Yokohama Publishers
ISSN 1349-8169 ONLINE JOURNAL

Pacific Journal of Optimization
Vol. 19, No. 4, pp. 645–674, 2023

ON LAGRANGIAN DUALITY FOR MULTIPARAMETRIC PROGRAMS*

J. Leverenz, H. Lee, M.M. Wiecek†

Abstract: Multiparametric programming examines the relationship between the data and solutions of optimization problems. Current solution methods in the field focus on the primal problem, lack options for solving the dual problem, and are limited in their application to nondifferentiable problems. These shortcomings are addressed by advancing a duality theory for convex multiparametric nonlinear programs. Staying within a framework of Lagrangian duality, weak and strong duality theorems are proved under suitable assumptions. The obtained results motivate the use of a subgradient algorithm proposed for optimization in a function space to solve primal and dual multiparametric problems. The algorithm is implemented in a multiparametric setting and its effectiveness is demonstrated on illustrative example problems.

Key words: *convex programming, subgradient optimization, nonlinear programming, nondifferentiable optimization, strong duality, weak duality*

Mathematics Subject Classification: *90C31, 90C46*

1 Introduction

The area of multiparametric programming (MPP) has received increased attention in recent years as a means of examining the relationship between the data and solutions of optimization problems. In MPP there are two types of unknowns: the parameters that determine problem data (e.g., coefficients, exponents, righthand side values, etc.) and the optimization variables that determine solutions [34, 29]. A solution to an MPP problem consists of mappings from the parameter space to the space of optimal variable values and the space of optimal objective function values.

While nonparametric problems produce a static solution, in MPP the optimal solution and objective values are expressed as functions of parameters. Solving an MPP problem requires three pieces of information [12]: (i) an exact or approximate expression for the optimal objective value as a function of the parameters; (ii) an exact or approximate expression for the optimal values of the optimization variables as functions of the parameters; (iii) the partition of the parameter space into critical regions for which the expressions in (i) and (ii) are valid. The virtue of a parametric solution arises from mapping a full range of solutions for sets of potential problem data represented by the parameters without the need

*This research has been supported by the National Science Foundation through grant number CMMI-1129969 and by the Office of Naval Research through grant number N00014-16-1-2725.

†Corresponding author

to repeatedly solve the same problem using every parameter combination. This approach differs from sensitivity analysis which is focused on posterior evaluation of a local region around a particular, static solution [9, 17]. By contrast, MPP methods produce a complete set of optimal solutions for the entire parameter space [38, 20].

MPP is used in model predictive control to construct control rules for automated systems [37, 22, 28, 39], is introduced to multidisciplinary design optimization (MDO) [29] to offer benefits in decomposition-based design, and is applied to solution methods for multiobjective programming (ϵ-constraint, weighted sum) [15, 19, 35], even in the presence of uncertainty [30, 11]. MPP is also used to solve multilevel hierarchical and decentralized problems [16, 27] and the multiparametric linear complementarity problem [2, 1].

Exact solutions can be obtained using the Karush-Kuhn-Tucker conditions when an MPP problem is linear or quadratic [38, 20, 2, 1]. Otherwise the solution must be approximated. Methods for constructing such an approximate solution may be broadly categorized in three areas: path following or homotopy methods [18, 21], space partition methods [7, 33], and problem approximation methods [14, 26, 13]. A survey of techniques in the latter two areas can be found in [12]. The fundamentals, recent algorithmic developments and applications are collected in [36].

The methods for solving an MPP problem focus on the primal problem: dual problems do not appear except in some cases to provide parametric bounds on solutions [17, 7] or in a conjugate duality framework [10]. Duality, however, plays an important role in optimization and exploring ways to solve the dual of an MPP problem opens up new possibilities to the field.

The goal of this paper is to advance duality theory and methodology for multiparametric nonlinear programs (mp-NLPs). Staying within a framework of Lagrangian duality and making suitable assumptions, weak and strong duality theorems are proved for mp-NLPs. The obtained results motivate the use of a subgradient algorithm, which had been designed for optimization in a function space, to solve primal and dual multiparametric programs. The algorithm is implemented in a multiparametric setting and its effectiveness is demonstrated on examples.

The outline of the paper is as follows. In Section 2, a standard mp-NLP formulation is stated and then changed to an equivalent one which serves as the primal mp-NLP. For this equivalent formulation a dual problem is defined and a weak duality theorem proved. Auxiliary formulations, which are used to support the eventual proof of strong duality, are stated in Section 3 and strong duality for these formulations is shown to hold. Section 4 builds on Sections 2 and 3 to examine the relationship between the primal and dual multiparametric programs and the auxiliary problems, leading to a strong duality theorem for the primal mp-NLP. Since subgradient methods are often paired with dual problems, a subgradient algorithm is identified in the literature and adapted in Section 5 for solving the primal and dual multiparametric programs. In Section 6, the algorithm is applied to two example problems and its performance is analyzed. The paper is concluded in Section 7.

2 Problem Formulation and Basic Concepts

This section consists of three parts. In the first one, an mp-NLP is formulated in the standard way with vectors as optimization variables and available results on the existence and properties of the optimal solutions are quoted. Since optimal solutions to MPP come in the form of functions of the parameters, in the second part the standard formulation is modified into an alternate mp-NLP with functions as optimization variables. In the third part, a Lagrangian dual problem is formulated and weak duality is proved.

Throughout this paper the following notation is used. Every mapping is denoted by a single letter without an input argument, e.g., $\mathbf{x}$, while the mapping with an argument, e.g., $\mathbf{x}(\mathbf{t})$, is used to represent the mapped image of $\mathbf{t}$ under $\mathbf{x}$.

$\boxed{2.1}$ The standard formulation and concept of optimality

The standard formulation of mp-NLP relies on the following basic assumptions.

Assumption 2.1. (i) Let $\mathbb{R}^n$ and $\mathbb{R}^p$ be Euclidean spaces and $\Omega \subset \mathbb{R}^p$ be a nonempty and compact subspace of $\mathbb{R}^p$. Ω is referred to as a parameter space.

(ii) Let $\mathcal{X} : \mathbb{R}^p \to \mathbb{R}^n$ be a continuous point-to-set mapping such that for all $\mathbf{t} \in \Omega$, $\mathcal{X}(\mathbf{t})$ is a nonempty subset of $\mathbb{R}^n$.

(iii) Let $f : \mathbb{R}^n \times \Omega \to \mathbb{R}$ be an objective function that is continuous in the arguments $\boldsymbol{\psi} \in \mathbb{R}^n$ and $\mathbf{t} \in \Omega$.

The mp-NLP problem is typically formulated as:

$$\begin{aligned} f^*(\mathbf{t}) = \min_{\boldsymbol{\psi}} \quad & f(\boldsymbol{\psi}; \mathbf{t}) \\ s.t. \quad & \boldsymbol{\psi} \in \mathcal{X}(\mathbf{t}) \\ & \mathbf{t} \in \Omega \end{aligned} \tag{2.1}$$

Definition 2.1. Let $\mathbf{t} \in \Omega$ be fixed, i.e., $\mathbf{t} = \bar{\mathbf{t}}$. A feasible vector $\bar{\boldsymbol{\psi}} \in \mathcal{X}(\bar{\mathbf{t}})$ is an optimal solution to mp-NLP (2.1) provided $f(\bar{\boldsymbol{\psi}}; \bar{\mathbf{t}}) \leq f(\boldsymbol{\psi}; \bar{\mathbf{t}})$ for every $\boldsymbol{\psi} \in \mathcal{X}(\bar{\mathbf{t}})$.

The concept of optimality for mp-NLP (2.1) for all $\mathbf{t} \in \Omega$ requires other notions. In mp-NLP (2.1), each element of the parameter space is mapped into an optimal solution set and an optimal objective value, and the result of this mapping is an optimal solution mapping $\mathcal{X}^*$ and an optimal value function f^*. The following properties hold for (2.1) [8, 17].

Theorem 2.2. *Let Assumption 2.1 hold and $\mathcal{X}^*$ be the optimal solution mapping for* (2.1) *whose image is defined as $\mathcal{X}^*(\mathbf{t}) = \{\boldsymbol{\psi} \in \mathcal{X}(\mathbf{t}) : f(\boldsymbol{\psi}; \mathbf{t}) = f^*(\mathbf{t})\}$. Then the optimal value function f^* is continuous on Ω and $\mathcal{X}^*$ is an upper semicontinuous mapping of Ω into $\mathbb{R}^n$.*

In general, the literature on mp-NLPs follows a theoretical line of investigation [8, 17, 4] and a methodological one that was recently summarized in [36]. In this paper these two aspects are combined: some advances in the theory of mp-NLPs lead to an algorithm to solve parametric dual problems. As explained in [4], to construct solution algorithms for mp-NLP (2.1), the existence of a selection function for the optimal solution mapping is needed.

Definition 2.3. A function $\mathbf{x} : \Omega \to \mathbb{R}^n$ is said to be a selection for $\mathcal{X}^*$ if $\mathbf{x}(\mathbf{t}) \in \mathcal{X}^*(\mathbf{t})\ \forall \mathbf{t} \in \Omega$.

The existence of a continuous selection function is stated in Theorem 2.4 and is ensured under additional assumptions on the optimal solution mapping [4].

Assumption 2.2. (i) $\mathcal{X}^*(\mathbf{t})$ is a nonempty, compact, and convex set $\forall \mathbf{t} \in \Omega$.

(ii) For every $(\boldsymbol{\psi}^0; \mathbf{t}^0) \in \mathcal{X}^*(\mathbf{t}^0) \times \Omega$ there exists a function $\mathbf{x} : \Omega \to \mathbb{R}^n$ that is continuous at $\mathbf{t}^0$ and such that $\mathbf{x}(\mathbf{t}) \in \mathcal{X}^*(\mathbf{t})\ \forall \mathbf{t} \in \Omega$, and $\mathbf{x}(\mathbf{t}^0) = \boldsymbol{\psi}^0$.

Theorem 2.4. *Let $\mathbb{C}(\Omega, \mathbb{R}^n)$ denote the set of continuous functions $\mathbf{x} : \Omega \to \mathbb{R}^n$ and let Assumptions 2.1 (i) (ii) and 2.2 hold. Then there exists a function $\mathbf{x} \in \mathbb{C}(\Omega, \mathbb{R}^n)$ that is a selection for $\mathcal{X}^*$.*

$\boxed{2.2}$ Reformulation

Since the focus of this study is on constrained problems, a particular structure of the feasible set in (2.1) is assumed. Let

$$\mathcal{X}(\mathbf{t}) = \{\boldsymbol{\psi} \in \mathcal{S} : \mathbf{g}(\boldsymbol{\psi}; \mathbf{t}) \leq \mathbf{0}, \mathbf{h}(\boldsymbol{\psi}; \mathbf{t}) = \mathbf{0}\} \quad \text{for} \quad \mathbf{t} \in \Omega \tag{2.2}$$

where $\mathcal{S} \subset \mathbb{R}^n$, $\mathbf{g} : \mathbb{R}^n \times \Omega \to \mathbb{R}^m$ is the inequality constraint function, and $\mathbf{h} : \mathbb{R}^n \times \Omega \to \mathbb{R}^l$ is the equality constraint function. The following convexity assumptions are needed to develop duality results and a solution algorithm.

Assumption 2.3. Let mp-NLP (2.1)-(2.2) satisfy the following assumptions.

(i) Ω is a convex subspace of $\mathbb{R}^p$.

(ii) $\mathcal{S}$ is a nonempty, compact and convex set in $\mathbb{R}^n$.

(iii) The function $\mathbf{g}$ is continuous and convex in $\boldsymbol{\psi}$ and $\mathbf{t}$ and bounded over its domain, and the function $\mathbf{h}$ is affine in $\boldsymbol{\psi}$ and $\mathbf{t}$.

(iv) The function f is convex in the arguments $\boldsymbol{\psi} \in \mathbb{R}^n$ and $\mathbf{t} \in \Omega$, and bounded over its domain, i.e., $|f(\boldsymbol{\psi}; \mathbf{t})| < \infty \ \ \forall \boldsymbol{\psi} \in \mathbb{R}^n$ and $\mathbf{t} \in \Omega$.

(v) $\mathcal{X}(\mathbf{t})$ is a nonempty and compact set $\ \forall \mathbf{t} \in \Omega$.

By Theorem 2.4, a function in $\mathbb{C}(\Omega, \mathbb{R}^n)$ is an optimal solution to mp-NLP (2.1)-(2.2). Since the standard formulation neither recognizes nor gives means to directly define or find such a function, (2.1)-(2.2) is reformulated. In the new formulation, the feasible set is assumed to be a set of functions that are candidates for an optimal solution. To accomplish this reformulation and make it useful for the theory that is presented in the subsequent sections, Assumption 2.2 (ii) is modified and extended for every $(\boldsymbol{\psi}^0; \mathbf{t}^0) \in \mathcal{X}(\mathbf{t}^0) \times \Omega$ and the feasible functions are assumed to be piecewise continuous and bounded.

Assumption 2.4. For every $(\boldsymbol{\psi}^0; \mathbf{t}^0) \in \mathcal{X}(\mathbf{t}^0) \times \Omega$ there exist a neighborhood of $\mathbf{t}^0$, $N_\delta(\mathbf{t}^0) \subset \Omega$, and a continuous and bounded function $\mathbf{x} : N_\delta(\mathbf{t}^0) \to \mathbb{R}^n$ such that $\mathbf{x}(\mathbf{t}) \in \mathcal{X}(\mathbf{t}) \ \forall \mathbf{t} \in N_\delta(\mathbf{t}^0)$ and $\mathbf{x}(\mathbf{t}^0) = \boldsymbol{\psi}^0$.

Definition 2.5. A partition $\{\Omega_i\}_{i=1}^N$ of the parameter space Ω is a collection of a finite number of compact subsets $\Omega_i, i = 1, \ldots, N$, of Ω that cover Ω and are pairwise disjoint, i.e., $\Omega = \bigcup_{i=1}^N \Omega_i$ and $\operatorname{int}(\Omega_i) \cap \operatorname{int}(\Omega_j) = \emptyset \ \forall i, j \in \{1, \ldots, N\}, i \neq j$.

Using the notion of partition, piecewise continuous functions on Ω are defined. Each such function is associated with a partition and is continuous on the subsets of this partition. The subsets are referred to as the subsets of continuity.

Definition 2.6. A function $\mathbf{x} : \Omega \to \mathbb{R}^n$ is said to be

1. piecewise continuous on Ω if there exists a natural number $N_{\mathbf{x}}$ and a partition $\{\Omega_i^{\mathbf{x}}\}_{i=1}^{N_{\mathbf{x}}}$ for $\mathbf{x}$ such that

   (a) $\mathbf{x}$ is continuous on each subset $\Omega_i^{\mathbf{x}}, i \in \{1, \ldots, N_{\mathbf{x}}\}$,

   (b) $\forall\, \bar{\mathbf{t}} \in \Omega_i^{\mathbf{x}} \cap \Omega_j^{\mathbf{x}}$ and $\forall i, j \in \{1, \ldots, N_{\mathbf{x}}\}, i \neq j, \ \ \mathbf{x}(\bar{\mathbf{t}}) = \lim_{\mathbf{t} \in \operatorname{int}(\Omega_i^{\mathbf{x}}), \mathbf{t} \to \bar{\mathbf{t}}} \mathbf{x}(\mathbf{t})$ or $\mathbf{x}(\bar{\mathbf{t}}) = \lim_{\mathbf{t} \in \operatorname{int}(\Omega_j^{\mathbf{x}}), \mathbf{t} \to \bar{\mathbf{t}}} \mathbf{x}(\mathbf{t})$;

2. bounded on Ω if there exists a $C > 0$ such that $\|\mathbf{x}(\mathbf{t})\|_\infty < C$ for all $\mathbf{t} \in \Omega$, where $\|\cdot\|_\infty$ denotes the maximum l^∞ norm in $\mathbb{R}^n$.

Let $\bar{\mathbb{C}}_p(\Omega, \mathbb{R}^n)$ denote the set of piecewise continuous and bounded functions $\mathbf{x} : \Omega \to \mathbb{R}^n$.

Given Assumption 2.4, a set of admissible functions is defined as

$$W = \{\mathbf{x} \in \bar{\mathbb{C}}_p(\Omega, \mathbb{R}^n) : \mathbf{x}(\mathbf{t}) \in \mathcal{S} \ \ \forall \mathbf{t} \in \Omega\} \tag{2.3}$$

and a set of feasible functions is defined as

$$X = \{\mathbf{x} \in W : \mathbf{x}(\mathbf{t}) \in \mathcal{X}(\mathbf{t}) \ \ \forall \mathbf{t} \in \Omega \text{ and Assumption 2.4 holds}\} \tag{2.4}$$

Given sets (2.3) and (2.4), mp-NLP (2.1)-(2.2) can be written as

$$\begin{aligned} f^*(\mathbf{t}) = \min_{\mathbf{x}} \quad & f(\mathbf{x}(\mathbf{t}); \mathbf{t}) \\ s.t. \quad & \mathbf{x} \in X = \{\mathbf{x} \in W : \mathbf{g}\,(\mathbf{x}(\mathbf{t}); \mathbf{t}) \le \mathbf{0}, \mathbf{h}\,(\mathbf{x}(\mathbf{t}); \mathbf{t}) = \mathbf{0} \ \ \forall \mathbf{t} \in \Omega\} \\ & \mathbf{t} \in \Omega \end{aligned} \tag{2.5}$$

Definition 2.7. Let Assumptions 2.1, 2.3, and 2.4 hold. A function $\mathbf{x}^* \in X$ is an optimal solution to mp-NLP (2.5) provided $f(\mathbf{x}^*(\mathbf{t}); \mathbf{t}) \le f(\mathbf{x}(\mathbf{t}); \mathbf{t}) \ \ \forall \mathbf{t} \in \Omega \ \ \forall \mathbf{x} \in X$, or equivalently, $f(\mathbf{x}^*(\mathbf{t}); \mathbf{t}) \le f(\boldsymbol{\psi}; \mathbf{t}) \ \ \forall (\boldsymbol{\psi}; \mathbf{t}) \in \mathcal{X}(\mathbf{t}) \times \Omega \ \ \forall \mathbf{t} \in \Omega$. If $\mathbf{x}^* \in X$ is an optimal solution function then $f^*(\mathbf{t}) = f(\mathbf{x}^*(\mathbf{t}), \mathbf{t}) \ \ \forall \mathbf{t} \in \Omega$.

