Sparse optimal control of networks with multiplicative noise via policy gradient *

Benjamin Gravell   **Yi Guo**   **Tyler Summers**

The University of Texas at Dallas, Richardson, TX 75080 USA

Abstract: We give algorithms for designing near-optimal sparse controllers using policy gradient with applications to control of systems corrupted by multiplicative noise, which is increasingly important in emerging complex dynamical networks. Various regularization schemes are examined and incorporated into the optimization by the use of gradient, subgradient, and proximal gradient methods. Numerical experiments on a large networked system show that the algorithms converge to performant sparse mean-square stabilizing controllers.

Keywords: Optimal control, multiplicative noise, networks, sensor & actuator placement

1. INTRODUCTION

Emerging highly distributed networked dynamical systems, such as critical infrastructure for power, water, and transportation, are high-dimensional and increasingly instrumented with new sensing, actuation, and communication technologies. A key problem is to design high performance control architectures that limit the number of actuators, sensors, and actuator-sensor communication links to reduce complexity and cost. Sparse control architectures may be crucial for managing complexity in emerging complex networks, but require solution of extremely difficult mixed combinatorial-continuous optimization problems.

There is a variety of performance metrics and optimization methodology for sparse control architecture design in the recent literature. Examples include structural rank conditions from Liu et al. (2011); Ruths and Ruths (2014); Olshevsky (2014), controllability and observability Gramians from Pasqualetti et al. (2014); Summers et al. (2016); Tzoumas et al. (2016); Jadbabaie et al. (2018), and optimal and robust control metrics from Hassibi et al. (1998); Polyak et al. (2013); Jovanović and Dhingra (2016); Summers (2016); Taha et al. (2019); Zare and Jovanović (2018), which are optimized via greedy algorithms, convex and mixed-integer optimization, and randomization.

Here we develop methods for sparse optimal control design in dynamical networks with multiplicative noise via policy gradient algorithms with sparsity-inducing regularization. Multiplicative noise arises in many networked systems when the weights of edges connecting nodes are stochastic in time. The noise is thus on the system parameters themselves and has a fundamentally different effect on the state evolution than additive noise, and indeed can lead to dramatic robustness issues. Specifically, a noise-ignorant classical optimal linear-quadratic (LQ) controller may actually *destabilize* a multiplicative noise system in the mean-square sense, even if the system was open-loop mean-square stable. Therefore noise-aware control is imperative

to network performance and robustness. Moreover, the policy gradient methods we propose here, which operate directly on policy parameters, facilitate data-driven sparse control design when the model is unknown, a topic we are exploring in ongoing work.

In Section 2 we formulate the problem and discuss a policy gradient approach to optimal control design for linear-quadratic systems with multiplicative noise. In Section 3 we propose several sparse control design methods for sensor and actuator selection and communication network design using gradient, subgradient, and proximal algorithms. In Section 4 we present numerical experiments to illustrate the results. In Section 5 we conclude.

2. PROBLEM FORMULATION

Consider the discrete-time linear quadratic regulator with multiplicative noise (LQRm) optimal control problem

$$
\begin{aligned}
\min_{\pi \in \Pi} \quad & J(\pi) = \mathbb{E} \sum_{t=0}^{\infty} (x_t^T Q x_t + u_t^T R u_t), \\
\text{s.t.} \quad & x_{t+1} = (A + \sum_{i=1}^{p} \delta_{it} A_i) x_t + (B + \sum_{j=1}^{q} \gamma_{jt} B_j) u_t
\end{aligned}
\tag{1}
$$

where $x_t \in \mathbb{R}^n$ is the system state, $u_t \in \mathbb{R}^m$ is the control input, x_0 is randomly distributed according to $\mathcal{P}$, expectation is with respect to $x_0, \delta_{it}, \gamma_{jt}$, and $Q \succeq 0$ and $R \succ 0$. The dynamics incorporate multiplicative noise terms modeled by the mutually independent and i.i.d. (over time) zero-mean random variables δ_{it} and γ_{jt}, which have variance α_i and β_j, respectively. The matrices $A_i \in \mathbb{R}^{n \times n}$ and $B_i \in \mathbb{R}^{n \times m}$ specify how each noise term affects the state-to-state and input-to-state dynamics matrices A and B respectively. The goal is to determine an optimal closed-loop feedback policy π with $u_t = \pi(x_t)$. We assume A, B, α_i, A_i, β_j, and B_j are such that a finite optimal value of the problem exists, which is ensured if the system is mean-square stabilizable.

