Bi-Level-Based Inverse Stochastic Optimal Control

Philipp Karg[1], Manuel Hess[1], Balint Varga[1] and Sören Hohmann[1]

Abstract— In this paper, we propose a new algorithm to solve the Inverse Stochastic Optimal Control (ISOC) problem of the linear-quadratic sensorimotor (LQS) control model. The LQS model represents the current state-of-the-art in describing goal-directed human movements. The ISOC problem aims at determining the cost function and noise scaling matrices of the LQS model from measurement data since both parameter types influence the statistical moments predicted by the model and are unknown in practice. We prove global convergence for our new algorithm and at a numerical example, validate the theoretical assumptions of our method. By comprehensive simulations, the influence of the tuning parameters of our algorithm on convergence behavior and computation time is analyzed. The new algorithm computes ISOC solutions nearly 33 times faster than the single previously existing ISOC algorithm.

I. INTRODUCTION

Inverse Optimal Control (IOC) methods [1] have gained significant research interest in the last years, from a theoretical- as well as application-oriented perspective. IOC methods answer the inverse question of OC problems: when are observed trajectories or control laws optimal, i.e. the result of an OC solution. Hence, IOC approaches aim at identifying the unknown cost function from observed trajectories, the so-called ground truth (GT) data, that are assumed to be optimal. An important application of IOC methods is the identification of goal-directed human movements. While there are strong indications (see e.g. [2], [3]) that goal-directed human movements can be described by an OC model, the corresponding cost function is unknown in practice and needs to be identified from measurement data to verify the model hypothesis or to use the identified model for the design of human-machine systems (e.g. prediction, classification or support of human movements). For deterministic OC models (see e.g. [3]), IOC methods are applied with these goals for example in [4]–[6]. However, deterministic OC models describe human movements by the so-called feedforward planning approach [7]. Via the OC model only the process of specifying a desired trajectory by the human is described. This movement planning stage follows a separated execution stage where the desired trajectory is tracked. Hence, sensory feedback on the movement trajectory is not taken into account adequately. Currently, feedback control approaches [7] based on stochastic optimal control (SOC) models are more promising since they consider the continuous sensory feedback of the human on its movement trajectories and are able to explain the stochastic nature of

them [2], [7]. The current main representative of these approaches is the linear-quadratic sensorimotor (LQS) control model, which is derived from the linear-quadratic Gaussian (LQG) model by adding a control-dependent noise process to the state and a state-dependent noise process to the output equation. Both extensions are used to reflect the special characteristics of the stochastic nature of human movements, i.e. faster movements are performed and perceived more inaccurately [2], [8][1].

As for deterministic OC models, for SOC models an identification method is needed to determine their unknown parameters from measurement/GT data in order to verify the model hypothesis in a specific task in a data-driven manner [9] or to use the identified models for the design of human-machine systems. Until very recently [10], an approach that solves this Inverse Stochastic Optimal Control (ISOC) problem for the LQS model was missing and only impractical special cases, like the LQG model [11] or well-selected parameters of the LQS model [12], were considered so far. For the LQS model, the ISOC problem consists of identifying the cost function and noise scaling matrices since both parameter types influence the predicted statistical moments of the system quantities. However, our previous method [10] still shows high computation times (nearly one day on a standard PC) and lacks a guaranteed, proved global convergence behavior. Both challenges are solved in this paper by proposing a new ISOC algorithm with proved global convergence, which solves the ISOC problem in a simulation example nearly 33 times faster than our previous method [10]. Based on the simulation example, we furthermore validate the theoretical assumptions of our new approach and analyze the influence of its tuning parameters on convergence behavior and computation time.

II. PROBLEM DEFINITION

As explained in Section I, when describing the human biomechanics with a state equation, we need to consider a control-dependent noise process $\{\sum_{i=1}^{c} \sigma_i^{\boldsymbol{u}} \varepsilon_t^{(i)} \boldsymbol{B} \boldsymbol{F}_i \boldsymbol{u}_t\}$:

$$\boldsymbol{x}_{t+1} = \boldsymbol{A}\boldsymbol{x}_t + \boldsymbol{B}\boldsymbol{u}_t + \boldsymbol{\Sigma}^{\boldsymbol{\alpha}}\boldsymbol{\alpha}_t + \sum_{i=1}^{c} \sigma_i^{\boldsymbol{u}} \varepsilon_t^{(i)} \boldsymbol{B} \boldsymbol{F}_i \boldsymbol{u}_t, \quad (1)$$

where $\boldsymbol{x} \in \mathbb{R}^n$ denotes the system state, $\boldsymbol{u} \in \mathbb{R}^m$ the control variable, $\{\boldsymbol{\alpha}_t\}$ a standard white Gaussian noise process in sample space $\mathbb{R}^p$ and $\{\boldsymbol{\varepsilon}_t\}$ ($\boldsymbol{\varepsilon}_t^{\mathsf{T}} = \begin{bmatrix} \varepsilon_t^{(1)} & \dots & \varepsilon_t^{(c)} \end{bmatrix}$) a standard white Gaussian noise process in sample space $\mathbb{R}^c$.

---

[1]All authors are with the Institute of Control Systems (IRS), Karlsruhe Institute of Technology (KIT), 76131 Karlsruhe, Germany. Corresponding author is Philipp Karg, `philipp.karg@kit.edu`.

[1]With the LQS model, the human hand is typically modeled as point mass since the biomechanics of the complete human arm are nonlinear in general.

Furthermore, $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{\Sigma^\alpha}$ and $\boldsymbol{F}_i$ are matrices of appropriate dimension, where $\boldsymbol{A}$ and $\boldsymbol{B}$ may depend on time. The random variables $\varepsilon_t^{(i)}$ are scaled with a constant $\sigma_i^{\boldsymbol{u}} \in \mathbb{R}$ and with $\boldsymbol{u}_t$ [8]. The stochastic process $\{\boldsymbol{x}_t\}$ is initialized with $\mathrm{E}\{\boldsymbol{x}_0\}$ and $\mathrm{cov}(\boldsymbol{x}_0, \boldsymbol{x}_0) = \boldsymbol{\Omega}_0^{\boldsymbol{x}}$. Human perception is described by an output equation with a state-dependent noise process $\{\sum_{i=1}^d \sigma_i^{\boldsymbol{x}} \epsilon_t^{(i)} \boldsymbol{H} \boldsymbol{G}_i \boldsymbol{x}_t\}$ in the LQS model:

$$\boldsymbol{y}_t = \boldsymbol{H}\boldsymbol{x}_t + \boldsymbol{\Sigma^\beta}\boldsymbol{\beta}_t + \sum_{i=1}^d \sigma_i^{\boldsymbol{x}} \epsilon_t^{(i)} \boldsymbol{H}\boldsymbol{G}_i\boldsymbol{x}_t, \qquad (2)$$

where $\boldsymbol{y} \in \mathbb{R}^r$ denotes the observed output, $\{\boldsymbol{\beta}_t\}$ a standard white Gaussian noise process in sample space $\mathbb{R}^q$ and $\{\boldsymbol{\epsilon}_t\}$ ($\boldsymbol{\epsilon}_t^\mathsf{T} = \begin{bmatrix} \epsilon_t^{(1)} & \dots & \epsilon_t^{(d)} \end{bmatrix}$) a standard white Gaussian noise process in sample space $\mathbb{R}^d$. Moreover, $\boldsymbol{H}$, $\boldsymbol{\Sigma^\beta}$ and $\boldsymbol{G}_i$ are matrices of appropriate dimension, where $\boldsymbol{H}$ may depend on time. Similarly to (1), the random variables $\epsilon_t^{(i)}$ are scaled by a constant $\sigma_i^{\boldsymbol{x}} \in \mathbb{R}$ and $\boldsymbol{x}_t$ [8]. The random variables $\boldsymbol{x}_t$, $\boldsymbol{\alpha}_t$, $\boldsymbol{\varepsilon}_t$, $\boldsymbol{\beta}_t$ and $\boldsymbol{\epsilon}_t$ at time t are assumed to be independent to each other. The performance criterion which drives the goal-directed human movements is given by

$$J = \mathrm{E}\left\{ \boldsymbol{x}_N^\mathsf{T} \boldsymbol{Q}_N \boldsymbol{x}_N + \sum_{t=0}^{N-1} \boldsymbol{x}_t^\mathsf{T} \boldsymbol{Q}_t \boldsymbol{x}_t + \boldsymbol{u}_t^\mathsf{T} \boldsymbol{R} \boldsymbol{u}_t \right\}, \qquad (3)$$

where $\boldsymbol{Q}_N = \sum_{i=1}^{S_N} s_{N,i} \boldsymbol{q}_{N,i} \boldsymbol{q}_{N,i}^\mathsf{T}$ ($\boldsymbol{q}_{N,i} \in \mathbb{R}^n$, $s_{N,i} \in \mathbb{R}$), $\boldsymbol{Q}_t = \sum_{i=1}^{S_Q} s_{Q,i} \boldsymbol{q}_{Q,t,i} \boldsymbol{q}_{Q,t,i}^\mathsf{T}$ ($\boldsymbol{q}_{Q,t,i} \in \mathbb{R}^n$, $s_{Q,i} \in \mathbb{R}$) and $\boldsymbol{R} = \sum_{i=1}^{S_R} s_{R,i} \boldsymbol{q}_{R,i} \boldsymbol{q}_{R,i}^\mathsf{T}$ ($\boldsymbol{q}_{R,i} \in \mathbb{R}^m$, $s_{R,i} \in \mathbb{R}$).

Solving the LQS optimal control Problem 1 models the optimization performed by the human in its perception-action cycle.

Problem 1: Find an admissible control strategy $\boldsymbol{u}_t = \boldsymbol{\pi}_t(\boldsymbol{u}_0, \dots, \boldsymbol{u}_{t-1}, \boldsymbol{y}_0, \dots, \boldsymbol{y}_{t-1})$ for the system defined by (1) and (2) such that (3) is minimal.

