

Baringhaus, Ludwig; Grübel, Rudolf

Article — Published Version

## The quarter median

Metrika

**Provided in Cooperation with:**

Springer Nature

*Suggested Citation:* Baringhaus, Ludwig; Grübel, Rudolf (2021) : The quarter median, Metrika, ISSN 1435-926X, Springer, Berlin, Heidelberg, Vol. 85, Iss. 4, pp. 419-458, <https://doi.org/10.1007/s00184-021-00836-z>

This Version is available at:

<https://hdl.handle.net/10419/286765>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*



<https://creativecommons.org/licenses/by/4.0/>



# The quarter median

Ludwig Baringhaus<sup>1</sup> · Rudolf Grübel<sup>1</sup>

Received: 12 January 2021 / Accepted: 28 July 2021 / Published online: 8 August 2021  
© The Author(s) 2021

## Abstract

We introduce and discuss a multivariate version of the classical median that is based on an equipartition property with respect to quarter spaces. These arise as pairwise intersections of the half-spaces associated with the coordinate hyperplanes of an orthogonal basis. We obtain results on existence, equivariance, and asymptotic normality.

**Keywords** Asymptotic normality · Consistency · Estimation of location · Euclidean motion equivariance · Equipartition · Multivariate median

**Mathematics Subject Classification** 62H12 · 62H11 · 62F35

## 1 Introduction

The classical one-dimensional median associated with a probability distribution  $P$  on the Borel subsets of the real line  $\mathbb{R}$  is any value  $\theta$  that satisfies the inequalities

$$P((-\infty, \theta]) \geq 1/2, \quad P([\theta, \infty)) \geq 1/2. \quad (1)$$

Under certain conditions, for example if  $P$  has a positive density, we have equalities in (1) and the median is unique. In general, the set of solutions is the bounded closed interval  $[\text{MED}_-(P), \text{MED}_+(P)]$  with

$$\begin{aligned} \text{MED}_-(P) &:= \inf\{x \in \mathbb{R} : P((-\infty, x]) \geq 1/2\}, \\ \text{MED}_+(P) &:= \sup\{x \in \mathbb{R} : P([x, \infty)) \geq 1/2\} \end{aligned}$$

---

✉ Ludwig Baringhaus  
lbaring@stochastik.uni-hannover.de  
Rudolf Grübel  
rgrubel@stochastik.uni-hannover.de

<sup>1</sup> Institute of Actuarial and Financial Mathematics, Leibniz Universität Hannover, Postfach 6009, 30060 Hannover, Germany

as the smallest and the largest median of  $P$ . Often, ‘the’ median of  $P$  is defined as the corresponding midpoint,

$$\text{MED}_m(P) := (\text{MED}_-(P) + \text{MED}_+(P))/2. \tag{2}$$

For a data set  $x_1, \dots, x_n$  of real numbers a (sample) median is any (distributional) median associated with the empirical distribution  $P_n := n^{-1} \sum_{k=1}^n \delta_{x_k}$ , where  $\delta_x$  denotes unit mass at  $x$ . With  $x_{(1)} \leq \dots \leq x_{(n)}$  as the values obtained by arranging  $x_1, \dots, x_n$  in increasing order,

$$\text{MED}_m(P_n) = \text{MED}_-(P_n) = \text{MED}_+(P_n) = x_{((n+1)/2)}$$

is the unique sample median if  $n$  is odd, and we get  $[x_{(n/2)}, x_{(n/2+1)}]$  as the set of sample medians if  $n$  is even.

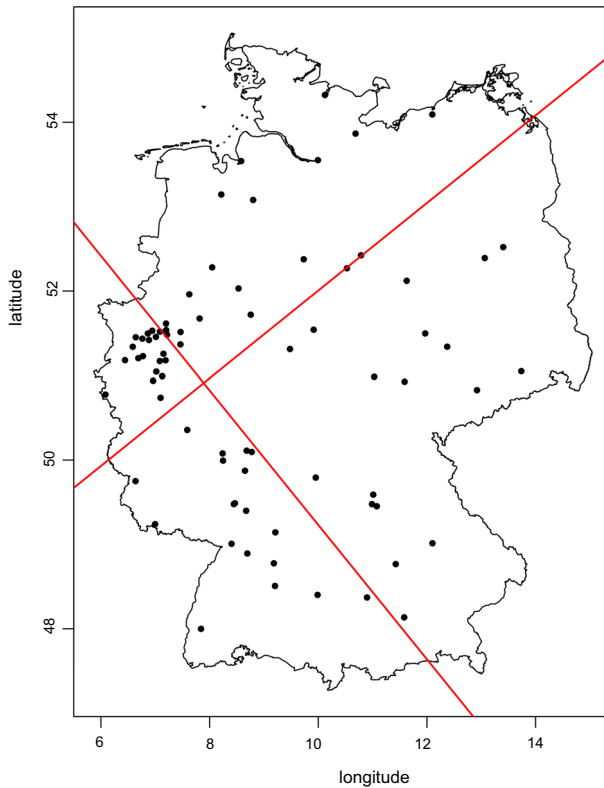
Historically, the first attempt to generalize this to dimensions  $d > 1$  was to use coordinates: If  $x_1, \dots, x_n$  is a subset of  $\mathbb{R}^d$  for some  $d \geq 2$  then a coordinatewise or *marginal median* of the empirical distribution  $P_n := n^{-1} \sum_{k=1}^n \delta_{x_k}$  is any vector whose  $i$ th coordinate is a (one-dimensional) median of the respective  $i$ th coordinates of the data vectors. This is the sample version; for a probability measure  $P$  on the Borel subsets  $\mathcal{B}^d$  of  $\mathbb{R}^d$  the push-forwards  $P^{\pi_i}$  of  $P$  under the coordinate projections  $\pi_i, i = 1, \dots, d$ , would be used. Formally, the marginal medians of  $P$  are elements of the  $d$ -dimensional interval  $[\text{MMED}_-(P), \text{MMED}_+(P)]$  where  $\text{MMED}_\pm(P)$  have  $\text{MED}_\pm(P^{\pi_i})$  as their  $i$ th component,  $i = 1, \dots, d$ , and

$$\text{MMED}_m(P) := (\text{MMED}_-(P) + \text{MMED}_+(P))/2 \tag{3}$$

is the midpoint marginal median. It is well known that such component-wise extensions of the one-dimensional median are not equivariant with respect to orthogonal transformations and thus depend on the coordinate system chosen for the representation of the data. One way to repair this is by ‘averaging out’ (symmetrization), see Grübel (1996).

Other generalizations of the one-dimensional median to higher dimensions start with some characterizing property of the classical version and then use a multidimensional analogue. For example, it is known that in the one-dimensional case any median  $\theta$  minimizes the function  $x \mapsto \int |y - x| P(dy)$ , which leads to the *spatial* or  $L^1$ -*median* of the data set, defined as a minimizer of  $x \mapsto \sum_{k=1}^n \|x - x_k\|$ , where  $\|\cdot\|$  denotes Euclidean distance. As this distance is invariant under orthogonal transformation, the spatial median does not depend on the chosen coordinates.

Closer to the above equipartition property are the various concepts related to data depth, where for each  $x \in \mathbb{R}^d$  and each affine hyperplane  $H$  with  $x \in H$  the number of data points to one side of  $H$  is the basic ingredient: A centerpoint median is defined as any such  $x$  with the property that the minimum number is at least a fraction  $n/(d + 1)$  of the data, and Tukey’s median is a point that maximizes these data depth values; see e.g. Donoho and Gasko (1992). This whole area has attracted the attention of many researchers, often with emphasis on robustness. The literature on multivariate



**Fig. 1** The city data (see text for details)

medians up to about 1990 is covered in the review of Small (1990), which includes an interesting discussion of the history of the concept. For more recent reviews of this subject we refer to the treatises of Oja (2013) and Dhar and Chaudhuri (2011).

In the present paper, following an idea of the late Dietrich Morgenstern, we again start with the marginal median as in the work of Grübel (1996), but then use a variant of the equipartition property instead of symmetrization. In dimension  $d = 2$ , for example, a coordinate system specifies four quarter spaces via the pairwise intersections of the half-spaces underlying the marginal median, and a quarter median of  $P$  may be defined as any vector  $\theta \in \mathbb{R}^2$  such that for some coordinate system centered at  $\theta$  all four of the respective quarter spaces have probability at least  $1/4$ . Figure 1 shows a simple example that Professor Morgenstern would have liked: The data are the locations of the cities in Germany with a population of at least 100,000 (in 2017). We have  $n = 76$  such cities in total, so we can aim for a partition with 19 cities in each open quarter space. One resulting quarter median is nearby the exit Freudenberg of the A45 (Sauerlandlinie), and with the North–South axis tilted by about  $39^\circ$  counterclockwise; see the lines in Fig. 1. The data (longitude/latitude) for drawing the German border are extracted from the file `gadm36_DEU_0_sp.rds` available at <https://gadm.org/data.html>.

Investigation of this concept leads to a number of interesting questions, starting with the obvious ones concerning existence and uniqueness. Next, what are the equivariance properties of the quarter median? Further, in contrast to the spatial median a quarter median comes with a system of associated hyperplanes, which may be regarded as a part of the estimator. In the two-dimensional case these are just two orthogonal lines and can be parametrized by a counterclockwise rotation of the horizontal axis, thus providing a connection to a real interval. In higher dimensions, however, specification of a topological notion of nearness of systems of coordinate hyperplanes is less straightforward, but it is needed in connection with consistency of the estimators. For asymptotic normality even a local linearization is required. Finally, there is the algorithmic issue: How do we find a quarter median for a given data set?

Section 2 contains our main results, on existence, equivariance, and large sample behavior. Elliptical distributions are treated in some detail and we compare the quarter median with other estimators for several specific families. We also consider algorithmic aspects together with the problem of measurable selection from the respective solution sets. Section 3 presents numerical comparisons with existing multivariate medians. Proofs are collected in Sect. 4.

## 2 Results

We first assemble some formal definitions and then give our results in separate subsections.

### 2.1 Preliminaries and basic definitions

We regard vectors  $x \in \mathbb{R}^d$  as column vectors and write  $\langle x, y \rangle$  for the inner product of  $x, y \in \mathbb{R}^d$ . The transpose of the vector  $x$  is  $x^t$ , so that  $\langle x, y \rangle = x^t y$ . Similarly,  $A^t$  denotes the transpose of a matrix  $A$ . We write  $\|\cdot\|$  for the Euclidean norm on  $\mathbb{R}^d$ ,  $\mathbb{S}_{d-1}$  is the unit sphere in  $\mathbb{R}^d$ ,  $\mathbb{O}(d)$  is the group of orthogonal  $d \times d$ -matrices, and  $\mathbb{SO}(d)$  denotes the group of rotations, consisting of those elements of  $\mathbb{O}(d)$  that have determinant 1.

If  $(\Omega, \mathcal{A}, P)$  is a probability space, if  $(\Omega', \mathcal{A}')$  is a measurable space, and if the mapping  $S: \Omega \rightarrow \Omega'$  is  $(\mathcal{A}, \mathcal{A}')$ -measurable then we write  $P^S$  for the push-forward of  $P$  under  $S$ , i.e.  $P^S(A') = P(S^{-1}(A'))$  for all  $A' \in \mathcal{A}'$ . If  $(\Omega, \mathcal{A}) = (\mathbb{R}^d, \mathcal{B}^d)$  and  $b \in \mathbb{S}_{d-1}$  or  $U \in \mathbb{O}(d)$  then  $P^b$  and  $P^U$  respectively refer to the push-forwards of  $P$  under the mappings  $x \mapsto \langle b, x \rangle$  and  $x \mapsto Ux$ .

For  $a \in \mathbb{R}^d$  and  $b \in \mathbb{S}_{d-1}$  we formally define the associated half-spaces by

$$H_-(a; b) := \{x \in \mathbb{R}^d : \langle b, x - a \rangle \leq 0\}, \quad H_+(a; b) := \{x \in \mathbb{R}^d : \langle b, x - a \rangle \geq 0\}.$$

Let  $e_1, \dots, e_d$  be the vectors of the canonical basis of  $\mathbb{R}^d$ . Then a marginal median for a probability measure  $P$  on  $(\mathbb{R}^d, \mathcal{B}^d)$  is any vector  $\theta \in \mathbb{R}^d$  with the property that

$$P(H_+(\theta; e_i)) \geq 1/2, \quad P(H_-(\theta; e_i)) \geq 1/2, \quad 1 \leq i \leq d. \quad (4)$$

For two orthogonal unit vectors  $b, b' \in \mathbb{S}_{d-1}$  and  $a \in \mathbb{R}^d$  the quarter spaces are the intersections of these half-spaces,

$$V_{\pm\pm}(a; b, b') := H_{\pm}(a; b) \cap H_{\pm}(a; b'),$$

with the obvious notational convention for the four possible combinations of plus and minus signs. We then define a *quarter median* of a probability measure  $P$  on  $(\mathbb{R}^d, \mathcal{B}^d)$  as any vector  $\theta \in \mathbb{R}^d$  with the property that, for some orthonormal basis  $B = \{b_1, \dots, b_d\}$  of  $\mathbb{R}^d$ ,

$$P(H_+(\theta; b_i)) \geq 1/2, \quad P(H_-(\theta; b_i)) \geq 1/2, \quad 1 \leq i \leq d, \tag{5}$$

and

$$P(V_{\pm\pm}(\theta; b_i, b_j)) \geq 1/4, \quad 1 \leq i < j \leq d, \tag{6}$$

for all combinations of plus and minus signs. If (5) and (6) hold then we call the pair  $(\theta, U) \in \mathbb{R}^d \times \mathbb{O}(d)$ , where  $U$  has rows  $b_1^t, \dots, b_d^t$ , a solution of the *quarter median problem* for the probability measure  $P$ . Again, the data versions of these notions refer to the empirical distribution  $P_n$ .

If  $P$  is absolutely continuous then hyperplanes have zero probability and (5) follows from (6). In general, this is not the case as the following simple example shows: Let  $d = 2$  and

$$P = \frac{1}{3}(\delta_{(0,0)^t} + \delta_{(1/2,1)^t} + \delta_{(1,1/2)^t}).$$

Then (6) holds with  $b_1 = e_1, b_2 = e_2$  and  $\theta = (1, 1/2)^t$ , but (5) is not true. Further, with this choice of  $b_1, b_2$  it is easily verified that the unique marginal median  $\theta = (1/2, 1/2)^t$  is a quarter median. This example also shows that the quarter median may not be unique. Further, if hyperplanes have probability zero, then the inequalities in (5) and (6) can be replaced by equalities, and with  $(\theta, U)$  as solution of (6) we then speak of a solution of the *equipartition problem* for  $P$ .

### 2.2 Existence and uniqueness

The existence of a median and thus of a marginal median follow immediately from the monotonicity of distribution functions. For the quarter median counting the number of variables and the number of constraints may give a first impression. For simplicity, we temporarily assume that hyperplanes have probability zero, so that we have an equipartition problem. If  $d = 2$ , there are then three unknowns, two for the location parameter and one for the angle of rotation, see also Remark 1 (b) below, and one of the four constraints is redundant, so that the number of unknowns is the same as the number of equations. For  $d > 2$  we first note that the conditions in (5) and (6) are not independent. Indeed, for a given basis  $\{b_1, \dots, b_d\}$  of  $\mathbb{R}^d$  it is enough to have probability  $1/2$  for  $H_+(\theta, b_i), i = 1, \dots, d$ , and probability  $1/4$  for  $V_{++}(\theta; b_i, b_j)$  for  $1 \leq i < j \leq d$ , so that  $d + d(d - 1)/2$  conditions remain. On the other hand, we have  $d(d - 1)/2$  parameters for the (special) orthogonal group, and there are  $d$  components of the location vector, hence the number of unknowns is again the same

as the number of independent constraints. This, of course, is just a heuristic argument, as the respective conditions are nonlinear.

Using a general tool from algebraic topology, the homotopy invariance of the Brouwer degree of continuous mappings, Makeev (2007) solved the equipartition problem under the additional condition that  $P$  is absolutely continuous. A smoothing argument leads to the extension needed here.

**Theorem 1** *A solution  $(\theta, U)$  of the quarter median problem exists for every probability measure  $P$  on  $(\mathbb{R}^d, \mathcal{B}^d)$ .*

This result raises an obvious question: Is there a stronger equipartition in dimension  $d > 2$ , such as an ‘octomedian’ if  $d = 3$ , with each of the octants receiving the same probability? This is certainly possible for specific distributions, such as the multivariate standard normal. However, if  $d = 3$  then we have three unknowns for the location vector, and three for the rotation, which can be parametrized by the Euler angles, and the condition that the octants all have probability  $1/8$  would similarly lead to seven constraints, which is one too many. Obviously, the difference would be even greater for dimension  $d > 3$  if all  $2^d$  intersections of coordinate half-spaces were to have the same probability. According to Makeev (2007, p. 554) this dimension counting argument already implies that a solution does not exist for a generic distribution.

The notion of uniqueness requires some attention. We say that a distribution  $P$  has a unique quarter median if  $\theta_1 = \theta_2$  for any two solutions  $(\theta_1, U_1)$  and  $(\theta_2, U_2)$  of the quarter median problem. Obviously, the full pair  $(\theta, U)$  will never be unique: Any linear hyperplane, i.e. a subspace  $H$  of  $\mathbb{R}^d$  of dimension  $d - 1$ , may be specified as the orthogonal complement of the one-dimensional space (line) generated by a unit vector  $b \in \mathbb{S}_{d-1}$  via  $H(b) := \{x \in \mathbb{R}^d : b^t x = 0\}$ , but obviously  $-b$  would lead to the same hyperplane. This simple observation already implies that we may restrict  $U$  to be an element of  $\mathbb{S}\mathbb{O}(d)$  when searching for a quarter median. Further, the conditions (5) and (6) are invariant under permutations of the rows of  $U$ . Conversely, given a set of coordinate hyperplanes, a corresponding basis is a set with elements that are unique up to a factor  $-1$ . Putting this together we write  $U \sim V$  for  $U, V \in \mathbb{O}(d)$  if  $U$  and  $V$  lead to the same set of linear hyperplanes; we then say that the solution of the quarter median problem is unique if for any two solutions  $(\theta_1, U_1)$  and  $(\theta_2, U_2)$  we have  $\theta_1 = \theta_2$  and  $U_1 \sim U_2$ .

