Sensor placement for the analysis of seismic surface waves: sources of error, design criterion and array design algorithms

Geophysical Journal International

Geophys. J. Int. (2014) **197,** 1566–1581
Advance Access publication 2014 April 25

doi: 10.1093/gji/ggt489

Sensor placement for the analysis of seismic surface waves: sources of error, design criterion and array design algorithms

Stefano Maranò,[1] Donat Fäh[1] and Yue M. Lu[2]

[1]*Swiss Seismological Service, ETH Zurich, CH-*8092 *Zürich, Switzerland. E-mail: stefano.marano@sed.ethz.ch*
[2]*School of Engineering and Applied Sciences, Harvard University, Cambridge, MA* 02138, *USA*

Accepted 2013 December 2. Received 2013 December 2; in original form 2013 May 21

SUMMARY

Seismic surface waves can be measured by deploying an array of seismometers on the surface of the earth. The goal of such measurement surveys is, usually, to estimate the velocity of propagation and the direction of arrival of the seismic waves. In this paper, we address the issue of sensor placement for the analysis of seismic surface waves from ambient vibration wavefields. First, we explain in detail how the array geometry affects the mean-squared estimation error of parameters of interest, such as the velocity and direction of propagation, both at low and high signal-to-noise ratios (SNRs). Secondly, we propose a cost function suitable for the design of the array geometry with particular focus on the estimation of the wavenumber of both Love and Rayleigh waves. Thirdly, we present and compare several computational approaches to minimize the proposed cost function. Numerical experiments verify the effectiveness of our cost function and resulting array geometry designs, leading to greatly improved estimation performance in comparison to arbitrary array geometries, both at low and high SNR levels.

Key words: Time series analysis; Fourier analysis; Surface waves and free oscillations.

1 INTRODUCTION

Sensor arrays are used in numerous applications, including radar, underwater source location, astronomical imaging and geophysical surveying. Since the geometry of the sensor array has a major impact on the performance of the array processing system, the design of optimal array geometries is an important task in many applications (Tokimatsu 1997; Van Trees 2002).

The motivation of this work arises from the analysis of seismic surface waves. In particular, our interest lies in the analysis of ambient vibrations from array recordings. Ambient vibrations span a broad range of frequencies and may have natural or anthropic origin (Bonnefoy-Claudet *et al.* 2006). Properties of the wavefield, such as the velocity of propagation and polarization, are used to infer a structural model for the site. This has application in microzonation and in geotechnical investigations (Tokimatsu 1997; Okada 2006).

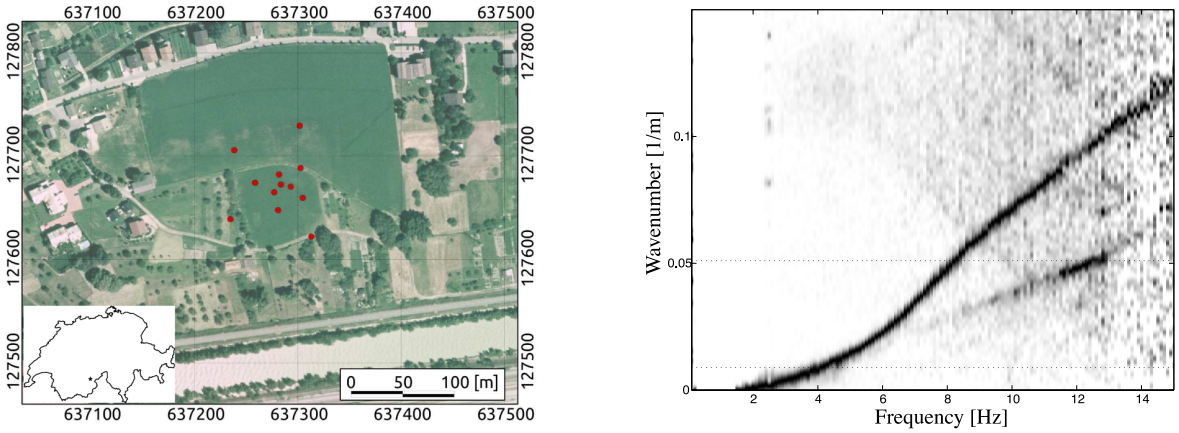
The wavefield of ambient vibrations is primarily composed of Love and Rayleigh waves. Array recordings of ambient vibrations are used to estimate the dispersion curve, i.e., the relationship between the velocity and frequency, of such waves.

Fig. 1(a) shows the location and the geometry of an array deployed by the Swiss Seismological Service near Brigerbad, in southwestern Switzerland. In this survey, the ground displacement produced by ambient vibrations is recorded for around two hours. A maximum likelihood (ML) method is used to estimate the wavenumbers of Love and Rayleigh waves. Additional details concerning the survey and the processing are given in Maranò *et al.* (2012). Fig. 1(b) depicts a large number of ML estimates of the wavenumber of Rayleigh waves at different frequencies. Darker regions indicate the presence of several wavenumber estimates having the same value. The dark curve extending across the whole figure, from bottom-left to top-right, identifies the dispersion curve of the fundamental mode. The first higher mode is also visible between 8 and 12 Hz, just below the fundamental mode.

An ML estimator suffers from two distinct types of errors, namely, *gross errors* (also known as global errors or outliers) and *fine errors* (also known as local errors; Vertatschitsch & Haykin 1991; Athley 2005). Both types of errors are influenced by array geometry. At low signal-to-noise ratio (SNR), the presence of local maxima in the likelihood function (LF) leads to large estimation errors. At high SNR, errors are smaller and accumulate around the global maximum of the LF.

In Fig. 1(b), it is possible to see that, towards higher frequencies and larger wavenumber, there is a significant amount of wavenumber estimates that do not belong to the dispersion curve. These are the gross errors or outliers. On the other hand, the thickness of the dispersion curve is related to fine errors, i.e., the variance of the estimator at high SNRs.

Using sensor arrays to study seismic wavefields has a long history and several different array geometries have been used. In Horike (1985), L-shaped and cross-shaped arrays with a regular sensor spacing have been employed. Irregularly spaced crosses were used



(a) (b)

Figure 1. Array deployment and Rayleigh wave dispersion curve from the Brigerbad survey. (a) Geometry of the sensor array. The inlet pinpoints the location of the array within Switzerland. The geographic coordinates are Swiss coordinates (CH1903). (b) Rayleigh wave dispersion curve. Fundamental mode and first higher mode are visible.

in Asten & Henstridge (1984), Milana *et al.* (1996), Ohori *et al.* (2002) and Rost & Thomas (2002). In other works, sensors were arranged as several triangles centred around a common point (Satoh *et al.* 2001a,b). In Gaffet *et al.* (1998) and Cornou *et al.* (2003), concentric circles were used.

The limitations of different array geometries have been investigated by different authors. In particular, the interest has been to identify a range of wavenumbers, or a related quantity, such as velocity or slowness, where the result of the array processing is more reliable. The largest and the smallest resolvable wavenumbers have been related either to the array aperture and the smallest interstation distance or to the height of the sidelobes of the array response function (Woods & Lintz 1973; Asten & Henstridge 1984; Tokimatsu 1997; Kind *et al.* 2005; Wathelet *et al.* 2008; Poggi & Fäh 2010).

The design of array geometries for the analysis of ambient vibrations has also been investigated by the community. Qualitative guidelines, based on empirical evidences, for array design are provided in Rost & Thomas (2002) and Kind *et al.* (2005). Haubrich (1968) proposes an approach for array design by considering properties of the coarray.

In this work, we present quantitative criteria and computational procedures for designing array geometries for measuring ambient vibrations. Our goal is to improve the performance of the ML estimator of wavefield parameters by optimizing the sensor positions. Given the nature of ambient vibrations and that we mostly rely on very noisy measurements, we are mainly interested in the low SNR regime, focusing primarily on reducing the occurrence of gross errors. In addition, we deal with small-scale arrays deployed at the earth surface, and thus optimize the geometry of a *planar* array, given a budget on the number of sensors and indication about the frequency support of the signals.

The contributions of this paper are threefold:

(1) We rigorously derive the relationship between array geometry and errors in parameter estimations. We show how the shape of the average LF is related to sensor positions through the Fourier transform of the sampling pattern.

(2) We propose a quantitative design criterion for improving estimation performance by means of sensor placement. By reformulating and relaxing the proposed optimization problem, we propose a practical array design algorithm based on a mixed integer program (MIP) with linear objective function and linear constraints. The proposed sensor placement algorithm generates arrays composed of simple regular geometries and are thus suitable for field deployment.

(3) We compare the proposed sensor placement algorithm with several other optimization techniques both in terms of array design and of estimation performance. We show through numerical experiments how the estimation performance can be increased by using the proposed array design criterion and algorithm.

The rest of this paper is organized as follows. In Section 2, wave equations, measurement model and an estimation method are presented. The distinction between gross and fine errors, together with a rigourous derivation of the relationship between sensor position and LF are given in Section 3. In Section 4, we outline the quantities relevant to the sensor placement problem and propose a design criterion. In Section 5, we consider array design methods and, in Section 6, we compare the results of different techniques. The findings of the paper are summarized in Section 7.

2 SYSTEM MODEL

Seismic surface waves propagate along the surface of the earth (Aki & Richards 1980) and can be measured using an array of seismometers. In seismic surveying, a typical goal is to estimate the velocity of propagation and the direction of arrival of such waves.

In this section, we briefly describe the model for seismic surface waves, the noise model for sensor measurements and the ML approach to parameter estimation.

Notation: The vector $\boldsymbol{\theta}$ indicates a generic value of the wavefield parameters. The vector $\check{\boldsymbol{\theta}}$ indicates the true, and possibly unknown, wavefield parameters. The vector $\hat{\boldsymbol{\theta}}$ indicates an estimate of $\check{\boldsymbol{\theta}}$. The same convention is used for each element of the wavefield parameter vector. When necessary, a superscript will specify whether the parameter vector describes a Love wave or a Rayleigh wave, i.e., $\boldsymbol{\theta}^{(\mathrm{L})}$ and $\boldsymbol{\theta}^{(\mathrm{R})}$, respectively.

2.1 Seismic surface waves

The perturbation induced on a ground particle by a seismic wave is described by a vector quantity. For a monochromatic, plane Love wave with frequency ω, the vector particle motion $\boldsymbol{u} = (u_x, u_y, u_z)^T$

is expressed as a function of location $\boldsymbol{p}$ and time t by

$$\begin{aligned} u_x(\boldsymbol{p},t) &= -\alpha \sin\psi \cos(\omega t - \boldsymbol{\kappa}\cdot\boldsymbol{p} + \varphi) \\ u_y(\boldsymbol{p},t) &= \alpha \cos\psi \cos(\omega t - \boldsymbol{\kappa}\cdot\boldsymbol{p} + \varphi) \\ u_z(\boldsymbol{p},t) &= 0\,, \end{aligned} \tag{1}$$

where ψ gives the direction of propagation, $\boldsymbol{\kappa} = \kappa(\cos\psi, \sin\psi)^{\mathrm{T}} = (\kappa_x, \kappa_y)^{\mathrm{T}}$ is the wave vector, κ being the wavenumber. The quantities α and φ relate to the amplitude and phase of the wave. We use a right-handed Cartesian coordinate system having the z-axis pointing upwards. The velocity of propagation is ω/κ. The azimuth ψ is measured counterclockwise from the x-axis. We say that a Love wave is parametrized by a wavefield parameter vector $\boldsymbol{\theta}^{(\mathrm{L})} = (\alpha, \varphi, \kappa, \psi)^{\mathrm{T}}$.

For a monochromatic, plane Rayleigh wave with frequency ω, the vector particle motion u is

$$\begin{aligned} u_x(\boldsymbol{p},t) &= \alpha \sin\xi \cos\psi \cos(\omega t - \boldsymbol{\kappa}\cdot\boldsymbol{p} + \varphi) \\ u_y(\boldsymbol{p},t) &= \alpha \sin\xi \sin\psi \cos(\omega t - \boldsymbol{\kappa}\cdot\boldsymbol{p} + \varphi) \\ u_z(\boldsymbol{p},t) &= \alpha \cos\xi \cos(\omega t - \boldsymbol{\kappa}\cdot\boldsymbol{p} + \pi/2 + \varphi)\,. \end{aligned} \tag{2}$$

The angle $\xi \in [-\pi/2, \pi/2)$ is called *ellipticity angle* of the Rayleigh wave and determines the eccentricity and the sense of rotation of the particle motion. The quantity $|\tan\xi|$ is known as the *ellipticity* of the Rayleigh wave. See Maranò *et al.* (2012) for a detailed description of this parametrization. A Rayleigh wave is parametrized by a wavefield parameter vector $\boldsymbol{\theta}^{(\mathrm{R})} = (\alpha, \varphi, \kappa, \psi, \xi)^{\mathrm{T}}$.

2.2 Scalar plane wave

We also consider a simpler scalar wave model. This model will be used in parts of this paper to introduce certain ideas before extending them to Love and Rayleigh wave models.

Let $u(\boldsymbol{p}, t)$ denote the scalar value of the wavefield at position $\boldsymbol{p}$ and time t. For a monochromatic source at frequency ω, the wavefield is

$$u(\boldsymbol{p},t) = \alpha_0 \cos(\omega t - \boldsymbol{\kappa}\cdot\boldsymbol{p} + \varphi_0)\,. \tag{3}$$

Suitable parametrization of α_0 and φ_0 makes the scalar wave model equivalent to a given component of a vector wave model as given in (1) or (2).

