

# Exploitation of the Quasi-structure of the Dual Hessian for distributed MPC with Non-delayed Couplings

Emil Klintberg · Sebastien Gros

Received: date / Accepted: date

**Abstract** This paper studies the quasi-banded structure in the Hessian of the dual function, originating from many distributed Model Predictive Control (MPC) problems. Due to the quasi-structure, we find a banded approximation of the dual Hessian, for which structure exploiting factorization methods can be used. Moreover, we show that the approximation can be used to form a quasi-Newton method. Especially for problems with a long horizon, the quasi-Newton method can find good approximations of the Newton direction, by only considering a fraction of the elements in the dual Hessian. The quasi-structure is present when both interior-point methods and active-set methods are used to solve the subproblems.

**Keywords** First keyword · Second keyword · More

## 1 Introduction

Recently there has been a great interest invested in controlling large-scale, networked systems with Model Predictive Control (MPC). Many approaches have emerged, with the common feature that they allow for distributed computations with a limited amount of communication. This is indeed an important property since computing the MPC solutions for a large-scale problem can

---

Emil Klintberg  
Chalmers Uni. of Tech., Signals and Systems  
Hörsalsvägen 11  
Göteborg  
E-mail: kemil@chalmers.se

Sebastien Gros  
Chalmers Uni. of Tech., Signals and Systems  
Hörsalsvägen 11  
Göteborg  
E-mail: grosse@chalmers.se

be intractable for a single computing unit. Moreover, when the system is geographically distributed, or when subsystems do not want to share sensitive information, it can be highly impractical to centralize the data of the problem.

Distributed MPC schemes have been proposed for various classes of control problems. In [20], a distributed multiple shooting method is proposed to tackle large-scale nonlinear MPC problems, in [29], a method to tackle distributed stochastic MPC problems is proposed, whereas in [4] distributed robust MPC is discussed.

Formulations of distributed MPC schemes can be divided into categories depending on the degree of communication between subsystems. Methods where no communication between the local controllers are allowed are commonly denoted as decentralized MPC. Methods belonging to this class offer great scalability, and can give a good performance when the subsystems are weakly coupled [1], [22]. However, when this is not fulfilled, the performance can be far from optimal.

Methods where communication between subsystems are allowed, are commonly referred to as cooperative/non-cooperative MPC depending on the cost function used in the local MPC controllers [34], [30], [32]. In terms of optimality, methods in this class often manage to perform better compared to cooperative MPC, but generally show a derated performance compared to the centralized formulations of the MPC problem.

Another promising approach to achieve distributed MPC is to formulate the MPC problem centrally and solve the underlying optimization problem with techniques stemming from the field of distributed optimization. A commonly used technique for this purpose is to use a *Lagrangian relaxation* to decompose the optimization problem into low-dimensional subproblems that can be solved independently.

Lagrangian relaxation is used in many different contexts to tackle convex large-scale problems, e.g. the authors in [5] propose a coordinate ascent approach to solve matrix problems. In [21], [12] a gradient method, whereas in [31], [24], [11], [10], [13] a fast gradient method is used in order to attain dual optimality. Moreover, in [23], [25], convergence rates for first-order methods are studied. All these methods make use of only first-order derivatives when finding a search direction.

Methods using second-order derivatives offer advantages in terms of convergence speed. In the context of active-set methods, this is exploited by the authors in [8], [18], [19], [9], [14]. However, the dual function is non-smooth and the dual Hessian can be singular if a poor initial guess is used for the dual variables, and hence possibly leading to an inconsistent Newton system. In [19], this issue was avoided by relaxing the local inequality constraints with an  $L_2$ -penalty.

In [26], it was proved that adding self-concordant barrier terms to the Lagrange function of a generic convex problem render a self-concordant dual function. Hence, in the context of interior-point methods, it is possible to use the Newton method to efficiently solve the dual problem. This result was used in [33], where a method based on inexact solutions of the local subproblems was

proposed. Moreover, in [16], a primal-dual interior-point method was proposed, which supports linear predictors and inexact solutions and showed a fast and consistent practical convergence.

Indeed, methods using second-order information offer a fast convergence in terms of iterations. However, this comes to the cost of factorizing the Hessian of the dual function at each iteration. Typically the cost of performing this factorization is cubic with respect to the dimensions of the dual space, and can hence be a bottleneck when solving large-scale problems. For specific applications, the dual Hessian can be structured and the cost of factorizing it can therefore be lowered. In the context of subproblems connected in a chain, this is exploited in [9]. However, most problems render a dense dual Hessian.

In this paper, we show that the dual Hessian of optimization problems originating from many MPC schemes have an inherent quasi-structure. The quasi-structure can be used to find a banded approximation of the dual Hessian, which offer a substantially lower cost of factorization. Hence, a quasi-Newton method can be formed which can solve many MPC problems to a lower overall complexity. The method performs especially well on problems with a very long horizon, which otherwise are challenging due to the size of the dual Hessian.

The quasi-structure arises when distributed systems, with non-delayed interactions between subsystems, are controlled with MPC. As an illustrative example, the structure will be analyzed in the context of the interior-point method described in [26]. It should however be understood that the dual Hessian of the methods described in [33] and [16] take the same form. For completeness, we also give a short description of the quasi-structure for active-set methods.

The paper is organized as follows. In Section 2, the considered class of problems will be introduced and the dual decomposition approach with second-order derivatives will be recalled, where emphasis will be placed on the computation of the Hessian of the dual function. In Section 3, the quasi-banded structure of the dual Hessian is investigated. It is shown that the quasi-structure can be used to form a banded approximation of the dual Hessian, for which structure exploiting factorization methods can be used. Moreover, a quasi-Newton method based on the approximation of the dual Hessian is proposed. In Section 4, the quasi-structure is briefly described in the context of active-set methods. In Section 5, numerical experiments are performed on two examples. Features of the quasi-structure are shown and convergence properties of the quasi-Newton method are presented. The paper is concluded in Section 6.

## 1.1 Contributions

This paper contain three contributions:

1. We establish that the Hessian of the dual function that is resulting from many MPC problems has a quasi-banded structure.

2. As a result of the quasi-structure, we show that the dual Hessian can be approximated by a banded matrix for which structure exploiting factorization methods can be used.
3. We show that the quasi-structure can be used to form a quasi-Newton method, where a good approximation of the Newton direction can be found by only considering a small fraction of the elements in the dual Hessian.

## 1.2 Notations

The notation  $\mathbb{R}_{++}^{n \times m}$  (and  $\mathbb{R}_+^{n \times m}$ ) defines the set of positive definite (and positive semidefinite) matrices with dimension  $n \times m$ . The set of symmetric matrices of dimension  $n \times n$  is denoted by  $\mathbb{S}^n$ .

For a function  $f$  with two arguments,  $t \in \mathbb{R}$  and  $x \in \mathbb{R}^n$ , we denote the partial derivative of  $f(x, t)$  w.r.t.  $t$  as  $f'(x, t)$  and w.r.t.  $x$  as  $\nabla f(x, t)$ .

## 2 Preliminaries

In this section, we present the considered class of problems and recall the second-order dual decomposition approach.

### 2.1 Problem formulation

In this section, we present the class of distributed Quadratic Programs (QPs) considered in this paper. Moreover, we show how problems belonging to this class arise when MPC is used to control multiple linear systems interacting via non-delayed couplings.

Let us consider  $M$  discrete time, possibly time varying, linear systems:

$$x_{k,i+1} = A_{k,i}x_{k,i} + B_{k,i}u_{k,i}, \quad k = 1, \dots, M \quad (1)$$

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  $k$  at time instance  $i$ . Moreover, each system is subject to local state and input constraints:

$$x_{k,i} \in \mathcal{X}_{k,i} \subseteq \mathbb{R}^{p_k}, \quad k = 1, \dots, M \quad (2a)$$

$$u_{k,i} \in \mathcal{U}_{k,i} \subseteq \mathbb{R}^{m_k}, \quad k = 1, \dots, M \quad (2b)$$

We assume, in this paper, that the systems interact with static, non-delayed couplings, i.e.:

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

Observe that with non-delayed couplings, we refer to interactions where states and inputs for a subsystem at time  $i$  can only have a direct impact on states and inputs of another subsystem at time  $i$ . The interactions between the systems

can accordingly not include dynamic relationships or time-delays. However, it should be understood that this can always be fulfilled by introducing extra states to handle the delays locally, even when delayed couplings are present in the original formulation.

