Imperial College London
Department of Electrical and Electronic Engineering

Sensing physical fields: Inverse problems for the diffusion equation and beyond

John Murray-Bruce

December 2016

Supervised by Professor Pier Luigi Dragotti

Submitted in part fulfilment of the requirements for the degree of
Doctor of Philosophy in Electrical and Electronic Engineering of Imperial College London
and the Diploma of Imperial College London

Declaration of Originality

I herewith certify that all material in this dissertation which is not my own work has been properly acknowledged.

John Murray-Bruce

Copyright Declaration

The copyright of this thesis rests with the author and is made available under a Creative Commons Attribution Non-Commercial No Derivatives licence. Researchers are free to copy, distribute or transmit the thesis on the condition that they attribute it, that they do not use it for commercial purposes and that they do not alter, transform or build upon it. For any reuse or redistribution, researchers must make clear to others the licence terms of this work.

In loving memory of my dear father,

Mr Edward Murray-Bruce

You will forever remain in my heart.

Acknowledgements

I am incredibly fortunate to have been a part of the Communications and Signal Processing group at Imperial college. And even more so, to have had Professor Pier Luigi Dragotti as my doctoral thesis advisor. Without whom none of this, in fact not even a single word in this thesis would have been possible. Your patience, vision and drive is inspiring. I'm sincerely grateful for the trust and belief you have in me and for being such a great mentor. It still puzzles me, how, despite your busy schedule you always found time to hold our weekly meetings. Thank you for making each and every one of them fun and immensely productive.

I would also like to extend my sincere gratitude to the examiners Dr. Wei Dai from Imperial College London and Professor Mounir Ghogho of Leeds University for their valuable feedback and suggestions towards the improvement of this thesis.

A special thanks to everyone that has, in some way, made my eight years at Imperial college such a memorable one. Titi and Ibiene for your friendship through the years. In the undergrad years: Duli, Faethon, Rapanos, Sengendo, Sid, Susana and of course my partners in crime, Taz and Damien, thanks for making that era—inside and outside the confines of Imperial college—so much fun. And in the postgrad years: Adam, Akin, Ali, Christine, Cooper, David, Enrico, Georgia, Hamza, Leonidas, Jelena, Lisa, Shanxiang, Segun, Thanos, Valentine, Ziggy. Many thanks also to Roberto, for your interest in my progress and making every effort to stay in touch; Steph, for listening to all my rants and being the best *desk neighbour* ever; Jon, for clearing up some of the FRI concepts for me, for ensuring that all the conferences we attended were done properly and above all for being such a *don*; Hamdi for our many discussions, both technical and non-technical, depressing and uplifting, silly and serious, and although you get the award for being pretty much the worst gym buddy ever, you made our late nights at the office fun and bearable. Basheer thanks for the encouragement in the final stages of this thesis, stay cool buddy! The ICSM basketball crew, Haz, Ranj, Mo, Kaveesh, etc... – I will miss you all. Finally, to the guys with whom I have shared every step of this Imperial experience, Sithan and Richard, thank you for your support through the years guys.

My fellow *rave ponies*, Soroush and Akil. Every so often when I begin to feel like you guys have been the best pair of friends one could ever ask for, you do things to remind me otherwise. Whether it be your great thesis writing tips like, 'increase the font size and you'll finish quicker,' offering non-stop to help me "write" it, reminding me of my apparent inability to grow convincing facial hair, being generally underwhelmed by anything I achieve, or having a phobia of celebrating my birthday. I can only hope you do

these things out of brotherly love. You guys are truly unique and have been there for me in every possible way imaginable. Thank you for all your encouraging words, for aborting business trips, holidays and other plans to offer your support during the tough times, as well as, all the breakfasts, lunches, dinners, drinks, relays, FIFA sessions, debates, discussions, arguments, shisha sessions, Christmas', NYEs, etc... we have shared over the years. Well, gentlemen! You will be happy to know that the long-standing question of which one of us three has the best dress sense has now been settled conclusively in this thesis. In fact, thanks to the results of Chapters 3 and 4, it can only be myself. Moreover in Chapter 6, you'll also find a corollary of this result which says that, I also taught you guys to dress.

A million thanks go to the Farah family, for all your kind and encouraging words, for making me feel like one of yours, for the swimming lessons in the sea, for the countless memories and moments we have shared from London to Cyprus (I am looking forward to making many more). Jude special thanks for taking such an interest in my PhD work and for being such an amazing bro. I am also grateful to *Papitchka* and *Mamitchka* for raising such a beautiful soul. Jany has been an absolute blessing to me.

To you my dearest, Janine, you have made every little aspect of my life infinitely beautiful ever since we met. Without you, this PhD ride would have been near impossible. For you believed in me even when I doubted myself and prayed for me during the tough times. Thank you for being my rock! أنا لحبيبي و حبيبي إلي

Finally, I would like to thank my beloved family. Daddy, although you are no longer with us not a single day passes by without thoughts of you. You were the very first to believe in me and you gave me every opportunity to express myself under your gentle guidance. Thank you for watching over me. Mum, you truly are the best! Your strength is astounding, you did the real hard work, and despite all the downs you still manage to stay lively and positive. Thank you Mum. Joy and Enny, my beautiful Sisters, and you too my *biggest* Brother, Stephen, you have all inspired me in so many different ways that words cannot convey. I'm deeply grateful to you guys for being such role models to me. I love you.

To God be all the Glory...

John Murray-Bruce
London, UK.

Abstract

Due to significant advances made over the last few decades in the areas of (wireless) networking, communications and microprocessor fabrication, the use of sensor networks to observe physical phenomena is rapidly becoming commonplace. Over this period, many aspects of sensor networks have been explored, yet a thorough understanding of how to analyse and process the vast amounts of sensor data collected, remains an open area of research. This work therefore, aims to provide theoretical, as well as practical, advances this area. In particular, we consider the problem of inferring certain underlying properties of the monitored phenomena, from our sensor measurements. Within mathematics, this is commonly formulated as an inverse problem; whereas in signal processing it appears as a (multidimensional) sampling and reconstruction problem. Indeed it is well known that inverse problems are notoriously ill-posed and very demanding to solve; meanwhile viewing it as the latter also presents several technical challenges. In particular, the monitored field is usually nonbandlimited, the sensor placement is typically non-regular and the space-time dimensions of the field are generally non-homogeneous. Furthermore, although sensor production is a very advanced domain, it is near impossible and/or extremely costly to design sensors with no measurement noise. These challenges therefore motivate the need for a stable, noise robust, yet simple sampling theory for the problem at hand.

In our work, we narrow the gap between the domains of inverse problems and modern sampling theory, and in so doing, extend existing results by introducing a framework for solving the inverse source problems for a class of some well-known physical phenomena. Some examples include: the reconstruction of plume sources, thermal monitoring of multicore processors and acoustic source estimation, to name a few. We assume these phenomena and their sources can be described using partial differential equation (PDE) and parametric source models, respectively. Under this assumption, we obtain a well-posed inverse problem.

Initially, we consider a phenomena governed by the two-dimensional diffusion equation – i.e. 2-D diffusion fields, and assume that we have access to its continuous field measurements. In this setup, we derive novel exact closed-form inverse formulae that solve the inverse diffusion source problem, for a class of localized and non-localized source models. In our derivation, we prove that a particular 1-D sequence of, so called, generalized measurements of the field is governed by a power-sum series, hence it can be efficiently solved using existing algebraic methods such as Prony's method. Next, we show how to obtain these generalized measurements, by using Green's second identity to combine the continuous diffusion field with a family of well-chosen sensing functions. From these new inverse formulae, we therefore develop novel noise robust centralized and distributed re-

construction methods for diffusion fields. Specifically, we extend these inverse formulae to centralized sensor networks using numerical quadrature; conversely for distributed networks, we propose a new physics-driven consensus scheme to approximate the generalized measurements through localized interactions between the sensor nodes. Finally we provide numerical results using both synthetic and real data to validate the proposed algorithms.

Given the insights gained, we eventually turn to the more general problem. That is, the two- and three-dimensional inverse source problems for any linear PDE with constant coefficients. Extending the previous framework, we solve the new class of inverse problems by establishing an otherwise subtle link with modern sampling theory. We achieved this by showing that, the desired generalized measurements can be computed by taking linear weighted-sums of the sensor measurements. The advantage of this is two-fold. First, we obtain a more flexible framework that permits the use of more general sensing functions, this freedom is important for solving the 3-D problem. Second, and remarkably, we are able to analyse many more physical phenomena beyond diffusion fields. We prove that computing the proper sequence of generalized measurements for any such field, via linear sums, reduces to approximating (a family of) exponentials with translates of a particular prototype function. We show that this prototype function depends on the Green's function of the field, and then derive an explicit formula to evaluate the proper weights. Furthermore, since we now have more freedom in selecting the sensing functions, we discuss how to make the correct choice whilst emphasizing how to retrieve the unknown source parameters from the resulting (multidimensional) Prony-like systems. Based on this new theory we develop practical, noise robust, sensor network strategies for solving the inverse source problem, and then present numerical simulation results to verify the performance of our proposed schemes.

Contents

Nomenclature **24**

1. Introduction **28**
1.1. Motivation: sensing reality 28
1.2. Outline of the thesis 30
1.3. Publications 31

2. Sampling theory: generalized sampling of multidimensional signals **33**
2.1. Classical sampling theory: uniform sampling of bandlimited signals 34
2.1.1. Non-bandlimited signals, aliasing and antialiasing 34
2.2. Modern sampling theory: sampling in shift-invariant spaces 36
2.2.1. Uniform and non-uniform sampling in shift-invariant spaces 37
2.2.2. Sampling finite rate of innovation signals 38
2.2.3. Sampling multidimensional FRI signals 43
2.3. Sensing physical fields: diffusion and beyond 45
2.3.1. Diffusion: from Brown to Fourier 45
2.3.2. Beyond diffusion 49
2.4. Summary 50

3. Inverse source problem for the diffusion equation **51**
3.1. Introduction: on inverse problems for the diffusion equation 51
3.2. Forward problem 53
3.2.1. Solution to the diffusion equation: Green's function method 53
3.3. The inverse diffusion source problem 54
3.3.1. The class of localized sources 55
3.3.2. The class of non-localized sources 56
3.4. Closed-form inversion of a class of inverse diffusion source problems 57
3.4.1. Diffusion fields of multiple instantaneous point sources 58
3.4.2. Diffusion fields of multiple non-instantaneous point sources 60
3.4.3. Estimating a single non-localized source 64
3.5. An extension to advection-diffusion fields 67
3.6. Summary and Conclusion 70

4. Estimating sources of diffusion fields from sensor network measurements 72
4.1. Sensor network model and problem formulation 73
4.1.1. Centralized network model 74
4.2. Generalized measurements from spatiotemporal samples: A centralized approach 74
4.3. Robust centralized source estimation algorithms 79
4.3.1. Handling noise and model mismatch: Subspace-based denoising of $\mathcal{Q}_\epsilon(k,r)$ 80
4.3.2. Non-instantaneous source activation time & decay coefficient estimation with Cadzow 80
4.3.3. Sequential estimation of multiple localized sources 81
4.4. Distributed sensor network model 85
4.5. Gossip schemes for distributed average consensus 88
4.5.1. Quantized gossip 89
4.6. Towards a distributed source recovery: a physics-driven consensus scheme . 91
4.6.1. Consensus-based estimation of the Generalized Measurements over Sensors Networks 91
4.6.2. Noise robust consensus-based estimation 95
4.7. Source estimation in communication constrained sensor networks 100
4.7.1. Quantization Noise 100
4.8. Summary and conclusions 101

5. Simulation results 103
5.1. Introduction 103
5.2. Synthetic data simulations and results 104
5.2.1. Centralized estimation of localized sources 104
5.2.2. Centralized estimation of non-localized sources 105
5.2.3. Approximation errors due to the integral discretization 110
5.2.4. Distributed estimation over sensor networks: noiseless channels . . . 110
5.2.5. Distributed estimation over sensor networks with noisy channels . . 113
5.2.6. Comparing performance: Centralized vs Distributed estimation . . . 116
5.2.7. Estimating field sources in rooms 119
5.3. Real data experiments 122
5.3.1. Acquiring the experimental data 122
5.3.2. Source estimation over centralized SNs: real data 122
5.3.3. Source estimation over distributed SNs with quantized channels: real data 125
5.4. Summary and conclusions 127

6. Universal framework for physics-driven inverse source problems 128
6.1. Introduction 128

6.2. Physics-driven inverse problems: Problem formulation 129
6.3. Choosing the sensing functions: source estimation from generalized measurements 131
6.4. Multidimensional generalized measurements from sensor data 133
6.4.1. Function spaces, generalized sampling and function approximation . 136
6.4.2. Exact and approximate Strang-Fix theory for exponential reproduction from uniform translates 137
6.4.3. The coefficients for the space-time fields 140
6.5. Source estimation using sensor networks 144
6.5.1. Centralized source estimation 144
6.5.2. Distributed source estimation 144
6.5.3. Filtering in the time-domain 145
6.5.4. A brief note on tackling sensor noise, and model mismatch and communication constraints 146
6.6. Applications and numerical results 146
6.6.1. Inverse source problem for the diffusion equation 147
6.6.2. Inverse source problem for Poisson's Equation 150
6.6.3. Acoustic source localization: inverse source problem for the wave equation 152
6.7. Summary and conclusions 154

7. Conclusion **156**
7.1. Summary 156
7.2. Comparisons 157
7.3. Outlook and future research 158

Appendices **161**

A. Appendix A **162**
A.1. The Cramer-Rao bound 162

B. Appendix B **164**
B.1. The Generalized Strang-Fix Conditions 164
B.2. Bilateral Laplace transforms 165
B.2.1. Green's function of the two-dimensional diffusion equation 165
B.2.2. Green's function of the three-dimensional wave equation 166
B.3. Approximate weights: using Green's second identity 166

Bibliography **182**

List of Tables

3.1. Summary of the source distributions of interest along with their corresponding generalized sequence to which Prony's method or matrix pencil method can be applied. 71

5.1. Normalized Mean absolute error for an instantaneous point source estimation using Algorithm 4.1 (500 independent trials). The field is induced by $M = 1$ source with unknowns $c_1 = 1, \tau_1 = 1.213s$ and $\boldsymbol{\xi}_1 = (0.1130, 0.2210)$, sampled at $f = \frac{1}{\Delta_t} = \frac{1}{0.5} = 2Hz$ with 63 arbitrarily placed sensors. For the estimation algorithm $K = 10$ and $r = 1$ is used. 106

5.2. Normalized Mean absolute error of single source parameter estimates using Algorithm 4.1 (500 independent trials). The field is induced by a single ($M = 1$) time-varying source with unknowns $\alpha_1 = -0.1, c_1 = 100, \tau_1 = 3.5s$ and $\boldsymbol{\xi}_1 = (0.1130, 0.2130)$ and is sampled at $\frac{1}{\Delta_t} = \frac{1}{0.5} = 2Hz$ with 63 arbitrarily placed sensors. For the estimation we use $K = 50$, $T_1 = 6s$, $T_{end} = 14s$, $P = 5$, $Q = 3$, and $\Delta T = 2s$. 107

5.3. **Normalized MAE of the centralized (Algorithm 4.1) and distributed (Algorithm 4.5) estimation algorithms.** The induced single ($M = 1$) source field, with c_1=1, τ_1=1.213, $\boldsymbol{\xi}_1$=(0.1130, 0.2210), is sampled at $\frac{1}{\Delta_t}$=1Hz over T=10s using N=45 sensors. The samples are corrupted with AWGN to obtain the specified SNR. We choose $r = 1$ and $K = 1$ and $K = 3$ for the estimation algorithms and present the normalized MAE statistic computed with results from 1000 independent trials at the chosen SNR levels. . 118

5.4. **MAE of centralized single source parameter estimates on real thermal data.** Field induced by point instantaneous source ($\tau = 1.7800s$ and $\boldsymbol{\xi} = (0.0594, 0.0298)m$) and its spatiotemporal samples are obtained at $f = \frac{1}{0.52} \approx 1.9231Hz$, over $T_{end} = 12s$. The MAEs are computed from 1000 independent trials. Here $K = 11$ and $r = 1$ is used for Algorithm 4.1. . 124

5.5. **Centralized source estimation results for six independent read data set-ups.** In each experiment the recording is downsampled so $f = \frac{1}{0.52}$Hz and $f = \frac{1}{0.5}$Hz for Expts. I–II and III–VI, respectively. 13 locations are chosen randomly as sensor positions. We estimate the unknowns from samples over $T_{end} = 12s$. The true parameters and corresponding estimates obtained by Algorithm 4.1 using $K = 11$ are shown. 124

6.1. Comparing the frameworks of Chapters 4 and 6. 155

List of Figures

2.1. **Spectrum of aliased signal.** The spectrum of the continuous signal is compactly supported on $[-W, W]$ and the sampling frequency is $f_s < 2W$. The spectrum of our sampled signal (shaded region) is a superposition of the original signal's spectrum and shifted copies, shifted by integer multiples of the sampling frequency f_s. Thus, Shannon's formula will return a spectrum that is both aliased and truncated to $[-f_s/2, f_s/2]$; the aliasing errors are the shaded areas above the dark (blue) curve. 35

2.2. **Spectrum of antialiased signal.** Prefiltering with an ideal antialiasing filter means the spectrum is now supported on $[-f_s/2, f_s/2]$ as such there is no overlap in the shifted copies obtained due to sampling, Shannon's formula will return the truncated but non-aliased baseband spectrum. 36

2.3. **Schematic of practical sampling process.** The continuous signal is prefiltered, sampled and then postfiltered to obtain the optimal least-squares approximation of $f(x)$ in the space of bandlimited functions. 36

2.4. **Examples of FRI signals.** (a) A superposition of four damped and compactly supported sinusoids with different locations and amplitudes, i.e. $f(x) = \sum_{m=1}^{4} c_m g(x - \xi_m)$ where the characteristic pulse shape $g(x) = (1.3 + \cos(x))e^{-\frac{x^2}{30}}(H(x + 50) - H(x - 50))$ and $H(x)$ is the Heaviside step function. (b) Superposition of five Dirac deltas with different amplitudes and locations, i.e. $f(x) = \sum_{m=1}^{5} c_m \delta(x - \xi_m)$. (c) The superposition of four truncated sinusoids with different amplitudes, frequencies, phases and locations, specifically $f(x) = \sum_{m=1}^{4} c_m \cos(\omega_m x + \phi_m)(H(x - \xi_m) - H(x - \xi_{m+1}))$. When the pulse shapes are known, these signals are fully specified once the corresponding pulse locations and amplitudes are found. 39

2.5. **Schematic of FRI sampling process.** The sampling kernel $g(x)$ now enjoys the freedom to be more general than the sinc function, however but it must have certain function reproducing properties, such as being polynomial and exponential reproducing. 40

2.6. **Estimating the thermal distribution of a two-dimensional plate.** Shown in (a) is the thermal distribution of a 2-D plate heated by four unknown localized and instantaneous sources. Using a network of 121 sensors the field is sensed at locations in (d); assuming ideal sensors the noiseless samples obtained are visualized in (e); whereas the noisy samples in a more realistic situation are visualized in (f). Using these noiseless and noisy samples, we obtain corresponding bandlimited reconstructions presented in (b) and (c) respectively. We notice significant errors in both reconstructions due to aliasing since the field is not bandlimited, also, we cannot apply an antialiasing filter in space. 47

4.1. **Sensor Placement.** An arbitrary sensor placement and the monitored domain Ω. 73

4.2. **A sensor network and its Delaunay triangulation.** The (approximate) monitored domain Ω divided into triangular meshes and the domain boundary divided into straight line segments (black solid lines). 77

4.3. **A sensor network example.** Links between sensors as modelled by a random geometric graph. 87

4.4. **A sensor network and its Delaunay triangulation.** The (approximate) monitored domain Ω divided into triangular meshes and the domain boundary divided into straight line segments (black solid lines). 88

4.5. **State evolution of quantized gossip schemes.** Shows at each gossip iteration the state evolution of the agents in the network assuming a 5-bit uniform quantizer is used; the desired average is denoted by the dashed grey line. We use in: (a) Naive Quantized Gossip (NQG); (b) Kashyap's Quantized Gossip (KQG); and (c) Symmetric Quantized Gossip (SQG). . . 91

4.6. **SDQNR of Quantized Gossip Schemes.** The realized SDQNR vs number of quantization bits (q-bits for $q = 2, \ldots, 10$) using (a) Frasca's Symmetric Quantized Gossip Scheme and (b) Kashyap's Quantized Gossip Scheme. This simulation uses $M = 2$, $K = 5$, $N = 90$, the network is a RGG, $\mathcal{G}(N, 0.6)$. The solid line in both plots shows the signal-to-noise ratio of the quantized version of ground truth Prony's sequence, i.e. $\{Q_q(\mathcal{R}(k))\}_k$ for each q. 101

5.1. **Estimation of $M = 3$ diffusion sources using randomly distributed sensors.** The spatiotemporal samples have SNR equal to 20dB and the results of 20 independent trials are shown. Intensities $c_1 = c_2 = c_3 = 1$; locations $\boldsymbol{\xi}_1 = (0.113, 0.221), \boldsymbol{\xi}_2 = (0.234, 0.175), \boldsymbol{\xi}_3 = (0.070, 0.100)$; and activation times $t_1 = 3.2s, t_2 = 8.1s, t_3 = 15.4s$. Field is sampled for $T_{end} = 20s$ at a frequency $\frac{1}{\Delta_t} = 2Hz$ and $K = 5$ i.e. $k = 0, 1, \ldots, 5$ for the test function family $\Psi_k(\mathbf{x}) = e^{-k(x_1+\mathrm{j}x_2)}$. The scatter-plot shows the true source locations (blue '+'), the estimated locations (red '×') and one typical realization of the sensor distribution (green '○') using Algorithm 4.4. We use in (a) $N = 45$ randomly distributed sensors; whilst in (b) $N = 63$ randomly distributed sensors. 106

5.2. **Estimation of $M = 1$ non-instantaneous diffusion source using 63 randomly distributed sensors.** The unknown source has intensity $c_1 = 100$, decay coefficient $\alpha_1 = -0.1$, location $\boldsymbol{\xi}_1 = (0.1130, 0.2130)$, and activation time $\tau_1 = 3.5s$. The field induced is sampled for $T = 14s$ at $\frac{1}{\Delta_t} = 2Hz$ and the samples obtained have SNR = 20dB. Moreover, we perform 15 independent trials using $K = 50$ for $\{\mathcal{R}(k)\}_{k=0}^{K}$ whilst $T_1 = 6s$, $\Delta T = 2s$ and $P = Q + 2 = 5$. The scatter-plot shows the true source locations (blue '+'), their estimated locations (red '×') and a typical realization of the sensor network distribution (green '○') Algorithm 4.1. 107

5.3. **Centralized estimation of non-localized sources using noiseless spatiotemporal measurements obtained by randomly placed sensors.** The field is sampled at $10Hz$ for $T = 7s$ using 45 sensors in (a) and 90 sensors in (b). In addition, for the centralized estimation parameters, we use in: (a) $R = 5$ and $K = 4$ for the line source algorithm, i.e. Algorithm 4.2, whereas (b) $R = 5$ and $K = 6$ for the triangular source algorithm, i.e. Algorithm 4.3. 108

5.4. **Line source estimation using noisy spatiotemporal measurements.** The field is sampled at $10Hz$ for $T = 10s$ using 45 arbitrarily placed sensors, with measurement SNR= 20dB. For the spatial and temporal sensing functions family, $K = 6$ and $R = 5$ respectively using Algorithm 4.2. For each trial we show in: (a) the estimated line source, and in (b) the corresponding activation time estimates. 108

5.5. **Triangular source estimation using noisy spatiotemporal measurements.** The field sampled at $10Hz$ for $T = 10s$ using 90 arbitrarily placed sensors, with measurement SNR= 35dB. For the spatial and temporal sensing functions, $K = 9$ and $R = 8$ respectively using Algorithm 4.3. For each trial we show in: (a) the vertices of the estimated triangular source, and in (b) the corresponding activation time estimates. 109

5.6. **Location estimation performance.** Standard deviation (500 trials) for localization estimates of 2 sources from 63 randomly distributed sensors, using Algorithm 4.1 with $K = 5$. $c_1 = c_2 = 1$; $\boldsymbol{\xi}_1 = (0.23, 0.15)$, $\boldsymbol{\xi}_2 = (0.15, 0.15)$ and $\tau_1 = \tau_2 = 1.2s$. $T_{end} = 10s$ and $\frac{1}{\Delta_t} = 1Hz$. 110

5.7. **Single diffusion source estimation using distributed SNs with noiseless links.** The evolution of the estimated source parameters with increasing number of pairwise message exchanges (Algorithm 4.5). Sensor measurements are assumed to be noiseless and we use $K = 1$ and $r = 1$ for the generalized measurements $\{\mathcal{Q}(k, r)\}_k$. (a) Evolution of location estimates; true source location is where the dashed lines intersect. (b) Evolution of intensity (top) and activation time (bottom) estimates. 111

5.8. **Multiple diffusion source estimation over distributed SNs with noiseless links.** The unknown source parameters estimated by each sensor using Algorithm 4.5 is shown for 10 independent trials. Communication links are ideal, and so messages are unquantized. Furthermore, the spatiotemporal samples are assumed noisy with SNR= 20dB. Here we use $K = 10$ and $r = 1$. (a) shows one realization of the SN (sensors and links), the scatterplot in (b) shows the true source locations (blue '+'), their estimates (red '×') and a realization of the sensor locations (green '•'), whilst (c) shows the estimated activation times. 112

5.9. **Single diffusion source estimation using distributed SNs with noisy links ($K = 1$).** We assume noiseless sensor measurements and Algorithm 4.5 with quantized gossiping is utilized for $K = 1$ and $r = 1$. In (a) $\mathcal{G}(45, 0.4)$ and in (b) a scatterplot of the location estimates at convergence. (c) Evolution of location estimates (directed grey lines) along with the true location (intersection of the dashed lines). (d) Evolution of source intensity (top) and activation time (bottom) estimates for three randomly chosen nodes. 114

5.10. **Single diffusion source estimation using distributed SNs with noisy links ($K = 5$).** We assume noiseless sensor measurements and Algorithm 4.5 with quantized gossiping is utilized for $K = 5$ and $r = 1$. In (a) $\mathcal{G}(45, 0.4)$ and in (b) a scatterplot of the location estimates at convergence. (c) Evolution of location estimates (directed grey lines) along with the true location (intersection of the dashed lines). (d) Evolution of source intensity (top) and activation time (bottom) estimates for three randomly chosen nodes. 115

5.11. **Multiple diffusion source estimation over SNs with noisy links (quantized gossip).** Noisy sensor measurements with SNR= 20dB, and noisy inter-sensor links are assumed. To estimate the unknowns we utilize Algorithm 4.5 under quantized gossiping for $K = 10$ and $r = 1$ and show results (at convergence) for 10 independent trials. (a) a realization of $\mathcal{G}(63, 0.4)$, whilst in (b) the scatterplot shows the true (blue '+') and estimated (red '×') source locations, along with a typical sensor distribution (green '•'). (c) Estimated activation times due to each sensor upon convergence. 116

5.12. **Performance of the centralized (Algorithm 4.1) and distributed (Algorithm 4.5) estimation algorithms.** A single source field, with c_1=1, τ_1=1.213, $\boldsymbol{\xi}_1$=(0.1130, 0.2210) is considered and its source parameters are estimated from the (noisy) spatiotemporal samples taken at $\frac{1}{\Delta_t}$=1Hz over T=10s using N=45 sensors, with the spatial sensing function family chosen such that $K = 1$. We show the MSE of the centralized and distributed algorithms computed using results from 5000 independent trials for varying signal-to-noise ratios: in (a) MSE of $\xi_{1,1}$, (b) MSE of $\xi_{2,1}$ and, (c) MSE τ_1. . 118

5.13. **Diffusion in bounded region.** Spatial field distribution of a single source in a bounded square region at different time instants after source activation. 119

5.14. **Multiple diffusion source estimation in a bounded region.** The $M = 3$ diffusion sources have $c_1 = c_2 = c_3 = 1$, $\boldsymbol{\xi}_1 = (0.113, 0.221), \boldsymbol{\xi}_2 = (0.234, 0.175), \boldsymbol{\xi}_3 = (0.070, 0.100)$, and $\tau_1 = 3.2s, \tau_2 = 8.1s, \tau_3 = 15.4s$. The field induced is sampled at $\frac{1}{\Delta_t} = 2Hz$ for $T_{end} = 25s$, using 45 sensors placed randomly inside the square room. The samples have SNR= 20dB. Estimation is performed using Algorithm 4.4 with $K = 11$ and $r = 1$. The estimated locations (left) and activation times (right) are shown. 120

5.15. **Real data experimental setup.** 121

5.16. **Measurements of two monitoring sensors obtained at different spatial locations.** The dotted vertical line in each plot indicates the instant of source activation. 121

5.17. **Centralized estimation of a single instantaneous heat source using real thermal data.** Samples are obtained from the thermal images by choosing 13 spatial locations (the circles '∘' in plots (a) and (b)) at random, and then downsampling so that $f = \frac{1}{0.52} \approx 1.9231Hz$ over the interval $T_{end} = 12s$. The true source parameters $\boldsymbol{\xi}_1 = (0.0594, 0.0298)m$ and $\tau_1 = 1.7800s$. We use Algorithm 4.1 and choose $K = 11$. (a) Thermal image immediately after activation, sensors '∘' and estimated source locations '×'. (b) Summary for 20 repetitions, estimated source locations ('×' left) and activation times (right). 123

5.18. **Distributed single source estimation with real thermal data.** The RGG, scatterplot of the locations and activation times estimated by the sensors are shown for each independent trial. The distributed estimation is performed by SNs with noisy communication links using Algorithm 4.5 under quantized gossiping, such that a 15−bit uniform quantizer with dynamic range $(-1, 1)$ is used. 125

5.19. **Sensor distribution, location estimates and the field reconstructions.** Source is located at $(0.06058, 0.03465)$ and activated at $\tau = 6.25s$. A 15−bit uniform quantizer with dynamic range $(-1, 1)$ is used, whilst the estimation is performed with Algorithm 4.5 under quantized gossiping for the choice $K = 5$ and $r = 1$. 126

6.1. **Exponential reproduction.** Reproduction of the 2-D sensing function $\Psi_{\boldsymbol{k}}(\mathbf{x}) = e^{\mathrm{j}k_1x_1+\mathrm{j}k_2x_2}$, assuming $N = 45$ uniformly placed sensors for the 2-D Green's function of Poisson's equation. 143

6.2. **Centralized estimation of non-localized sources using noisy measurements obtained by randomly placed sensors.** The field is sampled at $1Hz$ for $T = 25s$ using 41 uniformly placed sensors, samples assumed to have SNR = 15dB. (a) Location estimates (b) Estimated intensities and (c) Activation times. We use Algorithm 6.1 with $r = 1$ and $K_1 = K_2 = 15$. . . . 148

6.3. **Effects of sensor density and source separation on the performance of the localization algorithm.** The field induced by two sources is sampled at 1Hz for $T = 20s$ with varying sensor density and source separation. For each value sensor separation and source separation the obtained noisy measurements, with SNR= 20dB, are used to estimate the locations of the unknown diffusion sources. We utilize $K_1 = K_2 = 4$ and $r = 0$ with the estimation simultaneous recovery scheme summarized in Algorithm 6.1. The RMSE of the estimates obtained over 1000 independent trials are shown: i.e. in (a) RMSE of $\hat{\xi}_{1,1}$, in (b) RMSE of $\hat{\xi}_{2,1}$, in (c) RMSE of $\hat{\xi}_{1,2}$ and in (d) RMSE of $\hat{\xi}_{2,2}$. 149

6.4. **Centralized estimation of single localized source for Poisson's equation using uniformly placed sensors.** The field is sampled at $1Hz$ for $T = 1s$ using 27 randomly placed sensors, samples assumed to have SNR = 15dB. (a) Location estimates (b) Estimated intensities. Estimation is performed using Algorithm 6.1 with $r = 0$ and $K_1 = K_2 = 7$. 151

6.5. **Centralized estimation of non-localized sources for the Wave equation using uniformly placed sensors.** The field is sampled at $1Hz$ for $T = 20s$ using 27 randomly placed sensors, samples assumed to have SNR = 15dB. (a) Location estimates and (b) Estimated activation times. We use Algorithm 6.1 with $r = 1$ and $K_1 = K_2 = 9$. 153

7.1. **Comparing the source estimation frameworks.** Similarities and differences between the framework described in Chapter 4 and that described in Chapter 6. 158

Nomenclature

List of Acronyms

ADE	Advection-Diffusion Equation
EEG	Electroencephalogram
IADSP	Inverse Advection-Diffusion Source Problem
IDCP	Inverse Diffusion Coefficient Problem
IDS	Inverse Diffusion Source
IDSP	Inverse Diffusion Source Problem
ISP	Inverse Source Problem
KQG	Kashyap's Quantized Gossip
MAE	Mean Absolute Error
MSE	Mean Squared Error
NMAE	Normalized Mean Absolute Error
NQG	Naive Quantized Gossip
ODE	Ordinary Differential Equation
PDE	Partial Differential Equation
s.d.	Standard deviation
SNR	Signal to Noise Ratio
SN	Sensor Network
SQG	Symmetric Quantized Gossip

Numbers

$\mathbb{Z}$	Integers
$\mathbb{N}$	Non-negative integers (naturals)

$\mathbb{R}$	Real numbers
$\mathbb{R}_{\geq 0}$	Non-negative real numbers
$\mathbb{R}_{>0}$	Positive real numbers
$\mathbb{C}$	Complex numbers

Constants

μ	Diffusivity of diffusion medium
M	Total number of activated (point) sources
$\mathbf{v}_a$	Velocity (vector) of the advection current
N	Total number of monitoring sensors
I	Total number of boundary sensors, i.e. $I = \lvert\mathcal{S}_{\partial\Omega}\rvert$

Spaces

$\mathcal{L}^2$	Space of all square-integrable functions
ℓ^2	Space of all square-summable sequences
$V_{\mathbf{\Delta_x}}(g)$	Space spanned by integer (uniform) translates of the generator $g(\mathbf{x}/\mathbf{\Delta_x})$
$\mathcal{H}_C^p$	Space of all compactly supported functions whose derivatives up to and including $p \geq 0$ are in $\mathcal{L}^2$

Operators

$\frac{\partial}{\partial t}$	Temporal derivative, i.e. derivative w.r.t t
∇	Spatial derivative, i.e. $\nabla = \left(\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \ldots, \frac{\partial}{\partial x_d}\right)$
$*$	Spatiotemporal convolution, i.e. $(f * g)(\mathbf{x}, t) = \int_{\mathbf{x}' \in \mathbb{R}^d} \int_{t' \in \mathbb{R}_+} f(\mathbf{x}', t') g(\mathbf{x} - \mathbf{x}', t - t') \mathrm{d}\mathbf{x}' \mathrm{d}t'$
$\lvert\cdot\rvert$	Absolute value of a number, or the cardinality of a set
$\langle f, g \rangle$	The inner product, i.e. $\int_{\mathbf{x} \in \mathbb{R}^d} fg \mathrm{d}\mathbf{x}$
$\arg(\cdot)$	Argument of a complex number $z = x + \mathrm{j}y$, i.e. $\arg(z) = \arctan(y/x)$.
$\mathcal{DT}(\cdot)$	The Delaunay triangulation of a set of points (sensor locations)
$\mathcal{CH}(\cdot)$	The convex hull of a set of points (sensor locations), i.e. the convex hull boundary
$\mathrm{vec}(\cdot)$	Column-wise vectorization of a matrix

Signals, Functions, Distributions, Sets and Sequences

$\delta, \delta(\mathbf{x}, t)$	Dirac's delta function
$u(\mathbf{x}, t)$	Diffusion field intensity at position $\mathbf{x}$ and time t
$f(\mathbf{x}, t)$	Source distribution
$H(t)$	Heaviside's unit step function
$L(\mathbf{x})$	Straight line segment with endpoints $\boldsymbol{\xi}_1$ and $\boldsymbol{\xi}_2$
$F(\mathbf{x})$	M-sided polygon with vertices $\{\boldsymbol{\xi}_m\}_{m=1}^{M}$
$\Psi_k(\mathbf{x})$	Spatial sensing function, for e.g. $\Psi_k(\mathbf{x}) = e^{-k(x_1+\mathrm{j}x_2)}$
$\Gamma_r(t)$	Temporal sensing function, for e.g. $\Gamma_r(t) = e^{-\mathrm{j}rt/T}$
Ω	Monitored region/domain containing field sources of interest
$\partial\Omega$	Boundary of monitored region
$\mathcal{S}$	Set of all N sensor locations, i.e. $\mathcal{S} = \{\mathbf{x}_n\}_{n=1}^{N}$
$\mathcal{S}_{\partial\Omega}$	The cyclically ordered set of all I boundary sensor locations, i.e. $\mathcal{S}_{\partial\Omega} = \{\mathbf{x}_n \vert \mathbf{x}_n \in \partial\Omega, n = 1, \ldots, N\}$
$\varphi_n(t_l)$	Diffusion field measurement at sensor position $\mathbf{x}_n$ and time instant t_l, i.e. $\varphi_n(t_l) = u(\mathbf{x}_n, t_l)$
$\varphi_{j,j'}(t_l)$	Diffusion field measured at $\mathbf{v}_{j,j'}$ and instant t_l, i.e. $\varphi_{j,j'}(t_l) = u(\mathbf{v}_{j,j'}, t_l)$
$\Phi_{n,r}(T)$	Appears in (4.9): it is a quadrature-based approximation of the integral (4.4a), so $\Phi_{n,r}(T) \approx U_r(\mathbf{x}_n, T)$
$\dot{\Phi}_{n,r}(T)$	Appears in (4.10): it is a quadrature-based approximation of the integral (4.8), so that $\dot{\Phi}_{n,r}(T) \approx \dot{U}_r(\mathbf{x}_n, T)$
$(\Phi_{n,r}(T))_{x_i}$	Appears in (4.11): it is a first order (central finite difference) approximation of the derivative $\frac{\partial}{\partial x_i} U_r(\mathbf{x}, T)$ at $\mathbf{x} = \mathbf{x}_n$
$\mathcal{S}'_{\partial\Omega}$	The set of the nearest sensors to a boundary sensor location
$\epsilon_{n,l} \sim \mathcal{N}(0, \sigma^2)$	Additive white Gaussian noise sequence for $n = 1, \ldots, N$ and $l = 0, 1, \ldots, L$
$\epsilon(k, r)$	Coloured noise sequence generated from the AWGN sequence $\{\epsilon_{n,l}\}_{n,l}$, specifically $\epsilon(k, r) = \sum_{n,l} w_{n,l}(k, r)\epsilon_{n,l}$
$\mathcal{Q}(k, r)$	Sequence of generalized measurements indexed by k for a given r
$\mathcal{R}(k)$	1-D sequence of generalized measurements, specifically $\mathcal{R}(k) \stackrel{\text{def}}{=} \mathcal{Q}(k, 0)$.

$\mathcal{Q}(\boldsymbol{k}, r)$	Vector sequence of generalized measurements indexed by $\boldsymbol{k}$ for a given r
$G(\mathbf{s}_{\mathbf{x}}, s_t)$	The multidimensional Laplace transform of $g(\mathbf{x}, t)$, defined in (6.23)

Variables, Vectors and Matrices

t	Temporal variable, $t \in \mathbb{R}_+$
$\mathbf{x}$	Spatial variable, $\mathbf{x} = (x_1, x_2, \ldots, x_d) \in \mathbb{R}^d$ with $x_i \in \mathbb{R}$
t_l	Temporal sampling instant, $t_l = l\Delta_t \in \mathbb{R}_+$ with $l = 0, 1, \ldots, L$
$\mathbf{x}_n$	Spatial location of the n-th sensor, $\mathbf{x}_n = (x_{1,n}, \ldots, x_{d,n}) \in \Omega$
Δ_j, $\lvert\Delta_j\rvert$	Denotes the j-th triangle of a Delaunay triangulation, and its area respectively
$\Delta_{n,j}$, $\lvert\Delta_{n,j}\rvert$	Denotes the j-th triangle in a Delaunay triangulation with a vertex at location $\mathbf{x}_n$, and its area respectively.
$\mathbf{v}_{j,j'}$	Denotes the j'-th vertex of the j-th triangle in a Delaunay triangulation of the sensor locations.
$\mathbf{F}_{\text{pw}}$	Prewhitening filter (matrix)
$\mathbf{C}_\epsilon$	Noise covariance matrix, i.e. $\mathbf{C}_\epsilon = \mathbb{E}\left\{\mathbf{E}\mathbf{E}^{\mathsf{H}}\right\}$
$\mathbf{T}$	Toeplitz matrix constructed from the ideal generalized measurements, i.e. $\mathbf{T} = \text{Toep}(\{\mathcal{Q}(k, r)\}_k)$
$\mathbf{T}_\epsilon$	Toeplitz matrix generated from noisy generalized measurements, i.e. $\mathbf{T}_\epsilon = \text{Toep}(\{\mathcal{Q}_\epsilon(k, r)\}_k)$
$\mathbf{E}$	Toeplitz matrix generated from the (coloured) noise sequence $\{\epsilon(k, r)\}_k$, hence $\mathbf{E} = \text{Toep}(\{\epsilon(k, r)\}_k)$
$w_{n,l}(k, r)$	Coefficients in double sum approximation of generalized measurements (integrals)
$\mathbf{w}(k, r)$	Vectorization of the coefficients
$\mathbf{0}, \mathbf{1}$	The vectors with all entries equal to zeros, ones respectively
k, m, n, r	Integer-valued scalar indices
$\boldsymbol{k}, \boldsymbol{n}$	Vector-valued indices of integers

Chapter 1.

Introduction

As humans we are, in general, endowed with five senses through which we can perceive, experience and interact with our surrounding environment. Indeed having developed, refined and specialized this ability via complex evolutionary mechanisms over millions of years, our capability to sense our surroundings and react accordingly is key to survival in nature. However, certain phenomena remain imperceptible—to us—as such humankind have developed, and will continue to develop, external devices to sense these otherwise imperceptible phenomena; thereby further enriching our level of interaction with the surrounding environment. Through the years, these external sensing devices have assumed many different forms and complexities: ranging from the simple *cheugugi* which was the first rain gauge invented and used by the Joseon Dynasty of Korea to measure rainfall, to the more intricate ATLAS and CMS detectors of the famous large hadron collider, used to detect the elusive Higgs boson (aka *god particle*). In these examples, the information collected by the sensor is used to answer an important question, respectively: "how much rainwater has been absorbed into the ground?", or, "does the Higgs boson exist?" Consequently, a sensor is any device that detects, interacts or measures the physical reality by responding to it in some way.

Recently sensors, and the use thereof, have become ubiquitous in almost all aspects of life such as: healthcare, environmental monitoring, safety, entertainment, astronomy, communication and many more. Moreover, significant advances in sensor technologies leading to cheaper and more efficient devices, over the years, has necessitated the development of tools and techniques to store, process and even transmit the collected information. This new paradigm has led to an exponential surge in the amount of data collected, thus fuelling the growth of many fields like sampling & interpolation, compression, estimation and denoising. It is clear that signal processing holds the key to answering many of the questions posed by this big data paradigm.

1.1. Motivation: sensing reality

Most signals that we encounter in reality are, generally, continuous over space and time; hence in order to store, process, or transmit them using digital systems, and digital signal processing techniques, they have to be discretized. In general, a discretized version of a

continuous signal can be obtained by *sampling* it on a discrete set. For example a youtube video recording is simply a sequence of images obtained at uniform time intervals, and each image is acquired by sampling light intensities—of the scene—on a uniform (Cartesian) grid. Whilst this instance describes a uniform sampling setup, the sampling set needn't be uniform; specifically nonuniform polar and spiral sampling sets are common in computerized tomography and magnetic resonance imaging, for example. In these applications, and beyond, the natural sampling problem that arises presents two challenges. The first challenge is to specify the conditions under which it is possible to recover the original signal from different sets of its discrete samples; whilst, the second is to develop explicit signal reconstruction schemes for analysing and processing these samples.

Let the sampling sets be $\mathcal{S} = \{\mathbf{x}_n\}_{n\in\mathrm{N}} \subset \mathbb{R}^d$, where N is a countable index set and d is the number of (spatial) dimensions. In fact it is well known that, there are infinitely many functions that can result in the the same sampled values on $\mathcal{S}$, therefore in order to obtain a meaningful sampling problem, some a priori constraints must be imposed on the original continuous signal. The classical constraint is that the signal (or function) of interest belongs to a space of *bandlimited* functions – which means that it is limited to a known maximum frequency. In addition if $\mathcal{S}$ is a uniform sampling set, then the classical result of Shannon[135], Nyquist [114], Whittaker [155] and Kotel'nikov [85] provides an explicit reconstruction formula. On the other hand, for nonuniform sampling sets, the works of Paley and Weiner [156], Kadec [79], Beurling [21] and Landau [89] provide suitable conditions on $\mathcal{S}$ and corresponding reconstruction formulae. Furthermore, in the last few decades these results have been generalized to produce a more unified framework of sampling on *shift-invariant spaces*. Specifically, the class of bandlimited functions are a subspace of a shift-invariant space [8].

Whilst these elegant results have been extremely useful and important in signal processing, certain assumptions are an idealization. In reality, signals very rarely live in a bandlimited or shift-invariant space. Given a nonbandlimited signal, for example, it is common to artificially bandlimit it, by *prefiltering* with an ideal antialiasing filter, before sampling and performing a bandlimited reconstruction using Shannon's formula. Although, it has been shown that this process yields the "best" bandlimited approximation of the original signal, in certain practical situations (such as imaging and monitoring of physical fields) it is generally impossible to apply a spatial prefilter before sampling. Hence we either need to place an extremely large number of sensors in the monitored region or suffer the effects of aliasing errors when using a bandlimited reconstruction approach. In this work, we aim to solve the (inverse) problem relating to the inference of certain unknown parameters of some physical fields of interest from their temporal samples obtained at a discrete set of spatial sensor locations $\mathcal{S}$. Take for example, the Fukushima disaster where it is of interest to find the sources (i.e. the exact locations, release times and initial intensities) of the nuclear fallouts diffusing through space. In applications such as these, the induced field is nonbandlimited and the monitoring sensors are sparse. Therefore, it becomes difficult to extract useful information using standard techniques. Hence it becomes necessary to

develop new and sophisticated signal processing tools and techniques with which to process our sensor data.

In this thesis we will derive a sampling framework for analysing physical fields. In particular, we explore and present new robust signal processing algorithms to solve the resulting class of inverse source problems for diffusion fields, given sensor measurements of the field. Next we leverage from our results to develop a universal framework for analysing a wider class of physical fields driven by well-known partial differential equations.

1.2. Outline of the thesis

The remainder of this thesis is organised as follows. In Chapter 2 we begin our treatment of physics-driven inverse source problems with an overview of the present state of the art in sampling theory, along with a review of the necessary mathematical prerequisites.

Chapter 3 considers the inverse diffusion source problem for a class of source models. Specifically, we consider diffusion fields induced by a given class of parametric source models and derive simple, yet exact closed-form inversion formulae given continuous measurements of the field. In our treatment, we demonstrate through the use of Green's second identity, that by combining the diffusion field using a family of well-chosen sensing functions we can transform the inverse problem into one of estimating the amplitudes and frequencies of a superposition of sinusoids from a specific sequence, which we refer to as *generalized measurements*. In fact this new, frequency estimation problem is well-studied in applied mathematics and signal processing and can be solved using Prony's method and its variations to reveal the desired source parameters.

In Chapter 4 we adapt the exact inversion formulae obtained in Chapter 3 to realistic settings. In this new situation the diffusion field can only be measured using a network of digital sensors situated at arbitrary spatial locations in space; as such we would only have access to discrete spatiotemporal measurements of the field. We provide noise robust schemes that successfully solves this discrete inverse source problem for the diffusion equation when all the measurements are made available at a fusion centre. Next we show how to distribute the computation, in particular we derive a physics-based gossip scheme for solving the inverse source problem via localized communications between the monitoring sensors.

In Chapter 5 we present simulation results to validate the proposed centralized and distributed diffusion source estimation schemes. We demonstrate the performance of our algorithms on both synthetic and real data.

In Chapter 6 we demonstrate that it is possible to generalize the proposed approach to a larger class of physical fields governed by linear partial differential equations. Specifically we show that the inverse source problem, for this class of fields, can be mapped again to the same problem in Chapter 3 of estimating the amplitude and frequencies of a superposition of sinusoids from generalized measurements. We show how to compute these generalized measurements for arbitrary fields using recent results in modern sampling theory. We then

conclude the chapter with some numerical results for different fields.

Finally, in Chapter 7 we conclude the thesis by summarizing the main contributions, whilst also providing possible directions for future research studies.

1.3. Publications

The materials presented in this thesis have led to following publications:

To be Submitted

- J. Murray-Bruce and P. L. Dragotti. "A universal sampling framework for physics-driven inverse source problems." **To be submitted to:** *IEEE Transactions on Signal Processing.*

Peer-reviewed Journals

- J. Murray-Bruce and P. L. Dragotti. "Physics-driven quantized consensus for distributed diffusion source estimation using sensor networks." *EURASIP Journal on Advances in Signal Processing*, 2016(1), pp. 1–22, February 2016.

- J. Murray-Bruce and P. L. Dragotti. "Estimating localized sources of diffusion fields using spatiotemporal sensor measurements." *IEEE Transactions on Signal Processing*, 63(12), pp. 3018-3031, June 2015.

Peer-reviewed Conferences

- J. Murray-Bruce and P. L. Dragotti. "Solving inverse source problems for linear PDEs using sparse sensor measurements." *In Proc. 50th Asilomar Conference on Signals, Systems and Computers (Asilomar'16)*, California, United States, November 2016.

- J. Murray-Bruce and P. L. Dragotti. "Solving physics-driven inverse problems via structured least-squares." *In Proc. 2016 European Signal Processing Conference (EUSIPCO'16)*, Budapest, Hungary, August 2016.

- J. Murray-Bruce and P. L. Dragotti. "Reconstructing non-point sources of diffusion fields using sensor measurements." *In Proc. 41st IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'16)*, pp. 4004–4008, Shanghai, China, March 2016.

- J. Murray-Bruce and P. L. Dragotti. "Consensus for the distributed estimation of point diffusion sources in sensor networks." *In Proc. 40th IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'15)*, pp. 3262–3266, Brisbane, Australia, April 2015.

- J. Murray-Bruce and P. L. Dragotti. "Estimation of multiple instantaneous point sources of advection-diffusion fields from field samples." *In Proc. 10th IMA International Conference on Mathematics in Signal Processing*, Birmingham, UK, December 2014.

- J. Murray-Bruce and P. L. Dragotti. "Reconstructing diffusion fields sampled with a network of arbitrarily distributed sensors." *In Proc. 2014 European Signal Processing Conference (EUSIPCO'14)*, pp. 885–889, Lisbon, Portugal, September 2014.

- J. Murray-Bruce and P. L. Dragotti. "Spatio-temporal sampling and reconstruction of diffusion fields induced by point sources." *In Proc. 39th IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'14)*, pp. 31–35, Florence, Italy, May 2014.

Conference abstracts, Workshops and Invited talks

- J. Murray-Bruce and P. L. Dragotti. "A sampling and reconstruction framework for solving PDE-driven inverse problems." *In Proc. 2016 Workshop on Inverse Problems for PDEs (IP'16)*, Bremen, Germany, March 2016.

- J. Murray-Bruce and P. L. Dragotti. "Sensor networks and inverse problems of diffusion fields", *In Proc. 53rd Annual Allerton Conference on Communication, Control, and Computing (Allerton'15)*, Illinois, United States, September 2015 (Invited Talk - special session on Inverse Problems).

Chapter 2.

Sampling theory: generalized sampling of multidimensional signals

Sampling theory underpins the connection between the (continuous) real world and the (discrete) digital systems. This connection between the continuous and discrete worlds is established via *sampling* and *interpolation*. Broadly speaking sampling produces a numerical sequence from a function by observing or measuring it at specific points; whilst the converse of producing a continuous function from a sequence of numbers is known as interpolation. Sampling theory was heavily influenced by the pioneering works of Shannon, Nyquist, Whittaker and Kotel'nikov [85, 135, 155, 113], which led to the Whittaker-Kotel'nikov-Shannon sampling theorem. Their theorem describes a class of signals, called bandlimited signals, that can be perfectly reconstructed from their samples and provides an elegant reconstruction formula. In reality however, most signals do not fall within this class, hence several approaches have been proposed to generalized and extend this classical sampling theorem. The first half of this chapter will review the main generalizations. In particular, we will begin with an overview of uniform sampling using the classical sampling theory for bandlimited signals, and then discuss its generalization to the sampling of signals in shift-invariant spaces, as well as, the sampling of signals with finite rate of innovation whilst touching on some of the mathematical preliminaries that will be used throughout this thesis. Moreover, both of these modern sampling frameworks allow us to analyse more realistic signals, for example the bandlimitedness requirement in the classical framework can be completely eliminated. During our discussion, we will also describe their various extensions to multidimensional sampling and briefly mention some of the works done in nonuniform sampling of signals in shift-invariant spaces.

The second half of the chapter, describes the sampling problem associated with the *sensing of physical fields* using sensor networks. Specifically due to their ubiquity in modelling many real life situations, we first provide a brief history of diffusion fields along with an overview of the research by the signal processing community related to sampling diffusion fields; also, we use a numerical example to illustrate the detrimental effects of a bandlimited reconstruction from the field samples. Finally moving beyond diffusion, we conclude this chapter with a motivation for sampling other types of fields. To this end, we review the literature related to sampling potential and wave fields also.

2.1. Classical sampling theory: uniform sampling of bandlimited signals

Sampling theory aims to provide (theoretical) conditions and (practical) algorithms with which we can pass conveniently between the continuous and discrete worlds. To this end one of the most important works, in this area, entitled "Communications in the presence of noise" was published by Shannon in 1949 [135]. Therein, Shannon stated and provided a concise proof of the following theorem:

Theorem 2.1 (Shannon, 1949 [135])**.** *If a function $f(x)$ contains no frequencies higher than W cycles per second, it is completely determined by giving its ordinates at a series of points spaced $1/2W$ seconds apart.*

Noting that this theorem was "common knowledge in the communication art," he called the uniform interval $\Delta_x = 1/2W$ between the samples the Nyquist interval in recognition of Nyquist's previous contributions to the "communication art" [113, 114]. Shannon also provided an intuitive justification, arguing that if the signal contains no frequencies higher than W, then it is not possible for it to change to a substantially different value in an interval that is less than half its highest frequency, i.e. $1/2W$. Let the sampling frequency be $f_s = 1/\Delta_x$ Hz and denote the n-th sample of the signal by φ_n, such that:

$$\varphi_n = f(x)|_{x=n\Delta_x} \,. \tag{2.1}$$

If $f(t)$ is bandlimited—i.e. its Fourier transform vanishes outside $[-2\pi W, 2\pi W]$—then we can perfectly recover the continuous signal from the sequence $\{\varphi_n\}_{n=-\infty}^{+\infty}$ obtained at the Nyquist rate, using Shannon's reconstruction formula:

$$f(x) = \sum_{n=-\infty}^{+\infty} \varphi_n \operatorname{sinc}\left(\frac{x}{\Delta_x} - n\right), \tag{2.2}$$

where $\operatorname{sinc}(x) \stackrel{\text{def}}{=} \sin(\pi x)/\pi x$.

2.1.1. Non-bandlimited signals, aliasing and antialiasing

Shannon's result requires that the continuous signal be bandlimited. Now, if we decide to proceed with sampling it and then using (2.2) for reconstruction, we will obtain an aliased reconstruction (a numerical simulation which shows the detrimental effects of aliasing for physical fields will be presented later on). The origin of the aliasing errors can be explained diagrammatically by considering the spectrum of the signal as shown in Figure 2.1.

As seen in Figure 2.2, bandlimiting the continuous signal $f(t)$ prior to sampling will truncate its spectrum and so suppress the effect of aliasing, leading to an antialiased reconstruction. This reconstruction however is only a lowpass approximation of $f(x)$. In fact this process (summarized schematically in Figure 2.3) of, prefiltering the continuous signal with

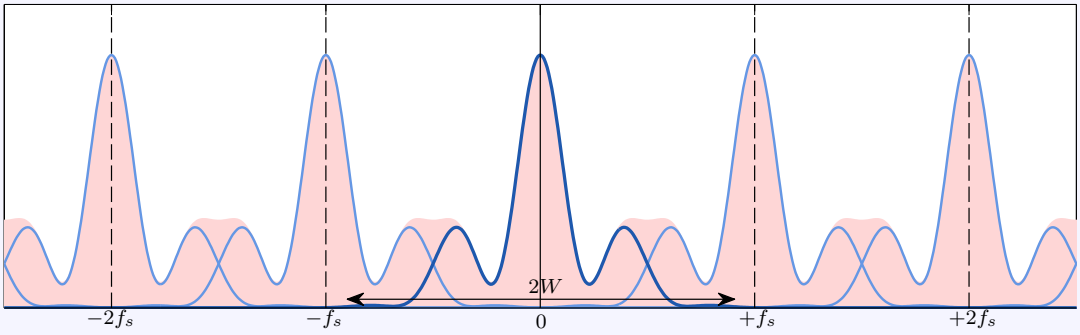


Figure 2.1.: **Spectrum of aliased signal.** The spectrum of the continuous signal is compactly supported on $[-W, W]$ and the sampling frequency is $f_s < 2W$. The spectrum of our sampled signal (shaded region) is a superposition of the original signal's spectrum and shifted copies, shifted by integer multiples of the sampling frequency f_s. Thus, Shannon's formula will return a spectrum that is both aliased and truncated to $[-f_s/2, f_s/2]$; the aliasing errors are the shaded areas above the dark (blue) curve.

the ideal lowpass *antialiasing filter* $h(x) = \text{sinc}\,(x/\Delta_x)$, sampling it at the Nyquist rate and then reconstructing it using (2.2), has been demonstrated to give the "best" approximation of $f(x)$—in a least-squares sense—in the bandlimited subspace.

Whilst the result of classical sampling theory presented so far, has played a crucial role in signal processing and communications, it is not without drawbacks. First, the reconstruction scheme (2.2) is difficult to realize in practice due to the very slow decay of the sinc function [119]. Second, real world signals are never exactly bandlimited [137] hence it is an idealization. Third, although one could attempt to use an *antialiasing filter* to bandlimit the signal as we have seen, such ideal lowpass filters are not realizable in practice. Moreover, it is generally impossible to achieve spatial lowpass filtering, in the case of sensor networks. As a direct consequence of these, Shannon's result is rarely used in practice. Indeed most practitioners resort to the use of simpler techniques such as linear interpolation.

We note that the classical, 1-D Shannon's sampling theorem was extended by Parzen [120], in 1956, to include the sampling of (bandlimited) functions $u(\mathbf{x})$ with d variables, i.e. $\mathbf{x} = (x_1, x_2, \ldots, x_d) \in \mathbb{R}^d$. Specifically, the multidimensional sampling theorem asserts that the continuous function can be reconstructed from its samples $\varphi_{\boldsymbol{n}} = u(\boldsymbol{n}\boldsymbol{\Delta}_{\mathbf{x}}) = u(n_1\Delta_{x_1}, n_2\Delta_{x_2}, \ldots, n_d\Delta_{x_d})$, where $\Delta_{x_i} = 2W_i$, using the formula:

$$f(\mathbf{x}) = \sum_{\boldsymbol{n}\in\mathbb{Z}^d} \varphi_{\boldsymbol{n}} \text{sinc}\left(\frac{\mathbf{x}}{\boldsymbol{\Delta}_{\mathbf{x}}} - \boldsymbol{n}\right), \tag{2.3}$$

where $\text{sinc}(\mathbf{x}) = \prod_{i=1}^{d} \text{sinc}(x_i)$, provided the function is bandlimited to the d-dimensional hypercube centred at the origin and with vertices $\{\pm W_1, \pm W_2, \ldots, \pm W_d\}$. Note that, with

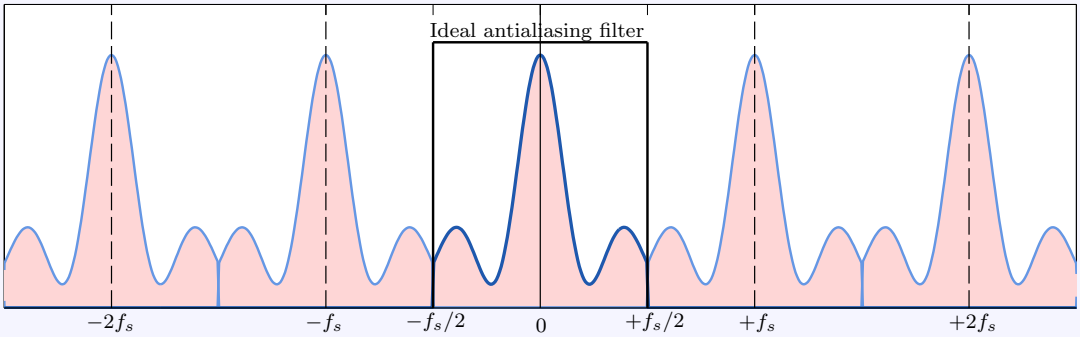


Figure 2.2.: **Spectrum of antialiased signal.** Prefiltering with an ideal antialiasing filter means the spectrum is now supported on $[-f_s/2, f_s/2]$ as such there is no overlap in the shifted copies obtained due to sampling, Shannon's formula will return the truncated but non-aliased baseband spectrum.

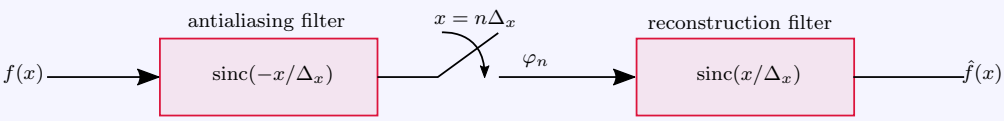


Figure 2.3.: **Schematic of practical sampling process.** The continuous signal is prefiltered, sampled and then postfiltered to obtain the optimal least-squares approximation of $f(x)$ in the space of bandlimited functions.

a slight abuse of notation, we take $\frac{\mathbf{x}}{\mathbf{\Delta_x}} = \left(\frac{x_1}{\Delta_{x_1}}, \frac{x_2}{\Delta_{x_2}}, \ldots, \frac{x_d}{\Delta_{x_d}}\right)$.

Furthermore, in 1977 Papoulis [118] generalized the sampling theorem to multichannel sampling, demonstrating that a bandlimited signal can be reconstructed from samples of response of m linear shift-invariant systems sampled at $1/m$ the reconstruction rate. This framework has been extended to multidimensional signals [33] also. Although Papoulis's scheme is conceptually straightforward, it has been shown that it is not always possible due to potential instabilities [32]. For a more comprehensive review of sampling theory up to 1977, we refer to the paper of Jerri's [77].

2.2. Modern sampling theory: sampling in shift-invariant spaces

Shannon's sampling theorem admits a more modern Hilbert space $\mathcal{L}_2$ interpretation [143], which views the sampling and interpolation process as a projection from the Hilbert space to the space of bandlimited functions – i.e. the space $V(\mathrm{sinc}) = \mathrm{span}\left(\{\mathrm{sinc}(\mathbf{x} - \boldsymbol{n})\}_{\boldsymbol{n}\in\mathbb{Z}^d}\right)$ spanned by uniform translates of the sinc function. It is this interpretation that allows us to see the classical sampling process as a special case of the so called: *sampling in shift-invariant spaces*. We provide a review of the main ideas of sampling in shift-invariant

spaces (SIS), our presentation is based on the review papers of Unser [143], as well as, Aldroubi and Gröchenig [8].

2.2.1. Uniform and non-uniform sampling in shift-invariant spaces

Sampling in shift-invariant spaces retains some of the basic flavours of the classical theory but generalizes the basis functions by replacing the sinc function with an arbitrary generator function $g(\mathbf{x})$. Consequently when we assume a uniform sampling setup, the corresponding approximation space below is obtained:

$$V(g) = \left\{ \sum_{\boldsymbol{n}\in\mathbb{Z}^d} w_{\boldsymbol{n}} g(\mathbf{x}-\boldsymbol{n}) \right\}, \tag{2.4}$$

where $w_{\boldsymbol{n}} \in \mathbb{R}$. The validity of this model relies of three conditions. First, for convergence the coefficients $w_{\boldsymbol{n}}$ must be square-summable. Second in order to ensure a stable and unique representation, the family of functions $\{g(\mathbf{x}-\boldsymbol{n})\}$ should form a Riesz basis of $V(g)$, meaning there must exist two constants $0 < a \leq A < +\infty$ such that:

$$a\|w\|_{l_2}^2 \leq \left\| \sum_{\boldsymbol{n}\in\mathbb{Z}^d} w_{\boldsymbol{n}} g(\mathbf{x}-\boldsymbol{n}) \right\|^2 \leq A\|w\|_{l_2}^2, \tag{2.5}$$

where $\|w\|_{l_2}^2 = \sum_{\boldsymbol{n}} |w_{\boldsymbol{n}}|^2$ and $\|f(\mathbf{x})\| = \sqrt{\langle f, f\rangle}$ is the norm induced by the usual $\mathcal{L}_2$-inner product,

$$\langle f, g\rangle = \int_{\mathbf{x}\in\mathbb{R}^d} f(\mathbf{x})g(\mathbf{x})\mathrm{d}\mathbf{x}. \tag{2.6}$$

The final condition, called the *partition of unity* condition:

$$\sum_{\boldsymbol{n}\in\mathbb{Z}^d} g(\mathbf{x}+\boldsymbol{n}) = 1,$$

ensures that the model can represent any function as closely as possible, by selecting sufficiently small sampling intervals. Having established an approximation space, the next logical step is to describe how to obtain the $w_{\boldsymbol{n}}$'s in (2.4) so as to obtain a faithful reconstruction of the original signal. The idea here is that we want to reconstruct $\hat{f}(\mathbf{x}) = \sum_{\boldsymbol{n}} w_{\boldsymbol{n}} g(\mathbf{x}-\boldsymbol{n})$, where $w_{\boldsymbol{n}}$ are the new samples chosen in such a way that $\hat{f}(\mathbf{x})$ is as close as possible to $f(\mathbf{x})$. The most straightforward approach is the so called *minimum error sampling*, which enforces an optimal solution in the least-squares sense via the projection:

$$P_{V(g)} f(\mathbf{x}) = \sum_{\boldsymbol{n}\in\mathbb{Z}^d} \langle f(\mathbf{x}), \tilde{g}(\mathbf{x}-\boldsymbol{n})\rangle g(\mathbf{x}-\boldsymbol{n})$$

where the projection operator $P_{V(g)} : \mathcal{L}_2 \to V(g)$ is the orthogonal projection of the space $\mathcal{L}_2$ onto $V(g)$, and the $\tilde{g}(\mathbf{x}-\boldsymbol{n}) \in V(g)$'s are the unique dual basis functions which can be computed using the biorthogonality condition $\langle \tilde{g}(\mathbf{x}-\boldsymbol{n}), g(\mathbf{x}-\boldsymbol{n}')\rangle = \delta(\boldsymbol{n}-\boldsymbol{n}')$ for

$\boldsymbol{n}, \boldsymbol{n}' \in \mathbb{Z}^d$. In the case where $a = A = 1$, then the generator g is orthonormal and so $\tilde{g}(\mathbf{x}) = g(\mathbf{x})$; an example of this is the sinc function. Other approaches for computing the $w_{\boldsymbol{n}}$'s based on *consistent sampling* [144], *interpolation* and so on also exist and can be found in [143].

