Attributing Aliasing Artifacts in Dirty Intensity Fields to their Parent Source

AUTHOR: MATTHIEU SIMEONI[1,2,*]
SUPERVISORS: PAUL HURLEY[2], **VICTOR PANARETOS**[1], **MARTIN VETTERLI**[1]

[1] *École Polytechnique Fédérale de Lausanne (EPFL), CH-1015 Lausanne, Switzerland*
[2] *IBM Zurich Research Laboratory, CH-8803 Rüschlikon, Switzerland*
[*] *Contact: matthieu.simeoni@epfl.ch*

Last Modification: February 1, 2016

Abstract:

Estimating intensity fields of stochastic phenomenons is of crucial interest in many scientific applications. Typical experimental setups involve an acquisition system, that subsequently filters and samples the probed intensity field. This equivalently defines a sampling operator, fully specified by the characteristics of the measuring device. Reconstruction of the intensity field can then be achieved by performing an interpolation step on the collected samples, with an interpolation operator *ideally matched* with the sampling operator. Sampling followed by interpolation can be shown to act as an orthogonal projection onto the unknown intensity field. For this reason, the reconstructed intensity field can in general differ quite substantially from the actual one, polluted by aliasing artifacts that complicate and often forbid the identification of features within it. Nonlinear algorithms such as CLEAN (acclamated in radio astronomy) have been proposed to recover the actual features from the aliased intensity field, but their convergence properties have not yet been fully assessed. In this work, we propose a novel method to locate sources within the recovered intensity field and attribute aliasing artifacts to their parent source, in the specific case of point source intensity fields. Rather than directly estimating the intensity field, we first reconstruct the underlying random field it characterises. We then estimate the second moment of this reconstructed random field, including its covariance function. We use the covariance function as a measure of resemblance of different parts of the intensity field. The problem is then formulated as a clustering problem on a graph: the intensity field is sampled on a fine grid, each grid point is interpreted as a node on a graph, and edges between nodes are weighted proportionally to their covariance. We then use spectral clustering to separate the artifacts in groups, and identify the actual sources in the field as the nodes with maximum intensity within each cluster. We conclude with an application of our method to the field of radio astronomy.

Keywords: *Intensity field, source detection, artifacts attribution, spectral clustering, aliasing, sampling, interpolation, CLEAN.*

1. Introduction

Many scientific applications [1, 2, 3] involve estimating the intensity field of a physical phenomenon modelled as a continuous stationary random field. In radio astronomy for example [2], the sky is modelled as a random field composed of uncorrelated point sources at unknown locations, emitting narrowband signals with random complex amplitudes [4, 5]. The goal is then to estimate the intensity field of this stochastic process, in order to recover the locations and intensities of the underlying sources. Similarly, in positron emission tomography (PET) the brain's metabolic activity is indirectly mapped by estimating the intensity field of positron emissions, modelled as a Poisson random point field [3].

In practice, data is collected through an acquisition system, which typically sense the unknown random field by subsequently filtering and sampling it [6, 7]. Associated with this acquisition system is a sampling operator, completely specified by the characteristics of the measuring device. When only the intensity field (variance of the random field at each point in the domain) is of interest, the whole system is often reinterpreted as acting on the intensity field rather than on the random field itself. Corresponding sampling and interpolation operators can equivalently be derived.

A natural reconstruction of the intensity field consists then in interpolating the sampled measurements with the interpolation operator *ideally matched* with the sampling operator [6]. In mathematical terms, the interpolation operator is the *generalized pseudoinverse* of the sampling operator.

Except from very specific input random fields, this process is in general lossy, as the composition of the sampling and interpolation operators can be shown to form an orthogonal projection operator [6]. Hence, the interpolated intensity field can in practice substantially differ from the actual one, potentially exhibiting strong aliasing artifacts. This is particularly an issue when one is interested in identifying features within the estimated intensity map. For example, if the true intensity field consisted in a set of point sources with various intensities, the estimated intensity field could exhibit complex extended features spanning over the entire domain, hence forbidding any direct identification of the sources. For this reason, the estimated intensity map is often called the *dirty map*. Deconvolution algorithms such as CLEAN [8, 9] in radio astronomy have been proposed to remove the interpolation artifacts, but they are heavily nonlinear and their convergence properties are not well understood yet [8, 5].

In this paper, we propose a novel method to locate sources within the recovered intensity field and attribute aliasing artifacts to their parent source, in the specific case of point source intensity fields. Rather than directly estimating the intensity field, we first reconstruct the underlying random field it characterises. The proposed method performs then in three subsequent steps. First, the random fields covariance function is estimated from multiple observations of the reconstructed random field. Then, we sample this covariance function on a grid over the domain of interest, and interpret each grid point as a node on a graph. Edges between the nodes are weighted according to the absolute value of their covariance. We exploit this graph structure, to cluster the artifacts together by means of spectral clustering. The true sources are finally identified as the nodes with maximum degree within each cluster. We conclude our study by demonstrating the benefits of our method for the specific example of estimating sources from dirty images in radio astronomy.

2. Data Model

In this section we introduce a special class of random fields, the *point source random fields*, for which we will have a special interest in all that follows. Then, we derive a model for the data collected by the acquisition system, and list the assumptions on which it is based. Finally, we show how one can define a virtual acquisition system and corresponding data model when only the intensity field is of interest.

2.1. Point Source Random Field

Even though the method described in this work could easily be extended to a wider class of random fields, we will restrict here our attention to the specific case of point source random fields, widely used in practice to model physical phenomena. We introduce in this section this specific class of random fields.

First, we need to formally define what we (loosely) call a *random field*, and which we should more accurately (but undoubtefully less conveniently) call a *continuous-spatio-temporal random field*.

Definition 2.1 — ***Continous-spatio-temporal Random Field.*** *Given some probability space $(\Omega, \mathcal{F}, \mathbb{P})$, a* ***$\mathbb{C}$-valued continuous-spatio-temporal random field*** *$\mathcal{S}$ is a collection of random variables $S(t, \boldsymbol{r}) : \Omega \to \mathbb{C}$, indexed both by a time parameter t and space parameter $\boldsymbol{r}$, taking values in some continuous sets $\mathcal{I}$ and $\mathcal{M}$:*

$$\mathcal{S} = \{S(t, \boldsymbol{r}) : \Omega \to \mathbb{C},\ (t, \boldsymbol{r}) \in \mathcal{I} \times \mathcal{M}\},$$

where $\mathcal{I} \subseteq \mathbb{R}$ is an interval of the real line and $\mathcal{M} \subseteq \mathbb{R}^n$ a manifold. The set $\mathcal{I} \times \mathcal{M}$ will be referred to as the ***indexing set*** *of the random field $\mathcal{S}$.*

In essence, random fields are extension of stochastic processes to more general indexing sets. A simple example of a continuous-spatio-temporal random field is given by the natural phenomenon governing temporal and spatial temperature fluctuations at the surface of the earth. Indeed, if we model these fluctuations as stochastic events, then for every point in space and time instant one can see a temperature record as the outcome of a certain random variable, hence defining a continuous-spatio-temporal random field. In that case, we have $t \in [0, +\infty[$ (assuming an origin of times) and $\boldsymbol{r} \in \mathbb{S}^2$ (assuming that the earth is a sphere).

In the context of this work, we will only consider random fields which are *strongly stationary through time*, meaning that their distribution does not change with time:

$$S(t_1, \boldsymbol{r}) \overset{dist}{=} S(t_2, \boldsymbol{r}), \quad \forall \boldsymbol{r} \in \mathcal{M},\ \forall (t_1, t_2) \in \mathcal{I}^2.$$

As a consequence of strong stationarity, the following holds:

$$\mathbb{E}[S(t, \boldsymbol{r})] = \mu(\boldsymbol{r}), \quad \forall (t, \boldsymbol{r}) \in \mathcal{I} \times \mathcal{M},$$

and

$$\mathrm{Var}(S(t, \boldsymbol{r})) = I(\boldsymbol{r}), \quad \forall (t, \boldsymbol{r}) \in \mathcal{I} \times \mathcal{M}.$$

We respectively call $\mu : \mathcal{M} \to \mathbb{C}$ and $I : \mathcal{M} \to \mathbb{R}_+$ the *mean field* and *intensity field* of the random field $\mathcal{S}$. Finally, the spatial coherency of the random field is measured by the *covariance function* $\kappa : \mathcal{M}^2 \to \mathbb{C}$ of the field, defined, for every $t \in \mathcal{I}$ and $(\boldsymbol{r}_1, \boldsymbol{r}_2) \in \mathcal{M}^2$ as:

$$\kappa(\boldsymbol{r}_1, \boldsymbol{r}_2) = \mathbb{E}\left[(S(t, \boldsymbol{r}_1) - \mu(\boldsymbol{r}_1))(S(t, \boldsymbol{r}_2) - \mu(\boldsymbol{r}_2))^*\right].$$

Once again, because of strong stationarity we observe that this function is time-invariant.

We are now ready to define a special class of random fields, called *point source random fields*. Roughly speaking, point source random fields can be pictured as streams of Diracs with time-varying random amplitudes.

Definition 2.2 — ***Point Source Random Field.*** *Let $\mathcal{S}$ be some continuous-spatio-temporal random field, with probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and indexing set $\mathcal{I} \times \mathcal{M}$. Then, this random field is called a* ***point source random field*** *if the indexed random variables in $\mathcal{S}$ can be written as*

$$S(t, \boldsymbol{r}) = \sum_{q=1}^{Q} \xi_q(t)\, \delta(\boldsymbol{r} - \boldsymbol{r}_q), \quad \forall (t, \boldsymbol{r}) \in \mathcal{I} \times \mathcal{M},$$

where $Q > 0$ is the number of sources composing the random field and $\{\boldsymbol{r}_1, \dots, \boldsymbol{r}_Q\} \subset \mathcal{M}$ are the locations of these sources. The continuous-time stochastic processes

$$\Xi_q = \{\xi_q(t) : \Omega \to \mathbb{C},\ t \in \mathcal{I}\}, \quad q = 1, \dots, Q,$$

are independent[a], strongly stationary and with indexing set $\mathcal{I} \subset \mathbb{R}$.

[a] We recall here that two stochastic processes $\{X(t),\ t \in T\}$ and $\{Y(t),\ t \in T\}$ are said to be independent if, for any sequence $t_1, \dots, t_n \in T$, $n > 0$, the random vectors $\boldsymbol{X} = (X(t_1), \dots, X(t_n))$ and $\boldsymbol{Y} = (Y(t_1), \dots, Y(t_n))$ are independent.

For a given source q, the associated stochastic process Ξ_q can be interpreted as describing the random fluctuations of the source's complex amplitude through time. We will assume here each stochastic process Ξ_q to be a white noise [10], with *complex, circularly symmetric Gaussian distribution*,

$$\xi_q(t) \sim \mathbb{C}\mathcal{N}(0, \sigma_q^2), \quad \forall t \in \mathcal{I},\ q = 1, \dots, Q.$$

The variance parameter $\sigma_q^2 > 0$ is called the **intensity** of the qth source. With this choice of distribution, the mean and intensity fields of the random field are given by $\mu(\boldsymbol{r}) = 0$ and $I(\boldsymbol{r}) = \sum_{q=1}^{Q} \sigma_q^2 \delta(\boldsymbol{r} - \boldsymbol{r}_q)$, $\forall \boldsymbol{r} \in \mathcal{M}$. Hence, the intensity field of a point source random field consists in a stream of Diracs with various amplitudes given by the respective variances of the stationary stochastic processes $\Xi_1, \dots, \Xi_Q$. Moreover, it is easy to see that $\kappa(\boldsymbol{r}_1, \boldsymbol{r}_2) = 0$, $\forall \boldsymbol{r}_1 \neq \boldsymbol{r}_2 \in \mathcal{M}$, meaning that the random field is *spatially uncorrelated*.

