IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received January 6, 2020, accepted January 20, 2020, date of publication January 30, 2020, date of current version February 10, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2970593

A Scheme for Simultaneous Optimal Tracking Control and Experiment Design

XUHUI FENG AND YUNING JIANG
School of Information Science and Technology, ShanghaiTech University, Shanghai 201210, China
Shanghai Institute of Microsyst and Information Technology, Chinese Academy of Sciences, Shanghai 200050, China
University of Chinese Academy of Sciences, Beijing 100049, China

Corresponding author: Xuhui Feng (fengxh@shanghaitech.edu.cn)

This work was supported in part by the National Science Foundation China (NSFC) under Grant 61473185, and in part by ShanghaiTech University under Grant F-0203-14-012.

ABSTRACT This paper is about the formulation and numerical solutions of simultaneous optimal experiment design and optimal tracking control problems. Our motivating example is a robot arm that is mounted on a kitchen wall over the hotplates in order to assist humans with cooking. The robot arm can take over simple tasks such as stirring, automatic seasoning, or adding ingredients to a pot following a given recipe. Here, one of the main challenges is that the robot has to learn about the mass, inertial and other properties of the objects it is picking up while satisfying control tasks. Thus, we are facing a classical dual control problem, where system excitation for the purpose of learning has to be trade-off with other objectives such as tracking performance. After reviewing existing approaches, we propose a new formulation which allows to implement a trade-off between experiment design as well as tracking objectives in a systematic way by exploiting recent ideas from the field of economic experiment design. The approach is tested with a dynamic robot arm model performing a simple but illustrative cooking maneuver where learning and control goals are in conflict.

INDEX TERMS Experiment design, tracking control, dual control.

I. INTRODUCTION

Feldbaum's seminal work on dual control [8], as originally published in the early 1960s, had a major influence on control theory development during the past 50 years [35]. This is due to the fact that in many model based closed-loop systems the task of designing optimal control inputs based on given state and parameter estimates is in conflict with the task of optimizing the accuracy of future state and parameter estimates. Especially, when a nonlinear process is affected by external disturbance that have to be estimated online and that vary rapidly, it is desirable that the controller takes the accuracy of the estimator into account in order to improve the overall performance of the closed system [34].

Although one might argue that standard optimal control formulation and its traditional variants lack of the ability to deal with the aforementioned control task, many other controllers from the field of adaptive control are designed with the consideration of both learning and control targets. There are in general two categories of adaptive control: one class adopts dual control perspective and the other belongs nondual adaptive controllers. The learning behavior in the non-dual controllers is ''accidental'' or ''passive'' [35]. For those nondual adaptive controllers, iterative learning control (ILC) is widely applied in robotic applications [6], [32], especially for tracking aim. The main idea behind ILC is to iteratively find an input sequence such that the output of the system is as close as possible to a desired output [24], [26]. A complete survey about the recent development of ILC can be referred to [1], [4]. In contrast to nondual controllers, an intentional probing signal is generated for dual controller. Such controllers from the field of adaptive control have a longer history of investigating how to achieve a tradeoff between nominal performance as well as learning objectives. One challenge is that numerical algorithms for solving the dual control problem in higher dimensional spaces are often based on approximate dynamic programming, which turn out to be rather computational expensive [21], [38]. There is thus a great need for searching different approximations that can lead to tractable formulations without losing dual features. For an overview about other attempts to solve the dual control problem approximately by using techniques from the field

The associate editor coordinating the review of this manuscript and approving it for publication was Sing Kiong Nguang.

of adaptive control, the reader is referred to the overview articles [35].

The question of how to design control inputs in such a way that the information content of the associated experiment is maximized, is addressed rigorously in the field of optimal experiment design [7], [10], [27]. The connections between experiment design and the intended model application have been analyzed in the linear system community and we refer to [12], [22] for an overview. This work has later been extended towards a joint design of identification and control as for example discussed in [13] and [15]. Notice that optimal experiment design problems have been analyzed for nonlinear dynamic systems, too. Here, we refer to [20], where nonlinear dynamic systems with unknown parameters are covered as well as to [33], where the case that nonlinear systems are affected by both unknown but time-invariant parameters as well as time-varying process noise is studied.

Concerning the real-time adaption strategies for online optimization of dynamic processes in the presence of model-plant mismatches we refer to [5]. Other process optimization techniques are discussed in [29], where an interesting variant of optimal experiment design is proposed. Furthermore, extensions of model predictive control towards dual control are suggested in a variety of articles [11], [16], [23], [30].

Similar techniques can sometimes—based on a slightly different assumptions and notation but very similar intentions—also be found under the name "output model predictive control", which can be interpreted as a variant of model predictive control where estimation and control aspects are simultaneously taken into account [14], [28].

This paper is about simultaneous experiment design and optimal tracking control. The main contribution of the paper is twofold:

1) We propose a novel formulation that combines an economic optimal experiment design objective with a tracking control objective in a systematic way. The proposed simultaneous tracking and experiment design problem formulation is invariant under affine transformations of the parameters.
2) We verify the effectiveness of proposed formulation via a challenging case study. Here, our motivating example is a cooking robot arm that picks up ingredients or tools with unknown mass, rotational inertia, and other physical properties. One goal is to measure these parameters whenever a new object is picked up which can be in conflict with the given ultimate cooking tasks.

The corresponding problem formulation and robot arm model are outlined in Section II. Section III reviews existing techniques from the field of optimal experiment design. Section IV reviews the concept of least-squares tracking control. The main contribution of this paper is presented in Section V, where we propose a novel formulation that combines an economic optimal experiment design objective with a tracking control objective in a systematic way. This formulation is inspired by recent developments in the field of economic optimal experiment design [18]. However, a major contribution of this paper with respect to [18] is that we specialize the algorithm for computing derivatives in the context of least-squares tracking problems and also explain how to combine economic OED objectives with a nominal tracking performance objective. Section VI presents a case study of a cooking robot, which has to perform a dual control task. Section VII concludes the paper.