Due to the multiplicative noise processes in (1) and (2), the separation theorem does not hold [8], [13]–[15]. Following the procedure proposed in [8], Lemma 1 shows the solution of Problem 1 considered throughout the paper which takes the interdependence between controller and filter into account.

Lemma 1: Let $\boldsymbol{R}$ and $\boldsymbol{\Sigma^\beta}\boldsymbol{\Sigma^\beta}^\mathsf{T}$ be positive definite. Furthermore, let the history of control and output values for the admissible control strategies $\boldsymbol{u}_t = \boldsymbol{\pi}_t(\boldsymbol{u}_0, \dots, \boldsymbol{u}_{t-1}, \boldsymbol{y}_0, \dots, \boldsymbol{y}_{t-1})$ (cf. Problem 1) be represented by the estimation $\hat{\boldsymbol{x}}_t$ of a linear filter:

$$\hat{\boldsymbol{x}}_{t+1} = \boldsymbol{A}\hat{\boldsymbol{x}}_t + \boldsymbol{B}\boldsymbol{u}_t + \boldsymbol{K}_t(\boldsymbol{y}_t - \boldsymbol{H}\hat{\boldsymbol{x}}_t) + \boldsymbol{\Sigma^\gamma}\boldsymbol{\gamma}_t, \qquad (4)$$

where $\boldsymbol{K}_t$ ($\forall t \in \{0, \dots, N-1\}$) are constant filter matrices of appropriate dimension and $\{\boldsymbol{\gamma}_t\}^2$ a standard white Gaussian noise process in $\mathbb{R}^l$ scaled by a constant matrix $\boldsymbol{\Sigma^\gamma}$ of appropriate dimension. Then, the optimal control law is

given by $\boldsymbol{u}_t = \boldsymbol{\pi}_t(\hat{\boldsymbol{x}}_t) = -\boldsymbol{L}_t\hat{\boldsymbol{x}}_t$ with[3]

$$\boldsymbol{L}_t = \left(\boldsymbol{R} + \boldsymbol{B}^\mathsf{T}\boldsymbol{Z}_{t+1}^{\boldsymbol{x}}\boldsymbol{B} \right.$$
$$\left. + \sum_i (\sigma_i^{\boldsymbol{u}})^2 \boldsymbol{F}_i^\mathsf{T}\boldsymbol{B}^\mathsf{T} \left(\boldsymbol{Z}_{t+1}^{\boldsymbol{x}} + \boldsymbol{Z}_{t+1}^{\boldsymbol{e}} \right) \boldsymbol{B}\boldsymbol{F}_i \right)^{-1} \boldsymbol{B}^\mathsf{T}\boldsymbol{Z}_{t+1}^{\boldsymbol{x}}\boldsymbol{A}. \tag{5}$$

Moreover, the optimal constant filter matrices $\boldsymbol{K}_t$ are given by[4]

$$\boldsymbol{K}_t = \boldsymbol{A}\boldsymbol{P}_t^{\boldsymbol{e}}\boldsymbol{H}^\mathsf{T}\left(\boldsymbol{H}\boldsymbol{P}_t^{\boldsymbol{e}}\boldsymbol{H}^\mathsf{T} + \boldsymbol{\Sigma^\beta}\boldsymbol{\Sigma^\beta}^\mathsf{T} \right.$$
$$\left. + \sum_i (\sigma_i^{\boldsymbol{x}})^2 \boldsymbol{H}\boldsymbol{G}_i \left(\boldsymbol{P}_t^{\boldsymbol{e}} + \boldsymbol{P}_t^{\hat{\boldsymbol{x}}} + \boldsymbol{P}_t^{\hat{\boldsymbol{x}}\boldsymbol{e}} + \boldsymbol{P}_t^{\boldsymbol{e}\hat{\boldsymbol{x}}} \right) \boldsymbol{G}_i^\mathsf{T}\boldsymbol{H}^\mathsf{T} \right)^{-1}. \tag{6}$$

Proof: With the assumption on a linear filter providing a sufficient statistic for the history of control and output values, we can derive a Bellman equation based on [16, Lemma 3.2, p. 261]. Minimizing J (3) w.r.t. the admissible control strategy $\boldsymbol{\pi}_0, \boldsymbol{\pi}_1, \dots$ yields

$$\min_{\boldsymbol{\pi}_0, \dots} J = \mathrm{E}\left\{ \min_{\boldsymbol{\pi}_0, \dots} \mathrm{E}\{\mathrm{E}\{J \mid \boldsymbol{x}_0, \hat{\boldsymbol{x}}_0\} \mid \hat{\boldsymbol{x}}_0\} \right\}. \tag{7}$$

By defining a value function

$$V_t(\boldsymbol{x}_t, \hat{\boldsymbol{x}}_t) = \mathrm{E}\left\{ \boldsymbol{x}_N^\mathsf{T}\boldsymbol{Q}_N\boldsymbol{x}_N \right.$$
$$\left. + \sum_{\tau=t}^{N-1} \boldsymbol{x}_\tau^\mathsf{T}\boldsymbol{Q}_\tau\boldsymbol{x}_\tau + \boldsymbol{u}_\tau^\mathsf{T}\boldsymbol{R}\boldsymbol{u}_\tau \mid \boldsymbol{x}_t, \hat{\boldsymbol{x}}_t \right\} \tag{8}$$

and its optimized version $V_t^*(\hat{\boldsymbol{x}}_t) = \min_{\boldsymbol{\pi}_t, \dots} \mathrm{E}\{V_t(\boldsymbol{x}_t, \hat{\boldsymbol{x}}_t) \mid \hat{\boldsymbol{x}}_t\}$, from (7) the recursive Bellman equation

$$V_t(\boldsymbol{x}_t, \hat{\boldsymbol{x}}_t) = \mathrm{E}\{\boldsymbol{x}_t^\mathsf{T}\boldsymbol{Q}_t\boldsymbol{x}_t + \boldsymbol{u}_t^\mathsf{T}\boldsymbol{R}\boldsymbol{u}_t$$
$$+ V_{t+1}(\boldsymbol{x}_{t+1}, \hat{\boldsymbol{x}}_{t+1}) \mid \boldsymbol{x}_t, \hat{\boldsymbol{x}}_t\} \tag{9}$$

follows. Now, it can be shown by induction (backwards, starting with $t = N$) that the value function V_t is quadratic, i.e. $V_t(\boldsymbol{x}_t, \hat{\boldsymbol{x}}_t) = \boldsymbol{x}_t^\mathsf{T}\boldsymbol{Z}_t^{\boldsymbol{x}}\boldsymbol{x}_t + \boldsymbol{e}_t^\mathsf{T}\boldsymbol{Z}_t^{\boldsymbol{e}}\boldsymbol{e}_t + z_t$ ($\boldsymbol{e}_t = \boldsymbol{x}_t - \hat{\boldsymbol{x}}_t$) [8]. Moreover, by computing $V_t^*(\hat{\boldsymbol{x}}_t)$ from $V_t(\boldsymbol{x}_t, \hat{\boldsymbol{x}}_t)$, the linear control law $\boldsymbol{\pi}_t(\hat{\boldsymbol{x}}_t) = -\boldsymbol{L}_t\hat{\boldsymbol{x}}_t$ with (5) results. The calculation of V_t^* and $\boldsymbol{\pi}_t$ is performed using (9) and assuming a given, fixed sequence of filter gains $\boldsymbol{K}_t$. In order to find the optimal filter matrices, J is minimized w.r.t. to $\boldsymbol{K}_0, \boldsymbol{K}_1, \dots$ with assumed fixed control matrices $\boldsymbol{L}_t$. This can be done via forward recursion $\boldsymbol{K}_t = \arg\min_{\boldsymbol{K}_t} \mathrm{E}\{V_{t+1}^*(\hat{\boldsymbol{x}}_{t+1})\}$ ($\forall t \in \{0, \dots, N-1\}$), which yields (6). $\blacksquare$

Since $\boldsymbol{L}_t$ is derived under the assumption of specific filter gains $\boldsymbol{K}_t$ and the filter gains vice versa, their calculation is done via an iterative procedure starting with a guess for $\boldsymbol{K}_t$. This leads to an alternating optimization (AO) [17]. In each AO optimization (with respect to $\boldsymbol{L}_t$ or $\boldsymbol{K}_t$) a unique global optimizer exists due to $\boldsymbol{R}$ and $\boldsymbol{\Sigma^\beta}\boldsymbol{\Sigma^\beta}^\mathsf{T}$ being positive definite. If in one AO optimization step the global optimal solution for $\boldsymbol{L}_t$ or $\boldsymbol{K}_t$ results, global convergence for the

---

[2]The random variables $\boldsymbol{\gamma}_t$ are independent to $\boldsymbol{x}_t$, $\boldsymbol{\alpha}_t$, $\boldsymbol{\varepsilon}_t$, $\boldsymbol{\beta}_t$ and $\boldsymbol{\epsilon}_t$.

[3]Recursive formulas for $\boldsymbol{Z}_t^{\boldsymbol{x}}$ and $\boldsymbol{Z}_t^{\boldsymbol{e}}$ can be found in [8].
[4]Recursive formulas for $\boldsymbol{P}_t^{\hat{\boldsymbol{x}}}$, $\boldsymbol{P}_t^{\boldsymbol{e}}$ and $\boldsymbol{P}_t^{\hat{\boldsymbol{x}}\boldsymbol{e}} = \boldsymbol{P}_t^{\boldsymbol{e}\hat{\boldsymbol{x}}}$ can be found in [8].

other parameter type follows [17, Theorem 2]. This behavior is guaranteed e.g. in the LQG case (separation theorem).

The statistical moments of $\{x_t\}$ according to Theorem 1 characterize the predictions of the LQS model.

