J Optim Theory Appl (2010) 147: 125–140
DOI 10.1007/s10957-010-9711-4

Duality and Exact Penalization for General Augmented Lagrangians

R.S. Burachik · A.N. Iusem · J.G. Melo

Published online: 11 May 2010

Abstract We consider a problem of minimizing an extended real-valued function defined in a Hausdorff topological space. We study the dual problem induced by a general augmented Lagrangian function. Under a simple set of assumptions on this general augmented Lagrangian function, we obtain strong duality and existence of exact penalty parameter via an abstract convexity approach. We show that every cluster point of a sub-optimal path related to the dual problem is a primal solution. Our assumptions are more general than those recently considered in the related literature.

Keywords Hausdorff topological spaces · Nonsmooth optimization · Nonconvex problem · General augmented Lagrangian · Duality · Abstract convexity

Communicated by X.Q. Yang.

The authors are very grateful to the referees for their careful reading and correction of the previous version of the manuscript. In particular, the example in Remark 3.2 was motivated by an example provided by one of the referees. J.G. Melo was supported by CNPq-Brazil. J.G. Melo would like to thank the School of Mathematics and Statistics at the University of South Australia, for providing excellent conditions and a stimulating environment for carrying out his research.

R.S. Burachik (✉)
School of Mathematics and Statistics, University of South Australia, Adelaide, Australia
e-mail: regina.burachik@unisa.edu.au

A.N. Iusem · J.G. Melo
Instituto de Matemática Pura e Aplicada, Rio de Janeiro, Brazil

A.N. Iusem
e-mail: iusp@impa.br

J.G. Melo
e-mail: jeferson@impa.br

1 Introduction

It is well known that augmented Lagrangian methods are useful for solving constrained (nonconvex) optimization problems. Rockafellar and Wets [1] considered a primal problem of minimizing an extended real-valued function and proposed and analyzed a dual approach via augmented Lagrangians. Strong duality and a criterion for exact penalty representation are shown in Rockafellar and Wets [1, Theorems 11.59 and 11.61]. Recently, this duality approach has been studied in a more general setting. In Huang and Yang [2] the convexity assumption on the augmenting function, which is assumed by Rockafellar and Wets [1], is relaxed to a level boundedness assumption. Nedic and Ozdaglar [3] considered a geometric dual approach and studied a zero-duality-gap property. Many efforts have been devoted to augmented Lagrangians with a valley-at-zero property on the augmenting function, see for example Burachik and Rubinov [4], Rubinov et al. [5], Zhou and Yang [6] and references therein. In Wang et al. [7, Sect. 3.1], an augmented Lagrangian type function is studied via an auxiliary coupling function, and a valley-at-zero type property is proposed in the derivative of the coupling function with respect to the penalty parameter. Penot and Rubinov [8] investigated the relationship between the Lagrangian multipliers and the generalized subdifferential of the perturbation function in ordered spaces.

In the present paper, we consider a primal problem of minimizing an extended real-valued function in a Hausdorff topological space. A main tool in our analysis is abstract convexity, which recently became a natural language to investigate duality-schemes via augmented Lagrangian type functions, see for example Burachik and Rubinov [4], Rubinov et al. [5], Penot and Rubinov [8], Nedić et al. [9] and Rubinov and Yang [10]. With abstract convexity tools, we propose and analyze a duality scheme induced by a general augmented Lagrangian function. We consider a valley-at-zero type property on the coupling (augmenting) function, which generalizes the valley-at-zero type property proposed in the related literature (e.g., Burachik and Rubinov [4] and references therein, Wang et al. [7, Sect. 3.1]), see Sect. 5 in the present paper. To obtain our results, we only need to assume continuity at a fixed point instead of at the whole space, the latter being a standard assumption in the literature (see, e.g., Burachik and Rubinov [4]). We show that our duality scheme has a zero-duality-gap property. A sub-optimal path related to the dual problem is considered, and we prove that all its cluster points are primal solutions.

A criterion for exact penalization was presented in Rockafellar and Wets [1, Theorem 11.61]. This criterion has been generalized, for instance, by Huang and Yang [2] and Burachik and Rubinov [4]. We also extend this criterion to our general setting. Since no linearity on the augmented Lagrangian is assumed, this allows us to consider our primal-dual scheme in Hausdorff topological spaces.

The main motivation for working in the framework of Hausdorff topological spaces is to develop a duality theory that can encompass different settings found in the literature, such as metric spaces (see e.g., [5, 11–13]) and Banach spaces with the weak topology (see e.g., [4, 6, 14]), which in general are not metrizable. It is also worthwhile to note that the general augmented Lagrangian, considered in the present paper, for which the valley-at-zero type property is assumed directly at the coupling function ρ (see Sect. 2), has not been considered in the literature even in finite dimensional spaces.

The outline of this manuscript is as follows. Section 2 contains basic definitions and assumptions. Also, our primal-dual scheme is stated. In Sect. 3, we show that our duality scheme provides strong duality, and a criterion to exact penalty representation is presented. In Sect. 4, we study the convergence properties of a sub-optimal path related to our dual problem. In the last section, we present some examples and compare our setting with the ones considered in Burachik and Rubinov [4], and Wang et al. [7, Sect. 3.1].

2 Statement of the Problem and Basic Assumptions

Let Y be an arbitrary (nonempty) set. Let X and Z be Hausdorff topological spaces. We consider the optimization problem

$$\text{minimize}\, \varphi(x) \quad \text{subject to} \quad x \text{ in } X, \tag{1}$$

where the function $\varphi : X \to \mathbb{R}_{+\infty} := \mathbb{R} \cup \{+\infty\}$ is a proper (i.e., $\operatorname{dom}\varphi \neq \emptyset$ and $\varphi > -\infty$) extended real-valued function. We fix a base point in Z and denote it by 0. In order to introduce our duality scheme, we consider a *duality parameterization* for φ, which is a function $f : X \times Z \to \bar{\mathbb{R}} := \mathbb{R} \cup \{\pm\infty\}$ satisfying $f(x, 0) = \varphi(x)$ for all $x \in X$. We also consider a perturbation function $\beta : Z \to \bar{\mathbb{R}}$, related to this duality parameterization, given by

$$\beta(z) := \inf_{x \in X} f(x, z).$$

Since φ is proper, $\beta(0) < +\infty$. The definition of the dual function and dual problem rely on the coupling function used in the conjugation. For instance, the classical conjugate duality (in the setting of Banach spaces) is defined using the coupling function $\rho_0 : X \times X^* \to \mathbb{R}$ given by $\rho_0(x, x^*) := x^*(x)$. For a set $V \subset Z$ we use the notation $V^c := Z \setminus V$.

In what follows, we consider a coupling function $\rho : Z \times Y \times \mathbb{R}_+ \to \mathbb{R}$ that satisfies the following basic assumptions:

(C_1) For any $(y, r) \in Y \times \mathbb{R}_+$ the function $\rho(\cdot, y, r)$ is upper semicontinuous at 0, and $\rho(0, y, r) = 0$.

