2017 IEEE 56th Annual Conference on Decision and Control (CDC)
December 12-15, 2017, Melbourne, Australia

Structured covariance completion via proximal algorithms

Armin Zare, Neil K. Dhingra, Mihailo R. Jovanović, and Tryphon T. Georgiou

Abstract— **This paper studies the interplay between dynamics and statistics of a stochastically driven dynamical system. Motivation is provided by applications in fluid flow modeling and control. In this context, second-order statistics around the mean velocity profile can be obtained, for a subset of variables, from experiments or numerical simulations. The basic idea is to determine a parsimonious perturbation of the generator of the linearized Navier-Stokes equations, together with directions of excitation sources, that can account for the observed statistics. This covariance completion problem is to determine minimum energy and low-rank perturbation of the linearized dynamics to reconcile them with the partially available second-order statistics – such models are valuable as tools for analysis and control purposes. The resulting optimization problem can be cast as a convex semidefinite program (SDP). However, general-purpose SDP solvers cannot handle typical problem-sizes that are of interest in fluid flows. We develop customized algorithms that allow handling such covariance completion problems for substantially larger scales. These algorithms exploit the structure of the problem and utilize the method of multipliers and the proximal augmented Lagrangian method.**

Index Terms— **Convex optimization, method of multipliers, proximal augmented Lagrangian, proximal methods, state covariances, structured matrix completion.**

I. INTRODUCTION

We are interested in *structured covariance completion problems*. These are of great value in fluid flow modeling and control. Indeed, it has been shown that stochastically-forced linearized Navier-Stokes (NS) equations around the mean velocity profile qualitatively replicate the structural features of shear flows [1]–[5]. This insight provides the basis and motivates the problem of determining suitable perturbations of linearized dynamics to reconcile data from measurements or numerical simulations and, thereby, identify consistent models for analysis and design.

A significant advance in utilizing stochastically-forced linearized NS equations for analysis and control, was achieved in recent years when the value of *nontrivial stochastic forcing* (i.e. colored-in-time noise as opposed to white) was recognized [6]–[10]. Indeed, it has been shown that colored-in-time noise can account for features of the flow field that white noise cannot. Interestingly, it has further been shown that colored-in-time excitation can be seen as *equivalent* to the effect of white noise *together* with a suitable perturbation of the system dynamics. In such perturbation, *no increase* in the state dimension is needed to generate time-correlations [9], [10]. Moreover, in a reverse direction, perturbation of the system dynamics reveals potentially important coupling between states in the form of dynamical interactions (state feedback) [10, Section 6.1].

Based on this last point, an optimal state-feedback synthesis problem was formulated to account for available statistical signatures in [11]. Parsimony in this approach dictates penalizing the size, directionality, and number of state-feedback couplings. Thus, it is natural to select, e.g., using suitable sparse/low-rank promoting functionals, a small subset of input "channels" through which suitable state-feedback can account for the available second-order statistics. To this end and to address the combinatorial complexity we present a convex formulation having roots in optimal sensor and actuator selection [12], [13]. The resulting convex optimization problem provides information about critical directions that have maximal effect in bringing model and statistics in agreement. More specifically, we formulate a minimum-control-energy covariance completion problem that can be cast as a semidefinite program (SDP). Similar problems have been considered by different complementing viewpoints, aiming at control of stochastic systems [14]–[17] and estimation of output covariances [18]–[21].

In the present paper, we exploit the problem structure and develop efficient customized algorithms for large-scale problems that generic SDP solvers cannot handle. Since the objective function in our optimization problem is nonsmooth, standard gradient-based methods are not applicable. The lack of differentiability can be addressed by splitting the smooth and nonsmooth components of the objective function over separate variables that are coupled via additional equality constraints. Indeed, this idea allows us to employ the recently developed proximal augmented Lagrangian method [22] as well as the well-known alternating direction method of multipliers [23]. We propose an alternative approach, which takes advantage of the structural constraint on the optimization variables. By expressing one variable in terms of the other, we remove one constraint and recast the problem into a form amenable to the standard proximal gradient method. Our numerical experiments demonstrate the efficacy and superior performance of our approach relative to splitting methods.

Our presentation is organized as follows. In Section II, we formulate the minimum energy covariance completion problem. In Section III, we present two customized algorithms for solving this optimization problem. In Section IV, we offer a motivating example and present results of numerical experiments. Finally, we provide concluding thoughts in Section V.

Financial support from the National Science Foundation under Award CMMI 1739243 and the Air Force Office of Scientific Research under Award FA9550-16-1-0009 is gratefully acknowledged.

A. Zare and M. R. Jovanović are with the Department of Electrical Engineering, University of Southern California, Los Angeles, CA 90089. N. K. Dhingra is with the Department of Electrical and Computer Engineering, University of Minnesota, Minneapolis, MN 55455. T. T. Georgiou is with the Department of Mechanical and Aerospace Engineering, University of California, Irvine, CA 92697. E-mails: armin.zare@usc.edu, dhin0008@umn.edu, mihailo@usc.edu, tryphon@uci.edu.

II. Problem formulation

Consider the linear time-invariant (LTI) system

$$\begin{aligned} \dot{x} &= A\,x \,+\, B_1\,u \,+\, B_2\,d \\ y &= Cx \end{aligned} \tag{1}$$

where $x(t) \in \mathbb{C}^n$ is the state vector, $y(t) \in \mathbb{C}^p$ is the output, $u(t) \in \mathbb{C}^m$ is the control input, $d(t)$ is a stationary zero-mean stochastic process, B_1 and $B_2 \in \mathbb{C}^{n\times m}$ are the input matrices with $m \leq n$, $C \in \mathbb{C}^{p\times n}$ is the output matrix, and (A, B_2) is a controllable pair.