The following formal approach will turn out to be useful. Let  $G_1$  be the subgroup of  $\mathbb{O}(d)$  that consists of the diagonal matrices with diagonal entries from  $\{-1, +1\}$ , i.e. the subgroup that represents the compositions of reflections at the coordinate axes, and let  $G_2$  be the subgroup of permutation matrices, which corresponds to coordinate permutations. We write  $G$  for the subgroup generated by  $G_1$  and  $G_2$ . This is the system of all  $d \times d$ -matrices  $A$  with entries  $a_{ij} \in \{-1, 0, 1\}$  and such that for some permutation  $\pi$  of  $\{1, \dots, d\}$  we have  $a_{ij} \neq 0$  if and only if  $j = \pi(i)$ . The above transition from equality to equivalence can be then regarded as a transition from the group  $\mathbb{O}(d)$  to the factor group

$$\mathbb{H}(d) := \mathbb{O}(d)/G, \quad (7)$$

with the equivalence classes  $[U] := \{V \in \mathbb{O}(d) : V \sim U\}$ ,  $U \in \mathbb{O}(d)$ , as elements. If we think of a solution of the quarter median problem as a pair  $(\theta, H) \in \mathbb{R}^d \times \mathbb{H}(d)$ , then uniqueness means that two solutions are equal.

- Remark 1** (a) The  $\theta$ -part of a solution  $(\theta, H)$  for the quarter median problem may be unique while the  $H$ -part is not. For example, if  $P$  has a density that depends on  $x \in \mathbb{R}^d$  only through  $\|x\|$  then  $0$  is the unique quarter median, but all  $H \in \mathbb{H}(d)$  lead to a solution  $(0, H)$ .
- (b) In dimension  $d = 2$  there is an (almost) canonical choice of an element  $U \in \mathbb{O}(d)$  from an equivalence class  $H \in \mathbb{H}(d)$  corresponding to a set of hyperplanes, here just two orthogonal lines  $L_1, L_2$  meeting at the point  $0 \in \mathbb{R}^2$ . With a suitable half-open interval  $I$  of length  $\pi/2$ , such as  $I = [-\pi/4, \pi/4)$ , there is exactly one  $\alpha \in I$  such that the lines are represented by

$$U_\alpha := \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix}, \tag{8}$$

which is the counterclockwise rotation of the canonical coordinate lines by the angle  $\alpha$ , in the sense that  $L_i = \{x \in \mathbb{R}^2 : b_i^t x = 0\}$ , where  $b_1^t, b_2^t$  are the rows of  $U$ . The corresponding topology would then be that of the quotient space  $\mathbb{R}/I$ .

- (c) While we work with the space  $\mathbb{R}^d$  of column vectors throughout the notions of quarter median and quarter median problem make sense for an arbitrary finite-dimensional Euclidean space  $(E, \langle \cdot, \cdot \rangle)$ . If  $(\theta, U)$  solves the quarter median problem for a distribution  $P$  on  $(\mathbb{R}^d, \mathcal{B}^d)$ , then  $\{b_2, \dots, b_d\}$  is an orthonormal basis for the linear space  $E = H(b_1)$ , which has dimension  $d - 1$ . Here  $b_i^t$  denotes the  $i$ th row of  $U$ ,  $i = 1, \dots, d$ ; equivalently,  $b_i$  is the  $i$ th column of  $U^t$ . If  $d \geq 3$ , this can be used to obtain a projectivity or consistency property that relates  $(\theta, U)$  to a solution  $(\theta_E, U_E)$  of the quarter median problem for the push-forward of  $P$  under the orthogonal projection  $\pi_E : \mathbb{R}^d \rightarrow E$ . In fact, let  $\mathbb{S}_{d-1}^E = \{y \in E : \|y\| = 1\}$  be the unit sphere in  $E$ . For  $c \in E$  and orthogonal unit vectors  $d, d' \in \mathbb{S}_{d-1}^E$  the associated half-spaces and quarter spaces are given by

$$H_-^E(c; d) := \{y \in E : \langle d, y - c \rangle \leq 0\}, \quad H_+^E(c; d) := \{y \in E : \langle d, y - c \rangle \geq 0\},$$

and

$$V_{\pm\pm}^E(c; d, d') := H_{\pm}^E(c; d) \cap H_{\pm}^E(c; d').$$

Using that a vector  $x \in \mathbb{R}^d$  can uniquely be written as  $x = \sum_{i=1}^d \langle x, b_i \rangle b_i$  we see that the orthogonal projection of  $x$  in  $E$  is  $x_E := \pi_E(x) = \sum_{i=2}^d \langle x, b_i \rangle b_i$ . It follows from  $\pi_E^{-1}(H_{\pm}^E(\theta_E; b_i)) = H_{\pm}(\theta; b_i)$  for  $2 \leq i \leq d$ , and  $\pi_E^{-1}(V_{\pm\pm}^E(\theta_E; b_i, b_j)) = V_{\pm\pm}(\theta; b_i, b_j)$  for  $2 \leq i < j \leq d$ , that  $P^{\pi_E}(H_{\pm}^E(\theta_E; b_i)) = P(H_{\pm}(\theta; b_i))$  for  $2 \leq i \leq d$ , and that  $P^{\pi_E}(V_{\pm\pm}^E(\theta_E; b_i, b_j)) = P(V_{\pm\pm}(\theta; b_i, b_j))$  for  $2 \leq i < j \leq d$ .



### 2.3 Equivariance

In the context of equivariance properties of location estimators for  $d$ -dimensional data we start with a family (often a group)  $\mathcal{S}$  of measurable transformations  $S : \mathbb{R}^d \rightarrow \mathbb{R}^d$  and we regard the location parameter  $\theta \in \mathbb{R}^d$  as a function  $\theta = T(P)$  of  $P$ . Recall that  $P^S$  is the push-forward of  $P$  under the transformation  $S$ . Then  $T$  is said to be  $\mathcal{S}$ -equivariant if

$$T(P^S) = (S \circ T)(P) \quad \text{for all } S \in \mathcal{S}.$$

In cases where the parameter is not unique this might better be expressed as  $S(\theta)$  being a parameter of this type for  $P^S$  if  $\theta$  is for  $P$ .

For example, with  $S$  being one of the groups  $G_1$  and  $G_2$  that were used in connection with (7), it is easy to check that

$$S^{-1}(\text{MMED}_m(P^S)) = \text{MMED}_m(P) \quad \text{for all } S \in \mathcal{S}, \tag{9}$$

i.e. the specific marginal median introduced in (3) is equivariant with respect to reflections and coordinate permutations. It is easily seen that we may even take  $S = G$ . This specific marginal median is also equivariant with respect to separate rescalings, where the corresponding group is the set of all regular diagonal matrices.

The general idea is of course that statistical inference should respect the structure of the data space. Equivariance of a location estimator with respect to shifts and orthogonal linear transformations means that the estimator interacts in this way with the full isometry group of Euclidean space.

At various stages the following function will be important,

$$\Psi(U, P) := U^t \text{MMED}_m(P^U), \tag{10}$$

with  $U \in \mathbb{O}(d)$  and  $P$  a probability measure on  $(\mathbb{R}^d, \mathcal{B}^d)$ ;  $\Psi(U, P)$  is the specific marginal median in the coordinate system associated with  $U$ , transformed back to canonical coordinates. By definition, a quarter median is always a marginal median in a specific coordinate system. Below, marginal medians in

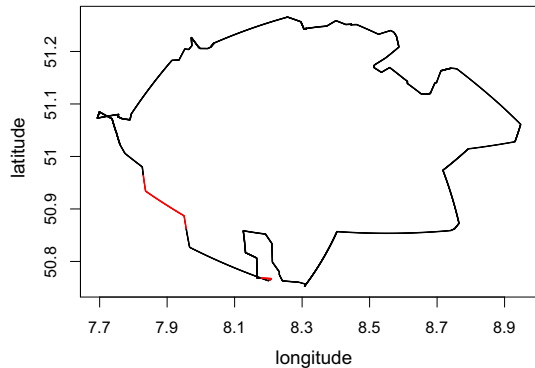
$$\mathcal{M}(P) := \bigtimes_{i=1}^d \left\{ \text{MED}_-(P^{\pi_i}), \text{MED}_m(P^{\pi_i}), \text{MED}_+(P^{\pi_i}) \right\} \tag{11}$$

will be of particular interest.

**Proposition 1** *Let  $P$  be a probability distribution on  $(\mathbb{R}^d, \mathcal{B}^d)$ , and let  $\mathcal{S}$  be the set of Euclidean motions.*

- (a) *If  $\theta$  is a quarter median for the distribution  $P$  then, for all  $S \in \mathcal{S}$ ,  $S(\theta)$  is a quarter median for  $P^S$ .*
- (b) *If  $U, V \in \mathbb{O}(d)$  are such that  $U \sim V$  then  $\Psi(U, P) = \Psi(V, P)$ .*
- (c) *If  $(\theta, U)$  solves the quarter median problem for  $P$ , then there is a marginal median  $\eta \in \mathcal{M}(P^U)$  such that  $(\tilde{\theta}, U)$  with  $\tilde{\theta} = U^t \eta$  is also a solution.*

**Fig. 2** The median curve for the city data



(d) If  $(\theta, U)$  solves the quarter median problem for  $P$  and  $S \in \mathcal{S}$ ,  $S(x) = Ax + b$ , then  $(A\theta + b, UA^t)$  solves the quarter median problem for  $P^S$ . Moreover, if  $P^b$  has a unique median for all  $b \in \mathbb{S}_{d-1}$ ,

$$\Psi(UA^t, P^S) = S(\Psi(U, P)). \tag{12}$$

(e) Suppose that  $P^b$  has a unique median for all  $b \in \mathbb{S}_{d-1}$ . Then, the function  $\Psi(\cdot, P) : \mathbb{O}(d) \rightarrow \mathbb{R}^d$  is continuous.

Part (a) is the equivariance on the set level. Part (b) shows that  $\Psi(\cdot, P)$  may be regarded as a function on  $\mathbb{H}(d)$ . Part (c) is of importance when developing an algorithm for searching a solution of empirical quarter median problems: For each  $U \in \mathbb{O}(d)$  the set of potential  $\theta$ -values in (11) consists of three vectors only, separately for each coordinate. It is tempting to think that this could be reduced further to the respective midpoint, but the uniform distribution on the six points  $(-2, -2)^t, (-1, 3)^t, (1, -1)^t, (2, 2)^t, (3, 4)^t, (4, 0)^t$  in the plane provides a counterexample: The pair  $((2, 2)^t, U)$  with  $U = \text{diag}(1, 1)$  solves the quarter median problem for this distribution, but the midpoint of the associated rectangle  $[1, 2] \times [0, 2]$  of multivariate medians would lead to  $((1.5, 1)^t, U)$ , which is not a solution. Part (d) notices the Euclidean motion equivariance of the quarter median for distributions  $P$  with the property that  $P^b$  has a unique median for all  $b \in \mathbb{S}_{d-1}$ , a condition that will be used repeatedly below.

Similar to the spatial median and the orthomedian introduced by Grübel (1996), the quarter median is not affine equivariant. In particular, it is not equivariant with respect to separate rescalings of the coordinates, in contrast to the special marginal median, but for distributions  $P$  with the property that  $P^b$  has a unique median for all  $b \in \mathbb{S}_{d-1}$ , equivariance does hold (and is used in the proofs below) for simultaneous rescalings, where the group  $\mathcal{S}$  consists of all positive scalar multiples of the identity matrix.

Any quarter median is a marginal median in a suitably chosen coordinate system, which we may take to be generated by an orthogonal matrix with determinant 1. Grübel (1996) introduced the orthomedian of  $P$  as the integral of the function  $\Psi(\cdot, P)$  over  $\mathbb{SO}(d)$  with respect to the Haar measure with total mass 1. Thus, the orthomedian may be seen as the expected value of this function with  $U$  chosen uniformly at random,

and its orthogonal equivariance is then an immediate consequence of the fact that the Haar measure is invariant under the group operations. In contrast, the quarter median picks a marginal median from the range of the set-valued function

$$U \mapsto \left\{ U^t \eta : \eta \in [\text{MMED}_-(P^U), \text{MMED}_+(P^U)] \right\}$$

via the partition constraint (6). Figure 2 shows the range of the function  $U \mapsto \Psi(U, P_n)$  for the empirical distribution  $P_n$  associated with the city data in Fig. 1, where  $n = 76$ . The parts colored red correspond to possible quarter medians.

### 2.4 Asymptotics

Throughout this section we assume that  $P$  is a probability measure on  $(\mathbb{R}^d, \mathcal{B}^d)$  with the property that

$$P^a \text{ has a unique median for all } a \in S_d, \tag{13}$$

and that  $X_1, X_2, \dots$  are independent copies of a random vector  $X$  with distribution  $P$ . We further assume that, for each  $n \in \mathbb{N}$ , the pair  $(\theta_n, U_n) \in \mathbb{R}^d \times \mathbb{O}(d)$  with  $\theta_n = \theta_n(X_1, \dots, X_n)$ ,  $U_n = U_n(X_1, \dots, X_n)$ , is a solution of the quarter median problem for the (random) empirical distribution  $P_n$  associated with  $X_1, \dots, X_n$ . We also assume that  $\theta_n$  and  $U_n$  are measurable with respect to the respective Borel  $\sigma$ -fields on  $(\mathbb{R}^d)^n$ . A discussion of the associated selection problem and an extension of the results to more general random elements is given in Sect. 2.6.

We deal with consistency first and then consider asymptotic normality. Recall that uniqueness and convergence refer not to the  $U$ -matrices themselves, but to the associated equivalence classes  $[U] \in \mathbb{H}(d)$ , hence we have to specify what convergence of the  $U$ -part means. On  $\mathbb{O}(d)$  we use the topology induced by  $\mathbb{R}^{d \times d}$  where a sequence of matrices converges if all entries converge individually. It is well known that this makes  $\mathbb{O}(d)$  a compact and separable Hausdorff space. We use the quotient topology on  $\mathbb{H}(d)$ , which is the finest topology with the property that the mapping  $\mathbb{S}\mathbb{O}(d) \rightarrow \mathbb{H}(d)$ ,  $U \mapsto [U]$ , is continuous. As we have  $\mathbb{H}(d) = \mathbb{O}(d)/G$  with some finite group  $G \subset \mathbb{O}(d)$ , the space  $\mathbb{H}(d)$  with this topology is again a compact and separable Hausdorff space.

**Theorem 2** *Suppose that (13) holds.*

- (a) *If the quarter median for  $P$  is unique and given by  $\theta \in \mathbb{R}^d$ , then  $\theta_n$  converges almost surely to  $\theta$  as  $n \rightarrow \infty$ .*
- (b) *If the solution of the quarter median problem for  $P$  is unique and given by  $(\theta, H) \in \mathbb{S}_{d-1} \times \mathbb{H}(d)$ , then  $(\theta_n, [U_n])$  converges almost surely to  $(\theta, H)$  as  $n \rightarrow \infty$ .*

In connection with distributional asymptotics we require (13) and also that  $P$  is absolutely continuous. Let  $M_a = M(a)$  be the (unique) median of  $P^a$ ,  $a \in \mathbb{S}_{d-1}$ . As in the paper of Grübel (1996) we further assume that  $P^a$  has a density  $f_a$  such that

$$(a, t) \rightarrow f_a(t) \text{ is continuous at each } (a, t) \in D := \{(a, M_a) : a \in \mathbb{S}_{d-1}\}, \tag{14}$$

$$f_a(M_a) > 0 \text{ for each } a \in \mathbb{S}_{d-1}, \tag{15}$$

and, in addition, that

$$\text{there is a point } \theta \in \mathbb{R}^d \text{ such that } M_a = a^t\theta \text{ for each } a \in \mathbb{S}_{d-1}. \tag{16}$$

For example, (16) is satisfied if  $P$  is symmetric about  $\theta$  in the sense that  $(X_1 - \theta)$  and  $-(X_1 - \theta)$  have the same distribution. Of course, if (13) and (16) hold then the quarter median for  $P$  is unique and given by  $\theta$ .

For  $U \in \mathbb{O}(d)$  with rows  $b_1^t, \dots, b_d^t$  let

$$\Delta(P, U) := \text{diag}\left(\frac{1}{f_{b_1}(M_{b_1})^2}, \dots, \frac{1}{f_{b_d}(M_{b_d})^2}\right). \tag{17}$$

Of course,  $f_a$  and  $M_a$  also depend on  $P$ . Finally, we write  $X_n \rightarrow_d Q$  for convergence in distribution of a sequence  $(X_n)_{n \in \mathbb{N}}$  of random variables to a random variable with distribution  $Q$ , and  $N_d(\mu, \Sigma)$  for the  $d$ -dimensional normal distribution with mean vector  $\mu$  and covariance matrix  $\Sigma$ .

**Theorem 3** *Suppose that the conditions (13)–(16) hold and that the solution of the quarter median problem for the absolutely continuous probability measure  $P$  is unique and given by  $(\theta, H)$ . Then, in the above setup, and with  $U \in H$ ,*

$$\sqrt{n}(\theta_n - \theta) \rightarrow_d N_d(0, \Sigma(P, U)) \text{ as } n \rightarrow \infty, \tag{18}$$

where  $\Sigma(P, U) := U^t \Delta(P, U)U$  and  $\Delta(P, U)$  is given by (17).

As part of the proof we will show that  $\Sigma(P, U)$  does not depend on the choice of the element  $U$  of  $H$ .