Problems (2.1)-(2.2) and (2.5) are equivalent in the sense that they yield the same optimal objective value for all $\mathbf{t} \in \Omega$.

Theorem 2.8. *Let $\boldsymbol{\psi}^*_{\bar{\mathbf{t}}} \in \mathcal{X}(\bar{\mathbf{t}})$ be an optimal solution to mp-NLP (2.1)-(2.2) for $\mathbf{t} = \bar{\mathbf{t}} \in \Omega$ and $\mathbf{x}^* \in X$ be an optimal solution to (2.5). Then $f(\boldsymbol{\psi}^*_{\bar{\mathbf{t}}}, \bar{\mathbf{t}}) = f(\mathbf{x}^*(\bar{\mathbf{t}}), \bar{\mathbf{t}}) \ \ \forall \bar{\mathbf{t}} \in \Omega$.*

Proof. Let $\boldsymbol{\psi}^*_{\bar{\mathbf{t}}} \in \mathcal{X}(\mathbf{t})$ be an optimal solution to mp-NLP (2.1)-(2.2) for $\mathbf{t} = \bar{\mathbf{t}} \in \Omega$. Then, by Definition 2.1, $f(\boldsymbol{\psi}^*_{\bar{\mathbf{t}}}, \bar{\mathbf{t}}) \le f(\boldsymbol{\psi}, \bar{\mathbf{t}}) \ \ \forall \boldsymbol{\psi} \in \mathcal{X}(\bar{\mathbf{t}})$. Since $\mathbf{x}^* \in X$, then $\mathbf{x}^*(\bar{\mathbf{t}}) \in \mathcal{X}(\bar{\mathbf{t}})$ and therefore $f(\boldsymbol{\psi}^*_{\bar{\mathbf{t}}}, \bar{\mathbf{t}}) \le f(\mathbf{x}^*(\bar{\mathbf{t}}), \bar{\mathbf{t}})$. Since $\bar{\mathbf{t}}$ is an arbitrary element in Ω, then

$$f(\boldsymbol{\psi}^*_{\bar{\mathbf{t}}}, \bar{\mathbf{t}}) \le f(\mathbf{x}^*(\bar{\mathbf{t}}), \bar{\mathbf{t}}) \ \ \forall \bar{\mathbf{t}} \in \Omega \tag{2.6}$$

On the other hand, let $\mathbf{x}^* \in X$ be an optimal solution to (2.5). Then by Definition 2.7,

$$f(\mathbf{x}^*(\mathbf{t}); \mathbf{t}) \le f(\boldsymbol{\psi}; \mathbf{t}) \ \ \forall (\boldsymbol{\psi}; \mathbf{t}) \in \mathcal{X}(\mathbf{t}) \times \Omega \ \ \forall \mathbf{t} \in \Omega \tag{2.7}$$

Inequalities (2.6) and (2.7) yield the final result. □

To check the conditions for the existence of an optimal solution function to mp-NLP (2.5), the following definitions are needed.

Definition 2.9. 1. All functions $\mathbf{x} \in X$ are said to be uniformly bounded in $\mathbb{R}^n$ if there exists a $C > 0$ such that $\|\mathbf{x}\|_\infty < C$ for all $\mathbf{x} \in X$ and $\mathbf{t} \in \Omega$.

2. A sequence of functions $\{\mathbf{x}_m\}$ in X is said to converge to $\mathbf{x}^*$ if $\|\mathbf{x}_m(\mathbf{t}) - \mathbf{x}^*(\mathbf{t})\|_\infty \to 0$ $\forall \mathbf{t} \in \Omega$ when $m \to \infty$.

Theorem 2.10. *If all $\mathbf{x} \in X$ are uniformly bounded functions on Ω then an optimal solution function $\mathbf{x}^* \in X$ to mp-NLP (2.5) exists.*

Proof. Let $\{\mathbf{x}_m\}$ be a minimizing sequence in X, i.e., assume

$$\lim_{m\to\infty} f(\mathbf{x}_m(\mathbf{t});\mathbf{t}) = \min_{\mathbf{x}\in X} f(\mathbf{x}(\mathbf{t});\mathbf{t}) \qquad \forall \mathbf{t}\in\Omega$$

Since each element $\mathbf{x}_i \in \{\mathbf{x}_m\}$ is uniformly bounded on Ω, there exists a convergent subsequence $\{\mathbf{x}_{m_k}\} \to \mathbf{x}^*$. The goal is to show that $\mathbf{x}^* \in X$. Since $\{\mathbf{x}_{m_k}\} \to \mathbf{x}^*$ and $\mathbf{x}_i$ is bounded $\forall i$, $\mathbf{x}^*$ is also bounded. Since $f(\mathbf{x}(\mathbf{t});\mathbf{t})$ is bounded, $f(\mathbf{x}^*(\mathbf{t});\mathbf{t})$ is bounded. By continuity $\mathbf{g}(\mathbf{x}_{m_k}(\mathbf{t});\mathbf{t}) \to \mathbf{g}(\mathbf{x}^*(\mathbf{t});\mathbf{t})\ \forall \mathbf{t}\in\Omega$ and $\mathbf{h}(\mathbf{x}_{m_k}(\mathbf{t});\mathbf{t}) \to \mathbf{h}(\mathbf{x}^*(\mathbf{t});\mathbf{t})\ \forall \mathbf{t}\in\Omega$ with $\mathbf{g}(\mathbf{x}^*(\mathbf{t});\mathbf{t}) \le \mathbf{0}$, $\mathbf{h}(\mathbf{x}^*(\mathbf{t});\mathbf{t}) = \mathbf{0}$ since $\{\mathbf{x}_m\} \subset X$. Therefore $\mathbf{x}^* \in X$.

Since f is continuous in $\mathbf{x}$, it is lower semicontinuous, and $\lim_{n_k\to\infty} f(\mathbf{x}_{m_k}(\mathbf{t});\mathbf{t}) > f(\mathbf{x}^*(\mathbf{t});\mathbf{t})\,\forall \mathbf{t}\in\Omega$. Using the fact that $\{\mathbf{x}_{m_k}\}$ is a minimizing sequence, $\min_{\mathbf{x}\in X} f(\mathbf{x}(\mathbf{t});\mathbf{t}) = \lim_{m\to\infty} f(\mathbf{x}_m(\mathbf{t});\mathbf{t}) = f(\mathbf{x}^*(\mathbf{t});\mathbf{t})\,\forall \mathbf{t}\in\Omega$ so that $\mathbf{x}^*$ is an optimal solution. □

The goal of the paper is to study (2.5) at two levels: (i) to develop its theoretical properties within a framework of Lagrangian duality, and (ii) to propose a solution algorithm. Based on Theorem 2.8, both will be applicable to (2.1)-(2.2). Weak duality is first examined in the next section.

2.3 Parametric Lagrangian dual problem

In this section, mp-NLP (2.5) is treated as a Primal Problem (PP) for which a Lagrangian Dual Problem (LDP) is stated and a weak duality result between the PP and LDP is established. As is the case for the PP, solutions to the LDP are functions of the parameter space. Defining an LDP therefore requires that the dual variables be defined as functions of $\mathbf{t}$, but is otherwise constructed in the usual way.

Let $\mathbf{y} = (\mathbf{u},\mathbf{v}) \in \bar{\mathbb{C}}_p(\Omega,\mathbb{R}^m\times\mathbb{R}^l)$ be the Lagrangian dual function composed of a pair of functions $\mathbf{u}\in\bar{\mathbb{C}}_p(\Omega,\mathbb{R}^m)$ and $\mathbf{v}\in\bar{\mathbb{C}}_p(\Omega,\mathbb{R}^l)$ that are associated with the functions $\mathbf{g}$ and $\mathbf{h}$ respectively. For every $\mathbf{x}\in W$, let $(\mathbf{x},\mathbf{u},\mathbf{v}) \in \bar{\mathbb{C}}_p(\Omega,\mathbb{R}^n\times\mathbb{R}^m\times\mathbb{R}^l)$. Let $\mathcal{L} : \mathbb{R}^n\times\mathbb{R}^m\times\mathbb{R}^l\times\Omega \to \mathbb{R}$ be the Lagrangian function defined as

$$\mathcal{L}(\mathbf{x}(\mathbf{t}),\mathbf{u}(\mathbf{t}),\mathbf{v}(\mathbf{t});\mathbf{t}) = f(\mathbf{x}(\mathbf{t});\mathbf{t}) + \mathbf{u}(\mathbf{t})^T\mathbf{g}(\mathbf{x}(\mathbf{t});\mathbf{t}) + \mathbf{v}(\mathbf{t})^T\mathbf{h}(\mathbf{x}(\mathbf{t});\mathbf{t}) \quad \forall \mathbf{t}\in\Omega \tag{2.8}$$

The Relaxed Primal Problem (RPP) takes the form:

$$\begin{aligned} \min_{\mathbf{x}} \quad & \mathcal{L}(\mathbf{x}(\mathbf{t}),\mathbf{u}(\mathbf{t}),\mathbf{v}(\mathbf{t});\mathbf{t}) \\ s.t. \quad & \mathbf{x}\in W \\ & \mathbf{t}\in\Omega \end{aligned} \tag{2.9}$$

Assumption 2.5. For every $\mathbf{x}\in W$ and each $\mathbf{t}^0\in\Omega$ consider $\mathbf{x}(\mathbf{t}^0)$, the image of $\mathbf{t}^0$ under $\mathbf{x}$. Assume there exist a function $\widetilde{\mathbf{x}}\in W$ and a neighborhood of $\mathbf{t}^0$, $N_\delta(\mathbf{t}^0)\subset\Omega$, such that $\widetilde{\mathbf{x}}$ is continuous and bounded on $N_\delta(\mathbf{t}^0)$ and $\widetilde{\mathbf{x}}(\mathbf{t}^0) = \mathbf{x}(\mathbf{t}^0)$.

Definition 2.11. A function $\mathbf{x}^*\in W$ is an optimal solution to RPP (2.9) provided $\mathcal{L}(\mathbf{x}^*(\mathbf{t}),\mathbf{u}(\mathbf{t}),\mathbf{v}(\mathbf{t});\mathbf{t}) \le \mathcal{L}(\mathbf{x}(\mathbf{t}),\mathbf{u}(\mathbf{t}),\mathbf{v}(\mathbf{t});\mathbf{t}) \quad \forall\mathbf{x}\in W \quad \forall\mathbf{t}\in\Omega$. If $\mathbf{x}^*\in W$ is an optimal solution function then θ denotes the optimal value function, $\theta : \mathbb{R}^m\times\mathbb{R}^l\times\Omega\to\mathbb{R}$, that is referred to as the dual function and is defined as

$$\begin{aligned} \theta(\mathbf{u}(\mathbf{t}),\mathbf{v}(\mathbf{t});\mathbf{t}) &= \min_{\mathbf{x}} \{\mathcal{L}(\mathbf{x}(\mathbf{t}),\mathbf{u}(\mathbf{t}),\mathbf{v}(\mathbf{t});\mathbf{t}) : \mathbf{x}\in W\} \quad \forall\mathbf{t}\in\Omega \\ &= \mathcal{L}(\mathbf{x}^*(\mathbf{t}),\mathbf{u}(\mathbf{t}),\mathbf{v}(\mathbf{t});\mathbf{t}) \ \ \forall\mathbf{t}\in\Omega \end{aligned} \tag{2.10}$$

A continuity assumption is needed for the optiml solution function $\mathbf{x}^* \in W$ to (2.9).

Assumption 2.6. Let $\mathbf{t} \in \Omega$. If $\mathbf{y} = (\mathbf{u}, \mathbf{v}) \in \bar{\mathbb{C}}_p(\Omega, \mathbb{R}^m \times \mathbb{R}^l)$ is continuous at $\mathbf{t}$ then the associated optimal solution function $\mathbf{x}^* \in W$ to (2.9) is also continuous at $\mathbf{t}$.

An LDP to PP (2.5) is formulated as:

$$\begin{aligned} \theta^*(\mathbf{t}) = \max_{\mathbf{u},\mathbf{v}} \quad & \theta(\mathbf{u}(\mathbf{t}), \mathbf{v}(\mathbf{t}); \mathbf{t}) \\ s.t. \quad & \mathbf{u}(\mathbf{t}) \geq \mathbf{0} \\ & \mathbf{t} \in \Omega \end{aligned} \tag{2.11}$$

A pair of functions $(\mathbf{u}, \mathbf{v}) \in \bar{\mathbb{C}}_p(\Omega, \mathbb{R}^m \times \mathbb{R}^l)$ is a feasible solution to LDP (2.11) provided $\mathbf{u}(\mathbf{t}) \geq \mathbf{0}\ \ \forall \mathbf{t} \in \Omega$.

Assumption 2.7. For every $\mathbf{y} = (\mathbf{u}, \mathbf{v}) \in \bar{\mathbb{C}}_p(\Omega, \mathbb{R}^m \times \mathbb{R}^l)$ and for each $\mathbf{t}^0 \in \Omega$ consider $\mathbf{y}(\mathbf{t}^0) = (\mathbf{u}(\mathbf{t}^0), \mathbf{v}(\mathbf{t}^0))$, the image of $\mathbf{t}^0$ under $\mathbf{y}$. Assume there exist a function $\widetilde{\mathbf{y}} = (\widetilde{\mathbf{u}}, \widetilde{\mathbf{v}}) \in \bar{\mathbb{C}}_p(\Omega, \mathbb{R}^m \times \mathbb{R}^l)$ and a neighborhood of $\mathbf{t}^0$, $N_\delta(\mathbf{t}^0) \subset \Omega$, such that $\widetilde{\mathbf{y}}$ is continuous and bounded on $N_\delta(\mathbf{t}^0)$ and $\widetilde{\mathbf{y}}(\mathbf{t}^0) = (\widetilde{\mathbf{u}}(\mathbf{t}^0), \widetilde{\mathbf{v}}(\mathbf{t}^0)) = (\mathbf{u}(\mathbf{t}^0), \mathbf{v}(\mathbf{t}^0)) = \mathbf{y}(\mathbf{t}^0)$ with $\widetilde{\mathbf{u}}(\mathbf{t}) \geq \mathbf{0}$ $\forall \mathbf{t} \in N_\delta(\mathbf{t}^0)$.

Definition 2.12. A pair of functions $(\mathbf{u}^*, \mathbf{v}^*), \mathbf{u}^*(\mathbf{t}) \geq \mathbf{0}\ \ \forall \mathbf{t} \in \Omega$, is an optimal solution to LDP (2.11) provided $\theta(\mathbf{u}^*(\mathbf{t}), \mathbf{v}^*(\mathbf{t}); \mathbf{t}) \geq \theta(\mathbf{u}(\mathbf{t}), \mathbf{v}(\mathbf{t}); \mathbf{t})\ \forall \mathbf{u}(\mathbf{t}) \geq \mathbf{0}\ \ \forall \mathbf{t} \in \Omega$. If $(\mathbf{u}^*, \mathbf{v}^*)$ is an optimal solution then θ^* denotes the optimal dual value function and $\theta^*(\mathbf{t}) = \theta(\mathbf{u}^*(\mathbf{t}), \mathbf{v}^*(\mathbf{t}); \mathbf{t})\ \ \forall \mathbf{t} \in \Omega$.

Assume that neither the Lagrangian function $\mathcal{L}$ nor the dual function θ takes on the value of $-\infty$ or ∞ respectively during optimization. It is straightforward to prove weak duality for PP (2.5) and LDP (2.11).

Theorem 2.13 (Weak Duality for mp-NLP)**.** *Let $\hat{\mathbf{x}} \in X$ and $(\hat{\mathbf{u}}, \hat{\mathbf{v}})$ be a feasible solution to LDP* (2.11)*. Then $f(\hat{\mathbf{x}}(\mathbf{t}); \mathbf{t}) \geq \theta(\hat{\mathbf{u}}(\mathbf{t}), \hat{\mathbf{v}}(\mathbf{t}); \mathbf{t})\ \ \forall \mathbf{t} \in \Omega$.*

Proof. $\forall \mathbf{t} \in \Omega$ we have:

$$\begin{aligned} \theta(\hat{\mathbf{u}}(\mathbf{t}), \hat{\mathbf{v}}(\mathbf{t}); \mathbf{t}) &= \min \{\mathcal{L}(\mathbf{x}(\mathbf{t}), \hat{\mathbf{u}}(\mathbf{t}), \hat{\mathbf{v}}(\mathbf{t}); \mathbf{t}), \mathbf{x} \in W\} \\ &= \min \left\{ f(\mathbf{x}(\mathbf{t}); \mathbf{t}) + \hat{\mathbf{u}}(\mathbf{t})^T \mathbf{g}(\mathbf{x}(\mathbf{t}); \mathbf{t}) + \hat{\mathbf{v}}(\mathbf{t})^T \mathbf{h}(\mathbf{x}(\mathbf{t}); \mathbf{t}) : \mathbf{x} \in W \right\} \\ &\leq f(\hat{\mathbf{x}}(\mathbf{t}); \mathbf{t}) + \hat{\mathbf{u}}(\mathbf{t})^T \mathbf{g}(\hat{\mathbf{x}}(\mathbf{t}); \mathbf{t}) + \hat{\mathbf{v}}(\mathbf{t})^T \mathbf{h}(\hat{\mathbf{x}}(\mathbf{t}); \mathbf{t}) \\ &\leq f(\hat{\mathbf{x}}(\mathbf{t}); \mathbf{t}) \end{aligned}$$

where the second inequality results from $\hat{\mathbf{u}}(\mathbf{t}) \geq \mathbf{0}, \mathbf{g}(\hat{\mathbf{x}}(\mathbf{t}); \mathbf{t}) \leq \mathbf{0}$, and $\mathbf{h}(\hat{\mathbf{x}}(\mathbf{t}); \mathbf{t}) = \mathbf{0}\ \ \forall \mathbf{t} \in \Omega$. □

A strong duality relationship is more complicated to establish. To do so, in Section 3 an auxiliary problem related to mp-NLP (2.5) is formulated and strong duality is shown to hold between the auxiliary problem and its Lagrangian dual. This strong duality relationship is then used in Section 4 to prove strong duality between mp-NLP (2.5) and its LDP (2.11).

