POLITECNICO DI TORINO
Repository ISTITUZIONALE

Robust parameter estimation based on the generalized log-likelihood in the context of Sharma-Taneja-Mittal measure

Original
Robust parameter estimation based on the generalized log-likelihood in the context of Sharma-Taneja-Mittal measure / Da Silva, S.L.E.F., Kaniadakis, G.. - In: PHYSICAL REVIEW. E. - ISSN 2470-0045. - ELETTRONICO. - 104:2(2021), p. 024107. [10.1103/PhysRevE.104.024107]

Availability:
This version is available at: 11583/2957167 since: 2022-03-02T20:51:47Z

Publisher:
American Physical Society

Published
DOI:10.1103/PhysRevE.104.024107

Terms of use:

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

Publisher copyright

(Article begins on next page)

26 November 2026

Robust parameter estimation based on the generalized log-likelihood in the context of Sharma-Taneja-Mittal measure

Sérgio Luiz E. F. da Silva [*]
Seismic Inversion and Imaging Group, Federal Fluminense University, 24210-346 Niterói, RJ, Brazil

G. Kaniadakis [†]
Department of Applied Science and Technology, Politecnico di Torino, 10129 Torino, Italy

(Received 17 June 2021; accepted 21 July 2021; published 6 August 2021)

The problem of obtaining physical parameters that cannot be directly measured from observed data arises in several scientific fields. In the classic approach, the well-known maximum likelihood estimation associated with a Gaussian distribution is employed to obtain the model parameters of a complex system. Although this approach is quite popular in statistical physics, only a handful of spurious observations (outliers) make this approach ineffective, violating the Gauss-Markov theorem. In this work, starting from the generalized logarithmic function associated to the Sharma-Taneja-Mittal (STM) information measure, we propose an outlier-resistant approach based on the generalized log-likelihood estimation. In particular, our proposal deforms the Gaussian distribution based on a two-parameter generalization of the ordinary logarithmic function. We have tested the effectiveness of our proposal considering a classic geophysical inverse problem with a very noisy data set. The results show that the task of obtaining physical parameters based on the STM measure from noisy data with several outliers outperforms the classic approach, and therefore, our proposal is a useful tool for statistical physics, information theory, and statistical inference problems.

DOI: 10.1103/PhysRevE.104.024107

I. INTRODUCTION

The task of estimating physical parameters that cannot be directly measured from empirical data arises in a wide variety of scientific problems [1–3]. For example, in electrical impedance tomography problems [4–6], the properties of the electrical conductivity (model parameters) of a conductive object (physical system) are obtained from electrical measurements (observed data) taken at different locations of that object. In this regard, an optimization problem is employed to finding the best physical model, m, that match the observed data, d^{obs}, to the modeled data, $d^{\text{mod}}(m) = G(m)$, in which G denotes the so-called forward operator [7]. In the electrical impedance tomography case, the forward operation is given by the solutions of Maxwell's equations. In this way, the electrical conductivity is obtained by minimizing the residual data (or error): $d^{\text{obs}} - d^{\text{mod}}$.

Usually, the optimization process used in the estimation of physical parameters (also called data inversion) is carried out in the least-squares sense, in which the objective function is given by the sum of the squares of the errors. From a probabilistic viewpoint, the least-squares method assumes that the errors are independent and identically distributed by a Gaussian probability distribution, which is closely linked to the maximization of the classic Boltzmann-Gibbs-Shannon (BGS) entropy (see, for instance, Sec. 2 of Ref. [8]). Although Gaussian statistics are very useful in solving several data inversion problems [9–13], the maximum likelihood estimator (MLE) associated with the classic approach is an unbiased estimator if the errors are non-Gaussian, violating the Gauss-Markov theorem [14,15].

In fact, a handful of spurious measures (outliers) is enough for the classic approach to fail [16,17]. Thus, the data inversion based on non-Gaussian criteria has been an alternative for robust inference of physical parameters. Among them we can mention the data inversion based on the Cauchy error criterion [18,19], Student's t distribution [20,21], hybrid criteria [22–24], and more recently, in the sense of Rényi [25], Tsallis [26–31], and Kaniadakis [32–35] statistics.

It is worth mentioning that all aforementioned approaches consider the method of maximum likelihood as a starting point. The maximum likelihood estimation is a powerful method to obtain model parameters from the measured data that is very useful in statistical physics, information theory, and statistical inference. However, if the assumed statistical model is not adequate for the distribution of the errors, then the classic approach is doomed to fail, which makes it difficult to perform automated tasks. In this regard, a long (and sometimes tedious) data processing is carried out before the data inversion to remove possible outliers, so that soon afterward the classic approach is employed. In fact, after heavy data processing, the distribution of the errors can be described by a Gaussian distribution obeying the well-known central-limit theorem [36]. Moreover, in addition to the time spent on data processing, it is worth mentioning that this process is not perfect.

[*] sergioluizsilva@id.uff.br
[†] giorgio.kaniadakis@polito.it

 ©2021 American Physical Society

In this work, in order to avoid a heavy data preprocessing and enable automatized approaches, we formulate the task of estimating physical parameters by generalizing the maximum likelihood method. In particular, we consider the generalized log-likelihood function associated with the Sharma-Taneja-Mittal (STM) information measure [37,38] and based on the κ, r-deformed product [39], which we call of maximum $L_{\kappa,w}$-likelihood estimation (M$L_{\kappa,w}$E) method after introducing a parameter $w = r/\kappa$. The M$L_{\kappa,w}$E is controlled by the entropic indexes κ and w that are related with the STM entropy. In this paper, we demonstrate that even if the observed data have outliers and assuming that the errors are Gaussian, the M$L_{\kappa,w}$E is robust for non-Gaussian errors if the entropic indexes are properly chosen.

This paper is organized as follows. In Sec. II, we present a brief review of the mathematical basis of STM information measure. In Sec. III, we present the theoretical foundations of the parameter estimation based in its classic framework. Then, in Sec. IV, we introduce the probabilistic maximum log-likelihood method in the sense of the STM information measure to derive an outlier-resistent objective function in order to perform robust inference on physical parameters of a complex system. In Sec. V, we demonstrate the robustness property of our proposal with several numerical experiments by employing a geophysical data-inverse problem. To conclude, we devote the Sec. VI to the final remarks on our proposal and its implications for statistical physics and information theory.

II. STM INFORMATION MEASURE

In Refs. [37,38], Mittal, Sharma, and Teneja have been introduced the κ, r-generalized entropy which can be defined in the following form [40]:

$$\mathcal{S}_{\kappa,r}[p] = -\frac{1}{2\kappa}\int_{\Omega}[p^{1+r+\kappa}(x) - p^{1+r-\kappa}(x)]dx, \tag{1}$$

where $x \in \Omega$ is a random variable with probability function $p(x)$, κ, and r are the entropic indexes.