We are interested in the setup where matrix A in (1) is known but due to experimental or numerical limitations, only partial correlations between a limited number of state components are available. Moreover, it is often the case that the origin and directionality of excitation sources that generate the statistics is unknown. It is desired to design an optimal feedback control law, $u = -Kx$, such that the closed loop system given by

$$\dot{x} \;=\; (A \,-\, B_2\,K)\,x \,+\, B_1\,d \tag{2}$$

accounts for the partially available statistics. On the other hand, the steady-state covariance matrix of the state

$$X \;:=\; \lim_{t\,\to\,\infty} \mathbf{E}\left(x(t)\,x^*(t)\right),$$

satisfies the following rank condition [24]:

$$\text{rank}\begin{bmatrix} A\,X \,+\, X\,A^* & B_2 \\ B_2^* & 0 \end{bmatrix} \;=\; \text{rank}\begin{bmatrix} 0 & B_2 \\ B_2^* & 0 \end{bmatrix}.$$

This implies that any $X \succ 0$ is admissible as a covariance of the LTI system (1) if the input matrix B_2 is full row rank [24], which eliminates the role of the dynamics inherent in A. It is thus important to limit the rank of the input matrix B_2 and thereby the number of input channels that can perturb the dynamical generator A in (2) [9], [10].

Based on this, the regularized minimum energy covariance completion problem

$$\begin{array}{ll} \underset{K,\,X}{\text{minimize}} & \text{trace}\,(K^* X\,K) \,+\, \gamma \displaystyle\sum_{i\,=\,1}^{n} w_i\,\|\mathrm{e}_i^*\,K\|_2 \\ \text{subject to} & (A - B_2K)X \,+\, X(A - B_2K)^* + V \;=\; 0 \\ & (CXC^*) \circ E \,-\, G \;=\; 0 \\ & X \,\succ\, 0, \end{array} \tag{3}$$

was proposed in [11] to obtain an optimal feedback gain K that minimizes the control energy in statistical steady-state and at the same time account for partially known statistics. In this problem, the performance index $\text{trace}\,(KXK^*)$ is augmented with a term that promotes row-sparsity of the feedback gain matrix K, $V \,:=\, B_1\,\Omega\,B_1^*$ with Ω as the covariance of the white noise process d, γ is a positive regularization parameter, w_i are nonzero weights, and e_i is the ith unit vector in $\mathbb{R}^m$. Matrices A, B_1, B_2, C, E, G, and Ω are problem data, and hermitian matrix $X \in \mathbb{C}^{n\times n}$ and feedback gain matrix $K \in \mathbb{C}^{m\times n}$ are optimization variables. Entries of matrix G represent partially available second-order statistics of the output y, the symbol $\circ$ denotes elementwise matrix multiplication, and E is the structural identity matrix,

$$E_{ij} \;=\; \begin{cases} 1, & \text{if } G_{ij} \text{ is available} \\ 0, & \text{if } G_{ij} \text{ is unavailable.} \end{cases}$$

When the ith row of K is identically equal to zero, the ith input channel in the matrix B_2 is not used. Therefore, problem (3) identifies a subset of critical input channels by promoting row-sparsity of K, thereby uncovering the precise dynamical feedback interactions that are required to reconcile the available covariance data with the given linear dynamics.

Since X is positive definite, the standard change of variables $Y := KX$ and the equivalence between the row-sparsity of K and the row-sparsity of Y [12] can be utilized to bring problem (3) into the following form

$$\begin{array}{ll} \underset{X,\,Y}{\text{minimize}} & \text{trace}\left(YX^{-1}Y^*\right) \,+\, \gamma \displaystyle\sum_{i\,=\,1}^{n} w_i\,\|\mathrm{e}_i^*\,Y\|_2 \\ \text{subject to} & A\,X \,+\, X\,A^* \,-\, B\,Y \,-\, Y^*B^* \,+\, V \;=\; 0 \\ & (CXC^*) \circ E \,-\, G \;=\; 0 \\ & X \,\succ\, 0, \end{array} \tag{CC}$$

which is SDP representable [25]. Finally, the optimal feedback gain matrix can be recovered as $K = YX^{-1}$. The convexity of (CC) follows from the convexity of its objective function and the convexity of the constraint set [26]. To solve problem (CC) for sizes that general-purpose SDP solvers cannot, we next develop customized algorithms based on the associated augmented Lagrangian.

III. Customized algorithms

We present two customized algorithms based on the method of multipliers (MM) and the recently developed Proximal Augmented Lagrangian (PAL) method [22]. The method of multipliers is widely used for solving constrained nonlinear programming problems [27]–[29]. Similar to the Alternating Direction Method of Multipliers (ADMM) [23], PAL utilizes the problem structure to split smooth and nonsmooth parts of the objective function over separate variables that are coupled via additional equality constraints. Relative to ADMM, PAL offers systematic update rules for the augmented Lagrangian parameter ρ which facilitates improved practical performance.

