Quasi-structure in the dual Hessian for distributed MPC with non-delayed couplings

Emil Klintberg · Sebastien Gros

Abstract Bla bla bla

1 Introduction

Bla bla bla

1.1 Motivation: Distributed MPC with non-delayed couplings

We consider M discrete time linear systems given by:

$$x_{1,i+1} = A_{1,i}x_{1,i} + B_{1,i}u_{1,i} \tag{1a}$$

$$\vdots \tag{1b}$$

$$x_{M,i+1} = A_{M,i}x_{M,i} + B_{M,i}u_{M,i} \tag{1c}$$

where $x_{k,i} \in \mathbb{R}^{p_k}$ and $u_{k,i} \in \mathbb{R}^{m_k}$ represents the state and input of system k. We also assume state and input constraints:

$$x_{k,i} \in X_{k,i} \subseteq \mathbb{R}^{p_k} \tag{2a}$$

$$u_{k,i} \in U_{k,i} \subseteq \mathbb{R}^{m_k} \tag{2b}$$

Emil Klintberg
first address
Tel.: +123-45-678910
Fax: +123-45-678910
E-mail: fauthor@example.com

Sebastien Gros
second address

Moreover, we assume non-delayed couplings between the systems:

$$h_i(x_{1,i}, u_{1,i}, \ldots, x_{M,i}, u_{M,i}) = 0, \quad i = 0, \ldots, N-1 \tag{3}$$

This means that we can state the MPC problem over the horizon N as:

$$\min_{x_{k,i}, u_{k,i}} \quad \sum_{k=1}^{M} \left(\sum_{i=0}^{N-1} l(x_{k,i}, u_{k,i}) + l_f(x_{k,N}) \right) \tag{4a}$$

$$\text{s.t.} \quad x_{k+1,i} = A_{k,i} x_{k,i} + B_{k,i} u_{k,i} \tag{4b}$$

$$h_k(x_{k,1}, u_{k,1}, \ldots, x_{k,M}, u_{k,M}) = 0 \tag{4c}$$

$$x_{k,i} \in X_{k,i}, \quad u_{k,i} \in U_{k,i} \tag{4d}$$

Furthermore, let us assume that the non-delayed couplings (4c) are linear, constraints (4d) are polyhedral and that the objective function (4a) is quadratic, and introduce the following notation: $z_{k,i} = [x_{k,i}^T \quad u_{k,i}^T]^T \in \mathbb{R}^{n_k}$. The MPC problem can then be stated as:

$$\min_{z} \quad \sum_{k=1}^{M} \sum_{i=0}^{N} \frac{1}{2} z_{k,i}^T H_{k,i} z_{k,i} + c_{k,i}^T z_{k,i} \tag{5a}$$

$$\text{s.t.} \quad \sum_{k=1}^{M} F_{k,i} z_{k,i} = e_i \tag{5b}$$

$$C_{k,i} z_{k,i} + D_{k,i+1} z_{k,i+1} = d_{k,i} \tag{5c}$$

$$G_{k,i} z_{k,i} \leq f_{k,i} \tag{5d}$$

where $H_{k,i} \in \mathbb{S}_{++}^{n_k \times n_k}$, $C_{k,i} \in \mathbb{R}^{l_k \times n_k}$, $D_{k,i+1} \in \mathbb{R}^{l_k \times n_k}$ and $d_{k,i} \in \mathbb{R}^{l_k}$ form the dynamics, $F_{k,i} \in \mathbb{R}^{r_i \times n_k}$ and $e_i \in \mathbb{R}^{r_i}$ yield the coupling constraints, $G_{k,i} \in \mathbb{R}^{t_k \times n_k}$ and $f_{k,i} \in \mathbb{R}^{t_k}$ form the local constraints.

