

Article

A Note on Shift Retrieval Problems

Cristian Rusu 

Faculty of Mathematics and Computer Science, University of Bucharest, 030018 Bucharest, Romania;
cristian.rusu@fmi.unibuc.ro

Abstract

In this paper, we discuss several shift retrieval problems, both classical and compressed, and provide connections between them using the general framework of circulant matrices. We review the properties of circulant matrices necessary for our calculations and then show how circular shifts can be recovered from as few measurements as possible in different scenarios. We treat several cases: circular shifts between two signals and between multiple pairs of signals, and linear combinations of circular shifts. In all these cases, we provide conditions under which the shift recovery is successful, we give explicit formulas, and we state convex optimization problems for the practical recovery of the shifts for both the noiseless and noisy measurement scenarios. Our goal is to accurately and robustly recover shift information from as few linear measurements as possible. Experimental results then validate the findings through simulation, where we compare the classic cross-correlation result with the proposed approaches.

Keywords: shift retrieval; compressed sensing; cross-correlation; phase correlation; Fourier transform; sparsity

MSC: 62M10; 94A12; 62J07; 15A23; 15A24

1. Introduction

The shift retrieval problem is fundamental to many areas of signal and image processing. Shift retrieval refers to the problem of estimating an unknown translation between two observations or signals. This problem appears wherever alignment or relative displacement must be identified, usually as a preprocessing step before further processing takes place.

The shift retrieval solutions often reveal patterns, delays, or matching features, which are of fundamental importance in pattern recognition, radar processing, or time series analysis.

There are two domains where shift retrieval is essential: time and space. When dealing with signals in time, the problem bears the name Time Delay Estimation (TDE) [1,2], while for space we have image registration [3] and alignment problems [4]. Several applications include synchronization for Global Positioning Systems (GPSs) and wireless communications [5], radar ranging [6] and sonar direction estimation [7], alignment and calibration for speaker localization [8], and augmented reality [9]. Very recent machine learning and multimedia cybersecurity applications include complex alignment and matching tasks in the context of robust feature representations for accurate human parsing in complex scenes [10] and robust watermarking techniques for light-field imaging [11]. Furthermore, modern image retrieval techniques such as the one proposed in [12], which is histogram-based on local neighborhood difference patterns (ELNDP), could benefit from preprocessing with alignment invariance techniques to add robustness to the retrieval process.



Academic Editor: Mingbo Zhao

Received: 27 December 2025

Revised: 22 January 2026

Accepted: 27 January 2026

Published: 2 February 2026

Copyright: © 2026 by the author.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and

conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

Previous work [13,14] has dealt with the problem of recovering the shift between signals given as few measurements as possible. They were the first to show that the shift can be recovered from Fourier measurements using a few samples and less computation compared to the classical cross-correlation setup. They showed that only one Fourier coefficient may suffice to recover the true shift. Subsequently, the work in [15] established a robust shift recovery method based on Bezout's identity and then also analyzed the recovery of weighted sums of two shifts.

Contribution. The shift retrieval problem is typically solved by maximizing the multiplicative cross-correlation between the two signals. Cross-correlation is a powerful tool for finding similarities between signals by systematically testing their alignment. In this paper, we consider different optimization problems that involve the minimization of quantities and whose solutions reduce to entry-wise division operations. In this setting, several known results can be contextualized together, and new results are naturally obtained. The contributions are summarized as follows:

- We give a unified, noiseless/noisy classic, and compressed view of previously known results on the shift retrieval problems (Results 1, 2 and 3; Remark 3);
- We make explicit the connections between the proposed approach and the classical circular convolution and phase cross-correlation theorems (Remark 2);
- We provide the exact necessary conditions under which scaling and shift information is retrieved from noiseless and noisy measurements (Remarks 1 and 4);
- We provide new explicit results that show the performance of the shift recovery accuracy depending on the noise level and the lengths of the signals (Remark 6);
- We provide new results regarding the generalization of shift retrieval problems to pairs of multiple signals (Results 4 and 5);
- We provide a new result related to the recovery of multiple shifts from measurements that are circular linear combinations (Result 6).

This paper is organized as follows: In Section 2, we provide the basic properties of circulant matrices and provide the classic cross-correlation shift retrieval method. Then, in Section 3, in different sub-sections, we analyze several shift retrieval problems and provide solutions based on optimization problems involving circulant matrices. In Section 4, we provide numerical simulation results for shift retrieval accuracy, and then we conclude the paper in Section 5.

Notation. Normal text presents scalars (real or complex valued), lowercase bold letters denote vectors, uppercase bold letters denote matrices, $(\cdot)^T$ is the transpose of a vector/matrix, $(\cdot)^H$ is the Hermitian transpose, $(\cdot)^*$ is the complex conjugate of a scalar/vector/matrix, we denote $j = \sqrt{-1}$, $\odot$ and $\oslash$ denote element-wise multiplication and division, respectively, between matrices of the same size, given a scalar x , then $|x|$ is the absolute value, given a vector $\mathbf{x}$ then $|\mathbf{x}|$ is the entry-wise absolute value. Given a vector $\mathbf{x}$, we denote its Fourier transform as $\tilde{\mathbf{x}}$ or explicitly $\text{FFT}(\mathbf{x})$. For a vector $\mathbf{x}$ of size n , the ℓ_p norms are defined as $\|\mathbf{x}\|_p = (\sum_{i=0}^{n-1} |x_i|^p)^{\frac{1}{p}}$, and for a matrix $\mathbf{X}$ of size $n \times N$, the squared Frobenius norm is defined as $\|\mathbf{X}\|_F^2 = \sum_{i,k} |X_{ik}|^2$. The matrix $\mathbf{I}_n$ is the identity of size $n \times n$, and $\mathbf{0}_{n \times N}$ and $\mathbf{1}_{n \times N}$ are the zero and one matrices of size $n \times N$, respectively. Symbols $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{R}$ and $\mathbb{C}$ denote the natural, integer, real and complex numbers, respectively, and $\Re(\cdot)$ denotes the real part of a complex number. Then $\mathbb{E}[w]$ denotes the expected value of a random variable w . For a vector $\mathbf{x}$, $\text{diag}(\mathbf{x})$ is the diagonal matrix with $\mathbf{x}$ on the diagonal, and for a square matrix $\mathbf{X}$, $\text{diag}(\mathbf{X})$ returns the diagonal vector of $\mathbf{X}$, and $\text{tr}(\mathbf{X}) = \sum_i X_{ii}$ is the trace. Finally, $\text{gcd}(\cdot, \cdot)$ denotes the greatest common divisor of two natural numbers, and $\mathbb{P}(\cdot)$ is the probability that an event will occur. Further, specific notation is explained in the place where it is used.

2. Classic Shift Retrieval Problems

In this section, we introduce several fundamental concepts to our analysis and provide an overview of the classic shift retrieval problem and solution.

2.1. Circulant Matrices Primer

Consider the square circulant matrices defined as

$$\begin{aligned} \mathbf{C} = \text{circ}(\mathbf{c}) &\stackrel{\text{def}}{=} \begin{bmatrix} c_0 & c_{n-1} & \dots & c_2 & c_1 \\ c_1 & c_0 & \dots & c_3 & c_2 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ c_{n-2} & c_{n-3} & \dots & c_0 & c_{n-1} \\ c_{n-1} & c_{n-2} & \dots & c_1 & c_0 \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{c} & \mathbf{P}\mathbf{c} & \mathbf{P}^2\mathbf{c} & \dots & \mathbf{P}^{n-1}\mathbf{c} \end{bmatrix} \in \mathbb{R}^{n \times n}. \end{aligned} \quad (1)$$

The matrix $\mathbf{P} \in \mathbb{R}^{n \times n}$ denotes the orthonormal circulant matrix that circularly down-shifts a target vector $\mathbf{c}$ by one position, i.e., $\mathbf{P} = \text{circ}(\mathbf{e}_2)$ where $\mathbf{e}_2$ is the second vector of the standard basis of $\mathbb{R}^n$. Notice that $\mathbf{P}^{q-1} = \text{circ}(\mathbf{e}_q)$ is also orthonormal circulant and denotes a circular down-shift by $q-1$. Positive powers of $\mathbf{P}$ perform the circular down-shift in a vector, while the negative powers perform the circular up-shift in a vector. Observe that $(\mathbf{P}^{q-1})^{-1} = (\mathbf{P}^{q-1})^T = \mathbf{P}^{1-q}$.

The eigenvalue factorization of circulant matrices reads

$$\mathbf{C} = \mathbf{F}^H \mathbf{\Sigma} \mathbf{F}, \quad \mathbf{\Sigma} = \text{diag}(\boldsymbol{\sigma}) \in \mathbb{C}^{n \times n}, \quad (2)$$

where $\mathbf{F} \in \mathbb{C}^{n \times n}$ is the unitary Fourier matrix ($\mathbf{F}^H \mathbf{F} = \mathbf{F} \mathbf{F}^H = \mathbf{I}_n$) and the diagonal $\boldsymbol{\sigma} = \sqrt{n} \mathbf{F} \mathbf{c}$, $\boldsymbol{\sigma} \in \mathbb{C}^n$. In some situations the factorization is presented as $\mathbf{C} = \mathbf{F}^{-1} \text{diag}(\mathbf{F} \mathbf{c}) \mathbf{F}$. The $\sqrt{n}$ factor is missing explicitly, as it is absorbed in the inverse. In this paper, we will omit the $\sqrt{n}$ term as it does not qualitatively change any of the results. It all depends on whether the Fourier matrix has its elements normalized by some factor $\frac{1}{\sqrt{n}}$, $\frac{1}{n}$, or none at all. Therefore, the quantities we compute hold up to scaling, in general. Different sources in the literature and software implementations use different conventions.

The multiplication with $\mathbf{F}$ is equivalent to the application of the Fast Fourier Transform, i.e., $\tilde{\mathbf{c}} = \mathbf{F} \mathbf{c} = \text{FFT}(\mathbf{c})$, while the multiplication with $\mathbf{F}^H$ is equivalent to the inverse Fourier transform, i.e., $\mathbf{F}^H \mathbf{c} = \mathbf{F}^{-1} \mathbf{c} = \text{IFFT}(\mathbf{c})$. Vectors that are Fourier transforms have the same name as their time-domain counterparts, but with an additional tilde to differentiate them. Of course, $\mathbf{F}^H \mathbf{F} \mathbf{c} = \text{IFFT}(\text{FFT}(\mathbf{c})) = \mathbf{c}$ and $\mathbf{F} \mathbf{F}^H \mathbf{c} = \text{FFT}(\text{IFFT}(\mathbf{c})) = \mathbf{c}$. Both fast transforms are applied in $O(n \log n)$ time and memory.

