
This is an electronic reprint of the original article.
This reprint may differ from the original in pagination and typographic detail.

Haasler, Isabel; Elvander, Filip

Multi-Frequency Tracking via Group-Sparse Optimal Transport

Published in:
IEEE Control Systems Letters

DOI:
[10.1109/LCSYS.2024.3405375](https://doi.org/10.1109/LCSYS.2024.3405375)

Published: 24/05/2024

Document Version
Peer-reviewed accepted author manuscript, also known as Final accepted manuscript or Post-print

Please cite the original version:
Haasler, I., & Elvander, F. (2024). Multi-Frequency Tracking via Group-Sparse Optimal Transport. *IEEE Control Systems Letters*, 8, 1048-1053. <https://doi.org/10.1109/LCSYS.2024.3405375>

This material is protected by copyright and other intellectual property rights, and duplication or sale of all or part of any of the repository collections is not permitted, except that material may be duplicated by you for your research use or educational purposes in electronic or print form. You must obtain permission for any other use. Electronic or print copies may not be offered, whether for sale or otherwise to anyone who is not an authorised user.

Multi-frequency tracking via group-sparse optimal transport

Isabel Haasler and Filip Elvander

Abstract—In this work, we introduce an optimal transport framework for inferring power distributions over both spatial location and temporal frequency. Recently, it has been shown that optimal transport is a powerful tool for estimating spatial spectra that change smoothly over time. In this work, we consider the tracking of the spatio-temporal spectrum corresponding to a small number of moving broad-band signal sources. Typically, such tracking problems are addressed by treating the spatio-temporal power distribution in a frequency-by-frequency manner, allowing to use well-understood models for narrow-band signals. This however leads to decreased target resolution due to inefficient use of the available information. We propose an extension of the optimal transport framework that exploits information from several frequencies simultaneously by estimating a spatio-temporal distribution penalized by a group-sparsity regularizer. This approach finds a spatial spectrum that changes smoothly over time, and at each time instance has a small support that is similar across frequencies. To the best of the authors’ knowledge, this is the first formulation combining optimal transport and sparsity for solving inverse problems. As is shown on simulated and real data, our method can successfully track targets in scenarios where information from separate frequency bands alone is insufficient.

Index Terms—Estimation; Optimization algorithms; Large-scale systems

I. INTRODUCTION

SPECTRAL estimation appears in a variety of control and signal processing applications, ranging from fault detection [1] to noise reduction and speech enhancement [2]. For wide-sense stochastic processes, commonly employed as a signal model, the (temporal) spectrum parametrizes the covariance function and describes the distribution of power over frequency [3], [4]. Analogously, in multi-sensor or array processing scenarios, the spatial spectrum gives the distribution of power on the spatial domain and parametrizes the array or spatial covariance matrix [5]. Commonly in applications, it is assumed that such spatial spectra correspond to temporal signals supported on a single carrier frequency, allowing for representing time-delays by wave-form phase-shifts. Although such narrow-band assumptions do not hold for scenarios with broad-band sources, approximations can be constructed by means of filter banks or the short-time Fourier transform [6].

This work was supported by the Knut and Alice Wallenberg Foundation under grant KAW 2021.0274.

I. Haasler is with Signal Processing Laboratory, LTS4, EPFL, Lausanne, Switzerland. isabel.haasler@epfl.ch

F. Elvander is with the Dept. of Information and Communications Engineering, Aalto University, Finland. filip.elvander@aalto.fi

However, as the resulting narrow-band signals are typically processed independently, information common for sets of frequencies, such as spectral coherence, is ignored [7].

In this letter, we consider the problem of spatial spectral estimation for broad-band sources or targets. Furthermore, we are interested in the case when the scene is observed at a sequence of time instances, between which the location of the targets change in a smooth fashion. In recent works, we have developed a framework for spectral estimation building on the concept of optimal transport (OT) [8], [9]. In this setting the geometric property of OT to capture smooth shifts in distributions, i.e., spectral energy content [10], has shown to be a powerful tool for target tracking. OT has also found applications in various other control applications [11], e.g., in control and estimation for multi-agent systems [12], [13], and uncertainty quantification [14].

In the case of spatial spectral estimation, our previous work has been limited to narrow-band scenarios, i.e., with signals supported on a single carrier frequency. In this work, we develop these concepts further, and in particular to scenarios with broad-band sources. Here, all available information, corresponding to several observation times and temporal frequencies, is used in estimation of the spatio-temporal spectrum, i.e., a distribution of power over both spatial domain and temporal frequency. We propose to achieve this information sharing between frequencies by imposing the assumption of spatial sparsity: as signal sources should be relatively few, the support of each spatial spectrum should be small. Furthermore, we present an efficient algorithm with linear convergence rate implementing our proposed estimator.

To the best of the authors’ knowledge, this is also the first time in which OT and sparsity-inducing penalties are used jointly for solving inverse problems.

II. BACKGROUND

A. Spatio-temporal estimation

Consider a scenario in which a superposition of broad-band signals, emitted by a set of spatially localized sources in the far-field, impinge on an array of $Q \in \mathbb{N}$ sensors. Let the corresponding sensor array signal be $\mathbf{y}(t) = [y_1(t) \ \dots \ y_Q(t)]^T \in \mathbb{C}^Q$, for $t \in \mathbb{R}$. Then, modeling the sources as wide-sense stationary processes, we seek a spatio-temporal spectrum describing the distribution of signal power over look-angle¹ and temporal frequency. That is,

¹To simplify the exposition, but without loss of generality, we here let the spatial domain correspond to direction-of-arrival (DoA).

