

# On robust portfolio and naïve diversification: mixing ambiguous and unambiguous assets

A. Burak Paç<sup>1</sup> · Mustafa Ç. Pınar<sup>1</sup>

Published online: 29 August 2017  
© Springer Science+Business Media, LLC 2017

**Abstract** Effect of the availability of a riskless asset on the performance of naïve diversification strategies has been a controversial issue. Defining an investment environment containing both ambiguous and unambiguous assets, we investigate the performance of naïve diversification over ambiguous assets. For the ambiguous assets, returns follow a multivariate distribution involving distributional uncertainty. A nominal distribution estimate is assumed to exist, and the actual distribution is considered to be within a ball around this nominal distribution. Complete information is assumed for the return distribution of unambiguous assets. As the radius of uncertainty increases, the optimal choice on ambiguous assets is shown to converge to the uniform portfolio with equal weights on each asset. The tendency of the investor to avoid ambiguous assets in response to increasing uncertainty is proven, with a shift towards unambiguous assets. With an application on the *CVaR* risk measure, we derive rules for optimally combining uniform ambiguous portfolio with the unambiguous assets.

**Keywords** Naïve diversification · Robust portfolio optimization · Ambiguous and unambiguous assets · Conditional Value-at-Risk · Worst-case risk measures

**Mathematics Subject Classification** 91G10 · 90C15 · 90C90

## 1 Introduction

Uncertainty in the asset return distribution has received considerable attention in the portfolio selection literature. Ambiguity in the type of distribution for the asset returns has been

---

✉ Mustafa Ç. Pınar  
mustafap@bilkent.edu.tr

A. Burak Paç  
burakpac@gmail.com

<sup>1</sup> Department of Industrial Engineering, Bilkent University, Ankara, Turkey

addressed in several studies, leading to distributionally robust portfolio selection. Often, the information regarding asset returns consists of the moment estimates, and parametric uncertainty approaches robustify the portfolio selection against uncertainty in the moment estimates. For a review on robust optimization and portfolio selection with coverage of parametric uncertainty applications, see e.g., [Bertsimas et al. \(2011\)](#), [Fabozzi et al. \(2007\)](#), [Fabozzi et al. \(2010\)](#), [Pachamanova \(2013\)](#), [Guidolin and Rinaldi \(2013\)](#). In this study, we adopt a non-parametric model of distribution uncertainty rather than uncertainty based on known moments, albeit with some error in the moment information. Rather than an ambiguous distribution type given mean/covariance information, we assume that a nominal multivariate probability distribution for asset returns is known. The uncertainty set is defined as a ball around this nominal distribution, and the probability metric used for defining the ball is the Kantorovich distance (also called the Wasserstein metric).

While adopting the uncertainty model introduced by [Pflug et al. \(2012\)](#), a specific investment strategy is in the focus of our studies: naïve diversification, i.e., investing in all assets with equal shares of the wealth. In [Pflug et al. \(2012\)](#), it is demonstrated that when uncertainty, i.e., the radius of the ball, increases, the worst-case risk minimizing portfolio converges to the uniform portfolio (also called the equally weighted,  $1/N$  or naïvely diversified portfolio), in which each asset receives equal portions of the wealth. This result is proven for any risk measure chosen from a class of convex, law invariant functionals. Following the result in [Pflug et al. \(2012\)](#), we investigate optimality of the uniform portfolio for assets involving distributional uncertainty (referred to as ambiguous assets/market), in the existence of an alternative market to invest in—a group of assets whose return distribution is known with full information (the unambiguous assets/market).

[Benartzi and Thaler \(2001\)](#) point out that naïve diversification is a common practice both as a general heuristic of choice and an investment strategy. Behavioral experiments indicate that subjects asked to choose multiple items from a list of possible selections simultaneously, tend to diversify their decisions, i.e., they pick as diverse a group of items as possible. Similar behavioral experiments investigate decisions on investment plans. Employees presented with a fictive mix of assets for their retirement saving plans predominantly use the naïve diversification strategy, where the ratio of the total amount invested in stocks/bonds is in strong correlation with the ratio of the number of stocks/bonds in the asset mix. These behavioral studies are supported by investor behavior elicited from archives of investment history: prevalence of the asset type in the portfolio determines the total allocation to that type. Such behavior is deemed to contradict rational choice, i.e., optimized diversification based on portfolio models, such as the mean-variance portfolio. In addition to contradicting optimized portfolio rules, the psychological bias towards naïve diversification can be considered irrational since it is difficult to imagine that a rational model fits the diverse preferences of people expressing this choice. However, [DeMiguel et al. \(2009b\)](#) point out that the anticipation of some form of uncertainty in the environment might be intuitively leading people to naïve diversification. The authors study 14 models, mostly the mean-variance model and its extensions, testing on 7 real market data sets. The result is fascinating: none of the 14 models investigated consistently outperform naïve diversification in terms of out-of-sample Sharpe ratio, certainty-equivalent return (CER) and turnover measures. Parameter estimation errors outweigh the theoretical gains promised by models, and the authors point out the need for an unrealistic amount of data for the models to perform better than the  $1/N$  rule. The result is significant, providing justification based on real market data for the fact that naïve diversification is not necessarily irrational, being hard to outperform. With growing interest in  $1/N$  portfolio, many studies follow both supporting this view, or challenging naïve diversification under different settings and with more sophisticated portfolio strategies.

Murtazashvili and Vozlyublennaya (2013) support the claim that naïve diversification outperforms enhancements on mean-variance portfolio rules, comparing out-of-sample mean to standard deviation ratios. A similar result is reported by Jacobs et al. for global equity diversification and diversification over asset classes (Jacobs et al. 2014). Again, out-of-sample performance of naïve diversification is either better or as good, compared to the shortfall minimizing portfolio (Haley 2016). Brown et al. (2013) also confirm that optimal diversification can not consistently outperform naïve diversification, considering out-of-sample Sharpe ratio, CER and turnover measures, but remark that performance of  $1/N$  comes at a cost: increase in the tail risk as measured by skewness and kurtosis, and increased concavity of the return distribution. Frahm et al. (2012) test the results in DeMiguel et al. (2009b) using out-of-sample multiple testing techniques, and combining naïve diversification with risk-free investment. The study also considers robust portfolio strategies and minimum variance portfolio rules. Extensions of minimum variance strategies are shown to (slightly) outperform trivial strategies (strategies formed by combination of  $1/N$  risky portfolio and the riskless asset in different ratios). The share riskless asset receives in the trivial strategy considerably affects the performance, and it is emphasized that including a riskless asset in the setting is critical for the comparison of optimized diversification strategies and naïve diversification. There are many studies countering the efficiency of naïve diversification, testing in different settings, or offering more sophisticated portfolio selection strategies. Cesarone et al. (2016) modify known portfolio models adding cardinality constraints to restrict number of assets in the portfolio. This seems to work, as the out-of-sample performance of (unrestricted)  $1/N$  is inferior in terms of return standard deviation and Sharpe ratio compared to most models and for most data sets. Similarly, Behr et al. (2013) devise a constrained minimum variance portfolio model with shrinkage. The model celebrates a consistent and significant improvement in out-of-sample Sharpe ratio over naïve diversification. Similarly, higher Sharpe ratio and Omega measures and lower volatility than the  $1/N$  portfolio indicates superiority of the sample based version of the Black–Litterman model (1992) developed by Bessler et al. (2017). Fugazza et al. (2015) emphasize two aspects that are critical for the superiority of naïve diversification: the investment horizon (1 month) and asset types in the portfolio. The linear vector autoregressive strategy offered can improve out-of-sample CER and Sharpe ratio when real estate assets are available and the investment horizon is longer than a year. DeMiguel et al. (2009a) provide a norm-constrained minimum-variance portfolio selection framework achieving higher out-of-sample Sharpe ratios. Tu and Zhou (2011) form strategies as a hybrid of naïve diversification and the three fund strategy of Kan and Zhou (2007), optimizing their combination. The resulting portfolio rule outperforms the pure  $1/N$  strategy. Kirby and Ostdiek (2012) show that the mean-variance strategies can be enhanced with timing strategies, and outperform naïve diversification. Fletcher (2011) tests the models in Tu and Zhou (2011) and Kirby and Ostdiek (2012) with U.K. stock return data and confirms superiority over  $1/N$ .

The study by Pflug et al. (2012) has a unique position in the debate. Studies aforementioned are empirical, focusing on the effect of parameter estimation errors and deciding based on out-of-sample statistics. Pflug et al. (2012), on the other hand, construct a mathematical framework to model uncertainty, and state that naïve diversification stands as an efficient portfolio strategy not because optimization is unviable under uncertainty, proving the optimality of  $1/N$  strategy in this case. Given a nominal distribution for asset returns, they define the distribution uncertainty by a ball of probability distributions around the nominal distribution. It is proven that as the radius of the ball, i.e. the level of uncertainty, increases, the optimal investment vector converges to the uniform investment. The choice of metric is important in this result. Kantorovich distance has stronger convergence properties (in

unbounded spaces), being bounded from below by the square of Prokhorov metric (Gibbs and Su 2002), which metrizes weak convergence on any separable metric space. The convergence of optimal investment to  $1/N$  based on Kantorovich distance does not generally imply the result if Prokhorov metric is used. Indeed, robust portfolio models assuming non-parametric distribution ambiguity are studied with other metrics. Calafiore (2007), defines uncertainty based on a nominal discrete distribution and the Kullback–Leibler divergence. Under this model, worst-case mean-variance and mean-absolute deviation risk measures are studied. In the former, an interior point barrier method in conjunction with an analytic center cutting plane technique is used, and in the latter case, a line search algorithm is incorporated to the solution procedure, additionally. In Erdoğan and Iyengar (2006), asset return distribution is considered to lie within a ball based on Prokhorov metric, and an ambiguous chance constrained problem is defined. The problem is approximated by robust sampling of probability distributions, which results in problem formulations having the same complexity as the nominal problem with certain distribution.

In this study, we adopt the framework in Pflug et al. (2012), and add a group of assets with known return distribution to the investment environment. The group of assets can be considered to represent an alternative market where the investor can invest part of the initial wealth, for instance a developed market allowing reliable estimation of the return distribution for a certain group of assets. The group of assets with ambiguous return distribution, on the other hand, can be considered to represent an emerging market with rapid growth promising high return rates, while the lack of consolidation and historical data might pose difficulty to the estimation of asset return information. Since the key factor for the optimality of the uniform portfolio is the increasing level of uncertainty, by introducing the alternative assets, we assess the willingness to invest in the ambiguous market under high uncertainty. We first extend the result in Pflug et al. (2012) showing that optimal investment in the ambiguous market converges to uniform portfolio for both positive and negative allocations to this market, after fixing the allocation to the unambiguous assets. While having this convergence effect, we show that increasing uncertainty may cause the investor to steer away from the ambiguous market. With this analysis we aim to shed light on the desirability of investing in environments where the naïve diversification heuristic becomes an optimal strategy. Indeed, in such an environment, allocation to ambiguous assets diminishes as the radius of uncertainty tends to infinity. Nevertheless, optimality of the uniform portfolio persists while the total amount allocated to the ambiguous assets diminishes. This mathematically describes investor behavior in avoiding uncertainty by pulling out of the ambiguous market, and explains the phenomenon addressed, for instance, by Kan et al. (2016): an alternative/riskless asset in the environment significantly changes the performance of optimal portfolio strategies and promotes them to consistently outperform naïve diversification. If naïve diversification is also combined with a riskless asset, such as the 25/50/75% naïve and riskless portfolio combinations in Frahm et al. (2012), a similar performance effect of exploiting a riskless asset is also observed on the naïve diversification strategy. Our study mathematically validates the optimality of a hybrid strategy combining riskless/unambiguous portfolio with uniform investment into ambiguous assets. Using *CVaR* as the risk measure, we derive rules to optimally combine the uniform ambiguous portfolio with the portfolio selection on unambiguous assets. With the uncertainty structure in Pflug et al. (2012), a riskless asset can not be defined in the asset collection, and with the extension in this study, it is possible to model riskless assets as a special case, to seek optimal combination of risky naïve portfolio and the riskless asset.

To the best of our knowledge, our study is unique in offering a mathematical model that brings together in an investment environment assets involving distribution ambiguity and assets with complete information on return distribution. In this environment, a decision

heuristic is promoted on one side, while conditions on the other side allow for optimized diversification. We prove that increasing uncertainty renders naïve diversification an optimal investment strategy for the ambiguous assets in combination with an optimized portfolio over the unambiguous assets. We prove divestment of the ambiguous market as the uncertainty level tends to infinity. While our convergence results are valid for a class of convex, law invariant risk functionals, we consider the Conditional Value-at-Risk (*CVaR*) measure as a core representative of this class and derive rules for optimally combining uniform ambiguous portfolio with optimal unambiguous portfolio. Our model covers the combination of a riskless asset with the ambiguous assets as a special case. This provides a clear justification for the optimality of portfolios that combine naïve diversification with a riskless asset. With the addition of a riskless asset to the environment, naïve diversification continues to be the optimal strategy on ambiguous assets, and this is a result based on a single factor: the level of uncertainty in the ambiguous market—not the existence of alternative assets in the environment.

The organization of this paper is as follows: In Sect. 2 the problem definition is provided, presenting the risk measure and uncertainty structure adopted in the model. In Sect. 3 we present the results establishing that investment in the ambiguous market converges to uniformity with increasing ambiguity level, fixing allocation to unambiguous assets. The result on diminishing investment in ambiguous market with increasing uncertainty follows. In Sect. 4, we derive rules for optimally combining unambiguous and uniform ambiguous portfolio, with (worst-case) *CVaR* as the risk measure. An example with the Markowitz functional and a riskless asset demonstrates the loss of interest in the ambiguous market when the uncertainty level is high enough to promote naïve diversification. For a more fluent presentation of the main results therein, the proofs of the results in Sect. 3 are collected in “Appendix”.

## 2 Problem definition

In a market involving ambiguity in asset return distributions, there are  $N$  assets, with return distribution  $Q_1$  on  $\mathbb{R}^N$ , which is known to be inside an open ball around a known nominal distribution  $P_1$  on  $\mathbb{R}^N$ . The definition of the ball of radius  $\kappa$ ,  $\mathcal{B}_\kappa(P_1)$ , is based on the Kantorovich/Wasserstein distance of degree  $p$ :

$$d_p(P_1, Q_1) = \inf_{\pi} \left\{ \left( \int_{\mathbb{R}^N \times \mathbb{R}^N} \|x - y\|_p^p d\pi(x, y) \right)^{\frac{1}{p}} : \text{proj}_1(\pi) = P_1, \text{proj}_2(\pi) = Q_1 \right\},$$

as defined for two measures  $P_1$  and  $Q_1$  on  $\mathbb{R}^N$  (for further discussion on Kantorovich/Wasserstein metrics see Chapter 6 in Villani (2008)). *proj* denotes projection of the measure, i.e., marginal distributions in the first and last  $N$  dimensions with  $P_1(A) = \pi(A \times \mathbb{R}^N)$ ,  $Q_1(B) = \pi(\mathbb{R}^N \times B)$  for all members  $A$  and  $B$  of the Borel  $\sigma$ -algebra  $\mathcal{B}(\mathbb{R}^N)$ .  $\pi$  is a measure on  $\mathbb{R}^N \times \mathbb{R}^N$ , with marginals  $P_1$  and  $Q_1$ . The infimum is known to be attained by some distribution  $\pi$  on  $\mathbb{R}^N \times \mathbb{R}^N$  (Villani 2003), and the minimizing distribution is called the optimal transportation plan between  $P_1$  and  $Q_1$ . When necessary, superscripts are used to clarify the space the Kantorovich distance is defined on, such as  $d_p^{\mathbb{R}^N}$  or  $d_p^{\mathbb{R}^{N+L}}$ .

Based on the uncertainty set  $\mathcal{B}_\kappa(P_1)$ , a robust portfolio selection problem arises, and is defined by Pflug et al. (2012) as follows:

$$\inf_{w \in \mathbb{R}^N} \sup_{Q_1 \in \mathcal{B}_\kappa(P_1)} \mathcal{R}(\langle X^{Q_1}, w \rangle) \tag{1}$$

$$s.t. \quad \langle \mathbb{1}, w \rangle = 1. \tag{2}$$

Here,  $X^{Q_1} : \Omega \rightarrow \mathbb{R}^N$  is a random variable in  $L^p(\Omega, \Sigma, \mu)$  with image measure  $Q_1$ , that is,  $\mu \circ (X^{Q_1})^{-1}(A) = Q_1(A)$  for all  $A \in \mathcal{B}(\mathbb{R}^N)$ .

The risk functional  $\mathcal{R} : L^p(\Omega, \Sigma, \mu) \rightarrow \mathbb{R}$  is convex, law invariant and can be stated in dual form as:

$$\mathcal{R}(X) = \max \{ \mathbb{E}(XZ) - R(Z) : Z \in L^q \}.$$

Here,  $R : L^q(\Omega, \Sigma, \mu) \rightarrow \mathbb{R}$  is a convex functional. Convexity in this context requires for random variables  $X, Y \in L^p(\Omega, \Sigma, \mu)$  that  $\mathcal{R}(\lambda X + (1 - \lambda) Y) \leq \lambda \mathcal{R}(X) + (1 - \lambda) \mathcal{R}(Y)$  for all  $\lambda \in (0, 1)$ . Law invariance, also known as version independence, means that the value of  $\mathcal{R}(X)$  depends on the law, i.e., the probability distribution, of  $X$  and not on the specific mapping of outcomes in  $\Omega$  to the range of  $X$ . For any other random variable  $Y$  with the same image measure  $\mu \circ Y^{-1} = \mu \circ X^{-1}$ , thus distribution,  $\mathcal{R}(Y)$  is equal to  $\mathcal{R}(X)$ . The domain of the risk functional  $\mathcal{R}$  determines the degree  $p$  of the Kantorovich metric. At each  $X \in L^p(\Omega, \Sigma, \mu)$ , by abuse of notation we denote  $\arg \max_{Z \in L^q} \{ \mathbb{E}(XZ) - R(Z) \}$  by the set of subgradients  $\partial \mathcal{R}(X)$ , although these sets might not coincide in general (see Pflug et al. 2012; Romisch and Pflug 2007).

