Closed-Form Preconditioner Design for Linear Predictive Control ⋆

Ian McInerney * **Eric C. Kerrigan** *,** **George A. Constantinides** *

* *Department of Electrical & Electronic Engineering, Imperial College London, SW7 2AZ London, UK, email: {i.mcinerney17,e.kerrigan,g.constantinides}@imperial.ac.uk*
** *Department of Aeronautics, Imperial College London, SW7 2AZ London, UK*

Abstract: Model Predictive Control (MPC) with linear models and constraints is extensively being utilized in many applications, many of which have low power requirements and limited computational resources. In these resource-constrained environments, many designers choose to utilize simple iterative first-order optimization solvers, such as the Fast Gradient Method. Unfortunately, the convergence rate of these solvers is affected by the conditioning of the problem data, with ill-conditioned problems requiring a large number of iterations to solve. In order to reduce the number of solver iterations required, we present a simple closed-form method for computing an optimal preconditioning matrix for the Hessian of the condensed primal problem. To accomplish this, we also derive spectral bounds for the Hessian in terms of the transfer function of the predicted system. This preconditioner is based on the Toeplitz structure of the Hessian and has equivalent performance to a state-of-the-art optimal preconditioner, without having to solve a semidefinite program during the design phase.

Keywords: predictive control, preconditioning, Toeplitz matrices, spectral properties

1. PRELIMINARIES

1.1 CLQR Formulation

In this work, we examine the input-constrained Linear Quadratic Regulator (CLQR) formulation of the MPC problem, which can be written as the following constrained quadratic programming problem

$$\min_{u,x} \; \frac{1}{2}x_N' P x_N + \frac{1}{2}\sum_{k=0}^{N-1} \begin{bmatrix} x_k \\ u_k \end{bmatrix}' \begin{bmatrix} Q & 0 \\ 0 & R \end{bmatrix} \begin{bmatrix} x_k \\ u_k \end{bmatrix} \tag{1a}$$

$$\text{s.t. } x_{k+1} = Ax_k + Bu_k, \; k = 0, \dots, N-1 \tag{1b}$$

$$x_0 = \hat{x}_0$$

$$Eu_k \le c_u, \; k = 0, \dots, N-1 \tag{1c}$$

where N is the horizon length, $x_k \in \mathbb{R}^n$ are the states, and $u_k \in \mathbb{R}^m$ are the inputs at sample instant k. $A \in \mathbb{R}^{n\times n}$ and $B \in \mathbb{R}^{n\times m}$ are the state-space matrices describing the discrete-time system $\mathcal{G}_s$, and $\hat{x}_0 \in \mathbb{R}^n$ is the current measured system state. $E \in \mathbb{R}^{l\times m}$ is the stage constraint matrix for the inputs, and the vector $c_u \in \mathbb{R}^l$ is the upper bound for the input constraints. The matrices $Q = Q' \in \mathbb{R}^{n\times n}, R = R' \in \mathbb{R}^{m\times m}, P = P' \in \mathbb{R}^{n\times n}$ are the weighting matrices for the system states, inputs and final states, respectively The weighting matrices are chosen such that P, Q and R are positive definite.

⋆ The support of the EPSRC Centre for Doctoral Training in High Performance Embedded and Distributed Systems (HiPEDS, Grant Reference EP/L016796/1) is gratefully acknowledged.

This problem can be condensed by removing the state variables from (1) to leave only the control inputs in the vector $u := \begin{bmatrix} u_0' & u_1' & \cdots & u_{N-1}' \end{bmatrix}'$. The optimization problem is then the inequality-constrained problem

$$\min_{u} \; \frac{1}{2}u' H_c u + \hat{x}_0' J' u \tag{2a}$$

$$\text{s.t. } Gu \le F\hat{x}_0 + g \tag{2b}$$

with $H_c := \Gamma'\bar{Q}\Gamma + \bar{R}$, $\bar{R} := I_N \otimes R$, $\bar{Q} := \begin{bmatrix} I_{N-1} \otimes Q & 0 \\ 0 & P \end{bmatrix}$,

$$\Gamma := \begin{bmatrix} B & 0 & 0 & 0 \\ AB & B & 0 & 0 \\ A^2B & AB & B & 0 \\ \vdots & & & \ddots & \vdots \\ A^{N-1}B & A^{N-2}B & A^{N-3}B & \cdots & B \end{bmatrix}.$$

