Pricing, Hedging and Optimally Designing Derivatives via Minimization of Risk Measures [1]

11th July 2005

Pauline Barrieu [2] and Nicole El Karoui [3]

The question of pricing and hedging a given contingent claim has a unique solution in a complete market framework. When some incompleteness is introduced, the problem becomes however more difficult. Several approaches have been adopted in the literature to provide a satisfactory answer to this problem, for a particular choice criterion. Among them, Hodges and Neuberger [72] proposed in 1989 a method based on utility maximization. The price of the contingent claim is then obtained as the smallest (resp. largest) amount leading the agent indifferent between selling (resp. buying) the claim and doing nothing. The price obtained is the indifference seller's (resp. buyer's) price. Since then, many authors have used this approach, the exponential utility function being most often used (see for instance, El Karoui and Rouge [51], Becherer [11], Delbaen et al. [39] , Musiela and Zariphopoulou [93] or Mania and Schweizer [89]...).

In this chapter, we also adopt this exponential utility point of view to start with in order to find the optimal hedge and price of a contingent claim based on a non-tradable risk. But soon, we notice that the right framework to work with is not that of the exponential utility itself but that of the certainty equivalent which is a convex functional satisfying some nice properties among which that of cash translation invariance. Hence, the results obtained in this particular framework can be immediately extended to functionals satisfying the same properties, in other words to convex risk measures as introduced by Föllmer and Schied [53] and [54] or by Frittelli and Gianin [57]. Starting with a utility maximization problem, we end up with an equivalent risk measure minimization in order to price and hedge this contingent claim.

Moreover, this hedging problem can be seen as a particular case of a more general situation of risk transfer between different agents, one of them consisting of the financial market. Therefore, we consider in this

[1] The authors would like to thank their respective husband, Laurent Veilex and Faycal El Karoui, for their constant support and incredible patience.

[2] L.S.E., Statistics Department, Houghton Street, London, WC2A 2AE, United Kingdom. (e-mail: p.m.barrieu@lse.ac.uk)
Research supported in part by the EPSRC, under Grant GR-T23879/01.

[3] C.M.A.P., Ecole Polytechnique, 91128 Palaiseau Cédex, France. (e-mail: elkaroui@cmapx.polytechnique.fr)

chapter the general question of optimal transfer of a non-tradable risk and specify the results obtained in the particular situation of an optimal hedging problem.

Both static and dynamic approaches are considered in this chapter, in order to provide constructive answers to this optimal risk transfer problem. Quite recently, many authors have studied dynamic version of static risk measures (see for instance, among many other references, Cvitanic and Karatzas [35], Scandolo [107], Weber [112], Artzner et al. [3], Cheridito, Delbaen and Kupper [29], Frittelli and Gianin [58], Gianin [61], Riedel [102] or Peng [98]). When considering a dynamic framework, our main purpose is to find a trade-off between static and very abstract risk measures as we are more interested in tractability issues and interpretations of the dynamic risk measures we obtain rather than the ultimate general results. Therefore, after introducing a general axiomatic approach to dynamic risk measures, we relate the dynamic version of convex risk measures to BSDEs. For the sake of a better understanding, a whole section in the second part is dedicated to some key results and properties of BSDEs, which are essential to this definition of dynamic convex risk measures.

Part I: Static Framework

In this chapter, we focus on the question of optimal hedging of a given risky position in an incomplete market framework. However, instead of adopting a standard point of view, we look at it in terms of an optimal risk transfer between different economic agents, one of them being possibly a financial market.

The risk that we consider here is not (directly) traded on any financial market. We may think for instance of a weather risk, a catastrophic risk (natural catastrophe, terrorist attack...) but also of any global insurance risk that may be securitized, such as the longevity of mortality risk...

First adopting a static point of view, we proceed in several steps. In a first section, we relate the notion of indifference pricing rule to that of transaction feasibility, capital requirement, hedging and naturally introduce convex risk measures. Then, after having introduced some key operations on convex risk measures, in particular the dilatation and the inf-convolution, we study the problem of optimal risk transfer between two agents. We see how the risk transfer problem can be reduced to an inf-convolution problem of convex functionals. We solve it explicitly in the dilated framework and give some necessary and sufficient conditions in the general framework.

1 Indifference Pricing, Capital Requirement and Convex Risk Measures

As previously mentioned in the introduction, since 1989 and the seminal paper by Hodges and Neuberger [72] indifference pricing based on a utility criterion has been a popular (academic) method to value claims in an incomplete market. Taking the buyer point of view, the indifference price corresponds to the maximal amount π, the agent having a utility function u is ready to pay for a claim X. In other words, π is determined as the amount the agent pays such that her expected utility remains unchanged when doing the transaction:

$$\mathbb{E}[u(X-\pi)] = u(0).$$

This price is not a transaction price. It gives an upper bound (for the buyer) to the price of this claim so that a transaction will take place. π also corresponds to the certainty equivalent of the claim payoff X. Certain properties this indifference price should have are rather obvious: first it should be an increasing function of X but also a convex function in order to take into account the diversification aspect of considering a portfolio of different claims rather than the sum of different individual portfolios. Another property which is rather interesting is the cash translation invariance property. More precisely, it seems natural to consider the situation where translating the payoff of the claim X by a non-risky amount m simply leads to a translation of the price π by the same amount. It is the case, as we will see for the exponential utility in the following subsection.

1.1 The Exponential Utility Framework

First, let us notice that exponential utility functions have been widely used in the financial literature. Several facts may justify their relative importance compared to other utility functions but, in particular, the absence of constraint on the sign of the future considered cash flows and its relationship with probability measures make them very convenient to use.

1.1.1 Indifference Pricing Rule

In this introductory subsection, we simply consider an agent, having an exponential utility function $U(x) = -\gamma \exp\big(-\frac{1}{\gamma}x\big)$, where γ is her risk tolerance coefficient. She evolves in an uncertain universe modelled by a standard probability space $(\Omega, \Im, \mathbb{P})$ with time horizon T. The wealth W of the agent at this future date T is uncertain, since W can be seen as a particular position on a given portfolio or as the book of the agent. To reduce her risk, she can decide whether or not to buy a contingent claim with a payoff X at time T. For the sake of simplicity, we neglect interest rate between 0 and T and assume that both random variables W and X are bounded.

In order to decide whether or not she will buy this claim, she will find the maximum price she is ready to pay for it, her indifference price $\pi(X)$ for the claim X given by the constraint $\mathbb{E}_{\mathbb{P}}\big[U(W+X-\pi(X))\big] = \mathbb{E}_{\mathbb{P}}\big[U(W)\big]$. Then,

$$\begin{aligned} \mathbb{E}_{\mathbb{P}}\big[\exp\big(-\frac{1}{\gamma}(W+X-\pi(X))\big)\big] &= \mathbb{E}_{\mathbb{P}}\big[\exp\big(-\frac{1}{\gamma}W\big)\big] \\ \Leftrightarrow \quad \pi(X|W) &= e_\gamma\big(W\big) - e_\gamma\big(W+X\big) \end{aligned}$$

where e_γ is the opposite of the certainty equivalent, defined for any bounded random variable Ψ

$$e_\gamma\big(\Psi\big) \triangleq \gamma \ln \mathbb{E}_{\mathbb{P}}\Big[\exp\big(-\frac{1}{\gamma}\Psi\big)\Big]. \tag{1}$$

The indifference pricing rule $\pi(X|W)$ has the desired property of increasing monotonicity, convexity and translation invariance: $\pi(X+m|W) = \pi(X|W) + m$. Moreover, the functional $e_\gamma(X) = -\pi(X|W=0)$ has similar properties; it is decreasing, convex and translation invariant in the following sense: $e_\gamma(\Psi + m) = e_\gamma(\Psi) - m$.

1.1.2 Some Remarks on the "Price" $\pi(X)$

$\pi(X)$ does not correspond to a transaction price but simply gives an indication of the transaction price range since it corresponds to the maximal amount the agent is ready to pay for the claim X and bear the associated risk given her initial exposure. This dependency seems quite intuitive: for instance, the considered agent can be seen as a trader who wants to buy the particular derivative X without knowing its price. She determines it by considering the contract *relatively* to her existing book.
For this reason and for the sake of a better understanding, we will temporarily denote it by $\pi^b(X|W)$, the upper-script "b" standing for "buyer". This heavy notation underlines the close relationship between the pricing rule and the actual exposure of the agent. The considered framework is symmetric since there is no particular requirement on the sign of the different quantities. Hence, it is possible to define by simple analogy the *indifference seller's price* of the claim X. Let us denote it by $\pi^s(X|W)$, the upper-script "s" standing for "seller". Both seller's and buyer's indifference pricing rules are closely related as

$$\pi^s(X|W) = -\pi^b(-X|W).$$

Therefore, the seller's price of X is simply the opposite of the buyer's price of $-X$.
Such an axiomatic approach of the pricing rule is not new. This was first introduced in insurance under the name of *convex premium principle* (see for instance the seminal paper of Deprez and Gerber [42] in 1985) and then developed in continuous time finance (see for instance El Karoui and Quenez [49]).
When adopting an exponential utility criterion to solve a pricing problem, the right framework to work with seems to be that of the functional e_γ and not directly that of utility. This functional, called entropic risk measure, holds some key properties of convexity, monotonicity and cash translation invariance. It is therefore possible to generalize the utility criterion to focus more on the notion of price keeping in mind these wished properties. The convex risk measure provides such a criterion as we will see in the following.

1.2 Convex Risk Measures: Definition and Basic Properties

Convex risk measures can have two possible interpretations depending on the representation which is used: they can be considered either as a pricing rule or as a capital requirement rule. We will successively present both of them in the following, introducing each time the vocabulary associated with this particular approach.

1.2.1 Risk Measure as an Indifference Price

We first recall the definition and some key properties of the convex risk measures introduced by Föllmer and Schied [53] and [54]. The notations, definitions and main properties may be found in this last reference [54]. In particular, we assume that uncertainty is described through a measurable space $(\Omega, \Im)$, and that risky positions belong to the linear space of bounded functions (including constant functions), denoted by $\mathcal{X}$.

Definition 1.1 *The functional $\rho : \mathcal{X} \to \mathbb{R}$ is a* (monetary) convex risk measure *if, for any Φ and Ψ in $\mathcal{X}$, it satisfies the following properties:*

a) Convexity: $\forall \lambda \in [0,1] \quad \rho\big(\lambda\Phi + (1-\lambda)\Psi\big) \le \lambda\rho(\Phi) + (1-\lambda)\rho(\Psi)$;

b) Monotonicity: $\Phi \le \Psi \Rightarrow \rho(\Phi) \ge \rho(\Psi)$;

c) Translation invariance: $\forall m \in \mathbb{R} \quad \rho(\Phi + m) = \rho(\Phi) - m$.

A convex risk measure ρ is coherent if it satisfies also:

d) Homogeneity : $\forall \lambda \in \mathbb{R}^+ \quad \rho\big(\lambda\,\Phi\big) = \lambda\rho\big(\Phi\big)$.

Note that the convexity property is essential: this translates the natural fact that diversification should not increase risk. In particular, any convex combination of "admissible" risks should be "admissible". One of the major drawbacks of the famous risk measure VAR (Value at Risk) is its failure to meet this criterion. This may lead to arbitrage opportunities inside the financial institution using it as risk measure as observed by Artzner, Delbaen, Eber and Heath in their seminal paper [2].

Intuitively, given the translation invariance, $\rho(X)$ may be interpreted as the amount the agent has to hold to completely cancel the risk associated with her risky position X since

$$\rho(X + \rho(X)) = \rho(X) - \rho(X) = 0. \tag{2}$$

$\rho(X)$ can be also considered as the opposite of the "buyer's indifference price" of this position, since when paying the amount $-\rho(X)$, the new exposure $X - (-\rho(X))$ does not carry any risk with positive measure, i.e. the agent is somehow indifferent using this criterion between doing nothing and having this "hedged" exposure.

The convex risk measures appear therefore as a natural extension of utility functions as they can be seen directly as an indifference pricing rule.

1.2.2 Dual Representation

In order to link more closely both notions of pricing rule and risk measure, the duality between the Banach space $\mathcal{X}$ endowed with the supremum norm $\|.\|$ and its dual space $\mathcal{X}'$, identified with the set $\mathbf{M}^{\text{ba}}$ of finitely additive set functions with finite total variation on $(\Omega, \Im)$, can be used as it leads to a dual representation. The properties of monotonicity and cash invariance allow to restrict the domain of the dual functional to the set $\mathbf{M}_{1,f}$ of all finitely additive measures (Theorem 4.12 in [54]). The following theorem gives an "explicit" formula for the risk measure (and as a consequence for the price) in terms of expected values:

Theorem 1.2 *Let $\mathbf{M}_{1,f}$ be the set of all finitely additive measures on $(\Omega, \Im)$, and $\alpha\big(\mathbf{Q}\big)$ the minimal penalty function taking values in $\mathbb{R} \cup \big\{+\infty\big\}$:*

$$\forall \mathbf{Q} \in \mathbf{M}_{1,f} \quad \alpha\big(\mathbf{Q}\big) = \sup_{\Psi \in \mathcal{X}} \left\{ \mathbb{E}_{\mathbf{Q}}[-\Psi] - \rho(\Psi) \right\} \qquad \big(\ge -\rho(0)\big). \tag{3}$$

$$\mathbf{Dom}(\alpha) = \{\mathbf{Q} \in \mathbf{M}_{1,f} |\ \alpha\big(\mathbf{Q}\big) < +\infty\} \tag{4}$$

The Fenchel duality relation holds :

$$\forall \Psi \in \mathcal{X} \quad \rho(\Psi) = \sup_{\mathbf{Q} \in \mathbf{M}_{1,f}} \big\{ \mathbb{E}_{\mathbf{Q}}[-\Psi] - \alpha\big(\mathbf{Q}\big) \big\} \tag{5}$$

Moreover, for any $\Psi \in \mathcal{X}$ there exists an optimal additive measure $\mathbf{Q}_\Psi \in \mathbf{M}_{1,f}$ such that

$$\rho(\Psi) = \mathbb{E}_{\mathbf{Q}_\Psi}[-\Psi] - \alpha\big(\mathbf{Q}_\Psi\big) = \max_{\mathbf{Q}\in\mathbf{M}_{1,f}} \big\{\mathbb{E}_{\mathbf{Q}}[-\Psi] - \alpha\big(\mathbf{Q}\big)\big\}.$$

Henceforth, $\alpha(\mathbf{Q})$ is the minimal penalty function, denoted by $\alpha_{\min}(\mathbf{Q})$ in [54].
The dual representation of ρ given in Equation (5) emphasizes the interpretation in terms of a worst case related to the agent's (or regulator's) beliefs.

Convex Analysis Point of view We start with Remark 4.17 and the Appendices 6 and 7 in [54]. The penalty function α defined in (5) corresponds to the Fenchel-Legendre transform on the Banach space $\mathcal{X}$ of the convex risk-measure ρ. The dual space $\mathcal{X}'$ can be identified with the set $\mathbf{M}^{\mathrm{ba}}$ of finitely additive set functions with finite total variation. Then the subset $\mathbf{M}_{1,f}$ of "finite probability measure" is weak*-compact in $\mathcal{X}' = \mathbf{M}^{\mathrm{ba}}$ and the functional $\mathbf{Q} \to \alpha(\mathbf{Q})$ is weak*-lower semi-continuous (or weak*-closed) as supremum of affine functionals. This terminology from convex analysis is based upon the observation that lower semi-continuity and the closure of the level sets $\{\phi \le c\}$ are equivalent properties. Moreover ρ is lower semi-continuous (lsc) with respect to the weak topology $\sigma(\mathcal{X}, \mathcal{X}')$ since any set $\{\rho \le c\}$ is convex and strongly closed given that ρ is strongly Lipschitz-continuous. Then, general duality theorem for conjugate functional yields to

$$\rho(\Psi) = \sup_{r\in\mathbf{M}^{\mathrm{ba}}} (r(\Psi) - \rho^*(r)), \qquad \rho^*(r) = \sup_{\Psi\in\mathcal{X}} (r(\Psi) - \rho(\Psi))$$

with the convention $r_{\mathbf{Q}}(\Psi) = \mathbb{E}_{\mathbf{Q}}[-\Psi]$ for $\mathbf{Q} \in \mathbf{M}_{1,f}$. We then use the properties of monotonicity and cash invariance of ρ to prove that when $\rho^*(r) < +\infty$, $-r \in \mathbf{M}_{1,f}$. Moreover by weak*-compacity of $\mathbf{M}_{1,f}$, the upper semi-continuous functional $\mathbb{E}_{\mathbf{Q}}[-\Psi] - \alpha(\mathbf{Q})$ attains its maximum on $\mathbf{M}_{1,f}$.
In the second part of this chapter, we will intensively used convex analysis point of view when studying dynamic convex risk measures.

Duality and Probability Measures We are especially interested in the risk measures that admit a representation (5) in terms of σ-additive probability measures $\mathbb{Q}$. In this paper, for the sake of simplicity and clarity, we use the notation $\mathbf{Q} \in \mathbf{M}_{1,f}$ when dealing with additive measures and $\mathbb{Q} \in \mathcal{M}_1$ when considering probability measures. So, we are looking for the following representation on $\mathcal{X}$

$$\rho(\Psi) = \sup_{\mathbb{Q}\in\mathcal{M}_1} \big\{\mathbb{E}_{\mathbb{Q}}[-\Psi] - \alpha\big(\mathbb{Q}\big)\big\}. \tag{6}$$

We can no longer expected that the supremum is attained without additional assumptions. Such representation on $\mathcal{M}_1$ is closely related to some continuity properties of the convex functional ρ (Lemma 4.20 and Proposition 4.21 in [54]).

Proposition 1.3 *i) Any convex risk measure ρ defined on $\mathcal{X}$ and satisfying (6) is continuous from above, in the sense that*

$$\Psi_n \searrow \Psi \quad \Longrightarrow \quad \rho\big(\Psi_n\big) \nearrow \rho(\Psi).$$

ii) The converse is not true in general, but holds under continuity from below assumption:

$$\Psi_n \nearrow \Psi \quad \Longrightarrow \quad \rho\big(\Psi_n\big) \searrow \rho(\Psi).$$

Then any additive measure $\mathbf{Q}$ *such that* $\alpha(\mathbf{Q}) < +\infty$ *is* σ*-additive and* (6) *holds true. Moreover, from i),* ρ *is also continuous by above.*

1.2.3 Risk Measures on $\mathbb{L}_\infty(\mathbb{P})$

The representation theory on $\mathbb{L}_\infty(\mathbb{P})$ was developed in particular by Delbaen [38] and extended by Frittelli and Gianin [57] and [58] and [54]. When a probability measure $\mathbb{P}$ is given, it is natural indeed to define risk measures ρ on $\mathbb{L}_\infty(\mathbb{P})$ instead of on $\mathcal{X}$ satisfying the compatibility condition:

$$\rho(\Psi) = \rho(\Phi) \quad \text{if} \quad \Psi = \Phi \quad \mathbb{P} - a.s. \tag{7}$$

Let us introduce some new notations: $\mathbf{M}_{1,ac}(\mathbb{P})$ is the set of finitely additive measures absolutely continuous w. r. to $\mathbb{P}$ and $\mathcal{M}_{1,ac}(\mathbb{P})$ is the set of probability measures absolutely continuous w. r. to $\mathbb{P}$.
We also define natural extension of continuity from below in the space $\mathbb{L}_\infty(\mathbb{P})$: ($\Psi_n \searrow \Psi \quad \mathbb{P} - a.s. \Longrightarrow \rho(\Psi_n) \nearrow \rho(\Psi)$), or continuity from above in the space $\mathbb{L}_\infty(\mathbb{P})$: ($\Psi_n \nearrow \Psi \quad \mathbb{P} - a.s. \Longrightarrow \quad \rho(\Psi_n) \searrow \rho(\Psi)$). These additional results on conjugacy relations are given in [54] Theorem 4.31 and in Delbaen [38] Corollary 4.35). Sometimes, as in [38], the continuity from above is called the Fatou property.

Theorem 1.4 *Let* $\mathbb{P}$ *be a given probability measure.*

1. *Any convex risk measure* ρ *on* $\mathcal{X}$ *satisfying* (7) *may be considered as a risk measure on* $\mathbb{L}_\infty(\mathbb{P})$*. A dual representation holds true in terms of absolutely continuous additive measures* $\mathbf{Q} \in \mathbf{M}_{1,ac}(\mathbb{P})$*.*
2. ρ *admits a dual representation on* $\mathcal{M}_{1,ac}(\mathbb{P})$*:*

$$\alpha(\mathbb{Q}) = \sup_{\Psi \in \mathbb{L}_\infty(\mathbb{P})} \{\mathbb{E}_\mathbb{Q}[-\Psi] - \rho(\Psi)\}, \qquad \rho(\Psi) = \sup_{\mathbb{Q} \in \mathcal{M}_{1,ac}(\mathbb{P})} \{\mathbb{E}_\mathbb{Q}[-\Psi] - \alpha(\mathbb{Q})\}$$

   if and only if one of the equivalent properties holds:
   - *a)* ρ *is continuous from above (Fatou property);*
   - *b)* ρ *is closed for the weak*-topology* $\sigma(\mathbb{L}_\infty, \mathbb{L}_1)$*;*
   - *c) the acceptance set* $\{\rho \leq 0\}$ *is weak*-closed in* $\mathbb{L}_\infty(\mathbb{P})$*.*
3. *Assume that* ρ *is a coherent (homogeneous) risk measure, satisfying the Fatou property. Then,*

$$\rho(\Psi) = \sup_{\mathbb{Q} \in \mathcal{M}_{1,ac}(\mathbb{P})} \{\mathbb{E}_\mathbb{Q}[-\Psi] \mid \alpha(\mathbb{Q}) = 0\} \tag{8}$$

   The supremum in (8) *is a* **maximum** *iff one of the following equivalent properties holds:*
   - *a)* ρ *is continuous from below;*
   - *b) the convex set* $\mathcal{Q} = \{\mathbb{Q} \in \mathcal{M}_{1,ac} | \, \alpha(\mathbb{Q}) = 0\}$ *is weakly compact in* $\mathbb{L}^1(\mathbb{P})$*.*

According to the Dunford-Pettis theorem, the weakly relatively compact sets of $\mathbb{L}^1(\mathbb{P})$ are sets of uniformly integrable variables and La Vallée-Poussin gives a criterion to check this property. Therefore, the subset $\mathcal{A}$ of $\mathbb{L}^1(\mathbb{P})$ is weakly relatively compact iff it is closed and uniformly integrable. Moreover, according to the La Vallée-Poussin criterion, an increasing convex continuous function $\Phi : \mathbb{R}_+ \to \mathbb{R}$, (also called Young's function) such that:

$$\lim_{x \to \infty} \frac{\Phi(x)}{x} = +\infty \qquad \text{and} \qquad \sup_{\mathbb{Q} \in \mathcal{A}} \mathbb{E}_\mathbb{P} \left[\Phi(\frac{d\mathbb{Q}}{d\mathbb{P}}) \right] < +\infty.$$

1.3 Comments on Measures of Risk and Examples

1.3.1 About Value at Risk

Risk measures, just as utility functions, go beyond the simple problem of pricing. Both are inherently a choice or decision *criterion*. More precisely, when assessing the risk related to a given position in order to define the amount of capital requirement, a first natural approach is based on the distribution of the risky position itself. In this framework, the most classical measure of risk is simply the *variance* (or the mean-variance analysis). However, it does not take into account the whole distribution's features (as asymmetry or skewness) and especially it does not focus on the "real" financial risk which is the downside risk. Therefore different methods have been developed to focus on the risk of losses: the most widely used (as it is recommended to bankers by many financial institutions) is the so-called *Value at Risk* (denoted by VAR), based on quantiles of the lower tail of the distribution. More precisely, the VAR associated with the position X at a level ε is defined as

$$VAR_\varepsilon(X) = \inf\big\{k : \mathbb{P}(X + k < 0) \leq \varepsilon\big\}.$$

The VAR corresponds to the minimal amount to be added to a given position to make it acceptable. Such a criterion satisfies the key properties of decreasing monotonicity, translation invariance since $\forall m \in \mathbb{R}$, $VAR_\varepsilon(X + m) = VAR_\varepsilon(X) - m$ and finally, the VAR is positive homogeneous as $\forall \lambda \geq 0$, $VAR_\varepsilon(\lambda X) = \lambda VAR_\varepsilon(X)$.

This last property reflects the linear impact of the size of the position on the risk measure. However, as noticed by Artzner et al. [2] this criterion fails to meet a natural consistency requirement: it is not a convex risk measure while the convexity property translates the natural fact that diversification should not increase risk. In particular, any convex combination of "admissible" risks should be "admissible". The absence of convexity of the VAR may lead to arbitrage opportunities inside the financial institution using such criterion as risk measure. Based on this logic, Artzner et al. [2] have adopted a more general approach to risk measurement. Their paper is essential as it has initiated a systematic axiomatic approach to risk measurement. A *coherent measure of risk* should be convex and satisfy the three key properties of the VAR

Conditional Value at Risk For instance, a coherent version of the Value at Risk is the so-called *Conditional Value at Risk* as observed by Rockafellar and Uryasev [104]. This risk measure is denoted by $CVAR_\varepsilon$ and defined as

$$CVAR_\lambda(X) = \inf_K \mathbb{E}\Big[\frac{1}{\lambda}(X - K)^- - K\Big].$$

This coincide with the *Expected Shortfall* under some assumptions for the X-distribution (for more details, see Corollary 5.3 in Acerbi and Tasche [1]). In this case, the $CVAR$ can be written as

$$CVAR_\lambda(X) = \mathbb{E}[-X|X + VAR_\lambda(X) < 0].$$

Moreover, the $CVAR$ also coincides with another coherent version of the VAR, called *Average Value at Risk* and denoted by $AVAR$. This risk measure is defined as:

$$AVAR_\lambda(\Psi) = \frac{1}{\lambda}\int_0^\lambda VAR_\epsilon(\Psi)d\epsilon.$$

For more details, please refer for instance to Föllmer and Schied [54] (Proposition 4.37).

More recently, the axiom of positive homogeneity has been questioned. Indeed, such a condition does not seem to be compatible with the notion of liquidity risk existing on the market as it implies that the size of the risky position has simply a linear impact on the risk measure. To tackle this shortcoming, Föllmer and Schied consider, in [53] and [54], *convex risk measures* as previously defined.

1.3.2 Risk Measures and Utility Functions

Entropic Risk Measure The most famous convex risk measure on $\mathbb{L}_\infty(\mathbb{P})$ is certainly the entropic risk measure defined as the functional e_γ in the previous section when considering an exponential utility framework. The dual formulation of this continuous from below functional justifies the name of *entropic risk measure* since:

$$\forall \Psi \in \mathbb{L}_\infty(\mathbb{P}) \qquad e_\gamma(\Psi) = \gamma \ln \mathbb{E}_{\mathbb{P}}\Big[\exp\big(-\frac{1}{\gamma}\Psi\big)\Big] = \sup_{\mathbb{Q}\in\mathcal{M}_1} \big\{\mathbb{E}_{\mathbb{Q}}[-\Psi] - \gamma h\big(\mathbb{Q}|\mathbb{P}\big)\big\}$$

where $h(\mathbb{Q}|\mathbb{P})$ is the relative entropy of $\mathbb{Q}$ with respect to the prior probability measure $\mathbb{P}$, defined by

$$h\big(\mathbb{Q}|\mathbb{P}\big) = \mathbb{E}_{\mathbb{P}}\Big[\frac{d\mathbb{Q}}{d\mathbb{P}} \ln \frac{d\mathbb{Q}}{d\mathbb{P}}\Big] \quad \text{if } \mathbb{Q} \ll \mathbb{P} \qquad \text{and} \qquad +\infty \qquad \text{otherwise.}$$

Since e_γ is continuous from below in $\mathbb{L}_\infty(\mathbb{P})$, by the previous theorem

$$\forall \Psi \in \mathbb{L}_\infty(\mathbb{P}) \qquad e_\gamma(\Psi) = \gamma \ln \mathbb{E}_{\mathbb{P}}\Big[\exp\big(-\frac{1}{\gamma}\Psi\big)\Big] = \mathbf{max}_{\mathbb{Q}\in\mathcal{M}_{1,ac}} \big\{\mathbb{E}_{\mathbb{Q}}[-\Psi] - \gamma h\big(\mathbb{Q}|\mathbb{P}\big)\big\}$$

As previously mentioned in Paragraph 1.1.1, this particular convex risk measure is closely related to the exponential utility function and to the associated indifference price. However, the relationships between risk measures and utility functions can be extended.

Risk Measures and Utility Functions More generally, risk measures and utility functions have close relationships based on the hedging and super-replication problem. It is however possible to obtain a more general connection between them using the notion of shortfall risk.
More precisely, any agent having a utility function U assesses her risk by taking the expected utility of the considered position $\Psi \in \mathbb{L}_\infty(\mathbb{P})$: $\mathbb{E}_{\mathbb{P}}[U(\Psi)]$. If she focuses on her "real" risk, which is the downside risk, it is natural to consider instead the *loss function* $\mathcal{L}$ defined by $\mathcal{L}(x) = -U(-x)$ ([54] Section 4.9). As a consequence, $\mathcal{L}$ is a convex and increasing function and maximizing the expected utility is equivalent to minimize the expected loss (also called the *shortfall risk*), $\mathbb{E}_{\mathbb{P}}[\mathcal{L}(-\Psi)]$.
It is then natural to introduce the following risk measure as the opposite of the indifference price:

$$\rho(X) = \inf\{m \in \mathbb{R} \big| \, \mathbb{E}_{\mathbb{P}}[\mathcal{L}(-\Psi - m)] \leq l(0)\}.$$

Moreover, there is an explicit formula for the associated penalty function given in terms of the Fenchel-Legendre transform $\mathcal{L}^*(y) = \sup\{-xy - l\mathcal{L}(x)\}$ of the convex function $\mathcal{L}$ ([54] Theorem 4.106):

$$\alpha(\mathbb{Q}) = \inf_{\lambda > 0} \Big\{\frac{1}{\lambda}\Big(\mathcal{L}(0) + \mathbb{E}_{\mathbb{P}}\big[\mathcal{L}^*\big(\lambda \frac{d\mathbb{Q}}{d\mathbb{P}}\big)\big]\Big)\Big\}.$$

1.4 Risk Measures and Hedging

In this subsection, we come back to the possible interpretation of the risk measure $\rho(X)$ in terms of capital requirement. This leads also to a natural relationship between risk measure and hedging. We then extend it to a wider perspective of super-hedging.

1.4.1 Risk Measure and Capital Requirement

Looking back at Equation (2), the risk measure $\rho(X)$ gives an assessment of the minimal *capital requirement* to be added to the position as to make it acceptable in the sense that the new position (X and the added capital) does not carry any risk with non-negative measure any more. More formally, it is natural to introduce the *acceptance set* $\mathcal{A}_\rho$ related to ρ defined as the set of all acceptable positions in the sense that they do not require any additional capital:

$$\mathcal{A}_\rho = \big\{\Psi \in \mathcal{X}, \quad \rho(\Psi) \leq 0\big\}. \tag{9}$$

Given that the epigraph of the convex risk measure ρ is $\text{epi}(\rho) = \{(\Psi, m) \in \mathcal{X} \times \mathbb{R}\big| \rho(\Psi) \leq m\} = \{(\Psi, m) \in \mathcal{X} \times \mathbb{R}\big| \rho(\Psi + m) \leq 0\}$, the characterization of ρ in terms of $\mathcal{A}_\rho$ is easily obtained

$$\rho(X) = \inf\big\{m \in \mathbb{R}; m + X \in \mathcal{A}_\rho\big\}.$$

This last formulation makes very clear the link between risk measure and capital requirement.
From the definition of both the convex risk measure ρ and the acceptance set $\mathcal{A}_\rho$ and the dual representation of the risk measure ρ, it is possible to obtain another characterization of the associated penalty function α as:

$$\alpha\big(\mathbf{Q}\big) = \sup_{\Psi\in\mathcal{A}_\rho} \mathbb{E}_{\mathbf{Q}}[-\Psi], \quad \text{if} \quad \mathbf{Q}{\in}\mathbf{M}_{1,f}, \qquad = +\infty, \quad \text{if not.} \tag{10}$$

$\alpha(\mathbf{Q})$ is the support function of $-\mathcal{A}_\rho$, denoted by $\Sigma^{\mathcal{A}_\rho}(\mathbf{Q})$. When $\mathcal{A}_\rho$ is a cone, i.e. ρ is a coherent (positive homogeneous) risk measure, then $\alpha(\mathbf{Q})$ only takes the values 0 and $+\infty$.
By definition, the set $\mathcal{A}_\rho$ is "too large" in the following sense: even if we can write $m + X \in \mathcal{A}_\rho$ as $m + X = \xi \in \mathcal{A}_\rho$, we cannot have an explicit formulation for ξ and in particular cannot compare $m + X$ with 0. Therefore, it seems natural to consider a (convex) class of variables $\mathcal{H}$ such that $m + X \geq H \in \mathcal{H}$. $\mathcal{H}$ appears as a natural (convex) set from which a risk measure can be generated.

