OPEN ACCESS

University of BRISTOL

Peer reviewed version

University of Bristol - Explore Bristol Research
General rights

Robust Optimal Control of Wave Energy Converters Based on Adaptive Dynamic Programming

Jing Na, *Member, IEEE,* Guang Li, *Member, IEEE,* Bin Wang, Guido Herrmann, *Senior Member, IEEE,* and Siyuan Zhan

Abstract—This paper presents a robust adaptive optimal control strategy for wave energy converters (WECs). We first propose a new estimator in a simple form to address modeling uncertainties and formulate the control of WECs as an optimal control problem. Then a novel energy maximization control strategy is developed based on the concept of adaptive dynamic programming (ADP), where a critic neural network (NN) is used to approximate the time-dependent optimal cost value. To achieve guaranteed convergence, a recently proposed adaptive law based on the parameter estimation error is further tailored to online update the weights of critic NN. Consequently, the critic NN output, e.g. the costate, can be used to compute the optimal feedback control. The proposed robust ADP WEC control method is not only effective in handling dynamic uncertainties, but also computationally efficient with a very fast online convergence rate for the weights of the critic NN (less than 20 seconds for irregular sea waves as demonstrated in the simulations). These advantages significantly enhance the real-time applicability of the proposed method. Simulation results show that this approach is robust to model uncertainties and has significantly reduced computational costs for implementation.

Index Terms—Wave energy converter, adaptive optimal control, adaptive dynamic programming, uncertainty estimator.

I. INTRODUCTION

Ocean waves provide vast, persistent and spatially concentrated energy compared with other renewable energy resources (e.g. solar and wind energies) [1], [2]. In the UK, between 200 and 300 MWs of wave and tidal energy may be harvested by 2020, and up to 27GWs by 2050 [3]. However, current wave technology is still immature for commercial purposes because the low energy extraction rate and the high risk of device damage can cause much higher unit cost of generated electricity than fossil fuels and even other relatively mature renewable energies [4]. Hence, appropriate control strategies need to be developed for WECs to solve the energy maximization and guarantee their safe operation.

Although a large number of wave energy converters (WECs) have been developed over the past decades, the corresponding control strategy development has lagged behind, which makes

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 E1 4NS, UK, (Email: g.li@qmul.ac.uk, s.zhan@qmul.ac.uk).

Guido Herrmann is with the Department of Mechanical Engineering, University of Bristol, BS8 1TR, UK, (Email:G.Herrmann@bristol.ac.uk).

the WECs' performance far from being optimal [5]. The WEC control problem is challenging since most conventional control strategies developed for tracking or regulation are not directly suitable for the energy maximization problem. Conventional control strategies for WECs were mainly developed using the impedance matching principle, that is, the output energy can be maximized if the dominant frequency of the incoming waves matches the WEC's resonance frequency, e.g. [6]–[8]. However, these control approaches are effective for idealized regular waves and can become very complicated to implement for realistic irregular sea waves.

Development of advanced control strategies for WECs is identified as one of the most promising cost-reduction pathways [4]. It is found that the control task of WECs is to maximize energy generation from sea waves and reduce the risk of device damage. The energy maximization problem can be formulated as a constrained optimal control problem, which is different from the conventional optimal control problems for reference tracking or regulation. Recent studies show that the WEC control problem can be potentially resolved using an economic model predictive control (MPC) strategy [9]–[12]. However, the main challenge for the MPC of WEC is the heavy computational burden for resolving the constrained optimization problem online. To reduce the computational burden, some alternative methods have been proposed, such as the approach based on a modified objective function leading to a convex optimization [13], linear noncausal optimal control [14], adaptive control [15], nonlinear MPC based on pseudospectral control [16], and nonlinear MPC based on a combination of the pseudospectral method and the differential flatness property [17].

The efficacy of the above WEC control strategies, especially those model-based ones, is highly affected by the accuracy of the WEC models. However, precise WEC modeling is not a trivial task. In particular, some parameters of the WEC hydrodynamic model, e.g. damping ratio, can vary significantly due to the change of sea conditions, and the nonlinearities from wave-structure interaction sometimes treated as unmodeled dynamic uncertainties can become more significant for large waves. Neglecting such uncertainties can degrade WEC control performance on the one hand; on the other hand, explicitly describing these uncertainties using highly complicated WEC models can dramatically increase the design complexities of the WEC controls. Therefore, it is highly promising to develop a control strategy that can handle the WEC model uncertainties efficiently.

Recently, reinforcement learning (RL) [18], [19] based on the principle of learning from experience coupled with the

reward and punishment for survival and growth has been used in many control designs, which subsequently stimulated the development of a new optimal control design methodology, named approximate dynamic programming (ADP). In the originally proposed actor-critic based ADP framework [20], neural networks (NNs) are trained to approximate the solution of the derived Hamilton-Jacobi-Bellman (HJB) equation and the control action [21], [22]. In recent years, substantial work has also been reported to address the online ADP control for continuous-time systems with unknown dynamics [23], [24]. However, note that these existing ADP methods have been mainly proposed to solve regulation or reference tracking problems and cannot be directly used to solve the energy maximization control problem of WECs.

In this paper, we exploit the applicability of ADP to the WEC control problem and present an alternative optimal WEC control method based on a new ADP scheme [24], [25]. The potential modeling errors are first online estimated using a simple uncertainty estimator, which only imposes filter operations on the measured system variables. We then reformulate the control of WEC systems as a constrained optimal control problem with a finite horizon, where the derived time-dependent HJB equation needs to be solved. For the purpose of online implementation, a critic NN with the current system states and the time-to-go [26], [27] being its inputs is constructed and used to approximate the solution of the HJB equation. Finally, a new adaptive law originally proposed in our previous work [28] is further tailored to update the weights of critic NN online to achieve convergence. In comparison to the existing ADP schemes, this adaptive law is used to directly estimate the unknown NN weights as in [24] rather than minimize the residual Bellman error as [23]. Thus, by means of the Lyapunov method, we prove that the obtained practical control action converges to the neighborhood around the optimal solution. The advantages of the proposed control approach are mainly in the following aspects: firstly, the proposed ADP approach can suitably cope with the modelling uncertainties of the WEC plant, where an online uncertainty estimator is employed; secondly, the ADP-based method is much more computationally efficient than other online optimal control approaches, e.g. dynamic programming (DP) [12] or MPC [9]–[11], [13], [17]. Although recent work [29] presented a RL-based WEC control which can be implemented online to retain the optimal WEC behavior, an offline learning phase with several hours is required. Different from the control method in [29], the ADP control approach proposed in this paper does not require the offline learning phase, which improves its computational efficiency in real time applications. A WEC model is needed for the proposed ADP approach, which is robust to modeling uncertainties.

