PARAMETRIC ESTIMATION OF RANDOMLY COMPRESSED FUNCTIONS

A Thesis
Presented to
The Academic Faculty

by

William Edward Mantzel Jr.

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
School of Electrical and Computer Engineering

Georgia Institute of Technology
August

PARAMETRIC ESTIMATION OF RANDOMLY COMPRESSED FUNCTIONS

Approved by:

Professor Justin Romberg, Advisor
School of Electrical and Computer Engineering
Georgia Institute of Technology

Professor Karim Sabra
Department of Mechanical Engineering
Georgia Institute of Technology

Professor Jim McClellan
School of Electrical and Computer Engineering
Georgia Institute of Technology

Professor Erik Verriest
School of Electrical and Computer Engineering
Georgia Institute of Technology

Professor Mark Davenport
School of Electrical and Computer Engineering
Georgia Institute of Technology

Date Approved: 20 Apr 2013

To my mother and father.

ACKNOWLEDGEMENTS

It has been an honor to work with so many talented people and to learn from and be influenced by so many innovative and resourceful minds. This thesis would not have been possible without you. First and foremost, I would like to thank my advisor Justin Romberg who nearly always made time to meet, who pushed me to do more than I thought was possible, who had a way of quickly elucidating fairly arcane concepts, who played a very active role in doing the heavy lifting in key parts of this thesis (sometimes emailing me at midnight with pages of typeset notes), and whose dedication and purposeful intensity for his work are truly inspiring. Next, I would also like to thank Karim Sabra for introducing me to many of the ideas, for patiently educating me on many of the aspects of acoustic localization, for providing a wealth of knowledge of a body of work I knew very little about, and generally for providing invaluable assistance and insight into what for me has been an unfamiliar territory.

Thanks also go to Mark Davenport, Vladimir Koltchinskii, Jim McCllelan, Chris Rozell, Mike Wakin, Alireza Aghasi, and Armin Eftekhari for stimulating conversations and for pointing me towards related work that helped to place this work in context. Many joyous thanks to my fellow students : Ali, Adam, Aditya, Aurèle, Chris, Darryl, Han, Kyle, Nicolas, Ning, Ross, Salman, Sam, and Steve. I've both enjoyed and benefited from our interaction. You have brought life to an otherwise dull office environment, and I will miss our diversionary chit-chats. Thanks also for reading more drafts of mine than you would have cared to. Finally, a warm thank you to my parents Bill and Sara, my brothers Nick and Kyle, my sister Mandy, and my girlfriend Kyla. Your patience, cheer, love, understanding and encouragement have sustained me through this tough experience.

TABLE OF CONTENTS

LIST OF FIGURES

SUMMARY

Within the last decade, a new type of signal acquisition has emerged called *Compressive Sensing* that has proven especially useful in providing a recoverable representation of sparse signals. This thesis presents similar results for *Compressive Parametric Estimation*. Here, signals known to lie on some unknown parameterized subspace may be recovered via randomized compressive measurements, provided the number of compressive measurements is a small factor above the product of the parametric dimension with the subspace dimension with an additional logarithmic term. In addition to potential applications that simplify the acquisition hardware, there is also the potential to reduce the computational burden in other applications, and we explore one such application in depth in this thesis.

Source localization by matched-field processing (MFP) generally involves solving a number of computationally intensive partial differential equations. We introduce a technique that mitigates this computational workload by "compressing" these computations. Drawing on key concepts from the recently developed field of compressed sensing, we show how a low-dimensional proxy for the Green's function can be constructed by backpropagating a small set of random receiver vectors. Then, the source can be located by performing a number of "short" correlations between this proxy and the projection of the recorded acoustic data in the compressed space. Numerical experiments in a Pekeris ocean waveguide are presented which demonstrate that this compressed version of MFP is as effective as traditional MFP even when the compression is significant. The results are particularly promising in the broadband regime where using as few as two random backpropagations per frequency performs almost as well as the traditional broadband MFP, but with the added benefit of

generic applicability. That is, the computationally intensive backpropagations may be computed offline independently from the received signals, and may be reused to locate any source within the search grid area.

This thesis also introduces a round-robin approach for multi-source localization based on Matched-Field Processing. Each new source location is estimated from the ambiguity function after nulling from the data vector the current source location estimates using a robust projection matrix. This projection matrix effectively minimizes mean-square energy near current source location estimates subject to a rank constraint that prevents excessive interference with sources outside of these neighborhoods. Numerical simulations are presented for multiple sources transmitting through a generic Pekeris ocean waveguide that illustrate the performance of the proposed approach which compares favorably against other previously published approaches. Furthermore, the efficacy with which randomized back-propagations may also be incorporated for computational advantage (as in the case of compressive parametric estimation) is also presented.

CHAPTER I

INTRODUCTION

Over the last half of a century, a powerful set of tools and insights in the field of digital signal processing have evolved that shape the way we look at modern challenges of inference, observation, and prediction. Advances in storage space, computational power, information transmission, algorithmic complexity, parallelism, sensor and acquisition hardware, fabrication costs, and energy efficiency have enticed innovative approaches to previously intractable problems. However, these virtues have not advanced at a uniform rate, and many applications place great emphasis some of these attributes while remaining relatively indifferent to others. These particular niche applications have motivated the evaluation of tradeoffs whereby some attributes are improved at the expense of others. For example, recent research in the field of "sensor networks" has produced computationally intensive distributed algorithms to overcome the limitations of a network of inexpensive battery-operated sensing devices [1, 2].

1.1 Compressive Sensing

More recently, research in the field of "compressive sensing" proposes an alternative data-acquisition method to solve a variety of inference and reconstruction challenges that were considered intractable only a decade ago. The problem is stated generally as follows. We observe some "compressive" measurements $y \in \mathbb{R}^M$ of some unknown signal $x \in \mathbb{R}^N$ corrupted by noise $e \in \mathbb{R}^M$:

$$y = \mathbf{\Phi} x + e, \tag{1}$$

where the "fat" measurement matrix $\mathbf{\Phi} \in \mathbb{R}^{M \times N}$ with $M < N$ yields an underdetermined system. That is, even in the absence of noise, the recovery of a general vector

x is ill-posed because for any potential solution, there exist arbitrarily many other solutions that differ from each other along the null space of the measurement matrix. It may nevertheless be the case that a restricted search within a specific signal class yields a unique solution. In fact, there are a remarkable number of signal classes for which efficient algorithms exist to recover approximations to x, along with associated guarantees of performance.

1.1.1 Sparse Signals and the Restricted Isometry Property

Seminal work in compressed sensing established the main results for the so-called "sparse" signal class model. A signal $x \in \mathcal{X}_S \subset \mathbb{R}^N$ is called S-sparse if at most S of its elements were nonzero. That is, $\mathcal{X}_S = \{x : \|x\|_0 \leq S\}$ where the pseudonorm ℓ_0 is defined as $\|x\|_0 \triangleq \sum_n I(x_n \neq 0)$. In this way, although a signal's "ambient dimension" is N, its "sparsity" or "intrinsic dimension" $S \ll N$ more closely represents the number of degrees of freedom it exhibits under such a characterization. The set $\mathcal{X}_S$ is then S-dimensional in the same sense that the surface of a standard cube is 2-dimensional (within an ambient dimension of 3). Examples of such sparse signals include photographs of the starry sky where most of the image is black, for example.

Effective recovery of such signals was tied directly to a specific property of the linear measurement matrix $\mathbf{\Phi}$. This operator $\mathbf{\Phi}$ is said to obey the restricted isometry property (RIP) over set $\mathcal{X}$ with parameter δ if this operator is nearly isometric over the domain $\mathcal{X}$. That is, for every $x \in \mathcal{X}$, we have:

$$(1 - \delta)\|x\|^2 \leq \|\mathbf{\Phi}x\|^2 \leq (1 + \delta)\|x\|^2, \tag{2}$$

where $\|.\|$ denotes the Euclidean norm. In this case, we say that $\mathbf{\Phi}$ *embeds* $\mathcal{X}$ in $\mathbb{R}^M$. This RIP can guarantee the uniqueness of a solution within the class $\mathcal{X}_S$ in the ideal (noiseless) case, even when the problem is underdetermined in general. For example, if the operator is isometric over $2S$-sparse signals with parameter $\delta < 1$, then the compressed representation of all S-sparse signals are unique. That is, any pair of

S-sparse vectors x and z satisfying $\mathbf{\Phi}(x - z) = 0$ (i.e., $\mathbf{\Phi}x = \mathbf{\Phi}z$) must also satisfy $(1 - \delta)\|x - z\|_2^2 \leq 0$ (i.e., $x = z$) by the RIP property (2).

Although it is difficult to verify the RIP condition for any *particular* matrix [3], random matrices generated with independent and identically distributed (i.i.d.) Gaussian entries entries with zero mean and variance $1/M$ have been shown using probabilistic methods to obey the RIP condition with overwhelming probability provided that $M \gtrsim S \log(N)$ [4].

1.1.2 Recovery Algorithms

Because of the uniqueness that RIP induces in the ideal case when $e = 0$, we could estimate x uniquely using the following ℓ_0 minimization:

$$\text{minimize} \quad \|x\|_0 \quad \text{s.t.} \quad \mathbf{\Phi}x = y, \tag{ℓ_0-min}$$

because x is the unique S-sparse solution satisfying $y = \mathbf{\Phi}x$. Unfortunately, solving this system directly essentially involves testing all $\binom{N}{S}$ combinations and is known to be NP-hard, essentially requiring $O((N/S)^S)$ computations.

The key insight that enabled compressed sensing to be of practical importance, rather than an academic curiosity, is that the ℓ_1 norm, defined as $\|x\|_1 \triangleq \sum_n |x_n|$, which is the closest convex norm to the ℓ_0 pseudo-norm, may be substituted in the above optimization as the following *basis pursuit*:

$$\text{minimize} \quad \|x\|_1 \quad \text{s.t.} \quad \mathbf{\Phi}x = y. \tag{BP}$$

In the more general case when the nonzero noise term satisfies $\|e\| \leq \epsilon$, we modify the optimization to yield the following basis pursuit de-noising:

$$\text{minimize} \quad \|x\|_1 \quad \text{s.t.} \quad \|y - \mathbf{\Phi}x\| \leq \epsilon. \tag{BPDN}$$

The surprising result is that the resulting estimate from latter convex optimizations, which can be obtained relatively easily using polynomial-time algorithms [5],

coincide remarkably often with the former (ℓ_0-min) minimization. This is formalized by the following theorem:

Theorem 1 *(Theorem 1.2 of [6]) Assume that* $\delta_{2S} \leq \sqrt{2} - 1$ *and* $\|e\|_2 \leq \epsilon$ *and let* x_S *be the best* S*-term approximation to the (not necessarily sparse) vector* x*. Then the solution* x^* *to* (BPDN) *obeys:*

$$\|x^* - x\|_2 \leq C_0 S^{-1/2} \|x - x_S\|_1 + C_1 \epsilon, \tag{3}$$

for some modest universal constants depending only on δ_{2S}*. For instance, when* $\delta_{2S} = 0.2$*, the bound holds with* $C_0 = 4.2$ *and* $C_1 = 8.5$*.*

This theorem not only guarantees perfect recovery in the noiseless case when x is S-sparse, but generalizes this result in a natural way to handle signals x that are not necessarily sparse, but are at least *compressible*, so that the magnitudes of their elements, when sorted in descending order, decay rapidly.

Also, there is a closely related formulation that generalizes the sparse model to the *dictionary-sparse model* a representation that models x as the weighted sum of S dictionary elements of some orthobasis:

$$x = \sum_{n=1}^{N} \Psi_n \alpha_n = \Psi \alpha, \tag{4}$$

where Ψ represents an orthobasis of $\mathbb{R}^N$ (e.g., the discrete cosine transform basis, or the discrete wavelet basis) so that $\Psi^T \Psi = I$, and $\alpha \in \mathcal{X}_S$ is an S-sparse vector. The natural extension of (BPDN), for example, simply solves for the sparse α as follows:

$$\text{minimize} \;\; \|\alpha\|_1 \quad \text{s.t.} \quad \|y - \mathbf{\Phi}\Psi\alpha\| \leq \epsilon. \tag{BPDN-Ψ}$$

In fact, the mechanics of recovery of this sparse vector are identical to the canonical case (i.e., when $\Psi = I$) by simply using a different observation matrix $\tilde{\mathbf{\Phi}} = \mathbf{\Phi}\Psi$ in the recovery procedure, which incidentally is equal in distribution to $\mathbf{\Phi}$ in the i.i.d. Gaussian case, due to rotational invariance.

Because of convex duality [7], the (BPDN) optimization is equivalent to its Lagrangian form, called LASSO [8]:

$$\text{minimize} \quad \|y - \mathbf{\Phi}x\|^2 + \lambda\|x\|_1, \tag{LASSO}$$

for some value of dual variable λ, and for the same reason is equivalent to the following optimization, for some value of L:

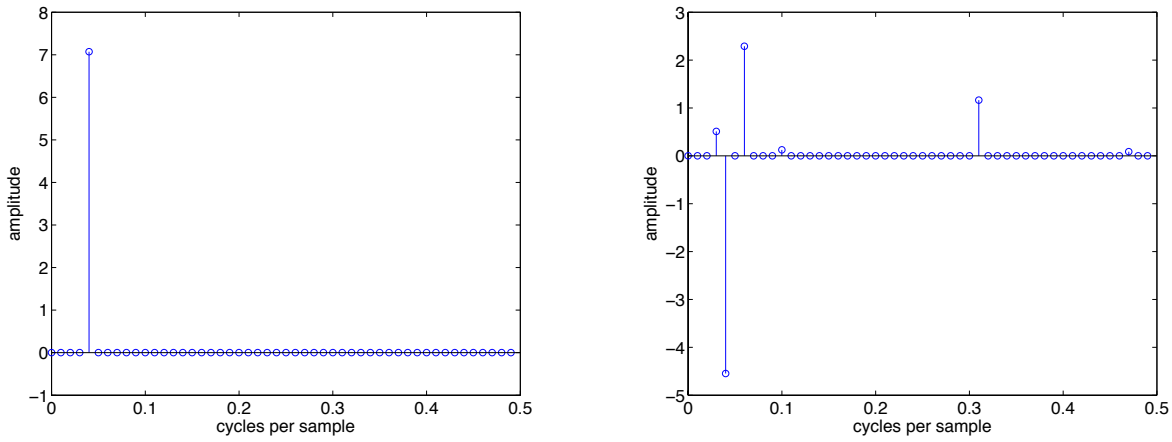
$$\text{minimize} \quad \|y - \mathbf{\Phi}x\|^2 \quad \text{s.t.} \quad x \in \mathcal{X}_L, \tag{5}$$

where $\mathcal{X}_L = \{x : \|x\|_1 \leq L\}$. Each of these equivalent formulations are advantageous at various times for building intuition, drawing connections with related work, and efficiently computing solutions. In particular, Eq. (5) takes the form that directly parallels the compressive parametric estimation defined below.

1.1.3 Applications and Limitations

The advent of compressive sensing has led to a series of innovations in a number of areas. In the field of medical imaging, compressed sensing techniques have been used to recover both magnetic resonance [9, 10, 11] and computed tomography [12, 13] images using fewer measurements or a simplified acquisition hardware compared to previous approaches. In the field of telecommunications, the insights of compressed sensing have been used to randomly generate an encoding operation for a transmitted signal to protect it against sparse additive errors [14], and also for the estimation of an unknown sparse channel [15, 16, 17, 18]. Additionally, novel imaging systems were prototyped, including the single-pixel camera [19] that have developed alongside related innovations from the computational photography community such as the coded aperture [20, 21] and the flutter shutter [22].

But there are subtle caveats to this field of compressed sensing, even in the ideal noiseless case. Consider, for example, the application of frequency estimation of a discrete sinusoid $x_n = \cos(8\pi n/N)$ for $0 \leq n \leq N-1$ from its compressed measurements $y = \mathbf{\Phi}x$, using an ambient dimension of, say, $N = 100$. By using as few



(a) (b)

Figure 1: The magnitudes of the recovered α^* vector using (BPDN-Ψ) when the underlying signal is (a) sparse ($x_n = \cos(8\pi n/N)$) and (b) compressible ($x_n = \cos(8\sqrt{2}\pi n/N)$) with respect to the Fourier basis so that the sorted coefficient decay as $O(1/n)$ in magnitude, resulting in poor recovery performance.

as 10 measurements out of the ambient dimension, (BPDN-Ψ) recovers the original signal correctly with high probability using the appropriate DCT basis Ψ (i.e., with $\Psi\Psi^T = I$) as shown on Fig. 1a.

On the other hand, by simply modifying the signal to have an irrational frequency, $x_n = \cos(8\sqrt{2}\pi n/N)$, the signal is no longer sparse in the DCT basis, and now the basis pursuit recovery gives an M-sparse signal with several nonzero elements but, at best, with its largest element close to the true frequency (as in Fig. 1b). In general, the largest element corresponds to a different frequency.

This result is unsatisfying, because it seems at least as though there is as much information in the observed vector y corresponding to the frequency $4/N$ cycles per sample as the one corresponding to the frequency $4\sqrt{2}/N$ cycles per sample. This sort of leakage phenomenon is not limited to frequency estimation, but shows up in many applications such as compressive target tracking, matched filtering, and the particular application we present in this thesis, passive acoustic localization. This type of artifact illustrates the perils of stretching the sparse signal model beyond its

intended use, and motivates the exploration of an alternative model.

1.2 Compressive Parametric Estimation

For all of the development in the field of sparse signal recovery, there is another class of signals that is not yet as well understood. This thesis aims to develop a general framework and corresponding theory for *compressive parametric estimation* (CPE), focusing particularly on the application of passive acoustic source and multi-source localization using compressive matched field processing [23].

In its most general form, parametric estimation involves searching for the closest function to h from within a parameterized set $\mathcal{F}$:

$$\text{minimize} \quad \|h - f\|^2 \quad \text{s.t.} \quad f \in \mathcal{F}. \tag{PE}$$

Its compressive counterpart simply finds the closest function with respect to some dimension-reducing linear operator $\mathbf{\Phi}$:

$$\text{minimize} \quad \|\mathbf{\Phi}(h - f)\|^2 \quad \text{s.t.} \quad f \in \mathcal{F}, \tag{CPE}$$

a constrained minimization that parallels Eq. (5). One of the main results of the thesis is that, for an appropriate choice of $\mathbf{\Phi}$, for a wide variety of parameter classes $\mathcal{F}$, and with high probability, (CPE) yields a solution that is characteristically similar to (PE).

Because this set $\mathcal{F}$ is not necessarily convex, we generally must use an exhaustive search over the entire parameter space to find the global optimum. Consequently, compressive parametric estimation generally lends itself well to problems with only a few parameters. For example, scanning a two-dimensional area for land mines, estimating a three-dimensional registration between a pair of images, searching for a one dimensional time-shift, or searching over range and depth for one or more acoustic sources – the application focused on in this thesis.

One example of nonlinear parametric estimation in signal processing is matched filtering where $\mathcal{F} = \{f_0(t-\theta) : \theta \in \Theta\}$ for some bounded set $\Theta = [a \quad b]$ and some base function f_0 where the parameter θ that is implicitly estimated corresponds to the best matching "shift" of the modeled base function f_0 to the observed function h. It is often taken for granted that the resulting parametric estimate $\bar{\theta}$ to (PE) is invariant to scalar multiplication of either h or f_0. That is, the solution is equivalent to the solution obtained via the set $\mathcal{F} = \{\alpha f_0(t-\theta) : \theta \in \Theta, \alpha \in \mathbb{R}\}$ for some compact parameter set $\Theta \subset \mathbb{R}^D$. This is a valuable feature since the scale of the received signal is often not known in advance. However, this feature depends entirely upon the fact that all $f \in \mathcal{F}$ have the same norm, and is unfortunately not shared with (CPE) since the $\mathbf{\Phi} f$ do not all share the same norm.

We can simultaneously overcome this shortcoming and generalize this parameterized set in an interesting way by explicitly defining $\mathcal{F}$ as a *parameterized collection of K-dimensional subspaces*:

$$\mathcal{F} = \{\mathbf{V}_\theta \alpha : \theta \in \Theta, \alpha \in \mathbb{R}^K\}, \tag{6}$$

where $\mathbf{V}_\theta : \mathbb{R}^K \to L_2$ represents an orthobasis spanning parameterized subspace $\mathcal{S}_\theta$.

Now, rather than modeling h as a shift of a function, we model h more generally from a class of parameterized subspaces. This generalization affords us a wide variety of applications that are especially well suited for inverse problems with a few number of nonlinear parameters and potentially many linear coefficients.

This generalization essentially amounts to a collection of least-squares problems, one for each fixed value of θ. We will discuss specific practical cases in Chapters 2 and 3 when solving this system is not only feasible but computationally advantageous. In such cases and others when CPE provides other advantages, it is important to weigh the cost, which is primarily due to the loss in accuracy.

To illustrate this approach, we apply CPE to the compressive tone estimation problem described above in section 1.1.3. On Figure 2, we compare the performance

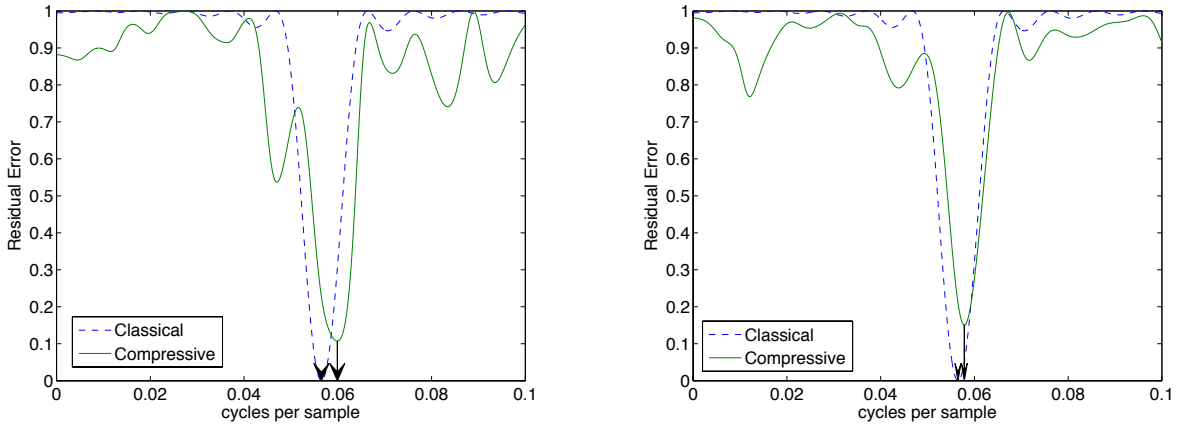


Figure 2: This figure shows that parametric estimation by minimizing the compressive proxy $\|\mathbf{\Phi}(h-f)\|^2$ over the set of all discrete sinusoids $f \in \mathcal{F}$ (i.e., (CPE), shown with the solid line) often gives a characteristically similar estimate to the classical approach that minimizes $\|h-f\|^2$ (i.e., (PE), shown with the dashed line) using only (a) $M = 10$ and (b) $M = 20$ compressive measurements. The arrows indicate the frequency estimates that result from these approaches, and show the similarity of these estimators.

of (CPE) with 10 and 20 compressive measurements to the classical estimator (PE). By modeling the functional set directly, we are able to achieve performance that more closely resembles the classical estimator than the ℓ_1 minimization discussed above.

In Chapter 4, after reviewing related work that has already applied specific instances and minor variations of CPE, we will give probabilistic performance bounds that depend only on the subspace dimension and the geometry or "regularity" of the parameterized set of subspaces. To wit, under mild conditions of regularity, we show that CPE performs favorably when the number M of compressive measurements taken are a small multiple of the product of the parameter dimension and subspace dimension, with an additional log factor in this subspace dimension and the parametric volume.

1.3 Compressive Matched Field Processing

The primary application of CPE that we consider in this thesis is compressive matched field processing (cMFP) for localizing underwater acoustic targets from passive sonar data. We give a short introduction here by giving the classic formulation of MFP and then continue to show how it may be modified in a straightforward way for computational advantage.

For the sake of brevity and simplicity, we discuss the simple canonical case where a single sound-source at location $\vec{r}_0 \in \mathbb{R}^2$ (containing range and depth) emits sound described by known frequency ω and unknown complex amplitude α. The thesis will discuss and show results for the broadband MFP where the advantages of random compression are much more salient.

The goal is to estimate the location of the source from the corresponding received complex amplitude at N receiver locations y_n ($n \in \{1, ..., N\}$), given by

$$y_n = \alpha g(\vec{r}_n, \vec{r}_0) + \eta_n, \tag{7}$$

where the Green's function $g(\vec{r}_n, \vec{r}_0)$ describes the acoustic frequency response between two locations and η_n is some noise term. Using the common assumption that η_n is independent and identically distributed Gaussian noise, the maximum likelihood solution for the source location gives rise to the familiar least-squares formulation:

$$\bar{r} = \arg \min_{\vec{r} \in \mathcal{R}} \min_{\beta \in \mathbb{C}} \ \|Y - \beta G(\vec{r})\|^2, \tag{8}$$

where $Y \in \mathbb{C}^N$ and $G(\vec{r}) \in \mathbb{C}^N$ (i.e., $G : \mathbb{R}^2 \to \mathbb{C}^N$) are the vectorized forms of y_n and $g(\vec{r_n}, \vec{r})$ over all $n \in \{1, ..., N\}$. Note that this formulation has the same form as Eq. (PE) with the range/depth vector $\vec{r}$ representing the parameter vector θ.

For any fixed location $\vec{r}$, the inner optimization problem over β (whose scalar value is of no intrinsic interest) is simply finding the closest point on the line spanned by $G(\vec{r})$ to the point Y. Plugging in the closed-form solution to this problem, the

problem above reduces to

$$\arg\max_{\vec{r}} \frac{|Y^H G(\vec{r})|^2}{\|G(\vec{r})\|^2}, \tag{9}$$

where Y^H denotes the Hermitian transpose. We designate this objective function as the ambiguity function $h(\vec{r})$ and show it in Figure 4.a.

In common practice, an unnormalized variation on Eq. (9) is solved by constructing $|Y^H G(\vec{r})|$ via a single back-propagation and identifying the maximizing $\vec{r}$. The construction of this unnormalized ambiguity function takes only as much computational effort as the evaluation of the full Green's frequency response $G(\vec{r})$ for a single point $\vec{r}$, though the quality of the estimate will suffer somewhat, as illustrated by the gap between the green and blue dashed lines on Figure 7a. In general, solving Eq. (9) requires knowledge of the full Green's function: the frequency response between any feasible source location and any of the N receivers. This process generally involves solving N computationally intensive PDEs to determine the frequency response between each of the N receivers and each candidate source location point.

By using a compressive approach, we can achieve comparable performance by solving only $M < N$ PDEs via a compression matrix $\mathbf{\Phi} \in \mathbb{R}^{M \times N}$. We are able to compute $\mathbf{\Phi} G(\vec{r})$ via M propagations of the form $G(\vec{r})^H \phi_m$, where ϕ_m is one of the M random rows of the matrix $\mathbf{\Phi}$. We refer to each one of these random back-propagations as a random measurement because it plays an analogous role to the random measurements taken in the traditional CS paradigm, and can be thought of as a measurement probe that gives some partial information about $G(\vec{r})$, though we note here that the term measurement is simply a useful fiction.

The application of the compressive parametric estimation as in Eq. (CPE) yields a least-squares problem in compressed space:

$$\arg\min_{\vec{r}} \min_{\beta} \ \|\mathbf{\Phi} Y - \beta \mathbf{\Phi} G(\vec{r})\|^2, \tag{10}$$

which reduces to

$$\text{(narrowband cMFP)} \qquad \arg\max_{\vec{r}} \frac{|Y^H \mathbf{\Phi}^H \mathbf{\Phi} G(\vec{r})|^2}{\|\mathbf{\Phi} G(\vec{r})\|^2}. \tag{11}$$

We designate this latter objective function $\tilde{h}(\vec{r})$ and show it in Figure 4.b and 4.c. It can be interpreted as a compressed version of the ambiguity function $h(\vec{r})$ in (9) shown in Figure 4.a.

Note that unlike the standard MFP, in this case the pre-computations give us direct access to the denominator $\|\mathbf{\Phi} G(\vec{r})\|^2$ (we simply take the norms of the columns of $\mathbf{\Phi} G(\vec{r})$), and so we leave it in the optimization program. This normalization term plays an important role in improving the source location estimation accuracy by up to a factor of two. Notice that an evaluation of (11) at a point $\vec{r}$ essentially only requires an inner product between the encoded observations $\mathbf{\Phi} Y$ and the M-vector $\mathbf{\Phi} G(\vec{r})$ (formed from the backpropagated fields at point $\vec{r}$ from all M random vectors) that can be effectively carried out with a matrix-vector multiply. This application is presented in greater detail in Chapter 3.

1.4 Multiple-Source Localization

This compressive approach to matched-field processing may be extended from a single source to multiple sources. In this thesis, we present a robust round-robin multiple-source localization method that uses a greedy algorithm similar to orthogonal matching pursuit [24]. To sum up, consider the following variation on Eq. (7). Suppose that instead of a single source $\vec{r}$ transmitting a narrowband pulse, we have $S > 1$ sources transmitting as:

$$Y = \sum_{s=1}^{S} \beta_s G(\vec{r}_s) + \eta, \tag{12}$$

for some noise term η. The localization approach that extends naturally from the earlier least-squares solution is then:

$$\arg\min_{\vec{r}_s \in \mathcal{R}} \min_{\beta \in \mathbb{C}} \; \|Y - \sum_{s=1}^{S} \beta_s G(\vec{r}_s)\|^2, \tag{13}$$

If not for computational constraints, we would solve this by maximizing directly over all joint combinations of source locations. When the $\vec{r}_s$ are fixed, this minimization amounts to linear least squares over the β_s. However, for the full minimization over all variables this approach is typically only computationally feasible for only 2 or 3 sources and suffers from the curse of dimensionality otherwise. Instead, we utilize a greedy approach that iteratively estimate each of the source location vectors $\vec{r}_s$ one at a time [24]. For an overview of these methods, refer to Appendix A.2.

1.5 Theoretical Analysis of CPE

The difference in performance between the classical and compressive parametric estimators primarily relate to the difference between the corresponding classical and compressive parametric-subspace-projection operators over a specific parameter set. In Chapter 4, we show the conditions under which this difference is small leading to favorable performance.

To this end, we leverage tools in empirical processes and random matrix theory to bound the performance of our proposed estimator, giving specific guarantees that depend only on the dimension of the subspaces and the geometry of the parameterized set. We do this by setting up the problem as the maximum deviation of an empirical process whose mean is precisely the classical parametric estimator, proceeding by bounding several processes of interest, incidentally showing the well-conditionedness of the compression operator with respect to the parameterized set. For these supremum bounds of random processes, we develop a chaining argument similar to and largely derived from the ones utilized by Talagrand [25] but tailored for our assumptions on the regularity of our parameter set to give specific probabilistic bounds. To make use of this chaining argument, it then only remains to establish the associated increment tail bounds for these various processes, showing that samples of these processes are "close" with high probability whenever the corresponding deterministic

parameterized subspaces are "close".

1.6 Contributions and Organization of the Thesis

The thesis presents a compressive approach to parametric estimation, focusing on the specific context of the passive localization of acoustic sources. The specific unique contributions are discussed in greater detail with the following overview of the thesis.

Chapter 2 introduces compressive matched field processing (cMFP), a computationally efficient method for passive acoustic source localization. While traditional approaches would construct a series of Green's vectors to match against by performing N time-reversed backpropagation partial differential equations (PDE) solutions across the N receivers, we show how a similar type of dimension-reduced templates may be constructed by time-reversing $M < N$ randomly chosen sets of initial conditions, and matching against the resulting templates. The application to the broadband case is discussed in terms of both the incoherent and coherent regimes where the source signal is either unknown or generally known to within an unknown scale factor, respectively. Simulated results employing a Pekeris ocean waveguide suggest only a modest sacrifice in accuracy, and are particularly promising in the broadband regime where using as few as two random backpropagations per frequency (resulting in an order of magnitude computational gain) performs almost as well as the traditional broadband MFP, but with the added benefit of generic applicability. That is, the computationally intensive backpropagations may be computed offline independently from the received signals, and may be reused to locate any source within the search grid area.

Chapter 3 extends this cMFP approach to multiple-source passive acoustic localization via matched-field processing (MFP). Here, we introduce a round-robin approach for multi-source localization. Each new source location is estimated from the ambiguity function after nulling from the data vector the current source location

estimates using a robust projection matrix. This projection matrix effectively minimizes mean-square energy near current source location estimates subject to a rank constraint that prevents excessive interference with sources outside of these neighborhoods. Numerical simulations are presented for multiple sources transmitting through a generic Pekeris ocean waveguide that illustrate the performance of the proposed approach which compares favorably against other previously published approaches.

Finally, Chapter 4 formalizes these parametric approaches to localization as a compressive parametrized subspace estimation problem. The problem formulation is somewhat more general than existing approaches in compressive parametric estimation, and favorable performance is claimed to depend only on the number of measurements and the condition of a specific type of geometric regularity. This type of regularity appears to be satisfied by a wide variety of parametric subspace classes, and is succinctly expressed in terms of an effective dimension and base covering number. Apart from proving this claim, this chapter focuses on validating this assumption of regularity for time-shifts of orthobases that are approximately compactly supported in both time and frequency. It is furthermore explained how compressive parametric estimation obeys this form of regularity.

1.6.1 Notation

This thesis will measure norms in a variety of different ways and will utilize the consistency between norms to make intuitive cases of the utility of CPE. Unless otherwise subscripted, all norms $\|.\|$ are Euclidean ℓ_2 norms for vectors, L_2 norms for functions, and operator 2-norms (spectral norms) for matrices. We will denote the Frobenius norm as $\|.\|_F$, which is equivalently the norm of the singular values of this matrix. We will also use a stochastic measure called the Orlicz norm, which will be denoted as $\|.\|_{\Psi_1}$ and defined later in Chapter 4.4.

Matrices and linear operators are generally capitalized and bold, e.g., $\mathbf{G}$. Scalars

are generally lowercase. C shall denote a universal constant, not necessarily the same at every occurrence. Subscripted constants (e.g., C_1) denote specific universal constants, generally with a known upper bound.

CHAPTER II

COMPRESSIVE MATCHED FIELD PROCESSING

2.1 Introduction

2.1.1 Background and Motivation

Matched field processing (MFP) continues to serve as one of the most widely used methods for localizing undersea targets acoustically. However, as the models governing undersea acoustic interactions become more sophisticated, often requiring fine-grain solutions to more complex partial differential equations, the tradeoff between run time and performance begins to worsen, perhaps unnecessarily. We will begin by discussing why this is the case and giving an overview of our approach to mitigate the problem.

MFP generalizes standard array beamforming methods (e.g. plane wave beamforming) for locating an acoustic source in a complex environment (such as a multipath shallow water waveguide). MFP has been studied extensively both theoretically and experimentally as described in several review articles [26, 27, 28]. MFP is usually implemented by systematically placing a test point source at each point of a spatial search grid of L candidate locations, computing the acoustic field (replicas) at all the elements of the receiver array and then correlating this modeled field with the data from the real point source whose localization is unknown to determine the best-fit location (see Fig. 3). This approach works well when that the computational replica environment is sufficiently accurate. However, this direct implementation of MFP using brute force search would require L computation runs which can become numerically cumbersome for large search space especially when simulating complex propagation environments.

One alternative to this direct implementation of MFP is to use a "backpropagation" algorithm (also referred to as "time-reversal imaging") to locate the unknown source. In this case, a time-reversed version of the recorded data is used as an initial waveform excitation along the array aperture using the principle of superposition, and then subsequently "backpropagated" numerically in the replica environment towards the grid search area [29]. The unknown source location is then estimated from the maximum of the distribution of the backpropagated peak amplitude (or energy) across the grid search. Consequently, when compared to the direct implementation first mentioned, this backpropagation approach appears attractive at first glance, since it requires one computational run per unknown source. Nevertheless, this backpropagation approach becomes computationally expensive if multiple sources need to be located repetitively over the *same* search grid as the number of required computational runs would grow proportionally. For instance, this may occur when one tries to locate a source moving along a long track throughout the search space. Indeed, in order to be able locate any source throughout the search space using N receivers, MFP would require computing N backpropagations by using sequentially each individual receiver as a backpropagation source [26, 27]. This would allow determining the full set of Green's functions associated with the channel between each search location and each receiver element. Alternatively, one could weight spatially the amplitude of the backpropagated signals along the receiver array using N different orthogonal codes (e.g. obtained from an Hadamard basis).

This article develops instead a compressive MFP formulation which reduces this computational burden by pre-computing the backpropagation of a number $M \ll N$ of random test signals. The results of these backpropagations effectively encode the Green's function associated with the channel, and they can be re-used in subsequent localizations without any additional computational cost. This approach is inspired by recent work in the field of compressed sensing [30, 31, 32], whose central message is

that random projections provide an effective encoding for sparse signals. The motivation for compressed sensing is typically concerned with reducing the cost of acquiring signals by shifting the workload from sensor hardware to software [33, 34, 35], and is natural in applications where physical measurements are expensive compared to numerical computations. Here we explore a variation on this theme: mitigating the computational workload in software instead of the sensing workload in hardware. The proposed compressive MFP allows us to estimate the underlying ambiguity function central to conventional MFP algorithms over the entire search space using only M computational runs instead of N, an effective speedup of a factor of N/M. In practice, these M simulations can be independently computed as a background process offline before the actual source signal is received.

2.1.2 Related Work

In this chapter, we effectively demonstrate how classical localization procedures under a least-squares framework such as matched-field processing (MFP) may be solved in a reduced-dimensional space even without *a-priori* knowledge of the "best" dimension-reducing transform. This property has been shown in similar forms in the mainstream canon of Compressed Sensing (CS) literature. Davenport et al. have described a number of useful variations on the theme of CS including a matched filtering detector [36]. They have also described the "smashed filter" that performs compressive parametric estimation inside of a generalized likelihood ratio test [37]. Wakin has also established some rigorous results on parameter estimation that relate the recovery properties of a general compressive estimation problem to the properties of the manifold that these parameters induce. Their work could be used to analyze this problem of acoustic localization via its manifold parameters [38].

Carin et al. have utilized CS principles to show how a Green's function of a scattering field that is compressible in the wavelet domain may be recovered from a

small set of measurements, though they use incoherence in their structured measurements to recover a scattering field, while we primarily care about the location of the source [39]. Likewise, Marengo et al. have applied compressed measurements to the scattering problem, utilizing the target-sparse model to improve their performance [40].

Our work may also be viewed in the context of randomized SVDs [41]. In this field of research, the idea is to apply the matrix $\boldsymbol{A}$ to a series of random vectors Φ_m as $\boldsymbol{A}\Phi_m$ in order to determine the range space of $\boldsymbol{A}$. For example, Chaillat et al. show how the inverse medium problem can be simplified using a dimension reducing random projection and solving the inverse problem in the reduced range-space [42]. Similarly to this field, we apply the time-reversal or adjoint of the Green's function $\mathbf{G}_\omega$ to random vectors in order to discover the range space of admissible ambiguity functions.

There is also a large amount of recent research performing multi-target tracking under the "target-sparse" assumption. That is, the methods propose to simultaneously localize several targets that lie on some grid (or generally some set of points) by solving an ℓ_1 minimization program. The recovered support resulting from this optimization corresponds to the grid points that the various targets are estimated to occupy. All of this work dovetails in very nicely with the main results of Compressed Sensing, which can be effectively leveraged to prove that the targets may be perfectly localized with high probability. Often, the painstaking effort in these papers involves showing that the Restricted Isometry Property (RIP) holds for the observation matrix. For example, Fannjiang et al. show the conditions under which a sufficiently small coherence is achieved for perfect recovery [43]. Gurbuz et al. show similar results for a *Compressive* beamformer, requiring a number of measurements on the order of the number of sources [44], but the application there is different in that they utilize a signal common to all sensors with an unknown time shift to localize their target

in angle (assuming free space propagation), and apply the compression operator in time per-sensor instead of applying the operator across the range of sensors as we do. Also from a communications perspective, Cevher et al. demonstrate the relatively low amount of information to be transmitted for purposes of localization when using a Compressed Sensing framework [45]. These "target-sparse" approaches depend on targets lying exactly on the grid points. Also, by necessity these grid points must be spaced sufficiently far away from one another to avoid coherence-inducing correlations in the observation matrix. This creates a restrictive model of limited applicability. When a target is somewhere in between a set of grid points, the necessary conditions for recovery may not even *approximately* hold, similarly to how a discrete sinusoid corresponding to an off-grid point in the DFT will not be sparse in the frequency domain (or any other basis for any standard transform for that matter) due to DFT leakage. In contrast to this approach, we do not require our target to lie on a grid point. However, instead of promising perfect recovery, we instead content ourselves to claim that our target may be localized to within a small neighborhood of the actual source location, or at least the location found via deterministic means.

2.1.3 Outline

The remainder of the chapter is organized as follows. Section 2.2 briefly describes conventional MFP formulation for locating both single-frequency (narrowband) and broadband sources. Section 2.3 presents the corresponding compressive MFP (cMFP) formulation for both cases. Section 2.4 presents numerical simulations in a Pekeris waveguide [46, pg 540–552] illustrating the performance of cMFP in comparison to the conventional MFP results including the effects of additive ambient noise to the data and model mismatch due to uncertain knowledge of the actual environment. Section 2.5 extends this compressive approach to adaptive MFP. Section 2.6 summarizes the findings and conclusions drawn from this study.