Additionally, to avoid an unnecessary heavy notation at places where we are only dealing with decomposition in space, we introduce the following augmented notations: $z_k = [z_{k,0}^T \ldots, z_{k,N-1}^T]^T \in \mathbb{R}^{n_k}$. The MPC problem (5) can then be expressed as:

$$\min_{z} \quad \sum_{k=1}^{M} \frac{1}{2} z_k^T H_k z_k + c_k^T z_k \tag{6a}$$

$$\text{s.t.} \quad \sum_{k=1}^{M} F_k z_k = e \tag{6b}$$

$$C_k z_k = d_k \tag{6c}$$

$$G_k z_k \leq f_k \tag{6d}$$

where $H_k \in \mathbb{S}_{++}^{(N+1)n_k \times (N+1)n_k}$, $C_k \in \mathbb{R}^{Nl_k \times (N+1)n_k}$ and $d_k \in \mathbb{R}^{Nl_k}$, $F_k \in \mathbb{R}^{(N+1)r \times (N+1)n_k}$ and $e \in \mathbb{R}^{(N+1)r}$, $G_k \in \mathbb{R}^{(N+1)t_k \times (N+1)n_k}$ and $f_k \in \mathbb{R}^{(N+1)t_k}$

and the matrices are accordingly given by:

$$H_k = \begin{bmatrix} H_{k,0} & & \\ & \ddots & \\ & & H_{k,N} \end{bmatrix},$$

$$C_k = \begin{bmatrix} C_{k,0} & D_{k,1} & & & \\ & C_{k,1} & D_{k,2} & & \\ & & \ddots & \ddots & \\ & & & C_{k,N-1} & D_{k,N} \end{bmatrix},$$

$$A_k = \begin{bmatrix} A_{k,0} & & \\ & \ddots & \\ & & A_{k,N} \end{bmatrix}.$$

2 Prelimenaries

2.1 Dual decomposition with second-order information

We introduce the dual variables $\lambda \in \mathbb{R}^{(N+1)r}$ corresponding to the coupling constraints (6b) and define the Lagrange function as

$$\mathcal{L}(z, \lambda) = \sum_{k=1}^{M} \left(\frac{1}{2} z_k^T H_k z_k + c_k^T z_k \right) + \lambda^T \left(\sum_{k=1}^{M} F_k z_k - e \right) \tag{8}$$

Note that $\mathcal{L}(z, \lambda)$ is separable in z, i.e.

$$\mathcal{L}(z, \lambda) = \sum_{k=1}^{M} \mathcal{L}_k(z_k, \lambda) \tag{9}$$

with

$$\mathcal{L}_k(z_k, \lambda) = \frac{1}{2} z_k^T H_k z_k + c_k^T z_k + \lambda^T \left(C_k z_k - \frac{1}{M} d \right) \tag{10}$$

The dual function $d(\lambda) = -\min_{z \in \mathcal{Z}} \mathcal{L}(z, \lambda)$ can thus be evaluated in parallel as:

$$d(\lambda) = -\sum_{k=1}^{M} \min_{z_k \in \mathcal{Z}_k} \mathcal{L}_k(z_k, \lambda) \tag{11}$$

Since (6) is strictly convex, (11) is convex and continuously differentiable, but not twice differentiable. However, the Hessian of $d(\lambda)$ is a piecewise constant matrix and change with the active-set [5].

The non-smoothness implies that $d(\lambda)$ is not self-concordant and the solution to the dual problem is hence not easily tracked with Newton's method.

However, if we relax the inequality constraints (6d) with a self-concordant log-barrier, according to:

$$\min_z \quad \sum_{k=1}^M \frac{1}{2} z_k^T H_k z_k + c_k^T z_k - \tau \sum_{i=1}^{(N+1)t_k} \log([s_k]_i) \tag{12a}$$

$$\text{s.t.} \quad \sum_{k=1}^M F_k z_k = e \tag{12b}$$

$$C_k z_k = d_k \tag{12c}$$

$$G_k z_k + s_k = f_k \tag{12d}$$

where $\tau > 0$ will be referred to as the *barrier parameter*, the resulting *relaxed dual function* $d(\lambda, \tau)$ is self-concordant [6]. This opens for the possibility of solving a sequence of dual problems $\{\min_\lambda d(\lambda, \tau)\}_{\tau \to 0}$ where each problem is self-concordant and therefore easily solved with Newton's method.