Definition 1. (Mean-square stability). The system in (1) is stable in the mean-square sense if $\lim_{t \to \infty} \mathbb{E}[x_t x_t^T] = 0$ for any given initial covariance $\mathbb{E}x_0 x_0^T$.

We are ultimately interested in the problem
$$\min_{\pi \in \Pi} \quad J(\pi) + \gamma J_{\text{reg}}(\pi) \tag{2}$$
where $J_{\text{reg}}(\pi)$ is a sparsity-promoting regularizer of the policy π and γ specifies the importance of sparsity. The regularizer ideally measures the number of actuators, sensors, or actuator-sensor links, but for computational tractability will be replaced by other functions. We begin by discussing the solution for $\gamma = 0$.

2.1 Optimal control via value iteration

Dynamic programming can be used to show that the optimal policy is linear state feedback with $u_t = K^* x_t$ where $K^* \in \mathbb{R}^{m \times n}$ and the resulting optimal cost for a fixed initial state is quadratic with $V_{K^*}(x_0) = x_0^T P x_0$ where $P \in \mathbb{R}^{n \times n}$ is a symmetric positive definite matrix. When the model parameters are known, there are several known ways to compute the optimal feedback gains and corresponding optimal cost. The optimal cost is found by solving the generalized algebraic Riccati equation (ARE) (see e.g. Damm (2004)) via e.g. value iteration:

$$P = Q + A^T P A + \sum_{i=1}^{p} \alpha_i A_i^T P A_i$$
$$- A^T P B (R + B^T P B + \sum_{j=1}^{q} \beta_j B_j^T P B_j)^{-1} B^T P A.$$

The corresponding optimal gain matrix is

$$K^* = -\Big(R + B^T P B + \sum_{j=1}^{q} \beta_j B_j^T P B_j \Big)^{-1} B^T P A.$$

2.2 Optimal control via policy gradient

For a fixed mean-square stabilizing linear state feedback policy $u_t = K x_t$, there exists a positive semidefinite cost matrix P_K which characterizes the cost by

$$J(K) = \mathbb{E}_{x_0} x_0^T P_K x_0$$

and is the solution to the generalized Lyapunov equation

$$P_K = Q + K^T R K + (A + BK)^T P_K (A + BK)$$
$$+ \sum_{i=1}^{p} \alpha_i A_i^T P_K A_i + \sum_{j=1}^{q} \beta_j K^T B_j^T P_K B_j K.$$

Furthermore, there exists a positive semidefinite infinite-horizon aggregate state covariance matrix

$$\Sigma_K = \mathbb{E}_{x_0, \delta_{it}, \gamma_i} \sum_{t=0}^{\infty} x_t x_t^T$$

which is the solution to the generalized Lyapunov equation

$$\Sigma_K = \Sigma_0 + (A + BK) \Sigma_K (A + BK)^T$$
$$+ \sum_{i=1}^{p} \alpha_i A_i \Sigma_K A_i^T + \sum_{j=1}^{q} \beta_j B_j K \Sigma_K K^T B_j^T,$$

where $\Sigma_0 = \mathbb{E}_{x_0} \left[x_0 x_0^T \right]$. Thus, we have

$$J(K) = \text{Tr}((Q + K^T R K)\Sigma_K) = \text{Tr}(P_K \Sigma_0).$$

This leads to the idea of performing gradient descent on J (i.e. policy gradient) to find the optimal gain matrix:

$$K \leftarrow K - \eta \nabla_K J(K),$$

for a fixed step size η. In this work we consider only the case where the model parameters are known, but the methods presented are immediately usable in the model-unknown case by estimating the gradient from trajectory data. The policy gradient for linear state feedback policies applied to the LQRm problem has the following form:

Lemma 2. The LQRm policy gradient $\nabla_K J(K)$ is

$$2\Big[\Big(R + B^T P_K B + \sum_{j=1}^{q} \beta_j B_j^T P_K B_j\Big)K + B^T P_K A\Big]\Sigma_K.$$

The proof can be found in our paper (Gravell et al. (2019)).

2.3 Gradient domination

It was shown recently by Fazel et al. (2018) that although the deterministic LQR cost is nonconvex, it is *gradient dominated*, i.e. satisfies the Polyak-Lojasiewicz inequality originally due to Polyak (1963). It is simple to show that if a function has a Lipschitz continuous gradient and satisfies this condition then performing gradient descent with a sufficiently small constant step size will result in asymptotic convergence to the optimal function value at a linear rate (see Karimi et al. (2016)). For the LQRm problem, so long as the initial controller is stabilizing the LQRm cost is continuously differentiable over the sublevel set associated with the initial controller's cost and thus the gradient possesses a local Lipschitz constant L on this set. Identifying L and the gradient domination constant is necessary for selection of a step size which is guaranteed to give convergence using gradient descent. Quantifying these constants is possible via lengthy chains of matrix inequalities as demonstrated by Fazel et al. (2018).

