Structural properties of optimal risk-aware controllers for spatially invariant systems

Juncal Arbelaiz[1], *Member, IEEE* Bassam Bamieh[2], *Fellow, IEEE* and Naomi Ehrich Leonard[3], *Fellow, IEEE*

***Abstract*— We analyze the optimal Linear Exponential Quadratic Gaussian (LEQG) control synthesis of a spatially distributed system with a shift invariance in its spatial coordinate, perturbed by additive white Gaussian noise. We refer to such a system as spatially invariant. The LEQG framework accounts for the risk attitude of the controller in its synthesis by appropriate selection of the value of a free parameter, providing the possibility to continuously tune the degree of risk-awareness of the controller. We prove important structural properties of the optimal LEQG control problem for spatially invariant systems, namely that: (i) the optimal LEQG control gain is spatially invariant itself; (ii) the LEQG control synthesis problem is equivalent to a family of decoupled LEQG optimization problems of smaller dimension; and (iii) under some further assumptions, the optimal LEQG control gain is spatially localized. Through a case study, we illustrate how the risk attitude of the controller tunes the degree of spatial localization of the optimal control gain. We argue that the proven structural properties can be leveraged to reduce the computational complexity of obtaining the optimal LEQG control gain in large-scale systems and to design distributed risk-aware controller implementations.**

***Index Terms*— control of networks, distributed control, optimal control.**

I. Introduction

We study the optimal risk-sensitive control synthesis for large-scale spatially invariant systems (SIS). SIS are characterized by the fact that their dynamics remain invariant to a notion of translation in the spatial coordinate [1]. For concreteness, we focus on linear circulant systems, a class of finite-dimensional SIS. Circulant networks of dynamical systems are of importance as they are utilized to model a wide variety of applications, including autonomous mobile agents in closed tracks [2], vibrations in mechanical structures [3], cell ensembles [4], synchronization of oscillators [5] and finite-dimensional approximations of continuum systems described by partial differential equations (PDEs) with periodic boundary conditions. These applications are often of large-scale and involve uncertainties that should be taken into consideration in controller synthesis. Some of such uncertainties (e.g., unknown environmental influences) are commonly modeled through stochastic disturbances perturbing the system [6]. Hence, the control designer might aim to make the controller concerned with large ("risky") values of the cost capturing the performance of the system rather than accounting only for its statistically typical behavior or expected value. We refer to such controllers as *risk-sensitive* or *risk-aware*. A synthesis framework to design risk-sensitive controllers is the Linear Exponential Quadratic Gaussian (LEQG), first introduced by Jacobson [7], [8] and extensively developed by Whittle [9]–[11]. In this framework, the value of a free parameter θ tunes the risk attitude of the controller (i.e., its degree of conservativeness towards uncertainty), providing the possibility to design risk-averse ($\theta > 0$), risk-neutral ($\theta = 0$) or even risk-seeking ($\theta < 0$) controllers. As $\theta \to 0$ the LEQG problem converges to the LQG, and for the maximum admissible risk-averse value of θ the LEQG criterion is equivalent to the minimum entropy $\mathcal{H}_\infty$ objective in steady-state [12].

Main Contributions: SIS provide a useful idealization to take a first step towards understanding the effect of risk-awareness on controllers for spatially distributed systems. We prove important structural properties of LEQG controllers for circulant systems, extending the seminal results in [1]: we show that *the optimal LEQG gain is circulant*, that the LEQG problem *decouples* into a family of LEQG problems of smaller dimension, and that under some further assumptions the *optimal LEQG control gain decays exponentially*. A case study illustrates how the risk attitude of the controller affects its degree of spatial localization. We discuss useful practical implications of these results to efficiently compute the optimal LEQG gain and to reduce the operational complexity of the controller through distributed implementations in large-scale systems.

Paper Structure: Section II introduces notation and mathematical background required for the paper. Section III describes SIS, their diagonalization through the Discrete Fourier Transform, and properties of the Fourier-transformed noise process perturbing the diagonalized dynamics. Section IV presents the LEQG problem formulation. Section V contains our main results on the structural properties of LEQG controllers for SIS and a case study evaluating the spatial localization of the optimal LEQG control gain. This section also discusses important practical implications of our structural results.

The work is supported by a Schmidt Science Postdoctoral Fellowship from Schmidt Futures & the Rhodes Trust.

[1]Juncal Arbelaiz is with the Dept. of Mechanical & Aerospace Engineering and the Center for Statistics and Machine Learning, Princeton University, NJ 08544, USA jarbelaiz@schmidtsciencefellows.org

[2]Bassam Bamieh is with the Dept. of Mechanical & Environmental Engineering, University of California Santa Barbara, CA 91106, USA bamieh@ucsb.edu

[3]Naomi Ehrich Leonard is with the Dept. of Mechanical & Aerospace Engineering, Princeton University, NJ 08544, USA naomi@princeton.edu

II. Mathematical Preliminaries

· *Notation:* x denotes a scalar and bold $\boldsymbol{x}$ a vector. Capitals denote matrices. Let A be a complex matrix; $A^\top$ denotes its transpose, $\bar{A}$ its complex conjugate, and A^* its transpose conjugate (i.e., $A^* = \overline{A^\top}$). $\boldsymbol{c}^{(l)}$ refers to the l-th entry of $\boldsymbol{c}$.

· *Group structure of the spatial coordinate [13]:* we work with SIS in which the spatial coordinate varies in the integers modulo n, $\mathbb{Z}_n := \{0, 1, \dots, n-1\}$. $\mathbb{Z}_n$ is an abelian group with modular arithmetic: k equals l modulo n if $k - l$ is an integer multiple of n: $k \equiv_n l \Leftrightarrow \exists\, m \in \mathbb{Z}$ s. t. $k - l = mn$. Any two integers k and l with $k - l$ a multiple of n are members of the same *equivalence class*. Two integers in the same equivalence class represent the same element in $\mathbb{Z}_n$.

· *Complex-valued jointly Gaussian random vectors [14], [15]:* a $\mathbb{C}^n$-valued jointly Gaussian random vector $\boldsymbol{z}$ has n complex-valued entries $\boldsymbol{z}_k$ $(k = 0, 1, \dots, n-1)$ that are jointly Gaussian: $\{\Re(\boldsymbol{z}^{(k)}), \Im(\boldsymbol{z}^{(k)})\}_{k=0}^{n-1}$ is a set of $2n$ jointly Gaussian (real-valued) random variables. In general, the probability density of $\boldsymbol{z}$ is *not* fully characterized by its mean $\boldsymbol{\mu_z} = \mathbb{E}[\boldsymbol{z}]$ and covariance $\Sigma_{\boldsymbol{z}} := \mathbb{E}[(\boldsymbol{z} - \boldsymbol{\mu_z})(\boldsymbol{z} - \boldsymbol{\mu_z})^*]$; the pseudo-covariance $\Gamma_{\boldsymbol{z}} := \mathbb{E}[(\boldsymbol{z} - \boldsymbol{\mu_z})(\boldsymbol{z} - \boldsymbol{\mu_z})^\top]$ must be specified as well. We use the notation $\boldsymbol{z} \sim \mathcal{CN}\big(\boldsymbol{\mu_z}, \Sigma_{\boldsymbol{z}}, \Gamma_{\boldsymbol{z}}\big)$ for complex Gaussians. $\boldsymbol{z}$ can be viewed as $2n$ real-valued random variables:

$$\boldsymbol{\zeta} := \begin{bmatrix} \Re(\boldsymbol{z}) \\ \Im(\boldsymbol{z}) \end{bmatrix}, \; \boldsymbol{\zeta} \sim \mathcal{N}(\boldsymbol{\mu_\zeta}, \Sigma_{\boldsymbol{\zeta}}), \tag{1}$$

where $\boldsymbol{\mu_\zeta} = \begin{bmatrix} \Re(\boldsymbol{\mu_z}) \\ \Im(\boldsymbol{\mu_z}) \end{bmatrix}$ and $\Sigma_{\boldsymbol{\zeta}} = \frac{1}{2}\begin{bmatrix} \Re(\Sigma_{\boldsymbol{z}} + \Gamma_{\boldsymbol{z}}) & \Im(\Gamma_{\boldsymbol{z}} - \Sigma_{\boldsymbol{z}}) \\ \Im(\Sigma_{\boldsymbol{z}} + \Gamma_{\boldsymbol{z}}) & \Re(\Sigma_{\boldsymbol{z}} - \Gamma_{\boldsymbol{z}}) \end{bmatrix}$.

· *Circulant matrices & the Discrete Fourier Transform*:

Definition 2.1 (Discrete Fourier Transform, DFT): Given $\boldsymbol{x} \in \mathbb{C}^n$, its *Discrete Fourier Transform* $\hat{\boldsymbol{x}} \in \mathbb{C}^n$ has entries $\hat{\boldsymbol{x}}^{(k)} := \frac{1}{\sqrt{n}} \sum_{j=0}^{n-1} \boldsymbol{x}^{(j)}\, e^{-i\frac{2\pi kj}{n}}$ $(k = 0, \dots, n-1)$, with i the imaginary unit. Define the matrix F with entries $F_{jk} := \frac{1}{\sqrt{n}}\, e^{-i\frac{2\pi kj}{n}}$. Then, $\hat{\boldsymbol{x}} = F\boldsymbol{x}$. F is unitary ($F^{-1} = F^*$) and symmetric ($F = F^\top$).

Proposition 2.1 (Conjugate symmetry): The DFT $\hat{\boldsymbol{x}} := F\boldsymbol{x}$ of $\boldsymbol{x} \in \mathbb{R}^n$ is conjugate symmetric, that is, $\hat{\boldsymbol{x}}^{(n-m)} = \overline{\hat{\boldsymbol{x}}^{(m)}}$, for $m = 0, 1, \dots, n-1$.

Remark 2.2: Proposition 2.1 implies that the first $\lfloor \frac{n}{2} \rfloor + 1$ entries of $\hat{\boldsymbol{x}}$ suffice to completely define $\hat{\boldsymbol{x}}$ when $\boldsymbol{x} \in \mathbb{R}^n$.

Definition 2.2 (Circulant right-shift operator, S): Let S denote the *circulant right-shift operator* defined by its action on vectors: $S(\boldsymbol{x}^{(0)}, \boldsymbol{x}^{(1)}, \dots, \boldsymbol{x}^{(n-1)})^\top := (\boldsymbol{x}^{(n-1)}, \boldsymbol{x}^{(0)}, \dots, \boldsymbol{x}^{(n-2)})^\top$. S^* is the circular left-shift operator and $S^*S = SS^* = I$.

Definition 2.3 (Circulant matrix): For $\boldsymbol{c} = (\boldsymbol{c}^{(0)}, \boldsymbol{c}^{(1)}, \dots, \boldsymbol{c}^{(n-1)})^\top \in \mathbb{C}^n$, the associated *circulant matrix* $C \in \mathbb{C}^{n\times n}$ is:

$$C = \begin{bmatrix} \boldsymbol{c}^{(0)} & \boldsymbol{c}^{(n-1)} & \dots & \boldsymbol{c}^{(1)} \\ \boldsymbol{c}^{(1)} & \boldsymbol{c}^{(0)} & \dots & \boldsymbol{c}^{(2)} \\ \vdots & & \ddots & \\ \boldsymbol{c}^{(n-1)} & \boldsymbol{c}^{(n-2)} & & \boldsymbol{c}^{(0)} \end{bmatrix} = \left[\; \boldsymbol{c} \mid S\boldsymbol{c} \mid \dots \mid S^{n-1}\boldsymbol{c} \;\right]. \tag{2}$$

(2) is completely defined by $\boldsymbol{c}$. A matrix C is *circulant* iff it commutes with the circular shift operator S: $CS = SC$.

Proposition 2.3: (Simultaneous diagonalization of circulant matrices, [13]) Given any circulant matrix C, $\Lambda_{\boldsymbol{c}} := FCF^*$ is diagonal. We denote the diagonal entries of $\Lambda_{\boldsymbol{c}}$ by $\boldsymbol{\lambda_c}$.

Corollary 2.4 (Eigenvalues of a circulant matrix, [13]): Let C be the circulant matrix generated by $\boldsymbol{c} \in \mathbb{R}^n$. Then, the eigenvalues $\boldsymbol{\lambda_c}$ of C are the entries of the DFT $\hat{\boldsymbol{c}}$ of $\boldsymbol{c}$, scaled by a factor $\sqrt{n}$: $\boldsymbol{\lambda_c}(j) = \sum_{l=0}^{n-1} \boldsymbol{c}^{(l)}\, e^{-i\frac{2\pi}{n}jl}$, $j = 0, \dots, n-1$. By Proposition 2.1, $\boldsymbol{\lambda_c} = \sqrt{n}\,\hat{\boldsymbol{c}}$ is conjugate symmetric.

III. Spatially invariant systems

We consider LTI dynamics in continuous time $t > 0$

$$\dot{\boldsymbol{x}}(t) = A\boldsymbol{x}(t) + B\boldsymbol{u}(t) + \boldsymbol{w}(t), \tag{3}$$

for $\boldsymbol{x}(0) \in \mathbb{R}^n$ given. The following assumptions hold:

Assumption 1: The state $\boldsymbol{x}(t)$, control input $\boldsymbol{u}(t)$, and noise $\boldsymbol{w}(t)$ are $\mathbb{R}^n$-valued. A and B are circulant[1] elements of $\mathbb{R}^{n\times n}$. The pair (A, B) is controllable.

Assumption 2: Process noise $\boldsymbol{w}(\cdot)$ is Gaussian white noise: $\mathbb{E}[\boldsymbol{w}(t)] = \boldsymbol{0}$ for all t and $\mathbb{E}[\boldsymbol{w}(t)\boldsymbol{w}(t-\tau)^\top] = \Sigma_{\boldsymbol{w}}\delta(\tau)$, where $\delta(\cdot)$ denotes the Dirac delta distribution; $\Sigma_{\boldsymbol{w}}$ is a positive definite circulant covariance matrix.

Under assumptions 1-2, we refer to the dynamics (3) as *spatially invariant* or *circulant*.

Assumption 3: For simplicity of exposition, we analyze the setting in which full state measurements are available for control, i.e., $\boldsymbol{y}(t) = \boldsymbol{x}(t), \forall t$.

