Einstein from Noise: Statistical Analysis

Amnon Balanov*[1], Wasim Huleihel[1], and Tamir Bendory[1]

[1]School of Electrical and Computer Engineering, Tel Aviv University, Tel Aviv 69978, Israel

March 10, 2026

Abstract

"Einstein from noise" (EfN) is a prominent example of the model bias phenomenon, where systematic errors in the statistical model lead to spurious but consistent estimates. In the EfN experiment, one falsely believes that a set of observations contains noisy, shifted copies of a template signal (e.g., an Einstein image), whereas in reality, it contains only pure noise observations. To estimate the signal, the observations are first aligned with the template using cross-correlation and then averaged. Although the observations contain nothing but noise, it was recognized early on that this process produces a signal that resembles the template signal! This model bias pitfall was at the heart of a central scientific controversy about validation techniques in structural biology.

This paper provides a comprehensive statistical analysis of the EfN phenomenon above. We show that the Fourier phases of the EfN estimator (namely, the average of the aligned noise observations) converge to the Fourier phases of the template signal, thereby explaining the observed structural similarity. Additionally, we prove that the convergence rate of Fourier phases is inversely proportional to the number of noise observations and, in the high-dimensional regime, to the Fourier magnitudes of the template signal. Moreover, in the high-dimensional regime, the EfN estimator converges to a scaled version of the template signal. This work not only deepens the theoretical understanding of the EfN phenomenon but also highlights potential pitfalls in template matching techniques and emphasizes the need for careful interpretation of noisy observations across disciplines in engineering, statistics, physics, and biology.

*Corresponding author: amnonba15@gmail.com

Contents

1 Introduction **4**

2 Problem Formulation and Notation **6**

3 Cryo-EM and Empirical Demonstration **8**

4 Main Results **11**
4.1 Finite-dimensional signal . . . 11
4.2 High-dimensional regime . . . 12

5 Extension to other noise statistics **15**
5.1 Positive correlation . . . 16
5.2 High-dimensional i.i.d. noise . . . 16
5.3 Circulant Gaussian process . . . 17

6 Discussion and outlook **19**
6.1 Extensions and implications . . . 19
6.2 Future work . . . 21

Appendices **28**

Appendix A Preliminaries **28**
A.1 Notations . . . 28
A.2 The convergence of the Einstein from Noise estimator . . . 29
A.3 Conditioning on the Fourier frequency noise component . . . 30
A.4 Uniqueness of the maximizer . . . 33
A.5 Positive probability of each maximizer event . . . 34
A.6 Auxiliary result for Proposition B.2 . . . 36

Appendix B Proof of Theorem 4.1 **41**
B.1 Convergence of the Fourier phases . . . 42
B.2 Convergence to non-vanishing signal . . . 43
B.3 Convergence rate in distribution of the Fourier phases . . . 44
B.4 Convergence rate in expectation of the Fourier phases . . . 45
B.5 Proof of Theorem 4.1 . . . 49

Appendix C High-dimensional argmax asymptotics **50**
C.1 Proof of Lemma C.2 . . . 51
C.2 Proof of Proposition C.1 . . . 53

Appendix D Proof of Theorem 4.3 **56**
D.1 Notations and auxiliary results . . . 56
D.2 High-dimensional limits . . . 58
D.3 Proof of Theorem 4.3 . . . 62

Appendix E Proof of Proposition 5.1 **63**

Appendix F Proof of Theorem 5.2: High-dimensional i.i.d. noise **65**
F.1 The functional CLT for DFT 66
F.2 Notations 66
F.3 Convergence of the real and imaginary parts of the EfN estimator 67
F.4 Proof of Theorem 5.2 69

Appendix G Proof of Proposition 5.4: Circulant Gaussian noise **71**
G.1 Preliminaries 71
G.2 The convergence of the Einstein from Noise estimator 72
G.3 Conditioning on the Fourier frequency noise component 73
G.4 Convergence of the Fourier phases 76
G.5 Convergence to non-vanishing signal 77
G.6 Proof of Proposition 5.4 78

1 Introduction

Model bias is a fundamental pitfall arising across a broad range of statistical problems, leading to consistent but inaccurate estimations due to systematic errors in the model. This paper focuses on the *Einstein from Noise* (EfN) experiment: a prototype example of model bias that appears in template matching techniques. Consider a scenario where scientists acquire observational data and genuinely believe their observations contain noisy, shifted copies of a known template signal. However, in reality, their data consists of pure noise with no actual signal present.

To estimate the (absent) signal, the scientists align each observation by cross-correlating it with the template and then average the aligned observations. Remarkably, empirical evidence has shown, multiple times, that the reconstructed structure from this process is structurally similar to the template, even when all the measurements are pure noise [23, 44, 46]. This phenomenon stands in striking contrast to the prediction of the unbiased model, that averaging pure noise signals would converge towards a signal of zeros, as the number of noisy observations diverges. Thus, the above EfN estimation procedure is biased towards the template signal.

While the EfN phenomenon has been analyzed in prior work (see Section 3 for more details), a comprehensive theoretical understanding of the EfN model remains limited. This work contributes to filling that gap by rigorously analyzing the relationship between the reconstructed signal and the underlying template. The term 'Einstein from Noise' was popularized in [44], where the authors illustrated the phenomenon using an image of Einstein as the template signal. However, the underlying effect had been observed earlier in the cryogenic electron microscope literature (see, for instance, [46], and further details in Section 3). In this work, we refer to the average of the aligned pure noise observations as the *EfN estimator*. A detailed formulation of the problem is provided in Section 2.

Figure 1 illustrates the EfN process, which consists of two key stages. First, the observations are aligned with the template signal to achieve optimal alignment. Then, the aligned observations are averaged. The result is the EfN estimator, which shares a structural resemblance to the template image, though it is not an identical reproduction.

Main results. The central results of this work are as follows. Our first result, stated in Theorem 4.1, shows that the Fourier phases of the EfN estimator converge to the Fourier phases of the template signal, as the number of noisy observations (denoted by M) converges to infinity. However, it is important to note that the EfN estimator's *Fourier magnitudes* do not necessarily converge to those of the template signal. We also show that the Fourier phases' mean squared error (MSE) decays to zero with a rate of $1/M$. Since the Fourier phases are responsible for the formation of geometrical image elements, such as contours and edges [36, 45], this clarifies why the resulting EfN estimator image exhibits a structural similarity to the template, but not necessarily a full recovery. Our second result, stated in Theorem 4.3, proves that in the high-dimensional regime, where the dimension of the signal diverges, the convergence rate of the Fourier phases is inversely proportional to the square of the Fourier magnitudes of the template signal. In this case, the Fourier magnitudes of the EfN estimator converge to a scaled version of the template's Fourier magnitudes.

While Theorems 4.1 and 4.3 are proved under the assumption of white Gaussian noise, we

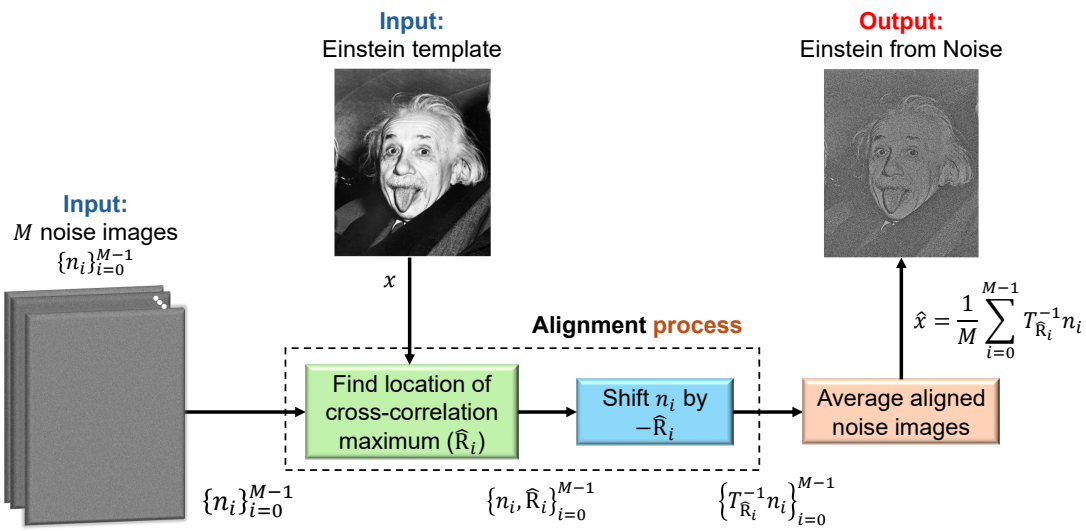

Figure 1: Einstein from Noise. The EfN estimator consists of three stages: (1) finding the index of the maximum of the cross-correlation ($\hat{\mathsf{R}}_i$) between the i-th noise signal (n_i) and the template signal (e.g., Einstein's image); (2) cyclically shifting the noise signal by $-\hat{\mathsf{R}}_i$; (3) averaging the shifted noise signals. In this paper, we characterize the relationship between the output of this process—*the EfN estimator*—and the template signal.

also extend our analysis to more general noise models. In particular, we show that, although the convergence results in Theorems 4.1 and 4.3 do not necessarily hold under arbitrary noise statistics, several structural properties of the EfN estimator persist. First, in Proposition 5.1, we show that the EfN estimator remains positively correlated with the template for arbitrary noise statistics, even when the Fourier phases do not converge. Since the correlation between images often implies visual resemblance, this explains why the EfN estimator still exhibits structural similarity to the template. Second, in Theorem 5.2, we show that in the high-dimensional limit, if the noise signal is independent and identically distributed (i.i.d) (not necessarily Gaussian), then the same phase convergence behavior observed in the white Gaussian case still holds. Finally, in Proposition 5.4, we demonstrate that if the noise signal is Gaussian with circular symmetry, then the conclusions of Theorem 4.1 remain valid, even though the noise is not white.

Organization. The remainder of this paper is organized as follows. Section 2 provides a detailed formulation of the problem. Section 3 discusses the connection between the EfN problem and single-particle cryo-electron microscopy (cryo-EM), the primary motivation for this work, and presents supporting empirical demonstrations. Our main theoretical results for white Gaussian noise observations, Theorems 4.1 and 4.3, are stated in Section 4. Extensions of these results to noise models beyond white Gaussian noise are presented in Section 5. Finally, we conclude with a discussion and outlook in Section 6.

2 Problem Formulation and Notation

This section outlines the probabilistic model behind the EfN experiment and delineates our main mathematical objectives. Although the EfN phenomenon is described typically for images, we will formulate and analyze it for one-dimensional signals, bearing in mind that the extension to two-dimensional images is straightforward (see Section 6 for more details).

Notations. Throughout the rest of this paper, we use $\xrightarrow{\mathcal{D}}$, $\xrightarrow{\mathcal{P}}$, $\xrightarrow{\text{a.s.}}$, and $\xrightarrow{\mathcal{L}^p}$, to denote the convergence of sequences of random variables in distribution, in probability, almost surely, and in $\mathcal{L}^p$ norm, respectively. Inner products in the Euclidean space between vectors a and b are written as either $a^\top b$ or $\langle a, b\rangle$.

Problem formulation. Consider a scenario where scientists collect a series of observations under the belief that each observation is a noisy, randomly shifted version of a known *template signal* $x \in \mathbb{R}^d$ (for example, an image of Einstein). Formally, the assumed postulated data model is given by:

$$\text{(Postulated model)} \quad y_i = \mathcal{T}_{\ell_i} \cdot x + n_i, \tag{2.1}$$

where $\mathcal{T}_\ell : \mathbb{R}^d \to \mathbb{R}^d$ is the cyclic shift operator defined by $[\mathcal{T}_\ell z]_r \triangleq z_{(r-\ell) \bmod d}$ for all $z \in \mathbb{R}^d$ and indices $0 \le r \le d-1$, and $n_i \sim \mathcal{N}(0, \sigma^2 I_{d\times d})$ are i.i.d. Gaussian noise vectors. In reality, however, there is no underlying signal: the observations consist entirely of white Gaussian noise. That is, the true data-generating process follows the underlying model:

$$\text{(Underlying model)} \quad y_0, y_1, \ldots, y_{M-1} \overset{\text{i.i.d.}}{\sim} \mathcal{N}(\mathbf{0}, \sigma^2 I_{d\times d}), \tag{2.2}$$

where M denotes the number of observations. Since the data consists purely of white Gaussian noise, we will explicitly write $y_i = n_i$ to emphasize this fact.

To estimate the (nonexistent) signal, the scientists align each observation to the template x using cross-correlation, and then average the aligned observations. Specifically, for each $i = 0, \ldots, M-1$, they compute the shift that maximizes the inner product with the template:

$$\hat{\mathsf{R}}_i \triangleq \operatorname*{arg\,max}_{0\le \ell < d} \langle n_i, \mathcal{T}_\ell x\rangle, \tag{2.3}$$

where n_i is the i-th noise observation, and $\hat{\mathsf{R}}_i$ defines the optimal cyclic shift that aligns the template signal x with the noise observation n_i in terms of cross-correlation.

Then, the EfN estimator is given by the average of the noise observations, but each is first aligned according to the above maximal shifts, i.e.,

$$\hat{x} \triangleq \frac{1}{M} \sum_{i=0}^{M-1} \mathcal{T}_{-\hat{\mathsf{R}}_i} n_i, \tag{2.4}$$

where $\mathcal{T}_{-\hat{\mathsf{R}}_i} n_i$ represents the noise observation n_i aligned by applying the inverse cyclic shift $-\hat{\mathsf{R}}_i$ to best match the template signal. Throughout the text, we refer to $\hat{x}$ as the EfN estimator.

The EfN phenomenon states that, at least empirically, $\hat{x}$ and x appear "close" in some sense; our goal is to understand this phenomenon mathematically. To that end, we will

consider the two asymptotic regimes: the first corresponds to the classical setting where the number of observations $M \to \infty$ while the dimension d is fixed (i.e., the number of observations diverges to infinity and the template vector dimension is fixed); the second is the high-dimensional regime, where $d \to \infty$ after $M \to \infty$. (i.e., both the number of observations and the dimension of the template signal diverge).

Fourier space notation. As will become clear in the next sections, it is convenient to work in the Fourier domain. Let $\phi_{\mathsf{Z}} \triangleq \sphericalangle \mathsf{Z}$ denote the phase of a complex number $\mathsf{Z} \in \mathbb{C}$, and recall that the discrete Fourier transform (DFT) of a d-length signal $y \in \mathbb{R}^d$ is given by,

$$\mathsf{Y}[k] \triangleq \mathcal{F}\{y\} = \frac{1}{\sqrt{d}} \sum_{\ell=0}^{d-1} y_\ell e^{-j\frac{2\pi}{d}k\ell}, \tag{2.5}$$

where $j \triangleq \sqrt{-1}$, and $0 \leq k \leq d-1$. Accordingly, we let X, $\hat{\mathsf{X}}$, and N_i, denote the DFTs of x, $\hat{x}$, and n_i, respectively, for $0 \leq i \leq M-1$. These DFT sequences can be equivalently represented in the magnitude-phase domain as follows,

$$\mathsf{X} = \{|\mathsf{X}[k]|\, e^{j\phi_{\mathsf{X}}[k]}\}_{k=0}^{d-1}, \quad \hat{\mathsf{X}} = \{|\hat{\mathsf{X}}[k]| e^{j\phi_{\hat{\mathsf{X}}}[k]}\}_{k=0}^{d-1}, \quad \mathsf{N}_i = \{|\mathsf{N}_i[k]|\, e^{j\phi_{\mathsf{N}_i}[k]}\}_{k=0}^{d-1}, \tag{2.6}$$

for $0 \leq i \leq M-1$, where $|\mathsf{X}[k]|$, $\left|\hat{\mathsf{X}}[k]\right|$, and $|\mathsf{N}_i[k]|$ are the k-th Fourier component magnitudes of the template signal, the EfN estimator, and the i-th noise observation, respectively. Similarly, $\phi_{\mathsf{X}}[k]$, $\phi_{\hat{\mathsf{X}}}[k]$, and $\phi_{\mathsf{N}_i}[k]$ represent the corresponding k-th Fourier phases. Note that the random variables $\{|\mathsf{N}_i[k]|\}_{k=0}^{d/2}$ and $\{\phi_{\mathsf{N}_i}[k]\}_{k=0}^{d/2}$ are two independent sequences of i.i.d. random variables, such that, $|\mathsf{N}_i[k]| \sim \mathsf{Rayleigh}\,(\sigma^2)$ has Rayleigh distribution, and the phase $\phi_{\mathsf{N}_i}[k] \sim \mathsf{Unif}[-\pi, \pi)$ is uniformly distributed over $[-\pi, \pi)$.

With the definitions above, we can express the estimation process in the Fourier domain. Since a shift in real-space corresponds to a linear phase shift in the Fourier space, it follows that,

$$\hat{\mathsf{X}}[k] = \frac{1}{M} \sum_{i=0}^{M-1} |\mathsf{N}_i[k]|\, e^{j\phi_{\mathsf{N}_i}[k]} e^{j\frac{2\pi k}{d}\hat{\mathsf{R}}_i}, \tag{2.7}$$

for $k = 0, 1, \ldots, d-1$, where $|\mathsf{N}_i[k]|$ and $\phi_{\mathsf{N}_i}[k]$ are defined in (2.6). It is important to note that the expression above will converge to zero without the last term that captures the dependency in $\hat{\mathsf{R}}_i$—the location of the maximum correlation. This term reflects the fundamental properties of the EfN process and its dependency on the template signal, as well as the connections between the different spectral components. We denote by $\mathbb{E}|\phi_{\hat{\mathsf{X}}}[k] - \phi_{\mathsf{X}}[k]|^2$ the MSE of the Fourier phases of the k-th spectral component.

Assumptions. Throughout this paper, we assume that the template signal x is normalized, i.e., $\|x\|_2^2 = 1$, where $\|\cdot\|_2$ is the Euclidean norm, and further assume that its Fourier transform in non-vanishing, except possibly at the DC (zero-frequency) component. The first assumption is used for convenience and does not alter (up to a normalization factor) our main results in Theorems 4.1 and 4.3. The second assumption is essential for the theoretical analysis of the EfN process and is expected to hold in many applications, including cryo-EM. A similar assumption is frequently taken in related work, e.g., [8, 38, 10]. It is

worth noting that since the Fourier transform of x is assumed to be non-vanishing, the maximizing shift $\hat{\mathsf{R}}_i$ in (2.3) is almost surely unique. In addition, without loss of generality, we assume that the signal length d is even.

3 Cryo-EM and Empirical Demonstration

Cryo-EM is a powerful tool of modern structural biology, offering advanced methods to visualize complex biological macromolecules with ever-increasing precision. One of its central advantages lies in its capability to resolve the structures of proteins that are hard to crystallize in traditional methods, especially in a near-physiological environment, see e.g., [35, 48]. This advantage enables researchers to delve into the dynamic behaviors of proteins and their complexes, shedding light on fundamental biological processes.

Single-particle cryo-EM uses electron microscopy to reconstruct 3D structures from 2D tomographic projection images [9]. Typically, the 3D reconstruction involves two main steps: detecting and extracting single particle images using a particle picking algorithm, [42, 22, 11, 21], and then reconstructing the 3D density map [41, 40]. Most detection algorithms use template-matching techniques, which can introduce bias if improper templates are chosen, especially in low signal-to-noise ratio (SNR) conditions, which is the standard scenario in cryo-EM.

The EfN controversy. A publication of the 3D structure of an HIV molecule in PNAS in 2013 [33] initiated a fundamental controversy about validation techniques within the cryo-EM community, published as four follow-up PNAS publications [23, 54, 52, 32]. The EfN pitfall played a central role in this discussion. The primary question of the discussion was whether the collected datasets contained informative biological data or merely pure noise images. The core of the debate emphasized the importance of exercising caution and implementing cross-validation techniques when fitting data to a predefined model. This precautionary approach aims to mitigate the risk of erroneous fittings, which could ultimately lead to inaccuracies in 3D density map reconstruction. Model bias is still a fundamental problem in cryo-EM, as highlighted by an ongoing debate concerning validation tools, see for example, [51, 44, 24, 16, 17, 25, 27, 50].

Empirical demonstration. As introduced in the previous section, the EfN phenomenon depends on several key parameters: (1) the number of observations, denoted by M; (2) the dimension of the signal, denoted as d (for example, the number of pixels in Einstein's image); and (3) the statistical properties of the template signal, and in particular its power spectral density (PSD). To demonstrate the dependency on these parameters and provide insight into our main results, Figures 2 and 3 show the convergence of the EfN estimator. Both figures were generated according to the procedure outlined in Section 2. When referring to Monte Carlo trials, it means that the EfN estimator procedure, as specified in (2.4), was executed multiple times (the number of Monte-Carlo trials), each trial with fresh data.

Figure 2 shows the EfN estimator as a function of M. Figure 2(a) shows that as M increases, the EfN estimator becomes more structurally similar to the template Einstein image. Indeed, Figure 2(b) shows that as the number of observations M grows, the MSE between the Fourier phases of the template image and the corresponding Fourier phases of

EfN estimator decreases. Figure 2(c) highlights that the convergence rate is proportional to $1/M$, with a faster convergence rate for stronger spectral components.

Figure 3 illustrates the impact of the template signal's PSD on the cross-correlation between the template signal and the EfN estimator. Notably, a flatter PSD (i.e., a faster decay of the auto-correlation) leads to a higher correlation between the template and the estimator signals. These empirical results are proved theoretically in Theorems 4.1 and 4.3.

More applications. The EfN phenomenon extends to various applications employing template matching, whether through a feature-based or direct template-based approach. For instance, template matching holds significance in computational anatomy, where it aids in discovering unknown diffeomorphism to align a template image with a target image [15]. Other areas include medical imaging processing [1], manufacturing quality control [3], and navigation systems for mobile robots [28]. This pitfall may also arise in the feature-based approach, which relies on extracting image features like shapes, textures, and colors to match a target image by neural networks and deep-learning classifiers [59, 34, 53, 30].

Previous work. The EfN phenomenon has been investigated in earlier studies. In particular, it was shown that the ratios between the expected values of the Fourier coefficients of the EfN estimator and those of the template are real-valued [60, Chapter 5]. In this work, we build upon and significantly extend these results. Specifically, we establish the convergence of the EfN estimator to a non-vanishing signal, derive its convergence rate, analyze its behavior in the high-dimensional regime, and generalize the analysis to encompass a broader class of noise models beyond white Gaussian noise.

A closely related work is that of Wang et al. [55], who conducted a rigorous statistical analysis of model bias in a different but complementary setting. They analyze the effects of selectively averaging only samples that exhibit the highest cross-correlation with a fixed reference signal (e.g., Einstein's image). This selection mechanism introduces a bias toward the reference, and their analysis reveals a phase transition in the resulting reconstruction, governed by the number of samples, the signal dimension, and the size of the selected subset. Notably, their results show that a structured image can emerge even when averaging purely noisy data. A related but statistically distinct selection-based mechanism is studied in [7] in the context of template matching for particle picking and extraction: candidate (pure-noise) observations are filtered by thresholding their cross-correlation with a bank of templates, and downstream averages inherit a strong template imprint.

In contrast, our work investigates the EfN estimator in the absence of any selection step: all pure-noise observations are first aligned to a fixed template and then averaged. This difference in mechanism leads to different statistical behavior. Whereas template-based selection can yield averages that closely resemble the templates up to a global scale factor, we show that the non-selective EfN procedure exhibits a more structured form of alignment-induced bias, manifested most prominently through phase locking: the Fourier phases of the estimator converge to those of the template, even though the resulting average is generally not identical to the template. Closely related alignment-induced artifacts and their connection to Fourier-phase behavior have also been discussed in the multireference alignment literature; see, e.g., [43].

