

A simple algorithm for optimal design in distributed fibre-optic sensing

Andreas Fichtner¹ and Coen Hofstede²

¹*Department of Earth Sciences, ETH Zurich, 8092, Zurich, Switzerland. E-mail: andreas.fichtner@erdw.ethz.ch*

²*Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research, 27570 Bremerhaven, Germany*

Accepted 2022 November 21. Received 2022 November 15; in original form 2022 October 10

SUMMARY

We present a basic algorithm for optimal experimental design in distributed fibre-optic sensing. It is based on the fast random generation of fibre-optic cable layouts that can be tested for their cost-benefit ratio. The algorithm accounts for the maximum available cable length, lets the cable pass through pre-defined points of interest, avoids obstacles that the cable must not traverse, permits the adaptation of geometric complexity of different cable segments and allows for the incorporation of topography. Furthermore, the algorithm can be combined with arbitrary measures of the cost-benefit ratio, and its simplicity enables easy adaptations to the needs of specific applications. In addition to a description of the basic concept, we provide examples that illustrate the circumnavigation of obstacles, the steering of geometric complexity and the cable layout optimization in the presence of topographic variations.

Key words: Inverse theory; Tomography; Seismic instruments; Theoretical seismology; Wave propagation.

1 INTRODUCTION

Optimal experimental design (OED) is a family of methods used to find the best possible balance between cost and benefit in the presence of limited resources. In the context of geophysics, OED frequently involves the positioning of a finite number of receivers or sources such that some measure of the obtainable information is close to its maximum. Practical OED methods have been developed, for instance, to design seismic networks (e.g. Kijko 1977; Hardt & Scherbaum 1994) and for various forms of seismic imaging (e.g. Curtis 1999a, b; Curtis & Maurer 2000; Maurer *et al.* 2009, 2010, 2017) and geoelectrical sounding (e.g. Stummer *et al.* 2004). Adaptations of geophysical OED have appeared recently in medical ultrasound tomography (Korta Martiartu *et al.* 2019).

The emergence of distributed acoustic sensing (DAS) has profound impact on seismic experiments, for example, in subsurface reservoir monitoring (e.g. Daley *et al.* 2013, 2016; Mateeva *et al.* 2014), volcanology (e.g. Klaasen *et al.* 2021, 2022; Fichtner *et al.* 2022b; Jousset *et al.* 2022), glaciology (e.g., Walter *et al.* 2020; Hudson *et al.* 2021), earthquake-induced ground motion analysis (e.g. Spica *et al.* 2020; Yang *et al.* 2021) and seismic event detection and monitoring (e.g. Lindsey *et al.* 2017; Martin *et al.* 2017; Thrastarson *et al.* 2021). The particularities of DAS add complications and desiderata to OED that are of lesser or no concern when individual receivers can be placed independently. These include the requirements to (1) account for the maximum available cable length, (2) traverse specific points of interest, such as defined start and endpoints of the cable, (3) avoid obstacles and areas that the

cable is not allowed to cross, (4) account for topography and (5) steer the geometric complexity of the cable layout as a function of anticipated data analyses or practical considerations of the cable deployment. Furthermore, an OED method for DAS should be simplistic in order to enable fast and easy adaptations to specific use cases. The introduction and illustration of such a basic algorithm is the primary objective of this work.

The development of an OED method for DAS was motivated by the actual need to design an active-source experiment around the EastGRIP ice core drilling project on the Greenland Ice Sheet, a site characterized by no-go areas and impassable infrastructure. Alluding to the Greenlandic town where the following concepts were developed, we refer to the method as the Kangerlussuaq algorithm.

2 THE BASIC KANGERLUSSUAQ ALGORITHM

The Kangerlussuaq algorithm rests on a simple concept for the random generation of cable layouts that fulfill the above-mentioned desiderata and can be tested for their cost-benefit ratio, however defined. The basic version of the algorithm may be modified and adapted depending on the needs of a specific application.

