Minimum Rényi entropy portfolios

Nathan Lassance & Frédéric Vrins

Annals of Operations Research

ISSN 0254-5330
Volume 299
Combined 1-2

Ann Oper Res (2021) 299:23-46
DOI 10.1007/s10479-019-03364-2

Springer

Annals of Operations Research (2021) 299:23–46
https://doi.org/10.1007/s10479-019-03364-2

S.I.: RECENT DEVELOPMENTS IN FINANCIAL MODELING AND RISK MANAGEMENT

Minimum Rényi entropy portfolios

Nathan Lassance[1] · **Frédéric Vrins**[2]

Published online: 14 September 2019

Abstract
Accounting for the non-normality of asset returns remains one of the main challenges in portfolio optimization. In this paper, we tackle this problem by assessing the risk of the portfolio through the "amount of randomness" conveyed by its returns. We achieve this using an objective function that relies on the exponential of *Rényi entropy*, an information-theoretic criterion that precisely quantifies the uncertainty embedded in a distribution, accounting for higher-order moments. Compared to Shannon entropy, Rényi entropy features a parameter that can be tuned to play around the notion of uncertainty. A Gram–Charlier expansion shows that it controls the relative contributions of the central (variance) and tail (kurtosis) parts of the distribution in the measure. We further rely on a non-parametric estimator of the exponential Rényi entropy that extends a robust sample-spacings estimator initially designed for Shannon entropy. A portfolio-selection application illustrates that minimizing Rényi entropy yields portfolios that outperform state-of-the-art minimum-variance portfolios in terms of risk-return-turnover trade-off. We also show how Rényi entropy can be used in risk-parity strategies.

Keywords Portfolio selection · Shannon entropy · Rényi entropy · Risk measure · Information theory · Higher-order moments · Risk parity

The authors are grateful to Kris Boudt, Victor DeMiguel, Mikael Petitjean and two anonymous reviewers for their comments and suggestions. The authors also thank participants of the *Actuarial and Financial Mathematics 2018 Conference*, the *35th Annual Conference of the French Finance Association (AFFI)* and the *2018 Belgian Financial Research Forum* for their feedback. This work was supported by the Fonds de la Recherche Scientifique (F.R.S.-FNRS) [Grant Number FC 17775].

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s10479-019-03364-2) contains supplementary material, which is available to authorized users.

✉ Nathan Lassance
nathan.lassance@uclouvain.be

Frédéric Vrins
frederic.vrins@uclouvain.be

[1] Louvain Finance, UCLouvain (BE), 34 Voie du Roman Pays, 1348 Louvain-la-Neuve, Belgium

[2] Louvain Finance, Center for Operations Research and Econometrics, UCLouvain (BE), 34 Voie du Roman Pays, 1348 Louvain-la-Neuve, Belgium

1 Introduction

In portfolio optimization, it is well known that a high sensitivity of the optimal portfolio weights to estimation errors in parameter inputs can render otherwise sound investment strategies largely suboptimal out of sample; see Kolm et al. (2014) and references therein. This is in particular the case of the mean-variance portfolio of Markowitz (1952): the optimal weights are very sensitive to estimation errors in the asset mean returns. To tackle this robustness issue, one can simply disregard the portfolio-mean-return constraint, leading to the risk-based-allocation framework; see, for instance, Ardia et al. (2017). *Minimum-risk portfolios* have in particular attracted investors' attention because "the minimum-variance portfolio usually performs better out of sample than any other mean-variance portfolio—even when the Sharpe ratio or other performance measures related to *both* the mean and variance are used for the comparison" ((DeMiguel and Nogales 2009) p. 560).

Minimum-risk portfolios are commonly built using the variance as risk measure. As the sample minimum-variance portfolio is still quite vulnerable to estimation errors, various robust alternatives have been developed; see Fabozzi et al. (2010) and Scutellà and Recchia (2013) for reviews. Shrinkage estimation, introduced by Ledoit and Wolf (2003, 2004a, b), has proved particularly appealing. Still, the problem remains that the variance is an adequate risk measure only for Gaussian distributions and is largely unaffected by increasing tail concentration (Vermorken et al. 2012). As a result, the minimum-variance portfolio does not account for the non-normality of asset returns. Two main alternative approaches can be employed to deal with non-normality. First, one can minimize a downside-risk measure, such as the minimum-VaR and CVaR portfolios. However, such portfolios coincide with a mean-risk approach (Fabozzi et al. 2010), producing robustness issues as for the mean-variance portfolio. Second, one can extend the Taylor series expansion of the utility function to include the portfolio-return third and/or fourth moments; see Harvey et al. (2010), Martellini and Ziemann (2010) and Adcock (2014). However, robustifying such higher-order portfolios is very challenging due to increased dimensionality and large sensitivity to outliers.

In this paper, we propose a new (albeit natural) way of designing robust minimum-risk portfolios that account for the non-normality of asset returns. We do so by minimizing the portfolio-return uncertainty measured via the exponential of *Rényi entropy*, estimated within a robust m-spacings framework. The optimal portfolios so obtained are called *minimum Rényi entropy portfolios*. Entropy is a well-known concept coming from information theory. It precisely aims to quantify the uncertainty/amount of randomness conveyed by a distribution, embedding all higher-order moments (Cover and Thomas 2006). As a result, it is not surprising to notice that Shannon entropy (the most standard definition of entropy) has been recognized as an appealing measure in finance (Sbuelz and Trojani 2008; Zhou et al. 2013; Ormos and Zibriczky 2014) portfolio management (Philippatos and Wilson 1972; Dionisio et al. 2006; Vermorken et al. 2012; Flores et al. 2017 and utility theory (Yang and Qiu 2005; Abbas 2006; Jose et al. 2008). However, when using entropy to *construct* optimal portfolios, the literature is so far limited to employing entropy as a penalty term besides a more standard cost function: one considers the weights as discrete probabilities, and uses their entropy as a penalty term to shrink them toward the equally-weighted solution; see Bera and Park (2008) and Zhou et al. (2013). Instead, we use the entropy of the portfolio-return *distribution* (not that of portfolio weights) as the risk-based cost function. Searching for the portfolio weights minimizing the latter amounts to minimize the return uncertainty and thus provides the minimum-risk portfolio in the sense of information theory.

In addition, we rely on Rényi entropy, an extension of Shannon entropy. It features a parameter $\alpha \in [0, \infty]$ that allows one to trade off the minimization of central and tail uncertainty. We argue in particular in favor of setting $\alpha \in [0, 1]$ as, then, Rényi entropy has natural connections with measures of spread (as shown by the extended Chebyshev's inequality of Campbell 1966) and with a minimum-variance-kurtosis objective (as shown by a novel Gram–Charlier expansion of the measure). The empirical results support that choice as well.

Our contribution is organized as follows. Section 2 explores the theoretical properties of the exponential Rényi entropy and makes the link with the notion of risk. Section 3 follows with the minimum Rényi entropy portfolio and its connections with higher-order moments. Section 4 derives a robust m-spacings estimator of the measure and studies its consistency and robustness. We design and perform an empirical out-of-sample performance study of the proposed method in Sect. 5. Minimum Rényi entropy portfolios are shown to outperform standard minimum-risk portfolios in terms of risk-adjusted performance, while achieving a reasonable level of turnover for values of α close to zero. Section 6 concludes.

The proofs of the different propositions are relegated to the Online Resources.

2 Exponential Rényi entropy and risk measurement

We start this section by introducing the *Rényi entropy*, a flexible measure that quantifies the uncertainty of a random variable from its distribution. It encompasses the well-known *Shannon entropy*, which is recovered as a special case. We then show how its exponential transform is closely connected to *deviation risk measure*. A discussion of the impact of Rényi's α parameter closes the section, arguing in particular that setting $\alpha \in [0, 1]$ should be favored in our portfolio-selection context.

In the sequel, we denote F_X and f_X the cdf and pdf of a random variable X, respectively. In our context, X represents a random asset or portfolio return. We are exclusively interested in continuous distributions.

2.1 Shannon and Rényi entropy

The entropy of a random variable X commonly refers to its Shannon entropy, first introduced by Shannon (1948), giving birth to a new scientific discipline: *information theory*. It is defined as

$$H(X) := H[f_X] := -\mathbb{E}\big(\ln f_X(X)\big). \tag{1}$$

This measure is known to quantify the amount of randomness embedded in X. For instance, when X is a continuous random variable with bounded support, this quantity is maximized for the uniform distribution, which is the most uncertain one. Shannon entropy embeds many important properties. We refer to Cover and Thomas (2006) for an extended treatment.

Rényi (1961) proposed a generalization of Shannon entropy in (1) with the help of a parameter $\alpha \in \mathbb{R}^+$. The idea was to consider a generalized averaging of $-\ln f_X$, leading to the following definition:

$$H_\alpha(X) := H_\alpha[f_X] := \frac{1}{1-\alpha} \ln \mathbb{E}\big(f_X^{\alpha-1}(X)\big), \tag{2}$$

whenever the expectation exists. Shannon entropy is recovered as a special case in the sense that

$$\lim_{\alpha \to 1} H_\alpha(X) := H_1(X) = H(X). \tag{3}$$

Just like Shannon entropy, Rényi entropy enjoys interesting properties. However, its exponential transform has more natural properties in the context of risk. The next section is dedicated to a more detailed analysis of the exponential Rényi entropy and its connection with deviation risk measures.

2.2 Exponential Rényi entropy

We denote by $H_\alpha^{\exp}$ the exponential Rényi entropy, which, from (2), reads as

$$H_\alpha^{\exp}(X) := \exp(H_\alpha(X)) = \left(\int (f_X(x))^\alpha dx \right)^{\frac{1}{1-\alpha}}. \tag{4}$$

This quantity was first introduced by Campbell (1966) who studied its relevance as a measure of spread of a distribution for $\alpha \in [0, 1]$. We come back to the link between $H_\alpha^{\exp}$ and measures of spread in Sect. 2.4. In this paper, we apply this measure to the construction of minimum-risk portfolios; see Sect. 3.

2.2.1 Properties

From the properties of Rényi entropy (see Koski and Persson 1992; Johnson and Vignat 2007; Pham et al. 2008), $H_\alpha^{\exp}$ obeys the below properties.

Proposition 1 *Let c be a real constant. $H_\alpha^{\exp}(X)$ satisfies the following properties:*

(i) *Translation-invariance*

$$H_\alpha^{\exp}(X + c) = H_\alpha^{\exp}(X);$$

(ii) *Scaling property*

$$H_\alpha^{\exp}(cX) = |c| H_\alpha^{\exp}(X);$$

(iii) *It is non-increasing and continuous in α.*

2.2.2 Connection with deviation risk measures

When quantifying uncertainty, it is appealing to use the exponential Rényi entropy as a deviation risk measure, introduced by Rockafellar et al. (2006).

