Local Convergence of Sequential Convex Programming for Nonconvex Optimization

Quoc Tran Dinh* and Moritz Diehl*

*Department of Electrical Engineering (SCD-ESAT) and OPTEC, Katholieke Universiteit Leuven, Kasteelpark Arenberg 10, 3001-Heverlee, Belgium
{quoc.trandinh, moritz.diehl}@esat.kuleuven.be

Summary. This paper introduces sequential convex programming (SCP), a local optimzation method for solving nonconvex optimization problems. A full-step SCP algorithm is presented. Under mild conditions the local convergence of the algorithm is proved as a main result of this paper. An application to optimal control illustrates the performance of the proposed algorithm.

1 Introduction and Problem Statement

Consider the following nonconvex optimization problem:

$$\begin{cases} \min\limits_{x} \; c^T x \\ \text{s.t. } \; g(x) = 0, \; x \in \Omega, \end{cases} \tag{P}$$

where $c \in \mathbf{R}^n$, $g : \mathbf{R}^n \to \mathbf{R}^m$ is non-linear and smooth on its domain, and Ω is a nonempty closed convex subset in $\mathbf{R}^n$.

This paper introduces *sequential convex programming* (SCP), a local optimization method for solving the nonconvex problem (P). We prove that under acceptable assumptions the SCP method locally converges to a KKT point[1] of (P) and the rate of convergence is linear.

Problems in the form of (P) conveniently formulate many problems of interest such as least squares problems, quadratically constrained quadratic programming, nonlinear semidefinite programming (SDP), and nonlinear second order cone programming problems (see, e.g., [1, 2, 5, 6, 10]). In nonlinear optimal control, by using direct transcription methods, the resulting problem is usually formulated as an optimization problem of the form (P) where the equality constraint $g(x) = 0$ originates from the dynamic system of an optimal control problem.

The main difficulty of the problem (P) is concentrated in the nonlinear constraint $g(x) = 0$ that can be overcome by linearizing it around the current

[1] KKT stands for "**K**arush-**K**uhn-**T**ucker".

iteration point and maintaining the remaining convexity of the original problem. This approach differs from sequential quadratic programming, Gauss-Newton or interior point methods as it keeps even nonlinear constraints in the subproblems as long as they are convex.

Optimization algorithms using convex approximation approaches have been proposed and investigated by Fares *et al.* [4] for nonlinear SDP and Jarre [8] for nonlinear programming. Recently, Lewis and Wright [12] introduced a proximal point method for minimizing the composition of a general convex function h and a smooth function c using the convex approximation of $h(c(\cdot))$.

1.1. Contribution. In this paper, we first propose a *full-step SCP algorithm* for solving (P). Then we prove the local convergence of this method. The main contribution of this paper is Theorem 1, which estimates the local contraction and shows that the *full-step SCP algorithm* converges linearly to a KKT point of the problem (P). An application in optimal control is implemented in the last section.

1.2. Problem Statement. Throughout this paper, we assume that g is twice continuously differentiable on its domain. As usual, we define the Lagrange function of (P) by $L(x, \lambda) := c^T x + \lambda^T g(x)$ and the KKT condition associated with (P) becomes

$$\begin{cases} 0 \in c + \nabla g(x)\lambda + N_\Omega(x), \\ 0 = g(x), \end{cases} \tag{1}$$

where $\nabla g(x)$ denotes the Jacobian matrix of g at x. The multivalued mapping

$$N_\Omega(x) := \begin{cases} \{w \in \mathbf{R}^n \mid w^T(y - x) \leq 0,\ y \in \Omega\} & \text{if } x \in \Omega, \\ \emptyset & \text{otherwise} \end{cases} \tag{2}$$

is the normal cone of the convex set Ω at x. A pair $z^* := (x^*, \lambda^*)$ satisfying (1) is called a KKT point and x^* is called a stationary point of (P). We denote by Γ^* and S^* the sets of the KKT and the stationary points of (P), respectively. Note that the first line of (1) includes implicitly the condition $x \in \Omega$ due to definition (2). Let us define $K := \Omega \times \mathbf{R}^m$ and introduce a new mapping φ as follows

$$\varphi(z) := \begin{pmatrix} c + \nabla g(x)\lambda \\ g(x) \end{pmatrix}, \tag{3}$$

where z stands for (x, λ) in $\mathbf{R}^{n+m}$. Then the KKT condition (1) can be regarded as a generalized equation:

$$0 \in \varphi(z) + N_K(z), \tag{4}$$

where $N_K(z)$ is the normal cone of K at z.

