

Risk-Aware Model Predictive Path Integral Control Using Conditional Value-at-Risk

Ji Yin, Zhiyuan Zhang, and Panagiotis Tsiotras

Abstract—In this paper, we present a novel Model Predictive Control method for autonomous robot planning and control subject to arbitrary forms of uncertainty. The proposed Risk-Aware Model Predictive Path Integral (RA-MPPI) control utilizes the Conditional Value-at-Risk (CVaR) measure to generate optimal control actions for safety-critical robotic applications. Different from most existing Stochastic MPCs and CVaR optimization methods that linearize the original dynamics and formulate control tasks as convex programs, the proposed method directly uses the original dynamics without restricting the form of the cost functions or the noise. We apply the novel RA-MPPI controller to an autonomous vehicle to perform aggressive driving maneuvers in cluttered environments. Our simulations and experiments show that the proposed RA-MPPI controller can achieve similar lap times with the baseline MPPI controller while encountering significantly fewer collisions. The proposed controller performs online computation at an update frequency of up to 80 Hz, utilizing modern Graphics Processing Units (GPUs) to multi-thread the generation of trajectories as well as the CVaR values.

I. INTRODUCTION

Model Predictive Control (MPC) is a control approach that optimizes the control input at the current time while taking the future evolution of the system states and the corresponding controls into account. Compared with more traditional approaches, MPC controllers can handle complicated interactions between the inputs and outputs while satisfying the given constraints. Moreover, the MPC's ability to predict future events based on the simulated trajectories makes it more responsive to dynamic environments than other methods. In recent years, MPC controllers have gained popularity in robotics as the massive computational power offered by modern computers allows the ability to perform fast computations on-the-fly.

MPC methods can be categorized into three types, namely, deterministic MPC (DMPC), stochastic MPC (SMPC), and robust MPC (RMPC), with increasing robustness levels. DMPC assumes that the dynamics are noise-free [1], [2], thus they lack the ability to analyze the uncertainties present in the planning phase and are especially vulnerable to unexpected disturbances. RMPC generates the safest solutions among the three types of controllers because it avoids the worst-case scenarios induced by potential disturbances, albeit at the cost of overly conservative solutions [3]. SMPC can handle dynamical systems with stochastic uncertainty subject to chance constraints [4], [5], [6]. SMPC generally produces more cost-efficient solutions compared to RMPC, and achieves more robust planning compared to the DMPC. Although SMPC

possesses a superior ability to utilize the probabilistic nature of uncertainties, most SMPC approaches suffer from several limitations when it comes to performance guarantees. First, the optimization process is usually designed for a specific form of stochastic noise [7]. As a result, robustness is greatly diminished when the environment changes and the assumptions about the form of the noise become invalid. Secondly, many SMPC approaches assume the dynamics to be linear and require linearization of the system during implementation [7], [8], [9], which slows down computation and imposes strong limitations on these controllers due to the fact that the linearized system can only capture the behavior of the real systems in a small region. Thirdly, the chance constraints used by most stochastic MPC approaches do not account for the seriousness of violating the constraints or the severity of potential accidents. In other words, chance constraints limit the probability of constraint violations, but they fail to distinguish the different costs at stake for violating the constraints.

In this paper, we propose the Risk-aware MPPI (RA-MPPI) algorithm that uses Conditional Value-at-risk (CVaR) [10] to solve the aforementioned problems of existing SMPC methods. The MPPI is a simulation-based nonlinear controller that solves optimal control problems in a receding horizon control setting. Compared with most other MPC methods, MPPI has fewer restrictions on the form of the objective functions and dynamics. Specifically, it can accept non-convex, and even gradient-free costs and nonlinear dynamics. MPPI samples thousands of control sequences by injecting Gaussian noise to a mean control sequence, then it rolls out simulated trajectories following the sampled controls. It then computes the optimal control sequence by taking the weighted average of the costs of the sampled trajectories. Using the parallel computing abilities of modern GPUs, MPPI can achieve on-line planning with a control frequency sufficient for time-critical tasks such as aggressive off-road autonomous driving [11].

Despite its attractive properties, MPPI is still a deterministic control design method, thus it does not account for dynamical noise, not to mention the risks induced by external disturbances. To alleviate this situation, Tube-MPPI [12] uses iLQG as an ancillary controller to track the nominal trajectory generated by the original MPPI. However, Tube-MPPI does not allow explicit assignment of risk levels during planning, making it difficult to specify the desired trade-off between risk and robustness. As a remedy, Robust MPPI (RMPPi) [13] improves upon Tube-MPPI by introducing an upper bound on the free-energy growth of the system to describe a task constraint satisfaction level, however, the

The authors are with the D. Guggenheim School of Aerospace Engineering, Georgia Institute of Technology, GA. E-mail: {jyin81, nickzhang, tsiotras}@gatech.edu

free-energy bound does not provide an intuitive interpretation of risk. Moreover, both Tube-MPPI and RMPPI use the linearized system dynamics to solve another ancillary optimization problem, which is problematic because linearizing the original dynamics is time-costly for real-time applications and is only accurate within a small region. RMPPI can alleviate this problem by replacing iLQG with Control Contraction Metric (CCM) [14], which is a feedback controller that provides exponential convergence guarantee for nonlinear systems [13]. However, CCM cannot be computed in real-time either. On the other hand, CVaR is a risk metric that has gained popularity in many engineering problems recently, due to its appealing ability to measure and prevent outcomes that hurt the most [15]. However, most existing CVaR optimization algorithms share the same problem with the current SMPC approaches, namely, they normally assume convex objective functions and constraints [16], [17], thereby limiting their applicability. Other approaches utilize dynamic programming to optimize time-consistent modifications of CVaR [18], [19], [20], which requires careful interpretation of their practical meanings.