These results extend readily to the LQRm problem with relevant quantities pertaining to Lipschitz continuity of the gradient and the gradient domination conditions suitably modified. In particular, multiplicative noise decreases the maximum step size (and thus convergence rate) that can be taken using gradient descent, and in the model-free setting increases the required number of samples and rollout length. The results and proofs also differ from that of Fazel et al. (2018) because mean-square stability complicates the state covariance propagation and commensurate analysis and because *generalized* Lyapunov equations must be solved to compute the gradient steps, which requires specialized solvers. We now state the relevant lemmas; the proofs can be found in Gravell et al. (2019).

Lemma 3. (Gradient domination). The LQRm cost $J(K)$ satisfies the gradient domination condition

$$J(K) - J(K^*) \le \frac{\|\Sigma_{K^*}\|}{4\sigma_{\min}(R)\sigma_{\min}(\Sigma_0)^2}\|\nabla_K J(K)\|_F^2$$

Lemma 4. (Gradient descent, convergence rate). Using the policy gradient step update

$$K^{(k+1)} = K^{(k)} - \eta \nabla_K J(K^{(k)})$$

with step size $0 < \eta \le c_{pg}$ gives global convergence to the optimal K^* at a linear rate described by

$$\frac{J(K^{(k+1)}) - J(K^*)}{J(K^{(k)}) - J(K^*)} \le 1 - \eta \frac{\sigma_{\min}(R)\sigma_{\min}(\Sigma_0)^2}{\|\Sigma_{K^*}\|}$$

where c_{pg} is a constant which is polynomial in the parameters A, B, B_j, Q, R, $J(K^{(0)})$.

3. SPARSE CONTROL DESIGN

Entrywise, row, and column sparsity in K correspond to actuator-sensor communication, actuator, and sensor sparsity respectively. With this in mind, we seek to solve the optimization problem of finding the sparsest set of entries, rows and/or columns of K that achieve some prescribed level of performance in terms of the LQRm cost. However this problem is a nonconvex combinatorial problem which is NP-hard; the number of independent problem instances which must be solved scales factorially with n and/or m. We instead turn to regularization as a heuristic to identifying good sparsity patterns.

3.1 Insufficiency of naïve hard thresholding

The most naïve method of inducing sparsity is hard thresholding of the ARE solution as $K_{ij} = 0$ if $|K_{ij}| < r$. However, in general this is not useful since the resulting gains may not be stabilizing. Consider the following example system without multiplicative noise:

$$A = \begin{bmatrix} 0.4 & 0.9 & -0.3 \\ 0.7 & -0.3 & -0.4 \\ -0.2 & 0.1 & -0.8 \end{bmatrix}, \quad B = \begin{bmatrix} 0.2 & -0.6 \\ -1.3 & -1.6 \\ -0.3 & -1.5 \end{bmatrix},$$

$$\mathcal{P} = \mathcal{N}(0, I_3), \quad Q = I_3, \quad R = I_2$$

where I_n is an $n \times n$ identity matrix. Imposing a hard threshold of 0.4 on the ARE solution results in

$$K = \begin{bmatrix} 0.503931 & -0.880322 & 0 \\ 0 & 0.614382 & -0.677758 \end{bmatrix}$$

which gives a closed-loop state transition matrix $A + BK$ with an eigenvalue of 1.048223 outside the unit circle. By contrast, by working with the regularized LQRm cost the optimal gains are always guaranteed to be stabilizing; even in the limit as the regularization weight $\to \infty$ the sparsity increases until the sparsest stabilizing solution is obtained. In practice, using a small step size helps ensure that each iterate remains inside the domain of $J(K)$.

3.2 Regularization

Certain types of regularization are well-known to be capable of inducing sparsity in the solutions to optimization problems. Perhaps the most basic and well-known is l_1-norm regularization which operates on a vector of decision variables; see Tibshirani (1996) for the seminal LASSO problem for sparse least-squares model selection and Hassibi et al. (1998) for sparse control design. In the case of a convex objective, increasing the regularization weight tends to increase sparsity by moving the global minimum onto the coordinate axes. Once the regularized problem has been solved, a sparsity pattern can easily be identified from the (near-)zero entries. Here we consider only the problem of identifying sparsity patterns, however an additional polishing step which re-solves the LQRm problem under a fixed sparsity pattern can be performed to further decrease the LQRm cost, as in Lin et al. (2013). Convergence using projected policy gradient descent was recently shown for deterministic LQR by Bu et al. (2019).