The entropy functional in Eq. (1) is commonly referred to as Sharma-Taneja-Mittal entropy (or κ, r entropy or STM entropy), and it may be written also in terms of the so-called κ, w-deformed logarithmic function (or κ, w logarithm), $\ln_{\kappa,w}$, after defining the parameter $w = \frac{r}{\kappa}$:

$$\mathcal{S}_{\kappa,w}[p] = -\int_{\Omega} p(x)\ln_{\kappa,w}[p(x)]dx, \tag{2}$$

in which

$$\ln_{\kappa,w}(x) = \frac{x^{\kappa(w+1)} - x^{\kappa(w-1)}}{2\kappa}. \tag{3}$$

We remark that in the limit $(\kappa, w) \to (0, 0)$, the κ, w-deformed function reduce to the ordinary logarithmic:

$$\lim_{(\kappa,w)\to(0,0)} \ln_{\kappa,w}(x) = \ln_{0,0}(x) := \ln(x), \tag{4}$$

and, consequently, note that the STM entropy [Eq. (1) or (2)] reduces to the classic BGS entropy at the same limit:

$$\begin{aligned}\lim_{(\kappa,w)\to(0,0)} \mathcal{S}_{\kappa,w}[p] &= \mathcal{S}_{0,0}[p] = -\int_{\Omega} p(x)\ln_{0,0}[p(x)]dx \\ \lim_{(\kappa,w)\to(0,0)} \mathcal{S}_{\kappa,w}[p] &= -\int_{\Omega} p(x)\ln[p(x)]dx := \mathcal{S}_{\mathrm{BGS}}.\end{aligned} \tag{5}$$

Furthermore, the κ, w logarithm [Eq. (3)] in the particular cases $w = 0$ and $w = 1$, correspond to the κ logarithm and to the q logarithm (with $q = 1 + 2\kappa$) functions, respectively. In this way, it is remarkable that the STM entropy [Eq. (1) or (2)] reduces the Kaniadakis κ entropy [41–46] and Tsallis q entropy [47–51] in the $w = 0$ and $w = 1$ cases.

III. CLASSIC MAXIMUM LIKELIHOOD ESTIMATOR

The principle of maximum likelihood is one of the most used inference methods to obtain the optimum values of the model parameters from the observations, through the so-called MLE. Let $\mathbf{x} = \{x_1, x_2, \ldots, x_N\}$ be an independent and identically distributed sample according to a probability function $p(x|m)$, with $m \in \mathcal{M}$ the model parameters, the MLE of m is obtained through likelihood function:

$$\underset{m\in\mathcal{M}}{\texttt{arg max}}\ L(m|\mathbf{x}) := \prod_{i=1}^{N} p(x_i|m), \tag{6}$$

or by employing the log-likelihood function:

$$\underset{m\in\mathcal{M}}{\texttt{arg max}}\ \mathcal{L}(m|\mathbf{x}) = \sum_{i=1}^{N} \ln[p(x_i|m)], \tag{7}$$

with $\mathcal{L} = \ln[L(m|\mathbf{x})]$. In this classic framework, the MLE estimates the model parameters in which the likelihood function [Eq. (6)] (or its logarithm [Eq. (7)]) have the maximum value.

Assuming that errors are independent and identically distributed according to a standard Gaussian distribution,

$$p(x) = \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}x^2\right), \tag{8}$$

the associated MLE is given, as expected, by a Gaussian likelihood function:

$$L(m|\mathbf{x}) = \left(\frac{1}{\sqrt{2\pi}}\right)^N \exp\left[-\frac{1}{2}\sum_{i=1}^{N} x_i^2(m)\right], \tag{9}$$

and, therefore, the Gaussian log-likelihood is given by:

$$\mathcal{L}(m|\mathbf{x}) = -\frac{N}{2}\ln(2\pi) - \frac{1}{2}\sum_{i=1}^{N} x_i^2(m). \tag{10}$$

Since that maximizing Eq. (10) is equivalent to minimizing its negative, the parameter estimation in this framework can be performed by minimizing the following function:

$$\phi(m) = \frac{1}{2}\sum_{i=1}^{N} x_i^2(m) + \frac{N}{2}\ln(2\pi), \tag{11}$$

where ϕ is the least-squares objective function, in which we call hereinafter of classic objective function.

Although the least-squares approach, and consequently data inversion based on Gauss's law of error, is quite popular, it is very sensitive to aberrant measures (outliers) [16]. This may be seen quantitatively by analyzing the influence function, which is defined as [52,53]:

$$\Upsilon := \frac{\partial\phi(m)}{\partial m}. \tag{12}$$

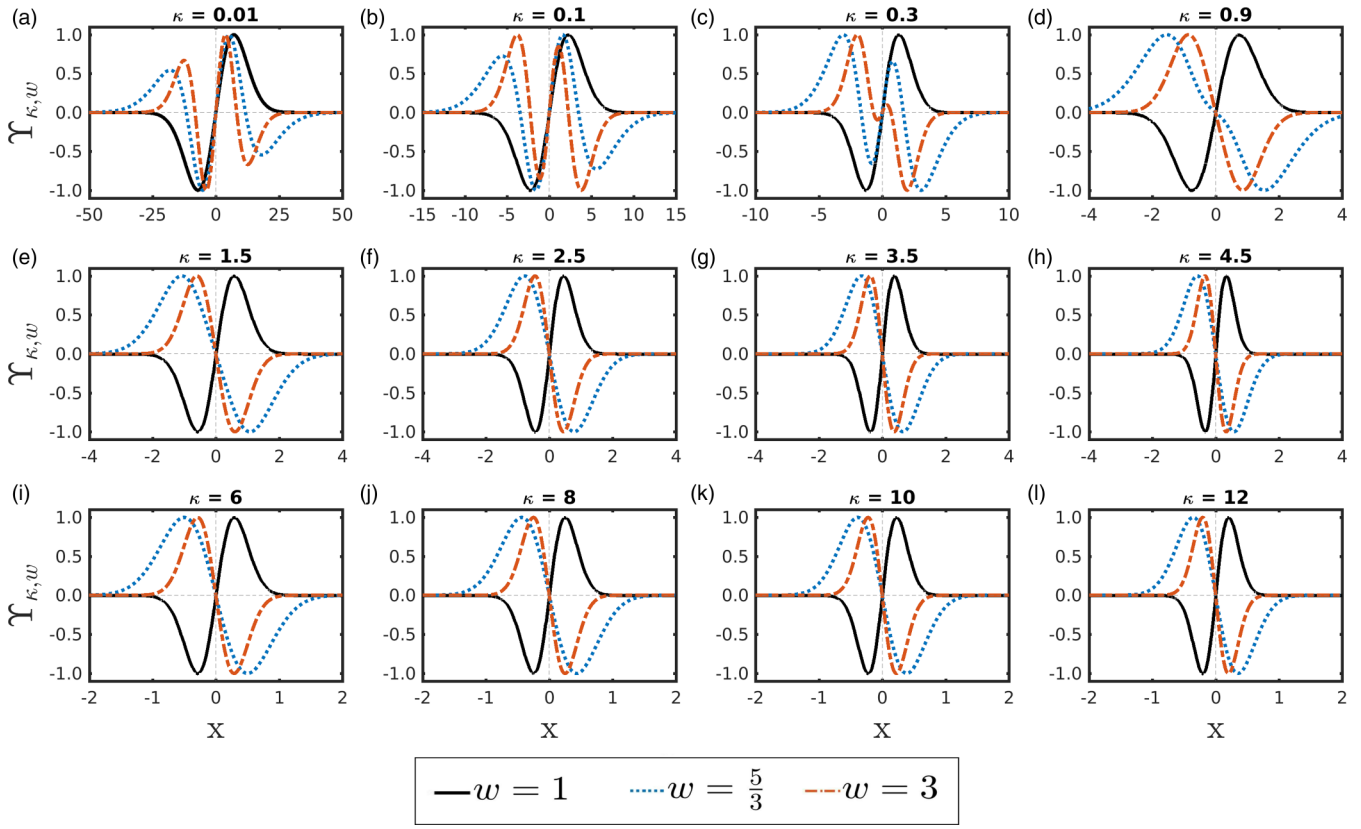


FIG. 1. The κ, w influence function plots [Eq. (19)] for typical for the (a) $\kappa = 0.01$, (b) $\kappa = 0.1$, (c) $\kappa = 0.3$, (d) $\kappa = 0.9$, (e) $\kappa = 1.5$, (f) $\kappa = 2.5$, (g) $\kappa = 3.5$, (h) $\kappa = 4.5$, (i) $\kappa = 6.0$, (j) $\kappa = 8.0$, (k) $\kappa = 10.0$, and (l) $\kappa = 12.0$ cases. The black solid, the blue dotted, and the dash-dotted curves represent, respectively, the $w = 1$, $w = 5/3$, and $w = 3$ cases.

