Stochastic Control - A Moment Approach to Sparse Control Design

Maryam Kamgarpour and Tyler Summers

***Abstract*— We consider design of sparse controllers for a stochastic linear system with infinite horizon quadratic objective. We formulate the non-sparse optimal solution through a semidefinite program for the second order moments of the states and inputs. Given that the centralized non-sparse controller solves a linear equation in these moments, we find sparse least squares approximate solutions to this linear equation. The performance of the approach is shown with several simulations.**

I. Background

A fundamental problem in multiagent systems is decentralized optimal control. Witsenhausen in his seminal paper showed that in a stochastic linear system with quadratic cost, when the information structure is not classic, as would be the case in decentralized control, the optimal controller need not be linear [1]. Papadimitriou and Tsitsiklis showed that the discrete version of the decentralized stochastic control is NP-complete and as a result the continuous version of the problem is also in general intractable [2]. Motivated by these results, several works in the past decades have been devoted to answer the following two problems. Given an information structure 1) under which conditions the optimal controller satisfying the given structure is linear? 2) under which conditions solving for an optimal structured controller is a convex problem?

In terms of the first problem, Ho and Chu derived sufficient conditions for optimality of linear controllers in a quadratic cost setting based on the notion of *nested information* structure [3]. Motee, Jadbabaie and Bamieh showed that if the state-space operator in the linear dynamics and the quadratic weighing matrix in the cost belong to an operator algebra, the optimal linear controller also belongs to this operator algebra and thus has the same information structure [4]. The answer to the second problem is completely characterized by the notion of *quadratic invariance* as discussed by Rotkowitz and Lall [5]. In particular, minimum norm structured linear optimal controller design is convex if and only if the structure and dynamics satisfy an algebraic condition referred to as quadratic invariance. Furthermore, the link between nested information structure and quadratic invariance is established [6].

Apart from the cases of nested information structure or quadratic invariance, solving for minimum-norm structured or sparse linear controller can be cast as a non-convex optimization problem. The work by Lin, Fardad and Jovanovic [7] addresses the $\mathcal{H}_2$ norm minimization of linear dynamical system given sparsity or structural constraint. The approach is to use the established semidefinite program (SDP) formulation for the optimal $\mathcal{H}_2$ controller, with an additional non-convex structural constraint. An approximation scheme is developed leveraging the alternating direction method of multipliers (ADMM) to find locally optimal structured/sparse linear feedback gains.

Our work is inspired by [7] in approximating the solution to the non-convex problem of optimal structured/sparse linear controllers. In particular, rather than attempting to find the optimal controller, which may have dynamic structure, we restrict our attention to static state feedback controller with sparsity objective on the gain matrix. Our formulation is an infinite horizon discrete-time linear quadratic Gaussian (LQG) control. Our approach is to characterize the solution of the standard centralized problem through an SDP formulation for the infinite horizon moments of the states and inputs. The optimal centralized controller solves a linear equation given by these moments. Thus, we include structural/sparsity objectives by solving a modified least squares problem. The SDP we consider for the second order moments of the states and inputs is the dual of that for the value function considered in [7]. By considering this dual approach, we hope to achieve a more complete picture of the (sub)optimality of these sparse controllers. The performance is analyzed with a number of case studies. Future work is discussed.

II. Problem Formulation

We consider the following stochastic linear system:

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

where $x \in \mathbb{R}^n$, $u \in \mathbb{R}^m$, $\omega \in \mathbb{R}^n$ are the state, input and the stochastic noise, respectively. We assume ω_t, $t = 0, 1, \ldots,$ are independent identically distributed

Maryam Kamgarpour is with the Automatic Control Laboratory at ETH Zürich, Switzerland. Email: mkamgar@control.ee.ethz.ch, Tyler Summers is with the Mechanical Engineering Department at the University of Texas at Dallas, USA. Email: tyler.summers@utdallas.edu.
The authors thank European Union Starting Grant for funding.