(C_2) For every neighborhood $V \subset Z$ of 0, and for every $(y, \bar{r}) \in Y \times \mathbb{R}_+$, it holds that

(i)

$$A^V_{y,\bar{r}}(r) := \inf_{z \in V^c} \{\rho(z, y, \bar{r}) - \rho(z, y, r)\} > 0, \quad \text{for all } r > \bar{r};$$

(ii)

$$\lim_{r \to \infty} A^V_{y,\bar{r}}(r) = \infty.$$

Remark 2.1 Condition C_2 is a valley-at-zero type property, which generalizes similar properties for augmenting functions recently introduced in the literature. Item (i) in condition C_2 ensures that the function $\rho(z, y, \cdot)$ is strictly decreasing for any fixed

$(y,z) \in Y \times Z \setminus \{0\}$. In particular, the function $A^V_{y,\bar{r}} : (\bar{r}, \infty) \to \mathbb{R}_+$ is nondecreasing, ensuring that $\lim_{r\to\infty} A^V_{y,\bar{r}}(r)$ exists. See Sect. 5 and references therein for more details on condition C_2, and its comparison with related assumptions in the literature.

The *augmented Lagrangian function* $\ell : X \times Y \times \mathbb{R}_+ \to \bar{\mathbb{R}}$, induced by the coupling function ρ, is defined as

$$\ell(x,y,r) := \inf_{z\in Z}\{f(x,z) - \rho(z,y,r)\}. \tag{2}$$

The dual function $q : Y \times \mathbb{R}_+ \to \bar{\mathbb{R}}$ is defined as $q(y,r) := \inf_{x\in X} \ell(x,y,r)$ and therefore the dual problem is stated as

$$\text{maximize}\, q(y,r) \quad \text{subject to} \quad (y,r) \text{ in } Y \times \mathbb{R}_+. \tag{3}$$

It is clear that $q(y,r) = \inf_{z\in Z}\{\beta(z) - \rho(z,y,r)\}$, where β is the perturbation function. We denote by $M_p := \inf_{x\in X} \varphi(x)$ the optimal value of the primal problem, and by $M_d := \sup_{(y,r)\in Y\times\mathbb{R}_+} q(y,r)$ the optimal value of the dual problem.

Since f is a parameterization function, condition C_1 easily implies the weak duality property for our scheme, that is, $M_d \le M_p$. In this section, we show that our duality scheme enjoys a strong duality property, that is to say, the zero-duality-gap property holds ($M_p = M_d$). Next we present some definitions related to abstract convexity. For a detailed presentation of this subject, see for example, Rubinov [15].

Definition 2.1 Let $g : Z \to \bar{\mathbb{R}}$. Given $\varepsilon \ge 0$, we say that (y,r) is an *ε-abstract subgradient of* g in $\bar{z}$ (with respect to ρ) iff

$$g(z) - \rho(z,y,r) \ge g(\bar{z}) - \rho(\bar{z},y,r) - \varepsilon \quad \text{for all } z \in Z. \tag{4}$$

The set of ε-abstract subgradients of g in $\bar{z}$, denoted by $\partial_{\rho,\varepsilon} g(\bar{z})$, is the *$\varepsilon$-abstract subdifferential of* g in $\bar{z}$ with respect to the coupling function ρ. The 0-abstract subdifferential in $\bar{z}$ is denoted by $\partial_\rho g(\bar{z})$, and is called *abstract subdifferential*.

Remark 2.2 It follows from C_1 and the definition of $\partial_{\rho,\varepsilon} g(0)$, that, if $(y,r_0) \in \partial_{\rho,\varepsilon} g(0)$, then $(y,r) \in \partial_{\rho,\varepsilon} g(0)$ for all $r \ge r_0$, using the fact that $\rho(z,y,\cdot)$ is decreasing.

The abstract conjugate and biconjugate functions of g with respect to the coupling function ρ are defined, respectively, by

$$g^\rho(y,r) = \sup_{z\in Z}\{\rho(z,y,r) - g(z)\}$$

and

$$g^{\rho\rho}(z) = \sup_{(y,r)\in Y\times\mathbb{R}_+}\{\rho(z,y,r) - g^\rho(y,r)\}.$$

Remark 2.3 It is easy to show that $\beta^{\rho\rho}(0) = M_d$, where β is the perturbation function. In particular, weak and strong duality are rewritten, respectively, as $\beta^{\rho\rho}(0) \leq \beta(0)$ and $\beta^{\rho\rho}(0) = \beta(0)$. In this context strong duality is related to abstract convexity of the function β with respect to the family of functions $H_\rho := \{\rho(\cdot, y, r) + c : (y, r, c) \in Y \times \mathbb{R}_+ \times \mathbb{R}\}$ at 0. For more details on the relationship between strong duality and abstract convexity at a point, see for example, Rubinov and Yang [10, Chap. 2 and Sect. 5.2].

Consider the set of functions H_ρ as in Remark 2.3. The set $\mathrm{Supp}(\beta, H_\rho) := \{h \in H_\rho : h \leq \beta\}$ is called the support set of β with respect to H_ρ. In next proposition we relate $\partial_\rho\beta(0)$, $\mathrm{Supp}(\beta, H_\rho)$, and $\mathrm{dom}\,\beta^\rho$ with the dual function q.

Proposition 2.1 *Take* $(\overline{y}, \overline{r}) \in Y \times \mathbb{R}_+$. *Then*

(i) $(\overline{y}, \overline{r}) \in \partial_\rho\beta(0)$, *if and only if* $q(\overline{y}, \overline{r}) = \beta(0)$;
(ii) $(\overline{y}, \overline{r}) \in \mathrm{dom}\,\beta^\rho$, *if and only if there exists* $\overline{c} \in \mathbb{R}$, *such that* $\rho(\cdot, \overline{y}, \overline{r}) + \overline{c} \in \mathrm{Supp}(\beta, H_\rho)$, *which in turn is equivalent to* $q(\overline{y}, \overline{r}) \geq \overline{c}$.

Proof Since weak duality holds, (i) follows from the following equivalences:

$$\begin{aligned}(\overline{y}, \overline{r}) \in \partial_\rho\beta(0) \quad &\Leftrightarrow \quad \beta(z) - \rho(z, \overline{y}, \overline{r}) \geq \beta(0) \quad (\forall z \in Z)\\ &\Leftrightarrow \quad \inf_z\{\beta(z) - \rho(z, \overline{y}, \overline{r})\} \geq \beta(0)\\ &\Leftrightarrow \quad q(\overline{y}, \overline{r}) \geq \beta(0).\end{aligned}$$

Since $q(\overline{y}, \overline{r}) = -\beta^\rho(\overline{y}, \overline{r})$, (ii) follows from the following equivalences:

$$\begin{aligned}\rho(\cdot, \overline{y}, \overline{r}) + \overline{c} \in \mathrm{Supp}(\beta, H_\rho) \quad &\Leftrightarrow \quad \beta(z) \geq \rho(z, \overline{y}, \overline{r}) + \overline{c} \quad (\forall z \in Z)\\ &\Leftrightarrow \quad \beta(z) - \rho(z, \overline{y}, \overline{r}) \geq \overline{c} \quad (\forall z \in Z)\\ &\Leftrightarrow \quad \inf_z\{\beta(z) - \rho(z, \overline{y}, \overline{r})\} \geq \overline{c}\\ &\Leftrightarrow \quad q(\overline{y}, \overline{r}) \geq \overline{c}.\end{aligned}$$

□

3 Strong Duality and Exact Penalty Representation

Next theorem ensures that, under mild assumptions, for every $\varepsilon > 0$ the ε-abstract subgradient of β at 0 is nonempty. As a consequence of this fact, we establish strong duality. In Example 4.1 we consider a constrained optimization problem for which the hypothesis of next theorem, regarding lower semicontinuity of β at 0, is satisfied. Proposition 4.1 presents some conditions, on the parameterization function, that guarantee the lower semicontinuity of β at 0.

