A dual control approach to solve exploration vs. exploitation trade-offs in the design of personalized physical exercise sessions

Roxanne R. Jackson[a], Damiano Varagnolo[b,c], and Steffi Knorn[a]

***Abstract*— Using biofeedback in medical therapies has proven to be effective for adapting patient behaviors while keeping the patients engaged and motivated in an exercise session. This paper considers general problems in personalized exercise sessions where the input is opportune biofeedback and the session goal is to maximize a particular exercise effect. Due to the individual differences between patients and their physiological signals, however, personalized patient models also need to be identified. With the two objectives: 1) maximize a training effect with minimal control effort, and 2) identify the individualized patient model, we have a typical exploration vs. exploration trade-off. Control problems of this form are called *dual control* problems. In this paper, we formulate a dual control problem for a personalized exercise session and test the approach against classical optimal control and optimal experimental design approaches in an illustrative example of performing Kegel exercises where the control and identification goals conflict with each other.**

***Index Terms*— Optimal control; Identification; Healthcare and medical systems**

I. Introduction

THE recent abundance of physiological data from easy-to-use medical devices has led to a surge in the use of biofeedback in therapy and physical exercise. Non-invasive, psycho-physiological sensors may provide real-time information to monitor and influence users' behavior, and thus perform biofeedback actions that can range from adapting the visual, tactile, or auditory environment of the user [1]. This approach has become an effective tool in the treatment of stress and anxiety-related disorders [2] as well as chronic pain [3], irritable bowel syndrome [4], and incontinence [5]. Recently, biofeedback has been adopted in personalizing training for aerobic fitness [6]. Physical exercising with the aid of biofeedback devices has the unique advantage of being a non-pharmaceutical intervention for enabling individuals to control their bodies in a non-invasive and low-risk way.

On the other hand, biofeedback by itself is not guaranteed to yield results in medical therapies and exercising, since improving one's condition is often achieved only after consistent commitment. The inconsistency of feedback to affirm the results of the treatment often leads to high therapy dropout rates and irregular exercising; in these cases, the exercising effect is effectively lost. Moreover, for those executing the exercises incorrectly, often no feedback is available to rectify their behavior. In essence, there seems to be a need for improving the efficacy of providing feedback both *during* an exercising session (to aid the user to adapt their own behavior within an individual session and to monitor the effectiveness of that single exercising session), and *across* the exercising season (to aid the user focus towards long-term goals and to monitor the combined effect of the time series of sessions).

Due to the stark individual differences between patients and physiological signals, the successful incorporation of biofeedback within and across exercise sessions requires a personalized patient model. Such individualized models shall then be able to account for the fact that, in time, exercising sessions may evolve, daily form may change, and sensor placement may vary. Such models shall moreover be identified typically with a minimal amount of samples. This highlights then a typical exploration vs. exploitation trade-off: design a series of exercising sessions whose medical effect is maximal within the shortest period of time, vs. collecting data maximally useful for model learning purposes.

Contribution: We propose a model structure that may aid in designing in-session and across-sessions biofeedback actions, and for which it is possible to frame the exploration vs. exploitation problem above as a Dual Control (DC) problem. We thus propose a model-based exercising sessions-design methodology for tackling such a trade-off between achieving a desired exercising effect and identifying the system to aid in the design of more efficient and effective exercise sessions.

Structure of the manuscript: We first formalize the structure of the proposed model and define the aims of the exercising session design as an optimal input design. We follow this by formalizing the sides of the trade-off above as two separate Optimal Control (OC) and Optimal Experimental Design (OED) problems. Next, we bring the two problems together and define a general cost function of interest where the trade-off is captured by a specific weight. Finally, we demonstrate the effect of trading off between the two objectives in a numerical example of a validated Kegel exercise model.

Notation: We use uppercase letters for random variables and the corresponding lowercase letters for realizations.

[a] Department of Control Engineering, Technical University of Berlin, Germany; {jackson | knorn}@tu-berlin.de

[b] Department of Engineering Cybernetics, Norwegian University of Science and Technology, Norway; damiano.varagnolo@ntnu.no

[c] Department of Information Engineering, University of Padua, Italy

II. Exercise Session Dynamics

a) Model structure: The majority of systems in medicine and biology often exhibit highly nonlinear and complex dynamical behavior. During exercise sessions, variables such as sensor placement, daily form, and concentration can cause time-variant behavior. Linear time-invariant models are commonly used in medical and biological systems, but non-linearities may be important in some cases, requiring linearizable or explicit non-linear system considerations. Notably, updating the system parameters is crucial to obtaining the best model for each exercise session.