1.4.2 Risk Measures Generated by a Convex Set

Risk Measures Generated by a Convex in $\mathcal{X}$ In this section, we study the generation of a convex risk measure from a general convex set.

Definition 1.5 *Given a non-empty convex subset $\mathcal{H}$ of $\mathcal{X}$ such that* $\inf\{m \in \mathbb{R} \mid \exists\xi \in \mathcal{H}, m \geq \xi\} > -\infty$, *the functional $\nu^{\mathcal{H}}$ on $\mathcal{X}$*

$$\nu^{\mathcal{H}}(\Psi) = \inf\big\{m \in \mathbb{R};\ \exists\xi \in \mathcal{H}, m + \Psi \geq \xi\big\} \tag{11}$$

is a convex risk measure. Its minimal penalty function $\alpha^{\mathcal{H}}$ is given by: $\alpha^{\mathcal{H}}(\mathbf{Q}) = \sup_{H\in\mathcal{H}} \mathbb{E}_{\mathbf{Q}}[-H]$.

The main properties of this risk measure are listed or proved below:

1. The acceptance set of $\nu^{\mathcal{H}}$ contains the convex subsets $\mathcal{H}$ and $\mathcal{A}_{\mathcal{H}} = \{\Psi \in \mathcal{X}, \exists \xi \in \mathcal{H},\ \Psi \geq \xi\}$. Moreover, $\mathcal{A}_{\nu^{\mathcal{H}}} = \mathcal{A}_{\mathcal{H}}$ if the last subset is closed in the following sense: For $\xi \in \mathcal{A}_{\mathcal{H}}$ and $\Psi \in \mathcal{X}$, the set $\{\lambda \in [0,1] > \mid \lambda\xi + (1-\lambda)\Psi \in \mathcal{A}_{\mathcal{H}}\}$ is closed in $[0,1]$ (see Proposition 4.6 in [54]).

2. The penalty function $\alpha^{\mathcal{H}}$ associated with $\nu^{\mathcal{H}}$ is the support function of $-\mathcal{A}_{\nu^{\mathcal{H}}}$ defined by $\alpha^{\mathcal{H}}(\mathbf{Q}) = \Sigma^{\mathcal{A}_{\nu^{\mathcal{H}}}}(\mathbf{Q}) = \sup_{X \in \mathcal{A}_{\nu^{\mathcal{H}}}} \mathbb{E}_{\mathbf{Q}}[-X]$. Let us show that $\alpha^{\mathcal{H}}$ is also nothing else but $\Sigma^{\mathcal{H}}$:
   For any $X \in \mathcal{A}_{\nu^{\mathcal{H}}}$ there exist $\epsilon > 0$ and $\xi \in \mathcal{H}$ such that $-X \leq -\xi + \epsilon$. Taking the "expectation" with respect to the additive measure $\mathbf{Q} \in \mathbf{M}_{1,f}$, it follows that $\mathbb{E}_{\mathbf{Q}}[-X] \leq \mathbb{E}_{\mathbf{Q}}[-\xi] + \varepsilon \leq \Sigma^{\mathcal{H}}(\mathbf{Q}) + \varepsilon$ where $\Sigma^{\mathcal{H}}(\mathbf{Q}) = \sup_{H \in \mathcal{H}} \mathbb{E}_{\mathbf{Q}}[-H]$. Taking the supremum with respect to $X \in \mathcal{A}_{\nu^{\mathcal{H}}}$ on the left hand side, we deduce that $\Sigma^{\mathcal{A}_{\nu^{\mathcal{H}}}} \leq \Sigma^{\mathcal{H}}$; the desired result follows from the observation that $\mathcal{H}$ is included in $\mathcal{A}_{\nu^{\mathcal{H}}}$.

3. When $\mathcal{H}$ is a cone, the corresponding risk measure is coherent (homogeneous). The penalty function $\alpha^{\mathcal{H}}$ is the indicator function (in the sense of the convex analysis) of the orthogonal cone $\mathbf{M}_{\mathcal{H}}$: $l^{\mathbf{M}_{\mathcal{H}}}(\mathbf{Q}) = 0$ if $\mathbf{Q} \in \mathcal{M}_{\mathcal{H}}$, $+\infty$ otherwise, where
$$\mathbf{M}_{\mathcal{H}} = \{\mathbf{Q} \in \mathbf{M}_{1,f}; \forall \xi \in \mathcal{H},\ \mathbb{E}_{\mathbf{Q}}[-\xi] \leq 0\}.$$
   The dual formulation of $\nu^{\mathcal{H}}$ is simply given for $\Psi \in \mathcal{X}$ by: $\nu^{\mathcal{H}}(\Psi) = \sup_{\mathbf{Q} \in \mathbf{M}_{\mathcal{H}}} \mathbb{E}_{\mathbf{Q}}[-\Psi]$.

It is natural to associate the convex indicator $l^{\mathcal{H}}$ on $\mathcal{X}$ with the set $\mathcal{H}$, $l^{\mathcal{H}}(X) = 0$ if $X \in \mathcal{H}$; $+\infty$ otherwise. This convex functional is not translation invariant, and therefore it is not a convex risk measure. Nevertheless, $l^{\mathcal{H}}$ and $\nu^{\mathcal{H}}$ are closely related as follows:

Corollary 1.6 *Let $l^{\mathcal{H}}$ be the convex indicator on $\mathcal{X}$ of the convex set $\mathcal{H}$.*
The risk measure $\nu^{\mathcal{H}}$, defined in Equation (11), is the largest convex risk measure dominated by $l^{\mathcal{H}}$ and it can be expressed as:
$$\nu^{\mathcal{H}}(\Psi) = \inf_{\xi \in \mathcal{X}} \{\rho_{\text{worst}}(\Psi - \xi) + l^{\mathcal{H}}(\xi)\}$$
where $\rho_{\text{worst}}(\Psi) = \sup_{\omega \in \Omega}\{-\Psi(\omega)\}$ is the worst case risk measure.

Proof: Let $\mathcal{L} = \{m \in \mathbb{R},\ \exists \xi \in \mathcal{H},\ m \geq \xi\}$. This set is a half-line with lower bound $\inf_{\xi \in \mathcal{H}} \sup_\omega \xi(\omega)$. Moreover, for any $m_0 \notin \mathcal{L}$, $m_0 \leq \inf_{\xi \in \mathcal{H}} \sup_\omega \xi(\omega)$. Therefore, $\nu^{\mathcal{H}}(0) = \inf_{\xi \in \mathcal{H}} \sup_\omega \xi(\omega) = \inf_\xi \rho_{\text{worst}}(-\xi)$. The same arguments hold for $\nu^{\mathcal{H}}(\Psi)$. $\square$

Therefore, $\nu^{\mathcal{H}}$ may be interpreted as the worst case risk measure ρ_{worst} reduced by the use of (hedging) variables in $\mathcal{H}$. This point of view would be generalized in Corollary 3.6 in terms of the inf-convolution $\nu^{\mathcal{H}} = \rho_{\text{worst}} \square l^{\mathcal{H}}$.

Risk Measures Generated by a Convex Set in $\mathbb{L}_\infty(\mathbb{P})$ Assume now $\mathcal{H}$ to be a convex subset of $\mathbb{L}_\infty(\mathbb{P})$. The functional $\nu^{\mathcal{H}}$ on $\mathbb{L}_\infty(\mathbb{P})$ is still defined by the same formula (11), in which the inequality has to be understood in $\mathbb{L}_\infty(\mathbb{P})$, i.e. $\mathbb{P} - a.s.$, with a penalty function only defined on $\mathbf{M}_{1,ac}(\mathbb{P})$ and given by $\alpha^{\mathcal{H}}(\mathbf{Q}) = \sup_{H \in \mathcal{H}} \mathbb{E}_{\mathbf{Q}}[-H]$.

The problem is then to give condition(s) on the set $\mathcal{H}$ to ensure that the dual representation holds on $\mathcal{M}_{1,ac}(\mathbb{P})$ and not only on $\mathbf{M}_{1,ac}(\mathbb{P})$. By Theorem 1.4, this problem is equivalent to the continuity from above of the risk measure $\nu^{\mathcal{H}}$ or equivalently to the weak*-closure of its acceptance set $\mathcal{A}_{\mathcal{H}}$. Properties of this kind are difficult to check and in the following, we will simply give some examples where this property holds.

1.5 Static Hedging and Calibration

In this subsection, we consider some examples motivated by financial risk hedging problems.

1.5.1 Hedging with a Family of Cash Flows

We start with a very simple model where it is only possible to hedge statically over a given period using a finite family of *bounded* cash flows $\{C_1, C_2, ..., C_d\}$, the (forward) price of which is known at time 0 and denoted by $\{\pi_1, \pi_2, ..., \pi_d\}$. All cash flows are assumed to be non-negative and non-redundant. Constants may be included and then considered as assets.
We assume that the different prices are **coherent** in the sense that

$$\exists\, \mathbb{Q}_0 \sim \mathbb{P}, s.t. \quad \forall i,\ \mathbb{E}_{\mathbb{Q}_0}[C_i] = \pi_i$$

Such an assumption implies in particular that any inequality on the cash flows is preserved on the prices. The quantities of interest are often the gain values of the basic strategies, $G_i = C_i - \pi_i$.
We can naturally introduce the non-empty set $\mathcal{Q}_e$ of equivalent *"martingale measures"* as

$$\mathcal{Q}_e = \{\ \mathbb{Q}|\ \mathbb{Q} \sim \mathbb{P}, s.t. \quad \forall i,\ \mathbb{E}_{\mathbb{Q}}[G_i] = 0\ \}$$

The different instruments we consider are very liquid; by selling or buying some quantities θ_i of such instruments, we define the family $\mathbf{\Theta}$ of gains associated with trading strategies θ:

$$\mathbf{\Theta} = \Big\{\ G(\theta) = \sum_{i=1}^{d} \theta_i\, G_i,\ \theta \in \mathbb{R}^d, \quad \text{with initial value} \quad \sum_{i=1}^{d} \theta_i\, \pi_i \Big\}$$

This framework is very similar to Chapter 1 in Föllmer and Schied [54] where it is shown that the assumption of coherent prices is equivalent to the absence of arbitrage opportunity in the market defined as

$$\textbf{(AAO)} \qquad G(\theta) \geq 0 \quad \mathbb{P}\ a.s. \quad \Rightarrow G(\theta) = 0 \quad \mathbb{P}\ a.s.$$

These strategies can be used to hedge a risky position Y. In the classical financial literature, a superhedging strategy is a par (m, θ) such that $m + G(\theta) \geq Y$, $a.s.$ This leads to the notion of superhedging (super-seller) price $\pi_{\uparrow}^{\text{sell}}(Y) = \inf\{\ m \mid \exists\, G(\theta)\ s.t.\ m + G(\theta) \geq Y\}$. In terms of risk measure, we are concerned with the static superhedging price of $-Y$. So, by setting $\mathcal{H} = -\mathbf{\Theta}$, we define the risk measure $\nu^{\mathcal{H}}$ as

$$\nu^{\mathcal{H}}(X) = \pi_{\uparrow}^{\text{sell}}(-X) = \inf\{m \in \mathbb{R},\ \exists \theta \in \mathbb{R}^d : m + X + G(\theta) \geq 0\}$$

Let us observe that the no arbitrage assumption implies that $\mathbb{E}_{\mathbb{Q}_0}[G(\theta)] = 0$. Hence, $\nu^{\mathcal{H}}(0) \geq \mathbb{E}_{\mathbb{Q}_0}[-X] > -\infty$. Moreover, the dual representation of the risk measure $\nu^{\mathcal{H}}$ in terms of probability measures is closely related to the absence of arbitrage opportunity as underlined in the following proposition (Chapter 4 in [54]):

Proposition 1.7 *i) If the market is arbitrage-free, i.e.* **(AAO)** *holds true, the convex risk measure* $\nu^{\mathcal{H}}$ *can be represented in terms of the set of equivalent "martingale" measures* $\mathcal{Q}_e$ *as*

$$\nu^{\mathcal{H}}(\Psi) = \sup_{\mathbb{Q}\in\mathcal{Q}_e} \mathbb{E}_{\mathbb{Q}}(-X), \quad \text{where} \quad \mathcal{Q}_e = \{\mathbb{Q} \sim \mathbb{P},\ \mathbb{E}_{\mathbb{Q}}(G_i) = 0,\ \forall i = 1...d\}. \tag{12}$$

By Theorem 1.2, this $\mathbb{L}_\infty(\mathbb{P})$*-risk measure is continuous from above.*
ii) Moreover, the market is arbitrage-free if $\nu^{\mathcal{H}}$ *is* **sensitive** *in the sense that* $\nu^{\mathcal{H}}(\Psi) > \nu^{\mathcal{H}}(0)$ *for all* Ψ *such that* $\mathbb{P}(X < 0) > 0$ *and* $\mathbb{P}(X \leq 0) = 1$.

1.5.2 Calibration Point of View and Bid-Ask Constraint

This point of view is often used on financial markets when cash flows depend on some basic assets $(S_1, S_2, ..., S_n)$, whose characteristics will be given in the next paragraph.
We can consider for instance (C_i) as payoffs of derivative instruments, sufficiently liquid to be used as *calibration tools and static hedging strategies.* So far, all agents having access to the market agree on the derivative prices, and do not have any restriction on the quantity they can buy or sell.
We now take into account some restrictions on the trading. We first introduce a bid-ask spread on the (forward) price of the different cash flows. We denote by $\pi_i^{ask}(C_i)$ the market buying price and by $\pi_i^{bid}(C_i)$ the market selling price. The price coherence is now written as

$$\exists\, \mathbb{Q}_0 \sim \mathbb{P}, \quad \forall i,\ \pi_i^{ask}(C_i) \leq \mathbb{E}_{\mathbb{Q}_0}[C_i] \leq \pi_i^{bid}(C_i)$$

To define the gains family, we need to make a distinction between cash flows when buying and cash flows when selling. To do that, we double the number of basic gains, by associating, with any given cash-flow C_i, both gains $G_i^{bid} = C_i - \pi_i^{bid}$ and $G_i^{ask} = \pi_i^{ask}(C_i) - C_i$. Henceforth, we do not make distinction of the notation and we still denote any gain by G_i. The price coherence is then expressed as

$$\exists\, \mathbb{Q}_0 \sim \mathbb{P}, \quad \forall\, i = 1....\mathbf{2d},\ \mathbb{E}_{\mathbb{Q}_0}[G_i] \leq 0.$$

The set of such probability measures, called *super-martingale measures*, is denoted by $\mathcal{Q}_e^s$. Note that the coherence of the prices implies that the set $\mathcal{Q}_e^s$ is non empty.
Using this convention, a strategy is defined by a $2d$-dimensional vector θ, the components of which are all non-negative. More generally, we can introduce more trading restriction on the size of the transaction by constraining θ to belong to a convex set $\mathcal{K} \subseteq \mathbb{R}_+^{2d}$ such that $0 \in \mathcal{K}$. Note that we can also take into account some limits to the resources of the investor, in such way the initial price $\langle\theta, \pi\rangle$ has an an upper bound. In any case, we still denote the set of admissible strategies by $\mathcal{K}$ and the family of associated gains by: $\mathbf{\Theta} = \Big\{\ G(\theta) = \sum_{i=1}^{2d} \theta_i G_i,\ \theta \in \mathcal{K}\Big\}$.
In this constrained framework, the relationship between price coherence and (**AAO**) on $\mathbf{\Theta}$ has been studied in details in Bion-Nadal [17] but also in Chapter 1 of [54].
More precisely, as above, the price coherence implies that the risk measure $\nu^{\mathcal{H}}$ related to $\mathcal{H} = -\mathbf{\Theta}$ is not identically $-\infty$.
A natural question is to extend the duality relationship (12) using the subset of *super-martingale measures.*

Using Paragraph 1.4.2, this question is equivalent to show that the minimal penalty function is infinite outside of the set of absolutely continuous probability measures and that $\nu^{\mathcal{H}}$ is continuous from above.
When studying the risk measure $\nu^{\mathcal{H}}$ (Definition 1.5 and its properties), we have proved that:

$$\forall \mathbb{Q} \in \mathcal{M}_{1,ac}(\mathbb{P}), \quad \alpha^{\mathcal{H}}(\mathbb{Q}) = \sup_{\xi \in \mathcal{H}} \mathbb{E}_{\mathbb{Q}}[-\xi] = \sup_{\theta \in \mathcal{K}} \mathbb{E}_{\mathbb{Q}}[G(\theta)].$$

In particular, since $0 \in \mathcal{K}$, if $\mathbb{Q} \in \mathcal{Q}_e^s$, then $\alpha(\mathbb{Q}) = 0$. Moreover, if Θ is a cone, then $\alpha^{\mathcal{H}}$ is the indicator function of $\mathcal{Q}_e^s$.
It remains to study the continuity from above of $\nu^{\mathcal{H}}$ and especially to relate it with the absence of arbitrage opportunity in the market. We summarize below the results Föllmer and Schied obtained in Theorem 4.95 and Corollary 9.30 [54].

Proposition 1.8 *Let the set $\mathcal{K}$ be a closed subset of $\mathbb{R}^d$. Then, the market is arbitrage-free if and only if the risk measure $\nu^{\mathcal{H}}$ is sensitive. In this case, $\nu^{\mathcal{H}}$ is continuous from above and admits the dual representation:*

$$\nu^{\mathcal{H}}(\Psi) = \sup_{\mathbb{Q} \in \mathcal{M}_{1,ac}} \left\{ \mathbb{E}_{\mathbb{Q}}[-\Psi] - \alpha^{\mathcal{H}}(\mathbb{Q}) \right\}.$$

1.5.3 Dynamic Hedging

A natural extension of the previous framework is the multi-period setting or more generally the continuous-time setting. We briefly present some results in the latter case. Note that we will come back to these questions, in the second part of this chapter, under a slightly different form, assuming that basic asset prices are Itô's processes.
We now consider a time horizon T, a filtration $(\mathcal{F}_t; t \in [0,T])$ on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and a financial market with n basic assets, whose (non-negative) vector price process S follows a special locally bounded semi-martingale under $\mathbb{P}$. To avoid arbitrage, we assume that:

(AAO) *There exists a probability measure* $\mathbb{Q}_0 \sim \mathbb{P}$ *such that* S is a $\mathbb{Q}_0 -$ *local-martingale.*

Let $\mathcal{Q}_{\text{ac}}$ be the family of absolutely continuous martingale measures: $\mathcal{Q}_{\text{ac}} = \{ \mathbb{Q} \mid \mathbb{Q} \ll \mathbb{P}, \ S \text{ is a } \mathbb{Q} \text{ local-martingale}\}$. **(AAO)** ensures that the set $\mathcal{Q}_{\text{ac}}$ is non empty. Then, as in Delbaen [38], $\mathcal{Q}_{ac}$ is a closed convex subset of $\mathbb{L}^1(\mathbb{P})$.
Let us now introduce dynamic strategies as predictable processes θ and their gain processes $G_t(\theta) = \int_0^t \langle \theta_u, dS_u \rangle = (\theta.S)_t$. We only consider bounded gain processes and define:

$$\Theta_T^S = \{G_T(\theta) = (\theta.S)_T \mid \theta.S \text{ is bounded}\}$$

Delbaen and Schachermayer have established in [40], as in the static case, the following duality relationship,

$$\sup\{ \mathbb{E}_{\mathbb{Q}}[-X] \mid \mathbb{Q} \in \mathcal{Q}_{\text{ac}}\} = \inf\{ m \mid \exists \ G_T(\theta) \in \Theta_T^S \ \ s.t. \quad m + X + G_T(\theta) \geq 0\}$$

Putting $\mathcal{H} = -\Theta_T^S$, this equality shows that $\nu^{\mathcal{H}}$ is a *coherent convex risk measure continuous from above.*

Constrained portfolios When constraints are introduced on the strategies, everything becomes more complex. Therefore, we refer to the course held by Schied [109] for more details.
We assume that hedging positions live in the following convex set:

$$\Theta_T^S = \{G_T(\theta) = (\theta.S)_T \mid \theta.S \text{ is bounded by below}, \theta \in \mathcal{K}\}$$

The set of constraints is closed in the following sense: the set $\{\int \theta dS \,|\, \theta \in \mathcal{K}\}$ is closed in the semi-martingale or Émery topology. The optional decomposition theorem of Föllmer and Kramkov [52] implies the following dual representation for the risk measure $\nu^{\mathcal{H}}$:

$$\nu^{\mathcal{H}}(\Psi) = \sup_{\mathbb{Q}\in\mathcal{M}_{1,ac}} \left\{\mathbb{E}_{\mathbb{Q}}[-\Psi] - \mathbb{E}_{\mathbb{Q}}[A_T^{\mathbb{Q}}]\right\}$$

where $A_.^{\mathbb{Q}}$ is the optional process defined by $A_0^{\mathbb{Q}} = 0$ and $dA_t^{\mathbb{Q}} = \operatorname{ess\,sup}_{\xi\in\mathcal{K}} \mathbb{E}_{\mathbb{Q}}[\theta_t dS_t|\mathcal{F}_t]$.
The penalty function $\alpha^{\mathcal{H}}$ of the risk measure $\nu^{\mathcal{H}}$ can be described as $\mathbb{E}_{\mathbb{Q}}[A_T^{\mathbb{Q}}]$ provided that $\mathbb{Q}$ satisfies the three following conditions:

- $\mathbb{Q}$ is equivalent to $\mathbb{P}$;
- Every process $\theta.S$ with $\theta \in \mathcal{K}$ is a special semi-martingale under $\mathbb{Q}$;
- $\mathbb{Q}$ admits the upper variation process $A^{\mathbb{Q}}$ for the set $\{\theta.S \mid \theta \in \mathcal{K}\}$.

We can set $\alpha^{\mathcal{H}}(\mathbb{Q}) = +\infty$ when one of these conditions does not hold.

Remark 1.9 *Note that there is a fundamental difference between static hedging with a family of cash flows and dynamic hedging. In the first case, the initial wealth is a market data: it corresponds to the (forward) price of the considered cash flows. The underlying logic is based upon calibration as the probability measures we consider have to be consistent with the observed market prices of the hedging instruments. In the dynamic framework, the initial wealth is a given data. The agent invests it in a self-financing admissible portfolio which may be rebalanced in continuous time.*
The problem of dynamic hedging with calibration constraints is a classical problem for practitioners. This will be addressed in details after the introduction of the inf-convolution operator. Some authors have been looking at this question (see for instance Bion-Nadal [16] or Cont [33]).

2 Dilatation of Convex Risk Measures, Subdifferential and Conservative Price

2.1 Dilatation: γ-Tolerant Risk Measures

For non-coherent convex risk measures, the impact of the size of the position is not linear. It seems therefore natural to consider the relationship between "risk tolerance" and the perception of the size of the position. To do so, we start from a given root convex risk measure ρ. The risk tolerance coefficient is introduced as a parameter describing how agents penalize compared with this root risk measure. More precisely, denoting by γ the risk tolerance, we define ρ_γ as:

$$\rho_\gamma(\Psi) = \gamma\rho(\frac{1}{\gamma}\Psi). \tag{13}$$

ρ_γ satisfies a *tolerance property* or a *dilatation property* with respect to the size of the position, therefore it is called the γ-*tolerant risk measure* associated with ρ (also called the *dilated risk measure* associated with ρ as in Barrieu and El Karoui [10]). A typical example is the entropic risk measure where e_γ is simply the γ-dilated of e_1. These dilated risk measures satisfy the following nice property:

Proposition 2.1 *Let $\big(\rho_\gamma, \gamma > 0\big)$ be the family of γ-tolerant risk measures issued of ρ. Then,*
(i) The map $\gamma \to (\rho_\gamma - \gamma\rho(0))$ is non-increasing,
(ii) For any $\gamma, \gamma' > 0$, $(\rho_\gamma)_{\gamma'} = \rho_{\gamma\,\gamma'}$.
(iii) The perspective functional defined on $]0,\infty[\times\mathcal{X}$ by

$$p_\rho(\gamma, X) = \gamma\rho(\frac{X}{\gamma}) = \rho_\gamma(X)$$

is a homogeneous convex functional, cash-invariant with respect to X (i.e. a coherent risk measure in X).

Proof: (i) We can take $\rho(0) = 0$ without loss of generality of the arguments. By applying the convexity inequality to $\frac{X}{\gamma}$ and 0 with the coefficients $\frac{\gamma}{\gamma+h}$ and $\frac{h}{\gamma+h}$ $(h > 0)$, we have, since $\rho(0) = 0$:

$$\rho(\frac{X}{\gamma+h}) \le \frac{\gamma}{\gamma+h}\rho(\frac{X}{\gamma}) + \frac{h}{\gamma+h}\rho(0) \le \frac{\gamma}{\gamma+h}\rho(\frac{X}{\gamma}).$$

(ii) is an immediate consequence of the definition and characterization of tolerant risk measures.
(iii) The perspective functional is clearly homogeneous. To show the convexity, let $\beta_1 \in [0,1]$ and $\beta_2 = 1-\beta_1$ two real coefficients, and (γ_1, X_1), (γ_2, X_2) two points in the definition space of p_ρ. Then, by the convexity of ρ,

$$\begin{aligned} p_\rho\big(\beta_1(\gamma_1, X_1) + \beta_2(\gamma_2, X_2)\big) &= (\beta_1\gamma_1 + \beta_2\gamma_2)\,\rho\Big(\frac{\beta_1 X_1 + \beta_2 X_2}{\beta_1\gamma_1 + \beta_2\gamma_2}\Big) \\ &\le (\beta_1\gamma_1 + \beta_2\gamma_2)\,\Big[\frac{\beta_1\gamma_1}{\beta_1\gamma_1 + \beta_2\gamma_2}\rho(\frac{X_1}{\gamma_1}) + \frac{\beta_2\gamma_2}{\beta_1\gamma_1 + \beta_2\gamma_2}\rho(\frac{X_2}{\gamma_2})\Big] \\ &\le \beta_1\,\rho_{\gamma_1}(X_1) \;+\; \beta_2\,\rho_{\gamma_2}(X_2). \end{aligned}$$

The other properties are obvious. $\Box$

So, we naturally are looking for the asymptotic behavior of the perspective risk measure when the risk tolerance either tends to $+\infty$ or tends to 0.

2.2 Marginal Risk Measures and Subdifferential

2.2.1 Marginal Risk Measure

Let us first observe that ρ is a *coherent* risk measure if and only if $\rho_\gamma \equiv \rho$. We then consider the behavior of the family of γ-tolerant risk measures when the tolerance becomes infinite.

Proposition 2.2 *Suppose that $\rho(0)=0$, or equivalently $\alpha(\mathbf{Q})\geq 0 \quad \forall \mathbf{Q}\in \mathbf{M}_{1,f}$.*
a) The marginal risk measure ρ_∞, defined as the non-increasing limit of ρ_γ when γ tends to infinity, is a coherent risk measure with penalty function $\alpha_\infty = \lim_{\gamma\to+\infty}(\gamma\alpha)$ that is:

$$\begin{array}{ll}\alpha_\infty(\mathbf{Q}) & := \sup_\Psi \big\{\mathbb{E}_{\mathbf{Q}}[-\Psi]-\rho_\infty(\Psi)\big\} = \ 0 \ \text{ if } \alpha(\mathbf{Q})=0 \ , \quad +\infty \text{ if not,}\\ \rho_\infty(\Psi) & = \sup_{\mathbf{Q}\in\mathbf{M}_{1,f}} \big\{\mathbb{E}_{\mathbf{Q}}[-\Psi] \ \big| \ \alpha\big(\mathbf{Q}\big)=0\big\}.\end{array}$$

b) Assume now that ρ is a $\mathbb{L}_\infty(\mathbb{P})$-risk measure such that $\rho(0)=0$.
If ρ is continuous from below, the ρ_∞ is continuous from below and admits a representation in terms of absolutely continuous probability measures as:

$$\rho_\infty(\Psi) = \max_{\mathbb{Q}\in\mathcal{M}_{1,ac}}\big\{\mathbb{E}_{\mathbb{Q}}[-\Psi] \ \big| \ \alpha\big(\mathbb{Q}\big)=0\big\},$$

and the set $\big\{\mathbb{Q}\in\mathcal{M}_{1,ac} \ \big| \ \alpha\big(\mathbb{Q}\big)=0\big\}$ is non empty, and weakly compact in $\mathbb{L}^1(\mathbb{P})$.

Proof: *a)* Thanks to Theorem 2.1, for any $\Psi\in\mathcal{X}$ $\rho_\gamma(\Psi)\searrow\rho_\infty(\Psi)$ when $\gamma\to+\infty$. Given the fact that $-m\geq\rho_\gamma(\Psi)\geq -M$ when $m\leq\Psi\leq M$, we also have $-m\geq\rho_\infty(\Psi)\geq -M$ and ρ_∞ is finite.
Convexity, monotonicity and cash translation invariance properties are preserved when taking the limit. Therefore, ρ_∞ is a convex risk measure with $\rho_\infty(0)=0$.
Moreover, given that $(\rho_\delta)_\gamma=\rho_{\delta\gamma}=(\rho_\gamma)_\delta$, we have that $(\rho_\delta)_\infty=\rho_\infty=(\rho_\infty)_\delta$ and ρ_∞ is a coherent risk measure.

Since $\alpha\geq 0$, the minimal penalty function is:

$$\begin{array}{ll}\alpha_\infty(\mathbf{Q}) & = \sup_\xi \big\{\mathbb{E}_{\mathbf{Q}}[-\xi]-\rho_\infty(\xi)\big\}\\ & = \sup_\xi \sup_{\gamma>0}\big\{\mathbb{E}_{\mathbf{Q}}[-\xi]-\gamma\rho(\frac{\xi}{\gamma})\big\}\\ & = \sup_{\gamma>0}\big\{\gamma\alpha(\mathbf{Q})\big\} = \ 0 \ \text{ if } \alpha(\mathbf{Q})=0 \ , \quad +\infty \text{ if not.}\end{array}$$

Moreover, α_∞ is not identically equal to $+\infty$ since the set $\big\{\mathbf{Q}\in\mathbf{M}_{1,f} \ \big| \ \alpha\big(\mathbf{Q}\big)=0\big\}$ is not empty given that $\rho(0)=0=\max\big\{-\alpha(\mathbf{Q})\big\}=-\alpha(\mathbf{Q}_0)$ for some additive measure $\mathbf{Q}_0\in\mathbf{M}_{1,f}$, from Theorem 1.2.
Assume now that ρ is continuous from below and consider a non-decreasing sequence $(\xi_n\in\mathcal{X})$ with limit $\xi\in\mathcal{X}$. By monotonicity,

$$\rho_\infty(\xi)=\inf_\gamma\rho_\gamma(\xi)=\inf_\gamma\inf_{\xi_n}\rho_\gamma(\xi_n)=\inf_{\xi_n}\inf_\gamma\rho_\gamma(\xi_n)=\inf_{\xi_n}\rho_\infty(\xi_n).$$

Then, ρ_∞ is also continuous from below.

b) When ρ is a $\mathbb{L}_\infty(\mathbb{P})$-risk measure, continuous from below, ρ is also continuous from above and the dual representation holds in terms of absolutely continuous probability measures. Using the same argument as above, we can prove that ρ_∞ is a coherent $\mathbb{L}_\infty(\mathbb{P})$-risk measure, continuous from below with minimal penalty function:

$$\begin{array}{ll}\alpha_\infty(\mathbb{Q}) & = 0 \ \text{ if } \ \alpha(\mathbb{Q})=0 \quad \text{and} \quad \mathbb{Q}\in\mathcal{M}_{1,ac}\\ & = +\infty \ \text{ otherwise.}\end{array}$$

Moreover, thanks to Theorem 1.4, the set $\big\{\mathbb{Q}\in\mathcal{M}_{1,ac} \ \big| \ \alpha\big(\mathbb{Q}\big)=0\big\}$ is non empty and weakly compact in $\mathbb{L}^1(\mathbb{P})$. □

To have some intuition about the interpretation in terms of marginal risk measure, it is better to refer to the risk aversion coefficient $\epsilon = 1/\gamma$. $\rho_\infty(\Psi)$ appears as the limit of $\frac{1}{\epsilon}\big(\rho(\epsilon\Psi)-\rho(0)\big)$, i.e. the right-derivative at 0 in the direction of Ψ of the risk measure ρ, or equivalently, the marginal risk measure. For instance $e_\infty(\Psi) = \mathbb{E}_{\mathbb{P}}(-\Psi)$.

In some cases, and in particular when the set $\mathcal{Q}_\infty^\alpha$ has a single element, the pricing rule $\rho_\infty(-\Psi)$ is a linear pricing rule and can be seen as an extension of the notion of marginal utility pricing and of the Davis price (see Davis [37] or Karatzas and Kou [76]).

2.2.2 Subdifferential and its Support Function

Subdifferential Let us first recall the definition of the subdifferential of a convex functional.

