TECHNISCHE UNIVERSITÄT WIEN

Operations Research and Control Systems

SWM ORCOS

Inexact Newton-Kantorovich Methods for Constrained Nonlinear Model Predictive Control

Asen L. Dontchev, Mike Huang,
Ilya V. Kolmanovsky, Marco M. Nicotra

Research Report 2018-06

October 2018

ISSN 2521-313X

Operations Research and Control Systems
Institute of Statistics and Mathematical Methods in Economics
Vienna University of Technology

Research Unit ORCOS
Wiedner Hauptstraße 8 / E105-4
1040 Vienna, Austria
E-mail: orcos@tuwien.ac.at

Inexact Newton-Kantorovich Methods for Constrained Nonlinear Model Predictive Control*

Asen L. Dontchev[1], Mike Huang[2], Ilya V. Kolmanovsky[3], Marco M. Nicotra[3]

***Abstract*—In this paper we consider Newton-Kantorovich type methods for solving control-constrained optimal control problems that appear in model predictive control. Conditions for convergence are established for an inexact version of the Newton-Kantorovich method applied to variational inequalities. Based on these results, two groups of algorithms are proposed to solve the optimality system. The first group includes exact and inexact Newton and Newton-Kantorovich implementations of the Sequential Quadratic Programming. In the second group, exact and inexact Newton and Newton-Kantorovich methods are developed for solving a nonsmooth normal map equation equivalent to the optimality system. Numerical simulations featuring examples from the aerospace and automotive domain are presented which show that inexact Newton-Kantorovich type methods can achieve significant reduction of the computational time.**

I. Introduction

In recent years, Constrained Nonlinear Model Predictive Control (CNMPC), see e.g. [1]–[4] and references therein, has been receiving an increasing amount of attention as a systematic approach to control nonlinear systems. In many current and envisioned applications, especially to complex autonomous systems operating in dynamically changing environments, CNMPC requires repeated numerical solution of constrained nonlinear optimal control problems. One of the main challenges for the successful implementation of CNMPC is that these optimal control problems must be solved at a faster timescale with respect to the dynamics of the controlled system. To address this challenge, a number of computational approaches for CNMPC have been developed over the last few decades. In particular, the Sequential Quadratic Programming (SQP) method, which is a popular Newton type method for solving constrained optimization problems, has been pursued for CNMPC. Specifically, in [5], [6], and [7], approximate differentiation, non-smooth penalty functions, and initial value embeddings have been used for reducing the computational cost of the individual SQP operations. Along similar lines, in [8] a trust-region feasibility-perturbed SQP algorithm developed in [9] is applied to CNMPC. In [10] a continuation strategy is proposed for solving the optimality system associated with the unconstrained optimal control problem. In [11], the SQP approach is combined with a neighboring extremal strategy to develop a predictor-corrector type algorithm for fast solution updates. An overview to computational strategies for CNMPC can be found in the book [2].

*This paper is supported by the National Science Foundation (NSF) Grant 156229, the Austrian Science Foundation (FWF) Grant P26640-N25, and the Australian Research Council (ARC) Project DP160100854.

[1] American Mathematical Society and the University of Michigan, email: dontchev@umich.edu.

[2] Toyota Motor North America, Research and Development, email: mike.huang@toyota.com

[3] Department of Aerospace Engineering, University of Michigan, email: ilya@umich.edu, mnicotra@umich.edu

Despite these encouraging developments and actual implementations in the process control industry (see, e.g., [12]), further enhancements are needed to accommodate applications to systems with fast dynamics and limited on-board computing power; such systems appear in, e.g., the aerospace and automotive domain.

In this paper, we highlight the potential for further reducing the computational effort by combining a version of the Newton method proposed by L. V. Kantorovich [13] with an inexact Newton iteration proposed by Dembo et al. [14] for solving nonlinear equations. To be specific, recall that the standard Newton method for solving the nonlinear equation $f(x) = 0$, where $f : \mathbb{R}^n \to \mathbb{R}^n$ is a continuously differentiable function, has the form

$$f(x^k) + \nabla f(x^k)(x^{k+1} - x^k) = 0, \quad k = 0, 1, \ldots, \qquad (1)$$

where $\nabla f(x)$ is the Jacobian of f at x. In numerical analysis textbooks one finds the classical result, sometimes attributed to Cauchy as in [15], that if the Jacobian $\nabla f(\bar{x})$ at a solution $\bar{x}$ of the equation is invertible, and if the starting point x^0 is close enough to $\bar{x}$, then the iteration (1) generates a sequence x^k which is convergent to $\bar{x}$ with a rate that depends on the properties of f. Kantorovich proposed a different approach to analyzing the convergence of the Newton method by imposing conditions at the point x^0, where the iterations begin, rather than at an (unknown) solution $\bar{x}$. Furthermore, he showed that, to achieve convergence, there is no need to calculate the Jacobian at each iteration but use the Jacobian evaluated at the starting point. That is, the iteration (1) is replaced by

$$f(x^k) + \nabla f(x^0)(x^{k+1} - x^k) = 0, \quad k = 0, 1, \ldots, \qquad (2)$$

which Kantorovich called the *modified Newton process*. In other words, the Kantorovich modification of the Newton method consists in computing the Jacobian only once, at the beginning, and reusing it in every successive iteration. This would clearly result in saving computational time — but most importantly, as discovered by Kantorovich, the method still has theoretically the same convergence properties as the Newton method. This makes the Kantorovich approach particularly attractive for solving numerically large-scale structured optimization problems such as those appearing in optimal control and CNMPC.

Kantorovich proved the so-called *semilocal* convergence of his modification of the Newton method by using a new approach, subsequently called the *majorization* technique, see, e.g., [16]. More recently, the iteration (2) was named the *Newton-Kantorovich method*, e.g., in the books [17] and [18], which is the name we use in this paper. In [19], and perhaps elsewhere, this method is called the *chord method*, however, using this name may cause some confusion since in other references the term is used for the secant method. Throughout

the years, the Kantorovich convergence theorem has been extended and modified in a number of ways by, in particular, utilizing various extensions of the majorization technique, but not only. We will not go into discussing these works and only mention an early version of the Kantorovich theorem due to Bartle [20], an extension of which is our main Theorem 1 presented in the next section.

Another remarkable development leading to savings of time for computations is the inexact Newton method proposed by Dembo, Eisenstat, and Steihaug in [14], who observed that, to achieve convergence, it is not necessary to solve exactly the linear equation (1) representing the Newton iteration. Specifically, Dembo at el. replaced the exact iteration (1) with

$$\|f(x^k)+\nabla f(x^k)(x^{k+1}-x^k)\| \le \zeta^k \|f(x^k)\|, \quad k=0,1,\dots, \tag{3}$$

where ζ^k is a *forcing* sequence of positive numbers convergent to zero. Most importantly, they proved that any sequence $\{x^k\}$ starting close enough to a solution $\bar{x}$ and generated by (3) is superlinearly convergent to $\bar{x}$. The key discovery here is that we can obtain convergence by using less computations at each iteration. Basic information about this method is given in the book of Kelley [19], where convergence and numerical implementations are discussed.

In this paper we combine the ideas of the Newton-Kantorovich method and the inexact Newton method, and we apply them to solving variational inequalities, representing in particular first-order necessary optimality conditions for optimal control problems. Our aim is to reduce the time for computations, an issue with crucial importance in Model Predictive Control.

Various Newton type methods for solving variational inequalities and optimization problems have been extensively studied since the 70s. In the last two decades, a number of new developments have appeared and are detailed in several books [21]–[27]. The extension of Kantorovich approach to variational inequalities goes back to basic works of S. M. Robinson, see [28], [29]. Kantorovich type theorems for a general inclusion in Banach spaces, which includes variational inequalities as a special case, are presented in [30]–[34].

Preliminary results pertaining to the implementation of the exact Newton-Kantorovich scheme to NMPC have been reported in our conference paper [35], which addressed the case of penalized constraints. The present paper provides a rigorous theoretical framework for applying inexact Newton-Kantorovich-like algorithms. In addition, we present extensive computational case studies, including exact and inexact Newton and Newton-Kantorovich strategies.

This paper is organized as follows: Section II of this paper addresses the convergence of an inexact version of the Newton-Kantorovich method applied to the following variational inequality: find a point $x \in C$ such that

$$\langle \varphi(x), y-x\rangle \ge 0 \quad \text{for all } y \in C, \tag{4}$$

where $\varphi : \mathbb{R}^n \to \mathbb{R}^n$ is a continuously differentiable function with Jacobian at x denoted by $\nabla\varphi(x)$, and C is a closed and convex set in $\mathbb{R}^n$. The main convergence theorem, Theorem 1, provides a theoretical basis for developing numerical strategies for solving variational systems of this kind, and its proof is given in the Appendix. Section III introduces a control-constrained optimal control problem and an associated optimality system involving the state and costate equations as well as a variational inequality for the control. This section discusses inexact and exact versions of the SQP method with Newton-Kantorovich updates and with Newton updates. Analogous versions of a nonsmooth Newton method applied to a nonsmooth equation equivalent to the optimality system are also presented. Section IV provides results of numerical experiments for two examples of applications to spacecraft attitude control and diesel engine control.

Throughout the paper we use the following notation: $\mathbb{R}^n$ denotes the standard Euclidean space of dimension n equipped with the usual scalar product $\langle\cdot,\cdot\rangle$ and the associated norm $\|\cdot\|$. The closed ball with center x and radius a is denoted as $\mathbb{B}_a(x)$. We use the operator norm for matrices. The distance from a point x to a set C in $\mathbb{R}^n$ is $d(x,C)=\inf_{v\in C}\|v-x\|$. Given a closed and convex set $C \subseteq \mathbb{R}^n$, its normal cone mapping N_C is defined as

$$N_C(x)=\begin{cases} \{y\in\mathbb{R}^n \mid \langle y, v-x\rangle \le 0 \text{ for all } v\in C\} & \text{if } x\in C,\\ \emptyset & \text{otherwise.}\end{cases}$$

In this notation, the variational inequality (4) can be written equivalently as the inclusion

$$\varphi(x)+N_C(x) \ni 0. \tag{5}$$

II. An inexact Newton-Kantorovich method for variational inequalities

In this section we study convergence of the following inexact version of the Newton-Kantorovich method for solving the variational inequality (5):

$$\begin{aligned}(\varphi(x^k)+(A+B^k)(x^{k+1}-x^k)\\ +N_C(x^{k+1}))\cap \mathbb{B}_{\zeta^k\sigma(x^k)}(0)\neq\emptyset\end{aligned} \tag{6}$$

for $k=0,1,\dots$, where A is an $n\times n$ matrix, B^k is a sequence of $n\times n$ matrices, ζ^k is a sequence of positive numbers convergent to zero, denoted $\zeta^k\searrow 0$, and σ is a function from $\mathbb{R}^n$ to itself which is locally Lipschitz continuous.

