DOI 10.1007/s10986-025-09663-5
Lithuanian Mathematical Journal, Vol. 65, No. 1, January, 2025, pp. 14–49

Existence and uniqueness of weighted generalized ψ-estimators

Mátyás Barczy [a, 1] **and Zsolt Páles** [b, 2, 3]

[a] HUN-REN–SZTE Analysis and Applications Research Group, Bolyai Institute, University of Szeged,
Aradi vértanúk tere 1, H–6720 Szeged, Hungary
[b] Institute of Mathematics, University of Debrecen, Pf. 400, H–4002 Debrecen, Hungary
(e-mail: barczy@math.u-szeged.hu; pales@science.unideb.hu)

Received July 6, 2024; revised January 15, 2025

Abstract. We introduce the notions of generalized and weighted generalized ψ-estimators as unique points of sign change of some appropriate functions and give necessary and sufficient conditions for their existence. We also derive a set of sufficient conditions under which the so-called ψ-expectation function has a unique point of sign change. We present several examples from statistical estimation theory, where our results are well applicable. For example, we consider the cases of empirical quantiles, empirical expectiles, some ψ-estimators that are important in robust statistics, and some examples from maximum likelihood theory. Further, we introduce Bajraktarević-type (in particular, quasiarithmetic-type) ψ-estimators. Our results specialized to ψ-estimators with a function ψ continuous in its second variable provide new results for (usual) ψ-estimators (also called Z-estimators).

MSC: 62F10, 62F35, 26E60, 26A48

Keywords: M-estimator, ψ-estimator, Z-estimator, likelihood equation, existence, uniqueness, Bajraktarević-type ψ-estimator, empirical quantile, empirical expectile, robust statistics

1 Introduction

In statistics, M-estimators play a fundamental role, and a special subclass, the class of ψ-estimators (also called Z-estimators), is also in the heart of investigations. The M-estimators (where the letter M refers to "maximum likelihood-type") were introduced by Huber [11, 12]. Let $(X, \mathcal{X})$ be a measurable space, let Θ be a Borel subset of $\mathbb{R}$, and let $\varrho : X \times \Theta \to \mathbb{R}$ be a function such that for each $t \in \Theta$, the function $X \ni x \mapsto \varrho(x, t)$ is measurable with respect to the sigma-algebra $\mathcal{X}$. Let $(\xi_k)_{k\geqslant 1}$ be a sequence of i.i.d. random variables with values in X such that the distribution of ξ_1 depends on an unknown parameter $\vartheta \in \Theta$. For each $n \geqslant 1$, Huber [11, 12] introduced an estimator of ϑ based on the observations $\xi_1, \ldots, \xi_n$ as a solution

---

[1] The author was supported by the project TKP2021-NVA-09. Project No. TKP2021-NVA-09 has been implemented with the support provided by the Ministry of Culture and Innovation of Hungary from the National Research, Development and Innovation Fund, financed under the TKP2021-NVA funding scheme.
[2] The author is supported by the K-134191 NKFIH grant.
[3] Corresponding author.

0363-1672/25/6501-0014

$\widehat{\vartheta}_n := \widehat{\vartheta}_n(\xi_1,\dots,\xi_n)$ of the following minimization problem:

$$\inf_{t\in\Theta}\sum_{i=1}^{n}\varrho(\xi_i,t), \tag{1.1}$$

provided that such a solution exists. We call $\widehat{\vartheta}_n$ an M-estimator of the unknown parameter $\vartheta\in\Theta$ based on the i.i.d. observations $\xi_1,\dots,\xi_n$. In stochastic optimization, $\widehat{\vartheta}_n$ is called a sample average approximation (SAA) of a solution of the classical risk neutral stochastic program $\inf_{t\in\Theta}\mathbf{E}(\varrho(\xi_1,t))$; see, for example, Shapiro et al. [23, Chap. 5]. For historical fidelity, we note that Huber [11] considered the special case where $X:=\mathbb{R}$, $\Theta:=\mathbb{R}$, and the function ϱ depends only on $x-t$, that is, $\varrho(x,t):=f(x-t)$, $x\in\mathbb{R}$, $t\in\Theta$, with some given nonconstant function $f:\mathbb{R}\to\mathbb{R}$. Turning back to the general case, under suitable regularity assumptions, the minimization problem (1.1) can be solved by setting the derivative of the objective function (with respect to the unknown parameter) equal to zero:

$$\sum_{i=1}^{n}\partial_2\varrho(\xi_i,t)=0,\quad t\in\Theta,$$

where $\partial_2\varrho$ denotes the (partial) derivative of ϱ with respect to its second variable. In the statistical literature, $\partial_2\varrho$ is often denoted by ψ, and hence in this case an M-estimator is often called a ψ-estimator, whereas other authors call it a Z-estimator (the letter Z refers to "zero"). For a detailed exposition of M-estimators and ψ-estimators (Z-estimators), see, for example, Kosorok [15, Sects. 2.2.5 and 13] or van der Vaart [25, Sect. 5].

Throughout this paper, let $\mathbb{N}=\mathbb{Z}_{++}$, $\mathbb{Z}_{+}$, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{R}_{+}$, and $\mathbb{R}_{++}$ denote the sets of positive integers, nonnegative integers, rational numbers, real numbers, nonnegative real numbers, and positive real numbers, respectively. For a real number $y\in\mathbb{R}$, its positive and negative parts as well as its upper and lower integer parts are denoted by y^{+} and y^{-} as well as by $\lceil y\rceil$ and $\lfloor y\rfloor$, respectively. For a subset $S\subseteq\mathbb{R}$, the convex hull of S (which is the smallest interval containing S) is denoted by $\operatorname{conv}(S)$. For each $n\in\mathbb{N}$, let us also introduce the set $\Lambda_n:=\mathbb{R}_{+}^{n}\setminus\{(0,\dots,0)\}$. All the random variables are defined on an appropriate probability space $(\Omega,\mathcal{A},\mathbf{P})$.

To the best of our knowledge, the topic of existence and uniqueness of ψ-estimators is less addressed in the statistical literature. In the present paper, we investigate two basic problems of this field. Roughly speaking, Problem 1 is about the existence and uniqueness of the newly introduced notions, generalized ψ-estimators and weighted generalized ψ-estimators. Problem 2 is devoted to the existence and uniqueness of a point of sign change for so-called ψ-expectation functions.

Problem 1. Let X be a nonempty set, and let Θ be a nonempty open interval of $\mathbb{R}$. Let $\Psi(X,\Theta)$ denote the class of real-valued functions $\psi: X\times\Theta\to\mathbb{R}$ such that for each $x\in X$, there exist $t_{+},t_{-}\in\Theta$ such that $t_{+}<t_{-}$ and $\psi(x,t_{+})>0>\psi(x,t_{-})$. Roughly speaking, a function $\psi\in\Psi(X,\Theta)$ satisfies the following property: for each $x\in X$, the function $t\ni\Theta\mapsto\psi(x,t)$ changes sign (from positive to negative) on the interval Θ at least once. Given a function $\psi\in\Psi(X,\Theta)$, $n\in\mathbb{N}$, and $\boldsymbol{x}=(x_1,\dots,x_n)\in X^n$, let us consider the equation

$$\psi_{\boldsymbol{x}}(t):=\sum_{i=1}^{n}\psi(x_i,t)=0,\quad t\in\Theta. \tag{1.2}$$

More generally, for $n\in\mathbb{N}$, $\boldsymbol{x}=(x_1,\dots,x_n)\in X^n$ and $\boldsymbol{\lambda}=(\lambda_1,\dots,\lambda_n)\in\Lambda_n$, we also consider the weighted equation

$$\psi_{\boldsymbol{x},\boldsymbol{\lambda}}(t):=\sum_{i=1}^{n}\lambda_i\psi(x_i,t)=0,\quad t\in\Theta. \tag{1.3}$$

The basic question we are going to investigate now is to find necessary and sufficient conditions for the unique solvability of Eqs. (1.2) and (1.3), respectively. In a broader context, we are going to find necessary and sufficient conditions for the existence of a point of sign change (see Definition 1) for the functions $\psi_{\boldsymbol{x}}$

and $\psi_{\boldsymbol{x},\boldsymbol{\lambda}}$ introduced by (1.2) and (1.3), respectively. It will turn out that the points of sign change in question are unique, provided that they exist, and we can call them a *generalized ψ-estimator* and *weighted generalized ψ-estimator*, respectively, for some unknown parameter in Θ based on the realization $(x_1, \ldots, x_n) \in X^n$ and weights $(\lambda_1, \ldots, \lambda_n) \in \Lambda_n$. In Proposition 1, we also study the measurability of (weighted) generalized ψ-estimators, provided that X is a measurable space, but we emphasize that in our general setup, X is not necessarily a measurable space; it can be an arbitrary nonempty set. Concerning Problem 1, it does not matter whether the random variables $\xi_1, \ldots, \xi_n$ of which $(x_1, \ldots, x_n)$ is a realization are i.i.d. or not. As future research, we could investigate the asymptotic properties of the (weighted) generalized ψ estimators based on $(\xi_1, \ldots, \xi_n)$ as $n \to \infty$, when the i.i.d. property for the sequence $(\xi_k)_{k\geqslant 1}$ could play a role.

Problem 2. Let $(X, \mathcal{X})$ be a measurable space, let Θ be a nonempty open interval of $\mathbb{R}$, and let $\psi : X \times \Theta \to \mathbb{R}$ be a measurable function in its first variable, that is, for each $t \in \Theta$, the mapping $X \ni x \mapsto \psi(x, t)$ is measurable with respect to the sigma-algebra $\mathcal{X}$. Further, let $\xi : \Omega \to X$ be a random variable defined on a probability space $(\Omega, \mathcal{A}, \mathbf{P})$ such that $\mathbf{E}(|\psi(\xi, t)|) < \infty$ for each $t \in \Theta$. We investigate the question of existence of a unique point of sign change (see Definition 1) for the function

$$\Theta \ni t \mapsto \mathbf{E}\big(\psi(\xi, t)\big). \tag{1.4}$$

In the literature, we could not find a name for the function (1.4); however, we may call it a ψ-expectation function. Under appropriate conditions, the ψ-estimator (Z-estimator) based on i.i.d. observations $\xi_1, \ldots, \xi_n$ (where ξ_1 has the same law as that of ξ) is supposed to "well estimate" the zero of the function (1.4), provided that it exists uniquely; for more detail, see, for example, Kosorok [15, Sects. 2.2.5 and 13].

In what follows, we discuss the connections between Problems 1 and 2. From a stochastic optimization point of view, Problem 1 can be considered as a sample average approximation (SAA) of Problem 2, provided that $(x_1, \ldots, x_n)$ is a realization of i.i.d. random variables $\xi_1, \ldots, \xi_n$ (where ξ_1 has the same law as that of ξ). Under appropriate conditions, the generalized ψ-estimator based on i.i.d. observations $\xi_1, \ldots, \xi_n$ (where ξ_1 has the same law as that of ξ) is supposed to "well estimate" the point of sign change of the function (1.4), provided that it exists. In this paper, we do not investigate this question; it could be a topic of future research. Further, note that if ξ is a simple random variable such that $\mathbf{P}(\xi = x_i) = p_i$, $i = 1, \ldots, n$, where $n \in \mathbb{N}$, $(x_1, \ldots, x_n) \in X^n$, and $p_1, \ldots, p_n \geqslant 0$, $p_1 + \cdots + p_n = 1$, then $\mathbf{E}(\psi(\xi, t)) = \sum_{i=1}^{n} p_i \psi(x_i, t)$, $t \in \Theta$, and hence, in this setting, Problem 2 is a particular case of Problem 1.

To mention some papers related to Problem 1, we can refer, for example, to Huber [11, Lemma 1], Tibshirani [24], and Ali and Tibshirani [1]. Tibshirani [24] considered the lasso (least absolute shrinkage and selection operator) problem, which is also known as the ℓ_1-penalized linear regression. The lasso estimator is a popular tool in the theory of sparse linear regression; mathematically, it is a solution of a not necessarily strictly convex minimization problem, where a penalty term being the ℓ_1-norm of the coefficient vector comes into play. Tibshirani [24] studied the question of uniqueness of the lasso estimator. Recently, Ali and Tibshirani [1] have studied the uniqueness of a generalized lasso estimator, where the penalty term in the corresponding minimization problem is the ℓ_1-norm of a (penalty) matrix times the coefficient vector. To mention further papers related to Problem 2, we can refer to Huber [11, Lemma 2], Clarke [6] (for details, see Remark 3.8 in the arXiv version [2] of this paper), Mathieu [18] (for details, see Examples 4 and 8), and to the very recent paper of Dimitriadis et al. [7, Props. S1, S2, and S3], in which the authors, in particular, considered solvability of the equation $\mathbf{E}(\psi(\zeta, \eta, t)) = 0$, $t \in \Theta$, where $\psi : \mathbb{R} \times \mathbb{R} \times \Theta \to \mathbb{R}$ is a measurable function, Θ is a (nonempty) open parameter set of $\mathbb{R}$, and (ζ, η) is a response-regression (covariate) pair in some regression model.

Section 2 is devoted to the existence and uniqueness of (weighted) generalized ψ-estimators (Problem 1). First, we introduce the required terminology: the notions of point of sign change and level of increase for a real-valued function defined on a nonempty open interval (see Definitions 1 and 3) and the notions of properties $[T_n]$ and $[T_n^{\boldsymbol{\lambda}}]$ for a function in $\Psi(X, \Theta)$ (see Definition 2, but we also present below). We say that a function $\psi \in \Psi(X, \Theta)$ *possesses the property* $[T_n^{\boldsymbol{\lambda}}]$ *for some* $n \in \mathbb{N}$ *and* $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_n) \in \Lambda_n$ if there exists a mapping $\vartheta^{\boldsymbol{\lambda}}_{n,\psi} : X^n \to \Theta$ such that, for all $\boldsymbol{x} = (x_1, \dots, x_n) \in X^n$ and $t \in \Theta$,

$$\psi_{\boldsymbol{x},\boldsymbol{\lambda}}(t) = \sum_{i=1}^{n} \lambda_i \psi(x_i, t) \begin{cases} > 0 & \text{if } t < \vartheta^{\boldsymbol{\lambda}}_{n,\psi}(\boldsymbol{x}), \\ < 0 & \text{if } t > \vartheta^{\boldsymbol{\lambda}}_{n,\psi}(\boldsymbol{x}). \end{cases}$$

Note that if there exists such a mapping $\vartheta^{\boldsymbol{\lambda}}_{n,\psi}$, then it is unique. In case of $\lambda_i = 1$, $i = 1, \dots, n$ (or equivalently, in case of equal positive weights), the property $[T^{\boldsymbol{\lambda}}_n]$ is called property $[T_n]$. In the first main result of our paper (see Theorem 1), necessary and sufficient conditions are given for the properties $[T_n]$ and $[T^{\boldsymbol{\lambda}}_n]$. If ψ is continuous in its second variable as well, then such conditions imply the unique existence of the corresponding (usual) ψ-estimator. After Theorem 1, we present some properties of the property $[T^{\boldsymbol{\lambda}}_n]$. For example, Proposition 3 is about a connection between the property $[T_n]$ and the strict $(1/n)$-increasingness of some appropriately defined functions, and in Proposition 4, we establish a "grouping" property of the property $[T^{\boldsymbol{\lambda}}_n]$. Examples 1, 2, and 3 highlight the role of the conditions in Theorem 1 and Proposition 4. We also introduce a class of Bajraktarević-type (in particular, quasiarithmetic-type) ψ-estimators (motivated by the representation of Bajraktarević means as special deviation means) for which our results are well applicable; see Definition 4 and Proposition 6.

Section 3 is devoted to study the existence and uniqueness of the point of sign change of the ψ-expectation function given in (1.4) (Problem 2). As the second main result of our paper, in Theorem 2, we give a set of sufficient conditions in order that the ψ-expectation function in question have a unique point of sign change. We apply our results for ψ-expectation functions when ψ is a Bajraktarević-type function (see Proposition 8) and when ψ has the form used by Mathieu [18] (see Example 4 and Proposition 9).

In Section 4, we present several examples from statistical estimation theory that demonstrate the applicability of our results in Sections 2 and 3. These examples may be divided into three main groups. The first group of examples includes several well-known descriptive statistics that can be considered as special ψ-estimators. Namely, the empirical median (Example 5), the empirical quantiles (Example 6), and the empirical expectiles (Example 7). In particular, in Proposition 10, we show that, given $n \geqslant 2$, the function ψ corresponding to the empirical α-quantile has the property $[T_n]$ if and only if $\alpha \notin \{1/n, \dots, (n-1)/n\}$. The second group of examples contains the class of ψ-estimators recently used by Mathieu [18] (Example 8) and some ψ-estimators that are important in robust statistics. In particular, in Proposition 4.4, we derive necessary and sufficient conditions under which the function ψ used by Mathieu [18] has the property $[T^{\boldsymbol{\lambda}}_n]$ for all $n \in \mathbb{N}$ and $\boldsymbol{\lambda} \in \Lambda_n$. We emphasize that in all these examples, we investigate the existence and uniqueness of (weighted) generalized ψ-estimators, compared to the existing results that addressed ψ-estimators (Z-estimators). The third group of examples demonstrates the applicability of Theorem 1 together with Proposition 2 for proving the existence and uniqueness of solutions of likelihood equations. In Example 9, we consider the maximum likelihood estimator (MLE) of one of the parameters of a normally distributed random variable supposing that its other parameter is known. In Example 10, we consider a mixture density function of the standard normally density function and the density function of a normally distributed random variable with mean $m \in \mathbb{R}$ and variance $\sigma^2 > 0$ with equal weights $1/2$, and we study the solutions of the likelihood equation for m, provided that σ is known.

Sections 5, 6, and 7 are devoted to the proofs of the results in Sections 2, 3, and 4, respectively.

Now we summarize the novelties of the paper. We extensively discuss that, up to our knowledge, only few results are available for the existence and uniqueness of ψ-estimators and of the roots of ψ-expectation functions, and our paper can be considered as a theoretical contribution to this field. Another important feature of our paper is that we present a broad variety of examples from statistical estimation theory, where our results can be well applied.

Finally, we mention that in the literature, we can find ψ-estimators where the function ψ depends on the sample size n as well; see, for example, Hampel, Hennig and Ronchetti [10, Sect. 2]. As future research, we might try to generalize the notion of (weighted) generalized ψ-estimators and our results to this more general case. Another possible direction for future research is to explore the extension of our setup and results from a one-dimensional parameter set Θ to a multidimensional one (note that in our present setup, Θ is supposed to be a nonempty open interval of $\mathbb{R}$).

2 Notions and results on the existence and uniqueness of weighted generalized ψ-estimators

To investigate Problem 1 presented in the Introduction, we introduce the required terminology. First, we introduce the notion of a point of sign change for real-valued functions defined on an open interval.

DEFINITION 1. Let Θ be a nonempty open interval of $\mathbb{R}$. For a function $f:\Theta\to\mathbb{R}$, consider the following three level sets:

$$\Theta_{f>0}:=\big\{t\in\Theta\colon\ f(t)>0\big\},\qquad \Theta_{f=0}:=\big\{t\in\Theta\colon\ f(t)=0\big\},\qquad \Theta_{f<0}:=\big\{t\in\Theta\colon\ f(t)<0\big\}.$$

We say that $\vartheta\in\Theta$ is a *point of sign change* (*of decreasing type*) *for f* if

$$f(t)>0\quad\text{for } t<\vartheta \quad\text{and}\quad f(t)<0\quad\text{for } t>\vartheta.$$

Remark 1. Note that if $\vartheta\in\Theta$ is a point of sign change for f, then $\Theta_{f>0}$ and $\Theta_{f<0}$ are nonempty sets, and $\sup\Theta_{f>0}=\inf\Theta_{f<0}=\vartheta$. Furthermore, there can exist at most one element $\vartheta\in\Theta$ that is a point of sign change for f. If f is continuous at a point ϑ of sign change, then $f(\vartheta)=0$; moreover, $\Theta_{f=0}=\{\vartheta\}$. Conversely, as an easy consequence of the Bolzano theorem, if $f:\Theta\to\mathbb{R}$ is continuous, then the sets $\Theta_{f>0}$ and $\Theta_{f<0}$ are nonempty, and f has a unique zero $\vartheta\in\Theta$, i.e., $\Theta_{f=0}=\{\vartheta\}$, and then ϑ is a point of sign change either for f or for $(-f)$. The continuity of f, however, is not necessary for the existence of a point of sign change for f. For example, if f is strictly decreasing and the sets $\Theta_{f>0}$ and $\Theta_{f<0}$ are nonempty, then it is easy to see that there exists a point of sign change for f.

DEFINITION 2. We say that a function $\psi\in\Psi(X,\Theta)$

(i) *possesses the property* $[C]$ (*briefly, ψ is a C-function*) if it is continuous in its second variable, that is, if for all $x\in X$, the mapping $\Theta\ni t\mapsto\psi(x,t)$ is continuous.

(ii) *possesses the property* $[T_n]$ (*briefly, ψ is a T_n-function*) *for some* $n\in\mathbb{N}$ if there exists a mapping $\vartheta_{n,\psi}:X^n\to\Theta$ such that for all $\boldsymbol{x}=(x_1,\dots,x_n)\in X^n$ and $t\in\Theta$,

$$\psi_{\boldsymbol{x}}(t):=\sum_{i=1}^{n}\psi(x_i,t)\begin{cases}>0 & \text{if } t<\vartheta_{n,\psi}(\boldsymbol{x}),\\ <0 & \text{if } t>\vartheta_{n,\psi}(\boldsymbol{x}),\end{cases}$$

that is, for all $\boldsymbol{x}\in X^n$, the value $\vartheta_{n,\psi}(\boldsymbol{x})$ is a point of sign change for the function $\psi_{\boldsymbol{x}}$. If there is no confusion, instead of $\vartheta_{n,\psi}$ we simply write ϑ_n. We may call $\vartheta_{n,\psi}(\boldsymbol{x})$ a generalized ψ-estimator for some unknown parameter in Θ based on the realization $\boldsymbol{x}=(x_1,\dots,x_n)\in X^n$. If for each $n\in\mathbb{N}$, ψ is a T_n-function, then we say that ψ possesses the property $[T]$ (briefly, ψ is a T-function).

