An Online Actor-Critic Identifier with Sampled Fatigue Measurements for Optimal Adaptive Control of FES and an Electric motor

Ashwin Iyer, Mayank Singh, Qiang Zhang, Ziyue Sun, Nitin Sharma*

***Abstract*— Cooperative control of functional electrical stimulation (FES) and electric motors in a hybrid exoskeleton may benefit from fatigue measurements and online model learning. Recent model-based cooperative control approaches rely on time-consuming offline system identification of a complex musculoskeletal system. Further, they may lack the ability to include measurements from muscle sensors that monitor the FES-induced muscle fatigue, which may hinder maintaining desired muscle fatigue levels. This paper develops an online adaptive reinforcement learning approach to control knee extension via an electric motor and FES. An optimal tracking control problem that uses an actor-critic identifier structure is formulated to approximate an optimal solution to the Hamiltonian-Jacobi-Bellman equation. The continuous controller provides asymmetrically saturated optimal control inputs of FES and the electric motor. Critic and identifier neural networks are designed to simultaneously estimate the reward function and the system dynamics based on sampled fatigue measurements and compute control actions. Importantly, simulation results show that a satisfactory joint angle tracking and actuator allocation can be obtained at multiple on-demand desired muscle fatigue levels and prolong FES utilization.**

I. Introduction

Spinal cord injuries (SCI) are debilitating and frequently result in complete or partial paraplegia that causes loss of essential lower limb functions such as walking, running, sitting-to-standing, etc. Several studies indicate that functional electrical stimulation (FES) can help restore lower limb function in persons with SCI. FES, which works by applying external electrical currents, artificially activates motor neurons that induce desired muscle contractions. However, the artificial nature of FES leads to unnatural motor neuron activation patterns, which causes a rapid onset of fatigue in the stimulated muscle. The onset of FES-induced muscle fatigue quickly deteriorates the desired limb movement, limiting its effectiveness. One of the proposed solutions to reduce the FES-induced muscle fatigue's effect is to utilize a hybrid approach that combines both electric motors and FES to facilitate walking or sitting to standing [1].

Ashwin Iyer, Qiang Zhang, Ziyue Sun and Nitin Sharma are with the Joint Department of Biomedical Engineering, North Carolina State University, Raleigh, NC, USA and University of North Carolina, Chapel Hill, NC, USA. Mayank Singh is with the Electrical and Computer Engineering Department at North Carolina State University, Raleigh, NC, USA. This work was funded by NSF Award # 1646009 and NSF CAREER Award # 2000261

*Corresponding author: Nitin Sharma (e-mail: nsharm23@ncsu.edu).

Model predictive control (MPC) methods were recently designed for cooperative and optimal allocation of desired torque between the motors and FES for sit-to-stand and leg extension tasks [2]. However, the MPC methods rely on exact model knowledge for optimal performance. Robust MPC methods have been designed to account for uncertainties in parameter estimation and disturbances by adding a feedback controller. Nonetheless, MPC relies on an offline model identification procedure to provide a nominal system model, which is likely a bit time-consuming for clinical implementation.

Reinforcement learning (RL) is a class of machine learning methods that modify an agent's actions through a reward function representing the system interactions with its environment. Recently, RL has excited control researchers to develop optimal control laws for dynamical systems described by ordinary differential equations. These continuous RL formulations lead to the Hamilton-Jacobi-Bellman (HJB) equation [3]. However, since the nonlinear HJB equation is harder to solve, policy-iteration algorithms have been developed that use neural networks trained to approximate a solution to the HJB equation. Specifically, a policy-iteration algorithm creates an actor-critic framework that employs neural networks to approximate value function and optimal control actions from system state measurements. These methods have been shown to regulate continuous time linear [4] and nonlinear systems [5]. Even an online identifier can be added to the framework to form an actor critic identifier (ACI) structure [6] to determine optimal adaptive control laws that can be potentially applied without a prior information of the system dynamics. Thus, we are motivated to explore ACI as a potential RL approach to determine optimal motor and FES control inputs for a hybrid exoskeleton while obviating the need to build a prior knowledge of musculoskeletal and exoskeleton dynamics.

Another desire is to include fatigue measurements in the control design rather than relying on a person-specific fatigue model to maintain desired muscle state or its fatigue level. Current fatigue measurement methods include include surface electromyography (sEMG) and force measurements. However, incorporating force measurements into an exoskeleton design and measuring joint torques are difficult. In addition, with force measurements it is difficult to isolate the fatigue of a specific muscle group within a larger set of muscles. On the other hand, sEMG based fatigue measurement is difficult due to FES-related stimulation artifacts and its sensitivity to cross-talk between muscle groups. Our previous work has shown that ultrasound (US) imaging-derived signals such as a cross-correlation-based strain measurement can estimate fatigue [7]. A drawback of US imaging-derived measures is that they are obtained offline due to a high computation cost of the image analysis. The US-derived measures may be fast processed using graphical processing units but still would be limited to low sampling rates. To incorporate the low-sampled measurement in FES based control design, in [8] we used a model-based sampled data observer to estimate FES-evoked muscle activation. However the observer as well as the control still relies on person specific neuromuscular parameters that are difficult to estimate.