2.1 Random generation of admissible cable geometries

At the initial stage, the interpolation points that the cable is required to pass are connected by straight lines of lengths ℓ_1, ℓ_2, \dots , as

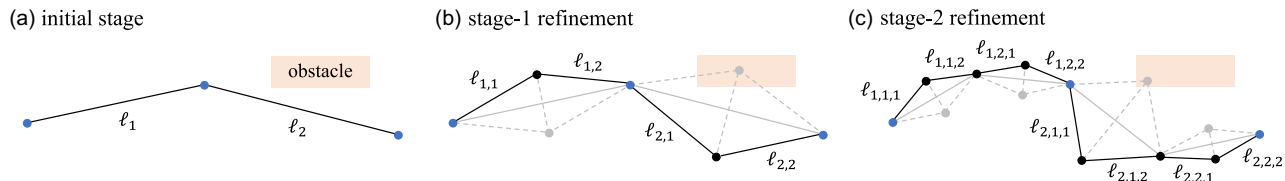


Figure 1. Illustration of successive refinements in cable geometry, starting in panel (a) with two straight cable segments that are required to pass through the blue interpolation points. The stage-1 refinement in panel (b) is based on the random drawing of new segment lengths $\ell_{i,j}$ from probability densities ρ_i and the assignment of new interpolation points in accord with the requirement to avoid any obstacles. The identical procedure is repeated in (c) for a stage-2 refinement.

illustrated in Fig. 1(a). The lengths of the segments must satisfy $\sum_i \ell_i < L$, where L is the maximum available cable length. Subsequently, the first segment of length ℓ_1 is refined randomly by drawing the lengths $\ell_{1,1} = \ell_{1,2}$ of two new segments from a probability distribution ρ_1 . These new segments subdivide ℓ_1 , as shown in Fig. 1(b). Two ways of subdividing ℓ_1 are possible (solid black and dashed grey), and one of them is chosen randomly (solid black). Similarly, ℓ_2 is subdivided by randomly drawing $\ell_{2,1} = \ell_{2,2}$ from a distribution ρ_2 . However, instead of choosing randomly between the two refinement options, only the one that does not cross the red shaded obstacle can be selected.

To address the optimal design problem, a large number N_1 of random stage-1 refinements, as described above, is generated and inserted into the objective function χ that evaluates the problem-specific cost-benefit ratio of a cable geometry. Typically, N_1 will be proportional to the number of initial segments, and a useful value will most likely be the result of some trial and error. The geometry with the lowest cost-benefit ratio is then selected for stage-2 refinements, illustrated in Fig. 1(c).

As in stage 1, the segments $\ell_{i,j}$ are refined by drawing the lengths of new segments $\ell_{i,j,k}$ from probability distributions $\rho_{i,j}$, where i ranges from 1 to the number of initial segments, and j and k are either 1 or 2. The choice of one or the other refinement option is made in accord with the requirement to not cross any obstacle. The algorithm proceeds by randomly drawing a large number N_2 of stage-2 refinements, finally selecting the one that minimizes χ . Depending on the complexity of the fibre geometry that is desirable or tolerable, the same procedure can be repeated with stage-3, stage-4, etc., refinements.

2.2 Probability densities for new fibre segments

The properties of the Kangerlussuaq algorithm are largely controlled by the choice of the probability densities $\rho_{i,\dots}$ that determine the lengths of new cable segments $\ell_{i,j,\dots}$. Creativity in choosing $\rho_{i,\dots}$ is only limited by few constraints on the support of these distributions. Taking the example of stage-1 refinements, the total length of a new segment $\ell_{i,j}$ must be at least half of ℓ_i . Denoting by ℓ_i^{\min} and ℓ_i^{\max} the lower and upper bounds within which ρ_i is non-zero, this translates to $\ell_i^{\min} = \ell_i/2$. Analogously, the maximum cumulative length of new segments must be less than the maximum available cable length. This implies $\sum_i \ell_i^{\max} \leq L/2$. At the same time, $\ell_i^{\max} > \ell_i^{\min}$ is required to ensure a non-empty support of ρ_i . One of infinitely many possibilities for the choice of ℓ_i^{\max} is $\ell_i^{\max} = L\ell_i/2\sum_i \ell_i$. These upper bounds produce new segments with lengths $\ell_{i,j}$ that are statistically proportional to the length of the original segment ℓ_i .

