Quadratic Truncated Random Return in Distributional LQR: Positive Definiteness, Density, and Log-Concavity

Ruyi Teng, Dan Wang, Wei Chen, and Yulong Gao

***Abstract*— Distributional linear quadratic regulator (LQR) is a new framework that integrates the distributional reinforcement learning and classical LQR, which offers a new way to study the random return instead of the expected cost. Unlike iterative approximation using dynamic programming in the DRL, a closed-form expression for the random return can be exactly characterized in the distributional LQR, which is defined over infinitely many random variables. In recent work [1, 2], it has been shown that this random return can be well approximated by a finite number of random variables, which we call truncated random return. In this paper, we study the truncated random return in the distributional LQR. We show that the truncated random return can be naturally expressed in the quadratic form. We develop a sufficient condition for the positive definiteness of the block symmetric matrix in the quadratic form and provide the lower and upper bounds on the eigenvalues of this matrix. We further show that in this case, the truncated random return follows a positively weighted non-central chi-square distribution if the random disturbances admits Gaussian, and its cumulative distribution function is log-concave if the probability density function of the random disturbances is log-concave.**

***Index Terms*— Distributional reinforcement learning; linear quadratic regulator; weighted non-central chi-square distribution; log-concavity**

I. Introduction

Linear Quadratic Regulator (LQR) [3] is one of the most fundamental problems in systems and control. Its importance can be reflected in a large number of practical applications, e.g., [4, 5]. When the additive disturbance is stochastic, the optimal controller that minimizes the expected quadratic cost is a linear feedback controller, where the optimal feedback gain can be obtained by solving the Riccati equation. Inspired by distributional reinforcement learning (DRL) [6], *distributional LQR* was developed in the recent work [1, 2] to take into account the randomness of total return, beyond its expectation.

The motivation of developing distributional LQR is the benefits of evaluating and exploiting the return distribution in the decision making that accounts for uncertainty, risk, and multi-modal outcomes [7, 2]. In the DRL, it has been shown or demonstrated that the use of return distribution not only improves resilience against unexpected deviations but also improves the efficiency of exploration during learning, guiding policies toward reduced variance and stabilized behavior [8]. A key result of distributional LQR is the closed-form expression of the random return [1, 2], which is intractable for other stochastic systems, including finite Markov decision processes.

The objective of this paper is to study the truncated random return in the distributional LQR. It was shown in [1, 2] that truncated random return converges pointwise in distribution to the un-truncated random return while providing improved computational tractability. We show that the truncated random return can be naturally expressed as a quadratic form in random variables, which is a key concept in statistics and applied mathematics [9, 10]. To the best of our knowledge, it is the first time to explore the positive definiteness of the block matrix in the quadratic truncated random return. This fundamental property reveals some important results: explicit from of density function and log-concavity. Our main contributions are summarized as follows.

(1) We prove that the symmetric block matrix in the quadratic form of the truncated random return is positive definite under a sufficient condition.
(2) We show that if the random disturbance is zero-mean Gaussian, the truncated random return follows a positively weighted non-central chi-square distribution.
(3) We prove that if the probability density function (PDF) of the random disturbances is log-concave, the cumulative distribution function of the truncated random return is also log-concave.

II. Preliminaries and Problem Statement

A. Notation

We denote by $\mathbb{R}$ the set of real numbers and $\mathbb{N}$ the set of natural numbers. For a symmetric matrix P, the notation $P \succ 0$ means that P is positive definite. For a matrix $Q \in \mathbb{R}^{n\times n}$, we denote by $\|Q\|$ the spectral norm. To indicate that two random variables Z_1 and Z_2 are equal in distribution, we use the notation $Z_1 \overset{D}{=} Z_2$. For a random variable Z, $\mathbb{E}[Z]$ denotes its expectation. For a real symmetric matrix H, $\lambda_{\min}(H)$ denotes its minimal eigenvalue.

Lemma 2.1: [11, Theorem 2] Let $H = [H_{ij}]$ be a block matrix where $H_{ij} \in \mathbb{R}^{n\times n}$, $i, j = 1, \ldots, N$. For each eigenvalue λ of H, there exists at least one $1 \le i \le N$

R. Teng and Y. Gao are with the Department of Electrical and Electronic Engineering, Imperial College London, United Kingdom. Email: ruyi.teng24@imperial.ac.uk, yulong.gao@imperial.ac.uk.

D. Wang is with the School of Robotics and Automation, Nanjing University, Suzhou, China. danwang@nju.edu.cn

W. Chen is with the Department of Mechanics and Engineering Science, Peking University, Beijing 100871, China, and also with the State Key Laboratory for Turbulence and Complex Systems, Peking University, Beijing 100871, China. Email: w.chen@pku.edu.cn.

such that

$$(\|(H_{ii}-\lambda I)^{-1}\|)^{-1} \leq \sum_{j=1,j\neq i}^{N} \|H_{ij}\|. \tag{1}$$

Definition 2.1: Let $H=[H_{ij}]$ be a block matrix where $H_{ij}\in\mathbb{R}^{n\times n}$, $i,j=1,\ldots,N$. If the diagonal submatrices H_{ii} are nonsingular and

$$\left(\|H_{ii}^{-1}\|\right)^{-1} \geq \sum_{j=1,j\neq i}^{N} \|H_{ij}\|, \forall i=1,\ldots,N, \tag{2}$$

then H is block diagonally dominant. If the strict inequality is valid, then H is block strictly diagonally dominant.