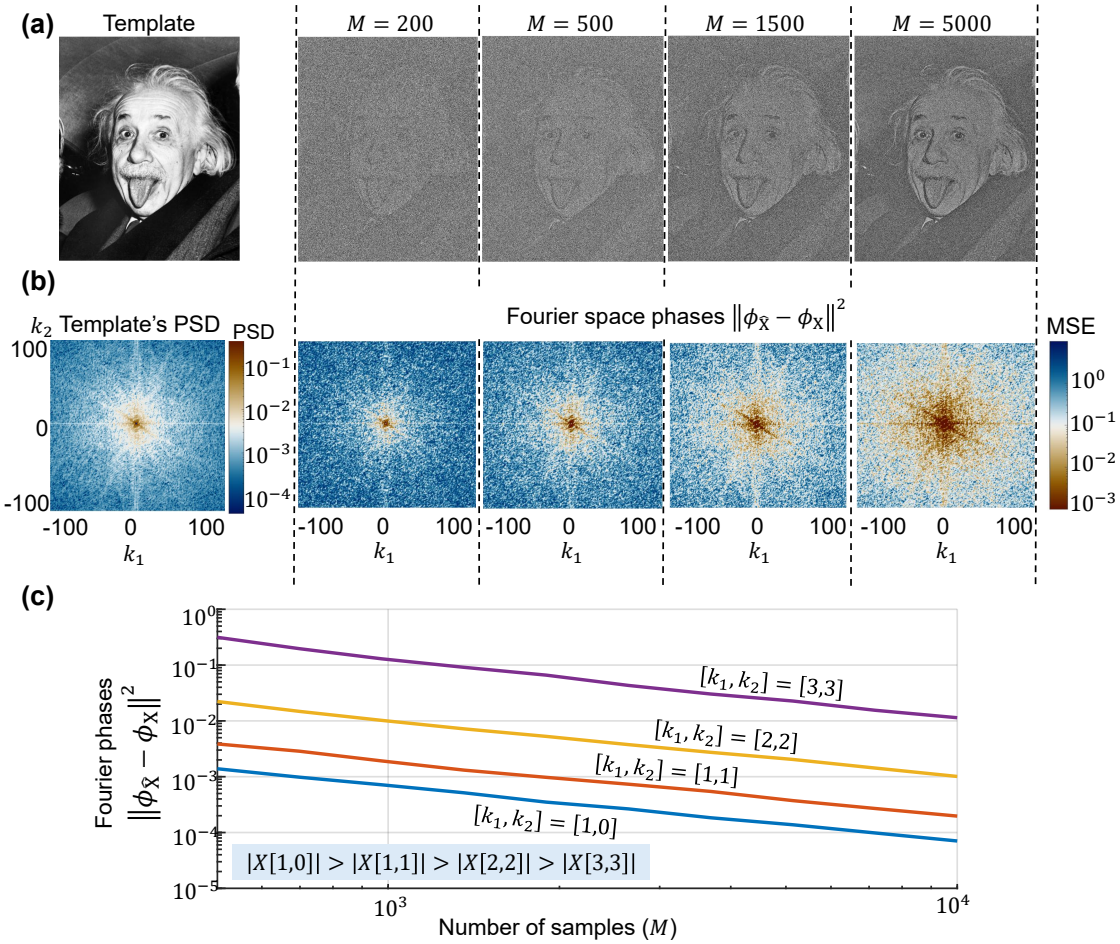

Figure 2: The impact of the number of noise observations on the EfN estimator. The EfN estimator is defined in real space by (2.4) and in Fourier space by (2.7). **(a)** The structural similarity between the EfN estimator and the template image increases as a function of the number of noise observations (M). **(b)** The mean-square-error (MSE) between the Fourier phases of the template image $\mathsf{X}[k_1, k_2]$ and the EfN estimator $\hat{\mathsf{X}}[k_1, k_2]$ for $-100 \leq k_1, k_2 \leq 100$, where k_1, k_2 are the indices of the 2D DFT. The colors in the left panel in (b) represent the power spectral density (PSD) of the Einstein image, while the colors in the four right panels represent the MSE between the Fourier phases of the Einstein image and the EfN estimator, for each spectral component, with a varying number of observations ($M = 200, 500, 1500, 5000$). An increase in the number of observations leads to a lower MSE of the Fourier phases between the EfN estimator and the template signal. A similar trend can be seen with respect to the strength of the spectral components, i.e., stronger spectral components lead to lower Fourier phases MSE. **(c)** The convergence rate of the MSE between the Fourier phases of the EfN estimator and the Fourier phases of the template signal as a function of the number of observations across different frequencies. The MSE decays as $1/M$. In addition, stronger spectral components lead to lower MSE. Figures **(b)** and **(c)** were generated through 200 Monte-Carlo trials of the EfN process defined in (2.4).

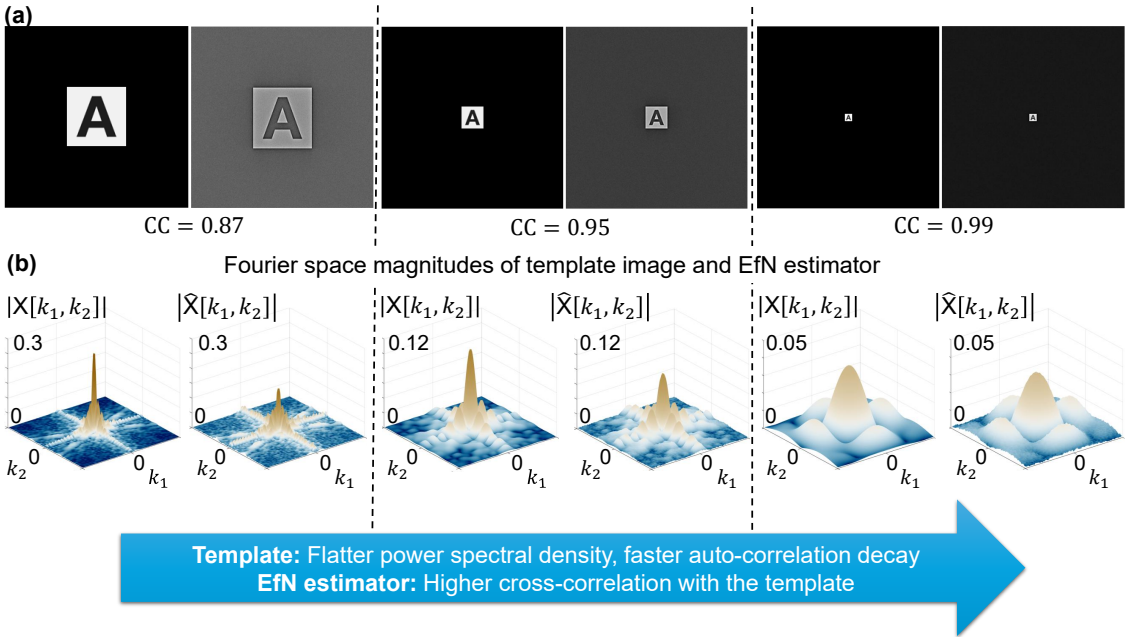

Figure 3: The influence of the power-spectral-density (PSD) of the template signal on the correlation between the template and the EfN estimator. (a) Three images of the letter A are shown, with an increasing zero-padding ratio. As the zero-padding ratio increases, the PSD flattens, and the cross-correlation (CC) between the template and the EfN estimator increases. This higher cross-correlation is evident in both the image background and the colors of the letter A. **(b)** Flatter PSDs lead to EfN estimators whose Fourier magnitudes are closer to those of the template image. The EfN estimators in these experiments were generated using $M = 10^5$ observations.

4 Main Results

We begin by analyzing the regime where $M \to \infty$ and the dimension of the signal d is fixed. In this setting, we show that the Fourier phases of the EfN estimator converge almost surely to those of the underlying template signal, and we characterize the convergence rate. We also analyze the behavior of the Fourier magnitudes. Then, we turn to the high-dimensional regime, where $d \to \infty$. Under additional assumptions, we derive refined asymptotic guarantees for both the phases and magnitudes. Throughout, we assume that the template signal $x \in \mathbb{R}^d$ has a unit norm and that its spectrum is non-vanishing, i.e., $\mathsf{X}[k] \neq 0$ for every $0 < k \leq d-1$, as discussed in the previous section.

4.1 Finite-dimensional signal

We begin with the case where the template signal has a fixed dimension d, as captured in the following result, whose proof is provided in Appendix B.5.

Theorem 4.1 (Fourier phases convergence for finite-dimensional signal)**.** *Fix $d \geq 2$ and assume that $\mathsf{X}[k] \neq 0$, for all $0 < k \leq d-1$.*

1. *For any* $0 \leq k \leq d-1$*, we have,*

$$\phi_{\hat{\mathsf{X}}}[k] \xrightarrow{\text{a.s.}} \phi_{\mathsf{X}}[k], \tag{4.1}$$

as $M \to \infty$*. Furthermore,*

$$\lim_{M\to\infty} \frac{\mathbb{E}|\phi_{\hat{\mathsf{X}}}[k] - \phi_{\mathsf{X}}[k]|^2}{1/M} = C_k, \tag{4.2}$$

for a finite constant $C_k < \infty$*.*

2. *For any* $0 \leq k \leq d-1$*, we have,*

$$|\hat{\mathsf{X}}[k]| \xrightarrow{\text{a.s.}} \mathbb{E}\left[|\mathsf{N}[k]| \cos\left(\frac{2\pi k}{d}\hat{\mathsf{R}}_1 + \phi_{\mathsf{N}}[k] - \phi_{\mathsf{X}}[k]\right)\right] > 0, \tag{4.3}$$

as $M \to \infty$*, where* $\hat{\mathsf{R}}_1$ *is defined in* (2.3)*.*

Theorem 4.1 captures two central properties. The first addresses the convergence of the EfN estimator's phases to those of the template signal. In addition, the corresponding convergence rate in MSE is proportional to $1/M$. The second result captures the convergence of the EfN estimator's magnitudes to the term given in the right-hand-side (r.h.s.) of (4.3), which is strictly greater than zero. Thus, the EfN estimator converges to a non-vanishing signal. Interestingly, this term is not necessarily proportional to the magnitudes $|\mathsf{X}[k]|$ of the template signal and, thus, the EfN estimator reproduces (asymptotically) only the phases of the template signal but not the magnitudes.

A central component of the proof of Theorem 4.1 is the circulant structure inherent in the alignment of the noise, which arises from the cyclic shift operations. This symmetry implies that the covariance matrix of the noise-aligned sum is circulant, corresponding to a cyclo-stationary Gaussian process. In particular, we apply the central limit theorem (CLT) and the strong law of large numbers (SLLN) for this setting, which yields

$$\phi_{\hat{\mathsf{X}}}[k] - \phi_{\mathsf{X}}[k] \xrightarrow{\mathcal{D}} \arctan(\mathsf{Q}_k),$$

as $M \to \infty$, where Q_k is a zero-mean Gaussian random variable with variance $\sigma^2_{\mathsf{Q}}[k] = C_k/M$, and the constant C_k admits a closed-form expression. By leveraging properties of cyclo-stationary Gaussian processes, which is justified by the circulant structure of the problem, we establish that $C_k < \infty$ for all $0 \leq k \leq d-1$. This directly leads to the results stated in (4.1)–(4.2).

4.2 High-dimensional regime

We now turn to the high-dimensional setting where $d \to \infty$, taken after the limit $M \to \infty$. In this regime, we impose additional technical conditions on the template signal, formalized in Assumption 4.2. Intuitively, these conditions reflect the empirical phenomenon illustrated in Figure 3, where a flatter PSD, which corresponds to a more rapidly decaying autocorrelation function, results in an improved alignment between the template and the estimator.

More precisely, Assumption 4.2 below requires control over the decay of both the autocorrelation function and the spectral magnitudes as functions of d. Specifically, for a length-d signal $x \in \mathbb{R}^d$, we define the (circular) autocorrelation in the time domain by

$$\mathsf{R}_{\mathsf{XX}}[\tau] \triangleq \frac{1}{\sqrt{d}} \sum_{n=0}^{d-1} x[n]\, x[n+\tau \pmod d], \tag{4.4}$$

for $\tau \in \{0, 1, \ldots, d-1\}$. By the discrete Wiener-Khinchin theorem, this is equivalent to taking the inverse discrete Fourier transform of the PSD, namely,

$$\mathsf{R}_{\mathsf{XX}}[\tau] = \mathcal{F}^{-1}\{\, |\mathsf{X}|^2 \,\}[\tau] = \frac{1}{\sqrt{d}} \sum_{k=0}^{d-1} |\mathsf{X}[k]|^2\, e^{j\frac{2\pi}{d}k\tau}, \tag{4.5}$$

where $\mathsf{X} = \mathcal{F}\{\mathsf{x}\}$ is the DFT of x. Assumption 4.2 requires that the autocorrelation $\mathsf{R}_{\mathsf{XX}}[\tau]$ decay faster than $1/\log d$, and that the maximum magnitude among nonzero Fourier components $\max_{k\neq 0} |\mathsf{X}[k]|$ decay faster than $1/\sqrt{\log d}$. In addition, we assume the DC component is vanishing, i.e., $|\mathsf{X}[0]| = 0$, to avoid degeneracies in alignment.

Assumption 4.2 (High-dimensional regularity of the template)**.** *Consider a sequence of template signals $\{x^{(d)}\}_{d\in\mathbb{N}}$ with $x^{(d)} \in \mathbb{R}^d$, and let $\mathsf{X}^{(d)} = \mathcal{F}\{x^{(d)}\}$ denote their DFT. Let $\mathsf{R}_{\mathsf{XX}}^{(d)}$ denote the autocorrelation of $x^{(d)}$, as defined in (4.4). When taking the limit $d \to \infty$, we assume the signals are normalized, namely $\|x^{(d)}\|_2 = 1$ for all $d \in \mathbb{N}$. We say that the template sequence $\{x^{(d)}\}_{d\in\mathbb{N}}$ satisfies Assumption 4.2 if the following hold:*

1. Autocorrelation decay. *The autocorrelation away from the zero lag satisfies*

$$\lim_{d\to\infty} \left(\max_{1\le\tau\le d-1} \left|\mathsf{R}_{\mathsf{XX}}^{(d)}[\tau]\right| \right) \cdot \log d = 0. \tag{4.6}$$

2. Spectral magnitude decay. *The non-DC Fourier magnitudes satisfy*

$$\lim_{d\to\infty} \left(\max_{1\le k\le d-1} \left|\mathsf{X}^{(d)}[k]\right| \right) \cdot \sqrt{\log d} = 0. \tag{4.7}$$

3. Vanishing DC component. *The signal's DC component is zero, i.e., $\left|\mathsf{X}^{(d)}[0]\right| = 0$.*

Although the conditions in Assumption 4.2 may seem technical, they are essential for establishing Theorem 4.3, which relies on classical limit theorems for the maxima of stationary Gaussian processes, most notably, convergence to the Gumbel distribution [29, 12, 2, 4]. Each part of the assumption plays a specific role: Part (1) ensures that the noise process lacks long-range dependencies, which corresponds to a sufficiently flat PSD; Part (2) guarantees that no individual Fourier component dominates the behavior of the EfN estimator. The final condition, requiring the DC component to vanish (i.e., $|\mathsf{X}^{(d)}[0]| = 0$), is not strictly necessary from an empirical standpoint but is introduced to streamline the theoretical analysis.

Theorem 4.3 (Fourier phases convergence for high-dimensional signal)**.** *Assume that $\mathsf{X}^{(d)}[k] \neq 0$, for all $0 < k \le d-1$, and that x satisfies Assumption 4.2. Then,*

1. *For any* $0 \le k \le d-1$*, we have,*

$$\lim_{d\to\infty}\lim_{M\to\infty} \frac{\mathbb{E}|\phi_{\hat{\mathsf{X}}^{(d)}}[k] - \phi_{\mathsf{X}^{(d)}}[k]|^2}{1/(M\log d)} \frac{1}{1/(4\,|\mathsf{X}^{(d)}[k]|^2)} = 1. \tag{4.8}$$

2. *For any* $0 \le k \le d-1$*, we have,*

$$\frac{1}{\sigma\sqrt{2\log d}} \frac{|\hat{\mathsf{X}}^{(d)}[k]|}{|\mathsf{X}^{(d)}[k]|} \xrightarrow{\text{a.s.}} 1, \tag{4.9}$$

as $M, d \to \infty$*.*

The proof of Theorem 4.3 is presented in Appendix D.3. Based on Theorem 4.3, as $M, d \to \infty$, the convergence rate of the Fourier phases of the EfN estimator is inversely proportional to the squared Fourier magnitude. In addition, unlike the fixed-d result in Theorem 4.1, the high-dimensional regime reveals an explicit dependence on d in the phase-error constant. Specifically, while Theorem 4.1 states that for each fixed dimension d, (4.2) holds with a finite constant C_k (which depends on d and on the template), Theorem 4.3 makes this dependence explicit when $d \to \infty$: comparing (4.8) with (4.2) shows that in the high-dimensional limit the constant scales as

$$C_k = \frac{1}{4\,|\mathsf{X}^{(d)}[k]|^2\,\log d}. \tag{4.10}$$

Intuitively, the $\log d$ factor arises from the alignment step: $\hat{\mathsf{R}}_i$ is chosen by maximizing a stationary Gaussian correlation process over d shifts, and the maximizer is governed by extreme-value statistics. In particular, the maximum of such a process grows on the scale $\sqrt{2\log d}$; see, e.g., [29, 12, 2, 4], and the proof in Appendix D.

Moreover, unlike Theorem 4.1 for a fixed d, Theorem 4.3 also shows that the Fourier magnitudes of the EfN estimator satisfy (4.9), namely they recover the template magnitudes up to the known normalization factor $\sigma\sqrt{2\log d}$. Therefore, when $d \to \infty$ under Assumption 4.2, the normalized estimate $\hat{x}^{(d)}$ recovers the template signal, which in turn implies that the normalized cross-correlation between the template and the EfN estimator approaches unity.

Empirically, we observe that Theorem 4.3 provides accurate predictions of the convergence behavior when Assumption 4.2 holds. As illustrated in Figure 4, the convergence rate is strongly influenced by the PSD of the template signal. In particular, Figure 4(b) shows that increasing the signal length and a flatter PSD lead to a stronger correlation between the EfN estimator and the true template. Furthermore, Figure 4(c) demonstrates that the convergence of the Fourier phases of the EfN estimator aligns closely with the theoretical predictions as the PSD becomes flatter. When the template violates Assumption 4.2 (e.g., if its autocorrelation decays too slowly), the predicted convergence rates become less accurate, highlighting the importance of the assumption for the theorem's formal guarantees. However, even when the spectral decay is moderate, and the assumption is not strictly met, we find that the analytical convergence rates still rather closely match empirical observations (Figure 4(c)). Notably, the key phenomenon that the convergence rate of the Fourier phases is inversely related to the magnitude of the corresponding spectral components remains robust beyond the regime where the theorem formally applies.

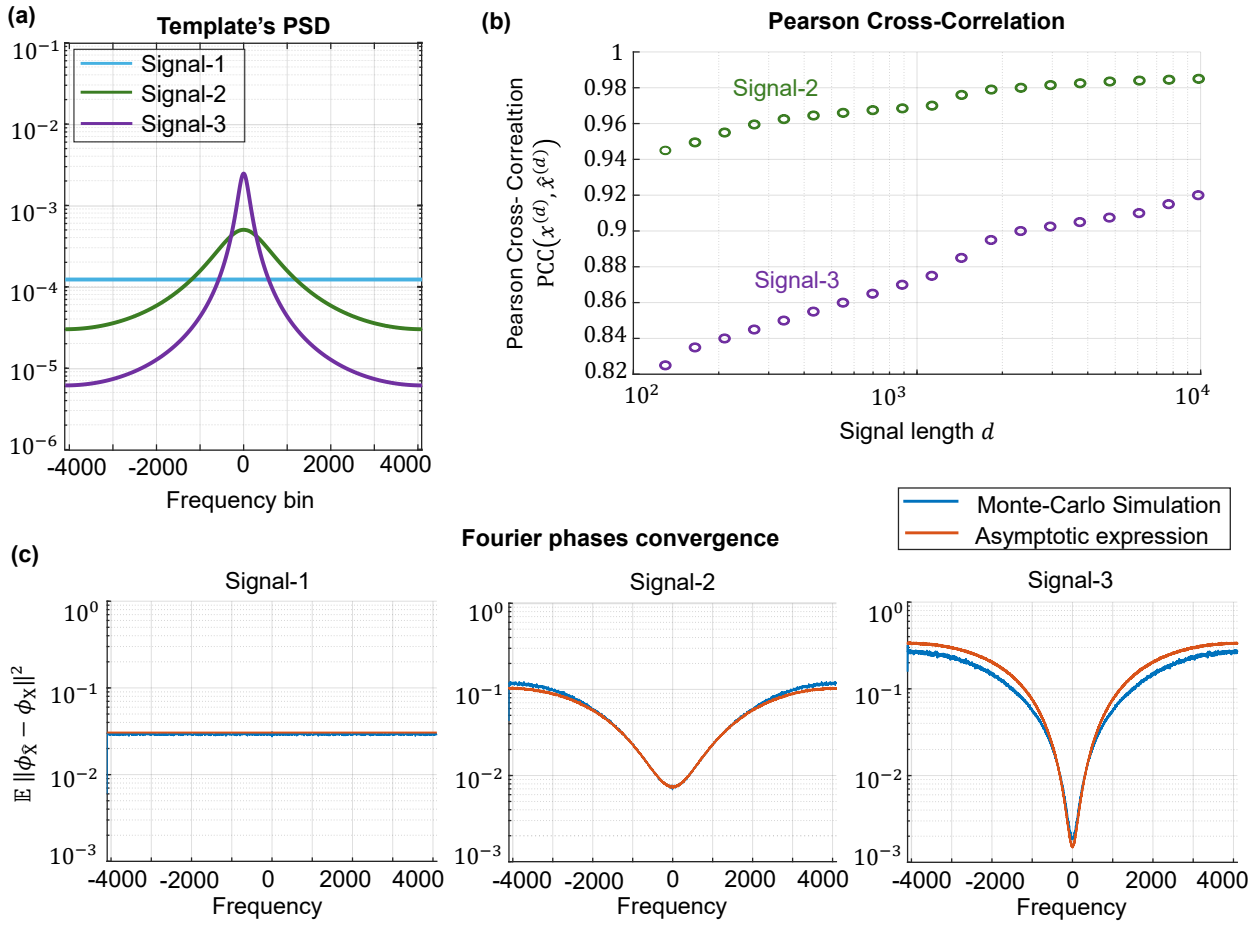

Figure 4: Comparison between analytic expression and Monte-Carlo simulations for high-dimensional signals, d, and for signals with varying power spectral densities. The analytic predictions for Fourier-phase convergence and Fourier-magnitude scaling are given by (4.8) and (4.9), respectively. **(a)** Template PSDs for three template families at a representative dimension $d = 8192$. For each dimension d, the template $x^{(d)} \in \mathbb{R}^d$ is generated directly at length d as an exponentially decaying signal, $x_\ell^{(d)}[m] \propto \exp(-m/\alpha_\ell)$, $m = 0, 1, \ldots, d-1$, with decay parameters $\alpha_\ell \in \{0.02, 2, 10\}$ (Signals 1-3, respectively), followed by mean removal and normalization. **(b)** Monte-Carlo estimates of the Pearson cross-correlation $\mathrm{PCC}(x_\ell^{(d)}, \hat{x}_\ell^{(d)})$ between the template $x_\ell^{(d)}$ and the EfN estimate $\hat{x}_\ell^{(d)}$ as a function of the signal length d (with fixed sample size $M = 10^4$). As d increases, the correlation increases, particularly for templates with slower-decaying PSDs. **(c)** Per-frequency phase mean-squared error $\mathbb{E}|\phi_{\hat{\mathsf{X}}_\ell^{(d)}}[k] - \phi_{\mathsf{X}_\ell^{(d)}}[k]|^2$ at $d = 8192$: Monte-Carlo estimates (blue) are compared with the asymptotic expression (red), i.e., the large-(M, d) closed-form approximation predicted by (4.8). All Monte-Carlo curves are averaged over 2000 independent trials.