As the uncertainty radius  $\kappa$  increases, the optimal solution of (1)–(2) either turns equal to or otherwise converges to the uniform portfolio  $w^u = \frac{1}{N} \mathbb{1}$  (we use  $\mathbb{1}$  to denote both characteristic functions and vectors composed of ones, discrimination to be made and dimension to be inferred by context). In this problem, the investor is forced to invest entire wealth to the market of  $N$  assets. The decision in this case is not informative on whether he is really interested in entering the market under high uncertainty. By introducing an alternative market with unambiguous asset return distribution, we try to assess whether the investor is willing to invest in, and with how large an allocation for the ambiguous market, under high uncertainty levels that theoretically prescribe naïve diversification therein. The alternative market consists of  $L$  assets with known return distribution. In this study, the correlation of the two markets is not incorporated in the model, thus the distribution  $Q$  of  $N + L$  assets together is product formed, i.e.,  $Q = Q_1 \times P_2$ , where  $P_2$  is the return distribution of the assets in the second market. The uncertainty set for the probability distribution is defined as:

$$\tilde{\mathcal{B}}_\kappa(P) = \left\{ Q : Q = Q_1 \times P_2, \quad d_p^{\mathbb{R}^{N+L}}(P, Q) \leq \kappa \right\},$$

where  $P = P_1 \times P_2$  is the nominal distribution for the entire collection of  $N + L$  assets,  $P, Q$  are Borel probability measures on  $\mathbb{R}^{N+L}$ ,  $P_1, Q_1$  are Borel probability measures on  $\mathbb{R}^N$ , and  $P_2$  is a Borel probability measure on  $\mathbb{R}^L$ . Since  $Q$  is product formed, the condition  $d_p^{\mathbb{R}^{N+L}}(P, Q) \leq \kappa$  is equivalent to  $d_p^{\mathbb{R}^N}(P_1, Q_1) \leq \kappa$  (as discussed in Lemma 3) but the preferred definition has advantages in representing the relationship between the distribution  $Q$  on  $(\mathbb{R}^{N+L}, \mathcal{B}(\mathbb{R}^{N+L}))$  and the random variable  $X^Q : \Omega \rightarrow \mathbb{R}^{N+L}$ . Our problem is redefined as follows:

$$\inf_{\substack{w \in \mathbb{R}^N \\ v \in \mathbb{R}^L}} \sup_{Q \in \tilde{\mathcal{B}}_\kappa(P)} \mathcal{R} \left( \left\langle X^Q, \begin{bmatrix} w \\ v \end{bmatrix} \right\rangle \right) \tag{3}$$

$$s.t. \quad \langle \mathbb{1}, w \rangle + \langle \mathbb{1}, v \rangle = 1. \tag{4}$$

We assume that the nominal problem corresponding to (3)–(4) defined as

$$\inf_{\substack{w \in \mathbb{R}^N \\ v \in \mathbb{R}^L}} \mathcal{R} \left( \left\langle X^P, \begin{bmatrix} w \\ v \end{bmatrix} \right\rangle \right) \tag{5}$$

$$s.t. \quad \langle \mathbb{1}, w \rangle + \langle \mathbb{1}, v \rangle = 1 \tag{6}$$

is well-posed, being bounded below with an optimal solution.

We first fix the amount  $v$  invested into the unambiguous assets, to prove that the optimal choice on ambiguous assets with the remaining wealth converges to the uniform portfolio  $w^{u,v} = \frac{1-\langle \mathbb{1}, v \rangle}{N} \mathbb{1}$ . Fixing  $v$  at a value in  $\mathbb{R}^L$ , we define an inner problem as follows:

$$\inf_{w \in \mathbb{R}^N} \sup_{Q \in \tilde{\mathcal{B}}_\kappa(P)} \mathcal{R} \left( \langle X^{Q_1}, w \rangle + \langle X^{Q_2}, v \rangle \right) \tag{7}$$

$$s.t. \quad \langle \mathbb{1}, w \rangle = 1 - \langle \mathbb{1}, v \rangle. \tag{8}$$

Given a distribution  $Q$  on  $(\mathbb{R}^{N+L}, \mathcal{B}(\mathbb{R}^{N+L}))$ , a random variable  $X^Q : (\Omega, \Sigma, \mu) \rightarrow (\mathbb{R}^{N+L}, \mathcal{B}(\mathbb{R}^{N+L}))$  exists with image measure  $Q$  (see Lemma 2 in Pflug et al. 2012), i.e., for all  $A \in \mathcal{B}(\mathbb{R}^{N+L})$ ,  $\mu \circ (X^Q)^{-1}(A) = Q(A)$ . Since  $\mathcal{R}$  is law invariant, we legitimately use random variables (e.g.,  $X^P$ ) in computations corresponding to Borel measures defined on  $\mathbb{R}^N, \mathbb{R}^L, \mathbb{R}^{N+L}$  (e.g.,  $P$ ). A remark regarding notation abuse is the usage of indexing in  $X^{Q_1}, X^{Q_2}$ . Former represents the first  $N$  entries of the  $N + L$  dimensional random variable  $X^Q$  and latter the last  $L$  entries. Then, image measures of  $X^{Q_1}$  and  $X^{Q_2}$  coincide with the marginals  $Q_1$  and  $Q_2$  of  $Q$  on the first  $N$  and last  $L$  dimensions, respectively (for all  $A \in \mathbb{R}^N$ ,  $\mu \circ (X^{Q_1})^{-1}(A) = Q(A \times \mathbb{R}^L) = Q_1(A) \times Q_2(\mathbb{R}^L) = Q_1(A)$ ). Via this notation we write  $\langle X^Q, \begin{bmatrix} w \\ v \end{bmatrix} \rangle = \langle X^{Q_1}, w \rangle + \langle X^{Q_2}, v \rangle$ . We denote by  $\sigma(X^{P_i}) \subset \Sigma$  the smallest sigma algebra with respect to which  $X^{P_i}$  is measurable,  $i = 1, 2$ , and by  $\mathcal{F}^\perp$  the largest sigma algebra independent from  $\sigma(X^{P_2})$ . Note that  $\sigma(X^{P_1}) \subset \mathcal{F}^\perp$ , if  $X^{P_1}$  and  $X^{P_2}$  are independent.

### 3 Uniform investment and divestment of ambiguous market

To prove (approximate) optimality of  $w^{u,v}$  for (7)–(8), some preliminary results are necessary. The first step for solution is to compute

$$\sup_{Q \in \tilde{\mathcal{B}}_\kappa(P)} \mathcal{R} \left( \langle X^{Q_1}, w \rangle + \langle X^{Q_2}, v \rangle \right)$$

given  $P, v$  and  $w$ . In the following proposition, we present the computation. Proofs are deferred until ‘‘Appendix’’ to present the main results in this section earlier and in a more fluent manner. Proofs contain similar lines to those in Pflug et al. (2012), and differences arise mainly due to the steps necessary to accommodate the additional unambiguous return term in the return/loss function and to assure that the ambiguous distribution  $Q$  is product formed.

**Proposition 1** *Let  $\mathcal{R} : L^P(\Omega, \Sigma, \mu) \rightarrow \mathbb{R}$  be a convex, law invariant risk measure. Let  $P = P_1 \times P_2$  be a probability measure on  $\mathbb{R}^{N+L}$  with  $P_1, P_2$  probability measures on  $\mathbb{R}^N, \mathbb{R}^L$ , respectively. Let  $X^P \in L^P(\Omega, \Sigma, \mu)$  be a random variable with image measure  $P$ , and*



$\mathcal{F}^\perp \subset \Sigma$  be the largest  $\sigma$ -algebra independent from  $\sigma(X^{P_2})$ . Let  $p = 1$  ( $q = \infty$  in this case) and assume

$$\|Z\|_{L^\infty} = C, \mu(\{\omega \in \Omega : |Z|(\omega) \notin \{0, C\}\}) = 0 \tag{9}$$

for all  $Z \in \partial\mathcal{R}(X)$ ,  $X \in L^1(\Omega, \Sigma, \mu)$ . In addition, assume for all  $\epsilon \in (0, \frac{1}{2})$  that there exists  $B \in \mathcal{F}^\perp$  such that  $\mu(B) > 0$ , and either

$$\mu(B \cap \{\omega \in \Omega : Z(\omega) = C\}) > (1 - \epsilon)\mu(B)$$

or

$$\mu(B \cap \{\omega \in \Omega : Z(\omega) = -C\}) > (1 - \epsilon)\mu(B)$$

holds.

Or, let  $1 < p < \infty$ ,  $q$  be such that  $\frac{1}{p} + \frac{1}{q} = 1$ , and assume

$$\|\mathbb{E}[Z|\mathcal{F}^\perp]\|_{L^q} = C \text{ for all } Z \in \bigcup_{X \in L^p} \partial\mathcal{R}(X) \text{ with } R(Z) < \infty.$$

Then it holds for every  $\kappa > 0$  and every  $(w, v) \in \mathbb{R}^N \times \mathbb{R}^L$  that

$$\sup_{Q \in \tilde{\mathcal{B}}_\kappa(P)} \mathcal{R}(\langle X^{Q_1}, w \rangle + \langle X^{Q_2}, v \rangle) = \mathcal{R}(\langle X^{P_1}, w \rangle + \langle X^{P_2}, v \rangle) + C\kappa\|w\|_q.$$

The assumptions in Proposition 1 for the case  $p = 1$  are not as restrictive as they might seem to be, particularly for our purposes. In Sect. 4, we will be working with the CVaR risk measure (Romisch and Pflug 2007), and it has the dual representation

$$CVaR_\alpha(X) = \sup \left\{ \mathbb{E}(XZ) : \mathbb{E}(Z) = 1, 0 \leq Z \leq \frac{1}{1 - \alpha} \right\}.$$

For  $A \in \Sigma$  with  $\mu(A) = 1 - \alpha$  and  $\{X \geq VaR_\alpha(X)\} \subset A$ ,  $Z \in \partial CVaR_\alpha(X)$  holds true for

$$Z(\omega) = \begin{cases} \frac{1}{1 - \alpha}, & \omega \in A \\ 0, & \text{otherwise} \end{cases}$$

and  $X \in L^1(\Omega, \Sigma, \mu)$ . With such  $Z \in \partial CVaR_\alpha(\langle X^{P_1}, w \rangle + \langle X^{P_2}, v \rangle)$ , if

$$\mu\left(\left\{\omega \in \Omega : \langle X^{P_1}(\omega), w \rangle < -k\right\}\right) > 0$$

for all  $k \in \mathbb{N}$  and  $w \neq 0$ , then, with increasing  $k$ , the ratio of

$$\mu\left(\left\{\omega \in \Omega : \langle X^{P_1}(\omega), w \rangle < -k\right\} \cap \left\{\omega \in \Omega : Z(\omega) = \frac{1}{1 - \alpha}\right\}\right)$$

to  $\mu(\{\omega \in \Omega : \langle X^{P_1}(\omega), w \rangle < -k\})$  converges to 1. Thus, any elliptical distribution commonly used for modeling asset returns (Owen and Rabinovitch 1983), such as multivariate Normal and t-distributions, and the Laplace distribution as well, satisfy the assumptions in Proposition 1, case  $p = 1$ . The assumptions for  $1 < p < \infty$  fixing  $\|\mathbb{E}[Z|\mathcal{F}^\perp]\|_{L^q}$  to a constant, on the other hand, might be very restrictive in certain cases. For the Markowitz functional  $M_\gamma(X) = \mathbb{E}[X] + \gamma\sqrt{Var(X)}$ , for instance, definition in dual form is (see Pflug et al. 2012 for derivations)

$$M_\gamma(X) = \sup \left\{ \mathbb{E}[XZ] : \mathbb{E}[Z] = 1, \|Z\|_{L^2} = \sqrt{1 + \gamma^2} \right\},$$



thus  $\|Z\|_{L^2}$  is constant for all  $Z \in \partial\mathcal{R}(X)$  and all  $X \in L^2$ . It holds that  $1 \leq \|\mathbb{E}[Z|\mathcal{F}^\perp]\|_{L^2} \leq \sqrt{1 + \gamma^2}$ , but unless further restrictions are posed on  $P_1$  and  $P_2$ ,  $\|\mathbb{E}[Z|\mathcal{F}^\perp]\|_{L^2} = C$  for all  $Z \in \partial\mathcal{R}(X)$  and some  $C \in \mathbb{R}$  does not hold true in general. For  $\|\mathbb{E}[Z|\mathcal{F}^\perp]\|_{L^2} = \|Z\|_{L^2} = C$  to hold true for all  $Z$ ,  $\sigma(X^{P_2}) = \{\emptyset, \Omega\}$ , thus a constant  $X^{P_2}$  is necessary, and  $P_2$  represents asset(s) with fixed return(s). Using a single riskless asset, it is possible to observe the willingness of the investor to enter the market constituted by ambiguous assets as the uncertainty radius increases and uniform investment in the ambiguous assets becomes optimal, as discussed in Sect. 4.

By Proposition 1, we transform the objective function of the inner problem (7)–(8) [also the main problem (3)–(4)], and restate without taking supremum:

$$\inf_{w \in \mathbb{R}^N} \mathcal{R}(\langle X^{P_1}, w \rangle + \langle X^{P_2}, v \rangle) + C\kappa \|w\|_q \tag{10}$$

$$s.t. \quad \langle \mathbb{1}, w \rangle = 1 - \langle \mathbb{1}, v \rangle. \tag{11}$$

With the result of Proposition 1, it is possible to state the worst-case risk measure as a functional whose input is based only on  $w, v$ , and the nominal distribution  $P$ . In this sense we are able to fix the distribution. Portfolio choice for the unambiguous assets,  $v$ , is also fixed in the inner problem, and we can now derive a bound for the change in the risk measure  $\mathcal{R}$  due to change in the ambiguous portfolio vector  $w$ . This bound is presented in Lemma 1, leading to the result in Lemma 2, and is critical to the results on convergence to the uniform portfolio for the ambiguous assets, hence the derivation of the optimality rules of the inner problem. From this point on, we do not restate the assumptions on  $\mathcal{R}$  and  $P$  for the results to be presented. Each result is presented for  $\mathcal{R}$  and  $P$  satisfying the conditions required by the appropriate case of Proposition 1 depending on the space  $\mathcal{R}$  is defined on ( $p = 1$  or  $1 < p < \infty$ ). We keep with the notation  $\mathcal{F}^\perp$  for the (largest)  $\sigma$ -algebra such that  $\mathcal{F}^\perp$  and  $\sigma(X^{P_2})$  are independent.

**Lemma 1** For all  $w^1, w^2 \in \mathbb{R}^N$ , and  $Z \in \partial\mathcal{R}(\langle X^{P_1}, w^1 \rangle + \langle X^{P_2}, v \rangle)$ ,

$$\begin{aligned} &\mathcal{R}(\langle X^{P_1}, w^1 \rangle + \langle X^{P_2}, v \rangle) - \mathcal{R}(\langle X^{P_1}, w^2 \rangle + \langle X^{P_2}, v \rangle) \\ &\leq C \|w^1 - w^2\|_q \mathbb{E} \left[ \|X^{P_1}\|_p^p \mathbb{1}_{\{\bar{Z} \neq 0\}} \right]^{\frac{1}{p}}. \end{aligned}$$

Here,  $\bar{Z} = \mathbb{E}[Z|\mathcal{F}^\perp]$ .

In Lemma 1, a bound is presented for the difference in risk measure caused by different ambiguous portfolio selections  $w^1$  and  $w^2$  with the same allocation for the unambiguous assets. Analyses that follow make clear the desirability of that bound being smaller, for reducing the threshold for the uncertainty parameter  $\kappa$  for the optimality of uniform ambiguous investment (for the case  $p = 1$ ), or reducing the upper bound reported for the distance of  $w^{\mu, v}$  to the optimal solution of (10)–(11) (for the case  $p = 2$ ). Notice how the characteristic function of  $\{\omega \in \Omega : \bar{Z}(w) \neq 0\}$  is used in (53) in the proof of Lemma 1 to reduce the final value derived. To continue with  $Z$  instead of  $\bar{Z}$  is an option in the derivations. If  $\|Z\|_{L^q}$  is also known to be constant on  $\partial\mathcal{R}(X)$  for all  $X \in L^p$ , say  $\bar{C}$ ,  $\mathbb{E}[\|X^{P_1}\|_p^p \mathbb{1}_{\{Z \neq 0\}}] \|w^1 - w^2\|_q \bar{C}$  is also a valid upper bound on the difference of risk and can be used instead of the bound derived in Lemma 1. For  $p = 1$ ,  $\|\mathbb{E}[Z|\mathcal{F}^\perp]\|_{L^\infty} = \|Z\|_{L^\infty} = C$  as noted in the remarks following Lemma 5 in ‘‘Appendix’’.  $L^p$ -norm is otherwise reduced by the conditional expectation operator. For instance, for the Markowitz functional  $M_\gamma$ , we can write  $1 \leq \|\mathbb{E}[Z|\mathcal{F}^\perp]\|_{L^2} \leq \|Z\|_{L^2} = \sqrt{1 + \gamma^2}$ . Thus, using  $\bar{Z} = \mathbb{E}[Z|\mathcal{F}^\perp]$  may lead to a

tighter bound in Lemma 1, but this is not certain, since the effect of the indicator functions  $\mathbb{1}_{\{\bar{Z} \neq 0\}}$  and  $\mathbb{1}_{\{Z \neq 0\}}$  can not be compared in general. Either form can be preferred considering the quality of the bound obtained or computational ease.

The objective function of (10)–(11) is composed of two terms, the term involving the risk functional and the term  $C\kappa\|w\|_q$ . In Lemma 1, an upper bound for the difference in the former term for two vectors  $w^1$  and  $w^2$  is given, which is based on  $\|w^1 - w^2\|_q$ . For the latter term, the difference is  $C\kappa\left(\|w^1\|_q - \|w^2\|_q\right)$ . In the following lemma,  $w^1$  will be fixed as  $w^{u,v}$ , and the ratio  $\frac{\|w - w^{u,v}\|_q}{\|w\|_q - \|w^{u,v}\|_q}$  will determine whether a solution  $w$  outperforms  $w^{u,v}$  in worst-case risk. This result critical to the optimality conditions of  $w^{u,v}$  for the inner problem (10)–(11) follows.

