2024 IEEE International Conference on Robotics and Automation (ICRA)
May 13-17, 2024. Yokohama, Japan

2024 IEEE International Conference on Robotics and Automation (ICRA) | 979-8-3503-8457-4/24/$31.00 ©2024 IEEE | DOI: 10.1109/ICRA57147.2024.10610765

Online-Learning-Based Distributionally Robust Motion Control with Collision Avoidance for Mobile Robots

Han Wang, Chao Ning*, Longyan Li, and Weidong Zhang

Abstract— **Collision-free navigation is a critical issue in robotic systems as the environment is often dynamic and uncertain. This paper investigates a data-stream-driven motion control problem for mobile robots to avoid randomly moving obstacles when the probability distribution of the obstacle's movement is partially observable through data and can be even time-varying. A data-stream-driven ambiguity set is firstly constructed from movement data by leveraging a Dirichlet process mixture model and is updated online using real-time data. Then we propose an Online-Learning-based Distributionally Robust Nonlinear Model Predictive Control (OL-DR-NMPC) approach for limiting the risk of collision through considering the worst-case distribution within the ambiguity set. To facilitate solving the OL-DR-NMPC problem, we reformulate it as a finite-dimensional nonlinear optimization problem. To cope with the bilinear matrix inequality constraints in the nonlinear problem, we develop a parabolic relaxation and a sequential algorithm, by which the problem is further transformed into polynomial-time solvable surrogates. The simulations using a quadrotor model are employed to demonstrate the effectiveness and advantages of the proposed method.**

I. Introduction

Safety is one of the most fundamental challenges for the motion control of mobile robots in uncertain and dynamic environments. The unexpected motion of obstacles typically poses a risk to the collision-free navigation of robots and brings about disastrous consequences. The key to ensuring safety is to estimate an accurate probability distribution of motion, which is a challenging task due to the limited amount of motion data. The goal of our work is to develop an optimization-based distributionally robust motion control framework that ensures safety when the underlying true distribution is only partially known via data streams.

There has been a plethora of risk-averse decision-making methods proposed to avoid collision for robot motion control in the existing literature [1]. Among them, chance-constrained methods are the most popular ones, ascribing to their capability of directly curbing the chance of collision [2]. However, the nonconvexity of chance constraints renders solving a chance-constrained optimization problem computationally expensive. A commonly used method is to leverage a computationally tractable risk measure. In particular, Conditional Value-at-Risk (CVaR) has recently stimulated considerable research interest in motion planning and control [3]. In addition to its computational merit, CVaR is effective in accounting for rare yet unsafe events. To leverage such advantages, we utilize CVaR to measure the risk of unsafety.

In practice, the underlying probability distribution is not perfectly known and can be partially inferred from data. To estimate the distribution, a straightforward way is to construct an empirical distribution using historical data [3]. However, the empirical distribution is likely to deviate from the actual distribution with a finite number of historical data; thus, the risk constraints are likely to be violated. One approach is to construct a set comprising all distributions within a Wasserstein ball centered at the empirical distribution. The CVaR constraints must hold for any distribution in this set [4]. However, choosing the radius of the Wasserstein ball online during the motion control is a challenging task. Worse still, without online learning, the motion control actions depend entirely on historical data, which means the violation of risk constraints is still likely to occur if the distribution is time-varying. Furthermore, some research papers utilize the Gaussian Processes Regression (GPR) to learn the distribution online [5]. Although GPR is quite effective, it can only obtain the global moment information rather than the fined-grained distribution information of multimodality. In some cases, this method might be over conservative. No existing method has been developed to automatically balance safety and control performance through a data lens.

Motivated by these considerations, we propose an online-learning-based distributionally robust motion control framework in such a situation where the probability distribution of the random motion is multimodal and\or time-varying. In order to learn the distribution and decipher its structural properties online, we construct a data-stream-driven ambiguity set for the random motion based on the Dirichlet Process Mixture Model (DPMM), and this set is devised as a weighted Minkovski sum of local basic ambiguity sets. In the offline planning stage, the initial ambiguity set is constructed using the historical data, and a collision-free reference trajectory is generated by using Rapidly exploring Random Tree* (RRT*) [6]. However, as the obstacles start to randomly move, the reference trajectory might not be safe to follow. To limit the risk of collision, a novel proposed Nonlinear Model Predictive Control (NMPC), called Online-Learning-based Distributionally Robust NMPC (OL-DR-NMPC), is employed in the online motion control stage. The OL-DR-NMPC is designed with CVaR constraints that must hold for any distribution in the ambiguity set. At each

This work was supported in part by the National Natural Science Foundation of China (62103264), Shanghai Pujiang Program (21PJ1406200), the National Key R&D Program of China (2022ZD0119903), Shanghai Science and Technology program (22015810300), Hainan Province Science and Technology Special Fund (ZDYF2021GXJS041), and the National Natural Science Foundation of China (U2141234).

*C. Ning is the corresponding author.

H. Wang, C. Ning, L. Li and W. Zhang are with the Department of Automation, Shanghai Jiao Tong University, and Key Laboratory of System Control and Information Processing, Ministry of Education of China, Shanghai, China. E-mail: {h.wang, chao.ning, lilongyan, wdzhang} @sjtu.edu.cn

sampling time, the robot observes the position of the obstacles and obtains the new data of the random motion, which is used to update the ambiguity set. To resolve the computational challenge of the infinite-dimensional OL-DR-NMPC problem, we first reformulate the CVaR constraint as the finite-dimensional Linear and Bilinear Matrix Inequality (LMI, BMI) constraints. Then, we further transform the problem into polynomial-time solvable surrogates by applying convex relaxation to BMI constraints and propose a sequential penalized relaxation algorithm to solve it.