B. Canonical discounted LQR

Consider a discrete time LTI system:

$$x_{t+1}=Ax_t+Bu_t+v_t$$

where $x_t\in\mathbb{R}^n$, $u_t\in\mathbb{R}^p$, and $v_t\in\mathbb{R}^n$ are the state, control input, and exogenous disturbance, respectively. The exogenous disturbances v_t are assumed to be independent and identically distributed (i.i.d.) with bounded moments, sampled from an arbitrary distribution $\mathcal{D}$.

The discounted LQR is to find a feedback controller $u_t=\pi(x_t)$ by minimizing

$$J(u)=\mathbb{E}\left[\sum_{t=0}^{\infty}\gamma^t\left(x_t^\top Qx_t+u_t^\top Ru_t\right)\right], \tag{3}$$

where $Q\in\mathbb{R}^{n\times n}$ and $R\in\mathbb{R}^{p\times p}$ are symmetric positive definite matrices, $\gamma\in(0,1)$ is the discount factor. Given an initial state $x_0=x$ and a linear feedback controller $u_t=Kx_t$, the value function is defined as $V_\pi(x):=\mathbb{E}\left[\sum_{t=0}^{\infty}\gamma^t\left(x_t^\top(Q+K^\top RK)x_t\right)\right]$, which fulfills the Bellman equation:

$$V^\pi(x)=x^\top(Q+K^\top RK)x+\gamma\mathbb{E}_{x'=(A+BK)x+v_0}\left[V^\pi(x')\right], \tag{4}$$

where $x'=(A+BK)x+v_0$ is the next random state from x. In this case, the solution to the Bellman equation can be explicitly characterized in a quadratic form:

$$V_\pi(x)=x^\top Px+q,$$

where q is a constant and P is the solution to the Lyapunov equation

$$P=Q+K^\top RK+\gamma(A+BK)^\top P(A+BK). \tag{5}$$

C. Distributional LQR

Unlike the traditional discounted LQR which primarily focuses on the expected return, the distributional LQR emphasizes the random return $G^\pi(x)$ with a given feedback controller $u_t=\pi(x_t)$ and an initial state $x_0=x$:

$$G^\pi(x)=\sum_{t=0}^{\infty}\gamma^t\left(x_t^\top Qx_t+u_t^\top Ru_t\right). \tag{6}$$

It has been shown in [1], if under the linear feedback controller $\pi(x_t)=Kx_t$ that K is stabilizing and satisfies $\|A_K\|<1$ with $A_K=A+BK$, the random return $G^K(x)$ that fulfills the random-variable Bellman equation

$$G^K(x)\overset{D}{=}x^\top Q_Kx+\gamma G^K(x'),\quad x'=A_Kx+v_0$$

can be exactly characterized as follows:

$$\begin{aligned}G^K(x)&=x^\top Px+2\sum_{k=0}^{\infty}\gamma^{k+1}w_k^\top PA_K^{k+1}x \\ &+\sum_{k=0}^{\infty}\gamma^{k+1}w_k^\top Pw_k+2\sum_{k=1}^{\infty}\gamma^{k+1}w_k^\top P\sum_{\tau=0}^{k-1}A_K^{k-\tau}w_\tau,\end{aligned} \tag{7}$$

where $w_k\sim\mathcal{D}$ are independent from each other for all $k\in\mathbb{N}$ and P is obtained from the Lyapunov equation (5).

Since $G^K(x)$ involves an infinite number of random variables w_k, one can approximate its distribution using a finite number of random variables:

$$\begin{aligned}G_N^K(x)&=x^\top Px+2\sum_{k=0}^{N-1}\gamma^{k+1}w_k^\top PA_K^{k+1}x \\ &+\sum_{k=0}^{N-1}\gamma^{k+1}w_k^\top Pw_k+2\sum_{k=1}^{N-1}\gamma^{k+1}w_k^\top P\sum_{\tau=0}^{k-1}A_K^{k-\tau}w_\tau,\end{aligned} \tag{8}$$

where $N\in\mathbb{N}$ is the number of random variables.

It has shown in [1] that the sequence $\{G_N^K(x)\}_{N\in\mathbb{N}}$ converges pointwise in distribution to $G^K(x)$, $\forall x\in\mathbb{R}^n$. More specifically, if the PDF of w_k is bounded, $\mathbb{E}[w_k^\top w_k]\leq\sigma^2$, for all $k\in\mathbb{N}$, and the feedback gain K satisfies $\|A_K\|<1$, the sup difference between the CDF F of $G^K(x)$ and the CDF F_N of $G_N^K(x)$ is bounded by

$$\sup_g|F(g)-F_N(g)|\leq c_0\gamma^N, \tag{9}$$

where c_0 is a constant that depends on the system parameters and the initial state value x.

In this paper, we will focus on the truncated random return $G_N^K(x)$ and investigate its fundamental properties (definite positive quadratic form, density function, and log-concavity) as well as their application to distributionally robust verification, which are detailed in the next section.

III. Main Results

In this section, we begin with the quadratic form of $G_N^K(x)$.

Proposition 3.1: Consider the truncated random return $G_N^K(x)$ in (8). Let $Z:=\left[w_0^\top,w_1^\top,w_2^\top,...,w_{N-1}^\top\right]^\top\in\mathbb{R}^{Nn\times1}$. The distribution of the truncated random return $G_N^K(x)$ is the same as the distribution of $S(Z)$

$$G_K^N(x)=S(Z):=Z^\top H_NZ+2L_N^\top Z+c \tag{10}$$

where $H_N\in\mathbb{R}^{Nn\times Nn}$ is a symmetric block matrix with

$$[H_N]_{ij}=\begin{cases}\gamma^iP, & \text{if } i=j,\\ \gamma^j\left(A_K^{j-i}\right)^\top P, & \text{if } i<j,\\ \gamma^iPA_K^{i-j}, & \text{if } i>j,\end{cases} \tag{11}$$

$L_N \in \mathbb{R}^{Nn\times 1}$ with $[L_N]_i = \gamma^i P A_K^i x$, and $c = x^\top P x$. Here $i, j = 1, \ldots, N$.