Entrywise sparsity is induced by the vector l_1-norm

$$\|K\|_1 = \sum_{i=1}^{m} \sum_{j=1}^{n} |K_{ij}|.$$

Row and column sparsity are induced by using matrix row and column norms respectively defined as

$$\|K\|_r = \sum_{i=1}^{m} \|K^{r,i}\|_\infty, \quad \text{and} \quad \|K\|_c = \sum_{i=1}^{n} \|K^{c,i}\|_\infty,$$

where $\|K^{r,i}\|_\infty$ and $\|K^{c,i}\|_\infty$ are the maximum absolute values of the ith row and column respectively of K. Row and column sparsity are also induced by the row and column group LASSO

$$\|K\|_{glr} = \sum_{i=1}^{m} \|K^{r,i}\|_2, \quad \text{and} \quad \|K\|_{glc} = \sum_{i=1}^{n} \|K^{c,i}\|_2,$$

where $\|K^{r,i}\|_2$ and $\|K^{c,i}\|_2$ are the vector l_2-norms of the ith row and column respectively of K. Combined row and column sparsity can be induced by the row and column sparse group LASSO

$$\|K\|_{sglr} = (1-\mu)\|K\|_1 + \mu\|K\|_{glr}$$
$$\|K\|_{sglc} = (1-\mu)\|K\|_1 + \mu\|K\|_{glc}$$

or by various other weighted combinations of entrywise, row, and column norms. We refer to $\|K\|_M$ as a generic nondifferentiable sparsity-inducing regularizer.

3.3 Stationary point characterization

Before proceeding, we must point out an important consequence of regularizing the LQRm cost. The sum of a convex function and a gradient dominated function is not gradient dominated in general, and in fact can have multiple local minima. For example, consider the scalar function

$$f(x) = x^2 + 4((x-8)^2 + 3\sin^2(x-8))$$

where x^2 is strongly convex and $4((x-8)^2 + 3\sin^2(x-8))$ is gradient dominated. But $f(x)$ has two local minima at $x = 5.372$ and $x = 7.459$ and therefore is not gradient dominated.

As a result any local first-order search procedure, such as those used by our algorithms, will not be guaranteed to find the global minimum. We conjecture that for the regularized LQRm problem there are at most two local minima, one associated with the LQRm cost and one associated with the regularization which tends to be more sparse. If this is so then choosing the initial point carefully may help the local search find the desired (sparser) local minimum. For open-loop mean-square systems, this motivates using zero gains as the initial condition. Likewise, in both the open-loop mean-square stable and unstable cases, an effective heuristic is to use the solution to a highly regularized problem instance to "warm start" another nearby problem instance with reduced regularization weight.

3.4 Algorithm design

We examine three variants of gradient descent which accommodate nondifferentiable regularization; subgradient descent, proximal gradient descent, and gradient descent with smooth Huber-type regularizer approximation. All three are natural ways a practitioner might solve the optimization problem, and our results reveal that empirically all three indeed converge to solutions of similar quality at similar computational cost.

Promising choices of step directions other than the gradient $\nabla_K J(K)$ are the natural gradient $\nabla_K J(K)\Sigma_K^{-1}$ and the Gauss-Newton step $R_K^{-1}\nabla_K J(K)\Sigma_K^{-1}$ as given by Fazel et al. (2018). When $\gamma = 0$, these step directions give faster convergence than the gradient step and in fact the convergence proofs are much simpler than that for the gradient step. Unfortunately, adding a regularizer makes these steps more difficult to calculate; it is not simply the sum of the gradient of the regularizer and the unregularized natural gradient or Gauss-Newton step of the LQRm cost. For this reason we restrict our attention to the standard (sub)gradient directions.

