Automatica 157 (2023) 111279

Contents lists available at ScienceDirect

Automatica

journal homepage: www.elsevier.com/locate/automatica

# Dealing with infeasibility in multi-parametric programming for application to explicit model predictive control[☆]

Alessandro Falsone [a,*], Federico Bianchi [b], Maria Prandini [a]

[a] Dipartimento di Elettronica, Informazione e Bioingegneria, Politecnico di Milano, Via Ponzio 34/5, 20133 Milano, Italy
[b] Department of Generation Technologies and Materials, Ricerca sul Sistema Energetico (RSE) S.p.A., Via Rubattino 54, 20134 Milano, Italy

ARTICLE INFO

*Article history:*
Received 23 January 2023
Received in revised form 19 June 2023
Accepted 31 July 2023
Available online 5 September 2023

*Keywords:*
Multi-parametric programming
Explicit model predictive control
Exact penalty

ABSTRACT

Motivated by explicit model predictive control, we address infeasibility in multi-parametric quadratic programming according to the exact penalty function approach, where some user-chosen parameter-dependent constraints are relaxed and the 1-norm of their violation is penalized in the cost function. We characterize the relation between the resulting multi-parametric quadratic program and the original one and show that, as the penalty coefficient grows to infinity, the solution to the former provides a piecewise affine continuous function, which is an optimal solution for the latter over the feasibility region, while it minimizes the 1-norm of the relaxed constraints violation over the infeasibility region.



*Notation.* We denote with $\mathbb{R}$ the set of real numbers, and with $\mathbb{R}_+$ the set of non-negative real numbers. Given a matrix $M \in \mathbb{R}^{n_r,n_c}$, $m^{[i]}$ denotes its $i$th row while $M^I$ denotes the submatrix containing the rows of $M$ indexed by $I \subseteq \{1,\dots,n_r\}$. The transpose of $M$ is denoted with $M^\top$. For a symmetric matrix $M = M^\top$, $M \succ 0$ ($M \succeq 0$) denotes that $M$ is positive (semi-)definite. The $n$-dimensional vector containing all ones/zeros are denoted with $1_n/0_n$, the subscript will be omitted when clear from the context. The $n$-by-$m$ zero matrix is denoted with $0_{n,m}$. For a vector $v \in \mathbb{R}^n$, $\|v\|_1$ denotes its 1-norm and $[v]_+ = \max\{v, 0_n\}$, where the maximum among vectors has to be intended as component-wise. Given two vectors $u$ and $v$, $u \le v$ means that each component of $u$ is less than or equal to the corresponding component of vector $v$, $u < v$ is the same but with the strict inequality, while we use $u \leqslant v$ to denote that some (possibly none or all, but not for sure) components of $u$ are strictly less than the corresponding component of $v$ and the rest are less than or equal to. This is useful in handling polyhedral partitions, whose regions are polyhedra described by linear inequalities, since the border can be assigned only to one of two neighboring regions leading to strict inequalities in the description of the other one.

## 1. Introduction

We consider the following multi-parametric Quadratic Program (mp-QP)

$$\begin{aligned}\min_z \quad & \tfrac{1}{2} z^\top Q z + (F\vartheta + c)^\top z \\ \text{s.t:} \quad & G_h z \le b_h + S_h \vartheta \\ & G_s z \le b_s + S_s \vartheta\end{aligned} \tag{1}$$

where $z \in \mathbb{R}^{n_z}$ is the decision vector, $\vartheta \in \mathbb{R}^{n_\vartheta}$ is the parameter vector, $0 \preceq Q = Q^\top \in \mathbb{R}^{n_z,n_z}$, $F \in \mathbb{R}^{n_z,n_\vartheta}$, and $c \in \mathbb{R}^{n_z}$ define the cost function, and $G_h \in \mathbb{R}^{n_h,n_z}$, $G_s \in \mathbb{R}^{n_s,n_z}$, $b_h \in \mathbb{R}^{n_h}$, $b_s \in \mathbb{R}^{n_s}$, $S_h \in \mathbb{R}^{n_h,n_\vartheta}$, and $S_s \in \mathbb{R}^{n_s,n_\vartheta}$ define the $n_c = n_h + n_s$ constraints, which are distinguished between constraints that cannot be relaxed (subscript $h$ as hard) and those that can be relaxed (subscript $s$ as soft). If equality constraints are present, they can be converted into double-sided inequalities to fit (1).

Our interest in problem (1) is motivated by Model Predictive Control (MPC) for discrete-time linear systems. In MPC (see, e.g., Mayne, 2014; Morari & Lee, 1999), one typically formulates a finite-horizon optimization problem with a quadratic cost function of the state and of the control input, subject to linear constraints on the input and the state. By unrolling the system dynamics initialized at the current state $\vartheta$, one gets an optimization problem of the form (1), where the decision vector $z$ collects the values of the control input over the entire finite-horizon. Typically, constraints on $z$ that cannot be relaxed encompass actuation, safety, and stability constraints, while constraints on $z$ that can be relaxed concern desired operating regions for the system related to comfort/performance. Moreover, we typically

☆ The material in this paper was not presented at any conference. This paper was recommended for publication in revised form by Associate Editor Martin Monnigmann under the direction of Editor Ian R Petersen.

* Corresponding author.
*E-mail addresses:* alessandro.falsone@polimi.it (A. Falsone), federico.bianchi@rse-web.it (F. Bianchi), maria.prandini@polimi.it (M. Prandini).

https://doi.org/10.1016/j.automatica.2023.111279

have $Q \succ 0$ because the control effort is weighted through a positive definite term. According to the receding horizon strategy, problem (1) is then solved and only the control action corresponding to the current time instant is applied, which results in a new initial state $\vartheta'$ to be used in place of $\vartheta$ in (1) when computing the control input at the next time instant, and so on. This results in a state feedback control law, which can thus counteract modeling errors and disturbances, while accounting for constraints on both input and state.

We impose the following assumption.

**Assumption 1** (*Well-Posedness*). The set $\Theta_f \subseteq \mathbb{R}^{n_\vartheta}$ of $\vartheta$ for which (1) is feasible is full dimensional. Moreover, the constraint set $\mathcal{Z}_\vartheta = \{z \in \mathbb{R}^{n_z} : G_h z \leq b_h + S_h\vartheta\}$ is bounded for some parameter value $\vartheta \in \mathbb{R}^{n_\vartheta}$. □

Note that since $\mathcal{Z}_\vartheta$ is a polyhedron and $\vartheta$ appears on the right hand side, if $\mathcal{Z}_\vartheta$ is bounded for one $\vartheta$, then it is bounded for all $\vartheta \in \mathbb{R}^{n_\vartheta}$. Also, since we are assuming that the feasibility set is full-dimensional, then, Assumption 1 is naturally satisfied in standard MPC problems where the state $\vartheta$ evolves in a full dimensional region and the actuation constraints limiting the control inputs $z$ in $\mathcal{Z}_\vartheta$ are independent of the state $\vartheta$ while the state-dependent hard constraints in $\mathcal{Z}_\vartheta$ are enforcing the state to be in a bounded terminal set.

Under Assumption 1, problem (1) admits an optimal solution given by a PieceWise Affine (PWA) map $z^\star : \Theta_f \to \mathcal{Z}$ defined over $\Theta_f$ and taking values in $\mathcal{Z} = \bigcup_{\vartheta\in\mathbb{R}^{n_\vartheta}} \mathcal{Z}_\vartheta$, i.e.,

$$z^\star(\vartheta) = K_z^i\vartheta + k_z^i, \quad \vartheta \in \Theta_i, \quad i = 1, \ldots, r, \tag{2}$$

where $\{\Theta_i\}_{i=1}^r$ is a polyhedral partition of $\Theta_f$, which is also polyhedral, Jones and Morrari (2006).

In MPC, the optimal map $z^\star(\cdot)$ in (2) provides the optimal control input to be applied as a function of the current state $\vartheta$, thus avoiding online re-computation (explicit MPC, Alessio and Bemporad (2009) and Bemporad, Borrelli, Morari et al. (2002)). This makes MPC viable also for those applications where limited computation power is available and those involving fast dynamics. Indeed, the optimal map is computed offline and, at each time step, one only needs to measure the current state, determine its position within the polyhedral regions (*point location* problem), and evaluate the corresponding affine function. Several approaches have been proposed regarding efficient computation and representation of the optimal map, see, e.g., Baotić, Borrelli, Bemporad, and Morari (2008), Bemporad, Morari, Dua and Pistikopoulos (2002), Bemporad, Oliveri, Poggi, and Storace (2011), Gupta, Bhartiya, and Nataraj (2011), Herceg, Jones, Kvasnica, and Morari (2015), Kvasnica and Fikar (2011), Nguyen, Gulan, Olaru, and Rodriguez-Ayerbe (2017), Oberdieck, Diangelakis, and Pistikopoulos (2017) and Tøndel, Johansen, and Bemporad (2003). Unfortunately, it may happen that the current state $\vartheta$ makes (1) infeasible ($\vartheta \notin \Theta_f$) so that the map $z^\star(\cdot)$ is not defined at that $\vartheta$ and no control action can be computed. To circumvent this issue, two mainstream strategies have been adopted in the literature, i.e., the *minimal time approach* and the *soft-constraint approach*, Scokaert and Rawlings (1999).