Definition 1 A *deviation risk measure* is any functional $\mathscr{D} : L^p(\Omega) \to [0, \infty]$ satisfying:[1]

(i) *Positivity* $\mathscr{D}(X) > 0$ for all non-constant X, and $\mathscr{D}(X) = 0$ for any constant X;
(ii) *Positive-homogeneity* $\mathscr{D}(cX) = c\mathscr{D}(X)\ \forall c > 0$;
(iii) *Translation-invariance* $\mathscr{D}(X + c) = \mathscr{D}(X)\ \forall c \in \mathbb{R}$;
(iv) *Sub-additivity* $\mathscr{D}(X + Y) \leqslant \mathscr{D}(X) + \mathscr{D}(Y)$.

[1] $L^p(\Omega)$ is the space of random variables defined on the support set Ω having finite pth moment.

Let us show that $H_\alpha^{\exp}$ fulfills the first three properties of deviation risk measures; sub-additivity is dealt with in next section.

Positive-homogeneity *(ii)* and translation-invariance *(iii)* result from the properties of $H_\alpha^{\exp}$ in Proposition 1. Regarding positivity *(i)*, $H_\alpha^{\exp}(X)$ is strictly positive if X is non-constant from the positivity of the density f_X. To see that it is null if X is constant, let us compute $H_\alpha^{\exp}(k)$ where k is a constant by computing the limit of $H_\alpha^{\exp}(k + cX)$ as c tends to zero for a given random variable X of finite entropy:

$$H_\alpha^{\exp}(k) = \lim_{c \downarrow 0} H_\alpha^{\exp}(k + cX) = \lim_{c \downarrow 0} c H_\alpha^{\exp}(X) = 0.$$

2.2.3 The sub-additivity property

In this section, we begin by underlining that, whereas $H_\alpha^{\exp}$ is expected to be sub-additive for most cases encountered in portfolio selection, it is not, *generally speaking*, sub-additive.

Proposition 2 *$H_\alpha^{\exp}$ is, generally speaking, not a sub-additive measure.*

Proof In the Online Resource 1, we give three analytical counter-examples to sub-additivity using pairs of random variables (X, Y) : a pair of independent one-sided (Lévy), a pair of independent two-sided bimodal (Gaussian mixtures) based on the entropy bounds derived in Vrins et al. (2007), and eventually a pair of comonotonic random variables. □

We stress that this proposition contradicts some statements recently made in the portfolio-selection literature; see Flores et al. (2017).[2]

It is however worth stressing that those counter-examples are atypical in portfolio-selection applications. The Lévy distribution for instance is extremely heavy-tailed: none of the moments are defined, and asset returns exhibit much lighter tails in practice (Cont 2001). Similarly, multimodal distributions and comonotonicity are behaviours that rarely arise in portfolio selection.

In fact, just like for the Value-at-Risk (Daníelsson et al. 2013), sub-additivity of the exponential Rényi entropy can be reasonably assumed to hold in the specific context of portfolio selection. For instance, the exponential Rényi entropy of $Z \sim \mathcal{N}(\mu, \sigma)$ collapses to $H_\alpha^{\exp}(Z) = \sigma\sqrt{2\pi\alpha^{1/(\alpha-1)}}$ (Koski and Persson 1992). From the stability of the Gaussian distribution, the sub-additivity property is equivalent to $\sigma_{X+Y} \leqslant \sigma_X + \sigma_Y$, which in turn holds from the sub-additivity of the standard deviation (Artzner et al. 1999).

However, this particular case is quite restrictive because asset returns are typically not well described by the Gaussian distribution. A more appealing candidate to model the fat tails observed in asset returns is the general class of *elliptical* distributions, which has many applications in mathematical finance and portfolio theory; see, for instance, Chen et al. (2011). The Gaussian distribution considered above is a special instance of the class of elliptical distributions, which also comprises more realistic distributions to model asset returns such as the Student t and Laplace distributions. As the next proposition shows, the exponential Rényi entropy is sub-additive when (X, Y) is distributed according to an elliptical distribution, providing a broader and more realistic sufficient condition for sub-additivity than the Gaussian setting.

[2] We are grateful to the authors of the aforementioned paper for discussions about the provided counter-examples.

Proposition 3 *Let* $(X, Y) \sim \mathrm{El}(\mu, \Sigma, g_2)$ *with* $\mathrm{El}(\boldsymbol{\mu}, \Sigma, g_2)$ *a bivariate elliptical distribution; that is,*

$$f_{X,Y}(x) = |\Sigma|^{-1/2} g_2\big((x-\mu)' \Sigma^{-1}(x-\mu)\big),$$

where g_2 *is a non-negative density generator function and* $|\Sigma|$ *is the absolute value of the determinant of* Σ*, the scaling matrix of* (X, Y)*. Then,* $H_\alpha^{\exp}$ *is sub-additive for the pair* (X, Y)*.*

Proof See Online Resource 2.1. □

Remark 1 Proposition 3 can be extended to any dimension, meaning that, if $X = (X_1, \ldots, X_n) \sim \mathrm{El}(\mu, \Sigma, g_n)$, then $H_\alpha^{\exp}\big(\sum_{i=1}^n X_i\big) \leqslant \sum_{i=1}^n H_\alpha^{\exp}(X_i)$. Combined with the positive-homogeneity property, this means that $H_\alpha^{\exp}$ is sub-additive at the portfolio level; that is, denoting $(w_1, \ldots, w_n)$ positive portfolio weights, we have $H_\alpha^{\exp}\big(\sum_{i=1}^n w_i X_i\big) \leqslant \sum_{i=1}^n w_i H_\alpha^{\exp}(X_i)$.

2.3 Exponential Rényi entropy as a flexible risk measure

This section explains how the parameter α allows to tune the relative contributions of the central and tail parts of the distribution, leading to different definitions of risk.

To show this, consider the two extreme cases $\alpha = 0$ and $\alpha = \infty$. As the next proposition shows, when $\alpha = 0$, $H_\alpha^{\exp}(X)$ measures the spread of X, while when $\alpha = \infty$, $H_\alpha^{\exp}(X)$ is given by the inverse of the supremum of f_X; see Koski and Persson (1992).

Proposition 4 *Let* X *be a continuous random variable, then* $H_0^{\exp}(X) := \lim_{\alpha \downarrow 0} H_\alpha^{\exp}(X)$ *and* $H_\infty^{\exp}(X) := \lim_{\alpha \to \infty} H_\alpha^{\exp}(X)$ *read as*

$$H_0^{\exp}(X) = \mathscr{L}(\Omega), \tag{5}$$

$$H_\infty^{\exp}(X) = 1/\sup f_X, \tag{6}$$

where $\mathscr{L}(\Omega)$ *is the Lebesgue measure of the support set of* X*,* $\Omega := \{x : f_X(x) > 0\}$*.*

As one can see, changing α modifies the way we measure entropy, that is uncertainty, and so the risk. Taking $\alpha = 0$ amounts to measure risk by the support of the distribution, while taking $\alpha = \infty$ amounts to measure risk by the maximal probability. By minimizing the portfolio-return entropy, as we propose in the next section, one can therefore minimize the density range on the x-axis with $\alpha = 0$ or maximize the density range on the y-axis with $\alpha = \infty$. $H_0^{\exp}$ focuses only on extreme values (low entropy = low distance between extreme values), while $H_\infty^{\exp}$ focuses only on the most likely outcomes (low entropy = high maximal probability), and so, in the symmetric unimodal case, on the center of the distribution.

From those two extreme cases, it results that, in portfolio-selection applications, taking α too large is not desirable because $H_\alpha^{\exp}$ will barely be affected by tail events, which is the criticism that is made about the variance. Conversely, by decreasing α, we assign more similar "weight" to all events, thus increasing the relative importance of tail events compared to events around the mode.

Example 1 To illustrate this effect of α, Fig. 1 shows how $H_\alpha^{\exp}(X)$, with $X \sim \mathrm{Student}(\nu)$, evolves with ν for different values of α. From Zografos and Nadarajah (2003), $H_\alpha^{\exp}(X)$ expresses as

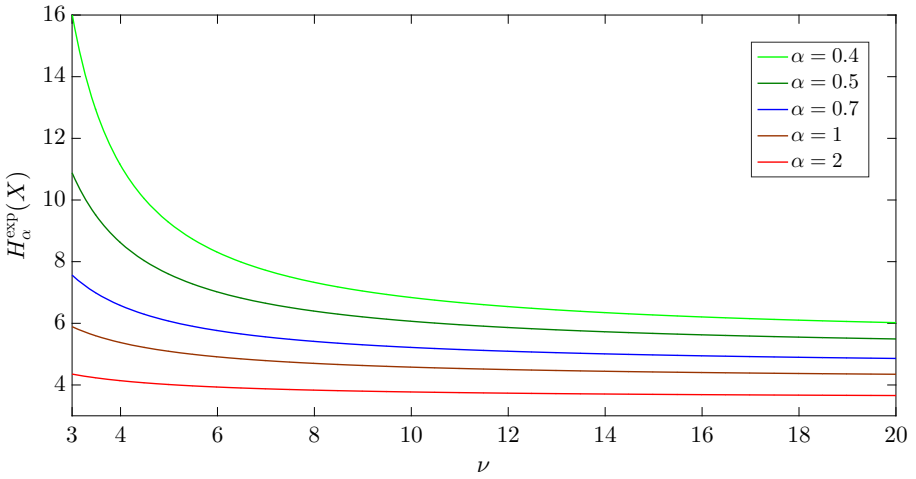

Fig. 1 The sensitivity of the exponential Rényi entropy $H_\alpha^{\exp}$ to tail uncertainty increases when its parameter α decreases. *Notes:* The figure depicts, for different values of α, the exponential Rényi entropy of the standardized t-Student distribution—$H_\alpha^{\exp}(X)$, $X \sim \text{Student}(\nu)$—as a function of the number of degrees of freedom ν. The analytical expression for $H_\alpha^{\exp}(X)$ is reported in Eq. (7)

$$H_\alpha^{\exp}(X) = (\pi\nu)^{\frac{1-\alpha}{2}} \left(\frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)} \right)^{\alpha} \frac{\Gamma\left(\frac{\alpha(\nu+1)}{2} - \frac{1}{2}\right)}{\Gamma\left(\frac{\alpha(\nu+1)}{2}\right)}. \tag{7}$$

As one can see, when going from $\alpha = 2$ to $\alpha = 0.4$, the sensitivity to the increase of tail uncertainty is indeed increasingly visible.