3 Integral Counterpart to mp-NLP

To further investigate duality for mp-NLP (2.5), a closely related auxiliary optimization problem is formulated in a function space. Weak and strong duality is established for this auxiliary problem.

Definition 3.1. Let

1. $L^2(\Omega, \mathbb{R}^n)$ be the Hilbert space of square-integrable, vector-valued functions on Ω defined as $L^2(\Omega, \mathbb{R}^n) = \{\mathbf{x} : \Omega \to \mathbb{R}^n \text{ such that } \|\mathbf{x}\|_2^2 = \int_\Omega |\mathbf{x}(\mathbf{t})|^2 \, dt < \infty\}$;
2. the inner product be defined $\langle \mathbf{x}_1, \mathbf{x}_2 \rangle = \int_\Omega \mathbf{x}_1(\mathbf{t})^T \mathbf{x}_2(\mathbf{t}) \, dt \quad$ for $\mathbf{x}_1, \mathbf{x}_2 \in L^2(\Omega, \mathbb{R}^n)$;
3. $\mathbf{x}_1 \leq \mathbf{x}_2$ denote $\mathbf{x}_1(\mathbf{t}) \leq \mathbf{x}_2(\mathbf{t}) \ \forall \mathbf{t} \in \Omega$ for $\mathbf{x}_1, \mathbf{x}_2 \in L^2(\Omega, \mathbb{R}^n)$.

Using this function space setting, an optimization problem is formulated on the feasible set of mp-NLP (2.5). Since the parameter space Ω is compact, $\bar{\mathbb{C}}_p(\Omega, \mathbb{R}^n) \subseteq L^2(\Omega, \mathbb{R}^n)$, and therefore $X \subseteq W \subset L^2(\Omega, \mathbb{R}^n)$. A functional $J : L^2(\Omega, \mathbb{R}^n) \to \mathbb{R}$ is defined and an integral counterpart to (2.5) is obtained:

$$\begin{aligned} J(\bar{\mathbf{x}}) = \quad \min_{\mathbf{x}} \quad & J(\mathbf{x}) = \int_\Omega f(\mathbf{x}(\mathbf{t}); \mathbf{t}) \, dt \\ s.t. \quad & \mathbf{x} \in X \subset L^2(\Omega, \mathbb{R}^n) \end{aligned} \tag{3.1}$$

Problem (3.1) is referred to as the integral primal problem (IPP). A similar formulation can be found in [24] (p. 250). Since f is continuous and convex in both arguments, J is continuous and convex in $\mathbf{x}$.

Definition 3.2. A function $\bar{\mathbf{x}} \in X$ is an optimal solution to IPP (3.1) provided $J(\bar{\mathbf{x}}) \leq J(\mathbf{x})$ for all $\mathbf{x} \in X$.

An optimal solution to (3.1) exists under assumptions similar to those of Theorem 2.10. The reader is referred to [40] for weaker conditions guaranteeing well-posedness of (3.1).

Duality notions are introduced for IPP in a similar way as for PP. By definition, the functions $\mathbf{g}, \mathbf{h}, \mathbf{u}, \mathbf{v}$ are elements of the Hilbert space and the integral Lagrangian function is defined $\mathcal{IL} : L^2(\Omega, \mathbb{R}^n) \times L^2(\Omega, \mathbb{R}^m) \times L^2(\Omega, \mathbb{R}^l)) \to \mathbb{R}$

$$\mathcal{IL}(\mathbf{x}, \mathbf{u}, \mathbf{v}) \; = \int_\Omega \left[f(\mathbf{x}(\mathbf{t}); \mathbf{t}) + \mathbf{u}(\mathbf{t})^T \mathbf{g}(\mathbf{x}(\mathbf{t}); \mathbf{t}) + \mathbf{v}(\mathbf{t})^T \mathbf{h}(\mathbf{x}(\mathbf{t}); \mathbf{t}) \right] \, dt \tag{3.2}$$

or equivalently

$$\mathcal{IL}(\mathbf{x}, \mathbf{u}, \mathbf{v}) \; = \int_\Omega [f(\mathbf{x}(\mathbf{t}); \mathbf{t})] \, dt + \langle \mathbf{u}, \mathbf{g}(\mathbf{x}) \rangle + \langle \mathbf{v}, \mathbf{h}(\mathbf{x}) \rangle \tag{3.3}$$

and the relaxed integral primal problem (RIPP) is formulated

$$\begin{aligned} \min_{\mathbf{x}} \quad & \mathcal{IL}(\mathbf{x}, \mathbf{u}, \mathbf{v}) \\ s.t. \quad & \mathbf{x} \in W \subset L^2(\Omega, \mathbb{R}^n) \end{aligned} \tag{3.4}$$

Definition 3.3. A function $\bar{\mathbf{x}} \in W$ is an optimal solution to RIPP (3.4) provided $\mathcal{IL}(\bar{\mathbf{x}}, \mathbf{u}, \mathbf{v}) \leq \mathcal{IL}(\mathbf{x}, \mathbf{u}, \mathbf{v}) \quad \forall \mathbf{x} \in W$. If $\bar{\mathbf{x}} \in W$ is an optimal solution function then Δ denotes the optimal value function, $\Delta : L^2(\Omega, \mathbb{R}^m) \times L^2(\Omega, \mathbb{R}^l) \to \mathbb{R}$, that is referred to as the integral dual function and is defined as

$$\begin{aligned} \Delta(\mathbf{u}, \mathbf{v}) &= \min_{\mathbf{x}} \{\mathcal{IL}(\mathbf{x}, \mathbf{u}, \mathbf{v}) : \mathbf{x} \in W\} \\ &= \mathcal{IL}(\bar{\mathbf{x}}, \mathbf{u}, \mathbf{v}) \end{aligned} \tag{3.5}$$

An Integral Dual Problem (IDP) to IPP assumes the form

$$\begin{aligned} \bar{\Delta} = \max_{\mathbf{u},\mathbf{v}} &\quad \Delta(\mathbf{u},\mathbf{v}) \\ s.t. &\quad \mathbf{u} \geq \mathbf{0} \end{aligned} \tag{3.6}$$

A pair of functions $(\mathbf{u},\mathbf{v})$ is a feasible solution to IDP (3.6) provided $\mathbf{u} \geq \mathbf{0}$.

Definition 3.4. A pair of functions $(\bar{\mathbf{u}},\bar{\mathbf{v}}), \bar{\mathbf{u}} \geq \mathbf{0}$, is an optimal solution to IDP (3.6) if $\Delta(\bar{\mathbf{u}},\bar{\mathbf{v}}) \geq \Delta(\mathbf{u},\mathbf{v}) \;\; \forall \mathbf{u} \geq \mathbf{0}$. If $(\bar{\mathbf{u}},\bar{\mathbf{v}})$ is an optimal solution then $\bar{\Delta}$ denotes the optimal value of the integral dual function.

Assume that neither the integral Lagrangian function $\mathcal{IL}$ nor the integral dual function Δ takes on the value of $-\infty$ or ∞ respectively during optimization. It is straightforward to show that weak duality holds between IPP (3.1) and IDP (3.6).

Theorem 3.5 (Weak Duality for Integral mp-NLP)**.** *Let $\hat{\mathbf{x}}$ be a feasible solution to IPP* (3.1) *and $(\hat{\mathbf{u}},\hat{\mathbf{v}})$ be a feasible solution to IDP* (3.6)*. Then $\int_\Omega f(\hat{\mathbf{x}}(\mathbf{t});\mathbf{t})\, dt \geq \Delta(\hat{\mathbf{u}},\hat{\mathbf{v}})$.*

Proof. Let $\hat{\mathbf{x}}$ be a feasible solution to (3.1) and $(\hat{\mathbf{u}},\hat{\mathbf{v}})$ be a feasible solution to (3.6). Then:

$$\begin{aligned} \Delta(\hat{\mathbf{u}},\hat{\mathbf{v}}) &= \min_{\mathbf{x}} \{\mathcal{IL}(\mathbf{x},\hat{\mathbf{u}},\hat{\mathbf{v}}) : \mathbf{x} \in W\} \\ &= \min_{\mathbf{x}} \left\{ \int_\Omega \left[f(\mathbf{x}(\mathbf{t});\mathbf{t}) + \hat{\mathbf{u}}(\mathbf{t})^T \mathbf{g}(\mathbf{x}(\mathbf{t});\mathbf{t}) + \hat{\mathbf{v}}(\mathbf{t})^T \mathbf{h}(\mathbf{x}(\mathbf{t});\mathbf{t}) \right] \, dt : \mathbf{x} \in W \right\} \\ &\leq \int_\Omega \left[f(\hat{\mathbf{x}}(\mathbf{t});\mathbf{t}) + \hat{\mathbf{u}}(\mathbf{t})^T \mathbf{g}(\hat{\mathbf{x}}(\mathbf{t});\mathbf{t}) + \hat{\mathbf{v}}(\mathbf{t})^T \mathbf{h}(\hat{\mathbf{x}}(\mathbf{t});\mathbf{t}) \right] \, dt \\ &\leq \int_\Omega f(\hat{\mathbf{x}}(\mathbf{t});\mathbf{t}) \, dt \end{aligned}$$

where the second inequality results from $\hat{\mathbf{u}}(\mathbf{t})^T \mathbf{g}(\hat{\mathbf{x}}(\mathbf{t});\mathbf{t}) \leq 0$ and $\hat{\mathbf{v}}(\mathbf{t})^T \mathbf{h}(\hat{\mathbf{x}}(\mathbf{t});\mathbf{t}) = 0 \;\; \forall \mathbf{t} \in \Omega$. □

Lemma 3.6 is needed to prove strong duality for IPP (3.1) and IDP (3.6) in Theorem 3.7. Both these results extend the results for NLPs in [5] (p.266).

Lemma 3.6. *Define the following two systems:*

$$\begin{aligned} &\textit{System 1: } J(\mathbf{x}) < 0, \; \mathbf{g}(\mathbf{x}(\mathbf{t});\mathbf{t}) \leq \mathbf{0}, \; \mathbf{h}(\mathbf{x}(\mathbf{t});\mathbf{t}) = \mathbf{0} \;\; \forall \mathbf{t} \in \Omega \;\; \textit{for some } \mathbf{x} \in W \\ &\textit{System 2: } u_0 J(\mathbf{x}) + \int_\Omega \mathbf{u}(\mathbf{t})^T \mathbf{g}(\mathbf{x}(\mathbf{t});\mathbf{t}) \, dt + \int_\Omega \mathbf{v}(\mathbf{t})^T \mathbf{h}(\mathbf{x}(\mathbf{t});\mathbf{t}) \, dt \geq 0 \\ &\qquad\qquad (u_0,\mathbf{u}(\mathbf{t})) \geq \mathbf{0}, \; (u_0,\mathbf{u}(\mathbf{t}),\mathbf{v}(\mathbf{t})) \neq \mathbf{0}, \;\; \forall \mathbf{t} \in \Omega \;\; \forall \mathbf{x} \in W \end{aligned}$$

If System 1 has no solution $\mathbf{x} \in W \;\; \forall \mathbf{t} \in \Omega$ then System 2 has a solution $(u_0,\mathbf{u}(\mathbf{t}),\mathbf{v}(\mathbf{t})) \in \mathbb{R} \times \mathbb{R}^m \times \mathbb{R}^l \;\; \forall \mathbf{t} \in \Omega$. The converse is true if $u_0 > 0$.

Proof. Let System 1 have no solution. Define the set

$$\Lambda = \{(p,\mathbf{q}(\mathbf{t}),\mathbf{r}(\mathbf{t})) \in \mathbb{R} \times \mathbb{R}^m \times \mathbb{R}^l : p > J(\mathbf{x}), \mathbf{q}(\mathbf{t}) \geq \mathbf{g}(\mathbf{x}(\mathbf{t});\mathbf{t}), \mathbf{r}(\mathbf{t}) = \mathbf{h}(\mathbf{x}(\mathbf{t});\mathbf{t}) \;\; \forall \mathbf{t} \in \Omega \text{ for some } \mathbf{x} \in W\}$$

Since J and $\mathbf{g}$ are convex and $\mathbf{h}$ is affine, Λ is a convex set. If System 1 has no solution then $(0,\mathbf{0},\mathbf{0}) \notin \Lambda$. By the Supporting Hyperplane Theorem [31] (Thm.2, p. 133), there exists a nonzero $(u_0, \mathbf{u}(\mathbf{t}), \mathbf{v}(\mathbf{t})) \in \mathbb{R} \times \mathbb{R}^m \times \mathbb{R}^l$ such that $u_0 p + \int_\Omega \mathbf{u}(\mathbf{t})^T \mathbf{q}(\mathbf{t})\ dt + \int_\Omega \mathbf{v}(\mathbf{t})^T \mathbf{r}(\mathbf{t})\ dt \geq 0\ \ \forall (p, \mathbf{q}(\mathbf{t}), \mathbf{r}(\mathbf{t})) \in \mathrm{cl}(\Lambda)$ and $\forall \mathbf{t} \in \Omega$. Now fix $\mathbf{x} \in W$. Since p and $\mathbf{q}(\mathbf{t})$ can be made arbitrarily large, it must be that $u_0 \geq 0$ and $\mathbf{u}(\mathbf{t}) \geq \mathbf{0}\ \ \forall \mathbf{t} \in \Omega$. Furthermore, $(p, \mathbf{q}(\mathbf{t}), \mathbf{r}(\mathbf{t})) = (J(\mathbf{x}), \mathbf{g}(\mathbf{x}(\mathbf{t}); \mathbf{t}), \mathbf{h}(\mathbf{x}(\mathbf{t}); \mathbf{t})) \in \mathrm{cl}(\Lambda)\ \ \forall \mathbf{t} \in \Omega$. Therefore, $u_0 J(\mathbf{x}) + \int_\Omega \mathbf{u}(\mathbf{t})^T \mathbf{g}(\mathbf{x}(\mathbf{t}); \mathbf{t})\ dt + \int_\Omega \mathbf{v}(\mathbf{t})^T \mathbf{h}(\mathbf{x}(\mathbf{t}); \mathbf{t})\ dt \geq 0$. Since $\mathbf{x}$ is arbitrary, the latter inequality holds $\forall \mathbf{x} \in W$ and System 2 has a solution.

For the converse, $\forall \mathbf{x} \in W$ and $\forall \mathbf{t} \in \Omega$ let $(u_0, \mathbf{u}(\mathbf{t}), \mathbf{v}(\mathbf{t}))$ be a solution to System 2 with $u_0 > 0$. Choose $\mathbf{x} \in W$ such that $\mathbf{g}(\mathbf{x}(\mathbf{t}); \mathbf{t}) \leq \mathbf{0}$ and $\mathbf{h}(\mathbf{x}; \mathbf{t}) = \mathbf{0}\ \ \forall \mathbf{t} \in \Omega$. Since $\mathbf{u}(\mathbf{t}) \geq \mathbf{0}$ and $\mathbf{g}(\mathbf{x}(\mathbf{t}); \mathbf{t}) \leq \mathbf{0}\ \ \forall \mathbf{t} \in \Omega$, it must be $u_0 J(\mathbf{x}) \geq 0$. Since $u_0 > 0$, $J(\mathbf{x}) \geq 0$ and System 1 does not have a solution. □

Strong duality for the integral mp-NLP can now be proved.

Theorem 3.7 (Strong Duality for Integral mp-NLP)**.** *Assume the following constraint qualification holds:* $\exists\ \hat{\mathbf{x}} \in W$ *such that* $\mathbf{g}(\hat{\mathbf{x}}(\mathbf{t}); \mathbf{t}) < \mathbf{0}$ *and* $\mathbf{h}(\hat{\mathbf{x}}(\mathbf{t}); \mathbf{t}) = \mathbf{0}\ \ \forall \mathbf{t} \in \Omega$*, and* $\mathbf{0} \in \mathit{int}\ \mathbf{h}(W)$*, where* $\mathbf{h}(W) = \{\mathbf{h}(\mathbf{x}(\mathbf{t}); \mathbf{t}) : \mathbf{x} \in W, \mathbf{t} \in \Omega\}$*. Let* $\bar{\mathbf{x}} \in X$ *be an optimal solution to IPP* (3.1) *and* $(\bar{\mathbf{u}}, \bar{\mathbf{v}})$ *with* $\bar{\mathbf{u}} \geq \mathbf{0}$ *be an optimal solution to IDP* (3.6)*. Then*

$$J(\bar{\mathbf{x}}) = \min_{\mathbf{x}} \left\{ \int_\Omega f(\mathbf{x}(\mathbf{t}); \mathbf{t})\ dt : \mathbf{x} \in X \right\} = \max_{\mathbf{u}, \mathbf{v}} \{\Delta(\mathbf{u}, \mathbf{v}) : \mathbf{u} \geq \mathbf{0}\} = \Delta(\bar{\mathbf{u}}, \bar{\mathbf{v}})$$

Proof. Consider the system

$$J(\mathbf{x}) - J(\bar{\mathbf{x}}) < 0,\ \mathbf{g}(\mathbf{x}(\mathbf{t}); \mathbf{t}) \leq \mathbf{0},\ \mathbf{h}(\mathbf{x}(\mathbf{t}); \mathbf{t}) = \mathbf{0},\ \ \forall \mathbf{t} \in \Omega\ \text{ for some }\ \mathbf{x} \in W$$

By definition of $J(\bar{\mathbf{x}})$, this system has no solution. By Lemma 3.6, there exists a nonzero $(u_0, \mathbf{u}(\mathbf{t}), \mathbf{v}(\mathbf{t}))$
$\in \mathbb{R} \times \mathbb{R}^m \times \mathbb{R}^l$ with $(u_0, \mathbf{u}(\mathbf{t})) \geq \mathbf{0}\ \ \forall \mathbf{t} \in \Omega$ such that

$$u_0 [J(\mathbf{x}) - J(\bar{\mathbf{x}})] + \int_\Omega \mathbf{u}^T(\mathbf{t}) \mathbf{g}(\mathbf{x}(\mathbf{t}); \mathbf{t})\ dt + \int_\Omega \mathbf{v}(\mathbf{t})^T \mathbf{h}(\mathbf{x}(\mathbf{t}); \mathbf{t})\ dt \geq 0\ \ \forall \mathbf{x} \in W$$