Different therapies may focus on different parts of the body, and thus focus on models that may be different since serving different needs. We assume that the particular model of interest is within the class of discrete-time systems writable as

$$x_{k+1} = f(x_k, u_k, \theta) + w_k, \tag{1}$$
$$y_k = h(x_k, u_k, \theta) + v_k, \tag{2}$$
$$z_k = \varphi(x_k, u_k)\,, \tag{3}$$

where $x_k \in \mathbb{R}^{n_x}$ are the states of the system (e.g., status of specific muscles), $u_k \in \mathcal{U} \subset \mathbb{R}^{n_u}$ are the constrained inputs (e.g., intensity of the physical activity), and $y_k \in \mathbb{R}^{n_y}$ are the measurements (e.g., heart rate, exerted muscular force). In addition, we assume the designer to have identified an **exercising effect** variable, denoted here as $z_k \in \mathbb{R}^{n_z}$ (e.g., some physical endurance index). z_k is assumed to be a non-measured output equal to a deterministic function of the current states and inputs. $\theta \in \Theta \subset \mathbb{R}^{n_\theta}$ is the vector of (unknown) system parameters. Accordingly, the function $f : \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \times \mathbb{R}^{n_\theta} \to \mathbb{R}^{n_x}$ is the state transition function, $h : \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \times \mathbb{R}^{n_\theta} \to \mathbb{R}^{n_y}$ is the measurement function; where the structure is known for both; and $\varphi : \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \to \mathbb{R}^{n_z}$ is the function defining the exercising effect z. We assume the system to be stochastic with process uncertainty $w_k \sim \mathcal{N}(0, Q)$ due to the initial condition uncertainty and the propagation of uncertainty from the parameter estimation. Similarly, $v_k \sim \mathcal{N}(0, R)$ is the measurement uncertainty. We assume here additive independent and identically distributed (i.i.d) Gaussian noise, which is legitimate for certain applications, but it should be carefully evaluated based on the specific physiological signals under consideration, see e.g. [7], [8].

We thus assume that the designer works with a specific instance of the model defined by (1)–(3), referred to in the remainder of the paper as the **exercising session system**. We show later how this may be used to design biofeedback actions to adapt the user's actions for a specific instance of the model, making it applicable for physiotherapy purposes.

b) System dynamics: We assume the initial value X_0 follows a known probability distribution, and at the beginning of each exercising session, the system parameters θ are unknown. Also, at any time, some data management system has collected the information up to and including time k. We denote this information via $\mathcal{D}_{0:k} = (X_0, U_0, \ldots, U_{k-1}, Y_1, Y_2, \ldots, Y_k)$ with $\mathcal{D}_0 = X_0$ and dimension $n_{\mathcal{D}_{0:k}}$.

c) Aim: Find an algorithm to recursively estimate, given $\mathcal{D}_{0:k-1}$, the inputs that maximize the current exercising effect z_k. Thus find the control policy sequence $U_{0:N-1} = (U_0, U_1, \ldots, U_{N-1})$ for $N \geq 1$ such that $U_{k+1} = \mu_k(\mathcal{D}_{0:k})$, where $\mu_k : \mathbb{R}^{n_{\mathcal{D}_{0:k}}} \to \mathcal{U}$.

Such an algorithm has to include the two ingredients of identifying the system and maximizing the exercising effect with minimal control effort, which correspond to two classical paradigms in the control literature, i.e., OC and OED:

Paradigm 1 (OC) *Design a minimal (control) effort u that leads to a session that maximizes the exercising effect z.*

Paradigm 2 (OED) *Design a maximally informative effort u that leads to a session for which the estimated exercising effect, $\hat{z}$, may be inferred as statistically accurate as possible.*

In the remainder of the paper we proceed as follows: detail our specific OC problem in Sec. III, formalize our type of OED problem in Sec. IV, and formulate in Sec. V a method to combine the two approaches into a single trading-off problem.

III. Designing an Exercise Session as an Optimal Control Problem

To formulate a session design problem as an optimal control one given model (1)-(3), one may consider a general formulation of the loss function as

$$\mathbb{E}\left[G_N(Z_N) + \sum_{i=k}^{N-1} G_i(Z_i, U_i)\right] =: \mathbb{E}[H(Z_k, U_{k:N-1})] \tag{4}$$

where $G_i : \mathbb{R}^{n_z} \times \mathbb{R}^{n_u} \to \mathbb{R}$ are the stage cost functionals, $G_N : \mathbb{R}^{n_z} \to \mathbb{R}$ is the terminal cost functional, and the expected loss is chosen to be a sensible criterion of interest since exercising effects Z_k are stochastic variables.

An optimal control approach would then be finding the admissible control strategy for system (1)–(3) minimizing (4). We can use Lemma 3.2 in [9] to find that we shall condition on the information collected until the decision process, i.e.,

$$\begin{aligned}
&\min_{U_{k:N-1}(\mathcal{D}_{0:k})} \mathbb{E}\left[G_N(Z_N) + \sum_{i=k}^{N-1} G_i(Z_i, U_i)\right] \\
&= \min_{u_{k:N-1}} \mathbb{E}\left[G_N(Z_N) + \sum_{i=k}^{N-1} G_i(Z_i, u_i) \middle| \mathcal{D}_{0:k} = d_{0:k}\right] \\
&= V(d_{0:k}, k)\,.
\end{aligned} \tag{5}$$

Since the dimension of $\mathcal{D}_{0:k}$ increases with k, and since that conditional expectation may become numerically intractable, we can simplify (5) by conditioning on our best state estimate given the data $\mathcal{D}_{0:k}$ instead. Thus, letting $\widehat{x}_{k|k} = \mathbb{E}\left[X_k | \mathcal{D}_{0:k} = d_{0:k}\right]$ and $\widehat{z}_k = \mathbb{E}\left[\varphi(X_k, U_k) | X_k = \widehat{x}_{k|k}\right]$, we consider the new objective

$$\begin{aligned}
&W(\widehat{x}_{k|k}, k) = \\
&\min_{u_{k:N-1}} \mathbb{E}\left[G_N(Z_N) + \sum_{i=k}^{N-1} G_i(\varphi(X_i, u_i), u_i) \middle| X_k = \widehat{x}_{k|k}\right].
\end{aligned} \tag{6}$$

From computational perspectives, $\widehat{x}_{k|k}$ has a constant dimension as k increases. We note that it is not necessary to restrict $\widehat{x}_{k|k}$ to be a specific type of filter.

