

Comment on “Convex Approach to Covariance Control with Application to Stochastic Low-Thrust Trajectory Optimization”

George Rapakoulias* and Panagiotis Tsiotras †

In the recent paper “Convex Approach to Covariance Control with Application to Stochastic Low-Thrust Trajectory Optimization” by B. Benedikter et al [1] the authors proposed a convex reformulation for the numerical solution to the non-convex discrete-time covariance steering problem in finite time via state feedback control. The main source of non-convexity in the state-feedback covariance steering problem stems from the covariance propagation equation constraint

$$P_{j+1} = A_j P_j A_j^\top + A_j Y_j^\top B_j^\top + B_j Y_j A_j^\top + B_j Y_j P_j^{-1} Y_j^\top B_j^\top + Q_j \quad (1)$$

which is not convex in terms of the optimization variables P_{j+1} , P_j and Y_j . The idea is to relax this non-convex constraint showing up in the optimization problem (B2)* with a convex *inequality* constraint of the form

$$P_{j+1} \geq A_j P_j A_j^\top + A_j Y_j^\top B_j^\top + B_j Y_j A_j^\top + B_j Y_j P_j^{-1} Y_j^\top B_j^\top + Q_j \quad (\text{B2b})$$

In [1] it is claimed by the authors that this convex relaxation is lossless, that is, the optimal solution of the relaxed convex problem is the same as the original non-convex problem. We identify two issues in their proof that invalidate the main conclusion of the paper. We also provide a counterexample to illustrate the fact that the convexification is indeed lossy. We also point to a recent work [2] that achieves lossless convexification of the exact state covariance steering problem bypassing the technical issues of [1].

Specifically, in Appendix B of [1] the authors use a contradiction argument to prove that the optimal solution of problem (B2) satisfies (B2b) with equality, and therefore problems (B1) and (B2) are equivalent.

The proof is based on the complementary slackness condition (B8)

$$\Omega_j^{(i)} \begin{cases} \geq 0 & \text{if } G_j^{(i)} = 0 \\ = 0 & \text{if } G_j^{(i)} < 0 \end{cases} \quad (\text{B8})$$

*Ph.D. student, School of Aerospace Engineering, Georgia Institute of Technology, Atlanta, GA, 30332, USA.

†David and Andrew Lewis Chair and Professor, School of Aerospace Engineering, and Institute for Robotics and Intelligent Machines, Georgia Institute of Technology, Atlanta, GA, 30332, USA.

*Please refer to [1] for the original equation numbering.

where

$$G_j^{(1)} = Y_j P_j^{-1} Y_j^\top - \tau_j^2 I_{n_u} \quad (\text{B11})$$

$$G_j^{(2)} = A_j P_j A_j^\top + A_j Y_j^\top B_j^\top + B_j Y_j A_j^\top + B_j Y_j P_j^{-1} Y_j^\top B_j^\top + Q_j - P_{j+1} \quad (\text{B12})$$

and where $\Omega_j^{(i)}$ ($i = 1, 2$) is the symmetric positive semidefinite Lagrange multiplier matrix corresponding to the matrix inequality constraint $G_j^{(i)} \leq 0$. However, for a Lagrange multiplier $\Omega_j^{(i)}$, the complementary slackness condition for this problem should read instead [3]

$$\text{tr}(\Omega_j^{(i)} G_j^{(i)}) = 0, \quad i = 1, 2. \quad (2)$$

Since $\Omega_j^{(i)} \geq 0$ and $G_j^{(i)} \leq 0$ it follows from (2) that $\Omega_j^{(i)} G_j^{(i)} = 0$, which implies that the ranges of $\Omega_j^{(i)}$ and $G_j^{(i)}$ are orthogonal [3]. To show, in particular, that $G_j^{(2)} = 0$ at optimality, the authors in [1] assume, ad absurdum, that $G_j^{(2)} < 0$ and proceed to show that $\Omega_j^{(2)} = 0$ as a result of the complementary slackness condition (B8). This argument is, however, incorrect since the negation of $G_j^{(2)} = 0$ is $G_j^{(2)} \leq 0$ and not $G_j^{(2)} < 0$, that is, $G_j^{(2)}$ is allowed to have at least one negative eigenvalue. The correct complementary slackness condition (2) then yields that $G_j^{(2)} \leq 0$ implies that $\Omega_j^{(2)}$ is merely singular, not the zero matrix [2]. Since showing that $G_j^{(2)} = 0$ is required for proving the equivalence of the optimization problems (B1) and (B2), the proposed convex relaxation is not lossless.

Furthermore, in the calculation of the derivative of the Lagrangian with respect to P_j in equation (B22), there is a term missing. The correct version of Equations (B22) and (B27) is

$$-P_j^{-1} Y_j \Omega_j^{(1)} Y_j^\top P_j^{-1} + A_j^\top \Omega_j^{(2)} A_j - P_j^{-1} Y_j B_j \Omega_j^{(2)} B_j^\top Y_j^\top P_j^{-1} - \Omega_{j-1}^{(2)} = 0. \quad (3)$$

The error occurs because the optimality condition is not

$$\frac{\partial L_j}{\partial \mathbf{u}_j^*} = 0, \quad \text{for } j = 0, \dots, N-2 \quad (\text{B16})$$

as stated in [1], but rather

$$\frac{\partial}{\partial \mathbf{u}_j^*} \left(\sum_{i=1}^{N-2} L_i \right) = 0, \quad \text{for } j = 0, \dots, N-2 \quad (4)$$

owing to the fact that, from (B18), the cost function to be minimized is

$$\hat{J} = L_0 + \phi_{N-1} - \langle \lambda_{N-1}, \mathbf{x}_{N-1} \rangle + \sum_{j=1}^{N-2} [L_j - \langle \lambda_j, \mathbf{x}_j \rangle] \quad (\text{B13})$$

The last term in (3) thus stems from the previous term of the sum in (B13). This term disappears when $\Omega_{j-1}^{(2)} = 0$, which,

however, does not follow from the contradiction argument used in the paper as detailed earlier.