In this paper we derive an online actor-critic approach to simultaneously compute control actions for motor and FES, estimate a value function, and identify unknown system dynamics during a knee extension task. A continuous optimal tracking control problem is formulated with a reward function that minimizes the effect of FES-induced muscle fatigue on the tracking while simultaneously satisfying asymmetric constraints on the FES input. We address the challenge of integrating sampled fatigue measurements with a continuous RL structure by designing a neural network (NN)-based identifier that accommodates intermittent measurements. Using a Lyapunov stability analysis and hybrid system approach, the identifier convergence error is shown to be ultimately bounded despite sampled fatigue measurements. Simulations show tracking performance and input allocation between motors and FES during the knee extension task under different manually selected desired muscle fatigue levels.

II. Dynamic Model of Knee Extension

The dynamics for a single degree of freedom musculoskeletal model are given as

$$J\ddot{q} + \tau_p + G(q) = \tau, \tag{1}$$

where $q, \dot{q}, \ddot{q} \in \mathbb{R}$ represent angular position, velocity, and acceleration of a limb joint respectively, $\tau \in \mathbb{R}$ is the total torque input required to move the limb joint, $J \in \mathbb{R}^+$ is the moment of inertia of the leg, $G(q) = mglsin(q + q_{eq})$ is a term that represents the gravitational torque, where $m \in \mathbb{R}^+$ is the mass of leg, $g \in \mathbb{R}^+$ is the gravitational acceleration constant, $l \in \mathbb{R}^+$ is the distance from the knee joint to the center of mass, and $q_{eq} \in \mathbb{R}^+$ is the equilibrium position of the lower leg with respect to the vertical, and τ_p is a passive torque based on person specific parameters defined in [9]. The total combined torque is composed as $\tau = \tau_m + \tau_f$, where $\tau_m, \tau_f \in \mathbb{R}$ represent the torque produced by the electric motor and and FES respectively and τ_f is modeled as

$$\tau_f = \rho(q, \dot{q})\varphi u_f, \tag{2}$$

where $\rho(q, \dot{q})$ contains force-length and force-velocity relations based on person specific parameters [9], $\varphi \in \mathbb{R}$ is the FES-induced muscle fatigue, and $u_f \in \mathbb{R}$ is the normalized FES current or pulse-width input. The FES-induced fatigue is modeled as

$$\dot{\varphi} = w_f(\varphi_{min} - \varphi)u_f + w_r(1 - \varphi)(1 - u_f), \tag{3}$$

where $w_f,\, w_r \in \mathbb{R}^+$ are time constants for fatigue and recovery of the muscle and φ_{min} is the minimum fatigue value for each person. Because u_f is the normalized FES input, the solution for (3) at the bounds $u_f = 0$ and $u_f = 1$ result in φ being bounded as $\varphi \in [\varphi_{\min}, 1]$ where a fatigue value of 1 means the muscle is fully rested and a fatigue value of φ_{min} means the muscle is completely fatigued. The fatigue dynamics in (3) can be rearranged as $\dot{\varphi} = g_\varphi(\varphi)u_f + f_\varphi(\varphi)$, where $g_\varphi(\varphi) = (w_f\varphi_{min} - w_f\varphi + w_r + w_r\varphi)$ and $f_\varphi(\varphi) = (w_r - w_r\varphi)$. By defining states $x_1 = q$, $x_2 = \dot{q}$, $x_3 = \varphi$ the state space representation of the dynamics in (1) can be written as

$$\dot{x} = f(x) + g_1(x)\tau_m + g_2(x)u_f, \tag{4}$$

where $\dot{x} = \begin{bmatrix} \dot{x}_1 & \dot{x}_2 & \dot{x}_3 \end{bmatrix}^T$ and $f(x) \in \mathbb{R}^{3\times1}, g_1(x) \in \mathbb{R}^{3\times1}$, and $g_2(x) \in \mathbb{R}^{3\times1}$ are defined as

$$f(x) = \begin{bmatrix} x_2 \\ -\frac{1}{J}G(x_2) - \frac{\tau_p}{J} \\ f_\varphi(x_3) + g_\varphi(x_3)u_f. \end{bmatrix}, g_1(x) = \begin{bmatrix} 0 \\ \frac{1}{J} \\ 0 \end{bmatrix},$$

$$g_2(x) = \begin{bmatrix} 0 \\ \frac{1}{J}\rho\varphi \\ g_\varphi(x_3)u_f \end{bmatrix}.$$

III. Determining Optimal Control Inputs

The control objective is to determine optimal motor and FES inputs that generate the required torque to track a prescribed limb angle trajectory by utilizing an ACI approach to solve the optimal tracking control problem. A tracking error $e \in \mathbb{R}^{3\times1}$ is defined as

$$e = x(t) - x_d(t), \tag{5}$$

where $x_d(t) \in \mathbb{R}^{3\times1}$ is a bounded desired trajectory for the position, velocity, and fatigue. It is assumed that $x_d(t)$ is bounded and there exists a continuous function $h_d(x_d(t)) \in \mathbb{R}^{3\times1}$ such that $\dot{x}_d(t) = h_d(x_d(t))$. Taking the derivative of (5) gives the error dynamics

$$\begin{aligned} \dot{e} &= \dot{x} - \dot{x}_d(t) \\ &= f(x) + g_1(x)\tau_m + g_2(x)u_f - h_d(x_d(t)), \end{aligned} \tag{6}$$