Observe that, having chosen a Gaussian distribution to model the random fluctuations of the sources' amplitudes, the point source random field can be put in bijection with a finite set of parameters

$$\mathcal{S} \leftrightarrow \{(\boldsymbol{r}_q, \sigma_q^2),\ q = 1, \dots, Q\}.$$

Indeed, the specification of the location and intensity parameters $(\boldsymbol{r}_q, \sigma_q^2)$ for the Q sources forming the point source random field is sufficient to fully characterise it. In practice, these parameters (as well as the number of sources Q) are unknown, and the goal will be to estimate them so as to characterise the underlying physical phenomenon.

2.2. The Acquisition System

In experimental setups, evidence about the stochastic phenomenon of interest is collected with an acquisition system [6]. Real-life acquisition systems act on the random field in two subsequent stages.

First, the random field $\mathcal{S}$ is *spatially filtered*, yielding what we call a **measurement field** $\mathcal{Y}$, with indexing set $\mathcal{I} \times \mathcal{P}$:

$$\mathcal{Y} = \{Y(t, \boldsymbol{p}) : \Omega \to \mathbb{C}, \quad (t, \boldsymbol{p}) \in \mathcal{I} \times \mathcal{P}\},$$

where

$$Y(t, \boldsymbol{p}) = \int_{\mathcal{M}} S(t, \boldsymbol{r}) \phi^*(\boldsymbol{r}, \boldsymbol{p})\, d\boldsymbol{r}, \quad \forall (t, \boldsymbol{p}) \in \mathcal{I} \times \mathcal{P},$$

and $\mathcal{P} \subset \mathbb{R}^n$ is the **measuring manifold**. This manifold is often related to the geometry of the acquisition system. For example, if one is measuring a sound field with an array of microphones disposed on a sphere, then the measuring manifold will be a sphere. The filtering kernel $\phi : \mathcal{M} \times \mathcal{P} \to \mathbb{C}$ is known and depends only on the characteristics of the acquisition system used. The measurement field $\mathcal{Y}$ inherits most of the properties of $\mathcal{S}$, except for its covariance function and intensity field that, as we will see, are modified by the filtering step.

The second stage of the acquisition consists in a *sampling* of the measurement field $\mathcal{Y}$, at locations $\boldsymbol{p}_1, \dots, \boldsymbol{p}_J \in \mathcal{P}$. These sampling locations are again often related to the underlying geometry of the measuring device, and can correspond for example to the positions of the various sensors composing the tool in the case of phased arrays. We finally obtain a continuous-time multivariate stochastic process $\boldsymbol{Y}(t) : \Omega \to \mathbb{C}^J$

$$\boldsymbol{Y}(t) = (Y(t, \boldsymbol{p}_1), \dots, Y(t, \boldsymbol{p}_J))^T, \quad \forall t \in \mathcal{I}. \tag{1}$$

Notice that, for every $i = 1, \dots, J$

$$Y(t, \boldsymbol{p}_i) = \int_{\mathcal{M}} S(t, \boldsymbol{r}) \phi^*(\boldsymbol{r}, \boldsymbol{p}_i)\, d\boldsymbol{r} = \langle S(t, \boldsymbol{r}), \phi_i(\boldsymbol{r}) \rangle_{\mathcal{L}^2(\mathcal{M})},$$

where $\phi_i(\boldsymbol{r}) = \phi(\boldsymbol{r}, \boldsymbol{p}_i)$ for $i = 1, \dots, J$. Hence, eq. (1) can be rewritten as

$$\boldsymbol{Y}(t) = \boldsymbol{\Phi}^* S(t, \boldsymbol{r}), \quad \forall t \in \mathcal{I}, \tag{2}$$

where $\boldsymbol{\Phi}^* : \mathcal{L}^2(\mathcal{M}) \to \mathbb{C}^J$ is the *analysis operator* [6] associated to the family $\{\phi_i(\boldsymbol{r}), i = 1, \dots, J\} \subset \mathcal{L}^2(\mathcal{M})$. We will call this operator the **sampling operator** in all that follows.

The data collected by the instrument during a certain time interval can hence be seen as various realisations of this J-dimensional strongly stationary stochastic process. The goal will then be to interpolate these samples of the measurement field with an appropriate interpolation operator, so as to estimate some summary statistics of the underlying random field $\mathcal{S}$.

When only the intensity field of the random field is of interest, it is possible (and sometimes simpler) to adopt a different perspective on the problem, and to define a somewhat virtual acquisition system that would directly act on the intensity field. Consider indeed the covariance matrix $\Sigma \in \mathbb{C}^{J \times J}$ of the multivariate stochastic process $\boldsymbol{Y}(t)$. Then, it is easy to show that, for every $i, k \in \{1, \dots, J\}$

$$\Sigma_{ik} = \mathbb{E}[Y_i(t) Y_k^*(t)] = \int_{\mathcal{M}} I(\boldsymbol{r}) \phi_i^*(\boldsymbol{r}) \phi_k(\boldsymbol{r}) d\boldsymbol{r} = \langle S(t, \boldsymbol{r}), \psi_{ik}(\boldsymbol{r}) \rangle,$$

where $\psi_{ik}(\boldsymbol{r}) = \phi_i(\boldsymbol{r}) \phi_k^*(\boldsymbol{r})$, $\forall \boldsymbol{r} \in \mathcal{M}$. Hence, elements of the covariance matrix Σ can be seen as samples of some measurement field. This (virtual) measurement field is actually the covariance function of the random field $\mathcal{Y}$

$$\begin{aligned} \kappa_{\mathcal{Y}}(\boldsymbol{p}, \boldsymbol{q}) &= \int_{\mathcal{M}} I(\boldsymbol{r}) \underbrace{\phi^*(\boldsymbol{r}, \boldsymbol{p}) \phi(\boldsymbol{r}, \boldsymbol{q})}_{:= \psi^*(\boldsymbol{r}, \boldsymbol{p}, \boldsymbol{q})} d\boldsymbol{r} \\ &= \int_{\mathcal{M}} I(\boldsymbol{r}) \psi^*(\boldsymbol{r}, \boldsymbol{p}, \boldsymbol{q}) d\boldsymbol{r}, \quad \forall p, q \in \mathcal{P}^2. \end{aligned}$$

With this perspective, the filtering kernel is given by $\psi : \mathcal{M} \times \mathcal{P}^2 \to \mathbb{C}$ and the measuring manifold is $\mathcal{P}^2$. Again, we can describe the action of the acquisition system onto the intensity field with a sampling operator. To this end, define $\boldsymbol{V} = \mathrm{vec}(\Sigma) \in \mathbb{C}^{J^2}$, where $\mathrm{vec} : \mathbb{C}^{J \times J} \to \mathbb{C}^{J^2}$ is the classical *vectorisation operator*, that converts a matrix in a vector by stacking the columns on top of one another. Then, we can show that

$$\boldsymbol{V} = \boldsymbol{\Psi}^* I, \tag{3}$$

where $\boldsymbol{\Psi}^* : \mathcal{L}^2(\mathcal{M}) \to \mathbb{C}^{J^2}$ is the analysis operator associated to the family $\{\psi_k(\boldsymbol{r}), k = 1, \dots, J^2\} \subset \mathcal{L}^2(\mathcal{M})$, where the functions are re-labeled and ordered according to the ordering induced by the vec operator on the data. Notice that the two synthesis operators $\boldsymbol{\Psi}$ and $\boldsymbol{\Phi}$ are linked by the following relationship

$$\boldsymbol{\Psi} = \overline{\boldsymbol{\Phi}} \otimes \boldsymbol{\Phi}, \tag{4}$$

where $\overline{\boldsymbol{\Phi}}$ is the conjugate of $\boldsymbol{\Phi}$ and $\otimes$ stands for the Kronecker product.

The formulation eq. (3) is often preferred in practice, as it only involves deterministic quantities and directly links the data with the intensity field to be estimated. In the next section, we will show that both formulations eq. (2) and eq. (3) are actually equivalent when it comes to estimating the intensity field. However, when one is interested in classifying features within the estimated intensity field, then the formulation eq. (2) presents a huge advantage over eq. (3), as it permits to exploit the extra information contained in the covariance function of the estimated random field. This will be the topic of section 4.

3. Recovery of the Intensity Field

As suggested by the preceding developments, two different procedures can be proposed to recover the least squares estimate of the intensity field from the data, depending on whether eq. (2) or eq. (3) is chosen as data model. As we will show in this section, both procedures yield the same intensity field estimate, which is intuitively reassuring. In all that follows, we will assume that both families associated with the analysis operators $\boldsymbol{\Phi}^*$ and $\boldsymbol{\Psi}^*$ are linearly independent, so that all the quantities involved in the computations are well-defined.

When it comes to estimating the intensity field (and *only* the intensity field), then eq. (3) seems like a better choice of a data model, as it explicitly links the samples $\boldsymbol{V} \in \mathbb{C}^{J^2}$ of the measuring field $\kappa_{\mathcal{Y}}$ with the underlying intensity field. Hence, we first describe the procedure resulting from the use of eq. (3) as data model.

The problem of recovering the intensity field from the measurements $\boldsymbol{V} \in \mathbb{C}^{J^2}$ is referred to as *interpolation*. Mathematically speaking, we would like to find a linear operator $\widetilde{\boldsymbol{\Psi}} : \mathbb{C}^{J^2} \to \mathcal{L}^2(\mathcal{M})$, that maps the *finite* measurements set $\boldsymbol{V}$ onto an *infinite*-dimensional object $I_D \in \mathcal{L}^2(\mathcal{M})$, hopefully constituting in a good

approximation of the true intensity field I from which the measurements were obtained. This can indeed be interpreted as an interpolation step, in the sense that a discrete set of samples from the measuring field is mapped to a continuous function [6]. If we choose $\widetilde{\boldsymbol{\Psi}}$ to be equal to the pseudo-inverse of $\boldsymbol{\Psi}$

$$\widetilde{\boldsymbol{\Psi}} = \boldsymbol{\Psi}\left(\boldsymbol{\Psi}^*\boldsymbol{\Psi}\right)^{-1} = \boldsymbol{\Psi}G_{\Psi}^{-1},$$

then we say that the interpolation operator $\widetilde{\boldsymbol{\Psi}}$ is *ideally matched* with the sampling operator $\boldsymbol{\Psi}^*$ (see [6]). The operator $\boldsymbol{\Psi} : \mathbb{C}^{J^2} \to \mathcal{L}^2(\mathcal{M})$, adjoint of $\boldsymbol{\Psi}^*$, is called the *synthesis* operator and is defined, for $\boldsymbol{x} \in \mathbb{C}^{J^2}$, as

$$(\boldsymbol{\Psi}\boldsymbol{x})(\boldsymbol{r}) = \sum_{k=1}^{J^2} x_k \psi_k(\boldsymbol{r}), \quad \forall \boldsymbol{r} \in \mathcal{M}.$$

The matrix $G_{\Psi} = \boldsymbol{\Psi}^*\boldsymbol{\Psi} \in \mathbb{C}^{J^2 \times J^2}$ is called the **Gram matrix** and is defined as

$$(G_{\Psi})_{ij} = \langle \psi_i, \psi_j \rangle_{\mathcal{L}^2(\mathcal{M})} = \int_{\mathcal{M}} \psi_i(\boldsymbol{r})\psi_j^*(\boldsymbol{r})\, d\boldsymbol{r}, \ \forall i,j \in \{1,\dots,J^2\}.$$

This matrix can either be computed numerically or, in some specific scenarios, analytically (see section 5 for an example). The invertibility of the Gram matrix comes from the assumption that the family of functions $\{\psi_k,\ k = 1,\dots,J^2\}$ is linearly independent. To prevent G_{Ψ} from being ill-conditioned (and hence facilitate its numerical inversion), we can further require $\{\psi_k,\ k = 1,\dots,J^2\}$ to be a *Riesz basis* for its span $E_{\Psi} = \text{span}\{\psi_k,\ k = 1,\dots,J^2\}$ (see [6]).