The generalized equation (4) can be considered as a basic tool for studying variational inequalities, complementarity problems, fixed point problems and mathematical programs with equilibrium constraints. In the landmark paper

[13], Robinson introduced a condition for generalized equation (4), which is called *strong regularity*. This assumption is then used to investigate the solution of (4) under the influence of perturbations. *Strong regularity* becomes a standard condition in variational analysis as well as in optimization. It is important to note that (see [3]) the generalized equation (4) is strongly regular at $z^* \in \Gamma^*$ *if and only if* the strong second order sufficient condition (SOSC) of (P) holds at this point whenever Ω is polyhedral and the LICQ condition[2] is satisfied. Many research papers which have studied the stability and sensitivity in parametric optimization and optimal control also used the strong regularity property (see, e.g., [11, 14]).

1.4. Sequential Convex Programming Framework. The *full-step sequential convex programming algorithm* for solving (P) is an iterative method that generates a sequence $\{z^k\}_{k\geq 0}$ as follows:

1. Choose an initial point x^0 inside the convex set Ω and λ^0 in $\mathbf{R}^m$. Set $k := 0$.
2. For a given x^k, solve the following convex subproblem:

$$\begin{cases} \min\limits_{x} c^T x \\ \text{s.t. } g(x^k) + \nabla g(x^k)^T(x - x^k) = 0, \\ \qquad x \in \Omega, \end{cases} \qquad (\mathrm{P}_{\mathrm{cvx}}(x^k))$$

   to obtain a solution $x_+(x^k)$ and the corresponding Lagrange multiplier $\lambda_+(x^k)$. Set $z_+(x^k) := (x_+(x^k), \lambda_+(x^k))$. If $\|z_+(x^k) - z^k\| \leq \varepsilon$ for a given tolerance $\varepsilon > 0$, then stop. Otherwise, set $z^{k+1} := z_+(x^k)$, increase k by 1 and go back to Step 2.

As we will show later, the iterative sequence $\{z^k\}$ generated by the *full-step SCP algorithm* converges to a KKT point z^* of the original problem (P), if it starts sufficiently close to z^* and the contraction property is satisfied (see Theorem 1 below).

In practice, this method should be combined with globalization strategies such as line search or trust region methods in order to ensure global convergence, if the starting point is arbitrary. Since Ω is convex, projection methods can be used to find an initial point x^0 in Ω.

Lemma 1. *If x^k is a stationary point of $P_{cvx}(x^k)$ then it is a stationary point of the problem (P).*

Proof. We note that x^k always belongs to Ω. Substituting x^k into the KKT condition of the subproblem $\mathrm{P}_{\mathrm{cvx}}(x^k)$, it collapses to (1).

2 Local convergence of SCP methods

Suppose that $x^k \in \Omega$, $k \geq 0$, is the current iteration associated with $\lambda^k \in \mathbf{R}^m$. Then the KKT condition of the convex subproblem $\mathrm{P}_{\mathrm{cvx}}(x^k)$ becomes

[2] LICQ stands for "**L**inear **I**ndependence **C**onstraint **Q**ualification".

$$\begin{cases} 0 \in c + \nabla g(x^k)\lambda + N_{\Omega}(x), \\ 0 = g(x^k) + \nabla g(x^k)^T(x - x^k), \end{cases} \tag{5}$$

where λ is the corresponding multiplier. Suppose that the Slater constraint qualification condition holds for $\mathrm{P}_{\mathrm{cvx}}(x^k)$, i.e.,

$$\mathrm{relint}\ \Omega \cap \{x \mid g(x^k) + \nabla g(x^k)^T(x - x^k) = 0\} \neq \emptyset,$$