Numerical simulations based on a typical WEC, called point absorber, are used to demonstrate the efficacy of the proposed ADP control. The point absorber has a relatively high energy conversion rate among many different designs, and is mainly studied in the WEC control community [30]. Simulation results show that the proposed ADP control algorithm is robust to modeling uncertainties and achieves stable energy output.

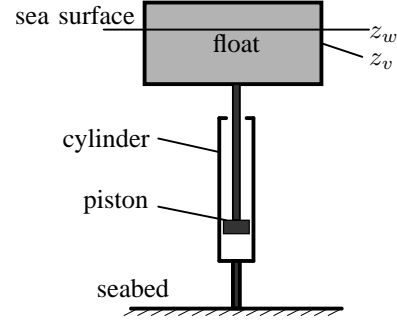The structure of this paper is as follows. The model of the

Fig. 1. Schematic diagram of the point absorber

selected WEC is established in Section II. The ADP approach with the uncertainty estimator is developed in Section III. The simulation results of the ADP are demonstrated in Section IV. The paper is concluded in Section V.

II. MODELLING OF WEC AND PROBLEM FORMULATION

A point absorber type of WEC to be studied in this paper has a float with a constant radius cylinder on the sea surface. Wave energy can be captured using different power take-off (PTO) mechanisms, for example the PTO based on a direct linear generator [31], or the one based on a hydraulic motor and converters [32]. Fig. 1 shows part of a possible hydraulic PTO design: a hydraulic cylinder is vertically installed below the float and is fixed to the bottom of the seabed; more details on this design can be found in [32]. In Fig. 1, z_w and z_v are the water level and the height of the mid-point of the float respectively. The generator's torque is proportional to the force f_u acting on the piston inside the cylinder. The extracted power is $P = -f_u v$, where the velocity on the piston is $v = \dot{z}_v$.

The dynamic equation for the float of a point absorber [33] can be established using Newton's second law

$$m_s \ddot{z}_v = f_s + f_r + f_e + f_u \qquad (1)$$

where m_s is the float mass, z_v is the heave motion of the float, f_r, f_e is the radiation force and excitation force, respectively; f_u is the PTO force acting on the piston. The buoyancy force f_s is calculated by

$$f_s = -K z_v \qquad (2)$$

with K being the stiffness coefficient calculated by $K = \rho g S$. Here, ρ is the water density; g is the gravitational acceleration; S is the water plane area of the floating body. The radiation force f_r is determined by

$$f_r = -\int_{-\infty}^{\infty} h_r(\tau)\dot{z}_v(t-\tau)d\tau - m_\infty \ddot{z}_v \qquad (3)$$

where m_∞ is the added mass, and h_r is the kernel of the radiation force, which can be calculated using hydraulic software packages. Define the integral term as $f_R := \int_{-\infty}^{\infty} h_r(\tau)\dot{z}_v(t-\tau)d\tau$, which can be approximated by a causal finite dimensional state-space model

$$\dot{x}_p = A_p x_p + B_p \dot{z}_v \qquad (4a)$$

$$y_p = C_p x_p \approx \int_{-\infty}^{t} h_r(\tau)\dot{z}_v(t-\tau)d\tau \qquad (4b)$$

with $x_p \in \mathbb{R}^{n_p}$ as the state and $y_p = f_R$.

The wave excitation force f_e can be expressed as

$$f_e = \int_{-\infty}^{\infty} h_f(\tau) z_w(t-\tau) d\tau \qquad (5)$$

which can be written as a state-space model (see [33] for more explanations)

$$\dot{x}_f = A_f x_f + B_f z_w \qquad (6a)$$

$$y_f = C_f x_f(t+t_c) \approx \int_{-\infty}^{t} h_f(\tau) z_w(t-\tau) d\tau \qquad (6b)$$

with $x_f \in \mathbb{R}^{n_f}$ is the state and $y_f = f_e$. Note that f_e is a noncausal term depending on the upstream wave measurement, and t_c in (6b) denotes the causalizing time shift. We use $\tilde{x}_f := x_f(t+t_c)$ to denote the noncausal information needed in the convolution term.

Substituting (2), (3), (4b) and (6b) into (1) gives

$$m\ddot{z}_v = -Kz_v - C_p x_p + C_f x_f + f_u \qquad (7)$$

with $m := m_s + m_\infty$ being the lumped mass.

By defining $x_1 := z_v$, $x_2 := \dot{z}_v$, $y := \dot{z}_v$ and $u := f_u$, we derive a state-space WEC model

$$\dot{x} = Ax + B_u u + B_w w + \xi \qquad (8a)$$

$$y = Cx \qquad (8b)$$

with $x := \begin{bmatrix} x_1, & x_2, & x_p^T, & \tilde{x}_f^\top \end{bmatrix}^\top$, $w = z_w$ and

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ -\frac{K}{m} & 0 & -\frac{C_p}{m} & \frac{C_f}{m} \\ 0 & B_p & A_p & 0 \\ 0 & 0 & 0 & A_f \end{bmatrix}, \quad B_w = \begin{bmatrix} 0 \\ 0 \\ 0 \\ B_f \end{bmatrix},$$

$$C = \begin{bmatrix} 0, & 1, & 0_{1\times n_p}, & 0_{1\times n_f} \end{bmatrix}, \quad B_u = \begin{bmatrix} 0, & \frac{1}{m}, & 0, 0 \end{bmatrix}^\top.$$

where the current wave evolution w is assumed to be measurable.