A. Diagonalization: modal dynamics

Circulant systems (3) are *diagonalized* by the DFT. Using $\hat{\boldsymbol{x}} := F\boldsymbol{x}$, (3) is equivalently written as $\dot{\hat{\boldsymbol{x}}}(t) = \Lambda_{\boldsymbol{a}}\hat{\boldsymbol{x}}(t) + \Lambda_{\boldsymbol{b}}\hat{\boldsymbol{u}}(t) + \hat{\boldsymbol{w}}(t)$, which is *decoupled* in n (scalar) subsystems

$$\hat{\boldsymbol{x}}^{(j)}(t) = \boldsymbol{\lambda}_{\boldsymbol{a}}^{(j)}\hat{\boldsymbol{x}}^{(j)}(t) + \boldsymbol{\lambda}_{\boldsymbol{b}}^{(j)}\hat{\boldsymbol{u}}^{(j)}(t) + \hat{\boldsymbol{w}}^{(j)}(t), \tag{4}$$

indexed by $j = 0, \dots, n-1$. We refer to (4) as the *modal dynamics*. Next, we derive some properties of the Fourier-transformed (complex) noise process $\hat{\boldsymbol{w}}$ which will be useful later on to prove the decoupling of the LEQG control problem in our spatially invariant setting.

Proposition 3.1 (Properties of $\hat{\boldsymbol{w}}$): Let $\hat{\boldsymbol{w}}(t) := F\,\boldsymbol{w}(t), \forall t$, where $\boldsymbol{w}(\cdot)$ satisfies assumptions 1 and 2. Then, $\hat{\boldsymbol{w}}(\cdot)$ is a complex Gaussian white noise process such that the first $\lfloor \frac{n}{2} \rfloor + 1$ entries of $\hat{\boldsymbol{w}}(t)$ are independent complex Gaussian random variables, $\forall t$. The real and imaginary parts of each such complex entries are independent as well.

Proof. See Appendix A.

IV. The Optimal LEQG Control Problem

We briefly review the Linear Exponential Quadratic Gaussian (LEQG) control synthesis problem formulation following the seminal works of Jacobson [7], [8] and Whittle [9]–[11]. We focus on the regulation problem. Define

$$J_T := \int_0^T \big(\boldsymbol{x}(t)^\top Q\boldsymbol{x}(t) + \boldsymbol{u}(t)^\top R\boldsymbol{u}(t)\big)\mathrm{d}t, \tag{5}$$

[1]Our main results in Section V generalize to the setting in which A and B are block-circulant by replacing F by the block-DFT in the proofs.

where $\boldsymbol{x}(\cdot)$ is given by (3) and the following holds:

Assumption 4: $Q \in \mathbb{R}^{n\times n}$ is circulant and $Q = Q^\top \succ 0$; $R \in \mathbb{R}^{n\times n}$ is circulant and $R = R^\top \succ 0$.

J_T in (5) is a random variable. The traditional LQG synthesizes $\boldsymbol{u}$ to minimize J_T's time-averaged expected value: $\boldsymbol{u}_{\text{LQG}}(\cdot) := \operatorname{argmin} \lim_{T\to\infty} \frac{1}{T}\mathbb{E}[J_T]$, a risk-neutral objective. Instead, we consider the control synthesis problem [10]:

$$\begin{aligned} \boldsymbol{u}_\theta(\cdot) := \operatorname{argmin} \lim_{T\to\infty} \frac{1}{\theta T} \ln \mathbb{E}\big[e^{\theta J_T}\big], \\ \text{s.t.} \quad \dot{\boldsymbol{x}}(t) = A\boldsymbol{x}(t) + B\boldsymbol{u}(t) + \boldsymbol{w}(t),\ \boldsymbol{x}(0) = \boldsymbol{x}_0 \end{aligned} \tag{6}$$

where J_T is as defined in (5) and $\theta \in \mathbb{R}$ is a free parameter. (6) is the infinite-horizon LEQG problem. The optimization objective in (6) accounts for moments of J_T beyond the expectation.

a) Risk attitude of the controller, θ: θ models the *risk attitude* of the controller. When $\theta \to 0$ the objective in (6) approximates that of the LQG, yielding a *risk-neutral* controller. When $\theta > 0$, the controller is *risk-averse*; it *assumes* that the noise $\boldsymbol{w}(\cdot)$ behaves as a non-cooperative "phantom" agent. When $\theta < 0$, the controller is *risk-seeking* and *assumes* that $\boldsymbol{w}(\cdot)$ behaves as a cooperative player in the cost minimization [7]. Thus, a risk-seeking (risk-averse) controller is *optimistic* (*pessimistic*) as it implicitly assumes that uncertainties will turn out to its advantage (disadvantage). Problem (6) does not have a stabilizing solution $\forall\theta \in \mathbb{R}$. Critical values $\bar{\theta}$ and $\underline{\theta}$ of extreme pessimism and optimism exist. At $\underline{\theta}$ the optimizer is too lax for stability; at $\bar{\theta}$, so pessimistic that it has the conviction of inability to control [10]. The equivalence between the LEQG at $\bar{\theta}$ and the $\mathcal{H}_\infty$ problem was established in [12].

b) Solution to the infinite horizon LEQG: The optimal risk-sensitive control input that solves the LEQG problem (6) is a linear function of the state $\boldsymbol{u}_\theta(t) = -K_\infty \boldsymbol{x}(t)$ with [7]

$$K_\infty = R^{-1}B^\top P_\infty. \tag{7}$$

P_∞ solves the Generalized Algebraic Riccati Equation (GARE)

$$Q + P_\infty A + A^\top P_\infty - P_\infty(BR^{-1}B^\top - \theta\,\Sigma_{\boldsymbol{w}})P_\infty = 0. \tag{8}$$

The following assumption guarantees the well-posedness of the LEQG problem (6) for the system (3):

Assumption 5: $BR^{-1}B^\top - |\theta|\,\Sigma_{\boldsymbol{w}} \succ 0$. We denote its principal matrix square root by B_θ. Such positive definiteness imposes lower $\underline{\theta}$ and upper $\bar{\theta}$ bounds on θ; in our circulant problem set-up, these are

$$\underline{\theta} := \max_j -\frac{|\boldsymbol{\lambda}_{\boldsymbol{b}}^{(j)}|^2}{\boldsymbol{\lambda}_{\boldsymbol{r}}^{(j)}\boldsymbol{\lambda}_{\boldsymbol{\sigma}}^{(j)}},\ \bar{\theta} := \min_j \frac{|\boldsymbol{\lambda}_{\boldsymbol{b}}^{(j)}|^2}{\boldsymbol{\lambda}_{\boldsymbol{r}}^{(j)}\boldsymbol{\lambda}_{\boldsymbol{\sigma}}^{(j)}}, \tag{9}$$

with $j = 0, \dots, n-1$. We work with $\underline{\theta} < \theta < \bar{\theta}$.