For notational compactness, we write the linear constraints in (CC) as

$$\begin{aligned} \mathcal{A}_1(X) \,-\, \mathcal{B}(Y) + V &= 0 \\ \mathcal{A}_2(X) \,-\, G &= 0 \end{aligned}$$

with linear operators $\mathcal{A}_1$: $\mathbb{C}^{n\times n} \to \mathbb{C}^{n\times n}$, $\mathcal{A}_2$: $\mathbb{C}^{n\times n} \to \mathbb{C}^{p\times p}$ and $\mathcal{B}$: $\mathbb{C}^{m\times n} \to \mathbb{C}^{n\times n}$ defined as

$$\begin{aligned} \mathcal{A}_1(X) &:= A\,X \,+\, X\,A^*, \\ \mathcal{A}_2(X) &:= (CXC^*) \circ E, \\ \mathcal{B}(Y) &:= B_2\,Y + Y^*B_2^*. \end{aligned}$$

A. Elimination of variable X

For any Y, there is a unique X that solves the equation

$$\mathcal{A}_1(X) \,-\, \mathcal{B}(Y) + V \;=\; 0$$

if and only if $\lambda_i + \bar{\lambda}_j \neq 0$ for all i and j where $\lambda_1, \lambda_2, \ldots, \lambda_n$ are eigenvalues of the matrix A [30]. Note that a stabilizing Y always has a unique solution X. Herein, we assume that $\lambda_i + \bar{\lambda}_j \neq 0$ which allows us to express the variable X as a linear function of Y,

$$X(Y) \;=\; \mathcal{A}_1^{-1}\left(\mathcal{B}(Y) \,-\, V\right), \tag{4}$$

and restate problem (CC) as

$$\begin{array}{ll} \underset{Y}{\text{minimize}} & f(Y) \,+\, \gamma\, g(Y) \\ \text{subject to} & \mathcal{A}_2(X(Y)) \,-\, G \;=\; 0 \\ & X(Y) \,\succ\, 0, \end{array} \tag{CC1}$$

where

$$\begin{aligned} f(Y) &\;:=\; \operatorname{trace}\left(Y X^{-1}(Y) Y^*\right) \\ g(Y) &\;:=\; \sum_{i\,=\,1}^{n} w_i \, \|\mathrm{e}_i^*\, Y\|_2. \end{aligned}$$

Since (CC1) is equivalent to (CC) constrained to the linear subspace defined by (4), it remains a convex problem.

B. Method of multipliers

The method of multipliers for the constrained optimization problem (CC1) is given by

$$Y^{k+1} \;:=\; \underset{Y}{\operatorname{argmin}}\; \mathcal{L}_\rho\left(Y;\, \Lambda^k\right) \tag{5a}$$

$$\Lambda^{k+1} \;:=\; \Lambda^k \,+\, \rho\left(\mathcal{A}_2(X(Y^{k+1})) \,-\, G\right), \tag{5b}$$

where $\mathcal{L}_\rho$ is the augmented Lagrangian,

$$\begin{aligned} \mathcal{L}_\rho(Y;\Lambda) \;=\; & f(Y) \,+\, \gamma\, g(Y) \,+\, \\ & \langle \Lambda, \mathcal{A}_2(X(Y)) \,-\, G\rangle \,+\, \frac{\rho}{2}\,\|\mathcal{A}_2(X(Y)) \,-\, G\|_F^2, \end{aligned}$$

$\Lambda \in \mathbb{C}^{p\times p}$ is the Lagrange multiplier, ρ is a positive scalar, $\langle\cdot,\cdot\rangle$ is the standard inner product between two matrices, and $\|\cdot\|_F$ is the Frobenius norm. The algorithm terminates when the primal and dual residuals are small enough

$$\|\mathcal{A}_2(X(Y^{k+1})) \,-\, G\|_F \;\leq\; \epsilon_1 \tag{6a}$$

$$\|Y^{k+1} \,-\, Y^k\|_F \;\leq\; \epsilon_2. \tag{6b}$$

1) Solution to the Y-minimization problem (5a)*:* For fixed Λ^k, minimizing the augmented Lagrangian with respect to Y amounts to finding the minimizer of

$$\begin{aligned} \mathcal{L}_\rho(Y;\Lambda^k) \;=\; & f(Y) \,+\, \gamma\, g(Y) \,+\, \\ & \left\langle \Lambda^k, \mathcal{A}_2(X(Y)) \,-\, G\right\rangle \,+\, \frac{\rho}{2}\,\|\mathcal{A}_2(X(Y)) \,-\, G\|_F^2. \end{aligned} \tag{7}$$

Since $g(Y)$ is nonsmooth, we cannot use standard gradient descent methods to find the update Y^{k+1}. This subproblem can be iteratively solved using the proximal gradient method

$$Y^{j+1} \;=\; \mathbf{prox}_{\beta g}\left(Y^j \,+\, \alpha^j\, \nabla F(Y^j)\right),$$

where α^j is the step-size, $\beta = \alpha^j \gamma$, j denotes the inner proximal gradient iterations, $\mathbf{prox}_{\beta g}(\cdot)$ is the proximal operator associated with the function g, and $F(Y)$ denotes the smooth part of (7),

$$\begin{aligned} F(Y) \;:=\; f(Y) \,+\, \left\langle \Lambda^k, \mathcal{A}_2(X(Y)) \,-\, G\right\rangle \,+\, \\ \frac{\rho}{2}\,\|\mathcal{A}_2(X(Y)) \,-\, G\|_F^2. \end{aligned}$$

