2020 59th IEEE Conference on Decision and Control (CDC)
Jeju Island, Republic of Korea, December 14-18, 2020

Risk-Constrained Linear-Quadratic Regulators

Anastasios Tsiamis, Dionysios S. Kalogerias, Luiz F. O. Chamon,
Alejandro Ribeiro and George J. Pappas

***Abstract*— We propose a new risk-constrained reformulation of the standard Linear Quadratic Regulator (LQR) problem. Our framework is motivated by the fact that the classical (risk-neutral) LQR controller, although optimal in expectation, might be ineffective under relatively infrequent, yet statistically significant (risky) events. To effectively trade between average and extreme event performance, we introduce a new risk constraint, which explicitly restricts the total expected predictive variance of the state penalty by a user-prescribed level. We show that, under rather minimal conditions on the process noise (i.e., finite fourth-order moments), the optimal risk-aware controller can be evaluated explicitly and in closed form. In fact, it is affine relative to the state, and is always internally stable regardless of parameter tuning. Our new risk-aware controller: i) pushes the state away from directions where the noise exhibits heavy tails, by exploiting the third-order moment (skewness) of the noise; ii) inflates the state penalty in riskier directions, where both the noise covariance and the state penalty are simultaneously large. The properties of the proposed risk-aware LQR framework are also illustrated via indicative numerical examples.**

I. Introduction

Achieving good performance in expectation is often insufficient in the design of stochastic control systems, especially when dealing with modern, critical applications. Examples appear naturally in many areas, including wireless industrial control [1], energy [2], [3], finance [4]–[6], robotics [7], [8], networking [9], and safety [10], [11], to name a few. Indeed, occurrence of less probable, non-typical or unexpected events might lead the underlying dynamical system to experience shocks with possibly catastrophic consequences, e.g., a drone diverging too much from a given trajectory in a hostile environment, or an autonomous vehicle crashing onto a wall or hitting a pedestrian. In such situations, design of effective *risk-aware* control policies is highly desirable, systematically compensating for those extreme events, at the cost of slightly sacrificing average performance under nominal conditions.

To highlight the usefulness of a risk-aware control policy, let us consider the following simple, motivating example. Let $x_{k+1} = x_k + u_k + w_{k+1}$ model an aerial robot, moving along a line. Assume that the process noise w_k is *i.i.d.* Bernoulli, taking the values $\beta > 2$ with probability $1/\beta$ and 0 with probability $1 - 1/\beta$. This noise represents shocks, e.g., wind gusts, that can occur with some small probability. We would then like to minimize the LQR cost $\mathbb{E}\sum_{t=0}^{N}\{x_t^2\}$, i.e., the total displacement of the robot over a horizon of N time steps. In this case, the LQR optimal controller is

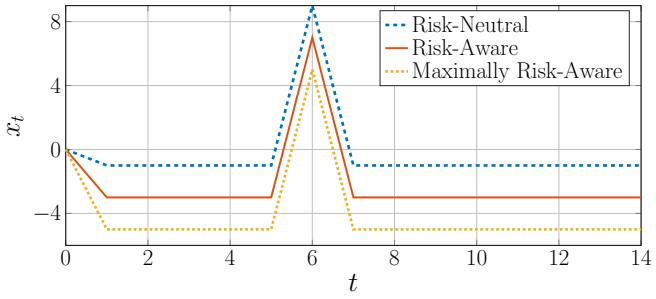


Fig. 1. Comparison between risk-neutral and risk-aware control performance, when the system experiences rare but large shocks—here the shock occurs at time 6. By sacrificing average behavior, the risk-aware controllers push the state away from the direction of the shock.

$u_k^{\mathrm{LQR}} = -x_k - 1$, where $-1 \equiv -\mathbb{E}w_k$ cancels the mean of the process noise. We see that the LQR solution is *risk-neutral*, as it does not account for the fact that the shock β could be arbitrarily large. On the other hand, the risk-aware LQR formulation proposed in this work results in a *family* of optimal controllers of the form

$$u_t^*(\mu) = -x_t - 1 - \frac{\mu}{1+2\mu}(\beta - 2), \quad \mu \geq 0,$$

where μ controls the trade-off between average performance and risk. As μ increases, we move from the risk-neutral to the *maximally risk-aware* controller $u_t^*(\infty) = -x_t - \beta/2$, which treats the noise as adversarial—see Fig. 1.

In both classical and recent literature in linear-quadratic problems, risk awareness in estimation and control is typically achieved by replacing the respective random cost with its exponentiation [12]–[21]. Yet, the resulting stochastic control problem might not be well-defined for general classes of noise distributions, as it requires the moment generating function of the cost to be finite. Thus, heavy-tailed or skewed distributions, which are precisely those exhibiting high risk, are naturally excluded. Also, even if the expectation of the exponential cost is finite, it does not lead to a general, closed-form and interpretable solution. A notable exception is that of a Gaussian process noise, also known as the Linear Exponential Quadratic Gaussian (LEQG) problem, which does enjoy a simple closed-form solution [13]. Apparently though, the Gaussian assumption is unable to capture distributions with asymmetric (skewed) structure, as in the above example.

In this paper, we propose a new risk-aware reformulation of the LQR problem, in which the standard LQR objective is minimized *subject to* an explicit and tunable risk constraint. Our contributions are as follows.

–New Risk Measure. We introduce the cumulative expected one-step predictive variance of the associated state penalty

The authors are with the Department of Electrical and Systems Engineering, University of Pennsylvania, Philadelphia, PA 19104 (email: {atsiamis,dionysis,luizf,aribeiro,pappasg}@seas.upenn.edu).

This work is supported by the AFOSR under grant FA9550-19-1-0265 (Assured Autonomy in Contested Environments), and the NSF under grant CPS 1837253.

Authorized licensed use limited to: University of Pennsylvania. Downloaded on October 29,2026 at 15:14:47 UTC from IEEE Xplore. Restrictions apply.

as a new risk measure for LQR control. In this way, our risk-constraint formulation ensures not only a small LQR cost, but also *guaranteed* statistical variability of the state penalty.

–Optimal Risk-Aware Controls & Stability. We show that our new risk-constrained formulation results in a quadratically constrained LQR problem (Proposition 1), which admits an explicit closed-form solution with a natural interpretation. The optimal risk-aware feedback controller is affine with respect to the system state (Theorems 2 & 3). The affine component of the control law pushes the state away from directions where the noise exhibits (skewed) heavy tails. Meanwhile, the state feedback gain satisfies a new risk-aware Riccati recursion, in which the state penalty is inflated in riskier directions, where both the noise covariance and state penalty are simultaneously larger. Further, we show that our optimal risk-aware controller is always stable, under standard LQR conditions (Corollary 1).

–Arbitrary Noise Model. Contrary to the LEQG approach, our results are valid for arbitrary noise distributions, provided the associated fourth-order moments are finite; thus, heavy-tailed or skewed noises are supported within our framework.

–Relation to LQR with Tracking. We show that, *by appropriate re-parameterization*, our risk-aware LQR problem is equivalent to a generalized risk-neutral LQR problem with a tracking objective. Essentially, this implies that risk-neutral LQR formulations can provide inherent risk-averse behavior, as long as the involved parameters are selected in a principled way, as presented herein.

Related Work: Risk-aware optimization has been studied in a wide variety of decision making contexts [11], [22]–[31]. The basic idea is to replace expectations by more general functionals, called risk measures [32], purposed to effectively quantify the statistical volatility of the involved random cost function, in addition to mean performance. Typical examples are mean-variance functionals [4], [32], mean-semideviations [26], and Conditional Value-at-Risk (CVaR) [33].