Theorem 3.1 *Assume that* C_1 *and* C_2 *hold, that* β *be lower semicontinuous* (lsc) *at* 0, *and that there exists* $(\overline{y}, \overline{r}) \in \mathrm{dom}\,\beta^\rho$. *Then* $\partial_{\rho,\varepsilon}\beta(0) \neq \emptyset$ *for all* $\varepsilon > 0$. *Moreover, for any* $\varepsilon > 0$, *there exists* r_0, *such that* $(\overline{y}, r) \in \partial_{\rho,\varepsilon}\beta(0)$ *for all* $r \geq r_0$.

Proof First, by assumption $\beta(0) < \infty$. Second, we have that $\beta(0) > -\infty$ by weak duality and the assumption that $\operatorname{dom} \beta^\rho$ is nonempty. Therefore, $\beta(0) \in \mathbb{R}$. Observe that we just need to prove the last statement of the theorem. To arrive at a contradiction suppose that there exists $\bar{\varepsilon} > 0$, such that for any $k > 0$ there exists $r_k \geq k$, $z_k \in Z$ satisfying:

$$\beta(z_k) - \rho(z_k, \overline{y}, r_k) < \beta(0) - \bar{\varepsilon}. \tag{5}$$

Suppose that $\{z_k\}_{k\in\mathbb{N}}$ converges to 0. Thus

$$\beta(0) - \bar{\varepsilon} > \beta(z_k) - \rho(z_k, \overline{y}, r_k) > \beta(z_k) - \rho(z_k, \overline{y}, \overline{r})$$

for all $k \geq k_0 > \overline{r}$. Hence, using C_1 and the lower semicontinuity of β at 0, we have

$$\beta(0) - \bar{\varepsilon} \geq \liminf_{k\to\infty}\{\beta(z_k) - \rho(z_k, \overline{y}, \overline{r})\} \geq \beta(0) - \rho(0, \overline{y}, \overline{r}) = \beta(0),$$

which is a contradiction. Therefore $\{z_k\}_{k\in\mathbb{N}}$ does not converge to 0, which implies that there exists some open neighborhood $V \subset Z$ of 0, and a subsequence $\{z_{k_j}\}_{j\in\mathbb{N}} \subset V^c$. Now, using (5) and the fact that there exists $\overline{c}$ such that $\rho(\cdot, \overline{y}, \overline{r}) + \overline{c} \in \operatorname{Supp}(\beta, H_\rho)$ (see Proposition 2.1(ii)), we have

$$\begin{aligned}\beta(0) - \bar{\varepsilon} &> \beta(z_{k_j}) - \rho(z_{k_j}, \overline{y}, r_{k_j})\\ &= \beta(z_{k_j}) - \rho(z_{k_j}, \overline{y}, \overline{r}) + \rho(z_{k_j}, \overline{y}, \overline{r}) - \rho(z_{k_j}, \overline{y}, r_{k_j})\\ &\geq \overline{c} + \inf_{z\in V^c}\{\rho(z, \overline{y}, \overline{r}) - \rho(z, \overline{y}, r_{k_j})\}.\end{aligned}$$

Henceforth,

$$A^V_{\overline{y},\overline{r}}(r_{k_j}) := \inf_{z\in V^c}\{\rho(z, \overline{y}, \overline{r}) - \rho(z, \overline{y}, r_{k_j})\} \leq \beta(0) - \bar{\varepsilon} - \overline{c},$$

which contradicts C_2, because $\lim_{j\to\infty} r_{k_j} = \infty$. The result follows. □

Next corollary, which extends [4, Proposition 4.2], shows that in order to check if the abstract subgradient of β at 0 is nonempty we just need to verify that there exists an element $(y, r) \in Y \times R_+$ satisfying the inequality (4) in a neighborhood of 0. As we will see in Theorem 3.3, under mild assumptions, this fact is equivalent to the existence of an exact penalty representation.

Corollary 3.1 *Suppose that the assumptions of Theorem 3.1 hold. Suppose also that there exists an open neighborhood $V \subset Z$ of 0 such that $\beta(z) - \rho(z, \overline{y}, \overline{r}) \geq \beta(0)$ for all $z \in V$, with $(\overline{y}, \overline{r}) \in \operatorname{dom} \beta^\rho$. Then there exists r_0, such that $\beta(z) - \rho(z, \overline{y}, r) \geq \beta(0)$ for all $z \in Z$ and $r \geq r_0$, i.e., $(\overline{y}, r) \in \partial_\rho \beta(0)$ for all $r \geq r_0$.*

Proof Take V as in the assumption. Consider $z \in V^c$ and $\varepsilon > 0$. By Theorem 3.1 there exists $r_\varepsilon > 0$, such that $(\overline{y}, r_\varepsilon) \in \partial_{\rho,\varepsilon}\beta(0)$. Thus

$$\begin{aligned}\beta(z) &\geq \beta(0) + \rho(z, \overline{y}, r_\varepsilon) - \varepsilon\\ &= \beta(0) + \rho(z, \overline{y}, r) + \rho(z, \overline{y}, r_\varepsilon) - \rho(z, \overline{y}, r) - \varepsilon\\ &\geq \beta(0) + \rho(z, \overline{y}, r) + \inf_{u\in V^c}\{\rho(u, \overline{y}, r_\varepsilon) - \rho(u, \overline{y}, r)\} - \varepsilon. \qquad (6)\end{aligned}$$

By C_2 there exists $r_1 > r_\varepsilon$ such that $\inf_{u\in V^c}\{\rho(u,\overline{y},r_\varepsilon) - \rho(u,\overline{y},r)\} > \varepsilon$ for all $r \geq r_1$. Using this estimate in (6) we obtain $\beta(z) \geq \beta(0) + \rho(z,\overline{y},r)$ for all $z \in V^c$ and $r \geq r_1$. Since, by assumption, $\beta(z) \geq \beta(0) + \rho(z,\overline{y},\overline{r})$ for all $z \in V$, the result follows by taking $r_0 = \max\{\overline{r}, r_1\}$ and observing that $\rho(z,y,\cdot)$ is nonincreasing in $\mathbb{R}_+$ for each $(z,y) \in Z \times Y$. □

Theorem 3.2 *Under the assumptions of Theorem 3.1, the zero-duality-gap property holds for the primal-dual pair of problems (1)–(3).*