For $A=0$, $B^k=\nabla\varphi(x^k)$, and $\sigma=0$, we obtain the standard exact Newton method for solving the variational inequality (5); for $C=\mathbb{R}^n$ it reduces to (1). For $A=\nabla\varphi(x^0)$, $B^k=0$, $C=\mathbb{R}^n$, and $\sigma=0$, we obtain the Newton-Kantorovich method for (2). Further, for $A=0$, $B^k=\nabla\varphi(x^k)$, $C=\mathbb{R}^n$, and $\sigma(x)=\|f(x)\|$ the iteration (6) reduces to (3). Another particular case of (6) is when $A=0$ and B^k represents quasi-Newton updates; this is an extensive subject which we will not discuss in this paper.

The following theorem, the proof of which is given in the Appendix, provides sufficient conditions for convergence of the iterative process (6) to a solution of (5).

Theorem 1. *Assume that the mapping*

$$\mathbb{R}^n \ni y \mapsto (A+N_C)^{-1}(y) \subset \mathbb{R}^n \tag{7}$$

is single-valued and Lipschitz continuous with a Lipschitz constant κ on $\mathbb{R}^n$ and let there exist a constant $\mu\in(0,1/(2\kappa))$ such that

$$\|B^k\|\le\mu \quad \textit{for all} \quad k\in\{0,1,\dots\}. \tag{8}$$

Then, given a sequence $\zeta^k \searrow 0$, for any starting point $x^0 \in \mathbb{R}^n$, there exists a unique sequence $\{x^k\}$ generated by the iteration (6). *Furthermore, assume that there exist constants $a > 0$, $\delta > 0$ and $l > 0$ such that, for*

$$t := \frac{\kappa}{1-\kappa\mu}\left(\delta + \mu + \zeta^0 l\right) < 1, \tag{9}$$

the following conditions hold:

(a) the starting point x^0 satisfies

$$d(0, \varphi(x^0) + N_C(x^0)) + \zeta^0\sigma(x^0) < at(1-t), \tag{10}$$

where the function σ is Lipschitz continuous on $\mathbb{B}_a(x_0)$ with Lipschitz constant l;

(b) along the sequence $\{x^k\}$ one has

$$\|\varphi(x^k) - \varphi(x^{k-1}) - A(x^k - x^{k-1})\| \leq \delta\|x^k - x^{k-1}\|. \tag{11}$$

Then $x^k \in \mathbb{B}_a(x^0)$ for all $k \in \{0, 1, \dots\}$ and the sequence $\{x^k\}$ converges r-linearly to a solution $\bar{x}$ of (5) *in $\mathbb{B}_a(x^0)$. Specifically, $\|x^k - \bar{x}\| \leq at^k$ for all $k \in \{1, 2, \dots\}$, where t is as in (9).*

Although in the case of practical large-scale structured optimization problems that appear in CNMPC it might not be always possible to verify all assumptions of Theorem 1 *a priori*, given estimates of the constants $\kappa, \mu, a, \delta, l$, one may check whether the starting point x^0 satisfies (10) and monitor conditions (8) and (11) during computations. This can be used as general guidelines to enhance the computations and provide a tool for troubleshooting in non-convergent cases.

Clearly, if $C = \mathbb{R}^n$, the assumption involving (7) reduces to the invertibility of A, in which case $\kappa = \|A^{-1}\|$. If C is a proper convex and closed subset of $\mathbb{R}^n$, this assumption is satisfied for example when there exists $\beta > 0$ such that

$$\langle Au, u\rangle \geq \beta\|u\|^2 \quad \text{for all} \quad u \in C - C, \tag{12}$$

where $C - C = \{v - w \mid v, w \in C\}$; then the Lipschitz constant is $\kappa = \beta^{-1}$. A short proof of this statement is given in the Appendix.

For $A = \nabla\varphi(\bar{x})$, where $\bar{x}$ is a solution of (5), $B^k = \nabla\varphi(x^k) - \nabla\varphi(\bar{x})$ and $\sigma = 0$, (6) becomes the exact Newton iteration

$$\varphi(x^k) + \nabla\varphi(x^k)(x^{k+1} - x^k) + N_C(x^{k+1}) \ni 0. \tag{13}$$

In this case, assumption (7) means that the mapping $A + N_C$ appearing in (5) is *globally strongly regular* at $\bar{x}$ for $-\varphi(\bar{x}) - A\bar{x}$. When we localize this condition to a neighborhood of $(-\varphi(\bar{x}) - A\bar{x}, \bar{x}) \in \mathrm{gph}(A + N_C)$, then we have the usual local strong regularity as defined by S. Robinson in [36]; for an extensive treatments of regularity properties of mappings in variational analysis and optimization, see the book [21]. Note that when C is a cone, then the local strong regularity at the origin is equivalent to the global one. In the path-breaking work of Josephy [37] and in numerous works on Newton type methods that have followed, the local strong regularity *at a solution* $\bar{x}$ has been utilized to establish superlinear convergence of the iteration (13) provided that the starting point is close to $\bar{x}$, a recent collection of results in that direction is given in [26]. In this paper we prefer not to involve an unknown solution in the assumptions. Another possibility is to assume local strong regularity at the starting point; this however leads to complications with controlling the neighborhoods during iterations which in turn may require other conditions; as a result, the local conditions become considerably more difficult to verify and less appealing for application to the large-scale structured optimization problems as those appearing in CNMPC. Still, it would be perhaps interesting to explore, at lest theoretically, possible ways to localize the strong regularity that would lead to meaningful results for computations; we put this topic aside for future research.

For $A = \nabla\varphi(x^0)$, $B^k = 0$, $C = \mathbb{R}^n$, and $\sigma = 0$, from Theorem 1 we obtain a version of the Kantorovich theorem as given in, e.g., [20]. In the Appendix we show that Theorem 1 implies not only a version of the original Kantorovich convergence theorem but, under an additional assumption, also yields the standard convergence result which requires the initial point x^0 be close to a solution $\bar{x}$. We also show how to obtain superlinear and quadratic convergence depending on the properties of the function φ. Note that, unlike the the bulk of convergence results in the literature concerning (13), the assumptions of Theorem 1 are stated only in terms of relations among constants that bound the expressions involved in the method. Another difference is that the assumptions of Theorem 1 do not involve any information regarding a solution, neither do we assume that the function φ is differentiable. As seen in Section III, the latter opens the door to extending the Newton-Kantorovich method to problems involving nonsmooth functions.

To incorporate the inexact method of Dembo et al. [14], the straightforward extension to variational inequalities would employ the residual, i.e $\sigma(x) = d(-\varphi(x), N_C(x))$. However, if we use this expression as a measure of inexactness, we may encounter difficulties with the assumption in Theorem 1 that σ is Lipschitz continuous since the normal cone mapping N_C may not even be continuous. A way to avoid this is to employ the so-called *normal map* associated with (5), in the form

$$\nu(z) := \varphi(P_C(z)) + z - P_C(z), \tag{14}$$

where P_C is the projection mapping on the set C. As shown in [29], the equations $x = P_C(x)$ and $z = x - \nu(x)$ give a one-to-one correspondence between the zeros of ν and the solutions of (5). Furthermore, since ν is locally Lipschitz continuous, the function $\sigma(x) = \|\nu(x)\|$ satisfies assumption (b) in Theorem 1.

To illustrate how to apply Theorem 1, consider the optimization problem

$$\text{Minimize } g(x) \quad \text{subject to} \quad x \in C, \tag{15}$$

for a twice continuously differentiable function g and a closed and convex set C. The Newton iteration (13) applied to the first-order necessary optimality condition $\nabla g(x) + N_C(x) \ni 0$ has the form

$$\nabla g(x^k) + \nabla^2 g(x^k)(x^{k+1} - x^k) + N_C(x^{k+1}) \ni 0 \tag{16}$$

for $k = 0, 1, \dots$. Let $\bar{x}$ be a solution of (15). Setting $A = \nabla^2 g(\bar{x})$, $B^k = \nabla^2 g(x^k) - \nabla^2 g(\bar{x})$, and $\sigma = 0$, we obtain (16) as a particular case of (6). In this case, condition (12), with $A = \nabla^2 g(\bar{x})$ is a strong form of the basic second-order optimality condition for (15), also called *coercivity*; this

condition holds for example when the function g is strongly convex If the variational inequality (16) is solved by first converting it to a quadratic programming problem and then solving that problem, we come to the familiar SQP method. Setting $\sigma(x) = \|\nabla g(P_C(z)) + z - P_C(z)\|$, we can formulate the conditions for the inexact SQP method.

III. The Newton-Kantorovich method for optimal control problems

This section specializes the results of the preceding section to the following discrete-time optimal control problem:

$$\text{Minimize} \quad J(x,u) = \sum_{i=0}^{N-1} \ell(x_i,u_i) + \Phi(x_N) \tag{17}$$

subject to the constraints

$$x_{i+1} = f(x_i,u_i), \quad i = 0,1,\ldots,N-1, \quad x_0 \quad \text{given}, \tag{18}$$

$$x_i \in \mathbb{R}^n, \quad u_i \in U_i, \quad i = 0,1,\ldots,N-1, \tag{19}$$

where $x = (x_1,\cdots,x_N) \in \mathbb{R}^{n\times N}, u = (u_0,\ldots,u_{N-1}) \in \mathbb{R}^{m\times N}$, $\ell : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$, $\Phi : \mathbb{R}^n \to \mathbb{R}$ and $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$, $x_i \in \mathbb{R}^n$ and $u_i \in \mathbb{R}^m$ are the state and control vectors, respectively, and i is the running discrete time. The sets U_i represents control constraints. Each U_i is assumed to be a convex and closed subset of $\mathbb{R}^m$. The functions ℓ, Φ and f are defined in respective Euclidean spaces and assumed to be twice continuously differentiable everywhere. Here and below, with some abuse of notation, all vectors are considered as column vectors although they may be not written as such.