2.4 Appeal of the $\alpha \in [0, 1]$ case in portfolio selection

The previous section argued how decreasing the value of α allows one to obtain a measure of entropy that is increasingly more affected by tail events. In this section, we show more specifically that investors should favor setting $\alpha \in [0, 1]$, in which case $H_\alpha^{\exp}$ provides an appealing extension of the variance as a risk criterion.

First, for $\alpha \in [0, 1]$, $H_\alpha^{\exp}$ has close connections with measures of spread. By minimizing the variance, investors ensure that most of the probability distribution of the portfolio return is concentrated in some small interval around the mean. This is established by Chebyshev's inequality which, given the set $A_k = \{x \in \Omega \mid |x - \mathbb{E}(X)| \geq k\}$, says that

$$\mathbb{P}(X \in A_k) \leq \frac{\mathbb{V}\text{ar}(X)}{k^2}.$$

Similarly, for $\alpha \in [0, 1]$, a small value of $H_\alpha^{\exp}(X)$ entails that most of the probability distribution of X is concentrated on a set of small Lebesgue measure. This is determined by Campbell (1966)'s extended Chebyshev's inequality.

Proposition 5 *Let X be a continuous random variable whose Rényi entropy is defined for $\alpha \in [0, 1]$. Then, given $\alpha \in [0, 1]$ and $A'_k = \{x \in \Omega \mid f_X(x) \leq k\}$, the following inequality holds:*

$$\mathbb{P}(X \in A'_k) \leq \left(k H_\alpha^{\exp}(X)\right)^{1-\alpha}. \tag{8}$$

This inequality is more general than Chebyshev's inequality because it does not only deal with the absolute deviation around the mean, but instead relates the spread in terms of the size of the set on which most of the probability density is situated. For an unimodal random variable X with $\Omega = \mathbb{R}$ and $f_X(x) \to 0$ as $x \to \pm\infty$, which is common for asset returns, then there are only two values of x for which $f_X(x) = k$ (provided $k < \text{mode}(X)$). Denoting them x_k^- and x_k^+, we have $x_k^- < \text{mode}(X) < x_k^+$ and (8) means that

$$\mathbb{P}\left(x_k^- < X < x_k^+\right) \geq 1 - \left(k H_\alpha^{\exp}(X)\right)^{1-\alpha} \to 1 \text{ as } H_\alpha^{\exp}(X) \to 0.$$

In other words, if $H_\alpha^{\exp}(X)$ is small, the probability that X is concentrated on a small interval around its mode is close to one.

A second argument in favor of setting $\alpha \in [0, 1]$ is related to the Gram–Charlier expansion of Rényi entropy derived in Sect. 3.2. The expansion will show that, when $\alpha \in [0, 1]$, the coefficient in front of the kurtosis of X is positive (and instead negative for $\alpha > 1$) and so that a decrease in kurtosis decreases the Rényi entropy, as desired by investors.

Third, the empirical results presented in Sect. 5.2 depict a largely superior out-of-sample performance of the minimum Rényi entropy portfolio when $\alpha \in [0, 1]$ as well.

3 Minimum Rényi entropy portfolio

Given the good match between the theoretical properties of $H_\alpha^{\exp}$ and the desirable features of portfolio-selection criteria, we use this measure as an objective function to design investment strategies. In particular, we propose to construct a minimum-risk portfolio, called the *minimum Rényi entropy (MRE) portfolio*, that minimizes the exponential Rényi entropy of the portfolio return. We denote by P the portfolio return such that $P := w'X = \sum_{i=1}^n w_i X_i$ where $w = (w_1, \ldots, w_n)'$ is the vector of portfolio weights and $X = (X_1, \ldots, X_n)'$ is the random vector of asset returns.

3.1 Definition

The MRE portfolio over an investment set of n assets for a given α is defined as

$$w_\alpha^\star := \arg\min_{\boldsymbol{w} \in \mathscr{W}} H_\alpha^{\exp}(P), \tag{9}$$

where $\mathscr{W}$ is a set of constraints on $\boldsymbol{w}$, including the full investment constraint $\mathbf{1}_n' \boldsymbol{w} = 1$.

Note that, being affected by higher-order moments that are non-convex functions of the portfolio weights (Jurczenko and Maillet 2006), the optimization program in (9) may not necessarily be convex; that is, feature only one local optimum. Thus, when solving for the MRE portfolio, one must ideally resort to global-optimization techniques rather than standard local optimizers. We come back to this matter in Sect. 5.1.4.

3.2 Connection with moment-based portfolios: a Gram–Charlier expansion

Traditional minimum-risk portfolios are built from specific moments of the portfolio return, typically the variance, leading to the minimum-variance portfolio, and possibly higher-order moments as in Harvey et al. (2010), Martellini and Ziemann (2010), Adcock (2014) and Vanduffel and Yao (2017), that we call *higher-moment portfolios*.

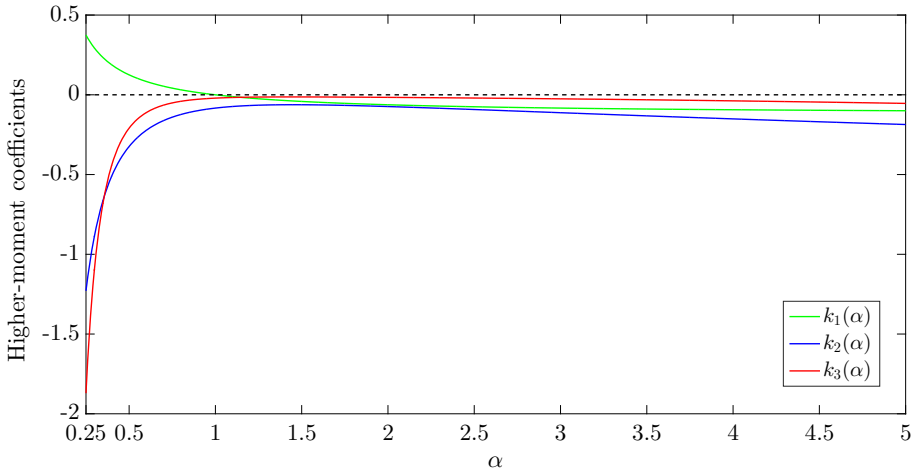

Fig. 2 Coefficients of the truncated Gram–Charlier expansion of Rényi entropy. *Notes:* The figure depicts the coefficients $k_1(\alpha)$, $k_2(\alpha)$ and $k_3(\alpha)$ in the truncated Gram–Charlier expansion of Rényi entropy as a function of α. The expression for the expansion and the coefficients are displayed in Eqs. (10) and (11)

In the classical Markowitz Gaussian setting, the MRE portfolio coincides with the minimum-variance portfolio because there is a one-to-one correspondence between $H_\alpha^{\exp}$ and the variance for Gaussian random variables.

In a more general setting however, the MRE portfolio is more attractive than the minimum-variance one because it accounts for the uncertainty coming from higher-order moments. To see this, we derive a truncated Gram–Charlier (GC) expansion of Rényi entropy.

Proposition 6 *Let $X \in L^4(\Omega)$ and note $\tilde{X} = (X - \mathbb{E}(X))/\sqrt{\mathbb{V}\mathrm{ar}(X)}$ its standardized copy. Define $\mathbb{S}\mathrm{kew}(X) = \mathbb{E}(\tilde{X}^3)$ and $\mathbb{K}\mathrm{urt}(X) = \mathbb{E}(\tilde{X}^4) - 3$. Then, the truncated GC expansion of $H_\alpha(X)$, denoted $H_\alpha^{GC}(X)$, is given by*

$$H_\alpha^{GC}(X) := H_\alpha\left[\mathcal{N}\left(0, \sqrt{\mathbb{V}\mathrm{ar}(X)}\right)\right] + k_1(\alpha)\mathbb{K}\mathrm{urt}(X) + k_2(\alpha)\mathbb{S}\mathrm{kew}(X)^2 + k_3(\alpha)\mathbb{K}\mathrm{urt}(X)^2, \tag{10}$$

with coefficients

$$\begin{aligned} k_1(\alpha) &= \frac{1-\alpha}{8\alpha}, \\ k_2(\alpha) &= -\frac{3\alpha^2 - 6\alpha + 5}{24\alpha^{3/2}}, \\ k_3(\alpha) &= -\frac{3\alpha^4 - 12\alpha^3 + 42\alpha^2 - 60\alpha + 35}{384\alpha^{5/2}}. \end{aligned} \tag{11}$$

Proof See Online Resource 2.2. □

The three coefficients $k_1(\alpha)$, $k_2(\alpha)$ and $k_3(\alpha)$ are depicted in Fig. 2.

Setting $\alpha = 1$, we get the GC expansion

$$H_1^{GC}(X) = H_1\left[\mathcal{N}\left(0, \sqrt{\mathbb{V}\mathrm{ar}(X)}\right)\right] - \frac{1}{12}\mathbb{S}\mathrm{kew}(X)^2 - \frac{1}{48}\mathbb{K}\mathrm{urt}(X)^2, \tag{12}$$

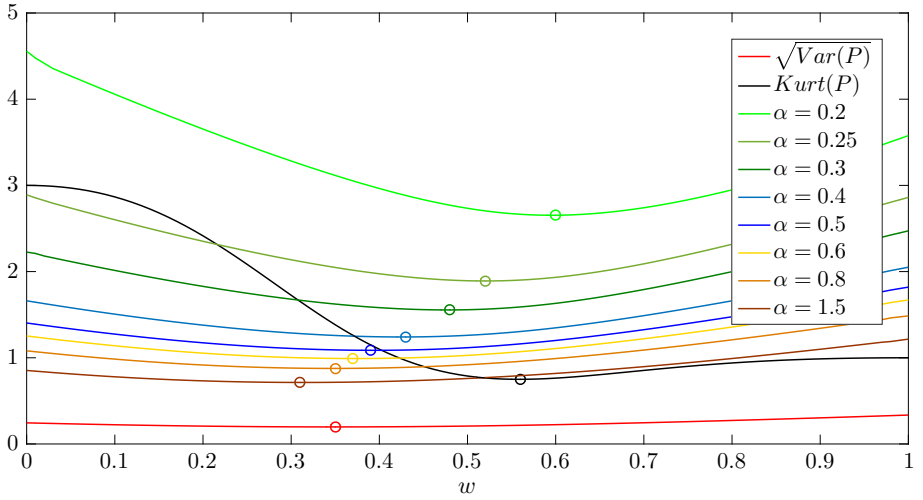

Fig. 3 The minimum Rényi entropy portfolio balances variance and kurtosis minimization. *Notes:* The figure depicts the standard deviation $\sqrt{\mathbb{V}\text{ar}(P)}$, excess kurtosis $\mathbb{K}\text{urt}(P)$ and exponential Rényi entropy $H_\alpha^{\exp}(P)$ of the portfolio return $P = wX + (1-w)Y$ as a function of $w \in [0, 1]$. The two asset returns $X \perp Y$ follow a zero-mean Student t distribution with $(\sigma_X, \nu_X) = (0.3, 10)$ and $(\sigma_Y, \nu_Y) = (0.2, 6)$

derived in Hyvärinen et al. (2001). Thus, the ability to control for kurtosis is a notable advantage of Rényi entropy over Shannon entropy because, in the latter case, $k_1(1) = 0$.