Overall, the main contribution of this paper is threefold: First, we propose the novel RA-MPPI controller, which improves the baseline MPPI controller by making it account for the control risks in terms of CVaR. The novel RA-MPPI controller neither limits the form of the state-dependent costs nor restricts the type of system noise. Second, the proposed RA-MPPI utilizes a novel, general control architecture that can be easily modified to adapt various risk metrics other than CVaR. For example, it can be integrated with chance constraints [21] or some other time-consistent risk metric [22], [23], [24]. Third, we demonstrate the proposed controller by testing it on challenging autonomous racing tasks in cluttered environments. Our simulations and experiments show that the proposed RA-MPPI runs on-line at about 80 Hz using a dynamic vehicle model with tire dynamics, and it outperforms the MPPI controller in terms of collisions while maintaining approximately the same lap time, given various assigned risk levels.

II. MPPI REVIEW

Consider a general discrete nonlinear dynamical system,

$$x_{k+1} = F(x_k, u_k), \quad (1)$$

where $x_k \in \mathbb{R}^{n_x}$ is the system state and $u_k \in \mathbb{R}^{n_u}$ is the control input. Within a control horizon K , we denote the trajectory $\mathbf{x} = [x_0^\top, \dots, x_K^\top]^\top \in \mathbb{R}^{n_x(K+1)}$, the mean control sequence $\mathbf{v} = [v_0^\top, \dots, v_{K-1}^\top]^\top \in \mathbb{R}^{n_u K}$, the injected Gaussian control noise $\epsilon = [\epsilon_0^\top, \dots, \epsilon_{K-1}^\top]^\top \in \mathbb{R}^{n_u K}$, and the disturbed control sequence $\mathbf{u} = \mathbf{v} + \epsilon$.

The original Model Predictive Path Integral control (MPPI) solves the following problem,

$$\min_{\mathbf{v}} J(\mathbf{v}) = \mathbb{E} \left[\phi(x_K) + \sum_{k=0}^{K-1} \left(q(x_k) + \frac{\lambda}{2} v_k^\top \Sigma_\epsilon^{-1} v_k \right) \right], \quad (2a)$$

subject to

$$x_{k+1} = F(x_k, v_k + \epsilon_k), \quad (2b)$$

$$x_0 = x(0), \quad \epsilon_k \sim \mathcal{N}(0, \Sigma_\epsilon), \quad (2c)$$

where $x(0)$ is the measured current state. Note that the state-dependent cost $q(x_k)$ can take an arbitrary form. The MPPI controller is derived by minimizing the KL-divergence between the current controlled trajectory distribution and the optimal distribution [12] to solve the problem (2). It samples a large number of simulated trajectories to synthesize the optimal control sequence during each optimization iteration. Assuming MPPI samples trajectories $m = 0, \dots, M$ in simulation at each iteration, the cost S_m of the m^{th} sample trajectory is evaluated as,

$$S_m = \phi(x_K^m) + \sum_{k=0}^{K-1} q(x_k^m) + \gamma (v_k^m)^\top \Sigma_\epsilon^{-1} (v_k^m + \epsilon_k^m), \quad (3)$$

where $x_k^m \in \mathbb{R}^{n_x}$ is the system state, $\epsilon_k^m \in \mathbb{R}^{n_u}$ is the control noise of the m^{th} sampled trajectory, and $\gamma \in [0, \lambda]$ is the weight for the control costs. The MPPI algorithm calculates the optimal control \mathbf{v}^+ by taking the weighted average of all sampled control sequences,

$$\mathbf{v}^+ = \sum_{m=1}^M \omega_m \mathbf{u}^m / \sum_{m=1}^M \omega_m, \quad (4)$$

where $\mathbf{u}^m = \mathbf{v} + \epsilon^m$ is the control sequence corresponding to the m^{th} simulated trajectory. The corresponding weight w_m for \mathbf{u}^m is determined by,

$$\omega_m = \exp \left(-\frac{1}{\lambda} (S_m - \beta) \right), \quad (5)$$

where $\beta = \min_{m=1, \dots, M} S_m$, is the smallest trajectory cost among the M sampled trajectories, and it is used to prevent numerical overflow while keeping the solution the same [13]. The mean control \mathbf{v} for the next receding horizon control iteration is then set to be $\mathbf{v} = \mathbf{v}^+$.

III. RISK-AWARE MPPI

The MPPI problem formulation in (2) only minimizes the cost for the expected performance of a system injected with some Gaussian control noise ϵ_k , and it follows a deterministic dynamical system (1) without considering risk induced by dynamical noise explicitly in the loop. To take risks into account, we consider a more general, stochastic system,

$$\tilde{x}_{k+1} = \tilde{F}(\tilde{x}_k, u_k, w_k), \quad \tilde{x}_0 = x_0, \quad (6)$$

where $\tilde{\mathbf{x}} = [\tilde{x}_0^\top, \dots, \tilde{x}_K^\top]^\top \in \mathbb{R}^{n_x(K+1)}$ is the trajectory realization following the disturbed dynamics (6), and w_k is some noise of arbitrary form. In this work, we are interested in minimizing the objective function (2a) while preventing worst-case, catastrophic scenarios. To this end, the proposed RA-MPPI solves the following problem,

$$\min_{\mathbf{v}} J(\mathbf{v}), \text{ subject to (2b), (2c) and,} \quad (7a)$$

$$\text{CVaR}_\alpha(L(\tilde{\mathbf{x}})) \leq C_u, \quad (7b)$$

$$\tilde{x}_{k+1} = \tilde{F}(\tilde{x}_k, v_k, w_k), \quad \tilde{x}_0 = x_0, \quad (7c)$$