2.3 Measurement model

To measure seismic waves, we deploy an array of N_{s} sensors on the surface of the earth positioned at locations $\{\boldsymbol{p}_n\}_{n=1,\ldots,N_{\mathrm{s}}}$. We restrict our interest to small aperture arrays and work with a flat earth model, thus consider planar arrays. The signal at each sensor component is sampled at K instants $\{t_k\}_{k=1,\ldots,K}$. In general, each sensor measures a vector quantity. Let L be the total number of channels recorded by the array. In the case of scalar sensors, then simply $L = N_{\mathrm{s}}$.

Each measurement $Y_k^{(\ell)}$ is corrupted by additive white Gaussian noise and is modelled as

$$Y_k^{(\ell)} = u_k^{(\ell)}(\boldsymbol{\theta}) + Z_k^{(\ell)}\,, \tag{4}$$

for each channel $\ell = 1, \ldots, L$, where $Z_k^{(\ell)} \sim \mathcal{N}(0, \sigma_\ell^2)$. The noise variance, σ_ℓ^2 is, in general, different on each channel. The quantities $u_k^{(\ell)}(\boldsymbol{\theta})$ are deterministic functions of wavefield parameters $\boldsymbol{\theta}$ as described in (1)–(3).

It follows that the joint probability density function (PDF) of the measurements is

$$p_Y(\boldsymbol{y}|\boldsymbol{\theta}) = \prod_{\ell=1}^{L}\prod_{k=1}^{K} \frac{1}{\sqrt{2\pi\sigma_\ell^2}} \exp\left(-\frac{(y_k^{(\ell)} - u_k^{(\ell)}(\boldsymbol{\theta}))^2}{2\sigma_\ell^2}\right)\,, \tag{5}$$

where we grouped the measurements as $\boldsymbol{Y} = \{Y_k^{(\ell)}\}_{\substack{\ell=1,\ldots,L \\ k=1,\ldots,K}}$.

2.4 Parameter estimation

Wavefield parameters can be found by using ML estimation. ML estimation is a useful method for estimating the parameters of a statistical model. It is a widely used estimation technique due to its broad applicability and to the optimal performances in many settings (Kay 1993). An implementation of this method for the estimation of wavefield parameters of surface waves has been proposed by Maranò *et al.* (2012).

The LF of the observations is obtained from the PDF of the measurements (5). We denote the LF by $p_Y(\tilde{\boldsymbol{y}}|\boldsymbol{\theta})$, where $\tilde{\boldsymbol{y}}$ are the observations and $\boldsymbol{\theta}$ is the vector of the wavefield parameters argument of the LF. We stress that the LF is a function of the model parameters $\boldsymbol{\theta}$ while the measurements $\tilde{\boldsymbol{y}}$ are fixed.

An ML estimate $\hat{\boldsymbol{\theta}}$ of the true wavefield parameters $\check{\boldsymbol{\theta}}$ is found by maximizing the LF, i.e.,

$$\hat{\boldsymbol{\theta}} = \operatorname*{argmax}_{\boldsymbol{\theta}} p_Y(\tilde{\boldsymbol{y}}|\boldsymbol{\theta})\,. \tag{6}$$

The LF can be thought of as a utility function which is legitimized by the statistical model of the observations. The point of maximum of the LF corresponds to the ML estimate of the parameters.

3 SOURCES OF ERROR

In this section, we describe in detail how the sensor positions affect the performance of the ML parameter estimation. We make a distinction between two types of errors: gross errors and fine errors. Array geometry affects both types of error.

Fig. 2 shows the typical behaviour of the mean-squared estimation error (MSEE) of an ML estimator. The figure is obtained by repeating the estimation of the wavenumber of an unknown wave with several different noise realization and for different SNRs. The

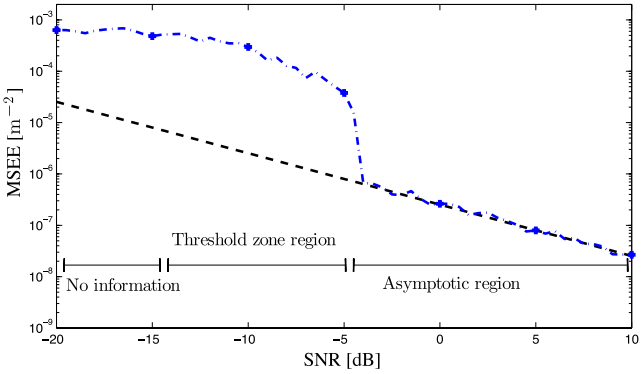


Figure 2. An example of the MSEE of a ML estimator. The MSEE is depicted with a blue dashed line. In the no information region the MSEE is very large and constrained by the implementation of the algorithm. In the threshold region, the occurrence of outliers keep the MSEE significantly larger than the Cramér–Rao bound (CRB). At last, in the asymptotic region, the MSEE is well described by the CRB, which is shown with the black dashed line.

figure should be understood as a qualitative depiction of how the MSEE decreases for increasing SNR. The SNR is defined as the ratio of signal power over noise power, i.e., SNR $= \alpha_0^2/2\sigma^2$. Three operation regions of the estimator are recognized at different SNR ranges. The approximate extent of the regions is shown in Fig. 2. At very low SNR, the noise dominates the signal of interest, and this is called the *no information* region. In this region, the estimates are completely random and carry no information about the value of the parameter estimated. At larger SNR, there is a region called the *threshold region*. In this region, the estimated value may be often close to the true value; however, the MSEE is still considerably large as gross estimation errors occur. Gross estimation errors are also known as global errors or outliers. Further increasing the SNR, we approach the *asymptotic region*. Fine estimation error occurs in this region and the MSEE of an ML estimator is well described by the Cramér–Rao bound (CRB; Kay 1993). Fine estimation errors are also known as local errors.

The abrupt increase in the MSEE below a certain SNR is known in literature as the *threshold effect* (or threshold phenomenon) and is due to a transition from fine errors to gross errors (Kay 1993; Van Trees 2001; Athley 2005).

3.1 Gross errors

Gross errors are due to the presence of local maxima (sidelobes) other than the true maximum in the LF. In this section, we establish the relationship between the LF and the array geometry.

Consider an array of N_s sensors at positions $\{\boldsymbol{p}_n\}_{n=1,\ldots,N_\mathrm{s}}$. The spatial sampling pattern is given by a sum of Dirac delta located at the sensor positions

$$h(x,y) = \sum_{n=1}^{N_\mathrm{s}} \delta(\boldsymbol{p}-\boldsymbol{p}_n)\,. \tag{7}$$

Its Fourier transform is given by

$$H(\kappa_x,\kappa_y) = \int_{\mathbb{R}^2} h(x,y)\mathrm{e}^{-i(\kappa_x,\kappa_y)\,\cdot\,p}\mathrm{d}\boldsymbol{p} \tag{8}$$

$$= \sum_{n=1}^{N_\mathrm{s}} \mathrm{e}^{-i(\kappa_x,\kappa_y)\,\cdot\,\boldsymbol{p}_n}\,, \tag{9}$$

where κ_x and κ_y are the wavenumber along the x and y coordinate axes, respectively.

We will show that $H(\kappa_x, \kappa_y)$ is important to explain the occurrence of gross errors and to establish a strategy to mitigate them.

It is possible to verify that the function $|H(\kappa_x, \kappa_y)|$ always exhibits a global maximum at $(\kappa_x, \kappa_y) = (0, 0)$. Other maxima also exist but, in general, have smaller amplitude. Moreover, the function is symmetric around the origin, i.e., $|H(\kappa_x, \kappa_y)| = |H(-\kappa_x, -\kappa_y)|$.

Similarly, the temporal sampling pattern g and its Fourier transform G are given by

$$g(t) = \sum_{k=1}^{K} \delta(t-t_k), \tag{10}$$

$$G(\omega) = \int_{\mathbb{R}} g(t)e^{-i\omega t}\mathrm{d}t = \sum_{k=1}^{K} e^{-i\omega t_k}\,, \tag{11}$$

where $\{t_k\}_{k=1,\ldots,K}$ are the sampling times.

After introducing these quantities, we return to our main interest, i.e. the analysis of the shape of the LF. To this aim, we compute the expectation of the LF from the model of (5).

We consider the observations $\tilde{\boldsymbol{y}}$ of a single scalar wave (*cf.* (3)) with true wavefield parameter vector $\check{\boldsymbol{\theta}} = (\check{\alpha}_0, \check{\varphi}_0, \check{\kappa}, \check{\psi})^\mathrm{T}$. We note that the ML estimates of $\check{\alpha}_0$ and $\check{\varphi}_0$ can be found explicitly as a function of the observations $\tilde{\boldsymbol{y}}$ and of the wave vector $\boldsymbol{\kappa}$ (see Appendix A1 for analytic expressions). Consequently, we can simplify the LF of the observations as a function of only $\boldsymbol{\kappa}$, as

$$p_Y(\tilde{\boldsymbol{y}}|\boldsymbol{\kappa}) = \max_{\alpha_0,\varphi_0} p_Y(\tilde{\boldsymbol{y}}|\boldsymbol{\theta})\,. \tag{12}$$

The maximization is achieved by inserting into $p_Y(\tilde{\boldsymbol{y}}|\boldsymbol{\theta})$ the ML estimates $\hat{\alpha}_0$ and $\hat{\varphi}_0$ (see Appendix A1 for more details).

Since we are interested only in the shape of the LF, we compute its logarithm, and drop multiplicative and additive constants. We use the expectation operator $\mathsf{E}\{\cdot\}$ to obtain an indication about the average shape of the LF. After some manipulations, explained in detail in Appendix A1, we obtain

$$\mathsf{E}\left\{\ln(p_Y(\boldsymbol{y}|\boldsymbol{\kappa}))\right\} \propto |G(\omega-\check{\omega})H(\boldsymbol{\kappa}-\check{\boldsymbol{\kappa}})|^2\,, \tag{13}$$

where the symbol $\propto$ denotes equality up to an affine transform, i.e., $f(x) \propto g(x)$ if $f(x) = C_1 g(x) + C_2$ and C_1, C_2 do not depend on x. From (13), we understand that the average shape of the log-likelihood function (LLF) is related to a translation of the Fourier transform of the sampling pattern.

The quantity $G(\cdot)$ is of marginal importance concerning the occurrence of gross errors. Indeed, it is usually possible to sample the signal with sufficiently small sampling time and with enough samples. We observe that for ω exactly known, $G(\omega-\check{\omega}) = G(0)$ is a real constant and thus does not change the shape of the LF. Hereafter, we will omit the factor $G(\cdot)$.

The quantity $|H(\boldsymbol{\kappa})|$, or quantities closely related to it, is known in literature with several names, such as array response (Woods & Lintz 1973; Asten & Henstridge 1984; Rost & Thomas 2002; Van Trees 2002; Wathelet *et al.* 2008), array factor (Bevelacqua & Balanis 2007) and array transfer function (Gaffet *et al.* 1998).

It is now clear how the Fourier transform of the sampling pattern $H(\boldsymbol{\kappa})$ affects the shape of the LF. At low SNR, outliers tend to accumulate around the local maxima of $|H(\boldsymbol{\kappa})|$. In order to reduce the occurrence of the gross errors, it is necessary to reduce the height of the local maxima (Vertatschitsch & Haykin 1991; Athley 2005).

As mentioned in the previous section, seismic waves are measured and modelled as vector quantities. Therefore, it is necessary to extend the findings concerning gross errors obtained for the scalar wave case to the vector wave case.

When considering vector measurements, the PDF of the observations needs to be augmented with the contribution of all the sensor components. With the assumption of independent observations, this is achieved by increasing L in (5) and by choosing the appropriate wave model $u_k^{(\ell)}$.

Concerning gross errors, the shape of the LF of observations of Love and Rayleigh waves is influenced differently by the different components of each sensor but remains a function of the Fourier transform of the sampling pattern $H(\boldsymbol{\kappa})$. The derivation of this relationships are detailed in Appendix A2 and in this section we only present the final results.

3.1.1 *Love wave*

For Love waves as in (1), the LLF is related to H as

$$\mathsf{E}\left\{\ln(p_Y(\boldsymbol{y}|\kappa,\psi))\right\} \propto f_\mathrm{L}(\psi,\check{\psi})\,|H(\boldsymbol{\kappa}-\check{\boldsymbol{\kappa}})|^2\,, \tag{14}$$

where

$$f_{\mathrm{L}}(\psi, \check{\psi}) = \cos^2(\psi - \check{\psi})\,. \tag{15}$$

From the previous expressions it is possible to understand that, similar to the scalar wave setting, the LLF for Love waves is directly related to the Fourier transform of the sampling pattern. The factor $f_{\mathrm{L}}(\psi, \check{\psi})$ influences the shape of the LLF as a function of the azimuth ψ and the true azimuth of the wave $\check{\psi}$.

3.1.2 *Rayleigh wave*

For Rayleigh waves as in (2), the LLF is related to H as

$$\mathsf{E}\left\{\ln(p_{\boldsymbol{Y}}(\boldsymbol{y}|\kappa, \psi, \xi))\right\} \propto f_{\mathrm{R}}(\psi, \check{\psi}, \xi, \check{\xi})\,|H(\boldsymbol{\kappa} - \check{\boldsymbol{\kappa}})|^2\ , \tag{16}$$

where

$$f_{\mathrm{R}}(\psi, \check{\psi}, \xi, \check{\xi}) = \left(\sin\xi \sin\check{\xi} \cos(\psi - \check{\psi}) + \cos\xi \cos\check{\xi}\right)^2\ . \tag{17}$$

We observe that for a fixed ellipticity angle ξ, the relationship is similar to the Love wave setting. For fixed azimuth ψ and wavenumber κ, the occurrence of local maxima is described by trigonometric functions of ξ. We observe that the local maxima due to (17) are independent of the sensor positions and, therefore, are outside the scope of this work.

3.2 Fine errors

At high SNR, the performance of the ML estimator is well described by the CRB. The CRB is a lower bound on the variance of all unbiased estimators. For example, Fig. 2 shows how the MSEE matches the CRB in the asymptotic region.