5 Extension to other noise statistics

So far, we have analyzed the setting in which the noise is white Gaussian. In this section, we extend the analysis to a broader class of noise distributions. Specifically, we now assume that the observations $y_0, y_1, \ldots, y_{M-1} \in \mathbb{R}^d$ are i.i.d. samples drawn from an arbitrary distribution

with zero mean and a fixed covariance matrix, namely,

$$\mathbb{E}[y_1] = \mathbf{0}, \quad \mathbb{E}[y_1 y_1^\top] = \Sigma, \tag{5.1}$$

where $\Sigma \succ 0$ is a positive-definite matrix with bounded operator norm, i.e., $\|\Sigma\| < \infty$. Notably, the entries of each sample y_i are not required to be independent or identically distributed.

5.1 Positive correlation

In general, the Fourier phase convergence property established under the white Gaussian assumption does not hold for arbitrary noise distributions, as demonstrated empirically in Figures 5 and 6. Nonetheless, we establish a positive correlation result between the EfN estimator and the underlying template signal.

Proposition 5.1 (Positive correlation)**.** *Let $d \geq 2$, and suppose the observations $\{y_i\}_{i=0}^{M-1}$ are drawn i.i.d. according to the model in (5.1). Let $x \in \mathbb{R}^d$ denote the template signal, and assume its discrete Fourier transform X satisfies $\mathsf{X}[k] \neq 0$ for all $1 \leq k \leq d-1$. Let $\hat{x}$ be the EfN estimator computed from the observations $\{y_i\}$. Then, as $M \to \infty$, the following inequality holds almost surely,*

$$\langle \hat{x}, x \rangle \geq \max_{0 \leq r_1, r_2 < d-1} \frac{1}{2} \mathbb{E}\left[|\langle y_1, \mathcal{T}_{r_1} x - \mathcal{T}_{r_2} x \rangle| \right] > 0. \tag{5.2}$$

The proof of Proposition 5.1 is provided in Appendix E. This result implies that the EfN estimator is positively correlated with the true template signal. Although this is a weaker guarantee than the Fourier phase convergence obtained under Gaussian white noise, it still ensures that the estimator retains meaningful structural information from the template.

5.2 High-dimensional i.i.d. noise

Our next result demonstrates that the Fourier phase convergence established in Theorem 4.1 for Gaussian white noise extends to a broader class of noise models in the high-dimensional regime. To this end, we impose an additional assumption that the entries of each observation vector $y_i \in \mathbb{R}^d$ are i.i.d. Namely, the covariance matrix Σ is diagonal.

Theorem 5.2 (High-dimensional i.i.d. noise)**.** *Let $\{y_i\}_{i=0}^{M-1}$ be i.i.d. observations drawn according to the model in (5.1), and assume further that the entries of each $y_i \in \mathbb{R}^d$ are i.i.d., with finite variance, and satisfy $\mathbb{E}[(y_i[\ell])^4] < \infty$, for all $\ell \in \{0, 1, \ldots, d-1\}$. Let $\hat{\mathsf{X}}$ denote the discrete Fourier transform of the EfN estimator under this noise model. Assume that the Fourier coefficients of the template x are non-vanishing, i.e., $\mathsf{X}[k] \neq 0$ for all $k \in \mathbb{N}^+$. Then, for any fixed $k \in \mathbb{N}^+$, we have,*

$$\phi_{\hat{\mathsf{X}}}[k] - \phi_{\mathsf{X}}[k] \xrightarrow{\text{a.s.}} 0, \tag{5.3}$$

as $M, d \to \infty$. Moreover,

$$\lim_{d \to \infty} \lim_{M \to \infty} \frac{\mathbb{E}\left[|\phi_{\hat{\mathsf{X}}}[k] - \phi_{\mathsf{X}}[k]|^2 \right]}{1/M} = C_k, \tag{5.4}$$

for some finite constant $C_k < \infty$. Finally, if x satisfies Assumption 4.2, then,

$$\lim_{d\to\infty}\lim_{M\to\infty}\frac{\mathbb{E}\left[|\phi_{\hat{\mathsf{X}}}[k]-\phi_{\mathsf{X}}[k]|^2\right]}{1/(M\log d)}\cdot\frac{1}{1/(4|\mathsf{X}[k]|^2)}=1. \tag{5.5}$$

The proof of Theorem 5.2 is given in Appendix F. In essence, this result extends the Fourier phase convergence of Theorem 4.1 to a broader class of noise distributions in the high-dimensional setting. The main idea of the proof is to apply the functional central limit theorem to the DFT coefficients [37, 14, 13]. As $d \to \infty$, the Fourier components of the noise converge in distribution to those of a circulant Gaussian random process, owing to the i.i.d. structure of the entries in y_i. This asymptotic Gaussianity enables us to apply the same analytical framework developed for the white noise case to establish convergence of the Fourier phases.

Empirical demonstration. Figure 5 provides empirical validation of Theorem 5.2 in settings where the noise distribution is non-Gaussian. In particular, we consider $y_i \in \mathbb{R}^d$ with i.i.d. entries drawn from either the uniform or Poisson distribution. As the figure shows, when d is relatively small, the Fourier phases fail to converge and instead plateau. However, as the dimension increases, phase convergence emerges at the predicted $1/M$ rate, aligning with our theoretical results.

5.3 Circulant Gaussian process

In this section, we consider the setting in which the noise exhibits correlations between entries. As previously noted, Fourier phase convergence does not generally hold under arbitrary noise models. However, we show that convergence is maintained when the noise follows a *circulant Gaussian* distribution, a structured class of Gaussian noise characterized by rotational symmetry.

Definition 5.3 (Symmetric circulant matrix)**.** *A matrix $\Sigma \in \mathbb{R}^{d\times d}$ is called* circulant *if each row is a right cyclic shift of the previous one. That is, there exists a vector $c = (c_0, c_1, \ldots, c_{d-1}) \in \mathbb{R}^d$ such that*

$$\Sigma = \mathrm{circ}(c) = \begin{bmatrix} c_0 & c_1 & c_2 & \cdots & c_{d-1} \\ c_{d-1} & c_0 & c_1 & \cdots & c_{d-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ c_1 & c_2 & c_3 & \cdots & c_0 \end{bmatrix}. \tag{5.6}$$

The matrix is said to be symmetric circulant *if $c_j = c_{d-j}$ for all $j = 1, \ldots, d-1$.*

Proposition 5.4 (Fourier phase convergence under circulant Gaussian noise)**.** *Let $d \geq 2$ be fixed, and suppose the observations $\{y_i\}_{i=0}^{M-1}$ are i.i.d samples drawn from the multivariate normal distribution $\mathcal{N}(0, \Sigma)$, where Σ is a symmetric circulant matrix as defined in Definition 5.3. Assume further that the eigenvalues of Σ are strictly positive, and that the template signal $x \in \mathbb{R}^d$ satisfies $\mathsf{X}[k] \neq 0$ for all $1 \leq k \leq d-1$. Let $\hat{x}$ denote the EfN estimator under this noise model. Then, for each $0 \leq k \leq d-1$:*

$$\phi_{\hat{\mathsf{X}}}[k] \xrightarrow{\text{a.s.}} \phi_{\mathsf{X}}[k], \tag{5.7}$$

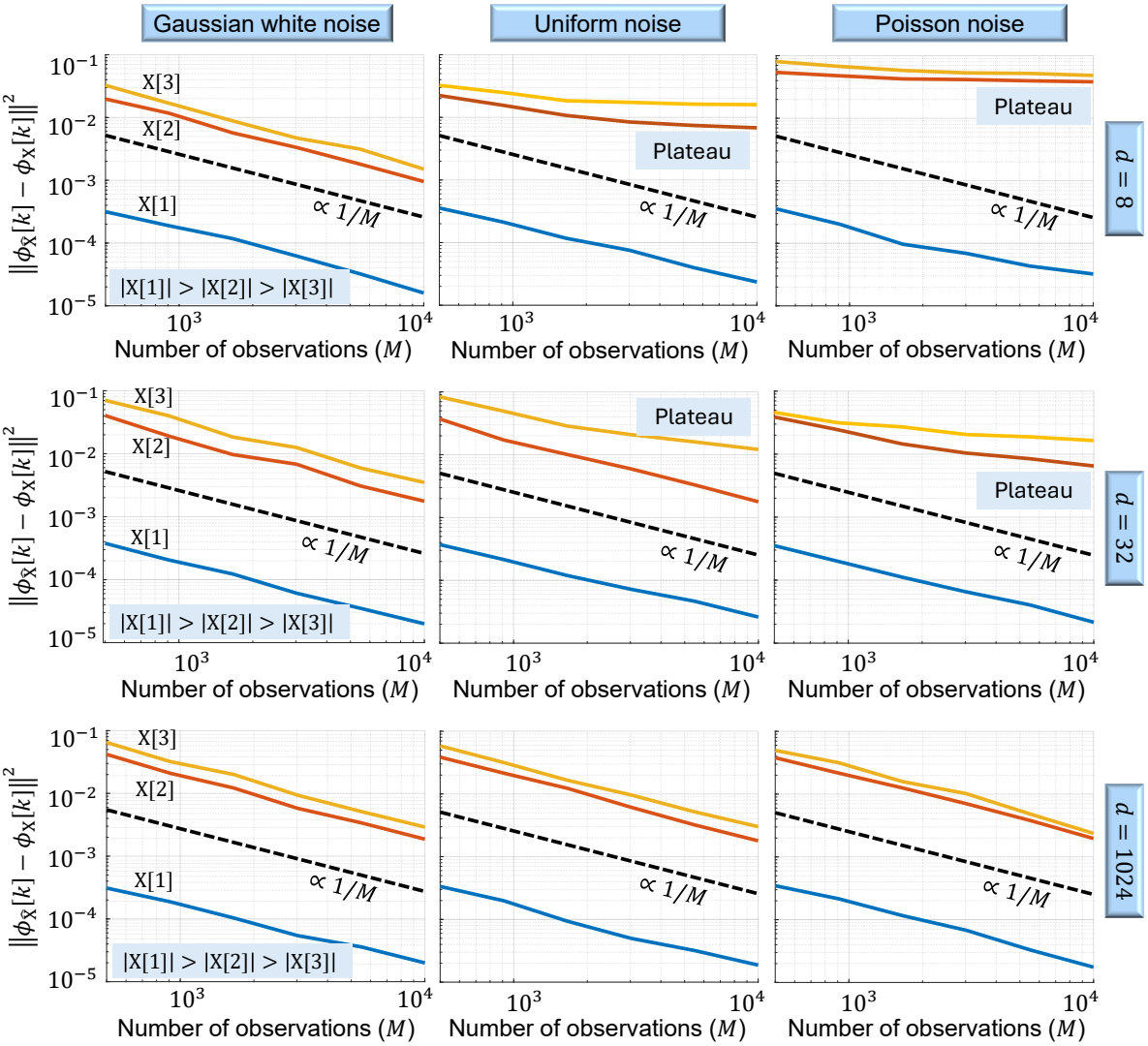

Figure 5: The impact of noise statistics and signal dimension (d) on Fourier phase convergence. Each panel displays the mean squared error (MSE) between the Fourier phases of the true template and those estimated by EfN, shown for three representative Fourier components. The dashed line represents the theoretical $1/M$ convergence rate. Columns correspond to different noise distributions: white Gaussian noise, i.i.d. noise drawn from a uniform distribution over the interval $[0, 1]$, and i.i.d. Poisson noise with parameter $\lambda = 10$. Rows correspond to increasing signal dimensions: $d = 8$, 32, and 1024. For white Gaussian noise, the Fourier phases converge at the expected $1/M$ rate across all signal dimensions, in agreement with Theorem 4.1. In contrast, under uniform and Poisson noise, the MSE plateaus at low dimensions. However, increasing the signal dimension restores convergence, even under non-Gaussian noise, consistent with the high-dimensional regime described in Theorem 5.2. Notably, for $d = 1024$, all three noise models produce similar MSE values across the selected Fourier components, suggesting that their phase noise statistics become nearly indistinguishable. Each data point represents an average of 300 Monte Carlo trials.

as $M \to \infty$. Moreover,

$$\lim_{M\to\infty} \frac{\mathbb{E}\left[|\phi_{\hat{\mathsf{X}}}[k] - \phi_{\mathsf{X}}[k]|^2\right]}{1/M} = C_k, \tag{5.8}$$

for some finite constant $C_k < \infty$.

The proof of Proposition 5.4 is given in Appendix G. In essence, this result serves as a generalization of Theorem 4.1, which considered the case of white Gaussian noise, to the broader setting of symmetric circulant Gaussian noise. Notably, white noise with covariance $\sigma^2 I_{d\times d}$ is a special case of circulant noise, making this extension a natural generalization. The critical insight here is that circulant covariance matrices remain diagonalizable in the Fourier basis, which preserves the independence of the DFT coefficients and enables phase convergence to proceed as in the white Gaussian case.

Empirical demonstration. Figure 6 presents an empirical comparison of the MSE of the Fourier phase estimates, as a function of the number of observations M, under three distinct noise models: (1) white Gaussian noise with covariance $\Sigma = \sigma^2 I_{d\times d}$; (2) Gaussian noise with a symmetric circulant covariance matrix, as defined in Definition 5.3; and (3) Gaussian noise with a Toeplitz (but non-circulant) covariance matrix. As shown in the figure, both the i.i.d. and circulant models exhibit the expected $1/M$ decay in the phase MSE curve, though the constants C_k differ, reflecting their distinct covariance structures. In contrast, under Toeplitz noise, the phase estimates do not converge: the MSE plateaus, and no $1/M$ scaling is observed. These results empirically confirm that the convergence of Fourier phases is tightly linked to the circulant structure of the noise covariance.

6 Discussion and outlook

In this work, we have shown that the Fourier phases of the EfN estimator converge to those of the template signal for an asymptotic number of observations. Since Fourier phases are crucial for perceiving image structure, the reconstructed image appears structurally similar to the template signal, even in cases where the estimator's spectral magnitudes differ from those of the template [36, 45]. We have also shown that the Fourier phases of spectral components with stronger magnitudes converge faster, leading to faster structural similarity in the overall image perception. In addition, we have extended our analysis beyond white Gaussian noise, examining other noise models. We have shown that the EfN estimator remains positively correlated with the template for arbitrary noise settings, and we have analyzed the Fourier phases convergence properties for high-dimensional i.i.d. noise (which is not necessarily Gaussian) and circulant Gaussian noise.

6.1 Extensions and implications

We anticipate that the findings of this paper will be beneficial in various fields. For example, the paper sheds light on a fundamental pitfall in template matching techniques, which may lead engineers and statisticians to misleading results. In addition, physicists and biologists working with data sets of low SNRs will benefit from understanding limitations and potential

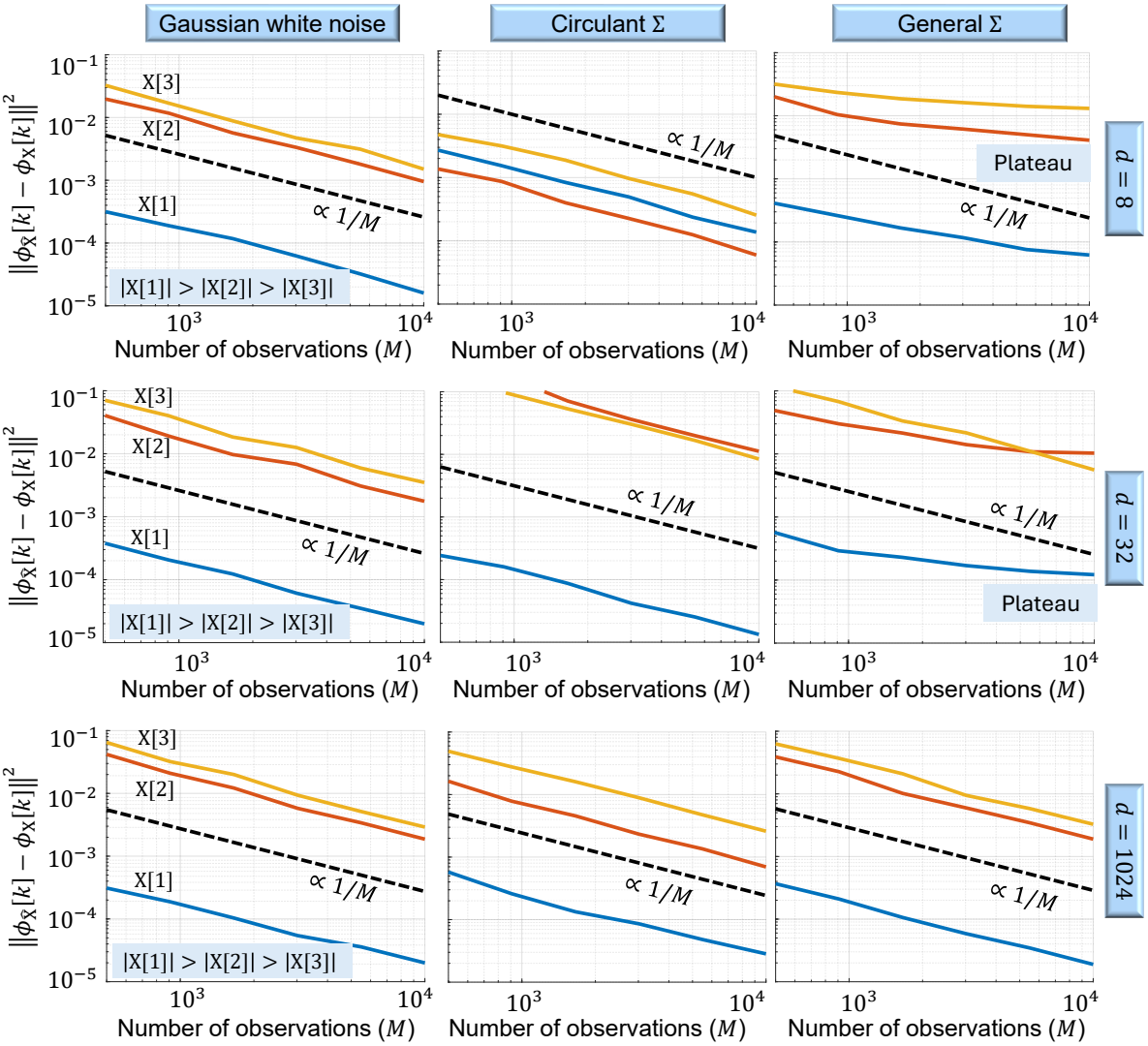

Figure 6: The impact of the noise covariance structure and signal dimension (d) on Fourier phase convergence. Each plot shows the mean squared error (MSE) between the Fourier phases of the ground-truth template and those estimated by EfN, evaluated across three spectral components. The dashed line indicates the theoretical $1/M$ convergence rate. Columns correspond to three types of noise: (1) white Gaussian noise with covariance $\Sigma = \sigma^2 I_{d\times d}$, (2) symmetric circulant covariance, and (3) a Toeplitz covariance matrix that is not circulant. Rows represent increasing signal dimensions: $d = 8$, 32, and 1024. Under white Gaussian noise, the Fourier phases converge at the expected $1/M$ rate, independent of the signal dimension (Theorem 4.1). A similar trend is observed when the noise has a circulant covariance structure: the same $1/M$ scaling holds, although the MSE is different compared to the white noise case. In contrast, for a Toeplitz covariance matrix that is not circulant, the MSE plateaus at small signal dimensions, indicating a failure of convergence. However, when the signal dimension increases to $d = 1024$, convergence at the $1/M$ rate is restored even under this more structured noise model. Each data point represents an average of 300 Monte Carlo trials.

biases introduced by template matching techniques. More generally, this work provides a cautionary framework for the broader scientific community, highlighting the importance of exercising care when interpreting noisy observations.

Extension to higher dimensions. While this paper focuses on one-dimensional signals, the analysis can be readily extended to higher dimensions. This extension involves replacing the one-dimensional DFT with its N-dimensional counterpart. The symmetry properties established in Theorem 4.1, including the results in Propositions B.3 and B.2, remain valid. For the high-dimensional case of Theorem 4.3, the conditions on the PSD adjust for the N-dimensional case. Specifically, the auto-correlation decay rate of the multidimensional array should be faster than $1/\log d$ in each dimension.

Implications to cryo-EM. The findings of Theorem 4.3 have practical implications for cryo-EM. Typically, protein spectra exhibit rapid decay at low frequencies (known as the Guinier plot) and remain relatively constant at high frequencies, a behavior characterized by Wilson in [58] and known as Wilson statistics. Wilson statistics is used to sharpen 3-D structures [47]. To mitigate the risk of model bias, we suggest using templates with reduced high frequencies, recommending filtered, smooth templates. This insight may also relate to or support the common practice of initializing the expectation-maximization (EM) algorithm for 3-D refinement with a smooth 3-D volume. Each iteration of the EM algorithm effectively applies a version of template matching multiple times, although projection images typically contain actual signal rather than pure noise, as in the EfN case.

The key message for the cryo-EM community is that, regardless of the specific setting, one should not rely solely on the raw alignment average when working with low-SNR data. Instead, robust validation practices, such as cross-validation, independent reconstructions, and other consistency checks, are essential to guard against artifact-driven effects like the EfN phenomenon. In this context, we mention a recent work suggesting that processing data in smaller mini-batches can help reduce the risk of EfN, offering a practical approach to mitigating model bias in such settings [6].

Noise statistics in cryo-EM. The results in Section 5 are particularly relevant to the noise characteristics commonly encountered in cryo-EM. While cryo-EM noise is often modeled as Poisson in nature, the standard practical assumption is that it follows a Gaussian distribution with a decaying power spectrum. These properties align well with the broader class of noise models considered in our analysis. Consequently, the conclusions of Theorem 5.2 can be extended to the cryo-EM setting, and we expect similar asymptotic phase convergence behavior to hold.

6.2 Future work

Here, we list open questions and directions for future work.

Extension to non-cyclic group actions. A natural direction for future work is to extend the EfN analysis beyond the simplified setting of cyclic translations, as defined in (2.1), to more general group actions, particularly those arising in practical applications such as cryo-EM. In this context, the relevant transformations are elements of the rotation group $\mathsf{SO}(3)$,

and the postulated observations are two-dimensional projections of a three-dimensional structure rather than simple translations of a one-dimensional signal. Recent empirical evidence suggests that template-induced bias can persist in such non-abelian settings as well, including under $\mathsf{SO}(3)$ group actions [57].