It is noted that in the WEC model (8a), we include an unknown vector ξ in the formulation, which denotes the effect of uncertainties, modeling errors and disturbances. This term will be explicitly considered in the following control design. Thus, the proposed control in this paper can be extended to more realistic WEC systems with unmodelled dynamic uncertainties.

The WEC control strategy is therefore to solve the following finite-horizon constrained optimal control problem:

$$\min_{u(t)} \int_0^T u(t) y(t) dt \qquad (9)$$
$$\text{s.t. (8a)}, \quad \forall t \in [t_0, +\infty)$$

In this paper, (9) is resolved based on the concept of ADP and the measured current wave evolution $w = z_w$, while the wave prediction algorithm is not needed. Compared with other noncausal optimal control strategies for WECs, which need future wave prediction (e.g. MPC), this ADP control provides a sub-optimal causal control solution. In some scenarios (sea states and devices), the loss of energy can be trivial [34], especially when compared to the benefit of computation load

reduction and the avoidance of the cost and maintenance of wave prediction hardware.

Remark 1: Note that the control performance of some existing model-based advanced control strategies for WECs, e.g. MPC [9], [10], [13] or dynamic programming [12], can heavily rely on the fidelity of the WEC models and have heavy computational cost. Moreover, the existing ADP methods (e.g. [23], [24] and references therein) have been developed for optimal regulation or tracking control problems only, and cannot be directly used for WECs, whose control objective is to achieve energy output maximization. Hence, the main contribution of this paper is to present a new fast online optimal control design methodology for WECs to maximize output energy and effectively handle modeling uncertainties. This has been achieved by introducing a simple uncertainty estimation and further tailoring the idea of ADP.

III. ADAPTIVE OPTIMAL CONTROL DESIGN

This section presents uncertainty estimation and optimal control design based on the ADP concept. The model uncertainties are addressed first by introducing an uncertainty estimator. Moreover, the input u cannot be too big since the torque of the generator has an upper limit, and the heave motion of the float z_v needs also to be limited to prevent damage. We cope with these constraints by tuning weights and adjusting the cost function appropriately.

A. Estimation of uncertainties

In the model (8a), the uncertainties are lumped into an additive uncertain variable ξ, whose effect may degrade control performance, and thus needs to be taken into account in the controller design. We first present a simple estimator to estimate ξ. For this purpose, the following filtered variables based on x, w, and u are defined as

$$\kappa \dot{x}_g + x_g = x, x_g(0) = 0, \qquad (10a)$$

$$\kappa \dot{w}_g + w_g = w, w_g(0) = 0, \qquad (10b)$$

$$\kappa \dot{u}_g + u_g = u, u_g(0) = 0, \qquad (10c)$$

where $\kappa > 0$ is a small positive constant. Then the following estimator can be given as

$$\hat{\xi} = \frac{x - x_g}{\kappa} - Ax_g - B_u u_g - B_w w_g. \qquad (11)$$

We show that the above estimator can achieve exponential error convergence.

Lemma 1: For system (8a) with estimator (11), the estimation error $e_\xi := \xi - \hat{\xi}$ is bounded by $\|e_\xi(t)\| \le \sqrt{e_\xi^2(0)e^{-\frac{t}{\kappa}} + \kappa^2 \varpi^2}$ with $\varpi = \sup_{t>0} \|\dot{\xi}\|$ being the upper bound of $\dot{\xi}$. This implies that $e_\xi(t)$ exponentially converges to a small set. Specifically, one can verify that $\hat{\xi} \to \xi$ for any $\kappa \to 0$ and/or $\varpi \to 0$.

Proof: One can apply a low-pass filter $1/(\kappa s + 1)$ as [28] on both sides of (8a), and then have

$$\dot{x}_g = Ax_g + B_u u_g + B_w w_g + \xi_g, \qquad (12)$$

where ξ_g is the filtered version of ξ given by $\kappa\dot{\xi}_g + \xi_g = \xi$ with $\xi_g(0) = 0$ (Note the variable ξ_g is used for analysis only since ξ is unknown).

Next, we can verify from (10) that $\dot{x}_g = (x - x_g)/\kappa$. This together with (12) imply that $\hat{\xi} = \xi_g$, which means that the proposed estimate given in (11) is equivalent to the filtered version of the unknown uncertainties ξ.

From the fact $\dot{\xi}_g = -\frac{1}{\kappa}\xi_g + \frac{1}{\kappa}\xi = \frac{1}{\kappa}e_\xi$, the estimation error can be given as

$$\dot{e}_\xi = \dot{\xi} - \dot{\xi}_g = \dot{\xi} + \frac{1}{\kappa}\xi_g - \frac{1}{\kappa}\xi = -\frac{1}{\kappa}e_\xi + \dot{\xi}. \quad (13)$$

We choose a Lyapunov function as $V = \frac{1}{2}e_\xi^\top e_\xi$, then its derivative with respect to time t can be calculated along (13) as

$$\dot{V} = e_\xi^\top \dot{e}_\xi \leq -\frac{1}{\kappa}e_\xi^\top e_\xi + \|e_\xi\|\varpi \leq -\frac{1}{\kappa}V + \frac{\kappa}{2}\varpi^2, \quad (14)$$

where ϖ denotes the bound of the lumped uncertainties. This further implies that $V(t) \leq e^{-\frac{t}{\kappa}}V(0) + \frac{\kappa^2}{2}\varpi^2$ and thus $\|e_\xi(t)\| \leq \sqrt{e_\xi^2(0)e^{-\frac{t}{\kappa}} + \kappa^2\varpi^2}$. Then the exponential convergence of $e_\xi(t)$ to zero can also be claimed for $\kappa \to 0$ and/or $\dot{\xi} \to 0$ (and hence $\varpi \to 0$). ∎

It is shown in the above lemma that the estimation $\hat{\xi}$ can exponentially converge to a small set around the true value of the lumped uncertainties ξ. In this case, we can reformulate the original system (8a) as

$$\dot{x} = Ax + B_u u + B_w w + \hat{\xi} + e_\xi, \quad (15)$$

where the estimation error e_ξ is vanishing for sufficiently small κ. In the following control design, we will use the model (15) instead of (8a).