Furthermore, let us assume that the constraints (2) are convex polytopes, and that the non-delayed couplings (3) are affine. In the following, we use the notation:  $z_{k,i} = [x_{k,i}^T \ u_{k,i}^T]^T \in \mathbb{R}^{n_k}$ , representing the optimization variables. Using a quadratic objective function, the *finite-time optimal control problem* over the horizon  $N$ , can then be stated as:

$$\min_z \quad \sum_{k=1}^M \sum_{i=0}^N \frac{1}{2} z_{k,i}^T H_{k,i} z_{k,i} + c_{k,i}^T z_{k,i} \quad (4a)$$

$$\text{s.t.} \quad \sum_{k=1}^M F_{k,i} z_{k,i} = e_i \quad (4b)$$

$$C_{k,i} z_{k,i} + D_{k,i+1} z_{k,i+1} = d_{k,i} \quad (4c)$$

$$G_{k,i} z_{k,i} \leq f_{k,i} \quad (4d)$$

where  $H_{k,i} \in \mathbb{S}_{++}^{n_k}$  and  $c_{k,i} \in \mathbb{R}^{n_k}$ , the dynamics (1) are described by  $C_{k,i} \in \mathbb{R}^{l_k \times n_k}$ ,  $D_{k,i+1} \in \mathbb{R}^{l_k \times n_k}$  and  $d_{k,i} \in \mathbb{R}^{l_k}$ ,  $F_{k,i} \in \mathbb{R}^{r_i \times n_k}$  and  $e_i \in \mathbb{R}^{r_i}$  yield the non-delayed coupling constraints (3),  $G_{k,i} \in \mathbb{R}^{t_k \times n_k}$  and  $f_{k,i} \in \mathbb{R}^{t_k}$  form the local state and input constraints (2).

Additionally, to avoid unnecessarily heavy notation, we introduce the following augmented notations:  $z_k = [z_{k,0}^T \dots z_{k,N}^T]^T \in \mathbb{R}^{(N+1)n_k}$ , which represents the collection of states and inputs of system  $k$  over the horizon  $N$ . The MPC problem (4) can then be expressed as:

$$\min_z \quad \sum_{k=1}^M \frac{1}{2} z_k^T H_k z_k + c_k^T z_k \quad (5a)$$

$$\text{s.t.} \quad \sum_{k=1}^M F_k z_k = e \quad (5b)$$

$$C_k z_k = d_k \quad (5c)$$

$$G_k z_k \leq f_k \quad (5d)$$

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  $f_k \in \mathbb{R}^{(N+1)t_k}$ .

The matrices in (5) possess banded structures given by:

$$H_k = \begin{bmatrix} H_{k,0} & & \\ & \ddots & \\ & & H_{k,N} \end{bmatrix}, \quad (6a)$$

$$F_k = \begin{bmatrix} F_{k,0} & & \\ & \ddots & \\ & & F_{k,N} \end{bmatrix}, \quad (6b)$$

$$C_k = \begin{bmatrix} C_{k,0} & D_{k,1} & & & \\ & C_{k,1} & D_{k,2} & & \\ & & \ddots & \ddots & \\ & & & C_{k,N-1} & D_{k,N} \end{bmatrix}, \quad (6c)$$

$$A_k = \begin{bmatrix} A_{k,0} & & \\ & \ddots & \\ & & A_{k,N} \end{bmatrix}. \quad (6d)$$

These banded structures will be key elements for the quasi-banded structure of the dual Hessian.

## 2.2 Dual decomposition with second-order information

In this section, we introduce the second-order dual decomposition approach. The aim is to present a constraint relaxation, first described in [26], in order to make the dual function self-concordant. Self-concordance is an important property for constructing polynomial time interior-point methods [27], and is hence a desirable feature since Newton steps in the dual space are considered. We detail next the construction of the dual Hessian, which is needed in order to compute a Newton step.

We introduce the dual variables  $\lambda \in \mathbb{R}^{(N+1)r}$  corresponding to the coupling constraints (5b) and define the Lagrange function as:

$$\mathcal{L}(z, \lambda) = \sum_{k=1}^M \left( \frac{1}{2} z_k^T H_k z_k + c_k^T z_k \right) + \lambda^T \left( \sum_{k=1}^M F_k z_k - e \right) \quad (7)$$

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

$$\mathcal{L}(z, \lambda) = \sum_{k=1}^M \mathcal{L}_k(z_k, \lambda), \quad (8)$$

with:

$$\mathcal{L}_k(z_k, \lambda) = \frac{1}{2} z_k^T H_k z_k + c_k^T z_k + \lambda^T \left( F_k z_k - \frac{1}{M} e \right), \quad (9)$$

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

$$d(\lambda) = -\sum_{k=1}^M \min_{z_k \in \mathcal{Z}_k} \mathcal{L}_k(z_k, \lambda) \quad (10)$$

Since (5) is strictly convex,  $d(\lambda)$  is convex and continuously differentiable, but not twice differentiable [19]. However, the Hessian of  $d(\lambda)$  is a constant matrix over any fixed active-set of (4) [19].

As a result,  $d(\lambda)$  is not self-concordant and the solution to the dual problem is not easily tracked with the Newton method. However, if the inequality constraints (5d) are relaxed with a self-concordant log-barrier, according to:

$$\min_z \quad \sum_{k=1}^M \frac{1}{2} z_k^T H_k z_k + c_k^T z_k - \tau \sum_{i=1}^{(N+1)t_k} \log([s_k]_i) \quad (11a)$$

$$\text{s.t.} \quad \sum_{k=1}^M F_k z_k = e \quad (11b)$$

$$C_k z_k = d_k \quad (11c)$$

$$G_k z_k + s_k = f_k \quad (11d)$$

where  $\tau > 0$  will be referred to as the *barrier parameter*, the resulting *relaxed dual function*  $d(\lambda, \tau)$  is self-concordant [26]. As a result, a sequence of dual problems  $\{\min_{\lambda} d(\lambda, \tau)\}_{\tau \rightarrow 0}$  can be solved, where each problem is self-concordant and therefore easily handled with the Newton method. The *relaxed dual function* can be made to be an arbitrary good approximation of (10) by making  $\tau$  arbitrary small.

In contrast to smoothing techniques based on the *augmented Lagrangian* [17], the separability of the dual function is preserved, and the relaxed dual function can be computed in parallel using:

$$d(\lambda, \tau) = -\sum_{k=1}^M \min_{z_k \in \mathcal{Z}_k} \left( \mathcal{L}_k(z_k, \lambda) - \tau \sum_{i=1}^{m_k} \log([s_k]_i) \right) \quad (12)$$

Observe that evaluating (12) involves solving  $M$  independent subproblems of the form:

$$\begin{aligned} \min_{z_k} \quad & \frac{1}{2} z_k^T H_k z_k + c_k^T z_k + \lambda^T F_k z_k - \tau \sum_{i=1}^{m_k} \log([s_k]_i) \\ \text{s.t.} \quad & C_k z_k = d_k \\ & G_k z_k + s_k = f_k \\ & s_k \geq 0 \end{aligned} \quad (13)$$

Strict convexity of (11) implies that the gradient of  $d(\lambda, \tau)$  with respect to  $\lambda$  is given by the residual of the coupling constraints [3], i.e.:

$$\nabla d(\lambda, \tau) = -\sum_{k=1}^M F_k z_k^*(\lambda, \tau) + e \quad (14)$$

where  $z_k^*(\lambda, \tau)$  is the solution to (13). The dual Hessian is then given by:

$$\nabla^2 d(\lambda, \tau) = - \sum_{k=1}^M F_k \frac{\partial z_k^*(\lambda, \tau)}{\partial \lambda} \quad (15)$$

Recall that in order to find the solution  $\lambda^*(\tau)$  to the dual problem  $\min_{\lambda} d(\lambda, \tau)$ ,  $\lambda$  is updated by taking Newton steps  $\Delta\lambda$ . A Newton step in the dual space can be obtained as a solution to the Newton system:

$$\nabla^2 d(\lambda, \tau) \Delta\lambda + \nabla d(\lambda, \tau) = 0 \quad (16)$$

Finding the Newton step requires knowledge of the gradient and the Hessian of the dual function  $d(\lambda, \tau)$ . Note that the dual gradient (14) can be easily computed from the solutions  $z_k^*(\lambda, \tau)$  to (13), while the Hessian (15) requires the sensitivities  $\frac{\partial z_k^*(\lambda, \tau)}{\partial \lambda}$ . In the following, we will describe how  $\frac{\partial z_k^*(\lambda, \tau)}{\partial \lambda}$  can be computed, for a detailed description we refer to [16].

By introducing  $y_k = \tau s_k^{-1} \in \mathbb{R}^{t_k(N+1)}$ , the primal-dual interior point KKT conditions to (13) can be stated as:

$$0 = H_k z_k^* + c_k + F_k^T \lambda + C_k^T \mu_k^* + G_k^T y_k^* \quad (17a)$$

$$0 = C_k z_k^* - d_k \quad (17b)$$

$$0 = G_k z_k^* + s_k^* - f_k \quad (17c)$$

$$0 = Y_k^* s_k^* - \tau \mathbf{1} \quad (17d)$$

$$s_k^* > 0, \quad y_k^* > 0 \quad (17e)$$

By differentiating (17), the following linear system is obtained [16]:

$$\begin{bmatrix} H_k & C_k^T & G_k^T & 0 \\ C_k & 0 & 0 & 0 \\ G_k & 0 & 0 & I \\ 0 & 0 & S_k & Y_k \end{bmatrix} \begin{bmatrix} \frac{\partial z_k^*}{\partial \lambda} \\ \frac{\partial \mu_k^*}{\partial \lambda} \\ \frac{\partial y_k^*}{\partial \lambda} \\ \frac{\partial s_k^*}{\partial \lambda} \end{bmatrix} = - \begin{bmatrix} F_k^T \\ 0 \\ 0 \\ 0 \end{bmatrix} \quad (18)$$

From the primal sensitivity,  $\frac{\partial z_k^*(\lambda, \tau)}{\partial \lambda}$ , the dual Hessian can then be formed according to (15).