Proof Take $\varepsilon > 0$. By Theorem 3.1, there exists $(\overline{y}, r_\varepsilon) \in \partial_{\rho,\varepsilon}\beta(0)$. Hence we have

$$\begin{aligned}\beta^{\rho\rho}(0) &= \sup_{(y,r)}\{\rho(0,y,r) - \beta^\rho(y,r)\} = \sup_{(y,r)} -\beta^\rho(y,r)\\ &\geq -\beta^\rho(\overline{y},r_\varepsilon) = \inf_z\{\beta(z) - \rho(z,\overline{y},r_\varepsilon)\} \geq \beta(0) - \varepsilon,\end{aligned}$$

using Condition C_1 in the second equality and the fact that $(\overline{y}, r_\varepsilon) \in \partial_{\rho,\varepsilon}\beta(0)$ in the last inequality. It follows that $\beta^{\rho\rho}(0) \geq \beta(0) - \varepsilon$. Since ε is arbitrary, we have that $\beta^{\rho\rho}(0) \geq \beta(0)$, and the reverse inequality is the weak duality property. We conclude that $\beta^{\rho\rho}(0) = \beta(0)$, i.e. the zero-duality-gap property holds. □

Remark 3.1 Corollary 3.1 and Theorem 3.2 generalize Burachik and Rubinov [4, Propositions 4.2 and 4.1], respectively. Observe also that we just use the lower semicontinuity of β at 0, while in Burachik and Rubinov [4] β is assumed to be lsc in all the space.

Exact penalty representation for augmented Lagrangian function was defined and studied in Rockafellar and Wets [1, Chap. 11]. A criterion for such a representation was presented in Rockafellar and Wets [1, Theorem 11.61]. This criterion has been studied for more general augmented Lagrangians, for instance, by Burachik and Rubinov [4], and Huang and Yang [2]. In next theorem we extend this criterion to our more general setting.

Definition 3.1 Consider the primal and dual problems (1)–(3). An element $\overline{y} \in Y$ is said to *support an exact penalty representation for problem* (1) iff there exists $r_0 \in \mathbb{R}_+$, such that for any $r > r_0$,

(E_1) $\beta(0) = q(\overline{y}, r)$;
(E_2) $\operatorname{argmin}_x \varphi(x) = \operatorname{argmin}_x l(x,\overline{y},r)$.

Theorem 3.3 *Assume that*

(a) *the parameterization function f satisfies: $f(x,\cdot)$ is* lsc *at* 0 *for every $x \in X$;*
(b) *the perturbation function β is lsc at* 0;
(c) *conditions C_1 and C_2 are satisfied;*
(d) *there exists $(\overline{y},\overline{r}) \in \operatorname{dom}\beta^\rho$.*

Then the following assertions are equivalent:

(i) *There exist an open neighborhood $V \subset Z$ of* 0 *and* $r_0 > 0$, *such that*

$$\beta(z) \geq \beta(0) + \rho(z, \overline{y}, r_0) \quad \textit{for all } z \in V;$$

(ii) $\overline{y}$ *supports an exact penalty representation for problem* (1).

Proof First, we prove that (ii) ⇒ (i). By E_1 there exists $r_0 > 0$ such that $\forall r \geq r_0$

$$\beta(0) = q(\overline{y}, r) = \inf_{z \in Z}\{\beta(z) - \rho(z, \overline{y}, r)\}.$$

In particular, for any open neighborhood $V \subset Z$ of 0, we have that $\beta(0) \leq \beta(z) - \rho(z, \overline{y}, r)$ for all $z \in V$ and $r \geq r_0$, which proves (i).

Let us now prove that (i) ⇒ (ii). By conditions (b), (i) and Corollary 3.1 we obtain that there exists r_1 such that $(\overline{y}, r) \in \partial_\rho \beta(0)$, for all $r \geq r_1$. From Proposition 2.1 we conclude that E_1 holds for all $r \geq r_1$. Take $r \geq r_1$. We prove now that E_2 holds.

($\subset$) Assume that $\operatorname{argmin}_x \varphi(x)$ is nonempty. Take $x^* \in \operatorname{argmin}_x \varphi(x)$. Then

$$\begin{aligned} l(x^*, \overline{y}, r) &= \inf_z \{f(x^*, z) - \rho(z, \overline{y}, r)\} \leq f(x^*, 0) - \rho(0, \overline{y}, r) = \varphi(x^*) \\ &= \beta(0) = q(\overline{y}, r) = \inf_x l(x, \overline{y}, r), \end{aligned}$$

where the second equality follows from C_1 and the fact that $f(x, 0) = \varphi(x)$ for all $x \in X$, and the fourth equality follows from E_1 (already proved). From these estimates we obtain that $x^* \in \operatorname{argmin}_x l(x, \overline{y}, \overline{r})$. Since x^* is arbitrary, we conclude that the announced inclusion holds.

($\supset$) Consider $r > r_1$ and take $x_r \in \operatorname{argmin}_x l(x, \overline{y}, r)$. We know that E_1 holds, and therefore

$$\begin{aligned} \beta(0) = q(\overline{y}, r) &= \inf_x l(x, \overline{y}, r) = l(x_r, \overline{y}, r) \\ &= \inf_z \{f(x_r, z) - \rho(z, \overline{y}, r)\} \\ &= \lim_{k \to \infty} \{f(x_r, z_k) - \rho(z_k, \overline{y}, r)\} \end{aligned} \tag{7}$$

for some minimizing sequence $\{z_k\}_{k \in \mathbb{N}}$. We analyze two possible cases:

(1) the sequence $\{z_k\}_{k \in \mathbb{N}}$ converges to 0;
(2) the sequence $\{z_k\}_{k \in \mathbb{N}}$ does not converge to 0.

In the first case we get from (7) that

$$\begin{aligned} \beta(0) &= \lim_{k \to \infty} \{f(x_r, z_k) - \rho(z_k, \overline{y}, r)\} \\ &= \liminf_{k \to \infty} \{f(x_r, z_k) - \rho(z_k, \overline{y}, r)\} \\ &\geq f(x_r, 0) - \rho(0, \overline{y}, r) = f(x_r, 0) = \varphi(x_r) \end{aligned}$$

where the inequality follows from (a) and C_1, included in (c), and the third equality also follows from C_1. We conclude that in this case $x_r \in \operatorname{argmin}_x \varphi(x)$. Since x_r is arbitrary, the proof will be complete if we prove that case (2) cannot occur. Suppose by contradiction that case (2) holds. Thus there exist an open neighborhood $V \subset Z$ of 0 and a subsequence $\{z_{k_j}\}_{j\in\mathbb{N}}$, such that $z_{k_j} \in V^c$ for all $j \in \mathbb{N}$. Then,

$$\begin{aligned} f(x_r, z_{k_j}) - \rho(z_{k_j}, \overline{y}, r) &= f(x_r, z_{k_j}) - \rho(z_{k_j}, \overline{y}, r_1) + \rho(z_{k_j}, \overline{y}, r_1) - \rho(z_{k_j}, \overline{y}, r) \\ &\ge \inf_z \{f(x_r, z) - \rho(z, \overline{y}, r_1)\} \\ &\quad + \inf_{z\in V^c} \{\rho(z, \overline{y}, r_1) - \rho(z, \overline{y}, r)\} \\ &\ge q(\overline{y}, r_1) + \inf_{z\in V^c} \{\rho(z, \overline{y}, r_1) - \rho(z, \overline{y}, r)\} \\ &= \beta(0) + \inf_{z\in V^c} \{\rho(z, \overline{y}, r_1) - \rho(z, \overline{y}, r)\}. \end{aligned}$$

