



**HAL**  
open science

# Fenchel-Moreau Conjugation Inequalities with Three Couplings and Application to Stochastic Bellman Equation

Jean-Philippe Chancelier, Michel de Lara

► **To cite this version:**

Jean-Philippe Chancelier, Michel de Lara. Fenchel-Moreau Conjugation Inequalities with Three Couplings and Application to Stochastic Bellman Equation. 2018. hal-01760462v1

**HAL Id: hal-01760462**

**<https://hal.science/hal-01760462v1>**

Preprint submitted on 6 Apr 2018 (v1), last revised 7 Sep 2018 (v2)

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Fenchel-Moreau Conjugation Inequalities with Three Couplings and Application to Stochastic Bellman Equation

Jean-Philippe Chancelier and Michel De Lara\*  
Université Paris-Est, CERMICS (ENPC)

April 6, 2018

## Abstract

Given two couplings between “primal” and “dual” sets, we prove a general implication that relates an inequality involving “primal” sets to a reverse inequality involving the “dual” sets. More precisely, let be given two “primal” sets  $\mathbb{C}, \mathbb{D}$  and two “dual” sets  $\mathbb{C}^\sharp, \mathbb{D}^\sharp$ , together with two coupling functions  $\mathbb{C} \overset{\Phi}{\leftrightarrow} \mathbb{C}^\sharp$  and  $\mathbb{D} \overset{\Psi}{\leftrightarrow} \mathbb{D}^\sharp$ . We define a new coupling  $\Phi \dagger \Psi$  between the “primal” product set  $\mathbb{C} \times \mathbb{D}$  and the “dual” product set  $\mathbb{C}^\sharp \times \mathbb{D}^\sharp$ . Then, we consider any bivariate function  $\mathcal{K} : \mathbb{C} \times \mathbb{D} \rightarrow [-\infty, +\infty]$  and univariate functions  $f : \mathbb{C} \rightarrow [-\infty, +\infty]$  and  $g : \mathbb{D} \rightarrow [-\infty, +\infty]$ , all defined on the “primal” sets. We prove that  $f(c) \geq \inf_{d \in \mathbb{D}} (\mathcal{K}(c, d) \dagger g(d)) \Rightarrow f^\Phi(c^\sharp) \leq \inf_{d^\sharp \in \mathbb{D}^\sharp} (\mathcal{K}^{\Phi \dagger \Psi}(c^\sharp, d^\sharp) \dagger g^{-\Psi}(d^\sharp))$ , where we stress that the Fenchel-Moreau conjugates  $f^\Phi$  and  $g^{-\Psi}$  are not necessarily taken with the same coupling. We study the equality case, after having established the classical Fenchel inequality but with a general coupling. We display several applications. We provide a new formula for the Fenchel-Moreau conjugate of a generalized inf-convolution. We obtain formulas with partial Fenchel-Moreau conjugates. Finally, we consider the Bellman equation in stochastic dynamic programming and we provide a “Bellman-like” equation for the Fenchel conjugates of the value functions.

## 1 Introduction

In convex analysis, the Fenchel conjugate plays a central part. It is involved in many equalities and inequalities, like the well known Fenchel (in)equalities or the Fenchel conjugate of an inf-convolution. The classical Fenchel conjugate was extended by J. J. Moreau [10], by replacing the bilinear pairing, between a vector space and its algebraic dual, with a more

---

\*delara@cermics.enpc.fr

general coupling. This gives the so-called Fenchel-Moreau conjugate (see Chapter 11L and the Commentary in [16] with a brief historical perspective and references). In abstract convexity [17, 20, 11, 8], affine functions are replaced by another class of functions (related to the coupling), and so are convex functions (replaced by so-called abstract convex functions), by taking the supremum. In this way, generalized Fenchel conjugation formulas are obtained, as well as duality for abstract convex functions. Generalized Fenchel conjugation also appears in the dual formulation of optimal transport problems [6, 18].

In this paper, we provide a main Fenchel-Moreau conjugation inequality with three couplings, and applications. In Sect. 2, we establish our main inequality. Then, we provide sufficient conditions for the equality case, after having established the classical Fenchel inequality, but with a general coupling. In Sect. 3, we display several applications. First, we provide a definition of a generalized inf-convolution, and new formulas for its Fenchel-Moreau conjugate (inequality and equality). Second, we obtain formulas with partial Fenchel-Moreau conjugates. Finally, we consider the Bellman equation in stochastic dynamic programming and we provide a “Bellman-like” equation for the Fenchel conjugates of the value functions.

## 2 Duality inequality with three Fenchel-Moreau conjugates

Given two couplings between “primal” and “dual” sets, we prove a general implication that relates an inequality involving “primal” sets to a reverse inequality involving the “dual” sets.

In what follows, we rely upon background on J. J. Moreau lower and upper additions and on Fenchel-Moreau conjugacy with respect to a coupling, that can be found in Appendix A.

### 2.1 Main duality inequality

Let be given two “primal” sets  $\mathbb{C}, \mathbb{D}$  and two “dual” sets  $\mathbb{C}^\sharp, \mathbb{D}^\sharp$ , together with two *coupling* functions

$$\Phi : \mathbb{C} \times \mathbb{C}^\sharp \rightarrow [-\infty, +\infty], \quad \Psi : \mathbb{D} \times \mathbb{D}^\sharp \rightarrow [-\infty, +\infty]. \quad (1)$$

We will call  $\mathbb{C}$  and  $\mathbb{D}$  “primal” sets, whereas  $\mathbb{C}^\sharp$  and  $\mathbb{D}^\sharp$  are “dual” sets.

We define the *sum coupling*  $\Phi \dot{+} \Psi$  — coupling the “primal” product set  $\mathbb{C} \times \mathbb{D}$  with the “dual” product set  $\mathbb{C}^\sharp \times \mathbb{D}^\sharp$  — by

$$\Phi \dot{+} \Psi : (\mathbb{C} \times \mathbb{D}) \times (\mathbb{C}^\sharp \times \mathbb{D}^\sharp) \rightarrow [-\infty, +\infty], \quad (2a)$$

$$((c, d), (c^\sharp, d^\sharp)) \mapsto \Phi(c, c^\sharp) \dot{+} \Psi(d, d^\sharp). \quad (2b)$$

With any bivariate function  $\mathcal{K} : \mathbb{C} \times \mathbb{D} \rightarrow [-\infty, +\infty]$ , defined on the “primal” product set  $\mathbb{C} \times \mathbb{D}$ , we associate the conjugate, with respect to the coupling  $\Phi \dot{+} \Psi$ , defined on the

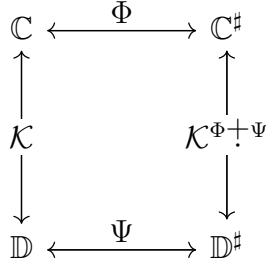


Figure 1: A kernel  $\mathcal{K}$ , two couplings  $\Phi$  and  $\Psi$ , and a new kernel

“dual” product set  $\mathbb{C}^\sharp \times \mathbb{D}^\sharp$ , by:

$$\mathcal{K}^{\Phi \dagger \Psi}(c^\sharp, d^\sharp) = \sup_{c \in \mathbb{C}, d \in \mathbb{D}} \left( (\Phi \dagger \Psi)((c, d), (c^\sharp, d^\sharp)) \dagger (-\mathcal{K}(c, d)) \right) \quad (3a)$$

$$\begin{aligned}
&= \sup_{c \in \mathbb{C}, d \in \mathbb{D}} \left( \Phi(c, c^\sharp) \dagger \Psi(d, d^\sharp) \dagger (-\mathcal{K}(c, d)) \right) \quad (3b) \\
&\quad \forall (c^\sharp, d^\sharp) \in \mathbb{C}^\sharp \times \mathbb{D}^\sharp.
\end{aligned}$$

In what follows, we will call the function  $\mathcal{K}$  a *kernel* (or a *potential*). Indeed, consider the expression in the left hand side assumption in (4). If we translate it from the  $(\min, +)$  algebra to the usual  $(+, \times)$  algebra, it stands as an integration with respect to a kernel.

**Theorem 1** *For any bivariate function  $\mathcal{K} : \mathbb{C} \times \mathbb{D} \rightarrow [-\infty, +\infty]$  and univariate functions  $f : \mathbb{C} \rightarrow [-\infty, +\infty]$  and  $g : \mathbb{D} \rightarrow [-\infty, +\infty]$ , all defined on the “primal” sets, we have that*

$$\begin{aligned}
f(c) &\geq \inf_{d \in \mathbb{D}} \left( \mathcal{K}(c, d) \dagger g(d) \right), \quad \forall c \in \mathbb{C} \Rightarrow \\
f^\Phi(c^\sharp) &\leq \inf_{d^\sharp \in \mathbb{D}^\sharp} \left( \mathcal{K}^{\Phi \dagger \Psi}(c^\sharp, d^\sharp) \dagger g^{-\Psi}(d^\sharp) \right), \quad \forall c^\sharp \in \mathbb{C}^\sharp. \quad (4)
\end{aligned}$$

Notice that the left hand side assumption in (4) is a rather weak inequality (upper bound for an infimum), whereas the right hand side assumption in (4) is a rather strong inequality (lower bound for an infimum).

**Proof.**

$$f^\Phi(c^\sharp) = \sup_{c \in \mathbb{C}} \left( \Phi(c, c^\sharp) \dagger (-f(c)) \right)$$

by definition (54) of the conjugate  $f^\Phi(c^\sharp)$

$$\leq \sup_{c \in \mathbb{C}} \left( \Phi(c, c^\sharp) \dagger \left( - \inf_{d \in \mathbb{D}} (\mathcal{K}(c, d) \dagger g(d)) \right) \right)$$

by the left hand side assumption in (4) and by the property (50b) that the operator  $\dot{+}$  is monotone [this inequality is an equality when the left hand side assumption in (4) is an equality]

$$= \sup_{c \in \mathbb{C}} \left( \Phi(c, c^\sharp) \dot{+} \sup_{d \in \mathbb{D}} \left( - (\mathcal{K}(c, d) \dot{+} g(d)) \right) \right)$$

by  $-\inf = \sup -$

$$= \sup_{c \in \mathbb{C}, d \in \mathbb{D}} \left( \Phi(c, c^\sharp) \dot{+} \left( - (\mathcal{K}(c, d) \dot{+} g(d)) \right) \right)$$

by the property (50e) that the operator  $\sup$  is linear in  $\dot{+}$

$$\leq \sup_{c \in \mathbb{C}, d \in \mathbb{D}} \left( \Phi(c, c^\sharp) \dot{+} \left( - (\mathcal{K}(c, d) \dot{+} g^{(-\Psi)(-\Psi)}(d)) \right) \right)$$

because  $g^{(-\Psi)(-\Psi)} \leq g$  by (56) and by the property (50b) that the operator  $\dot{+}$  is monotone [this inequality is an equality when  $g^{(-\Psi)(-\Psi)} = g$ ]

$$= \sup_{c \in \mathbb{C}, d \in \mathbb{D}} \left( \Phi(c, c^\sharp) \dot{+} \left( - \left( \mathcal{K}(c, d) \dot{+} \left( \sup_{d^\sharp \in \mathbb{D}^\sharp} \left( (-\Psi(d, d^\sharp)) \dot{+} (-g^{-\Psi}(d^\sharp)) \right) \right) \right) \right)$$

by definition (55) of the biconjugate  $g^{(-\Psi)(-\Psi)}$

$$\leq \sup_{c \in \mathbb{C}, d \in \mathbb{D}} \left( \Phi(c, c^\sharp) \dot{+} \left( - \sup_{d^\sharp \in \mathbb{D}^\sharp} \left( \mathcal{K}(c, d) \dot{+} (-\Psi(d, d^\sharp) \dot{+} (-g^{-\Psi}(d^\sharp))) \right) \right) \right)$$