Assume $u_0 = 0$, then $\int_\Omega \mathbf{u}(\mathbf{t})^T \mathbf{g}(\mathbf{x}(\mathbf{t}); \mathbf{t})\ dt + \int_\Omega \mathbf{v}(\mathbf{t})^T \mathbf{h}(\mathbf{x}(\mathbf{t}); \mathbf{t})\ dt \geq 0\ \forall \mathbf{x} \in W$. By assumption, there is an $\hat{\mathbf{x}} \in W$ such that $\mathbf{g}(\hat{\mathbf{x}}(\mathbf{t}); \mathbf{t}) < \mathbf{0}$ and $\mathbf{h}(\hat{\mathbf{x}}(\mathbf{t}); \mathbf{t}) = \mathbf{0}\ \ \forall \mathbf{t} \in \Omega$, which implies $\mathbf{v}(\mathbf{t})^T \mathbf{h}(\hat{\mathbf{x}}(\mathbf{t}); \mathbf{t}) = 0\ \ \forall \mathbf{t} \in \Omega$, and therefore $\int_\Omega \mathbf{u}^T(\mathbf{t}) \mathbf{g}(\hat{\mathbf{x}}(\mathbf{t}); \mathbf{t})\ dt \geq 0$. Because $\mathbf{u}(\mathbf{t}) \geq \mathbf{0}$ and $\mathbf{g}(\hat{\mathbf{x}}(\mathbf{t}); \mathbf{t}) < \mathbf{0}$ then $\int_\Omega \mathbf{u}(\mathbf{t})^T \mathbf{g}(\hat{\mathbf{x}}(\mathbf{t}); \mathbf{t})\ dt \leq 0$, which forces $\mathbf{u}(\mathbf{t}) = \mathbf{0}\ \ \forall \mathbf{t} \in \Omega$ and implies $\int_\Omega \mathbf{v}(\mathbf{t})^T \mathbf{h}(\mathbf{x}(\mathbf{t}); \mathbf{t})\ dt \geq 0\ \ \forall \mathbf{x} \in W$. Since $\mathbf{0} \in \mathrm{int}\ \mathbf{h}(W)$ there is an $\widetilde{\mathbf{x}} \in W$ such that $\mathbf{h}(\widetilde{\mathbf{x}}(\mathbf{t}); \mathbf{t}) = -\lambda \mathbf{v}(\mathbf{t})\ \ \ \forall \mathbf{t} \in \Omega$ and $\lambda > 0$. Calculate $\int_\Omega \mathbf{v}(\mathbf{t})^T \mathbf{h}(\widetilde{\mathbf{x}}(\mathbf{t}); \mathbf{t})\ dt = -\lambda \int_\Omega \|\mathbf{v}(\mathbf{t})\|^2\ dt \geq 0$, which implies $\mathbf{v}(\mathbf{t}) = \mathbf{0}\ \ \forall \mathbf{t} \in \Omega$. In effect, assuming $u_0 = 0$ makes $\mathbf{u}(\mathbf{t}) = \mathbf{0}$ and $\mathbf{v}(\mathbf{t}) = \mathbf{0}\ \ \forall \mathbf{t} \in \Omega$, a contradiction. Thus $u_0 > 0$ and

$$J(\mathbf{x}) + \int_\Omega \bar{\mathbf{u}}(\mathbf{t})^T \mathbf{g}(\mathbf{x}(\mathbf{t}); \mathbf{t})\ dt + \int_\Omega \bar{\mathbf{v}}(\mathbf{t})^T \mathbf{h}(\mathbf{x}(\mathbf{t}); \mathbf{t})\ dt \geq J(\bar{\mathbf{x}})\ \ \forall \mathbf{x} \in W$$

where $\bar{\mathbf{u}}(\mathbf{t}) = \mathbf{u}(\mathbf{t})/u_0$ and $\bar{\mathbf{v}}(\mathbf{t}) = \mathbf{v}(\mathbf{t})/u_0$. It must also be

$$\min_{\mathbf{x}} \left\{ J(\mathbf{x}) + \int_\Omega \bar{\mathbf{u}}(\mathbf{t})^T \mathbf{g}(\mathbf{x}(\mathbf{t}); \mathbf{t})\ dt + \int_\Omega \bar{\mathbf{v}}(\mathbf{t})^T \mathbf{h}(\mathbf{x}(\mathbf{t}); \mathbf{t})\ dt,\ \ \mathbf{x} \in W \right\} \geq J(\bar{\mathbf{x}})$$

or equivalently, $\Delta(\bar{\mathbf{u}}, \bar{\mathbf{v}}) \geq J(\bar{\mathbf{x}})$. However, from Theorem 3.5, $\Delta(\mathbf{u}, \mathbf{v}) \leq J(\mathbf{x})\ \forall (\mathbf{u}, \mathbf{v}),\ \mathbf{u} \geq \mathbf{0}$ and $\forall \mathbf{x} \in X$. Therefore, $\Delta(\bar{\mathbf{u}}, \bar{\mathbf{v}}) = J(\bar{\mathbf{x}})$ and $(\bar{\mathbf{u}}, \bar{\mathbf{v}})$ with $\bar{\mathbf{u}} \geq \mathbf{0}$ is an optimal solution to IDP (3.6). □

In the next section relationships between mp-NLP (2.5) and IPP (3.1) are examined.

4 Equivalence between mp-NLP and Integral mp-NLP

Both mp-NLP (2.5) and Integral mp-NLP (3.1) are now considered jointly, which leads to the main result of strong duality for (2.5). These two problems have the same feasible set with the same properties and are shown to be equivalent in the sense that they possess the same optimal solutions. The equivalence is shown at three levels: between the primal problems, relaxed primal problems, and dual problems. The equivalence between the primal problems is established in Theorem 4.1.

Theorem 4.1. *Consider mp-NLP* (2.5) *and its integral counterpart* (3.1) *and their respective assumptions. A function* $\mathbf{x}^* \in X$ *is an optimal solution to* (2.5) *if and only if* $\mathbf{x}^*$ *is an optimal solution to* (3.1).

Proof. Let $\mathbf{x}^* \in X$ be an optimal solution to (2.5). Then $\mathbf{x}^*$ is feasible to (2.5) and $f(\mathbf{x}^*(\mathbf{t});\mathbf{t}) \leq f(\mathbf{x}(\mathbf{t});\mathbf{t}) \;\; \forall \mathbf{t} \in \Omega \;\; \forall \mathbf{x} \in X$. This implies $\int_\Omega f(\mathbf{x}^*(\mathbf{t});\mathbf{t})\,dt \leq \int_\Omega f(\mathbf{x}(\mathbf{t});\mathbf{t})\,dt$ $\forall \mathbf{x} \in X$, or equivalently, $J(\mathbf{x}^*) \leq J(\mathbf{x}) \;\; \forall \mathbf{x} \in X$. Thus $\mathbf{x}^*$ is an optimal solution to (3.1).

Conversely, let $\mathbf{x}^*$ be an optimal solution to (3.1) and assume $\mathbf{x}^*$ is not an optimal solution to (2.5). Then, by Definition 2.7, there exist a function $\bar{\mathbf{x}} \in X, \bar{\mathbf{x}} \neq \mathbf{x}^*$ and $\bar{\mathbf{t}} \in \Omega$ such that $f(\bar{\mathbf{x}}(\bar{\mathbf{t}});\bar{\mathbf{t}}) < f(\mathbf{x}^*(\bar{\mathbf{t}});\bar{\mathbf{t}})$. Also, by Assumption 2.4, there exists a continuous and bounded function $\widetilde{\mathbf{x}}$ in a neighborhood of $\bar{\mathbf{t}}$, $N_\delta(\bar{\mathbf{t}})$, such that $\widetilde{\mathbf{x}}(\bar{\mathbf{t}}) = \bar{\mathbf{x}}(\bar{\mathbf{t}})$ and $\widetilde{\mathbf{x}}(\mathbf{t}) \in \mathcal{X}(\mathbf{t})$ $\forall \mathbf{t} \in N_\delta(\bar{\mathbf{t}})$. Since $\mathbf{x}^* \in \bar{\mathbb{C}}_p(\Omega, \mathbb{R}^n)$ by (2.4), there exists a neighborhood of $\bar{\mathbf{t}}$, $\widetilde{N}_\delta(\bar{\mathbf{t}})$, where both $\mathbf{x}^*$ and $\widetilde{\mathbf{x}}$ are continuous. Then the continuity of f in both arguments yields

$$f(\widetilde{\mathbf{x}}(\mathbf{t});\mathbf{t}) < f(\mathbf{x}^*(\mathbf{t});\mathbf{t}) \;\; \forall \mathbf{t} \in \widetilde{N}_\delta(\bar{\mathbf{t}}). \tag{4.1}$$

Construct a function $\hat{\mathbf{x}} : \Omega \to \mathbb{R}^n$ in the following way:

$$\hat{\mathbf{x}}(\mathbf{t}) = \begin{cases} \widetilde{\mathbf{x}}(\mathbf{t}) & \forall \;\; \mathbf{t} \in \widetilde{N}_\delta(\bar{\mathbf{t}}) \\ \mathbf{x}^*(\mathbf{t}) & \text{otherwise} \end{cases} \tag{4.2}$$

Then $\hat{\mathbf{x}} \in \bar{\mathbb{C}}_p(\Omega, \mathbb{R}^n)$ and therefore $\hat{\mathbf{x}} \in L^2(\Omega, \mathbb{R}^n)$, and also $(\hat{\mathbf{x}}(\mathbf{t});\mathbf{t}) \in \mathcal{X}(\mathbf{t}) \times \Omega \;\; \forall \mathbf{t} \in \Omega$, and therefore $\hat{\mathbf{x}} \in X$. Calculate $J(\hat{\mathbf{x}}) - J(\mathbf{x}^*) = \int_\Omega f(\hat{\mathbf{x}}(\mathbf{t});\mathbf{t})\,dt - \int_\Omega f(\mathbf{x}^*(\mathbf{t});\mathbf{t})\,dt = \int_{\widetilde{N}_\delta(\bar{\mathbf{t}})} f(\widetilde{\mathbf{x}}(\mathbf{t});\mathbf{t})\,dt - \int_{\widetilde{N}_\delta(\bar{\mathbf{t}})} f(\mathbf{x}^*(\mathbf{t});\mathbf{t})\,dt < 0$, where the inequality results from (4.1). This is a contradiction that $\mathbf{x}^*$ is optimal to (3.1). Thus $\mathbf{x}^*$ is an optimal solution to (2.1). ☐

The equivalence between RPP (2.9) and RIPP (3.4) is established in Theorem 4.2.

Theorem 4.2. *Consider RPP* (2.9) *and RIPP* (3.4) *and their respective assumptions. Let* $(\mathbf{u}(\mathbf{t}), \mathbf{v}(\mathbf{t}))$ *with* $\mathbf{u}(\mathbf{t}) \geq \mathbf{0}$ *be fixed* $\forall \mathbf{t} \in \Omega$*, i.e.,* $(\mathbf{u}(\mathbf{t}), \mathbf{v}(\mathbf{t})) = (\hat{\mathbf{u}}(\mathbf{t}), \hat{\mathbf{v}}(\mathbf{t})) \;\; \forall \mathbf{t} \in \Omega$*, and let* $(\mathbf{x}, \hat{\mathbf{u}}, \hat{\mathbf{v}}) \in \bar{\mathbb{C}}_p(\Omega, \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^l) \;\; \forall \mathbf{x} \in W$*. A function* $\mathbf{x}^* \in W$ *is an optimal solution to* (2.9) *if and only if* $\mathbf{x}^*$ *is an optimal solution to* (3.4).

Proof. Let $\mathbf{x}^* \in W$ be an optimal solution to RPP (2.9). $\forall \mathbf{t} \in \Omega$ we have:

$$\begin{aligned} \min_{\mathbf{x} \in W} \; \mathcal{L}(\mathbf{x}(\mathbf{t}), \hat{\mathbf{u}}(\mathbf{t}), \hat{\mathbf{v}}(\mathbf{t}); \mathbf{t}) = \; & \mathcal{L}(\mathbf{x}^*(\mathbf{t}), \hat{\mathbf{u}}(\mathbf{t}), \hat{\mathbf{v}}(\mathbf{t}); \mathbf{t}) \\ = \; & f(\mathbf{x}^*(\mathbf{t});\mathbf{t}) + \hat{\mathbf{u}}(\mathbf{t})^T \mathbf{g}(\mathbf{x}^*(\mathbf{t});\mathbf{t}) + \hat{\mathbf{v}}(\mathbf{t})^T \mathbf{h}(\mathbf{x}^*(\mathbf{t});\mathbf{t}) \\ \leq \; & f(\mathbf{x}(\mathbf{t});\mathbf{t}) + \hat{\mathbf{u}}(\mathbf{t})^T \mathbf{g}(\mathbf{x}(\mathbf{t});\mathbf{t}) + \hat{\mathbf{v}}(\mathbf{t})^T \mathbf{h}(\mathbf{x}(\mathbf{t});\mathbf{t}) \;\; \forall \mathbf{x} \in W \end{aligned}$$

Applying the integral $\int_\Omega dt$ to both sides of the inequality yields:

$$\mathcal{IL}(\mathbf{x}^*, \hat{\mathbf{u}}, \hat{\mathbf{v}}) \leq \mathcal{IL}(\mathbf{x}, \hat{\mathbf{u}}, \hat{\mathbf{v}}) \;\; \forall \mathbf{x} \in W$$

and thus $\mathbf{x}^*$ is an optimal solution to RIPP (3.4).

Conversely, let $\mathbf{x}^*$ be an optimal solution to RIPP (3.4), and assume $\mathbf{x}^*$ is not an optimal solution to RPP (2.9). By Definition 2.11, there exist a function $\bar{\mathbf{x}} \in W, \bar{\mathbf{x}} \neq \mathbf{x}^*$, and $\bar{\mathbf{t}} \in \Omega$ such that $\mathcal{L}(\bar{\mathbf{x}}(\bar{\mathbf{t}}), \hat{\mathbf{u}}(\bar{\mathbf{t}}), \hat{\mathbf{v}}(\bar{\mathbf{t}}); \bar{\mathbf{t}}) < \mathcal{L}(\mathbf{x}^*(\bar{\mathbf{t}}), \hat{\mathbf{u}}(\bar{\mathbf{t}}), \hat{\mathbf{v}}(\bar{\mathbf{t}}); \bar{\mathbf{t}})$. By Assumption 2.5, there exists a function $\widetilde{\mathbf{x}} \in W$ that is continuous and bounded in a neighborhood of $\bar{\mathbf{t}}$, $N_\delta(\bar{\mathbf{t}})$, such that $\widetilde{\mathbf{x}}(\bar{\mathbf{t}}) = \bar{\mathbf{x}}(\bar{\mathbf{t}})$. Since $(\mathbf{x}^*, \hat{\mathbf{u}}, \hat{\mathbf{v}}) \in \bar{\mathbb{C}}_p(\Omega, \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^l)$ by assumption, there exists a neighborhood of $\bar{\mathbf{t}}$, $\widetilde{N}_\delta(\bar{\mathbf{t}})$, where both $(\mathbf{x}^*, \hat{\mathbf{u}}, \hat{\mathbf{v}})$ and $(\widetilde{\mathbf{x}}, \hat{\mathbf{u}}, \hat{\mathbf{v}})$ are continuous. Then the inequality

$$\mathcal{L}(\widetilde{\mathbf{x}}(\mathbf{t}), \hat{\mathbf{u}}(\mathbf{t}), \hat{\mathbf{v}}(\mathbf{t}); \mathbf{t}) < \mathcal{L}(\mathbf{x}^*(\mathbf{t}), \hat{\mathbf{u}}(\mathbf{t}), \hat{\mathbf{v}}(\mathbf{t}); \mathbf{t}) \ \forall \mathbf{t} \in \widetilde{N}_\delta(\bar{\mathbf{t}}) \tag{4.3}$$

holds as all f, $\mathbf{g}$ and $\mathbf{h}$ are continuous in both arguments. Applying the integral $\int_{\widetilde{N}_\delta(\bar{\mathbf{t}})} dt$ to both sides of this inequality yields:

$$\int_{\widetilde{N}_\delta(\widetilde{\mathbf{t}})} \mathcal{L}(\widetilde{\mathbf{x}}(\mathbf{t}), \hat{\mathbf{u}}(\mathbf{t}), \hat{\mathbf{v}}(\mathbf{t}); \mathbf{t}) \, dt < \int_{\widetilde{N}_\delta(\bar{\mathbf{t}})} \mathcal{L}(\mathbf{x}^*(\mathbf{t}), \hat{\mathbf{u}}(\mathbf{t}), \hat{\mathbf{v}}(\mathbf{t}); \mathbf{t}) \, dt \tag{4.4}$$

Constructing a function $\hat{\mathbf{x}} : \Omega \to \mathbb{R}^n$ as in (4.2) yields $\hat{\mathbf{x}} \in W$ and $(\hat{\mathbf{x}}, \hat{\mathbf{u}}, \hat{\mathbf{v}}) \in \bar{\mathbb{C}}_p(\Omega, \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^l)$. A similar calculation to that in the proof of Theorem 4.1 gives

$$\mathcal{IL}(\hat{\mathbf{x}}, \hat{\mathbf{u}}, \hat{\mathbf{v}}) - \mathcal{IL}(\mathbf{x}^*, \hat{\mathbf{u}}, \hat{\mathbf{v}}) = \underset{<0}{\int_{\widetilde{N}_\delta(\bar{\mathbf{t}})}} \mathcal{L}(\widetilde{\mathbf{x}}(\mathbf{t}), \hat{\mathbf{u}}(\mathbf{t}), \hat{\mathbf{v}}(\mathbf{t}); \mathbf{t}) \, dt - \int_{\widetilde{N}_\delta(\bar{\mathbf{t}})} \mathcal{L}(\mathbf{x}^*(\mathbf{t}), \hat{\mathbf{u}}(\mathbf{t}), \hat{\mathbf{v}}(\mathbf{t}); \mathbf{t}) \, dt$$

where the inequality comes from (4.4). This is a contradiction that $\mathbf{x}^*$ is an optimal solution to RIPP (3.4). Thus $\mathbf{x}^*$ is an optimal solution to RPP (2.9).