The relaxed dual function is separable and can be evaluated in parallel as

$$d(\lambda, \tau) = -\sum_{k=1}^N \min_{z_k \in \mathcal{Z}_k} \left(\mathcal{L}_k(z_k, \lambda) - \tau \sum_{i=1}^{m_k} \log([s_k]_i) \right) \tag{13}$$

Hence, evaluating (13) involves solving local subproblems of the form

$$\begin{aligned} \min_{z_k} \quad & \frac{1}{2} z_k^T H_k z_k + c_k^T z_k + \lambda^T F_k z_k - \tau \sum_{i=1}^{m_k} \log([s_k]_i) \\ \text{s.t.} \quad & C_k z_k = d_k \\ & G_k z_k + s_k = f_k \\ & s_k \geq 0 \end{aligned} \tag{14}$$

The relaxed dual problem then reads

$$\min_\lambda d(\lambda, \tau) \tag{15}$$

from which solution, the solution to (12) can be recovered according to strong duality [2].

Strict convexity also implies that the gradient of $d(\lambda, \tau)$ is given by the residual of the coupling constraints [1], i.e.

$$\nabla d(\lambda, \tau) = -\sum_{k=1}^N F_k z_k^*(\lambda, \tau) + e \tag{16}$$

where $z_k^*(\lambda, \tau) = \arg\min_{z_k \in \mathcal{Z}_k} \mathcal{L}_k(z_k, s_k, \lambda, \tau)$. The dual Hessian is then given by

$$\nabla^2 d(\lambda, \tau) = -\sum_{k=1}^N F_k \frac{\partial z_k^*(\lambda, \tau)}{\partial \lambda} \tag{17}$$

A Newton direction $\Delta\lambda$ in the dual space can then be obtained as a solution to the Newton system

$$\nabla^2 d(\lambda, \tau) \Delta\lambda + \nabla d(\lambda, \tau) = 0 \tag{18}$$

3 Structure in the dual Hessian

3.1 Dual Hessian

By introducing $y_k = \tau/s_k \in \mathbb{R}^{m_k}$, the KKT conditions to (14) are given by

$$r_k(w_k^*, \lambda, \tau) = \begin{bmatrix} r_{Dk}(w_k^*, \lambda) \\ r_{Ek}(w_k^*) \\ r_{Ik}(w_k^*) \\ r_{Ck}(w_k^*, \tau) \end{bmatrix} = 0 \tag{19a}$$

$$s_k^* > 0, \quad y_k^* > 0 \tag{19b}$$

where we use the notation $w_k = [z_k^T, \mu_k^T, y_k^T, s_k^T]^T$ for the local primal-dual variables and $r_k(w_k, \lambda, \tau)$ is given by:

$$r_{D_k}(w_k, \lambda) = H_k z_k + c_k + F_k^T \lambda + C_k^T \mu_k + G_k^T y_k \tag{20a}$$

$$r_{Ek}(w_k) = C_k z_k - d_k \tag{20b}$$

$$r_{Ik}(w_k) = G_k z_k + s_k - f_k \tag{20c}$$

$$r_{Ck}(w_k, \tau) = Y_k s_k - \tau \mathbf{1} \tag{20d}$$

As we have seen in previous section, in order to form the dual Hessian we need to compute $\frac{\partial z_k^*}{\partial \lambda}$, where $z_k^* = z_k^*(\lambda)$ is the optimal primal solution and is hence fulfilling (19). By differentiating (19), the following linear system is obtained:

$$\begin{bmatrix} H_k & C_k^T & G_k^T & 0 \\ C_k & 0 & 0 & 0 \\ G_k & 0 & 0 & I \\ 0 & 0 & S_k & Y_k \end{bmatrix} \begin{bmatrix} \frac{\partial z_k^*}{\partial \lambda} \\ \frac{\partial \mu_k^*}{\partial \lambda} \\ \frac{\partial y_k^*}{\partial \lambda} \\ \frac{\partial s_k^*}{\partial \lambda} \end{bmatrix} = - \begin{bmatrix} F_k^T \\ 0 \\ 0 \\ 0 \end{bmatrix} \tag{21}$$