When $\theta \le 0$, [16, Lemma 4] together with assumptions 1 and 4 guarantee that the GARE (8) has a unique solution satisfying $P_\infty \in \mathbb{R}^{n\times n}$ and $P_\infty = P_\infty^\top \succ 0$, such that $A - (BR^{-1}B^\top - \theta\Sigma_{\boldsymbol{w}})P_\infty$ is Hurwitz. When $\theta > 0$, such a unique solution is guaranteed by assumption 5, which provides controllability of the pair (A, B_θ). In both settings, the stability of the optimal closed-loop is checked by defining the Lyapunov function $V(\boldsymbol{x}) := \boldsymbol{x}^\top P_\infty \boldsymbol{x}$, which under assumption 5 satisfies: $\dot{V}(\boldsymbol{x}) = \boldsymbol{x}^\top\big(P_\infty(A - BR^{-1}B^\top P_\infty) + (A - BR^{-1}B^\top P_\infty)^\top P_\infty\big)\boldsymbol{x} \overset{(8)}{=} \boldsymbol{x}^\top\big(-Q - P_\infty(BR^{-1}B^\top + \theta\Sigma_{\boldsymbol{w}})P_\infty\big)\boldsymbol{x} < 0,\ \forall\,\boldsymbol{x} \neq \boldsymbol{0}$. The optimal infinite horizon cost of the LEQG problem (6) is [12]

$$\min \lim_{T\to\infty} \frac{1}{T\theta} \ln \mathbb{E}\big[e^{\theta J_T}\big] = \operatorname{tr}\big(\Sigma_{\boldsymbol{w}} P_\infty\big). \tag{10}$$

V. Structural Properties of Optimal LEQG Controllers for Spatially Invariant Systems

We derive structural properties of optimal LEQG controllers for SIS. First, we prove that *the optimal LEQG control gain is spatially invariant* and provide explicit expressions for its eigenvalues. Second, we prove that the LEQG problem (6) for the circulant system (3) is *equivalent to a family of* $\lfloor \frac{n}{2} \rfloor + 1$ *decoupled scalar LEQG problems*. Finally, under some further assumptions we show that the optimal LEQG control gain is *spatially localized*; through an example, we illustrate how the degree of spatial localization is dependent on the risk-attitude (θ) of the system.

Theorem 5.1 (Spatial invariance of P_∞): Let assumptions 1 to 5 hold. Then, the stabilizing solution P_∞ of the GARE (8) is circulant and has eigenvalues $(j = 0, \dots, n-1)$:

$$\boldsymbol{\lambda}_{\boldsymbol{p}}^{(j)} = \frac{\Re(\boldsymbol{\lambda}_{\boldsymbol{a}}^{(j)}) + \sqrt{\Re(\boldsymbol{\lambda}_{\boldsymbol{a}}^{(j)})^2 + \boldsymbol{\lambda}_{\boldsymbol{q}}^{(j)}\Big(\frac{|\boldsymbol{\lambda}_{\boldsymbol{b}}^{(j)}|^2}{\boldsymbol{\lambda}_{\boldsymbol{r}}^{(j)}} - \theta\boldsymbol{\lambda}_{\boldsymbol{\sigma}}^{(j)}\Big)}}{\Big(\frac{|\boldsymbol{\lambda}_{\boldsymbol{b}}^{(j)}|^2}{\boldsymbol{\lambda}_{\boldsymbol{r}}^{(j)}} - \theta\boldsymbol{\lambda}_{\boldsymbol{\sigma}}^{(j)}\Big)}. \tag{11}$$

Proof: (i) Spatial invariance of P_∞: Left- and right-multiplying (8) by S and S^*, respectively, yields $S\big(Q + P_\infty A + A^\top P_\infty - P_\infty(BR^{-1}B^\top - \theta\Sigma_{\boldsymbol{w}})P_\infty\big)S^* = 0$. By assumptions 1, 2 and 4, A, Q, and $(BR^{-1}B^\top - \theta\Sigma_{\boldsymbol{w}})$ are circulant and thus, commute with S and S^*. Then,

$$\begin{aligned} Q + (SP_\infty S^*)A + A^\top(SP_\infty S^*) \\ - (SP_\infty S^*)(BR^{-1}B^\top - \theta\Sigma_{\boldsymbol{w}})(SP_\infty S^*) = 0. \end{aligned} \tag{12}$$

(12) shows that $SP_\infty S^*$ is a solution to the GARE (8). Furthermore, $SP_\infty S^* \succ 0$ because $P_\infty \succ 0$. Since assumption 5 holds, P_∞ is the unique (in the sense of Section IV) solution to (8). Then $P_\infty = SP_\infty S^*$ and thus, P_∞ is circulant.
(ii) Eigenvalues of P_∞: Using the DFT we write (8) as

$$\Lambda_{\boldsymbol{q}} + \Lambda_{\boldsymbol{p}}\Lambda_{\boldsymbol{a}} + \overline{\Lambda_{\boldsymbol{a}}}\Lambda_{\boldsymbol{p}} - \Lambda_{\boldsymbol{p}}\big(\Lambda_{\boldsymbol{b}}\Lambda_{\boldsymbol{r}}^{-1}\overline{\Lambda_{\boldsymbol{b}}} - \theta\Lambda_{\boldsymbol{\sigma}}\big)\Lambda_{\boldsymbol{p}} = 0, \tag{13}$$

where we used $\Lambda_{\boldsymbol{a}^\top} = \overline{\Lambda_{\boldsymbol{a}}}$ and $\Lambda_{(\boldsymbol{r}^{-1})} = \Lambda_{\boldsymbol{r}}^{-1}$. (13) is a system of n *decoupled* scalar GAREs; they are written in terms of eigenvalues $(j = 0, \dots, n-1)$ as:

$$\boldsymbol{\lambda}_{\boldsymbol{q}}^{(j)} + 2\,\boldsymbol{\lambda}_{\boldsymbol{p}}^{(j)}\Re(\boldsymbol{\lambda}_{\boldsymbol{a}}^{(j)}) - (\boldsymbol{\lambda}_{\boldsymbol{p}}^{(j)})^2\left(\frac{|\boldsymbol{\lambda}_{\boldsymbol{b}}^{(j)}|^2}{\boldsymbol{\lambda}_{\boldsymbol{r}}^{(j)}} - \theta\boldsymbol{\lambda}_{\boldsymbol{\sigma}}^{(j)}\right) = 0.$$

Solving for the positive solution yields (11). ∎

Remark 5.2: By Proposition 2.1, it suffices to solve for the first $\lfloor \frac{n}{2} \rfloor + 1$ eigenvalues (11) to completely determine P_∞.

Corollary 5.3 (Spatial invariance of K_∞ & performance): Let assumptions 1 to 5 hold. Then, the optimal risk-sensitive

gain K_∞ (7) for the system (3) is circulant with eigenvalues

$$\boldsymbol{\lambda}_{\boldsymbol{k}}^{(j)} = \frac{\overline{\boldsymbol{\lambda}}_{\boldsymbol{b}}^{(j)}}{\boldsymbol{\lambda}_{\boldsymbol{r}}^{(j)}} \boldsymbol{\lambda}_{\boldsymbol{p}}^{(j)},\ j = 0, \dots, n-1 \tag{14}$$

with $\boldsymbol{\lambda}_{\boldsymbol{p}}$ as in (11); the optimal LEQG performance (10) is $\min_{\boldsymbol{u}(\cdot)} \lim_{T\to\infty} \frac{1}{T\theta} \ln \mathbb{E}[e^{\theta J_T}] = \boldsymbol{\lambda}_{\boldsymbol{\sigma}}^\top \boldsymbol{\lambda}_{\boldsymbol{p}}$.

The spatial invariance of K_∞ implies that the same control algorithm is run at each actuator location, which eases controller implementation.

A. A decoupled family of scalar LEQG control problems

By Theorem 5.1, in a $n \times n$ circulant setting the GARE (8) decouples into a family of n scalar GAREs; the first $\lfloor \frac{n}{2} \rfloor + 1$ completely determine P_∞, and thus, K_∞. This suggests that in such a setup the LEQG problem (6) might decouple as well. We show next that it is equivalent to a family of $\lfloor \frac{n}{2} \rfloor + 1$ decoupled LEQG problems with scalar decision variable.