Definition 2.3 *Let ϕ be a convex functional on $\mathcal{X}$. The subdifferential of ϕ at X is the set*

$$\partial\phi(X) = \big\{\mathbf{q}\in\mathcal{X}' \mid \forall X\in\mathcal{X},\ \phi(X+Y)\geq\phi(X)+\mathbf{q}(-Y)\big\}$$

The subdifferential of a convex risk measure ρ with penalty function $\alpha(q)=\sup_Y\{q(-Y)-\rho(Y)\}$ is included in $\mathbf{Dom}(\alpha)$ since when $q\in\partial\rho(\xi)$, then $\alpha(q)-(q(-\xi)-\rho(\xi))\leq 0$. So, we always refer to finitely additive measure $\mathbf{Q}$ when working with risk measure subdifferential. In fact, we have the well-known characterization of the subdifferential:

$\mathbf{q}\in\partial\rho(\xi)$ *if and only if* $\mathbf{q}\in\mathbf{M}_{1,f}$ *is optimal for the maximization program* $\mathbb{E}_{\mathbf{Q}}[-\xi]-\alpha(\mathbf{Q}) \longrightarrow \max_{\mathbf{Q}\in\mathbf{M}_{1,f}}$.

We can also relate it with the notion of marginal risk measure, when the root risk measure is now centered around a given element $\xi\in\mathcal{X}$, i.e. $\rho_\xi(X)=\rho(X+\xi)-\rho(\xi)$, by defining:

$$\rho_{\infty,\xi}(\Psi)\equiv\lim_{\gamma\to+\infty}\gamma\Big(\rho(\xi+\frac{\Psi}{\gamma})-\rho(\xi)\Big).$$

Using Proposition 2.2, since the ρ_ξ penalty function is $\alpha_\xi(\mathbf{Q})\equiv\alpha(\mathbf{Q})-\mathbb{E}_{\mathbf{Q}}[-\xi]+\rho(\xi)$, $\rho_{\infty,\xi}$ is coherent and

$$\rho_{\infty,\xi}(\Psi)=\sup_{\mathbf{Q}\in\mathbf{M}_{1,f}}\big\{\mathbb{E}_{\mathbf{Q}}[-\Psi]\,\big|\,\rho(\xi)=\mathbb{E}_{\mathbf{Q}}[-\xi]-\alpha(\mathbf{Q})\big\}.$$

Proposition 2.4 *The coherent risk measure $\rho_{\infty,\xi}(\Psi)\equiv\lim_{\gamma\to+\infty}\gamma\Big(\rho(\xi+\frac{\Psi}{\gamma})-\rho(\xi)\Big)$ is the support function of the subdifferential $\partial\rho(\xi)$ of the convex risk measure ρ at ξ:*

$$\rho_{\infty,\xi}(\Psi)=\sup_{\mathbf{Q}\in\mathbf{M}_{1,f}}\big\{\mathbb{E}_{\mathbf{Q}}[-\Psi]\,\big|\,\rho(\xi)=\mathbb{E}_{\mathbf{Q}}[-\xi]-\alpha(\mathbf{Q})\big\}=\sup_{\mathbf{Q}\in\partial\rho(\xi)}\mathbb{E}_{\mathbf{Q}}[-\Psi]$$

Proof: From the definition of the subdifferential,

$$\begin{aligned}\partial\rho(\xi)&=\big\{\mathbf{q}\in\mathcal{X}'\mid\forall\Psi\in\mathcal{X},\ \rho(\xi+\Psi)\geq\rho(\xi)+\mathbf{q}(-\Psi)\big\}\\&=\big\{\mathbf{q}\in\mathcal{X}'\mid\forall\Psi\in\mathcal{X},\ \rho_{\infty,\xi}(\Psi)\geq\mathbf{q}(-\Psi)\big\}\\&=\partial\rho_{\infty,\xi}(0).\end{aligned}$$

But $q\in\partial\rho_{\infty,\xi}(0)$ iff $\alpha_{\infty,\xi}(\mathbf{Q_q})=0$. So the proof is complete. $\square$

The $\mathbb{L}_\infty(\mathbb{P})$ case: When working with $\mathbb{L}_\infty(\mathbb{P})$-risk measures, following Delbaen [38] (Section 8), the natural definition of the subdifferential is the following:

$$\partial\rho(\xi) = \big\{f \in \mathbb{L}^1(\mathbb{P}) \mid \forall \Psi \in \mathbb{L}_\infty(\mathbb{P}),\ \ \rho(\xi+\Psi) \geq \rho(\xi) + \mathbb{E}_{\mathbb{P}}[f(-\Psi)]\big\}$$

Using the same arguments as above, we can prove that every $f \in \partial\rho(\xi)$ is non-negative with a $\mathbb{P}$-expectation equal to 1. Since $\partial\rho(\xi)$ is also the subdifferential of $\rho_{\infty,\xi}(0)$, the properties of $\partial\rho(\xi)$ may be deduced from those of the coherent risk measure $\rho_{\infty,\xi}$, for which we have already shown that if ρ is continuous from below and $\rho(0) = 0$ then for any ξ, the effective domain of $\alpha_{\infty,\xi}$ is non empty. Then, under this assumption, $\partial\rho(\xi)$ is non empty and we have the same characterization of the subdifferential as:

$$\mathbb{Q} \in \partial\rho(\xi) \Longleftrightarrow \rho(\xi) = \mathbb{E}_{\mathbb{Q}}[-\xi] - \alpha(\mathbb{Q}).$$

We now summarize these results in the following proposition:

Proposition 2.5 *Let ρ be a $\mathbb{L}_\infty(\mathbb{P})$-risk measure, continuous from below. Then, for any $\xi \in \mathbb{L}_\infty(\mathbb{P})$, $\rho_{\infty,\xi}$ is the support function of the non empty subdifferential $\partial\rho(\xi)$, i.e.:*

$$\rho_{\infty,\xi}(\Psi) = \sup\big\{\mathbb{E}_{\mathbb{Q}}[-\Psi]\ ;\ \mathbb{Q} \in \partial\rho(\xi)\big\}.$$

and the supremum is attained by some $\mathbb{Q} \in \partial\rho(\xi)$.

2.3 Conservative Risk Measures and Super-Price

We now focus on the properties of the γ-tolerant risk measures when the risk tolerance coefficient tends to 0 or equivalently when the risk aversion coefficient goes to $+\infty$. The conservative risk measures that are then obtained can be reinterpreted in terms of super-pricing rules. Using vocabulary from convex analysis, these risk measures are related to recession (or asymptotic) functions.

Proposition 2.6 (a) *When γ tends to 0, the family of γ-tolerant risk measures (ρ_γ) admits a limit ρ_{0^+}, which is a coherent risk measure. This conservative risk measure ρ_{0^+} is simply the "super-price" of $-\Psi$:*

$$\rho_{0^+}(\Psi) = \lim_{\gamma\downarrow 0} \nearrow (\rho_\gamma(\Psi) - \gamma\rho(0)) = \sup_{\mathbf{Q}\in\mathbf{M}_{1,f}} \big\{\mathbb{E}_{\mathbf{Q}}[-\Psi] \big| \,\alpha\big(\mathbf{Q}\big) < \infty\big\}.$$

Its minimal penalty function is

$$\alpha_{0^+}(\mathbf{Q}) = 0\ \ if\ \ \alpha(\mathbf{Q}) < +\infty \quad \text{and} \quad = +\infty\ \text{ if not.}$$

(b) *If ρ is continuous from above on $\mathbb{L}^\infty(\mathbb{P})$, then ρ_{0^+} is continuous from above and*

$$\rho_{0^+}(\Psi) = \sup_{\mathbb{Q}\in\mathcal{M}_{1,ac}} \big\{\mathbb{E}_{\mathbb{Q}}[-\Psi] \big| \,\alpha\big(\mathbb{Q}\big) < \infty\big\}.$$

Proof: Let us first observe that $\rho_\gamma(\xi) = \gamma\big(\rho(\frac{\xi}{\gamma}) - \rho(0)\big) + \gamma\rho(0)$ is the sum of two terms. The first term is monotonic while the second one goes to 0.
The functional ρ_{0^+} is coherent (same proof as for ρ_∞) with the acceptance set $\mathcal{A}_{\rho_{0^+}} = \{\xi,\ \forall\lambda \geq 0,\ \ \lambda\xi \in$

$\mathcal{A}_\rho - \rho(0)\}$.
On the other hand, by monotonicity, the minimal penalty function $\alpha_{0^+} \geq \gamma\alpha \geq 0$; so, $\alpha_{0^+}(\mathbf{Q}) = 0$ on $\mathbf{Dom}(\alpha)$, and $\alpha_{0^+}(\mathbf{Q}) = +\infty$ if not. In other words, α_{0^+} is the convex indicator of $\mathbf{Dom}(\alpha)$.
If ρ is continuous from above on $\mathbb{L}^\infty(\mathbb{P})$, then the same type of dual characterization holds for ρ_{0^+} but in terms of $\mathcal{M}_{1,ac}$. So, $\alpha_{0^+}(\mathbb{Q}) = 0$ on $Dom(\alpha)$, and $\alpha_{0^+}(\mathbb{Q}) = +\infty$ if not.

We could have proved directly the continuity from above of ρ_{0^+}, since ρ_{0^+} is the non-decreasing limit of continuous from above risk measures $(\rho_\gamma - \gamma\rho(0))$. $\square$

Remark 2.7 *A nice illustration of this result can be obtained when considering the entropic risk measure e_γ. In this case, it comes immediately that $e_{0^+}(\Psi) = \sup_{\mathbb{Q}} \left\{\mathbb{E}_{\mathbb{Q}}[-\Psi] \middle| h(\mathbb{Q}\,|\,\mathbb{P}) < +\infty\right\} = \mathbb{P} - ess\sup(-\Psi) = \rho_{\max}(\Psi)$ where $\rho_{\max}$ is here the $\mathbb{L}_\infty(\mathbb{P})$-worst case measure. This also corresponds to the weak super-replication price as defined by Biagini and Frittelli in [15].*
Note that this conservative risk measure $e_{0^+}(\Psi)$ cannot be realized as $\mathbb{E}_{\mathbb{Q}_0}[-\Psi]$ for some $\mathbb{Q}_0 \in \mathcal{M}_{1,ac}$. It is a typical example where the continuity from below fails.

3 Inf-Convolution

A useful tool in convex analysis is the inf-convolution operation. While the classical convolution acts on the Fourier transforms by addition, the inf-convolution acts on Fenchel transforms by addition as we would see later.

3.1 Definition and Main Properties

The inf-convolution of two convex functionals ϕ_A and ϕ_B may be viewed as the functional value of the minimization program

$$\phi_{A,B}(X) = \inf_{H \in \mathcal{X}} \big\{\phi_A(X - H) + \phi_B(H)\big\}, \tag{14}$$

This program is the functional extension of the classical inf-convolution operator acting on real convex functions $f \Box g(x) = \inf_y\{f(x-y) + g(y)\}$.

Illustrative example: Let us assume that the risk measure ρ_A is the linear one $q_A(X) = \mathbb{E}_{\mathbb{Q}_A}[-X]$, whose the penalty function is the functional $\alpha_A(\mathbb{Q}) = 0$ if $\mathbb{Q} = \mathbb{Q}_A$, $= +\infty$ if not. Given a convex risk measure, ρ_B, with penalty functional α_B, we deduce from the definition of the inf-convolution that

$$q_A \Box \rho_B(X) = q_A(-X) - \alpha_B(\mathbb{Q}_A)$$

⋄ Then, $q_A \Box \rho_B$ is identically $-\infty$ if $\alpha_B(\mathbb{Q}_A) = +\infty$.
⋄ If it is not the case, the minimal penalty function $\alpha_{A,B}$ associated with this measure is:

$$\alpha_{A,B}(\mathbb{Q}) = \alpha_B(\mathbb{Q}_A) + \alpha_A(\mathbb{Q}) = \alpha_B(\mathbb{Q}) + \alpha_A(\mathbb{Q})$$

$\diamond$ Moreover, the infimum is attained in the inf-convolution program by any H^* such that

$$\alpha_B(\mathbb{Q}_A) = \mathbb{E}_{\mathbb{Q}_A}[-H^*] - \rho_B(H^*)$$

that is H^* is optimal for the maximization program defining the α_B.
In terms of subdifferential, we have the first order condition: $\quad \mathbb{Q}_A \in \partial\rho_B(H^*)$.

3.1.1 Inf-Convolution and Duality

In our setting, convex functionals are generally convex risk measures, but we have also been concerned by the convex indicator of convex subset, taking infinite values. In that follows, we already assume that convex functionals ϕ we consider are proper (i.e. not identically $+\infty$) and in general closed or lower semicontinuous (in the sense that the level sets $\{X|\, \phi_B(X) \leq c\},\ c \in \mathbb{R}$ are weak*-closed). To be consistent with the risk measure notations we define their Fenchel transforms on $\mathcal{X}'$ as

$$\beta(q) = \sup_{X\in\mathcal{X}} \{q(-X) - \phi(X)\}.$$

When the linear form q is related to an additive finite measure $\mathbf{Q} \in \mathbf{M}_{1,f}$, we use the notation $q_{\mathbf{Q}}(X) = \mathbb{E}_{\mathbf{Q}}[X]$. For a general treatment of inf-convolution of convex functionals, the interested reader may refer to the highlighting paper of Borwein and Zhu [19]. The following theorem extends these results to the inf-convolution of convex functionals whose one of them at least is a convex risk measure:

Theorem 3.1 *Let ρ_A be a convex risk measure with penalty function α_A and ϕ_B be a proper closed convex functional with Fenchel transform β. Let $\rho_A \Box \phi_B$ be the inf-convolution of ρ_A and ϕ_B defined as*

$$X \to \rho_A \Box \phi_B(X) = \inf_{H\in\mathcal{X}} \big\{\rho_A(X-H) + \phi_B(H)\big\} \tag{15}$$

and assume that $\rho_A \Box \phi_B(0) > -\infty$. Then,
- *$\rho_A \Box \phi_B$ is a convex risk measure which is finite for all $X \in \mathcal{X}$.*
- *The associated penalty function $\alpha_{A,B}$ takes the value $+\infty$ for any q outside of $\mathbf{M}_{1,f}$, and*

$$\begin{aligned} &\forall \mathbf{Q} \in \mathbf{M}_{1,f} \quad \alpha_{A,B}\big(\mathbf{Q}\big) = \alpha_A\big(\mathbf{Q}\big) + \beta_B\big(q_{\mathbf{Q}}\big), \\ \text{and} \quad &\exists \mathbf{Q} \in \mathbf{M}_{1,f} \quad \text{s.t.} \quad \alpha_A\big(\mathbf{Q}\big) + \beta_B\big(q_{\mathbf{Q}}\big) < \infty. \end{aligned}$$

- *Moreover, if the risk measure ρ_A is continuous from below, then $\rho_A \Box \phi_B$ is also continuous from below.*

Proof: We give here the main steps of the proof of this theorem.
$\diamond$ The monotonicity and translation invariance properties of $\rho_A \Box \phi_B$ are immediate from the definition, since at least one of the both functionals have these properties.
$\diamond$ The convexity property simply comes from the fact that, for any X_A, X_B, H_A and H_B in $\mathcal{X}$ and any $\lambda \in [0,1]$, the following inequalities hold as ρ_A and ρ_B are convex functionals,

$$\begin{aligned} \rho_A\big((\lambda X_A + (1-\lambda)X_B) - (\lambda H_A + (1-\lambda)H_B)\big) &\leq \lambda\rho_A\big(X_A - H_A\big) + (1-\lambda)\rho_A\big(X_B - H_B\big) \\ \phi_B\big(\lambda H_A + (1-\lambda)H_B\big) &\leq \lambda\phi_B(H_A) + (1-\lambda)\phi_B\big(H_B\big). \end{aligned}$$

By adding both inequalities and taking the infimum in H_A and H_B on the left-hand side and separately in H_A and in H_B on the right-hand side, we obtain:

$$\rho_A \Box \phi_B \big(\lambda X_A + (1-\lambda) X_B\big) \leq \lambda\, \rho_A \Box \phi_B (X_A) + (1-\lambda) \rho_A \Box \rho_B (X_B).$$

$\diamond$ Using Equation (3), the associated penalty function is given, for any $\mathbf{Q} \in \mathbf{M}_{1,f}$, by

$$\begin{aligned}
\alpha_{A,B}(\mathbf{Q}) &= \sup\nolimits_{X\in\mathcal{X}} \left\{\mathbb{E}_{\mathbf{Q}}[-X] - \rho_{A,B}(X)\right\} \\
&= \sup\nolimits_{\Psi\in\mathcal{X}} \left\{\mathbb{E}_{\mathbf{Q}}[-X] - \inf\nolimits_{H\in\mathcal{X}} \left\{\rho_A(X-H) + \phi_B(H)\right\}\right\} \\
&= \sup\nolimits_{X\in\mathcal{X}} \sup\nolimits_{H\in\mathcal{X}} \left\{\mathbb{E}_{\mathbf{Q}}\big[-(X-H)\big] + \mathbb{E}_{\mathbf{Q}}[-H] - \rho_A(X-H) - \phi_B(H)\right\}
\end{aligned}$$

by letting $\widetilde{X} \triangleq X - H \in \mathcal{X}$

$$= \sup\nolimits_{\widetilde{X}\in\mathcal{X}} \sup\nolimits_{H\in\mathcal{X}} \big(\mathbb{E}_{\mathbf{Q}}[-\widetilde{X}] - \rho_A(\widetilde{X}) + \mathbb{E}_{\mathbf{Q}}[-H] - \phi_B(H)\big) = \alpha_A(\mathbf{Q}) + \beta_B(q_{\mathbf{Q}}).$$

When $q \notin \mathbf{M}_{1,f}$, the same equalities hold true. Since ρ_A is a convex risk measure, $\alpha_A(q) = +\infty$, and since β is a proper functional, $\beta(q)$ is dominated from below; so, $\alpha_{A,B}(q) = +\infty$. This equality $\alpha_{A,B} = \alpha_A + \beta_B$ holds even when α_A and β_B they take infinite values.

$\diamond$ The continuity from below is directly obtained upon considering an increasing sequence of $(X_n) \in \mathcal{X}$ converging to X. Using the monotonicity property, we have

$$\begin{aligned}
\inf_n \rho_A \Box \phi_B (X_n) &= \inf_n \inf_H \left\{\rho_A (X_n - H) + \phi_B (H)\right\} \\
&= \inf_H \inf_n \left\{\rho_A (X_n - H) + \phi_B (H)\right\} = \inf_H \left\{\rho_A (X - H) + \phi_B (H)\right\} \\
&= \rho_A \Box \phi_B (X). \quad \Box
\end{aligned}$$

We can now give an inf-convolution interpretation of the convex risk measure $\nu^{\mathcal{H}}$ generated by a convex set $\mathcal{H}$ as in Corollary 1.6 as the inf-convolution of the convex indicator function of $\mathcal{H}$, and the worst-case risk measure. This regularization may be applied at any proper convex functional.

Proposition 3.2 [Regularization by inf-convolution with ρ_{worst}] Let $\rho_{\text{worst}}(X) = \sup_\omega(-X(\omega))$ be the worst case risk measure.

i) ρ_{worst} is a neutral element for the infimal convolution of convex risk measures.

ii) Let $\mathcal{H}$ be a convex set such that $\inf\{m \,|\, \exists\, \xi \in \mathcal{H}\} > -\infty$. The convex risk measure generated by $\mathcal{H}$, $\nu^{\mathcal{H}}$ is the inf-convolution of the convex indicator functional of $\mathcal{H}$ with the worst case risk measure,

$$\nu^{\mathcal{H}} = \rho_{\text{worst}} \Box l^{\mathcal{H}}$$

iii) More generally, let ϕ be a proper convex functional, such that for any H, $\phi(H) \geq -\sup_\omega H(\omega) - c$. The infimal convolution of ρ_{worst} and ϕ, $\rho_\phi = \rho_{\text{worst}} \Box \phi$ is the largest convex risk measure dominated by ϕ.

iv) Let β the penalty functional associated with ϕ. Then, the penalty functional associated with ρ_ϕ is the functional α_ϕ, restriction of β at the set $\mathcal{M}_{1,f}$,

$$\begin{aligned}
\alpha_\phi(q) &= \beta(q) + l^{\mathcal{M}_{1,f}}(q) \\
&= \beta(q_{\mathbf{Q}}) \quad \text{if} \quad q_{\mathbf{Q}} \in \mathcal{M}_{1,f}, \qquad +\infty \quad \text{if not.}
\end{aligned}$$

v) Given a general risk measure ρ_A such that $\rho_A\Box\phi(0) > -\infty$, then

$$\rho_A\Box\phi = \rho_A\Box\rho_{\text{worst}}\Box\phi = \rho_A\Box\rho_\phi.$$

Proof: We start by proving that $\rho\Box\rho_{\text{worst}} = \rho$. By definition,

$$\begin{aligned}\rho\Box\rho_{\text{worst}}(X) &= \inf_Y\{\sup_\omega(-Y(\omega)) + \rho(X-Y)\} = \inf_Y\left\{\rho\big(X - (Y - \sup_\omega(-Y(\omega)))\big)\right\}\\ &= \inf_{Y\geq 0}\{\rho(X-Y)\} = \rho(X)\end{aligned}$$

To conclude, we have used the cash invariance of ρ and the fact that $\rho(X-Y) \geq \rho(X)$ whenever $Y \geq 0$.
ii) has been proved in Corollary 1.6.
iii) By Theorem 3.1, $\rho_\phi = \rho_{\text{worst}}\Box\phi$ is a convex risk measure. Since ρ_{worst} is a neutral element for the inf-convolution of risk measure, any risk measure ρ dominated by ϕ is also dominated by $\rho_{\text{worst}}\Box\phi$ since $\rho = \rho_{\text{worst}}\Box\rho \leq \rho_{\text{worst}}\Box\phi = \rho_\phi$. Hence the result. $\Box$

Therefore, in the following, we only consider the infimal convolution of convex risk measures. The following result makes more precise Theorem 3.1 and plays a key role in our analysis.

Theorem 3.3 [Sandwich Theorem] Let ρ_A and ρ_B be two convex risk measures.
Under the assumptions of Theorem 3.1 (i.e. $\rho_{A,B}(0) = \rho_A\Box\rho_B(0) > -\infty$),
i) There exists $\mathbf{Q} \in \partial\rho_{A,B}(0)$ such that, for any X and any Y,

$$\rho_A\Box\rho_B(0) \leq \big(\rho_A(X) - \mathbb{E}_{\mathbf{Q}}[-X]\big) + \big(\rho_B(Y) - \mathbb{E}_{\mathbf{Q}}[-Y]\big).$$

ii) Assume $\rho_{A,B}(0) \geq c$. There is an affine function, $a_{\mathbf{Q}}(X) = -\mathbb{E}_{\mathbf{Q}}[-X] + r$, with $\mathbf{Q} \in \partial\rho_{A,B}(0)$, satisfying

$$\rho_A(.) \geq a_{\mathbf{Q}} \geq -\rho_B(-.) + c. \tag{16}$$

Moreover, for any $\overline{X}$ such that $\rho_A(\overline{X}) + \rho_B(-\overline{X}) = \rho_{A,B}(0)$, $\mathbf{Q} \in \partial\rho_B(-\overline{X}) \cap \partial\rho_A(\overline{X})$. The inf-convolution is said to be exact at $\overline{X}$.
iii) Interpretation of the Condition $\rho_A\Box\rho_B(0) > -\infty$.
The following properties are equivalent:
$\diamond$ $\rho_A\Box\rho_B(0) > -\infty$.
$\diamond$ The sandwich property (16) holds for some affine function $a_{\mathbf{Q}}(X) = -\mathbb{E}_{\mathbf{Q}}[-X] + r$.
$\diamond$ There exists $\mathbf{Q} \in \mathbf{Dom}(\alpha_A) \cap \mathbf{Dom}(\alpha_B)$. $\diamond$Let ρ^A_{0+} (resp. ρ^B_{0+}) be the conservative risk measure associated with ρ^A (resp. ρ^B). Then

$$\rho^A_{0+}(X) + \rho^B_{0+}(-X) \geq 0.$$

Before proving this Theorem, let us make the following comment: the inf-convolution risk measure $\rho_{A,B}$, given in Equation (15) may also be defined, for instance, as the value functional of the program

$$\rho_{A,B}(\Psi) = \rho_A\Box\rho_B(\Psi) = \rho_A\Box\nu^{\mathcal{A}_{\rho_B}}(\Psi) = \inf\left\{\rho_A(\Psi - H), H \in \mathcal{A}_{\rho_B}\right\},$$

where $\nu^{\mathcal{A}_{\rho_B}}$ is the risk measure with acceptance set $\mathcal{A}_{\rho_B}$. This emphasizes again the key role played the risk measures generated by a convex set, if needed.

Proof: *i)* By Theorem 3.1, the convex risk measure $\rho_{A,B}$ is finite; so its subdifferential $\partial\rho_{A,B}(0)$ is non empty. More precisely, there exists $\mathbf{Q}_0 \in \partial\rho_{A,B}(0)$ such that $\rho_{A,B}(X) \geq \rho_{A,B}(0) + \mathbb{E}_{\mathbf{Q}_0}(-X)$. In other words,

$$\rho_{A,B}(0) \leq \rho_{A,B}(X) + \mathbb{E}_{\mathbf{Q}_0}(X) \leq \rho_A(X-Y) - \mathbb{E}_{\mathbf{Q}_0}[-(X-Y)] + \rho_B(Y) - \mathbb{E}_{\mathbf{Q}_0}[-Y].$$

ii − a) Assume that $\rho_A \Box \rho_B(0) \geq c$. Applying the previous inequality at $Y = -Z$, and $X = U + Y = U - Z$, we have

$$\rho_A(U) - \mathbb{E}_{\mathbf{Q}_0}[-U] \geq -\rho_B(-Z) - \mathbb{E}_{\mathbf{Q}_0}[Z] + \rho_{A,B}(0).$$

Then,

$$-\alpha_A(\mathbf{Q}_0) := \inf_U\{\rho_A(U) - \mathbb{E}_{\mathbf{Q}_0}[-U]\} \geq \alpha_B(\mathbf{Q}_0) + \rho_{A,B}(0) := \sup_Z\{-\rho_B(-Z) - \mathbb{E}_{\mathbf{Q}_0}[Z] + \rho_{A,B}(0)\}.$$

By Theorem 3.1 this inequality is in fact an equality. Picking $r = \alpha_A(\mathbf{Q}_0)$, and defining $a_{\mathbf{Q}_0}(X) = \mathbb{E}_{\mathbf{Q}_0}[-X]+r$ yield to an affine function that separates ρ_A and $-\rho_B(-.) + c$.
ii−b) Finally, when $\rho_A(\overline{X})+\rho_B(-\overline{X}) = \rho_{A,B}(0)$, by the above inequalities, we obtain $-\rho_B(-\overline{X})-\mathbb{E}_{\mathbf{Q}_0}[-\overline{X}] \geq -\rho_B(-Z) - \mathbb{E}_{\mathbf{Q}_0}[Z]$. In other words, $\mathbf{Q}_0$ belongs to $\partial\rho_B(-\overline{X})$. By symmetry, $\mathbf{Q}_0$ also belongs to $\partial\rho_A(\overline{X})$.
iii) ⋄ The implication (1) ⇒ (2) ⇒ (3) is clear, using the results i) and ii) of this Theorem.
⋄ Very naturally, one obtains (3) ⇒ (2) and (3) ⇒ (1) as the existence of $\mathbf{Q}_0 \in \mathbf{Dom}(\alpha_A) \cap \mathbf{Dom}(\alpha_B)$ implies that for any X, $\rho_A(X) \geq \mathbb{E}_{\mathbf{Q}_0}[-X] - \alpha_A(\mathbf{Q})$ and $\rho_B(-X) \geq \mathbb{E}_{\mathbf{Q}_0}[X] - \alpha_B(\mathbf{Q})$. Considering $r = \sup\{\alpha_A(\mathbf{Q})); \alpha_B(\mathbf{Q})\}$, one obtains (2). Moreover, $\rho_A(X) + \rho_B(-X) \geq -(\alpha_A(\mathbf{Q}) + \alpha_B(\mathbf{Q}))$ and taking the infimum with respect to X, $\rho_A \Box \rho_B(0) > -\infty$, i.e. the property (1).
⋄ Let us now look at the following implication (2) ⇒ (4). We first observe that (2), i.e., $\rho_A(X) \geq -\mathbf{E}_{\mathbf{Q}_0}[-X] + r$ implies $\rho^A_{0+}(X) \geq -\mathbf{E}_{\mathbf{Q}_0}[-X]$, and $\rho_B(-X) \geq \mathbf{E}_{\mathbf{Q}_0}[-X] - r$ implies $\rho^B_{0+}(-X) \geq \mathbf{E}_{\mathbf{Q}_0}[-X]$. Therefore, we obtain (4) as $\rho^A_{0+}(X) + \rho^B_{0+}(-X) \geq 0$.
⋄ The converse implication (4) ⇒ (2) is obtained by applying the sandwich property (16) to ρ^A_{0+} and ρ^B_{0+}.
□

Remark 3.4 (On risk measures on $\mathbb{L}_\infty(\mathbb{P})$) *Let us consider the inf-convolution between two risk measures ρ_A and ρ_B, where one of them, for instance ρ_A, is continuous from below (and consequently from above) and therefore is defined on $\mathbb{L}_\infty(\mathbb{P})$. In this case, as the inf-convolution maintains the property of continuity from below (see Theorem 3.1), the risk measure $\rho_A \Box \rho_B$ is also continuous from below and therefore is a risk measure on $\mathbb{L}_\infty(\mathbb{P})$, having a dual representation on $\mathcal{M}_{1,ac}(\mathbb{P})$.*

3.1.2 γ-Tolerant Risk Measures and Inf-Convolution

In this subsection, we come back to the particular class of γ-tolerant convex risk measures ρ_γ to give an explicit solution to the exact inf-convolution. Recall that this family of risk measures is generated from a root risk measure ρ by the following transformation $\rho_\gamma(\xi_T) = \gamma\rho_\gamma\left(\frac{\xi_T}{\gamma}\right)$ where γ is the risk tolerance coefficient with respect to the size of the exposure. These risk measures satisfy the following semi-group property for the inf-convolution:

Proposition 3.5 *Let* $(\rho_\gamma, \gamma > 0)$ *be the family of* γ*-tolerant risk measures issued of* ρ*. Then, the following properties hold:*
i) For any $\gamma_A, \gamma_B > 0, \quad \rho_{\gamma_A} \Box \rho_{\gamma_B} = \rho_{\gamma_A+\gamma_B}$*.*
ii) Moreover, $F^* = \frac{\gamma_B}{\gamma_A+\gamma_B} X$ *is an optimal structure for the minimization program:*

$$\rho_{\gamma_A+\gamma_B}(X) = \rho_{\gamma_A} \Box \rho_{\gamma_B}(X) = \inf_F \big\{\rho_{\gamma_A}(X-F) + \rho_{\gamma_B}(F)\big\} = \rho_{\gamma_A}\big(X - F^*\big) + \rho_{\gamma_B}\big(F^*\big).$$

The inf-convolution is said to be exact at F^**.*
iii) Let ρ *and* ρ' *be two convex risk measures. Then, for any* $\gamma > 0$, $\rho_\gamma \Box \rho'_\gamma = (\rho \Box \rho')_\gamma$*.*
iv) Assume $\rho(0) = 0$ *and* $\rho'(0) = 0$*. When* $\gamma = +\infty$*, this relationship still holds:* $\rho_\infty \Box \rho'_\infty = (\rho \Box \rho')_\infty$*.*
v) If $\rho_{0^+} \Box \rho'_{0^+}(0) > -\infty$*, we also have* $\rho_{0^+} \Box \rho'_{0^+} = (\rho \Box \rho')_{0^+}$*.*

Proof: Both $i)$ and $iii)$ are immediate consequences of the definition of infimal convolution.
$ii)$ We first study the stability property of the functional ρ_γ by studying the optimization program $\rho_{\gamma_A}(X - F) + \rho_{\gamma_B}(F) \to \min_F$ restricted to the family $\{\alpha X, \alpha \in \mathbb{R}\}$. Then, given the expression of the functional ρ_γ, a natural candidate becomes $F^* = \frac{\gamma_B}{\gamma_A+\gamma_B} X$, since

$$\rho_{\gamma_A}\big(X - F^*\big) + \rho_{\gamma_B}\big(F^*\big) = (\gamma_A + \gamma_B)\rho\big(\frac{1}{\gamma_A + \gamma_B} X\big) = \rho_{\gamma_C}(X).$$

$iv)$ The asymptotic properties are based on the non increase of the map $\gamma \to \rho_\gamma$. Then, when γ goes to infinity, pass to the limit is equivalent to take the infimum w.r. of γ and change the order of minimization, in such way that pass to the limit is justified.
$v)$ When γ goes to 0, the problem becomes a minimax problem, and we only obtain the inequality.
When the finite assumption holds, by Theorem 3.1, the minimal penalty function of $\rho_{0^+} \Box \rho'_{0^+}$ is $\alpha_{0^+} + \alpha'_{0^+}$. By the properties of conservative risk measures, α_{0^+} is the convex indicator of $\mathbf{Dom}(\alpha)$. So, $\alpha_{0^+} + \alpha'_{0^+} = l^{\mathbf{Dom}(\alpha) \cap \mathbf{Dom}(\alpha)'}$. On the other hand, the minimal penalty function of $(\rho \Box \rho')_{0^+}$ is the indicator of $\mathbf{Dom}(\alpha + \alpha')$. Since, α is dominated by the same minimal bound $-\rho(0)$, $\mathbf{Dom}(\alpha + \alpha') = \mathbf{Dom}(\alpha) \cap \mathbf{Dom}(\alpha')$ Both risk measures have same minimal penalty functions. This completes the proof. $\Box$

3.1.3 An Example of Inf-Convolution: the Market Modified Risk Measure

We now consider a particular inf-convolution which is closely related to Subsection 1.4 as it also deals with the question of optimal hedging.
More precisely, the following minimization problem

$$\inf_{H \in \mathcal{V}_T} \rho(X - H)$$

can be seen as an hedging problem, where $\mathcal{V}_T$ corresponds to the set of hedging instruments. It somehow consists of restricting the risk measure ρ to a particular set of admissible variables and is in fact the inf-convolution $\rho \Box \nu^{\mathcal{V}_T}$. Using Proposition 3.2, it can also be seen as the inf-convolution $\rho \Box l^{\mathcal{V}_T} \Box \rho_{\text{worst}}$. The main role of ρ_{worst} is to transform the convex indicator $l^{\mathcal{V}_T}$, which is not a convex risk measure (in particular, it is not translation invariant), into the convex risk measure $\nu^{\mathcal{V}_T}$.