(iii) *possesses the property* $[Z_n]$ (*briefly, ψ is a Z_n-function*) *for some* $n\in\mathbb{N}$ if it is a T_n-function and

$$\psi_{\boldsymbol{x}}\big(\vartheta_{n,\psi}(\boldsymbol{x})\big)=\sum_{i=1}^{n}\psi\big(x_i,\vartheta_{n,\psi}(\boldsymbol{x})\big)=0\quad\text{for all } \boldsymbol{x}=(x_1,\dots,x_n)\in X^n.$$

If for each $n\in\mathbb{N}$, ψ is a Z_n-function, then we say that ψ possesses the property $[Z]$ (briefly, ψ is a Z-function).

(iv) *possesses the property* $[T_n^{\boldsymbol{\lambda}}]$ *for some* $n\in\mathbb{N}$ *and* $\boldsymbol{\lambda}=(\lambda_1,\dots,\lambda_n)\in\Lambda_n$ (*briefly, ψ is a $T_n^{\boldsymbol{\lambda}}$-function*) if there exists a mapping $\vartheta_{n,\psi}^{\boldsymbol{\lambda}}:X^n\to\Theta$ such that for all $\boldsymbol{x}=(x_1,\dots,x_n)\in X^n$ and $t\in\Theta$,

$$\psi_{\boldsymbol{x},\boldsymbol{\lambda}}(t):=\sum_{i=1}^{n}\lambda_i\psi(x_i,t)\begin{cases}>0 & \text{if } t<\vartheta_{n,\psi}^{\boldsymbol{\lambda}}(\boldsymbol{x}),\\ <0 & \text{if } t>\vartheta_{n,\psi}^{\boldsymbol{\lambda}}(\boldsymbol{x}),\end{cases}$$

that is, for all $\boldsymbol{x} \in X^n$, the value $\vartheta_{n,\psi}^{\boldsymbol{\lambda}}(\boldsymbol{x})$ is a point of sign change for the function $\psi_{\boldsymbol{x},\boldsymbol{\lambda}}$. If there is no confusion, instead of $\vartheta_{n,\psi}^{\boldsymbol{\lambda}}$, we simply write $\vartheta_{n}^{\boldsymbol{\lambda}}$. We may call $\vartheta_{n,\psi}^{\boldsymbol{\lambda}}(\boldsymbol{x})$ a weighted generalized ψ-estimator for some unknown parameter in Θ based on the realization $\boldsymbol{x} = (x_1, \dots, x_n) \in X^n$ and weights $(\lambda_1, \dots, \lambda_n) \in \Lambda_n$.

Given properties $[P_1], \dots, [P_q]$ introduced in Definition 2 (where $q \in \mathbb{N}$), the subclass of $\Psi(X, \Theta)$ consisting of elements possessing the properties $[P_1], \dots, [P_q]$ will be denoted by $\Psi[P_1, \dots, P_q](X, \Theta)$.

We call the attention to the fact that, given $\psi \in \Psi(X, \Theta)$ and $n \in \mathbb{N}$, by Remark 1 the function $\Theta \ni t \mapsto \psi_{\boldsymbol{x}}(t)$ can have at most one point of sign change for each $\boldsymbol{x} \in X^n$. Consequently, if $\psi \in \Psi[T_n](X, \Theta)$, then the generalized ψ-estimator $\vartheta_{n,\psi}(\boldsymbol{x})$ introduced in part (ii) of Definition 2 is unique for each $\boldsymbol{x} \in X^n$. A similar conclusion holds for the weighted generalized ψ-estimator introduced in part (iv) of Definition 2. Therefore, in our forthcoming results (for example, part (vi) of Theorem 1), when we establish that under some appropriate conditions a function $\psi \in \Psi(X, \Theta)$ satisfies the property $[T_n]$ for each $n \in \mathbb{N}$ or the property $[T_n^{\boldsymbol{\lambda}}]$ for each $n \in \mathbb{N}$ and $\boldsymbol{\lambda} \in \Lambda_n$, then it means that our result in question provides conditions under which the (weighted) generalized ψ-estimator exists uniquely. If ψ is continuous in its second variable as well, then such results provide conditions under which the usual ψ-estimator exists uniquely.

In the next proposition, we study the measurability of (weighted) generalized ψ-estimators, provided that X is a measurable space.

Proposition 1. *Let $(X, \mathcal{X})$ be a measurable space, let Θ be a nonempty open interval of $\mathbb{R}$, let $n \in \mathbb{N}$, and $\psi \in \Psi[Z_n](X, \Theta)$ be measurable in its first variable. Then $\vartheta_{n,\psi} : X^n \to \Theta$ is measurable with respect to the sigma-algebras $\mathcal{X}^n$ and $\mathcal{B}(\Theta)$.*

Proof. For all $r \in \Theta$, we have that

$$\begin{aligned} \vartheta_{n,\psi}^{-1}\big((-\infty, r)\big) &= \big\{(x_1, \dots, x_n) \in X^n\colon\ \vartheta_{n,\psi}(x_1, \dots, x_n) < r\big\} \\ &= \left\{(x_1, \dots, x_n) \in X^n\colon\ \sum_{i=1}^n \psi(x_i, r) < 0\right\}, \end{aligned} \tag{2.1}$$

where the second equality is a consequence of the property $[Z_n]$ of ψ. Further, for all $r \in \Theta$, the measurability of the mapping $X \ni x \mapsto \psi(x, r)$ implies the measurability of the mapping $X^n \ni (x_1, \dots, x_n) \mapsto (\psi(x_1, r), \dots, \psi(x_n, r))$ with respect to the sigma-algebras $\mathcal{X}^n$ and $\mathcal{B}(\Theta^n)$, and hence, since the summation $\Theta^n \ni (t_1, \dots, t_n) \mapsto t_1 + \dots + t_n$ is continuous, we have that $X^n \ni (x_1, \dots, x_n) \mapsto \sum_{i=1}^n \psi(x_i, r)$ is measurable as well. By (2.1) this implies that $\vartheta_{n,\psi}^{-1}((-\infty, r)) \in \mathcal{X}$ for all $r \in \Theta$. Since the sigma-algebra generated by the family $\{(-\infty, r) \cap \Theta, r \in \Theta\}$ of intervals coincides with the Borel sigma-algebra $\mathcal{B}(\Theta)$, we get the desired measurability of the generalized ψ-estimator $\vartheta_{n,\psi}$. □

As a consequence of Proposition 1, if $\xi_1, \dots, \xi_n$ are random variables on a probability space $(\Omega, \mathcal{A}, \mathbf{P})$, then $\vartheta_{n,\psi}(\xi_1, \dots, \xi_n)$ is a random variable (measurable with respect to the sigma-algebras $\mathcal{A}$ and $\mathcal{B}(\Theta)$), that is., it is a statistic in the language of mathematical statistics. A similar statement to Proposition 1 could be formulated for weighted generalized ψ-estimators as well.

Next, we present some basic facts about the properties $[T_n]$ and $[T_n^{\boldsymbol{\lambda}}]$ given in Definition 2, which can be easily checked.

Remark 2.

(i) If $n \in \mathbb{N}$ and $\psi \in \Psi[T_n](X, \Theta)$, then for each $x_1, \dots, x_n \in X$, Eq. (1.2) can have at most one solution.

(ii) If $n \in \mathbb{N}$ and $\psi \in \Psi[T_n](X, \Theta)$ is continuous in its second variable, then $t = \vartheta_n(x_1, \dots, x_n)$ is the unique solution to (1.2), called the ψ-estimator (Z-estimator) based on the observations $x_1, \dots, x_n \in X$. In particular, if $\psi \in \Psi[T_1](X, \Theta)$ is continuous in its second variable, then for each $x \in X$, the equation $\psi(x, t) = 0$, $t \in \Theta$, has a unique solution $\vartheta_1(x)$.

(iii) If $\lambda_1 = \dots = \lambda_n > 0$ for some $n \in \mathbb{N}$, then $\vartheta_{n,\psi}^{(\lambda_1, \dots, \lambda_n)} = \vartheta_{n,\psi}$.

In the next remark, we verify the invariance of the properties $[T_n]$ and $[T_n^{\boldsymbol{\lambda}}]$ given in Definition 2.

Remark 3. We introduce a notion of equivalence in $\Psi(X,\Theta)$ as follows. We say that maps $\psi,\varphi \in \Psi(X,\Theta)$ are *equivalent* (denoted as $\psi \sim \varphi$) if there exists a positive function $h:\Theta \to \mathbb{R}$ such that $\psi(x,t) = h(t)\varphi(x,t)$ for all $(x,t) \in X \times \Theta$. It is easy to see that $\sim$ is an equivalence relation on $\Psi(X,\Theta)$. Furthermore, the properties $[T_n]$ and $[T_n^{\boldsymbol{\lambda}}]$ are invariant with respect to this equivalence, that is, if $\psi \sim \varphi$ and φ possesses the property $[T_n]$ (or the property $[T_n^{\boldsymbol{\lambda}}]$), then ψ also enjoys this property, and $\vartheta_{n,\varphi} = \vartheta_{n,\phi}$ (resp., $\vartheta^{\boldsymbol{\lambda}}_{n,\varphi} = \vartheta^{\boldsymbol{\lambda}}_{n,\phi}$).

DEFINITION 3. Let Θ be a nonempty open interval of $\mathbb{R}$, and let $f:\Theta \to \mathbb{R}$. We say that $y \in \mathbb{R}$ is a *level of increase for* f if $u,v \in \Theta$ and $f(v) \leqslant y \leqslant f(u)$ imply $v \leqslant u$.

Remark 4.

(i) If $y \in \mathbb{R}$ is such that either $f(u) > y$ for all $u \in \Theta$ or $f(u) < y$ for all $u \in \Theta$, then y is automatically a level of increase for f.

(ii) If $y \in \mathbb{R}$ is a level of increase for f, then the inverse image $f^{-1}(\{y\})$ is either empty or a singleton. In general, the converse of the previous statement is not true. For a counterexample, let us consider the function $f:\Theta \to \mathbb{R}$ given by $f(\vartheta) := 1$ if $\vartheta \in \Theta$ is rational and $f(\vartheta) := 0$ if $\vartheta \in \Theta$ is irrational. Then $f^{-1}(\{1/2\}) = \emptyset$, but $1/2$ is not a level of increase for f.

(iii) Under the condition of Definition 3, $y \in \mathbb{R}$ is a level of increase for f if and only if the relations $u \in \Theta$ and $y \leqslant f(u)$ imply that $y < f(v)$ for all $v \in \Theta$ with $u < v$. Indeed, if $y \in \mathbb{R}$ is a level of increase for f, and $u,v \in \Theta$ are such that $u < v$ and $y \leqslant f(u)$, then $y < f(v)$, since otherwise $f(v) \leqslant y \leqslant f(u)$ would yield that $v \leqslant u$, leading us to a contradiction. Conversely, assume that $y \in \mathbb{R}$ is such that the relations $u \in \Theta$ and $y \leqslant f(u)$ imply that $y < f(v)$ for all $v \in \Theta$ with $u < v$. If $u,v \in \Theta$ are such that $f(v) \leqslant y \leqslant f(u)$, then $v > u$ cannot hold, since it would yield that $y < f(v)$, leading us to a contradiction.

In the following lemma, we establish a connection between the notions of point of sign change and level of increase.

Lemma 1. *Let Θ be a nonempty open interval of $\mathbb{R}$, and let $f:\Theta \to \mathbb{R}$. Then $y \in \mathbb{R}$ is a level of increase for f if and only if one of the following statements holds*:

(i) $y < f$ *on* Θ.

(ii) $y > f$ *on* Θ.

(iii) *There exists a point of sign change for the function* $y - f$.

The proof of Lemma 1 and that of all the forthcoming results in this section can be found in Section 5.

Lemma 2. *Let Θ be a nonempty open interval, and let $f:\Theta \to \mathbb{R}$. If the levels of increase for f form a dense subset in the convex hull of $f(\Theta)$, then f is increasing. The function f is strictly increasing if and only if every element of $f(\Theta)$ is a level of increase for f. Furthermore, if $g:H \to \mathbb{R}$ is a strictly increasing function, where H is a set containing $f(\Theta)$, and $y \in H$ is level of increase for f, then $g(y)$ is a level of increase for $g \circ f$.*

We recall also a definition due to Páles [20]: given a nonempty open interval Θ of $\mathbb{R}$ and $\varepsilon > 0$, a function $f:\Theta \to \mathbb{R}$ satisfying the inequality $f(u) \leqslant f(v) + \varepsilon$ for all $u < v$, $u,v \in \Theta$ is called *ε-increasing*. If the inequality is strict for all $u < v$, $u,v \in \Theta$, then f is said to be *strictly ε-increasing*. Note that Páles [20, Thm. 3] offers the following simple characterization: a function $f:\Theta \to \mathbb{R}$ is ε-increasing if and only if there exists an increasing function $g:\Theta \to \mathbb{R}$ such that $\|f-g\|_\infty := \sup_{u\in\Theta}|f(u)-g(u)| \leqslant \varepsilon/2$.

The next lemma describes a connection between levels of increase for a function $f:\Theta \to \mathbb{R}$ and its ε-increasingness property.

Lemma 3. *Let Θ be a nonempty open interval, let $n \in \mathbb{N}$, and let $y_0 < \dots < y_n$ be real numbers. Assume that $y_0,\dots,y_{n-1}$ are levels of increase for a function $f:\Theta \to \mathbb{R}$ and $f(\Theta) \subseteq [y_0,y_n]$. Then f is strictly ε-increasing with $\varepsilon := \max\{y_1-y_0,\dots,y_n-y_{n-1}\}$.*

Now we state our first main result by presenting necessary and sufficient conditions for the properties $[T_n]$ and $[T_n^{\boldsymbol{\lambda}}]$.

Theorem 1. *Let X be a nonempty set, let Θ be a nonempty open interval of $\mathbb{R}$, and let $\psi \in \Psi[T_1](X,\Theta)$.*

- (i) *If $\psi \in \Psi[T_2^{(\lambda_1,\lambda_2)}](X,\Theta)$ for some $(\lambda_1,\lambda_2) \in (0,\infty)^2$, then for all $x, y \in X$ with $\vartheta_1(x) < \vartheta_1(y)$, the numbers λ_1/λ_2 and λ_2/λ_1 are levels of increase for the function*

$$\big(\vartheta_1(x), \vartheta_1(y)\big) \ni t \mapsto -\frac{\psi(x,t)}{\psi(y,t)}. \tag{2.2}$$

- (ii) *If $\psi \in \Psi[T_n^{(\lambda_1,\ldots,\lambda_n)}](X,\Theta)$ for some $n \in \mathbb{N} \setminus \{1\}$ and $(\lambda_1,\ldots,\lambda_n) \in (0,\infty)^n$, then for all $x, y \in X$ with $\vartheta_1(x) < \vartheta_1(y)$, the numbers $(\lambda_1+\cdots+\lambda_k)/(\lambda_{k+1}+\cdots+\lambda_n)$ and $(\lambda_{k+1}+\cdots+\lambda_n)/(\lambda_1+\cdots+\lambda_k)$, $k \in \{1,\ldots,n-1\}$, are levels of increase for the function* (2.2).
- (iii) *If $\psi \in \Psi[T_n](X,\Theta)$ for some $n \in \mathbb{N} \setminus \{1\}$, then for all $x, y \in X$ with $\vartheta_1(x) < \vartheta_1(y)$, the elements of the set $\{k/(n-k),\ k \in \{1,\ldots,n-1\}\}$ are levels of increase for the function* (2.2).
- (iv) *If $\psi \in \Psi[T_n](X,\Theta)$ for infinitely many $n \in \mathbb{N}$, then for all $x, y \in X$ with $\vartheta_1(x) < \vartheta_1(y)$, the function* (2.2) *is increasing. In addition, if for each $m \in \mathbb{N}$, there exists $n \in \mathbb{N}$ such that m divides n and $\psi \in \Psi[T_n](X,\Theta)$, then for all $x, y \in X$ with $\vartheta_1(x) < \vartheta_1(y)$, every positive rational number is a level of increase for the function* (2.2).
- (v) *If $\psi \in \Psi[T_2^{\boldsymbol{\lambda}}](X,\Theta)$ for each $\boldsymbol{\lambda} \in \Lambda_2$, then for all $x, y \in X$ with $\vartheta_1(x) < \vartheta_1(y)$, the function* (2.2) *is strictly increasing.*
- (vi) *If $\psi \in \Psi[Z_1](X,\Theta)$ and for all $x, y \in X$ with $\vartheta_1(x) < \vartheta_1(y)$, the function* (2.2) *is strictly increasing, then $\psi \in \Psi[T_n^{\boldsymbol{\lambda}}](X,\Theta)$ for all $n \in \mathbb{N}$ and $\boldsymbol{\lambda} \in \Lambda_n$.*

We emphasize that for $\psi \in \Psi[T_1](X,\Theta)$, statement (vi) of Theorem 1 provides a sufficient condition for the existence and uniqueness of a weighted generalized ψ-estimator. If, in addition, ψ is continuous in its second variable, then it gives a sufficient condition for the existence and uniqueness of a (usual) ψ-estimator.

The following statement establishes three equivalent conditions under the property $[Z_1]$.

Corollary 1. *Let X be a nonempty set, let Θ be a nonempty open interval of $\mathbb{R}$, and let $\psi \in \Psi[Z_1](X,\Theta)$. Then the following statements are equivalent:*

- (i) *For all $x, y \in X$ with $\vartheta_1(x) < \vartheta_1(y)$, the function* (2.2) *is strictly increasing.*
- (ii) *For each $\boldsymbol{\lambda} \in \Lambda_2$, we have $\psi \in \Psi[T_2^{\boldsymbol{\lambda}}](X,\Theta)$.*
- (iii) *For all $n \in \mathbb{N}$ and $\boldsymbol{\lambda} \in \Lambda_n$, we have $\psi \in \Psi[T_n^{\boldsymbol{\lambda}}](X,\Theta)$.*

In part (ii) of the next proposition, we provide a sufficient condition (which does not involve the property $[Z_1]$) under which ψ has the property $[T_n^{\boldsymbol{\lambda}}]$ for all $n \in \mathbb{N}$ and $\boldsymbol{\lambda} \in \Lambda_n$. We also call the attention to the fact that our proof is elementary in the sense that it does not depend on Theorem 1.

Proposition 2. *Let X be a nonempty set, let Θ be a nonempty open interval of $\mathbb{R}$, and let $\psi \in \Psi[T_1](X,\Theta)$.*

- (i) *If for each $x \in X$, the function $\Theta \ni t \mapsto \psi(x,t)$ is (strictly) decreasing, then for all $x, y \in X$ with $\vartheta_1(x) < \vartheta_1(y)$, the function* (2.2) *is (strictly) increasing.*
- (ii) *If for each $x \in X$, the function $\Theta \ni t \mapsto \psi(x,t)$ is strictly decreasing, then $\psi \in \Psi[T_n^{\boldsymbol{\lambda}}](X,\Theta)$ for all $n \in \mathbb{N}$ and $\boldsymbol{\lambda} \in \Lambda_n$.*

The next proposition establishes a connection between the property $[T_n]$ and the strict $(1/n)$-increasingness of some appropriately defined functions.

Proposition 3. *Let X be a nonempty set, let $\varTheta$ be a nonempty open interval of $\mathbb{R}$, and let $\psi \in \varPsi[T_1](X,\varTheta)$. If $\psi \in \varPsi[T_n](X,\varTheta)$ for some $n \in \mathbb{N} \setminus \{1\}$, then for all $x, y \in X$ with $\vartheta_1(x) < \vartheta_1(y)$, the function*

$$\big(\vartheta_1(x), \vartheta_1(y)\big) \ni t \mapsto \frac{\psi(x,t)}{\psi(x,t) - \psi(y,t)} \tag{2.3}$$

is strictly $(1/n)$-increasing.

The following two results describe the hierarchy among the properties $([T_n])_{n\in\mathbb{N}}$ and establish a kind of "grouping" property of $[T_n^{\boldsymbol{\lambda}}]$.

Proposition 4. *Let X be a nonempty set, let $\varTheta$ be a nonempty open interval of $\mathbb{R}$, and let $\psi \in \varPsi(X,\varTheta)$. If $\psi \in \varPsi[T_n](X,\varTheta)$ for some $n \in \mathbb{N}$, then $\psi \in \varPsi[T_m](X,\varTheta)$ for all $m \in \{1,\dots,n\}$ that divide n.*

Proposition 5. *Let X be a nonempty set, let $\varTheta$ be a nonempty open interval of $\mathbb{R}$, let $\psi \in \varPsi[T_n^{\boldsymbol{\lambda}}](X,\varTheta)$ for some $n \in \mathbb{N}$, and let $\boldsymbol{\lambda} = (\lambda_1,\dots,\lambda_n) \in \varLambda_n$. Let $m \in \{1,\dots,n\}$, and let $H_1,\dots,H_m$ be nonempty pairwise disjoint subsets of $\{1,\dots,n\}$ such that $H_1 \cup \dots \cup H_m = \{1,\dots,n\}$. For each $\alpha \in \{1,\dots,m\}$, define $\mu_\alpha := \sum_{i\in H_\alpha} \lambda_i$. Then $\boldsymbol{\mu} := (\mu_1,\dots,\mu_m) \in \varLambda_m$ and $\psi \in \varPsi[T_m^{\boldsymbol{\mu}}](X,\varTheta)$.*

Note that Proposition 5 implies Proposition 4. Indeed, let $\psi \in \varPsi[T_n](X,\varTheta)$ for some $n \in \mathbb{N}$. If m is a divisor of n, then $n = km$ with some $k \in \mathbb{N}$, and hence $H_j := \{(j-1)k+1,\dots,jk\}$, $j = 1,\dots,m$, are nonempty pairwise disjoint subsets such that $H_1 \cup \dots \cup H_m = \{1,\dots,n\}$. With the choice $\boldsymbol{\lambda} := (\lambda_1,\dots,\lambda_n) := (1,\dots,1) \in \varLambda_n$, we have that ψ is a $T_n^{\boldsymbol{\lambda}}$-function, and $\mu_\alpha = k$, $\alpha \in \{1,\dots,m\}$. Hence Proposition 5 yields that ψ is a $T_m^{\boldsymbol{\mu}}$-function. Since μ_α, $\alpha \in \{1,\dots,m\}$, are all the same positive constant k, we have that ψ is a T_m-function as well. Further, we note that Proposition 5 may be useful to see, for example, that the property $[T_3^{(\lambda_1,\lambda_2,\lambda_3)}]$ of ψ implies the property $[T_2^{(\lambda_1+\lambda_2,\lambda_3)}]$ of ψ, where $(\lambda_1,\lambda_2,\lambda_3) \in \varLambda_3$.

