2023 62nd IEEE Conference on Decision and Control (CDC)
December 13-15, 2023. Marina Bay Sands, Singapore

Risk-Constrained Control of Mean-Field Linear Quadratic Systems

Masoud Roudneshin, Saba Sanami, and Amir G. Aghdam

***Abstract*— The risk-neutral LQR controller is optimal for stochastic linear dynamical systems. However, the classical optimal controller performs inefficiently in the presence of low-probability yet statistically significant (risky) events. The present research focuses on infinite-horizon risk-constrained linear quadratic regulators in a mean-field setting. We address the risk constraint by bounding the cumulative one-stage variance of the state penalty of all players. It is shown that the optimal controller is affine in the state of each player with an additive term that controls the risk constraint. In addition, we propose a solution independent of the number of players. Finally, simulations are presented to verify the theoretical findings.**

I. Introduction

The performance evaluation of dynamical systems in the optimal control framework has long been studied in the literature [1]–[3]. Specifically, in the linear quadratic regulator (LQR) with noisy inputs, the focus is on minimizing the expected cumulative time-average quadratic cost, also known as a risk-neutral setting [4]. However, such a risk-neutral framework often exhibits unsatisfactory performance in real-world control systems. For instance, there exists a rich body of research to address risk in different areas, including robotics [5], [6], financial systems [7], [8], power grids [9], [10], and multi-agent networks [12], [13]. Moreover, neglecting the effect of low-probability severe external events may lead to catastrophic consequences in dynamic systems, like crashing in a flock of UAVs or an autonomous vehicle hitting other vehicles and pedestrians.

There has been an increasing interest in the research community recently in the risk assessment of dynamical systems by deriving closed-form solutions for a single-agent setting [11], [16], [17]. Specifically, by solving a set of Riccati and fixed-point equations, one can obtain an affine form of the policy to meet the system's constraints. However, in the control of a large number of agents, such a method may not provide sufficient efficacy.

This research considers the problem of exchangeable agents (players) in a mean-field setting. In such a setting, all agents have similar dynamics, and the players' states evolve as a linear function of their previous states and the overall average state. Using the results in mean-field theory, we show that the required Riccati equation (whose size increases with the number of players) can be decomposed into two Riccati equations with the same dimension as the agents' states. Furthermore, we propose a primal-dual algorithm to solve the problem iteratively.

The rest of the paper is organized as follows. In Section II, we present some preliminaries and formulate the problem. The solution to the optimization problem is derived in Section III, followed by simulations to validate the results in Section IV. Finally, some concluding remarks and directions for future research are given in Section V.

This work has been supported by the Natural Sciences and Engineering Research Council of Canada (NSERC) under grant RGPIN-262127-17.

Masoud Roudneshin, Saba Sanami and Amir G. Aghdam are with the Department of Electrical and Computer Engineering, Concordia Univerity, Montreal, QC, Canada. Email: masoud.roudneshin.concordia.ca, saba.sanami@mail.concordia.ca, aghdam@ieee.org

II. Problem Formulation

Throughout the paper, $\mathbb{R}$, $\mathbb{R}_{>0}$ and $\mathbb{N}$ represent the sets of real, positive real and natural numbers, respectively. Given any $n \in \mathbb{N}$, $\mathbb{N}_n$, and $\mathbf{I}_{n\times n}$ denote the finite set $\{1,\dots,n\}$, and the $n\times n$ identity matrix, respectively. $\|\cdot\|$ is the spectral norm of a matrix, $\mathrm{Tr}(\cdot)$ is the trace of a matrix, $\tau_{\min}(\cdot)$ is the minimum singular value of a matrix, $\rho(\cdot)$ is the spectral radius of a matrix, and $\mathrm{diag}(\Lambda_1,\Lambda_2)$ is the block diagonal matrix $[\Lambda_1 \quad 0; 0 \quad \Lambda_2]$, and $\mathrm{diag}(\Lambda)_{i=1}^k$ denotes a bloack-diagonal matrix with k times repetition of the matrix Λ. For vectors x, y and z, $\mathrm{vec}(x,y,z) = [x^\intercal, y^\intercal, z^\intercal]^\intercal$ is a column vector, $x_{1:t}$ denotes the vector $(x_1,...,x_t)$ and the operator $\otimes$ denotes the Kronecker product between two matrices of appropriate size. Also, the rectified linear function is denoted by the operator $[x]_+ = \max\{0,x\}$.

A. General Form of the Problem

Given $n \in \mathbb{N}$ players, let $x_t^i \in \mathbb{R}^{d_x}$, $u_t^i \in \mathbb{R}^{d_u}$ and $w_t^i \in \mathbb{R}^{d_x}$ denote, respectively, the state, action and local noise of player $i \in \mathbb{N}_n$ at time $t \in \mathbb{N}$, where $d_x, d_u \in \mathbb{N}$. Define the mean-state of the players as $\bar{x}_t := \frac{1}{n}\sum_{i=1}^n x_t^i$. The initial states $\{x_0^1,\dots,x_0^n\}$ are random with finite covariance matrices. The evolution of the state of any player $i \in \mathbb{N}_n$ at time $t \in \mathbb{N}$ is given by:

$$x_{t+1}^i = Ax_t^i + Bu_t^i + \bar{A}\bar{x}_t + \bar{B}\bar{u}_t + w_t^i, \tag{1}$$

where $\{w_t^i\}_{t=0}^\infty$ is an independent and identically distributed (i.i.d.) zero-mean noise process with a finite covariance matrix.