3.5 Regularized policy subgradient descent

In order to use nondifferentiable regularizers we use subgradient methods which take steps in the direction of subgradients. It is known that using a constant step size gives convergence to a bounded neighborhood of the optimum and that a diminishing step size gives asymptotic, albeit slow, convergence (see Nesterov (2013)). One immediate issue is that subgradients are defined only for convex functions; since the LQRm cost is nonconvex, subgradients do not exist for the regularized LQRm cost. However we simply use the gradient of the LQRm cost plus the subgradient of the regularizer as the step direction. Thus our subgradient descent update is

Algorithm 1. (Policy subgradient update).
$$K^{(k+1)} = K^{(k)} - \eta(\nabla_K J(K^{(k)}) + \gamma \blacktriangledown_K \|K^{(k)}\|_M)$$
$$K_{\text{best}} = \arg\min \{C(K^{(k+1)}), C(K_{\text{best}})\}$$
where $C(K) = J(K) + \gamma\|K\|_M$ and $\blacktriangledown_K$ is a subgradient.

Another issue is that there is no guarantee of feasibility of each next step; it is possible to take a step so large that the next point is a mean-square unstable controller giving infinite objective cost. It is not straightforward to obtain restrictions on the step size to guarantee this feasibility. Gradient descent does not suffer from this problem since the gradient is guaranteed to be a true descent direction so there is always a sufficiently small step size to give a feasible next step. Nevertheless, in practice it is rare for a sufficiently small subgradient step to be infeasible.

3.6 Proximal policy gradient

Proximal gradient methods have become a preferred way to solve optimization problems of the form
$$\min_x f(x) + g(x)$$
where $f(x)$ has a Lipschitz continuous gradient and $g(x)$ is convex and nondifferentiable, as is the case when $g(x)$ is a sparsity-inducing regularizer. The proximal gradient method update is
$$x^{(k+1)} = \text{prox}_{\eta g}\left(x^{(k)} - \eta \nabla f\left(x^{(k)}\right)\right)$$
where the proximity operator is defined as
$$\text{prox}_{\eta g}(v) = \arg\min_x \left\{g(x) + (1/(2\eta))\|x - v\|_2^2\right\}.$$
Much of the existing literature examines the case where $f(x)$ is convex, in which case gradient descent is guaranteed to converge. The proximal operator has closed-form expressions for $\|K\|_1$ and $\|K\|_{glr}$ called *soft thresholding*

and *block soft thresholding* (see Parikh et al. (2014)). Thus to solve (2) we also use proximal policy gradient:

Algorithm 2. (Proximal policy gradient update).
$$K_s^{(k)} = K^{(k)} - \eta\nabla_K J(K^{(k)})$$
$$K^{(k+1)} = \text{prox}_{\eta\|K\|_M}\left(K_s^{(k)}\right)$$

A result from Hassan-Moghaddam and Jovanović (2018) guarantees convergence at a linear rate to the optimal function value using the proximal gradient method on a function satisfying a proximal gradient domination condition. This condition was shown to be equivalent by one given by Karimi et al. (2016) and an inequality from Kurdyka (1998). However, this condition is not guaranteed to hold when $f(x)$ is gradient dominated and $g(x)$ is convex; the full condition must be checked, which involves interaction between $f(x)$ and $g(x)$. It is nontrivial to verify that the condition is satisfied for the regularized LQRm cost. Empirically it seems the inequality may be satisfied since the proximal gradient method converged to solutions similar to those from our other two methods.

3.7 Regularized policy gradient descent

Another algorithm for solving (2) is gradient descent:

Algorithm 3. (Policy gradient update).
$$K^{(k+1)} = K^{(k)} - \eta\nabla_K(J(K^{(k)}) + \gamma\|K\|_M)$$

where we use differentiable Huber-type losses $\|K\|_{M,h,\phi}$ in place of nondifferentiable regularizers, which replace linear corners with quadratic tips for decision variable values smaller than a specified threshold. Although the solutions produced are not exactly sparse, in practice entries are sufficiently close to zero to identify the sparsity pattern. Furthermore, by iteratively decreasing the threshold the solutions can be made arbitrarily close to truly sparse.

We define the Huber function of a scalar a as
$$h_\phi(a) = \begin{cases} |a| - \dfrac{1}{2}\phi & \text{if } |a| > \phi \\ \dfrac{1}{2\phi}a^2 & \text{if } |a| \leq \phi \end{cases}$$
and the p-Huber function (like a p-norm) of a vector b as
$$h_{p,\phi}(b) = \begin{cases} \|b\|_p - \dfrac{1}{2}\phi & \text{if } \|b\|_p > \phi \\ \dfrac{1}{2\phi}\|b\|_p^2 & \text{if } \|b\|_p \leq \phi \end{cases}$$
We define the vector Huber loss as
$$\|K\|_{1,h,\phi} = \sum_{i=1}^m \sum_{j=1}^n h_\phi(K_{ij})$$
the Huber row and column norms as
$$\|K\|_{r,h,\phi} = \sum_{i=1}^m h_{\infty,\phi}(K^{r,i}), \quad \|K\|_{c,h,\phi} = \sum_{i=1}^n h_{\infty,\phi}(K^{c,i})$$
and the Huber row and column group LASSO as
$$\|K\|_{glr,h,\phi} = \sum_{i=1}^m h_{2,\phi}(K^{r,i}), \quad \|K\|_{glc,h,\phi} = \sum_{i=1}^n h_{2,\phi}(K^{c,i})$$
Subgradients of two regularizers and the gradients of their differentiable counterparts are given in Table 1.