An objective function is said to be robust (or outlier resistant) if Υ tends to zero when there are outliers in the observed data set. Otherwise, the objective function is nonrobust. The influence function related to the classic objective function [Eq. (11)] is

$$\Upsilon = \sum_{i=1}^{N} x_i(m) = \sum_{i=1}^{N} \left[d_i^{\text{obs}} - d_i^{\text{mod}}(m) \right], \tag{13}$$

where $x(m) = d^{\text{obs}} - d^{\text{mod}}(m)$ denotes the error. Let $d_{\text{out}}^{\text{obs}}$ be an outlier, analyzing the latter equation [Eq. (13)] we notice that $\Upsilon \to \infty$ under $d_{\text{out}}^{\text{obs}} \to \infty$ and therefore the parameter estimation based on the classic approach is nonrobust.

IV. MAXIMUM $L_{\kappa,w}$-LIKELIHOOD ESTIMATOR

In order to formulate a κ, r generalization of Gauss's law of error, Scarfone emphet al. has been proposed in Ref. [39], a κ, r generalization of the classic likelihood function by means of the so-called κ, r product. In this regard, the $L_{\kappa,w}$ likelihood function is obtained by replacing the ordinary product in the classic likelihood function [Eq. (6)] with the $\otimes_{\kappa,w}$ operator.

In this way, the M$L_{\kappa,w}$E of m is defined through the following optimization problem:

$$\begin{aligned} &\underset{m\in\mathcal{M}}{\texttt{arg max}} L_{\kappa,w}(m|\mathbf{x}) \\ &:= p(x_1|m) \underset{\kappa,w}{\otimes} p(x_2|m) \underset{\kappa,w}{\otimes} \cdots \underset{\kappa,w}{\otimes} p(x_N|m). \end{aligned} \tag{14}$$

We remark that in the limit $(\kappa, r) \to (0, 0)$, the M$L_{\kappa,w}$E [Eq. (14)] reduces to the MLE [Eq. (6)]. In addition, the $L_{\kappa,w}$ log-likelihood is obtained by taking the κ, w logarithm [Eq. (3)] of the both sides of Eq. (14), $\mathcal{L}_{\kappa,w} = \ln_{\kappa,w}(L_{\kappa,w})$:

$$\underset{m\in\mathcal{M}}{\texttt{arg max}} \mathcal{L}_{\kappa,w}(m|\mathbf{x}) = \sum_{i=1}^{N} \ln_{\kappa,w}[p(x_i|m)]. \tag{15}$$

Thus, if the estimator $\hat{m}$ exists, then the M$L_{\kappa,w}$E of m parameter is defined through the following optimization problem:

$$\hat{m} := \underset{m\in\mathcal{M}}{\texttt{arg max}} \ \mathcal{L}_{\kappa,w}(m|\mathbf{x}), \tag{16}$$

in which the classic approach is a particular case in the limit $(\kappa, w) \to (0, 0)$.

Again, based on the assumption that errors are independent and identically distributed according to a standard Gaussian distribution [Eq. (8)] the associated M$L_{\kappa,w}$E is given by the following κ, w-generalized log-likelihood function:

$$\mathcal{L}_{\kappa,w}(m|\mathbf{x}) = \sum_{i=1}^{N} \ln_{\kappa,w} \left\{ \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{1}{2} x_i^2(m) \right] \right\}. \tag{17}$$

We notice that maximizing Eq. (17) is equivalent to minimizing its negative. Thus, we define the κ, w objective function

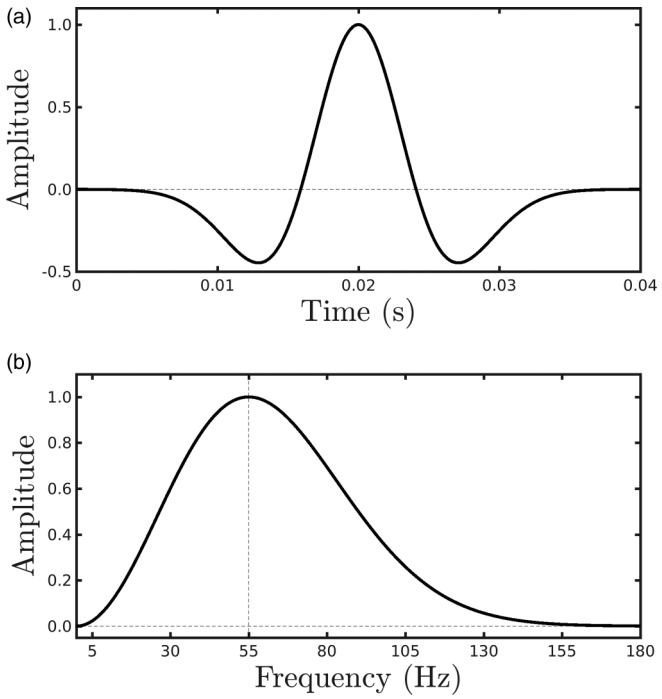


FIG. 2. (a) The 55-Hz Ricker wavelet and its (b) frequency spectrum.

as follows:

$$\phi_{\kappa,w}(m) = -\sum_{i=1}^{N} \ln_{\kappa,w}\left\{\frac{1}{\sqrt{2\pi}}\exp\left[-\frac{1}{2}x_i^2(m)\right]\right\}, \qquad (18)$$

in which it is remarkable that the classic objective function [Eq. (11)] is a particular case in the classic limit $(\kappa, w) \to (0, 0)$.

Thus, the κ, w influence function is

$$\Upsilon_{\kappa,w} = \sum_{i=1}^{N} \frac{x_i}{2}\left\{(2\pi)^{-\frac{\kappa}{2}(w+1)}(w+1)\exp\left[-\frac{\kappa}{2}(w+1)x_i^2\right]\right.$$
$$\left. - (2\pi)^{-\frac{\kappa}{2}(w-1)}(w-1)\exp\left[-\frac{\kappa}{2}(w-1)x_i^2\right]\right\}. \qquad (19)$$

It is worth noting that in classic limit $(\kappa, w) \to (0, 0)$ the κ, w influence function [Eq. (19)] reduces to the classic influence function [Eq. (13)]. Moreover, by analyzing the latter equation [Eq. (19)], we notice that the κ, w objective function is resistant to outlier for the case $\kappa > 0$ and $w \geqslant 1$ since $\Upsilon_{\kappa,w} \to 0$ under $d_{\text{out}}^{\text{obs}} \to \infty$ ($x_i \to \infty$). In this work, we consider the robust cases $w = 1$, $w = 5/3$, and $w = 3$, with $\kappa > 0$. Figure 1 shows the behavior of the κ, w influence function for typical κ values (as indicated in each panel), which indicates that the κ, w objective function for the cases $w = 1$, $w = 5/3$, and $w = 3$ cases are outlier resistant since $\Upsilon_{\kappa,w} \to 0$ under $x \to \pm\infty$. By analyzing this figure, it is remarkable that as the κ and w values increase, the κ, w influence function decays faster to zero, following an asymptotic behavior of the form: $\lim_{x\to\pm\infty}\Upsilon_{\kappa,w}(x) \propto \exp[-\frac{\kappa}{2}(w-1)x^2]$. In addition, we notice that for x close to zero in the cases $w = 5/3$ and $w = 3$, the κ, w influence function is oscillatory due to the misfit between the two terms inside of square brackets in Eq. (19), as depicted in Figs. 1(a)–1(e). Comparing the κ, w influence function [Eq. (19)] with the classic influence function [Eq. (13)], one can see that in our proposal the ith error x_i is weighted by the term in square brackets divided by 2. Since the term in square brackets is always positive and lower than one, the κ, w influence function is down-weighted by the magnitude of the ith error and the κ and w parameters. This robustness property is the key reason κ, w objective function is expected to be outlier-resistant in the parameter estimation task.