If block elimination of (21) is used, the *normal equations* can be formed [7],

$$\Lambda_k \frac{\partial \mu_k^*}{\partial \lambda} = -C_k \Phi_k^{-1} F_k^T \tag{22a}$$

$$\Phi_k \frac{\partial z_k^*}{\partial \lambda} = -F_k^T - C_k^T \frac{\partial \mu_k^*}{\partial \lambda} \tag{22b}$$

$$\frac{\partial s_k^*}{\partial \lambda} = -G_k \frac{\partial z_k^*}{\partial \lambda} \tag{22c}$$

$$\frac{\partial y_k^*}{\partial \lambda} = -S^{-1} Y \frac{\partial s_k^*}{\partial \lambda} \tag{22d}$$

where $\Phi_k = H_k + G_k^T S_k^{-1} Y_k G_k \in \mathbb{S}_{++}^{(N+1)n_k \times (N+1)n_k}$ and $\Lambda_k = C_k \Phi_k^{-1} C_k^T \in \mathbb{S}_{++}^{Nl_k \times Nl_k}$. By using (22a) and (22b), it can be obtained that

$$F_k \frac{\partial z_k}{\partial \lambda} = -F_k (\Phi_k^{-1} - \Phi_k^{-1} C_k^T \Lambda_k^{-1} C_k \Phi_k^{-1}) F_k^T \tag{23}$$

This means that the dual Hessian (17) can be written as

$$\nabla^2 d(\lambda, \tau) = \sum_{k=1}^{M} F_k(\Phi_k^{-1} - \Phi_k^{-1} C_k^T \Lambda_k^{-1} C_k \Phi_k^{-1}) F_k^T \qquad (24)$$

Let us now turn our attention to the structure coming from the time domain. We have already seen that C_k is block bidiagonal, and it can trivially be realized that Φ_k is block diagonal. This means that Λ_k has a block tridiagonal structure given by:

$$\Lambda_k = \begin{bmatrix} \Lambda_{11} & \Lambda_{12} & & \\ \Lambda_{12}^T & \Lambda_{22} & \ddots & \\ & \ddots & \ddots & \Lambda_{N-1,N} \\ & & \Lambda_{N-1,N}^T & \Lambda_{N,N} \end{bmatrix} \in \mathbb{S}_{++}^{Nl_k \times Nl_k} \qquad (25)$$

where

$$\Lambda_{i,i} = C_{k,i-1}\Phi_{k,i-1}^{-1}C_{k,i-1}^T + D_{k,i}\Phi_{k,i}^{-1}D_{k,i}^T \in \mathbb{R}^{l_k, \times l_k} \qquad (26a)$$

$$\Lambda_{i,i+1} = D_{k,i}\Phi_{k,i}^{-1}C_{k,i}^T \in \mathbb{R}^{l_k \times l_k} \qquad (26b)$$

Accordingly, all matrices in (24) are banded except Λ_k^{-1} which in general is dense.