B. Optimal control of WECs

In the control design, to add tuning flexibilities for the magnitude of the control input u and the system state x_1, a modified cost-to-go function is introduced

$$V(x, t) = \int_t^T \left(x_2(\tau)u(\tau) + \frac{\epsilon}{\delta - |x_1(\tau)|} + u(\tau)^\top Ru(\tau)\right)d\tau \quad (16)$$

where $\epsilon > 0$ is a small constant, and $R > 0$ is a positive definite matrix with appropriate dimension.

Remark 2: The first term in (16) represents the output energy; the second term is a barrier function aiming to handle the constraint imposed on the system state [12]; the third term aims to add a tuning parameter R to the control signal, so that a trade-off between the input effort and energy output can be adjusted.

The above optimal control can be solved based on the Pontryagin's minimum principle (PMP) and the dynamic programming method [12]. However, their heavy computational costs may be problematic for practical application when the WEC dynamics need to be described by a high order model. In the following, we present an efficient solution of the optimal control of WEC using the Hamiltonian method.

To solve this optimal control design, we first define a Hamiltonian as follows

$$H(x, u, V, t) = V_x^\top (Ax + B_w w + B_u u + \hat{\xi} + e_\xi) + x_2 u + \frac{\varepsilon}{\delta - |x_1|} + u^\top Ru \quad (17)$$

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

Suppose the optimal cost value under optimal control u^* is $V^*(x, t)$. Then the HJB equation associated with the finite-horizon cost function (16) can be derived (as (3) in [35])

$$V_t^* = -\min_u \left[l(x, u) + V_x^{*\top}(Ax + B_w w + B_u u + \hat{\xi} + e_\xi)\right] \quad (18)$$

where $l(x, u) := x_2 u + \frac{\epsilon}{\delta - |x_1|} + u^T Ru$, $V_t^* := \partial V^*(x(t), t)/\partial t$ and $V_x^* := \partial V^*(x(t), t)/\partial x$.

According to the optimization condition [21], we can solve $\partial H(x, u^*, V^*)/\partial u^* = 0$ for the optimal control

$$u^* = -\frac{1}{2}R^{-1}\left(x_2 + B_u^\top \frac{\partial V^*(x, t)}{\partial x}\right) \quad (19)$$

Since the optimal cost function (16) has a finite-horizon, we know that V^*, V_t^* and V_x^* are dependent on time t [26], [35], and V_t^* appears in the HJB equation (18) though it is not involved in the optimal control (19) explicitly. For this case, the HJB equation (18) is non-linear and with time-varying nature, and thus it is generally difficult or even not possible to find its analytical solution. In the next subsection, we will present an alternative solution by using the principle of ADP, where a critic NN is trained to approximate the optimal cost-to-go function.

C. Online adaptive optimal control via ADP

As shown in the last subsection, the above finite-horizon cost-to-go function $V^*(x, t)$ is time dependent. To solve this finite-horizon optimal control problem, the idea of ADP has recently been explored for specific systems [26], [27], [35]. In [35], a critic NN with time-dependent weights was introduced to approximate the time-dependent cost function (16). However, the time-dependent weights were calculated through a backward integration, which is time-consuming and computationally demanding. Alternatively, a critic NN with a time-varying activation function (taking time-to-go as its input) and constant weights was suggested in [26], [27] to derive an iterative algorithm. In this paper, to implement the resulting control algorithm online, we introduce a critic NN with a time-varying activation function as [26], [27] to avoid inclusion of future waves.

Assuming the optimal value function $V^*(x, t)$ is continuous on a compact set Ω, then as shown in [26], [27], we can represent $V^*(x, t)$ by a critic NN with a time-varying activation function within this compact set as:

$$V^*(x, t) = W^{*\top}\phi(x, T - t) + \varepsilon_n \quad (20)$$

where $W^* \in \mathbb{R}^l$ defines the unknown ideal NN weights, $\phi(x, T - t) = [\phi_1(x, T - t), \cdots, \phi_l(x, T - t)]^\top \in \mathbb{R}^l$ denotes

the time-varying activation function of the state x and the time-to-go $T - t$ as [26], [27]. l is the number of neurons, and ε_n is the NN approximation error.

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

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

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

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 x and t, respectively.

Similar to [22], [26], [27], we make the following assumption on the NN approximation:

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

Note that we can design the critic NN regressor $\phi(x, T - t)$ appropriately so that $\{\phi_i(x, T - t) : i = 1, \ldots, l\}$ can provide a completely independent basis. In this case, according to the Weierstrass theorem [22], the approximation errors of critic NN, e.g. $\varepsilon_n, \nabla\varepsilon_n, \nabla\varepsilon_{nt}$, can converge to zero for $l \to +\infty$.

Then substituting (21) into (19) gives the ideal optimal control u^*

$$u^* = -\frac{1}{2}R^{-1}\left[x_2 + B_u^\top(\nabla\phi^\top(x, T - t)W^* + \nabla\varepsilon_n)\right] \quad (23)$$

The optimal NN weight W^* is unknown, and thus the practical critic NN used to approximate the optimal value function is provided as

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

where $\hat{W}$ denotes the estimated NN weights of W^*. Then from the estimated cost function (24), the approximated optimal control u used for the practical control implementation is

$$u = -\frac{1}{2}R^{-1}(x_2 + B_u^\top\nabla\hat{V}_x(x,t)) \quad (25)$$

where $\nabla\hat{V}_x := \nabla\phi^\top(x, T - t)\hat{W}$ is defined as the partial derivative of $\hat{V}(x,t)$ with respect to x based on (24).

Now the remaining issue is to determine the estimated NN weights $\hat{W}$, which need to converge to the ideal NN weight W^*. In this paper, an adaptive law is developed to online update the NN weights $\hat{W}$, which can guarantee the convergence of $\hat{W}$ to a neighborhood of W^*. This can be achieved by further extending the idea of our previous work [24], [28] and considering the Hamiltonian. To facilitate the analysis, we substitute the critic NN (21) and (22) into the HJB equation (18) and obtain:

$$W^{*\top}\nabla\phi_t(x, T - t) + W^{*\top}\nabla\phi(x, T - t)[Ax + B_ww + B_uu$$
$$+ \hat{\xi}] + x_2u + \frac{\epsilon}{\delta - |x_1|} + u^\top Ru + \varepsilon_{HJB} = 0$$
$$(26)$$

We define $\varepsilon_{HJB} := \nabla\varepsilon_n^\top(Ax + B_ww + B_uu + \hat{\xi} + e_\xi) + W^{*\top}\nabla\phi(x, T - t)e_\xi + \nabla\varepsilon_{nt}$ to denote the residual HJB equation error. This error comes from the NN errors $\nabla\varepsilon_n$,

$\nabla\varepsilon_{nt}$ and estimation error e_ξ, which are all bounded. Hence, ε_{HJB} is also bounded.