□

Theorem 4.2 results in the following corollary.

Corollary 4.3. *Under the assumptions of Theorem 4.2, the following holds:*

$$\int_\Omega \min_{\mathbf{x} \in W} \mathcal{L}(\mathbf{x}(\mathbf{t}), \hat{\mathbf{u}}(\mathbf{t}), \hat{\mathbf{v}}(\mathbf{t}); \mathbf{t}) \, dt = \min_{\mathbf{x} \in W} \int_\Omega \mathcal{L}(\mathbf{x}(\mathbf{t}), \hat{\mathbf{u}}(\mathbf{t}), \hat{\mathbf{v}}(\mathbf{t}); \mathbf{t}) \, dt \tag{4.5}$$

Proof. Let $(\mathbf{u}(\mathbf{t}), \mathbf{v}(\mathbf{t})) = (\hat{\mathbf{u}}(\mathbf{t}), \hat{\mathbf{v}}(\mathbf{t})) \ \ \forall \mathbf{t} \in \Omega$ and $\mathbf{x}^* \in W$ be an optimal solution to RPP (2.9). By definition, $\min\limits_{\mathbf{x} \in W} \ \mathcal{L}(\mathbf{x}(\mathbf{t}), \hat{\mathbf{u}}(\mathbf{t}), \hat{\mathbf{v}}(\mathbf{t}); \mathbf{t}) = \mathcal{L}(\mathbf{x}^*(\mathbf{t}), \hat{\mathbf{u}}(\mathbf{t}), \hat{\mathbf{v}}(\mathbf{t}); \mathbf{t}) \ \ \forall \mathbf{t} \in \Omega$. Applying the integral $\int_\Omega dt$ to both sides of this equality yields

$$\begin{aligned}
\int_\Omega \min_{\mathbf{x} \in W} \mathcal{L}(\mathbf{x}(\mathbf{t}), \hat{\mathbf{u}}(\mathbf{t}), \hat{\mathbf{v}}(\mathbf{t}); \mathbf{t}) \, dt &= \int_\Omega \mathcal{L}(\mathbf{x}^*(\mathbf{t}), \hat{\mathbf{u}}(\mathbf{t}), \hat{\mathbf{v}}(\mathbf{t}); \mathbf{t}) \, dt \\
&= \mathcal{IL}(\mathbf{x}^*, \hat{\mathbf{u}}, \hat{\mathbf{v}}; \mathbf{t}) \\
&= \min_{\mathbf{x} \in W} \mathcal{IL}(\mathbf{x}, \hat{\mathbf{u}}, \hat{\mathbf{v}}) \\
&= \min_{\mathbf{x} \in W} \int_\Omega \mathcal{L}(\mathbf{x}(\mathbf{t}), \hat{\mathbf{u}}(\mathbf{t}), \hat{\mathbf{v}}(\mathbf{t})) \, dt
\end{aligned}$$

where the third equality holds because $\mathbf{x}^*$, by Theorem 4.2, is an optimal solution to RIPP (3.4). □

Under the assumptions of Theorems 3.7 and 4.2, a relationship is obtained between the optimal value of IDP (3.6) and the value of the integral of the dual function (2.10) for mp-NLP (2.5).

Corollary 4.4. *Let the assumptions of Theorems 3.7 and 4.2 hold and* $(\bar{\mathbf{u}}, \bar{\mathbf{v}})$ *be an optimal solution to IDP* (3.6)*. Then* $\bar{\Delta} = \int_\Omega \theta(\bar{\mathbf{u}}(\mathbf{t}), \bar{\mathbf{v}}(\mathbf{t}), \mathbf{t})\, dt$*.*

Proof. By definition of Δ,

$$\begin{aligned}\bar{\Delta} = \Delta(\bar{\mathbf{u}}, \bar{\mathbf{v}}) &= \min_{\mathbf{x}\in W}\ \mathcal{IL}(\mathbf{x}, \bar{\mathbf{u}}, \bar{\mathbf{v}})\\ &= \min_{\mathbf{x}\in W}\ \int_\Omega \mathcal{L}(\mathbf{x}(\mathbf{t}), \bar{\mathbf{u}}(\mathbf{t}), \bar{\mathbf{v}}(\mathbf{t}); \mathbf{t})\, dt\\ &= \int_\Omega \min_{\mathbf{x}\in W}\ \mathcal{L}(\mathbf{x}(\mathbf{t}), \bar{\mathbf{u}}(\mathbf{t}), \bar{\mathbf{v}}(\mathbf{t}); \mathbf{t})\, dt\\ &= \int_\Omega \theta(\bar{\mathbf{u}}(\mathbf{t}), \bar{\mathbf{v}}(\mathbf{t}); \mathbf{t})\, dt\end{aligned}$$

where the fourth equality follows from Corollary 4.3 and the last equality follows from (2.10). ☐

The equivalence between LDP and IDP can now be proved.

Theorem 4.5. *Consider LDP* (2.11) *and IDP* (3.6) *and their respective assumptions. A feasible pair of functions* $(\mathbf{u}^*, \mathbf{v}^*)$ *is an optimal solution to* (2.11) *if and only if* $(\mathbf{u}^*, \mathbf{v}^*)$ *is an optimal solution to* (3.6)*.*

Proof. Let $(\mathbf{u}^*, \mathbf{v}^*)$ be an optimal solution to LDP (2.11). By Definition 2.12,

$$\theta(\mathbf{u}^*(\mathbf{t}), \mathbf{v}^*(\mathbf{t}); \mathbf{t}) \geq \theta(\bar{\mathbf{u}}(\mathbf{t}), \bar{\mathbf{v}}(\mathbf{t}); \mathbf{t})\ \ \forall(\bar{\mathbf{u}}(\mathbf{t}), \bar{\mathbf{v}}(\mathbf{t})), \bar{\mathbf{u}}(\mathbf{t}) \geq \mathbf{0}\ \ \forall \mathbf{t} \in \Omega$$

or equivalently

$$\min_{\mathbf{x}\in W}\ \mathcal{L}(\mathbf{x}(\mathbf{t}), \mathbf{u}^*(\mathbf{t}), \mathbf{v}^*(\mathbf{t}); \mathbf{t}) \geq \min_{\mathbf{x}\in W}\ \mathcal{L}(\mathbf{x}(\mathbf{t}), \bar{\mathbf{u}}(\mathbf{t}), \bar{\mathbf{v}}(\mathbf{t}); \mathbf{t})\ \ \forall(\bar{\mathbf{u}}(\mathbf{t}), \bar{\mathbf{v}}(\mathbf{t})), \bar{\mathbf{u}}(\mathbf{t}) \geq \mathbf{0}\ \ \forall \mathbf{t} \in \Omega$$

Let $\mathbf{x}^* \in W$ be an optimal solution to RPP with $(\mathbf{u}^*, \mathbf{v}^*)$, and $\bar{\mathbf{x}} \in W$ be an optimal solution to RPP with $(\bar{\mathbf{u}}, \bar{\mathbf{v}})$. Then we obtain

$$\mathcal{L}(\mathbf{x}^*(\mathbf{t}), \mathbf{u}^*(\mathbf{t}), \mathbf{v}^*(\mathbf{t}); \mathbf{t}) \geq \mathcal{L}(\bar{\mathbf{x}}(\mathbf{t}), \bar{\mathbf{u}}(\mathbf{t}), \bar{\mathbf{v}}(\mathbf{t}); \mathbf{t})\ \ \forall(\bar{\mathbf{u}}(\mathbf{t}), \bar{\mathbf{v}}(\mathbf{t})), \bar{\mathbf{u}}(\mathbf{t}) \geq \mathbf{0}\ \ \forall \mathbf{t} \in \Omega \tag{4.6}$$

Applying the integral $\int_\Omega dt$ to both sides of (4.6) yields

$$\mathcal{IL}(\mathbf{x}^*, \mathbf{u}^*, \mathbf{v}^*) \geq \mathcal{IL}(\bar{\mathbf{x}}, \bar{\mathbf{u}}, \bar{\mathbf{v}}) \quad \forall(\bar{\mathbf{u}}, \bar{\mathbf{v}}), \bar{\mathbf{u}} \geq \mathbf{0}$$

By Theorem 4.2, $\mathbf{x}^* \in W$ is also an optimal solution to RIPP with $(\mathbf{u}^*, \mathbf{v}^*)$, and $\bar{\mathbf{x}} \in W$ is also an optimal solution to RIPP with $(\bar{\mathbf{u}}, \bar{\mathbf{v}})$, therefore

$$\min_{\mathbf{x}\in W}\ \mathcal{IL}(\mathbf{x}, \mathbf{u}^*, \mathbf{v}^*) \geq \min_{\mathbf{x}\in W}\ \mathcal{IL}(\mathbf{x}, \bar{\mathbf{u}}, \bar{\mathbf{v}})\ \ \forall(\bar{\mathbf{u}}, \bar{\mathbf{v}}), \bar{\mathbf{u}} \geq \mathbf{0}$$

or equivalently,

$$\Delta(\mathbf{u}^*, \mathbf{v}^*) \geq \Delta(\bar{\mathbf{u}}, \bar{\mathbf{v}}) \quad \forall(\bar{\mathbf{u}}, \bar{\mathbf{v}}), \bar{\mathbf{u}} \geq \mathbf{0}$$

and thus $(\mathbf{u}^*, \mathbf{v}^*)$ is an optimal solution to IDP (3.6).

Conversely, let $(\mathbf{u}^*, \mathbf{v}^*)$ be an optimal solution to IDP (3.6) and assume that $(\mathbf{u}^*, \mathbf{v}^*)$ is not an optimal solution to LDP (2.11). By Definition 2.12, there exist a pair of functions $(\bar{\mathbf{u}}, \bar{\mathbf{v}}) \in \bar{\mathbb{C}}_p(\Omega, \mathbb{R}^m \times \mathbb{R}^l)$, $\bar{\mathbf{u}} \geq \mathbf{0}, (\bar{\mathbf{u}}, \bar{\mathbf{v}}) \neq (\mathbf{u}^*, \mathbf{v}^*)$, and $\bar{\mathbf{t}} \in \Omega$ such that $\theta(\bar{\mathbf{u}}(\bar{\mathbf{t}}), \bar{\mathbf{v}}(\bar{\mathbf{t}}); \bar{\mathbf{t}}) > \theta(\mathbf{u}^*(\bar{\mathbf{t}}), \mathbf{v}^*(\bar{\mathbf{t}}); \bar{\mathbf{t}})$, or equivalently,

$$\mathcal{L}(\bar{\mathbf{x}}(\bar{\mathbf{t}}), \bar{\mathbf{u}}(\bar{\mathbf{t}}), \bar{\mathbf{v}}(\bar{\mathbf{t}}); \bar{\mathbf{t}}) > \mathcal{L}(\mathbf{x}^*(\bar{\mathbf{t}}), \mathbf{u}^*(\bar{\mathbf{t}}), \mathbf{v}^*(\bar{\mathbf{t}}); \bar{\mathbf{t}})$$

where $\bar{\mathbf{x}} \in W$ is an optimal solution to RPP with $(\bar{\mathbf{u}}, \bar{\mathbf{v}})$ and, as assumed above, $\mathbf{x}^* \in W$ is an optimal solution to RPP with $(\mathbf{u}^*, \mathbf{v}^*)$. By Assumption 2.7, there exists a pair of continuous and bounded functions $(\widetilde{\mathbf{u}}, \widetilde{\mathbf{v}})$ in a neighborhood of $\bar{\mathbf{t}}$, $N_\delta(\bar{\mathbf{t}})$, such that $(\widetilde{\mathbf{u}}(\bar{\mathbf{t}}), \widetilde{\mathbf{v}}(\bar{\mathbf{t}})) = (\bar{\mathbf{u}}(\bar{\mathbf{t}}), \bar{\mathbf{v}}(\bar{\mathbf{t}}))$ and $\widetilde{\mathbf{u}}(\mathbf{t}) \geq \mathbf{0} \ \ \forall \mathbf{t} \in N_\delta(\bar{\mathbf{t}})$. Let $\widetilde{\mathbf{x}} \in W$ be an optimal solution to RPP with $(\widetilde{\mathbf{u}}, \widetilde{\mathbf{v}})$. Since $(\mathbf{u}^*, \mathbf{v}^*) \in \bar{\mathbb{C}}_p(\Omega, \mathbb{R}^m \times \mathbb{R}^l)$, there exists a neighborhood of $\bar{\mathbf{t}}$, $\widetilde{N}_\delta(\bar{\mathbf{t}})$, where both $(\mathbf{u}^*, \mathbf{v}^*)$ and $(\widetilde{\mathbf{u}}, \widetilde{\mathbf{v}})$ are continuous. Then, by Assumption 2.6, the functions $\mathbf{x}^*$ and $\widetilde{\mathbf{x}}$ are also continuous on $\widetilde{N}_\delta(\bar{\mathbf{t}})$, and therefore

$$\mathcal{L}(\widetilde{\mathbf{x}}(\mathbf{t}), \widetilde{\mathbf{u}}(\mathbf{t}), \widetilde{\mathbf{v}}(\mathbf{t}); \mathbf{t}) > \mathcal{L}(\mathbf{x}^*(\mathbf{t}), \mathbf{u}^*(\mathbf{t}), \mathbf{v}^*(\mathbf{t}); \mathbf{t}) \ \forall \mathbf{t} \in \widetilde{N}_\delta(\bar{\mathbf{t}})$$

Applying the integral $\int_{N_\delta(\bar{\mathbf{t}})} dt$ to both sides of this inequality yields:

$$\int_{\widetilde{N}_\delta(\bar{\mathbf{t}})} \mathcal{L}(\widetilde{\mathbf{x}}(\mathbf{t}), \widetilde{\mathbf{u}}(\mathbf{t}), \widetilde{\mathbf{v}}(\mathbf{t}); \mathbf{t}) \, dt > \int_{\widetilde{N}_\delta(\bar{\mathbf{t}})} \mathcal{L}(\mathbf{x}^*(\mathbf{t}), \mathbf{u}^*(\mathbf{t}), \mathbf{v}^*(\mathbf{t}); \mathbf{t}) \, dt \tag{4.7}$$

Construct a pair of functions:

$$(\hat{\mathbf{u}}(\mathbf{t}), \hat{\mathbf{v}}(\mathbf{t})) = \begin{cases} (\widetilde{\mathbf{u}}(\mathbf{t}), \widetilde{\mathbf{v}}(\mathbf{t})) & \mathbf{t} \in \widetilde{N}_\delta(\bar{\mathbf{t}}) \\ (\mathbf{u}^*(\mathbf{t}), \mathbf{u}^*(\mathbf{t})) & \text{otherwise} \end{cases}$$

Then $\hat{\mathbf{u}} \in L^2(\Omega, \mathbb{R}^m)$, $\hat{\mathbf{u}} \geq \mathbf{0}$, and $\hat{\mathbf{v}} \in L^2(\Omega, \mathbb{R}^l)$, and therefore $(\hat{\mathbf{u}}, \hat{\mathbf{v}})$ is feasible for IDP (3.6). Calculate

$$\begin{aligned}
\Delta(\hat{\mathbf{u}}, \hat{\mathbf{v}}) - \Delta(\mathbf{u}^*, \mathbf{v}^*) &= \min_{\mathbf{x} \in W} \ \mathcal{IL}(\mathbf{x}, \hat{\mathbf{u}}, \hat{\mathbf{v}}) - \min_{\mathbf{x} \in W} \ \mathcal{IL}(\mathbf{x}, \mathbf{u}^*, \mathbf{v}^*) \\
&= \min_{\mathbf{x} \in W} \int_\Omega \mathcal{L}(\mathbf{x}(\mathbf{t}), \hat{\mathbf{u}}(\mathbf{t}), \hat{\mathbf{v}}(\mathbf{t}); \mathbf{t}) \, dt - \min_{\mathbf{x} \in W} \int_\Omega \mathcal{L}(\mathbf{x}(\mathbf{t}), \mathbf{u}^*(\mathbf{t}), \mathbf{v}^*(\mathbf{t}); \mathbf{t}) \, dt \\
&= \int_\Omega \min_{\mathbf{x} \in W} \mathcal{L}(\mathbf{x}(\mathbf{t}), \hat{\mathbf{u}}(\mathbf{t}), \hat{\mathbf{v}}(\mathbf{t}); \mathbf{t}) \, dt - \int_\Omega \min_{\mathbf{x} \in W} \mathcal{L}(\mathbf{x}(\mathbf{t}), \mathbf{u}^*(\mathbf{t}), \mathbf{v}^*(\mathbf{t}); \mathbf{t}) \, dt \\
&= \int_{\widetilde{N}_\delta(\bar{\mathbf{t}})} \mathcal{L}(\widetilde{\mathbf{x}}(\mathbf{t}), \widetilde{\mathbf{u}}(\mathbf{t}), \widetilde{\mathbf{v}}(\mathbf{t}); \mathbf{t}) \, dt - \int_{\widetilde{N}_\delta(\bar{\mathbf{t}})} \mathcal{L}(\mathbf{x}^*(\mathbf{t}), \mathbf{u}^*(\mathbf{t}), \mathbf{v}^*(\mathbf{t}); \mathbf{t}) \, dt \\
&> 0
\end{aligned}$$

The third equality results from Corollary 4.3 and the strict inequality is a consequence of (4.7). This is a contradiction that $(\mathbf{u}^*, \mathbf{v}^*)$ is an optimal solution to IDP (3.6). Thus $(\mathbf{u}^*, \mathbf{v}^*)$ is an optimal solution to LDP (2.11). □

The main theoretical result, that extends strong duality to mp-NLPs, can now be proved.