On the other hand, the problem of non-uniform sampling in SIS, i.e.

$$f(\mathbf{x}) = \sum_{\boldsymbol{n} \in \mathbb{Z}^d} w_{\boldsymbol{n}} g(\mathbf{x} - \mathbf{x}_{\boldsymbol{n}}) \tag{2.7}$$

is relatively more recent. As usual the sampling problem consists of:

(i) Finding, for some generator g, the conditions on the sampling points $\mathcal{S} = \{\mathbf{x}_{\boldsymbol{n}}\}_{\boldsymbol{n}}$, such that $f(\mathbf{x})$ can be stably and uniquely determined from its samples $\{f(\mathbf{x}_{\boldsymbol{n}})\}_{\boldsymbol{n}}$.

(ii) Designing practical algorithms for realizing the reconstruction.

In this setting however, it is generally very difficult to obtain closed form expressions for the desired coefficients $w_{\boldsymbol{n}}$ given an arbitrary set of sampling points $\mathcal{S}$. As a result, the earliest contributions [45] in the area focussed on perturbation of uniform sampling in SIS, whilst some recent results due to Aldroubi *et al* (see [8] and the references therein) resort to iterative algorithms [7]. In the latter, it is assumed that $\mathcal{S}$ is a so called *set of sampling* for $V(g)$. This means that there must exist two constants $0 < b \leq B < \infty$ that are independent of f, such that:

$$b\|f\|^2 \leq \left(\sum_{\mathbf{x}_{\boldsymbol{n}} \in \mathcal{S}} |f(\mathbf{x}_{\boldsymbol{n}})|^2 \right)^{1/2} \leq B\|f\|^2. \tag{2.8}$$

2.2.2. Sampling finite rate of innovation signals

Finite rate of innovation (FRI) signals [150, 26, 61] are a specific class of signals that extend the bandlimited formulation in (2.2)—to signals that are neither bandlimited nor reside in a fixed subspace—as follows:

$$f(x) = \sum_{m \in \mathbb{Z}} \sum_{r=0}^{R} \gamma_{m,r} g_r(x - \xi_m), \tag{2.9}$$

where the set of functions $\{g_r(x)\}_{r=0}^{R}$ is known a priori. Due to this fact, although nonbandlimited, the signal f is fully specified by the set of amplitudes $\gamma_{m,r}$ and locations ξ_m, which are called the degrees of freedom. Let $C_f(x_a, x_b)$ be a function that counts the degrees of freedom in $f(x)$ over some interval $[x_a, x_b]$, then the *rate of innovation* is defined as follows [150]:

$$\rho = \lim_{L \to \infty} \frac{1}{L} C_f(-\frac{L}{2}, \frac{L}{2}). \tag{2.10}$$

With this, it is now possible to give a precise definition of FRI signals:

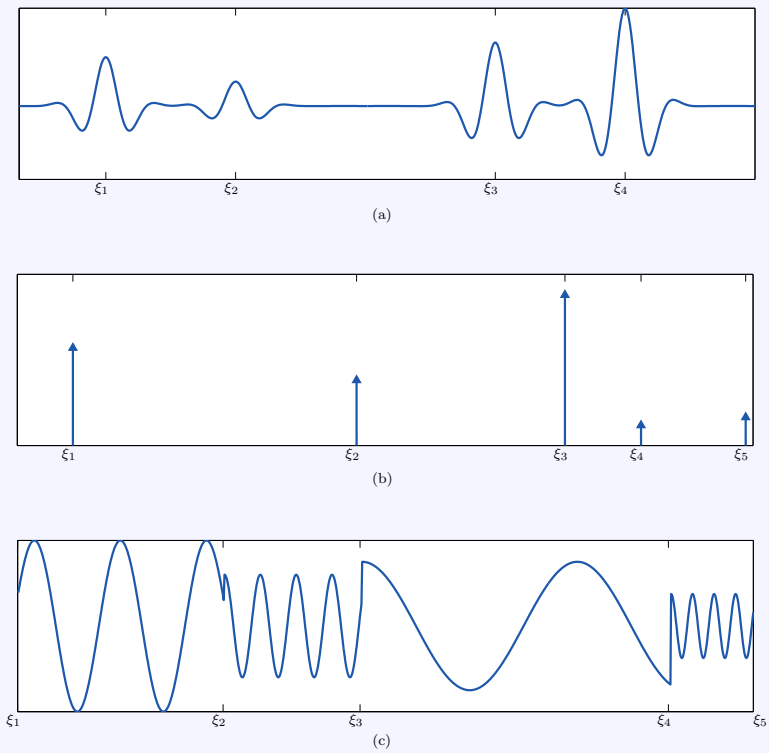


Figure 2.4.: **Examples of FRI signals.** (a) A superposition of four damped and compactly supported sinusoids with different locations and amplitudes, i.e. $f(x) = \sum_{m=1}^{4} c_m g(x - \xi_m)$ where the characteristic pulse shape $g(x) = (1.3 + \cos(x))e^{-\frac{x^2}{30}}(H(x+50) - H(x-50))$ and $H(x)$ is the Heaviside step function. (b) Superposition of five Dirac deltas with different amplitudes and locations, i.e. $f(x) = \sum_{m=1}^{5} c_m \delta(x - \xi_m)$. (c) The superposition of four truncated sinusoids with different amplitudes, frequencies, phases and locations, specifically $f(x) = \sum_{m=1}^{4} c_m \cos(\omega_m x + \phi_m)(H(x - \xi_m) - H(x - \xi_{m+1}))$. When the pulse shapes are known, these signals are fully specified once the corresponding pulse locations and amplitudes are found.

Definition 2.1 (FRI signals [150])**.** *A signal with finite rate of innovation is a signal whose parametric representation is given in* (2.9) *and with a finite* ρ*, as defined in* (2.10)*.*

In certain cases, it may be more convenient to consider a local rate of innovation with respect to some interval of size L. Note also that 1-D bandlimited signals can also be viewed as FRI signals, specifically we can say that their rate of innovation is W since they have as many degrees of freedom per unit of time. We show plots of some common examples of FRI signals in Figure 2.4.

The typical sampling setup for 1-D FRI signals is shown in Figure 2.5. Clearly, using this setup one obtains the samples given by:

$$\varphi_n = \langle f(x), g(x/\Delta_x - n) \rangle , \tag{2.11}$$

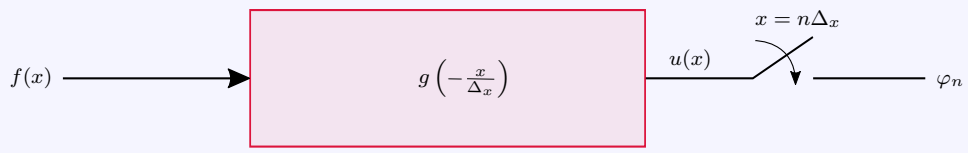


Figure 2.5.: **Schematic of FRI sampling process.** The sampling kernel $g(x)$ now enjoys the freedom to be more general than the sinc function, however but it must have certain function reproducing properties, such as being polynomial and exponential reproducing.

where $g(x)$ is the sampling kernel (which may either be imposed upon or chosen by us) with certain desirable properties. The main sampling question within this FRI framework is two-fold. First, what properties must $g(x)$ posses in other to allow a stable recovery of the parameters (amplitudes and locations) of $f(x)$ from the samples φ_n? And second, what are the (stable) reconstruction algorithms? We provide some brief answers to these questions in what follows, the interested reader may find a more in-depth treatment in [61, 145] and references therein.

There are a few choices of sampling kernels available in the FRI literature. A common choice which we will use to clearly illustrate the power of the FRI framework, in what follows, is the *exponential reproducing kernel*. These exponential reproducing kernels are so called, because they are able to reproduce real or complex exponentials and so satisfy the property:

$$\sum_{n\in\mathbb{Z}} w_{k,n} g\left(x/\Delta_x - n\right) = e^{\mathrm{j}\alpha_k x/\Delta_x}, \tag{2.12}$$

where $w_{k,n}$ are the exponential reproducing coefficients, more on this and how to compute it in Chapter 6. For notational convenience and without loss of generality, we will let $\Delta_x = 1$. Assume now that the samples φ_n are linearly combined using the exact exponential reproducing coefficients $w_{k,n}$ in (2.12), then we obtain:

$$\begin{aligned}\mathcal{R}(k) &= \sum_n w_{k,n} \left\langle f(x), g\left(x-n\right)\right\rangle \\ &= \left\langle f(x), \sum_n w_{k,n} g\left(x-n\right)\right\rangle \\ &= \left\langle f(x), e^{\mathrm{j}\alpha_k x}\right\rangle,\end{aligned} \tag{2.13}$$

where the second equality follows from the linearity of the inner product, whilst the final equality follows from the exponential reproduction property. Notice that $\mathcal{R}(k)$ can be seen as the Fourier transform of $f(x)$ evaluated at $-\alpha_k$. Moreover when $f(x)$ is a stream of M

Diracs, i.e. if $\sum_{m=1}^{M} c_m \delta(x - \xi_m)$, and $\alpha_k = k = 0, 1, \ldots, K$ then (2.13) reduces to:

$$\mathcal{R}(k) = \sum_{m=1}^{M} c_m e^{\mathrm{j}k\xi_m} = \sum_{m=1}^{M} c_m v_m^k, \tag{2.14}$$

where $v_m = e^{\mathrm{j}\xi_m}$. We have therefore seen that the FRI sampling scheme has allowed us to recast the problem of estimating a stream of Diracs as that of estimating the amplitudes and phases of a sum of sinusoids. This parameter estimation problem is well known in spectral estimation [140] and in fact we will frequently encounter it in our proposed approaches for solving the inverse source problem for physical fields, presented in subsequent chapters. We now consider the two main approaches for solving it.

Prony's method: the annihilating filter

The Prony system (2.14) and those which we will encounter in subsequent chapters can be written in the general form:

$$\mathcal{R}(k) = \sum_{m=1}^{M} c_m v_m^k, \tag{2.15}$$

where $c_m, v_m \in \mathbb{C}$ are unknowns. Such a system although linear in the unknown parameters c_m, is nonlinear in the parameters v_m. Hence there is some difficulty associated with finding these nonlinear parameters. Fortunately this problem is well studied and will be solved here by applying Prony's method. A brief overview of the method is given here, for a more in depth treatment see [140].

The method is based on the observation that when the input of a filter having zeros at v_m is the sequence $\mathcal{R}(k)$, then the output will be zero. This filter is called the *annihilating filter*, and has transfer function:

$$A(z) = \sum_{l=0}^{M} a(l) z^{-l} = \prod_{m=1}^{M} (1 - v_m z^{-1}), \tag{2.16}$$

where $a(k)$ is the impulse response of the filter $A(z)$. Specifically,

$$\begin{aligned}
(a * \mathcal{R})(k) &= \sum_{l=0}^{M} a(l) \mathcal{R}(k - l) \\
&= \sum_{l=0}^{M} a(l) \sum_{m=1}^{M} c_m v_m^{k-l} \\
&= \underbrace{\sum_{l=0}^{M} a(l) v_m^{-l}}_{=A(v_m)} \sum_{m=1}^{M} c_m v_m^k = 0,
\end{aligned} \tag{2.17}$$

since $A(z)|_{z=v_m} = 0$. Given the sequence $\mathcal{R}(k)$, the convolution between $(a * \mathcal{R})(k)$ may

be written in the matrix/vector form as $\mathbf{Ra} = \mathbf{0}$, such that:

$$\begin{bmatrix} \vdots & \vdots & \cdots & \vdots \\ \mathcal{R}(M) & \mathcal{R}(M-1) & \cdots & \mathcal{R}(0) \\ \mathcal{R}(M+1) & \mathcal{R}(M) & \cdots & \mathcal{R}(1) \\ \vdots & \vdots & \ddots & \vdots \\ \mathcal{R}(2M) & \mathcal{R}(2M-1) & \cdots & \mathcal{R}(M) \\ \vdots & \vdots & \cdots & \vdots \end{bmatrix} \begin{bmatrix} a(0) \\ a(1) \\ \\ \vdots \\ \\ a(M) \end{bmatrix} = \mathbf{0}. \tag{2.18}$$

The matrix $\mathbf{R}$ is rank deficient with rank M and is therefore overdetermined. Imposing $a(0) = 1$ enforces a unique solution—since there are now M coefficients of the filter to be found—therefore we need at least $2M$ consecutive terms of the sequence $\{\mathcal{R}(k)\}_{k=0}^{K}$; i.e. $K \geq 2M - 1$. Once $\mathbf{a}$ has been found, then the values of v_m are simply the roots of the polynomial $A(z)$. Finally the amplitudes c_m can be determined by simply taking any M equations in (2.15) and solving the corresponding Vandermonde system.

In the presence of model mismatch, (2.18) is no longer satisfied exactly, yet minimizing the Euclidean norm $\|\mathbf{Ra}\|^2$ subject to $\|\mathbf{a}\|^2 = 1$, gives a good estimate for $\mathbf{a}$ [26]. Hence, the Total Least-Squares (TLS) method is used to solve for $\mathbf{a}$, where $\mathbf{a}$ is chosen to be the eigenvector which corresponds to the smallest eigenvalue of the matrix $\mathbf{R}^\top\mathbf{R}$. More details of the TLS method can be found in [26].

Matrix pencil method

The matrix pencil method offers an alternative way to computing $\{c_m, v_m\}_m$. This method is based on subspace techniques for estimating generalized eigenvalues of matrix pencils [75]. The matrix pencil of order J for a set of square matrices $\mathbf{M}_0, \mathbf{M}_1, \ldots, \mathbf{M}_J$, is defined as

$$\mathbf{P}_J(x) = \mathbf{M}_0 + \mathbf{M}_1 x + \cdots + \mathbf{M}_J x^J.$$

Akin to the terminology used for polynomials, the pencil $\mathbf{M}_0 + \mathbf{M}_1 x$ is known as the linear matrix pencil. In order to solve the parameter estimation problem, the matrix pencil method capitalizes on the structure of the Toeplitz matrices constructed from $2M$ consecutive values of the sequence $\mathcal{R}(k)$. Specifically we construct the matrix $\mathbf{R} \in \mathbb{R}^{(M+1)\times M}$,

$$\mathbf{R} = \begin{bmatrix} \mathcal{R}(M-1) & \mathcal{R}(M-2) & \cdots & \mathcal{R}(0) \\ \mathcal{R}(M) & \mathcal{R}(M-1) & \cdots & \mathcal{R}(1) \\ \vdots & \vdots & \ddots & \vdots \\ \mathcal{R}(2M-1) & \mathcal{R}(2M-2) & \cdots & \mathcal{R}(M) \end{bmatrix}. \tag{2.19}$$

Next we form two new matrices $\overline{\mathbf{R}}$ and $\underline{\mathbf{R}}$, by deleting the first and last rows respectively of the Toeplitz matrix $\mathbf{R}$. Then it follows that these two $M \times M$ matrices form the linear matrix pencil $\overline{\mathbf{R}} - \sigma\underline{\mathbf{R}}$ with rank $M - 1$ for $\sigma = v_m$ but is otherwise full rank. Hence the unknowns $\{v_m\}_m$ can be computed by solving the generalized eigenvalue problem

$\overline{\mathbf{R}}\mathbf{v} = \sigma\underline{\mathbf{R}}\mathbf{v}$; by observing that

$$(\overline{\mathbf{R}} - \sigma\underline{\mathbf{R}})\mathbf{v} = \mathbf{0} \Leftrightarrow (\underline{\mathbf{R}}^{-1}\overline{\mathbf{R}} - \sigma\mathbf{I})\mathbf{v} = \mathbf{0},$$

it follows therefore that the v_m's we are after are the eigenvalues of $\underline{\mathbf{R}}^{-1}\overline{\mathbf{R}}$. Note that, the matrix $\underline{\mathbf{R}}$ is square and is also full rank, therefore it is invertible.

2.2.3. Sampling multidimensional FRI signals

The FRI framework has also been properly extended to sample and reconstruct certain parametric 2-D signals [104, 136, 40]. For these classes of two-dimensional FRI signals, it can be shown that finding the innovation parameters reduces to finding the amplitudes and phases of a 2-D coupled sum of exponentials, which is of the general form:

$$\mathcal{R}(\boldsymbol{k}) = \sum_{m=1}^{M} b_m u_m^{k_1} v_m^{k_2}, \tag{2.20}$$

where $\boldsymbol{k} = (k_1, k_2) \in \mathbb{Z}^2$. Given access to the 2-D sequence $\{\mathcal{R}(\boldsymbol{k})\}$ governed by (2.20), obtained using the 2-D FRI sampling principle, the next step is to devise practical algorithms for recovering the unknowns b_m, u_m, v_m from it. In theory one could consider $k_2 = 0$ for example and recover $\{u_m\}_m$ from $\{\mathcal{R}(k_1, 0)\}_{k_1=0}^{2M-1}$ using any of the previously discussed 1-D approaches and then repeat for $k_1 = 0$ to find $\{v_m\}_m$. However, we must then pair the solutions by performing an exhaustive search. To alleviate this costly approach, Vanpoucke *et al* [149] present an efficient algorithm—called the algebraically coupled matrix pencils (ACMP) method—to solve the parameter estimation problem by extending the matrix pencil method.

Algebraically coupled matrix pencils method

We provide a concise outline of the ACMP algorithm, and note that more details can be found [149]. First define the matrix:

$$\mathbf{R}_{i,j} = \begin{bmatrix} \mathcal{R}(i,j) & \mathcal{R}(i,j+1) & \cdots & \mathcal{R}(i,M+1) \\ \mathcal{R}(i+1,j) & \mathcal{R}(i+1,j+1) & \cdots & \mathcal{R}(i+1,M+1) \\ \vdots & \vdots & \ddots & \vdots \\ \mathcal{R}(M+1,j) & \mathcal{R}(M+1,j+1) & \cdots & \mathcal{R}(M+1,M+1) \end{bmatrix}.$$

(1) Construct the so called enhanced matrix $\mathbf{J} \in \mathbb{R}^{M(M+1)\times M(M+1)}$,

$$\mathbf{J} = \begin{bmatrix} \mathbf{R}_{1,1} & \mathbf{R}_{2,1} & \cdots & \mathbf{R}_{M,1} \\ \mathbf{R}_{1,2} & \mathbf{R}_{2,2} & \cdots & \mathbf{R}_{M,2} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{R}_{1,M} & \mathbf{R}_{2,M} & \cdots & \mathbf{R}_{M,M} \end{bmatrix},$$

which requires $2M \times 2M$ terms of the 2-D sequence to construct.

(2) Construct the following four sub-matrices of $\mathbf{J}$, by row-column deletion as follows:

- $\mathbf{J}_{\text{tl}} = \underline{\mathbf{J}|}$, i.e. delete the last row and last column of $\mathbf{J}$.
- $\mathbf{J}_{\text{tr}} = \underline{|\mathbf{J}}$, i.e. delete the last row and first column of $\mathbf{J}$.
- $\mathbf{J}_{\text{bl}} = \overline{\mathbf{J}|}$, i.e. delete the first row and last column of $\mathbf{J}$.

Theses matrices are related by the linear matrix pencils $\mathbf{J}_{\text{tr}} - \sigma\mathbf{J}_{\text{tl}}$ and $\mathbf{J}_{\text{bl}} - \lambda\mathbf{J}_{\text{tl}}$.

(3) Factorize the matrix $\mathbf{J}_{\text{tl}}$ using the singular value decomposition (SVD) to get,

$$\mathbf{J}_{\text{tl}} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^{\mathsf{H}}. \tag{2.21}$$

Moreover by using the matrices $\mathbf{U}$ and $\mathbf{U}$ from this decomposition, one can couple the pencils of $\mathbf{J}$ above, i.e. $\mathbf{J}_{\text{tr}} - \sigma\mathbf{J}_{\text{tl}}$ and $\mathbf{J}_{\text{bl}} - \lambda\mathbf{J}_{\text{tl}}$, as follows:

$$\begin{aligned} \mathbf{U}^{\mathsf{H}}\left(\mathbf{J}_{\text{tr}} - \sigma\mathbf{J}_{\text{tl}}\right)\mathbf{V} &= \mathbf{H}_{\text{tr}} - \sigma\mathbf{H}_{\text{tl}} \quad \text{and} \\ \mathbf{U}^{\mathsf{H}}\left(\mathbf{J}_{\text{bl}} - \lambda\mathbf{J}_{\text{tl}}\right)\mathbf{V} &= \mathbf{H}_{\text{bl}} - \lambda\mathbf{H}_{\text{tl}}. \end{aligned}$$

(4) From the decomposition and the new linear matrix pencils obtained above, compute $\mathbf{H}_{\text{tl}}$, $\mathbf{H}_{\text{tr}}$ and $\mathbf{H}_{\text{bl}}$ by evaluating:

- $\mathbf{H}_{\text{tl}} = \mathbf{\Sigma}$.
- $\mathbf{H}_{\text{tr}} = \mathbf{U}^{\mathsf{H}}\mathbf{J}_{\text{tr}}\mathbf{V}$.
- $\mathbf{H}_{\text{bl}} = \mathbf{U}^{\mathsf{H}}\mathbf{J}_{\text{bl}}\mathbf{V}$.

(5) Finally compute the poles, i.e. the M pairs $\{(u_m, v_m)\}_m$ by computing the eigenvalues of $\mathbf{H}_{\text{tl}}^{-1}\mathbf{H}_{\text{tr}}$ and $\mathbf{H}_{\text{tl}}^{-1}\mathbf{H}_{\text{bl}}$ respectively.

Once the unknowns $\{(u_m, v_m)\}_m$ have been found, b_m can be retrieved by solving the linear system formulated using (2.20).

Remark 2.1. *In subsequent chapters of this thesis, particularly in Chapters 3 and 4, we will sometimes encounter special situations where u_m and v_m in* (2.20) *are exponentials. As an example, consider the particular scenario where $u_m = e^{\mathrm{j}\xi_{1,m}}$ and $v_m = e^{\mathrm{j}\xi_{2,m}}$ so that our unknowns are the M-pairs $(\xi_{1,m}, \xi_{2,m})$. In this setup, notice that by setting $\boldsymbol{k} = (k_1, k_2) = (\mathrm{j}k, -k)$ we can obtain the following scalar indexed sequence*

$$\mathcal{R}(k) = \mathcal{R}(\mathrm{j}k, -k) = \sum_{m=1}^{M} b_m e^{-k(\xi_{1,m} + \mathrm{j}\xi_{2,m})} = \sum_{m=1}^{M} b_m \tilde{v}_m^k, \tag{2.22}$$

where $\tilde{v}_m = e^{-(\xi_{1,m} + \mathrm{j}\xi_{2,m})}$. The choice $\boldsymbol{k} = (\mathrm{j}k, -k)$ allows us to turn the 2-D problem into a coupled one-dimensional Prony system. Thus we can recover the unknowns $\{(\xi_{1,m}, \xi_{2,m})\}_m$ from $\{\mathcal{R}(k)\}_{k=0}^{K}$ by using either Prony's or matrix pencil method as described in Section 2.2.2, provided $K \geq 2M - 1$. Specifically, Prony's or matrix pencil gives $\tilde{v}_m = e^{-(\xi_{1,m} + \mathrm{j}\xi_{2,m})}$, from this we can immediately conclude that $\xi_{1,m} = -\Re(\log(v_m))$ and

$\xi_{2,m} = -\Im(\log(v_m))$, *where* $\Re(z)$ *and* $\Im(z)$ *are used to denote the real and imaginary parts of a complex number* z*, respectively.*

2.3. Sensing physical fields: diffusion and beyond

Due to several significant advances over the last few decades in the fields of (wireless) networking, communications and in the fabrication of microprocessors, the use of sensor networks (SNs) for sensing and monitoring physical fields has become a fast-growing area of research. During this period many aspects of SNs have been explored and developed [2, 87, 71, 162], alongside, a myriad of interesting applications in localization, tracking and parameter estimation have also been considered [102, 28, 96]. Typically, these SNs comprise of many cheap and low-powered nodes—capable of performing both sensing, communication and inference tasks—deployed over a region of interest. The sensor nodes obtain spatiotemporal samples of physical fields over the region of interest. There are several natural mechanisms that govern the propagation of physical fields through space and time. In most situations, these mechanisms typically involve the transportation of matter/particles or the transportation of energy, from one point to another and can be described by well-known partial differential equations. In the example of a factory leakage mentioned previously, the emitted substance comprises of tiny microscopic particles which, over time, propagate from the source thereby spreading throughout the factory. Therein, the mode of transport is well-known to be **diffusion** and the associated diffusion field is the concentration of the released substance over time, at each point in space.

Although the work presented in initial chapters of this thesis will focus on devising signal processing tools specifically for analysing diffusion fields and the related inverse source problems thereof, they will provide useful insights for sampling fields beyond diffusion. Notably in the penultimate chapter, we establish a certain link with generalized sampling theory that will enable us to extend these tools to also analyse other types of physical fields **beyond** diffusion. This is an important result since there are phenomena—examples include the propagation of sound and the variation of electric potential—that cannot be adequately described by diffusion.

We motivate our interest in sampling multidimensional physical fields using sensor networks in the following section: specifically, we start by giving a (very) concise tour of the history of diffusion, related problems and applications, as well as, a review of the current state-of-the-art. Finally we move beyond diffusion to some other popular examples of physical fields of interest, whilst providing our motivation thereof, along with an overview of recent research on the sampling of such fields.

2.3.1. Diffusion: from Brown to Fourier

Definition 2.2 (Diffusion)**.** *Diffusion is defined as a stochastic movement of a collection of particles from regions of high concentration to regions of lower concentration (until an*

equilibrium is established).

Specifically this stochastic movement is precisely a Brownian motion, named after its discoverer – botanist Robert Brown. In an experiment, Brown noticed a highly irregular but incremental path in the motion of pollen suspended in a fluid, whilst observing his experimental setup through a microscope and published his findings in 1828 in [34]. This highly irregular motion is due to collisions with other particles within the fluid. Later on in 1855, a partial differential equation—now known as the diffusion equation—that describes the diffusion field (i.e. the variation of the concentration of particles over space and time) was derived. This diffusion equation is a generalization of the heat equation which models the propagation of heat through a medium. It was during the study of heat flow and the heat equation, so as to solve it in the general case, that led Fourier to the discovery of his Fourier series [51]. Since then the forward problem of predicting the diffusion field induced by a source has been well-studied. More recently however, the reverse problem has received considerable research interests and will be the main focus of our work, albeit with a sensor network flavour. In particular, our goal is to be able to extract certain useful information (such as fully reconstructing the original field, or inferring its sources) from its spatiotemporal sensor measurements.

Sampling diffusion fields

In this thesis, we first consider processes governed by the diffusion equation. An efficient and robust sampling and reconstruction strategy for such fields will impact several real-life applications, from the detection of pollution and plume sources [63] in environmental monitoring to controlling the spread of fungal diseases in precision agriculture [90], as well as retracing the sources of biochemical and nuclear wastes and leakages [107, 44, 134]. Furthermore, understanding the distribution of hot and cold spots due to energy inefficiencies in processors [30, 126], as well as, large data centre clusters [23] can lead to better load balancing.

Consequently considerable research efforts, from the signal processing community, have recently concentrated on developing robust sensor data fusion schemes that aim to either, infer the sources inducing the field [59, 125, 124, 99, 100, 91] or to reconstruct directly the field [131, 129, 147]. These fields are typically non-bandlimited and hence require an extremely dense set of samples in order to achieve a faithful recovery using the classical linear bandlimited (BL) reconstruction framework – see [86] and also the 2-D spatial sampling example given in Figure 2.6. Ranieri and Vetterli [123] suggest that in some interesting cases, specifically when the initial field distribution is not important, a BL reconstruction is sufficient since the spatial bandwidth decays exponentially fast with time and frequency. To alleviate the limitations of a BL reconstruction, Reise *et al* [129, 128] propose the use of *hybrid shift-invariant spaces*, since these spaces allow the modeling of smooth non-BL fields without imposing strict band-limitation. They investigate the use of B-splines for static fields and extend their construction to time-varying fields using an iterative procedure. In

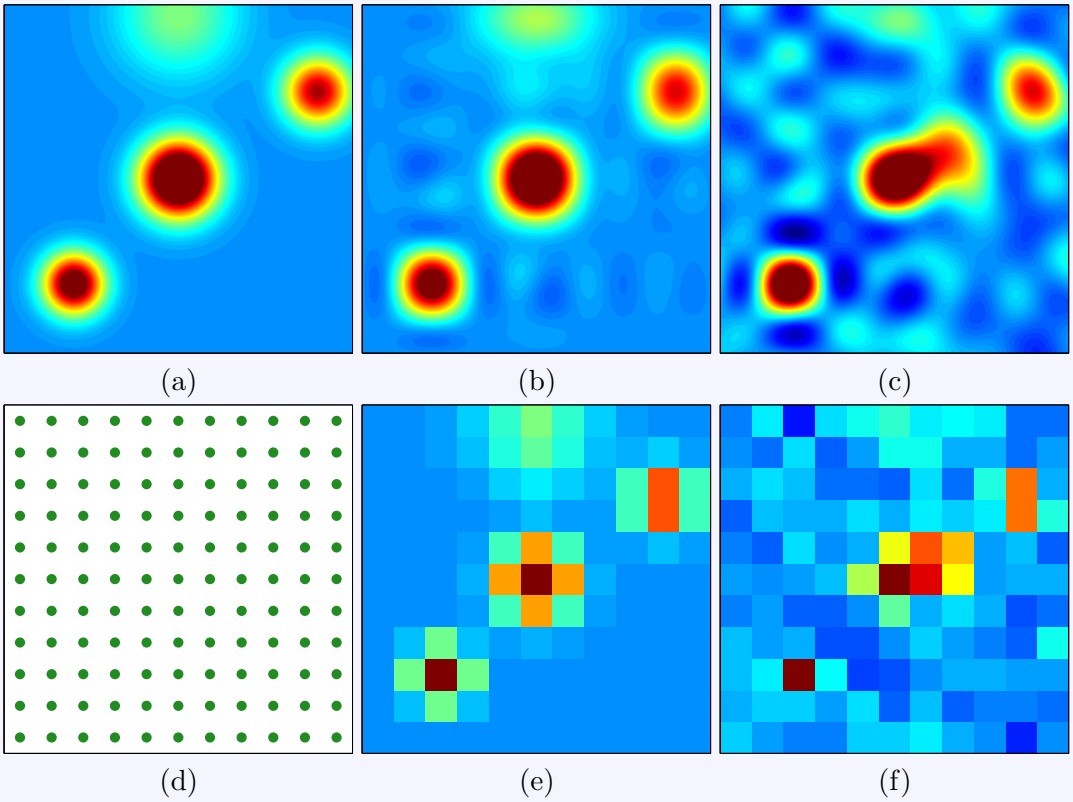


Figure 2.6.: **Estimating the thermal distribution of a two-dimensional plate.** Shown in (a) is the thermal distribution of a 2-D plate heated by four unknown localized and instantaneous sources. Using a network of 121 sensors the field is sensed at locations in (d); assuming ideal sensors the noiseless samples obtained are visualized in (e); whereas the noisy samples in a more realistic situation are visualized in (f). Using these noiseless and noisy samples, we obtain corresponding bandlimited reconstructions presented in (b) and (c) respectively. We notice significant errors in both reconstructions due to aliasing since the field is not bandlimited, also, we cannot apply an antialiasing filter in space.

[126] Ranieri *et al* also propose a subspace-based method for successfully recovering *thermal maps*; in this case, an optimal low-dimensional approximation is used, by first estimating the principal bases—*eigenmaps*—through an experiment carried out at design-time. Techniques based on the use of finite element method (FEM) [65, 29] to solve the static field reconstruction problem have also been researched. For example, van Waterschoot and Leus [147, 148] propose to combine the spatiotemporal samples with the PDE-based field model to achieve static field estimation at certain points of interest. Furthermore, several strategies based on compressive sensing (CS) [47] have been proposed. In particular, Yan *et al* [161] developed a CS-based environmental monitoring algorithm, wherein the sparse basis matrix is constructed from discretized versions of the Gaussian kernel, whereas [131] incorporates the diffusion equation in their work, thereby extending the approach of [124].

For non-static fields however, it is common to first estimate the sources of the field as this allows complete field reconstruction in space and time. Statistical estimation methods, see [112, 154, 163, 132] and the references therein, based on Bayesian estimation and Kalman filtering have been proposed. In addition, Le Niliot *et al* propose to estimate the sources using boundary element methods (BEM) [91] and validate their proposed iterative algorithm through real-life experiments [92]. In [107] Matthes *et al* develop a single source localization algorithm based on continuous concentration measurements of the field. Dokmanic *et al* [59] retrieve the single source parameters by approximating the single source field using a truncated Fourier series, whereas Lu *et al* demonstrate that by solving a set of linear equations the single source parameters can be estimated [99]. In addition, Lu and Vetterli propose two methods for source estimation, namely spatial super-resolution [98] and an adaptive scheme for sources with smooth spatial distributions [97]. We note that existing schemes based on FEM may require the use of dense meshes for a faithful recovery of the field, whilst compressed sensing-based schemes rely on uniform spatial sampling which is often difficult to achieve in practice. A more realistic assumption is a uniform placement of nodes but subjected to some random jitter [127]. Furthermore some of these existing methods make no assumptions on the temporal nature of the sources, and so are more generally applicable, but they may become unstable in the presence of noise or unable to fully reconstruct the entire field in both space and time.

In the SN setup many centralized solutions to this field/source estimation problem, such as [112, 98, 59, 124, 100, 147, 110, 125, 12, 84, 103], have been recently proposed. It is well-known however that centralized estimation strategies over sensor networks can be vulnerable to single point failures, for instance the network becomes inoperational if the fusion center (FC) fails. In addition, communicating with the FC typically involves long range transmission from the sensor nodes which can result in bottlenecks due to contention. To this end, some decentralized and fully distributed algorithms have also been proposed in the literature, where the aim has been to improve the network's robustness to node failures whilst also reducing transmission costs by relying only on local communication between nodes. Lu and Vetterli, for example, propose a distributed adaptive sampling scheme [97], van Waterschoot and Leus [148] develop a distributed scheme based on finite element method. A distributed field reconstruction method using hybrid shift-invariant spaces is proposed in [129], whilst a distributed extension of standard compressed sensing techniques is developed in [133].

These SNs often comprise of sensor nodes that are often battery powered and, as such, must adhere to strict communication and processing constraints for practical viability. However most of the current approaches violate these constraints due to high computational complexity; they also implicitly assume that the communication links between nodes are noiseless and, thus, of infinite capacity. Although a distributed sequential Bayesian estimation method which is suitable under strict power and computational constraints is proposed in [165], it also assumes that messages can be exchanged with infinite precision, i.e. over a noiseless channel. Moreover, a fast distributed detection, localization and esti-

mation algorithm was developed in [6], to solve the single point diffusion source problem, by constructing two non-linear least squares problems (locally and globally) and then solving them using sequential quadratic programming. This approach was further robustified in [5], through the use of the *Huber* loss function.

In our work, for the centralized approach, we first derive novel exact inversion formulas for our class of source distributions, through the use of Green's second identity and a family of sensing functions, to compute generalized field samples. These generalized samples can then be inverted using variations of existing algebraic methods such as Prony's method. Next, we develop a novel and robust reconstruction method for diffusion fields by properly extending these formulas to operate on the spatiotemporal samples of the field.

For the distributed approach, we demonstrate that computing the generalized measurements from the field samples can be distributed using a modification of the distributed gossip algorithms [27] for average consensus, such that each sensor in the network only needs to exchange some *properly modified versions* of its sensor measurements to its neighbouring nodes. This modification is based on the physics of the problem and thus allows each sensor to converge, through these localized interactions with its neighbours, to the desired generalized measurements. Hence the unknown source parameters can then be recovered as usual using Prony's method and its variations.

2.3.2. Beyond diffusion

There are of course a host of other physical phenomena beyond diffusion that can be described by PDEs; ranging from systems in such domains as fluid dynamics and finance to radio communications and tomography. These PDEs and their associated fields have been studied by engineers, mathematicians and physicists alike. However with the development of increasingly sophisticated measurement sensors for monitoring, new signal processing techniques for analysing such data has become a necessity. Consequently, in over the past few decades, the problem of solving these PDE-driven inverse problems—for many interesting applications—has received significant attention from the signal processing community at large. For example:

Potential fields: Poisson's equation

In biomedical engineering, the localization of sources of neuronal activity (also known as brain source imaging (BSI)) from electroencephalographic (EEG) signals [18, 73, 3, 157, 16] is of interest. For this problem the use of Poisson's equation has been ubiquitous, for modelling the brain activity, since this PDE accurately describes the relationship between the measured electrical potentials (i.e. the field) and the current dipoles (i.e. the sources). To solve this EEG-related inverse source problem, several approaches based on least-squares regularization have been proposed. For instance, the authors of [48, 121] formulate and solve a related optimization problem that includes a regularization term (in the objective function) based on the $\mathcal{L}_2$-norm of the signal measurements. These methods generally

lead to a blurred source distribution, and in order to alleviate this issue some authors have proposed sparsity-based regularization methods in [106, 95, 159]. More recently, techniques exploiting the use of cosparsity have been also been explored [110, 4], in these approaches the idea is to incorporate the physical model of the problem when deriving the analysis operator [84].

Furthermore, a number of methods based on Bayesian modelling have also been proposed [4, 70, 158], wherein a probabilistic model is assumed of the data, whilst the measurements and the sources are assumed to be random variables. Beamforming techniques (see [146, 94, 122] and references therein), as well as, subspace-based tools [3, 43] have been proposed for the source localization problem. More in-depth reviews of this EEG-related inverse problem can be found in [73, 18].

Wave and radiating fields: the wave equation

The collection of inverse problems governed by the wave equation continues to receive considerable research interests. This interest is primarily due to its importance in speech recognition [13], acoustic tomography [78], speech and sound enhancement [41], sound/wave source localization [103, 41], to name only a few. In most of these applications: fully reconstructing the wave field, or inferring the sources of the field is often of interest. To this end, several interesting techniques have been proposed. The classical and most common techniques for solving this problem is based on maximum likelihood estimation and beamforming (see [164, 42] and references therein). Recently, Dokmanič demonstrated how to solve the simultaneous localization and mapping problem for sound sources in rooms by making a crucial link between euclidean distance matrices and the acoustic echoes [60, 58]. Kitić *et al* [84] formulated an optimization problem based on the cosparse regularization framework, in order to solve the sound source localization problem; whilst the Finite Difference Time Domain method is used to achieve source localization and signal reconstruction of acoustic pressure fields in [12]. Furthermore, Doğan *et al* present a finite-rate-of-innovation framework to solve the inverse source problem for radiating fields given boundary only measurements [57]. In their approach, the authors also utilize the, so called, *reciprocity gap principle* to provide a relationship between the field measurements and the sources.

2.4. Summary

In this chapter we reviewed the classical sampling theorem for bandlimited signals along with its multidimensional extension. In addition, we also discussed the main ideas of sampling signals in shift-invariant spaces, as well as, the notion of sampling FRI signals, viewing them as generalizations of classical sampling. Finally we described the idea of sensing physical fields as a multidimensional sampling problem and presented an overview of related literature.

Chapter 3.

Inverse source problem for the diffusion equation

In this chapter we will consider an inverse problem for the diffusion equation and present a novel method for solving it. Specifically we will begin with a precise description of the particular inverse problem of interest—the so-called *inverse diffusion source problem*—wherein the goal is to find the source distribution inducing the field measurements. For this problem, we consider an interesting class of source distributions inspired by reality, and then present some novel closed-form solutions for the problem under the assumption that we have access to a set of continuous field measurements in some domain containing the sources. Such an assumption is common in the applied mathematics and physics communities [111, 31, 37, 1, 10, 160]. Although seemingly unusual in the sensor network setting, understanding and tackling the continuous problem can serve as a stepping-stone towards solving the related inverse source problem that arises when we have access only to noisy (discrete) spatiotemporal sensor network measurements – this will be the main focus of Chapter 4.

3.1. Introduction: on inverse problems for the diffusion equation

Inverse problems arising in nature, can broadly be divided into two different categories. Specifically for the diffusion equation, the problem either reduces to an *inverse diffusion source problem* (IDSP) or an *inverse diffusion coefficient problem* (IDCP). To outline the key differences between the IDSPs and the IDCPs, consider the PDE

$$\frac{\partial}{\partial t}u(\mathbf{x},t) = \nabla \cdot (\mu(\mathbf{x},t)\nabla u(\mathbf{x},t)) + f(\mathbf{x},t). \tag{3.1}$$

When the diffusivity $\mu(\mathbf{x},t) = \mu$, is a constant with respect to space and time we obtain the (popular) isotropic diffusion equation in (3.2). In fact, (3.1) is the more general model for the propagation of diffusion fields, and, is most appropriate for situations where the induced field is propagating through a non-isotropic medium.

Definition 3.1 (Inverse diffusion source problem)**.** *Given the diffusion field $u(\mathbf{x},t)$, governed by* (3.1)*, over some domain $\Omega \times [0,T]$ that is compactly supported on $\mathbb{R}^d \times \mathbb{R}_+$ and*

if the diffusivity $\mu(\mathbf{x}, t)$ is known for all $(\mathbf{x}, t) \in \mathbb{R}^d \times \mathbb{R}_+$, the problem of finding the source or inhomogeneous term $f(\mathbf{x}, t)$ compactly supported on $\mathbb{R}^d \times \mathbb{R}_+$ is known as the inverse diffusion source problem.

Definition 3.2 (Inverse diffusion coefficient problem)**.** *Given the diffusion field $u(\mathbf{x}, t)$, governed by* (3.1)*, over some domain $\Omega \times [0, T]$ that is compactly supported on $\mathbb{R}^d \times \mathbb{R}_+$, and given the source term $f(\mathbf{x}, t)$ over $(\mathbf{x}, t) \in \Omega \times [0, T]$, then the problem of finding the diffusivity/transport coefficient term $\mu(\mathbf{x}, t)$ for any $(\mathbf{x}, t) \in \mathbb{R}^d \times \mathbb{R}_+$ is known as the inverse diffusion coefficient problem.*

More details on the IDC problems can be found in [36, 130, 88, 81, 10] and their references. In the rest of this thesis, we will only concern ourselves with IDS problems with the constant diffusivity property, i.e. $\mu(\mathbf{x}, t) = \mu$ for any $\mathbf{x}, t$, since this assumption is sufficient for the sensor network applications which we are interested in.

Motivated by practical applications we consider diffusion fields induced by a finite number of spatially localized and non-localized sources and address the problem of estimating these sources from continuous measurements of the field induced by them. Within this framework, we consider two different time evolutions: the case where the sources are instantaneous, as well as, the case where the sources decay exponentially in time after activation. We derive novel exact inversion formulas for these source distributions of interest, through the use of Green's second identity and a family of spatial and temporal *sensing functions* that provide generalized samples of the field. These generalized samples can then be inverted using variations of existing algebraic methods such as Prony's method [54, 140] in order to exactly recover the unknown source distributions.

Naturally in any mathematical treatment of inverse problems, a useful and fundamental starting point is to, first of all, study and understand its corresponding forward problem. The key reason being that such an exercise can equip us with useful insights about the underlying phenomena which may be exploited when tackling the inverse problem. Therefore we will begin our discourse on the inverse diffusion source problem with a description of the forward problem and then present a well-known approach for solving it, before presenting our treatment of the inverse problem. In particular this will lead us to the main contribution of this chapter, i.e. the derivation of novel analytical solutions for the inverse diffusion source problem relating to an interesting class of sources.

The remainder of this chapter is organized as follows. In Section 3.2 the forward problem is presented along with a method of solving it, called the *Green function method*. Subsequently, in Section 3.3 we present and precisely formulate the inverse source problem (ISP) for the diffusion equation, therein we give a parametric representation along with practical justifications for the source models of interest. In Section 3.4, we present the derivation of our closed-form inversion formulas and present algorithms to summarize our proposed approach for solving the IDSP. In Section 3.5 we present an extension of our framework to the advection-diffusion equation, before concluding the present chapter with a summary in Section 3.6.

3.2. Forward problem

Consider the diffusion field induced by a real-valued space- and time-varying source distribution $f(\mathbf{x}, t)$ embedded in an infinite, homogeneous and isotropic medium such as free space. The induced real-valued field, denoted with $u(\mathbf{x}, t)$, propagates through space $\mathbf{x} \in \mathbb{R}^d$ and time $t \in \mathbb{R}_+$ according to the inhomogeneous diffusion equation:

$$\frac{\partial}{\partial t} u(\mathbf{x}, t) = \mu \nabla^2 u(\mathbf{x}, t) + f(\mathbf{x}, t), \tag{3.2}$$

where $\mu \in \mathbb{R}$ is the diffusivity of the medium through which the field propagates. The solution to this inhomogeneous diffusion equation, as it stands, is not unique. Specifically, notice that we can add to $u(\mathbf{x}, t)$ any solution to the homogeneous diffusion equation,

$$\frac{\partial}{\partial t} \upsilon(\mathbf{x}, t) = \mu \nabla^2 \upsilon(\mathbf{x}, t) \tag{3.3}$$

and still obtain a valid solution to (3.2). To guarantee uniqueness it is necessary to further specify *initial* and/or *boundary conditions* [109].

Consequently in what follows we describe the so called *Green function method*, which is a classical approach for solving the PDE (3.2) subject to certain initial/boundary conditions, in order to obtain the induced field $u(\mathbf{x}, t)$.

3.2.1. Solution to the diffusion equation: Green's function method

Otherwise known as the *method of fundamental solutions*, the Green function method is an approach that allows us to construct solutions to initial/boundary value problems from the so called Green functions/fundamental solutions of the PDE [62, 109, 153]. Qualitatively this means, we can generate the solution to any linear, inhomogeneous PDE with constant coefficients by taking the convolution between its Green's function and the source (or inhomogeneous) term in the PDE.[1] Mathematically,

Definition 3.3 (Green's function)**.** *Is a distribution g that satisfies the distributional equation*

$$\mathcal{W} g = \delta$$

where $\mathcal{W}$ denotes a linear, constant coefficient partial differential operator (for e.g. $\mathcal{W} = \frac{\partial}{\partial t} - \mu \nabla^2$ in (3.3)) and δ is the Dirac delta function situated at the origin.

Let $g(\mathbf{x}, t)$ denote the Green's function for the inhomogeneous diffusion equation (3.2); then for any smooth, compactly supported function $f = f(\mathbf{x}, t)$,

$$\left(\frac{\partial}{\partial t} - \mu \nabla^2 \right) (f * g) = f * \left(\frac{\partial}{\partial t} - \mu \nabla^2 \right) g$$

[1] This is however not true for non-constant coefficient PDEs. Intuitively, for non-constant coefficient equations we would require a different fundamental solution at each point and then proceed to compute their weighted sum, which is no longer a convolution.

$$= f * \mathcal{W}g = f * \delta$$
$$= f.$$

Therefore we can conclude that the PDE (3.2) has the solution:

$$u(\mathbf{x}, t) = (g * f)(\mathbf{x}, t). \tag{3.4}$$

We are now left with finding an explicit expression for $g(\mathbf{x}, t)$.

The Green function for the diffusion equation

Subject to the Sommerfeld radiation condition, i.e. a *quiescent condition at an initial time*

$$u(\mathbf{x}, t)|_{t=0} = \frac{\partial}{\partial t} u(\mathbf{x}, t)\bigg|_{t=0} = 0, \tag{3.5a}$$

and a *convergence condition at infinity*

$$u(\mathbf{x}, t)|_{\|\mathbf{x}\|\to\infty} = 0 \text{ and } \nabla u(\mathbf{x}, t)|_{\|\mathbf{x}\|\to\infty} = \mathbf{0}, \tag{3.5b}$$

it can be shown that the Green function for a d-dimensional field obeying (3.2) is given by [160, 152]:

$$g(\mathbf{x}, t) = \frac{1}{(4\pi\mu t)^{d/2}} e^{-\frac{\|\mathbf{x}\|^2}{4\mu t}} H(t) \tag{3.6}$$

where $H(t)$ denotes the Heaviside unit step function.

In our setting an implication of the Green function method is that, according to (3.4), the entire field $u(\mathbf{x}, t)$ can be perfectly reconstructed provided we know the source distribution $f(\mathbf{x}, t)$ exactly.

3.3. The inverse diffusion source problem

We are interested in solving the inverse diffusion source problem. Specifically we intend to find the source distribution $f(\mathbf{x}, t)$ inducing some known diffusion field $u(\mathbf{x}, t)$ governed by the PDE (3.2). This chapter will treat the case when the field $u(\mathbf{x}, t)$ is known *continuously* over a compactly supported spatiotemporal domain $\Omega \times [0, T]$, where $\Omega \subset \mathbb{R}^d$ and $0 < T < \infty$; whilst the case when the field is known only at a finite number of *discrete* points in $\Omega \times [0, T]$ is deferred to Chapter 4. Henceforth, we will assume that $d = 2$. Moreover, we note that both problems, in their present form, are ill-posed in the Hadamard sense [160, 101, 38, 66], in that—without a suitable structure imposed upon f—we could construct examples where two distinct source distributions (say f_1 and f_2) will lead to the same field measurements on $\Omega \times [0, T]$, hence implying non-uniqueness. Therefore to guarantee uniqueness of solutions in what follows, we impose a *parametric* structured model to the sources of diffusion fields that are of interest to us.

3.3.1. The class of localized sources

The localized source model assumes that the source is largely concentrated at a single infinitesimal point in space. This model is most suitable when each source of the field is several orders of magnitude smaller than the monitored region containing the field. Many typical examples of such situations are encountered in environmental monitoring applications, including the detection of plumes [63], gaseous leakages [112], and nuclear wastes as in the Fukushima disaster [44, 105]. For this class of sources we will be concerned with two temporal distributions,

1. *Instantaneous sources:* under this model the sources of interest are localized in both space and time. It describes a point that is active only for an extremely short burst of time compared to the measurement duration. The parameterization for a superposition of M such sources is given by

$$f(\mathbf{x}, t) = \sum_{m=1}^{M} c_m \delta(\mathbf{x} - \boldsymbol{\xi}_m, t - \tau_m), \tag{3.7}$$

where $c_m, \tau_m \in \mathbb{R}$ are the intensity and activation time of the m-th source respectively and $\boldsymbol{\xi}_m \in \Omega$ is the source location, specifically $\Omega \in \mathbb{R}^2$ and $\boldsymbol{\xi}_m = (\xi_{1,m}, \xi_{2,m})$. As a result of this source parameterization, the IDS problem becomes:

Problem 3.1. *Estimating the parameters* $\{c_m, \boldsymbol{\xi}_m, \tau_m : m = 1, \ldots, M\}$ *from continuous measurements* $u(\mathbf{x}, t)$ *of the diffusion field taken over the domain* $\Omega \times [0, T]$*, and induced by the distribution* (3.7).

2. *Non-instantaneous sources:* for time-varying emissions like factory leakages where, for example, a gas container gradually empties, a suitable model is an exponential decaying emission intensity. This is because leakages are generally modelled using first order ordinary differential equations (ODEs) which usually have an exponential family of solutions, making the decaying exponential function a suitable model for the time evolution of such sources. Given a superposition of M such time-varying point sources, the model imposed on f becomes

$$f(\mathbf{x}, t) = \sum_{m=1}^{M} c_m e^{\alpha_m (t - \tau_m)} \delta(\mathbf{x} - \boldsymbol{\xi}_m) H(t - \tau_m), \tag{3.8}$$

where $\alpha_m < 0$ is called the decay coefficient. For this source model the continuous IDSP is equivalent to the following problem:

Problem 3.2. *Estimating the parameters* $\{\alpha_m, c_m, \boldsymbol{\xi}_m, \tau_m\}_{m=1}^{M}$ *from continuous measurements* $u(\mathbf{x}, t)$ *of the diffusion field taken over the domain* $\Omega \times [0, T]$*, and induced by the distribution* (3.8).

3.3.2. The class of non-localized sources

Whilst the localized model remain ubiquitous in many real world applications, there are certain instances where the spatial support of the source is comparable to (either dimension of the) monitored region. These situations are encountered typically in applications including the monitoring of overheating servers in large data center clusters [23], or of individual cores in multi-core processors [30, 126] due to energy inefficiencies. In these applications a non-localized distribution may be more appropriate; this leads to the following interesting class of non-localized source models:

1. *Straight line source:* we parameterize these as follows

$$f(\mathbf{x}, t) = cL(\mathbf{x})\delta(t - \tau), \tag{3.9}$$

   where $c, \tau \in \mathbb{R}$ are the intensity and activation time respectively, and $L(\mathbf{x}) \subset \Omega$ describes a line coinciding with the position of the straight line source. In this case, clearly the geometry of the source is uniquely defined by its endpoints, i.e. the pair $\boldsymbol{\xi}_1, \boldsymbol{\xi}_2 \in \Omega$, with $\boldsymbol{\xi}_1 = (\xi_{1,1}, \xi_{2,1})$ and $\boldsymbol{\xi}_2 = (\xi_{1,2}, \xi_{2,2})$. We can now summarize the problem of interest:

   Problem 3.3. *Given continuous measurements of the diffusion field, $u(\mathbf{x}, t)$ for all $(\mathbf{x}, t) \in \Omega \times [0, T]$, induced by a line source with the parameterization* (3.9)*, the inverse diffusion source problem is to estimate the source intensity, activation time and endpoints $(c, \tau, \{\boldsymbol{\xi}_1, \boldsymbol{\xi}_2\})$.*

2. *Convex polygonal source:* these are characterized by the following model on f

$$f(\mathbf{x}, t) = cF(\mathbf{x})\delta(t - \tau), \tag{3.10}$$

   where $c, \tau \in \mathbb{R}$ are again the intensity and activation time respectively, and $F(\mathbf{x}) \subset \Omega$ is the region describing the location and shape of the convex polygonal diffusion source. By noticing that a convex polygon $F(\mathbf{x})$ is uniquely specified by its vertices, that is, the collection $\{\boldsymbol{\xi}_1, \boldsymbol{\xi}_2, \ldots, \boldsymbol{\xi}_M\}$, with $\boldsymbol{\xi}_m = (\xi_{1,m}, \xi_{2,m}) \in \Omega$, we can now state precisely the problem of interest as follows:

   Problem 3.4. *Given continuous measurements of the diffusion field, $u(\mathbf{x}, t)$ for all $(\mathbf{x}, t) \in \Omega \times [0, T]$, induced by an M-sided polygonal source with parameterization* (3.10)*, the inverse diffusion source problem is to estimate the source intensity, activation time and endpoints $(c, \tau, \{\boldsymbol{\xi}_m\}_{m=1}^M)$.*

Although the models (3.9) and (3.10) above are for instantaneous sources, it is straightforward to extend these models to the temporally non-instantaneous source case by replacing the time evolution accordingly with a decaying exponential. Specifically, the non-instantaneous line source would be:

$$f(\mathbf{x}, t) = cL(\mathbf{x})e^{\alpha(t-\tau)}H(t - \tau), \tag{3.11}$$

whilst the non-instantaneous (M-sided) polygonal source is modelled as:

$$f(\mathbf{x}, t) = cF(\mathbf{x})e^{\alpha(t-\tau)}H(t-\tau), \tag{3.12}$$

where $c, \alpha, \tau \in \mathbb{R}$ and $L(\mathbf{x}), F(\mathbf{x}) \subset \Omega$ take their usual definitions.

3.4. Closed-form inversion of a class of inverse diffusion source problems

In this section, we derive exact closed-form inversion formulas for diffusion fields induced by localized and non-localized sources as defined in (3.7) to (3.12). For the localized source models, we use Green's second identity to demonstrate that given access to generalized measurements of the form:

$$\mathcal{Q}(k, r) = \langle \Psi_k(\mathbf{x})\Gamma_r(t), f(\mathbf{x}, t)\rangle = \int_\Omega \int_t \Psi_k(\mathbf{x})\Gamma_r(t) f(\mathbf{x}, t)\mathrm{d}t\mathrm{d}V, \tag{3.13}$$

it is possible to uniquely determine the unknown source parameters in $f(\mathbf{x}, t)$. Here, V is the variable of integration performed over Ω (a surface in 2D), $\Psi_k(\mathbf{x})$ and $\Gamma_r(t)$ are properly chosen *sensing functions*. Specifically, we show that the generalized measurements $\mathcal{Q}(k, r)$ for $k, r \in \mathbb{Z}$ are given by a weighted-sum of complex exponentials. In fact, given a sum of exponentials of the form $\mathcal{Q}(k, r) = \sum_{m=1}^{M} c_m e^{-k(\xi_{1,m}+\mathrm{j}\xi_{2,m})} e^{-\mathrm{j}r\tau_m/T}$, where $\mathrm{j} = \sqrt{-1}$, we then demonstrate how to map uniquely the weights and exponents of this sum to the unknown source parameters in $f(\mathbf{x}, t)$ using Prony's method. This method is frequently encountered in spectral estimation [140] and in the finite rate of innovation (FRI) framework [150, 61, 26, 115]. For completeness, a brief overview was provided in Section 2.2.2.

The use of Green's second identity here allows us to relate, in a simple yet precise way, the boundary and interior measurements of the field, to the sources inducing the field. This is the basis of the *reciprocity gap* method [11] used in non-destructive testing of solids [11, 17]; it has also been exploited for the identification of heat sources from boundary measurements [14] and for estimating the sources of static fields governed by *Poisson's equation* in [80]. Herein we propose an extension of the reciprocity gap method to the identification of instantaneous and non-instantaneous localized sources of diffusion in both space and time, whilst also exploiting the use of stable sensing functions. In addition, we further extend this approach to non-localized source fields and solve the associated IDSP through a simple transformation of the generalized measurements $\mathcal{Q}(k, r)$. This transformation is obtained precisely using tools from the field of complex analysis.

Although the inversion formulas derived herein are based on continuous full-field measurements, which are generally inaccessible in practice, they provide insights on how to combine the discrete spatiotemporal measurements that we obtain using a sensor network in order to compute the generalized sequence $\mathcal{Q}(k, r)$, or at least an approximation of it, which then allows for source recovery using Prony's method.

3.4.1. Diffusion fields of multiple instantaneous point sources

We begin by relating the diffusion field $u(\mathbf{x}, t)$ inside Ω to the source parameters. Let Ψ_k be a twice differentiable function in Ω, then Green's second identity relates the boundary integral and the integral over the bounded region as follows:

$$\oint_{\partial\Omega} (\Psi_k \nabla u - u \nabla \Psi_k) \cdot \hat{\mathbf{n}}_{\partial\Omega} \mathrm{d}S = \int_{\Omega} \left(\Psi_k \nabla^2 u - u \nabla^2 \Psi_k\right) \mathrm{d}V, \tag{3.14}$$

where $\hat{\mathbf{n}}_{\partial\Omega}$ is the outward pointing unit normal vector to the boundary $\partial\Omega$ of Ω. Moreover, if Ψ_k satisfies

$$\frac{\partial \Psi_k}{\partial t} + \mu \nabla^2 \Psi_k = 0 \tag{3.15}$$

in Ω, then given that $u(\mathbf{x}, t)$ satisfies (3.2) we may substitute $\nabla^2 \Psi_k = -\frac{1}{\mu}\frac{\partial \Psi_k}{\partial t}$ and $\nabla^2 u = \frac{1}{\mu}\left(\frac{\partial u}{\partial t} - f\right)$ into the right hand side (RHS) of (3.14) to obtain:

$$\begin{aligned}
\oint_{\partial\Omega} (\Psi_k \nabla u - u \nabla \Psi_k) \cdot \hat{\mathbf{n}}_{\partial\Omega} \mathrm{d}S &= \frac{1}{\mu} \int_{\Omega} \Psi_k \left(\frac{\partial u}{\partial t} - f\right) + u \frac{\partial \Psi_k}{\partial t} \mathrm{d}V \\
&= \frac{1}{\mu} \int_{\Omega} \Psi_k \frac{\partial u}{\partial t} + u \frac{\partial \Psi_k}{\partial t} - \Psi_k f \mathrm{d}V \\
&= \frac{1}{\mu} \int_{\Omega} \frac{\partial}{\partial t}(u \Psi_k) - \Psi_k f \, \mathrm{d}V.
\end{aligned}$$

Finally multiplying through by μ and rearranging yields:

$$\int_{\Omega} \frac{\partial}{\partial t}(u \Psi_k) \mathrm{d}V - \mu \oint_{\partial\Omega} (\Psi_k \nabla u - u \nabla \Psi_k) \cdot \hat{\mathbf{n}}_{\partial\Omega} \mathrm{d}S = \int_{\Omega} \Psi_k f \mathrm{d}V. \tag{3.16}$$

This integral equation gives a simple yet interesting relationship between the source f and the induced field. Specifically, it is interesting to us because of the similarity between the right hand side (RHS) of (3.16) with the spatial integral in the RHS of (3.13). Given the source model (3.7), we can now establish how this expression can be used to recover the unknown source parameters:

Proposition 3.1. *For the instantaneous source parameterization* (3.7)*, providing* $\Psi_k(\mathbf{x})$ *is analytic and of the form* $e^{-k(x_1 + \mathrm{j}x_2)}$ $k = 0, 1, \ldots, K$ *with* $K \geq 2M - 1$ *and* $\Gamma_r(t) = e^{-\mathrm{j}rt/T}$, $r = 0, 1, \ldots, R$ *with* $R \geq 1$*, then the integral equation in* (3.16) *can be used to recover jointly the intensities, locations and activation times of the* M *instantaneous sources.*

Proof. Recall (3.16) and, for conciseness, denote its left hand side (LHS) by $Q_k(t)$; hence it follows that $Q_k(t) = \int_{\Omega} \Psi_k(\mathbf{x}) f(\mathbf{x}, t) \, \mathrm{d}V$. This identity holds true for any t, as such we can multiply both sides by some arbitrarily chosen window $\Gamma_r(t)$. Hence,

$$\Gamma_r(t) Q_k(t) = \Gamma_r(t) \int_{\Omega} \Psi_k(\mathbf{x}) f(\mathbf{x}, t) \, \mathrm{d}V, \quad \forall t.$$

Integrating this new expression over $t \in [0, T]$ yields:

$$\int_{t=0}^{T} \Gamma_r(t) Q_k(t)\,\mathrm{d}t = \int_{t=0}^{T} \Gamma_r(t) \int_{\Omega} \Psi_k(\mathbf{x}) f(\mathbf{x}, t)\,\mathrm{d}V\,\mathrm{d}t = \langle \Psi_k(\mathbf{x})\Gamma_r(t), f(\mathbf{x}, t)\rangle\,,$$

which is precisely the aforementioned generalized measurements (3.13) we seek.

Next we show that this generalized measurement relates the unknown source parameters to the known field measurements. To see this let $f(\mathbf{x}, t)$ be the instantaneous point source model (3.7), then

$$\begin{aligned}
\int_{t=0}^{T} \Gamma_r(t) Q_k(t)\mathrm{d}t &= \langle \Psi_k(\mathbf{x})\Gamma_r(t), f(\mathbf{x}, t)\rangle \\
&= \int_{t=0}^{T} \Gamma_r(t) \int_{\Omega} \Psi_k(\mathbf{x}) \sum_{m=1}^{M} c_m \delta(\mathbf{x} - \boldsymbol{\xi}_m, t - \tau_m)\,\mathrm{d}V\,\mathrm{d}t \\
&= \sum_{m=1}^{M} c_m \int_{t=0}^{T} \Gamma_r(t)\delta(t - \tau_m)\,\mathrm{d}t \int_{\Omega} \Psi_k(\mathbf{x})\delta(\mathbf{x} - \boldsymbol{\xi}_m)\mathrm{d}V \\
&= \sum_{m=1}^{M} c_m \Gamma_r(\tau_m)\Psi_k(\boldsymbol{\xi}_m).
\end{aligned}$$

Notice that because $Q_k(t)$ has been defined to be the LHS of (3.16), the LHS of the equation above is also equivalent to:[2]

$$\begin{aligned}
\int_{t=0}^{T} \Gamma_r(t) Q_k(t)\,\mathrm{d}t = &\int_{t=0}^{T} \Gamma_r(t) \int_{\Omega} \frac{\partial}{\partial t}(u(\mathbf{x}, t)\Psi_k)\,\mathrm{d}V\,\mathrm{d}t \\
&- \mu \int_{t=0}^{T} \Gamma_r(t) \oint_{\partial\Omega} (\Psi_k \nabla u - u \nabla \Psi_k) \cdot \hat{\mathbf{n}}_{\partial\Omega}\,\mathrm{d}S\,\mathrm{d}t.
\end{aligned}$$

Given the field measurement $u(\mathbf{x}, t)$ along with a proper choice of spatial and temporal sensing functions $\Psi_k(\mathbf{x})$ and $\Gamma_r(t)$, respectively, this integral can be evaluated exactly to produce an expression dependent only on k and r, for this reason we will denote it by $\mathcal{Q}(k, r)$. Specifically:

$$\begin{aligned}
\mathcal{Q}(k, r) \stackrel{\text{def}}{=} &\int_{t=0}^{T} \Gamma_r(t) \int_{\Omega} \frac{\partial}{\partial t}(u(\mathbf{x}, t)\Psi_k)\,\mathrm{d}V\,\mathrm{d}t \\
&- \mu \int_{t=0}^{T} \Gamma_r(t) \oint_{\partial\Omega} (\Psi_k \nabla u - u \nabla \Psi_k) \cdot \hat{\mathbf{n}}_{\partial\Omega}\,\mathrm{d}S\mathrm{d}t,
\end{aligned} \tag{3.17}$$

for $k = 0, 1, \ldots, K$ and $r = 0, 1, \ldots, R$. Therefore we have insofar shown that the sought after generalized measurement $\mathcal{Q}(k, r)$ is precisely the inner product $\langle \Psi_k(\mathbf{x})\Gamma_r(t), f(\mathbf{x}, t)\rangle$, which can be evaluated exactly using (3.17) above. Consequently, this means that $\mathcal{Q}(k, r)$

[2] To see this we simply substitute the expression $Q_k(t) \stackrel{\text{def}}{=} \int_{\Omega} \frac{\partial}{\partial t}(u\Psi_k)\mathrm{d}V - \mu \oint_{\partial\Omega} (\Psi_k \nabla u - u \nabla \Psi_k) \cdot \hat{\mathbf{n}}_{\partial\Omega}\mathrm{d}S$ into $\int_{t=0}^{T} \Gamma_r(t) Q_k(t)\,\mathrm{d}t$

is also related to the unknown source terms through,

$$\mathcal{Q}(k,r) = \langle \Psi_k(\mathbf{x})\Gamma_r(t), f(\mathbf{x},t)\rangle = \sum_{m=1}^{M} c_m \Gamma_r(\tau_m)\Psi_k(\boldsymbol{\xi}_m),$$

hence substituting the expressions $\Psi_k(\mathbf{x}) = e^{-k(x_1+\mathrm{j}x_2)}$ and $\Gamma_r(t) = e^{-\mathrm{j}rt/T}$ into the above immediately produces the Vandermonde system:

$$\mathcal{Q}(k,r) = \sum_{m=1}^{M} c_m e^{-k(\xi_{1,m}+\mathrm{j}\xi_{2,m})} e^{-\mathrm{j}r\tau_m/T}. \tag{3.18}$$

Now (3.18) allows us to uniquely and simultaneously retrieve c_m, τ_m and $\boldsymbol{\xi}_m$. Specifically, for joint location, intensity and activation time recovery given M instantaneous sources, notice that the sequence obtained by setting $r = 1$ in (3.18) for any $k \in \mathbb{N}$ is governed by the following Vandermonde system:

$$\mathcal{Q}(k,1) = \sum_{m=1}^{M} c_m e^{-\mathrm{j}\tau_m/T} e^{-k(\xi_{1,m}+\mathrm{j}\xi_{2,m})}. \tag{3.19}$$

where the exact values making up the sequence $\{\mathcal{Q}(k,1)\}_{k=0}^{K}$ are evaluated using (3.18) with $r = 1$.