1.2 Numerical Examples

Throughout this work we present numerical examples using the discrete-time system with four states and two inputs given in Jones and Morari (2008) with state equation and cost matrices

$$x^+ = \begin{bmatrix} 0.7 & -0.1 & 0.0 & 0.0 \\ 0.2 & -0.5 & 0.1 & 0.0 \\ 0.0 & 0.1 & 0.1 & 0.0 \\ 0.5 & 0.0 & 0.5 & 0.5 \end{bmatrix} x + \begin{bmatrix} 0.0 & 0.1 \\ 0.1 & 1.0 \\ 0.1 & 0.0 \\ 0.0 & 0.0 \end{bmatrix} u,$$

$$Q = \text{diag}(10, 20, 30, 40), \quad R = \text{diag}(10, 20).$$

We constrain the inputs of the system to be $|u_i| \le 0.5$.

2. SPECTRAL PROPERTIES

In order to effectively analyze and derive our closed-form preconditioner, we must first derive some spectral properties of the CLQR matrices Γ and H_c. Similar results to these were reported in Rojas and Goodwin (2004) and Section 11 of Goodwin et al. (2005), but our analysis applies to any positive definite Q matrix, does not require a special rearrangement of the Hessian, and is built upon the principles of mathematical Toeplitz theory instead of Fourier theory.

2.1 Prediction Matrix Γ

We start by analyzing the prediction matrix Γ and note that its diagonals are constant blocks, which means that the matrix is a truncated block Toeplitz matrix. Many properties of a Toeplitz matrix with blocks of size $m \times n$ are closely linked to properties of a matrix-valued function mapping $\mathbb{T} \to \mathbb{C}^{m\times n}$ (with $\mathbb{T} := \{z \in \mathbb{C} : |z| = 1\}$ the unit circle in the complex plane), which is called its matrix symbol. For this work, we focus on matrix symbols that are contained inside $\tilde{\mathcal{C}}_{2\pi}$ — the space of continuous 2π-periodic functions inside $\mathcal{L}^\infty$ (the space of matrix-valued essentially bounded functions).

The diagonal blocks of the matrix give the spectral coefficients of the matrix symbol, so the symbol can be represented as a Fourier series with the coefficients given by the matrix blocks. For Γ, this Fourier series is given by

$$\mathcal{P}_\Gamma(z) = \sum_{i=0}^{k} A^i B z^{-i}, \quad \forall z \in \mathbb{T}, \tag{3}$$

which converges to the matrix symbol in Lemma 1 as $k \to \infty$, provided the discrete-time system is Schur-stable.

Lemma 1. For a Schur-stable system $\mathcal{G}_s$, the prediction matrix Γ has the matrix symbol $\mathcal{P}_\Gamma \in \tilde{\mathcal{C}}_{2\pi}$ with

$$\mathcal{P}_\Gamma(z) := z(zI - A)^{-1}B = z\mathcal{G}_s(z), \qquad \forall z \in \mathbb{T},$$

where $\mathcal{G}_s(\cdot)$ is the transfer function matrix for the system $\mathcal{G}_s$.

Proof. The diagonals of Γ are composed of constant blocks of the form A^iB, where i is the diagonal number (0 is the main diagonal). The matrix symbol for Γ is formed using the trigonometric polynomial of the Fourier series with the diagonal blocks as the coefficients, as given in (3). Since B is a constant matrix, B can be factored out of the summation leaving $\sum_{i=0}^{\infty} A^i z^{-i}$. For a system that is Schur-stable, this summation is a Neumann series that converges to $z(zI - A)^{-1}$ (Peterson and Pederson, 2012, §3.4). Substituting this into (3) then produces the matrix symbol $z(zI - A)^{-1}B$. The spectral coefficients are absolutely summable, so $\mathcal{P}_\Gamma$ is in the Wiener class, meaning that $\mathcal{P}_\Gamma \in \mathcal{L}^\infty$ and is continuous and 2π-periodic, leading to $\mathcal{P}_\Gamma \in \tilde{\mathcal{C}}_{2\pi}$. $\square$

For Γ, the resulting matrix symbol is a time-shifted version of the dynamical system. The assumption of Schur-stability of the system is required for Lemma 1, since if the A matrix were to have eigenvalues outside the unit circle, the trigonometric polynomial would no longer converge and the matrix symbol would be unbounded.