However, extending the theoretical analysis to non-abelian groups presents more substantial challenges. In particular, the property of circular Gaussian statistics, which underpins the EfN analysis for cyclic groups, does not naturally extend to the non-abelian setting. Preliminary simulations for the non-abelian dihedral group (not shown here) indicate that the convergence of the EfN estimator's Fourier phases observed in the abelian setting does not directly carry over. We nonetheless expect that alignment over a richer group may still induce systematic structural biases, potentially governed by the group's representation-theoretic decomposition. This broader perspective is consistent with recent observations of confirmation bias effects even in unstructured latent-variable models: for Gaussian mixture models with pure-noise data, a single iteration of population k-means or EM initialized at hypothesized centroids produces estimates that remain positively correlated with the initialization [5], suggesting that algorithmic bias can persist beyond group-alignment settings, albeit often in weaker forms than the phase-locking phenomenon characterized here.

Hard assignment algorithms and the EM algorithm. One promising avenue for future research involves examining hard-assignment algorithms. These algorithms iteratively refine estimates of an underlying signal from noisy observations, where the signal is obscured by high noise (unlike the pure noise scenario in EfN). The process begins by aligning observations with a template signal in the initial iteration and averaging them to improve the template for subsequent iterations. A central objective is to understand and characterize the model bias introduced throughout this iterative process, specifically, how the final output depends on the initial template. Notably, the results presented in this work can be interpreted as describing a single iteration of a hard-assignment algorithm in the limit as the SNR approaches zero.

Another important direction is investigating the EM algorithm, a cornerstone of cryo-EM algorithms [41, 40]. EM maximizes the likelihood function of models incorporating nuisance parameters [19], a topic of significant recent interest [18, 56]. Unlike hard-assignment algorithms, EM operates iteratively as a soft assignment algorithm, assigning probabilities to various possibilities and computing a weighted average rather than selecting a single optimal alignment per observation. Further exploration of EM could provide deeper insights into iterative methodologies in cryo-EM and their associated model biases.

Extension to the non-i.i.d. case. While Theorem 5.2 assumes that the noise entries within each observation vector y_i are independent and identically distributed, an important direction for future research is to extend these results to more general noise settings. Specifically, the analysis could be broadened to cover cases where the noise entries are independent but not identically distributed, provided that their variances remain uniformly bounded and a Lindeberg-type condition is fulfilled [20]. Moreover, the framework may apply to noise that exhibits certain weak dependence structures, such as mixing conditions, allowing the use of functional central limit theorems and ensuring asymptotic Gaussianity of the Fourier components [37, 14, 13].

Asymptotic regimes. In this work, we analyzed two asymptotic regimes: (1) $M \to \infty$ with fixed d (Theorem 4.1), and (2) $M \to \infty$ followed by $d \to \infty$ (Theorem 4.3). These regimes capture distinct theoretical and practical scenarios. Our approach relies on classical probabilistic tools in the first limit ($M \to \infty$), such as SLLN and CLT, and results from the theory of Gaussian extremes (e.g., convergence to the Gumbel distribution) in the second ($d \to \infty$).

Other challenging asymptotic regimes merit further investigation. In particular, it is of interest to understand the behavior in the joint high-dimensional regime where both $M, d \to \infty$ with a fixed ratio, i.e., $\frac{d}{M} \to c \in (0, \infty)$. This regime, common in modern high-dimensional statistics, differs from the sequential limits we analyze. More broadly, other asymptotic behaviors of (M, d) are possible. When both $M = M_n$ and $d = d_n$ vary according to general sequences, a variety of additional regimes may arise, each potentially requiring different analytical techniques. Typically, in such settings, classical limit theorems may no longer apply directly, and new challenges arise, such as subtle phase transitions in statistical behavior and the breakdown of averaging effects when d and M grow at comparable rates. Addressing these phenomena typically requires more advanced tools from high-dimensional probability. We view the analysis of further asymptotic settings as a valuable direction for future research.

Statistical inference. While the present work establishes the asymptotic consistency of the EfN estimator's Fourier phases, an important direction for future research is to investigate their behavior in the finite-sample regime. In particular, developing tools for statistical inference, such as confidence intervals or non-asymptotic error bounds, would enhance the practical utility of the analysis. Addressing these questions may require the use of sharper probabilistic techniques beyond classical limit theorems, such as Berry–Esseen-type results, concentration inequalities, or non-asymptotic deviation bounds tailored to the specific structure of the problem.

Acknowledgment

T.B. is supported in part by BSF under Grant 2020159, in part by NSF-BSF under Grant 2024791, and in part by ISF under Grant 1924/21. W.H. is supported by ISF under Grant 1734/21.

References

[1] Ashley Aberneithy. Automatic detection of calcified nodules of patients with tuberculous. *University College, London*, 2007.

[2] Robert J Adler and Jonathan E Taylor. *Random fields and geometry*. Springer Science & Business Media, 2009.

[3] MS Aksoy, Orhan Torkul, and Ismail Hakki Cedimoglu. An industrial visual inspection system that uses inductive learning. *Journal of Intelligent Manufacturing*, 15:569–574, 2004.

[4] Jean-Marc Azaïs and Mario Wschebor. *Level sets and extrema of random processes and fields*. John Wiley & Sons, 2009.

[5] Amnon Balanov, Tamir Bendory, and Wasim Huleihel. Confirmation bias in Gaussian Mixture Models. *IEEE Transactions on Information Theory*, 71(11):8871–8898, 2025.

[6] Amnon Balanov, Wasim Huleihel, and Tamir Bendory. Expectation-maximization for low-SNR multi-reference alignment. *arXiv preprint arXiv:2505.21435*, 2026.

[7] Amnon Balanov, Alon Zabatani, and Tamir Bendory. Structure from noise: Confirmation bias in particle picking in structural biology. *arXiv preprint arXiv:2507.03951*, 2025.

[8] Afonso S Bandeira, Ben Blum-Smith, Joe Kileel, Jonathan Niles-Weed, Amelia Perry, and Alexander S Wein. Estimation under group actions: recovering orbits from invariants. *Applied and Computational Harmonic Analysis*, 66:236–319, 2023.

[9] Tamir Bendory, Alberto Bartesaghi, and Amit Singer. Single-particle cryo-electron microscopy: Mathematical theory, computational challenges, and opportunities. *IEEE signal processing magazine*, 37(2):58–76, 2020.

[10] Tamir Bendory, Nicolas Boumal, Chao Ma, Zhizhen Zhao, and Amit Singer. Bispectrum inversion with application to multireference alignment. *IEEE Transactions on signal processing*, 66(4):1037–1050, 2017.

[11] Tristan Bepler, Andrew Morin, Micah Rapp, Julia Brasch, Lawrence Shapiro, Alex J Noble, and Bonnie Berger. Positive-unlabeled convolutional neural networks for particle picking in cryo-electron micrographs. *Nature methods*, 16(11):1153–1160, 2019.

[12] Simeon M Berman. Limit theorems for the maximum term in stationary sequences. *The Annals of Mathematical Statistics*, pages 502–516, 1964.

[13] David R Brillinger. *Time series: data analysis and theory*. SIAM, 2001.

[14] Clément Cerovecki and Siegfried Hörmann. On the CLT for discrete fourier transforms of functional time series. *Journal of multivariate analysis*, 154:282–295, 2017.

[15] Gary E Christensen, Richard D Rabbitt, and Michael I Miller. Deformable templates using large deformation kinematics. *IEEE transactions on image processing*, 5(10):1435–1447, 1996.

[16] Jon Cohen. Is high-tech view of HIV too good to be true?, 2013.

[17] Pilar Cossio. Need for cross-validation of single particle cryo-EM. *Journal of Chemical Information and Modeling*, 60(5):2413–2418, 2020.

[18] Constantinos Daskalakis, Christos Tzamos, and Manolis Zampetakis. Ten steps of EM suffice for mixtures of two gaussians. In *Conference on Learning Theory*, pages 704–710. PMLR, 2017.

[19] Arthur P Dempster, Nan M Laird, and Donald B Rubin. Maximum likelihood from incomplete data via the EM algorithm. *Journal of the royal statistical society: series B (methodological)*, 39(1):1–22, 1977.

[20] Rick Durrett. *Probability: theory and examples*, volume 49. Cambridge university press, 2019.

[21] Amitay Eldar, Keren Mor Waknin, Samuel Davenport, Tamir Bendory, Armin Schwartzman, and Yoel Shkolnisky. Object detection under the linear subspace model with application to cryo-EM images. *arXiv preprint arXiv:2405.00364*, 2024.

[22] Ayelet Heimowitz, Joakim Andén, and Amit Singer. APPLE picker: Automatic particle picking, a low-effort cryo-EM framework. *Journal of structural biology*, 204(2):215–227, 2018.

[23] Richard Henderson. Avoiding the pitfalls of single particle cryo-electron microscopy: Einstein from noise. *Proceedings of the National Academy of Sciences*, 110(45):18037–18041, 2013.

[24] Richard Henderson, Andrej Sali, Matthew L Baker, Bridget Carragher, Batsal Devkota, Kenneth H Downing, Edward H Egelman, Zukang Feng, Joachim Frank, Nikolaus Grigorieff, et al. Outcome of the first electron microscopy validation task force meeting. *Structure*, 20(2):205–214, 2012.

[25] J Bernard Heymann. Validation of 3D EM reconstructions: The phantom in the noise. *AIMS biophysics*, 2(1):21, 2015.

[26] Olav Kallenberg. *Foundations of modern probability*. Springer, 1997.

[27] Gerard J Kleywegt, Paul D Adams, Sarah J Butcher, Catherine L Lawson, Alexis Rohou, Peter B Rosenthal, Sriram Subramaniam, Maya Topf, Sanja Abbott, Philip R Baldwin, et al. Community recommendations on cryoEM data archiving and validation. *IUCrJ*, 11(2), 2024.

[28] Theocharis Kyriacou, Guido Bugmann, and Stanislao Lauria. Vision-based urban navigation procedures for verbally instructed robots. *Robotics and Autonomous Systems*, 51(1):69–80, 2005.

[29] Malcolm R Leadbetter, Georg Lindgren, and Holger Rootzén. *Extremes and related properties of random sequences and processes*. Springer Science & Business Media, 2012.

[30] Yuhai Li, Jian Liu, Jinwen Tian, and Hongbo Xu. A fast rotated template matching based on point feature. In *MIPPR 2005: SAR and Multispectral Image Processing*, volume 6043, pages 453–459. SPIE, 2005.

[31] Sergio I Lopez and Leandro PR Pimentel. On the location of the maximum of a process: L'evy, gaussian and multidimensional cases. *arXiv preprint arXiv:1611.02334*, 2016.

[32] Youdong Mao, Luis R Castillo-Menendez, and Joseph G Sodroski. Reply to subramaniam, van heel, and henderson: Validity of the cryo-electron microscopy structures of the HIV-1 envelope glycoprotein complex. *Proceedings of the National Academy of Sciences*, 110(45):E4178–E4182, 2013.

[33] Youdong Mao, Liping Wang, Christopher Gu, Alon Herschhorn, Anik Désormeaux, Andrés Finzi, Shi-Hua Xiang, and Joseph G Sodroski. Molecular architecture of the uncleaved HIV-1 envelope glycoprotein trimer. *Proceedings of the National Academy of Sciences*, 110(30):12438–12443, 2013.

[34] Amit Moscovich and Saharon Rosset. On the cross-validation bias due to unsupervised preprocessing. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 84(4):1474–1502, 2022.

[35] Eva Nogales. The development of cryo-EM into a mainstream structural biology technique. *Nature methods*, 13(1):24–27, 2016.

[36] Alan V Oppenheim and Jae S Lim. The importance of phase in signals. *Proceedings of the IEEE*, 69(5):529–541, 1981.

[37] Magda Peligrad and Wei Biao Wu. Central limit theorem for fourier transforms of stationary processes. 2010.

[38] Amelia Perry, Jonathan Weed, Afonso S Bandeira, Philippe Rigollet, and Amit Singer. The sample complexity of multireference alignment. *SIAM Journal on Mathematics of Data Science*, 1(3):497–517, 2019.

[39] Leandro PR Pimentel. On the location of the maximum of a continuous stochastic process. *Journal of Applied Probability*, 51(1):152–161, 2014.

[40] Ali Punjani, John L Rubinstein, David J Fleet, and Marcus A Brubaker. cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nature methods*, 14(3):290–296, 2017.

[41] Sjors HW Scheres. RELION: implementation of a bayesian approach to cryo-EM structure determination. *Journal of structural biology*, 180(3):519–530, 2012.

[42] Sjors HW Scheres. Semi-automated selection of cryo-EM particles in relion-1.3. *Journal of structural biology*, 189(2):114–122, 2015.

[43] Vahid Shahverdi, Emanuel Ström, and Joakim Andén. Moment Constraints and Phase Recovery for Multireference Alignment. *arXiv preprint arXiv:2409.04868*, 2024.

[44] Maxim Shatsky, Richard J Hall, Steven E Brenner, and Robert M Glaeser. A method for the alignment of heterogeneous macromolecules from electron microscopy. *Journal of structural biology*, 166(1):67–78, 2009.

[45] Yoav Shechtman, Yonina C Eldar, Oren Cohen, Henry Nicholas Chapman, Jianwei Miao, and Mordechai Segev. Phase retrieval with application to optical imaging: a contemporary overview. *IEEE signal processing magazine*, 32(3):87–109, 2015.

[46] Fred J Sigworth. A maximum-likelihood approach to single-particle image refinement. *Journal of structural biology*, 122(3):328–339, 1998.

[47] Amit Singer. Wilson statistics: derivation, generalization and applications to electron cryomicroscopy. *Acta Crystallographica Section A: Foundations and Advances*, 77(5):472–479, 2021.

[48] Amit Singer and Fred J Sigworth. Computational methods for single-particle electron cryomicroscopy. *Annual review of biomedical data science*, 3:163–190, 2020.

[49] E. Slutsky. Über stochastische Asymptoten und Grenzwerte. 1925.

[50] Carlos OS Sorzano, JL Vilas, Erney Ramírez-Aportela, J Krieger, D Del Hoyo, David Herreros, Estrella Fernandez-Giménez, D Marchán, JR Macías, I Sánchez, et al. Image processing tools for the validation of CryoEM maps. *Faraday Discussions*, 240:210–227, 2022.

[51] Alex Stewart and Nikolaus Grigorieff. Noise bias in the refinement of structures derived from single particles. *Ultramicroscopy*, 102(1):67–84, 2004.

[52] Sriram Subramaniam. Structure of trimeric HIV-1 envelope glycoproteins. *Proceedings of the National Academy of Sciences*, 110(45):E4172–E4174, 2013.

[53] Itamar Talmi, Roey Mechrez, and Lihi Zelnik-Manor. Template matching with deformable diversity similarity. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 175–183, 2017.

[54] Marin van Heel. Finding trimeric HIV-1 envelope glycoproteins in random noise. *Proceedings of the National Academy of Sciences*, 110(45):E4175–E4177, 2013.

[55] Shao-Hsuan Wang, Yi-Ching Yao, Wei-Hau Chang, and I-Ping Tu. Quantification of model bias underlying the phenomenon of "Einstein from noise". *Statistica Sinica*, 31:2355–2379, 2021.

[56] Ji Xu, Daniel J Hsu, and Arian Maleki. Global analysis of expectation maximization for mixtures of two gaussians. *Advances in Neural Information Processing Systems*, 29, 2016.

[57] Sheng Xu, Amnon Balanov, Amit Singer, and Tamir Bendory. Bayesian perspective for orientation determination in cryo-EM with application to structural heterogeneity analysis. *bioRxiv*, pages 2024–12, 2025.

[58] SH Yü. Determination of absolute from relative X-ray intensity data. *Nature*, 150(3796):151–152, 1942.

[59] Richard Zhang, Phillip Isola, Alexei A Efros, Eli Shechtman, and Oliver Wang. The unreasonable effectiveness of deep features as a perceptual metric. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 586–595, 2018.

[60] Andy Zhu. *Multireference Alignment via Semidefinite Programming.* Phd thesis, Princeton University, 2013.

Appendix

Appendix organization. Appendix A provides general preliminaries used throughout the paper, including notation and common technical tools. Appendix B presents the auxiliary lemmas required for Theorem 4.1, whose full proof appears in Appendix B.5. For Theorem 4.3, the necessary supporting results are given in Appendices C, with the proof provided in Appendix D. Appendix E contains the proof of Proposition 5.1, establishing the positive correlation property. Appendix F proves Theorem 5.2, which extends the results to high-dimensional settings with i.i.d. noise that is not necessarily Gaussian. Finally, Appendix G provides the proof of Proposition 5.4, addressing the case of structured noise with a circulant Gaussian covariance.

A Preliminaries

Before we delve into the proofs, we fix notations and definitions and prove auxiliary results that will be used in the proofs.

A.1 Notations

Recall the definitions of the Fourier transforms of x and n_i from (2.6), and recall that the signal length d is assumed to be even. Note that since x and n_i are real-valued, their Fourier coefficients satisfy the conjugate-symmetry relation:

$$\mathsf{X}[k] = \overline{\mathsf{X}[d-k]}, \quad \mathsf{N}_i[k] = \overline{\mathsf{N}_i[d-k]}. \tag{A.1}$$

In particular, $|\mathsf{N}_i[k]| = |\mathsf{N}_i[d-k]|$ and $\phi_{\mathsf{N}_i}[k] = -\phi_{\mathsf{N}_i}[d-k]$, which implies that only the first $d/2+1$ components of $\mathsf{N}[k]$ are statistically independent.

The definition of the maximal correlation in (2.3) can be represented in the Fourier domain as follows,

$$\hat{\mathsf{R}}_i \triangleq \operatorname*{arg\,max}_{0\le r\le d-1} \langle n_i, \mathcal{T}_r x\rangle \tag{A.2}$$

$$= \operatorname*{arg\,max}_{0\le r\le d-1} \langle \mathcal{F}\{n_i\}, \mathcal{F}\{\mathcal{T}_r x\}\rangle \tag{A.3}$$

$$= \operatorname*{arg\,max}_{0\le r\le d-1} \sum_{k=0}^{d-1} |\mathsf{X}[k]|\,|\mathsf{N}_i[k]| \cos\left(\frac{2\pi kr}{d} + \phi_{\mathsf{N}_i}[k] - \phi_{\mathsf{X}}[k]\right). \tag{A.4}$$

To simplify notation, we define

$$\mathsf{S}_i[r] \triangleq \sum_{k=0}^{d-1} |\mathsf{X}[k]|\,|\mathsf{N}_i[k]| \cos\left(\frac{2\pi kr}{d} + \phi_{\mathsf{N}_i}[k] - \phi_{\mathsf{X}}[k]\right), \tag{A.5}$$

for $0 \leq r \leq d-1$, and therefore, $\hat{\mathsf{R}}_i = \arg\max_{0\leq r\leq d-1} \mathsf{S}_i[r]$. We note that for any $0 \leq i \leq M-1$, the random vector $\mathsf{S}_i \triangleq (\mathsf{S}_i[0], \mathsf{S}_i[1], \ldots, \mathsf{S}_i[d-1])^T$ is Gaussian distributed, with zero mean vector, and a circulant covariance matrix; therefore, it is a cyclo-stationary random process.

Our goal is to investigate the phase and magnitude of the estimator $\hat{\mathsf{X}}$ in (2.7). Simple manipulations reveal that, for any $0 \leq k \leq d-1$, the estimator's phases are given by,

$$\phi_{\hat{\mathsf{X}}}[k] = \phi_{\mathsf{X}}[k] + \arctan\left(\frac{\sum_{i=0}^{M-1} |\mathsf{N}_i[k]| \sin\left(\phi_{e,i}[k]\right)}{\sum_{i=0}^{M-1} |\mathsf{N}_i[k]| \cos\left(\phi_{e,i}[k]\right)}\right), \tag{A.6}$$

where we define,

$$\phi_{e,i}[k] \triangleq \frac{2\pi k \hat{\mathsf{R}}_i}{d} + \phi_{\mathsf{N}_i}[k] - \phi_{\mathsf{X}}[k], \tag{A.7}$$

and

$$|\hat{\mathsf{X}}[k]| = \frac{1}{M}\left|\sum_{i=0}^{M-1} |\mathsf{N}_i[k]|\, e^{j\phi_{e,i}[k]}\right|. \tag{A.8}$$

A.2 The convergence of the Einstein from Noise estimator

Recall the definition of $\phi_{e,i}[k]$ in (A.7). Then, following (2.7), and simple algebraic manipulation,

$$\hat{\mathsf{X}}[k] = \frac{1}{M}\sum_{i=0}^{M-1} |\mathsf{N}_i[k]|\, e^{j\phi_{\mathsf{N}_i}[k]} e^{j\frac{2\pi k}{d}\hat{\mathsf{R}}_i} \tag{A.9}$$

$$= \frac{e^{j\phi_{\mathsf{X}}[k]}}{M}\sum_{i=0}^{M-1} |\mathsf{N}_i[k]|\, e^{j\phi_{\mathsf{N}_i}[k]} e^{j\frac{2\pi k}{d}\hat{\mathsf{R}}_i} e^{-j\phi_{\mathsf{X}}[k]} \tag{A.10}$$

$$= \frac{e^{j\phi_{\mathsf{X}}[k]}}{M}\sum_{i=0}^{M-1} |\mathsf{N}_i[k]|\, e^{j\phi_{e,i}[k]}. \tag{A.11}$$

By applying the strong law of large numbers (SLLN) on the right-hand-side of (A.11), for $M \to \infty$, we have,

$$\hat{\mathsf{X}}[k] e^{-j\phi_{\mathsf{X}}[k]} = \frac{1}{M}\sum_{i=0}^{M-1} |\mathsf{N}_i[k]|\, e^{j\phi_{e,i}[k]} \tag{A.12}$$

$$\xrightarrow{\text{a.s.}} \mathbb{E}\left[|\mathsf{N}_1[k]| \cos\left(\phi_{e,1}[k]\right)\right] + j\mathbb{E}\left[|\mathsf{N}_1[k]| \sin\left(\phi_{e,1}[k]\right)\right], \tag{A.13}$$

where we have used the fact that the sequences of random variables $\{|\mathsf{N}_i[k]| \sin\left(\phi_{e,i}[k]\right)\}_{i=0}^{M-1}$ and $\{|\mathsf{N}_i[k]| \cos\left(\phi_{e,i}[k]\right)\}_{i=0}^{M-1}$ are i.i.d. with finite mean and variances.

We denote for every $0 \leq k \leq d-1$:

$$\mu_{\mathsf{A},k} \triangleq \mathbb{E}\left[|\mathsf{N}_1[k]| \sin(\phi_{e,1}[k])\right], \tag{A.14}$$

$$\mu_{\mathsf{B},k} \triangleq \mathbb{E}\left[|\mathsf{N}_1[k]|\cos(\phi_{e,1}[k])\right], \tag{A.15}$$

the imaginary and real parts of the right-hand-side of (A.13), respectively. In addition, we denote:

$$\sigma^2_{\mathsf{A},k} \triangleq \mathsf{Var}\left(|\mathsf{N}_1[k]|\sin(\phi_{e,1}[k])\right), \tag{A.16}$$

$$\sigma^2_{\mathsf{B},k} \triangleq \mathsf{Var}\left(|\mathsf{N}_1[k]|\cos(\phi_{e,1}[k])\right). \tag{A.17}$$

In Theorem 4.1, we prove that $\mu_{\mathsf{A},k} = 0$ while $\mu_{\mathsf{B},k} > 0$. Consequently, by (A.13), as $M \to \infty$, the EfN estimator converges to a non-vanishing signal, and its Fourier phases converge those of the template (Einstein).