We have then an algorithm to estimate the intensity field from the data:

Algorithm 1 | Recovering the Intensity Field with data model eq. (3)

procedure INTERP1($\boldsymbol{V}$) ▷ *INPUT: Samples $\boldsymbol{V}$ of the covariance function of $\mathcal{Y}$, the measurement field of the instrument*

Compute the Gram matrix $G_{\Psi} \in \mathbb{C}^{J^2 \times J^2}$

$$G_{\Psi} \leftarrow \boldsymbol{\Psi}^*\boldsymbol{\Psi}.$$

Apply the inverse of G_{Ψ} to the samples $\boldsymbol{V} \in \mathbb{C}^{J^2}$

$$\widetilde{\boldsymbol{V}} \leftarrow G_{\Psi}^{-1}\boldsymbol{V}.$$

Interpolate $\widetilde{\boldsymbol{V}}$ with the synthesis operator $\boldsymbol{\Psi}$

$$I_D(\boldsymbol{r}) = (\boldsymbol{\Psi}\widetilde{\boldsymbol{V}})(\boldsymbol{r}) = \sum_{k=1}^{J^2} \widetilde{V}_k \psi_k(\boldsymbol{r}), \quad \forall \boldsymbol{r} \in \mathcal{M}. \tag{5}$$

return $I_D \in \mathcal{L}^2(\mathcal{M})$ ▷ *OUTPUT: L_2-Estimate of the intensity field*

Note that applying first G_{Ψ}^{-1} to the samples $\boldsymbol{V}$ and then using the synthesis operator $\boldsymbol{\Psi}$ to interpolate the modified samples $\widetilde{\boldsymbol{V}}$ is mathematically equivalent to directly interpolating the data $\boldsymbol{V}$ with $\widetilde{\boldsymbol{\Psi}}$, but more convenient for practical implementations. The estimated intensity field $I_D \in \mathcal{L}^2(\mathcal{M})$ obtained with algorithm 1 can be shown [6] to be the orthogonal projection of the actual intensity field I onto the subspace $E_{\Psi} = \text{span}\{\psi_k,\ k = 1,\dots,J^2\}$. As such, we have

$$I_D = \underset{I_E \in E_{\Psi}}{\text{argmin}} \|I - I_E\|_2^2, \text{ and } I - I_D \in E_{\Psi}^{\perp}. \tag{6}$$

Hence, I_D is also the least-squares estimate of I in E_{Ψ}. In a noiseless scenario and without any a priori belief about the intensity field, this is the best approximation we could ever hope for. Indeed, any information that would be contained in the residual intensity field $I_R = I - I_D \in E_{\Psi}^{\perp}$ is anyway discarded by the acquisition system, and hence not available to us:

$$\boldsymbol{V} = \boldsymbol{\Psi}^*(I_D + I_R) = \boldsymbol{\Psi}^* I_D + \underbrace{\boldsymbol{\Psi}^* I_R}_{=0} = \boldsymbol{\Psi}^* I_D,$$

as $I_R \in E_{\Psi}^{\perp} = \mathcal{R}(\boldsymbol{\Psi})^{\perp} = \mathcal{N}(\boldsymbol{\Psi}^*)$. The procedure is hence lossy, and the recovered intensity field may in general significantly differ from the original intensity field. In practice, we often observe artifacts in the estimated intensity field, forbidding features' identification. For this reason, the estimated intensity field is sometimes called the *dirty* intensity field or *dirty* map. In what follows, we will refer to the artifacts as *aliasing* artifacts, by analogy with the aliasing artifacts introduced when interpolating the samples of a non-bandlimited signal.

An alternative estimation procedure can be proposed when eq. (2) is used as data model. In that case, the sampling operator $\boldsymbol{\Phi}^*$ is acting on the random field $\mathcal{S}$ and not on the intensity field. Hence, we first interpolate the samples of the measurement field $\mathcal{Y}$ to estimate the random field, and then compute the intensity field associated with the recovered random field. This yields an alternative procedure to compute the intensity field:

Algorithm 2 | Recovering the Intensity Field with data model eq. (2)

procedure INTERP2($\boldsymbol{Y}(t)$) ▷ *INPUT: Spatial samples $\boldsymbol{Y}(t)$ of the measurement field $\mathcal{Y}$*

Compute the Gram matrix $G_{\Phi} \in \mathbb{C}^{J \times J}$

$$G_{\Phi} \leftarrow \boldsymbol{\Phi}^*\boldsymbol{\Phi}.$$

Apply the inverse of G_{Φ} to the samples $\boldsymbol{Y}(t) : \Omega \to \mathbb{C}^J$

$$\widetilde{\boldsymbol{Y}}(t) \leftarrow G_{\Phi}^{-1}\boldsymbol{Y}(t).$$

Interpolate $\widetilde{\boldsymbol{Y}}(t) : \Omega \to \mathbb{C}^J$ with the synthesis operator $\boldsymbol{\Phi}$

$$\mathcal{S}_D = \left\{ S_D(t,\boldsymbol{r}) = \sum_{k=1}^{J} \widetilde{Y}_k(t)\phi_k(\boldsymbol{r}), \quad \forall (t,\boldsymbol{r}) \in \mathcal{I} \times \mathcal{M} \right\}.$$

Compute the intensity field of $\mathcal{S}_D$

$$I_{\mathcal{S}_D}(\boldsymbol{r}) = \sum_{k,j=1}^{J} \mathbb{E}[\widetilde{Y}_k(t)\widetilde{Y}_j^*(t)]\phi_k(\boldsymbol{r})\phi_j^*(\boldsymbol{r}), \quad \forall \boldsymbol{r} \in \mathcal{M}. \tag{7}$$

return $I_{\mathcal{S}_D} \in \mathcal{L}^2(\mathcal{M})$ ▷ *OUTPUT: L_2-Estimate of the intensity field*

Again, we can show that $\mathcal{S}_D$ is the least-squares estimate of $\mathcal{S}$ in $E_{\Phi} = \text{span}\{\phi_k,\ k = 1,\dots,J\}$, in the sense that, for all $t \in \mathcal{I}$

$$S_D(t,\boldsymbol{r}) = \underset{S_E(\cdot,\boldsymbol{r}) \in E_{\Phi}}{\text{argmin}} \|S(t,\boldsymbol{r}) - S_E(t,\boldsymbol{r})\|_2^2, \text{ and } S - S_D \in E_{\Phi}^{\perp}.$$

In section 4, we will use the covariance function $\kappa_{\mathcal{S}_D} : \mathcal{M}^2 \to \mathbb{C}$ of $\mathcal{S}_D$ in order to classify the artifacts polluting the estimated intensity field and attribute them to their parent source. This covariance function is given by, $\forall (\boldsymbol{r}_1,\boldsymbol{r}_2) \in \mathcal{M}^2$,

$$\kappa_{\mathcal{S}_D}(\boldsymbol{r}_1,\boldsymbol{r}_2) = \text{cov}(S_D(t,\boldsymbol{r}_1), S_D(t,\boldsymbol{r}_2)) = \sum_{k,j=1}^{J} \mathbb{E}[\widetilde{Y}_k(t)\widetilde{Y}_j^*(t)]\phi_k(\boldsymbol{r}_1)\phi_j^*(\boldsymbol{r}_2). \tag{8}$$

Disposing of two alternative procedures to estimate the intensity field, it is natural to wonder if both output the same estimate. This would be intuitively reassuring, but seems like a difficult question to answer, at least just by looking at eq. (5) and eq. (7). Theorem 3.2 establishes this result. The proof of this theorem relies on the following lemma.

Lemma 3.1 *Let $\mathbf{\Phi} : \mathbb{C}^J \to \mathcal{L}^2(\mathcal{M})$ and $\mathbf{\Psi} : \mathbb{C}^{J^2} \to \mathcal{L}^2(\mathcal{M})$ be two synthesis operators, such that $\mathbf{\Psi} = \overline{\mathbf{\Phi}} \otimes \mathbf{\Phi}$. Consider further the Gram matrices $G_\Phi \in \mathbb{C}^{J\times J}$ and $G_\Psi \in \mathbb{C}^{J^2\times J^2}$ associated with these operators. We assume G_Φ and G_Ψ to be invertible. Then, we have*

$$G_\Psi^{-1} = G_\Phi^{-T} \otimes G_\Phi^{-1}.$$

■ **Proof 3.1** The proof relies on properties of the Kronecker product. Indeed, we have

$$\begin{aligned} G_\Psi &= \mathbf{\Psi}^*\mathbf{\Psi} \\ &= \left(\overline{\mathbf{\Phi}} \otimes \mathbf{\Phi}\right)^* \left(\overline{\mathbf{\Phi}} \otimes \mathbf{\Phi}\right) \\ &= \left(\overline{\mathbf{\Phi}}^* \otimes \mathbf{\Phi}^*\right) \left(\overline{\mathbf{\Phi}} \otimes \mathbf{\Phi}\right) \\ &= \overline{\left(\mathbf{\Phi}^*\mathbf{\Phi}\right)} \otimes \left(\mathbf{\Phi}^*\mathbf{\Phi}\right) \\ &= G_\Phi^T \otimes G_\Phi, \end{aligned}$$

as $G_\Phi = \mathbf{\Phi}^*\mathbf{\Phi}$ is Hermitian symmetric, and hence $\overline{G_\Phi} = G_\Phi^T$. Moreover, another property of the Kronecker product tells us that $G_\Psi = G_\Phi^T \otimes G_\Phi$ is invertible if and only if G_Φ^T and G_Φ are invertible, in which case we have

$$G_\Psi^{-1} = G_\Phi^{-T} \otimes G_\Phi^{-1}.$$

As G_Ψ and G_Φ are assumed invertible, this concludes the proof. ■

We are now ready to state the theorem:

Theorem 3.2 Let $I_D \in \mathcal{L}^2(\mathcal{M})$ and $I_{\mathcal{S}_D} \in \mathcal{L}^2(\mathcal{M})$ be the respective outputs of algorithms 1 and 2, estimates of some intensity field I. Then, we have equality of the two estimates

$$I_D(\boldsymbol{r}) = I_{\mathcal{S}_D}(\boldsymbol{r}), \quad \forall \boldsymbol{r} \in \mathcal{M}.$$

Hence, $I_{\mathcal{S}_D}$ is also the least-squares estimate of I in $E_\Psi = \text{span}\{\psi_k,\, k = 1, \dots, J^2\}$, and we have

$$I_{\mathcal{S}_D} = \underset{I_E \in E_\Psi}{\text{argmin}} \|I - I_E\|_2^2, \quad \text{and} \quad I - I_{\mathcal{S}_D} \in E_\Psi^\perp.$$

■ **Proof 3.2** Let us first rewrite eq. (5):

$$I_D = \mathbf{\Psi}\widetilde{\boldsymbol{V}} = \widetilde{\mathbf{\Psi}}\boldsymbol{V} = \mathbf{\Psi} G_\Psi^{-1} \text{vec}(\Sigma) = \left(\overline{\mathbf{\Phi}} \otimes \mathbf{\Phi}\right) \left(G_\Phi^{-T} \otimes G_\Phi^{-1}\right) \text{vec}(\Sigma), \tag{9}$$

where $\boldsymbol{V} = \text{vec}(\Sigma) \in \mathbb{C}^{J^2}$ and $\Sigma \in \mathbb{C}^{J\times J}$ is the covariance matrix of the multivariate stochastic process $\boldsymbol{Y}(t) : \Omega \to \mathbb{C}^J$, obtained when sampling the measurement field $\mathcal{Y}$ at locations $\boldsymbol{p}_1, \dots, \boldsymbol{p}_J \in \mathcal{P}$ (see section 2.2). The last equality eq. (9) comes from lemma 3.1 applied to the operator $\mathbf{\Psi} = \overline{\mathbf{\Phi}} \otimes \mathbf{\Phi}$ (see eq. (4)).