Remark *In [9], it is shown for Linear Quadratic Gaussian (LQG) problems, that $\widehat{x}_{k|k}$ is a sufficient statistic given the* *data $\mathcal{D}_{0:k}$, i.e., $V(\mathcal{D}_{0:k}, k) = W(\widehat{x}_{k|k}, k)$; however, equality does not hold generally.*

IV. Designing an optimal experimental design for an exercising session

If one wishes to be able to compute estimates $\widehat{z}$ as precisely as possible, one may wish to promote data informativity by means of OED techniques. This means that, given an appropriate model structure and measurement data, $y_{0:N}$, one wishes to ensure low uncertainty about some Key Performance Indicator (KPI). Given our model choice, we consider the classical case of additive, normally distributed, uncorrelated measurement error, for which the maximum likelihood estimate

$$\begin{aligned}\widehat{\theta}_{\mathrm{ML}} &= \arg\max_{\theta} p_{Y|\theta}(y_{1:N}|\widehat{x}_{0:N}, u_{0:N-1}, \theta) \\ &= \arg\min_{\theta} \frac{N}{2}\log(2\pi) + \frac{1}{2}\log|R| + \frac{1}{2}\sum_{k=1}^{N} e_k^\top R^{-1} e_k \,,\end{aligned} \tag{7}$$

where $e_k = y_k - h(x_k, u_k, \theta)$ and $p_{Y|\theta}$ is the conditional probability density function (pdf) of the measurements given the parameters, is of statistical relevance [10].

A. Information matrix and optimality criteria

Given the parametric estimation problem of estimating θ in (1)-(2), we consider the classical focus on the optimal, unbiased estimate with minimal covariance. In other words, letting θ^* be the (assumed existing) true parameters vector and $M\left(\widehat{\theta}, u\right)$ be the Fisher Information Matrix (FIM)

$$M\left(\widehat{\theta}, u\right) = \mathbb{E}\Big[\, L^\top L\big|_{\theta=\widehat{\theta}} \Big] \,, \qquad L := \frac{\partial \log p_{Y|\theta}(y_{1:N}|\theta)}{\partial\theta} \,, \tag{8}$$

the Cramer-Rao inequality [10]

$$\mathbb{E}\left[\left(\widehat{\theta} - \theta^*\right)^\top \left(\widehat{\theta} - \theta^*\right)\right] \geq M\left(\widehat{\theta}, u\right)^{-1} \,, \tag{9}$$

follows, stating that the covariance of the estimation error cannot be smaller (in a positive definite sense) than the inverse of the FIM. To achieve the lower limit – the minimum variance unbiased estimator – several optimality criteria based on the FIM have been proposed, including the so-called A-criteria given by the trace $\Phi_A\left(M\left(\widehat{\theta}, u\right)\right) = \mathrm{tr}\left(M\left(\widehat{\theta}, u\right)^{-1}\right)$, the D-criteria, given by the determinate $\Phi_D\left(M\left(\widehat{\theta}, u\right)\right) = \det\left(M\left(\widehat{\theta}, u\right)^{-1}\right)$, and the modified E-criteria as $\Phi_{mE}\left(M\left(\widehat{\theta}, u\right)\right) = \frac{\lambda_{\max}\left(M\left(\widehat{\theta}, u\right)^{-1}\right)}{\lambda_{\min}\left(M\left(\widehat{\theta}, u\right)^{-1}\right)}$, where $\lambda_{\max}$ and $\lambda_{\min}$ are the maximum and minimum eigenvalues of the inverse of the FIM, respectively.

B. Computing the Fisher information matrix in practice

As noticed in [11] the criteria above may suffer from issues due to variable scaling, leading to poor estimates of parameters with larger values. We thus propose to consider solving the scaled OED variant

$$\min_{u_{0:N-1}} \Phi\left(\mathrm{diag}\left(\widehat{\theta}\right) M\left(\widehat{\theta}, u\right) \mathrm{diag}\left(\widehat{\theta}\right)\right) \,. \tag{10}$$

Moreover from (7) it follows that (8) can be written as

$$M\left(\widehat{\theta}, u\right) = \sum_{k=0}^{N-1}\left[\left(\frac{\mathrm{d}h_k}{\mathrm{d}\theta}\right)^\top R^{-1}\left(\frac{\mathrm{d}h_k}{\mathrm{d}\theta}\right)\right]\Bigg|_{\theta=\widehat{\theta}} \,, \tag{11}$$

where $h_k = h(x_k, u_k, \theta)$. It is then possible to define the sensitivities of the states and outputs as

$$s_k = \left.\frac{\mathrm{d}x_k}{\mathrm{d}\theta}\right|_{\theta=\widehat{\theta}} \,, \qquad \frac{\mathrm{d}h_k}{\mathrm{d}\theta} = \left.\frac{\partial h}{\partial x}\frac{\mathrm{d}x_k}{\mathrm{d}\theta}\right|_{\theta=\widehat{\theta}} + \left.\frac{\partial h}{\partial\theta}\right|_{\theta=\widehat{\theta}} \,. \tag{12}$$