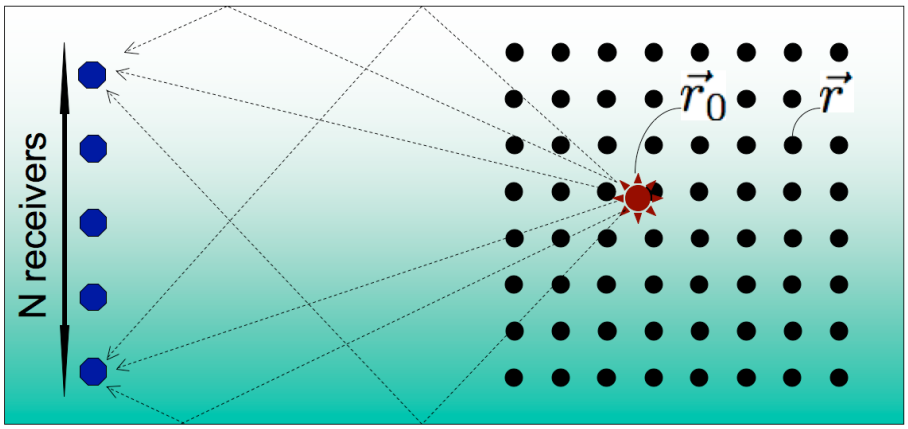


Figure 3: Schematic of a matched-field processing implementation in an ocean waveguide. The signal transmitted by a source (star symbol) located at an unknown location $\vec{r}_0$ is recorded along a N elements receiver array after multipath propagation. Using a computational model of the original ocean waveguide, the location $\vec{r}_0$ may be inferred by matching the actual received signals with the simulated replica waveforms obtained from varying the test source location (dot symbols) $\vec{r}$ throughout the search grid area.

2.2 Conventional MFP

A brief summary of the conventional MFP formulation is presented hereafter based on the standard solution of the linearized wave equation. The acoustic pressure field $y(\vec{r},t)$ at a fixed point $\vec{r}$ and time t produced by a point source located at $\vec{r}_0$ satisfies:

$$\frac{1}{c^2(\vec{r})}\frac{\partial^2 y(\vec{r},t)}{\partial t^2} - \nabla^2 y(\vec{r},t) = \alpha(t)\delta(\vec{r}-\vec{r}_0) \tag{14}$$

where $c(\vec{r})$ is the speed of sound and $\alpha(t)$ is the signal emitted by the source. The time-domain Green's function for the same environment $g(\vec{r},\vec{r}_0,t)$ is, by definition, the solution of Eq. (14) for a impulsive point source (i.e. for $\alpha(t) = \delta(t)$) that satisfies all boundary conditions [46, pg 540–552]. Using Eq. (14) (and assuming that the radiation condition applies as $\|\vec{r}\| \to \infty$) the Fourier transform of the recorded pressure field at $\vec{r}_n$, the n^{th} element of a receiver array ($n = 1..N$) (see Fig. 3), is denoted $y_\omega(\vec{r}_n)$ and given by:

$$y_\omega(\vec{r}_n) = \alpha_\omega g_\omega(\vec{r}_n,\vec{r}_0) \tag{15}$$

where ω is the frequency. The variables α_ω and $g_\omega(\vec{r}_n, \vec{r}_0)$ denote respectively the Fourier transform of the source signal and time-domain Green's function. Using vector notation, Eq. (15) can be restated as:

$$Y_\omega = \alpha_\omega G_\omega(\vec{r_0}), \tag{16}$$

where Y_ω is a $(N \times 1)$ column vector obtained by stacking the complex amplitudes $y_\omega(\vec{r}_n)$ measured along the receiver array. Similarly, the $(N \times 1)$ column vector $G_\omega(\vec{r})$ contains Green's functions $g_\omega(\vec{r}_n, \vec{r})$ between the N receiver array elements and a source located at $\vec{r_0}$. Note that the position vectors are written in lowercase letters with arrows and the column vectors are written with capital letters in the remainder of this article.

2.2.1 Single-Frequency MFP

We start by considering the simplest MFP that works from measurements at a single frequency ω (as in (16)), known as the harmonic (or narrowband) formulation. Given a set of measurements $Y_\omega \in \mathbb{C}^N$ across the N receivers at frequency ω, we search for the location $\vec{r}$ in our region of interest $\mathcal{R}$ (and complex source amplitude β) that best accounts for these measurements by solving the least-squares problem

$$\arg\min_{\vec{r}\in\mathcal{R}} \min_{\beta\in\mathbb{C}} \|Y_\omega - \beta G_\omega(\vec{r})\|^2. \tag{17}$$

With the location $\vec{r}$ fixed, the inner optimization problem is simply finding the closest point on the line spanned by $G_\omega(\vec{r})$ to the point Y_ω. Plugging in the closed-form solution to this problem (see Appendix A.1), the problem above reduces to:

$$\arg\min_{\vec{r}} \|Y_\omega\|^2 - \frac{|Y_\omega^H G_\omega(\vec{r})|^2}{\|G_\omega(\vec{r})\|^2} \;=\; \arg\max_{\vec{r}} \frac{|Y_\omega^H G_\omega(\vec{r})|^2}{\|G_\omega(\vec{r})\|^2}, \tag{18}$$

(where Y_ω^H denotes the Hermitian transpose) which we will refer to as the *normalized ambiguity function*, and will refer to its maximization as *normalized Matched Field Processing* (nMFP). We show an example of the normalized ambiguity function in Fig. 4.a.

The term $Y_\omega^H G_\omega(\vec{r})$ can be computed at every location $\vec{r}$ in an efficient manner using time-reversal. Precise values for $\|G_\omega(\vec{r})\|^2$ are typically not available when computing the backpropagation $Y_\omega^H G_\omega(\vec{r})$. However it is often the case (and we will assume this here) that these energies either do not vary much across our locations of interest, or vary predictably (e.g. cylindrical spreading of the field amplitude). Dropping the denominator yields the so-called *unnormalized ambiguity function* (alternatively the unnormalized Bartlett formulation) [26, 27, 28], the objective function used for estimating the source location:

$$\hat{\vec{r}} = \arg\max_{\vec{r}} |h(\vec{r})|^2 \quad \text{where} \quad h(\vec{r}) = Y_\omega^H G_\omega(\vec{r}), \tag{19}$$

2.2.2 Broadband MFP

Now suppose that most of the energy of the source signal occupies some continuous bandwidth $[\omega_{\min} \quad \omega_{\max}]$, known as the broadband formulation. Ideally, we would solve (17) over a continuum of ω values. However, for the sake of source localization, it is computationally advantageous to sample this bandwidth at K frequencies $\omega_1, \omega_2, \ldots, \omega_K$, yielding K measurement vectors Y_{ω_k} where $k \in \{1, 2, ...K\}$. In this way, we can achieve a computational complexity at most K times the single frequency case, without sacrificing much precision.

We now search for the location $\vec{r}$ that jointly matches the joint behavior of the measurements Y_ω over multiple frequencies $\omega_1, \omega_2, \ldots, \omega_K$. The least-squares problem from (17) becomes

$$\arg\min_{\vec{r}} \min_{\beta_{\omega_1},\ldots,\beta_{\omega_K}} \sum_{k=1}^{K} \|Y_{\omega_k} - \beta_{\omega_k} G_{\omega_k}(\vec{r})\|^2. \tag{20}$$

The inner optimization problem is separable over the β_{ω_k}, and so the above is equivalent to

$$\arg\min_{\vec{r}} \sum_{k=1}^{K} \min_{\beta_{\omega_k}} \|Y_{\omega_k} - \beta_{\omega_k} G_{\omega_k}(\vec{r})\|^2 \;=\; \arg\max_{\vec{r}} \sum_{k=1}^{K} \frac{|Y_{\omega_k}^H G_{\omega_k}(\vec{r})|^2}{\|G_{\omega_k}(\vec{r})\|^2}. \tag{21}$$

As before, if the energies $\|G_{\omega_k}(\vec{r})\|^2$ are homogenous across space and frequency, then a reasonable unnormalized approximation to the above is

$$\arg\max_{\vec{r}} \sum_{k=1}^{K} |Y_{\omega_k}^H G_{\omega_k}(\vec{r})|^2 \;=\; \sum_{k=1}^{K} |h_{\omega_k}(\vec{r})|^2. \tag{22}$$

The formulation in (22) assumes that the source amplitudes β_{ω_k} are unknown. If we have knowledge of the source signal's complex amplitudes, that is we know them up to a common amplitude and phase, then (20) can be refined to

$$\arg\min_{\vec{r}} \min_{\beta\in\mathbb{C}} \left\| \begin{bmatrix} Y_{\omega_1} \\ Y_{\omega_2} \\ \vdots \\ Y_{\omega_K} \end{bmatrix} - \beta \begin{bmatrix} \alpha_{\omega_1} G_{\omega_1}(\vec{r}) \\ \alpha_{\omega_2} G_{\omega_2}(\vec{r}) \\ \vdots \\ \alpha_{\omega_K} G_{\omega_K}(\vec{r}) \end{bmatrix} \right\|^2 \tag{23}$$

where the source amplitudes α_{ω_k} are fixed and known. Applying again the results from Appendix A.1, the inner optimization program can be solved in closed form, and so (23) is equivalent to

$$\arg\max_{\vec{r}} \frac{\left|\sum_{k=1}^{K} \alpha_{\omega_k} Y_{\omega_k}^H G_{\omega_k}(\vec{r})\right|^2}{\sum_{k=1}^{K} |\alpha_{\omega_k}|^2 \|G_{\omega_k}(\vec{r})\|^2}, \tag{24}$$

as shown in Fig. 4.b, which we can approximate (by removing the denominator) as its unnormalized counterpart

$$\arg\max_{\vec{r}} \left| \sum_{k=1}^{K} \alpha_{\omega_k} h_{\omega_k}(\vec{r}) \right|^2. \tag{25}$$

Hereafter, we will refer to (21) and (22) as the *incoherent* MFP formulation, and (24) and (25) as the *coherent* MFP formulation.

2.3 Compressive MFP

In this section, we describe Compressive Matched Field Processing (cMFP). This is an efficient method for acquiring a *compressed* version of the Green's function operator $G_\omega(\vec{r})$ that exhibits a behavior in some regards similar to the dimension-reduced

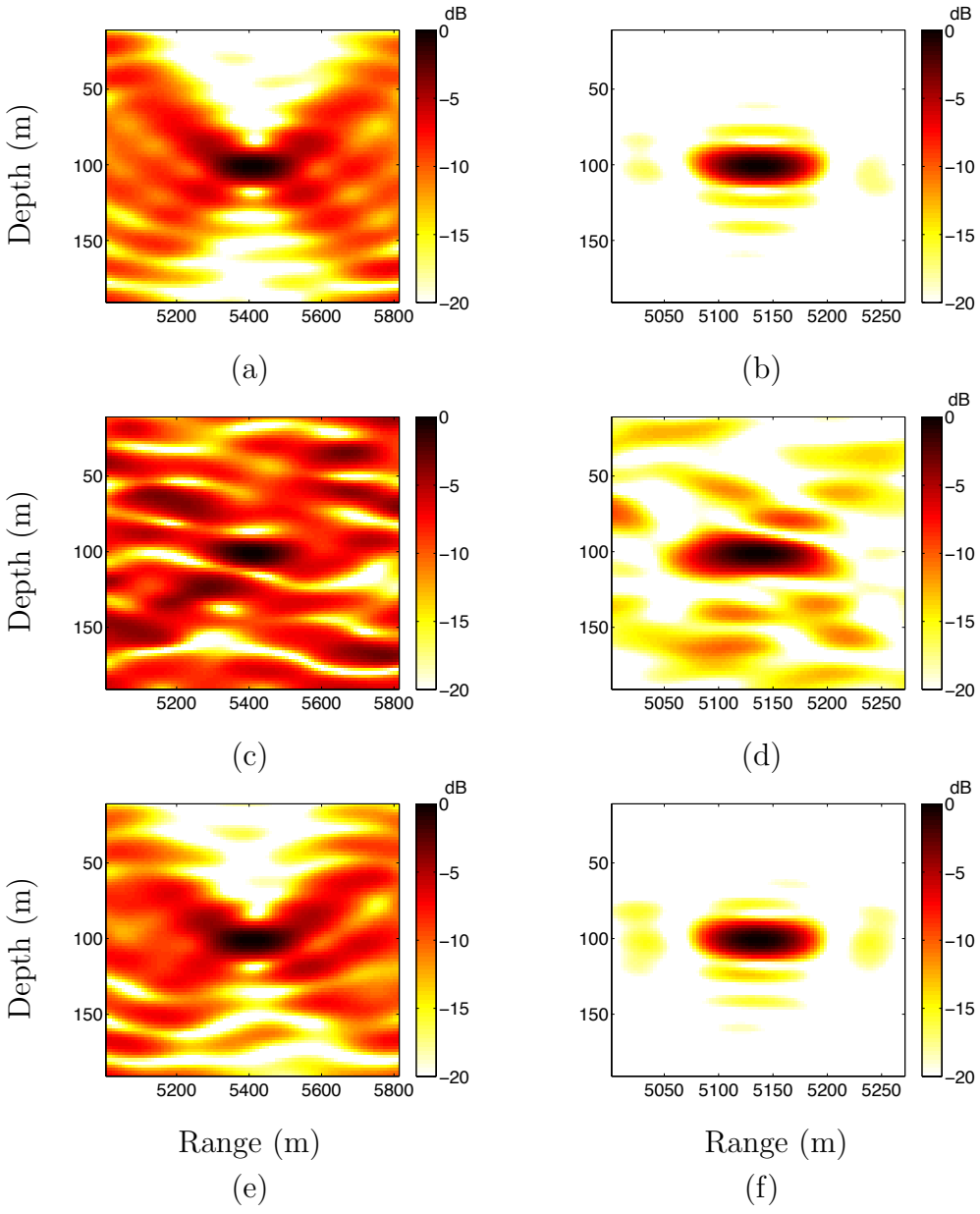


Figure 4: Single-frequency (left column) and broadband-coherent (right column) ambiguity functions. These ambiguity functions shown on the dB scale ($20\log_{10}(\cdot)$) for: (a, b) the standard MFP as described in Eqs. (18) and (24), and (c, d) cMFP as described in Eqs. (28) and (35) for the single-frequency case with $M = 10$ and broadband coherent case with $M = 2$ measurements per frequency, and (e, f) cMFP for the single-frequency case with $M = 30$ and broadband coherent case with $M = 20$ measurements per frequency.

counterpart achieved via Principal Component Analysis, but may be obtained with only incomplete knowledge of the Green's function $G_\omega(\vec{r})$. Our approach works by precomputing the backpropagation of a small number of hypothetical received signals to construct a dimension-reduced proxy for the Green's function. Then, given the actual observed data Y_ω we localize the source by finding the closest match between the received signal and the Green's function in the compressed domain. With the compressed version of $G_\omega(\vec{r})$ in hand, locating the source only requires computing a series of short inner products. In addition, the compressed version of $G_\omega(\vec{r})$ is independent of the received signal, and so can be pre-computed and re-used for later observations. As we will demonstrate in Section 2.4, this cMFP strategy is effective even when the number of pre-computed compressive measurementsis far fewer than what would be required for a full acquisition of $G_\omega(\vec{r})$ over the whole search grid area.

2.3.1 Single-Frequency cMFP

We start by discussing the single-frequency case in detail. First, we compute the *compressed* Green's function $\mathbf{\Phi} G_\omega(\vec{r})$, where $\mathbf{\Phi}$ is a $M \times N$ *encoding matrix*. Note that matrices are written in boldface letters in the remainder of this article. We construct $\mathbf{\Phi} G_\omega(\vec{r})$ by backpropagating (i.e. applying G_ω^H to) a series of *test vectors* $\Phi_1, \ldots, \Phi_M \in \mathbb{C}^N$ — we will discuss how the Φ_m are chosen in the next section.

The result of one of these computations $\Phi_m^H G_\omega(\vec{r})$ is a complex-valued acoustic field over $\vec{r}$ and requires as much effort to compute as the ambiguity function $h(\vec{r})$. We stack up the results of these precomputations as rows in the ensemble

$$\begin{bmatrix} \Phi_1^H G_\omega(\vec{r}) \\ \Phi_2^H G_\omega(\vec{r}) \\ \vdots \\ \Phi_M^H G_\omega(\vec{r}) \end{bmatrix} = \begin{bmatrix} \Phi_1 & \Phi_2 & \ldots & \Phi_M \end{bmatrix}^H G_\omega(\vec{r}) = \mathbf{\Phi} G_\omega(\vec{r}). \tag{26}$$

This ensemble gives us access to an indirect, dimension-reduced version of $G_\omega(\vec{r})$.

Given observations Y_ω, we search for the $\vec{r}$ that best explains these compressive measurementsin the compressed space. The least-squares program (17) becomes

$$\arg\min_{\vec{r}} \min_{\beta} \ \|\mathbf{\Phi} Y_\omega - \beta \mathbf{\Phi} G_\omega(\vec{r})\|^2, \tag{27}$$

which, again using the results from Appendix A.1, reduces to

$$\text{(narrowband cMFP)} \qquad \arg\max_{\vec{r}} \frac{|Y_\omega^H \mathbf{\Phi}^H \mathbf{\Phi} G_\omega(\vec{r})|^2}{\|\mathbf{\Phi} G_\omega(\vec{r})\|^2} \quad \text{where} \quad \tilde{h}(\vec{r}) = Y_\omega^H \mathbf{\Phi}^H \mathbf{\Phi} G_\omega(\vec{r}). \tag{28}$$

The function $\tilde{h}(\vec{r})$ is shown in Fig. 4.c and 4.e, and can be interpreted as a compressed version of the ambiguity function $h(\vec{r})$ in (19) shown in Fig. 4.a. The cross sections in range and depth of these ambiguity functions are shown in Fig. 5.a and 5.b.

Note that unlike the standard MFP, in this case the precomputations give us direct access to the denominator $\|\mathbf{\Phi} G_\omega(\vec{r})\|^2$ (we simply take the norms of the columns of $\mathbf{\Phi} G_\omega(\vec{r})$), and so we leave it in the optimization program. As shown in the results section, this normalization term plays an important role in improving the source location estimation when the magnitude of the Green's function varies significantly across the search grid area.

Notice that an evaluation of (28) at a point $\vec{r}$ essentially only requires an inner product between the encoded observations $\mathbf{\Phi} Y_\omega$ and the M-vector $\mathbf{\Phi} G_\omega(\vec{r})$ formed from the backpropagated fields at point $\vec{r}$ from all M test vectors.

2.3.2 Random Projections

The question remains as to how to choose the encoding matrix $\mathbf{\Phi}$ so that solution to the cMFP (28) is the same (or close to) the solution to the standard MFP (19). The corresponding least-squares problems are

$$\text{standard MFP :} \qquad \arg\min_{\vec{r},\beta} \|Y_\omega - \beta G_\omega(\vec{r})\|^2 \tag{29}$$

$$\text{cMFP :} \qquad \arg\min_{\vec{r},\beta} \|\mathbf{\Phi}\left(Y_\omega - \beta G_\omega(\vec{r})\right)\|^2. \tag{30}$$

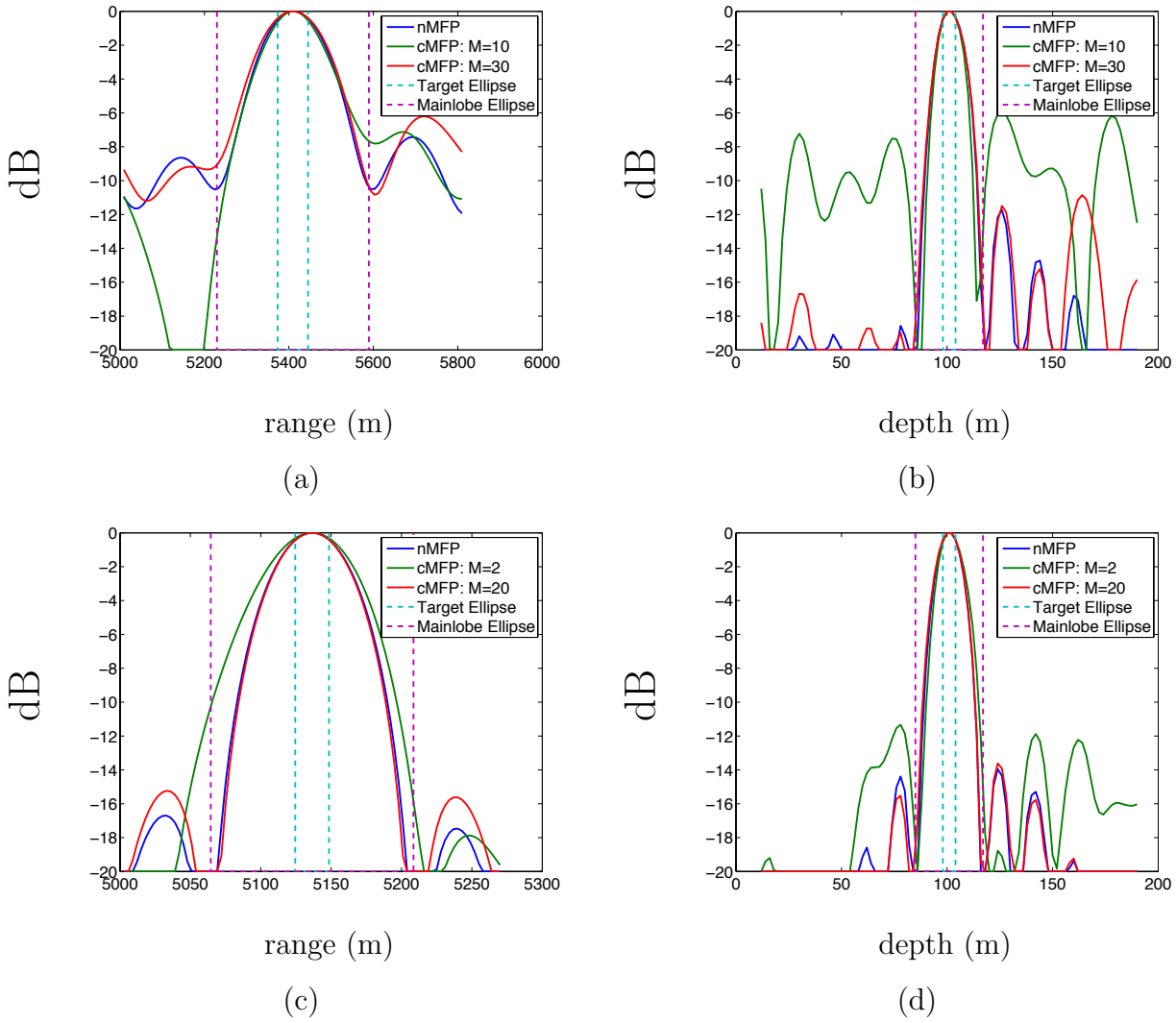


Figure 5: Cross-sections of the ambiguity functions displayed on Fig. 4: (a, b) single-frequency case described by Eqs. (18) and Eqs. (28); (c, d) broadband coherent case Eqs. (24) and (35); range (left column) and depth (right column). Here we show the normalized standard MFP (nMFP) and the cMFP (cMFP) for various values of M. The dashed lines show the boundaries for the main lobe and region of uncertainty that we are able to localize within under the presence of modest noise.

These two programs will have similar solutions if their functionals are close to one another for all values of β and $\vec{r}$. If $Y_\omega = \alpha G_\omega(\vec{r_0})$, then the performance of the cMFP will match that of the standard MFP when $\mathbf{\Phi}$ preserves the energy of the differences between the observations Y_ω and all scalar multiples of the Green's function at different points:

$$\|\mathbf{\Phi}(F_1 - F_2)\|^2 \approx \|F_1 - F_2\|^2 \quad \text{for all} \quad F_1, F_2 \in \mathcal{F} := \{F : F = \alpha G_\omega(\vec{r}),\ \alpha \in \mathbb{C};\ \vec{r} \in \mathcal{R}\}. \tag{31}$$

Essentially, we want $\mathbf{\Phi}$ to *stably embed* (i.e. preserve the distances between members of) the set $\mathcal{F}$ into $\mathbb{C}^M$.

We propose taking $\mathbf{\Phi}$ to be a random linear mapping. This choice is inspired both by classical results in theoretical computer science and from the recently developed theory of compressive sensing. In the mid-1980s, Johnson and Lindenstrauss [47] demonstrated that the distances within a finite set of n points are essentially preserved through a random projection into a space of dimension $\sim \log n$ (see also [48, 49]). Recently it has been shown that this same type of projection also embeds sparse signals into a low-dimensional subspace [4], a result which plays a key role in compressive sampling [31, 50], and are effective at reducing the dimensionality of certain types of manifolds [51].

We will discuss the particular the case where $\mathbf{\Phi}$ is a *random orthoprojection*, although the results will be almost identical for many different choices of random $\mathbf{\Phi}$ (e.g. with entries that are independent and identically distributed Gaussian or ± 1 random variables). To generate $\mathbf{\Phi}$, we simply draw an $M \times N$ matrix of independent Gaussian random variables with unit variance, orthonormalize the rows using the Gram-Schmidt (or QR) algorithm, and then multiply by $\sqrt{N/M}$. For an arbitrary

fixed vector F, the random orthoprojection $\mathbf{\Phi}$ obeys two properties[48]:

$$\mathrm{E}\left[\|\mathbf{\Phi}F\|^2\right] = \|F\|^2 \tag{32}$$

$$\mathrm{P}\left\{\left|\, \|\mathbf{\Phi}F\|^2 - \|F\|^2 \,\right| > \epsilon\right\} \leq 2e^{-\frac{M}{2\|F\|^2}\left(\epsilon^2/2 - \epsilon^3/3\right)}. \tag{33}$$

This allows us to interpret the compressed energy functional $\|\mathbf{\Phi}(Y_\omega - \beta G_\omega(\vec{r}))\|^2$ in (30) as a random process, indexed by β and $\vec{r}$, whose mean is the standard energy functional $\|Y_\omega - \beta G_\omega(\vec{r})\|^2$ in (29). At a fixed point $\beta, \vec{r}$, this random process is concentrated around its mean roughly like a Gaussian random variable with standard deviation $\sqrt{2/M}\|Y_\omega - \beta G_\omega(\vec{r})\|$. The larger we make M (the more random vectors we precompute backpropagations for), the tighter the concentration. By construction, when $M = N$, $\mathbf{\Phi}^H\mathbf{\Phi} = I$ and we have acquired a "lossless" version of the Green's function $G_\omega(\vec{r})$, meaning that the functionals are exactly equal to one another. In general, however, we will be interested in cases where there is a significant compression factor $M \ll N$ and benefit from the associated computational savings.

2.3.3 Broadband cMFP

The cMFP formulation can be readily extended to combine observations at multiple frequencies in both the incoherent and coherent cases. For frequencies $\omega_1, \omega_2, \ldots, \omega_K$, we generate a sequence of $M \times N$ random matrices $\mathbf{\Phi}_{\omega_1}, \mathbf{\Phi}_{\omega_2}, \ldots, \mathbf{\Phi}_{\omega_K}$ and backpropagate the rows of each (for a total of MK time-reversals) to acquire $\mathbf{\Phi}_{\omega_1}G_{\omega_1}(\vec{r}), \ldots, \mathbf{\Phi}_{\omega_K}G_{\omega_K}(\vec{r})$. Then given observations $Y_{\omega_1}, \ldots, Y_{\omega_K}$, we compress them by calculating $\mathbf{\Phi}_{\omega_1}Y_{\omega_1}, \ldots, \mathbf{\Phi}_{\omega_K}Y_{\omega_K}$,

and then using the compressed versions of the G_{ω_k}, we proceed as in (20) for the incoherent case

$$\begin{aligned}
\text{(incoherent cMFP)} \qquad & \arg\min_{\vec{r}} \min_{\beta_{\omega_1},\ldots,\beta_{\omega_K}} \sum_{k=1}^{K} \|\mathbf{\Phi}_{\omega_k} Y_{\omega_k} - \beta_{\omega_k} \mathbf{\Phi}_{\omega_k} G_{\omega_k}(\vec{r})\|^2 \\
& = \arg\min_{\vec{r}} \sum_{k=1}^{K} \min_{\beta_{\omega_k}} \|\mathbf{\Phi}_{\omega_k} Y_{\omega_k} - \beta_{\omega_k} \mathbf{\Phi}_{\omega_k} G_{\omega_k}(\vec{r})\|^2 \\
& = \arg\max_{\vec{r}} \sum_{k=1}^{K} \frac{|Y_{\omega_k}^H \mathbf{\Phi}_{\omega_k}^H \mathbf{\Phi}_{\omega_k} G_{\omega_k}(\vec{r})|^2}{\|\mathbf{\Phi}_{\omega_k} G_{\omega_k}(\vec{r})\|^2}, \\
& = \arg\max_{\vec{r}} \sum_{k=1}^{K} \frac{|\tilde{h}_{\omega_k}(\vec{r})|^2}{\|\mathbf{\Phi}_{\omega_k} G_{\omega_k}(\vec{r})\|^2}
\end{aligned} \tag{34}$$

and as in (23) for the coherent case:

$$\begin{aligned}
\text{(coherent cMFP)} \qquad & \arg\min_{\vec{r}} \min_{\beta} \left\| \begin{bmatrix} \mathbf{\Phi}_{\omega_1} Y_{\omega_1} \\ \mathbf{\Phi}_{\omega_2} Y_{\omega_2} \\ \vdots \\ \mathbf{\Phi}_{\omega_K} Y_{\omega_K} \end{bmatrix} - \beta \begin{bmatrix} \alpha_{\omega_1} \mathbf{\Phi}_{\omega_1} G_{\omega_1}(\vec{r}) \\ \alpha_{\omega_2} \mathbf{\Phi}_{\omega_2} G_{\omega_2}(\vec{r}) \\ \vdots \\ \alpha_{\omega_K} \mathbf{\Phi}_{\omega_K} G_{\omega_K}(\vec{r}) \end{bmatrix} \right\|^2 \\
& = \arg\max_{\vec{r}} \frac{\left| \sum_{k=1}^{K} \alpha_{\omega_k} Y_{\omega_k} \mathbf{\Phi}_{\omega_k}^H \mathbf{\Phi}_{\omega_k} G_{\omega_k}(\vec{r}) \right|^2}{\sum_{k=1}^{K} |\alpha_{\omega_k}|^2 \ \|\mathbf{\Phi}_{\omega_k} G_{\omega_k}(\vec{r})\|^2} \\
& = \arg\max_{\vec{r}} \frac{\left| \sum_{k=1}^{K} \alpha_{\omega_k} \tilde{h}_{\omega_k}(\vec{r}) \right|^2}{\sum_{k=1}^{K} |\alpha_{\omega_k}|^2 \ \|\mathbf{\Phi}_{\omega_k} G_{\omega_k}(\vec{r})\|^2}.
\end{aligned} \tag{35}$$

The incoherent and coherent case are respectively illustrated in Fig. 4.d and 4.f and in Fig. 5.c and 5.d. Note that in this coherent case, the optimization is identical in its structure to the single-frequency case. In particular, by concatenating:

$$G(\vec{r}) = \begin{bmatrix} \alpha_{\omega_1} G_{\omega_1}(\vec{r}) \\ \alpha_{\omega_2} G_{\omega_2}(\vec{r}) \\ \vdots \\ \alpha_{\omega_K} G_{\omega_K}(\vec{r}) \end{bmatrix} \quad Y = \begin{bmatrix} Y_{\omega_1} \\ Y_{\omega_2} \\ \vdots \\ Y_{\omega_K} \end{bmatrix} \quad \mathbf{\Phi} = \begin{bmatrix} \mathbf{\Phi}_{\omega_1} & & & \\ & \mathbf{\Phi}_{\omega_2} & & \\ & & \ddots & \\ & & & \mathbf{\Phi}_{\omega_K} \end{bmatrix}, \tag{36}$$

we see that the coherent broadband formulation (35) shares the same formulation as the single frequency case (28).

2.4 Numerical simulations

In this section, we present numerical experiments demonstrating that underwater acoustic sources can be localized from highly compressed versions of the Green's functions. Our cMFP results give locations estimates for single-frequency, incoherent broadband, and coherent broadband that are comparable with the traditional MFP. After the initial pre-computation (which consists of backpropagating the random codes at each frequency), the cMFP is substantially faster than the traditional MFP, requiring only a short inner product to be calculated at each search location.

The MATLAB code generating all the numerical results presented in this section is available online [1].

2.4.1 Numerical set-up

All numerical simulations were conducted using a 200m deep Pekeris waveguide and the Green's functions were computed using a standard normal mode code [46, pg 540–552]. The two dimensional search grid area in depth and range spans respectively [10m 190m], and [5000m 5810m] for the single frequency and broadband incoherent simulations. The range span for the broadband coherent simulations was reduced to [5000m 5270m] to keep constant the number of search locations over which the ambiguity functions are computed since the effective resolution of the ambiguity function in the coherent case was about 3 times higher in range (see Fig. 4 and Fig. 5). A uniformly spaced vertical line array with $N = 37$ elements spaced between 10 and 190 meters was used to sample the acoustic field. The Green's functions between each of the search locations and the receiver array (see Fig. 3) were calculated across $K = 20$ different frequencies between 141 Hz and 160 Hz (the narrowband configuration uses 150 Hz). Given the selected numerical set-up, the natural resolution in frequency of

[1] Download the code at http://users.ece.gatech.edu/~wmantzel3/cmfp/code.zip.

the computed Green's function is around 5 Hz; that is, $G_{\omega_1}(\vec{r})$ and $G_{\omega_2}(\vec{r})$ are essentially uncorrelated when $|\omega_1 - \omega_2| \geq 10\pi$. The selected sample spacing of 1 Hz falls well within this frequency resolution.

After selected a source location $\vec{r_0}$ inside the region of interest, observations at the K frequencies were simulated using the forward model, and uncorrelated zero-mean Gaussian noise was added to the result:

$$Y_{\omega_k} = \alpha_{\omega_k} G_{\omega_k}(\vec{r_0}) + Z_k, \quad Z_k \in \mathbb{C}^N, \quad Z_k \sim \text{Normal}(0, \sigma^2 I), \tag{37}$$

where each Z_k has i.i.d. Gaussian real and imaginary parts with variance $\sigma^2/2$. In all of our experiments, we set $\alpha_{\omega_k} = 1$ for all k. The signal-to-noise ratio (SNR) corresponding to noise variance σ^2 is

$$\text{SNR} \;=\; 10 \log_{10} \left(\frac{|\alpha_\omega|^2 \|G_\omega(\vec{r_0})\|^2}{N\sigma^2} \right) \tag{38}$$

in the single frequency case, and

$$\text{SNR} \;=\; 10 \log_{10} \left(\frac{\sum_{k=1}^K |\alpha_{\omega_k}|^2 \; \|G_{\omega_k}(\vec{r_0})\|^2}{KN\sigma^2} \right) \tag{39}$$

in the broadband case. Unless otherwise stated, we used an SNR of 16 dB.

Given a set of observations, we estimate the source location by solving (28) single frequency), (34) (broadband incoherent), or (35) (broadband coherent) and compare against the standard MFP formulations (19), (22), and (25) As stated, these optimizations problems are over a continuous variable $\vec{r}$ — in practice, we compute these functionals on a finite grid of points and choose the maximum from amongst these points. We used a 90×90 grid for the simulations presented below, which corresponds to 2m spacing in depth, a 9m spacing in range in the single-frequency and broadband incoherent cases, and 3m spacing in range in the broadband coherent case. We wish to emphasize that while our solution will of course lie on one of these grid points, the actual source location is chosen to be an arbitrary point.

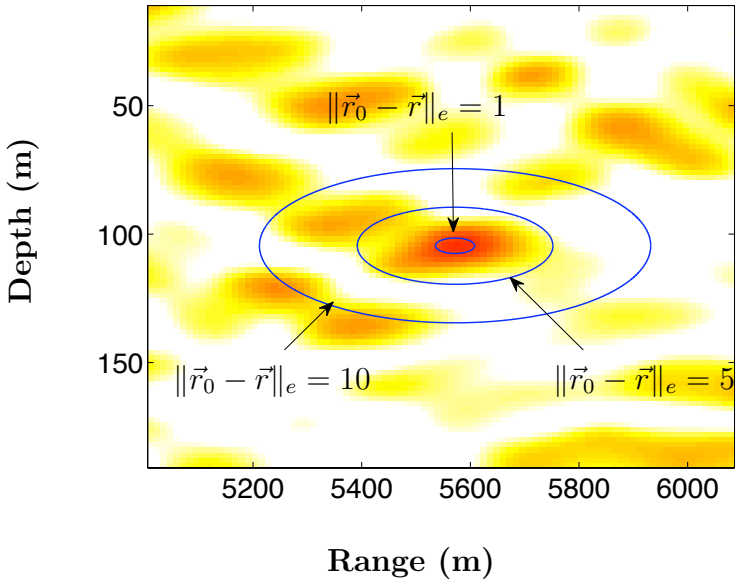


Figure 6: Definition of the elliptical distance metric used for the performance study of cMFP. Because range errors tend to be greater than depth errors in long-range localization estimates, our results use an elongated distance metric that gives greater weight to the depth than the range, leading to unit balls that are ellipses instead of circles as shown here (see Eq. (40)). The color scheme for this ambiguity surface $10\log_{10}(|h(\vec{r})|^2) = 10\log_{10}(|Y_\omega^H G_\omega(\vec{r})|^2)$ has been lightened somewhat to allow for better visibility for the overlaid ellipses.

The natural resolutions in depth and range of the ambiguity function $h_\omega(\vec{r})$ differ, as shown in Fig. 6 and Fig. 5 for a source located at $\vec{r_0} = (5540m, 100m)$ in (range,depth) for a single frequency $\omega = 300\pi$ rad/sec (150 Hz) for a source located at $\vec{r_0} = (5540, 100)$ in (range,depth). In this case, the main lobe has a width of ~ 360 m in range and ~ 32 m in depth. Again the grid spacing of $9m/2m$ in range/depth falls well within this resolution. The spatial resolution of the ambiguity surface in the selected multi-modal Pekeris waveguide is primarily a function of the source-receiver array configuration as well as the selected frequency band [26, 27] In light of these differing spatial resolutions, we use a weighted norm to report distance errors in most cases presented here in this section. The distance from a point $\vec{r_0} = (r_0^{\text{range}}, r_0^{\text{depth}})$ to the estimated source location $\hat{\vec{r}}$ is computed using the elliptical distance:

$$\|\vec{r_0} - \hat{\vec{r}}\|_{\text{e}} = \sqrt{\left(\frac{r_0^{\text{range}} - \hat{r}^{\text{range}}}{e^{\text{range}}}\right)^2 + \left(\frac{r_0^{\text{depth}} - \hat{r}^{\text{depth}}}{e^{\text{depth}}}\right)^2}. \tag{40}$$

We use $e^{\text{depth}} = 3$m and $e^{\text{range}} = 36$m for the single-frequency and incoherent cases, and $e^{\text{range}} = 12$m in the coherent case. The values of e^{depth} and e^{range} were chosen so that the contour $\{\vec{r} : \ \|\vec{r_0} - \vec{r}\|_{\text{e}} = 1\}$ was approximately the same as the isosurface of the ambiguity function at 0.9 of its maximum. Equidistant points from $\vec{r_0} = (5540, 100)$ for $\|\vec{r_0} - \vec{r}\|_{\text{e}} = 1, 5$, and 10 are shown in Fig. 6. For example, an error of 14.4 meters in range and 0.9 meters in depth translates to 0.5 units of distance error in the elleptical $\|\cdot\|_{\text{e}}$ norm.

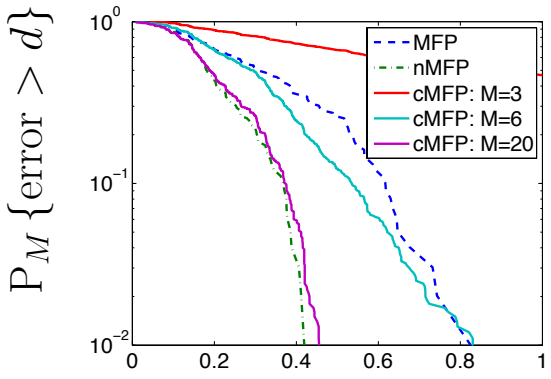
2.4.2 Localization performance of cMFP.

Fig. 7a compares the performance of cMFP (see Eq. (28)) and MFP (see Eq. (eq:amb-norm-eq:amb)) for locating a harmonic source ($f = 150Hz$). The SNR of the received data vector (see Eq. 37) was set to 16 dB. For a fixed M we aggregate performance statistics across 1000 simulations: 100 different source locations (chosen from $\mathcal{R}$ uniformly at random) and 10 different draws of the $\mathbf{\Phi}_\omega$ for each location. For each test

simulation, the error between the true and estimated target location was recorded in units of the target ellipse radius (see Fig. 6). From the results of the 1000 simulations, we calculated the empirical distance tail probability $P_M(d)$-for a given number of random backpropagations M- as the fraction of results that produces a location estimate $\hat{\vec{r}}$ with $\|\hat{\vec{r}} - \vec{r_0}\|_{\mathrm{e}} > d$. As shown, we are able to estimate the target within the unit ellipse more than 99% of the time from only $M = 6$ test vectors. Notice that the cMFP actually outperforms the unnormalized version of the MFP (from (19) above) when $M \approx 6$. This happens because the cMFP has an estimate of the normalizing factor in the denominator, as shown in (28). The cMFP is really an estimate of the normalized MFP in (18), and indeed that formulation is what the cMFP approaches as the number of random backpropagations M becomes equal to number of receivers N.

