

Optimality and Suboptimality of MPPI Control in Stochastic and Deterministic Settings

Hannes Homburger¹, Florian Messerer², Moritz Diehl³, *Member, IEEE*,
and Johannes Reuter⁴, *Member, IEEE*

Abstract—Model predictive path integral (MPPI) control has recently received a lot of attention, especially in the robotics and reinforcement learning communities. This letter aims to make the MPPI control framework more accessible to the optimal control community. We present three classes of optimal control problems and their solutions by MPPI. Further, we investigate the suboptimality of MPPI to general deterministic nonlinear discrete-time systems. Here, suboptimality is defined as the deviation between the control provided by MPPI and the optimal solution to the deterministic optimal control problem. Our findings are that in a smooth and unconstrained setting, the growth of suboptimality in the control input trajectory is second-order with the scaling of the MPPI exploration uncertainty. The results indicate that the suboptimality of the MPPI solution can be modulated by tuning the hyperparameters. We illustrate our findings using numerical examples.

Index Terms—Optimization, predictive control for nonlinear systems, stochastic optimal control.

I. INTRODUCTION AND OVERVIEW

MODEL predictive control (MPC) is an optimization-based state-feedback control technique that computes the control inputs concerning the state trajectory predicted over a finite horizon as the solution of an optimal control problem (OCP). To solve an OCP, often numerical methods are used that sequentially generate local approximations of the problem and then solve them efficiently [1]. These approaches lead to very fast convergence rates and can be applied even to high-dimensional problems. However, they can converge to local minima and have issues in the presence of strong nonlinearities. In contrast, zero-order solvers based on sampling have fewer restrictions on the problem and can be executed in parallel, but often lead to higher computational

effort by ignoring the first- or second-order information. Recently, a sample-based method called model predictive path integral (MPPI) control has been investigated and applied in numerous publications, e.g., see the survey [2]. Initially, this method, inspired by the path integral framework, was used to solve a special class of stochastic closed-loop OCPs [3], [4]. Later, it was extended based on the *information-theoretic* framework to general nonlinear systems with input noise – however, based on an open-loop problem formulation [5]. The certainty-equivalence principle [6] states that the solution for linear systems with independent noise and costs is independent of the uncertainty. However, in the literature, the uncertainty-aware MPPI method is often used for deterministic OCPs, and the suboptimality of the solution is tacitly ignored, see, e.g., [7], [8], [9], [10]. In this context, noise is intentionally introduced for exploration, representing a fundamental distinction from classical stochastic optimal control, where noise arises from the environment. Recently, an MPPI method for deterministic OCPs was presented [11].

A. Contribution and Outline

We review the MPPI control framework from a stochastic optimal control perspective and show that the problem classes considered in the MPPI literature are stochastically equivalent to input-affine stochastic systems with noise on the input channel. Subsequently, we investigate the suboptimality of standard MPPI to deterministic OCPs. We prove that for smooth and unconstrained problems, the suboptimality of the control trajectory is $\mathcal{O}(\beta^2)$ and the suboptimality of the value function is $\mathcal{O}(\beta^4)$, where $\beta \in \mathbb{R}^+$ is a factor that scales the standard deviation of the MPPI sampling distribution. Structure of this letter: in Section II, we define different classes of stochastic OCPs and discuss their solution via MPPI in Section III. The suboptimality is investigated in Section IV. The algorithm and a characterization are presented in Section V. Section VI presents illustrative examples, followed by a summarizing Section VII.

B. Notation and Preliminaries

The set of positive real numbers is $\mathbb{R}^+ = \{x \in \mathbb{R} \mid x > 0\}$ and the set of non-negative real numbers is $\mathbb{R}_0^+ = \{x \in \mathbb{R} \mid x \geq 0\}$. The concatenation of two column vectors $x \in \mathbb{R}^{n_x}$ and $y \in \mathbb{R}^{n_y}$ is denoted by $(x, y) := [x^\top, y^\top]^\top \in \mathbb{R}^{n_x+n_y}$. The partial derivative with respect to a variable x is represented as $\frac{\partial}{\partial x}$, indicating differentiation with respect to the explicit argument of the function. Arguments of a function may be specified using subscripts or parentheses. For instance, the function $f_\beta(x)$ has two arguments, β and

Received 5 March 2025; revised 30 April 2025; accepted 16 May 2025. Date of publication 27 May 2025; date of current version 24 June 2025. This work was supported in part by the Deutsche Forschungsgemeinschaft (DFG) under Project 504452366 (SPP 2364) and Project 525018088, and in part by the Bundesministerium für Wirtschaft und Klimaschutz (BMWK) under Grant 03EI4057A and Grant 03EN3054B. Recommended by Senior Editor S. Olaru. (Corresponding author: Hannes Homburger.)

Hannes Homburger and Johannes Reuter are with the Institute of System Dynamics, HTWG Konstanz–University of Applied Sciences, 78462 Konstanz, Germany (e-mail: hhomburg@htwg-konstanz.de).

Florian Messerer is with the Department of Microsystems Engineering, University of Freiburg, 79110 Freiburg, Germany.

Moritz Diehl is with the Department of Microsystems Engineering and the Department of Mathematics, University of Freiburg, 79104 Freiburg, Germany.

Digital Object Identifier 10.1109/LCSYS.2025.3574151

x . Gradients of functions are denoted by $\nabla_x f(x) = \frac{\partial}{\partial x} f(x)^\top$, representing the vector of partial derivatives with respect to x and $\nabla_x^2 f(x)$ denotes the Hessian. The multivariate normal distribution of a random variable $v \in \mathbb{R}^n$ is expressed as $v \sim \mathcal{N}(\mu, \Sigma)$, where $\mu \in \mathbb{R}^n$ is the mean, and $\Sigma \in \mathbb{R}^{n \times n}$ is the covariance matrix. The expectation of a random variable $W \in \mathbb{R}^n$ sampled from $\mathbb{Q}_\beta$ defined by its probability density function $q_\beta(W)$, is

$$\mathbb{E}_{W \sim \mathbb{Q}_\beta}[W] = \int_{\mathbb{R}^n} W q_\beta(W) dW.$$

For a function $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$, we write $f(x) = \mathcal{O}(g(x))$ if and only if there exists a constant $C \in \mathbb{R}^+$ and a neighborhood $\mathcal{A}$ of 0 such that $\forall x \in \mathcal{A}: \|f(x)\| \leq Cg(x)$.