### 3 Quasi-structure of the dual Hessian

In this section, we show that the dual Hessian decays exponentially towards its off-diagonal corners, and hence possess a quasi-banded structure. We derive theoretical bounds on the magnitude of its elements. Moreover, we propose a quasi-Newton method based on a banded approximation of the dual Hessian.



### 3.1 Quasi-banded

In Section 2.2, we presented a way of computing the dual Hessian. In this section, we will detail its properties by forming an explicit expression and characterizing its quasi-banded structure. The quasi-banded structure allows to approximating the dual Hessian by a banded matrix, for which structure exploiting factorization methods can be used.

The analysis of the structure will be based on the explicit expression of the dual Hessian, stated in the following proposition:

**Proposition 1** *The Hessian of the dual function  $d(\lambda, \tau)$  is given by:*

$$\nabla^2 d(\lambda, \tau) = \sum_{k=1}^M F_k (\Phi_k^{-1} - \Phi_k^{-1} C_k^T \Lambda_k^{-1} C_k \Phi_k^{-1}) F_k^T \quad (19)$$

where

$$\Phi_k = H_k + G_k^T S_k^{-1} Y_k G_k \quad (20)$$

$$\Lambda_k = C_k \Phi_k^{-1} C_k^T \quad (21)$$

*Proof* By using block elimination of (18), it can be obtained that:

$$\Lambda_k(s_k, y_k) \frac{\partial \mu_k^*}{\partial \lambda} = -C_k \Phi_k^{-1} F_k^T \quad (22a)$$

$$\Phi_k(s_k, y_k) \frac{\partial z_k^*}{\partial \lambda} = -F_k^T - C_k^T \frac{\partial \mu_k^*}{\partial \lambda} \quad (22b)$$

where

$$\Phi_k(s_k, y_k) = H_k + G_k^T S_k^{-1} Y_k G_k \in \mathbb{S}_{++}^{(N+1)n_k \times (N+1)n_k} \quad (23a)$$

$$\Lambda_k(s_k, y_k) = C_k \Phi_k^{-1}(s_k, y_k) C_k^T \in \mathbb{S}_{++}^{Nl_k \times Nl_k} \quad (23b)$$

By using (22a) to eliminate  $\frac{\partial \mu_k^*}{\partial \lambda}$  in (22b), the following is obtained:

$$F_k \frac{\partial z_k^*}{\partial \lambda} = -F_k (\Phi_k^{-1} - \Phi_k^{-1} C_k^T \Lambda_k^{-1} C_k \Phi_k^{-1}) F_k^T \quad (24)$$

where we, for the sake of brevity, have omitted the arguments and used  $\Phi_k = \Phi_k(s_k, y_k)$  and  $\Lambda_k = \Lambda_k(s_k, y_k)$ . As a consequence, the dual Hessian (15) can be written as

$$\nabla^2 d(\lambda, \tau) = \sum_{k=1}^M F_k (\Phi_k^{-1} - \Phi_k^{-1} C_k^T \Lambda_k^{-1} C_k \Phi_k^{-1}) F_k^T \quad (25)$$

□

Observe, that the relaxed dual Hessian of a problem of the form (5) can always be written in the form (19), regardless of the specific structure described by (4). However, by considering the specific structure in (4), it can be observed from (20) that  $\Phi_k$  is block diagonal with blocks  $\Phi_{k,i} = H_{k,i} + G_{i,k}^T S_{k,i}^{-1} Y_{k,i} G_{k,i}$ . Moreover, since the structure of  $\Lambda_k$  will be instrumental for our results, we establish it in the following proposition:

**Proposition 2** *Provided the structure in (4),  $\Lambda_k$  has a block tridiagonal structure given by:*

$$\Lambda_k = \begin{bmatrix} \Lambda_{11} & \Lambda_{12} & & & \\ \Lambda_{12}^T & \Lambda_{22} & \ddots & & \\ & \ddots & \ddots & \ddots & \\ & & & \Lambda_{N-1,N} & \\ & & & \Lambda_{N-1,N}^T & \Lambda_{N,N} \end{bmatrix} \in \mathbb{S}_{++}^{Nl_k \times Nl_k} \quad (26)$$

where

$$\Lambda_{i,i} = C_{k,i-1} \Phi_{k,i-1}^{-1} C_{k,i-1}^T + D_{k,i} \Phi_{k,i}^{-1} D_{k,i}^T \in \mathbb{R}^{l_k \times l_k} \quad (27a)$$

$$\Lambda_{i,i+1} = D_{k,i} \Phi_{k,i}^{-1} C_{k,i}^T \in \mathbb{R}^{l_k \times l_k} \quad (27b)$$

*Proof* Recall that  $\Lambda_k$  is given by (23b), i.e.:

$$\Lambda_k = C_k \Phi_k^{-1} C_k^T \quad (28)$$

Due to the block diagonal structure of  $\Phi_k$ ,  $\Phi_k^{-1}$  is block diagonal with blocks  $\Phi_{k,i}^{-1}$ . Moreover, recall the block bidiagonal structure of  $C_k$  from (6c). Accordingly, by performing the multiplication in (28), the results given in (26) and (27) are obtained.  $\square$

Note that all matrices in (19) are banded except  $\Lambda_k^{-1}$  which in general is dense, and consequently, makes also the dual Hessian dense. However, inverses of sparse matrices typically contain many elements that are small in magnitude [2]. In the following, we will show that  $\Lambda_k^{-1}$  possess a decaying property which will translate to a *quasi-banded structure* in  $\nabla^2 d(\lambda, \tau)$ . In the analysis we consider two mild assumptions. First, we assume boundedness of the problem data:

**Assumption 1** *The row and column absolute sums of Jacobians of equality constraints (i.e. (5b) and (5c)) are bounded. Hence:*

1.  $\|C_{k,i}\|_p \leq \gamma$
2.  $\|D_{k,i}\|_p \leq \gamma$
3.  $\|F_{k,i}\|_p \leq \gamma$

where  $p = 1, \infty$ .

For the sake of simplicity, a single constant  $\gamma$  is chosen to bound all problem data. It should be observed that Assumption 1 is not by any means restrictive, since a problem having unbounded data would be ill-posed. Furthermore, we assume boundedness of  $\Phi_k^{-1}$ :

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

$$\|\Phi_{k,i}^{-1}\|_p \leq \gamma_{\Phi_k^{-1}}, \quad \forall i \quad (29)$$

where  $p = 1, \infty$ .

*Remark 1* As a result of  $\Phi_k$  being symmetric,  $\|\Phi_{k,i}^{-1}\|_\infty = \|\Phi_{k,i}^{-1}\|_1$ .

A possible difficulty with interior-point methods stems from numerical difficulties due to ill-conditioning of the normal equations [35]. Indeed, since elements in  $S_k^{-1}Y_k$  can be close to zero in late iterations when  $\tau$  is small. However, since we assume that (4) is strongly convex, the eigenvalues of  $\Phi_k$  are 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  $\Phi_k^{-1}$ .

To characterize the structure in  $\Lambda_k^{-1}$ , we will use that inverses of sparse matrices are in general dense, but typically contain many elements that are small in magnitude [2]. Since  $\Lambda_k$  is banded, symmetric and positive definite, we will use the following classical result:

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

$$|[A^{-1}]_{ij}| < K\omega^{|i-j|}, \quad \forall i, j \quad (30)$$

where  $\sigma_{\min}(A)$  and  $\kappa(A)$  are the the smallest singular value and the condition number of  $A$  respectively,  $K = \max\{\sigma_{\min}^{-1}(A), K_0\}$ ,  $K_0 = (1 + \sqrt{\kappa(A)})$ ,  $\omega = \left(\frac{\sqrt{\kappa(A)}-1}{\sqrt{\kappa(A)}+1}\right)^{1/m}$ .

*Proof* A proof is given in [6]. □

Lemma 1 essentially states that the entries in  $A^{-1}$  are bounded by an exponentially decaying function along each row or column. 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 therefore result in a large  $K$  and  $\omega \approx 1$ , leading to a slow decay. A low condition number and a narrow band results in a rapid decay. Because of its block tridiagonal structure,  $\Lambda_k$  is  $3l_k$ -banded, i.e.  $m = 3l_k$ .

Accordingly, if we introduce the notation

$$\Lambda_k^{-1} = \begin{bmatrix} T_{11} & T_{21}^T & \cdots & T_{N+1,1}^T \\ T_{21} & T_{22} & \cdots & T_{N+1,2}^T \\ \vdots & \vdots & \ddots & \vdots \\ T_{N+1,1} & T_{N+1,2} & \cdots & T_{N+1,N+1} \end{bmatrix} \quad (31)$$

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

**Proposition 3** *The off-diagonal blocks (i.e.  $T_{i,j}$  where  $i - j > 0$ ) in  $\Lambda_k^{-1}$  satisfy the following bounds:*

$$\|T_{i,j}\|_p \leq K_{\Lambda_k} \omega_{\Lambda_k}^{i-j} \quad (32)$$

where  $p = 1, \infty$ ,  $\sigma_{\min}(\Lambda_k)$  and  $\kappa(\Lambda_k)$  are the the smallest singular value and the condition number of  $\Lambda_k$  respectively,  $K_{\Lambda_k} = \max\{\sigma_{\min}(\Lambda_k)^{-1}, 1 + \sqrt{\kappa(\Lambda_k)}\} l_k \omega_{\Lambda_k}^{1/l_k}$  and  $\omega_{\Lambda_k} = \left( \frac{\sqrt{\kappa(\Lambda_k)} - 1}{\sqrt{\kappa(\Lambda_k)} + 1} \right)^{\frac{1}{3}}$ .