The sequence $\{\mathcal{Q}(k,1)\}_{k=0}^{K}$ in (3.19) above is governed by a weighted sum of a finite number of complex exponentials and so we can use Prony's method to retrieve uniquely the pairs

$$\{c_m e^{-\mathrm{j}\tau_m/T}, \boldsymbol{\xi}_m\}_{m=1}^{M}$$

from the sequence, provided $K \geq 2M - 1$ (as outlined in the overview Section 2.2.3, Remark 2.1). Then given the pair $\{c_m e^{-\mathrm{j}\tau_m/T}, \boldsymbol{\xi}_m\}$ it is straightforward to retrieve c_m and τ_m as the absolute value and angle (multiplied by $-T$) of $c_m e^{-\mathrm{j}\tau_m/T}$ respectively. □

Remark 3.1. *The choice of $\Psi_k(\mathbf{x})$ and $\Gamma_r(t)$ here is important. Firstly, $\Psi_k(\mathbf{x})$ has to satisfy* (3.15) *in order to obtain* (3.16)*. This is why we pick $\Psi_k(\mathbf{x})$ to be analytic. Amongst the class of analytic functions, we choose $\Psi_k(\mathbf{x})$ to be the damped complex exponential for numerical stability. Similarly, whilst $\Gamma_r(t)$ can be any arbitrary function of time, again for stability reasons, we choose the exponential function with purely imaginary exponent.*

3.4.2. Diffusion fields of multiple non-instantaneous point sources

In what follows, we consider Problem 3.2 and solve it by extending the approach in Proposition 3.1. Specifically we derive closed-form expressions for simultaneous recovery of all unknown non-instantaneous localized source parameters.

Exact recovery of source locations

It was shown in Section 3.4.1 that the M triples $\{c_m, \tau_m, \boldsymbol{\xi}_m\}_{m=1}^M$ can be recovered when the diffusion field has been induced by a superposition of M localized and instantaneous sources. In this section however, we will be concerned with the IDSP for fields of non-instantaneous point sources. Indeed under the source model f in (3.8) we show that the Prony's system is preserved. As such the localization step (discussed in Section 3.4.1) can still reveal the locations, along with corresponding coefficients that we will refer to as the *generalized energies* of the non-instantaneous sources. The following proposition presents this new result.

Proposition 3.2. *For non-instantaneous source fields, with source parameterization* (3.8)*, providing* $\Psi_k(\mathbf{x})$ *is analytic, and is chosen such that* $\Psi_k(\mathbf{x}) = e^{-k(x_1+\mathrm{j}x_2)}$*, then the integral equation in* (3.16) *is governed by the following Vandermonde system:*

$$\mathcal{R}(k) = \sum_{m=1}^{M} c'_m e^{-k(\xi_{1,m}+\mathrm{j}\xi_{2,m})},\ k = 0, 1, \ldots, K \tag{3.20}$$

where $\mathcal{R}(k) = \mathcal{Q}(k, 0)$ *is used to denote the family of definite integrals* (3.17) *for* $k \in \mathbb{N}$ *and* $r = 0$*, whilst* $c'_m = \frac{c_m}{\alpha_m}\left(e^{\alpha_m(T-\tau_m)} - 1\right)$ *is the generalized energy of the* m*-th source.*

Proof. Firstly substitute $\Psi_k = e^{-k(x_1+\mathrm{j}x_2)}$ into (3.16) and integrate both sides of the resulting equation over $t \in [0, T]$, to obtain:

$$\int_\Omega (\Psi_k u)(\mathbf{x}, T)\mathrm{d}V - \mu \oint_{\partial\Omega} (\Psi_k(\mathbf{x})\nabla U(\mathbf{x}) - U(\mathbf{x})\nabla\Psi_k(\mathbf{x}))\cdot\hat{\mathbf{n}}_{\partial\Omega}\mathrm{d}S = \int_0^T \int_\Omega \Psi_k(\mathbf{x}) f(\mathbf{x}, t)\,\mathrm{d}V\mathrm{d}t, \tag{3.21}$$

where as before we denote $U(\mathbf{x}) = \int_0^T u(\mathbf{x}, t)\mathrm{d}t$. Notice that the left hand side of (3.21) coincides with the integral family $\mathcal{Q}(k, 0)$; we will henceforth denote it by $\mathcal{R}(k)$ for brevity. However, given the localized and non-instantaneous source parameterization (3.8) for f, the power-sum series for $\mathcal{R}(k)$ is different to that obtained for instantaneous sources, but can be easily obtained by substituting (3.8) into the right hand side of (3.21) as follows:

$$\begin{aligned}
\mathcal{R}(k) &= \int_{t=0}^T \int_\Omega \Psi_k(\mathbf{x}) f(\mathbf{x}, t)\mathrm{d}V\mathrm{d}t \\
&= \int_{\tau_m}^T \int_\Omega \Psi_k(\mathbf{x}) \sum_{m=1}^M c_m e^{\alpha_m(t-\tau_m)}\delta(\mathbf{x} - \boldsymbol{\xi}_m)\,\mathrm{d}V\mathrm{d}t \\
&= \sum_{m=1}^M c_m \int_\Omega \Psi_k(\mathbf{x})\delta(\mathbf{x} - \boldsymbol{\xi}_m)\,\mathrm{d}V \int_{\tau_m}^T e^{\alpha_m(t-\tau_m)}\mathrm{d}t \\
&= \sum_{m=1}^M c_m \Psi_k(\boldsymbol{\xi}_m)\left[\frac{1}{\alpha_m} e^{\alpha_m(t-\tau_m)}\right]_{\tau_m}^T \\
&= \sum_{m=1}^M \frac{c_m}{\alpha_m}\left(e^{\alpha_m(T-\tau_m)} - 1\right)\Psi_k(\boldsymbol{\xi}_m)
\end{aligned}$$

$$= \sum_{m=1}^{M} c'_m e^{-k(\xi_{1,m}+\mathrm{j}\xi_{2,m})}$$

where $c'_m = \frac{c_m}{\alpha_m}\left(e^{\alpha_m(T-\tau_m)} - 1\right)$ as required. □

Again the sequence $\{\mathcal{R}(k)\}_{k=0}^{K}$, governed by the weighted sum of exponentials (3.20) can be solved to recover the M locations $\{\boldsymbol{\xi}_m\}_{m=1}^{M}$ of the instantaneous sources, as well as their corresponding generalized energies c'_m using Prony's method.

Corollary 3.1. *From* (3.16)*, given any arbitrary temporal parameterization of M sources, it is possible to recover their locations by evaluating the integral expression $\mathcal{R}(k)$, for $k = 0, 1, \ldots, K$ and applying Prony's method on the resulting sequence $\{\mathcal{R}(k)\}_{k=0}^{K}$, as long as all M sources are localized in space. Specifically, one can show that $\mathcal{R}(k)$ will always take the form $\mathcal{R}(k) = \sum_{m=1}^{M} C_m e^{-k(\xi_{1,m}+\mathrm{j}\xi_{2,m})}$, $k = 0, 1, \ldots, K$, where C_m is the generalized energy given by $C_m = \int_0^T h_m(t)\mathrm{d}t$ for the generic source with parameterization $f(\mathbf{x}, t) = \sum_m h_m(t)\delta(\mathbf{x} - \boldsymbol{\xi}_m)$.*

We are now left with establishing novel expressions for directly recovering the remaining source parameters: α_m, τ_m and c_m from the generalized energy c'_m.

Exact recovery of decay coefficients, activation times and source intensities

Let us start by noting that $c'_m = \frac{c_m}{\alpha_m}\left(e^{\alpha_m(T-\tau_m)} - 1\right)$ for $m = 1, \ldots, M$ depends on the interval $[0, T]$ over which the time-integration, in (3.21), is performed; thus we may write,

$$c'_m(T) = \frac{c_m}{\alpha_m}\left(e^{\alpha_m(T-\tau_m)} - 1\right) \tag{3.22}$$

to emphasize this dependency. There are three unknowns (α_m, c_m, τ_m) in (3.22) hence in order to find them, we need at least three linearly independent expressions. The simplest way to achieve this is to assume that we can obtain the coefficients $\{c'_m(T) : m = 1, \ldots, M\}$ for the three distinct intervals $T = T_1$, $T_2 = T_1 + \Delta T$, and $T_3 = T_1 + 2\Delta T$, then:

$$c'_m(T_1) = \frac{c_m}{\alpha_m}\left(e^{\alpha_m(T_1-\tau_m)} - 1\right), \tag{3.23}$$

$$\begin{aligned} c'_m(T_2) &= \frac{c_m}{\alpha_m}\left(e^{\alpha_m(T_1+\Delta T-\tau_m)} - 1\right) \\ &= \frac{c_m}{\alpha_m}\left(e^{\alpha_m(T_1-\tau_m)}e^{\alpha_m\Delta T} - 1\right), \end{aligned} \tag{3.24}$$

$$\begin{aligned} c'_m(T_3) &= \frac{c_m}{\alpha_m}\left(e^{\alpha_m(T_2+\Delta T-\tau_m)} - 1\right) \\ &= \frac{c_m}{\alpha_m}\left(e^{\alpha_m(T_2-\tau_m)}e^{\alpha_m\Delta T} - 1\right). \end{aligned} \tag{3.25}$$

Fortunately it is now possible to solve these simultaneously in closed-form as follows: subtract (3.23) from (3.24) and similarly (3.24) from (3.25), to obtain

$$c'_m(T_2) - c'_m(T_1) = \frac{c_m}{\alpha_m}\left(e^{\alpha_m(T_1-\tau_m)}e^{\alpha_m\Delta T} - e^{\alpha_m(T_1-\tau_m)}\right)$$

$$
\begin{aligned}
&= \frac{c_m}{\alpha_m} e^{\alpha_m (T_1 - \tau_m)} \left(e^{\alpha_m \Delta T} - 1 \right) \\
&= \frac{2c_m}{\alpha_m} e^{\alpha_m \frac{\Delta T}{2}} e^{\alpha_m (T_1 - \tau_m)} \sinh(\alpha_m \Delta T/2),
\end{aligned} \tag{3.26}
$$

and

$$
c'_m(T_3) - c'_m(T_2) = \frac{2c_m}{\alpha_m} e^{\alpha_m \frac{\Delta T}{2}} e^{\alpha_m (T_2 - \tau_m)} \sinh(\alpha_m \Delta T/2), \tag{3.27}
$$

respectively. Dividing (3.26) by (3.27) gives,

$$
\frac{c'_m(T_2) - c'_m(T_1)}{c'_m(T_3) - c'_m(T_2)} = \frac{e^{\alpha_m (T_1 - \tau_m)}}{e^{\alpha_m (T_2 - \tau_m)}} = e^{\alpha_m (T_1 - T_2)}. \tag{3.28}
$$

Therefore,

$$
\alpha_m = \frac{1}{T_1 - T_2} \ln \left(\frac{c'_m(T_2) - c'_m(T_1)}{c'_m(T_3) - c'_m(T_2)} \right). \tag{3.29}
$$

Given α_m it is then possible to retrieve the activation time of the m-th source if we divide (3.23) by (3.24),

$$
\frac{c'_m(T_1)}{c'_m(T_2)} = \frac{\frac{c_m}{\alpha_m} \left(e^{\alpha_m (T_1 - \tau_m)} - 1 \right)}{\frac{c_m}{\alpha_m} \left(e^{\alpha_m (T_2 - \tau_m)} - 1 \right)} \tag{3.30}
$$

and re-arrange to obtain,

$$
e^{-\alpha_m \tau_m} \left(c'_m(T_1) e^{\alpha_m T_2} - c'_m(T_2) e^{\alpha_m T_1} \right) = c'_m(T_1) - c'_m(T_2).
$$

This yields

$$
\tau_m = \frac{1}{\alpha_m} \ln \left(\frac{c'_m(T_1) e^{\alpha_m T_2} - c'_m(T_2) e^{\alpha_m T_1}}{c'_m(T_1) - c'_m(T_2)} \right). \tag{3.31}
$$

Finally, with access to estimates of α_m and τ_m, it is straightforward to estimate c_m using any of (3.23), (3.24), and (3.25). In particular

$$
c_m = \frac{\alpha_m c'_m(T_i)}{e^{\alpha_m (T_i - \tau_m)} - 1}, \; i = 1, 2, 3 \tag{3.32}
$$

as required.

Remark 3.2. *An interesting point to notice about the proposed framework is that, for sources lying outside the region Ω their contribution to the integral expressions* (3.17) *and* (3.21) *is zero. We can leverage this fact to still recover the sources of diffusion fields in bounded regions, i.e. when the medium through which the field propagates is finite. Specifically, the method of image sources allows us to model reflections in bounded regions as an unbounded medium containing in addition to real sources, several virtual ones. However, our integrals enclose only the real sources, hence using our inverse formulae only these real sources will be recovered because the contributions of the virtual sources to the integrals will be zero. Consequently, the inversion formulas remain valid and we are still able to estimate the unknown source parameters as far as the impermeable (or semi-permeable) boundaries*

of the medium are outside Ω. We will demonstrate through simulations, using both synthetic and real data, that we are able to fully recover the unknown source distribution f in the case where the field propagates a finite region with impermeable boundaries.

Algorithm 3.1 Localized IDS problem for continuous measurements

Require: Field $u(\mathbf{x}, t)\,\forall(\mathbf{x}, t) \in \Omega \times [0, T]$, total number of sources M, diffusivity μ, source model SM.

Initialize $K \geq 2M - 1$.
if SM == '*instantaneous*' **then**
  Compute the sequence $\{\mathcal{Q}(k, 1)\}_{k=0}^{K}$ using (3.17).
  Apply Prony's method to $\{\mathcal{Q}(k, 1)\}_{k=0}^{K}$ to obtain M pairs of $(c_m e^{-\mathrm{j}\tau_m/T}, \boldsymbol{\xi}_m)$.
  For every m, $c_m = \left|c_m e^{-\mathrm{j}\tau_m/T}\right|$ and $\tau_m = T \arg\left(c_m e^{-\mathrm{j}\tau_m/T}\right)$.
  Return all M intensities, activation times and locations $\{(c_m, \tau_m, \boldsymbol{\xi}_m)\}_{m=1}^{M}$.
else
  Initialize ΔT so that $\Delta T \leq T/3$.
  for $i = 1 : 3$ **do**
    $T_i = T - (3 - i)\Delta T$.
    Using measurements over $\Omega \times [0, T_i]$ compute the sequence $\{\mathcal{R}(k) \stackrel{\text{def}}{=} \mathcal{Q}(k, 0)\}_{k=0}^{K}$ using (3.17) (equivalently (3.21)).
    Apply Prony's method to $\{\mathcal{R}(k)\}_{k=0}^{K}$ to obtain M pairs $(c'_m(T_i), \boldsymbol{\xi}_m)$.
  end for
  By matching the locations form M quadruples $(c'_m(T_1), c'_m(T_2), c'_m(T_3), \boldsymbol{\xi}_m)$.
  For each m recover α_m, τ_m, and c_m from $(c'_m(T_1), c'_m(T_2), c'_m(T_3), \boldsymbol{\xi}_m)$ using (3.29), (3.31) and (3.32) respectively.
  Return all M source decay coefficients, intensities, activation times and locations: $\{\alpha_m, c_m, \tau_m, \boldsymbol{\xi}_m\}_{m=1}^{M}$.
end if

Consequently, we can leverage from the results herein to develop a localized source estimation algorithm, shown in Algorithm 3.1, that solves the Problems 3.1 and 3.2.

3.4.3. Estimating a single non-localized source

We now outline our proposed schemes for recovering non-localized source distributions (3.9) and (3.10) respectively, assuming access to the following *generalized measurements*:

$$\mathcal{Q}(k, r) = \langle \Psi_k(\mathbf{x})\Gamma_r(t), f\rangle = \int_{\Omega}\int_{t} \Psi_k(\mathbf{x})\Gamma_r(t) f(\mathbf{x}, t)\mathrm{d}t\mathrm{d}V, \tag{3.33}$$

where $\Psi_k(\mathbf{x}) = e^{-k(x_1 + \mathrm{j}x_2)}$ and $\Gamma_r(t) = e^{-\mathrm{j}rt/T}$, with $k, r \in \mathbb{N}$. Specifically, we show that the above sequence of integral measurements when appropriately modified results in a new sequence that is governed by a power sum series, that when solved gives again the unknown source parameters.

Analytic recovery of instantaneous line sources

In what follows we demonstrate how to recover the unknown straight line source parameters $(c, \tau, \boldsymbol{\xi}_1, \boldsymbol{\xi}_2)$ from the generalized measurements $\mathcal{Q}(k, r)$.

Proposition 3.3. *Let $\Psi_k(\mathbf{x})$ be the analytic function $\Psi_k(\mathbf{x}) = e^{-k(x_1+jx_2)}$, where $k = 1, 2, \ldots, K$ with $K \geq 4$ and let $\Gamma_r(t) = e^{-\mathrm{j}rt/T}$, where $r = 0, 1, \ldots, R$ and $R \geq 1$, then the generalized measurements $\mathcal{Q}(k, r)$ in (3.13) can be used to recover jointly, the unknown line source intensity, location and activation time.*

Proof. We begin the proof by considering the expression (3.13) and substitute the source parameterization (3.9) as follows:

$$\begin{aligned}\mathcal{Q}(k,r) &= \langle \Psi_k(\mathbf{x})\Gamma_r(t), f\rangle = \int_\Omega \int_t \Psi_k(\mathbf{x})\Gamma_r(t) f(\mathbf{x},t)\mathrm{d}t\mathrm{d}V \\ &= c\int_t \Gamma_r(t)\delta(t-\tau)\mathrm{d}t \int_\Omega \Psi_k(\mathbf{x})L(\mathbf{x})\mathrm{d}V \\ &= c\Gamma_r(\tau)\int_{L(\mathbf{x})} \Psi_k(\mathbf{x})\mathrm{d}S, \end{aligned} \tag{3.34}$$

the last equality follows from the fact that $L(\mathbf{x})$ is only non-zero along the shortest line joining the endpoints $\boldsymbol{\xi}_1$ and $\boldsymbol{\xi}_2$. Next given a parametric representation of the line segment such as

$$L(\mathbf{x}(\theta)) : \begin{cases} x_1(\theta) = (1-\theta)\xi_{1,1} + \theta\xi_{1,2} \\ x_2(\theta) = (1-\theta)\xi_{2,1} + \theta\xi_{2,2} \end{cases}, \theta \in [0,1], \tag{3.35}$$

we have that

$$\begin{aligned}\int_{L(\mathbf{x})} \Psi_k(\mathbf{x})\mathrm{d}S &= \int_0^1 \Psi_k(\mathbf{x}(\theta))\sqrt{\left(\frac{\mathrm{d}x_1}{\mathrm{d}\theta}\right)^2 + \left(\frac{\mathrm{d}x_2}{\mathrm{d}\theta}\right)^2}\mathrm{d}\theta \\ &= \sqrt{(\xi_{1,2}-\xi_{1,1})^2 + (\xi_{2,2}-\xi_{2,1})^2}\int_0^1 \Psi_k(\mathbf{x}(\theta))\mathrm{d}\theta \\ &= -\frac{1}{k}\ell(\boldsymbol{\xi}_1, \boldsymbol{\xi}_2)\left(e^{-k(\xi_{1,2}+\mathrm{j}\xi_{2,2})} - e^{-k(\xi_{1,1}+\mathrm{j}\xi_{2,1})}\right), \end{aligned}$$

providing $k \neq 0$. Thus

$$\int_{L(\mathbf{x})} \Psi_k(\mathbf{x})\mathrm{d}S = -\frac{1}{k}\ell(\boldsymbol{\xi}_1, \boldsymbol{\xi}_2)\sum_{m=1}^{2}(-1)^m e^{-k(\xi_{1,m}+\mathrm{j}\xi_{2,m})}, \tag{3.36}$$

where

$$\ell(\boldsymbol{\xi}_1, \boldsymbol{\xi}_2) = \frac{\sqrt{(\xi_{1,2}-\xi_{1,1})^2 + (\xi_{2,2}-\xi_{2,1})^2}}{(\xi_{1,2}-\xi_{1,1}) + j(\xi_{2,2}-\xi_{2,1})}. \tag{3.37}$$

Substituting (3.36) into (3.34) and recalling that $\Gamma_r(t) = e^{-\mathrm{j}rt/T}$ yields the following power

sum series,

$$-k\mathcal{Q}(k,r) = \ell(\boldsymbol{\xi}_1,\boldsymbol{\xi}_2)ce^{-\mathrm{j}r\tau/T}\sum_{m=1}^{2}(-1)^m e^{-k(\xi_{1,m}+\mathrm{j}\xi_{2,m})}. \tag{3.38}$$

Applying Prony's method, as usual, to the sequence $\{-k\mathcal{Q}(k,1)\}_{k=1}^{K}$ given that $K \geq 4$ produces $\{(-1)^m\ell(\boldsymbol{\xi}_1,\boldsymbol{\xi}_2)ce^{-\mathrm{j}\tau/T},\boldsymbol{\xi}_m\}_{m=1}^{2}$. Using $\boldsymbol{\xi}_1$ and $\boldsymbol{\xi}_2$ we can compute exactly $\ell(\boldsymbol{\xi}_1,\boldsymbol{\xi}_2)$. Thus c and τ can then be directly recovered from $\{(-1)^m\ell(\boldsymbol{\xi}_1,\boldsymbol{\xi}_2)ce^{-\mathrm{j}\tau/T}\}_{m=1}^{2}$. □

Notice again that the choice $\Psi_k(\mathbf{x}) = e^{-k(x_1+\mathrm{j}x_2)}$ and $\Gamma_r(t) = e^{-\mathrm{j}rt/T}$ is key to obtaining the desired power sum series (3.38), once more highlighting the roles of the functions $\Psi_k(\mathbf{x})$ and $\Gamma_r(t)$ here as spatial and temporal sensing functions respectively.

Thus the algorithm that solves the continuous IDSP, Problem 3.3, using the above proposition is summarized in Algorithm 3.2.

Analytic recovery of instantaneous polygonal sources

We now outline how to estimate the unknown parameters c, τ and $\{\boldsymbol{\xi}_m\}_{m=1}^{M}$ for an M-sided convex polygonal diffusion source given access to the generalized measurements $\mathcal{Q}(k,r)$. We firstly state the following lemmas that will be useful in proving the scheme for recovering polygonal sources.

Lemma 3.1. *Let $\Psi_k(\mathbf{x}) = e^{-k(x_1+\mathrm{j}x_2)}$, then*

$$\Psi_k(\mathbf{x}) = \frac{1}{k^2}\Psi_k''(\mathbf{x}), \tag{3.39}$$

where $(\cdot)'$ is used to denote the derivative with respect to the complex variable '$x_1 + \mathrm{j}x_2$'.

Proof. Follows directly from two straightforward applications of the complex derivative (*w.r.t* the complex variable '$x_1 + \mathrm{j}x_2$'). □

Lemma 3.2. *Let $\Psi_k(\mathbf{x})$ be analytic inside the convex polygon $F(\mathbf{x})$, with vertices $\{\boldsymbol{\xi}_m\}_{m=1}^{M}$, then [49]*

$$\int_{F(\mathbf{x})}\Psi_k''(\mathbf{x})\mathrm{d}V = \sum_{m=1}^{M} a_m\Psi_k(\boldsymbol{\xi}_m), \tag{3.40}$$

where the coefficients a_m are related to the vertices of the polygon via,

$$a_m = \frac{\mathrm{j}}{2}\left(\frac{(\xi_{1,n-1}-\mathrm{j}\xi_{2,n-1})-(\xi_{1,n}-\mathrm{j}\xi_{2,n})}{(\xi_{1,n-1}+\mathrm{j}\xi_{2,n-1})-(\xi_{1,n}+\mathrm{j}\xi_{2,n})} - \frac{(\xi_{1,n}-\mathrm{j}\xi_{2,n})-(\xi_{1,n+1}-\mathrm{j}\xi_{2,n+1})}{(\xi_{1,n}+\mathrm{j}\xi_{2,n})-(\xi_{1,n+1}+\mathrm{j}\xi_{2,n+1})}\right). \tag{3.41}$$

Proof. For a proof of this lemma, see [49, 50, 64]. □

Proposition 3.4. *Let $\Psi_k(\mathbf{x})$ be the analytic function $\Psi_k(\mathbf{x}) = e^{-k(x_1+jx_2)}$, where $k = 1,2,\ldots,K$ with $K \geq 2M$ and let $\Gamma_r(t) = e^{-\mathrm{j}rt/T}$, where $r = 0,1,\ldots,R$ and $R \geq 1$, then the generalized measurements $\mathcal{Q}(k,r)$ in (3.13) can be used to recover jointly, the unknown intensity, vertices and activation time of an M-sided convex polygonal diffusion source.*

Proof. Consider now the expression (3.13) and substitute the source parameterization (3.10) as follows:

$$\begin{aligned}
\mathcal{Q}(k,r) &= \langle \Psi_k(\mathbf{x})\Gamma_r(t), f\rangle = \int_\Omega \int_t \Psi_k(\mathbf{x})\Gamma_r(t) f(\mathbf{x},t)\mathrm{d}t\mathrm{d}V \\
&= c\int_t \Gamma_r(t)\delta(t-\tau)\mathrm{d}t \int_\Omega \Psi_k(\mathbf{x})F(\mathbf{x})\mathrm{d}V \\
&= c\Gamma_r(\tau)\int_{F(\mathbf{x})} \Psi_k(\mathbf{x})\mathrm{d}V \\
&\overset{(i)}{=} c\Gamma_r(\tau)\int_{F(\mathbf{x})} \frac{1}{k^2}\Psi_k''(\mathbf{x})\mathrm{d}V \\
&\overset{(ii)}{=} \frac{1}{k^2}c\Gamma_r(\tau)\sum_{m=1}^{M} a_m \Psi_k(\boldsymbol{\xi}_m)
\end{aligned}$$

where the equality (i) follows from Lemma 3.1 and (ii) from Lemma 3.2. Multiplying through by $k^2, k \neq 0$ and substituting the expressions for $\Psi_k(\mathbf{x})$ and $\Gamma_r(t)$ gives

$$k^2\mathcal{Q}(k,r) = ce^{-\mathrm{j}r\tau/T}\sum_{m=1}^{M} a_m e^{-k(\xi_{1,m}+\mathrm{j}\xi_{2,m})} \tag{3.42}$$

which is again a coupled power sum series. Thus the unknowns of the M-sided polygonal diffusion source can be recovered from $\{\mathcal{Q}(k,1)\}_{k=1}^{K}$ using Prony's method provided $K \geq 2M$. □

Consequently, we can summarize the procedure for solving the IDSP of Problem 3.4 as shown in Algorithm 3.3.

In this section we considered and presented results for the IDSP for a certain class of source models. Interestingly, these results can be further extended to solve the inverse source problem for diffusion fields in the presence of an advection current.

3.5. An extension to advection-diffusion fields

The advection-diffusion equation (ADE) is the underlying mathematical model for transport in a range of biological and physical situations, specifically, when the transport is a combination of both diffusion and advection (also called convection or drift depending on the context). The dispersion of biochemical/nuclear contaminant released into the atmosphere in the presence of wind [139] and the flow of gases or liquids in porous soils and rocks [68, 69], are two typical examples. Indeed the identification of sources of advection-diffusion fields has been a topic of great interest in several environmental and industrial applications and as such has received a lot of research efforts. Herein, we define analogously to the IDSPs a class of inverse advection-diffusion source problems (IADSPs) and prove that the results for IDSPs can be extended to this new setting.

Algorithm 3.2 Solve Line IDS problem given continuous measurements

Require: Field $u(\mathbf{x}, t)\,\forall(\mathbf{x}, t) \in \Omega \times [0, T]$, total number of sources M, diffusivity μ, source model SM.

Initialize $K \geq 4$.
if SM == '*instantaneous*' **then**
  Compute the sequence $\{\mathcal{Q}(k, 1)\}_{k=1}^{K}$ using (3.17).
  From $\{\mathcal{Q}(k, 1)\}_{k=1}^{K}$, for each $k \in 1, \ldots, K$ form a new sequence $\{-k\mathcal{Q}(k, 1)\}_{k=1}^{K}$.
  Apply Prony's method to $\{-k\mathcal{Q}(k, 1)\}_{k=1}^{K}$ to obtain $(c\ell(\boldsymbol{\xi}_1, \boldsymbol{\xi}_2)e^{-\mathrm{j}\tau/T}, \boldsymbol{\xi}_1, \boldsymbol{\xi}_2)$.
  Compute the value of $\ell(\boldsymbol{\xi}_1, \boldsymbol{\xi}_2)$ using (3.37), hence $c = \left|\frac{c\ell(\boldsymbol{\xi}_1,\boldsymbol{\xi}_2)e^{-\mathrm{j}\tau/T}}{\ell(\boldsymbol{\xi}_1,\boldsymbol{\xi}_2)}\right|$ and $\tau = -T\arg\left(\frac{c\ell(\boldsymbol{\xi}_1,\boldsymbol{\xi}_2)e^{-\mathrm{j}\tau/T}}{\ell(\boldsymbol{\xi}_1,\boldsymbol{\xi}_2)}\right)$.
  Return intensity, activation time and endpoints $(c, \tau, \boldsymbol{\xi}_1, \boldsymbol{\xi}_2)$ of the line source.
else
  Initialize ΔT so that $\Delta T \leq T/3$.
  for $i = 1 : 3$ **do**
    $T_i = T - (3 - i)\Delta T$.
    Using measurements over $\Omega \times [0, T_i]$ compute the sequence $\{\mathcal{R}(k) \stackrel{\text{def}}{=} \mathcal{Q}(k, 0)\}_{k=1}^{K}$ using (3.17) (equivalently (3.21)).
    For each $k = 1, \ldots, K$ multiply corresponding term in $\{\mathcal{R}(k)\}_{k=1}^{K}$ by $-k$ to get $\{-k\mathcal{R}(k)\}_{k=1}^{K}$.
    Apply Prony's method to $\{-k\mathcal{R}(k)\}_{k=0}^{K}$ to obtain $(c'(T_i)\ell(\boldsymbol{\xi}_1, \boldsymbol{\xi}_2), \boldsymbol{\xi}_1, \boldsymbol{\xi}_2)$.
  end for
  Compute $\ell(\boldsymbol{\xi}_1, \boldsymbol{\xi}_2)$ using (3.37) and divide the above generalized energies by it to get $(c'(T_1), c'(T_2), c'(T_3))$.
  Recover α, τ, and c from $(c'(T_1), c'(T_2), c'(T_3))$ using (3.29), (3.31) and (3.32) respectively.
  Return the decay coefficient, intensity, activation time and endpoints $(\alpha, c, \tau, \boldsymbol{\xi}_1, \boldsymbol{\xi}_2)$ of the line source.
end if

As usual we begin with the forward model for the advection-diffusion phenomena. In particular, the field $u(\mathbf{x}, t)$ induced by the source distribution $f(\mathbf{x}, t)$ undergoing both *diffusion* and *advection* transport is governed by the advection-diffusion equation,

$$\frac{\partial}{\partial t}u(\mathbf{x}, t) = \mu\nabla^2 u(\mathbf{x}, t) - \mathbf{v}_a \cdot \nabla u(\mathbf{x}, t) + f(\mathbf{x}, t), \tag{3.43}$$

where $\mathbf{x} \in \mathbb{R}^d$ is the d-dimensional spatial domain, $t \in \mathbb{R}_+$ is the temporal domain, $\mu \in \mathbb{R}$ is the diffusivity of the medium through which the field propagates and $\mathbf{v}_a \in \mathbb{R}^d$ is the velocity of the advection current (wind velocity for instance).

Assuming $\Psi_k(\mathbf{x})$ is twice-differentiable in Ω and once again invoking Green's second identity (3.14), we can obtain the following corollary:

Corollary 3.2. *Let $u(\mathbf{x}, t)$ be an advection-diffusion field induced by an unknown source distribution $f(\mathbf{x}, t)$, i.e. $u(\mathbf{x}, t)$ satisfies the ADE (3.43) with a known advection velocity $\mathbf{v}_a$. Given that $\Psi_k(\mathbf{x})$ is chosen to be a complex analytic function of the form $\Psi_k(\mathbf{x}) = e^{-k(\xi_1+j\xi_2)}$ where $k = 0, \ldots, K$ with $K \geq 2M - 1$, and $\Gamma_r(t) = e^{-\mathrm{j}rt/T}$ for $r = 0, 1, \ldots, R$*

Algorithm 3.3 Solve M-gon IDS problem given continuous measurements

Require: Field $u(\mathbf{x}, t)\,\forall(\mathbf{x}, t) \in \Omega \times [0, T]$, total number of sources M, diffusivity μ, source model SM.

Initialize $K \geq 2M$.
if SM == '*instantaneous*' **then**
  Compute the sequence $\{\mathcal{Q}(k, 1)\}_{k=1}^{K}$ using (3.17).
  Multiply corresponding term in $\{\mathcal{Q}(k, 1)\}_{k=1}^{K}$ by k^2, for each $k = 1, 2, \ldots, K$ to get $\{k^2\mathcal{Q}(k, 1)\}_{k=1}^{K}$.
  Apply Prony's method to $\{k^2\mathcal{Q}(k, 1)\}_{k=1}^{K}$ to obtain M pairs $(ca_m e^{-\mathrm{j}\tau/T}, \boldsymbol{\xi}_m)$.
  Evaluate a_m using (3.41) and hence find $c = \left|\frac{ca_m e^{-\mathrm{j}\tau/T}}{a_m}\right|$ and $\tau = -T \arg\left(\frac{ca_m e^{-\mathrm{j}\tau/T}}{a_m}\right)$.
  Return the intensity, activation time and M vertices $\left(c, \tau, \{\boldsymbol{\xi}_m\}_{m=1}^{M}\right)$ of polygonal source.
else
  Initialize ΔT so that $\Delta T \leq T/3$.
  for $i = 1 : 3$ **do**
    $T_i = T - (3 - i)\Delta T$.
    Using measurements over $\Omega \times [0, T_i]$ compute the sequence $\{\mathcal{R}(k) \stackrel{\text{def}}{=} \mathcal{Q}(k, 0)\}_{k=1}^{K}$ using (3.17) (or equivalently (3.21)).
    For each $k = 1, \ldots, K$ multiply corresponding term in $\{\mathcal{R}(k)\}_{k=1}^{K}$ by k^2 to get $\{k^2\mathcal{R}(k)\}_{k=1}^{K}$.
    Apply Prony's method to $\{k^2\mathcal{R}(k)\}_{k=0}^{K}$ to obtain M pairs of $\{c'(T_i)a_m, \boldsymbol{\xi}_m\}_{m=1}^{M}$.
  end for
  Compute a_m using (3.41) and divide generalized energies above by it to get $(c'(T_1), c'(T_2), c'(T_3))$.
  Recover α, τ, and c from $(c'(T_1), c'(T_2), c'(T_3))$ using (3.29), (3.31) and (3.32) respectively.
  Return estimated decay coefficient, intensity, activation time and M vertices $\left(\alpha, c, \tau, \{\boldsymbol{\xi}_m\}_{m=1}^{M}\right)$ of the polygonal source.
end if

with $R \geq 1$, then given any source distribution $f(\mathbf{x}, t)$ admitting a parameterization (3.7), (3.8), (3.9), *or* (3.10) *it is possible to retrieve simultaneously the unknown source parameters from the corresponding continuous field measurements taken over $\Omega \times [0, T]$.*

Proof. It is clear that one would only need to show it is possible to obtain such a sequence $\{\mathcal{Q}_a(k, r)\}_{k,r}$, where $\mathcal{Q}_a(k, r) = \langle \Psi_k(\mathbf{x})\Gamma_r(t), f(\mathbf{x}, t)\rangle$, from the continuous measurements of the field on $\Omega \times [0, T]$. If we can obtain this sequence then all the properties of $\{\mathcal{Q}(k, r)\}_{k,r}$ (and its various modified versions) also hold. For example applying Prony's method to:

1. $\{\mathcal{Q}_a(k, 1)\}_{k=0}^{K}$ when $K \geq 2M - 1$ returns the unknown parameters for localized instantaneous sources undergoing advection-diffusion transport; or

2. to $\{\mathcal{R}_a(k) \stackrel{\text{def}}{=} \mathcal{Q}_a(k, 0)\}_{k=0}^{K}$, similarly to the IDSP non-instantaneous point sources, gives the unknown source parameters for the non-instantaneous point sources.

We proceed as follows: consider a complex analytic $\Psi_k(\mathbf{x})$ satisfying $\nabla^2\Psi_k(\mathbf{x}) = 0$, then

(3.14) reduces to

$$\oint_{\partial\Omega} (\Psi_k \nabla u - u \nabla \Psi_k) \cdot \hat{\mathbf{n}}_{\partial\Omega} \mathrm{d}S = \int_{\Omega} \Psi_k \nabla^2 u \mathrm{d}V. \tag{3.44}$$

Now substitute $\nabla^2 u = \frac{1}{\mu}\frac{\partial}{\partial t} u(\mathbf{x},t) + \frac{1}{\mu}\mathbf{v}_a \cdot \nabla u(\mathbf{x},t) - \frac{1}{\mu} f(\mathbf{x},t)$, i.e. (3.43), into (3.44) and rearrange to get

$$\int_{\Omega} \Psi_k \left(\frac{\partial u}{\partial t} + \mathbf{v}_a \cdot \nabla u \right) \mathrm{d}V - \mu \oint_{\partial\Omega} (\Psi_k \nabla u - u \nabla \Psi_k) \cdot \hat{\mathbf{n}}_{\partial\Omega} \mathrm{d}S = \int_{\Omega} \Psi_k f \mathrm{d}V. \tag{3.45}$$

Consider the time-integral of the product between $\Gamma_r(t)$ and (3.16) over the interval $t \in [0,T]$,

$$\overbrace{\int_t \Gamma_r(t) \int_{\Omega} \Psi_k \left(\frac{\partial u}{\partial t} + \mathbf{v}_a \cdot \nabla u \right) \mathrm{d}V - \mu \Gamma_r(t) \oint_{\partial\Omega} (\Psi_k \nabla u - u \nabla \Psi_k) \cdot \hat{\mathbf{n}}_{\partial\Omega} \mathrm{d}S \mathrm{d}t}^{\stackrel{\text{def}}{=} \mathcal{Q}_a(k,r)} = \int_t \Gamma_r(t) \int_{\Omega} \Psi_k f \mathrm{d}V \mathrm{d}t. \tag{3.46}$$

We notice that the left hand side of (3.46), denoted by $\mathcal{Q}_a(k,r)$, depends only on the field $u(\mathbf{x},t)$, the known advection velocity $\mathbf{v}_a$ and the sensing functions $\Psi_k(\mathbf{x})$ and $\Gamma_r(t)$, as such it can be evaluated exactly for any $k, r \in \mathbb{N}$, given the continuous field measurements $u(\mathbf{x},t)$ for all $(\mathbf{x},t) \in \Omega \times [0,T]$. This gives a sequence of measurements governed by $\mathcal{Q}_a(k,r) = \int_{t\in[0,T]} \int_{\Omega} \Gamma_r(t) \Psi_k(\mathbf{x}) f(\mathbf{x},t) \mathrm{d}V \mathrm{d}t$ as required. Therefore we could simply replace every occurrence of $\mathcal{Q}(k,r)$ in Algorithms 3.1 to 3.3 with the quantity $\mathcal{Q}_a(k,r)$ (computed using (3.46)) to obtain algorithms for solving the inverse source problems governed by the ADE for the respective source models of interest. □

Example 3.1. *As an example, under the instantaneous and localized source model* (3.7) *we obtain the usual sum of exponentials: i.e.* $\mathcal{Q}_a(k,r) = \sum_{m=1}^{M} c_m e^{-\mathrm{j}r\tau_m/T} e^{-k(\xi_{1,m}+\mathrm{j}\xi_{2,m})}$ *for each* $k = 0,1,\ldots,K$ *and* $r = 0,1,\ldots,R$.

3.6. Summary and Conclusion

In this chapter, we have studied inverse source problems related to the diffusion equation. More precisely, we studied the problems assuming a class of source models as defined in Section 3.3. These models although inspired by real world applications and scenarios, also play the role of parametric regularizers – allowing us to obtain meaningful and unique solutions for the inverse source problems considered.

Consequently, our proposed methodology results in a framework that provides a novel closed form expression, (3.17), for computing a set of generalized measurement sequence $\{\mathcal{Q}(k,r)\}_{k,r}$—from the continuous field $u(\mathbf{x},t)$—with a certain sum-of-exponentials structure, in the localized source setting. Although for non-localized source fields this sum-of-exponentials structure is no longer preserved, we demonstrated that it can be successfully restored through a simple multiplication of the terms in $\{\mathcal{Q}(k,r)\}_{k,r}$ by a suitable constant (that depends on k). We summarized the approach for solving the IDSP for point, line

Table 3.1.: Summary of the source distributions of interest along with their corresponding generalized sequence to which Prony's method or matrix pencil method can be applied.

	Source Model	$f(\mathbf{x}, t)$	Gen. sequence	K
Point	Instantaneous	$\sum_{m=1}^{M} c_m \delta(\mathbf{x} - \boldsymbol{\xi}_m, t - \tau_m)$	$\{\mathcal{Q}(k,1)\}_{k=0}^{K}$	$K \geq 2M - 1$
	Time-varying	$\sum_{m=1}^{M} c_m e^{\alpha_m (t-\tau_m)} \delta(\mathbf{x} - \boldsymbol{\xi}_m) H(t - \tau_m)$	$\{\mathcal{R}(k)\}_{k=0}^{K}$	$K \geq 2M - 1$
Line	Instantaneous	$cL(\mathbf{x})\delta(t - \tau)$	$\{k\mathcal{Q}(k,1)\}_{k=1}^{K}$	$K \geq 4$
	Time-varying	$cL(\mathbf{x})e^{\alpha(t-\tau)}H(t - \tau)$	$\{k\mathcal{R}(k)\}_{k=1}^{K}$	$K \geq 4$
M-polygon	Instantaneous	$cP(\mathbf{x})\delta(t - \tau)$	$\{k^2\mathcal{Q}(k,1)\}_{k=1}^{K}$	$K \geq 2M$
	Time-varying	$cP(\mathbf{x})e^{\alpha(t-\tau)}H(t - \tau)$	$\{k^2\mathcal{R}(k)\}_{k=1}^{K}$	$K \geq 2M$

and polygonal sources in Algorithms 3.1 to 3.3 respectively. In addition Table 3.1 provides a summary/reminder of the class of source models considered, along with corresponding proper generalized sequence.

Chapter 4.

Estimating sources of diffusion fields from sensor network measurements

This chapter deals with the problem of estimating the sources of a diffusion field from its spatiotemporal samples. Specifically we consider the scenario where spatiotemporal samples of the field are obtained using a sensor network, made of monitoring sensor nodes deployed over a region of interest (see Figure 4.1 for an example). In this scenario, our goal is then to estimate the unknown source distribution f—inducing the measured field $u(\mathbf{x}, t)$ satisfying (3.2)—from discrete spatial and temporal samples of $u(\mathbf{x}, t)$.

In stark contrast to the continuous problem treated in Chapter 3, in this new setting we are able to capture more realistic signal processing and sensor network scenarios. However, this comes at a cost; in particular, if we wish to leverage from the results and inverse formulae derived for the continuous measurements IDSP we will be faced with two main issues: first, not having access to continuous-field measurements means we must (stably) approximate them from the sensor measurements; second, the actual sensor measurements may be noisy, therefore the inversion formulas derived need to be adjusted in order to promote stability and robustness of our proposed estimation schemes. Consequently, we will pay particular attention to these issues as we develop the theory and algorithms for solving the IDSP given discrete field measurements.

The rest of this chapter is organized as follows. In Section 4.1 we give details of the centralized sensor network model and assumptions, allowing us to then define precisely the resulting class of discrete IDSP under our choice of source models. We develop a centralized strategy for estimating the generalized measurement $\mathcal{Q}(k, r)$ from spatiotemporal sensor measurements of the diffusion field in Section 4.2. Then, in Section 4.3 we present noise robust algorithms to solve the class of discrete IDSP. In Section 4.4 we describe the underlying assumptions for the distributed sensor network. Then in Section 4.5 we give an overview of gossip schemes for distributed average consensus, subsequently in Section 4.6, we develop a novel diffusion equation-driven gossip scheme that can solve our IDSPs in a distributed fashion over a sensor network.

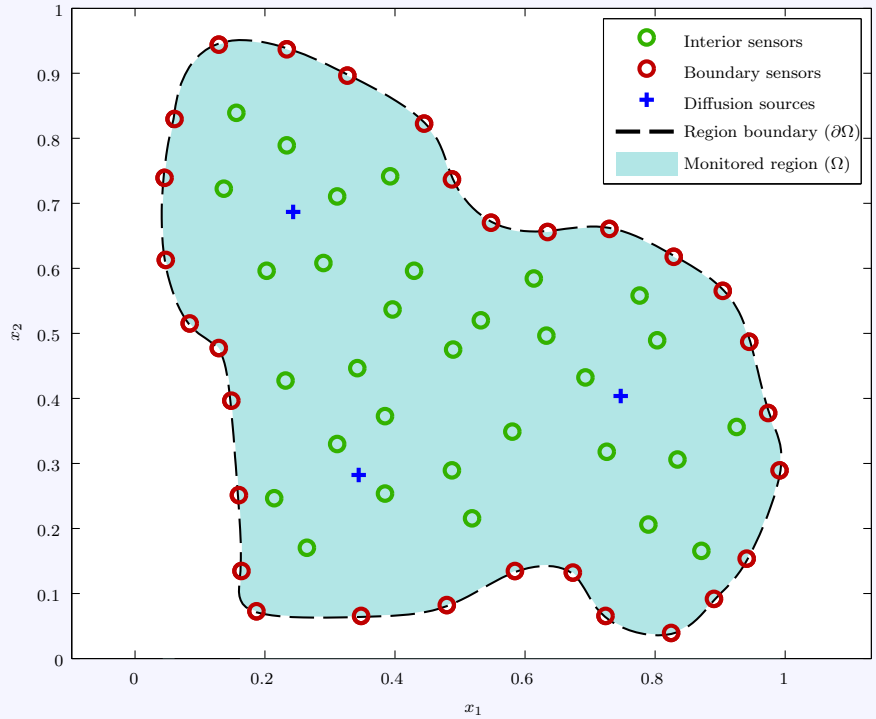


Figure 4.1.: **Sensor Placement.** An arbitrary sensor placement and the monitored domain Ω.

4.1. Sensor network model and problem formulation

Prior to formulating our class of inverse diffusion source problem for discrete spatiotemporal sensor measurements, we outline some model assumptions about the sensor network used to monitor the diffusion field, namely:

1. The sensor network comprises N sensor nodes that are deployed (uniformly or randomly) over the region of interest (Ω).

2. All N sensors are assumed to lie in the same plane.

3. The n-th sensor is positioned at a known location $\mathbf{x}_n \in \Omega$, where $n = 1, 2, \ldots, N$, and collects temporal samples of the field locally at time instants t_l for $l = 0, 1, \ldots, L$. Hence the noiseless field samples is simply the diffusion field evaluated at $\mathbf{x} = \mathbf{x}_n$ and $t = t_l$ as follows:
$$\varphi_n(t_l) = u(\mathbf{x}_n, t_l). \tag{4.1}$$

4. To model noisy sensor measurements, it is assumed that the spatiotemporal samples (4.1) above are further corrupted with zero mean additive white Gaussian noise (AWGN), i.e.:
$$\varphi^{\epsilon}_{n,l} = \varphi_n(t_l) + \epsilon_{n,l}, \tag{4.2}$$

where $\epsilon_{n,l} \sim \mathcal{N}(0, \sigma^2)$.

5. The sensor nodes are synchronized, as such the field samples obtained by the n-th sensor are $\{\varphi_n(t_l)\}_{l=0}^L$, and similarly $\{\varphi_n^\epsilon(t_l)\}_{l=0}^L$ in the noisy setting.

4.1.1. Centralized network model

In the centralized estimation scenario, the sensor network is assumed to comprise N *dumb* sensor nodes, that are without any data processing capabilities. Specifically, they are only able to: (i) sense/measure the field locally with a given temporal sampling frequency $f = \frac{1}{\Delta_t}$, so that $t_l = l\Delta_t$; as well as, (ii) communicate/forward their field measurements (spatiotemporal samples) directly to a central processing unit called the fusion centre (FC), via some (noiseless communication) channels. The FC however can perform the desired (batch) processing on the received data samples, in order to perform the inference tasks of recovering the unknown source distribution f.

Under this centralized sensor network model, we will first address the IDS problems given noiseless spatiotemporal samples, before considering the corresponding noisy data case.

We can now precisely state the new *discrete* IDS problems of interest for our class of source models (3.7), (3.8), (3.9) and (3.10), respectively, as follows:

Problem 4.1. *Given spatiotemporal samples $\varphi_n(t_l) = u(\mathbf{x}_n, t_l)$, at discrete spatial locations $\mathbf{x}_n \in \Omega$ with $n = 1, \ldots, N$ and time instants $t_l \in \mathbb{R}_+$ with $l = 0, 1, \ldots, L$, of a diffusion field $u(\mathbf{x}, t)$ induced by the instantaneous point source distribution* (3.7)*, the discrete IDSP for this source model is to estimate all M source intensities, locations and activation times $\{c_m, \boldsymbol{\xi}_m, \tau_m\}_{m=1}^M$ of the source distribution f.*

Problem 4.2. *The discrete IDSP for* (3.8) *is to estimate the M source decay coefficients, intensities, locations and activation times $\{\alpha_m, c_m, \boldsymbol{\xi}_m, \tau_m\}_{m=1}^M$ of the source distribution f, given its field samples $\varphi_n(t_l) = u(\mathbf{x}_n, t_l)$ as described in Problem 4.1.*

Problem 4.3. *For the instantaneous line source distribution* (3.9)*, the discrete IDSP reduces to estimating the intensity, activation time and endpoints $(c, \tau, \boldsymbol{\xi}_1, \boldsymbol{\xi}_2)$ of the source, from its field samples $\varphi_n(t_l) = u(\mathbf{x}_n, t_l)$ as described in Problem 4.1.*

Problem 4.4. *The discrete IDSP for* (3.10) *is to recover the intensity, activation time and vertices $(c, \tau, \{\boldsymbol{\xi}_m\}_{m=1}^M)$ of the M-sided polygonal source distribution f, from the samples of the induced field $\varphi_n(t_l) = u(\mathbf{x}_n, t_l)$ as described in Problem 4.1.*

4.2. Generalized measurements from spatiotemporal samples: A centralized approach

In the previous chapter (Section 3.4) we demonstrated the possibility of recovering the unknown source distribution f under a particular class of source models, from a sequence of so called generalized measurements $\{\mathcal{Q}(k, r)\}_{k,r}$. Therein, we obtained these generalized

measurements from the continuous field measurements taken over the domain $\Omega \times [0, T]$; specifically in that setting, we derived an exact expression for computing $\{\mathcal{Q}(k,r)\}_{k,r}$ from the field, i.e. (3.17), restated below, for ease of reference:

$$\mathcal{Q}(k,r) \stackrel{\text{def}}{=} \int_{t=0}^{T} \Gamma_r(t) \int_{\Omega} \frac{\partial}{\partial t}(u(\mathbf{x},t)\Psi_k)\,\mathrm{d}V\,\mathrm{d}t - \mu \int_{t=0}^{T} \Gamma_r(t) \oint_{\partial\Omega} (\Psi_k \nabla u - u \nabla \Psi_k)\cdot \hat{\mathbf{n}}_{\partial\Omega}\,\mathrm{d}S\mathrm{d}t.$$

In fact we can interchange the order of the time and space integrals to obtain

$$\begin{aligned}\mathcal{Q}(k,r) &= \int_{\Omega} \int_{t=0}^{T} \Gamma_r(t) \frac{\partial}{\partial t}(u(\mathbf{x},t)\Psi_k)\,\mathrm{d}t\,\mathrm{d}V - \mu \oint_{\partial\Omega} \int_{t=0}^{T} \Gamma_r(t)(\Psi_k \nabla u - u\nabla\Psi_k)\cdot\hat{\mathbf{n}}_{\partial\Omega}\,\mathrm{d}t\mathrm{d}S \\ &= \int_{\Omega} \Psi_k \dot{U}_r(\mathbf{x},T)\,\mathrm{d}V - \mu \oint_{\partial\Omega} (\Psi_k \nabla U_r(\mathbf{x},T) - U_r(\mathbf{x},T)\nabla\Psi_k)\cdot\hat{\mathbf{n}}_{\partial\Omega}\,\mathrm{d}S, \end{aligned} \tag{4.3}$$

where the last equality above follows from a single application of integration by parts, whilst $U_r(\mathbf{x},T)$ and $\dot{U}_r(\mathbf{x},T)$ are defined as follows:

$$U_r(\mathbf{x},T) = \int_0^T \Gamma_r(t) u(\mathbf{x},t)\mathrm{d}t, \quad \text{and,} \tag{4.4a}$$

$$\dot{U}_r(\mathbf{x},T) = \Gamma_r(T) u(\mathbf{x},T) - \int_0^T \frac{\partial \Gamma_r(t)}{\partial t} u(\mathbf{x},t)\mathrm{d}t. \tag{4.4b}$$

Analogously, the task for this section will be estimating these generalized measurements from the discrete field samples instead of the continuous data assumed of Chapter 3. In particular, given only spatiotemporal samples $\{\varphi_n(t_l)\}_{n,l}$ of the field, we propose a centralized strategy that approximates the integrals (3.17) (and (3.21)) using standard quadrature techniques [46]. We will focus on the estimation of $\mathcal{Q}(k,r)$, since $\mathcal{R}(k)$ is governed by (3.21) and so can be easily obtained from it. For the temporal integrals, a straightforward application of *trapezium rule* produces a good approximation. However the spatial integrals require some care; specifically, we are concerned with approximating:

- The family of path integrals along a boundary ($\partial\Omega$) of Ω given by:

$$\oint_{\partial\Omega} (\Psi_k(\mathbf{x}) \nabla U_r(\mathbf{x},T) - U_r(\mathbf{x},T)\nabla\Psi_k(\mathbf{x}))\cdot\hat{\mathbf{n}}_{\partial\Omega}\,\mathrm{d}S,$$

  as well as,

- The family of surface integrals on the bounded region Ω:

$$\int_{\Omega} \Psi_k(\mathbf{x}) \dot{U}_r(\mathbf{x},T)\,\mathrm{d}V.$$

Approximating these integrals with sums relies on obtaining non-overlapping subdivisions of the domain over which the integral is performed. Let us denote these subdivisions that make up the path ($\partial\Omega$) or surface (Ω) integrals, that is the line or triangular segments

by δl_i with $i = 1, \ldots, I$ and Δ_j with $j = 1, \ldots, J$, respectively. Hence for path integrals, providing $\delta l_i \bigcap \delta l_j = \varnothing$ for $i \neq j$ and $\bigcup_{i=1}^{I} \delta l_i = \partial\Omega$, a well-known approximation exists, namely the path integral of some function $h(\mathbf{x})$ along $\partial\Omega$ is approximated as follows:

$$\oint_{\partial\Omega} h(\mathbf{x})\mathrm{d}S \approx \sum_{i=1}^{I} \frac{[h(l_{i,1}) + h(l_{i,2})]}{2} \cdot |\delta l_i|, \tag{4.5}$$

where $l_{i,1} \in \mathbb{R}^2$ and $l_{i,2} \in \mathbb{R}^2$ denote the end points of the line segment δl_i and $|\delta l_i|$ is its length.

Moreover, with surface integrals, if these non-overlapping subdivisions $\{\Delta_j\}_{j=1}^{J}$ are triangular such that $\bigcup_{i=1}^{I} \Delta_i = \Omega$ and $\Delta_i \bigcap \Delta_j = \varnothing$ for $i \neq j$, the surface integral of $h(\mathbf{x})$ over a bounded region Ω is approximated by the sum [72]:

$$\int_{\Omega} h(\mathbf{x})\mathrm{d}V \approx \sum_{j=1}^{J} \frac{1}{3} \sum_{j'=1}^{3} h(\mathbf{v}_{j,j'})|\Delta_j|, \tag{4.6}$$

where $\{\mathbf{v}_{j,j'} : j' = 1, 2, 3\}$ are the vertices of Δ_j.

In our setup, these vertices coincide with the sensor locations, hence the triangular subdivisions depend directly on them.

Denote the collection of all sensor locations by $\mathcal{S} = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N\}$, we intend to construct non-overlapping triangular subdivisions given the set $\mathcal{S}$. This allows us to define the domains Ω and $\partial\Omega$ (its boundary) over which the surface and line integrals will be performed, respectively. In order to obtain stable approximations of these integrals, we will seek a triangulation that minimizes the occurrence of skinny triangles as these can sometimes introduce numerical instabilities. The so called *Delaunay triangulation* ($\mathcal{DT}$) [52] meets this requirement. Thus given $\mathcal{S}$, its Delaunay triangulation is denoted by $\mathcal{DT}(\mathcal{S}) = \{\Delta_j\}_{j=1}^{J}$, and the union of all these subdivisions gives the *Convex Hull* ($\mathcal{CH}$) of $\mathcal{S}$, i.e. $\mathcal{CH}(\mathcal{S})$. Therefore, for a given sensor distribution, we define the monitored region Ω to be $\mathcal{CH}(\mathcal{S})$ and the convex hull boundary to be $\partial\Omega$ as shown in Figure 4.2. Notice that all N sensors $\{\mathbf{x}_n\}_{n=1}^{N}$ are in $\mathcal{CH}(\mathcal{S})$. Whereas only a subset of sensors[1] lie on the boundary $\partial\Omega$; we call these the boundary sensors and denote the set of all boundary sensors for a given network by $\mathcal{S}_{\partial\Omega}$. Given this construction, we can then retrieve an approximation of the family of integrals in (3.17) and (3.21), as follows:

Again let $\{\mathbf{v}_{j,j'} : j' = 1, 2, 3\}$ be the vertices of the triangular element Δ_j. Then the family of surface integrals in (3.17) are approximated as follows:

$$\int_{\Omega} \Psi_k(\mathbf{x})\dot{U}_r(\mathbf{x}, T)\,\mathrm{d}V \approx \frac{1}{3} \sum_{j=1}^{J} \sum_{j'=1}^{3} \Psi_k(\mathbf{v}_{j,j'})\Phi_{j,j',r}(t_L)|\Delta_j|, \tag{4.7}$$

where $\Phi_{j,j',r}(t_L) \stackrel{\text{def}}{=} \dot{U}_r(\mathbf{v}_{j,j'}, t_L)$ is the properly weighted, time-integrated measurement of the sensor situated at the vertex $\mathbf{v}_{j,j'}$ over the time interval $t \in [0, T]$ with $T = t_L$. In

[1] The red ones in Figure 4.2 for instance.

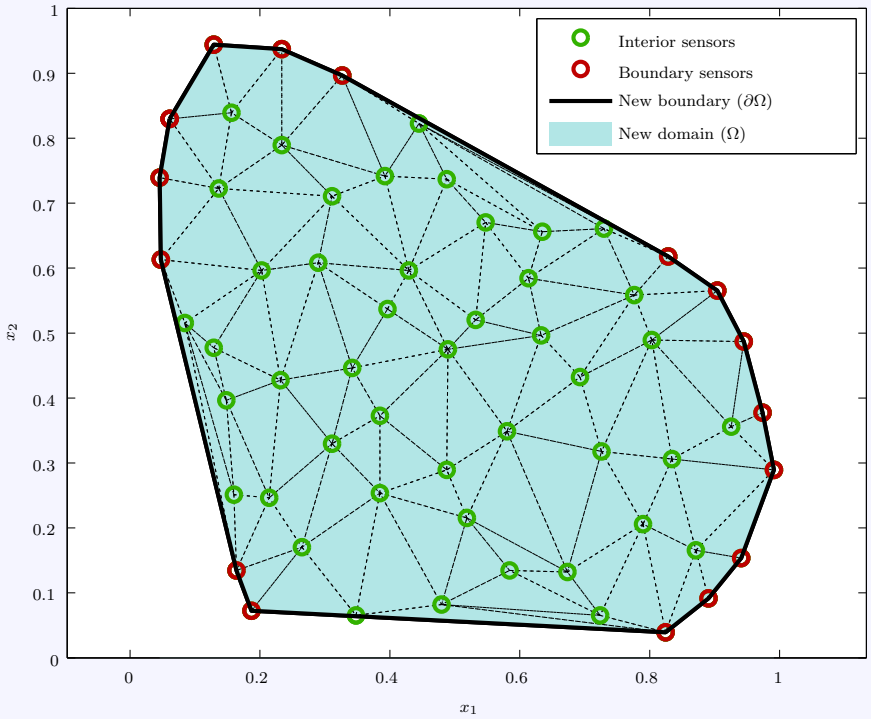


Figure 4.2.: **A sensor network and its Delaunay triangulation.** The (approximate) monitored domain Ω divided into triangular meshes and the domain boundary divided into straight line segments (black solid lines).

particular this specific time-integrated measurement is given by (4.4b) and can be simplified to,

$$\begin{aligned}\dot{U}_r(\mathbf{v}_n, T) &= \Gamma_r(T)u(\mathbf{x}_n, T) - \int_0^T \frac{\partial \Gamma_r(t)}{\partial t} u(\mathbf{x}_n, t)\mathrm{d}t \\ &= \Gamma_r(T)u(\mathbf{x}_n, T) + \frac{\mathrm{j}r}{T}\int_0^T \Gamma_r(t)u(\mathbf{x}_n, t)\mathrm{d}t \\ &= \Gamma_r(T)u(\mathbf{x}_n, T) + \frac{\mathrm{j}r}{T}U_r(\mathbf{x}_n, T). \end{aligned} \tag{4.8}$$

The second equality follows by recalling that for our temporal sensing function $\Gamma_r(t) = e^{-\mathrm{j}rt/T}$, then $\frac{\partial \Gamma_r(t)}{\partial t} = -\frac{\mathrm{j}r}{T}\Gamma_r(t)$ holds true. Whereas the last equality is obtained by substituting (4.4a). Moreover, we can approximate the time integral using trapezoidal rule for example,

$$\begin{aligned} U_r(\mathbf{x}_n, T)|_{T=t_L} &= \int_0^{t_L} \Gamma_r(t)u(\mathbf{x}_n, t)\mathrm{d}t \\ &\approx \left[\frac{1}{2}\left(\varphi_n(t_0)\Gamma_r(t_0) + \varphi_n(t_L)\Gamma_r(t_L)\right) + \sum_{l=1}^{L-1}\varphi_n(t_l)\Gamma_r(t_l)\right]\Delta_t. \end{aligned}$$

$$= \left[\frac{1}{2}\varphi_n(t_L)\Gamma_r(t_L) + \sum_{l=1}^{L-1}\varphi_n(t_l)\Gamma_r(t_l)\right]\Delta_t$$
$$\stackrel{\text{def}}{=} \Phi_{n,r}(t_L). \tag{4.9}$$

From this, an approximation for (4.8) (and of course (4.4b)) is,

$$\dot{U}_r(\mathbf{x}_n, T)\Big|_{T=t_L} \approx \Gamma_r(t_L)\varphi_n(t_L) + \frac{\mathrm{j}r}{t_L}\Phi_{n,r}(t_L) \stackrel{\text{def}}{=} \dot{\Phi}_{n,r}(t_L). \tag{4.10}$$

Similarly we can also derive a suitable approximation formula for the line integral, $\oint_{\partial\Omega}(\Psi_k\nabla U_r - U_r\nabla\Psi_k)\cdot\hat{\mathbf{n}}_{\partial\Omega}\,\mathrm{d}S$. First $\nabla U_r(\mathbf{x}_n, T) = \left[\frac{\partial}{\partial x_1}U_r(\mathbf{x}, T),\ \frac{\partial}{\partial x_2}U_r(\mathbf{x}, T)\right]^{\mathsf{T}}\Big|_{\mathbf{x}=\mathbf{x}_n}$ can be obtained using a first order central finite difference scheme, such that $\nabla U_r(\mathbf{x}_n, T) \approx \left[(\Phi_{n,r}(T))_{x_1}, (\Phi_{n,r}(T))_{x_2}\right]^{\mathsf{T}}$, where $(\Phi_{n,r}(T))_{x_1}$ and $(\Phi_{n,r}(T))_{x_2}$ are used to denote the approximation of the field's spatial derivatives at the specific sensor location $\mathbf{x}_n$. Moreover given our choice of $\Psi_k(\mathbf{x})$, $\nabla\Psi_k(\mathbf{x}_n) = -k\Psi_k(\mathbf{x}_n)\,[1, \mathrm{j}]^{\mathsf{T}}$. Hence,

$$\Psi_k(\mathbf{x}_n)\nabla U_r(\mathbf{x}_n, T) - U_r(\mathbf{x}_n, T)\nabla\Psi_k(\mathbf{x}_n) \approx \Psi_k(\mathbf{x}_n)\begin{bmatrix} (\Phi_{n,r}(T))_{x_1} + k\Phi_{n,r}(T) \\ (\Phi_{n,r}(T))_{x_2} + \mathrm{j}k\Phi_{n,r}(T)\end{bmatrix}. \tag{4.11}$$

Finally for $\hat{\mathbf{n}}_{\partial\Omega}\,\mathrm{d}S$, let us assume $\mathcal{S}_{\partial\Omega} = \{\mathbf{x}_1, \ldots, \mathbf{x}_{n-1}, \mathbf{x}_n, \ldots, \mathbf{x}_I\}$ is the set of boundary sensors such that $\mathbf{x}_n \in \partial\Omega$ for any $n = 1, \ldots, I$. Furthermore assume it is a cyclically ordered set, arranged in an anticlockwise order, then:

$$\hat{\mathbf{n}}_{\partial\Omega} \approx \frac{1}{\|\mathbf{x}_n - \mathbf{x}_{n-1}\|}\begin{pmatrix} x_{2,n} - x_{2,n-1} \\ x_{1,n-1} - x_{1,n}\end{pmatrix}$$

and

$$\mathrm{d}S \approx \|\mathbf{x}_n - \mathbf{x}_{n-1}\|.$$

Therefore,

$$\begin{aligned}&\oint_{\partial\Omega}(\Psi_k\nabla U_r - U_r\nabla\Psi_k)\cdot\hat{\mathbf{n}}_{\partial\Omega}\,\mathrm{d}S \\ &\approx \sum_{\mathbf{x}_n\in\mathcal{S}_{\partial\Omega}}\Psi_k(\mathbf{x}_n)\Big[\Big((\Phi_{n,r}(t_L))_{x_1} + k\Phi_{n,r}(t_L)\Big)(x_{2,n} - x_{2,n-1}) \\ &\qquad + \Big((\Phi_{n,r}(t_L))_{x_2} + \mathrm{j}k\Phi_{n,r}(t_L)\Big)(x_{1,n-1} - x_{1,n})\Big].\end{aligned} \tag{4.12}$$

Note that due to the cyclic ordering of the boundary sensors, $\mathbf{x}_0 = \mathbf{x}_I$.