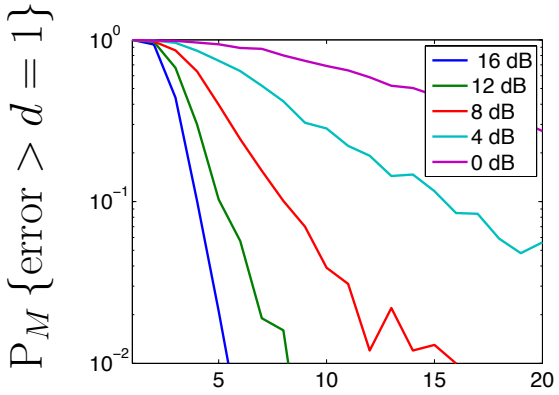
The cMFP was also tested in a variety of SNR for the single-frequency case. Fig. 7b shows the probability that the localization estimate is within the first ellipse (i.e. $d < 1$) as a function of the number of random backpropagations M. In all cases, the failure probability asymptotically decreases exponentially in the number of random backpropagations. Finally, Fig. 7c shows the tail probability of distance error for a fixed number of random backpropagations $M = 20$. As expected, the performance of cMFP gradually decrease as the SNR of the measurements is reduced from 16dB to 0dB, similarly to what occurs when using conventional MFP [28].

Fig. 8 and Fig. 9 show a similar performance study for respectively the broadband incoherent cMFP (see Eq. (34)) or broadband coherent cMFP(see Eq. (35)) formulations, including the influence of the SNR of the measurements as well as the number the number of random backpropagations M. Note that in Fig. 9, the horizontal axis is normalized differently than in the other two cases due to the different spatial resolutions of the ambiguity surfaces (see Eq. (40)). Our intentions are not to directly



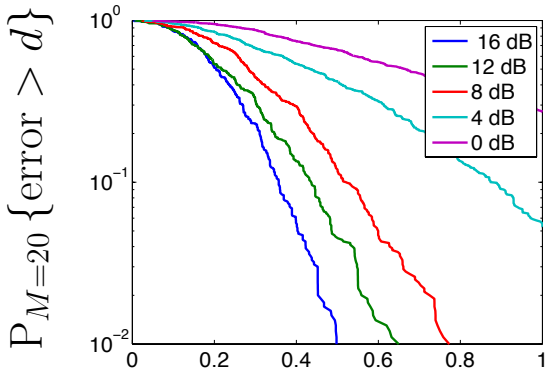
Distance d (in units of target ellipse radii)

(a)



Random Backpropagations: M

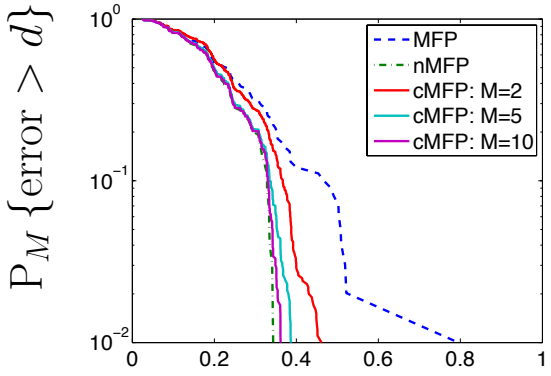
(b)



Distance d (in units of target ellipse radii)

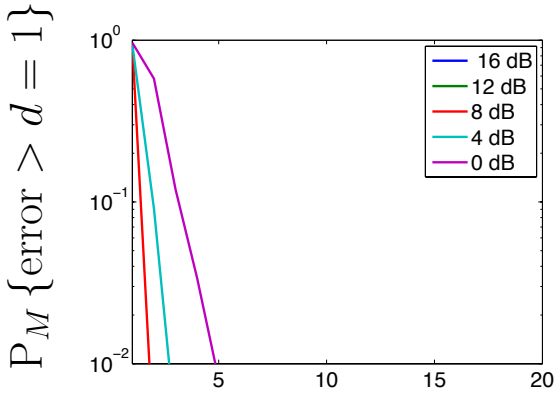
(c)

Figure 7: (a) Tail probability of distance error $\|\hat{\vec{r}} - \vec{r}_0\|_e$ (see Eq. (40)) for the single-frequency cMFP formulation (see Eq. (28)) at 150 Hz. $P_M(d)$ is the probability that the localization is worse than some distance d using M compressive measurements. The dashed lines indicate the performance under normalized and unnormalized MFP (Eq. (18) and Eq. (19)). The next two plots show results for $P_M(d)$ over various SNRs of the received signal with (b) fixing $d = 1$ and (c) fixing $M = 20$.



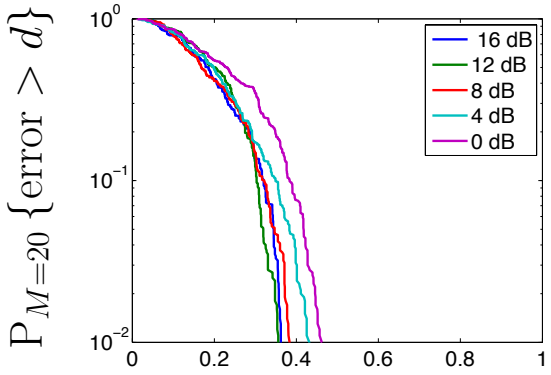
Distance d (in units of target ellipse radii)

(a)



Random Backpropagations: M

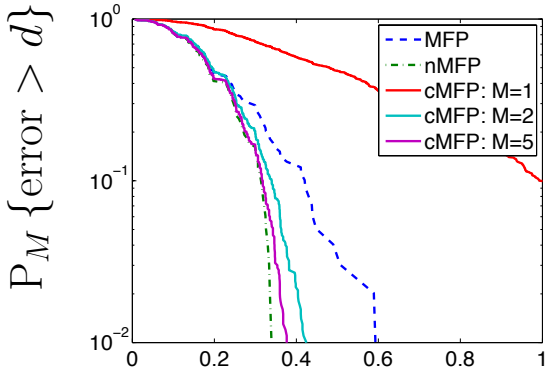
(b)



Distance d (in units of target ellipse radii)

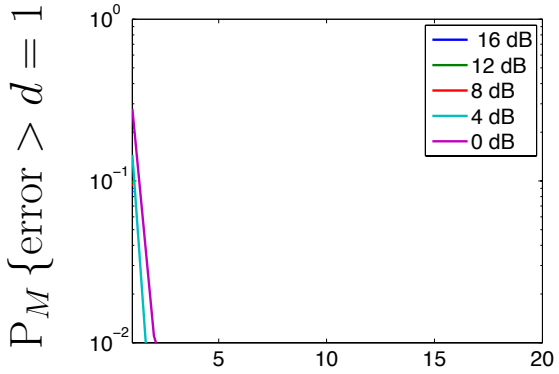
(c)

Figure 8: Same as Fig. 7 but using instead the incoherent broadband cMFP formulation (see Eq. (35))



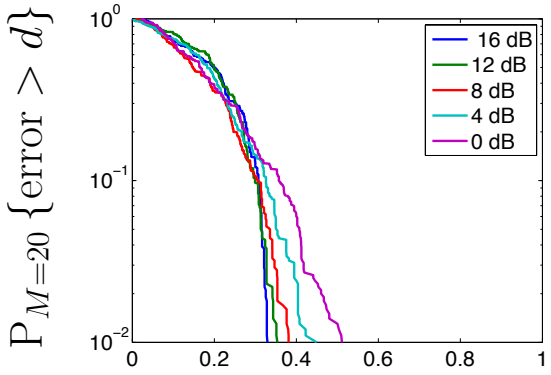
Distance d (in units of target ellipse radii)

(a)



Random Backpropagations: M

(b)



Distance d (in units of target ellipse radii)

(c)

Figure 9: Same as Fig. 7 but using instead the coherent broadband cMFP formulation (see Eq. (35))

compare the performance of each of the three cMFP formulations (the coherent localization being always better as expected), but rather to show that in each case the selected cMFP formulation performs as well as the corresponding normalized MFP formulation and better the corresponding unnormalized MFP formulation. This is especially true for the broadband coherent cMFP results as Fig. 9.a shows that with just $M = 1$ measurement per frequency, we achieve an error within 3 times what standard MFP gives us at least 90% of the time, and with $M = 2$, we fall within about 10% distance error of what MFP gives us about 99% of the time.

Furthermore, note that we do not show results for $M = 1$ for the broadband incoherent cMFP formulation (Fig. 8.a) as in this case $\mathbf{\Phi}_k G_{\omega_k}(\vec{r})$ is a scalar for each $\vec{r}$, and (34) reduces to

$$\arg\max_{\vec{r}} \sum_{k=1}^{K} \frac{|Y_{\omega_k}^H \mathbf{\Phi}_k^H \mathbf{\Phi}_k G_{\omega_k}(\vec{r})|^2}{\|\mathbf{\Phi}_k G_{\omega_k}(\vec{r})\|^2} = \arg\max_{\vec{r}} \sum_{k=1}^{K} \frac{|\mathbf{\Phi}_k Y_{\omega_k}|^2 |\mathbf{\Phi}_k G_{\omega_k}(\vec{r})|^2}{|\mathbf{\Phi}_k G_{\omega_k}(\vec{r})|^2} \tag{41}$$

$$= \arg\max_{\vec{r}} \sum_{k=1}^{K} |\mathbf{\Phi}_k Y_{\omega_k}|^2. \tag{42}$$

This optimization problem is ill-defined, as the functional does not depend on $\vec{r}$.

2.4.3 Evolution of the main lobe to side lobe ratio of the cMFP ambiguity surface.

Fig. 10 shows the logarithmic variations of the main lobe to side lobe ratio of the ambiguity surface obtained with the single frequency and broadband coherent cMFP formulations for increasing number of random backpropagations M. In each case, the displayed values represent the median value of the main lobe to side lobe ratios obtained from 1000 simulations for each value of M. Here the main lobe is defined as the maximum of the ambiguity surface $|h(\vec{r})|$ (obtained from the corresponding conventional MFP formulation, e.g. see Fig. 4a-b and Fig. 5) over the region of interest $\mathcal{R}$, and the side lobe as the maximum of $|h(\vec{r})|$ over the search area $\mathcal{R}$ excluding an ellipse E of the approximate size of the main lobe. We show the cross sections of

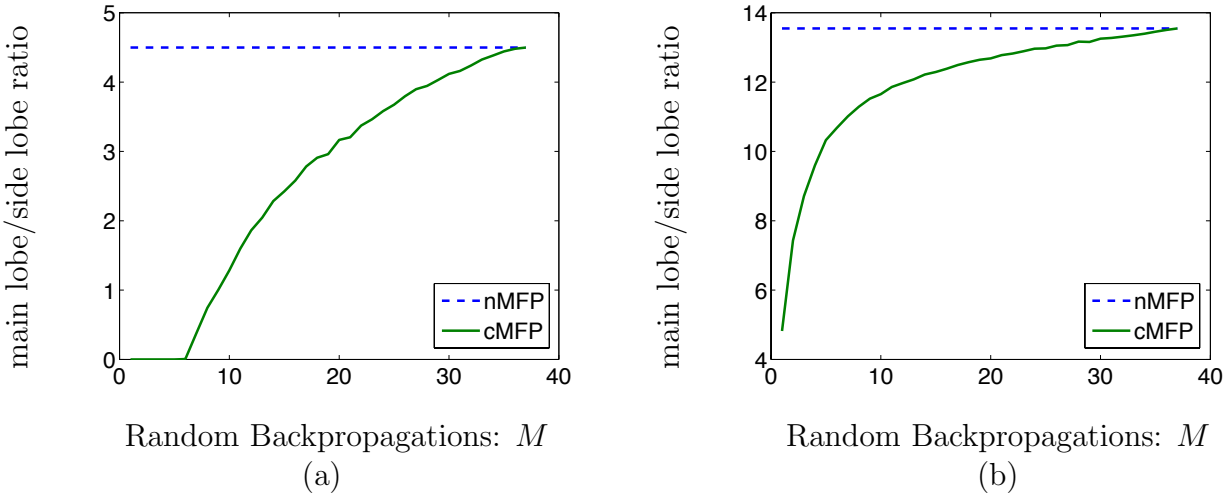


Figure 10: Evolution of the main lobe to side lobe ratio (in dB) of the estimated ambiguity surface (e.g. see Fig. 4) vs. number of random backpropagations M using either (a) the single frequency cMFP formulation at 150 Hz or (b) the broadband coherent MFP formulation (see Eq. (35)). Note that in each case the main lobe to side lobe ratio of the ambiguity surface obtained with cMFP reaches the main lobe to side lobe ratio value obtained using the corresponding nMFP formulation (dashed line) when $M = N = 37$.

the ambiguity function in Fig. 5 where we illustrate our choice of main lobe ellipse parameters that define our main lobe ellipse E. For the single frequency case, the ellipse has parameters $e^{\text{range}} = 180$ meters and $e^{\text{depth}} = 16$ meters (the broadband coherent case uses $e^{\text{range}} = 72$ meters and $e^{\text{depth}} = 16$ meters) as illustrated in Fig. 5. The logarithmic value of the main lobe to side lobe ratio is computed as:

$$20 \log_{10} \left(\frac{\max_{\vec{r} \in \mathcal{R}} |h(\vec{r})|}{\max_{\vec{r} \in \mathcal{R} \setminus E} |h(\vec{r})|} \right), \tag{43}$$

Note that for small M values, the cMFP side lobes may be significantly larger than their standard MFP counterparts. The concentration inequality (33) suggests that as M gets larger, the side lobes dampen. This behavior is observed in Fig. 10. Note that since the $\mathbf{\Phi}$ matrix is an isometry when $M = N$, the side lobes in this case are exactly the same as for the standard MFP.

2.4.4 Influence of model mismatch on the cMFP performance

Previous studies have shown extensively that one major liability of MFP is sensitivity to model mismatch which occurs when one has an incorrect model for the ocean waveguide (e.g. sound speed profile error) [27]. Since MFP exploits the knowledge of the environment (via the Green's functions), its numerical accuracy must be sufficiently accurate, to ensure accurate source localization. Here we simply ensure that the localization accuracy of cMFP remains comparable to conventional MFP in the presence of error in the sound speed value. To do so, a set of received signals with a set SNR of 16 dB were computed for a reference sound speed of 1520 m/s. The broadband coherent cMFP -using $M = 4$ random backpropagations per frequency (see (35)) and normalized MFP formulation (see (24)) were then implemented using backpropagations in a simulated environment with different nominal values for the sounds speed (between 1520 m/s and 1530 m/s) than the reference value of 1520 m/s. Fig. 11 shows that the cMFP performs substantially the same as traditional MFP, for better or for worse. We show the average distance error in actual Euclidean distance (meters) as

$$\sqrt{(r_0^{\text{range}} - r^{\text{range}})^2 + (r_0^{\text{depth}} - r^{\text{depth}})^2}, \tag{44}$$

instead of ellipse distance. The small localization error occurring even without modeling error is due to the fact that the true source location did not coincide exactly with one the grid search location $\vec{r}$.

Note also that the range error tends to dominate for sufficiently large modeling error: the slope of the displayed error values is roughly 5000/1520 (the nominal range divided by the nominal speed of sound) as we would expect because a 15 m/s error in the speed of sound of 1520 m/s causes a corresponding approximate 1% distortion in the apparent range, or 50 meters, i.e. $15/1520 \cdot 5000$.

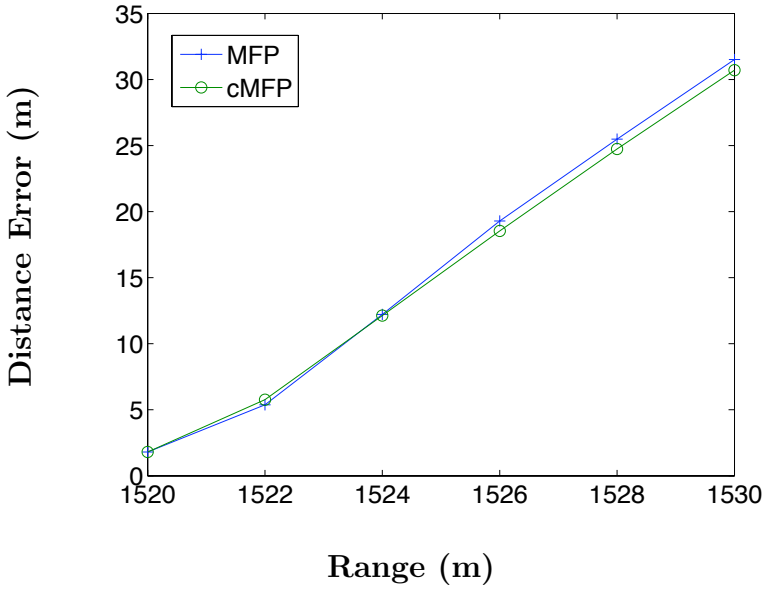


Figure 11: Evolution of the localization error for broadband coherent cMFP and corresponding conventional MFP a for increasing error of the modeled sound speed value. The correct sound speed value is $1520m/s$ here. Notice that the localization errors obtained from cMFP (circle symbols) match closely the localization errors obtained obtained from standard MFP (cross symbols).

2.4.5 Application of cMFP for tracking a moving source.

The advantage of cMFP over conventional MFP for locating a moving source along a long track is illustrated here. Fig. 12 displays the arbitrary path of a source moving along a parabolic trajectory (dashed lines). For the sake of simplicity, the Doppler effect is not accounted: this moving source scenario is simply simulated as 100 successive stationary sources located along the parabolic trajectory. For each positions, the SNR of the received signals at the vertical line array is constant and equal to 16 dB (Fig. 12.a) or 8 dB (Fig. 12.b). Conventional broadband coherent MFP is implemented by running 100 successive backpropagations per frequency over the search grid to estimate the source trajectory (see crosshair symbols). On the other hand, broadband coherent cMFP is implemented using $M = 2$ random backpropagations per frequency to estimate the same source trajectory (see cross symbols). The median

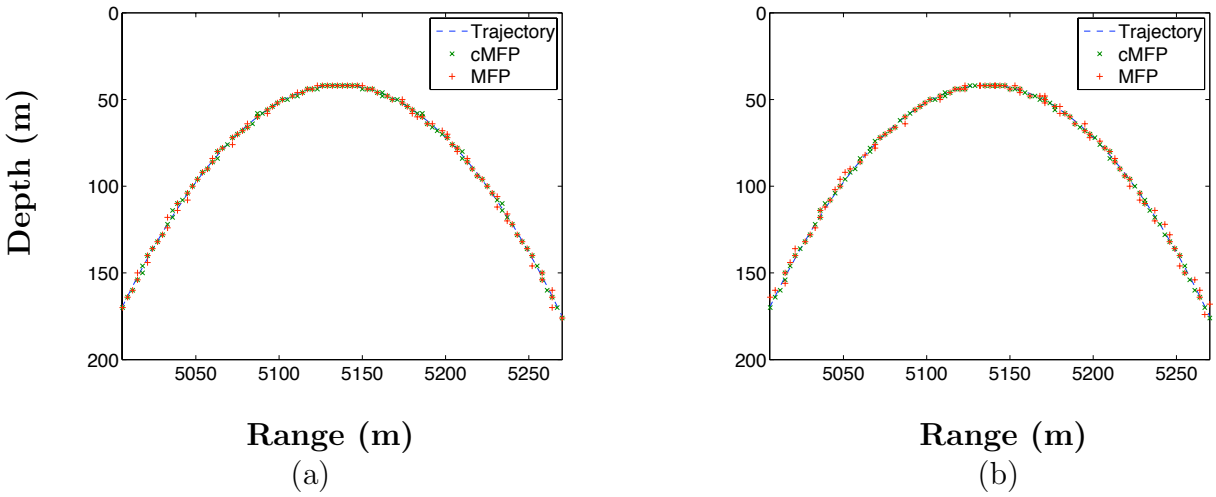


Figure 12: Tracking of a source moving along a parabolic source trajectory (dashed line) using either coherent broadband cMFP, implemented with $M = 2$ random back-propagations per frequency for the whole search grid, or using conventional broadband coherent MFP. For each of the 100 source positions, the SNR of the received signals at the vertical line array is constant and equal to (a) 16dB or (b) 8 dB.

value of the distance errors (computed from Eq. (44)) between the estimated and actual source trajectory is 1m when using both MFP and cMFP for a SNR of 16dB. A slightly higher error of 1.6m (resp. $1.1m$) for the cMFP (resp. MFP) was found for a SNR of 8dB. Overall, Fig. 12 indicates that cMFP can potentially achieve comparable source tracking performance with a significantly reduced number of simulations.

2.5 Extension to adaptive MFP

Several variants of the MFP algorithm have been proposed in the existing literature [26, 46] to enhance the robustness and performance of the basic Bartlett formulation presented above (see Eq. (19)). This can be especially beneficial in the presence of added coherent noise to the received data vector Y_ω (see Eq. (16)). To do so, these higher resolution MFP algorithms are data adaptive, but typically have also a high resolution in their environmental knowledge requirements. A commonly used adaptive MFP formulation is the Minimum Variance Distortionless Response (MVDR) formulation. The MVDR formulation adaptively constructs a replica (or weighting)

vector to yield a minimum mean square response to the recorded noise field along the receiver array while maintaining a constraint of unity processing gain for the incoming signal vector Y_ω [46, pg 540–552]:

$$|h_\omega^{MVDR}(\vec{r})|^2 = \left((G_\omega(\vec{r}))^H \mathbf{K}^{-1} G_\omega(\vec{r}) \right)^{-1}. \tag{45}$$

where $\mathbf{K}$ is the $N \times N$ is the empirical correlation matrix from multiple realizations of the noisy received data vector Y_ω:

$$\mathbf{K} = \sum_{l=1}^{L} Y_{\omega,l} Y_{\omega,l}^H. \tag{46}$$

The physical interpretation and performance analysis of the MVDR formulation (see Eq. (45)) over the simple Bartlett formulation (see Eq. (19)) have been discussed extensively in the previous literature [26, 46] and thus will not be further repeated in this article.

The previous cMFP formulation can be readily extended to handled adaptive variants of the simple Bartlett MFP algorithm as discussed in Section III.C. For instance, using Eq. (45) and by direct analogy to Eq. (28), the magnitude square of the compressive MVDR ambiguity surface is:

$$|\tilde{h}_\omega^{MVDR}(\vec{r})|^2 = \left((\mathbf{\Phi} G_\omega(\vec{r}))^H \left(\mathbf{\Phi K \Phi}^H \right)^{-1} \mathbf{\Phi} G_\omega(\vec{r}) \right)^{-1}. \tag{47}$$

So once we have computed the M tests measurements $\mathbf{\Phi} G_\omega(\vec{r})$, they can be readily applied to either the compressive adaptive MFP formulation (see Eq. (47)) or the simple Bartlett formulation (see Eq. (28)) to locate the unknown source.

2.6 Conclusions

We have shown here how dimension-reducing random projections can greatly reduce the computational cost involved with source localization via matched-field processing. When compared to the location of the maximum of the ambiguity surface obtained from conventional MFP using N distributed receivers, the localization error achieved

by cMFP scales down as square root of the number of random backpropagations M. The proposed cMFP formulation has also the added benefit to be able locate any source within the search grid area using only M random backpropagations, while conventional MFP would require at least N backpropagations to do the same. Thus cMFP provides an effective speedup factor of N/M per frequency, which can be significant when a large number of receivers N is available to locate a broadband source. Consequently this cMFP technique enables the ability to both broaden the search space and employ more sophisticated models of the Green's function, without introducing worries about sacrificing real-time performance

This compressive approach is not limited to source localization, and could be extended to a more general type of machine learning problem when matches are evaluated via inner products (or equivalently via Euclidean norms). This type of approach has the potential to substantially decrease computational complexity in these cases, while admitting a negligibly small probability of error.

CHAPTER III

MULTIPLE-SOURCE LOCALIZATION

3.1 Introduction

Matched field processing of the source signal (MFP) can be used to passively locate an acoustic source using time-series recordings taken from N receivers in the ocean [27]. The most likely position of the acoustic source is estimated by matching this acoustic response at the N receivers to the closest hypothetical modeled response generated from a candidate source location. The response at the N receivers from a candidate source location is determined by the corresponding Green's function for a given model of the ocean environment. However, localizing multiple sources in this way can present a challenge. It is often computationally prohibitive to jointly evaluate the plausibility of all combinations of potential source locations, and approaches that estimate the sources' locations independently rather than jointly can face challenges when these multiple sources "interfere" with each other, especially if the source locations are close to one another.

This chapter presents a round-robin multi-source localization scheme (ROMULO), which is outlined in Algorithm 1. This approach first makes initial estimates for all source locations, and then iteratively re-estimates each source's location in a round-robin fashion. Although the other source locations remain fixed as each source's location is re-estimated, their complex amplitude is jointly updated over all frequencies at each iteration to be maximally consistent with the observed data. A robust variation on this approach utilizes the uncertainty in the location estimate to null out a broad area around the source location estimates in an attempt to minimize their interference with the localization of the remaining sources. This is accomplished

by constructing a rank-restricted projection matrix via singular value decomposition (SVD) that effectively nulls a sum of correlation matrices. When this SVD is computationally prohibitive to compute over each iteration, a recently developed randomized method is utilized to rapidly construct an approximation of this projection matrix [41]. The main idea behind this randomized method is that for any low-rank matrix $\mathbf{Q} \in \mathbb{R}^{N \times N}$ and any i.i.d. Gaussian matrix $\mathbf{X} \in \mathbb{R}^{N \times M}$ with M less than N but larger than the rank of $\mathbf{Q}$, the range of $\mathbf{Q}$ is approximately preserved as the range of $\mathbf{QX}$, whose SVD is much easier to compute ($O(MN^2)$ total operations instead of $O(N^3)$). This round-robin multiple-source localization method is also conducive toward the use of randomly compressed Green's functions described in earlier work, and draws on the computational benefits of this approach while sacrificing only a small amount of accuracy [23].

With respect to the acoustics literature, ROMULO bears resemblance to an approach to multi-source MFP proposed by Song et al. that was inspired by the CLEAN algorithm [52, 53]. The essential differences are that they effectively keep their source estimates fixed instead of jointly estimating them at each iteration, and do not release any of the location estimates back into the residual (what they call the dirty image) for re-estimation. They also utilize a robust method that accounts for uncertainty in intermediate source estimates that is similar to ROMULO, but is a somewhat different method than the nulling projection used by ROMULO because it acts in ambiguity-function space instead of acting in data-model space. Kim et al. also propose a similar but non-iterative nulling-based approach for the 2-source case (loud source suppression in particular) using an objective function equivalent to Eq.(58) to estimate the weaker source. Earlier work by Mirkin and Sibul presents an alternating maximization approach that is substantially equivalent to the "point-nulling" version of ROMULO described below [54]. One minor distinction is that ROMULO immediately incorporates new source locations into its projection matrices, rather than

waiting until an outer loop has transpired (i.e., until all source locations have been re-estimated). Michalopoulou utilizes a Bayesian approach [55] to solve the multi-source localization problem that has essentially the same global objective function discussed in section 3.2.1 but solves it using a stochastic sampling approach. Although the focus discussed here is mainly on the standard least-squares (Bartlett) formulation of MFP, this approach is also conducive to other cost functions such as the one introduced by Westwood [56] and utilized by Neilsen [57] for multi-source localization. The latter approach uses simulated-annealing to identify sensitive parameters and jointly update all source locations, rather than updating each source's location individually. As is the case in general with simulated annealing algorithms utilizing gradient descent, there is still a risk of reaching a local minimum.

The remainder of the chapter is organized as follows. Section 3.2, after reviewing single-source matched-field processing, presents the natural extension of the objective function for the multi-source case, and illustrate how it could be approximately solved using a greedy iterative scheme such as the ones described in Appendix A.2. Then, after showing how it can be made more robust with respect to faulty intermediate source location estimates, the full algorithm is presented in the general broadband case. Then, it is shown how a compressive approach may be employed by substituting a randomized proxy Green's function for the actual Green's function. Section 3.3 shows simulated results from a shallow-water Pekeris waveguide to demonstrate the performance of this approach. Section 3.4, concludes with some remarks on implementation issues.

3.2 Matched-Field Processing

Matched-field processing estimates a sound source's location from acoustic data collected at N hydrophones by solving a parametric inverse problem, usually with least-squares (assuming Gaussian noise) [27]. For the sake of simplicity in illustrating the

crux of the matter and to avoid multiple subscripts, this exposition will begin by discussing the case where a single frequency is emitted from the source, and discuss the broadband extensions later in Section 3.2.3. In this case, a source located at some range and depth within the region of interest $\vec{r}_0 = (r_0, z_0) \in \mathcal{R}$ emits sound at a single frequency ω with unknown amplitude $\alpha \in \mathbb{C}$ (i.e., $\alpha e^{j\omega t}$) so that the received complex amplitudes at the N receivers, described by the data vector $Y \in \mathbb{C}^N$, is given as

$$Y = \alpha G(\vec{r}_0) + \eta, \tag{48}$$

where $\eta \in \mathbb{C}^N$ is a noise term, and the Green's function $G : \mathcal{R} \to \mathbb{C}^N$ is obtained from a model that approximately describes the frequency response between the source and the N receivers at frequency ω. Given a data vector Y, the source's location is estimated as a joint search for the source's amplitude $\beta \in \mathbb{C}$ and location $\vec{r} \in \mathcal{R}$ (all norms are Euclidean norms unless stated otherwise):

$$\bar{\vec{r}} = \arg\min_{\vec{r}} \min_{\beta} \|Y - \beta G(\vec{r})\|^2. \tag{49}$$

This approach generally gives an accurate estimate when the modeled Green's function G is accurate and the signal to noise ratio of the receiver is large.

Plugging in the closed-form solution to this problem with respect to β, reduces this to a maximization of the so-called Bartlett ambiguity function [27]:

$$\bar{\vec{r}} = \arg\max_{\vec{r}} \frac{|Y^H G(\vec{r})|^2}{\|G(\vec{r})\|^2}, \tag{50}$$

(where Y^H denotes the Hermitian transpose). Here and throughout, a normalized Green's function is used, so that $\|G(\vec{r})\|_2 = 1$. There is no loss of generality with this assumption, as Eq. (50) only uses the normalized version of the Green's function, making storage of the unnormalized version unnecessary.

3.2.1 Multiple Sources

The estimation of many source locations presents a challenge not present in the single-source case. This subsection discusses the nature of this challenge and how it may be

dealt with by estimating one source's location at a time.

The objective function given in Eq. (49) can readily be modified to deal with $S_0 > 1$ sources, yielding the following *global optimization* over all source amplitudes $\beta_1, ..., \beta_{S_0}$ and source locations $\vec{r}_1, ..., \vec{r}_{S_0}$:

$$\arg\min_{\vec{r}_s} \min_{\beta_s} \left\| Y - \sum_{s=1}^{S_0} \beta_s G(\vec{r}_s) \right\|^2 . \tag{51}$$

A generalization of this optimization for non-white noise was presented by Mirkin and Sibul [54].

If not for computational constraints, this would be solved by maximizing directly over all joint combinations of source locations. When the $\vec{r}_s$ are fixed, this minimization amounts to linear least squares over the β_s. However, for the full minimization over all variables this approach is typically only computationally feasible for only a few sources and suffers from the curse of dimensionality otherwise.

In the case of many sources, greedy methods may be utilized (such as orthogonal matching pursuit (OMP)) that iteratively estimate each of the source location vectors $\vec{r}_s$ one at a time [24]. For an overview of these methods, refer to Appendix A.2. The main idea as it applies here is that although it is computationally difficult to estimate all sources jointly, it is easy to estimate each source individually if the other sources' locations are known. For instance, given an initial estimate for the first source's location $\vec{r}_1$ by using standard MFP as in Eq. (49), the second source's location is estimated as:

$$\arg\min_{\vec{r}_2} \min_{\beta_2} \left\| (Y - \beta_1 G(\vec{r}_1)) - \beta_2 G(\vec{r}_2) \right\|^2 , \tag{52}$$

and then continuing onward similarly, estimate each subsequent source's location $\vec{r}_3, \vec{r}_4, ...$ sequentially using the following *greedy optimization* for all $S \geq 2$:

$$\arg\min_{\vec{r}_S} \min_{\beta_S} \left\| \left(Y - \sum_{s=1}^{S-1} \beta_s G(\vec{r}_s) \right) - \beta_S G(\vec{r}_S) \right\|^2 , \tag{53}$$

by substituting the first residual term $Y_S = Y - \sum_{s=1}^{S-1} \beta_s G(\vec{r}_s)$ for Y in Eq. (50).

Although a global search over all variables in Eq. (53) is usually intractable, one can utilize linear least squares to plug in a closed-form solution $\hat{\beta}_s$ ($s \leq S-1$) for all of the β_s that best match the existing source location estimates $\vec{r}_1$ through $\vec{r}_{S-1}$ by computing:

$$\hat{\beta} = \mathbf{G}_S^{\dagger} Y = (\mathbf{G}_S^H \mathbf{G}_S)^{-1} \mathbf{G}_S^H Y = \arg\min_{\beta} \|Y - \sum_{s=1}^{S-1} \beta_s G(\vec{r}_s)\|^2, \tag{54}$$

where $\mathbf{G}_S$ comprises the concatenation of Green's function column vectors $[G(\vec{r}_1)\ G(\vec{r}_2)\ ...\ G(\vec{r}_{S-1})]$ and $\mathbf{G}_S^{\dagger}$ is the pseudoinverse of this concatenation. In order for the residual term $Y_S = Y - \sum_{s=1}^{S-1} \beta_s G(\vec{r}_s) = Y - \mathbf{G}_S \mathbf{G}_S^{\dagger} Y$ to be of minimal norm, it must be orthogonal to $G(\vec{r}_1), ..., G(\vec{r}_{S-1})$, and therefore the projection $Y_S = \mathbf{P}_S Y$ is substituted for the smallest possible residual term as follows:

$$\arg\min_{\vec{r}_S} \min_{\beta_S} \|\mathbf{P}_S Y - \beta_S G(\vec{r}_S)\|^2. \tag{55}$$

where $\mathbf{P}_S = \mathbf{I} - \mathbf{G}_S \mathbf{G}_S^{\dagger}$ is a rank $N-(S-1)$ projection matrix satisfying $\mathbf{P}_S G(\vec{r}_s) = 0$ for all $s \leq S-1$. In effect, $\mathbf{P}_S$ attempts to "null out" from the data vector Y the influence of sources that are already estimated, so that the locations of remaining sources may be estimated with minimal interference. Here and throughout, $\mathbf{P}$ is a symmetric projection matrix acting on Green's functions or data vectors, but changes depending on the context (with subscripts added as appropriate to denote different constructions).

In a straightforward variation on Eq. (55), note that rather than fixing the β_s ($s \leq S-1$) while searching for the best β_S and $\vec{r}_S$ pair that accounts for the resulting residual term, ROMULO may implicitly include them in the optimization by substituting $\mathbf{P}_S G(\vec{r})$ for $G(\vec{r})$. This removes the portion of the candidate $G(\vec{r})$ vectors that can be accounted for by some other multiples of the existing $G(\vec{r}_s)$ terms ($s \leq S-1$), resulting in the formulation:

$$\arg\min_{\vec{r}_S} \min_{\beta_S} \|\mathbf{P}_S (Y - \beta_S G(\vec{r}_S))\|^2. \tag{56}$$

By utilizing Eq. (50), the ambiguity functions for Eq. (55) and Eq. (56) are given respectively by:

$$\arg\max_{\vec{r}} \frac{|Y_S^H \mathbf{P}_S G(\vec{r})|^2}{\|G(\vec{r})\|^2} \tag{57}$$

$$\arg\max_{\vec{r}} \frac{|Y_S^H \mathbf{P}_S G(\vec{r})|^2}{\|\mathbf{P}_S G(\vec{r})\|^2}, \tag{58}$$

and so differ only in the normalization of their denominator. We focus on the former formulation of Eq. (57) because it tends to work better in practice, but scenarios may exist where the latter version Eq. (58) outperforms, and that the mode-space version of Eq. (58) was presented in [58]. Note in particular that although the denominator of Eq. (58) may become very close to zero, potentially causing sharp singularities in the objective function, the magnitude of the denominator is always larger than the magnitude of the numerator (by the Cauchy-Schwarz inequality), so that the objective function remains bounded by 1.

Algorithm 1: Round-Robin Multi-Source Localization

ROMULO(Y,$G(\cdot)$))
Input: Data vector Y, Green's function $G(\vec{r})$ over domain $\vec{r} \in \mathcal{R}$
Output: Estimates $\{\vec{r}_1, \vec{r}_2, ..., r_{S_0}\}$ of the source locations

---

repeat
  for $S = 1$ **to** S_0

$\mathbf{Q}_S = \sum_{s \neq S} G(\vec{r}_s) G(\vec{r}_s)^H$ {Correlation Matrix}

$\mathbf{V}\Sigma^2\mathbf{V}^H = \mathbf{Q}_S$ {Eigenvalue Decomposition}

$\mathbf{P}_S \leftarrow \mathbf{I} - \mathbf{V}_{n_r}\mathbf{V}_{n_r}^H$ {Projection Matrix}

$\vec{r}_S = \arg\max_{\vec{r} \in \mathcal{R}} \frac{|Y^H \mathbf{P}_S G(\vec{r})|^2}{\|G(\vec{r})\|^2}$ {Least-Squares Estimation}

until (stopping criterion)

The algorithm is described in pseudocode in Algorithm 1 and illustrated on Fig. 16. Given some input data vector Y and corresponding model Green's function $G(\vec{r})$

defined over the region of interest $\vec{r} \in \mathcal{R}$, this approach seeks a set of source locations $\vec{r}_S$ so that the objective function in Eq. (51) is minimized. The outer loop allows each source to be re-estimated on subsequent passes. The stopping criterion may be designed to occur when the residual error (or its difference from the previous residual error) falls below a specified threshold. Alternatively, a small fixed number of outer-loop iterations are generally sufficient (five to ten, say). On the first pass through the outer loop, the sum over all $s \neq S$ is only carried out over the first $S-1$ terms $(\vec{r}_1, ..., \vec{r}_{S-1})$, and the first estimate of $\vec{r}_1$ is taken without projection (i.e., with $\mathbf{P} = \mathbf{I}$). The eigenvalue decomposition (EVD) of the rank-n_r correlation matrix $\mathbf{Q}_S$ returns a unitary eigenvector matrix $\mathbf{V} \in \mathbb{C}^{N \times N}$ [59]. The matrix $\mathbf{V}_{n_r}^H \in \mathbb{C}^{N \times n_r}$ is simply the "tall" sub-matrix of $\mathbf{V}$ corresponding to the first n_r columns (i.e., corresponding to the n_r largest eigenvalues, generally corresponding to the non-zero eigenvalues). This basic approach constructs the projection matrix through the matrix as $\mathbf{Q}_S$ given as:

$$\mathbf{Q}_S = \sum_{s \neq S} G(\vec{r}_s) G(\vec{r}_s)^H. \tag{59}$$

Constructing a projection matrix from the eigenvalue decomposition in this way is equivalent to the construction $\mathbf{P}_S = \mathbf{I} - \mathbf{G}_S \mathbf{G}_S^\dagger$ given above and will be referred to hereafter as the point-nulling method. This method is discussed in greater detail and generality in earlier work by Mirkin and Sibul [54].

This algorithm can also be viewed essentially as a continuous-time version of orthogonal matching pursuit (OMP) that has its own unique challenges. Even without noise in the signal, the source locations will generally not be accurately estimated during a first pass. For this reason, once location estimates exist for all sources by using either Eqs. (57) or (58), each estimate $\vec{r}_s$ is continuously improved in exactly the same way that the final $\vec{r}_S$ term was computed, by "nulling" out the other sources first. Because each new location estimate is at least as good as its previous estimate in terms of the residual of the objective function in Eq. (51), this leads to a monotonically decreasing error, resulting in convergence of the source location estimates.

This proposed approach is similar to greedy approaches found in compressive sensing literature used to solve sparse inverse problems, such as matching pursuit (MP) orthogonal matching pursuit (OMP) and compressive sampling matching pursuit (CoSaMP) [60, 24, 61, 62]. These algorithms choose a small number of vector elements drawn from a given dictionary whose weighted sum matches a given data vector. In fact, the first pass through Algorithm 1 to obtain initial source estimates is equivalent to OMP, because the minimization of the distance between the replica vector and the data post-projection is equivalent to a maximization of correlation between the Green's function and the least-squares residual of the data with respect to the current source location estimates.

CoSaMP [62] is especially well-suited for problems where signals x and y with disjoint support give rise to compressed vectors $\mathbf{\Phi}x$ and $\mathbf{\Phi}y$ (for random projection $\mathbf{\Phi}$) that are statistically independent from one another, an assumption whose analog is not met in this case. On the other hand, CoSaMP effectively handles much larger $\mathbf{\Phi}$ matrices than ROMULO requires (e.g., dimensions in the thousands) while the proposed approach is more tailored toward $\mathbf{\Phi}$ matrices whose dimensions are on the scale of the number of sources generally dealt with in acoustic localization (dozens rather than thousands, say). In particular, the computational complexity of ROMULO scales quadratically in the number of sources while theirs is somewhat faster, if not linear in the number of support elements. One advantage of this extra computational effort expended per-source (which is rather modest for common cases of interest) is the ability to revisit *all* source locations, rather than continuously focusing on those support elements which are contributing the least toward the minimization of the residual, a feature found (in simulations at least) to substantially improve the performance.

Here, we utilize a simulated shallow-water environment discussed in earlier work [23]. That is, a Pekeris waveguide with a depth of 200m, where the Green's functions

were computed using standard normal mode code [46] using 150 Hz in the single frequency case and 20 uniformly-spaced frequencies between 140 Hz and 160 Hz for the multi-frequency case. In this Green's function model, as with most others, the neighborhood of similar Green's functions $\{\vec{r}' : \|G(\vec{r}_0) - G(\vec{r}')\|^2 \leq \epsilon\}$ around any fixed location $\vec{r}_0$ is well-approximated by an ellipse, so the erroneous estimates from solving Eq. (49) with vector Y containing white additive noise tend to fall within an ellipse around the true source location [23]. For this reason, distance errors are reported according to an elliptical metric as later described by Eq. (74) in Sec. 3.3.