*Proof* **Emil:** Still sketchy

According to Lemma 1, the element in  $T_{i,j}$  with the largest bound is located in the top-right corner, and is hence the element  $[\Lambda_k^{-1}]_{il_k+1, j l_k}$  in  $\Lambda_k^{-1}$ . Moreover, recall that  $\Lambda_k$  is  $3l_k$ -banded. By directly applying Lemma 1 it follows that:

$$\max |[T_{i,j}]| \leq \max\{\sigma_{\min}(\Lambda_k)^{-1}, 1 + \sqrt{\kappa(\Lambda_k)}\} \left( \frac{\sqrt{\kappa(\Lambda_k)} - 1}{\sqrt{\kappa(\Lambda_k)} + 1} \right)^{\frac{1}{3l_k}((i-j)l_k+1)} \quad (33)$$

where  $\max |[T_{i,j}]|$  refers to the maximum absolute value of the entries in  $T_{i,j}$ . Moreover, since there are  $l_k$  elements in each row or column of a block  $T_{i,j}$ , we obtain:

$$\|T_{i,j}\|_p \leq l_k \max\{\sigma_{\min}(\Lambda_k)^{-1}, 1 + \sqrt{\kappa(\Lambda_k)}\} \left( \frac{\sqrt{\kappa(\Lambda_k)} - 1}{\sqrt{\kappa(\Lambda_k)} + 1} \right)^{\frac{1}{3l_k}((i-j)l_k+1)} \quad (34)$$

where  $p = 1, \infty$ . Observe that (34) is identical to (32).  $\square$

*Remark 2* The constants  $K_{\Lambda_k}$  and  $\omega_{\Lambda_k}$  in (32) depend heavily on the conditioning of  $\Lambda_k$ , which for interior point methods can be high for small values of  $\tau$ . Our experience is however that this is not a major issue in practice. This observation is also supported by commercial interior-point implementations [35] [7]. Moreover, it should be understood that strong convexity of (4), should improve the worst case conditioning of  $\Phi_k$  and hence also of  $\Lambda_k$ .

We have established that  $\Lambda_k^{-1}$  is decaying exponentially towards its off-diagonal corners. Moreover, since  $C_k$  is block bi-banded, we will see next that  $C_k^T \Lambda_k^{-1} C_k$ , which enters in the expression of  $\nabla^2 d(\lambda, \tau)$ , is also decaying exponentially towards the off-diagonal corners. Let us introduce the notation:

$$C_k^T \Lambda_k^{-1} C_k = \begin{bmatrix} V_{11} & V_{21}^T & \cdots & V_{N+1,1}^T \\ V_{21} & V_{22} & \cdots & V_{N+1,2}^T \\ \vdots & \vdots & \ddots & \vdots \\ V_{N+1,1} & V_{N+1,2} & \cdots & V_{N+1,N+1} \end{bmatrix} \quad (35)$$

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 4** *The off diagonal blocks (i.e.  $V_{i,j}$  where  $i-j > 0$ ) of  $C_k^T \Lambda_k^{-1} C_k$  are bounded by:*

$$\|V_{i,j}\|_p \leq \gamma^2 \bar{K}_{\Lambda_k} \omega_{\Lambda_k}^{i-j} \quad (36a)$$

$$(36b)$$

where  $\bar{K}_{\Lambda_k} = (2 + \omega_{\Lambda_k} + \omega_{\Lambda_k}^{-1})K_{\Lambda_k}$  and  $p = 1, \infty$ .

*Proof* By performing the multiplication  $C_k^T \Lambda_k^{-1} C_k$ , we find that:

$$\begin{aligned} V_{i,j} = & D_{k,i-1}^T T_{i,j-1} D_{k,j-1} + C_{k,i-1}^T T_{i-1,j-1} D_{k,j-1} + \\ & + D_{k,i-1}^T T_{i-1,j} C_{k,j-1} + C_{k,i-1}^T T_{i,j} C_{k,j-1} \end{aligned} \quad (37)$$

for  $i-j > 0$ . Using the Cauchy-Schwarz inequality, Assumption 1 and (32), each term in (37) can be bounded according to:

$$\|D_{k,i-1}^T T_{i,j-1} D_{k,j-1}\|_p \leq \gamma^2 K_{\Lambda_k} \omega_{\Lambda_k}^{i-j+1} \quad (38a)$$

$$\|C_{k,i-1}^T T_{i-1,j-1} D_{k,j-1}\|_p \leq \gamma^2 K_{\Lambda_k} \omega_{\Lambda_k}^{i-j} \quad (38b)$$

$$\|D_{k,i-1}^T T_{i-1,j} C_{k,j-1}\|_p \leq \gamma^2 K_{\Lambda_k} \omega_{\Lambda_k}^{i-j-1} \quad (38c)$$

$$\|C_{k,i-1}^T T_{i,j} C_{k,j-1}\|_p \leq \gamma^2 K_{\Lambda_k} \omega_{\Lambda_k}^{i-j} \quad (38d)$$

where  $p = 1, \infty$ . Hence, we can establish the following bound:

$$\|V_{i,j}\|_p \leq \gamma^2 (2 + \omega_{\Lambda_k} + \omega_{\Lambda_k}^{-1}) K_{\Lambda_k} \omega_{\Lambda_k}^{i-j} \quad (39)$$

□

Let us now introduce the notation:

$$\nabla^2 d(\lambda, \tau) = \begin{bmatrix} W_{11} & W_{21}^T & \cdots & W_{N+1,1}^T \\ W_{21} & W_{22} & \cdots & W_{N+1,2}^T \\ \vdots & \vdots & \ddots & \vdots \\ W_{N+1,1} & W_{N+1,2} & \cdots & W_{N+1,N+1} \end{bmatrix} \quad (40)$$

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

$$\|W_{i,j}\|_p \leq \sum_{k=1}^M \gamma^4 \gamma_{\Phi_k}^{-2} \bar{K}_{\Lambda_k} \omega_{\Lambda_k}^{i-j} \quad (41)$$

where  $p = 1, \infty$ .

*Proof* The dual Hessian is given by (19), i.e.:

$$\nabla^2 d(\lambda, \tau) = \sum_{k=1}^M F_k (\Phi_k^{-1} - \Phi_k^{-1} C_k^T \Lambda_k^{-1} C_k \Phi_k^{-1}) F_k^T \quad (42)$$

The matrices  $F_k \Phi_k^{-1} F_k^T$  are block-diagonal and do not contribute to the off-diagonal blocks in  $\nabla^2 d(\lambda, \tau)$ . Therefore using (40) and (42), for  $i - j > 0$ :

$$\|W_{i,j}\|_p = \left\| \sum_{k=1}^M F_{k,i} \Phi_{k,i}^{-1} V_{i,j} \Phi_{k,j}^{-1} F_{k,j} \right\|_p \leq \sum_{k=1}^M \left\| F_{k,i} \Phi_{k,i}^{-1} V_{i,j} \Phi_{k,j}^{-1} F_{k,j} \right\|_p \quad (43)$$

where  $p = 1, \infty$ . Moreover, by using Cauchy-Schwarz inequality, Assumption 1, 2 and Proposition 4, it can be verified that:

$$\|W_{i,j}\|_p \leq \sum_{k=1}^M \gamma^2 \gamma_{\Phi_k}^2 \|V_{i,j}\|_p \leq \sum_{k=1}^M \gamma^4 \gamma_{\Phi_k}^2 \bar{K}_{\Lambda_k} \omega_{\Lambda_k}^{i-j} \quad (44)$$

□

Proposition 5 concludes the characterization of the quasi-banded structure of  $\nabla^2 d(\lambda, \tau)$ .

### 3.2 A banded approximation of the dual Hessian

In Section 3.1, we established that the dual Hessian decays exponentially towards its off-diagonal corners. This suggests that a banded matrix can be used to form a good approximation of the dual Hessian, and accordingly be used as a surrogate of the exact dual Hessian in a quasi-Newton method. In this section, we will approximate the dual Hessian with its diagonal band, and characterize the error in the approximation.