Notation: Besides mathematical standard notation, we denote with $\mathbb{S}_+^n$ the set of symmetric and positive semi-define matrices. Moreover, the syntax $M^\dagger$ denotes the pseudo inverse of a given matrix M, which can also be written in the form

$$M^\dagger = (M^\top M)^{-1} M^\top$$

assuming that the matrix $M^\top M$ is invertible.

Throughout this paper, the following assumptions are used.

Assumption 1: We assume that all functions are three times continuously differentiable.

Assumption 2: We assume the Jacobian matrix of the right-hand-side of dynamic functions, given by

$$\frac{\partial f(z,u,p)}{\partial u},$$

to be invertible for all feasible inputs u such that the Fisher information matrix $\mathcal{F}(u,p)$ is well-defined.

II. A DYNAMIC MODEL FOR COOKING ROBOTS

Our motivating example is a cooking robot arm that picks up ingredients or tools with unknown mass, rotational inertia, and other physical properties. Notice that there exist several suggestions for how to design advanced cooking robot systems, as for example proposed in [31] or [36], [37]. However, in this paper, we concentrate on a simplified setting consisting of a single robot arm whose movements are restricted to a 2-dimensional working space as sketched in Figure 1.

In our model of the robot arm, the horizontal and vertical position of the robot arm is denoted by x and y, respectively. Moreover, we assume that the robot has grippers at the end of the arm, which can rotate. The corresponding rotation angle is denoted by φ. We install three motors on the robot arm, which can be used to control the forces in x- and y-direction as well as the torque at the grippers. The kinetic energy E_k of the system is

$$E_k = \frac{1}{2}(M+m)(\dot{x}^2+\dot{y}^2)+\frac{1}{2}I\dot{\varphi}^2+m\bar{l}[\dot{x}\dot{\varphi}\cos(\varphi)+\dot{y}\dot{\varphi}\sin(\varphi)],$$

where M and m are the mass of robot and object, respectively, and I is the rotational inertia of the gripper with respect to the center of rotation, i.e., we have

$$I = m\bar{l}^2 + I_{\text{object}},$$

where I_{object} denotes the rotational inertia of the object with respect to its center of mass. Notice that the mass of the rotating part of the arm, including the grippers, is neglected.

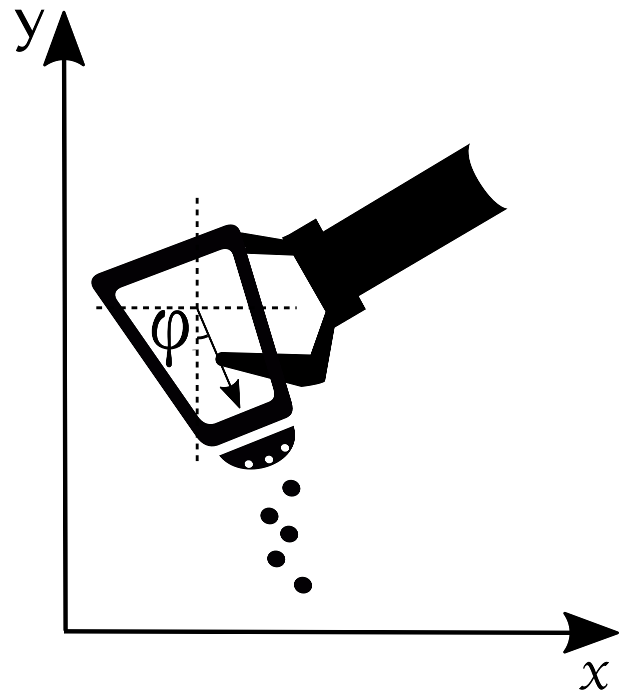


FIGURE 1. Sketch of the robot arm.

Here, $\bar{l}$ is the distance between the object's center of mass and the object's center of rotation. The potential energy E_p of the system is

$$E_p = Mgy + mgy - mg\bar{l}\cos(\varphi). \tag{1}$$

The Lagrangian function is given by

$$\begin{aligned} L &= E_k - E_p \\ &= \frac{1}{2}(M+m)(\dot{x}^2 + \dot{y}^2 - 2gy) + \frac{1}{2}I\dot{\varphi}^2 \\ &\quad + m\bar{l}\dot{\varphi}[\dot{x}\cos(\varphi) + \dot{y}\sin(\varphi)] + mg\bar{l}\cos(\varphi). \end{aligned} \tag{2}$$

In order to derive the equations of motion for the robot arm, we use the Lagrangian formalism, which yields

$$\begin{cases} \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}}\right) - \frac{\partial L}{\partial x} = u_x - \gamma_x \dot{x} \\ \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{y}}\right) - \frac{\partial L}{\partial y} = u_y - \gamma_y \dot{y} \\ \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{\varphi}}\right) - \frac{\partial L}{\partial \varphi} = u_\varphi - \gamma_\varphi \dot{\varphi}, \end{cases} \tag{3}$$

where u_x and u_y are the horizontal and vertical motor forces, and where u_ϕ is the torque of the third motor, while γ_x, γ_y, and γ_φ are the friction coefficients. Therefore, substituting equation (2) into equation (3) yields

$$\begin{cases} (M+m)\ddot{x} + m\bar{l}[\ddot{\varphi}\cos(\varphi) - \dot{\varphi}^2\sin(\varphi)] = u_x - \gamma_x\dot{x} \\ (M+m)(\ddot{y}+g) + m\bar{l}[\ddot{\varphi}\sin(\varphi) + \dot{\varphi}^2\cos(\varphi)] = u_y - \gamma_y\dot{y} \\ I\ddot{\varphi} + m\bar{l}[\ddot{x}\cos(\varphi) + \ddot{y}\sin(\varphi) + g\sin(\varphi)] = u_\varphi - \gamma_\varphi\dot{\varphi}. \end{cases}$$