The per-step cost of all players at time $t \in \mathbb{N}$ is given by:

$$c_t = (\bar{x}_t)^\intercal \bar{Q}\bar{x}_t + (\bar{u}_t)^\intercal \bar{R}\bar{u}_t + \frac{1}{n}\sum_{i=1}^n (x_t^i)^\intercal Q x_t^i + (u_t^i)^\intercal R u_t^i, \tag{2}$$

where Q, $\bar{Q}$, R, and $\bar{R}$ are symmetric matrices with appropriate dimensions.

Definition 1. *Let $h_t^i = \{x_0^i, u_0^i, ..., x_{t-1}^i, u_{t-1}^i, x_t^i\}$ denote the history trajectory of player $i \in \mathbb{N}_n$. Then, the per-step* *risk factor for the ith player is defined as*

$$d_t^i = \Big((x_t^i)^\intercal Q x_t^i - \mathbb{E}\big[(\mathbf{x}_t^i)^\intercal Q \mathbf{x}_t^i | h_t^i\big]\Big)^2.$$

Assumption 1. *It is assumed hereafter that the pair (A, B) is stabilizable, the pair $(A, Q^{\frac{1}{2}})$ is detectable, and matrices Q and R are positive semi-definite and positive definite, respectively.*

Assumption 2. *The local noises $w_t^1, ..., w_t^n$ have the same distribution.*

Assumption 3. *The noise w_t^i for every player $i \in \mathbb{N}_n$ has a finite fourth-order moment, i.e., $\mathbb{E}\|w_t^i\|^4 < \infty$.*

In this paper, we consider the infinite-horizon risk-constrained LQR for a team of cooperative players to minimize a common cost. Also, it is desired to constrain the cumulative per-step risk of all players. This leads to the following constrained optimization problem

$$\text{minimize} \quad J = \limsup_{T\to\infty} \frac{1}{T}\mathbb{E}\Big[\sum_{t=0}^{T} c_t\Big] \tag{3a}$$

$$\text{s.t.} \quad (1) \text{ and } \forall i \in \mathbb{N}_n, \tag{3b}$$

$$J_c = \frac{1}{n}\sum_{i=1}^{n} \limsup_{T\to\infty} \frac{1}{T}\mathbb{E}\Big[\sum_{t=0}^{T} d_t^i\Big] \le \Gamma, \tag{3c}$$

where $\Gamma > 0$ is a predefined risk tolerance of the user.

Remark 1. *From [14], [15], when player $i \in \mathbb{N}_n$ at any time $t \in \mathbb{N}$ observes its local state x_t^i and the mean state $\bar{x}_t$, i.e. $\{x_{1:t}^i, \bar{x}_{1:t}\}$, an information structure called deep state sharing (DSS) is considered.*

Definition 2. *Let the control input of player $i \in \mathbb{N}_n$ at time t be denoted by $u_t^i = \phi_t^i(x_{1:t}^i, \bar{x}_{1:t})$. Define $\Phi^i := \{\phi_t^i\}_{t=1}^{\infty}$ and $\mathbf{\Phi}_n := \{\mathbf{\Phi}^1, \ldots, \mathbf{\Phi}^n\}$ as the control strategy of player i and that of all players, respectively.*

We now present the main problem of this article.

Problem 1. *Consider the risk-constrained mean-field LQR problem in* (3). *Given the system dynamics* (1), *find an optimal control strategy $\mathbf{\Phi}^*$ such that for any arbitrary control law $\mathbf{\Phi}$, the cost function* (3a) *under the constraints* (3b) *and* (3c) *satisfies the following inequality*

$$J(\mathbf{\Phi}^*) \le J(\mathbf{\Phi}).$$

III. Main Results

In this section, we propose a step by step solution to the optimization problem (3).

A. Problem Reformulation

Define a new transformed state $\tilde{x}_t^i = x_t^i - \bar{x}_t$ for player $i \in \mathbb{N}_n$. Define also the mean control input of all players as $\bar{u}_t := \frac{1}{n}\sum_{i=1}^n u_t^i$, and the transformed control input of player $i \in \mathbb{N}_n$ as $\tilde{u}_t^i = u_t^i - \bar{u}_t$. It follows from [15] that

$$\begin{aligned}\tilde{x}_{t+1}^i &= A\tilde{x}_t^i + B\tilde{u}_t^i + \tilde{w}_t^i \\ \bar{x}_{t+1} &= \mathcal{A}\bar{x}_t + \mathcal{B}\bar{u}_t + \bar{w}_t,\end{aligned} \tag{4}$$

where $\mathcal{A} = A + \bar{A}$, $\mathcal{B} = B + \bar{B}$, $\bar{w}_t := \frac{1}{n}\sum_{i=1}^n w_t^i$ and $\tilde{w}_t^i = w_t^i - \bar{w}_t$.