II. PRELIMINARIES

A. Mobile Robot and Obstacle

In this paper, we consider a mobile robot modeled by the following time-invariant discrete-time system.

$$x(t+1) = Ax(t) + Bu(t), \quad y(t) = Cx(t) \tag{1}$$

where $x(t) \in \mathbb{R}^{n_x}$, $u(t) \in \mathbb{R}^{n_u}$ and $y(t) \in \mathbb{R}^{n_y}$ are the state, input, and output, respectively. The system output is defined as the Cartesian coordinates of the robot's Center of Mass (CoM) [5]. The system is subject to the following constraints.

$$x(t) \in \mathcal{X}, \quad u(t) \in \mathcal{U} \quad \forall t \geq 0 \tag{2}$$

where $\mathcal{X} \subseteq \mathbb{R}^{n_x}$ and $\mathcal{U} \subseteq \mathbb{R}^{n_u}$ are assumed to be convex sets.

The robot navigates an environment with L *randomly moving* obstacles. Let $O_\ell(t) \in \mathbb{R}^{n_y}$ denote the CoM of the ℓth obstacle at time t for $\ell = 1, \ldots, L$. The movement of the obstacle between two time periods is modeled by translations.

$$O_\ell(t+k) = O_\ell(t) + \omega_{\ell,t,k} \tag{3}$$

where $\omega_{\ell,t,k}$ is a random translation vector in $\mathbb{R}^{n_y}$, which is illustrated in Fig. 1. For safety concern, the robot should navigate within a region determined by the behaviors of obstacles. The safe region regarding obstacle ℓ is defined as the region outside of the open ball centered at the CoM of the obstacle with safe distance $r_\ell > 0$.

$$\mathcal{Y}^\ell(t) := \left\{ y(t) \in \mathbb{R}^{n_y} \,\middle|\, \operatorname{dist}\left(y(t), O_\ell(t)\right) \geq r_\ell \right\} \tag{4}$$

where $\operatorname{dist}\left(y(t), O_\ell(t)\right)$ is the Euclidean distance between the CoMs of the robot and the obstacle, defined by (5).

$$\operatorname{dist}\left(y(t), O_\ell(t)\right) := \left\| y(t) - O_\ell(t) \right\|_2 \tag{5}$$

An example of such a configuration is illustrated in Fig. 1. A quadrotor should navigate to avoid an obstacle. Both the quadrotor and the obstacle are approximated by the smallest balls enclosing them with radii r_r and r_ℓ, respectively [5]. Considering a safety margin r_s, the distance between the CoM of the quadrotor and the obstacle should be no smaller than the sum of all radii, as given in (6).

$$r_\ell = r_r + r_o^\ell + r_s \tag{6}$$

The safe region regarding all obstacles is defined as the intersection of all safe regions, as shown in (7).

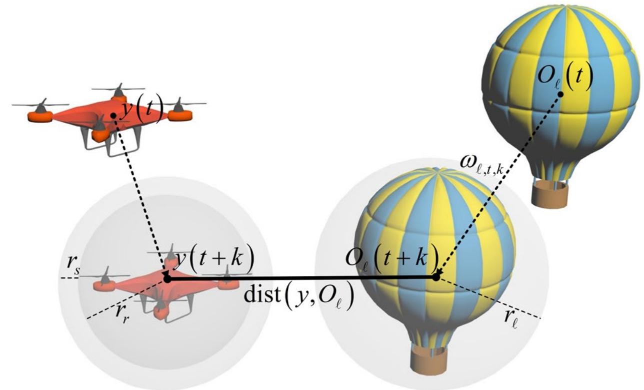


Figure 1. Quadrotor and randomly moving obstacle

$$\mathcal{Y}(t) := \bigcap_{\ell=1}^{L} \mathcal{Y}^\ell(t) \tag{7}$$

Note that the safe region is time-varying due to the random motion of the obstacles.

B. Reference Trajectory Planning

In the offline planning stage, a collision-free trajectory is generated using path-planning tools based on the initial position of obstacles. The resulting trajectory is traced as a reference trajectory in the online control stage. In our work, the RRT* is employed to generate a reference trajectory [6]. It is based on a rapidly growing tree graph, which expands by randomly sampling the configuration space. Further, it can provide an asymptotically optimal solution by using neighbor search and tree rewiring.

III. ONLINE-LEARNING-BASED DISTRIBUTIONALLY ROBUST MOTION CONTROL

Although the reference trajectory is collision-free for the initial configuration of the obstacles, it might be unsafe to follow when the obstacles begin to move. In order to curb the risk of collision, we develop a novel distributionally robust motion control framework, which can automatically decipher the structural property of the probability distribution of obstacles' movement using both historical and real-time data.