The expression for the gradient of $F(Y)$ is provided in the Appendix. The proximal operator of function g is defined as

$$\mathbf{prox}_{\tau g}(V) \;:=\; \underset{Y}{\operatorname{argmin}}\; g(Y) \,+\, \frac{1}{2\tau}\,\|Y \,-\, V\|_F^2,$$

and is determined by the soft-thresholding operator which acts on the rows of matrix V,

$$\mathcal{S}_\tau(\mathrm{e}_i^* V) \;=\; \begin{cases} (1 \,-\, \tau/\|\mathrm{e}_i^* V\|_2)\,\mathrm{e}_i^* V, & \|\mathrm{e}_i^* V\|_2 \,>\, \tau \\ 0, & \|\mathrm{e}_i^* V\|_2 \,\leq\, \tau. \end{cases}$$

At each iteration, we determine the step-size α^j via an adaptive Barzilai-Borwein step-size selection [31] to ensure sufficient descent of (7) and positive definiteness of $X(Y^{j+1})$; see [32, Theorem 3.1] for details.

2) Lagrange multiplier update and choice of step-size ρ in (5b)*:* We follow the procedure outlined in [29, Algorithm 17.4] for the adaptive update of the step-size ρ that is used to update the dual variable Λ. This procedure allows for inexact solutions of the subproblem (5a) and a more refined update of the Lagrange multiplier Λ through the adjustment of the convergence tolerances ϵ_1 and ϵ_2 in (6) and step-size ρ.

C. Proximal augmented Lagrangian method

We next employ the method developed in [22, Algorithm 1] to solve the constrained optimization problem (CC1).

By introducing an additional optimization variable Z, (CC1) can be written as

$$\begin{array}{ll} \underset{Y,\,Z}{\text{minimize}} & f(Y) \,+\, \gamma\, g(Z) \\ \text{subject to} & \mathcal{A}_2(X(Y)) \,-\, G \;=\; 0 \\ & Y \,-\, Z \;=\; 0 \\ & X(Y) \,\succ\, 0, \end{array} \tag{CC2}$$

The augmented Lagrangian associated with (CC2) is

$$\begin{aligned} \mathcal{L}_\rho(Y,Z;\Lambda_1,\Lambda_2) \;=\; & f(Y) \,+\, \gamma\, g(Z) \,+\, \\ & \langle \Lambda_1, \mathcal{A}_2(X(Y)) \,-\, G\rangle \,+\, \langle \Lambda_2, Y \,-\, Z\rangle \,+\, \\ & \frac{\rho}{2}\,\|\mathcal{A}_2(X(Y)) \,-\, G\|_F^2 \,+\, \frac{\rho}{2}\,\|Y \,-\, Z\|_F^2, \end{aligned}$$

where $\Lambda_1 \in \mathbb{C}^{p\times p}$ and $\Lambda_2 \in \mathbb{C}^{m\times n}$ are Lagrange multipliers. Through completion of squares and minimization with respect to Z we achieve an explicit expression for $Z^\star$,

$$Z^\star \;=\; \mathbf{prox}_{\beta g}\left(Y \,+\, \frac{1}{\rho}\Lambda_2\right),$$

where $\beta \,=\, \gamma/\rho$. Substitution of $Z^\star$ into the augmented Lagrangian yields the proximal augmented Lagrangian as

$$\begin{aligned} \mathcal{L}_\rho(Y;\Lambda_1,\Lambda_2) \;:=\; & \mathcal{L}_\rho(Y,Z^\star;\Lambda_1,\Lambda_2) \\ \;=\; & f(Y) \,+\, \langle \Lambda_1, \mathcal{A}_2(X(Y)) \,-\, G\rangle \,+\, \\ \frac{\rho}{2}\,\|\mathcal{A}_2(X(Y)) - G\|_F^2 \,+\, & M_{\beta g}(Y + \frac{1}{\rho}\Lambda_2) \,-\, \frac{1}{2\rho}\|\Lambda_2\|_F^2, \end{aligned} \tag{8}$$

which characterizes the augmented Lagrangian on the manifold corresponding to explicit minimization over variable Z. Here, $M_{\beta g}$ is the Moreau envelope associated with the proximal operator of function g [33], i.e.,

$$M_{\beta g}(V) \;:=\; \inf_{Z} \; \gamma\, g(Z) \;+\; \frac{\rho}{2}\, \|Z \,-\, V\|_F^2.$$

The Moreau envelope is a continuously differentiable function, even though g is not, and its gradient is given by

$$\nabla M_{\beta g}(V) \;=\; \rho\,(V \,-\, \mathbf{prox}_{\beta g}(V)).$$

The proximal augmented Lagrangian method solves problem (CC2) via a sequence of iterations in which $\mathcal{L}_\rho$ is minimized over the primal variable Y and maximized over the dual variables Λ_1 and Λ_2 with step-size ρ,

$$Y^{k+1} \;:=\; \underset{Y}{\operatorname{argmin}}\; \mathcal{L}_\rho\,(Y;\, \Lambda_1^k,\, \Lambda_2^k) \tag{9a}$$

$$\Lambda_1^{k+1} \;:=\; \Lambda_1^k \;+\; \rho\, \nabla_{\Lambda_1} \mathcal{L}_\rho(Y^{k+1}; \Lambda_1^k, \Lambda_2^k) \tag{9b}$$

$$\Lambda_2^{k+1} \;:=\; \Lambda_2^k \;+\; \rho\, \nabla_{\Lambda_2} \mathcal{L}_\rho(Y^{k+1}; \Lambda_1^k, \Lambda_2^k) \tag{9c}$$

where

$$\begin{aligned} \nabla_{\Lambda_1} \mathcal{L}_\rho &\;=\; \mathcal{A}_2(X(Y)) \,-\, G \\ \nabla_{\Lambda_2} \mathcal{L}_\rho &\;=\; Y^{k+1} \,-\, \mathbf{prox}_{\beta g}(Y^{k+1} \,+\, \tfrac{1}{\rho}\Lambda_2). \end{aligned}$$

Updates to the Lagrange multipliers $\{\Lambda_1^k, \Lambda_2^k\}$, and step-size ρ follow a similar adaptive procedure and termination criteria as Section III-B.2; see [22] for additional details.