by the property (51f) that the operator  $\sup$  is sublinear in  $\dot{+}$ , and by the property (51b) that the operator  $\dot{+}$  is monotone [this inequality is an equality when  $-\infty < \mathcal{K}(c, d)$  by (51g)]

$$= \sup_{c \in \mathbb{C}, d \in \mathbb{D}} \left( \Phi(c, c^\sharp) \dot{+} \inf_{d^\sharp \in \mathbb{D}^\sharp} - \left( \mathcal{K}(c, d) \dot{+} \left( (-\Psi(d, d^\sharp)) \dot{+} (-g^{-\Psi}(d^\sharp)) \right) \right) \right)$$

by  $-\sup = \inf -$

$$\leq \sup_{c \in \mathbb{C}, d \in \mathbb{D}} \inf_{d^\sharp \in \mathbb{D}^\sharp} \left( \Phi(c, c^\sharp) \dot{+} \left( - \left( \mathcal{K}(c, d) \dot{+} \left( (-\Psi(d, d^\sharp)) \dot{+} (-g^{-\Psi}(d^\sharp)) \right) \right) \right) \right)$$

by the property (50f) that the operator  $\inf$  is sublinear in  $\dot{+}$  [this inequality is an equality when  $\Phi(c, c^\sharp) < +\infty$  by (50g)]

$$= \sup_{c \in \mathbb{C}, d \in \mathbb{D}} \inf_{d^\sharp \in \mathbb{D}^\sharp} \left( \Phi(c, c^\sharp) \dot{+} (-\mathcal{K}(c, d)) \dot{+} \left( - \left( (-\Psi(d, d^\sharp)) \dot{+} (-g^{-\Psi}(d^\sharp)) \right) \right) \right)$$

by the correspondence (52b) between  $\dagger$  and  $\ddagger$  by means of  $a \mapsto -a$

$$= \sup_{c \in \mathbb{C}, d \in \mathbb{D}} \inf_{d^\# \in \mathbb{D}^\#} \left( \Phi(c, c^\#) \dagger (-\mathcal{K}(c, d)) \dagger \left( \Psi(d, d^\#) \ddagger g^{-\Psi}(d^\#) \right) \right)$$

by the correspondence (52b) between  $\dagger$  and  $\ddagger$  by means of  $a \mapsto -a$

$$= \sup_{c \in \mathbb{C}, d \in \mathbb{D}} \inf_{d^\# \in \mathbb{D}^\#} \left( \left( \Phi(c, c^\#) \dagger (-\mathcal{K}(c, d)) \right) \ddagger \left( \Psi(d, d^\#) \ddagger g^{-\Psi}(d^\#) \right) \right)$$

by associativity of  $\ddagger$

$$\leq \sup_{c \in \mathbb{C}, d \in \mathbb{D}} \inf_{d^\# \in \mathbb{D}^\#} \left( \left( \left( \Phi(c, c^\#) \dagger (-\mathcal{K}(c, d)) \right) \ddagger \Psi(d, d^\#) \right) \ddagger g^{-\Psi}(d^\#) \right)$$

by the inequality (52c) [this inequality is an equality when  $-\infty < \Phi(c, c^\#) \dagger (-\mathcal{K}(c, d)) < +\infty$  and  $-\infty < \Psi(d, d^\#) < +\infty$ ]

$$= \sup_{c \in \mathbb{C}, d \in \mathbb{D}} \inf_{d^\# \in \mathbb{D}^\#} \left( \left( \Phi(c, c^\#) \dagger (-\mathcal{K}(c, d)) \right) \ddagger \Psi(d, d^\#) \right) \ddagger g^{-\Psi}(d^\#) \quad (5a)$$

by associativity of  $\ddagger$

$$\leq \inf_{d^\# \in \mathbb{D}^\#} \sup_{c \in \mathbb{C}, d \in \mathbb{D}} \left( \left( \Phi(c, c^\#) \dagger (-\mathcal{K}(c, d)) \right) \ddagger \Psi(d, d^\#) \right) \ddagger g^{-\Psi}(d^\#) \quad (5b)$$

by  $\sup \inf \leq \inf \sup$

$$= \inf_{d^\# \in \mathbb{D}^\#} \left( \sup_{c \in \mathbb{C}, d \in \mathbb{D}} \left( \Phi(c, c^\#) \dagger (-\mathcal{K}(c, d)) \right) \ddagger \Psi(d, d^\#) \right) \ddagger g^{-\Psi}(d^\#)$$

by the property (50e) that the operator  $\sup$  is linear in  $\ddagger$

$$= \inf_{d^\# \in \mathbb{D}^\#} \left( \mathcal{K}^{\Phi \ddagger \Psi}(c^\#, d^\#) \ddagger g^{-\Psi}(d^\#) \right).$$

by the definition (3) of  $\mathcal{K}^{\Phi \ddagger \Psi}$ . This ends the proof.  $\square$

## 2.2 Fenchel inequality with a general coupling

In the next §2.3, we will provide sufficient conditions for the equality case in (4). However, we first prove two useful results: we establish the classical Fenchel inequality, but with a general coupling; we provide stronger estimates of the terms in the Fenchel inequality.

### 2.2.1 Fenchel inequality with a general coupling

Let be given two sets  $\mathbb{D}$ ,  $\mathbb{D}^\sharp$ , together with one coupling function

$$\Psi : \mathbb{D} \times \mathbb{D}^\sharp \rightarrow [-\infty, +\infty] . \quad (6)$$

**Proposition 2** *For any two functions  $h : \mathbb{D} \rightarrow [-\infty, +\infty]$  and  $g : \mathbb{D} \rightarrow [-\infty, +\infty]$ , we have that*

$$\sup_{d \in \mathbb{D}} \left( (-h(d)) \dot{+} (-g(d)) \right) \leq \inf_{d^\sharp \in \mathbb{D}^\sharp} \left( h^\Psi(d^\sharp) \dot{+} g^{-\Psi}(d^\sharp) \right) , \quad (7a)$$

or, equivalently, that

$$-\inf_{d \in \mathbb{D}} \left( h(d) \dot{+} g(d) \right) \leq \inf_{d^\sharp \in \mathbb{D}^\sharp} \left( h^\Psi(d^\sharp) \dot{+} g^{-\Psi}(d^\sharp) \right) , \quad (7b)$$

or, equivalently, that

$$0 \leq \inf_{d^\sharp \in \mathbb{D}^\sharp} \left( h^\Psi(d^\sharp) \dot{+} g^{-\Psi}(d^\sharp) \right) \dot{+} \inf_{d \in \mathbb{D}} \left( h(d) \dot{+} g(d) \right) . \quad (7c)$$

**Proof.** We take singleton sets  $\mathbb{C} = \{c\}$  and  $\mathbb{C}^\sharp = \{c^\sharp\}$ , with the null coupling  $\Phi(c, c^\sharp) = 0$ . Then, we use Theorem 1 with  $\mathcal{K}(c, d) = h(d)$  and  $f$  the constant function with value  $\inf_{d \in \mathbb{D}} \left( \mathcal{K}(c, d) \dot{+} g(d) \right) = \inf_{d \in \mathbb{D}} \left( h(d) \dot{+} g(d) \right)$ . We easily see that  $f^\Phi(c^\sharp) = -\inf_{d \in \mathbb{D}} \left( h(d) \dot{+} g(d) \right)$  and that  $\mathcal{K}^{\Phi \dot{+} \Psi}(c^\sharp, d^\sharp) = h^\Psi(d^\sharp)$ . We deduce Inequality (7b) from (4).

To end the proof, equivalence between the three inequalities follows from Equations (52b) and (52e).  $\square$

### 2.2.2 Stronger estimates in the Fenchel inequality

Now, we provide stronger estimates of the terms in the Fenchel inequality (2).

**Proposition 3** *Consider any two functions  $h : \mathbb{D} \rightarrow [-\infty, +\infty]$  and  $g : \mathbb{D} \rightarrow [-\infty, +\infty]$ . When any of the equivalent inequalities (7a), (7b) or (7c) in Proposition 2 is an equality, we have the following equality*

$$\sup_{d \in \mathbb{D}} \inf_{d^\sharp \in \mathbb{D}^\sharp} \left( \left( (-h(d)) \dot{+} \Psi(d, d^\sharp) \right) \dot{+} g^{-\Psi}(d^\sharp) \right) = \inf_{d^\sharp \in \mathbb{D}^\sharp} \sup_{d \in \mathbb{D}} \left( \left( (-h(d)) \dot{+} \Psi(d, d^\sharp) \right) \dot{+} g^{-\Psi}(d^\sharp) \right)$$

The proof is a straightforward consequence of the following Proposition.

**Proposition 4** For any two functions  $h : \mathbb{D} \rightarrow [-\infty, +\infty]$  and  $g : \mathbb{D} \rightarrow [-\infty, +\infty]$ , we have that

$$\sup_{d \in \mathbb{D}} \left( (-h(d)) \dot{+} (-g(d)) \right) \leq \sup_{d \in \mathbb{D}} \inf_{d^\# \in \mathbb{D}^\#} \left( \left( (-h(d)) \dot{+} \Psi(d, d^\#) \right) \dot{+} g^{-\Psi}(d^\#) \right), \quad (9a)$$

[with equality when  $-\infty < h$  and  $g = g^{(-\Psi)(-\Psi)}$ ], and

$$\inf_{d^\# \in \mathbb{D}^\#} \sup_{d \in \mathbb{D}} \left( \left( (-h(d)) \dot{+} \Psi(d, d^\#) \right) \dot{+} g^{-\Psi}(d^\#) \right) \leq \inf_{d^\# \in \mathbb{D}^\#} \left( h^\Psi(d^\#) \dot{+} g^{-\Psi}(d^\#) \right), \quad (9b)$$

[with equality when  $-\infty < g^{-\Psi}$ ].

If, in addition, we suppose that the functions  $h : \mathbb{D} \rightarrow [-\infty, +\infty]$  and  $g : \mathbb{D} \rightarrow [-\infty, +\infty]$  are such that<sup>1</sup>

$$-\infty < h, \quad g = g^{(-\Psi)(-\Psi)}, \quad -\infty < g^{(-\Psi)}, \quad (10)$$

then the two assumptions

$$\sup_{d \in \mathbb{D}} \left( (-h(d)) \dot{+} (-g(d)) \right) = \inf_{d^\# \in \mathbb{D}^\#} \left( h^\Psi(d^\#) \dot{+} g^{-\Psi}(d^\#) \right) \quad (11a)$$

and

$$\sup_{d \in \mathbb{D}} \inf_{d^\# \in \mathbb{D}^\#} \left( \left( (-h(d)) \dot{+} \Psi(d, d^\#) \right) \dot{+} g^{-\Psi}(d^\#) \right) = \inf_{d^\# \in \mathbb{D}^\#} \sup_{d \in \mathbb{D}} \left( \left( (-h(d)) \dot{+} \Psi(d, d^\#) \right) \dot{+} g^{-\Psi}(d^\#) \right) \quad (11b)$$

are equivalent.

