CrossMark

Source localization and denoising: a perspective from the TDOA space

Marco Compagnoni[1] · Antonio Canclini[2] · Paolo Bestagini[2] · Fabio Antonacci[2] · Augusto Sarti[2] · Stefano Tubaro[2]

Abstract In this manuscript, we formulate the problem of source localization based on Time Differences of Arrival (TDOAs) in the TDOA space, i.e., the Euclidean space spanned by TDOA measurements. More specifically, we show that source localization can be interpreted as a denoising problem of TDOA measurements. As this denoising problem is difficult to solve in general, our analysis shows that it is possible to resort to a relaxed version of it. The solution of the relaxed problem through linear operations in the TDOA space is then discussed, and its analysis leads to a parallelism with other state-of-the-art TDOA denoising algorithms. Additionally, we extend the proposed solution also to the case where only TDOAs between few pairs of microphones within an array have been computed. The reported denoising algorithms are all analytically justified, and numerically tested through simulative campaign.

Keywords TDOA space · TDOA denoising · TDOA redundancy · Source localization

✉ Antonio Canclini
   antonio.canclini@polimi.it

   Marco Compagnoni
   marco.compagnoni@polimi.it

   Paolo Bestagini
   paolo.bestagini@polimi.it

   Fabio Antonacci
   fabio.antonacci@polimi.it

   Augusto Sarti
   augusto.sarti@polimi.it

   Stefano Tubaro
   stefano.tubaro@polimi.it

[1] Dipartimento di Matematica, Politecnico di Milano, Piazza Leonardo da Vinci 32, 20133 Milan, Italy

[2] Dipartimento di Elettronica, Informazione e Bioingegneria, Politecnico di Milano, Via Ponzio 34/5, 20133 Milan, Italy

⨊ Springer

1 Introduction

Source localization is a research theme that has significantly grown in popularity in the past few decades, and whose interest ranges from audio to radar. As far as audio signal processing is concerned, several applications including teleconferencing (D'Arca et al. 2014), audio-surveillance (Valenzise et al. 2007) and human–machine interaction (Trifa et al. 2007) can benefit from the knowledge of the source location. Among the techniques that are available in the literature (Benesty and Huang 2004), those based on Time Difference of Arrival (TDOA) measurements are particularly appreciated for their modest computational requirements. TDOAs, in fact, are usually estimated through peak-picking on the Generalized Cross Correlation of the signals acquired at microphone pairs (Ianniello 1982; Knapp and Carter 1976), or on the whole set of microphones (Chen et al. 2002; Hu and Yang 2010). TDOAs can be easily converted to range differences (RD), once the sound speed is known. The source location is then found as the point in space that best fits the RD measurements according to properly defined cost functions (Beck et al. 2008; Hahn and Tretter 1973; Huang et al. 2001; Schau and Robinson 1987; Schmidt 1972; Stoica and Nehorai 1998). More recently, the widespread diffusion of sensor networks stemmed an interest in source localization also in other research communities, such as remote sensing and radar (Koch and Westphal 1995; Yimin et al. 2008). In this context range differences are obtained from TDOAs (Kehu et al. 2009), or from energy measurements (Hu and Li 2002).

The main drawback of TDOA-based localization techniques lies in their sensitivity to measurement noise. In particular, we can distinguish between additive noise (generally due to sampling in the time domain, circuit noise, but also other physical phenomena) and outlier measurements (produced by reverberation or interfering sources). Outlier identification and removal has been widely studied in the literature (see for instance Scheuing and Yang 2008 and references therein; Canclini et al. 2013, 2015). Therefore, applying one of these techniques it is possible to remove outliers from the pool of available measurements. Nonetheless, additive noise still impairs the localization accuracy.

In this manuscript we interpret the problem of source localization studying the effect of additive noise on TDOA measurements using the TDOA space formalism, i.e., a space in which a set of measured TDOAs is mapped into a point. The sensitivity to noise afflicts also range differences obtained from energy measurements. Indeed, the hostile propagation conditions yield a difference of the measured energy from the ideal free-field assumption. In the following we will specifically refer to the problem of localizing acoustic sources, but the theory can be readily applied also to other kinds of signals. The concept of TDOA space is not novel and was first introduced in Spencer (2007) for localization purposes. From that representation, a TDOA map (from the space of source locations to the space of TDOAs) was later introduced and analytically derived in Compagnoni et al. (2014), which proposed an exhaustive analytic study of the identifiability and invertibility of this map for the three-microphone case. In the most general case, given a set of $n + 1$ microphones, $q = n(n + 1)/2$ TDOAs can be computed considering all the possible microphone pairs. In a noiseless scenario, however, we can always find an independent set of n such TDOAs that we can compute all the other TDOAs from. This is why most TDOA-based algorithms define a reference microphone, with respect to which the n independent TDOAs are computed. In the TDOA space, this corresponds to the fact that TDOAs lie on a linear subspace V_n of the q-dimensional TDOA space. This subspace can be computed in closed form through simple considerations. Feasible TDOAs (i.e., points in the TDOA space that correspond to source locations) are bound to lie in a region $\Theta_n \subset V_n$. In Compagnoni et al. (2014) authors derive Θ_n in terms of real algebraic geometry.

Working in the TDOA space essentially means solving an estimation problem in its dual space. As typically done in estimation theory, using a dual domain enables to split a problem into two parts. In our case, using the TDOA space formalism, source localization can be interpreted as a two-step procedure: (1) a denoising operation, which consists in removing or attenuating part of the additive noise; (2) the application on the denoised TDOAs of a simple mapping from the TDOA space to the geometric space. Starting from this perspective, in this work we provide a deeper investigation on the geometrical characteristics of the TDOA space. More specifically, we first derive the correct denoising formulation that fully describes the source localization problem. As the denoising problem formulated this way is not easy to deal with, we resort to a relaxed version of it, which exploits the linear subspace V_n. In particular, we show that additive noise can be decomposed into the sum of two orthogonal components, and the relaxed problem formulation aims at reducing only one of them, still positively impacting on source localization.

This relaxed version of the problem was implicitly solved in So et al. (2008) and Schmidt (1996), where the authors derive closed-form expressions for converting the full set of TDOAs to the nonredundant one. However, authors in So et al. (2008) and Schmidt (1996) limited their analysis to simulations showing that this conversion is able to reduce the impact of noise in localization accuracy. Instead, working in the TDOA space paves the way to a deeper understanding of the impact of relaxed denoising on source localization. In particular we will: (1) analytically prove the positive effect of denoising on source localization through a set of solid theorems, thus also theoretically validating So et al. (2008) and Schmidt (1996); (2) find a solution to the relaxed denoising problem when some TDOA measurements are not available; (3) quantify analytically the improvement in localization accuracy brought by the use of the denoised TDOAs given a specific localization algorithm in use. We accomplish this analysis by means of the error propagation theory introduced in Compagnoni et al. (2012).

We test the presented algorithm also under different noise hypotheses to show that it works also if the underlying assumptions are not strictly verified. In particular, Monte Carlo simulations were carried out to show how different state-of-the-art techniques [the SRD-LS algorithm (Beck et al. 2008), least squares (Smith and Abel 1987)] and Gillette–Silverman (Gillette and Silverman 2008) methods benefit from denoising, approaching the RMSE Lower Bound (RLB) implied by the Cramer–Rao Lower Bound (CRLB) (Benesty and Huang 2004). We also show that it is possible to perform denoising on a set of TDOAs including $q-s$, $s > 0$ measurements, with an apparent advantage in terms of accuracy.

The rest of the manuscript is structured as follows. In Sect. 2 we deeply introduce the formalism of TDOA space. In Sect. 3 we interpret the problem of source localization within this context. In Sect. 4 we provide the denoising formulation of the source localization problem, also reporting the relaxed problem version and an algorithm for its solution. A parallelism with related state-of-the-art works is also provided. In Sect. 5 we analytically prove the positive impact of denoising on source localization, and provide additional simulative analysis. Section 6 is devoted to denoising problem formulation and solution in case some TDOAs are missing within the pool of measurements (i.e., we measure TDOAs using only a few pairs of microphones). Finally, Sect. 7 remarks some final conclusions highlighting possible open future research lines.

2 Theoretical background

In this section we offer the reader some background that will simplify the reading of this article. In particular, we first provide the formal definition of the TDOA space. Then, we

give the interpretation of noisy measurements and source localization problem in the TDOA space.

2.1 The TDOA space

The ideas of the TDOA space; the feasible set of TDOA measurements; and the TDOA map appeared in several manuscripts concerning multilateration (see for example Compagnoni et al. 2014; Grafarend and Shan 2002; Schmidt 1996; Spencer 2007). These concepts are the essential ingredients for the mathematical definition and analysis of many problems involving TDOA measurements, such as source localization, synchronization and calibration of the receivers. A recent example in this direction can be found in Alameda-Pineda and Horaud (2014), where the TDOA space formalism is used for defining a novel algorithm to estimate the TDOAs and concurrently locate the source. In the following, we present the basic definitions and properties regarding the TDOA space.

Let $\mathbf{m_i} = (x_i, y_i, z_i)^T$, $i = 0, \ldots, n$ be the sensor locations and $\mathbf{x}$ be the source position in the 3D Euclidean space $\mathbb{R}^3$. For notational simplicity, and with no loss of generality, in what follows we assume the sound speed to be equal to 1, so that the noiseless TDOAs correspond to the range differences. This way, given any pair of sensors $(\mathbf{m_j}, \mathbf{m_i})$, $n \geq j > i \geq 0$, the relative TDOA is a function of the source position $\mathbf{x}$ and it can be defined as

$$\tau_{ji} : \mathbb{R}^3 \longrightarrow \mathbb{R} \atop \mathbf{x} \longmapsto \tau_{ji}(\mathbf{x}), \tag{1}$$

where

$$\tau_{ji}(\mathbf{x}) = \|\mathbf{x} - \mathbf{m_j}\| - \|\mathbf{x} - \mathbf{m_i}\|. \tag{2}$$

If we collect the $q = \frac{n(n+1)}{2}$ range differences in a q-dimensional vector, we obtain the map

$$\boldsymbol{\tau_n^*} : \mathbb{R}^3 \longrightarrow \mathbb{R}^q \atop \mathbf{x} \longmapsto (\tau_{10}(\mathbf{x}), \tau_{20}(\mathbf{x}), \ldots, \tau_{n\,n-1}(\mathbf{x}))^T. \tag{3}$$

In Compagnoni et al. (2014), $\boldsymbol{\tau_n^*}$ has been called the complete TDOA map, while the resulting target set $\mathbb{R}^q$ of $\boldsymbol{\tau_n^*}$ is referred to as the TDOA space or τ-space. Clearly, a point in the TDOA space corresponds to any set of TDOA measurements. Moreover, in a noiseless scenario, the subset of the τ-space containing the TDOAs generated by all the potential source positions coincides with the image $\text{Im}(\boldsymbol{\tau_n^*})$ of the TDOA map, and we call it Θ_n. This means that any collection of noiseless TDOAs defines a point $\boldsymbol{\tau} = (\tau_{10}, \ldots, \tau_{n\,n-1})^T \in \Theta_n$ and viceversa.

2.2 The 2D case with n = 2

The study of the properties of $\boldsymbol{\tau_n^*}$ is a fundamental step towards a deeper understanding of the geometrical acoustics model for TDOA-based localization. However, since its inherent complexity, the full description of the general case of $\boldsymbol{\tau_n^*}$ goes beyond the scope of this manuscript. In this section, we summarize the main results contained in Compagnoni and Notari (2014) and Compagnoni et al. (2014) on the minimal case of two dimensional source localization, with three synchronized and calibrated sensors.[1] We report this analysis because this is the minimal non trivial case of TDOA-based localization, the only one that has been

---

[1] In order to simplify the presentation, we consider only the case with the microphones in general position on the plane, i.e., they do not lie on a line. The interested reader can find the complete analysis for every scenario and the proofs in the original manuscripts.

Fig. 1 The image of τ_2 is the *gray* subset of the hexagon P_2 with *continuous* and *dashed* sides. In the *light gray region* E^- the map τ_2 is 1-to-1, while in the *medium gray region* $U_0 \cup U_1 \cup U_2$ the map τ_2 is 2-to-1. Let us observe that $U_0 \cup U_1 \cup U_2 \subset C^+ \cap P_2$. The continuous part of the boundary of the hexagon and the *blue ellipse* E, together with the vertices R^i, are in the image, and there τ_2 is 1-to-1. The points $T_i^{\pm}$ and the *dashed boundaries* do not belong to $\mathrm{Im}(\tau_2)$ (Color figure online)
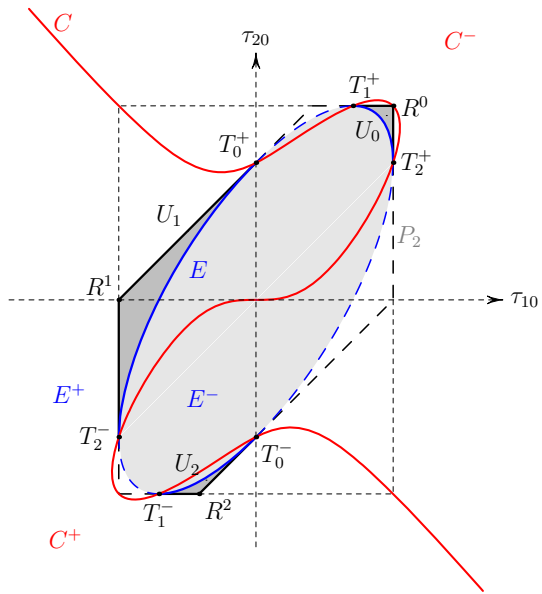
exhaustively studied and where one may observe some important features characterizing every localization model. We will return on this model in Sect. 4.