Theorem 5.4 (LEQG problem decoupling): Let assumptions 1 to 5 hold. W.l.g., let $\boldsymbol{u}(t) = -K\boldsymbol{x}(t)$, where $K \in \mathbb{R}^{n\times n}$ is circulant. Then, the infinite horizon LEQG control synthesis problem (6) for the system (3) is

$$\begin{aligned} &\min_{K} \lim_{T\to\infty} \frac{1}{\theta T} \ln\Big(\mathbb{E}\big[e^{\theta J_T}\big]\Big) \\ \text{s.t.}\quad &\dot{\boldsymbol{x}}(t) = A\boldsymbol{x}(t) + B\boldsymbol{u}(t) + \boldsymbol{w}(t),\ \boldsymbol{x}(0) = \boldsymbol{x}_0, \\ &\boldsymbol{u} = -K\boldsymbol{x},\ K \text{ circulant}, \end{aligned} \tag{15}$$

with J_T as in (5). Problem (15) is equivalent to $\lfloor \frac{n}{2} \rfloor + 1$ *decoupled scalar* LEQG control problems,

$$\begin{aligned} &\min_{\boldsymbol{\lambda}_{\boldsymbol{k}}^{(j)} \in \mathbb{C}} \lim_{T\to\infty} \frac{1}{\theta T} \ln\Big(\mathbb{E}\big[e^{\theta G_T^{(j)}}\big]\Big), \\ \text{s.t.}\quad &\dot{\hat{\boldsymbol{x}}}^{(j)}(t) = \big(\boldsymbol{\lambda}_{\boldsymbol{a}}^{(j)} - \boldsymbol{\lambda}_{\boldsymbol{b}}^{(j)} \boldsymbol{\lambda}_{\boldsymbol{k}}^{(j)}\big)\hat{\boldsymbol{x}}^{(j)}(t) + \hat{\boldsymbol{w}}^{(j)}(t), \end{aligned} \tag{16}$$

where $j = 0, \dots, \lfloor \frac{n}{2} \rfloor$ and

$$G_T^{(j)} := \begin{cases} J_T^{(j)}, & \text{if } j = 0 \text{ or } j = \frac{n}{2}, \\ 2J_T^{(j)}, & \text{otherwise}, \end{cases} \tag{17}$$

with $J_T^{(j)} := \big(\boldsymbol{\lambda}_{\boldsymbol{q}}^{(j)} + |\boldsymbol{\lambda}_{\boldsymbol{k}}^{(j)}|^2 \boldsymbol{\lambda}_{\boldsymbol{r}}^{(j)}\big) \int_0^T |\hat{\boldsymbol{x}}^{(j)}(t)|^2 \mathrm{d}t$. We refer to the family of problems (16) as the *modal LEQG problems*.

Proof. See Appendix B.

B. Spatial localization of the optimal LEQG control gain

For SIS over *unbounded* spatial domains with distributed sensing and actuation, [1] proved that quadratically-optimal control gains exhibit fast spatial decay, *asymptotically* exponential. Informally, this implies that for each spatial site, measurements from its neighborhood are more valuable for control than those further away. System's parameters determine the exponential asymptotic decay rate, which has been analytically characterized in some settings [17], [18]. In our circulant problem set-up the spatial coordinate is *bounded*. To analyze the spatial decay of the optimal risk-sensitive gain we use the arguments in [19], based on Bernstein's theorem from approximation theory [19, Theorem 3.3]. We are interested in the setting in which matrix A in (3) is symmetric and sparse. To this matrix, we associate a graph $\mathcal{G}(A) = (V, E)$, where V denotes the vertex set consisting of integers from 0 to $n-1$ and E is the edge set containing all pairs (i, j) with $A(i, j) \neq 0$. The distance $d_A(i, j)$ between $i, j \in V$ is the length of the shortest path connecting node i to node j. Assuming that the rest of the matrices in the LEQG problem are multiples of the identity, we provide an upper bound for the spatial decay rate of K_∞.

Theorem 5.5 (Spatial decay of the optimal control gain): Consider the optimal risk-sensitive control problem (6) for the system (3), subject to assumptions 1 to 3. Furthermore, assume that $B = bI, \Sigma = \sigma I, Q = qI$ and $R = rI$ with $q, b, r, \sigma \in \mathbb{R}_+$ and that θ satisfies the bounds in (9). Let A be such that $A = A^\top$ and $\big|A(i, j)\big| < \alpha\, \beta^{d_A(i,j)}$ with $\alpha > 0$ and $0 < \beta < 1$. Assume[2] $d_A(i, j) \geq d_{A^2}(i, j), \forall i, j$. Then, there exists constants $\mu > 0$ and $0 < \eta < 1$ such that $\forall\, i, j$ the optimal risk-sensitive gain (7) satisfies

$$\big|K_\infty(i, j)\big| < \mu\, \eta^{d_{A^2}(i,j)}. \tag{18}$$

Proof: Since all matrices in the GARE (8) are circulant and hence they commute, and $A = A^\top$, the optimal risk-sensitive control gain is $K_\infty = \frac{b}{b^2 - r\sigma\theta}\big(A + \sqrt{A^2 + \frac{q(b^2 - r\sigma\theta)}{r} I}\big)$, where we use $\sqrt{\cdot}$ to refer to the principal square root. Denote $\mathcal{R} := A^2 + \frac{q(b^2 - r\sigma\theta)}{r} I$. By assumption 5, $\mathcal{R} \succ 0$. Furthermore, $\mathcal{R}$ is circulant and thus diagonalized by the DFT, which has condition number $\kappa(F) = 1$. Taking the branch cut of $f(\cdot) := \sqrt{\cdot}$ to be the non-positive real axis, f is analytic in the open right-half complex plane, which contains the spectrum of $\mathcal{R}$. Thus, by [19, Theorem 3.4] constants $\xi > 0$ and $0 < \gamma < 1$ exist such that $|f(\mathcal{R})(i, j)| < \xi \gamma^{d_{A^2}(i,j)}$. Then, $|K_\infty(i, j)| < \frac{b}{b^2 - r\sigma\theta}\big(\alpha\beta^{d_A(i,j)} + \xi\gamma^{d_{A^2}(i,j)}\big) \leq \mu'(\beta^{d_A(i,j)} + \gamma^{d_{A^2}(i,j)}) \leq \mu'(\beta^{d_{A^2}(i,j)} + \gamma^{d_{A^2}(i,j)}) \leq \mu\eta^{d_{A^2}(i,j)}$, where $\mu' := \frac{b \max\{\alpha, \xi\}}{b^2 - r\sigma\theta}$, $\eta := \max\{\beta, \gamma\}$, and $\mu := 2\mu'$. ∎

Remark 5.6: The upper bound (18) for the spatial decay of K_∞ is meaningful when A^2 is sparse. This is often the case in large-scale systems (e.g., obtained by discretization of partial differential equations [20]).

Remark 5.7: The *spatial decay rate* η of the optimal LEQG control gain depends on the strength of the noise (σ) perturbing the plant and on the risk-attitude (θ) of the system. We illustrate the sensitivity to θ through a case study.