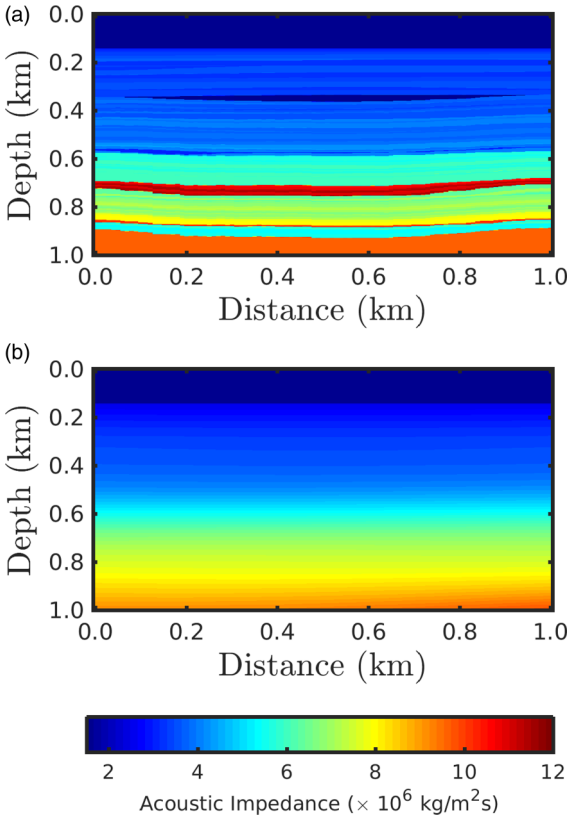


FIG. 3. (a) The portion of the Marmousi model considered as the true model. (b) Initial model used in all numerical simulations.

V. NUMERICAL EXAMPLE

To demonstrate the robustness of data inversion based on M$L_{\kappa,w}$E, we consider a classic geophysical imaging problem named poststack inversion (PSI) [54,55]. The PSI is routinely used to obtain a quantitative acoustic impedance model of the subsurface (model parameters) from seismic reflection data (observed data) (see, for instance, Refs. [56–59]). The PSI forward problem is based on the following convolutional model [60]:

$$d^{\text{mod}}(m, t) = \frac{1}{2}\int_{-\infty}^{\infty} s(\tau)\frac{\partial}{\partial t}[m(t-\tau)]d\tau, \qquad (20)$$

where s represents the seismic source employed in the seismic acquisition, $m = \ln(AI)$ denotes the model parameters with AI being the acoustic impedance, and t denotes the time. We remark that obtaining the acoustic impedance or its logarithm is completely equivalent [60].

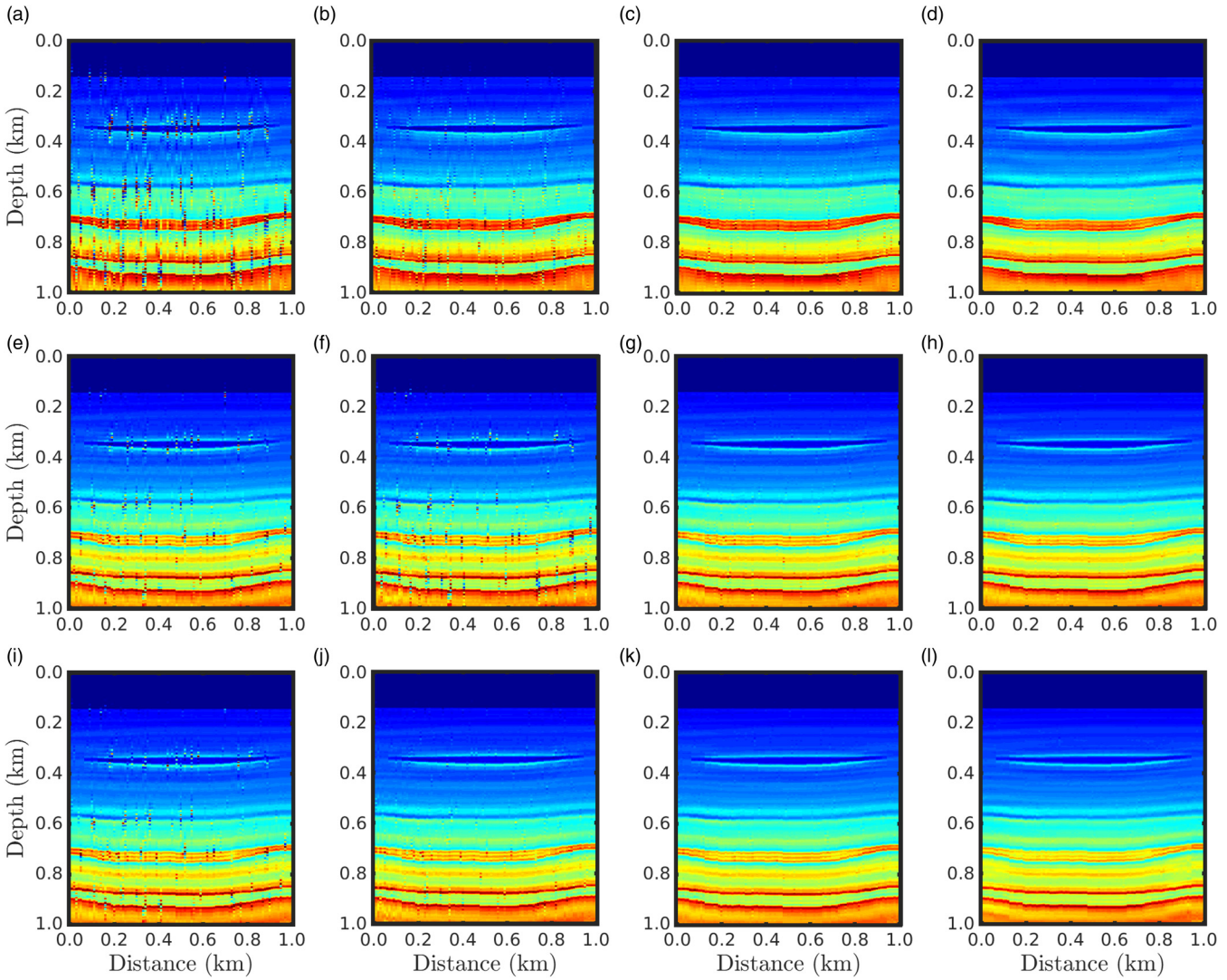


FIG. 4. Reconstructed acoustic impedance models for the %Spike = 1 case, in which panel (a) shows the PSI result by employing the classic approach $[(\kappa, w) \to (0, 1)]$. Panels (b)–(d) refer to PSI results based on M$L_{\kappa,w}$E with $w = 1$ and $\kappa = 0.9$, 4.5, and 12.0, respectively. Panels (e)–(h) show the PSI results based on M$L_{\kappa,w}$E $w = 5/3$ with $\kappa = 0.01$, 0.9, 4.5, and 12.0, while panels (i)–(l) show the $w = 3$ case with $\kappa = 0.01$, 0.9, 4.5, and 12.0.