Next, define the first and second-order moments (mean and covariance) of each player's local noise as $m_1 = \mathbb{E}[w_t^i]$ and $M_2 = \mathbb{E}[(w_t^i - m_1^i)(w_t^i - m_1^i)^\intercal]$, respectively. Furthermore, let the next two higher order moments of the local noise be defined as

$$\begin{aligned}M_3 &= \mathbb{E}[(w_t^i - m_1^i)(w_t^i - m_1^i)^\intercal Q (w_t^i - m_1^i)], \\ M_4 &= \mathbb{E}[(w_t^i - m_1^i)^\intercal Q (w_t^i - m_1^i) - \text{Tr}(M_2 Q)]^2.\end{aligned}$$

Also, for future reference, define $\mathbb{m}_1 = \mathbb{E}[\tilde{w}_t^i]$ and $\mathbb{M}_1 = \mathbb{E}[(\tilde{w}_t^i - \mathbb{M}_1)(\tilde{w}_t^i - \mathbb{m}_1)^\intercal]$.

Lemma 1. *The risk-constrained optimization problem in* (3) *can be reformulated as*

$$\begin{aligned}&\text{minimize} && J = \limsup_{T\to\infty} \frac{1}{T}\mathbb{E}\Big[\sum_{t=0}^{T} c_t^i\Big] \\ &\text{s.t.} && (4) \text{ and}, \quad \forall i \in \mathbb{N}_n, \\ &&& \tilde{J}_c = J_{\bar{c}} + \sum_{i=1}^{n} \tilde{J}_c^i \le \Lambda\end{aligned} \tag{5}$$

where

$$J_{\tilde{c}}^i = \lim_{T\to\infty} \frac{1}{T}\mathbb{E}\Big[\sum_{t=0}^{T} \frac{4}{n}(\tilde{x}_t^i)^\intercal Q M_2 Q \tilde{x}_t^i\Big],$$

$$J_{\bar{c}} = \limsup_{T\to\infty} \frac{1}{T}\mathbb{E}\Big[\sum_{t=0}^{T} 4(\bar{x}_t)^\intercal Q M_2 Q \bar{x}_t + 4(\bar{x}_t)^\intercal Q M_3\Big],$$

and $\Lambda = \Gamma - m_4 + Tr(M_2 Q)^2$.

Proof. Using the results in [16], the constraint in (3c) can be reformulated as

$$J_c = \frac{1}{n}\sum_{i=1}^{n} \limsup_{T\to\infty} \frac{1}{T}\mathbb{E}\sum_{t=0}^{T} 4(x_t^i)^\intercal Q M_2 Q x_t^i + 4(x_t^i)^\intercal Q M_3.$$

The proof follows immediately by rewriting the above equation as $x_t^i = \tilde{x}_t^i + \bar{x}_t$, and on noting that $\sum_{i=1}^n \tilde{x}_t^i = 0$. ■

B. Primal-Dual Approach

To solve the constrained optimization problem (5), we use $\lambda \ge 0$ as the Lagrange multiplier. The Lagrangian can then be expressed as

$$\mathcal{L}(\mathbf{\Phi}, \lambda) = J + \lambda(\tilde{J}_c - \Lambda). \tag{6}$$

Definition 3. *Define matrices $Q_c = \frac{4}{n} Q M_2 Q$, $Q_{\bar{c}} = 4 Q M_2 Q$, $R_c = \frac{1}{n}R$, $Q_\lambda = \frac{1}{n}Q + \lambda Q_c$, $Q_{\bar{\lambda}} = Q + \bar{Q} + \lambda Q_{\bar{c}}$, and $S_\lambda = 4\lambda Q M_3$.*

Lemma 2. *The Lagrangian in* (6) *can be reformulated as*

$$\mathcal{L}(\mathbf{\Phi}, \lambda) = \bar{\mathcal{L}} + \sum_{i=1}^{n} \mathcal{L}^i,$$

where

$$\mathcal{L}^i = \limsup_{T\to\infty} \frac{1}{T}\mathbb{E}\left[(\tilde{x}_t^i)^\intercal (Q_\lambda)\tilde{x}_t^i + (\tilde{u}_t^i)^\intercal \frac{1}{n} R\tilde{u}_t^i\right],$$
$$\bar{\mathcal{L}} = \limsup_{T\to\infty} \frac{1}{T}\mathbb{E}\left[\bar{x}_t^\intercal (Q_{\bar{\lambda}})\bar{x}_t + S_\lambda \bar{x}_t + \bar{u}_t^\intercal (R+\bar{R})\bar{u}_t\right].$$

Proof. The result follows directly from Lemma 1, the definition of the per-step cost in (2), and on noting that $\sum_{i=1}^n \tilde{x}_t^i = 0$ and $\sum_{i=1}^n \tilde{u}_t^i = 0$. ■

To solve for the optimal value of the Lagrangian $\mathcal{L}^*$ in (6), we find the general form of the policies for a constant multiplier.