1) A case study: consider the LEQG regulation problem of Theorem 5.5. We analyze the setting in which $A = -\texttt{toeplitz}([2\ -1\ 0\ \dots\ 0\ -1])$, the negative Laplacian of a cycle graph. We select $b = \sigma = q = r = 1$ and let θ vary in $-1 < \theta < 1$. The dynamics of the state are described by $\dot{\boldsymbol{x}}(t) = -\texttt{toeplitz}([2\ -1\ 0\ \dots\ -1])\,\boldsymbol{x}(t) + \boldsymbol{u}(t) + \boldsymbol{w}(t)$. The corresponding steady-state LEQG control gain is

$$K_\infty = \frac{1}{1-\theta}\big(A + \sqrt{A^2 + (1-\theta)I}\big). \tag{19}$$

[2]The condition $d_A(i, j) \geq d_{A^2}(i, j), \forall (i, j)$ is satisfied when A is a graph Laplacian: $A = dI - \mathsf{adj}(\mathsf{A})$, where $\mathsf{adj}(\mathsf{A})(i, j) := \begin{cases} 1, & \text{if } (i, j) \in E, \\ 0, & \text{otherwise}, \end{cases}$ is the adjacency matrix and $d > 0$ is the degree of the d-regular circulant graph $\mathcal{G}(A)$.

Since A is (circulant) banded and hence completely spatially localized, the spatial decay rate of (19) is determined by $\sqrt{A^2+(1-\theta)I}$; in particular, the decay rate depends on the spectrum of $\mathcal{R}(\theta) = A^2 + (1-\theta)I$ and the analyticity region of $f(\cdot) = \sqrt{\cdot}$. The eigenvalues of A are $\boldsymbol{\lambda_a}(j) = -2\big(1-\cos(\frac{2\pi j}{n})\big)$ (with $j = 0,\dots,n-1$). Hence, $\sigma_{\mathcal{R}}(\theta) \subset \mathcal{D}$, where $\sigma_{\mathcal{R}}(\theta)$ denotes the spectrum of $\mathcal{R}(\theta)$ and $\mathcal{D} := \{z \in \mathbb{C} : (\Re(z)-z_0)^2 + \Im(z)^2 \leq \rho\}$, with $z_0 = 9-\theta$ and $\rho = 8$. $f(\cdot) = \sqrt{\cdot}$ is analytic in the concentric open disk of radius $R_0 = 9-\theta$. Then, by [19, Theorems 3.3 and 3.4] the sensitivity of the spatial decay rate η of K_∞ to θ is $\frac{d\eta}{d\theta} = \frac{d(\rho/R_0)}{d\theta} = -\frac{8}{(9-\theta)^2}$. η monotonically decreases with θ: the more risk-averse the system is, the slower the spatial decay of K_∞. Such a monotonicity result implies that risk-seeking controllers are more amenable to distributed implementations. Fig. 1 shows the LEQG gain at different values of θ.

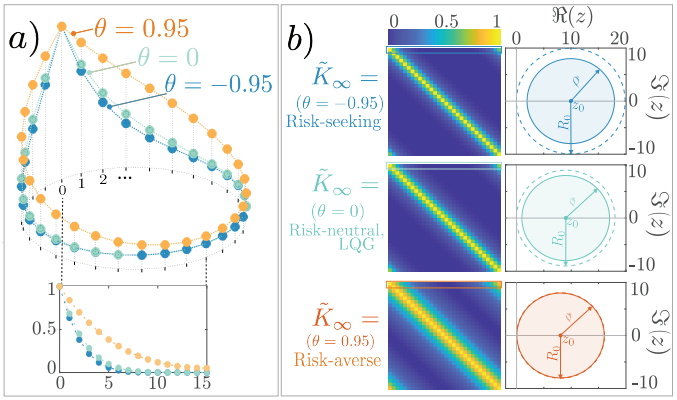

Fig. 1: Steady-state normalized optimal LEQG control gain $\tilde{K}_\infty := K_\infty/K_\infty(0,0)$ for the example in Section V-B with $n = 30$. $\tilde{K}_\infty$ is plotted for different θ, as indicated. a) $\tilde{\boldsymbol{k}}$ as a function in the discrete circle $\mathbb{Z}_{30}$. The change in the spatial decay rate of the gain at different values of θ is apparent in the lower panel. b) $\tilde{K}_\infty$ with entries colored according to their respective values, following the colorbar provided. The second column shows the closed disk $\mathcal{D}$ containing the spectrum $\sigma_{\mathcal{R}}(\theta)$ and the largest concentric open disk where $f(\cdot) := \sqrt{\cdot}$ is analytic in the complex plane.

C. Complexity reduction

The properties of LEQG controllers for circulant systems derived so far have important practical implications.

a) Solving large-scale AREs: Traditional solution methods for AREs require $\mathcal{O}(n^3)$ flops [21], and hence, are unsuitable for large-scale systems. In our circulant setting, the fact that K_∞ is circulant (Corollary 5.3) together with the efficient computation of the DFT provided by the Fast Fourier Transform (FFT) can be utilized to reduce the complexity of computing K_∞: computing the eigenvalues of the system $(A, B, \Sigma_{\boldsymbol{w}})$ and cost (Q, R) matrices has $\mathcal{O}(n\log(n))$ complexity using the FFT. Then, the eigenvalues of K_∞ can be independently computed using (14); $\boldsymbol{k}_\infty$ is obtained with $\mathcal{O}(n\log(n))$ complexity using the inverse FFT on $\boldsymbol{\lambda_k}$. These scalings hold even with a large number of control inputs.

b) Controller implementation: When the decay bound provided in Theorem 5.5 is meaningful, K_∞ exhibits fast spatial decay. The controller might be implemented in a distributed manner by dropping those entries of K_∞ that are below a certain threshold (as long as the closed-loop remains stable). Alternatively, an optimal circulant control gain with a desired sparsity pattern could be designed using the problem formulation in Theorem 5.4 and extending the arguments in [22]. These strategies reduce the complexity of controller implementation, as they only require local computation and communication between sensors and actuators, avoiding the need for a centralized processing unit.

Acknowledgements

The authors would like to thank Poorva Shukla (UCSB) for pointing out relevant literature.

References

[1] B. Bamieh, F. Paganini, and M. A. Dahleh, "Distributed Control of Spatially Invariant Systems," *IEEE Transactions on Automatic Control*, vol. 47, no. 7, pp. 1091–1107, Jul. 2002.

[2] M. R. Jovanović and B. Bamieh, "On the ill-posedness of certain vehicular platoon control problems," *IEEE Transactions on Automatic Control*, vol. 50, no. 9, pp. 1307–1321, 2005.

[3] B. J. Olson, S. W. Shaw, C. Shi, C. Pierre, and R. G. Parker, "Circulant matrices and their application to vibration analysis," *Applied Mechanics Reviews*, vol. 66, no. 4, 2014.

[4] I. Stewart and M. Parker, "Periodic dynamics of coupled cell networks II: Cyclic symmetry," *Dynamical Systems*, vol. 23, no. 1, pp. 17–41, 2008.

[5] A. Townsend, M. Stillman, and S. H. Strogatz, "Dense networks that do not synchronize and sparse ones that do," *Chaos*, vol. 30, no. 8, 2020.

[6] G. Dullerud and F. Paganini, *A Course in Robust Control Theory: A Convex Approach*. Springer, 2000.

[7] D. H. 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, 1973.

[8] J. L. Speyer, J. Deyst, and D. H. Jacobson, "Optimization of Stochastic Linear Systems with Additive Measurement and Process Noise Using Exponential Performance Criteria," *IEEE Transactions on Automatic Control*, vol. 19, no. 4, pp. 358–366, 1974.