Gaussian random variables and for all t, $E\{\omega_t\} = 0$ and $E\{\omega_t \omega_t^T\} = W$. The initial state is independent of the stochastic noise. We consider control policies of the form $\pi_t : X \to U$. Thus, control at time t is $u_t = \pi_t(x_t)$. We assume the pair (A, B) is controllable

The infinite horizon discrete-time linear quadratic Gaussian problem with average cost is given as:

$$\min_{\pi_0,\pi_1,\cdots\in\Pi} J := \lim_{T\to\infty} \frac{1}{T} E \sum_{t=0}^{T} (x_t^T Q x_t + u_t^T R u_t). \tag{D-LQG}$$

We assume $Q \in \mathbb{R}^{n\times n}$ is positive semidefinite, $R \in \mathbb{R}^{m\times m}$ is positive definite, and the pair (A, C) is observable, where $Q = C^T C$. Our objective is to find sparse controllers K that achieve minimal cost.

The optimal centralized solution to the above problem is static linear state feedback: $\pi_t^*(x_t) = Kx_t$. Furthermore, a semidefinite program (SDP) can be formulated to find the optimal linear gain K as follows:

$$\begin{aligned} \min_{Z_{xx},Z_{xu},Z_{uu}} \quad & tr(QZ_{xx}) + tr(RZ_{uu}) \qquad \text{(S-LQG)} \\ \text{s.t.} \quad & -Z_{xx} + W + AZ_{xx}A^T + AZ_{xu}B^T \\ & + BZ_{xu}^T A^T + BZ_{uu}B^T = 0_{n\times n}, \\ & \begin{bmatrix} Z_{xx} & Z_{xu} \\ Z_{xu}^T & Z_{uu} \end{bmatrix} \succeq 0. \end{aligned}$$

In the above, $tr(M)$ denotes trace of the matrix M. The matrices $Z_{xx} \in \mathbb{R}^{n\times n}, Z_{xu} \in \mathbb{R}^{n\times m}, Z_{uu} \in \mathbb{R}^{m\times m}$ denote the infinite horizon second order moments of the states and inputs. For example, $Z_{xu} = \lim_{T\to\infty} \frac{1}{T} E \sum_{t=0}^{T} (x_t u_t^T)$. The optimal linear policy is then found as $K = Z_{xu}^T (Z_{xx})^{-1}$. Note that problem (S-LQG) is the dual of the semidefinite program which solves for the matrix P associated to the Riccati equation and the optimal value function of the control problem [8]. The optimal policy can be equivalently computed as $K = (R + B^T PB)^{-1} B^T PA$. The equivalence of the moment based and Riccati based optimal policy follow from the strong duality of the SDP (S-LQG).

III. Sparse Control Solution Approach

As discussed above, the controller K is a nonlinear function of the optimization variables Z_{xx}, Z_{xu}. Consequently, including general linear constraints on K introduces a nonlinear constraint on the moments. Our approach is to first solve for the optimal centralized Z_{xx}, Z_{xu}, Z_{uu} in (S-LQG) and then to find sparse approximations of K by solving a least squares problem. In order to impose sparsity, we penalize the l_1 norm of the gain matrix [7] through a scalar $\gamma \in \mathbb{R}_+$. Thus, the optimization problem is as follows:

$$\min_{K} \quad \|KZ_{xx} - Z_{xu}^T\|_2^2 + \gamma\|K\|_1. \tag{L-LQG}$$

The optimization problem above can be solved with standard SDP solvers. If the required sparsity structure $\mathcal{S}$ is given, we set $\gamma = 0$ in the above and instead we include the constraint $K \in \mathcal{S}$. Since $\mathcal{S}$ is a linear subspace, this results in an additional linear constraint. By varying γ we hope to achieve a tradeoff between sparsity and performance of the controller K.