Theorem 1: Applying the optimal control strategy according to Lemma 1 to the system defined by (1) and (2) leads to the mean $\mathrm{E}\{x_t\}$ and covariance $\Omega_t^x = \mathrm{cov}(x_t, x_t)$:

$$\begin{bmatrix} \mathrm{E}\{x_{t+1}\} \\ \mathrm{E}\{\hat{x}_{t+1}\} \end{bmatrix} = \mathcal{A}_t \begin{bmatrix} \mathrm{E}\{x_t\} \\ \mathrm{E}\{\hat{x}_t\} \end{bmatrix}, \tag{10}$$

$$\begin{bmatrix} \Omega_{t+1}^x & \Omega_{t+1}^{x\hat{x}} \\ \Omega_{t+1}^{\hat{x}x} & \Omega_{t+1}^{\hat{x}} \end{bmatrix} = \mathcal{A}_t \begin{bmatrix} \Omega_t^x & \Omega_t^{x\hat{x}} \\ \Omega_t^{\hat{x}x} & \Omega_t^{\hat{x}} \end{bmatrix} \mathcal{A}_t^\mathsf{T}$$
$$+ \begin{bmatrix} \Omega^\xi & 0 \\ 0 & \Omega^\eta + K_t \Omega^\omega K_t^\mathsf{T} \end{bmatrix} + \begin{bmatrix} \bar{\Omega}_t^{\hat{x}} & 0 \\ 0 & \bar{\Omega}_t^x \end{bmatrix}, \tag{11}$$

with $\Omega_t^\xi = \Sigma^\alpha \Sigma^{\alpha\mathsf{T}}$, $\Omega_t^\eta = \Sigma^\gamma \Sigma^{\gamma\mathsf{T}}$, $\Omega_t^\omega = \Sigma^\beta \Sigma^{\beta\mathsf{T}}$, $\bar{\Omega}_t^{\hat{x}} = \sum_i (\sigma_i^u)^2 B F_i L_t \left(\Omega_t^{\hat{x}} + \mathrm{E}\{\hat{x}_t\} \mathrm{E}\{\hat{x}_t\}^\mathsf{T} \right) L_t^\mathsf{T} F_i^\mathsf{T} B^\mathsf{T}$, $\bar{\Omega}_t^x = \sum_i (\sigma_i^x)^2 K_t H G_i \left(\Omega_t^x + \mathrm{E}\{x_t\} \mathrm{E}\{x_t\}^\mathsf{T} \right) G_i^\mathsf{T} H^\mathsf{T} K_t^\mathsf{T}$,

$$\mathcal{A}_t = \begin{bmatrix} A & -BL_t \\ K_t H & A - K_t H - BL_t \end{bmatrix} \tag{12}$$

and $\mathrm{E}\{\hat{x}_0\} = \mathrm{E}\{x_0\}$, $\Omega_t^{\hat{x}x} = \Omega_t^{x\hat{x}} = \Omega_t^{\hat{x}} = 0$.

Proof: See [10, Lemma 2]. ∎

Although Theorem 1 gives recursive formulas to compute mean $\mathrm{E}\{x_t\}$ and covariance Ω_t^x of the system state, i.e. in context of human movement modeling the average behavior and the variability patterns of the movements, unknown model parameters need to be identified from human measurement data before the model can be used. This leads to the inverse problem (cf. Problem 2) of the LQS optimal control Problem 1. From GT data (cf. Definition 1) that are assumed to be measured realizations of the stochastic processes of (in general only some) system states resulting in the LQS model with unknown GT parameters θ^* (cf. Assumption 1), sufficient approximations of mean $\hat{m}_t$ and covariance values $\hat{\Omega}_t^{x^*}$ are computed. Then, the ISOC problem aims at finding parameters $\theta^\triangle$ that yield a stochastic process $\{x_t^\triangle\}$ with matching mean and covariance values to the GT data. Corollary 1 defines the parameter θ that influence the model predictions $\mathrm{E}\{x_t\}$ and Ω_t^x and thus, need to be identified. Identifying cost function and noise scaling matrices from GT data distinguishes the ISOC problem from deterministic IOC problems, which only search for cost function parameters.

Corollary 1: Applying the optimal control strategy according to Lemma 1 to the system defined by (1) and (2) leads to the mean $\mathrm{E}\{x_t\}$ (10) and covariance Ω_t^x (11), which both depend on $\theta = \begin{bmatrix} s^\mathsf{T} & \sigma^\mathsf{T} \end{bmatrix}^\mathsf{T} \in \mathbb{R}^\Theta$ with $s^\mathsf{T} = \begin{bmatrix} s_{N,1} & \cdots & s_{N,S_N} & s_{Q,1} & \cdots & s_{Q,S_Q} & s_{R,1} & \cdots & s_{R,S_R} \end{bmatrix}$ and $\sigma^\mathsf{T} = \begin{bmatrix} \mathrm{vec}(\Sigma^\alpha)^\mathsf{T} & \mathrm{vec}(\Sigma^\beta)^\mathsf{T} & \sigma_1^u & \cdots & \sigma_c^u & \sigma_1^x & \cdots & \sigma_d^x \end{bmatrix}$ [5].

Proof: The dependencies can directly be derived from Theorem 1: (10) and (11) depend on $\mathcal{A}_t$ which in turn depends on L_t; then, L_t on R and σ_i^u and via Z_t^x on Q_N, Q_t, σ_i^x and K_t, which finally depends on Σ^β and via P_t^e on Σ^α. ∎

---

[5] We assume $\Sigma^\gamma = 0$ throughout the paper. Its consideration would be straightforward.

Definition 1: The GT data are realizations of an unknown stochastic process $\{M\bar{x}_t\}$, where $\bar{x} \in \mathbb{R}^n$ and the system state x in (1) represent the same physical quantities. Furthermore, $M \in \mathbb{R}^{\bar{n} \times n}$ follows from the identity matrix by deleting rows that correspond to states/quantities that are not measured in the GT data. Finally, $\hat{m}_t$ and $\hat{\Omega}_t^{x^*}$ are mean and covariance values computed from the realizations of $\{M\bar{x}_t\}$.

Assumption 1: The GT data are realizations of the stochastic process $\{Mx_t^*\}$, where $\{x_t^*\}$ results from the application of the admissible control strategy that is optimal w.r.t. the GT parameters θ^* (cf. Lemma 1) to the system (1), (2): $\hat{m}_t \approx \mathrm{E}\{Mx_t^*\}$ and $\hat{\Omega}_t^{x^*} \approx M\Omega_t^{x^*}M^\mathsf{T}$ hold.

Assumption 2: The matrices A, B, F_i, H and G_i are known as well as the vectors $q_{N,i}$, $q_{Q,t,i}$ and $q_{R,i}$.

Problem 2: Find parameters $\theta^\triangle$ such that the stochastic process $\{x_t^\triangle\}$, which results from the application of the admissible control strategy that is optimal w.r.t. the parameters $\theta^\triangle$ (cf. Lemma 1) to the system (1), (2), yields $\mathrm{E}\{Mx_t^\triangle\} = \hat{m}_t$ and $M\Omega_t^{x^\triangle}M^\mathsf{T} = \hat{\Omega}_t^{x^*}$.

Remark 1: Although the stochastic processes $\{\alpha_t\}$, $\{\varepsilon_t\}$, $\{\beta_t\}$, $\{\epsilon_t\}$ and $\{\gamma_t\}$ in (1), (2) and (4) are Gaussian, $\{x_t\}$ is not, due to the multiplicative noise processes in (1) and (2) (see also [15]).

III. BI-LEVEL-BASED INVERSE STOCHASTIC OPTIMAL CONTROL

In this section, we propose our new ISOC algorithm to solve Problem 2 which is based on a bi-level structure. Bi-level-based algorithms are common to solve inverse problems of optimal control (see e.g. [5], [6], [9], [10], [18]). In general, in such methods a parameter optimization problem is defined where the objective function quantifies how well the trajectories, which are optimal w.r.t. a current guess of the cost function and the noise scaling parameters, match the observed trajectories in the GT data. Since the optimization of this objective function requires for each function evaluation the solution of a SOC problem, a so-called bi-level structure results. Typically, it is unknown or need to be verified if the GT trajectories are indeed a SOC solution, i.e. Assumption 1 cannot be guaranteed. In these situations, it is inevitable to have a bi-level-based ISOC approach, at least as baseline method for more advanced ones, since only bi-level-based algorithms with global convergence behavior guarantee the best possible parameters w.r.t. the chosen metric to define the objective function of the parameter optimization problem.

In Subsection III-A, we define the parameter optimization problem of our new ISOC algorithm. Then, we explain our new approach to solve it in Subsection III-B (so-called upper level optimization) and how we treat its non-convexity. In Subsection III-C, the solution to the SOC problem, which is a constraint of the parameter optimization problem defined in Subsection III-A and called lower level optimization, is discussed.

A. Parameter Optimization Problem

Problem 3 defines the parameter optimization derived from Problem 2. The objective function $J_{\mathrm{ISOC}}(\theta, \hat{m}_t, \hat{\Omega}_t^{x^*})$ is

based on the variance accounted for (VAF) metric, successfully applied for model regression with real data [9], [12]. The optimization w.r.t. $\boldsymbol{\theta}$ is performed on a feasible set $\mathcal{U}$ to guarantee the existence of a well-defined SOC solution in the lower level constituted by constraints (14c) and (14d).

Definition 2: The feasible set $\mathcal{U}$ is given by

$$\mathcal{U} = \{\boldsymbol{\theta} \in \mathbb{R}^{\Theta} : \theta_i \in [a_i, b_i], \forall i \in \{1, \dots, \Theta\}\}, \quad (13)$$

where $a_i, b_i \in \mathbb{R}$ and $b_i > a_i$ ($\forall i \in \{1, \dots, \Theta\}$), such that for every $\boldsymbol{\theta} \in \mathcal{U}$ Lemma 1 can be applied and such that $\boldsymbol{\theta}^* \in \mathcal{U}$ if Assumption 1 holds.

