t$, $\nabla\varepsilon_n = \partial\varepsilon_n/\partial x$ and $\nabla\varepsilon_{nt} = \partial\varepsilon_n/\partial t$ are defined as the partial derivatives of ϕ, ε_n regarding to x and t, respectively.

We make the assumption as used in [23] and other ADP references:

Assumption 1. *The ideal NN weight W^*, regressor ϕ and its derivatives $\nabla\phi, \nabla\phi_t$ of the critic NN are bounded by $\|W^*\| \leq W_N$, $\|\phi\| \leq \phi_N$, $\|\nabla\phi\| \leq \phi_M$, $\|\nabla\phi_t\| \leq \phi_t$. Moreover, the derivatives of approximation errors, e.g., $\nabla\varepsilon_n, \nabla\varepsilon_{nt}$, are bounded by $\|\nabla\varepsilon_n\| \leq \phi_{\varepsilon_n}$ and $\|\nabla\varepsilon_{nt}\| \leq \phi_{\varepsilon_{nt}}$.*

Based on (18) and (22), the suboptimal causal control u^* given by (18) is reformulated as

$$u^* = -u_{max}\tanh\left(\frac{1}{2Ru_{max}}\left[x_2 + B_u^\top(\nabla\phi^\top W^* + \nabla\varepsilon_n)\right]\right) \qquad (23)$$

In practical implementation, the ideal NN weight W^* is unknown. Hence, we can only use the estimate $\hat{W}$ of W^*, which is online updated by using appropriate adaptive algorithms. In this case, a practical critic NN with the estimated NN weight $\hat{W}$ is given by

$$\hat{V}(x, t) = \hat{W}^\top\phi(x, T - t) \qquad (24)$$

Substituting the estimated cost function (24) into (18) will yield the practical control as

$$u = -u_{max}\tanh\left(\frac{1}{2Ru_{max}}(x_2 + B_u^\top\nabla\phi^\top(x, T-t)\hat{W})\right) \qquad (25)$$

Next we investigate a new adaptive algorithm to online calculate the NN weight $\hat{W}$, which can guarantee that $\hat{W}$ converges to a small set around W^* even with the NN approximation error ε_n. Using this new adaptive law, the widely used actor NN (e.g. [21], [24]) can be avoided since the convergent critic NN weight $\hat{W}$ can be used to derive the practical control action in (25) directly. This results in a simplified ADP framework in comparison to most of the existing ADP results.

To design an adaptive law based on the HJB equation (20) and the critic NN (22), we have

$$0 = W^{*\top}\nabla\phi_t(t, T - t) + W^{*\top}\nabla\phi(x, T-t)(Ax + B_w w) + \frac{\varepsilon}{z_{max} - |x_1|} + u_{max}^2 R\ln\left[1 - \tanh^2(\Psi)\right] + \varepsilon_{HJB} \qquad (26)$$

where $\varepsilon_{HJB} = \nabla\varepsilon_n^\top(Ax + B_w w) + \nabla\varepsilon_{nt}$ is the residual error embedded in the HJB equation, which stems from the bounded NN approximation errors ε_n, $\nabla\varepsilon_n$ and $\nabla\varepsilon_{nt}$. In this case, ε_{HJB} is bounded. Specifically, this error can be arbitrarily small as the number of NN nodes $l \to +\infty$ [21], [23].

We can further represent (26) in a more compact form. Hence, define $\Xi := \nabla\phi_t + \nabla\phi(A_x + B_w w)$ and $\Theta := \frac{\varepsilon}{z_{max} - |x_1|} + u_{max}^2 R\ln\left[1 - \tanh^2(\Psi)\right]$, and then (26) can be rewritten as

$$\Theta = -W^{*\top}\Xi - \varepsilon_{HJB} \qquad (27)$$

Now, the unknown NN weight W^* is in a linearly parameterized form in (27) associated with the regressor vector Ξ. With this observation, we can further tailor the idea initially suggested in our recent work [22], [39] to design an adaptive law driven by the estimation error. Thus, the first step should be to extract the estimation error by defining a matrix M and a vector N by applying the following filter operations on the known dynamics Ξ, Θ as

$$\begin{cases} \dot{M} = -\iota M + \Xi\Xi^\top, M(0) = 0 \\ \dot{N} = -\iota N + \Xi\Theta, N(0) = 0 \end{cases} \qquad (28)$$

where $\iota > 0$ is a constant used to retain the boundedness of M, N.

Then another variable can be online calculated based on M and N as

$$\varpi = M\hat{W} + N \qquad (29)$$

We can prove that this introduced variable ϖ can represent the unknown estimation error between the unknown weight W^*

and the estimated weight $\hat{W}$. This can be summarized as the following lemma:

Lemma 1. *[22] The auxiliary variable ϖ calculated based on* (29) *with* (28) *can be represented as*

$$\varpi = -M\tilde{W} + \varphi \tag{30}$$

where $\tilde{W} = W^* - \hat{W}$ defines the estimation error of the NN weight, and $\varphi = -\int_0^t e^{-\iota(t-r)}\varepsilon_{HJB}(r)\Xi(r)dr$ denotes the effect of the residual HJB error, which is also bounded, i.e., $\|\varphi\| \leq \varepsilon_N$ for a positive constant $\varepsilon_N > 0$.

