



# Measures of kurtosis: inadmissible for asymmetric distributions?

Andreas Eberl<sup>1</sup>  · Bernhard Klar<sup>1</sup>

Received: 27 April 2023 / Accepted: 19 February 2024  
© The Author(s) 2024

## Abstract

Skewness and kurtosis are natural characteristics of a distribution. While it has long been recognized that they are more intrinsically entangled than other characteristics like location and dispersion, this has recently been made more explicit by Eberl and Klar (Stat Papers 65:415–433, 2024) with regard to orders of kurtosis. In this paper, we analyze the implications of this entanglement on kurtosis measures in general and for several specific examples. As a key finding, we show that kurtosis measures that are defined in the classical order-based way, which is analogous to measures of location, dispersion and skewness, do not exist. This raises serious doubts about the frequent application of such measures to skewed data. We then consider old and new proposals for kurtosis measures and evaluate under which additional conditions they measure kurtosis in a meaningful way. Some measures also allow more specific representations of the influence of skewness on the measurement of kurtosis than are available in a general setting. This works particularly well for a family of newly introduced density-based kurtosis measures.

**Keywords** Asymmetric distributions · Kurtosis · Skewness · Dispersion measures · Stochastic orders

## 1 Introduction

Pearson (1895) introduced the concept of kurtosis in the form of the fourth standardized moment. Contrary to location, dispersion and skewness, which can also be measured using (standardized) moments, it took a long time for kurtosis to be understood as a more general concept (Chissom 1970; Darlington 1970). Bickel and Lehmann (1975)

---

✉ Andreas Eberl  
andreas.eberl@kit.edu

Bernhard Klar  
bernhard.klar@kit.edu

<sup>1</sup> Institute of Stochastics, Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany

defined a measure of kurtosis as a “suitable ratio” of two measures of scale or dispersion without going into more detail. In the same paper, a general approach for the definition of characteristics of probability distributions was laid out using stochastic orderings. This idea was picked up by Oja (1981), who incorporated work on convex transformations by van Zwet (1964) to define a class of stochastic orderings that can be used as foundation for the measurement of location, dispersion, skewness and kurtosis. His crucial property for a kurtosis measure  $\kappa$  is that  $F \leq G$  for distribution functions  $F$  and  $G$  and a suitable stochastic order of kurtosis  $\leq$  implies  $\kappa(F) \leq \kappa(G)$ . Here, the usual choice for the order  $\leq$  is  $\leq_s$ , which means that the function  $F \circ G^{-1}$  is concave-convex, where  $F$  and  $G$  are assumed to be continuous and strictly increasing. In the years since, a number of alternative kurtosis measures have been proposed. Most of them are based on quantiles (see, e.g., Ruppert 1987; Balanda and MacGillivray 1988; Moors 1988 and Jones et al. 2011) with a notable exception being a measure based on L-moments by Hosking (1989).

Kurtosis measures are frequently used in different fields of applications, ranging from environmental science to finance and insurance. Here, it is commonplace that these measures are applied to skewed distributions, see Cooper (2020) and López-Martín et al. (2022) for specific examples and Eberl and Klar (2024) for a broader overview. In contrast, it has long been established in the statistical literature that a major problem arising in the treatment of kurtosis is its intrinsic entanglement with skewness. This is already implied by the fact that the orders of kurtosis by van Zwet (1964) and Oja (1981) are only defined for symmetric distributions. More recently, Blest (2003) and Jones et al. (2011) have presented approaches to construct skewness-invariant measures of kurtosis that are based on moments and quantiles, respectively.

Eberl and Klar (2024) examine the relationship between skewness and orders of kurtosis, which work as a foundation for the measurement of kurtosis. There, the drawbacks of earlier approaches by MacGillivray and Balanda (1988) and Balanda and MacGillivray (1990) are pointed out and the compatibility of two major types of kurtosis orders with asymmetric distributions is analyzed in detail. The first order is a generalization of  $\leq_s$ , and the second is  $\leq_3$ , which is defined via  $F \circ G^{-1}$  being convex of order three. This order is stronger than  $\leq_s$  and is a canonical choice in view of basic orders of location, dispersion and skewness. It is established that these kurtosis orders are not transitive for distributions with different skewness, which implies that these orders cannot be characterized by appropriate measures of kurtosis. In order to circumvent this problem, so-called transitivity sets are introduced that consist of equally skewed distributions and it is shown that the fundamental kurtosis order  $\leq_3$  is transitive on these sets.

In this paper, we examine and discuss the implications of these results for measures of kurtosis. Specifically, we show that no meaningful kurtosis measures exist that fulfil a slight modification of the usual order-based defining properties. This result follows directly from the non-transitivity of the aforementioned kurtosis orders shown by Eberl and Klar (2024). The emerging doubt as to whether kurtosis measures are still meaningful is then answered as follows: we consider a number of kurtosis measures considered in the literature as well as new proposals and analyze whether they are compatible with kurtosis orderings if they are restricted to transitivity sets. Subsequently, one of the first ideas concerning the systematic measurement of kurtosis

by Bickel and Lehmann (1975), namely considering ratios of dispersion measures, is evaluated. The analysis of the kurtosis measures is concluded by a comparison of their numerical values for a few specific distributions as well as the proposal of empirical counterparts along with some of their asymptotic properties.

## 2 Kurtosis orders and transitivity sets

In this section, we summarize several definitions and results from Eberl and Klar (2024). Throughout the paper, we assume that all distribution functions are three times differentiable and have interval support. We also adopt the notation from Eberl and Klar (2024) as we denote the interior of the support of a distribution function  $F$  by  $D_F$  and assume  $f = F'$  to be strictly positive on  $D_F$ . The set of all distribution functions that satisfy these assumptions is denoted by  $\mathcal{P}$ . Furthermore,  $X$  always denotes a random variable with distribution function  $F$ .

The following definition is a modified version of those made by Oja (1981) and Eberl and Klar (2024) whereas the term *relative inverse distributions function* is adopted from Müller and Stoyan (2002).

**Definition 1** Let  $F, G \in \mathcal{P}$  and  $k \in \mathbb{N}_0$ .

(a) The function

$$R_{FG} : D_F \rightarrow D_G, \quad x \mapsto G^{-1}(F(x))$$

is said to be the *relative inverse distribution function (RIDF)* from  $F$  to  $G$ . We also define the modification  $\Delta_{FG} : D_F \rightarrow \mathbb{R}, x \mapsto R_{FG}(x) - x$ .

(b)  $F$  is said to precede  $G$  in the *order of the  $k$ -th convex characteristic*, denoted by  $F \leq_k G$ , if  $\Delta_{FG}^{(k)} \geq 0$  holds. Here, we assume that the  $k$ -th derivative  $\Delta_{FG}^{(k)}$  of  $\Delta_{FG}$  exists.

Note that, for  $k \geq 2$ ,  $F \leq_k G$  is equivalent to  $R_{FG}^{(k)} \geq 0$ . For  $k = 0, 1, 2$ , the order of the  $k$ -th convex characteristic coincides with basic orders of location, dispersion and skewness:  $\leq_0$  is the usual stochastic order  $\leq_{st}$ ,  $\leq_1$  is the dispersive order  $\leq_{disp}$  and  $\leq_2$  is the convex transformation order  $\leq_c$  by van Zwet (1964) (see also Oja 1981). Although the kurtosis order  $\leq_3$  follows the same pattern, it is rarely mentioned in the literature. A much more popular kurtosis order is  $\leq_s$ , which was proposed by van Zwet (1964) and subsequently utilized by Oja (1981) and Groeneveld and Meeden (1984), among others. It is defined in the following, along with a generalization by Eberl and Klar (2024).

**Definition 2** Let  $F, G \in \mathcal{P}$  and  $t_0 \in \mathbb{R}$ .

(a)  $F$  is said to be *less kurtotic than  $G$  in concave-convex sense*, denoted by  $F \leq_s G$ , if

$$R''_{FG}(x) \begin{cases} \leq 0, & \text{if } x < 0, \\ \geq 0, & \text{if } x > 0. \end{cases}$$

(b)  $F$  is said to be *less kurtotic than*  $G$  in the *generalized concave-convex sense* with threshold  $t_0$ , denoted by  $F \leq_{gs}^{t_0} G$ , if there exists a  $p_{FG} \in [0, 1]$  such that

$$R''_{FG}(x) \begin{cases} \leq t_0, & \text{if } x < p_{FG}, \\ \geq t_0, & \text{if } x > p_{FG}. \end{cases}$$

Eberl and Klar (2024) proved the following results concerning the relationship between the kurtosis orders and their transitivity properties. Define

$$\gamma_D^p : \mathcal{P} \rightarrow \mathbb{R}, \quad F \mapsto \frac{f'(F^{-1}(p))}{(f(F^{-1}(p)))^2}$$

for  $p \in (0, 1)$ , which satisfies the crucial property of a skewness measure as it preserves the convex transformation order  $\leq_2$  (see Eberl and Klar 2024). Furthermore, define  $\mathcal{T}_{D,p}^t = \{F \in \mathcal{P} : \gamma_D^p(F) = t\}$  for  $p \in (0, 1), t \in \mathbb{R}$ . Note that  $\mathcal{S} \subseteq \mathcal{T}_{D,1/2}^0$  holds, where  $\mathcal{S}$  denotes the subset of  $\mathcal{P}$  containing all symmetric distributions.

**Theorem 3** For  $F, G \in \mathcal{P}$ ,  $F \leq_{gs}^{t_0} G$  for all  $t_0 \in \mathbb{R}$  is equivalent to  $F \leq_3 G$ .

**Theorem 4** (a) Neither the order  $\leq_3$  nor  $\leq_{gs}^{t_0}$  is transitive on  $\mathcal{P}$  for any  $t_0 \in \mathbb{R}$ .  
 (b) The order  $\leq_3$  is transitive on the set  $\mathcal{T}_{D,p}^t$  for all  $p \in (0, 1), t \in \mathbb{R}$ .

### 3 Measures of kurtosis: (non-)existence

Contrary to kurtosis orders, measures of kurtosis are often applied to asymmetric distributions (see, e.g., Wheeler 1975; Hanook et al. 2013 and the references in Eberl and Klar 2024). Measures are mostly chosen based on historical relevance and ease of use, typically resulting in the fourth standardized moment. Since this practice obviously lacks rigour, we start out by giving a general framework for the definition of kurtosis measures. Let  $\mathcal{Q} \subseteq \mathcal{P}$  be the subset of all distributions for which a given kurtosis measure candidate is defined. Oja (1981) called a mapping  $\kappa : \mathcal{Q} \rightarrow \mathbb{R}$  a measure of kurtosis, if it satisfies the following two properties:

- (K1)  $\kappa(aX + b) = \kappa(X)$  for all  $a \in \mathbb{R} \setminus \{0\}, b \in \mathbb{R}$  and  $F \in \mathcal{Q}$ ,  
 (K2)  $\kappa(F) \leq \kappa(G)$  for all  $F, G \in \mathcal{Q}$  such that  $F \leq_K G$  for some kurtosis order  $\leq_K$ .

This definition is in line with that of measures of central location, dispersion and skewness as given by, e.g., Oja (1981). The crucial property is (K2), requiring that the measure preserves a certain order with respect to the relevant characteristic. However, an evaluation of kurtosis measures with respect to this property has only been done under the assumption of symmetry in the literature (van Zwet 1964, Ruppert 1987 or Hosking 1989).