where $\mathrm{relint}\Omega$ is the set of the relative interior points of Ω. In other words, there exists a strictly feasible point of $\mathrm{P}_{\mathrm{cvx}}(x^k)$. Then by convexity of Ω, a point $(x_+(x^k), \lambda_+(x^k))$ is a KKT point of $\mathrm{P}_{\mathrm{cvx}}(x^k)$ if and only if $x_+(x^k)$ is a solution of (5) corresponding to the multiplier $\lambda_+(x^k)$. In the sequel, we use z for a pair (x, λ), z^* and $z_+(x^k)$ are a KKT point of (P) and $\mathrm{P}_{\mathrm{cvx}}(x^k)$, respectively. We denote by

$$\hat{\varphi}(z; x^k) := \begin{pmatrix} c + \nabla g(x^k)\lambda \\ g(x^k) + \nabla g(x^k)^T(x - x^k) \end{pmatrix}, \tag{6}$$

a linear mapping and $K := \Omega \times \mathbf{R}^m$. For each $x^* \in S^*$, we define a multivalued function:

$$L(z; x^*) := \hat{\varphi}(z; x^*) + N_K(z), \tag{7}$$

and $L^{-1}(\delta; x^*) := \{z \in \mathbf{R}^{n+m} : \delta \in L(z; x^*)\}$ for $\delta \in \mathbf{R}^{n+m}$ is its inverse mapping. To prove local convergence of the *full-step SCP algorithm*, we make the following assumptions:

(A1) The set of KKT points Γ^* of (P) is nonempty.
(A2) Let $z^* \in \Gamma^*$. There exists a neighborhood $U \subset \mathbf{R}^{n+m}$ of the origin and Z of z^* such that for each $\delta \in U$, $z^*(\delta) := L^{-1}(\delta; x^*) \cap Z$ is single valued. Moreover, the mapping $z^*(\cdot)$ is Lipschitz continuous on U with a Lipschitz constant $\gamma > 0$, i.e.,

$$\|z^*(\delta) - z^*(\delta')\| \leq \gamma \|\delta - \delta'\|, \quad \forall \delta, \delta' \in U. \tag{8}$$

(A3) There exists a constant $0 < \kappa < 1/\gamma$ such that $\|E_g(z^*)\| \leq \kappa$, where $E_g(z^*)$ is the Hessian of the Lagrange function L with respect to the argument x at $z^* = (x^*, \lambda^*)$ defined by

$$E_g(z) := \sum_{i=1}^{m} \lambda_i \nabla^2 g_i(x). \tag{9}$$

Remark 1. By definition of $\hat{\varphi}(\cdot;\cdot)$, we can refer to x^k as a parameter of this mapping and $\mathrm{P}_{\mathrm{cvx}}(x^k)$ can be considered as a parametric convex problem with respect to the parameter x^k.
i) It is easy to show that z^* is a solution to $0 \in \varphi(z) + N_K(z)$ if and only if it is a solution to $0 \in \hat{\varphi}(z; x^*) + N_K(z)$.
ii) Assumption (**A3**) implies that either the function g should be "weakly

nonlinear" (small second derivatives) in a neighborhood of a stationary point or the corresponding Lagrange multipliers are sufficiently small in the neighborhood of λ^*. The latter case occurs if the optimal objective value of (P) depends only weakly on perturbations of the nonlinear constraint $g(x) = 0$.
iii) Assumption (**A2**) is the strong regularity condition of the parametric generalized equation $0 \in \hat{\varphi}(z; x^k) + N_K(z)$ at (z^*, x^*) in the sense of Robinson [13].

For the assumption (**A2**), by linearity of $\hat{\varphi}$, we have $\hat{\varphi}(z; x^*) = \hat{\varphi}(z^*; x^*) + \nabla\hat{\varphi}(z^*; x^*)^T(z - z^*)$ where matrix $\nabla\hat{\varphi}(z)$ is defined by

$$\nabla\hat{\varphi}(z; x^*) := \begin{bmatrix} 0 & \nabla g(x^*) \\ \nabla g(x^*)^T & 0 \end{bmatrix}, \tag{10}$$

which may be singular even if $\nabla g(x^*)$ is full-rank. It is easy to see that $L(z; x^*)$ defined by (7) has the same form as $\hat{L}(z; x^*) := \hat{\varphi}(z^*, x^*) + \nabla\hat{\varphi}(z^*; x^*)(z - z^*) + N_K(z)$ a linearization of (4) at (z^*, x^*).
To make the strong regularity assumption clear in the sense of mathematical programming, for a given neighborhood U of 0 and Z of z^*, we define the following perturbed convex programming problem:

$$\begin{cases} \min\limits_{x} \ (c + \delta_{\mathrm{c}})^T(x - x^*) \\ \text{s.t.} \ \ g(x^*) + \delta_{\mathrm{g}} + \nabla g(x^*)^T(x - x^*) = 0, \\ \qquad x \in \Omega, \end{cases} \tag{$\mathrm{P_{cvx}}(x^*;\delta)$}$$

where $\delta = (\delta_{\mathrm{c}}, \delta_{\mathrm{g}})$ is a perturbation (or a parameter) vector. The Slater condition associated with $\mathrm{P_{cvx}}(x^*; \delta)$ becomes

$$\text{relint } \Omega \cap \{x \mid g(x^*) + \delta_{\mathrm{g}} + \nabla g(x^*)^T(x - x^*) = 0\} \neq \emptyset. \tag{11}$$

Then the assumption (A2) holds if and only if $z^*(\delta)$ is the unique KKT point of $\mathrm{P_{cvx}}(x^*; \delta)$, and this solution is Lipschitz continuous on U with a Lipschitz constant $\gamma > 0$ provided that (11) holds.

The *full-step SCP algorithm* is called to be well-defined if the convex subproblem $\mathrm{P_{cvx}}(x^k)$ has at least one KKT point $z_+(x^k)$ provided that z^k is sufficiently close to $z^* \in \Gamma^*$. In this case, the subproblem $\mathrm{P_{cvx}}(x^k)$ is said to be solvable.

Lemma 2. *Suppose that Assumptions (*__A1__*)-(*__A3__*) are satisfied, then the full-step SCP algorithm is well-defined.*

Proof. It follows from Remark 1 (i) that the parametric generalized equation $0 \in \hat{\varphi}(z; x^k) + N_K(z)$ is strongly regular at (z^*, x^*) according to Assumption (A2), where x^k is referred as a parameter. Applying Theorem 2.1 [13], we conclude that there exists a neighborhood X of x^* such that the generalized equation $0 \in \hat{\varphi}(z; x^k) + N_K(z)$ has unique solution $z_+(x^k)$ for all $x^k \in X$, which means that $z_+(x^k)$ is a KKT point of $\mathrm{P_{cvx}}(x^k)$. □

The main result of this paper is the following theorem.

Theorem 1. *[Local Contraction] Suppose that Assumptions (**A1**)-(**A3**) are satisfied. Suppose further for $z^* \in \Gamma^*$ that g is twice continuously differentiable on a neighborhood of x^*. Then the full-step SCP algorithm is well-defined and there exists $\rho > 0$ such that for all $z^k \in B(z^*, \rho)$ we have:*

$$\|z_+(x^k) - z^*\| \leq \alpha \|z^k - z^*\|, \tag{12}$$

where $\alpha \in (0,1)$ does not depend on z^k and $z_+(x^k)$. Thus, if the initial point z^0 is sufficiently close to z^* then the sequence $\{z^k\}$ generated by full-step SCP algorithm converges to z^* linearly.

Proof. Note that $\Gamma^* \neq \emptyset$ by (**A1**), take any $z^* \in \Gamma^*$. Then the well-definedness of the *full-step SCP algorithm* follows from Lemma 2. By assumption (**A3**) that $\gamma\kappa < 1$ we can choose $\varepsilon := \frac{(1-\gamma\kappa)}{(4\sqrt{22}+2\sqrt{3})\gamma} > 0$. Since g is twice continuously differentiable on a neighborhood X of x^* and $E(x, \lambda)$ defined by (9) is linear with respect to λ, it implies that, for a given $\varepsilon > 0$ defined as above, there exists a positive number $r_0 > 0$ such that $\|\nabla g(x) - \nabla g(x^k)\| \leq \varepsilon$, $\|\nabla g(x) - \nabla g(x^*)\| \leq \varepsilon$, $\|E_g(z) - E_g(z^*)\| \leq \varepsilon$ and $\|E_g(z) - E_g(z^k)\| \leq \varepsilon$ for all $z = (x, \lambda) \in B(z^*, r_0)$ and $z^k = (x^k, \lambda^k) \in B(z^*, r_0)$, where $B(z^*, r_0)$ is the closed ball of radius r_0 centered at z^*.