Combining (4.7) with (4.12) and using the discrete sensor measurements $\{\varphi_n(t_l)\}_{n,l}$ of

the field, we conclude that the desired generalized measurements are approximately:

$$\begin{aligned}\mathcal{Q}(k,r) \approx &\frac{1}{3}\sum_{j=1}^{J}\sum_{j'=1}^{3}\Psi_k(\mathbf{v}_{j,j'})\dot{\Phi}_{j,j',r}(t_L)|\Delta_j| \\ &- \mu \sum_{\mathbf{x}_n\in\mathcal{S}_{\partial\Omega}} \Psi_k(\mathbf{x}_n)\Big[\Big(\left(\Phi_{n,r}(t_L)\right)_{x_1} + k\Phi_{n,r}(t_L)\Big)(x_{2,n}-x_{2,n-1}) \\ &\qquad + \Big(\left(\Phi_{n,r}(t_L)\right)_{x_2} + \mathrm{j}k\Phi_{n,r}(t_L)\Big)(x_{1,n-1}-x_{1,n})\Big].\end{aligned} \tag{4.13}$$

4.3. Robust centralized source estimation algorithms

Having computed the sequence of generalized measurements $\{\mathcal{Q}(k,r)\}_{k,r}$ from $\{\varphi_n(t_l)\}_{n,l}$, using (4.13) for example, we could apply Prony's method to it (or its properly modified version depending on the source model assumed – see Table 3.1) in order to recover the unknown source parameters and hence solve the Problems 4.1 to 4.4. However in a realistic SN setting $\varphi_n(t_l)$ will be noisy, meaning that at the FC we will have access to:

$$\varphi^{\epsilon}_{n,l} = \varphi_n(t_l) + \epsilon_{n,l},$$

where we assume that $\epsilon_{n,l} \sim \mathcal{N}(0,\sigma^2)$. From these noisy samples we can only obtain a perturbed sequence $\{\mathcal{Q}_\epsilon(k,r)\}_k$ as the estimate of the true generalized measurements $\{\mathcal{Q}(k,r)\}_k$. The perturbed sequence $\mathcal{Q}_\epsilon(k,r)$ is precisely given by,

$$\begin{aligned}\mathcal{Q}_\epsilon(k,r) = &\frac{1}{3}\sum_{j=1}^{J}\sum_{j'=1}^{3}\Psi_k(\mathbf{v}_{j,j'})\dot{\Phi}^{\epsilon}_{j,j',r}(t_L)|\Delta_j| \\ &- \mu \sum_{\mathbf{x}_n\in\mathcal{S}_{\partial\Omega}} \Psi_k(\mathbf{x}_n)\Big[\Big(\left(\Phi^{\epsilon}_{n,r}(t_L)\right)_{x_1} + k\Phi^{\epsilon}_{n,r}(t_L)\Big)(x_{2,n}-x_{2,n-1}) \\ &\qquad + \Big(\left(\Phi^{\epsilon}_{n,r}(t_L)\right)_{x_2} + \mathrm{j}k\Phi^{\epsilon}_{n,r}(t_L)\Big)(x_{1,n-1}-x_{1,n})\Big],\end{aligned} \tag{4.14}$$

where

$$\begin{aligned}\Phi^{\epsilon}_{n,r}(t_L) &\stackrel{\text{def}}{=} \left[\frac{1}{2}\varphi^{\epsilon}_n(t_L)\Gamma_r(t_L) + \sum_{l=1}^{L-1}\varphi^{\epsilon}_n(t_l)\Gamma_r(t_l)\right]\Delta_t, \\ \dot{\Phi}^{\epsilon}_{n,r}(t_L) &\stackrel{\text{def}}{=} \Gamma_r(t_L)\varphi^{\epsilon}_n(t_L) + \frac{\mathrm{j}r}{t_L}\Phi^{\epsilon}_{n,r}(t_L),\end{aligned}$$

and $\left[\left(\Phi^{\epsilon}_{n,r}(T)\right)_{x_1}, \left(\Phi^{\epsilon}_{n,r}(T)\right)_{x_2}\right]^{\mathsf{T}}$ denotes the 2-D central finite difference approximation of $\nabla U_r(\mathbf{x},T)$ at $\mathbf{x}=\mathbf{x}_n$, from the noisy quantities $\{\Phi^{\epsilon}_{n,r}(t_L)\}_n$ found above.

Therefore since robustness and noise resilience are paramount we explore two preprocessing techniques, in the sequel, based on Cadzow's algorithm [35], to handle any noise and other sources of model mismatch (due, for example, to our approximation schemes), whilst estimating the unknown parameters from $\{\mathcal{Q}_\epsilon(k,r)\}_k$.

4.3.1. Handling noise and model mismatch: Subspace-based denoising of $\mathcal{Q}_\epsilon(k,r)$

Given access to $K > 2M-1$ consecutive terms of the generalized sequence, $\{\mathcal{Q}_\epsilon(k,r)\}_{k=0}^K$,[2] we denoise it using Cadzow's algorithm [35] for some fixed integer r ($r = 0$ or $r = 1$ for example). The basic idea of the method is as follows: when applying Prony's method on the sequence, we build a Toeplitz matrix $\mathbf{T}_\epsilon \in \mathbb{R}^{L_1 \times L_2}$ (which is as square as possible) of the form,

$$\mathbf{T}_\epsilon = \begin{pmatrix} \mathcal{Q}_\epsilon(L_2,r) & \mathcal{Q}_\epsilon(L_2-1,r) & \cdots & \mathcal{Q}_\epsilon(0,r) \\ \mathcal{Q}_\epsilon(L_2+1,r) & \mathcal{Q}_\epsilon(L_2,r) & \cdots & \mathcal{Q}_\epsilon(1,r) \\ \vdots & \vdots & \ddots & \vdots \\ \mathcal{Q}_\epsilon(K,r) & \mathcal{Q}_\epsilon(K-1,r) & \cdots & \mathcal{Q}_\epsilon(L_1,r) \end{pmatrix}, \tag{4.15}$$

where $L_2 = \lceil K/2 \rceil$ and $L_1 = K - L_2$. Moreover as highlighted in Section 2.2.2, in the absence of noise, the rank of $\mathbf{T}_\epsilon$ is exactly the number of sources/vertices M. However, noisy sensor measurements and the approximation of integrals using finite sums lead to model mismatches which subsequently makes $\mathbf{T}_\epsilon$ full rank. Cadzow's algorithm denoises $\mathbf{T}_\epsilon$ by first setting to zero the $L_1 - M$ smallest singular values of $\mathbf{T}_\epsilon$ which are typically due to noise so as to obtain a rank M matrix. The next step of the algorithm is to enforce the Toeplitz structure by averaging along the diagonals of the reconstructed low-rank matrix. The method is iterated a number of times. The end result of applying Cadzow to $\mathbf{T}_\epsilon$ is, therefore, to denoise $\{\mathcal{Q}_\epsilon(k,r)\}_{k=0}^K$ for a chosen integer r. We stress that we first of all fix r, say $r = 0$, before performing the denoising iterations.

4.3.2. Non-instantaneous source activation time & decay coefficient estimation with Cadzow

In this section, we outline a scheme to further improve certain estimates of the unknown non-instantaneous source parameters in the presence of perturbations. To achieve this we will need to notice that the generation of (3.23), (3.24) and (3.25) using equally spaced subintervals ΔT suggests deeper underlying connections with Prony's method and its variations. We will explore this relationship and show how to further improve noise robustness using Cadzow's denoising algorithm.

We begin by defining

$$\begin{aligned} D_p &\stackrel{\text{def}}{=} c'_m(T_{p+1}) - c'_m(T_p) \\ &= \frac{c_m}{\alpha_m}\left(e^{\alpha_m(T_p-\tau_m)}e^{\alpha_m \Delta T} - e^{\alpha_m(T_p-\tau_m)}\right) \\ &= \frac{2c_m}{\alpha_m} e^{\alpha_m \frac{\Delta T}{2}} e^{\alpha_m(T_p-\tau_m)} \sinh(\alpha_m \Delta T/2), \end{aligned} \tag{4.16}$$

[2] Similarly $\{-k\mathcal{Q}_\epsilon(k,r)\}_{k=1}^K$ with $K > 4$ for the line source; or $\{\mathcal{Q}_\epsilon(k,r)\}_{k=1}^K$ with $K \geq 2M$ for the M-sided polygonal source

for positive integer values of $p = 1, 2, \ldots, P$. Then it immediately follows that:

$$\begin{aligned}
\mathcal{A}(p,q) &\stackrel{\text{def}}{=} \frac{D_p}{D_{p+q}} = \frac{c'_m(T_{p+1}) - c'_m(T_p)}{c'_m(T_{p+q+1}) - c'_m(T_{p+q})} \\
&= \frac{\frac{2c_m}{\alpha_m} e^{\alpha_m \frac{\Delta T}{2}} e^{\alpha_m (T_p - \tau_m)} \sinh(\alpha_m \Delta T/2)}{\frac{2c_m}{\alpha_m} e^{\alpha_m \frac{\Delta T}{2}} e^{\alpha_m (T_{p+q} - \tau_m)} \sinh(\alpha_m \Delta T/2)} \\
&= \frac{e^{\alpha_m (T_p - \tau_m)}}{e^{\alpha_m (T_{p+q} - \tau_m)}} \\
&= e^{q \Delta T \alpha_m},
\end{aligned} \tag{4.17}$$

where $q = 0, 1, \ldots, P-2$.

We stress that the terms $\mathcal{A}(p,q)$ $\forall p, q$ are given by $\mathcal{A}(p,q) = \frac{c'_m(T_{p+1}) - c'_m(T_p)}{c'_m(T_{p+q+1}) - c'_m(T_{p+q})}$ where the $c'_m(T_i)$'s are computed by applying Prony's method to the sequence $\{\mathcal{R}(k)\}_k$ (satisfying (3.20)) computed for a given time window $t \in [0, T_i]$. Moreover, notice that for any fixed q the terms $\mathcal{A}(p,q)$, $p = 1, 2, \ldots, P-q-1$ are equal. This is true only in the ideal scenario when there exists no model mismatches in the system, but is false in the presence of noise and other model imperfections. Assuming that the perturbations can be modelled as an approximately i.i.d process, then taking the average should give a better estimate. As such we can form a new sequence as follows:

$$\mathcal{B}(q) = \frac{1}{P-q-1} \sum_{p=1}^{P-q-1} \mathcal{A}(p,q). \tag{4.18}$$

It is easy to see that $\mathcal{B}(q)$ still satisfies $\mathcal{B}(q) = e^{q \Delta T \alpha_m}$, where $q = 0, 1, \ldots, P-2$. This is again a sequence where we can apply Prony's method, but with a single unknown, hence it admits a solution when $P - 2 \geq 1 \Rightarrow P \geq 3$. Moreover, if $P > 3$ we can also apply Cadzow's denoising algorithm to the sequence $\{\mathcal{B}(q)\}_{q=0}^{P-2}$, in much the same way as we did in Section 4.3.1 on $\{\mathcal{Q}(k,r)\}_k$.

We can now outline the complete joint centralized estimation scheme with subspace denoising for localized sources in Algorithm 4.1, whilst the corresponding schemes for line and polygonal sources are listed in Algorithm 4.2 and Algorithm 4.3 respectively. These algorithms solve discrete IDSPs of interest to us.

4.3.3. Sequential estimation of multiple localized sources

Algorithm 4.1 can be readily used to jointly estimate multiple localized sources of diffusion fields from arbitrary field samples; and this approach works both in the case where sources become active simultaneously (i.e. sources at distinct locations with equal activation times) and also in the case of sequential activation. In the latter scenario, however, it is more effective to estimate one source at a time and then removing its contribution from the sensor measurements before estimating the next source. This is possible when: the sources have suitably distinct activation times; so that, the sampling interval is small enough to

Algorithm 4.1 Simultaneous Estimation of M point sources from field samples

Require: $\{\varphi_n(t_l)\}_{n=1,l=0}^{N,L}$, sensor locations $\{\mathbf{x}_n\}_n$, number of sources M, sampling interval Δ_t, diffusivity μ, source temporal model STM.

Retrieve the *convex hull* ($\mathcal{CH}$) of the set of points $\{\mathbf{x}_n\}$. $\mathcal{CH}$ and its boundary define Ω and $\partial\Omega$ respectively, in (3.17) and (3.21).
Initialize $K \geq 2M - 1$.
if SM == '*instantaneous*' **then**
    Compute the sequence $\{\mathcal{Q}(k,1)\}_{k=0}^{K}$ using (4.13) as explained in Section 4.2.
    Denoise $\{\mathcal{Q}(k,1)\}_{k=0}^{K}$ using Cadzow's algorithm.
    Apply Prony's method to $\{\mathcal{Q}(k,1)\}_{k=0}^{K}$ to obtain M pairs of $(c_m e^{-\mathrm{j}\tau_m/T}, \boldsymbol{\xi}_m)$.
    For each m, $c_m = \left|c_m e^{-\mathrm{j}\tau_m/T}\right|$ and $\tau_m = T \arg\left(c_m e^{-\mathrm{j}\tau_m/T}\right)$.
    return M intensities, activation times and locations $\{c_m, \tau_m, \boldsymbol{\xi}_m\}_{m=1}^{M}$.
else
    Initialize $P \geq 3$.
    Initialize ΔT so that $\Delta T \leq T/P$.
    for $i = 1 : P$ **do**
        $T_p = T - (P - p)\Delta T$.
        Using measurements over $\Omega \times [0, T_p]$ compute the sequence $\{\mathcal{R}(k) \stackrel{\text{def}}{=} \mathcal{Q}(k,0)\}_{k=0}^{K}$ using (4.13) with $r = 0$.
        Denoise $\{\mathcal{R}(k)\}_{k=0}^{K}$ using Cadzow's algorithm.
        Apply Prony's method to $\{\mathcal{R}(k)\}_{k=0}^{K}$ to obtain M pairs $(c'_m(T_p), \boldsymbol{\xi}_m)$.
    end for
    By matching the locations form M such sequences: $(\{c'_m(T_p)\}_{p=1}^{P}, \boldsymbol{\xi}_m)$.
    for $m = 1, \ldots, M$ **do**
        Construct $\{\mathcal{B}(q)\}$ from $\{c'_m(T_p)\}_{p=1}^{P}$ using (4.18).
        Denoise $\{\mathcal{B}(q)\}$ using Cadzow.
        For α_m apply Prony's method to the denoised sequence.
        Find τ_m and c_m using (3.31) and (3.32) respectively.
    end for
    return M source decay coefficients, intensities, activation times and locations: $\{\alpha_m, c_m, \tau_m, \boldsymbol{\xi}_m\}_{m=1}^{M}$.
end if

Algorithm 4.2 Single line source estimation algorithm from samples

Require: $\{\varphi_n(t_l)\}_{n=1,l=0}^{N,L}$, sensor locations $\{\mathbf{x}_n\}_n$, sampling interval Δ_t, diffusivity μ, source temporal model STM.

Retrieve the *convex hull* ($\mathcal{CH}$) of the set of points $\{\mathbf{x}_n\}$. $\mathcal{CH}$ and its boundary define Ω and $\partial\Omega$ respectively, in (3.17) and (3.21).
Initialize $K \geq 4$ and $T = t_L$.
if STM == '*instantaneous*' **then**
    Compute the sequence $\{\mathcal{Q}(k,1)\}_{k=1}^{K}$ using (4.13) as explained in Section 4.2.
    From $\{\mathcal{Q}(k,1)\}_{k=1}^{K}$, for each $k \in 1,\ldots,K$ form the new sequence $\{-k\mathcal{Q}(k,1)\}_{k=1}^{K}$.
    Denoise $\{-k\mathcal{Q}(k,1)\}_{k=1}^{K}$ using Cadzow's algorithm.
    Apply Prony's method to $\{-k\mathcal{Q}(k,1)\}_{k=1}^{K}$ to reveal $(c\ell(\boldsymbol{\xi}_1,\boldsymbol{\xi}_2)e^{-\mathrm{j}\tau/T},\boldsymbol{\xi}_1,\boldsymbol{\xi}_2)$.
    Compute the value of $\ell(\boldsymbol{\xi}_1,\boldsymbol{\xi}_2)$ using (3.37), hence $c = \left|\frac{c\ell(\boldsymbol{\xi}_1,\boldsymbol{\xi}_2)e^{-\mathrm{j}\tau/T}}{\ell(\boldsymbol{\xi}_1,\boldsymbol{\xi}_2)}\right|$ and $\tau = -T\arg\left(\frac{c\ell(\boldsymbol{\xi}_1,\boldsymbol{\xi}_2)e^{-\mathrm{j}\tau/T}}{\ell(\boldsymbol{\xi}_1,\boldsymbol{\xi}_2)}\right)$.
    return intensity, activation time and endpoints $(c,\tau,\boldsymbol{\xi}_1,\boldsymbol{\xi}_2)$ of the line source.
else
    Initialize $P \geq 3$.
    Initialize ΔT so that $\Delta T \leq T/P$.
    for $i = 1:P$ **do**
        $T_p = T - (P-p)\Delta T$.
        Using measurements over $\Omega \times [0,T_p]$ compute sequence $\{\mathcal{R}(k) \stackrel{\text{def}}{=} \mathcal{Q}(k,0)\}_{k=1}^{K}$ using (4.13) with $r = 0$.
        For each $k = 1,\ldots,K$ multiply corresponding term in $\{\mathcal{R}(k)\}_{k=1}^{K}$ by $-k$ to get $\{-k\mathcal{R}(k)\}_{k=1}^{K}$.
        Denoise $\{-k\mathcal{R}(k)\}_{k=1}^{K}$ using Cadzow's algorithm.
        Apply Prony's method to $\{-k\mathcal{R}(k)\}_{k=0}^{K}$ to obtain $(c'(T_p)\ell(\boldsymbol{\xi}_1,\boldsymbol{\xi}_2),\boldsymbol{\xi}_1,\boldsymbol{\xi}_2)$.
    end for
    Compute $\ell(\boldsymbol{\xi}_1,\boldsymbol{\xi}_2)$ using (3.37) and divide the above generalized energies by it to get $\{c'(T_1),c'(T_2),\ldots,c'(T_P)\}$.
    Construct $\{\mathcal{B}(q)\}_{q=0}^{P-2}$ from $\{c'(T_p)\}_{p=1}^{P}$ using (4.18).
    Denoise $\{\mathcal{B}(q)\}_{q=0}^{P-2}$ using Cadzow.
    Apply Prony's to the denoised sequence to obtain α.
    Recover τ, and c from $\{c'(T_p)\}_{p=1}^{P}$ using (3.31) and (3.32) respectively.
    return decay coefficient, intensity, activation time and endpoints $(\alpha,c,\tau,\boldsymbol{\xi}_1,\boldsymbol{\xi}_2)$ of the line source.
end if

Algorithm 4.3 Single M-sided polygonal source estimation from samples

Require: $\{\varphi_n(t_l)\}_{n=1,l=0}^{N,L}$, sensor locations $\{\mathbf{x}_n\}_n$, number of edges M, sampling interval Δ_t, diffusivity μ, source temporal model STM.

Retrieve the *convex hull* ($\mathcal{CH}$) of the set of points $\{\mathbf{x}_n\}$. $\mathcal{CH}$ and its boundary define Ω and $\partial\Omega$ respectively, in (3.17) and (3.21).
Initialize $K \geq 2M$ and $T = t_L$.
if STM == '*instantaneous*' **then**
    Compute the sequence $\{\mathcal{Q}(k,1)\}_{k=1}^{K}$ using (4.13) as explained in Section 4.2.
    Multiply corresponding term in $\{\mathcal{Q}(k,1)\}_{k=1}^{K}$ by k^2, for each $k = 1, 2, \ldots, K$ to get $\{k^2\mathcal{Q}(k,1)\}_{k=1}^{K}$.
    Denoise $\{k^2\mathcal{Q}(k,1)\}_{k=1}^{K}$ using Cadzow's algorithm.
    Apply Prony's method to $\{k^2\mathcal{Q}(k,1)\}_{k=1}^{K}$ to obtain M pairs $(ca_m e^{-\mathrm{j}\tau/T}, \boldsymbol{\xi}_m)$.
    Evaluate a_m.
    Hence $c = \frac{1}{M}\sum_{m=1}^{M}\left|\frac{ca_m e^{-\mathrm{j}\tau/T}}{a_m}\right|$ and $\tau = -\frac{T}{M}\sum_{m=1}^{M}\arg\left(\frac{ca_m e^{-\mathrm{j}\tau/T}}{a_m}\right)$.
    return intensity, activation time and M vertices $\left(c, \tau, \{\boldsymbol{\xi}_m\}_{m=1}^{M}\right)$ of polygonal source.
else
    Initialize $P \geq 3$.
    Initialize ΔT so that $\Delta T \leq T/P$.
    for $i = 1 : P$ **do**
        $T_p = T - (3-p)\Delta T$.
        Using measurements over $\Omega \times [0, T_p]$ compute the sequence $\{\mathcal{R}(k) \stackrel{\text{def}}{=} \mathcal{Q}(k,0)\}_{k=1}^{K}$ using (4.13) with $r = 0$.
        For each $k = 1, \ldots, K$ multiply corresponding term in $\{\mathcal{R}(k)\}_{k=1}^{K}$ by k^2 to get $\{k^2\mathcal{R}(k)\}_{k=1}^{K}$.
        Denoise $\{k^2\mathcal{R}(k)\}_{k=1}^{K}$ using Cadzow's algorithm.
        Apply Prony's method to $\{k^2\mathcal{R}(k)\}_{k=0}^{K}$ to obtain M pairs of $\{c'(T_i)a_m, \boldsymbol{\xi}_m\}_{m=1}^{M}$.
    end for
    Compute a_m and divide generalized energies above by it to get $\{c'(T_p)\}_{p=1}^{P}$.
    Construct $\{\mathcal{B}(q)\}_{q=0}^{P-2}$ from $\{c'(T_p)\}_{p=1}^{P}$ using (4.18).
    Denoise $\{\mathcal{B}(q)\}_{q=0}^{P-2}$ using Cadzow.
    Apply Prony's to the denoised sequence to obtain α.
    Recover τ, and c from $\{c'(T_p)\}_{p=1}^{P}$ using (3.31) and (3.32) respectively.
    return decay coefficient, intensity, activation time and vertices $\left(\alpha, c, \tau, \{\boldsymbol{\xi}_m\}_{m=1}^{M}\right)$ of the polygonal source.
end if

resolve the activation of any two consecutive sources. In such a scenario, we propose the following approach. Firstly, we find a time window over which only a single source is active. We can achieve this by examining the rank of the Toeplitz matrix constructed from $\{\mathcal{Q}_\epsilon(k,r)\}_k$. We then estimate the source parameters as described in Algorithm 4.1 (with $M = 1$). Given these preliminary estimates, a selection of sensor measurements collected in the neighbourbood of the estimated source location are used to obtain a more precise estimate for the activation time, as well as the decay coefficient if we are estimating a non-instantaneous source model. These sharper estimates are obtained by performing a simple local search around the initial estimates of τ_m (and α_m if necessary), as follows:

Instantaneous sources: activation time refinement

Given the initial estimate of the intensity, location and activation time, $\hat{c}$, $\hat{\boldsymbol{\xi}}$ and $\hat{\tau}$, respectively, consider the measurements $\{\varphi_n^\epsilon(t_l)\}_{l=0}^L$ collected by the n-th sensor (located at $\mathbf{x}_n$) and the re-synthesized sequence $\hat{\varphi}_n(t_l)^* = \hat{u}^*(\mathbf{x}_n, t_l) = \frac{\hat{c}}{4\pi\mu(t_l-\hat{\tau}^*)} e^{-\frac{\|\mathbf{x}_n-\hat{\boldsymbol{\xi}}\|^2}{4\mu(t_l-\hat{\tau}^*)}} H(t-\hat{\tau}^*)$. By comparing the normalized inner-product between the reconstructed sequence and the measurements $\{\varphi_n^\epsilon(t_l)\}_{l=0}^L$, we choose the $\hat{\tau}^* \in [\delta_\tau\hat{\tau}, \frac{1}{\delta_\tau}\hat{\tau}]$ where $\delta_\tau \in (0,1]$ that maximizes this normalized inner product – a modification of the Cauchy-Schwarz inequality for vectors.

Non-instantaneous sources: activation time & delay coefficient refinement

Again we assume a single source field and the initial estimates $\hat{c}, \hat{\alpha}, \hat{\boldsymbol{\xi}}$ and $\hat{\tau}$ for the source parameters. The measured field $\{\varphi_n^\epsilon(t_l)\}$ is compared with the reconstructed field $\{\hat{\varphi}_n^*(t_l)\}$ to obtain better estimates of α and τ. In this case however, we perform a local 2D search over $\hat{\tau}^* \in [\delta_\tau\hat{\tau}, \frac{1}{\delta_\tau}\hat{\tau}]$ and $\hat{\alpha}^* \in [\delta_\alpha\hat{\alpha}, \frac{1}{\delta_\alpha}\hat{\alpha}]$ where $\delta_\tau, \delta_\alpha \in (0,1]$ are some constants.

We perform the same search using the $\beta \in \mathbb{N}$ sensors closest to the estimated source location, and then obtain a final estimate for the activation time and decay coefficient by averaging the estimates obtained from the measurements due to the β selected sensors. The complete sequential method is summarized in Algorithm 4.4.

Remark 4.1. *The strategy of selecting the β closest sensors to the estimated field is implicitly noise reducing, as these sensors will, in general, have a higher* SNR *since the field intensity is greater at these locations (close to the source) whilst all sensors experience the same noise power.*

4.4. Distributed sensor network model

In contrast to the centralized network, the distributed SN in this work will comprise N *smart* sensor nodes that, in addition to local field sensing, can each perform typical mathematical operations on their field measurements (and on any received data). However due to power considerations the sensors no longer have long-range transmission capabilities.

Algorithm 4.4 Robust sequential estimation of M point sources

Require: $\{\varphi_n(t_l)\}_{n=1,l=0}^{N,L}$, sensor locations $\{\mathbf{x}_n\}_n$, number of edges M, sampling interval Δ_t, diffusivity μ, source temporal model STM.

1: Retrieve $\mathcal{CH}(\{\mathbf{x}_n\}_{n=1}^N)$.
2: Let $m \leftarrow 0$ and the number of valid sources $M_{vs} \leftarrow 0$.
3: **while** $m < M$ **do**
4: Construct $\{\mathcal{R}(k)\}$ with $K \geq 2M' - 1$ and $M' \geq 2$.
5: Estimate the M' generalized energy-location pairs $\{\sigma'_{m'}, \boldsymbol{\xi}'_{m'}\}_{m'=1}^{M'}$.
6: Set M_{vs} to be the number of pairs of $\{\sigma'_{m'}, \boldsymbol{\xi}'_{m'}\}$ having both $\sigma'_{m'}$ greater than some predefined threshold and $\boldsymbol{\xi}'_{m'} \in \mathcal{CH}$.
7: **if** $M_{vs} > 1$ **then**
8: Decrease window size T and Go to 4.
9: **else if** $M_{vs} < 1$ **then**
10: Increase window size T and Go to 4.
11: **else if** $M_{vs} == 1$ **then**
12: Estimate source parameters $c_m, \boldsymbol{\xi}_m, \hat{\tau}$ (and $\hat{\alpha}$ if appropriate) using Algorithm 4.1.
13: Select the $\beta \in \mathbb{N}$ nearest sensors to $\boldsymbol{\xi}_m$. For each of the β sensors, retrieve $\hat{\tau}_1^*, \ldots, \hat{\tau}_\beta^*$, and for time-varying sources $\hat{\alpha}_1^*, \ldots, \hat{\alpha}_\beta^*$ too, as described in Sections 4.3.3 and 4.3.3 respectively.
14: $\tau_m \leftarrow$Ave$\{\hat{\tau}_1^*, \ldots, \hat{\tau}_\beta^*\}$, $\alpha_m \leftarrow$Ave$\{\hat{\alpha}_1^*, \ldots, \hat{\alpha}_\beta^*\}$.
15: Reconstruct its field and adjust $\{\varphi_n(t_l)\}_{n,l}$.
16: $m \leftarrow m + 1$.
17: Increase window size T and Go to 4.
18: **end if**
19: **end while**
20: **return** source parameters $\{c_m, \alpha_m, \boldsymbol{\xi}, \tau_m\}_{m=1}^M$.

Therefore within the distributed setup, our intention is to estimate the source distribution f using sensor networks with underlying communication constraints; specifically we assume that: (i) each sensor can only communicate with a subset of neighbouring sensors; additionally, (ii) the communication links are assumed to have finite bandwidth. These constraints imply that what we seek here are distributed estimation strategies for solving our class of discrete IDS problems, so that each sensor performs local data acquisition (senses the diffusion field), and then through localized data processing and communications (i.e. exchanging properly modified versions of its measurements with its neighbors) can estimate the unknown source parameterization, f, inducing the measured field. Thus eliminating the need for a FC.

To prevent long-range (power hungry) transmissions we assume that each sensor node within the distributed network is able to communicate only with those sensors that are within some fixed radius of it. As such we can model the sensor network as a *connected* random geometric graph (RGG), denoted $\mathcal{G}(N, r_{\text{con}})$, with N sensor nodes and connectivity radius r_{con}. This system may be realized by placing N nodes uniformly at random over a square region and then placing an edge between a pair of nodes if their Euclidean distance is at most r_{con}, as shown in Figure 4.3. In this figure, an edge between a pair of nodes

depicts the existence of a communication link between them.

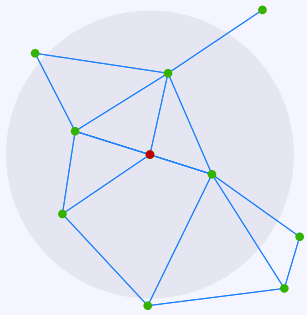


Figure 4.3.: **A sensor network example.** Links between sensors as modelled by a random geometric graph.

These communication links are assumed initially to be ideal; we then relax this assumption and consider the case when the links have finite capacity, so that the data exchanged between sensors need to be quantized. In this case, it is assumed that a capacity achieving communication scheme is used and the receiver is therefore able to recover the original message with zero error. Finally, we assume that upon deployment of the sensors a process is initiated whence:

(A) The sensors learn the topology of the network.

(B) They each compute the Delaunay triangulation, see Figure 4.4 for an example, such that we obtain a graph $\mathcal{G}_{del} = (\mathcal{V}, \mathcal{E})$, with the vertex set $\mathcal{V}$, corresponding to the locations of the sensors and $\mathcal{E}$ are the edges of the triangulation. Hence every sensor n knows if it lies, either on the convex hull boundary of the triangulation, or, in the interior of the convex hull.

(C) Given the triangulation, each sensor n also knows the total number (J_n) of triangles for which it is a vertex. Hence $\Delta_{n,j}$ is used to refer to the j-th triangle of the n-th sensor and $|\Delta_{n,j}|$ is the corresponding area of the triangle, with $j = 1, \ldots, J_n$.

The first assumption, (A), above means that each sensor knows its position relative to other sensors in the network, this is important when recovering the location information of the unknown diffusion sources. Whilst remaining assumptions '(B) and (C)', as we will demonstrate in Section 4.6, are required in order for the n-th sensor to be able to compute the weights with which to multiply its field measurements, before communicating its sum (i.e. a scalar value) to a neighbouring sensor. Under these assumptions, we wish to solve the discrete IDSPs stated in Problems 4.1 to 4.4—without the need for a fusion centre—through local in-network interactions between the monitoring sensors. To this end we first provide an overview of gossip algorithms for distributed average consensus in the following section and then demonstrate how to properly extend these gossip algorithms to solve our discrete IDSPs in Section 4.7.

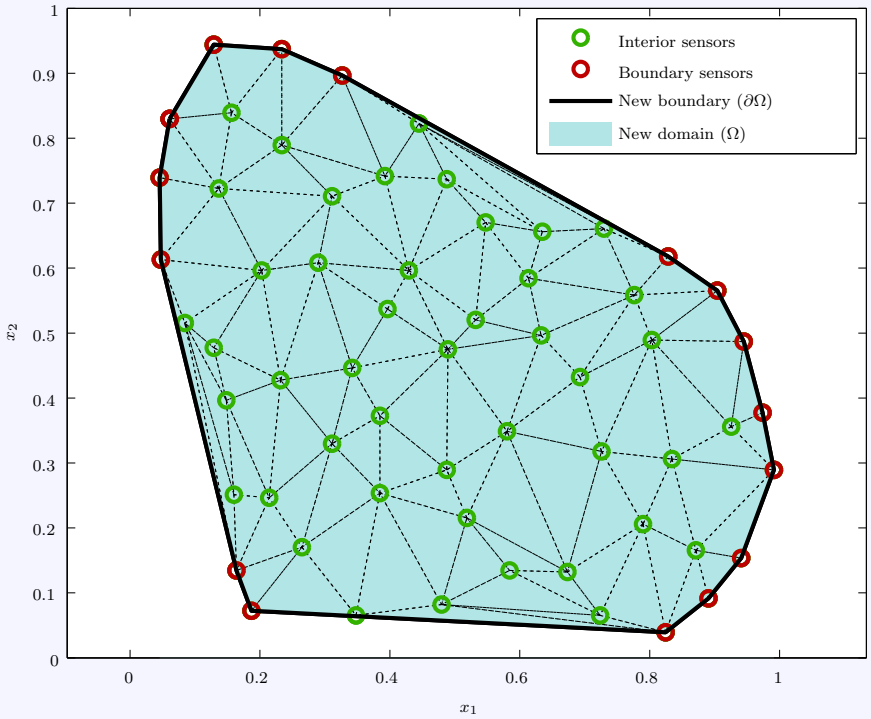


Figure 4.4.: **A sensor network and its Delaunay triangulation.** The (approximate) monitored domain Ω divided into triangular meshes and the domain boundary divided into straight line segments (black solid lines).

4.5. Gossip schemes for distributed average consensus

Gossiping [142, 55] is a distributed strategy for achieving consensus amongst agents in a network through a local exchange of data. Following the early works of [142] in the area, it has gained considerable interest for in-network processing in sensor networks as it mitigates the need for specialized routing protocols. In addition, gossip-based algorithms are robust to bottlenecks and link failures making it suitable for our distributed estimation problem. The results derived in this paper can be immediately extended to other gossiping schemes, such as those introduced in [15, 19]. An in depth survey of gossiping algorithms in sensor networks is given in [55], and analysis of general averaging sum-weight-like algorithms in WSNs can be found in [76].

For the purpose of demonstration, in our simulation results the archetypal *pairwise randomized gossip algorithm* [27] will be used. In this scheme, each node preserves an estimate of the sum and hence average of the node values. Let the value of node n after the i-th pairwise gossip round be $y_{n,i}$, hence $y_{n,0}$ is its initial value. In an iteration, a node n selected uniformly at random wakes up and contacts a randomly selected neighbor n' within its connectivity radius, and they both update their estimates by setting $y_{n,i+1} = y_{n',i+1} = (y_{n,i}+y_{n',i})/2$. Let $\mathbf{y}(i) = [y_{1,i}, y_{2,i}, \ldots, y_{N,i}]^\mathsf{T}$ be the vector of the values of the N agents in the network at the i-th gossip round, then this pairwise gossiping

operation can be summarized in the following way

$$\mathbf{y}(i+1) = \mathbf{P}(i)\mathbf{y}(i) \tag{4.19}$$

where $\mathbf{P}(i)$ are doubly stochastic matrices selected at random at the i-th iteration. For pairwise gossip algorithm, $\mathbf{P}(i)$ has entries such that elements $(n,n),(n,n'),(n',n)$ and (n',n') are equal to $1/2$ and $\mathbf{P}(i)$ is a diagonal identity elsewhere. Under this scheme it can be shown that, if a network (of N nodes) is connected and each pair of nodes communicate often enough the estimate at each node is guaranteed to converge to the global network average $\bar{y} = \frac{1}{N}\sum_{n=1}^{N} y_{n,0}$, i.e. $\lim_{i\to\infty} \mathbf{y}(i) = \mathbf{1}\bar{y}$. Performance guarantees and convergence results have also been studied (see [27] and the references therein).

The *localized interactions* in our field estimation setting will be based on the use of gossip algorithms for the distributed computation of a family of integrals whose final values can be used to reveal the unknown source parameters. In Section 4.6 we present our strategy for estimating these family of integrals, and hence recover the unknown source parameters, through the use of gossip.

4.5.1. Quantized gossip

When the interactions between the agents are over a channel with finite capacity, the messages exchanged needs to be quantized. It is then natural to wonder if the values of the agents can still achieve consensus and converge to $\bar{y}$ using the pairwise gossip scheme discussed previously. In what follows we briefly investigate this question and briefly overview some proposed solutions in the literature. The inter-agent communication is through a channel employing a uniform quantizer with quantization step δ_Q. This communication can be modelled by introducing the quantization map $Q_q : \mathbb{R} \mapsto \mathcal{Q}$ such that,

$$Q_q(y) = k\delta_Q, \quad \left(k - \frac{1}{2}\right)\delta_Q \leq y < \left(k + \frac{1}{2}\right)\delta_Q, \tag{4.20}$$

where $k \in \mathbb{Z}$ and $\mathcal{Q}$ is the set of permissible quantization levels, i.e. $\mathcal{Q} = \{Q_q(y) : k \in \mathbb{Z}\}$.

Average consensus via a Naive Quantized Gossip (NQG) scheme

This is a straightforward extension of the standard gossip consensus scheme, whereby after each (pairwise) exchange of messages the agents requantize their values and exchange this quantized value at the next iteration. It is summarized as follows,

$$\mathbf{y}(i+1) = \mathbf{P}(i)Q_q(\mathbf{y}(i)), \tag{4.21}$$

where i denotes the i-th iteration, Q_q denotes the mapping due to a uniform q-bit quantizer and $\mathbf{P}(i)$ is the transition matrix.

Under such a scheme, the agents in the network reach a consensus. However convergence results are poor in the sense that the consensus value can be far away from the true average

as is shown in Figure 4.5(a). The poor convergence result observed is known to be due to the loss of symmetry between neighbors in the use of received information. Consequently, there has been considerable research efforts put towards devising average consensus strategies for networks with finite communication capacity, where it becomes impossible for sensors to exchange real numbers and hence arrive at the real-valued network average $\bar{y}$.

We consider in what follows some common quantized consensus strategies which are well suited to our field estimation task.

Kashyap's Quantized consensus Gossip (KQG) scheme

In [82] Kashyap *et al* propose a quantized consensus algorithm that aims to preserve the network average at every iteration. They prove that the collection of values at each agent in the network will converge to a *quantized consensus distribution.* Specifically, each node will converge to either $\bar{Y}$ or $\bar{Y}+1$, where $\bar{Y} = \sum_{n=1}^{N} y_{n,0} \bmod N$, if the pair of communicating agents (n, n') use an update such as:

$$\begin{aligned} y_{n,i+1} &= \left\lceil \frac{y_{n,i} + y_{n',i}}{2} \right\rceil, \\ y_{n',i+1} &= \left\lfloor \frac{y_{n,i} + y_{n',i}}{2} \right\rfloor, \end{aligned} \tag{4.22}$$

where $y_{n,i} \leq y_{n',i}$, and $\lceil \cdot \rceil$, $\lfloor \cdot \rfloor$ is used to denote rounding up and down to the next quantizer level. We validate this approach through the simulation results shown in Figure 4.5(b). We refer to [82] for a theoretical analysis of convergence.

Average consensus via a Symmetric Quantized Gossip (SQG) scheme

Frasca *et al* [67] propose a scheme which aims to restore the symmetry lost by using an update scheme such as (4.21). This algorithm can be summarized as follows,

$$\mathbf{y}(i+1) = \mathbf{y}(i) + (\mathbf{P}(i) - \mathbf{I})Q_q(\mathbf{y}(i)) \tag{4.23}$$

where i denotes the i-th iteration, $Q_q(\cdot)$ is the mapping due to a uniform q-bit quantizer, $\mathbf{P}(i)$ is the diffusion (transition) matrix and $\mathbf{I}$ is the identity matrix.

Under such a scheme, the agents in the network achieve *quantized consensus* in the sense that the value at each node is at most one quantizer level away from the true average. Specifically, as well as the total sum of the states being preserved, the states of the agents in the network converge but to different values which are close to the true average, and in fact only differ from this true average by at most one bin. This can be seen in Figure 4.5(c) which shows the state evolution of the agents in the network assuming a 5-bit (32-Level) uniform quantizer.

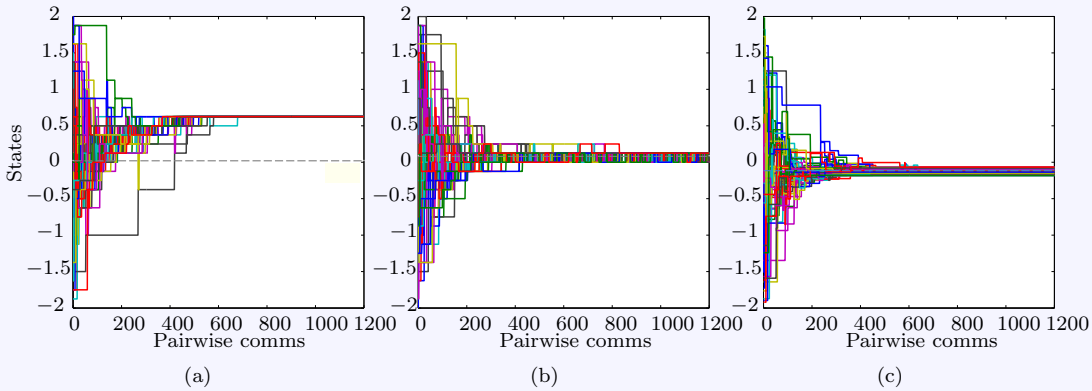


Figure 4.5.: **State evolution of quantized gossip schemes.** Shows at each gossip iteration the state evolution of the agents in the network assuming a 5-bit uniform quantizer is used; the desired average is denoted by the dashed grey line. We use in: (a) Naive Quantized Gossip (NQG); (b) Kashyap's Quantized Gossip (KQG); and (c) Symmetric Quantized Gossip (SQG).

4.6. Towards a distributed source recovery: a physics-driven consensus scheme

As highlighted in Section 4.4, we assume that the sensors know the topology of the network and that each sensor performs the Delaunay triangulation as shown in Figure 4.4.

4.6.1. Consensus-based estimation of the Generalized Measurements over Sensors Networks

With the aforementioned assumptions, we can now derive the consensus-based diffusion source estimation scheme and we focus here on the case of instantaneous and localized sources.

Firstly consider the surface integral contribution in (3.17). We have seen that if we partition the monitored domain Ω into non-overlapping triangular subdivisions $\{\Delta_j\}_{j=1}^{J}$ such that $\bigcup_{i=1}^{I} \Delta_i = \Omega$ and $\Delta_i \bigcap \Delta_j = \varnothing$ for $i \neq j$, as shown in Figure 4.4 for instance, then the integral can be approximated by the sum (4.7), restated below

$$\int_{\Omega} (\Psi_k \dot{U})(\mathbf{x}, T)\, \mathrm{d}V \approx \frac{1}{3} \sum_{j=1}^{J} \sum_{j'=1}^{3} \Psi_k(\mathbf{v}_{j,j',r}) \dot{\Phi}_{j,j'}(t_L) |\Delta_j|,$$

where $\mathbf{v}_{j,j'}$ is the j'-th vertex of triangle j-th and $\dot{\Phi}_{j,j',r}(t_L) = \dot{U}_r(\mathbf{v}_{j,j'}, t_L)$ is the measurement of the sensor situated at this vertex at time $t = t_L = T$. Moreover the double sum, in the above approximation (4.7), can be re-written in the following form:

$$\int_{\Omega} (\Psi_k \dot{U})(\mathbf{x}, T)\, \mathrm{d}V \approx \frac{1}{3} \sum_{j=1}^{J} \sum_{j'=1}^{3} \Psi_k(\mathbf{v}_{j,j',r}) \dot{\Phi}_{j,j'}(t_L) |\Delta_j|$$

$$
\begin{aligned}
&= \frac{1}{3}\sum_{j=1}^{J} |\Delta_j| \Big(\Psi_k(\mathbf{v}_{j,1})\dot{\Phi}_{j,1,r}(t_L) + \Psi_k(\mathbf{v}_{j,2})\dot{\Phi}_{j,2,r}(t_L) + \Psi_k(\mathbf{v}_{j,3})\dot{\Phi}_{j,3,r}(t_L) \Big) \\
&= \sum_{n=1}^{N} \underbrace{\left(\frac{1}{3}\Psi_k(\mathbf{x}_n) \sum_{j=1}^{J_n} |\Delta_{n,j}| \right)}_{:=A_n(k)} \dot{\Phi}_{n,r}(t_L) \qquad (4.24)
\end{aligned}
$$

where $\Delta_{n,j}$ and $|\Delta_{n,j}|$ are used to denote the j-th triangle and its area respectively, for which some node n is a vertex. The last equality follows by noticing that $\dot{\Phi}_{j,j',r}(t_L)$ is always weighted by the product of the area of triangle j and the term $1/3\Psi_k(\mathbf{v}_{j,j'})$. Hence letting the set of all triangles that share a common vertex n located at $\mathbf{x}_n$ be $\mathcal{T}_n = \{\Delta_{n,1}, \Delta_{n,2}, \ldots, \Delta_{n,J_n}\}$, the measurement $\dot{\Phi}_{n,r}(t_L) \stackrel{\text{def}}{=} \dot{U}_r(\mathbf{x}_n, t_L)$ is always weighted by the sum of the areas of its corresponding triangles and $1/3\Psi_k(\mathbf{x}_n)$. We denote this weight that directly depends on the sensing function and Delaunay triangulation (equivalently, the topology of the network) by $A_n(k)$.

For the boundary integral, the time-integrated field $U_r(\mathbf{x}_n, T)$ and its spatial derivative $\nabla U_r(\mathbf{x}_n, T) = \left[\frac{\partial U_r}{\partial x_1}, \frac{\partial U_r}{\partial x_2}\right]^{\mathsf{T}}$ are required. Recall that $\mathcal{S}_{\partial\Omega} = \{\mathbf{x}_1, \ldots, \mathbf{x}_{n-1}, \mathbf{x}_i, \ldots, \mathbf{x}_I\}$ denote the cyclically ordered set of the boundary sensors, that coincide with the vertices of the convex hull $\mathcal{CH}(\mathcal{S})$. Furthermore, assume the elements of $\mathcal{S}_{\partial\Omega}$ are in anticlockwise order, then: $\hat{\mathbf{n}}_{\partial\Omega}\mathrm{d}S \approx [x_{2,n} - x_{2,n-1}, x_{1,n-1} - x_{1,n}]^{\mathsf{T}}$. Under this construction, it has been shown that the boundary integral can be approximated by (4.12):

$$
\begin{aligned}
\oint_{\partial\Omega} (\Psi_k \nabla U_r - U_r \nabla \Psi_k) \cdot \hat{\mathbf{n}}_{\partial\Omega}\, \mathrm{d}S \approx \sum_{\mathbf{x}_n \in \mathcal{S}_{\partial\Omega}} \Psi_k(\mathbf{x}_n) \Big[&\Big((\Phi_{n,r}(t_L))_{x_1} + k\Phi_{n,r}(t_L) \Big) (x_{2,n} - x_{2,n-1}) \\
&+ \Big((\Phi_{n,r}(t_L))_{x_2} + \mathrm{j}k\Phi_{n,r}(t_L) \Big) (x_{1,n-1} - x_{1,n}) \Big].
\end{aligned}
$$

The first term in (4.12) depends on $\Phi_{n,r}(t_L)$, as well as, its spatial derivative which must be approximated from spatiotemporal samples of $u(\mathbf{x}, t)$. Recall that $\Phi_{n,r}(t_L)$, can be obtained by sensor n independently of the others using trapezium rule (see (4.9) on 78). However, $(\Phi_{n,r}(t_L))_{x_1}$ and $(\Phi_{n,r}(t_L))_{x_2}$ can only be estimated reliably using neighbouring sensor measurements; herein, we will use a polynomial fitting approach due to its simplicity and accuracy. Using this approach, the task at hand is to find the regression function $\Phi_{n,r}(t_L) = \alpha_n x_{1,n} + \beta_n x_{2,n} + \gamma_n$ by estimating $(\alpha_n, \beta_n, \gamma_n)$ for each boundary sensor $\mathbf{x}_n \in \mathcal{S}_{\partial\Omega}$, (where $n = 1, \ldots, I$), using measurements from its nearest neighbours (sensors closest to the location $\mathbf{x}_n$). We denote the set of all such sensors by $\mathcal{S}'_{\partial\Omega}$.

Let the n-th sensor located at $\mathbf{x}_n$ with measurement $\Phi_{n,r}(t_L)$ have as its two closest sensors $\mathbf{x}'_n$ and $\mathbf{x}''_n$ with their corresponding measurements $\Phi_{n',r}(t_L)$ and $\Phi_{n'',r}(t_L)$. Given these three pairs $\{\mathbf{x}_n, \Phi_{n,r}(t_L)\}$, $\{\mathbf{x}'_n, \Phi_{n',r}(t_L)\}$ and $\{\mathbf{x}''_n, \Phi_{n'',r}(t_L)\}$ we can estimate the

parameters $(\alpha_n, \beta_n, \gamma_n)$ by solving the linear system:

$$\begin{bmatrix} \Phi_{n'',r}(t_L) \\ \Phi_{n,r}(t_L) \\ \Phi_{n',r}(t_L) \end{bmatrix} = \begin{bmatrix} x''_{1,n} & x''_{2,n} & 1 \\ x_{1,n} & x_{2,n} & 1 \\ x'_{1,n} & x'_{2,n} & 1 \end{bmatrix} \begin{bmatrix} \alpha_n \\ \beta_n \\ \gamma_n \end{bmatrix}, \tag{4.25}$$

$$\mathbf{u}_n = \mathbf{X}_n \mathbf{d}_n. \tag{4.26}$$

The system admits a unique solution, if $\mathbf{x}''_n, \mathbf{x}_n$ and $\mathbf{x}'_n$ are not collinear. Therefore, the approximation $\left[(\Phi_{n,r}(t_L))_{x_1}, (\Phi_{n,r}(t_L))_{x_2}\right]^\mathsf{T}$ of the local spatial derivative $\nabla U_r(\mathbf{x}, T)|_{\mathbf{x}=\mathbf{x}_n}$ can be retrieved immediately from the solution to this system by noticing that the polynomial $p(\mathbf{x}) = \alpha x_1 + \beta x_2 + \gamma$ has the derivative $\nabla p(\mathbf{x}) = \langle \alpha, \beta \rangle$. Hence $\nabla U_r(\mathbf{x}_n, T) = \left[\frac{\partial}{\partial x_1} U_r(\mathbf{x}_n, T), \frac{\partial}{\partial x_2} U_r(\mathbf{x}_n, T)\right]^\mathsf{T}\Big|_{\mathbf{x}=\mathbf{x}_n} \approx \left[(\Phi_{n,r}(t_L))_{x_1}, (\Phi_{n,r}(t_L))_{x_2}\right]^\mathsf{T} = [\alpha_n, \beta_n]^\mathsf{T}$, where

$$\alpha_n = \frac{(x_{2,n} - x'_{2,n})\Phi_{n'',r}(t_L) + (x'_{2,n} - x''_{2,n})\Phi_{n,r}(t_L) + (x''_{2,n} - x_{2,n})\Phi_{n',r}(t_L)}{\det(\mathbf{X}_n)},$$
$$\beta_n = \frac{(x'_{1,n} - x_{1,n})\Phi_{n'',r}(t_L) + (x''_{1,n} - x'_{1,n})\Phi_{n,r}(t_L) + (x_{1,n} - x''_{1,n})\Phi_{n',r}(t_L)}{\det(\mathbf{X}_n)}.$$

Then substituting these back into (4.12) gives

$$\oint_{\partial\Omega} (\Psi_k \nabla U_r - U_r \nabla \Psi_k) \cdot \hat{\mathbf{n}}_{\partial\Omega} \, \mathrm{d}S \approx \sum_{\mathbf{x}_n \in \mathcal{S}_{\partial\Omega}}^{I} b''_n(k)\Phi_{n'',r}(t_L) + b_n(k)\Phi_{n,r}(t_L) + b'_n(k)\Phi_{n',r}(t_L), \tag{4.27}$$

where

$$b''_n(k) = \frac{\Psi_k(\mathbf{x}_n)}{\det(\mathbf{X}_n)} \begin{bmatrix} x_{1,n-1} - x_{1,n} \\ x_{2,n} - x_{2,n-1} \end{bmatrix}^\mathsf{T} \begin{bmatrix} x'_{1,n} - x_{1,n} \\ x_{2,n} - x'_{2,n} \end{bmatrix}, \tag{4.28a}$$

$$b_n(k) = \frac{\Psi_k(\mathbf{x}_n)}{\det(\mathbf{X}_n)} \begin{bmatrix} x_{1,n-1} - x_{1,n} \\ x_{2,n} - x_{2,n-1} \end{bmatrix}^\mathsf{T} \begin{bmatrix} x''_{1,n} - x'_{1,n} \\ x'_{2,n} - x''_{2,n} \end{bmatrix}$$
$$+ k\Psi_k(\mathbf{x}_n) \left[(x_{2,n} - x_{2,n-1}) + \mathrm{j}(x_{1,n-1} - x_{1,n})\right], \tag{4.28b}$$

$$b'_n(k) = \frac{\Psi_k(\mathbf{x}_n)}{\det(\mathbf{X}_n)} \begin{bmatrix} x_{1,n-1} - x_{1,n} \\ x_{2,n} - x_{2,n-1} \end{bmatrix}^\mathsf{T} \begin{bmatrix} x_{1,n} - x''_{1,n} \\ x''_{2,n} - x_{2,n} \end{bmatrix}. \tag{4.28c}$$

The terms $b''_n(k), b_n(k)$ and $b'_n(k)$ in (4.28) are dependent only on the topology of the network (specifically the locations of the sensors) and our choice of sensing function $\Psi_k(\mathbf{x})$. Indeed given the assumptions detailed in Section 4.4, these weights can be precomputed for every sensor in the network, such that

$$\oint_{\partial\Omega} (\Psi_k(\mathbf{x}) \nabla U_r(\mathbf{x}, T) - U_r(\mathbf{x}, T) \nabla \Psi_k(\mathbf{x})) \cdot \hat{\mathbf{n}}_{\partial\Omega} \, \mathrm{d}S \approx \sum_{\mathbf{x}_n \in \mathcal{S}} B_n(k)\Phi_{n,r}(t_L), \tag{4.29}$$

where $B_n(k)$ is non-zero only if the n-th sensor is a boundary sensor or if it is one of the two nearest sensors to a boundary sensor, i.e. $\mathbf{x}_n \in \mathcal{S}_{\partial\Omega} \cup \mathcal{S}'_{\partial\Omega}$. Otherwise $B_n(k)$ is zero. Finally, we can combine (4.24) and (4.29), to obtain the estimates for $\mathcal{Q}(k,r)$:

$$\mathcal{Q}(k,r) \approx \sum_{n\in\mathcal{N}} A_n(k)\dot{\Phi}_{n,r}(t_L) - \mu B_n(k)\Phi_{n,r}(t_L) \tag{4.30}$$

$$= \frac{1}{N}\sum_{n\in\mathcal{N}} y_n(k,r). \tag{4.31}$$

where

$$y_n(k,r) = N\Big(A_n(k)\dot{\Phi}_{n,r}(t_L) - \mu B_n(k)\Phi_{n,r}(t_L)\Big), \tag{4.32}$$

whilst the terms $\Phi_{n,r}(t_L)$ and $\dot{\Phi}_{n,r}(t_L)$ are approximations of the time-integrals $U_r(\mathbf{x}_n, t_L)$ and $\dot{U}_r(\mathbf{x}_n, t_L)$, in (4.4), obtained using (4.9) and (4.10) respectively.

Upon deployment of the sensors, each sensor can precompute its unique weights $\{A_n(k)\}_k$ and $\{B_n(k)\}_k$ for $k = 0, \ldots, K$. This is possible as a result of the assumptions stated in Section 4.4. After which they can start to monitor the region of interest Ω, by sensing the field locally. To initiate the estimation process, the sensor n exchanges its modified measurements $\{y_n(k,r)\}_k$ with a randomly chosen neighbor. This begins the gossip round, as detailed in Section 4.5, it continues until convergence to $\{\mathcal{Q}(k,r)\}_k$. All sensors in the network can now independently apply Prony's method to its current estimate of $\{\mathcal{Q}(k,r)\}_k$ in order to recover all unknown source parameters as described in Section 3.4. Consequently, we can now state the following proposition for solving the IDSP concerning instantaneous sources of diffusion fields:

Proposition 4.1. *By exchanging the measure* $y_n(k,r)$=$N\left[A_n(k)\dot{\Phi}_{n,r}(t_L) - \mu B_n(k)\Phi_{n,r}(t_L)\right]$ *which is a properly weighted sum of the sensor measurements* $\{\varphi_n(t_l)\}_l$ *of the* n*-th sensor, it is possible for each sensor in the network to converge to the same generalized measurements* $\{\mathcal{Q}(k,r)\}_k$ *of the centralized algorithm, and hence recover the unknown diffusion source parameters with the same estimation performance. Here* $A_n(k)$ *and* $B_n(k)$ *are dependent on* $\Psi_k(\mathbf{x})$ *and the topology of the network, whilst* $\Phi_{n,r}(t_L)$ *and* $\dot{\Phi}_{n,r}(t_L)$ *are approximations of the time integrals* (4.4a) *and* (4.4b) *respectively obtained by sensor* n *at location* $\mathbf{x} = \mathbf{x}_n$*, with* $T = t_L$*.*

The non-instantaneous source models require a more careful treatment, however. As usual for this class of models, we consider $\mathcal{R}(k) = \mathcal{Q}(k,0)$ and recover the unknown parameters from its field samples in a distributed manner as follows.

Recovery of non-instantaneous sources

It is clear from (3.20) that applying Prony's method directly on $\{\mathcal{R}(k) = \mathcal{Q}(k,0)\}_{k=0}^{K}$ where $K \geq 2M-1$, reveals the unknowns $\{(c'_m, e^{-(\xi_{1,m}+\mathrm{j}\xi_{2,m})})\}_m$, thus the locations of the M sources can be immediately recovered.

However to recover the remaining source parameter, we notice that c'_m and thus $\{\mathcal{R}(k)\}$ depend on T, i.e. the window over which the time-integrals are computed, hence we will write $c'_m(T)$ and $\{\mathcal{R}(k,T)\}$ respectively to emphasize this dependence (this is not true for instantaneous sources). By considering three sequences $\{\mathcal{R}(k,T_1)\}$, $\{\mathcal{R}(k,T_2)\}$ and $\{\mathcal{R}(k,T_3)\}$ obtained when we take the time integrals in (3.17) over three intervals $[0,T_1]$, $[0,T_2]$ and $[0,T_3]$, where $T_2 = T_1 + \Delta T$ and $T_3 = T_1 + 2\Delta T$, and applying Prony's method to the three sequence $\{\mathcal{R}(k,T_1)\}$, $\{\mathcal{R}(k,T_2)\}$ and $\{\mathcal{R}(k,T_3)\}$ leads to $\{c'_m(T_1)\}$, $\{c'_m(T_2)\}_m$ and $\{c'_m(T_3)\}_m$. From these, it has been shown that α_m, τ_m and c_m are given by (3.29), (3.31) and (3.32), respectively.

Given only spatiotemporal samples of the field the sequences $\{\mathcal{R}(k,T_1)\}$, $\{\mathcal{R}(k,T_2)\}$ and $\{\mathcal{R}(k,T_3)\}$ will need to be approximated directly from the field samples. In this situation, sensor n can compute locally $y_n^{(i)}(k,0)=N\left(A_n(k)\dot{\Phi}_{n,0}(T_i)-\mu B_n(k)\Phi_{n,0}(T_i)\right)$ for $i=1,2,3$ and $k=0,1,\ldots,K$. Gossiping can then begin: they all exchange locally and update their sequence $\{y_{n,0}^{(i)}(k)\}_k$ $i=1,2,3$, until convergence, to obtain the desired approximation for $\{\mathcal{R}(k,T_i)\}_k$ $i=1,2,3$. Where upon convergence each node will have the desired information to recover all unknown source parameters as described in Section 3.4.2, providing $K \geq 2M-1$.

This results in the following proposition.

Proposition 4.2. *By exchanging* $y_n^{(i)}(k,0)=N\left(A_n(k)\dot{\Phi}_{n,0}(T_i)-\mu B_n(k)\Phi_{n,0}(T_i)\right)$ *which is a properly weighted sum of the sensor measurements* $\{\varphi_n(t_l)\}_l$ *of the* n*-th sensor, for* $i = 1,2,3$ *and* $T_i = T-(3-i)\Delta T$*, it is possible for each sensor in the network to converge to the same generalized measurements* $\{\mathcal{Q}(k,0)\}_k$ *in the centralized algorithm, and hence recover the unknown diffusion source parameters with the same estimation performance. Here* $A_n(k)$ *and* $B_n(k)$ *are dependent on* $\Psi_k(\mathbf{x})$ *and the topology of the network, whilst* $\Phi_{n,0}(t_L)$ *and* $\dot{\Phi}_{n,0}(t_L)$ *are approximations of the time integrals* (4.4a) *and* (4.4b) *respectively obtained by sensor* n *at location* $\mathbf{x}=\mathbf{x}_n$*, with* $T=t_L$*.*

4.6.2. Noise robust consensus-based estimation

Proposition 4.1 presents a diffusion equation-driven gossiping strategy for recovering the unknown source model parameters from noiseless field samples. In a noisy scenario however, whilst the underlying gossiping strategy stays the same, some further processing may be required to enhance noise resilience. Specifically we will demonstrate how to design an in-network noise prewhitening filter, $\mathbf{F}_{\text{pw}}$, so that when applied to the noisy/perturbed generalized measurement sequence, $\{\mathcal{Q}_\epsilon(k,r)\}_k$, results in an improvement of the estimation performance of Prony's method and its variations.

We begin our exposition by recalling that: in Proposition 4.1 the sensors exchange their local measures $\{y_n(k,r)\}_k$ with neighbouring nodes in order to reach a consensus on the desired generalized measurements $\{\mathcal{Q}(k,r)\}_k$ (for a fixed r). In addition, for each non-negative integer k, $y_n(k,r) = N\left(A_n(k)\dot{\Phi}_{n,r}(t_L) - \mu B_n(k)\Phi_{n,r}(t_L)\right)$, this is simply a weighted sum of the temporal field measurements taken by sensor n, thus we can write

$y_n(k,r) = \sum_{l=0}^{L} w_{n,l}(k,r)\varphi_n(t_l)$ with $w_{n,l}(k,r) \in \mathbb{C}$, since $A_n(k), B_n(k) \in \mathbb{C}$. Therefore, in the noisy setting the sensors will instead exchange:

$$\begin{aligned} y_n^{\epsilon}(k,r) &= \sum_{l=0}^{L} w_{n,l}(k,r)\varphi_n^{\epsilon}(t_l) \\ &= \sum_{l=0}^{L} w_{n,l}(k,r)\left(\varphi_n(t_l) + \epsilon_{n,l}\right) \\ &= y_n(k,r) + \sum_{l=0}^{L} w_{n,l}(k,r)\epsilon_{n,l}. \end{aligned} \tag{4.33}$$

Consider now the consensus value. Upon convergence, each sensor in the network has the perturbed generalized measurement: $\{\mathcal{Q}_\epsilon(k,r) = \frac{1}{N}\sum_{n=1}^{N} y_n^{\epsilon}(k,r)\}_{k=0}^{K}$ from which they construct the Toeplitz matrix, i.e. $\text{Toep}\left(\{\mathcal{Q}_\epsilon(k,r)\}_{k=0}^{K}\right)$, in order to apply Prony's method (see Prony's and the matrix pencil methods in Section 2.2.2). In the noisy scenario they will instead construct,

$$\begin{aligned} \mathbf{T}_\epsilon &\stackrel{\text{def}}{=} \text{Toep}\left(\{\mathcal{Q}_\epsilon(k,r)\}_k\right) = \text{Toep}\left(\left\{\frac{1}{N}\sum_{n=1}^{N} y_n^{\epsilon}(k,r)\right\}_k\right) \\ &= \text{Toep}(\{\mathcal{Q}_\epsilon(k,r)\}_k) + \text{Toep}\left(\left\{\frac{1}{N}\sum_{n=1}^{N}\sum_{l=0}^{L} w_{n,l}(k,r)\epsilon_{n,l}\right\}_k\right) \\ &= \text{Toep}(\{\mathcal{Q}_\epsilon(k,r)\}_k) + \text{Toep}(\{\epsilon(k,r)\}_k) \\ &= \mathbf{T} + \mathbf{E}. \end{aligned} \tag{4.34}$$

It is clear that the noise term in $\text{Toep}\left(\{\mathcal{Q}^{\epsilon}(k,r)\}_{k=0}^{K}\right)$ is now coloured due to the weighted-sum. However Prony's method and its variations implicitly assume that the noise in the terms $\{\mathcal{Q}_\epsilon(k,r)\}_k$ are i.i.d, hence it must be prewhitened before estimating the unknown source parameters from it. To this end, before applying Prony's method, each sensor can design and apply (by post-multiplying the Toeplitz matrix $\mathbf{T}_\epsilon \stackrel{\text{def}}{=} \text{Toep}\left(\{\mathcal{Q}_\epsilon(k,r)\}_k\right)$) the pre-whitening filter $\mathbf{F}_{\text{pw}}$ given by:

$$\mathbf{F}_{\text{pw}} = \mathbf{C}_\epsilon^{\frac{\dagger}{2}} \tag{4.35}$$

where $(\cdot)^{\frac{\dagger}{2}}$, is the square root of the Moore-Penrose pseudoinverse, and

$$\mathbf{C}_\epsilon = \text{Cov}\{\mathbf{E}\} = \mathbb{E}\left\{\mathbf{E}\mathbf{E}^{\mathsf{H}}\right\}. \tag{4.36}$$

Notice that the covariance matrix, $\mathbf{C}_\epsilon$ depends directly on the variance of the sensor noise and on the weights $\{w_{n,l}(k,r)\}_{n,l}$ (hence the network topology), for a fixed r and any $k \in \{0,1,\ldots,K\}$. We shall now explicitly derive an expression for $\mathbf{C}_\epsilon$.