Let us now rewrite eq. (7) and see if we can arrive to a similar expression. We have

$$I_{\mathcal{S}_D}(\boldsymbol{r}) = \sum_{k,j=1}^{J} \mathbb{E}[\widetilde{Y}_k(t)\widetilde{Y}_j^*(t)]\phi_k(\boldsymbol{r})\phi_j^*(\boldsymbol{r}), \quad \forall \boldsymbol{r} \in \mathcal{M}.$$

This can be rewritten as

$$I_{\mathcal{S}_D} = \left(\overline{\mathbf{\Phi}} \otimes \mathbf{\Phi}\right) \text{vec}\left(\widetilde{\Sigma}\right),$$

where $\widetilde{\Sigma} \in \mathbb{C}^{J\times J}$ is the covariance matrix of the multivariate stochastic process $\widetilde{\boldsymbol{Y}}(t) = G_\Phi^{-1}\boldsymbol{Y}(t)$. We have then

$$\widetilde{\Sigma} = G_\Phi^{-1}\Sigma G_\Phi^{-H} = G_\Phi^{-1}\Sigma G_\Phi^{-1},$$

as G_Φ is Hermitian symmetric. Finally, a famous property of the Kronecker product tells us that

$$\text{vec}\left(\widetilde{\Sigma}\right) = \text{vec}\left(G_\Phi^{-1}\Sigma G_\Phi^{-1}\right) = \left(G_\Phi^{-T} \otimes G_\Phi^{-1}\right) \text{vec}\left(\Sigma\right),$$

and finally

$$I_{\mathcal{S}_D} = \left(\overline{\mathbf{\Phi}} \otimes \mathbf{\Phi}\right) \left(G_\Phi^{-T} \otimes G_\Phi^{-1}\right) \text{vec}(\Sigma),$$

which is in the exact same form as eq. (9). Hence, we have $I_D = I_{\mathcal{S}_D}$, and $\mathcal{S}_D$ is also the least-squares estimate of the intensity field I in E_Ψ (indeed, we showed the result for I_D before, see eq. (6)). ■

As a consequence of theorem 3.2, we can equivalently use algorithms 1 and 2 to produce the least-squares estimate of the intensity field. Even though algorithm 1 is mostly used in practice, we will prefer in this work the use of algorithm 2, as it permits to access to the covariance function of the reconstructed field $\mathcal{S}_D$, which we will use to classify the aliasing artifacts polluting the estimate.

4. Attributing Aliasing Artifacts to their Parent Source

In this section, we will see that the aliasing artifacts polluting the dirty intensity field I_D can actually be decomposed in Q different contributions from each of the sources composing the point source random field $\mathcal{S}$. Roughly speaking, aliasing artifacts can hence be seen as some energy "leakage" from the Q sources in the field. In this section, we will attribute artifacts to their parent source, defined as the source that mostly contributed to the dirty intensity field at a given location. We call this attribution task the parenting problem, which we will solve in different scenarios. Finally, we will reformulate the parenting problem as a graph clustering problem, and use spectral clustering to solve for the parenting problem in very general scenarios, with minimal assumptions on the underlying random field and acquisition system.

4.1. The Parenting Problem

Note that the interpolated random field $\mathcal{S}_D$ in algorithm 2 can alternatively be written as, $\forall (t, \boldsymbol{r}) \in \mathcal{I} \times \mathcal{M}$

$$S_D(t, \boldsymbol{r}) = \left(\widetilde{\mathbf{\Phi}}\boldsymbol{Y}(t)\right)(\boldsymbol{r}) = \sum_{k=1}^{J} Y_k(t)\widetilde{\phi}_k(\boldsymbol{r}) = \sum_{k=1}^{J} \langle S, \phi_k\rangle \widetilde{\phi}_k(\boldsymbol{r}), \tag{10}$$

where $\widetilde{\phi}_k = \sum_{j=1}^{J} \phi_j \left(G_\Phi^{-1}\right)_{jk}$, $\forall k = 1, \dots, J$. Since $\mathcal{S}$ is assumed to be a point source random field (see definition 2.2) we have, for $k = 1, \dots, J$,

$$\langle S, \phi_k\rangle = \sum_{q=1}^{Q} \xi_q(t)\langle \delta(\boldsymbol{r} - \boldsymbol{r}_q), \phi_k(\boldsymbol{r})\rangle = \sum_{q=1}^{Q} \xi_q(t)\phi_k^*(\boldsymbol{r}_q),$$

from the definition of the Dirac delta function. Hence, eq. (10) becomes

$$\begin{aligned} S_D(t, \boldsymbol{r}) &= \sum_{q=1}^{Q} \xi_q(t) \left(\sum_{k=1}^{J} \phi_k^*(\boldsymbol{r}_q)\widetilde{\phi}_k(\boldsymbol{r})\right) \\ &= \sum_{q=1}^{Q} \xi_q(t)\, \zeta_q(\boldsymbol{r}), \quad \forall \boldsymbol{r} \in \mathcal{M}, \end{aligned} \tag{11}$$

with $\zeta_q(\boldsymbol{r}) = \sum_{k=1}^{J} \phi_k^*(\boldsymbol{r}_q)\widetilde{\phi}_k(\boldsymbol{r})$, $\forall \boldsymbol{r} \in \mathcal{M}$, the **spread functions**. Then, the dirty intensity field can also be expressed in terms of the functions ζ_q as

$$I_D(\boldsymbol{r}) = \sum_{q=1}^{Q} \sigma_q^2\, \zeta_q^2(\boldsymbol{r}), \quad \forall \boldsymbol{r} \in \mathcal{M}. \tag{12}$$

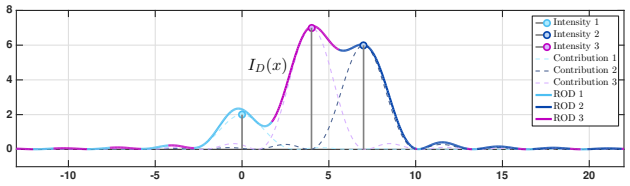


Figure 1 | Example of dirty intensity field (plain line). In this example, we chose $\mathcal{M} = \mathbb{R}$, $Q = 3$, $\sigma_q^2 = 2, 6, 7$, $\zeta_q(x) = \text{sinc}(x - n_q)$, and $n_q = 0, 7, 4$. The contributions (dashed lines) and regions of dominance (ROD) of the three sources (cyan, blue and magenta) are also displayed.

We observe that the value of the dirty intensity field I_D at a location $\boldsymbol{r} \in \mathcal{M}$ is the sum of Q contributions $\sigma_q^2 \zeta_q^2(\boldsymbol{r})$, convolution of each point source with the kernels $\zeta_q^2 : \mathcal{M} \to \mathbb{R}_+$ (see fig. 1). Even though every single source can theoritically contribute to the value of the dirty intensity field at a given location, in practice, for not too clustered sources, we often observe disjoint *regions of influence*. We say that a location $\boldsymbol{r} \in \mathcal{M}$ is in the **region of influence** of the q-th source, if the value of the dirty intensity field $I_D(\boldsymbol{r})$ at this location is almost completely explained by the single contribution of the q-th source. For practical purposes, it is more convenient to work with the concept of **region of dominance**, closely related to the concept of region of influence. We can define formally this concept by introducing the **parenting function**:

Definition 4.1 — ***Parenting Function & Region of Dominance.*** *Let $\mathcal{M}$ be some manifold, and $I_D : \mathcal{M} \to \mathbb{R}_+$ some intensity field defined over this manifold. We call the **parenting function** the function $\pi : \mathcal{M} \to \{1, \dots, Q\}$ defined as*

$$\pi(\boldsymbol{r}) = \underset{q=1,\dots,Q}{\text{argmax}} \{\sigma_q^2 \zeta_q^2(\boldsymbol{r})\}. \tag{13}$$

*Looking back at eq. (12), we observe that the parenting function associate each point $\boldsymbol{r} \in \mathcal{M}$ to its **parent source**, which contributed the most to the value of the dirty intensity field I_D observed at this location. The **region of dominance** $\mathcal{R}_q \subset \mathcal{M}$ for a given source $q \in \{1, \dots, Q\}$ is then defined as*

$$\mathcal{R}_q = \left\{ \boldsymbol{r} \in \mathcal{M} : \; q = \underset{k=1,\dots,Q}{\text{argmax}} \{\sigma_k^2 \zeta_k^2(\boldsymbol{r})\} \right\}.$$

By definition, we have

$$\mathcal{M} = \bigcup_{q=1,\dots,Q} \mathcal{R}_q, \text{ and } \mathcal{R}_q \cap \mathcal{R}_k = 0, \; \forall k \neq q.$$

Notice that the concept of region of dominance is a relaxation of the concept of region of influence, as we only require the parent source to have maximal contribution for every point of the region, which does not guarantee than its contribution alone would explain well enough the observed value of the intensity field in the whole region. However, this notion is easier to work with and well-suited for classification purposes, while still borrowing enough interpretability to the concept of region of influence to produce meaningful classifications. Our goal will hence be to reconstruct these regions of dominance, so as to classify the different aliasing artifacts polluting the dirty intensity field according to their respective parent source (see fig. 1). We will then identify within each region of dominance the most likely location of the parent source, exploiting known properties of the functions ζ_q^2.

Solving for this classification problem is equivalent to computing the parenting function eq. (13). For this reason, we will refer to it as the **parenting problem**. Of course, provided the knowledge of both σ_q^2 and ζ_q^2 for every source the problem would be trivial to solve. But these quantities are unavailable to us, as σ_q^2, $q = 1, \dots, Q$, are unknown and the spread functions ζ_q depend on the unknown locations $\boldsymbol{r}_q \in \mathcal{M}$ of the Q sources (see eq. (11)). Hence, we need first to estimate the parenting function.