2.3 Circumnavigating obstacles

Obstacles that block the direct path between two interpolation points generally lead to the requirement of longer cables. If the maximum available cable length is insufficient, the circumnavigation of the obstacle may not be possible. The Kangerlussuaq algorithm attempts to find new interpolation points that are permitted in the sense that none of the new cable segments traverses the obstacle. In case a randomly drawn $\ell_{i,\dots}$ can only place a new interpolation point inside an obstacle, the random trial is repeated at most n times. Failure to find a new interpolation point that avoids obstacles after n trials leads to the termination of the algorithm with the message that the maximum available cable length is insufficient given the set of required interpolation points and obstacles. Empirically, $n = 20$ is a useful choice.

3 EXAMPLES

We illustrate the basic Kangerlussuaq algorithm with a realistic example from the EastGRIP camp site, summarized in Fig. 2, where a cable of at most 7 km length and with 10 m receiver (DAS channel) spacing records waves emitted by 38 sources. Various obstacles are shown with grey shading. Three points through which the cable must pass are plotted in blue. For simplicity, we define $\rho_{i,\dots}$ as the uniform distribution in the interval $[\ell_{i,\dots}^{\min}, \ell_{i,\dots}^{\max}]$.

The goal of the active-source experiment was to constrain azimuthal anisotropy and velocity changes with depth. Since the ice stream can be considered a laterally homogeneous medium over length scales of few kilometres, this could be achieved through a sequence of 2-D tomographies, each with source–receiver pairs within a narrow azimuthal bin. To achieve comparable resolution, the offset coverage for each azimuth bin should be similar. However, azimuthal and offset coverage are antagonistic and mutually dependent. A straight cable offers maximum offset coverage but implies zero azimuthal coverage. A circular cable has the opposite effect. In order to find a suitable compromise, we note that each source–receiver pair is characterized by its offset d and azimuth ϕ . Their ensemble defines a discrete offset–azimuth distribution $\rho(d, \phi)$ for a given cable geometry. Aiming for a geometry that produces a broad and evenly distributed range of distances and azimuths, we define the objective function χ as the L_1 distance between $\rho(d, \phi)$ and the normalized uniform distribution $\rho_c = \text{const}$. As discussed in Section 5 other objective functions can easily be used within the Kangerlussuaq algorithm, depending on the requirements of specific applications.

Optimal geometries for different refinement stages are shown in Fig. 2(a) in different styles. By design, they avoid all obstacles without exceeding the allowable cable length, while homogenizing the offset–azimuth distribution as much as possible. Increasing refinement stages improve the distribution at the expense of a more