In the following, we use the notation  $[\nabla^2 d(\lambda, \tau)]_{\mathcal{M}}$  for the diagonal  $\mathcal{M}$ -block band of (40), which will be referred to as the *clipped dual Hessian*. We recall next Gershgorin's circle theorem, which will be instrumental in the following development:

**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 [15]. □

Using Gershgorin's theorem, we can establish a bound on the distance between the dual Hessian and its clipped counterpart:

**Lemma 2** *Let  $\mathcal{M}$  be an odd integer in the interval  $[1, 2N + 1]$ , then the following bound holds:*

$$\|[\nabla^2 d(\lambda, \tau)]_{\mathcal{M}} - \nabla^2 d(\lambda, \tau)\|_2 \leq \left(N - \frac{\mathcal{M} - 1}{2}\right) \sum_{k=1}^M \gamma^4 \gamma_{\Phi_k}^2 \tilde{K}_{\Lambda_k} \omega_{\Lambda_k}^{\frac{\mathcal{M}-1}{2}} \quad (45)$$

where  $\tilde{K}_{\Lambda_k} = \bar{K}_{\Lambda_k} \omega_{\Lambda_k} = (2 + \omega_{\Lambda_k} + \omega_{\Lambda_k}^{-1}) K_{\Lambda_k}$ .

*Proof* **Emil: Still sketchy**

Note that  $[\nabla^2 d(\lambda, \tau)]_{\mathcal{M}} - \nabla^2 d(\lambda, \tau)$  is symmetric, which implies that:

$$\|[\nabla^2 d(\lambda, \tau)]_{\mathcal{M}} - \nabla^2 d(\lambda, \tau)\|_2 = |\lambda_{\max}([\nabla^2 d(\lambda, \tau)]_{\mathcal{M}} - \nabla^2 d(\lambda, \tau))| \quad (46)$$

where  $\lambda_{\max}(\bullet)$  represents the largest eigenvalue of  $\bullet$ . We then use Gershgorin's circle theorem to bound the magnitude of the largest eigenvalue in (46). Since, the diagonal elements of  $[\nabla^2 d(\lambda, \tau)]_{\mathcal{M}} - \nabla^2 d(\lambda, \tau)$  are zero, all Gershgorin discs are centered at the origin. Since the maximum absolute row sum is equal to the largest Gershgorin radius, the following holds:

$$|\lambda_{\max}([\nabla^2 d(\lambda, \tau)]_{\mathcal{M}} - \nabla^2 d(\lambda, \tau))| \leq \|[\nabla^2 d(\lambda, \tau)]_{\mathcal{M}} - \nabla^2 d(\lambda, \tau)\|_{\infty} \quad (47)$$

Observe that each block row in  $[\nabla^2 d(\lambda, \tau)]_{\mathcal{M}} - \nabla^2 d(\lambda, \tau)$  has not more than  $N - \frac{\mathcal{M}-1}{2}$  nonzero blocks. Moreover, according to (41), the blocks  $W_{i,j}$  in  $[\nabla^2 d(\lambda, \tau)]_{\mathcal{M}} - \nabla^2 d(\lambda, \tau)$  with the largest bound on the maximum absolute row sum are located next to the diagonal  $\mathcal{M}$ -block band of zeros, i.e. where  $i - j = \frac{\mathcal{M}-1}{2} + 1$ . Hence, by using the width of the band and (41), we can establish that:

$$\|[\nabla^2 d(\lambda, \tau)]_{\mathcal{M}} - \nabla^2 d(\lambda, \tau)\|_{\infty} \leq \left(N - \frac{\mathcal{M} - 1}{2}\right) \sum_{k=1}^M \gamma^4 \gamma_{\Phi_k}^2 \bar{K}_{\Lambda_k} \omega_{\Lambda_k}^{\frac{\mathcal{M}-1}{2} + 1} \quad (48)$$

Accordingly, by using (46), (47) and (48), we can establish that:

$$\|[\nabla^2 d(\lambda, \tau)]_{\mathcal{M}} - \nabla^2 d(\lambda, \tau)\|_2 \leq \left(N - \frac{\mathcal{M} - 1}{2}\right) \sum_{k=1}^M \gamma^4 \gamma_{\Phi_k}^2 \bar{K}_{\Lambda_k} \omega_{\Lambda_k}^{\frac{\mathcal{M}-1}{2} + 1} \quad (49)$$

which is identical to (45).  $\square$

Observe that in (40),  $\nabla^2 d(\lambda, \tau)$  is partitioned into a block matrix with  $N + 1$  row partitions and  $N + 1$  column partitions. Accordingly, if  $\mathcal{M} = 2N + 1$ ,  $[\nabla^2 d(\lambda, \tau)]_{\mathcal{M}}$  is exactly the same as  $\nabla^2 d(\lambda, \tau)$ . This is reflected in the first factor of (45), which is zero for  $\mathcal{M} = 2N + 1$ .

It is well known that the use of an inexact Hessian will degrade the convergence of the Newton method to a linear rate provided that the inexact Hessian has full rank [28]. In the following, we will establish a criterion that guarantees a positive definite clipped dual Hessian. The criterion is based on the following proposition that follows from Weyl's inequality [15]:

**Proposition 6** *Let  $A, B \in \mathbb{S}^{n \times n}$ , then*

$$\lambda_{\min}(A) + \lambda_{\min}(B) \leq \lambda_{\min}(A + B) \quad (50)$$

where  $\lambda_{\min}$  represents the minimum eigenvalue.

*Proof* A proof is given in [15].

We can now establish a criterion for the non-singularity of  $\lfloor \nabla^2 d(\lambda, \tau) \rfloor_{\mathcal{M}}$ :

**Lemma 3** *If  $\mathcal{M}$  is chosen such that:*

$$\left(N - \frac{\mathcal{M} - 1}{2}\right) \sum_{k=1}^M \gamma^4 \gamma_{\Phi_k}^2 \tilde{K}_{A_k} \omega_{A_k}^{\frac{\mathcal{M}-1}{2}} < \lambda_{\min}(\nabla^2 d(\lambda, \tau)) \quad (51)$$

then  $\lfloor \nabla^2 d(\lambda, \tau) \rfloor_{\mathcal{M}}$  is positive definite.

*Proof* Observe that

$$\lfloor \nabla^2 d(\lambda, \tau) \rfloor_{\mathcal{M}} = \nabla^2 d(\lambda, \tau) + (\lfloor \nabla^2 d(\lambda, \tau) \rfloor_{\mathcal{M}} - \nabla^2 d(\lambda, \tau)) \quad (52)$$

Proposition 6 entails that:

$$\lambda_{\min}(\lfloor \nabla^2 d(\lambda, \tau) \rfloor_{\mathcal{M}}) \geq \lambda_{\min}(\nabla^2 d(\lambda, \tau)) + \lambda_{\min}(\lfloor \nabla^2 d(\lambda, \tau) \rfloor_{\mathcal{M}} - \nabla^2 d(\lambda, \tau)) \quad (53)$$

Accordingly,  $\lfloor \nabla^2 d(\lambda, \tau) \rfloor_{\mathcal{M}}$  is nonsingular if:

$$\lambda_{\min}(\nabla^2 d(\lambda, \tau)) > |\lambda_{\min}(\lfloor \nabla^2 d(\lambda, \tau) \rfloor_{\mathcal{M}} - \nabla^2 d(\lambda, \tau))| \quad (54)$$

We have already established that the eigenvalues of  $\lfloor \nabla^2 d(\lambda, \tau) \rfloor_{\mathcal{M}} - \nabla^2 d(\lambda, \tau)$  are located close to the origin and are not more negative than the largest Gershgorin radius given by (48). Hence, we can establish that:

$$\lambda_{\min}(\lfloor \nabla^2 d(\lambda, \tau) \rfloor_{\mathcal{M}} - \nabla^2 d(\lambda, \tau)) \geq - \left(N - \frac{\mathcal{M} - 1}{2}\right) \sum_{k=1}^M \gamma^4 \gamma_{\Phi_k}^2 \tilde{K}_{A_k} \omega_{A_k}^{\frac{\mathcal{M}-1}{2}} \quad (55)$$

Using (54) and (55) gives (51).  $\square$

*Remark 3* It should be noted that  $\lambda_{\min}(\nabla^2 d(\lambda, \tau)) > 0$ ,  $\forall \lambda$  [16], such that (51) can always be satisfied.

Observe that (51) is a sufficient and not a necessary condition. Moreover, according to our experiments, singularity is not a problem in practice, and (51) is hence very conservative.



### 3.3 Convergence of the quasi-Newton method

In the previous section, we have seen that the clipped dual Hessian is positive definite and close to the exact Hessian provided that the band  $\mathcal{M}$  is wide enough. In this section, we will consider issues related to the convergence and complexity of a quasi-Newton method based on the *clipped dual Hessian*.

Let us start by detailing how the clipped dual Hessian affect the convergence of the Newton method:

**Theorem 2** *If the dual function fulfills:*

$$\left\| \int_0^1 (\nabla^2 d(\lambda + t\Delta\lambda, \tau) - \nabla^2 d) dt \lfloor \nabla^2 d \rfloor_{\mathcal{M}}^{-1} \right\|_2 \leq \omega \|\nabla d\|_2 \quad (56)$$

*then the following bound holds:*

$$\|\nabla d(\lambda + \Delta\lambda, \tau)\|_2 \leq \omega \|\nabla d(\lambda, \tau)\|_2^2 + f(\mathcal{M}) \|\nabla d(\lambda, \tau)\|_2 \quad (57)$$

where

$$f(\mathcal{M}) = \|(\lfloor \nabla^2 d(\lambda, \tau) \rfloor_{\mathcal{M}} - \nabla^2 d(\lambda, \tau)) \lfloor \nabla^2 d(\lambda, \tau) \rfloor_{\mathcal{M}}^{-1}\|_2 \quad (58)$$

*Remark 4* We will further discuss assumption (56) in Proposition 7.