3.2.2 A Robust Variation on Multi-Source Localization

This section discusses some considerations that will motivate a variation on the minimization proposed in Eq. (56), and in particular the construction of a projection matrix that is robust against faulty location estimates.

First, consider the 2-source case where there is a source of primary interest at location $\vec{r}_2$ that is being obscured by a stronger source elsewhere at location $\vec{r}_1$ (e.g., a loud surface ship obscuring a weaker submerged source). Suppose the following observation is made:

$$Y = \alpha_1 G(r_1) + \alpha_2 G(r_2) + \eta. \tag{60}$$

While making the initial estimate of the loud source $\vec{r}_1$ (e.g., by maximizing the original ambiguity function Eq. (50)), the weaker source acts as an interferer whose energy contributes to the energy of the noise. This causes a mild error in the estimate of $\vec{r}_1$. During the iterations between estimating one source while attempting to null out the other, this procedure may end up with a situation where faulty estimates cause ROMULO to null out an insufficient amount of the interfering source, perhaps leading to an unsatisfying estimate of Eq. (56).

These considerations motivate a different construction of the projection matrix that takes into account the uncertainty inherent in the intermediate source location

estimates. Because interfering perturbations from the noise and other sources often cause each source's position to be estimated only within some ellipse around its true location, this fact is utilized in the proposed filter design. Given some estimate for $\vec{r}_1$ that is believed to be accurate to within some elliptical region of uncertainty E, and given that $\vec{r}_2$ could lie anywhere within the region of interest $\mathcal{R}$, the goal is to design a projection matrix $\mathbf{P}$ so that the projected data vector

$$\mathbf{P}Y = \alpha_1 \mathbf{P}G(r_1) + \alpha_2 \mathbf{P}G(r_2) + \mathbf{P}\eta, \tag{61}$$

contains a greatly diminished nuisance term $\mathbf{P}G(r_1)$ term while leaving the other two terms relatively unchanged so that $\vec{r}_2$ may be reliably estimated using this projected data vector.

To make this concrete, the "attenuation factor" is defined to represent the fraction of energy (between zero and one) leftover after the projection, as

$$0 < \frac{\|\mathbf{P}G(\vec{r})\|^2}{\|G(\vec{r})\|^2} = \|\mathbf{P}G(\vec{r})\|^2 < 1, \tag{62}$$

and will define the expected (or average) attenuation factor over a region A (e.g., where A is the ellipse of uncertainty E or the region of interest $\mathcal{R}$) as:

$$\begin{aligned} \mathrm{E}\left[\|\mathbf{P}\ G(\vec{x})\|^2\right] &= \mathrm{E}\left[\mathrm{Tr}(\mathbf{P}G(\vec{x})G(\vec{x})^H)\right] \\ &= \mathrm{Tr}(\mathbf{P}\,\mathrm{E}\left[G(\vec{x})G(\vec{x})^H\right]) \\ &= \mathrm{Tr}(\mathbf{P}\mathbf{Q}_A), \end{aligned}$$

where $\vec{x}$ is a uniform random variable over region A, and where:

$$\mathbf{Q}_A = |A|^{-1} \int_A G(\vec{r})G(\vec{r})^H d\vec{r}. \tag{63}$$

The specific goal is then to design a projection matrix that minimizes the expected energy of the nuisance term $\mathrm{E}\left[\|\mathbf{P}G(\vec{r}_1)\|^2\right] = \mathrm{Tr}(\mathbf{P}\mathbf{Q}_E)$ (i.e., by using $A = E$ in Eq. (63)) without significantly affecting the expected energy of the source of interest $\mathrm{E}\left[\|\mathbf{P}G(\vec{r}_2)\|^2\right] = \mathrm{Tr}(\mathbf{P}\mathbf{Q}_\mathcal{R})$. Put another way, the goal is to make the *local*

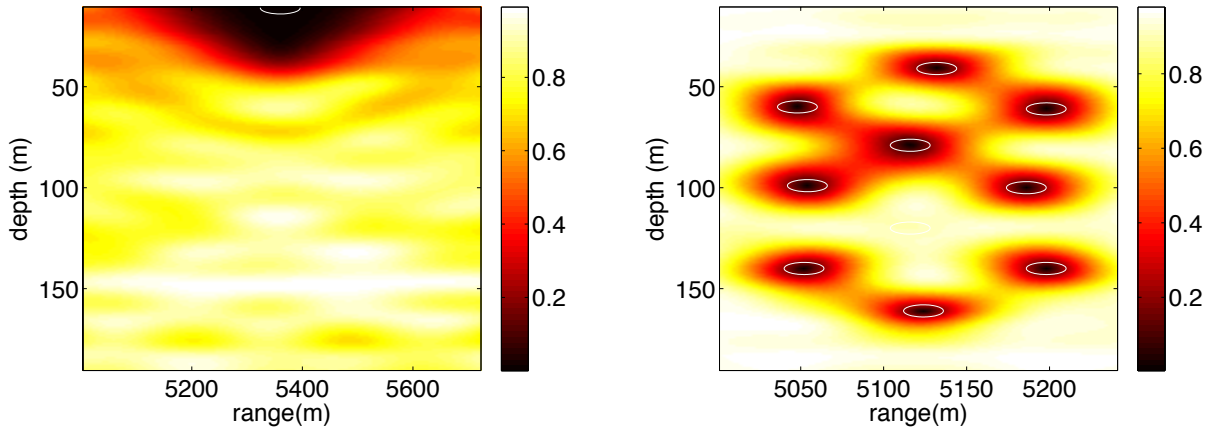


Figure 13: Attenuation factors $\|\mathbf{P}G(\vec{r})\|$ (see Eq. (62)) over range and depth for (a) single-frequency case when attempting to null out a single source located close to the ocean surface using a nulling rank n_r =5 to construct the projection matrix and (b) broadband-coherent case when attempting to null out nine sources distributed throughout the water column using a nulling rank n_r =20 to construct the projection matrix. The intended nulling region E for each source is indicated by a super-imposed line on the plot.

attenuation factor $\mathrm{Tr}(\mathbf{P}\mathbf{Q}_E)$ as small as possible while keeping the *global attenuation factor* $\mathrm{Tr}(\mathbf{P}\mathbf{Q}_{\mathcal{R}})$ as close to unity as possible.

The proposed construction of $\mathbf{P}$ takes the eigenvalue decomposition $\mathbf{V}\Sigma^2\mathbf{V}^H$ of $\mathbf{Q}_E$, and then defines the projection matrix $\mathbf{P} = \mathbf{I} - \mathbf{V}_{n_r}\mathbf{V}_{n_r}^H$ for some user-defined nulling rank n_r and where $\mathbf{V}_{n_r}$ contains the first n_r columns of $\mathbf{V}$. The two parameters that describe this projection are the size of the ellipse and the rank of the projection. The resulting attenuation factor over each point on the region of interest is shown in Fig. 13 for both the single-frequency case and the broadband coherent case discussed in greater detail later in Section 3.2.3. Here, a correlation matrix $\mathbf{Q}_E$ was constructed for a small ellipse near the top-center of the region of interest $\mathcal{R}$, then a projection matrix $\mathbf{P}$ was constructed using $n_r = 5$.

The remainder of this section considers the effect of this choice of projection matrix on the energy of the three terms in Eq. (61). As will be shown, this projection aggressively attenuates the $G(\vec{r}_1)$ term while indifferently affecting the other terms,

so that the $G(\vec{r}_2)$ and noise term η suffer only mildly from this collateral damage. First, note that this construction gives the rank $N - n_r$ projection matrix that best minimizes the expected energy of the nuisance term $G(\vec{r}_1)$:

$$\mathrm{E}\left[\|\mathbf{P}G(\vec{r}_1)\|^2\right] = \mathrm{Tr}(\mathbf{P}\mathbf{Q}_E) = \sum_{n=n_r+1}^{N} \sigma_n^2, \tag{64}$$

where the σ_n are the eigenvalues of the correlation matrix, $\mathbf{Q}_E$.

This expression summing the smallest eigenvalues represents the fraction of left-over energy after the projection averaged over all locations in the region E. The amount of energy that has been nulled out depends on the size of the region E. This fraction is illustrated in Figure 14 for several sizes of E (relative to some target ellipse) and for several choices of rank n_r. This approach is similar to the way that prolate spheroidal wave functions have been used to account for and isolate approximately time-limited and band-limited waveforms in related work [63].

For independent and identically distributed (i.i.d.) Gaussian noise, the noise term η from Eq. (61) can also be easily analyzed using rotational symmetry as follows:

$$\mathrm{E}\left[\|\mathbf{P}\eta\|^2\right] = \mathrm{E}\left[\| \begin{bmatrix} \mathbf{I}_{N-n_r} & \mathbf{0} \end{bmatrix} \eta\|^2\right] = \frac{N-n_r}{N}\,\mathrm{E}\left[\|\eta\|^2\right], \tag{65}$$

for rank $N - n_r$ projections, yielding an attenuation factor of $1 - n_r/N$. By rotational symmetry, we mean that $\mathbf{G}\eta$ is equal in distribution to η for any unitary $\mathbf{G}$.

To build a reasonable lower bound on the attenuation factor of the source to be estimated, $\mathbf{P}G(\vec{r}_2)$, a correlation matrix $\mathbf{Q}_{\mathcal{R}}$ over the entire region of interest $\mathcal{R}$ instead of the neighborhood E is constructed. A similar approach to the one used before is then used to bound the attenuation:

$$\begin{aligned} \mathrm{E}\left[\|\mathbf{P}\ G(\vec{r}_2)\|^2\right] &= \mathrm{Tr}(\mathbf{P}\mathbf{Q}_{\mathcal{R}}) \\ &= \mathrm{Tr}((\mathbf{V}_{\mathcal{R}}^H \mathbf{P} \mathbf{V}_{\mathcal{R}})\mathbf{\Sigma}_{\mathcal{R}}) \\ &\geq \sum_{n=n_r+1}^{N} \sigma_n^2 \\ &\geq 1 - n_r\sigma_1^2, \end{aligned} \tag{66}$$

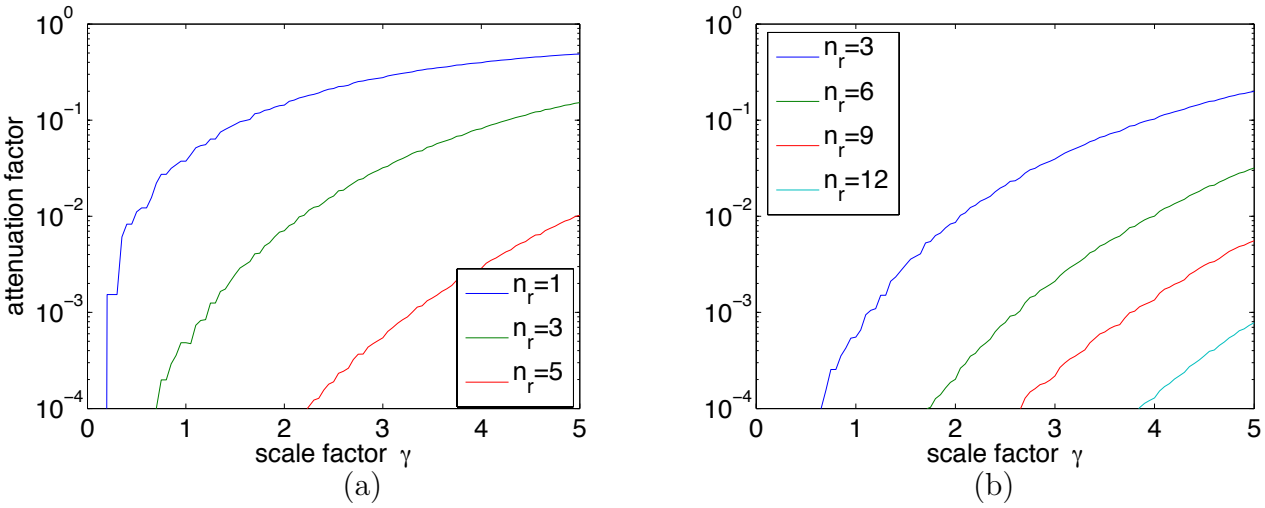


Figure 14: Local attenuation factor resulting vs. size of the ellipse-shaped nulling area for various values of the nulling rank n_r used to construct the projection matrix. The size of the nulling area, defined as $\{\vec{r} : d(\vec{r}, \vec{r}_1) \leq \gamma\}$ using the elliptical metric defined in Eq. (74), was quantified by a single factor γ which was used to scale up both major and minor axis of the ellipse. a) single-frequency case. b) broadband coherent case.

where the second-to-last inequality was changed from the equality used earlier, and the last inequality comes from the monotonicity of $\mathbf{Q}_{\mathcal{R}}$'s eigenvalues σ_n^2 and the fact that $\sum_{n=1}^{N} \sigma_n^2 = 1$ (because $\|G(\vec{r})\| = 1$). It turns out that this simple bound is actually quite accurate because most of the energy resides in eigenvalues that vary very little (well within a factor of 2 for the proposed model, empirically speaking). The quantities $\sum_{n=n_r+1}^{N} \sigma_n^2$ and $1 - n_r\sigma_1^2$ are compared in Fig. 15 using a projection ellipse of radius $\frac{1}{2}$ to illustrate the tightness of this approximation, including the broadband-coherent case discussed in more detail later in Section 3.2.3. Note that although the nominal dimension is $N = 37$ in the single-frequency case and $N = 37 * 20 = 740$ in the multiple-frequency (coherent) case (for the particular channel considered), the effective dimension is much lower, with 99% of the energy being contained in the first 12 and 54 eigenvalues for the single-frequency and coherent cases. Note also that this quantity is

Although the choice of construction for this projection matrix from the subspace

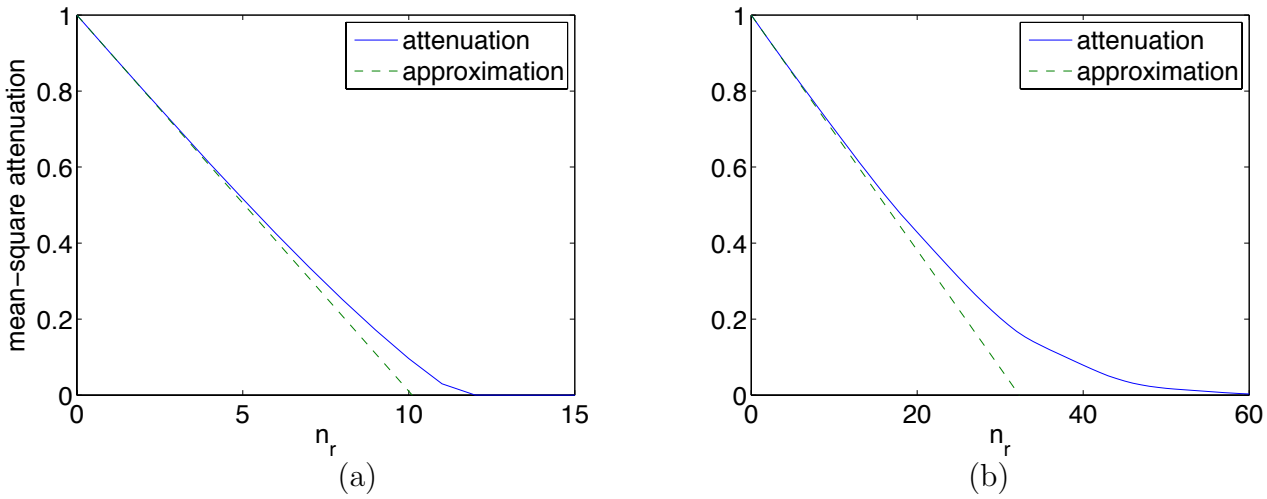


Figure 15: Shown in the solid line for the (a) single-frequency and (b) broadband-coherent cases is the attenuation factor $\mathrm{Tr}(\mathbf{PQ}_{\mathcal{R}})$ over the entire region of interest $\mathcal{R}$ under a rank-n_r nulling projection over the ellipse $\{\vec{r} : d(\vec{r}, \vec{r}_1) \leq \frac{1}{2}\}$ using the elliptical metric defined in Eq. (74). The dashed line shows the simple linear lower-bound approximation $1 - n_r\sigma_1$.

spanned by the first few principal components of the correlation matrix $\mathbf{Q}$ is relatively simple and intuitive, there is a stronger justification. This justification depends on assumptions that approximately hold in many cases that discussed in greater detail in Appendix A.3. In this Appendix section, we show how the apparently simple heuristic of eigenvalue-thresholding relates to more principled approaches such as the one proposed by Vaccaro et al. [64].

To summarize the robust variation of the general case where $S_0 \geq 2$, ROMULO continues to operate as described in Algorithm 1 with the following two modifications. First, the correlation matrix $\mathbf{Q}_S$ is defined over the union of all target ellipses except for the source location $\vec{r}_S$ currently being estimated:

$$\mathbf{Q}_S \triangleq \sum_{s \neq S} \mathbf{Q}_{E_s} = \int_{\{\bigcup_{s \neq S} E_s\}} G(\vec{r}) G(\vec{r})^H d\vec{r}, \tag{67}$$

where $\bigcup$ is the set-union operation. Second, the eigenvalue decomposition $\mathbf{V}\mathbf{\Sigma}^2\mathbf{V}^H$ becomes the eigenvalue decomposition of the closest rank-n_r matrix to $\mathbf{Q}_S$ (i.e., after eigenvalue truncation). Note that as the size of the ellipse approaches zero, this

reduces to the non-robust case as in Eq. (59) discussed in the previous section, so that this variation is a generalization of that approach.

3.2.3 Extension to Broadband MFP

The extension to the "broadband" case involves discretizing a frequency range of interest into a set of frequencies $\omega_1, ..., \omega_K$, observing a data vector $Y_k \in \mathbb{C}^N$ of complex amplitudes for each of the K frequencies, and solving a least-squares problem over the source's amplitudes β_k and location $\vec{r}$.

3.2.3.1 Incoherent MFP

In the general case, there is no prior information of the complex amplitudes of the source's signal (e.g. if the source is a a random radiator). Assuming When estimating the first source's location, the objective function is formed by incoherent summation over the K selected frequencies and thus becomes[27]:

$$\bar{\vec{r}} = \arg\min_{\vec{r}} \min_{\beta_k} \sum_{k=1}^{K} \|Y_k - \beta_k G_k(\vec{r})\|^2. \tag{68}$$

For each subsequent source's localization when some sources' locations are known, we utilize a series of projection matrices $\mathbf{P}_k$, constructed on a frequency-by-frequency basis exactly as in the single frequency case (e.g., by Eq. (67)). Then, Eq. (68) is modified as before by using these nulling projection matrices:

$$\arg\min_{\vec{r}\in\mathbb{R}^2} \min_{\beta_k\in\mathbb{C}} \sum_{k=1}^{K} \|\mathbf{P}_k Y_k - \beta_k G_k(\vec{r}))\|^2 = \arg\max_{\vec{r}\in\mathbb{R}^2} \sum_{k=1}^{K} \frac{|Y_k^H \mathbf{P}_k G_k(\vec{r})|^2}{\|G_k(\vec{r})\|^2}. \tag{69}$$

This objective function is referred to as broadband-incoherent MFP (or incoherent MFP for short).

3.2.3.2 Coherent MFP

There is the opportunity to do better than the incoherent case when the source complex amplitudes over the K frequencies are known up to some common multiplicative constant, known as broadband-coherent MFP (or coherent MFP for short)[27]. Here,

the measurement vectors and the Green's functions across all frequencies are simply stacked to achieve a much higher ambient dimension. Here, instead of the need to null for each frequency specifically, one may instead null across all frequencies jointly by constructing the correlation matrix in $NK \times NK$ space instead of $N \times N$ space, constructing the large $\mathbf{P}$ appropriately.

Specifically, when the source amplitudes $\alpha_1, ..., \alpha_K$ are known up to some unknown multiplicative constant β, the minimization is constructed as:

$$\arg\min_{\vec{r}} \min_{\beta} \left\| \mathbf{P} \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_K \end{bmatrix} - \beta \begin{bmatrix} \alpha_1 G_1(\vec{r}) \\ \alpha_2 G_2(\vec{r}) \\ \vdots \\ \alpha_K G_K(\vec{r}) \end{bmatrix} \right\|^2 = \arg\max_{\vec{r} \in \mathbb{R}^2} \frac{|\bar{Y}^H \mathbf{P} \bar{G}(\vec{r})|^2}{\|\bar{G}(\vec{r})\|^2}, \tag{70}$$

and note that this has an equivalent form to the single-frequency case in (55) when treating the stacked $\alpha_k G_k(r)$ terms as a single $\bar{G}(\vec{r})$ term and similarly treating the stacked Y_k terms as a single $\bar{Y}$ term (with dimension NK instead of N).

This coherent approach yields a higher dimension to start with (approximately 54 for this case, as shown in Fig. 15) so that a larger number of degrees of freedom are able to be removed via the nulling projection without suffering adverse effects. While degrees of freedom are scarce in the single-frequency case, in the coherent case, there are many extra degrees of freedom that may be used for nulling. The attenuation factors induced by the projection matrix (see Eq. (62)) at each location for the broadband coherent case are illustrated in Figure. 14b.

3.2.4 Extension to Compressive MFP

Earlier work demonstrated how a series of $M < N$ randomized backpropagations - using rows of a random projection matrix $\mathbf{\Phi} \in \mathbb{R}^{M \times N}$ as weighting vectors of the array elements for the backpropagation - could be used to construct a dimension-reduced proxy of the Green's function $\tilde{G}(\vec{r}) = \mathbf{\Phi} G(\vec{r})$[23]. This compressive MFP approach

allows for a more computationally efficient implementation of Eq. (49) by reducing the required number of backpropagations by a factor N/M . The entries of this random matrix are drawn independently from a Gaussian distribution, and then the matrix is constrained to be orthonormal, satisfying $\mathbf{\Phi}\mathbf{\Phi}^H = \mathbf{I}$ (e.g., by Q-R decomposition). The multi-source localization method discussed here (along with most other variations on MFP) can be easily adapted to compressive MFP (cMFP) by substituting the compressed proxies $\mathbf{\Phi}Y$, $\mathbf{\Phi}G(\vec{r})$, and M for their classical counterparts Y, $G(\vec{r})$, and N:

$$\mathbf{\Phi}Y \leftarrow Y \tag{71}$$

$$\mathbf{\Phi}G(\vec{r}) \leftarrow G(\vec{r}) \tag{72}$$

$$M \leftarrow N. \tag{73}$$

The single-frequency, incoherent, and coherent cases in Equations (56), (69), and (70) respectively are easily modified via this substitution. For example, the correlation matrix for a source believed to lie within ellipse E is constructed as: $\mathbf{Q}_E = \mathrm{E}\left[(\mathbf{\Phi}G(\vec{x}))(\mathbf{\Phi}G(\vec{x}))^H\right]$ for the vector $G(\vec{x})$ drawn randomly from the Green's function in that region. The entries of $\tilde{Y} = \mathbf{\Phi}Y$ are also called "compressed measurements", as if they were obtained by a random compressing projection $\mathbf{\Phi}$. However, this compression operation is actually introduced after the data vector Y has been measured in order to facilitate compressive matched field processing against the more easily obtained compressed Green's function $\tilde{G}(\vec{r}) = \mathbf{\Phi}G(\vec{r})$. In particular, $\mathbf{\Phi}G(\vec{r}) = (G(\vec{r})^H\mathbf{\Phi}^H)^H$ can be computed using M randomized backpropagations applying the adjoint operation G^H to each of the M rows of the $\mathbf{\Phi}$ matrix.

The primary source's location may be recovered via cMFP by nulling the interfering source below the additive noise level, provided that an extra number of randomized backpropagations are taken that will provide the buffer of necessary extra degrees of

freedom that may be removed during the nulling process. Then, in principle, ROMULO should be able to recover the primary source, where the attenuated interfering source acts as additive noise that combines with the existing white noise.

3.3 Numerical Results

This section demonstrates the efficacy of the proposed approach by simulating the localization of (1) two sources in the single-frequency case using Eq. (58), and (2) ten sources in the coherent case using Eq. (57). Specifically, numerical experiments are presented hereafter to quantify:

- the spatial resolution of the proposed approach for localizing two distinct sources accurately as they get closer together,
- the benefits of robust nulling over a region of the search area, the sacrifice in accuracy made when using the proposed greedy search method (ROMULO) instead of the relatively infeasible global search,
- the computational efficiency achieved when using randomized backpropagations (especially in the broadband regime) causing only relatively small loss of localization accuracy,
- the effectiveness of the greedy receiver-space nulling approach – used by the ROMULO algorithm – when compared to an alternative greedy ambiguity-space nulling approach. [52].

All numerical simulations were conducted using a 200m deep Pekeris waveguide and the Green's functions were computed using a standard normal mode code [46]. The configuration of the acoustic environment largely matches earlier work by Mantzel et al. [23]. A uniformly spaced vertical line array with $N = 37$ elements spaced between 10 and 190 meters was used to sample the acoustic field. The Green's

functions between each of the search locations and the receiver array were calculated across $K = 20$ different frequencies between 141 Hz and 160 Hz (the narrowband configuration uses 150 Hz). The region of interest (i.e. search area) was $\mathcal{R} = [5000\text{m} \;\; 5720\text{m}] \times [10\text{m} \;\; 190\text{m}]$ for the single frequency case, and was reduced to $\mathcal{R} = [5000\text{m} \;\; 5240\text{m}] \times [10\text{m} \;\; 190\text{m}]$ for the broadband coherent case due its higher spatial resolution for locating sound sources[27]. In both cases, $\mathcal{R}$ was discretized into 120 points in range and 180 points in depth so that the spatial resolution was 6 meters in range for the single-frequency case and 2 meters in range for the coherent case.

The following default parameters were used unless otherwise specified. The noise amplitude was 20 dB below the weakest source amplitude in the coherent case and 40 dB below the weakest source amplitude in the single-frequency case. All source amplitudes are set equal by default. The locations of all sources are independently drawn uniformly from $\mathcal{R}$. The coherent tests were taken with 10 sources with the fixed (but unknown to the algorithm) locations (depicted in Fig. 19) and the single-frequency tests were taken with 2 sources. In order to prevent the sources (whose pairwise distances were fixed) from having a fixed distance from the nearest gridpoint in all of the simulations, a small rigid random translation (on the order of magnitude of the grid spacing, roughly a meter) was added to the locations of the sources at each of the 1000 simulations. $M = 4$ randomized backpropagations per frequency were used in the coherent case and $M = 10$ randomized backpropagations were used in the single-frequency case. The case when $M = 37$ corresponds with the classical (uncompressed) approach. Regarding the parameters of the ROMULO algorithm, $10S_0$ iterations were used in a round-robin approach so that each source location was estimated in 10 total passes. For the projection matrices, $n_r = 2S_0$ in the coherent case and $n_r = \min(\lfloor M/2 \rfloor, 5)$ in the single-frequency case unless otherwise noted.

Fig. 16 illustrates the specific iterations of ROMULO (see Algorithm 1) for the coherent 10-source case using $M = 4$ randomized backpropagation per frequency.

The projection matrix was computed using a nulling rank of $n_r = 15$. Note that an approximate estimate for the location of five of the sources is obtained after 5 iterations as shown on Fig. 16c. The estimated locations for all $S_0 = 10$ sources are further refined after completing all 10 iterations as depicted in Fig. 16d.

Following earlier work [23], the distance $d(\vec{r}_1, \vec{r}_2)$ between two grid point location within the search area $\mathcal{R}$ is computed using an elliptical norm weighted with $(z_e = 3\text{m})^{-1}$ in the depth direction and $(r_e = 36\text{m})^{-1}$ in the range direction for the single-frequency case (or $(r_e = 12\text{m})^{-1}$ in the range direction for the broadband coherent case) such that:

$$d(\vec{r}_1, \vec{r}_2) = \sqrt{\left(\frac{r_1 - r_2}{r_e}\right)^2 + \left(\frac{z_1 - z_2}{z_e}\right)^2}, \tag{74}$$

where $\vec{r} = (r, z)$ is the ordered pair of range and depth and (r_e, z_e) are the ellipse parameters. The "target ellipse" E is then defined as the unit ball based on Eq. (74):

$$E = \{\vec{r} : d(\vec{r}, \vec{r}_s) < 1\}, \tag{75}$$

where $\vec{r}_s$ represents the actual source's location or its estimate. For example, the correlation matrices $\mathbf{Q}_{E_s}$ for a given source's location estimate $\vec{r}_s$ are constructed over the half-unit ball $\{\vec{r} : d(\vec{r}, \vec{r}_s) \leq 1/2\}$. When comparing all of the source estimates to ground truth, the aggregated distance error reported is the maximum distance $d(\vec{r}_s, \vec{r}_s{}')$ over all source indices s. Here, the labeling is chosen via a modified Hungarian algorithm that minimizes this maximum distance over all permutations of re-assignment [65].

3.3.1 Performance study

In general, it is difficult to localize a weak source that that is in the vicinity of another louder source. Fig. 17 illustrates this difficulty by showing the empirical probability of localization of the weak source as a function of its location. Here, the dominant source remains at a fixed location at the top-center of $\mathcal{R}$ ($(5360, 20)$ for the single-frequency

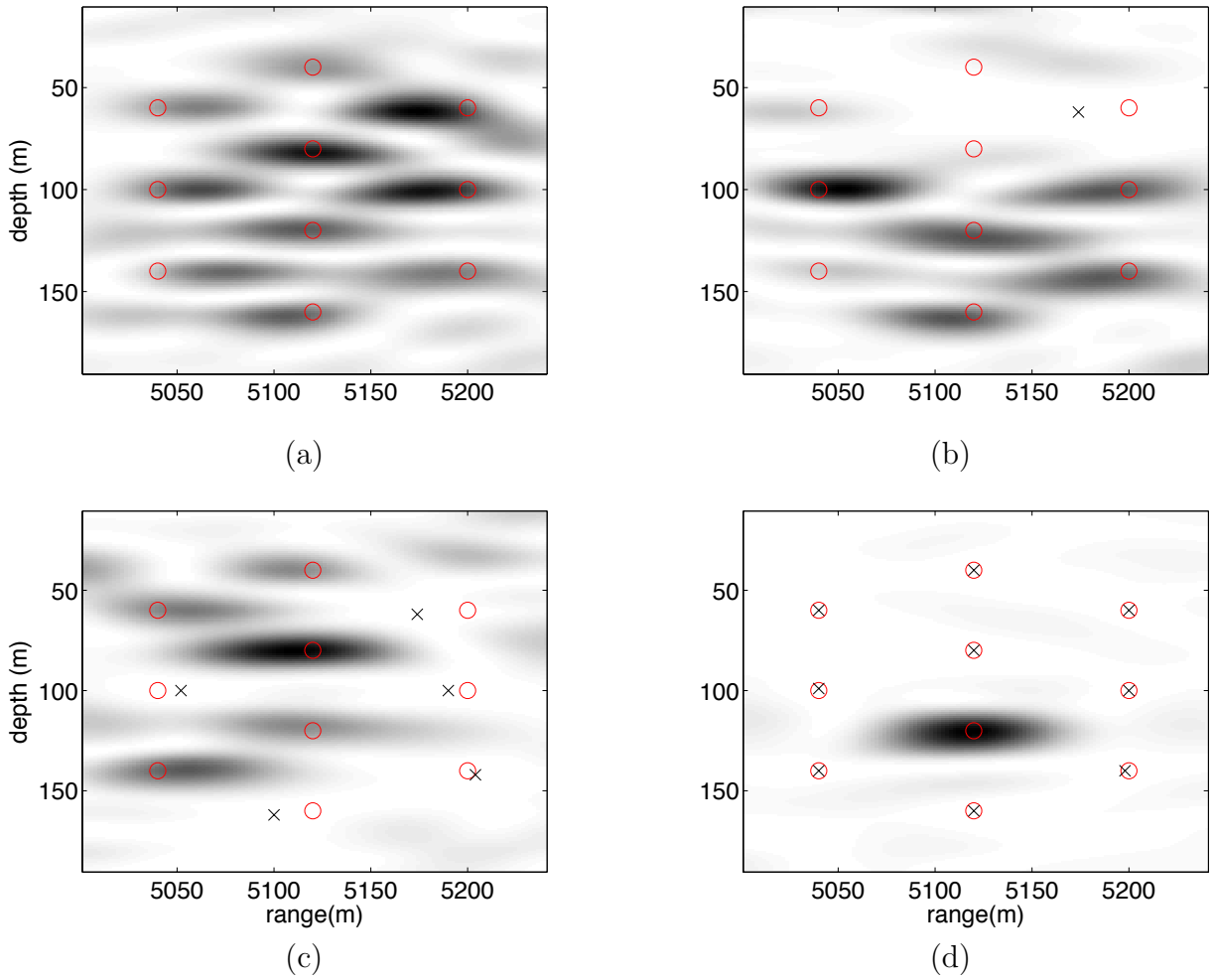


Figure 16: An illustration of the evolution of the source location estimates for the broadband-coherent case with a nulling rank of $n_r = 20$, $S_0 = 10$ acoustic sources, and $M = 4$ randomized backpropagations per frequency. Here the actual source locations are shown in circles and the estimates that are currently being nulled are shown using "x" symbols. First $\vec{r}_1$ is estimated using (a) the original ambiguity function in Eq. (50). Then, after constructing the appropriate projection $\mathbf{P}$ from $\vec{r}_1$, $\vec{r}_2$ is estimated using (b) the projected ambiguity function as in Eq. (58). The (c) pane shows this process after 5 iterations so that 5 sources are attempted to be nulled out, and pane (d) shows this process after 100 iterations, so that each of the 10 source locations have been estimated 10 times. In all cases, the compressed proxies $\mathbf{\Phi}Y$ and $\mathbf{\Phi}G(\vec{r})$ were used in place of Y and $G(\vec{r})$, corresponding to a compression ratio of $M/N = 4/37$.

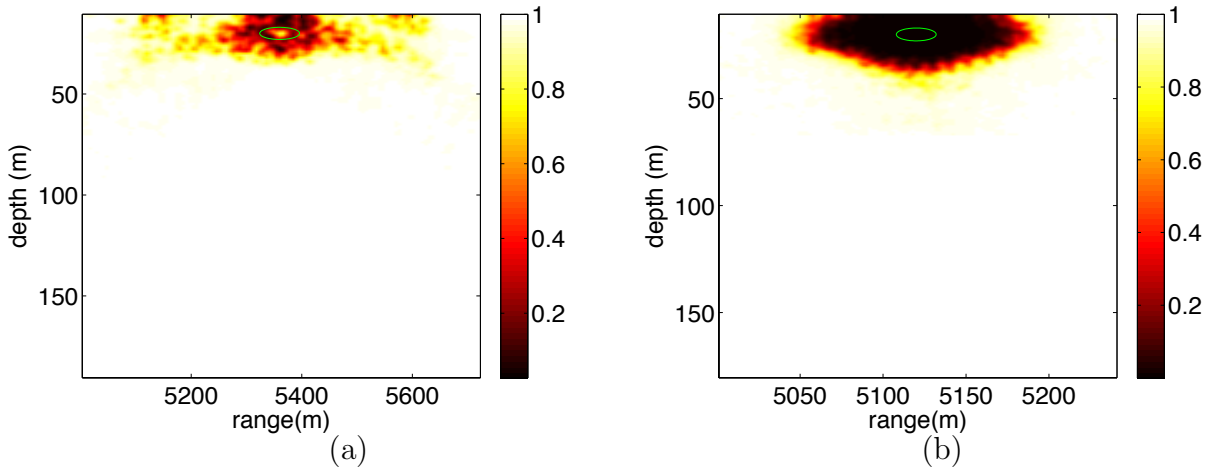


Figure 17: The probability of localizing both sources to within the target ellipse (shown superimposed) for the (a) single-frequency and (b) coherent cases as a function of the second source's location when the primary source is located in the top-center of their respective regions of interest $\mathcal{R}$ (i.e., (5360m, 20m) for (a) and (5120m, 20m) for (b)).

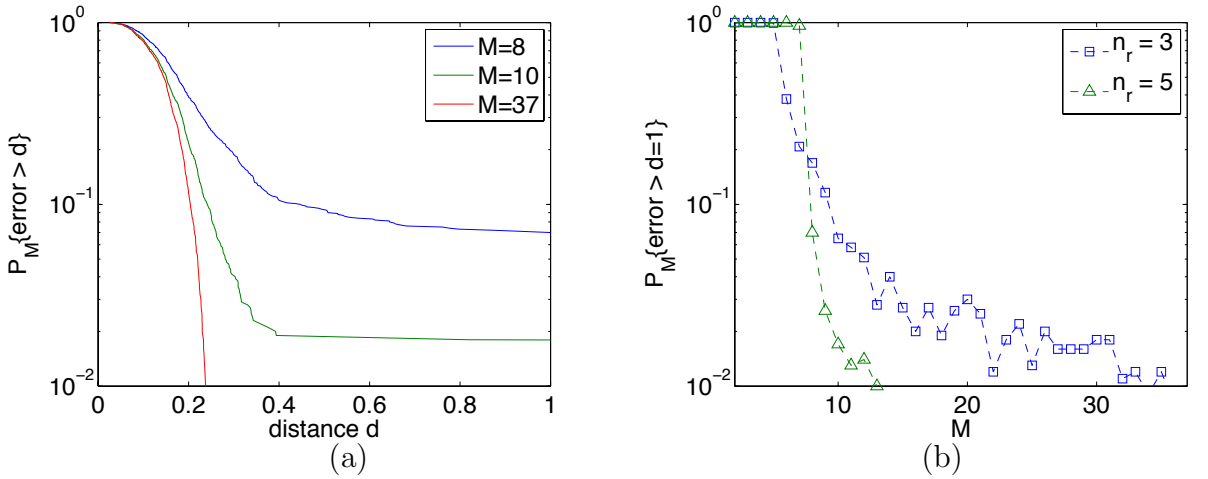


Figure 18: Single-frequency (a) tail probabilities of distance errors using a nulling rank $n_r = 5$ and $M \in \{8, 10, 37\}$ randomized backpropagations, and (b) probability that the location estimate is outside the target ellipse for increasing number of M randomized backpropagations, using two different values for the nulling rank $n_r \in \{3, 5\}$.

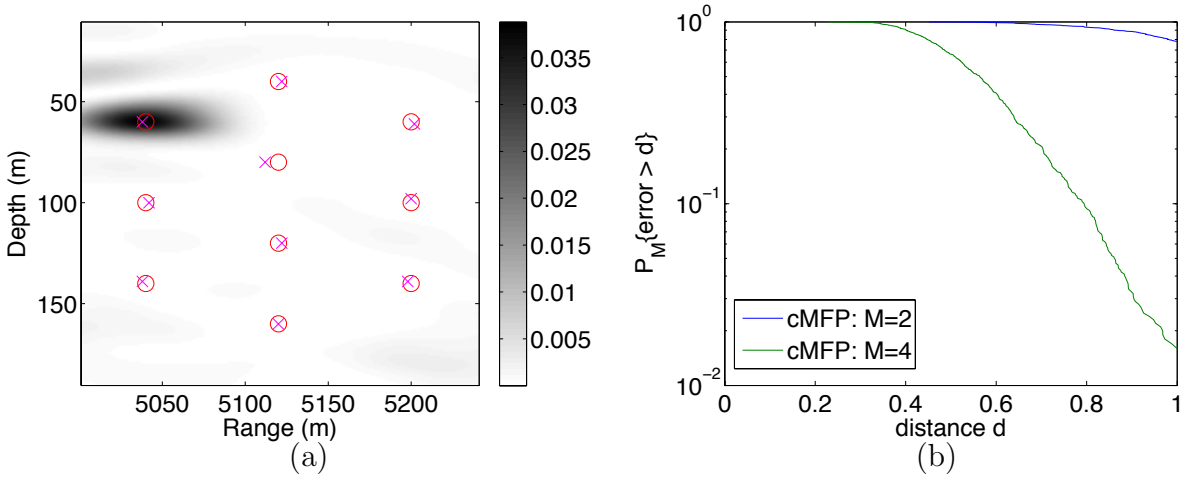


Figure 19: For a fixed pattern of 10 sources shown in (a), (b) shows the empirical tail probabilities (with respect to the randomness caused by the selection of the random matrix $\mathbf{\Phi}$ or the additive white noise η for each realization) for the localization of these sources in the broadband-coherent regime. Here, with only $M = 4$ randomized backpropagations per frequency, 10 sources may be localized to within the target ellipse with 98% probability.

case and $(5120, 20)$ for the coherent case as shown super-imposed in Fig.17). Then a weaker secondary source is placed within $\mathcal{R}$ and the empirical probability of detecting both to within their target ellipses is recorded. This weaker source has a signal to noise ratio of 40 dB in the single frequency case and 20 dB in the coherent case, and the dominant source has an SNR 20 dB higher in both cases, so that the loud source is almost always localized to within the target ellipse. These empirical probabilities were estimated by running one simulation per pixel (corresponding to a range and depth of the secondary source), and then averaging the results over a neighborhood of pixels. In both cases, a nulling rank of $n_r = 3$ was used. In both cases, the lower two thirds of the region of interest (i.e., $z > 70$) is assumed to have probability of recovery close to 1 though it was not explicitly tested for the sake of computational simplicity.

Fig. 18a shows the empirical tail probability of distance error for the single-frequency case using $M = 8$, 10, and 37 randomized backpropagations (where the latter coincides exactly with traditional MFP) using 1000 trials. In other words, this

figure shows, for a fixed d the fraction of simulations (out of 1000) that performed worse than a distance d, meaning that at least one of the sources was not localized to within a distance of d. Fig. 18b shows that same quantity as a function of M for a fixed target distance $d = 1$ (describing the so-called target ellipse) for nulling rank $n_r \in \{3, 5\}$. Note that $n_r = 3$ outperforms for low M, but $n_r = 5$ works slightly better when $M \gtrsim 8$, motivating the simple heuristic $n_r = \min(\lfloor M/2 \rfloor, 5)$ mentioned above.