Solution of the Y-minimization problem (9a)*:* Since (8) is once continuously differentiable, gradient-based methods, e.g., gradient descent, in conjunction with backtracking step-size selection can be used to find the solution to problem (9a). Herein, we employ the limited-memory Broyden-Fletcher-Goldfarb-Shanno (L-BFGS) method highlighted in [29, Algorithm 7.4] to estimate the Hessian inverse H_k and to compute the descent direction as $r \,=\, -H_k(\nabla \mathcal{L}_\rho)$. Our experiments show that this approach results in satisfactory performance for modest values of m (between 20-30).

For convex functions with Lipschitz continuous gradients, BFGS is guaranteed to converge [34]. Although L-BFGS requires *strong* convexity for guaranteed convergence [35], we find that it works well in practice.

D. Computational complexity

Computation of the gradient in the Y-minimization step of both MM and PAL involves computation of X from Y based on (4), a matrix inversion, and solution to the Lyapunov equation, which each take $O(n^3)$ operations, as well as an $O(mn^2)$ matrix-matrix multiplication. The proximal operator for the function g amounts to computing the 2-norm of all m rows of a matrix with n columns, which takes $O(mn)$ operations. These steps are embedded within an iterative backtracking procedure for selecting the step-size α. If the step-size selection takes q_1 inner iterations and it takes q_2 iterations for the gradient based method to converge, the total computation cost for a single iteration of MM or PAL is $O(q_1 q_2 n^3)$. In contrast, the worst-case complexity of standard SDP solvers is $O(n^6)$.

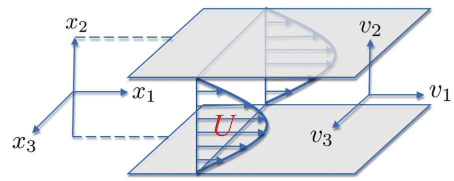

Fig. 1. Geometry of a three-dimensional pressure-driven channel flow.

E. Comparison with ADMM

In contrast to ADMM, the optimization algorithms considered in this paper do not handle the lack of differentiability of function g by fixing Z to minimize over Y. The customized method of multipliers algorithm in Section III-B uses the proximal gradient method to solve the nonsmooth problem in the Y-minimization step, and the proximal augmented Lagrangian method in Section III-C guarantees continuous differentiability of the proximal augmented Lagrangian $\mathcal{L}_\rho(Y; \Lambda_1, \Lambda_2)$ by constraining $\mathcal{L}_\rho(Y, Z; \Lambda_1, \Lambda_2)$ to the manifold resulting from explicit minimization over Z. Moreover, it is important to note that the rate of convergence in augmented Lagrangian-based methods is strongly influenced by the choice of ρ. While both MM and PAL follow adaptive rules for updating the step-size ρ in each iteration, ADMM does not have efficient step-size selection rules. Typically, in ADMM, either a constant step-size is selected or the step-size is adjusted to keep the norms of primal and dual residuals within a constant factor of one another [23].

IV. An example

In an incompressible channel-flow, we consider the dynamics of infinitesimal fluctuations around the parabolic mean velocity profile, $\bar{\mathbf{u}} \,=\, [\,U(x_2) \;\; 0 \;\; 0\,]^T$ with $U(x_2) \,=\, 1 \,-\, x_2^2$; see Fig. 1 for geometry. The streamwise, wall-normal and spanwise coordinates are represented by x_1, x_2, and x_3, respectively. Finite dimensional approximation of the differential operators in the linearized NS equations around $\bar{\mathbf{u}}$ results in the following state-space representation

$$\begin{aligned} \dot{x}(\mathbf{k},\, t) &\;=\; A(\mathbf{k})\, x(\mathbf{k},\, t) \;+\; \xi(\mathbf{k},\, t), \\ y(\mathbf{k},\, t) &\;=\; C(\mathbf{k})\, x(\mathbf{k},\, t). \end{aligned} \tag{10a}$$

Here, $x \,=\, [\, v_2^T \;\; \eta^T \,]^T \,\in\, \mathbb{C}^{2N}$ is the state, v_2 and $\eta \,=\, \partial_{x_3} v_1 - \partial_{x_1} v_3$ are the wall-normal velocity and vorticity, the output $y \,=\, [\, v_1^T \;\; v_2^T \;\; v_3^T \,]^T \,\in\, \mathbb{C}^{3N}$ denotes the fluctuating velocity vector, ξ is a stochastic forcing disturbance, $\mathbf{k} \,=\, [\, k_1 \;\; k_3 \,]^T$ denotes the vector of horizontal wavenumbers, and the input matrix is the identity $I_{2N \times 2N}$. The dynamical matrix $A \in \mathbb{C}^{2N \times 2N}$ and output matrix $C \in \mathbb{C}^{3N \times 2N}$ are described in [4]. We assume that the stochastic disturbance ξ is generated by the low-pass filter

$$\dot{\xi}(\mathbf{k}, t) \;=\; -\xi(\mathbf{k}, t) \;+\; w(t). \tag{10b}$$

Here, w denotes a zero mean unit variance white process.