Theorem 1. *For a fixed multiplier λ, the optimal policy for each player is affine, such that*

$$u_t^i = -\theta(\lambda)x_t^i - (\bar{\theta}(\lambda) - \theta(\lambda))\bar{x}_t + \tau(\lambda) + \bar{\tau}(\lambda), \quad (7)$$

in which

$$\theta = -(R + B^\intercal PB)^{-1}B^\intercal PA,$$
$$\bar{\theta} = -(\mathcal{R} + \mathcal{B}^\intercal \mathcal{P}\mathcal{B})^{-1}\mathcal{B}^\intercal \mathcal{P}\mathcal{A},$$

and

$$\tau = -\frac{1}{2}(R + B^\intercal PB)^{-1}B^\intercal (2P\mathbb{m}_1 + g),$$
$$\bar{\tau} = -\frac{1}{2}(\mathcal{R} + \mathcal{B}^\intercal \mathcal{P}\mathcal{B})^{-1}\mathcal{B}^\intercal (2\mathcal{P}m_1 + \mathbf{g}),$$

where P, $\mathcal{P}$, g and $\mathbf{g}$ are obtained by solving the following recursive equations

$$P = Q_\lambda + A^\intercal PA - A^\intercal PB(R + B^\intercal PB)^{-1}B^\intercal PA,$$
$$\mathcal{P} = Q_{\bar{\lambda}} + \mathcal{A}^\intercal \mathcal{P}\mathcal{A} - \mathcal{A}^\intercal \mathcal{P}\mathcal{B}(\mathcal{R} + \mathcal{B}^\intercal \mathcal{P}\mathcal{B})^{-1}\mathcal{B}^\intercal \mathcal{P}\mathcal{A},$$
$$g^\intercal = (2\mathbb{m}_1^\intercal P + g^\intercal)(A - B\theta),$$
$$\mathbf{g}^\intercal = (2m_1^\intercal \mathcal{P} + \mathbf{g}^\intercal)(\mathcal{A} - \mathcal{B}\bar{\theta}) + 4\lambda(QM_3)^\intercal.$$

Proof. Define the generalized state and action of all agents in an augmented form as $\mathbf{x}_t = [\text{vec}(\tilde{x}_t^i)_{i=1}^n, \bar{x}_t]$ and $\mathbf{u}_t = [\text{vec}(\tilde{u}_t^i)_{i=1}^n, \bar{u}_t]$, respectively. Then, it follows that

$$\mathbf{x}_{t+1} = \mathbf{A}\mathbf{x}_t + \mathbf{B}\mathbf{u}_t,$$

where

$$\mathbf{A} = \text{diag}(\text{diag}(A)_{i=1}^n, \bar{A}), \quad \mathbf{B} = \text{diag}(\text{diag}(B)_{i=1}^n, \bar{B}).$$

Define the finite-horizon Lagrangian as the value function V_T and note that the results in Theorem 2 of [16] imply that the Lagrangian has a quadratic form as

$$V_T = \mathbf{x}_t^\intercal \mathbf{P}\mathbf{x}_t + \mathbf{g}\mathbf{x}_t + \mathbf{z}_t.$$

Instead of solving for the optimal policy in the larger state-space of $\mathbf{x}_t$, from Lemma 2, the value function can also be decomposed into a set of smaller value functions such that

$$V_T = \bar{V}_T + \sum_{i=1}^n \tilde{V}_T^i.$$

Since the Lagrangians $\bar{\mathcal{L}}$ and $\tilde{\mathcal{L}}^i$ have complete square forms, the minimization can be carried out over the smaller state space of $\tilde{x}_t^i$ and $\bar{x}_t$. Therefore, by employing dynamic programming, we have the following two recursive optimality equations

$$\tilde{V}_T^i = \min_{\tilde{u}_t}\big((\tilde{x}_t^i)^\intercal Q_\lambda \tilde{x}_t^i + \frac{1}{n}(\tilde{u}_t^i)^\intercal R\tilde{u}_t^i + \tilde{V}_{T+1}^i\big),$$
$$\bar{V}_T = \min_{\bar{u}_t}(\bar{x}_t^\intercal Q_{\bar{\lambda}}\bar{x}_t + \bar{u}_t^\intercal (R+\bar{R})\bar{u}_t + \bar{V}_{T+1}).$$

The proof follows by taking the derivative with respect to $\tilde{u}_t^i$ and $\bar{u}_t$ and using backward dynamic programming. ■

Remark 2 (Strong Duality)**.** *Using the results established in Theorem 2 of [17] and [18], there exists an optimal multiplier λ^* such that the policy*

$$u_t^i = -\theta(\lambda^*)x_t^i - (\bar{\theta}(\lambda^*) - \theta(\lambda^*))\bar{x}_t + \tau(\lambda^*) + \bar{\tau}(\lambda^*)$$

is the optimal solution to (5).