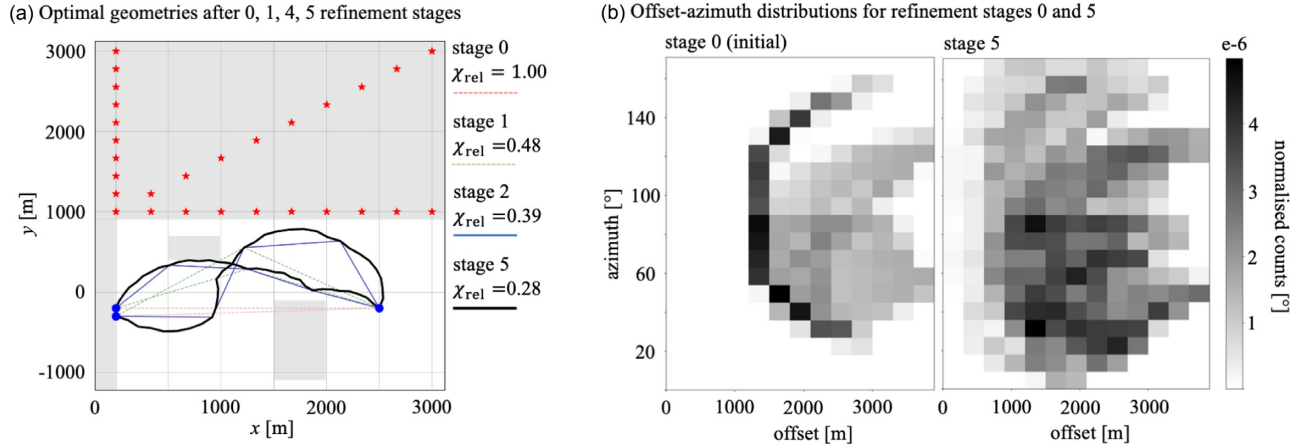


Figure 2. Optimal cable geometries as a function of refinement stage. (a) Experimental setup with sources marked by red stars and obstacles by grey shaded areas. Points through which the fibre must pass are shown in blue. The total allowable fibre length is 7 km. Optimal cable layouts for different refinement stages, including the initial stage, are shown in different colours and line styles. Relative objective function values χ_{rel} for stages 0, 1, 2 and 5 are indicated to the right. After five refinement stages, the optimal layout, shown as solid black curve, produces an objective function value of $\chi_{\text{rel}} = 0.28$, relative to the initial value, which is set to 1. (b) Offset-azimuth distribution of all surface ray paths between sources and DAS channels for the initial geometry and the optimal geometry after five refinement stages. By design, the final geometry leads to a more homogeneous distribution of offset-azimuth pairs.

complex layout, but after five stages, additional refinements do not lead to significant further improvements. This slow-down of convergence is typical for optimization algorithms that approach the optimum, though it may not generally occur after five stages. Using 5000 random realizations per stage, this example took less than 2 min compute time on an Intel® Core™ i9 processor. A comparison of the offset-azimuth distributions in the initial and the fifth stage is shown in Fig. 2(b). The initial stage, where two straight segments connect the three required points, produces a heterogeneous offset-azimuth distribution with few non-empty bins. In contrast, nearly all offset-azimuth bins are non-empty for the five-stage optimal solution from Fig. 2(a).

4 VARIATIONS OF THE THEME

The simplicity of the Kangerlussuaq algorithm permits easy adaptations to the needs of specific use cases. Below, we describe two of these, the steering of geometric complexity of different cable segments and the incorporation of topography.

4.1 Steering complexity

Controlling the geometric complexity of cable segments can be advantageous in various applications. For example, specific segments may need to be nearly linear in order to facilitate analysis techniques such as f - k filtering. Other segments may be given more flexibility in order to circumnavigate obstacles. Cable complexity can be controlled through the design of the probability densities $\rho_{i,\dots}$. To have more flexibility than in the examples of Section 3, we may generalize the uniform distribution in $[\ell_{i,\dots}^{\min}, \ell_{i,\dots}^{\max}]$ to the uniform distribution in $[\ell_{i,\dots}^{\min}, \ell_{i,\dots}^{\min} + w_{i,\dots} \Delta \ell_{i,\dots}]$, with $\Delta \ell_{i,\dots} = \ell_{i,\dots}^{\max} - \ell_{i,\dots}^{\min}$, and some weights $w_{i,\dots}$ between 0 and 1. Assigning numbers close to 0 to certain weights, forces the corresponding segments to be closer to a straight line, thereby giving more freedom to other segments.