letting Θ and \mathbb{T} denote the angle and frequency spaces respectively, we seek $\Phi \in \mathcal{M}_+(\Theta \times \mathbb{T})$. Furthermore, consider passing each sensor signal $y_k(t)$ through a filter bank² of F narrow-band filters with center frequencies ω_f , $f \in \llbracket F \rrbracket = \{1, \dots, F\}$, yielding the set of narrow-band sensor signals $\mathbf{y}_f(t) = [y_{1,f}(t) \dots y_{Q,f}(t)]^T \in \mathbb{C}^Q$, for $f \in \llbracket F \rrbracket$. With this, the so-called spatial covariance matrix for carrier frequency ω_f is given by

$$\mathbf{R}_f \triangleq \mathbb{E}(\mathbf{y}_f \mathbf{y}_f^H) = \int_{\Theta} \mathbf{a}_f(\theta) \mathbf{a}_f(\theta)^H \Phi_f(\theta) d\theta = G_f(\Phi_f) \quad (1)$$

where $\mathbb{E}(\cdot)$ denotes the expectation operation, and where we have defined the set of linear operators $G_f : \mathcal{M}_+(\Theta) \rightarrow \mathbb{C}^{Q \times Q}$, with $\mathcal{M}_+(\Theta)$ denoting the set of non-negative densities on Θ . Here, the vector functions $\mathbf{a}_f : \Theta \rightarrow \mathbb{C}^Q$ denotes the array response at carrier frequency ω_f , encoding the array geometry, as well as filter response and propagation properties of the space. Then, given \mathbf{R}_f , $f \in \llbracket F \rrbracket$, or estimates thereof, we seek to estimate Φ , or more precisely $\{\Phi_f\}_{f=1}^F = \{\Phi(\omega_f, \cdot)\}_{f=1}^F$.

B. Spectral tracking via optimal transport

Let $\Phi_0, \Phi_1 \in \mathcal{M}_+(X)$ for a space X . The optimal transport problem [15], [16] is to transform Φ_0 into Φ_1 in the most efficient way, where efficiency is measured in terms of a cost function $c : X \times X \rightarrow \mathbb{R}$, and where $c(x_0, x_1)$ denotes the cost for moving a unit mass from $x_0 \in X$ to $x_1 \in X$. OT finds a so-called transport plan, which is a bi-variate distribution $m \in \mathcal{M}_+(X \times X)$, that minimizes

$$\begin{aligned} S(\Phi^{(0)}, \Phi^{(1)}) &= \min_{m \in \mathcal{M}_+(X \times X)} \int_{X \times X} c(x_0, x_1) m(x_0, x_1) dx_0 dx_1 \\ \text{subject to } &\int_X m(x_0, x_1) dx_1 = \Phi^{(0)}(x_0) \\ &\int_X m(x_0, x_1) dx_0 = \Phi^{(1)}(x_1). \end{aligned} \quad (2)$$

The objective value of (2) can be interpreted as a measure of distance between the distributions $\Phi^{(0)}$ and $\Phi^{(1)}$, quantifying how much the mass has to be moved in order to transform $\Phi^{(0)}$ into $\Phi^{(1)}$. This property has recently proven useful in the setting of tracking spatial spectra based on covariance measurements as in (1), assuming that the spectral content varies slowly compared to the sampling frequency. Namely, with $X = \Theta$, in [8], [9] the OT distance (2) is used as a regularizing term to find spectral estimates whose mass moves smoothly between consecutive time points, which results in the formulation

$$\text{minimize}_{\Phi^{(t)}, t \in \llbracket T \rrbracket} \sum_{t=1}^T \|G(\Phi^{(t)}) - \mathbf{R}^{(t)}\|_2^2 + \sum_{t=1}^{T-1} S(\Phi^{(t)}, \Phi^{(t+1)}). \quad (3)$$

Moreover, the problem can equivalently be posed as a so-called multi-marginal OT problem over the product space $\mathcal{X} = \Theta^T := \Theta \times \dots \times \Theta$ [8]. In this setting, c and m are a T -variate function and distribution, respectively, where $c(x)$ and

$m(x)$ denote the cost and amount of transport associated with a tuple $x = (x_1, \dots, x_T)$. The multi-marginal formulation of (3) reads

$$\text{minimize}_{m \in \mathcal{M}_+(\mathcal{X})} \int_{\mathcal{X}} m(x) c(x) dx + \sum_{t=1}^T \|G(P^{(t)}(m)) - \mathbf{R}^{(t)}\|_2^2, \quad (4)$$

where the cost function decouples into pairwise interactions,

$$c(x_0, \dots, x_T) = \sum_{t=1}^{T-1} c(x_t, x_{t+1}), \quad (5)$$

and $P^{(t)}(m) \in \mathcal{M}_+(X)$ denotes projections of the transport plan, defined as

$$P^{(t)}(m) = \int_{X^{T-1}} m(x_1, \dots, x_T) dx_0 \dots dx_{t-1} dx_{t+1} \dots dx_T.$$

Analogously to standard OT problems [17], [18], an approximate solution to (4) can be found by adding an entropic regularization term to the discretized problem [8]. Although discretizing the multi-marginal optimization problem (4) results in a much larger optimization problem than discretizing (3), it turns out that utilizing the structure in the cost (5) reduces the computational complexity drastically to the same order of operations. Moreover, this approach results in sharper estimates of the distributions, which is a desirable property in many scenarios, including DoA estimation [8], [19].