**Proof.** First, we prove Inequality (9a). We have that

$$(-h(d)) \dot{+} (-g(d)) \leq (-h(d)) \dot{+} (-g^{(-\Psi)(-\Psi)}(d))$$

by the inequality (56) between a function and its biconjugate, and by the property (50b) that the operator  $\dot{+}$  is monotone

$$= (-h(d)) \dot{+} \left( - \sup_{d^\# \in \mathbb{D}^\#} \left( (-\Psi(d, d^\#)) \dot{+} \left( (-g^{(-\Psi)}(d^\#)) \right) \right) \right)$$

by definition (55) of the biconjugate  $g^{(-\Psi)(-\Psi)}$

$$= (-h(d)) \dot{+} \inf_{d^\# \in \mathbb{D}^\#} \left( \Psi(d, d^\#) \dot{+} g^{(-\Psi)}(d^\#) \right)$$

---

<sup>1</sup>Note that when  $g < +\infty$  and  $\Psi < +\infty$ , we have that  $-\infty < g^{(-\Psi)}$ .



by  $-\sup = \inf -$  and by the correspondence (52b) between  $+$  and  $\dot{+}$  by means of  $a \mapsto -a$

$$\leq \inf_{d^\# \in \mathbb{D}^\#} \left( (-h(d)) \dot{+} \left( \Psi(d, d^\#) \dot{+} g^{(-\Psi)}(d^\#) \right) \right)$$

by the property (50f) that the operator  $\inf$  is sublinear in  $\dot{+}$  [this inequality is an equality when  $(-h(d)) < +\infty$  by (50g)],

$$\leq \inf_{d^\# \in \mathbb{D}^\#} \left( \left( (-h(d)) \dot{+} \Psi(d, d^\#) \right) \dot{+} g^{-\Psi}(d^\#) \right)$$

by (52c) [this inequality is an equality when  $(-h(d)) < +\infty$ ].

Then, we take  $\sup_{d \in \mathbb{D}}$  on both sides of the obtained inequality to obtain Inequality (9a) [this inequality is an equality when  $g = g^{(-\Psi)(-\Psi)}$  and  $-\infty < h$ ].

Second, we prove Inequality (9b). We have that

$$h^\Psi(d^\#) \dot{+} g^{-\Psi}(d^\#) = \sup_{d \in \mathbb{D}} \left( (-h(d)) \dot{+} \Psi(d, d^\#) \right) \dot{+} g^{-\Psi}(d^\#)$$

by the definition (54) of the conjugate

$$\geq \sup_{d \in \mathbb{D}} \left( \left( (-h(d)) \dot{+} \Psi(d, d^\#) \right) \dot{+} g^{-\Psi}(d^\#) \right)$$

by the property (51f) that the operator  $\sup$  is sublinear in  $\dot{+}$ , and by the property (51b) that the operator  $\dot{+}$  is monotone [this inequality is an equality when  $-\infty < g^{-\Psi}(d^\#)$  by (51g)].

Then, we take  $\inf_{d^\# \in \mathbb{D}^\#}$  on both sides of the obtained inequality to obtain Inequality (9b) [this inequality is an equality when  $-\infty < g^{-\Psi}$ ].

The rest of the proof follows easily. □

## 2.3 The duality equality case

Let be given two coupling functions  $\Phi : \mathbb{C} \times \mathbb{C}^\# \rightarrow [-\infty, +\infty]$  and  $\Psi : \mathbb{D} \times \mathbb{D}^\# \rightarrow [-\infty, +\infty]$ .

The equality case in (4) is the property that

$$f(c) = \inf_{d \in \mathbb{D}} \left( \mathcal{K}(c, d) \dot{+} g(d) \right), \quad \forall c \in \mathbb{C} \Rightarrow$$

$$f^\Phi(c^\#) = \inf_{d^\# \in \mathbb{D}^\#} \left( \mathcal{K}^{\Phi \dot{+} \Psi}(c^\#, d^\#) \dot{+} g^{-\Psi}(d^\#) \right), \quad \forall c^\# \in \mathbb{C}^\#. \quad (12)$$

We will now provide sufficient conditions under which the equality case (12) holds true in different cases: with real-valued couplings and real-valued kernel; with extended couplings and extended kernel; with one bilinear coupling and extended kernel.

In [9], the equation  $f(c) = \inf_{d \in \mathbb{D}} \left( \mathcal{K}(c, d) \dot{+} g(d) \right)$  with unknown function  $g$  is studied.

### 2.3.1 With real-valued couplings and real-valued kernel

In that case, both the couplings and the kernel take real values, whereas all the other functions can take extended values.

**Corollary 5** *Consider any bivariate function  $\mathcal{K} : \mathbb{C} \times \mathbb{D} \rightarrow [-\infty, +\infty]$  and univariate functions  $f : \mathbb{C} \rightarrow [-\infty, +\infty]$  and  $g : \mathbb{D} \rightarrow [-\infty, +\infty]$ , all defined on the “primal” sets. Suppose that*

1.  $g^{(-\Psi)(-\Psi)} = g$ ;
2. *the equality between (5a) and (5b) holds true as, for instance, when the following function has a saddle point (or has no duality gap)*

$$((c, d), d^\sharp) \in (\mathbb{C} \times \mathbb{D}) \times \mathbb{D}^\sharp \mapsto \left( \Phi(c, c^\sharp) \dot{+} (-\mathcal{K}(c, d)) \dot{+} \Psi(d, d^\sharp) \right) \dot{+} g^{-\Psi}(d^\sharp), \quad (13)$$

3. *the two coupling functions  $\Phi : \mathbb{C} \times \mathbb{C}^\sharp \rightarrow \mathbb{R}$  and  $\Psi : \mathbb{D} \times \mathbb{D}^\sharp \rightarrow \mathbb{R}$ , and the kernel  $\mathcal{K} : \mathbb{C} \times \mathbb{D} \rightarrow \mathbb{R}$  all take finite values.*

Then, we have that the equality case (12) holds true.

**Proof.** Following the proof of Theorem 1, all but one inequality — namely  $\sup \inf \leq \inf \sup$  between (5a) and (5b) — become equalities when the functions  $\Phi : \mathbb{C} \times \mathbb{C}^\sharp \rightarrow \mathbb{R}$ ,  $\Psi : \mathbb{D} \times \mathbb{D}^\sharp \rightarrow \mathbb{R}$  and  $\mathcal{K} : \mathbb{C} \times \mathbb{D} \rightarrow \mathbb{R}$  take real values and when  $g^{(-\Psi)(-\Psi)} = g$ . Once we have the equality between (5a) and (5b), we obtain that the equality case (12) holds true.  $\square$

### 2.3.2 With extended couplings and extended kernel

In that case, the couplings, the kernel and all the other functions can take extended values.

**Corollary 6** *Consider any bivariate function  $\mathcal{K} : \mathbb{C} \times \mathbb{D} \rightarrow [-\infty, +\infty]$  and univariate functions  $f : \mathbb{C} \rightarrow [-\infty, +\infty]$  and  $g : \mathbb{D} \rightarrow [-\infty, +\infty]$ , all defined on the “primal” sets. We define*

$$\mathcal{K}_{c^\sharp}(d) = -\left( \mathcal{K}(\cdot, d)^\Psi(c^\sharp) \right) = \inf_{c \in \mathbb{C}} \left( (-\Phi(c, c^\sharp)) \dot{+} \mathcal{K}(c, d) \right), \quad \forall (c^\sharp, d) \in \mathbb{C}^\sharp \times \mathbb{D}. \quad (14)$$

Suppose that

1.  $g^{(-\Psi)(-\Psi)} = g$ ;
2.  $-\infty < g^{-\Psi}$ ;
3.  $-\infty < \mathcal{K}_{c^\sharp}(d)$ , for all  $(c^\sharp, d) \in \mathbb{C}^\sharp \times \mathbb{D}$ .

Then, we have that the equality case (12) holds true.

**Proof.** First, to prove the equality result (12), we start by giving a new proof of (4) in Theorem 1.

$$f^\Phi(c^\sharp) = \sup_{c \in \mathbb{C}} \left( \Phi(c, c^\sharp) \dagger (-f(c)) \right)$$

by definition (54) of the conjugate  $f^\Phi(c^\sharp)$

$$\leq \sup_{c \in \mathbb{C}} \left( \Phi(c, c^\sharp) \dagger \left( - \inf_{d \in \mathbb{D}} (\mathcal{K}(c, d) \dagger g(d)) \right) \right) \quad (15a)$$

by the left hand side assumption in (4) and by the property (50b) that the operator  $\dagger$  is monotone

$$= \sup_{c \in \mathbb{C}} \left( \Phi(c, c^\sharp) \dagger \sup_{d \in \mathbb{D}} \left( - (\mathcal{K}(c, d) \dagger g(d)) \right) \right)$$

by  $-\inf = \sup -$

$$= \sup_{c \in \mathbb{C}, d \in \mathbb{D}} \left( \Phi(c, c^\sharp) \dagger \left( - (\mathcal{K}(c, d) \dagger g(d)) \right) \right)$$

by the property (50e) that the operator  $\sup$  is linear in  $\dagger$

$$= \sup_{c \in \mathbb{C}, d \in \mathbb{D}} \left( \left( \Phi(c, c^\sharp) \dagger (-\mathcal{K}(c, d)) \right) \dagger (-g(d)) \right)$$

by (52b) and by associativity of  $\dagger$

$$= \sup_{d \in \mathbb{D}} \left( \sup_{c \in \mathbb{C}} \left( \Phi(c, c^\sharp) \dagger (-\mathcal{K}(c, d)) \right) \dagger (-g(d)) \right)$$

by the property (50e) that the operator  $\sup$  is linear in  $\dagger$

$$\leq \inf_{d^\sharp \in \mathbb{D}^\sharp} \left( \left( - \sup_{c \in \mathbb{C}} \left( \Phi(c, c^\sharp) \dagger (-\mathcal{K}(c, \cdot)) \right) \right)^\Psi(d^\sharp) \dagger g^{-\Psi}(d^\sharp) \right) \quad (15b)$$

by Fenchel inequality (7a) where  $h(d) = \mathcal{K}_{c^\sharp}(d) = - \sup_{c \in \mathbb{C}} \left( \Phi(c, c^\sharp) \dagger (-\mathcal{K}(c, d)) \right)$  by (14)

$$= \inf_{d^\sharp \in \mathbb{D}^\sharp} \left( \sup_{d \in \mathbb{D}} \left( \Psi(d, d^\sharp) \dagger \sup_{c \in \mathbb{C}} \left( \Phi(c, c^\sharp) \dagger (-\mathcal{K}(c, d)) \right) \right) \dagger g^{-\Psi}(d^\sharp) \right)$$

by definition of the Fenchel-Moreau  $\Psi$ -conjugate

$$= \inf_{d^\sharp \in \mathbb{D}^\sharp} \left( \sup_{c \in \mathbb{C}, d \in \mathbb{D}} \left( \Psi(d, d^\sharp) \dagger \Phi(c, c^\sharp) \dagger (-\mathcal{K}(c, d)) \right) \dagger g^{-\Psi}(d^\sharp) \right)$$

by the property (50e) that the operator  $\sup$  is linear in  $+$ , and by associativity of  $+$

$$= \inf_{d^\sharp \in \mathbb{D}^\sharp} \left( \mathcal{K}^{\Phi \dagger \Psi}(c^\sharp, d^\sharp) \dagger g^{-\Psi}(d^\sharp) \right)$$

by the definition (3) of  $\mathcal{K}^{\Phi \dagger \Psi}$ . This ends the new proof of Theorem 1.

Second, to end the proof of Corollary 6, we just check two points. That inequality (15a) is an equality, by the left hand side assumption in (12). That inequality (15b) is also an equality, as Proposition 4 applies under the assumptions of Corollary 6.  $\square$

### 2.3.3 With one bilinear coupling and extended kernel

In that case, one of the two couplings is bilinear, whereas the other coupling, the kernel and all the other functions can take extended values.

Let  $\mathbb{D}$  be a locally convex Hausdorff topological vector space over the real numbers  $\mathbb{R}$ , with its algebraic dual  $\mathbb{D}^\sharp$  made of linear forms on  $\mathbb{D}$ . The coupling is the duality bilinear form  $\langle \cdot, \cdot \rangle$ , and the conjugacy operator on functions is denoted by  $*$ . Let be given  $\mathbb{C}$  and  $\mathbb{C}^\sharp$  two sets and a coupling function  $\Phi : \mathbb{C} \times \mathbb{C}^\sharp \rightarrow [-\infty, +\infty]$ .