In the minimal time approach, the $\vartheta$-dependent smallest time instant after which constraints are satisfied over an infinite time-horizon is identified and constraints are only enforced afterwards, Rawlings and Muske (1993). Whilst the earliest possible constraint satisfaction is enforced, this comes at the price of possible large constraint violations in the transient. Furthermore, the control law is not guaranteed to be continuous in the state, which makes harder proving stability. In the soft-constraint approach, instead, those constraints that depend on $\vartheta$ are softened by removing them and penalizing their violation in the objective function. In the earliest works (Ricker, Subrahmanian, & Sim, 1988) and Zheng and Morari (1995), the peak of the violation is penalized through a quadratic term. In Scokaert and Rawlings (1999), penalization of the violation over the prediction horizon through a quadratic and/or a linear term is suggested to improve performance and ease the tuning of the penalization coefficient. Accordingly, Bemporad, Morari et al. (2002) penalize the squared violation, while Kerrigan and Maciejowski (2000) (the most relevant for us) penalize its 1-norm. Since we want to relax only soft constraints, this approach translates into solving the following problem

$$\begin{aligned} \min_z \quad & \tfrac{1}{2}z^\top Qz + (F\vartheta + c)^\top z + \mu\|[G_s z - b_s - S_s\vartheta]_+\|_1 \\ \text{s.t:} \quad & G_h z \leq b_h + S_h\vartheta, \end{aligned} \tag{3}$$

where $\mu > 0$ is some penalty coefficient.

Under Assumption 1, problem (3) is feasible for all values of $\vartheta \in \Theta$, where $\Theta = \{\vartheta \in \mathbb{R}^{n_\vartheta} : \mathcal{Z}_\vartheta \neq \emptyset\}$ is such that $\Theta_f \subseteq \Theta \subseteq \mathbb{R}^{n_\vartheta}$. If all hard constraints are non-parametric (i.e., $S_h = 0$), then, $\mathcal{Z}_\vartheta = \mathcal{Z}$, $\Theta = \mathbb{R}^{n_\vartheta}$, and $\Theta_f \subseteq \mathbb{R}^{n_\vartheta}$. Moreover, for any given $\vartheta$ for which the original constrained problem is feasible, one can properly set the penalty coefficient $\mu$ so as to compute the solution of the original constrained problem by solving the relaxed one (*exact penalty functions*, see, e.g., Bertsekas (2015, Proposition 1.5.1)). The authors of Kerrigan and Maciejowski (2000) discuss how to extend the guarantees of exact penalty functions to the multi-parametric case, stating that the condition on the penalty coefficient must be satisfied for all values of $\vartheta \in \Theta_f$, and propose a procedure to find a lower bound on the penalty coefficient. However, the proposed method is judged by the same authors to be computationally intensive as it requires to list all possible combination of active constraint. Hints on how to reduce the number of combinations to explore are given, but not thoroughly discussed. Moreover, nothing is said regarding how to set the penalty coefficient and what are the properties of the solution to the relaxed problem for those values of $\vartheta$ which make the constrained problem (1) infeasible (i.e., $\vartheta \notin \Theta_f$).

The main contribution of this paper is then to fully characterize the connection between (1) and its relaxed version (3) as the penalty parameter $\mu > 0$ grows, for all $\vartheta \in \Theta$, including those in the infeasibility set for (1). Note that, under Assumption 1, problem (3) is in fact feasible for all values of $\vartheta \in \Theta$. In particular, we shall show that

- relaxed problem (3) admits an optimal parametric solution $z(\vartheta, \mu)$, which is a PWA function defined over $\Theta \times \mathbb{R}_+$;
- as $\mu \to \infty$, any optimal map $z(\vartheta, \mu)$ approaches (pointwise) a map $\bar{z}(\vartheta)$, which is a PWA function defined over $\Theta$ and is continuous if $z(\vartheta, \mu)$ is continuous;
- for all $\vartheta \in \Theta$, any map $\bar{z}(\vartheta)$ achieves the minimization of the 1-norm of the soft constraint violation as a primary objective and then the minimization of the cost function of (1), as a secondary objective;
- for all $\vartheta \in \Theta_f$, any map $\bar{z}(\vartheta)$ coincides with an optimal map $z^\star(\vartheta)$ of (1).

The remainder of the paper is organized as follows. In Section 2, we present our main results on the properties of the relaxed problem. Section 3 provides an algorithm for its solution. Section 4 discusses the complexity of such solution. A numerical example is given in Section 5, and some concluding remarks are drawn in Section 6.

## 2. Solution of the relaxed problem: properties

First, note that (3) is a convex ($\mu > 0$) but non-linear and non-quadratic optimization problem due to the presence of the $[\,\cdot\,]_+$ operator and the $\|\cdot\|_1$ norm. The following lemma shows that (3) can nonetheless be posed as a mp-QP.

**Lemma 1.** *Fix $\vartheta \in \Theta$, $\mu > 0$, and consider the following optimization problem*

$$\begin{aligned}\min_{z,h}\quad & \tfrac{1}{2}z^\top Qz + (F\vartheta + c)^\top z + \mu\, 1^\top h \\ \text{s.t:}\quad & G_h z \le b_h + S_h\vartheta \\ & G_s z \le b_s + S_s\vartheta + h \\ & h \ge 0,\end{aligned} \tag{4}$$

*where $h \in \mathbb{R}^{n_s}$ is a vector of auxiliary variables. A pair $(z(\vartheta,\mu),\ h(\vartheta,\mu))$ is optimal for (4) if and only if $z(\vartheta,\mu)$ is optimal for (3) and $h(\vartheta,\mu) = \max\{G_s z(\vartheta,\mu) - b_s - S_s\vartheta, 0\}$.*

**Proof.** The term penalizing the violation of the relaxed constraints in the cost function of (3) can be rewritten as

$$\begin{aligned}\|[G_s z - b_s - S_s\vartheta]_+\|_1 &= \|\max\{G_s z - b_s - S_s\vartheta, 0\}\|_1 \\ &= \sum_{i=1}^{n_s} |\max\{G_s^{[i]} z - b_s^{[i]} - S_s^{[i]}\vartheta, 0\}| \\ &= \sum_{i=1}^{n_s} \max\{G_s^{[i]} z - b_s^{[i]} - S_s^{[i]}\vartheta, 0\} \\ &= 1^\top \max\{G_s z - b_s - S_s\vartheta, 0\},\end{aligned} \tag{5}$$

using the fact that $[v]_+ = \max\{v, 0\} \ge 0$ for any $v$ together with the maximum being intended component-wise. Using (5), we can turn (3) into a standard mp-QP via the so-called epigraphic reformulation. To this end, we introduce $n_s$ auxiliary variables $h^{[1]}, \dots, h^{[n_s]}$ and impose the additional constraint $h^{[i]} \ge \max\{G_s^{[i]} z - b_s^{[i]} - S_s^{[i]}\vartheta, 0\}$, for all $i = 1, \dots, n_s$. This leads to the following problem

$$\begin{aligned}\min_{z,h}\quad & \tfrac{1}{2}z^\top Qz + (F\vartheta + c)^\top z + \mu\, 1^\top h \\ \text{s.t:}\quad & G_h z \le b_h + S_h\vartheta \\ & \max\{G_s z - b_s - S_s\vartheta, 0\} \le h,\end{aligned} \tag{6}$$

where $h = [h^{[1]} \cdots h^{[n_s]}]^\top$. Problem (6) is clearly equivalent to (4) since $\max\{G_s z - b_s - S_s\vartheta, 0\} \le h$ if and only if $h \ge G_s z - b_s - S_s\vartheta$ and $h \ge 0$.

To ease the notation, let us drop the dependency of the optimal solutions from $\vartheta$ and $\mu$. For a pair $(\bar z, \bar h)$ to be optimal for (4) (or (6)), we must have that $\bar h = \max\{G_s\bar z - b_s - S_s\vartheta, 0\}$. Indeed, if there exists $i \in \{1, \dots, n_s\}$ such that $\bar h^{[i]} > \max\{G_s^{[i]}\bar z - b_s^{[i]} - S_s^{[i]}\vartheta, 0\}$, we could always select $(\bar z, \bar h')$ such that

$$\max\{G_s^{[i]}\bar z - b_s^{[i]} - S_s^{[i]}\vartheta, 0\} \le \bar h'^{[i]} < \bar h^{[i]}$$

and $\bar h'^{[j]} = \bar h^{[j]}$ for all $j \ne i$ as a new (feasible) solution for (4). Since $1^\top \bar h' < 1^\top \bar h$ and $\mu > 0$, $(\bar z, \bar h')$ would strictly decrease the cost function, thus rendering $(\bar z, \bar h)$ sub-optimal and causing a contradiction.

By definition of minimizer of (4),

$$\begin{aligned}&\tfrac{1}{2}\bar z^\top Q\bar z + (F\vartheta + c)^\top \bar z + \mu\, 1^\top \bar h \\ &\quad \le \tfrac{1}{2}z^\top Qz + (F\vartheta + c)^\top z + \mu\, 1^\top h\end{aligned}$$

for any $(z, h)$ such that $G_h z \le b_h + S_h\vartheta$ and $\max\{G_s z - b_s - S_s\vartheta, 0\} \le h$. Selecting $h = \max\{G_s z - b_s - S_s\vartheta, 0\}$ on the right hand side, using the fact that at optimality we have $\bar h = \max\{G_s\bar z - b_s - S_s\vartheta, 0\}$, and leveraging equivalence (5), yields

$$\begin{aligned}&\tfrac{1}{2}\bar z^\top Q\bar z + (F\vartheta + c)^\top \bar z + \mu\|[G_s\bar z - b_s - S_s\vartheta]_+\|_1 \\ &\quad \le \tfrac{1}{2}z^\top Qz + (F\vartheta + c)^\top z + \mu\|[G_s z - b_s - S_s\vartheta]_+\|_1\end{aligned}$$

for all $z$ such that $G_h z \le b_h + S_h\vartheta$, which means that $\bar z$ is a minimizer for (3), thus concluding the "only if" part.