By defining an augmented state as $x_a = \begin{bmatrix} e^T & x_d^T \end{bmatrix}^T \in \mathbb{R}^{6\times1}$ the system dynamics can be written as

$$\dot{x}_a = f_a(x_a) + g_{1a}(x_a)\tau_m + g_{2a}(x_a)u_f, \tag{7}$$

where the f_a, g_{1a},and g_{2a} matrices become

$$f_a(x_a) = \begin{bmatrix} f(e + x_d) - h_d \\ h_d(x_d(t)) \end{bmatrix} g_{1a}(x_a) = \begin{bmatrix} g_1(e + x_d) \\ 0 \end{bmatrix}$$

$$g_{2a}(x_a) = \begin{bmatrix} g_d(e + x_d) \\ 0 \end{bmatrix}.$$

In order to solve the optimal tracking control problem, an infinite horizon value function is defined as

$$V(x_a(t)) = \int_t^{\infty} e^{-\gamma(\tau - t)}[r(x_a(\tau), u(\tau))]d\tau, \tag{8}$$

where $r(x_a, u) = x_a^T Q x_a + U_1(\tau_m) + U_2(u_f)$ where $Q \in \mathbb{R}^{6\times 6}$, $R_1 \in \mathbb{R}^+$, $R_2 \in \mathbb{R}^+$ are positive weights, $U_1(\tau_m)$,$U_2(u_f)$ are defined as

$$\begin{array}{c} U_1(\tau_m) = R_1 \tau_m^2 \\ U_2(u_f) = 2\int_0^{2u_f - 1} tanh^{-1}(v) R_2 dv \end{array}, \tag{9}$$

and $\gamma > 0$ is a discount factor used to ensure the bound of the value function. It is noted that $U_2(u_f)$ is designed with a $tanh^{-1}(v)$ term similar to [10] with shifted bounds to satisfy an asymmetric constraint on u_f in (2). To obtain a closed form solution for the optimal control inputs, the Hamiltonian of the system is defined as

$$\begin{aligned} H(x_a, \tau_m, u_f, \frac{\partial V}{\partial x_a}) = \frac{\partial V}{\partial x_a}^T (f_a(x_a) + g_{1a}(x_a)\tau_m \\ + g_{2a}(x_a)u_f) + r(x_a, u(\tau)) - \gamma V(x_a). \end{aligned} \tag{10}$$

The optimal control inputs satisfy the Hamiltonian-Bellman-Jacobi equation which states that

$$H(x_a^*, \tau_m^*, u_f^*, \frac{\partial V}{\partial x_a}^*) = 0, \tag{11}$$

where V^* is the optimal value function defined as

$$V^*(x_a(t)) = \min_{\tau_m, u_f} \int_t^{\infty} r(x_a, u(x_a))d\tau. \tag{12}$$

The optimal control inputs can be solved using the stationary conditions $\frac{\partial H}{\partial \tau_m} = 0$, $\frac{\partial H}{\partial u_f} = 0$ and the definition of $U_1(\tau_m)$ and $U_2(u_f)$ in (9) giving the closed form solution

$$\tau_m^* = -\frac{1}{2} R_1^{-1} g_{1a}^T \frac{\partial V}{\partial x_a}^*, \tag{13}$$

$$u_f^* = \frac{1}{2}[1 - tanh(\frac{1}{4} R_2^{-1} g_{2a}^T \frac{\partial V}{\partial x_a}^*)]. \tag{14}$$

Based on (14), it is seen that the FES input is constrained as $0 < u_f < 1$.

IV. Online Actor-Critic System Identification and Value Function Approximation

From (13) and (14), it is obvious that knowledge of g_{1a}, g_{2a} and $\frac{\partial V}{\partial x_a}$ are required to solve for the optimal control inputs. Because the person specific parameters are uncertain, three Neural Networks are used to estimate the system dynamics of f_a, g_{1a} and g_{2a}. A fourth NN is developed to estimate $\frac{\partial V}{\partial x_a}$. The NN representation for the dynamics and the value function is given as

$$f_a(x_a) = W_0^T \phi(x_a) + \epsilon_0(x_a), \tag{15}$$

$$g_{1a}(x_a) = W_1^T \phi(x_a) + \epsilon_1(x_a), \tag{16}$$

$$g_{2a}(x_a) = W_2^T \phi(x_a) + \epsilon_2(x_a), \tag{17}$$

$$V(x_a) = W_3^T \phi(x_a) + \epsilon_3(x_a), \tag{18}$$

where $W_0, W_1, W_2 \in \mathbb{R}^{k\times 6}$, and $W_3 \in \mathbb{R}^{k\times 1}$ are ideal NN estimation weights, where $k \in \mathbb{R}^+$ is the number of neurons, $\phi \in \mathbb{R}^{k\times 1}$ is basis function vector, and $\epsilon_0 ... \epsilon_3$ are approximation errors, where $\epsilon_0, \epsilon_1, \epsilon_2 \in \mathbb{R}^{6\times 1}$ and $\epsilon_3 \in \mathbb{R}^1$. The gradient of the value function , $\partial V/\partial x_a$, is then given by