III. PROBLEM FORMULATION

In this work we consider the setting where a small number of targets are emitting broad-band signals that we measure at several different frequencies $\omega_1, \dots, \omega_F$. Thus, the spatial power spectra $\Phi_f(\theta)$, for $f \in \llbracket F \rrbracket$, to be estimated are expected to have supports concentrated on a small set of intervals of angles. Furthermore, as the targets are broad-band, these supports are similar across frequency. Herein, we propose to model this by requiring that the spatial sparsity measure

$$\int_{\Theta} \sup_{f \in \llbracket F \rrbracket} |\Phi_f(\theta)| d\theta \quad (6)$$

is small. In order to find power spectra that are spatially sparse and change smoothly over time, we propose to combine the tracking formulation (4) with the group-sparsity regularizer (6). More precisely, we seek a transport plan $m_f \in \mathcal{M}_+(\mathcal{X})$ for each frequency $f \in \llbracket F \rrbracket$. Note that its projections describe the power spectra at the discrete time instances, $P^{(t)}(m_f) = \Phi_f^{(t)}$ for $f \in \llbracket F \rrbracket$ and $t \in \llbracket T \rrbracket$. Moreover, let G_f and $\mathbf{R}_f^{(t)}$ denote the measurement operator and covariance measurements, defined in (1), at frequency f and time t . Then, we formulate the tracking problem for group-sparse spectra as

$$\begin{aligned} \text{minimize}_{m_f \in \mathcal{M}_+(\mathcal{X})} &\sum_{f=1}^F \sum_{t=1}^T \|G_f(P^{(t)}(m_f)) - \mathbf{R}_f^{(t)}\|_2^2 \\ &+ \alpha \sum_{f=1}^F \int_{\mathcal{X}} m_f(x) c(x) dx \\ &+ \beta \sum_{t=1}^T \int_{\Theta} \sup_{f \in \llbracket F \rrbracket} \{P^{(t)}(m_f)\} d\theta, \end{aligned} \quad (7)$$

²Equivalently, this can be performed as a decomposition using the short-time Fourier transform as is common in audio signal processing.

where the cost is structured as in (5), and $\alpha, \beta > 0$ are parameters that regulate the emphasis on smoothness over time and group-sparsity, respectively.

A. Discretization and entropic regularization

Following previous works [8], [18], [19], we solve (7) by discretizing and regularizing it with an entropic term. We discretize the angle space Θ into N grid points $\theta_1, \dots, \theta_N$. The cost and transport plans are then described by \mathcal{T} -mode tensors $\mathbf{C} \in \mathbb{R}^{N \times \dots \times N}$ and $\mathbf{M}_f \in \mathbb{R}_+^{N \times \dots \times N}$, where the elements are defined as $\mathbf{C}_{i_1, \dots, i_{\mathcal{T}}} = c(\theta_{i_1}, \dots, \theta_{i_{\mathcal{T}}})$ and similarly for \mathbf{M}_f . The discrete projection operator is defined as

$$[P^{(t)}(\mathbf{M})]_{i_t} = \sum_{i_1, \dots, i_{t-1}, i_{t+1}, \dots, i_{\mathcal{T}}} \mathbf{M}_{i_1, \dots, i_{t-1}, i_t, i_{t+1}, \dots, i_{\mathcal{T}}},$$

and $P^{(t)}(\mathbf{M}) \in \mathbb{R}^N$ is the discretization of the spectrum $\Phi_f^{(t)}$. Thus, the discrete version of the group-sparsity term (6) reads

$$\| [P^{(t)}(\mathbf{M}_1), \dots, P^{(t)}(\mathbf{M}_F)] \|_{\infty, 1} = \sum_{i=1}^N \sup_{f \in [F]} |[P^{(t)}(\mathbf{M}_f)]_i|.$$

Moreover, let $\mathbf{r}_f^{(t)} = \left[\text{vec}(\text{Re}(\mathbf{R}_f^{(t)}))^T \quad \text{vec}(\text{Im}(\mathbf{R}_f^{(t)}))^T \right]^T$ and let \mathbf{G}_f be the discrete counterpart of G_f . This lets us formulate the discretized and regularized group-sparse OT problem

$$\begin{aligned} \underset{\mathbf{M}_f \in \mathbb{R}_+^{N \times \dots \times N}}{\text{minimize}} \quad & \sum_{f=1}^F \langle \mathbf{C}, \mathbf{M}_f \rangle + \epsilon D(\mathbf{M}_f) \\ & + \sum_{t=1}^{\mathcal{T}} \sum_{f=1}^F \frac{\gamma}{2} \left\| \mathbf{r}_f^{(t)} - \mathbf{G}_f P^{(t)}(\mathbf{M}_f) \right\|_2^2 \\ & + \eta \sum_{t=1}^{\mathcal{T}} \| [P^{(t)}(\mathbf{M}_1), \dots, P^{(t)}(\mathbf{M}_F)] \|_{\infty, 1}, \end{aligned} \quad (8)$$

where we for convenience of the following exposition set $\gamma = 2/\alpha$, $\eta = \beta/\alpha$, and

$$D(\mathbf{M}) = \sum_{i_1, \dots, i_{\mathcal{T}}=1}^N (\mathbf{M}_{i_1, \dots, i_{\mathcal{T}}} \log(\mathbf{M}_{i_1, \dots, i_{\mathcal{T}}}) - \mathbf{M}_{i_1, \dots, i_{\mathcal{T}}} + 1)$$

is an entropic regularization term, and $\epsilon > 0$ is a small regularization parameter.

IV. METHOD

We solve the discretized and regularized problem (8) by a dual block coordinate descent, following the approach in [8], [18], [19].