The following corollary is an immediate extension of Theorem 3.1 as it establishes that the value functional of the problem, denoted by ρ^m, is a convex risk measure, called *market modified risk measure*.

Corollary 3.6 *Let $\mathcal{V}_T$ be a convex subset of $\mathbb{L}_\infty(\mathbb{P})$ and ρ be a convex risk measure with penalty function α such that* $\inf\{\rho(-\xi_T), \xi_T \in \mathcal{V}_T\} > -\infty$*. The inf-convolution of ρ and $\nu^{\mathcal{V}_T}$, $\rho^m \equiv \rho \square \nu^{\mathcal{V}_T}$, also defined as*

$$\rho^m(\Psi) \equiv \inf\{\rho(\Psi - \xi_T) \,|\, \xi_T \in \mathcal{V}_T\} = \rho \square l^{\mathcal{V}_T}(\Psi) \tag{17}$$

is a convex risk measure, called market modified risk measure, with minimal penalty function defined on $\mathbf{M}_{1,ac}(\mathbb{P})$, $\alpha^m(\mathbf{Q}) = \alpha(\mathbf{Q}) + \alpha^{\mathcal{V}_T}(\mathbf{Q})$.
Moreover, if ρ is continuous from below, ρ^m is also continuous from below.

This corollary makes precise the direct impact on the risk measure of the agent of the opportunity to invest optimally in a financial market.

Remark 3.7 *Note that the set $\mathcal{V}_T$ is rather general. In most cases, additional assumptions will be added and the framework will be similar to those described in Subsection 1.5.*

Acceptability and market modified risk measure: The market modified risk measure has to be related to the notion of acceptability introduced by Carr, Geman and Madan in [28]. In this paper, they relax the strict notion of hedging in the following way: instead of imposing that the final outcome of an acceptable position, suitably hedged, should always be non-negative, they simply require that it remains greater than an acceptable position. More precisely, using the same notations as in Subsection 1.5 and denoting by $\mathcal{A}$ a given acceptance set and by $\rho_{\mathcal{A}}$ its related risk measure, we can define the convex risk measure:

$$\bar{\nu}^{\mathcal{H}}(X) = \inf\{m \in \mathbb{R},\ \exists \theta \in \mathcal{K}\ \exists A \in \mathcal{A} \,:\, m + X + G(\theta) \geq A \ \mathbb{P}\, a.s.\}$$

To have a clearer picture of what this risk measure really is, let us first fix $G(\theta)$. In this case, we simply look at $\rho_{\mathcal{A}}(X + G(\theta))$. Then, the risk measure $\bar{\nu}^{\mathcal{H}}$ is defined by taking the infimum of $\rho_{\mathcal{A}}(X + G(\theta))$ with respect to θ,

$$\bar{\nu}^{\mathcal{H}}(X) = \inf_{\theta \in \mathcal{K}} \rho_{\mathcal{A}}(X + G(\theta)) = \inf_{H \in \mathcal{H}} \rho_{\mathcal{A}}(X - H)$$

Therefore, the risk measure $\bar{\nu}^{\mathcal{H}}$ is in fact the particular market modified risk measure $\rho^m = \nu^{\mathcal{H}} \square \rho_{\mathcal{A}}$. We obtain directly the following result of Föllmer and Schied [54] (Proposition 4.98): the minimal penalty function of this convex risk measure $\bar{\nu}^{\mathcal{H}}$ is given by

$$\bar{\alpha}^{\mathcal{H}}(\mathbf{Q}) = \alpha^{\mathcal{H}}(\mathbf{Q}) + \alpha(\mathbf{Q})$$

where $\alpha^{\mathcal{H}}$ is the minimal penalty function of $\nu^{\mathcal{H}}$ and α is the minimal penalty function of the convex risk measure with acceptance set $\mathcal{A}$.

4 Optimal Derivative Design

In this section, we now present our main problem, that of derivative optimal design (and pricing). The framework we generally consider involves two economic agents, at least one of them being exposed to a non-tradable risk. The risk transfer between both agents takes place through a structured contract denoted by F for an initial price π. The problem is therefore to design the transaction, in other words, to find the structure F and its price π. This transaction may occur only if both agents find some interest in doing this transaction. They express their satisfaction or interest in terms of the expected utility of their terminal wealth after the transaction, or more generally in terms of risk measures.

4.1 General Modelling

4.1.1 Framework

Two economic agents, respectively denoted by A and B, are evolving in an uncertain universe modelled by a standard measurable space $(\Omega, \Im)$ or, if a reference probability measure is given, by a probability space $(\Omega, \Im, \mathbb{P})$. In the following, for the sake of simplicity in our argumentation, we will make no distinction between both situations. More precisely, in the second case, all properties should hold $\mathbb{P} - a.s.$.

Both agents are taking part in trade talks to improve the distribution and management of their own risk. The nature of both agents can be quite freely chosen. It is possible to look at them in terms of a classical insured-insurer relationship, but from a more financial point of view, we may think of agent A as a market maker or a trader managing a particular book and of agent B as a traditional investor or as another trader. More precisely, we assume that at a future time horizon T, the value of agent A's terminal wealth, denoted by X_T^A, is sensitive to a non-tradable risk. Agent B may also have her own exposure X_T^B at time T. Note that by "terminal wealth", we mean the terminal value at the time horizon T of all capitalized cash flows paid or received between the initial time and T; no particular sign constraint is imposed. Agent A wants to issue a structured contract (financial derivative, insurance contract...) F with maturity T and forward price π to reduce her exposure X_T^A. Therefore, she calls on agent B. Hence, when a transaction occurs, the terminal wealth of the agent A and B are

$$W_T^A = X_T^A - F + \pi, \qquad W_T^B = X_T^B + F - \pi.$$

As before, we assume that all the quantities we consider belong to the Banach space $\mathcal{X}$, or, if a reference probability measure is given, to $\mathbb{L}_\infty(\mathbb{P})$.

The problem is therefore to find the optimal structure of the risk transfer (F, π) according to a given choice criterion, which is in our study a convex risk measure. More precisely, assuming that agent A (resp. agent B) assesses her risk exposure using a convex risk measure ρ_A (resp. ρ_B), agent A's objective is to choose the optimal structure (F, π) in order to minimize the risk measure of her final wealth

$$\rho_A(X_T^A - F + \pi) \rightarrow \inf_{F \in \mathcal{X}, \pi}.$$

Her constraint is then to find a counterpart. Hence, agent B should have an interest in doing this transaction. At least, the F-structure should not worsen her risk measure. Consequently, agent B simply compares the risk measures of two terminal wealth, the first one corresponds to the case of her initial exposure X_T^B and the second one to her new wealth if she enters the F-transaction,

$$\rho_B(X_T^B + F - \pi) \leq \rho_B(X_T^B).$$

4.1.2 Transaction Feasibility and Optimization Program

The optimization program as described above as

$$\inf_{F\in\mathcal{X},\pi} \rho_A(X_T^A - F + \pi) \qquad \text{subject to} \quad \rho_B(X_T^B + F - \pi) \leq \rho_B(X_T^B) \tag{18}$$

can be simplified using the cash translation invariance property. More precisely, binding the constraint imposed by agent B at the optimum and using the translation invariance property of ρ_B, we find directly the *optimal pricing rule* for a structure F:

$$\pi_B(F) = \rho_B\big(X_T^B\big) - \rho_B\big(X_T^B + F\big). \tag{19}$$

This pricing rule is an indifference pricing rule for agent B. It gives for any structure F the maximum amount agent B is ready to pay in order to enter the transaction.

Note also that this optimal pricing rule together with the cash translation invariance property of the functional ρ_A enable us to rewrite the optimization program (18) as follows, without any need for a Lagrangian multiplier:

$$\inf_{F\in\mathcal{X}} \big\{\rho_A\big(X_T^A - F\big) + \rho_B\big(X_T^B + F\big) - \rho_B\big(X_T^B\big)\big\}$$

or to within the constant $\rho_B(X_T^B)$ as:

$$R_{AB}(X_T^A, X_T^B) = \inf_{F\in\mathcal{X}} \{\rho_A\big(X_T^A - F\big) + \rho_B\big(X_T^B + F\big)\}. \tag{20}$$

Interpretation in Terms of Indifference Prices This optimization program (Program (20)) can be reinterpreted in terms of the indifference prices, using the notations introduced in the exponential utility framework in Subsection 1.1.2. To show this, we introduce the constants $\rho_A\big(X_T^A\big)$ and $\rho_B\big(X_T^B\big)$ in such a way that Program (20) is equivalent to:

$$\inf_{F\in\mathcal{X}} \{\rho_A\big(X_T^A - F\big) - \rho_A\big(X_T^A\big) + \rho_B\big(X_T^B + F\big) - \rho_B\big(X_T^B\big)\}.$$

Then, using the previous comments, it is possible to interpret $\rho_A\big(X_T^A - F\big) - \rho_A\big(X_T^A\big)$ as $\pi_A^s\big(F|X_T^A\big)$, i.e. the seller's indifference pricing rule for F given agent A's initial exposure X_T^A, while $\rho_B\big(X_T^B + F\big) - \rho_B\big(X_T^B\big)$ is simply the opposite of $\pi_B^b(F|X_T^B)$, the buyer's indifference pricing rule for F given agent B's initial exposure X_T^B. For agent A, everything consists then of choosing the structure as to minimize the difference between her (seller's) indifference price (given X_T^A) and the (buyer's) indifference price imposed by agent B:

$$\inf_{F\in\mathcal{X}} \big\{\pi_A^s\big(F|X_T^A\big) - \pi_B^b\big(F|X_T^B\big)\big\} \leq 0. \tag{21}$$

Note that for $F \equiv 0$, the spread between both transaction indifference prices is equal to 0. Hence, the infimum is always non-positive. This is completely coherent with the idea that the optimal transaction obviously reduces the risk of agent A. The transaction may occur since the minimal seller price is less than the maximal buyer price.

For agent A, everything can also be expressed as the following maximization program

$$\sup_{F\in\mathcal{X}} \big\{\pi_B^b\big(F|X_T^B\big) - \pi_A^s\big(F|X_T^A\big)\big\}. \tag{22}$$

The interpretation becomes then more obvious since the issuer has to optimally choose the structure in order to maximize the "ask-bid" spread associated with transaction.

Relationships with the Insurance Literature and the Principal-Agent Problem The relationship between both agents is very similar to a Principal-Agent framework. Agent A plays an active role in the transaction. She chooses the "payment structure" and then is the "Principal" in our framework. Agent B, on the other hand, is the "Agent" as she simply imposes a price constraint to the Principal and in this sense is rather passive.

Such a modelling framework is also very similar to an insurance problem: Agent A is looking for an optimal "insurance" policy to cover her risk (extending here the simple notion of loss as previously mentioned). In this sense, she can be seen as the "insured". On the other hand, Agent B accepts to bear some risk. She plays the same role as an "insurer" for Agent A. In fact, this optimal risk transfer problem is closely related to the standard issue of optimal policy design in insurance, which has been widely studied in the literature (see for instance Borch [18], Bühlman [23], [24] and [25], Bühlman and Jewell [27], Gerber [60], Raviv [99]). One of the fundamental characteristics of an insurance policy design problem is the sign constraint imposed on the risk, that should represent a loss. Other specifications can be mentioned as moral hazard or adverse selection problems that have to be taken into account when designing a policy (for more details, among a wide literature, refer for instance to the two papers on the relation Principal-Agent by Rees [100] and [101]). These are related to the potential influence of the insured on the considered risk.
Transferring risk in finance is somehow different. Risk is then taken in a wider sense as it represents the uncertain outcome. The sign of the realization does not a priori matter in the design of the transfer. The derivative market is a good illustration of this aspect: forwards, options, swaps have particular payoffs which are not directly related to any particular loss of the contract's seller.

4.2 Optimal Transaction

This subsection aims at solving explicitly the optimization Program (20):

$$R_{AB}(X_T^A, X_T^B) = \inf_{F \in \mathcal{X}} \{\rho_A\big(X_T^A - F\big) + \rho_B\big(X_T^B + F\big)\}.$$

The value functional $R_{AB}(X_T^A, X_T^B)$ can be seen as the *residual risk measure* after the F-transaction, or equivalently as a measure of the risk remaining after the transaction. It obviously depends on both initial exposures X_T^A and X_T^B since the transaction consists of an optimal redistribution of the respective risk of both agents.
Let us denote by $\widetilde{F} \equiv X_T^B + F \in \mathcal{X}$. The program to be solved becomes

$$R_{AB}\big(X_T^A, X_T^B\big) = \inf_{\widetilde{F} \in \mathcal{X}} \{\rho_A\big(X_T^A + X_T^B - \widetilde{F}\big) + \rho_B\big(\widetilde{F}\big)\},$$

or equivalently, using Section 3, it can be written as the following inf-convolution problem

$$R_{AB}\big(X_T^A, X_T^B\big) = \rho_A \square \rho_B (X_T^A + X_T^B). \tag{23}$$

As previously mentioned in Theorem 3.1, the condition $\rho_A \Box \rho_B(0) > -\infty$ is required when considering this inf-convolution problem. This condition is equivalent to $\forall \xi \in \mathcal{X}$, $\rho^A_{0^+}(\xi) + \rho^B_{0^+}(-\xi) \geq 0$ (Theorem 3.3 iii)). This property has a nice economic interpretation, since it says that the inf-convolution program makes sense if and only if *for any derivative ξ, the conservative seller price of the agent A, $-\rho^A_{0^+}(\xi)$, is less than the conservative buyer price of the agent B, $\rho^B_{0^+}(-\xi)$.*

In the following, we assume such a condition to be satisfied. The problem is not to study the residual risk measure as previously but to characterize the optimal structure $\tilde{F}^*$ or F^* such that the inf-convolution is exact at this point.
To do so, we first consider a particular framework where the optimal transaction can be explicitly identified. This corresponds to a well-studied situation in economics where both agents belong to the same family.

4.2.1 Optimal Transaction between Agents with Risk Measures in the Same Family

More precisely, we now assume that both agents have γ-tolerant risk measures ρ_{γ_A} and ρ_{γ_B} from the same root risk measure ρ with risk tolerance coefficients γ_A and γ_B, as introduced in Subsection 2.1. In this framework, the optimization program (23) is written as follows:

$$R_{AB}\big(X^A_T, X^B_T\big) = \rho_{\gamma_A} \Box \rho_{\gamma_B}(X^A_T + X^B_T)$$

In this framework, the optimal risk transfer is consistent with the so-called Borch's theorem. In this sense, the following result can be seen as an extension of this theorem since the framework we consider here is different from that of utility functions. In his paper [18], Borch obtained indeed, in a utility framework, optimal exchange of risk, leading in many cases to familiar linear quota-sharing of total pooled losses.

Theorem 4.1 (Borch [18]) *The residual risk measure after the transaction is given by:*

$$R_{AB}\big(X^A_T, X^B_T\big) = \inf_{F \in \mathcal{X}} \big\{\rho_{\gamma_A}\big(X^A_T - F\big) + \rho_{\gamma_B}\big(X^B_T + F\big)\big\} = \rho_{\gamma_C}(X^A_T + X^B_T) \quad \textit{with} \quad \gamma_C = \gamma_A + \gamma_B.$$

The optimal structure is given as a proportion of the initial exposures X^A_T and X^B_T, depending only on the risk tolerance coefficients of both agents:

$$F^* = \frac{\gamma_B}{\gamma_A + \gamma_B} X^A_T - \frac{\gamma_A}{\gamma_A + \gamma_B} X^B_T \qquad \textit{(to within a constant).} \tag{24}$$

The equality in the equation (24) has to be understood $\mathbb{P}\,a.s.$ if the space of structured products is $\mathbb{L}_\infty(\mathbb{P})$.

Proof: The optimization program (20) to be solved (with $\tilde{F} \equiv X^B_T + F \in \mathcal{X}$) is

$$R_{AB}\big(X^A_T, X^B_T\big) = \inf_{\tilde{F} \in \mathcal{X}} \big(\rho_{\gamma_A}\big(X^A_T + X^B_T - \widetilde{F}\big) + \rho_{\gamma_B}\big(\widetilde{F}\big)\big).$$

Using Proposition 3.5, the optimal structure $\tilde{F}^*$ is $\widetilde{F}^* = \frac{\gamma_B}{\gamma_A + \gamma_B}(X^A_T + X^B_T)$. The result is then obtained by replacing $\widetilde{F}^*$ by $F^* - X^B_T$. $\Box$

Comments and properties: *i*) Both agents are transferring a part of their initial risk according to their relative tolerance. The optimal risk transfer underlines the symmetry of the framework for both agents.

Moreover, even if the issuer, agent A has no exposure, a transaction will occur between both agents. The structure F enables them to exchange a part of their respective risk. Note that if none of the agents is initially exposed, no transaction will occur. In this sense, the transaction has a non-speculative underlying logic.

ii) Note also that the composite parameter γ_C is simply equal to the sum of both risk tolerance coefficients γ_A and γ_B. This may justify the use of risk tolerance instead of risk aversion where harmonic mean has to be used.

4.2.2 Individual Hedging as a Risk Transfer

In this subsection, we now focus on the individual hedging problem of agent A and see how this problem can be interpreted as a particular risk transfer problem. The question of optimal hedging has been widely studied in the literature under the name of "hedging in incomplete markets and pricing via utility maximization" in some particular framework. Most of the studies have considered exponential utility functions. Among the numerous papers, we may quote the papers by Frittelli [55], El Karoui and Rouge [51], Delbaen et al. [39], Kabanov and Stricker [75] or Becherer [11].

We assume that agent A assesses her risk using a ($\mathbb{L}_\infty(\mathbb{P})$) risk measure ρ_A. She can (partially) hedge her initial exposure X using instruments from a convex subset $\mathcal{V}_T^A$ (of $\mathbb{L}_\infty(\mathbb{P})$). Her objective is to minimize the risk measure of her terminal wealth.

$$\inf_{\xi\in\mathcal{V}_T^A} \rho_A\big(X_T^A - \xi\big). \tag{25}$$

The $\mathbb{L}_\infty(\mathbb{P})$ framework has been carefully described in Subsection 1.5. In particular, to have coherent transaction prices, we assume in the following that the market is arbitrage-free.

As already mentioned in Subsection 3.1.3, the opportunity to invest optimally in a financial market has a direct impact on the risk measure of the agent and transforms her initial risk measure ρ_A into the market modified risk measure $\rho_A^m = \rho_A \square \nu^A$.

This inf-convolution problem makes sense if the condition $\rho_A^m(0) > -\infty$ is satisfied. The hedging problem of agent A is identical to the previous risk transfer problem (20), agent B being now the financial market with the associated risk measure ν^A.

Existence of an Optimal Hedge The question of the existence of an optimal hedge can be answered using different approaches. One of them is based on analysis techniques and we present it in this subsection. In the following, however, when introducing dynamic risk measures, we will consider other methods leading to a more constructive answer.

In this subsection, we are interested in studying the existence of a solution for the hedging problem of agent A (Program (25)) or equivalently for the inf-convolution problem in $\mathbb{L}_\infty(\mathbb{P})$. The following of existence can be obtained:

Theorem 4.2 *Let $\mathcal{V}_T$ be a convex subset of $\mathbb{L}_\infty(\mathbb{P})$ and ρ be a convex risk measure on $\mathbb{L}_\infty(\mathbb{P})$ continuous from below, such that $\inf_{\xi\in\mathcal{V}_T}\rho(-\xi) > -\infty$.*

Assume the convex set $\mathcal{V}_T$ bounded in $\mathbb{L}^\infty(\mathbb{P})$. The infimum of the hedging program

$$\rho^m(X) \triangleq \inf_{\xi \in \mathcal{V}_T} \rho(X-\xi)$$

is "attained" for a random variable ξ_T^* in $\mathbb{L}^\infty(\mathbb{P})$, belonging to the closure of $\mathcal{V}_T$ with respect to the a.s. convergence.

Proof: First note that the proof of this theorem relies on arguments similar to those used by Kabanov and Stricker [75]. In particular, a key argument is the Komlos Theorem (Komlos [82]):

Lemma 4.3 (Komlos) *Let (ϕ_n) be a sequence in $\mathbb{L}^1(\mathbb{P})$ such that $\sup_n \mathbb{E}_\mathbb{P}(|\phi_n|) < +\infty$. Then there exists a subsequence $(\phi_{n'})$ of (ϕ_n) and a function $\phi^* \in L^1(\mathbb{P})$ such that for every further subsequence $\big(\phi_{n''}\big)$ of (ϕ_n), the Cesaro-means of these subsequences converge to φ^*, that is*

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n''=1}^{N} \phi_{n''}(\omega) = \phi^*(\omega) \qquad \textit{for almost every } \omega \in \Omega.$$

We first show that the set $S_r = \{\xi \in \mathbb{L}_\infty(\mathbb{P}) \big| \, \rho(X-\xi) \leq r\}$ is closed for the weak*-topology. To do that, by the Krein-Smulian theorem ([54] Theorem A.63), it is sufficient to show that $S_r \cap \{\xi; \|\xi\|_\infty \leq C\}$ is closed in $\mathbb{L}^\infty(\mathbb{P})$.
Let $(\xi_n \in \mathcal{V}_T)$ be a sequence bounded by C, converging in $\mathbb{L}^\infty$ to ξ^*. A subsequence still denoted by ξ_n converges a.s. to ξ^*. Since ρ_A is continuous from below, ρ is continuous w.r. to pointwise convergence of bounded sequences and then ξ^* belongs to S_r. S_r is weak*-closed.
Given the assumption that (ξ_n) is $\mathbb{L}^\infty$-bounded, we can apply Komlos lemma: therefore, there exists a subsequence $(\xi_{j_k} \in \mathcal{V}_T^A)$ such that the Cesaro-means, $\widetilde{\xi}_n \triangleq \frac{1}{n} \sum_{k=1}^{n} \xi_{j_k}$ converges almost surely to $\xi^* \in \mathbb{L}^\infty(\mathbb{P})$. Note that $\widetilde{\xi}_n$ belongs to $\mathcal{V}_T^A$ as a convex combination of elements of $\mathcal{V}_T^A$. So ξ^* belongs to the $a.s.$ closure of $\mathcal{V}_T^A$. Since ρ_A is continuous from below, ρ is continuous w.r. to pointwise convergence of bounded sequences.

$$\limsup_n \rho_A\big(X-\widetilde{\xi}_n\big) \leq \rho_A\big(X-\xi^*\big) = \rho_A\big(\lim_n \big(X-\widetilde{\xi}_n\big)\big) \leq \liminf_n \rho_A\big(X-\widetilde{\xi}_n\big).$$

Then, $\rho_A^m(X) \leq \rho_A(X-\xi^*) \leq \lim_n \inf \rho_A\big(\frac{1}{n}\sum_{k=1}^{n}(X-\xi_{j_k})\big) \leq \lim_n \inf \frac{1}{n}\sum_{k=1}^{n} \rho_A(X-\xi_{j_k})$ by Jensen inequality. Finally, given the convergence of $\rho_A\big(X-\xi_{j_k}\big)$ to $\rho_A^m(X)$, the Cesaro-means also converge and $\rho_A(X-\xi^*) = \inf_{\xi \in \mathcal{V}_T^A} \rho_A(X-\xi)$. $\square$

4.2.3 γ-Tolerant Risk Measures: Derivatives Design with Hedging Opportunities

We now consider the situation where both agents A and B have a γ-dilated risk measure, defined on $\mathbb{L}_\infty(\mathbb{P})$ and continuous from above. Moreover, they may reduce their risk by transferring it between themselves but also by investing in the financial market, choosing optimally their financial investments.
The investment opportunities of both agents are described by two convex subsets $\mathcal{V}_T^A$ and $\mathcal{V}_T^B$ of $\mathbb{L}_\infty(\mathbb{P})$. In order to have coherent transaction prices, we assume that the market is arbitrage-free. In our framework, this can be expressed as the existence of a probability measure which is **equivalent** to $\mathbb{P}$ in both sets of

probability measures $\mathcal{M}_{\mathcal{V}_T^i} = \{\mathbb{Q} \in \mathcal{M}_{1,e}(\mathbb{P}); \forall \xi \in \mathcal{V}_T^i, \mathbb{E}_{\mathbb{Q}}[-\xi] \leq 0\}$ for $i = A, B$. Equivalently,

$$\exists \mathbb{Q} \sim \mathbb{P} \quad \text{s.t.} \quad \mathbb{Q} \in \mathcal{M}_{\mathcal{V}_T^A} \cap \mathcal{M}_{\mathcal{V}_T^B}.$$

This opportunity to invest optimally in a financial market reduces the risk of both agents. To assess their respective risk exposure, they now refer to market modified risk measures $\rho^m_{\gamma_A}$ and $\rho^m_{\gamma_B}$ defined if $J = A, B$ as

$$\rho^m_{\gamma_A}(\Psi) = \rho_{\gamma_A} \Box \nu^A(\Psi) \qquad \text{and} \qquad \rho^m_{\gamma_B}(\Psi) = \rho_{\gamma_B} \Box \nu^B(\Psi).$$

Let us consider directly the optimal risk transfer problem with these market modified risk measures, i.e.

$$R^m_{AB}\big(X_T^A + X_T^B\big) = \inf_{F \in \mathcal{X}} \big\{\rho^m_{\gamma_A}\big(X_T^A - F\big) + \rho^m_{\gamma_B}\big(X_T^B + F\big)\big\} \tag{26}$$

The details of this computation will be given in the next subsection, when considering the general framework. The residual risk measure $R^m_{AB}\big(X_T^A + X_T^B\big)$ defined in equation (26) may be simplified using the commutativity property of the inf-convolution and the semi-group property of γ-tolerant risk measures:

$$\begin{aligned} R^m_{AB}\big(X_T^A + X_T^B\big) &= \rho^m_{\gamma_A} \Box \rho^m_{\gamma_B}\big(X_T^A + X_T^B\big) \\ &= \rho_{\gamma_A} \Box \nu^A \Box \rho_{\gamma_B} \Box \nu^B\big(X_T^A + X_T^B\big) \\ &= \rho_{\gamma_A} \Box \rho_{\gamma_B} \Box \nu^A \Box \nu^B\big(X_T^A + X_T^B\big) \\ &= \rho_{\gamma_C} \Box \nu^A \Box \nu^B\big(X_T^A + X_T^B\big). \end{aligned}$$

where ρ_{γ_C} is the γ-tolerant risk measure associated with the risk tolerance coefficient $\gamma_C = \gamma_A + \gamma_B$.

This inf-convolution program makes sense under the initial condition $\rho^m_{\gamma_A} \Box \rho^m_{\gamma_B}(0) > -\infty$. Such an assumption is made. The following theorem gives the optimal risk transfer in different situations depending on the access both agents have to the financial markets.

Theorem 4.4 *Let both agents have γ-tolerant risk measures with respective risk tolerance coefficients γ_A and γ_B.*
(a) If both agents have the same access to the financial market from a **cone**, *$\mathcal{V}_T$, then an optimal structure, solution of the minimization Program* (26) *is given by:*

$$F^* = \frac{\gamma_B}{\gamma_A + \gamma_B} X_T^A - \frac{\gamma_A}{\gamma_A + \gamma_B} X_T^B.$$

(b) Assume that both agents have different access to the financial market via **two convex sets** *$\mathcal{V}_T^A$ and $\mathcal{V}_T^B$. Suppose $\xi^* = \eta_A^* + \eta_B^*$ is an optimal solution of the Program* $\inf_{\xi \in \mathcal{V}_T^{(A+B)}} \rho_{\gamma_C}\big(X_T^A + X_T^B - \xi\big)$ *with $\eta_A^* \in \mathcal{V}_T^{(A)}$, $\eta_B^* \in \mathcal{V}_T^{(B)}$ and $\mathcal{V}_T^{(A+B)} = \{\xi_T^A + \xi_T^B \mid \xi_T^A \in \mathcal{V}_T^{(A)},\ \xi_T^B \in \mathcal{V}_T^{(B)}\}$. Then*

$$F^* = \frac{\gamma_B}{\gamma_A + \gamma_B} X_T^A - \frac{\gamma_A}{\gamma_A + \gamma_B} X_T^B - \frac{\gamma_B}{\gamma_A + \gamma_B} \eta_A^* + \frac{\gamma_A}{\gamma_A + \gamma_B} \eta_B^*$$

is an optimal structure. Moreover,
i) η_B^* is an optimal hedging portfolio of $\big(X_T^B + F^*\big)$ for Agent B

$$\frac{1}{\gamma_B} \rho_{\gamma_B}\big(X_T^B + F^* - \eta_B^*\big) = \frac{1}{\gamma_B} \inf_{\xi_B \in \mathcal{V}_T^{(B)}} \rho_{\gamma_B}\big(X_T^B + F^* - \xi_B\big) = \frac{1}{\gamma_C} \rho_{\gamma_C}\big(X_T^A + X_T^B - \xi^*\big).$$

ii) η_A^* *is an optimal hedging portfolio of* $\left(X_T^A - F^*\right)$ *for Agent A*

$$\frac{1}{\gamma_A}\rho_{\gamma_A}\left(X_T^A - \left(F^* + \eta_A^*\right)\right) = \frac{1}{\gamma_A}\inf_{\xi_A \in \mathcal{V}_T^{(A)}} \rho_{\gamma_A}\left(X_T^A - \left(F^* + \xi_A\right)\right) = \frac{1}{\gamma_C}\rho_{\gamma_C}\left(X_T^A + X_T^B - \xi^*\right).$$

Proof: To prove this theorem, we proceed in several steps:

Step 1:

Let us first observe that

$$R_{AB}^m\left(X_T^A + X_T^B\right) = \rho_{\gamma_C}\left(X_T^A + X_T^B - \xi^*\right) = \inf_{\widetilde{F}\in\mathcal{X}}\left(\rho_{\gamma_A}\left(X_T^A + X_T^B - \widetilde{F} - \xi^*\right) + \rho_{\gamma_B}\left(\widetilde{F}\right)\right),$$

where $\widetilde{F} = F + X_T^B - \xi_B$. Given Proposition 3.5, we obtain directly an expression for the optimal "structure" $\widetilde{F}^*$ as: $\widetilde{F}^* = \frac{\gamma_B}{\gamma_A+\gamma_B}\left(X_T^A + X_T^B - \xi^*\right) = \frac{\gamma_B}{\gamma_C}\left(X_T^A + X_T^B - \xi^*\right)$. Moreover, $\rho_{\gamma_B}(\widetilde{F}) = \frac{\gamma_B}{\gamma_C}(X_T^A + X_T^B - \xi^*)$.

Step 2:

Rewriting in the reverse order, we naturally set $F^* = \widetilde{F}^* - X_T^B + \eta_B^*$. We then want to prove that η_B^* is an optimal investment for agent B.

For the sake of simplicity in our notation, we consider $G^X\left(\xi_A,\xi_B,F\right) \triangleq \rho_{\gamma_A}\left(X_T^A - F - \xi_A\right) + \rho_{\gamma_B}\left(X_T^B + F - \xi_B\right)$.