2.2 Hessian H_c

The Hessian of the MPC problem formulation in (2) can be split into three distinct parts

$$H_c := H_Q + H_R + H_P \tag{4}$$

where H_Q, H_R and H_P are the parts that contain the matrices Q, R and P, respectively.

If $P = Q$, the term $\bar{H}_Q := H_Q + H_P = \Gamma^*(I_N \otimes Q)\Gamma$ can be used in (4) instead, forming $H_{cP} = \bar{H}_Q + H_R$. If instead, P is the solution to the discrete-time Lyapunov equation $A'PA + Q = P$ (denoted $P = \mathrm{DLYAP}(A, Q)$), the sum of H_Q and H_P will form a Toeplitz matrix that still has the same matrix symbol as $\bar{H}_Q$. This occurs because selecting the terminal cost matrix P to be the solution of the discrete-time Lyapunov equation captures the value of the cost after the prediction horizon, and extends the summations in each entry of H_Q to infinity. This means that as $N \to \infty$ the effect of H_P becomes concentrated in the lower-right corner of H_{cP} and eventually turns into $\bar{H}_Q$ at infinity. Since this occurs, we can perform the analysis assuming that $P = Q$ and trivially generalize to when P is the solution to the discrete-time Lyapunov equation.

Lemma 2. Let either $P = Q$ or $P = \mathrm{DLYAP}(A, Q)$ and $\mathcal{P}_\Gamma$ be the matrix symbol from Lemma 1 for a Schur-stable system. Then the matrix H_{cP} is a Toeplitz matrix with the matrix symbol $\mathcal{P}_{H_{cP}} \in \tilde{\mathcal{C}}_{2\pi}$, where

$$\mathcal{P}_{H_{cP}}(z) := \mathcal{P}_\Gamma(z)^* Q \mathcal{P}_\Gamma(z) + R, \quad \forall z \in \mathbb{T}.$$

Proof. $H_R := I_N \otimes R$ is a Toeplitz matrix with symbol $\mathcal{P}_R(z) := R$. Let $\tilde{Q} = I_N \otimes Q$. Using the assumptions on the value of P, we can say that $H_Q + H_P = \Gamma^*\tilde{Q}\Gamma$, which is a Toeplitz matrix as well. Since Γ is a lower-triangular matrix and Γ^* is an upper-triangular matrix, the product $\Gamma^*\tilde{Q}\Gamma$ is Toeplitz with generating symbol $\mathcal{P}_\Gamma^* Q \mathcal{P}_\Gamma$ (Gutiérrez-Gutiérrez and Crespo, 2012, Lemma 4.5). Additionally, Toeplitz structure is preserved over addition of two Toeplitz matrices, meaning matrix H_{cP} is then Toeplitz, with the symbol given in the Lemma. $\square$

Since H_{cP} is Toeplitz with the symbol in Lemma 2, we can estimate and bound its eigenvalues using the symbol.

Theorem 1. Let H_{cP} be the condensed Hessian for a Schur-stable system predicted over a horizon of length N with either $P = Q$ or $P = \mathrm{DLYAP}(A, Q)$, and the matrix symbol $\mathcal{P}_{H_{cP}}$ given in Lemma 2, then the following hold:

(a) $\lambda_{min}(\mathcal{P}_{H_{cP}}) \le \lambda(H_{cP}) \le \lambda_{max}(\mathcal{P}_{H_{cP}})$
(b) $\lim_{N\to\infty} \kappa(H_{cP}) = \kappa(\mathcal{P}_{H_{cP}})$

Proof.