**Corollary 7** *Consider any bivariate function  $\mathcal{K} : \mathbb{C} \times \mathbb{D} \rightarrow [-\infty, +\infty]$  and univariate functions  $f : \mathbb{C} \rightarrow [-\infty, +\infty]$  and  $g : \mathbb{D} \rightarrow [-\infty, +\infty]$ , all defined on the “primal” sets. Suppose that*

1. *the coupling  $\Psi : \mathbb{D} \times \mathbb{D}^\sharp \rightarrow \mathbb{R}$  is the duality bilinear form  $\langle \cdot, \cdot \rangle$  between  $\mathbb{D}$  and its algebraic dual  $\mathbb{D}^\sharp$ ,*
2. *the function  $g$  is a proper<sup>2</sup> convex function,*
3. *for any  $c^\sharp \in \mathbb{C}^\sharp$ , the function  $\mathcal{K}_{c^\sharp}$  in (14) is a proper convex function,*
4. *for any  $c^\sharp \in \mathbb{C}^\sharp$ , the function  $g$  is continuous at some point where  $\mathcal{K}_{c^\sharp}$  is finite.*

*Then, we have that the equality case (12) holds true.*

**Proof.** The equality case (12) follows by checking that the inequalities (15a) and (15b) turn out to be equalities, under the assumptions of Corollary 7. Indeed, the left hand equality in (12) gives an equality in (15a). The equality in (15b) is a consequence of the equality

$$\inf_{d \in \mathbb{D}} \left( \mathcal{K}_{c^\sharp}(d) \dagger g(d) \right) = - \inf_{d^\sharp \in \mathbb{D}^\sharp} \left( \mathcal{K}_{c^\sharp}^\Psi(d^\sharp) \dagger g^{-\Psi}(d^\sharp) \right),$$

which holds true by [14, Theorem 1], under the assumptions made on the functions  $g$  and  $\mathcal{K}_{c^\sharp}$ , where the coupling  $\Psi : \mathbb{D} \times \mathbb{D}^\sharp \rightarrow \mathbb{R}$  is the duality bilinear form  $\langle \cdot, \cdot \rangle$  between  $\mathbb{D}$  and its algebraic dual  $\mathbb{D}^\sharp$ .  $\square$

---

<sup>2</sup>The function  $g$  never takes the value  $-\infty$  and is not identically equal to  $+\infty$ .

**Remark 8** *We can weaken the assumptions in Corollary 7 in two ways.*

- *Some of the assumptions in item 3 in Corollary 7 — that bear on the marginal function  $\mathcal{K}_{c^\sharp}$  in (14) — can be obtained from assumptions on the basic elements  $\mathcal{K}$  and  $\Phi$ . Indeed, if both the functions  $\mathcal{K} : \mathbb{C} \times \mathbb{D} \rightarrow [-\infty, +\infty]$  and  $\Phi(\cdot, c^\sharp) : \mathbb{C} \rightarrow [-\infty, +\infty]$  are convex, for any  $c^\sharp \in \mathbb{C}^\sharp$ , then the function  $\mathcal{K}_{c^\sharp}$  in (14) is convex.*
- *The assumptions in item 4 in Corollary 7 can be replaced by the following assumptions (using [2, Proposition 15.13]):  $g$  is a proper l.s.c. (lower semi continuous) convex function, the marginal function  $\mathcal{K}_{c^\sharp}$  in (14) is a proper l.s.c. convex function, and  $0 \in \text{sri}(\text{dom}(\overline{\mathcal{K}}_{c^\sharp}) - \text{dom}(g))$ , for any  $c^\sharp \in \mathbb{C}^\sharp$ . Moreover, for any given  $c^\sharp \in \mathbb{C}^\sharp$ , the marginal function  $\mathcal{K}_{c^\sharp}$  in (14) is a l.s.c. convex function under the following assumptions (see [1]): the function  $\mathcal{K}$  is l.s.c. convex: the function  $\Phi(\cdot, c^\sharp)$  is proper u.s.c. (upper semi continuous), and the minimization in the definition (14) of  $\mathcal{K}_{c^\sharp}$  is performed on a compact set.*

### 3 Applications

We now display three applications of our main result in Theorem 1. We provide a new formula for the Fenchel-Moreau conjugate of a generalized inf-convolution. We obtain formulas with partial Fenchel-Moreau conjugates. Finally, we consider the Bellman equation in stochastic dynamic programming and we provide a “Bellman-like” equation for the Fenchel conjugates of the value functions.

#### 3.1 Fenchel-Moreau conjugate of generalized inf-convolution

We generalize the inf-convolution, and provide an inequality and an equality with Fenchel-Moreau conjugates involving three coupling functions.

**Definition 9** *Let be given three sets  $\mathbb{C}$ ,  $\mathbb{D}_1$  and  $\mathbb{D}_2$ . For any trivariate convoluting function*

$$\mathcal{I} : \mathbb{D}_1 \times \mathbb{C} \times \mathbb{D}_2 \rightarrow [-\infty, +\infty] , \quad (16)$$

*we define the  $\mathcal{I}$ -inf-convolution of two functions  $g_1 : \mathbb{D}_1 \rightarrow [-\infty, +\infty]$  and  $g_2 : \mathbb{D}_2 \rightarrow [-\infty, +\infty]$  by*

$$(g_1 \overset{\mathcal{I}}{\square} g_2)(c) = \inf_{d_1 \in \mathbb{D}_1, d_2 \in \mathbb{D}_2} \left( g_1(d_1) \dot{+} \mathcal{I}(d_1, c, d_2) \dot{+} g_2(d_2) \right) , \quad \forall c \in \mathbb{C} . \quad (17)$$

To any convoluting function  $\mathcal{I}$  in (16), we can easily attach

1. a coupling function  $\overline{\mathcal{I}} : \mathbb{C} \times (\mathbb{D}_1 \times \mathbb{D}_2) \rightarrow [-\infty, +\infty]$  between  $\mathbb{C}$  and  $\mathbb{D}_1 \times \mathbb{D}_2$  defined by

$$\overline{\mathcal{I}}(c, (d_1, d_2)) = \mathcal{I}(d_1, c, d_2) , \quad \forall (c, d_1, d_2) \in \mathbb{C} \times \mathbb{D}_1 \times \mathbb{D}_2 , \quad (18)$$

2. a kernel function  $\underline{\mathcal{I}} : \mathbb{C} \times \mathbb{D}_1 \times \mathbb{D}_2 \rightarrow [-\infty, +\infty]$  defined by

$$\underline{\mathcal{I}}(c, d_1, d_2) = \mathcal{I}(d_1, c, d_2), \quad \forall (c, d_1, d_2) \in \mathbb{C} \times \mathbb{D}_1 \times \mathbb{D}_2. \quad (19)$$

We provide an inequality with Fenchel-Moreau conjugates involving three coupling functions.

**Proposition 10** *Let be given three “primal” sets  $\mathbb{C}$ ,  $\mathbb{D}_1$ ,  $\mathbb{D}_2$  and three “dual” sets  $\mathbb{C}^\sharp$ ,  $\mathbb{D}_1^\sharp$ ,  $\mathbb{D}_2^\sharp$ , together with three coupling functions*

$$\Phi : \mathbb{C} \times \mathbb{C}^\sharp \rightarrow [-\infty, +\infty], \quad \Psi_1 : \mathbb{D}_1 \times \mathbb{D}_1^\sharp \rightarrow [-\infty, +\infty], \quad \Psi_2 : \mathbb{D}_2 \times \mathbb{D}_2^\sharp \rightarrow [-\infty, +\infty]. \quad (20)$$

*For any univariate functions  $f : \mathbb{C} \rightarrow [-\infty, +\infty]$ ,  $g_1 : \mathbb{D}_1 \rightarrow [-\infty, +\infty]$  and  $g_2 : \mathbb{D}_2 \rightarrow [-\infty, +\infty]$ , all defined on the “primal” sets, we have that*

$$f(c) \geq (g_1 \square^{\underline{\mathcal{I}}} g_2)(c), \quad \forall c \in \mathbb{C} \Rightarrow f^\Phi(c^\sharp) \leq (g_1^{(-\Psi_1)} \square^{\mathcal{I}^\sharp} g_2^{(-\Psi_2)})(c^\sharp), \quad \forall c^\sharp \in \mathbb{C}^\sharp, \quad (21)$$

where the convoluting function  $\mathcal{I}^\sharp$  on the “dual” sets is given by

$$\mathcal{I}^\sharp = \underline{\mathcal{I}}^{\Phi \dagger \Psi_1 \dagger \Psi_2}, \quad (22a)$$

that is, by

$$\mathcal{I}^\sharp(d_1^\sharp, c^\sharp, d_2^\sharp) = \sup_{(d_1, c, d_2) \in \mathbb{D}_1 \times \mathbb{C} \times \mathbb{D}_2} \left( \Phi(c, c^\sharp) \dagger \Psi_1(d_1, d_1^\sharp) \dagger \Psi_2(d_2, d_2^\sharp) \dagger (-\mathcal{I}(d_1, c, d_2)) \right). \quad (22b)$$

**Proof.** The left hand side assumption in (21) can be rewritten as

$$f(c) \geq \inf_{d_1 \in \mathbb{D}_1, d_2 \in \mathbb{D}_2} \left( \underline{\mathcal{I}}(c, d_1, d_2) \dagger (g_1(d_1) \dagger g_2(d_2)) \right), \quad \forall c \in \mathbb{C}. \quad (23)$$

Now, we apply Theorem 1 with

$$\begin{aligned} \mathbb{D} &= \mathbb{D}_1 \times \mathbb{D}_2, \quad \mathbb{D}^\sharp = \mathbb{D}_1^\sharp \times \mathbb{D}_2^\sharp \\ \Psi((d_1, d_2), (d_1^\sharp, d_2^\sharp)) &= \Psi_1(d_1, d_1^\sharp) \dagger \Psi_2(d_2, d_2^\sharp) \\ g(d_1, d_2) &= g_1(d_1) \dagger g_2(d_2) \quad \text{and} \quad \mathcal{K} = \underline{\mathcal{I}} \quad \text{by (19)}. \end{aligned}$$

We first prove that

$$g^{-(\Psi_1 \dagger \Psi_2)}(d_1^\sharp, d_2^\sharp) \leq \left( g_1^{(-\Psi_1)}(d_1^\sharp) \dagger g_2^{(-\Psi_2)}(d_2^\sharp) \right). \quad (24)$$

For this, we let the reader check that the following preliminary inequality always holds true

$$(- (u_1 \dagger u_2)) \dagger (- (v_1 \dagger v_2)) \leq ((-u_1) \dagger (-v_1)) \dagger ((-u_2) \dagger (-v_2)). \quad (25)$$

Then, we have that

$$\begin{aligned} g^{-(\Psi_1 \dagger \Psi_2)}(d_1^\#, d_2^\#) &= \sup_{d_1 \in \mathbb{D}_1, d_2 \in \mathbb{D}_2} \left( \left( -(\Psi_1(d_1, d_1^\#) \dagger \Psi_2(d_2, d_2^\#)) \right) \dagger \left( -(g_1(d_1) \dagger g_2(d_2)) \right) \right) \\ &\leq \sup_{d_1 \in \mathbb{D}_1} \left( \left( -\Psi_1(d_1, d_1^\#) \right) \dagger \left( -g_1(d_1) \right) \right) \dagger \sup_{d_2 \in \mathbb{D}_2} \left( \left( -\Psi_2(d_2, d_2^\#) \right) \dagger \left( -g_2(d_2) \right) \right) \end{aligned}$$

by the preliminary inequality (25)

$$= g_1^{(-\Psi_1)}(d_1^\#) \dagger g_2^{(-\Psi_2)}(d_2^\#).$$

We now obtain, by (4),

$$\begin{aligned} f^\Phi(c^\#) &\leq \inf_{d_1^\# \in \mathbb{D}_1^\#, d_2^\# \in \mathbb{D}_2^\#} \left( \underline{\mathcal{I}}^{\Phi \dagger (\Psi_1 \dagger \Psi_2)}(c^\#, d_1^\#, d_2^\#) \dagger g^{-(\Psi_1 \dagger \Psi_2)}(d_1^\#, d_2^\#) \right) \\ &\leq \inf_{d_1^\# \in \mathbb{D}_1^\#, d_2^\# \in \mathbb{D}_2^\#} \left( \underline{\mathcal{I}}^{\Phi \dagger (\Psi_1 \dagger \Psi_2)}(c^\#, d_1^\#, d_2^\#) \dagger \left( g_1^{(-\Psi_1)}(d_1^\#) \dagger g_2^{(-\Psi_2)}(d_2^\#) \right) \right) \quad (\text{by (24)}) \\ &= (g_1^{(-\Psi_1)} \square_{\mathcal{I}^\#} g_2^{(-\Psi_2)})(c^\#) \end{aligned}$$

by definition (17) of the generalized inf-convolution. This ends the proof.  $\square$

We check our result in Theorem 1 on the classical inf-convolution. Suppose that  $\mathbb{C} = \mathbb{D} = \mathbb{R}^n$  and  $\mathbb{C}^\# = \mathbb{D}^\# = \mathbb{R}^n$ , with coupling given by the scalar product  $\langle \cdot, \cdot \rangle$ . The conjugacy operator on functions is denoted by  $\star$ .