Conversely, by definition of minimizer of (3) together with equivalence (5),

$$\begin{aligned}&\tfrac{1}{2}\bar z^\top Q\bar z + (F\vartheta + c)^\top \bar z + \mu\, 1^\top \max\{G_s\bar z - b_s - S_s\vartheta, 0\} \\ &\quad \le \tfrac{1}{2}z^\top Qz + (F\vartheta + c)^\top z + \mu\, 1^\top \max\{G_s z - b_s - S_s\vartheta, 0\},\end{aligned}$$

for all $z$ such that $G_h z \le b_h + S_h\vartheta$. Defining $\bar h = \max\{G_s\bar z - b_s - S_s\vartheta, 0\}$ we obtain

$$\begin{aligned}&\tfrac{1}{2}\bar z^\top Q\bar z + (F\vartheta + c)^\top \bar z + \mu\, 1^\top \bar h \\ &\quad \le \tfrac{1}{2}z^\top Qz + (F\vartheta + c)^\top z + \mu\, 1^\top \max\{G_s z - b_s - S_s\vartheta, 0\} \\ &\quad \le \tfrac{1}{2}z^\top Qz + (F\vartheta + c)^\top z + \mu\, 1^\top h,\end{aligned}$$

for all $(z, h)$ such that $G_h z \le b_h + S_h\vartheta$ and $h \ge \max\{G_s z - b_s - S_s\vartheta, 0\}$, or, equivalently, $h \ge G_s z - b_s - S_s\vartheta$ and $h \ge 0$, which means that $(\bar z, \bar h)$ is a minimizer for (4), thus concluding the "if" part and the proof. □

Lemma 1 shows that problems (3) and (4) are equivalent in that, for any $\vartheta \in \Theta$ and $\mu > 0$, they admit the same optimal values for the $z$ variables. Note that the multi-parametric program (4) is again a mp-QP with parameter vector $p = [\vartheta^\top\ \mu]^\top$ as it can be reformulated as

$$\begin{aligned}\min_{z,h}\quad & \tfrac{1}{2}z^\top Qz + (F_0 p + c)^\top z + p^\top I_0 h \\ \text{s.t:}\quad & G_h z \le b_h + S_{h0} p \\ & G_s z \le b_s + S_{s0} p + h \\ & h \ge 0,\end{aligned} \tag{7}$$

with $F_0 = [F\ \ 0_{n_z}]$, $I_0 = [0_{n_s,n_\vartheta}\ \ 1_{n_s}]^\top$, $S_{h0} = [S_h\ \ 0_{n_h}]$, and $S_{s0} = [S_s\ 0_{n_s}]$.

The following corollary is therefore a straightforward application of the results in Jones and Morrari (2006) given Lemma 1 and the equivalence of problems (7) and (4).

**Corollary 1.** *Under Assumption 1, the multi-parametric program (3) admits an optimal solution $z : \Theta \times \mathbb{R}_+ \to \mathcal{Z}$, which is a finite collection of affine maps of $(\vartheta, \mu)$, each map being defined over an element of a finite collection of full-dimensional polyhedra that are not overlapping and such that the union of their closures covers $\Theta \times \mathbb{R}_+$.*

**Proof.** Given the fact that (7) is a mp-QP (which can be posed as a multi-parametric Linear Complementarity Problem, see Jones and Morrari (2006, Section II.C)) we can readily invoke the results from Jones and Morrari (2006) and conclude that, under Assumption 1 (which entails that $\Theta$ is full dimensional since $\Theta_f \subseteq \Theta$), (7) admits an optimal map $(z(p), h(p))$, which is a finite collection of affine maps of $p = [\vartheta^\top\ \mu]^\top$ (cf. Jones and Morrari (2006, Section II)) defined over a finite number of full-dimensional polyhedral regions that are not intersecting and such that the union of their closures covers $\Theta \times \mathbb{R}_+$ (cf. Jones and Morrari (2006, Theorem 2)). Given the equivalence between (4) and (7), and owing to the equivalence of minimizers between (3) and (4) granted by Lemma 1, the result readily follows. □

With a little abuse of notation, we shall call the optimal solution $z : \Theta \times \mathbb{R}_+ \to \mathcal{Z}$ in Corollary 1 a PWA optimal map. Note that if problem (1) is non-degenerate (cf. Bemporad, Morari et al. (2002, Section 4.1.1)), then, the optimal solution of the multi-parametric program (3) is indeed a PWA (continuous) map defined over a finite polyhedral partition of $\Theta \times \mathbb{R}_+$ (see Bemporad, Morari et al. (2002)).

The next result is less obvious and shows that any PWA optimal map $z(\vartheta, \mu)$ of (3) approaches a limit as $\mu$ grows. Note that the proof is constructive and, in practice, obtaining the limiting map does not involve the computation of any limit.

**Theorem 1.** *Under Assumption 1, any PWA optimal map $z : \Theta \times \mathbb{R}_+ \to \mathcal{Z}$ of the multi-parametric problem (3) satisfies*

$$\lim_{\mu \to \infty} z(\vartheta, \mu) = \bar{z}(\vartheta), \tag{8}$$

*where $\bar{z} : \Theta \to \mathcal{Z}$ is a PWA function defined over a polyhedral partition $\{\Theta_i\}_{i=1}^q$ of $\Theta$, which is continuous whenever $z(\vartheta, \mu)$ is continuous.*

*In particular, for each $i = 1, \dots, q$, $\Theta_i$ is related to the ith region $\mathcal{R}_i = \{(\vartheta, \mu) \in \Theta \times \mathbb{R}_+ : F_i\vartheta + f_i\mu \leqslant g_i\}$ of $z(\vartheta, \mu)$ receding along the $\mu$-axis (i.e., with $f_i \leq 0$) and is given by $\Theta_i = \{\vartheta \in \Theta : F_i^0\vartheta \leqslant g_i^0\}$, with $F_i^0$ and $g_i^0$ containing the rows of $F_i$ and $g_i$ corresponding to zero elements of $f_i$. Also, for each $i = 1, \dots, q$,*

$$\bar{z}(\vartheta) = K_z^i\vartheta + k_z^i, \quad \vartheta \in \Theta_i, \tag{9}$$

*is derived from the expression of $z(\vartheta, \mu)$ over $\mathcal{R}_i$, which is independent of $\mu$ and given by $z(\vartheta, \mu) = K_z^i\vartheta + k_z^i$, $(\vartheta, \mu) \in \mathcal{R}_i$.*

**Proof.** The proof is divided into three parts. We first study how the partition induced by any $z(\vartheta, \mu)$ onto the $\vartheta$-domain changes as $\mu \to \infty$ and we compute its limit in terms of polyhedral regions and affine laws. Then we show that the limit regions form a partition of the $\vartheta$-domain and, finally, we prove continuity of the limit map.

Consider any PWA optimal map $z(\vartheta, \mu)$ of (3). Let us start by noticing that $z(\vartheta, \mu)$ is defined over an unbounded domain $\Theta \times \mathbb{R}_+$. Under Assumption 1, by Corollary 1, $z(\vartheta, \mu)$ is defined over a finite number of polyhedral regions partitioning $\Theta \times \mathbb{R}_+$. Therefore, some regions of the partition have to be unbounded.

By Rockafellar (1970, Theorem 8.4), any unbounded region $\mathcal{U}_i$ admits at least one direction of recession, i.e., by Rockafellar (1970, p. 61 and Theorem 8.3), a vector $[d^\top \ 1]^\top$ such that

$$\begin{bmatrix} \vartheta \\ \mu \end{bmatrix} \in \mathcal{U}_i \implies \begin{bmatrix} \vartheta + \alpha d \\ \mu + \alpha \end{bmatrix} \in \mathcal{U}_i \quad \forall \alpha \in \mathbb{R}_+.$$

Moreover, each vector that is a direction of recession for the set $\Theta \times \mathbb{R}_+$ (we are interested in vector $[0_{n_\vartheta}^\top \ 1]^\top$ in particular) must be a direction of recession for at least one region, otherwise we could start at any point and then move along the direction that is not a direction of recession for any region and we would keep crossing different regions. But since there are a finite number of regions, this is not possible.