$$\frac{\partial V}{\partial x_a} = \nabla\phi^T W_3 + \nabla\epsilon_3, \tag{19}$$

where $\nabla\phi \in \mathbb{R}^{k\times 6}$ and $\nabla\epsilon_3 \in \mathbb{R}^{6\times 1}$. For online implementation, f_a, g_{1a}, and g_{2a} are approximated as $\hat{f}_a = \hat{W}_0^T \phi(\hat{x}_a)$, $\hat{g}_{1a} = \hat{W}_1^T \phi(\hat{x}_a)$, $\hat{g}_{2a} = \hat{W}_2^T \phi(\hat{x}_a)$ and the value function is estimated as $\hat{V} = \hat{W}_3^T \phi(x_a)$ where $\hat{W}_0$, $\hat{W}_1, \hat{W}_2$,and $\hat{W}_3$ are the estimated weights of the ideal NN weights.

The following assumptions are made about the ideal weights, basis function, and estimation error.

Assumption 1. The ideal NN weights are bounded such that $\|W_0\| ... \|W_3\| \le b_{wi}, i = 0, 1, 2, 3$

Assumption 2. The basis function ϕ is a sigmoid function that is bounded such that $0 < \|\phi(x_a)\| < 1$ and $||\nabla\phi(x_a)|| \le b_{\phi x}$

Assumption 3. The estimation errors and their gradients are bounded such that $\|\epsilon_0\| ... \|\epsilon_3\| \le b_{\epsilon i}$ and $\|\nabla\epsilon_0\| ... \|\nabla\epsilon_3\| \le b_{\epsilon xi}, i = 0, 1, 2, 3$

Using the estimates of g_{1a}, g_{2a} and $\frac{\partial V}{\partial x_a}$ in (16), (17) and (18) the optimal control inputs in (13) and (14) become

$$\tau_m = -\frac{1}{2} R_1^{-1} \hat{g}_{1a}^T \nabla\phi(x_a^T) \hat{W}_3, \tag{20}$$

$$u_f = \frac{1}{2}[1 - tanh(\frac{1}{4} R_2^{-1} \hat{g}_{2a}^T \nabla\phi(x_a)^T \hat{W}_3)]. \tag{21}$$

A. Online Value Function Approximation

It is difficult to obtain a solution for $\frac{\partial V}{\partial x_a}$ using the HJB equation due to its high nonlinearities in the value function derivative ($\frac{\partial V}{\partial x_a}$) and its dependence on full knowledge of system dynamics. Instead, a simplified policy iteration update as shown in [10] can be developed by noting that for any time interval $T > 0$, the value function satisfies

$$\begin{aligned} V(x_a(t - T)) = \int_{t-T}^{t} e^{-\gamma(\tau - t + T)}[x_a^T Q x_a + U_1(\tau_m) \\ + U_2(u_f)]d\tau + e^{-\gamma T} V(x_a(t)). \end{aligned} \tag{22}$$

It is noted that (22) does not depend on the system dynamics and is linear in $\frac{\partial V}{\partial x}$. Using the value function update in (18) and (22), a Bellman error due to the approximation of the value function can be defined as

$$\begin{aligned} e_b \triangleq \int_{t-T}^{t} e^{-\gamma(\tau - t + T)}[x_a^T Q x_a + U_1(\tau_m) \\ + U_2(u_f)]d\tau + W_3^T \Delta\phi. \end{aligned} \tag{23}$$

where $\Delta\phi = e^{-\gamma T}[\phi(x_a) - \phi(x_a(t - T))]$. Using the value function approximation in (23) gives

$$\begin{aligned} \hat{e}_b = \int_{t-T}^{T} e^{-\gamma(\tau - t + T)}[x_a^T Q x_a + U_1(u_1(\tau)) \\ + U_2(u_2(\tau))]d\tau + \hat{W}_3^T \Delta\phi. \end{aligned} \tag{24}$$

The update law for $\hat{W}_3$ can be found by minimizing the objective function

$$J_{w3} = \frac{1}{2}\hat{e}_b^2. \tag{25}$$

Using the gradient descent algorithm and chain rule, the update law is given as

$$\dot{\hat{W}}_3 = \frac{-\alpha\Delta\phi\hat{e}_b}{(1+\Delta\phi^T\Delta\phi)^2}, \tag{26}$$

where $\alpha \in \mathbb{R}^+$ is a positive gain that represents the learning rate, and $(1+\Delta\phi^T\Delta\phi)^2$ is used for normalization. To facilitate the closed loop stability analysis, (23) can be rewritten as

$$\int_{t-T}^{T} e^{-\gamma(\tau-t+T)}[x_a^T Q x_a + U_1(u_1(\tau)) \\ +U_2(u_2(\tau))]d\tau = -W_3^T\Delta\phi + e_b. \tag{27}$$