In the following, we stack all differential states into one vector $z = (x, y, \varphi, \dot{x}, \dot{y}, \dot{z})^T$ and write the above differential equation in the form of a standard ordinary differential equation of the form

$$\dot{z} = f(z, u, p),$$

where the vector

$$p = \left(m, \bar{l}, I_{\text{object}}\right)^T$$

contains the parameters, which are associated with the object that is currently carried by the robot. Thus, this parameter vector is not known in advance but has to be determined from measurements whenever a new object is picked up. Moreover, the right-hand side function f can be obtained by the above linear equation system with respect to the variables $\ddot{x}$, $\ddot{y}$, and $\ddot{\varphi}$. This elimination is always possible as the mass matrix

$$\begin{pmatrix} M+m & 0 & m\bar{l}\cos(\varphi) \\ 0 & M+m & m\bar{l}\sin(\varphi) \\ m\bar{l}\cos(\varphi) & m\bar{l}\sin(\varphi) & I \end{pmatrix}$$

is invertible, as we have $M, m > 0$ as well as $I > m\bar{l}^2$ for physical reasons. Notice that in our notation, the control inputs $u = (u_x, u_y, u_\varphi)^T$ from the three motors are stacked into one vector, too.

III. OPTIMAL EXPERIMENT DESIGN

Whenever the robot fetches a new object, we have to measure the parameters p associated with the object in order to carry it to a given target location. Following the traditional two-step approach of first solving the identification and then solving the control problem, we have to make a few experiments first. This means that we send control inputs to the robot arm and take measurements of the mechanical response using cameras. If $H : \mathbb{R}^{n_x} \to \mathbb{R}^{n_H}$ denotes the measurement function and η the measurements for a given control input $u : \mathbb{R} \to \mathbb{R}^{n_u}$, the parameters $p \in \mathbb{R}^{n_p}$ can be estimated by solving a Gaussian maximum likelihood estimation (MLE) problem,

$$\begin{aligned} &\min_{z,p} \sum_{i=1}^{N} \|H(z(t_i)) - \eta_i\|^2_{\Sigma^{-1}} \\ &\text{s.t.} \begin{cases} \dot{z}(t) = f(z(t), u(t), p), \ t \in [0, T] \\ z(0) = z_0(p). \end{cases} \end{aligned} \tag{4}$$

Here, $t_1, \ldots, t_N$ denote the time points at which measurements are taken and $\Sigma \in \mathbb{S}^{n_H}_{++}$ the variance-covariance matrix of the measurement error, which is assumed to be positive definite. In general, the initial value z_0 can be a function of the parameter p, too.

In order to simplify our syntax, we eliminate the differential states. We use the notation $\xi[t, u, p]$ to denote the solution of the differential equation system

$$\begin{aligned} \dot{z}(t) &= f(z(t), u(t), p), \ t \in [0, T] \\ z(0) &= z_0(p) \end{aligned} \tag{5}$$

for the state at time t. This syntax highlights the dependence of the state z on the control input u and the unknown parameter p. Next, we introduce the shorthand

$$h(u,p) := \left(H(\xi[t_1,u,p])^T, \ \ldots, \ H(\xi[t_N,u,p])^T \right)^T$$

such that the parameter estimation problem (4) can be written in the equivalent but unconstrained form

$$\min_p \ \|h(u,p) - \eta\|^2_{\Sigma^{-1}} .$$

where η is one large vector containing a stacked version of all measurements.

Now, the information content of the experiment can be quantified by computing the so-called Fisher information matrix [9], which we denote by

$$\mathcal{F}(u,p) = \nabla_p h(u,p) \Sigma^{-1} \nabla_p h(u,p)^T ,$$

where η is one large vector containing a stacked version of all measurements. In our context, the Fisher information matrix $\mathcal{F}(u,p)$ is regarded as a functional of the control input u.

It is well-known [22] that the inverse of the Fisher information matrix approximates the variance-covariance matrix which is used as a weighting in the cost function of the maximum likelihood estimate. However, $\mathcal{F}(u,p)$ is in general not invertible for all choices of u. For example, for our cooking robot model, we will only be able to measure the rotational inertia of the object by observing the states of the system, if we actually rotate the object. Otherwise, for example, if we only do up-and-down movements, we will not be able to collect information about this parameter, which leads to zero-entries in the associated rows and columns of the Fisher information matrix. This example illustrates that it is crucial to choose the control input u wisely in order to get maximum information about parameters.

In optimal experiment design, a suitable u is found by solving an optimization problem of the form

$$\min_u \ \Phi(\mathcal{F}(u,\hat{p})^{-1}) \quad \text{s.t.} \quad \underline{u} \le u \le \overline{u}. \tag{6}$$

Here, $\hat{p}$ denotes the current estimate of the parameter at which the Fisher information matrix is computed. The objective function $\Phi : \mathbb{S}^{n_p}_+ \to \mathbb{R}$ is called the optimal experiment design criterion, which measures the size of the variance matrix of the parameters estimate. Examples for famous criteria are the A-, E-, and D-criterion, where $\Phi(\mathcal{F}(u,\hat{p})^{-1})$ corresponds to the trace, maximum eigenvalue, or determinant of the matrix $\mathcal{F}(u,\hat{p})^{-1}$. In this context, we make use of an extended value definition, i.e., we set $\Phi(\mathcal{F}(u,\hat{p})^{-1}) = \infty$ whenever the matrix $\mathcal{F}(u,\hat{p})$ is not invertible. The vectors $\underline{u}$ and $\overline{u}$ are introduced in order to model bound constraints on the input u. Notice that optimal experiment design problem of this form are known in the literature for a long time [27] and many numerical algorithms for solving this type of problem have been suggested [20], [33]. The classical optimal experiment design scheme runs the following loop:

- S1 Solve problem (6) and perform an experiment at the optimizer $u^*(\hat{p})$.
- S2 Update a new parameter estimates $\hat{p}^+$ by solving the MLE (4)
- S3 If $\Phi(\mathcal{F}(u.\hat{p})^{-1}) \le \epsilon$, then stop; otherwise, set $\hat{p} \leftarrow \hat{p}^+$ and repeat.