Proof: By solving (28), one can obtain its solution as

$$\begin{cases} M(t) &= \int_0^t e^{-\iota(t-r)}\Xi(r)\Xi^\top(r)dr \\ N(t) &= \int_0^t e^{-\iota(t-r)}\Xi(r)\Theta(r)dr \end{cases} \tag{31}$$

Substituting (27) into (31), we can verify that

$$N = -MW^* + \varphi \tag{32}$$

Then from (29) and (32), it can be verified that (30) holds. On the other hand, the boundedness of φ can be proved by considering the definition φ and the fact that the regressor ϕ and the estimation error ε_n are all bounded. ⋄

Based on Lemma 1, we can find that the variable ϖ includes the estimation error $\tilde{W}$ perturbed by a bounded variable φ, which can be made arbitrarily small and assumed as a disturbance. Then based on the analysis in [22], the adaptive law driven by this estimation error can help to prove the convergence of $\hat{W}$ to W^* with fast convergence rate. Therefore, we can present the following adaptive law (33) driven by ϖ to online calculate the critic NN weight $\hat{W}$ as

$$\dot{\hat{W}} = -\Gamma\varpi \tag{33}$$

where the learning gain $\Gamma > 0$ can be set as a positive constant matrix.

Moreover, before proving the convergence of the proposed adaptive law, we need to investigate the positive definiteness of the matrix M. Define $\lambda_{\max}(\cdot)$, $\lambda_{\min}(\cdot)$ as the maximum and minimum matrix eigenvalues. Then we have the following lemma.

Lemma 2. *[39] The condition $\lambda_{\min}(M) > \sigma > 0$ holds for a constant $\sigma > 0$ (i.e., M is positive definite) provided that Ξ defined below* (26) *is persistently excited (PE).*

Based on (28), the matrix M is a filtered version of $\Xi\Xi^\top$ as shown in (31). Then by applying straightforward mathematical manipulations based on the definition of PE for Ξ, the positive definiteness of M can be proved. We refer to [39] for a similar proof.

The main results of this paper can be given as the following Theorem:

Theorem 2. *Consider nonlinear WEC system* (11) *with cost function* (13), *suboptimal causal control* (25) *and adaptive law* (33) *with the regressor Ξ being PE, then $\tilde{W}$ converges to a neighborhood around zero, and the derived control u given in* (25) *converges to a neighborhood around the optimal solution* (23). *Specifically, for the case when the NN errors are zero (i.e., $\varepsilon_n, \nabla\varepsilon_n, \nabla\varepsilon_{nt} = 0$), $\tilde{W}$ exponentially converges to zero, and the derived control u in* (25) *converges to the optimal solution in* (23).*

Proof: Based on Lemma 2, we can validate that if Ξ in (26) is PE, then the matrix M is positive definite, such that $\lambda_{\min}(M) > \sigma > 0$ is true for a constant $\sigma > 0$. Now, we select a Lyapunov function $V = \frac{1}{2}\tilde{W}^\top\Gamma^{-1}\tilde{W}$, then the derivative of V with respect to time t can be calculated along (33) as

$$\begin{aligned} \dot{V} &= \tilde{W}^\top\Gamma^{-1}\dot{\tilde{W}} = -\tilde{W}^\top M\tilde{W} + \tilde{W}^\top\varphi \\ &\leq -\sigma\|\tilde{W}\|^2 + \|\tilde{W}\|\varepsilon_N \\ &\leq -(\sigma - \frac{1}{2\eta})\|\tilde{W}\|^2 + \frac{\eta}{2}\varepsilon_N^2 \\ &\leq -\alpha V + \gamma \end{aligned} \tag{34}$$

where $\alpha = 2(\sigma - 1/2\eta)/\lambda_{\max}(\Gamma^{-1})$, $\gamma = \eta\varepsilon_N^2/2$ are positive constants for appropriately selected constant $\eta > 1/2\sigma$.

From the Lyapunov Theorem, we can claim from (34) that the NN weight error $\tilde{W}$ is bounded, and will exponentially converge to a set defined by $\Omega_w = \{\tilde{W}\|\|\tilde{W}\| \leq \sqrt{2\gamma/\alpha\lambda_{\min}(\Gamma^{-1})}\}$. It is shown that the size of this set is determined by the NN error ε_N, the learning gain Γ and the excitation level σ. Moreover, considering the continuous property of function $\tanh(\cdot)$ and comparing (25) with (23), it can be verified that the approximated control u in (25) can also converge to a neighborhood around the optimal solution u^* in (23).

In particular case when the NN approximation error is zero, i.e., $\varepsilon_n = 0$, one can validate that $\varphi = 0$ is true. Thus, (34) can be represented to

$$\dot{V} = -\tilde{W}^\top M\tilde{W} < -\sigma\|\tilde{W}\|^2 \leq -\mu V \tag{35}$$

where $\mu = 2\sigma/\lambda_{\max}(\Gamma^{-1})$ is a positive constant. Consequently, we can claim that $\tilde{W}$ exponentially converges to zero, and thus the proposed control u given in (25) converges to the optimal control u^* given in (23). This finishes the proof. ⋄

Remark 2. *The regressor $\phi(x, T-t)$ of the critic NN can be appropriately selected so that $\{\phi_i : i = 1, \ldots, l\}$ can provide a completely independent basis. Then from the Weierstrass theorem [23], the NN approximation error can be made sufficiently small by choosing sufficient neurons in the critic NN, which means $\varepsilon_n, \nabla\varepsilon_n, \nabla\varepsilon_{nt} \to 0$ as $l \to +\infty$. However, in practice, the selection of NN nodes should be taken as a trade-off between the convergence performance and the required computational costs.*