### 2.5 Elliptical distributions

We now consider a special family of distributions in more detail. Let  $h : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  be a function with the property

$$0 < c_d(h) := \int_{\mathbb{R}^d} h(\|x\|^2) dx < \infty. \tag{19}$$

Then, for each  $\mu \in \mathbb{R}^d$  and each positive definite matrix  $\Sigma \in \mathbb{R}^{d \times d}$ ,

$$f(x) = f(x; \mu, \Sigma) = \frac{1}{c_d(h)(\det \Sigma)^{1/2}} h((x - \mu)^t \Sigma^{-1}(x - \mu)), \quad x \in \mathbb{R}^d,$$

is the density of a probability measure  $P$  on  $(\mathbb{R}^d, \mathcal{B}^d)$ , the *elliptical distribution* with *location vector*  $\mu$  and *dispersion matrix*  $\Sigma$ . We abbreviate this to  $P = \text{Ell}_d(\mu, \Sigma; h)$ . If  $\mu = 0$  and  $\Sigma = I_d$  we write  $\text{Sym}_d(h)$  for  $\text{Ell}_d(0, I_d; h)$  and speak of a *spherically*

symmetric distribution with the density  $h(\|x\|^2)/c_d(h)$ ,  $x \in \mathbb{R}^d$ . A transformation to spherical coordinates leads to

$$c_d(h) = \frac{2\pi^{d/2}}{\Gamma(d/2)} \int_0^\infty r^{d-1} h(r^2) dr = \frac{\pi^{d/2}}{\Gamma(d/2)} \int_0^\infty r^{\frac{d}{2}-1} h(r) dr. \tag{20}$$

The function  $h$  is said to be the *density generator* of the elliptical distributions  $\text{Ell}_d(\mu, \Sigma; h)$ . For  $h(t) = \exp(-t/2)$  the associated constant is  $c_d(h) = (2\pi)^{d/2}$ , and we obtain the multivariate normal distributions  $N_d(\mu, \Sigma)$ .

For later purposes we list some properties of elliptical distributions. Clearly, if  $P = \text{Ell}_d(\mu, \Sigma; h)$  and if  $T : \mathbb{R}^d \rightarrow \mathbb{R}^d$ ,  $T(x) = Ax + b$  with  $A \in \mathbb{R}^{d \times d}$  regular,  $b \in \mathbb{R}^d$ , then  $P^T = \text{Ell}_d(\mu + b, A\Sigma A^t; h)$ . Further, for  $d \geq 2$  the univariate marginal densities of  $P = \text{Sym}_d(h)$  coincide and are given by

$$g_1(y) = g(y^2)/c_d(h), \quad y \in \mathbb{R},$$

where

$$\begin{aligned} g(y) &= \int_{\mathbb{R}^{d-1}} h(y + x_2^2 + \dots + x_d^2) dx_2 \dots dx_d \\ &= \frac{\pi^{\frac{d-1}{2}}}{\Gamma(\frac{d-1}{2})} \int_y^\infty h(t)(t-y)^{\frac{d-1}{2}-1} dt, \quad y \geq 0. \end{aligned}$$

The function  $g$  is nonincreasing and continuous on  $(0, \infty)$ . Hence the univariate marginal distributions are unimodal with mode 0 and their density is continuous on  $\mathbb{R} \setminus \{0\}$ ; see e.g. Fang et al. (1990, p. 37).

In what follows we deal with generators  $h$  that satisfy the condition

$$\sup_{0 \leq t \leq \epsilon} h(t) < \infty \quad \text{for some finite interval } [0, \epsilon]. \tag{21}$$

Then  $g$  and  $g_1$  are bounded and continuous. Additionally, with

$$c_k(h) := \frac{2\pi^{k/2}}{\Gamma(k/2)} \int_0^\infty r^{k-1} h(r^2) dr = \frac{\pi^{k/2}}{\Gamma(k/2)} \int_0^\infty r^{\frac{k}{2}-1} h(r) dr \tag{22}$$

for  $k = 1, \dots, d - 1$ , we have  $0 < c_k(h) < \infty$ ; especially,  $g(0) = c_{d-1}(h)$ . Finally, the transformation law mentioned above can be generalized to  $k \times d$ -matrices with rank  $k < d$  at the cost of changing the function  $h$ . In particular, the first marginal distribution associated with  $\text{Ell}_d(0, \Sigma; h)$  is  $\text{Ell}_1(0, \sigma_{11}; g)$  if  $\Sigma = (\sigma_{ij})_{i,j=1}^d$ , and has a density that is bounded, continuous and strictly positive at 0. More generally, if  $P = \text{Ell}_d(\mu, \Sigma; h)$  then the density  $f_a$  of  $P^a$  for  $a \in \mathbb{S}_{d-1}$  is given by

$$f_a(t) = \frac{1}{c_d(h)(a^t \Sigma a)^{1/2}} g\left(\frac{(t - a^t \mu)^2}{a^t \Sigma a}\right), \quad t \in \mathbb{R},$$

$M_a = a^t \mu$  is the unique median of  $P^a$ ,  $f_a(M_a)$  is positive for all  $a \in \mathbb{S}_{d-1}$ , and the function  $(a, t) \rightarrow f_a(t)$  is continuous. Hence the assumptions (13)–(16) of Theorem 3 are satisfied.

As  $\Sigma$  is positive definite and symmetric the principal axes theorem applies and gives the representation

$$\Sigma = U^t \text{diag}(\lambda_1, \dots, \lambda_d) U \tag{23}$$

with some  $U \in \mathbb{O}(d)$  and  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_d > 0$ , where the  $\lambda_i$ 's are the eigenvalues of  $\Sigma$ . The distribution  $P = \text{Ell}_d(\mu, \Sigma; h)$  is said to be *strictly elliptical* if no two of these eigenvalues coincide. The next theorem deals with the uniqueness of (the solution of) the quarter median (problem) for elliptical distributions. Note that condition (21) is not imposed there.

**Theorem 4** *If  $P = \text{Ell}_d(\mu, \Sigma; h)$ , then the quarter median of  $P$  is unique and given by  $\mu$ . Moreover, if  $P$  is strictly elliptical, then the solution of the quarter median problem is unique and given by  $(\mu, [U])$  with  $U$  as in (23).*

For samples from strictly elliptical distributions with density generators satisfying (21) we obtain asymptotic normality for a sequence of associated quarter medians  $\theta_n = \theta_n(X_1, \dots, X_n)$ . Part (a) of the following result is essentially a corollary to Theorem 3. In the symmetric case we only have weak uniqueness, hence we need to consider this situation separately. Intermediate cases, where only some of the eigenvalues coincide, can be treated similarly, but require a certain amount of bookkeeping.

**Theorem 5** *Let  $X_n, n \in \mathbb{N}$ , be independent random vectors with distribution  $P = \text{Ell}_d(\mu, \Sigma; h)$ , with  $h$  satisfying (21). Let  $c_{d-1}(h), c_d(h)$  be as in (19) and let*

$$\sigma_{\text{QMED}}^2 := c_d(h)^2 / (2c_{d-1}(h))^2.$$

(a) *Suppose that  $(\theta_n, U_n)_{n \in \mathbb{N}}$  is a sequence of random variables with values in  $\mathbb{R}^d \times \mathbb{O}(d)$  such that, for all  $n \in \mathbb{N}$ ,  $(\theta_n, U_n)$  solves the quarter median problem for  $P_n$ . Then, if  $P$  is strictly elliptical,*

$$\sqrt{n}(\theta_n - \mu) \rightarrow_d N_d(0, \sigma_{\text{QMED}}^2 \Sigma) \text{ as } n \rightarrow \infty. \tag{24}$$

(b) *Suppose that  $\Sigma = \lambda I_d$  for some  $\lambda > 0$ . Let  $(U_n)_{n \in \mathbb{N}}$  be an arbitrary sequence of elements of  $\mathbb{O}(d)$  and suppose that, for all  $n \in \mathbb{N}$ ,  $\theta_n$  is a marginal median of  $P_n^{U_n}$ . Then*

$$\sqrt{n}(\theta_n - \mu) \rightarrow_d N_d(0, \sigma_{\text{QMED}}^2 \Sigma) \text{ as } n \rightarrow \infty. \tag{25}$$

Remarkably, the covariance matrix of the limiting normal distribution always is a fixed multiple of the dispersion matrix  $\Sigma$ . This is in contrast to other familiar estimators of  $\mu$  such as the empirical spatial median  $\text{SMED}(X_1, \dots, X_1)$ , for example. In fact, with reference to Bai et al. (1990) where the asymptotics of the spatial median are considered in an even more general context, Somorčík (2006) states that in the  $d$ -dimensional case under some weak conditions

$$\text{SMED}(X_1, \dots, X_1) \rightarrow_d N_d(0, V), \tag{26}$$

where the asymptotic covariance matrix is given by

$$V = D_1^{-1} D_2 D_1^{-1} \tag{27}$$

with

$$D_1 = E \left( \frac{1}{\|X_1 - \mu\|} \left( I_d - \frac{(X - \mu)(X - \mu)^t}{\|X - \mu\|^2} \right) \right)$$

and

$$D_2 = E \left( \frac{(X - \mu)(X - \mu)^t}{\|X - \mu\|^2} \right).$$

In the special (spherical) symmetric case  $\Sigma = I_d$  the identity (27) simplifies to

$$V = \sigma_{\text{SMED}}^2 I_d,$$

where

$$\sigma_{\text{SMED}}^2 = \frac{d}{(d - 1)^2} \left( E \left( \frac{1}{\|X - \mu\|} \right) \right)^{-2}. \tag{28}$$

It was shown by Grübel (1996) that in this symmetric case the empirical orthomedian has the same asymptotic covariance matrix as the spatial median. In what follows, we give a comparison of the covariances of the limit distributions for some estimators of  $\mu$ , in particular for QMED( $X_1, \dots, X_n$ ) as the empirical quarter median, the sample mean  $\frac{1}{n} \sum_{j=1}^n X_j$ , and the maximum likelihood estimator ML( $X_1, \dots, X_n$ ), for some special distribution families, where our main focus is on strictly elliptical distribution families. In all cases, the limit distribution of the standardized estimators is a centered multivariate normal, and by asymptotic covariance we mean the respective covariance matrix.

**Example 1** We work out the details for some specific distribution families in the case of dimension  $d = 2$ .

(a) For normal distributions  $P = N_2(\mu, \Sigma)$  Theorem 5 (a) leads to

$$\sqrt{n}(\text{QMED}(X_1, \dots, X_n) - \mu) \rightarrow_d N_2\left(0, \frac{\pi}{2} \Sigma\right). \tag{29}$$

The maximum likelihood estimator of  $\mu$  is the sample mean; its (asymptotic) covariance is simply  $\Sigma$  itself, hence the asymptotic efficiency of the quarter median is about 64%. Of course, the quarter median is more robust with respect to gross outliers in the data. In the symmetric case  $\Sigma = I_2$  both orthomedian and spatial median have asymptotic covariance  $\frac{4}{\pi} I_2$ .

For non-symmetric two-dimensional normal distributions with  $\Sigma = \text{diag}(1, \lambda)$ ,  $0 < \lambda < 1$ , Brown (1983) and Grübel (1996) obtained expressions for the asymptotic covariance matrices of the spatial median and the orthomedian. Both are of diagonal form, with the two diagonal entries given by expressions involving infinite series and double integrals respectively. These can be used numerically, see Brown (1983, Table 1) and Grübel (1996, Table 1). For the second value, referring to the shorter axis, we

have the simple explicit formula  $\lambda\pi/2$  for the quarter median, which is smaller than the value for the spatial median if  $\lambda < 0.038$ , and for the orthomedian if  $\lambda < 0.1$ .

(b) Taking  $h(t) = \exp(-\frac{1}{2}\sqrt{t})$ ,  $t \geq 0$ , we obtain the bivariate doubly exponential distributions  $MDE_2(\mu, \Sigma)$ , with densities given by

$$f(x; \mu, \Sigma) = \frac{1}{8\pi} \frac{1}{\sqrt{\det \Sigma}} \exp\left(-\frac{1}{2} \left((x - \mu)^t \Sigma^{-1} (x - \mu)\right)^{\frac{1}{2}}\right), \quad x \in \mathbb{R}^2.$$

The asymptotic covariance matrix for the sample mean of independent two-dimensional random vectors  $X_1, \dots, X_n$  with the density  $f(\cdot; \mu, \Sigma)$  is given by  $\text{cov}(X_1) = 12 \Sigma$ ; see, e.g. (Gómez et al. 1998, Proposition 3.2 (iii)). Using

$$c_1(h) = 2 \int_0^\infty e^{-\frac{1}{2}t} dt = 4, \quad c_2(h) = 2\pi \int_0^\infty t e^{-\frac{1}{2}t} dt = 8\pi,$$

we obtain  $\sigma_{\text{QMED}}^2 = \pi^2$ , so that  $\pi^2 \Sigma$  is the asymptotic covariance matrix for the sample quarter median. For a given fixed symmetric positive definite  $\Sigma$  the Fisher information matrix can be calculated to be  $\frac{1}{8} \Sigma^{-1}$  (see, e.g. Mitchell 1989, formulas (3.2), (3.3)). So, by the general asymptotic theory of maximum likelihood estimation  $8\Sigma$  is the asymptotic covariance of the maximum likelihood estimator  $\text{ML}(X_1, \dots, X_n)$ . In the symmetric case  $\Sigma = I_d$ , it follows from (26) and (28) that

$$\text{SMED}(X_1, \dots, X_n) \rightarrow_d N_2(0, 8I_2), \tag{30}$$

which means that the asymptotic covariances of  $\text{SMED}(X_1, \dots, X_n)$  and  $\text{ML}(X_1, \dots, X_n)$  coincide in this case. This is not surprising, because  $\text{SMED}(X_1, \dots, X_n)$  is the maximum likelihood estimator of  $\mu$  in this symmetric case.

(c) With  $h(t) = (1 + t)^{-3/2}$  we obtain a class of bivariate Cauchy distributions  $\text{MC}_2(\mu, \Sigma)$  with densities

$$f(x; \mu, \Sigma) = \frac{1}{2\pi} \frac{1}{\sqrt{\det \Sigma}} \left(1 + (x - \mu)^t \Sigma^{-1} (x - \mu)\right)^{-\frac{3}{2}}, \quad x \in \mathbb{R}^2.$$

The associated marginal distributions are univariate Cauchy distributions. For this distribution family the sample mean is not even consistent, hence we do not take it into consideration as an estimator for  $\mu$ . Again we use (Mitchell 1989, formulas (3.2),(3.3)) to see that for given fixed positive definite  $\Sigma$  the Fisher information matrix is  $\frac{3}{5} \Sigma^{-1}$ . Thus, by the general asymptotic theory of maximum likelihood estimators,

$$\text{ML}(X_1, \dots, X_n) \rightarrow_d N_2\left(0, \frac{5}{3} \Sigma\right).$$

Using

$$c_1(h) = 2 \int_0^\infty (1 + t^2)^{-\frac{3}{2}} dt = 2, \quad c_2(h) = 2\pi \int_0^\infty t (1 + t^2)^{-\frac{3}{2}} dt = 2\pi$$



we arrive at the value  $\frac{\pi^2}{4} \Sigma$  for the asymptotic covariance matrix of the sample quarter median. In the symmetric case  $\Sigma = I_d$ , using (26) and (28), it follows that

$$\text{SMED}(X_1, \dots, X_n) \rightarrow_d N_2(0, 2I_2). \tag{31}$$

(d) With  $h(t) = (1 - t)^2$  for  $0 \leq t \leq 1$  and 0 elsewhere we have a special class of symmetric bivariate Pearson type II distributions  $\text{SMPII}_2(\mu, \Sigma)$  with densities

$$f(x; \mu, \Sigma) = \frac{3}{\pi} \frac{1}{\sqrt{\det \Sigma}} \max\{0, 1 - (x - \mu)^t \Sigma^{-1}(x - \mu)\}^2, \quad x \in \mathbb{R}^2.$$

The asymptotic covariance matrix for the sample mean of independent two-dimensional random vectors  $X_1, \dots, X_n$  with the density  $f(\cdot; \mu, \Sigma)$  is  $\text{cov}(X_1) = \frac{1}{8} \Sigma$ ; see Fang et al. (1990, p. 89). The centered bivariate limit normal distribution of the maximum likelihood estimator has the covariance matrix  $\frac{1}{12} \Sigma$ . Using

$$c_1(h) = 2 \int_0^1 (1 - t^2)^2 dt = \frac{16}{15}, \quad c_2(h) = 2\pi \int_0^1 t(1 - t^2)^2 dt = \frac{\pi}{3}$$

it follows that  $\frac{25}{1024} \pi^2 \Sigma$  is the asymptotic covariance matrix for the sample quarter median. In the symmetric case  $\Sigma = I_d$  we obtain

$$\text{SMED}(X_1, \dots, X_n) \rightarrow_d N_2\left(0, \frac{25}{128} I_2\right).$$

(e) Taking  $h(t) = \frac{\exp(-t)}{(1 + \exp(-t))^2}$ ,  $t \geq 0$ , we obtain the symmetric bivariate logistic distributions  $\text{SML}_2(\mu, \Sigma)$ , with densities given by

$$f(x; \mu, \Sigma) = \frac{1}{\pi \sqrt{\det \Sigma}} \frac{\exp(-(x - \mu)^t \Sigma^{-1}(x - \mu))}{(1 + \exp(-(x - \mu)^t \Sigma^{-1}(x - \mu)))^2}, \quad x \in \mathbb{R}^2.$$

For  $X \sim \text{SML}_2(0, I_2)$  the distribution of  $R^2 = \|X\|^2$  is the univariate half-logistic distribution with density  $2 \frac{\exp(-t)}{(1 + \exp(-t))^2}$ ,  $t \geq 0$ . Then,

$$m_{\text{SML}_2} := \frac{1}{2} E(R^2) = \int_0^\infty t \frac{\exp(-t)}{(1 + \exp(-t))^2} dt \approx 0.69314718 \tag{32}$$

and  $\text{cov}(X) = m_{\text{SML}_2} I_2$ . From this we deduce that the asymptotic covariance matrix for the sample mean of independent two-dimensional random vectors  $X_1, \dots, X_n$  with the above density  $f(\cdot; \mu, \Sigma)$  is  $\text{cov}(X_1) = m_{\text{SML}_2} \Sigma$ . With

$$k_{\text{SML}_2} := 4 \int_0^\infty t \left( \frac{1 - \exp(-t)}{1 + \exp(-t)} \right)^2 \frac{\exp(-t)}{(1 + \exp(-t))^2} dt \approx 1.59086291 \tag{33}$$

**Table 1** Relative asymptotic efficiencies

Distribution family	eff(Mean, ML)	eff(QMED, ML)
Bivariate normal	1	$\frac{2}{\pi}$
Bivariate doubly exponential	$\frac{2}{3}$	$\frac{8}{\pi^2}$
Bivariate Cauchy	–	$\frac{20}{3\pi^2}$
Bivariate Pearson type II	$\frac{2}{3}$	$\frac{256}{75\pi^2}$
Symmetric bivariate logistic	0.90686321	0.46253446

the Fisher information matrix is  $k_{\text{SML}_2} \Sigma^{-1}$ , and the centered bivariate limit normal distribution of the maximum likelihood estimator has the covariance matrix  $k_{\text{SML}_2}^{-1} \Sigma$ . Further, with

$$c_1(h) = \int_0^\infty t^{-\frac{1}{2}} \frac{\exp(-t)}{(1 + \exp(-t))^2} dt \approx 0.67371824,$$

$$c_2(h) = \pi \int_0^\infty \frac{\exp(-t)}{(1 + \exp(-t))^2} dt = \frac{\pi}{2}$$

we obtain that  $\frac{c_2(h)^2}{(2c_1(h))^2} \Sigma \approx 1.35901156 \Sigma$  is the covariance matrix of the centered limiting normal distribution of the quarter median. In the symmetric case  $\Sigma = I_d$  we get

$$\text{SMED}(X_1, \dots, X_n) \rightarrow_d N_2 \left( 0, \frac{1}{2c_1(h)^2} I_2 \right),$$

where  $\frac{1}{2c_1(h)^2} \approx 1.10157328$ . Numerical values for the ratios  $k_{\text{SML}_2}^{-1}/m_{\text{SML}_2}$  and  $k_{\text{SML}_2}^{-1}/(\frac{c_2(h)}{2c_1(h)})^2$  are shown in the last line of Table 1.