II. STOCHASTIC OPTIMAL CONTROL PROBLEMS

We consider the stochastic nonlinear discrete-time system

$$x^+ = f(x, u, w),$$

where the state is $x \in \mathbb{R}^{n_x}$, the control is $u \in \mathbb{R}^{n_u}$, the random disturbance is $w \in \mathbb{R}^{n_w}$, $w \sim \mathcal{N}(0, I)$, i.i.d., and without loss of generality unit matrix I as covariance matrix. Given an initial state $\bar{x}_0 \in \mathbb{R}^{n_x}$, an input trajectory $U = (u_0, u_1, \dots, u_{N-1})$, and a noise trajectory $W = (w_0, w_1, \dots, w_{N-1})$, the state trajectory is determined by

$$x_0(U, W, \bar{x}_0) = \bar{x}_0, \quad (1)$$

$$x_{k+1}(U, W, \bar{x}_0) = f(x_k(U, W, \bar{x}_0), u_k, w_k), \quad (2)$$

using forward simulation for $k = 0, 1, \dots, N-1$. The performance of a trajectory can be evaluated by the overall cost

$$J_{\bar{x}_0}(U, W) := \sum_{k=0}^{N-1} L(x_k(\cdot), u_k) + E(x_N(\cdot)), \quad (3)$$

where the states' dependencies given in (1) and (2) are omitted for readability, the stage cost function is $L: \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \rightarrow \mathbb{R}$, and $E: \mathbb{R}^{n_x} \rightarrow \mathbb{R}$ is the terminal cost function.

A. Closed-Loop Stochastic Optimal Control

The ideal problem is to minimize the expected overall cost for the closed-loop system. This means that each control u_k is selected with respect to the state $x_k(\cdot)$. The key point is that in $x_k(\cdot)$ all previous realizations of disturbances w_i and control actions u_i for $i = 0, \dots, k-1$ are included. This problem can be formalized in the recursive form

$$V_{\text{cls}}(\bar{x}_0) = \min_{u_0} \mathbb{E}_{w_0 \sim \mathcal{N}(0, \Sigma)} \dots \min_{u_{N-1}} \mathbb{E}_{w_{N-1} \sim \mathcal{N}(0, \Sigma)} J_{\bar{x}_0}(U, W), \quad (4)$$

that we call the *closed-loop stochastic optimal control problem* (CLS-OC). Note that this problem typically corresponds to optimization over policies.

B. Open-Loop Stochastic Optimal Control

Because the CLS-OC is typically hard to solve, it is sometimes reasonable to optimize the expected overall cost for the open-loop system instead. This means that the control inputs u_k are chosen independently of the previous realization of the disturbances, but in an optimal way concerning the prior known distribution of the disturbances. This problem is

$$V_{\text{ols}}(\bar{x}_0) = \min_U \mathbb{E}_{W \sim \mathcal{N}(0, \bar{\Sigma})} [J_{\bar{x}_0}(U, W)], \quad (5)$$

that we call the *open-loop stochastic optimal control problem* (OLS-OC), where $\bar{\Sigma} = \text{diag}(\Sigma, \Sigma, \dots, \Sigma) \in \mathbb{R}^{N n_w \times N n_w}$ is the mutual covariance matrix for a whole trajectory W with uncorrelated w_k . Usually, the solution to this problem requires less effort because the optimization concerns only one input trajectory. However, when applied in a receding horizon fashion, the controller's capability to react to realized disturbances is not modeled, leading to a suboptimal solution for the closed-loop system.

C. Deterministic Optimal Control

In this letter, we investigate the suboptimality of MPPI for solving the deterministic OCP given by

$$V_{\text{det}}(\bar{x}_0) = \min_U J_{\bar{x}_0}(U, 0), \quad (6)$$

where no uncertainty is considered. We call this problem the *deterministic optimal control problem* (DET-OC). If the computation power is limited or if there is no model of the model's uncertainty, practitioners often use this *nominal* setting. Note that in a smooth and unconstrained setting the DET-OC solution and the CLS-OC solution are similar for systems with small uncertainties [12].

III. MPPI SOLUTION METHODS

In this section, we review the MPPI methods for the OCPs stated in the previous section, and we show that these methods are inherently restricted to the overall cost structure

$$J_{\bar{x}_0}(U, W) = S_{\bar{x}_0}(U + W) + \frac{1}{2} \|U\|_2^2, \quad (7)$$

where the *path costs* are

$$S_{\bar{x}_0}(V) := E(x_N(V, 0, \bar{x}_0)) + \sum_{k=0}^{N-1} c(x_k(V, 0, \bar{x}_0)) \quad (8)$$

with $c: \mathbb{R}^{n_x} \rightarrow \mathbb{R}_0^+$ and $V := U + W$. Thus, the stage cost in (3) is restricted to the sum of a generic term dependent only on the state and a quadratic control cost.

A. MPPI Control for the CLS-OC

We consider the class of CLS-OCs, where the discrete-time system dynamics is restricted to input-affine systems

$$x_{\text{aff}}^+ = \check{f}(x) + B(x)u + G(x)w, \quad (9)$$

with $B: \mathbb{R}^{n_x} \rightarrow \mathbb{R}^{n_x \times n_u}$, $G: \mathbb{R}^{n_x} \rightarrow \mathbb{R}^{n_x \times n_w}$, $w \sim \mathcal{N}(0, I)$, and the stage cost

$$L(x, u) = c(x) + \frac{1}{2} u^\top R(x) u \quad (10)$$

with $R: \mathbb{R}^{n_x} \rightarrow \mathbb{R}^{n_u \times n_u}$, $R \succ 0$. The main theoretical results are derived under the following strong assumption.