Remark 3. *Compared to the existing ADP methods, e.g., [21], [24], where another actor NN are needed, this paper presents a new adaptive law* (33), *which can estimate the unknown parameter with guaranteed convergence. Thus, we can use the estimated critic NN weight $\hat{W}$ to calculate the control action directly without training another actor NN. This simplified ADP framework can help to reduce computational costs, and tackle the problem of 'curse of dimensionality' encountered in the classical dynamic programming methods.*

Remark 4. *Lemma 2 states that the well-known PE condition is sufficient for the required condition $\lambda_{\min}(M) > \sigma > 0$,*

which is required in the proof of the convergence of (33). This condition can be true for generic WEC systems since there is external sea wave input w in the system (11). In particular, it is possible to numerically online verify this condition by testing the minimum eigenvalue of M based on Lemma 2.

Remark 5. *To implement the proposed control (25), several parameters need to be selected by the designer. The regressor of critic NN is set based on the analysis shown in Remark 2. As shown in the proof of the above Theorem 2, a large learning gain Γ can enhance the convergence speed of the critic NN weight, while a too large adaptive gain can lead to oscillations. Hence, Γ can be chosen as a trade-off between the convergence rate and robustness. The filter constant ι in (28) is used to retain the boundedness of M, N since it can serve as a forgetting factor in the differential equation (28), and it also determines the convergence rate of M and N. Hence, ι cannot be set too large in general to avoid introducing too large DC gains in (28).*

The implementation of the proposed ADP based WEC control algorithm described in Section III can be given in Algorithm 1.

Algorithm 1 Implementation of the proposed WEC control

1: Measure the system state x and the current wave w.
2: Construct the regressor $\phi(x, T-t)$ for the critic NN, and then derive the HJB equation (27).
3: Calculate the auxiliary variables M, N as given in (28), and then online update the unknown weight $\hat{W}$ based on (30) and (33).
4: Derive the practical control action based on (25), and apply it on the WEC plant.
   Go back to Step 1.

IV. SIMULATIONS

This section presents numerical simulations to show the efficacy of the proposed control method. The model parameters of the used point absorber are given in Table I, and the radiation force matrices A_r, B_r are the same as those used in [7], which represent the dynamics of a medium-sized point absorber. The float diameter is $d = 9$ m and at a height of $h = 2.4$ m above the sea bottom, which leads to the float heave motion limits as $[-1.2, 1.2]$ m. This means the output constraint is given by $z_{\max} = 1.2$ m. The maximum allowable control input is $u_{\max} = 3 \times 10^5$ N. The dimensions of the float in Fig. 2 are $d_1 = 4.5$ m, $d_2 = 2$ m, $h_1 = 1.2$ m and $h_2 = 1$ m.

Numerical simulations of the proposed ADP control for nonlinear WEC system are run on a PC with Intel(R) Core (TM) i7 CPU @ 2.70 GHz, 8.00 GB memory, 64-bit OS. A simulator was built in Simulink MATLAB, where the sampling interval is set as 0.05 sec. The parameters used in the proposed ADP based control are chosen as: $\phi(x, T - t) = [x_1 t_n^4, x_2 t_n^5, 0.5 x_1^2 t_n^4, 0.5 x_2^2 t_n^4, x_1 x_2 t_n^4]^\top$ with $t_n = (T - t)/T, T = 50$ being the normalised time-to-go, $\hat{W}(0) = [0, 0, 0, 0, 0]^\top$, $R = 1/(3.8 \times 10^5)$, $\varepsilon = 1$, $z_{max} = 1.2$, $\Gamma = \text{diag}([1, 1, 1, 1, 1])$ and $\iota = 1.5$.

TABLE I
PARAMETERS USED FOR THE WEC MODEL

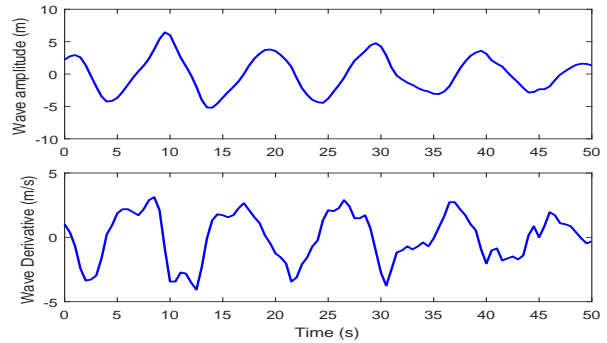| Description	Notation	Values
Density of sea water	ρ	1025 kg/m^3
Gravity	g	9.8 N/kg
Float radius	r	4.5 m
Damping (friction)	D_f	2×10^3 Nm/s
Float mass	m_s	1×10^4 kg
Added mass	m_∞	7×10^4 kg
Stiffness	D_f	6.39×10^5 N/m
Input force limit	$u_{\max}$	3×10^5 N

Fig. 3. Real sea wave profile No. 1: wave elevation magnitude and its first derivative with respect to time.