Given two real-valued matrices $\mathbf{X}$ and $\mathbf{Y}$, both $n \times N$, an immediate application [16,17] of the eigenvalue factorization with the Fourier matrix is to the solution of the problem

$$\underset{\boldsymbol{\sigma}}{\text{minimize}} \quad \|\mathbf{Y} - \mathbf{C}\mathbf{X}\|_F^2, \quad (3)$$

whose solution is given by

$$\sigma_0 = \frac{\tilde{\mathbf{x}}_0^T \tilde{\mathbf{y}}_0}{\|\tilde{\mathbf{x}}_0\|_2^2}, \sigma_i = \frac{\tilde{\mathbf{x}}_i^H \tilde{\mathbf{y}}_i}{\|\tilde{\mathbf{x}}_i\|_2^2}, \sigma_{n-i} = \sigma_i^*, i = 1, \dots, n-1, \quad (4)$$

where $\tilde{\mathbf{y}}_i^T$ and $\tilde{\mathbf{x}}_i^T$ are the rows of $\tilde{\mathbf{Y}} = \mathbf{F} \mathbf{Y}$ and $\tilde{\mathbf{X}} = \mathbf{F} \mathbf{X}$. In the special case when $N = 1$, we have the vectors $\mathbf{x}$ and $\mathbf{y}$ and their Fourier transforms $\tilde{\mathbf{x}}$ and $\tilde{\mathbf{y}}$, respectively. In this paper, we assume the working signals are real-valued.

2.2. Classic Shift Retrieval

Given two signals $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, assuming that $\mathbf{y}$ is a circular shift in $\mathbf{x}$ and that $\mathbf{x}$ is not periodic, in order to find the unique shift quantity, we maximize the inner product

$$\arg \max_q |\mathbf{y}^T \mathbf{P}^{q-1} \mathbf{x}|, \quad (5)$$

where $\mathbf{P}^{q-1} \in \mathbb{R}^{n \times n}$ denotes a circular shift. The calculations above explicitly find all inner products between $\mathbf{x}$ and all possible circular shifts in $\mathbf{y}$, or vice versa. Practically, to recover the shift, we use the circular cross-correlation theorem,

$$\arg \max |\text{IFFT}(\text{FFT}(\mathbf{x})^* \odot \text{FFT}(\mathbf{y}))|, \quad (6)$$

i.e., we take the index of the maximum-magnitude entry of the correlation vector.

The result follows directly from the factorization of Equation (2) by computing correlations between all circular shifts in $\mathbf{x}$ and the vector $\mathbf{y}$ as

$$\begin{aligned} \mathbf{C}^T \mathbf{y} &= \text{circ}(\mathbf{x})^T \mathbf{y} \\ &= \text{circ}(\mathbf{x})^H \mathbf{y} \\ &= \mathbf{F}^H \text{diag}(\mathbf{F}\mathbf{x})^H \mathbf{F}\mathbf{y} \\ &= \text{IFFT}(\text{FFT}(\mathbf{x})^* \odot \text{FFT}(\mathbf{y})). \end{aligned} \quad (7)$$

Therefore, the problem in Equation (6) becomes that of maximizing $\|\mathbf{C}^T \mathbf{y}\|_\infty$. The absolute value removes the distinction between positive and negative correlation of real-valued signals. Our proposed approaches start from a different point of view where the shift quantity q and the circular shift matrix $\mathbf{P}$ are explicitly used. These methods are detailed next.

3. A Circulant Matrix Perspective on the Shift Retrieval Problems

For the sake of clarity, we express and discuss each shift retrieval problem in separate subsections.

3.1. Noiseless Shift Retrieval

In this section, we provide a different view on the classic shift retrieval problem and give the following main result:

Result 1 (Noiseless shift retrieval). *We are given two signals $\mathbf{x}$ and $\mathbf{y}$, and we assume that there is a unique circular shift q between them like*

$$\mathbf{y} = \mathbf{P}^{q-1} \mathbf{x}, \quad (8)$$

then, assuming that $\tilde{x}_i \neq 0$, we have that

$$\text{IFFT}(\text{FFT}(\mathbf{y}) \oslash \text{FFT}(\mathbf{x})) = \mathbf{e}_q. \quad (9)$$

Proof. Assuming that $\mathbf{y} = \mathbf{P}^{q-1} \mathbf{x}$, with $\mathbf{P} = \text{circ}(\mathbf{e}_2)$, we consider the problem

$$\underset{q}{\text{minimize}} \|\mathbf{y} - \mathbf{P}^{q-1} \mathbf{x}\|_2^2. \quad (10)$$

Use $\mathbf{P}^{q-1} = \mathbf{F}^H \mathbf{\Sigma} \mathbf{F}$ and to develop $\|\mathbf{y} - \mathbf{P}^{q-1} \mathbf{x}\|_2^2 = \|\mathbf{y} - \mathbf{F}^H \mathbf{\Sigma} \mathbf{F} \mathbf{x}\|_2^2 = \|\mathbf{F} \mathbf{y} - \mathbf{\Sigma} \mathbf{F} \mathbf{x}\|_2^2 = \|\tilde{\mathbf{y}} - \mathbf{\Sigma} \tilde{\mathbf{x}}\|_2^2$, where $\mathbf{\Sigma} = \text{diag}(\sigma)$, $\sigma = \mathbf{F} \mathbf{e}_q = \mathbf{f}_q \in \mathbb{C}^n$ (the q^{th} column of the Fourier matrix).

If we relax the constraint and allow $\mathbf{P}^{q-1}$ to be any circulant matrix, to minimize the Frobenius norm, as the special case of Equation (3) for $N = 1$, we have $\sigma_i = \tilde{y}_i / \tilde{x}_i$, $\tilde{x}_i \neq 0$, and therefore $\mathbf{F}\mathbf{e}_q = \tilde{\mathbf{y}} \oslash \tilde{\mathbf{x}}$. $\square$

The assumption that $\tilde{x}_i \neq 0$ seems restrictive (and is missing in Equation (6)). We do not need to apply the inverse Fourier transform, but instead compute only σ_i where $\tilde{x}_i \neq 0$, and by inspection of all columns of $\mathbf{F}$ on the rows where this quantity was computed, we find the shift q .

If we want to recover the true shift between signals $\mathbf{x}$ and $\mathbf{y}$, then the following conditions need to hold:

Remark 1 (Necessary conditions for the recovery of the true shift). *In order to uniquely recover the true shift q from a single measurement i , the following properties need to hold:*

1. $\tilde{x}_i \neq 0$, so that the ratio σ_i is well defined;
2. $i \neq 0$, which is the DC component;
3. $\gcd(i, n) = 1$, which excludes the special case $i = \frac{n}{2}$ for n even.

Proof. Notice that σ_0 and $\sigma_{\frac{n}{2}}$ when n is even are $\{\pm 1\}$ for all columns of the Fourier matrix, and thus they cannot provide an unambiguous answer. In fact, σ_0 provides no information about the shift, while $\sigma_{\frac{n}{2}}$ only establishes the parity of the shift. On the other hand, in the best-case scenario, we need to compute a single σ_i to recover the true shift q . Given the ratio $\sigma_i = \frac{\tilde{y}_i}{\tilde{x}_i} = \exp\left(-j\frac{2\pi i}{n}q\right)$, two different shifts q and q' lead to the same quantity σ_i iff $\exp\left(-j\frac{2\pi i}{n}(q - q')\right) = 1$ and therefore $\frac{i}{n}(q - q') \in \mathbb{Z}$. To avoid having the exponential equal one for $q \neq q'$, we require $\gcd(i, n) = 1$. $\square$

By this remark, the complexity of the shift retrieval problem is $O(n)$, as also observed for the compressive shift retrieval result [14], which is discussed in the next subsection.

Remark 2 (Connection to the classical circular and phase cross-correlation theorems). *We can rewrite Equation (9) as*

$$\text{IFFT}(\text{FFT}(\mathbf{y}) \oslash \text{FFT}(\mathbf{x})) = \text{IFFT}(\text{FFT}(\mathbf{x})^* \odot \text{FFT}(\mathbf{y}) \oslash |\text{FFT}(\mathbf{x})|^2), \quad (11)$$

where $|\text{FFT}(\mathbf{x})|^2$ computes the square absolute values of each element of the Fourier transform of $\mathbf{x}$ which we assume are all non-zero. Notice that Equation (9) represents a weighted variant of Equation (6), a type of “whitened” cross-correlation quantity.

The expression in Equation (11) naturally relates to the well-known phase correlation formula

$$\text{IFFT}(\text{FFT}(\mathbf{x})^* \odot \text{FFT}(\mathbf{y}) \oslash (|\text{FFT}(\mathbf{x})^* \odot \text{FFT}(\mathbf{y})|)), \quad (12)$$

The normalization performed in the formula above removes the amplitude information, i.e., this is essentially a phase-only calculation because magnitude is removed by the division operation. In our scenario, we are interested in preserving and recovering the amplitude information as well.

If the two signals $\mathbf{x}$ and $\mathbf{y}$ are shifted versions of each other, then Equations (6) and (9) provide the same answer. If this is not the case, or the signals are noisy, then Equation (9) seems a weaker result in general since the minimizer $\mathbf{P}^{q-1}$ in Equation (10) might no longer be $\mathbf{P} = \text{circ}(\mathbf{e}_2)$, but some other circulant matrix that minimizes Equation (10). In this high-noise case, we might not be able to interpret that the signals are shifted versions of each other. The circulant cross-correlation theorem does not have this feature, as it will always provide the maximum correlation between the signals.

We note that the approach to maximize the quadratic form Equation (5) and that of norm minimization Equation (10) are equivalent since

$$\|\mathbf{y} - \mathbf{P}^{q-1}\mathbf{x}\|_2^2 = \|\mathbf{y}\|_2^2 + \|\mathbf{x}\|_2^2 - 2\mathbf{y}^T \mathbf{P}^{q-1}\mathbf{x}, \quad (13)$$

and then also

$$\begin{aligned} \|\mathbf{y} - \mathbf{P}^{q-1}\mathbf{x}\|_2^2 &= \|\mathbf{F}\mathbf{y} - (\text{diag}(\mathbf{F}\mathbf{e}_2))^{q-1}\mathbf{F}\mathbf{x}\|_2^2 \\ &= \|\tilde{\mathbf{y}} - \text{diag}(\mathbf{F}\mathbf{e}_q)\tilde{\mathbf{x}}\|_2^2 \\ &= \|\tilde{\mathbf{y}} - \text{diag}(\mathbf{f}_q)\tilde{\mathbf{x}}\|_2^2 \\ &= \|\tilde{\mathbf{y}}\|_2^2 + \|\tilde{\mathbf{x}}\|_2^2 - 2\Re(\tilde{\mathbf{y}}^H \text{diag}(\mathbf{f}_q)\tilde{\mathbf{x}}) \\ &= \|\tilde{\mathbf{y}}\|_2^2 + \|\tilde{\mathbf{x}}\|_2^2 - 2 \sum_{i=0}^{n-1} \tilde{y}_i^* \tilde{x}_i F_{iq}, \end{aligned} \quad (14)$$

where F_{iq} is an element from the Fourier matrix and the last quantity is real-valued due to the conjugate valued symmetries of the vectors $\tilde{\mathbf{x}}$ and $\tilde{\mathbf{y}}$, and of the columns $\mathbf{f}_q$. The last summation quantity is equivalent to Equation (6) for a fixed q . The result of Equation (9) is obtained by allowing the unknown to be the overall general circulant matrix denoted $\mathbf{P}^{q-1}$, not just the power q .