Two examples that illustrate the effect of the weights $w_{i,\dots}$ are shown in Fig. 3. The scenario is similar to the one in Fig. 2, but includes a fourth interpolation point that the cable is forced to traverse. Assigning the weights $w_{1,\dots} = 0$ to the first segment results in minimal complexity, that is, a straight line. The weights $w_{2,\dots}$

$= 0.3$ for the second segment permits more flexibility but still results in a nearly straight segment after five refinement stages. Finally, $w_{3,\dots} = 1$ grants maximum flexibility to the third segment. For comparison, letting all weights equal 1, gives full flexibility to all segments, thereby leading to a lower objective χ_{rel} at the cost of higher geometric complexity, which may be challenging during cable deployment.

4.2 Incorporating topography

The basic version of the Kangerlussuaq algorithm can be extended to cases with surface topography. While an exact version of such an algorithm would require the calculation of distances on curved surfaces, a pragmatic approximation requires only minor modifications. For this, we continue to connect points with randomly drawn distances by straight lines. In the early refinement stages, cable segments may therefore be located above or below the surface. However, as the algorithm proceeds, these misplacement artefacts become increasingly unimportant. This is illustrated in Fig. 4(a). An adaptation of the example from Fig. 2 to a scenario with topography is displayed in Fig. 4(b). Again, by construction, the cable avoids the obstacles defined in the x - y -plane without exceeding the maximum allowable length of 7 km, while homogenizing the offset-azimuth distribution.

5 DISCUSSION AND CONCLUSIONS

We presented an algorithm for OED in distributed fibre-optic sensing based on successive refinements of the cable geometry with increasing complexity. The algorithm meets basic desiderata of typical DAS experiments and is deliberately simplistic in order to enable easy application-specific adaptations.

In addition to adaptations that handle complexity and topography, more flexibility can be added by drawing start or endpoints randomly and looping the algorithm over these random locations. Increased flexibility in the construction of trial geometries may also be desirable. For example, allowing refined segments to have different length ($\ell_{1,1} \neq \ell_{1,2}$, etc.) would facilitate the avoidance of obstacles. However, this enhancement of the algorithm would

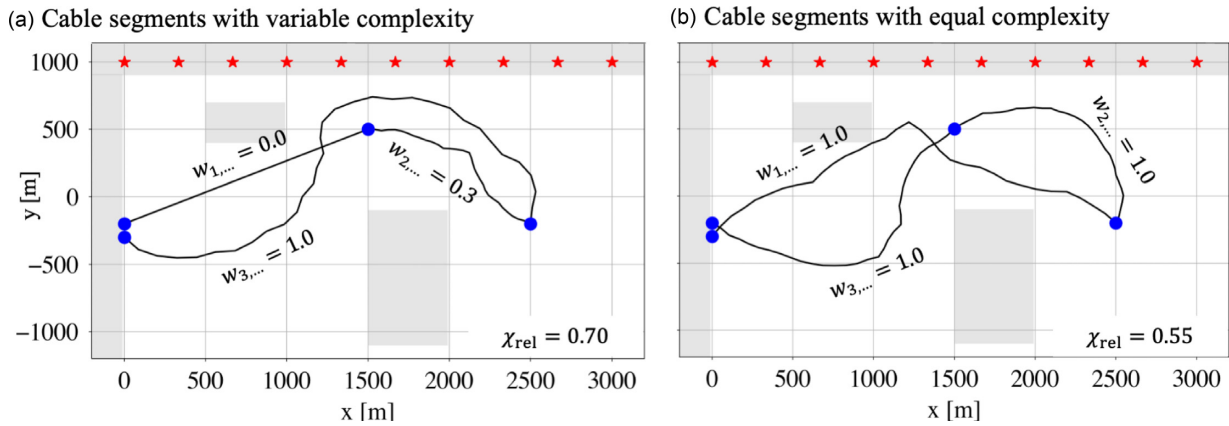
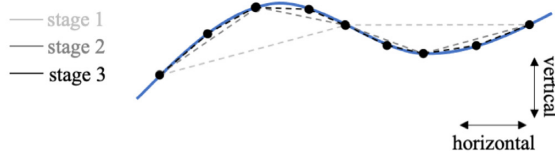