Theorem 4.6 (Strong Duality for mp-NLP). *Consider mp-NLP* (2.5) *and LDP* (2.11) *and their respective assumptions. Let the assumptions of Theorems 3.7 and 4.2 also hold. If* $\mathbf{x}^* \in X$ *is an optimal solution to* (2.5) *and* $(\mathbf{u}^*, \mathbf{v}^*), \mathbf{u}^* \geq \mathbf{0}$ *is an optimal solution to* (2.11), *then*

$$\begin{aligned}
f(\mathbf{x}^*(\mathbf{t}), \mathbf{t}) &= \min_{\mathbf{x}} \{ f(\mathbf{x}(\mathbf{t}); \mathbf{t}) : \mathbf{x} \in S, \mathbf{g}(\mathbf{x}(\mathbf{t}); \mathbf{t}) \leq \mathbf{0}, \mathbf{h}(\mathbf{x}(\mathbf{t}); \mathbf{t}) = \mathbf{0} \} \\
&= \max_{\mathbf{u}, \mathbf{v}} \{ \theta(\mathbf{u}(\mathbf{t}), \mathbf{v}(\mathbf{t}); \mathbf{t}) : \mathbf{u}(\mathbf{t}) \geq \mathbf{0} \} = \theta(\mathbf{u}^*(\mathbf{t}), \mathbf{v}^*(\mathbf{t}), \mathbf{t}) \ \ \forall \mathbf{t} \in \Omega
\end{aligned}$$

Proof. Let $\mathbf{x}^* \in X$ be an optimal solution to (2.5) and $(\mathbf{u}^*, \mathbf{v}^*), \mathbf{u}^* \geq \mathbf{0}$ be an optimal solution to (2.11). From Theorem 2.13

$$\theta(\mathbf{u}^*(\mathbf{t}), \mathbf{v}^*(\mathbf{t}); \mathbf{t}) \leq f(\mathbf{x}^*(\mathbf{t}); \mathbf{t}) \ \ \forall \mathbf{t} \in \Omega \tag{4.8}$$

By Theorems 4.1 and 4.5, $\mathbf{x}^*$ is also an optimal solution to IPP (3.1) and $(\mathbf{u}^*, \mathbf{v}^*)$ is also an optimal solution to IDP (3.6), and therefore Theorem 3.7 holds yielding

$$\int_\Omega f(\mathbf{x}^*(\mathbf{t}); \mathbf{t})\, dt = \Delta(\mathbf{u}^*, \mathbf{v}^*) \tag{4.9}$$

On the other hand, from Corollary 4.4,

$$\Delta(\mathbf{u}^*, \mathbf{v}^*) = \int_\Omega \theta(\mathbf{u}^*(\mathbf{t}), \mathbf{v}^*(\mathbf{t}); \mathbf{t})\, dt \tag{4.10}$$

From (4.9) and (4.10):

$$\int_\Omega f(\mathbf{x}^*(\mathbf{t}); \mathbf{t})\, dt = \int_\Omega \theta(\mathbf{u}^*(\mathbf{t}), \mathbf{v}^*(\mathbf{t}); \mathbf{t})\, dt \tag{4.11}$$

and from (4.8) and (4.11):

$$f(\mathbf{x}^*(\mathbf{t}); \mathbf{t}) = \theta(\mathbf{u}^*(\mathbf{t}), \mathbf{v}^*(\mathbf{t}); \mathbf{t}) \quad \forall \mathbf{t} \in \Omega \tag{4.12}$$

as desired. □

With the theoretical results established, solving the parametric primal and dual problems is now of interest.

5 Subgradient Algorithm

The presented theoretical results open up a possibility to solve the parametric primal and dual problems, mp-NLP (2.5) and LDP (2.11). Based on Theorems 4.1, 4.2, 4.5 in Section 4 presenting the three levels of equivalence between mp-NLP (2.5), RPP (2.9), LDP (2.11) and their integral counterparts, IPP (3.1), RIPP (3.4), IDP (3.6), an optimal solution to a problem in one group is also an optimal solution to the respective problem in the other group and vice versa. Given this equivalence, solving the integral counterparts can be a way to solve the parametric problems. More importantly, based on the strong duality result in Theorem 4.6, solving the parametric problems could be achieved jointly in a primal-dual algorithm. However, to benefit from these relationships, the first task is to be able to solve IPP (3.1).

Note that IPP (3.1) belongs to the class of optimization problems to minimize a convex functional on a Hilbert space for which solution algorithms have been developed. Since solving primal an dual problems is of interest in this study, subgradient optimization becomes an appropriate choice. In the literature, a subgradient algorithm to minimize a convex functional on a Hilbert space has been designed in [3] and, if appropriately implemented, it can immediately be applied to solving IPP (3.1). In this case, given the functional $J : L^2(\Omega, \mathbb{R}^n) \to \mathbb{R}$ as the objective function, the subdifferential of J at $\mathbf{x}$ is the set $\partial J(\mathbf{x})$ defined as

$$\partial J(\mathbf{x}) = \left\{ \boldsymbol{\xi} \in L^2(\Omega, \mathbb{R}^n) : J(\mathbf{z}) \geq J(\mathbf{x}) + \langle \boldsymbol{\xi}, \mathbf{z} - \mathbf{x} \rangle \quad \forall \mathbf{z} \in L^2(\Omega, \mathbb{R}^n) \right\} \tag{5.1}$$

where $\boldsymbol{\xi}$ is a subgradient of J at $\mathbf{x}$. In the algorithm, a subgradient-based direction is used as a direction for improving the values of J while the feasibility of each iterate is assured by projecting the new iterate on the feasible set X. The main convergence result addresses the case when the optimal solution set X^* is nonempty. Then the algorithm stops at iteration K, i.e., $\mathbf{x}^K \in X^*$, or it generates an infinite sequence which converges weakly to some $\bar{\mathbf{x}} \in X^*$. The reader is referred to [3] for the details and a convergence analysis.

Due to the equivalence discussed above, this algorithm can be adapted to directly solving mp-NLP (2.5). Additionally, after some modifications, it is also suitable for solving IDP (3.6) and therefore can also be adapted for (2.11). Since a literature review indicates that this algorithm has not been implemented while it serves the needs of this work, it has been implemented as a multiparametric subgradient algorithm (mp-SA) in two versions: Algorithm 1 to solve mp-NLP (2.5) and Algorithm 2 to solve LDP (2.11). Both algorithms are presented below in the context of the parametric problems they solve.

Algorithm 1 is implemented to solve mp-NLP (2.5) and its pseudocode is given below. Let P_X denote the projection operator onto the feasible set X so that $\mathbf{x}' = P_X(\hat{\mathbf{x}})$ is a minimizer of $\min\limits_{\mathbf{x}\in X} \|\mathbf{x} - \hat{\mathbf{x}}\|_2$. Let $\partial f(\mathbf{x})$ denote the subdifferential of the objective function in (2.5) at $\mathbf{x}$. At each iteration, in Line 1, a subgradient of the objective function is obtained at the current solution. If the subgradient is nonzero, in Lines 5 and 6, the next solution is computed by moving a prescribed step size λ_k along the direction $\mathbf{d}^k$ opposite to that of the normalized subgradient. This new solution in $L^2(\Omega, \mathbb{R}^n)$ is projected onto the feasible set X if necessary. In Line 7, the objective function is evaluated at the new solution. In Line 9, the current upper bound is updated with a smaller one if such has been found in Line 8. The process continues a prescribed number of times or until a convergence criterion is met (with K denoting the final iteration).

Algorithm 2 is implemented to solve LDP (2.11) and its pseudocode is also given below. Since this is a maximization problem, Algorithm 2 solves the LDP-equivalent problem of the form $\min_{\mathbf{u},\mathbf{v}} \; -\theta(\mathbf{u}(\mathbf{t}), \mathbf{v}(\mathbf{t}); \mathbf{t})$, $\mathbf{u}(\mathbf{t}) \geq \mathbf{0}, \mathbf{t} \in \Omega$, and provides primal and dual optimal solutions to mp-NLP (2.5). Let $P_{\geq}$ denote the projection operator onto the set of non-negative vectors in $\mathbb{R}^n$, and $\partial\theta(\mathbf{u}, \mathbf{v})$ denote the subdifferential of the objective function in (2.11) at $(\mathbf{u}, \mathbf{v})$. In Line 1, given the current dual solution $\left(\mathbf{u}^k, \mathbf{v}^k\right)$, RPP (2.9) is solved for a primal solution $\mathbf{x}^k$ and a subgradient of the dual function is obtained by evaluating $\boldsymbol{\xi}^k(\mathbf{t}) = -\left(\mathbf{g}\left(\mathbf{x}^k(\mathbf{t}); \mathbf{t}\right), \mathbf{h}\left(\mathbf{x}^k(\mathbf{t}); \mathbf{t}\right)\right) \quad \forall \mathbf{t} \in \Omega$. The main step proceeds in a similar way to Algorithm 1 with the difference that $-\theta(\mathbf{u}(\mathbf{t})$ is minimized with respect to $\left(\mathbf{u}^k, \mathbf{v}^k\right)$. As is the case in the nonparametric setting, evaluating θ to obtain a subgradient requires solving RPP (2.9) to obtain a primal solution used to evaluate $\mathbf{g}$ and $\mathbf{h}$. Since RPP (2.9) is also an mp-NLP, a suitable algorithm available in the literature is employed in Line 1. This is discussed in more detail in Section 6.3.

6 Applications

In Sections 6.1 and 6.2 two examples are presented to demonstrate the effectiveness of the subgradient algorithm mp-SA. Remarks on the numerical experience gained are included in Section 6.3. To support the duality theory developed in Sections 3 and 4, Algorithm 2 is applied to solve LDP (2.11) to each example problem. Each LDP (2.11) is created by relaxing a nonlinear constraint into the objective. For comparison purposes, for each example, mp-NLP (or PP) (2.5) is also solved using the multiparametric quadratic algorithm (mp-QA), a state-of-the-art quadratic approximation algorithm for mpNLPs [26, 12, 13]. The first example contains a single parameter and is used to provide additional insight to the steps at each iteration of Algorithm 2. The second example is a multiparametric extension of the Rosen-Suzuki problem [25].

All problems have been solved in Matlab on a 2.6 GHz Intel Core i7-6700HQ processor with 32 GB of RAM. The implementation of the mp-QA follows [12, 13], where the vertex-based implementation from the latter reference is used, and employs the Multi-Parametric Toolbox [23] to solve the resulting multiparametric quadratic program (mp-QP).

Input: initial primal feasible solution $\mathbf{x}^0$; sequence $\{\lambda_k\}$; initial upper bound UB_0 ($UB_0 = \infty$ is sufficient)

1 **Step k:** Given $\mathbf{x}^k$, obtain subgradient $\boldsymbol{\xi}^k \in \partial f(\mathbf{x}^k)$;
2 **if** $\boldsymbol{\xi}^k = \mathbf{0}$ **then**
3   stop ;
4 **else**
5   set $\mathbf{d}^k = -\frac{\boldsymbol{\xi}^k}{\|\boldsymbol{\xi}^k\|_2}$ and select step size λ_k ;
6   compute $\mathbf{x}^{k+1} = P_X(\hat{\mathbf{x}}^{k+1})$ where $\hat{\mathbf{x}}^{k+1} = \mathbf{x}^k + \lambda_k \mathbf{d}^k$;
7   evaluate $f_{k+1}(\mathbf{t}) = f(\mathbf{x}^{k+1}(\mathbf{t}); \mathbf{t})\ \forall \mathbf{t} \in \Omega$;
8   **if** $\int_\Omega f_{k+1}(\mathbf{t})\, dt < UB_k$ **then**
9   update $UB_{k+1} = \int_\Omega f_{k+1}(\mathbf{t})\, dt$;
10   **else**
11         $UB_{k+1} = UB_k$;
12   **end**
13   Set $k = k + 1$ and go to Line 1 ;
14 **end**

Output: primal solution $\mathbf{x}^K$, value function f_K, and upper bound UB_K on the integral of f_K

Algorithm 1: Subgradient algorithm mp-SA for mp-NLP (2.5)

Input: initial dual feasible solution $(\mathbf{u}^0, \mathbf{v}^0)$; sequence $\{\lambda_k\}$, initial upper bound UB_0 ($UB_0 = \infty$ is sufficient)

1 **Step k:** Given $(\mathbf{u}^k, \mathbf{v}^k)$, obtain primal solution $\mathbf{x}^k$ by evaluating $\theta(\mathbf{u}^k, \mathbf{v}^k; \mathbf{t})$, and obtain subgradient $\boldsymbol{\xi}^k \in \partial\theta(\mathbf{u}^k, \mathbf{v}^k)$;
2 **if** $\boldsymbol{\xi}^k = \mathbf{0}$ **then**
3   stop ;
4 **else**
5   set $\mathbf{d}^k = -\frac{\boldsymbol{\xi}^k}{\|\boldsymbol{\xi}^k\|_2}$ and select step size λ_k ;
6   compute $\mathbf{u}^{k+1} = P_{\geq \mathbf{0}}(\hat{\mathbf{u}}^{k+1})$ where $\hat{\mathbf{u}}^{k+1} = \mathbf{u}^k + \lambda_k \mathbf{d}^k$;
7   compute $\mathbf{v}^{k+1} = \mathbf{v}^k + \lambda_k \mathbf{d}^k$;
8   evaluate $\theta_{k+1}(\mathbf{t}) = \theta(\mathbf{u}^{k+1}(\mathbf{t}), \mathbf{v}^{k+1}(\mathbf{t}); \mathbf{t})\ \ \forall \mathbf{t} \in \Omega$;
9   **if** $\int_\Omega -\theta_{k+1}(\mathbf{t})\, dt < UB_k$ **then**
10   update $UB_{k+1} = \int_\Omega -\theta_{k+1}(\mathbf{t})\, dt$;
11   **else**
12         $UB_{k+1} = UB_k$;
13   **end**
14   Set $k = k + 1$ and go to Line 1;
15 **end**

Output: dual solution $(\mathbf{u}^K, \mathbf{v}^K)$, primal solution $\mathbf{x}^K$, value function $-\theta_K$, and upper bound UB_K on the integral of $-\theta_K$

Algorithm 2: Subgradient algorithm mp-SA for LDP (2.11)

6.1 Single-parametric example

The first example has a quadratic constraint but is otherwise linear.

$$\begin{aligned} \min_{x_1,x_2} \quad & f(\mathbf{x}(t);t) = -x_1(t) - x_2(t) \\ s.t. \quad & g_1(\mathbf{x}(t);t) = \ 2x_1(t) + x_2(t) - 1 - 5t \leq 0 \\ & g_2(\mathbf{x}(t);t) = \ x_1(t)^2 + x_2(t)^2 - 1 - t \leq 0 \\ & x_1(t), x_2(t) \geq 0,\ t \in [0,1] \end{aligned} \tag{6.1}$$

The analytic optimal solution is:

$$x_1^*(t) = \begin{cases} 0.4 + 2t - \sqrt{0.16 - 0.2t - t^2} & 0 \leq t \leq 0.28 \\ \sqrt{0.5(t+1)} & 0.28 < t \leq 1 \end{cases}$$

$$x_2^*(t) = \begin{cases} 0.2 + t + 2\sqrt{0.16 - 0.2t - t^2} & 0 \leq t \leq 0.28 \\ \sqrt{0.5(t+1)} & 0.28 < t \leq 1 \end{cases}$$

$$f^*(t) = \begin{cases} -.6 - 3t - \sqrt{.16 - .2t - t^2} & 0 \leq t \leq 0.28 \\ -2\sqrt{0.5(t+1)} & 0.28 < t \leq 1 \end{cases}$$

The Lagrangian dual function is created by relaxing g_2 into the objective.

$$\begin{aligned} \theta(u(t);t) = \min_{x_1,x_2} \quad & -x_1(t) - x_2(t) + u(t)(x_1(t)^2 + x_2(t)^2 - 1 - t) \\ s.t. \quad & 2x_1(t) + x_2(t) - 1 - 5t \leq 0 \\ & x_1(t), x_2(t) \geq 0,\ t \in [0,1] \end{aligned} \tag{6.2}$$

The mp-SA is applied to LDP (2.11)

$$\max_{u(t)\geq 0,\ t\in[0,1]} \theta(u(t);t) = -\min_{u(t)\geq 0,\ t\in[0,1]} -\theta(u(t);t) \tag{6.3}$$

The analytic dual optimal solution is:

$$u^*(t) = \begin{cases} \frac{1}{10\sqrt{0.16-0.2t-t^2}} & 0 \leq t \leq 0.28 \\ 0.5\sqrt{\frac{2}{(t+1)}} & 0.28 < t \leq 1 \end{cases}$$

Relaxing $g_2(\mathbf{x};t)$ into the objective allows problem (6.2) to be solved without the need to approximate any functions in the constraint set. However, (6.2) is not an mp-QP since $u(t)$ is not constant in t. Thus an exact solution to (6.2) cannot be computed and will instead be approximated using an appropriate algorithm as required in Line 1 of Algorithm 2 to obtain a subgradient. The objective function in (6.2) is then approximated with a quadratic function and the resulting mp-QP is solved with the mp-QA algorithm.

The initial dual solution u^0 is the linear interpolation of the dual optimal solutions to (6.1) for $t = 0$ and $t = 1$. This "warm start" is guaranteed to produce a dual feasible solution since it is a linear interpolation of feasible points in a convex space ($\mathbf{u} \geq \mathbf{0}$) and is easily generalized to higher dimensions. Since the dual variables only appear in the objective, the relaxed primal problem is guaranteed to have a feasible solution for every $\mathbf{t} \in \Omega$.