*Proof* An approximate Newton step based on the clipped dual Hessian is given by the following linear system:

$$\lfloor \nabla^2 d(\lambda, \tau) \rfloor_{\mathcal{M}} \Delta\lambda + \nabla d(\lambda, \tau) = 0 \quad (59)$$

In order to avoid heavy notations, we will omit arguments when they are obvious from the context and introduce the notation  $\nabla d^+ = \nabla d(\lambda + \Delta\lambda, \tau)$ . We then have:

$$\begin{aligned} \|\nabla d^+\|_2 &= \|\nabla d^+ - \nabla d - \lfloor \nabla^2 d \rfloor_{\mathcal{M}} \Delta\lambda\|_2 = \\ &= \left\| \int_0^1 (\nabla^2 d(\lambda + t\Delta\lambda, \tau) \Delta\lambda) dt - \lfloor \nabla^2 d \rfloor_{\mathcal{M}} \Delta\lambda \right\|_2 = \\ &= \left\| \left( \int_0^1 (\nabla^2 d(\lambda + t\Delta\lambda, \tau) - \nabla^2 d) dt + \nabla^2 d - \lfloor \nabla^2 d \rfloor_{\mathcal{M}} \right) \Delta\lambda \right\|_2 = \\ &= \left\| \left( \int_0^1 (\nabla^2 d(\lambda + t\Delta\lambda, \tau) - \nabla^2 d) dt + \nabla^2 d - \lfloor \nabla^2 d \rfloor_{\mathcal{M}} \right) \lfloor \nabla^2 d \rfloor_{\mathcal{M}}^{-1} \nabla d \right\|_2 \end{aligned} \quad (60)$$

where we have used (59) and the following result from calculus:

$$\int_0^1 \nabla^2 d(\lambda + t\Delta\lambda, \tau) \Delta\lambda dt = \nabla d^+ - \nabla d \quad (61)$$

Consequently, we can bound (60) according to:

$$\|\nabla d^+\|_2 \leq \omega \|\nabla d\|_2^2 + f(\mathcal{M}) \|\nabla d\|_2 \quad (62)$$

where we have introduced:

$$\left\| \int_0^1 (\nabla^2 d(\lambda + t\Delta\lambda, \tau) - \nabla^2 d) dt [\nabla^2 d]_{\mathcal{M}}^{-1} \right\|_2 \leq \omega \|\nabla d\|_2 \quad (63)$$

and

$$f(\mathcal{M}) = \|(\nabla^2 d - [\nabla^2 d]_{\mathcal{M}}) [\nabla^2 d]_{\mathcal{M}}^{-1}\|_2 \quad (64)$$

□

*Remark 5* Observe the  $f(\mathcal{M}) \geq 0$  is a function which is zero for  $\mathcal{M} = 2N + 1$ .

The condition (56) can be related to the Lipschitz continuity of the dual Hessian. This is established in the following:

**Proposition 7** *Let the dual Hessian be Lipschitz continuous with constant  $L$ , then:*

$$\omega = \frac{1}{2}L \|[\nabla^2 d(\lambda, \tau)]_{\mathcal{M}}^{-1}\|_2^2 \quad (65)$$

fulfills (56).

*Proof* By definition of Lipschitz continuity, the following criteria holds:

$$\|\nabla^2 d(\lambda + \Delta\lambda, \tau) - \nabla^2 d(\lambda, \tau)\|_2 \leq L\|\Delta\lambda\|_2 \quad (66)$$

Observe that:

$$\begin{aligned} & \left\| \int_0^1 (\nabla^2 d(\lambda + t\Delta\lambda, \tau) - \nabla^2 d) dt [\nabla^2 d]_{\mathcal{M}}^{-1} \right\|_2 \leq \\ & \leq \left\| \int_0^1 (\nabla^2 d(\lambda + t\Delta\lambda, \tau) - \nabla^2 d) dt \right\|_2 \|[\nabla^2 d]_{\mathcal{M}}^{-1}\|_2 \leq \\ & \leq \int_0^1 \|(\nabla^2 d(\lambda + t\Delta\lambda, \tau) - \nabla^2 d)\|_2 dt \|[\nabla^2 d]_{\mathcal{M}}^{-1}\|_2 \leq \\ & \leq \frac{1}{2}L \|[\nabla^2 d]_{\mathcal{M}}^{-1}\|_2 \|\Delta\lambda\|_2 \leq \frac{1}{2}L \|[\nabla^2 d]_{\mathcal{M}}^{-1}\|_2^2 \|\nabla d(\lambda, \tau)\|_2 \end{aligned} \quad (67)$$

□

In terms of iterations, the optimal choice would be to use the full dual Hessian and hence make  $f(\mathcal{M}) = 0$ . However, since a dense factorization method is more costly than a method exploiting sparsity, the overall complexity of solving a problem can be lower if a nonzero  $f(\mathcal{M})$  is used. As an illustrative example, the complexity of finding a Cholesky factorization of a dense matrix  $A \in \mathbb{S}_+^n$  is  $\frac{1}{3}n^3$  flops. On the other hand, if  $A$  is  $k$ -banded (where  $k \ll n$ ), the complexity is reduced to  $nk^2$  flops.

To further specify properties of  $f(\mathcal{M})$ , we state the following corollary that follows directly from Lemma 2:

**Corollary 1** *The following bound holds:*

$$f(\mathcal{M}) \leq \|\lfloor \nabla^2 d(\lambda, \tau) \rfloor_{\mathcal{M}}^{-1}\|_2 \left( N - \frac{\mathcal{M} - 1}{2} \right) \sum_{k=1}^M \gamma^4 \gamma_{\Phi_k}^2 \tilde{K}_{\Lambda_k} \omega_{\Lambda_k}^{\frac{\mathcal{M}-1}{2}} \quad (68)$$

*Proof* By using Cauchy Schwarz inequality, we have that:

$$f(\mathcal{M}) \leq \|\nabla^2 d - \lfloor \nabla^2 d \rfloor_{\mathcal{M}}\|_2 \|\lfloor \nabla^2 d \rfloor_{\mathcal{M}}^{-1}\|_2 \quad (69)$$

Moreover, by using (45), we can establish that

$$f(\mathcal{M}) \leq \|\lfloor \nabla^2 d(\lambda, \tau) \rfloor_{\mathcal{M}}^{-1}\|_2 \left( N - \frac{\mathcal{M} - 1}{2} \right) \sum_{k=1}^M \gamma^4 \gamma_{\Phi_k}^2 \tilde{K}_{\Lambda_k} \omega_{\Lambda_k}^{\frac{\mathcal{M}-1}{2}} \quad (70)$$

□

Accordingly, there is trade-off in the selection of  $\mathcal{M}$ . A small bandwidth of  $\lfloor \nabla^2 d(\lambda, \tau) \rfloor_{\mathcal{M}}$  results in more, but cheaper, iterations compared to a higher bandwidth.

#### 4 A brief note on Active-set methods

In previous sections, we have investigated the quasi-structure in the context of interior-point methods. In this section, we will make a short comment on the quasi-banded structure when the subproblems are solved with Active-set methods. Methods of this kind are proposed in [8], [18], [19].

In the context of Active-set methods, the *non-relaxed* dual function (10) is used, i.e.:

$$d(\lambda) = - \sum_{k=1}^M \min_{z_k \in \mathcal{Z}_k} \mathcal{L}_k(z_k, \lambda) \quad (71)$$

Observe that evaluating (71) involves solving  $M$  subproblems of the form:

$$\min_{z_k} \quad \frac{1}{2} z_k^T H_k z_k + c_k^T z_k + \lambda^T F_k z_k \quad (72a)$$

$$\text{s.t.} \quad C_k z_k = d_k \quad (72b)$$

$$G_k z_k \leq f_k \quad (72c)$$

Strict convexity of (72) implies that the gradient of  $d(\lambda)$  is given by [3]:

$$\nabla d(\lambda) = - \sum_{k=1}^M F_k z_k^*(\lambda) + e \quad (73)$$

where  $z_k^*(\lambda)$  is the solution to (72). Accordingly, the dual Hessian is given by:

$$\nabla^2 d(\lambda) = - \sum_{k=1}^M F_k \frac{\partial z_k^*(\lambda)}{\partial \lambda} \quad (74)$$

Observe that in order to compute the dual Hessian, the sensitivities  $\frac{\partial z_k^*(\lambda)}{\partial \lambda}$  are required. In the following, we will describe how  $\frac{\partial z_k^*(\lambda)}{\partial \lambda}$  can be computed.

