Integrated Path Planning and Tracking Control of Marine Current Turbine in Uncertain Ocean Environments

Arezoo Hasankhani, *Student Member, IEEE,* Ertugrul Baris Ondes, Yufei Tang, *Member, IEEE,*
Cornel Sultan, and James VanZwieten

Abstract— This paper presents an integrated path planning and tracking control of marine hydrokinetic energy harvesting devices. To address the highly nonlinear and uncertain oceanic environment, the path planner is designed based on a reinforcement learning (RL) approach by fully exploring the historical ocean current profiles. The planner will search for a path to optimize a chosen cost criterion, such as maximizing the total harvested energy for a given time. Model predictive control (MPC) is then utilized to design the tracking control for the optimal path command from the planner subject to problem constraints. The planner and the tracking control are accommodated in an integrated framework to optimize these two parts in a real-time manner. The proposed approach is validated on a marine current turbine (MCT) that executes vertical waypoint path searching to maximize the net power due to spatiotemporal uncertainties in the ocean environment, as well as the path following via an MPC tracking controller to navigate the MCT to the optimal path. Results demonstrate that the path planning increases harvested power compared to the baseline (i.e., maintaining MCT at an equilibrium depth), and the tracking controller can successfully follow the reference path under different shear profiles.

I. INTRODUCTION

Research in marine hydrokinetic energy harvesting devices, including wave energy converter (WEC) and marine current turbine (MCT), is motivated today in the academic community and industry [1]. Similar to an autonomous underwater vehicle (AUV), the MCT should autonomously navigate along the optimal path in the presence of a highly nonlinear and uncertain oceanic environment, accounting for the fact that the predicted ocean velocity is affected by perturbations. An effective real-time path control may entail several prerequisites, leading to a challenging problem. The trajectory control should track the optimal path to minimize the tracking error, which should ideally converge to zero. Meanwhile, the optimal path is acquired by optimizing an ultimate goal, i.e., maximizing harvested energy from an MCT in a given time period. Moreover, the optimal path

may change continuously due to the inherent spatiotemporal uncertainties in the oceanic environment [2]. Finally, the highly nonlinear dynamic model of the MCT that includes multi-physics couplings increases the problem complexity. Therefore, the challenges of real-time optimal path control is significant for this system.

The real-time path control is interpreted with a prohibitive large-scale control program, involving a complex dynamic model of the system, discrete inputs, discrete outputs that enforce waypoint tracking constraints while iteratively optimizing a chosen cost criterion via path planning. The path tracking problem has been addressed for a wide range of applications, including robotics [3], [4], unmanned aerial vehicles [5], [6], helicopters [7], autonomous underwater vehicles [8], and ocean current turbine [9]. For effective real-time path control design, the first task is to ensure that an optimal path is computed by the planner for the tracking controller as a reference path.

An integrated framework for planning and tracking control usually includes an upper-level planner to generate an optimal path and a lower-level tracker that is responsible for tracking the generated optimal path. For examples, the path planning and tracking control have been successfully proposed for airborne wind energy systems [10], [11], [12], [13], [14], [15], discrete manufacturing plants [16], autonomous vehicle [17], [18], unmanned aerial vehicle [19], [20], and autonomous underwater vehicle [21], [22]. In similar research, the trajectory planning and control have been solved for a buoyancy controlled device for an underwater vehicle to achieve its vertical maneuvering [23]. However, to be computationally tractable, the system model and tracking controller have been significantly simplified.

In this paper, to deal with the real-time path control for marine energy harvesting devices, we propose an integrated approach that iteratively plans the optimal path to maximize the harvested power and tracks the commanded path to minimize the tracking error. The motivation is clear – in order to maximize the energy harvesting of the device, the system vertical path should be real-time planned and tracked in the dynamic oceanic environment. Moreover, an integrated framework is intended as a simpler, computationally cheaper alternative to simultaneous path planning and tracking. Specifically, we develop a path planner that allocates an optimal path to maximize the harvested energy from an MCT using a deep reinforcement learning (DRL) approach. Model predictive control (MPC) is then developed to track the optimal reference path commanded from the

A. Hasankhani and Y. Tang are with the Department of Computer & Electrical Engineering and Computer Science, Florida Atlantic University, Boca Raton, FL 33431 USA. {ahasankhani2019, tangy}@fau.edu.

E. B. Ondes and C. Sultan are with the Department of Aerospace and Ocean Engineering, Virginia Tech, 460 Old Turner St., Blacksburg, VA 24061, USA. {ondes, csultan}@vt.edu.

J. VanZwieten is with the Department of Civil, Environmental, and Geomatics Engineering, Florida Atlantic University, Boca Raton, FL 33431, USA. jvanzwi@fau.edu.

planner subject to the MCT dynamic model and problem constraints. The ultimate goal of the tracking control is to safely navigate the turbine along the optimal path, considering the sluggish dynamics while avoiding any aggressive motion and instability.

The paper begins with Section II describes the modeling and overall structure of the integrated path planning and tracking control for the MCT. Section III presents detailed path planning and path tracking. Section IV shows the numerical results. Finally, Section V concludes the paper.

II. PROBLEM DESCRIPTION

In this section, the overall problem is formulated for an application of the MCT, discussing the model of the turbine, oceanic uncertain environment, and the proposed path planning and control architecture for the MCT system.

A. MCT Modeling

The investigated 700 kW MCT in this paper (see Fig. 1) is a representative design for the Gulf Stream, which consists of two variable buoyancy sections of a single variable buoyancy tank, variable pitch rotor, main body, and a 607 m mooring cable attached to the ocean floor at a depth of 325 m [24], [25], [26], [27]. The designed MCT system and its parameter roughly follow the prototype systems from IHI Corp. [28] and the University of Naples [29]; still, the MCT includes a single rotor, where the tanks designed to operate at a depth of 50 m in case of half-filled with ballast water at the ocean current speed of 1.6 m/s. A seven degrees-of-freedom (DOF) model is used to characterize the nonlinear dynamics of the MCT.

Kinematics and Coordinate Frame: In order to obtain the kinematics, the coordinate frames are defined for the MCT system: the inertial coordinate frame ($\mathcal{T}_I$), the body-fixed coordinate frame ($\mathcal{T}_B$), the momentum mesh coordinate frame ($\mathcal{T}_M$), the shaft coordinate frame ($\mathcal{T}_S$), and the rotor blade coordinate frame ($\mathcal{T}_R$). The transformation matrix from $\mathcal{T}_I$ to $\mathcal{T}_B$ is defined due to rotations about the yaw angle ψ, the pitch angle θ, and the roll angle ϕ:

$$T^{\mathcal{T}_B}_{\mathcal{T}_I} = \begin{bmatrix} c_\psi c_\theta & s_\psi c_\theta & -s_\theta \\ c_\psi s_\theta s_\phi - s_\psi c_\phi & c_\psi c_\phi + s_\psi s_\theta s_\phi & c_\theta s_\phi \\ c_\psi s_\theta c_\phi - s_\psi s_\phi & -c_\psi s_\phi + s_\psi s_\theta c_\phi & c_\theta c_\phi \end{bmatrix} \quad (1)$$

where $s_{(.)} = sin(.)$ and $c_{(.)} = cos(.)$.