To apply the Lagrange multiplier rule to (17), we need to make sure that the equality constraint given by the state equations (18) satisfies the constraint qualification condition. In our case this condition reduces to linear independence of the gradients of the constraints. Indeed, for any $\rho = (\rho_0,\ldots,\rho_{N-1}) \in \mathbb{R}^{n\times N}$, there always exists a solution $(\xi,\eta) \in \mathbb{R}^{n\times N} \times \mathbb{R}^{m\times N}$ that satisfies the difference relation

$$\xi_{i+1} - \nabla_x f(\bar{x}_i,\bar{u}_i)\xi_i - \nabla_u f(\bar{x}_i,\bar{u}_i)\eta_i = \rho_i, \quad i = 0,1,\ldots,N-1, \quad \xi_0 = 0.$$

Thus, from, e.g., [38, Theorem 6.14], the first-order necessary condition for problem (17) can be stated in terms of the Lagrangian

$$L(x,u,q) = J(x,u) + \sum_{i=0}^{N-1} \langle q_i, x_{i+1} - f(x_i,u_i)\rangle, \tag{20}$$

where $q = (q_0,\ldots,q_{N-1})$ is the vector of Lagrange multipliers, also referred to as the costates, and the state and control variables are grouped in the vectors $x = (x_0,\ldots,x_N)$ and $u = (u_0,\ldots,u_{N-1})$, respectively. Let $\bar{v} = (\bar{x},\bar{u})$ be a solution to (17). Specifically, by calculating the derivative of the Lagrangian (20) and making some rearrangements, we obtain the first-order necessary optimality condition for problem (17) as the following system which includes the state equations determined by a forward recursion, the costate equations determined by a backward recursion, and a variational inequality for the control:

$$\begin{aligned} x_{i+1} &= f(x_i,u_i),\ i = 0,1,\ldots,N-1,\ x_0 \text{ given},\\ q_{i-1} &= -\nabla_x H(x_i,u_i,q_i),\ i = 1,\ldots,N-1,\\ q_{N-1} &= -\nabla\Phi(x_N),\\ 0 &\in \nabla_u H(x_i,u_i,q_i) + N_{U_i}(u_i),\ i = 0,\ldots,N-1, \end{aligned} \tag{21}$$

where the Hamiltonian $H(x,u,q) = \ell(x,u) - \langle q, f(x,u)\rangle$. Problem (21) consists of a discrete-time two-point boundary value problem combined with a structured variational inequality. If the sets U_i are described by inequality and equality constraints, the optimality conditions can be written equivalently as a complementarity problem with associated Lagrange multipliers. In this paper, we avoid the direct usage of Lagrange multipliers which simplifies the notation, decreases the number of variables, and bypasses several issues, e.g. constraint qualification, normality, and complementary slackness. A potential drawback of this approach, however, is that we move away from techniques using, e.g. duality and multiplier methods, which could be beneficial in certain situations.

The remainder of this section is dedicated to the introduction of exact and inexact versions of standard Newton and Newton-Kantorovich type strategies for solving the optimal control problem (17).

A. Exact/inexact SQP/Kantorovich type methods

The exact version of the SQP-Kantorovich iteration is as follows: First, choose the starting point (x_i^0, q_i^0, u_i^0) and compute the matrices:

$$\begin{aligned} A_i &= \nabla_x f(x_i^0,u_i^0), & B_i &= \nabla_u f(x_i^0,u_i^0),\\ R &= \nabla^2\Phi(x_N^0) & T_i &= \nabla^2_{xx} H(x_i^0,q_i^0,u_i^0),\\ S_i &= \nabla^2_{xu} H(x_i^0,q_i^0,u_i^0), & V_i &= \nabla^2_{uu} H(x_i^0,q_i^0,u_i^0). \end{aligned} \tag{22}$$

Then, given the current iteration (x_i^k, q_i^k, u_i^k), compute the increments $(\xi_i^k, \eta_i^k, \chi_i^k)$ by solving the system of equations

$$\begin{cases} \xi_{i+1}^k - A_i\xi_i^k - B_i\chi_i^k + x_{i+1}^k - f(x_i^k,u_i^k) = 0,\\ \eta_{i-1}^k + A_i^T\eta_i^k + T_i\xi_i^k + S_i\chi_i^k + q_{i-1}^k + \nabla_x H(x_i^k,u_i^k,q_i^k) = 0,\\ \eta_{N-1}^k + R\xi_N^k + q_{N-1}^k + \nabla\Phi(x_N^k) = 0,\\ S_i^T\xi_i^k - B_i^T\eta_i^k + V_i\chi_i^k\\ \qquad + \nabla_u H(x_i^k,u_i^k,q_i^k) + N_{U_i}(u_i^k + \chi_i^k) \ni 0, \end{cases} \tag{23}$$

for $i = 1,\ldots,N-1$ and with $\xi_0^k = 0$. The next iteration is obtained as

$$x_i^{k+1} = x_i^k + \xi_i^k, \quad q_i^{k+1} = q_i^k + \eta_i^k \quad \text{and} \quad u_i^{k+1} = u_i^k + \chi_i^k.$$

Note that, for the standard SQP iteration, at each iteration the matrices in equation (23) would instead be computed as follows:

$$\begin{aligned} A_i^k &= \nabla_x f(x_i^k,u_i^k), & B_i^k &= \nabla_u f(x_i^k,u_i^k),\\ R^k &= \nabla^2\Phi(x_N^k) & T_i^k &= \nabla^2_{xx} H(x_i^k,q_i^k,u_i^k),\\ S_i^k &= \nabla^2_{xu} H(x_i^k,q_i^k,u_i^k), & V_i^k &= \nabla^2_{uu} H(x_i^k,q_i^k,u_i^k). \end{aligned} \tag{24}$$

The inexact SQP-Kantorovich iteration has the following form: given the current iterate (x_i^k, q_i^k, u_i^k), the kth increment is

obtained by solving the system

$$\begin{cases} \|\xi_{i+1}^k - A_i\xi_i^k - B_i\chi_i^k + x_{i+1}^k - f(x_i^k,u_i^k)\| \\ \qquad \le \zeta^k\sigma_x^k, \\ \|\eta_{i-1}^k + A_i^T\eta_i^k + T_i\xi_i^k + S_i\chi_i^k + q_{i-1}^k \\ \qquad +\nabla_x H(x_i^k,u_i^k,q_i^k)\| \le \zeta^k\sigma_q^k, \\ \|\eta_{N-1}^k + R\xi_N^k + q_{N-1}^k + \nabla\Phi(x_N^k)\| \le \zeta^k\sigma_N^k, \\ d(0, S_i^T\xi_i^k - B_i^T\eta_i^k + V_i\chi_i^k + \nabla_u H(x_i^k,u_i^k,q_i^k) \\ \qquad +N_{U_i}(u_i^k+\chi_i^k)) \le \zeta^k\sigma_u^k. \end{cases} \tag{25}$$

with $\xi_0^k = 0$, where

$$\sigma_x^k = \sum_{i=0}^{N-1} \|x_{i+1}^k - f(x_i^k,u_i^k)\|, \tag{26}$$

$$\sigma_q^k = \sum_{i=0}^{N-1} \|q_{i-1}^k + \nabla_x H(x_i^k,u_i^k,q_i^k)\|, \tag{27}$$

$$\sigma_N^k = \|q_{N-1}^k + \nabla\Phi(x_N^k)\|, \tag{28}$$

$$\sigma_u^k = \|\nabla_u H(x_i^k,P_U(u_i^k),q_i^k) + u_i^k - P_U(u_i^k)\|. \tag{29}$$

This provides a less stringent termination criteria for the underlying QP subsolver with respect to (23). Once again, the difference between the inexact SQP and the inexact SQP-Kantorovich is that the former computes the matrices in the quadratic problem at each iteration as in (24), whereas the latter computes them only at the starting point (x^0,q^0,u^0) as in (22).

B. Nonsmooth Newton type methods

Another approach to solving the optimality system (21) is to replace it by a (nonsmooth) equation and then apply a nonsmooth Newton method. This can be done in various ways which are described in detail in the books [22], [23], [24]. The main convergence result for this method states that if the function involved is *semismooth* and each element of the generalized Jacobian at the reference solution is nonsingular, then one has superlinear convergence provided that the iteration starts close enough to the solution. If the function is *strongly semismooth* then we have quadratic convergence. In this paper, the original variational inequality is reduced to a semismooth equationthe approach developed in [29], which is based on introducing the normal map (14) associated with (5). Without going into further details, we only mention that in our examples the constraint set C is polyhedral, therefore the projection mapping P_C is a piecewise linear, hence a semismooth (even strongly semismooth) function. Thus, the nonsmooth Newton iteration has the form,

$$\nu(z^k) + A\,\Delta z^k = 0,$$

where A is a square matrix which is an element of the Clarke generalized Jacobian $\partial\nu(z^k)$. Taking advantage of the composition rule $\partial\varphi(P_C(z)) = \nabla\varphi(P_C(z))\partial P_C(z)$, we obtain

$$\partial\nu(z^k) = \big(\nabla\varphi(P_C(z^k)) - I\big)\partial P_C(z^k) + I.$$

Abusing notation, the normal mapping for the optimality system (21) has the form

$$\nu(x,u,q) = \begin{pmatrix} x_0 \\ x_{i+1} - f(x_i, P_U(u_i)) \\ q_{i-1} + \nabla_x H(x_i, P_U(u_i), q_i) \\ q_{N-1} + \nabla\Phi(x_N) \\ \nabla_u H(x_i, P_U(u_i), q_i) + u_i - P_U(u_i) \end{pmatrix}$$

and the set $C = \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n \times U$. Therefore, given the current estimates (x_i^k, q_i^k, u_i^k), the nonsmooth Newton-Kantorovich method solves (21) by computing the increments $(\xi_i^k, \eta_i^k, \chi_i^k)$ obtained as solutions the following system of equations:

$$\begin{cases} \xi_{i+1}^k - A_i^P\xi_i^k - B_i^P\partial P_U(u_i^k)\chi_i^k + x_{i+1}^k - f(x_i^k, P_U(u_i^k)) = 0, \\ \eta_{i-1}^k + A_i^{P,T}\eta_i^k + T_i^P\xi_i^k + S_i^P\partial P_U(u_i^k)\chi_i^k + q_{i-1}^k \\ \qquad +\nabla_x H(x_i^k, P_U(u_i^k), q_i^k) = 0, \\ \eta_{N-1}^k + R\xi_N^k + q_{N-1}^k + \nabla\Phi(x_N^k) = 0, \\ S_i^{P,T}\xi_i^k - B_i^{P,T}\eta_i^k + \big((V_i^P - I)\partial P_U(u_i^k) + I\big)\chi_i^k \\ \qquad +\nabla_u H(x_i^k, P_U(u_i^k), q_i^k) + u_i^k - P_U(u_i^k) = 0, \end{cases} \tag{30}$$

with $\xi_0^k = 0$ and where the matrices $A_i^P, B_i^P, T_i^P, S_i^P, V_i^P$ are defined as in (22), with the only exception that u_i^0 should be replaced by $P_U(u_i^0)$. If U is a box, say $U = \{u \in \mathbb{R}^m \mid a_i \le u_i \le b_i\}$, the Clarke generalized Jacobian of P_U is easy to compute. Namely, $\partial P_U(u)$ is a diagonal matrix with

$$[\partial P_U(u)]_{ii} := \begin{cases} 1 & \text{if } \ a_i < u_i < b_i, \\ \text{any number in } [0,1] & \text{if } a_i = u_i \text{ or } u_i = b_i, \\ 0 & \text{otherwise.} \end{cases}$$

In the numerical simulations provided in this paper, we assign $\partial P_U(u_i) = 0.1$ whenever u_i belongs to the boundary of U.