Assumption 1: There exists a $\lambda \in \mathbb{R}^+$ such that

$$\lambda B(x)R(x)^{-1}B(x)^\top = G(x)G(x)^\top. \quad (11)$$

We show that this assumption can be interpreted as effectively requiring that the noise enters additively on the control. For the continuous time setting, a similar result has been obtained in [13, Sec. 3.3].

Proposition 1: Consider the CLS-OC (4) with input-affine dynamics (9), $w \sim \mathcal{N}(0, I)$, i.i.d., stage cost (10), and let Assumption 1 hold. Then, an equivalent CLS-OC exists with

dynamics $x^+ = \tilde{f}(x) + \tilde{B}(x)(\tilde{u} + \tilde{w})$, and stage cost $\tilde{L}(x, \tilde{u}) = c(x) + \frac{1}{2}\tilde{u}^\top \tilde{u}$, where $\tilde{w} \sim \mathcal{N}(0, \lambda I)$.

Proof: The solution of the CLS-OCF (4) can be characterized by the dynamic programming (DP) [14] recursion

$$V_k(x) = \min_u \left\{ c(x) + \frac{1}{2}u^\top R(x)u + \mathbb{E}_{w \sim \mathcal{N}(0, I)} [V_{k+1}(x_{\text{aff}}^+)] \right\},$$

$$k = N - 1, \dots, 0, \text{ with } V_N(\cdot) = E(\cdot) \text{ and } V_0(\cdot) = V_{\text{cls}}(\cdot).$$

Substituting $B(x) = \tilde{B}(x)R(x)^{\frac{1}{2}}$ and $u = R(x)^{-\frac{1}{2}}\tilde{u}$ yields

$$V_k(x) = \min_{\tilde{u}} \left\{ c(x) + \frac{1}{2}\tilde{u}^\top \tilde{u} + \mathbb{E}_w \{ V_{k+1}(\tilde{f}(x) + \tilde{B}(x)\tilde{u} + G(x)w) \} \right\}.$$

Applying the substitution of $B(x)$ also to (11), we obtain $\lambda \tilde{B}(x)\tilde{B}(x)^\top = G(x)G(x)^\top$. This allows us to write

$$\begin{aligned} \mathbb{E}_{w \sim \mathcal{N}(0, I)} \{ G(x)w \} &= \mathbb{E}_{v \sim \mathcal{N}(0, G(x)G(x)^\top)} \{ v \} \\ &= \mathbb{E}_{v \sim \mathcal{N}(0, \lambda \tilde{B}(x)\tilde{B}(x)^\top)} \{ v \} = \mathbb{E}_{\tilde{w} \sim \mathcal{N}(0, \lambda I)} \{ \tilde{B}(x)\tilde{w} \}. \end{aligned}$$

Using an analogous argument for the DP recursion yields

$$V_k(x) = \min_{\tilde{u}} \left\{ c(x) + \frac{1}{2}\tilde{u}^\top \tilde{u} + \mathbb{E}_{\tilde{w}} \left\{ V_{k+1}(\tilde{f}(x) + \tilde{B}(x)(\tilde{u} + \tilde{w})) \right\} \right\}.$$

This is exactly the DP recursion for the equivalent CLS-OCF as specified in the proposition statement. ■

The obtained dynamics are considered as a special case in [5, Sec. IV], with the solution

$$V_{\text{cls}}(\bar{x}_0) = -\lambda \log \mathbb{E}_{W \sim \mathcal{N}(0, \bar{\Sigma})} \left[\exp \left(-\frac{1}{\lambda} S_{\bar{x}_0}(W) \right) \right], \quad (12)$$

and according to [5, Eq. (28)] the optimal control for $k = 0$ is given by

$$u_{0, \text{cls}}^*(\bar{x}_0) = \frac{\mathbb{E}_{W \sim \mathcal{N}(0, \bar{\Sigma})} \left[\exp \left(-\frac{1}{\lambda} S_{\bar{x}_0}(W) \right) w_0 \right]}{\mathbb{E}_{W \sim \mathcal{N}(0, \bar{\Sigma})} \left[\exp \left(-\frac{1}{\lambda} S_{\bar{x}_0}(W) \right) \right]}. \quad (13)$$

Note the similarity between (13) and the *softargmin* function, which is well known in the field of machine learning. The expression (12) is derived from an exponential transformation of the stochastic Hamilton-Jacobi-Bellman equation [3], providing a closed form solution as an expectation via the Feynman-Kac Lemma [15, Th. 8.2.1]. Note that these theoretical foundations are based on a continuous-time setting [13]. As implicitly proposed in [5, Sec. IV], the Euler-Maruyama discretization method finally provides (12) and (13).

B. MPPI Control for the OLS-OCF

Now we consider a more general class of nonlinear discrete-time systems with additive noise on the input, given by

$$x^+ = f(x, u + w, 0), \quad w \sim \mathcal{N}(0, \Sigma), \quad (14)$$

quadratic costs on the inputs (10), and assumption $\lambda R^{-1} = \Sigma$, now constant. According to [5, Sec. 3.A], the optimal control sequence U_{ols}^* of the OLS-OCF (5) can be approximated by

$$U_{\text{ols}}^* \approx \tilde{U}_{\text{MPPI}}^* := \argmin_U \mathbb{E}_{W \sim \mathbb{Q}^*} \left[\log \left(\frac{q^*(W)}{q_U(W)} \right) \right], \quad (15)$$

sometimes called *information-theoretic* optimum. Here, $q_U(W)$ is the probability density function (PDF) of the normal

distribution $\mathcal{N}(U, \bar{\Sigma})$ and the *optimal distribution* $\mathbb{Q}^*$ is defined by its PDF

$$q^*(W) := \frac{1}{\eta} \exp \left(-\frac{1}{\lambda} S_{\bar{x}_0}(W) \right) p_{\bar{\Sigma}}(W), \quad (16)$$