Two real sea wave profiles gathered from the coast of Cornwall, UK are used in the simulations. In the first case, a wave profile with large amplitude (as shown in Fig.3) is used to validate the proposed ADP control. Simulation results are provided in Fig. 4 and Fig. 5. It can be observed from Fig. 4 that the constraints imposed on the control input u and the system output x_1 can be strictly fulfilled and the control signal is very smooth, which makes an easy implementation of this control input on hardware actuators. Moreover, Fig. 5 shows the extracted power and the energy output with the ADP control algorithm, which clearly illustrates stable energy output with the proposed ADP control. Finally, Fig. 6 gives the profile of the estimated critic NN weight $\hat{W}$, which illustrates the effectiveness of the introduced adaptive law for retaining the convergence as proved in the Theorem 2.

In the second case, a small wave profile used in [12] as shown in Fig. 7 is adopted for simulation. Comparative results
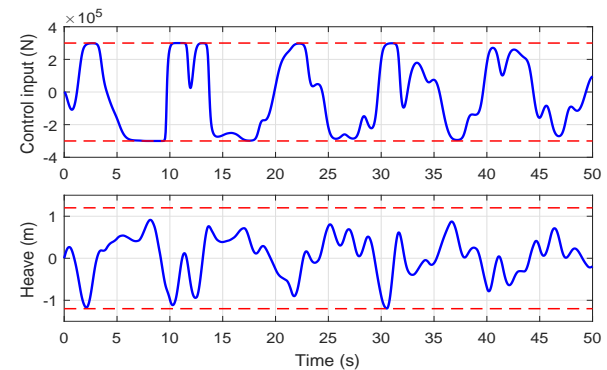

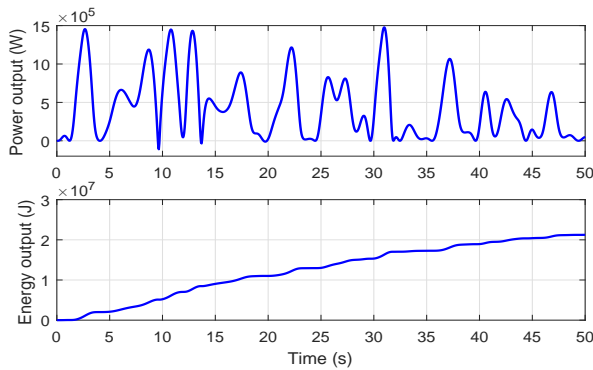Fig. 4. Control signal u and the heave motion x_1 for Wave No.1.

Fig. 5. The power output and the extracted energy with ADP control.
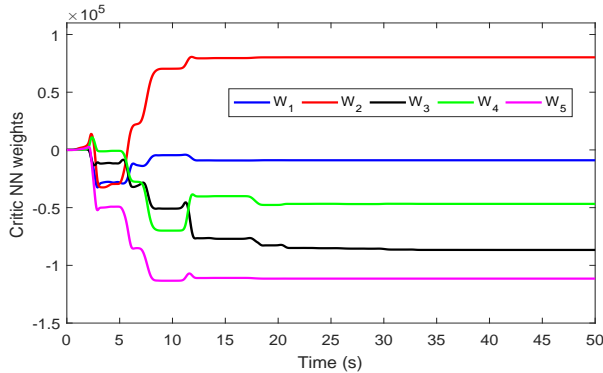

Fig. 6. Convergence of the critic NN weight $\hat{W}$.

of the MPC and ADP control are given in Fig.8 to Fig. 11. Fig. 8 shows the power outputs of the WEC controlled by the MPC and ADP control algorithms respectively when no model mismatch is assumed. Note that the simulation results of using MPC in [12] show the case when an extra constraint on unidirectional power flow (i.e., from WEC to grid) is imposed. This extra constraint in the MPC simulation is removed here for fair comparison purpose,since the unidirectional power flow cannot be guaranteed by the proposed ADP control. Most of the existing optimal WEC control strategies result in bidirectional energy flow (see [40], [41]) due to the nature of the WEC control problem: although the power flow can be negative (from grid to waves) at some instants, the energy output over a finite period can be more extracted compared to the case when the power flow is constrained always to be positive, as also demonstrated in [12]. As stated in [41], the bidirectional energy flow can be achieved by development of hydraulic PTO mechanisms (e.g. [31]) or electric generators (e.g. [30]). The PTO design for unidirectional energy flow may be less complicated than the case of bi-directional flow in some cases; it is thus an open question regarding the tradeoff between the increased energy output by allowing bidirectional flow and the PTO hardware cost for achieving this bidirectional flow, which depends on specific cases.

Because non-causal wave prediction information is not used, the proposed ADP control results in less energy output than MPC by about 11% as shown in Fig.9. However, the proposed ADP control shows better robustness performance against model uncertainties. When the WEC model used in the control design has model mismatches (e.g., added mass $m_d = 0.4m$ and radiation force $y_{rd} = 20y_r$), there is a

TABLE II
COMPARISONS BETWEEN MPC AND ADP FOR SIMULATION INTERVAL $T = 50$ S

Control method	Extracted Energy (J)	Time (sec)
ADP	1.39×10^7	0.92
MPC	1.56×10^7	42.59

significant decrease by 19% of the energy output from the WEC controlled by MPC, while the energy output decrease using the ADP control is negligible. Another advantage of the proposed ADP control is its ability to handle constraints. As shown in Fig.10, the constraints imposed on the heave motion x_1 can be retained for both controllers. Additionally, one may find from Fig.11 that the control signal of the ADP control is smoother than that from MPC [12], which is preferable for control implementation in practice.