In practices, the modeled seismic data [Eq. (20)] is computed in a discrete domain using the following expression: $\mathbf{d}^{\text{mod}} = \mathbf{Gm} = \mathbf{SDm}$, in which $\mathbf{S}$ is the so-called wavelet matrix:

$$\mathbf{S} = \begin{bmatrix} s_1 & 0 & \dots & 0 \\ \vdots & s_1 & \vdots & 0 \\ s_n & \vdots & \ddots & 0 \\ 0 & s_n & \vdots & s_1 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & \dots & s_n \end{bmatrix}, \tag{21}$$

and $\mathbf{D}$ denotes the first-order linear differentiation:

$$\mathbf{D} = \frac{1}{2}\begin{bmatrix} -1 & 1 & 0 & \dots & 0 \\ 0 & -1 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & 0 \\ 0 & \dots & \dots & -1 & 1 \end{bmatrix}, \tag{22}$$

being $\mathbf{s} = \{s_1, s_2, \dots, s_n\}$ the discretized seismic source. In this way, the residual data are defined by $x(\mathbf{m}) = \mathbf{Gm} - \mathbf{d}^{\text{obs}}$, where $\mathbf{d}^{\text{obs}}$ denotes the observed data.

Furthermore, we consider a Ricker wavelet as the seismic source [61] in all numerical simulations presented in this work, which is mathematically defined through the following constitutive relation [62]:

$$s(t) = \left(1 - 2\pi^2 \upsilon_p^2 t^2\right)\exp\left(-\pi^2 \upsilon_p^2 t^2\right), \tag{23}$$

where υ_p represents the most energetic frequency content (also known as peak frequency). In particular, we consider a Ricker wavelet with 55-Hz peak frequency, as depicted in Fig. 2.

To simulate realistic circumstances, we consider a section of the Marmousi geological model [63,64] as the true model [Fig. 3(a)], which is based on North Quenguela trough in the Kwanza Basin, Angola. The Marmousi model is widely used to perform numerical tests of new geophysical data inverse strategies.

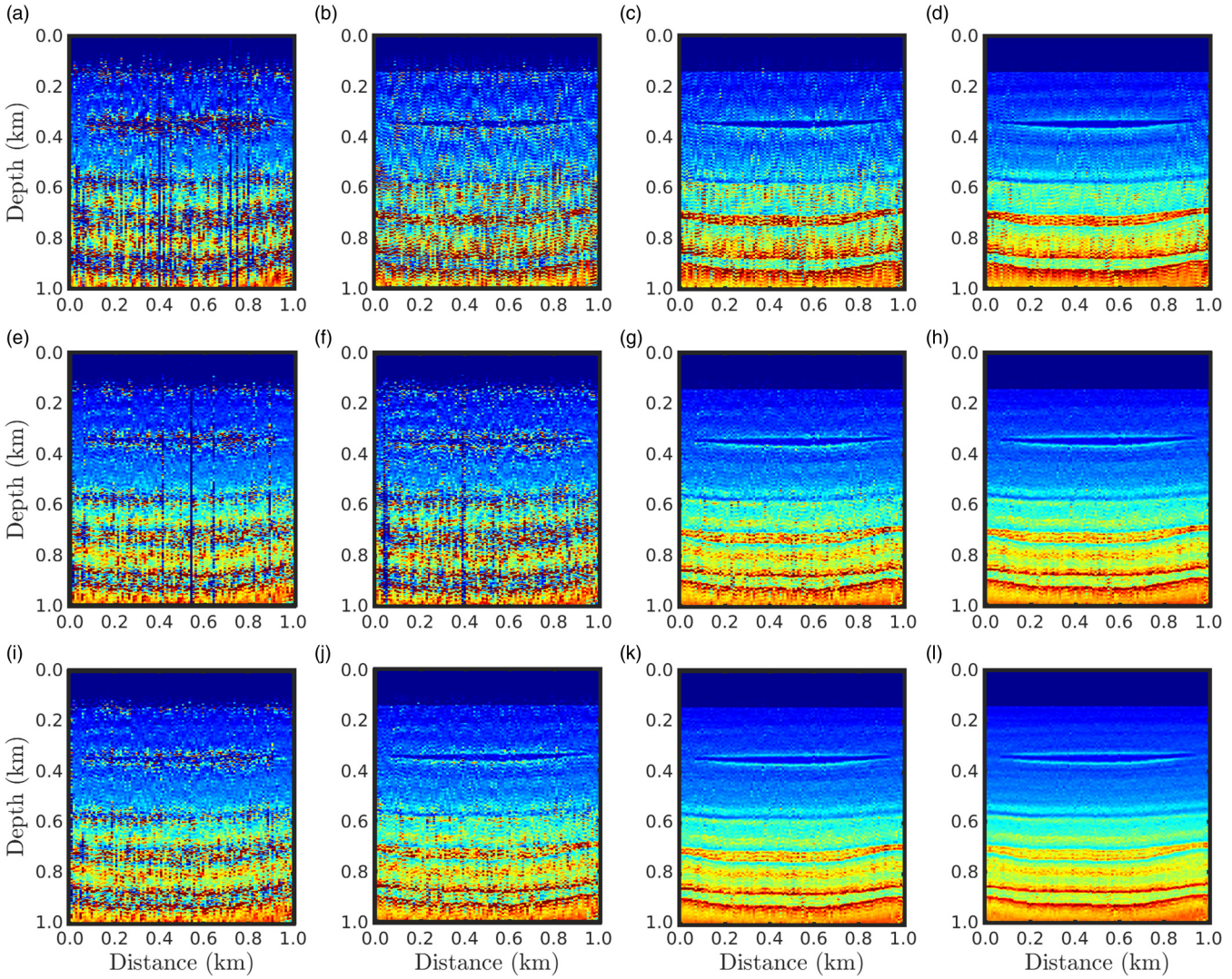


FIG. 5. Reconstructed acoustic impedance models for the %Spike = 30 case, in which the panel (a) shows the PSI result by employing the classic approach [$(\kappa, w) \to (0, 1)$]. Panels (b)–(d) refer to PSI results based on M$L_{\kappa,w}$E with $w = 1$ and $\kappa = 0.9$, 4.5, and 12.0, respectively. Panels (e)–(h) show the PSI results based on M$L_{\kappa,w}$E $w = 5/3$ with $\kappa = 0.01$, 0.9, 4.5, and 12.0, while panels (i)–(l) show the $w = 3$ case with $\kappa = 0.01$, 0.9, 4.5, and 12.0.

To minimize the κ, w objective function, we employ the quasi-Newton method named *limited-memory* Broyden-Fletcher-Goldfarb-Shanno (*L*-BFGS) [65], which updates the model parameters in the direction of decreasing of the gradient, $\nabla_m \phi(\mathbf{m})$:

$$\mathbf{m}_{j+1} = \mathbf{m}_j - \alpha_j \mathbf{H}_j^{-1} \nabla_m \phi(\mathbf{m}_j), \tag{24}$$

where $\alpha_j > 0$ is the step-length of the jth *L*-BGFS iteration and $\mathbf{H}^{-1}$ denotes the inverse of the Hessian matrix. In this work, we compute the step-length α according to the Wolfe conditions [66]. It is worth emphasizing that, in order to decrease the computational cost, *L*-BFGS algorithm computes an approximation of the inverse of the Hessian matrix based on previous gradient evaluations (see Ref. [67] for more details). Moreover, we use the same initial model [Fig. 3(b)] for all numerical simulations carried out in this work. It is worth mentioning that several numerical simulations have been carried out for the noiseless data circumstance, in which the reconstructed model from the classic approach and our proposal were satisfactory and very similar. For this reason, we will not present the results of this ideal case.