Let $K(\gamma)$ denote the optimal solution of this least squares problem for a given γ. In general, it is hard to connect the performance of $K(\gamma)$ with that of the centralized solution K. In particular, the closed loop system $A+BK(\gamma)$ may not even be stable. A sufficient condition for stability is based on the Gerschgorin circle theorem [9]. This theorem states that all eigenvalues of a matrix $M \in \mathbb{R}^{n\times n}$ lie in at least one of the discs centered at a diagonal entry with radius given by sum of absolute values of the corresponding non-diagonal entries. Let M_{ij} denote the ij-th entry of a matrix $M \in \mathbb{R}^{n\times n}$. This stability constraint is desirable since it results in addition of a linear constraint to (L-LQG) as follows:

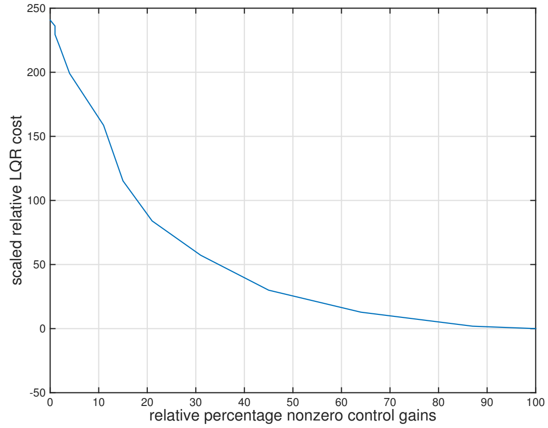
$$|(A + BK)_{ii} + \sum_{j\neq i} |(A + BK)ij| \,| < 1. \tag{2}$$

IV. Case Studies

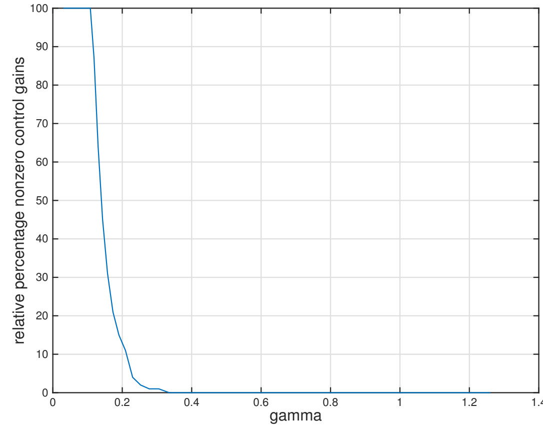
We run several simulations to study the performance of the proposed method. In all simulations the entries of the dynamics $A \in \mathbb{R}^{n\times n}$ were drawn from a normal distribution with variance 0.3. The input matrix was $B = I_{n\times n}$, that is, each individual coordinate has a control. The noise covariance was $W = I_{n\times n}$. The cost function was defined by $Q = I_{n\times n}$, $R = I_{n\times n}$. We let $n = 10$ and vary γ between $10^{-1.5} = 0.0316$ and $10^{\cdot 1} = 1.26$ in logarithmically spaced points. We illustrate three case studies corresponding to stable and unstable dynamics.

In the first two cases, matrix $A = A_1$ is stable. The closed loop system without any sparsity objectives has a full K matrix and a cost of $J = 14.8$. The results of the optimization problem (L-LQG) are shown in Figures 1a and 1b. Figure 1a on the left shows the percentage of LQG cost above its nominal centralized value as a function of the sparsity. Fig. 1b on the right shows the sparsity as a function of γ. As γ varies, the percentage of nonzero entries of the gain matrix K decreases. For $\gamma = 0.119$ there are 87 nonzero entries corresponding to the cost increase of 1.89%; for $\gamma = 0.143$, there are 45 nonzero entries with a cost increase of 30.0%. The results of the optimization problem (L-LQG) with the inclusion of Gerschgorin circle constraint (2) are shown in Figures 1c and 1d. In this case, the cost at $\gamma = .119$ increases only by 0.65%. Furthermore, as $\gamma \geq 0.37$ the sparsity remains at 91 nonzero entries and the cost remains at 1.43% increase above the centralized value.