The connection between the MRE and higher-moment portfolios is now made explicit. We have $H_\alpha\big[\mathscr{N}\big(0, \sqrt{\mathbb{V}\text{ar}(P)}\big)\big] = H_\alpha\big[\mathscr{N}(0,1)\big] + \frac{1}{2}\ln \mathbb{V}\text{ar}(P)$, which yields[3]

$$w_\alpha^\star \approx \arg\min_{\boldsymbol{w}\in\mathscr{W}} \frac{1}{2}\ln \mathbb{V}\text{ar}(P) + k_1(\alpha)\mathbb{K}\text{urt}(P) + k_2(\alpha)\mathbb{S}\text{kew}(P)^2 + k_3(\alpha)\mathbb{K}\text{urt}(P)^2. \tag{13}$$

When f_P is close to a Gaussian, the main contributing higher-order term will be $k_1(\alpha)\mathbb{K}\text{urt}(P)$. When $\alpha < 1$, $k_1(\alpha) > 0$ and so the MRE portfolio is similar to a minimum-variance-kurtosis portfolio, which, as noted by Martellini and Ziemann (2010), is a well-performing higher-moment portfolio because estimators for even moments are less noisy than estimators for odd moments. When $\alpha > 1$ however, $k_1(\alpha) < 0$ and so the effect is reversed. In line with investors' preferences for kurtosis, setting $\alpha \in [0, 1]$ is thus more natural, as we noted in Sect. 2.4.

Thus, in line with Sect. 2.3, by playing with α one trades off the minimization of the central (variance) and tail (kurtosis) uncertainty; that is, of the first two even moments.

Example 2 Consider $n = 2$ assets $X_1 = X \perp X_2 = Y$ that follow a zero-mean Student t distribution with $(\sigma_X, \nu_X) = (0.3, 10)$ and $(\sigma_Y, \nu_Y) = (0.2, 6)$. We build a portfolio $P = wX + (1-w)Y$ and evaluate $H_\alpha^{\exp}(P)$ by numerical integration. On Fig. 3, we display how $H_\alpha^{\exp}(P)$, $\sqrt{\mathbb{V}\text{ar}(P)}$ and $\mathbb{K}\text{urt}(P)$ depend on w. As we can see, when α is high enough, $w^\star$ is close to the minimum-variance portfolio because $\sigma_X > \sigma_Y$ and that mostly central events matter when α is high. However, the more α decreases, the more important is the impact of the fatter tails of Y ($\nu_Y < \nu_X$) and so the more $w^\star$ approaches the minimum-kurtosis portfolio.

[3] Minimizing $H_\alpha^{\exp}(P)$ or $H_\alpha(P)$ is equivalent as $\exp(x)$ is a monotonically increasing function.

Given that $k_2(\alpha)$ and $k_3(\alpha)$ are negative for all α, the two additional terms $k_2(\alpha)\mathbb{S}\text{kew}(P)^2$ and $k_3(\alpha)\mathbb{K}\text{urt}(P)^2$ can be interpreted as driving the solution away from the Gaussian skewness and kurtosis. This is intuitive because, for a fixed mean and variance, the Shannon entropy is maximized by the Gaussian distribution (Cover and Thomas 2006).

Finally, note that the optimization program in (9) can accommodate an additional constraint on the portfolio mean return of the form $\mathbb{E}(P) = w'\mu \geq \mu_0$ with μ the asset mean-return vector, to account for the fact that investors do not only look at the risk, but also at the reward. In light of the GC expansion, such a framework would be linked to higher-moment efficient frontiers studied by Adcock (2014) and Qi et al. (2017) among others. In the empirical study, we focus on risk minimization due to the difficulties inherent in estimating the vector μ (see Sect. 1) that yield significant loss in out-of-sample performance.

4 Robust *m*-spacings estimation of exponential Rényi entropy

In this section, we explain how, given a finite portfolio-return sample $P_t = \sum_{i=1}^{n} w_i X_{i,t}$, $1 \leqslant t \leqslant T$, one can estimate $H_\alpha^{\exp}(P)$ in a robust way. In particular, we propose to use an estimator based on sample-spacings and discuss its properties in terms of consistency and robustness.

4.1 Motivation and expression for the *m*-spacings estimator

To avoid making assumptions about the portfolio-return distribution, we are looking for a non-parametric estimator of $H_\alpha^{\exp}(P)$. There exists substantial research on non-parametric estimation of Shannon entropy; see Beirlant et al. (1997) for a review.

A natural way of estimating entropy is the plug-in estimate where a density estimator is plugged into the integral defining the entropy. One could for example choose the well-known Parzen (kernel) estimator. However, this estimator is known to be very sensitive to the bandwidth parameter, which can yield issues of stability for our portfolio-selection context.[4] Instead, m-spacings estimation of entropy is more reliable: Wachowiak et al. (2005) show that such estimators "are robust and accurate, compare favorably to the popular Parzen window method for estimating entropies, and, in many cases, require fewer computations than Parzen methods." Therefore, we rely on a robust m-spacings estimator of Rényi entropy that extends m-spacings estimator of Shannon entropy of Learned-Miller and Fisher (2003), a "consistent, rapidly converging and computationally efficient estimator of entropy which is robust to outliers."

The Online Resource 3 gives a detailed derivation of the estimator, of which we only report the final expression here for conciseness.

Proposition 7 *Let X be a continuous random variable of which we dispose of T i.i.d observations. Then, the m-spacings estimator of $H_\alpha^{\exp}(X)$ is given by*

$$\widehat{H}_\alpha^{\exp}(m,T) := \left(\frac{1}{T-m}\sum_{i=1}^{T-m}\left(\frac{T+1}{m}\left(X^{(i+m:T)} - X^{(i:T)}\right)\right)^{1-\alpha}\right)^{\frac{1}{1-\alpha}}, \tag{14}$$

[4] When applied to our empirical data in Sect. 5, the Parzen estimator (with Gaussian kernel) achieves a worse risk-adjusted performance than the m-spacings estimator considered here for a wide range of values of the bandwidth parameter.

where $X^{(1:T)} \leqslant X^{(2:T)} \leqslant \cdots \leqslant X^{(T:T)}$ *are the order statistics (the observations sorted by increasing order) and* $m \in [1, T-1]$ *is an integer parameter.*

Proof See Online Resource 3. □

The parameter m is a free parameter of great importance: increasing its value reduces the variance of the estimator by grouping more order statistics in each spacing $X^{(i+m:T)} - X^{(i:T)}$. As a consequence, it plays a crucial role as it determines the robustness of the estimator, and in turn of the MRE portfolio. We come back to this in Sect. 4.2.2.

Taking the limit where $\alpha \to 1$, we recover the exponential of the estimator of Learned-Miller and Fisher (2003):

$$\widehat{H}_1^{\exp}(m, T) := \exp\left(\frac{1}{T-m}\sum_{i=1}^{T-m} \ln\left(\frac{T+1}{m}\left(X^{(i+m:T)} - X^{(i:T)}\right)\right)\right). \tag{15}$$

4.2 Properties of the *m*-spacings estimator

m-spacings estimation of entropy has attracted numerous research (Beirlant et al. 1997) and dates a while back; see, for instance, Vasicek (1976). However, it has been considered mainly for Shannon entropy and, even for this specific case, only asymptotic behaviour has been studied. In the general case $\alpha \neq 1$, consistency has not been established. In this section, we first discuss the estimator's properties in terms of consistency, arguing that the estimator's asymptotic bias can be ignored for the sake of our portfolio-selection application. Second, we show how the parameter m determines the estimator's robustness.

4.2.1 Asymptotic bias

Let us first consider the case $\alpha = 1$ and denote $\widehat{H}_1(m, T) := \ln \widehat{H}_1^{\exp}(m, T)$. van Es (1992) proved that $\widehat{H}_1(m, T)$ is asymptotically biased but, interestingly, that the bias only depends on the fixed value of m and not on the density f_X:

$$\widehat{H}_1(m, T) - H_1(X) \to \psi(m) - \ln m \quad \text{a.s.}, \tag{16}$$

where $\psi(x) = \frac{d}{dx}\ln\Gamma(x)$ is the digamma function. Equation (16) means that we can simply subtract the bias to get a consistent estimator. Getting back to the exponential case, this means that

$$me^{-\psi(m)}\widehat{H}_1^{\exp}(m, T) \to H_1^{\exp}(X) \quad \text{a.s.} \tag{17}$$

Interestingly, as readily seen from (17), because the asymptotic bias depends only on m when $\alpha = 1$, using the bias-corrected estimator in (17) or the biased estimator in (15) is equivalent when searching the portfolio weights that minimize the portfolio-return entropy in (9). Ideally, we would want the same result to hold for all α; that is, the asymptotic bias to depend only on α and m. While such a result is not known, Hegde et al. (2005) note that "in many practical applications, [...] this bias does not affect the solution, since it is independent of the true data distribution [...]." Further, as we now show, the estimator's bias for X and $\tilde{X} = (X - \mu)/\sigma$ is the same; it does not depend on the specific location and scale of X.

Proposition 8 *Let* $\tilde{X} = (X - \mu)/\sigma$ *and* $\widehat{H}_\alpha(X; m, T) := \ln \widehat{H}_\alpha^{\exp}(X; m, T)$. *Then, the m-spacings estimator bias* $\mathbb{B}\big(\widehat{H}_\alpha(X; m, T)\big) := \mathbb{E}\big(\widehat{H}_\alpha(X; m, T)\big) - H_\alpha(X) = \mathbb{B}\big(\widehat{H}_\alpha(\tilde{X}; m, T)\big)$.