**Lemma 2** For fixed  $v \in \mathbb{R}^L$ , the uniform portfolio  $w^{u,v}$  is no worse than solutions in a set  $B \subset \mathbb{R}^N$ , i.e.:

$$\mathcal{R}\left(\langle X^{P_1}, w^{u,v} \rangle + \langle X^{P_2}, v \rangle\right) + C\kappa\|w^{u,v}\|_q \leq \mathcal{R}\left(\langle X^{P_1}, w \rangle + \langle X^{P_2}, v \rangle\right) + C\kappa\|w\|_q,$$

for all  $w \in B$ , if

$$\kappa \geq \frac{\|w - w^{u,v}\|_q}{\|w\|_q - \|w^{u,v}\|_q} \mathbb{E}\left(\|X^{P_1}\|_p^p \mathbb{1}_{\{\bar{Z} \neq 0\}}\right)^{\frac{1}{p}},$$

for all  $w \in B$ . Here,  $Z \in \partial\mathcal{R}\left(\langle X^{P_1}, w^{u,v} \rangle + \langle X^{P_2}, v \rangle\right)$ .

The result in Lemma 2 can be interpreted in two perspectives. Firstly, it defines the part of the feasible region that is outperformed by the uniform portfolio  $w^{u,v}$ , where the criteria is that the ratio  $\frac{\|w - w^{u,v}\|_q}{\|w\|_q - \|w^{u,v}\|_q}$  is less than a constant, in the sense that  $\kappa$  is a constant for the problem (10)–(11), and  $\mathbb{E}\left(\|X^{P_1}\|_p^p \mathbb{1}_{\{\bar{Z} \neq 0\}}\right)^{\frac{1}{p}}$  is again a constant computable given the knowledge of the nominal distribution  $P$ , the fixed value of  $v$ , and the intrinsic properties of the risk measure  $\mathcal{R}$  that define  $Z$  at  $\langle X^{P_1}, w^{u,v} \rangle + \langle X^{P_2}, v \rangle$ . Secondly, this result indicates that  $\kappa$ , the radius of uncertainty, is the threshold of acceptance determining the region outperformed by the uniform portfolio. This presents in three different shapes, as discussed in Proposition 2. For Kantorovich distance of degree  $p = 1$ , if  $\kappa$  is larger than a threshold value, the uniform portfolio is the exact optimal solution of the inner problem (10)–(11) regardless of the amount  $1 - \langle \mathbb{1}, v \rangle$  allocated to the first  $N$  assets. For  $p = 2$ , the rule of optimality rather directly reflects the result in Lemma 2:  $w^{u,v}$  is shown to be in the proximity of the optimal solution of (10)–(11) and outperform all the feasible region except a ball around itself. As the uncertainty radius  $\kappa$  increases, this ball becomes smaller and the optimal solution converges to  $w^{u,v}$ . The size of the ball is dependent on the allocation to the ambiguous assets, i.e.,  $1 - \langle \mathbb{1}, v \rangle$ . Structurally the result for  $p \neq 1, 2$  is similar to the case  $p = 2$ , that is, as  $\kappa$  increases, the region containing the optimal solution to (10)–(11) gets confined to a smaller neighborhood of  $w^{u,v}$ . Convergence of the optimal to the uniform portfolio as  $\kappa \rightarrow \infty$  is proven without reporting a relationship between the magnitude of  $\kappa$  and the radius of convergence. Proposition 2 follows, presenting the main result on the optimality of uniform portfolio for the inner problem (10)–(11).

**Proposition 2** Given fixed  $v \in \mathbb{R}^L$ , and considering  $N \geq 2$ ,

1. For  $p = 1$ ,  $w^{u,v}$  is the optimal solution to (10)–(11) if  $\kappa \geq \kappa^*$ , where:

$$\kappa^* = \begin{cases} (N - 1) \mathbb{E}\left[\|X^{P_1}\|_1 \mathbb{1}_{\{\bar{Z} \neq 0\}}\right] & \text{if } \langle \mathbb{1}, v \rangle \neq 1 \\ \mathbb{E}\left[\|X^{P_1}\|_1 \mathbb{1}_{\{\bar{Z} \neq 0\}}\right] & \text{if } \langle \mathbb{1}, v \rangle = 1. \end{cases} \tag{12}$$

2. For  $p = 2$ , the optimal portfolio for (10)–(11) lies in  $\{w \in \mathbb{R}^N : \|w - w^{u,v}\|_q < D\}$  if:

$$\kappa \geq \left( \left( \frac{(1 - \langle \mathbb{1}, v \rangle)^2}{ND^2} + 1 \right)^{\frac{1}{2}} + \frac{|1 - \langle \mathbb{1}, v \rangle|}{\sqrt{ND}} \right) \left[ \mathbb{E} \|X^{P_1}\|_2^2 \mathbb{1}_{\{\bar{Z} \neq 0\}} \right]^{\frac{1}{2}}. \tag{13}$$

3. For  $p \notin \{1, 2\}$ , for every  $\epsilon > 0$ , there is a  $\kappa_\epsilon$  such that for  $\kappa \geq \kappa_\epsilon$  the optimal solution  $w^*$  of (10)–(11) is inside  $\{w \in \mathbb{R}^N : \|w - w^{u,v}\|_q < \epsilon\}$ .

The conditions stated in Proposition 2 for the uncertainty parameter  $\kappa$  are localized, that is, specific to the value of  $v \in \mathbb{R}^L$ . For the uniform portfolio  $w^{u,v}$  to be set as the optimal or approximately optimal solution of the inner problem (10)–(11) at each  $v \in \mathbb{R}^L$ ,  $\kappa$  has to satisfy the condition required in Proposition 2 for all  $v \in \mathbb{R}^L$ . The terms in (12) and (13) that involve  $Z$  are critical in this regard, since  $Z \in \partial \mathcal{R}(\langle X^{P_1}, w^{u,v} \rangle + \langle X^{P_2}, v \rangle)$ , and being specific to the value of  $v$ , we now use the notation  $Z^v$ . For (12) to be satisfied at each  $v \in \mathbb{R}^L$ , we have to alleviate the dependency on  $v$  and modify the threshold as

$$\kappa^* = (N - 1) \sup_{v \in \mathbb{R}^L} \mathbb{E} \left[ \|X^{P_1}\|_1 \mathbb{1}_{\{\bar{Z}^v \neq 0\}} \right]. \tag{14}$$

Although it is presented with an exception for  $\langle \mathbb{1}, v \rangle = 1$  in Proposition 2, this is the actual threshold for  $\kappa$ , since we need the result of the inner problem to hold for all  $v \in \mathbb{R}^L$  to be able to proceed with the optimization for  $v$ . Once  $\kappa$  exceeds this exact threshold, we know that  $w^{u,v}$  is the optimal solution of the inner problem for all  $v \in \mathbb{R}^L$ . The situation is different with  $p = 2$ . The condition in (13) is flexible due to the term  $D$ .  $D$  can be picked as the smallest positive value satisfying (13), and the condition can be satisfied for all  $\kappa > 0$  and  $v \in \mathbb{R}^L$ , with different values of  $D$ . To keep  $D$  specific to  $v/Z^v$  as  $D^v$  is possible, but for practicality of computation it is reasonable to fix a value for  $D$  for all  $v \in \mathbb{R}^L$ . Since  $D$  determines the proximity of the optimal solution  $w^{*,v}$  to the uniform portfolio  $w^{u,v}$ , given  $\kappa$ , it is sensible to take  $D$  as the smallest positive number such that (13) is satisfied. Given  $D > 0$  that satisfies

$$\kappa = \left( \left( \frac{1}{ND^2} + 1 \right)^{\frac{1}{2}} + \frac{1}{\sqrt{ND}} \right) \sup_{v \in \mathbb{R}^L} \mathbb{E} \left[ \|X^{P_1}\|_2^2 \mathbb{1}_{\{\bar{Z}^v \neq 0\}} \right]^{\frac{1}{2}}, \tag{15}$$

for  $v \in \mathbb{R}^L$ ,  $\langle \mathbb{1}, v \rangle \neq 1$ ,  $w^{*,v} \in \{w \in \mathbb{R}^N : \|w - w^{u,v}\|_2 < |1 - \langle \mathbb{1}, v \rangle| D\}$  since (13) is satisfied at  $v$  with  $D$  replaced by  $|1 - \langle \mathbb{1}, v \rangle| D$ . Note that (15) fails if

$$\kappa < \sup_{v \in \mathbb{R}^L} \mathbb{E} \left[ \|X^{P_1}\|_2^2 \mathbb{1}_{\{\bar{Z}^v \neq 0\}} \right]^{\frac{1}{2}}.$$

As the magnitude of allocation  $|1 - \langle \mathbb{1}, v \rangle|$  increases, the ability of the uniform ambiguous investment to approximate the optimal reduces. Let’s consider the difference between the objective values attained in (10)–(11) by  $w^{*,v}$  and  $w^{u,v}$ :

$$\begin{aligned}
 0 &\leq \mathcal{R} \left( \langle X^{P_1}, w^{u,v} \rangle + \langle X^{P_2}, v \rangle \right) + C\kappa \|w^{u,v}\|_2 \\
 &\quad - \mathcal{R} \left( \langle X^{P_1}, w^{*,v} \rangle + \langle X^{P_2}, v \rangle \right) - C\kappa \|w^{*,v}\|_2 \\
 &\leq C \|w^{u,v} - w^{*,v}\|_2 \sup_{v \in \mathbb{R}^L} \mathbb{E} \left[ \|X^{P_1}\|_2^2 \mathbb{1}_{\{\bar{z}^v \neq 0\}} \right]^{\frac{1}{2}} + C\kappa (\|w^{u,v}\|_2 - \|w^{*,v}\|_2) \\
 &\leq C \|w^{u,v} - w^{*,v}\|_2 \sup_{v \in \mathbb{R}^L} \mathbb{E} \left[ \|X^{P_1}\|_2^2 \mathbb{1}_{\{\bar{z}^v \neq 0\}} \right]^{\frac{1}{2}} + C\kappa \|w^{u,v} - w^{*,v}\|_2 \\
 &< |1 - \langle \mathbb{1}, v \rangle| CD \sup_{v \in \mathbb{R}^L} \mathbb{E} \left[ \|X^{P_1}\|_2^2 \mathbb{1}_{\{\bar{z}^v \neq 0\}} \right]^{\frac{1}{2}} + |1 - \langle \mathbb{1}, v \rangle| CD\kappa,
 \end{aligned}$$

where the first three inequalities follow due to optimality of  $w^{*,v}$ , Lemma 1, and the triangle inequality, respectively, and the last inequality is due to the fact that  $\|w^{*,v} - w^{u,v}\|_2 < |1 - \langle \mathbb{1}, v \rangle| D$ . With this derivation, we can see that a band between two functions of  $v \in \mathbb{R}^L$  contains the optimal values for the inner problems (10)–(11). Once the optimal solution  $w^{*,v}$  of the inner problem is at hand for all  $v \in \mathbb{R}^N$ , the outer problem can be defined as

$$\inf_{v \in \mathbb{R}^L} \mathcal{R} \left( \langle X^{P_1}, w^{*,v} \rangle + \langle X^{P_2}, v \rangle \right) + C\kappa \|w^{*,v}\|_q, \tag{16}$$

which does not necessarily coincide with the problem that takes the uniform portfolio as the solution to the inner problem, unless  $p = 1$ :

$$\inf_{v \in \mathbb{R}^L} \mathcal{R} \left( \frac{1 - \langle \mathbb{1}, v \rangle}{N} \langle X^{P_1}, \mathbb{1} \rangle + \langle X^{P_2}, v \rangle \right) + C\kappa \frac{|1 - \langle \mathbb{1}, v \rangle|}{N}. \tag{17}$$

The objective function  $f^*(v)$  to be minimized in (16) is inside a band between the objective function  $f(v)$  of (17) and  $g(v)$  defined as:

$$g(v) = f(v) - |1 - \langle \mathbb{1}, v \rangle| CD \left( \sup_{v \in \mathbb{R}^L} \mathbb{E} \left[ \|X^{P_1}\|_2^2 \mathbb{1}_{\{\bar{z}^v \neq 0\}} \right]^{\frac{1}{2}} + \kappa \right).$$

$f(v) \geq f^*(v) > g(v)$ , for  $\langle \mathbb{1}, v \rangle \neq 1$ . The uniform portfolio  $w^{u,v}$  is at hand, and  $f(v)$  is an attainable objective value for all  $v \in \mathbb{R}$ . However, there are no guarantees that  $f^*$  gets any near  $g$ , indeed, it is possible that  $w^{*,v} = w^{u,v}$ , thus  $f^*(v) = f(v)$  for some  $v \in \mathbb{R}$ . The difference between  $f$  and  $g$  increases linearly with  $|1 - \langle \mathbb{1}, v \rangle|$ , along with the reduced quality of approximation of  $f^*$  by  $f$ .  $f - g$  is a bound—without guarantees—for the risk reduction opportunity, lost due to preferring  $w^{u,v}$  instead of searching for  $w^{*,v}$ . Nevertheless, with  $w^{u,v}$  and  $f$  at hand, it is possible to evaluate the performance of  $v$  based on an actual figure that estimates  $f^*$  within a certain error. While  $f$  and  $f^*$  are defined for  $p \notin \{1, 2\}$ , it is not possible to provide a lower bound such as  $g$  for the quality of approximation, since the relationship between the approximation radius  $\epsilon$  and the uncertainty parameter  $\kappa_\epsilon$  is not quantified. Nevertheless, the following result proves that the investor divests of the assets involving ambiguity as the uncertainty parameter  $\kappa$  tends to infinity, for all  $p \in [1, \infty)$ .

**Proposition 3** *As  $\kappa \rightarrow \infty$ , all solutions with  $\langle \mathbb{1}, v \rangle \neq 1$  turn suboptimal.*

For  $p \notin \{1, 2\}$ , Proposition 2 establishes the convergence to the uniform portfolio  $w^{u,v}$  but does not provide any quantitative relationship between the uncertainty parameter  $\kappa$  and the proximity of the uniform portfolio to the actual optimal of the inner problem (10)–(11). This reflects in Proposition 3, thus it is not possible to infer about the rate at which the investor is driven out of the ambiguous market. The figures in the proof of Proposition 3 for the cases

$p = 1$  and  $p = 2$  can be used to report the rate at which the total allocation to the ambiguous market diminishes. For instance, if  $v^*$  is the optimal solution of  $\inf_{\langle \mathbb{1}, v \rangle = 1} \mathcal{R}(\langle X^{P_2}, v \rangle)$ , and  $v^{*,s}$  is the optimal solution of  $\inf_{\langle \mathbb{1}, v \rangle = 1-s} \mathcal{R}(\langle X^{P_1}, w^{u,v} \rangle + \langle X^{P_2}, v \rangle)$ , then an allocation to the ambiguous market at a level  $s$  is ruled out by suboptimality when  $\kappa$  exceeds  $\frac{N}{C|s|} [\mathcal{R}(\langle X^{P_2}, v^* \rangle) - \mathcal{R}(\frac{s}{N} \langle X^{P_1}, \mathbb{1} \rangle + \langle X^{P_2}, v^{*,s} \rangle)]$ . Similarly, it is possible to derive approximate divestment rates for the case  $p = 2$ , as (15) can be solved for  $D$  given  $\kappa$ :

$$D = \frac{2\kappa\Pi}{\sqrt{N}(\kappa^2 - \Pi^2)},$$

for  $\kappa > \sup_{v \in \mathbb{R}^L} \left[ \mathbb{E} \|X^{P_1}\|_2^2 \mathbb{1}_{\{\bar{Z}^v \neq 0\}} \right]^{\frac{1}{2}} = \Pi$ . Then  $f^*(v)$  can be bounded below by

$$\mathcal{R} \left( \frac{s}{N} \langle X^{P_1}, \mathbb{1} \rangle + \langle X^{P_2}, v^{*,s} \rangle \right) + C\kappa \frac{|s|}{\sqrt{N}} \left( 1 - \frac{2\Pi^2}{\kappa^2 - \Pi^2} - \frac{2\Pi\kappa}{\kappa^2 - \Pi^2} \right),$$

and comparing this figure to  $\mathcal{R}(\langle X^{P_2}, v^* \rangle)$  provides an approximate rate for the divestment of ambiguous assets. However, these figures involve an abstract measure and optimal solutions of problems based on this measure, which do not provide practical insight. In the following section, we work with CVaR and Markowitz functionals as measures, and construct numerical examples to depict optimal ambiguous and unambiguous portfolio allocation combinations in practice.

### 4 Applications with the CVaR and Markowitz functionals

After the inner optimization problem (7)–(8) is solved, and (approximate) optimality of the uniform portfolio  $w^{u,v}$  corresponding to the fixed value of  $v$  is proven, it remains to solve for the optimal value of  $v$  that minimizes the risk measure  $\mathcal{R}$ , when combined with its corresponding uniform portfolio  $w^{u,v}$  for the ambiguous assets. Up to this point, the abstract definition of  $\mathcal{R}$ , with a dual formulation and the requirement of convexity and law invariance, was sufficient for deriving optimal portfolio rules. However, for characterizing the optimality rules for the investment into the assets that do not involve ambiguity, we will need to study specific risk measures. The Conditional Value-at-Risk (CVaR) (Rockafellar and Uryasev 2002) measure is one possible choice, satisfying all assumptions on  $\mathcal{R}$ , and popular in risk minimizing optimization applications due to its desirable characteristics such as translation equivariance, positive homogeneity and Lipschitz continuity. Despite its simple form and computational practicality, it is the core representative of the class of convex law invariant risk functionals as many risk functionals in this class can be defined as functions of CVaR (Romisch and Pflug 2007). Integrability of the loss function is sufficient for the computation of CVaR, thus we will work on  $L^1$ , and the assumptions in Proposition 1, case  $p = 1$ , and Lemma 5 apply in this case. We assume normal return distributions  $P_1$ , and  $P_2$ , while only  $P_1$  being a member of the family of elliptical distributions would be sufficient to satisfy the assumptions.