C. Solution of the Dual Problem with Subgradients

Since there is no optimality gap in the optimization problem (3), we can alternatively solve the following dual problem

$$\max_{\lambda\geq 0}\ D(\lambda) = \max_{\lambda\geq 0}\min_{\mathbf{u}}\ \mathcal{L}(\mathbf{u},\lambda)$$

which is also concave in λ. Let d denote the subgradient. Then, from the results in [19], [20], the subgradient of $D(\lambda)$ can be expressed as

$$d = \tilde{J}_c(\theta, \bar{\theta}, \lambda) - \Lambda.$$

The following theorem provides the explicit form of the constraints for deriving the subgradient vector.

Theorem 2. *Consider the stabilizing control input given by* (7)*. Then,*

$$J_{\tilde{c}}^i = \mathit{Tr}\left[P_{\tilde{c}}\big(\mathbb{M}_2 + (B\tau + \mathbb{m}_1)(B\tau + \mathbb{m}_1)^\intercal\big)\right] + g_{\tilde{c}}^\intercal (B\tau + \mathbb{m}_1),$$
$$J_{\bar{c}} = \mathit{Tr}\left[P_{\bar{c}}(\mathcal{B}\bar{\tau} + m_1)(\mathcal{B}\bar{\tau} + m_1)^\intercal\right] + g_{\bar{c}}^\intercal (\mathcal{B}\bar{\tau} + m_1),$$

where $P_{\tilde{c}}$ and P_c are the positive definite solutions of the following Lyapunov equations

$$P_{\tilde{c}} = \frac{4}{n}QM_2Q + (A - B\theta)^\intercal P_{\tilde{c}}(A - B\theta),$$
$$P_{\bar{c}} = 4QM_2Q + (\mathcal{A} - \mathcal{B}\bar{\theta})^\intercal P_{\bar{c}}(\mathcal{A} - \mathcal{B}\bar{\theta}),$$

where

$$g_{\tilde{c}}^\intercal = 2\big\{(B\tau + \mathbb{m}_1)^\intercal P_{\tilde{c}}(A - B\theta)\big\}\big(I - A + B\theta\big)^{-1},$$
$$g_{\bar{c}}^\intercal = 2\big\{(\mathcal{B}\bar{\tau} + m_1)^\intercal P_{\bar{c}}(\mathcal{A} - \mathcal{B}\bar{\theta}) + 2M_3^\intercal Q\big\}\big(I - \mathcal{A} + \mathcal{B}\bar{\theta}\big)^{-1}.$$

Proof. Define the relative value functions

$$V_{\tilde{c}}^i = \mathbb{E}\left[\sum_{t=0}^\infty \frac{4}{n}(\tilde{x}_t^i)^\intercal QM_2Q\tilde{x}_t^i - J_{\tilde{c}}^i\right],$$
$$V_{\bar{c}} = \mathbb{E}\left[\sum_{t=0}^\infty 4(\bar{x}_t)^\intercal QM_2Q\bar{x}_t + 4(\bar{x}_t)^\intercal QM_3 - J_{\bar{c}}\right].$$

Using backward dynamic programming, it can be shown that such value functions have a quadratic form, i.e. $V_{\tilde{c}}^i =$

$(\tilde{x}_t^i)^\intercal P_{\tilde{c}} \tilde{x}_t^i + g_{\tilde{c}}^\intercal \tilde{x}_t^i + z_{\tilde{c}}$ and $V_{\tilde{c}}^i = \bar{x}_t^\intercal P_{\tilde{c}} \bar{x}_t + g_{\tilde{c}}^\intercal \bar{x}_t + z_{\tilde{c}}$. Using the Bellman equation for $V_{\tilde{c}}^i$ one has

$$
\begin{aligned}
V_{\tilde{c}}^i &= (\tilde{x}_t^i)^\intercal P_{\tilde{c}} \tilde{x}_t^i + g_{\tilde{c}}^\intercal \tilde{x}_t^i + z_{\tilde{c}} = \frac{4}{n} (\tilde{x}_t^i)^\intercal Q M_2 Q \tilde{x}_t^i \\
&- J_{\tilde{c}}^i + \mathbb{E}[g_{\tilde{c}}^\intercal \big((A - B\theta)\tilde{x}_t^i + B\tau + \tilde{w}_t^i\big)] + z_{\tilde{c}} \\
&+ \mathbb{E}[(A - B\theta)\tilde{x}_t^i + B\tau + \tilde{w}_t^i]^\intercal P_{\tilde{c}} [(A - B\theta)\tilde{x}_t^i + B\tau + \tilde{w}_t^i] \\
&= (\tilde{x}_t^i)^\intercal \left[\frac{4}{n} Q M_2 Q + (A - B\theta)^\intercal P_{\tilde{c}} (A - B\theta)\right] (\tilde{x}_t^i) \\
&\left[2(B\tau + m_1)^\intercal P_{\tilde{c}} (A - B\theta) + g_{\tilde{c}}^\intercal (A - B\theta)\right] (\tilde{x}_t^i) \\
&- J_{\tilde{c}}^i + z_{\tilde{c}} + g_{\tilde{c}}^\intercal (B\tau + \mathbb{m}_1) + \\
&\text{Tr}\left[P_{\tilde{c}} (\mathbb{M}_2 + (B\tau + \mathbb{m}_1)(\mathbb{M}_2 + (B\tau + \mathbb{m}_1)^\intercal\right],
\end{aligned}
$$