A.3 Conditioning on the Fourier frequency noise component

Throughout the proofs, we condition the noise realization S_i (A.5) on the k-th Fourier coefficient $\mathsf{S}_i|\mathsf{N}_i[k]$, to capture the dependence of $\hat{\mathsf{R}}_i$ on the noise component. Specifically, we prove the following:

Lemma A.1. *Recall the definition of* S_i (A.5)*. Then, for every* $k \in \left\{1, 2, \ldots, \frac{d}{2}-1, \frac{d}{2}+1, \ldots d-1\right\}$,

$$\mathsf{S}_i|\mathsf{N}_i[k] \sim \mathcal{N}(\mu_{k,i}, \Sigma_{k,i}), \tag{A.18}$$

where,

$$\mu_{k,i}[r] \triangleq \mathbb{E}\left[\mathsf{S}_i[r]|\mathsf{N}_i[k]\right] = 2\,|\mathsf{X}[k]|\,|\mathsf{N}_i[k]|\cos\left(\frac{2\pi kr}{d} + \phi_{\mathsf{N}_i}[k] - \phi_{\mathsf{X}}[k]\right), \tag{A.19}$$

for $0 \le r \le d-1$*, and*

$$\begin{aligned}\Sigma_{k,i}[r,s] &\triangleq \mathbb{E}\left[\left(\mathsf{S}_i[r] - \mathbb{E}\mathsf{S}_i[r]\right)\left(\mathsf{S}_i[s] - \mathbb{E}\mathsf{S}_i[s]\right)|\mathsf{N}_i[k]\right] \\ &= \frac{\sigma^2}{2}\sum_{\ell=0}^{d-1}|\widetilde{\mathsf{X}}_k[\ell]|^2\cos\left(\frac{2\pi\ell}{d}(r-s)\right),\end{aligned} \tag{A.20}$$

for $0 \le r, s \le d-1$*, where* $\widetilde{\mathsf{X}}_k$ *is defined by:*

$$\widetilde{\mathsf{X}}_k[\ell] \triangleq \begin{cases} 0 & \text{if}\,\ell = k, d-k, \\ \mathsf{X}[\ell] & \text{if}\,\ell = 0, d/2, \\ \sqrt{2}\cdot\mathsf{X}[\ell] & \text{otherwise.}\end{cases} \tag{A.21}$$

Remark A.2. *In Lemma A.1, and throughout this work, we condition on* $\mathsf{S}_i|\mathsf{N}_i[k]$ *for all* $k \neq 0, d/2$*. Since the signals* x *and* n_i *lie in* $\mathbb{R}^d$*, their Fourier phases satisfy* $\phi_{\mathsf{X}}[0] = 0$ *and* $\phi_{\mathsf{X}}[d/2] = 0$*. Therefore, we restrict our analysis to the convergence of the Fourier phases for* $k \neq 0, d/2$*, as the convergence at* $k = 0$ *and* $k = d/2$ *is trivial.*

Note that the conditional process $\mathsf{S}_i|\mathsf{N}_i[k]$ is Gaussian because it is given by a linear transform of i.i.d. Gaussian variables. Also, since its covariance matrix is circulant and depends only on the difference between the two indices, i.e., $\Sigma_{k,i}[r,s] = \sigma_{k,i}[|r-s|]$, it is cycle-stationary with a cosine trend. The eigenvalues of this circulant matrix are given by the DFT of its first row, and thus its ℓ-th eigenvalue equals $|\widetilde{\mathsf{X}}_k[\ell]|^2$, for $0 \le \ell \le d-1$.

For simplicity of notation, whenever it is clear from the context, we will omit the dependence of the above quantities on the i-th observation and k-th frequency indices, and we will use $\mu[r]$ and $\Sigma[r,s]$, instead. Furthermore, for convenience, we assume that the template vector is normalized to unity, i.e. $\sum_{\ell=0}^{d-1} |\mathsf{X}[\ell]|^2 = 1$.

Proof of Lemma A.1. By definition of S_i (A.5), we have for every $k \neq 0, d/2$,

$$\begin{aligned}\mathsf{S}_i[r]\,|\mathsf{N}_i[k] =& 2\,|\mathsf{X}[k]|\,|\mathsf{N}_i[k]|\cos\left(\frac{2\pi kr}{d} + \phi_{\mathsf{N}_i}[k] - \phi_{\mathsf{X}}[k]\right) \\ &+ \sum_{\ell\neq k,d-k} |\mathsf{X}[\ell]|\,|\mathsf{N}_i[\ell]|\,\cos\left(\frac{2\pi \ell r}{d} + \phi_{\mathsf{N}_i}[\ell] - \phi_{\mathsf{X}}[\ell]\right),\end{aligned} \tag{A.22}$$

where we have used the property of $\mathsf{X}[k] = \overline{\mathsf{X}[d-k]}$, $\mathsf{N}_i[k] = \overline{\mathsf{N}_i[d-k]}$, (A.1). Clearly, as $\mathbb{E}\left[\mathsf{N}_i[\ell]\right] = 0$, for every $0 \le \ell \le d-1$, we have,

$$\mathbb{E}\left[|\mathsf{X}[\ell]|\,|\mathsf{N}_i[\ell]|\,\cos\left(\frac{2\pi \ell r}{d} + \phi_{\mathsf{N}_i}[\ell] - \phi_{\mathsf{X}}[\ell]\right)\right] = 0, \tag{A.23}$$

for every $0 \le \ell \le d-1$. Combining (A.22) and (A.23) results,

$$\mu_{k,i}[r] = \mathbb{E}\left[\mathsf{S}_i[r]|\mathsf{N}_i[k]\right] = 2\,|\mathsf{X}[k]|\,|\mathsf{N}_i[k]|\cos\left(\frac{2\pi kr}{d} + \phi_{\mathsf{N}_i}[k] - \phi_{\mathsf{X}}[k]\right), \tag{A.24}$$

proving the first result about the means.

The covariance term. In the following, we derive the covariance term,

$$\Sigma_{k,i}[r,s] \triangleq \mathbb{E}\left[\left(\mathsf{S}_i[r] - \mathbb{E}\mathsf{S}_i[r]\right)\left(\mathsf{S}_i[s] - \mathbb{E}\mathsf{S}_i[s]\right)|\mathsf{N}_i[k]\right]. \tag{A.25}$$

Denote,

$$\begin{aligned}\rho_{k,i}[r] &\triangleq \mathsf{S}_i[r] - \mathbb{E}\mathsf{S}_i[r] \\ &= \sum_{\ell\neq k,d-k} |\mathsf{X}[\ell]|\,|\mathsf{N}_i[\ell]|\,\cos\left(\frac{2\pi \ell r}{d} + \phi_{\mathsf{N}_i}[\ell] - \phi_{\mathsf{X}}[\ell]\right).\end{aligned} \tag{A.26}$$

Denote the set

$$\mathcal{I} = \{1, 2, \ldots k-1, k+1, \ldots, d/2-1\}, \tag{A.27}$$

which defines the indices of the Fourier coefficients, excluding $\{0, k, d/2\}$.

As the sequences $\{|\mathsf{N}_i[\ell]|\}_{\ell=0}^{d/2}$ and $\{\phi_{\mathsf{N}_i}[\ell]\}_{\ell=0}^{d/2}$ satisfy $\mathsf{N}_i[\ell] = \overline{\mathsf{N}_i[d-\ell]}$, as well as $\mathsf{X}[\ell] = \overline{\mathsf{X}[d-\ell]}$, we have,

$$\begin{aligned}\rho_{k,i}[r] &= \sum_{\ell\neq k,d-k} |\mathsf{X}[\ell]|\,|\mathsf{N}_i[\ell]| \cos\left(\frac{2\pi\ell r}{d} + \phi_{\mathsf{N}_i}[\ell] - \phi_{\mathsf{X}}[\ell]\right) = \\ &= \sum_{\ell\in\{0,d/2\}} |\mathsf{X}[\ell]|\,|\mathsf{N}_i[\ell]| \cos\left(\frac{2\pi\ell r}{d} + \phi_{\mathsf{N}_i}[\ell] - \phi_{\mathsf{X}}[\ell]\right) \\ &\quad + 2\cdot\sum_{\ell\in\mathcal{I}} |\mathsf{X}[\ell]|\,|\mathsf{N}_i[\ell]| \cos\left(\frac{2\pi\ell r}{d} + \phi_{\mathsf{N}_i}[\ell] - \phi_{\mathsf{X}}[\ell]\right),\end{aligned} \tag{A.28}$$

where each one of the terms in the sum is independent.

Since the terms in the sum on the right-hand side of (A.28) are independent, that is, $\mathbb{E}\left[\mathsf{N}_i[\ell_1]\,\overline{\mathsf{N}_i[\ell_2]}\right] = \mathbb{E}\left[|\mathsf{N}_i[\ell_1]|^2\right]\delta_{\ell_1,\ell_2}$, it follows that,

$$\begin{aligned}\Sigma_{k,i}[r,s] &= \mathbb{E}\left[\rho_{k,i}[r]\,\rho_{k,i}[s]\,|\mathsf{N}_i[k]\right] \\ &= \mathbb{E}\left[\sum_{\ell\in\{0,d/2\}} |\mathsf{X}[\ell]|^2\,|\mathsf{N}_i[\ell]|^2 \cos\left(\frac{2\pi\ell r}{d} + \phi_{\mathsf{N}_i}[\ell] - \phi_{\mathsf{X}}[\ell]\right)\cos\left(\frac{2\pi\ell s}{d} + \phi_{\mathsf{N}_i}[\ell] - \phi_{\mathsf{X}}[\ell]\right)\right] \\ &\quad + 4\cdot\mathbb{E}\left[\sum_{\ell\in\mathcal{I}} |\mathsf{X}[\ell]|^2\,|\mathsf{N}_i[\ell]|^2 \cos\left(\frac{2\pi\ell r}{d} + \phi_{\mathsf{N}_i}[\ell] - \phi_{\mathsf{X}}[\ell]\right)\cos\left(\frac{2\pi\ell s}{d} + \phi_{\mathsf{N}_i}[\ell] - \phi_{\mathsf{X}}[\ell]\right)\right].\end{aligned} \tag{A.29}$$

The expectation value in (A.29) is composed of the multiplications of cosines. Applying trigonometric identities, we obtain:

$$\begin{aligned}&\cos\left(\frac{2\pi\ell r}{d} + \phi_{\mathsf{N}_i}[\ell] - \phi_{\mathsf{X}}[\ell]\right)\cos\left(\frac{2\pi\ell s}{d} + \phi_{\mathsf{N}_i}[\ell] - \phi_{\mathsf{X}}[\ell]\right) \\ &\quad = \frac{1}{2}\cos\left(\frac{2\pi\ell(r-s)}{d}\right) + \frac{1}{2}\cos\left(\frac{2\pi\ell(r+s)}{d} + 2\left(\phi_{\mathsf{N}_i}[\ell] - \phi_{\mathsf{X}}[\ell]\right)\right).\end{aligned} \tag{A.30}$$

For $\ell \in \left\{1,\ldots,\frac{d}{2}-1\right\}$, the DFT coefficients $N_i[\ell]$ are i.i.d. circular complex Gaussian (under white Gaussian noise in the time domain), hence independent across ℓ, with $\phi_{N_i}[\ell] \sim \mathrm{Unif}[-\pi,\pi)$ independent of $|N_i[\ell]|$. Thus,

$$\begin{aligned}&\mathbb{E}\left[|\mathsf{N}_i[\ell]|^2\cos\left(\frac{2\pi\ell r}{d} + \phi_{\mathsf{N}_i}[\ell] - \phi_{\mathsf{X}}[\ell]\right)\cos\left(\frac{2\pi\ell s}{d} + \phi_{\mathsf{N}_i}[\ell] - \phi_{\mathsf{X}}[\ell]\right)\right] \\ &\quad = \frac{1}{2}\mathbb{E}\left[|\mathsf{N}_i[\ell]|^2\right]\cos\left(\frac{2\pi\ell(r-s)}{d}\right) = \frac{\sigma^2}{2}\cos\left(\frac{2\pi\ell(r-s)}{d}\right).\end{aligned} \tag{A.31}$$

Substituting (A.31) into (A.29) leads to,

$$\frac{2}{\sigma^2}\mathbb{E}\left[\rho_{k,i}[r]\,\rho_{k,i}[s]\,|\mathsf{N}_i[k]\right] = \sum_{\ell\in\{0,d/2\}} |\mathsf{X}[\ell]|^2\cos\left(\frac{2\pi\ell}{d}(r-s)\right)$$

$$+ 4 \cdot \sum_{\ell \in \mathcal{I}} |\mathsf{X}[\ell]|^2 \cos\left(\frac{2\pi\ell}{d}(r-s)\right). \tag{A.32}$$

As for every $\ell \in \mathcal{I}$, $|\mathsf{X}[\ell]| = |\mathsf{X}[d-\ell]|$, we have,

$$\sum_{\ell \in \mathcal{I}} 4\,|\mathsf{X}[\ell]|^2 \cos\left(\frac{2\pi\ell}{d}(r-s)\right) = \sum_{\ell \neq \{0,k,d/2,d-k\}} 2\,|\mathsf{X}[\ell]|^2 \cos\left(\frac{2\pi\ell}{d}(r-s)\right). \tag{A.33}$$

Substituting (A.33) into (A.32)

$$\mathbb{E}\left[\rho_{k,i}[r]\,\rho_{k,i}[s]\,|\mathsf{N}_i[k]\right] = \frac{\sigma^2}{2}\sum_{\ell=0}^{d-1} |\widetilde{\mathsf{X}}_k[\ell]|^2 \cos\left(\frac{2\pi\ell}{d}(r-s)\right),$$

for $\widetilde{\mathsf{X}}_k[\ell]$ defined in (A.21), which completes the proof. □

A.4 Uniqueness of the maximizer

Lemma A.3 (Uniqueness of the maximizer)**.** *Recall the definition of* S_i *from* (A.5)*. Assume* $d \geq 6$ *is even. Fix* $k \in \{1, \ldots, \frac{d}{2}-1, \frac{d}{2}+1, \ldots, d-1\}$ *and recall from Lemma A.1 that*

$$\mathsf{S}_i|\mathsf{N}_i[k] \sim \mathcal{N}(\mu_{k,i}, \Sigma_{k,i}), \qquad \Sigma_{k,i}[r,s] = \frac{\sigma^2}{2}\sum_{\ell=0}^{d-1} \left|\widetilde{\mathsf{X}}_k[\ell]\right|^2 \cos\left(\frac{2\pi\ell}{d}(r-s)\right). \tag{A.34}$$

Assume moreover that the template spectrum is non-vanishing, i.e. $\mathsf{X}[\ell] \neq 0$ *for all* $\ell \in \{0, \ldots, d-1\}$*. Then, for every* $r \neq s$*,*

$$\mathrm{Var}\Big((\mathsf{S}_i|\mathsf{N}_i[k])_r - (\mathsf{S}_i|\mathsf{N}_i[k])_s\Big) = \sigma^2 \sum_{\ell=0}^{d-1} \left|\widetilde{\mathsf{X}}_k[\ell]\right|^2 \Big(1 - \cos\big(\tfrac{2\pi\ell}{d}(r-s)\big)\Big) > 0. \tag{A.35}$$

Consequently, the maximizer $\hat{\mathsf{R}}_i = \arg\max_{0 \leq r \leq d-1} (\mathsf{S}_i|\mathsf{N}_i[k])_r$ *is unique almost surely.*

Proof of Lemma A.3. Fix $r \neq s$ and set $m \triangleq r - s \not\equiv 0 \pmod{d}$. Using the covariance formula and circulantness,

$$\begin{aligned}\mathrm{Var}\Big((\mathsf{S}_i|\mathsf{N}_i[k])_r - (\mathsf{S}_i|\mathsf{N}_i[k])_s\Big) &= \Sigma_{k,i}[r,r] + \Sigma_{k,i}[s,s] - 2\Sigma_{k,i}[r,s] \\ &= \sigma^2 \sum_{\ell=0}^{d-1} \left|\widetilde{\mathsf{X}}_k[\ell]\right|^2 \Big(1 - \cos\big(\tfrac{2\pi\ell}{d} m\big)\Big),\end{aligned} \tag{A.36}$$

which proves the identity in (A.35). Each summand is nonnegative.

It remains to show strict positivity. We show that there is at least one term in the sum that is strictly positive. Define

$$H_m \triangleq \{\ell \in \{0, \ldots, d-1\} : \ \ell m \equiv 0 \pmod{d}\}. \tag{A.37}$$

Then $|H_m| = \gcd(d, m) \le d/2$ since $m \not\equiv 0 \pmod d$. Because $d \ge 6$, we have

$$\big|\{0, \dots, d-1\} \setminus (H_m \cup \{k, d-k\})\big| \ge d - |H_m| - 2 \tag{A.38}$$

$$\ge d - \frac{d}{2} - 2 = \frac{d}{2} - 2 > 0. \tag{A.39}$$

Hence we may choose $\ell_0 \in \{0, \dots, d-1\} \setminus (H_m \cup \{k, d-k\})$. By construction, $\ell_0 \notin H_m$ implies $\ell_0 m \not\equiv 0 \pmod d$, i.e. $\cos(2\pi\ell_0 m/d) \ne 1$ and thus $1 - \cos(2\pi\ell_0 m/d) > 0$. Moreover, since $\widetilde{\mathsf{X}}_k[\ell] = 0$ only for $\ell \in \{k, d-k\}$ and $\mathsf{X}[\ell] \ne 0$ for all ℓ, we have $|\widetilde{\mathsf{X}}_k[\ell_0]|^2 > 0$. Therefore the ℓ_0-term in the sum is strictly positive, and the entire variance is strictly positive, proving (A.35).

Finally, for any $r \ne s$, the difference $(\mathsf{S}_i|\mathsf{N}_i[k])_r - (\mathsf{S}_i|\mathsf{N}_i[k])_s$ is a non-degenerate Gaussian, hence $\mathbb{P}\big((\mathsf{S}_i|\mathsf{N}_i[k])_r = (\mathsf{S}_i|\mathsf{N}_i[k])_s\big) = 0$. A union bound over finitely many pairs implies ties occur with probability 0, so the maximizer is unique almost surely. □

A.5 Positive probability of each maximizer event

Lemma A.4 (Positive probability of each strict maximizer event)**.** *Recall the definition of* S_i *from* (A.5)*. Assume* $d \ge 6$ *is even. Fix* $k \in \{1, \dots, \frac{d}{2}-1, \frac{d}{2}+1, \dots, d-1\}$ *and recall from Lemma A.1 that*

$$\mathsf{S}_i|\mathsf{N}_i[k] \sim \mathcal{N}(\mu_{k,i}, \Sigma_{k,i}), \qquad \Sigma_{k,i}[r,s] = \frac{\sigma^2}{2} \sum_{\ell=0}^{d-1} \big|\widetilde{\mathsf{X}}_k[\ell]\big|^2 \cos\Big(\frac{2\pi\ell}{d}(r-s)\Big). \tag{A.40}$$

Assume moreover that the template spectrum is non-vanishing, i.e. $\mathsf{X}[\ell] \ne 0$ *for all* $\ell \in \{0, \dots, d-1\}$*. Then, for every* $r \in \{0, \dots, d-1\}$*, the event*

$$\mathcal{C}_r \triangleq \Big\{ (\mathsf{S}_i|\mathsf{N}_i[k])_r > \max_{t \ne r} (\mathsf{S}_i|\mathsf{N}_i[k])_t \Big\} \tag{A.41}$$

has strictly positive probability:

$$\mathbb{P}\big(\mathcal{C}_r \,\big|\, \mathsf{N}_i[k]\big) > 0, \tag{A.42}$$

for almost every realization of $\mathsf{N}_i[k]$*.*

Proof of Lemma A.4. Fix k as in the statement and condition on a realization of $\mathsf{N}_i[k]$. Then $Y \triangleq \mathsf{S}_i|\mathsf{N}_i[k] \sim \mathcal{N}(m, \Sigma)$ with $m = \mu_{k,i}$ and $\Sigma = \Sigma_{k,i}$. By Lemma A.3, for every $r \ne s$, $\mathrm{Var}(Y_r - Y_s) > 0$, hence $Y_r - Y_s$ is a non-degenerate Gaussian and $\mathbb{P}(Y_r = Y_s) = 0$. In particular, ties occur with probability 0.

Since Σ is real, symmetric, and circulant, it is diagonalized by the DFT basis: there exist eigenvectors $\{f_\ell\}_{\ell=0}^{d-1}$ (the Fourier modes) such that the corresponding eigenvalues $\{\lambda_\ell\}_{\ell=0}^{d-1}$ are given by the DFT of the first row of Σ. In Model A.40, Lemma A.1 shows that these eigenvalues satisfy

$$\lambda_\ell \;\propto\; |\widetilde{\mathsf{X}}_k[\ell]|^2, \qquad \ell = 0, \dots, d-1. \tag{A.43}$$

Under the non-vanishing spectrum assumption $\mathsf{X}[\ell] \neq 0$ for all ℓ, and by the definition (A.21), we have

$$|\widetilde{\mathsf{X}}_k[\ell]|^2 > 0 \quad \text{for all } \ell \notin \{k, d-k\}, \tag{A.44}$$

while $|\widetilde{\mathsf{X}}_k[k]|^2 = |\widetilde{\mathsf{X}}_k[d-k]|^2 = 0$. Hence Σ has exactly two zero eigenvalues, corresponding to the $\ell = k$ and $\ell = d-k$ Fourier modes. In the real domain, these two modes are equivalently represented by the cosine and sine vectors $c, s \in \mathbb{R}^d$

$$c_t = \cos\Big(\tfrac{2\pi k}{d} t\Big), \qquad s_t = \sin\Big(\tfrac{2\pi k}{d} t\Big), \tag{A.45}$$

for $t = 0, \ldots, d-1$, so

$$\ker(\Sigma) = \operatorname{span}\{c, s\}. \tag{A.46}$$

Let $L \triangleq \operatorname{range}(\Sigma)$. Since Σ is symmetric, we have $L = (\ker \Sigma)^\perp$, and therefore the Gaussian vector $Y \sim \mathcal{N}(m, \Sigma)$ is supported on the affine subspace $m + L$.

We fix $r \in \{0, \ldots, d-1\}$ and consider the open cone

$$\mathcal{C}_r \triangleq \{y \in \mathbb{R}^d : \; y_r > \max_{t \neq r} y_t\}. \tag{A.47}$$

We claim that $\mathcal{C}_r \cap (m + L) \neq \emptyset$, i.e., the affine support contains at least one point whose unique largest coordinate is the r-th.

To build such a point, we start from the r-th standard basis vector $u = e_r$, whose maximizer is trivially at r (i.e., $e_r \in \mathcal{C}_r$), but which may fail to belong to $m + L$. We therefore remove from u its components along the two forbidden directions c and s (which span $\ker(\Sigma)$ (A.46)). We define

$$\alpha \triangleq \frac{\langle u, c\rangle}{\langle c, c\rangle}, \qquad \beta \triangleq \frac{\langle u, s\rangle}{\langle s, s\rangle}, \qquad v \triangleq u - \alpha\, c - \beta\, s. \tag{A.48}$$

Because $k \notin \{0, d/2\}$, the sine and cosine vectors are orthogonal and have equal energy:

$$\langle c, s\rangle = 0, \qquad \langle c, c\rangle = \langle s, s\rangle = d/2. \tag{A.49}$$

Hence $\langle v, c\rangle = \langle v, s\rangle = 0$, which means $v \in (\ker \Sigma)^\perp = L$.