To design an adaptive law based on (26), the known terms are denoted as $\Xi := \nabla\phi_t + \nabla\phi(Ax + B_ww + B_uu + \hat{\xi})$ and $\Theta := x_2u + \frac{\epsilon}{\delta - |x_1|} + u^\top Ru$, and then equation (26) can be represented as

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

As shown in (26), the critic NN weights W^* are linearly parameterized subject to the known dynamics Ξ. Then, we denote the auxiliary matrix P, auxiliary vector Q as follows

$$\begin{cases} \dot{P} &= -\gamma P + \Xi\Xi^\top, \quad P(0) = 0 \\ \dot{Q} &= -\gamma Q + \Xi\Theta, \quad Q(0) = 0 \end{cases} \quad (28)$$

where $\gamma > 0$ is a forgetting factor parameter.

Another auxiliary vector $M \in \mathbb{R}^l$ is defined based on P and Q as

$$M = P\hat{W} + Q \quad (29)$$

The adaptive law for updating the unknown NN weights $\hat{W}$ is proposed as

$$\dot{\hat{W}} = -\Gamma M \quad (30)$$

where $\Gamma > 0$ is the learning gain, which is set as a constant diagonal matrix.

To prove the convergence of the adaptive algorithm (30), the following lemma is needed:

Lemma 2: The auxiliary variable defined in (29) is equivalent to

$$M = -P\tilde{W} + \varphi \quad (31)$$

where $\tilde{W} = W - \hat{W}$ is the critic NN weight error, and $\varphi = -\int_0^t e^{-\gamma(t-r)}\varepsilon_{HJB}(r)\Xi(r)dr$ denotes the effect of the bounded HJB residual error, i.e. $\|\varphi\| \le \varepsilon_N$ holds for $\varepsilon_N > 0$.

Proof: One can solve the matrix equation (28) and obtain its solution as

$$\begin{cases} P(t) &= \int_0^t e^{-\gamma(t-r)}\Xi(r)\Xi^\top(r)dr \\ Q(t) &= \int_0^t e^{-\gamma(t-r)}\Xi(r)\Theta(r)dr \end{cases} \quad (32)$$

From (32), we have $Q = -PW^* + \varphi$. It is noted $\varphi = -\int_0^t e^{-\gamma(t-\tau)}\varepsilon_{HJB}(\tau)\Xi(\tau)d\tau$ is also bounded because the NN activation functions and the NN approximation errors are all bounded, that is $\|\varphi\| \le \varepsilon_N$. Then by substituting (32) into (29), one can obtain (31). ∎

It is clearly shown in Lemma 1 that the suggested auxiliary variable M in the adaptive law includes the NN weight error $\tilde{W}$ perturbing by a small bounded residual variable φ, such that the NN weight estimation convergence can be achieved using the adaptive law (30) driven by the variable M [24].

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

Lemma 3: [28] The condition $\lambda_{\min}(P) > \sigma > 0$ holds for any constant $\sigma > 0$ (i.e. P is positive definite) provided that Ξ defined in (26) is persistently excited (PE).

Now, we have the main theoretical result of this paper as follows:

Theorem 1: Consider the adaptive law (30) with Ξ being PE, then the weight error $\tilde{W}$ will ultimately converge to a neighborhood of zero. Specifically, when $\nabla\varepsilon_n, \nabla\varepsilon_{nt} = 0$ (i.e. the NN errors are zero), $\tilde{W}$ exponentially converges to zero. Moreover, the practical control (25) will converge to a neighborhood of the optimal solution (23).

Proof: We select a Lyapunov functional as $V = \frac{1}{2}\tilde{W}^\top\Gamma^{-1}\tilde{W}$, and calculate its time derivative along (30) as

$$\dot{V} = \tilde{W}^\top\Gamma^{-1}\dot{\tilde{W}} = -\tilde{W}^\top P\tilde{W} + \tilde{W}^\top\varphi \quad (33)$$

The residual error φ is bounded since the unknown NN errors and critic NN regressor are all bounded. Then the equation (33) can be presented as

$$\dot{V} = -\tilde{W}^\top P\tilde{W} + \tilde{W}^\top\varphi \leq -\|\tilde{W}\|(\sigma\|\tilde{W}\| - \varepsilon_N) \quad (34)$$

Based on the extended Lyapunov theorem, we can show that $\tilde{W}$ can uniformly converge to a set defined by $\Omega_c = \{\tilde{W}\|\|\tilde{W}\| \leq \varepsilon_N/\sigma\}$. whose size is determined by the bound of NN error ε_N and PE level σ. For the specific case with $\varepsilon_{HJB} = 0$ and thus $\varphi = 0$, we know (34) can be reduced to $\dot{V} = -\tilde{W}^\top P\tilde{W} < -\sigma\|\tilde{W}\|^2 \leq -\mu V$, which implies exponential convergence of $\tilde{W}$ to zero.

Based on the above analysis, one can find that the estimated NN weights $\hat{W}$ converge to a small set around the ideal NN weights W^*. Moreover, the NN estimation error $\nabla\varepsilon_n$ is also bounded. Consequently, it can be verified that the practical control (25) will converge to a neighborhood of the optimal solution (23). ∎

The above analysis based on Lyapunov's method for the convergence of the estimator and critic NN are conducted for the time going to infinity. However, as explained in [36] and other ADP references (e.g. [27]), this type of analysis can be used for finite horizon control problems though the final time is bounded in this case.