The subgradient step-size is $\lambda_k = \frac{1}{k+35}$ and Algorithm 2 runs for twenty iterations. The first iteration is described in detail to provide details on how the subgradient is obtained and the dual solution is updated. Inputs are: an initial dual solution $u^0(t) = 0.25t + 0.25$; a step size sequence $\{\lambda_k\} = \left\{\frac{1}{35+k}\right\}$, and an initial upper bound $\mathrm{UB}_0 = \infty$

Line 1: Evaluate $\theta_0(t) = \theta\left(u^0(t); t\right)$ to obtain an initial primal solution $\mathbf{x}^0$ and subgradient ξ^0. The evaluation of $\theta\left(u^0(t); t\right)$ involves solving the following mp-NLP:

$$\begin{aligned} \theta(u^0(t); t) = \min_{x_1, x_2} \quad & -x_1(t) - x_2(t) + (0.25t + 0.25)((x_1(t)^2 + x_2(t)^2 - 1 - t) \\ st \quad & 2x_1(t) + x_2(t) - 1 - 5t \leq 0 \\ & x_1(t), x_2(t) \geq 0,\ t \in [0, 1] \end{aligned} \tag{6.4}$$

The $(0.25t + 0.25)((x_1(t)^2 + x_2(t)^2 - 1 - t)$ term in (6.4) is cubic in terms of t and $\mathbf{x}$ so the mp-QA algorithm is used with a tolerance of 0.1 to solve (6.4) and evaluate $\theta(u^0(t); t)$. The approximate solution to (6.4) partitions the parameter space into two critical regions: $CR_1 = [0, 0.56818]$ and $CR_2 = [0.56818, 1]$. The initial primal solution is:

$$x_1^0(t) = \begin{cases} 2.2426t + 0.0196 & t \in CR_1 \\ -0.7556t + 1.7111 & t \in CR_2 \end{cases} \quad x_2^0(t) = \begin{cases} 0.5148t + 0.9608 & t \in CR_1 \\ -0.8222t + 1.7444 & t \in CR_2 \end{cases}$$

and the initial dual function is:

$$\theta_0(t) = \begin{cases} 2.4709t^2 - 3.0777t - 0.9692 & t \in CR_1 \\ -0.7176t^2 + 0.8440t - 2.1384 & t \in CR_2 \end{cases}$$

A subgradient is obtained by evaluating $-g_2$ at the primal solution: $\xi^0(t) = -g_2(\mathbf{x}^0(t); t)$. The subgradient is approximated as a linear interpolation of the value of g_2 at the vertices of each critical region, as shown in Table 1.

Vertex $\hat{t}$	$g_2(\mathbf{x}^0(t); \hat{t})$
0	-0.076398
0.56818	1.6766
1	-0.23642

Table 1: Evaluation of $g_2(\mathbf{x}^0(t); t)$ at vertices for Example 6.1

$$\xi^0(t) = \begin{cases} -3.0852t + 0.0764 & t \in CR_1 \\ 4.4989t - 4.2625 & t \in CR_2 \end{cases}$$

Line 2: Since $\xi^0 \neq 0$ the mp-SA proceeds to Line 5.

Line 5: The subgradient norm $\left\|\xi^0\right\|_2 = 0.9372$ is computed, the direction $d^0 = \xi^0 / \left\|\xi^0\right\|_2$ is set, and the step size $\lambda_0 = 1/35$ is chosen.

Line 6: The next dual solution is obtained via the subgradient update:

$$\begin{aligned} u^1(t) &= P_{\geq 0}\left(u^0(t) + \lambda_0 d^0(t)\right) \\ &= P_{\geq 0}\left(\begin{cases} 0.25t + 0.25 & t \in CR_1 \\ 0.25t + 0.25 & t \in CR_2 \end{cases} + \frac{1}{32.801}\begin{cases} -3.0852t + 0.0764 & t \in CR_1 \\ 4.4989t - 4.2625 & t \in CR_2 \end{cases}\right) \\ &= \begin{cases} 0.3415t + 0.2477 & t \in CR_1 \\ 0.1167t + 0.3763 & t \in CR_2 \end{cases} \end{aligned}$$

Lines 8 and 9: Evaluate

$$\begin{aligned} \int_\Omega -\theta_1(t)dt = &\int_0^{0.56818} -\left(2.4709t^2 - 3.0777t - 0.9692\right) dt \\ &+ \int_{0.56818}^{1} -\left(-0.7176t^2 + 0.8440t - 2.1384\right) dt = 1.7294 \end{aligned}$$

Since the value of the integral is an improvement on UB_0, set $\text{UB}_1 = 1.7294$.

Line 13: Set $k = 1$ and return to Line 1.

Algorithm 2 continues until the termination condition is met or 20 iterations are completed.

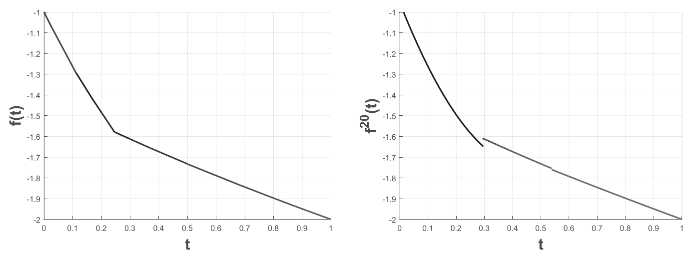
The results from solving (6.2) using the mp-SA and solving (6.1) directly using the mp-QA are listed in Table 2. The two methods are compared using the number of critical regions in the solution, CRs; the number of standard NLPs solved, NLPs; the number of multiparametric quadratic approximation problems solved, mp-QPs; the total objective function error, $\left\|f^* - f^{20}\right\|_2$; the total solution function error, $\left\|\mathbf{x}^* - \mathbf{x}^{20}\right\|_2$; and the total computational time.

	Tol = 0.1		Tol = 0.01		Tol = 0.001	
	mp-SA	**mp-QA**	**mp-SA**	**mp-QA**	**mp-SA**	**mp-QA**
CRs	7	16	36	43	16	122
NLPs	571	76	1527	233	1267	726
mp-QPs	177	10	473	32	403	111
$\left\|f^* - f^{20}\right\|_2$	0.0101	.0025	0.0027	.0004	0.0007	$3.6e-5$
$\left\|\mathbf{x}^* - \mathbf{x}^{20}\right\|_2$	0.0758	0.1817	0.0334	.0641	0.0272	0.0165
Time (sec)	48.9	5.6	164.3	14.4	104.1	41.6

Table 2: Example 6.1: comparison of results obtained with mp-SA and mp-QA

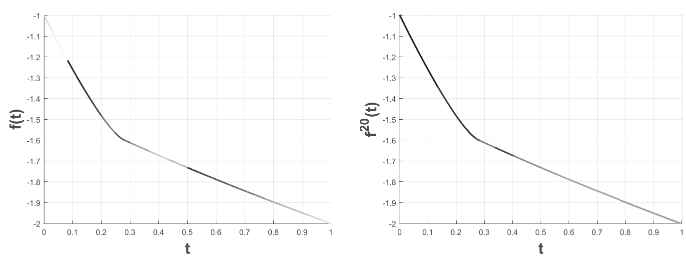
The number of NLP and mp-QP problems solved for the mp-SA is the sum over all iterations. In each case the mp-SA achieves comparable accuracy to the mp-QA and with fewer critical regions. The tradeoff for this advantage is the increased computational time due to the number of additional NLP and mp-QP problems that must be solved at each iteration. As the required accuracy increases however, the gap in the number of problems solved decreases.

The plots of the optimal value function produced be each algorithm are displayed in Figures 1 and 2. Figure 1 shows the value function obtained using an error tolerance of 0.1. There is a visible difference in these two graphs, with Figure 1(b) showing a discontinuity where none should exist, while Figure 1(a) is more accurate. Figure 2 shows the results for a 0.001 error tolerance and there is no distinguishable difference in the graphs of the two functions.



(a) mp-QA (b) mp-SA

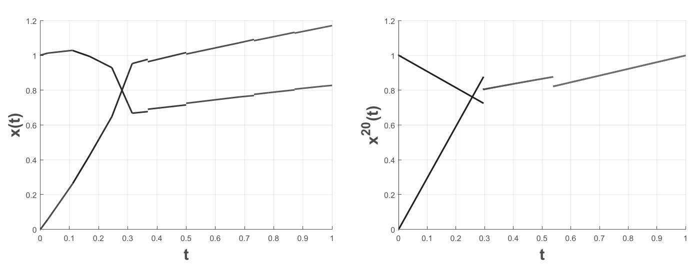
Figure 1: Example 6.1: optimal objective function, 0.1 tolerance



(a) mp-QA (b) mp-SA

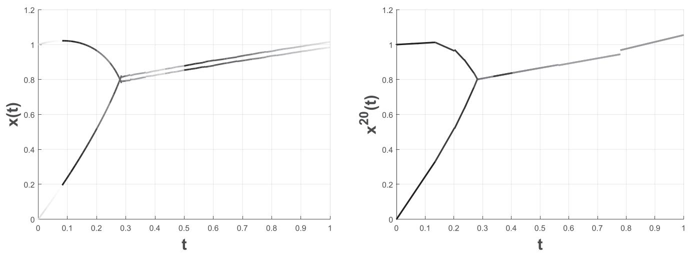
Figure 2: Example 6.1: optimal objective function, 0.001 tolerance

Figures 3 and 4 depict the graphs of the primal optimal solution functions produced by each algorithm for error tolerance levels 0.1 and 0.001, respectively, and allow for their comparison. The mp-QA algorithm produces a good approximation for the first critical region $[0, .28]$ but has a considerable trouble with the second region $[0.28, 1]$ where the approximations for x_1^* and x_2^* are not equal when they should be the same function. The gap between the two solutions diminishes as the error tolerance becomes more stringent, but is still present in Figure 4(a). The subgradient algorithm is much closer to the true solution in Figure 4(b), although the solution in Figure 3(b) is noticably worse.



(a) mp-QA (b) mp-SA

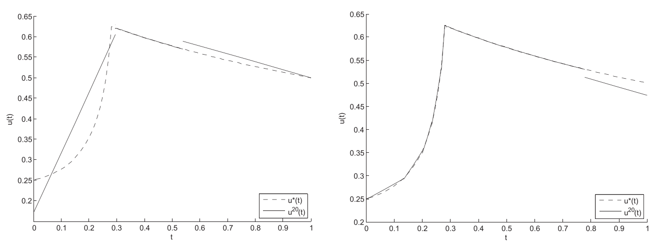
Figure 3: Example 6.1: optimal decision functions, 0.1 tolerance



(a) mp-QA (b) mp-SA

Figure 4: Example 6.1: optimal decision functions, 0.001 tolerance

In Figure 5, the graphs of the dual optimal solution function, u^{20}, produced by the mp-SA can be compared with the analytical solution, u^*, for error tolerances 0.1 and 0.001. Figure 5(a) clearly shows that the mp-SA solution after 20 iterations is not very good. However, Figure 5(b) shows a near-match between the approximated function and the optimal solution, with the only significant difference occuring at the end of the parameter space.



(a) 0.1 Tolerance (b) 0.001 Tolerance

Figure 5: Example 6.1: optimal dual function and its approximation

The difference between directly solving the primal problem (6.1) by means of the mp-QA versus solving the dual problem using a subgradient algorithm, such as the mp-SA, can be understood by examining how the constraint g_2 is handled by each method. Solving (6.1) requires that the quadratic g_2 be approximated by a set of linear inequalities. As the required accuracy becomes stricter, more inequalities must be used to satisfactorily approximate the feasible region. Many of these linear inequalities are nearly equivalent to each other, leading to problems related to linear dependence of the constraints. By contrast, solving the dual problem using a subgradient algorithm allows g_2 to be relaxed into the objective and the relaxed primal problem (6.2) to be solved instead. Using the mp-QA to solve (6.2) replaces the cubic function u_2g_2 with a single quadratic approximation. A new approximation of the objective function is made at each iteration of the mp-SA and hence the difference between the two methods can be summarized as a tradeoff between a *set* of linear approximations or a *sequence* of quadratic approximations to deal with the problematic constraint.

These results serve to illustrate the benefit that subgradient methods can bring to MPP. Combining them with other strategies, in this case mp-QA, can produce better results than the mp-QA can achieve by itself.

$\boxed{6.2}$ Multiparametric example

The next example is a parametric adaptation of the Rosen-Suzuki problem, a common test problem for subgradient algorithms [25]. A parameter t_1 is placed in the objective function and a second parameter t_2 is placed in the second constraint. The first parameter influences the value of x_1 which is present in all three constraints as both a quadratic and linear term. The second constraint is inactive at the optimal solution of the Rosen-Suzuki problem, so parameter t_2 is included to reduce the value of the righthand side, effectively tightening the constraint until it becomes active.

$$\begin{aligned}
\min_{\mathbf{x}} \quad & f(\mathbf{x}(\mathbf{t});\mathbf{t}) = \left(x_1(\mathbf{t}) - \frac{5}{2} + t_1\right)^2 + x_2(\mathbf{t})^2 + 2x_3(\mathbf{t})^2 + x_4(\mathbf{t})^2 - 5x_2(\mathbf{t}) - 21x_3(\mathbf{t}) + 7x_4(\mathbf{t}) \\
s.t. \quad & g_1(\mathbf{x}(\mathbf{t});\mathbf{t}) = x_1(\mathbf{t})^2 + x_2(\mathbf{t})^2 + x_3(\mathbf{t})^2 + x_4(\mathbf{t})^2 + x_1(\mathbf{t}) - x_2(\mathbf{t}) + x_3(\mathbf{t}) - x_4(\mathbf{t}) \le 8 \\
& g_2(\mathbf{x}(\mathbf{t});\mathbf{t}) = x_1(\mathbf{t})^2 + 2x_2(\mathbf{t})^2 + x_3(\mathbf{t})^2 + 2x_4(\mathbf{t})^2 - x_1(\mathbf{t}) - x_4(\mathbf{t}) \le 10 - t_2 \\
& g_3(\mathbf{x}(\mathbf{t});\mathbf{t}) = 2x_1(\mathbf{t})^2 + x_2(\mathbf{t})^2 + x_3(\mathbf{t})^2 + 2x_1(\mathbf{t}) - x_2(\mathbf{t}) - x_4(\mathbf{t}) \le 5 \\
& \mathbf{t} \in [-1,1] \times [0,3]
\end{aligned} \tag{6.5}$$

The dual function is created by relaxing all three nonlinear constraints into the objective, resulting in an unconstrained minimization problem.

$$\begin{aligned}
\theta(\mathbf{u}(\mathbf{t});\mathbf{t}) = \min_{x_1,x_2} \quad & \left(x_1(\mathbf{t}) - \frac{5}{2} + t_1\right)^2 \\
& + x_2(\mathbf{t})^2 + 2x_3(\mathbf{t})^2 + x_4(\mathbf{t})^2 - 5x_2(\mathbf{t}) - 21x_3(\mathbf{t}) + 7x_4(\mathbf{t}) \\
& + u_1(\mathbf{t})g_1(\mathbf{x}(\mathbf{t});\mathbf{t}) + u_2(\mathbf{t})g_2(\mathbf{x}(\mathbf{t});\mathbf{t}) + u_3(\mathbf{t})g_3(\mathbf{x}(\mathbf{t});\mathbf{t}) \\
& \mathbf{t} \in [-1,1] \times [0,3]
\end{aligned}$$

The primal and dual solutions to problem (6.6) are again approximated with piecewise affine functions of $\mathbf{t}$, using the mp-QA algorithm to evaluate $\theta(\mathbf{u}(\mathbf{t});\mathbf{t})$ in Line 1 of Algorithm 2. The subgradient step-size sequence is $\left\{\frac{2}{k+1}\right\}$ and Algorithm 2 is halted after twenty iterations. The initial dual solution $\mathbf{u}^0$ is the linear interpolation of the dual optimal solutions to problem (6.5) at the vertices of $[-1,1] \times [0,3]$.

The optimal solution to (6.5) obtained using Algorithm 2 is compared to the solution obtained using the mp-QA. Since an exact, analytic optimal solution is not available as it was for Example 6.1, the accuracy of each optimal solution is approximated by evaluating $\left\|f^* - f^{20}\right\|_2$ and $\left\|\mathbf{x}^* - \mathbf{x}^{20}\right\|_2$ using discretization with the Trapezoidal Rule and a 21×21 grid partition of Ω.

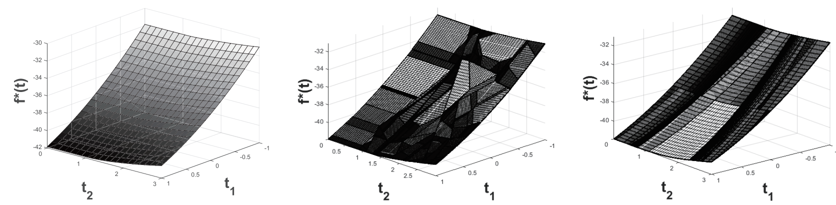
Three error tolerance levels are used and the results are reported in Table 3. The mp-QA failed when used with the strictest tolerance 0.001 due to partitioning Ω into critical regions with areas on the order of 10^{-8}. The algorithm creates these regions to maintain feasibility of the approximated solution, but when they become too small it leads to numerical failure. Algorithm 2 does not encounter this issue since all the nonlinear constraints are relaxed into the objective.

At each tolerance level the mp-SA requires fewer critical regions to approximate the solution than does the mp-QA. The tradeoff however is a substantial increase in the total number of NLPs and mp-QPs that must be solved by the mp-SA. The accuracy of the mp-QA approximation of f^* improves as the tolerance decreases while the mp-SA accuracy is unchanged. The mp-SA is capable of obtaining a solution at the 0.001 tolerance level while the mp-QA fails.

	Tol = 0.1		Tol = 0.01		Tol = 0.001	
	mp-SA	**mp-QA**	**mp-SA**	**mp-QA**	**mp-SA**	**mp-QA**
CRs	148	268	373	2257	804	-
NLPs	6890	1440	20898	11791	50998	-
mp-QPs	1378	47	4018	358	10195	-
$\left\lVert f^* - f^{20} \right\rVert_2$	0.0236	0.0631	0.0238	0.0067	0.0239	-
$\left\lVert \mathbf{x}^* - \mathbf{x}^{20} \right\rVert_2$	0.0645	0.1934	0.0647	0.0535	0.0646	-
Time (sec)	646.8	46.8	2479.2	284.3	7115.1	-

Table 3: Example 6.2: comparison of results obtained with mp-SA and mp-QA

The plots of f^* for each algorithm are shown in Figure 6 along with the plot of the optimal value function found through discretization. The optimal value function for (6.5) is close to planar and there is little to visually distinguish the solutions from each other. Both algorithms can achieve accurate solutions for this problem even for relatively large tolerances.