Next we show that v remains strongly peaked at the r-th coordinate. Since $u = e_r$, we have $\langle u, c\rangle = c_r$ and $\langle u, s\rangle = s_r$, and therefore

$$v_r = 1 - \frac{c_r^2}{\langle c, c\rangle} - \frac{s_r^2}{\langle s, s\rangle} = 1 - \frac{c_r^2 + s_r^2}{d/2} = 1 - \frac{2}{d}, \tag{A.50}$$

where we used $c_r^2 + s_r^2 = 1$ (A.45). For $t \neq r$, since $u = e_r$ is the r-th standard basis vector, we have $u_t = 0$; thus, using (A.45) and (A.49), we obtain

$$v_t = -\alpha\, c_t - \beta\, s_t = -\frac{c_r c_t + s_r s_t}{d/2} = -\frac{2}{d} \cos\Big(\tfrac{2\pi k}{d}(t - r)\Big), \tag{A.51}$$

so $|v_t| \leq 2/d$ for all $t \neq r$. Consequently, letting

$$\delta \triangleq v_r - \max_{t \neq r} v_t, \tag{A.52}$$

we obtain the uniform lower bound for $d > 4$

$$\delta \geq \left(1 - \frac{2}{d}\right) - \frac{2}{d} = 1 - \frac{4}{d} > 0. \tag{A.53}$$

Thus, within the direction $v \in L$, the r-th coordinate exceeds every other coordinate by at least δ.

Finally, because the affine support is $m + L$, any point of the form $y(\gamma) = m + \gamma v$ lies in the support. The offset m may change the ordering for small γ, but scaling γ makes the v-term dominate. In particular, let

$$M \triangleq \max_{t \neq r} |m_r - m_t|. \tag{A.54}$$

Then for any $t \neq r$,

$$y_r(\gamma) - y_t(\gamma) = (m_r - m_t) + \gamma(v_r - v_t) \;\geq\; -M + \gamma\delta. \tag{A.55}$$

Choosing $\gamma > M/\delta$ guarantees $y_r(\gamma) > y_t(\gamma)$ for all $t \neq r$, hence

$$y(\gamma) \in \mathcal{C}_r \cap (m + L), \tag{A.56}$$

proving $\mathcal{C}_r \cap (m + L) \neq \emptyset$.

Since $\mathcal{C}_r$ is open in $\mathbb{R}^d$, the intersection $\mathcal{C}_r \cap (m+L)$ contains a nonempty open subset of the affine support $m+L$. A Gaussian measure assigns positive probability to any nonempty open subset of its affine support [26]; therefore, $\mathbb{P}(Y \in \mathcal{C}_r) > 0$, or equivalently $\mathbb{P}\big(\mathcal{C}_r \,\big|\, \mathsf{N}_i[k]\big) > 0$, completing the proof. □

A.6 Auxiliary result for Proposition B.2

Let $\mathsf{S}^{(+)} \sim \mathcal{N}(\mu, \Sigma)$ and $\mathsf{S}^{(-)} \sim \mathcal{N}(-\mu, \Sigma)$ be two d-dimensional Gaussian vectors, where Σ is a real, symmetric, circulant covariance matrix. Define the maximizers

$$\hat{\mathsf{R}}^{(+)} = \arg\max_{0 \leq \ell \leq d-1} \mathsf{S}_\ell^{(+)}, \tag{A.57}$$

$$\hat{\mathsf{R}}^{(-)} = \arg\max_{0 \leq \ell \leq d-1} \mathsf{S}_\ell^{(-)}. \tag{A.58}$$

and assume they are unique almost surely. We define the entries of μ as,

$$\mu_\ell \triangleq [\mu]_\ell = \cos\left(\frac{2\pi k}{d}\ell + \varphi\right), \tag{A.59}$$

for $\varphi \in [0, 2\pi)$, and $0 \leq \ell \leq d-1$. Note that $-\mu_\ell = \cos\left(\frac{2\pi k}{d}\ell + \varphi + \pi\right)$, for $0 \leq \ell \leq d-1$.

Then, we have the following result.

Proposition A.5. *Consider the definitions in (A.59)–(A.58), and assume the maximizers in (A.58) are unique a.s. Moreover, assume that for every $r \in \{1, \dots, d-1\}$,*

$$\mathbb{P}\left(\mathsf{S}_r^{(+)} > \max_{t \neq r} \mathsf{S}_t^{(+)}\right) > 0 \qquad \textit{and} \qquad \mathbb{P}\left(\mathsf{S}_r^{(-)} > \max_{t \neq r} \mathsf{S}_t^{(-)}\right) > 0. \tag{A.60}$$

Fix $0 \leq \ell \leq d-1$. If $\mu_\ell > 0$, then,

$$\mathbb{P}\left[\hat{\mathsf{R}}^{(+)} = \ell\right] > \mathbb{P}\left[\hat{\mathsf{R}}^{(-)} = \ell\right], \tag{A.61}$$

otherwise, if $\mu_\ell < 0$, then,

$$\mathbb{P}\left[\hat{\mathsf{R}}^{(-)} = \ell\right] > \mathbb{P}\left[\hat{\mathsf{R}}^{(+)} = \ell\right]. \tag{A.62}$$

In particular, for any $\varphi \in [0, 2\pi)$ and $0 \leq k \leq d-1$,

$$\mathbb{E}\left[\cos\left(\frac{2\pi k}{d}\hat{\mathsf{R}}^{(+)} + \varphi\right)\right] + \mathbb{E}\left[\cos\left(\frac{2\pi k}{d}\hat{\mathsf{R}}^{(-)} + \varphi + \pi\right)\right] > 0. \tag{A.63}$$

Proof of Proposition A.5. By definition, it is clear that,

$$\mathbb{P}\left[\hat{\mathsf{R}}^{(+)} = \ell\right] = \mathbb{P}\left[\mathsf{S}_\ell^{(+)} \geq \max_{m \neq \ell} \mathsf{S}_m^{(+)}\right], \tag{A.64}$$

and,

$$\mathbb{P}\left[\hat{\mathsf{R}}^{(-)} = \ell\right] = \mathbb{P}\left[\mathsf{S}_\ell^{(-)} \geq \max_{m \neq \ell} \mathsf{S}_m^{(-)}\right], \tag{A.65}$$

for $0 \leq \ell \leq d-1$. Since $\mathsf{S}^{(+)}$ and $\mathsf{S}^{(-)}$ can be decomposed as $\mathsf{S}^{(+)} = \mathsf{Z} + \mu$ and $\mathsf{S}^{(-)} = \mathsf{Z} - \mu$, where Z is a cyclo-stationary process, and μ is defined in (A.59). Then,

$$\mathbb{P}\left[\mathsf{S}_\ell^{(+)} \geq \max_{m \neq \ell} \mathsf{S}_m^{(+)}\right] = \mathbb{P}\left[\mathsf{Z}_\ell + \mu_\ell \geq \max_{m \neq \ell} \mathsf{Z}_m + \mu_m\right], \tag{A.66}$$

and,

$$\mathbb{P}\left[\mathsf{S}_\ell^{(-)} \geq \max_{m \neq \ell} \mathsf{S}_m^{(-)}\right] = \mathbb{P}\left[\mathsf{Z}_\ell - \mu_\ell \geq \max_{m \neq \ell} \mathsf{Z}_m - \mu_m\right]. \tag{A.67}$$

We will show that for any ℓ such that $\mu_\ell > 0$, we have,

$$\mathbb{P}\left[\mathsf{Z}_\ell \geq \max_{m \neq \ell} \{\mathsf{Z}_m + \mu_m - \mu_\ell\}\right] > \mathbb{P}\left[\mathsf{Z}_\ell \geq \max_{m \neq \ell} \{\mathsf{Z}_m - \mu_m + \mu_\ell\}\right], \tag{A.68}$$

which in turn implies that $\mathbb{P}\{\hat{\mathsf{R}}^{(+)} = \ell\} > \mathbb{P}\{\hat{\mathsf{R}}^{(-)} = \ell\}$.

By definition, since Z is a zero-mean Gaussian, cyclo-stationary random process (i.e., with a real, symmetric, circulant covariance matrix), its cumulative distribution function F_{Z} is invariant under cyclic shifts, i.e.,

$$F_{\mathsf{Z}}\left(z_0, z_1, \dots, z_{d-1}\right) = F_{\mathsf{Z}}\left(z_\tau, z_{\tau+1}, \dots, z_{\tau+d-1}\right), \tag{A.69}$$

for any $\tau \in \mathbb{Z}$, with indices taken modulo d. Moreover, reversing the time indices does not affect the distribution; that is,

$$F_{\mathsf{Z}}\left(z_0, z_1, \ldots, z_{\ell-1}, z_\ell, z_{\ell+1}, \ldots, z_{d-1}\right) = F_{\mathsf{Z}}\left(z_{d-1}, z_{d-2}, \ldots, z_{\ell+1}, z_\ell, z_{\ell-1}, ..., z_0\right). \tag{A.70}$$

Combining (A.69) and (A.70) yields,

$$F_{\mathsf{Z}}\left(z_\ell, z_{\ell+1}, z_{\ell+2}, \ldots, z_{\ell-2}, z_{\ell-1}\right) = F_{\mathsf{Z}}\left(z_\ell, z_{\ell-1}, z_{\ell-2}, \ldots, z_{\ell+2}, z_{\ell+1}\right). \tag{A.71}$$

Accordingly, let us define the Gaussian vectors $\mathsf{Z}^{(1)}$ and $\mathsf{Z}^{(2)}$, such that their m-th entry is,

$$[\mathsf{Z}^{(1)}]_m = \mathsf{Z}_{\ell+m} - \mathsf{Z}_\ell, \tag{A.72}$$

$$[\mathsf{Z}^{(2)}]_m = \mathsf{Z}_{\ell-m} - \mathsf{Z}_\ell, \tag{A.73}$$

for $1 \leq m \leq d-1$. It is clear from (A.71) that $\mathsf{Z}^{(1)}$ and $\mathsf{Z}^{(2)}$ have the same cumulative distribution function, i.e.,

$$F_{\mathsf{Z}^{(1)}} = F_{\mathsf{Z}^{(2)}}. \tag{A.74}$$

Therefore, the following holds,

$$\begin{aligned}
&\mathbb{P}\left[\mathsf{Z}_\ell \geq \max_{m\neq 0}\left\{\mathsf{Z}_{\ell+m} + \mu_{\ell+m} - \mu_\ell\right\}\right] \\
&\qquad = \mathbb{P}\left[0 \geq \max_{m\neq 0}\left\{\mathsf{Z}_{\ell+m} - \mathsf{Z}_\ell + \mu_{\ell+m} - \mu_\ell\right\}\right] \\
&\qquad = \mathbb{P}\left[0 \geq \max_{m\neq 0}\left\{\mathsf{Z}_{\ell-m} - \mathsf{Z}_\ell + \mu_{\ell+m} - \mu_\ell\right\}\right] \\
&\qquad = \mathbb{P}\left[\mathsf{Z}_\ell \geq \max_{m\neq 0}\left\{\mathsf{Z}_{\ell-m} + \mu_{\ell+m} - \mu_\ell\right\}\right],
\end{aligned} \tag{A.75}$$

where the second equality follows from (A.74). Next, we note that for every $0 < m \leq d-1$ and $\mu_\ell > 0$,

$$\mu_{\ell-m} + \mu_{\ell+m} = 2\mu_\ell \cos\left(\frac{2\pi k}{d} m\right). \tag{A.76}$$

Therefore,

$$\mu_\ell - \mu_{\ell-m} + \mu_\ell - \mu_{\ell+m} = 2\mu_\ell\left(1 - \cos\left(\frac{2\pi k}{d} m\right)\right) \geq 0, \tag{A.77}$$

which implies

$$\mu_\ell - \mu_{\ell-m} \geq \mu_{\ell+m} - \mu_\ell, \tag{A.78}$$

or, equivalently,

$$\mu_\ell - \mu_{\ell+m} \geq \mu_{\ell-m} - \mu_\ell. \tag{A.79}$$

Remark A.6. *According to* (A.77)*, the inequalities in* (A.78) *and* (A.79) *are strict whenever* $\cos\left(\frac{2\pi k}{d}m\right) < 1$*, which holds for the majority of values of* m*. In particular, at least* $d/2$ *of the inequalities are strict for* $0 \leq m \leq d-1$*.*

Following from (A.78), (A.79), and the last remark, we have the following auxiliary Lemma, which we prove below.

Lemma A.7. *Assume the maximizers in* (A.58) *are unique a.s. and satisfy* (A.60)*. Then, for* $\mu_\ell > 0$*, we have,*

$$\mathbb{P}\left[\mathsf{Z}_\ell \geq \max_{m\neq 0}\left\{\mathsf{Z}_{\ell+m} + \mu_{\ell+m} - \mu_\ell\right\}\right] > \mathbb{P}\left[\mathsf{Z}_\ell \geq \max_{m\neq 0}\left\{\mathsf{Z}_{\ell-m} - \mu_{\ell-m} + \mu_\ell\right\}\right]. \tag{A.80}$$

Note that (A.80) is equivalent to the following expression, by a change of index notation:

$$\mathbb{P}\left[\mathsf{Z}_\ell \geq \max_{m\neq \ell}\left\{\mathsf{Z}_{m} + \mu_{m} - \mu_\ell\right\}\right] > \mathbb{P}\left[\mathsf{Z}_\ell \geq \max_{m\neq \ell}\left\{\mathsf{Z}_{m} - \mu_{m} + \mu_\ell\right\}\right], \tag{A.81}$$

which proves (A.68). A similar result can be obtained for the case where $\mu_\ell < 0$, i.e.,

$$\mathbb{P}\left[\mathsf{Z}_\ell \geq \max_{m\neq \ell}\left\{\mathsf{Z}_{m} + \mu_{m} - \mu_\ell\right\}\right] < \mathbb{P}\left[\mathsf{Z}_\ell \geq \max_{m\neq \ell}\left\{\mathsf{Z}_{m} - \mu_{m} + \mu_\ell\right\}\right], \tag{A.82}$$

which completes the proofs of (A.61)–(A.62).

Finally, we prove (A.63). By definition, it is clear that

$$\begin{aligned}\mathbb{E}\left[\cos\left(\frac{2\pi k}{d}\hat{\mathsf{R}}^{(+)} + \varphi\right)\right] + \mathbb{E}\left[\cos\left(\frac{2\pi k}{d}\hat{\mathsf{R}}^{(-)} + \varphi + \pi\right)\right] \\ = \sum_{\ell=0}^{d-1}\cos\left(\frac{2\pi k}{d}\ell + \varphi\right)\left[\mathbb{P}\left(\hat{\mathsf{R}}^{(+)} = \ell\right) - \mathbb{P}\left(\hat{\mathsf{R}}^{(-)} = \ell\right)\right],\end{aligned} \tag{A.83}$$

where we have used the fact that $\cos(\alpha+\pi) = -\cos\alpha$, for any $\alpha \in \mathbb{R}$.

By (A.61)–(A.62), as for any $0 \leq \ell \leq d-1$ such that $\mu_\ell = \cos\left(\frac{2\pi k}{d}\ell + \varphi\right) > 0$ it holds that $\mathbb{P}[\hat{\mathsf{R}}^{(+)} = \ell] > \mathbb{P}[\hat{\mathsf{R}}^{(-)} = \ell]$, otherwise, for $0 \leq \ell \leq d-1$ such that $\mu_\ell = \cos\left(\frac{2\pi k}{d}\ell + \varphi\right) < 0$, it holds that $\mathbb{P}[\hat{\mathsf{R}}^{(+)} = \ell] < \mathbb{P}[\hat{\mathsf{R}}^{(-)} = \ell]$. Therefore,

$$\sum_{\ell=0}^{d-1}\cos\left(\frac{2\pi k}{d}\ell + \varphi\right)\left[\mathbb{P}\left(\hat{\mathsf{R}}^{(+)} = \ell\right) - \mathbb{P}\left(\hat{\mathsf{R}}^{(-)} = \ell\right)\right] > 0, \tag{A.84}$$

which in light of (A.83) concludes the proof. □

It is left to prove Lemma A.7.

Proof of Lemma A.7. Using (A.78) and (A.79), we obtain the following inequalities for $\mu_\ell > 0$,

$$\max_{m\neq 0}\left\{\mathsf{Z}_{\ell-m} - \mu_{\ell+m} + \mu_\ell\right\} \geq \max_{m\neq 0}\left\{\mathsf{Z}_{\ell-m} + \mu_{\ell-m} - \mu_\ell\right\}, \tag{A.85}$$

and

$$\max_{m\neq 0}\left\{\mathsf{Z}_{\ell-m}-\mu_{\ell-m}+\mu_\ell\right\}\geq \max_{m\neq 0}\left\{\mathsf{Z}_{\ell-m}+\mu_{\ell+m}-\mu_\ell\right\}, \tag{A.86}$$

As a result, we also have the following probabilistic inequalities,

$$\mathbb{P}\left[\mathsf{Z}_\ell<\max_{m\neq 0}\left\{\mathsf{Z}_{\ell-m}-\mu_{\ell+m}+\mu_\ell\right\}\right]\geq \mathbb{P}\left[\mathsf{Z}_\ell<\max_{m\neq 0}\left\{\mathsf{Z}_{\ell-m}+\mu_{\ell-m}-\mu_\ell\right\}\right], \tag{A.87}$$

and,

$$\mathbb{P}\left[\mathsf{Z}_\ell<\max_{m\neq 0}\left\{\mathsf{Z}_{\ell-m}-\mu_{\ell-m}+\mu_\ell\right\}\right]\geq \mathbb{P}\left[\mathsf{Z}_\ell<\max_{m\neq 0}\left\{\mathsf{Z}_{\ell-m}+\mu_{\ell+m}-\mu_\ell\right\}\right]. \tag{A.88}$$

Next, we show that these probabilistic inequalities are in fact strict. Define the set of indices where the inequality in (A.79) is strict,

$$\mathcal{M}=\left\{m:\mu_\ell-\mu_{\ell+m}>\mu_{\ell-m}-\mu_\ell\right\}. \tag{A.89}$$

From Remark A.6, we know that $|\mathcal{M}|\geq d/2$. Now define the event,

$$\mathcal{C}_r=\left\{\mathsf{Z}_{\ell-r}-\mu_{\ell-r}+\mu_\ell=\max_{m\neq 0}\left\{\mathsf{Z}_{\ell-m}-\mu_{\ell-m}+\mu_\ell\right\}\right\}, \tag{A.90}$$

i.e., the event that $\mathsf{Z}_{\ell-r}-\mu_{\ell-r}+\mu_\ell$ attains the maximum in the expression above. By assumption (A.60), the event $\mathcal{C}_r$ for every r, i.e. $\mathbb{P}(\mathcal{C}_r)>0$. Thus, we may choose $r\in\mathcal{M}$ with $\mathbb{P}(\mathcal{C}_r)>0$, i.e.,

$$r\in\mathcal{M},\quad \text{and}\quad \mathbb{P}(\mathcal{C}_r)=\mathbb{P}\left(\max_{m\neq 0}\left\{\mathsf{Z}_{\ell-m}-\mu_{\ell-m}+\mu_\ell\right\}=\mathsf{Z}_{\ell-r}-\mu_{\ell-r}+\mu_\ell\right)>0. \tag{A.91}$$

Then, by the law of total probability,

$$\begin{aligned}\mathbb{P}\left[\mathsf{Z}_\ell<\max_{m\neq 0}\left\{\mathsf{Z}_{\ell-m}-\mu_{\ell-m}+\mu_\ell\right\}\right]&=\mathbb{P}\left[\mathsf{Z}_\ell<\max_{m\neq 0}\left\{\mathsf{Z}_{\ell-m}-\mu_{\ell-m}+\mu_\ell\right\}\mid\mathcal{C}_r\right]\mathbb{P}\left[\mathcal{C}_r\right]\\&\quad+\mathbb{P}\left[\mathsf{Z}_\ell<\max_{m\neq 0}\left\{\mathsf{Z}_{\ell-m}-\mu_{\ell-m}+\mu_\ell\right\}\mid\mathcal{C}_r^c\right]\mathbb{P}\left[\mathcal{C}_r^c\right].\end{aligned} \tag{A.92}$$

From (A.88), we have,

$$\mathbb{P}\left[\mathsf{Z}_\ell<\max_{m\neq 0}\left\{\mathsf{Z}_{\ell-m}-\mu_{\ell-m}+\mu_\ell\right\}\mid\mathcal{C}_r^c\right]\geq\mathbb{P}\left[\mathsf{Z}_\ell<\max_{m\neq 0}\left\{\mathsf{Z}_{\ell-m}+\mu_{\ell+m}-\mu_\ell\right\}\mid\mathcal{C}_r^c\right]. \tag{A.93}$$

Additionally, since $r\in\mathcal{M}$, it follows that,

$$\mathbb{P}\left[\mathsf{Z}_\ell<\max_{m\neq 0}\left\{\mathsf{Z}_{\ell-m}-\mu_{\ell-m}+\mu_\ell\right\}\mid\mathcal{C}_r\right]>\mathbb{P}\left[\mathsf{Z}_\ell<\max_{m\neq 0}\left\{\mathsf{Z}_{\ell-m}+\mu_{\ell+m}-\mu_\ell\right\}\mid\mathcal{C}_r\right]. \tag{A.94}$$

Substituting (A.93) and (A.94) into (A.92) yields,

$$\mathbb{P}\left[\mathsf{Z}_\ell < \max_{m\neq 0}\left\{\mathsf{Z}_{\ell-m} - \mu_{\ell-m} + \mu_\ell\right\}\right] > \mathbb{P}\left[\mathsf{Z}_\ell < \max_{m\neq 0}\left\{\mathsf{Z}_{\ell-m} + \mu_{\ell+m} - \mu_\ell\right\} \mid \mathcal{C}_r\right]\mathbb{P}\left[\mathcal{C}_r\right]$$
$$+\mathbb{P}\left[\mathsf{Z}_\ell < \max_{m\neq 0}\left\{\mathsf{Z}_{\ell-m} + \mu_{\ell+m} - \mu_\ell\right\} \mid \mathcal{C}_r^c\right]\mathbb{P}\left[\mathcal{C}_r^c\right]. \tag{A.95}$$

By the law of total probability, the right-hand-side of (A.95) is,

$$\mathbb{P}\left[\mathsf{Z}_\ell < \max_{m\neq 0}\left\{\mathsf{Z}_{\ell-m} + \mu_{\ell+m} - \mu_\ell\right\}\right] = \mathbb{P}\left[\mathsf{Z}_\ell < \max_{m\neq 0}\left\{\mathsf{Z}_{\ell-m} + \mu_{\ell+m} - \mu_\ell\right\} \mid \mathcal{C}_r\right]\mathbb{P}\left[\mathcal{C}_r\right]$$
$$+\mathbb{P}\left[\mathsf{Z}_\ell < \max_{m\neq 0}\left\{\mathsf{Z}_{\ell-m} + \mu_{\ell+m} - \mu_\ell\right\} \mid \mathcal{C}_r^c\right]\mathbb{P}\left[\mathcal{C}_r^c\right]. \tag{A.96}$$

Combining (A.95) and (A.96), we conclude,

$$\mathbb{P}\left[\mathsf{Z}_\ell < \max_{m\neq 0}\left\{\mathsf{Z}_{\ell-m} - \mu_{\ell-m} + \mu_\ell\right\}\right] > \mathbb{P}\left[\mathsf{Z}_\ell < \max_{m\neq 0}\left\{\mathsf{Z}_{\ell-m} + \mu_{\ell+m} - \mu_\ell\right\}\right]. \tag{A.97}$$

Equivalently, we can express (A.97) as a complementary event, and obtain,

$$\mathbb{P}\left[\mathsf{Z}_\ell \geq \max_{m\neq 0}\left\{\mathsf{Z}_{\ell+m} + \mu_{\ell+m} - \mu_\ell\right\}\right] > \mathbb{P}\left[\mathsf{Z}_\ell \geq \max_{m\neq 0}\left\{\mathsf{Z}_{\ell-m} - \mu_{\ell-m} + \mu_\ell\right\}\right], \tag{A.98}$$

which proves (A.80), and completes the proof. □

B Proof of Theorem 4.1

First, we prove several auxiliary statements needed in the proof of Theorem 4.1. Recall the definition of $\phi_{e,i}[k]$ in (A.7) and of $\mu_{\mathsf{A},k}, \mu_{\mathsf{B},k}, \sigma^2_{\mathsf{A},k}, \sigma^2_{\mathsf{B},k}$ in (A.14)–(A.17).