Given the optimality of $\xi^* = \eta_A^* + \eta_B^*$ and $\widetilde{F}^* = F^* + X_T^B - \eta_B^*$, we have

$$\begin{aligned} R_{AB}^m\left(X_T^A + X_T^B\right) &= G^X\left(\eta_A^*, \eta_B^*, F^*\right) \\ &= \inf_{F\in\mathcal{X},\xi_A\in\mathcal{V}_T^{(A)},\xi_B\in\mathcal{V}_T^{(B)}} G^X\left(\xi_A,\xi_B,F\right) \leq \inf_{\xi_B\in\mathcal{V}_T^{(B)}} G^X\left(\eta_A^*,\xi_B,F^*\right) \leq G^X\left(\eta_A^*,\eta_B^*,F^*\right). \end{aligned}$$

Then η_B^* is optimal for the problem $\rho_{\gamma_B}\left(F - \xi_B\right) \to \inf_{\xi_B\in\mathcal{V}_T^{(B)}}$. The optimality of η_A^* can be proved using the same arguments. $\square$

Remark 4.5 *(a) We first assume that both agents have the same access to the financial market from a cone* $\mathcal{H}$. *Given the fact that the risk measure generated by* $\mathcal{H}$ *is coherent and thus invariant by dilatation, the market modified risk measures of both agents are generated from the root risk measure* $\rho\square\nu^{\mathcal{H}} = \rho^{\mathcal{H}}$ *as* $\rho_A^{\mathcal{H}} = \rho_{\gamma_A}\square\nu^{\mathcal{H}} = \rho_{\gamma_A}\square\nu_{\gamma_A}^{\mathcal{H}} = \left(\rho\square\nu^{\mathcal{H}}\right)_{\gamma_A} = \rho_{\gamma_A}^{\mathcal{H}}$ *and* $\rho_B^{\mathcal{H}} = \rho_{\gamma_B}^{\mathcal{H}}$.

(b) In a more general framework, when both agents have different access to the financial market, the convex set $\mathcal{V}_T^{(A+B)}$ *associated with the risk measure* $\nu^{(A+B)} = \nu^A\square\nu^B$ *plays the same role as the set* $\mathcal{H}$ *above, since* $\rho_{\gamma_C}\square\nu^A\square\nu^B\left(X_T^A + X_T^B\right) = \rho_{\gamma_C}\square\nu^{(A+B)}\left(X_T^A + X_T^B\right)$.

Comments: Note that when both agents have the same access to the financial market, it is optimal to transfer the same proportion of the initial risk as in the problem without market. This result is very strong as it does not require any specific assumption either for the non-tradable risk or the financial market. Moreover, the optimal structure F^* does not depend on the financial market. The impact of the financial market is simply visible through the pricing rule, which depends on the market modified risk measure of agent B.

Standard diversification will also occur in exchange economies as soon as agents have proportional penalty functions. The regulator has to impose very different rules on agents as to generate risk measures with non-proportional penalty functions if she wants to increase the diversification in the market. In other words, diversification occurs when agents are very different one from the other. This result supports for instance the intervention of reinsurance companies on financial markets in order to increase the diversification on the reinsurance market.

4.2.4 Optimal Transaction in the General Framework

We now come back to our initial problem of optimal risk transfer between agent A and agent B, when now they both have access to the financial market to hedge and diversify their respective portfolio.

General framework As in the dilated framework, we assume that both their risk measures ρ_A and ρ_B are defined on $\mathbb{L}_\infty(\mathbb{P})$ and are continuous from above. The investment opportunities of both agents are described by two convex subsets $\mathcal{V}_T^A$ and $\mathcal{V}_T^B$ of $\mathbb{L}_\infty(\mathbb{P})$ and the financial market is assumed to be arbitrage-free.

1. This opportunity to invest optimally in a financial market reduces the risk of both agents. To assess their respective risk exposure, they now refer to market modified risk measures ρ_A^m and ρ_B^m defined if $J = A, B$ as $\rho_J^m(\Psi) \triangleq \inf_{\xi_J \in \mathcal{V}_T^{(J)}} \rho_J(\Psi - \xi_J)$. As usual, we assume that $\rho_J^m(0) > -\infty$ for the individual hedging programs to make sense. Thanks to Corollary 3.6,
$$\rho_A^m(\Psi) = \rho_A \Box \nu^A(\Psi) \qquad \text{and} \qquad \rho_B^m(\Psi) = \rho_B \Box \nu^B(\Psi).$$

2. Consequently, the optimization program related to the F-transaction is simply
$$\inf_{F,\pi} \rho_A^m\big(X_T^A - F + \pi\big) \qquad \text{subject to} \qquad \rho_B^m\big(X_T^B + F - \pi\big) \leq \rho_B^m\big(X_T^B\big).$$
As previously, using the cash translation invariance property and binding the constraint at the optimum, the pricing rule of the F-structure is fully determined by the buyer as
$$\pi^*(F) = \rho_B^m\big(X_T^B\big) - \rho_B^m\big(X_T^B + F\big). \tag{27}$$
It corresponds to an "indifference" pricing rule from the agent B's market modified risk measure.

3. Using again the cash translation invariance property, the optimization program simply becomes
$$\inf_F \left\{\rho_A^m\big(X_T^A - F\big) + \rho_B^m\big(X_T^B + F\big)\right\} - \rho_B^m\big(X_T^B\big) \triangleq R_{AB}^m\big(X_T^A + X_T^B\big) - \rho_B^m\big(X_T^B\big).$$
With the functional R_{AB}^m, we are in the framework of Theorem 3.1.
$$\begin{aligned} R_{AB}^m\big(X_T^A + X_T^B\big) &= \inf_F \left\{\rho_A^m\big(X_T^A - F\big) + \rho_B^m\big(X_T^B + F\big)\right\} && (28) \\ &= \inf_{\widetilde{F}} \left\{\rho_A^m\big(X_T^A + X_T^B - \widetilde{F}\big) + \rho_B^m\big(\widetilde{F}\big)\right\} = \rho_A^m \Box \rho_B^m\big(X_T^A + X_T^B\big) \\ &= \rho_A \Box \nu^A \Box \rho_B \Box \nu^B\big(X_T^A + X_T^B\big). && (29) \end{aligned}$$
The value functional R_{AB}^m of this program, resulting from the inf-convolution of four different risk measures, may be interpreted as the *residual risk measure* after all transactions. This inf-convolution problem makes sense if the initial condition $\rho_A^m \Box \rho_B^m(0) > -\infty$ is satisfied.

4. Using the previous Theorem 3.1 on the stability of convex risk measure, provided the initial condition is satisfied, R^m_{AB} is a convex risk measure with the penalty function $\alpha^m_{AB} = \alpha^m_A + \alpha^m_B = \alpha_A + \alpha_B + \alpha^{\mathcal{V}^A_T} + \alpha^{\mathcal{V}^B_T}$.

Comments: The general risk transfer problem can be viewed as a game involving four different agents if the access to the financial market is different for agent A and agent B (or three otherwise). As a consequence, we end up with an inf-convolution problem involving four different risk measures, two per agents.

Optimal design problem Our problem is to find an optimal structure F^* realizing the minimum of the Program (28):

$$R^m_{AB}\big(X^A_T + X^B_T\big) = \inf_F \left\{\rho^m_A\big(X^A_T - F\big) + \rho^m_B\big(X^B_T + F\big)\right\}$$

Let us first consider the following simple inf-convolution problem between a convex risk measure ρ_B and a linear function q_A as introduced in Subsection 3.1:

$$q_A \Box \rho_B(X) = \inf_F \{\mathbb{E}_{\mathbb{Q}_A}[-(X-F)] + \rho_B(F)\}. \tag{30}$$

Proposition 4.6 *The necessary and sufficient condition to have an optimal solution F^* to the linear inf-convolution problem (30) is expressed in terms of the subdifferential of ρ_B as $\mathbb{Q}_A \in \partial\rho_B(F^*)$.*

This necessary and sufficient corresponds to the first order condition of the optimization problem. More generally, the following result is obtained:

Theorem 4.7 (Characterization of the optimal) *Assume that $\rho^m_A \Box \rho^m_B(0) > -\infty$.*
The inf-convolution program

$$R^m_{AB}\big(X^A_T + X^B_T\big) = \inf_F \left\{\rho^m_A\big(X^A_T - F\big) + \rho^m_B\big(X^B_T + F\big)\right\}$$

is exact at F^* if and only if there exists $\mathbf{Q}^X_{AB} \in \partial R^m_{AB}(X^A_T + X^B_T)$ such that $\mathbf{Q}^X_{AB} \in \partial\rho^m_A(X^A_T - F^*) \cap \partial\rho^m_B(X^B_T + F^*)$.
In other words, the necessary and sufficient condition to have an optimal solution F^* to the inf-convolution program is that there exists an optimal additive measure $\mathbf{Q}^X_{AB}$ for $\big(X^A_T + X^B_T, R^m_{AB}\big)$ such that $X^B_T + F^*$ is optimal for $\big(\mathbf{Q}^X_{AB}, \alpha^m_B\big)$ and $X^A_T - F^*$ is optimal for $\big(\mathbf{Q}^X_{AB}, \alpha^m_A\big)$.

Both notions of optimality are rather intuitive as they simply translate the fact that the dual representations of the risk measure on the one hand, and of the penalty function on the other hand, are exact respectively at a given additive measure and at a given exposure.
A natural interpretation of this theorem is that both agents agree on the measure $\mathbf{Q}^X_{AB}$ in order to value their respective residual risk. This agreement enables the transaction.

Proof:
Let us denote by $\mathbf{Q}^X_{AB}$ the optimal additive measure for $\big(X^A_T + X^B_T, R^m_{AB}\big)$. In this case, $\mathbf{Q}^X_{AB} \in \partial R^m_{AB}(X^A_T + X^B_T)$. As mentioned in Subsection 1.2, the existence of such an additive measure is guaranteed as soon as the

penalty function is defined by Equation (10). This justifies the writing of the theorem in terms of additive measures rather than in terms of probability measures.
i) In the proof, we denote by $X \triangleq X_T^A + X_T^B$ and by Ψ^c, the centered random variable Ψ with respect to the given additive measure $\mathbf{Q}_{AB}^X$ optimal for $\big(X, R_{AB}\big)$: $\Psi^c = \Psi - \mathbb{E}_{\mathbf{Q}_{AB}^X}[\Psi]$. So, by definition,

$$\begin{aligned} -R_{AB}\big(X^c\big) &= \alpha_A\big(\mathbf{Q}_{AB}^X\big) + \alpha_B\big(\mathbf{Q}_{AB}^X\big) \\ &= \sup_F \big\{ -\rho_A\big(X^c - F^c\big)\big\} + \sup_F \big\{ -\rho_B\big(F^c\big)\big\} \\ &\geq -\inf_F \big\{\rho_A\big(X^c - F^c\big) + \rho_B\big(F^c\big)\big\} = -R_{AB}\big(X^c\big). \end{aligned}$$

In particular, all inequalities are equalities and

$$\sup_F \big\{ -\rho_A\big(X^c - F^c\big)\big\} + \sup_F \big\{ -\rho_B\big(F^c\big)\big\} = \sup_F \big\{ -\rho_A\big(X^c - F^c\big) - \rho_B\big(F^c\big)\big\}.$$

Hence, F^* is optimal for the inf-convolution problem, or equivalently for the program on the right-hand side of this equality, if and only if F^* is optimal for both problems $\sup_F \big\{-\rho_B\big(F^c\big)\big\}$ and $\sup_F \big\{-\rho_A\big(X^c - F^c\big)\big\}$. The second formulation is a straightforward application of Theorem 3.3 ii), considering the problem not at 0 but at $X_T^A + X_T^B$. $\square$

In order to obtain an explicit representation of an optimal structure F^*, some technical methods involving a localization of convex risk measures have to be used. This is the aim of the second part of this chapter, which is based upon some technical results on BSDEs. Therefore, before localizing convex risk measures and studying our optimal risk transfer in this new framework, we present in a separate section some quick recalls on BSDEs, which is essential for a good understanding of the second part on dynamic risk measures.

Part II: Dynamic Risk Measures

We now consider *dynamic convex risk measures*. Quite recently, many authors have studied dynamic version of static risk measures, focusing especially on the question of law invariance of these dynamic risk measures: among many other references, one may quote the papers by Cvitanic and Karatzas [35], Wang [111],Scandolo [107], Weber [112], Artzner et al. [3], Cheridito, Delbaen and Kupper [29] [30] or [31], Detlefsen and Scandolo [43], Frittelli and Gianin [58], Frittelli and Scandolo [59], Gianin [61], Riedel [102], Roorda, Schumacher and Engwerda [106] or the lecture notes of Peng [98]. Very recently, extending the work of El Karoui and Quenez [49], Klöppel and Schweizer have related dynamic indifference pricing and BSDEs in [78].

In this second part, we extend the axiomatic approach adopted in the static framework and introduce some additional axioms for the risk measures to be time-consistent. We then relate the dynamic version of convex risk measures to BSDEs. The associated dynamic risk measure is called g-conditional risk measure, where g is the BSDE coefficient. We will see how the properties of both the risk measure and the coefficient g are intimately connected. In particular, one of the key axioms in the characterization of the dynamic convex risk measure will be the translation invariance, as we will see in Section 6, and this will impose the g-coefficient of the related BSDE to depend only on z.

In the last two sections, we come back to the essential point of this chapter, the optimal risk transfer problem. We first derive some results on the inf-convolution of dynamic convex risk measures and obtain the optimal structure as a solution to the inf-convolution problem.

The idea behind our approach is to find a trade-off between static and very abstract risk measures as to obtain tractable risk measures. Therefore, we are more interested in tractability issues and interpretations of the dynamic risk measures we obtain rather than the ultimate general results in BSDEs.

5 Some recalls on Backward Stochastic Differential Equations

In the rest of the chapter, we take into account more information on the risk structure. In particular, we assume the σ-field $\mathcal{F}$ generated by a d-dimensional Brownian motion between $[0,T]$. Since any bounded $\mathcal{F}_T$-measurable variable is an stochastic integral w.r. to the Brownian motion, the risk measures of interest have to be robust with respect of this localization principle. To do that, we consider a family of risk measures described by backward stochastic differential equations (BSDE).
In this section, we introduce general BSDEs, defining them, recalling some key results on existence and uniqueness of a solution and presenting the comparison theorem. Complete proofs and additional useful results are given in the Chapter dedicated to BSDEs.

5.1 General Framework and Definition

Let $\big(\Omega,\mathcal{F},\mathbb{P}\big)$ be a probability space on which is defined a d-dimensional Brownian motion $W:=(W_t;t\leq T_H)$, where $T_H>0$ is the time horizon of the study. Let us consider the natural Brownian filtration $\mathcal{F}^0_t=\sigma\big(W_s;0\leq s\leq t;t\geq 0\big)$ and $(\mathcal{F}_t;t\leq T_H)$ its completion with the $\mathbb{P}$-null sets of $\mathcal{F}$.
Denoting by $\mathbb{E}$ the expected value with respect to $\mathbb{P}$, we introduce the following spaces which will be important in the formal setting of BSDEs. Since the time horizon may be sometimes modified, the definitions are referring to a generic time $T\leq T_H$.
• $L^2_n\big(\mathcal{F}_t\big)=\{\eta\,:\,\mathcal{F}_t-\text{measurable}\,\mathbb{R}^n-\text{valued random variable s.t. }\mathbb{E}(|\eta|^2)<\infty\}$.
• $\mathcal{P}_n(0,T)=\{(\phi_t;0\leq t\leq T)\,:\,\text{progressively measurable process with values in }\ \mathbb{R}^n\}$
• $\mathcal{S}^2_n(0,T)=\{(\phi_t;0\leq t\leq T)\,:\,\phi\in\mathcal{P}_n\,\text{s.t. }\mathbb{E}[\sup_{t\leq T}|Y_t|^2]<\infty\}$.
• $\mathcal{H}^2_n(0,T)=\{(\phi_t;0\leq t\leq T):\ \phi\in\mathcal{P}_n\ \text{s.t. }\mathbb{E}[\int_0^T|Z_s|^2ds]<\infty\}$.
• $\mathcal{H}^1_n(0,T)=\{(\phi_t;0\leq t\leq T):\ \phi\in\mathcal{P}_n\ \text{s.t. }\mathbb{E}[(\int_0^T|Z_s|^2ds)^{1/2}]<\infty\}$.

Let us give the definition of the one-dimensional BSDE; the multidimensional case is considered in the book's chapter dedicated to BSDEs.

Definition 5.1 *Let $\xi_T\in L^2(\Omega,\mathcal{F}_T,\mathbb{P})$ be a $\mathbb{R}$-valued terminal condition and g a coefficient $\mathcal{P}_1\otimes\mathcal{B}(\mathbb{R})\otimes\mathcal{B}(\mathbb{R}^d)$-measurable. A solution for the BSDE associated with (g,ξ_T) is a pair of progressively measurable processes $(Y_t,Z_t)_{t\leq T}$, with values in $\mathbb{R}\times\mathbb{R}^{1\times d}$ such that:*

$$\left\{\begin{array}{l}(Y_t)\in\mathcal{S}^2_1(0,T),\ (Z_t)\in\mathcal{H}^2_{1\times d}(0,T)\\ Y_t=\xi_T+\int_t^T g(s,Y_s,Z_s)ds-\int_t^T Z_s dW_s,\ 0\leq t\leq T\,.\end{array}\right.\tag{31}$$

The following differential form is also useful

$$-dY_t = g(t, Y_t, Z_t)dt - Z_t dW_t, \quad Y_T = \xi_T. \tag{32}$$

Conventional notation: To simply the writing of the BSDE, we adopt the following notations: the Brownian motion W is described as a column vector $(d, 1)$ and the Z vector is described as a row vector $(1, d)$ such that the notation ZdW has to be understood as a matrix product with (1,1)-dimension.

Remark 5.2 *If ξ_T and $g(t, y, z)$ are deterministic, then $Z_t \equiv 0$, and (Y_t) is the solution of ODE*

$$\frac{dy_t}{dt} = -g(t, y_t, 0), \qquad y_T = \xi_T\,.$$

If the final condition ξ_T is random, the previous solution is $\mathcal{F}_T$-measurable, and so non adapted. So we need to introduce the martingale $\int_0^t Z_s dW_s$ as a control process to obtain an adapted solution.

5.2 Some Key Results on BSDEs

Before presenting key results of BSDEs, we first summarize the results concerning the existence and uniqueness of a solution. The proofs are given in the Chapter dedicated to BSDEs with some complementary results.

5.2.1 Existence and Uniqueness Results

In the following, we always assume the necessary condition on the terminal condition $\xi_T \in L^2(\Omega, \mathcal{F}_T, \mathbb{P})$.

1. (H1): The standard case (uniformly Lipschitz): $(g(t, 0, 0); 0 \leq t \leq T)$ belongs to $\mathcal{H}^2(0, T)$ and g uniformly Lipschitz continuous with respect to (y, z), *i.e.* there exists a constant $C \geq 0$ such that

$$d\mathbb{P} \times dt - a.s. \quad \forall (y, y', z, z') \quad |g(\omega, t, y, z) - g(\omega, t, y', z')| \leq C(|y - y'| + |z - z'|).$$

   Under these assumptions, Pardoux and Peng [95] proved in 1990 the existence and uniqueness of a solution.

2. (H2) The continuous case with linear growth: there exists a constant $C \geq 0$ such that

$$d\mathbb{P} \times dt - a.s. \quad \forall (y, z) \quad |g(\omega, t, y, z)| \leq k(1 + |y| + |z|).$$

   Moreover we assume that $d\mathbb{P} \times dt$ *a.s.*, $g(\omega, t, ., .)$ is continuous in (y, z). Then, there exist a maximal and a minimal solutions (for a precise definition, please refer to the Chapter dedicated to BSDEs), as proved by Lepeltier and San Martin in 1998 [86].

3. (H3) The continuous case with quadratic growth in z: In this case, the assumption of square integrability on the solution is too strong. So we only consider bounded solution and obviously terminal condition $\xi_T \in L_\infty$. We also suppose that there exists a constant $k \geq 0$ such that

$$d\mathbb{P} \times dt - a.s. \quad \forall (y, z) \quad |g(\omega, t, y, z)| \leq k(1 + |y| + |z|^2).$$

Moreover we assume that $d\mathbb{P} \times dt - a.s.$, $g(\omega, t, ., .)$ is continuous in (y, z).

Then there exist a maximal and a minimal bounded solutions as first proved by Kobylansky [79] in 2000 and extended by Lepeltier and San Martin [86] in 1998. The uniqueness of the solution was proved by Kobylansky [79] under the additional conditions that the coefficient g is differentiable in (y, z) on a compact interval $[-K, K] \times \mathbb{R}^d$ and that there exists $c_1 > 0$ and $c_2 > 0$ such that:

$$\frac{\partial g}{\partial z} \leq c_1(1 + |z|), \qquad \frac{\partial g}{\partial y} \leq c_2(1 + |z|^2) \tag{33}$$

5.2.2 Comparison Theorem

We first present an important tool in the study of one-dimensional BSDEs: the so-called *comparison theorem.* It is the equivalent of the maximum principle when working with PDEs.

Theorem 5.3 (Comparison Theorem) *Let (ξ_T^1, g^1) and (ξ_T^2, g^2) be two pairs (terminal condition, coefficient) satisfying one of the above conditions (H1,H2,H3) (but the same for both pairs). Let (Y^1, Z^1) and (Y^2, Z^2) be the maximal associated solutions.*

(i) We assume that $\xi_T^1 \leq \xi_T^2$, $\mathbb{P} - a.s.$ and that $d\mathbb{P} \times dt - a.s.$ $\forall (y, z)$ $g^1(\omega, t, y, z) \leq g^2(\omega, t, y, z)$. Then we have

$$Y_t^1 \leq Y_t^2 \;\; a.s. \, \forall \, t \in [0, T]$$

(ii) **Strict inequality** *Moreover, under (H1), if in addition $Y_t^1 = Y_t^2$ on $B \in \mathcal{F}_t$, then a strict version of this result holds as*

$$a.s. \text{ on } B \quad \xi_T^1 = \xi_T^2 \,, \; \forall s \geq t, \; Y_s^1 = Y_s^2 \quad \text{and} \quad g^1(s, Y_s^1, Z_s^1) = g^2(s, Y_s^2, Z_s^2) \quad d\mathbb{P} \times ds - a.s. \text{ on } B \times [t, T]$$

6 Axiomatic Approach and g-Conditional Risk Measures

In this section, we give a general axiomatic approach for dynamic convex risk measures and see how they are connected to the existing notions of consistent convex price systems and non-linear expectations, respectively introduced by El Karoui and Quenez [50] and Peng [96]. Then, we relate the dynamic risk measures with BSDEs and focus on the properties of the solution of some particular BSDEs associated with a convex coefficient g, called g-conditional risk measures.

6.1 Axiomatic Approach

Following the study of static risk measures by Föllmer and Schied [53] and [54], we now propose a common axiomatic approach to dynamic convex risk measures, non-linear expectations and convex price systems and non-linear.

Definition 6.1 *Let $\big(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t; t \geq 0)\big)$ be a filtered probability space. A dynamic L^2-operator (L^∞-operator) $\mathcal{Y}$ with respect to $\big(\mathcal{F}_t; t \geq 0\big)$ is a family of continuous semi-martingales which maps, for any bounded stopping time T, a $L^2(\mathcal{F}_T)$ (resp. $L^\infty(\mathcal{F}_T)$) -variable ξ_T onto a process $\big(\mathcal{Y}_t(\xi_T); t \in [0, T]\big)$. Such an operator is said to be*

1. (P1) CONVEX*: For any stopping times $S \leq T$, for any (ξ_T^1, ξ_T^2), for any $0 \leq \lambda \leq 1$,*

$$\mathcal{Y}_S(\lambda\xi_T^1 + (1-\lambda)\xi_T^2) \leq \lambda\mathcal{Y}_S(\xi_T^1) + (1-\lambda)\mathcal{Y}_S(\xi_T^2) \quad \mathbb{P}-a.s.$$

2. (P2) MONOTONIC*: For any stopping times $S \leq T$, for any (ξ_T^1, ξ_T^2) such that $\xi_T^1 \geq \xi_T^2$ a.s.,*

   (P2+)*: the operator is increasing if* $\mathcal{Y}_S(\xi_T^1) \geq \mathcal{Y}_S(\xi_T^2)$ *a.s.*

   (P2-)*: the operator is decreasing* $\mathcal{Y}_S(\xi_T^1) \leq \mathcal{Y}_S(\xi_T^2)$ *a.s.*

3. (P3) TRANSLATION INVARIANT*: For any stopping times $S \leq T$ and any $\eta_S \in \mathcal{F}_S$, for any ξ_T,*

   (P3+) $\mathcal{Y}_S(\xi_T + \eta_S) = \mathcal{Y}_S(\xi_T) - \eta_S$ *a.s.*, (P3-) $\mathcal{Y}_S(\xi_T + \eta_S) = \mathcal{Y}_S(\xi_T) - \eta_S$ *a.s.*

4. (P4) TIME-CONSISTENT*: For $S \leq T \leq U$ three bounded stopping times, for any ξ_U*

   (P4+) $\mathcal{Y}_S(\xi_U) = \mathcal{Y}_S(\mathcal{Y}_T(\xi_U))$ *a.s.*, (P4-) $\mathcal{Y}_S(\xi_U) = \mathcal{Y}_S(-\mathcal{Y}_T(\xi_U))$ *a.s.*

5. (P5) ARBITRAGE-FREE*: For any stopping times $S \leq T$, and for any (ξ_T^1, ξ_T^2) such that $\xi_T^1 \geq \xi_T^2$,*

$$\mathcal{Y}_S(\xi_T^1) = \mathcal{Y}_S(\xi_T^2) \quad \text{on} \quad A_S = \{S < T\} \Longrightarrow \xi_T^1 = \xi_T^2 \quad a.s. \text{ on } A_S.$$

6. (P6) CONDITIONALLY INVARIANT*: For any stopping times $S \leq T$ and any $B \in \mathcal{F}_S$, for any ξ_T,*

$$\mathcal{Y}_S(\mathbf{1}_B\,\xi_T) = \mathbf{1}_B\mathcal{Y}_S(\xi_T) \quad a.s.$$

7. (P7) POSITIVE HOMOGENEOUS*: For any stopping times $S \leq T$, for any $\lambda_S \geq 0$ ($\lambda_S \in \mathcal{F}_S$) and for any ξ_T,*

$$\mathcal{Y}_S(\lambda_S\xi_T) = \lambda_S\mathcal{Y}_S(\xi_T) \quad a.s.$$

First, note that the property (P5) of no-arbitrage implies that the monotonicity property (P2) is strict.

Most axioms have two different versions, depending on the sign involved. Making such a distinction is completely coherent with the previous observations in the static part of this chapter about the relationship between price and risk measure: since the opposite of a risk measure is a price, the axioms with a "+" sign are related to the characterization of a price system, while the axioms with a "$-$" sign are related to that of a dynamic risk measure.

In [50], when studying pricing problems under constraints, El Karoui and Quenez defined a *consistent convex (forward) price system* as a convex (P1), increasing (P2+), time-consistency (P4+) dynamic operator $\mathcal{P}_t$, without arbitrage (P5). Time-consistency (P4+) may be view as a dynamic programming principle.

At the same period, Peng introduced the notion of *non-linear expectation* as a **translation invariance** (P3+) convex price system, satisfying the conditional invariance property (P6) which is very intuitive in this framework, (see for instance, Peng [96]). Note that (P6) of conditional invariance implies some additional assumptions on the operator $\mathcal{Y}$: in particular for any t, $\mathcal{Y}_t(0) = 0$. In the following, we denote the non-linear expectation by $\mathcal{E}$.

Now on, we focus on dynamic convex risk measures, where now only the properties with the "$-$" sign hold.

Definition 6.2 *A dynamic operator satisfying the axioms of convexity (P1), decreasing monotonicity (P2-), translation invariance (P3-), time-consistency (P4-) and arbitrage-free (P5) is said to be a dynamic convex risk measure. It will be denoted by $\mathcal{R}$ in the following.*
If $\mathcal{R}$ also satisfies the positive homogeneity property (P7), then it is called a dynamic coherent risk measure.

Note that a non-linear expectation defines a dynamic risk measure conditionally invariant and centered.

Remark 6.3 *It is not obvious to find a negligible set $\mathcal{N}$ such that for any bounded stopping time S and any bounded ξ_T, $\forall\,\omega \notin \mathcal{N}$, $\xi_T \to \mathcal{R}^g_S(\omega,\xi_T)$ is a static convex risk measure. The negligible sets may depend on the variable ξ_T itself.*

Dynamic Entropic Risk Measure A typical example is the *dynamic entropic risk measure* , obtained by conditioning the static entropic risk measure. For any ξ_T bounded:

$$e_\gamma(\xi_T) = \gamma \ln \mathbb{E}\big[\exp(-\frac{1}{\gamma}\xi_T)\big] \qquad \Rightarrow \qquad e_{\gamma,t}(\xi_T) = \gamma \ln \mathbb{E}\big[\exp(-\frac{1}{\gamma}\xi_T)|\mathcal{F}_t\big].$$

Since, ξ_T is bounded, $e_{\gamma,t}(\xi_T)$ is bounded for any t. Therefore, this dynamic operator defined on L^∞ satisfies the properties of dynamic convex risk measures. Convexity, decreasing monotonicity, translation invariance, no-arbitrage are obvious; the time-consistency property $(P4-)$ results from the transitivity of conditional expectation:

$$\forall t \geq 0\,,\ \forall h > 0, \quad e_{\gamma,t}\big(\xi_T\big) = e_{\gamma,t}\big(-e_{\gamma,t+h}\big(\xi_T\big)\big) \quad a.s.$$

We give the easy proof of this identity to help the reader to understand the $(-)$ sign in the formula.
Proof:

$$\begin{aligned}
Y_{t+h} &\equiv e_{\gamma,t+h}\big(\xi_T\big) = \gamma \ln \mathbb{E}\big[\exp(-\tfrac{1}{\gamma}\xi_T)|\mathcal{F}_{t+h}\big]\\
e_{\gamma,t}\big(-e_{\gamma,t+h}\big(\xi_T\big)\big) &= \gamma \ln \mathbb{E}\big[\exp(-\tfrac{1}{\gamma}(-Y_{t+h}))|\mathcal{F}_t\big]\\
&= \gamma \ln \mathbb{E}\Big[\exp(\tfrac{1}{\gamma}\gamma \ln \mathbb{E}\big[\exp(-\tfrac{1}{\gamma}\xi_T)|\mathcal{F}_{t+h}\big])|\mathcal{F}_t\Big]\\
&= \gamma \ln \mathbb{E}\Big[\mathbb{E}\big[\exp(-\tfrac{1}{\gamma}\xi_T)|\mathcal{F}_t\big]\Big]. \qquad \square
\end{aligned}$$

Moreover, it is possible to relate the dynamic entropic risk measure $e_{\gamma,t}$ with the solution of a BSDE, as follows:

Proposition 6.4 *The dynamic entropic measure $\big(e_{\gamma,t}(\xi_T); t \in [0,T]\big)$ is solution of the following BSDE with the quadratic coefficient $g\big(t,z\big) = \frac{1}{2\gamma}\big\|z\big\|^2$ and terminal bounded condition ξ_T.*

$$-de_{\gamma,t}\big(\xi_T\big) = \frac{1}{2\gamma}\big\|Z_t\big\|^2 dt - Z_t dW_t \qquad e_{\gamma,T}\big(\xi_T\big) = -\xi_T. \tag{34}$$

Proof: Let us denote by $M_t(\xi_T) = \mathbb{E}\Big[\exp\big(-\frac{1}{\gamma}\xi_T\big)|\mathcal{F}_t\Big]$. As M is a positive and bounded continuous martingale, one can use the multiplicative decomposition to get $dM_t = \frac{1}{\gamma}M_t\big(Z_t\,dW_t\big)$ where $\big(Z_t; t \geq 0\big)$ is a $1 \times d$ dimensional square-integrable process. By Itô's formula applied to the function $\gamma\ln(x)$, we obtain the Equation (34).
Note that the conditional expectation of the quadratic variation $\mathbb{E}\big[\int_t^T |Z_s|^2 ds|\mathcal{F}_t\big] = \mathbb{E}\big[e_{\gamma,t}(\xi_T) - \xi_T|\mathcal{F}_t\big]$ is bounded and conversely if the Equation (34) has the solution (Y,Z) such that Y_T and $\mathbb{E}\big[\int_t^T |Z_s|^2 ds|\mathcal{F}_t\big]$ are

bounded, then Y is bounded. This point will be detailed in Theorem 7.4. $\square$

This relationship between the dynamic entropic risk measure and BSDE can be extended to general dynamic convex risk measures as we will see in the rest of this section.

6.2 Dynamic Convex Risk Measures and BSDEs

This section is about the relationship between dynamic convex risk measures and BSDEs. More precisely, we are interested in the correspondence between the properties of the "BSDE" operator and that of the coefficient.

We consider the dynamic operator generated by the maximal solution of a BSDE:

Definition 6.5 *Let g be a standard coefficient. The g-dynamic operator, denoted by $\mathcal{Y}^g$, is such that $\mathcal{Y}^g_t(\xi_T)$ is the maximal solution of the BSDE(g,ξ_T).*

As a consequence, the adopted point of view is different from that of the section dedicated to recalls on BSDEs where the terminal condition of the BSDE was fixed.