In the planar case, the set Θ_2 is a surface embedded into $\mathbb{R}^3$, being the image of the restriction of the TDOA map τ_2^* to $\mathbb{R}^2$ (by abuse of notation, we continue to name it τ_2^*). Actually, one can interpret τ_2^* as a (radical) parameterization of Θ_2. Moreover, it is well known that the three TDOAs are not independent, since they satisfy the zero-sum condition (ZSC) (Scheuing and Yang 2006). Indeed, the linear relation $\tau_{21}(\mathbf{x}) = \tau_{20}(\mathbf{x}) - \tau_{10}(\mathbf{x})$ holds for each $\mathbf{x} \in \mathbb{R}^2$. Geometrically speaking, this means that three noiseless TDOAs are constrained on the plane

$$V_2 = \left\{ \boldsymbol{\tau} \in \mathbb{R}^3 \mid \tau_{10} - \tau_{20} + \tau_{21} = 0 \right\} \subset \mathbb{R}^3 \tag{4}$$

and so $\Theta_2 \subseteq V_2$.

Because of the above linear relation, in the literature it is customary to work with a reference microphone, for example $\mathbf{m_0}$, and to consider only the two TDOAs $\tau_{10}(\mathbf{x})$, $\tau_{20}(\mathbf{x})$. Mathematically speaking, let us define the reduced TDOA map

$$\begin{aligned} \boldsymbol{\tau_2} : \mathbb{R}^2 &\longrightarrow \mathbb{R}^2 \\ \mathbf{x} &\longrightarrow (\tau_{10}(\mathbf{x}), \tau_{20}(\mathbf{x})) \end{aligned} \tag{5}$$

and let us consider the projection map $p_3 : \mathbb{R}^3 \to \mathbb{R}^2$ forgetting the third coordinate τ_{21} of the τ-space. Then, we have $\boldsymbol{\tau_2} = p_3 \circ \boldsymbol{\tau_2^*}$ and p_3 is a natural bijection between $\mathrm{Im}(\boldsymbol{\tau_2^*})$ and $\mathrm{Im}(\boldsymbol{\tau_2})$. Hence, one can investigate the properties of the noiseless TDOA model by studying the simpler map $\boldsymbol{\tau_2}$. For the sake of simplicity, in Fig. 1 we draw $\mathrm{Im}(\boldsymbol{\tau_2})$ for the configuration of the microphone at $\mathbf{m_0} = (0,0)^T$, $\mathbf{m_1} = (1,0)^T$ and $\mathbf{m_2} = (1,1)^T$. Symbols defined therein are introduced in the next few paragraphs. Figure 2 shows its relation with $\mathrm{Im}(\boldsymbol{\tau_2^*})$.

Fig. 2 The image of τ_2^* is the *green* subset of the hexagon $\mathcal{P}_2 \subset V_2$, while the image of τ_2 is the *red* subset of P_2. There is a 1-to-1 correspondence between $\mathrm{Im}(\tau_2^*)$ and $\mathrm{Im}(\tau_2)$ via the projection map p_3. In the *lightly shaded regions*, the TDOA maps are 1-to-1, while in the *more darkly shaded regions* the maps are 2-to-1 (Color figure online)
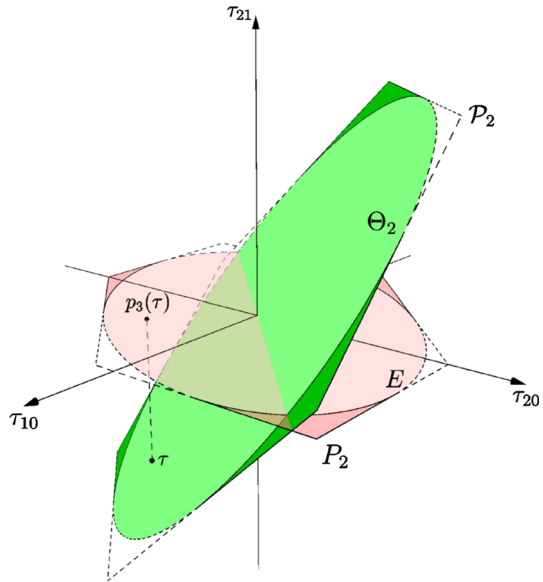
Let us define the displacement vectors $\mathbf{d_{ji}} = \mathbf{m_j} - \mathbf{m_i}$; their Euclidean norms $d_{ji} = \|\mathbf{d_{ji}}\|$, $i, j = 0, 1, 2$; the scalar $W = \det(\mathbf{d_{10}} \ \mathbf{d_{20}})$; and the matrix

$$\mathbf{H} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

First of all, $\mathrm{Im}(\tau_2)$ is contained into the hexagon P_2 defined by the triangle inequalities:

$$\begin{cases} -d_{10} \leq \tau_{10} \leq d_{10} \\ -d_{20} \leq \tau_{20} \leq d_{20} \\ -d_{21} \leq \tau_{20} - \tau_{10} \leq d_{21} \end{cases} \tag{6}$$

In particular, the vertices $R^0 = (d_{10}, d_{20})$, $R^1 = (-d_{10}, d_{21} - d_{10})$, $R^2 = (d_{21} - d_{20}, -d_{20})$ of P_2 correspond to the pairs of TDOAs associated to a source at $\mathbf{m_0}, \mathbf{m_1}, \mathbf{m_2}$, respectively. Then, by following the analysis contained in Sect. 6 of Compagnoni et al. (2014), for any $\tau = (\tau_{10}, \tau_{20}) \in \mathbb{R}^2$ we define the vectors

$$\mathbf{v}(\tau) = \mathbf{H}(\tau_{20}\mathbf{d_{10}} - \tau_{10}\mathbf{d_{20}}), \quad \mathbf{l_0}(\tau) = \mathbf{H}\frac{(d_{20}^2 - \tau_{20}^2)\mathbf{d_{10}} - (d_{10}^2 - \tau_{10}^2)\mathbf{d_{20}}}{2W} \tag{7}$$

and the polynomials

$$a(\tau) = \|\mathbf{v}(\tau)\|^2 - W^2, \quad b(\tau) = \mathbf{v}(\tau)^T \cdot \mathbf{l_0}(\tau), \quad c(\tau) = \|\mathbf{l_0}(\tau)\|^2. \tag{8}$$

$\mathrm{Im}(\tau_2)$ and the admissible source positions $\tau_2^{-1}(\tau)$ can be computed in terms of these polynomials.

We have the following facts:

- $a(\tau) = 0$ defines the unique ellipse E tangent to every facet of P_2. We name E^- and E^+ the interior and the exterior regions of E where $a(\tau) < 0$ and $a(\tau) > 0$, respectively;
- $b(\tau) = 0$ defines a cubic curve C. We name C^- and C^+ the regions where $b(\tau) < 0$ and $b(\tau) > 0$, respectively;

Fig. 3 The different localization regions and the curve $\tilde{E}$ in the x-plane. The microphones are the marked points $\mathbf{m_0} = (0, 0)$, $\mathbf{m_1} = (1, 0)$ and $\mathbf{m_2} = (1, 1)$. The *curve* $\tilde{E}$ separates the *light gray region* $\tilde{E}^-$, where the map τ_2 is 1–1 and it is possible to locate the source, and the *medium gray region* $\tilde{U}_0 \cup \tilde{U}_1 \cup \tilde{U}_2$, where τ_2 is 2–1 and the localization is not unique. On the *dashed lines* the localization is possible but very sensitive to the measurement noise
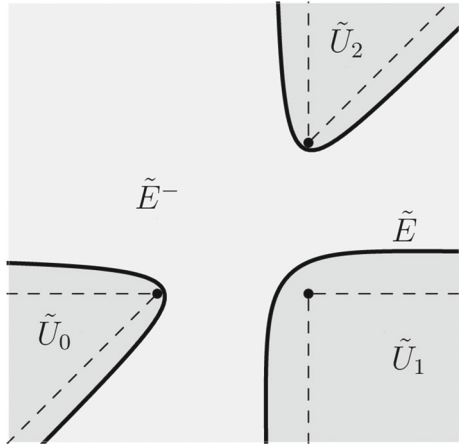
- $c(\tau)$ is a quartic non negative polynomial.

In Compagnoni et al. (2014) it has been proved that the image of τ_2 is the set

$$\text{Im}(\tau_2) = E^- \cup (C^+ \cap P_2) \cup R^0. \tag{9}$$

For each $\tau \in \text{Im}(\tau_2)$ we have at most two admissible source positions, whose coordinates are given by the formula

$$\mathbf{x}_\pm(\tau) = \mathbf{m_0} + \mathbf{l_0}(\tau) + \lambda_\pm(\tau)\mathbf{v}(\tau), \tag{10}$$

where $\lambda_\pm(\tau)$ are the solutions of the quadratic equation $a(\tau)\lambda^2 + 2b(\tau)\lambda + c(\tau) = 0$:

$$\lambda_\pm(\tau) = \frac{-b(\tau) \pm \sqrt{b(\tau)^2 - a(\tau)c(\tau)}}{a(\tau)}.$$

For $\tau \in (E^- \cup E \cup \partial P_2) \cap \text{Im}(\tau_2)$ we have to take only the $\mathbf{x}_+(\tau)$ solution and so we have uniqueness of localization. On the complementary set, that is the union of the three disjoint sets U_0, U_1 and U_2 depicted in medium gray in Fig. 1, the map τ_2 is 2-to-1 and there is an intrinsic ambiguity in the source position between the two solutions $\mathbf{x}_\pm(\tau)$.

For the sake of completeness, in Fig. 3 we depict the corresponding localization regions in the x-plane. Roughly speaking, we have the preimage of the interior of the ellipse $\tilde{E}^- = \tau_2^{-1}(E^-)$, where the TDOA map is 1-to-1 and the source localization is possible, and the preimages $\tilde{U}_i = \tau_2^{-1}(U_i)$, for $i = 0, 1, 2$, where the map is 2-to-1 and there is no way to uniquely locate the source. The region of transition is the *bifurcation curve* $\tilde{E} = \tau_2^{-1}(E)$, that is a quintic algebraic curve (Compagnoni and Notari 2014) consisting of three disjoint and unbounded arcs, one for each arc of E contained in $\text{Im}(\tau_2)$. As a point τ in one of the U_i gets close to E, the solution $\mathbf{x}_+(\tau)$ gets close to a point on $\tilde{E}$, while $\mathbf{x}_-(\tau)$ goes to infinity. The sets $\tilde{E}^-$, $\tilde{U}_0$, $\tilde{U}_1$, $\tilde{U}_2$ are open subsets of the x-plane, separated by the three arcs of $\tilde{E}$.

Finally, the union D of the six dashed half-lines outgoing from the receivers is the *degeneracy locus* of the TDOA map, where the rank of the Jacobian matrix of τ_2 drops. D is the preimage of the six segments in $\partial P_2 \cap \text{Im}(\tau_2)$. On D the two solutions $\mathbf{x}_\pm(\tau)$ are coincident, thus the TDOA map is 1-to-1. Furthermore, D divide each $\tilde{U}_i$ into two connected components and τ_2 is a bijection between each of them and the corresponding U_i.

2.3 The general case

As we said above, some of the properties we described in the minimal planar case are common to every TDOA-based localization model. In particular, we have the following proposition.

Proposition 1 *Let us take $n + 1$ sensors at $\mathbf{m_0}, \ldots, \mathbf{m_n}$ in $\mathbb{R}^3$, where $n \geq 2$. Then, Θ_n is a subset of the n-dimensional linear subspace $V_n \subset \mathbb{R}^q$ defined by equations*

$$- \tau_{i0} + \tau_{j0} - \tau_{ji} = 0, \quad 0 < i < j \leq n, \tag{11}$$

representing the ZSCs for all the microphone triplets containing $\mathbf{m_0}$.