Taking limits with $j \to \infty$ in the inequalities above and using (7), we obtain that

$$\beta(0) \ge \beta(0) + \inf_{z\in V^c} \{\rho(z, \overline{y}, r_1) - \rho(z, \overline{y}, r)\}. \tag{8}$$

We have that $\beta(0) \in \mathbb{R}$, because the primal function φ is proper and (d) holds. Since $r > r_1$, (8) contradicts condition C_2(i), included in (c). We conclude that case (2) cannot occur. For completing the proof of the theorem we just need to consider the case $\operatorname{argmin}_x \varphi(x) = \emptyset$. In this case the inclusion $\operatorname{argmin}_x \varphi(x) \subset \operatorname{argmin}_x l(x, \bar{y}, r)$ trivially holds. We need to prove that $\operatorname{argmin}_x l(x, \bar{y}, r) = \emptyset$ for every $r > r_1$. Supposing by contradiction that this set is nonempty, we can repeat the second part of the proof above and conclude that $\emptyset \ne \operatorname{argmin}_x l(x, \bar{y}, r) \subset \operatorname{argmin}_x \varphi(x)$, which is a contradiction, because $\operatorname{argmin}_x \varphi(x)$ is empty. The proof is complete. □

Remark 3.2 We mention that the assumption (a) in Theorem 3.3 does not imply assumption (b). Indeed, consider a primal problem $\min_{x\in\mathbb{R}} \varphi(x)$, where $\varphi(x) = x^2$. Let a continuous parameterization function $f(x, z) = x^2 + zx^3$. It follows that assumption (a) in Theorem 3.3 holds, but assumption (b) does not hold, because $\beta(0) = 0$ and $\beta(z) = -\infty$ for all $z \ne 0$. Proposition 4.1 presents some assumptions under which β is lsc at 0.

4 Sub-optimal Path

To obtain an exact solution of an optimization problem may, in general, be very hard or even impossible. However, when the optimal value of the problem is finite, approximate solutions always exist and they are, in principle, easier to find than exact solutions. In Wang et al. [16], the authors defined a sub-optimal path related with the dual problem and established some convergence results in finite dimensional spaces. In this section we consider a optimal path related to our duality scheme and analyze its convergence properties. This result is related to Burachik and Rubinov [4, Theorem 6.1], where the authors consider an optimal path in the sense that all the subproblems are supposed to be solved exactly. Also, as we will see in Sect. 5 that our duality scheme includes the one considered in Burachik and Rubinov [4].

Recall that the calculation of the dual function leads to the following problem:

$$\inf\{f(x,z)-\rho(z,y,r):(x,z)\in X\times Z\}. \tag{9}$$

Definition 4.1 Let $I\subset\mathbb{R}_+$ be unbounded above, and for each $r\in I$ take $\varepsilon_r\geq 0$. The set $\{(x_r,z_r)\}_{r\in I}\subset X\times Z$ is called a *sub-optimal path of problem* (9) iff

$$f(x_r,z_r)-\rho(z_r,y,r)\leq q(y,r)+\varepsilon_r \tag{10}$$

for all $r\in I$. When (x_r,z_r) satisfies (10) with $\varepsilon_r=0$ for all $r\in I$, the set $\{(x_r,z_r)\}_{r\in I}$ is called an *optimal path*.

In next theorem we analyze limit points of sub-optimal paths, where $\{\varepsilon_r\}_{r\in I}$ is assumed to satisfy $\lim_{r\to\infty}\varepsilon_r=0$.

Theorem 4.1 *Assume that*

(a) *there exists* $(\overline{y},\overline{r})\in\operatorname{dom}\beta^{\rho}$, *and conditions* C_1 *and* C_2 *hold*;
(b) *the parameterization function* f *is* lsc *at* $(x,0)$ *for each* $x\in X$, *and there exist an open neighborhood* $W\subset Z$ *of* 0, *a real number* $\alpha>\beta(0)$, *and a compact subset* $B\subset X$ *such that*

$$L_{f,W}(\alpha):=\{x\in X: f(x,z)\leq\alpha\}\subset B,\quad \textit{for all } z\in W.$$

Then

(i) *there exists a sub-optimal path* $\{(x_r,z_r)\}_{r\geq\overline{r}}$.
(ii) *Take a set* $I\subset\mathbb{R}_+$ *unbounded above and consider a sub-optimal path* $\{(x_r,z_r)\}_{r\in I}$ *satisfying* $\lim_{r\in I,r\to\infty}\varepsilon_r=0$. *Then* $\{z_r\}_{r\in I}$ *converges to* 0, *and the set of cluster points of* $\{x_r\}_{r\in I}$ *is a nonempty set contained in the primal optimal solution set.*

Proof Since $\rho(z,y,\cdot)$ is a nonincreasing function, we have that $q(\overline{y},\cdot)$ is nondecreasing. Thus, if $r\geq\overline{r}$ then $q(\overline{y},r)>-\infty$, by item (a) and Proposition 2.1. Thus the existence of a sub-optimal path is trivially ensured, which proves (i).

For proving (ii), let $\{(x_r,z_r)\}_{r\in I}$ be a sub-optimal path. Assume that $\lim_{r\in I,r\to\infty}\varepsilon_r=0$. Suppose by contradiction that $\{z_r\}_{r\in I}$ does not converge to 0 when $r\to\infty$. Thus there exist an open neighborhood $V\subset Z$ of 0 and $J\subset I$, unbounded above, such that $\{z_r\}_{r\in J}\subset V^c$ (for instance, we can take $J_k:=I\cap[k,\infty)$, for $k\in\mathbb{N}$, and hence there exists $r_k\in J_k$ such that $z_{r_k}\in V^c$; then $J=\{r_k\}_k$ is unbounded above). Therefore, we have

$$\begin{aligned}\beta(0)+\varepsilon_r\geq q(\overline{y},r)+\varepsilon_r&\geq f(x_r,z_r)-\rho(z_r,\overline{y},r)\\&=f(x_r,z_r)-\rho(z_r,\overline{y},\overline{r})+\rho(z_r,\overline{y},\overline{r})-\rho(z_r,\overline{y},r)\\&\geq q(\overline{y},\overline{r})+\inf_{z\in V^c}\{\rho(z,\overline{y},\overline{r})-\rho(z,\overline{y},r)\}.\end{aligned}$$