Figure 3. Steering cable complexity through the design of probability densities $\rho_{i,\dots}$. (a) The weights $w_{i,\dots}$ control the width of the support of $\rho_{i,\dots}$. Small values lead to more linear segments, whereas larger values provide more geometric flexibility. The scenario is the same as in Fig. 2, but with the addition of a fourth interpolation point. The map is cut above $y = 1100$ m. (b) Setting all weights to 1 grants maximum flexibility, which leads to a lower relative objective χ_{rel} than in the more constrained case with smaller weights.

(a) Schematic construction with topography



(b) Optimal cable layout in the presence of topography

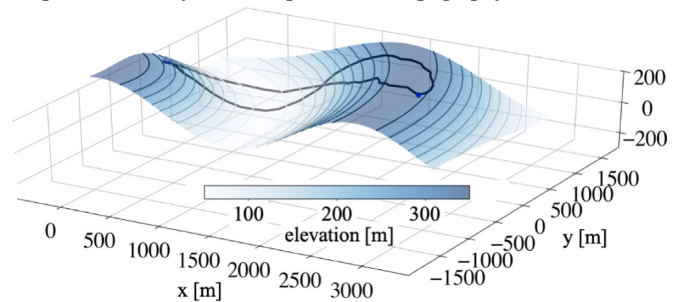


Figure 4. Incorporating topography into the Kangerlussuaq algorithm. (a) A cable that follows topography is approximated by straight lines. The approximation improves with increasing refinement stages. (b) Adaptation of the examples in Fig. 2 to a scenario with topography.

require an additional set of probability densities to control these lengths.

Variations of the Kangerlussuaq may be used for OED with fibre-optic sensing technologies that are not based on DAS. This includes, for instance, transmission fibre-optics systems that are characterized by a position-dependent measurement sensitivity that is controlled by local cable curvature (Marra *et al.* 2018; Mecozzi *et al.* 2021; Bowden *et al.* 2022; Bogris *et al.* 2022; Fichtner *et al.* 2022a). High-curvature segments along the cable thereby mimic a distributed sensing system that should be optimally designed in order to maximize resolution of Earth structure or earthquake source parameters.

The choice of a simple resolution proxy for the objective function χ in Sections 3 and 4 has the benefits of being independent of the *a priori* unknown medium properties and of providing an intuitive illustration that interested readers may reproduce within minutes. Alternatives may account for the directional sensitivity of DAS or gauge length effect. Conceptually similar resolution proxies based on various forms of ray coverage are known in traveltime tomography or array analysis (e.g. Rost & Thomas 2002; Schäfer *et al.* 2011). In applications where the inverse problem is sufficiently linear, objective functions may alternatively be defined in terms of the posterior model resolution or covariance matrices (e.g. Curtis 1999a; Stummer *et al.* 2004). Adaptations of the algorithm that maximize the value of information (e.g. W.J.Trainor-Guitton *et al.* 2014) may equally be imagined.

Actual field deployments will mostly be less accurate than suggested by the algorithm, for instance, due to limited navigation accuracy. Small deviations from the optimal layout will usually result in only small deviations in coverage, implying that the result would still be useful.

Along similar lines, we finally note that the Kangerlussuaq algorithm does not necessarily find the global optimal cable geometry, a property it shares with any other practical optimization algorithm. It finds, however, cable geometries that are useful and potentially better than an intuitive guess, especially in the presence of obstacles. Useful Experimental Design would certainly be the more reasonable technical term.

DATA STATEMENT

No data have been used for this research.