Proof In a configuration with $n + 1$ microphones, the maximum number of independent TDOAs is equal to n. In particular, if we take $\mathbf{m_0}$ as the reference microphone, the n TDOAs $\{\tau_{10}(\mathbf{x}), \ldots, \tau_{n0}(\mathbf{x})\}$ are independent, while the others satisfy Eq. (11), as can be easily verified using definition (2). These are $q - n$ independent homogeneous linear equations, therefore they define an n-dimensional linear subspace V_n of the τ-space and Θ_n is a subset of V_n. $\qquad\square$

This means that, also in the general case, the set of feasible TDOAs is contained into an n-dimensional linear subspace of the TDOA space $\mathbb{R}^q$. This property stays at the basis of the denoising procedure that we describe in the next sections. However, Θ_n is strictly contained in V_n. Indeed, since Θ_n is the image of $\mathbb{R}^3$ through the almost everywhere smooth function τ_n, its dimension is equal to 3. As above, one can consider τ_n as a radical parameterization of Θ_n. This way the feasible set becomes a topological manifold (possibly with a boundary, as for Θ_2), that is almost everywhere differentiable. Moreover, we can reasonably conjecture that Θ_n can be described again in terms of algebraic equations and inequalities, hence it is a so called semialgebraic variety (Basu et al. 2006).

3 Interpretation of source localization in the TDOA space

The TDOA space formalism that we just introduced can be used to provide a geometric interpretation of TDOA-based source localization problem. As a matter of fact, other geometric interpretations have been proposed in the literature (Bestagini et al. 2013). However, working in the TDOA space also highlights that source localization can be solved as a TDOA denoising problem. In this section, we provide such interpretation of source localization in the TDOA space, starting from a commonly used statistical noise model.

3.1 Statistical noise model

In the presence of measurement errors, we must resort to statistical modeling. In this manuscript we assume the TDOAs associated to a source in $\mathbf{x}$ to be described by

$$\hat{\tau}_n^*(\mathbf{x}) = \tau_n^*(\mathbf{x}) + \varepsilon, \tag{12}$$

where $\varepsilon \sim N(\mathbf{0}, \mathbf{\Sigma})$ is an additive Gaussian noise. Also techniques that compute range differences from other measurements, such as energy, are prone to additive noise. In this latter case, in particular, the magnitude of the additive noise becomes relevant. Under the assumption in (12), the probability density function (p.d.f.) of the TDOA set is (Benesty and Huang 2004)

$$p(\hat{\tau}; \tau_n^*(\mathbf{x}), \mathbf{\Sigma}) = \frac{e^{-\frac{1}{2}(\hat{\tau} - \tau_n^*(\mathbf{x}))^T \mathbf{\Sigma}^{-1}(\hat{\tau} - \tau_n^*(\mathbf{x}))}}{\sqrt{(2\pi)^q |\mathbf{\Sigma}|}}. \tag{13}$$

Since the covariance matrix $\boldsymbol{\Sigma}$ is symmetric and positive defined, from a geometric standpoint (see, for example, Amari and Nagaoka 2000) the Fisher matrix $\boldsymbol{\Sigma}^{-1}$ defines a scalar product on $\mathbb{R}^q$:

$$\langle \mathbf{v_1}, \mathbf{v_2} \rangle_{\boldsymbol{\Sigma}^{-1}} = \mathbf{v_1}^T \boldsymbol{\Sigma}^{-1} \mathbf{v_2}, \quad \mathbf{v_1}, \mathbf{v_2} \in \mathbb{R}^q. \tag{14}$$

This way, the TDOA space turns out to be equipped with a Euclidean structure, whose distance is known in the statistical literature as the Mahalanobis distance:

$$\|\mathbf{v}\|_{\boldsymbol{\Sigma}^{-1}} = \sqrt{\mathbf{v}^T \boldsymbol{\Sigma}^{-1} \mathbf{v}}, \quad \mathbf{v} \in \mathbb{R}^q. \tag{15}$$

With this setting, the p.d.f. (13) can be rewritten as

$$p(\hat{\boldsymbol{\tau}}; \boldsymbol{\tau}_n^*(\mathbf{x}), \boldsymbol{\Sigma}) = \frac{e^{-\frac{1}{2}\|\hat{\boldsymbol{\tau}} - \boldsymbol{\tau}_n^*(\mathbf{x})\|_{\boldsymbol{\Sigma}^{-1}}^2}}{\sqrt{(2\pi)^q |\boldsymbol{\Sigma}|}}, \tag{16}$$

which depends only on the Mahalanobis distance between $\hat{\boldsymbol{\tau}}$ and $\boldsymbol{\tau}_n^*(\mathbf{x})$.

3.2 Source localization

The first application of the TDOA space and map was in the study of the TDOA-based source localization (see Compagnoni et al. 2014; Spencer 2007). As a matter of fact, the fundamental questions in localization problems can be readily formulated in terms of $\boldsymbol{\tau}_n^*$. In a noiseless scenario, the analysis of the existence and uniqueness of localization is equivalent to the study of the set Θ_n and the invertibility of $\boldsymbol{\tau}_n^*$. As we saw in Sect. 2.2 for the minimal planar case, for a given $\boldsymbol{\tau} \in \mathbb{R}^q$ there exists a unique source at position $\boldsymbol{\tau}_n^{*-1}(\boldsymbol{\tau})$ if, and only if, $\boldsymbol{\tau}$ is a point lying on a region of Θ_n where the TDOA map is 1-to-1.

In case of noisy measurements, the data errors force the point $\hat{\boldsymbol{\tau}}$ not to lie on Θ_n. Therefore, to localize the source one needs an estimation procedure. Most of the algorithms proposed in the literature rely on the optimization of a cost function, based on some criterion that can be either statistically motivated [e.g., Maximum-Likelihood estimation (Benesty and Huang 2004)] or not [e.g., linear least squares (Gillette and Silverman 2008; Smith and Abel 1987), squared range-differences-based least squares estimation (Beck et al. 2008), etc.]. The source position is thus found as the point $\bar{\mathbf{x}}$ that minimizes a suitable non-negative cost function $f(\hat{\boldsymbol{\tau}}, \mathbf{x})$, defined so that its value is zero in noiseless conditions, i.e.,

$$f\left(\boldsymbol{\tau}_n^*(\mathbf{x}), \mathbf{x}\right) = 0 \tag{17}$$

for every $\mathbf{x} \in \mathbb{R}^3$. In mathematical terms, source localization is therefore formulated as

$$\bar{\mathbf{x}} = \underset{\mathbf{x} \in \mathbb{R}^3}{\operatorname{argmin}} \, f\left(\hat{\boldsymbol{\tau}}, \mathbf{x}\right). \tag{18}$$

In the TDOA space, the estimated source position is associated to the feasible point $\bar{\boldsymbol{\tau}} = \boldsymbol{\tau}_n^*(\bar{\mathbf{x}}) \in \Theta_n$. Therefore, any localization algorithm maps a noisy TDOA vector $\hat{\boldsymbol{\tau}}$ onto a feasible TDOA vector $\bar{\boldsymbol{\tau}}$. It is worth noticing that different algorithms may produce different source estimates, corresponding to mappings onto likewise different feasible TDOA vectors. A special case occurs when the input TDOA vector is feasible, i.e., when $\hat{\boldsymbol{\tau}} \in \Theta_n$. In this case, there exists a point $\bar{\mathbf{x}}$ so that $\hat{\boldsymbol{\tau}} = \boldsymbol{\tau}_n^*(\bar{\mathbf{x}})$. Thus, in virtue of (17), we expect any algorithm to produce the same estimate $\bar{\mathbf{x}} = \boldsymbol{\tau}_n^{*-1}(\hat{\boldsymbol{\tau}})$.

Let us now focus on the Maximum-Likelihood (ML) estimator, which is known to be optimal in the statistical sense. For the Gaussian noise model described above, the ML

localization problem can be formulated as (Benesty and Huang 2004)

$$\bar{\mathbf{x}}_{\text{ML}} = \underset{\mathbf{x} \in \mathbb{R}^3}{\text{argmax}} \; \frac{e^{-\frac{1}{2}\|\hat{\boldsymbol{\tau}} - \boldsymbol{\tau}_n^*(\mathbf{x})\|_{\boldsymbol{\Sigma}^{-1}}^2}}{\sqrt{(2\pi)^q |\boldsymbol{\Sigma}|}}. \tag{19}$$

By defining $f_{\text{ML}}(\hat{\boldsymbol{\tau}}, \mathbf{x}) = \|\hat{\boldsymbol{\tau}} - \boldsymbol{\tau}_n^*(\mathbf{x})\|_{\boldsymbol{\Sigma}^{-1}}^2$, we have

$$\bar{\mathbf{x}}_{\text{ML}} = \underset{\mathbf{x} \in \mathbb{R}^3}{\text{argmin}} \; f_{\text{ML}}(\hat{\boldsymbol{\tau}}, \mathbf{x}). \tag{20}$$

In the TDOA space framework, the ML estimator has a neat geometric interpretation. Indeed, solving the ML problem is equivalent to finding the point $\bar{\boldsymbol{\tau}}_{\text{ML}} = \boldsymbol{\tau}_n^*(\bar{\mathbf{x}}_{\text{ML}}) \in \Theta_n$ at minimum Mahalanobis distance from $\hat{\boldsymbol{\tau}}$:

$$\bar{\boldsymbol{\tau}}_{\text{ML}} = \underset{\boldsymbol{\tau} \in \Theta_n}{\text{argmin}} \; \|\hat{\boldsymbol{\tau}} - \boldsymbol{\tau}\|_{\boldsymbol{\Sigma}^{-1}}^2. \tag{21}$$

It trivially follows that the source position $\bar{\mathbf{x}}$ estimated by means of a generic localization algorithm leads to a point $\bar{\boldsymbol{\tau}} = \boldsymbol{\tau}_n^*(\bar{\mathbf{x}}) \in \Theta_n$ such that

$$\|\hat{\boldsymbol{\tau}} - \bar{\boldsymbol{\tau}}\|_{\boldsymbol{\Sigma}^{-1}} \geq \|\hat{\boldsymbol{\tau}} - \bar{\boldsymbol{\tau}}_{\text{ML}}\|_{\boldsymbol{\Sigma}^{-1}}. \tag{22}$$

In other words, the distance $\|\hat{\boldsymbol{\tau}} - \bar{\boldsymbol{\tau}}\|_{\boldsymbol{\Sigma}^{-1}}$ is bounded from below by the one obtained through ML estimation.

In the light of the above considerations, we can interpret source localization in the TDOA space as a two-step procedure:

1. mapping of the noisy TDOAs $\hat{\boldsymbol{\tau}}$ onto the corresponding feasible vector $\bar{\boldsymbol{\tau}} \in \Theta_n$, accordingly with the chosen optimization criterion;
2. recovering of the estimated source position as $\bar{\mathbf{x}} = \boldsymbol{\tau}_n^{*-1}(\bar{\boldsymbol{\tau}})$.

Note that step 1 represents a TDOA denoising operation. Moreover, once this step has been accomplished, step 2 is straightforward when the mapping $\boldsymbol{\tau}_n^*$ is 1-to-1, that is the standard case for a sufficiently large number n of sensors in general positions. From this perspective, source localization can be considered as a TDOA denoising problem.

4 Denoising of TDOAs

As discussed in Sect. 3, source localization can be solved in the TDOA space as a denoising problem. Given a noisy TDOA vector $\hat{\boldsymbol{\tau}}$, this corresponds to finding a feasible (i.e., denoised) TDOA vector belonging to Θ_n according to some criteria that depends on the chosen cost function $f(\hat{\boldsymbol{\tau}}, \mathbf{x})$. If we consider the Maximum-Likelihood formulation of the problem, the criterion is readily formulated as in (21). Therefore, statistically speaking, the best achievable denoising corresponds to finding the feasible TDOA vector at the minimum Mahalanobis distance from the noisy one.

Despite the error function $\|\hat{\boldsymbol{\tau}} - \boldsymbol{\tau}\|_{\boldsymbol{\Sigma}^{-1}}^2$ could suggest a standard weighted linear least-squares solution (Teunissen 2000), the problem in (21) can not be solved in a linear fashion. Indeed, the search space is Θ_n, which turns (21) in a difficult, in general non convex, optimization problem. However, in the following we investigate the possibility of relaxing the problem in (21), leveraging on the fact that Θ_n is contained in the linear subspace V_n of the TDOA space.