However, drawbacks of the traditional optimal experiment design criteria are that

1) the A- and E-criterion as well as many other exiting criteria [27] are not invariant under affine transformations of the parameter, i.e., the problem formulation is highly scaling dependent,
2) there is generally not much advice on how to choose Φ, but this choice can itself be a modelling problem,
3) the only goal of the problem formulation (6) is to measure parameters as accurately as possible, which may not be the only objective in practice.

In the following section we will propose a novel optimal experiment design problem formulation strategy in the context of control problems, which aims at fixing these drawbacks.

IV. OPTIMAL TRACKING CONTROL

In this section we assume that our ultimate goal is to solve an optimal control problem of the form

$$\min_{z,u} \int_0^T \frac{1}{2} \left\{ \|z(t) - z_{\mathrm{ref}}(t)\|_Q^2 + \|u(t)\|_R^2 \right\} \mathrm{d}t$$

$$\text{s.t.} \begin{cases} \dot{z}(t) = f(z(t), u(t), p), \ t \in [0,T] \\ z(0) = z_0(p), \end{cases} \tag{7}$$

where z_{ref} denotes a given reference trajectory that we want to follow (= task). The second term $\|u(t)\|_R^2$ minimizes the control effort. The weighting matrices Q and R are assumed to be given and positive semidefinite. Notice that if we would know the exact parameter p, we could solve this problem numerically by using direct methods [2], [3]. In order to briefly discuss this, we introduce the control parameterization

$$u(t) \approx \sum_{i=0}^{N-1} v_i \xi_i(t),$$

where $\xi_0, \ldots, \xi_{N-1} \in L^2[0,T]$ are given orthogonal functions while $v \in \mathbb{R}^{Nn_u}$ denotes the control coefficient vector. This notation includes the important special case that we use piecewise constant control approximations, which can be made arbitrarily accurate by choosing sufficiently large N. Next,

$$\zeta[t,v,p] = \xi \left[t, \sum_{i=0}^{N-1} v_i \xi_i, p \right]$$

denotes the solution of the differential equation recalling that ξ has already been defined in the previous section. By using this notation, we can approximate the tracking objective by an expression of the form

$$\int_0^T \frac{1}{2} \left\{ \|\zeta[t,v,p] - z_{\mathrm{ref}}(t)\|_Q^2 + \|u(t)\|_R^2 \right\} \mathrm{d}t \approx \frac{1}{2} \|g(v,p)\|_2^2 .$$

Here, we have introduced the shorthand $g = [g_1, \ldots, g_N]^T$ with

$$g_i(v,p) = \frac{1}{h}\begin{pmatrix} Q^{\frac{1}{2}}\left(\zeta[t_i, v, p] - z_{\mathrm{ref}}(t_i)\right) \\ R^{\frac{1}{2}}\sum_{i=0}^{N-1} v_i \xi_i(t_{i-1}) \end{pmatrix}$$

as well as the equidistant time mesh $t_i = ih$, $i \in \{1, \ldots, N\}$, with $h = \frac{T}{N}$. Notice that this reformulation has the advantage that the original continuous-time optimal control problem can be approximated by an unconstrained and finite dimensional nonlinear least-squares problem, which can be written as

$$\min_{v} \ \frac{1}{2}\,\|g(v,p)\|_2^2 \,. \tag{8}$$

For the case that the parameter p is known, finite dimensional optimization problems of this form can be solved with standard nonlinear programming solvers [25]. However, the main problem that we are facing in this paper is that we do not know p. In addition, in our robot example, measuring the unknown parameters whenever a new object is picked up can be in conflict with the given control tasks. Instead of adopting traditional identification-control scheme, we want to design a controller which can strike a balance between tracking control and identification. A simultaneous experiment design and optimal tracking control scheme will be discussed further in the section below.

V. SIMULTANEOUS EXPERIMENT DESIGN AND OPTIMAL TRACKING CONTROL

In this section, we are interested in economic optimal experiment design problems [18] for tracking control. Recall that our ultimate goal is to solve optimization problem (8), which can be interpreted as a discrete-time approximation of the original tracking optimal control problem (7). In the following, we denote with $v^*(p)$ a minimizer of problem (8) in dependence of the parameter p. Clearly, if we would know the exact parameter, denoted by p^*, we would control the system by implementing the optimal control input $v^*(p^*)$. However, if we do not know p^*, we can solve problem (8) based on an estimate p of the parameters only. This leads to a loss of optimality, which is given by

$$\Delta(p) = \frac{1}{2}\left\|g(v^*(p), p^*)\right\|_2^2 - \frac{1}{2}\left\|g(v^*(p^*), p^*)\right\|_2^2 \geq 0.$$

Notice that the loss of optimality $\Delta(p)$ is equal to zero for $p = p^*$ but in general positive if we have $p \neq p^*$. Let us introduce the shorthands

$$P(p) = \frac{\partial}{\partial p} g(v^*(p), p) \quad \text{and} \quad V(p) = \frac{\partial}{\partial v} g(v^*(p), p)$$

as well as

$$W(p) = P(p)^{\mathsf{T}} V(p) V(p)^{\dagger} P(p).$$

Here, the derivatives that are needed for computing $V(p)$ can be computed by using automatic differentiation as implemented in ACADO Toolkit [17]. The following result is based on this notation.