When using multiple frequencies coherently, several sources can be localized. Fig. 19 illustrates the coherent joint localization of 10 sources using $M = 2$ and $M = 4$ randomized backpropagations per frequency (i.e., 40 and 80 backpropagations total), yielding computational gains of $N/M = 37/2,\ 37/4$ respectively. The locations of these 10 sources are depicted in Fig. 19a and 16d, where at each simulation, a small random rigid translation is applied to all sources to create some degree of randomness (in addition to the randomness created by the additive noise and in the selection of the $\mathbf{\Phi}$ matrix) while keeping the distance between the sources fixed.

The performance of recovery depends on the noise level as well. Fig. 20 illustrates this dependency using $M = 10$ and $M = 4$ randomized backpropagations for the single-frequency and coherent cases respectively. The broadband-coherent case generally can perform under lower SNR thanks to the extra degrees of freedom available in this case.

3.3.2 Comparison of the ROMULO algorithm to previous variations and alternatives

Fig. 21 illustrates the benefit that the robust approach described in section 3.2.2 gives over the basic point-nulling approach given in Eq. (59). With 10 sources in the hexagonal pattern used above using coherent localization with $M \in \{2, 4\}$ randomized backpropagations, ROMULO searched for each source's location either by nulling out the other source location estimates directly (point-nulling) or by nulling out all sources

within a neighborhood of these location estimates via a more general correlation matrix. It is worth noting here that although the robust approach outperforms in this case, there may be other configurations or channel models when point-nulling works better, especially at high SNR.

Fig. 22 compares performance between ROMULO and the CLEAN-based algorithm proposed by Song et al. [52]. Note that with just $M = 4$ randomized backpropagations, the proposed algorithm gives better performance than the CLEAN algorithm does for $M = 37$ randomized backpropagations (the uncompressed case) while utilizing a compressed Green's function that requires nearly 10 times fewer backpropagation runs.

Finally, Fig. 23 compares ROMULO to the global optimum for the 2-source case in the broadband-coherent regime using $M = 2$ randomized backpropagations (i.e., using an exhaustive search as in Eq. (51)). Here, ROMULO performs roughly 3.5 more poorly than the global method in terms of distance error and roughly 3.7 times worse in residual error. However, finding this global solution involves solving least squares on the data Y using all possible combinations (pairs). To expedite this search, a simple heuristic was actually used to eliminate infeasible optima by examining the corresponding pairs of values of $Y^H G(\vec{r})$, as described in more detail in Appendix A.4.

3.4 Conclusions

This chapter presented a computationally efficient scheme for multi-source localization and shows how it is conducive to a compressive approach using randomized backpropagations, even when the number of degrees of freedom (as determined by the rank of the underwater channel's correlation matrix) in the acoustic field is a small factor of the number of sources present (a factor of three to five, say). The round-robin approach of releasing source estimates back into the residual for re-estimation in particular appears to improve localization estimates substantially compared to

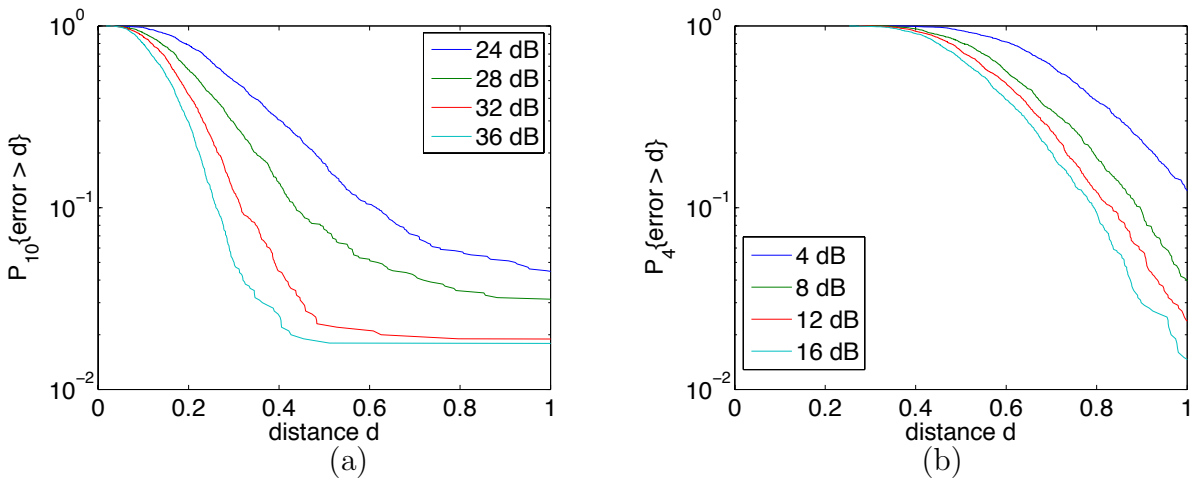


Figure 20: Tail probabilities under various noise levels in the (a) single-frequency ($M = 10$) and (b) coherent ($M = 4$) cases, where the signal to noise ratio is referenced to the weak source's amplitude and the ratio of amplitudes for the loud source to the weak source was fixed at 20 dB.

previous approaches. The simple design of the projection matrix, and in particular, the randomized approximation for larger dimensions (e.g., the broadband-coherent case) ensures computational feasibility. In practice, the compressive approach of the ROMULO algorithm provides a means to significantly reduce the dimensionality of the problem while the localization accuracy is gradually reduced when compared to results of the classical uncompressed approach

The major blind spot of ROMULO is the ability to localize sources that are close to one another, with strong correlations in the Green's function. However, this resolution problem is more or less fundamental to the multi-source localization problem, due to the ill-conditionedness of recovery. Furthermore, the compressive MFP approach inherits the known limitations of the classical MFP approach (e.g. to model mismatch of the actual environmental parameters).

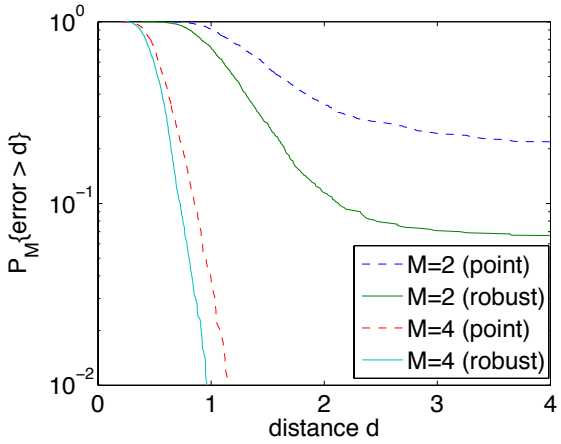


Figure 21: Localization performance (according to Eq. (55)) when the projection matrix is constructed from a correlation matrix via a neighborhood around current location estimates (solid lines) and when the correlation matrix is constructed from a neighborhood of size zero around those points, i.e., point nulling (dashed lines).

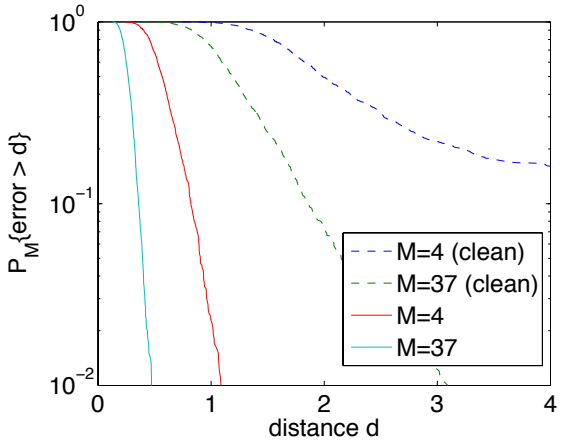


Figure 22: Comparison of the proposed ROMULO approach (solid lines) to previous implementation of the CLEAN algorithm[52] (dashed lines). Note that even in the compressed case with $M = 4$ randomized backpropagations per frequency, the ROMULO approach outperforms the CLEAN approach in the uncompressed case.

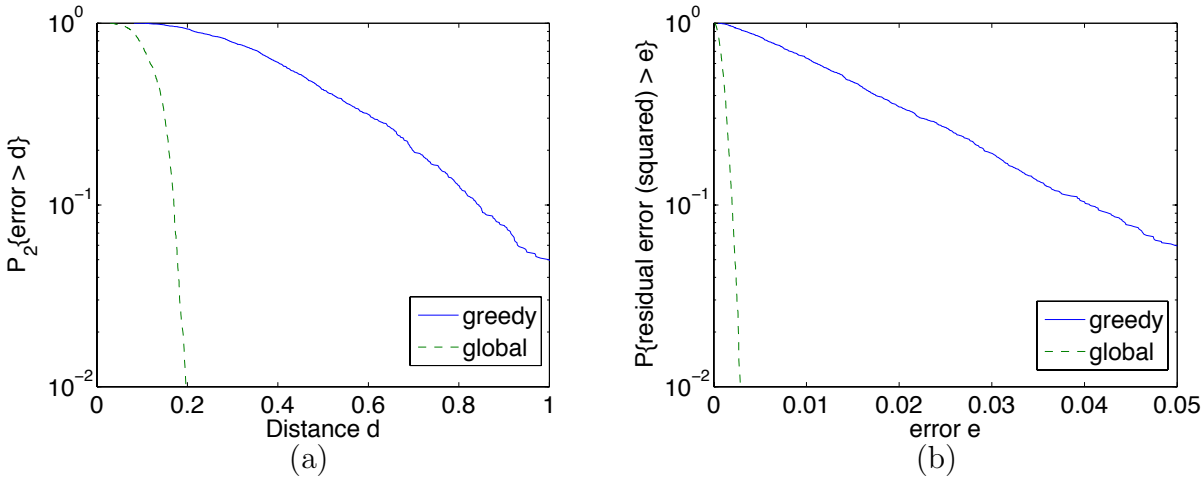


Figure 23: Comparison of the proposed greedy method (ROMULO) and the global optimum (solving Eq. (51) exhaustively) shown as (a) distance error and (b) squared-residual error $\|Y - \beta G(\vec{r})\|^2$.

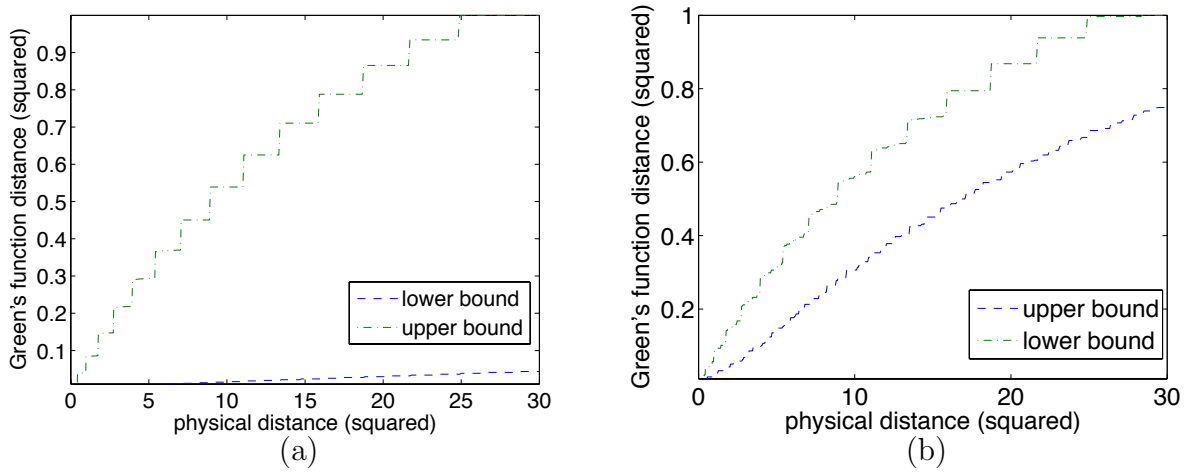


Figure 24: Relationship between physical distance and Green's function correlation. Shown here for the (a) single-frequency and (b) coherent cases are. In both panes, the upper bound $\epsilon_+(\delta)$ and the lower bound $\epsilon_-(\delta)$ on the Green's function $\|G(\vec{r}_1) - G(\vec{r}_2)\|^2$ are displayed as a function of the physical distance $\|\vec{r}_1 - \vec{r}_2\|^2$ under the elliptical distance described in Eq. (74). In particular, these functions satisfy both $\frac{1}{2}\|G(\vec{r}_1) - G(\vec{r}_2)\|^2 \leq \epsilon_+(\delta)$ for all $\|\vec{r}_1 - \vec{r}_2\|^2 \leq \delta$, and $\frac{1}{2}\|G(\vec{r}_1) - G(\vec{r}_2)\|^2 \geq \epsilon_-(\delta)$ for all $\|\vec{r}_1 - \vec{r}_2\|^2 \geq \delta$.

CHAPTER IV

COMPRESSIVE PARAMETRIC ESTIMATION

4.1 Introduction

In this chapter, we explore compressive parametric estimation more generally, where we search over a parameterized subspace for the closest match to the data:

$$\text{minimize} \quad \|h - f\|^2 \quad \text{s.t.} \quad f \in \mathcal{F}, \tag{PE}$$

where $\mathcal{F}$ is the parameterized collection of subspaces described by:

$$\mathcal{F} = \{\mathbf{V}_\theta \alpha : \theta \in \Theta, \alpha \in \mathbb{R}^K\}, \tag{76}$$

where $\mathbf{V}_\theta : \mathbb{R}^K \to L_2$ represents an orthobasis spanning parameterized subspace $\mathcal{S}_\theta$ and $\Theta \subset \mathbb{R}^D$ for compact parameter set Θ. (Whenever meaningful, we will express this functional class $\mathcal{F}$ as the union over over the set of subspaces $\{\mathcal{S}_\theta\}$, which is shorthand for $\{\mathcal{S}_\theta : \theta \in \Theta\}$.)

Its compressive counterpart simply finds the closest function with respect to some dimension-reducing linear operator $\mathbf{\Phi}$:

$$\text{minimize} \quad \|\mathbf{\Phi}(h - f)\|^2 \quad \text{s.t.} \quad f \in \mathcal{F}, \tag{CPE}$$

a constrained minimization that parallels Eq. (5). Note that estimating f is equivalent to estimating the D-dimensional parameter vector θ and the K-dimensional vector of linear coefficients α.

This formulation is a very natural approximation to the general parametric estimation problem in (CPE), given a constrained number of linear measurements $Y = \mathbf{\Phi}h \in \mathbb{R}^M$. This formulation essentially finds the closest member $f \in \mathcal{F}$ to h under some compression operator $\mathbf{\Phi}$. Provided that $\mathbf{\Phi}$ approximately preserves

distances between all close members of $\mathcal{F}$, it will be unlikely that the compressive estimate will be significantly different than its classical counterpart.

We denote the classical and compressive minimizers of (PE) and (CPE) as $\bar{\theta}$ and $\hat{\theta}$ respectively with associated errors:

$$\bar{E}^2 = \min_{\alpha} \|h - \mathbf{V}_{\bar{\theta}}\alpha\|^2 = \|h\|^2 - \|\mathbf{P}_{\bar{\theta}}h\|^2 \tag{77}$$

$$\hat{E}^2 = \min_{\alpha} \|h - \mathbf{V}_{\hat{\theta}}\alpha\|^2 = \|h\|^2 - \|\mathbf{P}_{\hat{\theta}}h\|^2, \tag{78}$$

where $\mathbf{P}_\theta = \mathbf{V}_\theta \mathbf{V}_\theta^T$. Throughout this chapter, we assume for simplicity and without loss of generality that $\|h\| = 1$. In the general case, $\bar{E}$ and $\hat{E}$ are defined as the relative errors:

$$\bar{E}^2 = \min_{\alpha} \frac{\|h - \mathbf{V}_{\bar{\theta}}\alpha\|^2}{\|h\|^2} = 1 - \frac{\|\mathbf{P}_{\bar{\theta}}h\|^2}{\|h\|^2} \tag{79}$$

$$\hat{E}^2 = \min_{\alpha} \frac{\|h - \mathbf{V}_{\hat{\theta}}\alpha\|^2}{\|h\|^2} = 1 - \frac{\|\mathbf{P}_{\hat{\theta}}h\|^2}{\|h\|^2}, \tag{80}$$

Note that the estimates of both $\bar{\theta}$ and $\hat{\theta}$ given in (PE) and (CPE) are invariant to scalar multiplication of h (or $\mathbf{\Phi}$ for that matter).

Here, rather than modeling h as a parametric function, we instead model it by a subspace with some small known of unknown parameters. This generalization affords us a wide variety of applications that are especially well suited to a small number of nonlinear parameters and a potentially large number of linear coefficients, problems involving a small number of unknown shifts or translations in particular. In Chapter 4.2.1, we discuss how this approach relates to the compressive approach to passive acoustic localization proposed in Chapter 2.

As in the case of compressed sensing of sparse signals, the performance of (CPE) is closely related to the restricted isometry property over our set $\mathcal{F}$, analogously defined as follows:

$$\delta_{\mathcal{F}} \triangleq \sup_{f \in \mathcal{F}} \left| \frac{\|\mathbf{\Phi} f\|^2}{\|f\|^2} - 1 \right| = \sup_{\theta \in \Theta} \|\mathbf{P}_\theta \mathbf{\Phi}^T \mathbf{\Phi} \mathbf{P}_\theta - \mathbf{P}_\theta\| \tag{81}$$

where the latter norm is a matrix spectral norm.

Obviously, the row-deficient $\mathbf{\Phi}$ has a null space, so $\mathcal{F}$ cannot be arbitrary if we require $\delta_{\mathcal{F}} < 1$. We need a useful measure of $\mathcal{F}$ (analogous to the "sparsity" measure of sparse sets) that will allow us to meaningfully relate the number of compressive measurements M to the accuracy of our compressive estimate.

4.1.1 Chaining Stochastic Processes

In order to build a meaningful measure of $\mathcal{F}$, it will be necessary to understand the behavior of its elements under random projection. For example, to understand the descriptor $\delta_{\mathcal{F}}$, we would like to estimate the probability that all subspaces parameterized by θ contain only functions whose norms differ from the norms of their compressive counterparts by at most some tolerance. Equivalently, we can estimate the supremum over θ of the random process $\|\mathbf{P}_\theta \mathbf{\Phi}^T \mathbf{\Phi} \mathbf{P}_\theta - \mathbf{P}_\theta\|$ as above in Eq. (81).

To bound the supremum of stochastic processes of this type, we will use a so-called "chaining" approach to analyze this continuous process [25]. The main idea of this approach is as follows. Consider a stochastic process $g(\theta)$. (For example, a stationary Gaussian process with zero mean and known autocorrelation function.) In order to bound the probability of a supremum over all $\theta \in \Theta$ of a process $g(\theta)$ from being too large, first define a sequence of finite but successively larger subsets T_j of Θ (with $j \geq 0$) with $T_0 = \{\theta_0\}$ containing only a single element (deterministically fixed but arbitrary) known as the reference point. Here, each subset T_j is a successively more "dense" approximation to Θ than its predecessor T_{j-1} (in the sense that the set of all fractions n/j with $0 \leq n < j \leq J$ become successively more dense on the interval $[0 \quad 1]$ as J increases). In this way, any element $\theta \in \Theta$ can be approximated at any "scale" j by the closest element in T_j. The operation $\pi_j(\theta)$ denotes projection of θ onto T_j – that is, the closest element in T_j to θ according to some metric that will be defined shortly below. Then, the deviation of any point from the reference point

$g(\theta) - g(\theta_0)$ can be bounded as the telescoping sum over all j of the "increments" $g(\pi_{j+1}(\theta)) - g(\pi_j(\theta))$. By defining increment "targets" δ_j, we can write:

$$\mathrm{P}\left\{\sup_{\theta\in\Theta} \; g(\theta) - g(\theta_0) \geq \delta\right\} \leq \sum_{j\geq 0} \mathrm{P}\left\{\sup_{\theta\in\Theta} \; g(\pi_{j+1}(\theta)) - g(\pi_j(\theta)) \geq \delta_j\right\} \tag{82}$$

where $\delta = \sum_{j\geq 0} \delta_j$. Now, the supremum of the continuous process can be reduced to a series of piecewise supremums over finite sets. This is because, for every fixed scale j, the corresponding increment can take only take a finite number of values, each corresponding to a pair of samples from the finite sets T_j and T_{j+1}. The maximum increment value for any given scale j can then be probabilistically bounded using a union bound over the finite set.

The goal of an effective chaining approach is then to simultaneously minimize both $\delta = \sum_j \delta_j$ and the sum of probabilities on the right hand side of Eq. (82). The remaining challenge then lies in defining the sets T_j and choosing the targets. Most commonly, these T_j are defined as ϵ-nets of the set Θ for some decreasing sequence ϵ_j. An ϵ-net T_j of Θ with radius ϵ_j is by definition a finite set of minimal cardinality satisfying $d(\theta, T_j) \leq \epsilon_j$ for all $\theta \in \Theta$ (where $d(\theta, T_j)$ denotes the smallest distance between θ and an element of T_j). This minimal cardinality is called the covering number and is defined as follows:

$$N(\Theta, \epsilon) = \min\{\mathrm{card}(T) \mid \sup_{\theta\in\Theta} d(\theta, T) \leq \epsilon, \; T \subset \Theta\}. \tag{83}$$

Evidently, the cardinality of the T_j described above is then $N(\Theta, \epsilon_j)$. The result of chaining then depends on the covering function of ϵ, $N(\Theta, \epsilon)$, for some appropriately defined metric.

4.1.2 Subspace Metric

Now that we have briefly overviewed the chaining approach, we next define an appropriate metric. There are many potential metrics on the parameter set Θ, but for

the purpose of this thesis, it will be useful and natural to view the structure of $\mathcal{F}$ according to the metric defined by the spectral norm of the difference of the projection operators:

$$d(\theta_1, \theta_2) = \|\mathbf{P}_{\theta_1} - \mathbf{P}_{\theta_2}\|. \tag{84}$$

Incidentally, this metric, sometimes called the Finsler distance, is equal to the sine of the largest principal angle between $\mathcal{S}_{\theta_1}$ and $\mathcal{S}_{\theta_2}$ [66]:

$$\|\mathbf{P}_{\theta_2} - \mathbf{P}_{\theta_1}\| = \sin(\gamma_K), \tag{85}$$

where the principal angles $\gamma_1, ..., \gamma_K$ are defined recursively via their cosines [59]:

$$\cos(\gamma_k) = \max\{\langle \mathbf{V}_{\theta_1} u_k, \mathbf{V}_{\theta_2} v_k \rangle \quad | \quad \|u_k\| = \|v_k\| = 1, \langle u_k, u_i \rangle = \langle v_k, v_i \rangle = 0 \ (\forall \ \leq i \leq k)\}. \tag{86}$$

Note that this metric is symmetric, satisfies $0 \leq d(\theta_1, \theta_2) \leq 1$, is zero if and only if $\mathcal{S}_{\theta_1} = \mathcal{S}_{\theta_2}$, and is equal to one if and only if there is a function $f_1 \in \mathcal{S}_{\theta_1}$ orthogonal to $\mathcal{S}_{\theta_2}$ (i.e., $\mathbf{P}_{\theta_2} f_1 = 0$).

In this thesis, using this metric, we explore the case when these covering numbers grow at most polynomially as ϵ decreases to zero. That is

$$N(\{\mathcal{S}_\theta\}, \epsilon) \leq N_0 \epsilon^{-d}, \tag{87}$$

for all $\epsilon \leq 1$. As it turns out, this assumption is met for many and perhaps even most cases of practical interest. In this way, $\{\mathcal{S}_\theta\}$ can be characterized by three scalars: the *base covering* N_0 (not necessarily an integer), *intrinsic dimension* d, and *subspace dimension* K. The values of N_0 and d need not be unique, and generally we will utilize any known and reasonably small pair of values that satisfy Eq. (87). A natural way to characterize the best pair of values is any feasible pair that minimizes $d + \log(N_0)$ for reasons that will become clear later. For the sake of analysis, we assume that both d and $\log(N_0)$ are at least 1.

4.1.3 Geometric Regularity

For a wide variety of applications in parametric estimation of subspaces, the signal model exhibits a great deal of structure, and we exploit one particular type of structure in this thesis. The class $\{\mathcal{S}_\theta\}_{\theta\in\Theta}$ is said to be *Lipchitz-regular* with respect to the descriptive parameterization Θ if the distances between subspaces are closely related to the Euclidean distances of the parameters that describe them. That is, for some scalar constant A,

$$\|\mathbf{P}_{\theta_2} - \mathbf{P}_{\theta_1}\| \leq A\|\theta_2 - \theta_1\|. \tag{88}$$

More generally, a set $\{\mathcal{S}_\theta\}_{\theta\in\Theta}$ is said to be *polynomial-regular* if this bound only holds up to some exponent $\alpha \leq 1$:

$$\|\mathbf{P}_{\theta_2} - \mathbf{P}_{\theta_1}\| \leq A\|\theta_2 - \theta_1\|^\alpha. \tag{89}$$

(Note that by redefining the parameter set Θ as its scaled counterpart, $A^{1/\alpha}\Theta$, then Eq. (89) is satisfied with a constant of unity.)

In either type of regularity, we can enforce the polynomial-growth bound on the covering numbers as in Eq. (87) to ensure that they do not grow too rapidly as $\epsilon \to 0$. In particular, we can link the structure of $\{\mathcal{S}_\theta\}$ to the simpler Euclidean structure of Θ to yield easy-to-compute estimates for N_0 and d as follows:

$$d = D/\alpha \tag{90}$$

$$N_0 = 3^d N(\Theta, 1, \|.\|) \tag{91}$$

by using the fact [67] that a D-dimensional unit ball B_D may be covered in Euclidean space as:

$$N(B_D, \epsilon) \leq \left(1 + \frac{2}{\epsilon}\right)^D. \tag{92}$$

In many cases, the base covering N_0 will be closer to $N(\Theta, 1, \|.\|)$. The extra 3^d term provides a loose buffer. This notion of regularity will expanded upon in the subsequent Section 4.3.

Under this characteristic structural regularity, we will show that, with high probability, the compressive estimate obtained via (CPE) is characteristically similar to the one obtained by (PE) provided that:

$$\boxed{M \gtrsim K(d + \log(KN_0))}, \tag{93}$$

and that the difference in residual errors between the two solutions decays asymptotically as the inverse-square-root of the oversampling factor. That is:

$$\hat{E}^2 - \bar{E}^2 \lesssim O\left(\left(\frac{M}{K(d + \log(KN_0))}\right)^{-1/2}\right), \tag{94}$$

When d and $\log(K)$ are less than $\log(N_0)$, as is often the case, this result essentially requires a number of measurements M on the order of magnitude of $K\log(N_0)$. This result is analogous to the classical compressive sensing result requiring a number of measurements M at least on the order of $K\log(N)$ to recover a K-sparse signal with ambient dimension N, but here the role of the ambient dimension N has been replaced by the base covering set descriptor N_0. There are $K + D$ degrees of freedom in the functional class $\mathcal{F}$, so there is little surprise that we require M at least on the order of K. The extra log factor is common to many randomized inverse problems for reasons relating to the coupon collector's problem [68].

4.1.4 Continuous Random Projection

Before stating the main result, we first will make the definition of $\mathbf{\Phi}$ concrete. In the classical case, $\mathbf{\Phi}$ is an $M \times N$ random matrix. In the continuous case, each mth element (for $1 \leq m \leq M$) is the inner product with a white noise process:

$$[\mathbf{\Phi} f]_m = \frac{1}{\sqrt{M}}\langle w_m, f\rangle, \tag{95}$$

where $w_m(t)$ is an independent sample of the white noise process $w(t)$ with zero mean and unit variance. Because we have $[\mathbf{\Phi} f]_m \sim \mathcal{N}(0, \|f\|^2/M)$, the M components of $\mathbf{\Phi} f$ are i.i.d. Gaussian with zero mean and variance $\|f\|^2/M$, so that $\mathrm{E}\left[\|\mathbf{\Phi} f\|^2\right] = \|f\|^2$, as is commonly characterized in the discrete case.

In the cases where the class of functions $\mathcal{F}$ is covered by some finite N-dimensional subspace (e.g., band-limited functions defined over some finite or periodic interval), this general continuous operation collapses in a very natural way to the standard $M \times N$ matrix-vector multiplication acting on the function's spanning coefficients. Also, by concatenating a finite dimensional orthoprojector $\boldsymbol{P}$ – e.g., the projection onto all piecewise constant functions defined on the interval $[0\ \ 1)$ – with this operator as $\boldsymbol{\Phi P} f$, we can reach broad variety of such operators. Although we focus on the Gaussian operator in this thesis, there is potential to extend to a wider variety of operators using similar techniques to those developed for traditional compressive sensing [69, 70].

4.1.5 Main Results

In this section, we state our main result, which essentially says that (PE) and (CPE) give comparable performance whenever M is sufficiently larger than K.

Theorem 2 *Let $\{\mathcal{S}_\theta\}$ be a polynomial-regular $(N_0,\ d)$ parameterization of K-dimensional subspaces. The difference in residual errors between classical and compressive parametric estimation (i.e., (77) and (78)) is probabilistically bounded above as:*

$$\mathrm{P}\left\{\hat{E}^2 - \bar{E}^2 > C\left(\sqrt{\frac{Kt}{M}} + \frac{K\log(2K)t}{M}\right)\right\} \leq KN_0 e^{d-t}. \tag{96}$$

The proof of this theorem is given in Chapter 4.4.6. Because the square root term dominates for sufficiently large M and t must be at least $d+\log(KN_0)$, this essentially says that the the difference in errors decays asymptotically with M as the inverse square root of the oversampling factor. That is,

$$\hat{E}^2 - \bar{E}^2 \leq O\left(\left(\frac{M}{K\log(KN_0)}\right)^{-1/2}\right) \tag{97}$$

For a comparable probabilistic bound, we also have a restricted isometry property over $\mathcal{F} = \bigcup_{\theta\in\Theta}\{\mathcal{S}_\theta\}$.

Theorem 3 *For some δ_G satisfying:*

$$\mathrm{P}\left\{\delta_G > C\left(\sqrt{\frac{Kt}{M}} + \frac{K\log(2K)t}{M}\right)\right\} \leq KN_0e^{d-t}. \tag{98}$$

and for all $f \in \mathcal{F}$, we have the restricted isometry property:

$$(1-\delta_G)\|f\|^2 \leq \|\mathbf{\Phi}f\|^2 \leq (1+\delta_G)\|f\|^2 \tag{99}$$

The proof of this theorem is also given in Section 4.4.6 and essentially depends on Lemma 15.

4.1.6 Related work

A closely related research area is the study of manifold embeddings, which may be parametric, or more generally, non-parametric. Baraniuk and Wakin made seminal progress in this area in a result that mirrored the isometric properties in a natural way [51]. Here, they relate the volume of the manifold, its ambient dimension and intrinsic dimension, its condition number, and its geodesic covering regularity to the number of compressive measurements needed to effectively preserve the pairwise distances of all members of the manifold within a prescribed distance. Clarkson later refined this result, removing the dependence on the ambient dimension, and robustly substituting an average curvature when the worst case curvature was previously used implicitly [71]. Yap et al. later published a variation on these earlier results for a variety of projection operators $\mathbf{\Phi}$ [72]. This thesis generalizes this prior work by guaranteeing performance of the compressive estimation of parameterized subspaces, and also utilizes set descriptors N_0, d, and K that can be readily and intuitively estimated for a wide variety of practical cases, providing relative ease of use.

In the field of empirical processes, Talagrand has recently stated necessary and sufficient conditions for the boundedness of arbitrary Gaussian processes via generic chaining arguments, improving upon earlier theorems involving Kolmogorov's metric entropy, and has established related theorems that are utilized in this thesis [25].

We develop a similar chaining framework to analyze the performance of CPE while focusing on probability bounds instead of expectations, but tailor it specifically for the relatively intuitive parameters N_0 and d that describe polynomial-regular parameter sets, lessening the need of the principled but somewhat arcane set descriptor γ_2 developed in the generic chaining. Mendelson et al. also prove a restricted isometry property for sparse vectors, but do so in a way that allows for generalizations to other sets via a set descriptor that can be analyzed via these generic chaining techniques [73]. However, it is not clear how this work can be applied to parametric estimation.

Researchers have already utilized this type of parametric estimation in specific applications, including work in compressive matched filtering, radar pulse signal acquisition, and compressive matched-field processing [74, 75, 23]. Mishali et al. have successfully designed hardware for analog to digital conversion at sub-Nyquist rates using a scheme they call Xampling [76]. This hardware appears to be conducive to the compressive estimation of subspaces from receiver data using parameters for each carrier frequency, although this has not been tested yet to the author's knowledge. In the field of radio estimation, Yoo et al. estimate the frequency, starting time and ending time of a sinusoidal pulse that has been sampled via a random block-diagonal sensing matrix [75]. Here, they utilize Discrete Prolate Spheroidal Sequences (DPSS) functions [63], and particularly the projection operator onto this basis in order to remove the contribution of each candidate sinusoid, as the procedure searches for multiple carrier frequencies.

Many researchers have performed multi-target tracking under the "target-sparse" assumption. That is, the methods propose to simultaneously localize several targets that lie on some grid (or generally some set of points) by solving an ℓ_1 minimization program. The recovered support resulting from this optimization corresponds to the grid points that the various targets are estimated to occupy. This leverages the main results of CS to prove that the targets may be perfectly localized with a

high probability. Often, the painstaking effort in these papers involves showing that the RIP holds for the observation matrix. For example, Fannjiang et al. show the conditions that guarantee a sufficiently small coherence, which in turn guarantees exact localization [43]. Gurbuz et al. show similar results for a compressive beamformer, requiring a number of measurements on the order of the number of sources [44], but the application there is different in that they utilize a signal common to all sensors with an unknown time shift to localize their target in angle (assuming free space propagation), and apply the compression operator in time per-sensor instead of applying the operator across the range of sensors as we do. In seminal work under a framework that predates compressed sensing somewhat, Fuchs proposed an algorithm for detecting multiple acoustic sources' direction of arrival [77]. Essentially they solve a sparse inverse problem of an underdetermined system. They consider a greedy algorithm such as matching pursuit [60], but instead opt to solve a regularized least squares problem that is equivalent via duality to basis pursuit denoising (BPDN) with an additional positivity constraint on the sparse vector, which can in turn be solved with a known polynomial time linear program solver. Also, from a communications perspective, Cevher et al. demonstrate the relatively low amount of information to be transmitted for purposes of localization when using a CS framework [45].

These "target-sparse" approaches depend on targets lying exactly on the grid points. Also, by necessity these grid points must be spaced sufficiently far away from one another to avoid coherence-inducing correlations in the observation matrix. This creates a restrictive model of limited applicability. When a target is somewhere in between a set of grid points, the necessary conditions for recovery may not even approximately hold, similarly to how a discrete sinusoid corresponding to an off-grid point in the DFT will not be sparse in the frequency domain (or any other basis for any standard transform for that matter) because of DFT leakage. In contrast to this approach, we do not require the target to lie on a grid point. However, instead of

promising perfect recovery, we instead make the softer claim that the target may be localized to within a small neighborhood of the actual source location.

Ekanadham et al. recognized this limitation of these target-sparse approaches and devised an innovative solution, called continuous basis pursuit (CBP) [78]. Here, the grid points are spaced far enough to avoid coherence issues, so to account for candidates lying between the grid points, they utilize a subspace of small dimension that accounts for local shifts of the base function. This work dovetails very nicely with the framework proposed in this thesis. For example, the Hermite functions discussed in Chapter 4.3.1 approximate the derivatives of a Gaussian function, and could be used to search for a fine-grained match with a Gaussian pulse, even using a coarse grid to do so.

Also, in 2011, Eftekhari et al. proposed a method for compressive matched filtering and provided probabilistic results to its performance. Here, they maximize the correlation between the compressed data vector and the compressed model over a bounded set of possible shift parameters, where the model signal being matched is band-limited and the compression operator measures frequencies uniformly at random inside this band. Lastly, Candès and Granda propose a novel analytical framework for super-resoution, defined as the approximate recovery of the sum of a small number of dirac point sources from a low-frequency (and consequently low-dimensional) observation [79]. This work differs from our own in that it measures the time series directly and deterministically instead of compressively and also does not explicitly estimate the parameteric shift of the point sources, but rather does so implicitly by recovering the function itself via a simple convex program that jointly minimizes the total variation of that function and a least-squares constraint.

The rest of this chapter is organized as follows. Section 4.3 illustrates practical methods for showing Eq. (87), including a few specific examples. Section 4.2 discusses

some potential applications that could benefit from a compressive approach, including compressive matched field processing (cMFP). Finally, Section 4.4 develops the analytical framework necessary to establish the main results.

4.2 Applications

Although this compressive approach in Eq. (CPE) tends to suffer a loss of accuracy relative to its classical counterpart in Eq. (PE), this loss may be small enough to be outweighed by other tangible advantages.

The flagship application presented in this thesis is compressive matched field processing. Here, the acquisition hardware is relatively simple, needing only sampling rates in the kilohertz, and the advantage of a compressive approach is an improvement on software, not hardware (see Section 2.1). However, the most commonly claimed advantage in compressive sensing involves a simplified data acquisition architecture. This can prove advantageous (for example) for low-power devices that must run on battery or solar power that must do very little data processing and communication, but are connected to machines with enormous computational power and storage capacity. ECG and transit detection. We now show how these applications may be analyzed within the proposed framework.

4.2.1 Compressive Matched-Field Processing

Compressive Matched-Field Processing, discussed in detail in Chapter 2, may be examined within the framework of the thesis. The complex-valued model for the Green's function for a given frequency:

$$\mathcal{F} = \{G(\vec{r})\alpha \in \mathbb{C}^N : \vec{r} \in \mathcal{R}, \alpha \in \mathbb{C}\} \tag{100}$$

can be readily adapted to a real-valued subspace-matching problem:

$$\{\mathcal{S}_{\vec{r}}\} = \{\mathbf{V}(\vec{r})\alpha : \vec{r} \in \mathcal{R}, \alpha \in \mathbb{R}^2\} \tag{101}$$

where the search parameters $\vec{r}$ constitute range and depth over some bounded search region of interest $\Theta = \mathcal{R}$ and the subspace model $V(\vec{r}) \in \mathbb{R}^{2N\times 2}$ is given as:

$$\mathbf{V}(\vec{r}) = \begin{bmatrix} \mathrm{Re}(G(\vec{r})) & -\mathrm{Im}(G(\vec{r})) \\ \mathrm{Im}(G(\vec{r})) & \mathrm{Re}(G(\vec{r})) \end{bmatrix}, \tag{102}$$

where it is assumed that the Green's function is normalized so that $\|G(\vec{r})\| = 1$. The natural extension of the projection metric under this model is:

$$d(\vec{r}_1, \vec{r}_2)^2 = 1 - |\langle G(\vec{r}_1), G(\vec{r}_2)\rangle|^2. \tag{103}$$

It only remains to show that this metric is Lipchitz-continuous. This condition holds at least empirically for the Pekeris waveguide employed in this thesis, as demonstrated on figure. 24. For the elliptical distance metric defined in Eq. (75), we have the following Lipchitz bounds for the single-frequency and coherent cases, respectively:

$$d_s(\vec{r}_1, \vec{r}_2) \leq \frac{1}{2}\|\vec{r}_1 - \vec{r}_2\|_s \tag{104}$$

$$d_c(\vec{r}_1, \vec{r}_2) \leq \frac{1}{3}\|\vec{r}_1 - \vec{r}_2\|_c \tag{105}$$

where $\|.\|_s$ and $\|.\|_c$ denote the elliptical norms defined for the single-frequency and coherent. For this reason, this parameterized set is polynomial-regular with $d = 2$ and $N_0 \simeq N(\mathcal{R}, 2, \|.\|_s)$ for the single-frequency case and $N_0 \simeq N(\mathcal{R}, 3, \|.\|_c)$ for the coherent case. In particular, the regions of interest $\mathcal{R}$ described in Chapter 2.4 correspond to $N_0 = 116$ in the single-frequency case where $\mathcal{R} = [10\text{m}\ \ 190\text{m}]\times[5000\text{m}\ \ 5720\text{m}]$ and $N_0 = 52$ in the broadband-coherent case where $\mathcal{R} = [10\text{m}\ \ 190\text{m}] \times [5000\text{m}\ \ 5240\text{m}]$ via a hexagonal covering argument discussed in more detail in Chapter 4.3.

4.2.2 Transit Detection

In the field of astronomy, the discovery of new planets in distant solar systems is indirectly possible by observing a drop in light intensity of a star caused by a temporarily occluding planet, a method known as *transit detection* [80]. The Kepler spacecraft,

launched in 2009, is focused on a star cluster NGC 6791 in an attempt to discover new planets using this method while mitigating atmospheric effects that encumber terrestrial observatories.

To date, 105 planets have been discovered using the Kepler spacecraft, but there are hardware constraints that limit the performance of this system. A 95-megapixel CCD array is sampled once every 6 seconds, and then further downsampled on-board by averaging over a thirty minute interval. Unfortunately, this results in the loss of high-frequency information so that a given transit event may only have dozens of samples describing it. But even at this low rate, there is too much information to send back to Earth, so only a small-number of pre-selected pixels are sent back to Earth, amounting to roughly 5% of the total pixel array.

Alternatively, to the extent that such transit patterns are well-described by a parametric model, it may be possible to improve performance for a comparable information budget using a compressive approach where this time-series information is randomly compressed, possibly by a common operation for all pixels over a given time interval. Rather than implicitly throwing away all high-frequency information via averaging, compressive measurements could be made across several frequency bands: some used for detection, and others used for more fine-grained timing estimation than would be afforded at the half-hour interval. A simple parametric model containing time-shift and dilation could suffice while the extra $K - 1$ degrees of freedom in the modeled subspace V_θ could account for variations in the nominal model.