From now on, we consider noisy seismic data to examine the robustness of our proposal in estimating physical parameters. In this way, we consider a data set contaminated by Gaussian errors, as the background noise, with a signal-to-noise ratio of 40 dB and a set of large errors (outliers). Concerning the outliers, we consider 80 different noisy scenarios in which each scenario is composed of a given amount of outliers contaminating the data. In this regard, in the first scenario, 1% of the seismic data samples are contaminated by outliers [named %Spike = 1]. In the second one, 2% of the data samples are corrupted by outliers [%Spike = 2], and so on, every 1% to maximum contamination of 80% [%Spike = 80] in the last one. We randomly add outliers to the seismic data to avoid biased samples. For that, we randomly select the samples through a uniform distribution, in which the observed data with Gaussian noise at ith outlier-position is calculated

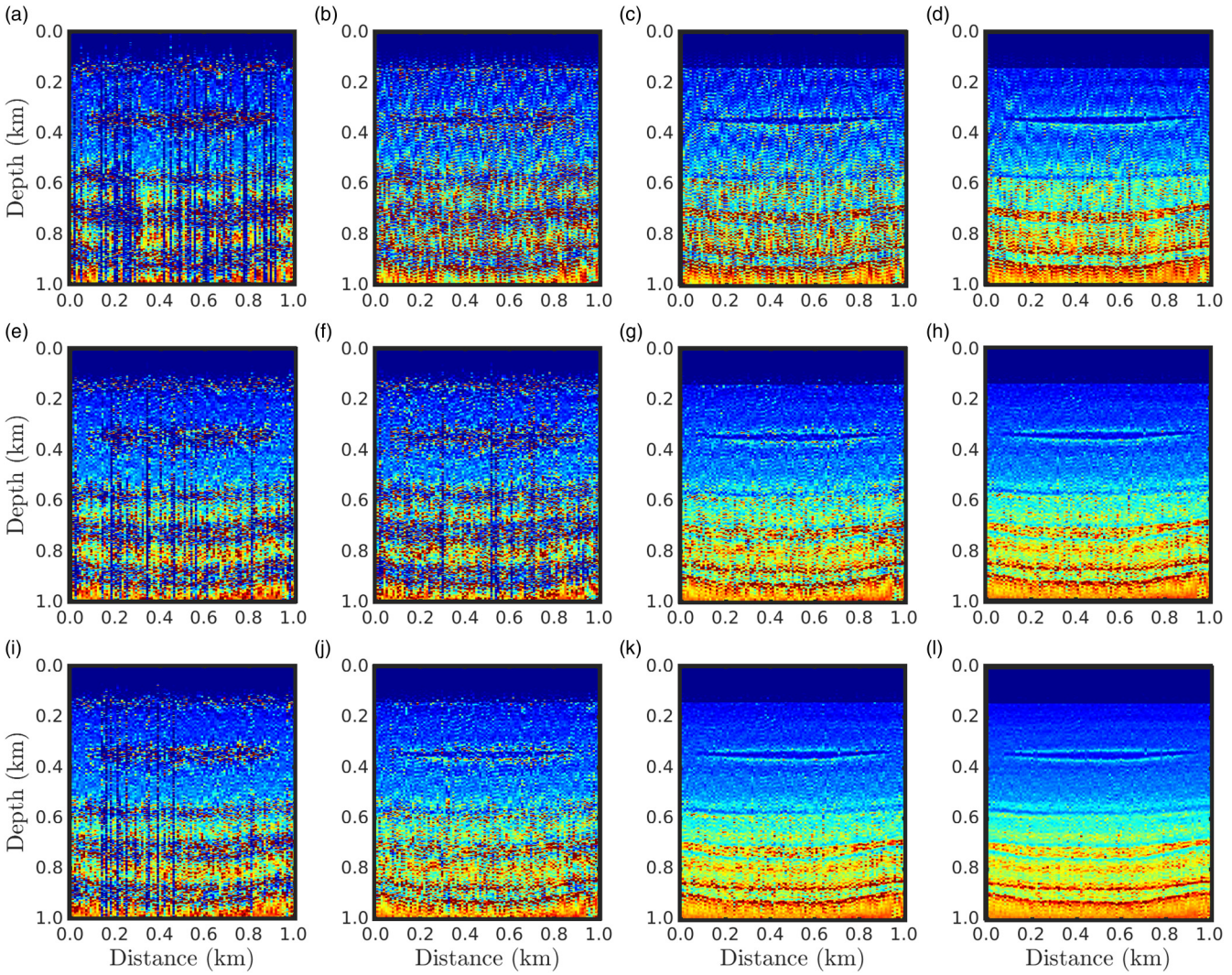


FIG. 6. Reconstructed acoustic impedance models for the %Spike = 80 case, in which the panel (a) shows the PSI result by employing the classic approach $[(\kappa, w) \rightarrow (0, 1)]$. Panels (b)–(d) refer to PSI results based on M$L_{\kappa,w}$E with $w = 1$ and $\kappa = 0.9$, 4.5, and 12.0, respectively. Panels (e)–(h) show the PSI results based on M$L_{\kappa,w}$E $w = 5/3$ with $\kappa = 0.01$, 0.9, 4.5, and 12.0, while panels (i)–(l) show the $w = 3$ case with $\kappa = 0.01$, 0.9, 4.5, and 12.0.

as: $d_i^{\text{obs}} = [d_i^{\text{obs}}]_{\text{Gaussian}_{\text{Noise}}} + \alpha \times \beta$, where $\alpha \in [-2; 2]$ and β follows a standard normal distribution.

Figure 4 shows the reconstructed model for the %Spike = 1 case, in which Fig. 4(a) depicts the classic approach $[(\kappa, w) \rightarrow (0, 1)]$ result. We notice that even with only 1% contamination by outliers, the classic approach generated an acoustic impedance model with a lot of artifacts, as can be seen through the vertical fringes in Fig. 4(a). On the other hand, the quality of the models generated through the data inversion based on the M$L_{\kappa,w}$E is sensitive to the κ and w parameters, as depicted in Figs. 4(b)–4(l). We notice that as the κ parameter increases the amount of artifacts in the reconstructed model decreases and, therefore, the result is more close to the true model [Fig. 3(a)]. Indeed, we already expected this behavior because as the κ value increases the major deviation of a Gaussian behavior is automatically considered.

Figures 5 and 6 show the reconstructed models for the %Spike = 30 and 80 cases, respectively, in which it is noticeable that the performance of the classic approach is very poor, generating wrong and biased models as depicted in panel (a) of these figures. Indeed, the classical framework is very sensitivity to the outliers in the seismic data set as expected [16]. In contrast, from a visual inspection, our proposal is outlier resistant especially for $\kappa > 4.5$, as depicted in panels (b)–(l) of Figs. 5 and 6.

In order to compare the data-inversion results in a quantitative way, we performed 97,441 numerical simulations considering $0 \leqslant \kappa \leqslant 12$ and $0 \leqslant$ %Spike $\leqslant 80$, and then we compute the linear correlation coefficient (Pearson's R) [68] between each reconstructed model and the true model [Fig. 3(a)]. In this paper, the Pearson's R measures the degree of linear relationship between the elements of the reconstructed model and the true model and expresses the degree of correlation through values between 0 (bad reconstructed model) and 1 (perfect reconstructed model).