In the case of control systems, apart form the aforementioned exponential approach, CVaR optimization techniques have also been considered for risk-aware constraint satisfaction [11]. Although CVaR captures variability and tail events well, CVaR optimization problems rarely enjoy closed-form expressions. Approximations are usually required to make computations tractable, e.g., process noise and controls are assumed to be finite-valued [11], thus excluding the LQR setting. Predictive variance constraints have also been used as a measure of risk in portfolio optimization [34]. Different from our paper, the noise is assumed to be Gaussian and the variance is with respect to linear stage costs.

Another related concept is that of robust control, where the system model is unknown. The objective is to optimally control the true system under worst case model uncertainty. When there is no model uncertainty, in the case of stochastic process noise, robust controllers usually reduce to their risk-neutral LQR counterparts [35], [36]. On the contrary, in risk-aware control, extreme noise events are part of the system model. Even if the system is exactly modeled, we would still need to consider risk-aware control if the process noise is heavy-tailed or highly variable. From this point of view, robustness and risk are orthogonal concepts. Interestingly, in the case of adversarial noise, there is a connection between robust and maximally risk-aware controllers [37].

Notation: The transpose operation is denoted by $(\cdot)'$. If $x_k, \dots, x_t$ is a sequence of vectors, then $x_{k:t}$ denotes the batch vector of all x_i for $k \le i \le t$. The σ-algebra generated by a random vector x is denoted by $\sigma(x)$.

II. Risk-Constrained LQR Formulation

Consider a discrete-time linear system in state-space form evolving according to the stochastic difference equation

$$x_{t+1} = Ax_t + Bu_t + w_{t+1}, \tag{1}$$

where $x_t \in \mathbb{R}^n$ is the state, $u_t \in \mathbb{R}^p$ is an exogenous control signal, A is the state transition matrix, and B is the input matrix. We assume that the initial value x_0 is deterministic and fixed. Signal $w_t \in \mathbb{R}^n$ is a random process noise (not necessarily Gaussian) and is assumed to be *i.i.d.* For $t \ge 0$, let $\mathscr{F}_t = \sigma(x_{0:t}, u_{0:t})$ be the σ-algebra generated by all observables up to time t, and let $\mathscr{F}_{-1}$ be the trivial σ-algebra. Based on this notation, x_t, u_t, w_t are $\mathscr{F}_t$-measurable, while w_{t+1} is independent of $\mathscr{F}_t$. We also make an additional assumption on the process noise, as follows.

Assumption 1 (Noise Regularity). *The process w_t has finite fourth-order moment, i.e., for every $t \in \mathbb{N}$, $\mathbb{E}\,\|w_t\|_2^4 < \infty$.*

The above condition is mild and satisfied by very general noise distributions, including many heavy-tailed ones. Denote the mean of the noise by $\bar{w} \triangleq \mathbb{E} w_k$ and its variance by $W \triangleq \mathbb{E}(w_k - \bar{w})(w_k - \bar{w})'$.

In the classical risk-neutral formulation of the LQR problem, one is interested in the multistage stochastic program

$$\begin{aligned} \min_{u} \quad & \mathbb{E}\left\{ x_N' Q x_N + \sum_{t=0}^{N-1} x_t' Q x_t + u_t' R u_t \right\}, \\ \text{s.t.} \quad & x_{t+1} = Ax_t + Bu_t + w_{t+1} \\ & u_t \in \mathcal{L}_2(\mathscr{F}_t),\ t = 0, \dots N-1 \end{aligned} \tag{2}$$

where $u = u_{0:N-1}$ are the inputs from time 0 up to time $N-1$, for some horizon $N \in \mathbb{N}$. For each t, the *causality constraint* on u_t restricts the inputs to the space of square-integrable $\mathscr{F}_t$-measurable vector-valued random elements of appropriate dimension, denoted as $\mathcal{L}_2(\mathscr{F}_t)$. It also guarantees that the optimization problem is well-defined and with finite cost. In order for the optimal LQR controller to be well-behaved and stable as the horizon N grows, we also make the following standard assumption.

Assumption 2 (LQR). *The pair (A, B) is stabilizable, the pair $(A, Q^{1/2})$ is detectable, matrix $Q \succeq 0$ is positive semi-definite and matrix $R \succ 0$ is positive definite.*

As mentioned above, the classical LQR problem is risk-neutral, since it optimizes performance only on average [38]. Still, even if the average performance is good, the state can grow arbitrarily large under less probable, yet extreme events. In other words, the state can exhibit large variability. To deal

Authorized licensed use limited to: University of Pennsylvania. Downloaded on October 29,2026 at 15:14:47 UTC from IEEE Xplore. Restrictions apply.

with this issue, we propose a risk-constrained formulation of the LQR problem, posed as

$$\begin{aligned}\min_{u} \quad & \mathbb{E}\left\{x_N' Q x_N + \sum_{t=0}^{N-1} x_t' Q x_t + u_t' R u_t\right\} \\ \text{s.t.} \quad & \mathbb{E}\left\{\sum_{t=1}^{N} \left[x_t' Q x_t - \mathbb{E}\left(x_t' Q x_t | \mathscr{F}_{t-1}\right)\right]^2\right\} \leq \epsilon \cdot \\ & x_{t+1} = Ax_t + Bu_t + w_{t+1} \\ & u_t \in \mathcal{L}_2(\mathscr{F}_t),\, t = 0, \ldots, N-1\end{aligned} \tag{3}$$

Here, the risk measure adopted is the cumulative expected predictive variance of the state cost. The predictive variance incorporates information about the tail *and* skewness of the penalty $x_t' Q x_t$. This forces the controller to take higher-order noise statistics into account, mitigating the effect of rare though large noise values. Hence, our risk-aware LQR formulation not only forces the state x_t to be close to zero, but also explicitly restricts its variability.

Problem (3) offers a simple and interpretable way to control the trade-off between average performance and risk. By simply decreasing ϵ, we increase risk-awareness. Inspired by standard risk-aware formulations, in the above optimization problem our risk definition is tied to the specific state penalty $x_t' Q x_t$ of the LQR. However, all of our results are still valid if we employ the predictive variance of a different quadratic form, e.g., the norm of the state, $\|x_t\|^2$, in the constraint—see Section V. Lastly, note that the initial state is fixed (for simplicity), so there is no associated risk term for $t = 0$.

In the next section, we show that the risk constraint of (3) can be rewritten in quadratic form. This will allow us to solve problem (3) using duality theory and obtain a closed-form solution, exploiting higher-order noise moments.

III. Optimal Risk-Aware LQR Controllers

The analysis of the risk-aware dynamic program (3) consists of the following steps. First, we ensure the well-definiteness of (3), also showing that (3) can be equivalently reexpressed as a *sequential variational Quadratically Contrained Quadratic Program (QCQP)*, or, more precisely, as a *Quadratically Constrained LQR (QC-LQR)* problem (Proposition 1). Then, we exploit Lagrangian duality (Theorem 1) to solve (3) *exactly and in closed form*. More specifically, we first derive an explicit expression for the optimal risk-aware controller (Theorem 2), given an arbitrary but fixed Lagrange multiplier. Then, we show how an optimal Lagrange multiplier may be efficiently discovered via trivial bisection (Theorem 3 and Proposition 2).