(a) $f^*(\mathbf{t})$ (b) mp-QA (c) mp-SA

Figure 6: Example 6.2: optimal objective value function

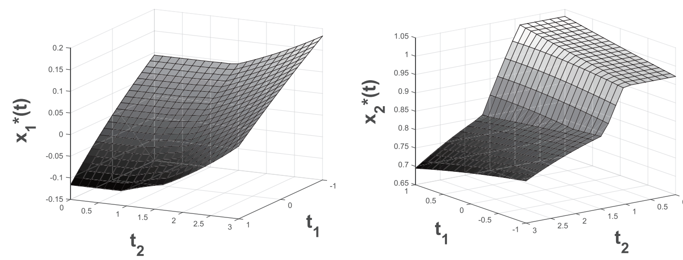
Figures 7 – 9 show the plots of the optimal x_1 and x_2 solution functions obtained by discretization as well as the functions returned by the two algorithms. These plots serve to illustrate - even more than the error calculations do in Table 3 - how much better the solution functions returned by the mp-SA are than those returned by the mp-QA. The constraint approximations used by the mp-QA produce discontinuous decision functions whereas the mp-SA maintains the continuity of the optimal solutions visible in Figure 7 thanks to the ability to relax the nonlinear constraints into the objective. This is a similar behavior to what is observed in the solution functions in Figures 3 and 4 for Example 6.1.

Examples 6.1 and 6.2 demonstrate that the mp-SA in conjunction with the dual problem produces solutions competitive with what a state-of-the-art mp-NLP algorithm can achieve solving the primal problem. Both algorithms yield accurate value functions, but the mp-SA distinguishes itself by producing solution functions that are more accurate and have fewer discontinuities than those yielded by the mp-QA. This is due to the ability to relax nonlinear constraints into the objective, bypassing the need to construct linear approximations for them when solving the relaxed primal problem at a given dual solution. In contrast, the mp-QA must approximate these constraints to maintain feasibility, resulting in partitioning the parameter space into ever smaller critical regions. As demonstrated in this example, if the error tolerance is too small it can lead to numerical problems that prevent the algorithm from returning a solution.

The downside to the mp-SA is the large number of total NLPs and mp-QPs it must solve over all its iterations. At smaller tolerances the difference in the number of problems solved by each method can shrink, as it does in Example 6.1, but this is not guaranteed to occur as Example 6.2 illustrates. If the value function is the only item of interest, the mp-QA is likely to be the more efficient solution method. If the solution functions are needed or if a strict tolerance is required, the mp-SA is the better choice.

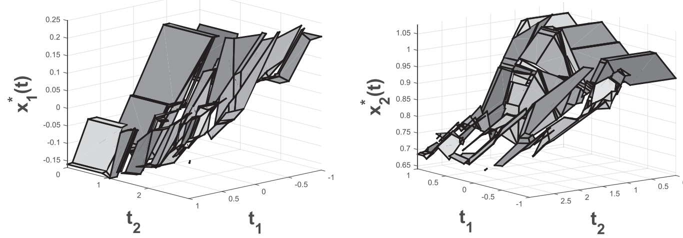
6.3 Implementation considerations

Several issues arise from implementing the mp-SA that do not exist or are more complicated to deal with than in the nonparametric setting. They involve representation of functions, obtaining a subgradient from the relaxed primal problem, and partitioning of the parameter space Ω into critical regions.



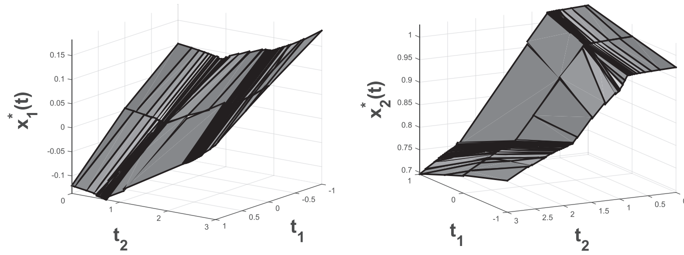
(a) $x_1^*(\mathbf{t})$ (b) $x_2^*(\mathbf{t})$

Figure 7: Example 6.2: optimal solution functions x_1, x_2



(a) $x_1(\mathbf{t})$ (b) $x_2(\mathbf{t})$

Figure 8: Example 6.2: optimal solution functions x_1, x_2 computed with mp-QA



(a) $x_1^{20}(\mathbf{t})$ (b) $x_2^{20}(\mathbf{t})$

Figure 9: Example 6.2: optimal solution functions x_1, x_2 computed with mp-SA

Every iteration of the mp-SA produces functions of the parameter that need to be represented in a tractable way. The solution functions and subgradient can be approximated with piecewise linear functions. Evaluating linear constraints at a solution is then straightforward while nonlinear constraints can be approximated by a linear interpolation of the constraint function at the vertices of each critical region. Keeping the approximations linear simplifies the process of obtaining the subgradient and making the update for the next iteration.

As mentioned earlier in Section 6.1, in Algorithm 2 obtaining a subgradient requires solving the relaxed primal problem to obtain the primal solution $\mathbf{x}^k$ associated with the dual solution $\left(\mathbf{u}^k, \mathbf{v}^k\right)$. Since the relaxed primal problem is itself an mp-NLP, a suitable solution method can be used to execute Line 1 in the algorithm. In Sections 6.1 and 6.2 the mp-QA [12, 13] is selected for this step.

Solving an mp-NLP results in partitioning Ω into critical regions for which mappings $\mathbf{t} \to \mathbf{x}^*$ are valid. The mp-SA generates many such mappings during its execution and it is likely that the intermediate solutions generate critical regions that differ from those of the optimal solution, particularly for Algorithm 2 which involves solving the relaxed primal problem for nonoptimal dual feasible solutions $(\mathbf{u}, \mathbf{v})$. Since the mp-SA iterates through several such solutions, it is possible that many such "extra" partitions could be introduced over the course of the algorithm. Because each iteration builds on the prior one, the "extra" partitions will remain in the solution unless they are explicitly addressed. The solution to Example 6.1 illustrates this outcome. During an early iteration in the execution of Algorithm 2, the primal solution associated with a nonoptimal dual solution generates critical regions $[0.2841, 0.2847]$ and $[0.2847, 0.2858]$. The components of $\mathbf{x}^k(\mathbf{t})$ differ between these two regions at that time, however at termination those same components converge to a common function as shown in Table 4.

	$CR_A = [0.2841, 0.2847]$	$CR_B = [0.2847, 0.2858]$
x_1	$0.3120t + 0.7127$	$0.3118t + 0.7127$
x_2	$0.3120t + 0.7126$	$0.3119t + 0.7127$
u	$-0.2429t + 0.6930$	$-0.2426t + 0.6929$

Table 4: Example 6.1: Critical regions A and B at termination

This outcome – a partition of Ω that is not required by the optimal solution – increases the amount of work performed by the algorithm since it must update the solution associated with every critical region in the partition of Ω. If some of the critical regions have the same solution, time can be saved by first consolidating them into a single region and then performing the update. The consolidation can be performed by periodically checking if two adjacent critical regions have identical mappings of $\mathbf{t} \to \mathbf{x}^*$ and $\mathbf{t} \to (\mathbf{u}^*, \mathbf{v}^*)$. One method to perform this check is to evaluate the normed difference of $\mathbf{x}$ (or $\mathbf{u}$ and $\mathbf{v}$) over the union of adjacent regions and consolidate them if the difference is within some tolerance ϵ:

$$\|\mathbf{x}_{CR_1} - \mathbf{x}_{CR_2}\|_2^2 < \int_{CR_1 \cup CR_2} \left(\mathbf{x}_{CR_1} - \mathbf{x}_{CR_2}\right)^T \left(\mathbf{x}_{CR_1} - \mathbf{x}_{CR_2}\right) < \epsilon$$

where $\mathbf{x}_{CR_j}$ is the part of the piecewise function $\mathbf{x}$ associated with critical region CR_j.

The benefits of this step are demonstrated in Table 5 where the mp-SA results given earlier in Table 2 are compared with the results obtained with mp-SA*, a version of mp-SA that is run for 20 iterations without the critical region unification step. In mp-SA, regions are checked for unification at every iteration and combined if $\left\|u_{CR_j} - u_{CR_k}\right\|_2 \leq 0.001$. These results indicate that performing the unification step can dramatically reduce the number of problems that must be solved without a corresponding loss in accuracy. The unification step reduces on average the number of critical regions by 48%, the number of NLPs that must be solved by 25%, and the number of mp-QPs that must be solved by 26%. Reducing the number of problems solved helps reduce the total time, though it is unclear why this does not happen for the 0.01 tolerance case. It is likely that the unification step does not need to be performed at every iteration and that performance gains can be made running it e.g., every five iterations. More research is needed however to determine desirable guidelines for applying this step.

	Tol = 0.1		Tol = 0.01		Tol = 0.001	
	mp-SA	**mp-SA***	**mp-SA**	**mp-SA***	**mp-SA**	**mp-SA***
CRs	7	20	36	57	16	36
NLPs	571	830	1527	1932	1267	1712
mp-QPs	177	264	473	608	403	552
$\left\lVert f^* - f^{20} \right\rVert_2$	0.0101	0.0101	0.0027	0.0027	0.0007	0.0009
$\left\lVert \mathbf{x}^* - \mathbf{x}^{20} \right\rVert_2$	0.0758	0.0758	0.0334	0.0334	0.0272	0.0273
Time (sec)	48.9	62.1	164.3	139.9	104.1	123.3

Table 5: Example 6.1: comparison of results obtained with mp-SA (cf. Table 2) and mp-SA* (i.e., mp-SA without critical region unification)

7 Conclusion

Results on Lagrangian duality for convex mp-NLPs are presented. The extension of theory has been possible due to redefining the feasible set of vectors in the standard mp-NLP formulation into a feasible set of functions and relating the resulting problem to its integral counterpart in a function space. Under suitable assumptions, weak and strong duality relationships are derived for the counterpart and an equivalence between the counterpart and the mp-NLP in the context of their primal and dual problems is proved. These results lead to proving strong duality for the mp-NLP.

Working in a function space allows to employ a subgradient algorithm available in the literature and designed to optimize a convex functional. The algorithm is recast to work for multiparametric optimization and is presented in two variants, primal and dual. Two examples illustrate how this algorithm, implemented with a simple choice of step size, can obtain primal and dual optimal solutions to mp-NLPs that are competitive to those of an existing method providing a primal optimal solution. In one instance, the algorithm obtains a solution at a tolerance level for which that other method fails. The obtained numerical experience is presented and the issues that affect the algorithm's performance are identified.

While a common step size condition is used in this implementation, other convergence results are available for subgradient algorithms in a Hilbert space which depend upon how the step size is chosen and the next iterate computed [3, 6]. This opens up possibilities for more numerical studies on the effectiveness of subgradient optimization for MPP. Additionally, the dual optimal solution may become useful in applications of parametric optimization analogously to the nonparametric counterpart.

References

[1] N. Adelgren, *Advancing Parametric Optimization: Theory and Solution Methodology for Multiparametric Linear Complementarity Problems with Parameters in General Locations*, SpringerBriefs on Optimization Series, Springer, 2021.

[2] N. Adelgren and M.M. Wiecek, A two-phase algorithm for the multiparametric linear complementarity problem, *Eur. J. Oper. Res.* 254 (2016) 715–738.

[3] Y.I. Alber, A.N. Iusem and M.V. Solodov, On the projected subgradient method for nonsmooth convex optimization in a Hilbert space, *Math. Program.* 81 (1998) 23–35.

[4] B. Bank, J. Guddat, D. Klatte, B. Kummer and K. Tammer, *Non-Linear Parametric Optimization*, Springer, Basel, 1983.

[5] M.S. Bazaara, H.D. Sherali and C.M. Shetty, *Nonlinear Programming: Theory and Algorithms*, Wiley, New York, 2006.

[6] J.Y. Bello Cruz and A.N. Iusem, A strongly convergent method for nonsmooth convex minimization in Hilbert spaces, *Numer. Funct. Anal. Optim.* 32 (2011) 1009–1018.

[7] A. Bemporad and C. Filippi, An algorithm for approximate multiparametric convex programming, *Comput. Optim. Appl.* 35 (2006) 87–108.

[8] C. Berge, *Topological Spaces*, Oliver & Boyd, Edinburgh, 1963.

[9] J.F. Bonnans and A. Shapiro, *Perturbation Analysis of Optimization Problems*, Series in Operations Research, Springer, 2000.

[10] R.I. Bot, V. Jeyakumar and G.Y. Li, Robust duality in parametric convex optimization, *Set-Valued Var. Anal.* 21 (2013) 177–189.

[11] V. Charitopoulos and V. Dua, A unified framework for model-based multi-objective linear process and energy optimisation under uncertainty, *Appl. Energ.* 186 (2017) 539–548.

[12] L.F. Domínguez, D.A. Narciso and E.N. Pistikopoulos, Recent advances in multiparametric nonlinear programming, *Comput. Chem. Eng.* 34 (2010) 707–716.

[13] L.F. Domínguez and E.N. Pistikopoulos, A quadratic approximation-based algorithm for the solution of multiparametric mixed-integer nonlinear programming problems, *AiChE J.* 59 (2013) 483–495.

[14] V. Dua and E.N. Pistikopoulos, Algorithms for the solution of multi-parametric mixed-integer nonlinear optimization problems, *Ind. Eng. Chem. Res.* 38 (1999) 3976–3987.

[15] M. Ehrgott, *Multicriteria Optimization*, Springer, 2nd edition, 2005.

[16] N. Faísca, P. Saraiva, B. Rustem and E.N. Pistikopoulos, A multi-parametric programming approach for multilevel hierarchical and decentralised optimisation problems, *Comput. Manage. Sci.* 6 (2009) 377–397.

[17] A.V. Fiacco, *Introduction to Sensitivity and Stability Analysis in Nonlinear Programming*, Mathematics in Science and Engineering, Academic Press, volume 165, 1983.

[18] C.B. Garcia and W.I. Zangwill, *Pathways to Solutions, Fixed Points, and Equilibria*, Series in Computational Mathematics, Prentice Hall, 1981.

[19] A. Geoffrion, Proper efficiency and the theory of vector maximization, *J. Math. Anal. Appl.* 22 (1968) 618–630.

[20] A. Grancharova and T.A. Johansen, *Explicit Nonlinear Model Predictive Control: Theory and Applications*, Springer, Berlin Heidelberg, 2012.

[21] J. Guddat, F.G. Vasquez and H.T. Jongen, *Parametric Optimization: Singularities, Pathfollowing, and Jumps*, John Wiley & Sons, 1990.

[22] Ø. Hegranæs, J. Gravdahl and P. Tøndel, Spacecraft altitude control using explicit model predictive control, *Automatica* 41 (2005) 2107–2114.

[23] M. Herceg, M. Kvasnica, C.N. Jones and M. Morari, Multi-Parametric Toolbox 3.0, in: *Proc. of the European Control Conference*, Zürich, Switzerland, 2013, pp. 502–510.

[24] M. Hestenes, *Optimization Theory: The Finite Dimensional Case*, Wiley, New York, 1975.

[25] W. Hock and K. Schittkowski, *Test Examples for Nonlinear Programming Codes*, Lecture Notes in Economics and Mathematical Systems, Springer, Berlin, volume 187, 1981.

[26] T.A. Johansen, On multi-parametric nonlinear programming and explicit nonlinear model predictive control, in: *Proc. 41st IEEE Conf. Decis. Control*, Las Vegas, Nevada, 2002, pp. 2768–2773.

[27] A.M. Kassa and S.M. Kassa, A branch-and-bound multi-parametric programming approach for non-convex multilevel optimization with polyhedral constraints, *J. Glob. Optim.* 64 (2016) 745–764.

[28] V. Kosmidis, A. Panga, V. Sakizlis, G. Charles, S. Kenchington, N. Bozinis and E.N. Pistikopoulos, Output feedback parametric controllers for an active valve train actuation system, in: *Proc. 45th IEEE Conf. Decis. Control*, San Diego, California, 2006, pp. 4520–4525.

[29] J. Leverenz, M. Xu and M.M. Wiecek, Multiparametric optimization for multidisciplinary engineering design, *Stuct. Multidisc. Optim.* 54 (2016) 795–810.

[30] L. Li and Z. Zabinsky, Incorporating uncertainty into a supplier selection problem, *Int. J. Prod. Econ.* 134 (2011) 344–356.

[31] D.G. Luenberger, *Optimization by Vector Space Methods*, Wiley, New York, 1969.

[32] D.G. Luenberger, *Linear and Nonlinear Programming*, Addison-Wesley, Reading, MA, 1984.

[33] D. Narciso *Developments in Nonlinear Multiparametric Programming and Control*, Ph.D. Thesis, Imperial College, London, UK, 2009.

[34] R. Oberdieck, N. Diangelakis, I. Nascu, M. Papathanasiou, S. Avraamidou and E.N. Pistikopoulos, On multi-parametric programming and its applications in process systems engineering, *Chem. Eng. Res. Des.* 116 (2016) 61–82.

[35] R. Oberdieck and E.N. Pistikopoulos, Multi-objective optimization with convex quadratic cost functions: A multi-parametric programming approach, *Comput. Chem. Eng.* 85 (2016) 36–39.

[36] E.N. Pistikopoulos, N.A. Diangelakis and R. Oberdieck, *Multi-Parametric Optimization and Control*, Series in Operations Research and Management Science, Wiley, 2021.

[37] E.N. Pistikopoulos, V. Dua, N. Bozinis, A. Bemporad and M. Morari, On-line optimization via off-line parametric optimization tools, *Comput. Chem. Eng.* 26 (2002) 175–184.

[38] E.N. Pistikopoulos, M.C. Georgiadis and V. Dua, *Multi-Parametric Programming: Theory, Algorithms, Applications*, Wiley, 2007.

[39] G. Takács and B. Rohal'-Ilkiv, Model predictive control algorithms for active vibration control: a study on timing, performance, and implementation properties, *J. Vib. Control* 20 (2014) 2061–2080.

[40] T. Zolezzi, Well-posed optimization problems for integral functionals, *J. Optimiz. Theory App.* 31 (1980) 417–430.

Manuscript received 5 January 2022
revised 14 June 2022
accepted for publication 16 August 2022

J. Leverenz
School of Mathematical and Statistical Sciences
Clemson University, Clemson, SC 29634, United States
E-mail address: jlevere@g.clemson.edu

H. Lee
School of Mathematical and Statistical Sciences
Clemson University, Clemson, SC 29634, United States
E-mail address: hklee@clemson.edu

M.M. Wiecek
School of Mathematical and Statistical Sciences
Clemson University, Clemson, SC 29634, United States
E-mail address: wmalgor@clemson.edu