Problem 3: Find a global optimizer $\boldsymbol{\theta}^{\triangle}$ of

$$\min_{\boldsymbol{\theta}} \quad J_{\text{ISOC}}(\boldsymbol{\theta}, \hat{\boldsymbol{m}}_t, \hat{\boldsymbol{\Omega}}_t^{\boldsymbol{x}^*}) =$$

$$\min_{\boldsymbol{\theta}} -\frac{\boldsymbol{w}_{\text{m}}^{\mathsf{T}} \boldsymbol{m}^{\text{VAF}}(\boldsymbol{\theta}, \hat{\boldsymbol{m}}_t) + \boldsymbol{w}_{\text{v}}^{\mathsf{T}} \text{vec}\left(\boldsymbol{\Omega}^{\text{VAF}}(\boldsymbol{\theta}, \hat{\boldsymbol{\Omega}}_t^{\boldsymbol{x}^*})\right)}{\|\boldsymbol{w}_{\text{m}}\|_1 + \|\boldsymbol{w}_{\text{v}}\|_1} \quad (14a)$$

$$\text{s.t.} \quad \boldsymbol{\theta} \in \mathcal{U} \quad (14b)$$

$$\boldsymbol{L}_t, \boldsymbol{K}_t = \arg\min_{\boldsymbol{L}_t, \boldsymbol{K}_t} J(\boldsymbol{\theta}) \text{ (Lemma 1)} \quad (14c)$$

$$\text{E}\{\boldsymbol{x}_t\} \text{ (10)}, \boldsymbol{\Omega}_t^{\boldsymbol{x}} \text{ (11) with } \boldsymbol{L}_t, \boldsymbol{K}_t, \boldsymbol{\theta} \text{ (Theorem 1)} \quad (14d)$$

where

$$m_i^{\text{VAF}} = \left(1 - \frac{\sum_{t=0}^{N}\left((\text{E}\{\boldsymbol{M}\boldsymbol{x}_t\})_i - \hat{m}_{i,t}\right)^2}{\sum_{t=0}^{N}\left(\hat{m}_{i,t} - \frac{1}{N+1}\sum_t \hat{m}_{i,t}\right)^2}\right), \quad (15)$$

$$\Omega_{ij}^{\text{VAF}} = \left(1 - \frac{\sum_{t=0}^{N}\left((\boldsymbol{M}\boldsymbol{\Omega}_t^{\boldsymbol{x}}\boldsymbol{M}^{\mathsf{T}})_{ij} - \hat{\Omega}_{ij,t}^{\boldsymbol{x}^*}\right)^2}{\sum_{t=0}^{N}\left(\hat{\Omega}_{ij,t}^{\boldsymbol{x}^*} - \frac{1}{N+1}\sum_t \hat{\Omega}_{ij,t}^{\boldsymbol{x}^*}\right)^2}\right) \quad (16)$$

with $i, j \in \{1, \dots, \bar{n}\}$ and $\boldsymbol{w}_{\text{m}} \in \mathbb{R}^{\bar{n}}$ as well as $\boldsymbol{w}_{\text{v}} \in \mathbb{R}^{\bar{n}\bar{n}}$ are arbitrary weighting vectors.

In general, $J_{\text{ISOC}}(\boldsymbol{\theta}, \hat{\boldsymbol{m}}_t, \hat{\boldsymbol{\Omega}}_t^{\boldsymbol{x}^*})$ is non-convex (cf. Assumption 3). Furthermore, Assumption 4[6] motivates the use of derivative-based optimization solvers with numerically approximated derivatives since analytical expressions of the derivatives cannot be given due to the bi-level structure (see (14c) and (14d)). Finally, Lemma 2 characterizes the set $\mathcal{T}$ of global optimizers of (14).

Assumption 3: $J_{\text{ISOC}}(\boldsymbol{\theta}, \hat{\boldsymbol{m}}_t, \hat{\boldsymbol{\Omega}}_t^{\boldsymbol{x}^*})$ is non-convex w.r.t. $\boldsymbol{\theta}$.

Assumption 4: $J_{\text{ISOC}}(\boldsymbol{\theta}, \hat{\boldsymbol{m}}_t, \hat{\boldsymbol{\Omega}}_t^{\boldsymbol{x}^*})$ is (at least) twice continuously differentiable w.r.t. $\boldsymbol{\theta}$.

Definition 3: Let $\mathcal{T} \subseteq \mathcal{U}$ be the set of global optimizers of (14).

Lemma 2: Let Assumption 4 hold. Then, $\mathcal{T} \neq \emptyset$. Moreover, if Assumption 1 holds additionally, every $\boldsymbol{\theta} \in \mathcal{T}$ solves Problem 2 and for every $\boldsymbol{\theta} \in \mathcal{T}$ $J_{\text{ISOC}}(\boldsymbol{\theta}, \hat{\boldsymbol{m}}_t, \hat{\boldsymbol{\Omega}}_t^{\boldsymbol{x}^*}) = -1$.

Proof: Since $\mathcal{U}$ is compact and $J_{\text{ISOC}}(\boldsymbol{\theta}, \hat{\boldsymbol{m}}_t, \hat{\boldsymbol{\Omega}}_t^{\boldsymbol{x}^*})$ continuous w.r.t. $\boldsymbol{\theta}$ (cf. Assumption 4), (at least) one global maximizer and minimizer exists for $J_{\text{ISOC}}(\boldsymbol{\theta}, \hat{\boldsymbol{m}}_t, \hat{\boldsymbol{\Omega}}_t^{\boldsymbol{x}^*})$ on

---

[6]In references on bi-level-based algorithms (see e.g. [5], [6], [10], [18]), the necessity of derivative-free optimization solvers is postulated. However, with our definition of the optimization problem (14), we show in Subsection IV-B numerically that $J_{\text{ISOC}}(\boldsymbol{\theta}, \hat{\boldsymbol{m}}_t, \hat{\boldsymbol{\Omega}}_t^{\boldsymbol{x}^*})$ has indeed sufficient differentiability characteristics.

$\mathcal{U}$. Hence, $\mathcal{T} \neq \emptyset$ follows. Now, let Assumption 1 hold additionally. First, we have $J_{\text{ISOC}}(\boldsymbol{\theta}^*, \hat{\boldsymbol{m}}_t, \hat{\boldsymbol{\Omega}}_t^{\boldsymbol{x}^*}) = -1$, which is the global minimal value due to the used VAF metric for m_i^{VAF} (15) and Ω_{ij}^{VAF} (16), and since $\boldsymbol{\theta}^* \in \mathcal{U}$ (see Definition 2), $\boldsymbol{\theta}^* \in \mathcal{T} \neq \emptyset$ follows and every $\boldsymbol{\theta} \in \mathcal{T}$ needs to yield $J_{\text{ISOC}}(\boldsymbol{\theta}, \hat{\boldsymbol{m}}_t, \hat{\boldsymbol{\Omega}}_t^{\boldsymbol{x}^*}) = -1$. ∎

Remark 2: If Assumption 1 does not hold, $\mathcal{T} \neq \emptyset$ still holds according to Lemma 2. However, $J_{\text{ISOC}}(\boldsymbol{\theta}, \hat{\boldsymbol{m}}_t, \hat{\boldsymbol{\Omega}}_t^{\boldsymbol{x}^*}) = -1$ cannot be guaranteed for $\boldsymbol{\theta} \in \mathcal{T}$. The elements of $\mathcal{T}$ yield the best possible fit of the LQS model to $\hat{\boldsymbol{m}}_t$ and $\hat{\boldsymbol{\Omega}}_t^{\boldsymbol{x}^*}$ computed with the GT data. Assumption 1 is not fulfilled if either too few realizations are measured for sufficient approximations of mean and covariance values or the LQS model is not valid to describe the GT data.

B. Upper Level Optimization: Threshold Random Linkage Algorithm with Approximation of Regions of Attraction

In this subsection, we propose two algorithms, together with their convergence proofs, to solve Problem 3. Lemma 2 shows that under Assumption 1 these solutions solve Problem 2 as well. Algorithm 2 is our main algorithm, influenced by [19], [20] and called Threshold Random Linkage Algorithm with Approximation of Regions of Attraction (TRLwARoA). The main idea of our algorithms is to use a derivative-based constrained local optimization solver Υ according to Definition 4 that converges to a local minimizer of (14). Then, this local solver is combined with a globalization strategy to account for the non-convexity of $J_{\text{ISOC}}(\boldsymbol{\theta})$.

Definition 4: Let Assumption 4 hold and let a feasible set $\mathcal{U}$ (cf. Definition 2) be given. A constrained local optimization solver $\Upsilon : \mathcal{U} \to \mathcal{U}$ yields $\boldsymbol{\theta}_{\min} = \Upsilon(\boldsymbol{\theta}_0)$, where $\boldsymbol{\theta}_{\min} \in \Theta_{\min}$ and $\boldsymbol{\theta}_0 \in \Theta_{\text{RoA}}(\Theta_{\min})$. Here, $\Theta_{\min}$ denotes a connected subset of all local minimizers $\{\boldsymbol{\theta}_{\min} \in \mathcal{U} : \exists \mathcal{N}(\boldsymbol{\theta}_{\min}) \text{ such that } J_{\text{ISOC}}(\boldsymbol{\theta}) \geq J_{\text{ISOC}}(\boldsymbol{\theta}_{\min}), \forall \boldsymbol{\theta} \in \mathcal{N}(\boldsymbol{\theta}_{\min})\}$ (with $\mathcal{N}(\boldsymbol{\theta}_{\min})$ as neighborhood of $\boldsymbol{\theta}_{\min}$) and $\Theta_{\text{RoA}}(\Theta_{\min})$ the Region of Attraction (RoA) of $\Theta_{\min}$, i.e. $\Theta_{\text{RoA}}(\Theta_{\min}) = \{\boldsymbol{\theta} \in \mathcal{U} : \exists \boldsymbol{\theta}_{\min} \in \Theta_{\min} \text{ such that } \forall \boldsymbol{\theta} \in \mathcal{N}(\boldsymbol{\theta}_{\min}) \exists \boldsymbol{\theta}'_{\min} \in \Theta_{\min} : \boldsymbol{\theta}'_{\min} = \Upsilon(\boldsymbol{\theta})\}$.