Now let $\mathcal{R}_i = \{(\vartheta, \mu) : F_i\vartheta + f_i\mu \leqslant g_i\}$ be a polyhedral region with $[0_{n_\vartheta}^\top \ 1]^\top$ as direction of recession. If $(\vartheta, \mu + \alpha) \in \mathcal{R}_i$ for all $\alpha \in \mathbb{R}_+$, then $F_i\vartheta + f_i(\mu + \alpha) \leqslant g_i$ for all $\alpha \in \mathbb{R}_+$, which implies $f_i \leq 0$. We can thus better describe the region $\mathcal{R}_i$ as

$$\mathcal{R}_i = \{(\vartheta, \mu) \in \Theta \times \mathbb{R}_+ : F_i^0\vartheta \leqslant g_i^0 \ \wedge \ F_i^-\vartheta + f_i^-\mu \leqslant g_i^-\}, \tag{10}$$

where $F_i^-, f_i^-$, and $g_i^-$ contain the rows of $F_i, f_i$, and $g_i$ corresponding to negative elements of $f_i$ (possibly none), and $F_i^0$ and $g_i^0$ those corresponding to zero elements of $f_i$ (possibly none).

Let

$$\begin{aligned}\Theta_i(\mu) &= \{\vartheta \in \Theta : (\vartheta, \mu) \in \mathcal{R}_i\} \\ &= \{\vartheta \in \Theta : F_i^0\vartheta \leqslant g_i^0 \wedge F_i^-\vartheta \leqslant g_i^- - f_i^-\mu\}\end{aligned}$$

be a slice of $\mathcal{R}_i$ at a given $\mu \in \mathbb{R}_+$ and

$$\Theta_i = \{\vartheta \in \Theta : F_i^0\vartheta \leqslant g_i^0\}. \tag{11}$$

Since $-f_i^- > 0$, $\Theta_i(\mu) \subseteq \Theta_i(\mu')$ for any $\mu' \geq \mu \geq 0$, meaning that the sequence of sets $\{\Theta_i(\mu)\}_{\mu \in \mathbb{R}_+}$ is non-decreasing as $\mu \to \infty$ and thus has a limit. Moreover, we have that

$$\lim_{\mu \to \infty} \Theta_i(\mu) = \bigcup_{\mu \in \mathbb{R}_+} \Theta_i(\mu) = \Theta_i, \tag{12}$$

because, on one hand, $\bigcup_{\mu \in \mathbb{R}_+} \Theta_i(\mu) \subseteq \Theta_i$ since $\Theta_i(\mu) \subseteq \Theta_i$ for any $\mu \in \mathbb{R}_+$, and, on the other hand, $\Theta_i \subseteq \bigcup_{\mu \in \mathbb{R}_+} \Theta_i(\mu)$ since any $\vartheta \in \Theta_i$ satisfies $\vartheta \in \Theta_i(\mu)$ for $\mu$ sufficiently high and, hence, $\vartheta \in \bigcup_{\mu \in \mathbb{R}_+} \Theta_i(\mu)$.

Consider again the polyhedral region $\mathcal{R}_i$ in (10). By Corollary 1, within $\mathcal{R}_i$ the optimal map $z(\vartheta, \mu)$ is affine in $\vartheta$ and $\mu$, i.e.,

$$z(\vartheta, \mu) = K_z^i\vartheta + M_z^i\mu + k_z^i,$$

for all $(\vartheta, \mu) \in \mathcal{R}_i$, with $K_z^i \in \mathbb{R}^{n_z, n_\vartheta}$, $M_z^i \in \mathbb{R}^{n_z}$, and $k_z^i \in \mathbb{R}^{n_z}$. Let $(\vartheta, \mu) \in \mathcal{R}_i$. Since $\mathcal{R}_i$ recedes along $[0_{n_\vartheta}^\top \ 1]^\top$, then $(\vartheta, \mu + \alpha) \in \mathcal{R}_i$ for all $\alpha \in \mathbb{R}_+$. The optimal map evaluated at $(\vartheta, \mu + \alpha)$ is then

$$z(\vartheta, \mu + \alpha) = K_z^i\vartheta + M_z^i\mu + M_z^i\alpha + k_z^i,$$

for all $\alpha \in \mathbb{R}_+$. However, under Assumption 1, $z(\vartheta, \mu + \alpha)$ must belong to the compact set $\mathcal{Z}_\vartheta$, for all $\alpha \in \mathbb{R}_+$. This is possible only if $M_z^i = 0$, therefore,

$$z(\vartheta, \mu) = \bar{z}(\vartheta) = K_z^i\vartheta + k_z^i, \tag{13}$$

for all $(\vartheta, \mu) \in \mathcal{R}_i$, or, equivalently, for all $(\vartheta, \mu)$ such that $\vartheta \in \Theta_i(\mu)$. Since (13) does not depend on $\mu$, owing to (12), we have

$$\lim_{\mu \to \infty} z(\vartheta, \mu) = \bar{z}(\vartheta) = K_z^i\vartheta + k_z^i, \quad \vartheta \in \Theta_i. \tag{14}$$

The discussion above clearly holds for any unbounded region $\mathcal{R}_i$ receding along $[0_{n_\vartheta}^\top \ 1]^\top$. Let $q$ be the number of unbounded polyhedral regions receding along $[0_{n_\vartheta}^\top \ 1]^\top$, which is finite since, by Corollary 1, the number of regions partitioning $\Theta \times \mathbb{R}_+$ is finite. We next show that $\{\Theta_i\}_{i=1}^q$ is a polyhedral partition of $\Theta$. By definition, each $\Theta_i$ is a non-empty polyhedral subset of $\Theta \subseteq \mathbb{R}^{n_\vartheta}$. Let us now assume, for the sake of contradiction, that the collection $\{\Theta_i\}_{i=1}^q$ does not cover $\Theta$, meaning that there exists a $\bar{\vartheta} \notin \bigcup_{i=1}^q \Theta_i$, or, equivalently, there exists a $\bar{\vartheta}$ such that there is no $\mu \in \mathbb{R}_+$ for which $(\bar{\vartheta}, \mu) \in \bigcup_{i=1}^q \mathcal{R}_i$. Since by Corollary 1 the map $z(\vartheta, \mu)$ is defined over a polyhedral partition of $\Theta \times \mathbb{R}_+$, for any $\mu \in \mathbb{R}_+$, then there must exist a region $\mathcal{N}_t$ (not receding along $[0_{n_\vartheta}^\top \ 1]^\top$) containing $(\bar{\vartheta}, \mu)$. Since $\mathcal{N}_t$ does not recede along $[0_{n_\vartheta}^\top \ 1]^\top$, there exists an $\alpha > 0$ such that $(\bar{\vartheta}, \mu + \alpha) \notin \mathcal{N}_t$ and either

(i) $(\bar{\vartheta}, \mu + \alpha) \in \mathcal{R}_j$ for some region $\mathcal{R}_j$ receding along $[0_{n_\vartheta}^\top \ 1]^\top$, or

(ii) $(\bar{\vartheta}, \mu + \alpha) \in \mathcal{N}_{t'}$ for some region $\mathcal{N}_{t'}$, $t' \neq t$, not receding along $[0_{n_\vartheta}^\top \ 1]^\top$.

In case (i), if $(\bar{\vartheta}, \mu + \alpha) \in \mathcal{R}_j$, then $\bar{\vartheta} \in \Theta_j(\mu + \alpha) \subseteq \Theta_j$, which contradicts the assumption that $\bar{\vartheta} \notin \bigcup_{i=1}^q \Theta_i$. In case (ii), we can select an $\alpha' > \alpha$ such that $(\bar{\vartheta}, \mu + \alpha') \notin \mathcal{N}_{t'} \cup \mathcal{N}_t$ and we are faced with the same alternatives. Since we only have a finite number of regions, we eventually end up in case (i), thus leading to a contradiction and showing that $\bigcup_{i=1}^q \Theta_i = \Theta$.

Now, either $q = 1$ and the collection $\{\Theta_i\}_{i=1}^q$ is trivially a partition of $\Theta$, or $q > 1$. In the latter case, consider two sets $\Theta_i$ and $\Theta_j$ with $i \neq j$.

Their intersection is given by

$$\begin{aligned}\Theta_i \cap \Theta_j &= \{\vartheta \in \Theta : \exists \mu_i \in \mathbb{R}_+ : \vartheta \in \Theta_i(\mu), \forall \mu \geq \mu_i\} \\ &\quad \cap \{\vartheta \in \Theta : \exists \mu_j \in \mathbb{R}_+ : \vartheta \in \Theta_j(\mu), \forall \mu \geq \mu_j\} \\ &= \{\vartheta \in \Theta : \exists \mu \in \mathbb{R}_+ : \vartheta \in \Theta_i(\mu) \cap \Theta_j(\mu)\} \\ &= \{\vartheta \in \Theta : \exists \mu \in \mathbb{R}_+ : (\vartheta, \mu) \in \mathcal{R}_i \cap \mathcal{R}_j\},\end{aligned}$$

where the first equality is due to (12) and the fact that $\Theta_i(\mu)$ and $\Theta_j(\mu)$ are non-decreasing as $\mu \to \infty$, the second is straightforward and the last one is by definition of $\Theta_i(\mu)$. By Corollary 1, the sets $\{\mathcal{R}_i\}_{i=1}^q$ are pairwise disjoint (as they are members of a partition), hence $\Theta_i \cap \Theta_j = \emptyset$ for all $i, j = 1, \ldots, q$ and $i \neq j$, meaning that also the sets $\{\Theta_i\}_{i=1}^q$ are pairwise disjoint, which, together with the fact that they are non-empty and their union covers $\Theta$, make the collection $\{\Theta_i\}_{i=1}^q$ a partition of $\Theta$.

The latter statement together with (11), (12) and (14) proves the result, except for the continuity of the PWA function $\bar{z} : \Theta \to \mathcal{Z}$ defined over the polyhedral partition $\{\Theta_i\}_{i=1}^q$, when the map $z(\vartheta, \mu)$ is continuous.