A. Dual problem

Theorem 1: The unique optimal transport plans \mathbf{M}_f are represented as

$$\mathbf{M}_f = \mathbf{U}_f \odot \mathbf{K} \odot \mathbf{V}_f, \quad f \in [F] \quad (9)$$

where

$$\begin{aligned} \mathbf{U}_f &= \mathbf{u}_f^{(1)} \otimes \dots \otimes \mathbf{u}_f^{(T)} \text{ with } \mathbf{u}_f^{(t)} = \exp\left(\frac{1}{\epsilon} \mathbf{G}_f^T \boldsymbol{\lambda}_f^{(t)}\right) \\ \mathbf{V}_f &= \mathbf{v}_f^{(1)} \otimes \dots \otimes \mathbf{v}_f^{(T)} \text{ with } \mathbf{v}_f^{(t)} = \exp\left(\frac{1}{\epsilon} \boldsymbol{\psi}_f^{(t)}\right) \\ \mathbf{K} &= \exp\left(-\frac{1}{\epsilon} \mathbf{C}\right) \end{aligned}$$

where exponentiation is applied element-wise, and where $\boldsymbol{\lambda}_f^{(t)}$ and $\boldsymbol{\psi}_f^{(t)}$ solve the dual problem

$$\begin{aligned} \underset{\boldsymbol{\lambda}_f^{(t)}, \boldsymbol{\psi}_f^{(t)}}{\text{minimize}} \quad & \sum_{f=1}^F \epsilon \langle \mathbf{U}_f, \mathbf{V}_f \odot \mathbf{K} \rangle + \sum_{t=1}^{\mathcal{T}} \sum_{f=1}^F \frac{1}{2\gamma} \left\| \boldsymbol{\lambda}_f^{(t)} - \gamma \mathbf{r}_f^{(t)} \right\|_2^2 \\ \text{subject to} \quad & \left\| [\boldsymbol{\psi}_1^{(t)}, \dots, \boldsymbol{\psi}_F^{(t)}] \right\|_{1, \infty} \leq \eta, \quad t \in [\mathcal{T}], \end{aligned} \quad (10)$$

where $\|\cdot\|_{1, \infty}$ is the dual norm³ of $\|\cdot\|_{\infty, 1}$.

Proof: Similarly to [20, Section 4.3], problem (8) can be written as one graph-structured multi-marginal OT problem. The result can then be derived from [20, Theorem 3.6]. In the appendix we provide a direct alternative proof. ■

Note that in the setting without sparsity penalty term in (8), i.e., $\eta = 0$, problem (10) decouples into F problems of the same form as the dual of the narrow-band tracking formulation in [8]. The main algorithmic contribution of our work thus concerns the treatment of the dual variables $\boldsymbol{\psi}_f^{(t)}$. We note that the first term in the objective of the dual (10) can be expressed in terms of the components in (9).

Lemma 1: With the decomposition (9) and cost function according to (5), we can for any $t = 1, \dots, \mathcal{T}$ write

$$\langle \mathbf{U}_f, \mathbf{V}_f \odot \mathbf{K} \rangle = \|\mathbf{u}_f^{(t)} \odot \mathbf{v}_f^{(t)} \odot \boldsymbol{\xi}_f^{(t)}\|_1,$$

where $\boldsymbol{\xi}_f^{(t)} = \mathbf{w}_f^{(t)} \odot \hat{\mathbf{w}}_f^{(t)}$, and

$$\hat{\mathbf{w}}_f^{(t)} \leftarrow \begin{cases} \mathbf{1} & , t = 1 \\ K^T \left(\mathbf{u}_f^{(t-1)} \odot \mathbf{v}_f^{(t-1)} \odot \hat{\mathbf{w}}_f^{(t-1)} \right) & , t > 1 \end{cases}, \quad (11)$$

$$\mathbf{w}_f^{(t)} \leftarrow \begin{cases} K \left(\mathbf{u}_f^{(t+1)} \odot \mathbf{v}_f^{(t+1)} \odot \mathbf{w}_f^{(t+1)} \right) & , t < \mathcal{T} \\ \mathbf{1} & , t = \mathcal{T} \end{cases}, \quad (12)$$

where $K \in \mathbb{R}_+^{N \times N}$ with elements $K_{ij} = \exp(-c(\theta_i, \theta_j)/\epsilon)$.

Proof: The results follows similarly to [8, Proof of Proposition 2]. ■

B. Algorithm

We propose to solve (8) by the means of a block coordinate descent in its dual problem (10). That is, we iteratively optimize (10) with respect to one set of variables, while keeping the other variables fixed. More precisely, we iterate the following steps

1) For $t \in [\mathcal{T}]$ and $f \in [F]$ let

$$\boldsymbol{\lambda}_f^{(t)} \leftarrow \arg \min_{\boldsymbol{\lambda}} \epsilon \langle e^{\mathbf{G}_f^{(t)T} \boldsymbol{\lambda}/\epsilon}, \mathbf{v}_f^{(t)} \odot \boldsymbol{\xi}_f^{(t)} \rangle + \frac{1}{2\gamma} \|\boldsymbol{\lambda}\|_2^2 - \langle \boldsymbol{\lambda}, \mathbf{r}_f^{(t)} \rangle \quad (13)$$

³ $\|\mathbf{A}\|_{1, \infty} = \max_{\ell} \sum_k |\mathbf{A}_{k, \ell}|$, for a matrix \mathbf{A} .

2) For $t \in \llbracket \mathcal{T} \rrbracket$ let $\{\psi_1^{(t)}, \dots, \psi_F^{(t)}\}$ be the solution of

$$\begin{aligned} & \underset{\psi_1, \dots, \psi_F}{\text{minimize}} \quad \sum_{f=1}^F \epsilon \langle e^{\psi_f / \epsilon}, \mathbf{u}_f^{(t)} \odot \boldsymbol{\xi}_f^{(t)} \rangle \\ & \text{subject to} \quad \|\llbracket \psi_1, \dots, \psi_F \rrbracket\|_{1, \infty} \leq \eta. \end{aligned} \quad (14)$$

Note that before each step the vector $\boldsymbol{\xi}_f^{(t)}$ must be computed as described in Lemma 1. Since iteratively updating the dual variables according to (13) and (14) is a block coordinate descent, and the dual problem (10) satisfies the assumptions of [21, Theorem 2.1]⁴, the iterates converge linearly to the optimal solution of (10). In the limit point, the optimal solution to the primal (8) can be constructed as described in (9). It turns out that the optimization problems (13) and (14) can be solved efficiently. First, we note that the minimizer $\boldsymbol{\lambda}$ of (13) solves

$$\mathbf{G}_f \left(\exp \left(\frac{1}{\epsilon} \mathbf{G}_f^T \boldsymbol{\lambda} \right) \odot \mathbf{v}_f^{(t)} \odot \boldsymbol{\xi}_f^{(t)} \right) + \frac{1}{\gamma} \boldsymbol{\lambda} - \mathbf{r}_f^{(t)} = 0, \quad (15)$$

which we solve by a Newton's method, as proposed in [8]. Secondly, (14) can be solved as described in the following.