Since $\lim_{r\in I, r\to\infty} \varepsilon_r = 0$, we conclude that there exists $r_0 \in J$ such that for all $r \geq r_0, r \in J$, we have

$$\beta(0) + 1 - q(\overline{y}, \overline{r}) \geq \inf_{z\in V^c} \{\rho(z, \overline{y}, \overline{r}) - \rho(z, \overline{y}, r)\},$$

which contradicts C_2, because J is unbounded above and $\beta(0), q(\overline{y}, \overline{r}) \in \mathbb{R}$. It follows that $\{z_r\}_{r\in I}$ converges to 0. Consider an open neighborhood $W \subset Z$ of 0 and $\alpha > \beta(0)$ as in assumption (b). Since $\{z_r\}_{r\in I}$ converges to 0, there exists $r_0 \in I$ such that $\{z_r\}_{r\geq r_0, r\in I} \subset W$. Take $t := \alpha - \beta(0) > 0$. The function $\rho(\cdot, \overline{y}, \overline{r})$ is upper semi-continuous at 0 by condition C_1. Thus there exists some $r_1 \geq \max\{r_0, \overline{r}\}$ such that $\rho(z_r, \overline{y}, \overline{r}) \leq \frac{t}{2}$ and $\varepsilon_r \leq \frac{t}{2}$ for all $r \geq r_1, r \in I$. Therefore, for all $r \geq r_1, r \in I$,

$$\begin{aligned}\beta(0) + \frac{t}{2} &\geq q(\overline{y}, r) + \varepsilon_r \geq f(x_r, z_r) - \rho(z_r, \overline{y}, r)\\ &\geq f(x_r, z_r) - \rho(z_r, \overline{y}, \overline{r}) \geq f(x_r, z_r) - \frac{t}{2}.\end{aligned}$$

Hence

$$f(x_r, z_r) \leq \beta(0) + t = \alpha, \quad \text{for all } r \geq r_1, r \in I,$$

that is to say $\{x_r\}_{r\geq r_1, r\in I} \subset L_{f,W}(\alpha)$. Assumption (b) implies that $\{x_r\}_{r\geq r_1, r\in I} \subset B$, where B is a compact set. In particular, since $\{z_r\}_{r\in I}$ converges to 0, the set of cluster points of the sub-optimal path $\{(x_r, z_r) : r \in I\}$ is nonempty. Moreover, every cluster point has the form $(x^*, 0)$. Let us prove that x^* is a primal optimal solution, where x^* is an arbitrary cluster point of $\{x_r\}_{r\in I}$. Take a subnet $\{x_{r_j}\}_{j\in J}$ converging to x^*, and $j_0 \in J$ satisfying $r_j \geq \overline{r}$ for all $j \geq j_0$, $j \in J$. Observe that $\{z_{r_j}\}_{j\in J}$ converges to 0. Thus

$$\begin{aligned}\beta(0) + \varepsilon_{r_j} &\geq q(\overline{y}, r_j) + \varepsilon_{r_j}\\ &\geq f(x_{r_j}, z_{r_j}) - \rho(z_{r_j}, \overline{y}, r_j)\\ &\geq f(x_{r_j}, z_{r_j}) - \rho(z_{r_j}, \overline{y}, \overline{r})\end{aligned}$$

for all $j \geq j_0$, $j \in J$. If we take the $\liminf_{j\in J}$ in these inequalities, we obtain

$$\beta(0) \geq f(x^*, 0) - \rho(0, \overline{y}, \overline{r}) = f(x^*, 0) = \varphi(x^*),$$

using conditions (b) and C_1. Thus x^* is a primal solution. The theorem is proved. □

Remark 4.1 In connection with the compactness assumption of Theorem 4.1, we mention that when X is an infinite dimensional reflexive Banach space with the weak topology (which is not metrizable), Banach-Alaoglu's Theorem implies that a set is weakly compact if and only if it is bounded and weakly closed. In particular, closed balls (in the strong topology) are weakly compact in such spaces. Thus, a parameterization function f such that some sub-level set of $f(\cdot, z)$ is uniformly bounded and weakly closed when z runs over a neighborhood of 0, provides an example for which assumption (b) of Theorem 4.1 holds. This situation is indeed a prototypical and nontrivial case to which Theorem 4.1 applies. We remind also that sub-level sets of convex and lsc functions are always weakly closed, so that in the convex case it suffices to check the uniform boundedness of the sub-level sets of $f(\cdot, z)$.

Remark 4.2 Theorem 4.1 is related to Burachik and Rubinov [4, Theorem 6.1], where the authors considered an optimal path (in a reflexive Banach space) instead of a sub-optimal path, and the compactness assumption on the sub-level sets of $f(\cdot, z)$ is assumed locally at all z, instead of just at $z = 0$, as assumed in the present paper. Also, in Burachik and Rubinov [4, Theorem 6.1] it is assumed that the compactness property holds for all sub-level sets of $f(\cdot, z)$, while Theorem 4.1 assumes compactness of just one of them, corresponding to $\alpha > \beta(0)$. Since we are not assuming convexity of $f(\cdot, z)$, compactness of just one sub-level set of $f(\cdot, z)$ is not equivalent to compactness of all of them.

Proposition 4.1 *Let $f : X \times Z \to \mathbb{R}$ be a function* lsc *at $(x, 0)$ for each $x \in X$. Take $\beta(z) := \inf_x f(x, z)$. Suppose that $\beta(0) \in \mathbb{R}$ and that there exist an open neighborhood $W \subset Z$ of 0, $\alpha \geq \beta(0)$ and a compact subset $B \subset X$, such that*

$$L_{f,W}(\alpha) := \{x \in X : f(x, z) \leq \alpha\} \subset B, \quad \textit{for all } z \in W.$$

Then the perturbation function β is lsc *at* 0.

Proof Let J be the set of all neighborhoods of 0. We know that J is a directed set with the partial order $V_1 \geq V_2$ iff $V_1 \subset V_2$. Suppose by contradiction that β is not lsc at 0. Then there exists $\varepsilon > 0$ such that

$$\sup_{V \in J} \inf_{v \in V} \beta(v) < \beta(0) - \varepsilon.$$

Thus, $\inf_{v \in V} \beta(v) < \beta(0) - \varepsilon$ for all $V \in J$. In particular for each $V \in J$ there exists $z_V \in V$ such that $\beta(z_V) < \beta(0) - \varepsilon$, which in turn implies that for each $V \in J$ there exists $x_V \in X$ satisfying

$$f(x_V, z_V) < \beta(0) - \varepsilon. \tag{11}$$

By construction the net $\{z_V\}_{V \in J}$ converges to 0. Taking W and α as in the assumption, it follows that $I := \{V \in J : V \geq W\}$ is a terminal subset of J such that $\{z_V\}_{V \in I} \subset W$ and $\{x_V\}_{V \in I} \subset L_{f,W}(\alpha) \subset B$, where B is the compact set given by hypothesis. Hence there exists a subnet $\{\eta_s\}_{s \in S}$ of $\{x_V\}_{V \in I}$ convergent to some $\overline{x}$. This means that $\eta_s = x_{g(s)}$, where S is a directed set and $g : S \to I$ is a function such that for every $U \in I$ there exists an $s_U \in S$ satisfying $g(s) \geq U$ for all $s \geq s_U$, $s \in S$. In particular, the set $\{t_s\}_{s \in S}$, where $t_s := z_{g(s)}$ for all $s \in S$, is a subnet of $\{z_V\}_{V \in I}$ converging to 0, and $f(\eta_s, t_s) < \beta(0) - \varepsilon$ for all $s \in S$, by (11). Therefore, using the lower semicontinuity of f in $(\overline{x}, 0)$ we obtain