Now, when we take

$$\mathcal{I}(d_1, c, d_2) = \delta_{d_1 + d_2}(c), \quad \Phi(c, c^\#) = \langle c, c^\# \rangle, \quad \Psi_i(d_i, d_i^\#) = -\langle d_i, d_i^\# \rangle, \quad i = 1, 2, \quad (27a)$$

we find that, by (22b)

$$\mathcal{I}^\#(d_1^\#, c^\#, d_2^\#) = \delta_{d_1^\#}(c^\#) \dagger \delta_{d_2^\#}(c^\#), \quad (g_1^{(-\Psi_1)} \square_{\mathcal{I}^\#} g_2^{(-\Psi_2)})(c^\#) = g_1^\star(c^\#) \dagger g_2^\star(c^\#). \quad (27b)$$

We conclude with (21) that we indeed obtain the well known property of the inf-convolution:  $f \geq g_1 \square g_2 \Rightarrow f^\star \leq (g_1 \square g_2)^\star = g_1^\star + g_2^\star$ .

To end up, we provide an expression of the inf-convolution as a Fenchel-Moreau conjugate, and we obtain an equality with Fenchel-Moreau conjugates involving three coupling functions.

**Proposition 11** *The  $\mathcal{I}$ -inf-convolution in (17) is given by*

$$g_1 \square_{\mathcal{I}} g_2 = -(g_1 \dagger g_2)^{-\overline{\mathcal{I}}}. \quad (28)$$

**Proof.** For any  $c \in \mathbb{C}$ , we have that

$$(g_1 \overset{\mathcal{I}}{\square} g_2)(c) = \inf_{d_1 \in \mathbb{D}_1, d_2 \in \mathbb{D}_2} \left( g_1(d_1) \dot{+} \mathcal{I}(d_1, c, d_2) \dot{+} g_2(d_2) \right)$$

by definition (17) of the generalized inf-convolution

$$= \inf_{d_1 \in \mathbb{D}_1, d_2 \in \mathbb{D}_2} \left( \overline{\mathcal{I}}(c, (d_1, d_2)) \dot{+} (g_1(d_1) \dot{+} g_2(d_2)) \right)$$

by definition (18) of the coupling function  $\overline{\mathcal{I}}$

$$\begin{aligned} &= - \sup_{d_1 \in \mathbb{D}_1, d_2 \in \mathbb{D}_2} \left( \left( -\overline{\mathcal{I}}(c, (d_1, d_2)) \right) \dot{+} \left( - (g_1(d_1) \dot{+} g_2(d_2)) \right) \right) \\ &= -((g_1 \dot{+} g_2)^{-\overline{\mathcal{I}}})(c). \end{aligned}$$

□

**Proposition 12** *If there exist two coupling functions*

$$\Gamma_1 : \mathbb{C}^\# \times \mathbb{D}_1 \rightarrow [-\infty, +\infty], \quad \Gamma_2 : \mathbb{C}^\# \times \mathbb{D}_2 \rightarrow [-\infty, +\infty], \quad (30)$$

*such that the  $\Phi$ -Fenchel-Moreau conjugate of the convoluting function  $\mathcal{I}$  splits as*

$$\mathcal{I}(d_1, \cdot, d_2)^\Phi(c^\#) = \Gamma_1(c^\#, d_1) \dot{+} \Gamma_2(c^\#, d_2), \quad \forall (c^\#, d_1, d_2) \in \mathbb{C}^\# \times \mathbb{D}_1 \times \mathbb{D}_2, \quad (31)$$

*then the  $\Phi$ -Fenchel-Moreau conjugate of the inf-convolution  $g_1 \overset{\mathcal{I}}{\square} g_2$  is given by a sum as*

$$(g_1 \overset{\mathcal{I}}{\square} g_2)^\Phi = g_1^{\Gamma_1} \dot{+} g_2^{\Gamma_2}. \quad (32)$$

**Proof.**

$$\begin{aligned} (g_1 \overset{\mathcal{I}}{\square} g_2)^\Phi(c^\#) &= \left( - (g_1 \dot{+} g_2)^{-\overline{\mathcal{I}}} \right)^\Phi(c^\#) && \text{(by (28))} \\ &= (g_1 \dot{+} g_2)^{(-\overline{\mathcal{I}}) \dot{+} \Phi}(c^\#) && \text{(by (60))} \\ &= \sup_{(d_1, c, d_2) \in \mathbb{D}_1 \times \mathbb{C} \times \mathbb{D}_2} \left( \left( -\mathcal{I}(d_1, c, d_2) \right) \dot{+} \Phi(c, c^\#) \dot{+} \left( - (g_1(d_1) \dot{+} g_2(d_2)) \right) \right) \\ &= \sup_{(d_1, d_2) \in \mathbb{D}_1 \times \mathbb{D}_2} \left( \sup_{c \in \mathbb{C}} \left( \left( -\mathcal{I}(d_1, c, d_2) \right) \dot{+} \Phi(c, c^\#) \right) \dot{+} \left( - (g_1(d_1) \dot{+} g_2(d_2)) \right) \right) \\ &= \sup_{(d_1, d_2) \in \mathbb{D}_1 \times \mathbb{D}_2} \left( \mathcal{I}(d_1, \cdot, d_2)^\Phi(c^\#) \dot{+} \left( - (g_1(d_1) \dot{+} g_2(d_2)) \right) \right) \\ &= \sup_{(d_1, d_2) \in \mathbb{D}_1 \times \mathbb{D}_2} \left( \left( \Gamma_1(c^\#, d_1) \dot{+} \Gamma_2(c^\#, d_2) \right) \dot{+} \left( - (g_1(d_1) \dot{+} g_2(d_2)) \right) \right) && \text{(by assumption)} \\ &= \sup_{(d_1, d_2) \in \mathbb{D}_1 \times \mathbb{D}_2} \left( \Gamma_1(c^\#, d_1) \dot{+} \Gamma_2(c^\#, d_2) \dot{+} (-g_1(d_1)) \dot{+} (-g_2(d_2)) \right) \\ &= g_1^{\Gamma_1}(c^\#) \dot{+} g_2^{\Gamma_2}(c^\#). \end{aligned}$$



□

### 3.2 Exchanging partial Fenchel-Moreau conjugates

Let be given two “primal” sets  $\mathbb{C}, \mathbb{D}$  and two “dual” sets  $\mathbb{C}^\#, \mathbb{D}^\#$ , together with two coupling functions

$$\Phi : \mathbb{C} \times \mathbb{C}^\# \rightarrow [-\infty, +\infty], \quad \Psi : \mathbb{D} \times \mathbb{D}^\# \rightarrow [-\infty, +\infty]. \quad (34)$$

In (4), all Fenchel-Moreau conjugates stand on the right side of the implication. We show formulas where they appear on both sides. For this purpose, for any *exchange function*

$$\mathcal{E} : \mathbb{C} \times \mathbb{D}^\# \rightarrow [-\infty, +\infty], \quad (35a)$$

we introduce the *partial Fenchel-Moreau conjugates*

$$(-\mathcal{E}(c, \cdot))^\Psi(d) = \sup_{d^\# \in \mathbb{D}^\#} \left( \Psi(d, d^\#) \dagger \mathcal{E}(c, d^\#) \right), \quad \forall (c, d) \in \mathbb{C} \times \mathbb{D}, \quad (35b)$$

$$\mathcal{E}(\cdot, d^\#)^\Phi(c^\#) = \sup_{c \in \mathbb{C}} \left( \Phi(c, c^\#) \dagger (-\mathcal{E}(c, d^\#)) \right), \quad \forall (c^\#, d^\#) \in \mathbb{C}^\# \times \mathbb{D}^\#. \quad (35c)$$

We prove the following implication.

**Proposition 13** *For any function  $\mathcal{E} : \mathbb{C} \times \mathbb{D}^\# \rightarrow [-\infty, +\infty]$ , we have that*

$$f(c) \geq \inf_{d \in \mathbb{D}} \left( \mathcal{E}(c, \cdot)^\Psi(d) \dagger g(d) \right), \quad \forall c \in \mathbb{C} \Rightarrow \\ f^\Phi(c^\#) \leq \inf_{d^\# \in \mathbb{D}^\#} \left( (-\mathcal{E}(\cdot, d^\#))^\Phi(c^\#) \dagger g^{-\Psi}(d^\#) \right), \quad \forall c^\# \in \mathbb{C}^\#. \quad (36)$$

**Proof.** We use the following Lemma 14. We apply Theorem 1 with the function  $\mathcal{K}(c, d) = (-\mathcal{E}(c, \cdot))^\Psi(d)$  defined by equality in the left hand side inequality in (37). Then, we insert the right hand side inequality in (37) into implication (4). □

**Lemma 14** *For any function  $\mathcal{E} : \mathbb{C} \times \mathbb{D}^\# \rightarrow [-\infty, +\infty]$ , we have that*

$$\mathcal{K}(c, d) \geq (-\mathcal{E}(c, \cdot))^\Psi(d), \quad \forall (c, d) \in \mathbb{C} \times \mathbb{D} \Rightarrow \\ \mathcal{K}^{\Phi \dagger \Psi}(c^\#, d^\#) \leq \mathcal{E}(\cdot, d^\#)^\Phi(c^\#), \quad \forall (c^\#, d^\#) \in \mathbb{C}^\# \times \mathbb{D}^\#. \quad (37)$$

**Proof.** Supposing that

$$\mathcal{K}(c, d) \geq (-\mathcal{E}(c, \cdot))^\Psi(d), \quad \forall (c, d) \in \mathbb{C} \times \mathbb{D}, \quad (38)$$

we calculate, for all  $(c^\sharp, d^\sharp) \in \mathbb{C}^\sharp \times \mathbb{D}^\sharp$ ,

$$\mathcal{K}^{\Phi \dagger \Psi}(c^\sharp, d^\sharp) = \sup_{c \in \mathbb{C}, d \in \mathbb{D}} \left( \Phi(c, c^\sharp) \dagger \Psi(d, d^\sharp) \dagger (-\mathcal{K}(c, d)) \right)$$

by the definition (3) of  $\mathcal{K}^{\Phi \dagger \Psi}$

$$\leq \sup_{c \in \mathbb{C}, d \in \mathbb{D}} \left( \Phi(c, c^\sharp) \dagger \Psi(d, d^\sharp) \dagger \mathcal{E}(c, \cdot)^\Psi(d) \right)$$

by inequality (38) for  $\mathcal{K}$

$$= \sup_{c \in \mathbb{C}} \left( \Phi(c, c^\sharp) \dagger \sup_{d \in \mathbb{D}} \left( \Psi(d, d^\sharp) \dagger \left( -(-\mathcal{E}(c, \cdot)^\Psi(d)) \right) \right) \right)$$

by by the property (50e) that the operator sup is linear in  $\dagger$

$$= \sup_{c \in \mathbb{C}} \left( \Phi(c, c^\sharp) \dagger (-\mathcal{E}(c, \cdot))^\Psi(d^\sharp) \right)$$

by definition (55) of the biconjugate

$$\leq \sup_{c \in \mathbb{C}} \left( \Phi(c, c^\sharp) \dagger (-\mathcal{E}(c, d^\sharp)) \right)$$

by the inequality (56) between a function and its biconjugate, and by the property (50b) that the operator  $\dagger$  is monotone

$$= \mathcal{E}(\cdot, d^\sharp)^\Phi(c^\sharp)$$

by the definition (35c) of partial Fenchel-Moreau conjugate. □

### 3.3 Fenchel conjugates of Bellman functions

We consider the Bellman equation in stochastic dynamic programming and we provide a “Bellman-like” equation for the Fenchel conjugates of the value functions.