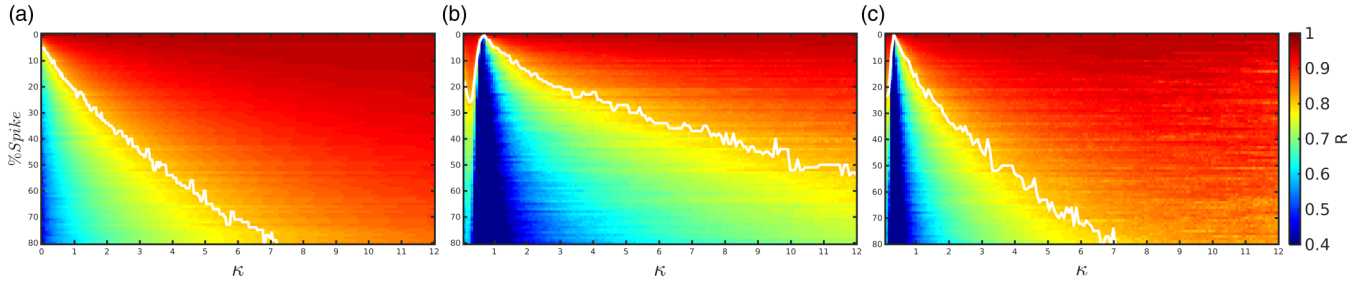


FIG. 7. Heatmaps generated from the Pearson's R between the reconstructed models and the true model [Fig. 3(a)] for all numerical simulations carried out in this study for the cases (a) $w = 1$, (b) $w = 5/3$, and (c) $w = 3$. The white solid line separates the reconstructed models strongly correlated ($R \geqslant 0.8$) with the true model, from the others ($R < 0.8$).

Due to the large number of data inversions carried out, we summarize the Pearson's R computed in each numerical simulation on heatmaps, as depicted in Fig. 7. On these maps, the reconstructed models with high similarity to the true model are represented by the hot colors (close to red scale) and the less similar ones by the cold colors (close to blue scale). In addition, the white line separates the reconstructed acoustic impedance models strongly correlated ($R \geqslant 0.8$) with the true model, following the strength-scale suggested by Ref. [69], from the others ($R < 0.8$). From a visual inspection of Fig. 7, we note that the models reconstructed in circumstances in which the number of spurious measurements (outliers) is less than 5% (%Spike < 5) are strongly correlated with the true model for any $\kappa \geqslant 0$ in the case $w = 1$ [Fig. 7(a)]. However, in situations where the seismic data are heavily corrupted by outliers, the best reconstructed models are associated with the highest κ values, especially for the $\kappa > 7$ case with $w = 1$ and $w = 3$ as depicted by the reddish area at the top of the white line in Figs. 7(a) and 7(c), respectively. The case $w = 5/3$, on the other hand, proved to be robust for contamination of up to 50% when $\kappa \to 12$.

Furthermore, the notable bluish region on the maps for cases $w = 5/3$ [Fig. 7(b)] and $w = 3$ [Fig. 7(c)] indicate that the reconstructed models are weakly correlated with the true model. In this regard, such unsatisfactory models were obtained just because under these parameters values the κ, w influence function is oscillatory around zero error [see Figs. 1(a)–1(e)], which is an undesirable behavior for an error law because, in some circumstances, positive and negative errors with similar magnitude have the same influence on the data-inversion process.

VI. FINAL REMARKS

In this work, we have presented a robust methodology to mitigate the sensitivity of the probabilistic maximum-likelihood method, for Gaussian distributions, to erratic measurements (outliers). Based on the κ, w-generalized algebra introduced by Ref. [39], we proposed a M$L_{\kappa,w}$E for Gaussian distributions which is robust to non-Gaussian errors. Furthermore, we analytically investigated the robustness properties of M$L_{\kappa,w}$E from the influence function analysis. In this way, we have discussed the role of the κ and w parameters for robust inference. In this regard, from the analyses performed in Sec. IV, we concluded that the M$L_{\kappa,w}$E is robust under outliers for the $\kappa > 0$ and $w \geqslant 1$ case and nonrobust for other situations. Indeed, the geophysical data-inversion example presented in Sec. V, confirmed that the objective function based on M$L_{\kappa,w}$E is a powerful tool for a reliable estimation of physical parameters from very noisy data. In particular, we have verified that a κ value close to 12 is a good choice to makes the κ, w objective function robust to a lot of erratic data.

It is worth emphasizing that the results presented in this study show that our proposal ignores the effect of outliers in the data-inversion process, which allows us to dispense a long work of data preprocessing. In addition, the M$L_{\kappa,w}$E-based objective function is a promising methodology to deal with low-quality data. Furthermore, it is worth noting that the M$L_{\kappa,w}$E-based objective function introduced in the present study may be easily adapted to inference physical parameters of any inverse problem. In fact, we notice that the requirement for employing the M$L_{\kappa,w}$E-based objective function is the definition of the residual data (the difference between the modeled data and the observed data). In other words, what differs between one application and another is the forward modeling process, which can be represented, for instance, by stochastic models in the description of biological systems [70–73], as well as by regularized problems in machine learning [74,75], epidemiological models [34,76], among many others in applied physics [77,78]. As a future perspective, since there are many other probability distributions, we intend to study to apply M$L_{\kappa,w}$E to other distributions and statistical physics applications.

[1] A. Turpin, V. Kapitany, J. Radford, D. Rovelli, K. Mitchell, A. Lyons, I. Starshynov, and D. Faccio, Phys. Rev. Lett. **126**, 174301 (2021).

[2] M. Razavy, *An Introduction to Inverse Problems in Physics* (World Scientific, Singapore, 2020).

[3] S. L. E. F. da Silva, J. Julià, and F. H. R. Bezerra, Bull. Seism. Soc. Am. **107**, 1495 (2017).

[4] M. Dolgin and P. D. Einziger, Phys. Rev. Lett. **93**, 148101 (2004).

[5] A. Adler and A. Boyle, IEEE Trans. Biomed. Eng. **64**, 2494 (2017).

[6] C. Putensen, B. Hentze, S. Muenster, and T. Muders, J. Clin. Med. **8**, 1176 (2019).

[7] A. Tarantola, *Inverse Problem Theory and Methods for Model Parameter Estimation* (SIAM, Philadelphia, PA, 2005).

[8] J. V. T. de Lima, S. L. E. F. da Silva, J. M. de Araújo, G. Corso, and G. Z. dos Santos Lima, Eur. Phys. J. Plus **136**, 269 (2021).

[9] R. Blume-Kohout, Phys. Rev. Lett. **105**, 200504 (2010).

[10] F. W. de Freitas Silva, S. L. E. F. da Silva, M. V. C. Henriques, and G. Corso, PLoS ONE **14**, e0213847 (2019).

[11] C. R. Contaldi, M. Pieroni, A. I. Renzini, G. Cusin, N. Karnesis, M. Peloso, A. Ricciardone, and G. Tasinato, Phys. Rev. D **102**, 043502 (2020).

[12] S. L. E. F. da Silva, P. T. C. Carvalho, C. A. N. da Costa, J. M. de Araújo, and G. Corso, PLoS ONE **15**, e0240999 (2020).

[13] A. Corral, I. Serra, and R. Ferrer-i-Cancho, Phys. Rev. E **102**, 052113 (2020).

[14] M. Kendall and A. Stuart, *The Advanced Theory of Statistics: Inference and Relationship* (Hodder Arnold, London, 1979).

[15] R. E. Thomson and W. J. Emery, Statistical methods and error handling, in *Data Analysis Methods in Physical Oceanography,* 3rd ed., edited by R. E. Thomson and W. J. Emery (Elsevier, Boston, 2014), pp. 219–311.

[16] J. F. Claerbout and F. Muir, Geophysics **38**, 826 (1973).

[17] J. A. Scales and A. Gersztenkorn, Robust methods in inverse theory, Inverse Probl. **4**, 1071 (1988).

[18] L. Amundsen, Geophysics **56**, 2027 (1991).

[19] P. Zhou, Y. Lv, H. Wang, and T. Chai, IEEE Trans. Ind. Electron. **64**, 7141 (2017).

[20] A. Y. Aravkin, M. P. Friedlander, F. J. Herrmann, and T. van Leeuwen, Math. Program. **135**, 101 (2012).

[21] A. Ubaidillah, K. A. Notodiputro, A. Kurnia, A. Fitrianto, and I. W. Mangku, IOP Conf. Ser.: Earth Environ. Sci. **58**, 012013 (2017).