Proof: The quadratic form in (10) follows from the definition of $G_N^K(x)$. ∎

Remark 3.1: The quadratic form in random variables is key in statistics [9, 10]. This form facilitates the characterization of probability distributions [12, 13]. Many standard distributions (especially chi-squared related distributions) involve quadratic forms. In addition, this form also plays an important role in related engineering fields, such as communication [14].

A. Positive definiteness

Given the quadratic form in Proposition 3.1, we are interested to explore the positive definiteness of the block symmetric matrix H. We first give the following preliminary result.

Lemma 3.1: Let $Q_K = Q + K^T R K$ and $\hat{A}_K = P^{\frac{1}{2}} A_K P^{-\frac{1}{2}}$. We have

(i) $\frac{1}{\gamma}\left(I - P^{-\frac{1}{2}} Q_K P^{-\frac{1}{2}}\right) = \hat{A}_K^\top \hat{A}_K$;

(ii) $\|\hat{A}_K\| = \left[\frac{1}{\gamma}\left(1 - \lambda_{\min}\left(P^{-\frac{1}{2}} Q_K P^{-\frac{1}{2}}\right)\right)\right]^{\frac{1}{2}}$.

Proof: Note that the matrix P is the solution to the Lyapunov function $P = Q_K + \gamma A_K^\top P A_K$. Since Q, R are positive definite matrices and A_K is stable, we have P is also positive definite. It follows that

$$\begin{aligned} & P^{\frac{1}{2}} P^{\frac{1}{2}} = Q_K + \gamma P^{\frac{1}{2}} P^{-\frac{1}{2}} A_k^\top P^{\frac{1}{2}} P^{\frac{1}{2}} A_K P^{-\frac{1}{2}} P^{\frac{1}{2}} \\ \Leftrightarrow\ & \frac{1}{\gamma}\left(I - P^{-\frac{1}{2}} Q_K P^{-\frac{1}{2}}\right) = \hat{A}_K^\top \hat{A}_K. \end{aligned}$$

Since $\hat{A}_K^\top \hat{A}_K \succeq 0$, we have $I - P^{-\frac{1}{2}} Q_K P^{-\frac{1}{2}} \succeq 0$, that is, the matrix $P^{-\frac{1}{2}} Q_K P^{-\frac{1}{2}}$ is stable. It further yields that

$$\begin{aligned} \|\hat{A}_K\|^2 &= \sigma_{\max}^2(\hat{A}_K) = \lambda_{\max}(\hat{A}_K^T \hat{A}_K) \\ &= \frac{1}{\gamma}\lambda_{\max}(I - P^{-\frac{1}{2}} Q_K P^{-\frac{1}{2}}) \\ &= \frac{1}{\gamma}\left(1 - \lambda_{\min}(P^{-\frac{1}{2}} Q_K P^{-\frac{1}{2}})\right), \end{aligned}$$

which proves (ii). ∎

Let

$$\bar{\lambda} = \left[\frac{1}{\gamma}\left(1 - \lambda_{\min}\left(P^{-\frac{1}{2}} Q_K P^{-\frac{1}{2}}\right)\right)\right]^{\frac{1}{2}}, \quad (12)$$

$$\phi_i = \sum_{j=1}^{i} \bar{\lambda}^j + \sum_{j=1}^{N-i} (\gamma\bar{\lambda})^j. \quad (13)$$

Before proving the positive definiteness of H, we make the following assumption on ϕ_i.

Assumption 3.1: Assume that $\phi_i < 1$ for all $i = 1, \ldots, N$.

Theorem 3.1: Consider the block matrix H in (11). Under Assumption 3.1, the following results hold:

(i) H is positive definite, i.e., $H \succ 0$;

(ii) the eigenvalues of H are bounded by

$$\min_{i=1,\ldots,N} \gamma^i \|P\|(1-\phi_i) \le \lambda(H) \le \max_{i=1,\ldots,N} \gamma^i \|P\|(1+\phi_i).$$

Proof: First, we observe that

$$H = \Psi \hat{H} \Psi$$

where $\hat{H}$ are defined in (14) and $\Psi = \mathrm{diag}\{P^{\frac{1}{2}}, \ldots, P^{\frac{1}{2}}\}$. More specifically, we can write $\hat{H} = [\hat{H}_{ij}]$ where

$$\hat{H}_{ij} = \begin{cases} \gamma^i I, & \text{if } i = j, \\ \gamma^j P^{-\frac{1}{2}} \left(A_K^{j-i}\right)^\top P^{\frac{1}{2}}, & \text{if } i < j, \\ \gamma^i P^{\frac{1}{2}} A_K^{i-j} P^{-\frac{1}{2}}, & \text{if } i > j. \end{cases} \quad (15)$$

Recall $\hat{A}_K = P^{\frac{1}{2}} A_K P^{-\frac{1}{2}}$, which implies that $\hat{A}_K^j = P^{\frac{1}{2}} A_K^j P^{-\frac{1}{2}}$. Then, $\hat{H}_{ij}$ can be equivalently written as

$$\hat{H}_{ij} = \begin{cases} \gamma^i I, & \text{if } i = j, \\ \gamma^j (\hat{A}_K^{j-i})^\top, & \text{if } i < j, \\ \gamma^i \hat{A}_K^{j-i}, & \text{if } i > j. \end{cases} \quad (16)$$