#### 3.3.1 Basic sets and couplings

Let  $(\Omega, \mathcal{F}, \mathbb{P})$  be a probability space. Let  $1 \leq p < +\infty$  and  $q$  be defined by  $1/p + 1/q = 1$ . Adopting the notation of Sect. 2, we put  $\mathbb{C} = \mathbb{X} = \mathbb{R}^{n_x}$  and  $\mathbb{D} = \mathbb{L}^p((\Omega, \mathcal{F}, \mathbb{P}), \mathbb{R}^{n_x})$  the space of  $p$ -integrable random variables with values in  $\mathbb{R}^{n_x}$ . Elements of  $\mathbb{D}$ , that is,  $p$ -integrable random variables with values in  $\mathbb{X}$ , will be denoted by bold letters like  $\mathbf{X}$  and elements of  $\mathbb{D}^\sharp = \mathbb{L}^q((\Omega, \mathcal{F}, \mathbb{P}), \mathbb{R}^{n_x})$  by  $\mathbf{X}^\sharp$ .

The coupling  $\Phi$  between  $\mathbb{C} = \mathbb{X} = \mathbb{R}^{n_x}$  and  $\mathbb{C}^\sharp = \mathbb{X}^\sharp = \mathbb{R}^{n_x}$  is the usual scalar product  $\langle \cdot, \cdot \rangle$ . The coupling  $\Psi$  between  $\mathbb{D} = \mathbb{L}^p((\Omega, \mathcal{F}, \mathbb{P}), \mathbb{R}^{n_x})$  and  $\mathbb{D}^\sharp = \mathbb{L}^q((\Omega, \mathcal{F}, \mathbb{P}), \mathbb{R}^{n_x})$  is naturally derived in such a way that

$$\Phi(x, x^\sharp) = \langle x, x^\sharp \rangle, \quad \Psi(\mathbf{X}, \mathbf{X}^\sharp) = \mathbb{E}[\langle \mathbf{X}, \mathbf{X}^\sharp \rangle]. \quad (40)$$

In that case, the conjugates  $f^\Phi, g^\Psi, g^{-\Psi}$  and  $\mathcal{K}^{\Phi \dagger (-\Psi)}$  are denoted by  $f^*, g^*, g^{(-*)}$  and  $\mathcal{K}^{*\dagger (-*)}$ . One can find such a difference coupling in [21].

### 3.3.2 Bellman functions and Bellman equation

Let time  $t = 0, 1, \dots, T$  be discrete, with  $T \in \mathbb{N}^*$ . Consider a stochastic optimal control problem with state space  $\mathbb{X} = \mathbb{R}^{n_x}$ , control space  $\mathbb{U} = \mathbb{R}^{n_u}$ , noise process  $\{\mathbf{W}_t\}_{t=1, \dots, T}$  taking values in  $\mathbb{W} = \mathbb{R}^{n_w}$  and defined over the probability space  $(\Omega, \mathcal{F}, \mathbb{P})$ .

For each time  $t = 0, 1, \dots, T-1$ , we have a dynamics  $F_t : \mathbb{X} \times \mathbb{U} \times \mathbb{W} \rightarrow \mathbb{X}$  and an instantaneous cost  $L_t : \mathbb{X} \times \mathbb{U} \times \mathbb{W} \rightarrow ]-\infty, +\infty]$ ; we also have a final cost  $K : \mathbb{X} \rightarrow ]-\infty, +\infty]$ . These two costs can take the value  $+\infty$ , so that we can easily handle state and control constraints.

**Assumption 15** *We make the following assumptions:*

1. for any  $(x, u) \in \mathbb{X} \times \mathbb{U}$ , the  $\mathbb{R}^{n_x}$ -valued random variable  $F_t(x, u, \mathbf{W}_{t+1})$  belongs to  $\mathbb{L}^p((\Omega, \mathcal{F}, \mathbb{P}), \mathbb{R}^{n_x})$ ,
2. the instantaneous costs  $L_t : \mathbb{X} \times \mathbb{U} \times \mathbb{W} \rightarrow [0, +\infty]$ , for  $t = 0, \dots, T-1$ , and the final cost  $K : \mathbb{X} \rightarrow [0, +\infty]$  are nonnegative<sup>3</sup> measurable functions.

By item 2 in Assumption 15, we can define *Bellman functions* by, for all  $x \in \mathbb{X}$ ,

$$V_T(x) = K(x), \quad (41a)$$

$$V_t(x) = \inf_{\mathbf{X}, \mathbf{U}} \mathbb{E} \left[ \sum_{s=t}^{T-1} L_s(\mathbf{X}_s, \mathbf{U}_s, \mathbf{W}_{s+1}) \dagger K(\mathbf{X}_T) \right], \quad t = T-1, \dots, 0, \quad (41b)$$

where  $\mathbf{X}_t = x \in \mathbb{X}$ ,  $\mathbf{X}_{s+1} = F_s(\mathbf{X}_s, \mathbf{U}_s, \mathbf{W}_{s+1})$  and  $\sigma(\mathbf{U}_s) \subset \sigma(\mathbf{X}_s)$ , for  $s = t, \dots, T-1$ . In addition, the Bellman functions are nonnegative.

**Assumption 16** *We suppose that the Bellman functions in (41) are measurable and satisfy the backward Bellman equation*

$$V_t(x) = \inf_{u \in \mathbb{U}} \mathbb{E} \left[ L_t(x, u, \mathbf{W}_{t+1}) \dagger V_{t+1}(F_t(x, u, \mathbf{W}_{t+1})) \right], \quad t = T-1, \dots, 0. \quad (42)$$

*This is the case when the noise process  $\{\mathbf{W}_t\}_{t=1, \dots, T}$  is a white noise and under technical assumptions [3, 4].*

---

<sup>3</sup>We could also consider functions that are uniformly bounded below. However, for the sake of simplicity, and without loss of generality, we will deal with nonnegative functions.

### 3.3.3 Fenchel conjugates of the Bellman functions

Now, we provide a ‘‘Bellman-like’’ equation for the Fenchel conjugates of the value functions (see [12] for related considerations).

**Proposition 17** *The Bellman functions in (41) satisfy the backward equalities*

$$V_t(x) = \inf_{\mathbf{X} \in \mathbb{D}} \left( \inf_{u \in \mathbb{U}} \left( (-\mathcal{H}(x, u, \cdot))^{(-\star)}(\mathbf{X}) \right) \dot{+} \mathbb{E}[V_{t+1}(\mathbf{X})] \right), \quad \forall t = T - 1, \dots, 0, \quad (43)$$

where the Hamiltonian  $\mathcal{H}$  is defined by

$$\mathcal{H}(x, u, \mathbf{X}^\sharp) = \mathbb{E}[L_t(x, u, \mathbf{W}_{t+1}) \dot{+} \langle F_t(x, u, \mathbf{W}_{t+1}), \mathbf{X}^\sharp \rangle], \quad \forall (x, u, \mathbf{X}^\sharp) \in \mathbb{C} \times \mathbb{U} \times \mathbb{D}^\sharp. \quad (44)$$

Moreover, letting  $\{V_t^\star\}_{t=0,1,\dots,T}$  be the Fenchel conjugates of the Bellman functions, we have, for all  $x^\sharp \in \mathbb{X}^\sharp$ ,

$$V_t^\star(x^\sharp) \leq \inf_{\mathbf{X}^\sharp} \left( \sup_{u \in \mathbb{U}} \left( \mathcal{H}(\cdot, u, \mathbf{X}^\sharp)^\star(x^\sharp) \right) \dot{+} \mathbb{E}[V_{t+1}^\star(\mathbf{X}^\sharp)] \right), \quad \forall t = T - 1, \dots, 0. \quad (45)$$

**Proof.** In what follows, we will manipulate mathematical expectations of random variables that are either nonnegative (by item 2 in Assumption 15), or nonpositive (by taking the opposite), or integrable (by item 1 in Assumption 15, giving random variables resulting from a scalar product between an element of  $\mathbb{L}^p((\Omega, \mathcal{F}, \mathbb{P}), \mathbb{R}^{n_x})$  and one of  $\mathbb{L}^q((\Omega, \mathcal{F}, \mathbb{P}), \mathbb{R}^{n_x})$ ). We will be careful to remain in the conditions where the usual rules of algebra apply [7].

By the Bellman equation (42), we have that

$$\begin{aligned} V_t(x) &= \inf_{\mathbf{X} \in \mathbb{D}, u \in \mathbb{U}} \mathbb{E}[L_t(x, u, \mathbf{W}_{t+1}) \dot{+} V_{t+1}(\mathbf{X})] \\ &\quad \text{s.t. } \mathbf{X} = F_t(x, u, \mathbf{W}_{t+1}) \\ &= \inf_{\mathbf{X} \in \mathbb{D}, u \in \mathbb{U}} \sup_{\mathbf{X}^\sharp \in \mathbb{D}^\sharp} \left( \mathbb{E}[L_t(x, u, \mathbf{W}_{t+1}) \dot{+} V_{t+1}(\mathbf{X}) \dot{+} \langle F_t(x, u, \mathbf{W}_{t+1}) - \mathbf{X}, \mathbf{X}^\sharp \rangle] \right) \end{aligned}$$

by using item 1 in Assumption 15

$$= \inf_{\mathbf{X} \in \mathbb{D}, u \in \mathbb{U}} \left( \sup_{\mathbf{X}^\sharp \in \mathbb{D}^\sharp} \left( \mathbb{E}[L_t(x, u, \mathbf{W}_{t+1}) \dot{+} \langle F_t(x, u, \mathbf{W}_{t+1}) - \mathbf{X}, \mathbf{X}^\sharp \rangle] \right) \dot{+} \mathbb{E}[V_{t+1}(\mathbf{X})] \right)$$

by (51g) with  $0 \leq \mathbb{E}[V_{t+1}(\mathbf{X})]$  since the Bellman functions are nonnegative, and as  $\mathbb{E}$  and  $\dot{+}$  commute because, by Assumption 15, all terms inside the expectation  $\mathbb{E}$  are either nonnegative or integrable

$$= \inf_{\mathbf{X} \in \mathbb{D}, u \in \mathbb{U}} \left( \sup_{\mathbf{X}^\sharp \in \mathbb{D}^\sharp} \left( \mathbb{E}[\langle -\mathbf{X}, \mathbf{X}^\sharp \rangle] \dot{+} \mathcal{H}(x, u, \mathbf{X}^\sharp) \right) \dot{+} \mathbb{E}[V_{t+1}(\mathbf{X})] \right)$$