Finally, note that for real-valued $\mathbf{x}$ and $\mathbf{y}$ we have that $\text{IFFT}(\text{FFT}(\mathbf{x})^* \odot \text{FFT}(\mathbf{y}))$ is equivalent to $\text{FFT}(\text{FFT}(\mathbf{x}) \odot \text{FFT}(\mathbf{y})^*)$, up to a normalization factor depending on n .

Remark 3 (Calculation of the circular shift from one measurement). Notice that Equation (9) is equivalent to

$$\text{FFT}(\mathbf{y}) \odot \text{FFT}(\mathbf{x}) = \mathbf{f}_q, \quad (15)$$

where $\mathbf{f}_q$ is the q^{th} column of the Fourier matrix $\mathbf{F}$. We can find the shift by computing a single entry $\sigma_i = \tilde{y}_i / \tilde{x}_i$ and then inspecting the entries of only the i^{th} row of the Fourier matrix. In this case, the recovery of the shift is performed via

$$q^* = -i^{-1} \arg(\sigma_i) \frac{n}{2\pi} \mod n, \quad (16)$$

where i^{-1} is the modular inverse of i modulo n .

This is the result previously developed by the work in [13,14]. Essentially, these results are a consequence of the well-known shift theorem for the Discrete Fourier Transform (DFT), which makes the connection between multiplication by pure phase factors and circular shifts in vectors, i.e., circular shift in a vector corresponds to multiplying its DFT by a linear phase. Treatments of this classic result can be found in fundamental signal processing sources such as ([18], Section 8.6.2).

More generally, when the measurement $\mathbf{y}$ has a different scale and mean than the signal $\mathbf{x}$, we have the following remark:

Remark 4 (Necessary conditions for the recovery of the true shift, scale, and mean). When the measurements are given by $\mathbf{y} = \alpha \mathbf{P}^{q-1}\mathbf{x} + \beta \mathbf{1}$, where $\alpha, \beta \in \mathbb{R}$ are scalars such that $\alpha \neq 0$ and $\sum_{i=0}^{n-1} x_i \neq 0$, then

$$\text{IFFT}(\tilde{\mathbf{y}} \odot \tilde{\mathbf{x}}) = \alpha \mathbf{e}_q + \left(\beta / \sum_{i=0}^{n-1} x_i \right) \mathbf{1}, \quad (17)$$

where $\mathbf{1} \in \mathbb{R}^n$ is the ones vector. In general, we need three measurements to recover all parameters (q, α, β) .

Proof. The mean component is easy to identify by using the previously uninformative DC component. Notice that for a single non-DC component such that $\tilde{x}_i \neq 0$ we have that

$$\sigma_i = \frac{\tilde{y}_i}{\tilde{x}_i} = \alpha \exp\left(-j \frac{2\pi i}{n} q\right). \quad (18)$$

When $\alpha > 0$, there is no change in the phase, and therefore the scale and shift can be recovered simultaneously. For any integer $\alpha \neq 0$, then we have that the phase of σ_i is not modified (if α is positive) or is modified by π (if α is negative). As a consequence of this phase change we cannot distinguish between the shift-scale pair (q, α) and $(q + \frac{n}{2i}, -\alpha)$ when $\frac{n}{2i}$ is a valid shift integer in $\{0, 1, \dots, n-1\}$. In general, this ambiguity cannot be resolved from a single non-DC measurement. To uniquely recover all three parameters, we will require two non-DC components and the DC component (for the calculation of the mean β). Given two distinct non-DC measurements $k_1 \neq k_2$,

$$\sigma_{k_1} = \frac{\tilde{y}_{k_1}}{\tilde{x}_{k_1}} = \alpha \exp\left(-j \frac{2\pi k_1}{n} q\right), \quad (19)$$

$$\sigma_{k_2} = \frac{\tilde{y}_{k_2}}{\tilde{x}_{k_2}} = \alpha \exp\left(-j \frac{2\pi k_2}{n} q\right). \quad (20)$$

Now take the ratio of the two quantities to eliminate α and reach

$$\frac{\sigma_{k_1}}{\sigma_{k_2}} = \exp\left(-j \frac{2\pi(k_1 - k_2)}{n} q\right). \quad (21)$$

Analogously to Remark 1, the necessary condition for the recovery of the true shift is that $\gcd(k_1 - k_2, n) = 1$. We call the best estimated shift

$$q^* = -(k_1 - k_2)^{-1} \text{round}\left(\arg\left(\frac{\sigma_{k_1}}{\sigma_{k_2}}\right) \frac{n}{2\pi}\right) \mod n. \quad (22)$$

Then, we compute the other two parameters,

$$\alpha^* = \frac{\sigma_{k_1}}{\exp\left(-j \frac{2\pi k_1}{n} q^*\right)}, \quad (23)$$

$$\beta^* = \frac{\tilde{y}_0 - \alpha^* \tilde{x}_0}{n}. \quad (24)$$

□

Just as the classic shift retrieval of Equation (5) is indifferent to the sign of the correlation between signals, Remark 4 provides the result indifferent to the sign of the correlation, allowing for both positive and negative α and therefore positive and negative correlations.

The results developed in this section assume no noise is present in the measurements. As already explained in Remark 2, we expect noisy measurements to significantly degrade the performance of the shift retrieval and make impractical the recovery mechanism from a single measurement.

3.2. Noisy Shift Retrieval

We are again given two signals $\mathbf{x}$ and $\mathbf{y}$ such that there is a noisy circular shift q between them,

$$\mathbf{y} = \mathbf{P}^{q-1} \mathbf{x} + \mathbf{w}, \quad (25)$$

where $\mathbf{w} \sim \mathcal{N}(\mathbf{0}_{n \times 1}, \zeta^2 \mathbf{I}_n)$ is an i.i.d. zero-mean Gaussian noise vector of size n . Unlike the noiseless scenario, here we will assume the number of measurements $m > 1$ and try to recover the most likely shift parameter. Then we have the following result.

Result 2 (The noisy recovery of the true shift, scale, and mean). *When the measurements are given by*

$$\mathbf{y} = \alpha \mathbf{P}^{q-1} \mathbf{x} + \beta \mathbf{1} + \mathbf{w}, \quad (26)$$

where $\alpha, \beta \in \mathbb{R}$ are scalars such that $\alpha \neq 0$, then given m measurements with indices in the set $\mathcal{K} = \{k_i\}_{i=1}^{m-1}$, such that $\gcd(k_1 - k_2, n) = 1$ for any two distinct $k_1, k_2 \in \mathcal{K}$ and $m - 1 \geq 2$, plus the DC component, which is treated separately, we have the following estimates:

$$q^* = \arg \max_q \left[\Re \left\{ \sum_{i=1}^{m-1} |\tilde{x}_{k_i}|^2 \sigma_{k_i} \exp \left(j \frac{2\pi k_i}{n} q \right) \right\} \right]^2, \quad (27)$$

$$\alpha^* = \frac{\Re \left\{ \sum_{i=1}^{m-1} |\tilde{x}_{k_i}|^2 \sigma_{k_i} \exp \left(j \frac{2\pi k_i}{n} q^* \right) \right\}}{\sum_{i=1}^{m-1} |\tilde{x}_{k_i}|^2}, \quad (28)$$

$$\beta^* = \frac{\tilde{y}_0 - \alpha^* \tilde{x}_0}{n}. \quad (29)$$

Proof. Assuming $\tilde{x}_{k_i} \neq 0$ for $k_i \in \mathcal{K}$, define $\sigma_{k_i} = \tilde{y}_{k_i} / \tilde{x}_{k_i}$ and the following measurement vector $(\tilde{\mathbf{y}})_{\mathcal{K}}$, which is of length $m - 1$ and contains the Fourier transform of the vector $\mathbf{y}$ restricted to the indices from the set $\mathcal{K}$. Define also the vector $\mathbf{s} = (\tilde{\mathbf{x}})_{\mathcal{K}} \odot \boldsymbol{\omega}$, where $\boldsymbol{\omega}$ is again a vector of size $m - 1$ whose elements are the quantities $\omega_i = \exp(-j \frac{2\pi k_i}{n} q)$ for $k_i \in \mathcal{K}$. The parameter β is estimated again from the DC component by minimizing $(\tilde{y}_0 - \alpha \tilde{x}_0 - \beta n)^2$ for β . The scale is estimated by minimizing $\|(\tilde{\mathbf{y}})_{\mathcal{K}} - \alpha \mathbf{s}\|_2^2$ for α which leads to $\alpha^* = \frac{\Re\{\mathbf{s}^H(\tilde{\mathbf{y}})_{\mathcal{K}}\}}{\|\mathbf{s}\|_2^2} = \frac{\Re\{\sum_{i=1}^{m-1} \tilde{x}_{k_i}^* \tilde{y}_{k_i} \exp(j 2\pi k_i q/n)\}}{\|(\tilde{\mathbf{x}})_{\mathcal{K}}\|_2^2}$, note that the result is a function of q . Then, in terms of the shift, the minimum residual quantity is given by $\|(\tilde{\mathbf{y}})_{\mathcal{K}} - \alpha \mathbf{s}\|_2^2 = \|(\tilde{\mathbf{y}})_{\mathcal{K}}\|_2^2 - \frac{[\Re\{\mathbf{s}^H(\tilde{\mathbf{y}})_{\mathcal{K}}\}]^2}{\|\mathbf{s}\|_2^2}$. Then, the optimal q^* is given by maximizing the numerator in the residual quantity. $\square$

As we increase the number of measurements m , we expect to improve the accuracy of the recovery. It is of interest to investigate if there are ways to choose the index set $\mathcal{K}$ to improve accuracy. The following remark speaks to this choice.

Remark 5 (Selection criteria for the Fourier measurements). *Given noisy measurements $\mathbf{y} = \mathbf{P}^{q-1} \mathbf{x} + \mathbf{w}$, in order to recover the shift quantity q while reducing the effects of the noise, one should choose indices k_i such that $|\tilde{x}_{k_i}|$ have the largest values.*

Proof. The observation stems from the ratio σ_{k_i} whose variance is proportional to $\frac{1}{|\tilde{x}_{k_i}|^2}$. This follows from

$$\sigma_{k_i} = \frac{\alpha \tilde{y}_i + \tilde{w}_{k_i}}{\tilde{x}_{k_i}} = \alpha \exp \left(-j \frac{2\pi k_i}{n} q \right) + \frac{\tilde{w}_{k_i}}{\tilde{x}_{k_i}}, \quad (30)$$

where $\tilde{\mathbf{w}}$ is the Fourier transform of the noise vector. Because $\mathbf{F}$ is unitary and the noise is zero mean and i.i.d., we have that the variance is

$$\mathbb{E} \left[\left| \frac{\tilde{w}_{k_i}}{\tilde{x}_{k_i}} \right|^2 \right] = \frac{\mathbb{E}[|\tilde{w}_{k_i}|^2]}{|\tilde{x}_{k_i}|^2} = \frac{\zeta^2}{|\tilde{x}_{k_i}|^2}. \quad (31)$$

To maximally reduce this quantity we take the largest $|\tilde{x}_{k_i}|$. $\square$

In the simulation results section, we investigate the impact of this index choice on the accuracy performance of recovering the shift from the noisy measurements. While choosing the highest entries $|\tilde{x}_{k_i}|$ is desirable, this might be difficult to achieve in general without computing the whole spectrum. Without any information on the spectrum of the signal, the Goertzel algorithm [19] can be applied to compute a few Fourier coefficients, still obeying $\gcd(k_i, n) = 1$. Still, the indices of these coefficients have to be given a priori. Keeping $\gcd(k_i, n) = 1$ means that the set of eligible indices has size $\varphi(n)$, the Euler totient function. If no information is available about the spectrum then selecting the coefficients randomly is a last method of choice. Approaches such as the Fastest Fourier Transform in the West (FFTW) [20] allow for the calculation of pruned FFT, which compute only a subset of outputs of the FFT at the cost of $O(n \log s)$ for s Fourier coefficients.