Finally, to show the necessity for using a nonsquare function $U(u(\tau))$ (14) in the cost function (13) to strictly guarantee the input constraint $|u| \leq u_{max}$, we also implement the ADP control in [26], where only a quadratic term $u^\top R u$ is used. Note the adopted WEC model and wave profile in this paper are different to [26], thus the weight parameter R of the ADP control [26] is retuned. First, we tune the parameter $R = 2.9 \times 10^{-7}$ to get a similar energy output as the ADP control with input constraint introduced in this paper. Fig. 12 shows the generated energy and the required control actions, from which we can see that the control input via the ADP without input constraint in [26] exceeds the control saturation, which in turn affects the safe operation of WEC system. Following the analysis in [26], we then increase the parameter to $R = 4.9 \times 10^{-7}$ for the ADP control in [26] to reduce the amplitude of the control input to fulfill the saturation condition. However, as shown in Fig. 13 the energy output via the ADP without input constraint [26] decreases about 36.2%. From these comparative simulations, it is evident that explicitly incorporating constraints into the cost function for the ADP control design can obtain better performance in terms of the satisfaction of the control input constraint and the generated energy.

Apart from the smooth control signal, strictly guaranteed constraints and the robustness, a major attractive feature of the proposed ADP control compared to MPC lies in its reduced computational cost and significantly improved computational efficiency in the online implementation. To further show this feature explicitly, we compare the simulation time and the generated energy of the proposed ADP control with the MPC method in [12]. Table II summarizes the results for the simulation interval $T = 50$ s, which shows that the computational time of MPC used in the simulation is clearly much longer than that of the ADP control. This is very attractive in practical real-time application, in particular for high-order WEC systems, where the resulting WEC optimal control problem becomes intractable.

V. CONCLUSIONS

In this paper, we propose an efficient adaptive suboptimal causal control strategy for WEC systems subject to nonlinear-
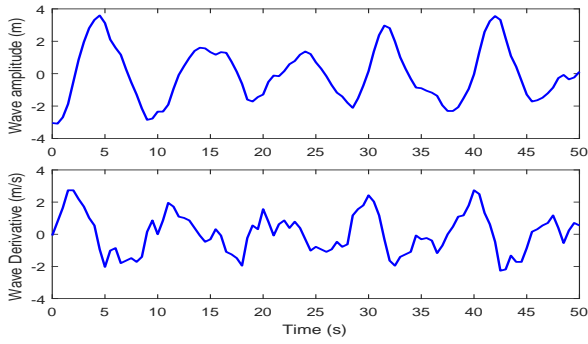
Fig. 7. Real sea wave profile No. 2: wave elevation magnitude and its first derivative with respect to time.
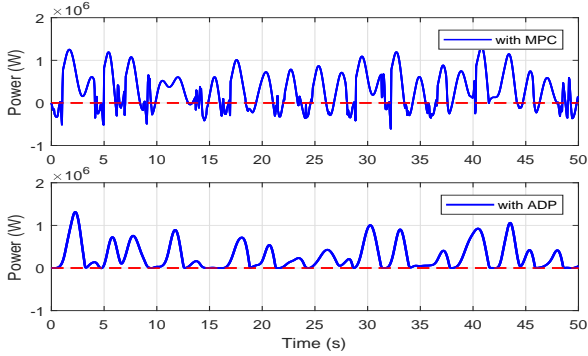

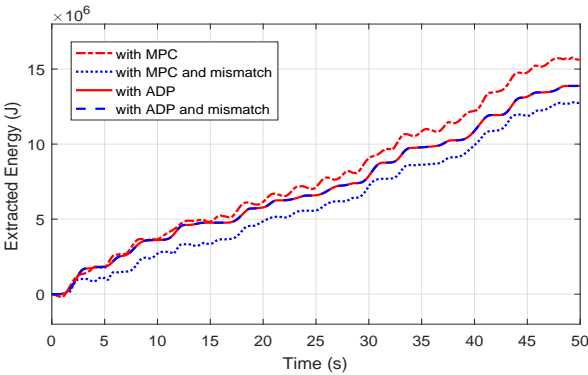
Fig. 8. Power outputs of MPC and ADP control.

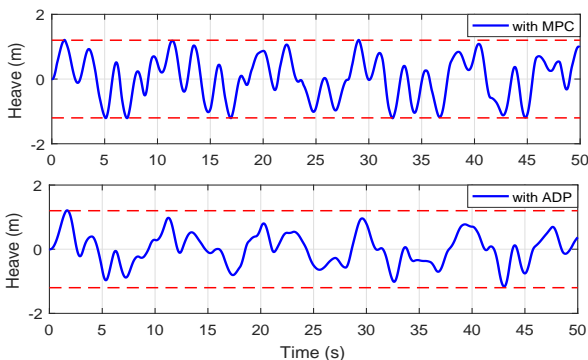Fig. 9. Extracted energy of MPC [12] and ADP control.

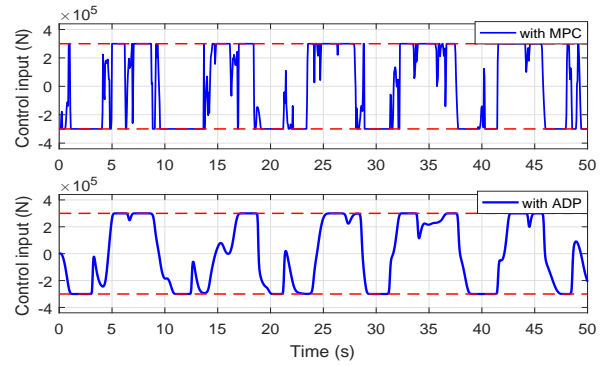Fig. 10. Heave motion x_1 of MPC and ADP control.