As in the preceding cases, the difference between the nonsmooth Newton method and the nonsmooth Newton-Kantorovich method is that in the former the matrices A_i^P, B_i^P, S_i^P, T_i^P, V_i^P are recomputed at each iteration in analogy to (24), whereas in the latter they are only computed at the starting iteration. It is important to note that the result in Theorem 1 is still applicable if we take as the matrix A any nonsingular element of Clarke's generalized Jacobian. To satisfy condition (11), one may assume that the latter has a small radius and then apply [21, Proposition 6F.3].

For the inexact nonsmooth Kantorovich iteration: given the current iterate (x_i^k, q_i^k, u_i^k), the kth increment is obtained by solving the system

$$\begin{cases} \|\xi_{i+1}^k - A_i^P\xi_i^k - B_i^P\partial P_U(u_i^k)\chi_i^k + x_{i+1}^k \\ \qquad -f(x_i^k, P_U(u_i^k))\| \le \zeta^k\sigma_x^k, \\ \|\eta_{i-1}^k + A_i^{P,T}\eta_i^k + T_i^P\xi_i^k + S_i^P\partial P_U(u_i^k)\chi_i^k + q_{i-1}^k \\ \qquad +\nabla_x H(x_i^k, P_U(u_i^k), q_i^k)\| \le \zeta^k\sigma_q^k, \\ \|\eta_{N-1}^k + R\xi_N^k + q_{N-1}^k + \nabla\Phi(x_N^k)\| \le \zeta^k\sigma_N^k, \\ \|S_i^{P,T}\xi_i^k - B_i^{P,T}\eta_i^k + \big((V_i^P - I)\partial P_U(u_i^k) + I\big)\chi_i^k \\ \qquad +\nabla_u H(x_i^k, P_U(u_i^k), q_i^k) + u_i^k - P_U(u_i^k)\| \le \zeta^k\sigma_u^k, \end{cases} \tag{31}$$

with $\xi_0^k = 0$ given and with σ_x^k, σ_q^k, σ_N^k, and σ_u^k defined as in (26)-(29). This equation can be solved by using the same linear solver that would be used for equation (30), but utilizing the inequalities (31) to modifying the termination criteria rather than imposing the standard fixed tolerance value.

IV. Numerical Case Studies

In this section we present results of numerical simulations with the goal of comparing the convergence rate of the standard SQP and nonsmooth Newton methods with their inexact and Kantorovich counterparts. This will be done by monitoring the computational time and number of iterations of the different solvers depending on the prediction horizon length N. The numerical simulations are performed on two practically relevant nonlinear systems from the aerospace and the automotive domain. The first example addresses the attitude control of a spacecraft, whereas the second addresses the control of a turbocharged diesel engine. In both case studies, the model predictive controller is implemented by formulating an appropriate optimal control problem which is iteratively solved at each time step. For the sake of comparison, the termination criteria is always $e^k \leq \varepsilon$, where $\varepsilon > 0$ is the allowed tolerance value and e^k is the residual of the optimality system (21), computed as

$$\begin{aligned} e^k = &\sum_{i=0}^{N-1} \|x_{i+1}^k - f(x_i^k, u_i^k)\| \\ &+ \sum_{i=0}^{N-1} \|q_{i-1}^k + \nabla_x H(x_i^k, u_i^k, q_i^k)\| + \|q_{N-1}^k + \nabla\Phi(x_N^k)\| \\ &+ \sum_{i=0}^{N-1} \|u_i^k - P_U(u_i^k - \nabla_u H(x_i^k, u_i^k, q_i))\|. \end{aligned}$$

Given a solution $(x^\star, q^\star, u^\star)$, the control input issued to the system is $u_0^\star$. The solution is then used to initialize the Newton-like iterations of the following time step. In both case studies, the following methods are used to solve the optimal control problem:

- SQP: Each iterate (23) is formulated as a Quadratic Programming (QP) subproblem and is solved to a fixed tolerance $\varepsilon_2 \ll \varepsilon$ using the smoothed Fischer-Burmeister method detailed in [39]. The second order terms in (23) are updated at each iteration according to (24).
- Inexact SQP: The QP subproblems are generated and solved as in the "SQP" method, but the tolerance of the subsolver is a function of the current residue. This is done in accordance to (25).
- SQP-Kantorovich: Analogous to the "SQP" method, but the second order terms are computed only at the first iteration using (22). The QP subproblems are solved to a fixed tolerance $\varepsilon_2 \ll \varepsilon$.
- Inexact SQP-Kantorovich: This method combines the "Inexact SQP" and the "SQP-Kantorovich" formulations by both computing the second order terms only at the first iteration and varying the tolerance of the QP subsolver on the basis of the current residue.
- NSN: The Nonsmooth Newton iteration in equation (30) is solved by implementing the GMRES method (see e.g. [40]) to a fixed tolerance $\varepsilon_2 \ll \varepsilon$. The second order terms are updated at each iteration according to (24);
- Inexact NSN: The iterations are generated and solved for the NSN method, but the tolerance of the linear solver is a function of the current residue. This is done in accordance to (31);
- NSN-Kantorovich: Analogous to the NSN method, but the second order terms are computed only at the first iteration using (22). The linear equality is solved to a fixed tolerance $\varepsilon_2 \ll \varepsilon$.
- Inexact NSN-Kantorovich: This method combines the Inexact NSN and the NSN-Kantorovich methods by both computing the second order terms only at the first iteration and varying the tolerance of the linear solver on the basis of the current residue.

Note that, in many industrial applications, models may not lend themselves easily to equation-based representation and automated differentiation (e.g. they may contain look-up tables, etc.). To take this into account, the numerical experiments compute the first and second order derivatives numerically rather than analytically. For a given function $f(a)$, the partial derivatives are computed using the central difference method

$$\frac{\partial f}{\partial a_j} \overset{\Delta}{\approx} g_i(a) = \frac{f(a + he_i) - f(a)}{h},$$

$$\frac{\partial^2 f}{\partial a_i \partial a_j} \overset{\Delta}{\approx} \frac{g_i(a + he_j) - g_i(a)}{2h} + \frac{g_j(a + he_i) - g_j(a)}{2h},$$

where $h = 10^{-6}$ is the step size of the approximation and e_i is a vector of the same length as a such that $[e_i]_i = 1$ and $[e_i]_j = 0,\ \forall j \neq i$. The proposed solvers are compared in terms of both maximum and mean computation time obtained in Matlab/Simulink running on an ASUS Notebook N550J (2.4GHz and 8GB RAM). Note that, although the reported computation times are suitable for comparison of different approaches with each other, they are likely not to be indicative (in an absolute sense) of what would be observed for practical implementations based on a faster optimized C code albeit running in slower target micro-controllers. The maximum and mean number of iterations performed by each solver is also reported.

A. Spacecraft Attitude Control

As a first example, consider the spacecraft model described in [41]. The vehicle attitude is parameterized using Euler angles, where $[\theta]_1$ is the roll angle, $[\theta]_2$ is the pitch angle and $[\theta]_3$ is the yaw angle. Given the vector of angular velocities $\omega \in \mathbb{R}^3$ and the vector of control torques $u \in \mathbb{R}^3$, the attitude dynamics are given by the following state-space equations

$$\begin{cases} \dot{\theta} = E(\theta)\omega \\ \dot{\omega} = \hat{\omega} J \omega + u, \end{cases} \tag{32}$$

where[1]

$$E(\theta) = \frac{1}{\cos([\theta]_2)} \begin{bmatrix} c([\theta]_2) & s([\theta]_1)s([\theta]_2) & c([\theta]_1)s([\theta]_2) \\ 0 & c([\theta]_1)c([\theta]_2) & -s([\theta]_1)c([\theta]_2) \\ 0 & s([\theta]_1) & c([\theta]_1) \end{bmatrix}, \tag{33}$$

$$\hat{\omega} = \begin{bmatrix} 0 & [\omega]_3 & -[\omega]_2 \\ -[\omega]_3 & 0 & [\omega]_1 \\ [\omega]_2 & -[\omega]_1 & 0 \end{bmatrix}, \tag{34}$$

and $J = \mathrm{diag}([120\ 100\ 80])\ kg \cdot m^2$ is the inertia matrix. Given a desired attitude reference $r \in \mathbb{R}^3$, the system is

[1]For the sake of shortening the notation, $c(x)$, $s(x)$ denote $\cos(x)$, $\sin(x)$.

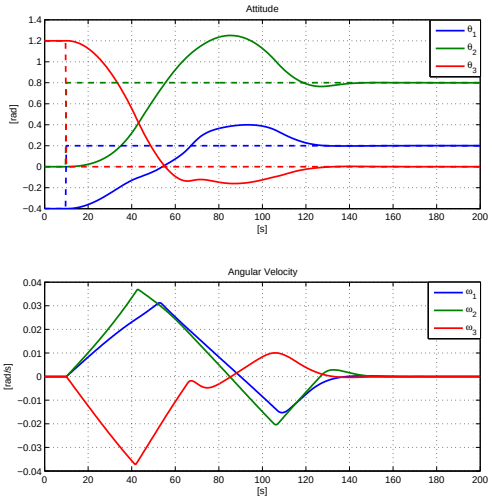

Fig. 6.1. Closed-loop response of the satellite (solid) given a change in the desired reference (dashed).

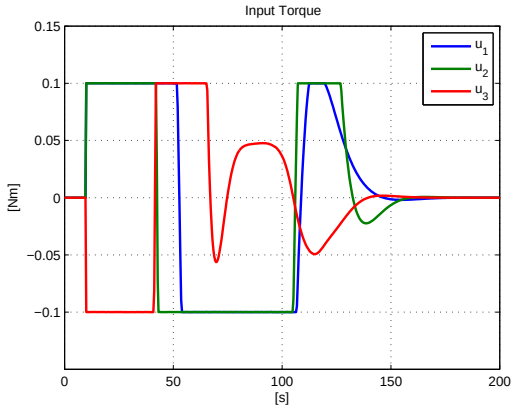

Fig. 6.2. Control input generated by the Model Predictive Control law, given a prediction horizon $N = 20$.