4.2. Solving the Parenting Problem for Orthogonal Spread Functions

In order to gain insight on the problem, it is instructive to temporary restrict our attention to the case of orthogonal functions ζ_q. In this section, we will hence assume that

$$\langle \zeta_q, \zeta_k \rangle_{\mathcal{L}^2(\mathcal{M})} = \delta_{kq}, \quad k, q = 1, \dots, Q. \tag{14}$$

With this assumption, we observe that the spread functions ζ_q are the eigenfunctions of the **covariance operator** $T_{\kappa_d} : \mathcal{L}^2(\mathcal{M}) \to \mathcal{L}^2(\mathcal{M})$, defined as

$$\left(T_{\kappa_D} \, f(\boldsymbol{r})\right)(\boldsymbol{s}) = \int_{\mathcal{M}} \kappa_D(\boldsymbol{s}, \boldsymbol{r}) f(\boldsymbol{r}) \, d\boldsymbol{r}, \quad \forall \boldsymbol{s} \in \mathcal{M},$$

where $\kappa_D : \mathcal{M}^2 \to \mathbb{C}$ is the covariance function eq. (8) of the random field $\mathcal{S}_D$, which can be rewritten as

$$\kappa_D(\boldsymbol{s}, \boldsymbol{r}) = \sum_{q=1}^{Q} \sigma_q^2 \zeta_q(\boldsymbol{s}) \zeta_q^*(\boldsymbol{r}), \quad \forall (\boldsymbol{s}, \boldsymbol{r}) \in \mathcal{M}^2.$$

Indeed, we have $\forall q = 1, \dots, Q$

$$\begin{aligned} \left(T_{\kappa_D} \, \zeta_q(\boldsymbol{r})\right)(\boldsymbol{s}) &= \int_{\mathcal{M}} \kappa_D(\boldsymbol{s}, \boldsymbol{r}) \zeta_q(\boldsymbol{r}) \, d\boldsymbol{r} \\ &= \sum_{k=1}^{Q} \sigma_k^2 \zeta_k(\boldsymbol{s}) \left(\int_{\mathcal{M}} \zeta_k(\boldsymbol{r}) \zeta_q^*(\boldsymbol{r}) \, d\boldsymbol{r} \right)^* \\ &= \sigma_q^2 \zeta_q(\boldsymbol{s}), \quad \forall \boldsymbol{s} \in \mathcal{M}, \end{aligned}$$

where the last equality has been obtained by using the orthogonality assumption eq. (14). Moreover, notice that $\mathcal{R}(T_{\kappa_D}) = \text{span}\{\zeta_q, \; q = 1, \dots, Q\}$ and hence $\dim\left(\mathcal{R}(T_{\kappa_D})\right) = Q$. Thus, T_{κ_D} has exactly Q eigenpairs, given by (σ_q^2, ζ_q), $q = 1, \dots, Q$. This means that the spread functions ζ_q and intensities σ_q^2, necessary for the computation of the parenting function eq. (13), can be obtained by computing the eigenpairs of the operator T_{κ_D}.

This observation inspires the following procedure to estimate the parenting function in practice:

- Define a regular grid $\Delta_N = \{\boldsymbol{r}_i, \, i = 1, \dots, N\} \subset \mathcal{C}$ over a large enough bounded region $\mathcal{C}$ of $\mathcal{M}$, with N large.
- Evaluate the covariance function over the fine grid:

  $$K_{ij} = \kappa_D(\boldsymbol{r}_i, \boldsymbol{r}_j), \quad \forall i, j = 1, \dots, N.$$

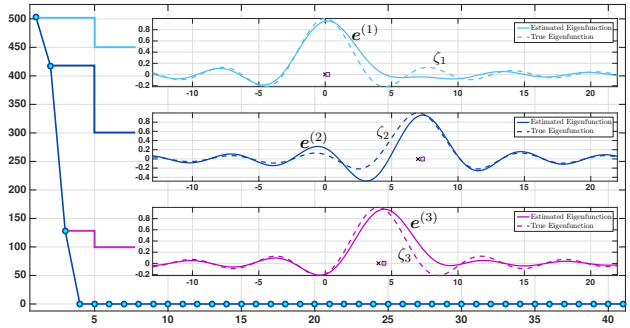
  This defines a matrix $K \in \mathbb{C}^{N \times N}$, discrete version of κ_D.
- Compute the N eigenpairs of the matrix K, $(\lambda_i, \boldsymbol{e}^{(i)}) \in \mathbb{R}_+ \times \mathbb{C}^N$, $i = 1, \dots, N$. Select the eigenpairs with eigenvalues significantly greater than zero. In the noiseless case and for large enough N, we will have exactly Q of these eigenpairs, discrete approximations of the eigenpairs (σ_q^2, ζ_q) evaluated on the grid Δ_N:

  $$e_i^{(q)} \simeq \zeta_q(\boldsymbol{r}_i), \quad \& \quad \lambda_q \simeq \sigma_q^2, \; q = 1, \dots, Q, \; i = 1, \dots, N.$$

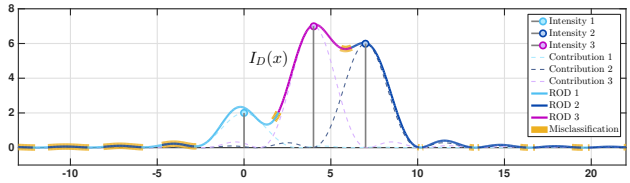
- Solve the parenting problem for the grid points in Δ_N by computing the discrete parenting vector $\boldsymbol{\pi} \in \{1, \dots, Q\}^N$, with coefficients given by:

  $$\pi_i = \underset{q=1,\dots,Q}{\text{argmax}} \left\{ \lambda_q \left(e_i^{(q)} \right)^2 \right\}.$$

  This vector associates each point of the grid to its parent source. We have $\pi_i \simeq \pi(\boldsymbol{r}_i)$, $i = 1, \dots, N$.



(a) Spectrum of the covariance matrix K and eigenvectors $\boldsymbol{e}^{(q)}$ (plain lines) compared to the true eigenfunctions ζ_q(dashed lines) for the three positive eigenvalues.



(b) Estimates of the regions of dominance for each source, using the three estimated eigenpairs. The misclassified portions are highlighted in orange ($\simeq 30\%$ of the total region).

Figure 2 | Estimating the regions of dominance for orthogonal spread functions (to be compared with fig. 1).

In fig. 2, we estimated the regions of confidence with the above procedure, for the scenario described in fig. 1 (in which the spread functions are indeed orthogonal). We simulated $N_s = 8000$ noiseless realisation of the dirty random field eq. (11), which we used to estimate the covariance function eq. (8) and its discrete version K, for a grid of 800 points, evenly spaced on the interval $[-13, 22]$. Notice that the eigenvectors $\boldsymbol{e}^{(q)}$ of K do not perfectly approximate the spread functions ζ_q. As a result, the accuracy of the obtained regions of confidence is not very satisfying, with only $\simeq 70\%$ of the grid points correctly classified. Finally, simulations show that this classification is very sensitive to the noise, as small perturbations in K can significantly change its spectral representation.

Hence, although mathematically elegant, the previously described procedure is unfortunately not suitable for practical purposes, being too restrictive and sensitive to the noise. It is however very insightful, as it suggests that the covariance function eq. (8) can be leveraged to solve for the parenting problem eq. (13), a fact that we will use in the next section.

4.3. The Parenting Problem & Spectral Clustering

In this section, we propose an alternative way of estimating the regions of dominance by reinterpreting the problem as a graph clustering problem. Again, consider a bounded subset $\mathcal{C}$ of $\mathcal{M}$, in which we would like to estimate the regions of dominance, and a fine grid of points $\Delta_N \subset \mathcal{C}$, uniformly distributed over the domain.

The idea is then to interpret these points as the *vertices* (or *nodes*) of some undirected graph $\mathcal{G} = (V, E)$, where $V = \Delta_N$ is the *vertex set* and $E \subset V \times V$ is the *edge set*. The network structure described by the edge set E is constructed so as to reflect potential similarities existing between the nodes, presumably inherited from common parent sources. Starting from a fully connected graph $E = V \times V$, a weight $w_{ij} \in \mathbb{R}_+$ is attributed to each edge $e_{ij} = (\boldsymbol{r}_i, \boldsymbol{r}_j) \in E$, according to the *similarity* of the associated nodes $\boldsymbol{r}_i, \boldsymbol{r}_j \in V$,

$$w_{ij} = \sigma(\boldsymbol{r}_i, \boldsymbol{r}_j) \geq 0, \quad i, j = 1, \dots, N,$$

where pair-wise similarities are computed with a suitable **similarity measure** $\sigma : \mathcal{M}^2 \to \mathbb{R}_+$. This similarity measure is, in some sense, assessing the *degree of kinship* between any two nodes in the field: if two nodes have the same parent source, then they will have strong similarity, and respectively for different parent sources. Edges with null weights are not related, and hence discarded from the edge set. The goal is then to cluster this kinship network in order to reconstruct the regions of dominance and identify the parent sources in each of them. Not surprisingly, the success of this operation will heavily depend on the chosen similarity measure [11]. Given the crucial role played by the covariance function in section 4.2, it seemed natural to us to define the similarity of any two points in $\mathcal{M}$ as the modulus of their covariance

$$\sigma(\boldsymbol{s}, \boldsymbol{r}) = |\kappa_D(\boldsymbol{s}, \boldsymbol{r})|, \quad \forall (\boldsymbol{s}, \boldsymbol{r}) \in \mathcal{M}^2.$$

Having specified the similarity measure, the undirected weighted graph $\mathcal{G}$ is entirely determined by its **affinity matrix** $A \in \mathbb{R}^{N \times N}$, defined as

$$A_{ij} = \begin{cases} w_{ij}, & \text{if } (\boldsymbol{r}_i, \boldsymbol{r}_j) \in E, \\ 0, & \text{otherwise,} \end{cases} \qquad i, j = 1, \dots, N.$$

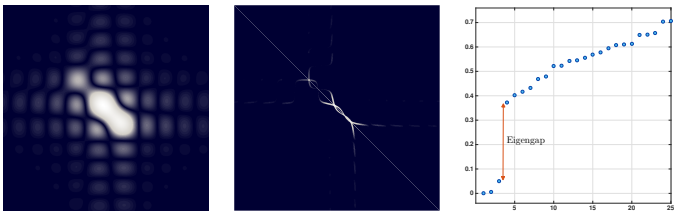
The graph clustering step is then performed by means of spectral clustering (see [11] for more details on the algorithm). Roughly speaking, this algorithm aims at partitionning the graph in K connected components of comparable size, such that the sum of the weights of the inter-components edges is minimised. In mathematical terms, the K connected components resulting from this operation, called **clusters**, are given by

$$\left\{\tilde{V}_1, \dots, \tilde{V}_K\right\} = \underset{\substack{V_1, \dots, V_K \subset V, \\ \cup_{i=1}^K V_i = V}}{\operatorname{argmin}} \left\{ \frac{1}{2} \sum_{i=1}^{K} \frac{W(V_i, \overline{V}_i)}{|V_i|} \right\}, \qquad (15)$$

where $|V_i|$ is the cardinality of V_i, $\overline{V}_i = V \backslash V_i$ and $W(V_i, \overline{V}_i)$ is the sum of the weights of the edges linking V_i and $\overline{V}_i$. The terms $1/|V_i|$ are introduced in the objective function in order to obtained well-balanced clusters, and hence avoid degenerate clusters with a single or very few vertices. This makes the optimisation problem eq. (15) NP hard (see [11]), and hence spectral clustering is only solving for a relaxed version of this problem. The key idea of the algorithm is to find a good embedding for the nodes in $\mathbb{R}^K$, which would greatly facilitate the clustering step by pre-clustering the data, hence allowing the use of simple clustering algorithms such as k-means to obtain the final clusters. This natural embedding for the graph is obtained from the rows of the eigenvectors associated with the K smallest eigenvalues of the graph Laplacian $L \in \mathbb{R}^{N \times N}$, linked to the affinity matrix by

$$L = D - A,$$

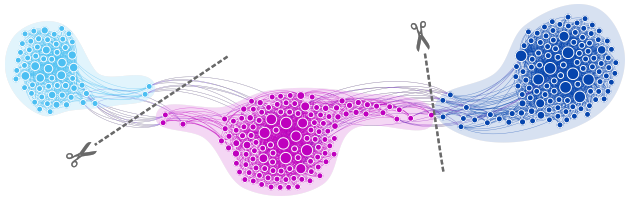
where $D \in \mathbb{R}^{N \times N}$ is a diagonal matrix with the weighted degree of each node on the diagonal $D_{ii} = \sum_{k=1}^{N} w_{ik}$, $i = 1, \dots, N$. Various heuristics can be used to estimate the number of cluster K, but the most popular one is certainly the *eigengap heuristic* [11], where K is chosen such that the K smallest eignevalues $\lambda_!, \dots, \lambda_K$ of L are very small and comparable, but λ_{K+1} is relatively large (see fig. 3). A theoretical justification for this heuristic is based on perturbation theory [11], as it can be shown that in the ideal case of K disconnected clusters, the eigenvalue 0 has multiplicity K. Finally, the intensity of the eigengap $\lambda_{K+1} - \lambda_K$ also affects the performances of spectral clustering, and large eigengaps will lead to better separated clusters [11]. Sometimes, when the graph is too densely connected and the similarity measure not local enough, the eigengap can be very subtle to detect, and the graph difficult to cluster. In that case, sparsifying the graph by discarding "superfluous"



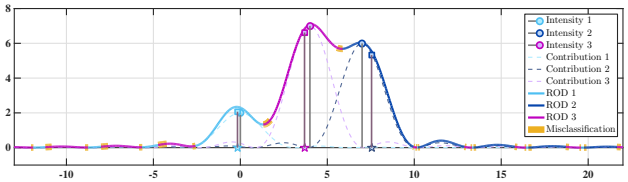
(a) Affinity matrix A of the graph.