$CVaR_\alpha(X)$  for  $X \in L^1$  is defined for loss functions, hence we negate the return variable.  $CVaR_\alpha(X)$  is the expected loss given that the loss exceeds the Value-at-Risk,  $VaR_\alpha(X)$ , which is the (minimum) amount of loss above which a loss can occur with probability no more than  $1 - \alpha$ .  $\alpha$  is generally chosen around 0.95.  $CVaR_\alpha$  can be defined as [see Rockafellar and Uryasev (2002) for details and the definition of Value-at-Risk (VaR)]

$$CVaR_\alpha(X) = \frac{1}{1-\alpha} \int_{\{X > VaR_\alpha(X)\}} X d\mu.$$

If  $\kappa$  exceeds (14), then the solution for the inner problem (10)–(11) is known for all  $v \in \mathbb{R}^L$ , and there remains to solve the outer problem:

$$\inf_{v \in \mathbb{R}^L} CVaR_\alpha \left( -\langle X^{P_1}, w^{u,v} \rangle - \langle X^{P_2}, v \rangle \right) + C\kappa \|w^{u,v}\|_\infty.$$

Plugging in  $w^{u,v} = \frac{1-\langle \mathbb{1}, v \rangle}{N} \mathbb{1}$  with the addition of a variable  $s$  for simplification, we have the equivalent formulation :

$$\begin{aligned} \inf_{\substack{s \in \mathbb{R} \\ v \in \mathbb{R}^L}} CVaR_\alpha \left( -\frac{s}{N} \langle \mathbb{1}, X^{P_1} \rangle - \langle X^{P_2}, v \rangle \right) + C\kappa \frac{|s|}{N} \\ s.t. \quad s + \langle \mathbb{1}, v \rangle = 1. \end{aligned}$$

For a normal loss function  $X \sim \mathcal{N}(\mu, \sigma)$ , one can easily compute  $CVaR_\alpha(X)$  to be  $\sigma \frac{\phi(\Phi^{-1}(\alpha))}{1-\alpha} + \mu$ . Here,  $\phi$  is the standard normal density function, and  $\Phi$  is the standard normal cumulative distribution function. Thus we redefine the objective function:

$$f(v, s) = \sigma \frac{\phi(\Phi^{-1}(\alpha))}{1-\alpha} + \mu + C\kappa \frac{|s|}{N},$$

since the loss function is the sum of two linearly transformed normal random variables, and is itself normal. Let  $X^{P_1}$  and  $X^{P_2}$  have mean vector and invertible covariance matrices  $\mu_1, \Gamma_1$  and  $\mu_2, \Gamma_2$ , respectively. Then  $\mu = -\frac{s}{N} \langle \mu_1, \mathbb{1} \rangle - \langle \mu_2, v \rangle$  and  $\sigma = \sqrt{\frac{s^2}{N^2} \mathbb{1}^T \Gamma_1 \mathbb{1} + v^T \Gamma_2 v}$ , and we have the following mathematical programming formulation:

$$\begin{aligned} \inf_{\substack{s \in \mathbb{R} \\ v \in \mathbb{R}^L}} a\sqrt{bs^2 + v^T \Gamma_2 v} - cs - \langle \mu_2, v \rangle + d|s| \\ s.t. \quad s + \langle \mathbb{1}, v \rangle = 1, \end{aligned}$$

or with the linearization of the absolute deviation term:

$$\inf_{\substack{s \in \mathbb{R} \\ \bar{s} \in \mathbb{R} \\ v \in \mathbb{R}^L}} a\sqrt{bs^2 + v^T \Gamma_2 v} - cs - \langle \mu_2, v \rangle + d\bar{s} \tag{18}$$

$$s.t. \quad s + \langle \mathbb{1}, v \rangle = 1 \tag{19}$$

$$s - \bar{s} \leq 0 \tag{20}$$

$$-s - \bar{s} \leq 0, \tag{21}$$

where  $a = \frac{\phi(\Phi^{-1}(\alpha))}{1-\alpha}$ ,  $b = \frac{1}{N^2} \mathbb{1}^T \Gamma_1 \mathbb{1}$ ,  $c = \frac{1}{N} \langle \mu_1, \mathbb{1} \rangle$ ,  $d = \frac{1}{N} C\kappa$ , and  $\bar{s}$  is the auxiliary variable for modeling  $|s|$ . Since the constraints are linear, and the objective function is convex (by convexity of the  $CVaR$  measure), Slater’s condition is trivially satisfied, and by KKT conditions we derive the optimal solution of (18)–(21). We have the KKT conditions

$$\nabla f + \lambda_1 \nabla f_1 + \lambda_2 \nabla f_2 + \lambda_3 \nabla f_3 = 0, \tag{22}$$

with  $\lambda_1 \in \mathbb{R}$ ,  $\lambda_2, \lambda_3 \in \mathbb{R}_+$ , and

$$\lambda_i f_i(v, s, \bar{s}) = 0, \quad i = 2, 3. \tag{23}$$

Here,  $f_1(v, s, \bar{s})$  is the function corresponding to the equality constraint in (18)–(21),  $f_2(v, s, \bar{s})$  and  $f_3(v, s, \bar{s})$  correspond to the inequality constraints.  $f$ , with its modified definition as a function of  $v, s$  and  $\bar{s}$ , is the objective function. With  $\mathbf{0}, \mathbf{1} \in \mathbb{R}^L$ , the gradients are:

$$\nabla f = \frac{a}{\sigma} \begin{bmatrix} \Gamma_2 v \\ bs \\ 0 \end{bmatrix} + \begin{bmatrix} -\mu_2 \\ -c \\ d \end{bmatrix}, \quad \nabla f_1 = \begin{bmatrix} \mathbf{1} \\ 1 \\ 0 \end{bmatrix}, \quad \nabla f_2 = \begin{bmatrix} \mathbf{0} \\ 1 \\ -1 \end{bmatrix}, \quad \nabla f_3 = \begin{bmatrix} \mathbf{0} \\ -1 \\ -1 \end{bmatrix}.$$

The last row in (22),  $d = \lambda_2 + \lambda_3$ , along with the complementary slackness constraints (23) imply that  $d = \lambda_2$  if  $s > 0$  ( $\bar{s} = s$ ) and  $d = \lambda_3$  if  $s < 0$  ( $\bar{s} = -s$ ). From the first two rows of (22), we have  $v = \frac{\sigma}{a} \Gamma_2^{-1} (\mu_2 - \lambda_1 \mathbf{1})$  and  $s = \frac{\sigma}{ab} (c - \lambda_1 - \lambda_2 + \lambda_3)$ . Since  $-\lambda_2 + \lambda_3 = -\text{sign}(s)d$ , we define  $e = c - \text{sign}(s)d$ , and write  $s = \frac{\sigma}{ab} (e - \lambda_1)$ . Plugging  $v$  and  $s$  into  $\sigma(v, s)$ :

$$\frac{\sigma^2}{a^2 b} (e - \lambda_1)^2 + \frac{\sigma^2}{a^2} \left( \mu_2^T \Gamma_2^{-1} \mu_2 - 2\lambda_1 \mu_2^T \Gamma_2^{-1} \mathbf{1} + \lambda_1^2 \mathbf{1}^T \Gamma_2^{-1} \mathbf{1} \right) = \sigma^2,$$

we obtain the quadratic equation of  $\lambda_1$ :

$$\lambda_1^2 (1 + bD) + \lambda_1 (-2e - 2bB) + (e^2 + bA - a^2 b) = 0,$$

where  $A = \mu_2^T \Gamma_2^{-1} \mu_2, B = \mu_2^T \Gamma_2^{-1} \mathbf{1}, D = \mathbf{1}^T \Gamma_2^{-1} \mathbf{1}$ . Solutions for the quadratic equation are:

$$\lambda_1 = \frac{e + bB \pm \sqrt{\Delta}}{1 + bD},$$

where  $\Delta = b^2 (B^2 - DA + a^2 D) + b (2eB - A - e^2 D + a^2)$ . Having  $\lambda_1$ , we check (19):

$$\frac{\sigma}{ab} (e - \lambda_1) + \frac{\sigma}{a} (B - \lambda_1 D) = 1,$$

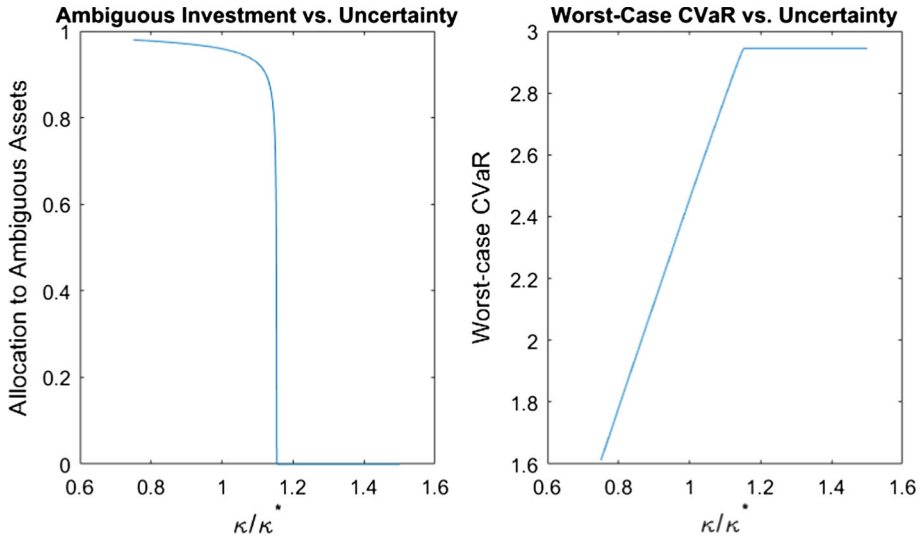
thus  $\sigma = \frac{ab}{\mp \sqrt{\Delta}}$ . Since  $\sigma = \sqrt{\sigma^2}$ , we pick  $\lambda_1 = \frac{e + bB - \sqrt{\Delta}}{1 + bD}$ . Finally, we plug the solution of  $\sigma$  into the solution of (18)–(21):

$$v^* = \frac{b}{\sqrt{\Delta}} \Gamma_2^{-1} \left( \mu_2 - \frac{e + bB - \sqrt{\Delta}}{1 + bD} \mathbf{1} \right)$$

$$s^* = \frac{1}{\sqrt{\Delta}} \left( \frac{ebD - bB + \sqrt{\Delta}}{1 + bD} \right).$$

There are two points to check in this solution, one is the positivity of the discriminant  $\Delta$ . The other is the compliance of the sign of the solution for  $s$  with the value of  $e$ . Either  $e = c - d$ , the respective value for  $\Delta$  is positive, and  $s^* > 0$  is a solution, or  $e = c + d$ , and the respective value for  $\Delta$  is positive, and  $s^* < 0$  is a solution. Thus, if  $\Delta > 0$  with  $e = c - d$ , and the nominator in the solution of  $s^*$ , i.e.,  $bcD - bdD - bB + \sqrt{\Delta}$ , is positive then (18)–(21) has a solution with  $s^* > 0$ , or if  $\Delta > 0$  with  $e = c + d, bcD + bdD - bB + \sqrt{\Delta} < 0$  then a solution exists with  $s^* < 0$ . Since the latter nominator term is larger (note the increase in both  $ebD$  and  $\Delta$  when  $e = c + d$ ), (18)–(21) has at most one solution, in accordance with the convexity of the problem. By convexity (thus continuity) of the objective function and convexity of the feasible region, if a solution such that  $s^* < 0$  or  $s^* > 0$  exists, then it





**Fig. 1** Allocation for the ambiguous market and worst-case CVaR for different levels of  $\kappa$ .  $\Gamma_2$  is held larger than  $\Gamma_1$  in this case. For the reader who wants to replicate the figures, we can supply data upon request

dominates the solution of the KKT equations with  $s = 0$ . One can similarly check with the KKT conditions that if  $s^* = 0$ , then

$$v^* = \frac{1}{\sqrt{B^2 - D(A - a^2)}} \Gamma_2^{-1} \left( \mu_2 - \frac{B - \sqrt{B^2 - D(A - a^2)}}{D} \mathbf{1} \right)$$

is the optimal solution given  $B^2 - D(A - a^2) > 0$ . If  $B^2 - D(A - a^2) \leq 0$ , there is no KKT solution with  $s = 0$ . If both discriminants are non-positive for  $s \neq 0$ , or there is no sign consistency between  $s$  and  $e$  on both sides, then there is no KKT solution with  $s \neq 0$ . If there are no KKT points, the problem does not have an optimal solution and might also be unbounded. As we assume (5)–(6), thus (17) is a well-posed problem, we consider examples/problems whose optimal solutions exist, without presenting further details on the setting of parameters for the existence of a solution.

In Fig. 1, left, the change in  $s^*$ , the allocation to the ambiguous assets in the optimal combined portfolio, in response to  $\kappa$  is depicted. The solution is valid after  $\kappa$  exceeds  $\kappa^*$ , and soon the investor pulls out of the ambiguous market. Until that point, the ambiguous market receives the major portion of the wealth, persistently. Here, naïve diversification combines with optimized portfolio choice to attain better results than any of the two strategies alone could achieve. On the right, we observe that the worst-case risk increases steadily until the investor immunizes portfolio against increasing uncertainty in the ambiguous market by moving out.

In a second example, we will consider the Markowitz functional  $M_\gamma(X) = \mathbb{E}[X] + \gamma\sqrt{\text{Var}(X)}$  for the case  $p = 2$ , whose natural domain is  $L^2(\Omega, \Sigma, \mu)$ . As noted in the remarks following Proposition 1, the conditions for Proposition 1,  $p > 1$ , are restrictive on the choice of return distributions  $P_1$  and  $P_2$ . In this setting, these conditions are satisfied by a choice of atomic  $P_2$  with a single atom  $\{\eta\}$ ,  $\eta \in \mathbb{R}^L$ , thus, when  $X^{P_2}$  is constant. In this case, it is sensible to take  $L = 1$ , since a single asset dominates when returns are constant. Hence,

we work on  $\mathbb{R}^N \times \mathbb{R}$ ,  $\begin{bmatrix} X^{P_1} \\ X^{P_2} \end{bmatrix} : \Omega \rightarrow \mathbb{R}^N \times \mathbb{R}$ ,  $X^{P_2}(\omega) = \eta$  for all  $\omega \in \Omega$ , and  $\mathbb{E}[X^{P_1}] = \mu \in \mathbb{R}^N$ ,  $\mathbb{E}\left[\left(X^{P_1} - \mu\right)^T \left(X^{P_1} - \mu\right)\right] = \Gamma \in \mathbb{R}^{N \times N}$  is sufficient information on  $P_1$  for an application with the Markowitz functional. Since the dual formulation of the Markowitz functional is  $M_\gamma(X) = \sup\left\{\mathbb{E}[XZ] : \mathbb{E}[Z] = 1, \|Z\|_{L^2} = \sqrt{1 + \gamma^2}\right\}$ , the constant  $C$  in Proposition 1 and the derivations that follow has the value  $\sqrt{1 + \gamma^2}$  (by the choice of  $P_2$ ,  $\mathcal{F}^\perp = \Sigma$  and  $\mathbb{E}[Z|\mathcal{F}^\perp] = Z$  for all  $X \in L^2$ ). Let us assume that the uncertainty radius  $\kappa$  exceeds  $\Pi = \sup_{v \in \mathbb{R}^L} \left[\mathbb{E}\|X^{P_1}\|_2^2 \mathbb{1}_{\{\bar{Z}^v \neq 0\}}\right]^{\frac{1}{2}}$  and we set  $D = \frac{2\kappa\Pi}{\sqrt{N}(\kappa^2 - \Pi^2)}$  to satisfy (15).

As discussed in Sect. 3, in this case  $w^{u,v}$  and  $f(v)$  stand as approximations. We work with  $f$  for an approximate solution for  $v$ , while keeping an eye on the approximation quality with  $g$ . Plugging in the Markowitz functional for  $\mathcal{R}$ , again negating the return random variable for loss,  $f$  takes the form:

$$\begin{aligned} f(v) &= \mathbb{E}\left[-\langle X^{P_1}, w^{u,v} \rangle - \eta v\right] + \gamma\sqrt{\text{Var}\left(-\langle X^{P_1}, w^{u,v} \rangle - \eta v\right)} + \sqrt{1 + \gamma^2}\kappa\|w^{u,v}\|_2 \\ &= -\langle \mu, w^{u,v} \rangle - \eta v + \gamma\sqrt{(w^{u,v})^T \Gamma w^{u,v}} + \sqrt{1 + \gamma^2}\kappa \frac{|1 - v|}{\sqrt{N}} \\ &= -\frac{1 - v}{N} \langle \mu, \mathbb{1} \rangle - \eta v + \frac{|1 - v|}{N} \gamma\sqrt{\mathbb{1}^T \Gamma \mathbb{1}} + \sqrt{1 + \gamma^2}\kappa \frac{|1 - v|}{\sqrt{N}}, \end{aligned}$$

and  $g$ , the lower bound for  $f^*$ , becomes:

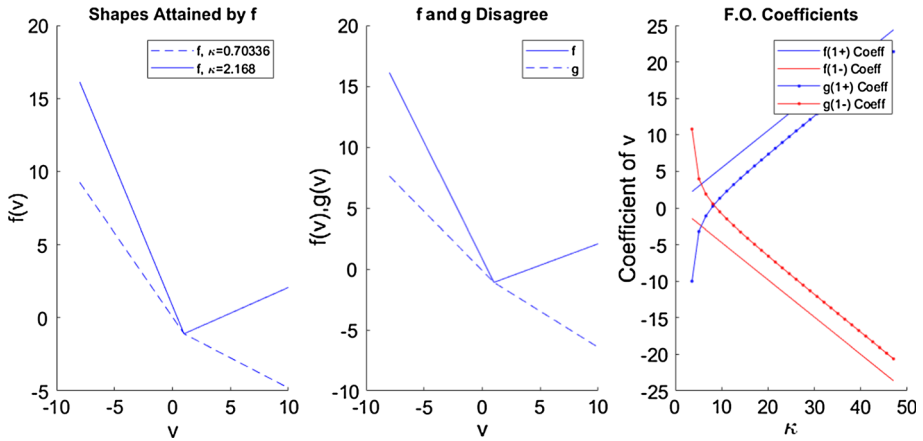
$$\begin{aligned} g(v) &= f(v) - \sqrt{1 + \gamma^2}D(\Pi + \kappa)|1 - v| \\ &= f(v) - 2\left(\frac{\Pi^2 + \kappa\Pi}{\kappa^2 - \Pi^2}\right) \frac{\sqrt{1 + \gamma^2}\kappa}{\sqrt{N}} |1 - v|. \end{aligned}$$