Take any $z \in B(z^*, r_0) \subseteq Z$ and define the residual quantity

$$\delta(z; x^*, x^k) := \hat{\varphi}(z; x^*) - \hat{\varphi}(z; x^k). \tag{13}$$

This quantity can be expressed as

$$\begin{aligned}
\delta(z; x^*, x^k) &= [\hat{\varphi}(z; x^*) - \varphi(z^*)] + [\varphi(z^*) - \varphi(z)] \\
&\quad + [\varphi(z) - \varphi(z^k)] + [\varphi(z^k) - \hat{\varphi}(z; x^k)] \\
&= \int_0^1 M(z_t^k; x^k)(z - z^k)dt - \int_0^1 M(z_t^*; x^*)(z - z^*)dt \\
&= \int_0^1 [M(z_t^k; x^k) - M(z_t^*; x^*)](z - z^k)dt \\
&\quad - \int_0^1 M(z_t^*; x^*)(z^k - z^*)dt,
\end{aligned} \tag{14}$$

where $z_t^* := z^* + t(z - z^*)$, $z_t^k := z^k + t(z - z^k)$ with $t \in [0,1]$, and the matrix M is defined by

$$M(\tilde{z}; \hat{x}) := \begin{bmatrix} E_g(\tilde{z}) & \nabla g(\tilde{x}) - \nabla g(\hat{x}) \\ \nabla g(\tilde{x})^T - \nabla g(\hat{x})^T & 0 \end{bmatrix}. \tag{15}$$

Since $t \in [0,1]$, the points z_t^k and z_t^* must belong to $B(z^*, r_0)$. Using the following inequalities

$$\|E_g(z_t^k) - E_g(z_t^*)\| \leq \|E_g(z_t^k) - E_g(z^*)\| + \|E_g(z_t^*) - E_g(z^*)\| \leq 2\varepsilon,$$
$$\|\nabla g(x_t^k) - \nabla g(x_t^*)\| \leq \|\nabla g(x_t^k) - \nabla g(x^*)\| + \|\nabla g(x_t^*) - \nabla g(x^*)\| \leq 2\varepsilon,$$
$$\text{and } \|\nabla g(x^k) - \nabla g(x^*)\| \leq \varepsilon,$$

it follows that

$$\begin{aligned}\|M(z_t^k; x^k) - M(z_t^*; x^*)\|^2 &\leq \|E_g(z_t^k) - E_g(z_t^*)\|^2 \\ &+ 2[\|\nabla g(x_t^k) - \nabla g(x_t^*)\| + \|\nabla g(x^k) - \nabla g(x^*)\|]^2 \\ &\leq 22\varepsilon^2.\end{aligned}$$

This inequality implies that

$$\|M(z_t^*; x^*) - M(z_t^k; x^k)\| \leq \sqrt{22}\varepsilon. \tag{16}$$

Similarly, using Assumption (**A3**), we can estimate

$$\begin{aligned}\|M(z_t^*; x^*)\|^2 &\leq \|E_g(z_t^*)\|^2 + 2\|\nabla g(x_t^*) - \nabla g(x^*)\|^2 \\ &\leq 2\varepsilon^2 + [\|E_g(z_t^*) - E_g(z^*)\| + \|E_g(z^*)\|]^2 \\ &\leq 2\varepsilon^2 + (\varepsilon + \kappa)^2 \\ &\leq (\kappa + \sqrt{3}\varepsilon)^2.\end{aligned} \tag{17}$$

Combining (14), (16) and (17) together we obtain

$$\|\delta(z, x^*, x^k\| \leq (\kappa + \sqrt{3}\varepsilon)\|z^k - z^*\| + \sqrt{22}\varepsilon\|z - z^k\|. \tag{18}$$