To compute the CRB, we first need to introduce the notion of Fisher information (FI). The FI conveys the amount of information about a statistical parameter carried by the PDF of the observations (Kay 1993).

For a statistical model with multiple parameters, the FI matrix (FIM) is given by

$$\boldsymbol{\mathcal{I}}(\boldsymbol{\theta}) = \mathsf{E}\left\{-\frac{\partial^2 \ln p_{\boldsymbol{Y}}(\boldsymbol{y}|\boldsymbol{\theta})}{\partial \boldsymbol{\theta}^2}\right\}\ , \tag{18}$$

where $p_{\boldsymbol{Y}}(\boldsymbol{y}|\boldsymbol{\theta})$ may be exactly (5). The matrix $\boldsymbol{\mathcal{I}}$ is a square symmetric matrix with as many columns as the elements in the vector $\boldsymbol{\theta}$. The diagonal terms of the matrix correspond to the FI of each parameter in the parameter vector $\boldsymbol{\theta}$. These elements should be interpreted with care, as they disregard the uncertainty due to the other model parameters being unknown. The off-diagonal terms are sometimes referred to as cross-information terms.

The information inequality states that the MSEE of an unbiased estimator is lower bounded as

$$\mathsf{E}\{(\hat{\boldsymbol{\theta}} - \mathsf{E}\{\hat{\boldsymbol{\theta}}\})(\hat{\boldsymbol{\theta}} - \mathsf{E}\{\hat{\boldsymbol{\theta}}\})^T\} \succeq (\boldsymbol{\mathcal{I}}(\boldsymbol{\theta}))^{-1}\,, \tag{19}$$

where $\mathbf{A} \succeq \mathbf{B}$ means that the matrix $\mathbf{A} - \mathbf{B}$ is positive semi-definite (Kay 1993). The left-hand side of (19) represents the covariance matrix of the vector $\hat{\boldsymbol{\theta}}$ and the right-hand side is the matrix inverse of the FIM. Following the information inequality, we are interested in the diagonal elements of $\boldsymbol{\mathcal{I}}^{-1}$ as they provide a lower bound on the MSEEs of the corresponding parameters.

The CRB on wavenumber for the scalar wave model is obtained using (3) and (5). The FIM is derived and then inverted analytically as in (18) and (19). From the corresponding entry of $\boldsymbol{\mathcal{I}}^{-1}$, the MSEE of wavenumber is lower bounded as

$$\mathsf{E}\left\{(\hat{\kappa} - \mathsf{E}\{\hat{\kappa}\})^2\right\} \geq \left(\frac{\alpha_0^2 K}{2\sigma^2}\left(Q_{aa}(\psi) - \frac{Q_{ab}^2(\psi)}{Q_{bb}(\psi)}\right)\right)^{-1}\ . \tag{20}$$

The CRB is directly proportional to noise power σ^2, inversely proportional to the amplitude of the wave α_0 and to the number of samples K. The CRB depends on the sensor positions through Q_{aa}, Q_{bb} and Q_{ab}. We also observe that the wavenumber CRB is independent of the temporal frequency ω, thus we expect fine errors to have comparable variance at any frequency.

The quantities Q_{aa}, Q_{bb} and Q_{ab} are called moment of inertia of the array. These quantities are independent of array translations, but are in general dependent on the azimuth. We introduce the coordinate system (a, b), which is related to the coordinate system (x, y) as

$$\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} \cos\psi & \sin\psi \\ -\sin\psi & \cos\psi \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}\ , \tag{21}$$

where the angle of rotation is the azimuth ψ. Therefore, a is the axis along the direction of propagation of the wave and b the axis perpendicular to it. The sensor positions in the rotated coordinate system are $\{(a_n, b_n)\}_{n=1,\ldots,N_{\mathrm{s}}}$. The moment of inertia (MOI) of the array in the coordinate system (a, b) are defined as

$$Q_{aa}(\psi) = \sum_{n=1}^{N_{\mathrm{s}}} (a_n - \bar{a})^2, \tag{22}$$

$$Q_{bb}(\psi) = \sum_{n=1}^{N_{\mathrm{s}}} (b_n - \bar{b})^2, \tag{23}$$

$$Q_{ab}(\psi) = \sum_{n=1}^{N_{\mathrm{s}}} (a_n - \bar{a})(b_n - \bar{b})\,, \tag{24}$$

where $\bar{a} = \frac{1}{N_{\mathrm{s}}} \sum_{n=1}^{N_{\mathrm{s}}} a_n$ and $\bar{b} = \frac{1}{N_{\mathrm{s}}} \sum_{n=1}^{N_{\mathrm{s}}} b_n$ define the *phase centre* of the array (Dogandžić & Nehorai 2001).

With reference to the CRB in (20), large Q_{aa} and Q_{bb} are desirable in order to reduce the CRB and thus the MSEE in the asymptotic region. Also a small Q_{ab}^2 is advantageous.

A large Q_{aa} can in general be obtained with a large aperture array. However, observe that a large aperture may invalidate the plane wave assumption, which is of critical importance in practical applications. Moreover, it is possible to choose an array geometry such that $Q_{ab} = 0$ and thus eliminating the term $-Q_{ab}^2/Q_{bb}$ from (20).

3.2.1 *Love wave and Rayleigh wave*

In the vector case, the derivation of the CRB in Section 3.2 needs to be extended. However, for translational sensors the dependence of the CRB on array geometry remain similar to the one presented in (20) and, therefore, we do not review this aspect. A detailed analysis of the CRB of parameters of Love wave and Rayleigh wave measured with translational and/or rotational sensors is found in Maranò & Fäh (2014).

4 PROBLEM STATEMENT AND DESIGN CRITERION

The aim of sensor placement is to improve the performance of parameter estimation, by an appropriate choice of the array geometry.

In particular, we consider the setting where two design requirements are given:

(1) A budget of N_s sensors to be placed on a plane surface.
(2) An indication about the spatial bandwidth of the signals of interest.

In several applications, only a limited number of sensors is available. This is often the case for the analysis of seismic wavefields where each individual instrument can be quite expensive and/or difficult to install at the measurement site.

Spatial bandwidth is defined by the spatial frequency content of the wavefield. Particularly important is the largest wavenumber present in the wavefield, denoted by $\kappa_{\max}$. Knowledge of spatial bandwidth is a reasonable assumption in many applications. In the seismic case, the wavenumber is the goal of the estimation process and thus an exact knowledge of the signal bandwidth is not available. However, prior knowledge on the geology of the site or previous surveys may lead to a rough estimate of $\kappa_{\max}$ and thus the spatial bandwidth of the signal.

We have seen that two different criteria applies at low and high SNR in order to reduce gross errors and fine errors, respectively. At low SNR, array design focuses on the design of the LF in order to reduce the height of local maxima and thus the occurrence of gross errors. We have shown how the Fourier transform of the spatial sampling pattern $H(\boldsymbol{\kappa})$ plays a central role in understanding the shape of the LF. At high SNR, array design targets the reduction of fine errors. This can be achieved by designing an array geometry that reduces the CRB in (20).

4.1 Reduction of gross errors

It is shown in the literature that the probability of occurrence of gross errors is related to the amplitude of the local maxima relative to the true maximum, see, for example, Athley (2005). Therefore, in order to decrease gross errors it is necessary to reduce the amplitude of local maxima of the LF over a certain region of the $\kappa_x\,\kappa_y$-plane. We choose this region as an *annulus* defined by $\kappa_{\min}$ and $2\kappa_{\max}$, i.e., the region bounded by two concentric circles with radii $\kappa_{\min}$ and $2\kappa_{\max}$, respectively. We now explain the rationale behind this design choice.

The circular symmetry of the considered region is motivated by the final application, the analysis of ambient vibration wavefields. In particular, by the fact that seismic waves may traverse the array of sensors from any direction-of-arrival (DOA).

We choose a value of $\kappa_{\min}$ related to the smallest spacing, in the $\kappa_x\,\kappa_y$-plane, between two signals that we wish to resolve. The smallest spacing between two resolvable signals is known in literature as the *Rayleigh resolution limit* (Van Trees 2001). The exact Rayleigh resolution limit can be computed from a given sampling pattern and it is in general slightly different from $\kappa_{\min}$. In practice, the quantity $\kappa_{\min}$ is also related to the smallest wavenumber that can be reliably estimated in field measurements.

We recall that the LF is related to the Fourier transform of the sampling pattern $H(\boldsymbol{\kappa})$ through a translation as in (13). We also observe that the ML estimate of the wave vector is found as the largest value of the LF in a disk of radius $\kappa_{\max}$ centred in the origin. A geometrical argument suggests that, in order to reduce gross errors, it is necessary to reduce the height of the local maxima of $|H(\boldsymbol{\kappa})|$ in a disk of radius $2\kappa_{\max}$.

We illustrate this argument with an example. In Fig. 3, we show graphically how the LF is related to a translation of the Fourier transform of the sampling pattern. Fig. 3(a) shows the magnitude of the Fourier transform of an optimized sampling pattern (the array layout is not shown). The two magenta circles have radii $\kappa_{\min}$ and $2\kappa_{\max}$, respectively. In this example $\kappa_{\min} = 1/4$ and $\kappa_{\max} = 1$. We observe the presence of the global maximum at (0, 0) inside the inner circle, that the local maxima in the annulus region are small in amplitude and that outside the outer ring there are several local maxima with large amplitude.

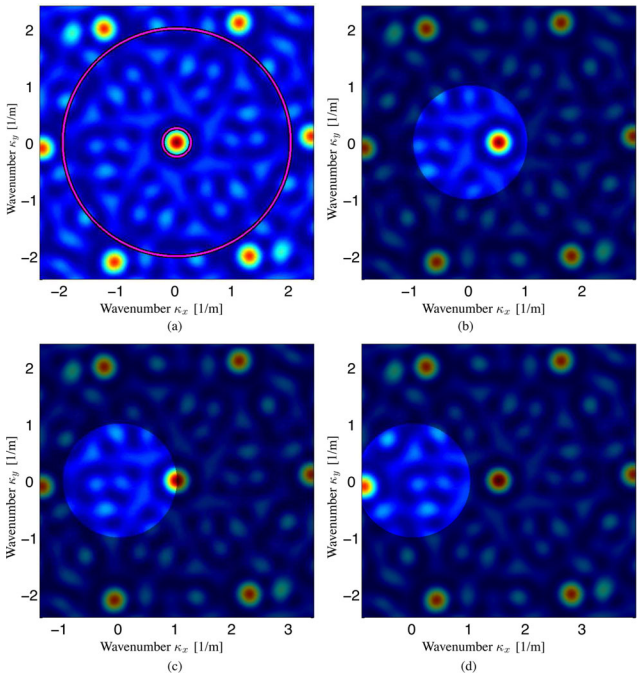


Figure 3. Illustration of the relationship between the magnitude of the Fourier transform of the sampling pattern and expected log-likelihood functions (LLFs) of different waves. Colours towards red (blue) represent large (small) values. In (a) the magnitude of the Fourier transform of an optimized sampling pattern is shown. The two magenta circles of radii $\kappa_{\min}$ and $2\kappa_{\max}$ limit the region where the local maxima need to be reduced. The highlighted disks in (b)–(d) depict the LLFs of the same array for different incoming waves. See text for a detailed explanation. (a) Fourier transform of an optimized sampling pattern. (b) Expected LLF, $\check{\boldsymbol{\kappa}} = (0.5, 0)$. (c) Expected LLF, $\check{\boldsymbol{\kappa}} = (1, 0)$. (d) Expected LLF, $\check{\boldsymbol{\kappa}} = (1.5, 0)$.

Fig. 3(b) depicts the expected LLF of a wave with wavenumber $\check{\boldsymbol{\kappa}} = (0.5, 0)$ as a highlighted disk of radius $\kappa_{\max} = 1$. Colours towards red represent high likelihood and colours towards blue low likelihood. The global maxima of the expected LLF is located at (0.5, 0). Observe that the shaded region of Fig. 3(b) is exactly the Fourier transform of the sampling pattern shown in Fig. 3(a). Fig. 3(c) depicts the expected LLF of a wave with wavenumber $\check{\boldsymbol{\kappa}} = (1, 0)$. This wave has the largest admissible wavenumber since $\|\check{\boldsymbol{\kappa}}\| = \kappa_{\max}$ and, therefore, the maxima of the LLF lies at the edge of the parameter space. Fig. 3(d) depicts the expected LLF of a wave with wavenumber $\check{\boldsymbol{\kappa}} = (1.5, 0)$. This wave has a wavenumber larger than what was assumed for the array design since $\|\check{\boldsymbol{\kappa}}\| > \kappa_{\max}$. The global maximum of the expected LLF corresponds to a local maxima of $|H(\boldsymbol{\kappa})|$ outside the minimization region. As results, the ML estimate of the wave vector is not corresponding to the true wave vector.

We formulate the following minimization program, aimed at minimizing the largest value of $|H(\boldsymbol{\kappa})|$ over the annulus defined

by $\kappa_{\min} \leq \|\boldsymbol{\kappa}\|_2 \leq 2\kappa_{\max}$ as

$$\min_{\boldsymbol{p}_1,\boldsymbol{p}_2,\ldots,\boldsymbol{p}_{N_s}} \max_{\boldsymbol{\kappa}} \left|H(\boldsymbol{\kappa},\boldsymbol{p}_1,\boldsymbol{p}_2,\ldots,\boldsymbol{p}_{N_s})\right| \tag{25a}$$

$$\text{s.t.} \quad \kappa_{\min} \leq \|\boldsymbol{\kappa}\|_2 \leq 2\kappa_{\max}\,, \tag{25b}$$

where we emphasized the dependence of H on sensor positions. In this optimization problem, the minimization variables are the sensor positions $\{\boldsymbol{p}_n\}_{n=1,\ldots,N_s}$. The target of the minimization, $\max_{\boldsymbol{\kappa}} |H(\boldsymbol{\kappa},\boldsymbol{p}_1,\boldsymbol{p}_2,\ldots,\boldsymbol{p}_{N_s})|$, is a function of the sensor positions. To the best of our knowledge, the use of such cost function is new to the seismological community.