Considering then that

$$\frac{\mathrm{d}}{\mathrm{d}\theta}x_{k+1} = \frac{\mathrm{d}}{\mathrm{d}\theta}f(x_k, u_k, \theta) \,, \tag{13}$$

and the chain rule it follows that

$$s_{k+1} = \frac{\partial f}{\partial x}s_k + \frac{\partial f}{\partial\theta} \,. \tag{14}$$

We note that these sensitivities shall be used not just for computing both FIM, but also the sensitivity of the loss of optimality in the following section IV-C.

C. An adapted OED approach, specific for model (1)–(3)

Standard OED formulations aim to design inputs that maximize the accuracy of parameter estimators. Our most relevant estimand is, though, a function of the states, i.e., the exercising effect z. Given this specific focus, we adapt the standard OED formulation. With respect to the existing literature, the proposed approach has contact points with [12], where authors formulate a weighted A-criteria approach for simultaneously solving (offline) an experimental design and trajectory tracking problem. The α weight in [12] is, though, tailored specifically for trajectory tracking. Our work connects also with [13], proposing a weighted G-criterion for a dual Nonlinear Model Predictive Control (NMPC) problem, and presenting a statistical method for tuning the parameter α specifically for NMPC formulations. [12] and [13] are meaningful approaches, but not appropriate for a setup (as the one we consider) where *a)* no trajectory for the wished exercise effect is defined, but rather a terminal maximization of the effect, and *b)* high model uncertainty is expected, something that may render NMPC formulations too sensitive to such modeling errors and thus perform poorly. The dual control approach is yet to be applied to an exercise session dynamical systems of our formulation.

In our approach, we form a loss of optimality for the OC problem in (6), which is specific for the exercise session system in (1)–(3) where we focus on the exercise effect z.

Consider then the proposed optimal control problem in Section III, focusing on finding $u^*_{k:N-1}$ (e.g., the current intensity of a suggested physical activity) that solves (6). However, since the parameters defining model (1)–(3) are only estimates, (6) is actually solved using $\widehat{\theta}$ and not the (assumed existing) true parameters. This leads to a hypothetical loss of optimality, i.e.,

$$\Delta H_k\left(\widehat{\theta}\right) = H\left(Z_k, u^*, \widehat{\theta}\right) - H\left(Z_k, u^*, \theta^*\right) \,. \tag{15}$$

From intuitive standpoints, we let the proposed OED approach try to minimize the expected loss with respect to the state estimates, i.e., $\mathbb{E}\left[\Delta H_k\left(\widehat{\theta}\right)\right]$. Then, Taylor-expanding this expectation around the unknown true parameters $\widehat{\theta} = \theta^*$ yields

$$\begin{aligned}\mathbb{E}\left[\Delta H_k\left(\widehat{\theta}\right)\right] \approx \mathbb{E}\Bigg[&\Delta H_k\left(\theta^*\right) + \left.\frac{\mathrm{d}\Delta H_k\left(\widehat{\theta}\right)}{\mathrm{d}\widehat{\theta}}\right|_{\widehat{\theta}=\theta^*} \mathrm{d}\theta \\ &+ \frac{1}{2}\mathrm{d}\theta^\top \left.\frac{\mathrm{d}^2\Delta H_k\left(\widehat{\theta}\right)}{\mathrm{d}\widehat{\theta}^2}\right|_{\widehat{\theta}=\theta^*} \mathrm{d}\theta\Bigg] .\end{aligned} \tag{16}$$

The first term in (16) is $H\left(Z_k, u^*, \theta^*\right) - H\left(Z_k, u^*, \theta^*\right) = 0$. As for the second term, first-order optimality conditions (assuming smoothness of such cost and optima that belong to the interior of the domain) require $\left.\frac{\mathrm{d}H\left(Z_k,u^*,\widehat{\theta}\right)}{\mathrm{d}\widehat{\theta}}\right|_{\widehat{\theta}=\theta^*} = 0$. Here we can use (15) to compute the sensitivity of the loss to be

$$\begin{aligned}\frac{\mathrm{d}H(Z_k,u^*,\theta)}{\mathrm{d}\theta} &= \frac{\mathrm{d}G_N(Z_N)}{\mathrm{d}\theta} + \sum_{i=k}^{N}\frac{\mathrm{d}G_i(Z_i,u_i^*)}{\mathrm{d}\theta} \\ &= \sum_{i=k}^{N}\frac{\mathrm{d}G_i(\varphi(X_i,u_i^*,\theta),u_i^*)}{\mathrm{d}\theta} \\ &= \sum_{i=k}^{N-1}\frac{\partial G_i}{\partial\varphi}\frac{\partial\varphi}{\partial X}\frac{\mathrm{d}X_i}{\mathrm{d}\theta} = \sum_{i=k}^{N-1}\frac{\partial G_i}{\partial\varphi}\frac{\partial\varphi}{\partial X}S_i ,\end{aligned} \tag{17}$$