Let us introduce the notation:

$$A_k z_k = b_k \quad (75)$$

for the augmentation of the equality constraints (72b) and the active inequality constraints (72c). Furthermore, assume that the rows of (75) are ordered to preserve a banded structure. The KKT conditions of (72) can then be stated as:

$$\begin{aligned} H_k z_k^*(\lambda) + c_k + F_k^T \lambda + A_k^T \mu_k^*(\lambda) &= 0 \\ A_k z_k^*(\lambda) - b_k &= 0 \end{aligned} \quad (76)$$

where the dual variables  $\mu_k$  corresponding to (75) have been introduced. By differentiating (76) with respect to  $\lambda$ , the following linear system is obtained:

$$\begin{bmatrix} H_k & A_k^T \\ A_k & 0 \end{bmatrix} \begin{bmatrix} \frac{\partial z_k^*(\lambda)}{\partial \lambda} \\ \frac{\partial \mu_k^*(\lambda)}{\partial \lambda} \end{bmatrix} = - \begin{bmatrix} F_k^T \\ 0 \end{bmatrix} \quad (77)$$

By using block elimination of (77), it can be obtained that:

$$\frac{\partial z_k^*(\lambda)}{\partial \lambda} = - (H_k^{-1} - H_k^{-1} A_k^T (A_k H_k^{-1} A_k^T)^{-1} A_k H_k^{-1}) C_k^T \quad (78)$$

Accordingly, by using (74) and (78), the dual Hessian can be expressed as:

$$\nabla^2 d(\lambda) = \sum_{k=1}^M C_k (H_k^{-1} - H_k^{-1} A_k^T (A_k H_k^{-1} A_k^T)^{-1} A_k H_k^{-1}) C_k^T \quad (79)$$

By comparing (79) and (19), it can be observed that the dual Hessian in the contexts of interior-point methods (19), is obtained if  $H_k$  in (79) is replaced by  $\Phi_k$ . Accordingly, a similar analysis as in Section 3 can be performed to show that the dual Hessian has a quasi-banded structure also when the subproblems are solved with active-set methods.

## 5 Numerical experiments

In this section, we aim at investigating the off-diagonal decay in numerical experiments. The parameter that seems to have the strongest impact on the quasi-structure of the dual Hessian is the horizon, where a long horizon generally render a strong decay. Therefore, we will illustrate the decaying properties with an example with a relatively long horizon ( $N = 30$ ). To illustrate the relative effects of horizon length, we will also show another example with similar dimensions but with a shorter horizon ( $N = 10$ ). There is an obvious decay also in this case, but the gain of using a banded factorization is smaller.

Preliminary numerical results indicate that the quasi-banded structure is consistent for different values of  $\tau$  and for different problems with the same dimensions.

### 5.1 A long horizon example

As an illustrative examples we will look at a randomly generated problem of the form (4), with  $M = 4$  subproblems each with a horizon  $N = 30$ , 6 states, 4 controls, and  $4(N + 1) = 124$  inequality constraints. At each time instance the subproblems are coupled via 4 coupling constraints, resulting in a dual space with  $4(N + 1) = 124$  dimensions. Moreover, the elements are generated with an even distribution in the interval  $[-1000, 1000]$ .

The absolute values of the elements in the dual Hessian at the solution  $\lambda^*(\tau)$  for  $\tau = 1$ , are visualized in Figure 1. It can be seen that elements with a large magnitude are present along the diagonal, while elements towards the off-diagonal corners are small in absolute value.

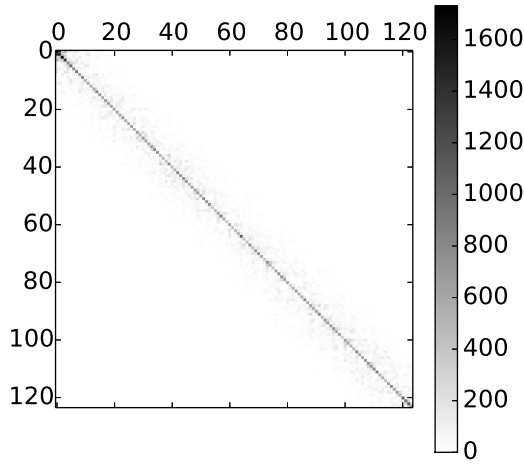


Fig. 1: Illustration of the dual Hessian evaluated at  $\lambda^*(\tau)$ , for  $\tau = 1$ .

In Section 3.3, it was proposed that the decaying property of the dual Hessian can be used to form a quasi-Newton method where the *clipped dual Hessian* is used. The convergence for an  $m$ -banded clipped dual Hessian, is visualized in Figure 2. Observe that even the widest band,  $m = 41$ , represents a narrow band compared to the full dual Hessian. Details of the structural zeros in the clipped dual Hessian are summarized in Table 1. Observe that, for  $m = 41$ , the clipped dual Hessian can be factorized by ignoring 69.7% of the elements in the full dual Hessian, at the cost of only a few extra iterations. For a smaller  $m$ , more elements can be ignored at the expense of performing more iterations.

The optimal trade-off between the number of iterations and the complexity of factorizing the clipped dual Hessian depends on many parameters. However, in general we can expect that it is beneficial to aim for few iterations when it is costly to solve the subproblems compared to factorizing the dual Hessian. On the contrary, when it is costly to factorize the dual Hessian, we can expect that a smaller  $m$  is beneficial.

Table 1: Overview of structural zeros in the clipped dual Hessian, for the long horizon example

	Full Hessian	$m = 41$	$m = 21$	$m = 11$
Structural zero elements	0	10712	12882	14042
Structural nonzero elements	15376	4664	2494	1334
Ratio of structural zeros	0%	69.7%	83.8%	91.3%

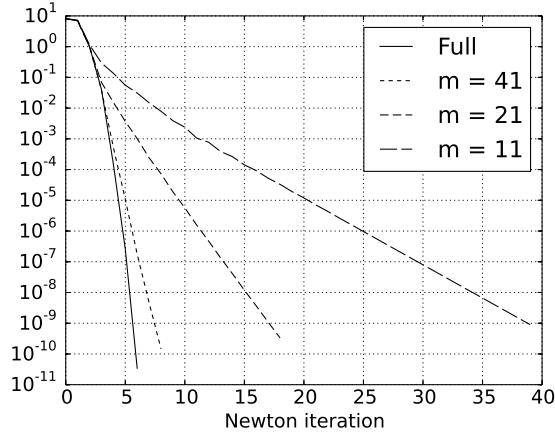


Fig. 2: Convergence of Newton and quasi-Newton methods for  $\tau = 1$ .

Let us investigate the consistency of the decay with respect to  $\lambda$ . The maximal absolute value of the elements in  $\nabla^2 d(\lambda, \tau)$  for given distances from its diagonal, are visualized in Figure 3 for 50 randomly selected values of  $\lambda$ . More specifically, the measure used in Figure 3 is:

$$\begin{aligned} \max_{i,j} \quad & |[\nabla^2 d(\lambda, \tau)]_{i,j}| \\ \text{s.t.} \quad & |i - j| = \text{constant} \end{aligned} \quad (80)$$

Observe that for this example, at all evaluated points  $\lambda$ , the dual Hessian has a clear decaying trend towards its off-diagonal corners. The same observation has been made on all tested problems.

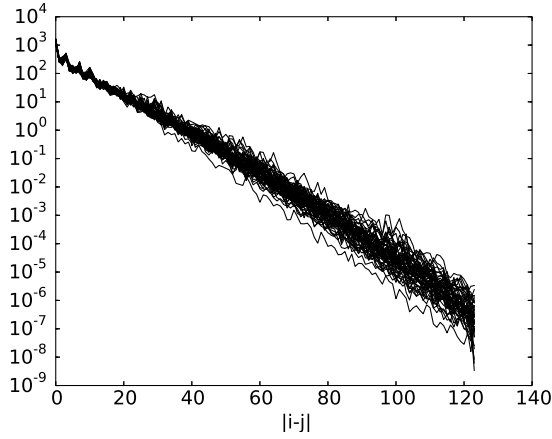


Fig. 3: Illustration of the off-diagonal decay. Each line corresponds to a different  $\lambda$ .

## 5.2 A short horizon example

The parameter that seems to have the strongest impact on the decaying is the horizon length. To illustrate this, we generate another random problem with  $N = 10$ , and all other dimensions being the same as before.

The absolute values of the elements in the dual Hessian at the solution  $\lambda^*(\tau)$ , for  $\tau = 1$ , are visualized in Figure 4. Indeed, there is an evident decay towards the off-diagonal corners, but the relative decay is weaker compared to Figure 1. As a consequence, the relative bandwidth of the clipped dual Hessian cannot be made as small before the convergence is seriously hindered.

As a comparison to Figure 2, the convergence for an  $m$ -banded clipped dual Hessian is visualized, for different values of  $m$ , in Figure 5. Observe that for a certain bandwidth  $m$ , the convergence is similar for the long and short horizon examples. The difference is, however, that the band is relatively wider in the short horizon case, and accordingly the gain in complexity for using a banded factorization method is smaller. Details of the structural zeros in the clipped dual Hessian for different values of  $m$  are summarized in Table 2.