Furthermore, in some applications where we expect the spectrum to be very sparse (many Fourier coefficients will be zero, or close to zero), the Sparse Fourier Transform (SFT) [21,22] can be applied to compute the s largest magnitude Fourier coefficients with complexity $O(s \log n)$ if the signal is exactly s -sparse or $O(s \log n \log(n/s))$ if the signal is s -sparse plus noise. The SFT is probabilistic, so usually, we do not compute just one coefficient, but a few $s > 1$ and select the largest in magnitude. An interesting point here is to note that these SFTs work on the principle of intentionally allowing aliasing to occur in order to group Fourier coefficients in the same bins, allowing conflicts that are subsequently resolved. This is relevant in our case because, by carefully choosing the aliasing, we could group the calculations in bins such that Fourier coefficients that are not of interest are grouped in the same bins, and then in separate bins, we group only the coefficients of interest. Therefore, resolving the collisions occurs only in the bins of interest.

The next remark addresses the issue of comparing, in the noisy case, the classic shift recovery method against the one measurement approach as a function of n and the signal-to-noise ratio.

Remark 6 (Expected accuracy performance and comparison between the circular convolution theorem and the one measurement model). *Given Result 2, we expect the shift recovery performance to follow:*

1. *The recovery of the shift by the circular convolutional theorem improves as the length of the signals n increases or the variance of the noise ζ^2 decreases, i.e., it is easier to recover the correct shift under these circumstances;*
2. *For shift recovery from a single measurement, as n increases, the noise variance ζ^2 for which the recovery accuracy approaches 100% decreases, i.e., it is harder to recover the correct shift under these circumstances;*
3. *For shift recovery from a single measurement, as the variance decreases to zero, i.e., $\zeta^2 \rightarrow 0$, the probability of correct recovery is bounded by $\frac{1}{\gcd(i,n)}$.*

Proof. Assuming the measurement model in Equation (25), for two different shifts q and q' such that $q \neq q'$, and the respective shifted vectors $\mathbf{x}_q$ and $\mathbf{x}_{q'}$, the idea is to compute the probability of choosing the wrong shift q' over the correct q as $\mathbb{P}(\|\mathbf{y} - \mathbf{x}_{q'}\|_2^2 \leq \|\mathbf{y} - \mathbf{x}_q\|_2^2)$. The inequality is reduced from $\|\mathbf{y} - \mathbf{x}_{q'}\|_2^2 \leq \|\mathbf{y} - \mathbf{x}_q\|_2^2$ to the equivalent

$\|\mathbf{x}_{q'} - \mathbf{x}_q\|_2^2 \leq 2\mathbf{w}^T(\mathbf{x}_{q'} - \mathbf{x}_q)$. Because $\mathbf{w} \sim \mathcal{N}(\mathbf{0}_{n \times 1}, \zeta^2 \mathbf{I}_n)$ we have that $\mathbf{w}^T(\mathbf{x}_{q'} - \mathbf{x}_q) \sim \mathcal{N}(0, \zeta^2 \|\mathbf{x}_{q'} - \mathbf{x}_q\|_2^2)$. Finally, note that

$$\begin{aligned} \|\mathbf{x}_{q'} - \mathbf{x}_q\|_2^2 &= \|\mathbf{P}^{q'-1} \mathbf{x} - \mathbf{P}^{q-1} \mathbf{x}\|_2^2 \\ &= \|\mathbf{x} - (\mathbf{P}^{q'-1})^T \mathbf{P}^{q-1} \mathbf{x}\|_2^2 \\ &= \|\mathbf{x} - \mathbf{P}^{1-q'} \mathbf{P}^{q-1} \mathbf{x}\|_2^2 \\ &= \|\mathbf{x} - \mathbf{P}^{q-q'} \mathbf{x}\|_2^2 \\ &= \|\mathbf{x} - \mathbf{x}_{q-q'+1}\|_2^2. \end{aligned} \quad (32)$$

Therefore, the results depend only on the nonzero differences between the two shifts we consider. Then, it follows that for any shift difference $d = q - q' + 1 \neq 0$ we have that

$$\|\mathbf{x} - \mathbf{x}_d\|_2^2 = \|\mathbf{x}\|_2^2 + \|\mathbf{x}_d\|_2^2 - 2\mathbf{x}^T \mathbf{x}_d = 2\|\mathbf{x}\|_2^2(1 - \rho_d), \quad (33)$$

where ρ_d is the normalized circular autocorrelation coefficient for delay d . We assume $|\rho_d| < 1$ for all $d \neq 0$, i.e., there is no periodicity in the signal $\mathbf{x}$. Then, the probability of error is given by

$$\begin{aligned} \mathbb{P}(\|\mathbf{y} - \mathbf{x}_{q'}\|_2^2 \leq \|\mathbf{y} - \mathbf{x}_q\|_2^2) &= Q\left(\frac{\|\mathbf{x} - \mathbf{x}_d\|_2}{2\zeta}\right) \\ &= Q\left(\sqrt{\frac{(1 - \rho_d)\|\mathbf{x}\|_2^2}{2\zeta^2}}\right) \\ &= Q\left(\sqrt{\frac{n(1 - \rho_d)\text{SNR}}{2}}\right). \end{aligned} \quad (34)$$

If we would allow $\rho_d = 1$ then $Q(0) = \frac{1}{2}$, describing the inherent uncertainty in the shift recovery between q and q' .

We have defined the Q-function as the tail distribution function of the standard normal distribution and $\text{SNR} = \frac{1}{n} \frac{\|\mathbf{x}\|_2^2}{\zeta^2}$. Finally, for q^* given by Result 2, by a union bound we have

$$\begin{aligned} \mathbb{P}(q^* \neq q) &\leq \sum_{d=1}^{n-1} Q\left(\sqrt{\frac{n(1 - \rho_d)\text{SNR}}{2}}\right) \\ &= (n-1)Q\left(\sqrt{\frac{n(1 - \rho_{\max})\text{SNR}}{2}}\right), \end{aligned} \quad (35)$$

where $\rho_{\max} = \max_d \rho_d$ is the highest autocorrelation coefficient. This shows that increasing n or SNR leads to a lower expected error upper bound. Naturally, high magnitude coefficients in the autocorrelation increase the probability of error. In fact, $\mathbb{P}(q^* \neq q) \rightarrow 0$ as $\text{SNR} \rightarrow \infty$ or $n \rightarrow \infty$.

In the case of a single measurement, the main difficulty is the angular separation between adjacent phases on the unit circle, which is $2\pi/n$. As n increases, this separation decreases, so at any fixed measurement SNR, the success probability of recovery decreases. Assuming that we use the classic approximation $2\sin(\pi/n) \approx 2\pi/n$ instead of the arc length distance between neighboring phases, correct identification occurs when the phase error in absolute value stays within half the grid spacing, i.e., π/n . Therefore, by Remark 5, the probability of error for a single measurement whose estimate q^* is given by Equation (16) is

$$\mathbb{P}_1(q^* \neq q) \approx 2Q\left(\sqrt{2 \frac{|\tilde{x}_i|^2}{\|\mathbf{x}\|_2^2} \text{SNR} \frac{\pi}{n}}\right). \quad (36)$$

Note there that the relationship with SNR is the same as in Equation (35), but the relationship with n is inverse, i.e., larger n decreases the probability of correct shift recovery from a single measurement. According to Equation (36), for signals of length $2n$ we need an SNR approximately $10 \log_{10}(4) \approx 6$ dB larger to achieve the same shift recovery accuracy as for n .

Finally, note that because $i(q - q') = 0 \pmod n$ has exactly $\gcd(i, n)$ solutions, and assuming that q is uniformly distributed among all possible shifts, then the maximum possible probability of exact recovery is at most $\frac{1}{\gcd(i, n)}$. This is because $\gcd(i, n)$ shifts the map to the same phase and only $\frac{n}{\gcd(i, n)}$ distinct phases. Therefore, the probability of Equation (36) asymptotically tends to one as the $\text{SNR} \rightarrow \infty$ iff $\gcd(i, n) = 1$. $\square$

Next, we look at the compressed sensing extension of the shift retrieval problem and several other, more general shift retrieval problems.

3.3. The Compressive Shift Retrieval Problem

The compressive shift retrieval problem has been previously introduced [13,14]. In this section, we show how this result can also be described in the overall structure developed in this paper.

Define the sensing matrix $\mathbf{A} \in \mathbb{C}^{m \times n}$, $m \leq n$, and the compressed measurement signals $\mathbf{z} = \mathbf{A}\mathbf{y} \in \mathbb{C}^m$ and $\mathbf{v} = \mathbf{A}\mathbf{x} \in \mathbb{C}^m$. Assuming that $\mathbf{y}$ is a circular shift in $\mathbf{x}$, the goal is to determine the shift from $\mathbf{z}$ and $\mathbf{v}$. Similarly to Equation (5), consider the test (Corollary 2 in [13]):

$$\arg\max_q \Re\{\mathbf{z}^H \tilde{\mathbf{P}}^{q-1} \mathbf{v}\}, \quad (37)$$

where $\tilde{\mathbf{P}}^{q-1} = \mathbf{A}\mathbf{P}^{q-1}\mathbf{A}^H$. It has been shown that when $\mathbf{A}$ is taken to be a partial Fourier matrix, then ([13], Corollary 4):

$$\max_{q \in \{0, \dots, n-1\}} \Re\left\{\sum_{i=1}^m z_i^* v_i e^{\frac{-2\pi j k_i q}{n}}\right\}, \quad (38)$$

recovers the true shift if there exists $i \in \{1, \dots, m\}$ such that $\tilde{x}_{k_i} \neq 0$ (the k_i^{th} coefficient of the Fourier transform of $\mathbf{x}$) and $\{1, \dots, n-1\} \frac{k_i}{n}$ contain no integers. The set $\mathcal{K} = \{k_i\}_{i=1}^m$ contains the indices of the rows contained in the partial Fourier matrix $\mathbf{A}$. Following ([13], Theorem 1), we assume that the sensing matrix $\mathbf{A}$ obeys: $\mathbf{A}^H \mathbf{A} \mathbf{P}^{q-1} = \mathbf{P}^{q-1} \mathbf{A}^H \mathbf{A}$, $\exists \gamma \in \mathbb{R}$ so that $\gamma \mathbf{A} \mathbf{A}^H = \mathbf{I}_m$ and all columns of $\mathbf{A} \text{circ}(\mathbf{x})$ are different so that there is no ambiguity in the shift in the measurements. Without loss of generality, assume $\gamma = 1$.