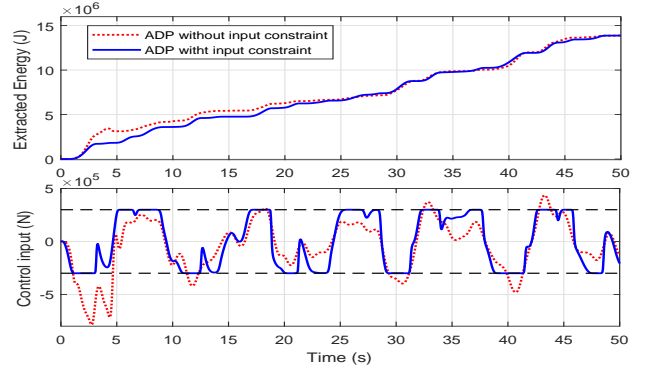Fig. 11. Control inputs u of MPC and ADP control.

Fig. 12. Energy output and control inputs of ADP control with/without input constraint with $R = 2.9 \times 10^{-7}$.

ities and constraints. We propose a modified optimization cost function to solve the problem of maximizing the generated energy, the control saturation and the constraint imposed on the system state. A constructive method is investigated to online solve the derived HJB equation by using the principle of ADP, where a critic NN is used to approximate the optimal cost function and calculate the control action. A new adaptive law is developed to online update the critic NN weight with guaranteed convergence. It has been proved that the input constraint can be strictly guaranteed, and the obtained control can converge to a small set around the optimal solution. This leads to a simplified ADP framework with very high computational efficiency. Comparative simulations demonstrate that the proposed online ADP control achieves energy output less than that of MPC but yields smoother
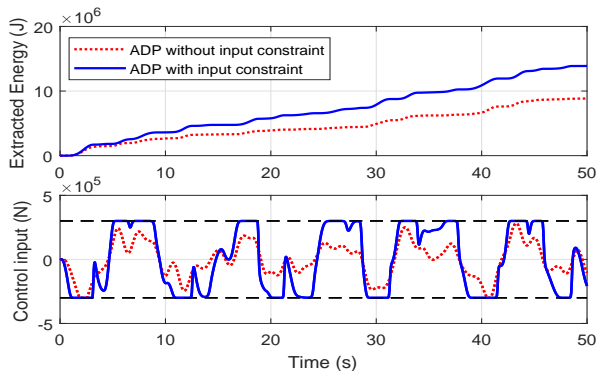

Fig. 13. Energy output and control inputs of ADP control with/without input constraint with $R = 4.9 \times 10^{-7}$

control input, much faster computational time and exhibits more robustness. Although the proposed control is causal and can only achieve a suboptimal solution compared with MPC, the significantly reduced computational burden can be a major benefit to promote its practical implementation especially for WECs with complicated models. The proposed approach has the potential to be extended to the control of other types of WECs, and even other energy maximization control problems. Future work will focus on modeling and control of full wave energy harvest system including the electricity generation subsystem.

REFERENCES

[1] A. Clément and et al, "Wave energy in europe: current status and perspectives," *Renewable and Sustainable Energy Reviews*, vol. 6, no. 5, pp. 405 – 431, 2002.

[2] U. A. Korde, "Latching control of deep water wave energy devices using an active reference," vol. 29, no. 11, pp. 1343–1355, 2002.

[3] J. Falnes, *Ocean Waves and Oscillating Systems*. Cambridge University Press, 2002.

[4] J. Cretel, A. W. Lewis, G. Lightbody, and G. P. Thomas, "An application of model predictive control to a wave energy point absorber," in *IFAC Conference on Control Methodologies and Technology for Energy Efficiency*, Vilamoura, Portugal, 2010.

[5] J. A. M. Cretel, G. Lightbody, G. P. Thomas, and A. W. Lewis, "Maximisation of energy capture by a wave-energy point absorber using model predictive control," in *Proc. 18th IFAC World Congress*, Milano, 2011.

[6] J. Hals, J. Falnes, and T. Moan, "Constrained optimal control of a heaving buoy wave-energy converter," *J. of Offshore Mechanics and Arctic Engineering*, vol. 133, no. 1, p. 011401, 2011.

[7] G. Li, G. Weiss, M. Mueller, S. Townley, and M. B. Belmont, "Wave energy converter control by wave prediction and dynamic programming," *Renewable Energy*, vol. 48, pp. 392–403, 2012.

[8] J. Ringwood, G. Bacelli, and F. Fusco, "Energy-maximizing control of wave-energy converters: The development of control system technology to optimize their operation," *Control Systems, IEEE*, vol. 34, no. 5, pp. 30–55, 2014.

[9] G. Li and M. R. Belmont, "Model predictive control of sea wave energy converters–Part I: A convex approach for the case of a single device," *Renewable Energy*, vol. 69, pp. 453–463, 2014.

[10] J. Davidson, R. Genest, and J. V. Ringwood, "Adaptive control of a wave energy converter," *IEEE Transactions on Sustainable Energy*, vol. PP, no. 99, pp. 1–1, 2018.