This minimization problem becomes harder as the extent of the annulus is increased. In particular, we observe that large values of the ratio $\kappa_{\max}/\kappa_{\min}$ lead to harder optimization problems.

4.2 Reduction of fine errors

Fine errors can be reduced by decreasing the CRB. The CRB derived in (20) depends on sensor positions $\{\boldsymbol{p}_n\}_{n=1,\ldots,N_s}$ and on the azimuth ψ. Indeed, in general, different MSEE values are expected for different DOAs of the incoming wave.

We choose to maximize the inverse of the CRB for the DOA leading to the worst performance

$$\max_{\boldsymbol{p}_1,\boldsymbol{p}_2,\ldots,\boldsymbol{p}_{N_s}} \min_{\psi} \left(Q_{aa}(\psi,\boldsymbol{p}_1,\ldots,\boldsymbol{p}_{N_s}) - \frac{Q_{ab}^2(\psi,\boldsymbol{p}_1,\ldots)}{Q_{bb}(\psi,\boldsymbol{p}_1,\ldots)} \right). \tag{26}$$

Such maximization problem is unbounded. Indeed, it is possible to verify that a uniform circular array (UCA) with sufficiently large radius can make the objective function arbitrarily large. Moreover, an array with a very large aperture could void the assumption of planar wave fronts.

4.3 Discussion

Elaborating on the two distinct sources of errors discussed in Section 3, we showed how two different sensor placement criteria exist. Depending on the application, interest may lie in reducing one or the other type of error. It is also possible to combine both optimization problems (25) and (26) in order to account for both types of errors.

We emphasize that the choice of performing the analysis in the wavenumber domain is important since it allows us to disregard the temporal frequency ω. Indeed, the conclusions drawn in this section are independent of the temporal frequency and the proposed design criteria are valid at every frequency.

The findings in this section are in general applicable to beamforming methods. The conditions for the equivalence between the ML estimation method and classical beamforming are stated in Reller *et al.* (2011).

For ambient vibration wavefields, typically the waves have very small amplitudes compared to the level of the background noise. Therefore, we choose to primarily focus on improving performance at low SNR and thus on the reduction of gross errors.

Although we choose not to address fine errors explicitly, we consider arrays with $Q_{ab} = 0$. This additional constraint is beneficial for reducing the CRB of (20) and it also ensures that the resulting array has no preferential DOA, i.e., $Q_{aa}(\psi)$ is equal for every azimuth.

5 ARRAY DESIGN METHODS

To the best of our knowledge, there is no algorithm with polynomial complexity that can always find the global optimal solution to the program in (25). The objective function (25a) has many local minima and any solution found by, for example, a gradient descent method will strongly depend on the initial starting point. In this section, we propose an algorithm using a MIP, which is shown empirically to be effective. We also compare it with two other possible approaches addressing the same optimization problem.

5.1 Mixed integer program

The program in (25) is difficult to solve because the minimization variables $\{\boldsymbol{p}_n\}_{n=1,\ldots,N_s}$ are arguments of the complex exponentials. To overcome this limitation, instead of representing the sensor positions with continuous variables we consider a finite number of possible sensor positions. We represent the discrete sensor positions with the binary vector $\boldsymbol{x} \in \{0,\,1\}^N$. An analogous minimization problem in this new variable is

$$\min_{\boldsymbol{x}} \|\mathbf{F}\boldsymbol{x}\|_\infty \tag{27a}$$

$$\text{s.t.} \quad \mathbf{A}_s\boldsymbol{x} = \boldsymbol{b}_s \tag{27b}$$

$$\boldsymbol{x} \in \{0,1\}^N\,, \tag{27c}$$

where $\|\boldsymbol{v}\|_\infty = \max(|v_1|, |v_2|, \ldots, |v_M|)$.

The vector $\boldsymbol{x}$ is a binary vector representing the presence or the absence of a sensor at given spatial locations on the plane. Fig. 4(a) shows a possible choice of N feasible spatial locations for positioning the N_s available sensors. The linear operator $\mathbf{F} : \mathbb{R}^N \to \mathbb{C}^M$ computes the two-dimensional Fourier transform of the array positions, restricted to M discrete spatial frequencies of interest. The operator $\mathbf{F}$ can be thought as a discretized version of $H(\boldsymbol{\kappa})$. The M frequencies lie in the annulus defined by $\kappa_{\min} \leq \|\boldsymbol{\kappa}\|_2 \leq 2\kappa_{\max}$. Fig. 4(b) shows a possible choice of the M frequencies computed by $\mathbf{F}$. Observe that because of the symmetry $|H(\boldsymbol{\kappa})| = |H(-\boldsymbol{\kappa})|$ the choice of the M frequencies can be limited only to half of the $\kappa_x\,\kappa_y$-plane.

We choose to arrange both the N possible sensor positions and and the M spatial frequencies on circles around the origin, as in Fig. 4. This arbitrary choice is supported by the radial symmetry of the problem as expressed in (25). Concerning the sensor positions, this choice also makes it easier to enforce numerically any constraint on the MOIs.

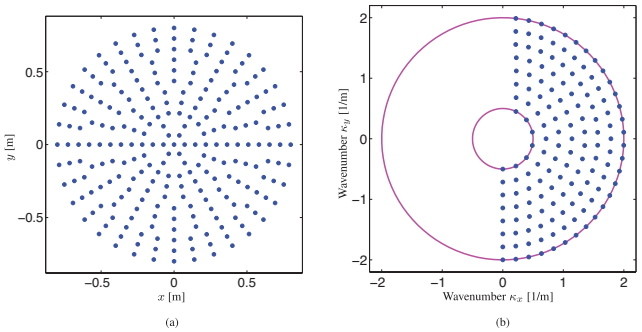


Figure 4. A possible choice of the N spatial locations and of the M spatial frequencies used in the construction of the operator $\mathbf{F}$. (a) Possible N locations available in the $x\,y$-plane for the placement of N_s sensors. (b) Possible positions of the M spatial frequencies in the $\kappa_x\,\kappa_y$-plane. The magenta circles depict $\kappa_{\min}$ and $2\kappa_{\max}$.

The equalities in (27b) enforce a number of linear constraints. The number of sensors is enforced to be equal to N_s. The vector $\boldsymbol{x}$ has exactly N_s elements equal to one, corresponding to the sensor positions. The remaining elements are zero. As an extension, using linear equalities is also possible to enforce the presence or the absence of a sensor in a specific position.

In (27b), we also enforce the linear constraint $Q_{ab} = 0$. This constraint allows to eliminate the term $-Q_{ab}^2/Q_{bb}$ from the expression of the CRB in (20), thus lowering the CRB. It also ensures that the array performance, in terms of fine errors, have no preferential direction, i.e., Q_{aa} and Q_{bb} are constant for every azimuth.

Comparing the original problem in (25) and the discretized formulation in (27), we observe that the discretization of the sensor positions causes the optimal value of (25) to be smaller than or equal to the optimal value of (27). This is especially important when the vector $\boldsymbol{x}$ has small dimension N and the possible sensor positions are coarse.

Observe that the objective function (27a) is a convex function of the minimization variable $\boldsymbol{x}$. However, due to the binary constraint (27c), the minimization is an integer programming problem. Integer programs are generally considered to be NP-hard (non-deterministic polynomial-time hard), i.e., there is no polynomial time algorithm to solve them.

We relax the optimization problem in (27) to make it more tractable for implementation. We replace the convex objective function with a linear objective function as

$$\|\mathbf{F}\boldsymbol{x}\|_\infty \rightarrow \left\| \begin{pmatrix} |\mathrm{Re}\,\mathbf{F}\boldsymbol{x}| \\ |\mathrm{Im}\,\mathbf{F}\boldsymbol{x}| \end{pmatrix} \right\|_\infty .$$

This modification enables us to formulate the problem as a MIP with linear objective function and linear constraints

$$\min_y \; y \tag{28a}$$

$$\text{s.t.} \quad \mathbf{A}_s\boldsymbol{x} = \boldsymbol{b}_s \tag{28b}$$

$$\begin{pmatrix} \mathrm{Re}\,\mathbf{F} \\ \mathrm{Im}\,\mathbf{F} \\ -\mathrm{Re}\,\mathbf{F} \\ -\mathrm{Im}\,\mathbf{F} \end{pmatrix} \boldsymbol{x} \preceq \mathbf{1}y \tag{28c}$$

$$y \in \mathbb{R} \tag{28d}$$

$$\boldsymbol{x} \in \{0,1\}^N , \tag{28e}$$

where $\mathbf{1}$ is a vector of ones of size $4M \times 1$ and $\preceq$ denotes element-wise $\leq$. In this program, there are N binary variables and one continuous real variable. This can be addressed using general purpose MIP algorithms (Gurobi Optimization, Inc. 2014). It is in theory possible to find the optimal solution to (28) using the branch and bound algorithm (Land & Doig 1960). However, for a large number of possible sensor positions N, finding the optimal solution becomes not always practical.

5.2 Genetic algorithm

We also attempt a direct minimization of (25) using genetic algorithm (GA; Goldberg 1989). Such algorithm attempts to find good solutions using some random search pattern and there is no guarantee to find the optimal solution.

In our implementation, we do not enforce any constraint on the MOIs. The constraint on the $Q_{ab} = 0$ is non-linear in the variables $\{\boldsymbol{p}_n\}_{n=1,\ldots,N_s}$. We do not enforce such constraint in the considered GA technique since we observed that it makes the minimization considerably harder.

5.3 Uniform circular array

In addition, we compare with the best UCA. In an UCA, sensors are uniformly spaced on a ring of radius r. A line search over the possible r allows us to obtain the radius of the best UCA for given $\kappa_{\min}$, $\kappa_{\max}$ and N_s

$$\min_r \max_{\boldsymbol{\kappa}} \left| H(\boldsymbol{\kappa}, \boldsymbol{p}_1, \boldsymbol{p}_2, \ldots, \boldsymbol{p}_{N_s}) \right| \tag{29a}$$

$$\text{s.t.} \quad \boldsymbol{p}_n = r(\cos(2\pi n/N_s), \sin(2\pi n/N_s)) \tag{29b}$$

$$\kappa_{\min} \leq \|\boldsymbol{\kappa}\|_2 \leq 2\kappa_{\max} . \tag{29c}$$

The objective function is the same of the original problem (25). Sensors are restricted to be uniformly spaced on a circle of radius r.

We observe that all UCA with $N_s \geq 3$ have $Q_{ab} = 0$. Moreover, Q_{aa} and Q_{bb} are constants for all azimuths.

6 NUMERICAL RESULTS

In this section, we compare the arrays designed using the considered techniques and we quantify the impact different geometries have on the estimation problem. In Section 6.1, the outcomes of the different array design methods are compared in terms of the design criteria presented in Section 4. In Section 6.2, the estimation performance achieved using different array layouts is analysed by means of Monte Carlo simulations.

6.1 Array design

The three array design techniques considered are: (i) the proposed approach, i.e., the MIP of (28); (ii) the direct minimization of (25) using a GA; and (iii) the best UCA obtained from the program in (29).

6.1.1 Gross errors

In terms of sensor design and minimization of the original problem (25), the goodness of a solution is quantified with the amplitude of the largest local maxima of $|H(\boldsymbol{\kappa})|$ compared to the central maximum $|H(0, 0)|$. Therefore, we consider the quantity

$$H_{\max} = \max_{\boldsymbol{\kappa}} \left(\frac{|H(\boldsymbol{\kappa})|^2}{N_s^2} \right) \tag{30}$$

for $\boldsymbol{\kappa}_{\min} \leq \|\boldsymbol{\kappa}\|_2 \leq 2\boldsymbol{\kappa}_{\max}$. Observe that the quantity $H_{\max}$ is smaller than or equal to 1, since the largest value of $|H(\boldsymbol{\kappa})|$ is N_s.

Fig. 5 shows the value of $H_{\max}$ achieved with different design techniques for different number of sensors N_s and $\kappa_{\max}/\kappa_{\min}$. In general, the value of $H_{\max}$ decreases for increasing number of sensors N_s. In particular, such decrease is steeper for few sensors. This indicates that the marginal benefit of an additional sensor is greater for few sensors.