The next example demonstrates that the property $[T_2]$ can already fail to hold for a T_1-function.

Example 1. Let $m \in \mathbb{N}$, $X := \{x_1,\dots,x_m\}$, $\varTheta := \mathbb{R}$, and let $w_1,\dots,w_m > 0$. Define $\psi : X \times \varTheta \to \mathbb{R}$ by

$$\psi(x_i,t) := \begin{cases} w_i & \text{if } t < i, \\ -w_i & \text{if } t \geqslant i, \end{cases} \qquad i \in \{1,\dots,m\}.$$

Then $\psi \in \varPsi[T_1](X,\varTheta)$, and $\vartheta_1(x_i) = i$ for all $i \in \{1,\dots,m\}$. We can easily see that the property $[T_2]$ holds if and only if $w_i \neq w_j$ for all distinct $i, j \in \{1,\dots,m\}$, Indeed, if $1 \leqslant i < j \leqslant m$, then

$$\psi(x_i,t) + \psi(x_j,t) = \begin{cases} w_i + w_j > 0 & \text{if } t < i, \\ -w_i + w_j & \text{if } i \leqslant t < j, \\ -w_i - w_j < 0 & \text{if } j \leqslant t. \end{cases}$$

This function has a point of sign change if and only if $w_i \neq w_j$, as desired. We also get that if $\psi \in \varPsi[T_2](X,\varTheta)$, then

$$\vartheta_2(x_i,x_j) = \vartheta_2(x_j,x_i) = \begin{cases} i & \text{if } w_i > w_j, \\ j & \text{if } w_i < w_j. \end{cases}$$

Furthermore, $\vartheta_2(x_i,x_i) = \vartheta_1(x_i) = i$ for all $i \in \{1,\dots,m\}$.

We now give an example to point out that in part (vi) of Theorem 1 the assumption that $\psi(x,\vartheta_1(x)) = 0$, $x \in X$, cannot be omitted.

Example 2. Let $X := \{x_1, x_2\}$ (with $x_1 \neq x_2$) and $\Theta := \mathbb{R}$. Let

$$\psi(x_1,t) := \begin{cases} 2 & \text{if } t < 1, \\ -t & \text{if } 1 \leqslant t \leqslant 2, \\ -2 & \text{if } t > 2, \end{cases} \quad \text{and} \quad \psi(x_2,t) := \begin{cases} 1 & \text{if } t < 2, \\ 2 & \text{if } t = 2, \\ -1 & \text{if } t > 2. \end{cases}$$

Then $\psi \in \Psi[T_1](X,\Theta)$ with $\vartheta_1(x_1) = 1$ and $\vartheta_1(x_2) = 2$, and $\psi(x_i, \vartheta_1(x_i)) \neq 0$ for $i \in \{1,2\}$. We also note that the function

$$\bigl(\vartheta_1(x_1), \vartheta_1(x_2)\bigr) = (1,2) \ni t \mapsto \frac{\psi(x_1,t)}{\psi(x_2,t)} = -t$$

is strictly decreasing. However, ψ is not a T_2-function, since

$$\psi(x_1,t) + \psi(x_2,t) = \begin{cases} 2+1 = 3 > 0 & \text{if } t < 1, \\ -1+1 = 0 & \text{if } t = 1, \\ -t+1 < 0 & \text{if } 1 < t < 2, \\ -2+2 = 0 & \text{if } t = 2, \\ -2-1 = -3 < 0 & \text{if } t > 2, \end{cases}$$

which shows that $\mathbb{R} \ni t \mapsto \psi(x_1,t) + \psi(x_2,t)$ does not have a point of sign change.

In what follows, as an application of Proposition 2, we present an example of a large class of functions $\psi : X \times \Theta \to \mathbb{R}$ that possesses the property $[T_n^{\boldsymbol{\lambda}}]$ for all $n \in \mathbb{N}$ and $\boldsymbol{\lambda} \in \Lambda_n$ and for which the point of sign change $\vartheta_{n,\psi}^{\boldsymbol{\lambda}}(\boldsymbol{x})$ (where $\boldsymbol{x} \in X^n$) has an explicit form. This class of functions may be called the class of Bajraktarević-type functions, motivated by the representation of Bajraktarević means as a special deviation means. For the description of this class of functions, we need to recall the notion of generalized left inverse of a strictly monotone (but not necessarily continuous) function defined on a nonempty open interval of $\mathbb{R}$; see, for example, Gasiński and Papageorgiou [8, Prop. 1.55 and the subsequent comment] and Grünwald and Páles [9, Lemma 1]. The notion of generalized left inverse in question is likely to be well known, and its properties are established for a while, but we could not trace the roots, and therefore we refer to the recent treatments in [8] and [9].

Lemma 4. *Let Θ be a nonempty open interval of $\mathbb{R}$, and let $f : \Theta \to \mathbb{R}$ be a strictly increasing function. Then there exists a uniquely determined monotone function $g : \operatorname{conv}(f(\Theta)) \to \Theta$ that is the left inverse of f, that is,*

$$(g \circ f)(x) = x, \quad x \in \Theta.$$

Furthermore, g is monotone in the same sense as f, is continuous, and the following relation holds:

$$(f \circ g)(y) = y, \quad y \in f(\Theta).$$

The function $g : \operatorname{conv}(f(\Theta)) \to \Theta$ described in Lemma 4 is called the *generalized left inverse of the strictly increasing function $f : \Theta \to \mathbb{R}$* and is denoted by $f^{(-1)}$. In fact, by the proof of Lemma 1 in Grünwald and Páles [9] it also turns out that

$$g(y) = \sup\bigl\{u \in \Theta\colon\ f(u) \leqslant y\bigr\} = \inf\bigl\{u \in \Theta\colon\ f(u) \geqslant y\bigr\}, \quad y \in \operatorname{conv}\bigl(f(\Theta)\bigr).$$

It is clear that the restriction of $f^{(-1)}$ to $f(\Theta)$ is the inverse of f in the standard sense. Therefore $f^{(-1)}$ is the continuous and monotone extension of the inverse of f to the smallest interval containing the range of f, that is, to the convex hull of $f(\Theta)$.

DEFINITION 4. Let X be a nonempty set, let Θ be a nonempty open interval of $\mathbb{R}$, let $f:\Theta\to\mathbb{R}$ be a strictly increasing function, and let $p:X\to\mathbb{R}_{++}$ and $\varphi:X\to\operatorname{conv}(f(\Theta))$. In terms of these functions, define $\psi: X\times\Theta\to\mathbb{R}$ by

$$\psi(x,t) := p(x)\big(\varphi(x)-f(t)\big), \quad x\in X,\ t\in\Theta. \tag{2.4}$$

The function ψ defined by (2.4) is said to be of Bajraktarević type. In particular, if $p=1$ is a constant function, then ψ is said to be of quasiarithmetic type.

Proposition 6. *Under the assumptions of Definition* 4, $\psi\in\Psi[T_n^{\boldsymbol{\lambda}}](X,\Theta)$ *for all* $n\in\mathbb{N}$ *and* $\boldsymbol{\lambda}=(\lambda_1,\dots,\lambda_n)\in\Lambda_n$, *and*

$$\vartheta_{n,\psi}^{\boldsymbol{\lambda}}(\boldsymbol{x}) = f^{(-1)}\left(\frac{\lambda_1 p(x_1)\varphi(x_1)+\dots+\lambda_n p(x_n)\varphi(x_n)}{\lambda_1 p(x_1)+\dots+\lambda_n p(x_n)}\right) \tag{2.5}$$

for all $\boldsymbol{x}=(x_1,\dots,x_n)\in X^n$. *In particular, we have the equality* $\vartheta_{1,\psi}=f^{(-1)}\circ\varphi$.

We may call the value $\vartheta_{n,\psi}^{\boldsymbol{\lambda}}(\boldsymbol{x})$ given by (2.5) as a Bajraktarević-type ψ-estimator of some unknown parameter in Θ based on the realization $\boldsymbol{x}=(x_1,\dots,x_n)\in X^n$ and weights $\boldsymbol{\lambda}=(\lambda_1,\dots,\lambda_n)\in\Lambda_n$ corresponding to the Bajraktarević-type function given by (2.4). In particular, if $p=1$ is a constant function in (2.4), then we speak about a quasiarithmetic-type ψ-estimator.

Note that in case of $X:=\Theta$ and $\varphi:=f$, Proposition 6 reduces to Theorem 3 in Grünwald and Páles [9] for Bajraktarević means. In addition, if $p=1$ is a constant function, then Proposition 6 is about generalized quasiarithmetic means (here we use the term "generalized", since for usual quasiarithmetic means, the function f is not only strictly increasing, but continuous as well).

In the next example, we point out that, in general, we cannot omit the restriction that m divides n in Proposition 4. For another example, see the case of empirical median discussed in Example 5.

Example 3. Let X be an arbitrary set with at least two distinct elements, let $X_1,X_2\subseteq X$ be nonempty disjoint subsets such that $X_1\cup X_2=X$, and let $w_1,w_2>0$. Define $\psi:X\times\Theta\to\mathbb{R}$ by

$$\psi(x,t) := \begin{cases} w_i, & t<i,\\ -w_i, & t\geqslant i,\end{cases} \qquad x\in X_i.$$

Then $\psi\in\Psi[T_1](X,\Theta)$, and, for $i\in\{1,2\}$ and $x\in X_i$, we have $\vartheta_1(x)=i$.

Let $n,k\in\mathbb{N}$ be such that k is not a divisor of n (which implies that $k\geqslant 2$). Then

$$\frac{1}{k}\notin\left\{\frac{1}{n},\dots,\frac{n-1}{n}\right\}. \tag{2.6}$$

Indeed, on the contrary, if the inclusion were valid, then $1/k$ would be of the form m/n for some $m\in\{1,\dots,n-1\}$, yielding that $n=km$, which contradicts the assumption $k\nmid n$.

Assuming that $w_1=k-1$ and $w_2=1$, we prove that ψ is a T_n-function but it is not a T_k-function.

To show that ψ is a T_n-function, let $\boldsymbol{y}:=(y_1,\dots,y_n)\in X^n$. For $j\in\{1,2\}$, define the set $S_j:=\{i\in\{1,\dots,n\}\colon\ y_i\in X_j\}$. Then $\{S_1,S_2\}$ forms a partition of $\{1,\dots,n\}$. Let n_j denote the cardinality of S_j, $j\in\{1,2\}$. Then $n=n_1+n_2$, and

$$\psi_{\boldsymbol{y}}(t) := \sum_{i=1}^{n}\psi(y_i,t) = \begin{cases} n_1w_1+n_2w_2>0 & \text{if } t<1,\\ -n_1w_1+n_2w_2 & \text{if } 1\leqslant t<2,\\ -n_1w_1-n_2w_2<0 & \text{if } 2\leqslant t.\end{cases}$$

Using condition (2.6) and $k \geqslant 2$, we have

$$-n_1 w_1 + n_2 w_2 = -n_1(k-1) + (n - n_1) = n - n_1 k \neq 0.$$

Therefore the point of sign change for the function $\psi_{\boldsymbol{y}}$ equals 1 if $-n_1 w_1 + n_2 w_2 < 0$ and equals 2 if $-n_1 w_1 + n_2 w_2 > 0$. This proves that $\psi \in \Psi[T_n](X, \Theta)$.

To verify that ψ is not a T_k-function, fix $x_1 \in X_1$ and $x_2 \in X_2$, and let $\boldsymbol{z} := (x_1, x_2, \dots, x_2) \in X^k$. Then

$$\psi_{\boldsymbol{z}}(t) := \psi(x_1, t) + (k-1)\psi(x_2, t) = \begin{cases} w_1 + (k-1)w_2 > 0 & \text{if } t < 1, \\ -w_1 + (k-1)w_2 = 0 & \text{if } 1 \leqslant t < 2, \\ -w_1 - (k-1)w_2 < 0 & \text{if } 2 \leqslant t. \end{cases}$$

Therefore the function $\psi_{\boldsymbol{z}}$ does not have a point of sign change, and, consequently, ψ is not a T_k-function, as desired.

3 Existence and uniqueness of the point of sign change of ψ-expectation functions

In this section, we investigate Problem 2 presented in the Introduction.

As it was mentioned in the Introduction, in case of simple random variables, Problem 2 is a particular case of Problem 1. More precisely, if $\psi \in \Psi(X, \Theta)$ and ξ is a simple random variable such that $\mathbf{P}(\xi = x_i) = p_i$, $i = 1, \dots, n$, where $n \in \mathbb{N}$, $(x_1, \dots, x_n) \in X^n$ and $p_1, \dots, p_n \geqslant 0$, $p_1 + \dots + p_n = 1$, then $\mathbf{E}(\psi(\xi, t)) = \sum_{i=1}^n p_i \psi(x_i, t)$, $t \in \Theta$. In addition, if ψ is a $T_n^{(p_1,\dots,p_n)}$-function, then by definition the function (1.4) has a unique point of sign change. Further, Theorem 1 provides some necessary and sufficient conditions under which ψ possesses the property $[T_n^{(p_1,\dots,p_n)}]$. In case of a general (not necessarily discrete) random variable ξ and $\psi \in \Psi(X, \Theta)$, in our forthcoming results, Theorem 2 and Proposition 7, we derive sufficient conditions on ξ and ψ under which there exists a unique point of sign change of the corresponding ψ-expectation function.

Next, we present our second main result, in which we give a set of sufficient conditions in order that the function given by (1.4) have a unique point of sign change.

Theorem 2. *Let $(X, \mathcal{X})$ be a measurable space, let Θ be a nonempty open interval of $\mathbb{R}$, let $\psi : X \times \Theta \to \mathbb{R}$, and let $\xi : \Omega \to X$ be a random variable defined on a probability space $(\Omega, \mathcal{A}, \mathbf{P})$. Let us suppose that*

(i) $\psi \in \Psi[Z_1](X, \Theta)$,
(ii) *for each $x, y \in X$ with $\vartheta_1(x) < \vartheta_1(y)$, the function* (2.2) *is strictly increasing,*
(iii) *ψ is measurable in its first variable,*
(iv) $\mathbf{E}(|\psi(\xi, t)|) < \infty$ *for each $t \in \Theta$,*
(v) *there exist $s_0, t_0 \in \Theta$ such that* $\mathbf{E}(\psi(\xi, s_0)) \geqslant 0$ *and* $\mathbf{E}(\psi(\xi, t_0)) \leqslant 0$.

Then the map $\Theta \ni t \to \mathbf{E}(\psi(\xi, t))$ admits a unique point of sign change in Θ.

The proofs of Theorem 2 and of all the forthcoming results in this section can be found in Section 6.

Next, we provide a set of sufficient conditions (which does not involve the condition $\psi(x, \vartheta_1(x)) = 0$ for each $x \in X$) under which the map $\Theta \ni t \to \mathbf{E}(\psi(\xi, t))$ also has a unique point of sign change.

Proposition 7. *Let $(X, \mathcal{X})$ be a measurable space, let Θ be a nonempty open interval of $\mathbb{R}$, let $\psi : X \times \Theta \to \mathbb{R}$, and let $\xi : \Omega \to X$ be a random variable defined on a probability space $(\Omega, \mathcal{A}, \mathbf{P})$. Let us suppose that*

(i) *for each $x \in X$, the function $\Theta \ni t \mapsto \psi(x, t)$ is strictly decreasing,*
(ii) *ψ is measurable in its first variable,*
(iii) $\mathbf{E}(|\psi(\xi, t)|) < \infty$ *for each $t \in \Theta$,*
(iv) *there exist $s_0, t_0 \in \Theta$ such that* $\mathbf{E}(\psi(\xi, s_0)) \geqslant 0$ *and* $\mathbf{E}(\psi(\xi, t_0)) \leqslant 0$.

Then the function $\Theta \ni t \to \mathbf{E}(\psi(\xi, t))$ admits a unique point of sign change in Θ.

Next, we formulate a corollary of Theorem 2, which is in fact part (vi) of Theorem 1.

Corollary 2. *Let* X *be a nonempty set, let* Θ *be a nonempty open interval of* $\mathbb{R}$*, and let* $\psi \in \Psi[Z_1](X, \Theta)$*. If for all* $x, y \in X$ *with* $\vartheta_1(x) < \vartheta_1(y)$*, the function* (2.2) *is strictly increasing, then* $\psi \in \Psi[T_n^{\boldsymbol{\lambda}}](X, \Theta)$ *for all* $n \in \mathbb{N}$ *and* $\boldsymbol{\lambda} \in \Lambda_n$*.*

In what follows, we present a particular case of Proposition 7, which can be considered as a counterpart of Proposition 6 for Bajraktarević-type functions ψ.

Proposition 8. *Let* $(X, \mathcal{X})$ *be a measurable space, let* Θ *be a nonempty open interval of* $\mathbb{R}$*, let* $f : \Theta \to \mathbb{R}$ *be a strictly increasing function, and let* $p : X \to \mathbb{R}_{++}$ *and* $\varphi : X \to \operatorname{conv}(f(\Theta))$ *be measurable functions. Define* $\psi : X \times \Theta \to \mathbb{R}$ *by* (2.4)*. Further, let* $(\Omega, \mathcal{A}, \mathbf{P})$ *be a probability space, let* $\xi : \Omega \to X$ *be a random variable such that* $\mathbf{E}(p(\xi)|\varphi(\xi)|) < \infty$ *and* $\mathbf{E}(p(\xi)) < \infty$*. Then the function* $\Theta \ni t \to \mathbf{E}(\psi(\xi, t))$ *admits a unique point of sign change in* Θ*, which is given by*

$$f^{(-1)}\left(\frac{\mathbf{E}(p(\xi)\varphi(\xi))}{\mathbf{E}(p(\xi))}\right).$$

The following auxiliary result is instrumental for the proof of Proposition 8.

Lemma 5. *Let* $(X, \mathcal{X})$ *be a measurable space, and let* $p : X \to \mathbb{R}_{++}$ *and* $\varphi : X \to \mathbb{R}$ *be measurable functions. Further, let* $\xi : \Omega \to X$ *be a random variable on a probability space* $(\Omega, \mathcal{A}, \mathbf{P})$ *such that* $\mathbf{E}(p(\xi)) < \infty$ *and* $\mathbf{E}(p(\xi)|\varphi(\xi)|) < \infty$*. Then*

$$\frac{\mathbf{E}(p(\xi)\varphi(\xi))}{\mathbf{E}(p(\xi))} \in \operatorname{conv}\big(\varphi(X)\big).$$

In the next example, we consider a particular form of ψ, which has been recently investigated by Mathieu [18]: namely, let $\psi : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ be defined as

$$\psi(x, t) := \operatorname{sign}(x - t) f\big(|x - t|\big), \quad x, t \in \mathbb{R}, \tag{3.1}$$

where $f : \mathbb{R}_+ \to \mathbb{R}_+$. Mathieu [18, Lemma 2] has derived some sufficient conditions on f and ξ under which the equation $\mathbf{E}(\psi(\xi, t)) = 0$, $t \in \mathbb{R}$, has a unique solution; for more detail and our new results in this particular case, see the next example and Proposition 9.

Example 4. Let $X := \mathbb{R}$, $\Theta := \mathbb{R}$ and $\psi : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ be given by (3.1). Given a random variable ξ, Mathieu [18] has recently considered the problem of finding a unique element $t_0 \in \Theta$ such that $\mathbf{E}(\psi(\xi, t_0)) = 0$, where ψ has the form given in (3.1) with f admitting the following properties (called Assumption 2 in Mathieu [18]):

(a) f is continuous and differentiable Lebesgue almost everywhere,
(b) $f(0) = 0$,
(c) f is concave,
(d) there exist $\beta, \gamma > 0$ such that $\gamma \mathbf{1}_{\{x \leqslant \beta\}} \leqslant f'(x) \leqslant 1$ Lebesgue a.e. $x \geqslant 0$.

For historical fidelity, we note that Mathieu [18] investigated a more general setup. He considered a random variable ξ with values in a Hilbert space $\mathcal{H}$ and a function $\psi : \mathcal{H} \times \mathcal{H} \to \mathcal{H}$, $\psi(x, t) := ((x - t)/\|x - t\|) \times f(\|x - t\|)$ for $x \neq t$, $x, t \in \mathcal{H}$, where $\|\cdot\|$ is the norm of the Hilbert space $\mathcal{H}$ (the value of ψ at (x, x), $x \in \mathcal{H}$, was not specified by Mathieu [18]).

Mathieu [18, Lemma 2] has shown that (formulating his result only in the case of $\mathcal{H} = \mathbb{R}$) if f admits the properties (a)–(d), $\mathbf{E}(|\xi|) < \infty$, and $\mathbf{E}(\varrho(|\xi - \mathbf{E}(\xi)|)) < \varrho(\beta)$, where $\varrho(x) := \int_0^x f(u)\,\mathrm{d}u$, $x \in \mathbb{R}_+$, then there exists a unique element $t_0 \in \Theta$ such that $\mathbf{E}(\psi(\xi, t_0)) = 0$. Mathieu [18] also noted that the assumptions under which the existence and uniqueness of a solution in question was established are not the minimal ones, but he has not searched for possible minimal assumptions.

Note that we can rewrite ψ given by (3.1) as $\psi(x,t) = \widetilde{f}(x-t)$, $x,t\in\mathbb{R}$, where $\widetilde{f}:\mathbb{R}\to\mathbb{R}$ denotes the odd extension of $f:\mathbb{R}_+\to\mathbb{R}_+$ to $\mathbb{R}$ given by

$$\widetilde{f}(z) := \begin{cases} f(z) & \text{if } z>0,\\ 0 & \text{if } z=0,\\ -f(-z) & \text{if } z<0. \end{cases}$$

As a new result, we have the following proposition.

Proposition 9. *If $f:\mathbb{R}_+\to\mathbb{R}_+$ is continuous and strictly increasing with $f(0)=0$ and $\lim_{z\to\infty} f(z)\in(0,\infty)$, then for any random variable ξ, we have that $\mathbf{E}(|\psi(\xi,t)|)<\infty$, $t\in\mathbb{R}$, and the equation*

$$\mathbf{E}\bigl(\psi(\xi,t)\bigr) = \mathbf{E}\bigl(\widetilde{f}(\xi-t)\bigr) = 0$$

has a unique solution with respect to $t\in\mathbb{R}$.