Equations of Motion: The motion of the buoyancy controlled MCT is modeled through the 7-DOF, consisting of 6-DOF motion of the main body and 1-DOF of rotor's rotation about the X-axis [30], namely:

$$f_x = m^v(\dot{u} - vr + wq) - m^v x_{cg}(q^2 + r^2) + m^v_b z^v_{cg_b}(p_b r + \dot{q}) \quad (2)$$

$$f_y = m^v(\dot{v} + ur) - w(m^v_b p_b + m^v_r p_r) + m^v_b z^v_{cg_b}(q_r - \dot{p}_b) \\ + m^v_b x^v_{cg_b} qp_b + m^v_r x^v_{cg_r} qp_r + m^v x^v_{cg}\dot{r} \quad (3)$$

$$f_z = m^v(\dot{w} - uq) + v(m^v_b p_b + m^v_r p_r) + m^v_b z^v_{cg_b}(p_b^2 + q^2) \\ + m^v_b x^v_{cg_b} rp_b + m^v_r x^v_{cg_r} rp_r + m^v x^v_{cg}\dot{q} \quad (4)$$

$$M_{x_b} + \tau_{em} = -m^v_b z^v_{cg_b}(\dot{v} - wp_b + ur) + I^v_{x_r}\dot{p}_r + qr(I^v_{z_r} - I^v_{y_r}) \quad (5)$$
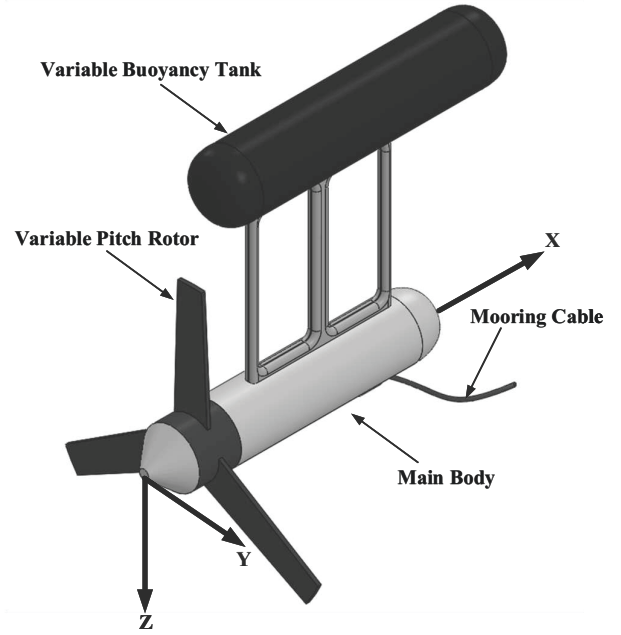

Fig. 1: Schematic of the investigated MCT in this paper [24].

$$M_y = I^v_y \dot{q} + rp_b(I^v_{x_b} - I^v_{z_b}) + rp_r(I^v_{x_r} - I^v_{z_r}) + I^v_{xz_b}(p_b^2 - r^2) \\ + m^v_b z^v_{cg_b}(\dot{u} - vr + wq) - m^v x^v_{cg}(\dot{w} - uq) + m^v_b x^v_{cg_b} vp_b \quad (6) \\ - m^v_r x^v_{cg_r} vp_r$$

$$M_z = I^v_z \dot{r} + qp_b(I^v_{y_b} - I^v_{x_b}) + qp_r(I^v_{y_r} - I^v_{x_r}) + I^v_{xz_b}(rq - \dot{p}_b) \\ + m^v x^v_{cg}(\dot{v} + ur) - m^v_b x^v_{cg_b} wp_b - m^v_r x^v_{cg_r} wp_r \quad (7)$$

$$\dot{p}_r = \frac{M_{x_r} - M_{x_s} - qr(I^v_{z_r} - I^v_{y_r})}{I^v_{x_r}} \quad (8)$$

where m is the mass; u, v, w are the linear velocities of the MCT main body in $\mathcal{T}_B$; p_b, q, r are the angular velocities of the MCT main body in $\mathcal{T}_B$; p_r denotes angular velocity of the MCT rotor in $\mathcal{T}_B$; x, y, z are location of the origin of the MCT main body; ϕ, θ, ψ are the Euler angles of the MCT main body; $f_{(.)}$ is the force about (.); $M_{(.)}$ is the moment about (.); $I_{(.)}$ denotes the mass moment; $(.)^v$ denotes the virtual; τ_{em} is the electromechanical shaft torque; $(.)_{cg}$ denotes the center of gravity; Finally, $(.)_r$ and $(.)_b$ denote the rotor and the main body, respectively.

The total external load f, which is shown about $x-$, $y-$, and $z-$ axes in the equations of motion, is the sum of gravitational and buoyancy force f_{gb}, rotor force f_r, main body forces f_b, and cable force f_c, i.e., $f = f_{gb} + f_r + f_b + f_c$. Cable forces are calculated using a finite element lumped mass cable model and rotor forces are calculated using a blade element momentum based rotor model, with a dynamic wake inflow algorithm (see [30] for details).

Linear Model: The linearization process averages the MCT dynamics over the rotor rotation to remove the dependence on rotor azimuth angle and cable node states as suggested in [31] and is expanded to account for variable buoyancy control as proposed in [32] (with the justification between the linear and nonlinear models). The nonlinear dynamic model of the MCT is linearized around the equilibrium

points, namely:

$$\Delta \dot{x} = A\Delta x + B\Delta u$$
$$y = C\Delta x \tag{9}$$

where $A = \frac{\partial f}{\partial x}\big|_{u=u_{eq}}^{x=x_{eq}}$, $B = \frac{\partial f}{\partial u}\big|_{u=u_{eq}}^{x=x_{eq}}$, $C = \frac{\partial g}{\partial u}\big|_{u=u_{eq}}^{x=x_{eq}}$, $\Delta(.) \triangleq (.) - (.)_{eq}$, with *eq* denoting the equilibrium values. In order to maintain the simplicity, Δ is not shown for the remainder of the paper. $x \in \mathbb{R}^{13}$, $u \in \mathbb{R}^3$, and $y \in \mathbb{R}$ are defined as $x = [u\ v\ w\ p_b\ p_r\ q\ r\ x\ y\ z\ \phi\ \theta\ \psi]$, $u = [B_f\ B_a\ \tau_{em}]$, and $y = [z]$. The control inputs are determined by two buoyancy tank fill fractions B_f and B_a denoting the fill fractions of the forward tank and the fill fraction of the aft tank, respectively, introducing into equations of motion by gravitational and buoyancy force f_{gb}.