where S_i are the sensitivities of the states. We will denote the sensitivities of the loss in (17) as $\nabla_\theta H_k(\theta)$. Since traces have the cyclic property and are equal to the scalar if applied to a scalar, the expected loss of optimality is

$$\begin{aligned}&\mathbb{E}\left[\Delta H\left(\widehat{\theta}\right)\right] \approx \\ &\quad\approx \mathbb{E}\left[\frac{1}{2}\left(\widehat{\theta}-\theta^*\right)^\top \nabla_\theta H_0\left(\theta^*\right)^\top \nabla_\theta H_0\left(\theta^*\right)\left(\widehat{\theta}-\theta^*\right)\right] \\ &\quad= \mathbb{E}\left[\frac{1}{2}\mathrm{tr}\left(\left(\widehat{\theta}-\theta^*\right)^\top \nabla_\theta H_0\left(\theta^*\right)^\top \nabla_\theta H_0\left(\theta^*\right)\left(\widehat{\theta}-\theta^*\right)\right)\right] \\ &\quad= \mathrm{tr}\left(\nabla_\theta H_0\left(\theta^*\right)\mathbb{E}\left[\left(\widehat{\theta}-\theta^*\right)\left(\widehat{\theta}-\theta^*\right)^\top\right]\nabla_\theta H_0\left(\theta^*\right)^\top\right) .\end{aligned} \tag{18}$$

Since the true parameter values are unknown, we suggest using the approximation $\nabla_\theta H_0\left(\theta^*\right) \approx \nabla_\theta H_0\left(\widehat{\theta}\right)$ and inequality (9) to formulate the approximate expected loss of optimality

$$\mathbb{E}\left[\Delta H\left(\widehat{\theta}\right)\right] \approx \mathrm{tr}\left(\nabla_\theta H_0\left(\widehat{\theta}\right)\widetilde{M}(\widehat{\theta},u)^{-1}\nabla_\theta H_0\left(\widehat{\theta}\right)^\top\right) . \tag{19}$$

This leads to the possibility of finding the sought inputs via the optimization problem

$$\min_{u_{0:N-1}} \mathrm{tr}\left(\nabla_\theta H_0\left(\widehat{\theta}\right)\widetilde{M}(\widehat{\theta},u)^{-1}\nabla_\theta H_0\left(\widehat{\theta}\right)^\top\right) . \tag{20}$$

Remark *The modified OED problem in* (20) *is a scaled version of the A-criteria in the standard OED problem.*

V. Combining the Optimal Control and Optimal Experimental Design Approaches

We now combine the formulations in Sections III and IV into an approach that trades off the two Paradigms 1 and 2 by means of an optimization problem that combines the OC cost in (6) and the associated loss of optimality in (19), i.e.,

$$\begin{aligned}&W_b\left(x_{0|0}, 0\,;\,\alpha\right) = \\ &\quad= \min_{u_{0:N-1}} \mathbb{E}\left[G_N(Z_N) + \sum_{i=0}^{N-1} G_i(Z_i,u_i)\,\middle|\, X_0 = x_{0|0}\right] \\ &\qquad+ \alpha\mathrm{tr}\left(\nabla_\theta H_0\left(\widehat{\theta}\right)\widetilde{M}(\widehat{\theta},u)^{-1}\nabla_\theta H_0\left(\widehat{\theta}\right)^\top\right) .\end{aligned} \tag{21}$$

with α acting as a weighting or trading-off between the two paradigms.

This formulation can then be laddered to formulate different exercise design algorithms. For instance, Algorithm 1 exemplifies how to use information from a certain physical exercise session to do a batch design of the next one.

Algorithm 1 Batch-designing a physical exercising session

1: Assume the input output data $\left(u^-_{0:N-1}, y^-_{1:N}\right)$ from a previous exercising session is available
2: Compute the maximum likelihood estimate $\widehat{\theta}^-$ of the model before starting the next session via (7), and the corresponding state estimates $\widehat{x}^-_{1:N} = \left\{\widehat{x}^-_{k|k}\right\}_{k=1,\ldots,N}$ via any filter of choice (e.g., an EKF)
3: Compute the corresponding sensitivities of the states $\widehat{s}^-_{1:N}$ via (12)
4: Compute the loss sensitivities $\nabla_\theta H_0\left(\widehat{\theta}^-\right)$ using $u^-_{0:N-1}$ in (17)
5: Solve the dual optimization problem in (21) given the state estimates $\widehat{x}^-_{1:N}$ and obtain the new exercising input profile $u^*_{0:N-1}$

The presented algorithm works for any general, possibly nonlinear, system of the form in (1)–(3); importantly, for systems where some exercise effect, rather than the actual states themselves, is of interest. We continue by applying the proposed algorithm, Algorithm 1, to a case study of interest where the exercise session system is an affine-input system.