On substituting (27) into (24), we get

$$\hat{e}_b = -\tilde{W}_3\Delta\phi + e_b, \tag{28}$$

where $\tilde{W}_3 = W_3 - \hat{W}_3$. Using (28) in (26), we get

$$\dot{\tilde{W}}_3 = -\dot{\hat{W}}_3 = -\alpha\Delta\bar{\phi}\Delta\bar{\phi}^T\tilde{W}_3 + \alpha\frac{\Delta\bar{\phi}}{m}e_b, \tag{29}$$

where $m = 1+\Delta\phi^T\Delta\phi$ and $\Delta\bar{\phi} = \frac{\Delta\phi}{(1+\Delta\phi^T\Delta\phi)}$. By using (26) (Theorem 3 in [10]), if $\Delta\bar{\phi}$ is persistently exciting (PE), i.e.,

$$\gamma_1 I \leq \int_t^{t+T_1} \bar{\Delta}\phi(\tau)\bar{\Delta}\phi(\tau)d\tau \leq \gamma_2 I \tag{30}$$

where $\gamma_1, \gamma_2 \in \mathbb{R}^+$ are constants and $I \in \mathbb{R}^{k\times k}$ is an identity matrix is satisfied, the NN weight approximation error converges exponentially to zero if $e_b = 0$ or converges exponentially to a residual set for a bounded bellman error. Due to the PE condition, the boundedness of $\Delta\bar{\phi}^T\tilde{W}_3$ implies that $\tilde{W}_3$ is bounded. This property will be used in Theorem 2 to show stability of the closed loop error system with a neural network value function approximation.

B. Online Actor NN Design

To guarantee stability of the closed-loop system, an actor NN is used for both the FES and motor inputs and is designed as

$$\hat{\tau}_m = -\frac{1}{2}R_1^{-1}\hat{g}_{1a}^T\nabla\phi(x_a^T)\hat{W}_4, \tag{31}$$

$$\hat{u}_f = \frac{1}{2}[1 - tanh(\frac{1}{4}R_2^{-1}\hat{g}_{2a}^T\nabla\phi(x_a)^T\hat{W}_5)], \tag{32}$$

where $\hat{W}_4, \hat{W}_5 \in \mathbb{R}^{k\times 1}$ are NN weight estimates for the actors. The update laws for $\hat{W}_4$ and $\hat{W}_5$ are determined using a gradient descent approach to minimize the error between the control input with only the value function approximation and the input with the actor NNs with additional terms added to ensure stability of the closed-loop system in Theorem 2.

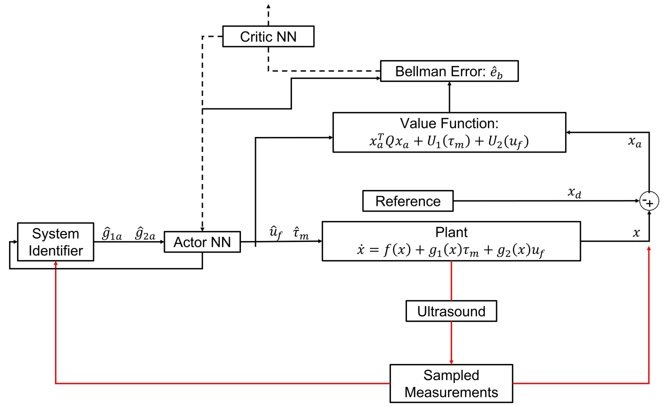

Figure 1. Combined control architecture with the system identifier and sampled measurements used to determine the control inputs in (31) and (32).

The error between the control input with the critic and actor is defined as

$$e_m = \tau_m - \hat{\tau}_m \\ = -\frac{1}{2}R_1^{-1}\hat{g}_{1a}^T\nabla\phi^T\hat{W}_3 + \frac{1}{2}R_1^{-1}\hat{g}_{1a}^T\nabla\phi^T\hat{W}_4, \tag{33}$$

$$e_f = u_f - \hat{u}_f \\ = -tanh(\frac{1}{4}R_1^{-1}\hat{g}_{2a}^T\nabla\phi^T\hat{W}_3) + tanh(\frac{1}{4}R_1^{-1}\hat{g}_{2a}^T\nabla\phi^T\hat{W}_4), \tag{34}$$

for both motor and FES inputs respectively.