4.2.3 ECG Monitoring

Similar compressive techniques have already been applied in remote monitoring of Electrocardiogram (ECG) data [81]. Here, Garudadri et al. present an efficient hardware design for random projections that has a side benefit that it is inherently resilient

against packet losses of the transmitted compressed data, since each compressed measurement is essentially as valuable as any other one and contains redundant information. Here, power efficiency enables a wearable form factor and a long battery life. Although they did not explore compressive parametric estimation, their hardware is conducive towards such a system.

The ultimate goal is not reconstruction per se, but rather the robust detection of irregular heartbeats. One approach, given a set of randomly compressed measurements, would be to model both healthy heartbeats and irregular heartbeats via some known basis (e.g., from PCA) with an unknown shift, and then classify each heartbeat according to false alarm criteria. The parameterization may include only a simple time shift, or could also include a parameterized deformation between a healthy heartbeat and an unhealthy one.

4.3 Parametric Regularity

In this section, we will discuss some parameterized subspaces $\mathcal{S}_\theta$, and show the conditions under which these classes are polynomial-regular, satisfying:

$$N(\{\mathcal{S}_\theta\}, \epsilon) \leq N_0 \epsilon^{-d}, \tag{106}$$

where $\{\mathcal{S}_\theta\}$ is shorthand for the set of parameterized subspaces $\{\mathcal{S}_\theta : \theta \in \Theta\}$.

We will do this primarily by showing that the projection matrices are Hölder-continuous with respect to parameterized transformation. That is:

$$\|\mathbf{P}_{\theta_1} - \mathbf{P}_{\theta_2}\| \leq \|A(\theta_1 - \theta_2)\|^\alpha, \tag{107}$$

for some constants α and A. When possible, we will redefine the set Θ to the "normalized" parametric set $A\Theta$ so that this scalar term A will be unnecessary.

This property in Eq. (107) will allow us to directly tie the covering numbers of the set class $\mathcal{F}$ to the covering numbers of the parameter class Θ:

$$\|\mathbf{P}_{\theta_1} - \mathbf{P}_{\theta_2}\| \leq \|\theta_1 - \theta_2\|^\alpha \;\longrightarrow\; N(\{\mathcal{S}_\theta\}, \epsilon) \;\leq\; N(\Theta, \epsilon^{1/\alpha}) \leq N_0 \epsilon^{-D/\alpha} \tag{108}$$

for some N_0. For example, the following lemma shows that $N_0 \leq 3^D N(\Theta, 1)$.

Lemma 1 *[67, Lemma 3.18] Let B_D denote the unit ball in $\mathbb{R}^D$. For any $0 < \epsilon < 1$, we have the following upper bound on the covering number of ball of radius ϵ:*

$$N(B_D, \epsilon) \leq \left(1 + \frac{2}{\epsilon}\right)^D \leq \left(\frac{3}{\epsilon}\right)^D \tag{109}$$

Alternatively, suppose that $\Theta \subset \theta_0 + RB_D$, where B_D is the unit ball in $\mathbb{R}^D$. (The boundedness of Θ follows from its compactness). Then Lemma 1 gives us, for any $\epsilon \leq R$:

$$N(\Theta, \epsilon) \leq \left(\frac{3R}{\epsilon^{1/\alpha}}\right)^D, \tag{110}$$

so that Eq. (106) is satisfied with base covering $N_0 = (3R)^D$ and effective dimension $d = D/\alpha$.

A simple volumetric argument shows that the covering numbers must be at least the ratio of the volume $|\Theta|$ to the volume of the covering ball. For example, for $\Theta \subset \mathbb{R}^2$, we have $N(\Theta, \epsilon) \geq \frac{|\Theta|}{\pi\epsilon^2}$, yielding $N_0 \geq \pi^{-1}|\Theta|$. On the other hand, the base covering term N_0 is often not much bigger than this minimum number. In the two-dimensional case, there exists a hexagonal lattice sampling which is only sub-optimal by a factor of $\vartheta_2 = \frac{2\pi}{3\sqrt{3}}$ (the ratio of the area of a circle to the area of a the hexagon with the same radius), yielding an asymptotic covering number of $N(\Theta, \epsilon) \simeq |\Theta|\frac{2}{3\sqrt{3}}\epsilon^{-2}$ as ϵ becomes smaller. This so-called *covering density* $\vartheta_D \geq 1$ is the asymptotic ratio of the covering area to the set area, and is known to be at most $CD\log^3(D)$ in the D dimensional case for some constant C [82, pg. 19], so that $N(\Theta, \epsilon) \gtrsim \frac{\vartheta_D|\Theta|}{|B_D|\epsilon^D}$ for small ϵ (where $|B_D|$ is the volume of the D-dimensional unit ball), yielding a lower bound of and often a decent approximation to N_0. The non-asymptotic case generally requires only a mild buffer for edge effects. For example, when $\Theta = [0\ a] \times [0\ b]$ for some dimensions $a,\ b$ at least 3, then $N(\Theta, \epsilon) \leq ab\epsilon^{-2}$ so that $N_0 = ab$ and $d = 2$, a factor of roughly 3 above the lower bound of N_0.

There are other ways that the polynomial condition may be applied to obtain good bounds on the base covering and effective dimension. In particular, if the set Θ has a great deal of variation in some of its dimensions, but relatively small variation in others, the following application of composite parametric operations becomes useful.

Lemma 2 *Let $\Theta_1, ..., \Theta_J$ be parameter sets such that the subspaces parameterized by them are polynomial-regular with base coverings $N_0^{(j)}$ and effective dimensions $d^{(j)}$. Then the functional class of their product: $\{\mathcal{S}_\theta\} = \{f^{(1)} \cdots f^{(J)} : f^{(j)} \in \mathcal{S}^{(j)}_{\theta^{(j)}},\ \theta^{(j)} \in \Theta^{(j)}\}$ is also polynomial-regular with base covering $N_0 = J^d \prod_j N_0^{(j)}$ with effective dimension $d = \sum_j d^{(j)}$.*

Proof Because:

$$\left\| \left(f^{(1)} \cdots f^{(J)}\right) - \left(f^{(1)} \cdots f^{(j)'} \cdots f^{(J)}\right) \right\| \leq \left\| f^{(j)} - f^{(j)'} \right\|, \tag{111}$$

we have:

$$N(\{\mathcal{S}_\theta\}, \epsilon) \leq N\left(\prod_j \{\mathcal{S}^{(j)}_{\theta^{(j)}}\}, \epsilon\right) \leq \prod_j N\left(\{\mathcal{S}^{(j)}_{\theta^{(j)}}\}, \epsilon/J\right) \leq \prod_j J^{d^{(j)}} N_0^{(j)} \epsilon^{-d^{(j)}}. \tag{112}$$

■

Likewise, it may be easier to consider a partition of the parameterized set, as follows.

Lemma 3 *Let $\mathcal{F}_1, ..., \mathcal{F}_J$ be parameterized functional classes that are polynomial-regular with base coverings $N_0^{(j)}$ and effective dimensions $d^{(j)}$. Then the functional class of their union: $\mathcal{F} = \bigcup_j \mathcal{F}_j$ has base covering $N_0 \leq \sum_j N_0^{(j)}$ with effective dimension $d \leq \max\{d^{(j)}\}$.*

Proof Simply adding the covering numbers gives:

$$N(\{\mathcal{S}_\theta\}, \epsilon) \leq \sum_j N(\{\mathcal{S}_\theta\}_j, \epsilon) \leq \sum_j N_0^{(j)} \epsilon^{-d^{(j)}} \leq N_0 \epsilon^{-d}. \tag{113}$$

■

4.3.1 Orthobasis Analysis

Now that we've established some basic properties of covering numbers and the parameters describing polynomial-regular sets, we proceed by showing Eq (107) for specific cases of interest. This bound is implied from a similar bound on the basis matrix $\mathbf{V}$ instead of the projection matrix $\mathbf{P}$ for the following reason.

Lemma 4 *Let $\mathbf{V}_1$ and $\mathbf{V}_2$ be K-dimensional orthogonal bases with corresponding rank-K projection matrices $\mathbf{P}_1 = \mathbf{V}_1\mathbf{V}_1^T$ and $\mathbf{P}_2 = \mathbf{V}_2\mathbf{V}_2^T$ onto corresponding subspaces $\mathcal{S}_1$ and $\mathcal{S}_2$. Then we have:*

$$\|\mathbf{P}_1 - \mathbf{P}_2\| \leq 2\|\mathbf{V}_1 - \mathbf{V}_2\| \leq 2\|\mathbf{V}_1 - \mathbf{V}_2\|_F, \tag{114}$$

where $\|.\|_F$ denotes the Frobenius norm [59]. Furthermore, there exist orthobases $\mathbf{V}_a$ and $\mathbf{V}_b$ for these subspaces (i.e., such that $\mathbf{P}_1 = \mathbf{V}_a\mathbf{V}_a^T$, $\mathbf{P}_2 = \mathbf{V}_b\mathbf{V}_b^T$) satisfying:

$$\|\mathbf{V}_a - \mathbf{V}_b\| \leq \|\mathbf{P}_1 - \mathbf{P}_2\|. \tag{115}$$

Proof For the first claim, we write:

$$\begin{aligned}
\|\mathbf{P}_1 - \mathbf{P}_2\| &= \frac{1}{2}\|(\mathbf{V}_1 - \mathbf{V}_2)(\mathbf{V}_1 + \mathbf{V}_2)^T + (\mathbf{V}_1 + \mathbf{V}_2)(\mathbf{V}_1 - \mathbf{V}_2)^T\| && (116)\\
&\leq 2\|\mathbf{V}_1 - \mathbf{V}_2\| && (117)\\
&\leq 2\|\mathbf{V}_1 - \mathbf{V}_2\|_F. && (118)
\end{aligned}$$

For the second claim, consider an arbitrary orthobasis $\tilde{\mathbf{V}}_1$ for $\mathcal{S}_1$ and $\tilde{\mathbf{V}}_2$ for $\mathcal{S}_2$ with singular value decomposition $\tilde{\mathbf{V}}_1^T\tilde{\mathbf{V}}_2 = \mathbf{U}\mathbf{\Sigma}\mathbf{Y}$. Defining $\sigma = \mathbf{\Sigma}_{K,K}$, $\mathbf{V}_a = \tilde{\mathbf{V}}_1\mathbf{U}$, and $\mathbf{V}_b = \tilde{\mathbf{V}}_2\mathbf{Y}$ gives $\tilde{\mathbf{V}}_1^T\tilde{\mathbf{V}}_2 = \mathbf{\Sigma}$, so that:

$$\begin{aligned}
\|\mathbf{V}_a - \mathbf{V}_b\|^2 &= \max_{\|x\|=1} \|\mathbf{V}_a x\|^2 + \|\mathbf{V}_b x\|^2 - 2\langle \mathbf{V}_a x,\ \mathbf{V}_b x\rangle && (119)\\
&= 1 - \sigma && (120)\\
&\leq 1 - \sigma^2 && (121)\\
&\leq \|(\mathbf{P}_1 - \mathbf{P}_2)\mathbf{V}_{aK}\|^2 && (122)\\
&\leq \|\mathbf{P}_1 - \mathbf{P}_2\|^2. && (123)
\end{aligned}$$

where $\mathbf{V}_{aK}$ is the Kth column of $\mathbf{V}_a$. ■

In light of this, it only remains to be shown that the individual unit-norm functions defining the orthobasis satisfy polynomial-regularity – i.e., that $\|\psi_1 - \psi_2\| \leq \|\theta_1 - \theta_2\|^\alpha$ – and then use the Frobenius norm as in Lemma 4 to establish bounds on N_0 and d.

Now, we consider the polynomial-regularity of shifts of orthobases that are approximately compactly supported. Note that a signal cannot have limited support in both time and frequency according to the Weyl-Heisenberg principle, but there are classes of signals that are more concentrated than others, and here we will consider three such orthobases (Hermite, LOT, and DPSS) that could be used to account for a signal that is well-localized in both time and frequency, but where the temporal support (and perhaps frequency support) of the signal is unknown. In particular, we show show how their bounded total variation ensures that shifts of these bases will be polynomial-regular.

The examples here are meant to be illustrative more than precise, and in particular we note that low order polynomials of d and K in N_0 do not substantially affect the $\log(KN_0)$ term in Theorem (2), which effectively already contains additive terms in d and $\log(K)$. Also, in all cases the base covering N_0 must be at least 1, so in all cases when writing $N_0 = (\cdot)$ we implicitly mean $N_0 = \max\{1, (\cdot)\}$. We begin by relating the polynomial-regularity of time shifts of an orthobasis to the bounded total variation of its individual orthobasis functions.

Lemma 5 *Specifically, suppose we have a base function whose total variation is bounded as:*

$$\int |\psi'(t)| dt \leq L. \tag{124}$$

Then we have:

$$\|\psi_{\theta_1} - \psi_{\theta_2}\| \leq L \|\theta_1 - \theta_2\|^{1/2}. \tag{125}$$

Proof The L_2 norm may be bounded as:

$$\begin{aligned}\int_{-\infty}^{\infty}(\psi(t)-\psi(t-\theta_s))^2dt &= \int_0^{\theta_s}\sum_{k=-\infty}^{\infty}(\psi(t+k\theta_s)-\psi(t+(k-1)\theta_s))^2 &(126)\\ &\leq \int_0^{\theta_s}\left(\sum_{k=-\infty}^{\infty}|\psi(t+k\theta_s)-\psi(t+(k-1)\theta_s)|\right)^2 &(127)\\ &\leq L^2\theta_s. &(128)\end{aligned}$$

■

Note that in all cases, the norm of the difference of functions only depends on the relative shift $\theta_s = \theta_2 - \theta_1$:

$$\|\psi(t-\theta_2)-\psi(t-\theta_1)\| = \|\psi(t-\theta_s)-\psi(t)\|, \tag{129}$$

and also that there is a sort of invariance of polynomial-regularity to dilation of the form:

$$\sqrt{\nu}\psi(\nu t) \tag{130}$$

provided that the set Θ is dilated to $\nu\Theta$. That is,

$$\|\psi(t-\theta_2)-\psi(t-\theta_1)\| = \|\sqrt{\nu}\psi(\nu t-\nu\theta_2)-\sqrt{\nu}\psi(\nu t-\nu\theta_1)\|, \tag{131}$$

and naturally the orthogonality of $\mathbf{V}$ is preserved under this dilation as well. For this reason, the properties stated for the canonical orthobases below generalize naturally to other scales of these orthobases.

Hermite Polynomials

Now we consider the properties of the specific orthobases, beginning with the Hermite polynomials [83]. The Hermite polynomials are defined as:

$$\begin{aligned}H_0(t) &= 1 &(132)\\ H_1(t) &= t &(133)\\ &... &(134)\\ H_{k+1}(t) &= tH_k(t)-kH_{k-1}(t), &(135)\end{aligned}$$

or alternatively are given explicitly as:

$$H_K(t) = K! \sum_{k=0}^{\lfloor K/2 \rfloor} \frac{(-1)^k}{k!(K-2k)!}(2t)^{K-2k}, \tag{136}$$

and are orthogonal under the inner product:

$$\langle H_k, H_{k'} \rangle = \int H_k(t) H_{k'}(t) w(t) dt = I(k = k') k! \sqrt{2\pi}, \tag{137}$$

where the window function:

$$w(t) = e^{-t^2/2}. \tag{138}$$

A natural choice of orthobasis in standard Euclidean space is then

$$\psi_k(t) = \sqrt{\frac{w(t)}{k!\sqrt{2\pi}}} H_k(t), \tag{139}$$

with the corresponding parameterized K-dimensional subspace $\mathcal{S}_\theta = \{\sum_{k=0}^{K-1} \alpha_k \psi_k(t-\theta) : \alpha \in \mathbb{R}^K\}$. Using the facts that $|\psi_k(t)| \leq 1$ and that ψ_k has exactly $k+1$ local extrema, we have $\int |\psi_k'(t)| dt \leq 2(k+2)$, so applying Lemmas 4 and 5, we have:

$$\|\mathbf{P}_1 - \mathbf{P}_2\|^2 \leq 64K^3 \|\theta_1 - \theta_2\|, \tag{140}$$

and

$$N(\{\mathcal{S}_\theta\}, \epsilon) \leq N(64K^3[t_a \;\; t_b], \epsilon^2) \leq 64K^3 |t_b - t_a| \epsilon^{-2}, \tag{141}$$

so that $\{\mathcal{S}_\theta : \theta \in \Theta = [a \; b]\}$ is polynomial-regular with $d = 2$ and $N_0 = 64K^3|t_b - t_a|$.

Lapped Orthogonal Transform

Next, we discuss the lapped orthogonal transform (LOT) [84]. These are defined via a window function $g(t)$, defined as:

$$g(t) = \begin{cases} 0: & t < -\eta \\ \beta(\frac{t}{\eta}): & -\eta \leq t < \eta \\ 1: & \eta \leq t < 1-\eta \\ \beta(\frac{1-t}{\eta}): & 1-\eta \leq t < 1+\eta \\ 0: & 1+\eta \leq t \end{cases}$$

for some $\eta \leq 1/2$ and base function $\beta(t)$ satisfying: $\beta(t)^2 + \beta(-t)^2 = 1$ on the domain $t \in [-1 \;\; 1]$, the most common example being:

$$\beta(t) = \sin\left(\frac{\pi}{4}(1+t)\right). \tag{142}$$

The orthobasis functions ψ_k for the lapped orthogonal transform are then defined via this window $g(t)$ as:

$$\psi_k(t) = g(t)\sqrt{2}\cos\left(\pi\left(k+\frac{1}{2}\right)t\right). \tag{143}$$

Similarly to the case of the Hermite orthobasis, we note that $|\psi_k(t)| \leq \sqrt{2}$ and also that $\psi_k(t) = 0$ everywhere except $t \in (-\eta \;\; 1+\eta)$, a domain containing at most $2k+1$ extrema, so that $\int |\psi_k'(t)|dt \leq 4\sqrt{2}(k+1)$. Using Lemmas 4 and 5 as before, we have:

$$\|\mathbf{P}_1 - \mathbf{P}_2\|^2 \leq 128K^3\|\theta_1 - \theta_2\|, \tag{144}$$

and

$$N(\{\mathcal{S}_\theta\}, \epsilon) \leq N(128K^3[t_a \;\; t_b], \epsilon^2) \leq 64K^3|t_b - t_a|\epsilon^{-2}, \tag{145}$$

so that $\{\mathcal{S}_\theta : \theta \in \Theta = [a \; b]\}$ is polynomial-regular with $d = 2$ and $N_0 = 128K^3|t_b - t_a|$.

Prolate Spheroidal Functions

Prolate spheroidal wave functions (PSWFs) (also called Slepian functions) have recently proven their utility at bridging the gap between theoretical compressed sensing and practical applications such as signal reconstruction and channel identification [85, 63]. The associated K-dimensional subspace $\mathcal{S}$ essentially contains the band-limited functions that are the most approximately time-limited. That is, this subspace $\mathcal{S}$ satisfies:

$$\min_{f\in\mathcal{S}} \frac{\|\mathbf{P}_\Omega \mathbf{P}_T f\|^2}{\|f\|^2} \geq \min_{f\in\mathcal{S}'} \frac{\|\mathbf{P}_\Omega \mathbf{P}_T f\|^2}{\|f\|^2}. \tag{146}$$

for all other K-dimensional subspaces $\mathcal{S}'$, where $\mathbf{P}_T$ denotes a projection onto the time interval $[-\frac{T}{2} \;\; \frac{T}{2}]$ via multiplication by a rectangular window, and $\mathbf{P}_\Omega$ denotes a projection onto the frequency interval $[-\frac{\Omega}{2} \;\; \frac{\Omega}{2}]$ via an ideal low-pass filter. The basis

functions $\psi_k(t)$ $(1 \leq k \leq K)$ spanning this space $\mathcal{S}$ are simply the eigenfunctions of the symmetrized operator $\mathbf{P}_T\mathbf{P}_\Omega\mathbf{P}_T$ with corresponding eigenvalues λ_k so that $\min_{f\in\mathcal{S}} \frac{\|\mathbf{P}_\Omega\mathbf{P}_T f\|^2}{\|f\|^2} = \lambda_K$. Roughly speaking, this successive time-limiting bandwidth-limiting operation has a rank of approximately $T\Omega$, so that $\lambda_k \simeq 1$ for $k \lesssim T\Omega$ and $\lambda_k \simeq 0$ for $k \gtrsim T\Omega$, so taking $K \lesssim T\Omega$ is a natural choice. In any rate, we will assume that K is chosen so that $\lambda_K \geq 1/2$.

Unlike the previous two examples, here we are able to show polynomial-regularity directly without the use of Lemma 5. Because the orthobasis functions $\psi_k(t)$ are band-limited, we have for any $\theta_s = \theta_1 - \theta_2$:

$$\begin{aligned} \|\psi_{k;\theta_1} - \psi_{k;\theta_2}\|^2 &= \int_{-\Omega/2}^{\Omega/2} \hat{\psi}_k(\omega)^2 |e^{j\omega\theta_s/2} - e^{-j\omega\theta_s/2}|^2 d\omega && (147) \\ &\leq 4\sin(\Omega\theta_s/4)^2 && (148) \\ &\leq (\Omega\theta_s/2)^2, && (149) \end{aligned}$$

where $\hat{\psi}_k$ is the Fourier transform of ψ_k. Therefore, the set of parameterized subspaces $\{\mathcal{S}_\theta\} = \{\{\sum_k \alpha_k \psi_k(t-\theta) : \alpha \in \mathbb{R}^K\} : \theta \in [t_a \;\; t_b]\}$ is polynomial-regular with base covering $N_0 = |t_b - t_a|\Omega$ and effective dimension $d = 1$. Note that because of the shift invariance property of function norms as in Eq. (129), we have the same regularity for any projection $\mathbf{P}_T$ onto an interval of length T, not only the canonical one. Indeed, these shifted subspaces are equivalent to those obtained via the operator $\mathbf{P}_{T;\theta}$ that truncates the signal to the interval $[\theta - T/2 \;\; \theta + T/2]$.

Continuing in this vein, suppose now that the parameter θ affects changes in frequency via a modulation with the function $e^{j\theta}$ (or equivalently, a shift in the frequency domain). We can use a similar approach as before to show polynomial-regularity. The ψ_k functions are not time-limited, but are closely related to functions

that are. We write:

$$\begin{aligned}
\|\psi_{k;\theta_1} - \psi_{k;\theta_2}\|^2 &= \lambda_k^{-2}\|\mathbf{P}_\Omega \mathbf{P}_T(\psi_{k;\theta_1} - \psi_{k;\theta_2})\|^2 && (150)\\
&\leq 4\|\mathbf{P}_T(\psi_{k;\theta_1} - \psi_{k;\theta_2})\|^2 && (151)\\
&\leq 4\int_{-T/2}^{T/2} (\mathbf{P}_T\psi_k(t))^2 |e^{j\theta_s t/2} - e^{-j\theta_s t/2}|^2 dt && (152)\\
&\leq 4\int_{-T/2}^{T/2} 4\sin(\theta_s T/4)^2 && (153)\\
&\leq (\theta_s T)^2, && (154)
\end{aligned}$$

so that as before, we can cover a frequency range of $\Theta = [\omega_a \;\; \omega_b]$ with base covering $N_0 = |\omega_b - \omega_a|T$ and effective dimension $d = 1$. Also, as before with the time-shifted case, we can equivalently apply this result to projections over a variable frequency range of bandwidth Ω.

Now suppose we construct a parameterized set with parameters controlling both time-shift and frequency shift, so that the parametric estimation essentially jointly searches for the time interval and frequency interval that best contains the compressed signal. In this case, we can bound the difference between basis functions as:

$$\begin{aligned}
& \|e^{j\theta_1^{(1)}}\psi_k(t-\theta_1^{(2)}) - e^{j\theta_2^{(1)}}\psi_k(t-\theta_2^{(2)})\| \\
\leq\; & \|e^{j\theta_1^{(1)}}\psi_k(t-\theta_1^{(2)}) - e^{j\theta_1^{(1)}}\psi_k(t-\theta_2^{(2)})\| + \|e^{j\theta_1^{(1)}}\psi_k(t-\theta_2^{(2)}) - e^{j\theta_2^{(1)}}\psi_k(t-\theta_2^{(2)})\| \\
\leq\; & |\theta_2^{(1)} - \theta_1^{(1)}|T + \Omega|\theta_2^{(2)} - \theta_1^{(2)}|/2
\end{aligned}$$

where the parameter vector $\theta = \begin{bmatrix}\theta^{(1)} \\ \theta^{(2)}\end{bmatrix}$ describes changes in both frequency and time. Consequently, because of the consistency between ℓ_1 and ℓ_2 norms, we have polynomial-regularity with base covering $N_0 = \sqrt{2}(|\omega_b - \omega_a|T + \Omega|t_b - t_a|/2)$ and effective dimension $d = 2$ for the class of parameterized subspaces with basis functions $e^{j\theta^{(1)}t}\psi_k(t - \theta^{(2)})$ for $\theta \in \Theta = [\omega_a \;\; \omega_b] \times [t_a \;\; t_b]$.

4.4 Analysis

In this section, we establish the theoretical framework that will help us to prove the main results. We start by establishing definitions and conventions, then proceed to use a chaining argument to establish uniform probabilistic bounds on processes related to the quantities of interest.

4.4.1 Definitions and Conventions

The following definitions will be used throughout this section:

$$\mathbf{P}_\theta = \mathbf{V}_\theta \mathbf{V}_\theta^T \tag{155}$$

$$\tilde{\mathbf{P}}_\theta = (\mathbf{\Phi V}_\theta)((\mathbf{\Phi V}_\theta)^T(\mathbf{\Phi V}_\theta))^{-1}(\mathbf{\Phi V}_\theta)^T \tag{156}$$

$$W_\theta = \|\tilde{\mathbf{P}}_\theta \mathbf{\Phi} h\|^2 - \|\mathbf{P}_\theta h\|^2 \tag{157}$$

$$G_\theta = \mathbf{P}_\theta - \mathbf{P}_\theta \mathbf{\Phi}^T \mathbf{\Phi} \mathbf{P}_\theta \tag{158}$$

$$\delta_G = \sup_{\theta \in \Theta} \|\mathbf{P}_\theta - \mathbf{P}_\theta \mathbf{\Phi}^T \mathbf{\Phi} \mathbf{P}_\theta\| = \sup_{\theta \in \Theta} \|G_\theta\|. \tag{159}$$

All parameter distances here are defined with respect to their corresponding projection operators:

$$d(\theta_1, \theta_2) \triangleq \|\mathbf{P}_{\theta_1} - \mathbf{P}_{\theta_2}\| \tag{160}$$

In order to bound the difference between $\bar{E}$ and $\hat{E}$, we will relate their difference to the W process as follows:

$$\begin{aligned}
\hat{E}^2 - \bar{E}^2 &= (\|h\|^2 - \|\mathbf{P}_{\hat{\theta}} h\|^2) - (\|h\|^2 - \|\mathbf{P}_{\bar{\theta}} h\|^2) && (161)\\
&\leq (\|\tilde{\mathbf{P}}_{\hat{\theta}} \mathbf{\Phi} h\|^2 - \|\mathbf{P}_{\hat{\theta}} h\|^2) - (\|\tilde{\mathbf{P}}_{\bar{\theta}} \mathbf{\Phi} h\|^2 - \|\mathbf{P}_{\bar{\theta}} h\|^2) && (162)\\
&\leq \sup_{\theta \in \Theta} W_\theta - W_{\bar{\theta}} && (163)\\
&= \sup_{\theta \in \Theta} \langle h, (\mathbf{\Phi}^T(\tilde{\mathbf{P}}_\theta - \tilde{\mathbf{P}}_{\bar{\theta}})\mathbf{\Phi} - (\mathbf{P}_\theta - \mathbf{P}_{\bar{\theta}}))h \rangle. && (164)
\end{aligned}$$

To this end, we begin by establishing a chaining argument that will help us to analyze this W process.

4.4.2 Chaining

As discussed earlier in Chapter 4.1.1, we can analyze the supremum of a stochastic process using finite union bounds over progressively denser subsets The following chaining argument utilizes polynomial-regularity to construct a supremum bound characteristically similar to the increment tail bound.

Proposition 1 *Suppose we have an increment bound of the following form on the random process $L(\theta)$:*

$$\mathrm{P}\left\{\|L(\theta_2) - L(\theta_1)\| \geq l(u)d(\theta_1, \theta_2)\right\} \leq C_a e^{-u}, \tag{165}$$

for some concave function $l(u)$ and constant C_a. Then the following supremum bound holds:

$$\mathrm{P}\left\{\sup_{\theta\in\Theta} \|L(\theta) - L(\theta_0)\| \geq 3l(u)\right\} \leq C_a N_0^2 8^d e^{-u+1}, \tag{166}$$

for any fixed $\theta_0 \in \Theta$.

Proof This proof adapts a similar one of a more general form from Talagrand [25, Theorem 1.2.7]. We first define $T_0 = \{\theta_0\}$, and $\{T_j\}_{j\geq 1}$ as a series of ϵ-nets of Θ with radius 2^{-j} with respect to projection distance d, so that the cardinality of the jth ϵ-net is at most $N_0 2^{jd}$. For any $\theta \in \Theta$, we define $\pi_j(\theta)$ as the closest member in T_j to the parameter θ, so that $d(\theta, \pi_j(\theta)) \leq 2^{-j}$. Consequently, we have $d(\pi_{j+1}(\theta), \pi_j(\theta)) \leq d(\pi_{j+1}(\theta), \theta) + d(\theta, \pi_j(\theta)) \leq \frac{3}{2} \cdot 2^{-j}$. By defining the sequence:

$$a_j \;\; = \;\; \log(2)(2j+1)d + 2\log(N_0) + j$$

then we have, by Eq. (165):

$$\mathrm{P}\left\{|L(\pi_{j+1}(\theta)) - L(\pi_j(\theta))| \geq \frac{3}{2} 2^{-j} l(u + a_j)\right\} \leq C_a \exp(-a_j - u). \tag{167}$$

Also, because l is concave, we have:

$$\sum_{j=0}^{\infty} 2^{-j-1} l(u + a_j) \;\; \leq \;\; l(3\log(2)d + 2\log(N_0) + u + 1). \tag{168}$$

This is because, by definition, $l(\lambda x_1 + (1-\lambda)x_2) \geq \lambda l(x_1) + (1-\lambda)l(x_2)$, and by inductive substitution we have $l(\sum_j \lambda_j x_j) \geq \sum_j \lambda_j l(x_j)$ for some convex combination defined by the λ_j (i.e., with $\lambda_j \geq 0$ and $\sum_j \lambda_j = 1$). Now, by summing these probabilities over all scales j and over all L_j links for each scale, we have:

$$\begin{aligned}
& \mathrm{P}\left\{\sup_{\theta\in\Theta} |L(\theta) - L(\theta_0)| \geq 3l(3\log(2)d + 2\log(N_0) + 1 + u)\right\} \\
\leq\ & \mathrm{P}\left\{\sup_{\theta\in\Theta} |L(\theta) - L(\theta_0)| \geq \sum_{j\geq 0} \frac{3}{2} 2^{-j} l(u + a_j)\right\} \\
\leq\ & \sum_{j\geq 0} \mathrm{P}\left\{\sup_{\theta\in\Theta} |L(\pi_{j+1}(\theta)) - L(\pi_j(\theta))| \geq \frac{3}{2} 2^{-j} l(u + a_j)\right\} \\
\leq\ & \sum_{j\geq 0} C_a L_j e^{-a_j} e^{-u} \leq C_a e^{-u}\left(\frac{1}{N_0} + \sum_{j\geq 1} e^{-j}\right) \leq C_a e^{-u},
\end{aligned}$$

where the number of combinatorial "links" L_j between the elements of T_j and T_{j+1} is the product of their cardinalities, at most $e^{a_j - j}$. ■

4.4.3 Matrix Bernstein and Orlicz Norms

In order to build the chaining argument for various processes, we will first need a tail bound on its increment of the form (165). To this end, we will utilize the matrix Bernstein (or non-commutative Bernstein) inequality that depends on the Orlicz norm, defined as follows.

Definition 1 *The Ψ_1 Orlicz norm of a random matrix X is defined as:*

$$\|X\|_{\Psi_1} \triangleq \inf\{c : \mathrm{E}\left[\exp\left(\frac{\|X\|}{c}\right)\right] \leq 2\}. \tag{169}$$

where $\|.\|$ is the spectral norm.

There are generalizations of this Orlicz norm, but in this thesis, we will work exclusively with this particular Ψ_1 norm, which generalizes the Orlicz norm of scalar random variables to matrices, dealing exclusively with the spectral norms on these matrices. The following two lemmas show that a finite Orlicz norm is consistent with a sub-exponential tail bound.

Lemma 6 *[86, page 96] Let X be a random matrix with finite Ψ_1 norm. Then we have:*

$$\mathrm{P}\left\{\|X\| \geq \tau \|X\|_{\Psi_1}\right\} \leq 2e^{-\tau}. \tag{170}$$

Proof Using Markov's inequality, we have:

$$\mathrm{P}\left\{\|X\| \geq \tau \|X\|_{\Psi_1}\right\} = \mathrm{P}\left\{e^{\|X\|/\|X\|_{\Psi_1}} \geq e^{\tau}\right\} \leq \frac{\mathrm{E}\left[e^{\|X\|/\|X\|_{\Psi_1}}\right]}{e^{\tau}} \leq 2e^{-\tau}. \tag{171}$$

■

Conversely, an exponential tail bound of this type shows that $\|X\|_{\Psi_1}$ is finite.

Lemma 7 *[86, Lemma 2.2.1] Let X be a random matrix with $\mathrm{P}\left\{\|X\| > x\right\} \leq K_1 e^{-x/K_2}$ for every x, for constants K_1 and K_2. Then its Orlicz norm satisfies $\|X\|_{\Psi_1} \leq (K_1 + 1)K_2$.*

Proof By Fubini's theorem

$$\mathrm{E}\left[e^{\|X\|/D} - 1\right] = \mathrm{E}\left[\int_0^{\|X\|} D^{-1} e^{s/D} ds\right] = \int_0^{\infty} \mathrm{P}\left\{\|X\| \geq s\right\} D^{-1} e^{s/D} ds. \tag{172}$$

Now, insert the inequality on the tails of $\|X\|$ and obtain the explicit upper bound $K_1 K_2/(D - K_2)$. This is less than or equal to 1 for D greater than or equal to $(1 + K_1)K_2$. ■

Proposition 2 *(Matrix Bernstein, Orlicz norm version [87]) Let $X_1, ..., X_M$ be independent self-adjoint random matrices with dimension K with $\mathrm{E}\left[X_m\right] = 0$ with bounded Ψ_1 norms:*

$$\|X_m\|_{\Psi_1} \leq B. \tag{173}$$

Then, for all $t \geq 0$ and some universal constant C_B at most 4,

$$\mathrm{P}\left\{\left\|\sum_m X_m\right\| \geq C_B \max\left\{\sigma\sqrt{t}, 2Bt\log\left(\frac{2B\sqrt{M}}{\sigma}\right)\right\}\right\} \leq 2Ke^{-t} \quad \textit{where} \quad \sigma^2 := \left\|\sum_m \mathrm{E}\left[X_m^2\right]\right\|. \tag{174}$$

We give a proof of this in Section A.5 that follows the more general proof given in [87], which mirrors the derivation of classical Bernstein's inequality similarly to other works (e.g., [88, 89, 90]), and is repeated there for convenience. The corollary via matrix dilation to the non-symmetric case follows.

Proposition 3 *(Non-Symmetric Matrix Bernstein, Orlicz norm version [91]) Let* $X_1, ..., X_M$ *be independent random matrices with dimensions* $K_1 \times K_2$ *with* $\mathrm{E}[X_m] = 0$ *with bounded* Ψ_1 *norms:*

$$\|X_m\|_{\Psi_1} \leq B. \tag{175}$$

Then, for all $t \geq 0$ *and some universal constant* C_B *at most* 4*, we have*

$$\mathrm{P}\left\{\left\|\sum_m X_m\right\| \geq C_B \max\left\{\sigma\sqrt{t}, 2Bt\log\left(\frac{2B\sqrt{M}}{\sigma}\right)\right\}\right\} \leq 2(K_1+K_2)e^{-t}, \tag{176}$$

where

$$\sigma^2 := \max\left\{\left\|\sum_{m=1}^{M}\mathrm{E}\left[X_m^T X_m\right]\right\|, \left\|\sum_{m=1}^{M}\mathrm{E}\left[X_m X_m^T\right]\right\|\right\} \tag{177}$$

We will also use the fact that Orlicz norms of chi-squared random variables are well-approximated by their mean.

Lemma 8 *Let* X *be a chi-squared random variable with* M *degrees of freedom. Then the following consistency bound relates its mean to its Orlicz norm:*

$$\frac{2}{\log(4)} \leq \frac{\|X\|_{\Psi_1}}{\mathrm{E}[X]} = \frac{2}{M(1-4^{-1/M})} \leq 8/3. \tag{178}$$

Naturally, the bound holds under scalar multiplication of the random variable.

Proof This result follows directly from the monotonicity of the moment generating function of a chi-squared random variable with M degrees of freedom:

$$M_X(t) \triangleq \mathrm{E}[\exp(tX)] = (1-2t)^{-M/2}. \tag{179}$$

Now, we have $\mathrm{E}[X] = M$ and $\|X\|_{\Psi_1}^{-1} = M_X^{-1}(2) = (1-4^{-1/M})/2$, establishing the lemma. ∎

Lemma 9 *The Orlicz norm of the geometric mean of two independent random variables X and Y is at most the geometric mean of their Orlicz norms:*

$$\|\sqrt{XY}\|_{\Psi_1} \leq \sqrt{\|X\|_{\Psi_1}\|Y\|_{\Psi_1}} \tag{180}$$

Proof Let $X' = X/\|X\|_{\Psi_1}$ and $Y' = Y/\|Y\|_{\Psi_1}$. Then:

$$\begin{aligned} \mathrm{E}\left[\exp(\sqrt{\frac{XY}{\|X\|_{\Psi_1}\|Y\|_{\Psi_1}}})\right] &= \mathrm{E}\left[e^{\sqrt{X'Y'}}\right] && (181)\\ &\leq \mathrm{E}\left[e^{(X'+Y')/2}\right] && (182)\\ &= \mathrm{E}\left[e^{X'/2}\right]\mathrm{E}\left[e^{Y'/2}\right] && (183)\\ &\leq \sqrt{\mathrm{E}\left[e^{X'}\right]\mathrm{E}\left[e^{Y'}\right]} \leq 2. && (184) \end{aligned}$$

■

Lemma 10 *The compressed inner product is concentrated around its expectation:*

$$\mathrm{P}\left\{\frac{\langle \mathbf{\Phi}f, \mathbf{\Phi}g\rangle - \langle f, g\rangle}{\|f\|\|g\|} \geq C_1 u + C_2\sqrt{u}\right\} \leq 2e^{-u} \tag{185}$$

for some universal constants C_1 and C_2 at most 51 and $4\sqrt{2}$, respectively.

The proof of this is given in Section A.5.

4.4.4 Increment Bounds

In this section, we establish bounds on various increments (e.g., $W_{\theta_{j+1}} - W_{\theta_j}$) that may be used in conjunction with Proposition 1 to bound the maximal deviation of a stochastic process. For the sake of brevity, we write $\mathbf{P}_2$, $\mathbf{P}_1$ for $\mathbf{P}_{\theta_{j+1}}$, $\mathbf{P}_{\theta_j}$ etc. and also write $\Delta := \|\mathbf{P}_2 - \mathbf{P}_1\|$.

Lemma 11 *The G increment may be bounded as follows:*

$$\mathrm{P}\left\{\|G_2 - G_1\| \geq C_B g(t)\|\mathbf{P}_2 - \mathbf{P}_1\|\right\} \leq 2K\exp(-t) \tag{186}$$

where

$$g(t) = \sqrt{\frac{8t(K+1)}{M}} + \frac{t(320K+6)\log(116\sqrt{K})}{M}. \tag{187}$$

Proof We write the increment as a sum of independent random matrices, and then use the matrix Bernstein inequality to establish a tail bound. We have

$$\begin{aligned} G_2 - G_1 &= \mathbf{P}_1 - \mathbf{P}_2 - \mathbf{P}_1\mathbf{\Phi}^{\mathrm{T}}\mathbf{\Phi}\mathbf{P}_1 + \mathbf{P}_2\mathbf{\Phi}^{\mathrm{T}}\mathbf{\Phi}\mathbf{P}_2 \\ &= \sum_{m=1}^{M} \frac{1}{M}\mathbf{P}_1 - \frac{1}{M}\mathbf{P}_2 - \mathbf{P}_1\phi_m\phi_m^{\mathrm{T}}\mathbf{P}_2 + \mathbf{P}_2\phi_m\phi_m^{\mathrm{T}}\mathbf{P}_2 \\ &=: \sum_{m=1}^{M} X_m, \end{aligned}$$

where the ϕ_m are rows of $\mathbf{\Phi}$.