3.2 Decaying of the dual Hessian

Let us first assume that the problem data is bounded. In other words, let us assume the following:

Assumption 1 *The row and column absolute sums of Jacobians of equality constraints (i.e. (6b) and (6c)) are bounded. Hence:*

1. $\|C_{k,i}\|_\infty \leq \gamma$ *and* $\|C_{k,i}\|_1 \leq \gamma$
2. $\|D_{k,i}\|_\infty \leq \gamma$ *and* $\|D_{k,i}\|_1 \leq \gamma$
3. $\|F_{k,i}\|_\infty \leq \gamma$ *and* $\|F_{k,i}\|_1 \leq \gamma$

It should be observed that Assumption 1 is not by any means restrictive, since any solver would struggle with a problem where it is not fulfilled.

Possible problems with interior-point methods are related to numerical difficulties due to ill-conditioning []. This might occur since elements in $S_k^{-1}Y_k$ can be close to zero in late iterations when τ is small. Our experience is however that this is not a major issue in practice, which is also supported by commercial interior-point implementations []. Moreover, since we assume that (5) is strongly convex, the eigenvalues of Φ_k is lower bounded by the smallest eigenvalue of H_k, even when $S_k^{-1}Y_k$ is singular. According to this reasoning it is not restrictive to assume boundedness of Φ_k^{-1} which we do next:

Assumption 2 *The row and column absolute sums of Φ_k^{-1} are bounded. Hence:*

$$\|\Phi_{k,i}^{-1}\|_\infty = \|\Phi_{k,i}^{-1}\|_1 \leq \gamma_{\Phi^{-1}} \qquad (27)$$

Inverses of sparse matrices are in general dense, but individual elements are often small in absolute value. Since Λ_k is banded, symmetric and positive definite, we will relay our analysis on the following classical result:

Lemma 1 *If A is Hermitian positive definite and m-banded ($[A]_{ij} = 0$ if $|i - j| > m$), the entries of A^{-1} satisfy the following bound:*

$$|[A^{-1}]_{ij}| < K\omega^{|i-j|}, \quad \forall i, j \qquad (28)$$

where $[a, b]$ is the smallest interval containing the spectrum $\sigma(A)$ of A, $K = \max\{a^{-1}, K_0\}$, $K_0 = (1 + \sqrt{\kappa})$, $\omega = q^{1/m}$, $q = \frac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1}$, $\kappa = \frac{b}{a}$.

Proof A proof is given in []. □

This means that the entries of A^{-1} are bounded by an exponentially decaying function along each row or column. However, the bound depends on the condition number and the bandwidth of the matrix. Matrices with a high condition number and/or a high bandwidth can accordingly result in a large K and $\omega \approx 1$, leading to a slow decay. The opposite, i.e. a low condition number and a small band, would result in a rapid decay.

Observe that due to the block tridiagonal structure of Λ_k, it is $3l_k$-banded. Moreover, if we introduce the notation

$$\Lambda_k^{-1} = \begin{bmatrix} T_{11} & T_{21}^T & \cdots & T_{N+1,1}^T \\ T_{21} & T_{22} & \cdots & T_{N+1,2}^T \\ \vdots & \vdots & \ddots & \vdots \\ T_{N+1,1} & T_{N+1,2} & \cdots & T_{N+1,N+1} \end{bmatrix} \qquad (29)$$

where $T_{i,j} \in \mathbb{R}^{l_k \times l_k}$, we can establish the following proposition:

Proposition 1 *The off-diagonal blocks (i.e. $T_{i,j}$ where $i - j > 0$) in Λ_k^{-1} satisfy the following bounds:*

$$\|T_{i,j}\|_\infty \leq l_k K_k \omega_k^{(i-j)l_k+1} \qquad (30a)$$

$$\|T_{i,j}\|_1 \leq l_k K_k \omega_k^{(i-j)l_k+1} \qquad (30b)$$

where K_k and ω_k are defined according to Lemma 1.