Different values of $\kappa_{\max}/\kappa_{\min}$ are also considered. Figs 5(a), (b) and (c) depict the minimization results for $\kappa_{\max} = 1$ and $\kappa_{\min}$ equal

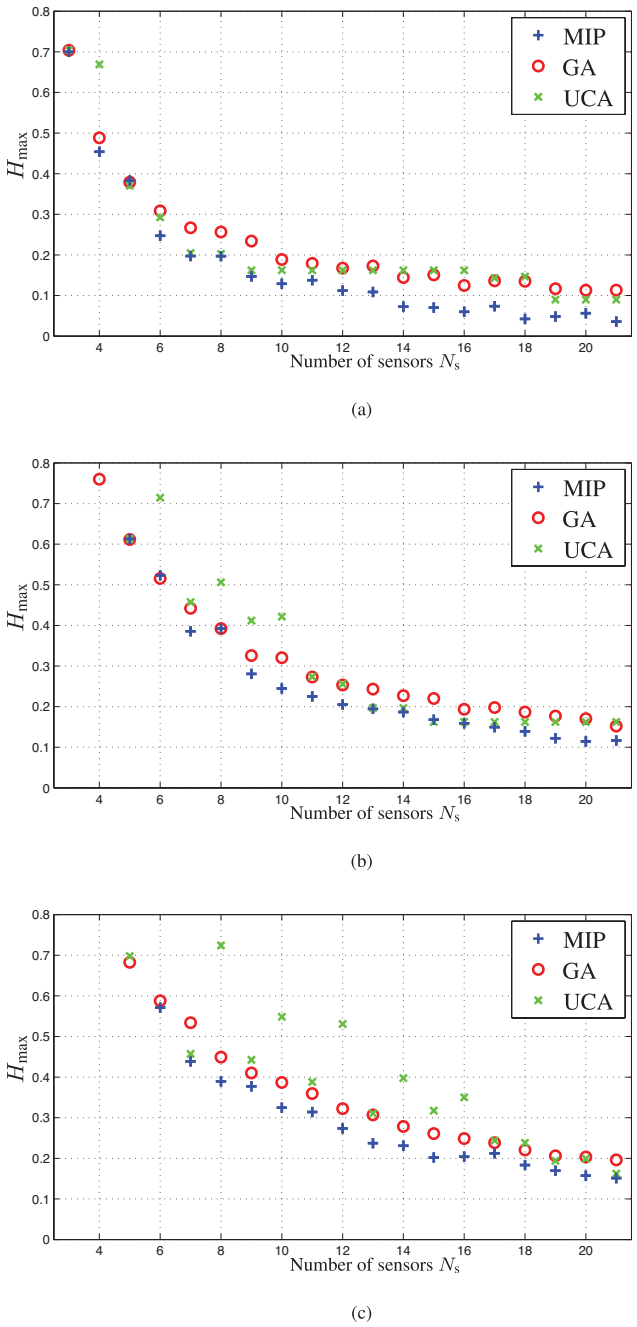


Figure 5. Values of H_{max} achieved with different array design strategies for different number of sensors. Different $\kappa_{max}/\kappa_{min}$ are considered. A small H_{max} is desirable in order to reduce the occurrence of gross errors. (a) $\kappa_{min} = 1/2$ and $\kappa_{max} = 1$. (b) $\kappa_{min} = 1/4$ and $\kappa_{max} = 1$. (c) $\kappa_{min} = 1/6$ and $\kappa_{max} = 1$.

to 1/2, 1/4 and 1/6, respectively. A large value of $\kappa_{max}/\kappa_{min}$ corresponds to a larger extension of the annulus involved in the minimization. For larger values of $\kappa_{max}/\kappa_{min}$, the minimization problem is harder and the H_{max} values found are indeed larger.

In Fig. 5, we observe that the MIP technique typically achieves values of H_{max} smaller than the other design techniques, for all the considered $\kappa_{max}/\kappa_{min}$.

The GA technique is able to consistently decrease H_{max} for increasing number of sensors. On the contrary, the UCA technique exhibits, for certain number of sensors, larger H_{max} than for fewer sensors. This behaviour is also observed in the MIP technique, but to a much smaller extent. The GA solutions are in general worse than the MIP solutions in terms of H_{max}.

In terms of minimization, an intrinsic advantage of the GA technique is that the minimization variables $\{\boldsymbol{p}_n\}_{n=1,\ldots,N_s}$ are continuous, this may help the algorithm to optimize the solution, at least locally. The GA solution presented is the best out of hundreds of runs performed with different initialization of the GA algorithm.

The UCA's technique performs similarly or worse than the other techniques depending on the specific design parameters. We observe that for $\kappa_{max}/\kappa_{min} = 2$, there is little or no decrease in the value of H_{max} for increasing number of sensors above 10. For $\kappa_{max}/\kappa_{min}$ equals to 4 and 6, the UCA is often largely outperformed by the other two design techniques.

The MIP technique exhibits H_{max} values smaller than the other techniques in most of the scenarios considered.

6.1.2 *Fine errors*

At high SNR, the performance of the ML estimator is well characterized by the CRB. From (20) it is clear how the CRB performance depends on the azimuth. We define the quantity Q_{min}, related to the CRB at the azimuth ψ exhibiting the worst performance as

$$Q_{min} = \min_{\psi} \left(Q_{aa}(\psi) - \frac{Q_{ab}^2(\psi)}{Q_{bb}(\psi)} \right) . \tag{31}$$

Note that Q_{min} is not explicitly taken into account in any of the strategies considered for array design. However, the MIP and the UCA geometries are guaranteed to have $Q_{ab} = 0$ for any azimuth.

Fig. 6 depicts the value of Q_{min} for the same arrays considered in Fig. 5(b). As expected, the value of Q_{min} is generally an increasing function of the number of sensors.

6.1.3 *Array geometry*

Example array geometries are depicted in Fig. 7. The array geometries obtained for $N_s = 14$ and $\kappa_{min} = 1/4$ with different techniques are compared together with a representation of the corresponding $|H(\boldsymbol{\kappa})|^2$.

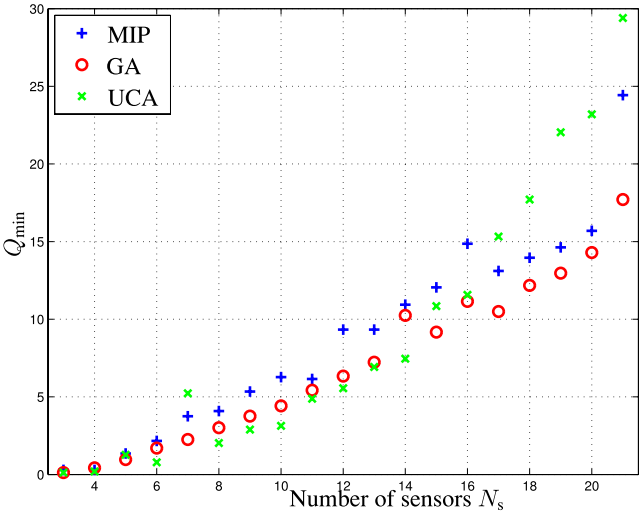


Figure 6. Values of Q_{min} obtained from different array design strategies for different number of sensors and $\kappa_{max}/\kappa_{min} = 4$. A large Q_{min} is desirable in order to reduce the MSEE at high SNR.

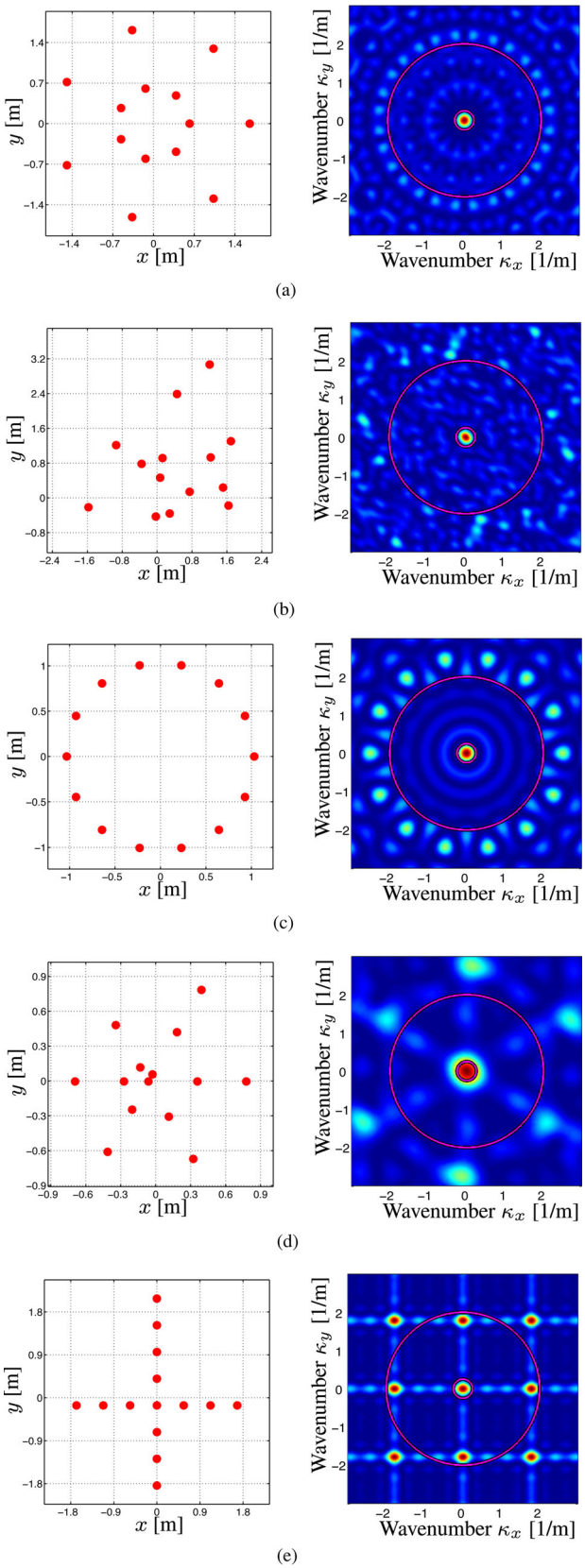


Figure 7. Array geometries $h(x, y)$ and the corresponding $|H(\kappa_x, \kappa_y)|^2$. The arrays are obtained by choosing $N_s = 14$, $\kappa_{min} = 0.25$ and $\kappa_{max} = 1$. The magenta circles delimit the annulus $\kappa_{min} < \|\boldsymbol{\kappa}\|_2 < 2\kappa_{max}$. (a) Mixed integer program (MIP) array. (b) Genetic algorithm (GA) array. (c) Uniform circular array (UCA) array. (d) Spiral array. (e) Cross-array.

The array obtained using the proposed MIP method, in Fig. 7(a), exhibits symmetry around the origin due to the constraint $Q_{ab} = 0$. The GA array of Fig. 7(b) exhibit a completely irregular pattern since the sensor positions are unconstrained. The UCA array is shown in Fig. 7(c). A spiral-shaped array is shown in Fig. 7(d). From the large width of the central lobe, a large MSEE at high SNR is to be expected for the spiral array. A cross-shaped array is shown in Fig. 7(e). Due to the regular geometry, the cross-array exhibits many maxima having the same magnitude as the central lobe. In these two latter arrays, sensor spacings are chosen arbitrarily and the arrays are not designed to comply with the proposed criteria.

Values of H_{max} and Q_{min} for the arrays shown in Figs 7(a)–(c) can be found in Figs 5(b) and 6, respectively.

The symmetry of the MIP array and the irregular pattern of the GA array are also observed for other sensor numbers and values of $\kappa_{max}/\kappa_{min}$. This fact suggests that the actual deployment in the field of a MIP array may be significantly easier that the deployment of a GA array. Additional array layouts obtained with the MIP technique are shown in Appendix B.

6.1.4 About MIP algorithm execution time

As mentioned in Section 5.1, it is not always practically possible to find the optimal solution to a MIP problem. For certain choices of N_s and arrangement of the possible spatial locations, the globally optimal solution was found. For large N_s and N the algorithm was often terminated after a given time limit and a suboptimal solution was considered.

We observed that the quality of the MIP solutions and the time necessary to find good solutions depend on the choice of the N possible sensor locations, *cf.* Fig. 4(a). In particular, we observed that the number of possible sensor locations on each concentric circle has to be related to N_s. One reason is that due to the $Q_{ab} = 0$ constraint and the arrangement of the possible sensor locations, the MIP solutions exhibit sensors placed on concentric circles. Therefore, certain choices of the number of possible locations on each circle may turn the MIP problem infeasible, and other choices may be more convenient. In addition, design requirements with large N_s or large $\kappa_{max}/\kappa_{min}$ required more time.

The actual execution time varied between hours and few tens of hours. Array design was performed on a machine with 20 cores at 2 GHz.

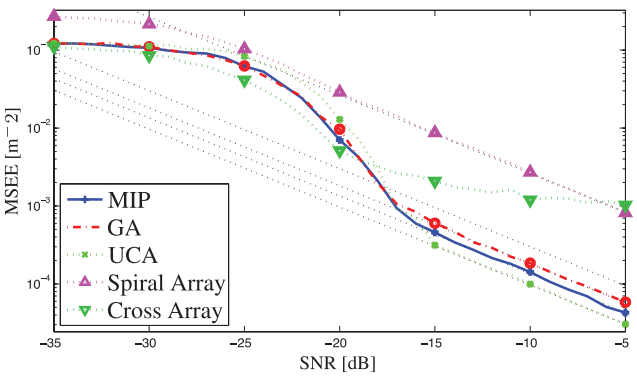
6.2 Estimator performance

In order to quantify the impact of the different arrays geometry on the actual estimation problem, we resort to the analysis of the MSEE using Monte Carlo simulations.

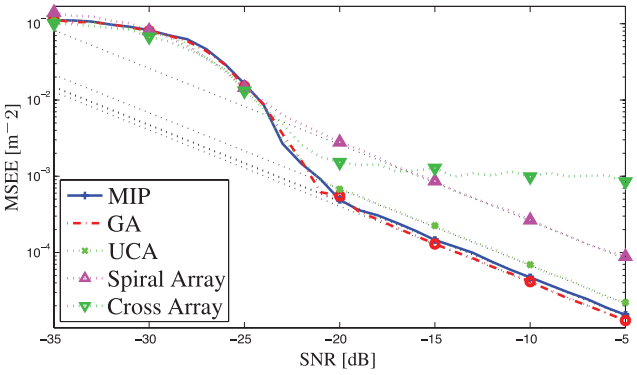
Arrays obtained with the sensor placement techniques considered in the previous section are compared. Moreover, we compare with an array having sensors arranged on a spiral and with another array with sensors arranged on a cross. These two latter arrays are designed without following the design criterion proposed in this work.