We summarize some special results presented in Example 1 for the respective strictly elliptical cases in Table 1, where the relative asymptotic efficiencies  $\text{eff}(\text{Mean, ML})$ , and  $\text{eff}(\text{QMED, ML})$  of the mean and the quarter median with respect to the maximum likelihood estimator are shown; these are the ratios  $\sigma_{\text{ML}}^2/\sigma_{\text{Mean}}^2$  and  $\sigma_{\text{ML}}^2/\sigma_{\text{QMED}}^2$ .

In contrast to the other multivariate versions of the median that appear in the above example the quarter median goes beyond providing a location estimate as it comes with an equipartition basis. In the elliptical case  $P = \text{Ell}_d(\mu, \Sigma; h)$  the basis elements are the unit vectors proportional to the eigenvectors of  $\Sigma$  and thus contain information about the dispersion parameter. In Theorem 3 the asymptotic behavior of  $(U_n)_{n \in \mathbb{N}}$  is important. In the strictly elliptical case we have consistency and we may assume that convergence takes place in  $\mathbb{S}\mathbb{O}(d)$ ; see also the beginning of the proof of Theorem 3. We aim at a second order statement about the distributional asymptotics of this sequence.

In the following theorem we only consider  $d = 2$ , but see the ensuing remarks. We can then use the parametrization of  $\mathbb{S}\mathbb{O}(2)$  given in Remark 1 (b) by an angle from an half-open interval  $I$  of length  $\pi/2$ . We assume that  $U = U_\alpha$  with  $I$  chosen such that  $\alpha$  is in its interior. Then consistency implies that  $U_n = U_{\alpha_n}$  with  $I$ -valued random

variables  $\alpha_n$  that converge to  $\alpha$  with probability 1. It then makes sense to consider the distributional asymptotics of  $\sqrt{n}(\alpha_n - \alpha)$  as  $n \rightarrow \infty$ .

**Theorem 6** *Let  $P = \text{Ell}_2(\mu, \Sigma; h)$  and let  $X_n, \theta_n, U_n, n \in \mathbb{N}$ , be as in Theorem 5 (a). Let  $\lambda_1$  and  $\lambda_2$  be the eigenvalues of  $\Sigma$ , with  $\lambda_1 > \lambda_2$ . Then, with  $\alpha$  and  $\alpha_n$  as defined above,*

$$\sqrt{n}(\alpha_n - \alpha) \rightarrow_d N(0, \sigma^2(\lambda_1, \lambda_2)),$$

where

$$\sigma^2(\lambda_1, \lambda_2) = \frac{c_2(h)^2 \lambda_1 \lambda_2}{4c_1(h)^2 (\sqrt{\lambda_1} - \sqrt{\lambda_2})^2}. \tag{34}$$

Moreover,  $\alpha_n$  and  $\theta_n$  are asymptotically independent.

In particular, the asymptotic variance in (34) is large if the eigenvalues are close to each other. This is to be expected as  $\text{Ell}_2(\mu, \Sigma; h)$  is then close to a symmetric distribution, where  $U$  is not unique.

In dimensions higher than two a possible approach could be based on the representation of the Lie group  $\mathbb{S}\mathbb{O}(d)$  by its Lie algebra  $\mathfrak{so}(d)$ , which consists of the skew-symmetric matrices  $A \in \mathbb{R}^{d \times d}$ , where  $U \in \mathbb{S}\mathbb{O}(d)$  is written as the matrix exponential  $U = \exp(A)$  of some  $A \in \mathfrak{so}(d)$ ; see Muirhead (1982, Theorem A9.7) for a proof of the latter assertion, and Hall (2004) for an elementary introduction to Lie groups, Lie algebras, and representations. This leads to an asymptotic distribution  $Q$  for  $(U^t U_n)^{\sqrt{n}}$  as  $n \rightarrow \infty$ , where  $Q$  is a probability measure on (the Borel subsets of)  $\mathbb{S}\mathbb{O}(d)$ , and it avoids any ambiguities caused by parametrization. In fact,  $Q$  is the distribution of the random orthogonal matrix  $\exp(A)$ , where  $A$  is a random skew symmetric matrix with jointly normal subdiagonal entries. If  $d = 2$ , as in the above theorem, then

$$U_\alpha = \exp(A(\alpha)) \quad \text{with } A := \begin{pmatrix} 0 & -\alpha \\ \alpha & 0 \end{pmatrix},$$

and  $A$  is then specified by its one subdiagonal entry, which has a central normal distribution with variance given by (34).

### 2.6 Measurability, selection and algorithms

Let  $S$  be a topological space with Borel  $\sigma$ -field  $\mathcal{B}(S)$ . We write  $C_b(S)$  for the set of bounded continuous functions  $f : S \rightarrow \mathbb{R}$ . Let  $Z, Z_1, Z_2, \dots$  be  $S$ -valued random variables, i.e.  $\mathcal{B}(S)$ -measurable functions on some probability space (which may depend on the respective variable). In its classical form, convergence in distribution  $Z_n \rightarrow_d Z$  of  $Z_n$  to  $Z$  as  $n \rightarrow \infty$  means that

$$\lim_{n \rightarrow \infty} Ef(Z_n) = Ef(Z) \quad \text{for all } f \in C_b(S). \tag{35}$$

This is the notion that we used in Sect. 2.4. An extension of this concept, due to Hoffmann–Jørgensen, can be applied if the  $Z_n$ 's are not measurable with respect to the full Borel  $\sigma$ -field on  $S$ ; roughly, the expectations  $Ef(Z_n)$  in (35) are then replaced by outer expectations  $E^*f(Z_n)$ . Similarly, almost sure convergence now refers to

outer probabilities. The research monographs of Dudley (1999) and van der Vaart and Wellner (1996) give an in-depth treatment of this circle of ideas, together with a variety of applications. This extension appears in our proof of Theorem 3 in connection with empirical processes; it can also be used if the estimators  $(\theta_n, U_n)$  are not measurable. We refer the reader to the paper of Kuelbs and Zinn (2013), where this is worked out in detail for the related situation of quantile processes.

While this avoids the problem of choosing an estimator from the respective solution set in a measurable way, it is of independent interest whether such a selection is possible. For the classical one-dimensional median and the marginal median this can obviously be done by choosing the midpoint of the respective (component) interval, but no such simple rule seems to exist for the quarter median.

We think of estimators as functions of the random variables  $X_i, i \in \mathbb{N}$ , and it is then enough to establish measurability for these functions. To be precise, for a given  $n \in \mathbb{N}$  and an  $n$ -tuple  $(x_1, \dots, x_n) \in (\mathbb{R}^d)^n$  we denote by  $(\theta_n(x_1, \dots, x_n), U_n(x_1, \dots, x_n)) \in \mathbb{R}^d \times \text{SO}(d)$  a solution of the quarter median problem for the empirical distribution  $P_{n;x_1, \dots, x_n} = n^{-1} \sum_{k=1}^n \delta_{x_k}$ .

**Proposition 2** *There exists a function  $\tau : (\mathbb{R}^d)^n \rightarrow \mathbb{R}^d \times \text{SO}(d)$  that is measurable with respect to the Borel  $\sigma$ -algebra on  $(\mathbb{R}^d)^n$  and the Borel  $\sigma$ -algebra on  $\mathbb{R}^d \times \text{SO}(d)$  such that, for each  $(x_1, \dots, x_n) \in (\mathbb{R}^d)^n$ ,*

$$(\theta_n(x_1, \dots, x_n), U_n(x_1, \dots, x_n)) := \tau(x_1, \dots, x_n)$$

*is a solution of the quarter median problem for  $P_{n;x_1, \dots, x_n}$ .*

*Moreover,  $\tau$  may be chosen to be permutation invariant in the sense that*

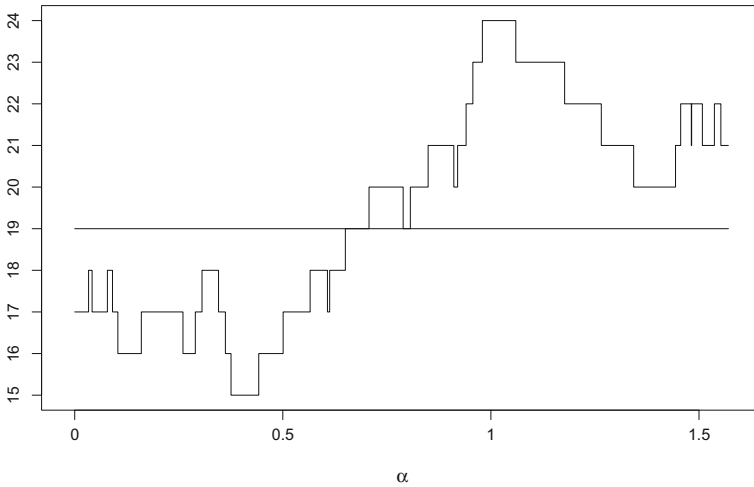
$$\tau(x_{\pi(1)}, \dots, x_{\pi(n)}) = \tau(x_1, \dots, x_n) \tag{36}$$

*for all permutations  $\pi$  of  $\{1, \dots, n\}$ .*

Permutation invariance means that the estimates depend on the data only through the respective empirical distribution.

Selecting a solution in a measurable way from the respective set of all solutions may seem as a corollary to establishing an algorithm that returns an estimate for every data input  $x_1, \dots, x_n$ . Indeed, without such an algorithm a statistical procedure would seem to be of limited use. In dimension two, and using the parametrization in Remark 1 (b), we obtain a real function on the interval  $[0, \pi/2)$  by counting the number of data points in the upper right quadrant specified by  $U_\alpha$  and  $\text{MMED}_m(P^{U_\alpha})$ ; see Fig. 3 for a plot of this function for the city data in Fig. 1, where the counts refer to the open quarter space. By construction, all half-spaces contain at least half of the data points, so each angle that leads to a count of  $n/4$  (if  $n$  is divisible by 4) would give a quarter median. For absolutely continuous distributions this approach, together with the intermediate value theorem, leads to a proof of Theorem 1 in dimension 2, and this may be used numerically via bisection. Such a bisection argument also works for functions that are piecewise constant and have jumps of size  $\pm 1$  only, provided a solution exists.

In what follows, still considering the case  $d = 2$ , we present a different algorithm. Let  $x_1, \dots, x_n$  be  $n \geq 2$  pairwise distinct data points in  $\mathbb{R}^2$ . For  $1 \leq i < j \leq n$  let



**Fig. 3** Number of cities in the upper right quadrant,  $\alpha \in [0, \pi/2)$

$b_{ij} = \frac{x_j - x_i}{\|x_j - x_i\|} \in \mathbb{S}_1$  be the direction of the line through  $x_i$  and  $x_j$ , and let  $b'_{ij} \in \mathbb{S}_1$  be orthogonal to  $b_{ij}$ . With the empirical distribution  $P_n = n^{-1} \sum_{k=1}^n \delta_{x_k}$  we associate the finite set  $\mathcal{L}(P_n)$  of pairs  $(\theta_{ij}, U_{ij})$ , with  $U_{ij} \in \mathbb{O}(2)$  as the matrix with row vectors  $b_{ij}$  and  $b'_{ij}{}^t$ , and  $\theta_{ij} = U_{ij}^t \eta$  with some  $\eta \in \mathcal{M}(P_n^{U_{ij}})$ .

**Theorem 7** *There is an element  $(\theta_{ij}, U_{ij}) \in \mathcal{L}(P_n)$  that solves the quarter median problem for  $P_n$ .*

Hence, checking successively the conditions (5) and (6) for all  $(\theta_{ij}, U_{ij}) \in \mathcal{L}(P_n)$ ,  $1 \leq i < j \leq n$ , provides a solution of the quarter median problem for  $P_n$ . This procedure, however, may lead to an estimator that is not permutation invariant in the sense of (36), meaning that it would not be a function of the empirical distribution  $P_n$ . For our introductory data example, see Fig. 1, this happens for both of the above algorithms: Fig. 3 shows that different solutions appear depending on whether the angles are scanned clockwise or counterclockwise. Similarly, the algorithm based on Theorem 7 could lead to the  $(i, j)$ -pair corresponding to the cities Berlin and Frankfurt am Main, or München and Dortmund, for example.

### 3 Numerical comparison

We present a small simulation study on the performance of the quarter median (QMED) as a location estimator, comparing it with that of three other procedures—the spatial median (SMED); Oja’s simplicial median (OMED), see Oja (1983); and Tukey’s half-space median (TMED), see Tukey (1975). For a detailed description of these classical estimators we refer to the survey papers of Small (1990) and Oja (2013). A Monte Carlo study on the performance of the three (and some other bivariate location) estimators is given by Massé and Plante (2003).

Samples are drawn from the strictly elliptical distributions  $N_2(\mu, \Sigma)$  (bivariate normal),  $MDE_2(\mu, \Sigma)$  (bivariate doubly exponential),  $SMPII_2(\mu, \Sigma)$  (symmetric bivariate Pearson type II), and  $SML_2(\mu, \Sigma)$  (symmetric bivariate logistic) introduced in Example 1. The bivariate Cauchy distributions  $MC_2(\mu, \Sigma)$  also discussed there are not taken into consideration for the simulation study, as the Oja median does not exist for this distribution family. In fact, for a given distribution  $P$  on  $\mathbb{R}^2$ , independent  $X_j \sim P, j = 1, 2$ , and  $\theta \in \mathbb{R}^2$ , let  $\Delta(X_1, X_2, \theta)$  be the area of the triangle in  $\mathbb{R}^2$  with vertices  $X_1, X_2, \theta$ . If the expectation  $\gamma_P(\theta) := E(\Delta(X_1, X_2, \theta))$  is finite for all  $\theta \in \mathbb{R}^2$ , the Oja median is a point  $\theta_{Oja}(P) \in \mathbb{R}^2$  that minimizes the function  $\theta \mapsto \gamma_P(\theta)$ . Obviously,  $\gamma_P \equiv \infty$  for  $P = MC_2(0, \Sigma)$ , while for the other elliptical distributions  $P$  under consideration  $\theta_{Oja}(P) = \mu$ .

Without loss of generality we take  $\mu = 0$ . The dispersion matrix chosen is  $\Sigma = \begin{pmatrix} 1 & 0 \\ 0 & \lambda \end{pmatrix}$ , with  $\lambda \in \{0.01, 0.1, 0.5, 0.9\}$ . Repeatedly, with  $r = 10,000$  replications, samples of size  $n = 100$  are drawn from the underlying distribution. Let  $SMED_{n,i}, OMED_{n,i}, TMED_{n,i}, QMED_{n,i}$  be the observed values of the corresponding estimators obtained in the  $i$ th repetition. Table 2 shows the numerical values of the components  $\hat{m}_1, \hat{m}_2$  of the empirical means and the numerical values of the eigenvalues  $\hat{l}_1, \hat{l}_2$  of the empirical covariance matrices of  $\sqrt{n}SMED_{n,i}, \sqrt{n}OMED_{n,i}, \sqrt{n}TMED_{n,i}, \sqrt{n}QMED_{n,i}, i = 1, \dots, r$ . The simulations are conducted by using the statistical software environment **R**; the calculation of the estimators  $SMED, OMED$ , and  $TMED$  is easily done by applying the corresponding functions `med(..., method="Spatial")`, `med(..., method="Oja", ...)` and `TukeyMedian(...)` provided with the additional **R** software packages `depth` and `TukeyRegion`. Note that, in the later package, the Tukey median is defined to be the barycenter of the Tukey region.

Roughly, for the bivariate distributions considered the Oja median seems to have the best power performance; it is followed by the Tukey median, the spatial median and the quarter median (in this order). But note the exceptional case  $\lambda = 0.01$ , with the smaller values of  $\hat{l}_2$  and  $\hat{l}_1\hat{l}_2$  for  $QMED$  compared to that of  $SMED$ . Note also that as we already pointed out above, the quarter median contains information about the dispersion parameter and so goes beyond providing a location estimate.