(a) The spectrum of a Toeplitz matrix with its symbol in $\tilde{\mathcal{C}}_{2\pi}$ is bounded by the extremes of the spectrum of its symbol (Gutiérrez-Gutiérrez and Crespo, 2012, Theorem 4.4).
(b) Note that H_{cP} is a Hermitian matrix, which means that it is also normal (Horn and Johnson, 2013, §4.1). Since it is both normal and positive semi-definite, $\sigma(H_{cP}) = \lambda(H_{cP})$ (Horn and Johnson, 1994, §3.1), resulting in the condition number becoming $\kappa(H_{cP}) = \frac{\lambda_n(H_{cP})}{\lambda_1(H_{cP})}$. Taking the limit of both sides in

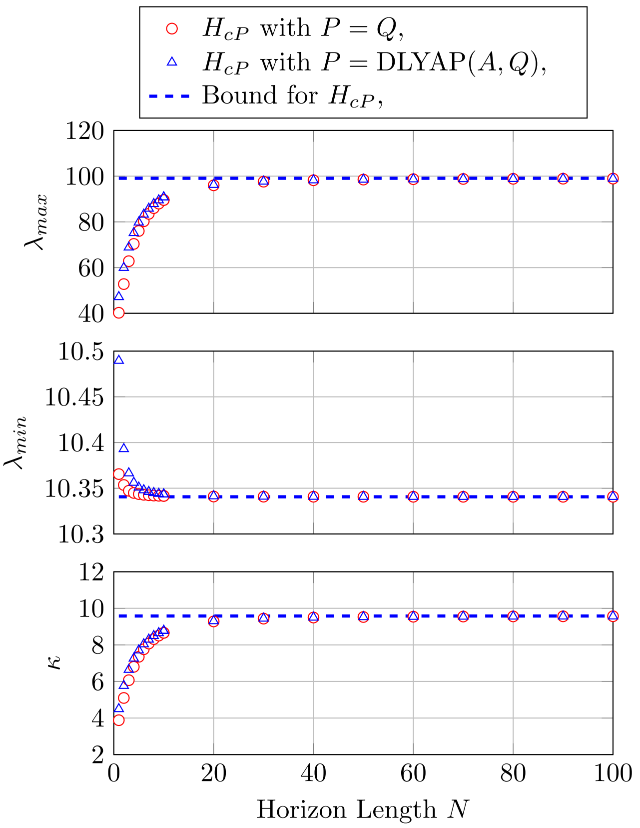

Fig. 1. Spectral properties of the condensed Hessian. The lines represent the bounds computed using Theorem 1, and the markers represent the values of the condensed matrix at that horizon length.

conjunction with the spectral bounds from part (a) gives

$$\lim_{N\to\infty} \kappa(H_{cP}) = \kappa(\mathcal{P}_{H_{cP}}).$$

□

Essentially, these results say that the spectrum for the condensed Hessian will always be contained inside the interval defined by the maximum and minimum eigenvalues of the matrix symbol in Lemma 2. Additionally, as $N \to \infty$ the extremal eigenvalues of H_{cP} will converge asymptotically to the maximum and mimimum eigenvalues of its symbol. This can be seen in Figure 1, which plots the maximum eigenvalues, minimum eigenvalues, and condition number for both H_{cP} with $P = Q$ and P the solution to the discrete-time Lyapunov function.

3. PRECONDITIONING

The spectral results presented in Section 2.2 can be readily extended to analyze the case of a preconditioned Hessian, as well as to help design new preconditioners.

3.1 Analysis of the Preconditioned Hessian

For simplicity of discussion, we focus on the case when H_c is symmetrically preconditioned as $L_N^{-1} H_c (L_N^{-1})'$ with a block-diagonal preconditioner L_N that has N copies of the block L on its diagonal, thus guaranteeing that the preconditioned matrix is Toeplitz. This case is fairly standard in the MPC literature for first-order methods, since it guarantees that the structure of the feasible set is preserved over the preconditioning operation and that the preconditioned Hessian is symmetric (Richter et al., 2012). Results can also be derived for non-block-diagonal preconditioners using Miranda and Tilli (2000, Theorem 4.3) with $M^{-1}H_c$ where $M := L_N L_N'$, but we do not discuss this extension.

Since the preconditioner matrix L_N is block-diagonal with only L on its main diagonal, its matrix symbol is simply L. The results in Section 2.2 can then be extended to the preconditioned matrix H_L by simply replacing $\mathcal{P}_{H_{cP}}$ in Theorem 1 with $\mathcal{P}_{H_L}$ given by

$$\mathcal{P}_{H_L} := \bar{L} \mathcal{P}_{H_{cP}} \bar{L}', \tag{5}$$

where $\bar{L} := L^{-1}$.