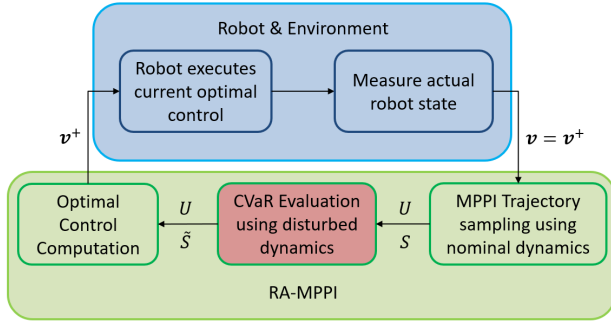


Fig. 1: The proposed RA-MPPI control architecture. The set $U = \{\mathbf{u}^0, \dots, \mathbf{u}^M\}$ contains all the sampled control sequences, and the set $S = \{S_0, \dots, S_M\}$ includes their corresponding trajectory costs. The set $\tilde{S} = \{\tilde{S}_0, \dots, \tilde{S}_M\}$ contains the modified trajectory costs which are obtained by adding the CVaR constraint violation costs evaluated using the disturbed dynamics (6) to the original MPPI sampled trajectory costs S .

where the parameter $\alpha \in (0, 1)$ is the CVaR confidence level, and C_u determines the upper bound of the CVaR constraint of the state-dependent risk cost $L(\tilde{\mathbf{x}})$ defined by (8) in Section III-A. Figure 1 illustrates the RA-MPPI control architecture. Section III-A discusses the evaluation of CVaR, and Section III-B introduces the details of applying CVaR to filter sampled trajectories such that the proposed RA-MPPI generates risk-aware controls satisfying the constraint (7b).

A. Monte-Carlo CVaR Evaluation

We define the trajectory risk cost $L(\tilde{\mathbf{x}})$ in (7b) as,

$$L(\tilde{\mathbf{x}}) = \sum_{k=0}^{K-1} \ell(\tilde{x}_k), \quad (8)$$

where $\ell(\tilde{x}_k)$ is the step risk cost. The Conditional Value-at-Risk is then defined as,

$$\text{CVaR}_\alpha(L(\tilde{\mathbf{x}})) = \mathbb{E}[L(\tilde{\mathbf{x}}) | L(\tilde{\mathbf{x}}) \geq \text{VaR}_\alpha(L(\tilde{\mathbf{x}}))], \quad (9)$$

where the value at risk (VaR) is defined as [15],

$$\text{VaR}_\alpha(L(\tilde{\mathbf{x}})) = \min\{t | \mathbb{P}(L(\tilde{\mathbf{x}}) \leq t) \geq \alpha\}, \quad (10)$$

that is, the risk cost $L(\tilde{\mathbf{x}})$ has a probability of $1 - \alpha$ to be greater than $\text{VaR}_\alpha(L(\tilde{\mathbf{x}}))$. It follows that the CVaR value (9) can be interpreted as the expected risk cost of the trajectory $\tilde{\mathbf{x}}$ among the worst $1 - \alpha$ quantile of the cost distribution $L(\tilde{\mathbf{x}})$. In addition to sampling M noise-free trajectories using the nominal system (1) as the original MPPI, the proposed RA-MPPI controller quantifies the risk of the m^{th} noise-free trajectory by simulating another N trajectories following the disturbed system (6), for $m = 0, \dots, M - 1$.

For simplicity, we drop the brackets in (9), (10), and use CVaR_α^m , VaR_α^m to denote the CVaR and VaR values for the m^{th} noise-free trajectory. To this end, CVaR_α^m can be approximated from,

$$\text{CVaR}_\alpha^m = \frac{1}{N_o} \sum_{n=0}^{N-1} \mathbf{1}_{\text{CVaR}_\alpha}(L(\tilde{\mathbf{x}}^{m,n})) L(\tilde{\mathbf{x}}^{m,n}), \quad (11)$$

where the simulated trajectories $X^m = \{\tilde{\mathbf{x}}^{m,0}, \dots, \tilde{\mathbf{x}}^{m,N-1}\}$ are sampled using the m^{th} sampled control sequence \mathbf{u}^m and

following the disturbed dynamics (6), given the initial state $\tilde{\mathbf{x}}_0^{m,n} = x_0$. In (11), N_o is the number of trajectories among X^m with cost $L(\tilde{\mathbf{x}}^{m,n}) \geq \text{VaR}_\alpha^m$, and the indicator function is defined as,

$$\mathbf{1}_{\text{CVaR}_\alpha}(L(\tilde{\mathbf{x}}^{m,n})) = \begin{cases} 1, & \text{if } L(\tilde{\mathbf{x}}^{m,n}) \geq \text{VaR}_\alpha^m, \\ 0, & \text{otherwise.} \end{cases} \quad (12)$$

In case α approaches 1, the method in [25] can be used to improve the sampling efficiency for the CVaR evaluation. The proposed control architecture is illustrated in Figure 1. Note that the red block in the figure evaluates the CVaR risk metric using the Monte-Carlo sampling method, and thus can be easily integrated with other risk metrics such as chance constraints.