Based on Eberl and Klar (2024), the order  $\leq_3$  seems to be superior to the generalized concave-convex order  $\leq_{gs}^{t_0}$  for three major reasons. First, while  $\leq_3$  is unambiguous,  $\leq_{gs}^{t_0}$  depends on the threshold  $t_0$  for which there exists no obvious choice if the distribution is asymmetric. Second,  $\leq_3$  is stronger than  $\leq_{gs}^{t_0}$  for any threshold  $t_0$ , meaning

that the former imposes a more fundamental requirement on corresponding kurtosis measures. Third,  $\leq_3$  is transitive if it is restricted to sets of constant density-based skewness, which could not be shown for  $\leq_{gs}^{t_0}$ . The order  $\leq_3$  is also the kurtosis analogue of the fundamental orders  $\leq_0$ ,  $\leq_1$  and  $\leq_2$  for location, dispersion and skewness. However, by refining the counterexample for the general transitivity of  $\leq_3$  in Eberl and Klar (2024), we observe that even this most basic requirement for a kurtosis measure cannot be satisfied in a meaningful way. For that, define the strict version  $<$  of any stochastic order  $\leq$  by  $F < G$ , if  $F \leq G$  and  $G \not\leq F$ .

- Theorem 5** (a) *There exists no mapping  $\kappa : \mathcal{P} \rightarrow \mathbb{R}$  such that  $F \leq_3 G$  implies  $\kappa(F) \leq \kappa(G)$  and  $F <_3 G$  implies  $\kappa(F) < \kappa(G)$  for all  $F, G \in \mathcal{P}$ .*  
 (b) *Let  $t_0 \in \text{int}(R''_{FG}(D_F))$ . There exists no mapping  $\kappa : \mathcal{P} \rightarrow \mathbb{R}$  such that  $F \leq_{gs}^{t_0} G$  implies  $\kappa(F) \leq \kappa(G)$  and  $F <_{gs}^{t_0} G$  implies  $\kappa(F) < \kappa(G)$  for all  $F, G \in \mathcal{P}$ .*

**Proof** (a) As in Example 5 in Eberl and Klar (2024), we consider the following distributions:

$$\begin{aligned} F &: [0, 1] \rightarrow [0, 1], \quad t \mapsto t^3, \\ G &: [0, 1] \rightarrow [0, 1], \quad t \mapsto t, \\ H &: [0, 1] \rightarrow [0, 1], \quad t \mapsto 1 - \sqrt[3]{1-t}. \end{aligned}$$

Note that  $F, G, H \in \mathcal{P}$ . Eberl and Klar (2024) proved that  $F \leq_3 G$  and  $G \leq_3 H$  holds, but also  $F \not\leq_3 H$ . Note that this does not imply  $H \leq_3 F$  as  $\leq_3$  is not a total relation.

Now we assume that there exists a mapping  $\kappa : \mathcal{P} \rightarrow \mathbb{R}$  that preserves the order  $\leq_3$ . It follows that  $\kappa(F) \leq \kappa(G) \leq \kappa(H)$ . We now contradict this by showing  $H \leq_3 F$ , which then implies  $H <_3 F$  and  $\kappa(H) < \kappa(F)$ . To this end, it holds that

$$R_{HF}(t) : [0, 1] \rightarrow [0, 1], \quad t \mapsto F^{-1}(H(t)) = \left(1 - (1-t)^{1/3}\right)^{1/3}.$$

It follows that, for  $t \in [0, 1]$ ,

$$\begin{aligned} R'_{HF}(t) &= \frac{1}{9} (1 - (1-t)^{1/3})^{-2/3} (1-t)^{-2/3}, \\ R''_{HF}(t) &= \frac{2}{27} (1 - (1-t)^{1/3})^{-2/3} (1-t)^{-5/3} - \frac{2}{81} (1 - (1-t)^{1/3})^{-5/3} (1-t)^{-4/3}, \\ R'''_{HF}(t) &= \frac{10}{81} (1 - (1-t)^{1/3})^{-2/3} (1-t)^{-8/3} - \frac{4}{81} (1 - (1-t)^{1/3})^{-5/3} (1-t)^{-7/3} \\ &\quad + \frac{10}{729} (1 - (1-t)^{1/3})^{-8/3} (1-t)^{-2}. \end{aligned}$$

Now,

$$\begin{aligned}
H \leq_3 F &\Leftrightarrow R''_{HF}(t) \geq 0 \quad \forall t \in [0, 1] \\
&\Leftrightarrow \frac{10}{81} \left(1 - (1-t)^{1/3}\right)^2 - \frac{4}{81} \left(1 - (1-t)^{1/3}\right) (1-t)^{1/3} \\
&\quad + \frac{10}{729} (1-t)^{2/3} \geq 0 \quad \forall t \in [0, 1] \\
&\Leftrightarrow (1-t)^{2/3} - \frac{27}{17} (1-t)^{1/3} + \frac{45}{68} \geq 0 \quad \forall t \in [0, 1] \\
&\Leftrightarrow \left((1-t)^{1/3} - \frac{27}{34}\right)^2 \geq -\left(\frac{3}{17}\right)^2 \quad \forall t \in [0, 1].
\end{aligned}$$

Since the last inequality is obviously true for all  $t \in [0, 1]$ , this concludes the proof.

- (b) Let  $F_i$  denote the cdf of a Weibull distributed random variable with shape parameter  $i > 0$  and consider the triple  $(F_j, F_k, F_\ell)$  with  $2\ell < k$  and  $\ell \in (\frac{j}{2}, j)$ . It follows that the conditions  $j \notin (k, 2k)$ ,  $k \notin (\ell, 2\ell)$  and  $\ell \in (\frac{j}{2}, j)$  are satisfied. According to the consideration of the Weibull distribution in Section 4.1 of Eberl and Klar (2024), these three conditions are, in order, equivalent to  $F_j \leq_{gs}^{t_0} F_k$ ,  $F_k \leq_{gs}^{t_0} F_\ell$  and  $F_\ell <_{gs}^{t_0} F_j$ , thus concluding the proof by contradiction.  $\square$

Obviously, the statement of Theorem 5 is also valid if the set  $\mathcal{P}$  is replaced by any other set of distributions that includes the three used for the counterexample or any other triple of distributions that poses an analogous contradiction.

Since there is an additional assumption made in Theorem 5, we cannot conclude that there is no kurtosis measure that satisfies property (K2) on a sufficiently rich set of distributions. This is due to the fact that the mapping  $\kappa_T \equiv k$  for some  $k \in \mathbb{R}$  is a trivial kurtosis measure which satisfies properties (K1) and (K2). Likewise, the mapping  $\kappa_T \equiv 0$  is a measure of central location, dispersion and skewness. To exclude the trivial measure, property (K2) could be extended to

- (K2')  $\kappa(F) \leq \kappa(G)$  for all  $F, G \in \mathcal{Q}$  such that  $F \leq_K G$  and  $\kappa(F) < \kappa(G)$  for all  $F, G \in \mathcal{Q}$  such that  $F <_K G$  for some kurtosis order  $\leq_K$  with strict version  $<_K$ .

Then, Theorem 5 proves that there exist no kurtosis measures based on the kurtosis orders  $\leq_3$  or  $\leq_{gs}^{t_0}$ . Furthermore, we can conclude that there exist no kurtosis measures based on any kurtosis order which is weaker than  $\leq_3$  in both the strict and non-strict version.

Since the proof of Theorem 5 is based upon the intransitivity of  $\leq_3$ , the question arises how kurtosis measures behave on the transitivity sets of  $\leq_3$ . However, that question needs to be addressed separately for each candidate, which is done in the following section. There, we focus on the fulfilment of (K2) for  $\leq_3$  and the generalized concave-convex order.

## 4 Measures of kurtosis: examples and properties

### 4.1 The standardized fourth moment

The earliest attempt at measuring kurtosis is attributed to Pearson (1895) and is given by the standardized fourth moment

$$\kappa_M : \mathcal{L}^4 \rightarrow \mathbb{R}, \quad X \mapsto \mathbb{E} \left[ \left( \frac{X - \mu_X}{\sigma_X} \right)^4 \right]$$

(denoted by  $\beta_2$  by Pearson), where  $\mathcal{L}^n \subseteq \mathcal{P}$  denotes the set of all distributions with finite  $n$ -th moment for  $n \in \mathbb{N}$ . Ever since, the concept of kurtosis and what it describes has been much discussed. However, its oldest measure is still its most prominent one and is often understood as synonymous with the notion of kurtosis itself (see, e.g., McAlevey and Stent 2018 or Crack 2022). The use of  $\kappa_M$  as a kurtosis measure is in agreement with the use of the first three (standardized) moments as measures of central location, dispersion and skewness, respectively. The fourth moment, however, is the first that cannot be standardized with respect to a lower moment, namely the third one. The fact that  $\kappa_M$  is not invariant to skewness (in terms of the standardized third moment  $\gamma_M$ ) is reflected in the inequality

$$\kappa_M(X) \geq (\gamma_M(X))^2 + 1, \quad F \in \mathcal{L}^4 \quad (1)$$

(Pearson 1916, p. 432). The inequality states that a distribution that is strongly skewed in either direction necessarily has a higher kurtosis than less skewed distributions. For example, consider a normally distributed random variable  $Z$ . Since  $\kappa_M(Z) = 3$ , it is less kurtotic with respect to  $\kappa_M$  than any random variable that is sufficiently skewed to satisfy  $|\gamma_M(X)| > \sqrt{2}$  (like, e.g.,  $X \sim \text{Exp}(\lambda)$ ,  $\lambda > 0$ , with  $\gamma_M(X) = 2$ ). Hence, distributions are generally not comparable with respect to  $\kappa_M$  if they exhibit a large difference in skewness. For symmetric distributions, the measure  $\kappa_M$  preserves the order  $\leq_s$  (van Zwet 1964, pp. 20–21). The fact that  $\leq_{gs}^0$  is generally weaker than  $\leq_3$  and equivalent to  $\leq_s$  for symmetric distributions gives the following result.

**Theorem 6** *If the mapping  $\kappa_M$  is restricted to the domain  $\mathcal{S}$ , it satisfies property (K2) for the kurtosis orders  $\leq_3$  and  $\leq_{gs}^0$ .*

In fact, the result by van Zwet includes more than that: it states that every even standardized moment higher than the second satisfies property (K2) if restricted to symmetric distributions. This is related to the fact that the  $2k$ -th moment,  $k \geq 2$ , can only be standardized with respect to the first two moments and not with respect to higher moments. Analogously, it is easy to show that the generalization of the standard deviation to  $\sqrt[2k]{\mathbb{E}[(X - \mu_X)^{2k}]}$  is a measure of dispersion for all  $k \geq 1$ , since, contrary to the kurtosis measure, the dispersion is not standardized out of the measure. Because there is no way of standardizing with respect to kurtosis, the difference in terms of kurtosis is still represented in higher-order standardized even moments.

### 4.2 Quantile-based approaches to measuring kurtosis

Without any differentiability assumptions, convexity of order  $k$  can be defined based on differences: Oja (1981) considers determinants of matrices, while Eberl and Klar (2024) consider divided differences. Both approaches are equivalent (see, e.g., Nørlund 1926). The connection of quantile-based measures of location, dispersion and skewness with the corresponding orders of convex characteristics is best established using these difference-based definitions of convexity (see, e.g., Oja 1981). Thus, we also follow this approach for measures and orders of kurtosis. Usage of divided differences yields that  $F \leq_3 G$  is equivalent to

$$\frac{\frac{G^{-1}(p_3)-G^{-1}(p_2)}{F^{-1}(p_3)-F^{-1}(p_2)} - \frac{G^{-1}(p_2)-G^{-1}(p_1)}{F^{-1}(p_2)-F^{-1}(p_1)}}{F^{-1}(p_3) - F^{-1}(p_1)}} - \frac{\frac{G^{-1}(p_2)-G^{-1}(p_1)}{F^{-1}(p_2)-F^{-1}(p_1)} - \frac{G^{-1}(p_1)-G^{-1}(p_0)}{F^{-1}(p_1)-F^{-1}(p_0)}}{F^{-1}(p_2) - F^{-1}(p_0)}} \geq 0 \tag{2}$$

for all  $0 < p_0 < p_1 < p_2 < p_3 < 1$ . This equivalence has previously been established by Eberl and Klar (2024), who used it to define a quantile-based kurtosis functional. Contrary to a conventional measure, this functional quantifies the difference in kurtosis between two given distributions and allows the functional to be compatible with the order  $\leq_3$  in the sense that it is non-negative for arbitrary distribution functions  $F, G$  with  $F \leq_3 G$ .