Now we compare the assumptions of Lemma 2 in Mathieu [18] and those of Proposition 9. Note that if a function $f:\mathbb{R}_+\to\mathbb{R}_+$ admits properties (a)–(d) of Example 4, then it is not necessarily strictly increasing (for example, it may happen that $f(x)=f(\beta)$ for $x\geqslant\beta$; see, for example, the Huber function (4.10)), so we cannot say that Proposition 9 is a generalization of Lemma 2 in Mathieu [18]. However, the conditions of Proposition 9 might be checked more easily than those of Lemma 2 in Mathieu [18] to prove that the equation $\mathbf{E}(\psi(\xi,t))=0$ has a unique solution with respect to $t\in\mathbb{R}$. For example, if $f:\mathbb{R}_+\to\mathbb{R}$, $f(z):=z/\sqrt{1+z^2/2}$, $z\in\mathbb{R}_+$, then f is a continuous strictly increasing function starting from 0, and $\lim_{z\to\infty} f(z)=\sqrt{2}$. Indeed, we have $f'(z)=(1+z^2/2)^{-3/2}>0$ for each $z\in\mathbb{R}_+$. This special choice of f plays a role in robust statistics; for more detail, see, for example, Rey [22, Sect. 6.4] or Example 8.

We note that in the arXiv version of this paper [2, Rem. 3.8], we also compare Theorem 2 with Theorem 3.2 in Clarke [6].

4 Examples from statistical estimation theory

In this section, we present several examples from statistical estimation theory, where our results in Sections 2 and 3 can be well applied. For example, we consider the cases of the empirical median, the empirical quantiles, the empirical expectiles, ψ-estimators recently used by Mathieu [18], some ψ-estimators that are important in robust statistics, and some examples from maximum likelihood theory. The proof of the results in this section (Propositions 10 and 12) can be found in Section 7.

Example 5 [*Empirical median*]. Let $X:=\mathbb{R}$, $\Theta:=\mathbb{R}$, and $\psi:\mathbb{R}\times\mathbb{R}\to\mathbb{R}$, $\psi(x,t):=\operatorname{sign}(x-t)$, $x,t\in\mathbb{R}$. For each $x\in\mathbb{R}$, the function $\mathbb{R}\ni t\mapsto\psi(x,t)$ is decreasing, but not strictly decreasing. Then for all $n\in\mathbb{N}$ and $x_1,\dots,x_n\in\mathbb{R}$, Eq. (1.2) takes the form

$$\sum_{i=1}^{n}\operatorname{sign}(x_i-t)=0,\quad t\in\mathbb{R}. \tag{4.1}$$

In this particular case, the function ψ is not continuous in its second variable, i.e., $\psi\notin\varPsi[Z](\mathbb{R},\mathbb{R})$, and the corresponding equation (4.1) has an important role in statistics. Namely, we can check that $\operatorname{Med}_n:\mathbb{R}^n\to\mathbb{R}$,

$$\begin{aligned}\operatorname{Med}_n(x_1,\dots,x_n) &:= \frac{1}{2}\bigl(x^*_{\lceil n/2\rceil}+x^*_{\lfloor n/2+1\rfloor}\bigr)\\ &= \begin{cases} x^*_{k+1} & \text{if } n=2k+1,\\ \frac{1}{2}(x^*_k+x^*_{k+1}) & \text{if } n=2k, \end{cases}\qquad k\in\mathbb{Z}_+, \end{aligned}\tag{4.2}$$

is a solution of Eq. (4.1), where $x_1^* \leqslant x_2^* \leqslant \cdots \leqslant x_n^*$ is the ordered sample of $x_1, \ldots, x_n \in \mathbb{R}$. Of course, if $n = 2k$, then there are other solutions of Eq. (4.1). For example, if $x_1 < x_2 < \cdots < x_{2k}$, then we have

$$\left\{ t \in \mathbb{R} \colon \sum_{i=1}^{2k} \operatorname{sign}(x_i - t) = 0 \right\} = [x_k, x_{k+1}], \tag{4.3}$$

where $[x_k, x_{k+1}]$ is not a singleton. Note that $\operatorname{Med}_n(x_1, \ldots, x_n)$ is nothing else but the well-known empirical median of $x_1, \ldots, x_n$.

Further, we have that $\psi \in \Psi[T_n](\mathbb{R}, \mathbb{R})$ for each $n = 2k + 1$, $k \in \mathbb{Z}_+$, with $\vartheta_n(\boldsymbol{x}) = x_{k+1}^*$, $\boldsymbol{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$. Note also that $\psi \notin \Psi[T_n](\mathbb{R}, \mathbb{R})$ for all $n = 2k$, $k \in \mathbb{N}$. Indeed, if $x_1 < x_2 < \cdots < x_{2k}$, then (4.3) implies that the function $\psi_{(x_1, \ldots, x_{2k})}$ does not have a point of sign change. Furthermore, for all $x, y \in \mathbb{R}$ with $\vartheta_1(x) = x < y = \vartheta_1(y)$, the function (2.3) takes the form

$$(x, y) \ni t \mapsto \frac{\psi(x, t)}{\psi(x, t) - \psi(y, t)} = \frac{\operatorname{sign}(x - t)}{\operatorname{sign}(x - t) - \operatorname{sign}(y - t)} = \frac{-1}{-1 - 1} = \frac{1}{2}.$$

This function is a rational constant, in particular, strictly $(1/\ell)$-increasing for each $\ell \in \mathbb{N}$. It underlines the fact that the strictly $(1/n)$-increasing property of the function (2.3) in Proposition 3 is only a necessary but not a sufficient condition for $\psi \in \Psi[T_n](\mathbb{R}, \mathbb{R})$. Indeed, in the present example, the function (2.3) is strictly $(1/\ell)$-increasing for each $\ell \in \mathbb{N}$, but $\psi \notin \Psi[T_n](\mathbb{R}, \mathbb{R})$ for $n = 2k$, $k \in \mathbb{N}$. Moreover, for all $x, y \in \mathbb{R}$ with $\vartheta_1(x) = x < y = \vartheta_1(y)$, the function (2.2) takes the form

$$(x, y) \ni t \mapsto -\frac{\psi(x, t)}{\psi(y, t)} = -\frac{\operatorname{sign}(x - t)}{\operatorname{sign}(y - t)} = 1.$$

This function is a rational constant, in particular, increasing but not strictly increasing. It is in accordance with part (iv) of Theorem 1 and part (i) of Proposition 2 as well, since $\psi \in \Psi[T_n](\mathbb{R}, \mathbb{R})$ for infinitely many $n \in \mathbb{N}$, namely, for each $n = 2k + 1$, $k \in \mathbb{Z}_+$; and for each $x \in \mathbb{R}$, the function $\mathbb{R} \ni t \mapsto \psi(x, t)$ is decreasing. Note also that it does not hold that for each $m \in \mathbb{N}$, there exists $n \in \mathbb{N}$ such that m divides n and $\psi \in \Psi[T_n](\mathbb{R}, \mathbb{R})$ (indeed, in case of an even $m \in \mathbb{N}$, we cannot choose such n). Moreover, we cannot apply part (vi) of Theorem 1. This underlines that the increasing property of the function (2.2) is only a necessary but not a sufficient condition in order that ψ be a T_n-function for each $n \in \mathbb{N}$. Finally, we mention that the present example also shows that in Proposition 4 the restriction that $m \in \{1, \ldots, n\}$ divides n cannot be removed in general.

Example 6 [*Empirical quantiles*]. Given $\alpha \in (0, 1)$, $n \in \mathbb{N}$, and $x_1, \ldots, x_n \in \mathbb{R}$, an empirical α-quantile based on $x_1, \ldots, x_n$ is defined as any solution of the minimization problem

$$\begin{aligned} \min_{t \in \mathbb{R}} \sum_{i=1}^{n} \varphi_\alpha(x_i - t) &= \min_{t \in \mathbb{R}} \sum_{i=1}^{n} \big(\alpha \mathbf{1}_{\{x_i \geqslant t\}} + (\alpha - 1)\mathbf{1}_{\{x_i < t\}}\big)(x_i - t) \\ &= \min_{t \in \mathbb{R}} \sum_{i=1}^{n} \frac{1}{2}\big(|x_i - t| + (2\alpha - 1)(x_i - t)\big), \end{aligned} \tag{4.4}$$

where $\varphi_\alpha : \mathbb{R} \to \mathbb{R}$ is the so-called α-quantile check function given by

$$\begin{aligned} \varphi_\alpha(x) &:= |\alpha - \mathbf{1}_{\{x<0\}}||x| = (\alpha \mathbf{1}_{\{x \geqslant 0\}} + (\alpha - 1)\mathbf{1}_{\{x<0\}})x \\ &= \frac{1}{2}\big(|x| + (2\alpha - 1)x\big), \quad x \in \mathbb{R}; \end{aligned}$$

see, for example, Koenker and Bassett [14, Sect. 3]. Some authors call a solution of (4.4) a geometric α-quantile; see, for example, Passeggeri and Reid [21, Sect. 2].

It is known that $q \in \mathbb{R}$ is an empirical α-quantile based on $x_1, \ldots, x_n$ if and only if

$$\frac{1}{n} \sum_{i:\, x_i < q} 1 \leqslant \alpha \quad \text{and} \quad \alpha \leqslant \frac{1}{n} \sum_{i:\, x_i \leqslant q} 1; \tag{4.5}$$

see, for example, Lange [17, Problems 12.12/13.].

It is also known that given $\alpha \in (0,1)$, $n \geqslant 2$, $n \in \mathbb{N}$, and $x_1, \ldots, x_n \in \mathbb{R}$, an empirical α-quantile given as a solution of the minimization problem (4.4) is uniquely defined if and only if $\alpha \notin \{1/n, 2/n, \ldots, (n-1)/n\}$, and in case of uniqueness, we have that it is given by

$$q_n^{(\alpha)}(x_1, \ldots, x_n) := \frac{1}{2}\left(x^*_{\lceil n\alpha \rceil} + x^*_{\lfloor n\alpha + 1 \rfloor}\right), \tag{4.6}$$

where $x_1^* \leqslant x_2^* \leqslant \cdots \leqslant x_n^*$ is the ordered sample of $x_1, \ldots, x_n \in \mathbb{R}$; see Passeggeri and Reid [21, Lemma 4.1]. We also mention an interesting result of Passeggeri and Reid [21, Lemma 4.2], which states that given $\alpha \in (0,1)$ and $n \geqslant 2$, $n \in \mathbb{N}$, the function $q_n^{(\alpha)} : \mathbb{R}^n \to \mathbb{R}$ given by (4.6) is Lipschitz continuous:

$$\left| q_n^{(\alpha)}(x_1, \ldots, x_n) - q_n^{(\alpha)}(y_1, \ldots, y_n) \right| \leqslant \max_{j \in \{1, \ldots, n\}} |x_j - y_j|$$

for all $x_1, \ldots, x_n, y_1, \ldots, y_n \in \mathbb{R}$.

Note that if $\alpha = 1/2$, then the empirical median $\mathrm{Med}_n(x_1, \ldots, x_n)$ of $x_1, \ldots, x_n$ given in (4.2) is a solution of the minimization problem

$$\min_{t \in \mathbb{R}} \sum_{i=1}^{n} |x_i - t|.$$

Indeed, if $\alpha = 1/2$, then $\varphi_{1/2}(x) = |x|/2$, $x \in \mathbb{R}$, and hence the minimization problem (4.4) with $\alpha = 1/2$ is equivalent to $\min_{t \in \mathbb{R}} \sum_{i=1}^{n} |x_i - t|$, and, as we have recalled, $q \in \mathbb{R}$ is a solution of this minimization problem if and only if inequalities (4.5) hold for q with $\alpha = 1/2$. Further, we can easily check that the empirical median $\mathrm{Med}_n(x_1, \ldots, x_n)$ based on $x_1, \ldots, x_n$ satisfies inequalities (4.5) with $\alpha = 1/2$. Moreover, for $\alpha = 1/2$ and $n \geqslant 2$, $n \in \mathbb{N}$, we have that $\alpha \notin \{1/n, 2/n, \ldots, (n-1)/n\}$ if and only if $n = 2k+1$ with some $k \in \mathbb{N}$, and in this case,

$$\begin{aligned} q_n^{(1/2)}(x_1, \ldots, x_n) &= \frac{1}{2}\left(x^*_{\lceil (2k+1)/2 \rceil} + x^*_{\lfloor (2k+1)/2 + 1 \rfloor}\right) = \frac{1}{2}\left(x^*_{k+1} + x^*_{k+1}\right) \\ &= x^*_{k+1} = \mathrm{Med}_n(x_1, \ldots, x_n) \end{aligned}$$

for all $x_1, \ldots, x_n \in \mathbb{R}$, where $\mathrm{Med}_n(x_1, \ldots, x_n)$ is defined in (4.2).

Motivated by the minimization problem (4.4), we investigate the function $\psi : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$,

$$\psi(x,t) := \begin{cases} \alpha & \text{if } x > t, \\ 0 & \text{if } x = t, \\ \alpha - 1 & \text{if } x < t. \end{cases} \tag{4.7}$$

For each $x \in \mathbb{R}$, the function $\mathbb{R} \ni t \mapsto \psi(x,t)$ is decreasing but not strictly decreasing, and not continuous. By choosing $X := \Theta := \mathbb{R}$, we have that ψ is a T_1-function with $\vartheta_1(x) := x$, $x \in \mathbb{R}$. Analogously to the result of Passeggeri and Reid [21, Lemma 4.1], we show the following result.

Proposition 10. *Given $\alpha \in (0,1)$, for each $n \geqslant 2$, the function ψ defined by* (4.7) *has the property $[T_n]$ if and only if $\alpha \notin \{1/n, \dots, (n-1)/n\}$. Further, in this case, for all $x_1, \dots, x_n \in \mathbb{R}$, we have that*

$$\vartheta_n(x_1, \dots, x_n) = x^*_{\lceil n\alpha \rceil} = x^*_{\lfloor n\alpha+1 \rfloor},$$

and hence (4.6) *is also valid, where $x_1^* \leqslant x_2^* \leqslant \dots \leqslant x_n^*$ is the ordered sample of $x_1, \dots, x_n$.*

Example 7 [*Expectiles*]. Let $\alpha \in (0,1)$, $n \in \mathbb{N}$, and $x_1, \dots, x_n \in \mathbb{R}$. The empirical α-expectile based on $x_1, \dots, x_n$ is defined as any solution of the minimization problem

$$\min_{t\in\mathbb{R}} \sum_{i=1}^n \widetilde{\varphi}_\alpha(x_i - t) = \min_{t\in\mathbb{R}} \sum_{i=1}^n \big(\alpha \mathbf{1}_{\{x_i \geqslant t\}} + (1-\alpha)\mathbf{1}_{\{x_i < t\}}\big)(x_i - t)^2,$$

where $\widetilde{\varphi}_\alpha : \mathbb{R} \to \mathbb{R}$ is given by

$$\widetilde{\varphi}_\alpha(x) := |\alpha - \mathbf{1}_{\{x<0\}}|x^2 = (\alpha \mathbf{1}_{\{x\geqslant 0\}} + (1-\alpha)\mathbf{1}_{\{x<0\}})x^2, \quad x \in \mathbb{R};$$

see, for example, Newey and Powell [19]. Expectiles are also called smoothed versions of quantiles or least asymptotically weighted squares estimators.

Motivated by this minimization problem, we may investigate the applicability of Theorem 1 for the function $\psi : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$,

$$\psi(x,t) := \alpha(x-t)^+ - (1-\alpha)(x-t)^- = \begin{cases} \alpha(x-t) & \text{if } x > t, \\ 0 & \text{if } x = t, \\ (1-\alpha)(x-t) & \text{if } x < t. \end{cases} \tag{4.8}$$

For each $x \in \mathbb{R}$, the function $\mathbb{R} \ni t \mapsto \psi(x,t)$ is strictly decreasing. By choosing $X := \mathbb{R}$ and $\Theta := \mathbb{R}$ we have that ψ is a T_1-function with $\vartheta_1(x) := x$, $x \in \mathbb{R}$, and for all $x, y \in \mathbb{R}$ with $\vartheta_1(x) = x < y = \vartheta_1(y)$, the function (2.2) takes the form

$$(x,y) \ni t \mapsto -\frac{\psi(x,t)}{\psi(y,t)} = -\frac{(1-\alpha)(x-t)}{\alpha(y-t)} = -\frac{1-\alpha}{\alpha}\left(1 - \frac{y-x}{y-t}\right),$$

which is a strictly increasing function. It is in accordance with part (i) of Proposition 2. Since $\psi(x, \vartheta_1(x)) = 0$, $x \in \mathbb{R}$, by part (vi) of Theorem 1,we have $\psi \in \Psi[T_n^{\boldsymbol{\lambda}}](X,\Theta)$ for all $n \in \mathbb{N}$ and $\boldsymbol{\lambda} \in \Lambda_n$. In particular, since ψ is continuous in its second variable, we also have, for all $n \in \mathbb{N}$ and $x_1, \dots, x_n \in \mathbb{R}$, that the equation $\sum_{i=1}^n \psi(x_i, t) = 0$, $t \in \mathbb{R}$, has a unique solution; see part (ii) of Remark 2.

Now, using Proposition 7, we show that the ψ-expectation function with given ψ and a random variable ξ such that $\mathbf{E}(|\xi|) < \infty$ has a unique zero, which is known to be the α-expectile of ξ; see, for example, Bellini et al. [3, Sect. 2].

Proposition 11. *Let $\psi \in \Psi(\mathbb{R},\mathbb{R})$ be defined by* (4.8)*, let $\alpha \in (0,1)$, and let ξ be a random variable on a probability space $(\Omega, \mathcal{A}, \mathbf{P})$ such that $\mathbf{E}(|\xi|) < \infty$. Then the equation $\mathbf{E}(\psi(\xi,t)) = 0$, $t \in \mathbb{R}$, has a unique solution, which is known to be the α-expectile of ξ.*

We note that Proposition 11 also follows from Lemma A.1 in Krätschmer and Zähle [16], where it was shown that the mapping $\mathbb{R} \ni t \mapsto \mathbf{E}(\psi(\xi,t))$ is real-valued, continuous, strictly decreasing, and satisfying $\lim_{t\to\pm\infty} \mathbf{E}(\psi(\xi,t)) = \mp\infty$. Nonetheless, we give a proof of Proposition 11 using Proposition 7, since we would like to demonstrate the applicability of our result. Note also that if $n \in \mathbb{N}$ and $x_1, \dots, x_n \in \mathbb{R}$, then by choosing ξ as a random variable such that $\mathbf{P}(\xi = x_i) = 1/n$, $i \in \{1, \dots, n\}$ we have $\mathbf{E}(\psi(\xi,t)) =$

$(1/n)\sum_{i=1}^n \psi(x_i,t)$, $t\in\mathbb{R}$, so Proposition 11 yields that the equation $\sum_{i=1}^n \psi(x_i,t)=0$, $t\in\mathbb{R}$, has a unique solution.

Next, we recall a function ψ that has been recently used for constructing M-estimators by Mathieu [18] (see also Example 4), and in case of the Huber function, we take advantage of Proposition 4.

Example 8. Let $X:=\mathbb{R}$, $\Theta:=\mathbb{R}$, and $\psi:\mathbb{R}\times\mathbb{R}\to\mathbb{R}$,

$$\psi(x,t):=\operatorname{sign}(x-t)f\big(|x-t|\big),\quad x,t\in\mathbb{R},\tag{4.9}$$

where $f:\mathbb{R}_+\to\mathbb{R}$. Note that the function ψ given in (4.9) has the same form as that given in (3.1). The only difference is that the function f can take negative values in case of (4.9) but not in case of (3.1). Then $\psi(x,t)=f((x-t)^+)-f((t-x)^+)$, $x,t\in\mathbb{R}$. Note that if f is continuous and $f(0)=0$, then ψ is continuous in its second variable, that is, $\psi\in\Psi[C](\mathbb{R},\mathbb{R})$. We now recall some known special choices for the function f appearing in (4.9) such that the corresponding ψ-estimator has an important role in (robust) statistics:

(i) The Huber function $f_H:\mathbb{R}_+\to\mathbb{R}$,

$$f_H(z):=z\mathbf{1}_{\{z\leqslant\beta\}}+\beta\mathbf{1}_{\{z>\beta\}},\quad z\in\mathbb{R}_+,\tag{4.10}$$

where $\beta>0$ (see Huber [11]), which is a continuous and increasing (but not strictly increasing) function starting from 0. Then the function $\psi:\mathbb{R}\times\mathbb{R}\to\mathbb{R}$ given in (4.9) takes the form

$$\psi(x,t)=\begin{cases}x-t & \text{if } |x-t|\leqslant\beta,\\ \beta\operatorname{sign}(x-t) & \text{if } |x-t|>\beta,\end{cases}\qquad x,t\in\mathbb{R}.$$

In general, ψ is not a T_2-function. Indeed, for example, if $\beta:=1$, $x_1:=0$, and $x_2:=3$, then

$$\psi_{(x_1,x_2)}(t)=\sum_{i=1}^2\psi(x_i,t)=\begin{cases}1+1=2>0 & \text{if } t\leqslant-1,\\ -t+1>0 & \text{if } -1<t\leqslant1,\\ -1+1=0 & \text{if } 1<t\leqslant2,\\ -1+3-t=2-t<0 & \text{if } 2<t<4,\\ -1-1=-2<0 & \text{if } t\geqslant4.\end{cases}$$

Consequently, the function $\psi_{(x_1,x_2)}$ does not have a point of sign change, and thus ψ is not a T_2-function, and by Proposition 4 this yields that ψ is not a T_{2k}-function for any $k\in\mathbb{N}$. On the other hand, for each $k\in\mathbb{N}$, ψ is a T_{2k+1}-function. Indeed, let $k\in\mathbb{N}$ and $x_1,\dots,x_{2k+1}\in\mathbb{R}$ be arbitrary. Then the function $\mathbb{R}\ni t\mapsto\sum_{i=1}^{2k+1}\psi(x_i,t)$ is decreasing (because each of its terms is decreasing), and, on the contrary, let us assume that it does not have a point of sign change. Then it is constant on a proper subinterval I of $\mathbb{R}$. In this case, each term in question must be also constant on this subinterval, and by taking into account the form of ψ it must be equal to β or to $-\beta$ on I. However, a $(2k+1)$-term sum whose terms are equal to β or to $-\beta$ cannot be equal to zero, which leads us to a contradiction. All in all, ψ is a T_n-function if $n\in\mathbb{N}$ is odd, and ψ is not a T_n-function if $n\in\mathbb{N}$ is even.