Denoting  $\langle \mu, \mathbb{1} \rangle$  by  $\bar{\mu}$  and  $\sqrt{\mathbb{1}^T \Gamma \mathbb{1}}$  by  $\bar{\sigma}$ , and rearranging terms, we have the following form:

$$f(v) = -\frac{\bar{\mu}}{N}(1 - v) - \eta v + \frac{\gamma\bar{\sigma}}{N}|1 - v| + \frac{\sqrt{1 + \gamma^2}\kappa}{\sqrt{N}}|1 - v|.$$

The functions  $f$  and  $g$  have piecewise linear form, and since no limits are posed on short selling, lead to unbounded investments if the problem is not well-posed:

$$f(v) = \begin{cases} v\left(\frac{\bar{\mu}}{N} - \eta - \frac{\gamma\bar{\sigma}}{N} - \frac{\sqrt{1 + \gamma^2}\kappa}{\sqrt{N}}\right) - \frac{\bar{\mu}}{N} + \frac{\gamma\bar{\sigma}}{N} + \frac{\sqrt{1 + \gamma^2}\kappa}{\sqrt{N}} & \text{if } v < 1 \\ -\eta & \text{if } v = 1 \\ v\left(\frac{\bar{\mu}}{N} - \eta + \frac{\gamma\bar{\sigma}}{N} + \frac{\sqrt{1 + \gamma^2}\kappa}{\sqrt{N}}\right) - \frac{\bar{\mu}}{N} - \frac{\gamma\bar{\sigma}}{N} - \frac{\sqrt{1 + \gamma^2}\kappa}{\sqrt{N}} & \text{if } v > 1, \end{cases}$$



**Fig. 2** Possible shapes that  $f$  attains (left),  $f$  and  $g$  disagreeing in whether to enter the ambiguous market (middle), and convergence of the first order coefficients of  $g$  to those of  $f$  with increasing  $\kappa$  (right)

$$g(v) = \begin{cases} v \left( \frac{\bar{\mu}}{N} - \eta - \frac{\gamma\bar{\sigma}}{N} - \frac{\sqrt{1+\gamma^2\kappa}}{\sqrt{N}} \left( 1 - 2 \left( \frac{\Pi^2 + \kappa\Pi}{\kappa^2 - \Pi^2} \right) \right) \right) - \frac{\bar{\mu}}{N} + \frac{\gamma\bar{\sigma}}{N} + \frac{\sqrt{1+\gamma^2\kappa}}{\sqrt{N}} \left( 1 - 2 \left( \frac{\Pi^2 + \kappa\Pi}{\kappa^2 - \Pi^2} \right) \right) & \text{if } v < 1 \\ -\eta & \text{if } v = 1 \\ v \left( \frac{\bar{\mu}}{N} - \eta + \frac{\gamma\bar{\sigma}}{N} + \frac{\sqrt{1+\gamma^2\kappa}}{\sqrt{N}} \left( 1 - 2 \left( \frac{\Pi^2 + \kappa\Pi}{\kappa^2 - \Pi^2} \right) \right) \right) - \frac{\bar{\mu}}{N} - \frac{\gamma\bar{\sigma}}{N} - \frac{\sqrt{1+\gamma^2\kappa}}{\sqrt{N}} \left( 1 - 2 \left( \frac{\Pi^2 + \kappa\Pi}{\kappa^2 - \Pi^2} \right) \right) & \text{if } v > 1. \end{cases}$$

Noting the larger coefficient of  $v$  to the right of 1,  $f$  extends with a higher slope towards the right at the point the pieces join.  $f$  can be monotone decreasing or v-shaped, but not monotone increasing (Fig. 2, left), excluding the cases with either first order coefficient equal to 0—there is divergence on the other side in this case. For  $X \in L^1, \{X \neq \mathbb{E}[X]\} = \{Z \neq 0\}$  for  $Z \in \mathcal{R}(X)$  (see Pflug et al. 2012), and  $\Pi$  can be simply taken as  $\mathbb{E} \left[ \|X^{P_1}\|_2^2 \right]^{\frac{1}{2}} = \left( (\mu, \mu) + \sum_{i=1}^N \Gamma_{ii} \right)^{\frac{1}{2}}$ . Since  $\|\mathbb{1}\|_2 = \sqrt{N}$ , with an application of Cauchy-Schwarz inequality it can be proven that  $\frac{\Pi}{\sqrt{N}} > \frac{\bar{\mu}}{N}$  so that the first order coefficient of  $f$  for  $v < 1$  is always negative unless  $\eta < 0$ . We observe that investment on the buying side is already prohibited in the ambiguous market when  $\kappa$  reaches the threshold for uniformity. For  $v > 1$ , the first order coefficient in  $f$  is positive unless mean returns are low in the ambiguous market, and the riskless rate  $\eta$  is high enough to compensate for the risk terms (the last two) in the coefficient and allow risk reduction by short selling of ambiguous assets (Fig. 2, left, dashed curve). This is quite unexpected of a riskless asset and a favorable market. High  $\bar{\sigma}$ ,  $\kappa$  and  $\gamma$  are prohibitive for both buying and short selling in the ambiguous market, since the first two terms represent risk, and the third represents risk aversion. For the Markowitz functional case, we clearly observe that investing in the ambiguous assets is not an optimal option for the worst-case risk minimizer at the uncertainty levels justifying uniform ambiguous portfolio (Fig. 2, left, solid curve).

The extra term in the function  $g$  diminishes as  $\kappa$  increases, and  $g$  improves as a lower bound for  $f^*$  while it gets closer to  $f$ . Figure 2, right, depicts the convergence of first order coefficients of  $g$  to those of  $f$ , and difference in coefficients indicate that uniform portfolio is a very weak approximation when the uncertainty level is low. Lowering both sides, the extra term in  $g$  might cause an upward pointing side of  $f$  to turn downwards in  $g$ . If  $f$  and  $g$  agree on the sign of the slope on both sides of  $v = 1$ , then  $g$  confirms the overall allocation decision to the ambiguous and unambiguous markets, while being a reminder and a bound for the risk reduction opportunity that might be lost by using  $w^{u,v}$  instead of  $w^{*,v}$ . The only interesting case of  $g$  flipping below the x-axis is when  $f$  is v-shaped (Fig. 2, middle). While  $f$  prescribes avoiding the ambiguous market,  $g$  indicates possible opportunities there with the possibility of investments performing higher than  $w^{u,v}$ . The investor might prefer further investigations and research in the ambiguous market to allow for a better decision than naïve diversification.  $g$  has no guarantees, therefore if the uniform solution is the best solution at hand,  $g$  will not be a decision changer for the worst-case risk minimizing investor.

## 5 Conclusion

A previous result in Pflug et al. (2012) on the optimality of uniform portfolio under increasing uncertainty is a solid evidence supporting the rationality of naïve diversification strategies. Adopting this framework, we extend the model to investigate validity of naïve diversification when there is an alternative group of risky assets with known distribution. We prove for a division of wealth to the two groups of assets that naïve diversification persists as the optimal strategy for the allocation into the ambiguous assets, despite the diminishing total allocation therein as the level of uncertainty increases. Taking *CVaR* as a representative of the class of convex and law invariant risk measures, we derive rules for efficient combinations of optimal diversification in the unambiguous assets with naïve diversification in the ambiguous assets. Modeling the existence of a riskless asset as a special case, we describe the attitude of the investor towards taking risk in an uncertain environment where naïve diversification is justified to be an optimal strategy. Indeed, while it renders uniform investment to risky assets optimal, increasing uncertainty radius drives the investor out of the ambiguous market.

Similar to the manner the strength of naïve diversification strategy under uncertainty is somewhat an expected result pointed out by previous studies (DeMiguel et al. 2009b), the result in this study provides the rationale for an expected behavior: divestment of ambiguous assets as uncertainty increases and naïve diversification becomes optimal. For instance, in Pinar and Paç (2014) and Paç and Pinar (2014) parametric uncertainty for return means is modeled, and their solutions display diminishing risky investment as the uncertainty radius for the mean return estimate increases. Nevertheless, in this study we were able to prove persistence of the uniform portfolio while diminishing in total allocation, providing formal mathematical justification for this behavior under a non-parametric model of uncertainty.

A shortcoming in this study is absence of correlation between the ambiguous market and the unambiguous market. Since global financial trends can affect financial markets on a parallel basis or investor movements can have effects in opposite directions, this would be a valuable aspect of the model. A research direction is to extend the framework in this study to incorporate correlation between the returns for the ambiguous and unambiguous asset groups.

**Acknowledgements** Many thanks to our anonymous referees. The manuscript greatly benefited from their constructive comments regarding both presentation of ideas and depth of content.

### Appendix: Intermediate results and proofs

We begin by proving an upper bound for the absolute deviation in the risk measure caused by the difference in distributions  $P$  and  $Q$ , while holding the ambiguous portfolio selection vector  $w$  fixed. The bound is the product of a constant related to the risk measure  $\mathcal{R}$ , the norm of the ambiguous portfolio  $\|w\|_q$ , and the distance between the two measures  $P$  and  $Q$ , which is shown to be equal to the distance between marginals  $P_1$  and  $Q_1$ .

**Lemma 3** *Let  $\mathcal{R} : L^p(\Omega, \Sigma, \mu) \rightarrow \mathbb{R}$  be a convex, law invariant risk measure with dual representation as discussed in Sect. 2, where  $1 \leq p < \infty$  and  $\frac{1}{p} + \frac{1}{q} = 1$ . Let  $P = P_1 \times P_2$  and  $Q = Q_1 \times P_2$  be product measures on  $(\mathbb{R}^{N+L}, \mathcal{B}(\mathbb{R}^{N+L}))$ , for arbitrary Borel probability measures  $P_1, Q_1$  on  $(\mathbb{R}^N, \mathcal{B}(\mathbb{R}^N))$  and  $P_2$  on  $(\mathbb{R}^L, \mathcal{B}(\mathbb{R}^L))$ . We denote by  $\mathcal{F}^\perp$  the largest sigma algebra independent from  $\sigma(X^{P_2})$ . Note that  $\sigma(X^{P_1}) \subset \mathcal{F}^\perp \perp \sigma(X^{P_2})$ , since  $X^{P_1}$  and  $X^{P_2}$  are independent. Then,*

$$d_p^{\mathbb{R}^{N+L}}(P, Q) = d_p^{\mathbb{R}^N}(P_1, Q_1),$$

and

$$\begin{aligned} & \left| \mathcal{R} \left( \langle X^{P_1}, w \rangle + \langle X^{P_2}, v \rangle \right) - \mathcal{R} \left( \langle X^{Q_1}, w \rangle + \langle X^{P_2}, v \rangle \right) \right| \\ & \leq \sup_{Z: R(Z) < \infty} \|\mathbb{E}[Z | \mathcal{F}^\perp]\|_{L^q} \|w\|_q d_p(P, Q). \end{aligned} \tag{24}$$

*Proof* Let  $\hat{\pi}_1$  be the optimal transportation plan, i.e., the minimizing distribution giving the Kantorovich distance between  $P_1$  and  $Q_1$ . We will define a transportation plan  $\hat{\pi}$  on  $(\mathbb{R}^{N+L} \times \mathbb{R}^{N+L}, \mathcal{B}(\mathbb{R}^{N+L} \times \mathbb{R}^{N+L}))$  between  $P$  and  $Q$  as follows:

$$\hat{\pi}(A \times B \times C \times D) = \hat{\pi}_1(A \times C) \times P_2(B \cap D), \tag{25}$$

where  $A, C \in \mathcal{B}(\mathbb{R}^N)$  and  $B, D \in \mathcal{B}(\mathbb{R}^L)$ . The  $\pi$ -system defined as the product  $A \times B \times C \times D$  of Borel sets generates  $\mathcal{B}(\mathbb{R}^{N+L} \times \mathbb{R}^{N+L})$ , and (25) uniquely defines a measure on  $(\mathbb{R}^{N+L} \times \mathbb{R}^{N+L}, \mathcal{B}(\mathbb{R}^{N+L} \times \mathbb{R}^{N+L}))$ .  $\hat{\pi}$  is a transportation plan between  $P$  and  $Q$ , as projections  $\hat{\pi}(A \times B \times \mathbb{R}^{N+L}) = \hat{\pi}_1(A \times \mathbb{R}^N) \times P_2(B) = P_1(A) \times P_2(B)$  and  $\hat{\pi}(\mathbb{R}^{N+L} \times C \times D) = \hat{\pi}_1(\mathbb{R}^N \times C) \times P_2(D) = Q_1(C) \times P_2(D)$  coincide with  $P$  and  $Q$ , respectively on the  $\pi$ -system of measurable rectangles in  $\mathcal{B}(\mathbb{R}^N) \times \mathcal{B}(\mathbb{R}^L)$ .  $\hat{\pi}$  is a horizontal transportation plan, in the sense that it adopts the plan indicated by  $\hat{\pi}_1$  at each  $(x_{N+1}, \dots, x_{N+L}) \in \mathbb{R}^L$  to redistribute the weight between  $P$  and  $Q$ , and does not shift the distributional weight between two locations  $(\bar{x}_{N+1}, \dots, \bar{x}_{N+L}) \neq (x_{N+1}, \dots, x_{N+L})$  to reach  $Q$  from  $P$ .

We begin by showing  $d_p^{\mathbb{R}^N}(P_1, Q_1) \leq d_p^{\mathbb{R}^{N+L}}(P, Q)$ , denoting by  $\pi$  the optimal transportation plan between  $P$  and  $Q$ , and by  $\pi_1$  its projection  $\pi(\cdot \times \mathbb{R}^L \times \cdot \times \mathbb{R}^L)$  (that this is a transportation plan between  $P_1$  and  $Q_1$  can be checked as above):

$$\begin{aligned} d_p^{\mathbb{R}^{N+L}}(P, Q) &= \left( \int_{\mathbb{R}^{N+L} \times \mathbb{R}^{N+L}} \sum_{i=1}^{N+L} |x_i - y_i|^p d\pi(x, y) \right)^{\frac{1}{p}} \\ &\geq \left( \int_{\mathbb{R}^{N+L} \times \mathbb{R}^{N+L}} \sum_{i=1}^N |x_i - y_i|^p d\pi(x, y) \right)^{\frac{1}{p}} \\ &= \left( \int_{\mathbb{R}^N \times \mathbb{R}^N} \sum_{i=1}^N |x_i - y_i|^p d\pi_1(x, y) \right)^{\frac{1}{p}} \end{aligned} \tag{26}$$

$$\begin{aligned} &\geq \left( \int_{\mathbb{R}^N \times \mathbb{R}^N} \sum_{i=1}^N |x_i - y_i|^p d\hat{\pi}_1(x, y) \right)^{\frac{1}{p}} \\ &= d_p^{\mathbb{R}^N}(P_1, Q_1). \end{aligned} \tag{27}$$

In the above derivations, (26) follows since the integrand is constant with respect to  $x_{N+1}, \dots, x_{N+L}, y_{N+1}, \dots, y_{N+L}$  and (27) follows due to the optimality of  $\hat{\pi}_1$  among the transportation plans between  $P_1$  and  $Q_1$ . The reverse, that is,  $d_p^{\mathbb{R}^N}(P_1, Q_1) \geq d_p^{\mathbb{R}^{N+L}}(P, Q)$ , follows in a similar fashion. In the derivations, we evaluate the integral with respect to  $\hat{\pi}$  separately over complementary sets  $S = \{(x_1, x_2, x_3, x_4) \in \mathbb{R}^N \times \mathbb{R}^L \times \mathbb{R}^N \times \mathbb{R}^L : x_2 = x_4\}$  and  $\bar{S} = \{(x_1, x_2, x_3, x_4) \in \mathbb{R}^N \times \mathbb{R}^L \times \mathbb{R}^N \times \mathbb{R}^L : x_2 \neq x_4\}$ , which consist of ordered pairs in  $\mathbb{R}^{N+L} \times \mathbb{R}^{N+L}$  that agree on the last  $L$  coordinates and those that disagree, respectively.  $S$  is closed,  $\bar{S}$  is open and can be considered as a countable union of open sets of the form  $O = O_1 \times O_2 \times O_3 \times O_4$ , where  $O_1, O_3 \subset \mathbb{R}^N$  and  $O_2, O_4 \subset \mathbb{R}^L$ .  $O \subset \bar{S}$  implies  $O_2 \cap O_4 = \emptyset$ , since otherwise  $O \neq \emptyset$  would contain  $(x_1, x_2, x_3, x_4) \in \mathbb{R}^{N+L} \times \mathbb{R}^{N+L}$  such that  $x_2 = x_4$ . By measure subadditivity,  $\hat{\pi}(\bar{S}) = 0$ , i.e.,  $\bar{S}$  is a  $\hat{\pi}$ -negligible set, leading to following computations:

$$\begin{aligned} d_p^{\mathbb{R}^{N+L}}(P, Q) &= \left( \int_{\mathbb{R}^{N+L} \times \mathbb{R}^{N+L}} \sum_{i=1}^{N+L} |x_i - y_i|^p d\pi(x, y) \right)^{\frac{1}{p}} \\ &\leq \left( \int_{\mathbb{R}^{N+L} \times \mathbb{R}^{N+L}} \sum_{i=1}^{N+L} |x_i - y_i|^p d\hat{\pi}(x, y) \right)^{\frac{1}{p}} \end{aligned} \tag{28}$$

$$\begin{aligned} &= \left( \int_S \sum_{i=1}^{N+L} |x_i - y_i|^p d\hat{\pi}(x, y) + \int_{\bar{S}} \sum_{i=1}^{N+L} |x_i - y_i|^p d\hat{\pi}(x, y) \right)^{\frac{1}{p}} \\ &= \left( \int_S \sum_{i=1}^{N+L} |x_i - y_i|^p d\hat{\pi}(x, y) \right)^{\frac{1}{p}} \end{aligned} \tag{29}$$

$$\begin{aligned} &= \left( \int_{\mathbb{R}^{N+L} \times \mathbb{R}^{N+L}} \sum_{i=1}^N |x_i - y_i|^p d\hat{\pi}(x, y) \right)^{\frac{1}{p}} \\ &= \left( \int_{\mathbb{R}^N \times \mathbb{R}^N} \sum_{i=1}^N |x_i - y_i|^p d\hat{\pi}_1(x, y) \right)^{\frac{1}{p}} \end{aligned} \tag{30}$$

$$= d_p^{\mathbb{R}^N}(P_1, Q_1). \tag{31}$$

Here, (28) follows since  $\pi$  is the optimal transportation plan for  $P$  and  $Q$ , (29) follows from the  $\hat{\pi}$ -negligibility of  $\bar{S}$  and equality of  $(x, y) \mapsto \sum_{i=1}^N |x_i - y_i|^p$  and  $(x, y) \mapsto \sum_{i=1}^{N+L} |x_i - y_i|^p$  on  $S$ , (30) follows since  $\sum_{i=1}^N |x_i - y_i|^p \mathbb{1}_S = \sum_{i=1}^{N+L} |x_i - y_i|^p$  almost everywhere. (31) holds since the integrand is constant with respect to  $x_{N+1}, \dots, x_{N+L}, y_{N+1}, \dots, y_{N+L}$ . We have that  $d_p^{\mathbb{R}^N}(P_1, Q_1) = d_p^{\mathbb{R}^{N+L}}(P, Q)$ .