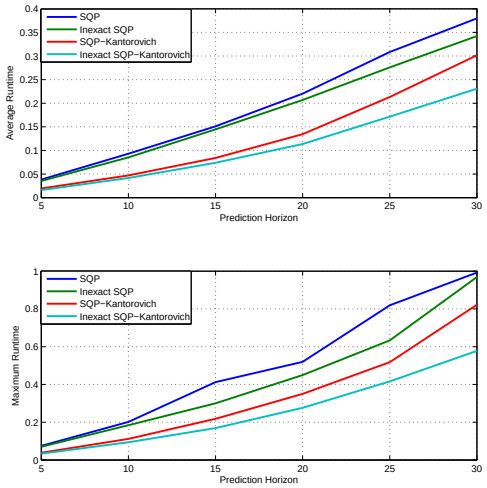

Fig. 6.3. Comparison of the average and maximum runtime required by each SQP solver.

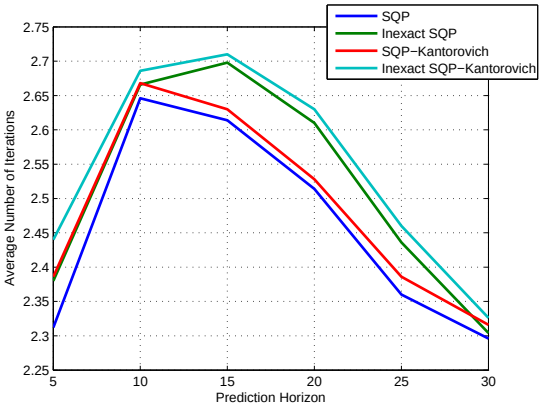

Fig. 6.4. Comparison of the average number of iterations performed by each SQP solver.

stabilized using a model predictive control law with a finite prediction horizon N, which will be varied. This is done by solving at each time step the following optimal control problem

$$\min \begin{bmatrix} \theta_N - r \\ \omega_N \end{bmatrix}^T P \begin{bmatrix} \theta_N - r \\ \omega_N \end{bmatrix} + \sum_{i=0}^{N-1} \begin{bmatrix} \theta_i - r \\ \omega_i \end{bmatrix}^T Q \begin{bmatrix} \theta_i - r \\ \omega_i \end{bmatrix} + u_i^T R u_i, \tag{35}$$

subject to the dynamics

$$\begin{bmatrix} \theta_{i+1} \\ \omega_{i+1} \end{bmatrix} = \begin{bmatrix} \theta_i \\ \omega_i \end{bmatrix} + T_s \begin{bmatrix} E(\theta_i)\omega_i \\ \hat{\omega}_i J \omega_i + u_i, \end{bmatrix} \tag{36}$$

where $T_s = 1s$ is the sampling time, and subject to the input constraints

$$u_{min} \leq [u]_h \leq u_{max}, \quad h = 1, 2, 3, \tag{37}$$

with $u_{min} = -0.1\ Nm$ and $u_{max} = 0.1\ Nm$. The proposed cost matrices for equation (35) are $Q = \text{diag}([1\ 1\ 1\ 10\ 10\ 10])$ and $R = \text{diag}([0.1\ 0.1\ 0.1])$. To ensure local stability of the MPC scheme, the terminal cost matrix P is computed by linearizing equation (36) and solving the discrete-time algebraic Riccati equation

$$P = Q + A^T P A - A^T P B (R + B^T P B)^{-1} B^T P A. \tag{38}$$

Figures 6.1-6.2 illustrate the behavior of the closed-loop system subject to the proposed Model Predictive Control law for a prediction horizon of $N = 20$. We note that, regardless of the method used to solve the optimal control problem, the closed-loop response remains practically unchanged. Given a step change in the system reference, the control inputs are initially operated at their full saturation values in a way that is somewhat similar to a bang-bang control for a double integrator. At the end of the transient, the spacecraft is successfully stabilized to the desired attitude.

The simulations are performed using $\varepsilon = 10^{-4}$ for the termination criteria, $\varepsilon_2 = 10^{-10}$ for the fixed-tolerance subsolvers, and $\zeta^k = 10^{-3} \cdot 2^{-k}$ for computing the tolerance of the inexact subsolvers. Figure 6.3 compares both the average and the maximum runtime achieved by the SQP type solvers depending

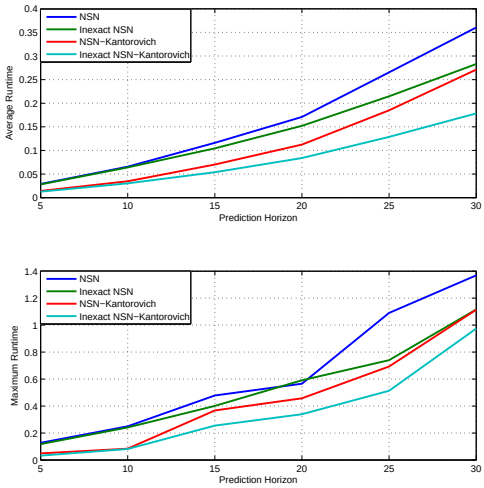

Fig. 6.5. Comparison of the average and maximum runtime required by each semismooth Newton solver.

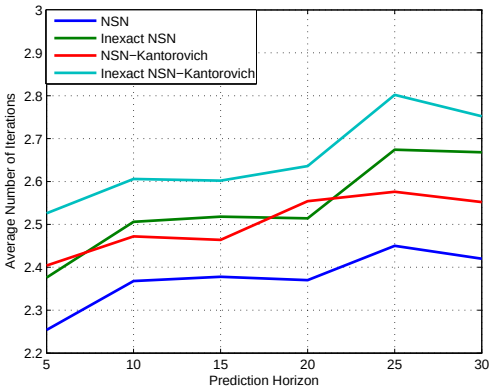

Fig. 6.6. Comparison of the average number of iterations performed by each semismooth Newton solver.

on the length of the prediction horizon N, whereas Figure 6.4 reports the average number of iterations. All SQP-solvers had a maximum number of iterations equal to 5. It is interesting to note that even though the inexact and the Kantorovich versions use more iterations than the basic SQP method, their overall runtime is shorter. This is likely due to the fact that the computational time needed at each iteration is shorter because the Jacobian does not need to be recomputed. Please note that if the Jacobian were computationally inexpensive to compute, the Kantorovich method would likely be slower than the classic implementation due to the higher number of iterations. Figures 6.5-6.6 report the same data for the nonsmooth Newton solvers. The behavior is similar in that the inexact Kantorovich method; that is, we have shorter computational time than the standard implementation. In these simulations, the maximum number of iterations performed by all the nonsmooth Newton methods is 13.

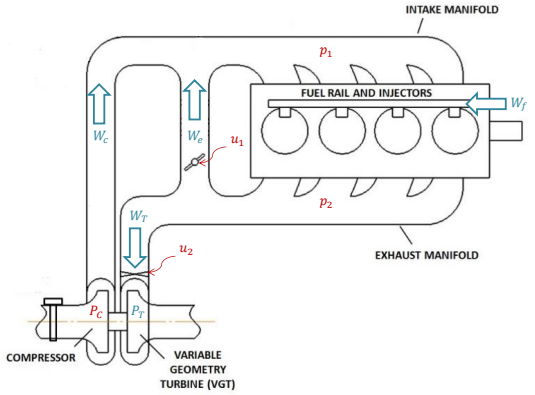

Fig. 6.7. Simplified diesel engine model.

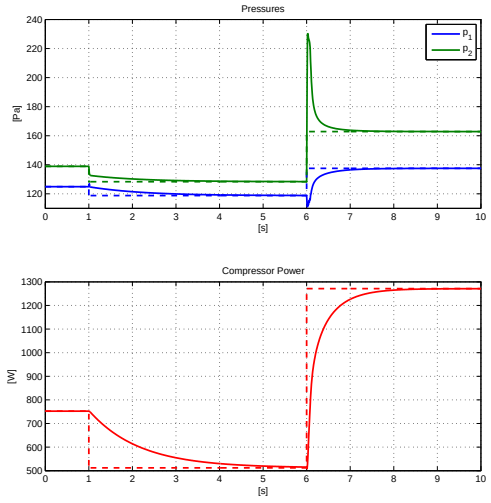

Fig. 6.8. Closed-loop response of the diesel engine states for a prediction horizon of $N = 20$. The dashed lines represent the desired equilibrium $\bar{x}$, whereas the solid lines provide the actual state trajectories.

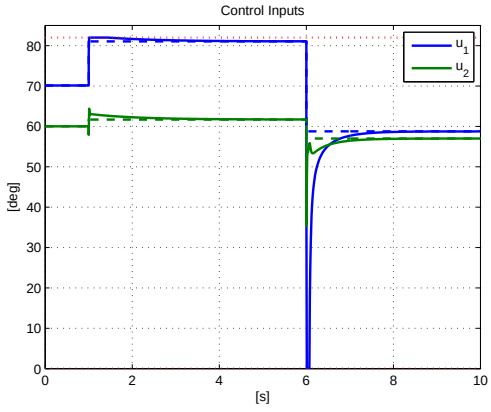

Fig. 6.9. Closed-loop response of the diesel engine inputs for a prediction horizon of $N = 20$. The dashed lines represent the target input at steady-state $\bar{u}$, whereas the solid lines provide the actual inputs issued by the MPC.

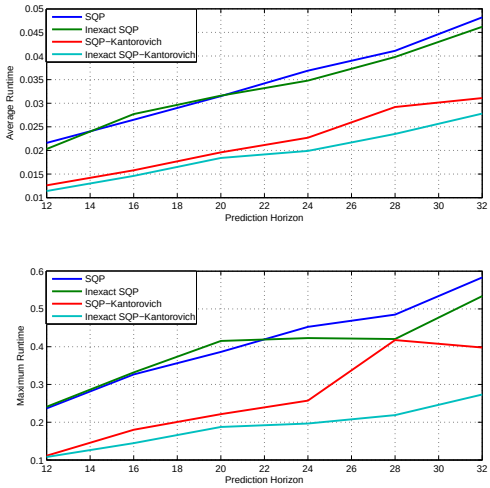

Fig. 6.10. Comparison of the average and maximum runtime required by each SQP solver.

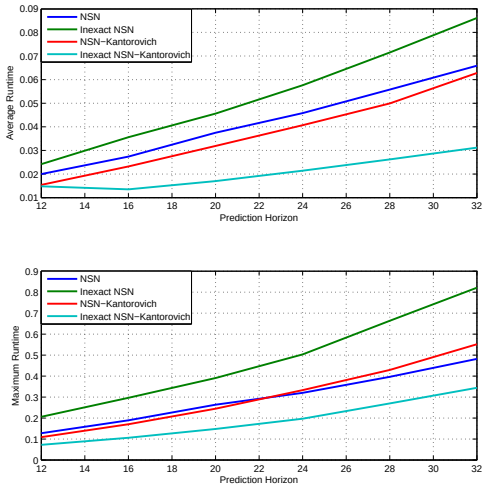

Fig. 6.12. Comparison of the average and maximum runtime required by each semismooth Newton solver.