Theorem 1: Let the function g be three times continuously differentiable. If we can achieve perfect tracking with exact parameters, i.e., if $g(v^*(p^*), p^*) = 0$, and if the matrix $V(p^*)$ has full rank, then the second order Taylor expansion of the loss of optimality is given by

$$\Delta(p) = \frac{1}{2}(p - p^*)^{\mathsf{T}} W(p^*)(p - p^*) + \mathbf{O}(\|p - p^*\|_2^3).$$

Proof: The first order optimality condition for the original least-squares minimization problem (8) can be written in the form

$$g(v^*(p), p)^{\mathsf{T}} \frac{\partial g(v^*(p), p)}{\partial v} = 0 \tag{9}$$

for all p, which implies

$$\frac{\partial}{\partial p}\Delta(p^*) = g(v^*(p^*), p^*)^{\mathsf{T}} \frac{\partial g(v^*(p^*), p^*)}{\partial v}\frac{\partial v^*(p^*)}{\partial p} = 0.$$

We apply the implicit function theorem to the stationarity equation (9), which yields

$$\frac{\partial v^*(p^*)}{\partial p} = -V(p^*)^{\dagger} P(p^*)$$

recalling that the matrix $V(p^*)$ has full-rank. Next, we can work out the second order derivative of the function Δ finding

$$\begin{aligned}\frac{\partial^2}{\partial p^2}\Delta(p^*) &= \frac{\partial v^*(p^*)}{\partial p}^{\mathsf{T}} V(p)^{\mathsf{T}} V(p) \frac{\partial v^*(p^*)}{\partial p} \\ &= P(p^*) V(p^*) V(p^*)^{\dagger} P(p^*) \ = \ W(p^*). \end{aligned} \tag{10}$$

The statement of the theorem is a consequence of Taylor's theorem. ◇

An important consequence of the above theorem is that the expected loss of optimality can be approximated by an expression of the form

$$\begin{aligned}\mathbb{E}_p\left\{\Delta(p)\right\} &\approx \mathbb{E}_p\left\{\frac{1}{2}(p - p^*)^{\mathsf{T}} W(p^*)(p - p^*)\right\} \\ &= \frac{1}{2}\mathrm{Tr}\left(\mathbb{E}_p\left\{(p - p^*)(p - p^*)^{\mathsf{T}}\right\} W(p^*)\right) \\ &\approx \frac{1}{2}\mathrm{Tr}\left(\mathcal{F}\left(\sum_i v_i \xi_i, \hat{p}\right)^{-1} W(\hat{p})\right). \end{aligned}$$

Next, we propose to minimize the sum of the nominal least-squares tracking term and its associated loss of optimality, which leads us to the optimization problem

$$\min_{v} \ \frac{1}{2}\left\|g(v, \hat{p})\right\|_2^2 + \frac{1}{2}\mathrm{Tr}\left(\mathcal{F}\left(\sum_{i=0}^{N-1} v_i \xi_i, \hat{p}\right)^{-1} W(\hat{p})\right). \tag{11}$$

This optimization problem can be solved to local optimality with standard nonlinear programming solvers. In this paper, we use SQP methods [25]. Notice that in contrast to the standard optimal experiment design framework from

Section III, the optimization problem takes both the nominal tracking performance

$$\frac{1}{2}\|g(v,p)\|_2^2$$

as well as an experiment design objective into account.

Remark 1: In order to differentiate our proposed formulation (11) with existing method, we would first introduce the formulation

$$\min_v \frac{1}{2}\|g(v,\hat{p})\|_2^2 + \frac{1}{2}\mathrm{Tr}\left(\alpha\cdot\mathcal{F}\left(\sum_{i=0}^{N-1} v_i\xi_i,\hat{p}\right)^{-1}\right),$$

where $\alpha \in \mathbb{S}_+^{n_x}$ is a weighting matrix. In general, the method that is widely adopted in practice attempts to tune the weighting matrix α in order to achieve a good tradeoff between control and learning performance. In contrast to different tuning methods, our proposed scheme provides a systematic way to construct this weighting matrix α tailored for tracking control. Specifically, the term

$$\frac{1}{2}\mathrm{Tr}\left(\mathcal{F}\left(\sum_i v_i\xi_i,\hat{p}\right)^{-1} W(\hat{p})\right)$$

can be interpreted as a weighted A-criterion. However, this weighted A-criterion has a meaningful physical interpretation, as it corresponds to the expected loss of optimality. Moreover, in contrast to many existing optimal experiment design criteria [27], the proposed simultaneous tracking and experiment design problem formulation (11) is invariant under affine transformations of the parameters.

Under the assumption of unbiased and uncorrelated Gaussian noise, the inverse of (approximate) Fisher information matrix approximates the parameter estimation variance-covariance matrix, i.e.,

$$\mathcal{F}\left(\sum_{i=0}^{N-1} v_i\xi_i,\hat{p}\right)^{-1} \approx \mathcal{S}(T,v,\hat{p}).$$

Here, $\mathcal{S}(t,v,\hat{p})$ denotes the solution of forward propagating Riccati equation of the following form at time t

$$\begin{aligned}\dot{S}(t) &= A(t,\hat{p})S(t) + S(t)A(t,\hat{p})^\mathsf{T}\\ &\quad -S(t)C(t)^\mathsf{T}\Sigma^{-1}C(t)S(t),\\ S(0) &= \hat{S},\end{aligned} \tag{12}$$

where $S(t)$ denotes the variance-covariance matrix of uncertain parameters at time $t \in [0,T]$ and $\hat{S}$ the initial guess of variance-covariance matrix. Moreover, the notation $A(t,\hat{p})$ and $C(t)$ are given by

$$A(t,\hat{p}) = \frac{f(z(t),u(t),\hat{p})}{z(t)} \quad\text{and}\quad C(t) = \frac{H(z(t))}{z(t)}, \tag{13}$$

respectively. Therefore, the optimization problem (11) is equivalent to solving the following dual control problem of the form

$$\min_v \int_0^T \frac{1}{2}\|g(v,\hat{p})\|_2^2 + \frac{1}{2}\mathrm{Tr}\left(W(\hat{p})\mathcal{S}(T,v,\hat{p})\right). \tag{14}$$

TABLE 1. Parameter reference.