From Definition 2.1, $\hat{H}$ is block strictly diagonally dominant if $\hat{H}_{ii}$ are nonsingular and

$$\left(\left\|\hat{H}_{ii}^{-1}\right\|\right)^{-1} > \sum_{j=1, j\neq i}^{N} \left\|\hat{H}_{ij}\right\|, \forall i = 1, \ldots, N. \quad (17)$$

Here, $\hat{H}_{ii} = \gamma^i I$, which is naturally nonsingular. From (16), the inequality (17) can be rewritten as

$$\gamma^i > \gamma^i \sum_{j=1}^{i-1} \left\|\hat{A}\right\|^{i-j} + \gamma^i \sum_{j=i+1}^{N} \gamma^{j-i} \left\|\hat{A}\right\|^{j-i}, \forall i = 1, \ldots, N. \quad (18)$$

According to Lemma 3.1, $\left\|\hat{A}\right\| = \bar{\lambda}$ as defined in (12). Then, a sufficient condition for the block strictly diagonal dominance of $\hat{H}$ is $\phi_i < 1$ for all $i = 1, \ldots, N$, where ϕ_i is defined in (13). It follows from Lemma 2.1 that for each eigenvalue λ of $\hat{H}$, there exists at least one $1 \le i \le N$ such that

$$|\lambda - \gamma^i| \le \gamma^i \phi_i \Leftrightarrow \gamma^i(1 - \phi_i) \le \lambda \le \gamma^i(1 + \phi_i).$$

Note that $\hat{H}^\top = \hat{H}$. Thus, if $\phi_i < 1$ for all $i = 1, \ldots, N$, all the eigenvalues of $\hat{H}$ is positive, which implies that $\hat{H} \succ 0$. Since $H = \Psi \hat{H} \Psi$ and $\Psi \succ 0$, we have that $H \succ 0$. It directly yields the bounds of eigenvalues of H in (ii). ∎

Theorem 3.1 implies that the positive definite property of the block matrix H relies on Assumption 3.1, which may restrict the system parameters. The following proposition shows that Assumption 3.1 is not required to establish the positive definiteness of H for scalar systems.

Proposition 3.2: If the system is a scalar system, i.e., $n = 1$, then H in (11) is positive definite for any N.

Proof: For a scalar system, A_K and P are scalars with $|A_K| < 1$ and $P > 0$. The positiveness of H directly follows from that its principal minor of order i are positive, i.e.,

$$\Delta_i = P^i \gamma^{\frac{i(i+1)}{2}} (1 - \gamma A_k^2)^{i-1} > 0, \ \forall i = 1, \ldots, N,$$

which completes the proof. ∎

Remark 3.2: In [1], $\|A_K\| < 1$ is sufficient to ensure good approximation property of the truncated random return $G_N^K(x)$ where the error bound exponentially decays with respect to N. This condition has to be sharper, as shown

$$\hat{H} = \begin{bmatrix} \gamma I & \gamma^2 P^{-\frac{1}{2}} A_K^\top P^{\frac{1}{2}} & \gamma^3 P^{-\frac{1}{2}} (A_K^2)^\top P^{\frac{1}{2}} & \gamma^4 P^{-\frac{1}{2}} (A_K^3)^\top P^{\frac{1}{2}} & \cdots & \gamma^N P^{-\frac{1}{2}} (A_K^{N-1})^\top P^{\frac{1}{2}} \\ \gamma^2 P^{\frac{1}{2}} A_K P^{-\frac{1}{2}} & \gamma^2 I & \gamma^3 P^{-\frac{1}{2}} A_K^\top P^{\frac{1}{2}} & \gamma^4 P^{-\frac{1}{2}} (A_K^2)^\top P^{\frac{1}{2}} & \cdots & \gamma^N P^{-\frac{1}{2}} (A_K^{N-2})^\top P^{\frac{1}{2}} \\ \gamma^3 P^{\frac{1}{2}} A_K^2 P^{-\frac{1}{2}} & \gamma^3 P^{\frac{1}{2}} A_K P^{-\frac{1}{2}} & \gamma^3 I & \gamma^4 P^{-\frac{1}{2}} A_K^\top P^{\frac{1}{2}} & \cdots & \gamma^N P^{-\frac{1}{2}} (A_K^{N-3})^\top P^{\frac{1}{2}} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \gamma^N P^{\frac{1}{2}} A_K^{N-1} P^{-\frac{1}{2}} & \gamma^N P^{\frac{1}{2}} A_K^{N-2} P^{-\frac{1}{2}} & \gamma^N P^{\frac{1}{2}} A_K^{N-3} P^{-\frac{1}{2}} & \gamma^N P^{\frac{1}{2}} A_K^{N-4} P^{-\frac{1}{2}} & \cdots & \gamma^N I \end{bmatrix} \tag{14}$$

in Lemma 3.1 and Assumption 3.1, for proving the positive definiteness of H, since we need to enforce $\|\hat{A}_K\| = \|P^{\frac{1}{2}} A_K P^{-\frac{1}{2}}\| < 1$.