(b) Affinity matrix after k-nearest neighbours with $k = 6$.

(c) The smallest 25 eigenvalues of L. We have $K = Q = 3$ clusters.



(d) The kinship graph after sparsification and the clusters obtained with spectral clustering (cyan, magenta and blue). The sum of the weights of the edges (in purple here) than need to be removed to separate the graph in three disconnected components is minimized.



(e) Estimates of the regions of dominance for each source, using spectral clustering. The misclassified portions are highlighted in orange ($\simeq$ 10% of the total region). The centers of the clusters (coloured pentagrams) as well as the estimated intensities of the parent sources (coloured squares) are also displayed.

Figure 3 | Estimating the regions of dominance with the kinship graph formulation and spectral clustering (to be compared with figs. 1 and 2).

edges can greatly help spectral clustering. Different sparsifying techniques have been proposed in the literature, and we will use in this work the k-nearest neighbours strategy [11], that links each node of the graph to its k most similar neighbours. The value of k to choose is application dependent, and is usually chosen according to the eigengap heuristic [11]. In fig. 3, we constructed the kinship graph for the scenario described in fig. 1, applied k-nearest neighbours with $k = 6$, and finally performed spectral clustering to recover the regions of dominance. The eigengap heuristic for this example yields $K = Q = 3$ clusters, or exactly as much as the number of point sources in the field (as expected). We observe that the accuracy of the classification obtained with this method is much more satisfying than in fig. 2. Indeed, more than 90% of the points on the grid Δ_N are correctly classified in their actual region of dominance (which is much better than in fig. 2).

Disposing of an algorithm to recover the respective regions of dominance, we are now interested in finding their associated parent source. This task can be accomplished by analysing the connection network within each cluster. Indeed, it makes intuitive sense that the parent source should be very *connected* with its children nodes, or mathematically speaking have the highest *kinship degree* within its corresponding region of dominance. We then propose to recover the parent sources $\boldsymbol{p}_k \in V$ of each cluster ROD_k by solving the following K optimisation problems

$$\boldsymbol{p}_k = \underset{\boldsymbol{r}_j \in ROD_k}{\operatorname{argmax}} \sum_{\boldsymbol{r}_i \in ROD_k} w_{ij}, \quad k = 1, \dots, K. \tag{16}$$

The vertex set V being in bijection with the physical grid $\Delta_N \subset \mathcal{M}$, we have then equivalently recovered the positions of each of the sources in the field. This knowledge completely characterises the spread functions ζ_k, which are thus available to us for computation. We can then estimate the corresponding intensities $\sigma_1^2, \dots, \sigma_K^2 \in \mathbb{R}_+$ of the respective parent sources by solving for the following linear system

$$\begin{pmatrix} \zeta_1^2(\boldsymbol{p}_1) & \cdots & \zeta_K^2(\boldsymbol{p}_1) \\ \vdots & \ddots & \vdots \\ \zeta_1^2(\boldsymbol{p}_K) & \cdots & \zeta_K^2(\boldsymbol{p}_K) \end{pmatrix} \begin{pmatrix} \sigma_1^2 \\ \vdots \\ \sigma_K^2 \end{pmatrix} = \begin{pmatrix} I_D(\boldsymbol{p}_1) \\ \vdots \\ I_D(\boldsymbol{p}_K) \end{pmatrix}. \tag{17}$$

In fig. 3 we used eqs. (16) and (17) to recover both the sources' locations and their respective intensity, for the specific example described in fig. 1. We observe that the resulting stream of Diracs follows quiet closely the actual stream of Diracs composing the underlying intensity field.

5. Application to Radio Astronomy

In this section, we apply the methodology described in section 4.3 to the problem of identifying sources in dirty images from radio astronomy. For the simulations, we use the layout of the LOw Frequency ARray (LOFAR) [1], the current largest radio interferometer.

In radio astronomy, the sky is often modelled as a point source random field $\mathcal{S}$ of the form of 2.2, with indexing set $\mathcal{I} \times \mathbb{S}^2$, where $\mathcal{I} \subset \mathbb{R}_+$ and $\mathbb{S}^2 \subset \mathbb{R}^3$ is the unit sphere, also called in this context the *celestial sphere* [4]. Assuming the sources to lie on this notional celestial sphere is a mathematical conveniency, that simply reflects the fact that the sources are in the far field [4]. This random field is then spatially sampled by an array of antennas, that record the incoming field at various locations on the earth. Modern radio telescope such as LOFAR are composed of thousands of antennas, spread all over Europe [1]. It is then crucial to reduce the data rate sent to the central processor as far up the chain as possible. For this reason, hierarchical designs have been proposed, where the antennas are grouped geographically in **stations** and the data from the individual antennas combined together at the station level by **beamforming** [12]. This effectively creates an array of virtual antennas (the stations) with potentially higher sensitivity towards certain regions of the sky [12, 5]. For an array of M stations and L antennas, the sensitivity pattern of the ith station is described by its **beamshape**

$$b_i(\boldsymbol{r}) = \sum_{k=1}^{L} \omega_k^{(i)} e^{j2\pi \langle \boldsymbol{r}, \boldsymbol{p}_k^{(i)} \rangle}, \quad \forall \boldsymbol{r} \in \mathbb{S}^2,\ i = 1, \dots, M.$$

where $\omega_k^{(i)} \in \mathbb{C}$, $k = 1, \dots, L$ are the known **beamforming weights** (see [13]) for station i and $\boldsymbol{p}_k^{(i)} \in \mathbb{R}^3$, $k = 1, \dots, L$ are the positions of the L antennas composing station i. This beamshape, different for each station, acts as a spatial filter [13] on the point source random field $\mathcal{S}$. Hence, the beamformed output $Y_i(t) : \Omega \to \mathbb{C}$ of station i is a continuous-time stochastic process, given by

$$Y_i(t) = \int_{\mathbb{S}^2} S(t, \boldsymbol{r})\, b_i^*(\boldsymbol{r})\, d\boldsymbol{r} \ + \ n_i(t), \quad \forall t \in \mathcal{I}, \quad i = 1, \dots, M, \tag{18}$$

with $n_i(t) : \Omega \to \mathbb{C}$ some thermal noise, assumed uncorrelated across stations and usually modelled as a complex, circularly symmetric Gaussian variable [2, 4], with variance $\sigma_n^2 > 0$

$$n_i(t) \sim \mathbb{C}\mathcal{N}(0, \sigma_n^2), \quad \forall t \in \mathcal{I},\, i = 1, \dots, M.$$

Notice that, if we define the station layout vector $\boldsymbol{p}^{(i)} = \left(\boldsymbol{p}_1^{(i)}, \dots, \boldsymbol{p}_L^{(i)}\right) \in \mathbb{R}^{3L}$, then eq. (18) can be seen as a noisy spatial sample $Y\left(t, \boldsymbol{p}^{(i)}\right)$ of the measurement field

$$\mathcal{Y} = \{Y(t, \boldsymbol{p}) : \Omega \to \mathbb{C}, \quad (t, \boldsymbol{p}) \in \mathcal{I} \times \mathbb{R}^{3L}\},$$

with

$$Y(t, \boldsymbol{p}) = \int_{\mathbb{S}^2} S(t, \boldsymbol{r}) b^*(\boldsymbol{r}, \boldsymbol{p})\, d\boldsymbol{r}, \quad \forall (t, \boldsymbol{p}) \in \mathcal{I} \times \mathbb{R}^{3L},$$

and, $\forall \boldsymbol{r} \in \mathbb{S}^2$, and $\boldsymbol{p} = (\boldsymbol{p}_1, \dots, \boldsymbol{p}_L) \in \mathbb{R}^{3L}$,

$$b(\boldsymbol{r}, \boldsymbol{p}) = \sum_{k=1}^{L} \omega\left(\boldsymbol{p}_k^{(i)}\right) e^{j2\pi\langle \boldsymbol{r}, \boldsymbol{p}_k^{(i)}\rangle},$$

where $\omega(\boldsymbol{q}) = \sum_{i,k=1}^{L,M} \omega_k^{(i)} \delta\left(\boldsymbol{q} - \boldsymbol{p}_k^{(i)}\right)$, $\forall \boldsymbol{q} \in \mathbb{R}^3$.

As in section 2.2, the M-dimensional continuous-time stochastic process $\boldsymbol{Y}(t) : \Omega \to \mathbb{C}^M$ composed by the M beamformed outputs $\boldsymbol{Y}(t) = (Y_1(t), \dots, Y_M(t))$ can be expressed in terms of the underlying point source random field $\mathcal{S}$ as

$$\boldsymbol{Y}(t) = \boldsymbol{\Phi}^* S(t, \boldsymbol{r}) \; + \; \boldsymbol{n}(t), \quad \forall t \in \mathcal{I},$$

where $\boldsymbol{\Phi}^* : \mathcal{L}^2(\mathbb{S}^2) \to \mathbb{C}^M$ is the analysis operator associated with the family of function $\mathcal{B} = \{b_i(\boldsymbol{r}),\, i = 1, \dots, M\} \subset \mathcal{L}^2(\mathbb{S}^2)$ and $\boldsymbol{n}(t) = (n_1(t), \dots, n_M(t)) \in \mathbb{C}^M$. In all that follows, we assume the family $\mathcal{B}$ to be linearly independent (verified in practice for the LOFAR layout).