Name	Symbol	value
mass of robot	M	2(kg)
gravitational acceleration	g	9.81 ($\frac{\mathrm{m}}{\mathrm{s}^2}$)
coefficient of friction	γ_x, γ_y	0.1 ($\frac{\mathrm{Ns}}{\mathrm{m}}$)
coefficient of friction	γ_ϕ	0.1 ($\frac{\mathrm{Nms}}{\mathrm{rad}}$)
estimate of the object's mass	$\hat{m}$	1 (kg)
estimate of the object's moment of inertia	$\hat{I}_{\mathrm{object}}$	0.02 (kg · m^2)
estimate of the distance	$\hat{\bar{l}}$	0.1 (m)
standard deviation of x-measurement	σ_x	0.02 (m)
standard deviation of y-measurement	σ_y	0.02 (m)
standard deviation of φ-measurement	σ_x	0.1 (rad)

Now, we run the following loop:

S1 Compute $W(\hat{p})$ by using Equation (10).

S2 Solve problem (14) and perform an experiment at the optimizer $u^*(\hat{p})$.

S3 Update a new parameter estimates $\hat{p}^+$ by solving the optimization problem (4)

S4 If $\frac{1}{2}\mathrm{Tr}\left(W(\hat{p})\mathcal{S}(T,v,\hat{p})\right) \le \epsilon$, then stop; otherwise, set $\hat{p} \leftarrow \hat{p}^+$ and repeat.

Finally, we mention that problem (11) cannot only be used to design optimal input profiles offline, but it could also be employed as an economic objective of a model predictive controller (MPC) [28], which would solve optimization problems of the form (11) at every time step based on the current parameter estimate in order to not only predict tracking performance but also improve the quality of future parameter estimates. However, a discussion of such dual MPC control schemes is beyond the scope of this paper and left for future investigations.

Remark 2: With the aim of guaranteeing the safety in some critical cases, we suggest directly including the safety constraints into the optimization problem. However, a direct inclusion of such constraints is at the expense of performance loss of proposed dual controller, since the controller cannot excite the system as much as it can do to achieve the best learning effect.

VI. SIMULTANEOUS EXPERIMENT DESIGN AND CONTROL OF COOKING ROBOTS

In this section, we solve Problem (11) for our cooking robot model that has been introduced in Section II. In case of the occure of large deviation, the safety constraint $x \ge -0.2$(m) is introduced. All model parameters are listed in Table 1. Notice that in our case study, we assume that the parameter vector

$$p = \left(m, \bar{l}, I_{\mathrm{object}}\right)^T$$

is unknown. Initial estimates $\hat{m}$, $\hat{\bar{l}}$, and $\hat{I}_{\mathrm{object}}$ for the object's mass, distance between rotation center and center of mass, as well as the object's moment of inertia are given in Table 1. The true parameter values are in our case study equal to these

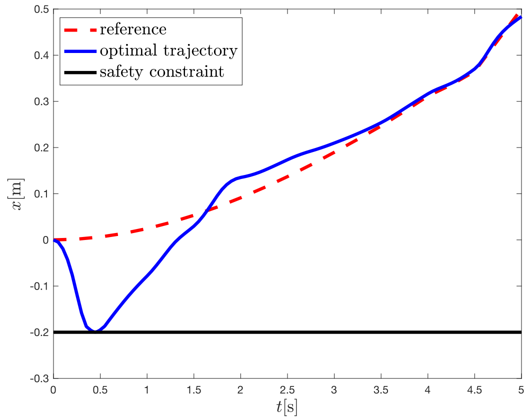


FIGURE 2. Numerical result for the reference (dashed) and optimal trajectory (solid) for the state $x(t)$.

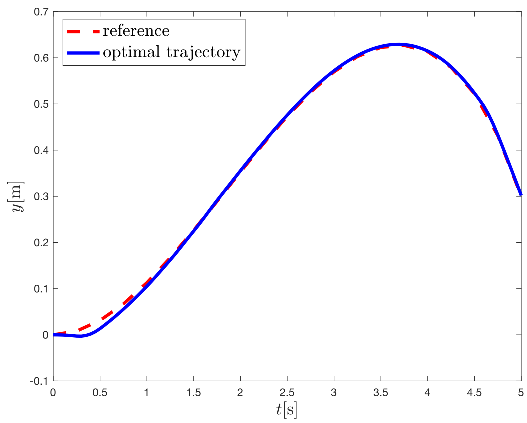


FIGURE 3. Numerical result for the reference (dashed) and optimal trajectory (solid) for the state $y(t)$.

estimates, but the controller does not know this. Moreover, we assume that the position x, y, as well as the angle φ can be measured. The corresponding variance matrix of the sensors has the form

$$\Sigma = \begin{pmatrix} \sigma_x^2 & 0 & 0 \\ 0 & \sigma_y^2 & 0 \\ 0 & 0 & \sigma_\varphi^2 \end{pmatrix},$$

where the standard deviations σ_x, σ_y, and σ_φ are given in Table 1, too. Recall that the function $g(v, p)$ is obtained by discretizing the least-squares objective function in the optimal control problem (7). The corresponding weighting matrices Q and R in the optimal control problem are in our case study given by $R = 0 \in \mathbb{S}_+^3$ as well as

$$Q = \mathrm{diag}\,(1, 1, 1, 0, 0, 0) \in \mathbb{S}_+^6.$$

The reference trajectory z_{ref} is pre-computed and shown in the form of the dashed lines in Figures 2, 3, and 4, respectively. We are using ACADO Toolkit [17] in combination with a piecewise constant control discretization using 20 pieces in order to solve Problem (11), numerically.