Example 3.1: We consider an idealized data center cooling example [15, 16] with the dynamics $x_{t+1} = Ax_t + Bu_t + v_t$, where

$$A = \begin{bmatrix} 1.01 & 0.01 & 0 \\ 0.01 & 1.01 & 0.01 \\ 0 & 0.01 & 1.01 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

We select $Q = I$ and $R = I$. The controller is given by

$$K = -0.015 \begin{bmatrix} 56.19 & 0.7692 & 0.0027 \\ 0.7692 & 56.20 & 0.7692 \\ 0.0027 & 0.7692 & 56.19 \end{bmatrix},$$

and γ is set to 0.6. According to (12), the parameter $\bar{\lambda} = 0.1692$. We choose different N and compute the corresponding ϕ_i in (13). As shown in Table I, $\max_{i=1,\ldots,N} \phi_i$ is less than 1, which implies that Assumption 3.1 holds for the chosen N. We further compute the minimal and maximal eigenvalues of the matrix H, both of which are positive, indicating the positive definiteness of H. In addition, we further provide the lower and upper bounds in Theorem 3.1(ii), validating the soundness of these bounds, in comparison with the minimal and maximal eigenvalues of the matrix H.

TABLE I: The values of ϕ_i and eigenvalues of H

N	$\max_i \phi_i$	$\lambda_{\min}(H)$	$\lambda_{\max}(H)$	LB	UB
2	0.2708	0.5916	1.0848	0.3043	1.3583
3	0.2994	0.3547	1.0857	0.3043	1.3583
5	0.3146	0.1277	1.0857	0.1094	1.3595
10	0.3167	0.0099	1.0857	0.0085	1.3596
15	0.3167	0.0008	1.0857	0.0007	1.3596
20	0.3167	6.003×10^{-5}	1.0857	5.145×10^{-5}	1.3596
25	0.3167	4.668×10^{-6}	1.0857	4.001×10^{-6}	1.3596

B. Probability density function

Thanks to the positive definite quadratic form, we will show that when the disturbances $w_k \sim \mathcal{N}(0, \Sigma)$, we can characterize the density function of the truncated random return $G_N^K(x)$.

Proposition 3.3: Consider the truncated random return $G_N^K(x)$ in (8) and let $Z := \left[w_0^\top, w_1^\top, w_2^\top, ..., w_{N-1}^\top\right]^\top \in \mathbb{R}^{Nn \times 1}$. Assume that $w_k \sim \mathcal{N}(0, \Sigma)$. Then, the distribution of $G_N^K(x)$ (or $S(Z)$ in (10)) is a positively weighted non-central chi-square distribution with Nn degrees of freedom

$$\begin{aligned} G_N^K(x) = S(Z) = \sum_{i=1}^{Nn} \lambda_i \left(Y_i + \eta_i\right)^2 + x^\top P x - \eta^\top \Lambda \eta \\ + d^\top (\Sigma_Z^{\frac{1}{2}} H \Sigma_Z^{\frac{1}{2}} - H) d, \end{aligned} \tag{19}$$

where $\Sigma_Z = I_N \otimes \Sigma$, $\Sigma_Z^{\frac{1}{2}} = I_N \otimes \Sigma^{\frac{1}{2}}$, $Y \in \mathbb{R}^{Nn \times 1}$ with $Y_i \sim \mathcal{N}(0, 1)$. Here $\lambda_i > 0$, $i = 1, \ldots, Nn$, $\Lambda = \text{diag}\{\lambda_i\} \in \mathbb{R}^{Nn \times Nn}$, $\eta \in \mathbb{R}^{Nn \times 1}$, and $d \in \mathbb{R}^{Nn \times 1}$ depend on H_N, L_N, and c in (10).

Proof: Recall (10), we have

$$\begin{aligned} S(Z) &= Z^\top H_N Z + L_N^\top Z + c = (Z + d)^\top H_N (Z + d) + c' \\ &= F^\top H_N F + c', \end{aligned}$$

where $d = \frac{1}{2} H_N^{-1} L_N$, $c' = c - d^\top H_N d$, $F \in \mathbb{R}^{Nn \times 1}$ and $F \sim \mathcal{N}(d, \Sigma_Z)$. Then F can be be normalized by subtracting d and left multiply $\Sigma_Z^{-\frac{1}{2}}$, which yields

$$\begin{aligned} S(Z) &= \left[\Sigma_Z^{\frac{1}{2}} (W + d)\right]^\top H_N \left[\Sigma_Z^{\frac{1}{2}} (W + d)\right] + c' \\ &= W^\top \Sigma_Z^{\frac{1}{2}} H \Sigma_Z^{\frac{1}{2}} W + 2 d^\top \Sigma_Z^{\frac{1}{2}} H \Sigma_Z^{\frac{1}{2}} W + c'', \end{aligned}$$

where $W = \Sigma_Z^{-\frac{1}{2}} (F - d)$, $W \sim \mathcal{N}(\mathbf{0}, I_{Nn})$, and $c'' = d^\top \Sigma_Z^{\frac{1}{2}} H \Sigma_Z^{\frac{1}{2}} d + c'$. Let $\Gamma = \Sigma_Z^{\frac{1}{2}} H_N \Sigma_Z^{\frac{1}{2}}$, $\Phi^\top = d^\top \Sigma_Z^{\frac{1}{2}} H_N \Sigma_Z^{\frac{1}{2}}$. Choose the orthogonal matrix U such that $\Gamma = U \Lambda U^\top$ with Λ being a diagonal matrix. Then, we have

$$\begin{aligned} S(Z) &= W^\top \Gamma W + 2 \Phi^\top W + c' = W^\top U \Lambda U^\top W \\ &\quad + 2 \Phi^\top W + c'' = Y^\top \Lambda Y + 2 \Phi^\top U Y + c'' \end{aligned}$$