[11] A. P. L. Merigaud and J. V. Ringwood, "Improving the computational performance of non-linear pseudo-spectral control of wave energy converters," *IEEE Transactions on Sustainable Energy*, vol. PP, no. 99, pp. 1–1, 2017.

[12] G. Li, "Nonlinear model predictive control of a wave energy converter based on differential flatness parameterisation," *International Journal of Control*, vol. 90, no. 1, pp. 68–77, 2017.

[13] B. Xu, D. Yang, Z. Shi, Y. Pan, B. Chen, and F. Sun, "Online recorded data based composite neural control of strict-feedback systems with application to hypersonic flight dynamics," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 29, no. 8, p. 3839, 2018.

[14] J. Na, Q. Chen, X. Ren, and Y. Guo, "Adaptive prescribed performance motion control of servo mechanisms with friction compensation," *IEEE Transactions on Industrial Electronics*, vol. 61, no. 1, pp. 486–494, 2014.

[15] J. Na, Y. Huang, X. Wu, G. Gao, G. Herrmann, and J. Z. Jiang, "Active adaptive estimation and control for vehicle suspensions with prescribed performance," *IEEE Transactions on Control Systems Technology*, vol. PP, no. 99, pp. 1–15, 2017.

[16] R. S. Sutton and A. G. Barto, *Reinforcement learning: An introduction*. MIT press Cambridge, 1998, vol. 1, no. 1.

[17] F. L. Lewis, D. L. Vrabie, and V. L. Syrmos, "Reinforcement learning and optimal adaptive control," *Optimal Control, Third Edition*, 2012.

[18] P. J. Werbos, "A menu of designs for reinforcement learning over time," *Neural networks for control*, pp. 67–95, 1990.

[19] D. Wang, D. Liu, Q. Wei, D. Zhao, and N. Jin, "Optimal control of unknown nonaffine nonlinear discrete-time systems based on adaptive dynamic programming," *Automatica*, vol. 48, no. 8, pp. 1825–1832, 2012.

[20] Q. Yang, J. B. Vance, and S. Jagannathan, "Control of nonaffine nonlinear discrete-time systems using reinforcement-learning-based linearly parameterized neural networks," *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on*, vol. 38, no. 4, pp. 994–1001, 2008.

[21] K. G. Vamvoudakis and F. L. Lewis, "Online actor critic algorithm to solve the continuous-time infinite horizon optimal control problem," *Automatica*, vol. 46, no. 5, pp. 878 – 888, 2010.

[22] J. Na and G. Herrmann, "Online adaptive approximate optimal tracking control with simplified dual approximation structure for continuous-time unknown nonlinear systems," *Automatica Sinica, IEEE/CAA Journal of*, vol. 1, no. 4, pp. 412–422, 2014.

[23] M. Abu-Khalaf and F. L. Lewis, "Nearly optimal control laws for nonlinear systems with saturating actuators using a neural network HJB approach," *Automatica*, vol. 41, no. 5, pp. 779–791, 2005.

[24] H. Zhang, L. Cui, X. Zhang, and Y. Luo, "Data-driven robust approximate optimal tracking control for unknown general nonlinear systems using adaptive dynamic programming method," *Neural Networks, IEEE Transactions on*, vol. 22, no. 12, pp. 2226–2236, 2011.

[25] D. Wang and C. Mu, "Adaptive-critic-based robust trajectory tracking of uncertain dynamics and its application to a spring-mass-damper system," *IEEE Transactions on Industrial Electronics*, vol. PP, no. 99, pp. 1–1, 2017.

[26] J. Na, G. Li, B. Wang, G. Herrmann, and S. Zhan, "Robust optimal control of wave energy converters based on adaptive dynamic programming," *IEEE Transactions on Sustainable Energy*, vol. PP, no. 99, 2018.

[27] A. I. Vakis and J. S. Anagnostopoulos, "Mechanical design and modeling of a single-piston pump for the novel power take-off system of a wave energy converter," *Renewable Energy*, vol. 96, pp. 531–547, 2016.

[28] A. F. D. O. Falco, "Modelling and control of oscillating-body wave energy converters with hydraulic power take-off and gas accumulator," *Ocean Engineering*, vol. 34, no. 14, pp. 2021–2032, 2007.

[29] G. Weiss, G. Li, M. Mueller, S. Townley, , and M. R. Belmont, "Optimal control of wave energy converters using deterministic sea wave prediction," in *The Energy & Materials Research Conference*, Torremolinos, Malaga, Spain, 2012.

[30] D. O'Sullivan, J. Griffiths, M. G. Egan, and A. W. Lewis, "Development of an electrical power take off system for a sea-test scaled offshore wave energy device," *Renewable Energy*, vol. 36, no. 4, pp. 1236 – 1244, 2011.

[31] R. Costello, J. V. Ringwood, and J. Weber, "Comparison of two alternative hydraulic pto concepts for wave energy conversion," in *Proc. 9th Eur. Wave and Tidal Energy Conf. (EWTEC)*, Southampton, U.K., 2011.

[32] V. Franzitta, D. Curto, and D. Rao, "Energetic sustainability using renewable energies in the mediterranean sea," *Sustainability*, vol. 8, no. 11, p. 1164, 2016.

[33] V. Franzitta, P. Catrini, and D. Curto, "Wave energy assessment along sicilian coastline, based on deim point absorber," *Energies*, vol. 10, no. 3, p. 376, 2017.