Furthermore, since the real-time path controller is designed in discrete-time with future objectives of implementation in a real plant (e.g., an MCT-based power plant), the linearized model is discretized using a sampling time, T_s, determined by the Nyquist-Shannon's sampling theorem [33]. The corresponding discrete linear time-invariant (DLTI) plant model at discrete-time instant k is:

$$x(k+1) = A_d x(k) + B_d u(k)$$
$$y(k) = C_d x(k) \tag{10}$$

with equilibrium points of $x_{eq} = [0\ 0\ 0\ 0\ 1.49\ 0\ 0\ 554.50\ 0.38\ 50\ 0.01\ 0.00\ 3.14]$ and $u_{eq} = [0.4677\ 0.4677\ -188280]$. It should be noticed regarding the determination of the equilibrium points that the non-linear simulation is run until all accelerations are zero and the MCT is no longer moving.

B. Ocean Environment Modeling

To model the spatiotemporal uncertainties in the ocean environment, it is important to use real data. The spatial and temporal fluctuations in the current flow are resulted from turbulence, waves, and lower frequency flow structures. In this paper, we use the field measured data by a 75 kHz acoustic Doppler current profiler (ADCP) presented in [34], which were recorded at a latitude of $26.09°N$ and longitude of $-79.80°E$, as shown in Fig. 2. The measurement resolution was 5 m within 400 m water depth, where various current flow data (i.e., current speed, northward current velocity, eastward current velocity, etc.) were recorded. These measured current data were filtered to remove bad data as described in [34], which were mostly measured above a depth of 50 m.

MCT will operate spatiotemporally in this dynamic environment to harness the ocean current energy. To address its real-time path planning and tracking, it is necessary to model and predict the uncertain ocean environment using the recorded ADCP data. In this paper, we leverage *Gaussian process (GP)* modeling to characterize a statistical model of the environment. The GP models an input vector of spatiotemporal uncertain environment η by $e = f(\eta) + \varepsilon$, where e denotes the observations from the environment, and $\varepsilon \sim \mathcal{N}(0, \sigma^2)$ specifies the Gaussian noise:

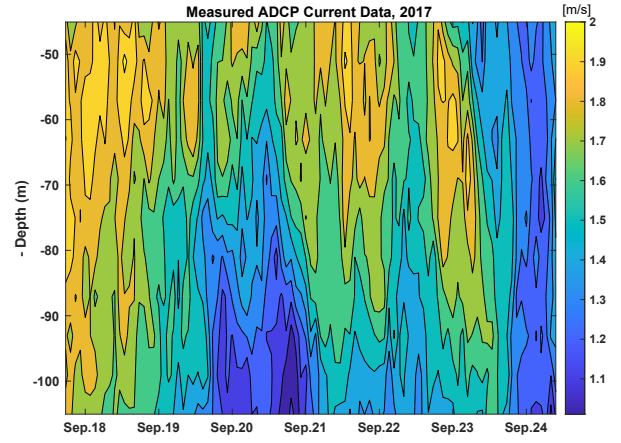$$f(\eta) \sim \mathcal{N}(\mathcal{M}(\eta), \mathcal{K}(\eta, \eta')) \tag{11}$$

Fig. 2: Time histories of the current flow (i.e., the water velocity magnitude) recorded by a 75 kHz ADCP at a latitude of $26.09°N$ and longitude of $-79.80°E$ represented for a sample one-week period.

$$\mathcal{M}(\eta) = \mathbb{E}[f(\eta)] \tag{12}$$

$$\mathcal{K}(\eta, \eta') = \mathbb{E}[(f(\eta) - \mathcal{M}(\eta))(f(\eta') - \mathcal{M}(\eta'))] \tag{13}$$

where $\mathcal{M}(\eta)$ is the mean function, and $\mathcal{K}(\eta, \eta')$ is the covariance function between η and new envionmental data denoted as η'. Note that GP modeling ensures an *accurate* model in the presence of a *sufficiently rich* database [35], denoted by $D = [(\eta, e)|\eta \in H, e \in E]$, from historical ocean current profile observations.

C. Proposed Integrated Path Planning and Tracking Control

To address the coupled MCT path planning/tracking control problem, we develop an *integrated control* framework. The goal of the proposed architecture is to plan a cost criterion-optimal path and to track this optimal path, generally formulated in discrete-time in the following. To ensure an admissible path while avoiding the computational complexity, the path planning is constrained with the major operational constraints of the MCT instead of considering the detailed plant dynamics, which is considered in the path tracking.

Path Planning:

$$J_p(x(.), u(.), y(.)) = \max_{y(.)}[l_f(x(N_p|k), u(N_p|k), y(N_p|k))$$
$$+ \sum_{i=k}^{k+N_p-1} l(x(i|k), u(i|k)), y(i|k)] \tag{14a}$$

$$e(i+1|k) \sim \mathcal{N}(\mathcal{M}(\eta), \mathcal{K}(\eta, \eta')), \ \forall i \in [0:N_p-1] \tag{14b}$$

$$h_{ip}(x(i), u(i), y(i)) \leq 0 \tag{14c}$$

$$h_{ep}(x(i), u(i), y(i)) = 0 \tag{14d}$$

Path Tracking:

$$J_t(x(.), u(.), y(.)) = \min_{u(.)}[|y(N_t|k) - y^*(N_t|k)|^2$$
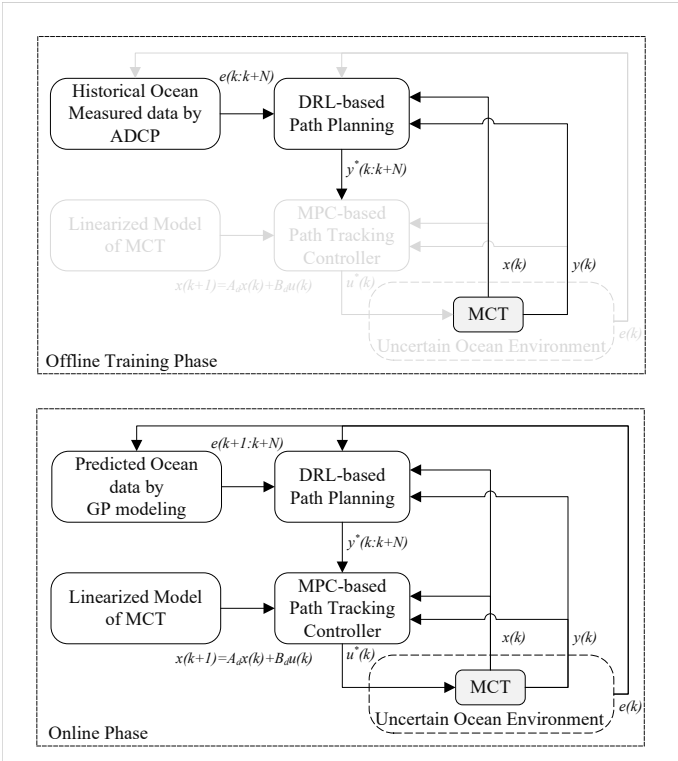$$+ \sum_{i=k}^{k+N_t-1} |y(i|k) - y^*(i|k)|^2] \tag{15a}$$