The compressive shift retrieval result is partly based on the fact that $\mathbf{A}^H \mathbf{A} \mathbf{P}^{q-1} = \mathbf{P}^{q-1} \mathbf{A}^H \mathbf{A}$. Notice that $\mathbf{A}^H \mathbf{A} = \mathbf{F}^H \mathbf{\Sigma} \mathbf{F}$ where the diagonal $\mathbf{\Sigma}$ contains $\{0, 1\}$ with ones on the positions where the rows of the Fourier matrix are selected (the set $\mathcal{K}$). Notice that $\mathbf{A}^H \mathbf{A}$ is a circulant and thus it commutes with $\mathbf{P}^{q-1}$ – they have the same eigenspace. Also, given a set $\mathcal{K}$ of indices, we define the operation $[\mathbf{a}]_{\mathcal{K}} = \mathbf{b}$ for vectors $\mathbf{a} \in \mathbb{C}^n$, $\mathbf{b} \in \mathbb{C}^m$, $m \leq n$, as equality between values $\mathbf{b}$ and positions $\mathcal{K}$ of $\mathbf{a}$, leaving the rest of the values of $\mathbf{a}$ indexed in $\{1, 2, \dots, n\} \setminus \mathcal{K}$ to zero.

Result 3 (Circulant compressive shift retrieval with a proof based on circulant matrices). Given $\mathbf{z} = \mathbf{A}\mathbf{y}$ and $\mathbf{v} = \mathbf{A}\mathbf{x}$ where $\mathbf{y} = \mathbf{P}^{q-1}\mathbf{x}$, assuming $v_i \neq 0, i = 1, \dots, m$, then

$$(\mathbf{F}\mathbf{e}_q)_{\mathcal{K}} = \mathbf{z} \oslash \mathbf{v}. \quad (39)$$

Proof. We start again from the least squares problem,

$$\underset{q}{\text{minimize}} \|\mathbf{z} - \mathbf{A}\mathbf{P}^{q-1}\mathbf{A}^H\mathbf{v}\|_2^2. \quad (40)$$

With the assumption that $\mathbf{y} - \mathbf{P}^{q-1}\mathbf{x} = \mathbf{0}_{n \times 1}$ the objective reaches the zero minimum,

$$\begin{aligned} \mathbf{A}\mathbf{y} - \mathbf{A}\mathbf{P}^{q-1}\mathbf{x} &= \mathbf{A}\mathbf{y} - \mathbf{A}\mathbf{A}^H\mathbf{A}\mathbf{P}^{q-1}\mathbf{x} \\ &= \mathbf{A}\mathbf{y} - \mathbf{A}\mathbf{P}^{q-1}\mathbf{A}^H\mathbf{A}\mathbf{x} \\ &= \mathbf{z} - \mathbf{A}\mathbf{P}^{q-1}\mathbf{A}^H\mathbf{v}, \end{aligned} \quad (41)$$

where we used the commutativity of circulant matrices and that $\mathbf{A}\mathbf{A}^H = \mathbf{I}_m$. To develop Equation (40), start again from Equation (2) and the expression of the matrix multiplication as $\text{vec}(\mathbf{A}\mathbf{F}^H\mathbf{\Sigma}\mathbf{F}\mathbf{A}^H\mathbf{v}) = ((\mathbf{F}\mathbf{A}^H\mathbf{v})^T \otimes (\mathbf{A}\mathbf{F}^H))\text{vec}(\mathbf{\Sigma})$. We finally obtain

$$\begin{aligned} \|\mathbf{z} - \mathbf{A}\mathbf{P}^{q-1}\mathbf{A}^H\mathbf{v}\|_2^2 &= \|\mathbf{z} - \mathbf{A}\mathbf{F}^H\mathbf{\Sigma}\mathbf{F}\mathbf{A}^H\mathbf{v}\|_2^2 \\ &= \|\text{vec}(\mathbf{z}) - \text{vec}(\mathbf{A}\mathbf{F}^H\mathbf{\Sigma}\mathbf{F}\mathbf{A}^H\mathbf{v})\|_2^2 \\ &= \|\mathbf{z} - ((\mathbf{F}\mathbf{A}^H\mathbf{v})^T \otimes (\mathbf{A}\mathbf{F}^H))\text{vec}(\mathbf{\Sigma})\|_2^2 \\ &= \|\mathbf{z} - \mathbf{V}\mathbf{F}\mathbf{e}_q\|_2^2, \end{aligned} \quad (42)$$

where the matrix $\mathbf{V}$ of size $m \times n$ contains only the columns of the Kronecker product that match the non-zero elements of the diagonal matrix $\mathbf{\Sigma}$. The matrix contains the elements of $\mathbf{v}$ in positions (k_i, i) . The second equality holds because the ℓ_2 norm is element-wise, and therefore applying the vec operator does not change the value. It follows that $\mathbf{V}\mathbf{F}\mathbf{e}_q = \mathbf{z}$ and

$$\begin{aligned} (\mathbf{F}\mathbf{e}_q)_{\mathcal{K}} &= \mathbf{V}^H(\mathbf{V}\mathbf{V}^H)^{-1}\mathbf{z} \\ &= \mathbf{V}^H(\mathbf{z} \oslash |\mathbf{v}|^2) \\ &= \mathbf{v}^* \odot \mathbf{z} \oslash |\mathbf{v}|^2 \\ &= \mathbf{z} \oslash \mathbf{v}. \end{aligned} \quad (43)$$

The compressive shift retrieval is equivalent to Equation (9), the regular shift retrieval, on the set of Fourier components $\mathcal{K}$. This is a unified view of the classic and compressed shift retrieval solutions. $\square$

In relation to Equation (38), we use the circulant structures to reach

$$\begin{aligned} \mathbf{z}^H\tilde{\mathbf{P}}^{q-1}\mathbf{v} &= \mathbf{z}^H\mathbf{A}\mathbf{F}^H\mathbf{\Sigma}\mathbf{F}\mathbf{A}^H\mathbf{v} \\ &= \text{vec}(\mathbf{z}^H\mathbf{A}\mathbf{F}^H\mathbf{\Sigma}\mathbf{F}\mathbf{A}^H\mathbf{v}) \\ &= ((\mathbf{F}\mathbf{A}^H\mathbf{v})^T \otimes (\mathbf{z}^H\mathbf{A}\mathbf{F}^H))\text{vec}(\mathbf{\Sigma}) \\ &= \mathbf{r}^T\mathbf{F}\mathbf{e}_q, \end{aligned} \quad (44)$$

where we expressed the matrix multiplications as a linear transformation on $\mathbf{\Sigma} = \text{diag}(\mathbf{F}\mathbf{e}_q)$ and $\mathbf{r} \in \mathbb{C}^n$ is the expression in the parenthesis with $[\mathbf{r}]_{\mathcal{K}} = \mathbf{z}^* \odot \mathbf{v}$. The matrix $\mathbf{F}\mathbf{A}^H \in \mathbb{R}^{n \times m}$ is a partial permutation matrix—only positions (k_i, i) are non-zero. The products with $\mathbf{v}$ and $\mathbf{z}$ produce extended vectors $[\mathbf{v}]_{\mathcal{K}}, [\mathbf{z}]_{\mathcal{K}} \in \mathbb{C}^n$. Thus, maximizing $\mathbf{z}^H\tilde{\mathbf{P}}^{q-1}\mathbf{v}$ reduces to the selection of $\mathbf{e}_q$.

Due to the natural appearance of the Fourier matrix $\mathbf{F}$ in the factorization of circulant matrices its rows are also the natural choice in the rows of the measurement matrix $\mathbf{A}$. Cancellations that occur because of this choice lead to the analytic results found. This shows a simple, but equivalent, alternative way to develop Equation (38) of [13].

3.4. The 1-to-N Shift Retrieval Problem

In the previous sections, we have assumed that the signals to be compared are singletons (we could call this the 1-to-1 shift retrieval problem). In this section, we explore what happens when we want to solve the shift retrieval problem between a signal $\mathbf{x}$ and a group of signals $\mathbf{Y} \in \mathbb{R}^{n \times N}$, i.e., find the shift for the signal $\mathbf{x}$ such that it aligns best with all N signals from $\mathbf{Y}$. Just as before, we can approach this problem as maximizing Equation (7) or like a minimization problem (Equation (10)).

In our case, the quantity in Equation (7) generalizes to

$$\arg \max_q \|\mathbf{Y}^T \mathbf{P}^{q-1} \mathbf{x}\|_1, \quad (45)$$

and this is equivalent to the approach:

$$\arg \max |\text{circ}(\mathbf{x})^T \mathbf{Y}|_{N \times 1}, \quad (46)$$

i.e., we take the index of the maximum entry of the $n \times 1$ argument vector. The argument we want is the index where the quantity $\|\text{circ}(\mathbf{x})^T \mathbf{Y}\|_\infty$ is achieved. This is the matrix ∞ -norm, i.e., $\|\mathbf{Z}\|_\infty = \max_i \sum_k |Z_{ik}|$. The next result provides the way to compute the optimum shift.

Result 4 (One-to-many shift retrieval). *We are given a signal $\mathbf{x}$ and a group of signals $\mathbf{Y}$, we aim to find the shift that achieves the highest correlation, in absolute value, between $\mathbf{x}$ and all the vectors $\mathbf{y}_i$ from $\mathbf{Y}$ in the sense of Equation (46). The shift q that maximizes this quantity is returned by*

$$\arg \max |\text{IFFT}(\text{diag}(\text{FFT}(\mathbf{x})^*) \text{FFT}(\mathbf{Y}))|_{N \times 1}, \quad (47)$$

i.e., we take the index of the maximum entry of the $n \times 1$ argument vector.

Proof. We use Equation (2) and expand the quantity in Equation (46),

$$\text{circ}(\mathbf{x})^T \mathbf{Y} = (\mathbf{F}^H \text{diag}(\mathbf{F}\mathbf{x})\mathbf{F})^H \mathbf{Y} = \mathbf{F}^H \text{diag}(\mathbf{F}\mathbf{x})^* \mathbf{F}\mathbf{Y}. \quad (48)$$

The matrix-vector product that follows computes the row-wise sums of the absolute value matrix. The computational complexity is dominated by $O(nN \log n)$ for the Fourier transforms and $O(nN)$ for the summations. $\square$

This result establishes the circular shift that, on average, aligns the data points as well as possible.

3.5. The N-to-N Shift Retrieval Problem

In the most general case of pairwise shifts, we are given two sets of signals $\mathbf{X} \in \mathbb{R}^{n \times N}$ and $\mathbf{Y} \in \mathbb{R}^{n \times N}$; the problem is to find a single shift such that each signal $\mathbf{x}_i$ aligns as best as possible with the corresponding signal $\mathbf{y}_i$. This can be seen as the generalization of the problem in the previous sections.