3.2 Preconditioner Design

The Toeplitz structure of the Hessian can also be exploited to design preconditioners. There is a rich literature of preconditioners for Toeplitz and circulant matrices, with a focus on designing the preconditioners independent of the size of the matrix (see Chan and Jin (2007) and references therein).

Chan (1988) proposes a closed-form expression for a circulant preconditioner that optimally approximates a given matrix in the Frobenius norm. This can then be used in designing a diagonal preconditioner for H_{cP}:

Theorem 2. Let H_{cP} be the condensed primal Hessian from Section 2.2 and P be the solution to the discrete-time Lyapunov equation $A^T P A + Q = P$. The matrix H_{cP} can be symmetrically preconditioned as $L_N^{-1} H_{cP} (L_N^{-1})'$, where the blocks L are the lower-triangular Cholesky decomposition of M with

$$M := B'PB + R.$$

Proof. Based on the work in Chan (1988), the optimal Circulant preconditioning matrix C for the matrix A will have entries

$$c_i = \frac{i a_{-(n-i)} + (n-i) a_i}{n} \tag{6}$$

where i is the diagonal number, and a_{-i}/a_i are the terms on the i^{th} diagonal below/above the main diagonal respectively. Since we wish to have a block-diagonal preconditioner, we focus only on $i = 0$. In this case, (6) will become the value on the diagonal of H_{cP},

$$c_0 = B'PB + R.$$

□

The block-diagonal preconditioner proposed in Theorem 2 is independent of the horizon length, and is computable for any Schur-stable system. The performance is also similar to that of the optimal preconditioner given in Richter et al. (2012), as shown in Figure 2. Note that the preconditioner from Richter et al. (2012) must be recalculated by solving a semidefinite program for each value of the horizon length, but the preconditioner in Theorem 2 does not need to be.

While the condition number of H_L is the same for both preconditioners, the actual eigenvalue distribution is different. Theorem 2 produces a lower minimum and maximum eigenvalue than the optimal preconditioner, which holds

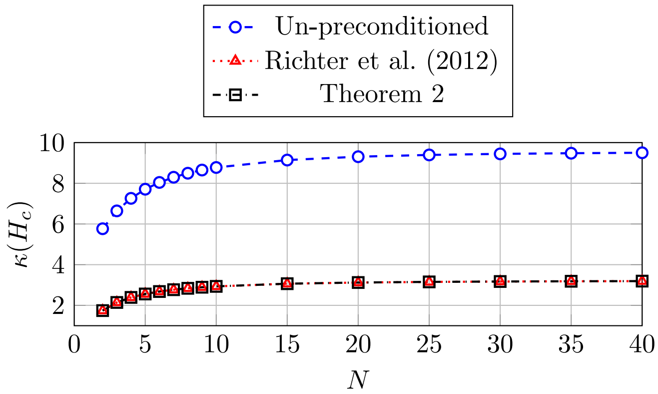

Fig. 2. The effect of preconditioning on the condition number of the condensed Hessian.