Fig. 3: Proposed integrated path planning and tracking control framework for marine current turbine.

$$x(i+1|k) = Ax(i|k) + Bu(i|k), \ \forall i \in [0 : N_t - 1] \quad (15b)$$

$$h_{it}(x(i), u(i), y(i)) \leq 0 \quad (15c)$$

$$h_{et}(x(i), u(i), y(i)) = 0 \quad (15d)$$

where $l \in \mathbb{R}$ and $l_f \in \mathbb{R}$ are the cost criterion and the terminal cost, respectively. Constraint (14b) represents the uncertain environment profile, where $e \in \mathbb{R}^o$ is the environment data vector, $h_{i.}$ and $h_{e.}$ are inequality and equality constraint functions, respectively. y and y^* denotes the true waypoint path and optimal waypoint path (reference waypoint path for path tracking) for the MCT, N_p and N_t denote the prediction horizon for path planning and prediction horizon for path tracking, respectively. Constraint (15b) represents the linearized model of the plant. The notation $x(i|k)$ represents the prediction of what the value of x will be at time i given the value of x at time k.

Fig. 3 visualizes the proposed integrated framework, which contains two phases. In the offline phase, the planner is trained due to the historical ocean environment profile to seek the best path. In the online phase, at each iteration, the path planner takes the previous iteration's state, $x(k)$, and environmental data, $e(k)$, to generate an optimal path $y^*(k)$. At each discrete-time instant, k, the path tracking controller follows the commanded optimal path. To be more specific, the path planner solves the optimization problem (14a) to maximize a user-defined cost criterion, where the optimal path satisfies the constraints corresponded to the uncertain ocean model (14b). A learning-based approach will be utilized to find the optimal path through learning from

the historical environment profile (details in Section III-A). The path tracker follows the optimal path assigned by the path planner and computes feasible control inputs, $u(k)$, to minimize the tracking error (15a) by taking into account the linearized model of the plant (15b) subject to the system constraints. An MPC-based approach will be applied to address the optimization problem as a quadratic programming problem using the linearized MCT model (details in Section III-B).

Remark. *The imposition of a learning-based approach, DRL, for the upper-level path planning is significant to address the highly uncertain environment through learning from the historical ocean environment profile. On the other hand, in the presence of the dynamic model of the plant, the selection of an analytical approach, MPC, for the lower-level path tracking ensures to successfully follow the assigned optimal path robustly, without violating the system's limitations. The result of this framework is an integrated control approach that satisfies path planning in an uncertain environment and path tracking using a detailed plant model.*

III. DETAILED PATH PLANNING AND PATH TRACKING

A. Path Planning through Reinforcement Learning

The objective of path planning is to optimize a chosen cost criterion, considering the plant in a dynamic and highly uncertain environment. Specifically, for the MCT, the major objective is to maximize the harvested power from the plant. To solve the real-time path planning, denoted as an iterative decision making (control) problem in a stochastic environment, a DRL approach is considered as a powerful tool to find the optimal path through learning from the historical data, which shows acceptable robustness to the environmental uncertainties in similar applications [36]. A brief review of the RL preliminaries is presented below.

Reinforcement Learning (RL): In RL, a Markov decision process (MDP), as a canonical formalism for the stochastic control problem, is modeled by a tuple $\mathcal{MDP} := < \mathcal{S}, \mathcal{A}, \mathcal{P}, R, \gamma >$, where $\mathcal{S}$ are the state and action spaces, respectively; $\mathcal{P}$ is the probability of state transition, with $\mathcal{P}(s'|s,a)$ denoting the probability of transiting to the state $s' \in \mathcal{S}$ given the current state $s \in \mathcal{S}$ after taking the action $a \in \mathcal{A}$, $R : \mathcal{S} \times \mathcal{A} \to \mathbb{R}$ is the reward function.

A policy $\pi : \mathcal{S} \to \mathcal{A}$ determines the action to take at state s. To estimate how good a particular policy will be with a given state, a value function $V_\pi(s)$ is used to measure the accumulated reward, which is defined by:

$$V_\pi(s) = \sum_{a \in \mathcal{A}} \pi(a|s) \sum_{s' \in \mathcal{S}} \mathcal{P}(s'|s,a)(r_k + \gamma V_\pi(s')) \quad (16)$$

where $\pi(a|s)$ denotes a policy in the RL to take an action $a \in \mathcal{A}$ at state $s \in \mathcal{S}$ and $r_k = R(s_k, a_k)$. Similarly, a Q-value function of a policy π is defined as follows while $\gamma \in [0,1)$ is the discount factor:

$$Q_\pi(s,a) = r_k + \gamma \sum_{s' \in \mathcal{S}} \mathcal{P}(s'|s,a)V_\pi(s') \quad (17)$$

The optimal Q-value function Q^* is defined as $Q^*(s,a) = \max_{\pi} Q_\pi(s,a)$. Given Q^*, the optimal policy π^* is then calculated by $\pi^* = \underset{\pi}{argmax}\, Q_\pi(s,a)$.

Q-learning approach, as one of a popular RL algorithms, is chosen to find $Q^*(s,a)$; however, a standard tabular Q-learning lacks to scale to problems with high-dimensional state and action spaces. To tame this issue, deep Q-learning is introduced to approximate the Q-value function for all possible actions with a deep neural network. To find the optimal Q-value Q^*, we use the DRL approach (specifically deep Q-network (DQN)). The DQN includes two deep neural networks named the *Q-network* $Q(.;\theta)$ and *target network* $Q(.;\theta^-)$, where θ and θ^- denote the parameters (weights) of these two networks. The main goal of the DQN is to optimize $Q(s,a)$ estimated by $Q(.;\theta)$ and its distance with the temporal difference target t_j estimated by the $\widehat{Q}(.;\theta^-)$, expressed as a gradient descent:

$$L(\theta) = [t_j - Q(s_j,a_j;\theta)]^2 \tag{18}$$

$$t_j := r_j + \gamma \max_{a'} Q(s_{j+1},a';\theta^-)) \tag{19}$$

To select an action, we follow an adaptive $\varepsilon-$greedy policy [37] as a method to choose random actions with uniform distribution from possible actions, namely:

$$a_k = \begin{cases} \underset{a_k}{argmax}\, Q(s_k,a_k), & 1-\varepsilon \\ \text{random } a \in A, & \varepsilon \end{cases} \tag{20}$$

$$\varepsilon = \varepsilon_{min} + (\varepsilon_{max} - \varepsilon_{min})e^{-d \times n_e} \tag{21}$$

where d is the decay factor and n_e is the episode number.