Notice that since $\mathrm{E}\left[\mathbf{P}_1\phi_m\phi_m^{\mathrm{T}}\mathbf{P}_1\right] = M^{-1}\mathbf{P}_1$ and similarly $\mathrm{E}\left[\mathbf{P}_2\phi_m\phi_m^{\mathrm{T}}\mathbf{P}_2\right] = M^{-1}\mathbf{P}_2$, the X_m are zero mean. The matrix Bernstein inequality depends on the variance term

$$\left\| \sum_{m=1}^{M} \mathrm{E}\left[X_m^2\right] \right\| = \frac{1}{M} \| \mathrm{E}\left[X^2\right] \|$$

where X is the random matrix

$$X = \mathbf{P}_1 - \mathbf{P}_2 - \mathbf{P}_1\phi\phi^{\mathrm{T}}\mathbf{P}_1 + \mathbf{P}_2\phi\phi^{\mathrm{T}}\mathbf{P}_2, \tag{188}$$

where ϕ is a random vector whose entries are Normal$(0, 1)$. We compute

$$\begin{aligned} \mathrm{E}\left[X^2\right] &= \mathrm{E}\left[(\mathbf{P}_2\phi\phi^T\mathbf{P}_2 - \mathbf{P}_1\phi\phi^T\mathbf{P}_1\right] - \mathrm{E}\left[\mathbf{P}_2\phi\phi^T\mathbf{P}_2 - \mathbf{P}_1\phi\phi^T\mathbf{P}_1])^2\right] \\ &= \mathrm{E}\left[(\mathbf{P}_2\phi\phi^T\mathbf{P}_2 - \mathbf{P}_1\phi\phi^{\mathrm{T}}\mathbf{P}_1)^2\right] - (\mathbf{P}_1 - \mathbf{P}_2)^2. \end{aligned}$$

To make computing the expectation above a little less unwieldy, we introduce the sum and difference matrices

$$\begin{aligned} S &= \mathbf{P}_1 + \mathbf{P}_2 \\ D &= \mathbf{P}_1 - \mathbf{P}_2, \end{aligned}$$

and so

$$\begin{aligned} (\mathbf{P}_1\phi\phi^{\mathrm{T}}\mathbf{P}_1 - \mathbf{P}_2\phi\phi^{\mathrm{T}}\mathbf{P}_2)^2 &= \frac{1}{4}\left(S\phi\phi^{\mathrm{T}}D + D\phi\phi^{\mathrm{T}}S\right)^2 \\ &= \frac{1}{4}\left(S\phi\phi^{\mathrm{T}}DS\phi\phi^{\mathrm{T}}D + S\phi\phi^{\mathrm{T}}D^2\phi\phi^{\mathrm{T}}S + D\phi\phi^{\mathrm{T}}S^2\phi\phi^{\mathrm{T}}D + D\phi\phi^{\mathrm{T}}SD\phi\phi^{\mathrm{T}}S\right) \end{aligned}$$

Using Lemma 14, we have

$$\begin{aligned}
\mathrm{E}\left[S\phi\phi^{\mathrm{T}}DS\phi\phi^{\mathrm{T}}D\right] &= S\,\mathrm{E}\left[\phi\phi^{\mathrm{T}}DS\phi\phi^{\mathrm{T}}\right]D \\
&= S\left(DS + SD + \mathrm{trace}(DS)I\right)D \\
&= SDSD + S^2D^2 + \mathrm{trace}(DS)SD \\
\mathrm{E}\left[D\phi\phi^{\mathrm{T}}SD\phi\phi^{\mathrm{T}}S\right] &= DSDS + D^2S^2 + \mathrm{trace}(DS)DS \\
\mathrm{E}\left[S\phi\phi^{\mathrm{T}}D^2\phi\phi^{\mathrm{T}}S\right] &= S(2D^2 + \mathrm{trace}(D^2)I)S \\
&= 2SD^2S + \mathrm{trace}(D^2)S^2 \\
\mathrm{E}\left[D\phi\phi^{\mathrm{T}}S^2\phi\phi^{\mathrm{T}}D\right] &= 2DS^2D + \mathrm{trace}(S^2)D^2,
\end{aligned}$$

and so

$$\begin{aligned}
\mathrm{E}\left[X^2\right] = \frac{1}{4}\big(&SDSD + S^2D^2 + \mathrm{trace}(DS)SD + DSDS + D^2S^2 + \\
&\mathrm{trace}(DS)DS + 2SD^2S + \mathrm{trace}(D^2)S^2 + 2DS^2D + \mathrm{trace}(S^2)D^2 - 4D^2\big) \\
= \frac{1}{4}\big(&SDSD + (S^2 - 4I)D^2 + \mathrm{trace}(DS)SD + DSDS + D^2S^2 + \\
&\mathrm{trace}(DS)DS + 2SD^2S + \mathrm{trace}(D^2)S^2 + 2DS^2D + \mathrm{trace}(S^2)D^2\big).
\end{aligned}$$

Defining $\Delta = \|D\| = \|P_1 - P_2\|$ and using the facts that

$$\begin{aligned}
\|S\| = \|P_1 + P_2\| &\le 2, \\
\mathrm{trace}(DS) &\le \|D\|_F\|S\|_F \le \sqrt{2K}\sqrt{2K}\Delta 2 = 4K\Delta, \\
\mathrm{trace}(D^2) &\le 2K\Delta^2, \\
\mathrm{trace}(S^2) &\le 8K,
\end{aligned}$$

we have the bound

$$\begin{aligned}
\|\,\mathrm{E}\left[X^2\right]\,\| &\le \frac{1}{4}\left(4\Delta^2 + 4\Delta^2 + 8K\Delta^2 + 4\Delta^2 + 4\Delta^2 + 8K\Delta^2 + 8\Delta^2 + 8K\Delta^2 + 8\Delta^2 + 8K\Delta^2\right) \\
&= 8(K+1)\Delta^2.
\end{aligned}$$

Thus

$$\left\|\sum_{m=1}^{M}\mathrm{E}\left[X_m^2\right]\right\| \le \frac{8(K+1)}{M}\|\mathbf{P}_1 - \mathbf{P}_2\|^2$$

The other ingredient for matrix Bernstein is a uniform bound on the Orlicz norms of the X_m (or equivalently, X). We know that

$$\begin{aligned}
\|X\| &= \|\mathbf{P}_1(I-\phi\phi^{\mathrm{T}})\mathbf{P}_1 - \mathbf{P}_2(I-\phi\phi^{\mathrm{T}})\mathbf{P}_2\| \\
&= \left\|\frac{1}{2}\left(S(I-\phi\phi^{\mathrm{T}})D + D(I-\phi\phi^{\mathrm{T}})S\right)\right\| \\
&\leq \|S(I-\phi\phi^{\mathrm{T}})D\| \\
&\leq \|S\|\,\|D\| + \|S\phi\|_2\|D\phi\|_2.
\end{aligned}$$

It is a standard result (see, e.g., [86, proposition A.2.1]) that

$$\begin{aligned}
\mathrm{P}\left\{\|S\phi\|_2 > u\right\} &\leq 2\exp\left(-\frac{u^2}{8\|S\|_F^2}\right) \leq 2\exp\left(-\frac{u^2}{64K}\right), \\
\mathrm{P}\left\{\|D\phi\|_2 > u\right\} &\leq 2\exp\left(-\frac{u^2}{8\|D\|_F^2}\right) \leq 2\exp\left(-\frac{u^2}{16K\Delta^2}\right),
\end{aligned}$$

and so

$$\begin{aligned}
\mathrm{P}\left\{\|S\phi\|_2\|D\phi\|_2 > t\right\} &\leq \mathrm{P}\left\{\|S\phi\|_2 > \sqrt{2t/\Delta}\right\} + \mathrm{P}\left\{\|D\phi\|_2 > \sqrt{\Delta t/2}\right\} \\
&\leq 4\exp\left(-\frac{t}{32K\Delta}\right).
\end{aligned}$$

Thus

$$\begin{aligned}
\|X\|_{\psi_1} &\leq \frac{2}{\log 2}\Delta + \|\|S\phi\|_2\|D\phi\|_2\|_{\psi_1} \\
&\leq 3\Delta + 160K\Delta,
\end{aligned}$$

and

$$\|X_m\|_{\psi_1} \leq \frac{(160K+3)}{M}\,\|\mathbf{P}_1 - \mathbf{P}_2\|.$$

Now applying Proposition 2 with these two bounds on the Orlicz norm and variance term, we have:

$$\mathrm{P}\left\{\|G_2 - G_1\| \geq C_B g(t)\|\mathbf{P}_2 - \mathbf{P}_1\|\right\} \leq 2K\exp(-t), \tag{189}$$

as desired. ■

Lemma 12

$$\mathrm{P}\left\{\|\mathbf{P}_2\mathbf{\Phi}^T\mathbf{\Phi}\mathbf{P}_2^{\perp}h - \mathbf{P}_1\mathbf{\Phi}^T\mathbf{\Phi}\mathbf{P}_1^{\perp}h\| \geq C_B g^{\perp}(t)\|\mathbf{P}_2 - \mathbf{P}_1\|\right\} \leq (4K+2)\exp(-t), \quad (190)$$

where

$$g^{\perp}(t) = \sqrt{\frac{8t(K+1)}{M}} + \frac{758\sqrt{K}t}{M}. \quad (191)$$

Proof

We can represent:

$$\mathbf{P}_2\mathbf{\Phi}^T\mathbf{\Phi}\mathbf{P}_2^{\perp}h - \mathbf{P}_1\mathbf{\Phi}^T\mathbf{\Phi}\mathbf{P}_1^{\perp}h = \frac{1}{M}\sum_m x_m, \quad (192)$$

where

$$x_m = \mathbf{P}_2\phi\phi^T\mathbf{P}_2^{\perp}h - \mathbf{P}_1\phi\phi^T\mathbf{P}_1^{\perp}h, \quad (193)$$

and then control this using Vector Bernstein (i.e., Proposition 3) yielding:

$$\mathrm{P}\left\{\left\|\frac{1}{M}\sum_m x_m\right\| \geq C_B\Delta\left(\sqrt{\frac{8t(K+1)}{M}} + 182\log(182/\sqrt{8})\frac{t\sqrt{K}}{M}\right)\right\} \leq (4K+2)e^{-t}, \quad (194)$$

as desired using the bound on the variance term $\sigma^2 \leq 8\Delta^2(K+1)/M$ and Orlicz norm $\|x_m\|_{\Psi_1} \leq 91\Delta\sqrt{K}/M$, which remain to be shown. Note also that the dimension utilized in Vector Bernstein is $2K \times 1$ by using the argument:

$$\|\mathbf{P}_2\mathbf{\Phi}^T\mathbf{\Phi}\mathbf{P}_2^{\perp}h - \mathbf{P}_1\mathbf{\Phi}^T\mathbf{\Phi}\mathbf{P}_1^{\perp}h\| = \|\mathbf{V}_{12}^T(\mathbf{P}_2\mathbf{\Phi}^T\mathbf{\Phi}\mathbf{P}_2^{\perp}h - \mathbf{P}_1\mathbf{\Phi}^T\mathbf{\Phi}\mathbf{P}_1^{\perp}h)\|, \quad (195)$$

where $\mathbf{V}_{12} : \mathbb{R}^{2K} \to L_2$ is an orthobasis for the direct sum of $\mathcal{S}_1$ and $\mathcal{S}_2$.

For the variance term, since the x_m are i.i.d. and zero mean, we have that

$$\left\|\sum_{m=1}^{M}\mathrm{E}\left[x_m x_m^T\right]\right\| \leq \mathrm{trace}(\mathrm{E}\left[x_m x_m^T\right]) = \left\|\sum_{m=1}^{M}\mathrm{E}\left[x_m^T x_m\right]\right\| = \frac{1}{M}\mathrm{E}\left[\|x\|_2^2\right], \quad (196)$$

where

$$x = \frac{1}{2}\left(D\phi\phi^T Th - S\phi\phi^T Dh\right), \quad (197)$$

and ϕ is a Gaussian random vector whose entries are independent and have unit variance, and

$$\begin{aligned} D &= \mathbf{P}_1 - \mathbf{P}_2 \quad (\text{and also } D = \mathbf{P}_2^{\perp} - \mathbf{P}_1^{\perp}) \\ S &= \mathbf{P}_1 + \mathbf{P}_2 \\ T &= \mathbf{P}_1^{\perp} + \mathbf{P}_2^{\perp}. \end{aligned}$$

To begin, note that

$$\|x\|_2^2 = \frac{1}{4}\left(h^T T\phi\phi^T D^2 \phi\phi^T Th - 2h^T T\phi\phi^T DS\phi\phi^T Dh + h^T D\phi\phi^T S^2\phi\phi^T Dh\right).$$

Treating each of these terms separately,

$$\begin{aligned} \mathrm{E}\left[h^T T\phi\phi^T D^2\phi\phi^T Th\right] &= h^T T(2D^2 + \mathrm{trace}(D^2)I)Th \\ &\leq 2\Delta^2\|Th\|_2^2 + \mathrm{trace}(D^2)\|Th\|_2^2 \\ &\leq 8(K+1)\Delta^2, \end{aligned}$$

since $\|Th\|_2^2 \leq 4\|h\|_2^2 = 4$. For the second term,

$$\begin{aligned} \left|\mathrm{E}\left[h^T T\phi\phi^T DS\phi\phi^T Dh\right]\right| &= \left|h^T T(DS + SD + \mathrm{trace}(DS)I)Dh\right| \\ &\leq (2\|S\|\,\|D\| + \mathrm{trace}(DS))\,\|Th\|_2\|Dh\|_2 \\ &\leq (4\Delta + 4K\Delta)\,2\Delta \\ &= 8(K+1)\Delta^2. \end{aligned}$$

Finally for the third term,

$$\begin{aligned} \mathrm{E}\left[h^T D\phi\phi^T S^2\phi\phi^T Dh\right] &= h^T D(2S^2 + \mathrm{trace}(S^2)I)Dh \\ &\leq (8 + 8K)\|Dh\|_2^2 \\ &\leq 8(K+1)\Delta^2. \end{aligned}$$

Collecting these results means that the variance is

$$\left\|\sum_{m=1}^{M} \mathrm{E}\left[x_m^T x_m\right]\right\| \;\leq\; \frac{8(K+1)}{M}\,\|\mathbf{P}_1 - \mathbf{P}_2\|^2. \tag{198}$$

Next we need to bound the Orlicz-1 norm of the x_m. Note that

$$\|x_m\|_{\psi_1} = \frac{\|x\|_{\psi_1}}{M}, \tag{199}$$

where $x = D\phi\phi^T Th - S\phi\phi^T Dh$ as above. We will bound each part of x separately. For the first term, we note that $\phi^T Th$ is a Gaussian random scalar with variance $\|Th\|_2^2$, and so

$$\mathrm{P}\left\{|\phi^T Th| > u\right\} \;\leq\; \exp\left(-\frac{u^2}{2\|Th\|_2^2}\right) \;\leq\; \exp\left(-\frac{u^2}{8}\right). \tag{200}$$

Since $D\phi$ is itself a Gaussian random vector, we have the bound

$$\begin{aligned}\mathrm{P}\left\{\|D\phi\|_2 > u\right\} &\leq\; 2\exp\left(-\frac{u^2}{8\|D\|_F^2}\right) \\ &\leq 2\exp\left(-\frac{u^2}{16K\Delta^2}\right).\end{aligned}$$

Then for any $u > 0$,

$$\mathrm{P}\left\{\|\langle\phi, Th\rangle D\phi\|_2 > t\right\} \leq \exp\left(-\frac{u^2}{8}\right) + 2\exp\left(-\frac{t^2}{u^2 16K\Delta^2}\right),$$

and taking $u = t^{1/2}(2K\Delta^2)^{-1/4}$ yields

$$\mathrm{P}\left\{\|\langle\phi, Th\rangle D\phi\|_2 > t\right\} \leq 3\exp\left(-\frac{t}{8\sqrt{2K}\Delta}\right),$$

and so

$$\|\langle\phi, Th\rangle D\phi\|_{\psi_1} \leq 32\sqrt{2K}\Delta.$$

For the second part of x, we have

$$\mathrm{P}\left\{|\phi^T Dh| > u\right\} \leq \exp\left(-\frac{u^2}{2\|Dh\|_2^2}\right) \;\leq\; \exp\left(-\frac{u^2}{2\Delta^2}\right),$$

and

$$\mathrm{P}\left\{\|S\phi\|_2 > u\right\} \leq 2\exp\left(-\frac{u^2}{8\|S\|_F^2}\right) \;\leq\; 2\exp\left(-\frac{u^2}{64K}\right),$$

and so

$$\begin{aligned}\mathrm{P}\left\{\|\langle\phi, Dh\rangle\|S\phi\|_2 > t\right\} &\leq \mathrm{P}\left\{|\langle\phi, Dh\rangle| > \frac{\sqrt{\Delta t}}{2(2K)^{1/4}}\right\} + \mathrm{P}\left\{\|S\phi\|_2 > \frac{2(2K)^{1/4}\sqrt{t}}{\sqrt{\Delta}}\right\} \\ &\leq 3\exp\left(-\frac{t}{8\sqrt{2K}\Delta}\right),\end{aligned}$$

and so

$$\|\langle\phi, Dh\rangle S\phi\|_{\psi_1} \leq 32\sqrt{2K}\Delta,$$

and finally

$$\|x_m\|_{\psi_1} \;\leq\; \frac{91\sqrt{K}}{M}\,\|\mathbf{P}_1 - \mathbf{P}_2\|.$$

■

4.4.4.1 W increment

Lemma 13 *The increment on W is as follows:*

$$\mathrm{P}\left\{W_2 - W_1 \geq w(t)\|\mathbf{P}_2 - \mathbf{P}_1\|\right\} \leq (4K+6)e^{-t}. \tag{201}$$

where

$$w(t) \;=\; 4(C_1 t/M + C_2\sqrt{t/M}) + C_B g^{\perp}(t)/(1-\delta_G) \tag{202}$$

and $g^{\perp}(t)$ is defined above in Eq. (191)*.*

Proof First, we break h into a pair of orthogonal decompositions as follows: $h = h_1 + h_1^{\perp} = h_2 + h_2^{\perp}$ where $h_1 = \mathbf{P}_1 h$ and $h_2 = \mathbf{P}_2 h$. Then, using the fact that $\tilde{\mathbf{P}}\mathbf{\Phi}\mathbf{P} = \mathbf{\Phi}\mathbf{P}$, we have:

$$\begin{aligned} W_2 - W_1 &= \langle\mathbf{\Phi}h, (\tilde{\mathbf{P}}_2 - \tilde{\mathbf{P}}_1)\mathbf{\Phi}h\rangle - \langle h, (\mathbf{P}_2 - \mathbf{P}_1)h\rangle && (203)\\ &= \|\mathbf{\Phi}h_1\|^2 - \|h_1\|^2 - (\|\mathbf{\Phi}h_2\|^2 - \|h_2\|^2) && (204)\\ &\quad +2(\langle\mathbf{\Phi}(h_2 - h_1), \mathbf{\Phi}h\rangle - \langle h_2 - h_1, h\rangle) && (205)\\ &\quad +\|\tilde{\mathbf{P}}_2\mathbf{\Phi}h_2^{\perp}\|^2 - \|\tilde{\mathbf{P}}_1\mathbf{\Phi}h_1^{\perp}\|^2 && (206)\end{aligned}$$

The first two terms may be dealt with using Lemma 10 as follows:

$$\begin{aligned} \mathrm{P}\left\{ \frac{\|\mathbf{\Phi} h_1\|^2 - \|h_1\|^2 - (\|\mathbf{\Phi} h_2\|^2 - \|h_2\|^2)}{\|h_2 + h_1\| \|h_2 - h_1\|} \geq C_1 u + C_2 \sqrt{u} \right\} &\leq 2e^{-u} \\ \mathrm{P}\left\{ \frac{\langle \mathbf{\Phi}(h_2 - h_1), \mathbf{\Phi} h \rangle - \langle h_2 - h_1, h \rangle}{\|h_2 - h_1\|} \geq C_1 u + C_2 \sqrt{u} \right\} &\leq 2e^{-u}. \end{aligned}$$

The last term may be bounded as:

$$\begin{aligned} & \|\tilde{\mathbf{P}}_2 \mathbf{\Phi} h_2^{\perp}\|^2 - \|\tilde{\mathbf{P}}_1 \mathbf{\Phi} h_1^{\perp}\|^2 && (207) \\ \leq\ & (\|\tilde{\mathbf{P}}_2 \mathbf{\Phi} h_2^{\perp}\| - \|\tilde{\mathbf{P}}_1 \mathbf{\Phi} h_1^{\perp}\|)(\|\tilde{\mathbf{P}}_2 \mathbf{\Phi} h_2^{\perp}\| + \|\tilde{\mathbf{P}}_1 \mathbf{\Phi} h_1^{\perp}\|) && (208) \\ \leq\ & (1 - \delta_G)^{-1/2} (\|\mathbf{P}_2 \mathbf{\Phi}^T \mathbf{\Phi} h_2^{\perp}\| - \|\mathbf{P}_1 \mathbf{\Phi}^T \mathbf{\Phi} h_1^{\perp}\|)(\|\tilde{\mathbf{P}}_2 \mathbf{\Phi} h_2^{\perp}\| + \|\tilde{\mathbf{P}}_1 \mathbf{\Phi} h_1^{\perp}\|) && (209) \\ \leq\ & \frac{2\sqrt{1+\delta_h} + 2\sqrt{1+\delta_G}}{\sqrt{1-\delta_G}} (\|\mathbf{P}_2 \mathbf{\Phi}^T \mathbf{\Phi} h_2^{\perp} - \mathbf{P}_1 \mathbf{\Phi}^T \mathbf{\Phi} h_1^{\perp}\|). && (210) \end{aligned}$$

By Lemma 12, this can be bounded as:

$$\mathrm{P}\left\{ \|\tilde{\mathbf{P}}_2 \mathbf{\Phi} h_2^{\perp}\|^2 - \|\tilde{\mathbf{P}}_1 \mathbf{\Phi} h_1^{\perp}\|^2 \geq \frac{4\sqrt{8} C_B}{1 - \delta_G} \|\mathbf{P}_2 - \mathbf{P}_1\| g^{\perp}(t) \right\} \leq (4K+2) \exp(-t). \quad (211)$$

Combining these gives the bound as desired. ■

Lemma 14 *Let $\phi \in \mathbb{R}^N$ be a random vector with $\phi[n] \sim \mathrm{Normal}(0,1)$, and let $\mathbf{A}$ be an arbitrary $N \times N$ matrix. Then*

$$\mathrm{E}\left[\phi \phi^{\mathrm{T}} \mathbf{A} \phi \phi^{\mathrm{T}}\right] = \mathbf{A} + \mathbf{A}^{\mathrm{T}} + \mathrm{trace}(\mathbf{A}) \cdot \mathbf{I}. \quad (212)$$

Proof Let $Q = \mathrm{E}\left[\phi \phi^{\mathrm{T}} \mathbf{A} \phi \phi^{\mathrm{T}}\right]$ be the matrix in question. An entry of Q can be written as

$$Q(j,k) = \mathrm{E}\left[(\phi^{\mathrm{T}} \mathbf{A} \phi) \phi(j) \phi(k)\right] = \sum_{n_1=1}^{N} \sum_{n_2=1}^{N} \mathbf{A}(n_1, n_2) \,\mathrm{E}\left[\phi(n_1) \phi(n_2) \phi(j) \phi(k)\right].$$

For an off-diagonal term, $j \neq k$, the expectation $\mathrm{E}\left[\phi(n_1) \phi(n_2) \phi(j) \phi(k)\right]$ is nonzero only when

$$(n_1 = j \text{ and } n_2 = k) \quad \text{or} \quad (n_1 = k \text{ and } n_2 = j). \quad (213)$$

Under either of these conditions (which do not overlap, since $j \neq k$), $\mathrm{E}\left[\phi(n_1)\phi(n_2)\phi(j)\phi(k)\right] = \mathrm{E}\left[|\phi(j)|^2\right]\mathrm{E}\left[|\phi(k)|^2\right] = 1$, and so

$$Q(j,k) = \mathbf{A}(j,k) + \mathbf{A}(k,j), \qquad j \neq k. \tag{214}$$

On the diagonal, when $j = k$, we have

$$Q(k,k) = \sum_{n_1=1}^{N}\sum_{n_2=1}^{N} \mathbf{A}(n_1,n_2)\,\mathrm{E}\left[\phi(n_1)\phi(n_2)\phi^2(k)\right]. \tag{215}$$

For the expectation in the expression above to be non-zero, we need $n_1 = n_2$, and so

$$\begin{aligned} Q(k,k) &= \sum_{n=1}^{N} \mathbf{A}(n,n)\,\mathrm{E}\left[\phi^2(n)\phi^2(k)\right] \\ &= \sum_{n=1}^{N} \mathbf{A}(n,n) - \mathbf{A}(k,k) + \mathrm{E}\left[\phi^4(k)\right]\mathbf{A}(k,k) \\ &= \mathrm{trace}(\mathbf{A}) + 2\mathbf{A}(k,k), \end{aligned} \tag{216}$$

since $\mathrm{E}\left[\phi^4(k)\right] = 3$. Combining (214) and (216) establishes the lemma. ■

4.4.5 Chaining the Processes

Applying Proposition 1 to the increment bounds established in the preceding section, we are now able to bound δ_G and consequently, $\sup_{\theta\in\Theta} W_\theta - W_{\bar{\theta}}$ using the following lemmas.

Lemma 15 *The uniform bound on δ_G is given as:*

$$\mathrm{P}\left\{\delta_G > 4C_B g(t)\right\} \leq 2K(8^d N_0^2 e + 1)e^{-t}. \tag{217}$$

for the $g(t)$ defined in Eq. (187)*.*

Proof A straightforward application of Proposition 1 with the increment bound in Lemma 11 yields:

$$\mathrm{P}\left\{\sup_{\theta} \|G_\theta - G_{\bar{\theta}}\| \geq 3C_B g(t)\right\} \;\leq\; 2K8^d N_0^2 e^{-t+1}, \tag{218}$$

In light of this, it only remains to prove:

$$\mathrm{P}\left\{\|G_{\bar{\theta}}\| \geq C_B(\sqrt{\frac{t(K+1)}{M}} + \frac{t(K+2)\log(2K+4)}{M}\right\} \leq 2Ke^{-t}, \tag{219}$$

because:

$$\sqrt{\frac{t(K+1)}{M}} + \frac{t(K+2)\log(2K+4)}{M} \leq g(t). \tag{220}$$

Similarly to Lemma 11, we employ the Orlicz norm version of Matrix Bernstein, by noting that $\mathbf{V}_{\bar{\theta}}^T \mathbf{\Phi}^T \mathbf{\Phi} \mathbf{V}_{\bar{\theta}} - \mathbf{I}$ is equal in distribution to the sum of M independently drawn copies of:

$$\mathbf{X} = (\mathbf{V}_{\bar{\theta}}^T \phi (\mathbf{V}_{\bar{\theta}}^T \phi)^T - \mathbf{I})/M. \tag{221}$$

Using a similar of line of reasoning to Lemma 11 yields $\| \mathrm{E}\left[\mathbf{X}^2\right] \| = \frac{K+1}{M}$ and $\|\mathbf{X}\|_{\Psi_1} \leq \frac{K+2}{M}$, yielding the requisite bound via Proposition 2 and the fact that $\|G_{\bar{\theta}}\| = \|\mathbf{V}_{\bar{\theta}}^T \mathbf{\Phi}^T \mathbf{\Phi} \mathbf{V}_{\bar{\theta}} - \mathbf{I}\|$. ■

Lemma 16 *The maximal difference of the W process is given probabilistically as:*

$$\mathrm{P}\left\{\sup_{\theta} \|W_\theta - W_{\bar{\theta}}\| \geq 3w(t)\right\} \leq (4K+6)N_0^2 8^d e^{-t+1} \tag{222}$$

where, as before:

$$w(t) = 4C_1 \frac{t}{M} + 4C_2 \sqrt{\frac{t}{M}} + \frac{C_B g^{\perp}(t)}{1-\delta_G}. \tag{223}$$

4.4.6 Main Theorems and their Proofs

With these tools established in the preceding sections, we are now able to prove the main results.

Proof (of Theorem 2)

The specific version of this inequality that is proven takes the following form:

$$\mathrm{P}\left\{\hat{E}^2 - \bar{E}^2 \geq C_0 \sqrt{\frac{(K+1)t}{M}} + C_0^2 \frac{K\log(116K)t}{M}\right\} \leq (8K+6)(8^d N_0^2 e + 1)e^{-t}.$$

where C_0 is a universal constant at most 100.

First, as above, we relate the difference in errors to the W process:

$$\begin{aligned}
\hat{E}^2 - \bar{E}^2 &= (\|h\|^2 - \|\mathbf{P}_{\hat{\theta}} h\|^2) - (\|h\|^2 - \|\mathbf{P}_{\bar{\theta}} h\|^2) && (224)\\
&\leq (\|\tilde{\mathbf{P}}_{\hat{\theta}} \mathbf{\Phi} h\|^2 - \|\mathbf{P}_{\hat{\theta}} h\|^2) - (\|\tilde{\mathbf{P}}_{\bar{\theta}} \mathbf{\Phi} h\|^2 - \|\mathbf{P}_{\bar{\theta}} h\|^2) && (225)\\
&\leq \sup_{\theta \in \Theta} W_\theta - W_{\bar{\theta}}. && (226)
\end{aligned}$$

Combining lemmas 15 and 16, while noting that $\hat{E}^2 - \bar{E}^2$ is always less than one, we have, with probability at least $1 - (8K+6)(8^d N_0^2 e + 1)e^{-t}$:

$$\begin{aligned}
\hat{E}^2 - \bar{E}^2 &\leq \min\left\{1,\ 12C_2\sqrt{\frac{t}{M}} + 12C_1\frac{t}{M} + \frac{C_B(\sqrt{\frac{8t(K+1)}{M}} + \frac{758\sqrt{K}t}{M})}{1 - 4C_B(\sqrt{\frac{8t(K+1)}{M}} + \frac{t(320K+6)\log(116\sqrt{K})}{M})}\right\}\\
&\leq 48\sqrt{\frac{t(K+1)}{M}} + \min\left\{1,\ 612\frac{tK}{M} + \frac{12\sqrt{\frac{t(K+1)}{M}} + \frac{3040\sqrt{K}t}{M}}{1 - 16(\sqrt{\frac{8t(K+1)}{M}} + \frac{t(320K+6)\log(116\sqrt{K})}{M})}\right\}\\
&\leq 100\sqrt{\frac{(K+1)t}{M}} + 10000\frac{K\log(116K)t}{M}
\end{aligned}$$

■

Proof (of Theorem 3)

There is at least one θ such that $\mathbf{P}_\theta f = f$. For this $\mathbf{P}_\theta$, we then have, using the same δ_G from Lemma 15:

$$|\|\mathbf{\Phi} f\|^2 - \|f\|^2| = |f^T(\mathbf{P}_\theta \mathbf{\Phi}^T \mathbf{\Phi} \mathbf{P}_\theta - \mathbf{P}_\theta) f| \leq \delta_G \|f\|^2. \quad (227)$$

■

CHAPTER V

CONCLUSIONS

Compressive sensing has opened up many avenues for new applications in sparse acquisition and underdetermined inverse problems beyond what had been previously thought possible. This thesis explores a variation on classical compressive sensing to establish properties of parametric subspace estimation, where these parametric subspaces exhibit a specific type of structure. This characteristic structure, called Hölder-regularity, is evidently common among many types of parametric estimation problems, and is sufficiently described in terms of an effective dimension and base covering. These two descriptors can often be intuitively estimated, providing immediate insight on the dependencies of the accuracy of the compressive estimator upon the parameters of the problem.

The exploration of this work was inspired primarily through the application of compressive matched field processing (cMFP), discussed in Chapter 2. Here, for a set of N hydrophones, we demonstrated how a series of $M < N$ randomized back-propagations could greatly reduce the computational complexity by reducing the number of necessary partial differential equations from N to M without significantly reducing the accuracy of the estimator. In this way, we essentially provide an easily-implemented tradeoff between accuracy and computational complexity. The simulations we ran indicated a logarithmic dependency on the area of the region of interest by the number of measurements M for comparable performance as well as an inverse-square-root relationship of M on the accuracy, an observation validated by work in Chapter 4.

This compressive approach also extends to the localization of multiple sources, as

discussed in Chapter 3. There, we presented a novel algorithm for multi-source localization that utilized a variation on OMP, but where the least-squares re-calculation step was performed with respect to a projection matrix that sought to minimize the interfering energy of the sources whose localizations had already been estimated, while taking into consideration the uncertainty in the existing estimates of those source locations. The design of this projection matrix ties in closely with related work on surface source suppression, but is much easier to compute, which is a vital property given Romulo's iterative nature. We also utilized recent work in randomized subspace sensing by Tropp to more rapidly compute our projection matrices, speeding the computation of this matrix up by more than a factor of 10.

The algorithm Romulo is essentially a greedy approximation to the type of compressive parametric estimation we discuss in this thesis, so performance guarantees of this algorithm have yet to be established. This is an area of future potential work, and may benefit from similar analysis on other greedy algorithms [61]. On the other hand, the global optimization discussed in Appendix A.4 falls under the framework studied in Chapter 4, but is only posed for the two-source case. Extending this heuristic proposed in Appendix A.4 to handle dozens or even hundreds of sources would be broadly useful and is another area of future potential interest.

In Chapter 4, we discussed the generalized compressive parametric estimation as a K-dimensional subspace estimation problem. We established novel results on the accuracy of this estimator under a natural assumption of regularity of the parameterized subspace, and showed how such assumptions were met for several practical cases of interest, including cMFP. This work essentially rests upon an application of a chaining argument along with the use of a recently developed version of matrix-Bernstein to prove the key result: that effective estimation becomes feasible when the number of compressive measurements M is taken to be at least a small factor of $K(d + \log(KN_0)$ for effective dimension d and base covering N_0 of the parameterized

class of subspaces. Although this work was shown specifically for i.i.d. Gaussian measurement operators that are in some sense ideal, much of this work could extend to more practical operators such as block-diagonal operators and Bernoulli operators as well, a topic of potential interest.

APPENDIX A

APPENDIX

A.1 Closest point on a line

For fixed vectors $U, V \in \mathbb{C}^n$, the following optimization program finds the closest point on the line spanned by v to u,

$$\min_{\beta \in \mathbb{C}} \ \|U - \beta V\|^2.$$

The functional above attains its minimum value of

$$\|U\|^2 - \frac{|V^H U|^2}{\|V\|^2}$$

when

$$\beta = \frac{V^H U}{\|V\|^2}.$$

This fact can be verified by differentiating $\|U - \beta V\|^2$ with respect to the real and imaginary parts of β, and solving for value of β that makes them both equal to zero.

A.2 Matching Pursuits

This section overviews matching pursuit (MP) and orthogonal matching pursuit (OMP), two seminal approaches in the field of greedy sparse approximation [60, 24].

The nature of the problem to be solved is as follows. Given a dictionary of column vectors $\mathbf{A} = [A_1 \ A_2 \ ... \ A_N] \in \mathbb{R}^{M \times N}$, usually of unit norm with $M \ll N$, the goal is to find $K < M$ columns $\mathbf{A}_K = [A_{n_1} \ A_{n_2} \ A_{n_3} ... A_{n_K}]$ and an associated model vector $x \in \mathbb{R}^K$ to approximate a data vector b. That is, that minimizes:

$$\|b - \mathbf{A}_K x\|^2 = \sum_{m=1}^{M} (b_m - \sum_{k=1}^{K} x_k A_{m,n_k})^2. \tag{228}$$

Essentially, matching pursuit loops the following steps ranging the variable k from 1 to K, after initializing the residual $r = b$:

1. $n_k = \arg\max |\langle r, A_n \rangle|$

2. $x_k = |\langle r, A_{n_k} \rangle|$

3. $r = b - \sum_{k'}^{k} x_{k'} A_{n_{k'}}$.

Orthogonal matching pursuit improves upon this approach by re-estimating *all* x_k at every iteration, instead of just the current one. To wit, step 2 is replaced by the following step

$$x = \arg \min_{x_1,\ldots,x_k} \|b - \sum_{k'}^{k} x_{k'} A_{n_{k'}}\|^2. \tag{229}$$

In this way, all elements of the model vector x are updated through joint re-estimation.

A.3 Matrix Filter Analysis and Comparisons

Although the proposed choice of construction for the projection matrix $\mathbf{P}$ from the subspace spanned by the first few principal components of the correlation matrix $\mathbf{Q}$ is relatively simple and intuitive, there is a stronger justification that depends on assumptions that approximately hold in many cases. In particular, when the attenuation factor over the region of interest scales linearly with the rank of the projection (i.e., proportional to $N - n_r$ for rank $N - n_r$ projections), then this choice of projection $\mathbf{P}$ coincides with the solution to the following minimization for certain values of z:

$$\text{minimize} \quad |E|^{-1} \int_E \|\mathbf{P}G(\vec{x})\|^2 d\vec{x} \tag{230}$$

$$\text{subject to} \quad |\mathcal{R}|^{-1} \int_{\mathcal{R}} \|\mathbf{P}G(\vec{x})\|^2 d\vec{x} \geq z \tag{231}$$

$$\sigma_{\max}(\mathbf{P}) \leq 1 \tag{232}$$

$$\mathbf{P} = \mathbf{P}^H \succeq 0, \tag{233}$$

where the objective function attempts to minimize $\mathrm{E}\left[\|\mathbf{P}\ G(\vec{r}_1)\|^2\right]$, the first constraint keeps from minimizing $\mathrm{E}\left[\|\mathbf{P}\ G(\vec{r}_2)\|^2\right]$ too much via parameter z, the second constraint requires $\mathbf{P}$ to be passive (i.e., so that $\|\mathbf{P}G(\vec{r})\| \leq 1$ for all $\vec{r} \in \mathcal{R}$), and where

the last constraint restricts the unnecessary extra degrees of freedom in $\mathbf{P}$ by forcing it to be symmetric and positive semidefinite. To see why this last constraint may be imposed without loss of generality, suppose that a matrix $\mathbf{P}$ with SVD decomposition $\mathbf{P} = \mathbf{G}\mathbf{\Sigma}\mathbf{V}^H$ is a minimizer of the objective function (230) under the first two constraints (231) and (232). Then it must also be the case that $\mathbf{P}_* = \mathbf{V}\mathbf{\Sigma}\mathbf{V}^H = \mathbf{V}\mathbf{G}^H\mathbf{P}$ is also a minimizer under those constraints because of the invariance of Euclidean norms and singular values under unitary transformation, and in particular under multiplication by unitary matrix $\mathbf{V}\mathbf{G}^H$.

This optimization is similar to the one proposed by Vaccaro et al. [64]. In fact, the main difference is that they replaced the average region attenuation factor with the minimum region attenuation factor. Although, they solve for this matrix-filter using an iterative log-barrier scheme, there now additionally exist specialized software packages that solve this problem as a semi-definite program (SDP) [7, 92, 93].

This optimization may be explicitly cast as an SDP as follows:

$$\begin{aligned} \text{minimize} \quad & \langle \mathbf{Q}_E, \mathbf{P} \rangle & & \\ \text{subject to} \quad & \langle \mathbf{Q}_{\mathcal{R}}, \mathbf{P} \rangle & \geq & \ z \\ & \mathbf{P} & \preceq & \ \mathbf{I} \\ & \mathbf{P} & \in & \ S_+^N, \end{aligned} \tag{234}$$

where the matrix inner products is given as $\langle \mathbf{A}, \mathbf{B} \rangle \triangleq \text{Tr}(\mathbf{A}^H\mathbf{B})$ and where S_+^N is the cone of symmetric positive semi-definite matrices of size $N \times N$.

When the attenuation factor $\text{Tr}(\mathbf{P}\mathbf{Q}_{\mathcal{R}})$ falls linearly with respect to the nulling rank n_r (see for example Fig. 15), then the matrix $\mathbf{Q}_{\mathcal{R}}$ is well-approximated as $\mathbf{Q}_{\mathcal{R}} \simeq$

$N_0^{-1}\mathbf{G}_{\mathcal{R}}\mathbf{G}_{\mathcal{R}}^H$ for some unitary $\mathbf{G}_{\mathcal{R}} \in \mathbb{C}^{N \times N_0}$. In this case, the optimization may be well-characterized by the following dimension-reduced minimization over $\hat{\mathbf{P}} = \mathbf{G}_{\mathcal{R}}^H \mathbf{P} \mathbf{G}_{\mathcal{R}}$:

$$\begin{aligned} \text{minimize} \quad & \langle \hat{\mathbf{Q}}_E, \hat{\mathbf{P}} \rangle \\ \text{subject to} \quad & \langle N_0^{-1}\mathbf{I}, \hat{\mathbf{P}} \rangle \geq z \\ & \hat{\mathbf{P}} \preceq \mathbf{I} \\ & \hat{\mathbf{P}} \in S_+^{N_0}, \end{aligned} \tag{235}$$

where $\hat{\mathbf{Q}}_E = \mathbf{G}_{\mathcal{R}}^H \mathbf{Q}_E \mathbf{G}_{\mathcal{R}}$. Here, the constraints essentially require that the singular values σ_n of $\hat{\mathbf{P}}$ be between zero and one and average at least z. In the case that $\hat{\mathbf{Q}}_E$ and $\hat{\mathbf{P}}$ share the same singular vectors, this essentially leads to a linear program (LP) over the vector σ containing these singular values:

$$\begin{aligned} \text{minimize} \quad & \langle \hat{\sigma}_E, \sigma \rangle \\ \text{subject to} \quad & 0 \preceq \sigma \preceq 1 \\ & \tfrac{1}{N_0} \textstyle\sum_{n=1}^{N_0} \sigma_n \geq z. \end{aligned} \tag{236}$$

When $z = 1 - \frac{n_r}{N_0}$ for some positive integer $n_r \leq N_0$, the minimizing σ has n_r leading zeros followed by $N - n_r$ ones, and the corresponding $\mathbf{P}$ follows the simple principal-component construction exactly. In the general case when z falls in between two such values, the resulting matrix $\mathbf{P}$ is a convex combination of the two corresponding projection matrices, so that it has one of its eigenvalues between zero and one. This parameter z defines a tradeoff between the aggressiveness of the nulling in the ellipse E and the ability to preserve most of the energy of the other locations.