Theorem 2: The solution $\{\psi_1^{(t)}, \dots, \psi_F^{(t)}\}$ of (14) can be constructed for each index $i = 1, \dots, N$ separately by performing the following steps.

- 1) Let $\mathbf{z} = \left[(\mathbf{u}_1^{(t)})_i (\boldsymbol{\xi}_1^{(t)})_i, \dots, (\mathbf{u}_F^{(t)})_i (\boldsymbol{\xi}_F^{(t)})_i \right]^T \in \mathbb{R}_+^F$.
- 2) Sort the vector \mathbf{z} in ascending order.
- 3) Identify f^* such that $g(z_{f^*}) > 0$ and $g(z_{f^*+1}) \leq 0$, where

$$g(z) \triangleq -(F - k) \log z - \frac{\eta}{\epsilon} + \sum_{\ell=k+1}^F \log z_\ell. \quad (16)$$

- 4) For $f \in \llbracket F \rrbracket$, let

$$(\psi_f^{(t)})_i = -\epsilon \max \left(0, \log z_f + \frac{1}{F - f^*} \left(\frac{\eta}{\epsilon} - \sum_{\ell=f^*+1}^F \log z_\ell \right) \right).$$

Proof: See appendix. ■

The full method is summarized in Algorithm 1. The algorithm sweeps forward and backwards through the time index $t \in \llbracket \mathcal{T} \rrbracket$. By storing previous results for the vectors $\hat{\mathbf{w}}_f^{(t)}$ and $\mathbf{w}_f^{(t)}$, in each iteration only one of these vectors has to be updated for all $f \in \llbracket F \rrbracket$. This requires F matrix-vector multiplications as in (11)-(12), where the matrix is of size $N \times N$, and is thus of complexity $\mathcal{O}(FN^2)$. The update of $\boldsymbol{\lambda}_f^{(t)}$ requires finding the root of (15) by Newton's method. We observe that after a few outer Sinkhorn iterations, the Newton method typically converges within one step, and thus requires solving only one system of Q^2 linear equations. Finally, for the updates of $\psi_f^{(t)}$ we need to perform the steps listed in Theorem 2 for the N elements in $\psi_f^{(t)}$. The most computationally expensive operation here is the sorting in step 2) which requires $\mathcal{O}(F \log F)$ operations. One update of $\psi_f^{(t)}$ for a given t and $f \in \llbracket F \rrbracket$ has thus complexity $\mathcal{O}(NF \log F)$.

⁴The theorem requires standard assumptions on the optimization problem, e.g., strict convexity.

Algorithm 1 Group-sparse multi-marginal OT

```

while Not converged do
  for  $t \in \{1, \dots, \mathcal{T}\}$  do
    if forward sweep then
      Update  $\hat{\mathbf{w}}_f^{(t)}$  for  $f \in \llbracket F \rrbracket$  as in (11)
    else
      Update  $\mathbf{w}_f^{(t)}$  for  $f \in \llbracket F \rrbracket$  as in (12)
    end if
     $\boldsymbol{\xi}_f^{(t)} \leftarrow \mathbf{w}_f^{(t)} \odot \hat{\mathbf{w}}_f^{(t)}$ ,  $f \in \llbracket F \rrbracket$ 
     $\boldsymbol{\lambda}_f^{(t)} \leftarrow$  solution to (15),  $f \in \llbracket F \rrbracket$ 
     $\mathbf{u}_f^{(t)} \leftarrow \exp \left( \frac{1}{\epsilon} \mathbf{G}_f^T \boldsymbol{\lambda}_f^{(t)} \right)$ ,  $f \in \llbracket F \rrbracket$ 
     $\psi_f^{(t)}$ ,  $f \in \llbracket F \rrbracket \leftarrow$  as in Theorem 2,
     $\mathbf{v}_f^{(t)} \leftarrow \exp \left( \frac{1}{\epsilon} \psi_f^{(t)} \right)$ ,  $f \in \llbracket F \rrbracket$ 
  end for
end while

```

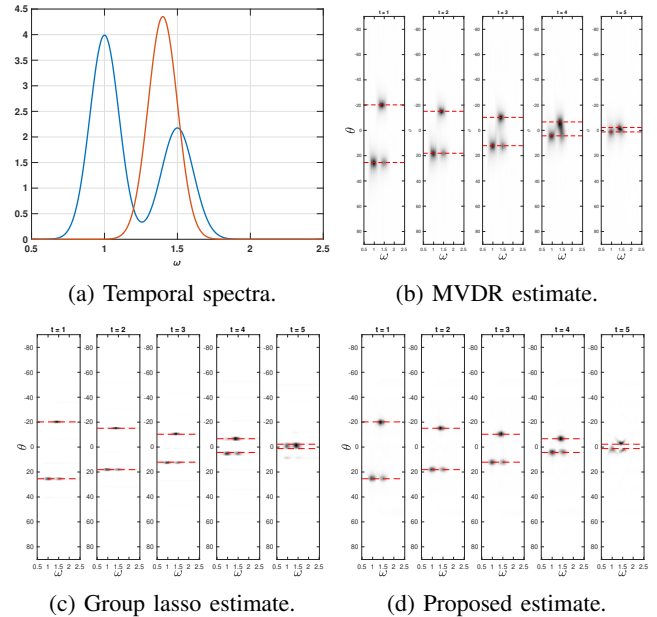


Fig. 1: Estimated spatio-temporal spectra in tracking scenario with two targets.