Having defined the sampling operator $\boldsymbol{\Phi}^*$ associated with our specific acquisition system we can then use algorithm 2 to compute the dirty intensity field $I_D \in \mathcal{L}^2(\mathbb{S}^2)$ as well as the covariance function $\kappa_{\mathcal{S}_D} \in \mathcal{L}^2(\mathbb{S}^2 \times \mathbb{S}^2)$ of the dirty random field $\mathcal{S}_D$, least-squares estimate of the random field $\mathcal{S}$ in $E_\Phi = \text{span}\,\{b_i,\, i = 1, \dots, M\}$. The computation of the Gram matrix $G_\Phi = \boldsymbol{\Phi}^*\boldsymbol{\Phi} \in \mathbb{C}^{M \times M}$ in algorithm 2 requires to evaluate the pair-wise inner products between the beamshapes. This can be computed analytically, as

$$\begin{aligned}
(G_\Phi)_{ik} &= \langle b_i, b_k \rangle_{\mathcal{L}^2(\mathbb{S}^2)} \\
&= \int_{\mathbb{S}^2} b_i(\boldsymbol{r}) b_k^*(\boldsymbol{r})\, d\boldsymbol{r}, \\
&= \sum_{h,g=1}^{L} \omega_h^{(i)} \omega_g^{(i)*} \int_{\mathbb{S}^2} e^{j2\pi\langle \boldsymbol{r}, \boldsymbol{p}_h^{(i)} - \boldsymbol{p}_g^{(k)}\rangle} d\boldsymbol{r}, \\
&= 4\pi \sum_{h,g=1}^{L} \omega_h^{(i)} \omega_g^{(i)*} \text{sinc}\left(2\pi \|\boldsymbol{p}_h^{(i)} - \boldsymbol{p}_g^{(k)}\|_2\right),
\end{aligned}$$

where the last equality was obtain by applying the known identity [14, p. 154]

$$\int_{\mathbb{S}^2} e^{j2\pi\langle \boldsymbol{r}, \boldsymbol{p}\rangle} d\boldsymbol{r} = 4\pi \text{sinc}(2\pi\|\boldsymbol{p}\|_2), \quad \forall \boldsymbol{p} \in \mathbb{R}^3.$$

The Gram matrix being definite-positive, its inversion is performed by means of the Cholesky factorization followed by backward and forward substitution (see [15, Chapter 5] for more details on those concepts).

We finally apply the methodology described in section 4.3 to recover the regions of dominance together with the corresponding parent sources' locations and intensities. To reduce the computational requirements, this task was only performed on regions of the intensity field significantly different from zero (see [5] for a description of the statistical test used).

In figs. 4 to 6, we investigated simple scenarios with respectively two, three and five sources in the field of view, with various intensities. For a given sky model, we simulated $N_s = 800$ time samples from the random field $S(t, \boldsymbol{r})$, from which we built the M matched-beamformed [12] outputs $Y_i(t) : \Omega \to \mathbb{C}$, of the 24 core stations from the LOFAR telescope, each composed of 48 antennas. We have indeed, for a point source random field with Q sources

$$Y_i(t) = \sum_{q=1}^{Q} S(t, \boldsymbol{r}_q) b_i^*(\boldsymbol{r}_q) \; + \; n_i(t), \quad t \in \mathcal{I},\, i = 1, \dots, M.$$

In accordance with experimental conditions, the thermal noise variance was chosen much higher than the sources intensities, with a resulting peak signal to noise ratio around -25 dB for the three experiments. Finally, the covariance matrix $\widetilde{\Sigma} \in \mathbb{C}^{M \times M}$ of the stochastic process $\widetilde{\boldsymbol{Y}}(t) : \Omega \to \mathbb{C}^M$ in algorithm 2 was estimated from the 800 time samples with the classical maximum likelihood estimate of a covariance matrix.

The results from the three experiments are very satisfying, despite the very high noise level. The regions of dominance were recovered with an accuracy oscillating between 86.6% and 98.1%. The corresponding sources locations and intensities were also recovered almost perfectly, with a recovered sky almost indistinguishable from the actual sky with the human eye.

6. Conclusion

We proposed a novel method to locate sources within dirty intensity fields by attributing aliasing artifacts to their parent source. Fundamentally different from the deconvolution algorithms (such as CLEAN) proposed in the literature, our procedure recovers the *random field* from the data rather than directly targeting the intensity field. We then leverage the covariance function of this reconstructed random field as a measure of similarity between different points in the field. This allowed us to construct a kinship network between the different points in the field of view, on which we performed spectral clustering to learn the regions of dominance of the respective sources, and identify the actual sources in the field together with their respective intensity. Simulations results were very satisfying and the method was shown to be applicable to realistic scenarios from the field of radio astronomy.

References

[1] M. Van Haarlem, M. Wise, A. Gunst, G. Heald, J. McKean, J. Hessels, A. De Bruyn, R. Nijboer, J. Swinbank, R. Fallows, *et al.*, "LOFAR: The low-frequency array," Astronomy & astrophysics **556**, A2 (2013).

[2] A.-J. van der Veen and S. J. Wijnholds, "Signal processing tools for radio astronomy," in *Handbook of Signal Processing Systems*, pp. 421–463 (Springer, 2013).

[3] M. Simeoni, "Statistical Inference in Positron Emission Tomography," Tech. rep., EPFL (2014).

[4] O. Ocal, "Data Reduction Algorithms for Radio Astronomy Antenna Stations," Master's thesis, EPFL (Spring 2014).

[5] M. Simeoni, "Towards More Accurate and Efficient Beamformed Radio Interferometry Imaging," Master's thesis, EPFL (Spring 2015).

[6] M. Vetterli, J. Kovačević, and V. K. Goyal, *Foundations of signal processing* (Cambridge University Press, 2014).

[7] M. Vetterli, P. Marziliano, and T. Blu, "Sampling signals with finite rate of innovation," Signal Processing, IEEE Transactions on **50**(6), 1417–1428 (2002).

[8] U. Schwarz, "Mathematical-statistical description of the iterative beam removing technique (method CLEAN)," Astronomy and Astrophysics **65**, 345 (1978).

[9] D. Steer, P. Dewdney, and M. Ito, "Enhancements to the deconvolution algorithm'CLEAN'," Astronomy and Astrophysics **137**, 159–165 (1984).

[10] H.-H. Kuo, *White noise distribution theory*, vol. 5 (CRC press, 1996).

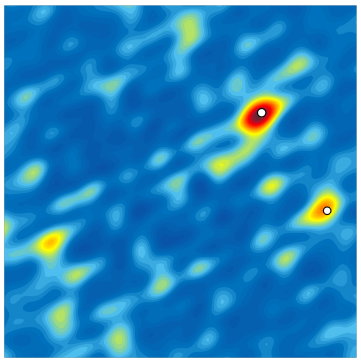
[11] U. Von Luxburg, "A tutorial on spectral clustering," Statistics and computing **17**(4), 395–416 (2007).

[12] O. Ocal, P. Hurley, G. Cherubini, and S. Kazemi, "Collaborative Randomized Beamforming for Phased Array Radio Interferometers," in *International Conference on Acoustics, Speech and Signal Processing (ICASSP), IEEE* (2015).

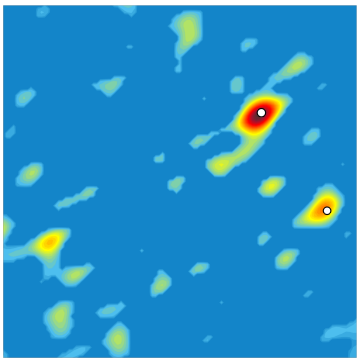
[13] P. Hurley and M. Simeoni, "Flexibeam: analytic spatial filtering by beamforming," in *International Conference on Acoustics, Speech and Signal Processing (ICASSP), IEEE* (March 2016 (to appear)).

[14] E. M. Stein and G. L. Weiss, *Introduction to Fourier analysis on Euclidean spaces*, vol. 1 (Princeton university press, 1971).

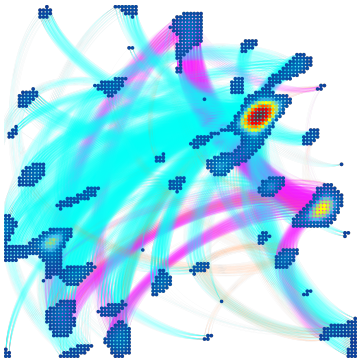
[15] L. N. Trefethen and D. Bau III, *Numerical linear algebra*, vol. 50 (Siam, 1997).



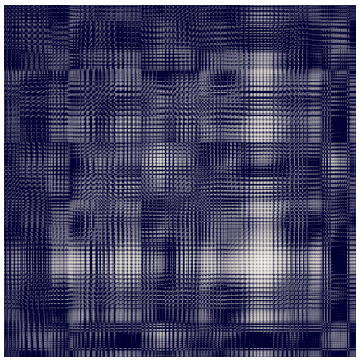
(a) Dirty intensity field. Actual sources are overlayed (white dots).



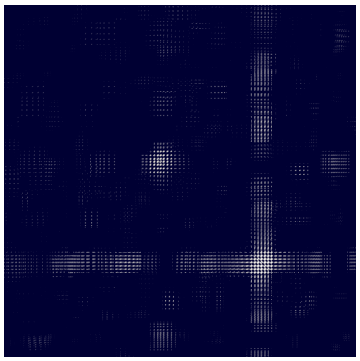
(b) Portions of the field significantly positive.



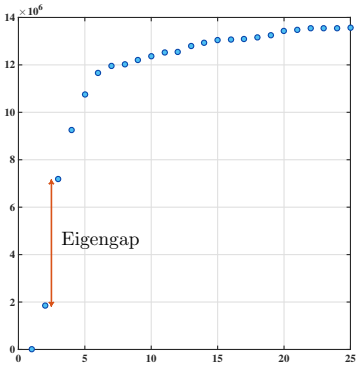
(c) Kinship graph of the significant portions.



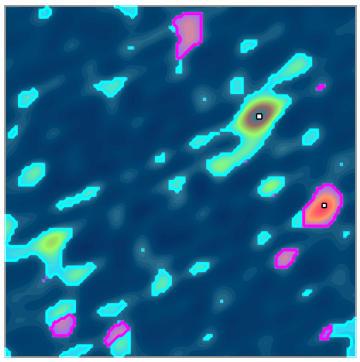
(d) Affinity matrix of the kinship graph.



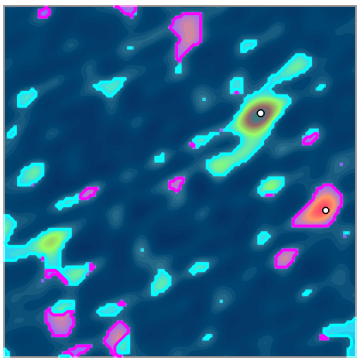
(e) Affinity matrix after k-nearest neighbours with $k = 38$.



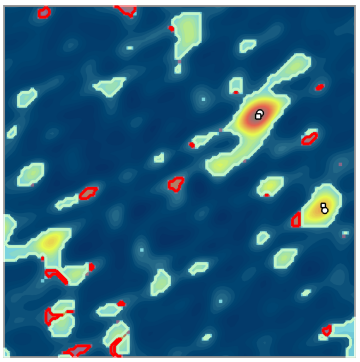
(f) The smallest 25 eigenvalues of L. We have $K = Q = 2$ clusters.



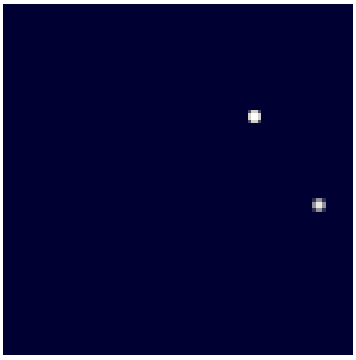
(g) Regions of dominance (cyan and magenta) recovered by spectral clustering. Centers of the clusters are overlayed (white squares).



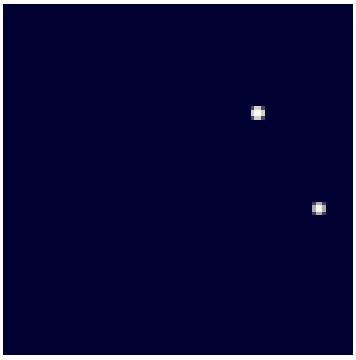
(h) Actual regions of dominance (cyan and magenta) and parent sources (white dots).



(i) Correctly (green) and wrongly (red) classified portions of the field (accuracy rate of $\simeq$ 98.1%); plus parent sources (dots) and their estimates (squares).