B. Soft Trajectory Filtering using CVaR

The RA-MPPI filters the MPPI sampled trajectories by adding their corresponding CVaR values as penalty costs to the original MPPI trajectory costs evaluated by (3), thus penalizing MPPI sampled trajectories with high risk. Let us define the CVaR constraint violation cost J_C^m for the m^{th} noise-free trajectory,

$$J_C^m = A \text{CVaR}_\alpha^m \mathbf{1}_{C_u}(\text{CVaR}_\alpha^m), \quad (13)$$

where,

$$\mathbf{1}_{C_u}(\text{CVaR}_\alpha^m) = \begin{cases} 1, & \text{if } \text{CVaR}_\alpha^m > C_u, \\ 0, & \text{otherwise.} \end{cases} \quad (14)$$

Note that the coefficient A in (13) adjusts the magnitude of the CVaR cost, and the parameter C_u in (14) determines how strong the CVaR constraint (7b) is. Inspired by [26], we obtain the modified cost for the m^{th} trajectory by adding (13) to (3), to fulfill the constraint (7b),

$$\tilde{S}_m = J_C^m + \phi(x_K^m) + \sum_{k=0}^{K-1} q(x_k^m) + \gamma v_k^\top \Sigma_\epsilon^{-1} u_k^m. \quad (15)$$

It follows from (5) that the updated weight for the m^{th} sampled control sequence \mathbf{u}^m is,

$$\tilde{w}_m = \exp\left(-\frac{1}{\lambda}(\tilde{S}_m - \tilde{\beta})\right), \quad (16)$$

where $\tilde{\beta} = \min_{m=1, \dots, M} \tilde{S}_m$ and the resulting control sequence can be obtained by using (4).

C. Risk Cost Sensitivity Scaling

In case J_C^m is relatively small compared to \tilde{S}_m in (15), a change in the CVaR_α^m value in (13) has limited impact on the relative change of \tilde{S}_m , thus the resulting controls could become insensitive to risk variations. Increasing the coefficient A in (13) can potentially alleviate the situation, however, adjusting only the value of A may cause J_C^m to dominate (15), leading to overly conservative solutions. A simple remedy to this problem is to scale up the variance of the risk costs while maintaining their mean value. To this end, given a set of risk costs $L^m = \{L(\tilde{\mathbf{x}}^{m,0}), \dots, L(\tilde{\mathbf{x}}^{m,N-1})\}$ in (11) for the m^{th} noise-free trajectory, we can obtain the updated risk costs with adjusted variance by,

$$L_a(\tilde{\mathbf{x}}^{m,n}) = B(L(\tilde{\mathbf{x}}^{m,n}) - \bar{L}^m) + \bar{L}^m, \quad (17)$$

where \bar{L}^m is the mean of the original set L^m , given by,

$$\bar{L}^m = \sum_{n=0}^{N-1} L(\tilde{\mathbf{x}}^{m,n}), \quad (18)$$

and where B is a scaling factor. Increasing B will make the proposed RA-MPPI more sensitive to risk level changes. We can then use the adjusted risk costs L_a^m instead of L^m in (11) to evaluate CVaR_α^m .

IV. THE RISK-AWARE MPPI ALGORITHM

In this section, we introduce the RA-MPPI algorithm. In Algorithm 1, line 2 gets the current estimate of the system state x_0 . Line 4 to 15 samples the control sequences, rolls out, and evaluates the simulated trajectories following the deterministic system (1). Line 16 to line 23 samples the simulated trajectories used for evaluating the CVaR values following the disturbed dynamics (6), then line 24 to line 29 compute CVaR values and carry out trajectory filtering. More specifically, line 8 produces sampled controls by adding control noise to mean controls. Line 10 introduces some zero-mean controls to help smooth the resulting controls [11]. Line 24 scales up the sensitivity of the risk costs following (17) and (18). Line 25 finds the set P that contains the upper $1 - \alpha$ quantile of the risk cost set $\{L(\tilde{\mathbf{x}}^{m,0}), \dots, L(\tilde{\mathbf{x}}^{m,N-1})\}$, and line 26 calculates CVaR using (11). Line 28 adds the CVaR constraint (7b) violation costs to the trajectory costs S_m to obtain the modified trajectory costs \tilde{S}_m using (13) and (15). Based on the trajectories $\tilde{S} = \{\tilde{S}_0, \dots, \tilde{S}_M\}$ and the control sequence set $U = \{\mathbf{u}^0, \dots, \mathbf{u}^M\}$, line 31 computes the optimal control sequence \mathbf{v}^+ using (5) and (4). Line 32 sends the first control v_0^+ to the actuators, and line 33 sets the optimal controls to be the mean controls for the next optimization iteration.

V. SIMULATION AND EXPERIMENTS

In this section, we demonstrate the proposed RA-MPPI controller and compare it with the baseline MPPI controller on an autonomous racing platform. We show that the novel RA-MPPI significantly improves robustness against disturbances in challenging autonomous driving tasks by using a variety of simulation and experiment examples. All the simulation and experiment results are obtained by running the controllers in real-time.