Table 2: Overview of structural zeros in the clipped dual Hessian, for the short horizon example

	Full Hessian	$m = 41$	$m = 21$	$m = 11$
Structural zero elements	0	552	1122	1482
Structural nonzero elements	1936	1384	814	454
Ratio of structural zeros	0%	28.5%	58.0%	76.5%

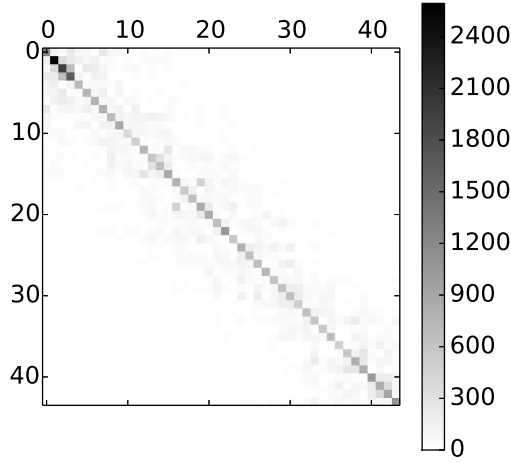


Fig. 4: Illustration of the dual Hessian evaluated at  $\lambda^*(\tau)$ , for  $\tau = 1$ .

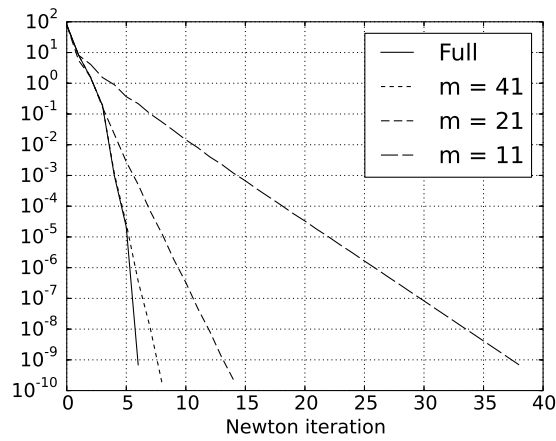


Fig. 5: Convergence of Newton and quasi-Newton methods for  $\tau = 1$ .

## 6 Conclusions

This paper is devoted to showing that the Hessian of the dual function originating from many distributed MPC problems possess a quasi-banded structure. Due to the quasi-structure, a banded approximation of the Hessian can be formed, for which structure exploiting factorization methods can be used. We show that the structure can be used to form a quasi-Newton method, which can



solve problems of the considered class to a lower overall complexity compared to a Newton method.

The quasi-structure is analyzed in the context of the interior-point methods presented in [26], [33] and [16]. However, we also show that the structure is present in active-set methods.

By numerical experiments, we showed that the quasi-structure is strong especially for MPC problems with a very long horizon. In those cases, it is often possible to find a good approximation of the Newton direction by only considering a small fraction of the elements in the dual Hessian. This should be very appealing for systems with fast and slow dynamics, e.g. power grids with storage capabilities.

## References

1. Alessio, A., Barcelli D. Bemporad, A.: Decentralized model predictive control of dynamically coupled linear systems. *Journal of Process Control* **21**, 705–714 (2011)
2. Benzi, M.: Preconditioning techniques for large linear systems: a survey. *Journal of Computational Physics* **182**, 418–477 (2002)
3. Bertsekas, D., Tsitsiklis, J.N.: *Parallel and distributed computation: Numerical methods*. Prentice Hall (1989)
4. Conte, C., Zeilinger, M., Morari, M., Jones, C.: Robust distributed model predictive control of linear systems. In: *European Control Conference* (2013)
5. Cottle, R.W., Duvall, S.G., Zikan, K.: A lagrangean relaxation algorithm for the constrained matrix problem. *Naval Research Logistics Quarterly* **33**(1), 55–76 (1986). DOI 10.1002/nav.3800330106. URL <http://dx.doi.org/10.1002/nav.3800330106>
6. Demko, S., Moss, W., Smith, P.: Decay rates for inverses of band matrices. *Mathematics of Computation* **43**, 491–499 (1986)
7. Domahidi, A., Zraggen, 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)
8. Ferreau, H., Kozma, A., Diehl, M.: A parallel active-set strategy to solve sparse parametric quadratic programs arising in MPC. In: *Proceedings of the 4th IFAC Nonlinear Model Predictive Control Conference*, Noordwijkerhout, The Netherlands (2012)
9. Frasch J. and Vukov, M., Ferreau H. and Diehl, M.: A new quadratic programming strategy for efficient sparsity exploitation in sqp-based nonlinear mpc and mhe. In: *Proceedings of the 19th IFAC World Congress* (2014)
10. Giselsson, P.: Improved dual decomposition for distributed model predictive control. In: *Proceedings of the 19th IFAC World Congress* (2014)
11. Giselsson, P., Doan, M.D., Keviczky, T., Schutter, B.D., Rantzer, A.: Accelerated gradient methods and dual decomposition in distributed model predictive control. *Automatica* **49**(3), 829 – 833 (2013). DOI <http://dx.doi.org/10.1016/j.automatica.2013.01.009>. URL <http://www.sciencedirect.com/science/article/pii/S0005109813000101>
12. Giselsson, P., Rantzer, A.: Distributed model predictive control with suboptimality and stability guarantees. In: *Decision and Control (CDC)*, 2010 49th IEEE Conference on, pp. 7272–7277 (2010). DOI 10.1109/CDC.2010.5717026
13. Giselsson P. and Boyd, S.: Preconditioning in fast dual gradient methods. In: *IEEE Conference on Decision and Control* (2014)
14. Gros, S.: A distributed algorithm for nmmpc-based wind farm control. In: *Conference on Decision and Control* (2014)
15. Horn, R.A., Johnson, C.R.: *Matrix Analysis* - 2nd ed. Cambridge University Press (2013)
16. Klintberg, E., Gros, S.: A primal-dual newton method for distributed quadratic programming. In: *Conference on Decision and Control* (2014)

17. Kozma, A., Conte, C., Diehl, M.: Benchmarking large scale distributed convex quadratic programming algorithms. *Optimization Methods & Software* (2014). (submitted for publication)
18. Kozma, A., Frasch, J.V., Diehl, M.: A Distributed Method for Convex Quadratic Programming Problems Arising in Optimal Control of Distributed Systems. In: *Proceedings of the 52nd Conference on Decision and Control (CDC)* (2013)
19. 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)
20. Kozma, A., Savorgnan, C., Diehl, M.: Distributed Model Predictive Control Made Easy, *Intelligent Systems, Control and Automation: Science and Engineering*, vol. Vol 69., chap. Distributed multiple shooting for large scale nonlinear systems, p. 604. Springer (2014)
21. Lasdon, L.S.: *Optimization theory for Large Systems*. Dover (1970)
22. Magni, L., Scattolini, R.: Stabilizing decentralized model predictive control of nonlinear systems. *Automatica* **42**, 1231–1236 (2006)
23. Necoara, I., Nedelcu, V.: Rate analysis of inexact dual first-order methods application to dual decomposition. *IEEE Trans. Automat. Control* **59**, 1232–1243 (2014)
24. Necoara, I., Nedelcu, V., Dumitrache, I.: Parallel and distributed optimization methods for estimation and control in networks. *Journal of Process Control* **21**(5), 756 – 766 (2011). DOI <http://dx.doi.org/10.1016/j.jprocont.2010.12.010>. URL <http://www.sciencedirect.com/science/article/pii/S095915241000257X>. Special Issue on Hierarchical and Distributed Model Predictive Control
25. Necoara, I., Patrascu, A.: Iteration complexity analysis of dual first order methods for convex programming. *Journal of Optimization Theory and Applications* (Submitted) (2014)
26. Necoara, I., Suykens, J.: An interior-point Lagrangian decomposition method for separable convex optimization. *J. Optim. Theory and Appl.* **143**(3), 567–588 (2009)
27. Nesterov, Y., Nemirovski, A.: *Interior-point Polynomial Algorithms in Convex Programming*. Society for Industrial Mathematics (1994)
28. Nocedal, J., Wright, S.: *Numerical Optimization*, 2 edn. Springer Series in Operations Research and Financial Engineering. Springer (2006)
29. Perizzato, A., Farina, M., Scattolini, R.: Stochastic distributed predictive control of independent systems with coupling constraints. In: *Conference on Decision and Control* (2014)
30. Rawlings, J., Stewart, B.: Coordinating multiple optimization-based controllers: New opportunities and challenges. *Journal of Process Control* **18**, 839–845 (2008)
31. Richter, S., Morari, M., Jones, C.: Towards computational complexity certification for constrained MPC based on Lagrange Relaxation and the fast gradient method. In: *Decision and Control and European Control Conference (CDC-ECC)*, 2011 50th IEEE Conference on, pp. 5223 –5229 (2011). DOI 10.1109/CDC.2011.6160931
32. Stewart, B., Rawlings, J., Wright, S.: Hierarchical cooperative distributed model predictive control. In: *American Control Conference* (2010)
33. Tran, D.Q., Necoara, I., Savorgnan, C., Diehl, M.: An inexact perturbed path-following method for lagrangian decomposition in large-scale separable convex optimization. *SIAM Journal on Optimization* **23**, 95–125 (2013)
34. Venkat, A., Rawlings, J., Wright, S.: Stability and optimality of distributed model predictive control. In: *Conference on Decision and Control* (2005)
35. Wright, M.: Ill-conditioning and computational error in interior methods for nonlinear programming. *SIAM Journal on Optimization* **9**, 84–111 (1998)