A. Risk Constraint for Safety using CVaR

To begin with, we define the *loss of safe* at each prediction time $t + k$, evaluated at t, with respect to obstacle ℓ in (8).

$$\mathcal{L}_{t,k}(y, O_\ell) = -\left\| y(t+k) - O_\ell(t+k) \right\|_2^2 \tag{8}$$

According to (4), the robot navigates in the safe region $\mathcal{Y}^\ell(t)$ if and only if $\mathcal{L}_t(y, O_\ell) + r_\ell^2 \leq 0$. However, due to the uncertain movement of the obstacle, employing such a hard constraint might be too conservative. Therefore, we take a risk-averse approach by using CVaR to resolve this issue. Our method considers the following risk constraint.

$$\mathrm{CVaR}_\alpha\left[\mathcal{L}_{t,k}(y, O_\ell)\right] + r_\ell^2 \leq 0 \tag{9}$$

where $\mathrm{CVaR}_\alpha[X] := \min_{z \in \mathbb{R}} \mathbb{E}\left[z + (X - z)^+ / (1-\alpha)\right]$ measures the conditional expectation of the loss within the $(1-\alpha)$ worst-case quantile.

B. Online Learning the Movements using DPMM

To compute the CVaR of the loss of safe in the risk constraint (9), we need the distribution of $\omega_{\ell,t,k}$. However, it is challenging to obtain a perfectly accurate distribution in real practice. In many cases, there is only a limited amount of data available for inferring the underlying true distribution. Furthermore, the distribution might be multimodal, and can even be time-varying in nature.

Fortunately, the position of the obstacles can be observed by the robot using sensors like cameras [7], and the observations can be instrumental in online learning the movements of the obstacles. Specifically, as the prediction horizon moves forward, the real-time data of $\omega_{\ell,t,k}$ are gathered for learning the distribution online. To decipher the structural property of the distribution of $\omega_{\ell,t,k}$, we employ the DPMM, which is one of the most popular nonparametric Bayesian models.

In order to learn from the translation vector data streams, we utilize an online variational inference algorithm. The reader is referred to [8] for the detail of the algorithm. Based on the online variational inference results up to time t, we develop an ambiguity set for $\omega_{\ell,t,k}$ as follows.

$$\mathbb{D}_{\ell,t,k} := \sum_{j=1}^{m_{\ell,t,k}} \gamma_{\ell,t,k}^{(j)} \mathbb{D}_{\ell,t,k}^{(j)}\left(\mathbb{W}, \mu_{\ell,t,k}^{(j)}, \Sigma_{\ell,t,k}^{(j)}\right) \tag{10a}$$

$$\mathbb{D}_{\ell,t,k}^{(j)}\left(\mathbb{W}, \mu_{\ell,t,k}^{(j)}, \Sigma_{\ell,t,k}^{(j)}\right) := \left\{ \rho \in \mathcal{P}(\mathbb{W}) \left| \begin{array}{l} \int_{\mathbb{W}} \rho(d\omega) = 1 \\ \int_{\mathbb{W}} \omega \cdot \rho(d\omega) = \mu_{\ell,t,k}^{(j)} \\ \int_{\mathbb{W}} \omega\omega^T \cdot \rho(d\omega) \le \Sigma_{\ell,t,k}^{(j)} + \mu_{\ell,t,k}^{(j)}\left(\mu_{\ell,t,k}^{(j)}\right)^T \end{array} \right. \right\} \tag{10b}$$

where $\mathbb{W} := \left\{ \omega \in \mathbb{R}^{n_y} \middle| H\omega \le h \right\}$ is the support set of $\omega_{\ell,t,k}$, $m_{\ell,t,k}$ is the number of mixture components, and $\gamma_{\ell,t,k}^{(j)}$ is the mixing weight indicating the occurring probability of the jth mixture component. $\mathbb{D}_{\ell,t,k}$ represents the ambiguity set, which is devised as a Minkowski sum of $m_{\ell,t,k}$ basic ambiguity sets $\mathbb{D}_{\ell,t,k}^{(j)}$. $\mu_{\ell,t,k}^{(j)}$ and $\Sigma_{\ell,t,k}^{(j)}$ denote the mean and covariance estimate for the jth mixture component, respectively. $\mathcal{P}(\mathbb{W})$ denotes the set of positive Borel measures on the supports and ρ is a positive measure.

There are several advantages of the proposed ambiguity set. First, leveraging the local moment information, the proposed ambiguity set is more suitable for describing multimodal distributions compared with the ambiguity set based on global moment information, e.g., GPR-based ambiguity sets. Moreover, each basic ambiguity is devised using mean and covariance information, which endows the following OL-DR-NMPC with enormous computational benefits to meet the real-time requirements of robotic systems.

C. Online-Learning-Based Distributionally Robust MPC

Given the data-stream-driven ambiguity set of $\omega_{\ell,t,k}$, the risk-averse motion control problem is formulated as an OL-DR-NMPC problem.

$$\inf_{\boldsymbol{u},\boldsymbol{x},\boldsymbol{y}} J(x(t),\boldsymbol{u}) := \left\| y_K - y_K^{ref} \right\|_P^2 + \sum_{k=0}^{K-1} \left\| y_k - y_k^{ref} \right\|_Q^2 + \left\| u_k \right\|_R^2 \tag{11a}$$

$$s.t.\ x_{k+1} = \mathrm{A}x_k + Bu_k \tag{11b}$$

$$y_k = Cx_k \tag{11c}$$

$$x_0 = x(t) \tag{11d}$$

$$x_k \in \mathcal{X} \tag{11e}$$

$$u_k \in \mathcal{U} \tag{11f}$$

$$\sup_{\upsilon_{\ell,t,k} \in \mathbb{D}_{\ell,t,k}} \mathrm{CVaR}_\alpha^{\upsilon_{\ell,t,k}}\left[\mathcal{L}_{t,k}\left(y_k, O_\ell\right)\right] + r_\ell^2 \le 0 \tag{11g}$$