Proof According to Lemma 1, the element in $T_{i,j}$ with the largest bound is located in the top-right corner, and is hence element $[\Lambda_k^{-1}]_{il_k+1, jl_k}$. From Lemma 1 it then follows that:

$$\max|[T_{ij}]| \leq K_k \omega_k^{(i-j)l_k+1} \qquad (31)$$

where $\max |[T_{ij}]|$ refers to the maximum absolute value of the components in T_{ij}. Accordingly, we have that:

$$\|T_{i,j}\|_\infty \le l_k K_k \omega_k^{(i-j)l_k+1} \tag{32a}$$

$$\|T_{i,j}\|_1 \le l_k K_k \omega_k^{(i-j)l_k+1} \tag{32b}$$

□

Proposition 1 confirms our intuition that the row and column sum of T_{ij} are decreasing when $|i-j|$ increases. For interior point methods in general, the condition number of Λ_k can be very high for small values of τ. It should however be understood that strong convexity of (5), should improve the conditioning of Φ_k and hence also of Λ_k.

To gain some perspective, we have concluded that Λ_k^{-1} is decaying exponentially towards the off-diagonal corners and that all other matrices in (24) are banded. This suggests that also $\nabla_{\lambda\lambda}^2 d(\lambda, \tau)$ should decay towards the off-diagonal corners. To maintain a simple reasoning, one more stepping stone will be used before arriving at the main results of this section. Therefore, let us introduce the notation:

$$C_k^T \Lambda_k^{-1} C_k = \begin{bmatrix} V_{11} & V_{21}^T & \cdots & V_{N+1,1}^T \\ V_{21} & V_{22} & \cdots & V_{N+1,2}^T \\ \vdots & \vdots & \ddots & \vdots \\ V_{N+1,1} & V_{N+1,2} & \cdots & V_{N+1,N+1} \end{bmatrix} \tag{33}$$

where $V_{ij} \in \mathbb{R}^{l_k \times l_k}$, and look at the decay of $C_k^T \Lambda_k^{-1} C_k$.

Proposition 2 *The off diagonal blocks (i.e. $V_{i,j}$ where $i-j > 0$) of $C_k^T \Lambda_k^{-1} C_k$ satisfy the following bounds:*

$$\|V_{i,j}\|_\infty \le (2 + \omega_k^{l_k} + \omega_k^{-l_k})\gamma^2 l_k K_k \omega_k^{(i-j-1)l_k+1} \tag{34a}$$

$$\|V_{i,j}\|_1 \le (2 + \omega_k^{l_k} + \omega_k^{-l_k})\gamma^2 l_k K_k \omega_k^{(i-j-1)l_k+1} \tag{34b}$$

Proof Bla bla bla   □

Let us now introduce the notation:

$$\nabla^2 d(\lambda, \tau) = \begin{bmatrix} W_{11} & W_{21}^T & \cdots & W_{N+1,1}^T \\ W_{21} & W_{22} & \cdots & W_{N+1,2}^T \\ \vdots & \vdots & \ddots & \vdots \\ W_{N+1,1} & W_{N+1,2} & \cdots & W_{N+1,N+1} \end{bmatrix} \tag{35}$$

where $W_{i,j} \in \mathbb{R}^{r_i \times r_j}$. We can now establish a decay towards the off-diagonal corners of $\nabla^2 d(\lambda, \tau)$.

Proposition 3 *The off-diagonal blocks (i.e. $W_{i,j}$ where $i-j > 0$) of $\nabla^2 d(\lambda, \tau)$ satisfy the following bounds:*

$$\|W_{i,j}\|_\infty \leq \sum_{k=1}^{M}(2 + \omega_k^{l_k} + \omega_k^{-l_k})\gamma^4\gamma_{\Phi^{-1}}^2 l_k K_k \omega_k^{(i-j-1)l_k+1} \tag{36a}$$

$$\|W_{i,j}\|_1 \leq \sum_{k=1}^{M}(2 + \omega_k^{l_k} + \omega_k^{-l_k})\gamma^4\gamma_{\Phi^{-1}}^2 l_k K_k \omega_k^{(i-j-1)l_k+1} \tag{36b}$$

Proof To be done.