[9] P. Whittle, "Risk-Sensitive Linear / Quadratic / Gaussian Control," *Advanced Applied Probability*, vol. 13, no. 4, pp. 764–777, 1981.

[10] ——, "RISK SENSITIVITY, A STRANGELY PERVASIVE CONCEPT," *Macroeconomic Dynamics*, vol. 6, pp. 5–18, 2002.

[11] ——, *Risk Sensitive Optimal Control*. Wiley, 1990.

[12] K. Glover, "Minimum entropy and risk-sensitive control: The continuous time case," *Proceedings of the IEEE Conference on Decision and Control*, vol. 1, pp. 388–391, 1989.

[13] B. Bamieh, "Discovering transforms: A tutorial on circulant matrices, circular convolution, and the discrete fourier transform," *arXiv:1805.05533v2*, 2020. [Online]. Available: http://arxiv.org/abs/1805.05533

[14] R. G. Gallager, *Principles of Digital Communication Systems*. Cambridge University Press, 2008.

[15] ——, "Circularly-Symmetric Gaussian random vectors," 2008. [Online]. Available: http://www.rle.mit.edu/rgallager/documents/CircSymGauss.pdf

[16] J. C. Willems, "Least Squares Stationary Optimal Control and the Algebraic Riccati Equation," *IEEE Transactions on Automatic Control*, vol. 16, no. 6, pp. 621–634, 1971.

[17] J. Arbelaiz, B. Bamieh, A. E. Hosoi, and A. Jadbabaie, "Distributed Kalman filtering for spatially-invariant diffusion processes: The effect of noise on communication requirements," *Proceedings of the IEEE Conference on Decision and Control*, pp. 622–627, dec 2020.

[18] J. Arbelaiz, E. Jensen, B. Bamieh, A. E. Hosoi, A. Jadbabaie, and L. Lessard, "Information structures of the kalman filter for the elastic wave equation," *IFAC-PapersOnLine*, vol. 55, no. 13, pp. 1–6, 2022, 9th IFAC Conference on Networked Systems NECSYS 2022.

[19] M. Benzi and N. Razouk, "Decay bounds and O(n) algorithms for approximating functions of sparse matrices," *Electronic Transactions on Numerical Analysis*, vol. 28, pp. 16–39, 2007.

[20] A. Haber and M. Verhaegen, "Sparsity preserving optimal control of discretized PDE systems," *Computer Methods in Applied Mechanics and Engineering*, vol. 335, pp. 610–630, 2018.

[21] P. Benner, J. R. Li, and T. Penzl, "Numerical solution of large-scale Lyapunov equations, Riccati equations, and linear-quadratic optimal control problems," *Numerical Linear Algebra with Applications*, vol. 15, no. 9, pp. 755–777, 2008.

[22] J. Arbelaiz, B. Bamieh, A. E. Hosoi, and A. Jadbabaie, "Optimal structured controllers for spatially invariant systems: a convex reformulation," *Proceedings of the IEEE Conference on Decision and Control*, pp. 3374–3380, dec 2021.

APPENDIX

A. Proof of Proposition 3.1

By assumption 1, $\boldsymbol{w}(t)$ is $\mathbb{R}^n$-valued and $F \in \mathbb{C}^{n\times n}$, then $\hat{\boldsymbol{w}}(t)$ is $\mathbb{C}^n$-valued; write it as $\hat{\boldsymbol{w}}(t) = \Re(F)\boldsymbol{w}(t) + i\,\Im(F)\boldsymbol{w}(t)$ and define the $\mathbb{R}^{2n}$-valued random vector[3] $\hat{\boldsymbol{\omega}}(t) \quad \hat{\boldsymbol{\omega}}(t) := \begin{bmatrix}\Re[F] & \Im[F]\end{bmatrix}^\top \boldsymbol{w}(t)$. Since by assumption 2 $\boldsymbol{w}(t)$ is jointly Gaussian and linear transformations preserve Gaussian distributions [15], $\hat{\boldsymbol{\omega}}(t)$ is jointly Gaussian (and $\hat{\boldsymbol{w}}(t)$ is complex jointly Gaussian). $\boldsymbol{\mu}_{\hat{\boldsymbol{\omega}}} = \mathbb{E}[\hat{\boldsymbol{\omega}}(t)] = \mathbf{0}\ \forall t$, and

$$\mathbb{E}\big[\hat{\boldsymbol{\omega}}(t)\hat{\boldsymbol{\omega}}(t-\tau)^\top\big] = \underbrace{\left[\begin{array}{c|c} \Re[F]\,\Sigma_{\boldsymbol{w}}\,\Re[F] & \Re[F]\,\Sigma_{\boldsymbol{w}}\,\Im[F] \\ \hline \Im[F]\,\Sigma_{\boldsymbol{w}}\,\Re[F] & \Im[F]\,\Sigma_{\boldsymbol{w}}\,\Im[F] \end{array}\right]}_{=:\Sigma_{\hat{\boldsymbol{\omega}}}} \delta(\tau),$$

where we have used the symmetry of F. $\hat{\boldsymbol{\omega}}(\cdot)$ is a white Gaussian noise process. Next, we evaluate the (y,k)-th entry of each block of $\Sigma_{\hat{\boldsymbol{\omega}}}$. First, consider the *odd* n case. Let $\boldsymbol{\sigma} \in \mathbb{R}^n$ denote the first column of $\Sigma_{\boldsymbol{w}}$. Since $\Sigma_{\boldsymbol{w}} = \Sigma_{\boldsymbol{w}}^\top$ and circulant, $\boldsymbol{\sigma}^{(k)} = \boldsymbol{\sigma}^{(n-k)} = \boldsymbol{\sigma}^{(-k)}$ $(k = 0, 1, \dots, n-1)$, that is, $\boldsymbol{\sigma} = \begin{bmatrix}\boldsymbol{\sigma}^{(0)} & \boldsymbol{\sigma}^{(1)} & \boldsymbol{\sigma}^{(2)} & \dots & \boldsymbol{\sigma}^{(2)} & \boldsymbol{\sigma}^{(1)}\end{bmatrix}^\top \in \mathbb{R}^n$. The entries of the respective blocks of $\Sigma_{\hat{\boldsymbol{\omega}}}$ are given in (20). Since $\hat{\boldsymbol{\omega}}(t)$ is Gaussian, the sparsity pattern of the covariance matrix $\Sigma_{\hat{\boldsymbol{\omega}}}$ provided in (20) proves the claimed independence relations among its entries. When n is *even*, $\boldsymbol{\sigma} = \begin{bmatrix}\boldsymbol{\sigma}^{(0)} & \boldsymbol{\sigma}^{(1)} & \boldsymbol{\sigma}^{(2)} & \dots & \boldsymbol{\sigma}^{(\frac{n}{2})} & \dots & \boldsymbol{\sigma}^{(2)} & \boldsymbol{\sigma}^{(1)}\end{bmatrix}^\top \in \mathbb{R}^n$. (20) must be augmented with additional terms corresponding to the $\frac{n}{2}$-th entry of $\boldsymbol{\sigma}$, but the sparsity pattern of $\Sigma_{\hat{\boldsymbol{\omega}}}$ is preserved. ■

[3]By Proposition 2.1 some entries of $\hat{\boldsymbol{\omega}}(t)$ are identically zero, namely: the entry with index n; and the entry with index $3n/2$ when n is even. These correspond to the imaginary parts of entries of $\hat{\boldsymbol{w}}$ that are real. Such identically zero entries do not affect the off-diagonal sparsity pattern of $\Sigma_{\hat{\boldsymbol{\omega}}}$, and hence, they have not been removed.