where $\boldsymbol{u} := (u_0, \ldots, u_{K-1})$, $\boldsymbol{x} := (x_0, \ldots, x_K)$, $\boldsymbol{y} := (y_0, \ldots, y_K)$. Constraints (11b) and (11f) should hold for $k = 0, \ldots, K-1$, Constraints (11c) and (11e) should be imposed for $k = 0, \ldots, K$, and Constraint (11g) should be satisfied for $k = 1, \ldots, K$, $\ell = 1, \ldots, L$. $Q \succeq 0$, $R \succeq 0$ are the weighting matrices of the state and input, respectively. $P \succeq 0$ is chosen to ensure stability. y^{ref} is the reference trajectory generated in Section II-B. Constraint (11g) is the most important part. Since the distribution obtained by DPMM is imperfect, we cannot compute the CVaR of the loss of safety accurately. Thus, we propose the Distributionally Robust CVaR (DR-CVaR), which represents the worst-case value of CVaR over all distributions in the data-stream-driven ambiguity set.

IV. Finite-Dimensional Reformulation and Its Convex Relaxation

The OL-DR-NMPC problem in (11) is nontrivial to solve due to the DR-CVaR constraint (11g). This constraint is infinite-dimensional, because there involves an infinite number of distributions in the ambiguity set. In this section, we first reformulate it as LMI and BMI constraints which are finite-dimensional. Then, we propose a parabolic relaxation for the BMI to guarantee computational tractability.

A. Finite-Dimensional Reformulation

We reformulate the OL-DR-NMPC problem as a finite-dimensional problem. For ease of exposition, we simplify (11g) by suppressing the subscripts as follows.

$$\sup_{\upsilon \in \mathbb{D}} \mathrm{CVaR}_\alpha^{\upsilon}\left[\mathcal{L}(y, O+\omega)\right] + r^2 \le 0 \tag{12}$$

The DR-CVaR constraint (12) can be reformulated as LMI and BMI constraints [9], as given in the following theorem.

Theorem 1. Suppose that $\mathbb{W} := \left\{ \omega \in \mathbb{R}^{n_y} \middle| H\omega \le h \right\}$. Then the DR-CVaR in (12) has the following upper bound.

$$\begin{aligned} &\sup_{\upsilon \in \mathbb{D}} \mathrm{CVaR}_\alpha^{\upsilon}\left[\mathcal{L}(y, O+\omega)\right] \\ &\le \inf_{z \in \mathbb{R}} z + \tfrac{1}{1-\alpha} \sum_{j=1}^{m} \gamma_j \left\{ t_j + \mu_j^T \xi_j + \left(\Sigma_j + \mu_j \mu_j^T\right) \bullet \Omega_j \right\} \\ &s.t. \begin{pmatrix} \Omega_j + I & \tfrac{1}{2}\left(\xi_j + H^T\varphi_j\right) - (y - O) \\ \tfrac{1}{2}\left(\xi_j + H^T\varphi_j\right)^T - (y-O)^T & z + t_j - h^T\varphi_j + \left\| y - O \right\|_2^2 \end{pmatrix} \ge 0 \\ &\begin{pmatrix} \Omega_j & \tfrac{1}{2}\left(\xi_j + H^T\eta_j\right) \\ \tfrac{1}{2}\left(\xi_j + H^T\eta_j\right)^T & t_j - h^T\eta_j \end{pmatrix} \ge 0 \\ &\Omega_j \ge 0, \varphi_j \ge 0, \eta_j \ge 0 \end{aligned} \tag{13}$$

where all the constraints hold for $j=1,\ldots,m$ and I is the identity matrix.

Its proof is provided in the Appendix. Specifically, according to **Theorem 1**, the OL-DR-NMPC problem (11) can be reformulated as follows.

$$\inf_{\substack{\boldsymbol{u},\boldsymbol{x},\boldsymbol{y},\boldsymbol{z},\\ t,\xi,\Omega,\varphi,\lambda}} J\left(x(t),\boldsymbol{u}\right) := \left\|y_K - y_K^{ref}\right\|_P^2 + \sum_{k=0}^{K-1}\left\|y_k - y_k^{ref}\right\|_Q^2 + \left\|u_k\right\|_R^2 \quad (14a)$$

$$s.t. \quad x_{k+1} = \mathrm{A}x_k + Bu_k \quad (14b)$$

$$y_k = Cx_k \quad (14c)$$

$$x_0 = x(t) \quad (14d)$$

$$z_{\ell,k} + \tfrac{1}{1-\alpha}\sum_{j=1}^{m_{\ell,t,k}} \gamma_{\ell,t,k}^{(j)} \begin{Bmatrix} t_{\ell,k,j} + \mu_{\ell,t,k}^{(j)\,T}\xi_{\ell,k,j} + \\ \left(\Sigma_{\ell,t,k}^{(j)} + \mu_{\ell,t,k}^{(j)}\mu_{\ell,t,k}^{(j)\,T}\right)\bullet\Omega_{\ell,k,j} \end{Bmatrix} + r_\ell^2 \le 0 \quad (14e)$$