Table 1. Regularizer (sub)gradients

$\|K\|_M$	$\blacktriangledown_K \|K\|_M$	$\nabla_K \|K\|_{M,h,\phi}$
$\|K\|_{1,ij}$	$\mathrm{sgn}(K_{ij})$	$K_{ij}/\max(|K_{ij}|,\phi)$
$\|K\|_{glr}^{r,i}$	$K^{r,i}/\|K^{r,i}\|_2$	$K^{r,i}/\max(\|K^{r,i}\|_2,\phi)$

4. SIMULATION RESULTS

We considered an example system which represents diffusion dynamics on a particular undirected Erdős-Rényi random graph. It is well known that if $p_{ER} = (\log n + c)/n$ for constant $c \in \mathbb{R}$, then $\lim_{n\to\infty} P(G(n,p) \text{ connected}) = e^{-e^{-c}}$ so we chose $n = 51$, $c = 7$ and $p_{ER} = 0.2144$ and with probability $P = 0.999$ obtained a connected graph (see Bollobás and Béla (2001)). The graph was selected so it was connected, ensuring controllability. The first row and and column of the graph Laplacian were removed in order to fix the system's state reference to the first node which removed the zero eigenvalue otherwise present. The continuous time system was discretized using a standard bilinear transform (Tustin's approximation) which preserved the open-loop mean stability of this system. Two multiplicative noises acted each on A and B with entries drawn from a Gaussian distribution. The noise variances were set at two levels, low and high, so the system was open-loop mean-square stable and unstable, respectively.

For the subgradient and proximal gradient methods, we stopped iterating after the best iterate had been held for 100 iterations. For the gradient method, we stopped iterating when the Frobenius norm of the gradient of the cost function fell below a small threshold value, $0.1 \times \mathrm{card}(K)$. We swept through a range of sparsity levels by solving a problem with low γ then increasing γ and resolving the problem using the previous solution as the initial guess. The step size η was initialized at 10^{-5}. For the l_1-norm and row group LASSO γ was initialized at 10 and 100 respectively. For each successive problem, the regularization weight was multiplied by a ratio $r_\gamma = \sqrt{2}$ and the step size was multiplied by $r_\eta = r_\gamma^{-\sqrt[4]{2}}$. To determine sparsity patterns we considered a value to be sparse if it was less than 5% than the max value in K. For the l_1-norm the sparsity values were the absolute values of the entries. For the row group lasso norm the sparsity values were the the values are the l_2-norms of the rows and columns respectively. Sparsity patterns are presented in Figs. 1 and 2 with white cells representing (near-)zero entries. The LQRm costs given in Figs. 3 and 4 are for the sparse gains without any polishing step applied, which otherwise could significantly reduce the cost. We give the total "wall-clock" computation time in Fig. 4 to capture the aggregate computational expense of each algorithm. The main computational expense came from evaluating the LQRm gradient at each iteration, which required solving a generalized discrete Lyapunov equation.

As seen in Fig. 4, the first iteration had the longest compute time since successive iterations benefited from favorable initial conditions from warm-starting. The compute time increased as the regularization weight was increased and a larger number of smaller steps were required to accommodate the increasing gradient magnitude.
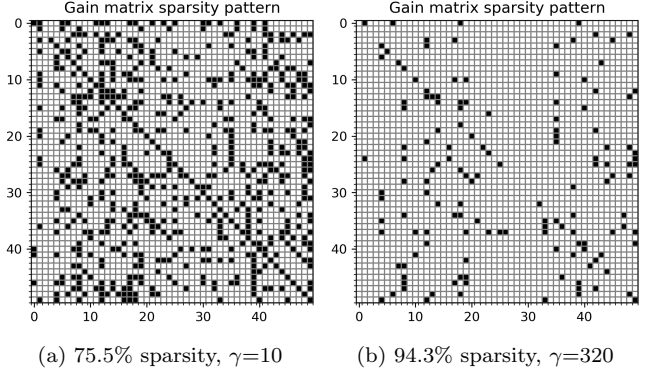