The objective functions to be minimized by the actor NNs for both motor and FES inputs are defined as $J_{w4} = \frac{1}{2}e_m^2$ and $J_{w5} = \frac{1}{2}e_f^2$ respectively. Using the gradient descent algorithm and chain rule, the update law for actor NN on the motor input becomes

$$\dot{\hat{W}}_4 = -\alpha_2(\frac{1}{2}R_1^{-1}\nabla\phi\hat{g}_{1a}e_m + \frac{k_1}{2}R_1^{-1}\nabla\phi\hat{g}_{1a} + Y_1\hat{W}_4), \tag{35}$$

where $Y_1, k_1 \in \mathbb{R}^+$ are gains to ensure stability. Similarly, the update law for the actor NN on the FES input becomes

$$\dot{\hat{W}}_5 = -\alpha_3(\frac{e_f}{4}R_2^{-1}\nabla\phi\hat{g}_{2a} - \frac{e_f}{4}R_2^{-1}\nabla\phi\hat{g}_{2a}tanh^2(\hat{P}) + Y_2\hat{W}_5) \tag{36}$$

where $\hat{P} = -\frac{1}{4}R_2^{-1}\phi^T\hat{W}_2\nabla\phi^T\hat{W}_5$ and $Y_2 \in \mathbb{R}^+$ is a gain to ensure stability. A flow chart of the the complete control architecture is presented in Fig. 1.

C. Online System Identifier Design with Sampled Fatigue Measurements

The goal of the system identifier is to generate a continuous estimate of the dynamic system given in (7) to be used in $\hat{g}_{1a}$ and $\hat{g}_{2a}$ in the control laws defined in (13) and (14) while simultaneously accounting for the sampled US fatigue measurements. The measurements available in real-time for the system defined in (4) can be written as $y = \begin{bmatrix} x_1(t) & x_2(t) & x_3(t_k) \end{bmatrix}^T$ where $x_1(t)$ and $x_2(t)$ are the angular position and velocity of the limb as previously defined and are measurable at a high sampling frequency by using either IMUs or encoders and $x_3(t_k)$ is the US-based fatigue measurements that are available at discrete instants t_k. It is assumed that the US-based measurement available at t_k is held constant until a subsequent measurement is available at time instant t_{k+1}. The sampling interval between two consecutive measurements at t_k and t_{k+1} is a positive constant denoted as T.

The dynamic system in (7) with optimal control inputs (20) and (21) can be represented by replacing the unknown functions f, g_{1a}, and g_{2a} with NN approximations as

$$\dot{\hat{x}}_a = \hat{f}_a(\hat{x}_a) + \hat{g}_{1a}(\hat{x}_a)\hat{\tau}_m + \hat{g}_{2a}(\hat{x}_a)\hat{u}_f + k\tilde{x}_a + \mu + \xi\varepsilon(t_k), \quad (37)$$

where $k, \xi \in \mathbb{R}^+$ are positive constant gains, and $\tilde{x}_a$ is the error between the estimated state and the augmented state x_a defined as $\tilde{x}_a = x_a - \hat{x}_a$, $\mu \in \mathbb{R}^{6\times 1}$ is an auxiliary variable defined in order to facilitate the stability analysis defined as

$$\mu \triangleq \hat{W}_0^T\tilde{\phi}(x_a) + \hat{W}_1^T\tilde{\phi}(x_a)\hat{\tau}_m + \hat{W}_2^T\tilde{\phi}(x_a)\hat{u}_f, \quad (38)$$

where $\tilde{\phi}(x_a) = \phi(x_a) - \phi(\hat{x}_a)$. To account for the sampled US measurements at a lower frequency, the proposed identifier for $\hat{x}_a$ is augmented with an update term: $\xi\varepsilon(t_k) \in \mathbb{R}^{6\times 1}$ for the sampled ultrasound (US) measurements defined as $\varepsilon(t_k) = \begin{bmatrix} 0 & 0 & \varepsilon_3(t_k) & 0 & 0 & 0 \end{bmatrix}^T$ where $\varepsilon_3(t_k) \in \mathbb{R}$ is defined as $\varepsilon_3(t_k) = x_3(t_k) - x_{d3}(t_k) - \hat{x}_{3a}(t_k) = \tilde{x}_{3a}$, where $x_3(t_k)$ and x_{d3} are the normalized US-based fatigue measurement and desired fatigue value at t_k and $\hat{x}_{3a}(t_k)$ is the estimated value of the fatigue error at t_k. $\varepsilon_3(t_k)$ has an upper bound of $\bar{\varepsilon}_3 \in \mathbb{R}^+$ and $\varepsilon_3(t_k)$ is a constant value during every time interval $[t_k, t_{k+1}]$ after which an US-based fatigue measurement is available. The estimation error dynamics can then be written as

$$\begin{aligned}\dot{\tilde{x}}_a &= \dot{x}_a - \dot{\hat{x}}_a \\ &= \tilde{f} + \tilde{g}_{1a}\hat{\tau}_m + \tilde{g}_{2a}\hat{u}_f - k\tilde{x}_a - \xi\varepsilon(t_k) - \mu + \epsilon, \end{aligned} \quad (39)$$

where $\tilde{f}_a = f_a - \hat{f}_a$, $\tilde{g}_{1a} = g_{1a} - \hat{g}_{1a}$, $\tilde{g}_{2a} = g_{2a} - \hat{g}_{2a}$, $\epsilon = \epsilon_0 + \epsilon_1\hat{\tau}_m + \epsilon_2\hat{u}_f$ represents the bounded NN estimation errors for W_0, W_1 and W_2 respectively. The update laws are designed based on the stability analysis in Theorem as 1