In Definition 4, we consider connected sets of local minima due to the well-known ambiguity of solutions of ISOC problems [1], [10]. For example, if Assumption 1 is fulfilled, every $\boldsymbol{\theta} = \begin{bmatrix} \lambda \boldsymbol{s}^{*\mathsf{T}} & \boldsymbol{\sigma}^{*\mathsf{T}} \end{bmatrix}^{\mathsf{T}} \in \mathcal{T}$ with $\lambda \in \mathbb{R}_{\geq 0}$ and such that $\boldsymbol{\theta} \in \mathcal{U}$. Hence, sufficient conditions for strict optimizers are normally not fulfilled in ISOC problems. The shape of the RoA Θ_{RoA} of a set $\Theta_{\min}$ depends on the concrete implementation of the used local solver. However, there are convergence results for basic implementations of SQP and interior-point (IP) algorithms (see e.g. [21, Theorem 18.3, Theorem 19.1]) that are applicable for (14). Later we use an IP algorithm as local solver. Since $J_{\text{ISOC}}(\boldsymbol{\theta})$ as well as the inequality constraints given by $\mathcal{U}$ are continuously differentiable and since the LICQ hold $\forall \boldsymbol{\theta} \in \mathcal{U}$ (cf. Lemma 3), from [21, Theorem 19.1] the convergence of IP algorithms to first-order optimal points follows.

Lemma 3: Let a feasible set $\mathcal{U}$ according to Definition 2 be given. The linear independence constraint qualification

Algorithm 1: Pure Multi-Start ISOC Algorithm.

Input: $\hat{m}_t$, $\hat{\Omega}_t^{x^*}$, $\mathcal{U}$, $k_{\max}$
Output: $\boldsymbol{\theta}^{\triangle}$, $J_{\text{ISOC}}(\boldsymbol{\theta}^{\triangle}, \hat{m}_t, \hat{\Omega}_t^{x^*})$
Draw $\boldsymbol{\theta}^{(1)}, \ldots, \boldsymbol{\theta}^{(k_{\max})}$ samples from uniform
  distribution in $\mathcal{U}$
Set $k = 1$ and $J_{\text{OPT}} = \infty$
while $k \leq k_{\max}$ **do**
  Compute $\boldsymbol{\theta}_{\min} = \Upsilon(\boldsymbol{\theta}^{(k)})$ with Υ applied to (14)
  if $J_{\text{ISOC}}(\boldsymbol{\theta}_{\min}, \hat{m}_t, \hat{\Omega}_t^{x^*}) < J_{\text{OPT}}$ **then**
    $\boldsymbol{\theta}^{\triangle} = \boldsymbol{\theta}_{\min}$
    $J_{\text{OPT}} = J_{\text{ISOC}}(\boldsymbol{\theta}_{\min}, \hat{m}_t, \hat{\Omega}_t^{x^*})$
  $k \leftarrow k + 1$
return $\boldsymbol{\theta}^{\triangle}$, $J_{\text{ISOC}}(\boldsymbol{\theta}^{\triangle}, \hat{m}_t, \hat{\Omega}_t^{x^*})$

(LICQ) holds $\forall \boldsymbol{\theta} \in \mathcal{U}$.

Proof: Since $\mathcal{U}$ is a hyperrectangle, all constraints are given by $\theta_i - a_i \geq 0$ and $-\theta_i + b_i \geq 0$ ($i \in \{1, \ldots, \Theta\}$). For the set of their gradients $\{\boldsymbol{e}_i, -\boldsymbol{e}_i, i \in \{1, \ldots, \Theta\}\}$ ($\boldsymbol{e}_i$ as i-th standard basis vector of $\mathbb{R}^{\Theta}$) follows. Now, since $b_i > a_i$, the linear dependent vectors $\boldsymbol{e}_i$ and $-\boldsymbol{e}_i$ cannot correspond to active constraints at the same point. Thus, from the independence of $\boldsymbol{e}_i$ and $\boldsymbol{e}_j$ ($\forall i, j \in \{1, \ldots, \Theta\}$ with $i \neq j$), we conclude the LICQ $\forall \boldsymbol{\theta} \in \mathcal{U}$. ∎

Suppose a constrained local optimization solver Υ according to Definition 4 is given, we can find a global optimizer $\boldsymbol{\theta}^{\triangle}$ of (14) by starting it from $k_{\max}$ uniformly sampled points in $\mathcal{U}$. This leads to Algorithm 1 and Theorem 2.

Theorem 2: Let Assumption 4 hold and let a feasible set $\mathcal{U}$ (cf. Definition 2) as well as a constrained local optimization solver Υ according to Definition 4 be given. Algorithm 1 yields $\boldsymbol{\theta}^{\triangle} \in \mathcal{T}$ (cf. Definition 3) if (at least) one sample point $\boldsymbol{\theta}^{(k)} \in \Theta_{\text{RoA}}(\mathcal{T})$.

Proof: From Definition 4, we have $\boldsymbol{\theta}^{\triangle} = \Upsilon(\boldsymbol{\theta}^{(k)})$ with $\boldsymbol{\theta}^{(k)} \in \Theta_{\text{RoA}}(\mathcal{T})$. ∎

The probability of choosing a uniformly sampled point $\boldsymbol{\theta}^{(k)} \in \Theta_{\text{RoA}}(\mathcal{T})$ is given by [22]

$$P(\boldsymbol{\theta}^{(k)} \in \Theta_{\text{RoA}}(\mathcal{T})) = 1 - \left(1 - \frac{\text{vol}(\Theta_{\text{RoA}}(\mathcal{T}))}{\text{vol}(\mathcal{U})}\right)^{k_{\max}}, \quad (17)$$

where $\text{vol}(\cdot)$ denotes the volume of the corresponding space. Hence, since $\mathcal{T} \neq \emptyset$ follows from Lemma 2, an appropriate local solver needs to guarantee $\text{vol}(\Theta_{\text{RoA}}(\mathcal{T})) \neq 0$. Then, $k_{\max} \to \infty$ guarantees $P(\boldsymbol{\theta}^{(k)} \in \Theta_{\text{RoA}}(\mathcal{T})) \to 1$. Starting a local solver from every sample point $\boldsymbol{\theta}^{(k)}$ results in impractical computation times. Therefore, we include two filters to determine if a local solver is started at the current sample point $\boldsymbol{\theta}^{(k)}$. First, a distance filter is implemented motivated by [19]. A local solver is only started from $\boldsymbol{\theta}^{(k)}$ if the distance

$$\delta^{(k)} = \begin{cases} \delta_c^{(k)} & \text{if } \exists i < k : J_{\text{ISOC}}(\boldsymbol{\theta}^{(i)}) < J_{\text{ISOC}}(\boldsymbol{\theta}^{(k)}) \\ \infty & \text{otherwise} \end{cases}, \quad (18)$$

where $\delta_c^{(k)} = \min_{i<k}\{\|\boldsymbol{\theta}^{(k)} - \boldsymbol{\theta}^{(i)}\|_2 : J_{\text{ISOC}}(\boldsymbol{\theta}^{(i)}) < J_{\text{ISOC}}(\boldsymbol{\theta}^{(k)})\}$, to previously checked sample points $\boldsymbol{\theta}^{(i)}$ ($i \in$

Algorithm 2: TRLwARoA ISOC Algorithm.

Input: $\hat{m}_t$, $\hat{\Omega}_t^{x^*}$, $\mathcal{U}$, $k_{\max}$, γ, v
Output: $\boldsymbol{\theta}^{\triangle}$, $J_{\text{ISOC}}(\boldsymbol{\theta}^{\triangle}, \hat{m}_t, \hat{\Omega}_t^{x^*})$
Draw $\boldsymbol{\theta}^{(1)}, \ldots, \boldsymbol{\theta}^{(k_{\max})}$ samples from uniform
  distribution in $\mathcal{U}$
Set $k = 1$, $l = 0$ and $J_{\text{OPT}} = \infty$
Compute α (19)
while $k \leq k_{\max}$ **do**
  Compute $\delta^{(k)}$ (18)
  if $\delta^{(k)} > \alpha \wedge \left\|\boldsymbol{\theta}^{(k)} - \boldsymbol{\theta}_{\min}^{(i)}\right\|_2 \geq v\delta_{\text{RoA}}^{(i)}, \forall i \in$
    $\{1, \ldots, l\}$ **then**
    $l \leftarrow l + 1$
    Compute $\boldsymbol{\theta}_{\min}^{(l)} = \Upsilon(\boldsymbol{\theta}^{(k)})$ with Υ to (14)
    $\delta_{\text{RoA}}^{(l)} = \left\|\boldsymbol{\theta}_{\min}^{(l)} - \boldsymbol{\theta}^{(k)}\right\|_2$
    if $J_{\text{ISOC}}(\boldsymbol{\theta}_{\min}^{(l)}, \hat{m}_t, \hat{\Omega}_t^{x^*}) < J_{\text{OPT}}$ **then**
      $\boldsymbol{\theta}^{\triangle} = \boldsymbol{\theta}_{\min}^{(l)}$
      $J_{\text{OPT}} = J_{\text{ISOC}}(\boldsymbol{\theta}_{\min}^{(l)}, \hat{m}_t, \hat{\Omega}_t^{x^*})$
  $k \leftarrow k + 1$
return $\boldsymbol{\theta}^{\triangle}$, $J_{\text{ISOC}}(\boldsymbol{\theta}^{\triangle}, \hat{m}_t, \hat{\Omega}_t^{x^*})$

$\{1, \ldots, k-1\}$) is greater than a threshold α. Sample points with $J_{\text{ISOC}}(\boldsymbol{\theta}^{(k)}) \leq J_{\text{ISOC}}(\boldsymbol{\theta}^{(i)})$ ($\forall i \in \{1, \ldots, k-1\}$) are always used as starting points ($\delta^{(k)} = \infty$). With the distance filter, a sample point density in $\mathcal{U}$ independent from the size of $\mathcal{U}$ is maintained by adapting α to the size of $\mathcal{U}$. Suppose $\mathcal{U}$ is divided into hyperspheres $\mathcal{S}$ with radius α. Then, α is chosen such that the number of hyperspheres fitting into $\mathcal{U}$ is constant: $\text{const} = \frac{\text{vol}(\mathcal{U})}{\text{vol}(\mathcal{S})}$. This leads to

$$\text{const} = \frac{\Gamma\left(\frac{\Theta}{2} + 1\right)\text{vol}(\mathcal{U})}{\pi^{\frac{\Theta}{2}}\alpha^{\Theta}}$$

$$\Rightarrow \alpha = \gamma\left(\Gamma\left(\frac{\Theta}{2} + 1\right)\prod_{i=1}^{\Theta}(b_i - a_i)\right)^{\frac{1}{\Theta}}, \quad (19)$$

where $\gamma = \pi^{-\frac{1}{2}}\text{const}^{-\frac{1}{\Theta}}$ is a tuning parameter for the distance filter and Γ denotes the gamma function.