$$\begin{pmatrix} \Omega_{\ell,k,j} + I & \tfrac{1}{2}\left(\xi_{\ell,k,j} + H^T\varphi_{\ell,k,j}\right) - \left(y_k - O_\ell\right) \\ \tfrac{1}{2}\left(\xi_{\ell,k,j} + H^T\varphi_{\ell,k,j}\right)^T - \left(y_k - O_\ell\right)^T & z_{\ell,k} + t_{\ell,k,j} - h^T\varphi_{\ell,k,j} + \left\|y_k - O_\ell\right\|_2^2 \end{pmatrix} \ge 0 \quad (14f)$$

$$\begin{pmatrix} \Omega_{\ell,k,j} & \tfrac{1}{2}\left(\xi_{\ell,k,j} + H^T\lambda_{\ell,k,j}\right) \\ \tfrac{1}{2}\left(\xi_{\ell,k,j} + H^T\lambda_{\ell,k,j}\right)^T & t_{\ell,k,j} - h^T\lambda_{\ell,k,j} \end{pmatrix} \ge 0 \quad (14g)$$

$$\Omega_{\ell,k,j} \ge 0, \varphi_{\ell,k,j} \ge 0, \lambda_{\ell,k,j} \ge 0 \quad (14h)$$

$$x_k \in \mathcal{X}, u_k \in \mathcal{U}, z_{\ell,k} \in \mathbb{R} \quad (14i)$$

where (14b) and $u_k \in \mathcal{U}$ in (14i) should hold for $k=0,\ldots,K-1$, and all the other constraints hold for $k=0,\ldots,K$, $\ell=1,\ldots,L$, and $j=1,\ldots,m_{\ell,t,k}$. Note that $m_{\ell,t,k}$ might be different for different time t and obstacle ℓ, which is obtained from the DPMM.

As desired, the optimization problem is reformatted as a finite-dimensional problem. However, the problem (14) is still nonconvex due to the BMI constraint (14f), and solving such a problem is a computationally burdensome task. To alleviate the computational issue, we further propose a parabolic relaxation to convexify the problem and incorporate a penalty term into the objective function to ensure feasibility.

B. Convexification

In order to relax the BMI constraint (14f), we introduce an auxiliary matrix variable $Y_k \in \mathbb{R}^{n_y \times n_y}$ to account for $y_k y_k^T$. Then Constraint (14f) is reformulated as follows.

$$\begin{pmatrix} \Omega_{\ell,k,j} + I & \tfrac{1}{2}\left(\xi_{\ell,k,j} + H^T\varphi_{\ell,k,j}\right) - \left(y_k - O_\ell\right) \\ \tfrac{1}{2}\left(\xi_{\ell,k,j} + H^T\varphi_{\ell,k,j}\right)^T & z_{\ell,k} + t_{\ell,k,j} - h^T\varphi_{\ell,k,j} \\ -\left(y_k - O_\ell\right)^T & +tr\left\{Y_k\right\} - 2y_k^T O_\ell + \left\|O_\ell\right\|_2^2 \end{pmatrix} \ge 0 \quad (15)$$

where $tr\{\bullet\}$ denote the trace. To recover the original problem, the following constraint is added.

$$Y_k = y_k y_k^T \quad (16)$$

It is clear that BMI constraint (14f) is transformed into LMI constraint (15), and the entire non-convexity is captured by the new constraint (16). Next, we handle the non-convexity of Y_k through convexification. Specially, we utilize the recently developed parabolic relaxation to generate a low-complexity convex formulation. Following [10], Constraint (16) is convexified as follows.

$$\begin{aligned} \left[Y_k\right]_{ii} + \left[Y_k\right]_{jj} + 2\left[Y_k\right]_{ij} &\ge \left\|\left[y_k\right]_i + \left[y_k\right]_j\right\|_2^2 \forall i,j \in n_y, \forall k \\ \left[Y_k\right]_{ii} + \left[Y_k\right]_{jj} - 2\left[Y_k\right]_{ij} &\ge \left\|\left[y_k\right]_i - \left[y_k\right]_j\right\|_2^2 \forall i,j \in n_y, \forall k \end{aligned} \quad (17)$$

where $[a]_i$ and $[A]_{ij}$ respectively indicate the ith element of a and (i, j)th element of A given vector a and matrix A.

Note that the presented relaxations are not necessarily exact, so we incorporate a penalty term into the objective function (14a) to recover the feasible points of the original problem. The new objective function is given as follows.

$$J\left(x(t),\boldsymbol{u}\right) + \eta\sum_{k=0}^{K-1} tr\left\{Y_k\right\} - 2\breve{y}_k^T y_k \quad (18)$$