where $\eta \in \mathbb{R}^+$ is a normalization factor and $p_{\bar{\Sigma}}(W)$ is the PDF of the normal distribution $\mathcal{N}(0, \bar{\Sigma})$. Note that the objective in (15) is the *Kullback-Leibler divergence* [16, Ch. 3] of the optimal distribution $\mathbb{Q}^*$ and the controlled distribution $\mathbb{Q}_U$. In contrast to the presentation in [5, Eq. (3) and (16)], as postulated in (15), MPPI only provides an approximation to the OLS-OCF solution, although the *optimal distribution* [5, Sec. 3] establishes a lower bound on the value function of the OLS-OCF [17]. This follows from the fact that in general, the *optimal distribution* cannot be reached by $\mathbb{Q}_U$. Consequently, the relationship between U_{ols}^* and $\tilde{U}_{\text{MPPI}}^*$ should be interpreted as a heuristic, and it is not clear how suboptimal the relation is. The minimizer of (15) is given by

$$\tilde{u}_{k, \text{MPPI}}^*(\bar{x}_0) = \frac{\mathbb{E}_{W \sim \mathcal{N}(0, \bar{\Sigma})} \left[\exp \left(-\frac{1}{\lambda} S_{\bar{x}_0}(W) \right) w_k \right]}{\mathbb{E}_{W \sim \mathcal{N}(0, \bar{\Sigma})} \left[\exp \left(-\frac{1}{\lambda} S_{\bar{x}_0}(W) \right) \right]} \quad (17)$$

for $k = 0, 1, \dots, N - 1$ [5, Sec. 3.C]. In contrast to the presentation in [5, Sec. 4], the equivalence of (13) and (17) for $k = 0$ implies an interpretation of (15) as an approximation of the CLS-OCF. Here, suboptimality presumably depends on a metric of how strongly the nonlinear dynamics (14) deviate from the input-affine case (9).

C. MPPI Control for the D-OCF

However, there is a special case in which MPPI via the *information-theoretic* framework converges to an exact solution: consider the general nonlinear and deterministic system $x^+ = f(x, u, 0)$ and stage costs given in the form of (10) containing quadratic costs on the control inputs and arbitrary state-dependent costs. The optimal deterministic input trajectory is given by

$$U_{\text{det}}^*(\bar{x}_0) = \argmin_U J_{\bar{x}_0}(U, 0) \quad (18)$$

and the k^{th} element of this solution is denoted by $u_{k, \text{det}}^*(\bar{x}_0)$. In MPPI, the system dynamics covariance $\bar{\Sigma}$ employed during the sampling process. To enable the solution of the DET-OCF within the MPPI framework, this covariance must be reduced. Accordingly, we modify equation (17) by introducing a scaling factor $\beta \in \mathbb{R}^+$, and replace λ and $\bar{\Sigma}$ with $\beta^2 \lambda$ and $\beta^2 \bar{\Sigma}$, respectively. Then we obtain

$$\begin{aligned} \tilde{u}_{k, \text{MPPI}}^*(\bar{x}_0, \beta) &= \frac{\mathbb{E}_{W \sim \mathcal{N}(0, \beta^2 \bar{\Sigma})} \left[\exp \left(-\frac{1}{\beta^2 \lambda} S_{\bar{x}_0}(W) \right) w_k \right]}{\mathbb{E}_{W \sim \mathcal{N}(0, \beta^2 \bar{\Sigma})} \left[\exp \left(-\frac{1}{\beta^2 \lambda} S_{\bar{x}_0}(W) \right) \right]}, \\ &= \mathbb{E}_{W \sim \mathbb{Q}^*(\beta)} [w_k], \end{aligned} \quad (19)$$

and in the limit of shrinking uncertainty this expression with

$$u_{k, \text{det}}^*(\bar{x}_0) = \lim_{\beta \rightarrow 0} \tilde{u}_{k, \text{MPPI}}^*(\bar{x}_0, \beta) \quad (20)$$

provides an exact solution to the DET-OCF for $k = 0, \dots, N - 1$. Note that the noise is intentionally introduced for exploration and is not part of the environment. The statement (20) is proven in Thm. 1 in Section IV.

IV. INVESTIGATION OF SUBOPTIMALITY

In this section, we present and prove Thm. 1, the principal result of this letter. It addresses the convergence of the deterministic MPPI method and establishes the order of suboptimality. Now we briefly introduce Erdélyi's formulation of Laplace's classical method that will be crucial later.

Lemma 1 (Adapted from [18, Th. 1.1]): For the integral

$$I(\lambda) = \int_a^b \exp(-\lambda f(x)) g(x) dx,$$

where $(a, b) \subseteq \mathbb{R}$ is a real interval, which may be finite or infinite, we assume that

- (i) $f(x) > f(a)$ for all $x \in (a, b)$, for every $\delta > 0$ the infimum of $f(x) - f(a)$ in $[a + \delta, b)$ is positive;
- (ii) f and g are scalar real analytic functions; and
- (iii) the integral $I(\lambda)$ converges absolutely for all sufficiently large λ .

Then

$$I(\lambda) \equiv \exp(-\lambda f(a)) \sum_{n=0}^{\infty} \tilde{c}_n \lambda^{-\frac{n}{2}},$$

where the coefficients $\tilde{c}_n$ are real values and the symbol $\equiv$ denotes an asymptotic expansion for $\lambda \rightarrow +\infty$.