$$\beta(0) \leq f(\overline{x}, 0) \leq \liminf_{s \in S} f(\eta_s, t_s) \leq \beta(0) - \varepsilon,$$

entailing a contradiction. □

Example 4.1 Consider the following constrained optimization problem

$$\text{minimize}\, h(x) \quad \text{subject to} \quad x \text{ in } C, \tag{12}$$

where $h : X \to \mathbb{R}$ is a lsc function such that $L_\alpha := \{x \in X : h(x) \leq \alpha\}$ is compact for some $\alpha > \inf_{x \in C} h(x)$, and C is a closed subset of X. Take a mapping $D : Z \rightrightarrows X$ such that $D(0) = C$ and suppose that D has a closed graph, that is, the set $\{(z, u) : u \in D(z), z \in Z\}$ is closed (in the case that $C := \{x : g_j(x) \leq 0,\ j = 1, \dots, m\}$, where $g_j : X \to \mathbb{R}$ is lsc for $j = 1, \dots, m$, a canonical such mapping is $D(z) = \{x : g_j(x) \leq z_j,\ j = 1, \dots, m\}$). A canonical dualizing parameterization function for problem (12) is $f(x, z) = h(x) + \delta_{D(z)}(x)$, where $\delta_V(x) = 0$ if $x \in V$ and $\delta_V(x) = \infty$ otherwise. It is not difficult to see that f satisfies the assumptions of Proposition 4.1. Thus the perturbation function $\beta(z) = \inf_{x \in D(z)} h(x)$ is lsc at 0. See also Proposition 5.2 in Burachik and Rubinov [4], where a similar result is stated.

Next, we show some examples of general augmented Lagrangians and compare our setting with the ones considered in Burachik and Rubinov [4] and Wang et al. [7, Sect. 3.1].

5 Augmented Lagrangians

Consider a coupling function $p : Z \times Y \times \mathbb{R}_{++} \to \mathbb{R}$ such that $p(z, y, \cdot)$ is differentiable. A valley-at-zero type property of $p'_r(z, y, \cdot)$ was introduced in Wang et al. [7, Sect. 3.1] where X, Y and Z are finite dimensional vector spaces. In addition, the primal problem is an inequality constrained problem, which is a particular case of the primal problem considered in the present paper. We state next the valley-at-zero property given in Wang et al. [7, Sect. 3.1] in our general setting.

(A_1) There exists $\alpha \in [0, 1)$ such that, for every open neighborhood $V \subset Z$ of 0, and $y \in Y$,

$$M_{V,\varepsilon} := \inf_{u \in V^c, \tau \geq \varepsilon} \tau^\alpha p'_r(u, y, \tau) > 0$$

for all $\varepsilon > 0$.

Remark 5.1 Wang et al. [7, Sect. 3.1] also assume that $p'_r(0, y, r) = 0$. We do not assume this condition. Regarding our condition C_1, it is a standard assumption, used also in [4] and [7]. Therefore, is enough for us to study the relationship between our condition C_2 and related assumptions in the aforementioned papers.

Wang et al. [7] use as coupling function $\rho := -p$ in the construction of the Lagrangian scheme.

Proposition 5.1 *Take a function p satisfying Condition* A_1. *Then the function $\rho := -p$ satisfies Condition* C_2.

Proof Fix an open neighborhood $V \subset Z$ of 0, $y \in Y$, and $\hat{r} > 0$. For every $z \in V^c$ and $r > \hat{r}$ there exists $\theta_r \in (\hat{r}, r)$ such that

$$p(z, y, r) - p(z, y, \hat{r}) = p'_r(z, y, \theta_r)(r - \hat{r}) \geq (r^{1-\alpha} - \hat{r}^{1-\alpha})\theta_r^{\alpha} p'_r(z, y, \theta_r), \quad (13)$$

where the inequality follows from the following estimates:

$$r > \theta_r \quad \Rightarrow \quad r = r^{1-\alpha} r^{\alpha} \geq r^{1-\alpha}\theta_r^{\alpha},$$

where $\alpha \in [0, 1)$ is given by A_1; analogously we have $\hat{r} = \hat{r}^{1-\alpha}\hat{r}^{\alpha} \leq \hat{r}^{1-\alpha}\theta_r^{\alpha}$. Thus we get

$$r - \hat{r} \geq r^{1-\alpha}\theta_r^{\alpha} - \hat{r}^{1-\alpha}\theta_r^{\alpha} = \theta_r^{\alpha}(r^{1-\alpha} - \hat{r}^{1-\alpha}).$$

Take $0 < \varepsilon < \hat{r}$. From (13) we obtain

$$\begin{aligned} p(z, y, r) - p(z, y, \hat{r}) &\geq (r^{1-\alpha} - \hat{r}^{1-\alpha}) \inf_{\tau \geq \varepsilon} \tau^{\alpha} p'_r(z, y, \tau) \\ &\geq (r^{1-\alpha} - \hat{r}^{1-\alpha}) \inf_{u \in V^c, \tau \geq \varepsilon} \tau^{\alpha} p'_r(u, y, \tau) \\ &= (r^{1-\alpha} - \hat{r}^{1-\alpha}) M_{V,\varepsilon} \end{aligned}$$

for all $z \in V^c$. Therefore

$$\inf_{z \in V^c} \{p(z, y, r) - p(z, y, \hat{r})\} \geq (r^{1-\alpha} - \hat{r}^{1-\alpha}) M_{V,\varepsilon}.$$

It is easy to see that C_2 follows from the last estimate above and A_1, observing that $\rho = -p$ and $\alpha \in [0, 1)$. □

The above result shows that our setting is more general than the one considered in Wang et al. [7]. In order to show that our setting is more general than the one considered in Burachik and Rubinov [4], we recall next their main assumptions.

Consider a function $s : \mathbb{R}^2 \to \mathbb{R}$ such that $s(0, 0) = 0$ and for every $a \in \mathbb{R}$ and $b_1 \geq b_2$, it satisfies

$$s(a, b_1) - s(a, b_2) \geq \psi(b_1 - b_2), \quad (14)$$

where $\psi : \mathbb{R}_+ \to \mathbb{R}_+$ is a strictly increasing function such that $\psi(0) = 0$ and ψ is coercive, that is, $\lim_{t\to\infty} \psi(t) = \infty$.

Let $\{\nu_\eta\}_{\eta \in U_1}$ be a family of upper semicontinuous functions satisfying

$$\nu_\eta(0) = 0 \quad \text{for all } \eta \in U_1, \quad (15)$$

and $\{\sigma_\mu\}_{\mu \in U_2}$ be a family of augmenting functions which have a valley-at-zero property, that is, for every $\mu \in U_2$, $\sigma_\mu : Z \to \mathbb{R}_+$ is proper, lsc and satisfies

$$\sigma_\mu(0) = 0 \quad \text{and} \quad \inf_{z \in V^c} \sigma_\mu(z) > 0, \quad (16)$$

for every open neighborhood $V \subset Z$ of 0.