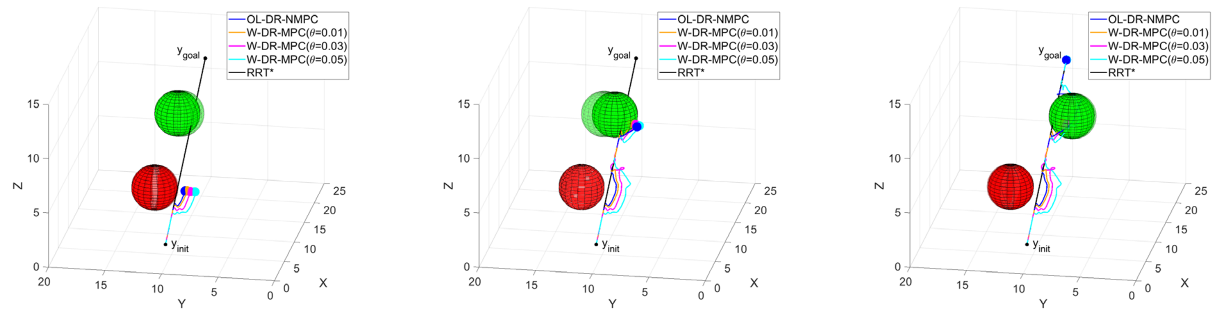
where $\breve{y}_k$ is an initial guess for the solution and $\eta > 0$ is a fixed parameter which is sufficiently large. It has been proved that if $\breve{y}_k$ is a feasible point for the original problem and satisfies Mangasarian-Fromovitz constraint qualification [11], the penalized convex relaxation preserves the feasibility of $\breve{y}_k$ and generates a solution with an improved objective value. It is also shown that even if the initial point $\breve{y}_k$ is not feasible for the original problem, but sufficiently close to the feasible set, the penalized convex relaxation is guaranteed to provide a feasible point. The reader is referred to [10] and [12] for more detail.

Motivated by [13], we develop a sequential penalized relaxation algorithm that generates a satisfactory solution of the OL-DR-NMPC problem by solving a sequence of the penalized relaxed problems. The proposed algorithm can start from an arbitrary initial point. Once a feasible point is found, the feasibility is preserved and the objective value is monotonically improved in the subsequent round.

The online-learning-based distributionally robust motion control algorithm including the sequential penalized relaxation algorithm is delineated in **Algorithm 1**.

Algorithm 1: OL-DR-NMPC algorithm

1: **Input:** x_0, $\breve{y}_k, k=0,\ldots,K$
2: **for** $t=0,1\ldots$ **do**
3: Observe the CoM of obstacles, and obtain $\omega_{\ell,t,k}$
4: Run online learning for DPMM with real-time data $\omega_{\ell,t,k}$
5: Update ambiguity set in (10)
6: **Sequential Penalized Relaxation:**
7: $i \leftarrow 0$, $y_k^0 \leftarrow \breve{y}_k$
8: **repeat**
9: $i \leftarrow i+1$
10: $y_k^i \leftarrow$ Solve the penalized relaxed problem
11: $\breve{y}_k \leftarrow y_k^i + \frac{i-1}{i+2}\left(y_k^i - y_k^{i-1}\right)$
12: **until** stopping criteria is met
13: Obtain u_k with regard to y_{k+1}^i
14: **end**
15: **return** $u(t) = u_0$



(a) $t = 30$ (b) $t = 68$ (c) $t = 100$

Figure 2. Trajectories of the quadrotor controlled by OL-DR-NMPC and W-DR-MPC with multiple θ's

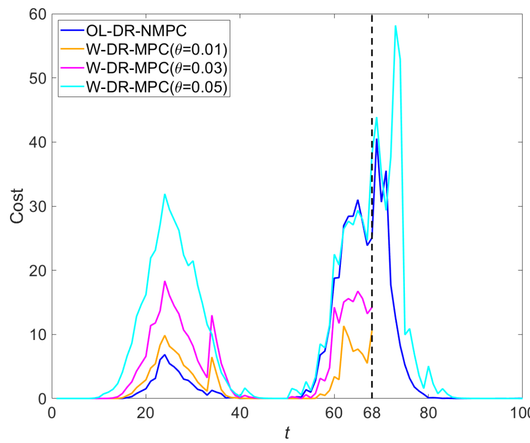


Figure 3. Cost with different controllers at each time t

TABLE I. OPERATION COST, SUCCESS RATE AND COMPUTATION TIME FOR THE QUADROTOR MOTION CONTROL WITH DIFFERENT CONTROLLERS

	OL-DR-NMPC	**W-DR-MPC (θ)**		
		0.01	0.03	0.05
Cost (before t = 50)	52.48	91.32	197.3	394.3
Cost (total)	470.85	-	-	978.4
Success Rate (%)	100	57	73	78
Average Time (sec)	2.67	9.77	12.76	18.45

V. SIMULATION RESULT

In this section, we conduct simulations to demonstrate the effectiveness and superiority of the proposed online-learning-based distributionally robust motion control method. In the simulations, we consider a quadrotor navigating in a dynamic 3D environment with the following linear dynamics.

$$\begin{aligned} &\ddot{x} = g\theta, \quad \ddot{y} = -g\phi, \quad \ddot{z} = \tfrac{1}{m_Q}u_1, \\ &\ddot{\phi} = \tfrac{l_Q}{I_{xx}}u_2, \quad \ddot{\theta} = \tfrac{l_Q}{I_{yy}}u_3, \quad \ddot{\psi} = \tfrac{l_Q}{I_{zz}}u_4, \end{aligned} \tag{19}$$

where g is the gravitational acceleration, m_Q is the mass of the quadrotor, l_Q is the distance between the CoM of the quadrotor and the rotor, and I_{xx}, I_{yy} and I_{zz} represent the area moments of inertia about the principle axes in the body frame. The states are $(x, y, z, \dot{x}, \dot{y}, \dot{z}, \phi, \theta, \psi, \dot{\phi}, \dot{\theta}, \dot{\psi}) \in \mathbb{R}^{12}$, and the outputs are taken as the CoM of the quadrotor (x, y, z).