Remark *In Algorithm 1 we propose to use an Extended Kalman Filter (EKF) for state estimation. For noise that is not i.i.d Gaussian noise and for highly nonlinear systems, one could use a sigma-point or particle filter.*

VI. A Case Study – Designing Kegel Exercises

We test the capabilities of our approach via simulations where we use the pelvic floor muscles model in [14] for modeling the dynamics of fatiguing in Kegel exercising. The proposed compartmental model is an extension of the work in [15] and involves active muscles m^{a}, fatigued muscles m^{f}, and resting muscles m^{r}, where the total number of muscular units is M. Note that the underlying true system is considered to be the more complex model derived in [16] where cramping muscles m^{c} are additionally used to capture a particular dynamic often observed when people perform Kegel exercises, which is a more realistic situation where the underlying system is more complex than the proposed model. We consider here, however, the original model in [14], where the discrete-time, time-invariant, control-affine system is given by

$$\begin{aligned} m^{\mathrm{f}}_{k+1} &= \phi^{\mathrm{fa}} m^{\mathrm{f}}_k + (1-\phi^{\mathrm{af}}) m^{\mathrm{a}}_k \\ m^{\mathrm{a}}_{k+1} &= \phi^{\mathrm{af}} m^{\mathrm{a}}_k + (1-\phi^{\mathrm{fa}}) m^{\mathrm{f}}_k + \phi^{\mathrm{ra}} m^{\mathrm{r}}_k u_k \\ &\quad - \phi^{\mathrm{ar}} m^{\mathrm{a}}_k (1-u_k) \\ m^{\mathrm{r}}_{k+1} &= M - m^{\mathrm{a}}_{k+1} - m^{\mathrm{f}}_{k+1} \\ y_k &= m^{\mathrm{a}}_k + w_k \, . \end{aligned} \tag{22}$$

Here the input u_k, called the *brain force*, is a normalized amount of effort (i.e., relative to some pre-defined maximum) exerted by the person performing the exercises at time k. Moreover, $w_k \sim \mathcal{N}(0, 0.25)$, ϕ^{fa} is the recovery rate from a fatigued to an active state, ϕ^{af} is the fatiguing rate, ϕ^{ar} is the rate of relaxation, and ϕ^{ra} is the activating rate from a rested state to an active state. We should note that the model in [14] is only validated for maximum effort or no effort, i.e., one or zero. However, Hill-type models are commonly used in biomechanics to model muscular contractions, where the normalized excitation input u can range between zero and one [15]. For notation, we will use $m_k = \begin{bmatrix} m^{\mathrm{f}}_k & m^{\mathrm{a}}_k \end{bmatrix}^\top \in \mathbb{R}^2$. To be able to use OED for parameter identification, we need to have an identifiable model. We know from [14] that $\phi = \begin{bmatrix} \phi^{\mathrm{fa}} & \phi^{\mathrm{af}} & \phi^{\mathrm{ar}} & \phi^{\mathrm{ra}} \end{bmatrix}^\top$ are identifiable parameters for known $M = 1$.

A. Defining a suitable exercising effect z

Our approach focuses on a generic exercising effect, a variable that is missing in the original model formulation. We thus augment (22) with the discounted accumulated output $z_k = \sum_{t=0}^{k} \gamma^k y_k$, a medically relevant extra state that intuitively corresponds to how much the person has been active during the physical exercising session (the discount $\gamma = 0.999$ having been added for stability reasons). We then consider the augmented system with $x_k = \begin{bmatrix} m_k & \bar{y}_k \end{bmatrix}^\top$ as

$$\begin{aligned} x_{k+1} &= \begin{bmatrix} \phi^{\mathrm{fa}} & (1-\phi^{\mathrm{af}}) & 0 \\ (1-\phi^{\mathrm{fa}}) & \phi^{\mathrm{af}} - \phi^{\mathrm{ar}} & 0 \\ 0 & T & (1-T\gamma) \end{bmatrix} x_k \\ &\quad + \begin{bmatrix} 0 \\ \phi^{\mathrm{ra}}(M - m^{\mathrm{f}}_k - m^{\mathrm{a}}_k) + \phi^{\mathrm{ar}} m^{\mathrm{a}}_k \\ 0 \end{bmatrix} u_k + \bar{w}_k \\ &= \mathcal{A}(\phi) x_k + \mathcal{B}(\phi, x_k) u_k + \bar{w}_k \, , \end{aligned} \tag{23}$$

where T is the length of the discretization period and $x_0 = \begin{bmatrix} 0 & 0 & 0 \end{bmatrix}^\top$. Moreover, the new measurements model is

$$y_k = m^{\mathrm{a}}_k = \begin{bmatrix} 0 & 1 & 0 \end{bmatrix} x_k + v_k = \mathcal{C} x_k + v_k \, . \tag{24}$$

The exercising effect is thus given by

$$z_k = \bar{y}_k = \begin{bmatrix} 0 & 0 & 1 \end{bmatrix} x_k = \mathcal{E} x_k \, , \tag{25}$$

while the state sensitivities may be computed via the discrete-time system

$$\begin{aligned} s_{k+1} &= \frac{\partial f(x_k, u_k, \phi)}{\partial x} s_k + \frac{\partial f(x_k, u_k, \phi)}{\partial \phi} \\ &= \mathcal{A}(\phi) s_k + \begin{bmatrix} m^{\mathrm{f}}_k & -m^{\mathrm{a}}_k & 0 & 0 \\ -m^{\mathrm{f}}_k & m^{\mathrm{a}}_k & b_{23} & b_{24} \\ 0 & 0 & 0 & 0 \end{bmatrix} \, , \end{aligned} \tag{26}$$

where $b_{23} = m^{\mathrm{a}}_k(1-u_k)$, $b_{24} = (M - m^{\mathrm{a}}_k - m^{\mathrm{f}}_k) u_k$.