The solid lines in figures 2, 3, and 4 show the optimal solution of Problem (11) for x, y, and φ, respectively. Notice

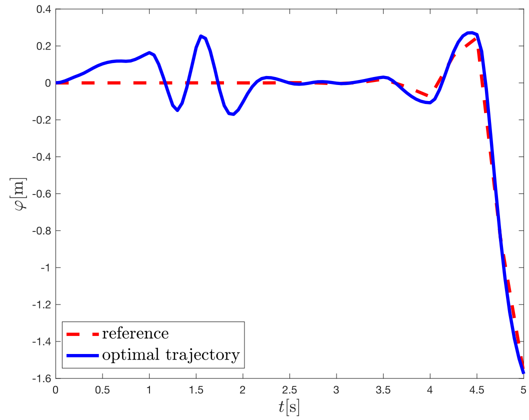


FIGURE 4. Numerical result for the reference (dashed) and optimal trajectory (solid) for the state $\varphi(t)$.

that it is optimal to deviate significantly from the reference trajectory. For example, our reference for the angle φ is close to 0 during the first three seconds. If we would follow this reference trajectory closely, we could gather almost no information about the parameter $\bar{l}$ and the rotational inertia I, as we have to actually rotate the object for identifying its rotational properties. Clearly, it is optimal to deviate from the reference and rotate (''shake'') the object carefully during the first seconds in order to be able to learn about its physical properties. A similar effect can be seen when studying the optimal trajectory for the x-coordinate: during the first seconds we deviate significantly from the reference in order to be able to identify the object's mass. Notice that such a behavior is typical for dual controllers, which apply excitations from the reference trajectory in order to learn.

VII. CONCLUSION

In this paper, we have presented a method for simultaneous optimal experiment design and optimal tracking control. After reviewing the standard framework for both optimal experiment design as well as optimal tracking control, we have proposed a novel way of formulating a dual control objective. Here, the goal is to minimize the sum of the nominal least-squares tracking performance and the expected loss of optimality that has to be taken into account due to unknown parameters that have to be learned from measurements. A crucial result for ensuring computational tractability has been presented in Theorem 1, which allows us to compute a second order expansion of the expected loss of optimality. Finally, we have analyzed a case study of a cooking robot, which is carrying objects with unknown mass and geometric properties. We have illustrated how the proposed formulation can be used to compute optimal control excitations that allow the robot to simultaneously identify parameters without deviating too much from the given tracking reference trajectory. Future work will investigate how the proposed formulation can be extended for model predictive control schemes, which re-optimize the control input based on the current parameter estimate and incoming measurements.

REFERENCES

[1] H.-S. Ahn, Y. Chen, and K. L. Moore, "Iterative learning control: Brief survey and categorization," *IEEE Trans. Syst., Man, Cybern. C, Appl. Rev.*, vol. 37, no. 6, pp. 1099–1121, Nov. 2007.

[2] L. T. Biegler and J. B. Rawlings, "Optimization approaches to nonlinear model predictive control," in *Proc. 4th Int. Conf. Chem. Process Control (CPC IV)*, W. H. Ray and Y. Arkun, Eds., 1991, pp. 543–571.

[3] H. G. Bock and K. J. Plitt, "A multiple shooting algorithm for direct solution of optimal control problems," in *Proc. 9th IFAC World Congr. Budapest*. New York, NY, USA: Pergamon, 1984, pp. 242–247.

[4] D. A. Bristow, M. Tharayil, and A. G. Alleyne, "A survey of iterative learning control," *IEEE Control Syst. Mag.*, vol. 26, no. 3, pp. 96–114, May 2006.

[5] B. Chachuat, B. Srinivasan, and D. Bonvin, "Adaptation strategies for real-time optimization," *Comput. Chem. Eng.*, vol. 33, no. 10, pp. 1557–1567, Oct. 2009.

[6] J. Y. Choi and J. S. Lee, "Adaptive iterative learning control of uncertain robotic systems," *IEE Proc.-Control Theory Appl.*, vol. 147, no. 2, pp. 217–223, 2000.

[7] D. Espie and S. Macchietto, "The optimal design of dynamic experiments," *AIChE J.*, vol. 35, no. 2, pp. 223–229, Feb. 1989.

[8] A. A. Feldbaum, "Dual-control theory," *Automat. Remote Control*, vol. 21, no. 9, pp. 874–880, 1961.

[9] R. A. Fisher, "Design of experiments," *Brit. Med. J.*, vol. 1, no. 3923, p. 554, 1936.

[10] G. Franceschini and S. Macchietto, "Model-based design of experiments for parameter precision: State of the art," *Chem. Eng. Sci.*, vol. 63, no. 19, pp. 4846–4872, Oct. 2008.

[11] H. Genceli and M. Nikolaou, "New approach to constrained predictive control with simultaneous model identification," *AIChE J.*, vol. 42, no. 10, pp. 2857–2868, Oct. 1996.

[12] M. Gevers and L. Ljung, "Optimal experiment designs with respect to the intended model application," *Automatica*, vol. 22, no. 5, pp. 543–554, Sep. 1986.

[13] M. Gevers, "Towards a joint design of identification and control?" *Progr. Syst. Control Theory*, vol. 14, pp. 111–151, 1993.

[14] P. J. Goulart and E. C. Kerrigan, "Output feedback receding horizon control of constrained systems," *Int. J. Control*, vol. 80, no. 1, pp. 8–20, 2007.

[15] H. Hjalmarsson, "From experiment design to closed-loop control," *Automatica*, vol. 41, no. 3, pp. 393–438, Mar. 2005.