Theorem 1: Assume $J_{\bar{x}_0} : \mathbb{R}^{Nn_u} \times \mathbb{R}^{Nn_u} \rightarrow \mathbb{R}_0^+$ is continuously differentiable, has a unique global minimum $J_{\bar{x}_0}(U_{\det}^*, 0)$, and second-order sufficient condition (SOSC) at $U_{\det}^*$. Then

$$\lim_{\beta \rightarrow 0} \tilde{U}_{\text{MPPI}}^*(\beta) = U_{\det}^* \quad (21)$$

and for small $\beta > 0$ the bias of the solution shrinks quadratically with the uncertainty,

$$\|\tilde{U}_{\text{MPPI}}^*(\beta) - U_{\det}^*\| = \mathcal{O}(\beta^2). \quad (22)$$

Proof: We start with the **scalar case** $Nn_u = 1$ and then extend this to the multivariate case. Replacing λ and $\bar{\Sigma}$ by their scaled versions, we can rewrite (16) as

$$\begin{aligned} q_\beta^*(W) &:= \frac{1}{\eta(\beta)} \exp\left(-\frac{1}{\beta^2 \lambda} S_{\bar{x}_0}(W)\right) p_{\beta^2 \bar{\Sigma}}(W), \\ &= \frac{\exp\left(-\frac{1}{\beta^2 \lambda} \left[S_{\bar{x}_0}(W) + \frac{\lambda}{2} \sum_{k=0}^{N-1} w_k^\top \Sigma^{-1} w_k\right]\right)}{\eta(\beta) Z(\beta)}, \\ &= \frac{1}{\tilde{\eta}(\beta)} \exp\left(-\frac{1}{\beta^2 \lambda} J_{\bar{x}_0}(W, 0)\right), \end{aligned} \quad (23)$$

where in the second line, the exponential kernel of the multivariate normal distribution is included in the exponent, and its normalization factor is denoted by $Z(\beta)$. In the third line, the term in square brackets is identified as the overall costs, and the normalization factors are combined to $\tilde{\eta}(\beta)$. Based on this representation of the PDF of the optimal distribution, the MPPI solution (19) is given by

$$\begin{aligned} \tilde{U}_{\text{MPPI}}^*(\beta) &= \mathbb{E}_{W \sim \mathbb{Q}^*(\beta)}[W] \\ &= \frac{\int_{\mathbb{R}} W \exp\left(-\frac{1}{\beta^2 \lambda} J_{\bar{x}_0}(W, 0)\right) dW}{\int_{\mathbb{R}} \exp\left(-\frac{1}{\beta^2 \lambda} J_{\bar{x}_0}(W, 0)\right) dW}, \end{aligned} \quad (24)$$

where the second line is the explicit representation of the expected value, where $q_\beta^*(W)$ is substituted. For small $\beta > 0$, both integrals are dominated by the neighborhood of $W = U_{\det}^*$, motivating to apply Laplace's method, since the integrand

becomes sharply peaked around this point, making the approximation accurate. Substituting $\beta^2 = \zeta$ in (24) and applying Laplace's method from Lemma 1 separately to the numerator and the denominator result in

$$\tilde{U}_{\text{MPPI}}^*(\zeta) = \frac{U_{\det}^* \alpha_1 \zeta^{\frac{1}{2}} + \alpha_2 \zeta^{\frac{3}{2}} + \mathcal{O}(\zeta^{\frac{5}{2}})}{\alpha_1 \zeta^{\frac{1}{2}} + \alpha_3 \zeta^{\frac{3}{2}} + \mathcal{O}(\zeta^{\frac{5}{2}})}, \quad (25)$$

where the structure follows from SOSC at $U_{\det}^*$ and $\alpha_1, \alpha_2, \alpha_3$ are scalar constants. The re-substitution of $\zeta = \beta^2$ and a subsequent Taylor expansion of (25) at $\beta = 0$ yields

$$\|\tilde{U}_{\text{MPPI}}^*(\beta) - U_{\det}^*\| = \frac{|\alpha_2 - U_{\det}^* \alpha_3|}{|\alpha_1|} \beta^2 + \mathcal{O}(\beta^4) = \mathcal{O}(\beta^2).$$

In the **multivariate case** of (24) the MPPI solution for all scalar elements of the solution trajectory $i = 0, \dots, \hat{N} - 1$ with $\hat{N} := Nn_u$ is given by

$$\begin{aligned} \tilde{U}_{\text{MPPI},i}^*(\beta) &= \mathbb{E}_{w_i \sim \mathbb{Q}^*(\beta)}[w_i] \\ &= \frac{\int_{\mathbb{R}} \cdots \int_{\mathbb{R}} \int_{\mathbb{R}} w_i \Xi(W, \beta) dw_0 dw_1, \dots, dw_{\hat{N}-1}}{\int_{\mathbb{R}} \cdots \int_{\mathbb{R}} \int_{\mathbb{R}} \Xi(W, \beta) dw_0 dw_1, \dots, dw_{\hat{N}-1}} \\ &= \frac{\int_{\mathbb{R}} w_i \hat{\Xi}(w_i, \beta) dw_i}{\int_{\mathbb{R}} 1 \hat{\Xi}(w_i, \beta) dw_i}, \end{aligned} \quad (26)$$

where $\Xi(W, \beta) := \exp(-\frac{1}{\beta^2 \lambda} J_{\bar{x}_0}(W, 0))$ and $\hat{\Xi}(w_i, \beta)$ is the marginal PDF for $i = 0, \dots, \hat{N} - 1$. The marginal PDF (26) is similar to (24) that we already treated via Laplace's Method. The same argumentation results in

$$\|\tilde{U}_{\text{MPPI},i}^*(\beta) - U_{\det,i}^*\| = \gamma_i \beta^2 + \mathcal{O}(\beta^4) = \mathcal{O}(\beta^2)$$

that holds for all scalar elements of input trajectory with index $i = 0, 1, \dots, \hat{N} - 1$ with $\gamma_i \in \mathbb{R}_0^+$ and

$$\|\tilde{U}_{\text{MPPI}}^*(\beta) - U_{\det}^*\| = \mathcal{O}(\beta^2) \quad (27)$$

follows. ■

Corollary 1: The corresponding optimal value function is reached in the limit and shrinks with fourth-order,

$$J_{\bar{x}_0}(\tilde{U}_{\text{MPPI}}^*(\beta), 0) - V_{\det}(\bar{x}_0) = \mathcal{O}(\beta^4). \quad (28)$$