Finally, let us bound the euclidean distance between $\nabla^2 d(\lambda, \tau)$ and a band along its diagonal. To do so, we start by recalling Gershgorin's circle theorem:

Theorem 1 *For $A \in \mathbb{R}^{n \times n}$ with elements a_{ij}, let $R_i = \sum_{j \neq i}|a_{ij}|$ be the sum of the absolute values of the non-diagonal entries in the row i. Let $D(a_{ii}, R_i)$ be the closed disc centered in a_{ii} with radius R_i, then every eigenvalue of A lies within at least one of the discs $D(a_{ii}, R_i)$.*

Proof A proof is given in []. $\qquad\square$

Finally, we can establish our main result:

Lemma 2 *Let $\lfloor\nabla^2 d(\lambda, \tau)\rceil_\mathcal{M}$ represent the diagonal $\mathcal{M}$-block band of $\nabla^2 d(\lambda, \tau)$. The following bound holds:*

$$\|\lfloor\nabla^2 d(\lambda, \tau)\rceil_\mathcal{M} - \nabla^2 d(\lambda, \tau)\|_2 \leq \sum_{k=1}^{M}(N+1-\mathcal{M})(2+\omega_k^{l_k}+\omega_k^{-l_k})\gamma^4\gamma_{\Phi^{-1}}^2 l_k K_k \omega_k^{\frac{\mathcal{M}-1}{2}l_k+1} \tag{37}$$

where...

Proof Since $\lfloor\nabla^2 d(\lambda, \tau)\rceil_\mathcal{M} - \nabla^2 d(\lambda, \tau)$ is symmetric, its singular values collapse into the absolute value of its eigenvalues. The problem of finding a bound on the 2-norm is hence reduced to the problem of bounding the magnitude of the largest eigenvalue.

First, observe that all diagonal elements of $\lfloor\nabla^2 d(\lambda, \tau)\rceil_\mathcal{M} - \nabla^2 d(\lambda, \tau)$ are zero. This means that every Gershgorin disc will be centered in the origin, and a bound on the largest eigenvalue can be found by finding the largest radius of a Gershgorin disc.

The blocks $W_{i,j}$ in $\lfloor\nabla^2 d(\lambda, \tau)\rceil_\mathcal{M} - \nabla^2 d(\lambda, \tau)$ with the largest bound is located next to the diagonal band of zeros, i.e. where $i - j = \frac{\mathcal{M}-1}{2} + 1$. Furthermore, at each block row there are $N + 1 - \mathcal{M}$ nonzero blocks. This means that we can establish a bound on the radius of the Gershgorin discs, and hence the following:

$$\|\lfloor\nabla^2 d(\lambda, \tau)\rceil_\mathcal{M} - \nabla^2 d(\lambda, \tau)\|_2 \leq \sum_{k=1}^{M}(N+1-\mathcal{M})(2+\omega_k^{l_k}+\omega_k^{-l_k})\gamma^4\gamma_{\Phi^{-1}}^2 l_k K_k \omega_k^{\frac{\mathcal{M}-1}{2}l_k+1} \tag{38}$$

$$\square$$

Say something about the principal behavior of the bound...

It is well known that the use of an inexact Hessian will degrade the convergence of Newton's method to a linear rate []. However, let us still look at the relative error in the Newton direction that we get from using $\lfloor \nabla^2 d(\lambda, \tau) \rceil_{\mathcal{M}}$ instead of $\nabla^2 d(\lambda, \tau)$. First, recall the following lemma:

Lemma 3 *If x is a solution to $Ax = b$ and $\hat{x}$ is a solution to the pertubed system $(A + F)\hat{x} = b + f$, then the following bound holds:*

$$\frac{\|\hat{x} - x\|_2}{\|x\|_2} \leq \kappa(A)(\rho_A - \rho_b) \tag{39}$$

where $\kappa(A)$ represents the condition number of A, $\rho_A = \frac{\|F\|_2}{\|A\|_2}$ and $\rho_b = \frac{\|f\|_2}{\|b\|_2}$.