The task is to control the quadrotor steering to the target position while avoiding the two randomly moving spherical obstacles as shown in Fig. 2. The random movements of the first (red) obstacle are sampled from the distribution composed of Gaussian mixture distributions. Additionally, the random movements of the second (green) obstacle are sampled from a time-varying distribution. Specifically, the standard deviation used to generate historical data is 0.01, while the standard deviation increases to 0.3 for the real-time data generation during the runtime of motion control. In the offline planning stage, a reference trajectory is generated by using RRT*. Then, the quadrotor follows the reference trajectory by solving the OL-DR-NMPC problem. The MPC horizon is set to $K = 10$ with sampling time $T_s = 0.1\text{s}$, and we set $\alpha = 0.95$.

We compare the proposed method with the Wasserstein Distributionally Robust MPC (W-DR-MPC) method with radius $\theta = 0.01$, 0.03, and 0.05 [3]. As shown in Fig. 2 (a), at t =30, the quadrotor controlled by OL-DR-NMPC and W-DR-MPC all pass the first obstacle. Obviously, compared with the W-DR-MPC, the trajectory with OL-DR-NMPC is closer to the reference trajectory. For the second obstacle, the random movements become violent because of the increased standard deviation. Without online learning, the quadrotor controlled by W-DR-MPC method with $\theta = 0.01$ and 0.03 collides with the second obstacle as shown in Fig. 2 (b). Although the W-DR-MPC with $\theta = 0.05$ does not cause collision, such large θ often leads to a very conservative result when facing a general situation, e.g., the first obstacle. In contrast, the OL-DR-NMPC can learn the changes of the distribution through online learning, and keeps a safe distance from the obstacle to avoid collision. The trajectories at t = 100 are shown in Fig. 2 (c).

Fig. 3 shows the cost $\sum_{k=0}^{K-1} r(x_k, u_k) + q(x_K)$ of the four controllers at each time. Table I shows the cost before t = 50, the total cost, the success rate of collision-free navigation for 100 simulations and the average computation time at each time. It can be observed that, the OL-DR-NMPC not only enjoys less conservative control performance by using the fined-grained distribution information, but also avoids collision even facing time-varying distributions. Table I shows that the OL-DR-NMPC can reduce the cost by an average of 51.88% compared with the W-DR-MPC with $\theta = 0.05$. Additionally, observing the result of 100 simulations, the proposed method can always steer to the target position without collision. In contrast, the W-DR-MPC method with $\theta = 0.05$ has 22% probability of colliding with the second obstacle without online learning, which is much more than the prescribed tolerance (5%) and cannot be tolerated in practice. Further, due to the efficiency of the online variational inference algorithm and the sequential penalized relaxation algorithm, the computation time for OL-DR-NMPC is lower than that for W-DR-MPC.

VI. Conclusion

In this paper, an OL-DR-NMPC based motion control framework was proposed for mobile robots to avoid collisions in environments with randomly moving obstacles. We reformulated the DR-CVaR constraints as LMI and BMI constraints and developed a sequential penalized relaxation algorithm for computational tractability. The simulation results demonstrated that the proposed method enjoyed less conservative control performance and lower computational overhead meanwhile ensuring safety even facing time-varying distribution.

Appendix. Proof of Theorem 1

Proof. By (8) and (9), the DR-CVaR in (12) can be rewritten.

$$\begin{aligned} &\sup_{\upsilon\in\mathbb{D}} \mathrm{CVaR}_{\alpha}^{\upsilon}\left[\mathcal{L}(y, O+\omega)\right] \\ &= \sup_{\upsilon\in\mathbb{D}} \inf_{z\in\mathbb{R}} \left(z + \tfrac{1}{1-\alpha}\mathbb{E}^{\upsilon}\left[\left(-\|y-O-\omega\|_2^2 - z\right)^{+}\right]\right) \\ &\le \inf_{z\in\mathbb{R}} \left(z + \tfrac{1}{1-\alpha}\sup_{\upsilon\in\mathbb{D}}\mathbb{E}^{\upsilon}\left[\left(-\|y-O-\omega\|_2^2 - z\right)^{+}\right]\right) \end{aligned} \tag{20}$$

where the inequality follows from the minimax inequality.

Based on the definition of the ambiguity set (10), the worst-case expectation problem $\sup_{\upsilon\in\mathbb{D}}\mathbb{E}^{\upsilon}\left[\left(-\|y-O-\omega\|_2^2 - z\right)^{+}\right]$ can be rewritten as the following optimization problem.

$$\begin{aligned} &\sup_{\rho_1,\ldots,\rho_m\in\mathcal{P}(\mathbb{W})} \sum_{j=1}^{m}\gamma_j \int_{\mathbb{W}} \max\left(-\|y-O-\omega\|_2^2 - z, 0\right)\rho_j(d\omega) \\ &s.t. \left.\begin{aligned} &\int_{\mathbb{W}}\rho_j(d\omega) = 1 \\ &\int_{\mathbb{W}}\omega\rho_j(d\omega) = \mu_j \\ &\int_{\mathbb{W}}\omega\omega^T\rho_j(d\omega) \le \Sigma_j + \mu_j\mu_j^T \end{aligned}\right\}\forall j = 1,\ldots,m \end{aligned} \tag{21}$$

The dual problem of (21) can be rewritten as follows.