(a) 75.5% sparsity, γ=10   (b) 94.3% sparsity, γ=320

Fig. 1. Sparsity patterns for low noise, subgradient descent on the l_1-norm regularized LQRm cost.
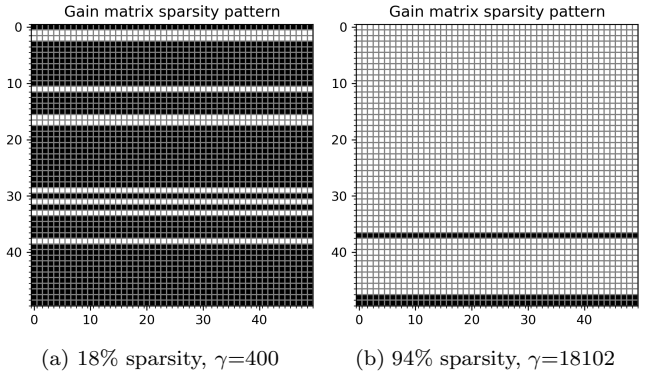

(a) 18% sparsity, γ=400   (b) 94% sparsity, γ=18102

Fig. 2. Sparsity patterns for low noise, subgradient descent on the row group LASSO regularized LQRm cost.

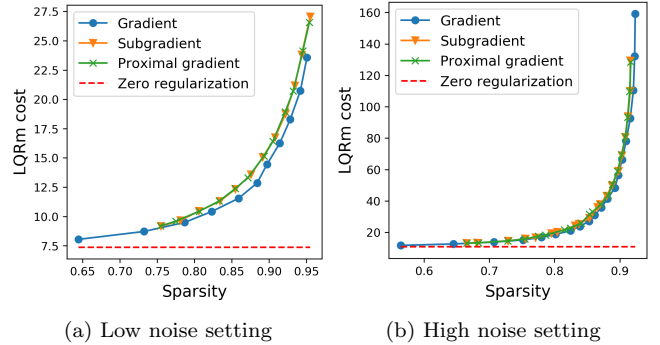(a) Low noise setting   (b) High noise setting

Fig. 3. LQRm cost vs. sparsity with l_1-norm regularization.

From our empirical studies, the three methods presented all gave very similar results with similar efficacy; arbitrarily entrywise and row sparse mean-square stabilizing solutions were obtained for the low noise setting after a reasonable amount of computation time. Similarly, very sparse solutions for the high noise setting were obtained.

5. CONCLUDING REMARKS

We developed three policy gradient algorithms for solving the sparse gain design problem for networked dynamical systems with multiplicative noise. We showed that the regularized LQR cost does not necessarily have a unique local minimum, hampering efforts to guarantee global con-

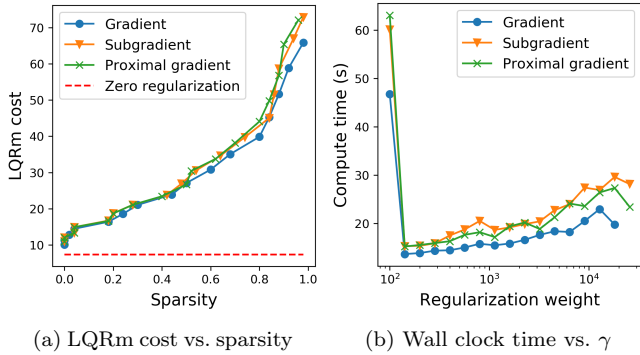(a) LQRm cost vs. sparsity   (b) Wall clock time vs. γ

Fig. 4. Algorithm comparisons for the low noise setting with row group LASSO regularization.

vergence of the algorithms. Nevertheless, efficacy of the algorithms is demonstrated empirically via computational simulations. Through various regularization functions we identified sparsity patterns for near-optimal actuator, sensor, and actuator-sensor link removal. This paves the way for data-driven control design in the model-free setting for such systems.

Future work will attempt to prove unique local minimization of the regularized LQR cost or provide a set of restrictions under which such a condition holds. A salient issue with policy gradient methods relates to scalability; for large systems the gradient calculation is computationally expensive. Hence we will explore low-rank approximations of the gradient and consequent effects on convergence. We will also extend this work to the unknown-model setting and explore alternative model-based learning schemes.

REFERENCES

Bollobás, B. and Béla, B. (2001). *Random graphs.* 73. Cambridge university press.