Then, it follows that

$$
J_{\tilde{c}}^i = \text{Tr}\left[P_{\tilde{c}} (\mathbb{M}_2 + (B\tau + \mathbb{m}_1)(\mathbb{M}_2 + (B\tau + \mathbb{m}_1)^\intercal\right] + g_{\tilde{c}}^\intercal (B\tau + \mathbb{m}_1).
$$

Using a similar argument, $V_{\bar{c}}$ can be written as

$$
\begin{aligned}
V_{\bar{c}} &= \bar{x}_t^\intercal P_{\bar{c}} \bar{x}_t + g_{\bar{c}}^\intercal \bar{x}_t = 4\bar{x}_t^\intercal Q M_2 Q \bar{x}_t + 4 M_3^\intercal Q \bar{x}_t \\
&- J_{\bar{c}} + \mathbb{E}[g_{\bar{c}}^\intercal \big((\mathcal{A} - \mathcal{B}\bar{\theta})\bar{x}_t + \mathcal{B}\tau + \bar{w}_t\big)] + z_{\bar{c}} \\
&+ \mathbb{E}[(\mathcal{A} - \mathcal{B}\bar{\theta})\bar{x}_t + \mathcal{B}\tau + \bar{w}_t]^\intercal P_{\bar{c}} [(\mathcal{A} - \mathcal{B}\bar{\theta}))\bar{x}_t + \mathcal{B}\tau + \bar{w}_t] \\
&= \bar{x}_t^\intercal \left[Q M_2 Q + (\mathcal{A} - \mathcal{B}\bar{\theta})^\intercal P_{\bar{c}} (\mathcal{A} - \mathcal{B}\bar{\theta})\right] \bar{x}_t \\
&\left[2(\mathcal{B}\tau + m_1)^\intercal P_{\bar{c}} (\mathcal{A} - \mathcal{B}\bar{\theta}) + 4 M_3^\intercal Q + g_{\bar{c}}^\intercal (\mathcal{A} - \mathcal{B}\bar{\theta})\right] \bar{x}_t \\
&+ \text{Tr}\left[P_{\bar{c}} (\mathcal{B}\tau + m_1)(\mathcal{B}\tau + m_1)^\intercal\right] + g_{\bar{c}}^\intercal (\mathcal{B}\tau + m_1) - J_{\bar{c}} + z_{\bar{c}},
\end{aligned}
$$

which yields

$$
J_{\bar{c}} = \text{Tr}\left[P_{\bar{c}} (\mathcal{B}\bar{\tau} + m_1)(\mathcal{B}\bar{\tau} + m_1)^\intercal\right] + g_{\bar{c}}^\intercal (\mathcal{B}\bar{\tau} + m_1).
$$

■

From Theorem 2, we can compute $\tilde{J}_c = J_{\bar{c}} + \sum_{i=1}^n J_{\tilde{c}}^i$ and then find the subgradients accordingly. Algorithm 1 describes the proposed primal-dual method to solve the optimization problem in (3).

Algorithm 1 Primal-Dual Algorithm for Risk-Constrained Mean-field LQR

Input: Initial λ_0, step size η

Iteration counter k
for $k = 1, 2, ...$ **do**
  Obtain $u_t = \text{argmin } \mathcal{L}(\mathbf{u}_t, \boldsymbol{\lambda}_k)$ from Theorem 1
  Compute d_k from Theorem 2
  Update the multiplier $\lambda_{k+1} = [\lambda_k + \eta_k . d_k]_+$
end for

Remark 3. *Since the policy in* (7) *is stabilizable, the subgradients' and multipliers' vectors have upper bounds [11].*

Remark 4. *Since the subgradients and multipliers are upper bounded, using an argument analogous to that in Theorem 3 in [17], Algorithm 1 converges to the optimal policy after sufficient iterations.*

IV. Simulations

We validate the proposed method using numerical simulations on a low-inertia microgrid (MG) system. Consider the load frequency problem (LFC) with risk constraints on the agents' frequency and mean state. The MGs exchange information with each other through the mean state of the system.

Consider microgrids in n areas. Let $\Delta P_{\text{tie},i}$ and Δf_i denote the power inflow and the frequency deviation corresponding to the ith microgrid. We assume that this power flow is proportional to the discrepancy between the frequency deviation of each area and the mean frequency deviation of all areas, i.e.

$$
\Delta P_{\text{tie},i} = \int K_{\text{tie},i} (\Delta f_i - \Delta \bar{f}) dt
$$

In addition, the control signal of the ith area is the sum of two terms given below

$$
\Delta u_{tot,i} = \Delta P_{f,i} + \Delta P_{C,i},
$$

where $\Delta P_{f,i} = -\frac{1}{R_i} \Delta f_i$, and $\Delta P_{C,i}$ denotes the automatic generation control (AGC). These two controls specify the output power of the microgrid at the ith area denoted by $\Delta P_{G,i}$. The other state variable is the area control error (ACE) denoted by $z_i := \beta \Delta f_i + \Delta P_{\text{tie},i}$ with the bias factor $\beta_i = D_i + \frac{1}{R_i}$.