4.1 From the complete to the relaxed denoising problem

We claimed above that solving the ideal denoising problem (21) is a non trivial task. However, by working in the TDOA space we can subdivide such problem in two distinct steps. Let us consider the orthogonal projection $\mathcal{P}(\hat{\boldsymbol{\tau}}; \boldsymbol{\Sigma})$ of $\hat{\boldsymbol{\tau}} \in \mathbb{R}^q$ on V_n, with respect to the scalar product $\langle \, , \, \rangle_{\boldsymbol{\Sigma}^{-1}}$. From the general properties of this map, we have the following equivalences:

$$
\begin{aligned}
\bar{\boldsymbol{\tau}}_{\mathrm{ML}} &= \underset{\boldsymbol{\tau} \in \Theta_n}{\operatorname{argmin}} \, \|\hat{\boldsymbol{\tau}} - \boldsymbol{\tau}\|_{\boldsymbol{\Sigma}^{-1}}^2 \\
&= \underset{\boldsymbol{\tau} \in \Theta_n}{\operatorname{argmin}} \|(\hat{\boldsymbol{\tau}} - \mathcal{P}(\hat{\boldsymbol{\tau}}; \boldsymbol{\Sigma})) + (P(\hat{\boldsymbol{\tau}}; \boldsymbol{\Sigma}) - \boldsymbol{\tau})\|_{\boldsymbol{\Sigma}^{-1}}^2 \\
&= \underset{\boldsymbol{\tau} \in \Theta_n}{\operatorname{argmin}} \left(\|\hat{\boldsymbol{\tau}} - \mathcal{P}(\hat{\boldsymbol{\tau}}; \boldsymbol{\Sigma})\|_{\boldsymbol{\Sigma}^{-1}}^2 + \|\mathcal{P}(\hat{\boldsymbol{\tau}}; \boldsymbol{\Sigma}) - \boldsymbol{\tau}\|_{\boldsymbol{\Sigma}^{-1}}^2 \right) \\
&= \underset{\boldsymbol{\tau} \in \Theta_n}{\operatorname{argmin}} \, \|\mathcal{P}(\hat{\boldsymbol{\tau}}; \boldsymbol{\Sigma}) - \boldsymbol{\tau}\|_{\boldsymbol{\Sigma}^{-1}}^2,
\end{aligned}
$$

where in the third equality we used the fact that $\hat{\boldsymbol{\tau}} - \mathcal{P}(\hat{\boldsymbol{\tau}}; \boldsymbol{\Sigma})$ and $\mathcal{P}(\hat{\boldsymbol{\tau}}; \boldsymbol{\Sigma}) - \boldsymbol{\tau}$ are orthogonal each other, thus

$$
\langle \hat{\boldsymbol{\tau}} - \mathcal{P}(\hat{\boldsymbol{\tau}}; \boldsymbol{\Sigma}), \mathcal{P}(\hat{\boldsymbol{\tau}}; \boldsymbol{\Sigma}) - \bar{\boldsymbol{\tau}}_{\mathrm{ML}} \rangle_{\boldsymbol{\Sigma}^{-1}} = 0.
$$

This means that the ML estimation gives exactly the same results if we start from the original data $\hat{\boldsymbol{\tau}}$ or the projected one $\mathcal{P}(\hat{\boldsymbol{\tau}}; \boldsymbol{\Sigma})$.

We explicitly show this fact in Fig. 4, for the case of planar localization with $n = 2$. In this case one has two different situations, exemplified for the measurement points $\hat{\boldsymbol{\tau}}_1$ and $\hat{\boldsymbol{\tau}}_2$, respectively. In the first case, the orthogonal projection $\mathcal{P}(\hat{\boldsymbol{\tau}}_1; \boldsymbol{\Sigma}) \in \Theta_2$, thus it coincides with its ML estimate $\bar{\boldsymbol{\tau}}_{\mathrm{ML},1}$. Differently, for the second point one has $\mathcal{P}(\hat{\boldsymbol{\tau}}_2; \boldsymbol{\Sigma}) \notin \Theta_2$. In this case, finding the corresponding ML solution $\bar{\boldsymbol{\tau}}_{\mathrm{ML},2}$ corresponds to finding the closest point to $\mathcal{P}(\hat{\boldsymbol{\tau}}_2; \boldsymbol{\Sigma})$ in Θ_2, which implies solving a complicated minimization problem in V_2. Indeed, as we explained in Sect. 2.2, the feasible set Θ_2 has a non trivial structure. In particular, its boundary is the union of six segments and three arcs of ellipse. If we define $\bar{\boldsymbol{\tau}}_{\mathrm{ML},2}$ as the closest point to $\hat{\boldsymbol{\tau}}_2$ lying on one of these sets, then it is necessary to develop an ad hoc algorithm for finding it. Another complication is that Θ_2 is not a closed set. This implies that if $\bar{\boldsymbol{\tau}}_{\mathrm{ML},2}$ lies on the ellipse, then it does not correspond to a true source position because it is not part of Θ_2. Conversely we should consider $\bar{\boldsymbol{\tau}}_{\mathrm{ML},2}$ as the TDOAs associated to a source placed at infinity.

Thanks to the results and discussion contained in Sect. 2.3, we can generalize the previous analysis. The denoising problem (21) can be subdivided into two subproblems.

1. The easiest part is the projection on the linear subspace V_n. We can call it the relaxed denoising problem

$$
\mathcal{P}(\hat{\boldsymbol{\tau}}; \boldsymbol{\Sigma}) = \underset{\boldsymbol{\tau} \in V_n}{\operatorname{argmin}} \, \|\hat{\boldsymbol{\tau}} - \boldsymbol{\tau}\|_{\boldsymbol{\Sigma}^{-1}}^2, \tag{23}
$$

   in comparison to the complete denoising problem (21). Being the search set in (23) a linear subspace, the problem admits a closed solution.

2. The hardest part is the projection of $\mathcal{P}(\hat{\boldsymbol{\tau}}; \boldsymbol{\Sigma})$ onto Θ_n :

$$
\bar{\boldsymbol{\tau}}_{\mathrm{ML}} = \underset{\boldsymbol{\tau} \in \Theta_n}{\operatorname{argmin}} \, \|\mathcal{P}(\hat{\boldsymbol{\tau}}; \boldsymbol{\Sigma}) - \boldsymbol{\tau}\|_{\boldsymbol{\Sigma}^{-1}}^2,
$$

   From the discussion contained in Sect. 2.3, the difficulties are twofold: we have not the analytic description of Θ_n and, in any case, the feasible set is a complicated three dimensional semialgebraic variety embedded in the n- dimensional linear subspace V_n.
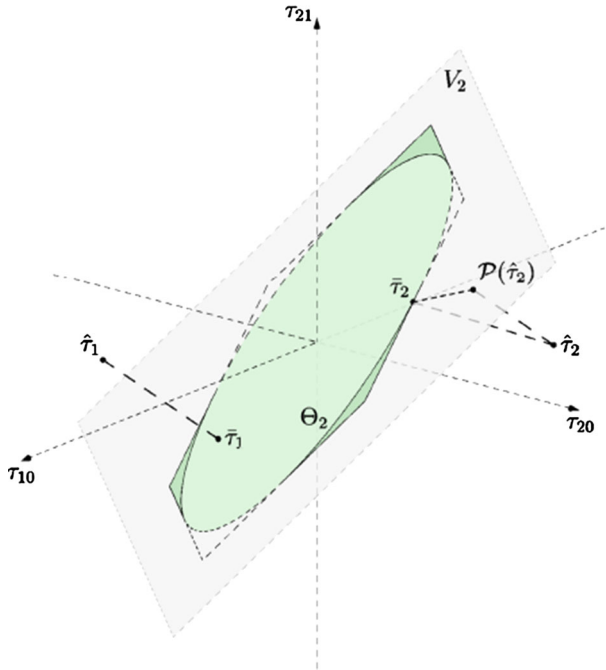
Fig. 4 Denoising in the TDOA space, for the case of planar localization with $n = 2$. A generic set of noisy TDOAs $\hat{\boldsymbol{\tau}}$ do not lie on Θ_2. The ML estimation finds the closest point $\bar{\boldsymbol{\tau}}_{\text{ML}} \in \Theta_2$ to $\hat{\boldsymbol{\tau}}$. The solution of the relaxed denoisig problem is $\mathcal{P}(\hat{\boldsymbol{\tau}}, \boldsymbol{\Sigma}) \in V_2$. For the point $\hat{\boldsymbol{\tau}}_1$ it lies on Θ_2, then the projection $\mathcal{P}(\hat{\boldsymbol{\tau}}_1, \boldsymbol{\Sigma})$ coincides with the ML solution $\bar{\boldsymbol{\tau}}_1$. For the point $\hat{\boldsymbol{\tau}}_2$ this is not true, thus $\mathcal{P}(\hat{\boldsymbol{\tau}}_1, \boldsymbol{\Sigma}) \neq \bar{\boldsymbol{\tau}}_2$. In any case, the projection gives a better estimate of the noiseless measurements $\boldsymbol{\tau}$, closest to $\bar{\boldsymbol{\tau}}_{\text{ML}}$ with respect to the original data point $\hat{\boldsymbol{\tau}}$

The previous observations are the original starting points for our interpretation and for the statistical justification of the relaxed denoising procedure. Indeed, observation 2 confirms the well known fact that solving ML estimation is a too difficult task, due to the non-linearity and non-convexity of the feasible set Θ_n. However, observation 1 states that ML is composed by two parts with different difficulties. We can easily envision that the projection of the measured TDOAs onto the linear subspace V_n leads to improvements in terms of localization accuracy, and this will be confirmed analytically in the forthcoming sections.

4.2 The relaxed denoising algorithm and its statistical analysis

We can now proceed with a rigorous formulation of the intuition discussed above and with a precise definition of the relaxed denoising algorithm.

4.2.1 The projection as a sufficient statistic

Theorem 1 *The orthogonal projection $\mathcal{P}(\hat{\boldsymbol{\tau}}; \boldsymbol{\Sigma})$ of $\hat{\boldsymbol{\tau}} \in \mathbb{R}^q$ on V_n, with respect to the scalar product $\langle \, , \, \rangle_{\boldsymbol{\Sigma}^{-1}}$, is a sufficient statistic for the underlying parameter $\mathbf{x}$.*

Proof Let us start from

$$\|\hat{\boldsymbol{\tau}} - \boldsymbol{\tau}_n^*(\mathbf{x})\|_{\boldsymbol{\Sigma}^{-1}}^2 = \|\hat{\boldsymbol{\tau}} - \mathcal{P}(\hat{\boldsymbol{\tau}}; \boldsymbol{\Sigma})\|_{\boldsymbol{\Sigma}^{-1}}^2 + \|\mathcal{P}(\hat{\boldsymbol{\tau}}; \boldsymbol{\Sigma}) - \boldsymbol{\tau}_n^*(\mathbf{x})\|_{\boldsymbol{\Sigma}^{-1}}^2 .$$

Therefore, the probability density function (13) can be rewritten as

$$p(\hat{\boldsymbol{\tau}}; \boldsymbol{\tau}_n^*(\mathbf{x}), \boldsymbol{\Sigma}) = \frac{e^{-\frac{1}{2}\|\hat{\boldsymbol{\tau}} - \mathcal{P}(\hat{\boldsymbol{\tau}}; \boldsymbol{\Sigma})\|_{\boldsymbol{\Sigma}^{-1}}^2} \; e^{-\frac{1}{2}\|\mathcal{P}(\hat{\boldsymbol{\tau}}; \boldsymbol{\Sigma}) - \boldsymbol{\tau}_n^*(\mathbf{x})\|_{\boldsymbol{\Sigma}^{-1}}^2}}{\sqrt{(2\pi)^q |\boldsymbol{\Sigma}|}}.$$

Then, the proof follows as a consequence of the Fisher–Neyman factorization theorem (Lehmann and Casella 1998). □

Theorem 1 states that the component of $\hat{\boldsymbol{\tau}}$ orthogonal to the linear subspace V_n does not carry information on the source location and that we can remove it without corrupting the data.

4.2.2 The algorithm

For any $\hat{\boldsymbol{\tau}} \in \mathbb{R}^q$, the projection $\mathcal{P}(\hat{\boldsymbol{\tau}}; \boldsymbol{\Sigma})$ can be computed in closed form in two steps:

I. If we group all the Eq. (11), we end up with the homogeneous equation system

$$\mathbf{C}\boldsymbol{\tau} = \mathbf{0}, \tag{24}$$

where $\mathbf{C}$ is the $(q - n, q)$ matrix

$$\mathbf{C} = \begin{pmatrix} -1 & 1 & 0 & \cdots & 0 & 0 & -1 & 0 & \cdots & 0 \\ -1 & 0 & 1 & \cdots & 0 & 0 & 0 & -1 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & -1 & 1 & 0 & 0 & \cdots & -1 \end{pmatrix}.$$

The solution of this linear system is $V_n = \ker(\mathbf{C})$. In particular, we can find an orthonormal basis $\{\mathbf{v_1}, \ldots, \mathbf{v_n}\}_{\boldsymbol{\Sigma}^{-1}}$ of V_n, if necessary by using Gram–Schmidt algorithm defined with respect to the scalar product (14).