Proof See Online Resource 2.3. □

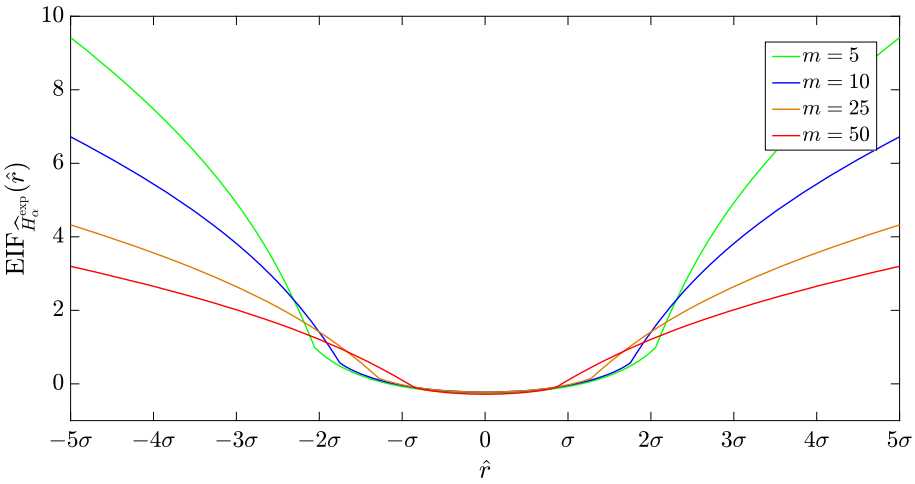

Fig. 4 Increasing m improves the robustness to outliers of the m-spacings estimator of the exponential Rényi entropy $\widehat{H}_\alpha^{\exp}(m, T)$. *Notes:* The figure depicts the empirical influence function (EIF) of the m-spacings estimator of the exponential Rényi entropy—$\mathrm{EIF}_{\widehat{H}_\alpha^{\exp}}(\hat{r})$—for $\alpha = 0.5$ and different values of m as a function of $\hat{r}$. We generate $T = 250$ observations from a Gaussian random variable $X \sim \mathcal{N}(\mu = 0, \sigma = 0.2)$ by setting $X_i = \mu + \sigma\Phi^{-1}\left(\frac{i}{T+1}\right)$ and we set $\hat{r} \in [-5\sigma, 5\sigma]$

4.2.2 Robustness to outliers

The parameter m acts as a smoothing parameter that controls the estimator's variance. This section shows that increasing m makes the m-spacings estimator more robust to outliers, which is crucial to ensure a solid out-of-sample performance of the MRE portfolio. Robustness conveys that a small perturbation from the true return distribution yields only a small change in the estimator's value.

In assessing the robustness of an estimator, the *Empirical Influence Function* (EIF) represents a useful tool; see Hampel et al. (1986). Given an estimator $\hat{\theta}(X_1, \ldots, X_T)$ of a quantity θ based on a sample of size T, $\mathrm{EIF}_{\hat{\theta}}(\hat{r})$ measures the sensitivity of the estimator $\hat{\theta}$ to the addition of a supplementary observation $\hat{r}$ in the sample:

$$\mathrm{EIF}_{\hat{\theta}}(\hat{r}) := (T+1)\big(\hat{\theta}(X_1, \ldots, X_T; \hat{r}) - \hat{\theta}(X_1, \ldots, X_T)\big). \tag{18}$$

Intuitively, the lower is $\mathrm{EIF}_{\hat{\theta}}(\hat{r})$, the more robust is the estimator $\hat{\theta}$. Figure 4 illustrates the EIF of the m-spacings estimator— $\mathrm{EIF}_{\widehat{H}_\alpha^{\exp}}(\hat{r})$—for $T = 250$ values from $X \sim \mathcal{N}(\mu = 0, \sigma = 0.2)$. Following Hampel et al. (1986), we set $X_i = \mu + \sigma\Phi^{-1}\left(\frac{i}{T+1}\right)$ to eliminate the random-sample variability. We consider $\hat{r} \in [-5\sigma, 5\sigma]$ and report the results for $\alpha = 0.5$ only because other values yield similar insights. One can indeed observe that the EIF decreases with m for large enough values of $\hat{r}$; that is, for outliers.

5 Out-of-sample empirical study

We finish the paper with an out-of-sample performance study of the MRE portfolio that aims at showing the practical interest of the proposed portfolio policy compared to several existing strategies. The study is performed on six datasets commonly used as benchmarks in

the portfolio-selection literature. Section 5.1 presents the methodology, Sect. 5.2 reports the results, and Sect. 5.3 considers an extension to the risk-parity strategy.

5.1 Methodology

We begin by detailing the methodology employed to generate the out-of-sample results reported in Sect. 5.2.

5.1.1 Strategies of comparison

The reported results compare the MRE portfolio with $\alpha \in \{0.3, 0.5, 0.7, 1, 1.5, 2\}$ to five different minimum-variance (MV) portfolios. The first four solve the quadratic optimization program

$$w^\star = \arg\min_{w \in \mathscr{W}} w' \Sigma w \tag{19}$$

by estimating Σ with the sample covariance matrix $\widehat{\Sigma}$ and the three robust shrinkage estimators developed by Ledoit and Wolf (2003, 2004a, b):

$$\begin{aligned} \widehat{\Sigma}_{CC} &:= \delta^\star \widehat{F}_{CC} + (1-\delta^\star)\widehat{\Sigma}, \quad \widehat{\Sigma}_{SF} := \delta^\star \widehat{F}_{SF} \\ &+(1-\delta^\star)\widehat{\Sigma}, \quad \widehat{\Sigma}_I := \delta^\star \widehat{F}_I + (1-\delta^\star)\widehat{\Sigma}, \end{aligned} \tag{20}$$

where $\delta^\star$ minimizes the Frobenius norm between the shrinkage estimator and the true matrix Σ. The three target matrices are based upon a constant correlation model ($\widehat{F}_{CC}$), a single-factor model ($\widehat{F}_{SF}$) and a scalar multiple of the identity matrix ($\widehat{F}_I$).

The fifth MV portfolio is the one-step M-portfolio (MP) of DeMiguel and Nogales (2009), given by

$$(w^\star, \mu^\star) = \arg\min_{w \in \mathscr{W},\, \mu} \frac{1}{T} \sum_{t=1}^{T} \rho(P_t - \mu), \tag{21}$$

where ρ is Huber robust loss function

$$\rho(x) := \begin{cases} x^2/2 & \text{if } |x| \leqslant c \\ c(|x| - c/2) & \text{if } |x| > c \end{cases}, \quad c = 1\%. \tag{22}$$

We note that we have also implemented the robust Bayes-Stein mean-variance portfolio of Jorion (1986) as well as the minimum-VaR portfolio using modified VaR as in Boudt et al. (2008). We do not report these results because, even though such criteria are positively affected by higher returns, they feature a lower risk-adjusted performance than the MRE portfolio due to their sensitivity to the portfolio mean return. The equally weighted portfolio has been considered as well but, while it naturally achieves the lowest turnover, it is largely outperformed by all the other strategies and so it is not reported either. It will be considered as a competitor when considering the risk-parity strategy in Sect. 5.3.

5.1.2 Datasets

We rely on six monthly returns datasets from the Kenneth French library that are extensively used as benchmarks in the literature to compare portfolio strategies; see, for instance, DeMiguel et al. (2009a, b), Behr et al. (2013) and Ardia et al. (2017). The datasets are listed in Table 1.

Table 1 List of datasets considered in the empirical study . *Source*: Kenneth French library

Datasets	Abb.	n	Time period
6 Fama-French port11folios of firms sorted by size and book-to-market	*6BTM*	6	07/1963–06/2016
25 Fama-French portfolios of firms sorted by size and book-to-market	*25BTM*	25	07/1963–06/2016
6 Fama-French portfolios of firms sorted by size and momentum	*6Mom*	6	07/1963–06/2016
25 Fama-French portfolios of firms sorted by size and momentum	*25Mom*	25	07/1963–06/2016
10 industry portfolios representing the US stock market	*10Ind*	10	07/1963–06/2016
17 industry portfolios representing the US stock market	*17Ind*	17	07/1963–06/201

All datasets are made of monthly returns. The value weighting scheme is considered for the industry portfolios

5.1.3 Dynamic rebalancing

We construct the portfolios by dynamic rebalancing. Rebalancing the weights not too frequently is important to ensure a satisfactory performance and turnover (Carroll et al. 2017), thus we set a rolling window of one year as in Behr et al. (2013). The estimation window is set to ten years ($T = 120$). We use as starting date 07/1963 as in DeMiguel et al. (2009b) and Behr et al. (2013), and rebalance the portfolios until 06/2016. This represents an out-of-sample period of 43 years.

5.1.4 Optimization

As pointed out in Sect. 3.1, the MRE optimization program is, generally speaking, not a convex problem. Therefore, to find the optimal weights, we rely on the global optimizer of Ugray et al. (2007) based on the Nelder-Mead algorithm. By doing so, we minimize the risk of getting stuck in a local minimum. We observe on the different datasets that recompiling the optimization several times yields essentially indistinguishable solutions, outlining that non-convexity is not a major issue.

5.1.5 Choice of *m*

As pointed out in Sect. 4.2.2, the value of m for the m-spacings estimator is of paramount importance as it determines the robustness of the MRE portfolio. To illustrate this, Fig. 5 reports, for the *10Ind* dataset, the time evolution of the MRE optimal weights for $\alpha = 0.5$ and $m = 2, 8, 24$. The portfolio weights are unconstrained; that is, $\mathscr{W} = \{\boldsymbol{w} \in \mathbb{R}^n \mid \mathbf{1}_n' \boldsymbol{w} = 1\}$. The weights of the M-portfolio are also reported for comparison. One can clearly observe that increasing m improves the stability of the optimal portfolio weights obtained.

As a first strategy, we have considered the *leave-one-out cross-validation* method, using as criteria maximum return, minimum variance and maximum Sharpe ratio. However, the results obtained were quite poor both in terms of performance and turnover. Allowing m to change at each rolling window seems to add instability to the procedure and is not recommended.

Therefore, as a second strategy, we have used the simple rule-of-thumb $m = \lfloor T^{2/3} \rfloor = 24$, which works well on the considered datasets. We observe actually that once m is high enough, the results display only a very minor sensitivity to the specific value of m that is chosen.