by definition (44) of the Hamiltonian

$$= \inf_{\mathbf{X} \in \mathbb{D}, u \in \mathbb{U}} \left( \sup_{\mathbf{X}^\sharp \in \mathbb{D}^\sharp} \left( \mathbb{E} \left[ \langle -\mathbf{X}, \mathbf{X}^\sharp \rangle \right] \dot{+} \mathcal{H}(x, u, \mathbf{X}^\sharp) \right) \dot{+} \mathbb{E}[V_{t+1}(\mathbf{X})] \right)$$

as Moreau upper and lower additions coincide above because  $-\infty < \mathbb{E} \left[ \langle -\mathbf{X}, \mathbf{X}^\sharp \rangle \right] < +\infty$  by definition of the spaces  $\mathbb{D}$  and  $\mathbb{D}^\sharp$  and of the coupling (40) between them

$$\begin{aligned} &= \inf_{\mathbf{X} \in \mathbb{D}, u \in \mathbb{U}} \left( \sup_{\mathbf{X}^\sharp \in \mathbb{D}^\sharp} \left( \mathbb{E} \left[ \langle -\mathbf{X}, \mathbf{X}^\sharp \rangle \right] \dot{+} \left( - \left( - \mathcal{H}(x, u, \mathbf{X}^\sharp) \right) \right) \right) \dot{+} \mathbb{E}[V_{t+1}(\mathbf{X})] \right) \\ &= \inf_{\mathbf{X} \in \mathbb{D}, u \in \mathbb{U}} \left( \left( - \mathcal{H}(x, u, \cdot) \right)^{(-\star)}(\mathbf{X}) \dot{+} \mathbb{E}[V_{t+1}(\mathbf{X})] \right) \end{aligned}$$

by definition of the Fenchel conjugate of  $\mathbf{X}^\sharp \mapsto -\mathcal{H}(x, u, \cdot)$  with respect to the opposite coupling  $(-\star)$  defined by  $(\mathbf{X}, \mathbf{X}^\sharp) \mapsto \mathbb{E} \left[ \langle -\mathbf{X}, \mathbf{X}^\sharp \rangle \right]$ , so that we have proven (43)

$$= \inf_{\mathbf{X} \in \mathbb{D}} \left( \inf_{u \in \mathbb{U}} \left( \left( - \mathcal{H}(x, u, \cdot) \right)^{(-\star)}(\mathbf{X}) \right) \dot{+} \mathbb{E}[V_{t+1}(\mathbf{X})] \right)$$

by the property (51e) that the operator  $\inf$  is linear in  $\dot{+}$

$$= \inf_{\mathbf{X} \in \mathbb{D}} \left( \inf_{u \in \mathbb{U}} \mathcal{K}_u(x, \mathbf{X}) \dot{+} \mathbb{E}[V_{t+1}(\mathbf{X})] \right),$$

where we have defined

$$\mathcal{K}_u(x, \mathbf{X}) = \left( - \mathcal{H}(x, u, \cdot) \right)^{(-\star)}(\mathbf{X}), \quad \forall u \in \mathbb{U}, \quad \forall (x, \mathbf{X}) \in \mathbb{C} \times \mathbb{D}. \quad (46)$$

By (37), we obtain that

$$\mathcal{K}_u^{\star \dot{+} (-\star)}(x^\sharp, \mathbf{X}^\sharp) \leq \mathcal{H}(\cdot, u, \mathbf{X}^\sharp)^\star(x^\sharp), \quad \forall u \in \mathbb{U}, \quad \forall (x^\sharp, \mathbf{X}^\sharp) \in \mathbb{C}^\sharp \times \mathbb{D}^\sharp. \quad (47)$$

Therefore, as we have just established that  $V_t(x) = \inf_{\mathbf{X} \in \mathbb{D}} \left( \inf_{u \in \mathbb{U}} \mathcal{K}_u(x, \mathbf{X}) \dot{+} \mathbb{E}[V_{t+1}(\mathbf{X})] \right)$ , we deduce from implication (4) that

$$\begin{aligned} V_t^\star(x^\sharp) &\leq \inf_{\mathbf{X}^\sharp} \left( \left( \inf_{u \in \mathbb{U}} \mathcal{K}_u \right)^{\star \dot{+} (-\star)}(x^\sharp, \mathbf{X}^\sharp) \dot{+} \mathbb{E}[V_{t+1}(\cdot)]^\star(\mathbf{X}^\sharp) \right) \\ &= \inf_{\mathbf{X}^\sharp} \left( \sup_{u \in \mathbb{U}} \left( \mathcal{K}_u^{\star \dot{+} (-\star)}(x^\sharp, \mathbf{X}^\sharp) \right) \dot{+} \mathbb{E}[V_{t+1}(\cdot)]^\star(\mathbf{X}^\sharp) \right) \end{aligned}$$

since  $(\inf_{u \in \mathbb{U}} \mathcal{K}_u)^{\star \dot{+} (-\star)} = \sup_{u \in \mathbb{U}} \mathcal{K}_u^{\star \dot{+} (-\star)}$  by the formula (58b)

$$\leq \inf_{\mathbf{X}^\sharp} \left( \sup_{u \in \mathbb{U}} \left( \mathcal{H}(\cdot, u, \mathbf{X}^\sharp)^\star(x^\sharp) \right) \dot{+} \mathbb{E}[V_{t+1}(\cdot)]^\star(\mathbf{X}^\sharp) \right)$$

as  $\mathcal{K}_u^{\star \dagger (-\star)}(x^\#, \mathbf{X}^\#) \leq \mathcal{H}(\cdot, u, \mathbf{X}^\#)^\star(x^\#)$  by (47)

$$\leq \inf_{\mathbf{X}^\#} \left( \sup_{u \in \mathbb{U}} \left( \mathcal{H}(\cdot, u, \mathbf{X}^\#)^\star(x^\#) \right) \dagger \mathbb{E}[V_{t+1}^\star(\mathbf{X}^\#)] \right),$$

as soon as we prove that  $\mathbb{E}[V_{t+1}(\cdot)]^\star(\mathbf{X}^\#) \leq \mathbb{E}[V_{t+1}^\star(\mathbf{X}^\#)]$ .

Indeed, we have that

$$\begin{aligned} \mathbb{E}[V_{t+1}(\cdot)]^\star(\mathbf{X}^\#) &= \sup_{\mathbf{X} \in \mathbb{D}} \left( \mathbb{E} \left[ \langle \mathbf{X}, \mathbf{X}^\# \rangle \right] \dagger (-\mathbb{E}[V_{t+1}(\mathbf{X})]) \right) \\ &= \sup_{\mathbf{X} \in \mathbb{D}} \left( \mathbb{E} \left[ \langle \mathbf{X}, \mathbf{X}^\# \rangle \dagger (-V_{t+1}(\mathbf{X})) \right] \right) \end{aligned}$$

because  $-\infty < \mathbb{E} \left[ \langle -\mathbf{X}, \mathbf{X}^\# \rangle \right] < +\infty$  by definition of the spaces  $\mathbb{D}$  and  $\mathbb{D}^\#$  and of the coupling (40) between them

$$\leq \mathbb{E} \left[ \sup_{\mathbf{X} \in \mathbb{D}} \left( \langle \mathbf{X}, \mathbf{X}^\# \rangle \dagger (-V_{t+1}(\mathbf{X})) \right) \right] = \mathbb{E}[V_{t+1}^\star(\mathbf{X}^\#)],$$

by definition of  $V_{t+1}^\star(\mathbf{X}^\#)$ . This ends the proof.  $\square$

Proposition 17 may be useful to obtain upper and lower estimates in approximations of Bellman functions. We just provide a sketch of the argument.

1. Suppose that the Bellman functions  $\{V_t\}_{t=0,1,\dots,T}$  satisfy the Bellman equation (42) and are convex l.s.c.. This is the case in *Stochastic Dual Dynamic Programming* (SDDP), when the dynamics  $F_t$  are jointly linear in state and control, the instantaneous costs  $L_t$  are jointly convex in state and control, the final cost  $K$  is convex, together with technical assumptions (see details in [19, 5] and references therein).
2. The Fenchel conjugates  $\{V_t^\star\}_{t=0,1,\dots,T}$  of the Bellman functions are convex l.s.c., by construction. Suppose that they satisfy a ‘‘Bellman like’’ equation

$$V_t^\star(x^\#) = \inf_{\mathbf{X}^\#} \left( \sup_{u \in \mathbb{U}} \left( \mathcal{H}(\cdot, u, \mathbf{X}^\#)^\star(x^\#) \right) \dagger \mathbb{E}[V_{t+1}^\star(\mathbf{X}^\#)] \right), \quad \forall t = T-1, \dots, 0, \quad (48)$$

which is (45), where the inequality is an equality. For this, one needs assumptions of the kind described in §2.3, as well as the equality  $\mathbb{E}[V_{t+1}(\cdot)]^\star(\mathbf{X}^\#) = \mathbb{E}[V_{t+1}^\star(\mathbf{X}^\#)]$  (see [13, 15]).

3. With the Bellman operators deduced from the Bellman equation (42) and ‘‘Bellman like’’ equation (45), one can easily produce piecewise linear lower bound functions  $\underline{V}_{t,(k)} \leq \underline{V}_{t,(k+1)} \leq V_t$  and  $\tilde{\underline{V}}_{t,(k)} \leq \tilde{\underline{V}}_{t,(k+1)} \leq V_t^\star$ , for  $k \in \mathbb{N}$ , by a proper algorithm (like the SDDP algorithm).

4. Since the Bellman functions  $\{V_t\}_{t=0,1,\dots,T}$  are convex l.s.c. and proper<sup>4</sup>, we deduce that

$$\underline{V}_{t,(k)} \leq \underline{V}_{t,(k+1)} \leq V_t = V_t^{**} \leq \tilde{V}_{t,(k+1)}^* \leq \tilde{V}_{t,(k)}^* . \quad (49)$$

Thus, we can control the evolution of the algorithm.

**Acknowledgements.** We want to thank Juan Enrique Martínez Legaz, Roger Wets, Lionel Thibault, Terry Rockafellar, Guillaume Carlier, Filippo Santambrogio, Jean-Paul Penot, Patrick L. Combettes, Marianne Akian, Stéphane Gaubert, Michel Théra, Michel Volle, for their comments on first versions of this work.

## A Appendix

### A.1 Background on J. J. Moreau lower and upper additions

When we manipulate functions with values in  $\overline{\mathbb{R}} = [-\infty, +\infty]$ , we adopt the following Moreau *lower addition* or *upper addition*, depending on whether we deal with sup or inf operations. We follow [10]. In the sequel,  $u, v$  and  $w$  are any elements of  $\overline{\mathbb{R}}$ .

#### Moreau lower addition

The Moreau *lower addition* extends the usual addition with

$$(+\infty) \dagger (-\infty) = (-\infty) \dagger (+\infty) = -\infty . \quad (50a)$$

With the *lower addition*,  $(\overline{\mathbb{R}}, \dagger)$  is a convex cone, with  $\dagger$  commutative and associative. The lower addition displays the following properties:

$$u \leq u' , \quad v \leq v' \Rightarrow u \dagger v \leq u' \dagger v' , \quad (50b)$$

$$(-u) \dagger (-v) \leq -(u \dagger v) , \quad (50c)$$

$$(-u) \dagger u \leq 0 , \quad (50d)$$

$$\sup_{a \in \mathbb{A}} f(a) \dagger \sup_{b \in \mathbb{B}} g(b) = \sup_{a \in \mathbb{A}, b \in \mathbb{B}} (f(a) \dagger g(b)) , \quad (50e)$$

$$\inf_{a \in \mathbb{A}} f(a) \dagger \inf_{b \in \mathbb{B}} g(b) \leq \inf_{a \in \mathbb{A}, b \in \mathbb{B}} (f(a) \dagger g(b)) , \quad (50f)$$

$$t < +\infty \Rightarrow \inf_{a \in \mathbb{A}} f(a) \dagger t = \inf_{a \in \mathbb{A}} (f(a) \dagger t) . \quad (50g)$$

#### Moreau upper addition

The Moreau *upper addition* extends the usual addition with

$$(+\infty) \dot{\dagger} (-\infty) = (-\infty) \dot{\dagger} (+\infty) = +\infty . \quad (51a)$$

---

<sup>4</sup>They are nonnegative and we exclude the degenerate case where they would have an empty domain.