V. NUMERICAL EXPERIMENTS

In this section, we illustrate the proposed method, and in particular the value of information sharing between frequencies and successive time-points as promoted by the group-sparsity promoting penalty of (8) and the OT distance, respectively. We do this in a simulated scenario as well as for real data measured on a hydrophone array. Throughout, we use a cost function according to (5) with $c(\theta_{i_t}, \theta_{i_{t+1}}) = (\theta_{i_t} - \theta_{i_{t+1}})^2$.

A. Simulated two-target scenario

Consider two broad-band point sources moving in angle space. The (constant) ground truth temporal spectra are shown in Figure 1a. Using a uniform linear array consisting of $Q = 11$ sensors, we for $F = 63$ frequencies uniform on $[0.5, 2.5]$ (in angular frequency) estimate the array covariance matrix at $\mathcal{T} = 5$ time instances by the sample covariance matrix

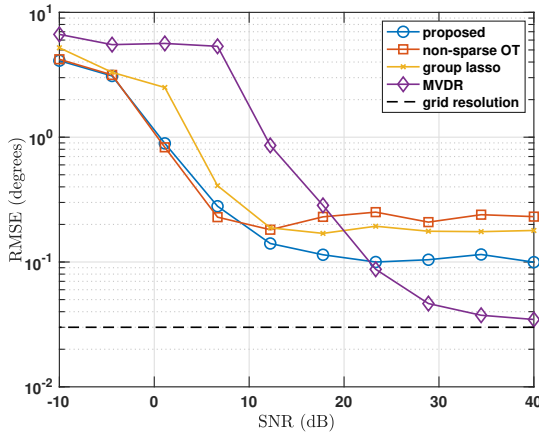


Fig. 2: Target angle estimation error as a function of SNR.

from 200 array snapshots. The array signals are contaminated by spatially and temporally white Gaussian noise as to yield a signal-to-noise ratio of 10 dB. The estimated sequence of spatio-temporal spectra are shown in Figure 1d, where ground-truth source locations are indicated by dashed lines. As can be seen, the proposed method is able to produce estimates indicating localized and well-separated sources. As reference, Figures 1b and 1c show estimates produced by the standard minimum-variance distortionsless response (MVDR/Capon) spatial spectral estimator [22], and non-negative group lasso (with each group being the frequencies corresponding to a spatial angle), respectively. Note here that the MVDR estimate treats each time instance and frequency independently, whereas the group lasso fuses information across frequency by means of its sparsity penalty. It may be noted that neither of the reference methods manages to accurately separate the targets in angle and frequency at $t = 5$ due to the limited array aperture and the frequency overlap of the sources.

B. Localization accuracy

For the same scenario, we study the accuracy of the target angle estimates as a function of the sensor noise. Figure 2 shows the (root) means squared error (RMSE) for the angle estimates at time point $t = 4$, averaged over the two targets, for varying SNR. For each SNR, 50 Monte Carlo simulations are performed, where the target angles are perturbed randomly as to avoid biasing effects caused by the discrete grid. The RMSE is displayed for the proposed method, as well as for the MVDR estimator, and the method from [8] that does not include any sparsity-promoting penalty. For all methods, the angle estimates are determined as the peaks of the spatial spectrum, i.e., the spatio-temporal spectrum averaged over frequency. As can be seen, the information sharing induced by the sparsity-promoting penalty leads to more accurate estimates as compared to only using the OT dynamics. It may also be noted that the OT-based methods incur a small bias, visible for the lowest noise levels, due to the tying together of consecutive time-points. For higher levels of noise, this is outweighed by the increased robustness.

C. Hydrophone array measurements

We here consider a real-world example with monitoring of a scene using an $Q = 8$ element non-uniform linear hydrophone

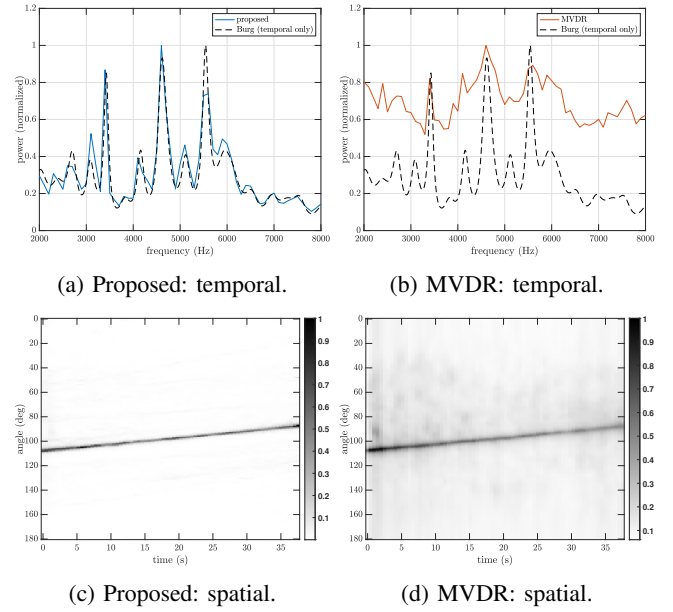


Fig. 3: Spectral estimates for hydrophone array data.