[22] P. J. Huber, Ann. Statist. **1**, 799 (1973).

[23] K. P. Bube and R. T. Langan, Geophysics **62**, 1183 (1997).

[24] S. L. E. F. da Silva, P. T. C. Carvalho, C. A. N. da Costa, J. M. de Araújo, and G. Corso, Misfit function for full waveform inversion based on Shannon entropy for deeper velocity model updates in *Proceedings of the SEG Technical Program Expanded Abstracts* (SEG Library, 2019), p. 1556.

[25] S. L. E. F. da Silva, G. Z. dos Santos Lima, J. M. de Araújo, and G. Corso, Physica A **563**, 125496 (2021).

[26] H. Suyari and M. Tsukada, IEEE Trans. Inf. Theory **51**, 753 (2005).

[27] Y. Hasegawa and M. Arita, Physica A **388**, 3399 (2009).

[28] D. Ferrari and Y. Yang, Ann. Statist. **38**, 753 (2010).

[29] S. L. E. F. da Silva, C. A. da Costa, P. T. C. Carvalho, J. M. de Araújo, L. dos Santos Lucena, and G. Corso, Physica A **548**, 124473 (2020).

[30] I. P. de Lima, S. L. E. F. da Silva, G. Corso, and J. M. de Araújo, Entropy **22**, 464 (2020).

[31] S. L. E. F. da Silva, C. A. N. da Costa, P. T. C. Carvalho, J. M. de Araújo, L. Lucena, and G. Corso, An objective function based on q-Gaussian distribution for full-waveform inversion, in *Proceedings of the EAGE 2020 Annual Conference & Exhibition Online* (European Association of Geoscientists & Engineers, 2020), p. 1.

[32] T. Wada and H. Suyari, Phys. Lett. A **348**, 89 (2006).

[33] Sergio Luiz E. F. da Silva, Pedro Tiago C. Carvalho, J. M. de Araújo, and G. Corso, Phys. Rev. E **101**, 053311 (2020).

[34] G. Kaniadakis, M. M. Baldi, T. S. Deisboeck, G. G. D. T. Hristopulos, A. M. Scarfone, A. Sparavigna, T. Wada, and U. Lucia, Sci. Rep. **10**, 19949 (2020).

[35] S. L. E. F. da Silva, G. Z. dos Santos Lima, E. V. Volpe, J. M. de Araújo, and G. Corso, Eur. Phys. J. Plus **136**, 518 (2021).

[36] J. L. Devore, *Probability and Statistics for Engineering and the Sciences*, 8th ed. (Cengage Learning, Boston, MA, 2011).

[37] D. P. Mittal, Metrika **22**, 35 (1975).

[38] B. D. Sharma and I. J. Taneja, Metrika **22**, 205 (1975).

[39] A. Scarfone, H. Suyari, and T. Wada, Cent. Eur. J. Phys. **7**, 414 (2009).

[40] G. Kaniadakis, M. Lissia, and A. M. Scarfone, Phys. Rev. E **71**, 046128 (2005).

[41] G. Kaniadakis, Physica A **296**, 405 (2001).

[42] G. Kaniadakis, Phys. Rev. E **66**, 056125 (2002).

[43] G. Kaniadakis, Phys. Rev. E **72**, 036108 (2005).

[44] G. Kaniadakis, Eurphys. Lett. **133**, 10002 (2021).

[45] B. Lei and J. Fan, Knowl. Based Syst. **225**, 107089 (2021).

[46] S. L. E. da Silva, Chaos Solitons Fractals **143**, 110622 (2021).

[47] C. Tsallis, J. Stat. Phys. **52**, 479 (1988).

[48] M. Gell-Mann and C. Tsallis, *Nonextensive Entropy—Interdisciplinary Applications* (Oxford University Press, New York, 2004), p. 440.

[49] S. L. E. F. da Silva, Physica A **565**, 125539 (2021).

[50] C. Tsallis, Nonextensive statistical mechanics and thermodynamics: Bibliography, http://tsallis.cat.cbpf.br/TEMUCO.pdf (2021).

[51] S. L. E. F. da Silva and G. Corso, Eur. Phys. J. B **94**, 25 (2021).

[52] F. R. Hampel, E. M. Ronchetti, P. J. Rousseeuw, and W. A. Stahel, *Robust Statistics: The Approach Based on Influence Functions*, 1st ed. (Wiley-Interscience, New York, 2005).

[53] F. R. Hampel, J. Am. Stat. Assoc. **69**, 383 (1974).

[54] B. Russell, *Introduction to Seismic Inversion Methods* (Birkhäuser Basel, Boston, MA, 1998), p. 178.

[55] B. Russell and D. Hampson, Comparison of Poststack Seismic Inversion Methods, in *Proceedings of the SEG Technical Program Expanded Abstracts* (SEG Library, 1991), p. 876.

[56] S. K. Ghosh, Geophysics **65**, 951 (2000).

[57] G. B. Madiba and G. A. McMechan, Geophysics **68**, 1460 (2003).

[58] A. Kumar Ray and S. Chopra, First Break **34**, 5 (2016).

[59] C. A. M. Assis, H. B. Santos, and J. Schleicher, Geophysics **84**, N15 (2019).

[60] M. K. Sen, *Seismic Inversion* (Society of Petroleum Engineers, Richardson, TX, 2006).

[61] N. Ricker, Bull. Seism. Soc. Am. **33**, 197 (1943).

[62] Y. Wang, Geophysics **80**, A31 (2015).

[63] R. Versteeg, Lead. Edg. **13**, 927 (1994).

[64] G. S. Martin, R. Wiley, and K. J. Marfurt, Lead. Edg. **25**, 156 (2006).

[65] J. Nocedal and S. J. Wright, *Numerical Optimization* (Springer, New York, 2006).

[66] P. Wolfe, SIAM Rev. **11**, 226 (1969).

[67] R. H. Byrd, P. Lu, J. Nocedal, and C. Zhu, J. Sci. Comput. **16**, 1190 (1995).

[68] K. Pearson and O. M. F. E. Henrici, Philos. Trans. Royal Soc. Lond. **187**, 253 (1896).

[69] J. D. Evans, *Straightforward Statistics for the Behavioral Sciences* (Brooks/Cole Publishing Company, Pacific Grove, CA, 1996).

[70] S. Spezia, L. Curcio, A. Fiasconaro, N. Pizzolato, D. Valenti, B. Spagnolo, P. L. Bue, and S. C. E. Peri, Eur. Phys. J. B **65**, 453 (2008).

[71] A. Giuffrida, D. Valenti, G. Ziino, B. Spagnolo, and A. Panebianco, Eur. Food Res. Technol. **228**, 767 (2009).

[72] G. Denaro, D. Valenti, B. Spagnolo, A. Bonanno, G. Basilone, S. Mazzola, S. W. Zgozi, and S. Aronica, PLoS ONE **8**, e66765 (2013).

[73] D. Valenti, G. Denaro, B. Spagnolo, S. Mazzola, G. Basilone, F. Conversano, C. Brunet, and A. Bonanno, Ecol. Complex. **27**, 84 (2016).

[74] F. Bauer, S. Pereverzev, and L. Rosasco, J. Complex. **23**, 52 (2007).

[75] M. Prato and L. Zanni, J. Phys. Conf. Ser. **135**, 012085 (2008).

[76] U. Tirnakli and C. Tsallis, Front. Phys. **8**, 557 (2020).

[77] A. J. Silva Neto and M. N. Özişik, J. Appl. Phys. **73**, 2132 (1993).

[78] E. Ayón-Beato, A. García, R. Mansilla, and C. A. Terrero-Escalante, Phys. Rev. D **62**, 103513 (2000).