Notation for convergence rate of the Fourier phases. Denote for every $0 \leq k \leq d-1$,

$$\mathsf{A}_{M,k} \triangleq \frac{1}{\sqrt{M}}\sum_{i=0}^{M-1}\left|\mathsf{N}_i[k]\right|\sin\left(\phi_{e,i}[k]\right), \tag{B.1}$$

and,

$$\mathsf{B}_{M,k} \triangleq \frac{1}{M}\sum_{i=0}^{M-1}\left|\mathsf{N}_i[k]\right|\cos\left(\phi_{e,i}[k]\right). \tag{B.2}$$

Note that $\mathsf{A}_{M,k}$ is the imaginary part of the EfN estimator, multiplied by the phase of the template signal as defined in (A.11), but is normalized by $1/\sqrt{M}$ instead of $1/M$, to facilitate the analysis of the convergence rate. Similarly, $\mathsf{B}_{M,k}$ corresponds to the real part in (A.11). Additionally, we define the following Gaussian random variable Q_k

$$\mathsf{Q}_k \sim \mathcal{N}\left(0, \frac{\sigma^2_{\mathsf{A},k}}{\mu^2_{\mathsf{B},k}}\right), \tag{B.3}$$

for every $0 \leq k \leq d-1$.

The main results of this section. Recall that by the SLLN (A.13), the EfN estimator converges to,

$$\hat{\mathsf{X}}[k] \xrightarrow{\text{a.s.}} e^{j\phi_{\mathsf{X}}[k]} \cdot \mathbb{E}\left[|\mathsf{N}_1[k]| \cos\left(\phi_{e,1}[k]\right)\right] + j\mathbb{E}\left[|\mathsf{N}_1[k]| \sin\left(\phi_{e,1}[k]\right)\right] \tag{B.4}$$

$$= e^{j\phi_{\mathsf{X}}[k]} \left(\mu_{\mathsf{B},k} + j\mu_{\mathsf{A},k}\right). \tag{B.5}$$

In Sections B.1 and B.2, we prove that $\mu_{\mathsf{A},k} = 0$ and $\mu_{\mathsf{B},k} > 0$. These results, combined with (B.5), imply that the EfN estimator converges to a non-vanishing signal, and its Fourier phases converge those of Einstein as $M \to \infty$. In Sections B.3 and B.4, we analyze the convergence rate of the Fourier phases, first establishing convergence rate in distribution to Q_k, and then proving convergence rate in expectation.

B.1 Convergence of the Fourier phases

Lemma B.1 (Convergence of the Fourier phases)**.** *Recall the definition of $\phi_{e,i}[k]$ in (A.7). Then we have,*

$$\mu_{\mathsf{A},k} \triangleq \mathbb{E}\left[|\mathsf{N}_1[k]| \sin(\phi_{e,1}[k])\right] = 0, \tag{B.6}$$

for every $0 \le k \le d-1$.

Proof of Lemma B.1. Let $\mathsf{D}[k] \triangleq \phi_{\mathsf{X}}[k] - \phi_{\mathsf{N}_1}[k]$, and recall the definition of $\hat{\mathsf{R}}_i$ in (A.4):

$$\hat{\mathsf{R}}_i = \underset{0 \le r \le d-1}{\arg\max} \sum_{k=0}^{d-1} |\mathsf{X}[k]|\, |\mathsf{N}_i[k]| \cos\left(\frac{2\pi k r}{d} + \phi_{\mathsf{N}_i}[k] - \phi_{\mathsf{X}}[k]\right). \tag{B.7}$$

Note that $\hat{\mathsf{R}}_i$ is a function of

$$\hat{\mathsf{R}}_i = \hat{\mathsf{R}}_i\left(\{|\mathsf{N}_i[k]|\}_{k=0}^{d-1}, \{|\mathsf{X}[k]|\}_{k=0}^{d-1}, \{\phi_{\mathsf{N}_i}[k]\}_{k=0}^{d-1}, \{\phi_{\mathsf{X}}[k]\}_{k=0}^{d-1}\right), \tag{B.8}$$

and it depends on $\phi_{\mathsf{N}_i}[k]$ and $\phi_{\mathsf{X}}[k]$ only through $\mathsf{D}[k]$. Accordingly, viewing $\hat{\mathsf{R}}_1$ as a function of $\mathsf{D}[k]$, for fixed $\{|\mathsf{N}_i[k]|\}_{k=0}^{d-1}, \{|\mathsf{X}[k]|\}_{k=0}^{d-1}$, we have,

$$\hat{\mathsf{R}}_1\left(-\mathsf{D}[0], -\mathsf{D}[1], \ldots, -\mathsf{D}[d-1]\right) = -\hat{\mathsf{R}}_1\left(\mathsf{D}[0], \mathsf{D}[1], \ldots, \mathsf{D}[d-1]\right). \tag{B.9}$$

Namely, from symmetry arguments, by flipping the signs of all the phases, the location of the maximum flips its sign as well. Then, by the law of total expectation,

$$\mu_{\mathsf{A},k} = \mathbb{E}\left[|\mathsf{N}_1[k]| \sin\left(\frac{2\pi k}{d}\hat{\mathsf{R}}_1 + \phi_{\mathsf{N}_1}[k] - \phi_{\mathsf{X}}[k]\right)\right]$$

$$= \mathbb{E}\left\{|\mathsf{N}_1[k]| \cdot \mathbb{E}\left[\sin\left(\frac{2\pi k}{d}\hat{\mathsf{R}}_1 + \phi_{\mathsf{N}_1}[k] - \phi_{\mathsf{X}}[k]\right) \middle| \{|\mathsf{N}_1[k]|\}_{k=0}^{d-1}\right]\right\}. \tag{B.10}$$

The inner expectation in (B.10) is taken w.r.t. the uniform distribution randomness of the phases $\{\phi_{\mathsf{N}_1}[k]\}_{k=0}^{d-1} \in [-\pi, \pi)$. However, due to (B.9), and since the sine function is odd around zero, the integration in (B.10) nullifies. Therefore,

$$\mathbb{E}\left[\sin\left(\frac{2\pi k}{d}\hat{\mathsf{R}}_1 + \phi_{\mathsf{N}_1}[k] - \phi_{\mathsf{X}}[k]\right) \middle| \{|\mathsf{N}_1[k]|\}_{k=0}^{d-1}\right] = 0, \tag{B.11}$$

and thus $\mu_{\mathsf{A},k} = 0$. □

B.2 Convergence to non-vanishing signal

Proposition B.2 (Convergence to non-vanishing signal)**.** *Recall the definition of $\phi_{e,i}[k]$ in (A.7). Fix $d \in \mathbb{N}$, and assume that $\mathsf{X}[k] \neq 0$ for all $0 < k \leq d-1$. Then, for any $0 \leq k \leq d-1$,*

$$\mu_{\mathsf{B},k} \triangleq \mathbb{E}[|\mathsf{N}_1[k]| \cos(\phi_{e,1}[k])] > 0. \tag{B.12}$$

Proof of Proposition B.2. By the law of total expectation, we have,

$$\begin{aligned}\mathbb{E}[|\mathsf{N}_1[k]| \cos(\phi_{e,1}[k])] &= \mathbb{E}\left[|\mathsf{N}_1[k]| \cdot \mathbb{E}\left(\cos(\phi_{e,1}[k]) | \, \mathsf{N}_1[k]\right)\right] \\ &= \mathbb{E}\left[|\mathsf{N}_1[k]| \cdot \mathbb{E}\left(\cos\left(\frac{2\pi k \hat{\mathsf{R}}_1}{d} + \phi_{\mathsf{N}_1}[k] - \phi_{\mathsf{X}}[k]\right) \middle| \, \mathsf{N}_1[k]\right)\right]. \end{aligned} \tag{B.13}$$

More explicitly, we can write,

$$\begin{aligned}&\mathbb{E}[|\mathsf{N}_1[k]| \cos(\phi_{e,1}[k])] = \\ &\frac{1}{2\pi}\int_0^\infty \mathrm{d}n n f_{|\mathsf{N}_1[k]|}(n) \int_{-\pi}^{\pi} \mathrm{d}\varphi \mathbb{E}\left[\cos\left(\frac{2\pi k}{d}\hat{\mathsf{R}}_1 + \varphi\right) \middle| \, |\mathsf{N}_1[k]| = n, \phi_{\mathsf{N}_1}[k] = \phi_{\mathsf{X}}[k] + \varphi\right]. \end{aligned} \tag{B.14}$$

Now, note that the inner integral can be written as,

$$\begin{aligned}&\int_{-\pi}^{\pi} \mathrm{d}\varphi \mathbb{E}\left[\cos\left(\frac{2\pi k}{d}\hat{\mathsf{R}}_1 + \varphi\right) \middle| \, |\mathsf{N}_1[k]| = n, \phi_{\mathsf{N}_1}[k] = \phi_{\mathsf{X}}[k] + \varphi\right] \\ &\quad = \int_0^{\pi} \mathrm{d}\varphi \; \mathbb{E}\left[\cos\left(\frac{2\pi k}{d}\hat{\mathsf{R}}_1 + \varphi\right) \middle| \, |\mathsf{N}_1[k]| = n, \phi_{\mathsf{N}_1}[k] = \phi_{\mathsf{X}}[k] + \varphi\right] + \\ &\quad + \int_0^{\pi} \mathrm{d}\varphi \; \mathbb{E}\left[\cos\left(\frac{2\pi k}{d}\hat{\mathsf{R}}_1 + \varphi + \pi\right) \middle| \, |\mathsf{N}_1[k]| = n, \phi_{\mathsf{N}_1}[k] = \phi_{\mathsf{X}}[k] + \varphi + \pi\right]. \end{aligned} \tag{B.15}$$

Now, we apply Proposition A.5 on the integrands in (B.15). Using its notation, we define the Gaussian process:

$$\mathsf{S}^{(+)} = \mathsf{S}_1 | \mathsf{N}_1[k], \tag{B.16}$$

where the right-hand side is defined as in (A.18). By (A.19), the mean vector of $\mathsf{S}_1|\mathsf{N}_1[k]$ has a cosine trend, as assumed in Proposition A.5 in (A.59). Additionally, $\mathsf{S}_1|\mathsf{N}_1[k]$ is a Gaussian cyclo-stationary process, as described in (A.20). The final condition to verify is (A.60), which is satisfied by Lemma A.4.

Since the conditional distribution of $\hat{\mathsf{R}}_1$ given $\{|\mathsf{N}_1[k]| = n, \phi_{\mathsf{N}_1}[k] = \phi_{\mathsf{X}}[k] + \varphi\}$ matches that of $\hat{\mathsf{R}}^{(+)}$ in (A.57), and similarly, given $\{|\mathsf{N}_1[k]| = n, \phi_{\mathsf{N}_1}[k] = \phi_{\mathsf{X}}[k] + \varphi + \pi\}$, it matches $\hat{\mathsf{R}}^{(-)}$ in (A.58), the sum of the integrands on the right-hand side of (B.15) equals the left-hand side of (A.63). By Proposition A.5, this sum is positive for all $\varphi \in [0, \pi]$. Together with (B.14), this completes the proof of Proposition B.2.

□

B.3 Convergence rate in distribution of the Fourier phases

Proposition B.3 (Convergence in distribution of the Fourier phases)**.** *Fix $d \in \mathbb{N}$. Then, for any $0 \leq k \leq d-1$,*

$$\sqrt{M} \cdot \tan\left(\phi_{\hat{\mathsf{X}}}[k] - \phi_{\mathsf{X}}[k]\right) \xrightarrow{\mathcal{D}} \mathsf{Q}_k, \tag{B.17}$$

as $M \to \infty$, where Q_k is defined in (B.3).

Proof of Proposition B.3. Recall the definition of $\mathsf{A}_{M,k}, \mathsf{B}_{M,k}$ in (B.1), (B.2), of Q_k in (B.3), and of $\phi_{e,i}[k]$ in (A.7). Then, following (A.6), the left-hand-side of (B.17) is given by,

$$\sqrt{M} \cdot \tan\left[\phi_{\hat{\mathsf{X}}}[k] - \phi_{\mathsf{X}}[k]\right] = \frac{\frac{1}{\sqrt{M}} \sum_{i=0}^{M-1} |\mathsf{N}_i[k]| \sin\left(\phi_{e,i}[k]\right)}{\frac{1}{M} \sum_{i=0}^{M-1} |\mathsf{N}_i[k]| \cos\left(\phi_{e,i}[k]\right)} \triangleq \frac{\mathsf{A}_{M,k}}{\mathsf{B}_{M,k}}. \tag{B.18}$$

Since $\{\mathsf{N}_i\}_{i=0}^{M-1}$ is an i.i.d. sequence of random variables, and because each $\phi_{e,i}$ depends on N_i solely (in particular, independent of N_j, for $j \neq i$), we have that $\{|\mathsf{N}_i[k]| \sin\left(\phi_{e,i}[k]\right)\}_{i=0}^{M-1}$ and $\{|\mathsf{N}_i[k]| \cos\left(\phi_{e,i}[k]\right)\}_{i=0}^{M-1}$ are two sequences of i.i.d. random variables. Recall the definition of $\mu_{\mathsf{A},k}$, and $\sigma^2_{\mathsf{A},k}$, in (A.14), (A.16), respectively:

$$\mu_{\mathsf{A},k} \triangleq \mathbb{E}\left[|\mathsf{N}_1[k]| \sin(\phi_{e,1}[k])\right], \tag{B.19}$$

$$\sigma^2_{\mathsf{A},k} \triangleq \mathsf{Var}\left(|\mathsf{N}_1[k]| \sin(\phi_{e,1}[k])\right), \tag{B.20}$$

which are the mean value and variance of $\mathsf{A}_{M,k}$, as defined in (B.1). Then, by the CLT:

$$\left(\mathsf{A}_{M,k} - \sqrt{M}\mu_{\mathsf{A},k}\right) \xrightarrow{\mathcal{D}} \mathsf{A}_k, \tag{B.21}$$

where $\mathsf{A}_k \sim \mathcal{N}(0, \sigma^2_{\mathsf{A},k})$. In particular, by Lemma B.1, $\mu_{\mathsf{A},k} = 0$.

Next, we analyze the denominator in (B.18). Specifically, we already saw that $\{|\mathsf{N}_i[k]| \cos\left(\phi_{e,i}[k]\right)\}_{i=0}^{M-1}$ form a sequence of i.i.d. random variables, and thus by the SLLN we have $\mathsf{B}_{M,k} \xrightarrow{\text{a.s.}} \mu_{\mathsf{B},k}$, where,

$$\mu_{\mathsf{B},k} \triangleq \mathbb{E}\left[|\mathsf{N}_1[k]| \cos(\phi_{e,1}[k])\right]. \tag{B.22}$$

By Proposition B.2, $\mu_{\mathsf{B},k} > 0$. Thus, applying Slutsky's Theorem on the ratio $\frac{\mathsf{A}_{M,k}}{\mathsf{B}_{M,k}}$, we obtain,

$$\frac{\mathsf{A}_{M,k}}{\mathsf{B}_{M,k}} \xrightarrow{\mathcal{D}} \mathcal{N}\left(0, \frac{\sigma^2_{\mathsf{A},k}}{\mu^2_{\mathsf{B},k}}\right) = \mathsf{Q}_k, \tag{B.23}$$

which concludes the proof. □

B.4 Convergence rate in expectation of the Fourier phases

Proposition B.4 (Convergence rate of the Fourier phases)**.** *Recall the definitions of* $\mu_{\mathsf{B},k}$*, and* $\sigma^2_{\mathsf{A},k}$ *in* (A.15)*, and* (A.17)*, respectively. Assume that* $\mathsf{X}[k] \neq 0$*, for all* $0 < k \leq d-1$*. Then, as* $M \to \infty$*,*

$$\lim_{M\to\infty} \frac{\mathbb{E}|\phi_{\hat{\mathsf{X}}}[k] - \phi_{\mathsf{X}}[k]|^2}{1/M} = \frac{\sigma^2_{\mathsf{A},k}}{\mu^2_{\mathsf{B},k}}. \tag{B.24}$$

Proof of Proposition B.4. Recall the definitions of $\mathsf{A}_{M,k}$ and $\mathsf{B}_{M,k}$ in (B.1) and (B.2), respectively, and of Q_k in (B.3). Then, using the phase difference expression in (A.6), it follows that establishing (B.24) is equivalent to proving the following:

$$\lim_{M\to\infty} \frac{\mathbb{E}\left[\arctan^2\left(\frac{1}{\sqrt{M}}\frac{\mathsf{A}_{M,k}}{\mathsf{B}_{M,k}}\right)\right]}{\frac{1}{M}\mathbb{E}\left[\mathsf{Q}_k^2\right]} = 1, \tag{B.25}$$

for every $0 \leq k \leq d-1$. Recall by the definition of Q_k in (B.3) that $\mathbb{E}\left[\mathsf{Q}_k^2\right] = \sigma^2_{\mathsf{A},k}/\mu^2_{\mathsf{B},k}$, which is equivalent to the right-hand-side of (B.24).

For brevity, we fix k, and denote $\mathsf{A}_M = \mathsf{A}_{M,k}$, $\mathsf{B}_M = \mathsf{B}_{M,k}$, $\mu_{\mathsf{B}} = \mu_{\mathsf{B},k}$, $\sigma^2_{\mathsf{A}} = \sigma^2_{\mathsf{A},k}$. Using (A.6) it is clear that,

$$\sqrt{M} \cdot \tan\left[\phi_{\hat{\mathsf{X}}}[k] - \phi_{\mathsf{X}}[k]\right] = \frac{\frac{1}{\sqrt{M}}\sum_{i=0}^{M-1} |\mathsf{N}_i[k]| \sin\left(\phi_{e,i}[k]\right)}{\frac{1}{M}\sum_{i=0}^{M-1} |\mathsf{N}_i[k]| \cos\left(\phi_{e,i}[k]\right)} \triangleq \frac{\mathsf{A}_M}{\mathsf{B}_M}, \tag{B.26}$$

It is important to note that the denominator B_M can be zero with positive probability, implying that the expression in (B.26) may diverge with non-zero probability. Therefore, it is necessary to control the occurrence of such events. To this end, $\mathsf{B}_M \xrightarrow{\text{a.s.}} \mu_{\mathsf{B}}$, by SLLN (see Section A.2), where μ_{B} is defined in (A.15). Fix $0 < \epsilon < \mu_{\mathsf{B}}$, and proceed by decomposing as follows:

$$\begin{aligned}\mathbb{E}\left[\arctan^2\left(\frac{1}{\sqrt{M}}\frac{\mathsf{A}_M}{\mathsf{B}_M}\right)\right] &= \mathbb{E}\left[\arctan^2\left(\frac{1}{\sqrt{M}}\frac{\mathsf{A}_M}{\mathsf{B}_M}\right)\mathbb{1}_{|\mathsf{B}_M|>\epsilon}\right] \\ &\quad + \mathbb{E}\left[\arctan^2\left(\frac{1}{\sqrt{M}}\frac{\mathsf{A}_M}{\mathsf{B}_M}\right)\mathbb{1}_{|\mathsf{B}_M|<\epsilon}\right].\end{aligned} \tag{B.27}$$

The next lemma shows that the second term at the r.h.s. of (B.27) converges to zero with rate $O(1/M^2)$.

Lemma B.5. *The following inequality holds,*

$$\mathbb{E}\left[\arctan^2\left(\frac{1}{\sqrt{M}}\frac{\mathsf{A}_M}{\mathsf{B}_M}\right)\mathbb{1}_{|\mathsf{B}_M|<\epsilon}\right] \leq \frac{D}{M^2}, \tag{B.28}$$

for a finite $D > 0$*.*

In addition, we have the following asymptotic relation for the last term in (B.27).

Lemma B.6. *The following asymptotic relation hold,*

$$\lim_{M\to\infty} \frac{\mathbb{E}\left[\arctan^2\left(\frac{1}{\sqrt{M}}\frac{\mathsf{A}_M}{\mathsf{B}_M}\right)\mathbb{1}_{|\mathsf{B}_M|>\epsilon}\right]}{\frac{1}{M}\mathbb{E}\left[\mathsf{Q}_k^2\right]} = 1. \tag{B.29}$$

We prove these lemmas below. Substituting (B.28) and (B.29) in (B.27), leads to (B.25), and completing the proof of the proposition.