A. Controller Setup

In our simulations, both the proposed RA-MPPI controller and the baseline MPPI controller evaluate the costs of their sampled trajectories, using (15) and (3), respectively. For both controllers, we use the following running cost,

$$q(x) = c_1 \mu_{\text{bdry}}(x) + c_2 \mu_{\text{obs}}(x) + c_3 q_{\text{dev}}(x), \quad (19)$$

where in the simulations the weights are $c_1 = 2, c_2 = 1, c_3 = 0.1$, and the cost $q_{\text{dev}}(x) = e(x)^2$ penalizes the lateral deviation $e(x)$ between the vehicle's CoM and the track centerline. The differentiable boundary cost $\mu_{\text{bdry}}(x)$ is,

$$\mu_{\text{bdry}}(x) = \max \left\{ 0, \frac{\tan^{-1}(-100d(x))}{\pi} + \frac{1}{2} \right\}, \quad (20)$$

Algorithm 1: Risk-Aware MPPI Algorithm

Given: RA-MPPI Parameters $\gamma, \eta, \alpha, A, B, C_u$;
Input: Initial control sequence \mathbf{v}

```

1 while task not complete do
2    $x_0 \leftarrow \text{GetStateEstimate}()$ ;
3   for  $m \leftarrow 0$  to  $M - 1$  in parallel do
4      $x_0^m \leftarrow x_0, S_m \leftarrow 0$ ;
5     Sample  $\epsilon^m \leftarrow \{\epsilon_0^m, \dots, \epsilon_{N-1}^m\}$ ;
6     for  $k \leftarrow 0$  to  $K - 1$  do
7       if  $m < (1 - \eta)M$  then
8          $u_k^m \leftarrow v_k + \epsilon_k^m$ ;
9       else
10         $u_k^m \leftarrow \epsilon_k^m$ ;
11      end
12       $x_{k+1}^m \leftarrow F(x_k^m, u_k^m)$ ;
13       $S_m \leftarrow S_m + q_k(x_k^m) + \gamma v_k^\top \Sigma_\epsilon^{-1} u_k^m$ ;
14    end
15     $S_m \leftarrow S_m + \phi(x_N^m)$ ;
16    for  $n \leftarrow 0$  to  $N - 1$  in parallel do
17       $\tilde{x}_0^{m,n} \leftarrow x_0, L(\tilde{\mathbf{x}}^{m,n}) \leftarrow 0$ ;
18      Sample  $\mathbf{w}^{m,n} \leftarrow \{w_0^{m,n}, \dots, w_{N-1}^{m,n}\}^\top$ ;
19      for  $k \leftarrow 0$  to  $K - 1$  do
20         $\tilde{x}_{k+1}^{m,n} \leftarrow \tilde{F}(\tilde{x}_k^{m,n}, u_k^m, w_k^{m,n})$ ;
21         $L(\tilde{\mathbf{x}}^{m,n}) \leftarrow L(\tilde{\mathbf{x}}^{m,n}) + l(\tilde{x}_k^{m,n})$ ;
22      end
23    end
24     $L_a^m \leftarrow \text{Scal}(\{L(\tilde{\mathbf{x}}^{m,0}), \dots, L(\tilde{\mathbf{x}}^{m,N-1})\}, B)$ ;
25     $P \leftarrow \text{Max}(L_a^m, 1 - \alpha)$ ;
26     $\text{CVaR}_\alpha^m \leftarrow \text{Average}(P)$ ;
27    if  $\text{CVaR}_\alpha^m > C_u$  then
28       $\tilde{S}_m \leftarrow S_m + A \cdot \text{CVaR}_\alpha^m$ ;
29    end
30  end
31   $\mathbf{v}^+ \leftarrow \text{CalculateOptimalControl}(\tilde{S}, U)$ ;
32  ExecuteCommand( $v_0^+$ );
33   $\mathbf{v} \leftarrow \mathbf{v}^+$ ;
34 end

```

and the discrete obstacle cost $\mu_{\text{obs}}(x)$ is given by,

$$\mu_{i,\text{obs}}(x) = \begin{cases} 1, & \text{if } d_i(x) < r_i, \\ 0, & \text{otherwise.} \end{cases} \quad (21)$$

Notice that (20) is a smooth approximation of the unit step function that penalizes collisions with the track boundaries, and $d(x)$ is the signed distance from the nearest track boundary to the vehicle's CoM, which is a positive number if the vehicle stays on the track. The term $d_i(x)$ in (21) measures the distance between the vehicle's CoM to the center of the i th circular obstacle with radius r_i . The terminal costs in both (3) and (15) are set the same as,

$$\phi(x) = c_4 - c_5 s(x), \quad (22)$$

where we use the weights $c_4 = 0.6$ and $c_5 = 2$ such that $\phi(x) \geq 0$ for all x in the simulations, and $s(x)$ is the distance traveled by the m^{th} noise-free sampled trajectory by the end of the control horizon along the track centerline.

Figure 2 demonstrates the physical meaning of the variables $e(x)$, $s(x)$, $d(x)$, $d_i(x)$ in (19)-(22) with a track schematic.

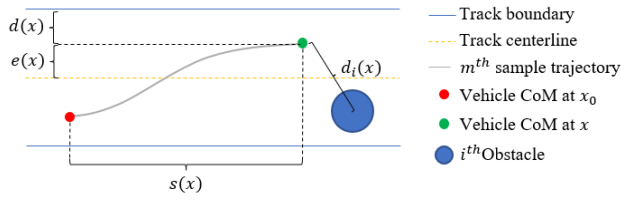


Fig. 2: Track schematic.