In this case, the quantity in Equation (45) further generalizes to

$$\arg \max_q \text{trace}(|\mathbf{Y}^T \mathbf{P}^{q-1} \mathbf{X}|), \quad (49)$$

We state the following result, as a generalization of Result 4.

Result 5 (Many-to-many shift retrieval). *We are given the signals $\mathbf{X}$ and $\mathbf{Y}$, we aim to find the shift that achieves the highest correlation, in absolute value, between all pairs $\mathbf{x}_i$ and $\mathbf{y}_i$ in the sense of Equation (49). The shift q that maximizes this quantity is returned by*

$$\arg \max |\text{IFFT}(\text{FFT}(\mathbf{X})^* \odot \text{FFT}(\mathbf{Y}))| \mathbf{1}_{N \times 1}, \quad (50)$$

i.e., we take the index of the maximum entry of the $n \times 1$ argument vector.

Proof. We use Equation (2), Result 4 and expand the quantity

$$\begin{aligned} \text{diag}(\mathbf{Y}^T \mathbf{P}^{q-1} \mathbf{X}) &= \text{diag}(\mathbf{Y}^T \mathbf{F}^H \text{diag}(\mathbf{F} \mathbf{e}_q) \mathbf{F} \mathbf{X}) \\ &= \text{diag}(\tilde{\mathbf{Y}}^H \text{diag}(\mathbf{f}_q) \tilde{\mathbf{X}}) \\ &= (\tilde{\mathbf{Y}}^H \odot \tilde{\mathbf{X}}^T) \mathbf{f}_q. \end{aligned} \quad (51)$$

The last equality leads to the expression in the result statement, as the trace is the sum of the diagonal vector entries. The matrix-vector product that follows computes the row-wise sums in absolute value. The computational complexity is dominated by $O(nN \log n)$ for the Fourier transforms and $O(nN)$ for the summations. $\square$

Notice how Results 4 and 5 are generalizations of the multiplicative cross-correlation formula from Equation (6). In these cases, when we do not expect alignment to be performed exactly, the division approach taken in Result 1 is not appropriate. In the context of these results, if indeed signals are circular shifts in each other, then $N = 1$ is enough to recover the true shift. Thus, these methods actually recover an average shift that maximally aligns the data point pairs.

3.6. Linear Combinations of a Known Circularly Shifted Signal

In all previous sections, our objective was to recover a single shift that maximally aligns data points, either 1-to-1, 1-to- N , or N -to- N . Now, we consider a scenario where a single signal is circularly shifted in multiple positions, and we take linear combinations of these. The task is to recover all the shifts performed and their weights from the minimum number of measurements. Consider the following result, a generalization of Result 1:

Result 6 (Recovery of linear combinations of circular shifts). *We are given a signal $\mathbf{x}$ and the measurement $\mathbf{y}$, which we assume is a linear combination of an unknown number of weighted circular shifts in $\mathbf{x}$ such that*

$$\mathbf{y} = \sum_{q=1}^n \alpha_q \mathbf{P}^{q-1} \mathbf{x}, \quad (52)$$

then, stacking the real-valued weights α_q in the vector $\boldsymbol{\alpha} \in \mathbb{R}^n$, and assuming $\tilde{x}_i \neq 0$ holds for all indices, we have that

$$\text{IFFT}(\text{FFT}(\mathbf{y}) \oslash \text{FFT}(\mathbf{x})) = \boldsymbol{\alpha}. \quad (53)$$

Proof. We start by solving the optimization problem

$$\underset{\boldsymbol{\alpha}}{\text{minimize}} \left\| \mathbf{y} - \sum_{q=1}^n \alpha_q \mathbf{P}^{q-1} \mathbf{x} \right\|_2^2. \quad (54)$$

Note that the optimization variables are the weights α_q , not the shifts. If a circular shift is missing in the linear combination, then the corresponding weight is zero. We develop the objective function value

$$\begin{aligned}
 \|\mathbf{y} - \sum_{q=1}^n \alpha_q \mathbf{P}^{q-1} \mathbf{x}\|_2^2 &= \|\mathbf{y} - \sum_{q=1}^n \alpha_q \mathbf{F}^H \text{diag}(\mathbf{f}_q) \mathbf{F} \mathbf{x}\|_2^2 \\
 &= \|\mathbf{F} \mathbf{y} - \sum_{q=1}^n \alpha_q \text{diag}(\mathbf{f}_q) \mathbf{F} \mathbf{x}\|_2^2 \\
 &= \|\tilde{\mathbf{y}} - \sum_{q=1}^n \alpha_q \text{diag}(\mathbf{f}_q) \tilde{\mathbf{x}}\|_2^2 \\
 &= \|\tilde{\mathbf{y}} - \text{diag}(\tilde{\mathbf{x}}) \sum_{q=1}^n \alpha_q \mathbf{f}_q\|_2^2 \\
 &= \|\tilde{\mathbf{y}} - \text{diag}(\tilde{\mathbf{x}}) \mathbf{F} \boldsymbol{\alpha}\|_2^2,
 \end{aligned} \tag{55}$$

here we have used that $\text{diag}(\mathbf{f}_q) \tilde{\mathbf{x}} = \mathbf{f}_q \odot \tilde{\mathbf{x}} = \text{diag}(\tilde{\mathbf{x}}) \mathbf{f}_q$. Assuming that Equation (52) holds, we have $\tilde{\mathbf{y}} = \text{diag}(\tilde{\mathbf{x}}) \mathbf{F} \boldsymbol{\alpha}$ and finally $\tilde{\mathbf{y}} \oslash \tilde{\mathbf{x}} = \mathbf{F} \boldsymbol{\alpha}$, to reach the desired result. $\square$

This result establishes that weighted linear combinations of a single known signal, which is circularly shifted, can be efficiently recovered from noiseless linear measurements. Note that we need not use all possible shifts, but only a subset—equivalent to having a sparse weight vector $\boldsymbol{\alpha}$.

A natural question might be what is the minimum number of measurements needed to recover the weights, and what happens when noise is added to the measurements? In general, we will need all n measurements $\tilde{\mathbf{y}} \oslash \tilde{\mathbf{x}}$, but when the weight vector $\boldsymbol{\alpha}$ is sparse, then well-known results from the signal processing literature provide better insights. First, note that for $\tilde{x}_i \neq 0$, the problem can be seen as a linear measurement problem of the type:

$$\tilde{\mathbf{y}} \oslash \tilde{\mathbf{x}} = \mathbf{F} \boldsymbol{\alpha}. \tag{56}$$

Assuming sparsity $s \in \mathbb{N}, 1 \leq s \ll n$ for $\boldsymbol{\alpha}$, this is now a standard problem in Compressed Sensing (CS) [23] where we ask how many Fourier measurements we need, from the total n available, in order to correctly recover the weights $\boldsymbol{\alpha}$. To understand this problem and its solution, we make use of the following well-established results from the literature:

- In the noiseless case, we know that in order to recover an exactly s -sparse vector $\boldsymbol{\alpha}$, we need at least $m = 2s$ consecutive Fourier measurements. This result is described in ([23], Theorem 2.15) via a Prony-type reconstruction procedure. Note that this is consistent with the findings of Remark 4 for $s = 1$, where it is established that two non-DC components are required;
- In the noisy measurements case, Prony-type methods cannot be used, as they are not robust against noise. Now, the stable recovery of an s -sparse $\boldsymbol{\alpha}$ of length n needs, with high probability, order $m \approx s \text{ polylog}(n)$ random Fourier measurements. The recovery of $\boldsymbol{\alpha}$ is performed via the ℓ_1 optimization problem Basis Pursuit Denoizing (BPD),

$$\underset{\boldsymbol{\alpha} \in \mathbb{R}^n}{\text{minimize}} \|\boldsymbol{\alpha}\|_1 \text{ subject to } \|(\tilde{\mathbf{y}} \oslash \tilde{\mathbf{x}})_{\mathcal{K}} - (\mathbf{F})_{\mathcal{K}} \boldsymbol{\alpha}\|_2 \leq \epsilon, \tag{57}$$

where positive $\epsilon \in \mathbb{R}$ is given and depends on the expected noise level. We have denoted here $(\mathbf{F})_{\mathcal{K}}$ the $m \times n$ sub-matrix of the $n \times n$ Fourier matrix consisting of all the columns and only the rows indexed in the set $\mathcal{K}$. For the technical details on this result, the reader can consult ([23], Chapter 11). In the case of $s = 1$, note that

the solution is computed by finding the largest absolute value correlation between the columns of $(\mathbf{F})_{\mathcal{K}}$ and $(\tilde{\mathbf{y}} \odot \tilde{\mathbf{x}})_{\mathcal{K}}$. Normalizing the columns of $(\mathbf{F})_{\mathcal{K}}$ and defining the mutual coherence denoted $0 \leq \mu((\mathbf{F})_{\mathcal{K}}) \leq 1$, the shift retrieval for $s = 1$ and the recovery of the single weight α is robust against noise whenever $|\alpha| \geq \frac{2\epsilon}{1-\mu((\mathbf{F})_{\mathcal{K}})}$. In general, $\mu((\mathbf{F})_{\mathcal{K}})$ decreases with increasing m [24].

When the signal $\mathbf{x}$ is unknown and we try to recover both the shifts and the signal itself, the problem is much more difficult, as it requires some alternating optimization strategy, in general. This is related to the circulant dictionary learning problem [16,17,25].

4. Experimental Results

Some results described in this paper are algebraic in nature, and therefore, beyond their proofs, simulation experiments do not bring any further significant insights. In this section, we check numerically Result 2 and the noisy variant of Result 6, as in these cases, estimation accuracy needs to be computed and insights verified empirically.

4.1. Shift Recovery from Multiple Noisy Measurements

In the first experimental setting, we want to validate the findings in Result 2, Remarks 5 and 6. We will generate random signals $\mathbf{x}$ of size n , circularly shift them by a uniformly random quantity q , and then try to recover them from the $m \leq n$ noisy measurements from $\mathbf{y}$, as described in Equation (25). We also want to validate the intuition and findings of Remark 5 by proposing two ways of selecting the Fourier measurements: uniformly at random and then such that the m chosen Fourier coefficients are from the first half of the spectrum (to avoid duplicates due to conjugation) and have maximum ℓ_2 norm, i.e., maximum sum squared magnitude. We expect the latter to perform much better in experiments. Performance is measured in two ways: the percentage of average correct shift recovery and as the Root Mean Squared Error (RMSE) between the true shifts and the estimates obtained, always modulo n .