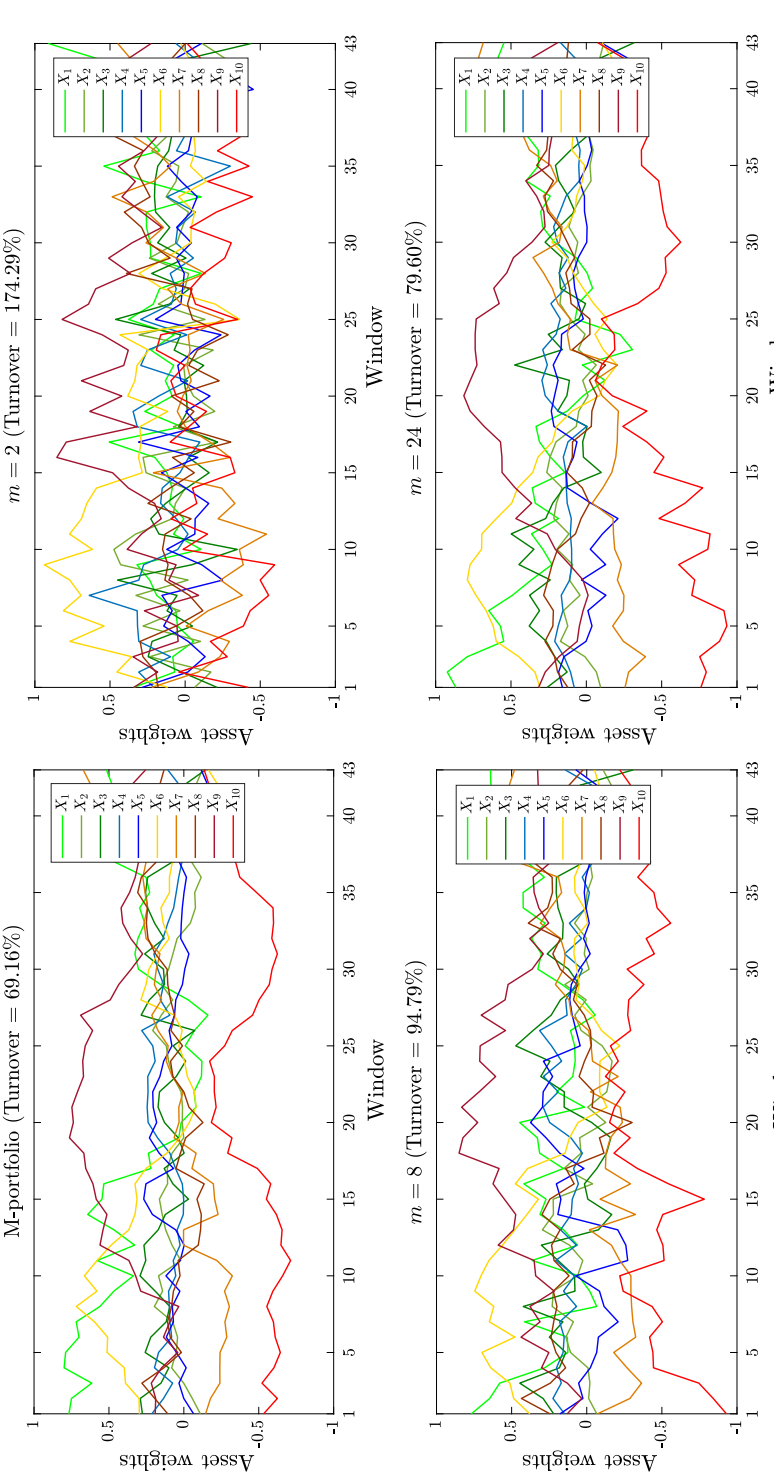

Fig. 5 Increasing m improves the stability of the MRE portfolio weights. *Notes:* The figure depicts, for the *10Ind* dataset, the time evolution of the portfolio weights for the M-portfolio of DeMiguel and Nogales (2009) and the MRE portfolio with $\alpha = 0.5$ and $m \in \{2, 8, 24\}$. The weights are unconstrained; that is, w is restricted to $\mathscr{W} = \{w \in \mathbb{R}^n \mid \mathbf{1}_n' w = 1\}$

Thus, one can set $m = 24$ without fearing that another different but close value would yield dissimilar results.[5]

5.1.6 Portfolio-weight constraints

To alleviate estimation errors, it is common in portfolio optimization to restrict the solution space $\mathscr{W}$. This has the effect of improving the stability of the portfolio weights obtained and in turn the out-of-sample performance. This is easily understood from Fig. 5, in which all the portfolios feature a significant turnover in the unconstrained case, even though $n = 10$ is relatively low in comparison to the sample size $T = 120$.

Therefore, we optimize the different portfolios subject to a constraint on the portfolio weights. We implement the global variance-based constraint devised by Levy and Levy (2014), which reads as

$$\sum_{i=1}^{n} \left(w_i - \frac{1}{n} \right)^2 \frac{\sigma_i}{\bar{\sigma}} \leq \delta, \tag{23}$$

where $\sigma_i := \sqrt{\mathbb{V}\text{ar}(X_i)}$ and $\bar{\sigma} := \frac{1}{n} \sum_{i=1}^{n} \sigma_i$. The underlying rationale is "to impose more stringent constraints on stocks with relatively high standard deviations, as the estimation errors for these stocks' parameters, and hence the potential economic loss, are larger than for stocks with relatively low standard deviations" (Levy and Levy 2014 p. 375). Using a U.S. industry portfolio dataset, the authors observe largely improved out-of-sample Sharpe ratios compared to several robust portfolio-selection strategies for a wide range of values of δ. In particular, the results are stable for δ between 10 and 25%. In the sequel, we set δ at the higher hand, $\delta = 25\%$, because too low values make it difficult to distinguish between the different portfolios as they are too close to the equally-weighted one, which corresponds to $\delta = 0$. The conclusions of the empirical study remain the same for $\delta = 20\%$ and $\delta = 15\%$, though naturally less strikingly.

5.1.7 Performance measures

We measure the out-of-sample performance and stability of the considered portfolio strategies with four criteria:

1. The Sharpe ratio, defined as

$$SR := (\mathbb{E}(P) - r_f)/\sqrt{\mathbb{V}\text{ar}(P)} \tag{24}$$

and estimated using sample estimators, which is the most common performance measure used in the asset allocation literature. For simplicity, we assume that $r_f = 0$ as in DeMiguel et al. (2009b); that is, we report the reciprocal of the coefficient of variation. However, given that the appeal of the MRE portfolio compared to the minimum-variance portfolio is to account for higher-order uncertainty, using the Sharpe ratio alone is not sufficient to assess the merit of our portfolio policy. Thus, the next two performance criteria also account for higher-order moments.

[5] Specifically, the results in Table 2 for $m = 18$ and $m = 35$ yield a very similar performance. The only changes are in terms of turnover, which is higher for $m = 18$ and nearly identical for $m = 35$.

2. The adjusted Sharpe ratio of Pézier (2004), defined as

$$ASR := SR\left(1 + \frac{\mathbb{S}\text{kew}(P)}{3!}SR - \frac{\mathbb{K}\text{urt}(P)}{4!}SR^2\right), \tag{25}$$

estimated using sample moment estimators.

3. The modified Value-at-Risk introduced by Favre and Galeano (2002), which approximates the Value-at-Risk via the first four moments via a second-order Cornish-Fisher expansion. It is defined as

$$\begin{aligned} MVaR := &-\mathbb{E}(P) - \sqrt{\mathbb{V}\text{ar}(P)}\Big[q_\varepsilon + \frac{1}{6}(q_\varepsilon^2 - 1)\mathbb{S}\text{kew}(P) + \frac{1}{24}(q_\varepsilon^3 - 3q_\varepsilon)\mathbb{K}\text{urt}(P) \\ &- \frac{1}{36}(2q_\varepsilon^3 - 5q_\varepsilon)\mathbb{S}\text{kew}(P)^2\Big], \end{aligned} \tag{26}$$

where q_ε is the quantile at confidence level ε of the standard Gaussian. We fix $\varepsilon = 10\%$. It is estimated using sample moment estimators.

4. To assess the stability and associated transaction costs of the portfolios, we report the turnover, defined as

$$\text{Turnover} := \frac{1}{R-1}\sum_{t=1}^{R-1}\sum_{i=1}^{n}|w_{i,t+1} - w_{i,t^+}|, \tag{27}$$

where $R = 43$ is the number of rebalancing periods, $w_{i,t+1}$ is the desired weight of asset i at time $t+1$ and w_{i,t^+} is its weight before rebalancing at $t+1$.

The criteria are expressed on an annual basis, except for the *MVaR* that is expressed on a monthly basis.

5.2 Out-of-sample results

The results are reported in Table 2. Several interesting observations can be made.

First, comparing the six MRE portfolios, one can clearly observe that $\alpha = 1.5$ and $\alpha = 2$ yield by far the worst performance. This is consistent with the Gram–Charlier expansion in (10), which shows that setting $\alpha > 1$ favors solutions with more kurtosis because as $k_1(\alpha) < 0$ when $\alpha > 1$. Therefore, both from a theoretical and empirical perspective, it is recommended to set $\alpha \leqslant 1$, as argued in Sect. 2.4.

Second, the turnover of the MRE portfolio systematically increases with α. It does not increase too much from $\alpha = 0.3$ to $\alpha = 0.7$ but then quickly increases dramatically. This effect can be explained by the fact that the convexity of $H_\alpha^{\exp}(P)$ (as a function of $\boldsymbol{w}$) decreases as α increases. This is for example visible in Fig. 3. This makes that the minimum is more bound to change largely from one rolling window to another because the objective function is flatter around the minimum the higher the value of α.

Third, it is appealing to observe that, for $\alpha \in \{0.3, 0.5, 0.7\}$, the performance of the MRE portfolio is very stable with respect to α. This is easily observed by looking at the average SR, ASR and $MVaR$ across the six datasets. This parameter robustness is an appealing behaviour for the decision-maker. For $\alpha = 1$, the SR and MSR remain very similar, but the $MVaR$ is quite substantially worse. Combined with the fact that the turnover increases with α, this means that choosing α quite low, in this case around $\alpha = 0.3$, yields the best trade-off between risk, return and turnover.[6]

[6] For completeness, we have checked the results for $\alpha \in \{0.05, 0.1, 0.2\}$ as well. The turnover barely decreases compared to $\alpha = 0.3$, and the performance measures remain nearly identical.