DRL-based Path Planning: For the MCT application, at each iteration, the DRL-based approach decides on whether to keep the MCT at its current position (i.e., water depth) or to change the position that has a large flow speed to maximize the harvested power in the planning prediction horizon N_p (see (14a)). The state $\mathscr{S}$, action $\mathscr{A}$, and reward $\mathscr{R}$ notations should be defined according to the MCT-specific application. The MCT can observe its current states, $x(k)$, and the ocean current velocity $e(k) \triangleq v_c(k)$; hence, the state of this system at time k is $s(k) = [x(k), v_c(k)]$. To define the action space, we consider the possible action that MCT can take to maximize the generated power, where the position change constitutes the action of MCT at k (i.e., $a(k) \triangleq y(k)$).

The major concern for the path planning problem for the MCT system arises from its complex power equation and how to define the DRL's reward in relation to that power expression. Since the ultimate objective of path planning is to maximize the generated power, a positive reward, α, is gained if the position change fulfills the power increase, and a zero reward is given to the wrong position change:

$$R(s_k,a_k) = \begin{cases} \alpha, & P(z(k+1)) > P(z(k)) \\ 0, & \text{else} \end{cases} \tag{22}$$

The total harvested power $P(z(k))$ from the MCT comprises of three terms of generated power P^G, consumed

power to hold depth P^{HD}, and consumed power to change depth P^{CD}, namely:

$$P(z(k)) = P^G - P^{HD} - P^{CD} \tag{23}$$

$$P^G = \frac{1}{2}\rho A c_p v_c^3(z(k)) \tag{24}$$

$$P^{HD} = \begin{cases} 0, & v_c(z(k+1)) - v_c(z(k)) < 0 \\ \frac{\zeta_1}{T_p}[v_c(z(k+1)) - v_c(z(k))], & v_c(z(k+1)) - v_c(z(k)) > 0 \end{cases} \tag{25}$$

$$P^{CD} = \begin{cases} 0, & z(k+1) - z(k) > 0 \\ \frac{\zeta_2}{T_p}[z(k+1) - z(k)], & z(k+1) - z(k) < 0 \end{cases} \tag{26}$$

where ρ denotes the water density, A is the swept area by the MCT rotor, v_c denotes the ocean velocity (denoted as e in our framework presented in Fig. 3), c_p is the average power coefficient, ζ_1 and ζ_2 denote the coefficients for the power equations, and T_p denotes the path planning sampling time.

According to the power equation (23), the optimal depth (denoted the optimal power) is a compromise between the maximum water current and both the power consumption to hold depth and power consumption to change the depth. On the other hand, note that the magnitude of a non-zero reward α significantly impacts the training procedure. We formulate the path planning problem to increase the total reward gained by the position change, namely:

$$\pi^* = \underset{\pi}{argmax}\,[r_k + \gamma \sum_{s' \in \mathscr{S}} \mathscr{P}(s'|s,a)V_\pi(s')] \tag{27}$$

where the MCT learns the optimal policy π^* using DQN approach to find the optimal waypoint path by solving (27).

To do the path planning, we should first construct the DQN through offline training using the historical ocean environmental data, and the constructed DQN is then used to find the optimal path in an online phase. Algorithm 1 details the offline training process.

B. Path Tracking through Model Predictive Control

The path tracking will minimize the distance between the plant's current path and the optimal path commanded by the path planner. MPC, as a powerful model-based algorithm, is chosen to address path tracking due to the complex dynamics consisting of multiple states and the availability of a detailed MCT model.

Model Predictive Control (MPC): To design an MPC for the DLTI system defined in (10), the following cost function must be minimized:

$$\begin{aligned} \tilde{V} \equiv \frac{1}{2} \Bigg(&\sum_{i=k+1}^{k+N_t-1} \Big((y(i|k) - y^*(i))^T Q_{\text{part}}(y(i|k) - y^*(i)) \Big) \\ &+ \sum_{i=k}^{k+N_c-1} \Big((u(i|k) - u(i-1|k))^T R_{\Delta\text{part}}(u(i|k) - u(i-1|k)) \\ &+ (u(i|k) - d(i))^T R_{\text{part}}(u(i|k) - d(i)) \Big) \\ &+ (y(k+N_t|k) - y^*(k+N_t))^T Q_{\text{term}}(y(k+N_t|k) - y^*(k+N_t)) \Bigg) \end{aligned} \tag{28a}$$

$$u^{\min}(i) \leq u(i) \leq u^{\max}(i) \tag{28b}$$

$$\Delta u^{\min}(i) \leq \Delta u(i) \leq \Delta u^{\max}(i) \tag{28c}$$

$$y^{\min}(i) \leq y(i) \leq y^{\max}(i) \tag{28d}$$

Algorithm 1 Deep reinforcement learning for path planning
1: Initialize recorded training data sample of an uncertain environment;
2: Initialize replay memory M and mini-batch size;
3: Initialize action-value function Q and target action-value function $\widehat{Q}$ with random weights θ and $\theta^- = \theta$;
4: **for** episode $i = 1\,to\,N_{\text{episode}}$ **do**
5:    Initialize initial state;
6:    **for** time step $k = 1\,to\,N_p$ **do**
7:       Select action a_k using ε-greedy policy (20);
8:       Take action a_k and observe r_k and s_{k+1};
9:       Store transition (s_k, a_k, r_k, s_{k+1}) in M;
10:      Sample random mini-batch of transitions (s_j, a_j, r_j, s_{j+1}) from M;
11:      Set $t_j = r_j + \gamma \max_{a'} Q(s_{j+1}, a'; \theta^-))$;
12:      Perform a gradient descent step on $L(\theta) = [t_j - Q(s_j, a_j; \theta)]^2$ with respect to θ;
13:      Update $\widehat{Q}$ every b steps and set $\theta^- = \theta$;
14:    **end for**
15: **end for**
16: Output offline trained optimal deep Q-network Q^*;