$$\dot{\hat{W}}_0 = proj(\Gamma_0(\phi(x_a)\tilde{x}_a^T - \iota_0\hat{W}_0)), \quad (40)$$

$$\dot{\hat{W}}_1 = proj(\Gamma_1(-\frac{1}{2}\phi(x_a)R_1^{-1}\phi(x_a)^T\hat{W}_1\nabla\phi_3^T\hat{W}_3\tilde{x}_a^T - \iota_1\hat{W}_1)), \quad (41)$$

$$\dot{\hat{W}}_2 = proj(\Gamma_2(\phi(x_a)\hat{u}_f\tilde{x}_a^T - \iota_2\hat{W}_2)), \quad (42)$$

where $\Gamma_0, \Gamma_1, \Gamma_2 \in \mathbb{R}^{k\times k}$ are positive gain matrices, $\iota_0, \iota_1, \iota_2 \in \mathbb{R}^+$ are positive constants and $proj()$ is a smooth operator that bounds the NN weights [11]. It is noted that based on the projection operator, the optimal control laws in (20) and (21), and assumptions 1-3, $\hat{\tau}_m, \hat{u}_f \in \mathcal{L}_\infty$.

Theorem 1. *The identifier designed in (37) along with weight updates in (40),(41) and (42) ensure that the sampled identifier error $\tilde{x}_a$ is globally uniformly ultimately bounded (G.U.U.B).*

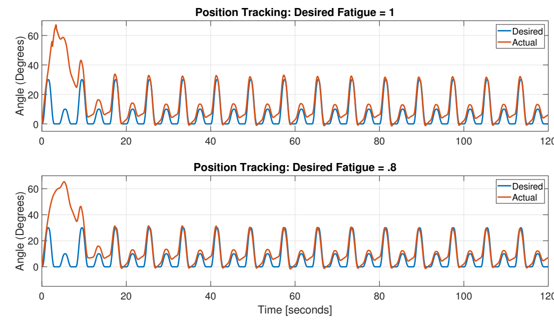

Figure 2. Position tracking at both discrete fatigue levels.

Proof: The proof to Theorem 1 is available upon request. ■

Remark 1: The proof of Theorem 1 shows that the sampled estimation error in (39) is bounded. However the boundedness of $\tilde{x}_a$ does not ensure convergence of the neural network weights. In order to ensure convergence of the weights, $\tilde{W}_0$, $\tilde{W}_1$ and $\tilde{W}_2$ should satisfy persistence of excitation conditions. Future versions of this analysis will address the convergence of the neural network weights.

Theorem 2. *The control inputs for motor and FES defined in (31) and (32) along with the NN update law in (26), (35), and (36) make the closed loop error system defined in (7) and the weight estimation error for the value function GUUB.*

Proof: The proof to Theorem 2 is available upon request. ■

V. Simulation Results

Simulations were implemented for a leg extension system in MATLAB under two different on-demand fatigue levels: 1 and 0.8. The desired position and velocity trajectories were designed using a third order polynomial for knee extension patterns during the gait cycle and the trajectory was simulated for a total duration of 2 minutes of continuous knee extension using simulated discrete US-based measurements that were updated every 5 seconds. To ensure the PE condition in the simulations, a probing noise modeled as $n(t) = sin^2(t)cos(t) + sin^2(2t)cos(5t)$ was added to the system. The model parameters for the combined fatigue and leg extension system were determined experimentally for an able body participant using the approach described in [9]. Fig. 2 shows the joint position tracking at both desired fatigue levels for the two minute duration. It is clear that after two knee extension cycles, the position tracking reaches steady state. When the desired fatigue level was 1 the steady state root mean squared error (RMSE) was 3.80 degrees. In comparison the steady state RMSE when the desired fatigue level was 0.8 was 4.09 degrees. Fig. 3 shows the motor torque along with normalized FES input and its corresponding torque generation calculated using (2) under both desired fatigue conditions. To maintain the desired fatigue at 1, most of the control effort is taken up by motors. However it is seen that as the desired fatigue level is adjusted from 1 to

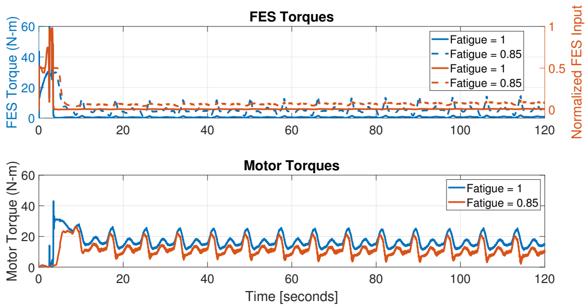

Figure 3. Motor torques, normalized FES input, and its corresponding torque generation under both desired fatigue conditions. The torque generated by FES was calculated using the dyanmic model based on pre-identifed person specific parameters.