Now fix an arbitrary $\theta \in \Theta$, denote with $\overline{\mathcal{B}}_\theta$ the intersection between $\Theta$ and a ball $\mathcal{B}_\theta$ centered in $\theta$, and let $\{\Theta_{i_j}\}_{j=1}^{q_\theta}$ with $q_\theta \leq q$ be the collection of all sets in the partition $\{\Theta_i\}_{i=1}^q$ with a nonempty intersection with $\mathcal{B}_\theta$, so that

$$\overline{\mathcal{B}}_\theta = \mathcal{B}_\theta \cap \Theta = \mathcal{B}_\theta \cap \bigcup_{i=1}^q \Theta_i = \bigcup_{i=1}^q \mathcal{B}_\theta \cap \Theta_i = \bigcup_{j=1}^{q_\theta} \mathcal{B}_\theta \cap \Theta_{i_j},$$

where the second equality is due to $\bigcup_{i=1}^q \Theta_i = \Theta$ and the last equality is due to $\mathcal{B}_\theta \cap \Theta_i = \emptyset$ for all $i \notin \{i_1, \ldots, i_{q_\theta}\}$.

Now fix $j$ and recall the definition of $\Theta_{i_j}(\mu)$. Let $\mu_j \in \mathbb{R}_+$ be such that

$$F_{i_j}^- \vartheta - g_{i_j}^- \leqslant -f_{i_j}^- \mu_j \quad \forall \vartheta \in \mathcal{B}_\theta \cap \Theta_{i_j},$$

which always exists since $-f_{i_j}^- > 0$ and the left hand side of the inequality is finite as an effect of $\mathcal{B}_\theta \cap \Theta_{i_j}$ being bounded. Then $\vartheta \in \mathcal{B}_\theta \cap \Theta_{i_j}$ implies $\vartheta \in \Theta_{i_j}(\mu_j)$ and, hence, $\mathcal{B}_\theta \cap \Theta_{i_j} \subseteq \Theta_{i_j}(\mu_j)$, which clearly holds for all $j = 1, \ldots, q_\theta$.

Now let $\bar{\mu} = \max_{j=1,\ldots,q_\theta} \mu_j$. By monotonicity, $\mathcal{B}_\theta \cap \Theta_{i_j} \subseteq \Theta_{i_j}(\mu_j) \subseteq \Theta_{i_j}(\bar{\mu})$ and thus

$$\begin{aligned}\overline{\mathcal{B}}_\theta &= \bigcup_{j=1}^{q_\theta} \mathcal{B}_\theta \cap \Theta_{i_j} \\ &\subseteq \bigcup_{j=1}^{q_\theta} \Theta_{i_j}(\bar{\mu}) \\ &= \bigcup_{j=1}^{q_\theta} \{\vartheta \in \Theta : (\vartheta, \bar{\mu}) \in \mathcal{R}_{i_j}\} \\ &= \{\vartheta \in \Theta : (\vartheta, \bar{\mu}) \in \mathcal{S}\}, \end{aligned} \tag{15}$$

where $\mathcal{R}_{i_j}$, $j = 1, \ldots, q_\theta$, are unbounded regions receding along the positive $\mu$-axis and $\mathcal{S} = \bigcup_{j=1}^{q_\theta} \mathcal{R}_{i_j}$.

According to (13), $\bar{z}(\vartheta) = z(\vartheta, \mu)$ for all $(\vartheta, \mu)$ belonging to $\bigcup_{i=1}^q \mathcal{R}_i$. Since $\mathcal{S} \subseteq \bigcup_{i=1}^q \mathcal{R}_i$, then $\bar{z}(\vartheta) = z(\vartheta, \mu)$ for any $(\vartheta, \mu) \in \mathcal{S}$ and, by (15), $\bar{z}(\vartheta) = z(\vartheta, \bar{\mu})$ for any $\vartheta \in \overline{\mathcal{B}}_\theta$. Therefore, if we assume $z : \Theta \times \mathbb{R}_+ \to \mathcal{Z}$ to be continuous over $\overline{\mathcal{B}}_\theta$, then $\bar{z}(\cdot)$ is continuous over $\overline{\mathcal{B}}_\theta$. Since $\theta$ is arbitrary, this shows that if $z : \Theta \times \mathbb{R}_+ \to \mathcal{Z}$ is continuous, then also $\bar{z} : \Theta \to \mathcal{Z}$ is continuous and concludes the proof. □

Let us highlight once again that the proof of Theorem 1 is constructive, as it shows how to obtain the map $\bar{z}(\vartheta)$ by computing (only) the regions of the optimal map of (3) receding along the $\mu$-axis (i.e., with $f_i \leq 0$), and their associated affine law. Furthermore, in contrast with Kerrigan and Maciejowski (2000), there is no need to tune the penalty parameter $\mu$ appropriately either.

Next, we establish a connection between any limit map $\bar{z}(\vartheta)$ and the original mp-QP in (1).

**Theorem 2.** *Under Assumption 1, for all $\vartheta \in \Theta$, any limit map $\bar{z}(\vartheta)$ obtained via (8) satisfies*

$$\begin{aligned}\bar{z}(\vartheta) \in \arg\min_z \quad & \tfrac{1}{2} z^\top Q z + (F\vartheta + c)^\top z \\ \text{s.t:} \quad & G_h z \leq b_h + S_h \vartheta \\ & \|[G_s z - b_s - S_s \vartheta]_+\|_1 \leq \bar{v}(\vartheta),\end{aligned} \tag{16}$$

*with*

$$\begin{aligned}\bar{v}(\vartheta) = \min_z \quad & \|[G_s z - b_s - S_s \vartheta]_+\|_1 \\ \text{s.t:} \quad & G_h z \leq b_h + S_h \vartheta.\end{aligned} \tag{17}$$

**Proof.** To ease the notation, set:

$$f(z, \vartheta) = \tfrac{1}{2} z^\top Q z + (F\vartheta + c)^\top z,$$
$$v(z, \vartheta) = \|[G_s z - b_s - S_s \vartheta]_+\|_1,$$

and recall that $\mathcal{Z}_\vartheta = \{z \in \mathbb{R}^{n_z} : G_h z \leq b_h + S_h \vartheta\}$. Let $z(\vartheta, \mu)$ be an optimal map of (3), which can be rewritten as

$$z(\vartheta, \mu) \in \arg\min_{z \in \mathcal{Z}_\vartheta} f(z, \vartheta) + \mu\, v(z, \vartheta)$$

based on the introduced shorthand notations. By definition of optimality, we have

$$f(z(\vartheta, \mu), \vartheta) + \mu\, v(z(\vartheta, \mu), \vartheta) \leq f(z, \vartheta) + \mu\, v(z, \vartheta), \tag{18}$$

for all $z \in \mathcal{Z}_\vartheta$, $\vartheta \in \Theta$, and $\mu \in \mathbb{R}_+$. Now fixing $\vartheta \in \Theta$, setting $z = \tilde{z}(\vartheta)$ with

$$\begin{aligned}\tilde{z}(\vartheta) \in \arg\min_{z \in \mathcal{Z}_\vartheta} \quad & f(z, \vartheta) \\ \text{s.t:} \quad & v(z, \vartheta) \leq \bar{v}(\vartheta),\end{aligned} \tag{19}$$

and bringing the $\mu$ terms on the left hand side of (18), yields

$$f(z(\vartheta, \mu), \vartheta) + \mu[v(z(\vartheta, \mu), \vartheta) - v(\tilde{z}(\vartheta), \vartheta)] \leq f(\tilde{z}(\vartheta), \vartheta),$$

or, equivalently,

$$f(z(\vartheta, \mu), \vartheta) + \mu[v(z(\vartheta, \mu), \vartheta) - \bar{v}(\vartheta)] \leq f(\tilde{z}(\vartheta), \vartheta), \tag{20}$$

since $\bar{v}(\vartheta) \leq v(\tilde{z}(\vartheta), \vartheta) \leq \bar{v}(\vartheta)$ by definition of $\bar{v}(\vartheta) = \min_{z \in \mathcal{Z}_\vartheta} v(z, \vartheta)$ in (17) and by the constraint in (19). For any fixed $\vartheta \in \Theta$, $\tilde{z}(\vartheta) \in \mathcal{Z}_\vartheta$ and $\bar{v}(\vartheta)$ exist and are finite under Assumption 1 and, by definition of $f(z, \vartheta)$, also $f(\tilde{z}(\vartheta), \vartheta)$ is finite. Moreover, under Assumption 1, by Theorem 1, $\lim_{\mu \to \infty} z(\vartheta, \mu) = \bar{z}(\vartheta)$, which is finite, and, by continuity of $f(z, \vartheta)$ and $v(z, \vartheta)$, $\lim_{\mu \to \infty} f(z(\vartheta, \mu), \vartheta) = f(\bar{z}(\vartheta), \vartheta)$ and $\lim_{\mu \to \infty} v(z(\vartheta, \mu), \vartheta) = v(\bar{z}(\vartheta), \vartheta)$, both finite. Taking now the limit as $\mu \to \infty$ on both sides of (20) yields

$$f(\bar{z}(\vartheta), \vartheta) + \lim_{\mu \to \infty} \mu[v(z(\vartheta, \mu), \vartheta) - \bar{v}(\vartheta)] \leq f(\tilde{z}(\vartheta), \vartheta). \tag{21}$$