The second filter to further reduce the number of local solver starts is a RoA filter. If a local minimum is found, i.e. $\boldsymbol{\theta}_{\min}^{(l)} = \Upsilon(\boldsymbol{\theta}^{(k)})$, the RoA $\Theta_{\text{RoA}}(\boldsymbol{\theta}_{\min}^{(l)})$ is approximated by a hypersphere $\{\boldsymbol{\theta} \in \mathcal{U} : \left\|\boldsymbol{\theta} - \boldsymbol{\theta}_{\min}^{(l)}\right\|_2 < v\delta_{\text{RoA}}^{(l)}\}$, where $\delta_{\text{RoA}}^{(l)} = \left\|\boldsymbol{\theta}^{(k)} - \boldsymbol{\theta}_{\min}^{(l)}\right\|_2$ and v denotes a tuning parameter guaranteeing Assumption 5.

Overall, our TRLwARoA is given by Algorithm 2 and in Theorem 3, its global convergence is shown.

Definition 5: Let $\mathcal{M}^{J_{\text{ISOC}}}(m) = \{\boldsymbol{\theta} \in \mathcal{U} : J_{\text{ISOC}}(\boldsymbol{\theta}) \leq m\}$ with $m \in [-1, \infty)$ be the m-level set of J_{ISOC}. The set $\mathcal{M}^{J_{\text{ISOC}}}(m^*)$ is the optimal level set where m^* is the highest value such that $\mathcal{T}$ are the only local minima in $\mathcal{M}^{J_{\text{ISOC}}}(m^*)$.

Assumption 5: For all $\boldsymbol{\theta}_{\min}^{(l)}$ found by Algorithm 2 $\{\boldsymbol{\theta} \in \mathcal{U} : \left\|\boldsymbol{\theta} - \boldsymbol{\theta}_{\min}^{(l)}\right\|_2 < v\delta_{\text{RoA}}^{(l)}\} \subseteq \Theta_{\text{RoA}}(\boldsymbol{\theta}_{\min}^{(l)})$.

Theorem 3: Let Assumptions 4 and 5 hold and let a feasible set $\mathcal{U}$ (cf. Definition 2) as well as a constrained local optimization solver Υ according to Definition 4 be given. Algorithm 2 yields $\boldsymbol{\theta}^{\triangle} \in \mathcal{T}$ (cf. Definition 3) if (at least) one sample point $\boldsymbol{\theta}^{(k)} \in \Theta_{\mathrm{RoA}}(\mathcal{T}) \cap \mathcal{M}^{J_{\mathrm{ISOC}}}(m^*)$.

Proof: We need to verify that $\boldsymbol{\theta}^{(k)} \in \Theta_{\mathrm{RoA}}(\mathcal{T}) \cap \mathcal{M}^{J_{\mathrm{ISOC}}}(m^*)$ is not filtered by the distance and RoA filter. Then, $\boldsymbol{\theta}^{\triangle} \in \mathcal{T}$ follows from Definition 4. First, since $\boldsymbol{\theta}^{(k)}$ cannot be in the RoA of different local minima, $\boldsymbol{\theta}^{(k)}$ is not filtered by the RoA filter due to Assumption 5 and $\boldsymbol{\theta}^{(k)} \in \Theta_{\mathrm{RoA}}(\mathcal{T})$. Furthermore, $\boldsymbol{\theta}^{(k)}$ is not filtered by the distance filter since $\forall \boldsymbol{\theta} \in \mathcal{M}^{J_{\mathrm{ISOC}}}(m^*)$ $J_{\mathrm{ISOC}}(\boldsymbol{\theta}) \leq J_{\mathrm{ISOC}}(\boldsymbol{\theta}^{(i)})$ with $\boldsymbol{\theta}^{(i)} \notin \mathcal{M}^{J_{\mathrm{ISOC}}}(m^*)$ and uniformly sampled in $\mathcal{U}$. ∎

Remark 3: Typically, $\mathcal{M}^{J_{\mathrm{ISOC}}}(m^*) \subseteq \Theta_{\mathrm{RoA}}(\mathcal{T})$[7]. Hence, by reducing the number of local solver starts with the distance filter, the requirement on the sampling procedure (at least one sample point $\boldsymbol{\theta}^{(k)} \in \Theta_{\mathrm{RoA}}(\mathcal{T}) \cap \mathcal{M}^{J_{\mathrm{ISOC}}}(m^*)$) gets more restrictive. However, by lowering γ this more restrictive assumption can be relaxed and for $\gamma \to 0$ $\boldsymbol{\theta}^{(k)} \in \Theta_{\mathrm{RoA}}(\mathcal{T})$ is sufficient (cf. Theorem 2). Thus, in case of problems with small $\mathrm{vol}\left(\mathcal{M}^{J_{\mathrm{ISOC}}}(m^*)\right)$ γ can be used to maintain convergence under the best possible performance improvement of the distance filter.

Remark 4: Theorem 2 and 3 show that the main requirement on the local solver Υ (cf. Definition 4) is to guarantee $\mathrm{vol}\left(\Theta_{\mathrm{RoA}}(\mathcal{T})\right) \neq 0$, which can be motivated by the convergence of e.g. basic IP methods to KKT points of (14) (see before). Note that the algorithms do not depend on $\mathrm{vol}\left(\Theta_{\mathrm{RoA}}(\Theta_{\min})\right) \neq 0$ for every $\Theta_{\min}$. Furthermore, most implementations of constrained local optimization solvers will even guarantee convergence to local minima, e.g. due to corrections on Hessian approximations, which is beneficial for Algorithm 2 due to the RoA filter.

C. Lower Level Optimization

The lower level of our bi-level-based ISOC algorithm is given by the constraints (14c) and (14d) of (14) and for their computation Assumption 2 is needed.

In general, the assumption on a linear non-adaptive filter in Lemma 1 leads to suboptimal solutions to SOC Problem 1. However, exact solutions are still an open question. They exist for the fully observable case[8] ($\boldsymbol{y}_t = \boldsymbol{x}_t$) [8], [23] but in case of partial state information only suboptimal solutions were proposed (see e.g. [8], [13]–[15]).

IV. NUMERICAL EXAMPLE

After introducing an example system and the implementation of Algorithms 1 and 2 in Subsection IV-A, in Subsection IV-B, Assumptions 3 and 4 are validated numerically. Furthermore, Lemma 1 is compared to other (suboptimal) control strategies for Problem 1 and we show that $\mathrm{vol}\left(\Omega_{\mathrm{RoA}}(\mathcal{T})\right) \neq 0$ for our used IP implementation. Finally, in Subsection IV-C and IV-D the new TRLwARoA

---

[7]From Lemma 2, $\mathcal{T} \neq \emptyset$ and due to the assumed continuity of $J_{\mathrm{ISOC}}(\boldsymbol{\theta})$ (cf. Assumption 4), we have $\mathrm{vol}\left(\mathcal{M}^{J_{\mathrm{ISOC}}}(m^*)\right) \neq 0$

[8]The setup of (14) for the fully observable case is straightforward and all our algorithms and statements can be applied as well.

ISOC algorithm is analyzed regarding its tuning parameters and compared to State-of-the-Art (SoA) algorithms.

A. Example System and Implementation

Our simulation example is given by a planar point-to-point human hand movement, where the human hand is modeled as point mass. The exact definition can be found in [10]. The GT parameters $\boldsymbol{\theta}^*$ are given by: $\boldsymbol{\theta}^* = \begin{bmatrix} \boldsymbol{s}^{*\mathsf{T}} & \boldsymbol{\sigma}^{*\mathsf{T}} \end{bmatrix}^{\mathsf{T}}$ with $\boldsymbol{s}^* = \begin{bmatrix} 1 & 1 & 0.04 & 0.04 & 0.0004 & 0.0004 & \frac{1}{42}10^{-5} & \frac{1}{42}10^{-5} \end{bmatrix}^{\mathsf{T}}$ and $\boldsymbol{\sigma}^* = \begin{bmatrix} \mathbf{0}_{1\times8} & 0.02 & 0.02 & 0.2 & 0.2 & 1 & 1 & 0.5 & 0.1 \end{bmatrix}^{\mathsf{T}}$. From the GT parameters, $\hat{\boldsymbol{m}}_t$ and $\hat{\boldsymbol{\Omega}}_t^{\boldsymbol{x}^*}$ follow according to Lemma 1 and Theorem 1. Hence, due to the simulation scenario Assumptions 1 and 2 are fulfilled. Lastly, the lower bounds of the feasible set $\mathcal{U}$ are defined by $a_i = 0$ ($\forall i \in \{1, \ldots, 6, 9, \ldots, 16, 23, 24\}$) and $a_i = 10^{-10}$ ($\forall i \in \{7, 8, 17, \ldots, 22\}$) and $\boldsymbol{w}_{\mathrm{m}} = \begin{bmatrix} 0.9 & 0.9 & 0.9 & 0.9 \end{bmatrix}^{\mathsf{T}}$ and $\boldsymbol{w}_{\mathrm{v}} = \begin{bmatrix} 0.1 & \mathbf{0}_{1\times4} & 0.1 & \mathbf{0}_{1\times4} & 0.1 & \mathbf{0}_{1\times4} & 0.1 \end{bmatrix}^{\mathsf{T}}$ are used to set up (14).