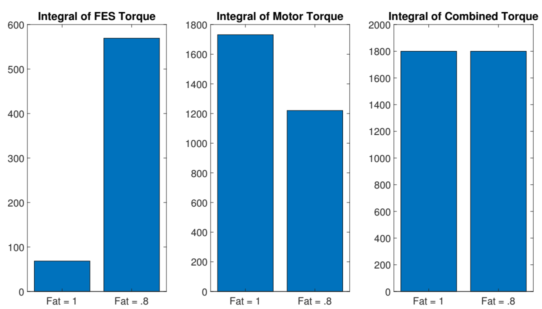

Figure 4. Integral of steady state FES and motor inputs along with total combined torque under both desired fatigue scenarios.

0.8 the motor torque decreases while the torque generated from the FES input increases while tracking performance is maintained. This is further highlighted by taking the integral of the steady state motor torque and normalized FES input as seen in Fig. 4. Clearly, as the desired fatigue increase more control effort is placed on the FES in comparison to the motor while the total torque generated remains the same in both scenarios. The simulated US fatigue measurements for each scenario are shown in Fig. 5 and it is seen that the fatigue approaches is desired value during each scenario. Thus the developed actor-critic system identifier scheme has the ability to determine the optimal allocation of motor torques and FES with using sampled US fatigue measurements to produce cyclic knee extension for an extended duration at different desired fatigue profiles while generating the similar steady-state tracking performance.

VI. Conclusion

In this paper, an ACI approach is proposed to solve the optimal tracking control problem with asymmetric FES constraints and US-based fatigue measurements sampled at a low frequency to optimally allocate motor and FES torque in a hybrid exoskeleton system. An NN-based system identifier was designed to estimate the unknown dynamics despite the low sampled fatigue measurement and simulation results show knee extension tracking performance and torque allocation at two on-demand fatigue levels. Further, it is seen that total

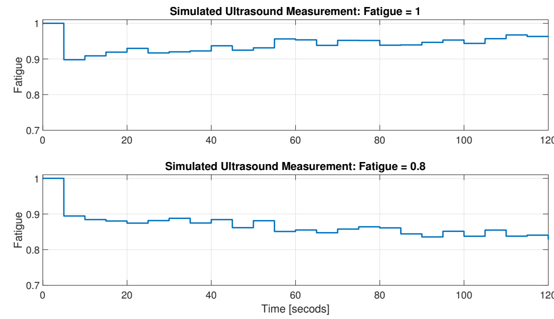

Figure 5. Simulated US measurements when desired fatigue was set to 1 and 0.8 respectively.

torque administered to the system torque remains consistent while the desired fatigue is manually adjusted and FES and motor torques are allocated accordingly. Our preliminary results demonstrates that the ACI approach is an effective and promising method to automatically allocate electric motor and FES . Potentially, optimal adaptive tracking control of a hybrid exoskeleton seems feasible without prior knowledge of the system dynamics and desired fatigue levels may be maintained based on low-sampled fatigue measurements.

References

[1] K. H. Ha, S. A. Murray, and M. Goldfarb, "An approach for the cooperative control of fes with a powered exoskeleton during level walking for persons with paraplegia," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 24, no. 4, pp. 455–466, 2015.

[2] N. A. Kirsch, X. Bao, N. A. Alibeji, B. E. Dicianno, and N. Sharma, "Model-based dynamic control allocation in a hybrid neuroprosthesis," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 26, no. 1, pp. 224–232, 2017.

[3] F. L. Lewis, D. Vrabie, and V. L. Syrmos, *Optimal control*. John Wiley & Sons, 2012.

[4] D. Vrabie, O. Pastravanu, M. Abu-Khalaf, and F. L. Lewis, "Adaptive optimal control for continuous-time linear systems based on policy iteration," *Automatica*, vol. 45, no. 2, pp. 477–484, 2009.

[5] 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.

[6] S. Bhasin, R. Kamalapurkar, M. Johnson, K. G. Vamvoudakis, F. L. Lewis, and W. E. Dixon, "A novel actor–critic–identifier architecture for approximate optimal control of uncertain nonlinear systems," *Automatica*, vol. 49, no. 1, pp. 82–92, 2013.

[7] Z. Sheng, N. Sharma, and K. Kim, "Quantitative assessment of changes in muscle contractility due to fatigue during nmes: An ultrasound imaging approach," *IEEE Transactions on Biomedical Engineering*, vol. 67, no. 3, pp. 832–841, 2019.

[8] Q. Zhang, A. Iyer, Z. Sun, A. Dodson, and N. Sharma, "Sampled-data observer based dynamic surface control of delayed neuromuscular functional electrical stimulation," in *Dynamic Systems and Control Conference*, vol. 84270. American Society of Mechanical Engineers, 2020, p. V001T14A003.

[9] N. Kirsch, N. Alibeji, and N. Sharma, "Nonlinear model predictive control of functional electrical stimulation," *Control Engineering Practice*, vol. 58, pp. 319–331, 2017.

[10] H. Modares and F. L. Lewis, "Optimal tracking control of nonlinear partially-unknown constrained-input systems using integral reinforcement learning," *Automatica*, vol. 50, no. 7, pp. 1780–1792, 2014.

[11] M. Krstic, P. V. Kokotovic, and I. Kanellakopoulos, *Nonlinear and adaptive control design*. John Wiley & Sons, Inc., 1995.