Proof: Due to the continuously differentiable setting, the continuous differentiability of the optimal value function is preserved [12]. Therefore, we can use the Taylor series

$$\begin{aligned} J_{\bar{x}_0}(U, 0) - V_{\det}(\bar{x}_0) &= \frac{1}{2} \Delta U^\top \nabla_U^2 J_{\bar{x}_0}(U_{\det}^*, 0) \Delta U \\ &\quad + \mathcal{O}(\|\Delta U\|^3), \end{aligned} \quad (29)$$

where $\Delta U := U - U_{\det}^*$ denotes the difference to the deterministic solution trajectory. Since, with (22), the error of the computed control trajectory shrinks with $\mathcal{O}(\beta^2)$, and with (29) there exists a local quadratic model of the overall cost function, the uncertainty manifests itself with fourth-order in the value function. ■

Algorithm 1 Deterministic MPPI Solving (20)

Require: Initial system state $\bar{x}_0$, initial control sequence $\hat{U}$, number of iterations K , number of trajectory samples M , shrinking factor ν , initial covariance $\bar{\Sigma}_0$, initial temperature λ_0

- 1: **for** $j \in \{0, 1, \dots, K-1\}$ **do**
- 2: $\beta \leftarrow \nu^j$; $\lambda \leftarrow \beta^2 \lambda_0$; $\bar{\Sigma} \leftarrow \beta^2 \bar{\Sigma}_0$ $\triangleright$ *reducing uncertainty*
- 3: **for** $m \in \{0, 1, \dots, M-1\}$ **do** $\triangleright$ *in parallel*
- 4: $W^m \sim \mathcal{N}(0, \bar{\Sigma})$
- 5: $S^m \leftarrow S_{\bar{x}_0}(\hat{U} + W^m) + \lambda W^{m\top} \bar{\Sigma}^{-1} \hat{U}$; $\triangleright$ *2nd term is correction because $\hat{U} \neq 0$; see Section V*
- 6: $\psi \leftarrow \min_{m \in \{0, 1, \dots, M-1\}} S^m$ $\triangleright$ *offset only for numerics*
- 7: $\eta \leftarrow \sum_{m=0}^{M-1} \exp(-\frac{1}{\lambda}(S^m - \psi))$
- 8: **for** $m \in \{0, 1, \dots, M-1\}$ **do** $\triangleright$ *in parallel*
- 9: $\omega^m \leftarrow \frac{1}{\eta} \exp(-\frac{1}{\lambda}(S^m - \psi))$ $\triangleright$ *compute weights*
- 10: **for** $k \in \{0, 1, \dots, N-1\}$ **do**
- 11: $\hat{u}_k \leftarrow \hat{u}_k + \sum_{m=1}^M \omega^m w_k^m$ $\triangleright$ *solution to (19)*
- 12: **return** $U_{\text{det}}^* \leftarrow \hat{U}$ $\triangleright$ *solution to (20)*

V. ALGORITHM AND CHARACTERIZATION

The expectation operator regarding the distribution of the uncontrolled system is part of the MPPI solution to the CLS-OCP (13), the OLS-OCP (17), and the DET-OCP (20). MPPI algorithms typically approximate this expectation operator using Monte Carlo estimation [2]. However, the sample efficiency using the uncontrolled distribution is strongly limited. To increase sample efficiency with *importance sampling* the sample distribution is changed to a *proposal distribution*. In the same course, a correction term must be added in the expectation to preserve its value. To change the mean of the normal distribution from 0 to $\hat{U}$ and sample from $\mathcal{N}(\hat{U}, \bar{\Sigma})$ instead of $\mathcal{N}(0, \bar{\Sigma})$, a correction term presented in [5, Sec. 3.C] is added to Line 5 of Algorithm 1, which is adapted from [11]. While K is the number of shrinking iterations and the final exploitation covariance is $\nu^{2K} \bar{\Sigma}_0$, the limit (21) is reached for $K \rightarrow \infty$ and, for $K = 1$, Algorithm 1 reduces to standard MPPI presented in [5] computing (13) resp. (17). It is important to note that importance sampling with this correction term does not change the value of the expectation [5, Sec. 3.C]. The suboptimality caused by the Monte Carlo estimation with a finite number of samples is investigated in [19], [20], [21] and is not part of this letter. In comparison to other **Monte Carlo optimization** approaches to solve the DET-OCP, like reward-weighted regression [22], cross-entropy methods [23], covariance matrix adaptation evolution strategy (CMA-ES) [24], MPPI-Generic [10], or CoVo-MPC [25], the deterministic MPPI method considered in this letter is characterized by the predefined shrinking rate of the sampling covariance, favorable for real-time applications [11]. However, although it is not part of this letter, it should be noted that Monte Carlo approaches often require excessive computational effort. Methods of **Newton-type optimization** can be tailored to solve DET-OCPs efficiently [1, Sec. 8]. To apply Newton-type methods on the OLS-OCP or the CLS-OCP, an approximation of the problem is typically required. This takes the form of either sampling the uncertainty vector, cf. e.g., the scenario approach [26] and the sample average approximation [21], or based on linearization, cf. [27]. For the CLS-OCP, the latter approach requires, e.g., a policy parametrization such as linear state [28] or disturbance feedback [29].

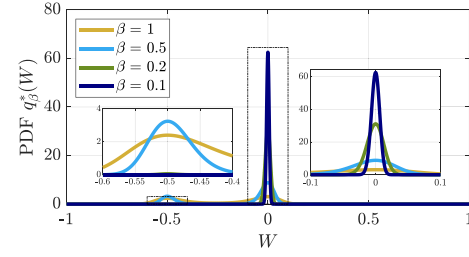


Fig. 1. PDF of optimal distribution $Q_\beta^*(W)$ for different β . The two additional interior plots show zoomed sections of the same PDF.