For ease of exposition, we define the matrix of coefficients $\mathbf{W}(k,r) \in \mathbb{C}^{N\times(L+1)}$, such that its entries are given by $[\mathbf{W}(k,r)]_{n,l+1} = w_{n,l}(k,r)$. Hence we can define the vector

of coefficients $\mathbf{w}(k,r) = \text{vec}(\mathbf{W}(k,r))$ to be the vectorization of $\mathbf{W}(k,r)$ – i.e. $\mathbf{w}(k,r) \in \mathbb{C}^{N(L+1)}$ is formed by stacking the columns of the matrix $\mathbf{W}(k,r)$ into a single column vector. Similarly, let $\boldsymbol{\epsilon} \in \mathbb{R}^{N(L+1)}$ be the vector formed from the i.i.d noise sequence $\{\epsilon_{n,l}\}_{n,l}$ in the same way. Then $\epsilon(k,r) = \frac{1}{N}\sum_{n=1}^{N}\sum_{l=0}^{L} w_{n,l}(k,r)\epsilon_{n,l} = \frac{1}{N}\mathbf{w}^{\mathsf{T}}(k,r)\boldsymbol{\epsilon}$. From this, if for instance $\mathbf{E} \in \mathbb{R}^{N_1 \times N_2}$, it follows that

$$\mathbf{E} = \begin{bmatrix} \epsilon(N_2{-}1,r) & \epsilon(N_1{-}2,r) & \cdots & \epsilon(0,r) \\ \epsilon(N_2,r) & \epsilon(N_2{-}1,r) & \cdots & \epsilon(1,r) \\ \vdots & \vdots & \ddots & \vdots \\ \epsilon(N_1{+}N_2{-}2,r) & \epsilon(N_1{+}N_2{-}3,r) & \cdots & \epsilon(N_1{-}1,r) \end{bmatrix}.$$

To reduce notational burden we will drop, without loss of generality, the dependence on r in the remainder of our derivation. Thus we may write:

$$\begin{aligned}
\mathbf{C}_\epsilon &= \mathbb{E}\left\{\mathbf{E}\mathbf{E}^{\mathsf{H}}\right\} \\
&= \mathbb{E}\left\{\begin{bmatrix} \epsilon(N_2{-}1) & \epsilon(N_1{-}2) & \cdots & \epsilon(0) \\ \epsilon(N_2) & \epsilon(N_2{-}1) & \cdots & \epsilon(1) \\ \vdots & \vdots & \ddots & \vdots \\ \epsilon(N_1{+}N_2{-}2) & \epsilon(N_1{+}N_2{-}3) & \cdots & \epsilon(N_1{-}1) \end{bmatrix} \begin{bmatrix} \epsilon(N_2{-}1) & \epsilon(N_2) & \cdots & \epsilon(N_1{+}N_2{-}2) \\ \epsilon(N_1{-}2) & \epsilon(N_2{-}1) & \cdots & \epsilon(N_1{+}N_2{-}3) \\ \vdots & \vdots & \ddots & \vdots \\ \epsilon(0) & \epsilon(1) & \cdots & \epsilon(N_1{-}1) \end{bmatrix}\right\} \\
&= \mathbb{E}\left\{\begin{bmatrix} \sum_{i=0}^{N_2-1}\epsilon(i)\epsilon^{\mathsf{H}}(i) & \sum_{i=0}^{N_2-1}\epsilon(i)\epsilon^{\mathsf{H}}(i{+}1) & \cdots & \sum_{i=0}^{N_2-1}\epsilon(i)\epsilon^{\mathsf{H}}(i{+}N_1{-}1) \\ \sum_{i=0}^{N_2-1}\epsilon(i{+}1)\epsilon^{\mathsf{H}}(i) & \sum_{i=0}^{N_2-1}\epsilon(i{+}1)\epsilon^{\mathsf{H}}(i{+}1) & \cdots & \sum_{i=0}^{N_2-1}\epsilon(i{+}1)\epsilon^{\mathsf{H}}(i{+}N_1{-}1) \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{i=0}^{N_2-1}\epsilon(i{+}N_1{-}1)\epsilon^{\mathsf{H}}(i) & \sum_{i=0}^{N_2-1}\epsilon(i{+}N_1{-}1)\epsilon^{\mathsf{H}}(i{+}1) & \cdots & \sum_{i=0}^{N_2-1}\epsilon(i{+}N_1{-}1)\epsilon^{\mathsf{H}}(i{+}N_1{-}1) \end{bmatrix}\right\}.
\end{aligned}$$

Taking the expectation operator inside the matrix and noting that each element,

$$\begin{aligned}
\mathbb{E}\left\{\sum_{i=0}^{N_2-1}\epsilon(i{+}j)\epsilon^{\mathsf{H}}(i)\right\} &= \mathbb{E}\left\{\sum_{i=0}^{N_2-1}\epsilon(i)\epsilon^{\mathsf{H}}(i{+}j)\right\} \\
&= \frac{1}{N^2}\sum_{i=0}^{N_2-1}\mathbb{E}\left\{\mathbf{w}^{\mathsf{T}}(i)\boldsymbol{\epsilon}\boldsymbol{\epsilon}^{\mathsf{H}}\mathbf{w}^{*}(i{+}j)\right\} \\
&= \frac{1}{N^2}\sum_{i=0}^{N_2-1}\mathbf{w}^{\mathsf{T}}(i)\left(\sigma^2\mathbf{I}\right)\mathbf{w}^{*}(i{+}j) \\
&= \frac{\sigma^2}{N^2}\sum_{i=0}^{N_2-1}\mathbf{w}^{\mathsf{T}}(i)\mathbf{w}^{*}(i{+}j),
\end{aligned}$$

allows us to write the covariance matrix (4.36) as,

$$\mathbf{C}_\epsilon = \frac{\sigma^2}{N^2}\begin{bmatrix} \sum_{i=0}^{N_2-1}\|\mathbf{w}(i)\|^2 & \sum_{i=0}^{N_2-1}\mathbf{w}^\mathsf{T}(i)\mathbf{w}^*(i+1) & \cdots & \sum_{i=0}^{N_2-1}\mathbf{w}^\mathsf{T}(i)\mathbf{w}^*(i+N_1-1) \\ \sum_{i=0}^{N_2-1}\mathbf{w}^\mathsf{T}(i+1)\mathbf{w}^*(i) & \sum_{i=0}^{N_2-1}\|\mathbf{w}(i+1)\|^2 & \cdots & \sum_{i=0}^{N_2-1}\mathbf{w}^\mathsf{T}(i+1)\mathbf{w}^*(i+N_1-1) \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{i=0}^{N_2-1}\mathbf{w}^\mathsf{T}(i+N_1-1)\mathbf{w}^*(i) & \sum_{i=0}^{N_2-1}\mathbf{w}^\mathsf{T}(i+N_1-1)\mathbf{w}^*(i+1) & \cdots & \sum_{i=0}^{N_2-1}\|\mathbf{w}(i+N_1-1)\|^2 \end{bmatrix}. \tag{4.37}$$

We can restore the dependence of the covariance matrix on r by replacing $\mathbf{w}(\cdot)$, in (4.37) above with, $\mathbf{w}(\cdot, r)$.

Note that since all sensors know the network topology, they are actually able to compute $\{w_{n,l}(k,r)\}_{n,l}$ for all $k = 0, 1, \ldots, K$ and a fixed r. From this sequence they can construct the column vector $\mathbf{w}(k,r)$ and compute the desired covariance matrix $\mathbf{C}_\epsilon$ and so obtain the desired prewhitening filter $\mathbf{F}_{\text{pw}} = \mathbf{C}_\epsilon^{\frac{\dagger}{2}}$ as required.

To further improve the noise robustness of our source estimation approach, we may also adapt the sequential local search technique introduced in 4.3.3. This sequential local search technique for robustifying the estimation of τ_m (and α_m for our non-instantaneous source models), is performed independently and does not require the FC or data with other nodes, hence can be implemented directly in a distributed network.

This modification is based on searching for a time interval over where only a single diffusion source is active. When such an interval is found, the active source parameters are estimated and its contribution to the spatiotemporal field measurements is removed before the next source can be estimated. In summary, this is achieved by considering measurements only from an initial time interval $[0, T]$ and initiating the following steps:

1. Assume there are $M' \geq 2$ unknown sources. Approximate $\{\mathcal{R}_\epsilon(k)\}_{k=0}^{2M'-1}$ using the gossiping approach as described in Section 5.1.

2. Estimate the M' source intensities $\{\sigma'_{m'}\}_{m'=1}^{M'}$ and locations $\{\boldsymbol{\xi}'_{m'}\}_{m'=1}^{M'}$ using Prony's method.

3. Each sensor can then check if their estimated sources are valid, i.e. the pair $(\sigma'_{m'}, \boldsymbol{\xi}'_{m'})$ is valid if both conditions:

   a) c_m greater than some predefined threshold, and

   b) $\boldsymbol{\xi}'_{m'}$ is inside the monitored region Ω,

   are simultaneously satisfied. Let M_{vs} be the number of valid sources.

4. There are now three cases:

   a) $M_{\text{vs}} > 1$: the time window $[0, T]$ is reduced and steps (i)-(iv) are repeated.

b) $M_{\text{vs}} < 1$: the time window $[0, T]$ is increased and steps (i)-(iv) are repeated.

c) $M_{\text{vs}} = 1$: sensors estimate all unknowns, say $(c_1, \boldsymbol{\xi}_1, \tau_1)$, for this active source as described in Section 4.6.1, and then perform a local line search around their estimates of τ_1 to refine this estimate.

5. Output the estimates $(c_1, \boldsymbol{\xi}_1, \tau_1)$.

The resulting algorithm that combines prewhitening and the sequential local search with the gossiping strategy of Proposition 4.1 is described in the inset Algorithm 4.5. Under the assumptions outlined in Section 4.4, Algorithm 4.5 guarantees that the unknown source parameters can be recovered in a distributed way (by performing local exchanges of the measurements $\{y_n(k, r)\}_n$ using gossiping and then applying Prony's method on convergence) with the same estimation performance as the centralized estimation algorithm.

In this section, we have seen how to implement a distributed estimation of the unknown source models f using gossiping in a sensor network. Implicitly however, we have assumed that the communication links between sensors are ideal. In the following section, we develop techniques for adapting our approach to operate properly even when this assumption does not hold.

Algorithm 4.5 Robust Distributed Estimation of M sources

Require: $\{\varphi_n(t_l)\}_l$, sampling interval Δt, μ, $\{A_n(k)\}$ and $\{B_n(k)\}$.

1: The sensor n starts the process at some time T.
2: Sensors compute their $\{y_n(k)\}$ over window $t \in [0, T]$.
3: Then the gossip round begins, to compute $\{\mathcal{Q}(k, 1)\}$ with $K \geq 2M' - 1$ and $M' \geq 2$ using consensus.
4: Compute and apply the prewhitening filter $\mathbf{F}_{\text{pw}}$.
5: Each sensor applies Prony's to its estimate of $\{\mathcal{Q}(k, 1)\}$ to recover M' pairs of $\{\sigma'_{m'}, \boldsymbol{\xi}'_{m'}\}_{m'=1}^{M'}$.
6: Each sensor sets $M_{vs}(n)$ as the number of pairs $\{\sigma'_{m'}, \boldsymbol{\xi}'_{m'}\}$ having both $\sigma'_{m'}$ greater than some threshold and $\boldsymbol{\xi}'_{m'} \in \Omega$.
7: They can now gossip $M_{vs}(n)$ to get $\hat{M}_{vs}$, or a sensor is elected at random to decide $\hat{M}_{vs}$. There now arises three cases:
8: **if** $\hat{M}_{vs} > 1$ **then**
9: Decrease window size $[0, T]$ and Go to 2.
10: **end if**
11: **if** $\hat{M}_{vs} < 1$ **then**
12: Increase window size $[0, T]$ and Go to 2.
13: **end if**
14: **if** $\hat{M}_{vs} == 1$ **then**
15: Sensors apply Prony's to its estimate of $\{\mathcal{R}(k)\}$ to recover $\{c_m, \boldsymbol{\xi}_m, \tau_m\}$.
16: Sensors perform a Local Line Search around τ_m, to refine this estimate.
17: Sensors reconstruct local field and adjusts $\{\varphi_n(t_l)\}_l$.
18: $m \leftarrow m + 1$.
19: Increase size of window $[0, T]$ and Go to 2, if $m < M$, else go to 18.
20: **end if**
21: **return** $\{c_m\}_{m=1}^M$, $\{\alpha_m\}_{m=1}^M$, $\{\boldsymbol{\xi}_m\}_{m=1}^M$ and $\{\tau_m\}_{m=1}^M$.

4.7. Source estimation over communication constrained sensor networks: physics-driven quantized consensus

A limited communication bandwidth is a common constraint on many sensor networks built for estimation/inference tasks. This limitation necessitates the quantization of data transmitted over such networks. In this section we consider the distributed estimation of the sources of a diffusion field using a bandwidth constrained network of sensors. In particular, we propose an adaptation of the distributed diffusion source estimation algorithm outlined in Section 4.6.1, which implicitly assumed that the measurement sequence $\{y_n(k)\}_k$ of the n-th sensor can be transmitted with infinite precision to its neighbours. Our adaptation is based on the use of quantized gossip strategies. Although there exist a variety of quantized gossip strategies some are much more suited to our estimation task. To this end, through simulations we use a distortion measure which we introduce in the next section, to guide our choice as to which quantized gossip strategy is the most suitable for our estimation task when the sequence $\{y_n(k,r)\}_k$ is quantized using a uniform q-bit quantizer.

4.7.1. Quantization Noise

Communicating over a quantized channel leads, in general, to a distortion in the transmitted signals. Therefore, we will study through simulations the extent (and severity) of the distortion introduced in the estimates of $\{\mathcal{R}(k) = \mathcal{Q}(k,0)\}_k$ obtained at each node, upon convergence to a quantized consensus resulting from localized exchanges of quantized versions of $\{y_n(k,0)\}_k$, $\forall n \in \{1,\ldots,N\}$. The metric used here is called the signal-to-discretization and-quantization noise ratio (SDQNR), and it measures the level of distortion (compared to the ground truth values) present in the consensus values. The SDQNR at the n-th sensor is defined as

$$\mathrm{SDQNR}(n) = 10\log_{10}\left(\frac{\sum_{k=0}^{K}\left|\sum_{m=0}^{M} c_m e^{-k(\xi_{1,m}+j\xi_{2,m})}\right|^2}{\sum_{k=0}^{K}\left|\lim_{i\to\infty}\tilde{y}_{n,i}(k) - \sum_{m=0}^{M} c_m e^{-k(\xi_{1,m}+j\xi_{2,m})}\right|^2}\right), \tag{4.38}$$

where $\tilde{y}_{n,i}(k)$ is the quantized. Both quantized gossip schemes [82, 67] considered in Sections 4.5.1 and 4.5.1 respectively are known to achieve quantized consensus, such that all sensors in the network converge to values which are within one bin of the true average. Both schemes lead to a distribution of values upon convergence, these values are such that the initial network average is preserved. Figure 4.6 shows the SDQNR for both schemes, this simulation is for two sources with intensities ($c_1 = c_2 = 1$), locations $\boldsymbol{\xi}_1 = (0.113, 0.221), \boldsymbol{\xi}_2 = (0.234, 0.085)$ and activation times $\tau_1 = 1.213, \tau_2 = 5.126$. For statistical significance, we perform 100 trials, where each new trial uses a different realization of the random graph (i.e. a new RGG). In so doing, we obtain an SDQNR for each sensor and each independent trial and so Figure 4.6 shows distribution of SDQNR of

the quantized consensus states for a range of q-bit uniform quantizers using a box plot. It can be seen that the scheme proposed by Frasca *et al* results in consistently higher SNR compared to that of Kashyap *et al*. Consequently we propose the use of the quantized

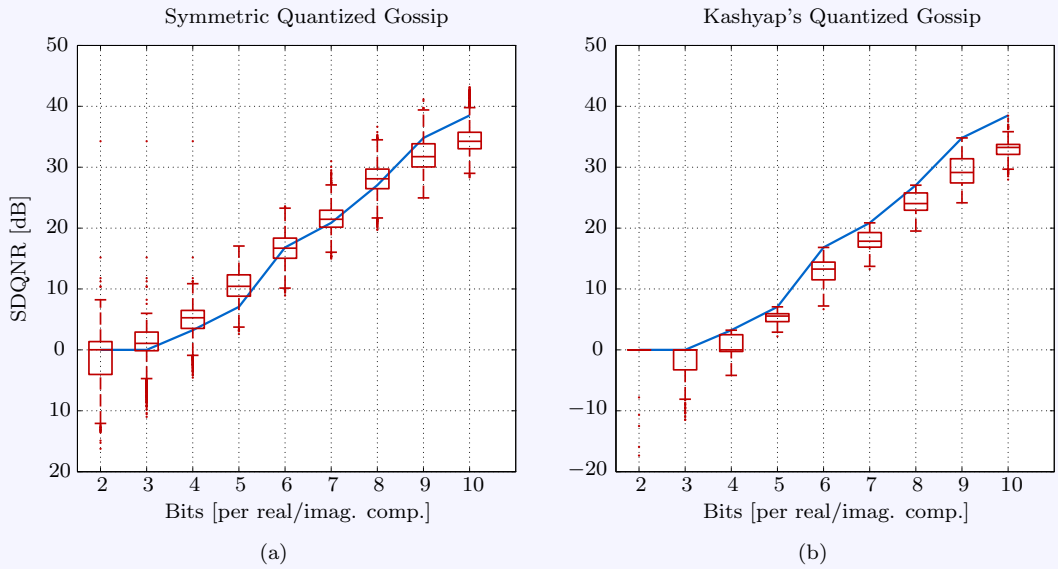


Figure 4.6.: **SDQNR of Quantized Gossip Schemes.** The realized SDQNR vs number of quantization bits (q-bits for $q = 2, \ldots, 10$) using (a) Frasca's Symmetric Quantized Gossip Scheme and (b) Kashyap's Quantized Gossip Scheme. This simulation uses $M = 2$, $K = 5$, $N = 90$, the network is a RGG, $\mathcal{G}(N, 0.6)$. The solid line in both plots shows the signal-to-noise ratio of the quantized version of ground truth Prony's sequence, i.e. $\{Q_q(\mathcal{R}(k))\}_k$ for each q.

pairwise gossip scheme summarized by (4.23) during the gossip stage as this results in a sequence that is closer to the ground truth sequence of $\{\mathcal{R}(k)\}_k$.

4.8. Summary and conclusions

In this chapter we discuss both centralized and distributed schemes for solving a class of discrete inverse diffusion source problems. Specifically, with the assumption that the field of interest is monitored via a network of *dumb* and *smart* sensors, in the centralized and distributed settings respectively, we show how to properly adapt the inversion formulae obtained in Chapter 3 to solve the problem. In the centralized setting all measurements are forwarded to a FC for processing, whereas in the distributed setting, a diffusion-equation driven gossip algorithm is devised where we show how to properly modify the sensor measurements using the mathematical model of the field, as well as the network topology, such that each sensor node in the network is able to recover the unknown source parameters

through *localized* and *quantized* interactions with other nodes alone. This allows us to eliminate the need for a fusion centre whilst also managing the communications costs. In both of these centralized and distributed settings we also presented several iterative and non-iterative approaches for robustifying the proposed methods under measurement noise and other model mismatches.

Chapter 5.

Simulation results

5.1. Introduction

Chapter 3 presented a precise framework for solving a class of IDSPs given continuous measurements of the induced field. Specifically, using this framework we derived a set of closed form inverse formulae from which we can recover the unknown source parameters. Chapter 4 then demonstrated how to properly adapt these formulae to the case of discrete spatiotemporal field measurements obtained through the use of a sensor network and derived results for both the centralized and distributed cases. In the present chapter, our main aim is to evaluate and understand the validity, robustness and limitations of the proposed centralized and distributed algorithms for solving the class of discrete IDSPs, proposed in Chapter 4.

In order for us to carry out proper numerical validations of the algorithms, we must first obtain the field measurements to operate on. To this end we begin by considering numerical results when we have access firstly, to synthetic data in Section 5.2 before considering the case of real data experiments in Section 5.3. Particularly, we assess the performance of our centralized schemes by:

(i) presenting estimation results in Sections 5.2.1 and 5.2.2, given noisy field measurements, for fields induced by our class of source models; and

(ii) investigating which effect the error of approximating our family of integrals using quadrature techniques has on our algorithms, in Section 5.2.3.

On the other hand, Sections 5.2.4 to 5.2.6 focus on the distributed algorithms for sensor networks having either ideal or noisy links. In these distributed scenarios, we establish:

(i) the robustness of our distributed approaches to measurement noise for both single and multiple source fields; and

(ii) that if the communication links are ideal the distributed estimation scheme achieves the same performance as the centralized counterpart; and finally

(iii) the performance of the quantized distributed scheme can be improved by increasing K (– the number of generalized measurements used).

In Section 5.2.7 we present results that corroborate the claim outlined in Remark 3.2 concerning source estimation in bounded regions.

Section 5.3 validates both the centralized and distributed schemes for point sources on real thermal data obtained experimentally and additionally demonstrates that we can also achieve full field reconstruction having estimated the source inducing the field. We then conclude the chapter in Section 5.4.

5.2. Synthetic data simulations and results

In this section, the synthetic 2D diffusion field measurements used have been simulated numerically in MATLAB using (3.4) with (3.7) and (3.8) for fields induced by the class of localized source models.[1] Whereas COMSOL multiphysics software is used to simulate the field measurements induced by our class of non-localized sources given in Section 3.3.2. To simulate noisy measurements, zero mean additive white Gaussian noise (AWGN), $\epsilon_{n,l} \sim \mathcal{N}(0, \sigma^2)$, is directly introduced to the ideal spatiotemporal samples as follows:

$$\varphi_n^{\epsilon}(t_l) = \varphi_n(t_l) + \epsilon_{n,l}, \tag{5.1}$$

such that the noise power σ^2 is the same for all sensors. Hence, the sensors closest to the source will experience a higher SNR compared to sensors placed further away. As such the sensor measurements will have an "average" signal-to-noise ratio (SNR) of

$$\text{SNR} \stackrel{\text{def}}{=} 10 \log_{10} \left(\frac{\sum_{n=1}^{N} \sum_{l=0}^{L} |\varphi_n(t_l)|^2}{N(L+1)\sigma^2} \right). \tag{5.2}$$

Finally, these spatiotemporal samples are obtained by a collection of N sensors (randomly) deployed over a square region. All results involving multiple independent trials, unless otherwise specified, have been performed using both a new arbitrary placement of sensors, as well as, a new realization of the AWGN process for each independent trial.

5.2.1. Centralized estimation of localized sources

In what follows, we examine the estimation accuracy and performance of our centralized source estimation algorithms (i.e. Algorithms 4.2 to 4.4):

Instantaneous sources

Figure 5.1 shows results of the sequential source estimation algorithm (Algorithm 4.4); it shows that our algorithm is able to recover the desired parameters with high accuracy even

[1] Consider an instantaneous source field for example, substituting (3.6) & (3.7) into (3.4) gives the closed form expression for the field, $u(\mathbf{x}, t) = \sum_{m=1}^{M} \frac{c_m}{4\pi\mu(t-\tau_m)} e^{-\frac{\|\mathbf{x}-\boldsymbol{\xi}_m\|^2}{4\mu(t-\tau_m)}} H(t-\tau_m)$. This expression can therefore be evaluated explicitly at the sensor locations $\{\mathbf{x}_n\}_n$ and sampling instants $\{t_l\}_l$ to obtain the desired spatiotemporal sensor measurements without resorting to a grid.

in the noisy setting. We show results separately for $N = 45$ (top) and $N = 63$ (bottom) arbitrarily placed sensors respectively. Furthermore we present in Table 5.1 a summary of the variation of the normalized mean absolute error (NMAE)[2] for the activation time and intensity estimates with noise, in the single source setting. The NMAE decreases with increasing SNR, as expected; moreover at SNR= $10dB$ we achieve activation time resolution NMAE that is much less than the sampling period, for both sensor densities. When decreasing the SNR further, we notice the threshold effect characteristic of Prony's method, in that a large jump in the normalized MAE of the estimates is observed.

Non-instantaneous sources

Figure 5.2 shows our algorithm (Algorithm 4.1) operating on the spatiotemporal samples of a single non-instantaneous source field. We compare the estimated parameters against their true values and the plots demonstrate that our algorithm can successfully recover the unknown source parameters. We also summarize the normalized MAE of the estimation algorithm in Table 5.2, to highlight the algorithm's performance.

5.2.2. Centralized estimation of non-localized sources

The field due to the non-localized sources have been simulated using COMSOL Multiphysics and spatiotemporal samples are obtained at arbitrary spatial locations. We show results for noiseless data samples in Figure 5.3(a) and (b) which shows almost perfect spatial reconstruction for line and triangular sources, respectively. In both cases the sources are assumed to be instantaneous in time, to this end we use COMSOL's built-in Gaussian pulse function '`gp1()`' with a standard deviation $\sigma = 0.0025$ to simulate a delta. Then, on this data we apply the proposed algorithms to reconstruct the non-localized diffusion sources. Furthermore in order to demonstrate the noise resilience of the proposed algorithms, we artificially corrupt the measurements with AWGN and employ the proposed non-point source reconstruction schemes (i.e. Algorithms 4.2 and 4.3) to recover the line and polygonal sources, respectively. We perform 10 independent trials on noisy data utilizing a new realization of sensor placement and a new sensor noise process for each trial. The obtained results are presented in Figures 5.4 and 5.5, in which we notice that the unknown source parameters of interest (specifically the vertices and source activation times) are recovered fairly reliably in this noisy setting.

[2]For I total independent estimates $\hat{x}_i$ of some parameter x, the normalized mean absolute error of x is defined here as: $\text{NMAE}(x) = \frac{1}{x} \cdot \frac{\sum_{i=1}^{I} |x - \hat{x}_i|}{I}$.

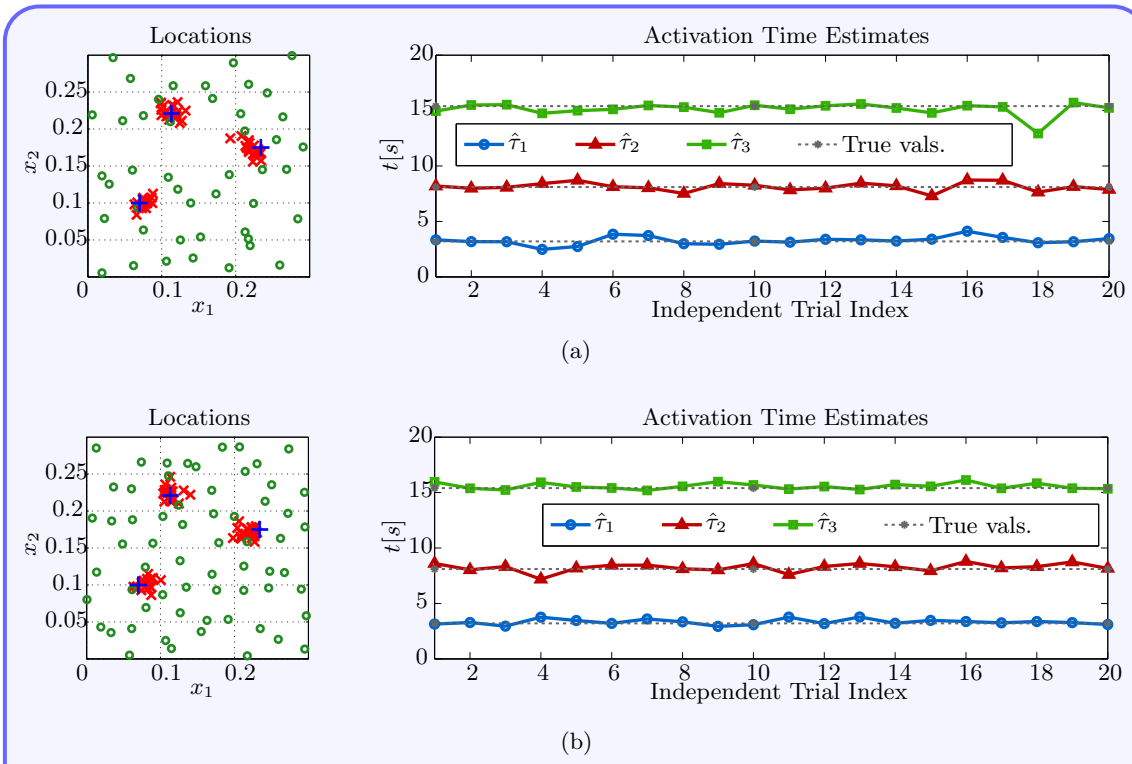


Figure 5.1.: **Estimation of $M = 3$ diffusion sources using randomly distributed sensors.** The spatiotemporal samples have SNR equal to 20dB and the results of 20 independent trials are shown. Intensities $c_1 = c_2 = c_3 = 1$; locations $\boldsymbol{\xi}_1 = (0.113, 0.221), \boldsymbol{\xi}_2 = (0.234, 0.175), \boldsymbol{\xi}_3 = (0.070, 0.100)$; and activation times $t_1 = 3.2s, t_2 = 8.1s, t_3 = 15.4s$. Field is sampled for $T_{end} = 20s$ at a frequency $\frac{1}{\Delta_t} = 2Hz$ and $K = 5$ i.e. $k = 0, 1, \ldots, 5$ for the test function family $\Psi_k(\mathbf{x}) = e^{-k(x_1 + \mathrm{j}x_2)}$. The scatter-plot shows the true source locations (blue '+'), the estimated locations (red '×') and one typical realization of the sensor distribution (green 'o') using Algorithm 4.4. We use in (a) $N = 45$ randomly distributed sensors; whilst in (b) $N = 63$ randomly distributed sensors.

Table 5.1.: Normalized Mean absolute error for an instantaneous point source estimation using Algorithm 4.1 (500 independent trials). The field is induced by $M = 1$ source with unknowns $c_1 = 1, \tau_1 = 1.213s$ and $\boldsymbol{\xi}_1 = (0.1130, 0.2210)$, sampled at $f = \frac{1}{\Delta_t} = \frac{1}{0.5} = 2Hz$ with 63 arbitrarily placed sensors. For the estimation algorithm $K = 10$ and $r = 1$ is used.

		SNR (dB)				
		5	10	20	30	40
τ_m	45 Sensors	0.6351	0.3112	0.1653	0.1345	0.1319
	63 Sensors	0.5138	0.2267	0.0922	0.0707	0.0725
c_m	45 Sensors	0.2163	0.1500	0.1202	0.1071	0.1071
	63 Sensors	0.1610	0.1140	0.0766	0.0707	0.0698

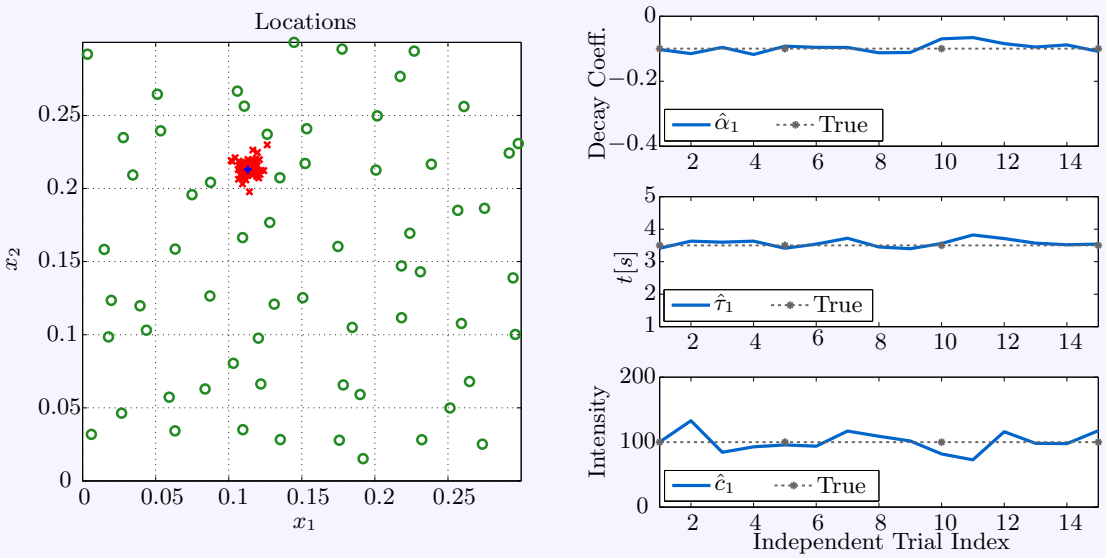


Figure 5.2.: **Estimation of $M = 1$ non-instantaneous diffusion source using 63 randomly distributed sensors.** The unknown source has intensity $c_1 = 100$, decay coefficient $\alpha_1 = -0.1$, location $\boldsymbol{\xi}_1 = (0.1130, 0.2130)$, and activation time $\tau_1 = 3.5s$. The field induced is sampled for $T = 14s$ at $\frac{1}{\Delta_t} = 2Hz$ and the samples obtained have SNR = 20dB. Moreover, we perform 15 independent trials using $K = 50$ for $\{\mathcal{R}(k)\}_{k=0}^{K}$ whilst $T_1 = 6s$, $\Delta T = 2s$ and $P = Q + 2 = 5$. The scatter-plot shows the true source locations (blue '+'), their estimated locations (red '×') and a typical realization of the sensor network distribution (green '○') Algorithm 4.1.

Table 5.2.: Normalized Mean absolute error of single source parameter estimates using Algorithm 4.1 (500 independent trials). The field is induced by a single ($M = 1$) time-varying source with unknowns $\alpha_1 = -0.1, c_1 = 100, \tau_1 = 3.5s$ and $\boldsymbol{\xi}_1 = (0.1130, 0.2130)$ and is sampled at $\frac{1}{\Delta_t} = \frac{1}{0.5} = 2Hz$ with 63 arbitrarily placed sensors. For the estimation we use $K = 50$, $T_1 = 6s$, $T_{end} = 14s$, $P = 5$, $Q = 3$, and $\Delta T = 2s$.

	SNR (dB)			
	10	20	30	40
α_m	0.9739	0.3263	0.2665	0.1285
τ_m	0.1493	0.0525	0.0329	0.0258
c_m	1.4016	0.3721	0.2819	0.1677

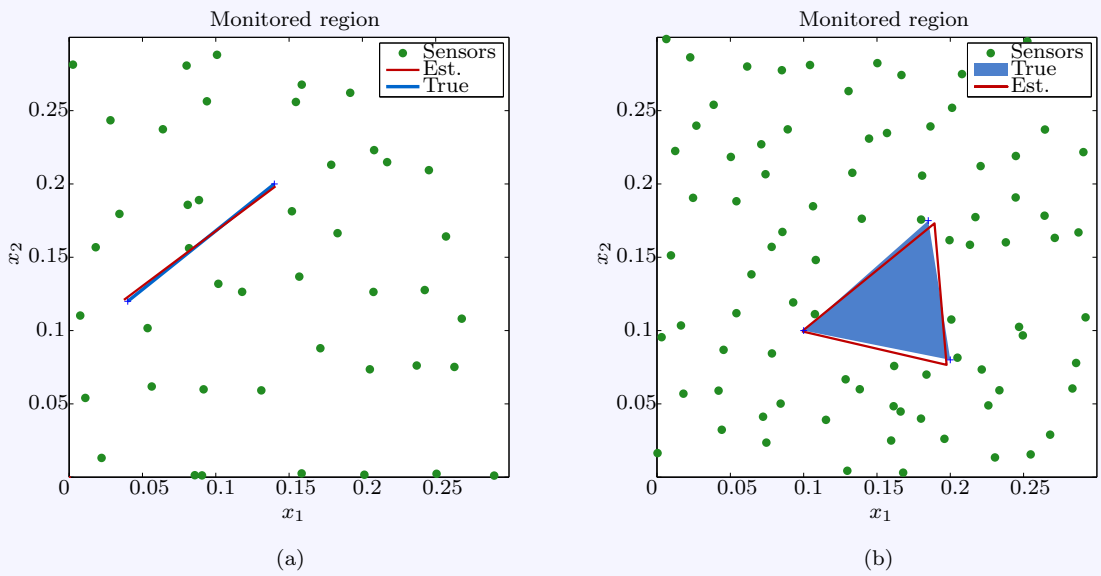


Figure 5.3.: **Centralized estimation of non-localized sources using noiseless spatiotemporal measurements obtained by randomly placed sensors.** The field is sampled at $10Hz$ for $T = 7s$ using 45 sensors in (a) and 90 sensors in (b). In addition, for the centralized estimation parameters, we use in: (a) $R = 5$ and $K = 4$ for the line source algorithm, i.e. Algorithm 4.2, whereas (b) $R = 5$ and $K = 6$ for the triangular source algorithm, i.e. Algorithm 4.3.

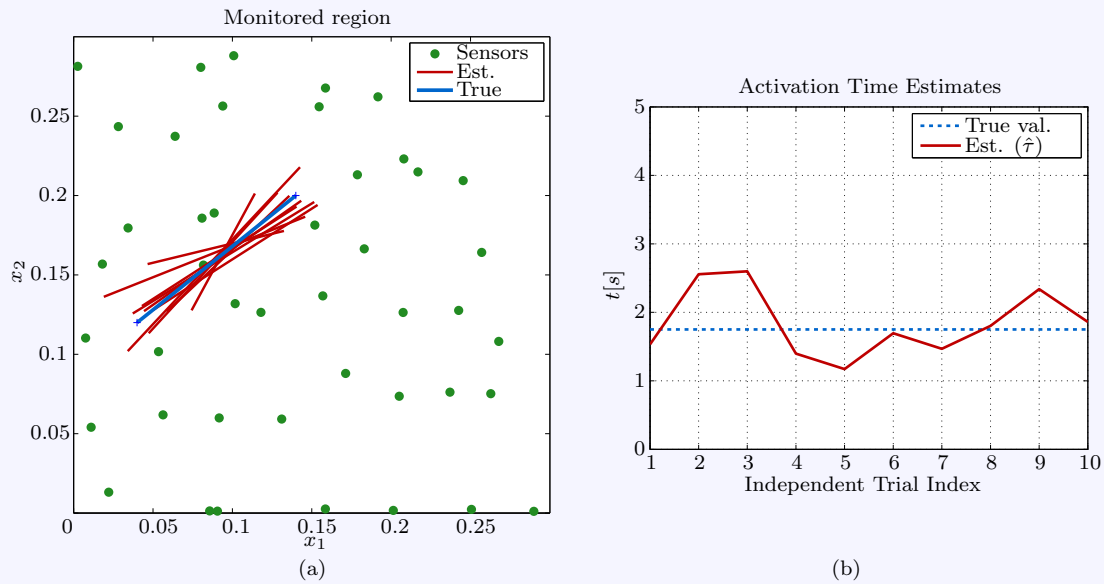


Figure 5.4.: **Line source estimation using noisy spatiotemporal measurements.** The field is sampled at $10Hz$ for $T = 10s$ using 45 arbitrarily placed sensors, with measurement SNR= 20dB. For the spatial and temporal sensing functions family, $K = 6$ and $R = 5$ respectively using Algorithm 4.2. For each trial we show in: (a) the estimated line source, and in (b) the corresponding activation time estimates.

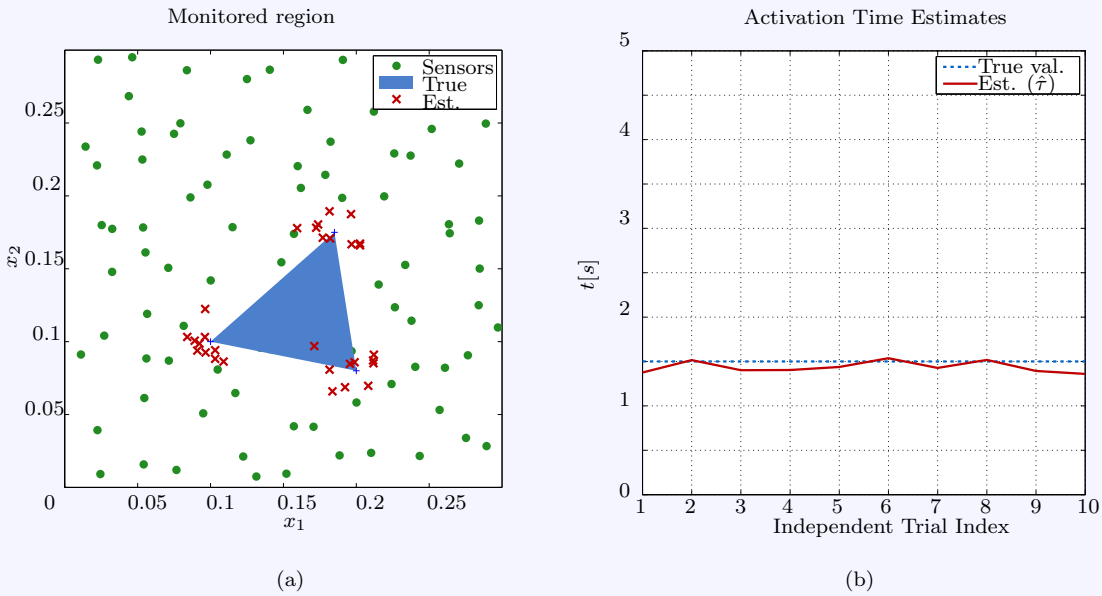


Figure 5.5.: **Triangular source estimation using noisy spatiotemporal measurements.** The field sampled at $10Hz$ for $T = 10s$ using 90 arbitrarily placed sensors, with measurement SNR= 35dB. For the spatial and temporal sensing functions, $K = 9$ and $R = 8$ respectively using Algorithm 4.3. For each trial we show in: (a) the vertices of the estimated triangular source, and in (b) the corresponding activation time estimates.

5.2.3. Approximation errors due to the integral discretization

In this section we examine the effectiveness of the quadrature techniques used when obtaining the generalized measurements from spatiotemporal samples. Specifically, we provide simulation results which suggests that the error due to approximating the integrals in (3.17) (and (3.21)) does not affect the estimation of the sources of the field, in that, even at fairly high SNRs, the noise in the sensor measurements dominates the errors in the reconstruction. To see this, we compare the localization results obtained by applying Prony's method on two sequences, namely: $\{\mathcal{R}_\epsilon(k)\}_{k=0}^K$ and $\{\tilde{\mathcal{R}}_\epsilon(k)\}_{k=0}^K$. The first sequence, i.e. $\{\mathcal{R}_\epsilon(k)\}_k$ is obtained from noisy samples in the usual fashion; whilst $\{\tilde{\mathcal{R}}_\epsilon(k)\}_k$ is obtained by adding an equivalent noise process to the exact power-sum series $\{\mathcal{R}(k)\}_k$ (here $\{\mathcal{R}_\epsilon(k)\}_k$ represents the noisy continuous measurement scenario, where the integrals in (3.17) are known precisely). Figure 5.6 shows the standard deviation of the estimated spatial locations of $M = 2$ sources, separated by $0.08m$, using $\{\mathcal{R}_\epsilon(k)\}_k$ (dashed lines) and $\{\tilde{\mathcal{R}}_\epsilon(k)\}_k$ (solid lines) respectively. Observe that for realistic SNRs of interest, i.e. 30dB and below, the performance of the location recovery coincides with that of the ideal, full-field measurement, case.

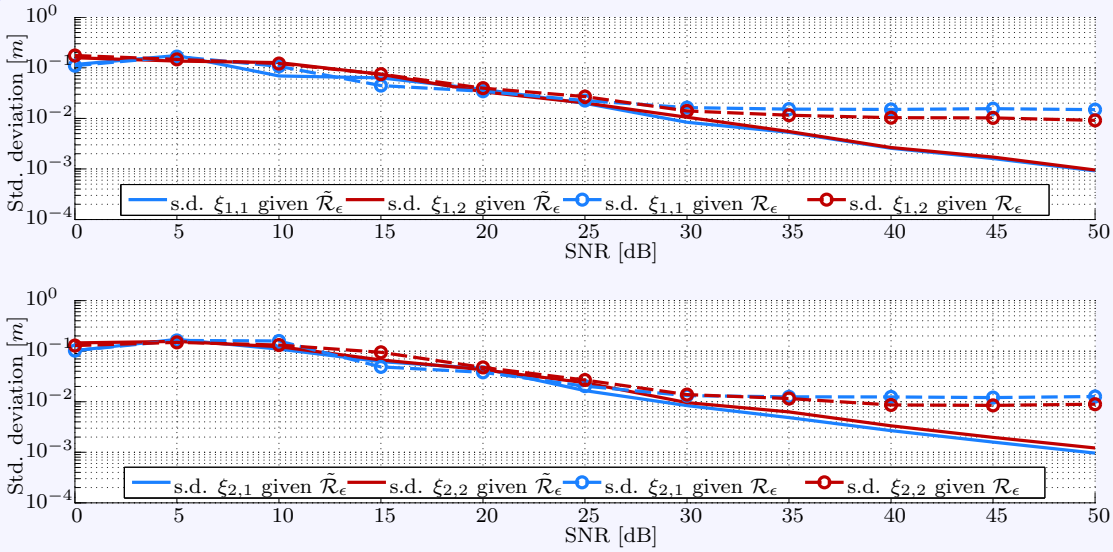


Figure 5.6.: **Location estimation performance.** Standard deviation (500 trials) for localization estimates of 2 sources from 63 randomly distributed sensors, using Algorithm 4.1 with $K = 5$. $c_1 = c_2 = 1$; $\boldsymbol{\xi}_1 = (0.23, 0.15)$, $\boldsymbol{\xi}_2 = (0.15, 0.15)$ and $\tau_1 = \tau_2 = 1.2s$. $T_{end} = 10s$ and $\frac{1}{\Delta_t} = 1Hz$.

5.2.4. Distributed estimation over sensor networks: noiseless channels

We present numerical simulation results aimed at evaluating our proposed distributed source estimation algorithms assuming that sensors in the distributed network can communicate locally over noiseless channels. The sensor distribution is modelled by the RGG

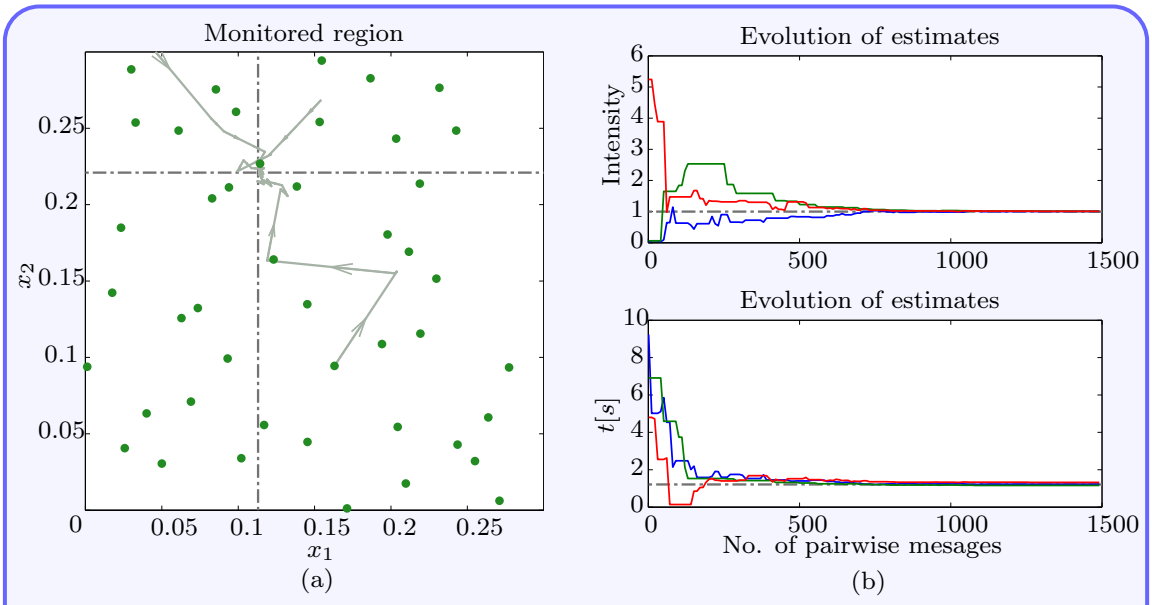


Figure 5.7.: **Single diffusion source estimation using distributed SNs with noiseless links.** The evolution of the estimated source parameters with increasing number of pairwise message exchanges (Algorithm 4.5). Sensor measurements are assumed to be noiseless and we use $K = 1$ and $r = 1$ for the generalized measurements $\{\mathcal{Q}(k, r)\}_k$. (a) Evolution of location estimates; true source location is where the dashed lines intersect. (b) Evolution of intensity (top) and activation time (bottom) estimates.

$\mathcal{G}(N, r_{\text{con}})$, where N and r_{con} are stated for each simulation. We consider the following scenarios:

Single source estimation: convergence of estimates

Figure 5.7 shows the performance of the distributed estimation algorithm in the single source setting given noiseless spatiotemporal measurements. Specifically $M = 1$ and the true parameters of the source inducing the field are: $c_1 = 1$, $\boldsymbol{\xi}_1 = (0.113, 0.221)$ and $\tau_1 = 1.213s$. The induced field is sampled, assuming no measurement noise, using $N = 45$ randomly placed sensors at a rate $\frac{1}{\Delta_t} = 1Hz$ for $T_{end} = 10s$. Furthermore, the SN and associated communication links are modelled by $\mathcal{G}(N, 0.4)$. The communication between the sensors is assumed to be unquantized, hence simple pairwise gossiping is used. In particular Figure 5.7(a) shows the evolution of the location estimates (solid directed grey lines), for three randomly chosen sensors, with increasing number of pairwise message exchanges. As we expect the location estimates progressively tend towards the true source location. Figure 5.7(b) shows similarly the evolution of the intensity (top) and activation time (bottom) estimates with each pairwise exchange. The three curves in each plot show the evolution of the estimates due to (the same) three randomly chosen sensors. Again we clearly notice, as expected, that the estimates converge to the desired values after several

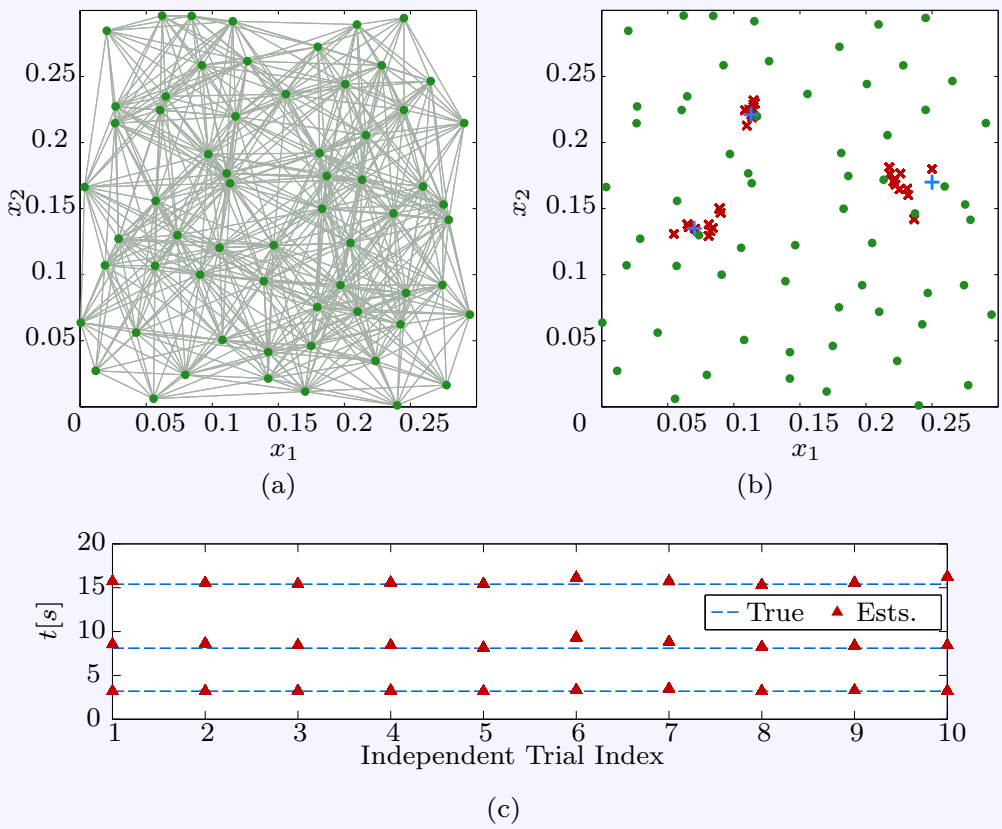


Figure 5.8.: **Multiple diffusion source estimation over distributed SNs with noiseless links.** The unknown source parameters estimated by each sensor using Algorithm 4.5 is shown for 10 independent trials. Communication links are ideal, and so messages are unquantized. Furthermore, the spatiotemporal samples are assumed noisy with SNR= 20dB. Here we use $K = 10$ and $r = 1$. (a) shows one realization of the SN (sensors and links), the scatterplot in (b) shows the true source locations (blue '+'), their estimates (red '×') and a realization of the sensor locations (green '•'), whilst (c) shows the estimated activation times.

pairwise message exchanges.

Multiple source estimation: robustness

Next we evaluate the proposed distributed estimation algorithm in the case of multiple sources given noisy sensor measurements. Specifically, field induced by $M = 3$ sources with parameters $c_1 = c_2 = c_3 = 1$, $\boldsymbol{\xi}_1 = (0.113, 0.221), \boldsymbol{\xi}_2 = (0.250, 0.170), \boldsymbol{\xi}_3 = (0.070, 0.135)$ and $\tau_1 = 3.2s, \tau_2 = 8.1s, \tau_3 = 15.4s$, is sampled at a rate $\frac{1}{\Delta_t} = 2Hz$ for $T_{end} = 25s$, in the presence of measurement noise (SNR = 20dB) using a network of $N = 63$ nodes given by the RGG $\mathcal{G}(N, 0.4)$. In Figure 5.8, we present results from 10 independent trials for this multiple source setting. Specifically, Figure 5.8(a) shows one realization of the SN, the scatterplot in Figure 5.8(b) shows estimates of the source locations, whilst Figure 5.8(c) indicates the

estimated activation times at convergence. In line with the theory, the unknown source parameters are recovered reliably even when the sensor measurements are noisy.

5.2.5. Distributed estimation over sensor networks with noisy channels

The results presented here examine the proposed distributed diffusion source estimation algorithms for communication constrained inter-sensor channels.

Single source estimation

The single source field with $c_1 = 1$, $\boldsymbol{\xi}_1 = (0.113, 0.221)$ and $\tau_1 = 1.213s$, is considered. It is sampled (assuming no measurement noise) with $N = 45$ randomly placed sensors at $\frac{1}{\Delta_t} = 1Hz$ for $T_{end} = 10s$. Furthermore, a 10-bit uniform quantizer with dynamic range [-20,20] is assumed. In Figure 5.9 $K = 1$, whereas $K = 5$ in Figure 5.10 for $\{\mathcal{Q}(k, 1)\}_{k=0}^{K}$.

Recall that under pairwise symmetric quantized gossip, as discussed in Section 4.5.1, the sensors in the network converge to a distribution of values. Hence we expect that the sensor n obtains, in general, a (slightly) different estimate—from the other sensors in the network—for each estimated source parameter. In the scatterplots of Figure 5.9(b) and Figure 5.10(b), we plot the source location estimates for each of the 45 sensors in the network, obtained at convergence, with $K = 1$ and $K = 5$ respectively. Observe that the variance of these location estimates is smaller when $K = 5$, this is due to the fact that Prony's method and its variations are more robust to noise when K is large. Furthermore, we also show, with each gossip iteration, the evolution of the source location estimates by three randomly chosen sensors for $K = 1$ in Figure 5.9(c) and for $K = 5$ in Figure 5.10(c), where each trajectory corresponds to evolution of estimates due to one of the three chosen sensors. Similarly, Figure 5.9(d) and Figure 5.10(d) shows the evolution of the intensity (top) and activation time (bottom) estimates with increasing number of pairwise message exchanges, for $K = 1$ and $K = 5$ respectively. In both cases the estimates converge with the variance of the latter being smaller, again due to the increase in estimation accuracy of Prony's method for higher values of K.

Multiple source estimation

In the following results, we aim to evaluate the robustness of the proposed distributed source estimation algorithm for quantized inter-sensor communication. Specifically we consider the effectiveness of the algorithm at estimating multiple sources in the presence of measurement noise. The field induced by $M = 3$ sources—with $c_1 = c_2 = c_3 = 1$, $\boldsymbol{\xi}_1 = (0.113, 0.221), \boldsymbol{\xi}_2 = (0.250, 0.170), \boldsymbol{\xi}_3 = (0.070, 0.135)$ and $\tau_1 = 3.2s, \tau_2 = 8.1s, \tau_3 = 15.4s$—is sampled at a rate $\frac{1}{\Delta_t} = 2Hz$ for $T_{end} = 25s$, assuming measurement noise with SNR = 20dB using $N = 63$ randomly deployed sensors. Moreover, the SN is the RGG $\mathcal{G}(N, 0.4)$ and a 15-bit uniform quantizer with dynamic range [-20,20] is assumed.

Figure 5.11 shows results of 10 independent trials for this multiple source scenario. Specifically, in Figure 5.11(a) we show one realization of the SN, the scatterplot in Figure 5.11(b)

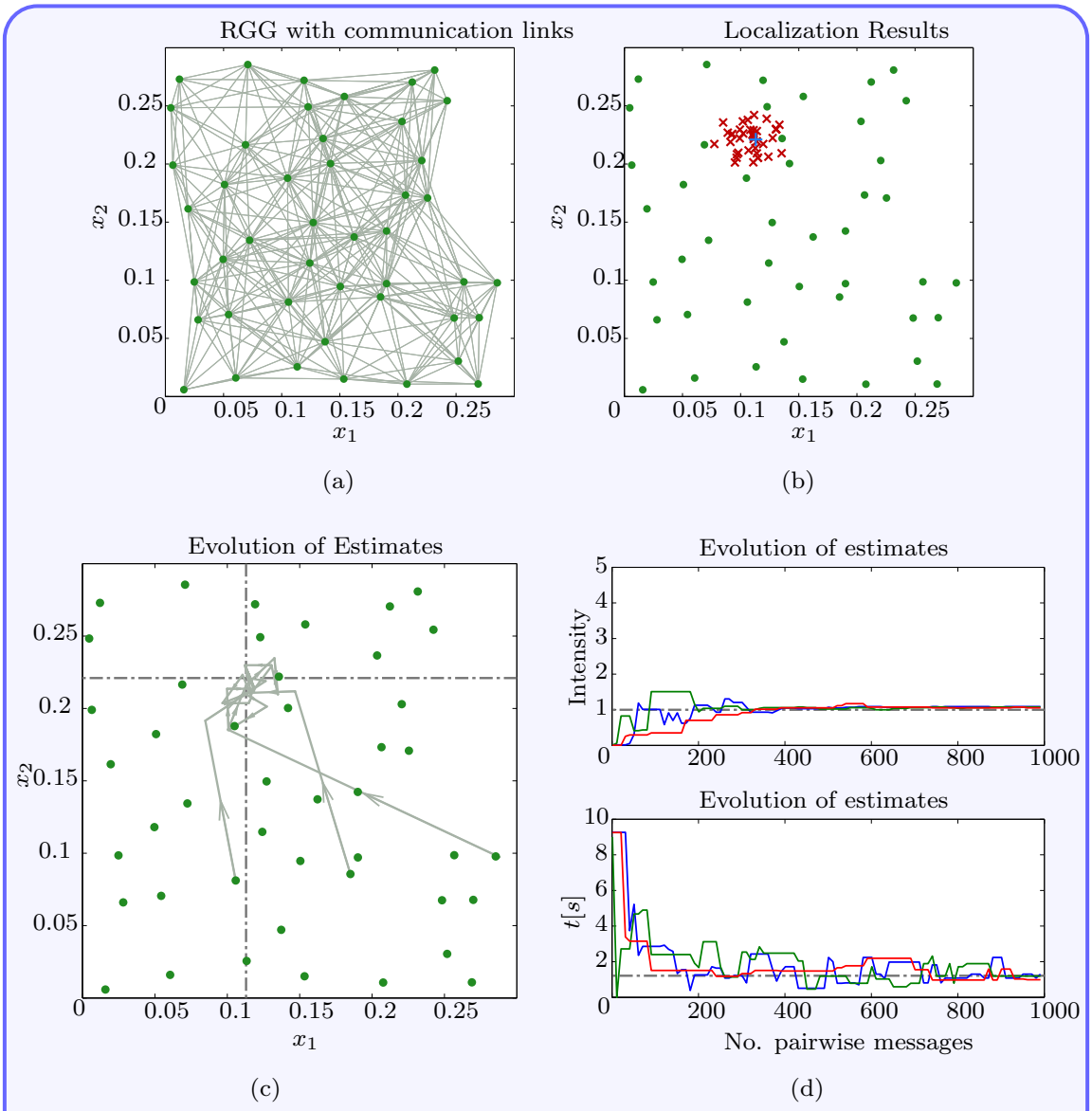


Figure 5.9.: **Single diffusion source estimation using distributed SNs with noisy links ($K = 1$).** We assume noiseless sensor measurements and Algorithm 4.5 with quantized gossiping is utilized for $K = 1$ and $r = 1$. In (a) $\mathcal{G}(45, 0.4)$ and in (b) a scatterplot of the location estimates at convergence. (c) Evolution of location estimates (directed grey lines) along with the true location (intersection of the dashed lines). (d) Evolution of source intensity (top) and activation time (bottom) estimates for three randomly chosen nodes.

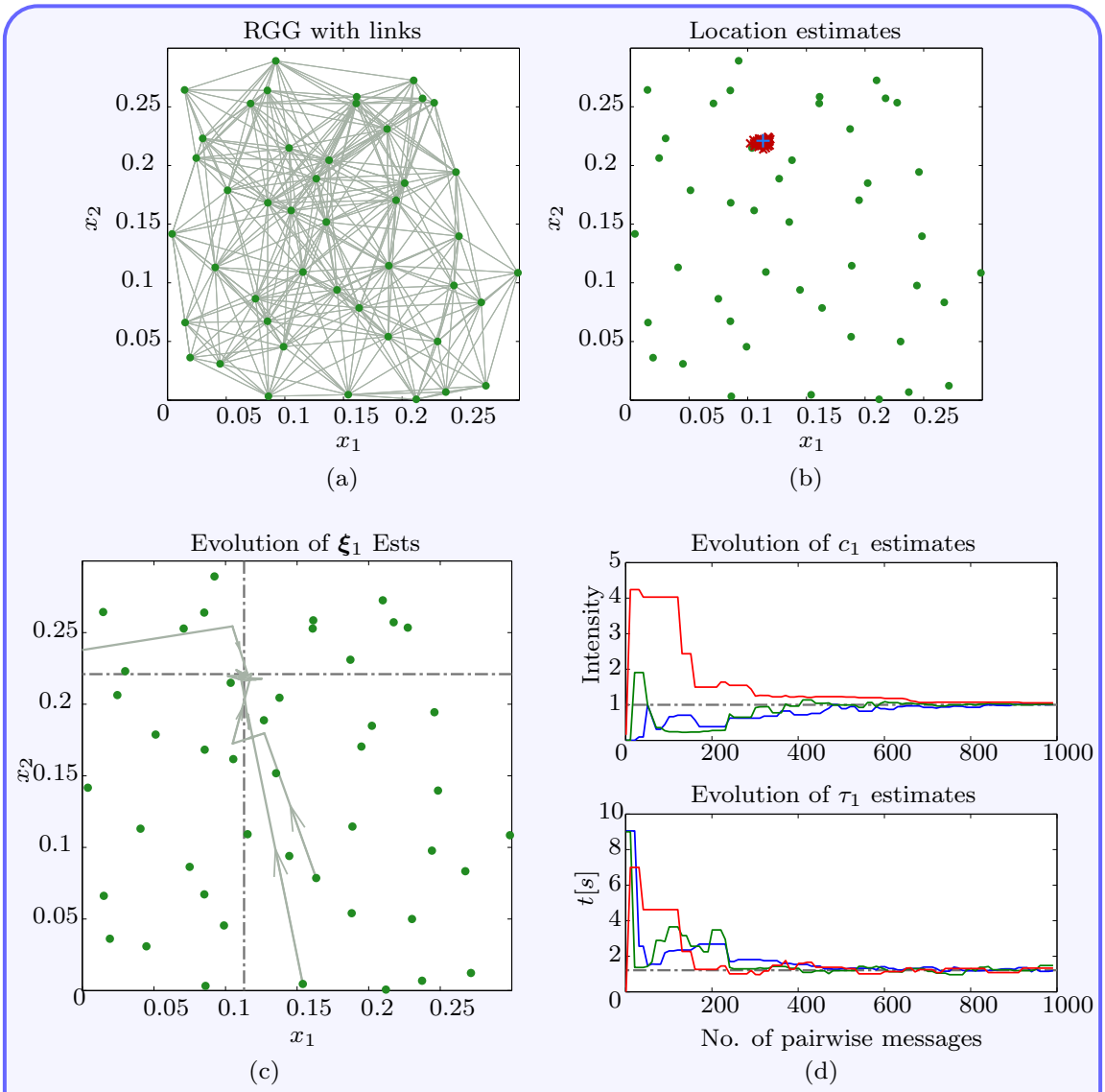


Figure 5.10.: **Single diffusion source estimation using distributed SNs with noisy links ($K = 5$).** We assume noiseless sensor measurements and Algorithm 4.5 with quantized gossiping is utilized for $K = 5$ and $r = 1$. In (a) $\mathcal{G}(45, 0.4)$ and in (b) a scatterplot of the location estimates at convergence. (c) Evolution of location estimates (directed grey lines) along with the true location (intersection of the dashed lines). (d) Evolution of source intensity (top) and activation time (bottom) estimates for three randomly chosen nodes.