For samples taken from bivariate normal distributions with dispersion matrix  $\Sigma = \begin{pmatrix} 1 & 0 \\ 0 & \lambda \end{pmatrix}$  the limit distributions of  $SMED, OMED$ , and  $QMED$  as  $n \rightarrow \infty$  are the centered bivariate normal distributions with the diagonal covariance matrices  $\Sigma_{SMED} = \begin{pmatrix} l_{1,SMED} & 0 \\ 0 & l_{2,SMED} \end{pmatrix}$ , see Brown (1983), where also expressions for the diagonal elements  $l_{1,SMED}, l_{2,SMED}$  and their numerical values for some special values of  $\lambda$  are given,  $\Sigma_{OMED} = \begin{pmatrix} 4/\pi & 0 \\ 0 & 4\lambda/\pi \end{pmatrix}$ , see Oja and Niinimaa (1985), and  $\Sigma_{QMED} = \begin{pmatrix} \pi/2 & 0 \\ 0 & \pi\lambda/2 \end{pmatrix}$ . The limit distribution of  $TMED$  is the distribution of  $\arg \sup_{t \in \mathbb{R}^2} \inf_{u \in \mathbb{S}_1} \left( Z(u) - \frac{1}{2\pi\lambda^{1/2}} u^t t \right)$ , where  $(Z(u), u \in \mathbb{S}_1)$  is a centered Gaussian process with the covariance function  $\text{cov}(Z(u), Z(v)) = P(V_{++}(0; u, v) -$

**Table 2** Components  $\hat{m}_1, \hat{m}_2$  of the empirical means, and eigenvalues  $\hat{l}_1, \hat{l}_2$  of the empirical covariance matrices of the  $\sqrt{n}$ -scaled estimators SMED, OMED, TMED, QMED obtained by simulations with 10,000 replications for strictly elliptical distributions with dispersion matrix  $\Sigma = \begin{pmatrix} 1 & 0 \\ 0 & \lambda \end{pmatrix}, \lambda \in \{0.01, 0.1, 0.5, 0.9\}$ ; sample size  $n = 100$

	$\hat{m}_1$	$\hat{m}_2$	$\hat{l}_1$	$\hat{l}_2$
$\lambda = 0.01$				
$N_2(0, \Sigma)$				
SMED	0.009648	-0.002215	1.453955	0.019903
OMED	0.001726	-0.001769	1.282321	0.012935
TMED	-0.000700	-0.001661	1.311815	0.013135
QMED	0.010454	-0.001571	1.558638	0.015973
$MDE_2(0, \Sigma)$				
SMED	0.062843	0.000282	9.651512	0.132048
OMED	0.047608	-0.001115	8.376979	0.084180
TMED	0.041778	-0.001260	8.548116	0.085507
QMED	0.066367	0.000427	10.363242	0.102538
$SMPII_2(0, \Sigma)$				
SMED	-0.002203	0.000669	0.227956	0.003093
OMED	-0.000128	0.000219	0.201077	0.001999
TMED	0.000433	0.000230	0.205588	0.002041
QMED	-0.003445	0.000139	0.243375	0.002450
$SML_2(0, \Sigma)$				
SMED	-0.004630	-0.000134	1.245813	0.017274
OMED	-0.014100	-0.000211	1.102237	0.010843
TMED	-0.014277	-0.000292	1.127860	0.011012
QMED	-0.001921	-0.000723	1.326942	0.013266
$\lambda = 0.1$				
$N_2(0, \Sigma)$				
SMED	0.003782	-0.002045	1.347796	0.139426
OMED	0.002807	-0.000513	1.290243	0.127393
TMED	0.004400	-0.001163	1.316496	0.129837
QMED	0.001002	-0.002581	1.550401	0.157899
$MDE_2(0, \Sigma)$				
SMED	-0.024886	-0.014379	8.864224	0.937143
OMED	-0.025370	-0.011321	8.434326	0.854891
TMED	-0.020137	-0.009642	8.564605	0.873420
QMED	-0.031408	-0.003320	10.250833	1.052323
$SMPII_2(0, \Sigma)$				
SMED	-0.008639	-0.002740	0.204793	0.021491
OMED	-0.006975	-0.002715	0.197094	0.019512
TMED	-0.006893	-0.002416	0.201236	0.019897
QMED	-0.011107	-0.001635	0.232204	0.023828

Table 2 continued

	$\hat{m}_1$	$\hat{m}_2$	$\hat{l}_1$	$\hat{l}_2$
SML <sub>2</sub> (0, $\Sigma$ )				
SMED	-0.008358	-0.006214	1.175185	0.120428
OMED	-0.011263	-0.004974	1.119404	0.110270
TMED	-0.013367	-0.005026	1.146522	0.112307
QMED	-0.007513	-0.001273	1.351009	0.136288
$\lambda = 0.5$				
$N_2(0, \Sigma)$				
SMED	0.011107	-0.007306	1.274190	0.639970
OMED	0.011333	-0.006804	1.274619	0.636957
TMED	0.011808	-0.005540	1.309882	0.650354
QMED	0.009552	-0.007997	1.551297	0.767945
MDE <sub>2</sub> (0, $\Sigma$ )				
SMED	-0.022503	-0.017123	8.333408	4.103622
OMED	-0.019896	-0.016413	8.340335	4.101557
TMED	-0.018584	-0.020229	8.551098	4.201095
QMED	-0.022966	-0.005561	10.177384	5.088397
SMPII <sub>2</sub> (0, $\Sigma$ )				
SMED	-0.005249	-0.001641	0.197165	0.097735
OMED	-0.005218	-0.001727	0.196625	0.097539
TMED	-0.005330	-0.002090	0.200822	0.099652
QMED	-0.006836	-0.000873	0.237224	0.118714
SML <sub>2</sub> (0, $\Sigma$ )				
SMED	0.008670	0.013723	1.099552	0.563826
OMED	0.009137	0.013315	1.101891	0.561546
TMED	0.006853	0.015428	1.125076	0.573525
QMED	0.012909	0.018256	1.337041	0.684311
$\lambda = 0.9$				
$N_2(0, \Sigma)$				
SMED	0.010573	0.002084	1.284202	1.157361
OMED	0.010942	0.001677	1.291926	1.166358
TMED	0.009962	0.002722	1.314541	1.191291
QMED	0.007804	0.005679	1.553846	1.414504
MDE <sub>2</sub> (0, $\Sigma$ )				
SMED	0.036521	0.070980	8.360825	7.417609
OMED	0.037581	0.072080	8.417271	7.462682
TMED	0.037717	0.070611	8.586432	7.609654
QMED	0.007530	0.087122	10.274479	9.154238



**Table 2** continued

	$\hat{m}_1$	$\hat{m}_2$	$\hat{l}_1$	$\hat{l}_2$
SMPII <sub>2</sub> (0, $\Sigma$ )				
SMED	-0.006974	-0.000672	0.193231	0.178529
OMED	-0.006773	-0.001132	0.194627	0.178877
TMED	-0.006907	-0.001222	0.198142	0.183445
QMED	-0.005604	-0.002042	0.234378	0.220361
SML <sub>2</sub> (0, $\Sigma$ )				
SMED	-0.003381	0.003081	1.078231	0.982265
OMED	-0.005205	0.002348	1.082106	0.989597
TMED	-0.002501	0.003175	1.107970	1.007464
QMED	-0.010836	-0.000977	1.321225	1.190672

**Table 3** Bivariate normal case: Empirical eigenvalues  $\hat{l}_1, \hat{l}_2$  and eigenvalues  $l_1, l_2$  of  $\Sigma_{\text{SMED}}, \Sigma_{\text{OMED}},$  and  $\Sigma_{\text{QMED}}$

	$\hat{l}_1$	$\hat{l}_2$	$l_1$	$l_2$
$\lambda = 0.01$				
SMED	1.453955	0.019903	1.460156	0.019491
OMED	1.282321	0.012935	1.273240	0.012732
QMED	1.558638	0.015973	1.570796	0.015708
$\lambda = 0.1$				
SMED	1.347796	0.139426	1.347542	0.140825
OMED	1.290243	0.127393	1.273240	0.127324
QMED	1.550401	0.157899	1.570796	0.157080
$\lambda = 0.5$				
SMED	1.274190	0.639970	1.281958	0.641807
OMED	1.274619	0.636957	1.273240	0.636620
QMED	1.551297	0.767945	1.570796	0.785398
$\lambda = 0.9$				
SMED	1.284202	1.157361	1.273457	1.146117
OMED	1.291926	1.166358	1.273240	1.145916
QMED	1.553846	1.414504	1.570796	1.413717

1/4,  $u, v \in \mathbb{S}_1$ ; see Nolan (1999). To the best of our knowledge, nothing is known about the type of this distribution. In this respect, see for example also the interesting results on the asymptotic behaviour of the empirical Tukey depth process given by Massé (2004).

Table 3 shows the values of  $\hat{l}_1, \hat{l}_2$  and, for comparison, that of the diagonal elements  $l_1, l_2$  of  $\Sigma_{\text{SMED}}, \Sigma_{\text{OMED}},$  and  $\Sigma_{\text{QMED}}$ .

### 4 Proofs

#### 4.1 Proof of Theorem 1

Using Makeev’s result, we only need to remove the smoothness assumption.

For an arbitrary probability measure  $P$  on  $(\mathbb{R}^d, \mathcal{B}^d)$  let  $P_\epsilon = P * N_d(0, \epsilon I_d)$  be the convolution of  $P$  with a centered  $d$ -dimensional normal distribution with independent components that all have variance  $\epsilon > 0$ . Then  $P_\epsilon$  is absolutely continuous with an everywhere positive density, so that a solution  $(\theta_\epsilon, U_\epsilon)$  of the equipartition problem exists. Let  $\|\cdot\|$  be a matrix norm on  $\mathbb{R}^{d \times d}$  that is compatible with the Euclidean norm  $\|\cdot\|$  on  $\mathbb{R}^d$ . From  $\theta_\epsilon = U_\epsilon^t U_\epsilon \theta_\epsilon$  it then follows that

$$\|\theta_\epsilon\| \leq \|U_\epsilon^t\| \|U_\epsilon \theta_\epsilon\| \leq s \|U_\epsilon \theta_\epsilon\|, \tag{37}$$

where  $s = \sup_{U \in \mathbb{S}\mathbb{O}(d)} \|U^t\| < \infty$ . Let  $t = \limsup_{\epsilon \rightarrow 0} \|U_\epsilon \theta_\epsilon\|$ . Hence there exists a sequence  $(\epsilon_n)_{n \in \mathbb{N}}$  with  $\epsilon_n \rightarrow 0$  such that  $\|U_{\epsilon_n} \theta_{\epsilon_n}\| \rightarrow t$  and  $U_{\epsilon_n} \rightarrow U \in \mathbb{O}(d)$  as  $n \rightarrow \infty$ . As a consequence,  $P_{\epsilon_n}^{U_{\epsilon_n}} \rightarrow P^U$  weakly. Due to the fact that  $U_{\epsilon_n} \theta_{\epsilon_n}$  is a marginal median of  $P_{\epsilon_n}^{U_{\epsilon_n}}$  we deduce from this that the sequence  $(\|U_{\epsilon_n} \theta_{\epsilon_n}\|)_{n=1}^\infty$  is bounded; hence, by (37), the sequence  $(\theta_{\epsilon_n})_{n=1}^\infty$  is also bounded. Therefore,  $\theta_{\epsilon_{n'}} \rightarrow \theta \in \mathbb{R}^d$  and  $U_{\epsilon_{n'}} \rightarrow U \in \mathbb{O}(d)$  along some subsequence  $(\epsilon_{n'})$  of  $(\epsilon_n)$ . To see that  $(\theta, U)$  solves the quarter median problem for  $P$  we use that  $P_{\epsilon_{n'}}^{U_{\epsilon_{n'}}} \rightarrow P^U$  weakly and argue as follows: With  $b_{\epsilon_{n'}, 1}^t, \dots, b_{\epsilon_{n'}, d}^t$  and  $b_1^t, \dots, b_d^t$  as the row vectors of  $U_{\epsilon_{n'}}$  and  $U$ , respectively, we have that, for each  $1 \leq i < j \leq d$  and each  $\delta > 0$ ,

$$\begin{aligned} \frac{1}{4} &\leq \limsup_{n' \rightarrow \infty} P_{\epsilon_{n'}}(V_{+-}(\theta_{\epsilon_{n'}}; b_{\epsilon_{n'}, i}, b_{\epsilon_{n'}, j})) \\ &\leq \limsup_{n' \rightarrow \infty} P_{\epsilon_{n'}}(b_{\epsilon_{n'}, i}^t x \geq b_i^t \theta - \delta, b_{\epsilon_{n'}, j}^t x \leq b_j^t \theta + \delta) \\ &\leq P(b_i^t x \geq b_i^t \theta - \delta, b_j^t x \leq b_j^t \theta + \delta). \end{aligned}$$

Let  $\delta \downarrow 0$  to obtain  $P(V_{+-}(\theta; b_i, b_j)) \geq \frac{1}{4}$ . The other inequalities stated in (5) and (6) can be verified in the same way.

#### 4.2 Proof of Proposition 1

(a) Let  $\theta$  be a quarter median for  $P$  with  $b_1, \dots, b_d$  an associated orthonormal basis of  $\mathbb{R}^d$  such that the constraints (5) and (6) are satisfied. Further, let  $S(x) = Ax + c$  with  $A \in \mathbb{O}(d)$ ,  $c \in \mathbb{R}^d$ , be a Euclidean motion. Then  $b'_1, \dots, b'_d$  with  $b'_i := Ab_i$ ,  $1 \leq i \leq d$ , is again an orthonormal basis for  $\mathbb{R}^d$ , and it is easily checked that

$$\begin{aligned} S^{-1}(H_\pm(A\theta + c; b'_i)) &= H_\pm(\theta; b_i), \quad 1 \leq i \leq d, \\ S^{-1}(V_{\pm\pm}(A\theta + c; b'_i, b'_j)) &= V_{\pm\pm}(\theta; b_i, b_j), \quad 1 \leq i < j \leq d. \end{aligned}$$

In view of the definition of push-forwards this implies that  $S(\theta)$  is a quarter median for  $P^S$ , with  $b'_1, \dots, b'_d$  an associated equipartition basis.

(b) If  $V = AU$  with  $A \in G$  then

$$\Psi(V, P) = (AU)^t \text{MMED}_m((P^U)^A) = U^t A^t A \text{MMED}_m(P^U) = \Psi(U, P),$$

where we have used (9) and  $G \subset \mathbb{O}(d)$ .

(c) Let  $b^t_1, \dots, b^t_d$  be the row vectors of  $U$ . With

$$\theta_U = (\theta_{U,1}, \dots, \theta_{U,d})^t := U\theta = (b^t_1\theta, \dots, b^t_d\theta)^t$$

it holds that  $\theta = U^t\theta_U = \sum_{i=1}^d \theta_{U,i}b_i$  and that, by definition,  $\theta_U$  is a marginal median of  $P^U$ . We can now choose an element  $\tilde{\theta}_U = (\tilde{\theta}_{U,1}, \dots, \tilde{\theta}_{U,d})^t \in \mathcal{M}(P^U)$  such that

$$P(b^t_i x \geq \theta_{U,i}) = P(b^t_i x \geq \tilde{\theta}_{U,i}), \quad P(b^t_i x \leq \theta_{U,i}) = P(b^t_i x \leq \tilde{\theta}_{U,i})$$

for  $1 \leq i \leq d$ , and

$$\begin{aligned} P(b^t_i x \geq \theta_{U,i}, b^t_j x \geq \theta_{U,j}) &= P(b^t_i x \geq \tilde{\theta}_{U,i}, b^t_j x \geq \tilde{\theta}_{U,j}) \\ P(b^t_i x \geq \theta_{U,i}, b^t_j x \leq \theta_{U,j}) &= P(b^t_i x \geq \tilde{\theta}_{U,i}, b^t_j x \leq \tilde{\theta}_{U,j}) \\ P(b^t_i x \leq \theta_{U,i}, b^t_j x \geq \theta_{U,j}) &= P(b^t_i x \leq \tilde{\theta}_{U,i}, b^t_j x \geq \tilde{\theta}_{U,j}) \\ P(b^t_i x \leq \theta_{U,i}, b^t_j x \leq \theta_{U,j}) &= P(b^t_i x \leq \tilde{\theta}_{U,i}, b^t_j x \leq \tilde{\theta}_{U,j}) \end{aligned}$$

for  $1 \leq i < j \leq d$ . Then putting  $\tilde{\theta} = U^t\tilde{\theta}_U = \sum_{i=1}^d \tilde{\theta}_{U,i}b_i$  we have

$$P(b^t_i x \geq b^t_i\theta) = P(b^t_i x \geq b^t_i\tilde{\theta}), \quad P(b^t_i x \leq b^t_i\theta) = P(b^t_i x \leq b^t_i\tilde{\theta})$$

for  $1 \leq i \leq d$ , and

$$\begin{aligned} P(b^t_i x \geq b^t_i\theta, b^t_j x \geq b^t_j\theta) &= P(b^t_i x \geq b^t_i\tilde{\theta}, b^t_j x \geq b^t_j\tilde{\theta}) \\ P(b^t_i x \geq b^t_i\theta, b^t_j x \leq b^t_j\theta) &= P(b^t_i x \geq b^t_i\tilde{\theta}, b^t_j x \leq b^t_j\tilde{\theta}) \\ P(b^t_i x \leq b^t_i\theta, b^t_j x \geq b^t_j\theta) &= P(b^t_i x \leq b^t_i\tilde{\theta}, b^t_j x \geq b^t_j\tilde{\theta}) \\ P(b^t_i x \leq b^t_i\theta, b^t_j x \leq b^t_j\theta) &= P(b^t_i x \leq b^t_i\tilde{\theta}, b^t_j x \leq b^t_j\tilde{\theta}) \end{aligned}$$

for  $1 \leq i < j \leq d$ . Thus,  $(\tilde{\theta}, U)$  solves the quarter median problem for  $P$ .