The solution to the Lyapunov equation

$$\tilde{A}\,\Sigma \;+\; \Sigma\, \tilde{A}^* \;+\; \tilde{B}\, \tilde{B}^* \;=\; 0$$

TABLE I

COMPARISON OF DIFFERENT ALGORITHMS (IN SECONDS) FOR DIFFERENT NUMBER OF DISCRETIZATION POINTS N AND $\gamma = 10$.

N	CVX	MM	PAL	PAL + L-BFGS	ADMM
11	9.3	0.19	10.02	14.06	3.10
21	97.67	5.6	127.28	101.16	113.4
31	899.99	7.19	321.03	295.24	574.44
51	-	34.76	–	4225	–
101	-	146.51	–	–	–

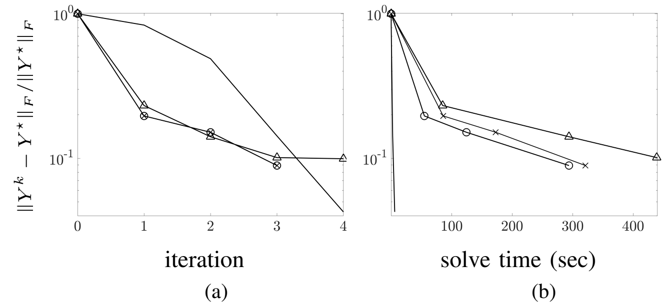

Fig. 2. Convergence curves showing performance of MM (−), PAL (×), PAL + L-BFGS (○), and ADMM (△) versus (a) the number of outer iterations; and (b) solve times for $N = 31$ collocation points in the wall-normal direction x_2 and $\gamma = 10$. Here, $Y^\star$ is the optimal value for Y.

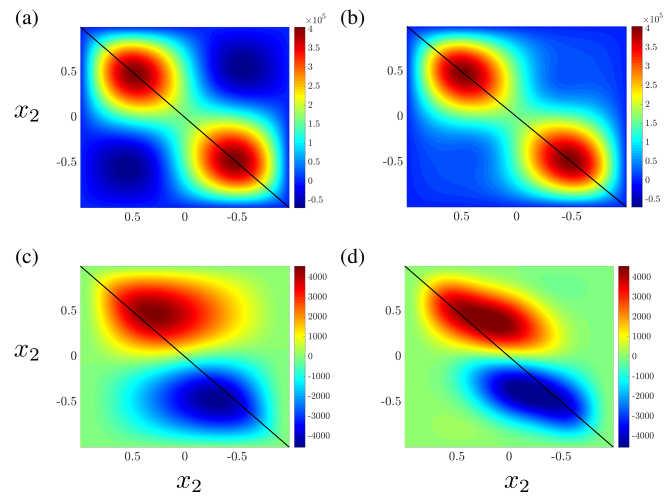

Fig. 3. True covariance matrices of the output velocity field (a, c), and covariance matrices resulting from solving problem (CC) (b, d) with $\gamma = 10$ and $N = 51$. (a, b) Streamwise Φ_{11}, and (c, d) streamwise/wall-normal Φ_{12} two-point correlation matrices at $\mathbf{k} = (0, 1)$. One-point correlation profiles that are used as problem data are marked along the main diagonals.

represents the steady-state covariance of system (10). Here,

$$\tilde{A} = \begin{bmatrix} A & I \\ O & -I \end{bmatrix}, \quad \tilde{B} = \begin{bmatrix} 0 \\ I \end{bmatrix},$$

and

$$\Sigma = \begin{bmatrix} \Sigma_{xx} & \Sigma_{x\xi} \\ \Sigma_{\xi x} & \Sigma_{\xi\xi} \end{bmatrix}.$$

The sub-covariance Σ_{xx} denotes the state covariance of system (10a). At any horizontal wavenumber pair $\mathbf{k}$, the steady-state covariance matrices of the output y and the state x are related by

$$\Phi(\mathbf{k}) = C(\mathbf{k}) \, \Sigma_{xx}(\mathbf{k}) \, C^*(\mathbf{k}).$$

In this example, we set the covariance of white noise disturbances to the identity ($\Omega = I$) and assume knowledge of one-point velocity correlations, or diagonal entries of the streamwise Φ_{11}, wall-normal Φ_{22}, spanwise Φ_{33}, and the streamwise/wall-normal Φ_{12} two-point correlation matrices. To account for the available statistics, we solve (CC) for a state covariance X which agrees with the available statistics.

Numerical experiments were conducted for a channel flow with Reynolds number $Re = 10^3$, $(k_1, k_3) = (0, 1)$, for various number of collocation points N in the wall-normal direction x_2 (state dimension $n = 2N$), and for various values of the regularization parameter γ. We initialize our algorithms with $Y^0 = K_c X_c$, where K_c and X_c solve the algebraic Riccati equation which specifies the optimal centralized controller. This guarantees that $X(Y^0) \succ 0$. Iterations were run for each method until the primal and dual residuals satisfy a certain tolerance; cf. (6). For $\epsilon_1, \epsilon_2 = 10^{-2}$ and $\gamma = 10$, Table I compares various methods based on run times (sec). For $N = 51$ and 101, CVX failed to converge and PAL and ADMM did not converge in a resonable time. Clearly, MM outperforms PAL and ADMM for large problems. This can also be deduced from Fig. 2, which shows convergence curves of MM and PAL for $N = 31$ and $\gamma = 10$.