$$\min_{t_j,\xi_j,\Omega_j} \sum_{j=1}^{m}\gamma_j\left\{t_j + \mu_j^T\xi_j + \left(\Sigma_j + \mu_j\mu_j^T\right)\bullet\Omega_j\right\} \tag{22a}$$

$$\begin{aligned} s.t.\ & \|y-O-\omega\|_2^2 + z + t_j \\ & + \omega^T\xi_j + \omega^T\Omega_j\omega \ge 0, \forall\omega\in\mathbb{W}, \forall j \end{aligned} \tag{22b}$$

$$t_j + \omega^T\xi_j + \omega^T\Omega_j\omega \ge 0, \forall\omega\in\mathbb{W}, \forall j \tag{22c}$$

$$\Omega_j \ge 0, \forall j \tag{22d}$$

where t_j, ξ_j and Ω_j are the dual variables corresponding to the constraints in the ambiguity set. For matrices A and B, $A\bullet B$ represents their Frobenius inner product. Constraints (22b) and (22c) are further reformulated follow.

Constraints (22b) can be rewritten as follows.

$$\min_{\omega\in\mathbb{W}}\left\{\begin{aligned} &\|y-O\|_2^2 + z + t_j \\ &+\omega^T\left[\xi_j - 2(y-O)\right] + \omega^T\left(\Omega_j + I\right)\omega \end{aligned}\right\} \ge 0, \forall j \tag{23}$$

Based on the duality of convex quadratic programs and Schur complements [14], Constraint (23) can be reformulated as the following semidefinite constraint.

$$\begin{pmatrix} \Omega_j + I & \tfrac{1}{2}\left(\xi_j + H^T\varphi_j\right) - (y-O) \\ \tfrac{1}{2}\left(\xi_j + H^T\varphi_j\right)^T - (y-O)^T & z + t_j - h^T\varphi_j + \|y-O\|_2^2 \end{pmatrix} \ge 0, \forall j \tag{24}$$

where $\varphi_j \ge 0$ is the dual variables corresponding to the constraint $H\omega \le h$. Similarly, Constraint (22c) can be reformulated as semidefinite constraint below.

$$\begin{pmatrix} \Omega_j & \tfrac{1}{2}\left(\xi_j + H^T\lambda_j\right) \\ \tfrac{1}{2}\left(\xi_j + H^T\lambda_j\right)^T & t_j - h^T\lambda_j \end{pmatrix} \ge 0, \forall j \tag{25}$$

where $\lambda_j \ge 0$ is the dual variables corresponding to the constraint $H\omega \le h$. Substituting (22), (24) and (25) into (20), we can obtain the result of the theorem. ☐

References

[1] G. Zhang, Y. Deng, W. Zhang and C. Huang, "Novel DVS guidance and path-following control for underactuated ships in presence of multiple static and moving obstacles," *Ocean Engineering*, vol. 170, pp. 100-110, 2018.

[2] L. Blackmore, M. Ono, and B. C. Williams, "Chance-constrained optimal path planning with obstacles," *IEEE Transactions on Robotics*, vol. 27, no. 6, pp. 1080–1094, 2011.

[3] A. Hakobyan, G. C. Kim, and I. Yang, "Risk-aware motion planning and control using CVaR-constrained optimization," *IEEE Robotics and Automation Letters*, vol. 4, no. 4, pp. 3924–3931, 2019.

[4] A. Hakobyan, and I. Yang, "Wasserstein distributionally robust motion control for collision avoidance using conditional value-at-risk," *IEEE Transactions on Robotics*, vol. 38, no. 2, pp. 939-957, 2022.

[5] A. Hakobyan, and I. Yang. "Distributionally robust risk map for learning-based motion planning and control: A semidefinite programming approach," *IEEE Transactions on Robotics*, vol. 39, no. 1, pp. 718-737, 2023.

[6] S. Karaman and E. Frazzoli, "Sampling-based algorithms for optimal motion planning," *The International Journal of Robotics Research*, vol. 30, no. 7, pp. 846–894, 2011.

[7] R. Funada, M. Santos, and T. Gencho, "Visual Coverage Maintenance for Quadcopters Using Nonsmooth Barrier Functions," in *2020 IEEE International Conference on Robotics and Automation (ICRA)*, 2020.

[8] R. Gomes, M. Welling, and P. Perona, "Incremental Learning of Nonparametric Bayesian Mixture Models," in *2008 IEEE Conference on Computer Vision and Pattern Recognition*, 2008.

[9] C. Ning, and F. You. "Online learning based risk-averse stochastic MPC of constrained linear uncertain systems," *Automatica*, vol. 125, pp. 109402, 2021.

[10] M. Kheirandishfard, F. Zohrizadeh, and R. Madani, "Convex relaxation of bilinear matrix inequalities part i: Theoretical results, " in *2018 IEEE Conference on Decision and Control (CDC)*, 2018.

[11] A. Shapiro, "First and second order analysis of nonlinear semidefinite programs," *Math. Program*, vol. 77, no. 1, pp. 301–320, 1997.

[12] R. Madani, M. Kheirandishfard, and J. Lavaei, "Penalized semidefinite programming for quadratically-constrained quadratic optimization," *Journal of Global Optimization*, vol. 78, pp. 423-451, 2020.

[13] M. Kheirandishfard, F. Zohrizadeh, and R. Madani, "Convex relaxation of bilinear matrix inequalities part ii: Applications to optimal control synthesis," in *2018 IEEE Conference on Decision and Control*, 2018.

[14] S. Boyd, and L. Vandenberghe. *Convex optimization*. Cambridge university press, 2004.