[16] T. A. N. Heirung, B. Foss, and B. E. Ydstie, "MPC-based dual control with online experiment design," *J. Process Control*, vol. 32, pp. 64–76, Aug. 2015.

[17] B. Houska, H. J. Ferreau, and M. Diehl, "ACADO toolkit—An open-source framework for automatic control and dynamic optimization," *Optim. Control Appl. Methods*, vol. 32, no. 3, pp. 298–312, 2011.

[18] B. Houska, D. Telen, F. Logist, M. Diehl, and J. F. Van Impe, "An economic objective for the optimal experiment design of nonlinear dynamic processes," *Automatica*, vol. 51, pp. 98–103, Jan. 2015.

[19] J. Kiefer and J. Wolfowitz, "Optimum designs in regression problems," *Ann. Math. Statist.*, vol. 30, pp. 271–294, Jun. 1959.

[20] S. Körkel, E. Kostina, H. G. Bock, and J. P. Schlöder, "Numerical methods for optimal control problems in design of robust optimal experiments for nonlinear dynamic processes," *Optim. Methods Softw.*, vol. 19, nos. 3–4, pp. 327–338, Jun. 2004.

[21] J. M. Lee and J. H. Lee, "An approximate dynamic programming based approach to dual adaptive control," *J. Process Control*, vol. 19, no. 5, pp. 859–864, May 2009.

[22] L. Ljung, *System Identification: Theory for the User*. Upper Saddle River, NJ, USA: Prentice-Hall, 1999.

[23] G. Marafioti, "Enhanced model predictive control: Dual control approach and state estimation issues," Ph.D. dissertation, Norwegian Univ. Sci. Technol., Trondheim, Norway, 2010.

[24] O. Markusson, "Model and system inversion with applications in nonlinear system identification and control," Ph.D. dissertation, Signaler Sensorer Och System, 2001.

[25] J. Nocedal and S. Wright, *Numerical Optimization*. New York, NY, USA: Springer-Verlag, 2000.

[26] S. Preitl, R. E. Precup, Z. Preitl, S. Vaivoda, S. Kilyeni, and J. K. Tar, "Iterative feedback and learning control servo systems applications," *IFAC Proc.* vol. 40, no. 8, pp. 16–27, 2007.

[27] F. Pukelsheim, *Optimal Design of Experiments*. New York, NY, USA: Wiley, 1993.

[28] J. B. Rawlings and D. Q. Mayne, *Model Predictive Control: Theory and Design*. San Francisco, CA, USA: Nob Hill, 2009.

[29] S. Recker, N. Kerimoglu, A. Harwardt, O. Tkacheva, and W. Marquardt, "On the integration of model identification and process optimization," in *Proc. ESCAPE*, vol. 23, 2013, pp. 1021–1026.

[30] M. S. Shouche, H. Genceli, and M. Nikolaou, "Effect of on-line optimization techniques on model predictive control and identification (MPCI)," *Comput. Chem. Eng.*, vol. 26, no. 9, pp. 1241–1252, 2002.

[31] Y. Sugiura, D. Sakamoto, A. Withana, M. Inami, and T. Igarashi, "Cooking with robots: Designing a household system working in open environments," in *Proc. SIGCHI Conf. Hum. Factors Comput. Syst.*, 2010, pp. 2427–2430.

[32] A. Tayebi, "Adaptive iterative learning control for robot manipulators," *Automatica*, vol. 40, no. 7, pp. 203–1195, 2004.

[33] D. Telen, B. Houska, F. Logist, E. Van Derlinden, M. Diehl, and J. Van Impe, "Optimal experiment design under process noise using Riccati differential equations," *J. Process Control*, vol. 23, no. 4, pp. 613–629, 2013.

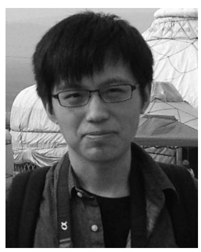
[34] B. Wittenmark and C. Elevitch, "An adaptive control algorithm with dual features," in *Proc. IFAC Symp. Identificat. Syst. Parameter Estimation*, 1985, pp. 587–592.

[35] B. Wittermark, "Adaptive dual control methods: An overview," in *Proc. 5th IFAC Symp. Adapt. Syst. Control Signal Process.*, Budapest, Hungary, 1995, pp. 67–72.

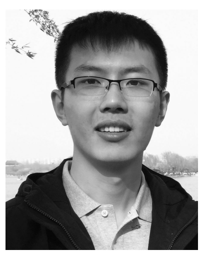
[36] J. Zhai, W. Yan, Z. Fu, and Y. Zhao, "Kinematic analysis of a dual-arm humanoid cooking robot," in *Proc. Int. Conf. Mechatronics Autom. (ICMA)*, 2012, pp. 249–254.

[37] J.-X. Xu, "A survey on iterative learning control for nonlinear systems," *Int. J. Control*, vol. 84, no. 7, pp. 1275–1294, 2011.

[38] Y. Zhu, D. Zhao, H. He, and J. Ji, "Event-triggered optimal control for partially unknown constrained-input systems via adaptive dynamic programming," *IEEE Trans. Ind. Electron.*, vol. 64, no. 5, pp. 4101–4109, May 2017.



XUHUI FENG is currently pursuing the Ph.D. degree with the School of Information Science and Technology, ShanghaiTech University. He was a Visiting Scholar with UC Berkeley, in 2017. He worked as a Research Intern with Mitsubishi Electrical Research Laboratories (MERL), Boston, in 2019. His research interests include model-based control, robust and self-reflective control, and global optimal control algorithms.



YUNING JIANG received the B.S. degree in electronic engineering from Shandong University, Jinan, China, 2014. He is currently pursuing the Ph.D. degree with the School of Information Science and Technology, ShanghaiTech University. His research interests include distributed optimization and model predictive control, in particular for power flow optimization, intelligent transportation, stochastic control, and fast MPC algorithms.