Figures 3(b,d) show the streamwise, and the streamwise/wall-normal two-point correlation matrices resulting from solving (CC) with $\gamma = 10$. Even though only one-point velocity correlations along the main diagonal of these matrices were used as problem data in (CC), we observe reasonable recovery of off-diagonal terms of the full two-point velocity correlation matrices.

V. CONCLUDING REMARKS

We have examined a class of covariance completion problems in which the problem of accounting for partially known second-order statistics via stochastically forced linear systems is formulated as a state-feedback synthesis problem. Parsimony in the arising optimization problems is dictated by penalizing the control effort, in addition to the number of control input channels, which is achieved by promoting row-sparsity of the feedback gain matrix. To efficiently solve covariance completion problems of large size we develop customized algorithms using augmented Lagrangian-based methods. Our algorithms utilize the Lyapunov-like constraint to express one variable in terms of the other. In addition, proximal methods are used to handle the lack of differentiability of the objective function. Numerical experiments show that the method of multipliers, which utilizes the proximal gradient method to solve its subproblem, performs much better than the splitting methods considered in this paper. In ongoing work, we will investigate solving this problem via recently developed second order methods [36], [37] and quasi-Newton methods [35] for nonsmooth composite problems.

In this work, row sparsity is promoted by penalizing a weighted sum of row norms of the feedback gain matrix. While we note that iterative reweighting [38] can improve the row-sparsity patterns determined by this convex approximation of cardinality, the efficacy of more refined approximations, namely low-rank inducing norms [39], [40], for which proximal operators can be efficiently computed, is a subject of future research.

APPENDIX

At the kth iteration the gradient of F with respect to Y is given by,

$$\nabla F(Y) \;=\; 2\,Y^k X^{-1} \;-\; 2\,B^*(W_2 \,+\, \rho\,W_3 \,-\, W_1),$$

where W_1, W_2, and W_3 are solutions to the following Lyapunov equations

$$\begin{aligned} A^* W_1 \;+\; W_1 A \;+\; X^{-1} Y^{k*} Y^k X^{-1} \;&=\; 0 \\ A^* W_2 \;+\; W_2 A \;+\; \mathcal{A}_2^\dagger \left(\Lambda^k\right) \;&=\; 0 \\ A^* W_3 \;+\; W_3 A \;+\; \mathcal{A}_2^\dagger \left(\mathcal{A}_2 \left(X(Y^k)\right) \,-\, G\right) \;&=\; 0 \end{aligned}$$

Here, X^{-1} denotes the inverse of $X(Y^k)$ and the adjoint of the operator $\mathcal{A}_2$ is given by

$$\mathcal{A}_2^\dagger \left(\Lambda\right) \;:=\; C^* \left(E \circ \Lambda\right) C.$$

REFERENCES

[1] B. Bamieh and M. Dahleh, "Energy amplification in channel flows with stochastic excitation," *Phys. Fluids*, vol. 13, no. 11, pp. 3258–3269, 2001.

[2] M. R. Jovanović and B. Bamieh, "The spatio-temporal impulse response of the linearized Navier-Stokes equations," in *Proceedings of the 2001 American Control Conference*, 2001, pp. 1948–1953.

[3] M. R. Jovanović and B. Bamieh, "Modelling flow statistics using the linearized Navier-Stokes equations," in *Proceedings of the 40th IEEE Conference on Decision and Control*, 2001, pp. 4944–4949.

[4] M. R. Jovanović and B. Bamieh, "Componentwise energy amplification in channel flows," *J. Fluid Mech.*, vol. 534, pp. 145–183, July 2005.

[5] R. Moarref and M. R. Jovanović, "Model-based design of transverse wall oscillations for turbulent drag reduction," *J. Fluid Mech.*, vol. 707, pp. 205–240, September 2012.

[6] Y. Chen, M. R. Jovanović, and T. T. Georgiou, "State covariances and the matrix completion problem," in *Proceedings of the 52nd IEEE Conference on Decision and Control*, 2013, pp. 1702–1707.

[7] A. Zare, M. R. Jovanović, and T. T. Georgiou, "Completion of partially known turbulent flow statistics," in *Proceedings of the 2014 American Control Conference*, 2014, pp. 1680–1685.

[8] A. Zare, M. R. Jovanović, and T. T. Georgiou, "Alternating direction optimization algorithms for covariance completion problems," in *Proceedings of the 2015 American Control Conference*, 2015, pp. 515–520.

[9] A. Zare, Y. Chen, M. R. Jovanović, and T. T. Georgiou, "Low-complexity modeling of partially available second-order statistics: theory and an efficient matrix completion algorithm," *IEEE Trans. Automat. Control*, vol. 62, no. 3, pp. 1368–1383, March 2017.

[10] A. Zare, M. R. Jovanović, and T. T. Georgiou, "Colour of turbulence," *J. Fluid Mech.*, vol. 812, pp. 636–680, February 2017.