Here, we use (2) as a basis for conventional kurtosis measures. Since (2) cannot generally be rewritten in a way that is symmetric in  $F$  and  $G$ , we assume that  $F$  is a symmetric distribution function, and choose  $0 < \alpha < \eta < 1/2$ . Further, put  $p_0 = \alpha, p_1 = \eta, p_2 = 1 - \eta, p_3 = 1 - \alpha$ , and define  $c = F^{-1}(\eta) - F^{-1}(\alpha) = F^{-1}(1 - \alpha) - F^{-1}(1 - \eta)$  and  $d = F^{-1}(1 - \eta) - F^{-1}(\eta)$ . Then, (2) takes the specific form

$$\frac{1}{c} \left( G^{-1}(1 - \alpha) - G^{-1}(\alpha) \right) - \left( \frac{2}{d} + \frac{1}{c} \right) \left( G^{-1}(1 - \eta) - G^{-1}(\eta) \right) \geq 0.$$

This is equivalent to

$$\frac{F^{-1}(1 - \alpha) - F^{-1}(\alpha)}{F^{-1}(1 - \eta) - F^{-1}(\eta)} \leq \frac{G^{-1}(1 - \alpha) - G^{-1}(\alpha)}{G^{-1}(1 - \eta) - G^{-1}(\eta)}. \tag{3}$$

As a consequence, for  $0 < \alpha < \eta < 1/2$ , the mapping

$$\kappa_Q^{\alpha,\eta} : \mathcal{P} \rightarrow \mathbb{R}, \quad F \mapsto \frac{F^{-1}(1 - \alpha) - F^{-1}(\alpha)}{F^{-1}(1 - \eta) - F^{-1}(\eta)} \tag{4}$$

preserves the order  $\leq_3$  on the subset  $\mathcal{S} \subseteq \mathcal{P}$  of symmetric distributions. The same is true for the alternative mapping

$$\kappa_{QA}^{\alpha,\eta} : \mathcal{P} \rightarrow \mathbb{R}, \quad F \mapsto \frac{F^{-1}(1 - \alpha) - 3F^{-1}(1 - \eta) + 3F^{-1}(\eta) - F^{-1}(\alpha)}{F^{-1}(1 - \eta) - F^{-1}(\eta)},$$

which is structurally more similar to the kurtosis functional by Eberl and Klar (2024).  $\kappa_{QA}^{\alpha,\eta}(F) \leq \kappa_{QA}^{\alpha,\eta}(G)$  can be equivalently transformed into (3) by adding 3 on either side. As  $\leq_3, \leq_{gs}^0$  is also preserved by both quantile-based measures, as the following result shows.

**Theorem 7** *Let  $0 < \alpha < \eta < 1/2$ . If the mappings  $\kappa_Q^{\alpha,\eta}$  and  $\kappa_{QA}^{\alpha,\eta}$  are restricted to the domain  $\mathcal{S}$ , they both satisfy property (K2) for the kurtosis orders  $\leq_3$  and  $\leq_{gs}^0$ .*

**Proof** Let  $F, G \in \mathcal{S}$  satisfy  $F \leq_{gs}^0 G$ . It is sufficient to show  $\kappa_Q^{\alpha,\eta}(F) \leq \kappa_Q^{\alpha,\eta}(G)$ . Since  $R_{FG}$  is antisymmetric, it is concave on  $\text{supp}(F) \cap (-\infty, 0]$  and convex on  $\text{supp}(F) \cap [0, \infty)$ . Because of  $F^{-1}(\frac{1}{2}) = 0$ , this is equivalent to

$$\frac{G^{-1}(p_2) - G^{-1}(p_1)}{F^{-1}(p_2) - F^{-1}(p_1)} - \frac{G^{-1}(p_1) - G^{-1}(p_0)}{F^{-1}(p_1) - F^{-1}(p_0)} \begin{cases} \leq 0 & , \text{ if } 0 < p_0 < p_1 < p_2 \leq \frac{1}{2}, \\ \geq 0 & , \text{ if } \frac{1}{2} \leq p_0 < p_1 < p_2 < 1. \end{cases} \tag{5}$$

Because of the symmetry of  $F$  and  $G$ , one of the two inequalities is redundant and we limit ourselves to the upper inequality. By applying  $p_0 = \alpha, p_1 = \eta, p_2 = \frac{1}{2}$  to the upper part of (5), we obtain

$$\begin{aligned} \frac{F^{-1}(1 - \alpha) - F^{-1}(\alpha)}{F^{-1}(1 - \eta) - F^{-1}(\eta)} &= \frac{F^{-1}(\frac{1}{2}) - F^{-1}(\alpha)}{F^{-1}(\frac{1}{2}) - F^{-1}(\eta)} = \frac{F^{-1}(\eta) - F^{-1}(\alpha)}{F^{-1}(\frac{1}{2}) - F^{-1}(\eta)} + 1 \\ &\leq \frac{G^{-1}(\eta) - G^{-1}(\alpha)}{G^{-1}(\frac{1}{2}) - G^{-1}(\eta)} + 1 = \frac{G^{-1}(\frac{1}{2}) - G^{-1}(\alpha)}{G^{-1}(\frac{1}{2}) - G^{-1}(\eta)} = \frac{G^{-1}(1 - \alpha) - G^{-1}(\alpha)}{G^{-1}(1 - \eta) - G^{-1}(\eta)} \end{aligned}$$

The outer identities follow from the symmetry of  $F$  and  $G$ , which yields  $F^{-1}(1 - \alpha) - F^{-1}(\frac{1}{2}) = F^{-1}(\frac{1}{2}) - F^{-1}(\alpha)$  (and the same if  $F$  is replaced by  $G$  or  $\alpha$  is replaced by  $\eta$ ). □

In fact, Theorem 7 is still true, if only the less kurtotic distribution  $F$  is assumed to be symmetric. This is already clear for  $\leq_3$  from the derivation of (3), but the following explanation for  $\leq_{gs}^0$  offers more insight into how that relates to the relationship between kurtosis and skewness.

Let  $0 < \alpha < \eta < \frac{1}{2}$  and  $F \in \mathcal{S}, G \in \mathcal{P}$  with  $F \leq_{gs}^0 G$ . Then, (5) still holds. By choosing  $p_0 = \alpha, p_1 = \eta, p_2 = \frac{1}{2}$  in the upper case of (5) and choosing  $p_0 = \frac{1}{2}, p_1 = 1 - \eta, p_2 = 1 - \alpha$  in the lower case, we obtain

$$\begin{aligned} \frac{G^{-1}(\frac{1}{2}) - G^{-1}(\alpha)}{G^{-1}(\frac{1}{2}) - G^{-1}(\eta)} &\geq \frac{F^{-1}(\frac{1}{2}) - F^{-1}(\alpha)}{F^{-1}(\frac{1}{2}) - F^{-1}(\eta)} \\ &= \frac{F^{-1}(1 - \alpha) - F^{-1}(\frac{1}{2})}{F^{-1}(1 - \eta) - F^{-1}(\frac{1}{2})} \leq \frac{G^{-1}(1 - \alpha) - G^{-1}(\frac{1}{2})}{G^{-1}(1 - \eta) - G^{-1}(\frac{1}{2})}. \end{aligned} \tag{6}$$

Since  $\kappa_Q^{\alpha,\eta}(F)$  and  $\kappa_Q^{\alpha,\eta}(G)$  are weighted averages of the one-sided quantities in (6), it follows

$$\kappa_Q^{\alpha,\eta}(F) \leq \min \left\{ \frac{G^{-1}(\frac{1}{2}) - G^{-1}(\alpha)}{G^{-1}(\frac{1}{2}) - G^{-1}(\eta)}, \frac{G^{-1}(1 - \alpha) - G^{-1}(\frac{1}{2})}{G^{-1}(1 - \eta) - G^{-1}(\frac{1}{2})} \right\} \quad (7)$$

$$\begin{aligned} &\leq \frac{G^{-1}(\frac{1}{2}) - G^{-1}(\eta)}{G^{-1}(1 - \eta) - G^{-1}(\eta)} \frac{G^{-1}(\frac{1}{2}) - G^{-1}(\alpha)}{G^{-1}(\frac{1}{2}) - G^{-1}(\eta)} \\ &\quad + \frac{G^{-1}(1 - \eta) - G^{-1}(\frac{1}{2})}{G^{-1}(1 - \eta) - G^{-1}(\eta)} \frac{G^{-1}(1 - \alpha) - G^{-1}(\frac{1}{2})}{G^{-1}(1 - \eta) - G^{-1}(\frac{1}{2})} \\ &= \kappa_Q^{\alpha,\eta}(G). \end{aligned} \quad (8)$$

Inequality (7) is basically the same inequality that was used to prove Theorem 7 and its tightness is influenced by both  $F$  and  $G$ . However, the tightness of inequality (8) only depends on  $G$  with equality holding if  $G$  is symmetric, suggesting that an increase of asymmetry for  $G$  tends to decrease the tightness of the inequality  $\kappa_Q^{\alpha,\eta}(F) \leq \kappa_Q^{\alpha,\eta}(G)$ .

The measures  $\kappa_Q^{\alpha,\eta}$  and  $\kappa_{QA}^{\alpha,\eta}$  have appeared quite often in the literature, see Ruppert (1987), Balanda and MacGillivray (1988), and Jones et al. (2011), where further references can be found. Specific choices of the parameters in the literature are  $\eta = \frac{1}{4}$  and  $\alpha = 0.05$  or  $\alpha = 0.01$ . Alternative parameter choices can be obtained through equidistant evaluation of the quantile function. For example, the quintile-based measure  $\kappa_{QA}^{1/5,2/5}$  was introduced in Jones et al. (2011, p. 90), motivated by the analogy with Bowley's skewness measure  $\gamma_Q^{1/4}$ , which takes second differences instead of third ones. Furthermore, Moors (1988, p. 26) defined the octile-based measure

$$M = \frac{F^{-1}(\frac{7}{8}) - F^{-1}(\frac{5}{8}) + F^{-1}(\frac{3}{8}) - F^{-1}(\frac{1}{8})}{F^{-1}(\frac{6}{8}) - F^{-1}(\frac{2}{8})}.$$

Noting

$$M = \kappa_Q^{1/8,2/8} - \kappa_Q^{3/8,2/8} = \kappa_Q^{1/8,2/8} - \left( \kappa_Q^{2/8,3/8} \right)^{-1},$$

it follows that  $M$  preserves the orders  $\leq_3$  and  $\leq_{gs}$  for symmetric distributions as well.

### 4.3 Density-based approaches to measuring kurtosis

In this section, we consider kurtosis measures based on the density and its derivatives. The proposed quantities at times bear a resemblance to those used in Critchley and Jones (2008), where a similar idea is addressed using a different approach.

Eberl and Klar (2024) showed that

$$\gamma_D^p : \mathcal{P} \rightarrow \mathbb{R}, \quad F \mapsto - \frac{f'(F^{-1}(p))}{(f(F^{-1}(p)))^2}$$

satisfies the crucial property of a skewness measure for all  $p \in (0, 1)$ , i.e. it preserves the order  $\leq_2$ . In fact, it is easy to see that  $R''_{FG}(x) \geq 0$  for all  $x \in D_F$ , and thereby also  $F \leq_2 G$ , is equivalent to  $\gamma_D^p(F) \leq \gamma_D^p(G)$  for all  $p \in (0, 1)$ . The defining inequalities  $\Delta_{FG}(x) \geq 0$  and  $\Delta'_{FG}(x) \geq 0$  of the orders  $\leq_0$  and  $\leq_1$  of location and dispersion can be transformed in a similar way. By defining  $v_D^p(F) = F^{-1}(p)$ ,  $p \in (0, 1)$ , as a location measure and  $\tau_D^p(F) = (f(F^{-1}(p)))^{-1}$  as a dispersion measure, it can be shown that  $F \leq_0 G$  is equivalent to  $v_D^p(F) \leq v_D^p(G)$  for all  $p \in (0, 1)$  and that  $F \leq_1 G$  is equivalent to  $\tau_D^p(F) \leq \tau_D^p(G)$  for all  $p \in (0, 1)$ .