The MSEE is computed as follows. A wave vector is drawn randomly from the uniform distribution having as support a disk with radius κ_{max}. For each considered array geometry and SNR, random noise is added to the wavefield as in (5). The ML estimation method of Maranò *et al.* (2012) is used to estimate the parameters of the wave and the MSEE is obtained by repeating the procedure 3000 times. The MSEEs are also compared with the CRBs.

In Fig. 8, we compare the MSEE of the ML estimate of the wavenumber for different array geometries. The MIP, GA and UCA



(a)



(b)

Figure 8. Comparison of the MSEE of the wavenumber obtained with different array geometries. The CRBs are also depicted with black thin dotted lines. (a) $N_s = 7$. (b) $N_s = 14$.

arrays are designed with $\kappa_{max}/\kappa_{min} = 4$. In general, the MSEE decreases as the SNR increases and it achieves the CRB for sufficiently large SNR.

Fig. 8(a) shows the MSEE for an array of $N_s = 7$ sensors. The MIP, the GA and the UCA exhibit similar performances in the threshold region. At high SNR, in the asymptotic region, certain arrays perform better than others. We observe how the best performing arrays are those with larger Q_{min}, *cf.* Fig. 6. The spiral array shows a large MSEE over all the SNR range considered. The cross-array does not achieve its CRB because the local maxima of the LLF are indistinguishable from the right maximum. This ambiguity is due to the regular array geometry, *cf.* Fig. 7(e).

Fig. 8(b) shows the MSEE for an array of $N_s = 14$ sensors. The geometry of these arrays is shown in Fig. 7. The MIP, the GA and the UCA arrays exhibit similar performance both at low and high SNR. The spiral array and the cross-array again exhibit the larger MSEE.

In Fig. 9, the MSEE of the ML estimate of the wavenumber for different array geometries and number of sensors is depicted. Only the MIP geometries, with $\kappa_{max}/\kappa_{min} = 4$, are considered for different values of the SNR. As expected, the MSEE decreases when the number of sensors is increased. For SNR $= -25$ dB, the MSEE is constantly well above the CRB because the SNR is low and the estimator operates in the threshold region, thus gross errors occur. For SNR $= -20$ dB, the MSEE achieves the CRB for a sufficiently large number of sensors. At last, a SNR of -15 dB is sufficient for the estimator to operate in the asymptotic region for any number of sensors.

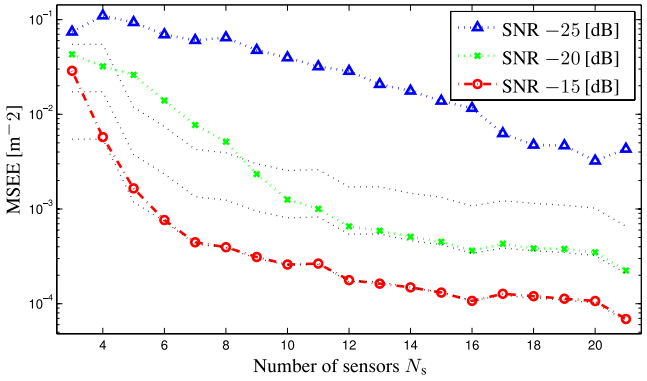


Figure 9. Comparison of the MSEE of the wavenumber obtained with different array geometries having a different number of sensors. The CRBs are also depicted with black thin dotted lines.

In Fig. 9, it is also observed that the marginal benefit of an additional sensor also depends on the SNR. When the estimator operates in the asymptotic region (e.g. SNR $= -15$ dB), most of the MSEE reduction is achieved with the first 8 sensors. However, for SNR $= -20$ dB, the MSEE reduction is still significant until $N_s = 11$ because the threshold effect of the estimator. These considerations may also vary for different $\kappa_{max}/\kappa_{min}$ as suggested by Fig. 5.

7 CONCLUSIONS

In this paper, we described in detail the occurrence of gross errors and fine errors in the estimation of parameters of Love and Rayleigh waves. We derived a relationship between the Fourier transform of the sampling pattern and the average LF of the observations. Sensor placement criteria for the reduction of gross errors and fine errors are then proposed.

An array design algorithm employing the proposed criterion with linear objective function and linear constraints is formulated as a MIP. The proposed algorithm is compared with the GA technique and with the UCA technique using the same design criterion. The proposed algorithm achieves superior design for most choices of the sensor number and spatial frequency requirements. In addition, the proposed sensor placement algorithm generates arrays composed of simple regular geometries and are thus suitable for field deployment.

The MSEE of the ML estimator using different array geometries are compared by means of Monte Carlo simulations. We find that the proposed design criterion is suitable for the optimized placement of seismic sensors for the analysis of Love and Rayleigh waves. We show that the sensor placement techniques considered can greatly reduce the MSEE when compared with non optimized arrays.

We emphasize that prior knowledge of the spatial frequency support of the wavefield is useful to design an array able to achieve better performance. From our findings, we suggest that a minimum number of around 10 sensors is desirable.

ACKNOWLEDGEMENTS

This work is supported in part by the Swiss Commission for Technology and Innovation under project 9260.1 PFIW-IW and with funds of the Swiss Seismological Service. The first author would like to thank Dr. Michel Baes and Dr. Michael Bürgisser for the helpful discussions.

REFERENCES

Aki, K. & Richards, P.G., 1980. *Quantitative Seismology,* W. H. Freeman and Company.

Asten, M.W. & Henstridge, J.D., 1984. Array estimators and the use of microseisms for reconnaissance of sedimentary basins, *Geophysics,* **49**(11), 1828–1837.

Athley, F., 2005. Threshold region performance of maximum likelihood direction of arrival estimators, *IEEE Trans. Signal Process.,* **53**(4), 1359–1373.

Bevelacqua, P.J. & Balanis, C.A., 2007. Optimizing antenna array geometry for interference suppression, *IEEE Trans. Antennas Propag.,* **55**(3), 637–641.

Bonnefoy-Claudet, S., Cotton, F. & Bard, P.-Y., 2006. The nature of noise wavefield and its applications for site effects studies: a literature review, *Earth-Sci. Rev.,* **79**(3–4), 205–227.

Cornou, C., Bard, P.-Y. & Dietrich, M., 2003. Contribution of dense array analysis to the identification and quantification of basin-edge-induced waves, part II: application to Grenoble basin, French alps, *Bull. seism. Soc. Am.,* **93**(6), 2624–2648.

Dogandžić, A. & Nehorai, A., 2001. Cramér-Rao bounds for estimating range, velocity, and direction with an active array, *IEEE Trans. Signal Process.,* **49**(6), 1122–1137.

Gaffet, S., Larroque, C., Deschamps, A. & Tressols, F., 1998. A dense array experiment for the observation of waveform perturbations, *Soil Dyn. Earthq. Eng.,* **17,** 475–484.

Goldberg, D.E., 1989. *Genetic Algorithms in Search, Optimization, and Machine Learning,* Addison-Wesley Professional.

Gurobi Optimization, Inc., 2014. *Gurobi Optimizer Reference Manual*. http://www.gurobi.com.

Haubrich, R.A., 1968. Array design, *Bull. seism. Soc. Am.,* **58**(3), 977–991.

Horike, M., 1985. Inversion of phase velocity of long-period microtremors to the S-wave-velocity structure down to the basement in urbanized areas., *J. Phys. Earth.,* **33**(2), 59–96.

Kay, S.M., 1993. *Fundamentals of Statistical Signal Processing: Estimation Theory,* Prentice Hall.

Kind, F., Fäh, D. & Giardini, D., 2005. Array measurements of S-wave velocities from ambient vibrations, *Geophys. J. Int.,* **160,** 114–126.

Land, A.H. & Doig, A.G., 1960. An automatic method of solving discrete programming problem, *Econometrica,* **28**(3), 497–520.

Maranò, S. & Fäh, D., 2014. Processing of translational and rotational motions of surface waves: performance analysis and applications to single sensor and to array measurements, *Geophys. J. Int.,* **196**(1), 317–339.

Maranò, S., Reller, C., Loeliger, H.-A. & Fäh, D., 2012. Seismic waves estimation and wavefield decomposition: application to ambient vibrations, *Geophys. J. Int.,* **191**(1), 175–188.

Milana, G., Barba, S., Pezzo, E.D. & Zambonelli, E., 1996. Site response from ambient noise measurements: new perspectives from an array study in central Italy, *Bull. seism. Soc. Am.,* **86,** 320–328.

Ohori, M., Nobata, A. & Wakamatsu, K., 2002. A comparison of ESAC and FK methods of estimating phase velocity using arbitrarily shaped microtremor arrays, *Bull. seism. Soc. Am.,* **92**(6), 2323–2332.

Okada, H., 2006. Theory of efficient array observations of microtremors with special reference to the SPAC method, *Explor. Geophys.,* **37,** 73–85.

Poggi, V. & Fäh, D., 2010. Estimating Rayleigh wave particle motion from three-component array analysis of ambient vibrations, *Geophys. J. Int.,* **180**(1), 251–267.

Reller, C., Loeliger, H.-A. & Maranò, S., 2011. Multi-sensor estimation and detection of phase-locked sinusoids, in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing,* Prague, Czech Republic, pp. 3872–3875.

Rost, S. & Thomas, C., 2002. Array seismology: methods and applications, *Rev. Geophys.,* **40**(3), 2-1–2-27.

Satoh, T., Kawase, H., Iwata, T., Higashi, S., Sato, T., Irikura, K. & Huang, H.-C., 2001a. S-wave velocity structure of the Taichung basin, Taiwan, estimated from array and single-station records of microtremors, *Bull. seism. Soc. Am.,* **91,** 1267–1282.

Satoh, T., Kawase, H. & Matsushima, S., 2001b. Estimation of S-wave velocity structures in and around the Sendai basin, Japan, using array records of microtremors, *Bull. seism. Soc. Am.,* **91**(2), 206–218.

Stoica, P., Moses, R., Friedlander, B. & Söderstrom, T., 1989. Maximum likelihood estimation of the parameters of multiple sinusoids from noisy measurements, *IEEE Trans. Acoust., Speech, Signal Process.,* **37**(3), 378–392.

Tokimatsu, K., 1997. Geotechnical site characterization using surface waves, in *Proceedings of the 1st International Conference on Earthquake Geotechnical Engineering,* Balkema, pp. 1333–1368.

Van Trees, H.L., 2001. *Detection, Estimation, and Modulation Theory,* John Wiley & Sons, Inc.

Van Trees, H.L., 2002. *Optimum Array Processing,* John Wiley & Sons, Inc.

Vertatschitsch, E.J. & Haykin, S., 1991. Impact of linear array geometry on direction-of-arrival estimation for a single source, *IEEE Trans. Antennas Propag.,* **39**(5), 576–584.

Wathelet, M., Jongmans, D., Ohrnberger, M. & Bonnefoy-Claudet, S., 2008. Array performances for ambient vibrations on a shallow structure and consequences over v_s inversion, *J. Seismol.,* **12,** 1–19.

Woods, J.W. & Lintz, P.R., 1973. Plane waves at small arrays, *Geophysics,* **38**(6), 1023–1041.

APPENDIX A: RELATIONSHIP BETWEEN LIKELIHOOD FUNCTION AND SAMPLING PATTERN

In this section, we show the relationship between the LF $p_Y(\tilde{\boldsymbol{y}}|\boldsymbol{\theta})$ and the Fourier transform of the sampling pattern $H(\kappa_x, \kappa_y)$.

A1 Scalar case

We consider the scalar wave equation of (3) and the measurement model of (4). We parametrize the wave equation as a function of $\{\alpha_1, \alpha_2\}$ instead of $\{\alpha_0, \varphi_0\}$ using

$$\alpha_1 = \alpha_0 \cos\varphi_0$$

$$\alpha_2 = -\alpha_0 \sin\varphi_0 .$$

With respect to ML estimation, this parametrization is equivalent because the transformation between the two parameters pairs is bijective.

The measurement $Y_k^{(n)}$, at the time instant t_k and location $\boldsymbol{p}_n$ is therefore

$$\begin{aligned} Y_k^{(n)} &= u(\boldsymbol{p}_n, t_k) + Z_k^{(n)} \\ &= \alpha_0 \cos(\omega t - \boldsymbol{\kappa}\cdot\boldsymbol{p} + \varphi_0) + Z_k^{(n)} \\ &= \alpha_1 \cos(\omega t - \boldsymbol{\kappa}\cdot\boldsymbol{p}) + \alpha_2 \sin(\omega t - \boldsymbol{\kappa}\cdot\boldsymbol{p}) + Z_k^{(n)} . \end{aligned}$$

For a given $\boldsymbol{\kappa}$, the ML estimates $\hat{\alpha}_1$ and $\hat{\alpha}_2$ can be found analytically and are

$$\hat{\alpha}_1 = \frac{2}{N_s K} \sum_{n=1}^{N_s} \sum_{k=1}^{K} Y_k^{(n)} \cos(\omega t_k - \boldsymbol{\kappa}\cdot\boldsymbol{p}_n),$$

$$\hat{\alpha}_2 = -\frac{2}{N_s K} \sum_{n=1}^{N_s} \sum_{k=1}^{K} Y_k^{(n)} \sin(\omega t_k - \boldsymbol{\kappa}\cdot\boldsymbol{p}_n),$$

and their expected values are

$$\mathsf{E}\{\hat{\alpha}_1\} = \frac{2}{N_\mathrm{s}K}\sum_{n=1}^{N_\mathrm{s}}\sum_{k=1}^{K}(\check{\alpha}_1\cos(\check{\omega}t_k - \check{\boldsymbol{\kappa}}\cdot\boldsymbol{p}_n)\cos(\omega t_k - \boldsymbol{\kappa}\cdot\boldsymbol{p}_n) + \check{\alpha}_2\sin(\check{\omega}t_k - \check{\boldsymbol{\kappa}}\cdot\boldsymbol{p}_n)\cos(\omega t_k - \boldsymbol{\kappa}\cdot\boldsymbol{p}_n))\,,$$

$$\mathsf{E}\{\hat{\alpha}_2\} = -\frac{2}{N_\mathrm{s}K}\sum_{n=1}^{N_\mathrm{s}}\sum_{k=1}^{K}(\check{\alpha}_1\cos(\check{\omega}t_k - \check{\boldsymbol{\kappa}}\cdot\boldsymbol{p}_n)\sin(\omega t_k - \boldsymbol{\kappa}\cdot\boldsymbol{p}_n) + \check{\alpha}_2\sin(\check{\omega}t_k - \check{\boldsymbol{\kappa}}\cdot\boldsymbol{p}_n)\sin(\omega t_k - \boldsymbol{\kappa}\cdot\boldsymbol{p}_n))\,,$$

where the superscript $\check{}$ denotes the true value of the parameter and the lack of superscript denotes search parameters, for example, the argument of the LF. The expectations are taken with respect to $Z_k^{(\ell)}$.