[34] M. Penalba and J. V. Ringwood, "A review of wave-to-wire models for wave energy converters," *Energies*, vol. 9, no. 7, 2016.

[35] S. R. Nielsen, Q. Zhou, M. M. Kramer, B. Basu, and Z. Zhang, "Optimal control of nonlinear wave energy point converters," *Ocean engineering*, vol. 72, pp. 176–187, 2013.

[36] T. Chen, F. L. Lewis, and M. Abu-Khalaf, "A neural network solution for fixedfinal time optimal control of nonlinear systems," *Automatica*, vol. 43, pp. 482–490, 2007.

[37] A. Heydari and S. N. Balakrishnan, "Finite-horizon control-constrained nonlinear optimal control using single network daptive critics," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 24, no. 1, pp. 145–157, 2013.

[38] Q. Zhao, H.Xu, and S. Jagannathan, "Neural network-based finite-horizon optimal control of uncertain affine nonlinear discrete-time systems," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 26, no. 3, p. 486499, 2015.

[39] J. Na, M. N. Mahyuddin, G. Herrmann, X. Ren, and P. Barber, "Robust adaptive finite-time parameter estimation and control for robotic systems," *International Journal of Robust and Nonlinear Control*, vol. 25, no. 16, pp. 3045 – 3071, 2015.

[40] N. Faedo, S. Olaya, and J. V. Ringwood, "Optimal control, mpc and mpc-like algorithms for wave energy systems: An overview," *IFAC Journal of Systems and Control*, vol. 1, pp. 37 – 56, 2017.

[41] F. Fusco and J. V. Ringwood, "A simple and effective real-time controller for wave energy converters," *IEEE Transactions on Sustainable Energy*, vol. 4, no. 1, pp. 21–30, 2013.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIE.2018.2880728, IEEE Transactions on Industrial Electronics

IEEE TRANSACTIONS ON INDUSTRIAL ELECTRONICS

Jing Na (M'15) received the B.Sc. and Ph.D. degrees from the School of Automation, Beijing Institute of Technology, Beijing, China, in 2004 and 2010, respectively. From 2011 to 2013, he was a Monaco/ITER Postdoctoral Fellow at the ITER Organization, Saint-Paul-lès-Durance, France. From 2015 to 2017, he was a Marie Curie Intra-European Fellow with the Department of Mechanical Engineering, University of Bristol, U.K. Since 2010, he has been with the Faculty of Mechanical and Electrical Engineering, Kunming University of Science and Technology, Kunming, China, where he became a Professor in 2013. He is currently an Associate Editor of Neurocomputing and has served as an international program committee Chair of ICMIC 2017. His current research interests include intelligent control, adaptive parameter estimation, nonlinear control and applications. Dr Na has been awarded the Best Application Paper Award of the 3rd IFAC International Conference on Intelligent Control and Automation Science (IFAC ICONS 2013), and the 2017 Hsue-shen Tsien Paper Award.

Wei He (M'12-SM'16) received the B.Eng. and M.Eng. degrees in control engineering from the College of Automation Science and Engineering, South China University of Technology, Guangzhou, China, in 2006 and 2008, respectively, and the Ph.D. degree in control engineering from the Department of Electrical and Computer Engineering, National University of Singapore, Singapore, in 2011. He is currently a Full Professor with the School of Automation and Electrical Engineering, University of Science and Technology Beijing, Beijing, China. He has coauthored 2 book published in Springer and authored or coauthored more than 100 international journal and conference papers. His current research interests include robotics, distributed parameter systems, and intelligent control systems. He was the recipient of a Newton Advanced Fellowship from the Royal Society, U.K., the IEEE SMC Society Andrew P. Sage Best Transactions Paper Award in 2017. He is the Chair of IEEE SMC Society Beijing Capital Region Chapter. He is an Associate Editor of the IEEE TRANSACTIONS ON NEURAL NETWORKS AND LEARNING SYSTEMS, the IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS: SYSTEMS, and IEEE ACCESS, and an Editor of the IEEE/CAA JOURNAL OF AUTOMATICA SINICA, Neurocomputing, and Journal of Intelligent and Robotic Systems. He is the member of the IFAC TC on Distributed Parameter Systems, IFAC TC on Computational Intelligence in Control, and IEEE CSS TC on Distributed Parameter Systems.

Bin Wang received the B.S. from Beijing Technology and Business University, Beijing, China, in 2016. He is currently pursuing the M.S. degree in Kunming University of Science and Technology, Kunming, China. His current research interests include adaptive dynamic programming, optimal control, parameter identification and prescribed performance control.

Guang Li (M'09) received his PhD degree in Electrical and Electronics Engineering, specialized in control systems, from the University of Manchester. He is currently a Senior Lecturer in dynamics modelling and control in Queen Mary University of London, UK. His current research interests include constrained optimal control, model predictive control, adaptive robust control and control applications including renewable energies and energy storage, etc.

Siyuan Zhan was born in Yingtan, China, in 1992. He received the bachelors degree from Shanghai Jiao Tong University, Shanghai, China, in 2013, and the masters degree from the University of Pennsylvania, Philadelphia, PA, USA, in 2014. He is currently pursuing the Ph.D. degree with the Queen Mary University of London, London, U.K. His current research interests include constrained optimal control, model predictive control, and con- trol applications in renewable energy.