Corresponding to the transportation plan  $\hat{\pi}$ , there exists a random variable  $Y : (\Omega, \sigma, \mu) \rightarrow \mathbb{R}^{N+L} \times \mathbb{R}^{N+L}$  with image measure  $\mu \circ Y^{-1} = \hat{\pi}$ . We denote the first  $N + L$  components of  $Y$  by  $X^P$ , and the latter part by  $X^Q$ . It can be checked, as above, that the image measure of  $X^P$  and  $X^Q$  coincide with marginals  $P$  and  $Q$ , respectively.

Note that  $\mu(\{\omega \in \Omega : X^{Q_2}(\omega) \neq X^{P_2}(\omega)\}) = \hat{\pi}(\bar{S}) = 0$ , thus  $X^{Q_2} = X^{P_2}$  almost everywhere, and we take  $X^Q = \begin{bmatrix} X^{Q_1} \\ X^{P_2} \end{bmatrix}$  in our calculations. When  $Z$  is chosen from  $\partial\mathcal{R}(\langle X^{P_1}, w \rangle + \langle X^{P_2}, v \rangle)$ , we have:

$$\begin{aligned} &\mathcal{R}(\langle X^{P_1}, w \rangle + \langle X^{P_2}, v \rangle) - \mathcal{R}(\langle X^{Q_1}, w \rangle + \langle X^{P_2}, v \rangle) \\ &\leq \mathbb{E}[\langle \langle X^{P_1}, w \rangle + \langle X^{P_2}, v \rangle, Z \rangle] - R(Z) - \mathbb{E}[\langle \langle X^{Q_1}, w \rangle + \langle X^{P_2}, v \rangle, Z \rangle] + R(Z) \\ &= \mathbb{E}[\langle X^{P_1} - X^{Q_1}, w \rangle, Z] \end{aligned} \tag{32}$$

$$\leq \mathbb{E}[\langle X^{P_1} - X^{Q_1}, w \rangle, Z | \mathcal{F}^\perp] \tag{33}$$

$$\leq \left\| \langle X^{P_1} - X^{Q_1}, w \rangle \right\|_{L^p} \left\| \mathbb{E}[Z | \mathcal{F}^\perp] \right\|_{L^q} \tag{34}$$

$$= \left\| \mathbb{E}[Z | \mathcal{F}^\perp] \right\|_{L^q} \|w\|_q \left( \int_{\mathbb{R}^{N+L} \times \mathbb{R}^{N+L}} \sum_{n=1}^N |x_n - y_n|^p d\hat{\pi}(x, y) \right)^{1/p} \tag{35}$$

$$= \left\| \mathbb{E}[Z | \mathcal{F}^\perp] \right\|_{L^q} \|w\|_q \left( \int_{\mathbb{R}^N \times \mathbb{R}^N} \sum_{n=1}^N |x_n - y_n|^p d\hat{\pi}_1(x, y) \right)^{1/p}$$

$$= \left\| \mathbb{E}[Z | \mathcal{F}^\perp] \right\|_{L^q} \|w\|_q d_p^{\mathbb{R}^N}(P_1, Q_1)$$

$$\leq \sup_{Z: R(Z) < \infty} \left\| \mathbb{E}[Z | \mathcal{F}^\perp] \right\|_{L^q} \|w\|_q d_p^{\mathbb{R}^{N+L}}(P, Q).$$

Equality in (32) follows since  $\langle X^{P_1} - X^{Q_1}, w \rangle$  is measurable with respect to  $\mathcal{F}^\perp$ , as  $X^{P_1}$  and  $X^{Q_1}$  are both independent from  $X^{P_2}$ , i.e.,  $\sigma(X^{P_1}) \perp \sigma(X^{P_2})$  and  $\sigma(X^{Q_1}) \perp \sigma(X^{P_2})$ , and  $\sigma(X^{P_1}) \vee \sigma(X^{Q_1}) \subset \mathcal{F}^\perp$ , where  $\sigma(X^{P_1}) \vee \sigma(X^{Q_1})$  is the smallest  $\sigma$ -algebra containing  $\sigma(X^{P_1}) \cup \sigma(X^{Q_1})$ . Inequality (33) follows from the application of Hölder’s Inequality on the functions  $\langle X^{P_1} - X^{Q_1}, w \rangle$  and  $\mathbb{E}[Z | \mathcal{F}^\perp]$ , and (34) follows from Hölder’s Inequality on vectors  $X^{P_1} - X^{Q_1}$  and  $w$  in  $\mathbb{R}^N$ . The arguments are based on a random variable  $Z$  specific to maximizing  $\mathbb{E}[\langle \langle X^{P_1}, w \rangle + \langle X^{P_2}, v \rangle, Z \rangle] - R(Z)$ , but in the final step, by using the supremum over all  $Z \in \partial\mathcal{R}(X)$  for  $X$  in the problem domain, this dependence is alleviated. Repeating the arguments for  $\mathcal{R}(\langle X^{Q_1}, w \rangle + \langle X^{P_2}, v \rangle) - \mathcal{R}(\langle X^{P_1}, w \rangle + \langle X^{P_2}, v \rangle)$ , we reach the result.  $\square$

It is now proven that the absolute difference in risk measures due to different return distributions is bounded from above, by the product of a constant intrinsic to the risk measure  $\mathcal{R}$ ,  $\|w\|_q$ , and the distance between the distributions. We will prove now that this bound is tight and attained, that is, given  $\kappa > 0$ , for every distribution  $P$ , there is a distribution  $Q$  at distance  $\kappa$  such that the difference in risk is equal to the bound in Lemma 3. In the two lemmas that follow, by restrictions on possibly  $\mathcal{R}$ ,  $P_1$  and  $P_2$ , it is possible to assume  $\sup_{Z: R(Z) < \infty} \left\| \mathbb{E}[Z | \mathcal{F}^\perp] \right\|_{L^q}$  is a constant  $C$ , and for two separate cases (the domain  $L^p$  of  $\mathcal{R}$  being defined for  $p \in (1, \infty)$  or  $p = 1$ ) it is proven that the bound stated in the previous lemma (that can now be pronounced as  $C\kappa\|w\|_q$ ) is tight and attainable. Namely, for every distribution  $P = P_1 \times P_2$  and random variable  $X^P$  with image measure  $P$ , there exists another random variable  $X^Q$  whose image measure  $Q = Q_1 \times P_2$  is at a distance  $d_p(P, Q) = \kappa$  to



$P$ , such that the difference in risk when investing  $(w, v)$  is exactly  $C\kappa\|w\|_q$  for  $P$  and  $Q$ . We begin with the case  $p \in (1, \infty)$ .

**Lemma 4** *Let the risk functional  $\mathcal{R}$  be as defined above. Let  $1 < p < \infty$  and  $q$  be such that  $\frac{1}{p} + \frac{1}{q} = 1$ . Let  $P = P_1 \times P_2$  be a probability measure on  $\mathbb{R}^{N+L}$  with  $P_1, P_2$  probability measures on  $\mathbb{R}^N, \mathbb{R}^L$ , respectively. Let  $X^P = \begin{bmatrix} X^{P_1} \\ X^{P_2} \end{bmatrix}$  be a random variable with image measure  $P$ , and  $\mathcal{F}^\perp \subset \Sigma$  be the largest  $\sigma$ -algebra independent from  $\sigma(X^{P_2})$ . Assume that:*

$$\mathbb{E} \left[ Z|\mathcal{F}^\perp \right] \|_{L^q} = C \text{ for all } Z \in \bigcup_{X \in L^P} \partial \mathcal{R}(X) \text{ with } R(Z) < \infty.$$

Then it holds that for every  $\kappa > 0$  and every  $(w, v) \in \mathbb{R}^N \times \mathbb{R}^L$ , there are measures  $Q_1$  on  $\mathbb{R}^N$  and  $Q = Q_1 \times Q_2 = Q_1 \times P_2$  on  $\mathbb{R}^{N+L}$  such that  $d_p(P, Q) = \kappa$  and

$$\left| \mathcal{R} \left( \langle X^{Q_1}, w \rangle + \langle X^{Q_2}, v \rangle \right) - \mathcal{R} \left( \langle X^{P_1}, w \rangle + \langle X^{P_2}, v \rangle \right) \right| = C\kappa\|w\|_q,$$

i.e., the bound of Lemma 3 holds with equality.

*Proof* Fix a  $Z \in \partial \mathcal{R}(\langle X^{P_1}, w \rangle + \langle X^{P_2}, v \rangle)$  with  $R(Z) < \infty$ . We set  $\bar{Z} = \mathbb{E}[Z|\mathcal{F}^\perp]$  and define a random variable  $X^Q$  as follows:

$$\begin{aligned} X_n^Q &= X_n^P + c_1(n) |w_n|^{\frac{q}{p}} \quad \text{with} \\ c_1(n) &= \frac{\text{sign}(w_n) \text{sign}(\bar{Z}) c_2}{\|w\|_q^{\frac{q}{p}}} |\bar{Z}|^{\frac{q}{p}} \end{aligned}$$

for  $n \in \{1, \dots, N\}$  and a constant  $c_2 > 0$ . For  $n \in \{N + 1, \dots, N + L\}$ , we let  $X_n^Q = X_n^P$ . Setting  $c_1 = |c_1(n)|$ , it follows that:

$$c_1^p |w_n|^q = \left| X_n^Q - X_n^P \right|^p, \quad \forall n : 1 \leq n \leq N. \tag{36}$$

Also,

$$\begin{aligned} \left| \sum_{n=1}^N w_n \left( X_n^Q - X_n^P \right) \right|^p &= \left| \sum_{n=1}^N w_n c_1(n) |w_n|^{\frac{q}{p}} \right|^p \\ &= \left| \sum_{n=1}^N \frac{\text{sign}(\bar{Z}) c_2}{\|w\|_q^{\frac{q}{p}}} |\bar{Z}|^{\frac{q}{p}} |w_n|^q \right|^p \\ &= \frac{c_2^p}{\|w\|_q^{pq}} |\bar{Z}|^q \|w\|_q^{pq} = c_2^p |\bar{Z}|^q. \end{aligned} \tag{37}$$

$c_2$  is a parameter for adjusting the distance between the distributions so that  $d_p(P, Q) = \kappa$ , which is achieved unless  $Z = 0$  (which would require  $C = 0$ , leading to a triviality where  $\mathcal{R}$  is constant). Note that  $X^{Q_1}$  is  $\mathcal{F}^\perp$ -measurable since  $X^{P_1}$  is  $\sigma(X^{P_1})$ -measurable ( $\sigma(X^{P_1}) \subset \mathcal{F}^\perp$ ) and  $\text{sign}(\bar{Z}) |\bar{Z}|^{\frac{q}{p}}$  is  $\mathcal{F}^\perp$ -measurable. The independence of  $\mathcal{F}^\perp$  and  $\sigma(X^{P_2})$  implies the product form of  $Q = \mu \circ (X^Q)^{-1} = \mu \circ (X^{Q_1})^{-1} \times \mu \circ (X^{P_2})^{-1} = Q_1 \times P_2$ . The result is obtained as follows:

$$\begin{aligned}
 & \left| \mathcal{R} \left( \langle X^{Q_1}, w \rangle + \langle X^{Q_2}, v \rangle \right) - \mathcal{R} \left( \langle X^{P_1}, w \rangle + \langle X^{P_2}, v \rangle \right) \right| \\
 & \geq \mathbb{E} \left( \langle X^{Q_1}, w \rangle + \langle X^{Q_2}, v \rangle \right) - \mathbb{E} \left( \langle X^{P_1}, w \rangle + \langle X^{P_2}, v \rangle \right) \\
 & = \mathbb{E} \left( \langle X^{Q_1} - X^{P_1}, w \rangle Z \right) \\
 & = \mathbb{E} \left( \left( \sum_{n=1}^N (X_n^Q - X_n^P) w_n \right) \bar{Z} \right) \\
 & = \mathbb{E} \left( \left| \sum_{n=1}^N (X_n^Q - X_n^P) w_n \right| |\bar{Z}| \right) \tag{38}
 \end{aligned}$$

$$= \left( \mathbb{E} \left( \left| \sum_{n=1}^N (X_n^Q - X_n^P) w_n \right|^p \right) \right)^{1/p} \|\bar{Z}\|_{L^q} \tag{39}$$

$$= \left( \mathbb{E} \left( \sum_{n=1}^N |X_n^Q - X_n^P| |w_n| \right)^p \right)^{1/p} \|\bar{Z}\|_{L^q} \tag{40}$$

$$= \left( \mathbb{E} \left( \|X^{Q_1} - X^{P_1}\|_p \|w\|_q \right)^p \right)^{1/p} \|\bar{Z}\|_{L^q} \tag{41}$$

$$= \|\bar{Z}\|_{L^q} \|w\|_q \left( \int_{\Omega} \sum_{n=1}^N |X_n^Q - X_n^P|^p d\mu \right)^{1/p} \tag{42}$$

$$\geq \|\bar{Z}\|_{L^q} \|w\|_q d_p^{\mathbb{R}^N}(P_1, Q_1) \tag{43}$$

$$= \|\bar{Z}\|_{L^q} \|w\|_q \kappa.$$

Transitions to (38) and (40) are possible due to the incorporation of  $sign(w_n)$  and  $sign(Z)$  in  $X^{Q_1} - X^{P_1}$ , which makes all terms in the sum non-negative. (39) and (41) are applications of Hölder’s Inequality, where conditions for equality in Hölder’s are assured by (37) and (36), respectively. Equation (39) is possible since  $\left| \sum_{n=1}^N (X_n^Q - X_n^P) w_n \right|^p$  is equal to  $|\bar{Z}|^q$ , when constant multipliers set aside. Similarly, (36) implies the condition for equality on transition to (41). Inequality in (43) follows since  $\tilde{\pi} = \mu \circ \left( \begin{bmatrix} X^{P_1} \\ X^{Q_1} \end{bmatrix} \right)^{-1}$  is a transportation plan between  $P_1$  and  $Q_1$ , and the integral in (42) is equivalent to  $\int_{\mathbb{R}^N \times \mathbb{R}^N} \sum_{n=1}^N |x_n - y_n|^p d\tilde{\pi}(x, y)$ , while  $d_p(P_1, Q_1) = \kappa$  is given by the optimal transportation plan between  $P_1$  and  $Q_1$ .  $\square$

A similar result on the tightness of the bound given in Lemma 3 follows for the case  $p = 1$ .

**Lemma 5** *Let  $\mathcal{R}$  be a risk measure as defined above. Let  $P = P_1 \times P_2$  be a probability measure on  $\mathbb{R}^{N+L}$  where  $P_1, P_2$  are probability measures on  $\mathbb{R}^N, \mathbb{R}^L$ , respectively, and  $X^P \in L^1(\Omega, \Sigma, \mu)$  a random variable with image measure  $P$ . Let  $\mathcal{F}^\perp \subset \Sigma$  be the largest  $\sigma$ -algebra independent from  $\sigma(X^{P_2})$ . Assume:*

$$\|Z\|_{L^\infty} = C, \quad \mu(\{\omega \in \Omega : |Z|(\omega) \notin \{0, C\}\}) = 0 \tag{44}$$

for all  $Z \in \partial\mathcal{R}(X), X \in L^1(\Omega, \Sigma, \mu)$ . In addition, assume for all  $\epsilon \in (0, \frac{1}{2})$  that there exists  $B \in \mathcal{F}^\perp$  such that  $\mu(B) > 0$ , and either

$$\mu(B \cap \{\omega \in \Omega : Z(\omega) = C\}) > (1 - \epsilon) \mu(B)$$

or

$$\mu(B \cap \{\omega \in \Omega : Z(\omega) = -C\}) > (1 - \epsilon) \mu(B)$$

holds. Then for every  $\kappa > 0$ , there is a probability measure  $Q = Q_1 \times P_2$  on  $\mathbb{R}^{N+L}$  with  $Q_1$  a probability measure on  $\mathbb{R}^N$  such that  $d_1^{\mathbb{R}^{N+L}}(P, Q) = \kappa$  and

$$\left| \mathcal{R} \left( \langle X^{Q_1}, w \rangle + \langle X^{Q_2}, v \rangle \right) - \mathcal{R} \left( \langle X^{P_1}, w \rangle + \langle X^{P_2}, v \rangle \right) \right| = C\kappa \|w\|_\infty.$$

*Proof* Take  $Z \in \partial \mathcal{R}(\langle X^{P_1}, w \rangle + \langle X^{P_2}, v \rangle)$ . For  $\epsilon \in (0, \frac{1}{2})$ , we consider the case where there exists  $B \in \mathcal{F}^\perp$  with  $\mu(B \cap \{\omega \in \Omega : Z(\omega) = C\}) > (1 - \epsilon) \mu(B)$ . For the alternative case, the result follows in a similar manner with a change of sign in  $c_1(n)$  defined below. Let

$$X_n^Q = X_n^P + c_1(n) \tag{45}$$

$$c_1(n) = \begin{cases} c_2 \text{sign}(w_n) \mathbb{1}_B, & \text{if } |w_n| = \|w\|_\infty \\ 0, & \text{otherwise} \end{cases} \tag{46}$$

for  $n \in \{1, \dots, N\}$  and a constant  $c_2 > 0$ . Again,  $c_2$  is a constant for adjusting the distance between  $P$  and  $Q$  to  $\kappa$ . Let  $X_n^Q = X_n^P$  for  $n \in \{N + 1, \dots, N + L\}$ . Let us label the number of entries in  $w$  that set its norm, i.e.,  $|\{n : |w_n| = \|w\|_\infty\}|$  by  $\chi$ .