(d) The first statement may be regarded as an equivariance property of solution pairs for the quarter median problem; its proof proceeds as in (a). For the second statement we note that  $A \in \mathbb{O}(d)$  and then calculate, using the shift equivariance of the marginal median,

$$\Psi(UA^t, P^S) = (UA^t)^t \text{MMED}_m((P^S)^{UA^t})$$

$$\begin{aligned}
 &= AU^t(\text{MMED}_m(P^{UA^tA}) + UA^tb) \\
 &= A\Psi(U, P) + b = S(\Psi(U, P)).
 \end{aligned}$$

(e) Let  $U \in \mathbb{O}(d)$  and  $(U_n)_{n=1}^\infty$  be a sequence of elements in  $\mathbb{O}(d)$  converging to  $U$ . Due to  $P^{U_n} \rightarrow P^U$  weakly and the fact that  $P^b$  has a unique median for all  $b \in S_d$ , it follows that  $\text{MMED}_m(P^{U_n}) \rightarrow \text{MMED}_m(P^U)$ . Then, from

$$\begin{aligned}
 &\|U_n^t \text{MMED}_m(P^{U_n}) - U^t \text{MMED}_m(P^U)\| \\
 &= \|U_n^t \text{MMED}_m(P^{U_n}) - U_n^t \text{MMED}_m(P^U) + U_n^t \text{MMED}_m(P^U) - U^t \text{MMED}_m(P^U)\| \\
 &\leq \|U_n^t\| \|\text{MMED}_m(P^{U_n}) - \text{MMED}_m(P^U)\| + \|U_n^t - U^t\| \|\text{MMED}_m(P^U)\|
 \end{aligned}$$

and the compactness of  $\mathbb{O}(d)$  we deduce that  $U_n^t \text{MMED}_m(P^{U_n}) \rightarrow U^t \text{MMED}_m(P^U)$ .

### 4.3 Proof of Theorem 2

We will make use of the fact that if  $m_n, n \in \mathbb{N}$ , are medians of univariate distributions that converge weakly as  $n \rightarrow \infty$  to a univariate distribution with unique median  $m$ , then  $\lim_{n \rightarrow \infty} m_n = m$ . Further, we only prove the second part of the theorem as the arguments for (a) are similar.

Let  $(\Omega, \mathcal{A}, \mathbb{P})$  be a probability space on which the random vectors  $X_1, X_2, \dots$  are defined, so that  $\mathbb{P}^{X_j} = P$  for all  $j \in \mathbb{N}$ . Then there is a  $\mathbb{P}$ -null set  $N \in \mathcal{A}$  such that for each  $\omega \in N^c$

$$P_{n;\omega} := \frac{1}{n} \sum_{j=1}^n \delta_{X_j(\omega)} \rightarrow P \text{ weakly.}$$

Fix  $\omega \in N^c$ ; until further notice we omit the argument  $\omega$  below. By definition, the quarter median  $\theta_n$  of  $P_n$  can be written as  $\theta_n = U_n^t \eta_n$  with some  $U_n \in \mathbb{O}(d)$  and some  $\eta_n \in [\text{MMED}_-(P_n^{U_n}), \text{MMED}_+(P_n^{U_n})]$  (both may depend on  $\omega \in N^c$ ). Let  $(U_{n'})$  be a subsequence of  $(U_n)$ . As  $\mathbb{O}(d)$  is compact a subsubsequence  $(U_{n''})$  converges to some  $V \in \mathbb{O}(d)$ , and we have weak convergence  $P_{n''}^{U_{n''}} \rightarrow P^V$  by the extended continuous mapping theorem; see Billingsley (1968, Theorem 5.5). As the distribution  $P^b$  has a unique median for all  $b \in S_{d-1}$ , the marginal median  $\eta_{n''}^i$  of  $P_{n''}^{U_{n''}}$  converges to the marginal median  $\eta^i$  of  $P^V$ . Now fix  $i, j$  with  $1 \leq i < j \leq d$ . Let  $a_{n''}^i, b_{n''}^i, a, b$  be the corresponding rows in  $U_{n''}^i$  and  $V$  respectively. Using the same arguments as in the proof of Theorem 1 we then obtain for each  $\epsilon > 0$ , and with  $m$  instead of  $n''$ ,

$$\begin{aligned}
 \frac{1}{4} &\leq \limsup_m P_m(a_m^t x \leq a_m^t \theta_m, b_m^t x \geq b_m^t \theta_m) \\
 &\leq \limsup_m P_m(a_m^t x \leq a^t \eta + \epsilon, b_m^t x \geq b^t \eta - \epsilon) \\
 &\leq P(a^t x \leq a^t \eta + \epsilon, b^t x \geq b^t \eta - \epsilon),
 \end{aligned}$$

where in the first line we have used that  $(\theta_m, U_m)$  solves the quarter median problem for  $P_m$ . Letting  $\epsilon \downarrow 0$  we get  $P(V_{-+}(\eta; a, b)) \geq 1/4$ . More generally, and using the

same argument, we obtain  $P(V_{\pm\pm}(\eta; a, b)) \geq 1/4$  as well as  $P(H_{\pm}(\eta; a)) \geq 1/2$  and  $P(H_{\pm}(\eta(\omega); b)) \geq 1/2$ . Recall that some of the quantities depend on the  $\omega$  chosen above so that, for example,

$$P(H_+(\eta; a)) = P(H_+(\eta(\omega); a(\omega))) = P(\{x \in \mathbb{R}^d : \langle a(\omega), x - \eta(\omega) \rangle \geq 0\}).$$

In summary, we have shown that every limit point  $(\eta(\omega), V(\omega))$  of the sequence  $(\theta_n(\omega), U_n(\omega))_{n \in \mathbb{N}}$  is a solution of the quarter median problem for  $P$ . Hence, if the solution is unique and given by  $(\theta, H) \in \mathbb{R}^d \times \mathbb{H}$ , then, on a set of probability 1,  $\eta = \theta$  and  $[V] = H$ .

### 4.4 Proof of Theorem 3

Invariance with respect to shifts means that we may assume  $\theta = 0$ . As already noticed in Sect. 4.3, the quarter median  $\theta_n$  can be written as  $\theta_n = U_n^t \eta_n$  with  $\eta_n \in [\text{MMED}_-(P_n^{U_n}), \text{MMED}_+(P_n^{U_n})]$ . Theorem 2 implies that  $([U_n])_{n \in \mathbb{N}}$  converges almost surely to  $H$  with respect to the quotient topology on  $\mathbb{H}(d)$ . The mapping  $p : \mathbb{O}(d) \rightarrow \mathbb{H}(d), U \mapsto [U]$ , associates to each  $H \in \mathbb{H}(d)$  a finite fiber  $p^{-1}(\{H\})$ , and  $\mathbb{O}(d)$  is a covering space for  $\mathbb{H}(d)$ . General results from algebraic topology, see e.g. tom Dieck (1991, Section II.6), imply that  $p$  may locally be inverted to obtain homeomorphisms to open subsets of the individual sheets of the covering space. In particular, we may assume that the matrices  $U_n$  converge in  $\mathbb{O}(d)$  to some  $U \in H$  almost surely as  $n \rightarrow \infty$ .

For each  $a \in \mathbb{S}_{d-1}$  and  $n \in \mathbb{N}$  let  $M_n^\pm(a) := \text{MED}_\pm(P_n^a)$ , and let  $M(a)$  be the uniquely determined median of  $P^a$ . We define two stochastic processes  $Y_n^\pm = (Y_n^\pm(a), a \in \mathbb{S}_{d-1})$  with index set  $\mathbb{S}_{d-1}$  by

$$Y_n^\pm(a) = \sqrt{n}(M_n^\pm(a) - M(a)) \quad \text{for all } a \in \mathbb{S}_{d-1},$$

and will use results presented by Grübel (1996) and adopt arguments used by Kuelbs and Zinn (2013) to prove that

$$Y_n^\pm \rightarrow_d Y \quad \text{as } n \rightarrow \infty. \tag{38}$$

Here  $Y = (Y_a, a \in \mathbb{S}_{d-1})$  is a centered Gaussian process with continuous paths and covariance function

$$\text{cov}(Y_a, Y_b) = \frac{K(a, b)}{f_a(M(a))f_b(M(b))}, \tag{39}$$

with

$$K(a, b) = P(\{x \in \mathbb{R}^d : a^t x \leq M(a), b^t x \leq M(b)\}) - \frac{1}{4}. \tag{40}$$

We regard the  $Y_n^\pm$  as processes with paths in  $\ell_\infty(\mathbb{S}_{d-1})$ , the space of real-valued bounded functions on  $\mathbb{S}_{d-1}$  endowed with the supremum norm, and the symbol ' $\rightarrow_d$ ' refers to weak or distributional convergence in the sense of Hoffmann–Jørgensen; see e.g. van der Vaart and Wellner (1996) and Dudley (1999) for details.

For the proof of (38) we start with a functional central limit theorem for the empirical processes and then use an almost sure representation in order to be able to work with individual paths in an  $\epsilon$ - $\delta$  style.

Let  $F(a, \cdot) := F_a(\cdot)$  be the distribution function of  $a^t X$  and let

$$F_n(a, \cdot) = F_{a,n}(\cdot) = \frac{1}{n} \sum_{j=1}^n 1(a^t X_j \leq \cdot)$$

be the (random) distribution function associated with  $P_n^a$ ,  $a \in \mathbb{S}_{d-1}$ . Here  $1(\cdot)$  denotes the indicator function of its (logical) argument. We introduce the empirical process  $Z_n = (\sqrt{n} (F_n(a, t) - F(a, t)), (a, t) \in \mathbb{S}_{d-1} \times \mathbb{R})$  as a process with paths in  $\ell_\infty(\mathbb{S}_{d-1} \times \mathbb{R})$ . We obtain a semimetric  $d$  on  $\mathbb{S}_{d-1} \times \mathbb{R}$  via

$$d((a, s), (b, t))^2 = E\left(\left(1(a^t X \leq s) - P(a^t X \leq s)\right) - \left(1(b^t X \leq t) - P(b^t X \leq t)\right)\right)^2,$$

$(a, s), (b, t) \in \mathbb{S}_{d-1} \times \mathbb{R}$ . As noticed in Grübel (1996, Proof of Theorem 1), we then have  $Z_n \rightarrow_d Z_0$ , where  $Z_0 = (Z_0(a, t), (a, t) \in \mathbb{S}_{d-1} \times \mathbb{R})$  is a centered Gaussian process with covariance function

$$E(Z_0(a, s)Z_0(b, t)) = E\left(\left(1(a^t X \leq s) - P(a^t X \leq s)\right)\left(1(b^t X \leq t) - P(b^t X \leq t)\right)\right),$$

$(a, s), (b, t) \in \mathbb{S}_{d-1} \times \mathbb{R}$ , and sample paths that are bounded and continuous with respect to  $d$ . The assumptions on  $P$  ensure that the covariance function is continuous with respect to the usual topology on  $\mathbb{S}_{d-1} \times \mathbb{R}$ . This implies that the process  $Z_0$  has continuous sample paths.

For the remainder of the proof we abbreviate ‘almost surely’ to ‘a.s.’, and convergence refers to  $n \rightarrow \infty$  unless specified otherwise.

By the Skorokhod–Dudley–Wichura representation theorem (Dudley 1999, Theorem 3.5.1) there exists a probability space  $(\tilde{\Omega}, \tilde{\mathcal{A}}, \tilde{\mathbb{P}})$  together with perfect and measurable maps  $g_n : \tilde{\Omega} \rightarrow \Omega$ ,  $n \in \mathbb{N}_0$ , such that  $\tilde{\mathbb{P}}^{g_n} = \mathbb{P}$  for all  $n \in \mathbb{N}_0$ , and such that, with  $\tilde{Z}_n := Z_n \circ g_n$ ,

$$\sup_{(a,t) \in \mathbb{S}_{d-1} \times \mathbb{R}} |\tilde{Z}_n(a, t) - \tilde{Z}_0(a, t)| \rightarrow 0 \quad \tilde{\mathbb{P}}\text{-a.s.} \tag{41}$$

As  $F_n$  is a deterministic function of  $Z_n$  it follows that  $\tilde{F}_n := F_n \circ g_n$  has the same distribution under  $\tilde{\mathbb{P}}$  as  $F_n$  under  $\mathbb{P}$ ,  $n \in \mathbb{N}$ . This extends to the associated one-dimensional projections and their medians, and to the  $Y_n$ -processes, where  $\tilde{M}_n^\pm(a) := \text{MED}_\pm(\tilde{F}_n(a, \cdot))$  and  $\tilde{Y}_n^\pm := Y_n^\pm \circ g_n$ ,  $n \in \mathbb{N}$ .

Arguing as in Grübel (1996, Proof of Theorem 1) we obtain that

$$\tilde{Y}_n^\pm(a) \rightarrow \tilde{Y}(a) := -f_a(M(a))^{-1} \tilde{Z}_0(a, M(a)) \quad \tilde{\mathbb{P}}\text{-a.s.}$$

for all  $a \in \mathbb{S}_{d-1}$ , and that

$$\sup_{a \in \mathbb{S}_{d-1}} |\tilde{Y}_n^\pm(a)| = O(1) \quad \tilde{\mathbb{P}}\text{-a.s.} \tag{42}$$

For the integral of the marginal medians needed by Grübel (1996) this was enough, but here we need the stronger stronger statement

$$\sup_{a \in \mathbb{S}_{d-1}} |\tilde{Y}_n^\pm(a) - \tilde{Y}(a)| \rightarrow 0 \quad \tilde{\mathbb{P}}\text{-a.s.} \tag{43}$$

For the proof of (43) we adopt some arguments given in Kuelbs and Zinn (2013, Proof of Proposition 1). First we note that  $F(a, M(a)) = 1/2$  and then write

$$F(a, M(a) + h) - 1/2 = f_a(M(a))h + r(a, h)h, \tag{44}$$

where  $r(a, h) = f_a(M(a) + \zeta_{a,h}h) - f_a(M(a))$  with some  $\zeta_{a,h} \in [0, 1]$ . Using compactness of  $\mathbb{S}_{d-1}$  we see that the assumption (14) implies

$$\sup_{a \in \mathbb{S}_{d-1}} |r(a, h)| \leq \sup_{a \in \mathbb{S}_{d-1}} \sup_{|s| \leq |h|} |f_a(M(a) + s) - f_a(M(a))| \rightarrow 0 \text{ as } h \rightarrow 0.$$

Further, (42) yields

$$T_n := \sup_{a \in \mathbb{S}_{d-1}} |\tilde{M}_n^\pm(a) - M(a)| \rightarrow 0 \quad \tilde{\mathbb{P}}\text{-a.s.}$$

in view of the definition of the  $Y$ -processes. Taken together this gives,  $\tilde{\mathbb{P}}\text{-a.s.}$

$$\sup_{a \in \mathbb{S}_{d-1}} |r(a, \tilde{M}_n^\pm(a) - M(a))| \leq \sup_{a \in \mathbb{S}_{d-1}} \sup_{|s| \leq T_n} |f_a(M(a) + s) - f_a(M(a))| \rightarrow 0,$$

and using (42) again we get

$$\tilde{B}_n^\pm := \sup_{a \in \mathbb{S}_{d-1}} |\tilde{Y}_n^\pm(a)| |r(a, \tilde{M}_n^\pm(a) - M(a))| \rightarrow 0 \quad \tilde{\mathbb{P}}\text{-a.s.}$$

From (41) and the continuity of  $a \rightarrow \tilde{Z}_0(a, M(a))$  it follows that, again  $\tilde{\mathbb{P}}\text{-a.s.}$ ,

$$\sup_{a \in \mathbb{S}_{d-1}} \left| \sqrt{n} \left( \tilde{F}_n(a, \tilde{M}_n^\pm(a)) - F(a, \tilde{M}_n^\pm(a)) \right) - \tilde{Z}_0(a, M(a)) \right| \rightarrow 0. \tag{45}$$

Absolute continuity of  $P$  implies that samples from  $P$  are in general position with probability 1, meaning that at most  $d$  points are on a hyperplane. Hence at most  $d$  of the variables  $a^t X_1, \dots, a^t X_n$  coincide, and thus, given that the range of  $\tilde{F}_n = F_n \circ g_n$  is contained in the range of  $F_n$ ,

$$|\tilde{F}_n(a, \tilde{M}_n^\pm(a)) - 1/2| \leq d/n.$$

with probability 1. Using this and (45) we get

$$\sup_{a \in \mathbb{S}_{d-1}} \left| \sqrt{n} \left( F(a, \tilde{M}_n^\pm(a)) - 1/2 \right) + \tilde{Z}_0(a, M(a)) \right| \rightarrow 0 \quad \tilde{\mathbb{P}}\text{-a.s.}$$

From

$$\sqrt{n} \left( F(a, \tilde{M}_n^\pm(a)) - 1/2 \right) = \tilde{Y}_n^\pm(a) \left( f_a(M(a)) + r(a, \tilde{M}_n^\pm(a) - M(a)) \right),$$

which holds by (44), we obtain that,  $\tilde{\mathbb{P}}\text{-a.s.}$ ,

$$\tilde{A}_n^\pm := \sup_{a \in \mathbb{S}_{d-1}} \left| \tilde{Y}_n^\pm(a) \left( f_a(M(a)) + r(a, \tilde{M}_n^\pm(a) - M(a)) \right) + \tilde{Z}_0(a, M(a)) \right| \rightarrow 0.$$

Thus,

$$\sup_{a \in \mathbb{S}_{d-1}} \left| \tilde{Y}_n^\pm(a) f_a(M(a)) + \tilde{Z}_0(a, M(a)) \right| \leq \tilde{A}_n^\pm + \tilde{B}_n^\pm \rightarrow 0 \quad \tilde{\mathbb{P}}\text{-a.s.},$$

which finishes the proof of (43) as  $a \mapsto f_a(M(a))$  is bounded away from 0 on  $\mathbb{S}_{d-1}$ .

Returning to the un-tilde variables we see that this establishes the functional limit theorems in (38).