The fact that the relaxation proposed in [1] is not lossless can be confirmed using the following counterexample.

Counterexample

Following the notation of [1], consider the discrete time-invariant dynamic system of the form

$$x_{j+1} = Ax_j + Bu_j + c + Gw_j \quad (5)$$

where

$$A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad c = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

and $G_j w_j$ is a Gaussian disturbance with covariance $Q = 0.01I_3$. Consider the covariance steering problem with horizon $N = 20$ steps and boundary conditions

$$P_i = 4I_3, \quad P_f = 2I_3, \quad \mu_0 = [0 \ 0 \ 0]^\top, \quad \mu_f = [0 \ 0 \ 0]^\top.$$

The optimization problem to be solved according to [1] is given as

$$\min_{P_j, Y_j, \rho_j} J = \sum_{j=1}^{N-1} \rho_j \quad (6a)$$

such that, for all $j = 0, 1, \dots, N-1$,

$$P_{j+1} \geq AP_j A^\top + AY_j^\top B^\top + BY_j A^\top + BY_j P_j^{-1} Y_j^\top B^\top + Q \quad (6b)$$

$$\begin{bmatrix} \rho_j I_2 & Y_j \\ Y_j^\top & P_j \end{bmatrix} \geq 0 \quad (6c)$$

$$P_1 = P_i \quad (6d)$$

$$P_N \leq P_f \quad (6e)$$

Where $\rho_j \in \mathbb{R}$, $P_j \in \mathbb{R}^{3 \times 3}$, $Y_j \in \mathbb{R}^{2 \times 3}$ and (6b) is converted to an LMI via the Schur complement as:

$$\begin{bmatrix} P_j & Y_j^\top B^\top \\ BY_j & \Pi_j \end{bmatrix} \geq 0, \quad \Pi_j = P_{j+1} - (AP_j A^\top + AY_j^\top B^\top + BY_j A^\top + Q).$$

Solving this problem numerically using MOSEK and YAMLP [4] one can easily verify that the constraint (6b) is not strict for multiple time steps. For example, at time step $j = 3$, the corresponding matrix is

$$AP_3A^\top + AY_3^\top B^\top + BY_3A^\top + BY_3P_3^{-1}Y_3^\top B^\top + Q - P_4 = \begin{bmatrix} -13.86 & 2.50 & -0.33 \\ 2.50 & -0.87 & 0.17 \\ -0.33 & 0.17 & 0.04 \end{bmatrix} \quad (7)$$

which has eigenvalues $\{-14.34, -0.43, 0\}$. The matrix is therefore singular but not the zero matrix.

Lossless Convexification of Covariance Equation Propagation

One way to achieve lossless convexification to the original non-convex covariance steering problem is to solve

$$\min_{P_j, Y_j, \Phi_j} J = \sum_{j=1}^{N-1} \text{tr}(\Phi_j) \quad (8a)$$

such that, for all $j = 0, 1, \dots, N-1$,

$$Y_j P_j^{-1} Y_j^\top \leq \Phi_j \quad (8b)$$

$$P_{j+1} = A_j P_j A_j^\top + B_j Y_j A_j^\top + A_j Y_j^\top B_j^\top + B_j \Phi_j B_j^\top + Q \quad (8c)$$

$$P_1 = P_i \quad (8d)$$

$$P_N \leq P_f \quad (8e)$$

where the constraint (8b) can be written in terms of the LMI

$$\begin{bmatrix} \Phi_j & Y_j \\ Y_j^\top & P_j \end{bmatrix} \geq 0. \quad (9)$$

For more details please see [2]. As shown in [2] the constraint (8b) is satisfied with equality at the optimal solution, yielding $\Phi_j = Y_j P_j^{-1} Y_j^\top$. Furthermore, rather than penalizing the spectral radius of the control effort covariance $\rho(Y_j P_j^{-1} Y_j)$ as in (6a) and (6c), $\text{tr}(\Phi_j)$ is penalized instead. This difference is crucial when proving that (8b) is satisfied with equality at optimality.

We also point out that the idea of relaxing the covariance equation to an inequality is not new and it has been used before; see, for example, [5] where a similar relaxation technique is used for a system with multiplicative noise. However, the authors in [5] do not claim that this relaxation is lossless.

References

- [1] Benedikter, B., Zavoli, A., Wang, Z., Pizzurro, S., and Cavallini, E., “Convex Approach to Covariance Control with Application to Stochastic Low-Thrust Trajectory Optimization,” *Journal of Guidance, Control, and Dynamics*, Vol. 45, No. 11, 2022, pp. 2061–2075.
- [2] Liu, F., Rapakoulias, G., and Tsiotras, P., “Optimal Covariance Steering for Discrete-Time Linear Stochastic Systems,” *arXiv preprint arXiv:2211.00618*, 2022.
- [3] Vandenberghe, L., and Boyd, S., “Semidefinite programming,” *SIAM review*, Vol. 38, No. 1, 1996, pp. 49–95.
- [4] ApS, M., “Mosek optimization toolbox for matlab,” *User’s Guide and Reference Manual, Version*, Vol. 4, 2019.
- [5] Primbs, J. A., and Sung, C. H., “Stochastic receding horizon control of constrained linear systems with state and control multiplicative noise,” *IEEE transactions on Automatic Control*, Vol. 54, No. 2, 2009, pp. 221–230.