B. Control Problem Setup

To maximize the exercising effect and penalize the control effort (in this case the overall physical demand required to the person), we consider the LQG cost

$$\begin{aligned} W(x_0, 0) &= \min_{u_{0:N-1}} \mathbb{E}\Big[- Z_N^\top G_N Z_N \\ &\quad + \sum_{i=0}^{N-1} -\varphi(X_i, u_i, \theta)^\top G_z \varphi(X_i, u_i, \theta) + u_i^\top G_u u_i \Big] \, , \end{aligned} \tag{27}$$

with $G_N \succ 0$, $G_z \succ 0$, $G_u \succ 0$, and where the fact that the expectation is conditioned on the estimated states is omitted for brevity. Moreover, the sensitivities of the loss (17) for the exercising effect in (25) are

$$\nabla_\theta H_0(\theta) = \sum_{i=0}^{N-1} -G_z (\mathcal{E} X_i)^\top \mathcal{E} S_i \tag{28}$$

and the FIM in (11) is

$$\begin{aligned} M(\theta, u) &= \sum_{i=0}^{N-1} \left(\frac{\partial h}{\partial X} \frac{\mathrm{d} X_i}{\mathrm{d}\theta} \right)^\top R^{-1} \left(\frac{\partial h}{\partial X} \frac{\mathrm{d} X_i}{\mathrm{d}\theta} \right) \\ &= \sum_{i=0}^{N-1} (\mathcal{C} S_i)^\top R^{-1} (\mathcal{C} S_i) \, . \end{aligned} \tag{29}$$

From the ingredients above one may then build the final scaled cost $W_b\left(x_{0|0}, 0 \, ; \, \alpha\right)$ as in (21).

C. Simulation Results

Our goal is to compare the results one obtains by varying α in $W_b\left(x_{0|0}, 0 \, ; \, \alpha\right)$, i.e., investigate which exploration vs. exploitation trade-offs may emerge in the design of personalized physical exercise sessions via the proposed methodology with $G_N = G_z = 10$ and $G_u = 1$ in (27).

We thus use a discretization period length of $T = 0.1$ seconds for a session length of $N = 1201$ (two minutes), and compare simulation results for the following situations over 10 runs:

1) performing a pure OC without any OED objective,
2) performing a standard OED with the A-criteria for the first half of the experiment to re-estimate the parameters, followed by performing OC for the last half of the experiment with the new parameter update,
3) performing the DC problem in (21) with $\alpha = 0.1, 1, 10$, to test different levels of tradeoffs.

Table I presents a selection of results from the simulations: the terminal exercise effect, the final expected loss of optimality, the total cost, the control effort, and the estimation error.

TABLE I

A comparison of the different control schemes for the achieved exercise effect, final expected loss of optimality, the total cost, the control effort, and the estimation error. The values are given as means and standard deviations over the ten runs.

Control type	final exercise effect	final loss of optimality	total cost	control effort	estimation error
1) OC (without noise)	19.57 (1.49)	3.34 (1.14)	−6.11 (0.218)	3.38 (0.769)	9.27 (3.12)
2) $\frac{1}{2}$ OED $\frac{1}{2}$ OC	20.08 (2.31)	7.21 (16.02)	−6.07 (0.400)	4.16 (0.695)	8.36 (3.25)
3.1) DC $\alpha = 0.1$	12.79 (2.26)	0.866 (0.321)	−1.44 (1.74)	2.10 (0.16)	3.26 (1.92)
3.2) DC $\alpha = 1$	9.31 (1.00)	0.338 (0.073)	1.34 (1.39)	1.85 (0.549)	1.40 (0.578)
3.3) DC $\alpha = 10$	8.11 (2.29)	0.219 (0.184)	29.08 (65.53)	2.72 (1.06)	0.766 (0.732)

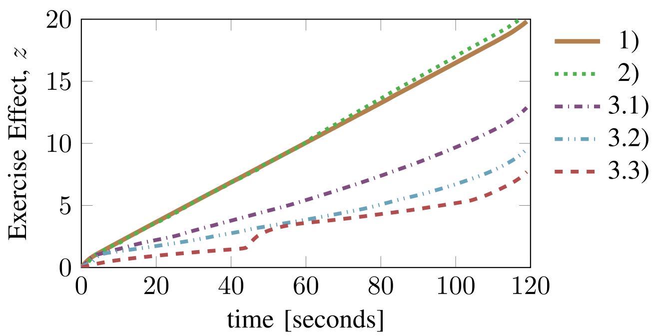

Fig. 1. The exercise effect z over the simulation period for the different control scenarios. Only one run was chosen from each control scenario for illustrative purposes.