ACKNOWLEDGMENTS

The authors gratefully acknowledge Marie Kirk and Iben Koldtoft, logistics coordinators of the EastGRIP project, for putting us into the same room at Kangerlussuaq International Science Support (KISS), thereby forcing us to talk about science continuously for 3 d in a row (including all breakfasts and dinners)! Constructive comments by Sara Klaasen, Patrick Paitz, Daniel Bowden, Pascal Edme, Cédric

Schmelzbach, Robert Mellors and an anonymous reviewer helped us to improve the manuscript. The Kangerlussuaq algorithm is freely available in the form of Python code at <https://github.com/afichtner/Kangerlussuaq>.

REFERENCES

- Bogris, A. *et al.*, 2022. Sensitive seismic sensors based on microwave frequency fiber interferometry in commercially deployed cables, *Sci. Rep.*, **12**, doi:10.1038/s41598-022-18130-x.
- Bowden, D.C. *et al.*, 2022. Linking distributed and integrated fiber-optic sensing, *Geophys. Res. Lett.*, **49**, doi:10.1029/2022GL098727.
- Curtis, A., 1999a. Optimal experimental design: cross-borehole tomographic examples, *Geophys. J. Int.*, **136**, 637–650.
- Curtis, A., 1999b. Optimal design of focused experiments and surveys, *Geophys. J. Int.*, **139**, 205–215.
- Curtis, A. & Maurer, H., 2000. Optimizing the design of geophysical experiments: is it worthwhile?, *Leading Edge*, **19**, 1058–1062.
- Daley, T.M., Pevzner, R., Shulakova, V., Kashikar, S., Miller, D.E., Goetz, J. & S. Lueth, J.H., 2013. Field testing of fiber-optic distributed acoustic sensing (DAS) for subsurface seismic monitoring, *Leading Edge*, **June 2013**, 936–942.
- Daley, T.M., Miller, D.E., Dodds, K., Cook, P. & Freifeld, B.M., 2016. Field testing of modular borehole monitoring with simultaneous distributed acoustic sensing and geophone vertical seismic profiles at Citronelle, Alabama, *Geophys. Prosp.*, **64**, 1318–1334.
- Fichtner, A. *et al.*, 2022a. Theory of phase transmission fibre-optic sensing, *Geophys. J. Int.*, in press, doi:10.1093/gji/ggac237.
- Fichtner, A., Klaasen, S., Thrastarson, S., Çubuk-Sabancı, Y., Paitz, P. & Jónsdóttir, K., 2022b. Fiber-optic observation of volcanic tremor through floating ice-sheet resonance, *Seismic Record*, **2**, 148–155.
- Hardt, M. & Scherbaum, F., 1994. The design of optimum networks for aftershock recordings, *Geophys. J. Int.*, **117**, 716–726.
- Hudson, T.S. *et al.*, 2021. Distributed Acoustic Sensing (DAS) for natural microseismicity studies: a case study from Antarctica, *J. geophys. Res.*, **126**, doi:10.1029/2020JB021493.
- Jousset, P. *et al.*, 2022. Fibre optic distributed acoustic sensing of volcanic events, *Nat. Comm.*, **13**, doi:10.1038/s41467-022-29184-w.
- Kijko, A., 1977. An algorithm for the optimum distribution of a regional seismic network—I, *Pure appl. Geophys.*, **115**, 999–1009.
- Klaasen, S., Paitz, P., Lindner, N., Dettmer, J. & Fichtner, A., 2021. Distributed Acoustic Sensing in volcano-glacial environments—Mount Meager, British Columbia, *J. geophys. Res.*, **159**, doi:10.1029/2021JB022358.
- Klaasen, S., Thrastarson, S., Fichtner, A., Çubuk-Sabancı, Y. & Jónsdóttir, K., 2022. Sensing Iceland’s most active volcano with a “buried hair”, *EOS*, **103**, doi:10.1029/2022EO220007.