where $Y \sim \mathcal{N}(0, I_{Nn})$. Let $\zeta^\top = \Phi^\top U$, then

$$\begin{aligned} S(Z) &= Y^\top \Lambda Y + 2 \Phi^\top U Y + c'' = Y^\top \Lambda Y + 2 \zeta^\top Y + c'' \\ &= (Y + \eta)^\top \Lambda (Y + \eta) + c''' \\ &= \sum_{i=1}^{Nn} \lambda_i \left(Y_i + \eta_i\right)^2 + c'' - \eta^\top \Lambda \eta \\ &= \sum_{i=1}^{Nn} \lambda_i \left(Y_i + \eta_i\right)^2 + d^\top \Sigma_Z^{\frac{1}{2}} H_N \Sigma_Z^{\frac{1}{2}} d + c' - \eta^\top \Lambda \eta \\ &= \sum_{i=1}^{Nn} \lambda_i \left(Y_i + \eta_i\right)^2 + d^\top (\Sigma_Z^{\frac{1}{2}} H_N \Sigma_Z^{\frac{1}{2}} - H_N) d - \eta^\top \Lambda \eta \\ &\quad + x^\top P x, \end{aligned}$$

where $\eta = -\Lambda^{-1} \zeta$, $c''' = c'' - \eta^\top \Lambda \eta$, which yields a positively weighted non-central chi-square distribution. ■

An immediate result is the explicit PDF of $G_N^K(x)$ in the following corollary.

Corollary 3.1: [17, 18] Consider the truncated random return $G_N^K(x)$ in (8) and assume that $w_k \sim \mathcal{N}(0, \Sigma)$. The PDF of G_N^K can be explicitly written in the form

$$f_G(g) = \sum_{k=0}^{\infty} a_k \frac{q^{\alpha+k-1} e^{-\frac{q}{2\beta}}}{(2\beta)^{\alpha+k}\Gamma(\alpha+k)} = \sum_{k=0}^{\infty} a_k \, f_\Omega\big(g; \alpha+k, 2\beta\big),$$

$f_\Omega\big(g; \alpha+k, 2\beta\big)$ is the PDF of the gamma distribution of a random variable with $\alpha+k$ shape parameter evaluated at g. The coefficients are defined recursively:

$$a_0 = e^{-\frac{\eta}{2}} \prod_{i=1}^{Nn} \left(\frac{\beta}{\lambda_i}\right)^{\frac{1}{2}}, \; a_k = (2k)^{k-1} \sum_{l=0}^{k-1} b_{k-l} \, a_l, \quad k \geq 1,$$

$$b_k = k\,\beta \sum_{i=1}^{Nn} \eta_i \, \lambda_i^{c_i} \; + \; \sum_{i=1}^{Nn} c_i^k, \quad k \geq 1,$$

$$\eta = \sum_{i=1}^{Nn} \eta_i^2, \quad \alpha = \frac{Nn}{2}, \quad c_i = 1 - \frac{\beta}{\lambda_i}.$$

Here β is an arbitrary constant satisfying $|1 - \frac{\beta}{\lambda_i}| < 1$ to guarantee the b_k's convergence and λ_i's are defined in Proposition 3.3.

Example 3.2: Let us continue to consider the same system in Example 3.1. We assume that the i.i.d random disturbances w_k are zero-mean Gaussian random vector with covariance matrix $\Sigma = \begin{bmatrix} 1 & 0.5 & 0.3 \\ 0.5 & 2 & 0.4 \\ 0.3 & 0.4 & 2 \end{bmatrix}$. Let $N = 20$ and $x = [3 \; 3 \; 3]^\top$. In this case, the PDF of $G_N^K(x)$ is shown Fig. 1, consistent with the shape of positively weighted non-central chi-square distribution.

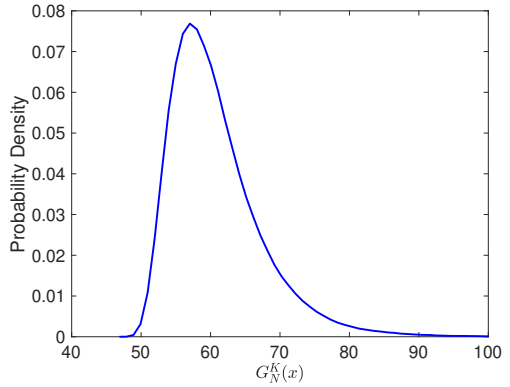


Fig. 1: Probability density values of $G_N^K(x)$ where the random disturbance is zero-mean Gaussian.

Remark 3.3: The positively weighted non-central chi-square distribution is a special non-central generalized chi-square distribution, which governs the the performance index of the finite-horizon LQR problem [19]. This kind of distribution plays a significant role in advanced statistical theory [20] and applications, e.g., portfolio theory [21].

C. Log-concavity of CDF

In the light of the positive definite quadratic form of $S(Z)$, we have the following result.

Lemma 3.2: Consider the $S(Z)$ in (10) and define the sub-level set of this mapping S at the level g:

$$\mathcal{L}_g = \{z \in \mathbb{R}^{Nn} \mid S(z) \leq g\}. \tag{20}$$

The set $\mathcal{L}_g$ is convex and compact for any $g \in \mathbb{R}$.

Proof: It directly follows from the positive definite quadratic form of $S(Z)$ which is convex in Z. ■

Proposition 3.4: Consider the truncated random return $G_N^K(x)$ in (8) and let $Z := \big[w_0^\top, w_1^\top, w_2^\top, ..., w_{N-1}^\top\big]^\top \in \mathbb{R}^{Nn\times 1}$. Suppose that the PDF $f_Z(z)$ of Z is log-concave. Then, the CDF $F_{x,N}^K$ of the truncated random return $G_N^K(x)$ is log-concave.