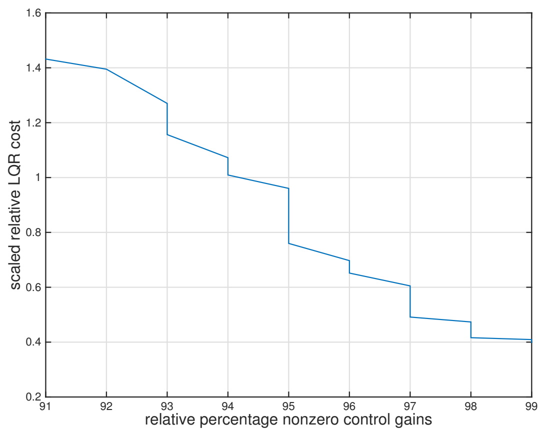
In the third case, the matrix $A = A_2$ is unstable. Without inclusion of Gerschgorin constraint (2) in



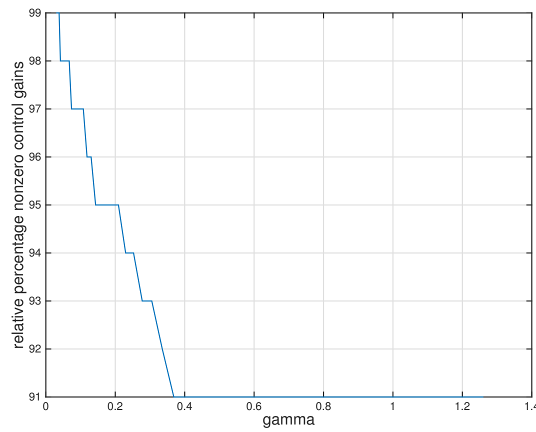
(a) Stable A_1 matrix, Gerschgorin not included.



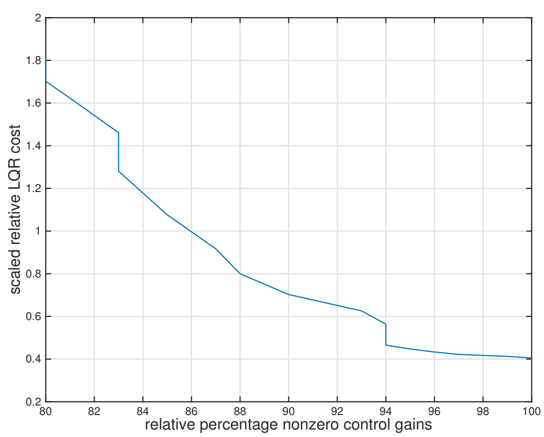
(b) Stable A_1 matrix, Gerschgorin not included.



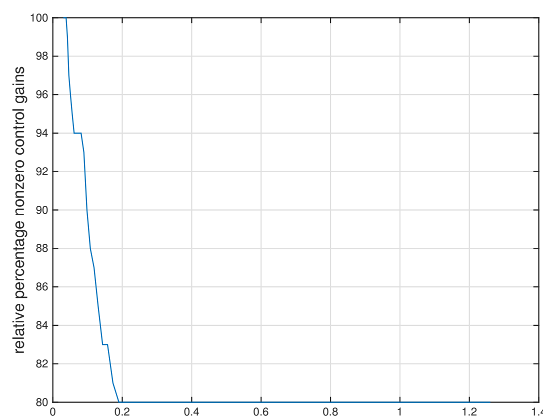
(c) Stable A_1 matrix, Gerschgorin included.



(d) Stable A_1 matrix, Gerschgorin included.



(e) Unstable A_2 matrix, Gerschgorin included.



(f) Unstable A_2 matrix, Gerschgorin included.

Fig. 1: The dynamics are given by the randomly generated stable matrix A_1 in the first four figures and by randomly generated unstable matrix A_2 in the last two figures. In case of unstable A_2, with the inclusion of Gerschgorin circle constraint we can find stable closed loop dynamics but the sparsity in the control gain remains above a threshold.