In order to use the same methodology for measures of kurtosis, we first calculate

$$R'''_{FG}(t) = \left( \frac{1}{g(R_{FG}(t))} \right)^5 \cdot \left[ f''(t)(g(R_{FG}(t)))^4 - 3f(t)f'(t)(g(R_{FG}(t)))^2g'(R_{FG}(t)) \right. \\ \left. + 3(f(t))^3(g'(R_{FG}(t)))^2 - (f(t))^3g(R_{FG}(t))g''(R_{FG}(t)) \right].$$

Based on this, we find

$$F \leq_3 G \Leftrightarrow f''(t)(g(R_{FG}(t)))^4 - (f(t))^3g(R_{FG}(t))g''(R_{FG}(t)) \\ \geq 3 \left[ f(t)f'(t)(g(R_{FG}(t)))^2g'(R_{FG}(t)) - (f(t))^3(g'(R_{FG}(t)))^2 \right] \quad \forall t \in D_F \\ \Leftrightarrow \frac{f''(t)}{(f(t))^3} - \frac{g''(R_{FG}(t))}{(g(R_{FG}(t)))^3} \geq 3 \frac{f'(t)}{(f(t))^2} \frac{g'(R_{FG}(t))}{(g(R_{FG}(t)))^2} - 3 \frac{(g'(R_{FG}(t)))^2}{(g(R_{FG}(t)))^4} \\ \forall t \in D_F \\ \Leftrightarrow \frac{f''(F^{-1}(p))}{(f(F^{-1}(p)))^3} - \frac{g''(G^{-1}(p))}{(g(G^{-1}(p)))^3} \\ \geq 3 \frac{g'(G^{-1}(p))}{(g(G^{-1}(p)))^2} \left( \frac{f'(F^{-1}(p))}{(f(F^{-1}(p)))^2} - \frac{g'(G^{-1}(p))}{(g(G^{-1}(p)))^2} \right) \quad \forall p \in (0, 1). \quad (9)$$

By plugging in the definition of  $\gamma_D^p$ ,  $p \in (0, 1)$ ,  $F \leq_3 G$  is equivalent to

$$\frac{f''(F^{-1}(p))}{(f(F^{-1}(p)))^3} - \frac{g''(G^{-1}(p))}{(g(G^{-1}(p)))^3} \geq 3\gamma_D^p(G)(\gamma_D^p(F) - \gamma_D^p(G)) \quad \forall p \in (0, 1).$$

If we now assume that both  $F$  and  $G$  are symmetric,  $\gamma_D^{1/2}(F) - \gamma_D^{1/2}(G) = 0$  holds for the specific choice  $p = \frac{1}{2}$ . In that case,  $F \leq_3 G$  implies

$$-\frac{f''(F^{-1}(\frac{1}{2}))}{(f(F^{-1}(\frac{1}{2})))^3} \leq -\frac{g''(G^{-1}(\frac{1}{2}))}{(g(G^{-1}(\frac{1}{2})))^3}.$$

This suggests that a reasonable choice for a density-based kurtosis measure is obtained as the special case  $p = \frac{1}{2}$  from the following class of mappings

$$\kappa_D^p : \mathcal{P} \rightarrow \mathbb{R}, \quad F \mapsto -\frac{f''(F^{-1}(p))}{(f(F^{-1}(p)))^3}, \quad (10)$$

where  $p \in (0, 1)$ . By plugging this into (9), we see that  $F \leq_3 G$  is equivalent to

$$\kappa_D^p(G) - \kappa_D^p(F) \geq 3\gamma_D^p(G)(\gamma_D^p(F) - \gamma_D^p(G)) \quad \forall p \in (0, 1). \quad (11)$$

Additionally, swapping the roles of  $F$  and  $G$  yields that  $G \leq_3 F$  is equivalent to

$$\kappa_D^p(G) - \kappa_D^p(F) \leq 3\gamma_D^p(F)(\gamma_D^p(F) - \gamma_D^p(G)) \quad \forall p \in (0, 1). \quad (12)$$

Now, if  $\kappa_D^p$  is heuristically understood to measure kurtosis and  $\gamma_D^p$  is understood to measure skewness, this statement can be interpreted as follows: if two distributions are equally skewed, the right sides of both (11) and (12) vanish, meaning that the kurtosis comparison between  $F$  and  $G$  is centred around zero.  $G$  exhibits more kurtosis than  $F$ , if  $\kappa_D^p(G) \geq \kappa_D^p(F)$ , and vice versa, with both cdf's being equivalent in terms of kurtosis, if the two quantities are equal. This behaviour can be summarized as follows.

**Theorem 8** *Let  $t \in \mathbb{R}$  and  $p \in (0, 1)$ . If  $F, G \in \mathcal{T}_{D,p}^t$  satisfy  $F \leq_3 G$ , then  $\kappa_D^p(F) \leq \kappa_D^p(G)$ .*

It can be shown that  $\kappa_D^p$  satisfies (K1), if and only if  $p = \frac{1}{2}$ . Thus, for any  $t \in \mathbb{R}$ ,  $\kappa_D^{1/2}$  restricted to the set  $\mathcal{T}_D^t$  is a measure of kurtosis in the sense of (K1) and (K2). Note that  $\mathcal{T}_D^0$  is a superset of  $\mathcal{S}$ .

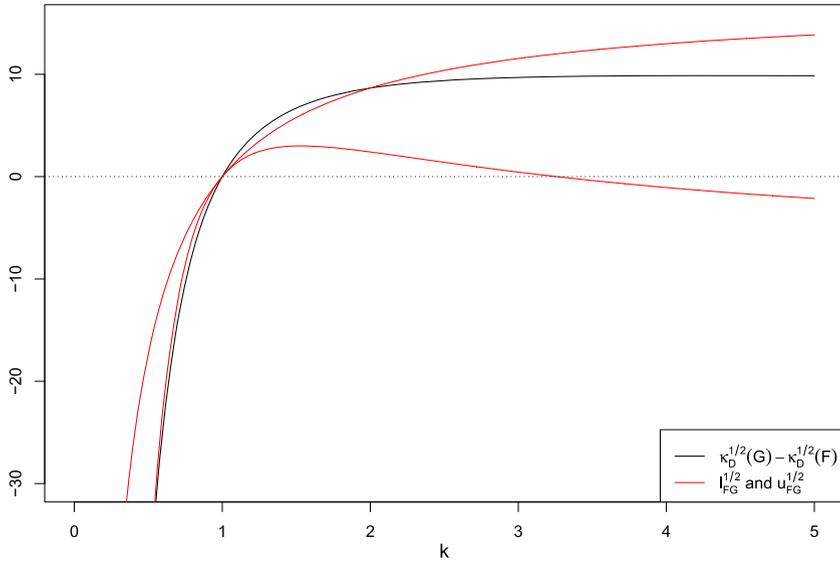
However, the situation gets more complex if  $F$  and  $G$  differ in terms of skewness. Define  $\ell_{FG}^p = 3\gamma_D^p(G)(\gamma_D^p(F) - \gamma_D^p(G))$  as the lower limit for  $F \leq_3 G$  and  $u_{FG}^p = 3\gamma_D^p(F)(\gamma_D^p(F) - \gamma_D^p(G))$  as the upper limit for  $G \leq_3 F$ , both for all  $p \in (0, 1)$ . Since

$$u_{FG}^p - \ell_{FG}^p = 3(\gamma_D^p(F) - \gamma_D^p(G))^2 \geq 0,$$

the centre of their comparison in terms of skewness extends to an interval of length  $3(\gamma_D^p(F) - \gamma_D^p(G))^2$ . Hence, the bigger the difference in skewness is between the two distributions, the larger the interval that is associated with equivalence with respect to the kurtosis order  $\leq_3$ . The aforementioned centre of the kurtosis comparison not only extends to an interval but also shifts, depending on the concrete values of  $\gamma_D^p(F)$  and  $\gamma_D^p(G)$  and, more specifically, on their signs. It is easy to see from the definitions that  $\ell_{FG}^p$  and  $u_{FG}^p$  have the same sign, if and only if  $\gamma_D^p(F)$  and  $\gamma_D^p(G)$  have the same sign. Note that  $\ell_{FG}^p$  and  $u_{FG}^p$  having the same sign means that  $\kappa_D^p(F) = \kappa_D^p(G)$  for all  $p \in (0, 1)$  implies  $F <_3 G$  if this sign is negative and  $G <_3 F$  if it is positive. Furthermore, 0 is contained in the interval  $[\ell_{FG}^p, u_{FG}^p]$ , around which the kurtosis comparison is centred, if and only if  $\gamma_D^p(F)$  and  $\gamma_D^p(G)$  have differing signs (including the case that either is zero).

The interplay of the difference of kurtosis measures and the upper and lower limits with respect to the order  $\leq_3$  is considered in the following for the example of Weibull distributions.

**Example 9** Let  $X \sim \text{Weib}(1)$  with cdf  $F$  and let  $Y \sim \text{Weib}(k)$  with cdf  $G$  for  $k > 0$ , where  $\text{Weib}(k)$  denotes the Weibull distribution with shape parameter  $k$  and scale



**Fig. 1** Graph of the difference of density-based kurtosis measures for the Weib(1) and the Weib( $k$ )-distribution

parameter 1. The difference  $\kappa_D^p(G) - \kappa_D^p(F)$  for  $p = \frac{1}{2}$  is plotted in Fig. 1 as a function of  $k$ . Additionally, the plot shows in red the lower and upper limits  $\ell_{FG}^{1/2}$  and  $u_{FG}^{1/2}$  from (11) and (12). All three graphs are obviously zero at  $k = 1$  since  $F = G$  holds in that case. For  $1 < k < 2$ , we have  $\kappa_D^{1/2}(G) - \kappa_D^{1/2}(F) > u_{FG}^{1/2}$ . This observation is in line with  $F <_3 G$  holding for exactly those values of  $k$  as mentioned in Eberl and Klar (2024) because of the equivalence

$$F <_3 G \Leftrightarrow F \leq_3 G \text{ and } G \not\leq_3 F \Leftrightarrow \kappa_D^p(G) - \kappa_D^p(F) \geq \ell_{FG}^p \quad \forall p \in (0, 1) \text{ and} \\ \exists p_0 \in (0, 1) : \kappa_D^{p_0}(G) - \kappa_D^{p_0}(F) > u_{FG}^{p_0}. \tag{13}$$

In this case, the latter inequality holds for  $p_0 = \frac{1}{2}$ . The observation  $\kappa_D^{1/2}(G) - \kappa_D^{1/2}(F) < \ell_{FG}^{1/2}$  for  $\frac{1}{2} < k < 1$  is in accordance with the fact that  $G <_3 F$  holds in a similar way. This is not immediately obvious from Fig. 1 but can be validated by rescaling the plot window. The fact that  $\kappa_D^{1/2}(G) - \kappa_D^{1/2}(F) \in [\ell_{FG}^{1/2}, u_{FG}^{1/2}]$  holds for  $k < \frac{1}{2}$  and  $k > 2$  is an implication of  $F =_3 G$ . For the same reason,  $\kappa_D^p(G) - \kappa_D^p(F) \in [\ell_{FG}^p, u_{FG}^p]$  follows for all  $p \in (0, 1)$ .

Note that the family of Weibull distributions decreases in skewness with respect to  $\leq_2$  as  $k$  increases. Hence, the sign change of both  $\ell_{FG}^{1/2}$  and  $u_{FG}^{1/2}$  at  $k = 1$  stems from the fact that  $\gamma_D^{1/2}(G) - \gamma_D^{1/2}(F)$  changes its sign from positive to negative. The additional sign change of  $\ell_{FG}^{1/2}$  at  $k \approx 3.26$  stems from  $\gamma_D^{1/2}(G)$  changing sign. The

Weibull distribution changes from right-skewed to left-skewed around that value of  $k$  with the exact value being determined by the utilized skewness measure.