Via the inverse image of  $X^Q$ , we obtain a distribution  $Q = \mu \circ (X^Q)^{-1}$  on  $\mathbb{R}^{N+L}$ . Again, since  $B \in \mathcal{F}^\perp$  and  $\mathcal{F}^\perp$  and  $\sigma(X^{P_2})$  are independent, the product form of  $Q = Q_1 \times Q_2$  follows. We have:

$$\begin{aligned} & \mathcal{R} \left( \langle X^{Q_1}, w \rangle + \langle X^{Q_2}, v \rangle \right) - \mathcal{R} \left( \langle X^{P_1}, w \rangle + \langle X^{P_2}, v \rangle \right) \\ & \geq \mathbb{E} \left( \langle X^{Q_1} - X^{P_1}, w \rangle Z \right) \end{aligned} \tag{47}$$

$$\begin{aligned} & = \mathbb{E} \left( \langle X^{Q_1} - X^{P_1}, w \rangle Z \mathbb{1}_B \right) \\ & = \mathbb{E} \left( \langle X^{Q_1} - X^{P_1}, w \rangle Z \mathbb{1}_{\{Z > 0\} \cap B} \right) + \mathbb{E} \left( \langle X^{Q_1} - X^{P_1}, w \rangle Z \mathbb{1}_{\{Z \leq 0\} \cap B} \right) \\ & \geq \mathbb{E} \left( \langle X^{Q_1} - X^{P_1}, w \rangle Z \mathbb{1}_{\{Z > 0\} \cap B} \right) - c_2 C \chi \|w\|_\infty \mu(\{Z \leq 0\} \cap B) \\ & \geq \mathbb{E} \left( \langle X^{Q_1} - X^{P_1}, w \rangle Z \mathbb{1}_{\{Z > 0\} \cap B} \right) - \epsilon c_2 C \chi \|w\|_\infty \mu(B) \\ & = \mathbb{E} \left( \left| \langle X^{Q_1} - X^{P_1}, w \rangle \mathbb{1}_{\{Z > 0\} \cap B} \right| |Z| \right) - \epsilon c_2 C \chi \|w\|_\infty \mu(B) \\ & = \mathbb{E} \left( \left| \langle X^{Q_1} - X^{P_1}, w \rangle \mathbb{1}_{\{Z > 0\} \cap B} \right| \|Z\|_{L^\infty} - \epsilon c_2 C \chi \|w\|_\infty \mu(B) \right) \end{aligned} \tag{48}$$

$$\begin{aligned} & = c_2 \chi \|w\|_\infty \mu(\{Z > 0\} \cap B) \|Z\|_{L^\infty} - \epsilon c_2 C \chi \|w\|_\infty \mu(B) \\ & > (1 - \epsilon) c_2 \chi \|w\|_\infty \mu(B) C - \epsilon c_2 C \chi \|w\|_\infty \mu(B) \\ & = (1 - 2\epsilon) c_2 \chi \|w\|_\infty \mu(B) C \\ & = (1 - 2\epsilon) \mathbb{E} \left( \int_\Omega \sum_{n=1}^N |c_1(n)| d\mu \right) \|w\|_\infty C \\ & = (1 - 2\epsilon) \mathbb{E} \left( \|X^{Q_1} - X^{P_1}\|_1 \right) \|w\|_\infty C \\ & \geq (1 - 2\epsilon) d_1^{\mathbb{R}^N}(P_1, Q_1) \|w\|_\infty C \end{aligned} \tag{49}$$

$$\begin{aligned}
 &= (1 - 2\epsilon) \kappa \|w\|_\infty C & (50) \\
 &= (1 - 2\epsilon) \kappa \|w\|_\infty C.
 \end{aligned}$$

As  $X^{Q_2} = X^{P_2}$ , and  $Z$  is possibly not in  $\partial\mathcal{R}(\langle X^{Q_1}, w \rangle + \langle X^{Q_2}, v \rangle)$ , (47) follows. Equality (48) is assured since the  $L^1$  function has value 0 on  $\{Z \neq \|Z\|_{L^\infty}\}$ . Equation (49) follows since  $X^{Q_1}$  and  $X^{P_1}$  jointly have an image measure which is a transportation between  $Q_1$  and  $P_1$  (which is not necessarily optimal), and (50) since  $d_1^{\mathbb{R}^N}(P_1, Q_1) = d_1^{\mathbb{R}^{N+L}}(P, Q) = \kappa$ . With the above holding for all  $\epsilon \in (0, \frac{1}{2})$ , the result is established.  $\square$

The perturbation used in Lemma 4 to obtain  $X^Q$  from  $X^P$  was based on the conditional expectation  $\mathbb{E}[Z|\mathcal{F}^\perp]$  of the random variable  $Z \in \partial\mathcal{R}(\langle X^{P_1}, w \rangle + \langle X^{P_2}, v \rangle)$ , whereas in Lemma 5 we used the characteristic function of the set  $B$  defined in the assumptions. However, the two results are analogous, and agree the figure in Lemma 3, since the assumptions in Lemma 5 imply that  $\|Z\|_\infty = \|\mathbb{E}[Z|\mathcal{F}^\perp]\|_\infty$  for all  $Z \in \partial\mathcal{R}(X)$ ,  $X \in L^1(\Omega, \Sigma, \mu)$ . To observe this, one can check that the value of  $\mathbb{E}[Z|\mathcal{F}^\perp]$  on the set  $B$ , as referred to in Lemma 5, is either inside  $((1 - 2\epsilon) C, C]$  or  $[-C, -(1 - 2\epsilon) C)$ .

With the proofs of Lemmas 4 and 5, we show that the bound in Lemma 3 is tight and attained, and also we are ready to prove Proposition 1.

*Proof of Proposition 1* With fixed portfolio selection  $(w, v)$ , the maximum (absolute) difference in the risk measure due to distributions  $P$  and  $Q$  ( $d_p(P, Q) \leq \kappa$ ) is  $C\kappa\|w\|_q$ , by Lemma 3. By assumption,  $\mathcal{R}$  and  $P$  satisfy the assumptions in Lemma 4 or 5, as necessitated by the domain of the risk measure, i.e., the value of  $p$ . Moreover, the deviation in the risk measure due to the perturbed measure  $Q = Q_1 \times P_2$  in the proofs of Lemmas 4 and 5 is in the positive direction, therefore we can write:

$$\begin{aligned}
 &\sup_{\bar{P} \in \tilde{\mathcal{E}}_\kappa(P)} \mathcal{R}(\langle X^{\bar{P}_1}, w \rangle + \langle X^{\bar{P}_2}, v \rangle) \\
 &= \mathcal{R}(\langle X^{Q_1}, w \rangle + \langle X^{Q_2}, v \rangle) & (51) \\
 &= \mathcal{R}(\langle X^{P_1}, w \rangle + \langle X^{P_2}, v \rangle) + C\kappa\|w\|_q. & (52)
 \end{aligned}$$

$\square$

*Proof of Lemma 1* Taking arbitrary  $Z \in \partial\mathcal{R}(\langle X^{P_1}, w^1 \rangle + \langle X^{P_2}, v \rangle)$ ,

$$\begin{aligned}
 &\mathcal{R}(\langle X^{P_1}, w^1 \rangle + \langle X^{P_2}, v \rangle) - \mathcal{R}(\langle X^{P_1}, w^2 \rangle + \langle X^{P_2}, v \rangle) \\
 &\leq \mathbb{E}[\langle \langle X^{P_1}, w^1 \rangle + \langle X^{P_2}, v \rangle, Z \rangle] - \mathbb{E}[\langle \langle X^{P_1}, w^2 \rangle + \langle X^{P_2}, v \rangle, Z \rangle] \\
 &= \mathbb{E}[\langle X^{P_1}, w^1 - w^2 \rangle, Z] \\
 &= \mathbb{E}[\langle X^{P_1}, w^1 - w^2 \rangle, \bar{Z}] \\
 &= \mathbb{E}[\langle X^{P_1}, w^1 - w^2 \rangle \mathbb{1}_{\{\bar{Z} \neq 0\}} \bar{Z}] & (53)
 \end{aligned}$$

$$\leq \mathbb{E}[\langle X^{P_1}, w^1 - w^2 \rangle \mathbb{1}_{\{\bar{Z} \neq 0\}}]^p]^{\frac{1}{p}} \|\bar{Z}\|_{L^q} & (54)$$

$$\leq \mathbb{E}[\langle \|X^{P_1}\|_p \|w^1 - w^2\|_q \rangle^p \mathbb{1}_{\{\bar{Z} \neq 0\}}]^{\frac{1}{p}} \|\bar{Z}\|_{L^q} & (55)$$

$$= \mathbb{E}[\langle \|X^{P_1}\|_p^p \mathbb{1}_{\{\bar{Z} \neq 0\}} \rangle]^{\frac{1}{p}} \|w^1 - w^2\|_q C. & (56)$$

Inequalities (54) and (55) follow due to Hölder’s Inequality applied on functions  $\langle X^{P_1}, w^1 - w^2 \rangle \mathbb{1}_{\{\bar{Z} \neq 0\}} \in L^p$  and  $\bar{Z} \in L^q$ , and vectors in  $\mathbb{R}^N$ , respectively. The result, (56), follows due to the assumptions in Lemmas 4 and 5 that imply  $\|\bar{Z}\|_{L^q} = C$  for  $Z \in \partial\mathcal{R}(\langle X^{P_1}, w^1 \rangle + \langle X^{P_2}, v \rangle)$ .  $\square$

*Proof of Lemma 2*  $\kappa \geq \frac{\|w - w^{u,v}\|_q}{\|w\|_q - \|w^{u,v}\|_q} \mathbb{E} \left[ \|X^{P_1}\|_p^p \mathbb{1}_{\{\bar{Z} \neq 0\}} \right]^{\frac{1}{p}}$  implies, by multiplying both sides by  $C(\|w\|_q - \|w^{u,v}\|_q)$ :

$$\begin{aligned} C\kappa (\|w\|_q - \|w^{u,v}\|_q) &\geq C (\|w\|_q - \|w^{u,v}\|_q) \frac{\|w - w^{u,v}\|_q}{\|w\|_q - \|w^{u,v}\|_q} \mathbb{E} \left[ \|X^{P_1}\|_p^p \mathbb{1}_{\{\bar{Z} \neq 0\}} \right]^{\frac{1}{p}} \\ &= C \|w - w^{u,v}\|_q \mathbb{E} \left[ \|X^{P_1}\|_p^p \mathbb{1}_{\{\bar{Z} \neq 0\}} \right]^{\frac{1}{p}} \\ &\geq \mathcal{R}(\langle X^{P_1}, w^{u,v} \rangle + \langle X^{P_2}, v \rangle) - \mathcal{R}(\langle X^{P_1}, w \rangle + \langle X^{P_2}, v \rangle), \end{aligned}$$

where the last inequality follows by Lemma 1. Regrouping terms above, we have:

$$\mathcal{R}(\langle X^{P_1}, w \rangle + \langle X^{P_2}, v \rangle) + C\kappa \|w\|_q \geq \mathcal{R}(\langle X^{P_1}, w^{u,v} \rangle + \langle X^{P_2}, v \rangle) + C\kappa \|w^{u,v}\|_q,$$

for all  $w \in B$  ( $\|w\|_q \neq \|w^{u,v}\|_q$  is assumed, which holds true for  $w \neq w^{u,v}$  on a hyperplane of fixed  $\langle \mathbb{1}, w \rangle = \langle \mathbb{1}, w^{u,v} \rangle$ , since  $w^{u,v}$  uniquely minimizes  $\|\cdot\|_q$  on that hyperplane).  $\square$

*Proof of Proposition 2* Case  $\langle \mathbb{1}, v \rangle = 1, 1 \leq p < \infty$

In this case  $w^{u,v} = 0$ , hence  $\|w - w^{u,v}\|_q = \|w\|_q - \|w^{u,v}\|_q = \|w\|_q$ . Thus the condition in Lemma 2 is satisfied by all  $w \neq 0$  if

$$\kappa \geq \mathbb{E} \left[ \|X^{P_1}\|_p^p \mathbb{1}_{\{\bar{Z} \neq 0\}} \right]^{\frac{1}{p}},$$

and the optimality of  $w^{u,v} = 0$  follows.

Case  $\langle \mathbb{1}, v \rangle > 1, p = 1$

In this case,  $1 - \langle \mathbb{1}, v \rangle < 0$ , and  $w^{u,v} = \frac{1 - \langle \mathbb{1}, v \rangle}{N} \mathbb{1}$  is composed of negative entries. For any  $w$  such that  $\langle \mathbb{1}, w \rangle = 1 - \langle \mathbb{1}, v \rangle$ ,  $w \neq w^{u,v}$ , we let  $n^* = \arg \max_{1 \leq n \leq N} \left| w_n - \frac{1 - \langle \mathbb{1}, v \rangle}{N} \right|$ . If  $w_{n^*} < \frac{1 - \langle \mathbb{1}, v \rangle}{N}$ ,  $\|w - w^{u,v}\|_\infty = \left| w_{n^*} - \frac{1 - \langle \mathbb{1}, v \rangle}{N} \right| = \|w\|_\infty - \|w^{u,v}\|_\infty$ . Otherwise,  $w_{n^*} > \frac{1 - \langle \mathbb{1}, v \rangle}{N}$ ,  $w_{n^*} = \max_{1 \leq n \leq N} w_n$ , and

$$\begin{aligned} \min_{1 \leq n \leq N} w_n &\leq \frac{1 - \langle \mathbb{1}, v \rangle}{N} - \frac{w_{n^*} - \frac{1 - \langle \mathbb{1}, v \rangle}{N}}{N - 1} \\ &= \frac{1 - \langle \mathbb{1}, v \rangle}{N} - \frac{\|w - w^{u,v}\|_\infty}{N - 1}. \end{aligned}$$

Since  $\min_{1 \leq n \leq N} w_n \geq -\|w\|_\infty$ ,  $\frac{1 - \langle \mathbb{1}, v \rangle}{N} - \frac{\|w - w^{u,v}\|_\infty}{N - 1} \geq -\|w\|_\infty$ , and noting that  $\frac{1 - \langle \mathbb{1}, v \rangle}{N} = -\|w^{u,v}\|_\infty$ , we have:

$$\|w\|_\infty - \|w^{u,v}\|_\infty \geq \frac{\|w - w^{u,v}\|_\infty}{N - 1},$$

equivalently:

$$\frac{\|w - w^{u,v}\|_\infty}{\|w\|_\infty - \|w^{u,v}\|_\infty} \leq N - 1.$$

If  $\kappa \geq (N - 1) \mathbb{E} \left[ \|X^{P_1}\|_1 \mathbb{1}_{\{\bar{Z} \neq 0\}} \right]$ , then  $\kappa \geq \frac{\|w - w^{u,v}\|_\infty}{\|w\|_\infty - \|w^{u,v}\|_\infty} \mathbb{E} \left[ \|X^{P_1}\|_1 \mathbb{1}_{\{\bar{Z} \neq 0\}} \right]$  holds for all  $w \neq w^{u,v}$  such that  $\langle \mathbb{1}, w \rangle = 1 - \langle \mathbb{1}, v \rangle$ , and  $w^{u,v}$  is optimal to (10)–(11).

Case  $\langle \mathbb{1}, v \rangle < 1, p = 1$

In this case,  $w^{u,v} = \frac{1 - \langle \mathbb{1}, v \rangle}{N} \mathbb{1}$  is composed of positive entries. Similar to the case where  $\langle \mathbb{1}, v \rangle > 1$ , we set  $n^* = \arg \max_{1 \leq n \leq N} \left| w_n - \frac{1 - \langle \mathbb{1}, v \rangle}{N} \right|$ . If  $w_{n^*} > \frac{1 - \langle \mathbb{1}, v \rangle}{N}$ ,  $\|w - w^{u,v}\|_\infty = \left( w_{n^*} - \frac{1 - \langle \mathbb{1}, v \rangle}{N} \right) = \|w\|_\infty - \|w^{u,v}\|_\infty$ . Otherwise,  $w_{n^*} < \frac{1 - \langle \mathbb{1}, v \rangle}{N}$ ,  $w_{n^*} = \min_{1 \leq n \leq N} w_n$ , and

$$\begin{aligned} \max_{1 \leq n \leq N} w_n &\geq \frac{1 - \langle \mathbb{1}, v \rangle}{N} + \frac{w_{n^*} - \frac{1 - \langle \mathbb{1}, v \rangle}{N}}{N - 1} \\ &= \frac{1 - \langle \mathbb{1}, v \rangle}{N} + \frac{\|w\|_\infty - \|w^{u,v}\|_\infty}{N - 1}. \end{aligned}$$

Since  $\|w\|_\infty \geq \max_{1 \leq n \leq N} w_n$ , and  $\|w^{u,v}\|_\infty = \frac{1 - \langle \mathbb{1}, v \rangle}{N}$ , we can write:

$$\begin{aligned} \|w\|_\infty - \|w^{u,v}\|_\infty &\geq \frac{\|w - w^{u,v}\|_\infty}{N - 1} \\ N - 1 &\geq \frac{\|w - w^{u,v}\|_\infty}{\|w\|_\infty - \|w^{u,v}\|_\infty}. \end{aligned}$$

Again, if  $\kappa \geq (N - 1) \mathbb{E} \left[ \|X^{P_1}\|_1 \mathbb{1}_{\{\bar{Z} \neq 0\}} \right]$ , the condition for Lemma 2 is satisfied in the feasible region and  $w^{u,v}$  is optimal to (10)–(11).