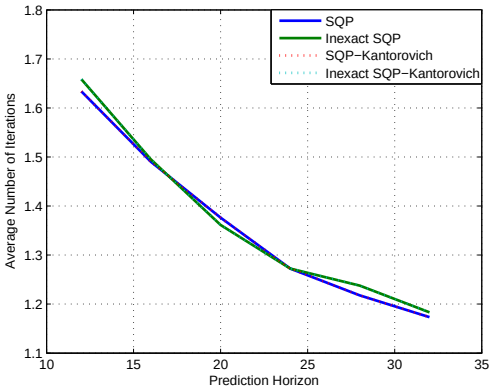

Fig. 6.11. Comparison of the average number of iterations performed by each SQP solver.

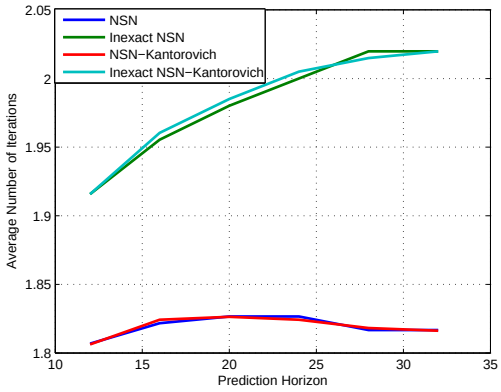

Fig. 6.13. Comparison of the average number of iterations performed by each semismooth Newton solver.

B. Diesel Engine

As a second case study, we consider a (simplified) turbo-diesel engine model detailed in [42]. Referring to Figure 6.7, the state vector $x = [p_1 \ p_2 \ P_c]^{\mathrm{T}}$ describes the inlet pressure p_1, the exhaust pressure p_2 and the compressor power P_c. The control input $[u]_1$ corresponds to the opening angle of the exhaust gas recirculation valve, whereas $[u]_2$ reflects the guide vane position of the variable geometry turbine. The state space model is given by

$$\begin{aligned} \dot{p}_1 &= k_1(W_c + W_e - k_e p_1), \\ \dot{p}_2 &= k_2(k_e p_1 + W_f - W_e - W_t), \\ \dot{P}_c &= (\eta_m P_t - P_c)/\tau, \end{aligned} \tag{39}$$

where W_f is the mass flow rate of fuel, $P_t = \eta_t c_p T_2 (1 - (p_a/p_2)^{\mu}) W_t$ is the turbine power,

$$\begin{aligned} W_e &= A_e \tfrac{p_2}{\sqrt{R_a T_2}} (1 - \cos([u]_1)) \, \Psi\left(\tfrac{p_1}{p_2}\right), \\ W_t &= A_e \tfrac{p_2}{\sqrt{R_a T_2}} (1 - \cos([u]_2)) \, \Psi\left(\tfrac{p_a}{p_2}\right), \\ W_c &= \tfrac{\eta_c}{T_a c_p} \tfrac{1}{(p_1/p_a)^{\mu} - 1} P_c, \end{aligned} \tag{40}$$

are the exhaust gas recirculation mass flow rate, the turbine mass flow rate, and the compressor mass flow rate, respectively. The function Ψ is a linear approximation of the orifice flow as a function of the pressure ratio,

$$\Psi(r) = \begin{cases} \psi(a), & \text{if } r \geq a, \\ -\psi(a), & \text{if } r \leq a^{-1}, \\ -\psi(a) + 2\psi(a)\frac{r - a^{-1}}{a - a^{-1}}, & \text{if } r \in (a^{-1}, a), \end{cases}$$

where

$$a = \left(\frac{2}{\gamma + 1}\right)^{1/\mu},$$

and

$$\psi(a) = \sqrt{\frac{2\gamma}{\gamma-1}\left(a^{\frac{2}{\gamma}} - a^{\frac{\gamma+1}{\gamma}}\right)}.$$

The remaining parameters are positive physical constants as detailed in [42].

Given a fuel mass flow rate W_f imposed by the driver torque request, the desired air mass flow rate $\bar{W}_c$ and desired EGR mass flow rate $\bar{W}_e$ that maximize engine efficiency subject to NO_x emissions constraints are tabulated by the engine manufacturer. The steady-state equilibrium that corresponds to the desired mass flows can then be obtained from equations (39)-(40), and represents the reference point for the control law

$$\begin{aligned}
\bar{p}_1 &= \tfrac{\bar{W}_c+\bar{W}_e}{k_e},\\
\bar{p}_2 &= p_a\left(1+\tfrac{\bar{W}_c}{\bar{W}_c+W_f}\tfrac{T_a}{T_2}\tfrac{1-(\bar{p}_1/p_a)^{\mu}}{\eta_c\eta_t\eta_m}\right)^{-1/\mu},\\
\bar{P}_c &= \eta_m\eta_t c_p T_2\left(1-\left(\tfrac{p_a}{\bar{p}_2}\right)^{\mu}\right)(\bar{W}_c+W_f).
\end{aligned}$$

The steady-state control input then follows from the first two equations of (40),

$$\begin{aligned}
[\bar{u}]_1 &= \arccos\left(1-\tfrac{\bar{W}_e}{A_e}\tfrac{\sqrt{R_aT_2}}{\bar{p}_2}\Psi\left(\tfrac{\bar{p}_1}{\bar{p}_2}\right)^{-1}\right),\\
[\bar{u}]_2 &= \arccos\left(1-\tfrac{\bar{W}_c+W_f}{A_e}\tfrac{\sqrt{R_aT_2}}{\bar{p}_2}\Psi\left(\tfrac{p_a}{\bar{p}_2}\right)^{-1}\right).
\end{aligned}$$

Given the desired operating point corresponding to the state vector, $\bar{x} = [\bar{p}_1 \;\; \bar{p}_2 \;\; \bar{P}_c]^{\mathrm{T}}$, and $\bar{u}$, the Model Predictive Controller is tasked with solving the following optimal control problem: minimize the cost function

$$\begin{aligned}
&(x_N-\bar{x})^T P(x_N-\bar{x})\\
&\quad+\textstyle\sum_{i=0}^{N-1}(x_i-\bar{x})^T Q(x_i-\bar{x})+(u_i-\bar{u})^T R(u_i-\bar{u}),
\end{aligned} \tag{41}$$

subject to the dynamics

$$x_{i+1} = x_i + T_s\begin{bmatrix} k_1(W_c(x_i)+W_e(x_i,u_i)-k_e p_{1,i})\\ k_2(k_e p_{1,i}+W_{f,i}-W_e(x_i,u_{1,i})-W_t(x_i,u_i))\\ (\eta_m P_t(x_i)-P_{c,i})/\tau, \end{bmatrix}, \tag{42}$$

where $T_s = 0.01\ s$ is the sampling time, and the control constraints

$$u_{min} \le [u]_h \le u_{max}, \quad h = 1, 2, \tag{43}$$

with $u_{min} = 0^{\circ}$ and $u_{max} = 82^{\circ}$. The weight matrices for equation (41) are $R = \mathrm{diag}([4\ 4])$ and $Q = \mathrm{diag}([5.4\cdot 10^{-5} \;\; 4.3\cdot 10^{-5} \;\; 5.9\cdot 10^{-7}])$. To ensure closed-loop stability, the terminal cost matrix P is computed by linearizing equation (42) and solving the discrete-time algebraic Riccati equation (38). Figures 6.8-6.9 illustrate the closed-loop behavior of the controlled system given a sudden decrease and successive increase in W_f. Given a decrease in W_f at time $t = 1\ s$, both EGR valve and VGT are opened to reduce the intake and exhaust pressures. During the initial transient, the control input $[u]_1$ is operated at the upper saturation limit $[u]_1 = 82^{\circ}$. Given an increase in W_f at time $t = 6\ s$, both the EGR valve and VGT are closed to increase the intake and exhaust pressures. Interestingly enough, the control input $[u]_1$ is briefly operated at the lower saturation limit $[u]_1 = 0^{\circ}$, which causes a spike in the exhaust pressure p_2. This speeds up the engine dynamics since both the recirculation flow W_e and the air intake flow W_c increase substantially for higher values of p_2.

The simulations are performed using $\varepsilon = 10^{-2}$ for the termination criteria, $\varepsilon_2 = 10^{-8}$ for the fixed-tolerance subsolvers, and $\zeta^k = 10^{-3}\cdot 2^{-k}$ for computing the tolerance of the inexact subsolvers. Figure 6.10 compares both the average and the maximum runtime achieved by the SQP type solvers depending on the length of the prediction horizon N, whereas Figure 6.11 reports the average number of iterations. All SQP-solvers had a maximum number of iterations that decreased from 23 to 13 as the length of the prediction horizon increased. Once again, we note that the inexact and the Kantorovich use shorter overall runtime than the basic SQP method although they need more iterations. Note that, when the prediction horizon increases, the number of iterations decreases yet the time per iteration increases. The reduction in the number of iteration is fairly interesting and could be due to the fact that, when the horizon length increases, the quality of the initial guess is also improved. The increase in the time per iteration is expected since the size of the subproblems scale with the length of the horizon. A similar behavior can be seen in Figures 6.12-6.13 which pertain to the nonsmooth Newton method and its variants.

V. Concluding Remarks

In this paper we develop a theoretical framework to support the implementation of inexact Newton-Kantorovich type methods for solving optimal control problems with control constraints. In our main Theorem 1 we present sufficient conditions for linear convergence of such methods. This theorem can be easily extended to infinite dimensional (Banach) spaces when the normal cone mapping N_C is replaced by a set-valued mapping with closed graph.

While our primary application is for constrained nonlinear model predictive control, our results are also relevant to solving offline or online various kinds of trajectory optimization problems. Overall, the computational experiments indicate that applying inexact Newton-Kantorovich methods has a potential for significant computational time savings without degrading the accuracy. In our future work we will focus on problems with both state and control constraints, and on the interplay between continuous-time models and their discretized counterpart.

References

[1] G. Goodwin, M. Seron, and J. de Dona, *Constrained Control and Estimation: An Optimization Approach*, Springer, 2006.

[2] L. Grüne, and J. Pannek, *Nonlinear Model Predictive Control*, Second Edition, Springer, 2017.

[3] D. Q. Mayne, J. B. Rawlings, C. V. Rao, and P. O. Scokaert, "Constrained model predictive control: Stability and optimality", *Automatica*, vol. 36(6), pp. 789–814, 2000.