We note in passing that  $\sup_{a \in \mathbb{S}_{d-1}} |\tilde{Y}_n^+(a) - \tilde{Y}_n^-(a)| \rightarrow 0 \quad \tilde{\mathbb{P}}\text{-a.s.}$ , which implies

$$\sup_{a \in \mathbb{S}_{d-1}} \left| \sqrt{n} \left( M_n^+(a) - M_n^-(a) \right) \right| \rightarrow 0 \quad \text{in } \mathbb{P}\text{-probability.} \tag{46}$$

This shows that the volume of the interval  $[\text{MMED}_-(P_n^{U_n}), \text{MMED}_+(P_n^{U_n})]$  shrinks fast enough so that the choice of a vector from this interval is irrelevant for the distributional asymptotics.

Now let  $b_n(i)^t$  and  $b(i)^t$  be the  $i$ th row of  $U_n$ ,  $n \in \mathbb{N}$ , and  $U$  respectively, for  $i = 1, \dots, d$ . In view of  $U_n \rightarrow U$  in  $\mathbb{O}(d)$  we have  $b_n(i) \rightarrow b(i)$   $\mathbb{P}$ -a.s. for  $i = 1, \dots, d$ . Let  $W_n := \sqrt{n}\eta_n = U_n\sqrt{n}(\theta_n - \theta)$  and

$$W_n^\pm := \sqrt{n} \begin{pmatrix} M_n^\pm(b_n(1)) \\ \vdots \\ M_n^\pm(b_n(d)) \end{pmatrix}, \quad W := \begin{pmatrix} Y(b(1)) \\ \vdots \\ Y(b(d)) \end{pmatrix}.$$

Note that  $W_n$  is an element of the  $d$ -dimensional interval  $[W_n^-, W_n^+]$  with probability 1. As the paths of the limit process are continuous the functional limit theorems (38) now imply  $W_n^\pm \rightarrow_d W$ , hence

$$W_n \rightarrow_d W, \tag{47}$$

with  $W \sim N_d(0, \mathcal{E})$ , where the entries of  $\mathcal{E}$  can be calculated from (39) and (40). Indeed, as  $M \equiv 0$  and as  $(0, U)$  solves the quarter median problem for  $P$ , we obtain



$K(b_i, b_i) = 1/4$  and  $K(b_i, b_j) = 0$  if  $i \neq j$ , so that  $\mathcal{E} = \Delta(P, U)$  with  $\Delta(P, U)$  the diagonal matrix given in (17).

We now apply  $U_n^t$  on the left hand side,  $U^t$  on the right hand side of (47) and use (Billingsley 1968, Theorem 5.5) again to obtain the claimed asymptotic normality of  $\sqrt{n}(\theta_n - \theta)$ , together with the mean and covariance matrix of the limit distribution.

It remains to show that  $\Sigma(P, U)$  does not depend on the choice of  $U$  from the set  $H$ . This follows easily from the fact that, for diagonal matrices  $\Delta$ , we always have  $P^t \Delta P = \Delta$  if  $P$  corresponds to the permutation of two rows or the multiplication by  $-1$  of some row.

For the proof the assumption (16) is important. Without such an assumption we would need the decomposition of  $\sqrt{n}(M_n^\pm(b_n(i))) - M(b(i))$  into two terms  $\sqrt{n}(M_n^\pm(b_n(i))) - M(b_n(i))$  and  $\sqrt{n}(M(b_n(i))) - M(b(i))$ . The first could still be analyzed with the above arguments, but for the second term the local behaviour of the function  $a \mapsto M(a)$  and distributional asymptotics of the estimates  $U_n$  would become important; see also Theorem 6.

Further, (46) and the ensuing comment show that the original estimate  $\theta_n$  need not be permutation invariant. This may seem surprising as the proof relies on empirical process theory, which deals with the data through their empirical distribution. Here it turned out to be enough that the estimate can be squeezed between two other estimates that are such functions of the empirical distribution.

### 4.5 Proof of Theorem 4

The transformation law for  $d$ -dimensional elliptical distributions, together with the equivariance from Proposition 1 (a), means that we may assume that  $\mu = 0$  and that  $\Sigma = \text{diag}(\lambda_1, \dots, \lambda_d)$ . Recall that  $e_1, \dots, e_d$  is the canonical basis of  $\mathbb{R}^d$ . For a subset  $I$  of  $\{1, \dots, d\}$  let

$$A_I := \{x \in \mathbb{R}^d : x^t e_i \geq 0 \text{ for } i \in I, x^t e_i \leq 0 \text{ for } i \notin I\}.$$

The assumptions imply that  $P^T = P$  for all reflections  $T : \mathbb{R}^d \rightarrow \mathbb{R}^d$  about hyperplanes  $H(e_i)$ , hence  $P(A_I)$  does not depend on  $I$ ; also,  $\sum_{I \subset \{1, \dots, d\}} P(A_I) = 1$ . A quarter space such as  $\{x \in \mathbb{R}^d : x^t e_i \geq 0, x^t e_j \geq 0\}$  with  $i \neq j$  can be written as the union of  $2^{d-2}$  such sets  $A_I$ , where the intersections are all  $P$ -null sets, hence all of these have probability  $2^{d-2} 2^{-d} = 1/4$ . This shows that  $(0, \Sigma)$  solves the quarter median problem for  $P$ .

In order to see that the quarter median is unique it is enough to show that  $\text{MMED}(P^U) = 0$  for all  $U \in \mathbb{O}(d)$  under the above assumptions on  $P$ . This follows from the fact that  $\text{Ell}_d(0, \Sigma; h)$  is invariant under the reflection  $x \mapsto -x, x \in \mathbb{R}^d$ .

It remains to show that the solution of the quarter median is unique in the strictly elliptical case. To prove this, we recall that  $\mu = 0$  and that  $\Sigma = \text{diag}(\lambda_1, \dots, \lambda_d)$ . Assume that  $(0, [U])$  with  $U \in \mathbb{O}(d)$  is a solution of the quarter median problem. Let  $b_1^t, \dots, b_d^t$  be the row vectors of  $U$ . Then  $P^U = \text{Ell}_d(0, \tilde{\Sigma}; h)$ , where  $\tilde{\Sigma} = (\tilde{\sigma}_{ij})_{1 \leq i, j \leq d} = U \Sigma U^t$ . With  $U_{ij}$  as the  $2 \times d$  matrix with the rows  $b_i^t, b_j^t, 1 \leq i < j \leq d$ , we see that  $P_{ij} := P^{U_{ij}}$  is the elliptical distribution  $\text{Ell}_2(0, \tilde{\Sigma}_{ij}; h_2)$  with

dispersion matrix

$$\tilde{\Sigma}_{ij} = U_{ij} \Sigma U_{ij}^t = \begin{pmatrix} \tilde{\sigma}_{ii} & \tilde{\sigma}_{ij} \\ \tilde{\sigma}_{ji} & \tilde{\sigma}_{jj} \end{pmatrix}$$

and density generator  $h_2$  which is equal to  $h$  if  $d = 2$  and given by

$$h_2(u) = \frac{\pi^{d/2-1}}{\Gamma(d/2 - 1)} \int_u^\infty (t - u)^{d/2-2} h(t) dt \text{ for } u > 0,$$

if  $d \geq 3$ ; see, e.g. Fang et al. (1990, Section 2.2.3). According to Grübel (1996, p. 1466) the probability assigned by  $P_{ij}$  to the left lower quadrant is given by

$$\frac{1}{4} + \frac{1}{2\pi} \arcsin \left( \tilde{\sigma}_{ij} / (\tilde{\sigma}_{ii} \tilde{\sigma}_{jj})^{1/2} \right),$$

which is equal to  $\frac{1}{4}$  if and only if  $\tilde{\sigma}_{ij} = 0$ , i.e. if and only if  $\tilde{\Sigma}_{ij}$  is a diagonal matrix with positive diagonal elements. Therefore,  $\tilde{\Sigma} = U \Sigma U^t$  is a diagonal matrix,  $\tilde{\Sigma} = \text{diag}(\tilde{\lambda}_1, \dots, \tilde{\lambda}_d)$  say, or equivalently,  $\Sigma = U^t \tilde{\Sigma} U = U^t \text{diag}(\tilde{\lambda}_1, \dots, \tilde{\lambda}_d) U$ . It follows from this that the set of eigenvalues  $\{\tilde{\lambda}_1, \dots, \tilde{\lambda}_d\}$  of  $\tilde{\Sigma}$  and the set of eigenvalues  $\{\lambda_1, \dots, \lambda_d\}$  of  $\Sigma$  coincide and, as a consequence, that there is a permutation matrix  $\Pi$  such that  $\Pi^t \Sigma \Pi = \tilde{\Sigma}$ , which in turn implies  $U^t \Pi^t \Sigma U \Pi = \Sigma$ . Thus, putting  $U^t \Pi^t =: Q = (q_{ij})$  we have  $Q \Sigma = \Sigma Q^t$ , i.e.  $q_{ij} \lambda_j = \lambda_i q_{ij}$ . As the  $\lambda_j$  are pairwise distinct this gives  $q_{ij} = 0$  for  $i \neq j$ . Thus  $Q$  is a diagonal matrix with the entries  $\pm 1$  in the diagonal so that  $U \sim I_d$ . Taken together this shows that  $(0, [I_d])$  is the unique solution of the quarter median problem.

### 4.6 Proof of Theorem 5

(a) We first assume that  $\mu = 0$  and that  $\Sigma = \text{diag}(\lambda_1, \dots, \lambda_d)$ . Using Theorem 3 we only need to evaluate the diagonal elements of  $\Delta(P, U)$ , and because of Theorem 4 we may take  $U$  to be the identity matrix. In particular,  $b_i = e_i$  for  $i = 1, \dots, d$ . Then  $f_{e_i}$  is the  $i$ th marginal density of  $P$ , and the calculations preceding Theorem 4 lead to

$$f_{e_i}(M(e_i)) = \frac{c_{d-1}(h)}{\sqrt{\lambda_i} c_d(h)}, \quad i = 1, \dots, d.$$

Putting this together we see that, for such  $P$ , the asymptotic covariance matrix is indeed the specified multiple of  $\Sigma$ .

We now use the equivariance properties of the quarter median to ‘bootstrap’ this to general strictly elliptical distributions. Hence suppose that  $P = \text{Ell}_d(\mu, \Sigma; h)$  and that  $\lambda_1 > \lambda_2 > \dots > \lambda_d$  are the eigenvalues of  $\Sigma$ . Then, for some  $U \in \mathbb{O}(d)$ ,  $\Sigma = U^t \text{diag}(\lambda_1, \dots, \lambda_d) U$ . If  $X_i, i \in \mathbb{N}$ , are independent with distribution  $P$  then the random variables  $Y_i := U(X_i - \mu), i \in \mathbb{N}$ , are independent with distribution  $\text{Ell}_d(0, \text{diag}(\lambda_1, \dots, \lambda_d); h)$ . As  $U(\theta_n - \mu)$  is a quarter median of  $Y_1, \dots, Y_n$  the first

part of the proof leads to

$$\sqrt{n} U(\theta_n - \mu) \rightarrow_d N_d(0, \mathcal{E}),$$

with  $\mathcal{E} = \sigma_{\text{QMED}}^2 \text{diag}(\lambda_1, \dots, \lambda_d)$ . Using the well-known properties of weak convergence and the behavior of multivariate normal distributions under affine transformations we finally obtain (24).

(b) Suppose that the eigenvalues of  $\Sigma$  coincide and let  $(U_n)_{n \in \mathbb{N}}$  be as in the statement of the second part of the theorem. Then, for any subsequence  $(U_{n'})$  there exists a subsubsequence  $(U_{n''})$  such that  $U_{n''} \rightarrow U_0$  for some  $U_0 \in \mathbb{O}(d)$  as  $n'' \rightarrow \infty$ . With  $m$  for  $n''$  and  $b_m(i)^t, b(i)^t$  the  $i$ th row of  $U_m$  respectively  $U_0$ , and using the same arguments as for (47), we obtain

$$\sqrt{m} \begin{pmatrix} M_m(b_m(1)) \\ \vdots \\ M_m(b_m(d)) \end{pmatrix} \rightarrow_d \begin{pmatrix} Y(b(1)) \\ \vdots \\ Y(b(d)) \end{pmatrix} \text{ as } m \rightarrow \infty.$$

The symmetry of  $P$  implies that the limit distribution does not depend on  $U_0$ .

### 4.7 Proof of Theorem 6

As the result is somewhat tangential to the topic of multivariate location estimation we only provide the main ideas together with some arguments specific to the quarter median application.

The general approach consists in writing the parameter  $\theta = \theta(P)$  as the solution of an equation  $\Psi_P(\theta) = z$  with a suitably chosen function  $\Psi_P$  and known  $z$ . Let  $\Psi_n$  be the empirical version of  $\Psi_P$  and suppose that  $\Psi_n(\theta_n)$  sufficiently close to  $z$  for large  $n \in \mathbb{N}$ . Then, under certain conditions, properties of the estimator sequence  $(\theta_n)_{n \in \mathbb{N}}$  can be obtained by localizing  $\Psi_P$  near  $\theta$  and then using the delta method. An excellent general exposition of this approach can be found in the textbook of van der Vaart (1998), see also van der Vaart (2002). Specifically, van der Vaart(1998, Theorem 5.9) together with van der Vaart (2002, Theorem 6.17) may serve as a guide towards filling in the details. Ultimately, this would also lead to an alternative proof for Theorem 5 in the case  $d = 2$  that does not make use of the results presented by Grübel (1996).

For  $a, b > 0$  let

$$E(a, b) := \left\{ (x, y)^t \in \mathbb{R}^2 : \frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 \right\} = \{ (a \cos(t), b \sin(t))^t : 0 \leq t < 2\pi \}$$

be the boundary of the two-dimensional centered ellipse with main half-axes parallel to the unit vectors  $e_1$  and  $e_2$  and of length  $a$  and  $b$  respectively. On this set we consider the push-forward  $Q(a, b) := \text{unif}(\mathbb{S}_1)^{T_{a,b}}$  of the uniform distribution on  $\mathbb{S}_1 = E(1, 1)$  under the mapping  $T_{a,b} : \mathbb{R}^2 \rightarrow \mathbb{R}^2, (x, y)^t \mapsto (ax, by)^t$ . If  $a \neq b$  then the solution of the quarter median problem for  $Q(a, b)$  is unique and given by  $\mu = (0, 0)^t$  and  $U = \text{diag}(1, 1) = U_\alpha$  with  $\alpha = 0$ .

Let  $P$  be a probability measure on the Borel subsets of  $\mathbb{R}^2$ . We introduce the function  $\Psi_P : \mathbb{R} \times \mathbb{R} \times [-\pi/4, \pi/4] \rightarrow \mathbb{R}^3$ , where the components  $\Psi_{P,i}$  are the respective probabilities of the upper right, upper left and lower left quarter spaces associated with the argument. Formally, and with  $b_\alpha, b'_\alpha$  the columns of  $U_\alpha^t$ ,

$$\Psi_{P,1}(x, y, \alpha) = P(V_{++}((x, y)^t; b_\alpha, b'_\alpha)),$$

and  $\Psi_{P,2}(x, y, \alpha) = P(V_{-+}(\dots)), \Psi_{P,3}(x, y, \alpha) = P(V_{--}(\dots))$ .

Suppose that  $P = Q(a, b)$  with  $a \neq b$ . Then elementary geometric considerations lead to the following matrix of partial derivatives of  $\Psi_P$  at  $(x, y, z) = (0, 0, 0)$ ,

$$M(a, b) := \frac{1}{2\pi ab} \begin{pmatrix} -b & b & b \\ -a & -a & a \\ b-a & a-b & b-a \end{pmatrix}. \tag{48}$$

For later use we note that

$$M(a, b)^{-1} := \pi \begin{pmatrix} 0 & -b & ab/(b-a) \\ a & -b & 0 \\ a & 0 & ab/(b-a) \end{pmatrix}.$$

In order to obtain the derivative for more general  $P$  we use some structural properties of symmetric and elliptical distributions. Note that  $Q(r, r)$  is the uniform distribution on the boundary of the sphere with radius  $r$ . If  $X \sim \text{Sym}_2(h)$  then  $\|X\|$  has density

$$f_{\|X\|}(r) = \frac{2\pi}{c_2(h)} r h(r^2), \quad r > 0,$$

and given  $\|X\| = r$ ,  $X$  is uniformly distributed on  $\{x \in \mathbb{R}^2 : \|x\| = r\}$ . Taken together this implies the mixture representation

$$\text{Sym}_2(h) = \frac{2\pi}{c_2(h)} \int_0^\infty Q(r, r) r h(r^2) dr.$$

Elliptical distributions are affine transformations of symmetric distributions. In particular,

$$P := \text{Ell}_2((0, 0)^t, \text{diag}(a^2, b^2); h) = \text{Sym}_2(h)^{T_{a,b}},$$

so that

$$P = \frac{2\pi}{c_2(h)} \int_0^\infty Q(ar, br) r h(r^2) dr.$$

Inserting quarter spaces we obtain a relation between  $\Psi_P$  and the  $\Psi$ -functions corresponding to the measures  $Q(a, b)$ ,  $a, b > 0$ . This in turn can be used to relate the derivative  $D(P)$  of  $\Psi_P$  at  $(0, 0, 0)$  to the derivatives obtained earlier for  $Q(a, b)$ , and we arrive at

$$D(P) = \frac{2\pi}{c_2(h)} \int_0^\infty M(ar, br) r h(r^2) dr = \frac{\pi c_1(h)}{c_2(h)} M(a, b),$$

and thus

$$D(P)^{-1} = \frac{c_2(h)}{c_1(h)} \begin{pmatrix} 0 & -b & ab/(b-a) \\ a & -b & 0 \\ a & 0 & ab/(b-a) \end{pmatrix}.$$