Relation (21), together with the fact that $\mu$ and $[v(z(\vartheta, \mu), \vartheta) - \bar{v}(\vartheta)]$ are both non-negative and that $f(\bar{z}(\vartheta), \vartheta)$ and $f(\tilde{z}(\vartheta), \vartheta)$ are both finite, implies that

$$0 \leq \lim_{\mu \to \infty} \mu[v(z(\vartheta, \mu), \vartheta) - \bar{v}(\vartheta)] < \infty. \tag{22}$$

In turn, (22), implies

$$\bar{v}(\vartheta) = \lim_{\mu \to \infty} v(z(\vartheta, \mu), \vartheta) = v(\bar{z}(\vartheta), \vartheta) \tag{23}$$

and

$$f(\bar{z}(\vartheta), \vartheta) \leq f(\tilde{z}(\vartheta), \vartheta) \leq f(z, \vartheta), \tag{24}$$

for all $z \in \mathcal{Z}_\vartheta$ such that $v(z, \vartheta) \le \bar{v}(\vartheta)$, the second inequality being due to (19). Since $\bar{z}(\vartheta) \in \mathcal{Z}_\vartheta$ by definition and, by (23), satisfies $v(\bar{z}(\vartheta), \vartheta) \le \bar{v}(\vartheta)$, then $\bar{z}(\vartheta)$ is feasible for the problem in (19) and, by (24), it has a cost no-greater than the optimal one, thus showing (16) and concluding the proof. □

The result in Theorem 2 states that, in the limit as $\mu \to \infty$, solving (3) for a given $\vartheta$ is equivalent to first finding all the $z \in \mathcal{Z}_\vartheta$ that minimize the 1-norm of the soft constraint violation and then, among these ones, those that minimize the original cost function. This is both intuitive, as $\mu \to \infty$ increasingly penalizes the constraint violation while keep minimizing the original cost function, and practical, as minimizing the constraint violation first and then the cost is the best one can hope for, if the alternative is not having a solution at all. This observation immediately leads to the following result.

**Corollary 2.** *Under Assumption 1, for any limit map $\bar{z}(\vartheta)$ obtained via (8), there exists an optimal map $z^\star(\vartheta)$ of (1) such that*

$$\bar{z}(\vartheta) = z^\star(\vartheta), \quad \forall \vartheta \in \Theta_f, \tag{25}$$

*where $\Theta_f$ is the set of $\vartheta$ for which (1) is feasible.*

**Proof.** If $\vartheta \in \Theta_f$, then problem (1) is feasible, meaning that there exist $\tilde{z} \in \mathcal{Z}_\vartheta$ such that $G_s\tilde{z} \le b_s + S_s\vartheta$. Therefore

$$\|[G_s\tilde{z} - b_s - S_s\vartheta]_+\|_1 = 1^\top \max\{G_s\tilde{z} - b_s - S_s\vartheta, 0\} = 0,$$

thus $\bar{v}(\vartheta) = 0$, and the problem in (16) reduces to

$$\begin{aligned} \min_z \quad & \tfrac{1}{2} z^\top Q z + (F\vartheta + c)^\top z \\ \text{s.t:} \quad & G_h z \le b_h + S_h\vartheta \\ & 1^\top \max\{G_s z - b_s - S_s\vartheta, 0\} \le 0, \end{aligned}$$

which is equivalent to (1) since $1^\top \max\{G_s z - b_s - S_s\vartheta, 0\} \le 0$ if and only if $G_s z - b_s - S_s\vartheta \le 0$. Under Assumption 1, by Theorem 2, any $\bar{z}(\vartheta)$ obtained via (8) is an optimal map of (16), which was just shown to be equivalent to (1) for any $\vartheta \in \Theta_f$, thus concluding the proof. □

Corollary 2 shows that $\bar{z}(\vartheta)$ is the object to look for, as it is equal to the optimal map $z^\star(\vartheta)$ of (1) for any value of the parameter for which (1) is feasible.

## 3. Solution of the relaxed problem: algorithm

Now that we have established that $\bar{z}(\vartheta) : \Theta \to \mathcal{Z}$ is the map we are interested in, we shift our attention on how to actually find it. As mentioned before Theorem 2, the proof of Theorem 1 is constructive and shows how to build a $\bar{z}(\vartheta)$ starting from any $z(\vartheta, \mu)$ without the need of computing any limit. The high-level steps that we need to follow are provided in Algorithm 1 and described next.

First, in Step 1, we need to compute an optimal map $z(\vartheta, \mu)$ of (3), which can be obtained feeding problem (4) (which, by Lemma 1, has the same optimal $z$-map of (3)) to any available mp-QP solver. The optimal map $z(\vartheta, \mu)$ is stored as a set $\mathcal{P}_r$ of 6-tuples, the $i$th element $(F_i, f_i, g_i, K_z^i, M_z^i, k_z^i) \in \mathcal{P}_r$ encoding that $z(\vartheta, \mu) = K_z^i\vartheta + M_z^i\mu + k_z^i$ for all $(\vartheta, \mu)$ such that $F_i\vartheta + f_i\mu \leqslant g_i$. In Step 2 we initialize the set $\mathcal{P}$ that will similarly encode the optimal map $\bar{z}(\vartheta)$. Then, for each element associated to the optimal map $z(\vartheta, \mu)$ (cf. Step 3) we check if the positive $\mu$-axis is a direction of recession for the corresponding region $\{(\vartheta, \mu) : F_i\vartheta + f_i\mu \leqslant g_i\}$ (cf. Step 4). If this is the case, we first find those constraints defining the current region that do not depend on $\mu$ (cf. Steps 5 and 6) and we then add the 4-tuple $(F_i^0, g_i^0, K_z^i, k_z^i)$ to $\mathcal{P}$ (cf. Step 7) encoding the fact that $\bar{z}(\vartheta) = K_z^i\vartheta + k_z^i$ for all $\vartheta \in \Theta_i = \{\vartheta : F_i^0\vartheta \leqslant g_i^0\}$.

**Algorithm 1** Procedure to compute $\bar{z}(\vartheta)$

```
1: P_r ← optimal map z(ϑ, μ) of (4)
2: P ← ∅
3: for all (F_i, f_i, g_i, K_z^i, M_z^i, k_z^i) ∈ P_r do
4:     if f_i ≤ 0 then
5:         O ← {j : f_i^[j] = 0}
6:         F_i^0 = F_i^O and g_i^0 = g_i^O
7:         P ← P ∪ (F_i^0, g_i^0, K_z^i, k_z^i)
8:     end if
9: end for
```

Algorithm 1 builds $\bar{z}(\vartheta)$ as prescribed by Theorem 1 by sequentially exploring all the regions of the partition induced by any optimal map $z(\vartheta, \mu)$ of (3). Since, by Corollary 1, the number of regions is finite, Algorithm 1 terminates and returns $\bar{z}(\vartheta)$.

Clearly, Algorithm 1 has the downside of requiring the computation of the whole optimal map $z(\vartheta, \mu)$ of (3) and checking if all associated regions recede along the $\mu$-axis. On the positive side, we can rely on any available mp-QP solver to have a solution to (1) also for those $\vartheta \notin \Theta_f$, which is the optimal one among those minimizing the violation (cf. Theorem 2). Algorithm 1 thus inherits the computational complexity of the (user-chosen) solver used to compute $\mathcal{P}_r$ in Step 1.

The investigation of an algorithm that is able to compute $\bar{z}(\vartheta)$ without the need of solving (3), at least not entirely, goes beyond the scope of this paper and is left for future work.

Finally, let us recall that, according to Theorem 1, the continuity property of the $\bar{z}(\vartheta)$ map directly depends on the continuity property of the obtained optimal map $z(\vartheta, \mu)$ of (4). If, for example, the problem is degenerate but a continuous optimal map is known to exist, then one may resort to a method which is able to compute such a continuous optimal map $z(\vartheta, \mu)$ in Step 1.

## 4. Solution of the relaxed problem: complexity

In this section we discuss the complexity of the limit map $\bar{z}(\vartheta)$ and compare it to the complexity of the optimal map $z^\star(\vartheta)$, measured in terms of the number of regions, $q$ and $r$ respectively, needed to define it.

The number of regions of the solution of an mp-QP is given by the number of optimal combinations of active constraints, which is known to depend on the number of optimization variables, the number of constraints, and the dimension of the parameter space, Bemporad, Morari et al. (2002). The number of optimal combinations of active constraints is clearly bounded by the number of (not necessarily optimal) combinations of active constraints, which is therefore used to characterize the complexity of the solution of the mp-QP, Pistikopoulos, Diangelakis, and Oberdieck (2020, p. 51).

Let us focus on problem (1) first. If we assume non-degeneracy, since the problem is convex, the number of active constraints is no-larger than the number of decision variables. Assuming $n_z \le n_c$, the number of regions $r$ of $z^\star(\vartheta)$ satisfies

$$r \le \sum_{i=0}^{n_z} \binom{n_c}{i} = \sum_{i=0}^{n_z} \binom{n_h + n_s}{i}. \tag{26}$$

As for the number $q$ of regions of $\bar{z}(\vartheta)$, according to Theorem 1, it is no-greater than the number of unbounded regions of $z(\vartheta, \mu)$.