It is easy to deduce properties of the g-dynamic operator from those of the coefficient g. The converse is more complex and this study has been initiated by Peng when considering g-expectations ([96]). Our characterization is based upon the following lemma:

Lemma 6.6 (Coefficient Uniqueness) *Let g^1 and g^2 be two regular coefficients, such that uniqueness of solution for the BSDE(g^1) holds. Let $\mathcal{Y}^{g^i}$ be g^i-dynamic operator($i=1,2$). Assume that*

$$\forall(T,\xi_T),\quad d\mathbb{P}\times dt-a.s.\qquad \mathcal{Y}^{g^1}_t(\xi_T)=\mathcal{Y}^{g^2}_t(\xi_T).$$

a) If the coefficients g^1 and g^2 simply depend on t and z, then $d\mathbb{P}\times dt\ -a.s.\ \ \forall z\ \ g^1(t,z)=g^2(t,z)$.

b) In the general case, the same identity holds provided the coefficients are continuous w.r. to t.

$$d\mathbb{P}\times dt\ -a.s.\quad \forall(y,z)\quad g^1(t,y,z)=g^2(t,y,z).$$

Proof: Suppose that both coefficients g^1 and g^2 generate the same solution Y (but a priori different processes Z^1 and Z^2) for the BSDEs (g^1,ξ_T) and (g^2,ξ_T), for **any** ξ_T in the appropriate space (L^2 or L^∞). Given the uniqueness of the decomposition of the semimartingale Y, the martingale parts and the finite variation processes of the both decompositions of Y are indistinguishable. In particular, $\int_0^t Z^1_s dW_s=\int_0^t Z^2_s dW_s=\int_0^t Z_s dW_s$, $a.s.$ and $\int_0^t g^1(s,Y_s,Z^1_s)ds=\int_0^t g^2(s,Y_s,Z^2_s)ds$. Therefore, $\int_0^t g^1(s,Y_s,Z_s)ds=\int_0^t g^2(s,Y_s,Z_s)ds$. A priori, these equalities only hold for processes (Y,Z) obtained through BSDEs.

a) Assume that g^1 and g^2 do not depend on y. Given a bounded adapted process Z, we consider the following locally bounded semimartingale U as $dU_t=g^1(t,Z_t)dt-Z_tdW_t$; $U_0=u_0$. (U,Z) is the solution of the BSDE$(g^1,U_{T\wedge\tau})$ where τ is a stopping time s.t. $U_{T\wedge\tau}$ is bounded. By uniqueness, $\int_0^{t\wedge\tau} g^1(s,Z_s)ds=\int_0^{t\wedge\tau} g^2(s,Z_s)ds$. As shown in b) below, this equality implies that $g^1(s,Z_s)=g^2(s,Z_s)$, $a.s.\,ds\times d\mathbb{P}$. Thanks to the continuity of g^1 and g^2 w.r. to z, we can only consider denumerable rational z to show that, with $dt\times d\mathbb{P}$ probability one, for any z, $g^1(s,z)=g^2(s,z)$.

b1) In the general case, given a bounded process Z, we consider **a** solution of the following forward stochastic differential equation Y as $dY_t = g^1(t, Y_t, Z_t)dt - Z_t dW_t$; $Y_0 = y_0$ and the stopping time τ_N defined as the first time, when $|Y|$ crosses the level N.
The pair of processes $(Y_{\tau_N \wedge t}, Z_t \mathbf{1}_{]0,\tau_N]}(t))$ is solution of the BSDE with bounded terminal condition $\xi_T = Y_{\tau_N \wedge T}$. Thanks to the previous observation, the pair of processes $(Y_{\tau_N \wedge t}, Z_t \mathbf{1}_{]0,\tau_N]}(t))$ is also solution of the BSDE$(g^2, Y_{\tau_N \wedge T})$. Hence, both processes $\int_0^t g^1(s, Y_s, Z_s)ds = \int_0^t g^2(s, Y_s, Z_s)ds$ are indistinguishable on $]0, \tau_N \wedge T]$. Since τ_N goes to infinity with N, the equality holds at any time, for any bounded process Z.
b2) Assume $g^1(s, y, z)$ and $g^2(s, y, z)$ continuous w.r. to s. Let z be a given vector. Let Y^z_{t+h} be a forward perturbation of a general solution Y, at the level z between t and $t + h$,

$$Y^z_u = Y_t + \int_t^u g^1(s, Y^z_s, z)ds - \int_t^u z dW_s \qquad \forall u \in [t, t + h].$$

By assumption, (Y^z_u, z) is also solution of the BSDE(g^2, Y^z_{t+h}), for $u \in [t, t + h]$, and $\int_t^u g^1(s, Y^z_s, z)ds = \int_t^t g^2(s, Y^z_s, z)ds$. Hence, by continuity, $\frac{1}{h}\mathbb{E}\Big[Y^{z,i}_{t+h} - Y_t|\mathcal{F}_t\Big]$ goes in L^1 to $g^i(t, Y_t, z)$ $(i = 1, 2)$ with $h \to 0$. Then $g^1(t, Y_t, z) = g^2(t, Y_t, z)$ for any solution Y_t of the BSDE, i.e. for any v.a $\mathcal{F}_t$-measurable. $\square$

Comments Peng ([96] and [98]) and Briand et al. [21] have been among the first to look at the dynamic operators to deduce local properties through the coefficient g of the associated BSDE, when considering non-linear expectations. More recently, Jiang has considered the applications of g-expectations in finance in his PhD thesis [74].
ii) In [21], Briand et al. proved a more accurate result for g Lipschitz. More precisely, let g be a standard coefficient such that $\mathbb{P}$-a.s., $t \longmapsto g(t, y, z)$ is continuous and $g(t, 0, 0) \in \mathcal{S}^2$. Let us fix $(t, y, z) \in [0, T] \times \mathbb{R} \times \mathbb{R}^d$ and consider for each $n \in \mathbb{N}^*$, $\{(Y^n_s, Z^n_s); s \in [t, t_n = t + \frac{1}{n}]\}$ solution of the BSDE (g, X_n) where the terminal condition X_{t_n} is given by $X_{t_n} = y + z\big(W_{t_n} - W_t\big)$. Then for each $(t, y, z) \in [0, T] \times \mathbb{R} \times \mathbb{R}^d$, we have

$$L^2 - \lim_{n\to\infty} n\big(Y^n_t - y\big) = g(t, y, z).$$

Some properties automatically hold for the dynamic operator $\mathcal{Y}^g$ simply because it is the maximal solution of a BSDE. Some others can be obtained by imposing conditions on the coefficient g:

Theorem 6.7 *Let $\mathcal{Y}^g$ be the g-dynamic operator.*
a) Then, $\mathcal{Y}^g$ is increasing monotonic (P2+), time-consistent (P4+) and arbitrage-free (P5).
b) Moreover, under the assumptions of Lemma 6.6,

1. *$\mathcal{Y}^g$ is conditionally invariant (P6) if and only if for any $t \in [0, T], z \in \mathbb{R}^n$, $g(t, 0, 0) = 0$.*
2. *$\mathcal{Y}^g$ is translation invariant (P3+) if and only if g does not depend on y.*
3. *$\mathcal{Y}^g$ is homogeneous if and only if g is homogeneous;*

c) For properties related to the order, the following implications simply hold:

1. *If g is convex, then $\mathcal{Y}^g$ is convex (P1).*
2. *If $g^1 \leq g^2$, then $\mathcal{Y}^{g^1} \leq \mathcal{Y}^{g^2}$.*

Therefore, if g is a convex coefficient depending only on z, $\mathcal{R}^g(\xi_T) \equiv \mathcal{Y}^g(-\xi_T)$ is a dynamic convex risk measure, called g*-conditional risk measure.*

Note that $\mathcal{Y}^g$ is a consistent convex price system and moreover, if for any $t \in [0,T]$, $g(t,0)=0$, then $\mathcal{Y}^g$ is a non-linear expectation, called g-expectation.

Proof: a) • The strict version of the comparison Theorem 5.3 leads immediately to both properties (P2+) and (P5).

• Up to now, we have defined and considered BSDEs with a terminal condition at a fixed given time T. It is always possible to consider it as a BSDE with a time horizon $T_H \geq T$, even if T_H is a bounded stopping time. Obviously, the coefficient g has to be extended as $g\mathbf{1}_{[0,T]}$ and the terminal condition $\xi_{T_H} = \xi_T$. Therefore the solution Y_t is constant on $[T, T_H]$.

To obtain the *time-consistency* property (P4), also called the flow property, we consider three bounded stopping times $S \leq T \leq U$ and write the solution of the BSDEs as function of the terminal date. With obvious notations, we want to prove that $Y_S(T, Y_T(U,\xi_U)) = Y_S(U,\xi_U)$ *a.s.*.

By simply noticing that:

$$\begin{aligned} Y_S(T, Y_T(U,\xi_U)) = Y_T(U,\xi_U) + \quad & \textstyle\int_S^T g(t,Z_t)dt - \int_S^T Z_t dW_t \\ = \xi_U + \textstyle\int_T^U g(t,Z_t)dt - \int_T^U Z_t dW_t \quad & + \textstyle\int_S^T g(t,Z_t)dt - \int_S^T Z_t dW_t, \end{aligned}$$

the process which is defined as $Y_t(T, Y_T(U,\xi_U))$ on $[0,T]$ and by $Y_t(U,\xi_U)$ on $]T,U]$ is the maximal solution of the BSDE (g, ξ_U, U). Uniqueness of the maximal solution implies (P4).

b) The three properties b1), b2) and b3) involve the same type of arguments to be proved, so we simply present the proof for b2).

Let $g_m(t,y,z) = g(t,y+m,z)$. We simply note that $Y^m_{.} = Y_{.}(\xi_T + m) - m$ is the maximal solution of the BSDE (g_m, ξ_T). The translation invariance property is equivalent to the indistinguishability of both processes Y and Y^m; by the uniqueness Lemma 6.6, this property is equivalent to the identity

$$g(t,y,z) = g_m(t,y,z) = g(t,y+m,z) \qquad a.s.$$

This implies that g does not depend on y.

c) • c1) For the *convexity* property, we consider different BSDEs: (Y^1_t, Z^1_t) is the (maximal) solution of (g, ξ^1_T) and (Y^2_t, Z^2_t) is the (maximal) solution of (g, ξ^2_T). Then, we look at $\widetilde{Y}_t = \lambda Y^1_t + (1-\lambda) Y^2_t$, with $\lambda \in [0,1]$. We have:

$$-d\widetilde{Y}_t = (\lambda g(t, Y^1_t, Z^1_t) + (1-\lambda) g(t, Y^2_t, Z^2_t))dt - (\lambda Z^1_t + (1-\lambda) Z^2_t) dW_t \quad ; \quad \widetilde{Y}_T = \lambda \xi^1_T + (1-\lambda)\xi^2_T.$$

Since g is convex, we can rewrite this BSDE as:

$$-d\widetilde{Y}_t = (g(t, \widetilde{Y}_t, \widetilde{Z}_t) + \alpha(t, Y^1_t, Y^2_t, Z^1_t, Z^2_t, \lambda))dt - \widetilde{Z}_t dW_t$$

where α is a a.s. non-negative process. Hence, using the comparison theorem, the solution $\widetilde{Y}_t$ of this BSDE is for any $t \in [0,T]$ a.s. greater than the solution Y_t of the BSDE $(g, \lambda\xi^1_T + (1-\lambda)\xi^2_T)$. It is a super-solution

in the sense of Definition 2.1 of El Karoui, Peng and Quenez [47].
• c2) is a direct consequence of the comparison Theorem 5.3. □

Some additional comments on the relationship between BSDE and dynamic operators Since 1995, Peng has focused on finding conditions on dynamic operators so that they are linear growth g-expectations. This difficult problem is solved in particular for dynamic operators satisfying a domination assumption introduced by Peng [96] in 1997 where $b_k(z) = k|z|$. For more details, please refer to his lecture notes on BSDEs and dynamic operators [98].

Theorem 6.8 *Let $(\mathcal{E}_t; 0 \leq t \leq T)$ be a non-linear expectation such that:*
There exists $|\lambda| \in \mathcal{H}^2$ and a sufficiently large real number $k > 0$ such that for any $t \in [0,T]$ and any $\xi_T \in L^2(\mathcal{F}_T)$:

$$\mathcal{E}_t^{-b_k+|\lambda|}(\xi_T) \leq \mathcal{E}_t(\xi_T) \leq \mathcal{E}_t^{b_k+|\lambda|}(\xi_T) \quad a.s.$$

and for any $(\xi_T^1, \xi_T^2) \in L^2(\mathcal{F}_T)$: $\mathcal{E}_t(\xi_T^1) - \mathcal{E}_t(\xi_T^2) \leq \mathcal{E}_t^{b_k}(\xi_T^1 - \xi_T^2)$ Then, there exists a function $g(t,y,z)$ satisfying assumption (H1) such that for any $t \in [0,T]$,

$$\forall\, \xi_T \in L^2(\mathcal{F}_T), \quad a.s., \quad \forall t, \quad \mathcal{E}_t(\xi_T) = \mathcal{E}_t^g(\xi_T)$$

For a proof of this theorem, please refer to Peng [97].

Infinitesimal Risk Management The coefficient of any g-conditional risk measure $\mathcal{R}^g$ can be naturally interpreted as the *infinitesimal risk measure* over a time interval $[t, t+dt]$ as:

$$\mathbb{E}_{\mathbb{P}}[d\mathcal{R}_t^g | \mathcal{F}_t] = -g(t, Z_t)dt,$$

where Z_t is the local volatility of the $g$conditional risk measure.
Therefore, choosing carefully the coefficient g enables to generate g-conditional risk measures that are locally compatible with the views and practice of the different agents in the market. In other words, knowing the infinitesimal measure of risk used by the agents is enough to generate a dynamic risk measure, locally compatible. In this sense, the g-conditional risk measure may appear more tractable than static risk measures. The following example gives a good intuition of this idea: the g-conditional risk measure corresponding to the mean-variance paradigm has a g-coefficient of the type $g(t,z) = -\lambda_t z + \frac{1}{2}z^2$. The process λ_t can be interpreted as the correlation with the market numéraire.

Therefore, g-Conditional risk measures are a way to construct a wide family of convex risk measures on a probability space with Brownian filtration, taking into account the ability to decompose the risk through inter-temporal local risk measures $g(t, Z_t)$.

In the following, to study g-conditional risk measures, we adopt the same methodology as in the static framework. In particular, we start by developing a dual representation for these dynamic risk measures, in terms of the "dual function" of their coefficient. This study requires some general properties of convex functions on $\mathbb{R}^n$.

7 Dual Representation of g-Conditional Risk Measures

Following the approach adopted in the first part of this chapter when studying static risk measures, we now focus on a dual representation for g-conditional risk measures. The main tool is the *Legendre-Fenchel transform* G of the coefficient g, defined by:

$$G(t,\mu) = \sup_{z \in \mathbf{Q}^{\mathbf{n}}_{\text{rational}}} \left\{ \langle \mu, -z \rangle - g(t,z) \right\}. \tag{35}$$

The convex function G is also called the *polar function* or the *conjugate* of g. Provided that g is continuous,

$$g(t,z) = \sup_{\mu \mathbf{Q}^{\mathbf{n}}_{\text{rational}}} \big(\langle \mu, -z \rangle - G(t,\mu) \big). \tag{36}$$

More precisely,

Definition 7.1 *A g-conditional risk $\mathcal{R}^g$ measure is said to have a dual representation if there exists a set $\mathcal{A}$ of admissible controls such that for any bounded stopping time $S \leq T$ and any ξ_T in the appropriate space*

$$\mathcal{R}^g_S(\xi_T) = \operatorname{ess\,sup}_{\mu \in \mathcal{A}} \mathbb{E}_{\mathbb{Q}^\mu} \Big[-\xi_T - \int_S^T G(t,\mu_t) dt \big| \mathcal{F}_S \Big] \tag{37}$$

where $\mathbb{Q}^\mu$ is a probability measure absolutely continuous with respect to $\mathbb{P}$.
The dual representation is said to be exact at $\bar{\mu}$ if the ess sup *is reached for $\bar{\mu}$.*

In order to obtain this representation, several intermediate steps are needed:

1. Refine results on Girsanov theorem and the integrability properties of martingales with respect to change of probability measures.

2. Refine results from convex analysis on the Legendre-Fenchel transform and the existence of an optimal control in both Formulae (35) and (36), including measurability properties,

The next paragraph gives a summary of the main results that are needed.

7.1 Girsanov Theorem and BMO-Martingales

Our main reference on Girsanov theorem and BMO-martingales is the book by Kazamaki [77]. The exponential martingale associated with the d-dimensional Brownian motion W, $\mathcal{E}(\int_0^t \mu_s dW_s) = \Gamma^\mu_t = \exp\Big(\int_0^t \mu_s dW_s - \frac{1}{2}|\mu_s|^2 ds \Big)$, solution of the forward stochastic equation

$$d\Gamma^\mu_t = \Gamma^\mu_t \mu^*_t dW_t \ , \quad \Gamma^\mu_0 = 1 \tag{38}$$

is a positive local martingale, if μ is an adapted process such that $\int_0^T |\mu_s|^2 ds < \infty$.
When Γ^μ is a uniformly integrable (u.i.) martingale, Γ^μ_T is the density (w.r. to $\mathbb{P}$) of a new probability measure denoted by $\mathbb{Q}^\mu$. Moreover, if W is a $\mathbb{P}$-Brownian motion, then $W^\mu_t = W_t - \int_0^t \mu_s ds$ is a $\mathbb{Q}^\mu$-Brownian motion.

Questions around Girsanov theorem are of two main types. They mainly consist of:

- first, finding conditions on μ so that Γ^μ is a u.i. martingale.
- second, giving so;e precision on the integrability properties that are preserved under the new probability measure.

The bounded case, that is recalled below, is well-known. The BMO case is less standard, so we give more details.

7.1.1 Change of Probability Measures with Bounded Coefficient

When μ is bounded, it is well-known that the exponential martingale belongs to all $\mathcal{H}^p$-spaces.
Moreover, if a process is in $\mathcal{H}^2(\mathbb{P})$, it is in $\mathcal{H}^{1+\epsilon}(\mathbb{Q}^\mu)$. In particular, if $M_t^Z = \int_0^t Z_s dW_s$ is a $\mathcal{H}^2(\mathbb{P})$-martingale, then $\widehat{M}_t^Z = \int_0^t Z_s dW_s^\mu$ is a u.i. martingale under $\mathbb{Q}^\mu$, with null $\mathbb{Q}^\mu$-expectation.

7.1.2 Change of Probability Measures with BMO-Martingale

The right extension of the space of bounded processes is the space of BMO processes defined as:

$$\mathrm{BMO}(\mathbb{P}) = \{\varphi \in \mathcal{H}^2 \quad s.t \quad \exists C \quad \forall t \quad \mathbb{E}\Big[\int_t^T |\varphi_s|^2 ds|\mathcal{F}_t\Big] \leq C \ a.s.\}$$

The smallest constant C such that the previous inequality holds is denoted by $C^* = ||\varphi||^2_{\mathrm{BMO}}$.
In terms of martingale, the stochastic integral $\int_0^t \varphi_s dW_s$ is said to be a BMO($\mathbb{P}$)-martingale if and only if the process φ belongs to BMO($\mathbb{P}$). The following deep result is proved in Kazamaki [77] (Section 3.3).

Theorem 7.2 *Let the adapted process μ be in* BMO($\mathbb{P}$). *Then*

1. *The exponential martingale Γ^μ is a u.i. martingale and defines a new equivalent probability measure $\mathbb{Q}^\mu$. Moreover, $W_t^\mu = W_t - \int_0^t \mu_s ds$ is a $\mathbb{Q}^\mu$-Brownian motion.*
2. *$M_t^\mu = \int_0^t \mu_s^* dW_s$, and more generally any* BMO($\mathbb{P}$)*-martingale $M_t^Z = \int_0^t Z_s dW_s$, are transformed into continuous processes $\widehat{M}_t^\mu = \int_0^t \mu_s^* dW_s^\mu$ and $\widehat{M}_t^Z = \int_0^t Z_s dW_s^\mu$ that are* BMO($\mathbb{Q}^\mu$)*-martingales.*
3. *The* BMO*-norms with respect to $\mathbb{P}$ and $\mathbb{Q}^\mu$ are equivalent:*

$$k||Z||_{\mathrm{BMO}(\mathbb{Q}^\mu)} \leq ||Z||_{\mathrm{BMO}(\mathbb{P})} \leq K||Z||_{\mathrm{BMO}(\mathbb{Q}^\mu)}.$$

The constants k and K only depend on the BMO*-norm of μ.*

Hu, Imkeller and Müller [73] were amongst the first to use the property that the martingale $dM_t^Z = Z_t dW_t$ which naturally appears in BSDEs associated with exponential hedging problems, is BMO. Since then, such a property has been used in different papers, mostly dealing with the question of dynamic hedging in an exponential utility framework (see for instance the recent papers by Mania, Santacroce and Tevzadze [88] and Mania and Schweizer [89]).
In the proposition below, we extend their results to general quadratic BSDEs.

Proposition 7.3 *Let (Y,Z) be the maximal solution of the quadratic (H3) BSDE with coefficient g, and $M^Z = \int_0^\cdot Z_s dW_s$ the stochastic integral Z.W*

$$dY_t = g(t, Z_t)dt - dM_t^Z, \quad Y_T = \xi_T.$$

Given that by assumption Y is bounded, and $|g(t,0)|^{1/2} \in \mathrm{BMO}(\mathbb{P})$, M^Z is a $\mathrm{BMO}(\mathbb{P})$-martingale

Proof: Let k be the constant such that $|g(t,z)| \leq |g(t,0)| + k|z|^2$.
Thanks to Itô's formula applied to the solution (Y,Z) and to the exponential function:

$$\begin{aligned}
\exp(\beta Y_t) &= \exp(\beta Y_T) + \beta \int_t^T \exp(\beta Y_s) g(s, Z_s) ds - \frac{\beta^2}{2} \int_t^T \exp(\beta Y_s)|Z_s|^2 ds \\
&- \beta \int_t^T \exp(\beta Y_s) Z_s dW_s \\
&= \exp(\beta Y_T) + \beta \int_t^T \exp(\beta Y_s)\Big(g(s, Z_s) - \frac{\beta}{2}|Z_s|^2\Big) ds - \beta \int_t^T \exp(\beta Y_s) Z_s dW_s.
\end{aligned}$$

Given that $\frac{\beta}{2}|Z_s|^2 - g(s,Z_s) \geq (\frac{\beta}{2} - k)|Z_s|^2 - |g(s,0)| \geq \varepsilon|Z_s|^2 - |g(s,0)|$ for $\beta \geq (k+\varepsilon)$ and taking the conditional expected value, we obtain:

$$\beta\,\varepsilon\,\mathbb{E}\Big[\int_t^T \exp(\beta Y_s)|Z_s|^2 ds|\mathcal{F}_t\Big] \leq C + \beta\mathbb{E}\Big[\int_t^T \exp(\beta Y_s)|g(s,0)|ds|\mathcal{F}_t\Big] \leq C$$

where C is a universal constant that may change from place to place. Since $\exp(\beta Y_s)$ is bounded both from below and from above, the property holds. $\square$

7.2 Some Results in Convex Analysis

Some key results in convex analysis are needed to obtain the dual representation of g-conditional risk measures. They are presented in the Appendix 9 to preserve the continuity of the arguments in this part. More details or proofs may be found in Aubin [4], Hiriart-Urruty and Lemaréchal [70] or Rockafellar [103].

7.3 Dual Representation of Risk Measures

We now study the dual representation of g-conditional risk measures. The space of admissible controls depends on the assumption imposed on the coefficient g. We consider successively both situations (H1) and (H3). There is no need to look separately at (H2), as, under our assumptions, the condition (H2) implies the condition (H1) (for more details, please refer to the Appendix 9.2.1). The (H1) case has been solved in [47] but the (H3) case is new.

Theorem 7.4 *Let g be a convex coefficient satisfying (H1) or (H3) and G be the associated polar process, $G(t,\mu) = \sup_{z \in \mathbf{Q}^{\mathbf{n}}_{rational}} \{\langle \mu, -z\rangle - g(t,z)\}$.*
i) For almost all (ω,t), the program $g(\omega,t,z) = \sup_{\mu \in \mathbf{Q}^{\mathbf{n}}_{rational}} [\langle \mu, -z\rangle - G(\omega,t,\mu)]$ has an optimal progressively measurable solution $\bar{\mu}(\omega,t)$ in the subdifferential of g at z, $\partial g(\omega,t,z)$.
ii) Then $\mathcal{R}^g$ has the following dual representation, exact at $\bar{\mu}$,

$$\mathcal{R}_t^g(\xi_T) = \mathrm{esssup}_{\mu \in \mathcal{A}} \mathbb{E}_{\mathbb{Q}^\mu}\Big[-\xi_T - \int_t^T G(s,\mu_s)ds|\mathcal{F}_t\Big] = \mathbb{E}_{\mathbb{Q}^{\bar{\mu}}}\Big[-\xi_T - \int_t^T G(s,\bar{\mu}_s)ds|\mathcal{F}_t\Big]$$

where:

1. *Under (H1) ($|g(t,z)| \leq |g(t,0)|| + k|z|$), $\mathcal{A}$ is the space of adapted processes μ bounded by k, and $\mathbb{Q}^\mu$ is the associated equivalent probability measure with density Γ^μ_T where Γ^μ is the exponential martingale defined in* (38).

2. *Under (H3), ($|g(t,z)| \leq |g(t,0)| + k|z|^2$), $\mathcal{A}$ is the space of BMO($\mathbb{P}$)-processes μ and $\mathbb{Q}^\mu$ is defined as above.*

iii) Let $g(t,.)$ be a strongly convex function (i.e. $g(t,z) - \frac{1}{2}C|z|^2$ is a convex function). Then the Fenchel-Legendre transform $G(t,\mu)$ has a quadratic growth in μ and the following dual representation holds true:

$$\mathbb{E}_{\mathbb{Q}^\mu}\Big[\int_t^T G(s,\mu_s)ds\Big|\mathcal{F}_t\Big] = \text{esssup}_{\xi_T}\mathbb{E}_{\mathbb{Q}^\mu}\big[\xi_T\big|\mathcal{F}_t\big] - \mathcal{R}^g_t(\xi_T) = \mathbb{E}_{\mathbb{Q}^\mu}\big[\bar{\xi}_T\big|\mathcal{F}_t\big] - \mathcal{R}^g_t(\bar{\xi}_T)$$

Proof: *i)* Since g is a proper function, the dual representation of g with its polar function G is exact at $\bar{\mu} \in \partial g(z)$:

$$g(t,z) = \sup_{\mu \in \mathbf{Q}^{\mathbf{n}}_{\text{rational}}} [\langle \mu, -z\rangle - G(t,\mu)\,] = \langle \bar{\mu}, -z\rangle - G(t,\bar{\mu}),$$

using classical results of convex analysis, recalled in the Appendix 9.
The measurability of $\bar{\mu}$ is separately studied in Lemma 7.5 just after this proof.

ii) a) Let us first consider a coefficient g with linear growth (H1); so, $g(t,0)$ is in $\mathcal{H}^2$. By definition, $-G(t,\mu_t)$ is dominated from above by the square integrable process $g(t,0)$. Then, let $\mathcal{R}^g_t(\xi_T) := Y_t$ be the solution of the BSDE $(g, -\xi_T)$,

$$-dY_t = g(t,Z_t)dt - Z_t dW_t = (g(t,Z_t) - \langle \mu_t, -Z_t\rangle)dt - Z_t dW^\mu_t, \quad Y_T = -\xi_T. \tag{39}$$

By Girsanov Theorem (Theorem 7.2), for $\mu \in \mathcal{A}$ the exponential martingale Γ^μ is u.i. and defines a probability measure $\mathbb{Q}^\mu$ on $\mathcal{F}_T$ such that the process $W^\mu = W - \int_0^\cdot \mu_s ds$ is a $\mathbb{Q}^\mu$-Brownian motion. Moreover, since $M^Z = \int_0^\cdot Z_s dW_s$ is in $\mathcal{H}^2(\mathbb{P})$, $\widehat{M}^Z = \int_0^\cdot Z_s dW^\mu_s$ is a u.i. $\mathbb{Q}^\mu$-martingale. Moreover, since μ is bounded and g uniformly Lipschitz, the process $(g(t,Z_t) - Z_t\mu_t)$ belongs to $\mathcal{H}^2(\mathbb{P})$ but also to $\mathcal{H}^{1+\epsilon}(\mathbb{Q}^\mu)$. So we can use an integral representation of the BSDE (39) in terms of

$$Y_t = \mathbb{E}_{\mathbb{Q}^\mu}\Big[-\xi_T + \int_t^T (g(s,Z_s) - \langle \mu_s, -Z_s\rangle)ds\Big|\mathcal{F}_t\Big] \geq \mathbb{E}_{\mathbb{Q}^\mu}\Big[-\xi_T - \int_t^T G(s,\mu_s)ds\Big|\mathcal{F}_t\Big]. \tag{40}$$

We do not need to prove that the last term is finite. It is enough to recall that $(-G(s,\mu_s))^+$ is dominated from above by the $d\mathbb{Q} \times ds$ integrable process $(g(s,0))^+$.

b) Let $\bar{\mu}$ be an optimal control, bounded by k, such that $g(t,Z_t) = \langle \bar{\mu}_t, -Z_t\rangle - G(t,\bar{\mu}_t)$ (see Lemma 7.5 for measurability results). Then the process $-G(t,\bar{\mu}_t)$ belongs to $\mathcal{H}^2(\mathbb{P})$ and so to $\mathcal{H}^{1+\epsilon}(\mathbb{Q}^{\bar{\mu}})$. By the previous result, $Y_t = \mathbb{E}_{\mathbb{Q}^{\bar{\mu}}}\Big[-\xi_T - \int_t^T G(s,\bar{\mu}_s)ds\Big|\mathcal{F}_t\Big]$. So the process Y is the value function of the maximization dual problem $Y_t = \text{esssup}_{\mu \in \mathcal{A}}\mathbb{E}_{\mathbb{Q}^\mu}\Big[-\xi_T - \int_t^T G(s,\mu_s)ds\Big|\mathcal{F}_t\Big]$.

c) We now consider a coefficient g with quadratic growth (H3) and bounded solution Y_t. Using the same notation, we know by Girsanov Theorem 7.2 that if $\mu \in$ BMO($\mathbb{P}$), Γ^μ is a u.i. martingale and the probability measure $\mathbb{Q}^\mu$ is well-defined. The proof of the dual representation is very similar to that of the previous case, after solving some integrability questions. It is enough to notice that

- by assumption, $|g(.,0)|^{\frac{1}{2}}$ is BMO($\mathbb{P}$),
- by Proposition 7.3, Z is BMO($\mathbb{P}$),
- by Girsanov Theorem 7.2, for any $\mu \in$ BMO($\mathbb{P}$), the processes μ, Z and $|g(.,0)|^{\frac{1}{2}}$ are in BMO($\mathbb{Q}^\mu$).

So $|g(t,Z_t)|^{\frac{1}{2}}$ and $|\mu_t Z_t|^{\frac{1}{2}}$ are in BMO($\mathbb{Q}^\mu$). Moreover, the process $(-G(t,\mu))^+$ which is dominated from above by $|g(t,0)|$ is a $\mathbb{Q}^\mu \times dt$-integrable process. Then the inequality (40) holds.
d) Let $\bar{\mu}$ be an optimal control, such that $g(t,Z_t) = \langle \bar{\mu}_t, -Z_t\rangle - G(t,\bar{\mu}_t)$. Given that $g(t,.)$ has quadratic growth, the polar function $G(t,.)$ satisfies the following inequality, $G(t,\bar{\mu}_t) \geq -|g(t,0)| + \frac{1}{4k}|\bar{\mu}_t|^2$. Then, for small $\varepsilon < \frac{1}{4k}$, $(\frac{1}{4k} - \varepsilon)|\bar{\mu}_t|^2 \leq G(t,\bar{\mu}_t) + |g(t,0)| - \varepsilon|\bar{\mu}_t|^2 \leq |g(t,0)| - g(t,Z_t) + \langle \bar{\mu}_t, -Z_t\rangle - \varepsilon|\bar{\mu}_t|^2 \leq |g(t,0)| - g(t,Z_t) + \frac{1}{4\varepsilon}|Z_t|^2$. Since both processes $|g(t,Z_t)|^{1/2}$ and Z are BMO($\mathbb{P}$), $\bar{\mu}$ is also BMO($\mathbb{P}$), and the other processes hold nice integrability properties with respect to both probability measures $\mathbb{P}$ and $\mathbb{Q}^\mu$ and the integral representation follows.
iii) Let $h(t,z) = g(t,z) - \frac{1}{2}C|z|^2$ be the convex function associated with g. Since g is the sum of two convex functions h and $\frac{1}{2}C|.|^2$, its Fenchel-Legendre transform G is the inf-convolution of the Fenchel-Legendre transforms of both h and $\frac{1}{2}C|.|^2$. But the Fenchel-Legendre transform of the quadratic function $\frac{1}{2}C|.|^2$ is still a quadratic function, $\frac{1}{2C}|\mu|^2$ and G has a quadratic growth (as the inf-convolution of a convex function H with a quadratic function). Therefore, for a given $\mu \in$ BMO($\mathbb{P}$), there exists $\bar{Z} \in$ BMO($\mathbb{P}$) such that $G(t,\mu_t) = \langle \mu_t, -\bar{Z}_t\rangle - g(t,\bar{Z}_t)$ (in other words, $\mu \in \partial(\bar{Z})$).
We now introduce the penalty function α^μ defined by $\alpha^\mu_t = \mathbb{E}_{\mathbb{Q}^\mu}\Big[\int_t^T G(s,\mu_s)ds\Big|\mathcal{F}_t\Big]$. Using the above duality result, we have:

$$\alpha^\mu_t = \mathbb{E}_{\mathbb{Q}^\mu}\Big[\int_t^T (\langle \mu_s, -\bar{Z}_s\rangle - g(s,\bar{Z}_s))ds\Big|\mathcal{F}_t\Big].$$

Since $\bar{\xi}_T = \int_0^T(\langle \mu_s, -\bar{Z}_s\rangle - g(s,\bar{Z}_s))ds + \int_0^T \bar{Z}_s dW^\mu_s$ and $\mathcal{R}^g_t(\bar{\xi}_T) = \int_0^t(\langle \mu_s, -\bar{Z}_s\rangle - g(s,\bar{Z}_s))ds + \int_0^t \bar{Z}_s dW^\mu_s$, we finally deduce that:

$$\alpha^\mu_t = \mathbb{E}_{\mathbb{Q}^\mu}\big[\bar{\xi}_T\big|\mathcal{F}_t\big] - \mathcal{R}^g_t(\bar{\xi}_T).$$

Moreover, using Equation (40), we have:

$$\alpha^\mu_S \geq \operatorname*{ess\,sup}_{\xi_T} \mathbb{E}_{\mathbb{Q}^\mu}\big[\xi_T\big|\mathcal{F}_S\big] - \mathcal{R}^g(\xi_T).$$

Hence, the result. $\square$

The question of the measurability of the optimal solution(s) $\bar{\mu}$ is considered in the following lemma.