[4] D. Q. Mayne, "Model predictive control: Recent developments and future promise", *Automatica*, vol. 50(2), pp. 2967–2986, 2014.

[5] M. Diehl, H.G. Bock, J.P. Schlöder, R. Findeisen, Z. Nagy, and F. Allgower, "Real-time optimization and nonlinear model predictive control of processes governed by differential-algebraic equations", *J. of Process Control*, vol. 12, pp. 577–585, 2002.

[6] M. Diehl, H.G. Bock, and J.P. Schlöder, "A real-time iteration scheme for nonlinear optimization in optimal feedback control", *SIAM Journal on Control and Optimization*, vol. 43(5), pp. 1714–1736, 2005.

[7] M. Diehl, R. Findeisen, H.G. Bock, J.P. Schlöder, and F. Allgower, "Nominal stability of the real-time iteration scheme for nonlinear model predictive control", *IEE Control Theory and Applications,* vol. 152(3), pp. 296–308, 2005.

[8] M. J. Tenny, S. J. Wright, and J. B. Rawlings, "Nonlinear model predictive control via feasibility-perturbed sequential quadratic programming", *Computational Optimization and Applications*, vol. 28, pp. 87–121, 2004.

[9] M. J. Tenny, and S. J. Wright, "A feasible trust-region sequential quadratic programming algorithm", *SIAM J. on Optimization*, vol. 14, pp. 1074–1105, 2004.

[10] T. Ohtsuka, "A continuation/GMRES method for fast computation of nonlinear receding horizon control", *IEEE Trans. on Automatic Control*, vol. 44(3), pp. 648–654, 1999.

[11] R. Ghaemi, J. Sun, and I. V. Kolmanovsky, "An integrated perturbation analysis and sequential quadratic programming approach for model predictive control", *Automatica*, vol. 45(10), pp. 2412–2418, 2009.

[12] ABB, "BoilerMax: Optimize power plant start-up", *ABB AG*, 2013.

[13] L. V. Kantorovich, "On Newton's method for functional equations" (Russian), *Doklady Akad. Nauk SSSR* (N.S.), vol. 59, pp. 1237–1240, 1948.

[14] R. S. Dembo, S. C. Eisenstat, T. Steihaug, "Inexact Newton methods", *SIAM J. Numer. Anal.*, vol. 19, pp. 400–408, 1982.

[15] J. M. Ortega, W. C. Rheinboldt, *Iterative Solution of Nonlinear Equations in Several Variables*, Academic Press, 1970.

[16] F. A. Potra, and V. Pták, "Nondiscrete induction and iterative processes", *Research Notes in Mathematics*, 1984.

[17] I. Argyros, *Convergence and Applications of Newton-type Iterations*, Springer, 2008.

[18] J. A. Ezquerro Fernández, M. A. Hernández Verón, *Newton's Method: an Updated Approach of Kantorovich's Theory,* Birkhäuser/Springer, 2017.

[19] C. T. Kelley, "Solving nonlinear equations with Newton's method", *Fundamentals of Algorithms*, SIAM, 2003.

[20] R. G. Bartle, "Newton's method in Banach spaces", *Proc. Amer. Math. Soc.* vol. 6, pp. 827–831, 1955.

[21] A. L. Dontchev, R. T. Rockafellar, *Implicit Functions and Solution Mappings*, Second Edition, Springer, 2014.

[22] F. Facchinei, and J.-S. Pang, *Finite-Dimensional Variational Inequalities and Complementarity Problems*, Springer, 2003.

[23] M. Hintermüller, *Semismooth Newton Methods and Applications*, Department of Mathematics, Humboldt-University of Berlin, 2010.

[24] K. Ito, and K. Kunisch, *Lagrange Multiplier Approach to Variational Problems and Applications*, SIAM, 2008.

[25] D. Klatte, and B. Kummer, *Nonsmooth Equations in Optimization. Regularity, Calculus, Methods and Applications,* Kluwer, 2002.

[26] A. F. Izmailov, and M. V. Solodov, *Newton-type Methods for Optimization and Variational Problems*, Springer, 2014.

[27] M. Ulbrich, *Semismooth Newton Methods for Variational Inequalities and Constrained Optimization Problems in Function Spaces*, SIAM, 2011.

[28] S. M. Robinson, "Extension of Newton's method to nonlinear functions with values in a cone", *Numer. Math.*, vol. 19, pp. 341–347, 1972.

[29] S. M. Robinson, "Newton's method for a class of nonsmooth functions", *Set-Valued and Variational Analysis*, vol. 2(1), pp. 291–305, 1994.

[30] A. L. Dontchev, "Local analysis of a Newton-type method based on partial linearization", *Lectures in Applied Math*, vol. 32, pp. 295–306, 1996.

[31] S. Adly, R. Cibulka, and H. Van Ngai, "Newton's method for solving inclusions using set-valued approximations", *SIAM J. on Optimization*, vol. 25(1), pp. 159–184, 2015.

[32] S. Adly, H. Van Ngai, and V. V. Nguyen, "Newton's method for solving generalized equations: Kantorovich's and Smale's approaches", *J. of Mathematical Analysis and Application*, vol. 439, pp. 396–418, 2016.

[33] O. P. Ferreira, and B. F. Svaiter, "A robust Kantorovich's theorem on the inexact Newton method with relative residual error tolerance", *J. Complexity*, vol. 28, pp. 346–363, 2012.

[34] A. L. Dontchev, R. T. Rockafellar, "Convergence of inexact Newton methods for generalized equations" *Math. Programming*, vol. 139(1-2), pp. 115–137, 2013.

[35] K. Butts, A. Dontchev, M. Huang, and I. V. Kolmanovsky, "A perturbed chord (Newton-Kantorovich) method for constrained nonlinear model predictive control", *IFAC-Papers Online*, vol. 49(18), pp. 253–258, 2016.

[36] S. M. Robinson, "Strongly regular generalized equations", *Math. Oper. Res.* vol. 5, pp 43–62, 1980.

[37] N. H. Josephy, Newton's method for generalized equations and the PIES energy model, Ph.D. Dissertation, Department of Industrial Engineering, University of Wisconsin-Madison. 1979.

[38] R. T. Rockafellar, and R. J-B Wets, *Variational analysis*, Springer, 2009.

[39] M. Huang, H. Nakada, K. Butts, and I. V. Kolmanovsky, "Nonlinear model predictive control of a diesel engine airpath: A comparison of constraint handling and computational strategies", IFAC-PapersOnLine, vol. 48(23), pp. 372–379, 2015.

[40] C.T. Kelley, *Iterative Methods for Linear and Nonlinear Equations*, SIAM, 1995.

[41] A. Giuggiani, I. V. Kolmanovsky, P. Patrinos, and A. Bemporad. "Fixed point constrained Model Predictive Control of spacecraft attitude", *Proc. of the American Control Conference*, pp. 2317–2322, 2015.

[42] M. Jankovic, M. Jankovic, and I.V. Kolmanovsky, "Constructive Lyapunov control design for turbocharged diesel engines", *Control Systems Technology, IEEE Transactions on*, vol. 8(2), pp. 288–299, 2000.

APPENDIX

Proof of the implication (12) $\Rightarrow$ (7)

Recall that a set-valued mapping $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ is said to be monotone if

$$\langle y - y', x - x' \rangle \geq 0 \text{ whenever } (x, y), (x', y') \in \operatorname{gph} F.$$

A monotone mapping F is said to be maximal monotone whenever its graph cannot be enlarged to become the graph of another monotone mapping, and is strongly monotone if there exists a scalar $\sigma > 0$ such that

$$\langle y - y', x - x' \rangle \geq \sigma \|x - x'\|^2 \text{ whenever } (x, y), (x', y') \in \operatorname{gph} F.$$

Using these definitions, the following statement, also valid in Hilbert spaces, is well-known. Let $F : X \rightrightarrows X$ be maximal monotone and strongly monotone with constant $\sigma > 0$. Then, the inverse F^{-1} is single-valued and Lipschitz continuous on X with Lipschitz constant σ^{-1}.

To show that (12) implies (7), observe first that under (12) the mapping $F := A + N_C$ is maximal monotone. Further, $y \in Ax + N_C(x)$ is the same as having

$$\langle y - Ax, v - x \rangle \leq 0 \quad \text{for all} \quad v \in C.$$

Let $(x, y), (x', y') \in \operatorname{gph}(A + N_C)$. Then

$$\langle y - Ax, x' - x \rangle \leq 0 \quad \text{and} \quad \langle -y' + Ax', x' - x \rangle \leq 0.$$

Adding these inequalities we get

$$\langle y - y' + A(x' - x), x' - x \rangle \leq 0$$

and then

$$\langle A(x'-x), x'-x\rangle \le \langle y'-y, x'-x\rangle.$$

But then $A+N_C$ will be strongly monotone provided that there is $\beta>0$ such that

$$\beta\|x-x'\|^2 \le \langle A(x'-x), x'-x\rangle \quad \text{for all} \quad x, x' \in C.$$

This is condition (12).

Proof of Theorem 1

In the proof, we utilize the following simplified version of Robinson's inverse function theorem, see [21, Theorem 5G.3] with $a=b=\infty$:

Lemma 1. *Let $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ be a set-valued mapping whose inverse F^{-1} is single-valued and Lipschitz continuous everywhere with a Lipschitz constant κ. Then, for any $n\times n$ matrix B such that $\|B\| < 1/\kappa$, the mapping $(B+F)^{-1}$ is a Lipschitz continuous function with Lipschitz constant $\kappa/(1-\kappa\|B\|)$.*

Proof of Theorem 1. Utilizing Lemma 1, we get that the mapping $D^k=(A+B^k+N_C)^{-1}$ is single-valued everywhere, hence the iteration (6) generates a unique sequence $\{x^k\}$; in addition, each mapping D^k is a Lipschitz continuous function with Lipschitz constant $\kappa/(1-\kappa\mu)$. From (10), there exists $y^0\in\varphi(x^0)+N_C(x^0)$ such that $\|y^0\|+\|\sigma(x^0)\|\le at(1-t)$.

We will show by induction that the sequence $\{x^k\}$ generated by the method (6) has the following properties:

(i) $\|x^k-x^0\|\le a$;

(ii) $\|x^k-x^{k-1}\|\le at^k$ for each natural number k.