Algorithm 2 Model predictive control for path tracking
1: Discretize the continuous linear model based on Shannon's Theorem;
2: Initialize the linear model inside the MPC solver;
3: Initialize the constraints for the actuators and the system states;
4: Initialize N_t, N_c, Q_{part}, Q_{term}, $R_{\Delta\text{part}}$ and R_{part};
5: Initialize the initial state and control inputs;
6: **for** $i = 1\,to\,N_{\text{operation}}$ **do**
7:    Initialize the reference trajectory based on the optimal waypoint path from the DRL;
8:    **for** $j = 1\,to\,N_{\text{operation-int}}$ **do**
9:      Apply the interior-point method, select u^* for N_c over N_t;
10:     Solve the cost function $\tilde{V}$ in (28a) and obtain optimal $[u^*(i|N_c), ..., u^*(i|i+N_c-1)]$;
11:     Apply $u^*(i|N_c)$ to the DLTI plant and obtain the next system state $x(i+1)$;
12:    **end for**
13:    Update the position error and send a feedback to DRL;
14: **end for**

where N_t and N_c are the prediction horizon and the control horizon, respectively; $d(i)$ is a vector of the arbitrary user-selected target controls at time step i, Q_{part} and Q_{term} are the weight factors for the states and the final state point, respectively. Finally, $R_{\Delta\text{part}}$ and R_{part} are the weight factors for the control input changes and the desired control inputs, respectively. Constraints (28b)-(28c) impose that control input and control input change should stay within a predefined bound. Constraint (28d) captures the need to bound the system output. Here, $(.)^{\min}(i)$ and $(.)^{\max}(i)$ are the minimum and maximum values of $(.)(i)$.

After the algebraic manipulations as given in [7], the cost function is converted to the quadratic form, which is then minimized using the interior-point method, subject to the constraints given in (28b)-(28d).

MPC-based Path Tracking: In our problem, we will leverage the path tracking algorithm to obligate the plant to follow the optimal path with a minimum error subject to system constraints. Considering the MCT with a high number of coupled states and the requirement of handling many constraints, MPC provides a controller that meets the desired objectives in a guaranteed manner. Also, the user-defined weights bring flexibility about choosing the more important states in the plant to be followed while giving less priority to the other states that are relatively less crucial for the desired performance. MPC is designed to minimize the cost function as given in (28a), in which the weights are defined to prioritize the states to follow the trajectory with the minimum error. The reference path is defined in the vector of target outputs, $y^*(i)$, and at each sampling time, the updated reference is received from the DRL-based path planner. Algorithm 2 describes the MPC-based path tracking. The constraints of the MPC solver are defined based on the

MCT design and specifications. Control inputs saturations and the slew rates constraints are defined as:

$$[|B_f|, |B_a|, |\tau_{em}|] \leq [\%50, \%50, 0.2 \times \tau_{em}^{eq}\ \text{Nm}] \tag{29}$$

$$[|\dot{B}_{aft}|, |\dot{B}_{front}|] \leq [7.45 \times 10^{-4}, 7.45 \times 10^{-4}]\%/s \tag{30}$$

Since the desired performance is to track the depth reference, the related state, z, has the only weight defined in Q_{part} and the rest is set to be zero as $Q_{\text{part}} = \text{diag}(0,0,0,0,0,0,0,0,1,0,0,0)$. The other rates are defined as $Q_{\text{term}} = I_{13}, R_{\text{part}} = 0$, and $R_{\Delta\text{part}} = I_3$, where I_n is the $n \times n$ identity matrix.

IV. NUMERICAL RESULTS

We evaluate the proposed path planning and path tracking for the MCT with the total mass of 4.98×10^5 kg, and mass moments of $I_{xx} = 1.35 \times 10^7$ kgm^3, $I_{yy} = 4.74 \times 10^7$ kgm^3, and $I_{zz} = 3.45 \times 10^7$ kgm^3 [24]. The other parameters for the power equations include $\rho = 1030$ kg/m^3, $A = 100\pi$, $c_p = 41.5\%$, $\zeta_1 = 9.113$, and $\zeta_2 = -0.0365$. For the path planner, we select a DQN with two hidden layers with a buffer size of $5e5$, a batch size of 64, $\gamma = 0.5$, $\varepsilon_{\min} = 0.01$, and $\varepsilon_{\max} = 1$. The experimental tests show that setting the reward value to $\alpha = +1$ justifies a successful training in our problem. For the tracking controller, the horizons are selected as $N_t = 40$, and $N_c = 20$ considering the system's performance with highly-limited actuator slew rates. For the path planning and DQN training, the experimental tests are implemented in Python 3.7 and Tensorflow 2.4, where each training step takes 58 ms within 3000 steps and 100 episodes. To find the sampling time, the eigenvector analysis is carried
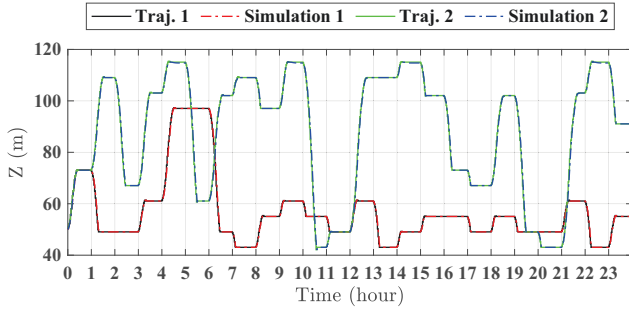
Fig. 4: Trajectory followed by the MCT along with the optimal reference trajectory for two ocean shear profiles.

out and the eigenvalues of the linear MCT model are:

$$\begin{bmatrix} -0.2731 \pm 1.2585i & -0.1121 \pm 0.1549i \\ -0.2588 \pm 0.9618i & -0.0033 \pm 0.0021i \\ -0.2647 \pm 0.3564i & -0.0483 \\ -0.1754 \pm 0.3793i & \end{bmatrix} \quad (31)$$

According to this analysis, the maximum frequency of the MCT system is calculated $\omega_{n,max} = 1.29$ rad/s so that we opt for a sampling time of $T_s = 2$ s satisfying the minimum requirement based on the Shannon sampling theorem. The path tracking is implemented in MATLAB, and the simulation time for solving (28a) per iteration is 140 ms.

The goal is to maximize the net power of the MCT through modifying the operating water depth z (i.e., vertical path planning), where the tracking controller tracks the optimal depth with the minimized error. After the path planning, the z trajectory is sent to the tracking controller, where the error is minimized through (28a). Fig. 4 shows the trajectory followed by the MCT along with the optimal reference trajectory for two shear profiles, verifying that the proposed path tracking controller successfully follows the assigned optimal path by the path planner for the entire trajectory. It is shown that the path planner is able to find the optimal trajectory over low and high shear profiles, where the MCT also navigates the commanded reference paths. Results show the efficiency of the path planner in finding the optimal path to maximize the net power, where the total harvested power after 24 hours shows an over 9 % increase compared to the baseline case (the MCT maintains at $z_{eq} = 50\,m$ without path planning). Note that the baseline case is selected based on the equilibrium depth of the MCT, which can produce the highest power while keeping a constant depth. Fig. 5 demonstrates that the Euler angles of the MCT remain within the acceptable bounds ($\pm 6°$), and Fig. 6 illustrates the control inputs applied to the MCT to track the reference trajectory. We can see that the control inputs do not violate the constraints presented in (29).