Proposition 1 (QCQP Reformulation). *Let Assumption 1 be in effect, and define the higher-order weighted statistics*

$$M_3 \triangleq \mathbb{E}\left\{(w_i - \bar{w})(w_i - \bar{w})' Q (w_i - \bar{w})\right\} \in \mathbb{R}^n \text{ and}$$
$$m_4 \triangleq \mathbb{E}\left\{\left[(w_i - \bar{w})' Q (w_i - \bar{w}) - \operatorname{tr}(WQ)\right]^2\right\} \geq 0.$$

Then, the risk-constrained LQR problem (3) *is well-defined and equivalent to the sequential variational QCQP*

$$\begin{aligned}\min_{u} \quad & J(u) \triangleq \mathbb{E}\left\{x_N' Q x_N + \sum_{t=0}^{N-1} x_t' Q x_t + u_t' R u_t\right\} \\ \text{s.t. } & J_R(u) \triangleq \mathbb{E}\left\{\sum_{t=1}^{N} 4 x_t' Q W Q x_t + 4 x_t' Q M_3\right\} \leq \bar{\epsilon}\,, \\ & x_{t+1} = Ax_t + Bu_t + w_{t+1} \\ & u_t \in \mathcal{L}_2(\mathscr{F}_t), \quad t = 0, \ldots N-1,\end{aligned} \tag{4}$$

where $\bar{\epsilon} \triangleq \epsilon - N m_4 + 4N \operatorname{tr}\left\{(WQ)^2\right\}$.

The proof can be found in the Appendix. Proposition 1 is critical because it shows that our risk constraint is quadratic with respect to the state and control inputs. This enables us to apply duality theory, as discussed next.

A. Lagrangian Duality

To tackle problem (3), we now consider the *variational Lagrangian* $\mathcal{L} : \mathcal{L}_2(\mathscr{F}_0) \times \cdots \times \mathcal{L}_2(\mathscr{F}_{N-1}) \times \mathbb{R}_+ \to \mathbb{R}$ of the sequential QCQP (4), defined as

$$\mathcal{L}(u, \mu) \triangleq J(u) + \mu J_R(u) - \mu \bar{\epsilon}, \tag{5}$$

where $\mu \in \mathbb{R}_+$ is a multiplier associated with the variational risk constraint of (4). Hereafter, problem (4) will be called the *primal problem*. Accordingly, the *dual function* $D : \mathbb{R}_+ \to [-\infty, \infty)$ is additionally defined as

$$D(\mu) \triangleq \inf_{u \in \mathcal{U}_0} \mathcal{L}(u, \mu), \tag{6}$$

where the *implicit feasible set* $\mathcal{U}_0$ obeys ($k \leq N-1$)

$$\mathcal{U}_k \triangleq \left\{ u_{k:N-1} \in \prod_{t=k}^{N-1} \mathcal{L}_2(\mathscr{F}_t) \,\middle|\, x_{t+1} = Ax_t + Bu_t + w_{t+1} \right\},$$

and contains the constraints of (4) that have not been dualized in the construction of the Lagrangian in (5). Note that it is always the case that $D \leq J^*$ on $\mathbb{R}_+$, where $J^* \in [0, \infty]$ denotes the optimal value of the primal problem (4). Then, the optimal value of the always concave *dual problem*

$$\sup_{\mu \geq 0} D(\mu) \equiv \sup_{\mu \geq 0} \inf_{u \in \mathcal{U}_0} \mathcal{L}(u, \mu), \tag{7}$$

$D^* \triangleq \sup_{\mu \geq 0} D(\mu) \in [-\infty, \infty]$, is the tightest under-estimate of J^*, when knowing only D.

Leveraging Lagrangian duality, we may now state the following result, which provides sufficient optimality conditions for the QCQP (4). The proof is omitted, as it follows as direct application of [39, Theorem 4.10] .

Theorem 1 (Optimality Conditions). *Let Assumption 1 be in effect. Suppose that there exists a feasible policy-multiplier pair* $(u^*, \mu^*) \in \mathcal{U}_0 \times \mathbb{R}_+$ *such that*

1) $\mathcal{L}(u^*(\mu^*), \mu^*) = \min_{u \in \mathcal{U}_0} \mathcal{L}(u, \mu^*) = D(\mu^*)$;
2) $J_R(u^*) \leq \bar{\epsilon}$, *i.e., the dualized variational risk constraint of* (4) *is satisfied by control policy* u^*;
3) $\mu^*(J_R(u^*) - \bar{\epsilon}) = 0$, *i.e., complementary slackness holds.*

Then, u^* *is optimal for both the primal problem* (4) *and the initial problem* (3), μ^* *is optimal for the dual problem* (7), *and* (4) *exhibits zero duality gap, that is,* $D^* \equiv P^* < \infty$.

Authorized licensed use limited to: University of Pennsylvania. Downloaded on October 29,2026 at 15:14:47 UTC from IEEE Xplore. Restrictions apply.

Theorem 1 will be serving as the backbone of our analysis towards the solution to problem (4), presented as follows.

B. Optimal Risk-Aware Control Policies

Let $\mu \geq 0$ be arbitrary but fixed. First, we may simplify the form of the Lagrangian $\mathcal{L}$ and express it within a canonical dynamic programming framework. In this respect, we have the following, almost obvious, but important result.

Lemma 1 (Lagrangian Reformulation). *Let Assumption 1 be in effect and define the inflated state penalty matrix*

$$Q_\mu \triangleq Q + 4\mu QWQ.$$

Then, for every $u_t \in \mathcal{L}_2(\mathscr{F}_t)$, $t \leq N-1$, the Lagrangian function $\mathcal{L}$ can be expressed as

$$\mathcal{L}(u,\mu) = \mathbb{E}\left\{ g_N(x_N,\mu) + \sum_{t=0}^{N-1} g_t(x_t,u_t,\mu) \right\} + g(\mu), \quad (8)$$

where, in the notation of Proposition 1,

$$g_N(x_N,\mu) \triangleq x_N' Q_\mu x_N + 4\mu M_3' Q x_N,$$
$$g_t(x_t,u_t,\mu) \triangleq x_t' Q_\mu x_t + 4\mu M_3' Q x_t + u_t' R u_t,\ t \leq N-1,$$
$$\text{and} \quad g(\mu) \triangleq -\mu\bar{\epsilon} - 4\mu x_0' QWQ x_0 - 4\mu M_3' Q x_0.$$

Proof. It follows from Proposition 1 and the form of $\mathcal{L}$. □

Remark 1 (*Relation to generalized LQR with tracking*)**.** The Lagrangian (8) has the structure of a generalized LQR problem with a tracking objective. By completing the squares we can rewrite the stage cost $g_t(x_t,u_t,\mu)$ as

$$g_t(x_t,u_t,\mu) = (x_t + 2\mu M_3)' Q (x_t + 2\mu M_3) + x_t'(4\mu QWQ)x_t + u_t' R u_t - 4\mu^2 M_3' Q M_3,$$

i.e., the state penalty is quadratic and consists of two distinct terms. The first one, i.e., $(x_t + 2\mu M_3)' Q (x_t + 2\mu M_3)$ is a tracking error term that forces the state to be close to the static target $-2\mu M_3$. Informally, in the case of skewed noise, by tracking $-2\mu M_3$ we pre-compensate for directions in which the distribution of the noise has heavy tails. This decreases the statistical variability of the predicted stage cost. The second term, $x_t'(4\mu QWQ)x_t$, is a standard quadratic penalty term; notice that, contrary to the risk-neutral case, the covariance of the noise W now affects the penalty term. Informally, this term penalizes state directions which not only lead to high cost but are also more sensitive to noise, as captured by the product QWQ. Hence, the risk-neutral LQR framework can exhibit inherent risk-averse properties, provided that its parameters are selected in a principled way. Of course, selecting those parameters *a priori* is not trivial.