Alternatively, we first shrink $B(z^*, r_0)$, if necessary, such that $\delta(z, x^*; x^k) \in U$ and then apply Assumption (**A2**) to imply that there exists $\tilde{z}(\delta) = (\tilde{x}(\delta), \tilde{\lambda}(\delta)) \in B(z^*, r_0)$ a solution of $\delta \in L(\cdot; z^*)$ for all $\delta \in U$ satisfying

$$\|\tilde{z}(\delta) - z^*\| \leq \gamma\|\delta\|. \tag{19}$$

If we recall $z_+(x^k)$ a KKT point of $\mathrm{P}_{\mathrm{cvx}}(x^k)$, one has $0 \in \hat{\varphi}(z_+(x^k); x^k) + N_K(z_+(x^k))$ which implies $\delta(z_+(x^k); x^*, x^k) \in \hat{\varphi}(z_+(x^k); x^*) + N_K(z_+(x^k))$ by definition of δ. Therefore, it follows from (19) that

$$\|z_+(x^k) - z^*\| \leq \gamma\|\delta(z_+(x^k); x^*, x^k)\|. \tag{20}$$

Substituting z by $z_+(x^k)$ into (18) and then merging with (20) we get

$$\|z_+(x^k) - z^*\| \leq (\gamma\kappa + \sqrt{3}\gamma\varepsilon)\|z^k - z^*\| + \sqrt{22}\gamma\varepsilon\|z_+(x^k) - z^k\|. \tag{21}$$

Using the triangle inequality $\|z_+(x^k) - z^k\| \leq \|z_+(x^k) - z^*\| + \|z^k - z^*\|$ for the right hand side of (21), after a simple rearrangement, the inequality (21) implies

$$\|z_+(x^k) - z^*\| \leq \frac{[\gamma\kappa + (\sqrt{22}+\sqrt{3})\gamma\varepsilon]}{1-\sqrt{22}\gamma\varepsilon}\|z^k - z^*\|. \tag{22}$$

Let us denote $\alpha := \frac{[\gamma\kappa+(\sqrt{22}+\sqrt{3})\gamma\varepsilon]}{1-\sqrt{22}\gamma\varepsilon}$. From the choice of ε, it is easy to show that

$$\alpha = \frac{(3\sqrt{22}+\sqrt{3})\gamma\kappa + \sqrt{22}+\sqrt{3}}{3\sqrt{22}+2\sqrt{3}+\sqrt{22}\gamma\kappa} \in (0,1). \tag{23}$$

Thus the inequality (22) is rewritten as

$$\|z_+(x^k) - z^*\| \leq \alpha\|z^k - z^*\|, \quad \alpha \in (0,1), \tag{24}$$

which proves (12).

If the starting point $z^0 \in B(z^*, r_0)$ then we have $\|z^1 - z^*\| \leq \alpha\|z^0 - z^*\| \leq \|z^0 - z^*\|$, which shows that $z^1 \in B(z^*, r_0)$. By induction, we conclude that the whole sequence $\{z^k\}$ is contained in $B(z^*, r_0)$. The remainder of the theorem follows directly from (12).

Remark 2. It is easy to see from (23) that $\alpha \in (\gamma\kappa, 1)$.

3 Numerical Results

In this section, we apply the SCP method to the optimal control problem arising from the optimal maneuvers of a rigid asymmetric spacecraft [7, 9]. The Euler equations for the angular velocity $\omega = (\omega_1, \omega_2, \omega_3)^T$ of the spacecraft are given by

$$\begin{cases} \dot{\omega}_1 = -\frac{(I_3-I_2)}{I_1}\omega_2\omega_3 + \frac{u_1}{I_1}, \\ \dot{\omega}_2 = -\frac{(I_1-I_3)}{I_2}\omega_1\omega_3 + \frac{u_2}{I_2}, \\ \dot{\omega}_3 = -\frac{(I_2-I_1)}{I_3}\omega_1\omega_2 + \frac{u_3}{I_3}, \end{cases} \tag{25}$$

where $u = (u_1, u_2, u_3)^T$ is the control torque; $I_1 = 86.24$ kg.m^2, $I_1 = 85.07$ kg.m^2 and $I_3 = 113.59$ kg.m^2 are the spacecraft principal moments of inertia. The performance index to be minimized is given by (see [7]):

$$J := \frac{1}{2}\int_0^{t_f} \|u(t)\|^2 dt. \tag{26}$$

The initial condition $\omega(0) = (0.01, 0.005, 0.001)^T$, and the terminal constraint is