Remark 3: Lemma 3 states that the well-known PE condition is sufficient to guarantee the required condition $\lambda_{\min}(P) > \sigma > 0$, which is necessary to the proof of convergence of (30). This condition can be true for generic WEC systems since there is external sea wave input w perturbing the system (8a). In particular, it is possible to numerically online verify this condition by testing the minimum eigenvalue of P as shown in [28].

Remark 4: The stability of WEC control is an important but not fully addressed problem in the literature. In the case study, this approximated optimal solution can stabilize the WEC system within the reasonable ranges of device parameters, because the WEC system (8a) in our study is inherently stable (i.e. the matrix A is Hurwitz), and the sea wave w is bounded.

IV. SIMULATIONS

This section presents simulation results based on a point absorber to validate the proposed online optimal control strategies. The nominal values of the WEC model's parameters are exactly the same with those in [33], which are summarized in Table I. We adopt this WEC model from [33] because it has a relatively high energy conversion rate among many

TABLE I
NOMINAL SYSTEM PARAMETERS.

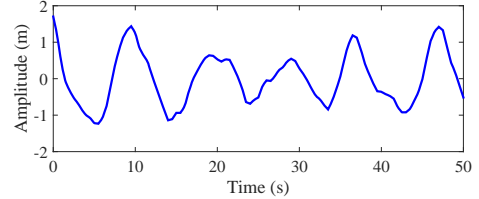| Parameters	Values
r	0.35 (m)
m_s	242 (kg)
m_∞	83.5 (kg)
K	3775.3 (N/s)
ρ	1000 (kg/m^3)
g	9.81 (m/s^2)

Fig. 2. A 50 seconds period of wave profile (wave elevation).

different designs, and has been widely studied in the WEC control community.

The state-space description of the impulse function for calculating the wave excitation force is

$$A_f = \begin{bmatrix} 0 & 0 & 0 & 0 & -400 \\ 1 & 0 & 0 & 0 & -459 \\ 0 & 1 & 0 & 0 & -226 \\ 0 & 0 & 1 & 0 & -64 \\ 0 & 0 & 0 & 1 & -9.96 \end{bmatrix} \quad B_f = \begin{bmatrix} 1549886 \\ -116380 \\ 24748 \\ -644 \\ 19.3 \end{bmatrix}$$

$$C_f = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

The state-space description of the impulse functions for the radiation force is

$$A_p = \begin{bmatrix} 0 & 0 & -17.9 \\ 1 & 0 & -17.7 \\ 0 & 1 & -4.41 \end{bmatrix} \quad B_p = \begin{bmatrix} 36.5 \\ 394 \\ 75.1 \end{bmatrix} \quad C_p = \begin{bmatrix} 0 & 0 & 1 \end{bmatrix}$$

A realistic sea wave heave trajectory gathered from the coast of Cornwall, UK, is used in the simulations. For demonstration purpose, only 50 secs of the collected wave data are used, as shown in Fig. 2. The magnitude of this wave profile is small and its maximum is less than 2 m.

In the simulations, the parameters for the ADP algorithm were chosen as: $\phi(x, t_n) = [x_1 t_n^3, x_2 t_n^3, 0.5x_1^2 t_n^5, 0.5x_2^2 t_n^5, x_1 x_2 t_n^5]^\top$ with $t_n = (T - t)/T$ and $T = 50$ being the normalised time-to-go, $\hat{W}(0) = [0, 0, 0, 0, 0]^\top$, $\varepsilon = 1$, $\delta = 1$, $\Gamma = \text{diag}([0.5, 5, 5, 5, 5])$ and $\gamma = 10$. The design of the activation function of the critic NN is usually based on engineering experience as [26], [27]. The parameter R is used for tuning the magnitude of control input depending on the limit of the PTO control force. In general, a big value of R can penalize the required control actions, which in turn helps to retain safe operations of a WEC with a small limit of PTO control force and a large magnitude of sea wave profile. To demonstrate this effect, we choose different values for R in our simulations for comparison purpose.

Moreover, to verify the efficacy of the proposed estimator and test the robustness of the ADP controller, we assume

model mismatch exists between the WEC model and the model used in the ADP controller. In this case, the WEC model is designed based on the above nominal parameters, while these parameters in the ADP controller model have different values. In the simulations, the model mismatch due to the variation of the mass m_∞ depending on the radiation force variation is taken into account. We can also assume parameter variation for the damping ratio in a similar way.

The simulations were conducted in the following 3 cases.

Case 1: No model mismatch between the WEC model and the controller model. Different R values are tested.

Case 2: With model mismatch. The added mass used in the control is 18.4 kg which is 78% less than the nominal value of WEC added mass. Thus the total mass used in the control design is $242 + 18.4 = 260.4$ kg, which is 20% less than the nominal value of WEC total mass. $R = 1 \times 10^{-4}$ is chosen.

Case 3: With model mismatch and estimator. The added mass used in the control is 18.4 kg and the estimator (11) is used to estimate and compensate for the parametric uncertainties in the control. $R = 1 \times 10^{-4}$ is chosen.

Fig. 3 shows the power outputs and energy outputs of Case 1) with $R = 1 \times 10^{-4}$. It is shown that the proposed control algorithm can lead to stable energy output, which indicates its efficacy. Moreover, the generated power approaches to negative values at some time instant, which indicates a power flow from the grid to the ocean may happen. This bi-directional power flow can be achieved by some hardware design, e.g. [32], and it can generate more energy than the one-directional power flow from ocean to the grid, as demonstrated in e.g. [17]. However, since in these simulations, the negative power flow only happens occasionally, this means that if we restrict the PTO to unidirectional power flow, the energy loss cannot be significant.

To specifically show the effect of the tuning parameter R on the control response, Fig. 4 gives the tendency of the maximum control input ($|u|$), maximum float heave ($|x_1|$) and energy output with different R, from which, one may find that with the increase of R the maximum control input decreases, which indicates that the control constraints on u can be satisfied using large R. However, the maximum float heave increases for larger R, which may cause device damage (in this case, the parameter δ should be increased). Hence, the trade-off between the control constraints and operation safety should be considered in the control implementation. Moreover, one may find from Fig.4 that the generated energy increases with R when $R < 1 \times 10^{-4}$, but decreases when $R > 1 \times 10^{-4}$. This is because the addressed WEC optimal control problem with the given optimal cost function may be nonconvex for small R. The proposed optimal control using the concept of ADP can solve this issue even when R is small, which cannot be easily resolved using conventional MPC method, e.g. [13].