When $\boldsymbol{\kappa} = \check{\boldsymbol{\kappa}}$, then $\mathsf{E}\{\hat{\alpha}_1\} \approx \check{\alpha}_1$ and $\mathsf{E}\{\hat{\alpha}_2\} \approx \check{\alpha}_2$. The symbol $\approx$ denotes an approximation. We used the following trigonometric approximations

$$\sum_{k=1}^{K}\cos^2(\omega k + \gamma) \approx \frac{K}{2},$$

$$\sum_{k=1}^{K}\sin^2(\omega k + \gamma) \approx \frac{K}{2},$$

$$\sum_{k=1}^{K}\sin(\omega k + \gamma)\cos(\omega k + \gamma) \approx 0\,,$$

which are valid for ω being not near 0 or $1/2$ and are exact when $\omega = \frac{2\pi m}{K}$ $m \in \mathbb{Z}$ [see example 3.14 in Kay (1993) or Stoica *et al.* (1989)].

The second moments of $\hat{\alpha}_1$ and $\hat{\alpha}_2$ are

$$\mathsf{E}\left\{\hat{\alpha}_1^2\right\} \approx \mathsf{E}\{\hat{\alpha}_1\}^2 + \frac{2}{N_\mathrm{s}K}\sigma^2$$

and

$$\mathsf{E}\left\{\hat{\alpha}_2^2\right\} \approx \mathsf{E}\{\hat{\alpha}_2\}^2 + \frac{2}{N_\mathrm{s}K}\sigma^2\,,$$

respectively.

From the previous derivations, it is clear to see the dependence on $\boldsymbol{\kappa}$ of $\hat{\alpha}_1(\boldsymbol{\kappa})$ and $\hat{\alpha}_2(\boldsymbol{\kappa})$. Therefore, given observations $\tilde{\boldsymbol{y}}$ and plugging $\hat{\alpha}_1$ and $\hat{\alpha}_2$ into the PDF (5), it is possible to rewrite the LF of the observation as a sole function of the $\boldsymbol{\kappa}$

$$p_Y(\tilde{\boldsymbol{y}}|\boldsymbol{\kappa}) = \max_{\alpha_0,\varphi_0} p_Y(\tilde{\boldsymbol{y}}|\boldsymbol{\theta}) = p_Y(\tilde{\boldsymbol{y}}|\hat{\alpha}_1(\boldsymbol{\kappa}), \hat{\alpha}_2(\boldsymbol{\kappa}), \boldsymbol{\kappa})\,.$$

In order to investigate the shape of the LF, we take the natural logarithm of the LF and drop all the additive and multiplicative constants not depending on $\boldsymbol{\kappa}$

$$\ln(p_Y(\boldsymbol{y}|\boldsymbol{\kappa})) \propto -\sum_{n=1}^{N_\mathrm{s}}\sum_{k=1}^{K}\left(Y_k^{(n)} - u_k^{(n)}(\boldsymbol{\theta})\right)^2 = -\sum_{n=1}^{N_\mathrm{s}}\sum_{k=1}^{K}\left(u_k^{(n)}(\check{\boldsymbol{\theta}}) + Z_k^{(n)} - u_k^{(n)}(\boldsymbol{\theta})\right)^2,$$

where again we distinguish between the true and unknown wavefield $\check{u}_k^{(n)} = u_k^{(n)}(\check{\boldsymbol{\theta}})$, and the wavefield as a function of the search parameters $u_k^{(n)} = u_k^{(n)}(\boldsymbol{\theta})$. The symbol $\propto$ denotes equality up to an affine transform.

In order to get insight about the average shape of the function, we take the expectation of such quantity

$$\begin{aligned}\mathsf{E}\{\ln(p_Y(\boldsymbol{y}|\boldsymbol{\kappa}))\} &\propto \mathsf{E}\left\{-\sum_{n=1}^{N_\mathrm{s}}\sum_{k=1}^{K}\left(\check{u}_k^{(n)} + Z_k^{(n)} - u_k^{(n)}\right)^2\right\}\\ &\propto \mathsf{E}\{\hat{\alpha}_1\}^2 + \mathsf{E}\{\hat{\alpha}_2\}^2\\ &= |\mathsf{E}\{\hat{\alpha}_1\} + i\mathsf{E}\{\hat{\alpha}_2\}|^2\\ &= \left|\sum_{n=1}^{N_\mathrm{s}}\sum_{k=1}^{K}\check{u}_k^{(n)}e^{-i(\omega t_k - \boldsymbol{\kappa}\cdot\boldsymbol{p}_n)}\right|^2\\ &= \left|\check{\alpha}_0 e^{-i\check{\varphi}_0}G(\omega - \check{\omega})H(\boldsymbol{\kappa} - \check{\boldsymbol{\kappa}})\right|^2,\end{aligned}$$

where we dropped the quantities $\mathsf{E}\{\sum_{n=1}^{N_\mathrm{s}}\sum_{k=1}^{K}(\check{u}_k^{(n)})^2\}$ and $\mathsf{E}\{\sum_{n=1}^{N_\mathrm{s}}\sum_{k=1}^{K}(Z_k^{(n)})^2\}$ since they are constants with respect to $\boldsymbol{\kappa}$. Moreover, we used

$$\begin{aligned}\sum_{n=1}^{N_\mathrm{s}}\sum_{k=1}^{K}u_k^{(n)}\check{u}_k^{(n)} &= \sum_{n=1}^{N_\mathrm{s}}\sum_{k=1}^{K}(\hat{\alpha}_1\cos(\omega t_k - \boldsymbol{\kappa}\cdot\boldsymbol{p}_n) + \hat{\alpha}_2\sin(\omega t_k - \boldsymbol{\kappa}\cdot\boldsymbol{p}_n))\\ &\quad\cdot(\check{\alpha}_1\cos(\check{\omega}t_k - \check{\boldsymbol{\kappa}}\cdot\boldsymbol{p}_n) + \check{\alpha}_2\sin(\check{\omega}t_k - \check{\boldsymbol{\kappa}}\cdot\boldsymbol{p}_n))\\ &\approx \frac{N_\mathrm{s}L}{2}(\hat{\alpha}_1\mathsf{E}\{\hat{\alpha}_1\} + \hat{\alpha}_2\mathsf{E}\{\hat{\alpha}_2\})\end{aligned}$$

and

$$\begin{aligned}\sum_{n=1}^{N_\mathrm{s}}\sum_{k=1}^{K}\left(u_k^{(n)}\right)^2 &= \sum_{n=1}^{N_\mathrm{s}}\sum_{k=1}^{K}(\hat{\alpha}_1\cos(\omega t_k - \boldsymbol{\kappa}\cdot\boldsymbol{p}_n) + \hat{\alpha}_2\sin(\omega t_k - \boldsymbol{\kappa}\cdot\boldsymbol{p}_n))^2\\ &\approx \frac{N_\mathrm{s}L}{2}\left(\hat{\alpha}_1^2 + \hat{\alpha}_2^2\right).\end{aligned}$$

A2 Vector case

In the vector case, each sensor component may experience a different amplitude scalings or phase delay on each component, as shown in (1) and (2). The amplitude scalings $\{\beta^{(\ell)}\}_{\ell=1,\ldots,L}$ and the phase delays $\{\gamma^{(\ell)}\}_{\ell=1,\ldots,L}$ are, in general, functions of the wavefield parameters $\boldsymbol{\theta}$ except for α and φ. The measurement $Y_k^{(\ell)}$, from the ℓth channel, at the time instant t_k and at position $\boldsymbol{p}_\ell$ is

$$\begin{aligned}Y_k^{(\ell)} &= u(\boldsymbol{p}_\ell, t_k) + Z_k^{(\ell)}\\ &= \alpha_0\beta^{(\ell)}\cos(\omega t_k - \boldsymbol{\kappa}\cdot\boldsymbol{p}_\ell + \varphi_0 + \gamma^{(\ell)}) + Z_k^{(\ell)}\\ &= \alpha_1\beta^{(\ell)}\cos(\omega t_k - \boldsymbol{\kappa}\cdot\boldsymbol{p}_\ell + \gamma^{(\ell)})\\ &\quad + \alpha_2\beta^{(\ell)}\sin(\omega t_k - \boldsymbol{\kappa}\cdot\boldsymbol{p}_\ell + \gamma^{(\ell)}) + Z_k^{(\ell)}.\end{aligned}$$

With a suitable parametrization of $\{\beta^{(\ell)}, \gamma^{(\ell)}\}_{\ell=1,\dots,L}$, this model is able to capture all the amplitude scaling and phase delays as in the seismic wave equations of (1) and (2).

Similar to the scalar case, it is possible to obtain analytic expressions of both ML estimates $\hat{\alpha}_1$ and $\hat{\alpha}_2$ as

$$\hat{\alpha}_1 = \frac{2}{K} \frac{\sum_{\ell=1}^{L} \sum_{k=1}^{K} Y_k^{(\ell)} \beta^{(\ell)} \cos(\omega t_k - \boldsymbol{\kappa} \cdot \boldsymbol{p}_\ell + \gamma^{(\ell)})}{\sum_{\ell=1}^{L} (\beta^{(\ell)})^2}.$$

Here and in what follows, we omit the formulae concerning $\hat{\alpha}_2$ because of their similarity with the derivations of $\hat{\alpha}_1$. Details on the estimation of $\hat{\alpha}_1$ and $\hat{\alpha}_2$ in this setting can be found in Reller *et al.* (2011).

The first and the second moments of $\hat{\alpha}_1$ are

$$\mathsf{E}\{\hat{\alpha}_1\} \approx \frac{2/K}{\sum_{\ell=1}^{L} (\beta^{(\ell)})^2} \sum_{\ell=1}^{L} \sum_{k=1}^{K} \beta^{(\ell)} \check{\beta}^{(\ell)} \left(\check{\alpha}_1 \cos(\check{\omega} t_k - \check{\boldsymbol{\kappa}} \cdot \boldsymbol{p}_\ell + \check{\gamma}^{(\ell)}) \right.$$
$$\left. + \check{\alpha}_2 \sin(\check{\omega} t_k - \check{\boldsymbol{\kappa}} \cdot \boldsymbol{p}_\ell + \check{\gamma}^{(\ell)}) \right) \cos(\omega t_k - \boldsymbol{\kappa} \cdot \boldsymbol{p}_\ell + \gamma^{(\ell)}),$$

$$\mathsf{E}\{\hat{\alpha}_1^2\} \approx \mathsf{E}\{\hat{\alpha}_1\}^2 + \frac{2}{N_s K} \sigma^2.$$

As for the scalar case, we are interested in the average shape. Therefore, we take the logarithm of the LF, take the expectation and drop multiplicative and additive constants to get

$$\mathsf{E}\{\ln(p_{\boldsymbol{Y}}(\boldsymbol{y}|\boldsymbol{\theta}))\} \propto \mathsf{E}\{\hat{\alpha}_1\}^2 + \mathsf{E}\{\hat{\alpha}_2\}^2$$
$$\approx \left| \frac{\sum_{\ell=1}^{L} \sum_{k=1}^{K} \beta^{(\ell)} \check{u}_k^{(n)} e^{-i(\omega t_k - \boldsymbol{\kappa} \cdot \boldsymbol{p}_\ell + \gamma^{(\ell)})}}{\sum_{\ell=1}^{L} (\beta^{(\ell)})^2} \right|^2.$$

We observe that, in the wave models of (1) and (2) the amplitude scalings and the phase delays are identical on the same component of different sensors. We denote with β_x, γ_x, β_y, γ_y, and β_z, γ_z the scalings and delays on the x, y and z components, respectively. Moreover, the quantity $\sum_{\ell=1}^{L} (\beta^{(\ell)})^2$ is constant with respect to $\boldsymbol{\kappa}$.

The expression can be further simplified by grouping identical scalings

$$\mathsf{E}\{\ln(p_{\boldsymbol{Y}}(\boldsymbol{y}|\boldsymbol{\theta}))\} \propto \left| \left(\beta_x \check{\beta}_x e^{i(\check{\gamma}_x - \gamma_x)} + \beta_y \check{\beta}_y e^{i(\check{\gamma}_y - \gamma_y)} + \beta_z \check{\beta}_y e^{i(\check{\gamma}_z - \gamma_z)} \right) \right.$$
$$\left. \cdot \check{\alpha} e^{-i\check{\varphi}} G(\omega - \check{\omega}) H(\boldsymbol{\kappa} - \check{\boldsymbol{\kappa}}) \right|^2.$$

APPENDIX B: MIP ARRAY LAYOUTS

In this section, we provide the layouts found using the MIP method for different number of sensors. We explain how it is possible to use the optimized geometries by considering the largest wavenumber in the actual wavefield in practical applications.