II. The projection map $\mathcal{P}$ is defined on $\hat{\boldsymbol{\tau}} \in \mathbb{R}^q$ as

$$\mathcal{P}(\hat{\boldsymbol{\tau}}; \boldsymbol{\Sigma}) = \langle \hat{\boldsymbol{\tau}}, \mathbf{v_1} \rangle_{\boldsymbol{\Sigma}^{-1}} \mathbf{v_1} + \cdots + \langle \hat{\boldsymbol{\tau}}, \mathbf{v_n} \rangle_{\boldsymbol{\Sigma}^{-1}} \mathbf{v_n}. \tag{25}$$

Let $\mathbf{e_{ji}}$, $n \geq j > i \geq 1$ be the vectors in the standard basis $\mathcal{B}_q$ of $\mathbb{R}^q$. With respect to $\mathcal{B}_q$, the projection is represented by the (q, q) matrix

$$\mathbf{P} = \begin{bmatrix} \mathcal{P}(\mathbf{e_{10}}; \boldsymbol{\Sigma}) & \cdots & \mathcal{P}(\mathbf{e_{n-1\,n}}; \boldsymbol{\Sigma}) \end{bmatrix}. \tag{26}$$

Consequently, the set of denoised TDOAs is obtained as

$$\mathcal{P}(\hat{\boldsymbol{\tau}}; \boldsymbol{\Sigma}) = \mathbf{P}\,\hat{\boldsymbol{\tau}}. \tag{27}$$

4.2.3 The analysis of the noise reduction

We can now compute the noise reduction on the TDOAs due to the relaxed denoising procedure defined above. Preliminarily, we prove the following Lemma.

Lemma 1 *The covariance matrix $\boldsymbol{\Sigma}$ defines a Euclidean structure on $\mathbb{R}^q$ and the matrix $\mathbf{P}^T$ represents an orthogonal projection with respect to the scalar product $\langle \, , \, \rangle_{\boldsymbol{\Sigma}}$.*

Proof By construction $\boldsymbol{\Sigma}$ is symmetric and positive defined, therefore it defines a scalar product on $\mathbb{R}^q$. To prove that $\mathbf{P}^T$ represents an orthogonal projection with respect to this Euclidean structure, we have to show that $\mathbf{P}^T \mathbf{P}^T = \mathbf{P}^T$ and $\boldsymbol{\Sigma}\mathbf{P}^T = \mathbf{P}\boldsymbol{\Sigma}$. On the other hand, we know that the matrix $\mathbf{P}$ represents an orthogonal projection with respect to $\langle \, , \, \rangle_{\boldsymbol{\Sigma}^{-1}}$, hence $\mathbf{P}\mathbf{P} = \mathbf{P}$ and $\boldsymbol{\Sigma}^{-1}\mathbf{P} = \mathbf{P}^T \boldsymbol{\Sigma}^{-1}$. By using these identities, we have:

- $\mathbf{P}^T \mathbf{P}^T = (\mathbf{P}\mathbf{P})^T = \mathbf{P}^T$,
- $\mathbf{P}\boldsymbol{\Sigma} = \boldsymbol{\Sigma}\,\boldsymbol{\Sigma}^{-1}\mathbf{P}\boldsymbol{\Sigma} = \boldsymbol{\Sigma}\mathbf{P}^T\boldsymbol{\Sigma}^{-1}\boldsymbol{\Sigma} = \boldsymbol{\Sigma}\mathbf{P}^T$,

which completes the proof. □

A consequence of Lemma 1 is that $\|\mathbf{v}\|_{\boldsymbol{\Sigma}}^2 \geq \|\mathbf{P}^T\mathbf{v}\|_{\boldsymbol{\Sigma}}^2$ for any $\mathbf{v} \in \mathbb{R}^q$. This is useful for proving the next Theorem.

Theorem 2 *Let $\boldsymbol{\Sigma}$ be the covariance matrix of $\hat{\boldsymbol{\tau}}$. Then:*

1. *the covariance matrix of $\mathcal{P}(\hat{\boldsymbol{\tau}}; \boldsymbol{\Sigma})$ is $\boldsymbol{\Sigma}' = \mathbf{P}\boldsymbol{\Sigma}\mathbf{P}^T$;*
2. *we have $\boldsymbol{\Sigma} \succeq \boldsymbol{\Sigma}'$, i.e., $\boldsymbol{\Sigma} - \boldsymbol{\Sigma}'$ is positive semidefinite.*

Proof the first claim follows from the general transformation rule of the covariance matrix under linear mapping of $\hat{\boldsymbol{\tau}}$. Using this fact and Lemma 1, we have that $\mathbf{v}^T(\boldsymbol{\Sigma} - \boldsymbol{\Sigma}')\mathbf{v} = \|\mathbf{v}\|_{\boldsymbol{\Sigma}}^2 - \|\mathbf{P}^T\mathbf{v}\|_{\boldsymbol{\Sigma}}^2 \geq 0$ for any $\mathbf{v} \in \mathbb{R}^q$, which completes the proof. □

Theorem 2 states that the relaxed denoising procedure always reduces the noise on the TDOA dataset and it gives the way to quantify such reduction.

4.3 Relation to state-of-the-art algorithms

As already mentioned in the Introduction, the algorithm presented to solve the relaxed denoising problem is a different interpretation of other methods proposed for reducing the noise on TDOAs, exploiting data redundancy. In particular, authors in Schmidt (1996) and So et al. (2008) use the constraints in (11) to relate the full set of q TDOAs $\boldsymbol{\tau}$ to n nonredundant TDOAs referred to a sensor. Selecting the first sensor as the reference one, the nonredundant set is $\boldsymbol{\tau}_{\text{NR}} = (\tau_{10}, \tau_{20}, \ldots, \tau_{n0})^T$. The linear relation between the two sets is given by

$$\boldsymbol{\tau} = \mathbf{G}\boldsymbol{\tau}_{\text{NR}}, \quad \text{with } \mathbf{G} = \begin{bmatrix} \mathbf{I}_n \\ \mathbf{Y} \end{bmatrix} \text{ where } \mathbf{Y} = \begin{pmatrix} -1 & 1 & 0 & \cdots & 0 & 0 \\ -1 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & -1 & 1 \end{pmatrix},$$

and $\mathbf{I}_n$ is the identity matrix of order n. Given the measured TDOAs $\hat{\boldsymbol{\tau}}$, in Schmidt (1996) the nonredundant set is estimated in the least squares sense as

$$\hat{\boldsymbol{\tau}}_{\text{NR}} = (\mathbf{G}^T\mathbf{G})^{-1}\mathbf{G}^T\hat{\boldsymbol{\tau}}.$$

In So et al. (2008) this result is generalized accounting for the covariance structure of noise. To this purpose, the noise covariance matrix $\boldsymbol{\Sigma}$ is introduced in the weighted least squares solution

$$\hat{\boldsymbol{\tau}}_{\text{NR}} = (\mathbf{G}^T\boldsymbol{\Sigma}^{-1}\mathbf{G})^{-1}\mathbf{G}^T\boldsymbol{\Sigma}^{-1}\hat{\boldsymbol{\tau}}. \tag{28}$$

Note that the two procedures coincide if $\boldsymbol{\Sigma} = \sigma^2\mathbf{I}$.

The nonredundant TDOAs $\hat{\boldsymbol{\tau}}_{\text{NR}}$ can be considered as the denoised measurements corresponding to the original set $\hat{\boldsymbol{\tau}}$. As a matter of fact, $\hat{\boldsymbol{\tau}}_{\text{NR}}$, computed as in (28), coincides with the first n components of the orthogonal projection $\mathcal{P}(\hat{\boldsymbol{\tau}}; \boldsymbol{\Sigma})$. At this respect, we can be more precise.

Proposition 2 *Given the Euclidean structure on $\mathbb{R}^q$ defined by $\boldsymbol{\Sigma}^{-1}$, the matrix $\mathbf{P}$ that represents the projection map $\mathcal{P}$, with respect to the canonical basis $\mathcal{B}_q$, is*

$$\mathbf{P} = \mathbf{G}(\mathbf{G}^T\boldsymbol{\Sigma}^{-1}\mathbf{G})^{-1}\mathbf{G}^T\boldsymbol{\Sigma}^{-1}.$$

It follows that $\hat{\boldsymbol{\tau}}_{\text{NR}}$ is a sufficient statistics for $\hat{\boldsymbol{\tau}}$.

Proof First of all, we show that $\text{Im}(\mathbf{P}) = V_n$.[2] We begin by observing that the matrix $\mathbf{G}^T \boldsymbol{\Sigma}^{-1} \mathbf{G}$ is invertible. Indeed, its kernel is given by the vectors $\mathbf{x} \in \mathbb{R}^n$ such that $\mathbf{G}^T \boldsymbol{\Sigma}^{-1} \mathbf{G} \mathbf{x} = \mathbf{0}$. In this case, we have

$$\mathbf{x}^T \mathbf{G}^T \boldsymbol{\Sigma}^{-1} \mathbf{G} \mathbf{x} = \|\mathbf{G} \mathbf{x}\|_{\boldsymbol{\Sigma}^{-1}}^2 = 0,$$

hence $\mathbf{G} \mathbf{x} = \mathbf{0}$. Being $\text{rank}(\mathbf{G}) = n$, it follows that $\ker(\mathbf{G}^T \boldsymbol{\Sigma}^{-1} \mathbf{G}) = \ker(\mathbf{G}) = \{\mathbf{0}\}$.

From the previous point, we have

$$\text{rank}\left((\mathbf{G}^T \boldsymbol{\Sigma}^{-1} \mathbf{G})^{-1} \mathbf{G}^T \boldsymbol{\Sigma}^{-1}\right) = \text{rank}(\mathbf{G}^T) = n$$

and so $\text{Im}((\mathbf{G}^T \boldsymbol{\Sigma}^{-1} \mathbf{G})^{-1} \mathbf{G}^T \boldsymbol{\Sigma}^{-1}) = \mathbb{R}^n$. It follows that $\text{Im}(\mathbf{P}) = \text{Im}(\mathbf{G}) = V_n$.

Now, it is straightforward to check that $\mathbf{P}^2 = \mathbf{P}$ and $\boldsymbol{\Sigma}^{-1} \mathbf{P} = \mathbf{P}^T \boldsymbol{\Sigma}^{-1}$, thus $\mathbf{P}$ represents the projection map on $\text{Im}(\mathbf{P}) = V_n$.

The last statement follows from the 1-to-1 relation $\mathcal{P}(\hat{\boldsymbol{\tau}}; \boldsymbol{\Sigma}) = \mathbf{G}\hat{\boldsymbol{\tau}}_{\text{NR}}$. $\qquad\square$

In the light of the above considerations, all our theorems and simulative results can be considered as further validations of state-of-the-art works in Schmidt (1996) and So et al. (2008). In the following sections we will prove, both analytically and numerically, that any source localization algorithm benefits from denoised TDOAs. Moreover, we will see that this holds also when the full TDOA set is not completely available.

5 Impact on source localization

To this point we have shown how it is possible to reduce the noise on a TDOA set $\hat{\boldsymbol{\tau}}$, thus approaching $\bar{\boldsymbol{\tau}}_{\text{ML}}$. The goal of this section is to prove that source location estimate benefits from the use of a denoised TDOA set. To this purpose, we first provide a solid theoretical analysis that demonstrates our claim. Then, we further investigate the effect on source localization by means of an extensive simulative campaign.

5.1 Impact on sub-optimal localization algorithms

The present analysis is mainly based on the results in Sect. 4.2.3. In great generality, let us consider a given localization algorithm based on the minimization of a certain cost function $f(\mathbf{x}, \mathbf{c})$, where $\mathbf{c}$ are the input TDOA data. Theorem 2 allows us to compare the accuracy of $\bar{\mathbf{x}} = \arg\min f(\mathbf{x}, \hat{\boldsymbol{\tau}})$ and $\bar{\mathbf{x}}' = \arg\min f(\mathbf{x}, \mathcal{P}(\hat{\boldsymbol{\tau}}; \boldsymbol{\Sigma}))$. Indeed, the first order error propagation formula given in Compagnoni et al. (2012) relates the covariance matrices $\boldsymbol{\Sigma}$, $\boldsymbol{\Sigma}'$ to $\boldsymbol{\Sigma}_{\bar{\mathbf{x}}}$, $\boldsymbol{\Sigma}'_{\bar{\mathbf{x}}}$, respectively:

$$\boldsymbol{\Sigma}_{\bar{\mathbf{x}}} = \mathbf{A}(\mathbf{x}) \boldsymbol{\Sigma} \mathbf{A}(\mathbf{x})^T \quad \text{and} \quad \boldsymbol{\Sigma}'_{\bar{\mathbf{x}}} = \mathbf{A}(\mathbf{x}) \boldsymbol{\Sigma}' \mathbf{A}(\mathbf{x})^T, \tag{29}$$

where $\mathbf{A}(\mathbf{x})$ is defined by equation (26) from Compagnoni et al. (2012). Then, by easily adapting the proof of Theorem 2, we end up with

Corollary 1 *at first order approximation,* $\boldsymbol{\Sigma}_{\bar{\mathbf{x}}} \succeq \boldsymbol{\Sigma}'_{\bar{\mathbf{x}}}$.

We remark that Corollary 1 is valid for every choice of the cost function $f(\mathbf{x}, \mathbf{c})$. In particular, it includes the special cases of $f(\mathbf{x}, \mathbf{c})$ explicitly depending only on n TDOAs, the ones relative to a reference microphone. In the following paragraph we will report concrete examples of these facts.