Overall, (11) and (12) suggest that the comparison of distributions in terms of kurtosis has to account for possible differences in skewness. The same observation is made in Sects. 4.1 and 4.2. In the first case, the evidence comes in the form of the well-known inequality (1). In the second case, it is shown that the tightness of the inequality  $\kappa_Q^{\alpha,\eta}(F) \leq \kappa_Q^{\alpha,\eta}(G)$  is reduced if  $F$  or  $G$  are markedly skew or if they exhibit a large difference in skewness.

A similar statement can be made about the density-based measurement of kurtosis in this section. However, compared to the ideas from Sect. 4.2, the inequalities (11) and (12) represent the influence of skewness on the measurement of kurtosis in a more precise way. Since these inequalities also characterize the kurtosis order  $\leq_3$ , this helps to understand why it is difficult to meaningfully apply the order to distributions with a significant difference in skewness.

Similarly to the quantile-based kurtosis functional  $\kappa_{QF}^\alpha, \alpha \in (0, \frac{1}{2})$ , by Eberl and Klar (2024), which preserves the order  $\leq_3$  in the sense that  $F \leq_3 G$  implies  $\kappa_{QF}^\alpha(F, G) \geq 0$ , we can now define a density-based kurtosis functional by

$$\kappa_{DF}^p(F, G) = \kappa_D^p(G) - \kappa_D^p(F) - 3\gamma_D^p(G) (\gamma_D^p(F) - \gamma_D^p(G))$$

for all  $p \in (0, 1)$ . An advantage  $\kappa_{DF}^p$  has over  $\kappa_{QF}^\alpha$  is that it characterizes the order  $\leq_3$  in the sense that  $F \leq_3 G$  is equivalent to  $\kappa_{DF}^p(F, G) \geq 0$  for all  $p \in (0, 1)$ . However, neither  $\kappa_{DF}^p$  nor  $\kappa_{QF}^\alpha$  is symmetric in their two arguments  $F$  and  $G$ .

Another connection between density- and quantile-based measures is that the former can be obtained as a limiting case of the latter. This is also true for corresponding measures of dispersion and skewness as discussed in the following remark.

**Remark 10** Let  $F \in \mathcal{P}$ .

- (a) The canonical quantile-based measures of location are quantiles themselves, which can be seen as density-based measures of location, as stated in the beginning of this section.
- (b) The canonical quantile-based measure of dispersion is the interquantile distance (see, e.g., Müller and Stoyan 2002). In order to obtain a symmetric quantity, we define  $\tau_Q^\alpha(F) = F^{-1}(1-\alpha) - F^{-1}(\alpha), \alpha \in (0, \frac{1}{2})$ . Obviously,  $\lim_{\alpha \nearrow \frac{1}{2}} \tau_Q^\alpha(F) = 0$  holds. In order to obtain a derivative, we multiply with a factor that only depends on  $\alpha$  and not on  $F$  itself, which vanishes for  $\alpha \nearrow \frac{1}{2}$ . This yields

$$\begin{aligned} \lim_{\alpha \nearrow \frac{1}{2}} \frac{1}{1-2\alpha} \cdot \tau_Q^\alpha(F) &= \lim_{\alpha \nearrow \frac{1}{2}} \frac{F^{-1}(1-\alpha) - F^{-1}(\alpha)}{(1-\alpha) - \alpha} \\ &= (F^{-1})' \left( \frac{1}{2} \right) = \frac{1}{f(F^{-1}(\frac{1}{2}))} = \tau_D(F). \end{aligned}$$

- (c) A quantile-based skewness measure that is structurally similar to the previously presented quantities is given by

$$\gamma_D^\alpha(F) = \frac{F^{-1}(1 - \alpha) - 2F^{-1}(\frac{1}{2}) + F^{-1}(\alpha)}{F^{-1}(1 - \alpha) - F^{-1}(\alpha)}$$

(see, e.g., Groeneveld and Meeden 1984). After multiplying with a similar  $\alpha$ -dependent factor as before, we now obtain the limiting value for  $\alpha \nearrow \frac{1}{2}$  by rewriting the resulting term as the ratio of a second-order and a first-order difference quotient

$$\begin{aligned} \lim_{\alpha \nearrow \frac{1}{2}} \frac{4}{1 - 2\alpha} \cdot \gamma_Q^\alpha(F) &= 4 \cdot \lim_{\alpha \nearrow \frac{1}{2}} \frac{\frac{F^{-1}(1-\alpha) - 2F^{-1}(\frac{1}{2}) + F^{-1}(\alpha)}{(1-2\alpha)^2}}{\frac{F^{-1}(1-\alpha) - F^{-1}(\alpha)}{1-2\alpha}} = \frac{(F^{-1})''(\frac{1}{2})}{(F^{-1})'(\frac{1}{2})} \\ &= -\frac{f'(F^{-1}(\frac{1}{2}))}{(f(F^{-1}(\frac{1}{2})))^2} = \gamma_D(F), \end{aligned}$$

where we used that

$$(F^{-1})''(p) = \left( \frac{1}{f(F^{-1}(p))} \right)' = -\frac{f'(F^{-1}(p))}{(f(F^{-1}(p)))^3}$$

for  $p \in (0, 1)$ .

- (d) The quantile-based kurtosis measure that seems to be most closely related to  $\tau_Q^\alpha$  and  $\gamma_Q^\alpha$  is  $\kappa_{QA}^{\alpha,\eta}$  because it has a higher-order difference in its numerator. Because  $\kappa_{QA}^{\alpha,\eta}$  is dependent upon two parameters  $0 < \alpha < \eta < \frac{1}{2}$ , further assumptions have to be made to calculate a meaningful limit. In order to obtain a natural continuation of the previous limits,  $\alpha$  and  $\eta$  are chosen in such a way that the evaluation points of the quantile function are equidistant. Hence, we choose  $\eta = \frac{1}{2} - \beta$  and  $\alpha = \frac{1}{2} - 3\beta$  for a  $\beta \in (0, \frac{1}{6})$ , meaning that the distance between to neighbouring evaluation points is fixed to  $2\beta$ . After slightly altering the previously used rescaling factor, the following limit is obtained for  $\beta \searrow 0$

$$\begin{aligned} \lim_{\beta \searrow 0} \frac{2}{\beta^2} \cdot \kappa_{QA}^{\frac{1}{2}-3\beta, \frac{1}{2}-\beta}(F) &= 8 \cdot \lim_{\beta \searrow 0} \frac{\frac{F^{-1}(\frac{1}{2}+3\beta) - 3F^{-1}(\frac{1}{2}+\beta) + 3F^{-1}(\frac{1}{2}-\beta) - F^{-1}(\frac{1}{2}-3\beta)}{(2\beta)^3}}{\frac{F^{-1}(\frac{1}{2}+\beta) - F^{-1}(\frac{1}{2}-\beta)}{2\beta}} \\ &= \frac{(F^{-1})'''(\frac{1}{2})}{(F^{-1})'(\frac{1}{2})} = -\frac{f''(F^{-1}(\frac{1}{2}))}{(f(F^{-1}(\frac{1}{2})))^3} + 3 \left( \frac{f'(F^{-1}(\frac{1}{2}))}{(f(F^{-1}(\frac{1}{2})))^2} \right)^2 \\ &= \kappa_D^{1/2}(F) + 3(\gamma_D^{1/2}(F))^2. \end{aligned} \tag{14}$$

In contrast to the two previous parts, we do not obtain the already known density-based measure as the limiting value. However, if the limit above is defined as an alternative density-based kurtosis measure, most of the results concerning  $\kappa_D$  can be replicated.

Particularly, if we define

$$\kappa_{DA}^p : \mathcal{P} \rightarrow \mathbb{R}, \quad F \mapsto \kappa_D^p(F) + 3(\gamma_D^p(F))^2,$$

we obtain that  $F \leq_3 G$  is equivalent to

$$\kappa_{DA}^p(G) - \kappa_{DA}^p(F) \geq 3\gamma_D^p(F)(\gamma_D^p(G) - \gamma_D^p(F)) \quad \forall p \in (0, 1). \quad (15)$$

This lower bound coincides with  $\ell_{GF}^p$ , the lower bound for the corresponding difference of measures  $\kappa_D^p$  with the roles of  $F$  and  $G$  reversed. Therefore,  $\kappa_{DA}^p$  also preserves the order  $\leq_3$ , if  $F$  and  $G$  are equally skewed in the sense of  $\gamma_D^p(F) = \gamma_D^p(G)$ , or if one of the cdf's (in this case  $F$ ) is symmetric and  $p = \frac{1}{2}$ .

#### 4.4 Expectile-based approaches to measuring kurtosis

Eberl and Klar (2022, 2023) presented methods for the quantification of dispersion and skewness using expectiles. Expectiles are generalized quantiles and possess a number of desirable quantities. In the cited papers, expectile-based measures of dispersion and skewness were defined analogously to the quantile-based measures  $\tau_Q^\alpha$  and  $\gamma_Q^\alpha$  in the present paper by simply substituting the quantile function for the corresponding expectile function. In the context of this paper, we denote these measures by  $\tau_E^\alpha$  and  $\gamma_E^\alpha$ . Since the resulting quantities are at times difficult to analyze, their limits analogous to those obtained in Remark 10 were considered since they are easier to handle; they are given by  $\tau_{EL}(X) = \mathbb{E}[|X - \mu_X|]$  (the mean absolute deviation) and by  $\gamma_{EL}(F) = 2F(\mu_F) - 1$ . In this section, we proceed similarly for measures of kurtosis.

For any cdf  $F \in \mathcal{L}^1$  with  $e_F$  denoting the corresponding expectile function, we define an expectile-based kurtosis measure by

$$\kappa_E^{\alpha,\eta}(F) = \frac{e_F(1-\alpha) - 3e_F(1-\eta) + 3e_F(\eta) - e_F(\alpha)}{e_F(1-\eta) - e_F(\eta)}, \quad 0 < \alpha < \eta < \frac{1}{2}.$$

We use the analogue to  $\kappa_{QA}^{\alpha,\eta}$  instead of the equivalent analogue to  $\kappa_Q^{\alpha,\eta}$  because the former is better suited for determining its limiting value, as noted in Remark 10. In contrast to the expectile-based measures considered in Eberl and Klar (2022, 2023), the expectile kurtosis  $\kappa_E^{\alpha,\eta}$  has the major problem that it is not known whether it preserves any kurtosis order on some subset of distributions like  $\mathcal{S}$ . Thus, its limiting value is needed to draw a connection to kurtosis orders. We consider the same limit as in (14) and replace the quantile function with the corresponding expectile function, yielding

$$\lim_{\beta \searrow 0} \frac{2}{\beta^2} \cdot \kappa_E^{\frac{1}{2}-3\beta, \frac{1}{2}-\beta}(F) = 8 \cdot \lim_{\beta \searrow 0} \frac{\frac{e_F(\frac{1}{2}+3\beta) - 3e_F(\frac{1}{2}+\beta) + 3e_F(\frac{1}{2}-\beta) - e_F(\frac{1}{2}-3\beta)}{(2\beta)^3}}{\frac{e_F(\frac{1}{2}+\beta) - e_F(\frac{1}{2}-\beta)}{2\beta}}$$

$$= \frac{(e_F)'''(\frac{1}{2})}{(e_F)'(\frac{1}{2})}. \tag{16}$$

The first two derivatives of the expectile function  $e_F$ , evaluated at  $1/2$ , are given by

$$\begin{aligned} e'_F(1/2) &= 2\tau_{EL}(F), \\ e''_F(1/2) &= 8\tau_{EL}(F) \cdot (2F(\mu_F) - 1) \end{aligned}$$

(see Eberl and Klar 2022). To shorten the following calculations, we use the abbreviation  $\delta_F = \tau_{EL}(F)$ . The third derivative at  $1/2$  can be calculated similarly to the second, which utilizes the general form of the first derivative, given by

$$e'_F(\alpha) = \frac{\mathbb{E}[|X - e_F(\alpha)|]}{(1 - \alpha)F(e_F(\alpha)) + \alpha(1 - F(e_F(\alpha)))}, \quad \alpha \in (0, 1)$$