Proof: From Lemma 3.2, the sub-level set $\mathcal{L}(g)$ in (20) of the mapping S is convex and compact for any $g \in \mathbb{R}$. The CDF of G can be defined using the set $\mathcal{L}(g)$:

$$F_N(g) = \mathbb{P}[G_K^N(x) \leq g] = \int_{\mathcal{L}(g)} f_Z(z)\, dz.$$

The Prékopa's theorem has shown that, if a function $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ is log-concave, then $h(x) = \int f(x,y)\, dy$ is a log-concave function of x. To apply the Prékopa's Theorem, we construct an auxiliary function $f(z,g) \in \mathbb{R}^N \times \mathbb{R}$:

$$f(z,g) = \begin{cases} f_Z(z), & S(z) \leq g, \\ 0, & S(z) > g. \end{cases}$$

Therefore, if we can prove that $f(z,g)$ is log-concave on $(z,g) \in \mathbb{R}^{Nn} \times \mathbb{R}$, we can apply the Prékopa's Theorem to show that $F_N(g)$ is log-concave on $\mathbb{R}$. Since $f(z,g) \equiv 0$ when $S(z) > g$, We consider two scenarios by selecting distinct z_1, z_2, g_1, g_2:

1) **Case 1:** $f(z_1, g_1) = 0$ or $f(z_2, g_2) = 0$. The following trivially holds for any $\theta \in [0,1]$:

$$\begin{aligned} &f\big[(1-\theta)(z_1,g_1) + \theta(z_2,g_2)\big] \\ &\geq 0 = [f(z_1,g_1)]^{1-\theta} \cdot [f(z_2,g_2)]^{\theta}. \end{aligned}$$

2) **Case 2:** $F(z_1,g_1) \neq 0$, $f(z_2,g_2) \neq 0$. Let $\big(z_\theta, g_\theta\big) = \big((1-\theta)z_1 + \theta z_2, \, (1-\theta)g_1 + \theta g_2\big)$. Since $S(Z)$ is convex,

$$\begin{aligned} S(z_\theta) = S\big((1-\theta)z_1 + \theta z_2\big) &\leq (1-\theta)S(z_1) + \theta S(z_2) \\ &\leq (1-\theta)g_1 + \theta g_2 = g_\theta. \end{aligned}$$

As a result, $S(z_\theta) \leq g_\theta$. For $f(z_\theta, g_\theta)$, we have

$$f\big((1-\theta)(z_1,g_1) + \theta(z_2,g_2)\big) = f(z_\theta, g_\theta) = f_Z(z_\theta).$$

Since $f_Z(z)$ is log-concave on $\mathbb{R}^{Nn}$, we have that for any $\theta \in (0,1)$, $f_Z\big(z_\theta\big) \geq \big(f_Z(z_1)\big)^{1-\theta}\big(f_Z(z_2)\big)^{\theta}$. Recalling that $f(z_1,g_1) = f_Z(z_1)$, $f(z_2,g_2) = f_Z(z_2)$ when $f \neq 0$,

$$\begin{aligned} &f\big[(1-\theta)(z_1,g_1) \; + \; \theta(z_2,g_2)\big] = f(z_\theta, g_\theta) \\ &= f_Z(z_\theta) \geq \big(f_Z(z_1)\big)^{\,1-\theta}\,\big(f_Z(z_2)\big)^{\,\theta} \\ &= \big(f(z_1,g_1)\big)^{1-\theta}\,\big(f(z_2,g_2)\big)^{\theta}. \end{aligned}$$

To this end, we show that $f(z, g)$ is log-concave. According to the Prékopa's theorem, we further conclude that the CDF $F_N(g)$ is log-concave. ■

Remark 3.4: The log-concavity is an important property in statistical inference and optimization. In particular, the log-concavity of CDF is useful in reliability theory [22], while the log-concavity of PDF implies uni-modality of the density function and allows for efficient MCMC sampling [23].

Example 3.3: We verify our proof of log-concavity of CDF on a discretized linear model for vehicle steering, which is borrowed from [24, 25]. The system matrices are $A = [1\ 0.2; 0\ 1]$ and $B = [0.06; 0.20]$. We select $Q = 10I$ and $R = 1$. We consider five different random disturbances w_k whose PDFs are log-concave. We choose $x = [0; 0]$, $N = 20$, $\gamma = 0.6$, and $K = [-2.11\ \ -2.56]$. We plot the second derivative of the CDF $F_N(g)$ of the random return $G_N^K(x)$, shown in Fig. 2. We see that for all these different log-concave input disturbances, the second derivatives of $\log(F_N(g))$ are negative, validating its concavity.

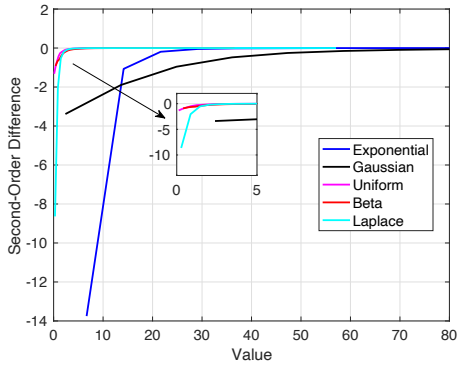


Fig. 2: Second derivative of $\log(F_N(g))$ under five different random disturbances whose PDFs are log-concave

IV. Conclusion

In this paper, we have studied the truncated random return in the distributional LQR. We showed that the truncated random return can be naturally expressed as a quadratic form in random variable. We proved the positive definiteness of the block symmetric matrix in the quadratic form. We further showed that the truncated random return follows a positively weighted non-central chi-square distribution if the random disturbances are Gaussian, and its cumulative distribution function is log-concave if the probability density distribution of the random disturbances is log-concave.