The implementation of the algorithms was done in Matlab 2021b on a standard PC with a Ryzen 9 5950X. As local optimization solver Υ, the IP method of the `fmincon` solver of the Matlab environment is used. First and second order derivatives are numerically approximated. Both ISOC algorithms were implemented via parallel computing on the 16 cores available on the CPU (scalable to more cores) by starting local optimizations on these cores in parallel.

B. Numerical Validation of the Assumptions

First, we motivate the validity of Assumption 4. Hereto, we compute the 276 projections of $J_{\mathrm{ISOC}}(\boldsymbol{\theta})$ on the θ_i-θ_j-planes ($i, j \in \{1, \ldots, 24\}$, $j > i$). The value of $J_{\mathrm{ISOC}}(\boldsymbol{\theta})$ is calculated $\forall \boldsymbol{\theta} = \begin{bmatrix} \ldots & \theta_{i-1}^* & \theta_i & \theta_{i+1}^* & \ldots & \theta_{j-1}^* & \theta_j & \theta_{j+1}^* & \ldots \end{bmatrix}^{\mathsf{T}}$ with $\theta_{i/j} = \lambda \theta_{i/j}^*$ and λ chosen in 101 equidistant steps from $[0.5, 1.5]$ if $\theta_{i/j}^* \neq 0$. If $\theta_{i/j}^* = 0$, 101 equidistant steps for $\theta_{i/j}$ from $[0, 1]$ are chosen. Finally, an evaluation of $J_{\mathrm{ISOC}}(\boldsymbol{\theta})$ at equidistant grid points for every θ_i-θ_j-plane follows (see red crosses in Fig. 1a for θ_1-θ_2-plane). Now, by fitting 2D polynomials of order five to the grid points of $J_{\mathrm{ISOC}}(\boldsymbol{\theta})$ for every θ_i-θ_j-plane, we achieve R^2 values of > 0.999 ($R^2 = 0.99986$ for θ_1-θ_2-plane in Fig. 1a, minimal value of all polynomial fits). The quantitative and qualitative results (see surface plot in Fig. 1a) of this fitting procedure and the approximated shape of $J_{\mathrm{ISOC}}(\boldsymbol{\theta})$ through the evaluation at the grid points (the results for the other θ_i-θ_j-planes look very similar) provide strong indications that $J_{\mathrm{ISOC}}(\boldsymbol{\theta})$ is an analytical function like the polynomial fits. Thus, $J_{\mathrm{ISOC}}(\boldsymbol{\theta})$ can be assumed to be (at least) twice continuously differentiable (cf. Assumption 4). Fig. 1b depicts the 1D projection of $J_{\mathrm{ISOC}}(\boldsymbol{\theta})$ at $\boldsymbol{\theta} = \boldsymbol{\theta}^* + \lambda(\boldsymbol{\theta}_{\min}^{(l)} - \boldsymbol{\theta}^*)$, where $\boldsymbol{\theta}_{\min}^{(l)}$ denotes a local minimum. It clearly shows the assumed non-convexity of $J_{\mathrm{ISOC}}(\boldsymbol{\theta})$ (cf. Assumption 3).

Next, we look at $\Theta_{\mathrm{RoA}}(\mathcal{T})$. By applying Algorithm 1 with $k_{\max} = 20000$ and $b_i = 2$ ($\forall i \in \{1, \ldots, 24\}$) to the GT data, 692 local solver runs converge to a $\boldsymbol{\theta}^{\triangle} \in \mathcal{T}$. Hence, $\mathrm{vol}\left(\Theta_{\mathrm{RoA}}(\mathcal{T})\right) \neq 0$ follows for the used IP implementation
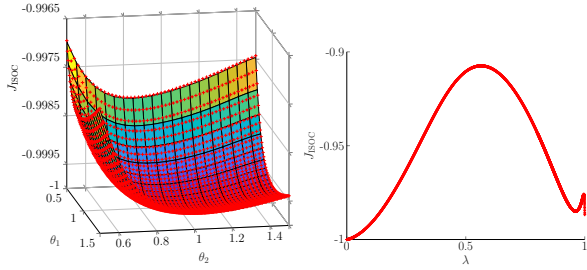
(a) Evaluation of $J_{\text{ISOC}}(\boldsymbol{\theta})$ at equidistant grid points in $\boldsymbol{\theta}$ = θ_1-θ_2-plane together with surface of polynomial fit.

(b) Evaluation of $J_{\text{ISOC}}(\boldsymbol{\theta})$ at $\boldsymbol{\theta} = \boldsymbol{\theta}^* + \lambda(\boldsymbol{\theta}_{\min}^{(l)} - \boldsymbol{\theta}^*)$.

Fig. 1: Validation of Assumptions 3 and 4.

TABLE I: Mean, maximal and minimal relative change $\triangle J$ of performance criterion (3) with different control strategies. Strategy of Lemma 1 (non-adaptive $\boldsymbol{K}_t$) serves as the reference value.

	Lemma 1 with adaptive $\boldsymbol{K}_t$ [8] vs. non-adaptive $\boldsymbol{K}_t$	[14] vs. Lemma 1 with non-adaptive $\boldsymbol{K}_t$
$\triangle J^{\text{mean}}$	$-1.1\,\%$	$1.4\,\%$
$\triangle J^{\text{max}}$	$2.9\,\%$	$5.7\,\%$
$\triangle J^{\text{min}}$	$-7.0\,\%$	$-3.8\,\%$

and furthermore, we approximate $\frac{\text{vol}(\Theta_{\text{RoA}}(\mathcal{T}))}{\text{vol}(\mathcal{U})} \approx \frac{692}{20000} = 0.0346$ and have with $k_{\max} = 10000$ $P(\boldsymbol{\theta}^{(k)} \in \Theta_{\text{RoA}}(\mathcal{T})) \approx 1$ according to (17).

Finally, following the discussion on the suboptimality of the solution to the lower level optimization in Subsection III-C, Table I shows the comparison of two other suboptimal control strategies to the one of Lemma 1. The comparison is performed for 20 parameters $\boldsymbol{\theta}$: $\boldsymbol{\theta}^*$ and $\boldsymbol{\theta}$ with $\theta_i = \lambda\theta_i^*$ (λ randomly from $[0.5, 1.5]$) if $\theta_i^* \neq 0$, θ_i randomly from $[0, 1]$ otherwise ($\forall i \in \{1, \ldots, 24\}$). Both strategies can lead to an improvement as well as a worsening of J with negligible relative change in average. Hence, Lemma 1 is the best choice since it furthermore enables the use of Theorem 1 due to its non-adaptiveness.

C. Analysis of the Tuning Parameters of Algorithm 2

In the following, we analyze the convergence behavior and the computation time of our new TRLwARoA ISOC Algorithm 2 w.r.t. to different choices of its tuning parameters, i.e. upper bounds b_i of feasible set $\mathcal{U}$, tuning parameters γ of the distance filter and v of the RoA filter. We set $k_{\max} = 10000$ since it is sufficient for $P(\boldsymbol{\theta}^{(k)} \in \Theta_{\text{RoA}}(\mathcal{T})) \approx 1$ (see Subsection IV-B). Table II shows the parameter combinations evaluated, where b_i is chosen equally $\forall i \in \{1, \ldots, 24\}$. With every tuning parameter combination, Algorithm 2 was executed 10 times. As evaluation metrics, the average computation time $t_{\text{comp}}^{\text{mean}}$, the average number of started local solvers $\#\Upsilon^{\text{mean}}$ and the worst objective function value $J_{\text{ISOC}}^{\text{max}}$ achieved during these 10 runs are used. Furthermore, we define a run of Algorithm 2 as converged to a global optimizer if for the achieved J_{ISOC} value $J_{\text{ISOC}} \leq$

TABLE II: Tuning parameter analysis for Algorithm 2.

			$\gamma = 0.6$	$\gamma = 0.7$	$\gamma = 0.8$
$v = 0$	$b_i = 2$	$\#\Upsilon^{\text{mean}}$	719.8	96.3	24.6
		$\#l$	**10**	**9**	**7**
		$t_{\text{comp}}^{\text{mean}}$	258.2 min	42.1 min	24.1 min
	$b_i = 20$	$\#\Upsilon^{\text{mean}}$	680.7	86.4	24.2
		$\#l$	**10**	**8**	**3**
		$t_{\text{comp}}^{\text{mean}}$	313.2 min	78.2 min	48.9 min
$v = 0.7$	$b_i = 2$	$\#\Upsilon^{\text{mean}}$	386.9	69.1	20.2
		$\#l$	**10**	**9**	**5**
		$t_{\text{comp}}^{\text{mean}}$	137.0 min	36.8 min	14.3 min
	$b_i = 20$	$\#\Upsilon^{\text{mean}}$	419.8	67.3	20.5
		$\#l$	**10**	**7**	**1**
		$t_{\text{comp}}^{\text{mean}}$	173.9 min	39.2 min	15.6 min
$v = 0.9$	$b_i = 2$	$\#\Upsilon^{\text{mean}}$	33.3	18.1	9.1
		$\#l$	**3**	**1**	**1**
		$t_{\text{comp}}^{\text{mean}}$	15.8 min	14.4 min	9.1 min
	$b_i = 20$	$\#\Upsilon^{\text{mean}}$	33.4	13.3	9.4
		$\#l$	**0**	**1**	**2**
		$t_{\text{comp}}^{\text{mean}}$	17.9 min	9.4 min	8.0 min

-0.999 holds. The number of converged runs is denoted as $\#l$. Table II shows $\#\Upsilon^{\text{mean}}$, $\#l$ and $t_{\text{comp}}^{\text{mean}}$ for the evaluated tuning parameter combinations. With $v = 0$ Assumption 5 is guaranteed and for $\gamma = 0.6$ all 10 runs converge. By increasing v to $v = 0.7$, Assumption 5 can still be considered as fulfilled ($\#l = 10$) and due to the RoA filter, performance is improved significantly: $\#\Upsilon^{\text{mean}}$ and $t_{\text{comp}}^{\text{mean}}$ are nearly halved[9]. Now, by increasing γ to $\gamma = 0.7$, $\#\Upsilon^{\text{mean}}$ and $t_{\text{comp}}^{\text{mean}}$ are reduced by more than $80\,\%$ and $70\,\%$, respectively (for $v = 0$ and $v = 0.7$). However, with $\gamma = 0.7$ convergence does only occur in $7-9$ of the 10 runs, but a local minimum with an objective function value close to the global one is always achieved in the not-converged runs: $J_{\text{ISOC}}^{\text{max}} < -0.98$. Due to these near global minima, $\text{vol}\left(\mathcal{M}^{J_{\text{ISOC}}}(m^*)\right)$ is small in our case and we need to choose γ according to Remark 3, sufficiently small but as large as possible. For $\gamma = 0.8$ or $v = 0.9$, $\#l \leq 5$ except for $\gamma = 0.8$, $v = 0$ and $b_i = 2$ with $\#l = 7$. Hence, in these cases either the RoA filter ($v = 0.9$) or the distance filter ($\gamma = 0.8$) hinders convergence due to not fulfilled Assumption 5 or Remark 3. Noticeably, for $\gamma = 0.8$ with $v = 0$ or $v = 0.7$ and for $\gamma = 0.6$ with $v = 0.9$, still $J_{\text{ISOC}}^{\text{max}} < -0.98$. Overall, the best tuning parameters for our example system are $v \leq 0.7$ and $\gamma = 0.6 - 0.7$ independent from the choice of upper bounds b_i ($\forall i \in \{1, \ldots, 24\}$).