- Korta Martiartu, N., Boehm, C., Hapla, V., Maurer, H., Balic, I.J. & Fichtner, A., 2019. Optimal experimental design for joint reflection-transmission ultrasound breast imaging: From ray-to wave-based methods, *J. acoust. Soc. Am.*, **146**, 1252–1264.
- Lindsey, N.J., *et al.* 2017. Fiber-optic network observations of earthquake wavefields, *Geophys. Res. Lett.*, **44**, 11792–11799.
- Marra, G. *et al.*, 2018. Ultrastable laser interferometry for earthquake detection with terrestrial and submarine cables, *Science*, **361**, 486–490.
- Martin, E.R., *et al.*, 2017. Seismic monitoring leveraging existing telecom infrastructure at the SDASA: active, passive, and ambient-noise analysis, *Leading Edge*, **36**, 1025–1031.
- Mateeva, A. *et al.*, 2014. Distributed acoustic sensing for reservoir monitoring with vertical seismic profiling, *Geophys. Prosp.*, **62**, 679–692.
- Maurer, H., Greenhalgh, S. & Latzel, S., 2009. Frequency and spatial sampling strategies for crosshole seismic waveform spectral inversion experiments, *Geophysics*, **74**, WCC11–WCC21.
- Maurer, H., Curtis, A. & Boerner, D.E., 2010. Recent advances in optimized geophysical survey design, *Geophysics*, **75**, A177–A195.
- Maurer, H., *et al.*, 2017. Optimized experimental design in the context of seismic full waveform inversion and seismic waveform imaging, *Adv. Geophys.*, **58**, 1–45.
- Mecozzi, A., Cantono, M., Castellanos, J.C., Kamalov, V., Muller, R. & Zhan, Z., 2021. Polarization sensing using submarine optical cables, *Optica*, **8**, doi:10.1364/OPTICA.424307.
- Rost, S. & Thomas, C., 2002. Array seismology: methods and applications, *Rev. Geophys.*, **40**, doi:10.1029/2000RG000100.
- Schäfer, J., Boschi, L. & Kissling, E., 2011. Adaptively parametrized surface wave tomography: methodology and a new model of the European upper mantle, *Geophys. J. Int.*, **186**, 1431–1453.
- Spica, Z.J., Perton, M., Martin, E.R., Beroza, B.C. & Biondi, B., 2020. Urban seismic site characterization by fiber-optic seismology, *J. geophys. Res.*, **125**, doi:10.1029/2019JB018656.
- Stummer, P., Maurer, H. & Green, A.G., 2004. Experimental design: Electrical resistivity data sets that provide optimum subsurface information, *Geophysics*, **69**, 120–139.
- Thrastarson, S., Torfason, R., Klaasen, S., Paitz, P., Çubuk-Sabancı, Y., Jónsdóttir, K. & Fichtner, A., 2021. Detecting seismic events with computer vision: applications for fiber-optic sensing, *Earth Space Sci. Open Archive*, doi:10.1002/essoar.10509693.1.
- Trainor-Guitton, W.J., Ramirez, G.M.H. A., Roberts, J., Juliusson, E., Key, K. & Mellors, R., 2014. The value of spatial information for determining well placement: a geothermal example, *Geophysics*, **79**, W27–W41.
- Walter, F., Gräff, D., Lindner, F., Paitz, P., Köpfl, M., Chmiel, M. & Fichtner, A., 2020. Distributed Acoustic Sensing of microseismic sources and wave propagation in glaciated terrain, *Nat. Comm.*, **11**, doi:10.1038/s41467-020-15824.
- Yang, Y., Atterholt, J.W., Shen, Z., Muir, J.B., Williams, E.F. & Zhan, Z., 2021. Sub-kilometer correlation between near-surface structure and ground motion measured with distributed acoustic sensing, *Geophys. Res. Lett.*, **49**(1), doi:10.1029/2021GL096503.