B. Proof of Theorem 5.4

We note that

$$\begin{aligned}
\mathbb{E}[e^{\theta J_T}] &= \mathbb{E}\Big[e^{\theta\int_0^T \left(\boldsymbol{x}(t)^\top Q\boldsymbol{x}(t) + \boldsymbol{u}(t)^\top R\boldsymbol{u}(t)\right)\mathrm{d}t}\Big] \\
&= \mathbb{E}\Big[e^{\theta\int_0^T \left(\hat{\boldsymbol{x}}(t)^* \Lambda_{\boldsymbol{q}} \hat{\boldsymbol{x}}(t) + \hat{\boldsymbol{u}}(t)^* \Lambda_{\boldsymbol{r}} \hat{\boldsymbol{u}}(t)\right)\mathrm{d}t}\Big] \\
&\overset{(a)}{=} \mathbb{E}\Big[e^{\theta\sum_{j=0}^{n-1}\left(\boldsymbol{\lambda}_{\boldsymbol{q}}^{(j)} + |\boldsymbol{\lambda}_{\boldsymbol{k}}^{(j)}|^2 \boldsymbol{\lambda}_{\boldsymbol{r}}^{(j)}\right)\int_0^T |\hat{\boldsymbol{x}}^{(j)}(t)|^2 \mathrm{d}t}\Big] \\
&\overset{(b)}{=} \begin{cases} \mathbb{E}\Big[e^{\theta J_T^{(0)}} \cdot e^{2\theta\sum_{j=1}^{\frac{n-1}{2}} J_T^{(j)}}\Big], & \text{if } n \text{ odd}, \\ \mathbb{E}\Big[e^{\theta J_T^{(0)}} \cdot e^{\theta J_T^{(\frac{n}{2})}} \cdot e^{2\theta\sum_{j=1}^{\frac{n}{2}-1} J_T^{(j)}}\Big], & \text{if } n \text{ even}, \end{cases} \\
&\overset{(c)}{=} \prod_{j=0}^{\lfloor \frac{n}{2} \rfloor} \mathbb{E}\Big[e^{\theta G_T^{(j)}}\Big], \qquad (21)
\end{aligned}$$

where $\overset{(a)}{=}$ follows using $\boldsymbol{u}(t) = -K\boldsymbol{x}(t)$, with K circulant; $\overset{(b)}{=}$ follows by defining $J_T^{(j)}$ as in Theorem 5.4 and using the conjugate symmetry of the DFT of a real vector (see Proposition 2.1); $\overset{(c)}{=}$ follows by assumption 2 and Proposition 3.1, together with the observation that $J_T^{(j)}$ $(j = 0, \dots, \lfloor \frac{n}{2} \rfloor)$ only depends on the corresponding stochastic process $\hat{\boldsymbol{w}}^{(j)}$; using $G_T^{(j)}$ as defined in (17), we write (21). Substitution of (21) in the objective of problem (6) yields $\lim_{T\to\infty} \frac{1}{\theta T} \ln\Big(\mathbb{E}\big[e^{\theta J_T}\big]\Big) = \lim_{T\to\infty} \frac{1}{\theta T} \sum_{j=0}^{\lfloor \frac{n}{2} \rfloor} \ln\Big(\mathbb{E}\big[e^{\theta G_T^{(j)}}\big]\Big)$. Together with the modal dynamics (4), the family of decoupled *modal optimization problems* (16) is obtained. ■

$$\begin{aligned}
\Re[F]\,\Sigma_{\boldsymbol{w}}\,\Re[F](y,k) &= \frac{1}{n}\sum_{z=0}^{n-1} \cos\left(\frac{2\pi kz}{n}\right) \sum_{j=0}^{n-1} \cos\left(\frac{2\pi yj}{n}\right) \boldsymbol{\sigma}^{(j-z)} \overset{(a)}{=} \frac{1}{n}\sum_{p=0}^{n-1}\sum_{z=0}^{n-1} \cos\left(\frac{2\pi kz}{n}\right) \cos\left(\frac{2\pi y(p+z)}{n}\right) \boldsymbol{\sigma}^{(p)} \\
&\overset{(b)}{=} \frac{1}{n}\sum_{z=0}^{n-1} \cos\left(\frac{2\pi kz}{n}\right) \cos\left(\frac{2\pi yz}{n}\right) \boldsymbol{\sigma}^{(0)} + \frac{1}{n}\sum_{p=1}^{\frac{n-1}{2}}\sum_{z=0}^{n-1} \cos\left(\frac{2\pi kz}{n}\right) \left[\cos\left(\frac{2\pi y(p+z)}{n}\right) + \cos\left(\frac{2\pi y(z-p)}{n}\right)\right] \boldsymbol{\sigma}^{(p)} \\
&= \frac{1}{n}\sum_{z=0}^{n-1} \cos\left(\frac{2\pi kz}{n}\right) \cos\left(\frac{2\pi yz}{n}\right) \boldsymbol{\sigma}^{(0)} + \frac{2}{n}\sum_{p=1}^{\frac{n-1}{2}} \cos\left(\frac{2\pi yp}{n}\right) \sum_{z=0}^{n-1} \cos\left(\frac{2\pi kz}{n}\right) \cos\left(\frac{2\pi yz}{n}\right) \boldsymbol{\sigma}^{(p)} \\
&\overset{(c)}{=} \big(\chi(k-y) + \chi(k+y)\big)\left[\frac{1}{2}\boldsymbol{\sigma}^{(0)} + \sum_{p=1}^{\frac{n-1}{2}} \cos\left(\frac{2\pi yp}{n}\right) \boldsymbol{\sigma}^{(p)}\right]. \qquad (20)
\end{aligned}$$

$\overset{(a)}{=}$ follows defining $p := j - z$, noticing that for a given z: $\sum_{p=-z}^{n-1-z} \cos\big(\frac{2\pi y(p+z)}{n}\big)\boldsymbol{\sigma}^{(p)} = \sum_{p=0}^{n-1} \cos\big(\frac{2\pi y(p+z)}{n}\big)\boldsymbol{\sigma}^{(p)}$, and rearranging the sums; $\overset{(b)}{=}$ follows using the symmetry of $\boldsymbol{\sigma}$ and the fact that n is odd; and $\overset{(c)}{=}$ follows using $\sum_{j=0}^{n-1} \cos\big(\frac{2\pi jk}{n}\big) \cos\big(\frac{2\pi jy}{n}\big) = \frac{n}{2}\big(\chi(k-y) + \chi(k+y)\big)$, where $\chi(k) := \begin{cases} 1, & \text{if } k \equiv_n 0 \\ 0, & \text{if } k \not\equiv_n 0 \end{cases}$. Similarly, we obtain that $\Im[F]\,\Sigma_{\boldsymbol{w}}\,\Im[F](y,k) = \big(\chi(k-y) - \chi(k+y)\big)\left[\frac{1}{2}\boldsymbol{\sigma}^{(0)} + \sum_{p=1}^{\frac{n-1}{2}} \cos\left(\frac{2\pi yp}{n}\right)\boldsymbol{\sigma}^{(p)}\right]$ and $\Im[F]\,\Sigma_{\boldsymbol{w}}\,\Re[F](y,k) = 0,\ \forall\, y, k$.