Lemma 7.5 *Let g be a convex coefficient satisfying (H1) or (H3) and G be the associated polar function. There exists an progressively measurable optimal solution $\bar{\mu}$ such that $g(t,Z_t) = \langle \bar{\mu}_t, -Z_t\rangle - G(t,\bar{\mu}_t)$ a.s.$d\mathbb{P}\times dt$.*

Proof: For each $(\omega, t) \in \Omega \times [0, T]$, the sets given by: $\{\mu \in \mathbb{R}^n : g(\omega, t, Z_t) = Z_t\mu - G(\omega, t, \mu)\}$ are nonempty. Hence, by a measurable selection theorem (see for instance Dellacherie and Meyer [41] or Benes [14]), there exists a $\mathbb{R}^n$-valued progressively measurable process $\bar{\mu}$ such that: $g(\omega, t, Z_t) = \langle \bar{\mu}_t, -Z_t \rangle - G(\omega, t, \bar{\mu}_t) \quad d\mathbb{P} \times dt - a.s.$. $\square$

7.4 g-Conditional γ-Tolerant Risk Measures and Asymptotics

In this subsection, we pursue our presentation and study of g-conditional risk measures using an approach similar to that we have adopted in the static framework.

7.4.1 g-Conditional γ-Tolerant Risk Measures

As in the static framework, we can define dynamic versions for both coherent and γ-tolerant risk measures based on the properties of their coefficients using the uniqueness Lemma 6.6.
More precisely, let $\gamma > 0$ be a risk-tolerance coefficient. As in the static framework, where the γ-dilated of any static convex risk measure ρ is defined by $\rho_\gamma(\xi_T) = \gamma\rho\Big(\frac{1}{\gamma}\xi_T\Big)$we can define the g-conditional risk measure, $\mathcal{R}^g_\gamma$, *γ-tolerant of* $\mathcal{R}^g$, as the risk measure associated with the coefficient g_γ, which is the γ-dilated of g: $g_\gamma(t, z) = \gamma g(\frac{1}{t,\gamma} z)$.
Note that if g is Lipschitz continuous (H1), g_γ also satisfies (H1), and if g is continuous with quadratic growth (H3) with parameter k, then g also satisfies (H3), but with parameter $\frac{k}{\gamma}$. Note also that the dual function of g_γ, G_γ, can be expressed in terms of G, the dual function of g as $G_\gamma(\mu) = \gamma G(\mu)$.
A standard example of g-conditional γ-tolerant risk measure is certainly the dynamic entropic risk measure $e_{\gamma,t}(\xi_T) = \gamma \ln \mathbb{E}\big[\exp(-\frac{1}{\gamma}\xi_T)|\mathcal{F}_t\big]$, which is the γ-tolerant of $e_{1,t}$.

Asymptotic behaviour of entropic risk measure Let us look more closely at the dynamic entropic risk measure. Letting γ go to $+\infty$, the BSDE-coefficient $q_\gamma(z) = \frac{1}{2\gamma}|z|^2$ tends to 0 and we directly obtain the natural extension of the static case, $e_{\infty,t}(\xi_T) = \mathbb{E}_\mathbb{P}[-\xi_T|\mathcal{F}_t]$.
Letting γ tend to 0, the BSDE coefficient explodes if $|z| \neq 0$ and intuitively the martingale of this BSDE has to be equal to 0. More precisely, since by definition $\exp(e_{\gamma,t}(\xi_T)) = \mathbb{E}\big[\exp(-\frac{1}{\gamma}\xi_T)|\mathcal{F}_t\big]^\gamma$, $\lim_{\gamma\to 0}\exp(e_{\gamma,t}(\xi_T)) = ||\exp(-\xi_T)||^\infty_t = \inf\{Y \in \mathcal{F}_t : Y_t \geq \exp(-\xi_T)\}$. So we have $e_{0^+,t}(\xi_T) = ||-\xi_T||^\infty_t$. This conditional risk measure is a g-conditional risk measure associated with the indicator function of $\{0\}$. Let us also observe that $e_{0^+,t}(\xi_T)$ is an adapted non-increasing process without martingale part.

7.4.2 Marginal Risk Measure

In the general γ-tolerant case, assuming that the g-conditional risk measures are centered (equivalently $g(t, 0) = 0$ equivalently $G(t, .) \geq 0$), the same type of results can be obtained concerning the asymptotic behavior of the γ-dilated coefficient and the duality. Then, the limit of g_γ when $\gamma \to +\infty$ is the derivative of g at the origin in the direction of z.
$\mathcal{R}^g_\infty$ is the non-increasing limit of $\mathcal{R}^g_\gamma$ defined by its dual representation $\mathcal{R}^g_{\infty,t}(\xi_T) = \operatorname{ess\,sup}_{\mu\in\mathcal{A}}\{\mathbb{E}_{\mathbb{Q}^\mu}\big[-\xi_T\big|\mathcal{F}_t\big]\big|\, G(u, \mu_u) = 0,\ \forall u \geq t,\ -a.s.\}$; in some cases (in particular, in the quadratic case when the polar

function G has a unique 0, i.e. $G(u,0)=0$ is unique), $-\mathcal{R}^g_\infty$ is a linear pricing rule and can be seen as an extension of the Davis price (see Davis [37]).

7.4.3 Conservative Risk Measures and Super Pricing

We now focus on the properties of the g-conditional γ-tolerant risk measures when the risk tolerance coefficient goes to zero. To do so, we need some results in convex analysis regarding the so-called *recession function*, defined for any $z \in \mathrm{Dom(g)}$ by $g_{0^+}(z) := \lim_{\gamma\downarrow 0}\gamma g\big(\frac{1}{\gamma}z\big) = \lim_{\gamma\downarrow 0}\gamma\big(g(y+\frac{1}{\gamma}z)-g(y)\big)$. The key properties of this function are recalled in the Appendix 9.2.1.

Conservative Risk Measure • Under assumption $(H1)$, we may assume that $g(t,0)=0$. Therefore, the polar function G is non negative. Since $g(t,.)$ has a linear growth with constant k, the recession function $g_{0^+}(t,.)$ is finite everywhere with linear growth, and the domain of the dual function G is bounded by k. The BSDE(g_{0^+},ξ_T) has a unique solution $Y^0_t(\xi_T) \geq \mathcal{R}^{g_\gamma}_t(\xi_T)$. Using their dual representation through their polar functions $l_{\mathrm{Dom(G)}}$ and γG,

$$\begin{aligned}
Y^0_t(\xi_T) &= \operatorname{ess\,sup}_{\mu\in\mathcal{A}_k}\mathbb{E}_{\mathbb{Q}^\mu}\big[-\xi_T-\textstyle\int_t^T l_{\mathrm{Dom(G)}}(u,\mu_u)du\big|\mathcal{F}_t\big],\\
\mathcal{R}^g_{\gamma,t}(\xi_T) &= \operatorname{ess\,sup}_{\mu\in\mathcal{A}_k}\mathbb{E}_{\mathbb{Q}^\mu}\big[-\xi_T-\gamma\textstyle\int_t^T G(u,\mu_u)du\big|\mathcal{F}_t\big].
\end{aligned}$$

we can take the non decreasing limit in the second line and show that

$$\begin{aligned}
\mathcal{R}^g_{0^+,t}(\xi_T) &= \lim_{\gamma\downarrow 0}\mathcal{R}^g_{\gamma,t}(\xi_T) = Y^0_t(\xi_T)\\
&= \operatorname{ess\,sup}_{\mu\in\mathcal{A}_k}\big\{\mathbb{E}_{\mathbb{Q}^\mu}\big[-\xi_T\big|\mathcal{F}_t\big]\big|\, G(u,\mu_u)<\infty\ \forall u\geq t,\ du-a.s..\big\}\\
&= \operatorname{ess\,sup}_{\mu\in\mathcal{A}_k\cap\mathrm{Dom}(G)}\mathbb{E}_{\mathbb{Q}^\mu}\big[-\xi_T\big|\mathcal{F}_t\big].
\end{aligned}$$

• When the coefficient g has a quadratic growth $(H3)$, the recession function may be infinite on a set with positive measure and the BSDE(g_{0^+},ξ_T) is not well-defined. However, we can still take the limit in the dual representation of $\mathcal{R}^g_{\gamma,t}$, obtain the same characterization of $\mathcal{R}^g_{0^+,t}$, and consider $\mathcal{R}^g_{0^+}$ as a generalized solution of BSDE whose the coefficient g_{0^+} may be take infinite values. In particular if, as in the entropic case, $g_{0^+}=l_{\{0\}}$, G is finite everywhere and any equivalent probability measure associated with BMO coefficient, said to be in $\mathcal{Q}(\mathrm{BMO})$, is admissible. Then,

$$\mathcal{R}^{l_{\{0\}}}_{0^+,t}(\xi_T) = \operatorname{ess\,sup}_{\mathbb{Q}\in\mathcal{Q}(\mathrm{BMO})}\mathbb{E}_{\mathbb{Q}}\big[-\xi_T\big|\mathcal{F}_t\big] = ||-\xi_T||^\infty_t = e_{0^+,t}(\xi_T).$$

Super Price System Note that the conservative risk-measure $\mathcal{R}^g_{0^+,t}(\xi_T) = \operatorname{ess\,sup}_{\mu\in\mathcal{A}\cap\mathrm{Dom}(G)}\mathbb{E}_{\mathbb{Q}^\mu}\big[-\xi_T\big|\mathcal{F}_t\big]$ is the equivalent of the super-pricing rule of $-\xi_T$ (this notion was first introduced by El Karoui and Quenez [49] under the name "upper hedging price"). When the λ_t-translated of $\mathrm{Dom}(G)_t$ is a vector space, the recession function $g_{0^+}(t,z)$ is the indicator function of the orthogonal vector space $\mathrm{Dom}(G)_t^\top$ plus a linear function $\langle z,-\lambda_t\rangle$. Then, $\mathcal{R}^g_{0^+,t}(-\xi_T)$ is exactly the upper-hedging price associated with hedging portfolios constrained to live in $\mathrm{Dom}(G)^\top$.
The conservative measure is the smallest of coherent risk measure such that $\mathcal{R}^g_t(-\xi_T)-\mathcal{R}^g_t(-\eta_T) \leq \mathcal{R}^{\mathrm{coh}}_t(-\xi_T+\eta_T)$ for any (ξ_T,η_T) in the appropriate space.

Volume Perspective Risk Measure It is also possible to associate a coherent risk measure $\mathcal{R}^{\tilde{g}}$ with any convex risk measure $\mathcal{R}^g$, using the perspective function $\tilde{g}$ of the coefficient g, which is assumed to be normalized for the sake of simplicity ($g(t,0)=0$). The perspective function $\tilde{g}$ is defined as:

$$\tilde{g}(t,\gamma,z)=\begin{cases}\gamma g(t,\frac{z}{\gamma}) & \text{if } \gamma>0\\ \lim_{\gamma\to 0}\gamma g(t,\frac{z}{\gamma})=g_{0^+}(t,z) & \text{if } \gamma=0\end{cases}$$

More details about $\tilde{g}$ can be found in the Appendix 9.2.2. As a direct consequence, the $\tilde{g}$-conditional risk measure $\mathcal{R}^{\tilde{g}}$ is a coherent risk measure.

8 Inf-Convolution of g-Conditional Risk Measures

In this section, we come back to inf-convolution of risk measures, when they are g-conditional risk measures. This study is based upon the inf-convolution of their respective coefficients.
More precisely, we will study for any t the inf-convolution of the g-conditional risk measures $\mathcal{R}_t^A$ and $\mathcal{R}_t^B$ defined as

$$\left(\mathcal{R}^A\square\mathcal{R}^B\right)_t(\xi_T)=\operatorname{ess\,inf}_{F_T}\left\{\mathcal{R}_t^A\left(\xi_T-F_T\right)+\mathcal{R}_t^B\left(F_T\right)\right\} \tag{41}$$

where both ξ_T and F_T are taken in the appropriate space and show that this new dynamic risk measure is under mild assumptions the (maximal solution) $\mathcal{R}^{A,B}$ of the BSDE $(g^A\square g^B,-\xi_T)$ where $(g^A\square g^B)(.,t,z)=\operatorname{ess\,inf}_z(g(.,t,x-z)+g(.,t,z))$. Then, the next step is to characterize the optimal transfer of risk between both agents A and B, agent A being exposed to ξ_T at time T. Some key results on the inf-convolution of convex functions are recalled in the Appendix 9.3, the main argument being summarized in the proposition below:

Proposition 8.1 *Let g^A and g^B be two convex functions of z. Under the following condition*

$$g^A_{0^+}(t,z)+g^B_{0^+}(t,-z)>0,\quad \forall z\neq 0$$

then $g^A\square g^B$ is exact for any z as the infimum is attained by some x^*:

$$g^A\square g^B(z)=\inf_x\{g^A(z-x)+g^B(x)\}=g^A(z-x^*)+g^B(x^*).$$

8.1 Inf-convolution and Optima

We now focus on our main problem of inf-convolution of g-conditional risk measures as expressed in Equation (41). The following theorem gives us an explicit characterization of an optimum for the inf-convolution problem provided such an optimum exists:

Theorem 8.2 *Let g^A and g^B be two convex coefficients depending only on z and satisfying the condition of Proposition 8.1. For a given ξ_T in the appropriate space (either $\mathbb{L}^2$ or $\mathbb{L}_\infty$), let $(\mathcal{R}_t^{A,B}(\xi_T),Z_t)$ be the maximal solution of the BSDE $(g^A\square g^B,-\xi_T)$ and $\widehat{Z}_t^B$ be a measurable process such that $\widehat{Z}_t^B=\arg\min_x\Big\{g^A\big(t,Z_t-x\big)+g^B\big(t,x\big)\Big\}$ $dt\times d\mathbb{P}-a.s.$.*

Then, the following results hold:

(1) *For any $t \in [0,T]$ and for any F_T such that both $\mathcal{R}_t^A(\xi_T - F_T)$ and $\mathcal{R}_t^B(F_T)$ are well defined:*

$$\mathcal{R}_t^{A,B}(\xi_T) \leq \mathcal{R}_t^A(\xi_T - F_T) + \mathcal{R}_t^B(F_T) \quad \mathbb{P} \ - a.s.$$

(2) *If the process $\widehat{Z}^B$ is admissible, then for any $t \in [0,T]$*

$$\mathcal{R}_t^{A,B}(\xi_T) = (\mathcal{R}^A \Box \mathcal{R}^B)_t(\xi_T) \quad \mathbb{P} - a.s.$$

and the structure F_T^* defined by the forward equation

$$F_T^* = \int_0^T g^B(t, \widehat{Z}_t^B)dt - \int_0^T \widehat{Z}_t^B dW_t$$

is an optimal solution for the inf-convolution problem:

$$(\mathcal{R}^A \Box \mathcal{R}^B)_t(\xi_T) = \mathcal{R}_t^A(\xi_T - F_T^*) + \mathcal{R}_t^B(F_T^*).$$

Proof: (1) First, note that the existence of such a measurable process $\widehat{Z}_t^B$ is guaranteed by Theorem 8.1. In the following, we consider any F_T such that both $\mathcal{R}_t^A(\xi_T - F_T)$ and $\mathcal{R}_t^B(F_T)$ are well defined.
Let us now focus on $\mathcal{R}_t^A(\xi_T - F_T) + \mathcal{R}_t^B(F_T)$. It satisfies

$$\begin{aligned} -d\big(\mathcal{R}_t^A(\xi_T - F_T) + \mathcal{R}_t^B(F_T)\big) &= \big(g^A(t, Z_t^A) + g^B(t, Z_t^B)\big)dt - \big(Z_t^A + Z_t^B\big)dW_t \\ &= \big(g^A(t, Z_t - Z_t^B) + g^B(t, Z_t^B)\big)\,dt - Z_t dW_t, \end{aligned}$$

and at time T, $\mathcal{R}_T^A(\xi_T - F_T) + \mathcal{R}_T^B(F_T) = -\xi_T$.
Therefore, $(\mathcal{R}_t^A(\xi_T - F_T) + \mathcal{R}_t^B(F_T), Z_t)$ is solution of the BSDE with terminal condition $-\xi_T$, which is also the terminal condition of the BSDE $(g^A \Box g^B, -\xi_T)$, and a coefficient g written in terms of the solution Z_t^B of the BSDE (g^B, F_T) as: $g(t,z) = g^A(t, z - Z_t^B) + g^B(t, Z_t^B)$. Using the definition of the inf-convolution, this coefficient is then always greater than $g^A \Box g^B$. Thus, we can compare $\mathcal{R}_t^A(\xi_T - F_T) + \mathcal{R}_t^B(F_T)$ with the solution of the BSDEs $(g^A \Box g^B, -\xi_T)$ using the comparison Theorem (5.3) and obtain the desired inequality.
(2) Let now assume that the process $\widehat{Z}_t^B$ is admissible, using different notions of admissibility when either (H1) or (H3) (square integrability or BMO).
Thanks to Theorem 8.1, we can show that both dynamic risk measures coincide.
We now introduce the structure F_T^* defined by the forward equation $F_t^* = \int_0^t g^B(s, \widehat{Z}_s^B)ds - \int_0^t \widehat{Z}_s^B dW_s$.
Note first that thanks to the admissibility of the process $\widehat{Z}_t^B$, such a structure is well-defined and belongs to the appropriate space (either $L^2(\mathcal{F}_T)$ or $L^\infty(\mathcal{F}_T)$).
Let us also observe that $-F_t^*$ is also solution of the BSDE $(g^B, -F_T^*)$ since $-F_t^* = -F_T^* + \int_t^T g^B(u, \widehat{Z}_u^B)dt - \int_t^T \widehat{Z}_u^B dW_u$. By uniqueness, this process is $\mathcal{R}_t^B(F_T^*)$.

Since $\mathcal{R}_t^A(\xi_T - F_T^*) + \mathcal{R}_t^B(F_T^*)$ is solution of the BSDE with coefficient written as $g^A(t, Z_t - \widehat{Z}_t^B) + g^B(t, \widehat{Z}_t^B)$ and terminal condition $-\xi_T$ and given that $\big(g^A \Box g^B\big)(t, Z_t) = g^A\big(t, Z_t - \widehat{Z}_t^B\big) + g^B\big(t, \widehat{Z}_t^B\big)$, by uniqueness, we also have $\forall t \geq 0$, $\mathcal{R}_t^{A,B}(\xi_T) = (\mathcal{R}^A \square \mathcal{R}^B)_t(\xi_T) \quad \mathbb{P}\ a.s.$

The proof also gives the optimality for the Problem (41) of the structure $F_T^* = \int_0^T g^B(t, \widehat{Z}_t^B)dt - \int_0^T \widehat{Z}_t^B dW_t$.

$\square$

Remark 8.3 (On uniqueness on the optimum) *Note that the optimal structure F_T^* is determined to within a constant because of the translation invariance property (P3-) satisfied by both risk measures $\mathcal{R}_t^A$ and $\mathcal{R}_t^B$ since:*

$$\begin{aligned} \operatorname{ess\,inf}_{F_T} \{\mathcal{R}_t^A(\xi_T - (F_T + m)) + \mathcal{R}_t^B(F_T + m)\} &= \operatorname{ess\,inf}_{F_T} \{\mathcal{R}_t^A(\xi_T - F_T) + m + \mathcal{R}_t^B(F_T) - m\} \\ &= (\mathcal{R}^A \square \mathcal{R}^B)_t(\xi_T). \end{aligned}$$

Note also that F_T^* is optimal for all the optimal structure problems for all stopping times S such that $0 \leq S \leq T \quad a.s.$.

The following Theorem gives some sufficient conditions ensuring the admissibility of the process $\widehat{Z}_t^B$:

Theorem 8.4 [Exact Inf-convolution] *Let g^B be a strongly convex coefficient. For any convex function g^A, the inf-convolution $g^A \square g^B$ is convex with quadratic growth (H3), so in particular, if g^A satisfies (H3).*
In this case, the process $\widehat{Z}_t^B$, defined in Theorem 8.2, is in BMO($\mathbb{P}$).

Note that in this case, the optimal structure F_T^*, defined in Theorem 8.2, is quasi-bounded as it belongs to the BMO-closure of $\mathbb{L}_\infty$ as defined by Kazamaki [77] (chapter 3).

Proof: From the duality Theorem 7.4, the optimal control μ^* of $G^{A,B}$, the polar function of $g^A \square g^B$, is in BMO($\mathbb{P}$). From the inf-convolution, we deduce that this is also the optimal control for G^A and G^B in the following sense:

$$g^A(t, Z_t - \widehat{Z}_t^B) = \langle \mu_t^*, -(Z_t - \widehat{Z}_t^B)\rangle - G^A(t, \mu_t^*),$$
$$g^B(t, \widehat{Z}_t^B) = \langle \mu_t^*, -\widehat{Z}_t^B\rangle - G^B(t, \mu_t^*).$$

Therefore, both $Z_t - \widehat{Z}_t^B$ and $\widehat{Z}_t^B$ are in BMO($\mathbb{P}$) (from Proposition 7.3) and the process $\widehat{Z}_t^B$ is admissible.

$\square$

Comments:

(i) Just as in the static framework, we obtain the same result when considering g-conditional γ-tolerant risk measures. The Borch theorem is therefore extremely robust since the quota sharing of the initial exposure remains an optimal way of transferring the risk between different agents.

(ii) Under some particular assumptions, the underlying logic of the transaction is non-speculative since there is no interest for the first agent to transfer some risk or equivalently to issue a structure if she is not initially exposed. This result is completely consistent with the result we have already obtained in the static framework.

8.2 Hedging Problem

As in subsection 4.2.2, we consider the hedging problem of a single agent. She wants to hedge her terminal wealth X_T by optimally investing on financial market and assesses her risk using a general g-conditional risk measure $\mathcal{R}^g$.

8.2.1 Framework

We consider the same framework as that introduced in Subsection 1.5.3 when looking at the question of dynamic hedging in the static part. More precisely, we assume that d basic securities are traded on the market. Their forward (non-negative) vector price process S follows an Itô semi-martingale with a uniformly bounded drift coefficient and an invertible and bounded volatility matrix σ_t. Under $\mathbb{P}$,

$$\frac{dS_t}{S_t} = \sigma_t(dW_t + \lambda_t dt) \quad ; \quad S_0 \quad \text{given.} \tag{42}$$

To avoid arbitrage, we assume **(AAO)**: there exists a probability measure $\mathbb{Q}$, equivalent to $\mathbb{P}$, such that S is a $\mathbb{Q}$-local martingale. From the completeness of this basic arbitrage-free market, we deduce the uniqueness of $\mathbb{Q}$, which is usually called the risk-neutral probability measure.

The agent can invest in dynamic strategies θ, i.e. d-predictable processes and $(G_t(\theta) = (\theta.S)_t)$ denotes the associated gain process.

We assume that not all strategies are admissible and that, for instance, the agent has some restriction imposed on the transaction size. These constraints create some market incompleteness in the framework we consider. $\Theta_T^S = \{G_T(\theta) \,|\, \theta.S \text{ is bounded by below }, \theta \in \mathcal{K}\}$ is the set of admissible hedging gain processes. $\mathcal{K}$ is a convex subset of BMO($\mathbb{P}$) such that any admissible strategies θ is in $\mathcal{K}$ (equivalently, $\forall\, t,\ \theta_t \in \mathcal{K}_t$).

8.2.2 Hedging Problem

At time 0, the hedging problem of the agent can be expressed as the determination of an optimal admissible strategy θ as to minimize the initial g-conditional risk measure of her terminal wealth

$$\inf_{\theta \in \mathcal{K}} \mathcal{R}_0\big(X_T - G_T(\theta)\big). \tag{43}$$

The value functional of this program is the dynamic market modified risk measure of agent A, denoted by $\mathcal{R}^m$. Using the previous results, we can obtain the following proposition:

Proposition 8.5 *i) Let $l_{\sigma_t^*}(\mathcal{K}_t) = l_{\widehat{\mathcal{K}_t}}$ be the indicator function of the convex set $\widehat{\mathcal{K}_t} = \sigma_t^* \mathcal{K}_t$. Provided that the inf-convolution $g \square l_{\sigma_t^*}(\mathcal{K}_t)(Z_t)$ is well-defined, the residual risk measure $\mathcal{R}^m$ is given as the maximal solution of the following BSDE:*

$$-d\mathcal{R}_t^m(X) = g^m(t, Z_t)dt - Z_T dW_t \quad ; \quad \mathcal{R}_T^m(X) = -X_T$$

where g^m is the restriction of the coefficient g to the admissible set: $g^m(t, Z_t) = g \square l_{\sigma_t^*}(\mathcal{K}_t)(Z_t)$.
ii) If g is strongly convex, then this hedging problem has a solution.
In particular, in the entropic case, $g^m(t,z) = \frac{1}{2\gamma} d_{\frac{1}{\gamma}}(z, \widehat{K_t})^2$ where γ is the risk tolerance coefficient and $d_{\frac{1}{\gamma}}(z, \mathcal{K})$ is the distance function to $\mathcal{K}$. The optimal investment strategy $\theta^\star$ is the projection on $\mathcal{K}$ of Z_t,

solution of the BSDE $(g^m, -X_T)$.

The terminal value $G_T(\theta^\star)$ *of the associated portfolio is given by:*

$$G_T(\theta^\star) = x + \int_0^T (\theta_t^\star)^* \sigma_t \lambda_t dt + \int_0^T (\theta_t^\star)^* \sigma_t dW_t.$$

8.2.3 Comments

Generalized BSDEs: In the static framework, we expressed the hedging problem as an inf-convolution between the seminal risk measure of the agent and the risk measure $\nu^{\mathcal{H}}$ generated by $\mathcal{H}$, the convex set of constrained terminal gains, or more generally the inf-convolution between the seminal risk measure of the agent and the convex indicator of $\mathcal{H}$ (Proposition 3.2).

From a dynamic point of view, the set $\mathcal{H}$ can be seen as the set of all dynamic terminal values of portfolios with some constraint on the strategies. Everything can be formulated in the same way. Note that the natural candidate for $\mathcal{R}^{\mathcal{H}}$ would be the inf-convolution between the dynamic worst case risk measure and the convex indicator of $\mathcal{H}$: $l^{\mathcal{H}} \square \lim_{\gamma\to\infty}(\frac{1}{2\gamma}|z|^2)$. This infimum is always strictly positive. Moreover, it is an increasing process at the limit. To model this "limit BSDE", an increasing process has to be introduced (for more details, please refer to El Karoui and Quenez [50] and Cvitanic and Karatzas [35]). As a consequence, the dynamic version of the risk measure generated by $\mathcal{H}$ cannot be seen exactly as the solution of a standard BSDE, as previously defined, in the sense that the coefficient can take infinite values.

This is however not such a problem here as we really focus on the inf-convolution. Therefore, we can simply consider the restriction of the seminal risk measure to a particular set. The powerful regularization impact of the inf-convolution is again visible here.

Hedging problem at any time t: Solving the hedging problem at time 0 leads to the characterization of a particular probability measure, which can be called *calibration probability measure* as the prices of any hedging instruments made with respect to this measure coincide with the observed market prices on which all agents agree.

Solving the hedging problem at any time t is equivalent to solving the same problem at time 0 as soon as the prices of these hedging instruments at this time t are given as the expected value of their discounted future cash flows under the optimal calibration probability measure determined at time 0. This optimal probability measure is very robust as it remains the pricing measure for hedging instruments between 0 and T.

Therefore, we can introduce the same problem at any time t:

$$\operatorname*{ess\,inf}_{\theta\in\mathcal{K}} \mathcal{R}_t\big(X_T - G_T(\theta)\big) = \mathcal{R}_t^m(X_T).$$

BSDEs time-consistency and uniqueness are key arguments to show that if θ is optimal for the problem at time 0, then θ is optimal for the optimization program at any time t.

Dynamic Entropic Framework The entropic hedging problem, lying at the core of this book, has been intensively studied in the literature. But only a few papers are using a BSDEs framework. After the seminal paper by El Karoui and Rouge [51], different authors have used BSDEs to solve this problem under various assumptions (see in particular Sekine [110], Mania et al. [88] and more recently Hu, Imkeller and Müller [73] or Mania and Schweizer [89]).

Another approach, different from what we have mentioned above, has been used to solve the hedging problem involves the dual representation for the dynamic entropic risk measure as given by Theorem 7.4:

$$e_{\gamma,t}(\Psi) = \sup_{\mu\in\mathcal{A}^q} \mathbb{E}_{\mathbb{Q}^\mu}\Big[-\Psi - \gamma\int_t^T \frac{|\mu_s|^2}{2}ds|\mathcal{F}_t\Big].$$

Therefore, the hedging problem at any time t can be rewritten as:

$$\operatorname*{ess\,inf}_{\theta\in\mathcal{K}} \operatorname{ess\,sup}_{\mu\in\mathcal{A}^q}\Big\{\mathbb{E}_{\mathbb{Q}^\mu}[-X_T + G_T(\theta)|\mathcal{F}_t] - \gamma h(\mathbb{Q}^\mu|\mathbb{P})\Big\}$$

and it may be solved by using dynamic programming arguments.

9 Appendix: Some Results in Convex Analysis

We now present some key results in convex analysis that will be useful to obtain the dual representation of g-conditional risk measures. More details or proofs may be found in Aubin [4], Hiriart-Urruty and Lemaréchal [70] or Rockafellar [103].
All the notations and definitions we introduce are consistent with the notations of risk measures. They may differ from the standard framework of convex analysis (especially regaring the sign).

Even if the coefficient of the BSDE is finite, we are also interested in convex functions taking infinite values. The main motivation is the definition of its convex polar function G. In that follows, as in [70], we always assume that the considered functions are not identically $+\infty$ and are bounded from below by a affine function (note that this assumption is rather general and does not necessarily require that the functions are convex). The domain of a function g is defined as the nonempty set $\mathrm{Dom}(g) = \{z\ :\ g(z) < +\infty\}$. The epigraph of convex function is the subset of $\mathbb{R}^n \times \mathbb{R}$ as: $\mathrm{epi}\,\mathrm{g} = \{(x,\lambda)\,|\,g(x) \leq \lambda\}$. When the convex functions are lower semicontinuous (lsc), epig is closed, and they are said to be closed.

9.1 Duality

9.1.1 Legendre-Fenchel Transformation

Let g be a convex function. The polar function G is defined on $\mathbb{R}^n$ by

$$G(\mu) = \sup_z(\langle\mu, z\rangle - g(z)) = \sup_{z\in\mathrm{Dom}(g)}(\langle\mu, -z\rangle - g(z)). \tag{44}$$

The function G is a closed convex function, which can take infinite values. The conjugacy operation induces a symmetric one-to-one correspondence in the class of all closed convex functions on $\mathbb{R}^n$:

$$g(z) = \sup_\mu(\langle\mu, -z\rangle - G(\mu)), \quad G(\mu) = \sup_z(\langle\mu, -z\rangle - g(z)).$$

Convex set and duality Given a nonempty subset $S \subset \mathbb{R}^n$, the *indicator function* (in the convex analysis terminology) of S, $l_S : \mathbb{R}^n \to \mathbb{R}^+ \cup \{+\infty\}$, is defined by:

$$l_S(z) = 0 \quad \text{if} \quad z \in S \qquad \text{and} \qquad +\infty \quad \text{if not.}$$

l_S is convex (closed), iff S is convex (closed) since $\operatorname{epi} l_S = S \times \mathbb{R}^+$.
The polar function of l_S is the *support function* of $-S$:

$$\sigma_S(z) := \sup_{s \in S} \langle s, -z \rangle = \sup_s \{ \langle s, -z \rangle - l_S(s) \}.$$

The support function is closed, convex, homgeneous function: $\sigma_S(\lambda z) = \lambda \sigma_S(z)$ for all $\lambda > 0$. Its epigraph and its domain are convex cones.