(see Holzmann and Klar 2016). If we denote the numerator and the denominator by  $u(\alpha)$  and  $v(\alpha)$ , respectively, we obtain

$$e'''_F(\frac{1}{2}) = \lim_{\alpha \nearrow \frac{1}{2}} \frac{u''(\alpha)(v(\alpha))^2 - 2u'(\alpha)v'(\alpha)v(\alpha) + 2u(\alpha)(v'(\alpha))^2 - u(\alpha)v''(\alpha)v(\alpha)}{(v(\alpha))^3}.$$

This can be calculated using

$$\begin{aligned} u(\alpha) &= \mathbb{E}[|X - e_F(\alpha)|] \xrightarrow{\alpha \rightarrow \frac{1}{2}} \delta_F, \\ u'(\alpha) &= e'_F(\alpha)(2F(e_F(\alpha)) - 1) \xrightarrow{\alpha \rightarrow \frac{1}{2}} e'_F(\frac{1}{2})(2F(\mu_F) - 1) = 2\delta_F\gamma_{EL}(F), \\ u''(\alpha) &= e''_F(\alpha)(2F(e_F(\alpha)) - 1) + 2(e'_F(\alpha))^2 f(e_F(\alpha)) \xrightarrow{\alpha \rightarrow \frac{1}{2}} 8\delta_F((\gamma_T(F))^2 + \delta_F f(\mu_F)) \end{aligned}$$

and

$$\begin{aligned} v(\alpha) &= (1 - \alpha)F(e_F(\alpha)) + \alpha(1 - F(e_F(\alpha))) \xrightarrow{\alpha \rightarrow \frac{1}{2}} \frac{1}{2}, \\ v'(\alpha) &= (1 - 2F(e_F(\alpha))) + (1 - 2\alpha)f(e_F(\alpha))e'_F(\alpha) \xrightarrow{\alpha \rightarrow \frac{1}{2}} -\gamma_{EL}(F), \\ v''(\alpha) &= (1 - 2\alpha)f'(e_F(\alpha))(e'_F(\alpha))^2 + (1 - 2\alpha)f(e_F(\alpha))e''_F(\alpha) - 4f(e_F(\alpha))e'_F(\alpha) \\ &\xrightarrow{\alpha \rightarrow \frac{1}{2}} -8f(\mu_F)\delta_F \end{aligned}$$

(see also Eberl and Klar 2022). Hence,

$$\begin{aligned} e'''_F(\frac{1}{2}) &= 8[2\delta_F((\gamma_{EL}(F))^2 + \delta_F f(\mu_F)) + 2\delta_F(\gamma_{EL}(F))^2 + 2\delta_F(\gamma_{EL}(F))^2 + 4f(\mu_F)\delta_F^2] \\ &= 48\delta_F(\delta_F f(\mu_F) + (\gamma_{EL}(F))^2). \end{aligned}$$

By plugging this into (16), we obtain the limiting expectile-based kurtosis measure

$$\begin{aligned} \lim_{\beta \searrow 0} \frac{2}{\beta^2} \cdot \kappa_E^{\frac{1}{2}-3\beta, \frac{1}{2}-\beta}(F) &= \frac{(e_F)'''(\frac{1}{2})}{(e_F)'(\frac{1}{2})} = \frac{48\delta_F(\delta_F f(\mu_F) + (\gamma_{EL}(F))^2)}{2\delta_F} \\ &= 24\left(\delta_F f(\mu_F) + (\gamma_{EL}(F))^2\right). \end{aligned}$$

Disregarding the factor 24, we define the mapping

$$\kappa_{EL} : \mathcal{L}^1 \rightarrow \mathbb{R}, \quad F \mapsto \tau_{EL}(F) \cdot f(\mu_F) + (\gamma_{EL}(F))^2. \tag{17}$$

If we restrict that mapping to symmetric distributions, the second summand vanishes, as then,  $\gamma_T(F) = 0$ . In that case,  $\kappa_{EL}$  satisfies the requirements for a kurtosis measure.

**Theorem 11** *If the mapping  $\kappa_{EL}$  is restricted to the domain  $\mathcal{S}$ , it satisfies (K1) and (K2) for the kurtosis orders  $\leq_3$  and  $\leq_{gs}^0$ .*

**Proof** Let  $F \in \mathcal{S}$ . We start by showing that  $\kappa_{EL}$  satisfies (K1) in two steps. To prove that  $\kappa_{EL}(aX + b) = \kappa_{EL}(X)$  for all  $a > 0$  and  $b \in \mathbb{R}$ , note that  $\tau_{EL}(aX + b) = a\tau_{EL}(X)$  and

$$f_{aX+b}(\mathbb{E}[aX + b]) = \frac{1}{a} f\left(\frac{\mathbb{E}[aX + b] - b}{a}\right) = \frac{1}{a} f(\mu_F).$$

For (K1), it remains to be shown that  $\kappa_{EL}(-X) = \kappa_{EL}(X)$ , which is implied by  $\tau_{EL}(-X) = \tau_{EL}$  and  $f_{-X}(\mathbb{E}[-X]) = f(-\mathbb{E}[-X]) = f(\mu_F)$ .

Now, assume  $F, G \in \mathcal{S}$  such that  $F \leq_{gs}^0 G$ . Since  $\kappa_{EL}$  satisfies (K1), it is invariant under shifts and rescaling, and we may assume without restriction that  $\mu_F = \mu_G = 0$  and  $f(\mu_F) = g(\mu_G)$ . It follows from  $F \leq_{gs}^0 G$  as well as  $F, G \in \mathcal{S}$  that  $R_{FG}$  is concave on  $(-\infty, 0] \cap D_F$  and convex on  $[0, \infty) \cap D_F$ . Thus,  $R'_{FG}(t) = f(t)/g(R_{FG}(t))$ ,  $t \in D_F$ , is decreasing for  $t \leq 0$  and increasing for  $t \geq 0$ . Therefore, the function  $R'_{FG}$  reaches its global minimum  $f(\mu_F)/g(\mu_G) = f(0)/g(0) = 1$  at 0. (Since both  $F$  and  $G$  are symmetric, their medians and means coincide, yielding  $R_{FG}(0) = 0$ .) It follows that  $R'_{FG}(t) \geq 1$  for all  $t \in D_F$ , which is equivalent to  $F \leq_1 G$ . Since  $\tau_{EL}$  is a dispersion measure (see Hürlimann 2002, p. 15),  $\tau_{EL}(F) \leq \tau_{EL}(G)$  holds, concluding the proof.  $\square$

In fact, the result above holds for any mapping  $\kappa : \mathcal{S} \rightarrow \mathbb{R}$ ,  $F \mapsto \tau(F)f(\nu(F))$ , where  $\nu$  is an arbitrary measure of central location and  $\tau$  is an arbitrary measure of dispersion. The second generalization is valid because the fact that  $\tau_{EL}$  is a dispersion measure is the only property of  $\tau_{EL}$  that is utilized in the proof of Theorem 11. The mean can be replaced by any other location measure since all location measures are equal on a symmetric distribution.

We make use of this flexibility by centring the kurtosis measure around the median instead of around the mean. The resulting mapping

$$\kappa_{EM} : \mathcal{L}^1 \rightarrow \mathbb{R}, \quad F \mapsto \mathbb{E}[|X - F^{-1}(\frac{1}{2})|] \cdot f(F^{-1}(\frac{1}{2})), \tag{18}$$

involves the mean absolute deviation around the median and preserves the order  $\leq_3$  under weaker assumptions.

**Theorem 12** *Let  $t \in \mathbb{R}$ . If  $F, G \in \mathcal{T}_{D,1/2}^t$  satisfy  $F \leq_3 G$ , then  $\kappa_{EM}(F) \leq \kappa_{EM}(G)$ .*

**Proof** Since  $\kappa_{EM}$  is invariant to affine linear transformations (for analogous reasons as  $\kappa_{EL}$  on  $\mathcal{S}$ ), we can assume without restriction that  $F^{-1}(\frac{1}{2}) = G^{-1}(\frac{1}{2}) = 0$  and  $f(F^{-1}(\frac{1}{2})) = g(G^{-1}(\frac{1}{2}))$ . According to Proposition 12 in Eberl and Klar (2024),  $1/2$  is a so-called inflection value for all pairs  $F, G \in \mathcal{T}_{D,1/2}^t$  with  $F \leq_3 G$ , meaning that  $R_{FG}$  is concave on  $(-\infty, F^{-1}(1/2)] \cap \text{supp}(F)$  and convex on  $[F^{-1}(1/2), \infty) \cap \text{supp}(F)$ . Noting  $F^{-1}(1/2) = 0$  and that the mapping  $F \mapsto \mathbb{E}[|X - F^{-1}(1/2)|]$  is also a dispersion measure (see Hürlimann 2002, p. 15), the remainder of the proof is analogous to that of Theorem 11.  $\square$

### 4.5 Ratios of dispersion measures as measures of kurtosis

As mentioned in the introduction, one of the first attempts to construct kurtosis measures in a general way was made by Bickel and Lehmann (1975, pp. 469–470), where they announced to define a kurtosis measure as a ‘suitable’ ratio of two (possibly rescaled) dispersion measures. Of the measures we have discussed in Sect. 4, several fit this description.

The most obvious instance is the quantile-based measure  $\kappa_Q^{\alpha,\eta}$ , which is defined as a wider interquantile range divided by a more narrow interquantile range. Any interquantile range is by definition a measure of dispersion. The situation is similar for the expectile kurtosis  $\kappa_E$ , which is equivalent to a ratio of two interexpectile ranges  $\tau_E$ . While any interexpectile range is a dispersion measure (see Eberl and Klar 2023), the ratio could not be shown to preserve the order  $\leq_3$  on any notable subset of distributions.

Another example is the expectile limit measure around the median, which can be rewritten as

$$\kappa_{EM}(F) = \frac{\mathbb{E}[|X - F^{-1}(\frac{1}{2})|]}{\tau_D(F)}.$$

Finally, the moment-based measure  $\kappa_M$  is equivalent to the monotonically transformed

$$\sqrt[4]{\kappa_M(F)} = \frac{\sqrt[4]{\mathbb{E}[(X - \mu_F)^4]}}{\sigma_F}, \quad F \in \mathcal{L}^4.$$

Here, two measures  $\kappa_1$  and  $\kappa_2$  are said to be equivalent, if  $\kappa_1(F) \leq \kappa_1(G)$  is equivalent to  $\kappa_2(F) \leq \kappa_2(G)$  for all cdf’s  $F$  and  $G$ . Since the generalized standard deviation  $\sqrt[2k]{\mathbb{E}[(X - \mu_F)^{2k}]}$  is also a dispersion measure for all  $k \in \mathbb{N}$  (see Sect. 4.1),  $\sqrt[4]{\kappa_M}$  is also a ratio of two dispersion measures. The other kurtosis measures in this work have no obvious representation as a ratio of two dispersion measures. Hence, while the rather vague definition from Bickel and Lehmann (1975) has some merit, it does not seem to coincide with the order-based approach.