Moreover, Figure 1 illustrates the difference in the exercise effect over the optimal exercising session for the different control scenarios. The OC and half OED / half OC scenarios have the best overall exercising effect but have high associated control efforts (i.e., demands to the person that is exercising) and expected loss of optimality. The different DC trade-offs act as expected, where the control effort is higher for a higher weight α on the informativity of the data, additionally yielding estimation error and a lower loss of optimality. The trade-off α acts also as an excitement variable since the OED aspect pushes the input to be more exciting leading to less predictable inputs and more engaging biofeedback for the user.

VII. Conclusion

We presented a dual control (simultaneous OC and OED) method for the design of physical exercising sessions suitable when the goal is to simultaneously estimate a personalized model of the exerciser while maximizing the medical effect the sessions are having on the person. The strategy considers that in medical settings one may define the overall exercising effect as a specific KPI. We thus provided the theoretical background for solving OC and OED problems on such a KPI, based on investigating the sensitivity of the loss in the OC problem as opposed to the sensitivity of the output. We have then been proposing to include a weight α in the problem formulation, that can be used to trade-off between these two factors as wished. We then applied the technique to a case study for Kegel exercises where we considered the unique objective of maximizing accumulated muscle activation, thereby addressing a previously unexplored aspect of exercise science.

The formulated approach considers, for now, batch offline designs. Using recursive parameter identification and a receding horizon control, one may easily make the strategy for computing the inputs recursively. The problem of adapting α in time, i.e., prioritizing OC versus OED more and more throughout the experiments, is a non-trivial extension. In other words, an adaptive α may intuitively prioritize the OED when the person starts their therapy, improve the quality of the parameter estimates, and later in the treatments prioritize maximizing their effect.

We also note that the current formulation has no constraints on unrealistic control inputs. Physical exercise is indeed constrained by biological factors (e.g., how fast one may run and recovery times). In other words, the approach does not account for what physiotherapists may consider optimal from a medical perspective. Another extension may, thus, enable users and physiotherapy constraints to influence how much the control can excite the system.

References

[1] O. M. Giggins, U. M. C. Persson, and B. Caulfield, “Biofeedback in rehabilitation,” *Journal of NeuroEngineering and Rehabilitation*, 2013.

[2] L. Kennedy and S. H. Parker, “Biofeedback as a stress management tool: a systematic review,” *Cognition, Technology and Work*, vol. 21, 2019.

[3] R. Sielski, W. Rief, and J. A. Glombiewski, “Efficacy of Biofeedback in Chronic back Pain: a Meta-Analysis,” *International Journal of Behavioral Medicine*, vol. 24, no. 1, 2017.

[4] J. Z. Goldenberg, M. Brignall, M. Hamilton, J. Beardsley, R. D. Batson, J. Hawrelak, B. Lichtenstein, and B. C. Johnston, “Biofeedback for treatment of irritable bowel syndrome,” *Cochrane Database of Systematic Reviews*, no. 11, 2019.

[5] K. Larsen Burgio, J. Robinson, and B. T. Engel, “The role of biofeedback in Kegel exercise training for stress urinary incontinence,” *American Journal of Obstetrics and Gynecology*, vol. 154, no. 1, 1986.

[6] S. Jiménez Morgan and J. A. Molina Mora, “Effect of Heart Rate Variability Biofeedback on Sport Performance, a Systematic Review,” *Applied Psychophysiology Biofeedback*, vol. 42, 2017.

[7] A. Facchinetti, S. Del Favero, G. Sparacino, J. R. Castle, W. K. Ward, and C. Cobelli, “Modeling the glucose sensor error.” *IEEE transactions on bio-medical engineering*, vol. 61, no. 3, pp. 620–629, 2014.

[8] S. Donnet and A. Samson, “A review on estimation of stochastic differential equations for pharmacokinetic/pharmacodynamic models.” *Advanced drug delivery reviews*, vol. 65, no. 7, pp. 929–939, 2013.

[9] K. J. Åström, *Introduction to Stochastic Control Theory*, K. J. Åström, Ed. New York and London: Academic Press, 1970, vol. 70.

[10] G. C. Goodwin and R. L. Payne, *Dynamic System Identification: Experiment Design and Data Analysis*, 1st ed. Academic Press, 1977.

[11] H. Chernoff, “Locally Optimal Designs for Estimating Parameters,” *The Annals of Mathematical Statistics*, vol. 24, no. 4, 1953.

[12] X. Feng and Y. Jiang, “A Scheme for Simultaneous Optimal Tracking Control and Experiment Design,” *IEEE Access*, vol. 8, 2020.

[13] H. C. La, A. Potschka, J. P. Schlöder, and H. G. Bock, “Dual Control and Online Optimal Experimental Design,” *SIAM Computational Methods in Science and Engineering*, vol. 39, no. 4, 2017.

[14] N. Kask, D. M. Budgett, J. Kruger, P. M. F. Nielsen, D. Varagnolo, and S. Knorn, “Data-driven modelling of fatigue in pelvic floor muscles when performing Kegel exercises,” in *IEEE CDC*, 2019.

[15] S. A. Ross, N. Nigam, and J. M. Wakeling, “A modelling approach for exploring muscle dynamics during cyclic contractions,” *PLOS Computational Biology*, vol. 14, no. 4, 2018.

[16] R. Doshmanziari, R. R. Jackson, D. Varagnolo, and S. Knorn, “Data-driven modeling of fatigue effects following repeated muscular contractions,” in *IEEE CDC*, 2022.