With the *upper addition*,  $(\overline{\mathbb{R}}, \dot{+})$  is a convex cone, with  $\dot{+}$  commutative and associative. The upper addition displays the following properties:

$$u \leq u', \quad v \leq v' \Rightarrow u \dot{+} v \leq u' \dot{+} v', \quad (51b)$$

$$(-u) \dot{+} (-v) \geq -(u \dot{+} v), \quad (51c)$$

$$(-u) \dot{+} u \geq 0, \quad (51d)$$

$$\inf_{a \in \mathbb{A}} f(a) \dot{+} \inf_{b \in \mathbb{B}} g(b) = \inf_{a \in \mathbb{A}, b \in \mathbb{B}} (f(a) \dot{+} g(b)), \quad (51e)$$

$$\sup_{a \in \mathbb{A}} f(a) \dot{+} \sup_{b \in \mathbb{B}} g(b) \geq \sup_{a \in \mathbb{A}, b \in \mathbb{B}} (f(a) \dot{+} g(b)), \quad (51f)$$

$$-\infty < t \Rightarrow \sup_{a \in \mathbb{A}} f(a) \dot{+} t = \sup_{a \in \mathbb{A}} (f(a) \dot{+} t). \quad (51g)$$

### Joint properties of the Moreau lower and upper addition

We obviously have that

$$u \dot{+} v \leq u + v. \quad (52a)$$

The Moreau lower and upper additions are related by

$$-(u \dot{+} v) = (-u) \dot{+} (-v), \quad -(u + v) = (-u) \dot{+} (-v). \quad (52b)$$

They satisfy the inequality

$$(u \dot{+} v) \dot{+} w \leq u \dot{+} (v \dot{+} w). \quad (52c)$$

with

$$(u \dot{+} v) \dot{+} w < u \dot{+} (v \dot{+} w) \iff \begin{cases} u = +\infty \text{ and } w = -\infty, \\ \text{or} \\ u = -\infty \text{ and } w = +\infty \text{ and } -\infty < v < +\infty. \end{cases} \quad (52d)$$

Finally, we have that

$$u \dot{+} (-v) \leq 0 \iff u \leq v \iff 0 \leq v \dot{+} (-u). \quad (52e)$$

## A.2 Background on Fenchel-Moreau conjugacy with respect to a coupling

Let be given two sets  $\mathbb{C}$  and  $\mathbb{C}^\sharp$ . Consider a *coupling* function  $\Phi : \mathbb{C} \times \mathbb{C}^\sharp \rightarrow [-\infty, +\infty]$ . We also use the notation  $\mathbb{C} \overset{\Phi}{\leftrightarrow} \mathbb{C}^\sharp$  for a coupling, so that

$$\mathbb{C} \overset{\Phi}{\leftrightarrow} \mathbb{C}^\sharp \iff \Phi : \mathbb{C} \times \mathbb{C}^\sharp \rightarrow [-\infty, +\infty]. \quad (53)$$

**Definition 18** *The Fenchel-Moreau conjugate of a function  $f : \mathbb{C} \rightarrow [-\infty, +\infty]$ , with respect to the coupling  $\Phi$  in (53), is the function  $f^\Phi : \mathbb{C}^\sharp \rightarrow [-\infty, +\infty]$  defined by*

$$f^\Phi(c^\sharp) = \sup_{c \in \mathbb{C}} \left( \Phi(c, c^\sharp) \dot{+} (-f(c)) \right), \quad \forall c^\sharp \in \mathbb{C}^\sharp. \quad (54)$$



The Fenchel-Moreau biconjugate is the function  $f^{\Phi\Phi} : \mathbb{C} \rightarrow [-\infty, +\infty]$  defined by

$$f^{\Phi\Phi}(c) = \sup_{c^\sharp \in \mathbb{C}^\sharp} \left( \Phi(c, c^\sharp) \dagger (-f^\Phi(c^\sharp)) \right), \quad \forall c \in \mathbb{C}. \quad (55)$$

The following property is well known.

**Proposition 19** *For any function  $f : \mathbb{C} \rightarrow [-\infty, +\infty]$ , we have that*

$$f^{\Phi\Phi}(c) \leq f(c). \quad (56)$$

**Proof.** We prove (56) as follows.

$$\begin{aligned} f^{\Phi\Phi}(c) \dagger (-f(c)) &= \sup_{c^\sharp \in \mathbb{C}^\sharp} \left( \Phi(c, c^\sharp) \dagger (-f^\Phi(c^\sharp)) \right) \dagger (-f(c)) && \text{(by (55) and (50e))} \\ &= \sup_{c^\sharp \in \mathbb{C}^\sharp} \left( \left( \Phi(c, c^\sharp) \dagger (-f^\Phi(c^\sharp)) \right) \dagger (-f(c)) \right) && \text{(by (50e))} \\ &= \sup_{c^\sharp \in \mathbb{C}^\sharp} \left( \Phi(c, c^\sharp) \dagger (-f^\Phi(c^\sharp)) \dagger (-f(c)) \right) && \text{(by associativity of } \dagger \text{)} \\ &= \sup_{c^\sharp \in \mathbb{C}^\sharp} \left( \Phi(c, c^\sharp) \dagger (-f(c)) \dagger (-f^\Phi(c^\sharp)) \right) && \text{(by commutativity of } \dagger \text{)} \\ &\leq \sup_{c^\sharp \in \mathbb{C}^\sharp} \left( \sup_{c \in \mathbb{C}} \left( \Phi(c, c^\sharp) \dagger (-f(c)) \right) \dagger (-f^\Phi(c^\sharp)) \right) && \text{(by (50b))} \\ &= \sup_{c^\sharp \in \mathbb{C}^\sharp} \left( f^\Phi(c^\sharp) \dagger (-f^\Phi(c^\sharp)) \right) && \text{(by (54))} \\ &\leq 0. && \text{(by (50d))} \end{aligned}$$

We have obtained that  $f^{\Phi\Phi}(c) \dagger (-f(c)) \leq 0$ . Now, using (52e), we obtain (56). This ends the proof.  $\square$

The following properties are easy to establish.

**Proposition 20** *For any family  $\{f_u\}_{u \in \mathbb{U}}$  of functions  $f_u : \mathbb{C} \rightarrow [-\infty, +\infty]$ , we have that*

$$\left( \inf_{u \in \mathbb{U}} f \right)^\Phi(c) = \sup_{u \in \mathbb{U}} f^\Phi(c) \quad (58a)$$

$$-\left( \inf_{u \in \mathbb{U}} f \right)^\Phi(c) = \inf_{u \in \mathbb{U}} (-f^\Phi(c)). \quad (58b)$$

**Proposition 21** *Let be given two “primal” sets  $\mathbb{C}, \mathbb{D}$  and two “dual” sets  $\mathbb{C}^\sharp, \mathbb{D}^\sharp$ , together with two coupling functions*

$$\Phi : \mathbb{C} \times \mathbb{C}^\sharp \rightarrow [-\infty, +\infty], \quad \Psi : \mathbb{D} \times \mathbb{D}^\sharp \rightarrow [-\infty, +\infty]. \quad (59)$$

*For any bivariate function  $\mathcal{K} : \mathbb{C} \times \mathbb{D} \rightarrow [-\infty, +\infty]$ , we have that*

$$\left( - (c \mapsto \mathcal{K}(c, \cdot)^\Psi) \right)^\Phi = \mathcal{K}^{\Phi \dagger \Psi} = \left( - (d \mapsto \mathcal{K}(\cdot, d)^\Phi) \right)^\Psi. \quad (60)$$

## References

- [1] J.-P. Aubin and H. Frankowska. *Set-valued analysis*. Modern Birkhäuser Classics. Birkhäuser Boston, Inc., Boston, MA, 2009.
- [2] H. H. Bauschke and P. L. Combettes. *Convex analysis and monotone operator theory in Hilbert spaces*. CMS Books in Mathematics/Ouvrages de Mathématiques de la SMC. Springer-Verlag, second edition, 2017.
- [3] D. P. Bertsekas and S. E. Shreve. *Stochastic Optimal Control: The Discrete-Time Case*. Athena Scientific, Belmont, Massachusetts, 1996.
- [4] P. Carpentier, J.-P. Chancelier, G. Cohen, and M. De Lara. *Stochastic Multi-Stage Optimization. At the Crossroads between Discrete Time Stochastic Control and Stochastic Programming*. Springer-Verlag, Berlin, 2015.
- [5] P. Girardeau, V. Leclere, and A. B. Philpott. On the convergence of decomposition methods for multistage stochastic convex programs. *Mathematics of Operations Research*, 40(1):130–145, 2014.
- [6] V. L. Levin. Abstract convexity and the Monge-Kantorovich duality. In *Generalized convexity and related topics*, volume 583 of *Lecture Notes in Econom. and Math. Systems*, pages 33–72. Springer, Berlin, 2007.
- [7] M. Loève. *Probability Theory I*. Springer-Verlag, New York, fourth edition, 1977.
- [8] J. E. Martínez-Legaz. Generalized convex duality and its economic applications. In S. S. Hadjisavvas N., Komlósi S., editor, *Handbook of Generalized Convexity and Generalized Monotonicity. Nonconvex Optimization and Its Applications*, volume 76, pages 237–292. Springer-Verlag, 2005.
- [9] M.-L. Mazure and M. Volle. Equations inf-convolutives et conjugaison de moreau-fenchel. *Ann. Fac. Sc. Toulouse*, 12(1):103–125, 1991.
- [10] J.-J. Moreau. Inf-convolution, sous-additivité, convexité des fonctions numériques. *J. Math. Pures Appl. (9)*, 49:109–154, 1970.
- [11] D. Pallaschke and S. Rolewicz. *Foundations of mathematical optimization*, volume 388 of *Mathematics and its Applications*. Kluwer Academic Publishers Group, Dordrecht, 1997.
- [12] T. Pennanen and A.-P. Perkkiö. Shadow price of information in discrete time stochastic optimization. *Mathematical Programming*, 168(1):347–367, Mar 2018.
- [13] R. Rockafellar. Integrals which are convex functionals. *Pacific J. Math.*, 24(3):525–539, 1968.

- [14] R. T. Rockafellar. Extension of Fenchel's duality theorem for convex functions. *Duke Math. J.*, 33:81–89, 1966.
- [15] R. T. Rockafellar. Integrals which are convex functionals, II. *Pacific J. Math.*, 39(2):439–469, 1971.
- [16] R. T. Rockafellar and R. J.-B. Wets. *Variational Analysis*. Springer-Verlag, Berlin, 1998.
- [17] A. Rubinov. *Abstract convexity and global optimization*, volume 44 of *Nonconvex Optimization and its Applications*. Kluwer Academic Publishers, Dordrecht, 2000.
- [18] F. Santambrogio. *Optimal Transport for Applied Mathematicians: Calculus of Variations, PDEs, and Modeling*. Progress in Nonlinear Differential Equations and Their Applications. Birkhäuser, 2015.
- [19] A. Shapiro. Analysis of stochastic dual dynamic programming method. *European J. Oper. Res.*, 209(1):63–72, 2011.
- [20] I. Singer. *Abstract convex analysis*. Canadian Mathematical Society Series of Monographs and Advanced Texts. John Wiley & Sons, Inc., New York, 1997.
- [21] M. Volle. Multiapplications duales et problèmes d'optimisation en dualité. *Comptes Rendus de l'Académie des Sciences Paris, Série I*(t. 296):11–14, 1983.