Figs 5-7 illustrate the control input trajectories, float heave trajectories, and energy outputs of the studied three cases, respectively. To guarantee energy output, we use $R = 1 \times 10^{-4}$ with maximum energy as shown in Fig.4. This group of comparisons are dedicated to validate the effectiveness of the proposed uncertainty estimator. One may find that the trajectories of Case 1 and Case 3 are very close because
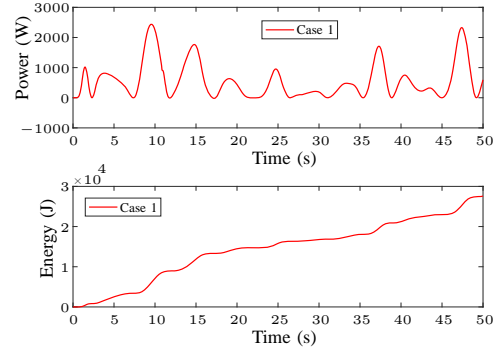

Fig. 3. The energy outputs of Case 1.

TABLE II
COMPARISON OF CONTROL PERFORMANCE OF ADP IN DIFFERENT SCENARIOS.

	Case 1	Case 2	Case 3
Energy output (J)	27598	26609	27354
Maximum input (N)	3321	3005	3294
Minimum input (N)	-3746	-3615	-3585
Maximum heave (m)	1.00	0.96	0.99
Minimum heave (m)	-0.96	-0.91	-0.95

the modeling uncertainty due to the variation of the added mass can be precisely estimated using the proposed uncertainty estimator (as shown in Fig.8). Hence, the mismatch between the nominal WEC plant and the model used in the control design can be compensated when the estimated uncertainty is incorporated into the ADP control implementation. Consequently, the generated power and energy for Case 1 and Case 3, as shown in Fig.7, are very similar. However, for Case 2 where the model mismatch is not considered (the uncertainty estimator is turned off) in the control design, the generated power and energy are smaller than both Case 1 and Case 3. This result clearly indicates that the proposed estimator can improve the robust performance of the controller. Finally, Fig. 9 provides the profile of the critic NN weights, which can achieve convergence after a transient period, illustrating the efficacy of the proposed adaptive law.

The above mentioned conclusions can also be demonstrated using a quantitative analysis. For this purpose, Table II shows the maximum and minimum values of the control inputs and float heaves, and the energy outputs of the three cases, respectively. Again, it is shown that the ADP control with estimator (Case 3) outperforms Case 2 when the estimator is not used.

In the simulation, we also note that the computational burden for implementing the developed ADP control is very small even for this 10th-order WEC system, since it can be implemented online without any offline learning phases. The low computational burden of the proposed ADP control is one of the main advantages over the MPC controllers developed for the WEC control problem. Moreover, since the ADP algorithm does not require the future wave information, which requires a wave prediction technique, economic hardware is sufficient for the implementation of the proposed algorithm in this paper.
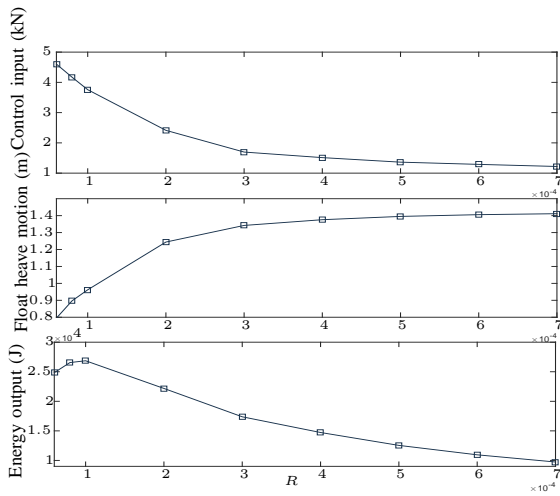
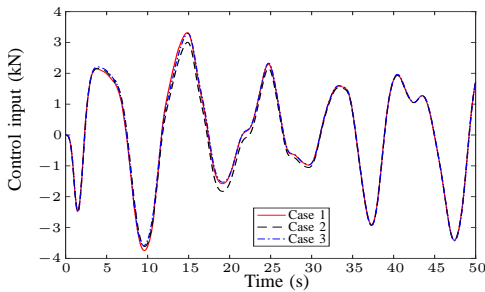Fig. 4. The maximum input, float heave and energy output with different R.
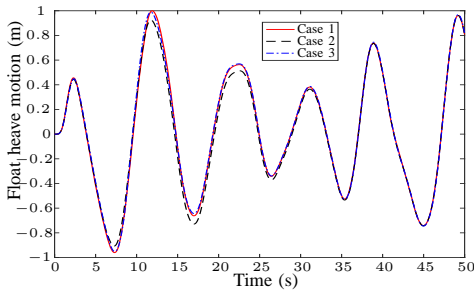

Fig. 5. The input trajectories.
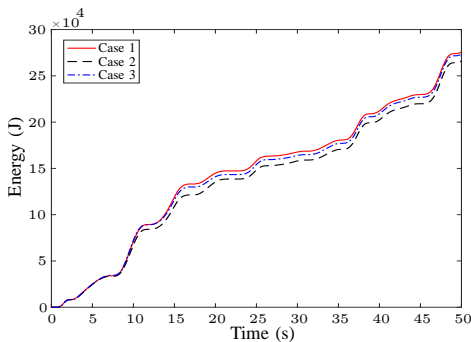

Fig. 6. The float heave motion trajectories.
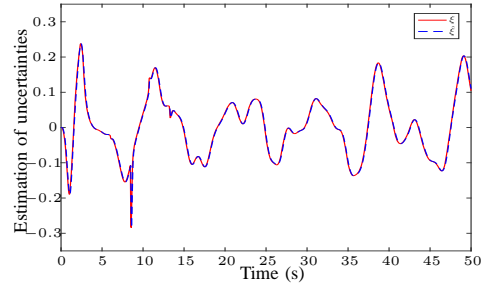

Fig. 7. Energy outputs.