array with a total aperture of 2.08 meters. The data consists of a 35 seconds long recording, with a sampling frequency of 32 kHz. The signal source is a surface vessel moving in shallow water. We construct the band-pass signals by means of the STFT using a Hann window of length 0.2 seconds with 50% overlap. The array covariance matrix is estimated in each frame using exponential averaging, resulting in a sequence of $\mathcal{T} = 73$ observation time points. We apply the proposed method using $F = 61$ frequencies in the interval 2 kHz – 8 kHz. The array response vectors \mathbf{a}_f are constructed under the assumption of free-field propagation and targets in the far-field, with an assumed speed of sound in water of 1480 m/s. The resulting estimates are shown in Figure 3. Here, the estimated spatial spectrum over time (averaged over the frequencies) is shown in Figure 3c, whereas the estimated (assumed stationary) temporal spectrum is shown in Figure 3a. As can be seen, the spatial spectrum is well resolved, showing a single target trajectory. Overlaid in Figure 3a is the Burg temporal spectral estimate, computed from one of the hydrophone channels under assumption of stationarity. As can be seen, the proposed estimate coincides well with the Burg estimate. It may here be noted that no annotated ground truth for the data is available. The corresponding MVDR estimates are shown in Figures 3b and 3d. As can be seen, the spatial spectrum is less well-resolved and contains more noise, and the corresponding temporal spectrum does not give an accurate representation of the signal’s frequency content.

VI. CONCLUSION

In this work, we modeled multi-sensor broad-band signals by means of the concept of a spatio-temporal spectrum, and proposed to track the evolution of time-varying spatio-temporal spectra by a group-sparse optimal transport formulation. This allows us to fuse information across both separate time-instances and across frequency, and numerical experiments show that we achieve accurate estimates of the frequency content and location of broad-band signal sources.

APPENDIX

A. Proof of Theorem 1

Introduce auxiliary variables $\Delta_f^{(t)} = \mathbf{G}_f P^{(t)}(\mathbf{M}_f) - \mathbf{r}_f^{(t)}$, $\Phi_f^{(t)} = P^{(t)}(\mathbf{M}_f)$, $f \in [F]$, $t \in [T]$, to define the Lagrangian

$$\begin{aligned} & \sum_{f=1}^F \langle \mathbf{C}, \mathbf{M}_f \rangle + \epsilon D(\mathbf{M}_f) + \eta \sum_{t=1}^T \left\| [\Phi_1^{(t)}, \dots, \Phi_F^{(t)}] \right\|_{\infty,1} \\ & + \sum_{t=1}^T \sum_{f=1}^F \left(\frac{\gamma}{2} \left\| \Delta_f^{(t)} \right\|_2^2 + (\lambda_f^{(t)})^T \left(\Delta_f^{(t)} - \mathbf{G}_f P^{(t)}(\mathbf{M}_f) + \mathbf{r}_f^{(t)} \right) \right. \\ & \quad \left. + (\psi_f^{(t)})^T \left(\Phi_f^{(t)} - P^{(t)}(\mathbf{M}_f) \right) \right). \end{aligned}$$

Minimizing this Lagrangian with respect to \mathbf{M}_f and $\Delta_f^{(t)}$ yields (9) and $\Delta_f^{(t)} = -\frac{1}{\gamma} \lambda_f^{(t)}$ [8, Proof of proposition 1]. For a fixed t , minimizing the Lagrangian with respect to $\Phi_f^{(t)}$ for $f \in [F]$ requires solving

$$\underset{\Phi_f^{(t)}, f \in [F]}{\text{minimize}} \quad \eta \left\| [\Phi_1^{(t)}, \dots, \Phi_F^{(t)}] \right\|_{1,\infty} + \sum_{f=1}^F (\psi_f^{(t)})^T \Phi_f^{(t)}.$$

Note that this term can be treated separately for each element of the involved vectors. Thus, for each $i \in [N]$, this becomes

$$\begin{aligned} & \inf_{[\Phi_f^{(t)}]_i, f \in [F]} \eta \sup_{f \in [F]} \{ [\Phi_f^{(t)}]_i \} + \sum_{f=1}^F ([\psi_f^{(t)}]_i) [\Phi_f^{(t)}]_i \\ & = \begin{cases} 0, & \text{if } \sum_{f=1}^F ([\psi_f^{(t)}]_i) \leq \eta \\ -\infty, & \text{otherwise.} \end{cases} \end{aligned}$$

Thus, in order for the dual to be bounded, the dual variables $\psi_f^{(t)}$ must satisfy $\left\| [\psi_1^{(t)}, \dots, \psi_F^{(t)}] \right\|_{1,\infty} \leq \eta$. Plugging the optimal \mathbf{M}_f and $\Delta_f^{(t)}$ in the Lagrangian results in the stated dual problem, excluding additive constants.

B. Proof of Theorem 2

Note that problem (14) decouples in the vector indices. For each $i = 1, \dots, N$, we solve a problem of the form

$$\underset{\mathbf{x} \in \mathbb{R}^F}{\text{minimize}} \quad \exp(-\mathbf{x})^T \mathbf{z}, \text{ subject to } \|\mathbf{x}\|_1 \leq p, \quad (17)$$

where $\mathbf{x} = -1/\epsilon [(\psi_1^{(t)})_i, \dots, (\psi_F^{(t)})_i]^T$, $p = \eta/\epsilon$, and \mathbf{z} is defined as in step 1) of the theorem. As \mathbf{z} constructed in this way is non-negative, it is clear that the minimizer of (17) is non-negative and satisfies $\|\mathbf{x}\|_1 = p$. Consider the Lagrangian of (17) with Lagrangian multiplier $\nu > 0$, given by $\exp(-\mathbf{x})^T \mathbf{z} + \nu(\|\mathbf{x}\|_1 - p)$. Let \mathbf{s} be a subgradient of $\|\mathbf{x}\|_1$, then the Lagrangian is minimized with respect to \mathbf{x} if $-\exp(-\mathbf{x}) \odot \mathbf{z} + \nu \mathbf{s} = 0$. From this we can conclude that the optimal \mathbf{x} is elementwise of the form $x_f = \max(\log(z_f) - \log(\nu), 0)$. We now find ν such that the optimal \mathbf{x} satisfies $\|\mathbf{x}\|_1 = p$. Therefore, note that the function