For both the RA-MPPI and the MPPI controllers, we set the control horizon to $K = 30$, the inverse temperature λ in (5) and (16) to be $\lambda = 0.35$, and the portion of the zero-mean control trajectories η in Algorithm 1, Line 7 to be $\eta = 0.2$. For the proposed RA-MPPI controller, we set the number of sampled noise-free trajectories to be $M = 1024$, and for each such trajectory, we sample $N = 300$ to evaluate its corresponding CVaR value using (11). As a result, the RA-MPPI samples 3.072×10^5 trajectories in total at each time step. We further let the weight A of the CVaR constraint violation cost in (13) be $A = 10$, and the running risk cost $\ell(\tilde{x}_k)$ in (8) be the same as the state-dependent cost (19), such that $\ell(\tilde{x}_k) = q(\tilde{x}_k)$. The number of sampled trajectories for the baseline MPPI is set to be equal to the total number of sampled trajectories of the RA-MPPI, in order to make a fair comparison between the controllers.

B. Aggressive Driving in Cluttered Environments Subject to Disturbances

Our simulations use a 0.6 m wide track with a centerline of length 10.9 m, and each corner of the track has a turning radius 0.3 m. There are 10 circular obstacles of the same radius $r_i = 0.1$ m scattered randomly on the track. The track is shown in Figure 3. In addition, we use a combination of the discrete obstacle cost (21) and the continuous boundary cost (20) for the track to test the robustness of the controllers, and the controllers are set up to drive the vehicle as fast as possible.

We run both the proposed RA-MPPI and the original MPPI controllers with system (6) using Gaussian noise $w_k \sim \mathcal{N}(0, 0.2I)$, in order to compare their performance in the presence of dynamical disturbances. Each controller runs the vehicle for 100 laps and Figure 3 shows the resulting trajectories.

The RA-MPPI trajectories show strong awareness to avoid collisions in this aggressive driving setting, resulting in 80% fewer collisions than the MPPI controller, while achieving approximately the same lap time. This result indicates that the RA-MPPI can prevent worst-case scenarios without sacrificing the average performance of the controller.

As the proposed control architecture in Figure 1 utilizes Monte-Carlo sampling to evaluate CVaR, the RA-MPPI controller accepts arbitrary forms of disturbances w_k in (6). To this end, we examined the robustness of the proposed RA-MPPI controller by repeating the simulations in Figure 3 using various forms of w_k . Specifically, in addition to the

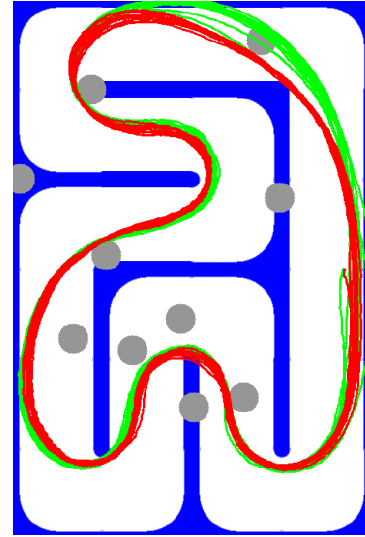


Fig. 3: Samples of RA-MPPI and MPPI trajectories on the race track. The trajectories in red are generated by the proposed RA-MPPI controller, and the trajectories in green are generated by the baseline MPPI controller.

normally distributed disturbance $w_k \sim \mathcal{N}(0, 0.2I)$, we also use uniformly distributed noise $w_k \sim \mathcal{U}[-0.2, 0.2]$ and a form of impulse/jump noise disturbance w_k that has a 2% chance to give the system state \tilde{x}_{k+1} in (6) a sudden change of magnitude 0.45 in arbitrary directions. We define a collision to be a situation when the vehicle state is inside some infeasible region, such as the space occupied by obstacles or areas outside of the track. The simulation results are shown in Figure 4.

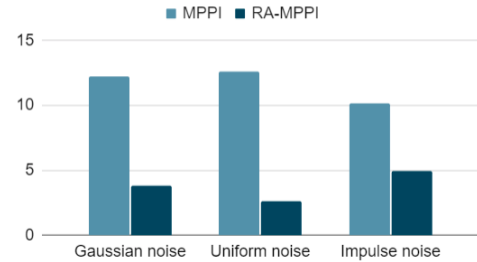


Fig. 4: The average number of collisions per lap for the RA-MPPI and MPPI controllers subject to various forms of noise. The RA-MPPI controller outperforms the MPPI controller by reducing the number of collisions by at least 50% for all of the three selected types of noise.

To further investigate the performance of the proposed RA-MPPI controller with different user-specified risk levels, we performed a grid search by varying the CVaR constraint upper bound C_u , and the CVaR confidence level α in (7b). The results are summarized in Figure 5. In the figure, the proposed RA-MPPI algorithm tends to achieve fewer collisions as the CVaR confidence level α increases, despite some outliers resulting from the randomness of the simulations. Moreover, as we tighten the CVaR constraint by decreasing the constraint upper bound value C_u , the RA-MPPI shows more robust performance and experiences fewer collisions. Due to the fact that the CVaR is a time-inconsistent risk metric [23], applying an overly strict CVaR constraint such as setting $C_u = 0.5$ and $\alpha = 0.9$ in (7b) puts a large

weight on preventing risk evaluated at the current time while aggravating the overall performance [23]. To alleviate this situation, we can simply loosen the constraint by using $C_u = 0.6$ and $\alpha = 0.7$, or use a time-consistent modification of CVaR [22].