The structure of the Lagrangian as suggested by Lemma 1 enables us to derive both a closed-form expression for its minimum and an explicit optimal control policy. To this end, define the *optimal cost-to-go* at stage $k \leq N-1$ as

$$\mathcal{L}_k^*(x_k,\mu) \triangleq \inf_{u_{k:N-1} \in \mathcal{U}_k} \mathbb{E}\left\{ g_N(x_N,\mu) + \sum_{t=k}^{N-1} g_t(x_t,u_t,\mu) \,\middle|\, \mathscr{F}_k \right\},$$

where we omit the constant components of the Lagrangian. Under this definition, it is true that

$$D(\mu) \equiv \inf_{u \in \mathcal{U}_0} \mathcal{L}(u,\mu) = \mathcal{L}_0^*(x_0,\mu) + g(\mu).$$

We may now derive the complete solution to (6), which is one of the main results of this paper, and provides optimal risk-aware control policies for every fixed multiplier $\mu \geq 0$.

Theorem 2 (Optimal Risk-Aware Controls). *Let Assumption 1 be in effect, choose $\mu \geq 0$, and adopt the notation of Lemma 1. For $t \leq N-1$, the optimal cost-to-go $\mathcal{L}_t^*(x_t,\mu)$ may be expressed as*

$$\mathcal{L}_t^*(x_t,\mu) = x_t' V_t x_t + 4\mu M_3' S_t' x_t + 2\bar{w}' T_t' x_t + c_t, \quad (9)$$

where the quantities V_t, S_t, T_t and c_t are evaluated through the backward recursions

$$V_{t-1} = A' V_t A + Q_\mu - A' V_t B (B' V_t B + R)^{-1} B' V_t A, \quad (10)$$
$$K_{t-1} = -(B' V_t B + R)^{-1} B' V_t A, \quad (11)$$
$$S_{t-1} = (A + BK_{t-1})' S_t + Q, \quad (12)$$
$$T_{t-1} = (A + BK_{t-1})'(T_t + V_t), \quad (13)$$
$$l_{t-1} = -2\mu (B' V_t B + R)^{-1} B' S_t M_3, \quad (14)$$
$$h_{t-1} = -(B' V_t B + R)^{-1} B' (V_t + T_t)\bar{w} \quad \text{and} \quad (15)$$
$$c_{t-1} = c_t + \operatorname{tr}(WV_t) + \bar{w}'(2T_t' + V_t)\bar{w} + 4\mu M_3' S_N' \bar{w} - (l_{t-1} + h_{t-1})'(B' V_t B + R)(l_{t-1} + h_{t-1}), \quad (16)$$

with terminal values $V_N = Q_\mu$, $S_N = Q$, $T_N = 0$ and $c_N = 0$. Additionally, an optimal control policy that achieves the dual value in (6) *may be expressed as*

$$u_t^*(\mu) = K_t x_t + l_t + h_t \in \mathcal{L}_2(\mathscr{F}_t),\ \forall t \leq N-1, \quad (17)$$

and is unique up to sets of probability measure zero.

Proof. By using dynamic programming and assuming (temporarily) that involved measurability issues are resolved ([40], Appendix A), we have, for every $k \leq N-1$, the recursive optimality condition (i.e., the Bellman equation)

$$\mathcal{L}_k^*(x_k,\mu) = \inf_{u_k} \mathbb{E}\left\{ g_k(x_k,u_k,\mu) + \mathcal{L}_{k+1}^*(x_{k+1},\mu) \,\middle|\, \mathscr{F}_k \right\},$$

where minimization is taken *pointwise* (i.e., over constants) over the control space $\mathbb{R}^p$, and with $\mathcal{L}_N^*(x_N,\mu) = x_N' Q x_N$. We will prove the result by induction. The base case $k = N$ is immediate. Assume that (9) is true for $k = t+1$; we will show that this implies the same for $k = t$. By Lemma 1, and after standard algebraic manipulations, we obtain

$$\begin{aligned} &\mathbb{E}\left\{ g_t(x_t,u_t,\mu) + \mathcal{L}_{t+1}^*(x_{t+1},\mu) \,\middle|\, \mathscr{F}_t \right\} = u_t'(B' V_{t+1} B + R) u_t \\ &+ 2\left\{ x_t' A' V_{t+1}' + \bar{w}'(V_{t+1} + T_{t+1}') + 2\mu M_3' S_{t+1}' \right\} B u_t \\ &+ x_t'(A' V_{t+1} A + \bar{Q}) x_t + c_{t+1} \\ &+ 2\left\{ 2\mu M_3' S_{t+1}' + \bar{w}'(V_{t+1} + T_{t+1}') \right\} A x_t \\ &+ \bar{w}'(2T_t' + V_t)\bar{w} + 4\mu M_3' S_N' \bar{w} + \operatorname{tr}(WV_{t+1}), \end{aligned} \quad (18)$$

where we have exploited the identities $x_{t+1} = Ax_t + Bu_t + w_{t+1}$, $\mathbb{E}(w_{t+1}|\mathscr{F}_t) = \bar{w}$ and $\mathbb{E}(w_{t+1} - \bar{w})(w_{t+1} - \bar{w})' = W$. The reader may also verify that all measurability issues are

Authorized licensed use limited to: University of Pennsylvania. Downloaded on October 29,2026 at 15:14:47 UTC from IEEE Xplore. Restrictions apply.

now resolved in a recursive way, retrospectively. The unique stationary point of the convex quadratic (18) is

$$u_t^* = K_t x_t + l_t + h_t,$$

which may be readily verified to lie in $\mathcal{L}_2(\mathscr{F}_t)$, as well. Replacing u_t^* into (18) yields the optimal cost-to-go (9). □

As suggested by Remark 1, it turns out that the optimal controller (17) is affine with respect to the state. The noise-dependent term ℓ_t forces the state to track the reference $-2\mu M_3$, which points away from heavy-tailed regions of the noise distribution. Meanwhile, the state-feedback term accounts for the internal dynamics. Similar to the risk-neutral case, the controller's behavior is governed the Riccati difference equation (10). However, we now have an inflated stage cost matrix $\bar{Q} = Q + 4\mu QWQ$, instead of the original. As suggested by the product QWQ, the risk-aware control gain becomes more strict in directions that are simultaneously more costly and prone to noise, as captured by the covariance W. Finally, the term h_t acts against the mean value of the noise–such a term also appears in risk-neutral LQR.

As a corollary, from standard LQR theory, we immediately obtain that for any $\mu \geq 0$, the optimal controller (17) will be internally stable, i.e., the spectral radius will be bounded $\rho(A + BK_t) < 1$, as the horizon N grows to infinity.

Corollary 1 (Internal Stability). *Let Assumptions 1 and 2 be in effect, and adopt the notation of Lemma 1. For fixed $\mu \geq 0$, consider the control policy $u^*(\mu)$, as defined in* (17)*. As $N \to \infty$, V_t converges exponentially fast to the unique stabilizing solution*[1] *of the algebraic Riccati equation*

$$V = A'VA + Q_\mu - A'VB(B'VB + R)^{-1}B'VA.$$

As a result, for every $t \geq 0$, it is true that, as $N \to \infty$,

$$\begin{aligned}
K_t &\to K \triangleq -(B'VB + R)^{-1}B'VA,\\
S_t &\to S \triangleq (I - (A + BK)')^{-1}Q,\\
T_t &\to T \triangleq (I - (A + BK)')^{-1}(A + BK)'V,\\
l_t &\to l \triangleq -2\mu(B'VB + R)^{-1}B'SM_3 \quad \textit{and}\\
h_t &\to h \triangleq -(B'VB + R)^{-1}B'(V + T)\bar{w},
\end{aligned}$$

exponentially fast, and the closed-loop matrix $A + BK$ is stable (spectral radius $\rho(A + BK) < 1$).

Proof. Since $Q_\mu \succeq Q$ and $(A, Q^{1/2})$ is detectable, the pair $(A, Q_\mu^{1/2})$ is also detectable. Since (A, B) is stabilizable, $(A, Q_\mu^{1/2})$ is detectable, and $R \succ 0$, the exponential convergence of V_t and K_t to V and K respectively, and the stability of $A+BK$ follow from standard LQR theory [41, Chapter 4]. The proof of the convergence of the remaining terms follows similar steps and is omitted due to space constraints. □

Up to now we have discussed the properties of the optimal controller given a fixed $\mu \geq 0$. In what follows, we show how to compute an optimal multiplier μ^*, which will also satisfy the sufficient conditions for optimality suggested by Theorem 1, as stated previously.