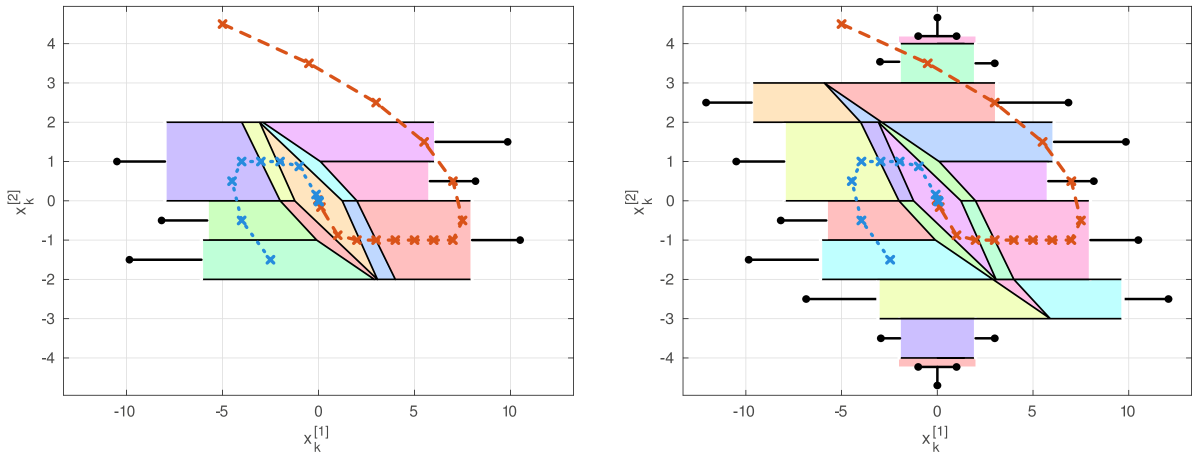


**Fig. 1.** Partition induced by the optimal solution $z^\star$ of (1) (left) and by $\bar{z}$ in (9) (right), whiskers indicating directions in which each region (colored areas, different colors denote different regions) extends to infinity. Lines with crosses denote a 20-step-long trajectory of system (28) initialized at $x_0 = [-2.5\ \ -1.5]^\top$ (dotted blue) or $x_0 = [-5\ 4.5]^\top$ (dashed red) and with $u_k(x_k) = \bar{z}^{[1]}(x_k)$ computed either from (16) (left) or by Algorithm 1 (right), $k = 0, \ldots, 19$.

Let us then consider problem (4) and find an upper bound for the number $q'$ of regions of $z(\vartheta, \mu)$.

Problem (4) is also a convex problem and, under non-degeneracy, it has a number of active constraints, out of its $n_c + n_s$ constraints, that is at most equal to the number of decision variables $n_z + n_s$. Therefore, under the assumption $n_z \leq n_c$, the number of regions $q'$ of $z(\vartheta, \mu)$ satisfies

$$q' \leq \sum_{i=0}^{n_z+n_s} \binom{n_h + 2n_s}{i}.$$

This estimate can, however, be reduced as follows. By Lemma 1, for any $\vartheta \in \Theta$ and $\mu > 0$, a pair $(z(\vartheta, \mu), h(\vartheta, \mu))$ is optimal for (4) if and only if $z(\vartheta, \mu)$ is optimal for (3) and $h(\vartheta, \mu) = \max\{G_s z(\vartheta, \mu) - b_s - S_s\vartheta, 0\}$. Since (3) is a convex problem, under the non-degeneracy assumption, it has at most $n_z$ active constraints and, therefore, at most $n_z$ constraints among $G_h z \leq b_h + S_h\vartheta$ can be active at $z = z(\vartheta, \mu)$. Moreover, $h(\vartheta, \mu) = \max\{G_s z(\vartheta, \mu) - b_s - S_s\vartheta, 0\}$ and therefore at least $n_s$ of the remaining $2n_s$ constraints $G_s z \leq b_s + S_s\vartheta$ and $h \geq 0$ must be active at $z = z(\vartheta, \mu)$ and $h = h(\vartheta, \mu)$. We thus need to consider all the possible choices of: i) exactly $n_s$ active constraints from $G_s z \leq b_s + S_s\vartheta$ and $h \geq 0$ and ii) at most $n_z$ active constraints from the $n_h + n_s$ constraints among $G_h z \leq b_h + S_h\vartheta$ and in those among $G_s z \leq b_s + S_s\vartheta$ and $h \geq 0$ that are remaining after having chosen the $n_s$ ones in i), i.e.,

$$q' \leq \binom{2n_s}{n_s} \sum_{i=0}^{n_z} \binom{n_h + n_s}{i}. \tag{27}$$

Since $q \leq q'$, by comparing (26) with (27) we have that, in the worst case, $q$ and $r$ scale in the same way except for an additional factor $\binom{2n_s}{n_s}$ for the number of regions in $\bar{z}(\vartheta)$, which depends only on the number of relaxed constraints $n_s$. Finally, note that, by Corollary 2, any region of $z^\star(\vartheta)$ is also a region of $\bar{z}(\vartheta)$, so that we expect $q \geq r$.

## 5. Numerical example

In this section we validate the theoretical findings developed in Section 2 and we test Algorithm 1 on a slightly modified version of the double integrator example presented in Bemporad, Morari et al. (2002).

Let $k$ be a discrete-time index and let us consider the dynamical system

$$x_{k+1} = \underbrace{\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}}_{A} x_k + \underbrace{\begin{bmatrix} 0 \\ 1 \end{bmatrix}}_{B} u_k, \tag{28}$$

with state $x_k \in \mathbb{R}^2$ and input $u_k \in \mathbb{R}$. The control objective is to regulate the system to the origin of the state space subject to constraints on the actuation capability $u_k \in [-1, 1]$. Differently from Bemporad, Morari et al. (2002), we impose the additional requirement that the second state should be within the interval $[-1, 1]$. To this end we formulate the following finite-horizon optimal control problem parametric in $x_k$

$$\begin{aligned} \min_{\{u_t, x_t\}_t} \quad & \sum_{t=k}^{k+N-1} x_{t+1}^\top H x_{t+1} + u_t^\top R u_t && \qquad (29) \\ \text{s.t:} \quad & -1 \leq u_t \leq 1 && t = k, \ldots, k+N-1 \\ & x_{t+1} = Ax_t + Bu_t && t = k, \ldots, k+N-1 \\ & -1 \leq x_t^{[2]} \leq 1 && t = k+1, \ldots, k+N, \end{aligned}$$

with $H = \text{diag}([1\ 0])$ and $R = 0.1$ (set according to Bemporad, Morari et al., 2002), and apply the returned optimal policy $\{u_t^\star\}_t$ in a receding-horizon fashion, i.e., at a generic instant $k$ we solve (29) with the current $x_k$, we apply $u_k^\star$ to the system, we move to $k + 1$ and we repeat the process.

The problem can easily fit the structure of (1) using the dynamic equation (28) to express each $x_t$ as a function of $x_k$ and $u_t$, $t = k, \ldots, t-1$, for each $t = k, \ldots, k+N$. In particular $z = [u_k\ \cdots\ u_{k+N-1}]^\top \in \mathbb{R}^N$, $\vartheta = x_k \in \mathbb{R}^2$, $G_h$ and $b_h$ model the actuation constraints $u_t \in [-1, 1]$, $t = k, \ldots, k+N-1$, with $S_h = 0$ as they depend on $z$ only, while $G_s$, $b_s$, and $S_s$ model the state constraints $x_t^{[2]} \in [-1, 1]$, $t = k+1, \ldots, k+N$, which are the ones we are willing to relax. From now on, when we refer to problems introduced in the previous sections we mean instances of those problems related to the introduced example and (29).

If we solve (1) as a mp-QP (using the MPT3 Toolbox Herceg, Kvasnica, Jones, & Morari, 2013), we obtain a map $u_k^\star(x_k) =$

$z^{\star[1]}(x_k)$ which we can apply at each $k$ without the need of solving (29) at every step. We report in Fig. 1 (left) the partition of the state space induced by the map $z^\star$ when $N = 3$. It is easy to see that problem (29) (and thus (1)) is feasible only when $x_k^{[2]} \in [-2, 2]$ because only in such a case there exists a $u_k \in [-1, 1]$ able to bring $x_k^{[2]} + u_k = x_{k+1}^{[2]} \in [-1, 1]$. Accordingly, the partition in Fig. 1 (left) covers the band $x_k^{[2]} \in [-2, 2]$ only.

If we now run Algorithm 1 (Step 1 was performed again using the MPT3 Toolbox Herceg et al., 2013), we can compute the map $\bar{z}(\vartheta)$, which is guaranteed by Theorem 1 to be the limit of the optimal map $z(\vartheta, \mu)$ of (3) as $\mu \to \infty$. We report in Fig. 1 (right) the partition of the state space induced by $\bar{z}$ for comparison. As it can be seen from the picture, the induced partition is indeed polyhedral and is defined over the whole state space $\Theta = \mathbb{R}^2$, since $S_h = 0$. Moreover, by comparing the two partitions, it can be easily noticed that the regions inside the $x_k^{[2]} \in [-2, 2]$ band have the same shape, as a first indication about the validity of Theorem 2.