V. CONCLUSIONS

In this paper, we presented an integrated path planning and tracking control framework for the MCT wherein the planner was designed using a DRL approach to develop a strategy in order to optimize a cost criterion (maximizing the harvested
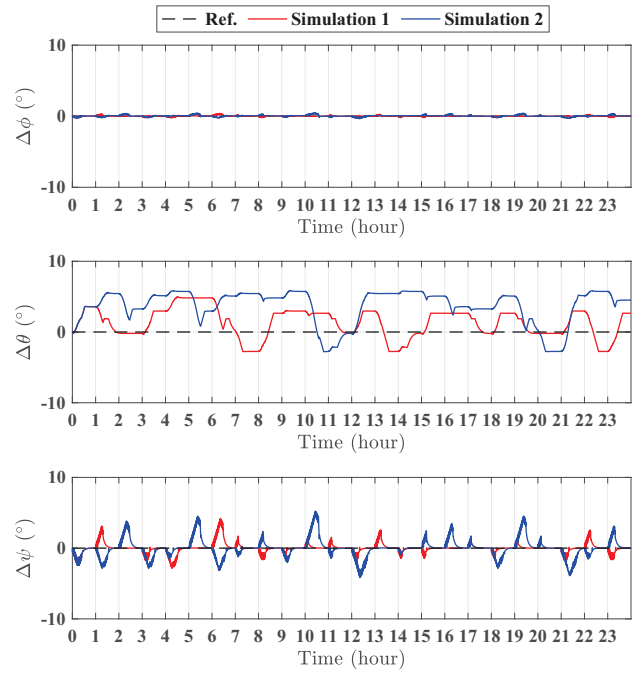

Fig. 5: Euler angles of the MCT system for two ocean shear profiles.

power from the MCT). The tracking control was addressed through an MPC approach considering the linear model of the system obtained from the nonlinear dynamic model. The proposed approach was applied to a buoyancy controlled MCT, and the numerical results verified that our proposed approach was able to maximize the net power through path planning compared with the baseline. Results also demonstrated that the tracking controller could effectively track the optimal path by considering the sluggish dynamics while avoiding any aggressive motion and instability.

REFERENCES

[1] Z. Zhou, M. Benbouzid, J.-F. Charpentier, F. Scuiller, and T. Tang, "Developments in large marine current turbine technologies–a review," *Renewable and Sustainable Energy Reviews*, vol. 71, pp. 852–858, 2017.

[2] P. F. Lermusiaux, C.-S. Chiu, G. G. Gawarkiewicz, P. Abbot, A. R. Robinson, R. N. Miller, P. J. Haley, W. G. Leslie, S. J. Majumdar, A. Pang *et al.*, "Quantifying uncertainties in ocean predictions," HARVARD UNIV CAMBRIDGE MA, Tech. Rep., 2006.

[3] Y. Kanayama, Y. Kimura, F. Miyazaki, and T. Noguchi, "A stable tracking control method for an autonomous mobile robot," in *Proceedings., IEEE International Conference on Robotics and Automation*. IEEE, 1990, pp. 384–389.

[4] D. Verscheure, B. Demeulenaere, J. Swevers, J. De Schutter, and M. Diehl, "Time-optimal path tracking for robots: A convex optimization approach," *IEEE Transactions on Automatic Control*, vol. 54, no. 10, pp. 2318–2327, 2009.

[5] Y. Wang and D. Wang, "Non-singular moving path following control for an unmanned aerial vehicle under wind disturbances," in *2017 IEEE 56th Annual Conference on Decision and Control (CDC)*. IEEE, 2017, pp. 6442–6447.

[6] T. Oliveira, A. P. Aguiar, and P. Encarnacao, "Moving path following for unmanned aerial vehicles with applications to single and multiple target tracking problems," *IEEE Transactions on Robotics*, vol. 32, no. 5, pp. 1062–1078, 2016.
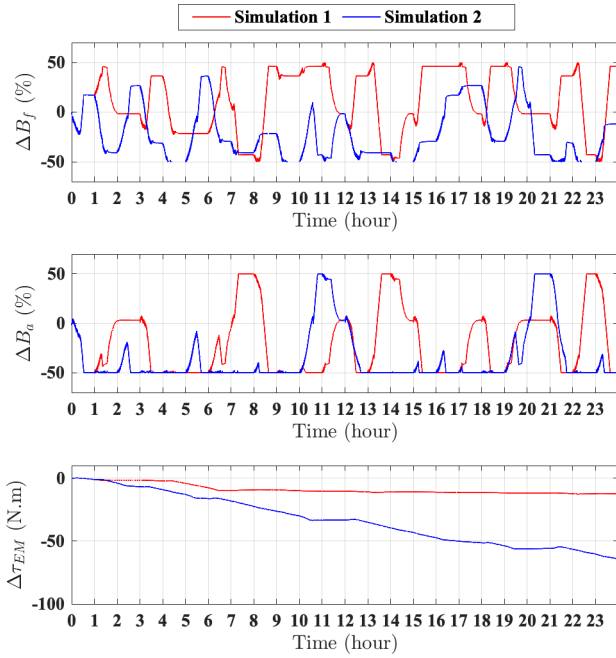
Fig. 6: Control inputs applied to the MCT system to track the reference trajectory for two ocean shear profiles.

[7] W. B. Greer and C. Sultan, "Shrinking horizon model predictive control method for helicopter–ship touchdown," *Journal of Guidance, Control, and Dynamics*, vol. 43, no. 5, pp. 884–900, 2020.

[8] P. Encarnacao and A. Pascoal, "3d path following for autonomous underwater vehicle," in *Proceedings of the 39th IEEE conference on decision and control (Cat. No. 00CH37187)*, vol. 3. IEEE, 2000, pp. 2977–2982.

[9] T. D. Ngo, C. Sultan, J. H. VanZwieten, and N. I. Xiros, "Model predictive control for moored ocean current turbines," in *2017 American Control Conference (ACC)*. IEEE, 2017, pp. 875–880.

[10] M. Cobb, M. Wu, K. Barton, and C. Vermillion, "Flexible-time economic iterative learning control: A case study in airborne wind energy," in *2019 IEEE 58th Conference on Decision and Control (CDC)*. IEEE, 2019, pp. 5580–5586.

[11] D. Todeschini, L. Fagiano, C. Micheli, and A. Cattano, "Control of vertical take off, dynamic flight and landing of hybrid drones for airborne wind energy systems," in *2019 American control conference (ACC)*. IEEE, 2019, pp. 2177–2182.