[1]A stabilizing solution renders $A + BK$ stable.

C. Recovery of Primal-Optimal Solutions

From Theorem 2, we know how to compute the relaxed optimal input $u^*(\mu)$, for any given multiplier $\mu \geq 0$. But the risk constraint of the primal problem (4) is the only one that has been dualized in the construction of the Lagrangian in (5). Then, it turns out that we can also compute an optimal multiplier μ^* via bisection, thus providing a complete solution to the primal problem. We exploit the fact that, under the relaxed optimal policy $u^*(\cdot)$, both the LQR cost $J(u^*(\cdot))$ and the risk functional $J_R(u^*(\cdot))$ are monotone functions.

Theorem 3 (Primal-Optimal Solution). *Let Assumption 1 be in effect, and consider the control policy $u^*(\mu), \mu \geq 0$, as defined in* (17)*. Then, the following statements are true:*

1) *The LQR cost $J(u^*(\mu))$ is increasing with $\mu \geq 0$, while the risk constraint functional $J_R(u^*(\mu))$ is decreasing.*
2) *Define the multiplier*

$$\mu^* \triangleq \inf\{\mu \geq 0 :\ J_R(u^*(\mu)) \leq \bar{\epsilon}\}. \tag{19}$$

If μ^* is finite, then the policy $u^*(\mu^*)$ is optimal for the primal problem (4)*, and this is the case as long as* (4) *satisfies Slater's condition.*

Proof. To prove part 1), let $\mu_2 > \mu_1 \geq 0$. From the definition of the Lagrangian and optimality of the controller $u^*(\mu)$, we obtain the inequalities

$$\begin{aligned}
J(u^*(\mu_1)) + \mu_1 J_R(u^*(\mu_1)) &\leq J(u^*(\mu_2)) + \mu_1 J_R(u^*(\mu_2))\\
J(u^*(\mu_1)) + \mu_2 J_R(u^*(\mu_1)) &\geq J(u^*(\mu_2)) + \mu_2 J_R(u^*(\mu_2)).
\end{aligned}$$

By subtracting, we get

$$(\mu_2 - \mu_1)\{J_R(u^*(\mu_1)) - J_R(u^*(\mu_2))\} \geq 0,$$

which shows that $J_R(u^*(\mu_1)) \geq J_R(u^*(\mu_2))$. The proof of $J(u^*(\mu_1)) \leq J(u^*(\mu_2))$ is similar.

To prove part 2), we first show that, whenever $\mu^* < \infty$, $\mu^*(J_R(u^*(\mu^*)) - \bar{\epsilon}) = 0$, i.e., complementary slackness holds. We have two cases: either $\mu^* = 0$, where complementary slackness is satisfied trivially; or $\mu^* > 0$, $J_R(u^*(\mu^*)) \leq \bar{\epsilon}$. Therefore, it will be sufficient to show that in the latter case we can only have $J_R(u^*(\mu^*)) = \bar{\epsilon}$. Since $\mu^* > 0$, it is true that $J_R(u^*(0)) > \bar{\epsilon}$. From Theorem 2 and Proposition 2, it also follows that the function $J_R(u^*(\mu))$ is continuous with respect to μ (all matrix inverses in (10) are continuous since $R \succ 0$). Now, assume that $J_R(u^*(\mu^*)) < \bar{\epsilon}$. Then by continuity, there exists a $0 < \bar{\mu} < \mu^*$ such that $J_R(u^*(\bar{\mu})) = \bar{\epsilon}$, contradicting the definition of μ^*. Hence, we can only have $J_R(u^*(\mu^*)) = \bar{\epsilon}$, which shows that complementary slackness is satisfied.

Now, complementary slackness, along with the trivial fact that $J_R(u^*(\mu^*)) \leq \bar{\epsilon}$ imply that the policy-multiplier pair $(u^*(\mu^*), \mu^*) \in \mathcal{U}_0 \times \mathbb{R}_+$ satisfies the sufficient conditions for optimality provided by Theorem 1. Enough said.

To prove the last claim of part 2), suppose that (4) satisfies Slater's condition, i.e., that there is an admissible policy $u^\dagger$ in $\mathcal{U}_0$ such that $J_R(u^\dagger) - \bar{\epsilon} < 0$. For every $\mu \geq 0$, we have

$$\begin{aligned}
&D(\mu) \leq J(u^\dagger) + \mu(J_R(u^\dagger) - \bar{\epsilon})\\
\implies\ &D(\mu) - \mu(J_R(u^\dagger) - \bar{\epsilon}) \leq J(u^\dagger) < \infty.
\end{aligned}$$

Authorized licensed use limited to: University of Pennsylvania. Downloaded on October 29,2026 at 15:14:47 UTC from IEEE Xplore. Restrictions apply.

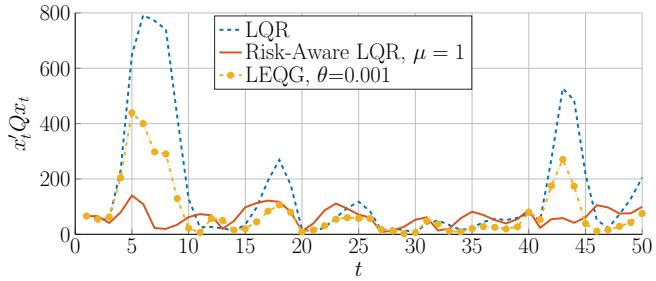


Fig. 2. Evolution of the state penalties $x_k' Q x_k$, over the first 50 steps. Notice that our risk-aware LQR controller indeed limits the variability of $x_k' Q x_k$. In fact, it sacrifices performance under small wind forces, but protects the system against large wind gusts, for example at time $5-10$.

Next, suppose that, for every $\mu \geq 0$, $J_R(u^*(\mu)) - \bar{\epsilon} \geq 0$. Because $J(u^*(\cdot))$ is increasing on $\mathbb{R}_+$, it must be true that

$$\begin{aligned} J(u^\dagger) &\geq \sup_{\mu \geq 0} D(\mu) - \mu(J_R(u^\dagger) - \bar{\epsilon}) \\ &= \sup_{\mu \geq 0} J(u^*(\mu)) + \mu(J_R(u^*(\mu)) - \bar{\epsilon}) - \mu(J_R(u^\dagger) - \bar{\epsilon}) \\ &= \infty, \end{aligned}$$

which contradicts the fact that $J(u^\dagger) < \infty$. Therefore, there must exist $\mu^\dagger \geq 0$, such that $J_R(u^*(\mu^\dagger)) - \bar{\epsilon} < 0$. But $J_R(u^*(\cdot))$ is decreasing on $\mathbb{R}_+$ and, consequently, it must be the case that $\mu^* \in [0, \mu^\dagger)$. The proof is now complete. □

Theorem 3 implies that we can find an optimal multiplier satisfying the optimality conditions on Theorem 1 from (19), by performing simple bisection on μ. Of course, this requires evaluating the risk constraint functional $J_R(u^*(\mu))$ for different values $\mu \geq 0$. The evaluation may be performed in a recursive fashion, as the following result suggests.