We use induction, starting with $k=1$. First, note that, by the iteration (6), there exists $r^0\in\mathbb{R}^n$ with $\|r^0\|\le\zeta^0\sigma(x^0)$ such that $x^1=D^0(-\varphi(x^0)+(A+B^0)x^0-r^0)$. Since $x^0=D^0(y^0-\varphi(x^0)+(A+B^0)x^0)$, from condition (a) we obtain

$$\|x^1-x^0\|\le\kappa(\|y^0\|+\|r^0\|)\le at(1-t)\le a.$$

Assume that for some natural number $k>1$ and for all $j=0,1,2,\dots,k-1$ there exists $r^j\in\mathbb{R}^n$ such that $\|r^j\|\le\zeta^j\sigma(x^j)$,

$$\|r^j-r^{j-1}\|\le\zeta^j l\|x^j-x^{j-1}\| \quad \text{for} \quad j=1,2,\dots,k-1 \quad (44)$$

and for $j=1,2,\dots,k-1$ the iterates

$$x^j=D^{j-1}(-\varphi(x^{j-1})+(A+B^{j-1})x^{j-1}-r^{j-1}) \quad (45)$$

satisfy conditions (i)–(ii) above. Clearly, from (b), the set-valued mapping $x\mapsto I\!B_{\zeta^j\sigma(x)}(0)$ is Lipschitz continuous in the Hausdorff metric, for any j, hence there exists $r^k\in\mathbb{R}^n$ such that $\|r^k\|\le\zeta^k\sigma(x^k)$ and also (44) holds for $j=k$. As for the case $j=k$, the left side of (45) implies

$$(A+B^{k-1})x^k+N_C(x_k)\ni-\varphi(x^{k-1})+(A+B^{k-1})x^{k-1}-r^{k-1}.$$

By adding and subtracting B^kx^k and rearranging the terms, we have

$$(A+B^k)x^k+N_C(x_k)\ni-\varphi(x^{k-1})+(A+B^{k-1})x^{k-1}+B^kx^k-B^{k-1}x^k-r^{k-1}.$$

As a result, (45) with $j=k$ can be re-written as

$$x^k=D^k(-\varphi(x^{k-1})+Ax^{k-1}+B^kx^k-B^{k-1}x^k+B^{k-1}x^{k-1}-r^{k-1}).$$

Subtracting this last equality from 45 with $j=k+1$, we get

$$x^{k+1}-x^k=D^k(-\varphi(x^k)+\varphi(x^{k-1})+A(x^k-x^{k-1})+B^{k-1}(x^k-x^{k-1})-r^k+r^{k-1}).$$

Utilizing (8), (11) and (44), we obtain

$$\begin{aligned}\|x^{k+1}-x^k\|&\le\frac{\kappa}{1-\kappa\mu}(\|-\varphi(x^k)+\varphi(x^{k-1})\\&\qquad+A(x^k-x^{k-1})\|\\&\quad+\|B^{k-1}(x^k-x^{k-1})\|+\|r^k-r^{k-1}\|\\&\le\frac{\kappa}{1-\kappa\mu}(\delta+\mu+\zeta^0l)\|x^k-x^{k-1}\|\\&=t\|x^k-x^{k-1}\|.\end{aligned}$$

Hence,

$$\|x^{k+1}-x^k\|\le t^{k-1}\|x^1-x_0\|\le t^ka(1-t).$$

Furthermore,

$$\|x^{k+1}-x^0\|\le\sum_{j=0}^{k}\|x^{j+1}-x^j\|\le a(1-t)\sum_{j=0}^{\infty}t^j=a.$$

This completes the induction step and hence (i) and (ii) hold for all k.

We shall prove now that the sequence $\{x^k\}$ is convergent. By condition (b), for any natural numbers k and p we have

$$\begin{aligned}\|x^{k+p+1}-x^k\|&\le\textstyle\sum_{j=k}^{k+p}\|x^{j+1}-x^j\|\le\\&\le\textstyle\sum_{j=k}^{k+p}a(1-t)t^j\le t^k\sum_{j=0}^{\infty}a(1-t)t^j=at^k.\end{aligned}$$

Hence, $\{x^k\}$ is a Cauchy sequence and is therefore convergent. Let $\bar x=\lim_{k\to\infty}x^k$. Since $B^k(x^{k+1}-x^k)\to 0$ and $\zeta^k\sigma(x^k)\to 0$, passing to the limit in (6) we obtain that $\bar x$ is a solution of (5). Also, by (a), $\|\bar x-x^0\|\le a$. Using (b), for any natural number k we have

$$\begin{aligned}\|x^k-\bar x\|&=\lim_{m\to\infty}\|x^k-x^{k+m}\|\le\\&\le\lim_{m\to\infty}\textstyle\sum_{i=k}^{k-1+m}\|x^i-x^{i+1}\|\le\\&\le\lim_{m\to\infty}\textstyle\sum_{i=k}^{k-1+m}a(1-t)t^i\le at^k.\end{aligned}$$

This completes the proof.

Proof of the considerations on Theorem 1

Consider now the exact Newton method (13) obtained from (6) by setting $A=\nabla\varphi(\bar x)$ for a solution $\bar x$ of (5), $B^k=\nabla\varphi(x^k)-\nabla\varphi(\bar x)$ and $\sigma\equiv 0$. We will now show how to derive from Theorem 1 the classical (Cauchy) convergence theorem for the method under the additional assumption that the function $x\mapsto d(0,\varphi(x)+N_C(x))$ is upper semicontinuous at $x=\bar x$. This assumption is strong but it holds when, for instance, C is a ball, a hyperplane, or the entire $\mathbb{R}^n$. However, it may be violated when C is a polyhedron and $\bar x$ is a vertex of it.

Suppose that the mapping $(A+N_C)^{-1}$ is a Lipschitz continuous function (as noted in Section I, this assumption can be localized as in the original version of the Josephy theorem) with a Lipschitz constant κ. This condition implies that $\bar x$ is the only solution of (5) in a neighborhood of $\bar x$ and, moreover, for every starting point x^0 there exists a unique sequence generated by (13). Choose a positive $\mu<1/(2\kappa)$; then there exists $\alpha>0$ such that

$$\|\nabla\varphi(x)-\nabla\varphi(\bar x)\|\le\mu \quad \text{for all} \quad x\in I\!B_\alpha(\bar x).$$

Choose $\delta > 0$ such that

$$t := \frac{\kappa(\delta + \mu)}{1 - \kappa\mu} < 1 \tag{46}$$

and make α smaller if necessary to obtain

$$\|\varphi(x) - \varphi(x') - \nabla\varphi(\bar{x})(x - x')\| \leq \delta\|x - x'\| \tag{47}$$

for all $x, x' \in I\!B_\alpha(\bar{x})$. By assumption, the function $x \mapsto d(0, \varphi(x)+N_C(x))$ is upper semicontinuous at $\bar{x}$, therefore we can make α smaller if necessary so that $d(0, \varphi(x)+N_C(x)) < t(1-t)$ for all $x \in I\!B_\alpha(\bar{x})$ and, moreover, $\bar{x}$ is the only solution of (5) in $I\!B_\alpha(\bar{x})$. Now choose $x^0 \in I\!B_\alpha(\bar{x})$. Then from Theorem 1 we obtain that the sequence generated by the iteration (13) converges linearly to a solution of (5) in $I\!B_\alpha(x^0)$ and this solution is of course $\bar{x}$. Furthermore, after adapting (45) to the changed notations, we have

$$x^{k+1} = D^k(-\varphi(x^k) + \nabla\varphi(x^k)x^k).$$

and

$$\bar{x} = D^k(\nabla\varphi(x^k)\bar{x} - \varphi(\bar{x})).$$

Hence,

$$\begin{aligned} &\|x^{k+1} - \bar{x}\| \leq \\ &\tfrac{\kappa}{1-\kappa\mu}\|-\varphi(x^k) + \varphi(\bar{x}) + \nabla\varphi(x^k)(x^k - \bar{x})\| \times \\ &\qquad \|x^k - \bar{x}\|. \end{aligned} \tag{48}$$

Taking into account (46) and (47) in the last inequality, we obtain that the sequence $\{x^k\}$ is linearly convergent to $\bar{x}$. Furthermore, from (48), taking into account that

$$\|-\varphi(x^k) + \varphi(\bar{x}) + \nabla\varphi(x^k)(x^k - \bar{x})\| \tag{49}$$

converges to zero with $k \to \infty$, we obtain that the convergence is actually superlinear. In addition, if we assume that the derivative $\nabla\varphi$ is Lipschitz continuous around $\bar{x}$, then the mean value theorem applied to (49), and the result used in (48), yields quadratic convergence.

Asen L. Dontchev received the PhD degree in control from Warsaw University of Technology and the PhD in Mathematical Sciences from the Bulgarian Academy of Sciences. He is Associate Editor at Mathematical Reviews and is also employed by the University of Michigan. His research interests include estimating the effect of perturbations and approximations on solutions of variational problems, in particular problems in optimization, control, differential equations, numerical analysis and approximation theory. He is the author of *Perturbations, approximations and sensitivity analysis of optimal control systems* Springer-Verlag 1983, *Well-posed optimization problems* (with T. Zolezzi) Springer-Verlag 1993, and *Implicit functions and solution mappings* (with R. T. Rockafellar), Springer 2014. He has served as a Program Director at the Division of Mathematical Sciences of NSF and in the boards of several journals including Journal of Mathematical Analysis and Applications, SIAM Journal on Control and Optimization, SIAM Journal on Optimization, Computational Optimization and Applications and Journal of Optimization Theory and Applications.

Mike Huang received the B.S. and M.S. degrees in electrical engineering and Ph.D. degree in aerospace engineering from the University of Michigan, Ann Arbor, MI, USA, in 2010, 2012, and 2015 respectively. He is currently a Senior Engineer with the Model Based Development Department, Toyota Motor North America R&D. His research interests include optimal control and numerical methods for automotive applications.

Ilya V. Kolmanovsky received the M.S. and Ph.D. degrees in aerospace engineering and the M.A. degree in mathematics from the University of Michigan, Ann Arbor, MI, USA, in 1993, 1995, and 1995, respectively. He was with Ford Research and Advanced Engineering, Dearborn, MI, USA. He is currently a Full Professor with the Department of Aerospace Engineering, University of Michigan. He is also an Inventor of 92 U.S. patents. His current research interests include control theory for systems with state and control constraints, and control applications to aerospace and automotive systems. Dr. Kolmanovsky was a recipient of the Donald P. Eckman Award of American Automatic Control Council.

Marco M. Nicotra completed a T.I.M.E. double degree program in 2012, receiving the M.S. degrees in mechanical and electromechanical engineering from Politecnico di Milano and Université Libre de Bruxelles, respectively. In 2016 he received the Ph.D. degree in control engineering under the joint supervision of the Université Libre de Bruxelles and the University of Bologna. In 2014 he received a honorable mention for the Young Author Prize of the 19th IFAC World Congress in Cape Town, South Africa. He is currently a Research Fellow at the Aerospace Department of the University of Michigan. His current research focuses on the constrained control of nonlinear systems and unmanned aerial vehicle applications.