Proof A proof is given in [].

Accordingly, by combining Lemma 2 and Lemma 3 we can establish the following:

Lemma 4

$$\frac{\|\Delta\hat{\lambda} - \Delta\lambda\|_2}{\|\Delta\lambda\|_2} \leq \tag{40}$$

4 Numerical experiments

4.1 Decaying of the dual Hessian

4.1.1 A small problem

As an example of a small problem, we use a randomly generated problem with 4 subproblems (i.e. $M = 4$), where each subproblem has $N = 10$, 6 states and 4 controls and hence $n_k = 10$, and each time instance has 4 coupling constraints.

According to Section 3.2, we expect a slower decay when the conditioning of $S_k^{-1} Y_k$ gets worse, i.e. when we reduce τ. Surprisingly, this is something that does not seem to have a big influence in practice. To illustrate this, Figure 2 shows the dual Hessian for the same problem as in Figure 1, but with $\tau = 10^{-5}$.

4.1.2 A large problem

4.2 Newton steps with inexact Hessian

References

1. Bertsekas, D., Tsitsiklis, J.N.: Parallel and distributed computation: Numerical methods. Prentice Hall (1989)
2. Boyd, S., Vandenberghe, L.: Convex Optimization. University Press, Cambridge (2004)
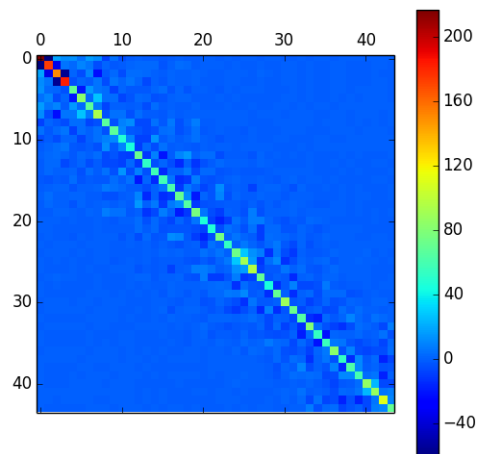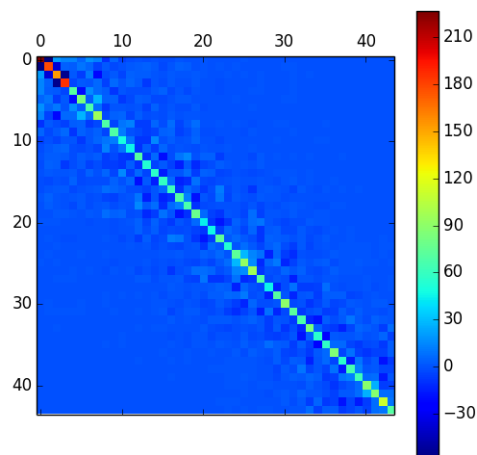
Fig. 1 $\tau = 1$

Fig. 2 $\tau = 10^{-5}$

3. Domahidi, A., Zgraggen, A., Zeilinger, M., Morari, M., Jones, C.: Efficient Interior Point Methods for Multistage Problems Arising in Receding Horizon Control. In: IEEE Conference on Decision and Control (CDC), pp. 668 – 674. Maui, HI, USA (2012)
4. Klintberg, E., Gros, S.: A primal-dual newton method for distributed quadratic programming. In: Conference on Decision and Control (2014)

5. Kozma, A., Klintberg, E., Gros, S., Diehl, M.: An improved distributed dual Newton-CG method for convex quadratic programming problems. In: American Control Conference (2014)
6. Necoara, I., Suykens, J.: An interior-point Lagrangian decomposition method for separable convex optimization. J. Optim. Theory and Appl. **143**(3), 567–588 (2009)
7. Wright, S.: Primal-Dual Interior-Point Methods. SIAM Publications, Philadelphia (1997)