D. Comparison to State-of-the-Art-Algorithms

Table III shows the comparison of the TRLwARoA ISOC algorithm to different SoA methods. First, the results of our grid-search (GS) and bi-level-based ISOC algorithm [10] are included and second, we designed an additional new bi-level-based ISOC algorithm where the upper level optimization problem is solved by the Matlab implementation of the algorithm proposed in [24]. Here, promising starting points for local solver runs in $\mathcal{U}$ are heuristically determined by a scatter search (ScS) algorithm. We set $k_{\max} = 10000$, $\gamma =$

---

[9]Mean number of local solver starts $\#\Upsilon^{\text{mean}}$ is nearly independent from the choice of bounds due to the adaptive threshold α (19).

TABLE III: Comparison of the TRLwARoA algorithm to the grid-search-based (GS) method [10] and the algorithm based on the scatter search (ScS) procedure in [24].

	TRLwARoA	ScS	GS [10]
$\#l$	**10**	**8**	−
$J_{\mathrm{ISOC}}^{\max}$	−0.9993	−0.9986	−0.9987
$t_{\mathrm{comp}}^{\mathrm{mean}}$	36.6 min	114.2 min	1194.0 min
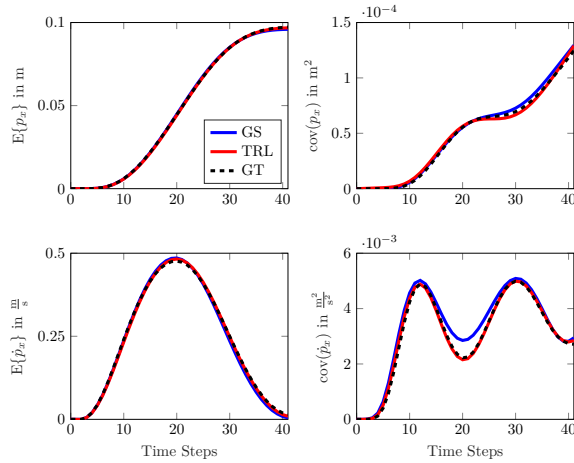
Fig. 2: Mean and covariance of position p_x and velocity $\dot{p}_x$ of the human hand in the example system. Values achieved with GT parameters and parameters identified with the TRLwARoA algorithm and the method in [10] are shown.

0.7 and $v = 0$ for the TRLwARoA and $k_{\max} = 5000$ (best performance of ScS) for the ScS algorithm. The parameters for the GS-based method as well as the upper bounds of the feasible set $\mathcal{U}$ are the same as in [10]. For the stochastic algorithms, TRLwARoA and ScS, 10 runs were performed. Table III highlights that only the TRLwARoA algorithm yields convergence in all 10 runs. Moreover, it computes ISOC solutions ca. 3 and 33 times faster than the ScS and GS, respectively. Finally, Fig. 2 illustrates the match with the GT data achieved by the parameters $\boldsymbol{\theta}^{\triangle}$ identified with the TRLwARoA algorithm ($\boldsymbol{\theta}^{\triangle}$ with $J_{\mathrm{ISOC}}^{\max}$ chosen).

V. CONCLUSION

In this paper, we present a new algorithm to solve the ISOC problem for the LQS control model. The ISOC problem consists of determining cost function and noise scaling parameters from measurement (GT) data. Both parameter types influence the model predictions (statistical moments of the system quantities) of the LQS model. We overcome convergence problems and high computation times of our previous method [10] by proving global convergence and achieving computation times that are nearly 33 times smaller. Based on simulations, the tuning parameters of our new algorithm are analyzed and shown to be intuitively tuneable. Furthermore, the theoretical assumptions are validated numerically. In the future, we are going to apply our newly developed algorithm to real human measurement data.

REFERENCES

[1] T. L. Molloy, J. Inga, S. Hohmann, and T. Perez, *Inverse Optimal Control and Inverse Noncooperative Dynamic Game Theory. A Minimum Principle Approach.* Cham: Springer Nature, 2022.

[2] E. Todorov and M. I. Jordan, "Optimal feedback control as a theory of motor coordination," *Nat. Neurosci.*, vol. 5, no. 11, pp. 1226–1235, 2002.

[3] Y. Uno, M. Kawato, and R. Suzuki, "Formation and control of optimal trajectory in human multijoint arm movement," *Biol. Cybern.*, vol. 61, pp. 89–101, 1989.

[4] W. Jin, D. Kulic, J. F.-S. Lin, S. Mou, and S. Hirche, "Inverse optimal control for multiphase cost functions," *IEEE Trans. Rob.*, vol. 35, no. 6, pp. 1387–1398, 2019.

[5] O. S. Oguz, Z. Zhou, S. Glasauer, and D. Wollherr, "An inverse optimal control approach to explain human arm reaching control based on multiple internal models," *Sci. Rep.*, vol. 8, no. 1, 2018.

[6] B. Berret, E. Chiovetto, F. Nori, and T. Pozzo, "Evidence for composite cost functions in arm movement planning: an inverse optimal control approach," *PLoS Comput. Biol.*, vol. 7, no. 10, 2011.

[7] J. P. Gallivan, C. S. Chapman, D. M. Wolpert, and J. R. Flanagan, "Decision-making in sensorimotor control," *Nat. Rev. Neurosci.*, vol. 19, no. 9, pp. 519–534, 2018.

[8] E. Todorov, "Stochastic optimal control and estimation methods adapted to the noise characteristics of the sensorimotor system," *Neural Comput.*, vol. 17, pp. 1084–1108, 2005.

[9] P. Karg, S. Stoll, S. Rothfuß, and S. Hohmann, "Validation of stochastic optimal control models for goal-directed human movements on the example of human driving behavior," *IFAC World Congress 2023*, 2023.

[10] ——, "Inverse stochastic optimal control for linear-quadratic gaussian and linear-quadratic sensorimotor control models," *61st IEEE Conf. Decis. Control*, 2023.

[11] M. C. Priess, J. Choi, and C. Radcliffe, "The inverse problem of continuous-time linear quadratic gaussian control with application to biological systems analysis," *ASME 2014 Dynamic Systems and Control Conference*, 2014.

[12] S. Kolekar, W. Mugge, and D. Abbink, "Modeling intradriver steering variability based on sensorimotor control theories," *IEEE Trans. Hum.-Mach. Syst.*, vol. 48, no. 3, pp. 291–303, 2018.

[13] X. Liang, A. Zhang, Z. Liu, Y. Du, and J. Qiu, "Suboptimal linear output feedback control of discrete-time systems with multiplicative noises," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 70, no. 4, pp. 1480–1484, 2023.

[14] J. B. Moore, X. Y. Zhou, and A. E. B. Lim, "Discrete time lqg controls with control dependent noise," *Syst. Control Lett.*, vol. 36, pp. 199–206, 1999.

[15] S. Joshi, "On optimal control of linear systems in the presence of multiplicative noise," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 12, no. 1, pp. 80–85, 1976.

[16] K. J. Aström, *Introduction to Stochastic Control Theory.* New York: Academic Press, Inc., 1970.

[17] J. C. Bezdek and R. J. Hathaway, "Some notes on alternating optimization," *AFSS International Conference on Fuzzy Systems*, 2002.

[18] K. Mombaur, A. Truong, and J.-P. Laumond, "From human to humanoid locomotion—an inverse optimal control approach," *Auton. Robots*, vol. 28, no. 3, pp. 369–383, 2010.

[19] M. Locatelli and F. Schoen, "Random linkage: a family of acceptance/rejection algorithms for global optimisation," *Math. Program.*, vol. 85, no. 2, pp. 379–396, 1999.

[20] A. H. G. Rinnooy Kan and G. T. Timmer, "Stochastic global optimization methods part ii: Multi level methods," *Math. Program.*, vol. 39, no. 1, pp. 57–78, 1987.

[21] J. Nocedal and S. J. Wright, *Numerical Optimization*, 2nd ed. New York: Springer, 2006.

[22] A. H. G. Rinnooy Kan and G. T. Timmer, "Stochastic global optimization methods part i: Clustering methods," *Math. Program.*, vol. 39, no. 1, pp. 27–56, 1987.

[23] A. Beghi, "Discrete-time optimal control with control-dependent noise and generalized riccati difference equations," *Automatica*, vol. 34, no. 8, pp. 1031–1034, 1998.

[24] Z. Ugray, L. Lasdon, J. Plummer, F. Glover, J. Kelly, and R. Martí, "Scatter search and local nlp solvers: A multistart framework for global optimization," *INFORMS J. Comput.*, vol. 19, no. 3, pp. 328–340, 2007.