There are several interesting future directions to be explored. The first one is whether we can develop a necessary and sufficient condition for proving the positive definiteness of the block matrix in the quadratic form. The second one is the use of these properties for policy improvement in the risk-averse control. In addition, thanks to the importance of quadratic form in random variables in statistics and other related fields of engineering, we are interested in the role of random return in decision making from a statistical perspective.

References

[1] Z. Wang, Y. Gao, S. Wang, M. M. Zavlanos, A. Abate, and K. H. Johansson, "Policy evaluation in distributional LQR," *arXiv preprint arXiv:2401.10240*, 2023.

[2] ——, "Policy evaluation in distributional LQR," in *Learning for Dynamics and Control Conference*. PMLR, 2023, pp. 1245–1256.

[3] B. D. Anderson and J. B. Moore, *Optimal Control: Linear Quadratic Methods*. Courier Corporation, 2007.

[4] C. Vandamme, V. Otlet, R. Ronsse, and F. Crevecoeur, "Model of gait control in parkinson's disease and prediction of robotic assistance," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 31, pp. 1374–1383, 2023.

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

[6] M. G. Bellemare, W. Dabney, and R. Munos, "A distributional perspective on reinforcement learning," in *International Conference on Machine Learning*. PMLR, 2017, pp. 449–458.

[7] Y. Chow, M. Ghavamzadeh, L. Janson, and M. Pavone, "Risk-constrained reinforcement learning with percentile risk criteria," *Journal of Machine Learning Research*, vol. 18, no. 167, pp. 1–51, 2018.

[8] W. Dabney, M. Rowland, M. Bellemare, and R. Munos, "Distributional reinforcement learning with quantile regression," in *AAAI Conference on Artificial Intelligence*, vol. 32, no. 1, 2018, pp. 2892–2901.

[9] A. M. Mathai and S. Provost, *Quadratic Forms in Random Variables: Theory and Applications*. Dekker, 1992.

[10] D. E. Varberg, "Convergence of quadratic forms in independent random variables," *The Annals of Mathematical Statistics*, vol. 37, no. 3, pp. 567–576, 1966.

[11] D. G. Feingold and R. S. Varga, "Block diagonally dominant matrices and generalizations of the gerschgorin circle theorem," *Pacific Journal of Mathematics*, vol. 12, no. 4, pp. 1241–1250, 1962.

[12] B. Baldessari, "The distribution of a quadratic form of normal random variables," *The Annals of Mathematical Statistics*, vol. 38, no. 6, pp. 1700–1704, 1967.

[13] C. Zhang and D. Zhang, "Probability and moment inequalities for quadratic forms in independent random variables with fat tails," *Statistics & Probability Letters*, vol. 217, p. 110290, 2025.

[14] T. Y. Al-Naffouri and B. Hassibi, "On the distribution of indefinite quadratic forms in gaussian random variables," in *2009 IEEE International Symposium on Information Theory*, 2009, pp. 1744–1748.

[15] S. Dean, H. Mania, N. Matni, B. Recht, and S. Tu, "On the sample complexity of the linear quadratic regulator," *Foundations of Computational Mathematics*, vol. 20, no. 4, pp. 633–679, 2020.

[16] F. A. Yaghmaie, F. Gustafsson, and L. Ljung, "Linear quadratic control using model-free reinforcement learning," *IEEE Transactions on Automatic Control*, vol. 68, no. 2, pp. 737–752, 2023.

[17] H. Ruben, "Probability content of regions under spherical normal distributions, IV: The distribution of homogeneous and non-homogeneous quadratic functions of normal variables," *The Annals of Mathematical Statistics*, vol. 33, no. 2, pp. 542–570, 1962.

[18] F. G. Aldo Gardini and C. Trivisano, "The mellin transform to manage quadratic forms in normal random variables," *Journal of Computational and Graphical Statistics*, vol. 31, no. 4, pp. 1416–1425, 2022.

[19] F. Qian, J. Gao, and D. Li, "Complete statistical characterization of discrete-time LQG and cumulant control," *IEEE Transactions on Automatic Control*, vol. 57, no. 8, pp. 2110–2115, 2011.

[20] H. Liu, Y. Tang, and H. H. Zhang, "A new chi-square approximation to the distribution of non-negative definite quadratic forms in non-central normal variables," *Computational Statistics & Data Analysis*, vol. 53, no. 4, pp. 853–856, 2009.

[21] Y. Guo, "Applications of the chi-square distribution in quantitative finance: A statistical perspective," *Available at SSRN 5048181*, 2024.

[22] M. Bagnoli and T. Bergstrom, "Log-concave probability and its applications," in *Rationality and Equilibrium: A Symposium in Honor of Marcel K. Richter*. Springer, 2006, pp. 217–241.

[23] R. Dwivedi, Y. Chen, M. J. Wainwright, and B. Yu, "Log-concave sampling: Metropolis-hastings algorithms are fast," *Journal of Machine Learning Research*, vol. 20, no. 183, pp. 1–42, 2019.

[24] K. J. Åström and R. Murray, *Feedback Systems: An Introduction for Scientists and Engineers*. Princeton University Press, 2021.

[25] M. Kishida and A. Cetinkaya, "Risk-aware linear quadratic control using conditional value-at-risk," *IEEE Transactions on Automatic Control*, vol. 68, no. 1, pp. 416–423, 2022.