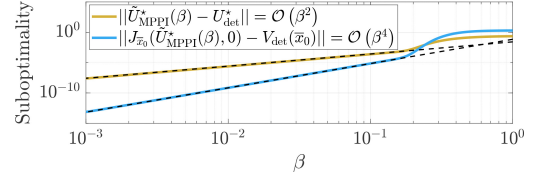


Fig. 2. Suboptimality of the MPPI solution to the DET-OCP of the controls (gold) and the value function (blue).

Since the problem (18) solved by deterministic MPPI follows a single-shooting formulation [1, Sec. 8.5.1], which is commonly addressed using Newton-type optimization methods, we compare both approaches in terms of **execution time**. The computational cost of evaluating a function F is denoted as $\text{cost}(F)$. Under assumption of ideal parallelization, due to the parallel nature of sampling, particularly when executed on a GPU [10], the execution time of Algorithm 1 is given by $\text{cost}(\text{Algorithm 1}) \approx K \cdot \text{cost}(S_{\bar{x}_0})$, as the evaluation of $S_{\bar{x}_0}$ is computationally expensive. However, we emphasize that this cost model assumes ideal parallelization, which is not achievable in practice due to non-parallelizable components, such as hardware limitations or memory access overhead. These factors can introduce additional execution time, particularly for large sample counts or limited processing units. In contrast, the execution time of computing the gradient of $S_{\bar{x}_0}$ using the reverse mode of algorithmic differentiation (AD) is $\text{cost}(\nabla_U S_{\bar{x}_0}) \approx 3 \cdot \text{cost}(S_{\bar{x}_0})$ [1, Sec. 8.4.5], usually sequentially evaluated in multiple iterations.

From Thm. (1) in combination with the exponential shrinkage of β (Algorithm 1; Line 2) a q -linear convergence follows for deterministic MPPI, while advanced Hessian approximations [1, Sec. 8.6] can lead to faster convergence rates of Newton-type methods in the neighborhood of the optimum [1, Th. 8.7]. Although MPPI is generally considered computationally inefficient, largely due to its omission of informative 1st- and 2nd-order derivatives, and in general yields suboptimal solutions for nonlinear systems due to artificial noise (Thm. 1), it possesses several notable strengths. These include inherent parallelizability, problem-agnostic applicability, robustness to poor initialization, and effectiveness in handling highly nonlinear and even non-differentiable cost functions. These properties lead to widespread adoption in the fields of robotics and reinforcement learning [2].

VI. ILLUSTRATIVE EXAMPLES

In this section, we provide two illustrative examples, while more complex applications and additional technical details are presented in our previous work [11], although without a theoretical investigation of the method's properties.

A) Consider the scalar polynomial value function $J(U) = \frac{1}{24}c_4U^4 + \frac{1}{6}c_3U^3 + \frac{1}{2}c_2U^2$ with $U \in \mathbb{R}$ and $c_2, c_3, c_4 \in \mathbb{R}$

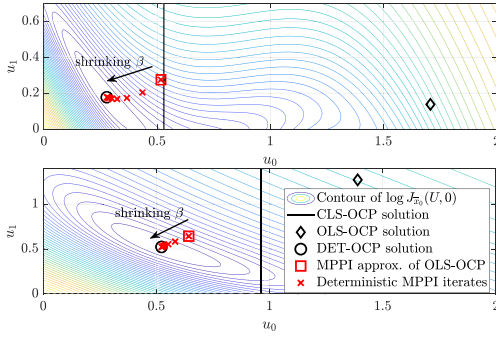


Fig. 3. Numerical solutions to the three different problems DET-OCP, OLS-OCP, and CLS-OCP (black), all with overall cost function (30) based on input-affine dynamics f_{af} (top) and nonlinear dynamics f_{nl} (bottom) alongside the iterates of *deterministic MPPI*.

selected to determine a global minimum at $J(0) = 0$, a local minimum at $J(-\frac{1}{2}) = \frac{1}{4}$, and the local maximum $J(-\frac{1}{4}) = 2$. The PDF of the optimal distribution of this function is plotted in Fig. 1 for different values of $\beta > 0$. The suboptimality of the MPPI solution is plotted as a function of β in Fig. 2. For small β , the influence of the local minimum at $W = -0.5$ vanishes in Fig. 1 and the progressions of the error of the MPPI solution show the theoretically predicted behavior.

B) Consider the input-affine dynamics $f_{af} = x - \frac{1}{2} \sin(3x) + u + w$ and the nonlinear dynamics $f_{nl} = x + \arctan(u + w)$ with $x, u, w \in \mathbb{R}$, where $w \sim \mathcal{N}(0, \beta^2 \Sigma)$, the stage cost function $L(x, u) = \frac{1}{2} Ru^2$, the asymmetric terminal cost function $E(x) = (x-1)^6 + 6x$, and a horizon with $N = 2$ steps. For both settings, the overall cost function has the structure

$$J_{\bar{x}_0}(U, W) = L(x_0(\cdot), u_0) + L(x_1(\cdot), u_1) + E(x_2(\cdot)) \quad (30)$$

and can be substituted in the CLS-OCP formulation (4), the OLS-OCP formulation (5), and the nominal DET-OCP formulation (6). Note that the CLS-OCP solution is visualized as a set because u_1 is a policy in this setting. The solutions are plotted with the $K = 10$ iterates of deterministic MPPI in Fig. 3. We choose $\bar{x}_0 = -1$, $\hat{u}_0 = \hat{u}_1 = 0$, $v = \frac{\sqrt{2}}{2}$, and an arbitrary high number of $M = 10^7$ samples. The solutions differ significantly due to the asymmetry of the problem. As considered in Section III-A, only in the case of input-affine dynamics, standard MPPI provides a reasonable solution to the CLS-OCP. In both cases, standard MPPI provides weak approximations to the OLS-OCPs; however, the deterministic MPPI iterations converge q -linearly to the D-OCP solution.

GitHub Repo: github.com/hhomb/DetMPPI

VII. SUMMARY

This letter provides a brief review of different stochastic OCP formulations and their solution methods from the MPPI control framework and simplifies their notation. Further, we show that in a smooth and unconstrained setting the suboptimality of the MPPI solution to deterministic OCP shrinks smoothly as the level of remaining exploration uncertainty decreases. This proves the convergence of the previously presented deterministic MPPI algorithm and shows that the suboptimality of the MPPI solution can be modulated by appropriately tuning hyperparameters.