The overall state of each microgrid is

$$
x^i = [\Delta f_i, \Delta P_{G,i}, \Delta P_{\text{tie},i}, \int z_i]^\intercal.
$$

The dynamics of the system is

$$
x_{t+1}^i = A x_t^i + \bar{A} \bar{x}_t + B u_t^i,
$$

where

$$
A = \begin{bmatrix} -\frac{1}{T_p} & \frac{K_p}{T_p} & -\frac{K_p}{T_p} & 0 \\ -\frac{K_t}{RT_t} & -\frac{1}{T_t} & 0 & 0 \\ 0 & 0 & 0 & 1 \\ \beta & 0 & 1 & 0 \end{bmatrix}
$$

TABLE I
Simulation Parameters from [21]

Damping Factor	D	16.66	MW/Hz
Speed Droop	R	1.2×10^{-3}	Hz/MW
Turbine Static Gain	K_t	1	MW/MW
Turbine Time Constant	T_t	0.3	s
Area Static Gain	K_p	0.06	Hz/MW
Area Time Constant	T_p	24	s
Tie-line Coefficient	K_{tie}	850	MW/Hz

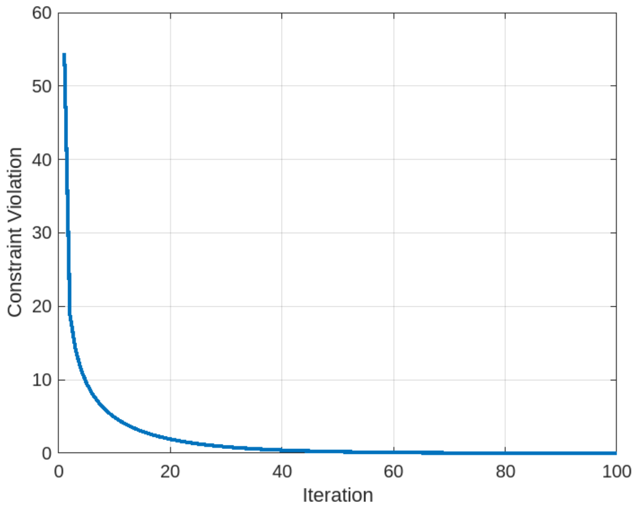

Fig. 1. Constraint violation with iterations for the microgrid problem

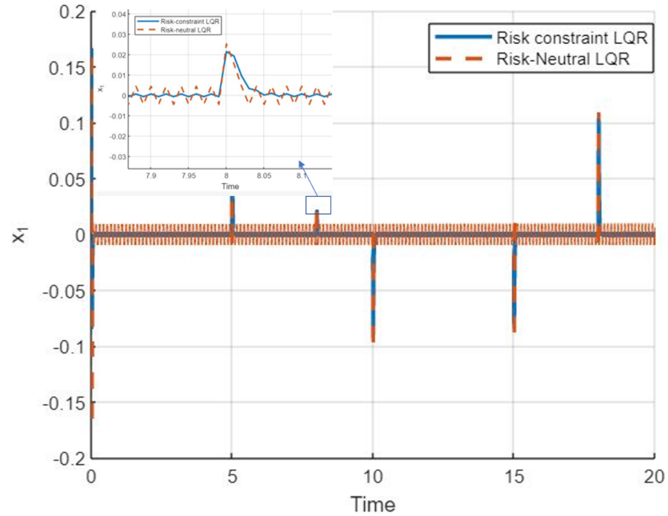

Fig. 2. Comparison of system state using the risk-neutral controller and the proposed one

$$\bar{A} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ K_{tie} & 0 & 0 & 0 \\ \beta & 0 & 0 & 0 \end{bmatrix}$$

$$B = \begin{bmatrix} 0 \\ 0 \\ \frac{K_t}{T_t} \\ 0 \end{bmatrix}$$

We use the parameters in Table I from [21]. Also, we select $Q = \text{diag}(800, 80, 80, 4000)$, $R = 5$, $\Lambda = 100$ and $\eta = 0.05$. Fig. 1 shows the constraint violation, where it is observed that as the number of iterations grows, the constraint violation tends to zero. In other words, the control law resulting from the algorithm minimizes the common cost function of players while not violating the system's constraint.Fig. 2 illustrates the variation of the first state, i.e. Δf_i, with high-amplitude disturbances at different time instants. We compare the performance of the proposed method with that of the risk-neutral control approach. It is observed that our method results in less state fluctuations and smaller overshoot in the presence of high-amplitude disturbances, confirming the results developed in Theorems 1 and 2.

V. Conclusions

We proposed a computationally-efficient method to tackle the problem of risk-constrained control of mean-field linear quadratic systems. The method only requires the solution of two Riccati equations and is independent of the number of players. This is a feature that is essential in controlling a multi-agent system of large size. The application of policy gradient methods as an alternative approach and considering individual constraints for the players are two interesting topics for the extension of the current research.