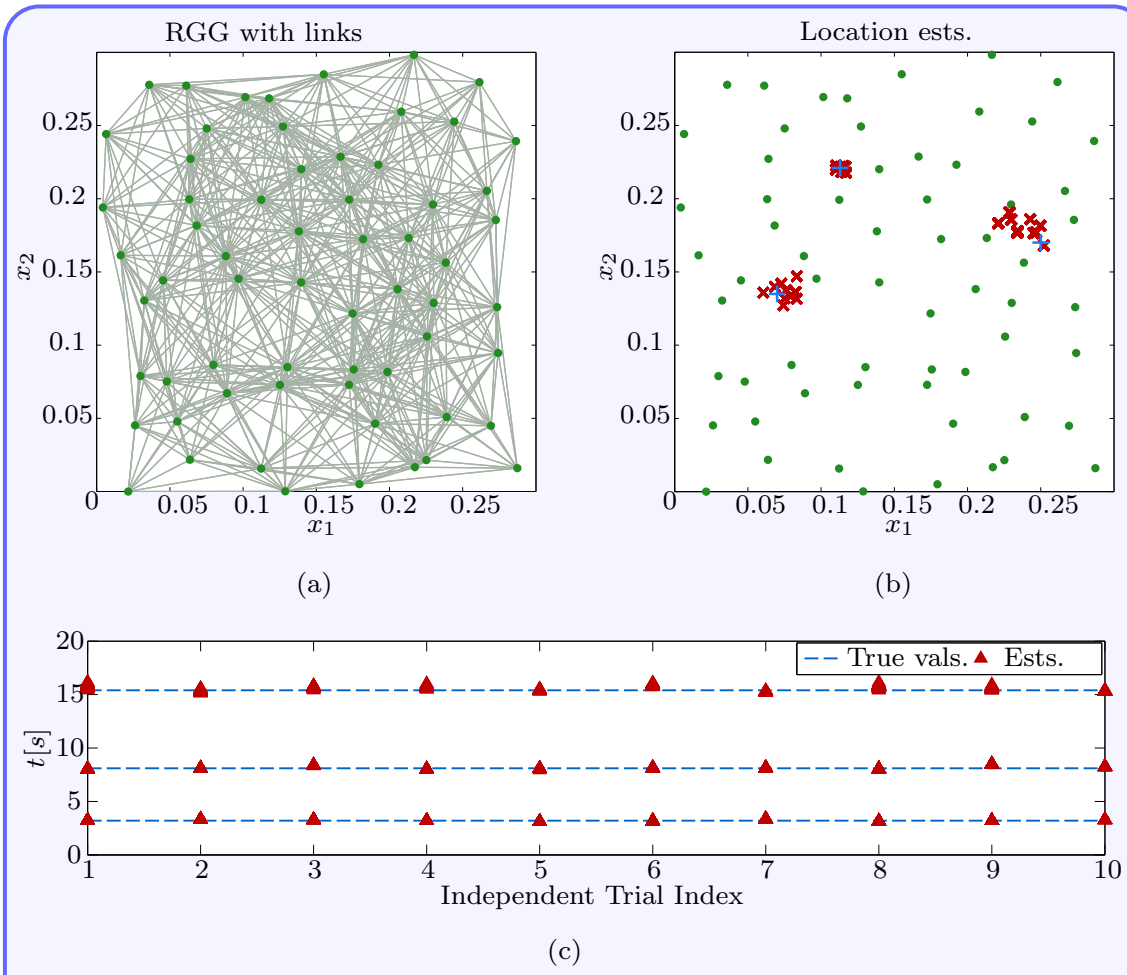


Figure 5.11.: **Multiple diffusion source estimation over SNs with noisy links (quantized gossip).** Noisy sensor measurements with SNR= 20dB, and noisy inter-sensor links are assumed. To estimate the unknowns we utilize Algorithm 4.5 under quantized gossiping for $K = 10$ and $r = 1$ and show results (at convergence) for 10 independent trials. (a) a realization of $\mathcal{G}(63, 0.4)$, whilst in (b) the scatterplot shows the true (blue '+') and estimated (red '×') source locations, along with a typical sensor distribution (green '•'). (c) Estimated activation times due to each sensor upon convergence.

shows distribution of the estimates of the source locations, whilst Figure 5.11(c) indicates the distribution of the estimated activation times upon convergence. In these plots we notice that the estimated source parameters are close to the desired as expected.

5.2.6. Comparing performance: Centralized vs Distributed estimation

We now present some statistical results in order to compare quantitatively the distributed algorithm with its centralized counterpart.

In Table 5.3, we show the normalized MAE of the estimated diffusion source parameters

obtained using the centralized and both distributed (unquantized and quantized communication links) schemes. To obtain the reported statistics, we simulate the field induced by a single source (M=1), with $c_1 = 1, \tau_1 = 1.23s$ and $\boldsymbol{\xi}_1 = (0.1130, 0.2210)$, for T=10s and sample it with N=45 randomly placed sensors at a frequency $\frac{1}{\Delta_t} = 1$Hz. Next, for the noisy scenarios, we first corrupt the spatiotemporal samples with AWGN before applying the centralized, unquantized distributed and quantized distributed algorithms on the samples to recover the unknown source parameters. We perform the experiment for different sensor noise levels (i.e. noiseless and SNR= $\{10, 20\}$dB), and repeat it 1000 times (for statistical significance). In each experiment we use a new arbitrary sensor placement and a new sensor noise realization for the noisy case; the MAE of each parameter is then computed from the estimates over the 1000 trials. We repeat this for $K = 1$ and $K = 3$ (and $r = 1$ for $\{\mathcal{Q}_\epsilon(k, r)\}_{k=0}^{K}$), to show that the average error decreases gracefully with increasing K. Moreover, the centralized and unquantized distributed estimation schemes both perform similarly in all scenarios. In the noiseless case, for instance, when $K = 1$ the average percentage error in the intensity, activation time and location estimates are around $8.1\%, 12.3\%$ and $(4.4\%, 2.8\%)$ respectively. The MAEs for all estimates are greater under quantized communications (q=10 bits and dynamic range of quantizer is $[-20, 20]$) than the unquantized counterpart. This observation is unsurprising and is a result of the quantization.[3] In this case given noiseless sensor measurements, when $K = 1$, we observe average percentage errors of $10.1\%, 41.3\%$ and $(11.5\%, 5.7\%)$ in the estimates of c_1, τ_1 and $(\xi_{1,1}, \xi_{1,2})$ respectively, and a lot of improvement can be gained by increasing K. Going from $K = 1$ to $K = 3$, the errors are roughly halved for all parameters to about $7.4\%, 18.5\%$ and $(6.5\%, 3.1\%)$ respectively. Given noisy spatiotemporal samples the NMAE, and hence the percentage errors, for all algorithms increase but this increase can be largely compensated by increasing K. As an example, the average percentage error at 20dB for the activation time estimate using the quantized distributed estimation drops from around 48% (at $K = 1$) to around 28% (at $K = 3$).

Furthermore, we now compare more closely the performance of the unquantized distributed and the centralized schemes. Particularly, we use the mean squared error (MSE) of the estimates produced by each algorithm as a metric for evaluating performance. Moreover through numerical simulations, we demonstrate that the performance of the unquantized distributed and centralized schemes coincide (as claimed in Proposition 4.1). To obtain the MSE, a single source field is simulated with the same setup as the MAE simulations above. The field samples are then corrupted with AWGN and the desired source parameters are estimated using both algorithms (centralized and unquantized distributed) with $K = 1$. We repeat the experiment 5000 times, with each new trial using a new sensor noise realization (but fix the topology). The experiment is repeated for several SNRs and the MSEs of the estimate are computed and displayed in Figure 5.12 along with the Cramer-Rao bound (CRB), see Appendix A.1 for an analytic expression of the CRB for this particular single

[3]This includes quantization errors and in addition since the quantized gossip scheme used only allows the sensors converge to a distribution of values rather than a single value, this will also introduce errors.

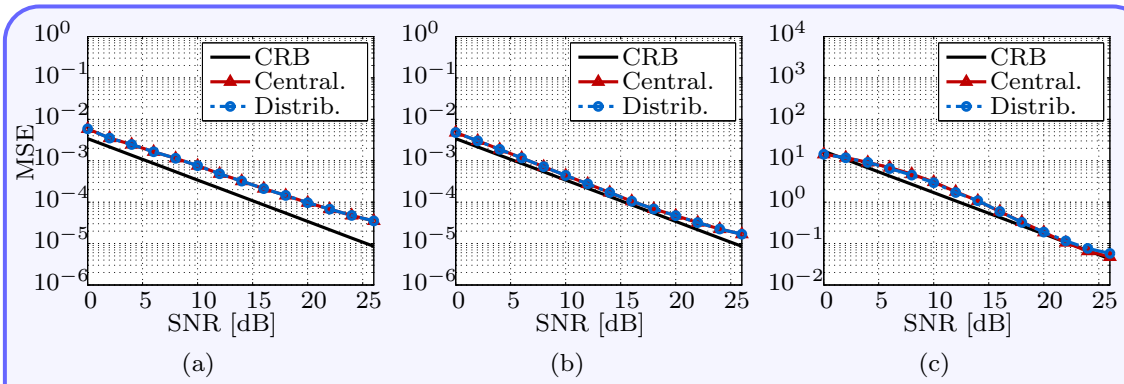


Figure 5.12.: **Performance of the centralized (Algorithm 4.1) and distributed (Algorithm 4.5) estimation algorithms.** A single source field, with c_1=1, τ_1=1.213, $\boldsymbol{\xi}_1$=(0.1130, 0.2210) is considered and its source parameters are estimated from the (noisy) spatiotemporal samples taken at $\frac{1}{\Delta_t}$=1Hz over T=10s using N=45 sensors, with the spatial sensing function family chosen such that $K = 1$. We show the MSE of the centralized and distributed algorithms computed using results from 5000 independent trials for varying signal-to-noise ratios: in (a) MSE of $\xi_{1,1}$, (b) MSE of $\xi_{2,1}$ and, (c) MSE τ_1.

Table 5.3.: **Normalized MAE of the centralized (Algorithm 4.1) and distributed (Algorithm 4.5) estimation algorithms.** The induced single ($M = 1$) source field, with c_1=1, τ_1=1.213, $\boldsymbol{\xi}_1$=(0.1130, 0.2210), is sampled at $\frac{1}{\Delta_t}$=1Hz over T=10s using N=45 sensors. The samples are corrupted with AWGN to obtain the specified SNR. We choose $r = 1$ and $K = 1$ and $K = 3$ for the estimation algorithms and present the normalized MAE statistic computed with results from 1000 independent trials at the chosen SNR levels.

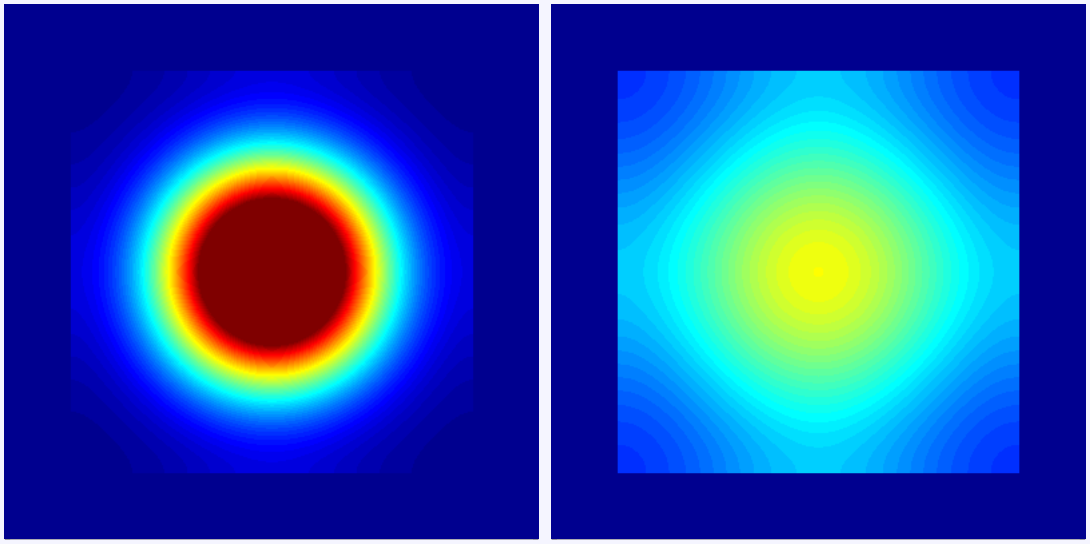
		Centralized		**Unquantized (distr.)**		**Quantized (distr.)**	
		K=1	K=3	K=1	K=3	K=1	K=3
No noise	c_m	0.0811	0.0683	0.0811	0.0684	0.1017	0.0740
	τ_m	0.1375	0.1277	0.1231	0.1128	0.4127	0.1883
	$\xi_{1,m}$	0.0439	0.0306	0.0439	0.0306	0.1151	0.0644
	$\xi_{2,m}$	0.0285	0.0219	0.0285	0.0219	0.0571	0.0313
20dB	c_m	0.4385	0.1025	0.4385	0.1026	0.5939	0.1030
	τ_m	0.2794	0.2565	0.2583	0.2267	0.4928	0.2887
	$\xi_{1,m}$	0.0861	0.0611	0.0861	0.0620	0.1893	0.0767
	$\xi_{2,m}$	0.0600	0.0415	0.0600	0.0415	0.1040	0.0451
10dB	c_m	0.2417	0.2027	0.2417	0.2031	0.2785	0.2171
	τ_m	0.7132	0.7310	0.6952	0.7102	0.8321	0.7240
	$\xi_{1,m}$	0.2081	0.1808	0.2081	0.1874	0.2627	0.1951
	$\xi_{2,m}$	0.1135	0.0990	0.1135	0.0995	0.1310	0.1037

source estimation problem. The location plots for each dimension are shown in Figure 5.12 (a) and (b) respectively, whilst the MSE for the activation time with its CRB is shown in Figure 5.12(c). We can see that the CRB is approximately achieved using both schemes and more importantly the performance (MSEs) of both schemes exactly coincide for the range of SNRs of interest.

5.2.7. Estimating field sources in rooms

We simulate the diffusion field in a square region, using the well-documented method of *image sources*, which is based on the introduction of virtual sources to model reflections due to the edges of the room. The edges of the bounded region are assumed to be *perfectly insulating*, so that the field incident upon them is reflected back (i.e. it does not leak out of the region through its walls). The resulting field therefore diffuses through the square medium as shown in Figure 5.13.

We present results in Figure 5.14 which shows that our algorithm is able to recover the multiple diffusion field sources even when they are in a bounded region.



(a) Time $t = 20s$ (b) Time $t = 50s$

Figure 5.13.: **Diffusion in bounded region.** Spatial field distribution of a single source in a bounded square region at different time instants after source activation.

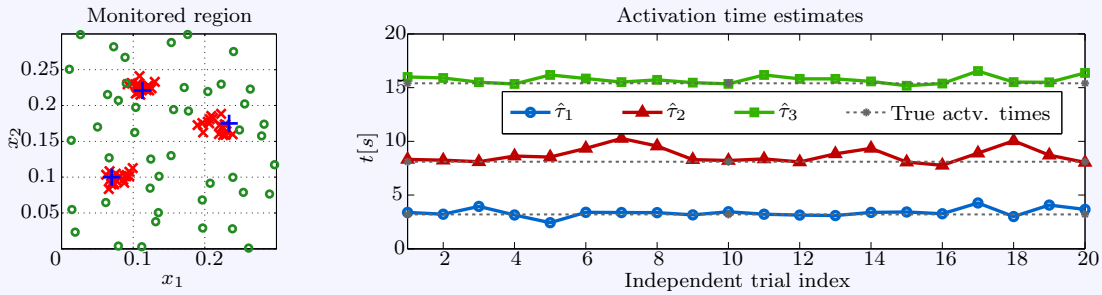


Figure 5.14.: **Multiple diffusion source estimation in a bounded region.** The $M = 3$ diffusion sources have $c_1 = c_2 = c_3 = 1$, $\boldsymbol{\xi}_1 = (0.113, 0.221), \boldsymbol{\xi}_2 = (0.234, 0.175), \boldsymbol{\xi}_3 = (0.070, 0.100)$, and $\tau_1 = 3.2s, \tau_2 = 8.1s, \tau_3 = 15.4s$. The field induced is sampled at $\frac{1}{\Delta_t} = 2Hz$ for $T_{end} = 25s$, using 45 sensors placed randomly inside the square room. The samples have SNR= 20dB. Estimation is performed using Algorithm 4.4 with $K = 11$ and $r = 1$. The estimated locations (left) and activation times (right) are shown.

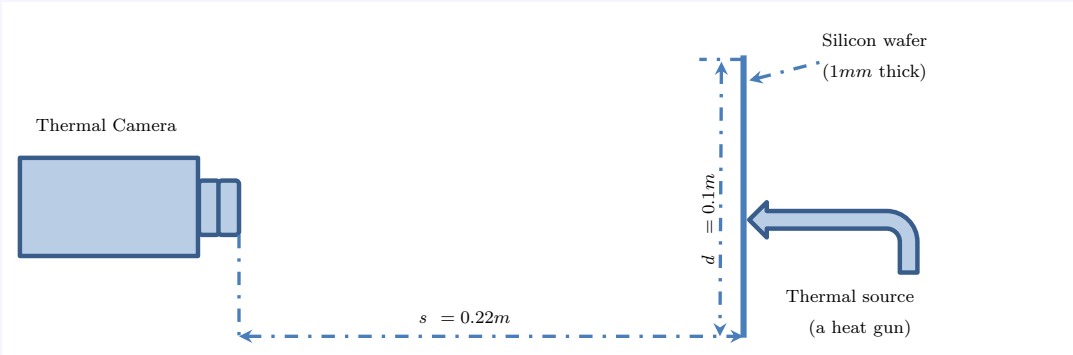


Figure 5.15.: **Real data experimental setup.**

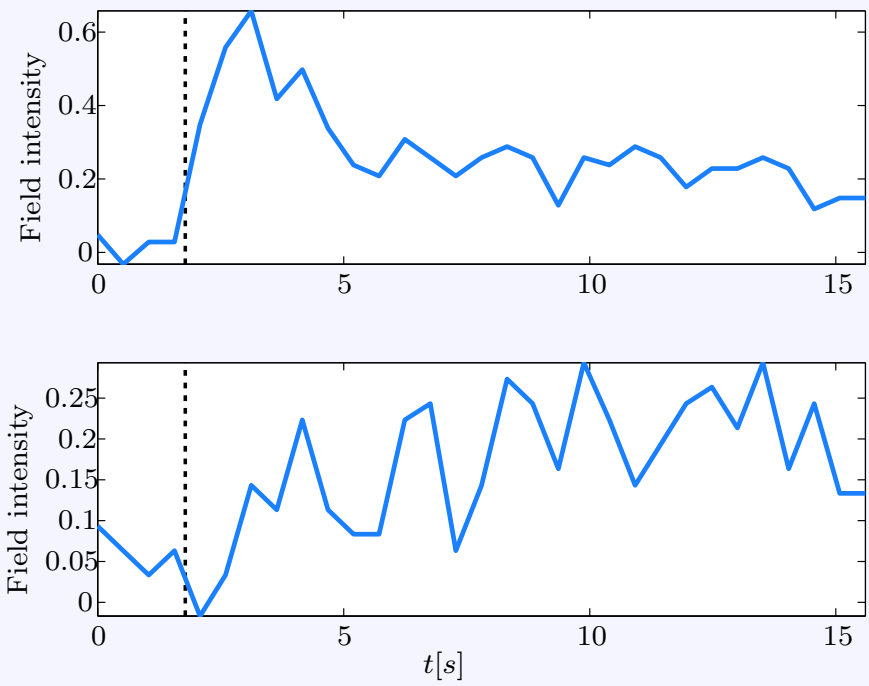


Figure 5.16.: **Measurements of two monitoring sensors obtained at different spatial locations.** The dotted vertical line in each plot indicates the instant of source activation.

5.3. Real data experiments

In this section, we utilize real temperature data measurements obtained using a thermal imaging camera to further validate the proposed source estimation algorithms. Firstly, we outline the experimental set-up by providing a brief overview of how the thermal spatiotemporal samples are obtained and then we present some estimation results obtained when the proposed algorithms are applied on the measured data.

5.3.1. Acquiring the experimental data

In this experiment, a silicon wafer disc of diameter $d = 0.1m$ is used as the diffusion medium. The wafer is placed $s = 0.22m$ from a thermal camera, with the disc lying on the focal plane of the camera; this arrangement, shown in Figure 5.15, allows us to measure the entire surface temperature of the silicon plate. We obtain thermal recordings of the disc, at specified frame rates (we use 10Hz and 25Hz). Then a heat gun, with a $1mm$ nozzle, is used to apply a localized and instantaneous initial heat source on the reverse side of the silicon plate (i.e. the face opposite that seen by the camera). We continue recording the thermal images for $15s$. We pre-process the recordings by averaging the first few frames and subtracting this average from all frames in the video. This has the effect of imposing the initial temperature distribution of $\approx 0°C$ at all spatial locations at time $t = 0s$.[4]

The camera is properly calibrated so that true xy-locations can be assigned to the 384×288 pixels of each frame. Then spatial sampling, in our setup, corresponds to obtaining samples at a few (specified) pixel locations, and these spatial locations are chosen randomly. The temporal evolution of two such sensors are shown in Figure 5.16. Moreover, the true value of the source location is the center of the region where the heat source is first observed. For the true activation time, since the frame rate is known, we assume that the source is activated at the frame where we first observe a hot region minus half the sampling interval.

5.3.2. Source estimation over centralized SNs: real data

The results of our experimentation with real thermal data are summarized in what follows. Figure 5.17(a) shows the complete temperature distribution of the monitored region immediately after source activation with the hottest (light) region of the map indicating the true source location. Furthermore, we show the estimated source location using a red cross '×'; this estimate has been obtained by applying our proposed algorithm on spatiotemporal measurements obtained at the 13 locations marked by black circles '∘'. Note that the sampling frequency $f = \frac{1}{0.52}Hz$, of the sensors in this experiment is much lower than the frame rate of the camera. This is achieved by downsampling the actual time measurements.

To demonstrate the robustness of the algorithm to the choice of sensor locations, we draw randomly a new set of 13 locations and apply Algorithm 4.1 (with $M = 1$) on the

[4]Note that, due to external factors, we obtain noisy non-zero measurements as seen in Figure 5.16 and Figure 5.17(a).

new spatiotemporal samples. This experiment is repeated 20 times and a scatterplot of the estimated source location and the activation time estimates is shown in Figure 5.17(b). The obtained estimates vary marginally about the true values. For statistical significance, we repeat this experiment 1000 times and present the MAE of the location and activation time estimates in Table 5.4. For the location estimates the MAEs are small compared to the dimensions of the monitored region, and also smaller than the average inter-sensor separation. Similarly, the normalized MAE of the activation time is around 0.0867, which is almost an order of magnitude smaller than the temporal sampling interval ($0.52s$). Hence on average we observe an absolute error of around 8.67% on the activation time estimates.

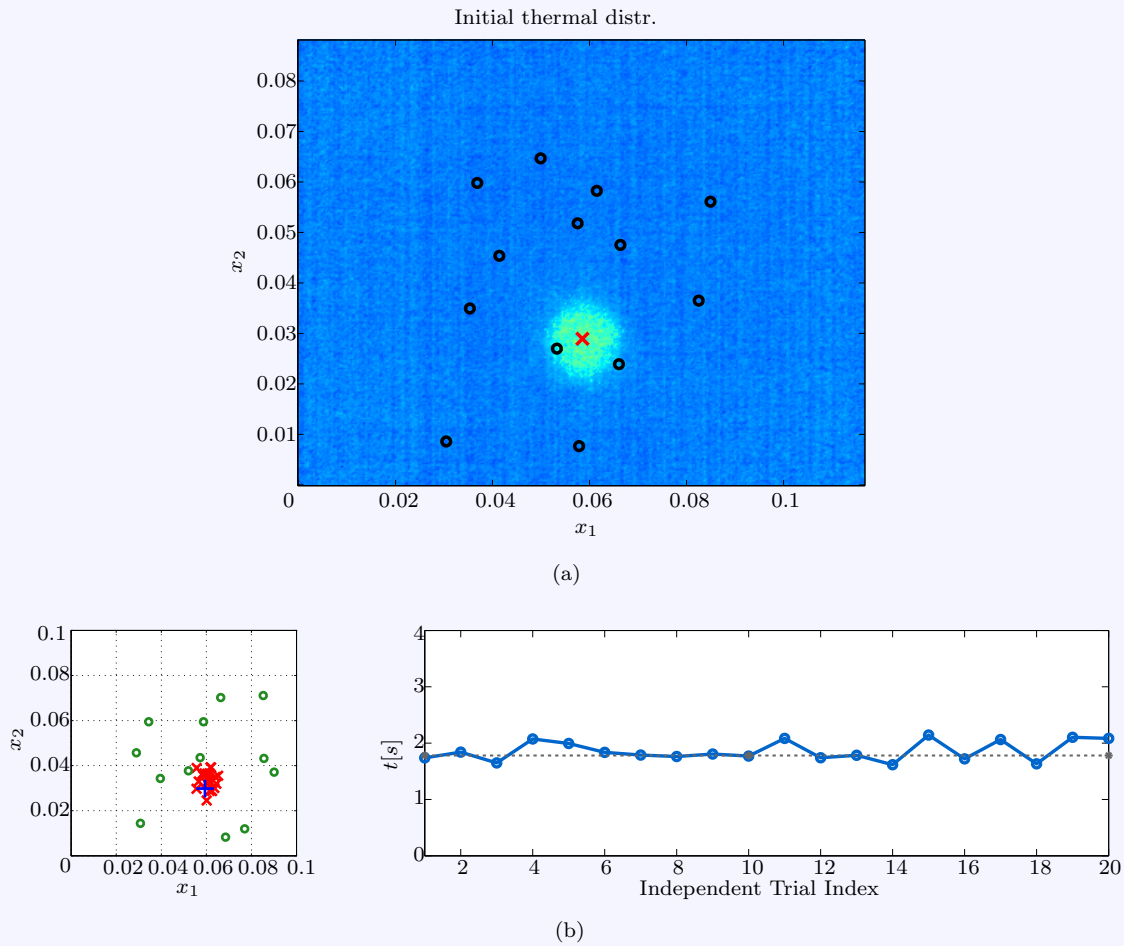


Figure 5.17.: **Centralized estimation of a single instantaneous heat source using real thermal data.** Samples are obtained from the thermal images by choosing 13 spatial locations (the circles '∘' in plots (a) and (b)) at random, and then downsampling so that $f = \frac{1}{0.52} \approx 1.9231Hz$ over the interval $T_{end} = 12s$. The true source parameters $\boldsymbol{\xi}_1 = (0.0594, 0.0298)m$ and $\tau_1 = 1.7800s$. We use Algorithm 4.1 and choose $K = 11$. (a) Thermal image immediately after activation, sensors '∘' and estimated source locations '×'. (b) Summary for 20 repetitions, estimated source locations ('×' left) and activation times (right).

We now consider recordings of different source set-ups. Spatiotemporal measurements are taken for different source activation times and locations; then we attempt to recover the source parameters for each data set using our method. The estimates are presented, alongside their true values, in Table 5.5, we observe that for each new experiment the parameter estimates remain close to the true values.

Table 5.4.: **MAE of centralized single source parameter estimates on real thermal data.** Field induced by point instantaneous source ($\tau = 1.7800s$ and $\boldsymbol{\xi} = (0.0594, 0.0298)m$) and its spatiotemporal samples are obtained at $f = \frac{1}{0.52} \approx 1.9231Hz$, over $T_{end} = 12s$. The MAEs are computed from 1000 independent trials. Here $K = 11$ and $r = 1$ is used for Algorithm 4.1.

	Source Parameter		
	ξ_1	ξ_2	τ
MAE	0.0036	0.0050	0.1544

Table 5.5.: **Centralized source estimation results for six independent read data set-ups.** In each experiment the recording is downsampled so $f = \frac{1}{0.52}$Hz and $f = \frac{1}{0.5}$Hz for Expts. I–II and III–VI, respectively. 13 locations are chosen randomly as sensor positions. We estimate the unknowns from samples over $T_{end} = 12s$. The true parameters and corresponding estimates obtained by Algorithm 4.1 using $K = 11$ are shown.

		Ground truth	Estimate	Absolute error
Experiment I	ξ_1	0.0697	0.0664	0.0033
	ξ_2	0.0310	0.0361	0.0051
	τ	8.7800	8.7883	0.0083
Experiment II	ξ_1	0.0594	0.0590	0.0004
	ξ_2	0.0298	0.0347	0.0049
	τ	1.7800	1.8933	0.1133
Experiment III	ξ_1	0.0600	0.0589	0.0011
	ξ_2	0.0377	0.0451	0.0074
	τ	5.3500	5.5767	0.2267
Experiment IV	ξ_1	0.0606	0.0596	0.0010
	ξ_2	0.0347	0.0323	0.0024
	τ	6.2500	6.4667	0.2167
Experiment V	ξ_1	0.0582	0.0671	0.0089
	ξ_2	0.0359	0.0371	0.0012
	τ	5.3500	5.8917	0.5417
Experiment VI	ξ_1	0.0585	0.0629	0.0044
	ξ_2	0.0365	0.0359	0.0006
	τ	4.9500	5.1833	0.2333

5.3.3. Source estimation over distributed SNs with quantized channels: real data

Figures 5.18 and 5.19 both summarize the performance of our proposed distributed source recovery method over quantized communication channels. In particular, Figure 5.18 shows a summary of the results when multiple independent experiments, with different sensor placements, are carried out. Figure 5.18(a) visualizes one realization of the SN. The scatterplot in Figure 5.18(b) shows the estimated source locations of each trial, whilst Figure 5.18(c) shows the estimated activation times. We see that estimates vary only marginally around the true values, suggesting that the proposed recovery algorithm remains robust to the sensor distribution. Furthermore, in Figure 5.19(a) the complete thermal distribution of the monitored region is shown immediately after source activation, where the epicentre of the hot (lighter) region indicates the true source location. We use measurements from the chosen sensors to estimate the source parameters and then reconstruct the field (from these estimates). For comparison the reconstructed field is shown alongside the noisy recorded field for some specific time instants in Figure 5.19.

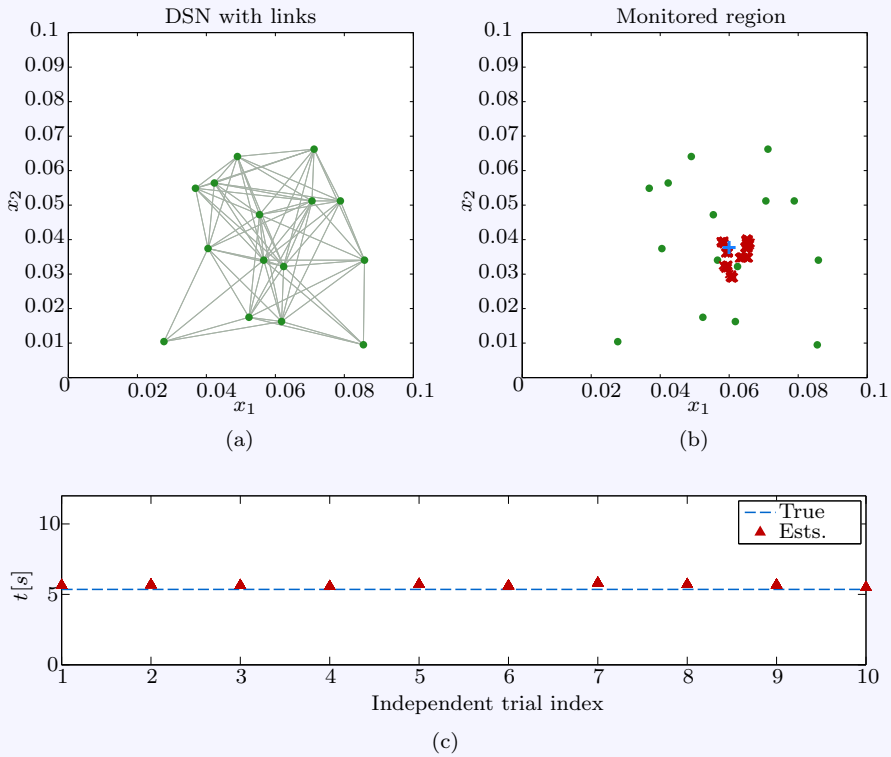
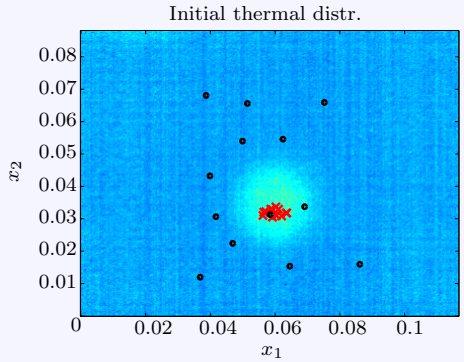
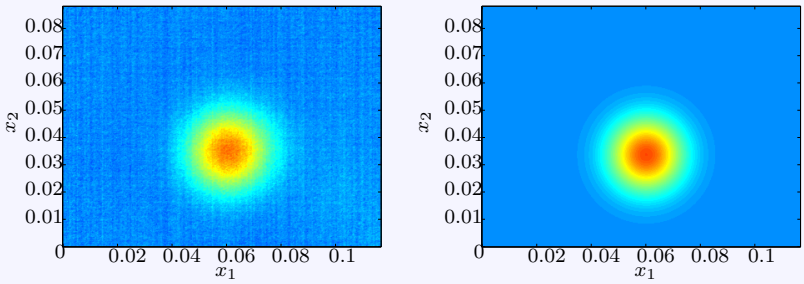


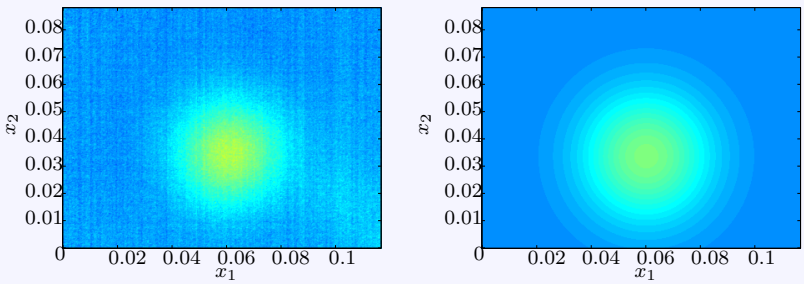
Figure 5.18.: **Distributed single source estimation with real thermal data.** The RGG, scatterplot of the locations and activation times estimated by the sensors are shown for each independent trial. The distributed estimation is performed by SNs with noisy communication links using Algorithm 4.5 under quantized gossiping, such that a 15−bit uniform quantizer with dynamic range $(-1, 1)$ is used.



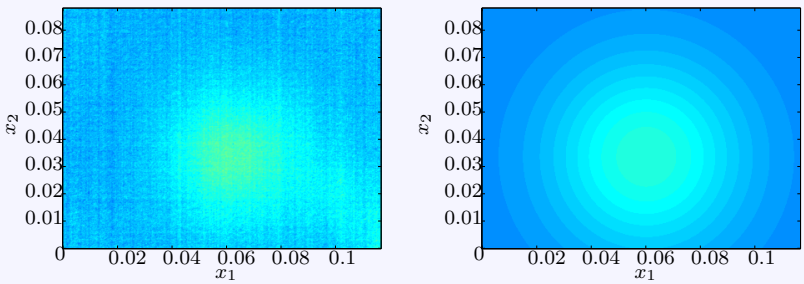
(a) Thermal distribution (just after activation) and $\boldsymbol{\xi}_1$ estimates



(b) Real field (left) and reconstruction (right) at 7.1s



(c) Real field (left) and reconstruction (right) at 8.2s



(d) Real field (left) and reconstruction (right) at 10.2s

Figure 5.19.: **Sensor distribution, location estimates and the field reconstructions.** Source is located at $(0.06058, 0.03465)$ and activated at $\tau = 6.25s$. A 15−bit uniform quantizer with dynamic range $(-1, 1)$ is used, whilst the estimation is performed with Algorithm 4.5 under quantized gossiping for the choice $K = 5$ and $r = 1$.

5.4. Summary and conclusions

The main content of this chapter was dedicated to the verification and assessment of the sensor network algorithms proposed in Chapter 4 using numerical simulations. In particular we presented numerical simulations results carried out on synthetic data to demonstrate the robustness of the proposed method even in the presence of noise and other model mismatches. Furthermore, we have also validated our algorithm using real thermal data obtained experimentally. Specifically, we have shown that we can also successfully recover the unknown source parameters given real thermal data as well as achieve an alias-free reconstruction of the entire field in space and time.

Chapter 6.

Universal framework for physics-driven inverse source problems

6.1. Introduction

In Chapter 3 we showed that given a proper sequence of generalized measurements it is possible to recover the unknown parameters for a specific class of source distributions. We will revisit this concept in the current chapter, herein however, our motive will be to devise a generalization of the the previous approach to one that is capable of solving a more "general" class of inverse source problems. Doing this, as we will see, allows us to: firstly, develop a unified framework for solving the higher dimensional inverse source problem; specifically we will be able to transition from solving just the two-dimensional problem to also solving the three-dimensional one. Secondly and more importantly, we obtain a powerful result which generalizes our previous source estimation approach to solve the inverse source problem for a wider class of PDEs—such as (6.1) and (6.5) for instance—rather than just the two-dimensional diffusion equation (3.2).

The source estimation schemes proposed in Chapter 3 can be essentially summarized in the following two steps:

1) compute the sequence of generalized measurements $\{\mathcal{Q}(k,r)\}_k$, for some fixed r by evaluating a specific family of integrals (3.17); and then,

2) apply Prony's method to $\{\mathcal{Q}(k,r)\}_k$ (or some properly modified version of it) to retrieve the unknown source parameters.

The first step, which involves evaluating integrals, can generally be interpreted as taking properly weighted sums of the sensor measurements. Exploring this new idea enables us to show that the weighted sums yield the desired generalized measurements, when these weights coincide with those that reproduce exponentials from weighted translates of a certain prototype function. As we will see this prototype function is exactly the space- and time-reversed Green's function of the PDE model for the monitored physical phenomenon. This new interpretation highlights an otherwise subtle link between inverse source problems and modern sampling theory, and so paves the way for the main contributions of the present chapter, which we organise as follows.

In Section 6.2 we give a brief account of some common linear, constant coefficient PDEs encountered in reality and state the related inverse source problem. Next, we begin Section 6.3 by arguing that in order to fully exploit the new interpretation, we will need to refine our choice of sensing functions. Hence we discuss and state explicitly our new choice of "multidimensional" spatial sensing functions for the 2-D and 3-D problems, i.e. $d = \{2, 3\}$. Additionally, we argue that these new sensing functions directly invalidate the family of integrals (3.17) and hence evaluating them no longer gives the desired generalized measurements. Consequently, inspired by certain results from modern sampling theory, we explore a new approach for computing the desired generalized measurements in Section 6.4, using proper linear combinations of the sensor data. In particular we demonstrate that taking these proper linear combinations leads to the well-known problem—of exponential reproduction using translates of a prototype function—encountered in approximation theory and in the finite rate of innovation (FRI) framework; where the prototype function, also called the *generator*, coincides with the space- and time-reversed Green's function of the physical field. Therefore leveraging from results in those domains, in particular the exact and approximate Strang-Fix conditions, we provide corresponding conditions on the Green's function under which this exponential reproduction problem can either be solved exactly or approximately. We will see that the approximate exponential reproduction framework leads to a more flexible way to analyse and solve inverse source problems relating to a wider class of PDE-driven phenomena, therefore, extending the framework of Chapter 4, which is suitable only for IDSPs. In Section 6.5, we properly adapt this new framework to solving ISPs using sensor networks, considering both uniform and nonuniform sensor placements. Moreover we develop explicit centralized and distributed estimation strategies whilst taking into account the usual limitations of typical SNs, such as noise and the impossibility of realising a spatial antialiasing filter. However, since filtering in the time domain is allowed, we show how to exploit this property, such that we may alter the underlying Green's function in a way that is more amenable to our new approach. Then in Section 6.6, we provide numerical simulation results to validate the new framework. Finally we conclude this chapter in Section 6.7 by drawing some comparisons between the framework of Chapter 4 and the present one.

6.2. Physics-driven inverse problems: Problem formulation

We use the term physics-driven to refer to phenomena, in particular physical fields, that are governed by linear PDEs with constant coefficients. Typical examples of these fields and their respective PDE models are:

1. **Poisson's equation:** is encountered frequently in many situations arising in physics such as the fields of electrostatics, (Newtonian) gravitation and so on. Mathematically, it can be written as
$$\nabla^2 u(\mathbf{x}, t) = f(\mathbf{x}, t). \tag{6.1}$$

The Green's function for this PDE in 2D (i.e. $d = 2$) is,

$$g(\mathbf{x}) = \frac{1}{2\pi} \log(\|\mathbf{x}\|)\,. \tag{6.2}$$

Whilst for $d = 3$ the Green's function becomes,

$$g(\mathbf{x}) = -\frac{1}{4\pi\|\mathbf{x}\|}. \tag{6.3}$$

2. **Diffusion equation:** as previously discussed, this PDE is of the form $\frac{\partial}{\partial t}u(\mathbf{x},t) = \mu\nabla^2 u(\mathbf{x},t)+f(\mathbf{x},t)$, with the following Green's function when $d = \{2, 3\}$:

$$g(\mathbf{x}, t) = \frac{1}{(4\pi\mu t)^{d/2}} e^{-\frac{\|\mathbf{x}\|^2}{4\mu t}} H(t), \tag{6.4}$$

where $H(t)$ is the unit step function.

3. **Wave equation:** describes many situations arising in acoustics, electromagnetism and so on. The wave equation is given by:

$$\nabla^2 u(\mathbf{x}, t) - \frac{1}{c^2}\frac{\partial^2}{\partial t^2} u(\mathbf{x}, t) = f(\mathbf{x}, t), \tag{6.5}$$

where the wave field $u(\mathbf{x}, t)$ induced by the source distribution $f(\mathbf{x}, t)$ propagates through the medium at a speed c. The Green's function for the 2-D wave equation (i.e. $d = 2$) is given by:

$$g(\mathbf{x}, t) = \frac{c}{2\pi\sqrt{c^2t^2 - \|\mathbf{x}\|^2}} H(ct - \|\mathbf{x}\|). \tag{6.6}$$

Moreover, when $\mathbf{x} \in \mathbb{R}^3$ (i.e. $d = 3$), it can be shown that:

$$g(\mathbf{x}, t) = \frac{1}{4\pi\|\mathbf{x}\|}\delta(t - \|\mathbf{x}\|/c). \tag{6.7}$$

The corresponding Green's functions stated for the examples above assume a Sommerfeld radiation condition, i.e. a *quiescent condition* at an initial time such that $u(\mathbf{x}, t)|_{t=0} = \frac{\partial}{\partial t}u(\mathbf{x}, t)\big|_{t=0} = 0$ and a *convergence condition* at infinity meaning that $u(\mathbf{x}, t)|_{\|\mathbf{x}\|\to\infty} = \frac{\partial}{\partial x_1}u(\mathbf{x}, t)\Big|_{\|\mathbf{x}\|\to\infty} = \frac{\partial}{\partial x_2}u(\mathbf{x}, t)\Big|_{\|\mathbf{x}\|\to\infty} = 0$. See, for example, [152, 62] for the derivation of these expressions.

In the present chapter, we study the discrete inverse source problems for such phenomena. In particular, given access only to spatiotemporal samples of such fields (obtained using an appropriate sensor network[1]), we will present a universal framework for recovering the unknown sources inducing the field, therefore solving the corresponding discrete ISP.

Recall that for a d-dimensional homogeneous and isotropic region $\Omega \subset \mathbb{R}^d$, we can write

[1]For instance using a suitable microphone array for audio fields.

the following expression for linear PDEs with constant coefficients:

$$u(\mathbf{x}, t) = (g * f)(\mathbf{x}, t), \tag{6.8}$$

where $g(\mathbf{x}, t)$ is the Green's function of the PDE model for the underlying phenomena such as, but not limited to, those stated above.

We can now precisely state the (inverse source) problem considered here:

Problem 6.1. *Let $\mathcal{S} = \{\mathbf{x}_n\}_{n=1}^N$ denote a network of N sensors, so that the n-th sensor situated at $\mathbf{x}_n$ collects samples $\varphi_n(t_l) = u(\mathbf{x}_n, t_l)$ of the field u, at times t_l for $l = 0, 1, \ldots, L$. Given these spatiotemporal samples, and knowledge of the Green's function of the field, we intend to estimate the unknown source distribution $f(\mathbf{x}, t)$.*

6.3. Choosing the sensing functions: source estimation from generalized measurements

Consider now the following scenario.

Example 6.1. *We wish to estimate a point source in 3-D, such that $\mathbf{x} = (x_1, x_2, x_3) \in \mathbb{R}^3$ and $\boldsymbol{\xi}_m = (\xi_{1,m}, \xi_{2,m}, \xi_{3,m}) \in \mathbb{R}^3$ with the usual parametrization, using the inner product $\langle f(\mathbf{x}, t), \Gamma_r(t)\Psi_k(\mathbf{x})\rangle$. We immediately realize that this might be problematic, since $\langle f(\mathbf{x}, t), \Gamma_r(t)\Psi_k(\mathbf{x})\rangle = \sum_m c_m e^{-\mathrm{j}r\tau_m/T} e^{-k(\xi_{1,m}+\mathrm{j}\xi_{2,m})}$ does not contain $\xi_{3,m}$; hence we stand no chance of recovering it from $\{\mathcal{Q}(k, r)\}_k$. To resolve this we may simply replace $\Psi_k(\mathbf{x})$ with $\Psi_{k_1,k_2}(\mathbf{x}) = e^{-k_1(x_1+\mathrm{j}x_2)-\mathrm{j}k_2x_3}$, so the inner product becomes $\langle f(\mathbf{x}, t), \Gamma_r(t)\Psi_{k_1,k_2}(\mathbf{x})\rangle = \sum_m c_m e^{-\mathrm{j}r\tau_m/T} e^{-k_1(\xi_{1,m}+\mathrm{j}\xi_{2,m})-\mathrm{j}k_2\xi_{3,m}}$, thus for each value of r we obtain the 2-D sequence $\{\mathcal{Q}(k_1, k_2, r)\}_{k_1,k_2}$. This new sequence is governed by a multidimensional Prony-like system and can be solved to find all the unknown parameters.*[2]

Thus the problem of inferring the unknown source parameters from the generalized measurements can be ill-posed if we do not select the (spatial) sensing functions properly. Consequently, we anticipate that our particular choice(s) will be two-dimensional with respect to its index, i.e. it will be indexed by the 2-D vector $\boldsymbol{k} = (k_1, k_2) \in \mathbb{N}^2$ just like we did in the example.

We therefore end up with the following multidimensional sequence of generalized measurements,

$$\mathcal{Q}(\boldsymbol{k}, r) = \langle f(\mathbf{x}, t), \Psi_{\boldsymbol{k}}(\mathbf{x})\Gamma_r(t)\rangle_{\mathbf{x},t}\,, \tag{6.9}$$

where $\{\Psi_{\boldsymbol{k}}(\mathbf{x})\}_{\boldsymbol{k}\in\mathbb{N}^2}$ and $\{\Gamma_r(t)\}_{r\in\mathbb{N}}$, are families of properly chosen *spatial* and *temporal sensing functions* respectively. In what follows, we first discuss how to choose the sensing functions and then we show that the family of integrals discussed in Chapter 3 are no longer

[2]For example, we can set $k_2 = 0$ and solve the resulting system using Prony's method to obtain $\{c_m, \tau_m, \xi_{1,m}, \xi_{2,m}\}_m$ and then repeat with $k_1 = 0$ to get $\{c_m, \tau_m, \xi_{3,m}\}_m$. Moreover, we may also use the ACMP method described in Section 2.2.3.

valid, for computing the generalized measurements, under these new choices of sensing functions.

Herein, the focus will predominantly be on fields induced by localized and instantaneous sources. For the reader's convenience, we restate the model describing M such sources, i.e.:

$$f(\mathbf{x}, t) = \sum_{m=1}^{M} c_m \delta(\mathbf{x} - \boldsymbol{\xi}_m, t - \tau_m), \tag{6.10}$$

where $c_m, \tau_m \in \mathbb{R}$ are the intensity and activation time of the m-th source respectively, situated at $\boldsymbol{\xi}_m = (\xi_{i,m})_{i=1}^d \in \mathbb{R}^d$. The extension to non-instantaneous point sources, (3.8), is reasonably straightforward by using $\mathcal{Q}(\boldsymbol{k}, 0)$ as described in Section 3.4.2. Although the case of non-localized sources, defined in Section 3.3.2, is not fully explored and is deferred to future works in the area.

The problem of recovering such sources now becomes one of estimating all M triples $\{c_m, \tau_m, \boldsymbol{\xi}_m\}_{m=1}^M$. We first observe that for this source distribution, the inner product (6.9) reduces to:

$$\mathcal{Q}(\boldsymbol{k}, r) = \sum_{m=1}^{M} c_m \Psi_{\boldsymbol{k}}(\boldsymbol{\xi}_m) \Gamma_r(\tau_m). \tag{6.11}$$

Hence, our current task is to select $\Psi_{\boldsymbol{k}}(\mathbf{x})$ and $\Gamma_r(t)$ so that we can recover $\{c_m, \tau_m, \boldsymbol{\xi}_m\}_{m=1}^M$ from $\{\mathcal{Q}(\boldsymbol{k}, r)\}_{\boldsymbol{k},r}$. As usual we propose the use of exponentials as this choice promotes stability and also results in an algebraically coupled power-sum series, which can be solved efficiently using multidimensional variations of Prony's method as discussed in Section 2.2.3 (see also [140, 104]). We examine explicitly the $d = 2$ and $d = 3$ cases below:

1. **Sensing in time and 2-D space:** In this case $t \in \mathbb{R}_+$ and $\mathbf{x} = (x_1, x_2) \in \mathbb{R}^2$. A valid temporal sensing function is $\Gamma_r(t) = e^{\mathrm{j}rt/T}$, where $T = t_L$ i.e. the instant at which the sensors measure the last sample of the field. Whereas for $\boldsymbol{k} \stackrel{\text{def}}{=} (k_1, k_2) \in \mathbb{Z}^2$, we choose the spatial sensing function $\Psi_{\boldsymbol{k}}(\mathbf{x}) = e^{\mathrm{j}k_1 x_1 + \mathrm{j}k_2 x_2}$. This choice turns (6.11) into:

$$\mathcal{Q}(\boldsymbol{k}, r) = \sum_{m=1}^{M} c_m e^{\mathrm{j}r\tau_m/T} e^{\mathrm{j}k_1\xi_{1,m} + \mathrm{j}k_2\xi_{2,m}}. \tag{6.12}$$

   Notice that if we allow k_1 to be imaginary and impose, $k_1 = -\mathrm{j}k_2 = \mathrm{j}k$ we obtain the usual $\mathcal{Q}(k, r)$, so this new sensing function is a generalization.

   In fact, when we use this particular *sensing function* in 2-D, then the results obtained for restoring the sum of exponentials structure of $\mathcal{Q}(k, r)$ for our non-localized source models (see Section 3.4.3) still hold in this new framework. This is not true however, for general 2-D sensing function choices.

2. **Sensing in time and 3-D space:** In this case $\mathbf{x} = (x_1, x_2, x_3) \in \mathbb{R}^3$, $\Gamma_r(t) = e^{\mathrm{j}rt/T}$

as before, but $\Psi_{\boldsymbol{k}}(\mathbf{x}) = e^{k_1(x_1+\mathrm{j}x_2)+\mathrm{j}k_2x_3}$. Given this choice,

$$\mathcal{Q}(\boldsymbol{k}, r) = \sum_{m=1}^{M} c_m e^{\mathrm{j}r\tau_m/T} e^{k_1(\xi_{1,m}+\mathrm{j}\xi_{2,m})+\mathrm{j}k_2\xi_{3,m}}. \tag{6.13}$$

We remark that the following: $\Psi_{\boldsymbol{k}}(\mathbf{x})=e^{\mathrm{j}k_1x_1+k_2(x_3+\mathrm{j}x_2)}$, $\Psi_{\boldsymbol{k}}(\mathbf{x})=e^{k_1(\xi_{1,m}+\mathrm{j}\xi_{3,m})+\mathrm{j}k_2\xi_{2,m}}$ and $\Psi_{\boldsymbol{k}}(\mathbf{x})=e^{k_1(x_1+\mathrm{j}x_2)+k_2x_3}$ would also be valid choices.

Notice now that, for some fixed $r \neq 0$, in particular $r = 1$, (6.12) and (6.13) are of the form:

$$\mathcal{Q}(\boldsymbol{k}, 1) \stackrel{\text{def}}{=} \mathcal{Q}(k_1, k_2, 1) = \sum_{m=1}^{M} b_m u_m^{k_1} v_m^{k_2},$$

where in the case of (6.12) $b_m = c_m e^{\mathrm{j}\tau_m/T}$, $u_m = e^{\mathrm{j}\xi_{1,m}}$ and $v_m = e^{\mathrm{j}\xi_{2,m}}$, whilst for (6.13) $b_m = c_m e^{\mathrm{j}\tau_m/T}$, but now $u_m = e^{(\xi_{1,m}+\mathrm{j}\xi_{2,m})}$ and $v_m = e^{\mathrm{j}\xi_{3,m}}$. Both systems are therefore coupled (multi-dimensional) Prony-like systems, which can be solved using Algebraically Coupled Matrix Pencil (ACMP) method (see Section 2.2.2 or the references [104, 149]) to find jointly $\{c_m, \tau_m, \boldsymbol{\xi}_m\}_{m=1}^{M}$ from $\{\mathcal{Q}(\boldsymbol{k}, 1)\}_{\boldsymbol{k}}$ with $k_1 =, 1, \ldots, K_1$ and $k_2 =, 1, \ldots, K_2$, when $K_1, K_2 \geq 2M$.

Having seen how we may estimate the unknown point source parameters from the generalized measurements, we can now focus on computing these generalized measurements from the sensor data. In Chapter 3, it was shown that these sequence of generalized measurements are governed by a corresponding family of definite integrals (3.17) (see Proposition 3.1). Therein, the proof of the proposition relied on the analyticity of the spatial sensing function, specifically we required that $\nabla^2\Psi_k(\mathbf{x}) = 0$. Unfortunately for these new and more general choice of spatial sensing functions the above analyticity property is no longer valid, i.e.

$$\nabla^2\Psi_{\boldsymbol{k}}(\mathbf{x}) \neq 0.$$

Therefore we cannot compute $\mathcal{Q}(\boldsymbol{k}, r)$ from (3.17) by simply replacing the old sensing function with the new ones. Consequently we present a new approach for computing these generalized measurements from the sensor data in the next section.

6.4. Multidimensional generalized measurements from sensor data

Although the family of definite integrals (derived in Chapter 3) and their quadrature-based approximations (in Chapter 4) no longer yield the correct generalized measurements, in this new multidimensional setting, they still highlight a very useful and important consideration. Specifically, the integrals themselves and their approximations, are linear in the field measurements; and so evaluating them can be seen simply as computing a proper linear combination of the field measurements. Therefore we may take advantage of this result to

write the following summation,

$$\sum_{n=1}^{N}\sum_{l=0}^{L} w_{n,l}(\boldsymbol{k},r)\varphi_n(t_l) = \mathcal{Q}(\boldsymbol{k},r), \tag{6.14}$$

where the goal is now to find the correct weights $\{w_{n,l}(\boldsymbol{k},r)\}_{n,l}$, such that our inner product structure (6.9) is still preserved, for each $\boldsymbol{k} \in \mathbb{N}^2$ and $r \in \mathbb{N}$. The following results from this consideration.

Proposition 6.1. *Computing the multidimensional sequence of generalized measurements $\{\mathcal{Q}(\boldsymbol{k},r)\}_{\boldsymbol{k}}$, in (6.9) for each $r \in \mathbb{N}$, using weighted linear combinations (6.14) of the sensor data $\varphi_n(t_l)$ is equivalent to reproducing the function $\Psi_{\boldsymbol{k}}(\mathbf{x})\Gamma_r(t)$ from space- and time-reversed translates of the Green's function $g(\mathbf{x},t)$ of the underlying field. Furthermore when $\Psi_{\boldsymbol{k}}(\mathbf{x})$ and $\Gamma_r(t)$ are chosen to be exponentials, this results in a multidimensional exponential reproduction problem.*

Proof. We commence this proof by noting that the sensor measurements, $u(\mathbf{x}_n,t_l)$, can be re-written, using (6.8), as follows:

$$\begin{aligned} u(\mathbf{x},t) &= f(\mathbf{x},t) * g(\mathbf{x},t) \\ &= \int_{\mathbf{x}'\in\mathbb{R}^d}\int_{t'\in\mathbb{R}} g(\mathbf{x}',t')f(\mathbf{x}-\mathbf{x}',t-t')\,\mathrm{d}t'\mathrm{d}\mathbf{x}' \\ &= \langle f(\mathbf{x}',t'), g(\mathbf{x}-\mathbf{x}',t-t')\rangle_{\mathbf{x}',t'}. \end{aligned}$$

where $\mathbf{x}' = (x_1', x_2', \ldots, x_d') \in \mathbb{R}^d$ and $\mathrm{d}\mathbf{x}' = \prod_{i=1}^{d}\mathrm{d}x_i'$. Consequently, the discrete measurement obtained by the n-th sensor (situated at $\mathbf{x}_n$) at some time instant $t_l \geq 0$ is

$$\varphi_n(t_l) = u(\mathbf{x}_n,t_l) = \langle f(\mathbf{x},t), g(\mathbf{x}_n-\mathbf{x},t_l-t)\rangle_{\mathbf{x},t}. \tag{6.15}$$

Substituting the expression (6.15) into the weighted sum (6.14) of the spatiotemporal sensor measurements, $\{\varphi_n(t_l)\}_{n,l}$, produces the expression:

$$\begin{aligned} \sum_{n\in\mathcal{N}}\sum_{l=0}^{L} w_{n,l}(\boldsymbol{k},r)\varphi_n(t_l) &= \sum_{n\in\mathcal{N}}\sum_{l=0}^{L} w_{n,l}(\boldsymbol{k},r)u(\mathbf{x}_n,t_l) \\ &= \sum_{n\in\mathcal{N}}\sum_{l=0}^{L} w_{n,l}(\boldsymbol{k},r)\,\langle f(\mathbf{x},t), g(\mathbf{x}_n-\mathbf{x},t_l-t)\rangle_{\mathbf{x},t} \\ &= \left\langle f(\mathbf{x},t), \sum_{n\in\mathcal{N}}\sum_{l=0}^{L} w_{n,l}(\boldsymbol{k},r)g(\mathbf{x}_n-\mathbf{x},t_l-t)\right\rangle_{\mathbf{x},t}, \end{aligned} \tag{6.16}$$

where $\{w_{n,l}(\boldsymbol{k},r)\}_{n,l} \in \mathbb{C}$ are the specific sequence of weights we wish to compute[3]. In

[3]In the last equality, we are able to pass the summation inside the inner product because it is finite. If it were infinite then we would require that the sum converges absolutely; which is ensured if g and its translates form a Riesz basis.

particular if we require this weighted sum of the sensor data to yield the exact multidimensional measurements, $\sum_{n\in\mathcal{N}}\sum_{l=0}^{L} w_{n,l}(\boldsymbol{k},r)\varphi_n(t_l) = \mathcal{Q}(\boldsymbol{k},r) = \langle f(\mathbf{x},t), \Psi_{\boldsymbol{k}}(\mathbf{x})\Gamma_r(t)\rangle_{\mathbf{x},t}$, then by comparing the inner products in (6.9) and (6.16), we realize that we must choose the sequence of weights $\{w_{n,l}(\boldsymbol{k},r)\}_{n,l}$ such that, for each $\boldsymbol{k}$ and r, the identity

$$\sum_{n\in\mathcal{N}}\sum_{l=0}^{L} w_{n,l}(\boldsymbol{k},r)g(\mathbf{x}_n-\mathbf{x},t_l-t) \equiv \Psi_{\boldsymbol{k}}(\mathbf{x})\Gamma_r(t) \tag{6.17}$$

is satisfied. This proves the first claim of the proposition.

The second statement follows immediately. If we impose the choice of sensing functions of Section 6.3, where: for $d=2$, $\Psi_{\boldsymbol{k}}(\mathbf{x}) = e^{\mathrm{j}k_1x_1+\mathrm{j}k_2x_2}$ and $\Gamma_r(t) = e^{\mathrm{j}rt/T}$, (6.17) reduces to

$$\sum_{n\in\mathcal{N}}\sum_{l=0}^{L} w_{n,l}(\boldsymbol{k},r)g(\mathbf{x}_n-\mathbf{x},t_l-t) = e^{\mathrm{j}k_1x_1+\mathrm{j}k_2x_2}e^{\mathrm{j}rt/T}; \tag{6.18}$$

whilst, for $d=3$, $\Psi_{\boldsymbol{k}}(\mathbf{x}) = e^{k_1(x_1+\mathrm{j}x_2)+\mathrm{j}k_2x_3}$ and $\Gamma_r(t) = e^{\mathrm{j}rt/T}$, (6.17) reduces to

$$\sum_{n\in\mathcal{N}}\sum_{l=0}^{L} w_{n,l}(\boldsymbol{k},r)g(\mathbf{x}_n-\mathbf{x},t_l-t) = e^{k_1(x_1+\mathrm{j}x_2)+\mathrm{j}k_2x_3}e^{\mathrm{j}rt/T}. \tag{6.19}$$

A closer inspection of (6.18) and (6.19) above, allows us to rewrite the 2-D and 3-D problems in a more compact vector form as follows:

$$\sum_{n,l} w_{n,l}(\boldsymbol{k},r)g(\mathbf{x}_n-\mathbf{x},t_l-t) = e^{\boldsymbol{\kappa}\cdot\mathbf{x}+\rho t}, \tag{6.20}$$

where $\boldsymbol{\kappa} = \mathrm{j}\boldsymbol{k}$ or $\boldsymbol{\kappa} = (k_1, \mathrm{j}k_1, \mathrm{j}k_2)$ respectively and $\rho = \mathrm{j}r/T$. In both cases, therefore, we observe that the required coefficients are those that reproduce the $(d+1)$-dimensional (space and time varying) exponentials by summing translates of $g(-\mathbf{x},-t)$ – the Green's function of the underlying field $u(\mathbf{x},t)$. This completes the proof. □

Thus if we had access to the precise coefficients $\{w_{n,l}(\boldsymbol{k},r)\}_{n,l}$, capable of reproducing exponentials from the translates of $g(\mathbf{x},t)$, all we would need to do is evaluate the sequence $\{\mathcal{Q}(\boldsymbol{k},r)\}_{\boldsymbol{k}}$ for a fixed $r\neq 0$ using (6.14) and from it extract the unknown source parameters as described in Section 6.3. Therefore the missing piece in our framework is how to obtain the exponential reproducing coefficients. For this missing piece, it is important to understand if the approximation is possible and how accurate it is, before deriving algorithms to find it. We devote the remainder of this section to these three points.

First, we will leverage from results on generalized sampling and approximation theory to outline the conditions for which an exact or an approximate reproduction is achievable. Second, for both (exact and approximate) scenarios, we derive closed-form formulae for computing the "best" weights $w_{n,l}(\boldsymbol{k},r)$ when the translates of the approximant are assumed to be on a regular lattice. In the sensor network setup this is equivalent to having uniform

spatiotemporal sampling with positive sampling intervals $\mathbf{\Delta_x} = (\Delta_{x_1}, \Delta_{x_2}, \ldots, \Delta_{x_d})$ and Δ_t. Finally in the nonuniform sampling case, where it is generally not possible to obtain simple closed expressions for the desired exponential reproducing coefficients $w_{n,l}(\boldsymbol{k}, r)$, we propose two approaches:

1. Formulating and solving the linear system that comes from discretizing (6.20).
2. Interpolating and then resampling the sensor data uniformly, so as to permit the use of the formulae derived in the case of uniform shifts.

6.4.1. Function spaces, generalized sampling and function approximation

Recall that, in the generalized (uniform) sampling paradigm—see Section 2.2 or [138, 9, 8, 143, 151] and references therein— the primary goal is to reconstruct some function of a continuous variable from a discrete set of measurements collected on a (uniform) set of grid points. Often, this reconstruction will be an approximation of the original signal in some signal/function spaces. Furthermore a desirable property of this approximation is such that, in the limit, as the "density" of the sampling grid points increases the approximation error should vanish, at least for a well-behaved function.

Precisely if a d-dimensional generating function s whose uniform translates yields the vector space $V_{\mathbf{\Delta_x}}(s) = \text{span}_{\boldsymbol{n}\in\mathbb{Z}^d}\{s(\mathbf{x}/\mathbf{\Delta_x} - \boldsymbol{n})\}$, then any function $\hat{h}(\mathbf{x})$ necessarily in the space $V_{\Delta_\mathbf{x}}(s) \subset \mathcal{L}^2$ is characterized by the sequence of coefficients $a_{\boldsymbol{n}}$, such that:

$$\hat{h}(\mathbf{x}) = \sum_{\boldsymbol{n}\in\mathbb{Z}^d} a_{\boldsymbol{n}} s(\mathbf{x}/\mathbf{\Delta_x} - \boldsymbol{n}). \tag{6.21}$$

In fact for this series to make sense, that is if we want to have this continuous-discrete representation[4], then:

(i) For convergence, $\{a_{\boldsymbol{n}}\}_{\boldsymbol{n}}$ must be square-summable.

(ii) For uniqueness and stability of this discrete representation, so that $V_{\mathbf{\Delta_x}}$ is a well defined subspace of $\mathcal{L}^2$, $\{s(\mathbf{x} - \boldsymbol{n})\}_{\boldsymbol{n}}$ must form a Riesz basis of, $V_{\mathbf{1}}(g)$, the space formed from integer shifts (i.e. $\Delta_{x_i} = 1$ for any $i = 1, \ldots, d$).

(iii) Finally, $s(\mathbf{x})$ must satisfy the partition of unity condition

$$\sum_{\boldsymbol{n}\in\mathbb{Z}} s(\mathbf{x} + \boldsymbol{n}) = 1, \tag{6.22}$$

for all $\mathbf{x} \in \mathbb{R}^d$; which guarantees that by choosing $\mathbf{\Delta_x}$ in (6.21) sufficiently small, we can approximate any function $\hat{h}(\mathbf{x})$ as closely as we want. A detailed proof of this fact can be found in [143, Appendix B].

[4] A representation that gives a useful correspondence, for the purposes of signal processing, between $s(\mathbf{x})$ a function over the continuous variable $\mathbf{x}$ and $a_{\boldsymbol{n}}$ defined over a discrete set $\boldsymbol{n}$.

If we instead want to reconstruct some signal $h(\mathbf{x}) \in \mathcal{L}^2 \backslash V_{\boldsymbol{\Delta}_\mathbf{x}}$, then the reconstruction $\hat{h}(\mathbf{x})$ as in (6.21) should give the best approximation of $h(\mathbf{x})$ in the space $V_{\boldsymbol{\Delta}_\mathbf{x}}$. In that, it is the approximation that minimizes the approximation error in the least-squares sense. Then we must pick the coefficients $a_{\boldsymbol{n}} = \langle h(\mathbf{x}), \tilde{s}\left(\mathbf{x}/\boldsymbol{\Delta}_\mathbf{x} - \boldsymbol{n}\right)\rangle$, where $\tilde{s}$ is the dual of s given by [25],

$$\hat{\tilde{s}}(\boldsymbol{\omega}) = \frac{\hat{s}(\boldsymbol{\omega})}{\sum_{\boldsymbol{n}} |\hat{s}(\boldsymbol{\omega} + 2\pi\boldsymbol{n})|^2},$$

and $\hat{s}(\boldsymbol{\omega}) = \mathcal{F}_\mathbf{x}\{s\} = \int_\mathbf{x} s(\mathbf{x}) e^{-\mathrm{j}\boldsymbol{\omega}\cdot\mathbf{x}} \mathrm{d}\mathbf{x}$ denotes the multidimensional Fourier transform of s. Please also note that in the rest of this thesis, we assume that all transforms are taken in the sense of distributions.