$$g(\nu) = \|\mathbf{x}(\nu)\|_1 - p = -p + \sum_{f=1}^F \max(\log(z_f) - \log(\nu), 0)$$

is continuous piecewise differentiable with non-differentiable points in z_1, \dots, z_F . Moreover, g is strictly decreasing and thus has a root in the interval $(0, \max\{z_f, f \in [F]\})$. Letting the elements in \mathbf{z} be sorted in ascending order, the root lies in

an interval $[z_f, z_{f+1}]$ for which $g(z_f) > 0$ and $g(z_{f+1}) \leq 0$. As g in this interval is given by (16), we get the root as

$$\nu = \exp \left(\frac{1}{F-f} \left(-c + \sum_{\ell=f+1}^F \log(z_\ell) \right) \right).$$

Step 4) in the theorem follows from plugging the root into $x_f = \max(\log(z_f) - \log(\nu), 0)$.

REFERENCES

- [1] Y. Trachi, E. Elbouchikhi, V. Choqueuse, and M. E. H. Benbouzid, "Induction machines fault detection based on subspace spectral estimation," *IEEE Trans. Ind. Electron.*, vol. 63, no. 9, pp. 5641–5651, 2016.
- [2] S. Gannot, E. Vincent, S. Markovich-Golan, and A. Ozerov, "A Consolidated Perspective on Multimicrophone Speech Enhancement and Source Separation," *IEEE/ACM Trans. Audio Speech Lang. Process.*, vol. 25, no. 4, pp. 692–730, 2017.
- [3] U. Grenander and G. Szegő, *Toeplitz Forms and Their Applications*. Los Angeles: University of California Press, 1958.
- [4] F. Elvander and J. Karlsson, "Variance Analysis of Covariance and Spectral Estimates for Mixed-Spectrum Continuous-Time Signals," *IEEE Trans. Signal Process.*, vol. 71, pp. 1395–1407, 2023.
- [5] H. L. Van Trees, *Detection, Estimation, and Modulation Theory, Part IV, Optimum Array Processing*. John Wiley and Sons, Inc., 2002.
- [6] J. F. Böhme, "Estimation of Spectral Parameters of Correlated Signals in Wavefields," *Signal Processing*, vol. 10, pp. 329–337, 1986.
- [7] S. Weiss and I. K. Proudler, "Comparing efficient broadband beamforming architectures and their performance trade-offs," in *14th Int. Conf. Digital Signal Process.*, 2002, pp. 417–423.
- [8] F. Elvander, I. Haasler, A. Jakobsson, and J. Karlsson, "Multi-marginal optimal transport using partial information with applications in robust localization and sensor fusion," *Signal Process.*, vol. 171, no. 107474, 2020.
- [9] —, "Tracking and sensor fusion in direction of arrival estimation using optimal mass transport," in *26th European Signal Process. Conf.*, 2018, pp. 1617–1621.
- [10] T. T. Georgiou, J. Karlsson, and M. S. Takyar, "Metrics for power spectra: an axiomatic approach," *IEEE Trans. Signal Process.*, vol. 57, no. 3, pp. 859–867, 2009.
- [11] Y. Chen, T. T. Georgiou, and M. Pavon, "Optimal transport in systems and control," *Annu. Rev. Control Robot. Auton. Syst.*, vol. 4, pp. 89–113, 2021.
- [12] V. Krishnan and S. Martínez, "Distributed optimal transport for the deployment of swarms," in *IEEE Conf. Decis. Control*, 2018, pp. 4583–4588.
- [13] I. Haasler, J. Karlsson, and A. Ringh, "Control and estimation of ensembles via structured optimal transport," *IEEE Control Syst. Mag.*, vol. 41, no. 4, pp. 50–69, 2021.
- [14] L. Aolaritei, N. Lanzetti, H. Chen, and F. Dörfler, "Distributional uncertainty propagation via optimal transport," *arXiv preprint arXiv:2205.00343*, 2022.
- [15] G. Peyré and M. Cuturi, "Computational optimal transport: With applications to data science," *Foundations and Trends® in Machine Learning*, vol. 11, no. 5-6, pp. 355–607, 2019.
- [16] C. Villani, *Topics in optimal transportation*. American Mathematical Soc., 2021, vol. 58.
- [17] M. Cuturi, "Sinkhorn distances: Lightspeed computation of optimal transport," in *Adv Neural Inf Process Syst*, 2013, pp. 2292–2300.
- [18] J.-D. Benamou, G. Carlier, M. Cuturi, L. Nenna, and G. Peyré, "Iterative bregman projections for regularized transportation problems," *SIAM J. Sci. Comput.*, vol. 37, no. 2, pp. A1111–A1138, 2015.
- [19] I. Haasler, A. Ringh, Y. Chen, and J. Karlsson, "Multimarginal optimal transport with a tree-structured cost and the Schrödinger bridge problem," *SIAM J. Control Optim.*, vol. 59, no. 4, pp. 2428–2453, 2021.
- [20] A. Ringh, I. Haasler, Y. Chen, and J. Karlsson, "Graph-structured tensor optimization for nonlinear density control and mean field games," *arXiv preprint arXiv:2112.05645*, 2021.
- [21] Z.-Q. Luo and P. Tseng, "On the convergence of the coordinate descent method for convex differentiable minimization," *J. Optim. Theory Appl.*, vol. 72, no. 1, pp. 7–35, 1992.
- [22] J. Capon, "High Resolution Frequency Wave Number Spectrum Analysis," *Proc. IEEE*, vol. 57, pp. 1408–1418, 1969.