(ii) The Catoni function $f_C:\mathbb{R}_+\to\mathbb{R}$,

$$f_C(z):=\ln\biggl(1+\frac{z}{b}+\frac{1}{2}\Bigl(\frac{z}{b}\Bigr)^2\biggr),\quad z\in\mathbb{R}_+,$$

where $b>0$ (see Catoni [4]). It is continuous, strictly increasing, and starting from 0.

(iii) The polynomial function $f_P : \mathbb{R}_+ \to \mathbb{R}$,

$$f_P(z) := \frac{z}{1 + (\frac{z}{\beta})^{1-1/p}}, \quad z \in \mathbb{R}_+,$$

where $p \in \mathbb{N}$ and $\beta > 0$. It is continuous, strictly increasing, and starting from 0.

(iv) Another Catoni-type function $\widetilde{f}_C : \mathbb{R}_+ \to \mathbb{R}$,

$$\widetilde{f}_C(z) := \ln\bigg(1 + z + \frac{z^\alpha}{\alpha}\bigg), \quad z \in \mathbb{R}_+,$$

where $\alpha \in (1, 2)$ (see Chen et al. [5]), which is continuous, strictly increasing, and starting from 0.

(v) $f : \mathbb{R}_+ \to \mathbb{R}$, $f(z) := z/\sqrt{1 + z^2/2}$, $z \in \mathbb{R}_+$, which is a continuous strictly increasing function starting from 0. Indeed, we have $f'(z) = (1 + z^2/2)^{-3/2} > 0$ for each $z \in \mathbb{R}_+$. Then the function $\psi : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ given in (4.9) takes the form

$$\psi(x, t) = \frac{x - t}{\sqrt{1 + \frac{(x-t)^2}{2}}}, \quad x, t \in \mathbb{R}.$$

In robust statistics, we call ψ the L_1-L_2 function; see, for example, Rey [22, Sect. 6.4].

(vi) $f : \mathbb{R}_+ \to \mathbb{R}$, $f(z) := z/(1 + z)$, $z \in \mathbb{R}_+$, which is a continuous and strictly increasing function starting from 0. Then the function $\psi : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ given in (4.9) takes the form

$$\psi(x, t) = \frac{x - t}{1 + |x - t|}, \quad x, t \in \mathbb{R},$$

which is called a "fair"-type function in robust statistics; see, for example, Rey [22, Sect. 6.4].

Next, we discuss the applicability of Theorem 1 for the function ψ given in (4.9). Namely, we prove the following statement.

Proposition 12. *Let $f : \mathbb{R}_+ \to \mathbb{R}$ be a function with $f(0) = 0$, and let ψ be given by* (4.9)*. Then we have:*

(a) $\psi \in \Psi[T_1](\mathbb{R}, \mathbb{R})$ *if and only if $f(z) > 0$ for all $z > 0$, and in this case, $\vartheta_1(x) = x$ and $\psi(x, \vartheta_1(x)) = 0$ for all $x \in \mathbb{R}$.*

(b) *if $\psi \in \Psi[T_n](\mathbb{R}, \mathbb{R})$ for infinitely many $n \in \mathbb{N}$, then f is increasing.*

(c) $\psi \in \Psi[T_2^{\boldsymbol{\lambda}}](\mathbb{R}, \mathbb{R})$ *for each $\boldsymbol{\lambda} \in \Lambda_2$ if and only if f is strictly increasing.*

(d) $\psi \in \Psi[T_n^{\boldsymbol{\lambda}}](\mathbb{R}, \mathbb{R})$ *for all $n \in \mathbb{N}$ and $\boldsymbol{\lambda} \in \Lambda_n$ if and only if f is strictly increasing.*

As a consequence of Proposition 12, if $f : \mathbb{R}_+ \to \mathbb{R}$ is a strictly increasing continuous function such that $f(0) = 0$, then ψ given in (4.9) is continuous in its second variable, and hence for all $n \in \mathbb{N}$ and $x_1, \ldots, x_n \in \mathbb{R}$, the equation $\sum_{i=1}^n \psi(x_i, t) = 0$, $t \in \mathbb{R}$ (i.e., Eq. (1.2)) has a unique solution.

In particular cases (ii)–(vi) of Example 8, by part (d) of Proposition 12 the corresponding function ψ given in (4.9) is a $T_n^{\boldsymbol{\lambda}}$-function for all $n \in \mathbb{N}$ and $\boldsymbol{\lambda} \in \Lambda_n$. In particular, since in cases (ii)–(vi), ψ is continuous in its second variable, for all $n \in \mathbb{N}$ and $x_1, \ldots, x_n \in \mathbb{R}$, Eq. (1.2) with the given function ψ has a unique solution; see part (ii) of Remark 2.

In the remaining part of this section, we investigate the applicability of Theorem 1 for finding solutions of likelihood equations in the theory of MLEs. Let Θ be a nonempty open interval of $\mathbb{R}$, and let $f : \mathbb{R} \times \Theta \to \mathbb{R}$ be a function such that for each $t \in \Theta$, the function $\mathbb{R} \ni x \mapsto f(x, t)$ is a density function. Let us introduce the set

$$\mathcal{X}_f := \big\{x \in \mathbb{R}\colon\ f(x, t) > 0\ \forall t \in \Theta\big\}$$

and suppose that $\mathcal{X}_f$ is nonempty. Note that, in general, it can happen that $\mathcal{X}_f = \emptyset$. For example, if $\Theta = (0,\infty)$ and $f : \mathbb{R} \times (0,\infty) \to \mathbb{R}$,

$$f(x,t) := \begin{cases} \frac{1}{t} & \text{if } x \in (0,t), \\ 0 & \text{if } x \notin (0,t), \end{cases} \qquad t \in (0,\infty),$$

that is, for each $t \in (0,\infty)$, the function $\mathbb{R} \ni x \mapsto f(x,t)$ is the density function of a uniformly distributed random variable on $(0,t)$, then $\mathcal{X}_f = \emptyset$. Turning back to the case where $\mathcal{X}_f \neq \emptyset$ and supposing that the (partial) derivative $\partial_2 f$ of f with respect to its second variable exists, Eq. (1.2) with $X := \mathcal{X}_f$ and the function $\psi : \mathcal{X}_f \times \Theta \to \mathbb{R}$ defined by

$$\psi(x,t) := \partial_2\bigl(\ln\bigl(f(x,t)\bigr)\bigr) = \frac{\partial_2 f(x,t)}{f(x,t)}, \quad (x,t) \in \mathcal{X}_f \times \Theta, \tag{4.11}$$

is nothing else but the likelihood equation based on the observations $x_1,\dots,x_n \in \mathcal{X}_f$ in the theory of MLEs. In some cases, we need to consider an appropriate Borel subset $\widetilde{\mathcal{X}}_f$ of $\mathcal{X}_f$ such that $\mathbf{P}(\xi_t \in \widetilde{\mathcal{X}}_f) = 1$ for all $t \in \Theta$, where ξ_t is a random variable having a density function $\mathbb{R} \ni x \mapsto f(x,t)$. For such a case, see the second part of Example 9.

In the next examples, we demonstrate the applicability of Theorem 1 together with Proposition 2 for proving the existence and uniqueness of a solution of the likelihood equation (1.2) corresponding to the function ψ given in (4.11).

Example 9. Let ξ be a normally distributed random variable with mean $m \in \mathbb{R}$ and variance σ^2, where $\sigma > 0$. Let $n \in \mathbb{N}$, and let $x_1,\dots,x_n \in \mathbb{R}$ be a realization of a sample of size n for ξ. Here by a sample of size n we mean independent and identically distributed random variables $\xi_1,\dots,\xi_n$ with common distribution as that of ξ. It is known that supposing that σ is known, there exists a unique MLE of m based on $x_1,\dots,x_n \in \mathbb{R}$, and it takes the form $\widehat{m}_n := (x_1 + \dots + x_n)/n$. We will establish the existence and uniqueness of a solution of the corresponding likelihood equation using Theorem 1 together with Proposition 2. In this case, we have $\Theta = \mathbb{R}$ and $f : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$,

$$f(x,m) := \frac{1}{\sqrt{2\pi}\sigma} \mathrm{e}^{-(x-m)^2/(2\sigma^2)}, \quad x,m \in \mathbb{R},$$

and consequently $\mathcal{X}_f = \mathbb{R}$. Then $\psi : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$,

$$\psi(x,m) = \frac{\frac{1}{\sigma^2}(x-m)f(x,m)}{f(x,m)} = \frac{1}{\sigma^2}(x-m), \quad x,m \in \mathbb{R}.$$

Hence $\psi \in \Psi[C, Z_1](\mathbb{R},\mathbb{R})$ with $\vartheta_1(x) := x$, $x \in \mathbb{R}$, and ψ is strictly decreasing in its second variable. Further, using Proposition 2 and Theorem 1 (with $X := \mathcal{X}_f = \mathbb{R}$), we can conclude that for all $n \in \mathbb{N}$ and $x_1,\dots,x_n \in \mathbb{R}$, the (likelihood) equation (1.2) with the given function ψ has a unique solution, which is equal to $\vartheta_n(x_1,\dots,x_n) = (x_1 + \dots + x_n)/n = \widehat{m}_n$, as desired.

It is also known that supposing that m is known, there exists a unique MLE of σ^2 based on $x_1,\dots,x_n \in \mathbb{R} \setminus \{m\}$, and it takes the form $\widehat{\sigma}_n^2 := (1/n)\sum_{i=1}^n (x_i - m)^2$. We will establish the existence and uniqueness of a solution of the corresponding likelihood equation using Theorem 1. In this case, we have $\Theta = (0,\infty)$, and $f : \mathbb{R} \times (0,\infty) \to \mathbb{R}$,

$$f(x,\sigma^2) := \frac{1}{\sqrt{2\pi\sigma^2}} \mathrm{e}^{-(x-m)^2/(2\sigma^2)}, \quad x \in \mathbb{R},\ \sigma^2 > 0,$$

and consequently $\mathcal{X}_f = \mathbb{R}$. Instead of $\mathcal{X}_f$, let us consider its subset $\widetilde{\mathcal{X}}_f := \mathcal{X}_f \setminus \{m\} = \mathbb{R} \setminus \{m\}$. Then $\mathbf{P}(\xi_{\sigma^2} \in \widetilde{\mathcal{X}}_f) = 1$ for all $\sigma^2 > 0$, where ξ_{σ^2} is a normally distributed random variable with mean m and variance σ^2. Then $\psi : \widetilde{\mathcal{X}}_f \times (0, \infty) \to \mathbb{R}$,

$$\begin{aligned}\psi(x, \sigma^2) &= \frac{1}{f(x, \sigma^2)}\left(\frac{1}{\sqrt{2\pi}}\left(-\frac{1}{2}\right)(\sigma^2)^{-3/2}\mathrm{e}^{-(x-m)^2/(2\sigma^2)} + \frac{1}{\sqrt{2\pi\sigma^2}}\mathrm{e}^{-(x-m)^2/(2\sigma^2)}\frac{(x-m)^2}{2}(\sigma^2)^{-2}\right)\\ &= \frac{1}{2(\sigma^2)^2}\big((x-m)^2 - \sigma^2\big), \quad x \in \widetilde{\mathcal{X}}_f,\ \sigma^2 > 0.\end{aligned}$$

Hence $\psi \in \Psi[C, Z_1](\widetilde{\mathcal{X}}_f, (0, \infty))$ with $\vartheta_1(x) := (x-m)^2$, $x \in \widetilde{\mathcal{X}}_f$. Note that $\vartheta_1(x) \in \Theta = (0, \infty)$ for all $x \in \widetilde{\mathcal{X}}_f$. (This explains the restriction of $\mathcal{X}_f = \mathbb{R}$ to $\widetilde{\mathcal{X}}_f = \mathbb{R} \setminus \{m\}$.) Further, for all $x, y \in \widetilde{\mathcal{X}}_f$ with $\vartheta_1(x) < \vartheta_1(y)$, that is, $(x-m)^2 < (y-m)^2$, we get that the function

$$\big((x-m)^2, (y-m)^2\big) \ni \sigma^2 \mapsto -\frac{\psi(x, \sigma^2)}{\psi(y, \sigma^2)} = -\frac{(x-m)^2 - \sigma^2}{(y-m)^2 - \sigma^2} = -1 + \frac{(y-m)^2 - (x-m)^2}{(y-m)^2 - \sigma^2}$$

is strictly increasing. Consequently, by part (vi) of Theorem 1 (with $X := \widetilde{\mathcal{X}}_f$) we conclude that for all $n \in \mathbb{N}$ and $x_1, \dots, x_n \in \widetilde{\mathcal{X}}_f$, the (likelihood) equation (1.2) with the given ψ has a unique solution, which is equal to $\vartheta_n(x_1, \dots, x_n) = (1/n)\sum_{i=1}^n (x_i - m)^2 = \widehat{\sigma}_n^2$, as desired. Finally, we present an alternative argument. Note that Eq. (1.2) with the given function ψ has a solution if and only if Eq. (1.2) with the function $\widetilde{\psi} : \widetilde{\mathcal{X}}_f \times (0, \infty) \to \mathbb{R}$,

$$\widetilde{\psi}(x, \sigma^2) := 2\big(\sigma^2\big)^2\psi\big(x, \sigma^2\big) = (x-m)^2 - \sigma^2, \quad x \in \widetilde{\mathcal{X}}_f,\ \sigma^2 > 0,$$

has a solution, and the two sets of solutions coincide. Further, $\widetilde{\psi}$ is a T_1-function, and, for each $x \in \widetilde{\mathcal{X}}_f$, the unique point of sign change of the function $(0, \infty) \ni \sigma^2 \mapsto \widetilde{\psi}(x, \sigma^2)$ is $(x-m)^2$. The function $\widetilde{\psi}$ is strictly decreasing in its second variable, and hence part (i) of Proposition 2 can be applied to $\widetilde{\psi}$. This, together with part (vi) of Theorem 1, yields that $\widetilde{\psi} \in \Psi[T_n](\widetilde{\mathcal{X}}_f, (0, \infty))$ for each $n \in \mathbb{N}$, and hence, trivially, $\psi \in \Psi[T_n](\widetilde{\mathcal{X}}_f, (0, \infty))$ for each $n \in \mathbb{N}$ as well, as expected.

Example 10. Let ξ be an absolutely continuous random variable with density function

$$f_\xi(x) := \frac{1}{2} \cdot \frac{1}{\sqrt{2\pi}}\mathrm{e}^{-x^2/2} + \frac{1}{2} \cdot \frac{1}{\sqrt{2\pi}\sigma}\mathrm{e}^{-(x-m)^2/(2\sigma^2)}, \quad x \in \mathbb{R},$$

where $m \in \mathbb{R}$ and $\sigma > 0$. Note that f_ξ is a mixture density function of the standard normal density function and the density function of a normally distributed random variable with mean m and variance σ^2 with equal $1/2$ weights. Let $n \in \mathbb{N}$, and let $x_1, \dots, x_n \in \mathbb{R}$ be a realization of a sample of size n for ξ. In what follows, we assume that σ is known, and using Theorem 1, we show that, in general, the corresponding likelihood equation for m may have more solutions. In this case, we have $\Theta = \mathbb{R}$ and $f : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$,

$$f(x, m) := \frac{1}{2} \cdot \frac{1}{\sqrt{2\pi}}\mathrm{e}^{-x^2/2} + \frac{1}{2} \cdot \frac{1}{\sqrt{2\pi}\sigma}\mathrm{e}^{-(x-m)^2/(2\sigma^2)}, \quad x, m \in \mathbb{R},$$

and consequently $\mathcal{X}_f = \mathbb{R}$. Then $\psi : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$,

$$\psi(x, m) = \frac{\frac{1}{2\sigma^3\sqrt{2\pi}}(x-m)\mathrm{e}^{-(x-m)^2/(2\sigma^2)}}{f(x, m)}, \quad x, m \in \mathbb{R}.$$

Hence $\psi \in \Psi[C, T_1](\mathbb{R}, \mathbb{R})$ with $\vartheta_1(x) := x$, $x \in \mathbb{R}$. Consequently, for all $n \in \mathbb{N}$ and $\boldsymbol{x} = (x_1, \dots, x_n) \in X^n$, the likelihood equation (1.2) has at least one solution in $\mathbb{R}$. Indeed, if $m < \min(x_1, \dots, x_n)$, then $\sum_{i=1}^n \psi(x_i, m) > 0$, and if $m > \max(x_1, \dots, x_n)$, then $\sum_{i=1}^n \psi(x_i, m) < 0$, and hence the continuity of ψ in

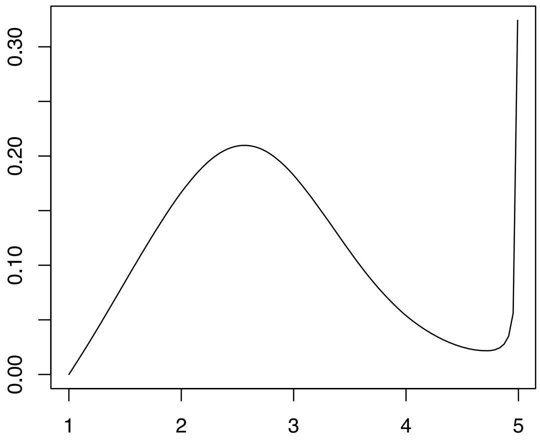

Figure 1. The function $(1,5) \ni m \mapsto -\psi(x,m)/\psi(y,m)$ with $\sigma = 1$.

its second variable, together with the Bolzano theorem, implies the existence of a solution of (1.2), as desired. Further, for each $x \in \mathbb{R}$, we have $\lim_{m\to\pm\infty} \psi(x,m) = 0$. In what follows, we check that it is not true that for all $x, y \in \mathbb{R}$ with $\vartheta_1(x) < \vartheta_1(y)$, that is,, $x < y$, the function (2.2) is increasing. For all $x, y \in \mathbb{R}$ with $x < y$, the function (2.2) takes the form

$$(x,y) \ni m \mapsto -\frac{\psi(x,m)}{\psi(y,m)} = \frac{m-x}{y-m} \cdot \frac{f(y,m)}{f(x,m)} \cdot \frac{\mathrm{e}^{(y-m)^2/(2\sigma^2)}}{\mathrm{e}^{(x-m)^2/(2\sigma^2)}} = \frac{(m-x)(\sigma \mathrm{e}^{-y^2/2+(y-m)^2/(2\sigma^2)}+1)}{(y-m)(\sigma \mathrm{e}^{-x^2/2+(x-m)^2/(2\sigma^2)}+1)},$$

and, for each $m \in (x,y)$, we can check that

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}m}\biggl(-\frac{\psi(x,m)}{\psi(y,m)}\biggr) &= \frac{1}{(y-m)^2(\sigma \mathrm{e}^{-x^2/2+(x-m)^2/(2\sigma^2)}+1)^2} \\ &\quad \times \biggl[(y-x)\bigl(\sigma \mathrm{e}^{-x^2/2+(x-m)^2/(2\sigma^2)}+1\bigr)\bigl(\sigma \mathrm{e}^{-y^2/2+(y-m)^2/(2\sigma^2)}+1\bigr) \\ &\quad - \frac{1}{\sigma}(m-x)(y-m)^2 \mathrm{e}^{-y^2/2+(y-m)^2/(2\sigma^2)}\bigl(\sigma \mathrm{e}^{-x^2/2+(x-m)^2/(2\sigma^2)}+1\bigr) \\ &\quad - \frac{1}{\sigma}(m-x)^2(y-m)\mathrm{e}^{-x^2/2+(x-m)^2/(2\sigma^2)}\bigl(\sigma \mathrm{e}^{-y^2/2+(y-m)^2/(2\sigma^2)}+1\bigr)\biggr]. \end{aligned} \tag{4.12}$$

If for all $x, y \in \mathbb{R}$ with $x < y$, the function (2.2) were increasing, then we would have that

$$\frac{\mathrm{d}}{\mathrm{d}m}\biggl(-\frac{\psi(x,m)}{\psi(y,m)}\biggr) \geqslant 0, \quad m \in (x,y),\ x<y,\ x,y \in \mathbb{R},$$

which by (4.12) is equivalent to

$$\sigma(y-x) \geqslant \frac{(m-x)(y-m)^2}{\sigma + \mathrm{e}^{y^2/2-(y-m)^2/(2\sigma^2)}} + \frac{(m-x)^2(y-m)}{\sigma + \mathrm{e}^{x^2/2-(x-m)^2/(2\sigma^2)}}, \quad m \in (x,y),\ x<y,\ x,y \in \mathbb{R}.$$

However, this inequality does not hold in general, since its left-hand side tends to 0 as $\sigma \downarrow 0$, but its right-hand side tends to ∞ as $\sigma \downarrow 0$. To give an example, on Fig. 1, we plotted the function $(x,y) \ni m \mapsto -\psi(x,m)/\psi(y,m)$ with $x = 1$, $y = 5$, and $\sigma = 1$, which is not increasing. In general, it is not true that for all $x, y \in \mathbb{R}$ with $x < y$, the function (2.2) is increasing. If the function (2.2) is not increasing for some $x, y \in \mathbb{R}$ with $x < y$, then by part (iv) of Theorem 1 (with $X := \mathcal{X}_f = \mathbb{R}$) we get that there exists $n_0 \in \mathbb{N}$ such that ψ is not a T_n-function for any $n \geqslant n_0$, $n \in \mathbb{N}$. In particular, this yields that there exists $n_0 \in \mathbb{N}$ such that for each $n \geqslant n_0$, there exist real numbers $x_1, \ldots, x_n \in \mathbb{R}$ such that the likelihood equation (1.2) based on $\boldsymbol{x} = (x_1, \ldots, x_n)$ has at least two solutions. In such a case, by part (v) of Theorem 1 we also get that there exists a $\boldsymbol{\lambda} \in \Lambda_2$ such that ψ is not a $T_2^{\boldsymbol{\lambda}}$-function.

A further example is presented in the arXiv version [2, Ex. 4.9] of this paper.

5 Proofs for Section 2

Proof of Lemma 1. Assume that y is a level of increase for f and (i) and (ii) are not valid. Define

$$A := \{v \in \Theta\colon y \geqslant f(v)\} \quad \text{and} \quad B := \{u \in \Theta\colon y \leqslant f(u)\}.$$

Then $A \cup B = \Theta$, and A and B are nonempty, since (i) and (ii) do not hold. If $v \in A$ and $u \in B$, then $v \leqslant u$ (since y is a level of increase for f). Consequently, $\sup A \leqslant \inf B$, and using that $A \cup B = \Theta$, we have that $\sup A = \inf B =: \vartheta$. If $t < \vartheta$, then $t \notin B$, which implies that $y > f(t)$, that is, $y - f(t) > 0$. Similarly, if $\vartheta < t$, then $t \notin A$, which yields $y < f(t)$, that is, $y - f(t) < 0$. Hence we have proved that ϑ is a point of sign change for the function $y - f$, that is, (iii) must hold.