In Figures 1 and 2, we show the experimental results for several n and as the number of Fourier measurements m increases. In Figure 1 we show the percentage of correct shift recovery by Result 2 over 10,000 realizations for SNR in the range $[-30, 60]$ dB. We recover the shift from signals of size n with an increasing number of Fourier measurements m for various SNR values. Largest magnitude entries are selected, as per Remark 5. As highlighted by Remark 6, notice how the performance of the circular convolutional approach improves as the length of the signals n increases, and also notice how the performance of the one-measurement approach degrades (top to bottom). For the bottom plot, where $n = 8192$, we expected near-perfect shift recovery from one measurement only around 65 dB SNR. Note that all other experiments, which take the number of measurements m to be a percentage of the signal size, also exhibit improved accuracy as n grows, from top to bottom. Also note that for $n = 8192$ we have almost perfect recovery by 0 dB SNR for $m = 25$ measurements (0.03% of n). For all n used in Figure 1, we have that $\varphi(n) = n/2$, since these are perfect powers of two, and therefore we have 50% of measurements that obey $\gcd(k_i, n) = 1$, i.e., the measurements with odd indices. Notice that when we approach 50% measurements, we reach the accuracy of the full circular convolution method. At the opposite end, for prime n we have $\varphi(n) = n - 1$, and therefore all indices would obey $\gcd(k_i, n) = 1$.

In Figure 2, we show RMSE for the shift recovery accuracy when dealing with signals of size $n \in \{300, 600, 900, 1200\}$ with an increasing number of Fourier measurements m . We show the two indices selection rules: uniformly at random and by largest magnitude, as per Remark 5. Results in all plots are averaged over 10,000 realizations. Observe that the index selection rule from Remark 5 significantly outperforms a random selection strategy and

quickly reaches a near-zero RMSE. Of course, in order to be able to select the top magnitude m entries in the spectrum, we have to compute the whole spectrum. This might not always be possible, and it comes with the cost of calculating the full Fourier transform at the cost of $O(n \log n)$, as opposed to $O(nm)$ when only m Fourier coefficients are needed.

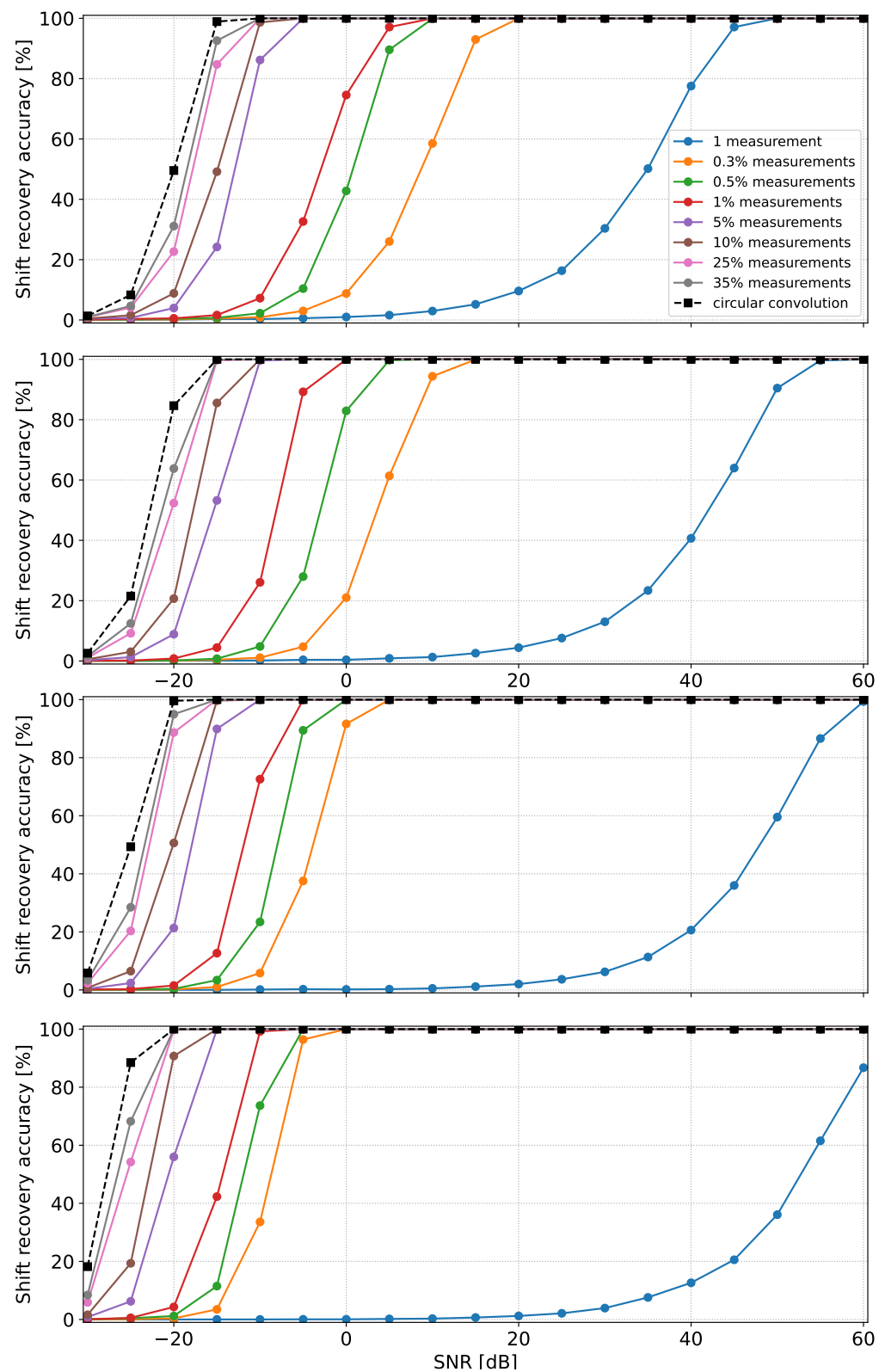


Figure 1. Average accuracy of correct shift recovery by Result 2. From top to bottom, the signal sizes are $n \in \{1024, 2048, 4096, 8192\}$.

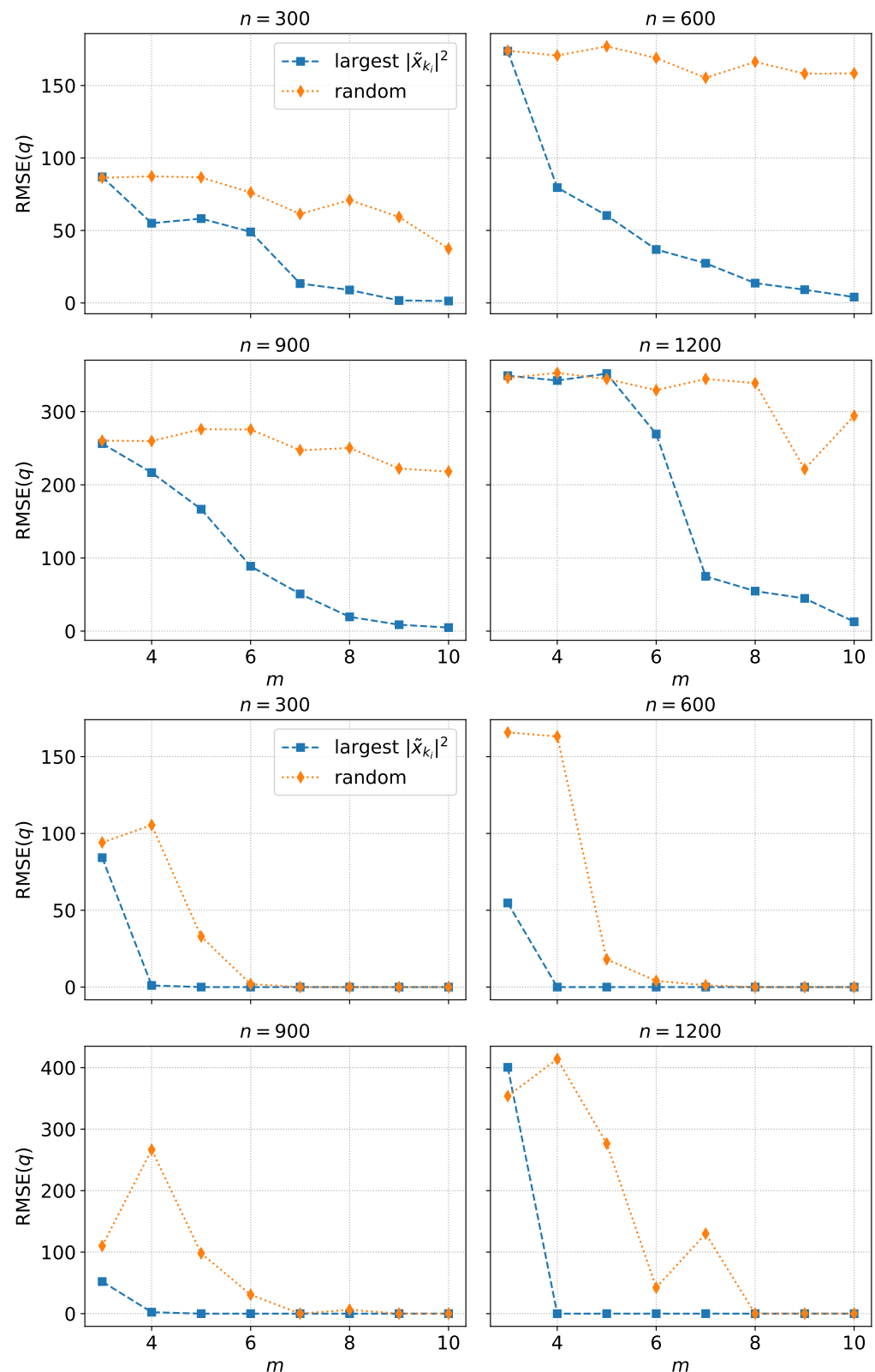


Figure 2. RMSE of recovered shift for increasing number of Fourier measurements m . The top four plots have SNR = 0 dB, while the bottom four plots have SNR = 15 dB.

4.2. Recovering Multiple Shifts via ℓ_1 Minimization

In the last experimental setup, we experimentally test the findings of Result 6. We generate a random signal $\mathbf{x}$ with entries from the standard Gaussian distribution. For a fixed sparsity level s , we generate the weight vector $\mathbf{a}$ whose support is generated uniformly

at random and whose values are from the standard Gaussian distribution. The shifts α are also generated uniformly at random in the feasible set. We acquire m noisy Fourier measurements, similarly to Equation (52),

$$\mathbf{y} = \sum_{q=1}^n \alpha_q \mathbf{P}^{q-1} \mathbf{x} + \mathbf{w}, \mathbf{w} \sim \mathcal{N}(\mathbf{0}_{n \times 1}, \zeta^2 \mathbf{I}_n). \quad (58)$$

Our goal is to recover the vector α , as its entries provide the weights, and the positions of the non-zero entries provide the shifts. To solve ℓ_1 optimization problems, we use the publicly available CVXPY library [26]. Recovery results, averaged over 100 realizations, are shown in Figure 3. For α , we show the RMSE, which expresses the accuracy of weight estimation, and the successful support recovery rate, which expresses the accuracy of shift retrieval. For the support recovery, we compute the positions of the s largest magnitude entries in the solutions to the ℓ_1 optimization problem, and we check the overlap with the true support. We report the recovery of support as a percentage. As expected, increasing the number of measurements leads to better performance, and of course, increasing the dimension n and the sparsity s degrades the recovery of both the weights and the support.