[11] A. Zare, M. R. Jovanović, and T. T. Georgiou, "Perturbation of system dynamics and the covariance completion problem," in *Proceedings of the 55th IEEE Conference on Decision and Control*, 2016, pp. 7036–7041.

[12] B. Polyak, M. Khlebnikov, and P. Shcherbakov, "An LMI approach to structured sparse feedback design in linear control systems," in *Proceedings of the 2013 European Control Conference*, 2013, pp. 833–838.

[13] N. K. Dhingra, M. R. Jovanović, and Z. Q. Luo, "An ADMM algorithm for optimal sensor and actuator selection," in *Proceedings of the 53rd IEEE Conference on Decision and Control*, 2014, pp. 4039–4044.

[14] A. Hotz and R. E. Skelton, "Covariance control theory," *Int. J. Control*, vol. 46, no. 1, pp. 13–32, 1987.

[15] K. Yasuda, R. E. Skelton, and K. M. Grigoriadis, "Covariance controllers: A new parametrization of the class of all stabilizing controllers," *Automatica*, vol. 29, no. 3, pp. 785–788, 1993.

[16] K. M. Grigoriadis and R. E. Skelton, "Alternating convex projection methods for covariance control design," *Int. J. Control*, vol. 60, no. 6, pp. 1083–1106, 1994.

[17] Y. Chen, T. T. Georgiou, and M. Pavon, "Optimal steering of a linear stochastic system to a final probability distribution, Part II," *IEEE Trans. Automat. Control*, 2015.

[18] F. Lin and M. R. Jovanović, "Least-squares approximation of structured covariances," *IEEE Trans. Automat. Control*, vol. 54, no. 7, pp. 1643–1648, July 2009.

[19] A. Ferrante, M. Pavon, and M. Zorzi, "A maximum entropy enhancement for a family of high-resolution spectral estimators," *IEEE Trans. Automat. Control*, vol. 57, no. 2, pp. 318–329, 2012.

[20] M. Zorzi and A. Ferrante, "On the estimation of structured covariance matrices," *Automatica*, vol. 48, no. 9, pp. 2145–2151, 2012.

[21] M. Zorzi, "Multivariate Spectral Estimation Based on the Concept of Optimal Prediction," *IEEE Trans. Automat. Control*, vol. 60, no. 6, pp. 1647–1652, 2015.

[22] N. K. Dhingra, S. Z. Khong, and M. R. Jovanović, "The proximal augmented Lagrangian method for nonsmooth composite optimization," *IEEE Trans. Automat. Control*, 2016, submitted; also arXiv:1610.04514.

[23] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Found. Trends Mach. Learn.*, vol. 3, no. 1, pp. 1–122, 2011.

[24] T. T. Georgiou, "The structure of state covariances and its relation to the power spectrum of the input," *IEEE Trans. Autom. Control*, vol. 47, no. 7, pp. 1056–1066, 2002.

[25] G. E. Dullerud and F. Paganini, *A course in robust control theory: a convex approach*. New York: Springer-Verlag, 2000.

[26] S. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge University Press, 2004.

[27] D. P. Bertsekas, *Constrained optimization and Lagrange multiplier methods*. New York: Academic Press, 1982.

[28] D. P. Bertsekas, *Nonlinear programming*. Belmont, MA: Athena Scientific, 1999.

[29] J. Nocedal and S. J. Wright, *Numerical Optimization*. Springer, 2006.

[30] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge University Press, 2012.

[31] J. Barzilai and J. M. Borwein, "Two-point step size gradient methods," *IMA J. Numer. Anal.*, vol. 8, no. 1, pp. 141–148, 1988.

[32] A. Beck and M. Teboulle, "A fast iterative shrinkage-thresholding algorithm for linear inverse problems," *SIAM J. Imaging Sci.*, vol. 2, no. 1, pp. 183–202, 2009.

[33] N. Parikh and S. Boyd, "Proximal algorithms," *Found. Trends Optim.*, vol. 1, no. 3, pp. 123–231, 2013.

[34] J.-F. Bonnans, J. C. Gilbert, C. Lemaréchal, and C. A. Sagastizábal, *Numerical optimization: theoretical and practical aspects*. Springer Science & Business Media, 2013.

[35] L. Stella, A. Themelis, and P. Patrinos, "Forward-backward quasi-Newton methods for nonsmooth optimization problems," *Comput. Optim. Appl.*, vol. 67, no. 3, pp. 443–487, 2017.

[36] J. D. Lee, Y. Sun, and M. A. Saunders, "Proximal Newton-type methods for minimizing composite functions," *SIAM J. Optim.*, vol. 24, no. 3, pp. 1420–1443, 2014.

[37] N. K. Dhingra, S. Z. Khong, and M. R. Jovanović, "A second order primal-dual method for nonsmooth convex composite optimization," *IEEE Trans. Automat. Control*, 2017, submitted; also arXiv:1709.01610.

[38] E. J. Candes, M. B. Wakin, and S. P. Boyd, "Enhancing sparsity by reweighted ℓ_1 minimization," *J. Fourier Anal. Appl.*, vol. 14, no. 5-6, pp. 877–905, 2008.

[39] C. Grussler, A. Zare, M. R. Jovanović, and A. Rantzer, "The use of the $r*$ heuristic in covariance completion problems," in *Proceedings of the 55th IEEE Conference on Decision and Control*, 2016, pp. 1978–1983.

[40] C. Grussler, "Rank reduction with convex constraints," Ph.D. dissertation, Lund University, 2017.