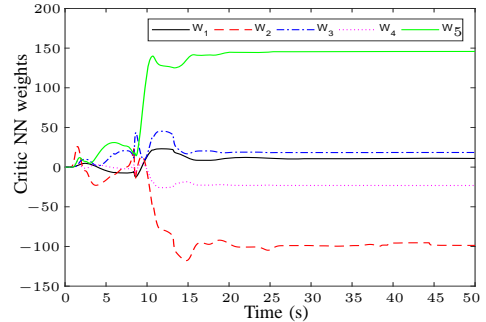Fig. 8. The uncertainty and estimation of Case 3.

Fig. 9. Convergence of the critic NN weight $\hat{W}$ of Case 3.

V. CONCLUSION

In this paper, a new optimal control strategy based on ADP is proposed for a particular WEC device (point absorber). The energy maximization control problem is solved by introducing a critic NN with time-to-go as its input to approximate the unknown time-varying cost function. Modeling uncertainties are online estimated using a simple yet robust uncertainty estimator, which is incorporated into the control design. To achieve a fast convergence rate, we use a new adaptive law to online update the critic NN weights, which is driven by the obtained NN weight error information. Numerical simulations demonstrate that the proposed ADP method can handle uncertainties effectively, which helps to maintain the control performance of the WECs. The proposed approach has the potential to be extended to the control of other types of WECs, and even other energy maximization control problems. It is also noted that the proposed approach has some tuning parameters to cope with input saturation. In the future work, we will explicitly incorporate the control constraints into the ADP optimization algorithm.

REFERENCES

[1] I. Glendenning, "Ocean wave power," *Applied Energy*, vol. 3, no. 3, pp. 197 – 222, 1977.
[2] 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.
[3] "The UK Renewable Energy Roadmap," *https://www.gov.uk/government/publications/renewable-energy-roadmap*, 2011.
[4] D. Bull and M. E. Ochs, "Technological cost-reduction pathways for attenuator wave energy converters in the marine hydrokinetic environment," Sandia National Laboratories, Tech. Rep., 2013, SAND2013-7207, Unlimited Release, Sandia Report.

[5] 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.

[6] K. Budal and J. Falnes, "Optimum operation of improved wave-power converter," *Marine Science Communciations*, vol. 3, no. 2, pp. 133–150, 1977.

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

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

[9] T. K. A. Brekken, "On model predictive control for a point absorber wave energy converter," in *PowerTech, 2011 IEEE Trondheim*, June 2011, pp. 1–8.

[10] 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.

[11] 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.

[12] 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.

[13] 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.

[14] S. Zhan and G. Li, "Linear optimal noncausal control of wave energy converters," *IEEE Transactions on Control Systems Technology*, pp. 1–11, 2018.

[15] R. Genest, J. Davidson, and J. V. Ringwood, "Adaptive control of a wave energy converter," *published online, IEEE Transactions on Sustainable Energy*, 2018.

[16] 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. 9, no. 3, pp. 1419–1426, 2018.

[17] G. Li, "Nonlinear model predictive control of a wave energy converter based on differential flatness parameterisation," *International Journal of Control*, vol. 90, pp. 68–77, 2017.

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

[19] S. G. Khan, G. Herrmann, F. L. Lewis, T. Pipe, and C. Melhuish, "Reinforcement learning and optimal adaptive control: An overview and implementation examples," *Annual Reviews in Control*, vol. 36, no. 1, pp. 42 – 59, 2012.

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

[21] F. L. Lewis, D. L. Vrabie, and V. L. Syrmos, "Reinforcement learning and optimal adaptive control," *Optimal Control, Third Edition*, pp. 461–517, 2012.

[22] 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.

[23] 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.

[24] 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.

[25] Y. Lv, J. Na, Q. Yang, X. Wu, and Y. Guo, "Online adaptive optimal control for continuous-time nonlinear systems with completely unknown dynamics," *International Journal of Control*, vol. 89, no. 1, pp. 99–112, 2016.

[26] 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.

[27] 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, pp. 486–499, 2015.

[28] 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.

[29] E. Anderlini, D. I. M. Forehand, P. Stansell, Q. Xiao, and M. Abusara, "Control of a point absorber using reinforcement learning," *IEEE Transactions on Sustainable Energy*, vol. 7, no. 4, pp. 1681–1690, Oct 2016.

[30] B. Drew, A. R. Plummer, and M. N. Sahinkaya, "A review of wave energy converter technology," *Proc. IMechE Part A: J. Power and Energy*, vol. 223, pp. 887–902, 2009.

[31] S.-R. Huang, H.-T. Chen, C.-H. Chung, C.-Y. Chu, G.-C. Li, and C.-C. Wu, "Multivariable direct-drive linear generators for wave energy," *Applied Energy*, vol. 100, pp. 112 – 117, 2012.

[32] 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.

[33] Z. Yu and J. Falnes, "State-space modelling of a vertical cylinder in heave," *Applied Ocean Research*, vol. 17, pp. 265–275, 1995.

[34] 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.

[35] 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.

[36] A. Heydari and S. N. Balakrishnan, "Closed form solution to finite-horizon suboptimal control of nonlinear systems," *International Journal of Robust and Nonlinear Control*, vol. 25, no. 15, pp. 2687–2704, 2015.

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.

Guang Li (M'09) received his PhD degree in Electrical and Electronics Engineering, specialized in control systems, from the University of Manchester, in 2007. He is currently a 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.

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.

Guido Herrmann (SM'06) received the Ph.D.degree from the University of Leicester, Leicester, U.K., in 2001. From 2001 to 2003, he was a Senior Research Fellow with the AStar Data Storage Institute in Singapore. From 2003 to 2007, he was a Research Associate, Fellow, and Lecturer at the University of Leicester. In 2007, he joined the University of Bristol, Bristol, U.K., as a Lecturer, became a Senior Lecturer in 2009, and then a Reader in Control and Dynamics at the Department of Mechanical Engineering in 2012. His current research interests include the development and application of novel, robust and nonlinear control systems. Dr. Herrmann is a Fellow of the IET and has been awarded together with Jing Na the 2017 Hsue-shen Tsien Paper Award.

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.