References

[1] Y.S. Wang, N. Matni, and J. C. Doyle, "Localized LQR Optimal Control," in Proceedings of the *53rd IEEE Conference on Decision and Control*, Dec. 2014, pp. 1661-1668.

[2] D. Kleinman, "Optimal Control of Linear Systems with Time-delay and Observation Noise," *IEEE Transactions on Automatic Control*, 1969, vol. 14, no. 5.

[3] P. P. Khargonekar and M. A. Rotea, "Multiple Objective Optimal Control of Linear Systems: The Quadratic Norm Case," *IEEE Transactions on Automatic Control*, 1991, vol. 36, no. 1.

[4] L. C. Evans, "An Introduction to Mathematical Optimal Control Theory," Lecture notes available at *http://math. berkeley.edu/ evans/control. course. pdf*, 1983.

[5] A. Majumdar, M. Pavone, "How Should a Robot Assess Risk? Towards an Axiomatic Theory of Risk in Robotics," in *Robotics Research*, 2020, pp. 75-84.

[6] A. A. Pereira, J. Binney, G. A. Hollinger, and G. S. Sukhatme, "Risk-Aware Path Planning for Autonomous Underwater Vehicles using Predictive Ocean Models," *Journal of Field Robotics,* Sep. 2013, vol. 30, no. 5, pp. 741–762.

[7] C. S Tapiero, "Applied Stochastic Models and Control for Finance and Insurance," *Springer Science & Business Media*, Dec. 2012.

[8] Z. Wu, "Using Machine Learning Approach to Evaluate the Excessive Financialization Risks of Trading Enterprises," *Computational Economics*. Apr. 2022, 59(4):1607-1625.

[9] I. Zografopoulos, J. Ospina, X. Liu, C. Konstantinou, "Cyber-Physical Energy Systems Security: Threat Modeling, Risk Assessment, Resources, Metrics, and Case Studies," *IEEE Access*. Feb. 2021, 9:29775-29818.

[10] S. Bruno, S. Ahmed, A. Shapiro, and A. Street, "Risk-Neutral and Risk-Averse Approaches to Multistage Renewable Investment Planning under Uncertainty," *European Journal of Operational Research*, May 2016, vol. 250, no. 3, pp. 979–989.

[11] F. Zhao, K.You, T. Basar, "Global Convergence of Policy Gradient Primal-Dual Methods for Risk-Constrained LQRs," *IEEE Transactions on Automatic Control*. Jan. 2023.

[12] V. Fathi, J. Arabneydi, A. G. Aghdam, "Reinforcement Learning in Linear Quadratic Deep Structured Teams: Global Convergence of Policy Gradient Methods," in Proceedings of the *59th IEEE Conference on Decision and Control*, Dec. 2020, pp. 4927-4932.

[13] N. Li, X. Li, Y. Shen, Z. Bi, M. Sun, "Risk Assessment Model Based on Multi-Agent Systems for Complex Product Design," *Information Systems Frontiers*, Apr. 2015, 17(2): 363-385.

[14] M. Roudneshin, J. Arabneydi, A. G. Aghdam, "Reinforcement Learning in Nonzero-sum Linear Quadratic Deep Structured Games: Global Convergence of Policy Optimization," in Proceedings of the

59th IEEE Conference on Decision and Control, Dec. 2020, pp. 512-517.

[15] J. Arabneydi and A. G. Aghdam, “Deep Teams: Decentralized Decision Making with Finite and Infinite Number of Agents,” *IEEE Transactions on Automatic Control*, Jan. 2020, 65(10): 4230-4245.

[16] A. Tsiamis, D. S. Kalogerias, L. F. Chamon, A. Ribeiro,G. J. Pappas, “Risk-constrained Linear-Quadratic Regulators,” in Proceedings of the *59th IEEE Conference on Decision and Control*. Dec. 2020, pp. 3040-3047.

[17] F. Zhao, K. You, T. Basar, “Infinite-horizon Risk-constrained Linear Quadratic Regulator with Average Cost,” in Proceedings of the *60th IEEE Conference on Decision and Control*, Dec. 2021, pp. 390-395.

[18] F. Zhao, K. You, T. Basar, “Global Convergence of Policy Gradient Primal-dual Methods for Risk-constrained LQRs,” *arXiv preprint arXiv:*2104.04901. Apr. 2021.

[19] Y. Nesterov, “Introductory Lectures on Convex Optimization: A Basic Course,” *Springer Science & Business Media*, Dec. 2013, vol. 87.

[20] S. P. Boyd, and L. Vandenberghe, “Convex Optimization,” *Cambridge University Press*, 2004.

[21] E. E. Vlahakis, L. D. Dritsas, G. D. Halikias, “Distributed LQR Design for Identical Dynamically Coupled Systems: Application to Load Frequency Control of Multi-area Power Grid,” in Proceedings of the *58th IEEE Conference on Decision and Control*. Dec. 2019, pp. 4471-4476.