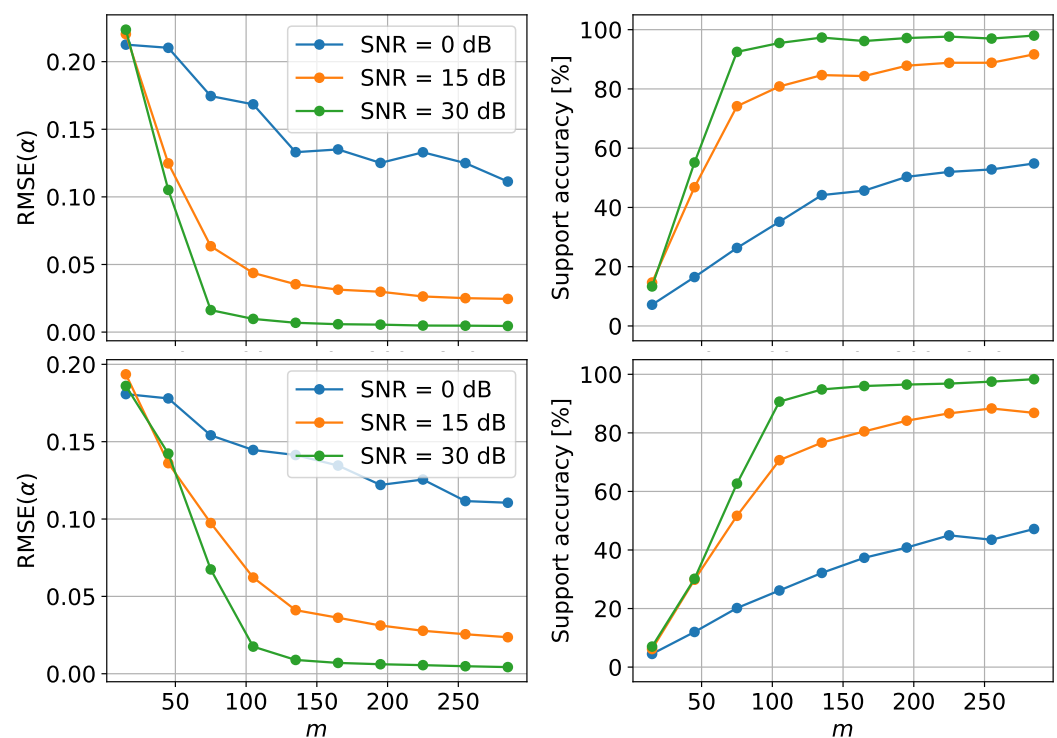


Figure 3. RMSE of recovered shifts and their weights according to Result 6. The top two plots have $n = 300$ and sparsity $s = 15$, and the bottom two plots have $n = 600$ and sparsity $s = 20$.

5. Conclusions

In this letter, we provide an overview of several shift retrieval problems based on optimization problems involving circulant matrices. We demonstrate that while the classic multiplicative cross-correlation method performs a perfectly adequate job in the shift retrieval problem, in many scenarios, the shift can be retrieved naturally from a few measurements by using a weighted correlation-like quantity. Our proposed approach also unifies several previously known results and methods under a single framework, providing natural generalizations. When appropriate, we successfully validate the algebraic results through numerical experimental simulations where the goal is to perform shift retrieval with as few measurements as possible.

Funding: This research was funded by Romanian Hub for Artificial Intelligence—HRIA, Smart Growth, Digitization and Financial Instruments Program, 2021–2027 (MySMIS no. 334906).

Data Availability Statement: The data presented in this study are openly available in GitHub at <https://github.com/cristian-rusu-research/shift-invariance> (accessed on 20 December 2025).

Conflicts of Interest: The author declare no conflicts of interest.

References

1. Carter, G.C. Time Delay Estimation. *Adaptive Methods in Underwater Acoustics*; Springer: Dordrecht, The Netherlands, 1985; pp. 175–196. https://doi.org/10.1007/978-94-009-5361-1_15.
2. Cao, H.; Chan, Y.T.; So, H.C. Compressive TDOA Estimation: Cramer-Rao Bound and Incoherent Processing. *IEEE Trans. Aerosp. Electron. Syst.* **2020**, *56*, 3326–3331. <https://doi.org/10.1109/TAES.2020.2966095>.
3. Tong, X.; Ye, Z.; Xu, Y.; Gao, S.; Xie, H.; Du, Q.; Liu, S.; Xu, X.; Liu, S.; Luan, K.; et al. Image Registration With Fourier-Based Image Correlation: A Comprehensive Review of Developments and Applications. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2019**, *12*, 4062–4081. <https://doi.org/10.1109/JSTARS.2019.2937690>.
4. Peng, Y.; Ganesh, A.; Wright, J.; Xu, W.; Ma, Y. RASL: Robust alignment by sparse and low-rank decomposition for linearly correlated images. In Proceedings of the 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, San Francisco, CA, USA, 13–18 June 2010; pp. 763–770. <https://doi.org/10.1109/CVPR.2010.5540138>.
5. Chandrakasan, A.P.; Lee, F.S.; Wentzloff, D.D.; Sze, V.; Ginsburg, B.P.; Mercier, P.P.; Daly, D.C.; Blazquez, R. Low-Power Impulse UWB Architectures and Circuits. *Proc. IEEE* **2009**, *97*, 332–352. <https://doi.org/10.1109/JPROC.2008.2008787>.
6. Amiri, R.; Behnia, F.; Noroozi, A. An Efficient Estimator for TDOA-Based Source Localization With Minimum Number of Sensors. *IEEE Commun. Lett.* **2018**, *22*, 2499–2502. <https://doi.org/10.1109/LCOMM.2018.2876525>.
7. Spiesberger, J.L. Finding the right cross-correlation peak for locating sounds in multipath environments with a fourth-moment function. *J. Acoust. Soc. Am.* **2000**, *108*, 1349–1352. <https://doi.org/10.1121/1.1287709>.
8. Nishiura, T.; Yamada, T.; Nakamura, S.; Shikano, K. Localization of multiple sound sources based on a CSP analysis with a microphone array. In Proceedings of the 2000 IEEE International Conference on Acoustics, Speech, and Signal Processing. Proceedings (Cat. No.00CH37100), Istanbul, Turkey, 5–9 June 2000; Volume 2, pp. II1053–II1056. <https://doi.org/10.1109/ICASSP.2000.859144>.
9. Huber, M.; Schlegel, M.; Klinker, G. Application of time-delay estimation to mixed reality multisensor tracking. *J. Virtual Real. Broadcast.* **2014**, *11*. <https://doi.org/10.20385/1860-2037/11.2014.3>.
10. Liu, Y.; Wang, C.; Lu, M.; Yang, J.; Gui, J.; Zhang, S. From Simple to Complex Scenes: Learning Robust Feature Representations for Accurate Human Parsing. *IEEE Trans. Pattern Anal. Mach. Intell.* **2024**, *46*, 5449–5462. <https://doi.org/10.1109/TPAMI.2024.3366769>.
11. Wang, C.; Zhang, Q.; Wang, X.; Zhou, L.; Li, Q.; Xia, Z.; Ma, B.; Shi, Y.Q. Light-Field Image Multiple Reversible Robust Watermarking Against Geometric Attacks. *IEEE Trans. Dependable Secur. Comput.* **2025**, *22*, 5861–5875. <https://doi.org/10.1109/TDSC.2025.3576223>.
12. Kelishadrokh, M.K.; Ghattaei, M.; Fekri-Ershad, S. Innovative local texture descriptor in joint of human-based color features for content-based image retrieval. *Signal Image Video Process.* **2023**, *17*, 4009–4017. <https://doi.org/10.1007/s11760-023-02631-x>.
13. Ohlsson, H.; Eldar, Y.C.; Yang, A.Y.; Sastry, S.S. Compressive Shift Retrieval. *IEEE Trans. Signal Process.* **2014**, *62*, 4105–4113. <https://doi.org/10.1109/TSP.2014.2332974>.
14. Ohlsson, H.; Eldar, Y.C.; Yang, A.Y.; Sastry, S.S. Compressive shift retrieval. In Proceedings of the 2013 IEEE International Conference on Acoustics, Speech and Signal Processing, Vancouver, BC, Canada, 26–31 May 2013; pp. 6034–6038. <https://doi.org/10.1109/ICASSP.2013.6638823>.
15. Clausen, M.; Kurth, F. Robust compressive shift retrieval in linear time. In Proceedings of the 2016 24th European Signal Processing Conference (EUSIPCO), Budapest, Hungary, 28 August–2 September 2016; IEEE: New York, NY, USA, 2016; pp. 364–368.
16. Rusu, C.; Dumitrescu, B.; Tsiftaris, S.A. Explicit Shift-Invariant Dictionary Learning. *IEEE Signal Process. Lett.* **2014**, *21*, 6–9. <https://doi.org/10.1109/LSP.2013.2288788>.
17. Rusu, C. On learning with shift-invariant structures. *Digit. Signal Process.* **2020**, *99*, 102654. <https://doi.org/10.1016/j.dsp.2019.102654>.
18. Oppenheim, A.V.; Schaffer, R.W.; Buck, J.R. *Discrete-Time Signal Processing*, 2nd ed.; Prentice Hall: Englewood Cliffs, NJ, USA, 1999.
19. Goertzel, G. An Algorithm for the Evaluation of Finite Trigonometric Series. *Am. Math. Mon.* **1958**, *65*, 34–35.
20. Frigo, M.; Johnson, S.G. The Design and Implementation of FFTW3. *Proc. IEEE* **2005**, *93*, 216–231.
21. Hassanieh, H.; Indyk, P.; Katabi, D.; Price, E. Nearly optimal sparse Fourier transform. In Proceedings of the Forty-Fourth Annual ACM Symposium on Theory of Computing (STOC '12), New York, NY, USA, 19–22 May 2012; pp. 563–578. <https://doi.org/10.1145/2213977.2214029>.

22. Hassanieh, H.; Indyk, P.; Katabi, D.; Price, E. Simple and practical algorithm for sparse Fourier transform. In Proceedings of the Twenty-Third Annual ACM-SIAM Symposium on Discrete Algorithms (SODA '12), Kyoto, Japan, 17–19 January 2012; pp. 1183–1194.
23. Foucart, S.; Rauhut, H. *A Mathematical Introduction to Compressive Sensing*; Applied and Numerical Harmonic Analysis; Birkhäuser: Basel, Switzerland, 2013; pp. I–XVIII, 1–625.
24. Rusu, C.; González-Prelcic, N.; Heath, R.W. Algorithms for the construction of incoherent frames under various design constraints. *Signal Process.* **2018**, *152*, 363–372. <https://doi.org/https://doi.org/10.1016/j.sigpro.2018.06.015>.
25. Pope, G.; Aubel, C.; Studer, C. Learning phase-invariant dictionaries. In Proceedings of the 2013 IEEE International Conference on Acoustics, Speech and Signal Processing, Vancouver, BC, Canada, 26–31 May 2013; pp. 5979–5983. <https://doi.org/10.1109/ICASSP.2013.6638812>.
26. Diamond, S.; Boyd, S. CVXPY: A Python-embedded modeling language for convex optimization. *J. Mach. Learn. Res.* **2016**, *17*, 2909–2913.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.