Conversely, if (i) or (ii) holds, then just using the definition, we have that y is a level of increase for f. Finally, assume that (iii) is valid, that is, there exists $\vartheta \in \Theta$ that is a point of sign change for the function $y - f$, and, on the contrary, suppose that y is not a level of increase for f. Then there exist $u, v \in \Theta$ such that $u < v$ and $f(v) \leqslant y \leqslant f(u)$. Therefore $y - f(v) \geqslant 0 \geqslant y - f(u)$, which, using the definition of a point of sign change, implies that $v \leqslant \vartheta \leqslant u$, contradicting $u < v$. □

Proof of Lemma 2. Assume that the levels of increase for f form a dense subset in the convex hull of $f(\Theta)$, but f is not increasing. Then there exist $u, v \in \Theta$ with $u < v$ such that $f(u) > f(v)$. The convex hull of $f(\Theta)$ contains the open interval $(f(v), f(u))$. Hence by the assumption we can find an element $y \in (f(v), f(u))$ that is a level of increase for f. Therefore, by definition, $v \leqslant u$, which contradicts $u < v$. The second statement of the lemma readily follows from the definitions of strictly increasing property and level of increase. For the last statement of the lemma, assume that $u, v \in \Theta$ satisfy the inequalities

$$g\big(f(v)\big) = (g \circ f)(v) \leqslant g(y) \leqslant (g \circ f)(u) = g\big(f(u)\big).$$

Since g is strictly increasing, we get that $f(v) \leqslant y \leqslant f(u)$. Using that y is a level of increase for f, it follows that $v \leqslant u$, implying the third statement of the lemma. □

Proof of Lemma 3. To verify the ε-increasingness property of f, let $u, v \in \Theta$ be arbitrary such that $u < v$. The intervals $J_1 := [y_0, y_1], \ldots, J_n := [y_{n-1}, y_n]$ cover the image $f(\Theta)$. Therefore, for some $i \in \{1, \ldots, n\}$, we have that $f(u) \in J_i$, which implies $y_{i-1} \leqslant f(u) \leqslant y_i$. Since y_{i-1} is a level of increase for f, by part (iii) of Remark 4 we have that $y_{i-1} < f(v)$. On the other hand, due to the definition of ε, it follows that $y_i \leqslant y_{i-1} + \varepsilon$. Thus we get

$$f(u) \leqslant y_i \leqslant y_{i-1} + \varepsilon < f(v) + \varepsilon,$$

which was to be proved. □

Proof of Theorem 1. (i) Let $x, y \in X$ be such that $\vartheta_1(x) < \vartheta_1(y)$. Since $\psi \in \Psi[T_1](X, \Theta)$, for all $t \in (\vartheta_1(x), \vartheta_1(y))$, we have $\psi(x, t) < 0$ and $\psi(y, t) > 0$.

To the contrary, assume that λ_2/λ_1 is not a level of increase for the function (2.2). Then there exist $u, v \in \mathbb{R}$ such that $\vartheta_1(x) < u < v < \vartheta_1(y)$ and

$$-\frac{\psi(x, v)}{\psi(y, v)} \leqslant \frac{\lambda_2}{\lambda_1} \leqslant -\frac{\psi(x, u)}{\psi(y, u)}.$$

Rearranging these inequalities, we get that

$$\lambda_1\psi(x,u)+\lambda_2\psi(y,u)\leqslant 0\leqslant \lambda_1\psi(x,v)+\lambda_2\psi(y,v).$$

In view of the property $[T_2^{(\lambda_1,\lambda_2)}]$ of ψ, this implies that

$$v\leqslant \vartheta_{2,\psi}^{(\lambda_1,\lambda_2)}(x,y)\leqslant u,$$

which contradicts the inequality $u<v$.

The argument related to λ_1/λ_2 is analogous.

(ii) Let $\psi\in\Psi[T_n^{(\lambda_1,\dots,\lambda_n)}](X,\Theta)$ for some $n\in\mathbb{N}\setminus\{1\}$ and $(\lambda_1,\dots,\lambda_n)\in(0,\infty)^n$. Then for each $k\in\{1,\dots,n-1\}$, we have ψ is a $T_2^{(\sum_{i=1}^{k}\lambda_i,\sum_{i=k+1}^{n}\lambda_i)}$-function, since for all $x,y\in X$ and $k\in\{1,\dots,n-1\}$,

$$\left(\sum_{i=1}^{k}\lambda_i\right)\psi(x,t)+\left(\sum_{i=k+1}^{n}\lambda_i\right)\psi(y,t)=\sum_{i=1}^{n}\lambda_i\psi(x_i,t),\quad t\in\Theta,$$

where $x_i:=x$, $i\in\{1,\dots,k\}$, and $x_i:=y$, $i\in\{k+1,\dots,n\}$. Consequently, the statement readily follows from part (i) of the present theorem.

(iii) If ψ is a T_n-function for some $n\in\mathbb{N}\setminus\{1\}$, then it is a $T_n^{(\lambda_1,\dots,\lambda_n)}$-function with $\lambda_1=\dots=\lambda_n:=1$, and thus the statement follows from part (ii) of the present theorem.

(iv) Let $(n_i)_{i\in\mathbb{N}}\subseteq\mathbb{N}$ be a strictly increasing sequence such that $\psi\in\Psi[T_{n_i}](X,\Theta)$ for all $i\in\mathbb{N}$. Let $x,y\in X$ be such that $\vartheta_1(x)<\vartheta_1(y)$. Then by statement (iii) of this theorem we have that the numbers

$$\left\{\frac{k}{n_i-k},\ k\in\{1,\dots,n_i-1\},\ i\in\mathbb{N}\right\}=\left\{\frac{1}{n_i-1},\frac{2}{n_i-2},\dots,\frac{n_i-2}{2},n_i-1,\ i\in\mathbb{N}\right\}$$

are levels of increase for the function (2.2). We are going to apply Lemma 2. The convex hull of the range of the function (2.2) is contained in $(0,\infty)$, so if we check that the set $\{k/(n_i-k)\ k\in\{1,\dots,n_i-1\},\ i\in\mathbb{N}\}$ is dense in $(0,\infty)$, then Lemma 2 will imply that the function (2.2) is increasing. Since the function $g:(0,\infty)\to(0,1)$, $g(u):=u/(u+1)$, $u>0$, is bijective, it suffices to check that the set

$$\left\{g\left(\frac{k}{n_i-k}\right),\ k\in\{1,\dots,n_i-1\},\ i\in\mathbb{N}\right\}=\left\{\frac{k}{n_i},\ k\in\{1,\dots,n_i-1\},\ i\in\mathbb{N}\right\}$$

is dense in $g((0,\infty))=(0,1)$. This readily follows, since if $(a,b)\subseteq(0,1)$ is an open interval, then there exists $i_0\in\mathbb{N}$ such that $1/n_{i_0}<b-a$ (due to $n_i\to\infty$ as $i\to\infty$), and hence there exists $k\in\{1,\dots,n_{i_0}-1\}$ such that $k/n_{i_0}\in(a,b)$.

Now we turn to prove the second statement of statement (iv). Let $\ell,m\in\mathbb{N}$. We show that ℓ/m is a level of increase for the function (2.2). By assumption there exists $i_0\in\mathbb{N}$ such that $m+\ell$ divides n_{i_0} and $\psi\in\Psi[T_{n_{i_0}}](X,\Theta)$. By part (iii) of this theorem we have that the elements of the set

$$\left\{\frac{k}{n_{i_0}-k},\ k\in\{1,\dots,n_{i_0}-1\}\right\}$$

are levels of increase for the function (2.2). By choosing $k:=\ell n_{i_0}/(m+\ell)\in\{1,\dots,n_{i_0}-1\}$ we have

$$\frac{k}{n_{i_0}-k}=\frac{\ell n_{i_0}}{(m+\ell)n_{i_0}-\ell n_{i_0}}=\frac{\ell}{m},$$

yielding that ℓ/m is a level of increase for the function (2.2), as desired.

(v) Assume that $\psi \in \Psi[T_2^{\boldsymbol{\lambda}}](X, \Theta)$ for each $\boldsymbol{\lambda} \in \Lambda_2$. Let $x, y \in X$ be such that $\vartheta_1(x) < \vartheta_1(y)$. By statement (i) it follows that λ_1/λ_2 is a level of increase for the function (2.2) for all $\lambda_1, \lambda_2 > 0$. Since λ_1 and λ_2 are arbitrary positive numbers, we get that each positive number is a level of increase for the positive function (2.2). In view of the second statement of Lemma 2, this implies that the function (2.2) is strictly increasing.

(vi) This statement is an immediate consequence of Theorem 2 as stated in Corollary 2, and therefore its proof is omitted here. □

Proof of Corollary 1. If (i) holds, then part (vi) of Theorem 1 yields that (iii) is valid as well. If (iii) holds, then (ii) is readily satisfied. Finally, if (ii) holds, then part (v) of Theorem 1 implies the validity of (i). □

Proof of Proposition 2. (i) First, let us suppose that for each $x \in X$, the function $\Theta \ni t \mapsto \psi(x, t)$ is decreasing. Let $\vartheta_1(x) < s < t < \vartheta_1(y)$. Then, since ψ is a T_1-function, we have

$$0 > \psi(x, s) \geqslant \psi(x, t) \quad \text{and} \quad \psi(y, s) \geqslant \psi(y, t) > 0.$$

Consequently, we get

$$0 < -\psi(x, s)\psi(y, t) \leqslant -\psi(x, t)\psi(y, s),$$

which is equivalent to

$$-\frac{\psi(x, s)}{\psi(y, s)} \leqslant -\frac{\psi(x, t)}{\psi(y, t)},$$

yielding that the function (2.2) is increasing. The case where the function $\Theta \ni t \mapsto \psi(x, t)$ is strictly decreasing for each $x \in X$ can be handled similarly.

(ii) Let us suppose that for each $x \in X$, the function $\Theta \ni t \mapsto \psi(x, t)$ is strictly decreasing. Let $n \in \mathbb{N}$, $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_n) \in \Lambda_n$, and $(x_1, \ldots, x_n) \in X^n$. Since $\lambda_i \geqslant 0$, $i \in \{1, \ldots, n\}$, and $\lambda_1 + \cdots + \lambda_n > 0$, we get that the function $\Theta \ni t \mapsto \sum_{i=1}^n \lambda_i \psi(x_i, t)$ is strictly decreasing. Using that $\psi \in \Psi[T_1](X, \Theta)$, we have that

$$\sum_{i=1}^n \lambda_i \psi(x_i, t) \begin{cases} > 0 & \text{if } t < \min(\vartheta_1(x_1), \ldots, \vartheta_1(x_n)), \\ < 0 & \text{if } t > \max(\vartheta_1(x_1), \ldots, \vartheta_1(x_n)). \end{cases}$$

Consequently, we have

$$\begin{aligned} t^* &:= \sup\left\{t \in \Theta\colon \sum_{i=1}^n \lambda_i \psi(x_i, t) > 0\right\} \leqslant \max\bigl(\vartheta_1(x_1), \ldots, \vartheta_1(x_n)\bigr), \\ t_* &:= \inf\left\{t \in \Theta\colon \sum_{i=1}^n \lambda_i \psi(x_i, t) < 0\right\} \geqslant \min\bigl(\vartheta_1(x_1), \ldots, \vartheta_1(x_n)\bigr), \end{aligned}$$

yielding that $t^* \leqslant t_*$. Using the definition of infimum and supremum and that the map $\Theta \ni t \mapsto \sum_{i=1}^n \lambda_i \psi(x_i, t)$ is strictly decreasing, we get $t^* = t_*$. Indeed, if $t^* < t_*$ were true, then, by the definition of infimum and supremum, $\sum_{i=1}^n \lambda_i \psi(x_i, t) = 0$, $t \in (t^*, t_*)$, would hold, contradicting the strictly decreasing property of the function $\Theta \ni t \mapsto \sum_{i=1}^{\lambda} \lambda_i \psi(x_i, t)$. All in all, we get $\vartheta_{n,\psi}^{\boldsymbol{\lambda}}(x_1, \ldots, x_n) = t^* = t_*$, and then ψ possesses the property $[T_n^{\boldsymbol{\lambda}}]$. □

Proof of Proposition 3. Let $x, y \in X$ be such that $\vartheta_1(x) < \vartheta_1(y)$. Since $\psi \in \Psi[T_1](X, \Theta)$, for all $t \in (\vartheta_1(x), \vartheta_1(y))$, we have $\psi(x, t) < 0$ and $\psi(y, t) > 0$, and hence the function (2.3) takes values in $(0, 1)$. By part (iii) of Theorem 1 we have that the elements of the set $\{k/(n-k),\ k \in \{1, \ldots, n-1\}\}$ are levels of increase for the function (2.2). Note that

$$\frac{\psi(x,t)}{\psi(x,t)-\psi(y,t)} = g\biggl(-\frac{\psi(x,t)}{\psi(y,t)}\biggr), \quad t \in \bigl(\vartheta_1(x), \vartheta_1(y)\bigr),$$

where $g : (0,\infty) \to \mathbb{R}$, $g(u) := u/(u+1)$, $u > 0$. Since g is strictly increasing and $(0,\infty)$ contains the range of the function (2.2), by Lemma 2 we get that the elements of the set

$$\biggl\{g\biggl(\frac{k}{n-k}\biggr),\ k \in \{1,\dots,n-1\}\biggr\} = \biggl\{\frac{k}{n},\ k \in \{1,\dots,n-1\}\biggr\}$$

are levels of increase for the function (2.3). Since the function (2.2) is positive, we readily have that 0 is also a level of increase for the function (2.2). Consequently, using that the function (2.3) takes values in $(0,1)$, the conditions of Lemma 3 are satisfied with the choices $y_k := k/n$, $k \in \{0,1,\dots,n\}$, and then we get that the function (2.3) is strictly ε-increasing with $\varepsilon := \max\{k/n-(k-1)/n,\ k \in \{1,\dots,n\}\} = 1/n$, as desired. □

Proof of Proposition 4. Now let us suppose that $m \in \{1,\dots,n\}$ is a divisor of n. Then there exists $k \in \mathbb{N}$ such that $n = km$. Consequently, for each $(y_1,\dots,y_m) \in X^m$, with the notation

$$(x_1,\dots,x_n) := (\underbrace{y_1,\dots,y_1}_{k}, \underbrace{y_2,\dots,y_2}_{k}, \dots, \underbrace{y_m,\dots,y_m}_{k}) \in X^n,$$

using that ψ is a T_n-function, we have that

$$k\sum_{i=1}^{m}\psi(y_i,t) = \sum_{i=1}^{n}\psi(x_i,t) \begin{cases} >0 & \text{if } t < \vartheta_n(x_1,\dots,x_n), \\ <0 & \text{if } t > \vartheta_n(x_1,\dots,x_n). \end{cases}$$

Hence $\psi \in \Psi[T_m](X,\Theta)$ with $\vartheta_m(y_1,\dots,y_m) := \vartheta_n(x_1,\dots,x_n)$. □

Proof of Proposition 5. For all $\boldsymbol{y} := (y_1,\dots,y_m) \in X^m$ and $t \in \Theta$, we have

$$\psi_{\boldsymbol{y},\boldsymbol{\mu}}(t) := \sum_{\alpha=1}^{m}\mu_\alpha\psi(y_\alpha,t) = \sum_{\alpha=1}^{m}\biggl(\sum_{i\in H_\alpha}\lambda_i\biggr)\psi(y_\alpha,t) = \sum_{j=1}^{n}\lambda_j\psi(x_j,t) = \psi_{\boldsymbol{x},\boldsymbol{\lambda}}(t), \tag{5.1}$$

where $\boldsymbol{x} := (x_1,\dots,x_n) \in X^n$ is such that $x_j := y_\alpha$ if $j \in H_\alpha$. By the assumption the value $\vartheta_n^{\boldsymbol{\lambda}}(x_1,\dots,x_n)$ is a point of sign change for the function $\psi_{\boldsymbol{x},\boldsymbol{\lambda}}$. Therefore by (5.1) we can see that the function $\psi_{\boldsymbol{y},\boldsymbol{\mu}}$ has the same point of sign change, and hence we have

$$\vartheta_m^{\boldsymbol{\mu}}(y_1,\dots,y_m) = \vartheta_n^{\boldsymbol{\lambda}}(x_1,\dots,x_n),$$

yielding that $\psi \in \Psi[T_m^{\boldsymbol{\mu}}](X,\Theta)$. □

Proof of Proposition 6. First, we check that $\psi \in \Psi[T_1](X,\Theta)$ with $\vartheta_1 = f^{(-1)} \circ \varphi$. Let $x \in X$ be fixed. If $t < (f^{(-1)} \circ \varphi)(x)$, $t \in \Theta$, then $\psi(x,t) > 0$, since otherwise $\psi(x,t) \leqslant 0$ would yield that $\varphi(x) \leqslant f(t)$, and hence by Lemma 4 we would have that $(f^{(-1)} \circ \varphi)(x) \leqslant (f^{(-1)} \circ f)(t) = t$, leading us to a contradiction. Similarly, if $t > (f^{(-1)} \circ \varphi)(x)$, $t \in \Theta$, then $\psi(x,t) < 0$, since otherwise $\psi(x,t) \geqslant 0$ would yield that $\varphi(x) \geqslant f(t)$, and hence by Lemma 4 we would have that $(f^{(-1)} \circ \varphi)(x) \geqslant (f^{(-1)} \circ f)(t) = t$, leading us to a contradiction. All in all, for each $x \in X$, we have that

$$\psi(x,t) \begin{cases} > 0 & \text{if } t < (f^{(-1)} \circ \varphi)(x), \\ < 0 & \text{if } t > (f^{(-1)} \circ \varphi)(x), \end{cases} \qquad t \in \Theta,$$

as desired.

Consequently, using also that for each $x \in X$, the function $\Theta \ni t \mapsto \psi(x,t) = p(x)(\varphi(x) - f(t))$ is strictly decreasing, part (ii) of Proposition 2 implies that $\psi \in \Psi[T_n^{\boldsymbol{\lambda}}](X,\Theta)$ for all $n \in \mathbb{N}$ and $\boldsymbol{\lambda} \in \Lambda_n$.

It remains to check that (2.5) holds. First, note that the right-hand side of (2.5) is well defined, since

$$\begin{aligned} &\frac{\lambda_1 p(x_1)\varphi(x_1) + \cdots + \lambda_n p(x_n)\varphi(x_n)}{\lambda_1 p(x_1) + \cdots + \lambda_n p(x_n)} \\ &\quad = \frac{\lambda_1 p(x_1)}{\lambda_1 p(x_1) + \cdots + \lambda_n p(x_n)}\varphi(x_1) + \cdots + \frac{\lambda_n p(x_n)}{\lambda_1 p(x_1) + \cdots + \lambda_n p(x_n)}\varphi(x_n) \\ &\quad \in \operatorname{conv}(\varphi(X)) \subseteq \operatorname{conv}(f(\Theta)), \end{aligned}$$

and $f^{(-1)}$ is defined on $\operatorname{conv}(f(\Theta))$ (see Lemma 4). Let $n \in \mathbb{N}$, $(x_1,\dots,x_n) \in X^n$, and $\boldsymbol{\lambda} \in \Lambda_n$ be fixed. If

$$t < f^{(-1)}\left(\frac{\lambda_1 p(x_1)\varphi(x_1) + \cdots + \lambda_n p(x_n)\varphi(x_n)}{\lambda_1 p(x_1) + \cdots + \lambda_n p(x_n)}\right), \quad t \in \Theta,$$

then $\sum_{i=1}^n \lambda_i \psi(x_i,t) > 0$, since otherwise $\sum_{i=1}^n \lambda_i \psi(x_i,t) \leqslant 0$ would yield that

$$\frac{\lambda_1 p(x_1)\varphi(x_1) + \cdots + \lambda_n p(x_n)\varphi(x_n)}{\lambda_1 p(x_1) + \cdots + \lambda_n p(x_n)} \leqslant f(t).$$

Hence by Lemma 4 we would have that

$$f^{(-1)}\left(\frac{\lambda_1 p(x_1)\varphi(x_1) + \cdots + \lambda_n p(x_n)\varphi(x_n)}{\lambda_1 p(x_1) + \cdots + \lambda_n p(x_n)}\right) \leqslant \big(f^{(-1)} \circ f\big)(t) = t,$$

leading us to a contradiction. Similarly, we can easily see that the inequality

$$t > f^{(-1)}\left(\frac{\lambda_1 p(x_1)\varphi(x_1) + \cdots + \lambda_n p(x_n)\varphi(x_n)}{\lambda_1 p(x_1) + \cdots + \lambda_n p(x_n)}\right), \quad t \in \Theta,$$

implies $\sum_{i=1}^n \lambda_i \psi(x_i,t) < 0$. These two properties, together with $\psi \in \Psi[T_n^{\boldsymbol{\lambda}}](X,\Theta)$, yield equality (2.5). □

6 Proofs for Section 3

Proof of Theorem 2. Define the sets $U, V \subseteq \Theta$ by

$$U := \big\{s \in \Theta\colon\ \mathbf{E}\big(\psi(\xi,s)\big) \geqslant 0\big\} \quad \text{and} \quad V := \big\{t \in \Theta\colon\ \mathbf{E}\big(\psi(\xi,t)\big) \leqslant 0\big\}.$$

Then, in view of assumption (v), we have that $s_0 \in U$ and $t_0 \in V$. In what follows, we show that $s \leqslant t$ for all $s \in U$ and $t \in V$. To the contrary, assume that $t < s$, and for any Borel subset $H \subseteq \Theta$, define

$$\Omega_H := \big\{\omega \in \Omega\colon\ \vartheta_1\big(\xi(\omega)\big) \in H\big\}.$$

Then $\mathit{\Omega}_H \in \mathcal{A}$ due to the measurability of $\vartheta_1 : X \to \Theta$ and $\xi : \mathit{\Omega} \to X$. Indeed, for each $r \in \Theta$, we have that

$$\vartheta_1^{-1}((-\infty, r)) = \bigl\{x \in X\colon \vartheta_1(x) < r\bigr\} = \bigl\{x \in X\colon \psi(x, r) < 0\bigr\} \in \mathcal{X},$$

where we used assumptions (i) and (iii) and that the sigma-algebra generated by the family $\{(-\infty, r) \cap \Theta,$ $r \in \Theta\}$ coincides with the Borel sigma-algebra on Θ.