(L-LQG) the resulting close loop system $A + BK(\gamma)$ is unstable, even for small γ. Thus, we include this constraint. Figures 1e and 1f show the results. With $\gamma = .0316$, there are 100 nonzero entries, but the cost is 0.35% higher than the centralized cost. This highlights that the Gerschgorin sufficient stability condition can be conservative. Similar to the previous two cases, with increasing γ the sparsity of the gain K increases and the cost degrades. At $\gamma = 0.190$, the cost increases by 1.70% corresponding to sparsity of 80%. The sparsity remains constant at 80% for all $\gamma \geq 0.190$.

In all simulations we conducted, in the case in which matrix A was stable, without inclusion of constraint (2) in (L-LQG) the closed loop dynamics remained stable for γ within the given bounds. However, in the case of unstable A, inclusion of constraint (2) was necessary to ensure stability of sparse solutions. When this stability condition was included in the SDP, the sparsity pattern was not affected too much by the variations in γ.

V. Conclusions and Future Work

We considered sparse controller design for an infinite horizon discrete-time LQG problem. Given that the optimal centralized controller solves a linear equation defined by the second order moments of the states and inputs, we searched for sparse approximate solutions of this linear equation (sparse least squares). Based on the Gerschgorin circle theorem, we included a sufficient condition for stability that also preserves convexity of the sparse least squares formulation. We verified the performance of the resulting sparse controllers in simulation. In the cases considered, a stabilizing sparse controller was ensured with the Gerschgorin circle constraint. This came at a price of sub-optimality of the controller.

Future work includes exploration of less conservative stability conditions in the sparsity constraint design. To do so one can formulate the SDP problem for the moments of the states and inputs, with the addition of non-convex sparse/structured constraint and then seek convex relaxation or suboptimal solutions of this problem. Furthermore, we hope to derive improved lower bounds on the performance of sparse and structured controllers. In particular, so far the measure of performance of the sparse controller has been sub-optimality with respect to the performance of the centralized controller. By improving the lower bound of performance we gain a better understanding of achievable performance of sparse linear controllers in the problems in which quadratic invariance condition does not hold. Finally, we hope to apply our algorithms to real-world complex problems arising in power system domain.

References

[1] H. S. Witsenhausen, "A counterexample in stochastic optimum control," *SIAM Journal on Control*, vol. 6, no. 1, pp. 131–147, 1968.

[2] C. H. Papadimitriou and J. Tsitsiklis, "Intractable problems in control theory," *SIAM journal on control and optimization*, vol. 24, no. 4, pp. 639–654, 1986.

[3] Y. C. Ho and K.-h. Chu, "Team decision theory and information structures in optimal control problems–part i," *Automatic Control, IEEE Transactions on*, vol. 17, no. 1, pp. 15–22, 1972.

[4] N. Motee, A. Jadbabaie, and B. Bamieh, "On decentralized optimal control and information structures," in *American Control Conference*. IEEE, 2008, pp. 4985–4990.

[5] M. Rotkowitz and S. Lall, "A characterization of convex problems in decentralized control," *Automatic Control, IEEE Transactions on*, vol. 51, no. 2, pp. 274–286, 2006.

[6] M. Rotkowitz, "On information structures, convexity, and linear optimality," in *Conference on Decision and Control*. IEEE, 2008, pp. 1642–1647.

[7] F. Lin, M. Fardad, and M. R. Jovanovic, "Design of optimal sparse feedback gains via the alternating direction method of multipliers," *Automatic Control, IEEE Transactions on*, vol. 58, no. 9, pp. 2426–2431, 2013.

[8] S. P. Boyd, L. El Ghaoui, E. Feron, and V. Balakrishnan, *Linear matrix inequalities in system and control theory*. SIAM, 1994, vol. 15.

[9] S. A. Gershgorin, "über die abgrenzung der eigenwerte einer matrix," *zv. Akad. Nauk. USSR Otd. Fiz.-Mat. Nauk*, no. 6, pp. 749–754, 1931.