Proposition 2 (Risk Functional Evaluation). *Let Assumption 1 be in effect, and adopt the notation of Lemma 1. For fixed $\mu \geq 0$, consider the control policy $u^*(\mu)$, as defined in* (17)*. With terminal values $P_N = 4QWQ$, $z_N = 4M_3' Q$, $r_N = 0$, consider the backward recursions*

$$\begin{aligned} P_{t-1} &= (A + BK_{t-1})' P_t (A + BK_{t-1}) + 4QWQ, \\ z_{t-1} &= (A + BK_{t-1})' z_t + 4QM_3 \\ &\quad + 2(A + BK_{t-1})' P_t (Bh_{t-1} + Bl_{t-1} + \bar{w}) \quad \textit{and} \\ r_{t-1} &= r_t + \mathrm{tr}(P_t W) + z_t'(\bar{w} + Bl_{t-1} + Bh_{t-1}) \\ &\quad + (Bl_{t-1} + Bh_{t-1} + \bar{w})' P_t (Bl_{t-1} + Bh_{t-1} + \bar{w}). \end{aligned}$$

Then, the risk constraint in problem (4) *may be evaluated by*

$$J_R(u^*(\mu)) = x_0' P_0 x_0 + z_0' x_0 + r_0.$$

Proof. Omitted; it is similar to that of Theorem 2. □

IV. Simulations And Discussion

Consider a flying robot that moves on a horizontal plane, i.e., the Euclidean space $\mathbb{R}^2$. We assume that its linearized dynamics can be abstracted by a double integrator as

$$x_{k+1} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix} x_k + \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{bmatrix} (v_k + d_k), \quad (20)$$

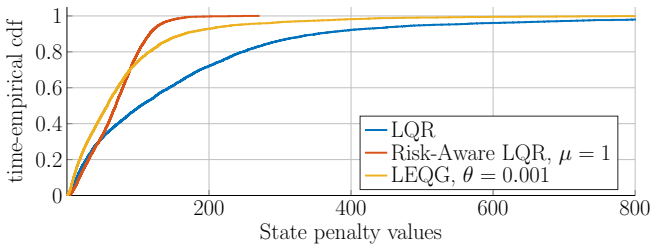


Fig. 3. The time-empirical cdf for the state penalties $x_k' Q x_k$, $k \leq N$, for the LQR (risk-neutral), our method, and LEQG. Our method sacrifices some average performance but exhibits much smaller variability for the state penalties. It also protects the system against rare but large wind gusts.

where $x_{k,1}$, $x_{k,3}$ are the position coordinates, $x_{k,2}$, $x_{k,4}$ the respective velocities and v_k is the acceleration input. Let d_k be a wind disturbance force that acts on the robot, which is modeled as follows: We assume that $d_{k,1}$ constitutes the dominant wind direction with non-zero mean and large variability, while the orthogonal direction $d_{k,2}$ is a weak wind direction with zero mean and small variability. We model $d_{k,1}$ as a mixture of two gaussians $\mathcal{N}(30, 30)$, $\mathcal{N}(80, 60)$ with weights 0.8 and 0.2, respectively. This bimodal distribution models the presence of infrequent but large wind gusts. The weak direction $d_{k,2}$ is modeled as zero-mean Gaussian $\mathcal{N}(0, 5)$. If we cancel the mean of d_k by applying $v_k = u_k - \mathbb{E}d_k$, then (20), can be written in terms of (1), where $w_k = B(d_k - \mathbb{E}d_k)$ is now a zero-mean disturbance.

Consider now the LQR problem with parameters

$$Q = \mathrm{diag}(1, 0.1, 2, 0.1) \quad \text{and} \quad R = I,$$

and a horizon of length $N = 5000$. We primarily compare our risk-aware LQR formulation with the classical, risk-neutral LQR via simulations. To tune our controller, we vary μ in (17) directly instead of varying ϵ. We also (heuristically) compare our controller with the exponential (LEQG) method, even though the noise is not Gaussian, by plugging in the second order statistics W. Let the tuning parameter of LEQG be θ. Note that the exponential problem is well defined only if $\theta < 0.001276$ (roughly), where the "neurotic breakdown" occurs [13]. For the purpose of comparison, we simulate all schemes under the same noise sequence $w_{0:N}$.

In Fig. 2, we see the evolution of the state penalty terms $x_k' Q x_k$, for the first 50 time steps, under the different control schemes. By slightly sacrificing performance under small wind forces, our risk-aware LQR controller forces the state to have less variability and, thus, protects the robot against large gusts. On the other hand, the state of the robot state can grow very large under the risk-neutral *and* LEQG schemes. This behavior is illustrated even more clearly in Fig. 3, where we present the time-empirical cumulative distribution of the state penalties for all N time steps. The time-empirical "probability" of suffering large state penalties is drastically smaller compared to LQR or LEQG.

To better illustrate how the proposed risk-aware controller works, we also discuss the evolution of the position $x_{k,1}$ and the input $u_{k,1}$, as shown in Fig. 4, for the first 50 steps. First, we observe that the controller pushes the state $x_{k,1}$ towards negative values, away from the direction of

Authorized licensed use limited to: University of Pennsylvania. Downloaded on October 29,2026 at 15:14:47 UTC from IEEE Xplore. Restrictions apply.

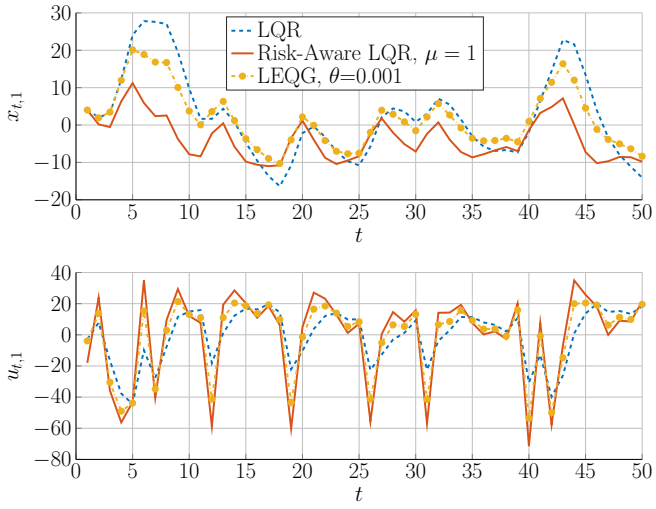


Fig. 4. Evolution of the state $x_{k,1}$, and the input $u_{k,1}$ over the first 50 steps. The controller pushes the state away from the direction of the large gusts, which helps the robot to avoid extreme perturbations. Meanwhile, by inflating the state penalty with the μQWQ, we force the state-feedback component to be more cautious with the state. Naturally, being more cautious with the state requires extra control effort.

the large gusts. Second, notice that we penalize $x_{k,3}$ more in Q. In fact, the risk-neutral LQR results in the steady state gains $K_{\text{LQR},11} = -0.697$, $K_{\text{LQR},12} = -1.201$, $K_{\text{LQR},23} = -0.925$, $K_{\text{LQR},24} = -1.376$, i.e., it is stricter with direction $x_{k,3}$. However, $x_{k,1}$ exhibits more variability due to the strong wind direction. In contrast, our risk-aware scheme adapts to the noise in a principled way. Due to the inflation term μQWQ, our scheme returns the steady-state gains $K_{11} = -2.1008$, $K_{12} = -2.2132$, $K_{23} = -1.1161$, $K_{24} = -1.5131$, which means that the risky direction $x_{k,1}$ is controlled more strictly. Naturally, being more cautious with the state leads to higher control effort. Lastly, although the LEQG controller is also more state-cautious, it is agnostic to the heavy tails of the wind distribution. Hence, it still suffers from large perturbation due to the wind gusts.

V. Conclusion and Future Work

We presented a new risk-aware reformulation of the classical LQR problem, where we introduce a new risk measure to be used as an explicit and tunable risk constraint, along with the standard LQR objective. By restricting the expected cumulative predictive variance of the state penalties, we can decrease the variability of the state at will, protecting the system against uncommon but strong random disturbances. The optimal controller enjoys a simple closed-form expression with clear interpretation, is always stable and is easy to tune. Our scheme works for arbitrary noise process distributions, as long as the corresponding fourth-order moments are finite.