---

[2] With slight abuse of notation, in the proof we identify a matrix with the associated linear map defined through matrix multiplication.

5.2 Numerical examples

We now analyze the effect of relaxed denoising[3] on simulated data. Some results were reported in Schmidt (1996) and So et al. (2008). In particular, in So et al. (2008) authors proved that the denoising procedure can be used to proficiently exploit the redundancy of the full set of TDOAs measured among all microphone pairs. Indeed, they numerically verified that ML source localization using the denoised TDOA set (denoted as optimum nonredundant set in So et al. 2008) meets the Cramer Rao Lower Bound (CRLB) (Benesty and Huang 2004) computed considering the full set of measurements. Note that the analytical proof of this fact is a direct consequence of Theorem 1.

Here we aim at providing a more comprehensive analysis, testing the effect of denoising on several source localization algorithms. In particular, we are interested in sub-optimal algorithms admitting an exact solution, i.e., for which the global optimum can be computed with no approximations. This feature is desirable in practice, as it prevents the risk of getting trapped in local minima, at the expense of obtaining a solution that does not meet the CRLB. To this end, we consider the following algorithms:

- *LS* unconstrained linear least squares estimator (Smith and Abel 1987), admitting a closed-form solution. It uses the nonredundant TDOA set, measured considering a reference sensor;
- *SRD-LS* squared-range-difference least squares estimator (Beck et al. 2008), a constrained version of the LS one, whose exact solution is computable efficiently. It is based on the nonredundant TDOA set;
- *GS* Gillette–Silverman algorithm (Gillette and Silverman 2008), an extension of the LS algorithm accommodating the case of multiple reference sensors. It can be applied to the full TDOA set.

Denoising was tested considering TDOAs corrupted by both Gaussian and non-Gaussian noise. The first case constitutes an ideal scenario, in which the noise model exactly meets the theoretical assumptions used for deriving the denoising procedure. The second case represents a non-ideal condition, useful to assess to what extent denoising is applicable when the assumptions are not obeyed.

5.2.1 TDOAs corrupted by Gaussian noise

We simulated a compact cross array composed by seven microphones in positions $(0, 0, 0)^T$, $(\pm 0.5, 0, 0)^T$, $(0, \pm 0.5, 0)^T$, $(0, 0, \pm 0.5)^T$ m. More than 500 sources are homogeneously distributed on a sphere centered at $(0, 0, 0)^T$, whose radius d ranges from 0.5 m to 2.5 m. The simulation setup is sketched in Fig. 5. For each source position, we computed the full set of $q = 21$ theoretical TDOAs τ. We corrupted the vectors τ with $I = 5000$ realizations of i.i.d. zero-mean Gaussian noise with standard deviation σ, leading to the noisy TDOAs $\hat{\tau}_i$. The covariance matrix thus resulted in $\Sigma = \sigma^2 \mathbf{I}_q$, where $\mathbf{I}_q$ is the identity matrix of order q. Monte-Carlo simulations were carried out considering the range $\sigma \in [0.5\,\text{cm}, 5\,\text{cm}]$. The corresponding denoised TDOAs $\mathcal{P}(\hat{\tau}_i; \Sigma)$ were computed using (27).

As a preliminary test, we computed the mean $\mu_{\tilde{\epsilon}}$ and the standard deviation $\sigma_{\tilde{\epsilon}}$ of the residual error $\epsilon_i = \tilde{\tau}_i - \tau$ left on denoised TDOAs. We first observed that $\mu_{\tilde{\epsilon}}$ is always negligible compared to $\sigma_{\tilde{\epsilon}}$, meaning that the denoising procedure does not introduce relevant

---

[3] For the sake of compactness, throughout this paragraph we will use the word *denoising* referring to *relaxed denoising*.
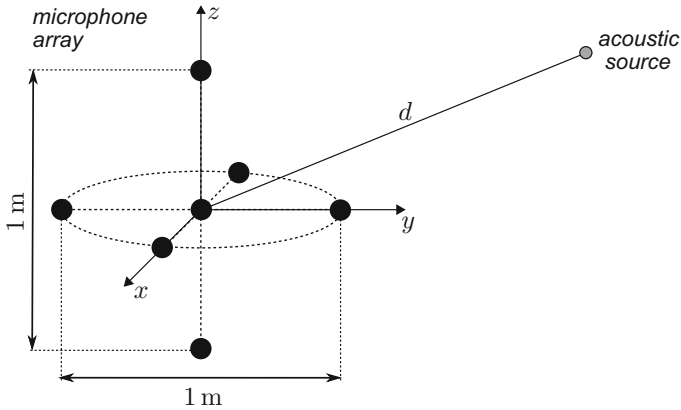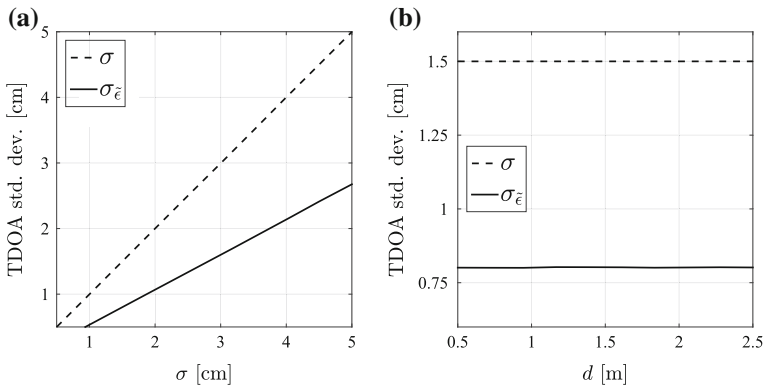
Fig. 5 Simulation setup

Fig. 6 TDOA residual error before and after denoising, as a function of: injected noise σ (**a**); source distance d (**b**)

bias on TDOAs. Figure 6a shows the standard deviation of denoised TDOAs $\sigma_{\tilde{\epsilon}}$ as a function of σ, averaged for sources located at a fixed distance $d = 1.5$ m. Similarly, Fig. 6b shows $\sigma_{\tilde{\epsilon}}$ for different distance values, fixing $\sigma = 1.5$ cm. In both the cases, we notice that $\sigma_{\tilde{\epsilon}}$ (solid line) is always significantly below the value of the standard deviation of the injected noise σ (dashed line), thus confirming the effectiveness of denoising. We also observe that $\sigma_{\tilde{\epsilon}} \approx \frac{1}{2}\sigma$, independently from the source position.

The localization performance was then evaluated in terms of Root Mean Square Error (RMSE), computed as

$$\text{RMSE}(\mathbf{x}) = \sqrt{\frac{1}{I} \sum_{i=1}^{I} \|\tilde{\mathbf{x}}_i - \mathbf{x}\|_2^2}, \tag{30}$$

where $\mathbf{x}$ is the nominal source position and $\tilde{\mathbf{x}}_i$ is its estimate at the ith Monte-Carlo run. The three algorithms were fed with both the measured TDOAs $\hat{\boldsymbol{\tau}}_i$ and the denoised TDOAs $\mathcal{P}(\hat{\boldsymbol{\tau}}_i; \boldsymbol{\Sigma})$. In particular, for LS and SRD-LS we considered only the $n = 6$ TDOAs measured with respect to the first sensor, selected as the reference one. For GS, we considered the full set of q TDOAs, before and after denoising. Results are reported in Fig. 7. As done before,
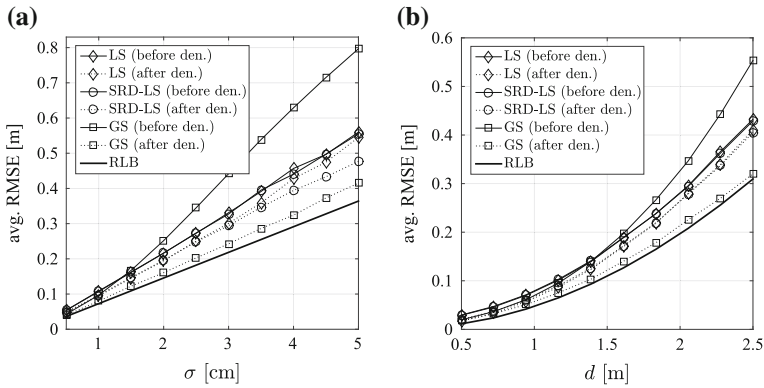
(a)

(b)

Fig. 7 Localization accuracy before and after denoising, as a function of: injected noise σ (**a**); source distance d (**b**)

we report results as a function of σ when $d = 1.5$ m (Fig. 7a); and varying the distance d when $\sigma = 1.5$ cm (Fig. 7b). For each tested algorithm, the figures show the average RMSE achieved before and after the denoising of the measured TDOAs. For the sake of comparison, we also report the RMSE Lower Bound (RLB) implied by the CRLB. As expected, all the algorithms exhibit improved localization accuracy when denoised TDOAs are used. It is interesting to observe the different behavior of the algorithms using only the nonredundant TDOAs (LS and SRD-LS) and that using the full TDOA set (GS). Indeed, denoising produces a moderate effect on LS and SRD-LS, while the benefit is impressive for GS. Despite using the full TDOA set, before denoising GS exhibits very poor performances compared to LS and SRD-LS, especially for distant sources and at high amount of injected noise on TDOAs. This behavior is not completely unexpected, as GS extends the LS approach, which is known to suffer from non negligible bias (Benesty and Huang 2004). We observed that, using the full TDOA set, the bias increases significantly, even if the variance of estimation is reduced due to the availability of more measurements. We preliminarily noticed that the bias is greatly reduced when GS is fed with the denoised TDOA set. Roughly speaking, this means that denoising allows GS to effectively exploit the data redundancy. This positively impacts on source localization with GS after denoising, producing an RMSE that approaches the RLB. On the other hand, the redundancy is only partially exploited by LS and SRD-LS, as they rely on the nonredundant denoised TDOA set.

5.2.2 TDOAs corrupted with non-Gaussian noise

Considering the simulation setup described in the previous paragraph, we repeated all the tests injecting different types of non-Gaussian noise in the ideal TDOAs. In particular, we focus on the following noise models:

- i.i.d. zero-mean uniform noise in the range $\left[-\frac{c}{2f_s}, \frac{c}{2f_s} \right]$, which mimics the sampling error caused by estimating the TDOA from the a Generalized Cross-Correlation (GCC) function (Knapp and Carter 1976) sampled at $f_s = 8$ kHz; $c = 343$ m/s is the speed of sound;
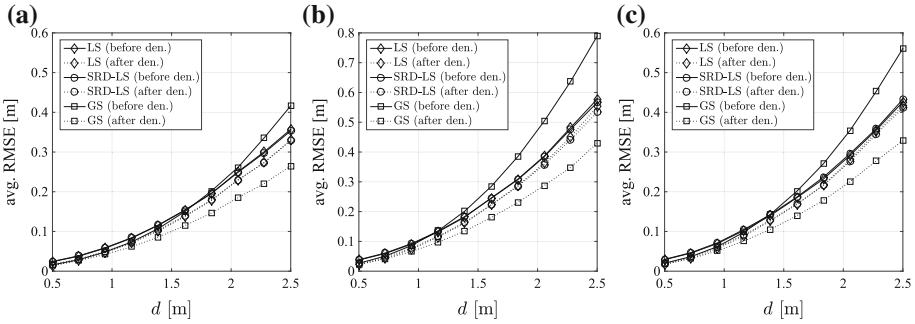
Fig. 8 Localization accuracy before and after denoising, as a function of distance d, for different TDOA noise models: uniformly distributed noise (**a**), mixture of uniform and Gaussian noise (**b**), Laplacian noise (**c**)

- a mixture of i.i.d. zero-mean uniform and Gaussian noise, the former distributed in the range $\left[-\frac{c}{2f_s}, \frac{c}{2f_s}\right]$; the latter with standard deviation $\sigma = 1.5$ cm. This mixture of models is suggested in Benesty and Huang (2004);
- i.i.d. zero-mean Laplacian noise with standard deviation $\sigma = 1.5$ cm.

The localization accuracy of the three algorithms exhibits the same trend obtained for the Gaussian noise model. For reasons of space, here we limit to report the average RMSE as a function of the source distance d. Figures 8a, b, c show the results for the three considered noise distributions, respectively. Even if the noise models differ from the Gaussian assumption underlying the denoising theory, all the algorithms improve their accuracy using denoised TDOAs. As before, this is especially true for GS, which exhibit the most noticeable improvement. This suggests that denoising can be applied when the TDOA error does not strictly meet the Gaussian assumption.

6 Dealing with an incomplete set of TDOAs

There exist many scenarios where not the whole set of TDOAs is available. For example, when computational cost is an issue, the computation of all the possible TDOAs is not feasible. In the following, we adapt our previous analysis in order to handle relaxed denoising also in such situations.