Consider the following partition of Θ, which is induced by t and s:

$$I := \Theta \cap (-\infty, t), \qquad J := [t, s], \qquad K := \Theta \cap (s, \infty).$$

Then using assumption (i), we have

$$\begin{aligned}\mathit{\Omega}_I &= \bigl\{\omega \in \mathit{\Omega}\colon \vartheta_1(\xi(\omega)) < t\bigr\} = \bigl\{\omega \in \mathit{\Omega}\colon \psi(\xi(\omega), t) < 0\bigr\},\\ \mathit{\Omega}_K &= \bigl\{\omega \in \mathit{\Omega}\colon \vartheta_1(\xi(\omega)) > s\bigr\} = \bigl\{\omega \in \mathit{\Omega}\colon \psi(\xi(\omega), s) > 0\bigr\}.\end{aligned}$$

We show that $\mathbf{P}(\mathit{\Omega}_I) > 0$ and $\mathbf{P}(\mathit{\Omega}_K) > 0$. Indeed, on the contrary, if $\mathbf{P}(\mathit{\Omega}_I) = 0$, then $\mathbf{P}(\psi(\xi, t) \geqslant 0) = 1$, which implies that $\mathbf{E}(\psi(\xi, t)) \geqslant 0$. By the inclusion $t \in V$ we also have that $\mathbf{E}(\psi(\xi, t)) \leqslant 0$ and hence $\mathbf{E}(\psi(\xi, t)) = 0$. Therefore $\mathbf{P}(\psi(\xi, t) = 0) = 1$, that is, $\mathbf{P}(\vartheta_1(\xi) = t) = 1$. It follows from the inequality $t < s$ that $\mathbf{P}(\vartheta_1(\xi) < s) = 1$, and hence $\mathbf{P}(\psi(\xi, s) < 0) = 1$. This implies that $\mathbf{E}(\psi(\xi, s)) < 0$, which contradicts that s belongs to U. The equality $\mathbf{P}(\mathit{\Omega}_K) = 0$ leads to a contradiction similarly.

The inequalities $\mathbf{P}(\mathit{\Omega}_I) > 0$ and $\mathbf{P}(\mathit{\Omega}_K) > 0$ imply that $\mathit{\Omega}_I \neq \emptyset$ and $\mathit{\Omega}_K \neq \emptyset$. Then, for all $\omega' \in \mathit{\Omega}_I$ and $\omega'' \in \mathit{\Omega}_K$, we have that

$$\vartheta_1\bigl(\xi(\omega')\bigr) < t < s < \vartheta_1\bigl(\xi(\omega'')\bigr).$$

Therefore, using assumption (ii) with $x := \xi(\omega')$ and $y := \xi(\omega'')$, the function (2.2) is strictly increasing, and hence we get

$$\frac{\psi(\xi(\omega'), s)}{\psi(\xi(\omega''), s)} < \frac{\psi(\xi(\omega'), t)}{\psi(\xi(\omega''), t)}.$$

Using that $\psi(\xi(\omega''), s) > 0$ and $\psi(\xi(\omega''), t) > 0$, we can obtain that

$$\psi\bigl(\xi(\omega'), s\bigr)\psi\bigl(\xi(\omega''), t\bigr) < \psi\bigl(\xi(\omega'), t\bigr)\psi\bigl(\xi(\omega''), s\bigr), \quad (\omega', \omega'') \in \mathit{\Omega}_I \times \mathit{\Omega}_K. \tag{6.1}$$

Integrating on $\mathit{\Omega}_I$ and then on $\mathit{\Omega}_K$ with respect to $\mathbf{P}$, it follows that

$$\begin{aligned}&\int\limits_{\mathit{\Omega}_I} \psi\bigl(\xi(\omega'), s\bigr)\,\mathrm{d}\mathbf{P}(\omega') \cdot \int\limits_{\mathit{\Omega}_K} \psi\bigl(\xi(\omega''), t\bigr)\,\mathrm{d}\mathbf{P}(\omega'')\\ &\qquad < \int\limits_{\mathit{\Omega}_I} \psi\bigl(\xi(\omega'), t\bigr)\,\mathrm{d}\mathbf{P}(\omega') \cdot \int\limits_{\mathit{\Omega}_K} \psi\bigl(\xi(\omega''), s\bigr)\,\mathrm{d}\mathbf{P}(\omega''),\end{aligned}$$

that is,

$$\mathbf{E}(\psi(\xi, s)\mathbf{1}_{\mathit{\Omega}_I}) \cdot \mathbf{E}(\psi(\xi, t)\mathbf{1}_{\mathit{\Omega}_K}) < \mathbf{E}(\psi(\xi, t)\mathbf{1}_{\mathit{\Omega}_I}) \cdot \mathbf{E}(\psi(\xi, s)\mathbf{1}_{\mathit{\Omega}_K}). \tag{6.2}$$

Inequality (6.2) is indeed strict because the left-hand side of (6.1) is strictly smaller than its right-hand side over the set $\mathit{\Omega}_I \times \mathit{\Omega}_K$ of positive measure with respect to the product probability $\mathbf{P} \otimes \mathbf{P}$.

Furthermore, using also that $t < s$, for all $\omega' \in \mathit{\Omega}_J$ and $\omega'' \in \mathit{\Omega}_K$, we have that $\psi(\xi(\omega'), s) \leqslant 0$, $\psi(\xi(\omega''), t) > 0$, $\psi(\xi(\omega'), t) \geqslant 0$, and $\psi(\xi(\omega''), s) > 0$. Therefore

$$\psi\bigl(\xi(\omega'), s\bigr)\psi\bigl(\xi(\omega''), t\bigr) \leqslant 0 \leqslant \psi\bigl(\xi(\omega'), t\bigr)\psi\bigl(\xi(\omega''), s\bigr), \quad (\omega', \omega'') \in \mathit{\Omega}_J \times \mathit{\Omega}_K.$$

Integrating on Ω_J and then on Ω_K with respect to $\mathbf{P}$, it follows that

$$\mathbf{E}\big(\psi(\xi,s)\mathbf{1}_{\Omega_J}\big)\cdot\mathbf{E}\big(\psi(\xi,t)\mathbf{1}_{\Omega_K}\big)\leqslant\mathbf{E}\big(\psi(\xi,t)\mathbf{1}_{\Omega_J}\big)\cdot\mathbf{E}\big(\psi(\xi,s)\mathbf{1}_{\Omega_K}\big). \tag{6.3}$$

Adding up inequalities (6.2) and (6.3), since Ω_I and Ω_J are disjoint, we get

$$\begin{aligned} A(s)B(t) &:= \mathbf{E}\big(\psi(\xi,s)\mathbf{1}_{\Omega_I\cup\Omega_J}\big)\cdot\mathbf{E}\big(\psi(\xi,t)\mathbf{1}_{\Omega_K}\big)\\ &< \mathbf{E}\big(\psi(\xi,t)\mathbf{1}_{\Omega_I\cup\Omega_J}\big)\cdot\mathbf{E}\big(\psi(\xi,s)\mathbf{1}_{\Omega_K}\big) =: A(t)B(s). \end{aligned}$$

Further, we have $A(t)+B(t)=\mathbf{E}(\psi(\xi,t))\leqslant 0$ and $A(s)+B(s)=\mathbf{E}(\psi(\xi,s))\geqslant 0$, since $t\in V$, $s\in U$, and $\Omega_I\cup\Omega_J$ and Ω_K are disjoint.

To summarize, $t,s\in\Theta$ are such that $t<s$ and the following inequalities hold:

$$A(s)B(t)<B(s)A(t),\qquad A(t)+B(t)\leqslant 0,\qquad A(s)+B(s)\geqslant 0. \tag{6.4}$$

Here $B(t)>0$ because it equals the integral of a positive function over the set Ω_K of positive measure with respect to the probability $\mathbf{P}$. On the other hand, $A(s)<0$, because

$$A(s)=\mathbf{E}\big(\psi(\xi,s)\mathbf{1}_{\Omega_I\cup\Omega_J}\big)=\mathbf{E}\big(\psi(\xi,s)\mathbf{1}_{\Omega_I}\big)+\mathbf{E}\big(\psi(\xi,s)\mathbf{1}_{\Omega_J}\big)$$

and the first term is negative as the integral of a negative function over the set Ω_I (of positive measure with respect to $\mathbf{P}$), and the second term is nonpositive as the integral of a nonpositive function over the set Ω_J.

Consequently, by the last two inequalities of (6.4) we get

$$0<B(t)\leqslant -A(t)\quad\text{and}\quad 0<-A(s)\leqslant B(s),$$

yielding that

$$0<-A(s)B(t)\leqslant -A(t)B(s),$$

that is, $A(s)B(t)\geqslant B(s)A(t)$. This contradicts to the first inequality in (6.4).

Consequently, we have that $s_0\leqslant u_0:=\sup U\leqslant\inf V=:v_0\leqslant t_0$. It remains to show that $u_0=v_0$. If, to the contrary, we assume that $u_0<v_0$, then for each $r\in(u_0,v_0)$, we get $r\notin U$ and $r\notin V$, yielding that $\mathbf{E}(\psi(\xi,r))<0$ and $\mathbf{E}(\psi(\xi,r))>0$, respectively, which is a contradiction.

All in all, $u_0=v_0$ is a unique point of sign change for the function $\Theta\ni t\mapsto\mathbf{E}(\psi(\xi,t))$, as desired. □

Proof of Proposition 7. By assumption (i), for each $\omega\in\Omega$, we have that the function $\Theta\ni t\mapsto\psi(\xi(\omega),t)$ is strictly decreasing. By the monotonicity of the expectation this implies that the function $\Theta\ni t\mapsto\mathbf{E}(\psi(\xi,t))$ is decreasing, and in fact, it is strictly decreasing. Indeed, if $t_1<t_2$, $t_1,t_2\in\Theta$, are such that $\mathbf{E}(\psi(\xi,t_1))=\mathbf{E}(\psi(\xi,t_2))$, then $\mathbf{E}(\psi(\xi,t_1)-\psi(\xi,t_2))=0$, where $\mathbf{P}(\psi(\xi,t_1)-\psi(\xi,t_2)\geqslant 0)=1$. Consequently, $\mathbf{P}(\psi(\xi,t_1)-\psi(\xi,t_2)=0)=1$, leading us to a contradiction, since $\psi(\xi(\omega),t_1)>\psi(\xi(\omega),t_2)$, $\omega\in\Omega$.

Define the sets $U,V\subseteq\Theta$ by

$$U:=\big\{s\in\Theta\colon\ \mathbf{E}\big(\psi(\xi,s)\big)\geqslant 0\big\}\quad\text{and}\quad V:=\big\{t\in\Theta\colon\ \mathbf{E}\big(\psi(\xi,t)\big)\leqslant 0\big\}.$$

By assumption (iv) we have that $s_0\in U$ and $t_0\in V$. Since the function $\Theta\ni t\mapsto\mathbf{E}(\psi(\xi,t))$ is strictly decreasing, we can easily deduce that $s\leqslant t$ for all $s\in U$, $t\in V$. Indeed, if for some $s\in U$ and $t\in V$, the inequality $s>t$ were true, then we would have that $0\leqslant\mathbf{E}(\psi(\xi,s))<\mathbf{E}(\psi(\xi,t))\leqslant 0$, leading us to a contradiction. Hence we have $s_0\leqslant u_0:=\sup U\leqslant\inf V=:v_0\leqslant t_0$. It remains to show that $u_0=v_0$. If, to the contrary, we assume that $u_0<v_0$, then for each $r\in(u_0,v_0)$, we get $r\notin U$ and $r\notin V$, yielding that $\mathbf{E}(\psi(\xi,r))<0$ and $\mathbf{E}(\psi(\xi,r))>0$, respectively, which is a contradiction. Consequently, $u_0=v_0$ is a unique point of sign change for the function $\Theta\ni t\to\mathbf{E}(\psi(\xi,t))$, as desired. □

Proof of Corollary 2. To verify the statement, we have to show that for all $n \in \mathbb{N}$, $x_1, \dots, x_n \in X$, and $\boldsymbol{\lambda} = (\lambda_1, \dots, \lambda_n) \in \Lambda_n$, the function $\Theta \ni t \mapsto \sum_{i=1}^{n} \lambda_i \psi(x_i, t)$ has a unique point of sign change in Θ. Without loss of generality, we may assume that $x_1, \dots, x_n$ are pairwise distinct elements of X and $\lambda_1, \dots, \lambda_n > 0$ with $\lambda_1 + \dots + \lambda_n = 1$.

Define the probability space $(\Omega, \mathcal{A}, \mathbf{P})$ by

$$\Omega := \{x_1, \dots, x_n\}, \qquad \mathcal{A} := 2^{\Omega}, \qquad \mathbf{P}\bigl(\{x_i\}\bigr) := \lambda_i \quad i \in \{1, \dots, n\},$$

and the random variable $\xi : \Omega \to \Omega$ by $\xi(\omega) := \omega$, $\omega \in \Omega$.

Then conditions (i) and (ii) of Theorem 2 follow from our assumptions. The measurability condition (iii) of Theorem 2 is trivial due to the fact that $\mathcal{A} = 2^{\Omega}$. Since

$$\mathbf{E}\bigl(\bigl|\psi(\xi, t)\bigr|\bigr) = \sum_{i=1}^{n} \lambda_i \bigl|\psi(x_i, t)\bigr|, \quad t \in \Theta,$$

condition (iv) of Theorem 2 is obviously valid. Finally, condition (v) of Theorem 2 is satisfied by

$$s_0 := \min\bigl\{\vartheta_1(x_1), \dots, \vartheta_1(x_n)\bigr\} \quad \text{and} \quad t_0 := \max\bigl\{\vartheta_1(x_1), \dots, \vartheta_1(x_n)\bigr\}.$$

Indeed, for each $i \in \{1, \dots, n\}$, we have that $\psi(x_i, s_0) \geqslant 0 \geqslant \psi(x_i, t_0)$, since ψ is a T_1-function and $\psi(x, \vartheta_1(x)) = 0$, $x \in X$. This implies that

$$\mathbf{E}\bigl(\psi(\xi, s_0)\bigr) = \sum_{i=1}^{n} \lambda_i \psi(x_i, s_0) \geqslant 0 \geqslant \sum_{i=1}^{n} \lambda_i \psi(x_i, t_0) = \mathbf{E}\bigl(\psi(\xi, t_0)\bigr).$$

Therefore, according to Theorem 2, the mapping $\Theta \ni t \mapsto \mathbf{E}(\psi(\xi, t)) = \sum_{i=1}^{n} \lambda_i \psi(x_i, t)$ has a unique point of sign change in Θ, as desired. □

Proof of Lemma 5. Let us define the probability measure $\mathbb{Q}$ on the measurable space $(\Omega, \mathcal{A})$ by

$$\mathbb{Q}(A) := \int\limits_A \frac{p(\xi)}{\mathbf{E}(p(\xi))} \, \mathrm{d}\mathbf{P} = \frac{\mathbf{E}(p(\xi)\mathbf{1}_A)}{\mathbf{E}(p(\xi))}, \quad A \in \mathcal{A}.$$

By denoting the expectation with respect to $\mathbb{Q}$ by $\mathbf{E}_{\mathbb{Q}}$ we have

$$\mathbf{E}_{\mathbb{Q}}\bigl(\bigl|\varphi(\xi)\bigr|\bigr) = \frac{\mathbf{E}(p(\xi)|\varphi(\xi)|)}{\mathbf{E}(p(\xi))},$$

and hence, by the assumptions, $\varphi(\xi)$ is integrable with respect to $\mathbb{Q}$, and we also get

$$\mathbf{E}_{\mathbb{Q}}\bigl(\varphi(\xi)\bigr) = \frac{\mathbf{E}(p(\xi)\varphi(\xi))}{\mathbf{E}(p(\xi))}.$$

Applying Lemma 1 in Janković and Merkle [13] (for integrable one-dimensional random variables), we have that $\mathbf{E}_{\mathbb{Q}}(\varphi(\xi)) \in \operatorname{conv}(\varphi(\xi(\Omega))) \subseteq \operatorname{conv}(\varphi(X))$, yielding the statement. □

Proof of Proposition 8. We apply Proposition 7. Assumptions (i), (ii), and (iii) of Proposition 7 readily hold.

To verify assumption (iv) of Proposition 7, we first show that for any $y \in J := \operatorname{conv}(f(\Theta))$, there exist $s_0, t_0 \in \Theta$ such that $f(s_0) \leqslant y \leqslant f(t_0)$. By the Carathéodory theorem on convex hulls there exist at most two elements $y_1, y_2 \in f(\Theta) \subseteq J$ with $y_1 \leqslant y_2$ such that y can be represented as a convex combination of y_1 and y_2. This also yields that $y_1 \leqslant y \leqslant y_2$, and therefore there exist $s_0, t_0 \in \Theta$ such that $f(s_0) \leqslant y \leqslant f(t_0)$. Now observe that

$$f^{(-1)}\left(\frac{\mathbf{E}(p(\xi)\varphi(\xi))}{\mathbf{E}(p(\xi))}\right)$$

is well defined, since by Lemma 5 we get that

$$\frac{\mathbf{E}(p(\xi)\varphi(\xi))}{\mathbf{E}(p(\xi))} \in \operatorname{conv}\big(\varphi(X)\big) \subseteq \operatorname{conv}\big(f(\Theta)\big),$$

and by Lemma 4, $f^{(-1)}$ is defined on $\operatorname{conv}(f(\Theta))$. Next, for $y := \mathbf{E}(p(\xi)\varphi(\xi))/\mathbf{E}(p(\xi))$, let us choose $s_0, t_0 \in \Theta$ as it was described above. Then

$$\mathbf{E}\big(\psi(\xi, s_0)\big) = \mathbf{E}\big(p(\xi)\varphi(\xi)\big) - f(s_0)\,\mathbf{E}\big(p(\xi)\big) = \mathbf{E}\big(p(\xi)\big)\big(y - f(s_0)\big) \geqslant 0,$$

and

$$\mathbf{E}\big(\psi(\xi, t_0)\big) = \mathbf{E}\big(p(\xi)\varphi(\xi)\big) - f(t_0)\,\mathbf{E}\big(p(\xi)\big) = \mathbf{E}\big(p(\xi)\big)\big(y - f(t_0)\big) \leqslant 0.$$

Therefore assumption (iv) of Proposition 7 holds as well, and, according to Proposition 7, we get that the function $\Theta \ni t \to \mathbf{E}(\psi(\xi, t))$ admits a unique point of sign change in Θ. It remains to check that this unique point of sign change takes the form given in the proposition.

If, for some $t \in \Theta$, we have $t < f^{(-1)}(y)$, then $\mathbf{E}(\psi(\xi, t)) > 0$, since otherwise $\mathbf{E}(\psi(\xi, t)) \leqslant 0$ would yield that $\mathbf{E}(p(\xi)\varphi(\xi)) \leqslant f(t)\,\mathbf{E}(p(\xi))$, that is, $y \leqslant f(t)$. Then by Lemma 4 we would get $f^{(-1)}(y) \leqslant f^{(-1)}(f(t)) = t$, leading us to a contradiction.

If for some $t \in \Theta$, we have $t > f^{(-1)}(y)$, then we can similarly argue to obtain that $\mathbf{E}(\psi(\xi, t)) < 0$.

Consequently, the unique point of sign change in question is $f^{(-1)}(y)$, as desired. □

Proof of Proposition 9. First, we give a direct proof. Denote the limit $\lim_{z\to\infty} f(z)$ by $f_\infty \in (0, \infty)$. In view of the increasingness of f, it follows that $0 \leqslant f(z) \leqslant f_\infty$ for all $z \in \mathbb{R}_+$. Therefore $|f(z)| \leqslant f_\infty$ for all $z \in \mathbb{R}$, which implies that $|\psi(x, t)| \leqslant f_\infty$ for all $x, t \in \mathbb{R}$. Hence, for any random variable ξ and any $t \in \mathbb{R}$, we have that $\mathbf{E}(|\psi(\xi, t)|) < \infty$.

Since for each $x \in \mathbb{R}$, the function $\mathbb{R} \ni t \mapsto \psi(x, t)$ is strictly decreasing, we have that the function $\mathbb{R} \ni t \mapsto \mathbf{E}(\psi(\xi, t))$ is strictly decreasing. Indeed, if $s < t$, $s, t \in \mathbb{R}$, then we have $\psi(\xi(\omega), s) > \psi(\xi(\omega), t)$, $\omega \in \Omega$, yielding that $\mathbf{E}(\psi(\xi, s)) \geqslant \mathbf{E}(\psi(\xi, t))$. Here the equality cannot hold, since otherwise $\mathbf{E}(\psi(\xi, s) - \psi(\xi, t)) = 0$ would be valid, yielding that $\mathbf{P}(\psi(\xi, s) - \psi(\xi, t) = 0) = 1$. This leads us to a contradiction, since $\psi(\xi(\omega), s) - \psi(\xi(\omega), t) > 0$, $\omega \in \Omega$. Since $\lim_{t\to\pm\infty} \psi(\xi(\omega), t) = \mp f_\infty$, $\omega \in \Omega$, and $|\psi(\xi(\omega), t)| \leqslant f_\infty$, $\omega \in \Omega$, $t \in \mathbb{R}$, the dominated convergence theorem implies that

$$\lim_{t\to\infty} \mathbf{E}\big(\psi(\xi, t)\big) = \mathbf{E}\Big(\lim_{t\to\infty} \psi(\xi, t)\Big) = \mathbf{E}(-f_\infty) = -f_\infty < 0 \tag{6.5}$$

and

$$\lim_{t\to-\infty} \mathbf{E}\big(\psi(\xi, t)\big) = \mathbf{E}\Big(\lim_{t\to-\infty} \psi(\xi, t)\Big) = \mathbf{E}(f_\infty) = f_\infty > 0. \tag{6.6}$$

Since f is continuous and $f(0) = 0$, we have that ψ is continuous in its second variable. Thus by the dominated convergence theorem it follows that the function $\mathbb{R} \ni t \mapsto \mathbf{E}(\psi(\xi, t))$ is also continuous. All in all, the function $\mathbb{R} \ni t \mapsto \mathbf{E}(\psi(\xi, t))$ is strictly decreasing, continuous, and changing sign, and hence there exists a unique $t_0 \in \mathbb{R}$ such that $\mathbf{E}(\psi(\xi, t_0)) = 0$, as desired.