$$\omega(t_f) = (0,0,0)^T \text{ (Case 1)} \;\text{ or }\; \omega(t_f)^T S_f \omega(t_f) \leq \rho_f \text{ (Case 2)}, \tag{27}$$

where matrix S_f is symmetric positive definite and $\rho_f > 0$. Matrix S_f is computed by using the discrete-time Riccati equation of the linearized form of (25) and ρ is taken by $\rho := 10^{-6} \times \lambda_{\max}(S_f)$, where $\lambda_{\max}(S_f)$ is the maximum eigenvalue of S_f. The additional inequality constraint is

$$\omega_1(t) - (5 \times 10^{-6} t^2 - 5 \times 10^{-4} t + 0.016) \leq 0, \tag{28}$$

for all $t \in [0, t_f]$ (see [7]).

In order to apply the SCP algorithm, we use the direct transcription method to transform the optimal control problem into a nonconvex optimization problem. The dynamic system is discretized based on the forward Euler scheme. With the time horizon $t_f = 100$, we implement the SCP algorithm for H_p (the number of the discretization points) from 100 to 500. The size (n, m, l) of the optimization problem goes from $(603, 300, 104)$ to $(3003, 1500, 504)$, where n is the number of variables, m is the number of equality constraints, and l is the number of inequality constraints.

We use an open source software (CVX) to solve the convex subproblems $\mathrm{P}_{\mathrm{cvx}}(x^k)$ and combine it with a line search strategy to ensure global convergence (not covered by this paper's theory). All the computational results are performed in Matlab 7.9.0 (2009) running on a desktop PC Pentium IV (2.6GHz, 512Mb RAM).

If we take the tolerance `TolX` $= 10^{-7}$ then the number of iterations goes from 3 to 6 iterations depending on the size of the problem. Note that the resulting convex subproblems in Case 1 are convex quadratic, while, in Case 2, they are quadratically constrained quadratic programming problems.

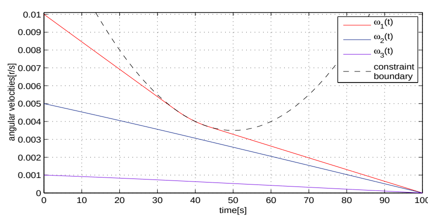

Fig1. Optimal angular velocities [Case 1]

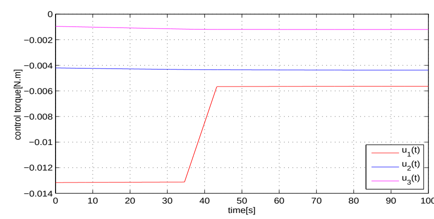

Fig2. Optimal control torques [Case 1]

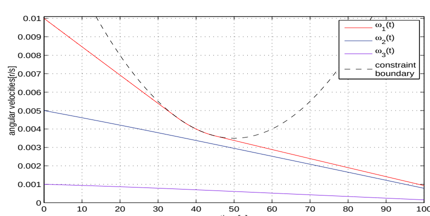

Fig3. Optimal angular velocities [Case 2]

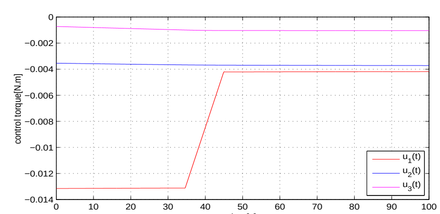

Fig4. Optimal control torques [Case 2]

Figure 1 (resp. Figure 3) shows the optimal angular velocity $\omega(t)$ of the rigid asymmetric spacecraft from 0 to $100s$ for Case 1 (resp. Case 2) with $H_p = 500$. The results show that $\omega_1(t)$ constrained by (28) touches its boundary around the point $t = 39s$ and $\omega(t)$ tends to zero at the end ($t = 100s$) identical to the results in [7]. Figure 2 (resp. Figure 4) shows the optimal torque $u(t)$ of the rigid asymmetric spacecraft for Case 1 (resp. Case 2). The rate of convergence is illustrated in Figures 5 and 6 for Case 1 and Case 2,