Moving forward, our framework opens up many directions for extensions and future research. First, we would like to note that our analysis does not depend on the constraints having the same matrix Q as in the cost. In fact, we can define our risk constraint as

$$\mathbb{E}\left\{\sum_{t=1}^{N}\left[x_t'Q_cx_t - \mathbb{E}\left(x_t'Q_cx_t|\mathscr{F}_{t-1}\right)\right]^2\right\} \leq \epsilon,$$

where Q_c is a design choice. Proposition 1 still holds, in the sense that the constraint can be rewritten as a quadratic one. This implies that, thanks to its simplicity, the predictive variance constraint can be easily incorporated in more general problems, e.g., MPC or classical constrained-LQR, adding a risk-aware flavor to them. Another possibility is to employ stage-wise constraints of the form

$$\mathbb{E}\left[x_t'Q_tx_t - \mathbb{E}\left(x_t'Q_tx_t|\mathscr{F}_{t-1}\right)\right]^2 \leq \epsilon_t, \quad t \leq N.$$

In this case, the optimal controller will be similar to (17), but will depend on multiple Lagrange multipliers that can be optimized with primal-dual algorithms. Lastly, our predictive variance constraint is based on one-step-ahead prediction. In some cases, careless selection of Q_c might make our controller more myopic. At the same time, though, increasing the prediction horizon might not always preserve the quadratic form of the constraint. In future work, we would also like to address this issue.

Appendix: Proof of Proposition 1

Let $\Delta_t \triangleq x_t'Qx_t - \mathbb{E}\left(x_t'Qx_t|\mathscr{F}_{t-1}\right)$ be the prediction error of the stage penalty at time t given $\mathscr{F}_{t-1}$. We proceed in two steps. First, we show that Δ_t is well-defined and belongs to $\mathcal{L}_2(\mathscr{F}_t)$. Second, we obtain the closed form expression for the expected predictive variance $\mathbb{E}\left\{\Delta_t^2\right\}$.

Step a). The state x_t of the system depends linearly on past inputs u_k as well as past noises w_{k+1}, for $k \leq t-1$. Under the constraint $u_k \in \mathcal{L}_2(\mathscr{F}_k)$, and since by Assumption 1 $w_k \in \mathcal{L}_2(\mathscr{F}_k)$, it also follows that $x_t \in \mathcal{L}_2(\mathscr{F}_t)$, for all $t \leq N-1$. As a result, the expectation of $x_t'Qx_t$ exists and any conditional expectation $\mathbb{E}\left(x_t'Qx_t|\mathscr{F}_{t-1}\right)$ is well-defined and finite almost everywhere, for all $t \leq N-1$. Define

$$\hat{x}_t \triangleq \mathbb{E}(x_t|\mathscr{F}_{t-1}) = Ax_{t-1} + Bu_{t-1} + \bar{w} \quad \text{and} \tag{21}$$

$$\delta_t \triangleq w_t - \bar{w}. \tag{22}$$

Note that $\hat{x}_t$ is well-defined since $x_t \in \mathcal{L}_2(\mathscr{F}_t)$. Replacing x_t with $\hat{x}_t + \delta_t$, we obtain the representation

$$x_t'Qx_t = \hat{x}_t'Q\hat{x}_t + 2\hat{x}_t'Q\delta_t + \delta_t'Q\delta_t.$$

All of the terms above are integrable since $\hat{x}_t$, δ_t are square-integrable. Since $\hat{x}_t$ is measurable with respect to $\mathscr{F}_{t-1}$, the expectation of $x_t'Qx_t$ conditioned on $\mathscr{F}_{t-1}$ is

$$\mathbb{E}\left(x_t'Qx_t|\mathscr{F}_{t-1}\right) = \hat{x}_t'Q\hat{x}_t + \operatorname{tr}(WQ).$$

Then, the difference of the above quantities is

$$\Delta_t = \delta_t'Q\delta_t - \operatorname{tr}(WQ) + 2\hat{x}_t'Q\delta_t.$$

Computing the squares of both sides leads to the expression

$$\begin{aligned}\Delta_t^2 = (\delta_t'Q\delta_t - \operatorname{tr}(WQ))^2 + 4\hat{x}_t'Q\delta_t\delta_t'Q\hat{x}_t \\ + 4\hat{x}_t'Q\delta_t(\delta_t'Q\delta_t - \operatorname{tr}(WQ)).\end{aligned} \tag{23}$$

Note that all of the above terms are integrable, hence $\mathbb{E}\left\{\Delta_t^2\right\}$ is well-defined and finite. Integrability of the first term follows from Assumption 1. Integrability of the second term comes from the fact that $\hat{x}_t$ and δ_t are square-integrable and independent of each other. Similarly, itegrability of the

Authorized licensed use limited to: University of Pennsylvania. Downloaded on October 29,2026 at 15:14:47 UTC from IEEE Xplore. Restrictions apply.

last term follows from integrability of $\hat{x}_t$, Assumption 1 and independence of $\hat{x}_t, \delta_t$.

Step b). From (23), it is true that

$$\mathbb{E}\left\{\Delta_t^2 | \mathscr{F}_{t-1}\right\} = 4\hat{x}_t' QWQ\hat{x}_t + m_4 + 4\hat{x}_t' QM_3.$$

Taking expectation again gives

$$\mathbb{E}\left\{\Delta_t^2\right\} = m_4 + \mathbb{E}(4\hat{x}_t' QWQ\hat{x}_t + 4\hat{x}_t' QM_3).$$

By orthogonality of $\hat{x}_t$, δ_t, and since $\mathbb{E}\delta_t = 0$, $\mathbb{E}\delta_t\delta_t' = W$,

$$\mathbb{E}\left\{\Delta_t^2\right\} = \mathbb{E}(4x_t' QWQx_t + 4x_t' QM_3) + m_4 - 4\,\mathrm{tr}\left\{(WQ)^2\right\}.$$

The result follows if we replace $\mathbb{E}\left\{\Delta_t^2\right\}$ with the right-hand side above in the risk constraint $\sum_{t=1}^{N} \mathbb{E}\left\{\Delta_t^2\right\} \leq \epsilon$. □

REFERENCES

[1] A. Ahlén, J. Akerberg, M. Eriksson, A. L. F. J. Isaksson, T. Iwaki, K. H. Johansson, S. Knorn, T. Lindh, and H. Sandberg, "Towards Wireless Control in Industrial Process Automation: A Case Study at a Paper Mill," *Control Systems, IEEE*, vol. 39, no. 5, pp. 36–57, 2019.

[2] 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*, vol. 250, no. 3, pp. 979–989, May 2016.

[3] S. Moazeni, W. B. Powell, and A. H. Hajimiragha, "Mean-Conditional Value-at-Risk Optimal Energy Storage Operation in the Presence of Transaction Costs," *IEEE Transactions on Power Systems*, vol. 30, no. 3, pp. 1222–1232, May 2015.

[4] H. Markowitz, "Portfolio Selection," *The Journal of Finance*, vol. 7, no. 1, pp. 77–91, Mar. 1952.

[5] H. Föllmer and A. Schied, "Convex Measures of Risk and Trading Constraints," *Finance and Stochastics*, vol. 6, no. 4, pp. 429–447, Oct. 2002.

[6] D. Shang, V. Kuzmenko, and S. Uryasev, "Cash Flow Matching with Risks Controlled by Buffered Probability of Exceedance and Conditional Value-at-Risk," *Annals of Operations Research*, vol. 260, no. 1-2, pp. 501–514, Jan. 2018.