Table 2 Out-of-sample Sharpe ratios, adjusted Sharpe ratios, modified Value-at-Risk and turnover of the MRE and MV portfolios for the six considered datasets

	MRE portfolios						*MV portfolios*				
	$\alpha = 0.3$	$\alpha = 0.5$	$\alpha = 0.7$	$\alpha = 1$	$\alpha = 1.5$	$\alpha = 2$	$\widehat{\Sigma}$	$\widehat{\Sigma}_{CC}$	$\widehat{\Sigma}_{SF}$	$\widehat{\Sigma}_{I}$	MP
Sharpe ratio (SR)											
6BTM	0.844	0.845	0.846	0.841	0.832	0.815	0.835	0.821	0.833	0.836	0.841
25BTM	0.985	0.980	0.973	0.961	0.937	0.906	0.953	0.913	0.951	0.957	0.956
6Mom	0.768	0.763	0.753	0.750	0.716	0.700	0.738	0.741	0.738	0.737	0.731
25Mom	0.936	0.948	0.957	0.960	0.954	0.928	0.902	0.920	0.904	0.903	0.916
10Ind	0.995	1.001	1.014	1.013	0.971	0.936	0.977	0.970	0.983	0.973	0.976
17Ind	0.938	0.947	0.936	0.965	0.905	0.891	0.936	0.924	0.935	0.935	0.918
Average	0.911	0.914	0.913	0.915	0.886	0.863	0.890	0.882	0.891	0.890	0.890
Adjusted Sharpe ratio (ASR)											
6BTM	0.829	0.830	0.831	0.826	0.816	0.800	0.822	0.809	0.821	0.824	0.826
25BTM	0.969	0.963	0.956	0.943	0.919	0.889	0.940	0.906	0.939	0.944	0.940
6Mom	0.759	0.753	0.744	0.741	0.708	0.694	0.730	0.733	0.729	0.729	0.723
25Mom	0.921	0.934	0.943	0.947	0.943	0.918	0.889	0.909	0.892	0.890	0.902
10Ind	0.990	0.995	1.005	1.004	0.965	0.927	0.973	0.966	0.979	0.968	0.972
17Ind	0.950	0.959	0.945	0.975	0.911	0.901	0.950	0.940	0.950	0.949	0.929
Average	0.903	0.906	0.904	0.906	0.877	0.855	0.884	0.877	0.885	0.884	0.882
Modified Value-at-Risk (MVaR)											
6BTM	3.71%	3.71%	3.75%	3.78%	3.91%	3.90%	3.73%	3.75%	3.73%	3.73%	3.74%
25BTM	3.12%	3.13%	3.17%	3.33%	3.49%	3.52%	3.23%	3.70%	3.19%	3.22%	3.17%
6Mom	3.97%	4.00%	4.07%	4.10%	4.09%	4.03%	3.94%	4.01%	3.95%	3.94%	4.11%
25Mom	3.37%	3.33%	3.23%	3.16%	3.35%	3.61%	3.47%	3.51%	3.43%	3.47%	3.38%
10Ind	3.16%	3.19%	3.09%	3.16%	3.28%	3.38%	3.18%	3.34%	3.19%	3.21%	3.10%
17Ind	2.98%	2.95%	3.12%	3.34%	3.25%	3.26%	2.99%	3.15%	3.04%	3.11%	3.10%
Average	3.39%	3.39%	3.41%	3.48%	3.56%	3.62%	3.42%	3.58%	3.42%	3.45%	3.43%
Turnover											
6BTM	0.184	0.182	0.183	0.189	0.218	0.266	0.171	0.167	0.170	0.168	0.168
25BTM	0.499	0.535	0.616	0.966	1.356	1.502	0.448	0.464	0.443	0.445	0.478
6Mom	0.167	0.171	0.180	0.191	0.247	0.284	0.168	0.168	0.167	0.168	0.178
25Mom	0.437	0.444	0.523	0.635	0.860	1.200	0.417	0.413	0.410	0.416	0.422
10Ind	0.348	0.350	0.378	0.450	0.600	0.776	0.283	0.276	0.274	0.286	0.321
17Ind	0.526	0.568	0.685	0.825	1.033	1.246	0.449	0.422	0.433	0.452	0.467
Average	0.360	0.375	0.427	0.542	0.719	0.879	0.323	0.318	0.316	0.322	0.339

The portfolios are constructed with the methodology presented in Sect. 5.1

Fourth, let us compare the MRE and MV portfolios, focusing only on $\alpha \in \{0.3, 0.5, 0.7\}$ following the above observations. One can definitely observe that the MRE portfolios improve upon the MV portfolios in terms of SR, ASR and $MVaR$. Indeed, averaging across the six datasets, each MRE portfolio displays a better SR, ASR and $MVaR$ than *all* the MV portfolios. In terms of turnover however, all the MRE portfolios display less stability than the MV portfolios. This is to be expected because the MRE portfolio is sensitive to the higher-order moments, which are more affected by outliers than the variance. That said, for $\alpha = 0.3$

and $\alpha = 0.5$, the increase in turnover remains quite modest (around four percentage points on average for $\alpha = 0.3$).

Therefore, we conclude that Rényi entropy provides a superior risk criterion than the variance in an asset-allocation context, especially for low values of α, specifically around $\alpha = 0.3$ for the datasets considered here.

5.3 Extension to the risk-parity strategy

As a final test of the appeal of Rényi entropy in portfolio selection, we apply it to the risk-parity strategy, which is an alternative risk-based strategy to the minimum-risk strategy considered in this paper. The risk-parity portfolio has been rigorously studied for the first time by Maillard et al. (2010), followed by many other studies. The underlying rationale of risk parity is to find the portfolio for which all the assets in the portfolio contribute equally to the portfolio-return risk.

To implement this portfolio, we need to compute the asset risk contributions to the total portfolio-return risk. For positive-homogeneous risk measures, it can be computed with the Euler risk contribution, which defines the risk contribution of the ith asset as w_i times the partial derivative of the portfolio-return risk with respect to w_i. However, without distributional assumptions, this partial derivative can not be computed analytically for the exponential Rényi entropy. Thus, we rely on an alternative more tractable approach proposed by Tasche (2008) called the *marginal risk contribution*. It is defined by starting from the following risk contribution:

$$H_\alpha^{\exp}(X_i|P) := H_\alpha^{\exp}(P) - H_\alpha^{\exp}(P - w_i X_i). \tag{28}$$

As shown by Tasche (2008), these risk contributions add up to less than the total portfolio-return risk for continuously differentiable, sub-additive and positive homogeneous risk measures. To circumvent this problem, Tasche (2008) proposes to normalize them, leading to the marginal contribution of the ith asset to $H_\alpha^{\exp}(P)$ defined as

$$MH_\alpha^{\exp}(X_i|P) := \frac{H_\alpha^{\exp}(X_i|P)}{\sum_{j=1}^n H_\alpha^{\exp}(X_j|P)}, \tag{29}$$

which fulfills

$$\sum_{i=1}^n MH_\alpha^{\exp}(X_i|P) = 1. \tag{30}$$

Then, the *Rényi entropy parity* (REP) portfolio is defined as the solution to

$$\min_{w\in\mathscr{W}} \sum_{i=1}^n \left(MH_\alpha^{\exp}(X_i|P) - 1/n\right)^2. \tag{31}$$

We report in Table 3 the out-of-sample performance of the REP portfolio with the same methodology than in Sect. 5.1. We have tested for values of $\alpha \in \{0.3, 0.5, 0.7, 1, 1.5, 2\}$, and observe that the performance criteria remain very stable with the choice of α but, as for the MRE portfolio, that the turnover systematically increases with α. Thus, for the sake of conciseness, we report the results for $\alpha = 0.3$ only. As shown by Maillard et al. (2010), the risk-parity portfolio achieves a trade-off between the equally weighted (EW) portfolio and the minimum-risk portfolio in terms of risk. Thus, in Table 3, we also compare the performance of the REP portfolio with that of the EW portfolio.

Table 3 Out-of-sample Sharpe ratios, adjusted Sharpe ratios, modified Value-at-Risk and turnover of the REP and EW portfolios for the six considered datasets

	6BTM	*25BTM*	*6Mom*	*25Mom*	*10Ind*	*17Ind*	*Average*
Sharpe ratio (SR)							
REP portfolio ($\alpha = 0.3$)	0.725	0.730	0.657	0.682	0.799	0.851	0.741
EW portfolio	0.708	0.708	0.637	0.654	0.749	0.803	0.710
Adjusted Sharpe ratio (ASR)							
REP portfolio ($\alpha = 0.3$)	0.713	0.717	0.649	0.674	0.788	0.848	0.731
EW portfolio	0.696	0.696	0.631	0.648	0.739	0.800	0.702
Modified Value-at-Risk (MVaR)							
REP portfolio ($\alpha = 0.3$)	4.26%	4.45%	4.61%	4.75%	3.57%	3.76%	4.23%
EW portfolio	4.43%	4.61%	4.77%	4.97%	3.89%	4.00%	4.45%
Turnover							
REP portfolio ($\alpha = 0.3$)	5.88%	6.89%	7.03%	7.58%	8.86%	10.33%	7.76%
EW portfolio	6.02%	6.67%	6.99%	7.63%	8.56%	9.68%	7.59%

The portfolios are constructed with the methodology presented in Sects. 5.1 and 5.3

We make two main conclusions. First, the REP portfolio clearly outperforms the EW portfolio on all three performance criteria, while being subject to a turnover that is only slightly larger (7.76% versus 7.59% on average over the six datasets). Second, comparing the results in Tables 2 and 3 , one can observe that the appeal of the REP portfolio compared to the MRE strategy is not in terms of performance: the REP portfolio is outperformed on all three performance criteria. Instead, the appeal of the REP portfolio is in terms of turnover, which is very substantially reduced: for $\alpha = 0.3$, the average turnover over the six datasets is 36.02% for the MRE portfolio, versus 7.76% for the REP portfolio. The greater stability of the REP portfolio strategy makes that it may be preferred by investors facing non-negligible transaction costs.

6 Conclusion

Many studies from the wide scientific literature suggest that minimum-risk portfolios exhibit solid out-of-sample performances in spite of the fact that there is no portfolio-mean-return constraint. Whereas variance—initially introduced by Markowitz—is a natural risk measure in a Gaussian framework, it fails to capture extreme events that arguably arise in real applications. In order to take this reality into account, various alternative risk measures have been put forward.

In this paper, we have proposed a natural uncertainty measure—the exponential Rényi entropy—as a higher-moment criterion for portfolio selection. Rényi entropy generalizes Shannon entropy, yielding a set of uncertainty measures. Its parameter α enables to tune the relative contributions of the central and tail parts of the distribution in the measure. Its exponential transform fulfills desirable properties as it is closely related to the class of deviation risk measures, as well as to measures of spread for $\alpha \in [0, 1]$.

Minimizing this measure yields the minimum Rényi entropy portfolio. A truncated Gram–Charlier expansion shows that this portfolio represents a higher-moment extension of the minimum-variance portfolio, with α controlling the trade-off between variance and kurtosis minimization.

In practical settings, the empirical study has demonstrated that the minimum Rényi entropy portfolio fares better out of sample compared to state-of-the-art robust minimum-variance portfolios in terms of trading off risk, return and turnover, especially for values of α close to zero. We have also exemplified how the exponential Rényi entropy can be used for other portfolio-selection strategies such as the popular risk-parity approach.

In this paper, we have concentrated on the single-period portfolio-selection problem. One important direction for future research is to extend the minimum Rényi entropy portfolio to the multiperiod setting. To that end, using the mathematical formulation proposed in Li and Ng (2000) one can consider an investor facing a finite number of future time periods during which he can invest, and who looks for the investment in the risky assets at each of the future time periods so as to minimize the exponential Rényi entropy of his terminal portfolio wealth. This portfolio would extend the multiperiod minimum-variance portfolio studied by Li and Ng (2000) and Pun (2018), among others. Without specific distributional assumptions, this problem is analytically intractable, just like for the single-period case considered in this paper. However, using our m-spacings estimator, it can be readily be solved using standard stochastic programming models, such as backward induction.