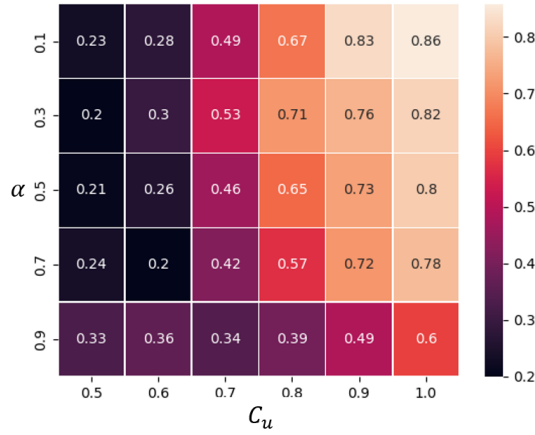


Fig. 5: RA-MPPI and MPPI performance comparison. The numbers in this chart shows the performance ratio (N_R/N_M), where N_R is the number of collisions of the proposed RA-MPPI, and N_M is the number of collisions of the MPPI controller. Each data point is collected by running 100 laps.

C. Autonomous Racing Experiment

To validate the proposed RA-MPPI controller in the real world, we implemented the algorithm on an autonomous racing platform and tested its performance. Our experiment uses the same track layout as the simulations described in Section V-B. Figure 6 shows a snapshot of our autonomous vehicle running on a real-world track. A schematic for this autonomous racing system setup is illustrated in Figure 7.

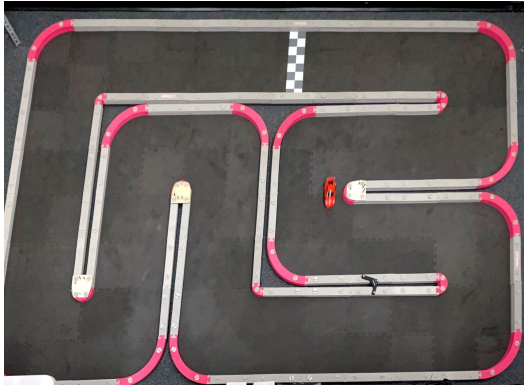


Fig. 6: The Autonomous vehicle racing on the track.

The system uses ten Optitrack cameras to observe the race track and the motion of a 1/28 scale race car on the track. These cameras pass raw data to the Optitrack computer, which then computes the state of the race car. The system utilizes the Optitrack visual tracking system for state updates of the vehicle. The state measurements are broadcasted on a local network to the control computer. As observed in the experiment video¹, the vehicle experiences disturbances,

¹<https://youtu.be/Ru6N8-qByGY>

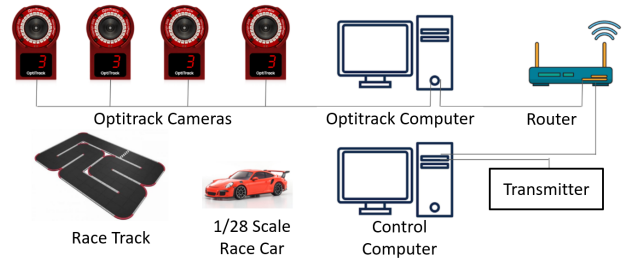


Fig. 7: Autonomous racing platform system schematic.

such as unstable network connection and unmodelled noises, which cause wiggling maneuvers. The collected results in Table I shows that the RA-MPPI experiences 57.7% fewer collisions with approximately the same lap time.

TABLE I: RA-MPPI and MPPI experimental results.

| Controller Type | No.Collision/lap | Avg. Laptime(s) |
|-----------------|------------------|-----------------|
| MPPI | 2.53 | 7.47 |
| RA-MPPI | 1.07 | 7.56 |

The performance of the proposed RA-MPPI Algorithm 1 requires a parallel implementation of the control architecture described in Figure 1. Compared to our previous implementation of MPPI in [11], we reduced the data exchange between the GPU and CPU to a minimum by moving all necessary computations to GPU. Specifically, our implementation of the RA-MPPI controller computes the dynamics propagation, including both the nominal and the disturbed dynamics, and evaluates CVaR and the trajectory costs in parallel on the GPU. On average, the resulting RA-MPPI can sample and evaluate 1.024×10^5 trajectories at 81.6 Hz, 3.072×10^5 trajectories at 44.3 Hz and 5.120×10^5 trajectories at 27.8 Hz, using an Nvidia GeForce RTX3090 GPU.

VI. CONCLUSIONS AND FUTURE WORK

In this paper, we proposed the RA-MPPI algorithm, a risk-aware model predictive control scheme for general nonlinear systems subject to disturbances. The proposed controller accepts arbitrary forms of dynamical disturbances and carries out risk-aware control by evaluating the CVaR values using Monte-Carlo sampling. Our simulations show that the RA-MPPI controller can effectively reduce the chances of catastrophic scenarios without compromising the overall control performance, even in time-critical applications such as aggressive autonomous driving.

In the future, we can integrate chance constraints [7] and time-consistent modifications of CVaR [22], [23] with the RA-MPPI control architecture. Since different risk metrics interpret risks from their own unique aspect, it would be interesting to compare experimentally these risk-aware controllers. In addition, we can further improve the RA-MPPI control architecture by applying ideas similar to [11] to achieve adjustable trajectory sampling distributions, hence increasing the sampling efficiency of the proposed algorithm.

VII. ACKNOWLEDGEMENT

This work was funded by NSF under awards CNS-2219755 and CCF-2238030.