[7] S.-K. Kim, R. Thakker, and A.-a. Agha-mohammadi, "Bi-directional Value Learning for Risk-aware Planning Under Uncertainty," *IEEE Robotics and Automation Letters*, vol. 4, no. 3, pp. 2493–2500, July 2019.

[8] 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*, vol. 30, no. 5, pp. 741–762, Sept. 2013.

[9] W.-J. Ma, C. Oh, Y. Liu, D. Dentcheva, and M. M. Zavlanos, "Risk-Averse Access Point Selection in Wireless Communication Networks," *IEEE Transactions on Control of Network Systems*, vol. 5870, no. c, pp. 1–1, 2018.

[10] S. Samuelson and I. Yang, "Safety-Aware Optimal Control of Stochastic Systems Using Conditional Value-at-Risk," in *2018 Annual American Control Conference (ACC)*, 2018, pp. 6285–6290.

[11] M. P. Chapman, J. Lacotte, A. Tamar, D. Lee, K. M. Smith, V. Cheng, J. F. Fisac, S. Jha, M. Pavone, and C. J. Tomlin, "A Risk-Sensitive Finite-Time Reachability Approach for Safety of Stochastic Dynamic Systems," in *2019 American Control Conference (ACC)*, 2019, pp. 2958–2963.

[12] D. Jacobson, "Optimal Stochastic Linear Systems with Exponential Performance Criteria and their Relation to Deterministic Differential Games," *IEEE Transactions on Automatic Control*, vol. 18, no. 2, pp. 124–131, 1973.

[13] P. Whittle, "Risk-Sensitive Linear/Quadratic/Gaussian Control," *Adv. Appl. Prob*, vol. 13, pp. 764–777, 1981.

[14] T. Başar, "Risk-Averse Designs: From Exponential Cost to Stochastic Games," in *System Theory*, pp. 131–143. Springer, 2000.

[15] K. D. Pham, *Linear-Quadratic Controls in Risk-Averse Decision Making: Performance-Measure Statistics and Control Decision Optimization*, Springer Science & Business Media, 2012.

[16] V. Roulet, M. Fazel, S. Srinivasa, and Z. Harchaoui, "On the Convergence of the Iterative Linear Exponential Quadratic Gaussian Algorithm to Stationary Points," 2019.

[17] J. Speyer, C.-H. Fan, and R. Banavar, "Optimal Stochastic Estimation with Exponential Cost Criteria," in *Proceedings of the 31st IEEE Conference on Decision and Control*. Aug. 1992, pp. 2293–2298, Institute of Electrical and Electronics Engineers (IEEE).

[18] S. Dey and J. B. Moore, "Finite-Dimensional Risk-Sensitive Filters and Smoothers for Discrete-Time Nonlinear Systems," *IEEE Transactions on Automatic Control*, vol. 44, no. 6, pp. 1234–1239, 1999.

[19] J. B. Moore, R. J. Elliott, and S. Dey, "Risk-Sensitive Generalizations of Minimum Variance Estimation and Control," *Journal of Mathematical Systems, Estimation, and Control*, vol. 7, no. 1, pp. 123–126, 1997.

[20] S. Dey and J. B. Moore, "Risk-Sensitive Filtering and Smoothing via Reference Probability Methods," *IEEE Transactions on Automatic Control*, vol. 42, no. 11, pp. 1587–1591, 1997.

[21] N. Bäuerle and U. Rieder, "More Risk-Sensitive Markov Decision Processes," *Mathematics of Operations Research*, vol. 39, no. 1, pp. 105–120, 2014.

[22] P. L. A. and M. Fu, "Risk-Sensitive Reinforcement Learning: A Constrained Optimization Viewpoint," *arXiv preprint, arXiv:1810.09126*, Oct. 2018.

[23] A. R. Cardoso and H. Xu, "Risk-Averse Stochastic Convex Bandit," in *International Conference on Artificial Intelligence and Statistics*, Apr. 2019, vol. 89, pp. 39–47.

[24] W. Huang and W. B. Haskell, "Risk-Aware Q-learning for Markov Decision Processes," in *2017 IEEE 56th Annual Conference on Decision and Control, CDC 2017*. Dec. 2018, vol. 2018-Janua, pp. 4928–4933, IEEE.

[25] D. R. Jiang and W. B. Powell, "Risk-Averse Approximate Dynamic Programming with Quantile-Based Risk Measures," *Mathematics of Operations Research*, vol. 43, no. 2, pp. 554–579, Nov. 2018.

[26] D. S. Kalogerias and W. B. Powell, "Recursive Optimization of Convex Risk Measures: Mean-Semideviation Models," *arXiv preprint, arXiv:1804.00636*, Apr. 2018.

[27] A. Tamar, Y. Chow, M. Ghavamzadeh, and S. Mannor, "Sequential Decision Making with Coherent Risk," *IEEE Transactions on Automatic Control*, vol. 62, no. 7, pp. 3323–3338, July 2017.

[28] C. A. Vitt, D. Dentcheva, and H. Xiong, "Risk-Averse Classification," *Annals of Operations Research*, Aug. 2019.

[29] L. Zhou and P. Tokekar, "An Approximation Algorithm for Risk-averse Submodular Optimization," *arXiv preprint, arXiv:1807.09358*, July 2018.

[30] A. Ruszczyński, "Risk-Averse Dynamic Programming for Markov Decision Processes," *Mathematical Programming*, vol. 125, no. 2, pp. 235–261, Oct. 2010.

[31] P. Sopasakis, D. Herceg, A. Bemporad, and P. Patrinos, "Risk-Averse Model Predictive Control," *Automatica*, vol. 100, pp. 281 – 288, 2019.

[32] A. Shapiro, D. Dentcheva, and A. Ruszczyński, *Lectures on Stochastic Programming: Modeling and Theory*, Society for Industrial and Applied Mathematics, 2nd edition, 2014.

[33] R. T. Rockafellar and S. Uryasev, "Optimization of Conditional Value-at-Risk," *Journal of Risk*, vol. 2, pp. 21–41, 1997.

[34] M. Abeille, A. Lazaric, X. Brokmann, et al., "LQG for Portfolio Optimization," *Available at SSRN: https://ssrn.com/abstract=2863925 or http://dx.doi.org/10.2139/ssrn.2863925*, 2016.

[35] I. Tzortzis, C. D. Charalambous, T. Charalambous, C. K. Kourtellaris, and C. N. Hadjicostis, "Robust Linear Quadratic Regulator for Uncertain Systems," in *2016 IEEE 55th Conference on Decision and Control (CDC)*, 2016, pp. 1515–1520.

[36] S. Dean, H. Mania, N. Matni, B. Recht, and S. Tu, "On the Sample Complexity of the Linear-Quadratic Regulator," *arXiv preprint arXiv:1710.01688*, 2017.

[37] K. Glover and J. C. Doyle, "State-Space Formulae for All Stabilizing Controllers that Satisfy an H_∞-norm Bound and Relations to Risk Sensitivity," *Systems & control letters*, vol. 11, no. 3, pp. 167–172, 1988.

[38] D. P. Bertsekas, *Dynamic Programming and Optimal Control*, vol. 1, Athena Scientific, 4th edition, 2017.

[39] A. Ruszczynski, *Nonlinear Optimization*, Princeton university press, 2011.

[40] D. P. Bertsekas, *Dynamic Programming and Optimal Control*, vol. 2: Approximate Dynamic Programming, Athena Scientific, 4th edition, 2012.

[41] B. Anderson and J. Moore, *Optimal Filtering*, Dover Publications, 2005.

Authorized licensed use limited to: University of Pennsylvania. Downloaded on October 29,2026 at 15:14:47 UTC from IEEE Xplore. Restrictions apply.