In the sensor network setting we are seeking the specific coefficients $\{w_{n,l}(\boldsymbol{k}, r)\}_{n,l}$ that reproduce the exponential function using shifted versions of the Green's function of the underlying physical field. Hence this is a special case of the above, where: (i) the signal we want to reconstruct is a specific $(d+1)$-dimensional exponential (i.e. $e^{\boldsymbol{\kappa}\cdot\mathbf{x}+\rho t}$), and (ii) the sampling kernel/acquisition device is fixed and dictated by the underlying PDE. In fact as we have shown in Proposition 6.1, it is precisely the Green's function of the PDE. Under these conditions we want to find the best discrete representation, of the exponentials, in the space spanned by the translates of the Green's functions. This scenario shares some similarities with problems which have been previously studied in approximation theory. Therefore we can borrow from tools found therein, at least in the case of uniform shifts.

6.4.2. Exact and approximate Strang-Fix theory for exponential reproduction from uniform translates

We summarize the specific results from Strang-Fix theory that will be of interest to us in the lemma stated after these definitions.

Definition 6.1 ($\mathcal{H}_C^p$ space)**.** *We denote by $\mathcal{H}_C^p$, the space of compactly supported functions whose derivatives up to (and including) the order $p \geq 0$ are in $\mathcal{L}^2$ (i.e. square-integrable).*

Definition 6.2 (Bilateral Laplace transform)**.** *The multidimensional bilateral Laplace transform $G(\mathbf{s}_\mathbf{x}, s_t)$ of a signal $g(\mathbf{x}, t)$ is given by,*

$$G(\mathbf{s}_\mathbf{x}, s_t) = \int_{t\in\mathbb{R}} \int_{\mathbf{x}\in\mathbb{R}^d} g(\mathbf{x}, t) e^{-(\mathbf{x},t)\cdot(\mathbf{s}_\mathbf{x}, s_t)} \mathrm{d}\mathbf{x}\mathrm{d}t. \tag{6.23}$$

Lemma 6.1 (Generalized Strang-Fix conditions [83])**.** *If a function $s(\mathbf{x}) \in \mathcal{H}_C^p$ then the following conditions are equivalent:*

1. *For any $|\alpha| \leq p$,*
$$S(\boldsymbol{\kappa}) \neq 0, \text{ whilst } S(\boldsymbol{\kappa} + \mathrm{j}2\pi\boldsymbol{n}) = 0, \tag{6.24}$$
if $\boldsymbol{n} \in \mathbb{Z}^d \setminus \{\mathbf{0}\}$, where $\mathbf{0} = (0, 0, \ldots, 0)$ is the zero d-vector..

2. *Exponential reproducing: for some coefficients* $a_{\boldsymbol{n}} \in \mathbb{C}$,

$$\sum_{\boldsymbol{n}\in\mathbb{Z}^d} a_{\boldsymbol{n}} s(\mathbf{x}-\boldsymbol{n}), \tag{6.25}$$

is an exponential in $\mathbf{x}$, *i.e.* $\sum_{\boldsymbol{n}\in\mathbb{Z}^d} a_{\boldsymbol{n}} s(\mathbf{x}-\boldsymbol{n}) = Ce^{\boldsymbol{\kappa}\cdot\mathbf{x}}$, *for some* $C \neq 0$.

Proof. See Section B.1 for a proof which is a modification of that given in [141] by extending the argument of [145] to multiple dimensions. □

The assumption of compactness here ensures that the bilateral Laplace transform exists, and so the conditions (6.24) are well-defined. It also ensures that (6.25) converges. However, it has been demonstrated that this assumption is sufficient, but not necessary for the equivalence in Lemma 6.1 to hold. Significant efforts have been made to extend the Strang-Fix results for noncompactly supported functions, for example, the authors of [93] replace this restriction with one that endows $s(\mathbf{x})$ with a suitable polynomial decay, whilst this was further relaxed in [53, 24] favouring even milder restrictions on the approximant so that the Strang-Fix result still holds. In fact current restrictions on the decay of the approximant are such that most functions are admissible, and for these the desired coefficients can be computed exactly. To treat our multidimensional problem, we now extend formally the one-dimensional formulae obtained in [145] to the multidimensional case, which is still valid even when the approximant is not separable with respect to its variables – i.e we do not require $s(\mathbf{x}) = \prod_{i=1}^d s_i(x_i)$. The absence of this separability property is of paramount importance for us, especially because the spatial and temporal dimensions for most non-static fields encountered, in reality, are neither separable nor homogeneous. We begin our derivation by recalling that we are after the coefficients $\{a_{\boldsymbol{n}}\}_{\boldsymbol{n}}$ that minimizes the approximation error in the least squares sense, such that

$$\sum_{\boldsymbol{n}\in\mathbb{Z}^d} a_{\boldsymbol{n}} s(\mathbf{x}-\boldsymbol{n}) = e^{\boldsymbol{\kappa}\cdot\mathbf{x}}. \tag{6.26}$$

According to generalized sampling theory, the sequence of weights that minimizes the approximation error in the least-squares sense is exactly that generated through,

$$\begin{aligned} a_{\boldsymbol{n}} &= \langle e^{\boldsymbol{\kappa}\cdot\mathbf{x}}, \tilde{s}(\mathbf{x}-\boldsymbol{n})\rangle_{\mathbf{x}} = \int_{\mathbf{x}\in\mathbb{R}^d} e^{\boldsymbol{\kappa}\cdot\mathbf{x}}\tilde{s}(\mathbf{x}-\boldsymbol{n})\mathrm{d}\mathbf{x} \\ &= \int_{\mathbf{x}'\in\mathbb{R}^d} e^{\boldsymbol{\kappa}\cdot(\mathbf{x}'+\boldsymbol{n})}\tilde{s}(\mathbf{x}')\mathrm{d}\mathbf{x}' = e^{\boldsymbol{\kappa}\cdot\boldsymbol{n}} \int_{\mathbf{x}'\in\mathbb{R}^d} e^{\boldsymbol{\kappa}\cdot\mathbf{x}'}\tilde{s}(\mathbf{x}')\mathrm{d}\mathbf{x}' \\ &= e^{\boldsymbol{\kappa}\cdot\boldsymbol{n}} a_{\mathbf{0}}, \end{aligned} \tag{6.27}$$

where the second line follows from the change of variable $\mathbf{x}' = \mathbf{x} - \boldsymbol{n}$. Thus finding $a_{\mathbf{0}}$ allows us to compute $a_{\boldsymbol{n}}$ for all $\boldsymbol{n} \in \mathbb{Z}^d$ using (6.27). To find $a_{\mathbf{0}}$ substitute (6.27) into (6.26) to get

$$a_{\mathbf{0}} \sum_{\boldsymbol{n}\in\mathbb{Z}^d} e^{\boldsymbol{\kappa}\cdot\boldsymbol{n}} s(\mathbf{x}-\boldsymbol{n}) = e^{\boldsymbol{\kappa}\cdot\mathbf{x}} \Leftrightarrow a_{\mathbf{0}} \sum_{\boldsymbol{n}\in\mathbb{Z}^d} e^{-\boldsymbol{\kappa}\cdot(\mathbf{x}-\boldsymbol{n})} s(\mathbf{x}-\boldsymbol{n}) = 1.$$

We apply Poisson summation formula on the lattice to the l.h.s. of this expression, which if $s(\mathbf{x})$ is well-behaved reduces to

$$a_{\mathbf{0}} \sum_{\boldsymbol{n}\in\mathbb{Z}^d} S(\boldsymbol{\kappa} + \mathrm{j}2\pi\boldsymbol{n})e^{\mathrm{j}2\pi\boldsymbol{n}\cdot\mathbf{x}} = 1.$$

Finally from the Strang-Fix condition (6.24) we get $a_{\mathbf{0}} = 1/S(\boldsymbol{\kappa})$, since all terms in the summation vanish for $\boldsymbol{n} \neq \mathbf{0}$. Thus for any $\boldsymbol{n} \in \mathbb{Z}^d$,

$$a_{\boldsymbol{n}} = \frac{e^{\boldsymbol{\kappa}\cdot\boldsymbol{n}}}{S(\boldsymbol{\kappa})}. \tag{6.28}$$

Approximate Strang-Fix in multidimensions and the approximation error

Note that in the derivation of (6.28) one needs to impose some regularity conditions on $s(\mathbf{x})$, especially for the l.h.s of the Poisson summation formula to make sense the function must decay sufficiently quickly. However the strongest constraint on $s(\mathbf{x})$ is due to (6.24), in particular the requirement that $S(\boldsymbol{\kappa} + \mathrm{j}2\pi\boldsymbol{n}) = 0$ when $\boldsymbol{n} \in \mathbb{Z}^d \setminus \{\mathbf{0}\}$.

For general physical fields of interest to us, the approximant s will be replaced by the corresponding Green's function g of the field. Whilst these will generally not satisfy the Strang-Fix condition (6.24), we still wish to approximately reproduce exponentials with them. Fortunately, however we can apply the so called approximate Strang-Fix method introduced in [145], which relaxes the assumptions on the generators (for the 1-D exponential case), such that we are now after the best set of coefficients that leads to approximate exponential reproduction, given any kernel s. Mathematically this means that we desire

$$\sum_{\boldsymbol{n}\in\mathbb{Z}^d} a_{\boldsymbol{n}} s(\mathbf{x} - \boldsymbol{n}) \approx e^{\boldsymbol{\kappa}\cdot\mathbf{x}}, \tag{6.29}$$

where s does not necessarily satisfy the generalized Strang-Fix conditions (6.24). There are a few possible choices one may make for the "best" approximation coefficients. For example, we may opt for the least-squares coefficients which ensures that the approximation error is orthogonal to the space $V_{\mathbf{1}}$, or the interpolation coefficients which interpolates the exponential at the exact locations $\mathbf{x} = \boldsymbol{\kappa}$ (see [145] for more details). However, in this work we will focus on the constant least squares coefficients of the form:

$$a_{\boldsymbol{n}}(\boldsymbol{\kappa}) = \frac{e^{\boldsymbol{\kappa}\cdot\boldsymbol{n}}}{S(\boldsymbol{\kappa})}, \tag{6.30}$$

for their simplicity and accuracy, in particular they provide a good approximation when the Fourier transform of the generator S decays quickly. Notice also that these coefficients also coincide with (6.28), since the error $\varepsilon(\mathbf{x}) = e^{\boldsymbol{\kappa}\cdot\mathbf{x}}\left(1 - a_{\mathbf{0}}\sum_{\boldsymbol{n}} S(\boldsymbol{\kappa} + \mathrm{j}2\pi\boldsymbol{n})e^{\mathrm{j}2\pi\boldsymbol{n}\cdot\mathbf{x}}\right)$ is minimized by this set of coefficients by requiring that: (i) the error $\varepsilon(\mathbf{x})$ be zero, meaning the term $1 - a_{\mathbf{0}}\sum_{\boldsymbol{n}} S(\boldsymbol{\kappa} + \mathrm{j}2\pi\boldsymbol{n})e^{\mathrm{j}2\pi\boldsymbol{n}\cdot\mathbf{x}} = 0$; and, (ii) that $S(\boldsymbol{\kappa} + \mathrm{j}2\pi\boldsymbol{n})$ decays quickly with $\boldsymbol{n}$.

Therefore for this choice of coefficients the approximation error is:

$$\varepsilon(\mathbf{x}) = e^{\boldsymbol{\kappa}\cdot\mathbf{x}}\left(1 - \frac{1}{S(\boldsymbol{\kappa})}\sum_{\boldsymbol{n}} S(\boldsymbol{\kappa} + \mathrm{j}2\pi\boldsymbol{n})e^{\mathrm{j}2\pi\boldsymbol{n}\cdot\mathbf{x}}\right), \tag{6.31}$$

which will be small if $S(\boldsymbol{\kappa} + \mathrm{j}2\pi\boldsymbol{n})$ decays quickly to zero as $|\boldsymbol{n}|$ increases.

6.4.3. The coefficients for the space-time fields

Uniform sampling

In this uniform sampling setup, we will denote our field samples as $\varphi_{\boldsymbol{n}}(t_l) = u(\boldsymbol{n}\boldsymbol{\Delta}_{\mathbf{x}}, l\Delta_t)$, where $\boldsymbol{n}\boldsymbol{\Delta}_{\mathbf{x}} = (n_1\Delta_{x_1}, n_2\Delta_{x_2}, \ldots, n_d\Delta_{x_d})$ and $n_i = 0, 1, \ldots, N_i - 1$ for $i = 1, \ldots, d$. Notice that we can reconcile the vector index sensor measurement $\varphi_{\boldsymbol{n}}(t_l)$ with the scalar indexed one, $\varphi_n(t_l)$, by simply taking the lexicographic ordering of the elements of $\{\boldsymbol{n}\Delta_{\mathbf{x}}\}_{\boldsymbol{n}\in\mathbb{N}^d}$ to give $\{\mathbf{x}_n\}_{n=1}^N$, where $N = \prod_{i=1}^d N_i$.

Moreover, for our physical fields with linear and constant coefficients we can see from (6.20) that the generator is the space- and time-reversed Green's function $s = g(-\mathbf{x}, -t)$ which has the bilateral Laplace transform $G(-\mathbf{s}_{\mathbf{x}}, -s_t)$. Therefore it follows that the desired coefficients (under integer translates i.e. $\boldsymbol{\Delta}_{\mathbf{x}} = \mathbf{1}$ and $\Delta_t = 1$) are given by

$$w_{\boldsymbol{n},l}(\boldsymbol{k}, r) = \frac{e^{(\boldsymbol{\kappa},\rho)\cdot(\boldsymbol{n},l)}}{G(-\boldsymbol{\kappa}, -\rho)}, \tag{6.32}$$

where $\rho = \mathrm{j}r/T$, whilst for $d = 2$ and $d = 3$, $\boldsymbol{\kappa} = \mathrm{j}(k_1, k_2)$ and $\boldsymbol{\kappa} = (k_1, \mathrm{j}k_1, \mathrm{j}k_2)$, respectively. The expression in (6.32) follows since $G(-\mathbf{s}_{\mathbf{x}}, -s_t)|_{(\mathbf{s}_{\mathbf{x}},s_t)=(\boldsymbol{\kappa},\rho)} = G(-\boldsymbol{\kappa}, -\rho)$, written in this way to highlight the dependence on $\boldsymbol{k}$ and r. Using these coefficients, we can obtain the desired multidimensional generalized measurements by substituting (6.32) into (6.14) for integer translates.

Furthermore when the translates take on non-integer values, i.e. if $(\boldsymbol{n}\boldsymbol{\Delta}_{\mathbf{x}}, \Delta_t) \neq \mathbf{1}$, so that the exponential reproduction problem is $\sum_{\boldsymbol{n}}\sum_l w_{\boldsymbol{n},l}(\boldsymbol{k}, r)g(\boldsymbol{n}\boldsymbol{\Delta}_{\mathbf{x}} - \mathbf{x}, l\Delta_t - t) \approx e^{(\boldsymbol{\kappa},\rho)\cdot(\mathbf{x},t)}$, then this problem is exactly the same as:

$$\sum_{\boldsymbol{n}}\sum_l w_{\boldsymbol{n},l}(\boldsymbol{k}, r)\bar{g}(\boldsymbol{n} - \frac{\mathbf{x}}{\boldsymbol{\Delta}_{\mathbf{x}}}, l - \frac{t}{\Delta_t}) \approx e^{(\bar{\boldsymbol{\kappa}},\bar{\rho})\cdot(\frac{\mathbf{x}}{\boldsymbol{\Delta}_{\mathbf{x}}},\frac{t}{\Delta_t})}, \tag{6.33}$$

where $(\bar{\boldsymbol{\kappa}}, \bar{\rho}) = (\boldsymbol{\Delta}_{\mathbf{x}}\boldsymbol{\kappa}, \Delta_t\rho)$ and $\bar{g}(\mathbf{x}, t) = g(\boldsymbol{\Delta}_{\mathbf{x}}\mathbf{x}, \Delta_t t)$. Consequently the problem (6.33) has the corresponding exponential reproducing coefficients,

$$\bar{w}_{\boldsymbol{n},l}(\boldsymbol{k}, r) = \frac{e^{(\bar{\boldsymbol{\kappa}},\bar{\rho})\cdot(\boldsymbol{n},l)}}{\bar{G}(-\bar{\boldsymbol{\kappa}}, -\bar{\rho})}. \tag{6.34}$$

From this we can obtain the proper coefficients for any Green's function g, and correspondingly any physics-driven phenomenon. In particular, using the fact that $\bar{g}(\mathbf{x}, t) = g(\boldsymbol{\Delta}_{\mathbf{x}}\mathbf{x}, \Delta_t t) \Leftrightarrow \bar{G}(\mathbf{s}_{\mathbf{x}}, s_t) = \frac{1}{\Delta_t\prod_{i=1}^d \Delta_{x_i}}G(\frac{\mathbf{s}_{\mathbf{x}}}{\boldsymbol{\Delta}_{\mathbf{x}}}, \frac{s_t}{\Delta_t})$, we can conclude that for any $\boldsymbol{\Delta}_{\mathbf{x}} \in$

$\mathbb{R}^d_{>0}, \Delta_t \in \mathbb{R}_{>0}$ then

$$w_{\boldsymbol{n},l}(\boldsymbol{k}, r) = \Delta_t \prod_{i=1}^{d} \Delta_{x_i} \frac{e^{(\boldsymbol{\Delta_x \kappa}, \Delta_t \rho)\cdot(\boldsymbol{n}, l)}}{G(-\boldsymbol{\kappa}, -\rho)}, \tag{6.35}$$

by substituting $\bar{G}(-\bar{\boldsymbol{\kappa}}, -\bar{\rho})$ into (6.34). We show, in Figure 6.1, the approximation of our 2-D spatial exponentials, i.e. $e^{\mathrm{j}k_1x_1+\mathrm{j}k_2x_2}$, using the 2-D Green's function of Poisson's equation (6.2) and the coefficients (6.35), for $r = 0$, $k_1 = 2$, and $k_2 = 0, 1, 2$.

Nonuniform sampling: linear systems

For non-uniformly placed sensors, it is generally not possible to find similar closed-form expressions for the desired coefficients $\{w_{n,l}(\boldsymbol{k}, r)\}_{n,l}$. However we may resort to formulating a linear system of equations to find $\{w_{n,l}(\boldsymbol{k}, r)\}_{n,l}$, since we know both the approximating function g, as well as, the exponential function $\Gamma_r(t)\Psi_{\boldsymbol{k}}(\mathbf{x})$ we wish to approximate. Our suggested approach is to discretize (6.17) as follows. First, for each l we can formulate the following linear system for some fixed time snapshot $t_j > 0$,

$$\begin{bmatrix} g(\mathbf{x}_1-\mathbf{x}'_1, t_l-t_j) & \cdots & g(\mathbf{x}_N-\mathbf{x}'_1, t_l-t_j) \\ g(\mathbf{x}_1-\mathbf{x}'_2, t_l-t_j) & \cdots & g(\mathbf{x}_N-\mathbf{x}'_2, t_l-t_j) \\ \vdots & \ddots & \vdots \\ g(\mathbf{x}_1-\mathbf{x}'_I, t_l-t_j) & \cdots & g(\mathbf{x}_N-\mathbf{x}'_I, t_l-t_j) \end{bmatrix} \begin{bmatrix} w_{1,l}(\boldsymbol{k}, r) \\ w_{2,l}(\boldsymbol{k}, r) \\ \vdots \\ w_{N,l}(\boldsymbol{k}, r) \end{bmatrix} = \begin{bmatrix} \Psi_{\boldsymbol{k}}(\mathbf{x}'_1)\Gamma_r(t_j) \\ \Psi_{\boldsymbol{k}}(\mathbf{x}'_2)\Gamma_r(t_j) \\ \vdots \\ \Psi_{\boldsymbol{k}}(\mathbf{x}'_I)\Gamma_r(t_j) \end{bmatrix}$$

$$\Rightarrow \mathbf{G}_{l,j}\mathbf{w}_l(\boldsymbol{k}, r) = \mathbf{p}_j(\boldsymbol{k}, r). \tag{6.36}$$

Solving this system gives the coefficients $\{w_{n,l}(\boldsymbol{k}, r)\}_n$ for a fixed l. We can however proceed by stacking (6.36) for each l and taking several t_j's, $j = 1, \ldots, J$ to get

$$\begin{bmatrix} \mathbf{G}_{0,1}\,\mathbf{G}_{1,1}\cdots\mathbf{G}_{L,1} \\ \mathbf{G}_{0,2}\,\mathbf{G}_{1,2}\cdots\mathbf{G}_{L,2} \\ \vdots \qquad \vdots \\ \mathbf{G}_{0,J}\mathbf{G}_{1,J}\cdots\mathbf{G}_{L,J} \end{bmatrix} \begin{bmatrix} \mathbf{w}_0(\boldsymbol{k}, r) \\ \mathbf{w}_1(\boldsymbol{k}, r) \\ \vdots \\ \mathbf{w}_L(\boldsymbol{k}, r) \end{bmatrix} = \begin{bmatrix} \mathbf{p}_1(\boldsymbol{k}, r) \\ \mathbf{p}_2(\boldsymbol{k}, r) \\ \vdots \\ \mathbf{p}_J(\boldsymbol{k}, r) \end{bmatrix}$$

$$\mathbf{G}\mathbf{w}(\boldsymbol{k}, r) = \mathbf{p}(\boldsymbol{k}, r), \tag{6.37}$$

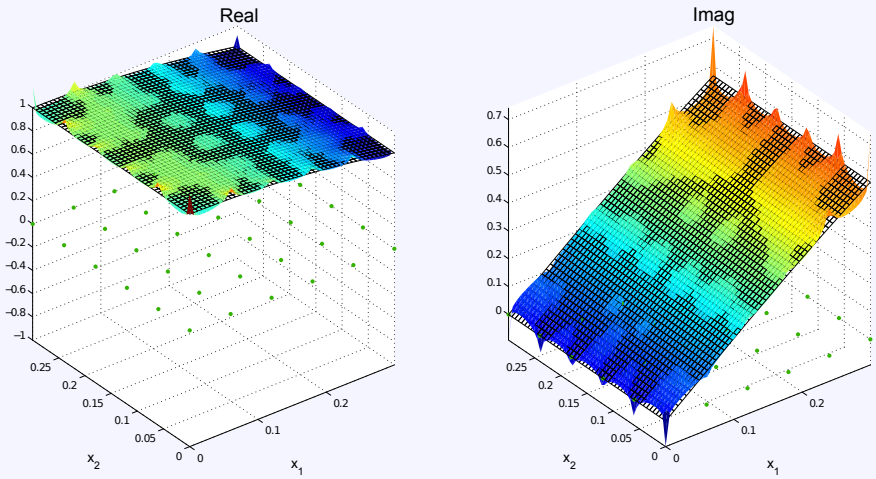
where $\mathbf{G} \in \mathbb{R}^{IJ\times N(L+1)}$ is a discretization of $g(\mathbf{x}, t)$, $\mathbf{p}(\boldsymbol{k}, r) \in \mathbb{R}^{IJ}$ are discretizations of the spatiotemporal sensing functions, whilst $\mathbf{w}(\boldsymbol{k}, r) \in \mathbb{R}^{N(L+1)}$ are the desired weights for each $\boldsymbol{k} \in \mathbb{R}^2$ and $r \in \mathbb{R}$. Consequently, in order to recover the desired field analysis coefficients, we would need to solve the system (6.37). In general, this system admits a least-squares solution if $IJ \geq N(L + 1)$, where the observation matrix $\mathbf{G}$ can be constructed from the Green's function of the problem at hand (i.e. (6.2), (6.3), (6.4), (6.6) and so on).

Although formulating this type of system to find the desired weights is rather straightforward and intuitive, it can sometimes lead to problems, in that, the conditioning of the

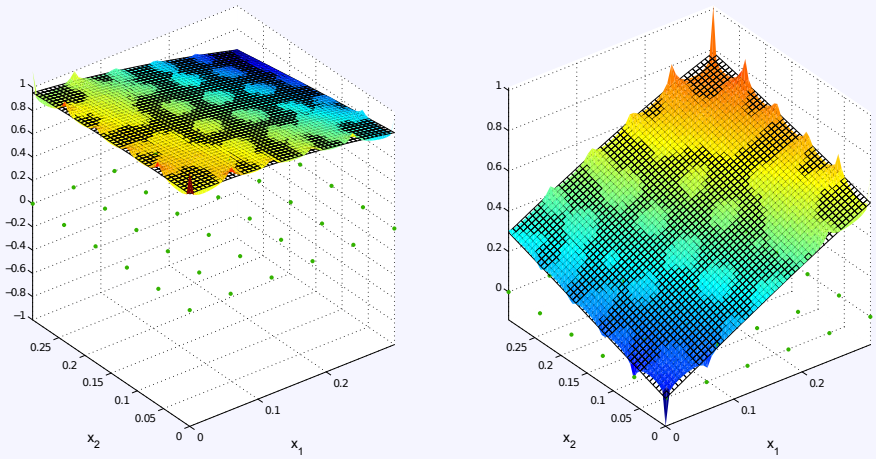
system can be rather poor. Specifically, the conditioning depends directly on the sensor locations $\mathbf{x}_n$, the sampling instants t_l and the Green's function $g(\mathbf{x}, t)$, of the underlying phenomena. Indeed for a Gaussian kernel, it is well known that the condition number of $\mathbf{G}$, i.e. $\mathrm{Cond}(\mathbf{G})$, can be very large and indeed this metric can be used as an indication for the estimation performance [125].

Nonuniform sampling: interpolation

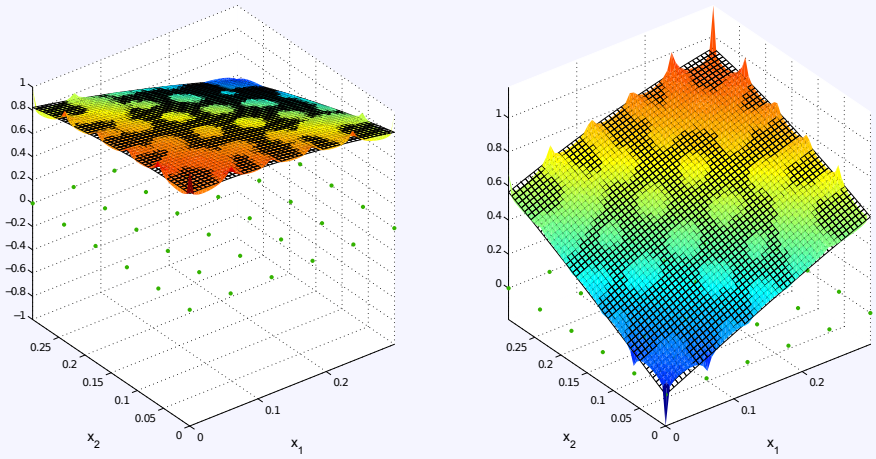
To enable us perform some meaningful estimation in situations where the Green's function and the sensor locations produce a badly conditioned matrix $\mathbf{G}$, we propose a second approach for the nonuniform spatial sampling setup that can still exploit the new universal framework proposed in this chapter. Essentially we want to return to a situation where the translates of the Green's function are on a uniform lattice. To do this, we assume that the spatial field samples are interpolated, on a uniform grid, using the interpolator $\gamma(\mathbf{x})$, such that $\hat{u}(\mathbf{x}, t_l) = \sum_{n=1}^{N} \varphi_n(t_l)\gamma(\mathbf{x} - \mathbf{x}_n)$. Then this new approximation of the underlying field is resampled uniformly at the new locations $\{\mathbf{x}_{\bar{\boldsymbol{n}}} = (\bar{n}_1\Delta_{x_1}, \bar{n}_2\Delta_{x_2}, \ldots, \bar{n}_d\Delta_{x_d})\}_{\bar{\boldsymbol{n}}\in\mathbb{N}^d}$ to obtain the corresponding data samples $\hat{\varphi}_{\bar{\boldsymbol{n}}}(t_l) = \hat{u}(\mathbf{x}_{\bar{\boldsymbol{n}}}, t_l)$. Hence the corresponding exponential reproducing coefficients can be recovered, using translates of the Green's functions at these new locations. Weighting the interpolated measurements by the obtained coefficients will produce an estimate for the desired sequence of generalized measurements. Besides avoiding the inversion of badly conditioned matrices, this new approach is also less intensive computationally, particularly when the matrix $\mathbf{G}$ is rather large.



(a) $\boldsymbol{k} = (2, 0)$



(b) $\boldsymbol{k} = (2, 1)$



(c) $\boldsymbol{k} = (2, 2)$

Figure 6.1.: **Exponential reproduction.** Reproduction of the 2-D sensing function $\Psi_{\boldsymbol{k}}(\mathbf{x}) = e^{\mathrm{j}k_1x_1+\mathrm{j}k_2x_2}$, assuming $N = 45$ uniformly placed sensors for the 2-D Green's function of Poisson's equation.

6.5. Source estimation using sensor networks

Based on the framework outlined in the chapter thus far, we now develop practical sensor network algorithms for estimating the sources of a physical field from its sensor measurements, and therefore solve the class of discrete ISPs driven by linear PDEs with constant coefficients. First, we propose a centralized estimation algorithm, which follows straightforwardly from Proposition 6.1, then we treat the distributed case by showing that, under the same SN assumptions outlined in Chapter 4 each sensor can compute the correct local measure. So that when the sensor nodes exchange and update their local measures, using a gossiping algorithm, they converge to the desired multidimensional generalized measurements $\mathcal{Q}(\boldsymbol{k}, r)$. As usual, from these generalized measurements they can each recover the unknown source parameters. In addition, we also briefly discuss how to deal with the usual SN nuisances, such as sensor noise and communication constraints.

6.5.1. Centralized source estimation

Assuming the source measurements have been made available at the fusion centre and that the PDE model—and Green's function—of the monitored phenomena are known, then the point and instantaneous source estimation scheme can be summarized as in Algorithm 6.1, for an arbitrary physical phenomena having PDEs with constant and linear coefficients.

Algorithm 6.1 Simultaneous Estimation of M point sources from field samples

Require: $\{\varphi_n(t_l)\}_{n=1,l=0}^{N,L}$, sensor locations $\{\mathbf{x}_n\}_n$, number of sources M, sampling interval Δ_t, diffusivity μ, spatial dimension d.

1: From $\{\mathbf{x}_n = (x_{1,n}, x_{2,n}, \ldots, x_{d,n})\}_n$ compute the sensor spacing $\boldsymbol{\Delta}_{\mathbf{x}}$.
2: Compute the Laplace transform $G(\boldsymbol{\omega}_{\mathbf{x}}, \omega_t)$ of the Green's function using (6.23).
3: Initialize $K_1, K_2 \geq 2M - 1$ and $r = 1$.
4: Compute coefficients $\{w_{\boldsymbol{n},l}(\boldsymbol{k}, 1)\}_{\boldsymbol{k}=(0,0)}^{(K_1,K_2)}$ using (6.35).
5: Compute the sequence $\{\mathcal{Q}(\boldsymbol{k}, 1) = \sum_{\boldsymbol{n},l} w_{\boldsymbol{n},l}(\boldsymbol{k}, 1)\varphi_{\boldsymbol{n}}(t_l)\}_{\boldsymbol{k}}$.
6: Denoise $\{\mathcal{Q}(\boldsymbol{k}, 1)\}_{\boldsymbol{k}}$ using Cadzow's algorithm.
7: Apply ACMP method to $\{\mathcal{Q}(\boldsymbol{k}, 1)\}_{\boldsymbol{k}}$ to obtain M pairs of $(c_m e^{-\mathrm{j}\tau_m/T}, \boldsymbol{\xi}_m)$.
8: For each m, $c_m = \left|c_m e^{-\mathrm{j}\tau_m/T}\right|$ and $\tau_m = T \arg\left(c_m e^{-\mathrm{j}\tau_m/T}\right)$.
9: **return** M intensities, activation times and locations $\{c_m, \tau_m, \boldsymbol{\xi}_m\}_{m=1}^{M}$.

6.5.2. Distributed source estimation

In the distributed set up we would like, as usual, for each node to estimate $\mathcal{Q}(\boldsymbol{k}, r)$ through localized interactions with its neighbours. We assume the same sensor network as described in Section 4.4, comprising of "smart" sensor nodes. In addition to knowing the Green's function of the monitored phenomena[5], these nodes are able to perform mathematical

[5]This comes for free since the SNs are designed to sense a particular phenomena: i.e. if we are sensing acoustic fields then we use the wave equation, for temperature and leakages we use the diffusion equation, and so on.

computations and also learn the network topology upon deployment (see assumption (A) in Section 4.4). Knowing the network topology and the Green's function of the underlying phenomena means that, each sensor n can compute independently its set of exponential reproducing coefficients $\{w_{\boldsymbol{n},l}\}_l$ using (6.32), if they are on a uniform grid. Otherwise, they can each solve the system (6.37). Then we propose the use of gossip, which involve message exchanges and updating, as outlined in Section 4.5. Specifically, gossiping is initiated when the n-th sensor contacts and exchanges its (complex) measure, given by

$$y_n(\boldsymbol{k}, r) = N \sum_{l=0}^{L} w_{n,l}(\boldsymbol{k}, r)\varphi_n(t_l), \tag{6.38}$$

with a neighbour, for example. Upon convergence, after several rounds of gossip, each sensor will have

$$\frac{1}{N}\sum_n y_n(\boldsymbol{k}, r) = \frac{1}{N}\sum_n N \sum_{l=0}^{L} w_{n,l}(\boldsymbol{k}, r)\varphi_n(t_l) = \mathcal{Q}(\boldsymbol{k}, r),$$

as desired. An appropriate (multidimensional) Prony-like method, such as ACMP, can now be applied independently by each sensor to estimate the unknown source parameters.

6.5.3. Filtering in the time-domain

Using the framework summarized by Proposition 6.1, we have been able to devise practical sensor network algorithms to solve our discrete ISP. During the sensing phase, although spatial prefiltering is generally not achievable we are still able to perform prefiltering in time. Under this fact the prefiltered samples, using a time-domain filter $h(t)$, obtained by the n-th sensor is: $\phi_n(t_l) = u(\mathbf{x}_n, t) \star h(t)|_{t=t_l} = f(\mathbf{x}, t) * g(\mathbf{x}, t) \star h(t)|_{\mathbf{x}=\mathbf{x}_n, t=t_l}$, where $\star$ is the time-convolution operator. Hence it may be helpful to pick a filter, that can improve our ability to reproduce exponentials. In light of this new formulation, the generator that will be used to reproduce exponentials from its space-time translates is

$$g_f(\mathbf{x}, t) = \int_{t'} g(-\mathbf{x}, -t - t')h(t')\mathrm{d}t', \tag{6.39}$$

which has the bilateral Laplace transform

$$G_f(\mathbf{s}_\mathbf{x}, s_t) = G(-\mathbf{s}_\mathbf{x}, -s_t)H(s_t). \tag{6.40}$$

We now have the freedom of designing $H(s_t)$ such that $G_f(\mathbf{s}_\mathbf{x}, s_t)$ has some desirable properties; in our framework it is favourable to choose $H(s_t)$ such that it $G_f(\mathbf{s}_\mathbf{x}, s_t)$ decays quickly enough, at least in the s_t-domain, in order to reduce approximation error when using the constant least-squares coefficients, as mentioned in (6.30).

6.5.4. A brief note on tackling sensor noise, and model mismatch and communication constraints

The presence of measurement noise in practical sensing devices is generally inevitable, as such the sensors measure a noisy version of the physical field, which we model here (using the multi-index notation $\boldsymbol{n}$) as: $\varphi^{\epsilon}_{\boldsymbol{n}}(t_l) = \varphi_{\boldsymbol{n}}(t_l) + \epsilon_{\boldsymbol{n},l}$ (cf. (4.2)) where $\epsilon_{\boldsymbol{n},l}$ is again AGWN. Consequently the FC (in the centralized SN) or the nodes themselves (in the distributed SN) will, in the process of computing $\mathcal{Q}(\boldsymbol{k}, r)$, end up with a noisy version $\mathcal{Q}_{\epsilon}(\boldsymbol{k}, r)$ of the true generalized measurements. As previously discussed, in Section 4.6.2 the noise in this sequence will now be coloured, however Prony-like methods perform better when the noise in $\mathcal{Q}_{\epsilon}(\boldsymbol{k}, r)$ is white. Consequently it is desirable to first prewhiten the sequence $\mathcal{Q}_{\epsilon}(\boldsymbol{k}, r)$. Certain multidimensional Prony-like methods (such as ACMP) are based on the construction of an enhanced matrix, of several stacked and cascaded Toeplitz matrices. Each Toeplitz matrix is formed using the terms in $\{\mathcal{Q}_{\epsilon}(\boldsymbol{k}, r)\}_{\boldsymbol{k}}$ for a fixed r. Consequently, we can prewhiten the noise in each block independently by designing and pre-multiplying by its corresponding filter, prior to recovering the unknown parameters. This can be achieved as described in Section 4.6.1, whilst paying close attention to the indices. For example, assuming we fix k_2 and r (say $k_2 = r = 0$ without loss of generality) and then utilize the resulting 1-D sequence, $\{\mathcal{Q}_{\epsilon}(\boldsymbol{k}, 0)\}_{\boldsymbol{k}=(0,0)}^{K_1,0}$, to construct a Toeplitz matrix. The correct prewhitening filter is $\mathbf{F}_{\text{pw}} = (\mathbf{C}_{\epsilon})^{\frac{\dagger}{2}}$, where $\mathbf{C}_{\epsilon}$ has been constructed using (4.37) with the corresponding weights $\{w_{n,l}(\boldsymbol{k}, 0)\}_{\boldsymbol{k}=(0,0)}^{(K_1,0)}$.

Furthermore, when the communication links in the network are noisy, the (local) inter-sensor messages can be quantized, to suppress the effects of noise. In this setting, we have seen in Section 4.5.1 that, these quantized communications must be performed correctly, so as to guarantee convergence to the true estimate of the generalized measurements. We propose adopting the same quantized gossip strategy in the case of communication over quantized links. Specifically, for our gossip-based algorithms over noisy channels, we utilize the quantized pairwise gossip scheme summarized by (4.23) during the gossip stage.

6.6. Applications and numerical results

In this section, we present some numerical results to validate the proposed framework for solving our linear PDE-driven inverse (source) problems. To examine several sampling setups, and applications, we present results for cases where the measured phenomena has been generated by:

(i) **Diffusion equation** – we assume a multiple source setting in 2-D with nonuniform sensor placement and then investigate the proposed approach based on field interpolation, as outlined in the latter part of Section 6.4.3. The results obtained are shown in Figure 6.2. Next we study the effects of, the sensor density (i.e. the uniform spatial sampling interval $\boldsymbol{\Delta}_{\mathbf{x}}$) and the source separation (i.e. the distance between two point sources) on the localization accuracy of the proposed framework. The results for this

simulation are shown in Figure 6.3

(ii) **Poisson's equation** – we consider a single source case in 3-D using nonuniform spatial samples, here we examine the linear systems approach of finding the desired coefficients. Presented in Figure 6.4.

(iii) **Wave equation** – we simulate a source field in 3-D with uniform sensor placement and we consider a situation where each sensor performs filtering in time using a cubic spline. Shown in Figure 6.5.

For the scenarios described above, i.e. the results presented in Figures 6.2, 6.4 and 6.5, the sensor measurements are simulated numerically using Matlab with the sensors distributed over a square region in 2-D (and equivalently a cubic region in 3-D). The measurements are then corrupted by AWGN, such that the SNR = 15dB as defined in (5.2), prior to applying our source estimation schemes. Furthermore for statistical significance, multiple independent trials are performed within each setup. As usual, each trial uses a new noise realization, and a new arbitrary sensor placement for the nonuniform sampling experiments. In Figures 6.2, 6.4 and 6.5, we denote with green '•' the sensor locations, whilst the estimated source location are shown as a red '×' and the true values by blue '+'.

6.6.1. Inverse source problem for the diffusion equation

We study the IDSP in light of this new universal framework. Specifically, we present some numerical results for the nonuniform sampling scenario, where the field is due to multiple localized and instantaneous sources. The performance of this new framework, which uses $r = 1, k_1 = k_2 = 0, 1, \ldots 15$ for the 2-D test function family $\{\Psi_{\boldsymbol{k}}(\mathbf{x})\Gamma_r(t) = e^{\mathrm{j}k_1x_1+\mathrm{j}k_2x_2}e^{\mathrm{j}rt/T}\}_{\boldsymbol{k},r}$, on the diffusion data is presented in Figure 6.2. In this simulation, we interpolate the field and then resample it on a regular grid, such that $\Delta_{x_1} = \Delta_{x_2} = 1/30$, before applying Algorithm 6.1 on this new (interpolated) field samples.

Additionally we use (6.35) to compute the desired coefficients for the weighted sum (6.14). Explicitly, since for the 2-D diffusion field, $g(\mathbf{x}, t) = \frac{1}{2\pi\mu t}e^{-\frac{\|\mathbf{x}\|^2}{4\mu t}}H(t)$ has the following multidimensional bilateral Laplace transform (see Appendix Section B.2.1):

$$G(\mathbf{s_x}, s_t) = \frac{1}{s_t - \mu\|\mathbf{s_x}\|^2}, \tag{6.41}$$

provided $\Re\big(s_t - \mu\|\mathbf{s_x}\|^2\big) > 0$, where $\Re(z)$ is used to denote the real part of a complex number z.

Remark 6.1. *The requirement that $\Re\big(s_t - \mu\|\mathbf{s_x}\|^2\big) > 0$ is a necessary condition for the transform integral to converge. This condition is almost always satisfied in reality. For instance, in 2-D where we use $s_t = \rho = \mathrm{j}r/T$ and $\mathbf{s_x} = -\boldsymbol{\kappa} = -\mathrm{j}(k_1, k_2) \Rightarrow \|\mathbf{s_x}\|^2 = -(k_1^2 + k_2^2)$. Given these choices, $\Re\big(\mathrm{j}r/T + \mu(k_1^2 + k_2^2)\big) > 0 \Rightarrow (k_1^2 + k_2^2) > 0$ which is satisfied for all $(k_1, k_2) \in \mathbb{Z} \setminus \{\mathbf{0}\}$. Similarly for the 3-D ISP, where $\mathbf{s_x} = -\boldsymbol{\kappa} = -(k_1, \mathrm{j}k_1, \mathrm{j}k_2)$, we end up with the condition that $k_2^2 > 0$ and is satisfied for any $k_2 \in \mathbb{Z} \setminus \{0\}$.*

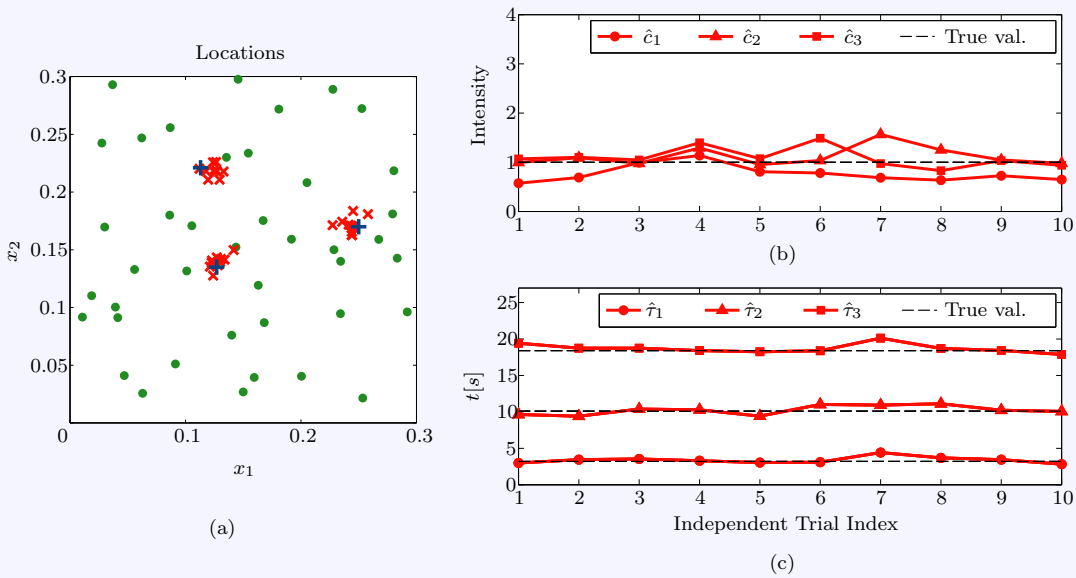


Figure 6.2.: **Centralized estimation of non-localized sources using noisy measurements obtained by randomly placed sensors.** The field is sampled at $1Hz$ for $T = 25s$ using 41 uniformly placed sensors, samples assumed to have SNR = 15dB. (a) Location estimates (b) Estimated intensities and (c) Activation times. We use Algorithm 6.1 with $r = 1$ and $K_1 = K_2 = 15$.

Therefore, by substituting (6.41) into (6.35), with $\boldsymbol{\kappa} = \mathrm{j}\boldsymbol{k} = \mathrm{j}(k_1, k_2)$ and $\rho = \mathrm{j}r/T$, we can write the desired coefficients as:

$$w_{\boldsymbol{n},l}(\boldsymbol{k}, r) = \Delta_{x_1}\Delta_{x_2}\Delta_t(\mu(k_1^2 + k_2^2) + \mathrm{j}r/T)e^{\mathrm{j}\Delta_{x_1}k_1n_1 + \mathrm{j}\Delta_{x_2}k_2n_2 + \mathrm{j}\Delta_t rl/T}. \tag{6.42}$$

All the estimated parameters have been recovered reliably, specifically the estimated locations (shown in Figure 6.2(a)) and activation times given in Figure 6.2(b) (usually the two main parameters of interest) are seen therein to vary only marginally around their true values, even when the measurement SNR is relatively low.

Effects of sensor and source separation on the location estimates

In this experiment, we consider the diffusion field induced by two point sources with $c_1 = c_2 = 1$, $\tau_1 = \tau_2 = 1$ and fix the location of the first source $\boldsymbol{\xi}_1 = (\xi_{1,1}, \xi_{2,1})$ whilst the location of the second source $\boldsymbol{\xi}_2 = (\xi_{1,2}, \xi_{2,2})$ is allowed to vary such that the separation between them $S_{\text{source}} = \|\boldsymbol{\xi}_2 - \boldsymbol{\xi}_1\| \in \{0.04 + \frac{0.26i}{11}\}_{i=0}^{11}$. Similarly, the sensor density also varies, such that the uniform spatial sampling interval $\boldsymbol{\Delta}_{\mathbf{x}} = (\Delta_{x_1}, \Delta_{x_2})$ where $\Delta_{x_1} = \Delta_{x_2} \in \{0.05, 0.06, 0.075, 0.015, 0.3\}$.[6] The field measurements, sampled at 1Hz for $T = 20$ seconds, by the sensors network are assumed to be noisy with fixed SNR= 20dB. Consequently, for

[6]This yields uniform 2−D sensor arrays of size $\{8 \times 8, 7 \times 7, \ldots, 3 \times 3\}$ respectively.

each fixed value of S_{source} and $\mathbf{\Delta_x}$, Algorithm 6.1 with $K_1 = K_2 = 4$ and $r = 0$ is used to recover the estimates $\hat{\boldsymbol{\xi}}_1$ and $\hat{\boldsymbol{\xi}}_2$ of the true source locations $\boldsymbol{\xi}_1$ and $\boldsymbol{\xi}_2$ respectively. The pairing $(\hat{\boldsymbol{\xi}}_m, \boldsymbol{\xi}_m)$ of the true value and its estimate is chosen to minimize the overall error (with respect to the Euclidean distance). The RMSE for each of the estimates, using 1000 independent trials of the experiment, is computed and provided in Figure 6.3.

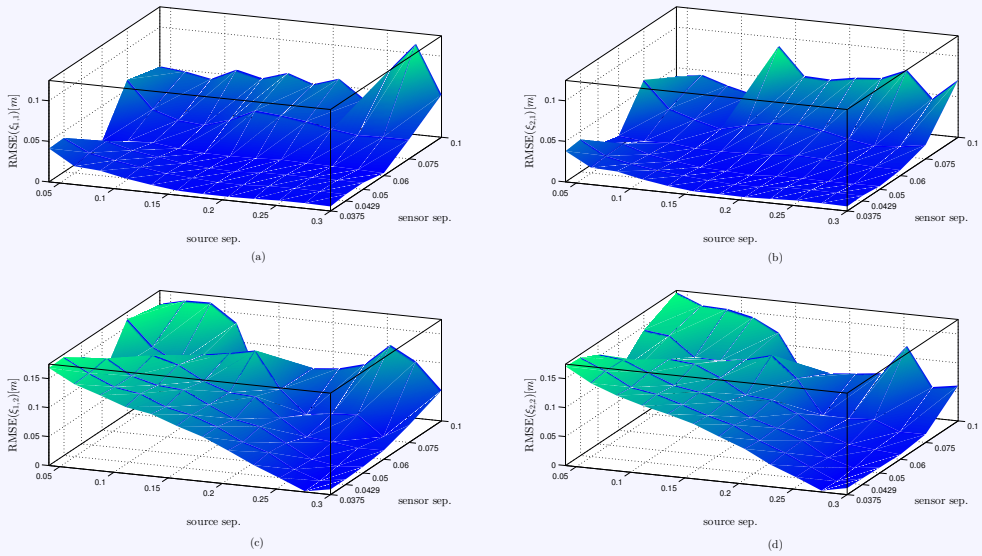


Figure 6.3.: **Effects of sensor density and source separation on the performance of the localization algorithm.** The field induced by two sources is sampled at 1Hz for $T = 20s$ with varying sensor density and source separation. For each value sensor separation and source separation the obtained noisy measurements, with SNR= 20dB, are used to estimate the locations of the unknown diffusion sources. We utilize $K_1 = K_2 = 4$ and $r = 0$ with the estimation simultaneous recovery scheme summarized in Algorithm 6.1. The RMSE of the estimates obtained over 1000 independent trials are shown: i.e. in (a) RMSE of $\hat{\xi}_{1,1}$, in (b) RMSE of $\hat{\xi}_{2,1}$, in (c) RMSE of $\hat{\xi}_{1,2}$ and in (d) RMSE of $\hat{\xi}_{2,2}$.

Observe in Figure 6.3 that, in line with our expectation, the performance of our estimation algorithm improves as the sensor density and separation between the two sources increases. In particular, the RMSE of the estimates for the first source—i.e. RMSE($\hat{\boldsymbol{\xi}}_1$) as shown in Figure 6.3(a) and (b)—decreases when the sensor density increases. This is a consequence of the reduction in the approximation error obtained in the exponential reproduction step, as $\mathbf{\Delta_x}$ decreases. Furthermore, the effect of the source separation on the localization performance becomes more noticeable as the sensor spacing decreases. For instance, when $\mathbf{\Delta_x} = (0.05, 0.05)$ we notice a gradual but steady decrease in the RMSE of $\hat{\boldsymbol{\xi}}_1$ in Figure 6.3(a) and (b). This increase in estimation performance is even higher for the estimates of the second source, $\hat{\boldsymbol{\xi}}_2$, as seen in Figure 6.3(c) and (d).

6.6.2. Inverse source problem for Poisson's Equation

Herein we study the ISP in relation to Poisson's equation under the assumption that the source of the field is localized. There is no time dependence here, so we only focus on recovering the source location and intensity. Specifically, since the field is stationary we only use single time measurements for the estimation. Although when several temporal measurements are available, we may simply average the estimates obtained at each snapshot.

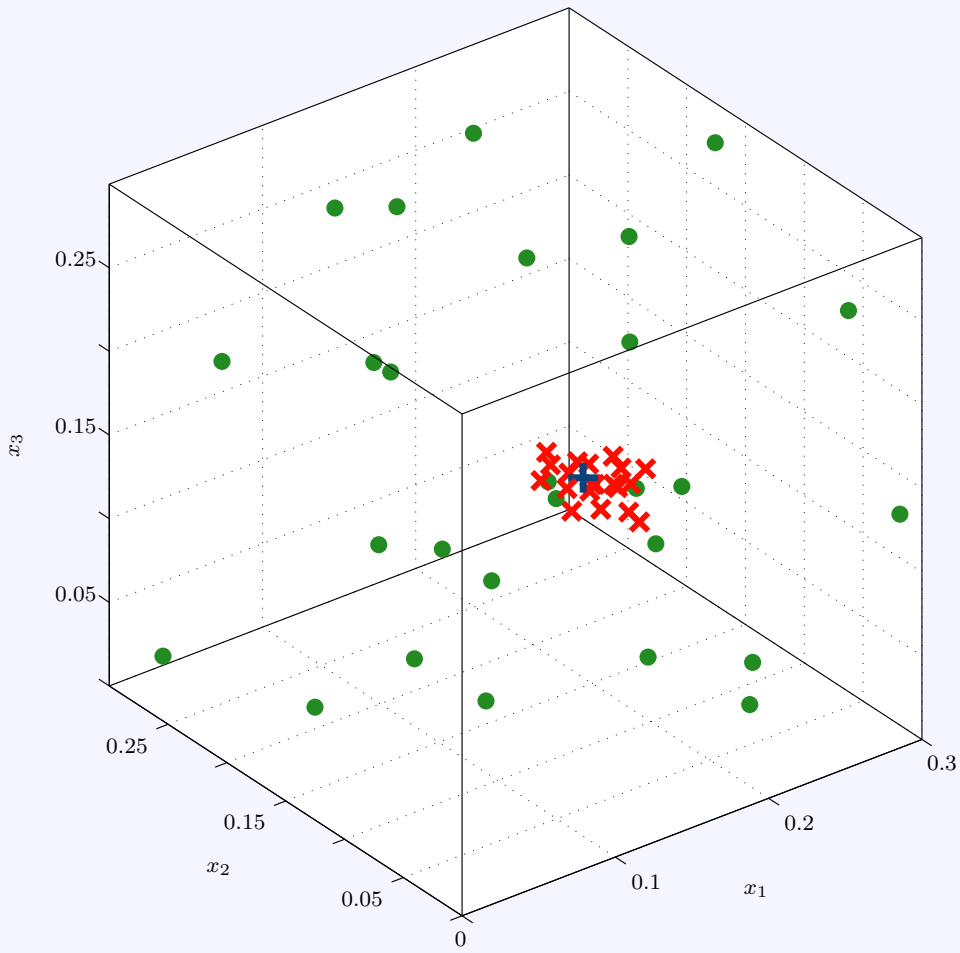
We simulate the 3-D field modelled by Poisson's equation using (6.1) by evaluating it at the randomly chosen sensor locations. These measurements are then corrupted with AWGN. For the estimation algorithm we use the linear system approach, discussed in Section 6.4.3, for computing the proper weights of the summation (6.14). Specifically, we use $r = 0$ and $K_1 = K_2 = 7$ to formulate the following linear system according to (6.36):

$$\begin{bmatrix} g(\mathbf{x}_1-\mathbf{x}'_1) & \cdots & g(\mathbf{x}_N-\mathbf{x}'_1) \\ g(\mathbf{x}_1-\mathbf{x}'_2) & \cdots & g(\mathbf{x}_N-\mathbf{x}'_2) \\ \vdots & \ddots & \vdots \\ g(\mathbf{x}_1-\mathbf{x}'_I) & \cdots & g(\mathbf{x}_N-\mathbf{x}'_I) \end{bmatrix} \begin{bmatrix} w_{1,l}(\boldsymbol{k},0) \\ w_{2,l}(\boldsymbol{k},0) \\ \vdots \\ w_{N,l}(\boldsymbol{k},0) \end{bmatrix} = \begin{bmatrix} \Psi_{\boldsymbol{k}}(\mathbf{x}'_1) \\ \Psi_{\boldsymbol{k}}(\mathbf{x}'_2) \\ \vdots \\ \Psi_{\boldsymbol{k}}(\mathbf{x}'_I) \end{bmatrix},$$

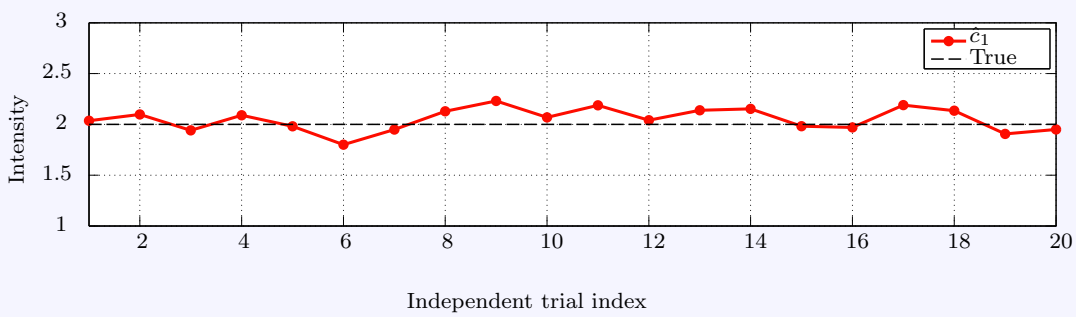
where $g(\mathbf{x}) = -\frac{1}{4\pi\|\mathbf{x}\|}$, for $I = 1000$ such that $\mathbf{x}'_i$ is obtained from a lexicographic ordering of $\{(i_1\delta_{x_1} + \varepsilon, i_2\delta_{x_2} + \varepsilon, i_3\delta_{x_3} + \varepsilon)\}_{i_1,i_2,i_3=0}^{9}$ and so $\delta_{x_1} = \delta_{x_1} = \delta_{x_1} = 0.03$.[7]

The simulation results are shown in Figure 6.4, showing the source location and intensity estimates for each of the 20 independent trials, the estimated 3-D locations (presented in the scatterplot Figure 6.4(a)) and the scalar intensities (plotted in the scatterplot Figure 6.4(b)) show small variations around the ground truth parameters. Particularly for the location estimates the overall variation is much smaller than the average sensor spacing. Furthermore the condition number for the matrix $\mathbf{G}$ was in the range $15 - 38$, which is reasonably small for our purposes.

[7]The slight shift $\varepsilon \geq 0$ is used here to avoid the singularity of $g(\mathbf{x})$ $\|\mathbf{x}\| = 0$.



(a)



(b)

Figure 6.4.: **Centralized estimation of single localized source for Poisson's equation using uniformly placed sensors.** The field is sampled at $1Hz$ for $T = 1s$ using 27 randomly placed sensors, samples assumed to have SNR = 15dB. (a) Location estimates (b) Estimated intensities. Estimation is performed using Algorithm 6.1 with $r = 0$ and $K_1 = K_2 = 7$.

6.6.3. Acoustic source localization: inverse source problem for the wave equation

We conclude this discussion of numerical results with a problem in acoustic source localization. Here the physical phenomena is assumed to be governed by the wave equation (6.5). The single instantaneous and localized source field is simulated numerically in the 3-D case using $N = 27$ uniformly placed sensors. Each of these sensors are assumed to acquire the field by filtering it (in time) at their precise location with the third order B-spline before sampling and then corrupting their observations with AWGN. We apply the proposed approach of Section 6.5.3 by first computing the desired coefficients $w_{\boldsymbol{n},l}(\boldsymbol{k}, r)$ from the filtered Green's function $g_f(\mathbf{x}, t)$, which has the bilateral Laplace transform, $G_f(\mathbf{s_x}, s_t) = G(\mathbf{s_x}, s_t)H(s_t)$, where

$$H(s_t) = \left(\frac{1 - e^{-s_t}}{s_t}\right)^4, \tag{6.43}$$

is the Laplace transform for the third order B-spline[8], whereas

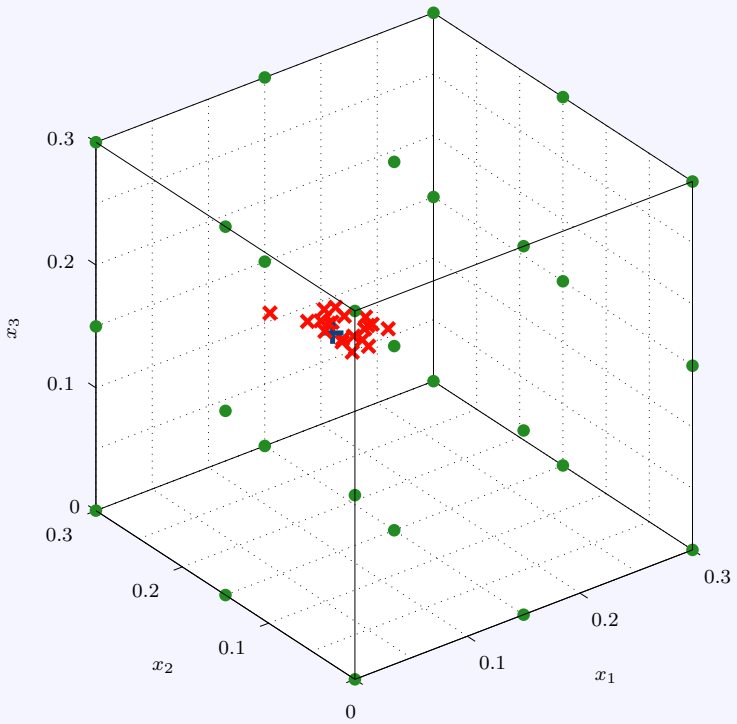
$$G(\mathbf{s_x}, s_t) = \frac{1}{\|\mathbf{s_x}\|^2 - (s_t/c)^2}. \tag{6.44}$$

A derivation of this expression is provided in Section B.2.2. Combining these gives us the desired coefficients,

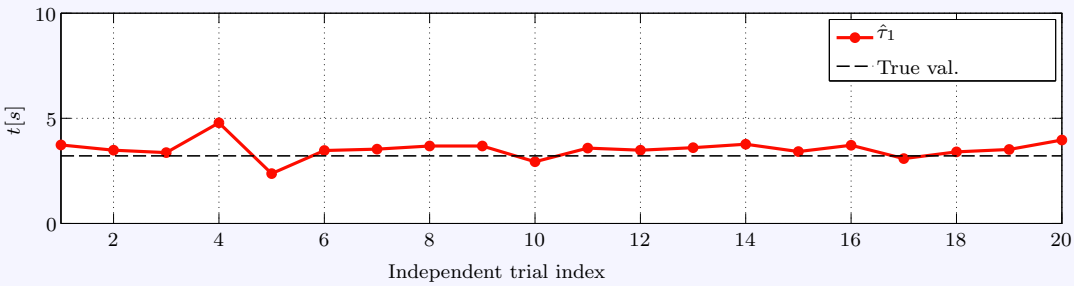
$$w_{\boldsymbol{n},l}(\boldsymbol{k}, r) = \Delta_{x_1}\Delta_{x_2}\Delta_{x_3}\Delta_t \frac{(r/T)^4\left(k_2^2 - \left(\frac{r}{cT}\right)^2\right)}{1 - e^{-\mathrm{j}r/T}} e^{k_1 n_1 \Delta_{x_1} + \mathrm{j}k_1 n_2 \Delta_{x_2} + \mathrm{j}k_2 n_3 \Delta_{x_3} + \mathrm{j}rl\Delta_t/T}, \tag{6.45}$$

where $\Delta_{x_1} = \Delta_{x_2} = \Delta_{x_3} = 0.1$ and $\Delta_t = 1$. When we chose $K_1, K_2 = 9$ and $r = 1$ for the sensing function family $\{\Psi_{\boldsymbol{k}}(\mathbf{x})\Gamma_r(t) = e^{k_1 x_1 + \mathrm{j}k_1 x_2 + \mathrm{j}k_2 x_3} e^{\mathrm{j}rt/T}\}_{\boldsymbol{k},r}$. The estimation results for this uniform sampling case is presented in Figure 6.5, where the estimates of the source location, in (a), and activation time, in (b), are plotted over the ground truth values. We notice that the estimates are reliable once more even when we have noisy measurements.

[8]The zero order B-spline is taken as the indicator function on [0,1], from this we define the n-th order B-spline as the spline generated by convolving $(n + 1)$ zero order B-splines.



(a)



(b)

Figure 6.5.: **Centralized estimation of non-localized sources for the Wave equation using uniformly placed sensors.** The field is sampled at $1Hz$ for $T = 20s$ using 27 randomly placed sensors, samples assumed to have SNR = 15dB. (a) Location estimates and (b) Estimated activation times. We use Algorithm 6.1 with $r = 1$ and $K_1 = K_2 = 9$.

6.7. Summary and conclusions

A novel framework for solving the inverse source problem relating to fields driven by linear PDEs with constant coefficients is discussed. This new framework was developed to extend on the previous method of Chapter 4, in such a way that we are able to obtain the desired sequence of generalized measurements—governed by Prony-like systems—not only for the diffusion equation but also for a wider class of PDE models. Besides this difference, both methods share certain subtle similarities, as well as differences which we highlight briefly in Table 6.1. Furthermore, for the approach of Chapter 4, we can also derive explicitly the desired coefficients, as shown in Section B.3.

Table 6.1.: Comparing the frameworks of Chapters 4 and 6.

	Framework of Chapter 4	Framework of Chapter 6
PDE models	Suitable only for the 2-D diffusion equation. A related approach has been outlined by the authors of [57, 56] for the wave equations for point sources localization alone.	Theoretically suitable for any ISP with arbitrary linear PDEs provided the underlying Green's function is *well-behaved*; ideally, it should satisfy the Strang-Fix conditions, or at least approximately with a sufficient decay for the Poisson summation formula to converge.
Source models admitted	Generalization to line and polygonal sources is straightforward and simulation results demonstrate the effectiveness of this approach in this setting.	Only point sources are considered presently for both instantaneous and non-instantaneous models. Possible extension to non-localized models is left as a future research direction.
Source recovery in rooms	Possible here and comes for free due to Green's second identity. This is enabled by the fact that the barriers can be modelled as image sources which are outside the monitored region.	This is possible in certain cases – primarily when the sources aren't very close to the walls. The reason here is due to leakage, i.e. the virtual source will also be close to the wall and so that when the Green's function is shifted to a sensor near the wall, its effect will leak into the generalized measurements. This border effect is well-known in the FRI literature for spike train recovery, see [116] for example.
Spatial dimensions	Presently a 2-D framework, although it is possibly to extend the approach to 3-D using the appropriate Green's second identity. The difficulty here is in designing a 3-D sensing function that satisfies the adjoint equation.	Can be easily generalized to many dimensions, as the sensing functions can assume very general structures, as long as we can solve the resulting Prony-like system.
Distributed Estimation	A distributed implementation of this approach is possible although non-trivial, as we saw in Chapter 4.	Can be implemented in a distributed network, this extension is relatively simple providing the nodes know the network topology.