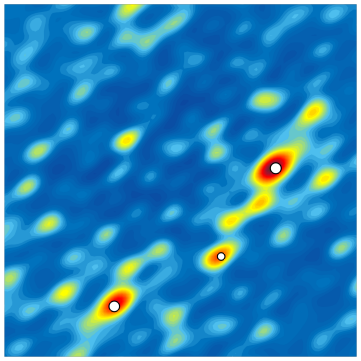


(j) Recovered sky.

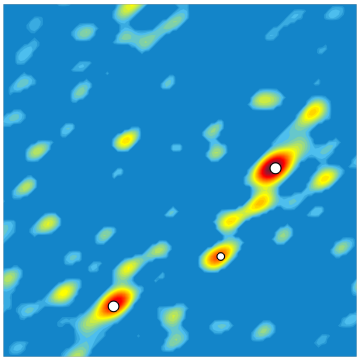


(k) Actual sky.

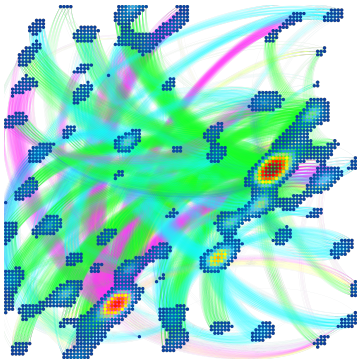
Figure 4 | Solving for the parenting problem with spectral clustering in radio-astronomy. Eight hundreds simulated samples from 24 antennas of the LOFAR telescope were used to estimate the intensity field and covariance function. The peak signal to noise ratio for this experiment is of -26.6 dB.



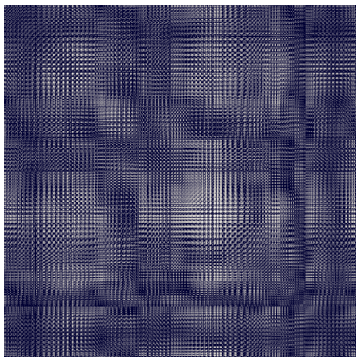
(a) Dirty intensity field. Actual sources are overlayed (white dots).



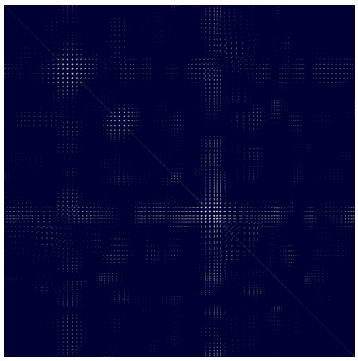
(b) Portions of the field significantly positive.



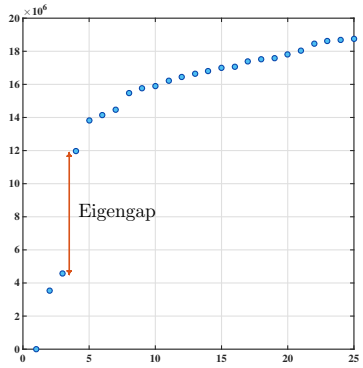
(c) Kinship graph of the significant portions.



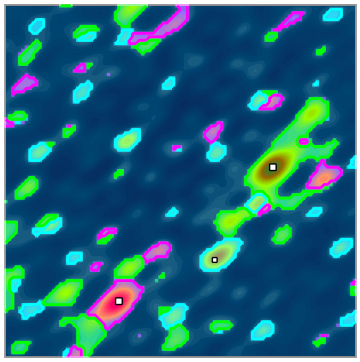
(d) Affinity matrix of the kinship graph.



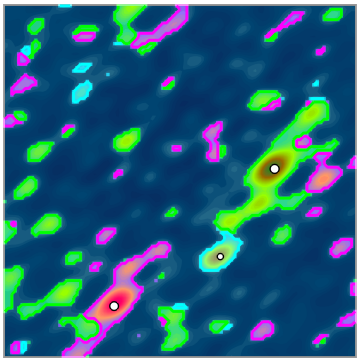
(e) Affinity matrix after k-nearest neighbours with $k = 28$.



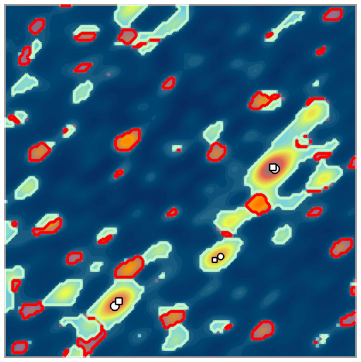
(f) The smallest 25 eigenvalues of L. We have $K = Q = 3$ clusters.



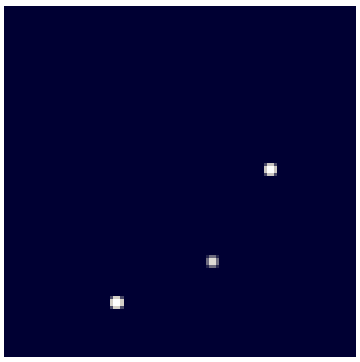
(g) Regions of dominance (cyan, magenta and green) recovered by spectral clustering. Centers of the clusters are overlayed (white squares).



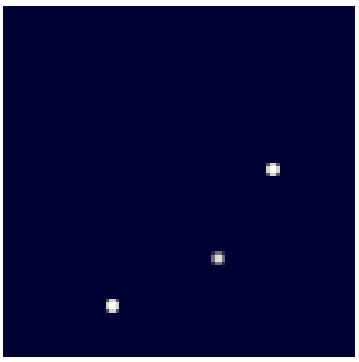
(h) Actual regions of dominance (cyan, magenta and green) and parent sources (white dots).



(i) Correctly (green) and wrongly (red) classified portions of the field (accuracy rate of $\simeq$ 86.8%); plus parent sources (dots) and their estimates (squares).

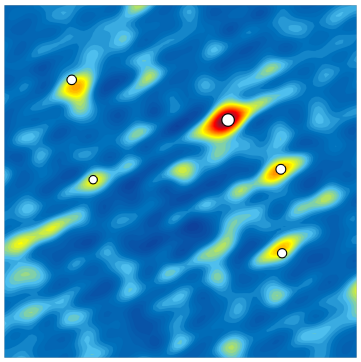


(j) Recovered sky.

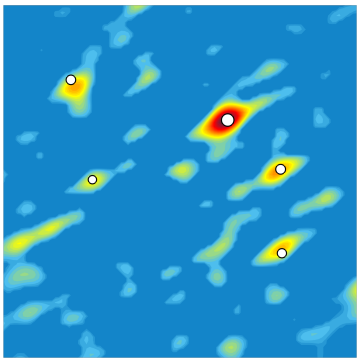


(k) Actual sky.

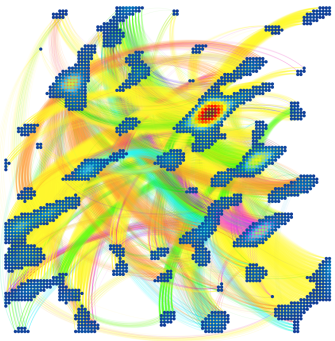
Figure 5 | Solving for the parenting problem with spectral clustering in radio-astronomy. Eight hundreds simulated samples from 24 antennas of the LOFAR telescope were used to estimate the intensity field and covariance function. The peak signal to noise ratio for this experiment is of -24.2 dB.



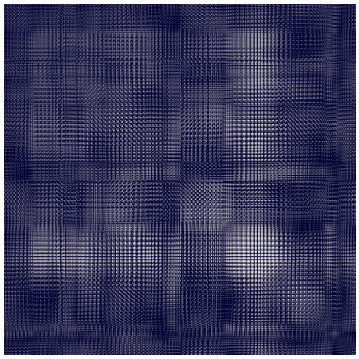
(a) Dirty intensity field. Actual sources are overlayed (white dots).



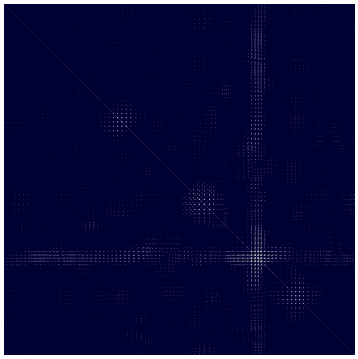
(b) Portions of the field significantly positive.



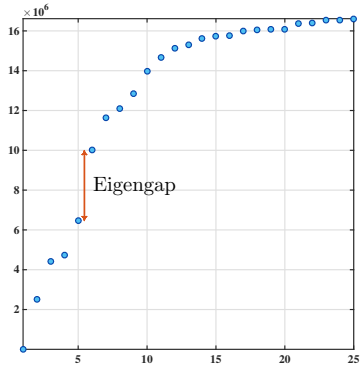
(c) Kinship graph of the significant portions.



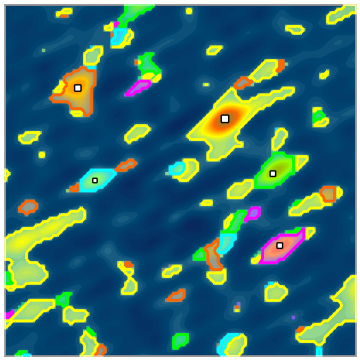
(d) Affinity matrix of the kinship graph.



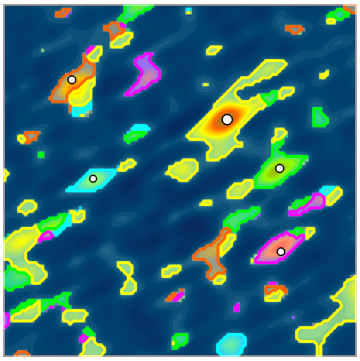
(e) Affinity matrix after k-nearest neighbours with $k = 20$.



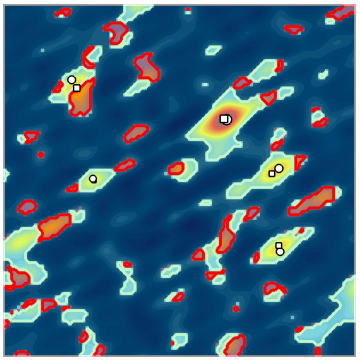
(f) The smallest 25 eigenvalues of L. We have $K = Q = 5$ clusters.



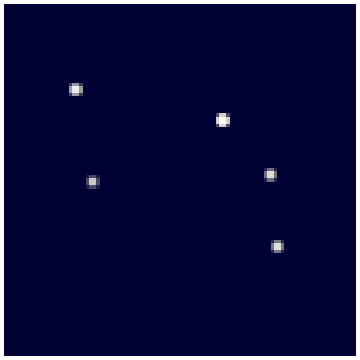
(g) Regions of dominance (cyan, magenta, green,orange and yellow) recovered by spectral clustering. Centers of the clusters are overlayed (white squares).



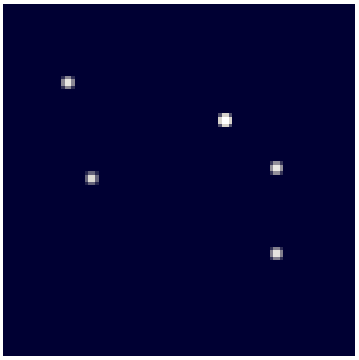
(h) Actual regions of dominance (cyan, magenta, green,orange and yellow) and parent sources (white dots).



(i) Correctly (green) and wrongly (red) classified portions of the field (accuracy rate of $\simeq$ 86.7%); plus parent sources (dots) and their estimates (squares).



(j) Recovered sky.



(k) Actual sky.

Figure 6 | Solving for the parenting problem with spectral clustering in radio-astronomy. Eight hundreds simulated samples from 24 antennas of the LOFAR telescope were used to estimate the intensity field and covariance function. The peak signal to noise ratio for this experiment is of -23 dB.