Let us assume that the TDOAs $\{\tau_{j_1 i_1}, \ldots, \tau_{j_s i_s}\}$, $0 \le s \le q$, are not available and let $S = \{(j_1, i_1), \ldots, (j_s, i_s)\}$ be the corresponding set of indices. In this setting, the proper TDOA map is

$$\boldsymbol{\tau}_{n,S}^{*} : \mathbb{R}^3 \longrightarrow \mathbb{R}^{q-s}$$
$$\mathbf{x} \longmapsto \boldsymbol{\tau}_{n,S}^{*}(\mathbf{x}) \,, \tag{31}$$

where

$$\boldsymbol{\tau}_{n,S}^{*}(\mathbf{x}) = (\tau_{10}(\mathbf{x}), \ldots, \widehat{\tau_{j_1 i_1}(\mathbf{x})}, \ldots, \widehat{\tau_{j_s i_s}(\mathbf{x})}, \ldots, \tau_{n\,n-1}(\mathbf{x}))^T$$

and $\widehat{\tau_{j_1 i_1}(\mathbf{x})}$ means that the item is missing. As before, we define the TDOA space as the target set $\mathbb{R}^{q-s}$ of $\boldsymbol{\tau}_{n,S}^{*}$ and the image $\text{Im}(\boldsymbol{\tau}_{n,S}^{*})$ as $\Theta_{n,S}$. Clearly, the TDOA map $\boldsymbol{\tau}_{n,S}^{*}$ is strictly related to $\boldsymbol{\tau}_{n}^{*}$. Indeed, let us consider the projection $p_S : \mathbb{R}^q \to \mathbb{R}^{q-s}$ that takes care of forgetting the s coordinates corresponding to the indices in S. Then, one has $\boldsymbol{\tau}_{n,S}^{*} = p_S \circ \boldsymbol{\tau}_{n}^{*}$ and $\Theta_{n,S} = p_S(\Theta_n)$, where the symbol $\circ$ denotes the function composition operator.

In the presence of measurement errors, we assume that the TDOAs are described by the statistical model

$$\hat{\tau}_{n,S}^{*}(\mathbf{x}) = \tau_{n,S}^{*}(\mathbf{x}) + \boldsymbol{\varepsilon}_S, \quad \text{where} \quad \boldsymbol{\varepsilon}_S \sim N(\mathbf{0}, \boldsymbol{\Sigma}_S). \tag{32}$$

As above, the Fisher matrix $\boldsymbol{\Sigma}_S{}^{-1}$ defines a Euclidean structure on the TDOA space $\mathbb{R}^{q-s}$ and the same discussion made in Sect. 2 holds in this situation.

The definition and the analysis of the relaxed denoising procedure are very similar to the ones made in Sect. 4.2. First of all, we adapt Theorem 1.

Theorem 3 *Let us take $n + 1$ microphones at $\mathbf{m_0}, \ldots, \mathbf{m_n}$ in $\mathbb{R}^3$, where $n \geq 2$. Then $\Theta_{n,S}$ is a subset of the linear subspace $V_S = p_S(V_n) \subset \mathbb{R}^{q-s}$. Moreover, the orthogonal projection $\mathcal{P}_S(\hat{\boldsymbol{\tau}}_S; \boldsymbol{\Sigma}_S)$ of $\hat{\boldsymbol{\tau}}_S \in \mathbb{R}^{q-s}$ on V_S, with respect to $\langle \ , \ \rangle_{\boldsymbol{\Sigma}_S^{-1}}$, is a sufficient statistic for the underlying parameter $\mathbf{x}$.*

Proof p_S is a linear map, therefore V_S is a linear subspace of $\mathbb{R}^{q-s}$. The other claims follow in the same way of Proposition 1 and Theorem 1. □

From Theorem 3 it follows that $\dim(V_S) \leq \dim(V_n) = n$, where dim denotes the dimension of a vector space. It is not difficult to prove that the equality holds if, and only if, the set of available TDOAs contains n independent measures, for example the n TDOAs calculated with respect to a reference microphone. In this case, the map p_S is a bijection between V and V_S. This means that it is possible to obtain the full set of q denoised TDOAs as $p_S^{-1}(\mathcal{P}_S(\hat{\boldsymbol{\tau}}_S; \boldsymbol{\Sigma}_S)) \in V$. Concretely, one has to use the linear equations

$$- \tau_{ik} + \tau_{jk} - \tau_{ji} = 0, \quad i \neq j \neq k, \tag{33}$$

to calculate the missing TDOAs with indices in S. It is important to remark that this operation does not increase the noise on the dataset, as indeed it would happen if we apply the same procedure on the original data by calculating $p_S^{-1}(\hat{\boldsymbol{\tau}}_S)$.

6.1 Denoising algorithm

In order to explicitly construct the projection map $\mathcal{P}_S$, let us take the $(q - s, q)$ matrix $\mathbf{I_S}$ defined by removing the s rows corresponding to the indexes in S from the (q, q) identity matrix. It can be easily proven that $\mathbf{I_S}$ represents the map p_S with respect to the standard basis $\mathcal{B}_q$ and $\mathcal{B}_{q-s}$ of $\mathbb{R}^q$ and $\mathbb{R}^{q-s}$, respectively. Then, given a generic basis $\{\mathbf{v_1}, \ldots, \mathbf{v_n}\}$ of $\ker(\mathbf{C})$, from Theorem 3 follows that the set $\{\mathbf{I_S v_1}, \ldots, \mathbf{I_S v_n}\}$ spans V_S. After reducing it to an independent set of vectors and subsequently applying the Gram–Schmidt algorithm, we can finally find an orthonormal basis of V_S with respect to the scalar product $\langle \ , \ \rangle_{\boldsymbol{\Sigma}_S^{-1}}$.

From this point on, one can proceed exactly as done in the previous sections. In particular, the map $\mathcal{P}_S$ is defined by the analogous of the formula (25) and it is represented, with respect to $\mathcal{B}_{q-s}$, by a $(q - s, q - s)$ matrix $\mathbf{P_S}$. Then, the denoised TDOAs are

$$\mathcal{P}_S(\hat{\boldsymbol{\tau}}_S; \boldsymbol{\Sigma}_S) = \mathbf{P_S}\,\hat{\boldsymbol{\tau}}_S. \tag{34}$$

6.2 Impact on source localization

We summarize the main facts on the denoising procedure in the following theorem.

Theorem 4 *let $\boldsymbol{\Sigma}_S$ be the covariance matrix of $\hat{\boldsymbol{\tau}}_S$ and $f(\mathbf{x}, \mathbf{c})$ be any given cost function, where $\mathbf{c}$ are the input TDOA data. Then:*

1. *the covariance matrix of $\mathcal{P}_S(\hat{\boldsymbol{\tau}}_S; \boldsymbol{\Sigma}_S)$ is $\boldsymbol{\Sigma}_S' = \mathbf{P}_S \boldsymbol{\Sigma}_S \mathbf{P}_S^T$;*
2. $\boldsymbol{\Sigma}_S \succeq \boldsymbol{\Sigma}_S'$;
3. *at first order approximation, the covariance matrices $\boldsymbol{\Sigma}_{S,\bar{\mathbf{x}}}$ and $\boldsymbol{\Sigma}_{S,\bar{\mathbf{x}}}'$ of the estimators $\bar{\mathbf{x}} = \arg\min f(\mathbf{x}, \hat{\boldsymbol{\tau}}_S)$ and $\bar{\mathbf{x}}' = \arg\min f(\mathbf{x}, \mathcal{P}_S(\hat{\boldsymbol{\tau}}_S; \boldsymbol{\Sigma}_S))$, respectively, satisfy $\boldsymbol{\Sigma}_{S,\bar{\mathbf{x}}} \succeq \boldsymbol{\Sigma}_{S,\bar{\mathbf{x}}}'$.*

Proof The proof is similar to the ones of Theorem 2 and Corollary 1. □

6.3 Numerical examples

In this paragraph we show some numerical examples, devoted to investigate the effect of relaxed denoising when the full TDOA set is not entirely available. To this end, we refer again to the simulation setup described in Sect. 5.2.1. However, we now consider the availability of the n TDOAs referred to the first microphone, along with $z \leq q - n$ additional TDOAs. Noisy TDOAs were obtained corrupting their nominal values with i.i.d. zero-mean Gaussian noise with standard deviation $\sigma = 1.5$ cm. In this scenario, denoising was accomplished using (34). In particular, we built the vector $\hat{\boldsymbol{\tau}}_S$ including the $n + z$ available TDOAs, and we computed the projection matrix $\mathbf{P}_S$ accordingly. For this test we considered sources at a fixed distance $d = 1.5$ m. We tested the denoising procedure considering values of z in the range from 1 and $q - n - 1 = 14$. For all the I Monte-Carlo runs, we generated all the possible combinations of z TDOAs extracted from the last $n - q$ entries of the vector $\hat{\boldsymbol{\tau}}_i$. As before, we considered the LS, SRD-LS and GS algorithms for source localization. The results, averaged among all the noise realizations and all the combinations, are reported in Fig. 9. In particular, Fig. 9a shows the residual error on TDOAs after denoising, while Fig. 9b highlights the impact of denoising on localization. Note that we added to the graphs the points at $z = 0$ (i.e., when only the n TDOAs referred to the first microphone are available) and at $z = q - n = 15$ (i.e., when all the TDOAs are used). It is worth noticing that the availability of just a few additional TDOAs leads to a relevant reduction of the TDOA standard deviation, with respect to $z = 0$. This reflects also on localization, as all the algorithms monotonically improve their accuracy increasing the number of available measurements. Also in this case, GS exhibits the best accuracy after denoising, while being characterized by an unstable behavior using the original TDOAs. Indeed, with the original data GS improves its accuracy when $z < 6$; for higher values of z, the error bias becomes relevant and the overall RMSE diverges.
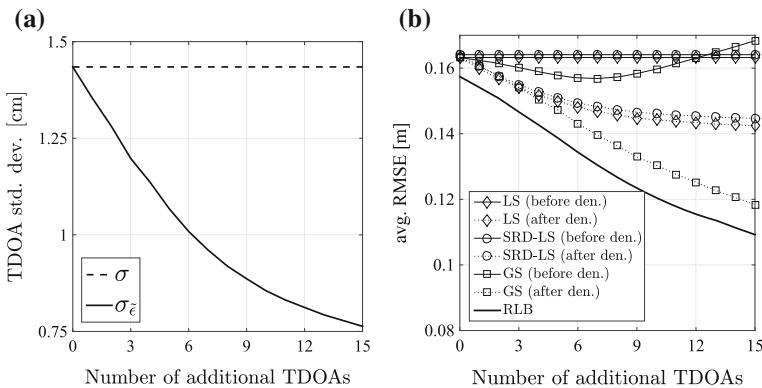

Fig. 9 TDOA residual error before and after denoising (**a**) and localization accuracy (**b**). Both plots are function of the number of additional TDOAs

It is important to highlight the practical implications of these results. Let us consider a microphone array composed by $(n + 1)$ sensors. The computational power requested for computing all the q TDOAs is mainly due to the calculation of GCCs between the pairs of microphone signals. In case of limited computational capabilities, a typical solution would be that of computing only the nonredundant TDOA set. However, it turns to be more convenient to compute z additional TDOAs in order to fully exploit the available computational power. This enables better localization, without modifying the array configuration.

7 Conclusions

In this manuscript we reformulated the problem of source localization in the TDOA space. This enabled us to show that source localization has a neat interpretation in terms of TDOA denoising. As a simple solution to the denoising problem is not available, we proved that it is possible to relax the problem to a linear one, whose solution is based on projecting TDOAs on a linear subspace of the TDOA space. Moreover, we also derived the problem solution for the case in which only a few TDOAs measurements are available.

The analysis performed in this manuscript does not limit to numerical simulations. Indeed, each choice behind the presented algorithm is fully justified by means of analytical proofs that further validate and justify the works in So et al. (2008) and Schmidt (1996). As a matter of fact, in this manuscript we proved that denoising has a positive effect on source localization from a theoretical perspective. Moreover, we tested the denoising algorithm using different noise models, thus highlighting that the method is still valid even when noise hypotheses are not completely fulfilled. Finally, we also made use of different cost functions to gain an interesting insight on how denoising impacts on different localization algorithms.

The extension of the relaxed denoising algorithm to the case of missing TDOAs have also interesting implications in real-world scenarios. As a matter of fact, it enables to fully exploit hardware computational capabilities in order to increase localization performance. As an example, by fixing the available computational complexity, one can tune the localization system in order to measure a given amount of TDOAs and fully take advantage of them in a synergistic fashion.

According to the results of this work, it is important to note that it is possible to envision the development of algorithms working in the TDOA space to solve the complex ML source localization problem in a easier way. This will be the scope of possible future works.

References

Alameda-Pineda, X., & Horaud, R. (2014). A geometric approach to sound source localization from time-delay estimates. *IEEE Transactions on Audio, Speech, and Language Processing*, *22*, 1082–1095. doi:10.1109/TASLP.2014.2317989.