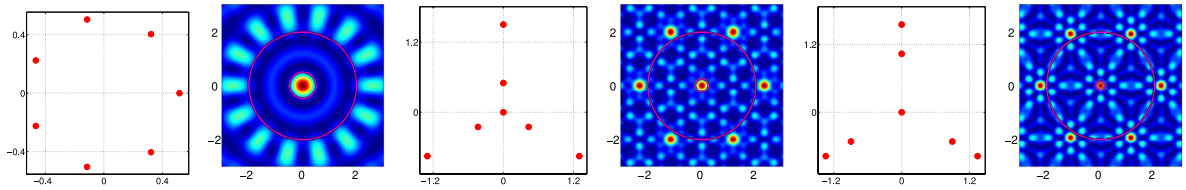
Using the scaling property of the Fourier transform it is possible to stretch or compress the array layout according to the largest wavenumber in the wavefield and, therefore, adapt the Fourier transform of the sampling pattern to the actual frequency content of the wavefield. The array layouts presented in this section are obtained using $\kappa_{\min} = 1$ and different values of $\kappa_{\min}$. Let $\kappa^*_{\max}$ be the largest wavenumber in the wavefield. Let $h^*(x, y)$ and $H^*(\kappa_x, \kappa_y)$ denote the desired sampling pattern and its Fourier Transform, respectively. They are related to the $h(x, y)$ and $H(\kappa_x, \kappa_y)$ provided in Figs B1 and B2 as

$$h^*(x, y) = h\left(x\kappa^*_{\max}, y\kappa^*_{\max}\right),$$

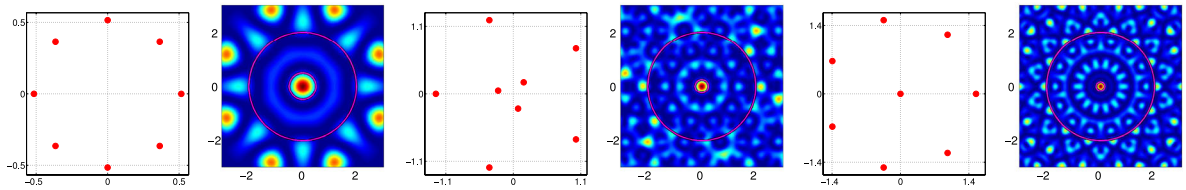
$$H^*(\kappa_x, \kappa_y) = \frac{1}{\kappa^*_{\max}} H^*\left(\frac{\kappa_x}{\kappa^*_{\max}}, \frac{\kappa_y}{\kappa^*_{\max}}\right).$$

The effective $\kappa^*_{\min}$ is also changed as $\kappa^*_{\min} = \kappa_{\min}\kappa^*_{\max}$. Choice of which $\kappa_{\min}$ to choose should be related in particular to the smallest wavenumber of interest in the analysis. In fact, we observe how a small $\kappa_{\min}$ leads to arrays with larger aperture.

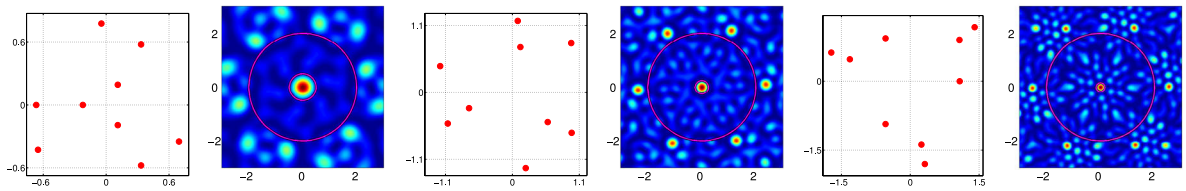
In practice, stretching the spatial sampling pattern by $1/\kappa^*_{\max}$ allows us to obtain a $H^*(\kappa_x, \kappa_y)$ with local maxima minimized in the annulus $\kappa_{\min}\kappa^*_{\max} \leq \|\boldsymbol{\kappa}\|_2 \leq 2\kappa^*_{\max}$.



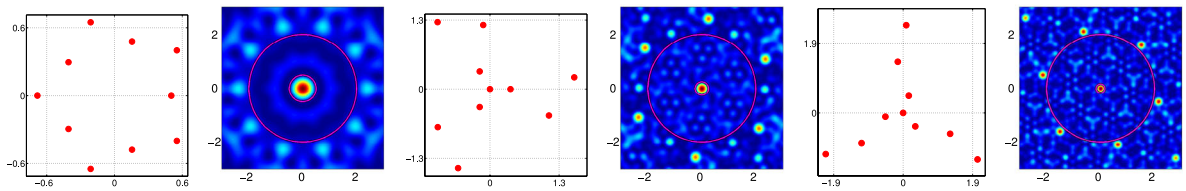
(a) $N_s = 7$, $\kappa_{max}/\kappa_{min} = 2$. (b) $N_s = 7$, $\kappa_{max}/\kappa_{min} = 4$. (c) $N_s = 7$, $\kappa_{max}/\kappa_{min} = 6$.



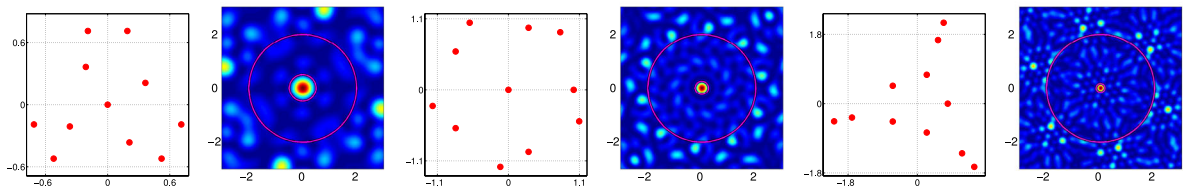
(d) $N_s = 8$, $\kappa_{max}/\kappa_{min} = 2$. (e) $N_s = 8$, $\kappa_{max}/\kappa_{min} = 4$. (f) $N_s = 8$, $\kappa_{max}/\kappa_{min} = 6$.



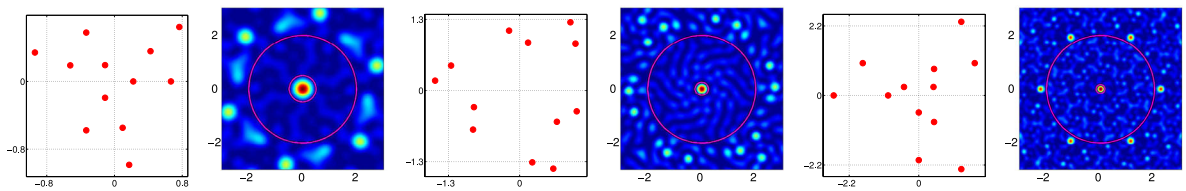
(g) $N_s = 9$, $\kappa_{max}/\kappa_{min} = 2$. (h) $N_s = 9$, $\kappa_{max}/\kappa_{min} = 4$. (i) $N_s = 9$, $\kappa_{max}/\kappa_{min} = 6$.



(j) $N_s = 10$, $\kappa_{max}/\kappa_{min} = 2$. (k) $N_s = 10$, $\kappa_{max}/\kappa_{min} = 4$. (l) $N_s = 10$, $\kappa_{max}/\kappa_{min} = 6$.

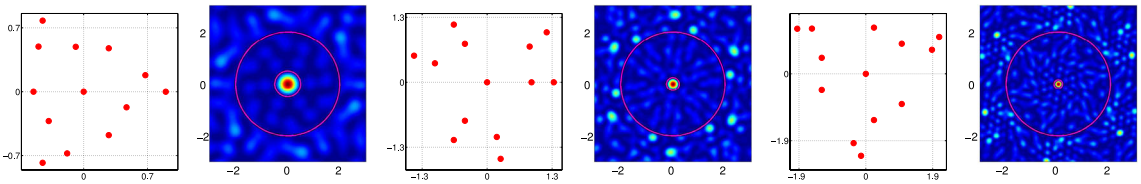


(m) $N_s = 11$, $\kappa_{max}/\kappa_{min} = 2$. (n) $N_s = 11$, $\kappa_{max}/\kappa_{min} = 4$. (o) $N_s = 11$, $\kappa_{max}/\kappa_{min} = 6$.

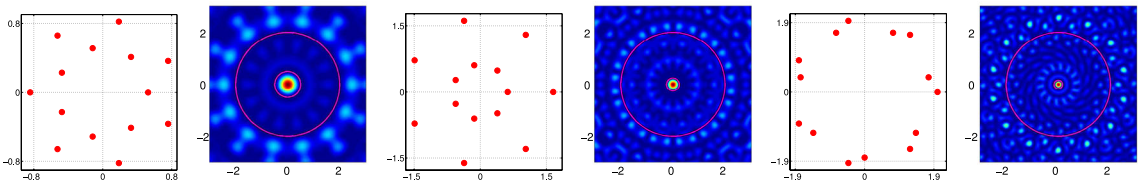


(p) $N_s = 12$, $\kappa_{max}/\kappa_{min} = 2$. (q) $N_s = 12$, $\kappa_{max}/\kappa_{min} = 4$. (r) $N_s = 12$, $\kappa_{max}/\kappa_{min} = 6$.

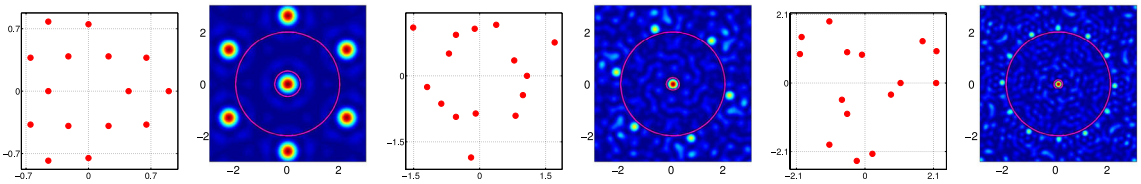
Figure B1. Array layouts found with the MIP method for $N_s = 7, \ldots, 12$.



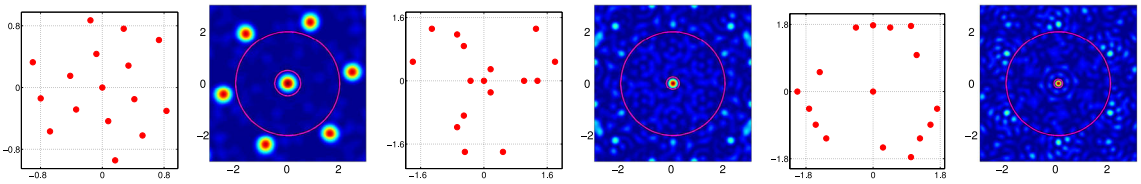
(a) $N_s = 13$, $\kappa_{max}/\kappa_{min} = 2$. (b) $N_s = 13$, $\kappa_{max}/\kappa_{min} = 4$. (c) $N_s = 13$, $\kappa_{max}/\kappa_{min} = 6$.



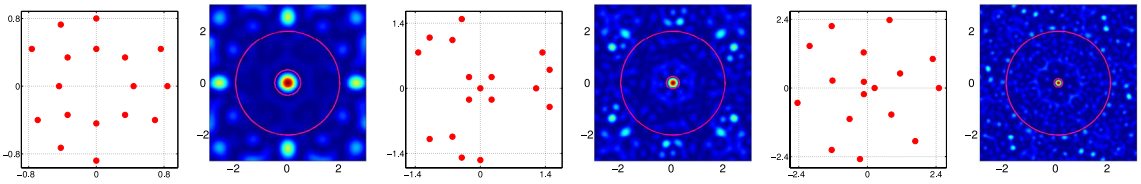
(d) $N_s = 14$, $\kappa_{max}/\kappa_{min} = 2$. (e) $N_s = 14$, $\kappa_{max}/\kappa_{min} = 4$. (f) $N_s = 14$, $\kappa_{max}/\kappa_{min} = 6$.



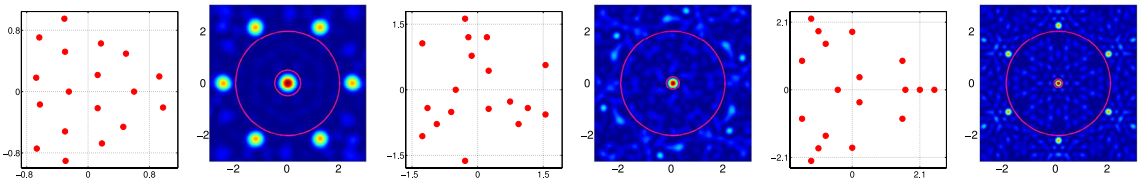
(g) $N_s = 15$, $\kappa_{max}/\kappa_{min} = 2$. (h) $N_s = 15$, $\kappa_{max}/\kappa_{min} = 4$. (i) $N_s = 15$, $\kappa_{max}/\kappa_{min} = 6$.



(j) $N_s = 16$, $\kappa_{max}/\kappa_{min} = 2$. (k) $N_s = 16$, $\kappa_{max}/\kappa_{min} = 4$. (l) $N_s = 16$, $\kappa_{max}/\kappa_{min} = 6$.



(m) $N_s = 17$, $\kappa_{max}/\kappa_{min} = 2$. (n) $N_s = 17$, $\kappa_{max}/\kappa_{min} = 4$. (o) $N_s = 17$, $\kappa_{max}/\kappa_{min} = 6$.



(p) $N_s = 18$, $\kappa_{max}/\kappa_{min} = 2$. (q) $N_s = 18$, $\kappa_{max}/\kappa_{min} = 4$. (r) $N_s = 18$, $\kappa_{max}/\kappa_{min} = 6$.

Figure B2. Array layouts found with the MIP method for $N_s = 13, \ldots, 18$.