ACKNOWLEDGMENT

The authors would like to thank Lothar Kiltz for asking about the convergence of deterministic MPPI and Katrin Baumgärtner for valuable discussions on Laplace's method.

REFERENCES

- [1] J. B. Rawlings et al., *Model Predictive Control: Theory, Computation, and Design*, 2nd ed. San Francisco, CA, USA: Nob Hill Publ., 2017.
- [2] M. Kazim et al., "Recent advances in path integral control for trajectory optimization: An overview in theoretical and algorithmic perspectives," *Annu. Rev. Control*, vol. 57, Jan. 2024, Art. no. 100931.
- [3] H. J. Kappen, "Path integrals and symmetry breaking for optimal control theory," *J. Statist. Mech. Theory Exp.*, vol. 2005, no. 11, 2005, Art. no. P11011.
- [4] E. A. Theodorou et al., "A generalized path integral control approach to reinforcement learning," *J. Mach. Learn. Res.*, vol. 11, pp. 3137–3181, Dec. 2010.
- [5] G. Williams et al., "Information-theoretic model predictive control: Theory and applications to autonomous driving," *IEEE Trans. Rob.*, vol. 34, no. 6, pp. 1603–1622, Dec. 2018.
- [6] B. D. O. Anderson and J. B. Moore, *Optimal Control: Linear Quadratic Methods*, Mineola, NY, USA: Dover Publ., 2007.
- [7] Y. Zhang et al., "Multi-modal MPPI and active inference for reactive task and motion planning," *IEEE Rob. Autom. Lett.*, vol. 9, no. 9, pp. 7461–7468, Sep. 2024.
- [8] H. Homburger et al., "MPPI control of a self-balancing vehicle employing subordinated control loops," in *Proc. Eur. Control Conf. (ECC)*, 2023, pp. 1–6.
- [9] L. L. Yan and S. Devasia, "Output-sampled model predictive path integral control (o-MPPI) for increased efficiency," in *Proc. IEEE Int. Conf. Rob. Autom. (ICRA)*, 2024, pp. 14279–14285.
- [10] B. Vlahov et al., "MPPI-generic: A CUDA library for stochastic optimization," 2024, *arXiv:2409.07563*.
- [11] P. Halder et al., "Trajectory planning with signal temporal logic costs using deterministic path integral optimization," in *Proc. IEEE Int. Conf. Rob. Autom. (ICRA)*, 2025, pp. 1–8.
- [12] F. Messerer et al., "Fourth-order suboptimality of nominal model predictive control in the presence of uncertainty," *IEEE Control Syst. Lett.*, vol. 8, pp. 508–513, 2024.
- [13] S. A. Thijssen, "Path integral control," Ph.D. dissertation, Radboud Univ. Nijmegen, Nijmegen, The Netherlands, 2016.
- [14] R. Bellman, *Dynamic Programming*. Princeton, NJ, USA: Princeton Univ. Press, 1957.
- [15] B. Øksendal, *Stochastic Differential Equations*. Berlin, Germany: Springer, 2003.
- [16] I. Goodfellow et al., *Deep Learning* (Adaptive Computation And Machine Learning series). Cambridge, MA, USA: MIT Press, 2016.
- [17] E. Theodorou, "Nonlinear stochastic control and information theoretic dualities: Connections, interdependencies and thermodynamic interpretations," *Entropy*, vol. 17, no. 5, pp. 3352–3375, 2015.
- [18] G. Nemes, "An explicit formula for the coefficients in laplace's method," *Construct. Approx.*, vol. 38, no. 3, pp. 471–487, 2013.
- [19] H.-J. Yoon et al., "Sampling complexity of path integral methods for trajectory optimization," in *Proc. Am. Control Conf. (ACC)*, 2022, pp. 3482–3487.
- [20] A. Patil et al., "Discrete-time stochastic LQR via path integral control and its sample complexity analysis," *IEEE Control Syst. Lett.*, vol. 8, pp. 1595–1600, 2024.
- [21] A. Shapiro et al., *Lectures on Stochastic Programming: Modelling and Theory*. Philadelphia, PA, USA: Soc. Ind. Appl. Math., 2009.
- [22] J. Peters and S. Schaal, "Reinforcement learning by reward-weighted regression for operational space control," in *Proc. Int. Conf. Mach. Learn.*, 2007, pp. 745–750.
- [23] Z. I. Botev et al., "The cross-entropy method for optimization," in *Handbook of Statistics: Machine Learning: Theory and Applications* (Handbook of Statistics), vol. 31. Amsterdam, The Netherlands: Elsevier, 2013, pp. 35–59.
- [24] N. Hansen, "The CMA evolution strategy: A comparing review," in *Towards a New Evolutionary Computation*, vol. 192. Berlin, Germany: Springer, 2006, pp. 75–102.
- [25] Z. Yi et al., "CoVO-MPC: Theoretical analysis of sampling-based MPC and optimal covariance design," in *Proc. Mach. Learn. Res.*, 2024, pp. 1–14.
- [26] G. C. Calafiore and M. C. Campi, "The scenario approach to robust control design," *IEEE Trans. Automat. Control*, vol. 51, no. 5, pp. 742–753, May 2006.
- [27] Z. Nagy and R. Braatz, "Robust nonlinear model predictive control of batch processes," *AIChE J.*, vol. 49, no. 7, pp. 1776–1786, 2003.
- [28] Z. Nagy and R. Braatz, "Open-loop and closed-loop robust optimal control of batch processes using distributional and worst-case analysis," *J. Process Control*, vol. 14, no. 4, pp. 411–422, 2004.
- [29] P. J. Goulart et al., "Optimization over state feedback policies for robust control with constraints," *Automatica*, vol. 42, no. 4, pp. 523–533, 2006.