Amari, S., & Nagaoka, H. (2000). *Methods of Information Geometry*. Providence, RI: American Mathematical Society.

Basu, S., Pollack, R., & Roy, M. (2006). *Algorithms in real algebraic geometry. Algorithms and computation in mathematics* (2nd ed.). Berlin: Springer. doi:10.1007/3-540-33099-2.

Beck, A., Stoica, P., & Li, J. (2008). Exact and approximate solutions of source localization problems. *IEEE Transactions on Signal Processing*, *56*, 1770–1778. doi:10.1109/TSP.2007.909342.

Benesty, J., & Huang, Y. (Eds.). (2004). *Audio signal processing for next-generation multimedia communication systems*. Berlin: Springer.

Bestagini, P., Compagnoni, M., Antonacci, F., Sarti, A., & Tubaro, S. (2013). TDOA-based acoustic source localization in the space-range reference frame. *Multidimensional Systems and Signal Processing*. doi:10.1007/s11045-013-0233-8.

Canclini, A., Antonacci, E., Sarti, A., & Tubaro, S. (2013). Acoustic source localization with distributed asynchronous microphone networks. *IEEE Transactions on Audio, Speech, and Language Processing*, *21*(2), 439–443. doi:10.1109/TASL.2012.2215601.

Canclini, A., Bestagini, P., Antonacci, F., Compagnoni, M., Sarti, A., & Tubaro, S. (2015). A robust and low-complexity source localization algorithm for asynchronous distributed microphone networks. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, *23*(10), 1563–1575. doi:10.1109/TASLP.2015.2439040.

Chen, J., Hudson, R., & Yao, K. (2002). Maximum-likelihood source localization and unknown sensor location estimation for wideband signals in the near-field. *IEEE Transactions on Signal Processing*, *50*(8), 1843–1854. doi:10.1109/TSP.2002.800420.

Compagnoni, M., & Notari, R. (2014). TDOA-based localization in two dimension: The bifurcation curve. *Fundamenta Informaticae*, *135*, 199–210.

Compagnoni, M., Bestagini, P., Antonacci, F., Sarti, A., & Tubaro, S. (2012). Localization of acoustic sources through the fitting of propagation cones using multiple independent arrays. *IEEE Transactions on Audio, Speech, and Language Processing*, *20*, 1964–1975. doi:10.1109/TASL.2012.2191958.

Compagnoni, M., Notari, R., Antonacci, F., & Sarti, A. (2014). A comprehensive analysis of the geometry of tdoa maps in localization problems. *Inverse Problems*, *30*(3), 035004. http://stacks.iop.org/0266-5611/30/i=3/a=035004.

D'Arca, E., Robertson, N., & Hopgood, J. (2014). Look who's talking: Detecting the dominant speaker in a cluttered scenario. In *IEEE International conference on acoustics, speech and signal processing, ICASSP '14*.

Gillette, M., & Silverman, H. (2008). A linear closed-form algorithm for source localization from time-differences of arrival. *IEEE Signal Processing Letters (SPL)*, *15*, 1–4.

Grafarend, E., & Shan, J. (2002). GPS solutions: Closed forms, critical and special configurations of P4P. *GPS Solutions*, *5*(3), 29–41.

Hahn, W., & Tretter, S. (1973). Optimum processing for delay-vector estimation in passive signal arrays. *IEEE Transactions on Information Theory*, *19*, 608–614. doi:10.1109/TIT.1973.1055077.

Hu, H. H., & Li, D. (2002). Energy based collaborative source localization using acoustic micro-sensor array. *2002 IEEE workshop on multimedia signal processing* (pp. 371–375).

Hu, J., & Yang, C. (2010). Estimation of sound source number and directions under a multisource reverberant environment. *EURASIP Journal on Advances in Signal Processing*, *2010*, 63.

Huang, Y., Benesty, J., Elko, G., & Mersereati, R. (2001). Real-time passive source localization: a practical linear-correction least-squares approach. *IEEE Transactions on Speech and Audio Processing*, *9*, 943–956. doi:10.1109/89.966097.

Ianniello, J. (1982). Time delay estimation via cross-correlation in the presence of large estimation errors. *IEEE Transactions on Acoustics, Speech and Signal Processing*, *30*, 998–1003.

Kehu, Y., Gang, W., & Luo, L. Z.-Q. (2009). Efficient convex relaxation methods for robust target localization by a sensor network using time differences of arrivals. *IEEE Transactions on Signal Processing*, *57*(7), 2775–2784.

Knapp, C., & Carter, G. (1976). The generalized correlation method for estimation of time delay. *IEEE Transactions on Acoustics, Speech and Signal Processing*, *24*, 320–327.

Koch, V., & Westphal, R. (1995). New approach to a multistatic passive radar sensor for air/space defense. *Aerospace and Electronic Systems Magazine, IEEE*, *10*(11), 24–32.

Lehmann, E. L., & Casella, G. (1998). *Theory of point estimation* (2nd ed.). New York: Springer.

Schau, H., & Robinson, A. (1987). Passive source localization employing intersecting spherical surfaces from time-of-arrival differences. *IEEE Transactions on Acoustics, Speech and Signal Processing*, *35*, 1223–1225. doi:10.1109/TASSP.1987.1165266.

Scheuing, J., & Yang, B. (2006). Disambiguation of TDOA estimates in multi-path multi-source environments (DATEMM). *IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP '06*, *4*, IV—IV. doi:10.1109/ICASSP.2006.1661099.

Scheuing, J., & Yang, B. (2008). Disambiguation of TDOA estimation for multiple sources in reverberant environments. *IEEE Transactions on Audio, Speech, and Language Processing*, *16*, 1479–1489. doi:10.1109/TASL.2008.2004533.

Schmidt, R. (1972). A new approach to geometry of range difference location. *IEEE Transactions on Aerospace and Electronic Systems*, *AES–8*, 821–835. doi:10.1109/TAES.1972.309614.

Schmidt, R. (1996). Least squares range difference location. *IEEE Transactions on Aerospace and Electronic Systems*, *AES–32*, 234–242. doi:10.1109/7.481265.

Smith, J., & Abel, J. S. (1987). Closed-form least-squares source location estimation from range-difference measurements. *IEEE Transactions on Acoustics, Speech, Signal Processing*, *ASSP–35*, 1661–1669.

So, H. C., Chan, Y. T., & Chan, F. K. W. (2008). Closed-form formulae for time-difference-of-arrival estimation. *IEEE Transactions on Signal Processing*, *56*, 2614–2620.

Spencer, S. (2007). The two-dimensional source location problem for time differences of arrival at minimal element monitoring arrays. *Journal of the Acoustical Society of America*, *121*, 3579–3594.

Stoica P, Nehorai A (1988) MUSIC, maximum likelihood and cramer-rao bound. In *IEEE international conference on acoustics, speech and signal processing, ICASSP '88*. doi:10.1109/ICASSP.1988.197097.

Teunissen, P. (2000). *Adjustment theory: An introduction. Series on mathematical geodesy and positioning*. Delft: Delft University Press.

Trifa, V., Koene, A., Moren, J., & Cheng, G. (2007). Real-time acoustic source localization in noisy environments for human-robot multimodal interaction. In *IEEE international symposium on robot and human interactive communication, RO-MAN '07*.

Valenzise, G., Gerosa, L., Tagliasacchi, M., Antonacci, F., & Sarti, A. (2007). Scream and gunshot detection and localization for audio-surveillance systems. In *IEEE conference on advanced video and signal based surveillance, AVSS '07*.

Yimin, Z., Amin, M., & Ahmad, F. (2008). Localization of inanimate moving targets using dual-frequency synthetic aperture radar and time-frequency analysis. *IEEE International Geoscience and Remote Sensing Symposium, 2008. IGARSS 2008*, *2*, II33–II36.

Marco Compagnoni was born in Tirano (Italy) on April 25, 1980. He received the Laurea cum laude degree in Theoretical Physics (Università di Milano Bicocca, 2004) and the Ph.D. degree in Applied Mathematics (Politecnico di Milano, 2009). He is currently a post-doc researcher at Geometry-Algebra In Application (GAIA) group in Dipartimento di Matematica at Politecnico di Milano. His research interests are applied math, especially in the application of geometry in mathematical physics and in statistics, in image processing and pattern recognition and in space-time processing of audio signals.

Antonio Canclini was born in Sondalo (Italy), on December 19, 1983. He received the M.Sc. degree (cum laude) in computer science and engineering (2008) and the Ph.D. degree (cum laude) in information technology (2012) from the Politecnico di Milano, Italy. He is currently a Post-Doc Researcher at the Image and Sound Processing Group, Dipartimento di Elettronica, Informazione e Bioingegneria, Politecnico di Milano. His research focuses on space-time audio signal processing (localization of acoustic sources and reflectors, spatial sound field rendering, radiation pattern estimation).

Paolo Bestagini was born in Novara (Italy) on February 22, 1986. He received the M.Sc. in Telecommunications Engineering and the Ph.D. in Information Technology at the Politecnico di Milano in 2010 and 2014, respectively. He is currently a Post-Doc researcher at the Image and Sound Processing Group, Politecnico di Milano. His research activity is focused on multimedia forensics and acoustic signal processing for microphone arrays.

Fabio Antonacci was born in Bari (Italy) on July 26, 1979. He received Laurea degree in 2004 in Telecommunication Engineering and Ph.D. in Information Engineering in 2008, both at Politecnico di Milano, Italy. He is currently assistant professor at Politecnico di Milano. His research focuses on space-time processing of audio signals, for both speaker and microphone arrays (source localization, acoustic scene analysis, rendering of spatial sound) and on modeling of acoustic propagation (visibility-based beam tracing). He is author of more than 50 articles in proceedings of international conferences and on peer-reviewed journals.

Augusto Sarti received the M.S. and the Ph.D. degrees in electronic engineering, both from the University of Padua, Italy, in 1988 and 1993, respectively, with research on nonlinear communication systems. His graduate studies included a joint graduate program with the University of California, Berkeley. In 1993, he joined the Politecnico di Milano, Milan, Italy, where he is currently an Associate Professor. In 2013, he also joined the University of California, Davis, as an Adjunct Professor. His research interests are in the area of multimedia signal processing, with particular focus on sound analysis, processing and synthesis; space-time audio processing; and music information processing and retrieval. He also worked on problems of multidimensional signal processing, image analysis and 3D vision; motion planning and nonlinear system theory. He coauthored well over 200 scientific publications on international journals and congresses as well as numerous patents in the multimedia signal processing area. He coordinates the activities of the Musical Acoustics Lab and of the Sound and Music Computing Lab of the Politecnico di Milano. He promoted and coordinated or contributed to numerous (20+) European projects. He is active in the IEEE Technical Committee on Audio and Acoustics Signal Processing, and with several editorial responsibilities for IEEE Journals. He was co-chairman of various international workshops, including the 2005 Edition of the IEEE International Conference on Advanced Video and Signal based Surveillance (AVSS); and the 2009 edition of the Digital Audio Effects conference, (DAFx). he has served in the organizing committees of numerous other conferences in the area of signal processing.

Stefano Tubaro was born in Novara in 1957. He completed his studies in Electronic Engineering at the Politecnico di Milano, Italy, in 1982. He then joined the Dipartimento di Elettronica, Informazione e Bioingegneria of the Politecnico di Milano, first as a researcher of the National Research Council, and then (in November 1991) as an Associate Professor. Since December 2004 he has been appointed as Full Professor of Telecommunication at the Politecnico di Milano. His current research interests are on advanced algorithms for video and sound processing. Stefano Tubaro authored over 180 scientific publications on international journals and congresses and he is co-author of more than 15 patents. In the past few years he has focused his interest on the development of innovative techniques for image and video tampering detection and, in general, for the blind recovery of the "processing history" of multimedia objects. Stefano Tubaro coordinates the research activities of the Image and Sound Processing Group (ISPG) at the Dipartimento di Elettronica, Informazione e Bioingegneria of the Politecnico di Milano. He had the role of Project Coordinator of the European Project ORIGAMI: A new paradigm for high-quality mixing of real and virtual and of the research project ICT-FET-OPEN REWIND: REVerse engineering of audio-VIsual coNtent Data. This last project was aimed at synergistically combining principles of signal processing, machine learning and information theory to answer relevant questions on the past history of such objects. Stefano Tubaro is a member the IEEE Multimedia Signal Processing Technical Committee and of the IEEE SPS Image Video and Multidimensional Signal Technical Commitee. He was in the organization committee of a number of international conferences: IEEE-MMSP-2004/2013, IEEE-ICIP-2005, IEEE-AVSS-2005/2009, IEEE-ICDSC-2009, IEEE-MMSP-2013, IEEE-ICME-2015 to mention a few. From May 2012 to April 2015 he was an Associate Editor of the IEEE Transactions on Image Processing, currently he is an Associate Editor of the IEEE Transactions on Information Forensics and Security.