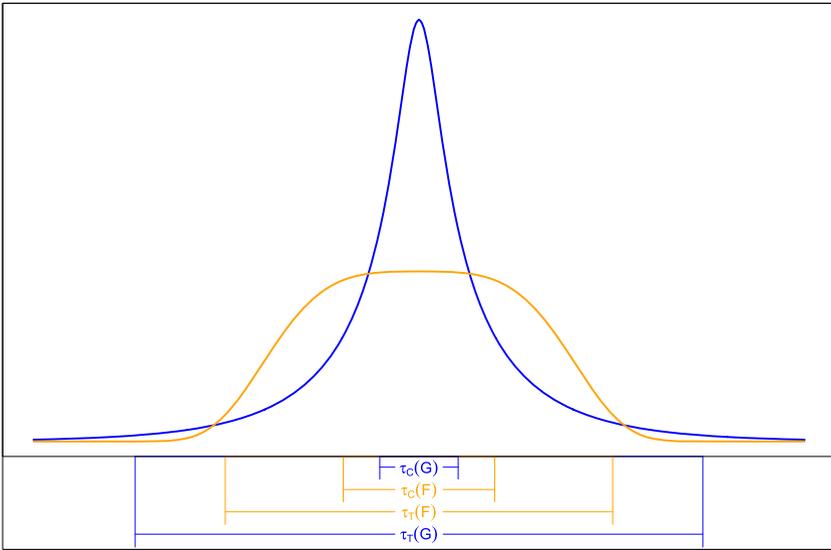
A pattern can be observed for kurtosis measures that are ratios of two dispersion measures. In all cases, the dispersion measure in the numerator puts more emphasis on the tails of the distribution relative to the measure in the denominator, which focuses more on the centre of the distribution. This begs the question as to why these kinds of constructions tend to preserve the order  $\leq_3$ , at least on the subset of symmetric distributions. While the notion of 'putting more emphasis on the tails' seems to be too vague to obtain any general rigorous result explaining this behaviour, the proof of Theorem 11 is fairly instructive for this kind of situation. If  $F, G \in \mathcal{S}$ ,  $F \neq G$  satisfy  $F \leq_{gs}^0 G$ , then the function  $R'_{FG}$  is decreasing up to 0 and increasing from there on. We now assume that the more centre-focused dispersion measures  $\tau_C$  in the denominator satisfies  $\tau_C(F) = \tau_C(G)$ . If  $\tau_C$  not only preserves the dispersive order  $\leq_1$  but also its strict version  $<_1$ , then neither  $R'_{FG} > 1$  nor  $R'_{FG} < 1$  is true. It follows that there exists a  $t_0 \in D_F \cap [0, \infty)$  such that  $R'_{FG}(t) > 1$  holds for  $|t| \geq t_0$ . If the more tail-focused dispersion measure  $\tau_T$  in the numerator is sufficiently similar to  $\tau_C$ , then  $\tau_T(F) \leq \tau_T(G)$  follows from the fact that the dispersions of  $F$  and  $G$  are similar around the centre (since  $\tau_C(F) = \tau_C(G)$ ) and that  $F$  is more dispersed than  $G$  on the tails (since  $R'_{FG}(t) > 1$  for large absolute values of  $t$ ). Overall, the mapping  $F \mapsto \tau_T(F)/\tau_C(F)$  then preserves the concave-convex order  $\leq_{gs}^0$  and therefore also the order  $\leq_3$ .

This similarity in the construction of a number of kurtosis measures strengthens the interpretation of the concept of kurtosis by Balanda and MacGillivray (1988), who describe kurtosis as "the location- and scale-free movement of probability mass from the shoulders of a distribution into its center and tails". The density of a typical symmetric and unimodal distribution that is very kurtotic has a sharp peak in the centre, declines steeply away from it, and has fat tails. Thus, the dispersion of the distribution mostly lies far away from its centre. If a distribution exhibits little kurtosis, the shoulders of its density are very prominent compared to its centre and tails. Here, the dispersion of the distribution is mostly close to the centre. Overall, a centre-focused dispersion measure  $\tau_C$  tends to take larger values for distributions with less kurtosis, and a tail-focused dispersion measure  $\tau_T$  tends to take larger values for more kurtotic distributions. This is illustrated in Fig. 2.

#### 4.6 Numerical comparison of the different kurtosis measures

Since measures of kurtosis are much more prevalent in applied sciences than in the mathematical literature, we briefly compare the numerical values of the measures considered in the previous subsections. For that, we use three families of distributions: Student's  $t$ -distribution, the beta distribution and the sinh-arcsinh distribution. To simplify the notation, we denote  $\kappa_Q = \kappa_Q^{0.05, 0.25}$ ,  $\kappa_D = \kappa_D^{0.5}$  and  $\kappa_E = \kappa_E^{0.05, 0.25}$  throughout. Furthermore, note that  $\kappa_{EL} = \kappa_{EM}$  holds for symmetric distributions.

First, we consider Student's  $t$ -distribution, which is a classical example for a one-parameter family of symmetric distributions that varies in terms of kurtosis. The values of all six kurtosis measures analyzed in this paper are plotted in Fig. 3. Instead of  $\kappa_E$  we consider  $\kappa_E + 3$ , which is structurally equivalent to  $\kappa_Q$  and always takes non-negative values. The different measures all behave similarly with large values and



**Fig. 2** Illustration: the centre-focused dispersion measure  $\tau_C$  is smaller for the more kurtotic distribution  $G$  in blue and larger for the less kurtotic distribution  $F$  in orange. The opposite is the case for the tail-focused dispersion measure  $\tau_T$ . (Here:  $\tau_C = \tau_Q^{0,25}$  and  $\tau_T = \tau_Q^{0,01}$ )

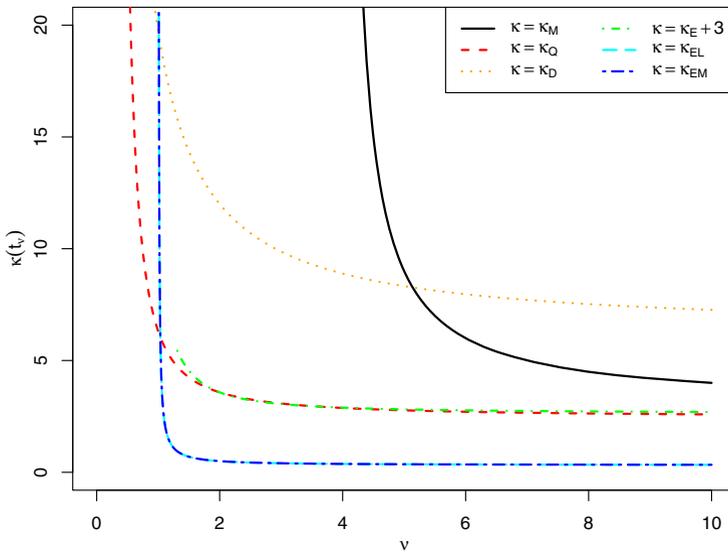
a steep decline for a small number of degrees of freedom. The curves then all flatten out as the parameter increases. The levels at which they flatten out as well as the exact shapes of the curves vary between the different measures. This can partly be explained by the fact that some of them are only defined for sufficiently high degrees of freedom.

Next, we use the beta distribution, which depends on two shape parameters  $\alpha, \beta > 0$ . We consider the symmetric case  $\alpha = \beta$ , as well as  $\alpha = 5\beta$ , i.e. asymmetric, left-skewed distributions. The values of the kurtosis measures in both scenarios are shown in Fig. 4. The curves in the symmetric case behave similarly to those for the  $t$ -distribution as they all have a similar shape and mostly differ in scale and shift. The fact that the distribution becomes more kurtotic for increasing parameter values coincides with its visual perception. For small values of  $\alpha, \kappa_D$  takes negative values because the density function is then left-curved at its center rather than right-curved as it usually is for symmetric unimodal distributions. In the asymmetric case, however, the six kurtosis measures behave very differently: some are increasing, some are decreasing and some are neither. This is because the skewness of the distribution interferes with the measurement of its kurtosis by introducing ambiguity.

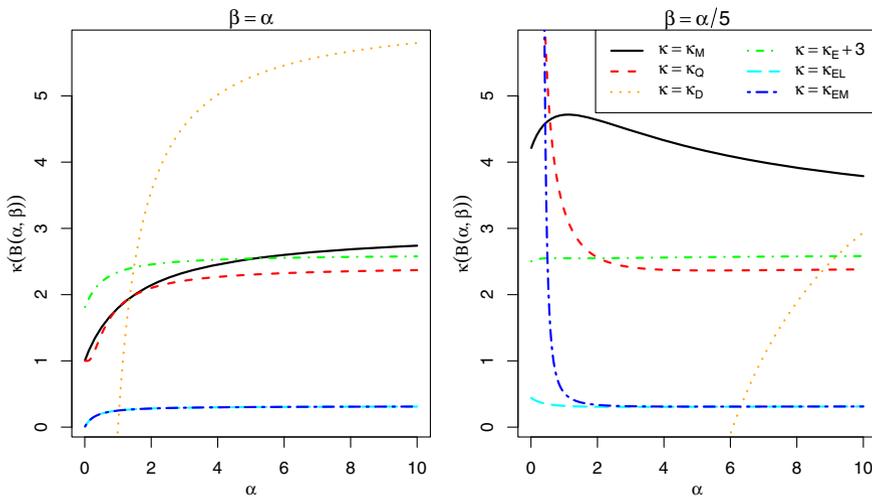
Finally, we consider the sinh-arsinh distribution introduced by Jones and Pewsey (2009). A random variable  $X$  is said to be sinh-arsinh-distributed with skewness parameter  $\nu \in \mathbb{R}$  and (inverse) kurtosis parameter  $\tau > 0$ , if

$$Z = S_{\nu, \tau}(X) = \sinh(\tau \cdot \operatorname{arsinh}(X) - \nu)$$

is standard normal. This distribution family is superior to other skewness-kurtosis-families for several reasons: its density, quantile function and moments can easily be

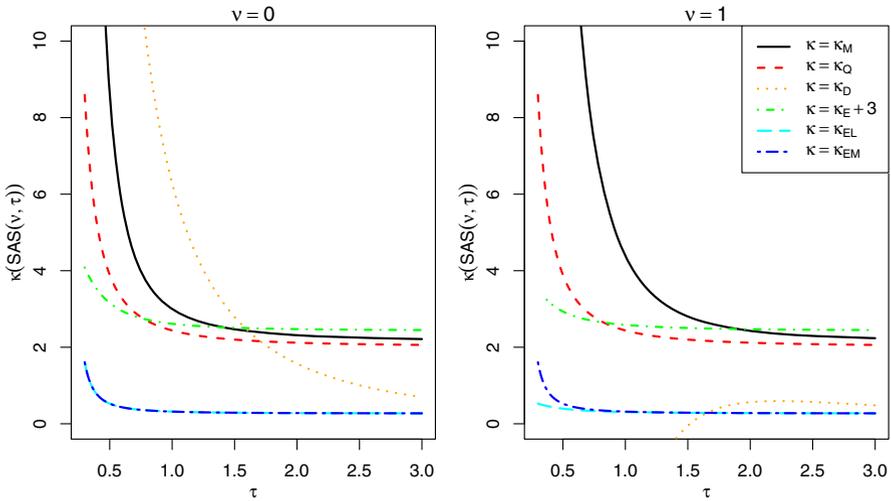


**Fig. 3** Values of the six considered kurtosis measures for Student's  $t_\nu$  with varying degrees of freedom  $\nu > 0$



**Fig. 4** Values of the six considered kurtosis measures the beta distribution  $B(\alpha, \beta)$  with varying shape parameters  $\alpha, \beta > 0$

calculated, it includes the standard normal distributions in the interior of the parameter space, and, most importantly, its kurtosis parameter exhibits strong invariance properties with respect to its skewness parameter (Eberl and Klar 2024). Besides the symmetric case  $\nu = 0$ , we consider the asymmetric, right-skewed case of  $\nu = 1$ . The corresponding values of the kurtosis measures are plotted in Fig. 5. In the symmetric case, all kurtosis measures again behave similarly. As one would expect, all curves



**Fig. 5** Values of the six considered kurtosis measures for the sinh-arcsinh distribution  $SAS(v, \tau)$  with varying skewness parameter  $v \in \mathbb{R}$  and (inverse) kurtosis parameter  $\tau > 0$

are decreasing. While the differences in the shapes of the curves are larger in the asymmetric case, all kurtosis measures except  $\kappa_D$  are still decreasing as a function of  $\tau$ . This behaviour persists if the degree of skewness is increased further, and we conjecture that this is related to the fact that the skewness and kurtosis parameters of this distribution family are not entangled (see Eberl and Klar 2024, Theorem 23). The peculiar behaviour of the density-based kurtosis measure is explained by the fact that it evaluates the distribution only in a local way. For symmetric unimodal distributions, the density evaluated at the median (which then coincides with the mode) is right-curved, resulting in  $\kappa_D > 0$ . However, in the right-skewed case, the median is typically larger than the mode and then lies within a long left-curved area of the density, leading to  $\kappa_D < 0$ . In order to understand what information is actually conveyed by this value of the kurtosis measure, one has to take the values of the corresponding density-based skewness measure  $\gamma_D$  into account in the sense of (11).