Finally, we present an alternative proof of Proposition 9 using Theorem 2. We check that the assumptions of Theorem 2 hold. Since $f(0) = 0$ and f is strictly increasing, assumption (i) of Theorem 2 holds with

$\vartheta_1(x) = x$, $x \in \mathbb{R}$. Using that f is strictly increasing, by part (d) of Proposition 12 we have that ψ is a $T_2^{\boldsymbol{\lambda}}$-function for all $\boldsymbol{\lambda} \in \Lambda_2$. Consequently, part (v) of Theorem 1 yields that for all $x, y \in \mathbb{R}$ with $\vartheta_1(x) < \vartheta_1(y)$, the function (2.2) is strictly increasing, that is, assumption (ii) of Theorem 2 holds. Assumption (iii) of Theorem 2 readily holds. The first part of the direct proof of the present proposition implies that assumption (iv) of Theorem 2 holds. Using (6.5) and (6.6), we have that assumption (v) of Theorem 2 holds as well. All in all, we can apply Theorem 2, which yields that the function $\mathbb{R} \ni t \mapsto \mathbf{E}(\psi(\xi, t)) = \mathbf{E}(\operatorname{sign}(\xi - t) f(|\xi - t|))$ has a (unique) point of sign change. Since the function $\mathbb{R} \ni t \mapsto \mathbf{E}(\psi(\xi, t))$ is strictly decreasing and continuous (see the direct proof), we have that the equation $\mathbf{E}(\psi(\xi, t)) = 0$ has a unique solution with respect to $t \in \mathbb{R}$, as desired. □

7 Proofs for Section 4

Proof of Proposition 10. Let $n \geqslant 2$ and $\psi \in \Psi[T_n](\mathbb{R}, \mathbb{R})$. Then, according to part (iii) of Theorem 1, for all $x, y \in \mathbb{R}$ with $x < y$, the number $(n-k)/k$ must be a level of increase for the function (2.2) if $k \in \{1, \ldots, n-1\}$. Note that for all $x, y \in \mathbb{R}$ with $x < y$, the function (2.2) takes the form

$$(x, y) \ni t \mapsto -\frac{\psi(x, t)}{\psi(y, t)} = \frac{1 - \alpha}{\alpha} > 0.$$

Therefore we have that

$$\frac{1 - \alpha}{\alpha} \neq \frac{n - k}{k} \quad \text{for each} k \in \{1, \ldots, n-1\},$$

which implies that $\alpha \neq k/n$ for each $k \in \{1, \ldots, n-1\}$.

Conversely, if $\alpha \notin \{1/n, \ldots, (n-1)/n\}$, where $n \geqslant 2$, then we have that

$$\frac{k}{n} < \alpha < \frac{k+1}{n} \tag{7.1}$$

for some $k \in \{0, \ldots, n-1\}$, and hence $k < n\alpha < k+1$. Let $x_1, \ldots, x_n \in \mathbb{R}$ be arbitrary. If $t \in \mathbb{R}$ with $t < x_{k+1}^*$, then we have that $\psi(x_i^*, t) \geqslant \alpha - 1$, $i = 1, \ldots, k$, and $\psi(x_i^*, t) = \alpha$, $i = k+1, \ldots, n$, yielding that

$$\begin{aligned} \sum_{i=1}^{n} \psi(x_i, t) &= \sum_{i=1}^{k} \psi(x_i^*, t) + \sum_{i=k+1}^{n} \psi(x_i^*, t) \geqslant k(\alpha - 1) + (n-k)\alpha \\ &= n\alpha - k > 0. \end{aligned}$$

If $t \in \mathbb{R}$ with $t > x_{k+1}^*$, then we have that $\psi(x_i^*, t) = \alpha - 1$, $i = 1, \ldots, k+1$, and $\psi(x_i^*, t) \leqslant \alpha$, $i = k+2, \ldots, n$, yielding that

$$\begin{aligned} \sum_{i=1}^{n} \psi(x_i, t) &= \sum_{i=1}^{k+1} \psi(x_i^*, t) + \sum_{i=k+2}^{n} \psi(x_i^*, t) \leqslant (k+1)(\alpha - 1) + (n-k-1)\alpha \\ &= n\alpha - k - 1 < 0. \end{aligned}$$

Therefore $\vartheta_n(x_1, \ldots, x_n)$ exists and equals x_{k+1}^*. This proves that ψ is indeed a T_n-function. Furthermore, using (7.1), we have $k < \alpha n < k+1$ and $k+1 < \alpha n + 1 < k+2$, yielding that $\lceil n\alpha \rceil = \lfloor n\alpha + 1 \rfloor = k+1$. Hence $x_{k+1}^* = x_{\lceil n\alpha \rceil}^* = x_{\lfloor n\alpha + 1 \rfloor}^*$, as desired. □

Proof of Proposition 11. Let us apply Proposition 7 with $X := \mathbb{R}$, $\mathcal{X} := \mathcal{B}(\mathbb{R})$, and $\Theta := \mathbb{R}$. Assumptions (i) and (ii) of Proposition 7 readily hold.

The validity of assumption (iii) of Proposition 7 can be seen from

$$\begin{aligned}\mathbf{E}\big(|\psi(\xi,t)|\big) &= \mathbf{E}\big(\alpha|\xi-t|\mathbf{1}_{\{\xi>t\}} + (1-\alpha)|\xi-t|\mathbf{1}_{\{\xi<t\}}\big)\\ &\leqslant \alpha\,\mathbf{E}\big(|\xi-t|\big) + (1-\alpha)\,\mathbf{E}\big(|\xi-t|\big)\\ &\leqslant \mathbf{E}\big(|\xi|\big) + |t| < \infty, \quad t\in\mathbb{R}.\end{aligned}\tag{7.2}$$

Finally, we verify assumption (iv) of Proposition 7. First, note that for all $t\in\mathbb{R}$, we have

$$\begin{aligned}\mathbf{E}\big(\psi(\xi,t)\big) &= \alpha\,\mathbf{E}\big((\xi-t)^+\big) - (1-\alpha)\,\mathbf{E}\big((\xi-t)^-\big)\\ &= \alpha\,\mathbf{E}(\xi-t) + (2\alpha-1)\,\mathbf{E}\big((\xi-t)^-\big)\end{aligned}\tag{7.3}$$

and, analogously,

$$\mathbf{E}\big(\psi(\xi,t)\big) = (2\alpha-1)\,\mathbf{E}\big((\xi-t)^+\big) + (1-\alpha)\,\mathbf{E}(\xi-t).\tag{7.4}$$

In case of $\alpha=1/2$, we have $\mathbf{E}(\psi(\xi,t)) = (\mathbf{E}(\xi)-t)/2$, which is positive if $t<\mathbf{E}(\xi)$ and is negative if $t>\mathbf{E}(\xi)$. This shows that assumption (iv) of Proposition 7 holds in case of $\alpha=1/2$.

In case of $\alpha\in(1/2,1)$, we have $2\alpha-1>0$, and hence (7.3) yields that $\mathbf{E}(\psi(\xi,t))\geqslant\alpha(\mathbf{E}(\xi)-t)$, which is positive if $t<\mathbf{E}(\xi)$. Further, for all $t>0$, we have

$$\begin{aligned}\mathbf{E}\big(\psi(\xi,t)\big) &\leqslant \alpha\big(\mathbf{E}(\xi)-t\big) + (2\alpha-1)\big(\mathbf{E}\big(|\xi|\big)+t\big)\\ &= \alpha\,\mathbf{E}(\xi) + (2\alpha-1)\,\mathbf{E}(|\xi|) - (1-\alpha)t,\end{aligned}$$

which is negative if

$$t > \max\left(0, \frac{\alpha\,\mathbf{E}(\xi) + (2\alpha-1)\,\mathbf{E}(|\xi|)}{1-\alpha}\right).$$

This shows that assumption (iv) of Proposition 7 holds in case of $\alpha\in(1/2,1)$.

In case of $\alpha\in(0,1/2)$, we have $2\alpha-1<0$, and hence (7.4) yields that $\mathbf{E}(\psi(\xi,t))\leqslant(1-\alpha)(\mathbf{E}(\xi)-t)$, which is negative if $t>\mathbf{E}(\xi)$. Further, for all $t<0$, we have

$$\begin{aligned}\mathbf{E}\big(\psi(\xi,t)\big) &\geqslant (2\alpha-1)\big(\mathbf{E}(|\xi|)+|t|\big) + (1-\alpha)\big(\mathbf{E}(\xi)-t\big)\\ &= (2\alpha-1)\,\mathbf{E}\big(|\xi|\big) + (1-\alpha)\,\mathbf{E}(\xi) - \alpha t,\end{aligned}$$

which is positive if

$$t < \min\left(0, \frac{(2\alpha-1)\,\mathbf{E}(|\xi|) + (1-\alpha)\,\mathbf{E}(\xi)}{\alpha}\right).$$

This shows that assumption (iv) of Proposition 7 holds in case of $\alpha\in(0,1/2)$.

Therefore assumption (iv) of Proposition 7 holds as well, and according to Proposition 7, we get that the function $\mathbb{R}\ni t\to\mathbf{E}(\psi(\xi,t))$ admits a unique point of sign change.

Using the dominated convergence theorem, we check that the function $\mathbb{R}\ni t\mapsto\mathbf{E}(\psi(\xi,t))$ is continuous. Let $(t_n)_{n\in\mathbb{N}}$ be a real sequence such that $t_n\to t_0$ as $n\to\infty$, where $t_0\in\mathbb{R}$. Then using that ψ is strictly decreasing in its second variable, we have that

$$\psi\Big(\xi, \sup_{m\in\mathbb{N}} t_m\Big) \leqslant \psi(\xi,t_n) \leqslant \psi\Big(\xi, \inf_{m\in\mathbb{N}} t_m\Big), \quad n\in\mathbb{N},$$

yielding that

$$\mathbf{E}\Big(\psi\Big(\xi, \sup_{m\in\mathbb{N}} t_m\Big)\Big) \leqslant \mathbf{E}(\psi(\xi,t_n)) \leqslant \mathbf{E}\Big(\psi\Big(\xi, \inf_{m\in\mathbb{N}} t_m\Big)\Big), \quad n\in\mathbb{N}.$$

Hence by (7.2) we get that

$$\mathbf{E}\big(|\psi(\xi,t_n)|\big) \leqslant \mathbf{E}\Big(\Big|\psi\Big(\xi,\inf_{m\in\mathbb{N}} t_m\Big)\Big|\Big) + \mathbf{E}\Big(\Big|\psi\Big(\xi,\sup_{m\in\mathbb{N}} t_m\Big)\Big|\Big) < \infty, \quad n \in \mathbb{N}.$$

Further, since ψ is continuous in its second variable, we have $\psi(\xi,t_n) \to \psi(\xi,t_0)$ as $n \to \infty$. Hence the dominated convergence theorem implies that $\mathbf{E}(\psi(\xi,t_n)) \to \mathbf{E}(\psi(\xi,t_0))$ as $n \to \infty$, as desired.

Consequently, the unique point of sign change of the function $\mathbb{R} \ni t \to \mathbf{E}(\psi(\xi,t))$ is nothing else but the unique solution of the equation $\mathbf{E}(\psi(\xi,t)) = 0$, $t \in \mathbb{R}$. □

Proof of Proposition 12. (a) It follows from the facts that for each $x \in \mathbb{R}$, we have $\psi(x,x) = 0$; $\psi(x,t) > 0$ for each $t < x$ if and only if $f(z) > 0$ for each $z > 0$; and $\psi(x,t) < 0$ for each $t > x$ if and only if $f(z) > 0$ for each $z > 0$.

(b) Let us suppose that $\psi \in \Psi[T_n](\mathbb{R},\mathbb{R})$ for infinitely many $n \in \mathbb{N}$. Then by Proposition 4 we have $\psi \in \Psi[T_1](\mathbb{R},\mathbb{R})$. By part (iv) of Theorem 1, for all $x,y \in \mathbb{R}$ with $\vartheta_1(x) < \vartheta_1(y)$, the function (2.2) is increasing. Since $\psi \in \Psi[T_1](\mathbb{R},\mathbb{R})$, by part (a) of the present proposition, we have $f(z) > 0$ for each $z > 0$ and $\vartheta_1(x) = x$. Consequently, for all $x < y$, $x,y \in \mathbb{R}$, the function (given by (2.2))

$$(x,y) \ni t \mapsto -\frac{\psi(x,t)}{\psi(y,t)} \tag{7.5}$$

is increasing. Hence the statement of part (b) follows by the following observation (which we check below): provided that $f(z) > 0$ for each $z > 0$, the function (7.5) is (strictly) increasing for all $x < y$, $x,y \in \mathbb{R}$ if and only if f is (strictly) increasing. Since the function ψ given in (4.9) depends only on $x - t$, it suffices to check that

$$\text{the function } (0,z) \ni t \mapsto \frac{\psi(0,t)}{\psi(z,t)} = -\frac{f(t)}{f(z-t)} \quad \text{is (strictly) decreasing for each } z > 0 \tag{7.6}$$

if and only if f is (strictly) increasing. Indeed, for all $x < y$, $x,y \in \mathbb{R}$, and $t \in (x,y)$, we have

$$\frac{\psi(x,t)}{\psi(y,t)} = -\frac{f(t-x)}{f(y-t)} = -\frac{f(t-x)}{f(y-x-(t-x))} = \frac{\psi(0,t-x)}{\psi(y-x,t-x)}, \quad t-x \in (0,\, y-x).$$

Thus property (7.6) holds if and only if

$$\frac{f(s)}{f(z-s)}\,(<)\!\leqslant \frac{f(t)}{f(z-t)} \quad \text{for all } s,t,z \in \mathbb{R} \text{ with } 0 < s < t < z,$$

which is equivalent to

$$f(s)f(z-t)\,(<)\!\leqslant f(t)f(z-s) \quad \text{for all } s,t,z \in \mathbb{R} \text{ with } 0 < s < t < z. \tag{7.7}$$

Using the nonnegativity of f, it yields that (7.6) holds if and only if

$$f(s)\,(<) \leqslant f(t) \qquad \text{for each } s,t \in \mathbb{R} \text{ with } 0 < s < t, \tag{7.8}$$

that is, f is (strictly) increasing on $\mathbb{R}_+$. Indeed, if (7.6) holds, then (7.7) holds as well, and by choosing $z = s + t$ we get that $f(s)^2(<) \leqslant f(t)^2$, which implies (7.8), since f is nonnegative. If (7.8) holds, then for all $s,t,z \in \mathbb{R}$ with $0 < s < t < z$, we have $0 \leqslant f(s)\,(<)\!\leqslant f(t)$ and $0 \leqslant f(z-t)\,(<)\!\leqslant f(z-s)$, implying (7.7), and hence (7.6) as well.

(c) and (d) Let us suppose that $\psi \in \Psi[T_2^{\boldsymbol{\lambda}}](\mathbb{R},\mathbb{R})$ for all $\boldsymbol{\lambda} \in \Lambda_2$. In particular, $\psi \in \Psi[T_1](\mathbb{R},\mathbb{R})$, and hence, by part (a) of the present proposition we have $f(z) > 0$ for each $z > 0$ and $\vartheta_1(x) = x$, $x \in \mathbb{R}$. Consequently, by part (v) of Theorem 1, for all $x < y$, $x, y \in \mathbb{R}$, the function (7.5) is strictly increasing. By the proof of part (b) it implies that f is strictly increasing, as desired. We give an alternative proof as well. Let $s, t \in \mathbb{R}_+$ with $s < t$, $x_1 \in \mathbb{R}$, $x_2 := x_1 + s + t$, and $r := x_1 + s$. Then $(x_1 + x_2)/2 = x_1 + ((s + t)/2) > x_1 + s = r$, yielding that

$$\psi\left(x_1, \frac{x_1 + x_2}{2}\right) + \psi\left(x_2, \frac{x_1 + x_2}{2}\right) = -f\left(\frac{s+t}{2}\right) + f\left(\frac{s+t}{2}\right) = 0.$$

Hence $\vartheta_{2,\psi}(x_1, x_2) = (x_1 + x_2)/2$, and since $r < (x_1 + x_2)/2$, we have that $\psi(x_1, r) + \psi(x_2, r) > 0$. Since $\psi(x_1, r) + \psi(x_2, r) = -f(s) + f(t)$, we get that $f(s) < f(t)$, as desired. Conversely, let us suppose that f is strictly increasing. Then for each $x \in \mathbb{R}$, the function $\mathbb{R} \ni t \mapsto \psi(x, t)$ is strictly decreasing. Hence Proposition 2 implies that $\psi \in \Psi[T_n^{\boldsymbol{\lambda}}](\mathbb{R},\mathbb{R})$ for all $n \in \mathbb{N}$ and $\boldsymbol{\lambda} \in \Lambda_n$ (in particular, $\psi \in \Psi[T_2^{\boldsymbol{\lambda}}](\mathbb{R},\mathbb{R})$-function for each $\boldsymbol{\lambda} \in \Lambda_n$), as desired. $\square$

Acknowledgment. We acknowledge the valuable suggestions from the two referees.

Funding. Open Access funding provided by University of Debrecen.

References

1. A. Ali and R.J. Tibshirani, The generalized lasso problem and uniqueness, *Electron. J. Stat.*, **13**(2):2307–2347, 2019, `https://doi.org/10.1214/19-EJS1569`.
2. M. Barczy and Zs. Páles, Existence and uniqueness of weighted generalized ψ-estimators, 2022, `https://doi.org/10.48550/arXiv.2211.06026`.
3. F. Bellini, B. Klar, and A. Müller, Expectiles, Omega ratios and stochastic ordering, *Methodol. Comput. Appl. Probab.*, **20**(3):855–873, 2018, `https://doi.org/10.1007/s11009-016-9527-2`.
4. O. Catoni, Challenging the empirical mean and empirical variance: a deviation study, *Ann. Inst. Henri Poincaré, Probab. Stat.*, **48**(4):1148–1185, 2012, `https://doi.org/10.1214/11-AIHP454`.
5. P. Chen, X. Jin, X. Li, and L. Xu, A generalized Catoni's M-estimator under finite α-th moment assumption with $\alpha \in (1, 2)$, *Electron. J. Stat.*, **15**(2):5523–5544, 2021, `https://doi.org/10.1214/21-ejs1911`.
6. B.R. Clarke, Uniqueness and Fréchet differentiability of functional solutions to maximum likelihood type equations, *Ann. Stat.*, **11**(4):1196–1205, 1983, `https://doi.org/10.1214/aos/1176346332`.
7. T. Dimitriadis, T. Fissler, and J. Ziegel, Characterizing M-estimators, *Biometrika*, **111**(1):339–346, 2024, `https://doi.org/10.1093/biomet/asad026`.
8. L. Gasiński and N.S. Papageorgiou, *Exercises in Analysis. Part 2: Nonlinear Analysis*, Probl. Books Math., Springer, Cham, 2016, `https://doi.org/10.1007/978-3-319-27817-9`.

9. R. Grünwald and Zs. Páles, On the equality problem of generalized Bajraktarević means, *Aequationes Math.*, **94**(4):651–677, 2020, `https://doi.org/10.1007/s00010-019-00670-9`.
10. F. Hampel, C. Hennig, and E. Ronchetti, A smoothing principle for the Huber and other location M-estimators, *Comput. Stat. Data Anal.*, **55**(1):324–337, 2011, `https://doi.org/10.1016/j.csda.2010.05.001`.
11. P.J. Huber, Robust estimation of a location parameter, *Ann. Math. Stat.*, **35**:73–101, 1964, `https://doi.org/10.1214/aoms/1177703732`.
12. P.J. Huber, The behavior of maximum likelihood estimates under nonstandard conditions, in *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability (June 21–July 18, 1965 and December 27, 1965–January 7, 1966), Vol. I: Statistics*, Univ. California Press, Berkeley, CA, 1967, pp. 221–233.
13. S. Janković and M. Merkle, A mean value theorem for systems of integrals, *J. Math. Anal. Appl.*, **342**(1):334–339, 2008, `https://doi.org/10.1016/j.jmaa.2007.12.012`.
14. R. Koenker and G. Bassett, Jr., Regression quantiles, *Econometrica*, **46**(1):33–50, 1978, `https://doi.org/10.2307/1913643`.
15. M.R. Kosorok, *Introduction to Empirical Processes and Semiparametric Inference*, Springer Ser. Stat., Springer, New York, 2008, `https://doi.org/10.1007/978-0-387-74978-5`.
16. V. Krätschmer and H. Zähle, Statistical inference for expectile-based risk measures, *Scand. J. Stat.*, **44**(2):425–454, 2017, `https://doi.org/10.1111/sjos.12259`.
17. K. Lange, *Numerical Analysis for Statisticians,* 2nd ed., Stat. Comput., Springer, New York, 2010, `https://doi.org/10.1007/978-1-4419-5945-4`.
18. T. Mathieu, Concentration study of M-estimators using the influence function, *Electron. J. Stat.*, **16**(1):3695–3750, 2022, `https://doi.org/10.1214/22-ejs2030`.
19. W.K. Newey and J.L. Powell, Asymmetric least squares estimation and testing, *Econometrica*, **55**(4):819–847, 1987, `https://doi.org/10.2307/1911031`.
20. Zs. Páles, On approximately convex functions, *Proc. Am. Math. Soc.*, **131**(1):243–252, 2003, `https://doi.org/10.1090/S0002-9939-02-06552-8`.
21. R. Passeggeri and N. Reid, A universal robustification procedure, 2022, `https://doi.org/10.48550/arXiv.2206.06998`.
22. W.J.J. Rey, *Introduction to Robust and Quasirobust Statistical Methods*, Universitext, Springer, Berlin, 1983, `https://doi.org/10.1007/978-3-642-69389-2`.
23. A. Shapiro, D. Dentcheva, and A. Ruszczyński, *Lectures on Stochastic Programming*, MPS/SIAM Ser. Optim., Vol. 9, SIAM, MPS, Philadelphia, 2009, `https://doi.org/10.1137/1.9780898718751`.
24. R. J. Tibshirani, The lasso problem and uniqueness, *Electron. J. Stat.*, **7**:1456–1490, 2013, `https://doi.org/10.1214/13-EJS815`.
25. A.W. van der Vaart, *Asymptotic Statistics, Vol. 3*, Camb. Ser. Stat. Probab. Math., Cambridge Univ. Press, Cambridge, 1998, `https://doi.org/10.1017/CBO9780511802256`.