Case  $\langle \mathbb{1}, v \rangle \neq 1, p = 2$

Let  $f_2, \dots, f_N$  be unit vectors orthogonal to each other and  $w^{u,v}$ . Then a unique selection of  $c_2, \dots, c_N \in \mathbb{R}$  gives  $w = w^{u,v} + \sum_{i=2}^N c_i f_i$  (since  $w^{u,v} = \frac{1 - \langle \mathbb{1}, v \rangle}{N} \mathbb{1}$ ,  $f_2, \dots, f_N$  are orthogonal to  $w^{u,v}$ , and  $\langle \mathbb{1}, w \rangle = 1 - \langle \mathbb{1}, v \rangle$ ,  $\langle \mathbb{1}, w^{u,v} \rangle = \langle \mathbb{1}, w \rangle$  implies that the coefficient of  $w^{u,v}$  in  $w$  is equal to 1). Then:

$$\begin{aligned} \frac{\|w - w^{u,v}\|_2}{\|w\|_2 - \|w^{u,v}\|_2} &= \frac{\|w - w^{u,v}\|_2}{\left( \frac{(1 - \langle \mathbb{1}, v \rangle)^2}{N} + \sum_{i=2}^N c_i^2 \right)^{\frac{1}{2}} - \frac{|1 - \langle \mathbb{1}, v \rangle|}{\sqrt{N}}} \\ &= \frac{\|w - w^{u,v}\|_2 \left[ \left( \frac{(1 - \langle \mathbb{1}, v \rangle)^2}{N} + \sum_{i=2}^N c_i^2 \right)^{\frac{1}{2}} + \frac{|1 - \langle \mathbb{1}, v \rangle|}{\sqrt{N}} \right]}{\frac{(1 - \langle \mathbb{1}, v \rangle)^2}{N} + \sum_{i=2}^N c_i^2 - \frac{(1 - \langle \mathbb{1}, v \rangle)^2}{N}} \\ &= \frac{\|w - w^{u,v}\|_2 \left[ \left( \frac{(1 - \langle \mathbb{1}, v \rangle)^2}{N} + \sum_{i=2}^N c_i^2 \right)^{\frac{1}{2}} + \frac{|1 - \langle \mathbb{1}, v \rangle|}{\sqrt{N}} \right]}{\|w - w^{u,v}\|_2^2} \\ &= \frac{\left( \frac{(1 - \langle \mathbb{1}, v \rangle)^2}{N} + \sum_{i=2}^N c_i^2 \right)^{\frac{1}{2}} + \frac{|1 - \langle \mathbb{1}, v \rangle|}{\sqrt{N}}}{\|w - w^{u,v}\|_2} \\ &= \left( \frac{(1 - \langle \mathbb{1}, v \rangle)^2}{N \|w - w^{u,v}\|_2^2} + 1 \right)^{\frac{1}{2}} + \frac{|1 - \langle \mathbb{1}, v \rangle|}{\sqrt{N} \|w - w^{u,v}\|_2}. \end{aligned}$$

Defining the set  $B$  as  $B : \{w \in \mathbb{R}^N : \|w - w^{\mu,v}\|_2 \geq D\}$ , the above equality implies that inside the set  $B$ ,

$$\frac{\|w - w^{\mu,v}\|_2}{\|w\|_2 - \|w^{\mu,v}\|_2} \leq \left( \frac{(1 - \langle \mathbb{1}, v \rangle)^2}{ND^2} + 1 \right)^{\frac{1}{2}} + \frac{|1 - \langle \mathbb{1}, v \rangle|}{\sqrt{ND}}.$$

If the value of  $\kappa$  satisfies

$$\kappa \geq \left[ \left( \frac{(1 - \langle \mathbb{1}, v \rangle)^2}{ND^2} + 1 \right)^{\frac{1}{2}} + \frac{|1 - \langle \mathbb{1}, v \rangle|}{\sqrt{ND}} \right] \mathbb{E} \left[ \|X^{P_1}\|_2^2 \mathbb{1}_{\{\tilde{z} \neq 0\}} \right]^{\frac{1}{2}}$$

then

$$\kappa \geq \frac{\|w - w^{\mu,v}\|_2}{\|w\|_2 - \|w^{\mu,v}\|_2} \mathbb{E} \left[ \|X^{P_1}\|_2^2 \mathbb{1}_{\{\tilde{z} \neq 0\}} \right]^{\frac{1}{2}}$$

for  $w \in B$ , and by Lemma 2, a solution with better objective value than  $w^{\mu,v}$  can only be inside  $\{w \in \mathbb{R}^N : \|w - w^{\mu,v}\|_2 < D\}$ .

Case  $p \notin \{1, 2\}$ ,  $\langle \mathbb{1}, v \rangle \neq 1$

In this case, we show that for an increasing sequence  $\kappa_n$ , the optimal solution gets to fall inside a smaller neighborhood surrounding  $w^{\mu,v}$  as  $\kappa_n \rightarrow \infty$ . We define the set

$$A_n = \{w \in \mathbb{R}^N : \langle \mathbb{1}, w \rangle = 1 - \langle \mathbb{1}, v \rangle\},$$

$$\mathcal{R}(\langle X^{P_1}, w^{\mu,v} \rangle + \langle X^{P_2}, v \rangle) + C\|w^{\mu,v}\|_q \kappa_n \geq \mathcal{R}(\langle X^{P_1}, w \rangle + \langle X^{P_2}, v \rangle) + C\|w\|_q \kappa_n.$$

Since  $C$  and  $\kappa_n$  are positive,  $\mathcal{R}(\langle X^{P_1}, \cdot \rangle + \langle X^{P_2}, v \rangle)$  and  $\|\cdot\|_q$  are convex functions of  $w$ ,  $A_n$  is a closed and convex set; and since  $\|w\|_q - \|w^{\mu,v}\|_q > 0$  for  $w \neq w^{\mu,v}$  with  $\langle \mathbb{1}, w \rangle = 1 - \langle \mathbb{1}, v \rangle$ ,  $A_n$  is monotone decreasing and  $\bigcap_{n=1}^\infty A_n = \{w^{\mu,v}\}$ .  $A_n$  is bounded, since,  $\mathcal{R}(\langle X^{P_1}, w \rangle + \langle X^{P_2}, v \rangle)$  is bounded from below (the nominal problem (5)–(6) is well-posed), and thus  $\|w\|_q$  is bounded from above. Being closed and bounded, compactness of  $A_n$  follows.

Let  $B_n^\epsilon = A_n \cap \{w \in \mathbb{R}^N : \|w - w^{\mu,v}\|_q \geq \epsilon\}$ .  $\bigcap_{n=1}^\infty B_n^\epsilon = \emptyset$ , since  $\bigcap_{n=1}^\infty A_n = \{w^{\mu,v}\}$  and  $w^{\mu,v} \notin B_n^\epsilon, \forall n \in \mathbb{N}$ . Compactness of  $B_n^\epsilon, n \in \mathbb{N}$ , implies that there is  $N^\epsilon \in \mathbb{N}$  such that  $B_{N^\epsilon}^\epsilon = \emptyset$ , that is, the optimal solution is inside  $\{w \in \mathbb{R}^N : \|w - w^{\mu,v}\|_q < \epsilon\}$  for  $\kappa \geq \kappa_{N^\epsilon}$ . □

*Proof of Proposition 3* Case  $p = 1$

When  $\kappa$  exceeds (14),  $w^{\mu,v}$  is the optimal solution for the inner problem for all  $v \in \mathbb{R}^L$ , and we are seeking the solution of the problem (17). For some  $s \neq 0$ , let us fix the total allocation to ambiguous assets as  $s$ , i.e.,  $\langle \mathbb{1}, w^{\mu,v} \rangle = s$ . The objective value for a solution  $v \in \mathbb{R}^L$  with  $\langle \mathbb{1}, v \rangle = 1 - s$  is  $\mathcal{R}(\frac{s}{N} \langle X^{P_1}, \mathbb{1} \rangle + \langle X^{P_2}, v \rangle) + C\kappa \frac{|s|}{N}$ . Here, the term on the left remains constant as  $\kappa \rightarrow \infty$ , and the term on the right tends to infinity. Let  $\tilde{v} \in \mathbb{R}^L, \langle \mathbb{1}, \tilde{v} \rangle = 1$ . The objective value of  $\tilde{v}$  in (17) is  $\mathcal{R}(\langle X^{P_2}, \tilde{v} \rangle)$ .  $\tilde{v}$  has the same objective value in (17) for all  $\kappa$ , and when  $\kappa$  exceeds  $\frac{N}{C|s|} [\mathcal{R}(\langle X^{P_2}, \tilde{v} \rangle) - \mathcal{R}(\frac{s}{N} \langle X^{P_1}, \mathbb{1} \rangle + \langle X^{P_2}, v \rangle)]$ ,  $v \in \mathbb{R}^L$  with  $\langle \mathbb{1}, v \rangle = 1 - s$  is suboptimal.

Case  $p = 2$



Again, for  $\tilde{v} \in \mathbb{R}^L$  with  $\langle \mathbb{1}, \tilde{v} \rangle = 1$ , the objective value in (16) is  $\mathcal{R}(\langle X^{P_2}, \tilde{v} \rangle)$ . For  $v \in \mathbb{R}^L$  with  $\langle \mathbb{1}, v \rangle = 1 - s, s \neq 0$ , the objective value is greater than

$$g(v) = \mathcal{R}\left(\frac{s}{N}\langle X^{P_1}, \mathbb{1} \rangle + \langle X^{P_2}, v \rangle\right) + C\kappa|s|\left(\frac{1}{\sqrt{N}} - D\left(\frac{1}{\kappa} \sup_{v \in \mathbb{R}^L} \mathbb{E}\left(\|X^{P_1}\|_2^2 \mathbb{1}_{\{Z_v \neq 0\}}\right)^{\frac{1}{2}} + 1\right)\right).$$

There exists  $k \in \mathbb{N}$  such that for all  $\kappa > k$ ,  $D$  can be picked small enough to assure  $\frac{1}{\sqrt{N}} - D\left(\frac{1}{\kappa} \sup_{v \in \mathbb{R}^L} \mathbb{E}\left(\|X^{P_1}\|_2^2 \mathbb{1}_{\{Z_v \neq 0\}}\right)^{\frac{1}{2}} + 1\right) > \delta$  for some  $\delta > 0$ . Then, as in the previous case,  $g(v)$  and hence the objective value of  $v$  tends to infinity as  $\kappa \rightarrow \infty$  while that of  $\tilde{v}$  stays constant.

Case  $p \notin \{1, 2\}$

When  $p \notin \{1, 2\}$ , the threshold of uncertainty  $\kappa_\epsilon$  for  $w^{*,v}$  to be inside  $\{w \in \mathbb{R}^N : \|w - w^{u,v}\|_q < \epsilon\}$  is not determined by a function of  $v \in \mathbb{R}^L$  or  $\epsilon$ , therefore, it is not possible to define a lower bound function such as  $g(v)$  for  $f^*(v)$ , the objective value of  $v \in \mathbb{R}^L$  (paired with  $w^{*,v}$ ) in the outer problem (16). The situation is otherwise similar to the case  $p = 2$ . Let us take arbitrary  $v \in \mathbb{R}^L$  such that  $\langle \mathbb{1}, v \rangle = 1 - s, s \neq 0$ , and again, note that for  $\tilde{v} \in \mathbb{R}^L$  with  $\langle \mathbb{1}, \tilde{v} \rangle = 1$ , the objective value is constant as the parameter  $\kappa$  increases. With similar calculations to the case  $p = 2$ , one can show that the objective value  $\mathcal{R}(\langle X^{P_1}, w^{*,v} \rangle + \langle X^{P_2}, v \rangle) + C\kappa\|w^{*,v}\|_q$  for  $v$  is larger than

$$f(v) - \epsilon C \left( \sup_{v \in \mathbb{R}^L} \mathbb{E} \left[ \|X^{P_1}\|_p^p \mathbb{1}_{\{Z_v \neq 0\}} \right]^{\frac{1}{p}} + \kappa \right) = \mathcal{R}\left(\frac{s}{N}\langle X^{P_1}, \mathbb{1} \rangle + \langle X^{P_2}, v \rangle\right) + C\kappa \frac{|s|}{N^{1-\frac{1}{q}}} - \epsilon C \left( \sup_{v \in \mathbb{R}^L} \mathbb{E} \left[ \|X^{P_1}\|_p^p \mathbb{1}_{\{Z_v \neq 0\}} \right]^{\frac{1}{p}} + \kappa \right),$$

given  $\kappa \geq \kappa_\epsilon$ . Picking  $\epsilon < \frac{|s|}{N^{1-\frac{1}{q}}}$ , the value of the above term increases once  $\kappa$  exceeds  $\kappa_\epsilon$  and tends to infinity with  $\kappa \rightarrow \infty$ . Thus, the arbitrary solution with total allocation  $s \neq 0$  turns suboptimal as  $\kappa \rightarrow \infty$ . □

### References

Behr, P., Guettler, A., & Miebs, F. (2013). On portfolio optimization: Imposing the right constraints. *Journal of Banking & Finance*, 37(4), 1232–1242.

Benartzi, S., & Thaler, R. H. (2001). Naive diversification strategies in defined contribution saving plans. *American Economic Review*, 91(1), 79–98.

Bertsimas, D., Brown, D. B., & Caramanis, C. (2011). Theory and applications of robust optimization. *SIAM Review*, 53(3), 464–501.

Bessler, W., Opfer, H., & Wolff, D. (2017). Multi-asset portfolio optimization and out-of-sample performance: An evaluation of Black–Litterman, mean-variance, and naïve diversification approaches. *The European Journal of Finance*, 23(1), 1–30.

Black, F., & Litterman, R. (1992). Global portfolio optimization. *Financial Analysts Journal*, 48(5), 28–43.

Brown, S. J., Hwang, I., & In, F. (2013). Why optimal diversification cannot outperform naive diversification: Evidence from tail risk exposure. Available at SSRN 2242694.

Calafiore, G. C. (2007). Ambiguous risk measures and optimal robust portfolios. *SIAM Journal on Optimization*, 18(3), 853–877.

Cesarone, F., Moretti, J., & Tardella, F. (2016). Optimally chosen small portfolios are better than large ones. *Economics Bulletin*, 36(4), 1876–1891.

- DeMiguel, V., Garlappi, L., Nogales, F. J., & Uppal, R. (2009a). A generalized approach to portfolio optimization: Improving performance by constraining portfolio norms. *Management Science*, 55(5), 798–812.
- DeMiguel, V., Garlappi, L., & Uppal, R. (2009b). Optimal versus naive diversification: How inefficient is the 1/N portfolio strategy? *Review of Financial Studies*, 22(5), 1915–1953.
- Erdogan, E., & Iyengar, G. (2006). Ambiguous chance constrained problems and robust optimization. *Mathematical Programming*, 107(1–2), 37–61.
- Fabozzi, F. J., Huang, D., & Zhou, G. (2010). Robust portfolios: Contributions from operations research and finance. *Annals of Operations Research*, 176(1), 191–220.
- Fabozzi, F. J., Kolm, P. N., Pachamanova, D., & Focardi, S. M. (2007). *Robust portfolio optimization and management*. New York: Wiley.
- Fletcher, J. (2011). Do optimal diversification strategies outperform the 1/n strategy in uk stock returns? *International Review of Financial Analysis*, 20(5), 375–385.
- Frahm, G., Wickern, T., & Wiechers, C. (2012). Multiple tests for the performance of different investment strategies. *ASTA Advances in Statistical Analysis*, 96(3), 343–383.
- Fugazza, C., Guidolin, M., & Nicodano, G. (2015). Equally Weighted vs. Long-Run Optimal Portfolios. *European Financial Management*, 21(4), 742–789.
- Gibbs, A. L., & Su, F. E. (2002). On choosing and bounding probability metrics. *International Statistical Review*, 70(3), 419–435.
- Guidolin, M., & Rinaldi, F. (2013). Ambiguity in asset pricing and portfolio choice: A review of the literature. *Theory and Decision*, 74(2), 183–217.
- Haley, M. R. (2016). Shortfall minimization and the naive (1/n) portfolio: An out-of-sample comparison. *Applied Economics Letters*, 23(13), 926–929.
- Jacobs, H., Müller, S., & Weber, M. (2014). How should individual investors diversify? An empirical evaluation of alternative asset allocation policies. *Journal of Financial Markets*, 19, 62–85.
- Kan, R., Wang, X., & Zhou, G. (2016). Optimal portfolio selection with and without risk-free asset. Available at SSRN.
- Kan, R., & Zhou, G. (2007). Optimal portfolio choice with parameter uncertainty. *Journal of Financial and Quantitative Analysis*, 42(03), 621–656.
- Kirby, C., & Ostdiek, B. (2012). It's all in the timing: Simple active portfolio strategies that outperform naive diversification. *Journal of Financial and Quantitative Analysis*, 47(02), 437–467.
- Murtazashvili, I., & Vozlyublennaya, N. (2013). Diversification strategies: Do limited data constrain investors? *Journal of Financial Research*, 36(2), 215–232.
- Owen, J., & Rabinovitch, R. (1983). On the class of elliptical distributions and their applications to the theory of portfolio choice. *The Journal of Finance*, 38(3), 745–752.
- Paç, A. B., & Pınar, M. Ç. (2014). Robust portfolio choice with cvar and var under distribution and mean return ambiguity. *Top*, 22(3), 875–891.
- Pachamanova, D. A. (2013). Robust portfolio selection. *Wiley Encyclopedia of Operations Research and Management Science*, 1–12.
- Pflug, G. C., Pichler, A., & Wozabal, D. (2012). The 1/n investment strategy is optimal under high model ambiguity. *Journal of Banking & Finance*, 36(2), 410–417.
- Pınar, M. Ç., & Paç, A. B. (2014). Mean semi-deviation from a target and robust portfolio choice under distribution and mean return ambiguity. *Journal of Computational and Applied Mathematics*, 259, 394–405.
- Romisch, W., & Pflug, G. C. (2007). *Modeling, measuring and managing risk*. Singapore: World Scientific.
- Rockafellar, R. T., & Uryasev, S. (2002). Conditional value-at-risk for general loss distributions. *Journal of Banking & Finance*, 26(7), 1443–1471.
- Tu, J., & Zhou, G. (2011). Markowitz meets talmud: A combination of sophisticated and naive diversification strategies. *Journal of Financial Economics*, 99(1), 204–215.
- Villani, C. (2003). *Topics in optimal transportation*. Providence: American Mathematical Soc.
- Villani, C. (2008). *Optimal transport: Old and new*. New York: Springer.