Beyond our application, this paper demonstrates the potential appeal of Rényi entropy in various operations-research problems as a powerful way of capturing higher-moment uncertainty, and of using entropy as an optimization criterion rather than simply an ad hoc evaluation measure. In the particular case of portfolio selection, Rényi entropy has been shown to be a powerful alternative to the variance as risk criterion, opening the door to other applications in higher-moment portfolio selection.

References

Abbas, A. (2006). Maximum entropy utility. *Operations Research*, *54*(2), 277–290.

Adcock, C. (2014). Mean-variance-skewness efficient surfaces, Stein's lemma and the multivariate extended skew-student distribution. *European Journal of Operational Research*, *234*(2), 392–401.

Ardia, D., Bolliger, G., Boudt, K., & Gagnon-Fleury, J. (2017). The impact of covariance misspecification in risk-based portfolios. *Annals of Operations Research*, *254*(1–2), 1–16.

Artzner, P., Delbaen, F., Eber, J., & Heath, D. (1999). Coherent measures of risk. *Mathematical Finance*, *9*(3), 203–228.

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

Beirlant, J., Dudewicz, E., Gyofi, L., & van der Meulen, E. (1997). Non-parametric entropy estimation: An overview. *International Journal of Mathematical and Statistical Sciences*, *6*(1), 17–39.

Bera, A., & Park, S. (2008). Optimal portfolio diversification using the maximum entropy principle. *Econometric Reviews*, *27*(4–6), 484–512.

Boudt, K., Peterson, B., & Croux, C. (2008). Estimation and decomposition of downside risk for portfolios with non-normal returns. *Journal of Risk*, *11*, 79–103.

Campbell, L. (1966). Exponential entropy as a measure of extent of a distribution. *Z Wahrsch*, *5*, 217–225.

Carroll, R., Conlon, T., Cotter, J., & Salvador, E. (2017). Asset allocation with correlation: A composite trade-off. *European Journal of Operational Research*, *262*(3), 1164–1180.

Chen, L., He, S., & Zhang, S. (2011). When all risk-adjusted performance measures are the same: In praise of the Sharpe ratio. *Quantitative Finance*, *11*(10), 1439–1447.

Cont, R. (2001). Empirical properties of asset returns: Stylized facts and statistical issues. *Quantitative Finance*, *1*, 223–236.

Cover, T., & Thomas, J. (2006). *Elements of information theory* (2nd ed.). New Jersey: Wiley.

Daníelsson, J., Jorgensenb, B., Samorodnitskyc, G., Sarma, M., & de Vries, C. (2013). Fat tails, VaR and subadditivity. *Journal of Econometrics*, *172*, 283–291.

DeMiguel, V., Garlappi, L., Nogales, F., & 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? *The Review of Financial Studies*, *22*(5), 1915–1953.
DeMiguel, V., & Nogales, F. (2009). Portfolio selection with robust estimation. *Operations Research*, *57*, 560–577.
Dionisio, A., Menezes, R., & Mendes, A. (2006). An econophysics approach to analyse uncertainty in financial markets: An application to the Portuguese stock market. *The European Physical Journal B*, *50*, 161–164.
Fabozzi, F., Huang, D., & Zhou, G. (2010). Robust portfolios: Contributions from operations research and finance. *Annals of Operations Research*, *176*(1), 191–220.
Favre, L., & Galeano, J. (2002). Mean-modified value-at-risk optimization with hedge funds. *Journal of Alternative Investments*, *5*(2), 21–25.
Flores, Y., Bianchi, R., Drew, M., & Trück, S. (2017). The diversification delta: A different perspective. *Journal of Portfolio Management*, *43*(4), 112–124.
Hampel, F., Ronchetti, E., Rousseeuw, P., & Stahel, W. (1986). *Robust statistics: The approach based on influence functions*. New York: Wiley.
Harvey, C. R., Liechty, J. C., Liechty, M. W., & Müller, P. (2010). Portfolio selection with higher moments. *Quantitative Finance*, *10*(5), 469–485.
Hegde, A., Lan, T., & Erdogmus, D. (2005). Order statistics based estimator for Rényi entropy. In *IEEE workshop on machine learning for signal processing* (pp. 335–339).
Hyvärinen, A., Karhunen, J., & Oja, E. (2001). *Independent component analysis*. New York: Wiley.
Johnson, O., & Vignat, C. (2007). Some results concerning maximum Rényi entropy distributions. *Annales de l'Institut Henri-Poincaré (B) Probab. Statist*, *43*(3), 339–351.
Jorion, P. (1986). Bayes–Stein estimation for portfolio analysis. *Journal of Financial and Quantitative Analysis*, *21*(3), 279–292.
Jose, V., Nau, R., & Winkler, R. (2008). Scoring rules, generalized entropy, and utility maximization. *Operations Research*, *56*(5), 1146–1157.
Jurczenko, E., & Maillet, B. (2006). *Multi-moment asset allocation and pricing models*. West Sussex: Wiley.
Kolm, P., Tütüncü, R., & Fabozzi, F. (2014). 60 years of portfolio optimization: Practical challenges and current trends. *European Journal of Operational Research*, *234*(2), 356–371.
Koski, T., & Persson, L. (1992). Some properties of generalized exponential entropies with application to data compression. *Information Sciences*, *62*, 103–132.
Learned-Miller, E., & Fisher, J. (2003). ICA using spacings estimates of entropy. *Journal of Machine Learning Research*, *4*, 1271–1295.
Ledoit, O., & Wolf, M. (2003). Improved estimation of the covariance matrix of stock returns with an application to portfolio selection. *Journal of Empirical Finance*, *10*, 603–621.
Ledoit, O., & Wolf, M. (2004a). A well-conditioned estimator for large-dimensional covariance matrices. *Journal of Multivariate Analysis*, *88*, 365–411.
Ledoit, O., & Wolf, M. (2004b). Honey, I shrunk the sample covariance matrix. *Journal of Portfolio Management*, *30*, 110–119.
Levy, H., & Levy, M. (2014). The benefits of differential variance-based constraints in portfolio optimization. *European Journal of Operational Research*, *234*, 372–381.
Li, D., & Ng, W. L. (2000). Optimal dynamic portfolio selection: Multiperiod mean-variance formulation. *Mathematical Finance*, *10*(3), 387–406.
Maillard, S., Roncalli, T., & Teiletche, J. (2010). On the properties of equally-weighted risk contributions portfolios. *Journal of Portfolio Management*, *36*(4), 60–70.
Markowitz, H. (1952). Portfolio selection. *The Journal of Finance*, *7*(1), 77–91.
Martellini, L., & Ziemann, V. (2010). Improved estimates of higher-order comoments and implications for portfolio selection. *Review of Financial Studies*, *23*(4), 1467–1502.
Ormos, M., & Zibriczky, D. (2014). Entropy-based financial asset pricing. *Plos One*, *9*(12), e115742.
Pézier, J. (2004). Risk and risk aversion. In C. Alexander & E. Sheedy (Eds.), *The professional risk managers' handbook*. Wilmington: PRMIA Publications.
Pham, D., Vrins, F., & Verleysen, M. (2008). On the risk of using Rényi's entropy for blind source separation. *IEEE Transactions on Signal Processing*, *56*(10), 4611–4620.
Philippatos, G., & Wilson, C. (1972). Entropy, market risk, and the selection of efficient portfolios. *Applied Economics*, *4*(3), 209–220.
Pun, C. S. (2018). Time-consistent mean-variance portfolio selection with only risky assets. *Economic Modelling*, *75*, 281–292.
Qi, Y., Steuer, R., & Wimmer, M. (2017). An analytical derivation of the efficient surface in portfolio selection with three criteria. *Annals of Operations Research*, *251*(1–2), 161–177.
Rényi, A. (1961). On measures of entropy and information. In *Fourth Berkeley symposium on mathematical statistics and probability* (pp. 547–561).

Rockafellar, R., Uryasev, S., & Zabarankin, M. (2006). Generalized deviations in risk analysis. *Finance and Stochastics*, *10*, 51–74.

Sbuelz, A., & Trojani, F. (2008). Asset prices with locally constrained-entropy recursive multiple-priors utility. *Journal of Economic Dynamics and Control*, *32*(11), 3695–3717.

Scutellà, M., & Recchia, R. (2013). Robust portfolio asset allocation and risk measures. *Annals of Operations Research*, *204*(1), 145–169.

Shannon, C. (1948). A mathematical theory of communication. *Bell Systems Technical Journal*, *27*, 379–423.

Tasche, D. (2008). Capital allocation to business units and sub-portfolios: The Euler principle. Pillar II in the new basel accord: The challenge of economic capital. Risk Books. In *Resti, A* (pp. 423–453).

Ugray, Z., Lasdon, L., Plummer, J., Glover, F., Kelly, J., & Martí, R. (2007). Scatter search and local NLP solvers: A multistart framework for global optimization. *INFORMS Journal on Computing*, *19*(3), 328–340.

van Es, B. (1992). Estimating functionals related to a density by a class of statistics based on spacings. *Scandinavian Journal of Statistics*, *19*(1), 61–72.

Vanduffel, S., & Yao, J. (2017). A stein type lemma for the multivariate generalized hyperbolic distribution. *European Journal of Operational Research*, *261*(2), 606–612.

Vasicek, O. (1976). A test for normality based on entropy. *Journal of the Royal Statistical Society Series B (Methodological)*, *38*(1), 54–59.

Vermorken, M., Medda, F., & Schroder, T. (2012). The diversification delta: A higher-moment measure for portfolio diversification. *Journal of Portfolio Management*, *39*(1), 67–74.

Vrins, F., Pham, D., & Verleysen, M. (2007). Mixing and non-mixing local minima of the entropy contrast for blind source separation. *IEEE Transactions on Information Theory*, *53*(3), 1030–1042.

Wachowiak, M., Smolikova, R., Tourassi, G., & Elmaghraby, A. (2005). Estimation of generalized entropies with sample spacing. *Pattern Analysis and Applications*, *8*, 95–101.

Yang, J., & Qiu, W. (2005). A measure of risk and a decision-making model based on expected utility and entropy. *European Journal of Operational Research*, *164*(3), 792–799.

Zhou, R., Cai, R., & Tong, G. (2013). Applications of entropy in finance: A review. *Entropy*, *15*, 4909–4931.

Zografos, K., & Nadarajah, S. (2003). Formulas for Rényi information and related measures for univariate distributions. *Information Sciences*, *155*(1–2), 119–138.