Chapter 7.

Conclusion

7.1. Summary

In this thesis we have considered the problem of sampling multidimensional physical fields governed by linear and constant coefficient PDEs, using sensor networks. We argued that these fields are spatially nonbandlimited and so a classical reconstruction approach will not be effective, since it is generally impossible to realize a spatial antialiasing filter before the field is sampled with a sensor network. To this end, we first developed a framework to analyse the spatiotemporal sensor measurements of 2-D diffusion fields induced by a specific class of source models. A closer look at our approach enabled us to establish a missing bridge between the analysis of ISPs for PDE-driven fields with generalized sampling theory. As a results, we developed a new sampling framework for analysing ISPs of 2-D and 3-D fields governed by linear, constant coefficient PDEs.

We commenced Chapter 2 with an overview of classical sampling theory and highlighted some of the issues with applying it to the multidimensional sampling problem at hand. Next we discussed two generalizations of classical sampling theory, namely: sampling in *shift-invariant spaces* and sampling signals with *finite rate of innovation*, highlighting also their corresponding reconstruction algorithms for single and multiple dimensions. Arguing qualitatively that, under certain conditions, the field reconstruction and source estimation problems are equivalent, we therefore concluded the chapter with an account of state-of-the-art methods for sampling diffusion, potential and wave fields in order to solve their resulting field or source estimation problem.

We considered the continuous IDSP in Chapter 3. First, we formulated a well-posed IDSP by assuming that the measured field was induced by sources with some known parametric models. Subsequently, the application of Green's second identity allowed us to show that the unknown parameters for the localized (and temporally instantaneous or non-instantaneous) source models can be found by estimating the amplitudes and frequencies of a sum of sinusoids, from a set of *generalized measurements*. These generalized measurements, as we showed, can be computed exactly from the continuous field measurements. For the class of non-localized sources the sum of exponentials structure is lost, however we showed how to restore it using tools from complex analysis. Finally the chapter concluded with some extensions to solving the IDSP in bounded regions, as well as, solving the ISP in the

presence of advection.

Using the insights gained from solving the continuous IDSP in Chapter 4, we developed explicit noise robust SN algorithms for solving discrete IDSP, given arbitrary spatiotemporal sensor measurements of the induced diffusion field. We achieved this by showing how to compute the desired generalized measurements from the SN samples, in a centralized or distributed fashion. For the centralized approach, we used numerical quadrature to obtain the generalized measurements; whereas for the distributed SN, a physics-driven consensus scheme is derived, such that the nodes of the network converge to the desired generalized measurements through local unquantized and quantized communications with neighbouring nodes. Hence given these generalized measurements, noise robust variations of Prony's method are proposed to estimate all unknown source parameters.

In Chapter 5 we gave extensive numerical simulations in order to study the proposed centralized and distributed algorithms. In our results we studied the performance and validity of our algorithms using both synthetic and real temperature field data. Notably, we showed that it is possible to recover the unknown sources, as well as, faithfully reconstruct the underlying diffusion field.

Finally we established a new result in Chapter 6, that connected the estimation approach derived in Chapter 4 with certain results from generalized sampling theory. This connection was based on computing the desired generalized measurements using properly weighted linear combinations of the sensor measurements. In particular this formulation reduced the ISP for the class of linear PDEs, with constant coefficients, to one of reproducing exponentials using shifted versions of a certain prototype function, which can be solved (approximately) using certain results from generalized sampling. We derived results and proposed practical sensor network algorithms for both uniform and nonuniform sampling setup. Finally, some numerical results were presented to demonstrate the validity of this universal framework.

7.2. Comparisons

We summarize some of the main similarities and differences between both techniques developed in this thesis for solving ISPs. Both techniques rely on the knowledge of the constant coefficient (e.g. the wave speed c or diffusivity μ for the wave and diffusion equations) present in the linear PDE considered. The assumption that these are constant is sufficient for many sensor network applications encountered in environmental monitoring, acoustic source localization and so on, where the medium is usually homogeneous or at most varies only marginally about a constant value. Another key assumption made is that the monitoring sensors are synchronized, this limitation is key for recovering the activation times; although using the framework developed in Chapter 4, the sensors do not necessarily have to sample the field at the same instant. What is important is that the time at which they start sampling the field is synchronized and similarly for the time at which they obtain the last temporal sample.

Other important similarities and differences are stated in the Venn diagram shown in Figure 7.1.

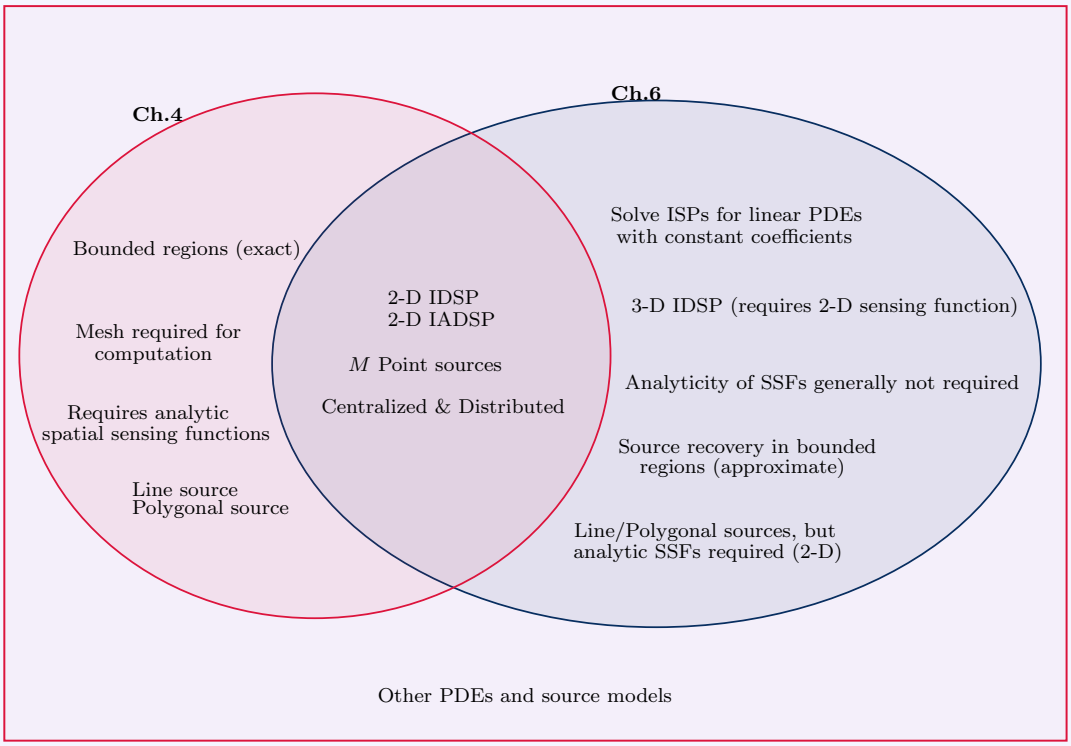


Figure 7.1.: **Comparing the source estimation frameworks.** Similarities and differences between the Green's second identity-based framework (**Ch.4**) and the universal sampling-based framework (**Ch.6**). Note that we abbreviated the phrase 'spatial sensing function' with SSF.

7.3. Outlook and future research

We conclude this thesis with some outlook on future research directions related to our work.

1. **Arbitrary source shapes** – We presented results for a specific class of spatially localized and non-localized source distributions, in our work. For the non-localized sources it was assumed that they are spatially well-modelled by a single straight line or convex polygon. In certain applications where these models may not hold but we still want to understand the precise shape of the sources, then a more tailored approach is needed. In other words, a framework for estimating sources of arbitrary shapes given samples of the induced field is still missing. One approach that may lead to promising new theory is to establish a link between a recent theory for sampling curves with FRI [117] and our results presented in Chapter 6 of this thesis.

2. **Extension to non-linear PDEs** – Despite the breadth of PDE models captured by

our framework, many other physical phenomena may be more accurately modelled using non-linear PDEs. Consider for example the Burgers equation (equivalently the Kardar-Parisi-Zhang (KPZ) equation which models the erosion of material from a solid surface) [20], it has been shown that this equation can be transformed into the heat equation using the Hopf-Cole transformation [74]. Therefore, for non-linear PDEs which can be transformed into linear ones, in this way, it is useful to understand the possibility/limitations of solving the associated ISP by applying our framework to the properly transformed sensor measurements.

3. **Complexity of algorithms** – In the sensor network setting, specifically in distributed sensor networks where resources are on a budget, understanding the complexity of the underlying estimation algorithms can be useful for projecting sensor network lifetime. This is even more useful when monitoring hostile environments which may be inaccessible to humans, since this information can assist with a better scheduling of maintenance (battery replacement) tasks. Furthermore, it will also enable us to compare the Green's second identity-based framework (of Chapter 4) with the universal sampling-based approach proposed in Chapter 6.

4. **Joint node localization and source estimation** – In the development of our source estimation framework, we assumed that the location of every monitoring node in the sensor network is known. In some very harsh situations this assumption may not hold. Hence the sensor locations, in addition to the field sources, must also be estimated. Typically, current SN localization methods can be classified as either active or passive localization. Active localization methods estimate the nodes using artificially induced signals, whereas passive methods do not. Consequently, passive localization methods are more attractive since they do not need additional infrastructure, therefore making them more cost effective. In light of this, a framework that can simultaneously localize the sensor nodes and estimate the unknown sources from the field measurements will have a significant positive impact in the area, by minimizing the costs of maintenance tasks for example.

5. **Atomic norm denoising for spectral estimation** – Atomic norm minimization [22] has emerged as a promising approach for solving line spectral estimation problems, such as the Prony-like system (2.7), motivated by recet work on atomic norms [39]. This approach exploits the sparsity property of the signal as in compressed sensing (sparse recovery) but, in contrast, operates directly on continuously parametrized dictionaries. In so doing it eliminates the need for gridding or discretizing the dictionaries/atoms, and the drawbacks thereof.

   Herein, the parsimonious signal model is given by:

$$\mathbf{r} = \sum_{m=1}^{M} c_m \mathbf{a}(\theta_m), \tag{7.1}$$

where the $\mathbf{r}$ coincides with a vectorization of our 1-D generalized measurements when we have $\mathbf{a}(\theta_m) = \left[e^{\mathrm{j}2\pi t_0\theta_m}, e^{\mathrm{j}2\pi t_1\theta_m}, \ldots, e^{\mathrm{j}2\pi t_K\theta_m}\right]^{\mathsf{T}}$, and $\theta_m \in [0, 1)$. Furthermore, by considering the dictionary or set of atoms $\mathcal{A} = \{\mathbf{a}(\theta) : \theta \in \Theta\}$, and defining the atomic norm $\|\cdot\|_{\mathcal{A}}$ as follows:

$$\|\mathbf{r}\|_{\mathcal{A}} = \inf\{\|\mathbf{c}\| : \mathbf{r} = \sum_{m=1}^{M} c_m \mathbf{a}(\theta_m)\} = \inf\{t > 0 : \mathbf{r} \in t\mathrm{conv}(\pm\mathcal{A})\}, \tag{7.2}$$

then one can formulate and solve convex optimization problems (see [22]) to recover the unknowns θ_m and c_m given measurements of the signal $\mathbf{r}$, possibly with missing or corrupt data. It may be useful to consider and investigate this new approach for recovering the unknown source parameters from the sequence of generalized measurements because it has been shown, under certain conditions, to compare favourably against classical spectral estimation methods [22].

6. **Space-time trade-off for the sampling density** – In many sensor network applications the placement of sensors remains an art. Related to this, due to the cost of sensors, the geometry of the monitored region or medium for example, practitioners are often constrained in terms of the density of sensors that can be placed, as well as where these sensors can be safely placed. Consequently, it may usually be more convenient to sample more densely in time to compensate for reducing the number of sensors (i.e. the spatial sampling density). In the context of our framework, the problem of understanding this space-time sampling density trade off is interesting but missing. Thus a thorough study of this trade off will serve to provide a more complete treatment of physics-driven inverse source problems under the framework provided in this thesis.

Appendices

Appendix A.

Chapter 4

A.1. The Cramer-Rao bound

Given noisy sensor measurements,

$$\mathcal{R}^{\epsilon}(k) = \mathcal{R}(k) + \frac{1}{N}\sum_{n,l} w_{n,l}(k)\epsilon_{n,l} \tag{A.1}$$

$$= \frac{1}{N}\sum_{n,l} w_{n,l}(k)\varphi_n(t_l) + \frac{1}{N}\sum_{n,l} w_{n,l}(k)\epsilon_{n,l}, \tag{A.2}$$

for $k = 0, 1, \ldots, K$. Let $\mathbf{r} = (\mathcal{R}^{\epsilon}(0), \mathcal{R}^{\epsilon}(1), \ldots, \mathcal{R}^{\epsilon}(K))^{\mathsf{T}}$ and as defined in Section 4.6.2, let $\mathbf{W}(k) \in \mathbb{C}^{N\times(L+1)}$ be the matrix with entries $[\mathbf{W}(k)]_{n,l+1} = w_{n,l}(k)$ and $\mathbf{w}(k) = \mathrm{vec}(\mathbf{W}(k))$. Furthermore, introducing the matrix $\widetilde{\mathbf{W}} = (\mathbf{w}(0), \mathbf{w}(1), \ldots, \mathbf{w}(K))^{\mathsf{T}}$ allows us to obtain the following matrix-vector expression for (A.2):

$$\mathbf{r} = \widetilde{\mathbf{W}}(\boldsymbol{\varphi} + \boldsymbol{\epsilon}) \tag{A.3}$$

where $\boldsymbol{\varphi}$ and $\boldsymbol{\epsilon}$ are formed in the same way as $\mathbf{w}(k)$. It can then be shown that the Fisher information matrix (FIM) is given by [108]:

$$\mathbf{I}(\boldsymbol{\theta}) = \frac{2}{\sigma^2}\mathrm{Re}\Big\{\mathbf{G}^{\mathsf{H}}(\widetilde{\mathbf{W}}\widetilde{\mathbf{W}}^{\mathsf{H}})^{-1}\mathbf{G}\Big\}, \tag{A.4}$$

where $\boldsymbol{\theta} = (c_1, \tau_1, \xi_{1,1}, \xi_{2,1}, c_2, \tau_2, \xi_{1,2}, \xi_{2,2}, \ldots, c_M, \tau_M, \xi_{1,M}, \xi_{2,M})^{\mathsf{T}}$, i.e. $\boldsymbol{\theta} \in \mathbb{R}^{4M}$, and

$$\mathbf{G} = \begin{bmatrix} \nabla_{\boldsymbol{\theta}}(0) \\ \nabla_{\boldsymbol{\theta}}(1) \\ \vdots \\ \nabla_{\boldsymbol{\theta}}(K) \end{bmatrix}. \tag{A.5}$$

For the single source estimation result shown in Figure 5.12, $\mathcal{R}(k) = c_1 e^{-\mathrm{j}\tau_1/T} e^{-k(\xi_{1,1}+\mathrm{j}\xi_{2,1})}$ with $M = 1$ and $k = 0, 1$, then

$$\mathbf{G} = \begin{bmatrix} e^{-\mathrm{j}\tau_1/T} & e^{-\mathrm{j}\tau_1/T-(\xi_{1,1}+\mathrm{j}\xi_{2,1})} \\ -\mathrm{j}\frac{c_1}{T} e^{-\mathrm{j}\tau_1/T} & -\mathrm{j}\frac{c_1}{T} e^{-\mathrm{j}\tau_1/T-(\xi_{1,1}+\mathrm{j}\xi_{2,1})} \\ 0 & -c_1 e^{-\mathrm{j}\tau_1/T-(\xi_{1,1}+\mathrm{j}\xi_{2,1})} \\ 0 & -\mathrm{j}c_1 e^{-\mathrm{j}\tau_1/T-(\xi_{1,1}+\mathrm{j}\xi_{2,1})} \end{bmatrix}^{\mathsf{T}} . \tag{A.6}$$

The Cramer-Rao bound (CRB) is then obtained by inverting the FIM.

Appendix B.

Chapter 6

B.1. The Generalized Strang-Fix Conditions

We first state the classical Strang-Fix condition [141], for multidimensional polynomial reproduction, in the following lemma.

Lemma B.1 (Strang-Fix condition)**.** *For any kernel $\psi(\mathbf{x}) \in \mathcal{H}_C^p$, the following statements are equivalent:*

1. $\hat{\psi}(\mathbf{0}) \neq 0$ *and* $\nabla^{\boldsymbol{\alpha}}\hat{\psi}(2\pi\boldsymbol{\ell}) = 0$, *where* $\hat{\psi}(\mathrm{j}\boldsymbol{\omega})$ *is the Fourier transform of* ψ *and* $\boldsymbol{\ell} \in \mathbb{Z}^d \setminus \{\mathbf{0}\}$

2. $\mathbf{x}^{\boldsymbol{\alpha}} = \sum_{\boldsymbol{n}\in\mathbb{Z}^d} c_{\boldsymbol{\alpha},\boldsymbol{n}}$.

Proof. The proof of this lemma depends on the Poisson summation formula – which connects the values of a function ψ on the lattice $\mathbb{Z}^d$ with its Fourier transform $\hat{\psi}$ on the lattice $2\pi\mathbb{Z}^d$. For which both sides of the summation converge (absolutely) when ψ and $\hat{\psi}$ decay sufficiently quickly. The proof is based on taking the Fourier transform of $\psi(\boldsymbol{n}) = \boldsymbol{n}^{\boldsymbol{\alpha}}\varphi(\mathbf{x}-\boldsymbol{n})$ and applying the frequency differentiation property to it. The implication 1. $\Rightarrow$ 2. follows immediately from this expression, whereas the converse follows by induction. A complete proof can be found in [141]. □

To demonstrate the multidimensional generalized Strang-Fix condition, we require that the function $\psi(\mathbf{x}) = e^{-\boldsymbol{\kappa}\cdot\mathbf{x}}s(\mathbf{x})$ is able to reproduce the polynomial $\mathbf{x}^{\boldsymbol{\alpha}}$. Consequently, we obtain

$$\mathbf{x}^{\boldsymbol{\alpha}} = \sum_{\boldsymbol{n}} c_{\boldsymbol{\alpha},\boldsymbol{n}} e^{-\boldsymbol{\kappa}\cdot(\mathbf{x}-\boldsymbol{n})} s(\mathbf{x}) \tag{B.1}$$

$$\Leftrightarrow \mathbf{x}^{\boldsymbol{\alpha}} e^{\boldsymbol{\kappa}\cdot\mathbf{x}} = \sum_{\boldsymbol{n}} c_{\boldsymbol{\alpha},\boldsymbol{n}} e^{\boldsymbol{\kappa}\cdot\boldsymbol{n}} s(\mathbf{x}), \tag{B.2}$$

such that $w_{\boldsymbol{n}} = c_{\boldsymbol{\alpha},\boldsymbol{n}} e^{\boldsymbol{\kappa}\cdot\boldsymbol{n}}|_{\boldsymbol{\alpha}=\mathbf{0}}$. Moreover, it follows that (B.2) holds true provided $\psi(\mathbf{x}) = \boldsymbol{n}^{\boldsymbol{\alpha}}\varphi(\mathbf{x}-\boldsymbol{n})$ satisfies the classical Strang-Fix condition with $\boldsymbol{\alpha} = \mathbf{0}$. This allows us to obtain the conditions on $s(\mathbf{x})$ as follows. First we relate the Fourier transform of ψ to the Laplace transform of $s(\mathbf{x})$, which is $\hat{\psi}(\boldsymbol{\omega}) = S(\boldsymbol{\kappa} + \mathbf{s}_{\mathbf{x}})|_{\mathbf{s}_{\mathbf{x}}=\mathrm{j}\boldsymbol{\omega}}$. Therefore when $\boldsymbol{\alpha} = \mathbf{0}$ and

$\boldsymbol{\omega} = 2\pi\boldsymbol{n}$, we obtain the desired conditions:

$$S(\boldsymbol{\kappa}) \neq 0, \text{ and } S(\boldsymbol{\kappa} + \mathrm{j}2\pi\boldsymbol{n}) = 0, \tag{B.3}$$

as required.

B.2. Bilateral Laplace transforms

B.2.1. Green's function of the two-dimensional diffusion equation

We derive expression (6.41), i.e. the multidimensional bilateral Laplace transform (6.23) for the Green's function of the d-dimensional diffusion equation. Recall that in d-dimensions, the diffusion has the following Green's function:

$$g(\mathbf{x}, t) = \frac{1}{(4\pi\mu t)^{d/2}} e^{-\frac{\|\mathbf{x}\|^2}{4\mu t}} H(t), \tag{B.4}$$

and so,

$$G(\mathbf{s_x}, s_t) = \int_{\mathbf{x}\in\mathbb{R}^3} \int_{t\in\mathbb{R}} \frac{1}{(4\pi\mu t)^{d/2}} e^{-\frac{\|\mathbf{x}\|^2}{4\mu t}} H(t) e^{-(\mathbf{x},t)\cdot(\mathbf{s_x},s_t)} \mathrm{d}t\mathrm{d}\mathbf{x}. \tag{B.5}$$

We first consider the spatial integral:

$$\begin{aligned}
\int_{\mathbf{x}\in\mathbb{R}^3} e^{-\frac{\|\mathbf{x}\|^2}{4\mu t}} e^{-\mathbf{x}\cdot\mathbf{s_x}} \mathrm{d}\mathbf{x} &= \int_{\mathbf{x}\in\mathbb{R}^3} e^{-\frac{\|\mathbf{x}\|^2 + 4\mu t \mathbf{x}\cdot\mathbf{s_x}}{4\mu t}} \mathrm{d}\mathbf{x} \\
&= \int_{\mathbf{x}\in\mathbb{R}^3} e^{-\frac{\|\mathbf{x}+2\mu t\mathbf{s_x}\|^2 - 4\mu^2 t^2 \|\mathbf{s_x}\|^2}{4\mu t}} \mathrm{d}\mathbf{x} \\
&= e^{\mu t \|\mathbf{s_x}\|^2} \int_{\mathbf{x}\in\mathbb{R}^3} e^{-\frac{\|\mathbf{x}+2\mu t\mathbf{s_x}\|^2}{4\mu t}} \mathrm{d}\mathbf{x} \\
&= e^{\mu t \|\mathbf{s_x}\|^2} (4\pi\mu t)^{d/2},
\end{aligned} \tag{B.6}$$

where the second equality follows by completing the square, whilst (B.6) follows immediatly from the fact that,

$$\int_{x=-\infty}^{\infty} e^{-\frac{x^2}{a}} \mathrm{d}x = \sqrt{a\pi},$$

and so,

$$\int_{\mathbf{x}} e^{-\frac{\|\mathbf{x}\|^2}{a}} \mathrm{d}\mathbf{x} = \prod_{i=1}^{d} \int_{x_i=-\infty}^{\infty} e^{-\frac{x_i^2}{a}} \mathrm{d}x_i = (\sqrt{a\pi})^d,$$

with $a = 4\mu t$. We can now substitute (B.6) back into (B.5) and proceed as follows:

$$\begin{aligned}
G(\mathbf{s_x}, s_t) &= \int_{t\in\mathbb{R}} e^{\mu t \|\mathbf{s_x}\|^2} H(t) e^{-s_t t} \mathrm{d}t \\
&= \int_{t\geq 0} e^{-(s_t - \mu\|\mathbf{s_x}\|^2)t} \mathrm{d}t \\
&= \frac{1}{s_t - \mu\|\mathbf{s_x}\|^2},
\end{aligned} \tag{B.7}$$

provided $\Re\big(s_t - \mu\|\mathbf{s_x}\|^2\big) > 0$, as required.

Remark B.1. *The requirement that $\Re\big(s_t - \mu\|\mathbf{s_x}\|^2\big) > 0$ is a necessary condition for the integral* (B.7) *to converge. This condition is almost always satisfied in reality. Take for example, $s_t = \rho = \mathrm{j}r/T$ and $\mathbf{s_x} = -\boldsymbol{\kappa} = -\mathrm{j}(k_1, k_2) \Rightarrow \|\mathbf{s_x}\|^2 = -(k_1^2+k_2^2)$. For these choices, $\Re\big(\mathrm{j}r/T + \mu(k_1^2 + k_2^2)\big) > 0 \Rightarrow (k_1^2 + k_2^2) > 0$ which is satisfied for all $(k_1, k_2) \in \mathbb{Z} \setminus \{\mathbf{0}\}$. Similarly for the 3-D ISP, where $\mathbf{s_x} = -\boldsymbol{\kappa} = -(k_1, \mathrm{j}k_1, \mathrm{j}k_2)$, we end up with the condition that $k_2^2 > 0$ and is satisfied for any $k_2 \in \mathbb{Z} \setminus \{0\}$.*

B.2.2. Green's function of the three-dimensional wave equation

We derive the multidimensional bilateral Laplace transform (6.23) of the Green's function for the wave equation (6.5). According to Definition 3.3, this Green's function $g(\mathbf{x}, t)$ must satisfy,

$$\nabla^2 g(\mathbf{x}, t) - \frac{1}{c^2}\frac{\partial^2}{\partial t^2} g(\mathbf{x}, t) = \delta(\mathbf{x}, t). \tag{B.8}$$

We begin by taking the Laplace transform of the PDE above, as follows:

$$\int_{\mathbf{x}\in\mathbb{R}^3}\int_{t\in\mathbb{R}}\left(\nabla^2 g(\mathbf{x}, t) - \frac{1}{c^2}\frac{\partial^2}{\partial t^2} g(\mathbf{x}, t)\right) e^{-(\mathbf{x},t)\cdot(\mathbf{s_x},s_t)}\mathrm{d}t\mathrm{d}\mathbf{x} = \int_{\mathbf{x}\in\mathbb{R}^3}\int_{t\in\mathbb{R}} \delta(\mathbf{x}, t) e^{-(\mathbf{x},t)\cdot(\mathbf{s_x},s_t)}\mathrm{d}t\mathrm{d}\mathbf{x}$$

$$\int_{\mathbf{x}\in\mathbb{R}^3}\left(\nabla^2 - \frac{s_t^2}{c^2}\right) g_1(\mathbf{x}, s_t) e^{-\mathbf{x}\cdot\mathbf{s_x}}\mathrm{d}\mathbf{x} = \int_{\mathbf{x}\in\mathbb{R}^3} \delta(\mathbf{x})\mathrm{d}\mathbf{x}$$

$$\big(\|\mathbf{s_x}\|^2 - (s_t/c)^2\big)\, G(\mathbf{s_x}, s_t) = 1.$$

From this we can rearrange to obtain the required multidimensional bilateral Laplace transform of g,

$$G(\mathbf{s_x}, s_t) = \frac{1}{(\|\mathbf{s_x}\|^2 - (s_t/c)^2)}.$$

B.3. Approximate weights: using Green's second identity

For the family of analytic spatial sensing functions, it can be shown using Green's second identity that (see (3.17)):

$$\mathcal{Q}(k, 1) = \langle f(\mathbf{x}, t), \Gamma(t)\Psi_k(\mathbf{x})\rangle_{\mathbf{x},t} = \int_\Omega (\Psi_k \dot{U})(\mathbf{x}, T)\mathrm{d}V - \mu \oint_{\partial\Omega} (\Psi_k \nabla U - U \nabla \Psi_k)\cdot \hat{\mathbf{n}}_{\partial\Omega}\mathrm{d}S,$$

where $\Gamma(t) = \Gamma_1(t)$, and

$$U(\mathbf{x}, T) = \int_0^T \Gamma(t) u(\mathbf{x}, t)\mathrm{d}t, \quad \text{and} \tag{B.9a}$$

$$\dot{U}(\mathbf{x}, T) = \Gamma(T) u(\mathbf{x}, T) - \int_0^T \frac{\partial \Gamma}{\partial t} u(\mathbf{x}, t)\mathrm{d}t. \tag{B.9b}$$

Hence $\mathcal{R}(k)$ can be computed exactly if $u(\mathbf{x}, t)$, $\Psi_k(\mathbf{x})$ and $\Gamma(t)$ are known over $\Omega \times [0, T]$. However, given real spatiotemporal sensor measurements of the field $u(\mathbf{x}, t)$, we can only

compute approximations of the quantities $U(\mathbf{x}, T)$ and $\dot{U}(\mathbf{x}, T)$ using quadrature schemes. These approximations, of the integral (3.17), can in general be written as a weighted sum of the field spatiotemporal measurements $\{\varphi_n(t_l)\}_{n,l}$, for $n = 1, \ldots, N$ and $l = 0, \ldots, L$. Thus,

$$\mathcal{R}(k) = \sum_{n=1}^{N} \sum_{l=0}^{L} w_{n,l}(k) \varphi_n(t_l). \tag{B.10}$$

To derive the *Green's weights*, we consider (3.17) and firstly consider approximating the time integrals for the n-th sensor and then combining these over all N sensors in the network.

To begin, let us denote the approximate values of the quantities (B.9a) and (B.9b)—to be obtained from the from temporal samples $\{\varphi_n(t_l)\}_{l=0}^{L}$ of the n-th sensor—at the location $\mathbf{x} = \mathbf{x}_n$ by $\Phi_n(t_L)$ and $\dot{\Phi}_n(t_L)$ respectively. Then for (B.9a):

$$\begin{aligned} U(\mathbf{x}_n, T)|_{T=t_L} &= \int_0^{t_L} \Gamma(t) u(\mathbf{x}, t) \mathrm{d}t \\ &\approx \Phi_n(t_L) = \frac{\Delta_t t}{2} \left(u(\mathbf{x}_n, t_0)\Gamma(t_0) + 2 \sum_{l=1}^{L-1} u(\mathbf{x}_n, t_l)\Gamma(t_l) + u(\mathbf{x}_n, t_L)\Gamma(t_L) \right) \end{aligned} \tag{B.11}$$

$$= \frac{\Delta_t t}{2} \left(\varphi_n(t_0) + 2 \sum_{l=1}^{L-1} \varphi_n(t_l) e^{-\mathrm{j}l/L} + \varphi_n(t_L) e^{-\mathrm{j}} \right). \tag{B.12}$$

Following similar lines for (B.9b), and noting that $\Gamma(t)|_{T=t_L} = e^{-\mathrm{j}t/t_L}$ and $\frac{\partial \Gamma(t)}{\partial t}\Big|_{T=t_L} = \frac{1}{\mathrm{j}t_L} e^{-\mathrm{j}t/t_L}$:

$$\begin{aligned} \dot{U}(\mathbf{x}_n, T)\Big|_{T=t_L} =& \Gamma(t_L) u(\mathbf{x}, t_L) - \int_0^{t_L} \frac{\partial \Gamma}{\partial t} u(\mathbf{x}, t) \mathrm{d}t \\ \approx& \dot{\Phi}_n(t_L) \\ =& u(\mathbf{x}_n, t_L) e^{-\mathrm{j}} - \frac{\Delta_t t}{\mathrm{j}2t_L} \left(e^{-\mathrm{j}t_0/t_L} u(\mathbf{x}_n, t_0) + 2 \sum_{l=1}^{L-1} e^{-\mathrm{j}t_l/t_L} u(\mathbf{x}_n, t_l) + e^{-\mathrm{j}} u(\mathbf{x}_n, t_L) \right) \end{aligned} \tag{B.13}$$

$$= \frac{\mathrm{j}\Delta_t t}{2t_L} \varphi_n(t_0) + \frac{\mathrm{j}\Delta_t t}{t_L} \sum_{l=1}^{L} e^{-\mathrm{j}l/L} \varphi_n(t_l) + \left(1 + \frac{\mathrm{j}\Delta_t t}{2t_L} \right) e^{-\mathrm{j}} \varphi_n(t_L). \tag{B.14}$$

Moreover, considering both space and time integrals together, it is possible to show that,

$$\begin{aligned} \mathcal{Q}(k, 1) &\approx \sum_{n \in \mathcal{S}} A_n(k) \dot{\Phi}_n(t_L) - \mu B_n(k) \Phi_n(t_L) \\ &= \frac{1}{N} \sum_{n \in \mathcal{S}} y_n(k). \end{aligned}$$

where $\{A_n(k)\}_n$ and $\{B_n(k)\}_n$ are constants, for each $k \in \mathbb{Z}$, that depend on the topology

of the sensor network. We can now substitute into the above, expressions (B.12) and (B.14) for $\Phi_n(t_L)$ and $\dot{\Phi}_n(t_L)$ respectively to get,

$$\begin{aligned} y_n(k) =& NA_n(k)\left(\frac{\mathrm{j}\Delta_t t}{2t_L}\varphi_n(t_0) + \frac{\mathrm{j}\Delta_t t}{t_L}\sum_{l=1}^{L} e^{-\mathrm{j}l/L}\varphi_n(t_l) + \left(1+\frac{\mathrm{j}\Delta_t t}{2t_L}\right)e^{-\mathrm{j}}\varphi_n(t_L)\right) \\ &- \frac{\mu N B_n(k)\Delta_t t}{2}\left(\varphi_n(t_0) + 2\sum_{l=1}^{L-1}\varphi_n(t_l)e^{-\mathrm{j}l/L} + \varphi_n(t_L)e^{-\mathrm{j}}\right) \\ =& \frac{\Delta_t t}{2t_L}\left(\mathrm{j}NA_n(k) - \mu N B_n(k) t_L\right)\varphi_n(t_0) + \frac{\Delta_t t}{t_L}\left(\mathrm{j}NA_n(k) - \mu N B_n(k) t_L\right)\sum_{l=1}^{L} e^{-\mathrm{j}l/L}\varphi_n(t_l) \\ &+ \left(NA_n(k)\big(1+\frac{\mathrm{j}\Delta_t t}{2t_L}\big) - \frac{\mu N B_n(k)\Delta_t t}{2}\right)e^{-\mathrm{j}}\varphi_n(t_L). \end{aligned} \tag{B.15}$$

Hence, the approximate weights using Green's second identity, i.e. the "Green's weights", are given by:

$$w_{n,l}(k) = \begin{cases} \frac{\mathrm{j}NA_n(k)\Delta_t t}{2t_L} - \frac{\mu N B_n(k)\Delta_t t}{2}, & \text{if } l = 0 \\ \left(\frac{\mathrm{j}NA_n(k)\Delta_t t}{t_L} - \mu N B_n(k)\Delta_t t\right)e^{-\mathrm{j}l/L}, & \text{if } l \in \{1,\ldots,L-1\} \\ \left(NA_n(k)\left(1+\frac{\mathrm{j}\Delta_t t}{2t_L}\right) - \frac{\mu N B_n(k)\Delta_t t}{2}\right)e^{-\mathrm{j}}, & \text{if } l = L \end{cases} \tag{B.16}$$

for any $n \in \mathcal{N}$.

Bibliography

[1] B. Abdelaziz, A. El Badia, and A. El Hajj. Direct algorithms for solving some inverse source problems in 2D elliptic equations. *Inverse Problems*, 31(10):105002, Sep 2015.

[2] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci. A survey on sensor networks. *IEEE Communications Magazine*, 40(8):102–114, Aug 2002.

[3] L. Albera, A. Ferréol, D. Cosandier-Rimélé, I. Merlet, and F. Wendling. Brain source localization using a fourth-order deflation scheme. *IEEE Transactions on Biomedical Engineering*, 55(2):490–501, 2008.

[4] L. Albera, S. Kitić, N. Bertin, G. Puy, and R. Gribonval. Brain source localization using a physics-driven structured cosparse representation of EEG signals. In *2014 IEEE International Workshop on Machine Learning for Signal Processing (MLSP)*, pages 1–6, Sep 2014.

[5] S. Aldalahmeh and M. Ghogho. Robust distributed detection, localization, and estimation of a diffusive target in clustered wireless sensor networks. In *2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 3012–3015, May 2011.

[6] S. Aldalahmeh, M. Ghogho, and A. Swami. Fast distributed detection, localization, and estimation of a diffusive target in wireless sensor networks. In *7th International Symposium on Wireless Communication Systems (ISWCS 2010)*, pages 882–886, Sept 2010.

[7] A. Aldroubi. Non-uniform weighted average sampling and reconstruction in shift-invariant and wavelet spaces. *Applied and Computational Harmonic Analysis*, 13(2): 151 – 161, 2002.

[8] A. Aldroubi and K. Gröchenig. Nonuniform sampling and reconstruction in shift-invariant spaces. *SIAM Review*, 43(4):585–620, 2001.

[9] A. Aldroubi and M. Unser. Sampling procedures in function spaces and asymptotic equivalence with Shannon's sampling theory. *Numerical Functional Analysis and Optimization*, 15(1-2):1–21, 1994.

[10] O. M. Alifanov. *Inverse heat transfer problems*. Springer Science & Business Media, 2012.

[11] S. Andrieux and A. B. Abda. Identification of planar cracks by complete overdetermined data: inversion formulae. *Inverse Problems*, 12(5):553, 1996.

[12] N. Antonello, T. van Waterschoot, M. Moonen, and P. A. Naylor. Source localization and signal reconstruction in a reverberant field using the FDTD method. In *Proceedings of the 22nd European on Signal Processing Conference (EUSIPCO'14)*, pages 301–305, Sept 2014.

[13] F. Asano, M. Goto, K. Itou, and H. Asoh. Real-time sound source localization and separation system and its application to automatic speech recognition. In *INTERSPEECH*, pages 1013–1016, 2001.

[14] N. Auffray, M. Bonnet, and S. Pagano. Identification of transient heat sources using the reciprocity gap. *Inverse Problems in Science and Engineering*, 21(4):721–738, 2013.

[15] T. C. Aysal, M. E. Yildiz, A. D. Sarwate, and A. Scaglione. Broadcast gossip algorithms for consensus. *IEEE Transactions on Signal Processing*, 57(7):2748–2761, 2009.

[16] S. Baillet, J. C. Mosher, and R. M. Leahy. Electromagnetic brain mapping. *IEEE Signal Processing Magazine*, 18(6):14–30, Nov 2001.

[17] T. Bannour, A. B. Abda, and M. Jaoua. A semi-explicit algorithm for the reconstruction of 3D planar cracks. *Inverse Problems*, 13(4):899, 1997.

[18] H. Becker, L. Albera, P. Comon, R. Gribonval, F. Wendling, and I. Merlet. Brain-source imaging: From sparse to tensor models. *IEEE Signal Processing Magazine*, 32 (6):100–112, Nov 2015.

[19] F. Bénézit, A. Dimakis, P. Thiran, and M. Vetterli. Gossip along the way: Order-optimal consensus through randomized path averaging. In *Proceedings of the Allerton Conference on Communication, Control, and Computing*, 2007.

[20] E. R. Benton and G. W. Platzman. A table of solutions of the one-dimensional burgers equation. *Quarterly of Applied Mathematics*, pages 195–212, 1972.

[21] A. Beurling. The collected works of Arne Beurling. In L. Carleson, P. Malliavin, J. Neuberger, and J. Wermer, editors, *Complex analysis*, volume 1. Birkhäuser, Boston, 1989.

[22] B. N. Bhaskar, G. Tang, and B. Recht. Atomic norm denoising with applications to line spectral estimation. *IEEE Transactions on Signal Processing*, 61(23):5987–5999, 2013.

[23] R. Bianchini and R. Rajamony. Power and energy management for server systems. *IEEE Computer*, 37(11):68–76, November 2004.

[24] T. Blu and M. Unser. Approximation error for quasi-interpolators and (multi-) wavelet expansions. *Applied and Computational Harmonic Analysis*, 6(2):219–251, 1999.

[25] T. Blu and M. Unser. Quantitative Fourier analysis of approximation techniques. Part I—Interpolators and projectors. *IEEE Transactions on Signal Processing*, 47 (10):2783–2795, Oct 1999.

[26] T. Blu, P. L. Dragotti, M. Vetterli, P. Marziliano, and L. Coulot. Sparse sampling of signal innovations. *Signal Processing Magazine, IEEE*, 25(2):31–40, March 2008.

[27] S. Boyd, A. Ghosh, B. Prabhakar, and D. Shah. Randomized gossip algorithms. *IEEE Transactions on Information Theory*, 52(6):2508–2530, 2006.

[28] P. Braca, S. Marano, and V. Matta. Enforcing consensus while monitoring the environment in wireless sensor networks. *IEEE Transactions on Signal Processing*, 56 (7):3375–3380, July 2008.

[29] S. C. Brenner and R. Scott. *The mathematical theory of finite element methods*, volume 15. Springer, 2008.

[30] D. Brooks, R. P. Dick, R. Joseph, and L. Shang. Power, thermal, and reliability modeling in nanometer-scale microprocessors. *Micro, IEEE*, 27(3):49–62, May 2007.

[31] R. A. Brooks and G. D. Chiro. Principles of computer assisted tomography (CAT) in radiographic and radioisotopic imaging. *Physics in Medicine and Biology*, 21(5): 689, 1976.

[32] J. L. Brown and S. D. Cabrera. On well-posedness of the Papoulis generalized sampling expansion. *IEEE Transactions on Circuits and Systems*, 38(5):554–556, May 1991.

[33] J. L. Brown and K. Sa-ngsari. Sampling reconstruction of n-dimensional band-limited images after multilinear filtering. *IEEE Transactions on Circuits and Systems*, 36(7): 1035–1038, Jul 1989.

[34] R. Brown. A brief account of microscopical observations made in the months of June, July and August 1827, on the particles contained in the pollen of plants; and on the general existence of active molecules in organic and inorganic bodies. *Philosophical Magazine Series 2*, 4(21):161–173, 1828.

[35] J. A. Cadzow. Signal enhancement — a composite property mapping algorithm. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 36(1):49–62, Jan 1988.

[36] J. R. Cannon and P. DuChateau. An inverse problem for a nonlinear diffusion equation. *SIAM Journal on Applied Mathematics*, 39(2):272–289, 1980.

[37] J. R. Cannon and S. P. Esteva. An inverse problem for the heat equation. *Inverse Problems*, 2(4):395, 1986.

[38] J. R. Cannon and S. Perez-Esteva. Uniqueness and stability of 3D heat sources. *Inverse Problems*, 7(1):57, 1991.

[39] V. Chandrasekaran, B. Recht, P. A. Parrilo, and A. S. Willsky. The convex geometry of linear inverse problems. *Foundations of Computational mathematics*, 12(6):805–849, 2012.

[40] C. Chen, P. Marziliano, and A. C. Kot. 2D finite rate of innovation reconstruction method for step edge and polygon signals in the presence of noise. *IEEE Transactions on Signal Processing*, 60(6):2851–2859, 2012.

[41] J. C. Chen, K. Yao, and R. E. Hudson. Source localization and beamforming. *IEEE Signal Processing Magazine*, 19(2):30–39, 2002.

[42] J. C. Chen, K. Yao, and R. E. Hudson. Acoustic source localization and beamforming: theory and practice. *EURASIP Journal on Advances in Signal Processing*, 2003(4): 1–12, 2003.

[43] P. Chevalier, A. Ferreol, and L. Albera. High-resolution direction finding from higher order statistics: The $2q$-music algorithm. *IEEE Transactions on Signal Processing*, 54(8):2986–2997, Aug 2006.

[44] M. Chino, H. Nakayama, H. Nagai, H. Terada, G. Katata, and H. Yamazawa. Preliminary estimation of release amounts of 131I and 137Cs accidentally discharged from the Fukushima Daiichi nuclear power plant into the atmosphere. *Journal of Nuclear Science and Technology*, 48(7):1129–1134, 2011.

[45] O. Christensen. Moment problems and stability results for frames with applications to irregular sampling and Gabor frames. *Applied and Computational Harmonic Analysis*, 3(1):82–86, 1996.

[46] G. Dahlquist and Å. Björck. *Numerical Methods in Scientific Computing: Vol. 1*. SIAM e-books. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2008.

[47] W. Dai and O. Milenkovic. Subspace pursuit for compressive sensing signal reconstruction. *IEEE Transactions on Information Theory*, 55(5):2230–2249, May 2009.

[48] A. M. Dale, A. K. Liu, B. R. Fischl, R. L. Buckner, J. W. Belliveau, J. D. Lewine, and E. Halgren. Dynamic statistical parametric mapping: combining fMRI and MEG for high-resolution imaging of cortical activity. *Neuron*, 26(1):55–67, 2000.

[49] P. J. Davis. Triangle formulas in the complex plane. *Mathematics of Computation*, 18(88):pp. 569–577, 1964.

[50] P. J. Davis. Plane regions determined by complex moments. *Journal of Approximation Theory*, 19(2):148–153, 1977.

[51] P. J. Davis and R. Hersh. *The mathematical experience.* Houghton Mifflin Harcourt, 1998.

[52] M. de Berg, O. Cheong, M. van Kreveld, and M. Overmars. *Computational Geometry: Algorithms and Applications.* Springer-Verlag, Santa Clara, CA, USA, 3rd ed. edition, 2008.

[53] C. De Boor, R. A. DeVore, and A. Ron. Approximation from shift-invariant subspaces of $l_2(\mathbb{R}^2)$. *Transactions of the American Mathematical Society*, 341(2):787–806, 1994.

[54] G. R. de Prony. Essai expérimental et analytique: sur les lois de la dilatabilité des fluides élastiques et sur celles de la force expansive de la vapeur de l'alkool, à différentes températures. *Journal de l'école Polytechnique*, 1, 1795.

[55] A. G. Dimakis, S. Kar, J. M. F. Moura, M. G. Rabbat, and A. Scaglione. Gossip algorithms for distributed signal processing. *Proceedings of the IEEE*, 98(11):1847–1864, Nov 2010.

[56] Z. Dogan, V. Tsiminaki, I. Jovanovic, T. Blu, and D. Van De Ville. Localization of point sources for systems governed by the wave equation. In *Proc. of the SPIE*, volume 8138, pages 81380P–1, 2011.

[57] Z. Doğan, T. Blu, and D. Van De Ville. Finite-rate-of-innovation for the inverse source problem of radiating fields. *Sampling Theory in Signal & Image Processing*, 13(3), 2014.

[58] I. Dokmanić. *Listening to Distances and Hearing Shapes: Inverse Problems in Room Acoustics and Beyond.* PhD thesis, École Polytechnique Fédérale de Laussane, 2015.

[59] I. Dokmanic, J. Ranieri, A. Chebira, and M. Vetterli. Sensor networks for diffusion fields: Detection of sources in space and time. In *49th Allerton Conference on Communication, Control, and Computing (Allerton'09)*, pages 1552–1558, Sept 2011.

[60] I. Dokmanić, L. Daudet, and M. Vetterli. From acoustic room reconstruction to SLAM. In *2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'16)*, pages 6345–6349. IEEE, 2016.

[61] P. L. Dragotti, M. Vetterli, and T. Blu. Sampling moments and reconstructing signals of finite rate of innovation: Shannon meets Strang-Fix. *IEEE Transactions on Signal Processing*, 55(5):1741–1757, 2007.

[62] D. G. Duffy. *Greens functions with applications.* Chapman and Hall/CRC Press, 2001.

[63] B. A. Egan and J. R. Mahoney. Numerical modeling of advection and diffusion of urban area source pollutants. *Journal of Applied Meteorology*, 11(2):312–322, 1972.

[64] M. Elad, P. Milanfar, and G. Golub. Shape from moments - an estimation theory perspective. *IEEE Transactions on Signal Processing*, 52(7):1814–1829, July 2004.

[65] H. C. Elman, D. J. Silvester, and A. J. Wathen. *Finite Elements and Fast Iterative Solvers: with Applications in Incompressible Fluid Dynamics.* Oxford University Press, 2005.

[66] A. Farcas and D. Lesnic. The boundary-element method for the determination of a heat source dependent on one variable. *Journal of Engineering Mathematics*, 54(4): 375–388, 2006.

[67] P. Frasca, R. Carli, F. Fagnani, and S. Zampieri. Average consensus on networks with quantized communication. *International Journal of Robust and Nonlinear Control*, 19(16):1787–1816, 2009.

[68] R. A. Freeze and J. A. Cherry. Groundwater, 1979.

[69] H. K. French, S. E. A. T. M. Van der Zee, and A. Leijnse. Prediction uncertainty of plume characteristics derived from a small number of measuring points. *Hydrogeology Journal*, 8(2):188–199, 2000.

[70] K. Friston, L. Harrison, J. Daunizeau, S. Kiebel, C. Phillips, N. Trujillo-Barreto, R. Henson, G. Flandin, and J. Mattout. Multiple sparse priors for the M/EEG inverse problem. *NeuroImage*, 39(3):1104–1120, 2008.

[71] M. Gastpar, M. Vetterli, and P. L. Dragotti. Sensing reality and communicating bits: a dangerous liaison. *Signal Processing Magazine, IEEE*, 23(4):70–83, 2006.

[72] K. Georg. Approximation of integrals for boundary element methods. *SIAM Journal on Scientific and Statistical Computing*, 12(2):443–453, 1991.

[73] R. Grech, T. Cassar, J. Muscat, K. P. Camilleri, S. G. Fabri, M. Zervakis, P. Xanthopoulos, V. Sakkalis, and B. Vanrumste. Review on solving the inverse problem in EEG source analysis. *Journal of NeuroEngineering and Rehabilitation*, 5(1):1–33, 2008.

[74] E. Hopf. The partial differential equation $u_t + uu_x = \mu_{xx}$. *Communications on Pure and Applied mathematics*, 3(3):201–230, 1950.

[75] Y. Hua and T. K. Sarkar. Matrix pencil method for estimating parameters of exponentially damped/undamped sinusoids in noise. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 38(5):814–824, May 1990.

[76] F. Iutzeler, P. Ciblat, and W. Hachem. Analysis of sum-weight-like algorithms for averaging in wireless sensor networks. *IEEE Transactions on Signal Processing*, 61 (11):2802–2814, 2013.

[77] A. J. Jerri. The Shannon sampling theorem—its various extensions and applications: A tutorial review. *Proceedings of the IEEE*, 65(11):1565–1596, Nov 1977.

[78] I. Jovanović, L. Sbaiz, and M. Vetterli. Acoustic tomography for scalar and vector fields: theory and application to temperature and wind estimation. *Journal of Atmospheric and Oceanic Technology*, 26(8):1475–1492, 2009.

[79] M. I. Kadec. Exact value of Paley-Wiener constant. *Doklady Akademii Nauk SSSR*, 155(6):1253, 1964.

[80] D. Kandaswamy, T. Blu, and D. Van De Ville. Analytic sensing: Noniterative retrieval of point sources from boundary measurements. *SIAM Journal on Scientific Computing*, 31(4):3179–3194, 2009.

[81] M. Karalashvili, S. Groß, W. Marquardt, A. Mhamdi, and A. Reusken. Identification of transport coefficient models in convection-diffusion equations. *SIAM Journal on Scientific Computing*, 33(1):303–327, 2011.

[82] A. Kashyap, T. Başar, and R. Srikant. Quantized consensus. *Automatica*, 43(7): 1192–1203, 2007.

[83] I. Khalidov, T. Blu, and M. Unser. Generalized L-spline wavelet bases. In *Proceedings of the SPIE Conference on Mathematical Imaging: Wavelet XI*, pages 1–8. SPIE, Aug 2005.

[84] S. Kitić, L. Albera, N. Bertin, and R. Gribonval. Physics-driven inverse problems made tractable with cosparse regularization. *IEEE Transactions on Signal Processing*, 64(2):335–348, Jan 2016.

[85] V. A. Kotel'nikov. On the transmission capacity of the ether and wire in telecommunications. In *Material for the First All-Union Conference on Questions of Communication, Izd. Red. Upr. Svyazi RKKA, Moscow*, volume 1, 1933.

[86] A. Kumar, P. Ishwar, and K. Ramchandran. On distributed sampling of smooth non-bandlimited fields. In *Proceedings of the 3rd international symposium on Information processing in sensor networks*, pages 89–98. ACM, 2004.

[87] S. Kumar, F. Zhao, and D. Shepherd. Collaborative signal and information processing in microsensor networks. *Signal Processing Magazine, IEEE*, 19(2):13–14, 2002.

[88] J. Kvasnica. The inverse problem for diffusion equations. *Czechoslovak Journal of Physics B*, 35(5):490–493, 1985.

[89] H. J. Landau. Necessary density conditions for sampling and interpolation of certain entire functions. *Acta Mathematica*, 117(1):37–52, 1967.

[90] K. Langendoen, A. Baggio, and O. Visser. Murphy loves potatoes: experiences from a pilot sensor network deployment in precision agriculture. In *Proc. 20th International Parallel and Distributed Processing Symposium (IPDPS'06)*, page 8, April 2006.

[91] C. Le Niliot and F. Lefevre. Multiple transient point heat sources identification in heat diffusion: application to numerical two-and three-dimensional problems. *Numerical Heat Transfer: Part B: Fundamentals*, 39(3):277–301, 2001.

[92] C. Le Niliot, F. Rigollet, and D. Petit. An experimental identification of line heat sources in a diffusive system using the boundary element method. *International Journal of Heat and Mass Transfer*, 43(12):2205 – 2220, 2000.

[93] W. A. Light and E. W. Cheney. Quasi-interpolation with translates of a function having noncompact support. *Constructive Approximation*, 8(1):35–48, 1992.

[94] T. Limpiti, B. D. Van Veen, and R. T. Wakai. Cortical patch basis model for spatially extended neural activity. *IEEE Transactions on Biomedical Engineering*, 53(9):1740–1754, 2006.

[95] F.-H. Lin, J. W. Belliveau, A. M. Dale, and M. S. Hämäläinen. Distributed current estimates using cortical orientation constraints. *Human brain mapping*, 27(1):1–13, 2006.

[96] C. G. Lopes and A. H. Sayed. Diffusion least-mean squares over adaptive networks: Formulation and performance analysis. *IEEE Transactions on Signal Processing*, 56 (7):3122–3136, 2008.

[97] Y. M. Lu and M. Vetterli. Distributed spatio-temporal sampling of diffusion fields from sparse instantaneous sources. In *3rd IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP'09)*, pages 205–208, Aruba, Dutch Antilles, 2009. IEEE.

[98] Y. M. Lu and M. Vetterli. Spatial super-resolution of a diffusion field by temporal oversampling in sensor networks. In *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'09)*, pages 2249–2252. IEEE, 2009.

[99] Y. M. Lu, P. L. Dragotti, and M. Vetterli. Localizing point sources in diffusion fields from spatiotemporal samples. In *Proc. 9th International Conference on Sampling Theory and Applications (SampTa'11)*, Singapore, 2011.

[100] Y. M. Lu, P. L. Dragotti, and M. Vetterli. Localization of diffusive sources using spatiotemporal measurements. In *Proc. 9th IEEE Annual Allerton Conference on Communication, Control, and Computing (Allerton'11)*, pages 1072–1076, Illinois, USA, 2011. IEEE.

[101] I. Malyshev. An inverse source problem for heat equation. *Journal of Mathematical Analysis and Applications*, 142(1):206 – 218, 1989.

[102] S. Marano, V. Matta, and P. Willett. Distributed estimation in large wireless sensor networks via a locally optimum approach. *IEEE Transactions on Signal Processing*, 56(2):748–756, 2008.

[103] S. Maranò, D. Fäh, and H.-A. Loeliger. A state-space approach for the analysis of wave and diffusion fields. In *Proc. 40th IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'15)*, Brisbane, Australia, apr 2015.

[104] I. Maravić and M. Vetterli. Exact sampling results for some classes of parametric nonbandlimited 2-D signals. *IEEE Transactions on Signal Processing*, 52(1):175–189, Jan 2004.

[105] M. Martinez-Camara, I. Dokmanic, J. Ranieri, R. Scheibler, M. Vetterli, and A. Stohl. The Fukushima inverse problem. In *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*, pages 4330–4334, May 2013.

[106] K. Matsuura and Y. Okabe. A robust reconstruction of sparse biomagnetic sources. *IEEE Transactions on Biomedical Engineering*, 44(8):720–726, Aug 1997.

[107] J. Matthes, L. Gröll, and H. B. Keller. Source localization by spatially distributed electronic noses for advection and diffusion. *IEEE Transactions on Signal Processing*, 53(5):1711–1719, May 2005.

[108] S. McKay. *Fundamentals of statistical signal processing: Estimation Theory*, volume I of *Signal Processing*. Prentice Hall PTR, Upper Saddle River, New Jersey, 2011.

[109] R. C. McOwen. *Partial Differential Equations: Methods and Applications*. Featured Titles for Partial Differential Equations Series. Prentice Hall, 2003.

[110] S. Nam and R. Gribonval. Physics-driven structured cosparse modeling for source localization. In *Acoustics, Speech and Signal Processing (ICASSP), 2012 IEEE International Conference on*, pages 5397–5400, March 2012.

[111] F. Natterer. *The mathematics of computerized tomography*, volume 32. SIAM, 1986.

[112] A. Nehorai, B. Porat, and E. Paldi. Detection and localization of vapor-emitting sources. *IEEE Transactions on Signal Processing*, 43(1):243–253, Jan 1995.

[113] H. Nyquist. Certain factors affecting telegraph speed. *Transactions of the American Institute of Electrical Engineers*, XLIII:412–422, Jan 1924.

[114] H. Nyquist. Certain topics in telegraph transmission theory. *Transactions of the American Institute of Electrical Engineers*, 47(2):617–644, 1928.

[115] J. Oñativia, S. R. Schultz, and P. L. Dragotti. A finite rate of innovation algorithm for fast and accurate spike detection from two-photon calcium imaging. *Journal of neural engineering*, 10(4):046017, 2013.

[116] J. Oñativia, J. A. Uriguen, and P. L. Dragotti. Sequential local FRI sampling of infinite streams of diracs. In *2013 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP),*, pages 5440–5444, 2013.

[117] H. Pan, T. Blu, and P. L. Dragotti. Sampling curves with finite rate of innovation. *IEEE Transactions on Signal Processing*, 62(2):458–471, Jan 2014. ISSN 1053-587X.

[118] A. Papoulis. Generalized sampling expansion. *IEEE transactions on circuits and systems*, 24(11):652–654, 1977.

[119] J. A. Parker, R. V. Kenyon, and D. E. Troxel. Comparison of interpolating methods for image resampling. *IEEE Transactions on medical imaging*, 2(1):31–39, 1983.

[120] E. Parzen. A simple proof and some extensions of the sampling theorem. Technical report, DTIC Document, 1956.

[121] R. D. Pascual-Marqui. Standardized low-resolution brain electromagnetic tomography (sLORETA): technical details. *Methods Findings Experimental Clinical Pharmacology*, 24(Suppl D):5–12, 2002.

[122] T. Piotrowski, D. Gutierrez, I. Yamada, and J. Żygierewicz. Reduced-rank neural activity index for EEG/MEG multi-source localization. In *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'14)*, pages 4708–4712. IEEE, 2014.

[123] J. Ranieri and M. Vetterli. Sampling and reconstructing diffusion fields in presence of aliasing. In *Proc. IEEE International Conference on Acoustics, Speech, And Signal Processing (ICASSP'13)*, Vancouver, Canada, May 2013.

[124] J. Ranieri, Y. Lu, A. Chebira, and M. Vetterli. Sampling and Reconstructing Diffusion Fields with Localized Sources. In *Proc. IEEE International Conference on Acoustics, Speech, And Signal Processing (ICASSP'11)*, pages 4016–4019, Prague, Czech Republic, May 2011. IEEE Service Center, NJ, USA.

[125] J. Ranieri, I. Dokmanic, A. Chebira, and M. Vetterli. Sampling and Reconstruction of Time-Varying Atmospheric Emissions. In *Proc. 37th IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP'12)*, Kyoto, Japan, Mar 2012. IEEE. ISBN 978-1-4673-0046-9.

[126] J. Ranieri, A. Vincenzi, A. Chebira, D. A. Alonso, and M. Vetterli. EigenMaps: Algorithms for Optimal Thermal Maps Extraction and Sensor Placement on Multicore

Processors. In *Proceedings of the 49th Design Automation Conference (DAC'12)*, Design Automation Conference (DAC'12), pages 636–641, San Francisco, 2012. ACM. ISBN 978-1-4503-1199-1.

[127] G. Reise and G. Matz. Clustered wireless sensor networks for robust distributed field reconstruction based on hybrid shift-invariant spaces. In *IEEE 10th Workshop on Signal Processing Advances in Wireless Communications (SPAWC'09)*, pages 66–70, 2009.

[128] G. Reise and G. Matz. Distributed sampling and reconstruction of non-bandlimited fields in sensor networks based on shift-invariant spaces. In *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'09)*, pages 2061–2064, Taipei, Taiwan, 2009. IEEE.

[129] G. Reise, G. Matz, and K. Grochenig. Distributed field reconstruction in wireless sensor networks based on hybrid shift-invariant spaces. *IEEE Transactions on Signal Processing*, 60(10):5426–5439, 2012.

[130] G. R. Richter. An inverse problem for the steady state diffusion equation. *SIAM Journal on Applied Mathematics*, 41(2):210–221, 1981.

[131] M. Rostami, N.-M. Cheung, and T. Q. S. Quek. Compressed sensing of diffusion fields under heat equation constraint. In *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'13)*, pages 4271–4274, Vancouver, Canada, May 2013.

[132] F. Sawo, K. Roberts, and U. D. Hanebeck. Bayesian estimation of distributed phenomena using discretized representations of partial differential equations. In *3rd International Conference on Informatics in Control, Automation and Robotics (ICINCO'06)*, pages 16–23. IEEE, Aug 2006.

[133] A. Schmidt and J. M. F. Moura. A distributed sensor fusion algorithm for the inversion of sparse fields. In *2009 Conference Record of the Forty-Third Asilomar Conference on Signals, Systems and Computers (Asolimar 2009)*, pages 1332–1336, Nov 2009.

[134] C. Shackelford and D. Daniel. Diffusion in saturated soil I: Background. *Journal of Geotechnical Engineering*, 117(3):467–484, 1991.

[135] C. E. Shannon. Communication in the presence of noise. *Proceedings of the IRE*, 37 (1):10–21, 1949.

[136] P. Shukla and P. L. Dragotti. Sampling schemes for multidimensional signals with finite rate of innovation. *IEEE Transactions on Signal Processing*, 55(7):3670–3686, July 2007.

[137] D. Slepian. On bandwidth. *Proceedings of the IEEE*, 64(3):292–300, March 1976.

[138] R. L. Stens. Error estimates for sampling sums based on convolution integrals. *Information and Control*, 45(1):37 – 47, 1980.

[139] J. M. Stockie. The mathematics of atmospheric dispersion modeling. *Siam Review*, 53(2):349–372, 2011.

[140] P. Stoica and R. L. Moses. *Introduction to spectral analysis*. Prentice Hall, Englewood Cliffs, NJ, 1997.

[141] G. Strang and G. Fix. A Fourier analysis of the finite element variational method. In G. Geymonat, editor, *Constructive Aspects of Functional Analysis*, pages 793–840. C.I.M.E., 1973.

[142] J. N. Tsitsiklis. *Problems in decentralized decision making and computation*. PhD thesis, Massachusetts Institute of Technology, 1984.

[143] M. Unser. Sampling-50 years after Shannon. *Proceedings of the IEEE*, 88(4):569–587, April 2000.

[144] M. Unser and A. Aldroubi. A general sampling theory for nonideal acquisition devices. *IEEE Transactions on Signal Processing*, 42(11):2915–2925, Nov 1994.

[145] J. A. Uriguen, T. Blu, and P. L. Dragotti. FRI sampling with arbitrary kernels. *IEEE Transactions on Signal Processing*, 61(21):5310–5323, Nov 2013.

[146] B. D. van Veen, W. van Drongelen, M. Yuchtman, and A. Suzuki. Localization of brain electrical activity via linearly constrained minimum variance spatial filtering. *IEEE Transactions on biomedical engineering*, 44(9):867–880, 1997.

[147] T. van Waterschoot and G. Leus. Static field estimation using a wireless sensor network based on the finite element method. In *4th IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP'11)*, pages 369–372, Dec 2011.

[148] T. van Waterschoot and G. Leus. Distributed estimation of static fields in wireless sensor networks using the finite element method. In *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'12)*, pages 2853–2856, Kyoto, Japan, March 2012.

[149] F. Vanpoucke, M. Moonen, and Y. Berthoumieu. An efficient subspace algorithm for 2-D harmonic retrieval. In *1994 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP'94)*, pages IV/461–IV/464, Apr 1994.

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

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

[152] K. Watanabe. *Integral transform techniques for green's function*, volume 71 of *Lecture Notes in Applied and Computational Mechanics.* Springer International Publishing, Switzerland, 2014.

[153] K. Watanabe. Green's functions for laplace and wave equations. In *Integral Transform Techniques for Green's Function*, volume 76 of *Lecture Notes in Applied and Computational Mechanics*, pages 33–76. Springer International Publishing, 2015.

[154] J. Weimer, B. Sinopoli, and B. H. Krogh. Multiple source detection and localization in advection-diffusion processes using wireless sensor networks. In *Proc. 30th IEEE Real-Time Systems Symposium (RTSS'09)*, pages 333–342, Washington DC, USA, 2009. IEEE.

[155] J. M. Whittaker. The Fourier theory of the cardinal function. *Proceedings of the Edinburgh Mathematical Society (Series 2)*, 1(03):169–176, 1928.

[156] N. Wiener and R. E. A. C. Paley. Fourier transforms in the complex domain. *Am. Math. Soc. Coll. Pub.*, 19, 1934.

[157] D. Wipf and S. Nagarajan. A unified Bayesian framework for MEG/EEG source imaging. *NeuroImage*, 44(3):947 – 966, 2009.

[158] D. P. Wipf, J. P. Owen, H. T. Attias, K. Sekihara, and S. S. Nagarajan. Robust Bayesian estimation of the location, orientation, and time course of multiple correlated neural sources using MEG. *NeuroImage*, 49(1):641 – 655, 2010.

[159] P. Xu, Y. Tian, H. Chen, and D. Yao. l_p norm iterative sparse solution for EEG source localization. *IEEE Transactions on Biomedical Engineering*, 54(3):400–409, 2007.

[160] L. Yan, C.-L. Fu, and F.-L. Yang. The method of fundamental solutions for the inverse heat source problem. *Engineering Analysis with Boundary Elements*, 32(3): 216 – 222, 2008.

[161] S. Yan, C. Wu, W. Dai, M. Ghanem, and Y. Guo. Environmental monitoring via compressive sensing. In *Proceedings of the Sixth International Workshop on Knowledge Discovery from Sensor Data*, SensorKDD '12, pages 61–68, New York, USA, 2012. ACM.

[162] J. Yick, B. Mukherjee, and D. Ghosal. Wireless sensor network survey. *Computer networks*, 52(12):2292–2330, 2008.

[163] Z. Yong, Z. Hui, G. Dongqiang, and W. Zhihua. Determination of chemical point source using distributed algorithm in sensors network. In *24th Chinese Control and Decision Conference (CCDC'12)*, pages 3373–3377, May 2012.

[164] C. Zhang, D. Florencio, D. E. Ba, and Z. Zhang. Maximum likelihood sound source localization and beamforming for directional microphone arrays in distributed meetings. *IEEE Transactions on Multimedia*, 10(3):538–548, April 2008.

[165] T. Zhao and A. Nehorai. Distributed sequential Bayesian estimation of a diffusive source in wireless sensor networks. *IEEE Transactions on Signal Processing*, 55(4): 1511–1524, 2007.