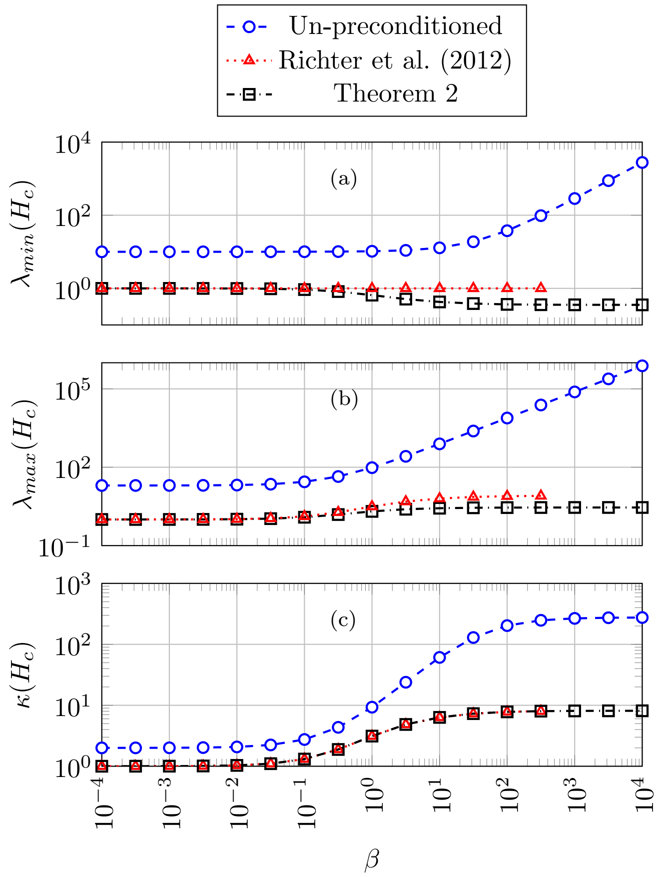

Fig. 3. The effect of the preconditoners on the extremal eigenvalues and the condition number of H_c when the Q matrix is scaled by β and R remains constant.

the lower eigenvalue constant at 1. This effect is most noticeable when the Q matrix dominates the Hessian, as shown in Figures 3a and 3b when the Q weight matrix is scaled by β and the R matrix is held constant. Another interesting thing to note is that the optimal preconditioner from Richter et al. (2012) actually becomes numerically difficult to solve for large ratios of Q to R. In the numerical example, the optimization problem from Richter et al. (2012) becomes infeasible for β above 400, but the proposed preconditioner in Theorem 2 is still calculable.

4. CONCLUSIONS

In this work we presented a closed-form solution for the optimal preconditioner of the condensed CLQR problem. To accomplish this, we also derived results relating the extrema of the spectrum for the condensed Hessian to the extrema of the spectrum for a complex-valued matrix symbol formed using the weighting matrices and the system's transfer function. Using this closed-form expression, we were able to compute a preconditoner that is as effective as the optimal preconditioner from Richter et al. (2012), but that also does not require recomputing when the horizon length changes.

This worked focused on the derivation of the preconditioner and an analysis of its spectral characteristics, and does not include an analysis of its usage. Further experiments should be run to show its effect on the iteration count and convergence of solvers for the condensed MPC problem. The relationship between the spectrum of the transfer function and the spectrum of the condensed Hessian also suggests that other system-theoretic preconditioning strategies may exist. Future work could explore developing preconditioner theory based on loop-shaping of the predicted system to reduce its condition number.

REFERENCES

Chan, R.H. and Jin, X.Q. (2007). *An Introduction to Iterative Toeplitz Solvers*. SIAM.

Chan, T.F. (1988). An Optimal Circulant Preconditioner for Toeplitz Systems. *SIAM Journal on Scientific and Statistical Computing*, 9(4), 766–771.

Goodwin, G.C., Seron, M.M., and De Dona, J.A. (2005). *Constrained Control and Estimation - An Optimisation Based Approach*. Springer, London, UK.

Gutiérrez-Gutiérrez, J. and Crespo, P.M. (2012). Block Toeplitz Matrices: Asymptotic Results and Applications. *Foundations and Trends in Communications and Information Theory*, 8(3), 179–257.

Horn, R.A. and Johnson, C.R. (1994). *Topics in Matrix Analysis*. Cambridge University Press, Cambridge, UK.

Horn, R.A. and Johnson, C.R. (2013). *Matrix Analysis*. Cambridge University Press, Cambridge, UK, second edition.

Jones, C.N. and Morari, M. (2008). The Double Description Method for the Approximation of Explicit MPC Control Laws. In *47th IEEE Conference on Decision and Control (CDC)*, 4724–4730. IEEE, Cancun, Mexico.

Miranda, M. and Tilli, P. (2000). Asymptotic spectra of Hermitian block Toeplitz matrices and preconditioning results. *SIAM Journal on Matrix Analysis and Applications*, 21(3), 867–881.

Peterson, K.B. and Pederson, M.S. (2012). The Matrix Cookbook.

Richter, S., Jones, C.N., and Morari, M. (2012). Computational Complexity Certification for Real-Time MPC With Input Constraints Based on the Fast Gradient Method. *IEEE Transactions on Automatic Control*, 57(6), 1391–1403.

Rojas, O.J. and Goodwin, G.C. (2004). On the asymptotic properties of the Hessian in discrete-time linear quadratic control. In *2004 American Control Conference (ACC)*, 239–244. IEEE, Boston, MA.