### 4.7 Empirical counterparts of the different measures of kurtosis

The asymptotic distribution of the empirical moment kurtosis, i.e. the fourth standardized empirical moment, can be found in Borroni and De Capitani (2022, Theorem 2). In the following, we briefly discuss the empirical counterparts of the remaining kurtosis measures. Details run along the lines of the corresponding skewness measures (Eberl and Klar 2020), and are omitted.

Let  $X_1, \dots, X_n$  be independent and identically distributed random variables with  $X_1 \sim F$ . Further, let  $\hat{q}_p$  denote the empirical  $p$ -quantile. Then, for  $0 < \alpha < \eta < 1/2$ , the obvious plug-in estimator of  $\kappa_Q^{\alpha, \eta}$  in (4) is  $\hat{\kappa}_Q^{\alpha, \eta} = (\hat{q}_{1-\alpha} - \hat{q}_\alpha) / (\hat{q}_{1-\eta} - \hat{q}_\eta)$ , and

standard arguments yield the asymptotic normality

$$\sqrt{n} \left( \hat{\kappa}_Q^{\alpha, \eta} - \kappa_Q^{\alpha, \eta} \right) \xrightarrow{D} N \left( 0, \sigma_{\kappa_Q}^2 \right),$$

where  $\sigma_{\kappa_Q}^2$  depends on  $\alpha$ ,  $\eta$  and  $F$ . Similarly, one can construct plug-in estimators of  $\kappa_{QA}^{\alpha, \eta}$  and  $M$ .

For estimating  $\kappa_D^{1/2}$  in (10), one needs estimates of the density  $f$  and its second derivative. This is commonly achieved by using a kernel density estimator  $\hat{f}_h$  and its second derivative  $\hat{f}_h''$ , where  $h$  denotes the bandwidth. Then, a possible estimator is

$$\hat{\kappa}_D^{1/2} = -\hat{f}_h''(\hat{q}_{1/2}) / (\hat{f}_h(\hat{q}_{1/2})^3).$$

Since the optimal bandwidth for estimation of the  $r$ th derivative is of order  $n^{-1/(2r+5)}$ , one can show that the AMISE converges to zero with rate  $n^{-4/(2r+5)}$  (Scott 1992). Hence, very large sample sizes are necessary compared to the rate  $n^{-1}$  for quantile-based measures to obtain a comparably accurate estimate.

Assuming a finite second moment and using the asymptotics of empirical expectiles (Holzmann and Klar 2016), it can be proved that the estimator  $\hat{\kappa}_E^{\alpha, \eta} = (\hat{e}_{1-\alpha} - 3\hat{e}_{1-\eta} + 3\hat{e}_\eta - \hat{e}_\alpha) / (\hat{e}_{1-\eta} - \hat{e}_\eta)$  of  $\kappa_E^{\alpha, \eta}$  in (16) satisfies

$$\sqrt{n} \left( \hat{\kappa}_E^{\alpha, \eta} - \kappa_E^{\alpha, \eta} \right) \xrightarrow{D} N \left( 0, \sigma_{\kappa_E}^2 \right),$$

where  $\sigma_{\kappa_E}^2$  depends on  $\alpha$ ,  $\eta$  and  $F$ . On the other hand, the plug-in estimator of  $\kappa_{EL}$  in (17), given by

$$\hat{\kappa}_{EL} = \hat{f}_h(\bar{X}) \sum_{j=1}^n |X_j - \bar{X}|/n + (2F_n(\bar{X}) - 1)^2,$$

again suffers from a rate of convergence slower than  $n^{-1}$ , but not to the same extent as  $\hat{\kappa}_D^{1/2}$ . The same remark applies to the plug-in estimator of  $\kappa_{EM}$  in (18). In summary, estimators for all measures of kurtosis are readily available. Estimators involving the estimation of the density or its derivative need larger sample sizes. This concerns  $\kappa_D^{1/2}$  in particular, which is why this measure should only be used for theoretical considerations.

## 5 Conclusion

As shown in this work, kurtosis measures in the traditional sense do not exist, due to the intrinsic entanglement between skewness and kurtosis. This implies that kurtosis cannot be measured in the same way as location, dispersion and skewness, and the same also holds for all higher convex characteristics. One possible solution to this problem is given by the kurtosis functional introduced by Eberl and Klar (2024),

which quantifies the difference in kurtosis between two given distributions. Another approach, which allows the kurtosis of a single distribution to be quantified, is pursued throughout Sect. 4 of this paper by restricting kurtosis measures to transitivity sets of the order  $\leq_3$ . However, most of the results are only valid if the measure is restricted to the specific transitivity set  $\mathcal{S}$  of symmetric distributions. In particular, this is true for the moment- and quantile-based measures, the two most well-known quantities in the literature. The only kurtosis measures that preserve  $\leq_3$  on more general transitivity sets are  $\kappa_D^p$  and  $\kappa_{EM}$ , both of which are computed by evaluating the density of the distribution in question. This makes it difficult to apply these measures to data.

Some of the results on the preservation of  $\leq_3$  by kurtosis measures on transitivity sets also shed some light on how notable differences in skewness influence the behaviour of these measures. An instructive example of this are formulas (11) and (12) on the relationship between the density-based measures of skewness and kurtosis. They can be interpreted as follows: as long as two distributions  $F$  and  $G$  are equally skewed, their kurtosis values are perfectly comparable. However, as the distributions start to differ in skewness, a window of indistinguishability in terms of kurtosis opens up. If the difference of the kurtosis values lies within this window, neither of the two distributions can be pointed out as being more kurtotic as  $F \equiv_3 G$  holds. If the two distributions are differently skewed in the same direction, zero is not even included in this window of indistinguishability. Thus, differences in skewness obscure the view on the comparison in terms of kurtosis.

Although the interaction with skewness is not equally explicit for other kurtosis measures, they do exhibit tendencies pointing in a similar direction. For the quantile-based measure  $\kappa_Q^{\alpha,\eta}$ , the kurtosis order is preserved by respective one-sided quantities that measure kurtosis on either side of the center of the distributions. However, the weights in their weighted average which gives the overall kurtosis measure are dependent on the skewness of the involved distributions. If the margin between the one-sided quantities is not wide enough relative to the skewness difference, no meaningful difference in terms of kurtosis can be established.

That considerable skewness in either direction interferes with the measurement of kurtosis is also represented in the fact that the square of a related skewness measure plays an additive role in a number of kurtosis measures. Obviously, this is the case in the well-known inequality between the moment-based measures of skewness and kurtosis. This is even more explicit in the kurtosis measures  $\kappa_{DA}^p$  and  $\kappa_{EL}$ , where the closely related skewness measure features as a squared summand in the definition.

A goal in future research is to find a general approach to construct kurtosis measures in a way that incorporates the intrinsic influence of skewness. Similarly, a more rigorous foundation underlying the idea of kurtosis measures as ratios of dispersion measures and how this incorporates skewness would be desirable.

**Funding** Open Access funding enabled and organized by Projekt DEAL.

## Declarations

**Conflict of interest** On behalf of all authors, the corresponding author states that there is no conflict of interest.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Balanda KP, MacGillivray HL (1988) Kurtosis: a critical review. *Am Stat* 42:111–119
- Balanda KP, MacGillivray HL (1990) Kurtosis and spread. *Can J Stat* 18:17–30
- Bickel PJ, Lehmann EL (1975) Descriptive statistics for nonparametric models: I: introduction. *Ann Stat* 3:1038–1044
- Blest DC (2003) A new measure of kurtosis adjusted for skewness. *Aust N Zealand J Stat* 45:175–179
- Borroni CG, De Capitani L (2022) Some measures of kurtosis and their inference on large datasets. *AStA Adv Stat Anal* 106:573–607
- Chissom BS (1970) Interpretation of the Kurtosis Statistic. *Am Stat* 24:19–22
- Cooper M (2020) Kurtosis and skew show longer males in centrobolus. *Arthropods* 9:21–26
- Crack T (2022) Foundations for scientific investing: capital markets intuition and critical thinking skills. Timothy F. Crack, 11th ed
- Critchley F, Jones MC (2008) Asymmetry and gradient asymmetry functions: density-based skewness and kurtosis. *Scand J Stat* 35:415–437
- Darlington RB (1970) Is kurtosis really “peakedness?”. *Am Stat* 24:19–22
- Eberl A, Klar B (2020) Asymptotic distributions and performance of empirical skewness measures. *Comput Stat Data Anal* 146:106939
- Eberl A, Klar B (2022) Expectile based measures of skewness. *Scand J Stat* 49:373–399
- Eberl A, Klar B (2023) Stochastic orders and measures of skewness and dispersion based on expectiles. *Stat Papers* 64:509–527
- Eberl A, Klar B (2024) Centre-free kurtosis orderings for asymmetric distributions. *Stat Papers* 65:415–433
- Groeneveld RA, Meeden G (1984) Measuring skewness and kurtosis. *J R Stat Soc Ser D* 33:391–399
- Hanook S, Shahbaz MQ, Mohsin M, Kibria BMG (2013) A note on beta inverse-weibull distribution. *Commun Stat Theory Methods* 42:320–335
- Holzmann H, Klar B (2016) Expectile asymptotics. *Electr J Stat* 10:2355–2371
- Hosking JRM (1989) Some theoretical results concerning l-moments. Research Report RC14492. IBM Research
- Hürlimann W (2002) On risk and price: stochastic orderings and measures. In: Transactions of the 27th international congress of actuaries
- Jones MC, Pewsey A (2009) Sinh-arcsinh distributions. *Biometrika* 96:761–780
- Jones MC, Rosco JF, Pewsey A (2011) Skewness-invariant measures of kurtosis. *Am Stat* 65:89–95
- López-Martín C, Arguedas-Sanz R, Muela SB (2022) A cryptocurrency empirical study focused on evaluating their distribution functions. *Int Rev Econ Financ* 79:387–407
- MacGillivray HL, Balanda KP (1988) The relationship between skewness and kurtosis. *Aust J Stat* 30:319–337
- McAlevy LG, Stent AF (2018) Kurtosis: a forgotten moment. *Int J Math Educ Sci Technol* 49:120–130
- Moors JJA (1988) A quantile alternative for kurtosis. *J R Stat Soc Ser D* 37:25–32
- Müller A, Stoyan D (2002) Comparison methods for stochastic models and risks. Wiley Series in Probability and Statistics, Chichester
- Nørlund NE (1926) Leçons sur les Séries d'Interpolation. Gauthier-Villars, Collection de Monographies sur la Théorie des Fonctions
- Oja H (1981) On location, scale, skewness and kurtosis of univariate distributions. *Scand J Stat* 8:154–168
- Pearson K (1895) Contributions to the mathematical theory of evolution: II: skew variation in homogeneous material. *Philosoph Trans R Soc London A* 186:343–414

- Pearson K (1916) Mathematical contributions to the theory of evolution: XIX—second supplement to a memoir on skew variation. *Philosoph Trans R Soc London A* 216:429–457
- Ruppert D (1987) What is kurtosis? An influence function approach. *Am Stat* 41:1–5
- Scott DW (1992) *Multivariate density estimation*. Wiley, New York, Theory, Practice and Visualization
- van Zwet WR (1964) *Convex transformations of random variables*. Mathematical Centre Tracts, Mathematisch Centrum Amsterdam
- Wheeler DJ (2014) An approximation for simulation of gamma distributions. *J Stat Comput Simul* 3:225–232

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.