In spite of the powerful existing software packages to solve the SDP, however, the work presented still opts for the approximate solution via the simpler version of SVD thresholding that is usually is more than a hundred times faster, especially since this matrix will need to be constructed repeatedly for each source location. Indeed, even this relatively fast method of SVD thresholding starts to become computationally

prohibitive when both dimensions of the matrix start to approach a thousand. In these cases however, this bottleneck can be reduced by an order of magnitude using a randomized SVD algorithm to construct an approximate projection matrix [41]. Specifically, the left singular vectors of $\mathbf{Q}_E\mathbf{X}$ are used for correlation matrix $\mathbf{Q}_E$ and i.i.d. Gaussian matrix $\mathbf{X} \in \mathbb{C}^{N\times N_x}$ $(N_x < N)$ to construct the projection matrix. Because each column of $\mathbf{X}$ is unitary invariant (in probability), this product can be expressed as $\sum_{n=1}^{N} g_n\sigma_n U_n$ where g_n is an i.i.d. Gaussian sequence (standard normal) and U_n and σ_n are the N singular vectors and values of $\mathbf{Q}_E$. In particular, the range of $\mathbf{Q}_E\mathbf{X}$ is the range of $\mathbf{Q}_E$ almost surely whenever N_x is taken to be at least the rank of $\mathbf{Q}_E$, so that projection matrices may be constructed appropriately.

At first glance, it appears that the attenuation constraint over the region of interest $\mathcal{R}$ should have been defined without the ellipse $\mathcal{R} \setminus E$. However, in light of the objective function, the two are functionally equivalent (for an appropriate modification of the z parameter). Additionally, there can be a computational convenience to defining the attenuation over the entire region of interest $\mathcal{R}$ instead of the intended "passband" $\mathcal{R} \setminus E$ (using language analogous to the design of a notch filter).

A.4 Necessary and Efficient Conditions

This section describes deterministic necessary conditions that may be used to quickly and efficiently rule out infeasible pairs of locations during the 2-source exhaustive search as in Eq. (51).

The goal is to search through the normalized columns of a matrix $\mathcal{A} \in \mathbb{C}^{M\times N}$ for the best pair of columns A_{n_1}, A_{n_2} of $\mathcal{A}$ that accounts for Y. (The N in this particular case is the number of potential source locations in the grid.) That is, to search for the pair that minimizes:

$$\min_{n_1,n_2,\beta_1,\beta_2} \|Y - \beta_1 A_{n_1} - \beta_2 A_{n_2}\|^2. \tag{237}$$

In this case, it may be feasible to solve such a system by maximizing the norm of the

pseudo-inverse of each pair of columns applied to the data vector Y:

$$\max_{n_1,n_2} \| \begin{bmatrix} A_{n_1} & A_{n_2} \end{bmatrix}^{\dagger} Y \|^2 \triangleq \max_{n' \leq \binom{N}{2}} \| \mathbf{A}_{n'}^{\dagger} Y \|^2, \tag{238}$$

where $\mathbf{A}_{n'}$ denotes a particular pair of columns (n_1, n_2), and where $\mathbf{A}_{n'}^{\dagger} = (\mathbf{A}_{n'}^H \mathbf{A}_{n'})^{-1} \mathbf{A}_{n'}^H$.

This exhaustive search would require $\binom{N}{2}$ pseudo-inverses for a total of $O(MN^2)$ computations. However, even this can be computationally prohibitive when N is large. The adjoint operation $\mathcal{A}^H Y$ may be computed with only $O(MN)$ computations (as may the normalization of the columns of $\mathcal{A}$), but the bottleneck lies in computing the correlation terms $\langle A_{n_1}, A_{n_2} \rangle$.

This bottleneck on the computation of $\langle A_{n_1}, A_{n_2} \rangle$ motivates the construction of a useful heuristic rule on the (easily computed) elements of $\mathcal{A}^H Y$. The main idea is that for any $\mathcal{A} \in \mathbb{C}^{M \times 2}$ satisfying $|\langle A_1, A_2 \rangle| \leq \gamma$ (for unit norm columns $\|A_1\| = \|A_2\| = 1$), and for any $x \in \mathbb{C}^2$:

$$\frac{\|\mathcal{A}^H \mathcal{A} x\|_p}{\|\mathcal{A} x\|_2} \geq \sqrt{1 - \gamma}, \tag{239}$$

where $p = \frac{2}{1+\gamma}$. (As before, all norms are Euclidean 2-norms by default unless stated otherwise.) This property is expressed more generally in the following lemma, which will be proved at the end of the Appendix.

Lemma 17 *Let $Y = \mathcal{A}x + \eta$, where $\mathcal{A} \in \mathbb{C}^{N \times 2}$, $\mathcal{A}^H \eta = 0$, $\|A_1\| = \|A_2\| = 1$. Then:*

$$\|\mathcal{A}^H Y\|_p \geq \sqrt{1 - \gamma} \|\mathcal{A} x\|, \tag{240}$$

where $\gamma = |\langle A_1, A_2 \rangle|$, $p = \frac{2}{1+\gamma}$.

Using this fact, one may efficiently eliminate candidate pairs (n_1, n_2) of Green's vectors (corresponding to candidate locations) from consideration of the search carried out in Eq. (238) in the following way. First, one chooses a value of γ (e.g., 1/2) and determine $p = 2/(1 + \gamma)$ accordingly. Then, one chooses a parameter L (discussed more later) in an attempt to estimate the $\|\mathcal{A}x\|$ term in Lemma 17. Then, after

normalizing the columns of $\mathcal{A}$ and Y (where $\mathcal{A}$ is a matrix containing the candidate Green's vectors), one computes the N values $Z_n = |\langle Y, A_n \rangle|^p$, and eliminate all pairs (n_1, n_2) from consideration if $Z_{n_1} + Z_{n_2} \leq (L\sqrt{1-\gamma})^p$, provided that this pair has correlation at most γ:

$$|\langle A_{n_1}, A_{n_2} \rangle| \leq \gamma, \tag{241}$$

This condition can be shown, at least empirically for the case of interest, in Fig. 24, where sources that are sufficiently far away from one another are guaranteed to have a correlation below some threshold γ. In particular $|\langle G(\vec{r}_1), G(\vec{r}_2) \rangle| \leq 1 - \epsilon_-(\delta)$ for all $d(\vec{r}_1, \vec{r}_2) \geq \delta$. The thresholding of the sum $Z_{n_1} + Z_{n_2}$ over all N pairs may be efficiently done by pre-sorting the Z_n in descending order. Then, Eq. (238) is maximized over all candidate pairs that remain, yielding optimum (n_1^*, n_2^*) with corresponding source amplitudes β_1^*, β_2^*.

Given this pair, it only remains to verify the assumptions of Lemma 17 to ensure that viable pairs were not accidentally eliminated. By virtue of the least squares solution, the orthogonality assumption is satisfied $\mathcal{A}_{n'}^H (Y - \mathcal{A}_{n'} \beta^*) = 0$ where $\mathcal{A}_{n'} = [A_{n_1^*} \quad A_{n_2^*}]$. It only remains to verify that $\|\mathcal{A}_{n'} \beta^*\| \geq L$. If this is not the case, then one simply reduces L and repeat the elimination procedure from the beginning. One sensible initial estimate for L is $\sqrt{1-2\sigma^2}$, for noise to signal ratio $\sigma^2 = \mathrm{E}\left[\|\eta\|^2\right] \|\alpha_1 G(\vec{r}_1) + \alpha_2 G(\vec{r}_2)\|^{-2}$, using the notation of Eq. (60).

Apparently, as evidenced by the relatively small $\epsilon_-(\delta)$ function in the single-frequency case as illustrated in Fig. 24, this approach is more effective on the coherent-case where the distance between pairs of Green's (replica) vectors are more closely tied to physical distance. In such cases with K frequencies, M randomized back-propagations, and N candidate source locations, this approach can potentially reduce an $O(MKN^2)$ operation to an $O(N^2)$ operation. For purposes of the results presented, it transformed a simulation that would have required weeks to complete into one that could be completed in less than a day.

Proof First, note that $\|\mathcal{A}^H Y\|_p = \|\mathcal{A}^H \mathcal{A} x\|_p$, so that it only remains to show that $\|\mathcal{A}^H \mathcal{A} x\|_p \geq \sqrt{1-\gamma}\|\mathcal{A} x\|$. Note that $\mathcal{A}^H \mathcal{A}$ is of the form:

$$\mathcal{A}^H \mathcal{A} = \begin{bmatrix} 1 & \gamma e^{j\theta_0} \\ \gamma e^{-j\theta_0} & 1 \end{bmatrix}, \tag{242}$$

and that x is of the form:

$$x = \begin{bmatrix} x_1 e^{j\theta_1} \\ x_2 e^{j\theta_2} \end{bmatrix}, \tag{243}$$

where $x_1, x_2 \geq 0$ and $\theta_0, \theta_1, \theta_2 \in [-\pi,\ \pi)$. Using this form, and using the fact that p-norms are invariant to element-wise phase shifts, gives:

$$\|\mathcal{A}^H \mathcal{A} x\|_p = \left\| \begin{bmatrix} 1 & \gamma e^{j\theta} \\ \gamma e^{-j\theta} & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \right\|_p = C\|\mathbf{B}_\theta x_z\|_p, \tag{244}$$

where $C = (x_1 + x_2)/2$, $\theta = \theta_0 + \theta_2 - \theta_1$, and:

$$\mathbf{B}_\theta = \begin{bmatrix} 1 & \gamma e^{j\theta} \\ \gamma e^{-j\theta} & 1 \end{bmatrix}, \quad z = \frac{x_1 - x_2}{x_1 + x_2}, \quad x_z = \begin{bmatrix} 1+z \\ 1-z \end{bmatrix}, \tag{245}$$

so that $[x_1 \ \ x_2]^T = C x_z$. For similar reasons,

$$\|\mathcal{A} x\|_2 = C\sqrt{x_z^T B_\theta x_z}. \tag{246}$$

For the last part of the proof, it will be shown that

$$\begin{aligned} \frac{\|\mathcal{A}^H \mathcal{A} x\|_p^p}{\|\mathcal{A} x\|_2^p} &= \left\| \frac{\mathbf{B}_\theta x_z}{\sqrt{x_z^T B_\theta x_z}} \right\|_p^p & (247) \\ &\geq \left\| \frac{\mathbf{B}_\pi x_z}{\sqrt{x_z^T B_\pi x_z}} \right\|_p^p & (248) \\ &\geq \left\| \frac{\mathbf{B}_\pi x_0}{\sqrt{x_0^T B_\pi x_0}} \right\|_p^p & (249) \\ &= 2\left(\frac{1-\gamma}{2}\right)^{p/2} \geq (1-\gamma)^{p/2}. & (250) \end{aligned}$$

It has already been established that the first equality is true for some z and θ (whose values depend on x and $\mathcal{A}$). The two following inequalities show that the worst case (smallest) value occurs when $\theta = \pi$ and $z = 0$ so that these values give a lower bound on the first expression.

To show that $\theta = \pi$ gives a lower bound for all $|z| \leq 1$, note that the magnitude-squared of the elements within the norm in Eq. (247) are:

$$1 - \frac{(1-z)^2(1-\gamma^2)}{(1+z)^2 + (1-z)^2 - 2\gamma(1+z)(1-z)\cos(\theta)} \tag{251}$$

$$1 - \frac{(1+z)^2(1-\gamma^2)}{(1+z)^2 + (1-z)^2 - 2\gamma(1+z)(1-z)\cos(\theta)}, \tag{252}$$

and both reach their minimum value at $\theta = \pi$, thereby establishing Eq. (248).

To show Eq. (250), the following function is defined as:

$$f(z) = \left\| \frac{\mathbf{B}_\pi x_z}{\sqrt{x_z^T B_\pi x_z}} \right\|_p^p, \tag{253}$$

so that $f(z) \geq f(0)$ will be shown by showing that $f^{'}(z) \geq 0$ for all $0 < z < 1$ and noting that f is an even function (by the definition of x_z).

First, an equivalent form on this domain, using the substitution $q = p - 1 = \frac{1-\gamma}{1+\gamma}$:

$$f(z) = \frac{((1-\gamma) + z(1+\gamma))^p + |(1-\gamma) - z(1+\gamma)|^p}{(2(1-\gamma) + 2z^2(1+\gamma))^{p/2}} \tag{254}$$

$$= (1+\gamma)^{p/2} \frac{(q+z)^p + |q-z|^p}{(2q+2z^2)^{p/2}}. \tag{255}$$

The term within the absolute value reaches its cusp at $z = q$. On the interval $q < z < 1$, both terms in the numerator are increasing faster than the denominator. On the interval $0 < z < q$, the first derivative can be expressed as:

$$f^{'}(z) = \frac{(1+\gamma)^{p/2} p}{(2q+2z^2)^{p/2+1}} \left((q+z)^q (2q + 2z^2 - 2z(q+z)) - (q-z)^q (2q + 2z^2 + 2z(q-z)) \right),$$

which can be shown positive on this domain by following this string of inequalities

(using the dummy variable $0 < w < 1$):

$$\begin{aligned}
(q &< 1) && (256)\\
\frac{q^2}{q^2-w^2} &> \frac{1}{1-w^2} && (257)\\
\int_0^z q\frac{-2q}{q^2-w^2}dw &< \int_0^z \frac{-2}{1-w^2}dw && (258)\\
q\log\left(\frac{q-z}{q+z}\right) &< \log\left(\frac{1-z}{1+z}\right), && (259)\\
\left(\frac{q-z}{q+z}\right)^q &< \frac{2q+2z^2-2z(q+z)}{2q+2z^2+2z(q-z)} && (260)\\
(q+z)^q(2q+2z^2-2z(q+z)) &> (q-z)^q(2q+2z^2+2z(q-z)), && (261)
\end{aligned}$$

as desired. ■

A.5 Constants

For reference, the following specific constants are used in this thesis

$$\begin{aligned}
C_0 &= 250 && (262)\\
C_1 &= 51 && (263)\\
C_2 &= 4\sqrt{2} && (264)\\
C_B &= 4 && (265)
\end{aligned}$$

These are understood to be upper bounds for these constants, and are expected to be at least somewhat loose. The following proofs of Proposition 2 and Lemma 10 support these particular numerical constants.

Proof (of Proposition 2)

This proof follows the more general proof given in [87], which mirrors the derivation of classical Bernstein's inequality similarly to other works (e.g., [88, 89, 90]), and is repeated here for convenience.

We wish to establish:

$$\mathrm{P}\left\{\left\|\sum_m X_m\right\| \geq C_B \max\left\{\sigma\sqrt{t}, 2Bt\log\left(\frac{2B\sqrt{M}}{\sigma}\right)\right\}\right\} \leq 2Ke^{-t}, \qquad (266)$$

where $\|X_m\|_{\Psi_1} \leq B$, and

$$\sigma^2 := \left\| \sum_m \mathrm{E}\left[X_m^2\right] \right\|. \tag{267}$$

Let $Y_M := X_1 + \cdots + X_M$. Note that $\|Y_M\| < t$ if and only if $-t\mathbf{I} < Y_M < t\mathbf{I}$. Therefore,

$$\mathrm{P}\left\{\|Y_M\| \geq t\right\} = \mathrm{P}\left\{Y_M \not\leq t\mathbf{I}\right\} + \mathrm{P}\left\{Y_M \not\geq -t\mathbf{I}\right\}. \tag{268}$$

The following bounds are straightforward by simple matrix algebra:

$$\mathrm{P}\left\{Y_M \not\leq t\mathbf{I}\right\} = \mathrm{P}\left\{e^{\lambda Y_M} \not\leq e^{\lambda t\mathbf{I}}\right\} \leq \mathrm{P}\left\{\mathrm{tr}\Big(e^{\lambda Y_M}\Big) \geq e^{\lambda t}\right\} \leq e^{-\lambda t}\,\mathrm{E}\left[\mathrm{tr}(e^{\lambda Y_M})\right]. \tag{269}$$

To bound the expected value in the right hand side, we use the well-known *Golden-Thompson inequality* (see, e.g., [94, pg. 94]):

$$\mathrm{tr}(e^{A+B}) \leq \mathrm{tr}(e^A e^B). \tag{270}$$

and the independence of random variables $X_1, \ldots, X_M$, yielding:

$$\mathrm{E}\left[\mathrm{tr}(e^{\lambda Y_M})\right] = \mathrm{E}\left[\mathrm{tr}\Big(e^{\lambda Y_{M-1} + \lambda X_M}\Big)\right] \leq \mathrm{E}\left[\mathrm{tr}\Big(e^{\lambda Y_{M-1}} e^{\lambda X_M}\Big)\right] = \mathrm{tr}\Big(\mathrm{E}\left[e^{\lambda Y_{M-1}} e^{\lambda X_M}\right]\Big) =$$

$$\mathrm{tr}\Big(\mathrm{E}\left[e^{\lambda Y_{M-1}}\right]\mathrm{E}\left[e^{\lambda X_M}\right]\Big) \leq \mathrm{E}\left[\mathrm{tr}\Big(e^{\lambda Y_{M-1}}\Big)\right]\Big\|\mathrm{E}\left[e^{\lambda X_M}\right]\Big\|.$$

By induction, we conclude that

$$\mathrm{E}\left[\mathrm{tr}(e^{\lambda Y_M})\right] \leq \mathrm{E}\left[\mathrm{tr}\Big(e^{\lambda X_1}\Big)\right]\Big\|\mathrm{E}\left[e^{\lambda X_2}\right]\Big\| \ldots \Big\|\mathrm{E}\left[e^{\lambda X_M}\right]\Big\|.$$

Since $\mathrm{E}\left[\mathrm{tr}\Big(e^{\lambda X_1}\Big)\right] = \mathrm{tr}\Big(\mathrm{E}\left[e^{\lambda X_1}\right]\Big) \leq K\Big\|\mathrm{E}\left[e^{\lambda X}\right]\Big\|$, we get

$$\mathrm{E}\left[\mathrm{tr}(e^{\lambda Y_M})\right] \leq K\Big\|\mathrm{E}\left[e^{\lambda X}\right]\Big\|^M. \tag{271}$$

It remains to bound the norm $\|\,\mathrm{E}\left[e^{\lambda X}\right]\,\|$. To this end, we use a Taylor expansion and the condition $\mathrm{E}\left[X\right] = 0$ to get

$$\mathrm{E}\left[e^{\lambda X}\right] = \mathbf{I} + \mathrm{E}\left[\lambda^2 X^2\left[\frac{1}{2!} + \frac{\lambda X}{3!} + \frac{\lambda^2 X^2}{4!} + \ldots\right]\right] \leq$$

$$\mathbf{I} + \lambda^2\,\mathrm{E}\left[X^2\left[\frac{1}{2!} + \frac{\lambda\|X\|}{3!} + \frac{\lambda^2\|X\|^2}{4!} + \ldots\right]\right] = \mathbf{I} + \lambda^2\,\mathrm{E}\left[X^2\left[\frac{e^{\lambda\|X\|} - 1 - \lambda\|X\|}{\lambda^2\|X\|^2}\right]\right].$$

Therefore, for all $\tau > 0$,

$$\left\| \mathrm{E}\left[e^{\lambda X}\right] \right\| \leq 1 + \lambda^2 \left\| \mathrm{E}\left[X^2 \left[\frac{e^{\lambda \|X\|} - 1 - \lambda \|X\|}{\lambda^2 \|X\|^2} \right] \right] \right\| \leq$$

$$1 + \lambda^2 \left\| \mathrm{E}\left[X^2\right] \right\| \left[\frac{e^{\lambda\tau} - 1 - \lambda\tau}{\lambda^2\tau^2} \right] + \lambda^2 \, \mathrm{E}\left[\|X\|^2 \left[\frac{e^{\lambda\|X\|} - 1 - \lambda\|X\|}{\lambda^2\|X\|^2} \right] I(\|X\| \geq \tau) \right].$$

Let $\tau := 2\log(\frac{4}{\sigma^2})$ and suppose that $\lambda \leq \tau^{-1} \leq 1/2$. Suppose also for now that $\|X\|_{\Psi_1} \leq 1$ so that $\mathrm{E}\left[e^{\|X\|}\right] \leq 2$. Then

$$\begin{aligned} \mathrm{E}\left[\|X\|^2 \left(\frac{e^{\lambda\|X\|} - 1 - \lambda\|X\|}{\lambda^2\|X\|^2} \right) I(\|X\| \geq \tau) \right] &\leq 4 \, \mathrm{E}\left[e^{\|X\|/2} I(\|X\| \geq \tau) \right] && (272) \\ &\leq 4\sqrt{\mathrm{E}\left[e^{\|X\|}\right] \mathrm{P}\left\{\|X\| \geq \tau\right\}} && (273) \\ &\leq 2\sigma^2, && (274) \end{aligned}$$

by Lemma 6 and Cauchy-Schwarz. As a result, we get the following bound:

$$\left\| \mathrm{E}\left[e^{\lambda X}\right] \right\| \leq 1 + \frac{\lambda^2\sigma^2}{M} \left[\frac{e^{\lambda\tau} - 1 - \lambda\tau}{\lambda^2\tau^2} \right] + \frac{2\lambda^2\sigma^2}{M} \leq 1 + \frac{e\lambda^2\sigma^2}{M} \tag{275}$$

Thus, for all λ satisfying the condition

$$\lambda \leq \frac{1}{2\log(\frac{4M}{\sigma^2})} \tag{276}$$

we have $\| \mathrm{E}\left[e^{\lambda X}\right] \| \leq \exp(e\lambda^2\sigma^2/M)$. This can be combined with Eqs. (268), (269), (271) to get

$$\mathrm{P}\left\{\|Y_M\| \geq t\right\} \leq 2K \exp(-\lambda t + e\lambda^2\sigma^2). \tag{277}$$

It remains now to minimize the last bound with respect to all λ satisfying Eq. (276) to get

$$\lambda_{\text{optimal}} = \min\left(\frac{1}{2\log(4M/\sigma^2)}, \frac{t}{2e\sigma^2} \right). \tag{278}$$

yielding

$$\mathrm{P}\left\{\|Y_M\| \geq t\right\} \leq 2K \exp\left(-\min\left(\frac{t^2}{4e\sigma^2}, \frac{t}{4\log(4M/\sigma^2)} \right) \right) \tag{279}$$

so that for $C_B = 4$ and $\|X_m\|_{\Psi_1} \leq 1$ we have

$$\mathrm{P}\left\{ \|Y_M\| \geq C_B \max(\sigma\sqrt{t}, \; 2t\log(2\sqrt{M}/\sigma)) \right\} \leq 2Ke^{-t}. \tag{280}$$

In the general case when $\|X_m\|_{\Psi_1} \leq B$, we simply substitute Y_M/B for Y_M and σ/B for σ, utilizing the homogenity of the σ parameter and B parameter with respect to scalar multiplication of the random matrices, yielding:

$$\mathrm{P}\left\{\|Y_M\| \geq C_B \max(\sigma\sqrt{t},\ 2Bt\log(2B\sqrt{M}/\sigma))\right\} \leq 2Ke^{-t}. \tag{281}$$

which immediately implies Eq.(266). ■

Proof (of Lemma 10)

$$\mathrm{P}\left\{\frac{\langle \mathbf{\Phi} f, \mathbf{\Phi} g\rangle - \langle f, g\rangle}{\|f\|\|g\|} \geq C_1\frac{u}{M} + C_2\sqrt{\frac{u}{M}}\right\} \leq 2e^{-u} \tag{282}$$

Note that it suffices to prove the case when $\|f\| = \|g\| = 1$. It will be useful to decompose g as:

$$g = \alpha f + \beta g^{\perp} \tag{283}$$

where $\alpha = \langle f, g\rangle$ and $\alpha^2 + \beta^2 = \|g^{\perp}\| = 1$, and write the quantity of interest as the sum of M i.i.d. copies of the random scalar:

$$X = \frac{\langle \phi f, \phi g\rangle - \alpha}{M} = \frac{\alpha(\|\phi f\|^2 - 1) + \beta\langle \phi f, \phi g^{\perp}\rangle}{M}, \tag{284}$$

where ϕ is an i.i.d. Gaussian row vector with zero mean and unit variance. Then:

$$\sigma^2 \leq 2/M \tag{285}$$

and

$$M\|X\|_{\Psi_1} \leq \frac{8}{3}(\alpha + \beta) \leq 8\sqrt{2}/3, \tag{286}$$

using Lemmas 8 and 9. Applying Matrix Bernstein via Proposition 2 gives

$$\mathrm{P}\left\{\frac{\langle \mathbf{\Phi} f, \mathbf{\Phi} g\rangle - \langle f, g\rangle}{\|f\|\|g\|} \geq C_B\left(\frac{16\sqrt{2}u}{3M}\log(16/3) + \sqrt{\frac{2u}{M}}\right)\right\} \leq 2e^{-u} \tag{287}$$

as desired.

■

Bibliography

[1] William Mantzel. "Distributed Alternating Localization-Triangulation of Camera Networks". MA thesis. Rice University, 2005.

[2] J. Polastre et al. "Analysis of wireless sensor networks for habitat monitoring". In: *Wireless sensor networks* (2004), pp. 399–423.

[3] A.S. Bandeira et al. "Certifying the restricted isometry property is hard". In: *arXiv preprint arXiv:1204.1580* (2012).

[4] R. G. Baraniuk et al. "A simple proof of the restricted isometry property for random matrices". In: *Constructive Approximation* 28.3 (2008), pp. 253–263.

[5] G.B. Dantzig, A. Orden, and P. Wolfe. "The generalized simplex method for minimizing a linear form under linear inequality restraints". In: *Pacific Journal of Mathematics* 5.2 (1955), pp. 183–195.

[6] E.J. Candès. "The restricted isometry property and its implications for compressed sensing". In: *Comptes Rendus Mathematique* 346.9 (2008), pp. 589–592.

[7] S.P. Boyd and L. Vandenberghe. *Convex optimization (pg. 167–174)*. Cambridge University Press, 2004, pp. 167–174. ISBN: 0521833787.

[8] R. Tibshirani. "Regression shrinkage and selection via the lasso". In: *Journal of the Royal Statistical Society. Series B (Methodological)* (1996), pp. 267–288.

[9] M. Lustig, D. Donoho, and J.M. Pauly. "Sparse MRI: The application of compressed sensing for rapid MR imaging". In: *Magnetic Resonance in Medicine* 58.6 (2007), pp. 1182–1195.

[10] I.F. Gorodnitsky, J.S. George, and B.D. Rao. "Neuromagnetic source imaging with FOCUSS: a recursive weighted minimum norm algorithm". In: *Electroencephalography and clinical Neurophysiology* 95.4 (1995), pp. 231–251.

[11] H. Jung et al. “k-t FOCUSS: A general compressed sensing framework for high resolution dynamic MRI”. In: *Magnetic Resonance in Medicine* 61.1 (2008), pp. 103–116.

[12] H. Yu and G. Wang. “Compressed sensing based interior tomography”. In: *Physics in medicine and biology* 54.9 (2009), p. 2791.

[13] J. Provost and F. Lesage. “The application of compressed sensing for photo-acoustic tomography”. In: *Medical Imaging, IEEE Transactions on* 28.4 (2009), pp. 585–594.

[14] E.J. Candes and T. Tao. “Decoding by linear programming”. In: *Information Theory, IEEE Transactions on* 51.12 (2005), pp. 4203–4215.

[15] Salman Asif, William Mantzel, and Justin Romberg. “Random Channel Coding and Blind Deconvolution”. In: *Allerton Conference*. Monticello, IL, 2009.

[16] Salman Asif, William Mantzel, and Justin Romberg. “Channel protection: Random coding meets sparse channels”. In: *Information Theory Workshop*. Taormina, Italy, 2009.

[17] W.U. Bajwa et al. “Compressed channel sensing: A new approach to estimating sparse multipath channels”. In: *Proceedings of the IEEE* 98.6 (2010), pp. 1058–1076.

[18] C.R. Berger et al. “Sparse channel estimation for multicarrier underwater acoustic communication: From subspace methods to compressed sensing”. In: *Signal Processing, IEEE Transactions on* 58.3 (2010), pp. 1708–1721.

[19] M.F. Duarte et al. “Single-pixel imaging via compressive sampling”. In: *Signal Processing Magazine, IEEE* 25.2 (2008), pp. 83–91.

[20] A. Levin et al. “Image and depth from a conventional camera with a coded aperture”. In: *ACM Transactions on Graphics (TOG)* 26.3 (2007), p. 70.

[21] A. Veeraraghavan et al. “Dappled photography: Mask enhanced cameras for heterodyned light fields and coded aperture refocusing”. In: *ACM Transactions on Graphics* 26.3 (2007), p. 69.

[22] R. Raskar, A. Agrawal, and J. Tumblin. “Coded exposure photography: motion deblurring using fluttered shutter”. In: *ACM Transactions on Graphics* 25.3 (2006), p. 795.

[23] William Mantzel, Justin Romberg, and Karim Sabra. “Compressive Matched-Field Processing”. In: *The Journal of the Acoustical Society of America* 132 (1 2012), pp. 90–102.

[24] Y.C. Pati, R. Rezaiifar, and PS Krishnaprasad. “Orthogonal matching pursuit: Recursive function approximation with applications to wavelet decomposition”. In: *Signals, Systems and Computers, 1993. 1993 Conference Record of The Twenty-Seventh Asilomar Conference on*. IEEE. 1993, pp. 40–44.

[25] M. Talagrand. “The generic chaining”. In: (2005).

[26] A. Tolstoy. “Applications of matched-field processing to inverse problems in underwater acoustics”. In: *Inverse Problems* 16 (2000), p. 1655.

[27] A.B. Baggeroer, W.A. Kuperman, and P.N. Mikhalevsky. “An overview of matched field methods in ocean acoustics”. In: *IEEE Journal of Oceanic Engineering* 18.4 (1993), pp. 401–424.

[28] AB Baggeroer, WA Kuperman, and H. Schmidt. “Matched field processing: Source localization in correlated noise as an optimum parameter estimation problem”. In: *The Journal of the Acoustical Society of America* 83 (1988), p. 571.

[29] F.K. Gruber, E.A. Marengo, and A.J. Devaney. “Time-reversal imaging with multiple signal classification considering multiple scattering between the targets”. In: *The Journal of the Acoustical Society of America* 115 (2004), p. 3042.

[30] E. Candès, J. Romberg, and T. Tao. "Stable signal recovery from incomplete and inaccurate measurements". In: *Communications on Pure and Applied Math.* 59.8 (2006), pp. 1207–1223.

[31] E. Candès and T. Tao. "Near-optimal signal recovery from random projections: Universal encoding strategies?" In: *IEEE Transactions on Information Theory* 52.12 (2006), pp. 5406–5245.

[32] D.L. Donoho. "Compressed sensing". In: *IEEE Transactions on Information Theory* 52.4 (2006), pp. 1289–1306.

[33] E. Candès and M. Wakin. "An introduction to compressive sampling". In: *IEEE Signal Processing Magazine* 25.2 (2008), pp. 21–30.

[34] J. Romberg. "Imaging via compressive sampling". In: *IEEE Signal Processing Magazine* (2008), pp. 14–20.

[35] D. Healy and D. J. Brady. "Compression at the physical interface". In: *IEEE Signal Processing Magazine* 25.2 (2008), pp. 67–71.

[36] M.A. Davenport et al. "Signal processing with compressive measurements". In: *Selected Topics in Signal Processing, IEEE Journal of* 4.2 (2010), pp. 445–460. ISSN: 1932-4553.

[37] M.A. Davenport et al. "The smashed filter for compressive classification and target recognition". In: *Computational Imaging V at SPIE Electronic Imaging* (2007).

[38] M.B. Wakin. "Manifold-based signal recovery and parameter estimation from compressive measurements". In: *Arxiv preprint arXiv:1002.1247* (2010).

[39] Lawrence Carin, Dehong Liu, and Bin Guo. "In situ compressive sensing". In: *Inverse Problems* 24.1 (2008), pp. 15–23.

[40] EA Marengo et al. "Compressive sensing for inverse scattering". In: *XXIX URSI General Assembly, Chicago, Illinois* (2008).

[41] N. Halko, P.G. Martinsson, and J.A. Tropp. "Finding structure with randomness: Stochastic algorithms for constructing approximate matrix decompositions". In: *SIAM Review, Survey and Review section* 53.2 (2011), pp. 217–288.

[42] S. Chaillat and G. Biros. *FalMS: A fast algorithm for the inverse medium problem with multiple frequencies and multiple sources for the scalar Helmholtz equation*. 2010.

[43] A. Fannjiang, P. Yan, and T. Strohmer. "Compressed remote sensing of sparse objects". In: *SIAM Journal on Imaging Sciences* 3.3 (2009), pp. 595–618.

[44] A.C. Gurbuz, J.H. McClellan, and V. Cevher. "A compressive beamforming method". In: *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)* (2008).

[45] V. Cevher, M.F. Duarte, and R.G. Baraniuk. "Distributed target localization via spatial sparsity". In: *European Signal Processing Conference (EUSIPCO)* (2008).

[46] Finn B. Jensen et al. *Computational Ocean Acoustics*. Ed. by R. T. Beyer. AIP Press, 2000.

[47] W. B. Johnson and J. Lindenstrauss. "Extensions of Lipschitz mappings into a Hilbert space". In: *Conference in Modern Analysis and Probability* (1984), pp. 189–206.

[48] S. Dasgupta and A. Gupta. "An elementary proof of a theorem of Johnson and Lindenstrauss". In: *Random Structures and Algorithms* 22.1 (2003), pp. 60–65.

[49] D. Achlioptas. "Database-friendly random projections: Johnson-Lindenstrauss with binary coins". In: *Journal of Computer and System Science* 66 (2003), pp. 671–687.

[50] E.J. Candès, J.K. Romberg, and T. Tao. "Stable signal recovery from incomplete and inaccurate measurements". In: *Communications on Pure and Applied Mathematics* 59.8 (2006), p. 1207.

[51] R. G. Baraniuk and M. B. Wakin. "Random projections of smooth manifolds". In: *Foundations of Computational Mathematics* 9.1 (2009), pp. 51–77.

[52] HC Song, J. De Rosny, and WA Kuperman. "Improvement in matched field processing using the CLEAN algorithm". In: *The Journal of the Acoustical Society of America* 113 (2003), p. 1379.

[53] JA Högbom. "Aperture synthesis with a non-regular distribution of interferometer baselines". In: *Astronomy and Astrophysics Supplement Series* 15 (1974), p. 417.

[54] A.N. Mirkin and L.H. Sibul. "Maximum likelihood estimation of the locations of multiple sources in an acoustic waveguide". In: *The Journal of the Acoustical Society of America* 95 (1994), p. 877.

[55] Z.H. Michalopoulou. "Multiple source localization using a maximum a posteriori Gibbs sampling approach". In: *The Journal of the Acoustical Society of America* 120 (2006), p. 2627.

[56] E.K. Westwood. "Broadband matched-field source localization". In: *The Journal of the Acoustical Society of America* 91.5 (1992), pp. 2777–2789.

[57] T.B. Neilsen. "Localization of multiple acoustic sources in the shallow ocean". In: *The Journal of the Acoustical Society of America* 118 (2005), p. 2944.

[58] K. Kim, W. Seong, and K. Lee. "Adaptive Surface Interference Suppression for Matched-Mode Source Localization". In: *Oceanic Engineering, IEEE Journal of* 35.1 (2010), pp. 120–130.

[59] G.H. Golub and C.F. Van Loan. *Matrix computations (pg. 69–75)*. Vol. 3. Johns Hopkins University Press, 1996, pp. 69–75. ISBN: 9780801830105.

[60] S.G. Mallat and Z. Zhang. "Matching pursuits with time-frequency dictionaries". In: *Signal Processing, IEEE Transactions on* 41.12 (1993), pp. 3397–3415.

[61] J.A. Tropp and A.C. Gilbert. "Signal recovery from random measurements via orthogonal matching pursuit". In: *Information Theory, IEEE Transactions on* 53.12 (2007), pp. 4655–4666.

[62] D. Needell and J.A. Tropp. "CoSaMP: Iterative signal recovery from incomplete and inaccurate samples". In: *Applied and Computational Harmonic Analysis* 26.3 (2009), pp. 301–321.

[63] M.A. Davenport and M.B. Wakin. "Compressive Sensing of Analog Signals Using Discrete Prolate Spheroidal Sequences". In: *Applied and Computational Harmonic Analysis* (2012).

[64] R.J. Vaccaro, A. Chhetri, and B.F. Harrison. "Matrix filter design for passive sonar interference suppression". In: *The Journal of the Acoustical Society of America* 115 (2004), p. 3010.

[65] H.W. Kuhn. "The Hungarian method for the assignment problem". In: *Naval research logistics quarterly* 2.1-2 (1955), pp. 83–97.

[66] P.A. Absil, A. Edelman, and P. Koev. "On the largest principal angle between random subspaces". In: *Linear algebra and its applications* 414.1 (2006), pp. 288–294.

[67] M. Ledoux. *The concentration of measure phenomenon*. Vol. 89. Amer Mathematical Society, 2001.

[68] Benjamin Recht. "A simpler approach to matrix completion". In: *arXiv preprint arXiv:0910.0651* (2009).

[69] Jae Young Park et al. "Concentration of measure for block diagonal matrices with applications to compressive signal processing". In: *Signal Processing, IEEE Transactions on* 59.12 (2011), pp. 5859–5875.

[70] Dimitris Achlioptas. "Database-friendly random projections". In: *Proceedings of the twentieth ACM SIGMOD-SIGACT-SIGART symposium on Principles of database systems*. ACM. 2001, pp. 274–281.

[71] K.L. Clarkson. "Tighter bounds for random projections of manifolds". In: *Proceedings of the twenty-fourth annual symposium on Computational geometry*. ACM. 2008, pp. 39–48.

[72] H.L. Yap, M.B. Wakin, and C.J. Rozell. "Stable manifold embeddings with operators satisfying the restricted isometry property". In: *Information Sciences and Systems (CISS), 2011 45th Annual Conference on*. IEEE. 2011, pp. 1–6.

[73] S. Mendelson, A. Pajor, and N. Tomczak-Jaegermann. "Uniform uncertainty principle for Bernoulli and subgaussian ensembles". In: *Constructive Approximation* 28.3 (2008), pp. 277–289.

[74] J. Yoo et al. "A compressed sensing parameter extraction platform for radar pulse signal acquisition". In: *IEEE Journal on Emerging and Selected Topics in Circuits and Systems* 2.3 (2012), pp. 626–638.

[75] A. Eftekhari, J. Romberg, and M.B. Wakin. "Matched filtering from limited frequency samples". In: *arXiv preprint arXiv:1101.2713* (2011).

[76] Moshe Mishali et al. "Xampling: Analog to digital at sub-Nyquist rates". In: *Circuits, Devices & Systems, IET* 5.1 (2011), pp. 8–20.

[77] J-J Fuchs. "On the application of the global matched filter to DOA estimation with uniform circular arrays". In: *Signal Processing, IEEE Transactions on* 49.4 (2001), pp. 702–709.

[78] Chaitanya Ekanadham, Daniel Tranchina, and Eero P Simoncelli. "Sparse decomposition of transformation-invariant signals with continuous basis pursuit". In: *Acoustics, Speech and Signal Processing (ICASSP), 2011 IEEE International Conference on*. IEEE. 2011, pp. 4060–4063.

[79] Emmanuel Candes and Carlos Fernandez-Granda. "Super-resolution from noisy data". In: *arXiv preprint arXiv:1211.0290* (2012).

[80] J.A. Carter and J.N. Winn. "Parameter estimation from time-series data with correlated errors: a wavelet-based method and its application to transit light curves". In: *The Astrophysical Journal* 704.1 (2009), p. 51.

[81] H. Garudadri et al. "Diagnostic grade wireless ECG monitoring". In: *Engineering in Medicine and Biology Society, EMBC, 2011 Annual International Conference of the IEEE*. IEEE. 2011, pp. 850–855.

[82] Claude Ambrose Rogers. *Packing and covering*. University Press, 1964.

[83] Charles Hermite. *Sur un nouveau développement en série des fonctions*. Gauthier-Villars, 1864.

[84] Stephane Mallat. *A wavelet tour of signal processing*. Academic press, 1999.

[85] Ervin Sejdic et al. "Channel estimation using DPSS based frames". In: *Acoustics, Speech and Signal Processing, 2008. ICASSP 2008. IEEE International Conference on*. IEEE. 2008, pp. 2849–2852.

[86] Aad Van der Vaart and Jon Wellner. *Weak convergence and empirical processes: with applications to statistics*. Springer, 1996.

[87] Vladimir Koltchinskii. "Von Neumann entropy penalization and low-rank matrix estimation". In: *The Annals of Statistics* 39.6 (2012), pp. 2936–2973.

[88] Rudolf Ahlswede and Andreas Winter. "Strong converse for identification via quantum channels". In: *Information Theory, IEEE Transactions on* 48.3 (2002), pp. 569–579.

[89] David Gross. "Recovering low-rank matrices from few coefficients in any basis". In: *Information Theory, IEEE Transactions on* 57.3 (2011), pp. 1548–1566.

[90] Benjamin Recht. "A simpler approach to matrix completion". In: *The Journal of Machine Learning Research* 12 (2011), pp. 3413–3430.

[91] V. Koltchinskii, K. Lounici, and A.B. Tsybakov. "Nuclear-norm penalization and optimal rates for noisy low-rank matrix completion". In: *The Annals of Statistics* 39.5 (2011), pp. 2302–2329.

[92] S.R. Becker, E.J. Candès, and M.C. Grant. "Templates for convex cone problems with applications to sparse signal recovery". In: *Mathematical Programming Computation* (2011), pp. 1–54.

[93] F. Alizadeh and D. Goldfarb. "Second-order cone programming". In: *Mathematical programming* 95.1 (2003), pp. 3–51.

[94] Barry Simon. *Trace ideals and their applications*. Vol. 120. American Mathematical Soc., 2010.