9.1.2 Subdifferential and Optimization

The sub-differential of the convex function g in z, whose the elements are called *subgradient* of g at z, is the set $\partial g(z)$ defined as:

$$\partial g(z) = \{ \mu \mid g(x) \geq g(z) - \langle \mu, x - z \rangle, \quad \forall x \} = \{ \mu \mid g(z) - \langle \mu, -z \rangle \geq G(\mu) \}. \tag{45}$$

If $z \notin Dom(g)$, $\partial g(z) = \emptyset$. But if z is in the interior of Dom(g), the subgradient $\partial g(z)$ is non-empty (see Section E in [70] or Chapter 23 in [103]); in fact, it is enough that z belongs to the relative interior of Dom(g), where ridom(g) is defined in Section A in [70] and in Chapter 6 in [103]. In particular, *if g is finite, then $\partial g(z)$ is nonempty for any z.* When $\partial g(z)$ is reduced to a single point, the function is said to be differentiable in z. Note that when the function g is the indicator function of the convex set C, the sub-differential of g in $z \in C$ is the positive normal cone $N_C^+(z)$ to C at z, $N_C^+(z) = \{ s \in \mathbb{R}^n \mid \forall y \in C \quad - \langle s, y - z \rangle \} \leq 0\}$.

Subgradients are solutions of minimization programs as $\inf_z (g(z) - \langle \mu, -z \rangle)$ $(= -G(\mu))$, or its dual program, $\inf_\mu (G(\mu) - \langle \mu, -z \rangle)$ $(= -g(z))$. The precise result is the following (see Section E in [70]): Let g be a closed convex function and G its polar function.

- $\widehat{\mu} \in \partial g(\widehat{z}) \iff \widehat{\mu}$ is the optimal for the following minimization program, that is $-g(z) = \inf_\mu (G(\mu) - \langle \mu, -z \rangle) = G(\widehat{\mu}) - \langle \widehat{\mu}, -z \rangle$.
- $\widehat{z} \in \partial G(\widehat{\mu}) \iff \widehat{z}$ is optimal for the following minimization program, that is in $-G(\mu) = \inf_z (g(z) - \langle \mu, -z \rangle) = g(\widehat{z}) - \langle \mu, -\widehat{z} \rangle$.

In the following, when working with BSDEs, we will denote by $z\mu$ the scalar product between the line vector z and the column vector μ.

9.2 Recession function

9.2.1 Recession Function

The recession function associated with a closed convex function g is the homogeneous convex function defined for $z \in \text{Dom}(g)$ by $g_{0^+}(z) := \lim_{\gamma \downarrow 0} \gamma g\big(\frac{1}{\gamma} z\big) = \lim_{\gamma \downarrow 0} \gamma \big(g(y + \frac{1}{\gamma} z) - g(y) \big)$. This function g_{0^+} is the smallest

homogeneous function h such that for any $z, y \in \mathrm{Dom}(g)$, $g(z) - g(y) \leq h(z-y)$. When $g(z) \leq c + k|z|$, $g_{0^+}(z) \leq k|z|$ is a finite convex function, and the function g is Lipschitz-continuous function with Lipschitz coefficient k since $g(z) - g(y) \leq k|z-y|$.
This property explains why any convex coefficient of BSDE satisfying the assumption $(H2)$ in fact satisfies $(H1)$.
Let G be the polar function of g. Using obvious notations, for any $\mu \in \mathrm{Dom}(g)$, $\mathrm{polar}\, g_{0^+}(\mu) = \lim_{\gamma \downarrow 0}(\gamma G(\mu)) = 0$. So, $\mathrm{polar}\, g_{0^+} = l_{\mathrm{dom}\,\mathrm{G}}$. By the conjugacy relationship applied to closed functions, g_{0^+} is the support function of $\mathrm{Dom}(G)$; so g_{0^+} is finite everywhere iff $\mathrm{Dom}(G)$ is bounded, or iff g is uniformly Lipschitz, or finally iff g has linear growth.
The recession function of the quadratic function $q_k(z) = c + k|z|^2$ is infinite except in $z = 0$, and its polar function is the null function. More generally, convex functions such that $g_{0^+} = l_{\{0\}}$ admit finite polar function G and this condition is sufficient.

9.2.2 Perspective Function

Let us consider a closed convex function g such that $g(0) = 0$. The perspective function associated with g is the function $\tilde{g}$ defined on $\mathbb{R}^+ \times \mathbb{R}^n$ as:

$$\tilde{g}(\gamma, z) = \begin{cases} \gamma g(\frac{z}{\gamma}) & \text{if } \gamma > 0 \\ \lim_{\gamma \to 0} \gamma g(\frac{z}{\gamma}) = g_{0^+}(z) & \text{if } \gamma = 0 \end{cases}$$

Note first that the perspective function of g corresponds to the γ-dilated of g, seen as a function of both variables z and γ, when $\gamma > 0$. It is prorogated for $\gamma = 0$ by the recession function g_{0^+}. Note that the risk tolerance coefficient is considered as a risk factor itself. $\tilde{g}$ is a positive homogeneous convex function (for more details, please refer to Part B [70]).
The dual function of $\tilde{g}$, defined on $\mathbb{R} \times \mathbb{R}^n$, is given by:

$$\tilde{G}(\theta, \mu) = 0 \quad \text{if } G(\mu) \leq -\theta\,, \quad \text{and} \quad +\infty \quad \text{otherwise.}$$

If $g(0) < \infty$, note that $G(\mu)$ is bounded.

9.3 Infimal Convolution of Convex Functions and Minimization Programs

Addition and inf-convolution of closed convex functions are two dual operations with respect to the conjugacy relation.
Let g^A and g^B be two closed convex functions from $\mathbb{R}^n \cup \{+\infty\}$. By definition, the infimal convolution of g^A and g^B is the function $g^A \Box g^B$ defined as:

$$\left(g^A \Box g^B\right)(z) = \inf_{y^A + y^B = z} (g^A(y^A) + g^B(y^B)) = \inf_y (g^A(z-y) + g^B(y)). \tag{46}$$

If $g^A \Box g^B \not\equiv \infty$, then its polar function, denoted by G^{AB}, is simply the sum of the polar functions of g^A and g^B:

$$G^{AB}(\mu) = G^A(\mu) + G^B(\mu)$$

9.3.1 Inf-Convolution as a Proper Convex Function

The function $g^A \Box g^B$ may take the value $-\infty$, which is contrary to the assumption made in Subsection 7.2. To avoid this difficulty, we assume that both functions g^A and g^B have a common affine minorant $\langle s, . \rangle - b$. This assumption may be expressed in terms of their recession functions, both of them being also bounded from below by $\langle s, . \rangle$. Therefore, $g^A_{0+}(z) + g^B_{0+}(-z) \geq 0$ for any z and consequently $\left(g^A_{0+} \Box g^B_{0+}\right)(0) \geq 0$. Note that this condition can also be expressed in terms of the polar functions of g^A and g^B as $\mathrm{dom}(G^A) \cap \mathrm{dom}(G^B) \neq \emptyset$.

9.3.2 Existence of Exact Inf-Convolution

We are interested in the existence of a solution to the inf-convolution problem (46). When a solution exists, the infimal convolution is said to be *exact*.

The previous conditions are almost sufficient, as proved in Rockafellar [103] since, if we assume

$$g^A_{0+}(z) + g^B_{0+}(-z) > 0, \quad \forall z \neq 0 \tag{47}$$

then $g^A \Box g^B$ is a closed convex function, and for any z, the infimum is attained by some x^*:

$$g^A \Box g^B(z) = \inf_x \{g^A(z-x) + g^B(x)\} = g^A(z - x^*) + g^B(x^*).$$

The condition (47) is satisfied if $\mathrm{intdom}(G^A) \cap \mathrm{intdom}(G^B) \neq \emptyset$ (in fact, the true interior corresponds to the relative interior defined in Section A by Hiriart-Urruty and Lemaréchal [70]).

Examples of exact inf-convolution: We now mention different cases where the inf-convolution has a solution.

- First, when both convex functions g^A and g^B are dilated, then their inf-convolution is exact without having to impose any particular assumption, as we have already noticed when working with static risk measures in the first part (see Proposition 3.5). More precisely, assume that g^A and g^B are dilated from a given convex function g such that $g^A = g_{\gamma_A}$ and $g^B = g_{\gamma_B}$, then $g^A \Box g^B = g_{\gamma_A + \gamma_B}$ and for any z, an optimal solution x^* to the inf-convolution problem is given by $x^* = \frac{\gamma_B}{\gamma_A + \gamma_B}$.
- More generally, if g^A is bounded from below and if g^B satisfies the qualification constraint ensuring that $\inf_z g^B(z)$ is reached for some z (in other words, g^B has a strictly positive recession function g^B_{0+}), then the condition (47) is satisfied and the inf-convolution $g^A \Box g^B$ has a non-empty compact set of solutions.

9.3.3 Characterization of Optima

We are now interested on the characterization of optima in the case of exact inf-convolution. This can be done in terms of the subdifferentials of the different convex functions involved. More precisely, let us consider z^A and z^B respectively in $\mathrm{dom}(g^A)$ and in $\mathrm{dom}(g^B)$ and $z = z^A + z^B$ in $\mathrm{dom}(g^A \Box g^B)$. Then, $\partial g^A(z^A) \cap \partial g^B(z^B) \subset \partial(g^A \Box g^B)(z)$.

Moreover, if $\partial g^A(z^A) \cap \partial g^B(z^B) \neq \emptyset$, then the inf-convolution $g^A \Box g^B$ is exact at $z = z^A + z^B$ and $\partial g^A(z^A) \cap$

$\partial g^B(z^B) = \partial(g^A \Box g^B)(z)$. (For more details, please refer to [70]).

In particular, as 0 belongs to the domain of g^A and g^B, if $\partial g^A(0) \cap \partial g^B(0) \neq \emptyset$, then $\partial(g^A \Box g^B)(0) = \partial g^A(0) \cap \partial g^B(0)$ and the inf-convolution is exact at 0.

Moreover, if both functions are centered, i.e. $g^A(0) = g^B(0) = 0$, then the inf-convolution is also centered as $(g^A \Box g^B)(0) = g^A(0) + g^B(0) = 0$.

9.3.4 Regularization by Inf-Convolution

As convolution, the infimal convolution is used in regularization procedures. The most famous regularizations are certainly, on the one hand, the *Lipschitz regularization* $g_{(k)}$ of g using the inf-convolution with the kernel $b_k(z) = k|z|$ and on the other hand, the *differentiable regularization*, also called *Moreau-Yosida regularization*, $g_{[k]}$ of g using the inf-convolution with the kernel $q_k(z) = \frac{k}{2}|z|^2$. Both regularizations do not have however the same "efficiency".

LIPSCHITZ REGULARIZATION We first consider the inf-convolution $g_{(k)}$ of g using the kernel $b_k(z) = k|z|$ or more generally using functions whose polar's domain is bounded (or equivalently with a finite recession function).

The function $g_{(k)}$ is finite, convex, non decreasing w.r. to k. Moreover, its inf-convolution $g_{(k)}$ is Lipschitz-continuous, with Lipschitz constant k. More generally, *the inf-convolution of two convex functions, one of them satisfying (H1), also satisfies (H1) without any condition on the other function.*

If $z_0 \in \text{int dom} g$, then $g_{(k)}(z_0) = g(z_0)$ for k large enough. When $g = l_C$ is the indicator function of a closed convex set C, $g_k = k\text{dist}(.,C)$.

This regularization is used in the book's chapter dedicated to BSDEs to show the existence of BSDE with continuous coefficient.

MOREAU-YOSIDA REGULARIZATION We now consider the inf-convolution $g_{[k]}$ of g using the kernel $q_k(z) = \frac{k}{2}|z|^2$. The function $g_{[k]}$ is finite, convex, non decreasing w.r. to k. Moreover, $g_{[k]}$ is differentiable and its gradient is Lipschitz-continuous with Lipschitz constant k. In other words, the polar function of $g_{[k]}$ is strongly convex with module k, equivalently $G_{[k]}(.) - \frac{k}{2}|.|^2$ is still a convex function (for more details, please refer to Cohen [32]).

There exists a point $J_k(z)$ that attains the minimum in the inf-convolution problem with q_k. The maps $z \to J_k(z)$ are Lipschitz continuous with a constant 1, independent of k and monotonic in the following sense $(J_k(z) - J_k(y))\ (z-y)^* \geq ||J_k(z) - J_k(y)||^2$. Moreover,$\nabla g_{[k]} = k(z - J_k(z))$.

More generally, *the inf-convolution of two convex functions, one of them being strongly convex, satisfies (H3) without any condition on the other function.*

References

[1] Acerbi, C., Tasche, D.: On the Coherence of Expected Shortfall, Journal of Banking and Finance 26, 1487-1503 (2002).

[2] Artzner, P., Delbaen, F., Eber, J.M., Heath,D.: Coherent Measures of Risk, Mathematical Finance 9, 203-228 (1999).

[3] Artzner, P., Delbaen, F., Eber, J.M., Heath,D., Ku, H.: Coherent Multiperiod risk adjusted values and Bellman's principle, Working Paper (2004) (http://www.math.ethz.ch/ delbaen/ftp/preprints/adehk-or.pdf|).

[4] Aubin, J.P.: Optima and Equilibria, Springer Verlag (1998).

[5] Bally, V., Caballero, M.E., Fernandez, B., El Karoui, N.: Reflected BSDE's, PDE's and Variational Inequalities, Preprint (Rapport de recherche INRIA No. 4455) (2002).

[6] Barrieu, P.: Structuration Optimale de Produits Financiers en Marché Illiquide et Trois Excursions dans d'autres Domaines des Probabilités, Thèse de doctorat, Université de Paris VI (2002).

[7] Barrieu, P., Cazanave, N., El Karoui, N.: Inf-convolution of BSDEs with Quadratic Growth and Applications to Dynamic Optimal Risk Transfer, Working Paper (2004).

[8] Barrieu, P., El Karoui, N.: Structuration Optimale de Produits Financiers et Diversification en Présence de Sources de Risque Non-Négociables, C. R. de l'Acad. Sci. Paris, Sér. I Math. 336, 493-498 (2003).

[9] Barrieu, P, El Karoui, N.: Optimal Derivatives Design under Dynamic Risk Measures, in Mathematics of Finance (eds: G. Yin, Q. Zhang), Contemporary Mathematics - A.M.S. Proceedings (2004).

[10] Barrieu, P, El Karoui, N.: Inf-convolution of Risk Measures and Optimal Risk Transfer, Finance and Stochastics 9, 269-298 (2005).

[11] Becherer, D.: Rational Hedging and Valuation with Utility-Based Preferences, PhD Thesis, Berlin University (2001).

[12] Becherer, D.: Rational Hedging and Valuation of Integrated Risks under Constant Absolute Risk Aversion, Insurance: Mathematics and Economics 33, 1-28 (2003).

[13] Bellini, F., Frittelli, M.: On the Existence of Minimax Martingale Measures, Mathematical Finance 12, 1-21 (2002).

[14] Benes, V.E.: Existence of Optimal Stochastic Control Law, SIAM Journal of Control 9, 446-472 (1971).

[15] Biagini, S., Frittelli, M.: On the Super Replication Price of Unbounded Claims, The Annals of Applied Probability 14, 1970-1991 (2004).

[16] Bion-Nadal, J.: Conditional risk measure and robust representation of convex conditional risk measures, Preprint Ecole Polytechnique, nb 557 (2005). (http://www.cmap.polytechnique.fr/preprint/repository/557.pdf).

[17] Bion-Nadal, J.: Pricing functions and risk measures in incomplete markets, Preprint Ecole Polytechnique, nb 577 (2005). (http://www.cmap.polytechnique.fr/preprint/repository/577.pdf).

[18] Borch, K.: Equilibrium in a Reinsurance Market, Econometrica 30, 424-444 (1962).

[19] Borwein, J., Zhu, Q.: Variational Methods in Convex Analysis and Techniques of Variational Analysis, CMS Springer Verlag Books, Springer Verlag, N.Y. (2004).

[20] Briand, P., Carmona, R.: BSDEs with polynomial growth generators, J. Appl. Math. stochastic Anal. 13, 207-238 (2000).

[21] Briand, P., Coquet, F., Hu, Y., Mémin, J., Peng, S.: A converse comparison theorem for BSDEs and related properties of g-expectations, Electron. Comm. Probab. 5, 101-117 (2000).

[22] Briand, P., Delyon, B., Hu, Y., Pardoux, E., Stoica, L.: L^p solutions of Backward Stochastic Diffrential Equations, Stochastic Processes and their Applications 108, 109-129 (2003).

[23] Bühlmann, H.: Mathematical Methods in Risk Theory, Berlin Heidelberg New York: Springer Verlag (1970).

[24] Bühlmann, H.: An Economic Premium Principle, Astin Bulletin 11, 52-60 (1980).

[25] Bühlmann, H.: The General Economic Premium Principle, Astin Bulletin 14, 13-21 (1984).

[26] Bühlmann, H., Delbaen, F., Embrechts, P., Shiryaev, A.: On Esscher Transforms in Discrete Finance Models, Working Paper (2003) (www.math.ethz.ch/~delbaen/ftp/preprints/Esscher-BDES-ASTIN.pdf).

[27] Bühlmann, H., Jewell, W.S.: Optimal Risk Exchanges, Astin Bulletin 10, 243-262 (1979).

[28] Carr, P., Geman, H., Madan, D.: Pricing and Hedging in Incomplete Markets, Journal of Financial Economics 62, 131-167 (2001).

[29] Cheridito, P., Delbaen, F., Kupper, M.: Coherent and Convex Risk Measures for Bounded Cádlaág Processes, Stochastic Processes and their Applications 112, 1-22 (2004).

[30] Cheridito, P., Delbaen, F., Kupper, M.: Dynamic Monetary Risk Measures for Bounded Discrete-time Processes, Working Paper (2004) (http://www.math.ethz.ch/ kupper).

[31] Cheridito, P., Delbaen, F., Kupper, M.: Coherent and Convex Risk Measures for Unbounded Cádlaág Processes, Working Paper (2004) (http://www.math.ethz.ch/ kupper).

[32] Cohen, G.: Convexité et Optimisation. Lecture Notes, ENPC (2000) (http://www-rocq.inria.fr/metalau/cohen/documents/Ponts-cours-A4-NB.pdf).

[33] Cont, R.: Model Uncertainty and its Impact on the Pricing of Derivative Instruments, to Appear in Mathematical Finance (2005).

[34] Coquet, F., Hu, Y., Mémin, J., Peng, S.: Filtration-consistent nonlinear expectations and related g-expectations, Probab. Theory Relat. Fields 123, 1-27 (2002).

[35] Cvitanic, J., Karatzas, I.: On Dynamic Risk Mesaures, Finance and Stochastics 3, 451-482 (1999).

[36] Darling, R.W.R., Pardoux, E.: Backward SDE with random terminal time and applications to semilinear elliptic PDE. Ann. Proba. 25, 1135-1159 (1997).

[37] Davis, M.: Option Pricing in Incomplete Markets; In: Dempster, M.A.H. et al. (eds.): Mathematics of Derivative Securities. Cambridge University Press 1997, pp 227-254.

[38] Delbaen, F.: Coherent Risk Measures, Lecture Notes, Cattedra Galileiana. Scuola Normale Superiore. Classe di Scienze, Pisa (2000).

[39] Delbaen, F., Grandits, P., Rheinländer, T., Samperi, D., Schweizer, M., Stricker, C.: Exponential Hedging and Entropic Penalities, Mathematical Finance 12, 99-123 (2002).

[40] Delbaen, F., Schachermayer, W.: Arbitrage and Free Lunch with Bounded Risk for Unbounded Continuous Processes, Mathematical Finance 4, 343-348 (1994).

[41] Dellacherie, C., Meyer, P.A.: Probabilités et Potentiel, Cahpitres V à VIII, Théorie des Martingales, Hermann (1980).

[42] Deprez, O., Gerber, H.U.: On Convex Principles of Premium Calculation, Insurance: Mathematics and Economics 4, 179-189 (1985).

[43] Detlefsen, K., Scandolo, G.: Conditional and Dynamic Convex Risk Measures, To appear in Finance and Stochastics (2005).

[44] Ekeland, I., Turnbull, T.: Infinite-Dimensional Optimization and Convexity, Chicago University Press (1983).

[45] El Karoui, N., Hamadène, S.: BSDE and Risk-Sensitive Control, Zero-Sum and Non-Zero-Sum Game Problems of Stochastic Functional Differential Equations, Stochastic Processes and their Applications 107, 145-169 (2003).

[46] El Karoui, N., Kapoudjian, C., Pardoux, E., Peng, S., Quenez, M.C.: Reflected Solutions of Backward SDE and Related Obstacle Problems for PDEs, Annals of Probability 25, 702-737 (1997).

[47] El Karoui, N., Peng, S., Quenez, M.C.: Backward Stochastic Differential Equations in Finance, Mathematical Finance 7, 1-71 (1997).

[48] El Karoui, N., Peng, S., Quenez, M.C.:A Dynamic Maximum Principle for the Optimization of Recursive Utilities under Constraints, Ann. Appl. Probab. 11, 664-693 (2001).

[49] El Karoui, N., Quenez, M.C.: Dynamic Programming and Pricing of Contingent Claims in an Incomplete Market, SIAM Journal of Control and Optimization 33, 29-66 (1995).

[50] El Karoui, N., Quenez, M.C.: Non-linear Pricing Theory and Backward Stochastic Differential Equations in Financial Mathematics (ed: W.J. Runggaldier) Lecture Notes in Mathematics 1656, Springer Verlag, 191-246 (1996).

[51] Rouge, R., El Karoui, N.: Pricing via Utility Maximization and Entropy, Mathematical Finance 10, 259-276 (2000).

[52] Föllmer, H., Kramkov, D.: Optional decompositions under constraints, Probability Theory and Related Fields 109, 1-25 (1997).

[53] Föllmer, H., Schied, A. Convex Measures of Risk and Trading Constraints, Finance and Stochastics 6, 429-447 (2002).

[54] Föllmer, H., Schied, A.: Stochastic Finance: An Introduction in discrete Time, De Gruyter Studies in Mathematics 2002 (revised edition 2004).

[55] Frittelli, M.: The Minimal Entropy Martingale Measure and the Valuation in Incomplete Markets, Mathematical Finance 10, 39-52 (2000).

[56] Frittelli, M.: Introduction to a Theory of Value Coherent with the No-Arbitrage Principle, Finance and Stochastics 4, 275-297 (2000).

[57] Frittelli, M., Gianin, E.R.: Putting Order in Risk Measures, Journal of Banking and Finance 26, 1473-1486 (2002).

[58] Frittelli, M., Gianin, E.R.:Dynamic Convex Risk Measures in Risk Measures for the 21st Century (ed: G. Szegö) Wiley Finance, John Wiley ans Sons (2004).

[59] Frittelli, M., Scandolo, G.: Risk Measures and Capital Requirement for Processes, Working Paper (2005) (http://www.gloriamundi.org/ShowTracking.asp?ResourceID=453057457").

[60] Gerber, H.U.: Risk Exchange Induced by an External Agent, Transactions of the International Congress of Actuaries, 385-392 (1980).

[61] Gianin, E.R.: Some Examples of Risk Measures via g-Expectations, Quaderno del Dipartemento di Metodi Quantitativi, Universita di Milano Bicocca (Italy) (July 2002).

[62] Grandits, P., Rheinländer, T.: On the Minimal Entropy Martingale Measure, Annals of Probability 30, 1003-1038 (2002).

[63] Hamadène, S.: Equation Différentielles Stochastiques Rétrogrades: Le Cas Localement Lipschitzien, Ann. Inst. Henri Poincaré 32, 645-659 (1996).

[64] Hamadène, S.: Reflected BSDE's with discontinuous Barrier and Applications, Stoch. and Stoc. Reports 74, 571-596 (2002).

[65] Hamadène, S., Lepeltier, J.P.: Zero-Sum Stochastic Differential Games and Backward Equations, Systems and Control letters 24, 259-263 (1995).

[66] Hamadène, S., Lepeltier, J.P.: Backward Equations, Stochastic Control and Zero-Sum Stochastic Differential Games, Stochastic and Stochastic Reports 54, 221-231 (1995).

[67] Hamadène, S., Lepeltier, J.P., Matoussi, A.: Double Barrier Backward SDEs with Continuous Coefficient, in Backward Stochastic Differential Equations (eds: N. El-Karoui and L. Mazliak) Pitman Research Notes in Mathematics Series 364 (1997).

[68] Hamadène, S., Lepeltier, J.P., Peng, S.: BSDE with Continuous Coefficients and Application to Markovian Nonzero Sum Stochastic Differential Games, in Backward Stochastic Differential Equations (eds: N. El-Karoui and L. Mazliak) Pitman Research Notes in Mathematics Series 364 (1997).

[69] Henderson, V.: Valuation of Claims on Non-Traded Assets using Utility Maximization, Mathematical Finance 12, 351-373 (2002).

[70] Hiriart-Urruty, J.B., Lemarchal, C.: Fundamentals of Convex Analysis, Springer Verlag, Grundlehren Text Edition (2001).

[71] Hobson, D.: Real Options, Non-Traded Assets and Utility Indifference Prices, Department of Mathematical Sciences, University of Bath (2003). (http://www.bath.ac.uk/ masdgh/publications.html).

[72] Hodges, S.D., Neuberger, A.: Optimal Replication of Contingent Claims under Transaction Costs, Review of Futures Markets 8, 222-239 (1989).

[73] Hu, Y., Imkeller, P., Müller, M.: Utility Maximization in Incomplete Markets, Working Paper (2003) (http://wws.mathematik.hu-berlin.de/~imkeller/research/papers/utilityHIM.pdf).

[74] Jiang, L.: Non-Linear - g-Expectation Theory and Its Applications in Finance, PhD Thesis, Shandong University, Jinan (2005).

[75] Kabanov, Y.M., Stricker, C.: On the Optimal Portfolio for the Exponential Utility Maximization: Remarks to the Six-Author. Mathematical Finance 12, 125-134 (2002).

[76] Karatzas, I., Kou, S.: On the pricing of contingent claims under constraints, Annals of Applied Probability 6, 321-369 (1996).

[77] Kazamaki, N.: Continuous Exponential Martingales and BMO, Lecture Notes 1579, Springer-Verlag (1994).

[78] Klöppel, S., Schweizer, M.: Dynamic Utility Indifference Valuation via Convex Risk Measures, Working Paper (2005) (http://www.nccr-finrisk.unizh.ch/media/pdf/wp/WP209-1.pdf).

[79] Kobylanski, M.: Backward Stochastic Differential Equations and Partial Differential Equations with Quadratic Growth, Annals of Probability 28, 558-602 (2000).

[80] Kobylanski, M., Lepeltier, J.-P., Quenez, M.C., Torres, S.: Reflected BSDE with Superlinear Quadratic Coefficient, Probab. and Mathematical Statis. 22, 51-83 (2002).

[81] Kohlman, M., Tang, S.: Minimization of Risk and Linear Quadratic Optimal Control Theory, SIAM Journal on Control and Optimization 42, 1118-1142 (2003).

[82] Komlos, J.: A generalisation of a problem of Steinhaus, Acta Math. Acad. Sci. Hungar. 18, 217-229 (1967).

[83] Kusuoka, S.: On Law Invariant Coherent Risk Measures in Advances in Mathematical Economics 3, 83-95, Springer Verlag (2001).

[84] Lepeltier, J.P., Matoussi, A., Xu M.: Reflected BSDEs under Monotonocity and General increasing Growth Conditions, Preprint Université du Maine, France (2004).

[85] Lepeltier, J.P., San Martin, J.: Backward Stochastic Differential Equations with Continuous Coefficient, Statist. Probab. lett. 32, 425-430 (1997).

[86] Lepeltier, J.P., San Martin, J.: Existence of BSDE with Superlinear Quadratic Coefficient, Stochastics and Stochastic Report 63, 227-240 (1998).

[87] Lepeltier, J.P., San Martin, J.: On the Existence or Non-Existence of Solutions for Certain Backward Stochastic Differential Equations, Bernoulli 8, 123-137 (2002).

[88] Mania, M., Santacroce, M., Tevzadze, R.: A Semimartingale BSDE Related to the Minimal Entropy Martingale Measure, Finance and Stochastics 7, 385-402 (2003).

[89] Mania, M., Schweizer, M.: Dynamic Exponential Utility Indifference Valuation, To appear in the Annals of Applied Probability (2005).

[90] Mania, M., Tevzade, R.: A Semimartingale Bellman Equation and the Variance-Optimal Martingale Measure, Georgian Mathematical Journal 7, 765-792; Correction, Georgian Mathematical Journal 9, 197-198 (2002).

[91] Matoussi, A. : Reflected solutions of Backward Stochastic Differential Equations, Statist. and Probab. Lett. 34, 347-354 (1997).

[92] Musiela, M., Zariphopoulou, T.: An Example of Indifference Prices under Exponential Preferences, Finance and Stochastics 8, 229-239 (2004).

[93] Musiela, M., Zariphopoulou, T.: A Valuation Algorithm for Indifference Prices in Incomplete Market, Finance and Stochastics 8, 399-414 (2004).

[94] Pardoux, E.: BSDEs, Weak Convergence and Homogenization of Semilinear PDEs, in Nonlinear Analysis, Differential Equations and Control (Montreal, QC, 1998), 503-549, NATO Sci. Ser. C Math. Phys. Sci., 528, Kluwer Acad. Publ., Dordrecht (1999).

[95] Pardoux, E., Peng, S.: Adapted Solution of a Backward Stochastic Differential Equation, Systems and Control Letters 14, 55-61 (1990).

[96] Peng, S.: Backward SDE and Related g-Expectations, in Backward Stochastic Diffential Equations in Backward Stochastic Differential Equations (eds: N. El-Karoui and L. Mazliak) Pitman Research Notes in Mathematics Series 364 (1997).

[97] Peng, S.: Monotonic Limit Theorem of BSDE and Nonlinear Decomposition Theorem of Doob-Meyer's Type, P.T.R.F. 38, 119-134 (1999).

[98] Peng, S.: Nonlinear Expectations, Nonlinear Evaluations and Risk Measures, Lecture Notes in Mathematics 1856, Springer Verlag (2004).

[99] Raviv, A.: The Design of an Optimal Insurance Policy, American Economic Review 69, 84-96 (1979).

[100] Rees, R.: The Theory of Principal and Agent, part I, Bulletin of Economic Research 37:1, 3-26 (1985).

[101] Rees, R.: The Theory of Principal and Agent, part II, Bulletin of Economic Research 37:2, 75-95 (1985).

[102] Riedel, F.: Dynamic Coherent Risk Measures, Stochastic Processes and their Applications 112, 185-200 (2004).

[103] Rockafellar, R.T.: Convex Analysis, Princeton Landmarks in Mathematics 1970.

[104] Rockafellar, R.T., Uryasev, S.: Optimization of Conditional Value at Risk, Journal of Risk 2, 21-41 (2000).

[105] Rockafellar, R.T., Uryasev, S.: Conditional Value at Risk for General Loss Distributions, Journal of Banking and Finance 26, 1443-1471 (2002).

[106] Roorda, B., Schumacher, H. Engwerda, J.: Coherent Acceptability Measures in Multiperiod Models. To Appear in Mathematical Finance (2005).

[107] Scandolo, G.: Risk Measures in a Dynamic Setting, PhD Thesis, Milano University (2003).

[108] Schachermayer, W.: Optimal Investment in Incomplete Financial Markets. In Geman, H. et al. (eds.): Mathematical Finance Bachelier Congress 2000. Berlin Heidelberg New York: Springer Verlag, 2002, pp 427-462.

[109] Schied, A.: Risk Measures and Robust Optimization Problems, Springer Verlag, Lecture Notes, To Appear in 2005.

[110] Sekine, J.: An Approximation for Exponential Hedging, Proceedings of the Symposium Stochastic Analysis and Related Topics, RIMS Tokyo, 2002.

[111] Wang, T.: A Class of Dynamic Risk Measures, Working Paper (1999) (http://finance.commerce.ubc.ca/research/papers/UBCFIN98-5.pdf).

[112] Weber, S.: Distribution-Invariant Dynamic Risk Measures, Working Paper (2003) (http://www.math.hu-berlin.de/ finance/papers/weber.pdf).