Bu, J., Mesbahi, A., Fazel, M., and Mesbahi, M. (2019). LQR through the lens of first order methods: Discrete-time case. *arXiv preprint arXiv:1907.08921.*

Damm, T. (2004). *Rational matrix equations in stochastic control*, volume 297. Springer Science & Business Media.

Fazel, M., Ge, R., Kakade, S., and Mesbahi, M. (2018). Global convergence of policy gradient methods for the linear quadratic regulator. In *Proceedings of the 35th International Conference on Machine Learning*, volume 80 of *Proceedings of Machine Learning Research*, 1467–1476. PMLR.

Gravell, B., Esfahani, P.M., and Summers, T. (2019). Learning robust control for LQR systems with multiplicative noise via policy gradient. URL https://arxiv.org/abs/1905.13547. Under review.

Hassan-Moghaddam, S. and Jovanović, M.R. (2018). On the exponential convergence rate of proximal gradient flow algorithms. In *2018 IEEE Conference on Decision and Control (CDC)*, 4246–4251. IEEE.

Hassibi, A., How, J., and Boyd, S. (1998). Low-authority controller design via convex optimization. In *Proceedings of the 37th IEEE Conference on Decision and Control (Cat. No.98CH36171)*, volume 1, 140–145 vol.1.

Jadbabaie, A., Olshevsky, A., and Siami, M. (2018). Deterministic and randomized actuator scheduling

with guaranteed performance bounds. *arXiv preprint arXiv:1805.00606.*

Jovanović, M.R. and Dhingra, N.K. (2016). Controller architectures: Tradeoffs between performance and structure. *European Journal of Control*, 30, 76–91.

Karimi, H., Nutini, J., and Schmidt, M. (2016). Linear convergence of gradient and proximal-gradient methods under the Polyak-Łojasiewicz condition. In *Machine Learning and Knowledge Discovery in Databases*, 795–811. Springer International Publishing, Cham.

Kurdyka, K. (1998). On gradients of functions definable in o-minimal structures. *Annales de l'institut Fourier*, 48(3), 769–783.

Lin, F., Fardad, M., and Jovanović, M.R. (2013). Design of optimal sparse feedback gains via the alternating direction method of multipliers. *IEEE Transactions on Automatic Control*, 58(9), 2426–2431.

Liu, Y.Y., Slotine, J.J., and Barabási, A.L. (2011). Controllability of complex networks. *Nature*, 473(7346), 167–173.

Nesterov, Y. (2013). *Introductory lectures on convex optimization: A basic course*, volume 87. Springer Science & Business Media.

Olshevsky, A. (2014). Minimal controllability problems. *IEEE Transactions on Control of Network Systems*, 1(3), 249–258.

Parikh, N., Boyd, S., et al. (2014). Proximal algorithms. *Foundations and Trends in Optimization*, 1(3), 127–239.

Pasqualetti, F., Zampieri, S., and Bullo, F. (2014). Controllability metrics, limitations and algorithms for complex networks. *IEEE Transactions on Control of Network Systems*, 1(1), 40–52.

Polyak, B., Khlebnikov, M., and Shcherbakov, P. (2013). An LMI approach to structured sparse feedback design in linear control systems. In *Proc. European Control Conference*, 833–838.

Polyak, B. (1963). Gradient methods for the minimisation of functionals. *USSR Computational Mathematics and Mathematical Physics*, 3(4), 864 – 878.

Ruths, J. and Ruths, D. (2014). Control profiles of complex networks. *Science*, 343(6177), 1373–1376.

Summers, T., Cortesi, F., and Lygeros, J. (2016). On submodularity and controllability in complex dynamical networks. *IEEE Transactions on Control of Network Systems*, 3(1), 91–101.

Summers, T. (2016). Actuator placement in networks using optimal control performance metrics. In *IEEE Conference on Decision and Control*, 2703–2708.

Taha, A.F., Gatsis, N., Summers, T., and Nugroho, S. (2019). Time-varying sensor and actuator selection for uncertain cyber-physical systems. *IEEE Transactions on Control of Network Systems. to appear.*

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society. Series B (Methodological)*, 58(1), 267–288.

Tzoumas, V., Rahimian, M.A., Pappas, G., and Jadbabaie, A. (2016). Minimal actuator placement with bounds on control effort. *IEEE Transactions on Control of Network Systems*, 3(1), 67–78.

Zare, A. and Jovanović, M.R. (2018). Optimal sensor selection via proximal optimization algorithms. In *2018 IEEE Conference on Decision and Control (CDC)*, 6514–6518. IEEE.