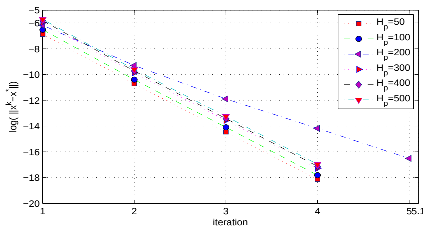

Fig5. Rate of Convergence [Case 1]

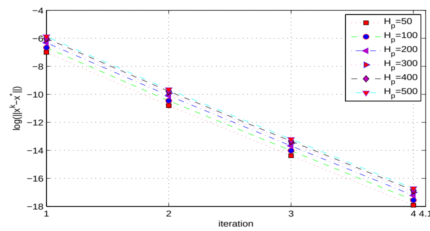

Fig6. Rate of Convergence [Case 2]

respectively. As predicted by the theoretical results in this paper, the rate of convergence shown in these figures is linear (with very fast contraction rate) for all the cases we implemented.

Acknowledgments. This research was supported by Research Council KUL: CoE EF/05/006 Optimization in Engineering(OPTEC), GOA AMBioRICS, IOF-SCORES4CHEM, several PhD/ postdoc & fellow grants; the Flemish Government via FWO: PhD/postdoc grants, projects G.0452.04, G.0499.04, G.0211.05, G.0226.06, G.0321.06, G.0302.07, G.0320.08 (convex MPC), G.0558.08 (Robust MHE), G.0557.08, G.0588.09, research communities (ICCoS, ANMMM, MLDM) and via IWT: PhD Grants, McKnow-E, Eureka-Flite+EU: ERNSI; FP7-HD-MPC (Collaborative Project STREP-grantnr. 223854), Erbocon, Contract Research: AMINAL, and Helmholtz Gemeinschaft: viCERP; Austria: ACCM, and the Belgian Federal Science Policy Office: IUAP P6/04 (DYSCO, Dynamical systems, control and optimization, 2007-2011).

The authors are very much thankful to the anonymous referees, who corrected numerous mistakes and suggested several improvements.

References

1. S. Boyd and L. Vandenberghe (2004). *Convex optimization*. University Press, Cambridge.
2. R. Correa and H. Ramirez C (2004). A global algorithm for nonlinear semidefinite programming. *SIAM J. Optim.*, 15(1):303–318.
3. A. L. Dontchev and T. R. Rockafellar (1996). Characterizations of strong regularity for variational inequalities over polyhedral convex sets. *SIAM J. Optim.*, 6(4):1087–1105.
4. B. Fares, D. Noll, and P. Apkarian (2002). Robust control via sequential semidefinite programming. *SIAM J. Control Optim.*, 40:1791–1820.
5. R. W. Freund, F. Jarre, and C. H. Vogelbusch (2007). Nonlinear semidefinite programming: sensitivity, convergence, and an application in passive reduced-order modeling. *Mathematical Programming*, Ser. B, 109:581–611.
6. M. Fukushima, Z.-Q. Luo, and P. Tseng (2003). A sequential quadratically constrained quadratic programming method for differentiable convex minimization. *SIAM J. Optimization*, 13(4):1098–1119.
7. H. Jaddu (2002). Direct solution of nonlinear optimal control problems using quasilinearization and Chebyshev polynomials. *Journal of the Franklin Institute*, 339:479–498.
8. F. Jarre (2003). On an approximation of the Hessian of the Lagrangian. *Optimization Online* (http://www.optimization−online.org/DB_HTML/2003/12/800.html).
9. J.L. Junkins and J.D. Turner (1986). *Optimal spacecraft rotational maneuvers*. Elsevier, Amsterdam.
10. C. Kanzow, C. Nagel, H. Kato and M. Fukushima (2005). Successive linearization methods for nonlinear semidefinite programs. *Computational Optimization and Applications*, 31:251–273.
11. D. Klatte and B. Kummer (2001). *Nonsmooth equations in optimization: regularity, calculus, methods and applications*. Springer-Verlag, New York.
12. A. S. Lewis and S. J. Wright (2008). A proximal method for composite minimization. http://arxiv.org/abs/0812.0423.
13. S. M. Robinson (1980). Strong regularity generalized equations, *Mathematics of Operation Research*, 5(1):43–62.
14. R. T. Rockafellar and R. J-B. Wets (1997). *Variational analysis*. Springer-Verlag, New York.