REFERENCES

- [1] W. Li and E. Todorov, "Iterative linear quadratic regulator design for nonlinear biological movement systems," *Proceedings of the 1st International Conference on Informatics in Control, Automation and Robotics, (ICINCO 2004)*, vol. 1, pp. 222–229, Jan 2004.
- [2] P. Falcone, M. Tufo, F. Borrelli, J. Asgari, and H. E. Tseng, "A linear time varying model predictive control approach to the integrated vehicle dynamics control problem in autonomous systems," in *46th IEEE Conference on Decision and Control*, 2007, pp. 2980–2985.
- [3] A. A. Jalali and V. Nadimi, "A survey on robust model predictive control from 1999–2006," in *International Conference on Computational Intelligence for Modelling Control and Automation and International Conference on Intelligent Agents Web Technologies and International Commerce (CIMCA'06)*, 2006, pp. 207–207.
- [4] M. Cannon, B. Kouvaritakis, and D. Ng, "Probabilistic tubes in linear stochastic model predictive control," *Systems & Control Letters*, vol. 58, pp. 747–753, 2009.
- [5] M. Cannon, B. Kouvaritakis, and X. Wu, "Probabilistic constrained MPC for multiplicative and additive stochastic uncertainty," *IEEE Transactions on Automatic Control*, vol. 54, no. 7, pp. 1626–1632, 2009.
- [6] D. Bernardini and A. Bemporad, "Scenario-based model predictive control of stochastic constrained linear systems," in *Proceedings of the 48th IEEE Conference on Decision and Control*, Shanghai, China, 2009, pp. 6333–6338, held jointly with the 28th Chinese Control Conference.
- [7] K. Okamoto and P. Tsotras, "Stochastic model predictive control for constrained linear systems using optimal covariance steering," 2019, arXiv:1905.13296.
- [8] L. Hewing and M. N. Zeilinger, "Stochastic model predictive control for linear systems using probabilistic reachable sets," in *IEEE Conference on Decision and Control (CDC)*, Dec 2018.
- [9] B. Kouvaritakis, M. Cannon, S. V. Raković, and Q. Cheng, "Explicit use of probabilistic distributions in linear predictive control," *Automatica*, vol. 46, no. 10, pp. 1719–1724, 2010.
- [10] R. T. Rockafellar, S. Uryasev *et al.*, "Optimization of conditional value-at-risk," *Journal of risk*, vol. 2, pp. 21–42, 2000.
- [11] J. Yin, Z. Zhang, E. Theodorou, and P. Tsotras, "Trajectory distribution control for model predictive path integral control using covariance steering," in *International Conference on Robotics and Automation (ICRA)*, 2022, pp. 1478–1484.
- [12] G. Williams, B. Goldfain, P. Drews, K. Saigol, J. Rehg, and E. Theodorou, "Robust sampling based model predictive control with sparse objective information," June 2018.
- [13] M. S. Gandhi, B. Vlahov, J. Gibson, G. Williams, and E. A. Theodorou, "Robust model predictive path integral control: Analysis and performance guarantees," *IEEE Robotics and Automation Letters*, vol. 6, no. 2, pp. 1423–1430, 2021.
- [14] I. R. Manchester and J.-J. E. Slotine, "Control contraction metrics: Convex and intrinsic criteria for nonlinear feedback design," *IEEE Transactions on Automatic Control*, vol. 62, no. 6, pp. 3046–3053, 2017.
- [15] Y. Chow, A. Tamar, S. Mannor, and M. Pavone, "Risk-sensitive and robust decision-making: a cvar optimization approach," 2015. [Online]. Available: <https://arxiv.org/abs/1506.02188>
- [16] P. Krokmal, J. Palmquist, and S. Uryasev, "Portfolio optimization with conditional value-at-risk objective and constraints," *Journal of Risk*, vol. 4, May 2003.
- [17] M. Ahmadi, U. Rosolia, M. D. Ingham, R. M. Murray, and A. D. Ames, "Risk-averse decision making under uncertainty," *arXiv preprint arXiv:2109.04082*, 2021.
- [18] Y.-L. Chow and M. Pavone, "A framework for time-consistent, risk-averse model predictive control: Theory and algorithms," in *2014 American Control Conference*. IEEE, 2014, pp. 4204–4211.
- [19] A. Ruszczyński, "Risk-averse dynamic programming for markov decision processes," *Mathematical programming*, vol. 125, no. 2, pp. 235–261, 2010.
- [20] P. Artzner, F. Delbaen, J.-M. Eber, D. Heath, and H. Ku, "Coherent multiperiod risk adjusted values and bellman's principle," *Annals of Operations Research*, vol. 152, no. 1, pp. 5–22, 2007.
- [21] K. Okamoto and P. Tsotras, "Optimal stochastic vehicle path planning using covariance steering," *IEEE Robotics and Automation Letters*, vol. 4, no. 3, pp. 2276–2281, 2019.
- [22] S. Samuelson and I. Yang, "Safety-aware optimal control of stochastic systems using conditional value-at-risk," *American Control Conference (ACC)*, pp. 6285–6290, 2018.
- [23] Y.-L. Chow and M. Pavone, "Stochastic optimal control with dynamic, time-consistent risk constraints," *American Control Conference*, pp. 390–395, June 2013.
- [24] O. Çavuş and A. Ruszczyński, "Risk-averse control of undiscounted transient markov models," *SIAM Journal on Control and Optimization*, vol. 52, no. 6, pp. 3935–3966, 2014.
- [25] A. Deo and K. Murthy, "Efficient black-box importance sampling for var and cvar estimation," in *Winter Simulation Conference (WSC)*. IEEE, 2021, pp. 1–12.
- [26] Z. Wang, O. So, K. Lee, and E. A. Theodorou, "Adaptive risk sensitive model predictive control with stochastic search," in *Learning for Dynamics and Control*. PMLR, 2021, pp. 510–522.