As a further test to validate our theoretical findings we initialize the state of the system at $x_0 = [-2.5 \;\; -1.5]^\top$ and then compute two trajectories by repeatedly applying either $u_k(x_k) = z^{\star[1]}(x_k)$ (which is defined, since $x_0$ is feasible for (1) and $x_k$ stays within the feasible region) or $u_k(x_k) = \bar{z}^{[1]}(x_k)$, with $\bar{z}$ provided by Algorithm 1, for $k = 0, \dots, 19$. The resulting trajectories are reported in Fig. 1 (left) and (right), respectively, as blue dotted lines. As it can be seen from the pictures, the two trajectories are equal, thus supporting the claim of Corollary 2, and thus the claim of Theorem 2 within the feasible region of (1).

As a final test to validate the statement of Theorem 2, we initialize the state of the system at $x_0 = [-5 \;\; 4.5]^\top$. Note that, in this case, $x_0$ makes (1) infeasible, thus $z^\star(x_k)$ is undefined. We then compute a first trajectory applying the control action $u_k(x_k)$ which minimizes the 1-norm of the constraint violation outside the feasible region and $u_k(x_k) = z^{\star[1]}(x_k)$ inside the feasible region (as prescribed by (16)) and a second trajectory applying again the control action $u_k(x_k) = \bar{z}^{[1]}(x_k)$, with $\bar{z}$ returned by Algorithm 1, both for $k = 0, \dots, 19$. The resulting trajectories are reported in Fig. 1 (left) and (right), respectively, as red dashed lines. As it can be seen from the pictures, the two trajectories are equal, thus supporting the claim of Theorem 2 that Algorithm 1 indeed returns a map $\bar{z}$ satisfying (16) with $\bar{v}(k)$, computed as in (17), positive only for $k = 0, 1, 2$ and zero for $k \geq 3$, meaning that for $k \geq 3$ we indeed recover the optimal policy $z^{\star[1]}(x_k)$.

## 6. Conclusions

Motivated by application to MPC, we addressed the infeasibility issue arising in multi-parametric quadratic programming. We showed that the soft-constrained approach is effective in providing a solution that is optimal for the original constrained optimization program in the region of feasibility and minimizes constraint violation in the region of infeasibility. Our current research effort is devoted to the investigation of a computationally efficient algorithm to compute such a solution and to extend the adopted exact penalty approach to the case of prioritized constraints.

## References


Alessio, Alessandro, & Bemporad, Alberto (2009). A survey on explicit model predictive control. In *Nonlinear model predictive control: Towards new challenging applications* (pp. 345–369). Springer.

Baotić, Mato, Borrelli, Francesco, Bemporad, Alberto, & Morari, Manfred (2008). Efficient on-line computation of constrained optimal control. *SIAM Journal on Control and Optimization*, *47*(5), 2470–2489.

Bemporad, Alberto, Borrelli, Francesco, Morari, Manfred, et al. (2002). Model predictive control based on linear programming~ the explicit solution. *IEEE Transactions on Automatic Control*, *47*(12), 1974–1985.

Bemporad, A., Morari, M., Dua, V., & Pistikopoulos, E. N. (2002). The explicit linear quadratic regulator for constrained systems. *Automatica*, *38*(1), 3–20.

Bemporad, Alberto, Oliveri, Alberto, Poggi, Tomaso, & Storace, Marco (2011). Ultra-fast stabilizing model predictive control via canonical piecewise affine approximations. *IEEE Transactions on Automatic Control*, *56*(12), 2883–2897.

Bertsekas, Dimitri P. (2015). *Convex optimization algorithms*. Athena Scientific.

Gupta, Arun, Bhartiya, Sharad, & Nataraj, P. S. V. (2011). A novel approach to multiparametric quadratic programming. *Automatica*, *47*(9), 2112–2117.

Herceg, Martin, Jones, Colin N., Kvasnica, Michal, & Morari, Manfred (2015). Enumeration-based approach to solving parametric linear complementarity problems. *Automatica*, *62*, 243–248.

Herceg, M., Kvasnica, M., Jones, C. N., & Morari, M. (2013). Multi-parametric toolbox 3.0. In *Proceedings of the 12th European Control Conference* ECC, (pp. 502–510).

Jones, Colin N., & Morrari, Manfred (2006). Multiparametric linear complementarity problems. In *Proceedings of the 45th IEEE Conference on Decision and Control* (pp. 5687–5692).

Kerrigan, Eric C., & Maciejowski, Jan M. (2000). Soft constraints and exact penalty functions in model predictive control. In *2000 UKACC International Conference on Control, Cambridge* (pp. 2319–2327).

Kvasnica, Michal, & Fikar, Miroslav (2011). Clipping-based complexity reduction in explicit MPC. *IEEE Transactions on Automatic Control*, *57*(7), 1878–1883.

Mayne, David Q. (2014). Model predictive control: Recent developments and future promise. *Automatica*, *50*(12), 2967–2986.

Morari, Manfred, & Lee, Jay H. (1999). Model predictive control: past, present and future. *Computers & Chemical Engineering*, *23*(4), 667–682.

Nguyen, Ngoc Anh, Gulan, Martin, Olaru, Sorin, & Rodriguez-Ayerbe, Pedro (2017). Convex lifting: Theory and control applications. *IEEE Transactions on Automatic Control*, *63*(5), 1243–1258.

Oberdieck, Richard, Diangelakis, Nikolaos A., & Pistikopoulos, Efstratios N. (2017). Explicit model predictive control: A connected-graph approach. *Automatica*, *76*, 103–112.

Pistikopoulos, E. N., Diangelakis, N. A., & Oberdieck, R. (2020). *Wiley series in operations research and management science, Multi-parametric optimization and control*. Wiley.

Rawlings, J. B., & Muske, K. R. (1993). The stability of constrained receding horizon control. *IEEE Transactions on Automatic Control*, *38*(10), 1512–1516.

Ricker, N. L., Subrahmanian, T., & Sim, T. (1988). Case studies of model-predictive control in pulp and paper production. *IFAC Proceedings Volumes*, *21*(4), 13–22, IFAC Workshop on Model Based Process Control, Atlanta, GA, USA, 13-14 June.

Rockafellar, R. Tyrrell (1970). *Convex analysis*. Princeton University Press.

Scokaert, Pierre O. M., & Rawlings, James B. (1999). Feasibility issues in linear model predictive control. *AIChE Journal*, *45*(8), 1649–1659.

Tøndel, Petter, Johansen, Tor Arne, & Bemporad, Alberto (2003). An algorithm for multi-parametric quadratic programming and explicit MPC solutions. *Automatica*, *39*(3), 489–497.

Zheng, Alex, & Morari, Manfred (1995). Stability of model predictive control with mixed constraints. *IEEE Transactions on Automatic Control*, *40*(10), 1818–1823.


**Alessandro Falsone** received his Master degree in Automation and Control Engineering in 2013, and his Ph.D. degree in Information Technology, in 2018, both with laude and from Politecnico di Milano. During his Ph.D. studies he spent three months as a visiting researcher at the University of Oxford. Since 2018, he is an assistant professor at the Dipartimento di Elettronica, Informazione e Bioingegneria at Politecnico di Milano. His current research interests include distributed optimization and control, data-based optimal control of stochastic systems, model predictive control, and nonlinear model identification. In 2018 he was the recipient of the Dimitris N. Chorafas Prize for his Ph.D. thesis. In 2019 he received the IEEE CSS Italy Chapter Best Young Author Journal Paper Award.

**Federico Bianchi** received the M.Sc. degree (2015, cum laude) in Computer Science and Engineering, and Ph.D. (February 2020) in Information Technology with the thesis "A randomized model structure selector for complex dynamical systems", both at the Politecnico di Milano, Department of Electronics, Information and Bioengineering. His main research interests include nonlinear and hybrid system identification, modeling and control for energy storage systems, and vehicle-to-grid applications. After two years as PostDoc temporary researcher in the Systems and Control group at Politecnico di Milano, DEIB, he is currently a tenured researcher at Ricerca sul Sistema Energetico (RSE) S.p.A., Milan, Italy, Department of Generation Technologies and Materials.

**Maria Prandini** received the Ph.D. degree in Information Technology from the University of Brescia, Italy, in 1998. She was a postdoctoral researcher at the University of California at Berkeley from 1998 to 2000. She also held visiting positions at Delft University of Technology (1998), Cambridge University (2000), UC Berkeley (2005), ETH Zurich (2006), and University of Oxford (2022). In 2002, she joined Politecnico di Milano, where she is Full Professor since 2018. She has been Chair of the Automation and Control Engineering Program for two consecutive triennia starting from 2019. She was elected Fellow of the IEEE in 2020 and IEEE Control Systems Society Distinguished Member in 2018. In 2017, she was August-Wilhelm Scheer Visiting Professor and Honorary fellow of the TUM Institute for Advanced Studied. She was nominated Visiting Professor in Engineering at the University of Oxford for the triennium 2022–2025. She has been active in the IEEE Control Systems Society (CSS), the International Federation of Automatic Control (IFAC), and the Association for Computing Machinery (ACM), contributing to their activities in different roles. She is IFAC President-elect for the triennium 2023–26. Previously, she was Vice-President for conference activities for IFAC (2020–23) and IEEE CSS (2016 and 2017), and a member of SIGBED Board of Directors (2019–21). Her research interests include stochastic hybrid systems, randomized algorithms, distributed and data-driven optimization, multi-agent systems, and the application of control theory to transportation and energy systems.