[12] E. Schmidt, M. D. L. C. de Oliveira, R. S. da Silva, L. Fagiano, and A. T. Neto, "In-flight estimation of the aerodynamics of tethered wings for airborne wind energy," *IEEE Transactions on Control Systems Technology*, vol. 28, no. 4, pp. 1309–1322, 2019.

[13] A. U. Zgraggen, L. Fagiano, and M. Morari, "Real-time optimization and adaptation of the crosswind flight of tethered wings for airborne wind energy," *IEEE Transactions on Control Systems Technology*, vol. 23, no. 2, pp. 434–448, 2014.

[14] J. Koenemann, P. Williams, S. Sieberling, and M. Diehl, "Modeling of an airborne wind energy system with a flexible tether model for the optimization of landing trajectories," *IFAC-PapersOnLine*, vol. 50, no. 1, pp. 11 944–11 950, 2017.

[15] C. Vermillion, M. Cobb, L. Fagiano, R. Leuthold, M. Diehl, R. S. Smith, T. A. Wood, S. Rapp, R. Schmehl, D. Olinger *et al.*, "Electricity in the air: Insights from two decades of advanced control research and experimental flight testing of airborne wind energy systems," *Annual Reviews in Control*, 2021.

[16] L. Fagiano, M. Tanaskovic, L. C. Mallitasig, A. Cataldo, and R. Scattolini, "Hierarchical routing control in discrete manufacturing plants via model predictive path allocation and greedy path following," in *2020 59th IEEE Conference on Decision and Control (CDC)*. IEEE, 2020, pp. 5546–5551.

[17] J. Ji, A. Khajepour, W. W. Melek, and Y. Huang, "Path planning and tracking for vehicle collision avoidance based on model predictive control with multiconstraints," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 2, pp. 952–964, 2016.

[18] A. Maitland and J. McPhee, "Towards integrated planning and control of autonomous vehicles using nested mpcs," in *ASME 2018 Dynamic Systems and Control Conference*. American Society of Mechanical Engineers Digital Collection, 2018.

[19] Y. V. Pant, H. Yin, M. Arcak, and S. A. Seshia, "Co-design of control and planning for multi-rotor uavs with signal temporal logic specifications," *arXiv preprint arXiv:2009.14363*, 2020.

[20] M. Jafari, H. Xu, and L. R. G. Carrillo, "Brain emotional learning-based path planning and intelligent control co-design for unmanned aerial vehicle in presence of system uncertainties and dynamic environment," in *2018 IEEE Symposium Series on Computational Intelligence (SSCI)*. IEEE, 2018, pp. 1435–1440.

[21] X. Wang, X. Yao, and L. Zhang, "Path planning under constraints and path following control of autonomous underwater vehicle with dynamical uncertainties and wave disturbances," *Journal of Intelligent & Robotic Systems*, pp. 1–18, 2020.

[22] F. Repoulias and E. Papadopoulos, "Planar trajectory planning and tracking control design for underactuated auvs," *Ocean engineering*, vol. 34, no. 11-12, pp. 1650–1667, 2007.

[23] W. Zuo, X. Yi, F. H. Ghorbel, and Z. Chen, "Optimal trajectory planning and control of buoyancy control device enabled by water electrolyzer," in *2019 IEEE 58th Conference on Decision and Control (CDC)*. IEEE, 2019, pp. 2120–2125.

[24] A. Hasankhani, J. VanZwieten, Y. Tang, B. Dunlap, A. De Luera, C. Sultan, and N. Xiros, "Modeling and numerical simulation of a buoyancy controlled ocean current turbine," *International Marine Energy Journal*, vol. 4, no. 2, pp. 47–58, 2021.

[25] Y. Tang, Y. Zhang, A. Hasankhani, and J. VanZwieten, "Adaptive super-twisting sliding mode control for ocean current turbine-driven permanent magnet synchronous generator," in *2020 American Control Conference (ACC)*. IEEE, 2020, pp. 211–217.

[26] Y. Tang, J. VanZwieten, B. Dunlap, D. Wilson, C. Sultan, and N. Xiros, "In-stream hydrokinetic turbine fault detection and fault tolerant control-a benchmark model," in *2019 American Control Conference (ACC)*. IEEE, 2019, pp. 4442–4447.

[27] A. Hasankhani, Y. Tang, A. Snyder, J. VanZwieten, and W. Qiao, "Control co-design for buoyancy-controlled mhk turbine: A nested optimization of geometry and spatial-temporal path planning," *arXiv preprint arXiv:2110.06470*, 2021.

[28] T. Ueno, S. Nagaya, M. Shimizu, H. Saito, S. Murata, and N. Handa, "Development and demonstration test for floating type ocean current turbine system conducted in kuroshio current," in *2018 OCEANS-MTS/IEEE Kobe Techno-Oceans (OTO)*. IEEE, 2018, pp. 1–6.

[29] D. Coiro, G. Troise, F. Scherillo, A. De Marco, G. Calise, and N. Bizzarrini, "Development, deployment and experimental test on the novel tethered system gem for tidal current energy exploitation," *Renewable Energy*, vol. 114, pp. 323–336, 2017.

[30] J. H. VanZwieten, N. Vanrietvelde, and B. L. Hacker, "Numerical simulation of an experimental ocean current turbine," *IEEE Journal of Oceanic Engineering*, vol. 38, no. 1, pp. 131–143, 2012.

[31] T. D. Ngo, C. Sultan, J. H. VanZwieten, and N. I. Xiros, "Constrained control of moored ocean current turbines with cyclic blade pitch variations," *IEEE Journal of Oceanic Engineering*, 2020.

[32] A. Hasankhani, , Y. Tang, J. VanZwieten, and C. Sultan, "Spatiotemporal optimization for path planning, with application to an ocean current turbine," *IEEE Transactions on Control Systems Technology*, Under revision, 2nd round.

[33] A. J. Jerri, "The shannon sampling theorem—its various extensions and applications: A tutorial review," *Proceedings of the IEEE*, vol. 65, no. 11, pp. 1565–1596, 1977.

[34] M. C. P. M. Machado, J. H. VanZwieten, and I. Pinos., "A measurement based analyses of the hydrokinetic energy in the gulf stream," *Journal of Ocean and Wind Energy*, vol. 3, no. 1, pp. 25–30, 2016.

[35] K.-C. Ma, L. Liu, and G. S. Sukhatme, "Informative planning and online learning with sparse gaussian processes," in *2017 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2017, pp. 4292–4298.

[36] Y. Lin, J. McPhee, and N. L. Azad, "Comparison of deep reinforcement learning and model predictive control for adaptive cruise control," *IEEE Transactions on Intelligent Vehicles*, 2020.

[37] X. Liang, X. Du, G. Wang, and Z. Han, "A deep reinforcement learning network for traffic light cycle control," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 2, pp. 1243–1253, 2019.