If  $(x, y, \alpha)$  solves the quarter median problem for an absolutely continuous  $P$ , then  $\Psi_P(x, y, \alpha) = (1/4, 1/4, 1/4)$ . For a sample of size  $n$  from  $P$  the random vector of observation counts in the quarter spaces associated with the solution of the quarter median problem for  $P$  has a multinomial distribution with parameters  $n$  and  $(1/4, 1/4, 1/4, 1/4)$ . The multivariate central limit theorem shows that the standardized counts in the three quarter spaces corresponding to the components of  $\Psi$  are asymptotically normal with mean vector 0 and covariance matrix

$$\Sigma_0 := \begin{pmatrix} 3/16 & -1/16 & -1/16 \\ -1/16 & 3/16 & -1/16 \\ -1/16 & -1/16 & 3/16 \end{pmatrix}.$$

We may now apply the delta method to obtain asymptotic normality for the triplet consisting of the quarter median coordinates and the rotation angle for a sequence of independent random variables with distribution  $P = \text{Ell}_2((0, 0)^t, \text{diag}(a^2, b^2); h)$ . The limiting normal distribution is centered and has covariance matrix

$$\Sigma_1 = (D(P)^{-1})^t \Sigma_0 D(P)^{-1} = \frac{c_2(h)^2}{4c_1(h)^2} \begin{pmatrix} a^2 & 0 & 0 \\ 0 & b^2 & 0 \\ 0 & 0 & a^2b^2/(a-b)^2 \end{pmatrix}.$$

It remains to extend this to  $P = \text{Ell}_2(\mu, \Sigma; h)$ . We use essentially the same argument as at the end of Sect. 4.6 (a). By assumption and Theorem 4, we have that  $\Sigma = U_\alpha^t \text{diag}(\lambda_1, \lambda_2) U_\alpha$ , where  $U_\alpha$  is the matrix representing the rotation by the angle  $\alpha$  in the interior of the half-open interval  $I$ , see (8), and that  $(\mu, U_\alpha)$  is the unique solution of the quarter median problem for  $P$ . Theorem 2 implies that  $\alpha_n - \alpha \in (-\pi/4, +\pi/4)$  with probability 1 for  $n$  sufficiently large. Let  $Y_i := U_\alpha^t(X_i - \mu), i \in \mathbb{N}$ . Then, with probability 1 for  $n$  large enough,  $(U_\alpha^t(\text{QMED}(X_1, \dots, X_n) - \mu), U_{\alpha_n(X_1, \dots, X_n)} - \alpha)$  is a solution of the quarter median problem for  $Y_1, \dots, Y_n$ . Noting that  $\lambda_1 = a^2$  and  $\lambda_2 = b^2$  we may now apply the above asymptotic normality result to the  $Y$ -sequence to obtain

$$\begin{pmatrix} U_\alpha^t & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \sqrt{n}(\text{QMED}(X_1, \dots, X_n) - \mu) \\ \sqrt{n}(\alpha_n(X_1, \dots, X_n) - \alpha) \end{pmatrix} \rightarrow_d N_3(0, \Sigma_1).$$

Consequently,

$$\begin{pmatrix} \sqrt{n}(\text{QMED}(X_1, \dots, X_n) - \mu) \\ \sqrt{n}(\alpha_n(X_1, \dots, X_n) - \alpha) \end{pmatrix} \rightarrow_d N_3 \left( 0, \sigma_{\text{QMED}}^2 \begin{pmatrix} \Sigma & 0 \\ 0 & \frac{\lambda_1 \lambda_2}{(\sqrt{\lambda_1} - \sqrt{\lambda_2})^2} \end{pmatrix} \right), \tag{49}$$

which is the statement of the theorem.

### 4.8 Proof of Proposition 2

We denote by  $\mathcal{C}$  the set of non-empty closed subsets of  $\mathbb{R}^d \times \mathbb{SO}(d)$ . For  $x_1, \dots, x_n \in \mathbb{R}^d$  let  $\Gamma(x_1, \dots, x_n)$  be the (by Theorem 1 non-empty) set of solutions  $(\theta, U) \in \mathbb{R}^d \times \mathbb{SO}(d)$  of the quarter median problem for  $P_{n;x_1, \dots, x_n}$ . As permutations of  $x_1, \dots, x_n$  do not change  $P_{n;x_1, \dots, x_n}$  we may regard  $\Gamma$  as a function on  $E := (\mathbb{R}^d)^n / \sim$ , where  $(x_1, \dots, x_n), (y_1, \dots, y_n) \in (\mathbb{R}^d)^n$  are equivalent if  $y_i = x_{\pi(i)}$  for  $i = 1, \dots, n$  for some permutation  $\pi$  of the set  $\{1, \dots, n\}$ . We aim to prove that the function  $\Gamma$  has values in  $\mathcal{C}$  and that it is measurable in the sense that the set

$$A(C) := \{z \in E : \Gamma(z) \cap C \neq \emptyset\} \tag{50}$$

is a Borel set for each  $C \in \mathcal{C}$ . Then, by the measurable selection theorem of Kuratowski and Ryll–Nardzewski (see, e.g. Rockafellar 1976, 1.C Corollary) there exists a measurable function  $\tau : E \rightarrow \mathbb{R}^d \times \mathbb{SO}(d)$  such that  $\tau(z) \in \Gamma(z)$  for all  $z \in E$ . In fact, we will prove that the set  $A(C)$  is closed whenever  $C \in \mathcal{C}$ .

Arguing as in the first part of the proof of Theorem 1 in Sect. 4.1 we see that if  $(\theta_\ell, U_\ell)_{\ell \in \mathbb{N}}$  is a sequence of elements in  $\Gamma(z)$  converging to  $(\theta, U) \in \mathbb{R}^d \times \mathbb{SO}(d)$  as  $\ell \rightarrow \infty$ , then  $(\theta, U) \in \Gamma(z)$ . This shows that the function  $\Gamma$  has values in  $\mathcal{C}$ .

Now, more generally, fix  $C \in \mathcal{C}$ , let  $A(C)$  be as in (50), and let  $(z_\ell)_{\ell \in \mathbb{N}}$  be a sequence of elements of  $A(C)$  that converges to some element  $z \in E$  as  $\ell \rightarrow \infty$ . As pointed out in connection with the transition from  $\mathbb{O}(d)$  to  $\mathbb{H}(d)$  in Sect. 4.4 we can find  $(x_1^{(\ell)}, \dots, x_n^{(\ell)}) \in z_\ell, l \in \mathbb{N}$ , and  $(x_1, \dots, x_n) \in z$  such that  $(x_1^{(\ell)}, \dots, x_n^{(\ell)})$  converges to  $(x_1, \dots, x_n) \in z$  in the space  $(\mathbb{R}^d)^n$ . Then for each  $\ell \in \mathbb{N}$  there exists an associated element  $(\theta_\ell, U_\ell)$  of  $\Gamma(x_1^{(\ell)}, \dots, x_n^{(\ell)})$ . Arguing as in Sect. 4.1 again, we see that along some subsequence  $(\ell')$  it holds that  $(\theta_{\ell'}, U_{\ell'}) \rightarrow (\theta, U) \in \mathbb{R}^d \times \mathbb{SO}(d)$  and

$$P_{n;x_1^{(\ell')}, \dots, x_n^{(\ell')}}^{U_{\ell'}} \rightarrow P_{n;x_1, \dots, x_n}^U \quad \text{weakly.}$$

With  $b_1^t, \dots, b_d^t$  as the row vectors of  $U$  this gives  $P_{n;x_1, \dots, x_n}(V_{\pm\pm}(\theta; b_i, b_j)) \geq \frac{1}{4}$ . Consequently,  $(\theta, U) \in \Gamma(z)$ ; also,  $z \in A(C)$  as  $C$  is closed.

### 4.9 Proof of Theorem 7

Let  $\theta$  be a quarter median of  $P_n$ , with associated orthogonal directions  $b, b' \in S_2$ . Let  $L_{\theta,b} = \{\theta + tb : t \in \mathbb{R}\}$ ,  $L_{\theta,b'} = \{\theta + tb' : t \in \mathbb{R}\}$  be the axes of the rectangular coordinate system with origin  $\theta$  and directions  $b, b'$ . We consider two cases.

(i)  $\theta \in \{x_1, \dots, x_n\}$ . We may assume without loss of generality that  $\theta = x_1$ , and define

$$\psi = \min(\min\{\arccos(|b_{1j}^t b|) : 2 \leq j \leq n\}, \min\{\arccos(|b_{1j}^t b'|) : 2 \leq j \leq n\}).$$

There are orthogonal directions  $b_*, b'_* \in S_1$  obtained from  $b, b'$  by the same clockwise or counterclockwise rotation with the angle  $\psi$ , where  $b_*$  or  $b'_*$  is an element of  $\{-b_{12}, b_{12}, \dots, -b_{1n}, b_{1n}\}$ . By construction, the number of data points  $x_j$  lying

in  $V_{\pm\pm}(\theta; b_*, b'_*)$  is greater than or equal to the number of data points  $x_j$  lying in  $V_{\pm\pm}(\theta; b, b')$ ; also the number of data points  $x_j$  lying in  $H_{\pm}(\theta; b_*)$  and the number of data points  $x_j$  lying in  $H_{\pm}(\theta; b'_*)$  is greater than or equal to the number of data points  $x_j$  lying in  $H_{\pm}(\theta; b)$  and the number of data points  $x_j$  lying in  $H_{\pm}(\theta; b')$ , respectively. It follows that  $(\theta, U_*)$  with  $U_*$  as the matrix with the row vectors  $b_*, b'_*$  solves the quarter median problem for  $P_n$ . Thus by Proposition 3 (c) there is an  $\eta \in \mathcal{M}(P_n^{U_*})$  such that  $(\theta_*, U_*)$  with  $\theta_* = U_*^t \eta$  also solves the quarter median problem for  $P_n$ . Note that  $\eta = \alpha b_* + \beta b'_*$ , where  $\alpha$  is median of  $P_n^{b_*}$  and  $\beta$  is a median of  $P_n^{b'_*}$ . Because  $b_*$  or  $b'_*$  is an element of  $\{-b_{12}, b_{12}, \dots, -b_{1n}, b_{1n}\}$ , and

$$\text{MED}_{\pm} \left( P_n^{-b_*} \right) = -\text{MED}_{\mp} \left( P_n^{b_*} \right), \quad \text{MED}_{\pm} \left( P_n^{-b'_*} \right) = -\text{MED}_{\mp} \left( P_n^{b'_*} \right), \quad (51)$$

it follows that  $(\theta_*, U_*) \in \mathcal{L}(P_n)$ .

(ii)  $\theta \notin \{x_1, \dots, x_n\}$ . With  $\tilde{\theta}$  as the intersection point of the line parallel to  $L_{\theta,b}$  through the data point  $\tilde{x} \in \{x_1, \dots, x_n\}$  with the shortest distance to  $L_{\theta,b}$ , i.e.  $\inf\{\|\tilde{x} - y\| : y \in L_{\theta,b}\} = \min\{\inf\{\|x_j - y\| : y \in L_{\theta,b}\} : 1 \leq j \leq n\}$ , and the line parallel to  $L_{\theta,b'}$  through the data point  $\tilde{\tilde{x}} \in \{x_1, \dots, x_n\}$  with the shortest distance to  $L_{\theta,b'}$ , we have that  $(\tilde{\theta}, U)$  with  $U$  as matrix with the row vectors  $b, b'$  is a solution of the quarter median problem for  $P_n$ . If  $\tilde{x} = \tilde{\tilde{x}}$  then  $\tilde{\theta} = \tilde{x} = \tilde{\tilde{x}}$  and we arrive at case (i). So, we can (and do) assume that there are two points,  $x_1, x_2$  say, such that  $x_1 \in L_{\theta,b}, x_2 \in L_{\theta,b'}$ . Now we put  $B := \{b_{1j} : 2 \leq j \leq n\} \cup \{b_{2j} : 3 \leq j \leq n\}$  and define

$$\psi = \min(\min\{\arccos(|d^t b|) : d \in B\}, \min\{\arccos(|d^t b'|) : d \in B\}).$$

There are orthogonal directions  $b_*, b'_* \in \mathbb{S}_1$  obtained from  $b, b'$  by the same clockwise or counterclockwise rotation with the angle  $\psi$ , where  $b_*$  or  $b'_*$  is an element of  $B \cup (-B)$ . With  $L_1 = \{x_1 + tb_* : t \in \mathbb{R}\}$  and  $L_2 = \{x_2 + tb'_* : t \in \mathbb{R}\}$  we obtain the axes of a new rectangular coordinate system. The origin  $\theta$  of the coordinate system with the axes  $L_{\theta,b}$  and  $L_{\theta,b'}$  and the origin  $\theta_*$  of the new system are both located on the Thales circle about the line segment connecting  $x_1$  and  $x_2$ . By construction, the number of data points  $x_j$  lying in  $V_{\pm\pm}(\theta_*; b_*, b'_*)$  is greater than or equal to the number of data points  $x_j$  lying in  $V_{\pm\pm}(\theta; b, b')$ . Further, the number of data points  $x_j$  lying in  $H_{\pm}(\theta_*; b_*)$  and the number of data points  $x_j$  lying in  $H_{\pm}(\theta_*; b'_*)$  are greater than or equal to the number of data points  $x_j$  lying in  $H_{\pm}(\theta; b)$  and the number of data points  $x_j$  lying in  $H_{\pm}(\theta; b')$ , respectively. Consequently,  $(\theta_*, U_*)$  with  $U_*$  as matrix with the row vectors  $b_*, b'_*$  solves the quarter median for  $P_n$ . Again, by Proposition 3 (c), there is an  $\eta \in \mathcal{M}(P_n^{U_*})$  such that  $(\theta_{**}, U_*)$ , with  $\theta_{**} = U_*^t \eta$  also solves the quarter median problem for  $P_n$ . Here,  $b_*$  or  $b'_*$  is an element of  $B \cup (-B)$ , so that, again by (51), we have  $(\theta_{**}, U_*) \in \mathcal{L}(P_n)$ .

**Acknowledgements** Dietrich Morgenstern (1924-2007) was the founding father of the Institut für Mathematische Stochastik at the then Technische Universität Hannover, which is now the Leibniz Universität Hannover, and of two similar institutes in Freiburg and Münster; see the obituary of Baringhaus et al. (2008). Discussions with him were the starting point for the present paper. The authors also thank the Editor and the Reviewer for useful suggestions.

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

## Declarations

**Conflict of interest** The authors declare that they have 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

- Bai ZD, Chen XR, Miao BQ, Rao CR (1990) Asymptotic theory of least distances estimate in multivariate linear models. *Statistics* 21:503–519
- Baringhaus L, Grübel R, Henze N (2008) Nachruf auf Dietrich Morgenstern. *Jahresber Deutsch Math Ver* 110:101–113
- Billingsley P (1968) *Convergence of probability measures*. Wiley, New York
- Brown BM (1983) Statistical uses of the spatial median. *J R Stat Soc Ser B* 45:25–30
- Dhar S, Chaudhuri P (2011) On the statistical efficiency of robust estimators of multivariate location. *Stat Methodol* 8:113–128
- Donoho DL, Gasko M (1992) Breakdown properties of location estimates based on halfspace depth and projected outlyingness. *Ann Stat* 20:1803–1827
- Dudley R (1999) *Uniform central limit theorems*. Cambridge University Press, Cambridge
- Fang TH, Kotz S, Ng KW (1990) *Symmetric multivariate and related distributions*. Chapman and Hall, London
- Gómez E, Gómez-Villegas MA, Marín JM (1998) A multivariate generalization of the power exponential family of distributions. *Commun Stat Theory Methods* 27:589–600
- Grübel R (1996) Orthogonalization of multivariate location estimators: the orthomedian. *Ann Stat* 24:1457–1473
- Hall BC (2004) *Lie groups, Lie algebras, and representations*. Springer, New York
- Kuelbs J, Zinn J (2013) Empirical quantile CLTs for time-dependent data. In: *High dimensional probability*. *Progr Probab* 66. vol VI. Birkhäuser/Springer, Basel, pp 167–194
- Makeev VV (2007) Equipartition of a continuous mass distribution. *J Math Sci* 140:551–557
- Massé JC (2004) Asymptotics for the Tukey depth process, with an application to a multivariate trimmed mean. *Bernoulli* 10:397–419
- Massé JC, Plante JF (2003) A Monte Carlo study of the accuracy and robustness of ten bivariate location estimators. *Comput Stat Data Anal* 42:1–26
- Mitchell AFS (1989) The information matrix, skewness tensor and  $\alpha$ -connections for the general multivariate elliptic distribution. *Ann Inst Stat Math* 41:289–304
- Muirhead RJ (1982) *Aspects of multivariate statistical theory*. Wiley, New York
- Nolan D (1999) On min–max majority and deepest points. *Stat Probab Lett* 43:325–333
- Oja H (1983) Descriptive statistics for multivariate distributions. *Stat Probab Lett* 1:327–333
- Oja H (2013) Multivariate median. In: Becker C, Fried R, Kuhnt S (eds) *Robustness and complex data structures*. Springer, Heidelberg, pp 3–15
- Oja H, Niinimaa A (1985) Asymptotic properties of the generalized median in the case of multivariate normality. *J R Stat Soc Ser B* 47:372–377
- Rockafellar RT (1976) Integral functionals, normal integrands and measurable selections. In: *Nonlinear operators and the calculus of variations* (Summer School, Univ Libre Bruxelles, Brussels, 1975). *Lecture Notes in Mathematics* 543, Springer, Berlin, pp 157–207



- Small CG (1990) A survey of multidimensional medians. *Int Stat Rev* 58:263–277
- Somorčík J (2006) Tests using spatial median. *Aust J Stat* 35:331–338
- tom Dieck T (1991) *Topologie*. Walter de Gruyter & Co., Berlin
- Tukey JW (1975) Mathematics and the picturing of data. In: *Proceedings of the 1974 international congress of mathematics, Vancouver, Vol. 2*. pp 523–531
- van der Vaart AW (1998) *Asymptotic statistics*. Cambridge University Press, Cambridge
- van der Vaart AW (2002) Semiparametric statistics. In: Bolthausen E, Perkins E, van der Vaart A (eds) *Lectures on probability theory and statistics. Lecture Notes in Mathematics 1781*. Springer, Berlin
- van der Vaart AW, Wellner JA (1996) *Weak convergence and empirical processes*. Springer, New York

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