The coupling function considered in Burachik and Rubinov [4] is ρ, given by

$$\rho(z,(\eta,\mu),r)=s(\nu_\eta(z),-r\sigma_\mu(z))$$

where s, $\{\nu_\eta\}_{\eta\in U_1}$ and $\{\sigma_\mu\}_{\mu\in U_2}$ satisfy (14)–(16).

Since we are not supposing any structure on the set Y, we can consider $Y := U_1 \times U_2$. In next proposition we show that our primal-dual scheme includes the one in Burachik and Rubinov [4].

Proposition 5.2 *Let $\rho(z,y,r) := s(\nu_\eta(z),-r\sigma_\mu(z))$, where $y=(\eta,\mu)$ and the functions s, $\{\nu_\eta\}_{\eta\in U_1}$ and $\{\sigma_\mu\}_{\mu\in U_2}$ satisfy* (14)–(16). *Then condition C_2 is satisfied.*

Proof Fix an open neighborhood $V\subset Z$ of 0, $y\in Y$ and $\bar r>0$. For all $r>\bar r$ and $z\in V^c$ we have

$$\begin{aligned}\rho(z,y,\bar r)-\rho(z,y,r)&=s(\nu_\eta(z),-\bar r\sigma_\mu(z))-s(\nu_\eta(z),-r\sigma_\mu(z))\\ &\geq\psi((r-\bar r)\sigma_\mu(z))\\ &\geq\psi((r-\bar r)M_V),\end{aligned}$$

where the first inequality follows from the property of the function s, and the second inequality follows from the fact that ψ is increasing and $M_V := \inf_{u\in V^c}\sigma_\mu(u)>0$. It follows that

$$\inf_{z\in V^c}\rho(z,y,\bar r)-\rho(z,y,r)\geq\psi((r-\bar r)M_V).$$

Using this last estimate and the property of the function ψ, we conclude that Condition C_2 is satisfied. □

Remark 5.2 The coercivity property $\lim_{t\to\infty}\psi(t)=\infty$ was not explicitly required in Burachik and Rubinov [4], but it was used in the proof of Burachik and Rubinov [4, Theorem 4.1], and this theorem is applied throughout the paper.

Example 5.1 Let Z be a reflexive Banach space. Take a coupling function $g: Y\times Z\to\mathbb{R}$ such that $g(y,\cdot)$ is weakly upper semicontinuous and $g(y,0)=0$ for each $y\in Y$. Let $\rho(z,y,r) := g(y,z)-r\sigma(z)$, where σ is an augmenting function with a valley-at-zero (i.e., σ satisfies (16)). In this case, we recover the augmented Lagrangian studied in Zhou and Yang [14]:

$$\ell(x,y,r)=\inf_z\{\phi(x,z)-g(y,z)+r\sigma(z)\}.$$

Example 5.2 Let Z be a Hilbert space. Consider a continuous and invertible map $A: Z\to Z$, and suppose that $Y=Z$. Let the coupling function ρ be defined by $\rho(z,y,r)=\langle y,Az\rangle-r\sigma(Az)$, where $\sigma: Z\to\mathbb{R}$ is an augmenting function, i.e. a proper, lsc and convex function satisfying:

$$\sigma(0)=0\quad\text{and}\quad\operatorname{Argmin}\sigma=\{0\}.$$

In this context our general augmented Lagrangian is the A-augmented Lagrangian proposed and studied in Yang and Zhang [17]:

$$\ell_A(x, y, r) = \inf_{z \in Z} \{\phi(x, z) - \langle y, Az \rangle + r\sigma(Az)\}.$$

The A-augmented Lagrangian was studied in finite dimensional space, and some additional conditions are imposed on the mapping A, see Yang and Zhang [17]. In particular, when $A = I$, that is, $Az = z$ for all $z \in Z$, we recover the classical augmented Lagrangian proposed in Rockafellar and Wets [1, Chap. 11], which is also an example of the augmented Lagrangians proposed in Burachik and Rubinov [4].

References

1. Rockafellar, R.T., Wets, R.J.-B.: Variational Analysis. Springer, Berlin (1998)
2. Huang, X.X., Yang, X.Q.: A unified augmented Lagrangian approach to duality and exact penalization. Math. Oper. Res. **28**, 533–552 (2003)
3. Nedić, A., Ozdaglar, A.: A geometric framework for nonconvex optimization duality using augmented Lagrangian functions. J. Global Optim. **40**, 545–573 (2008)
4. Burachik, R.S., Rubinov, A.M.: Abstract convexity and augmented Lagrangians. SIAM J. Optim. **18**, 413–436 (2007)
5. Rubinov, A.M., Huang, X.X., Yang, X.Q.: The zero duality gap property and lower semicontinuity of the perturbation function. Math. Oper. Res. **27**, 775–791 (2002)
6. Zhou, Y.Y., Yang, X.Q.: Augmented Lagrangian function, non-quadratic growth condition and exact penalization. Oper. Res. Lett. **34**, 127–134 (2006)
7. Wang, C.Y., Yang, X.Q., Yang, X.M.: Unified nonlinear Lagrangian approach to duality and optimal paths. J. Optim. Theory Appl. **135**, 85–100 (2007)
8. Penot, J.P., Rubinov, A.M.: Multipliers and general Lagrangians. Optimization **54**, 443–467 (2005)
9. Nedić, A., Ozdaglar, A., Rubinov, A.M.: Abstract convexity for nonconvex optimization duality. Optimization **56**, 655–674 (2007)
10. Rubinov, A.M., Yang, X.Q.: Lagrange-type Functions in Constrained Non-convex Optimization. Kluwer Academic, Dordrecht (2003)
11. Burachik, R.S., Rubinov, A.M.: On the absence of duality gap for Lagrange-type functions. J. Ind. Manage. Optim. **1**, 33–38 (2005)
12. Oettli, W., Yang, X.Q.: Modified Lagrangian and least root approaches for general nonlinear optimization problems. Acta Math. Appl. Sin. Engl. Ser. **18**, 147–152 (2002)
13. Wang, C.Y., Yang, X.Q., Yang, X.M.: Nonlinear Lagrange duality theorems and penalty function methods in continuous optimization. J. Global Optim. **27**, 473–484 (2003)
14. Zhou, Y.Y., Yang, X.Q.: Duality and penalization in optimization via an augmented Lagrangian function with applications. J. Optim. Theory Appl. **140**, 171–188 (2009)
15. Rubinov, A.M.: Abstract Convexity and Global Optimization. Kluwer Academic, Dordrecht (2000)
16. Wang, C.Y., Yang, X.Q., Yang, X.M.: Unified nonlinear Lagrangian approach to duality and optimal paths. J. Optim. Theory Appl. **135**(1), 85–100 (2007)
17. Zhang, L., Yang, X.Q.: An augmented Lagrangian approach with a variable transformation in nonlinear programming. Nonlinear Anal. **69**, 2095–2113 (2008)