□

Proof of Lemma B.5. Since $\arctan(x) \leq \frac{\pi}{2}$, for any $x \in \mathbb{R}$, we have

$$\mathbb{E}\left[\arctan^2\left(\frac{1}{\sqrt{M}}\frac{\mathsf{A}_M}{\mathsf{B}_M}\right)\mathbb{1}_{|\mathsf{B}_M|<\epsilon}\right] \leq \frac{\pi^2}{4}\cdot\mathbb{E}\left[\mathbb{1}_{|\mathsf{B}_M|<\epsilon}\right] \tag{B.30}$$

$$\leq \frac{\pi^2}{4}\cdot\mathbb{P}\left(\mathsf{B}_M<\epsilon\right) \tag{B.31}$$

$$= \frac{\pi^2}{4}\cdot\mathbb{P}\left(\mathsf{B}_M-\mu_{\mathsf{B}}<\epsilon-\mu_{\mathsf{B}}\right) \tag{B.32}$$

$$\leq \frac{\pi^2}{4}\cdot\mathbb{P}\left(|\mathsf{B}_M-\mu_{\mathsf{B}}|>\mu_{\mathsf{B}}-\epsilon\right). \tag{B.33}$$

Let us denote the summand in the denominator in (B.26) by $V_i \triangleq |\mathsf{N}_i[k]|\cos\left(\phi_{e,i}[k]\right)$, for $0 \leq i \leq M-1$. Then, we note that,

$$\mathbb{E}(V_i^4) = \mathbb{E}\left[|\mathsf{N}_i[k]|\cos\left(\phi_{e,i}[k]\right)\right]^4 \leq \mathbb{E}\left[|\mathsf{N}_i[k]|\right]^4 < \infty. \tag{B.34}$$

Thus, by Chebyshev's inequality,

$$\mathbb{P}\left(|\mathsf{B}_M-\mu_{\mathsf{B}}|>\mu_{\mathsf{B}}-\epsilon\right) \leq \frac{\mathbb{E}\left[\mathsf{B_M}-\mu_{\mathsf{B}}\right]^4}{\left(\mu_{\mathsf{B}}-\epsilon\right)^4}. \tag{B.35}$$

Now, by the definition of B_M, we have,

$$\mathbb{E}\left[\mathsf{B}_M-\mu_{\mathsf{B}}\right]^4 = \frac{1}{M^4}\sum_{i,j,k,l=0}^{M-1}\mathbb{E}\left[\left(V_i-\mu_{\mathsf{B}}\right)\left(V_j-\mu_{\mathsf{B}}\right)\left(V_k-\mu_{\mathsf{B}}\right)\left(V_l-\mu_{\mathsf{B}}\right)\right] \tag{B.36}$$

$$= \frac{1}{M^4}\left[M\cdot\mathbb{E}\left[V_1-\mu_{\mathsf{B}}\right]^4+3M(M-1)\left(\mathbb{E}\left[V_1-\mu_{\mathsf{B}}\right]^2\right)^2\right]. \tag{B.37}$$

Therefore, it is evident that there exists a constant D_1, which depends on the second and fourth moments of V_1, such that,

$$\frac{\mathbb{E}\left[\mathsf{B_M}-\mu_{\mathsf{B}}\right]^4}{\left(\mu_{\mathsf{B}}-\epsilon\right)^4} \leq \frac{D_1}{\left(\mu_{\mathsf{B}}-\epsilon\right)^4 M^2}. \tag{B.38}$$

Thus, plugging (B.35) and (B.38) into (B.33) leads to,

$$\mathbb{E}\left[\arctan^2\left(\frac{1}{\sqrt{M}}\frac{\mathsf{A}_M}{\mathsf{B}_M}\right)\mathbb{1}_{|\mathsf{B}_M|<\epsilon}\right] \leq \frac{\pi^2}{4}\cdot\frac{D_1}{\left(\mu_{\mathsf{B}}-\epsilon\right)^4 M^2}. \tag{B.39}$$

Thus, the second term at the r.h.s. of (B.27) indeed converges to zero as $1/M^2$. □

Proof of Lemma B.6. We analyze the first term at the r.h.s. of (B.27). We will show that,

$$1 \leq \lim_{M\to\infty} \frac{\mathbb{E}\left[\arctan^2\left(\frac{1}{\sqrt{M}}\frac{\mathsf{A}_M}{\mathsf{B}_M}\right)\mathbb{1}_{|\mathsf{B}_M|>\epsilon}\right]}{\frac{1}{M}\mathbb{E}\left[\mathsf{Q}_k^2\right]} \leq \frac{\mu_{\mathsf{B}}^2}{\epsilon^2}. \tag{B.40}$$

As this is true for every $\epsilon < \mu_{\mathsf{B}}$, it would imply that,

$$\lim_{M\to\infty} \frac{\mathbb{E}\left[\arctan^2\left(\frac{1}{\sqrt{M}}\frac{\mathsf{A}_M}{\mathsf{B}_M}\right)\mathbb{1}_{|\mathsf{B}_M|>\epsilon}\right]}{\frac{1}{M}\mathbb{E}\left[\mathsf{Q}_k^2\right]} = 1. \tag{B.41}$$

First, due to the monotonicity of $\arctan^2(\cdot)$,

$$\mathbb{E}\left[\arctan^2\left(\frac{1}{\sqrt{M}}\frac{\mathsf{A}_M}{\mathsf{B}_M}\right)\mathbb{1}_{|\mathsf{B}_M|>\epsilon}\right] \leq \mathbb{E}\left[\arctan^2\left(\frac{\mathsf{A}_M}{\epsilon\sqrt{M}}\right)\right]. \tag{B.42}$$

We decompose the right-hand-side of (B.42) into two events, as follows,

$$\begin{aligned}\mathbb{E}\left[\arctan^2\left(\frac{\mathsf{A}_M}{\epsilon\sqrt{M}}\right)\right] &= \mathbb{E}\left[\arctan^2\left(\frac{\mathsf{A}_M}{\epsilon\sqrt{M}}\right)\mathbb{1}_{|\mathsf{A}_M|>\epsilon\sqrt{M}}\right] \\ &\quad + \mathbb{E}\left[\arctan^2\left(\frac{\mathsf{A}_M}{\epsilon\sqrt{M}}\right)\mathbb{1}_{|\mathsf{A}_M|<\epsilon\sqrt{M}}\right].\end{aligned} \tag{B.43}$$

By the SLLN, $\mathsf{A}_M/\sqrt{M} \xrightarrow{\text{a.s.}} \mu_{\mathsf{A}}$ (see Section A.2), where $\mu_{\mathsf{A}} = 0$ (Lemma B.1). In addition, by Proposition B.3, we have,

$$\mathsf{A}_M \xrightarrow{\mathcal{D}} \mathcal{N}\left(0, \sigma_{\mathsf{A}}^2\right), \tag{B.44}$$

by the CLT. Then, by arguments similar to those used in Lemma B.5, the first term on the right-hand side of (B.43) satisfies:

$$\mathbb{E}\left[\arctan^2\left(\frac{\mathsf{A}_M}{\epsilon\sqrt{M}}\right)\mathbb{1}_{|\mathsf{A}_M|>\epsilon\sqrt{M}}\right] \leq \tilde{D}/M^2. \tag{B.45}$$

Namely, the first term at the r.h.s. of (B.43) converges to zero with rate $O(\frac{1}{M^2})$.

For the last term in the right-hand-side of (B.43), we prove the following:

$$\lim_{M\to\infty} \frac{\mathbb{E}\left[\arctan^2\left(\frac{\mathsf{A}_M}{\epsilon\sqrt{M}}\right)\mathbb{1}_{|\mathsf{A}_M|<\epsilon\sqrt{M}}\right]}{\frac{1}{M}\mathbb{E}\left[\left(\frac{\mathsf{A}_M}{\epsilon}\right)^2\mathbb{1}_{|\mathsf{A}_M|<\epsilon\sqrt{M}}\right]} = 1. \tag{B.46}$$

Since $[\arctan(x)]^2/x^2 \to 1$ as $x \to 0$, it follows that the Taylor expansion of $\arctan(x)$ around $x = 0$, which holds for $|x| < 1$, and is applicable on the event $\left\{\mathsf{A}_M < \epsilon\sqrt{M}\right\}$:

$$\mathbb{E}\left[\arctan^2\left(\frac{\mathsf{A}_M}{\epsilon\sqrt{M}}\right)\mathbb{1}_{|\mathsf{A}_M|<\epsilon\sqrt{M}}\right] = \mathbb{E}\left[\sum_{k=0}^{\infty}(-1)^k\frac{\left[\frac{1}{\sqrt{M}}\frac{\mathsf{A}_M}{\epsilon}\right]^{2k+1}}{2k+1}\mathbb{1}_{|\mathsf{A}_M|<\epsilon\sqrt{M}}\right]^2. \tag{B.47}$$

The right-hand-side of (B.47) can be decomposed to,

$$\mathbb{E}\left[\sum_{k=0}^{\infty}(-1)^k\frac{\left[\frac{1}{\sqrt{M}}\frac{\mathsf{A}_M}{\epsilon}\right]^{2k+1}}{2k+1}\mathbb{1}_{|\mathsf{A}_M|<\epsilon\sqrt{M}}\right]^2=\frac{1}{M}\mathbb{E}\left[\left(\frac{\mathsf{A}_M}{\epsilon}\right)^2\mathbb{1}_{|\mathsf{A}_M|<\epsilon\sqrt{M}}\right]$$
$$+\sum_{(k_1,k_2)\neq(0,0)}\frac{(-1)^{k_1+k_2}}{(2k_1+1)(2k_2+1)}\mathbb{E}\left[\left(\frac{\mathsf{A}_M}{\sqrt{M}\epsilon}\right)^{(2k_1+2k_2+2)}\mathbb{1}_{|\mathsf{A}_M|<\epsilon\sqrt{M}}\right]. \tag{B.48}$$

Now, since the term at the left-hand-side of (B.48) as well as the first term at the right-hand-side of (B.48), are bounded for every M and converges to zero, then also the last term at the right-hand-side of (B.48) is bounded for every M and converge to zero as $M\to\infty$. Specifically, we note that the last term converges to zero with rate $1/M^2$, while the first term in the right-hand-side converges to zero with rate $1/M$. Thus, (B.46) is satisfied. Finally, we have,

$$\lim_{M\to\infty}\frac{\frac{1}{M}\mathbb{E}\left[\left(\frac{\mathsf{A}_M}{\epsilon}\right)^2\mathbb{1}_{|\mathsf{A}_M|<\epsilon\sqrt{M}}\right]}{\frac{1}{M}\mathbb{E}\left[\left(\frac{\mathsf{A}_M}{\epsilon}\right)^2\right]}=1-\lim_{M\to\infty}\frac{\frac{1}{M}\mathbb{E}\left[\left(\frac{\mathsf{A}_M}{\epsilon}\right)^2\mathbb{1}_{|\mathsf{A}_M|>\epsilon\sqrt{M}}\right]}{\frac{1}{M}\mathbb{E}\left[\left(\frac{\mathsf{A}_M}{\epsilon}\right)^2\right]}. \tag{B.49}$$

As the probability of the event $\left\{|\mathsf{A}_M|>\epsilon\sqrt{M}\right\}$ is $O\left(1/M^2\right)$, it follows that,

$$\lim_{M\to\infty}\frac{\frac{1}{M}\mathbb{E}\left[\left(\frac{\mathsf{A}_M}{\epsilon}\right)^2\mathbb{1}_{|\mathsf{A}_M|>\epsilon\sqrt{M}}\right]}{\frac{1}{M}\mathbb{E}\left[\left(\frac{\mathsf{A}_M}{\epsilon}\right)^2\right]}=0. \tag{B.50}$$

Then, following (B.49),(B.50), we have,

$$\lim_{M\to\infty}\frac{\frac{1}{M}\mathbb{E}\left[\left(\frac{\mathsf{A}_M}{\epsilon}\right)^2\mathbb{1}_{|\mathsf{A}_M|<\epsilon\sqrt{M}}\right]}{\frac{1}{M}\mathbb{E}\left[\left(\frac{\mathsf{A}_M}{\epsilon}\right)^2\right]}=1. \tag{B.51}$$

By definition $\sigma_{\mathsf{A}}^2=\mathbb{E}\left[\mathsf{A}_M^2\right]$. Therefore substitution (B.51) into (B.46) leads to

$$\lim_{M\to\infty}\frac{\mathbb{E}\left[\arctan^2\left(\frac{\mathsf{A}_M}{\epsilon\sqrt{M}}\right)\mathbb{1}_{|\mathsf{A}_M|<\epsilon\sqrt{M}}\right]}{\frac{1}{M}\frac{\sigma_{\mathsf{A}}^2}{\epsilon^2}}=1. \tag{B.52}$$

Substituting (B.45), and (B.52) into (B.43) results,

$$\lim_{M\to\infty}\frac{\mathbb{E}\left[\arctan^2\left(\frac{1}{\sqrt{M}}\frac{\mathsf{A}_M}{\epsilon}\right)\mathbb{1}_{|\mathsf{B}_M|>\epsilon}\right]}{\frac{1}{M}\mathbb{E}\left[\mathsf{Q}_k^2\right]}=\frac{\mu_{\mathsf{B}}^2}{\epsilon^2}, \tag{B.53}$$

where $\mathbb{E}\left[\mathsf{Q}_k^2\right]=\sigma_{\mathsf{A}}^2/\mu_{\mathsf{B}}^2$. Then, substituting (B.53) into (B.42) results,

$$\lim_{M\to\infty}\frac{\mathbb{E}\left[\arctan^2\left(\frac{1}{\sqrt{M}}\frac{\mathsf{A}_M}{\mathsf{B}_M}\right)\mathbb{1}_{|\mathsf{B}_M|>\epsilon}\right]}{\frac{1}{M}\mathbb{E}\left[\mathsf{Q}_k^2\right]}\leq\frac{\mu_{\mathsf{B}}^2}{\epsilon^2}. \tag{B.54}$$

which proves the upper bound in (B.40).

Similarly, since $\mathsf{B}_M \xrightarrow{\text{a.s.}} \mu_{\mathsf{B}}$, for any $\epsilon_2 > 0$, we have,

$$\lim_{M\to\infty} \mathbb{E}\left[\left(\frac{\mathsf{A}_M}{\mathsf{B}_M}\right)^2 \mathbb{1}_{\{\mathsf{B}_M>\epsilon\}}\right] \geq \lim_{M\to\infty} \mathbb{E}\left[\left(\frac{\mathsf{A}_M}{\mathsf{B}_M}\right)^2 \mathbb{1}_{\{\mathsf{B}_M>\epsilon\}\wedge\{\mathsf{B}_M<\mu_{\mathsf{B}}+\epsilon_2\}}\right] \tag{B.55}$$

$$\geq \frac{\sigma_A^2}{(\mu_{\mathsf{B}}+\epsilon_2)^2}. \tag{B.56}$$

Since (B.56) is true for every $\epsilon_2 > 0$, we get the lower bound in (B.40), which concludes the proof of (B.41). □

B.5 Proof of Theorem 4.1

Convergence of the Fourier magnitudes. We start with the convergence of the estimator's magnitudes. Recall the definition of $\phi_{e,i}[k]$ in (A.7). According to (A.13), we have,

$$\left|\hat{\mathsf{X}}[k]e^{-j\phi_{\mathsf{X}}[k]}\right| \xrightarrow{\text{a.s.}} \left|\mathbb{E}\left[|\mathsf{N}_1[k]|\cos\left(\phi_{e,1}[k]\right)\right] + j\mathbb{E}\left[|\mathsf{N}_1[k]|\sin\left(\phi_{e,1}[k]\right)\right]\right|, \tag{B.57}$$

Clearly, $\left|e^{-j\phi_{\mathsf{X}}[k]}\right| = 1$. By Lemma B.1,

$$\mu_{\mathsf{A},k} = \mathbb{E}\left[|\mathsf{N}_1[k]|\sin\left(\phi_{e,1}[k]\right)\right] = 0. \tag{B.58}$$

By Proposition B.2,

$$\mu_{\mathsf{B},k} = \mathbb{E}\left[|\mathsf{N}_1[k]|\cos\left(\phi_{e,1}[k]\right)\right] > 0. \tag{B.59}$$

Combining (B.57), (B.58), and (B.59) proves the convergence of the estimator's magnitudes of (4.3).

Convergence of the Fourier phases. Next, we prove the Fourier phases convergence of Theorem 4.1, starting with (4.1). To this end, recall (A.6)

$$\phi_{\hat{\mathsf{X}}}[k] - \phi_{\mathsf{X}}[k] = \arctan\left(\frac{\sum_{i=0}^{M-1}|\mathsf{N}_i[k]|\sin\left(\phi_{e,i}[k]\right)}{\sum_{i=0}^{M-1}|\mathsf{N}_i[k]|\cos\left(\phi_{e,i}[k]\right)}\right), \tag{B.60}$$

Using the continuous mapping theorem, it is evident that it suffices to prove that,

$$\frac{\sum_{i=0}^{M-1}|\mathsf{N}_i[k]|\sin\left(\phi_{e,i}[k]\right)}{\sum_{i=0}^{M-1}|\mathsf{N}_i[k]|\cos\left(\phi_{e,i}[k]\right)} \xrightarrow{\text{a.s.}} 0. \tag{B.61}$$

This, however, follows by applying the SLLN,

$$\frac{\sum_{i=0}^{M-1}|\mathsf{N}_i[k]|\sin\left(\phi_{e,i}[k]\right)}{\sum_{i=0}^{M-1}|\mathsf{N}_i[k]|\cos\left(\phi_{e,i}[k]\right)} \xrightarrow{\text{a.s.}} \frac{\mu_{\mathsf{A},k}}{\mu_{\mathsf{B},k}}, \tag{B.62}$$

where $\mu_{\mathsf{A},k} \triangleq \mathbb{E}\left[|\mathsf{N}_1[k]|\sin(\phi_{e,1}[k])\right]$ and $\mu_{\mathsf{B},k} \triangleq \mathbb{E}\left[|\mathsf{N}_1[k]|\cos(\phi_{e,1}[k])\right]$, defined in (A.14), and (A.15), respectively. By Lemma B.1, $\mu_{\mathsf{A},k} = 0$, while by Proposition B.2, we have that $\mu_{\mathsf{B},k} > 0$, and thus their ratio converges a.s. to zero by the continuous mapping theorem. Thus, we proved that $\phi_{\hat{\mathsf{X}}}[k] \xrightarrow{\text{a.s.}} \phi_{\mathsf{X}}[k]$.

Finally, we prove the convergence rate, given in (4.2). According to Proposition B.4, we have,

$$\lim_{M\to\infty} \frac{\mathbb{E}|\phi_{\hat{\mathsf{X}}}[k] - \phi_{\mathsf{X}}[k]|^2}{1/M} = \frac{\sigma^2_{\mathsf{A},k}}{\mu^2_{\mathsf{B},k}}, \tag{B.63}$$

which completes the proof of the Theorem.

Remark B.7. *Note that the above result implies that C_k in (4.2) is given by,*

$$C_k \triangleq \frac{\sigma^2_{\mathsf{A},k}}{\mu^2_{\mathsf{B},k}} = \frac{\mathbb{E}\left(\left[|\mathsf{N}_1[k]|\sin(\phi_{e,1}[k])\right]^2\right)}{\left(\mathbb{E}[|\mathsf{N}_1[k]|\cos(\phi_{e,1}[k])]\right)^2}. \tag{B.64}$$

C High-dimensional argmax asymptotics

In this section, we present a key proposition that plays a central role in the proof of Theorem 4.3.

Proposition C.1 (High-dimensional argmax asymptotics)**.** *Let $\mathsf{S} \sim \mathcal{N}(\mu, \Sigma)$ be a d-dimensional Gaussian random vector, with mean μ and a covariance matrix Σ. Assume that $|\Sigma_{ij}| = \rho_{|i-j|}$, where $\{\rho_\ell\}_{\ell\in\mathbb{N}}$ is a sequence of real-valued numbers such that $\rho_0 = 1$, $\rho_\ell < 1$, and $\rho_\ell \log \ell \to 0$, as $\ell \to \infty$. Assume also that $\sqrt{\log d} \cdot \max_{1\le i\le d} |\mu_i| \to 0$, as $d \to \infty$, and let $\hat{\mathsf{R}} \triangleq \arg\max\{\mathsf{S}_0, \mathsf{S}_1, \ldots, \mathsf{S}_{d-1}\}$. Then, for a bounded deterministic function $f : \{0, 1, \ldots, d-1\} \to \mathbb{R}$, we have,*

$$\lim_{d\to\infty} \mathbb{E}[f(\hat{\mathsf{R}})] - \frac{\sum_{r=0}^{d-1} f(r) e^{\mu_r a_d}}{\sum_{r=0}^{d-1} e^{\mu_r a_d}} = 0, \tag{C.1}$$

where $a_d \triangleq \sqrt{2\log d}$.

The proof of Proposition C.1 is based on an auxiliary result, which we prove in Section C.1. To state this result, we introduce some additional notation. Let $\mathsf{S}(r)$, for $r \in \{0, 1, \ldots, d-1\}$, be a discrete stochastic process. We define the function $h^{(\alpha)}(r)$ as follows,

$$h^{(\alpha)}(r) \triangleq \mathsf{S}(r) + \alpha f(r), \tag{C.2}$$

where $f(r)$ is a bounded deterministic function, and $\alpha \in \mathbb{R}$. We further define,

$$\mathsf{M}_d(\alpha) \triangleq \max_r h^{(\alpha)}(r), \tag{C.3}$$

and

$$\hat{\mathsf{R}}(\alpha) \triangleq \arg\max_r h^{(\alpha)}(r). \tag{C.4}$$

Note that $\mathsf{M}_d(a)$ and $\hat{\mathsf{R}}(a)$ are random variables. Finally, we denote $\hat{\mathsf{R}} \triangleq \hat{\mathsf{R}}(0)$. We have the following result, which is proved in Appendix C.1.

Lemma C.2. *The following holds,*

$$\mathbb{E}[f(\hat{\mathsf{R}})] = \frac{\mathrm{d}}{\mathrm{d}\alpha}\mathbb{E}[\mathsf{M}_d(\alpha)]\bigg|_{\alpha=0}. \tag{C.5}$$

Lemma C.2 implies that finding the expected value of $f(\hat{\mathsf{R}})$ is related directly to the derivative of the expected value of the maximum around zero. Thus, the problem of finding the expected value of $f(\hat{\mathsf{R}})$ is related to finding the expected value of the maximum of the stochastic process. In our case, S will be a Gaussian vector with mean given by (A.19) and a covariance matrix given by (A.20). Thus, our goal now is to find the expected value of the maximum of S. For this purpose, we will recall some well-known results on the maximum of Gaussian processes.

It is known that for an i.i.d. sequence of normally distributed random variables $\{\xi_n\}$, the asymptotic distribution of the maximum $\mathsf{M}_n \triangleq \max\{\xi_1, \xi_2, ..., \xi_n\}$ is the Gumbel distribution, i.e., for any $x \in \mathbb{R}$,

$$\mathbb{P}\left[a_n(\mathsf{M}_n - b_n) \leq x\right] \to e^{-e^{(-x)}}, \tag{C.6}$$

as $n \to \infty$, where,

$$a_n \triangleq \sqrt{2\log n} \tag{C.7}$$

and,

$$b_n \triangleq \sqrt{2\log n} - \frac{1}{2}\frac{\log\log n + \log 4\pi}{\sqrt{2\log n}}. \tag{C.8}$$

It turns out that the above convergence result remains valid even if the sequence $\{\xi_n\}$ is not independent and normally distributed. Specifically, as shown in [29, Theorem 6.2.1], a similar result holds for Gaussian random variables $\{\xi_n\}$ with a covariance matrix that decays such that $\lim_{n\to\infty} \rho_n \cdot \log n = 0$, and with a mean vector whose maximum value decays faster than $\lim_{n\to\infty} \max_{0\leq m\leq n-1} |\mu_m| \cdot \sqrt{\log n} = 0$. These conditions precisely match those specified in Theorem 4.3.

C.1 Proof of Lemma C.2

The proof technique of Lemma C.2 is similar to the technique used in [39, 31], but with a non-trivial adaption to the discrete case. To prove this lemma, we will first establish a deterministic counterpart of (C.5). Specifically, we define,

$$h^{(\alpha)}(r) \triangleq X(r) + \alpha f(r), \tag{C.9}$$

where $r \in \{0, 1, \ldots, d-1\}$. The functions $X : \{0, 1, \ldots, d-1\} \to \mathbb{R}$, and $f : \{0, 1, \ldots, d-1\} \to \mathbb{R}$ are assumed bounded and deterministic. We further assume that X is injective, i.e., for $z_i \neq z_j$, we have $X(z_i) \neq X(z_j)$. Define,

$$s(\alpha) \triangleq \max_r\{h^{(\alpha)}(r)\}, \tag{C.10}$$

and note that $s(\alpha)$ is well-defined over the supports of X and f, and it is a continuous function of α around $\alpha = 0$. Finally, we let,

$$Z_{\max}^{(\alpha)} \triangleq \arg\max_r \{h^{(\alpha)}(r)\}. \tag{C.11}$$

We have the following result.

Lemma C.3. *The following relation holds,*

$$\frac{\mathrm{d}}{\mathrm{d}\alpha} s(\alpha)\bigg|_{\alpha=0} = f(Z_{\max}^{(0)}). \tag{C.12}$$

Proof of Lemma C.3. Note that,

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}\alpha} s(\alpha)\bigg|_{\alpha=0} &= \lim_{\alpha\to 0} \frac{s(\alpha) - s(0)}{\alpha} \\ &= \lim_{\alpha\to 0} \frac{\max_r [X(r) + \alpha f(r)] - \max_r X(r)}{\alpha}.\end{aligned} \tag{C.13}$$

By the definition of $Z_{\max}^{(\alpha)}$, we have,

$$\max_r [X(r) + \alpha f(r)] = X(Z_{\max}^{(\alpha)}) + \alpha f(Z_{\max}^{(\alpha)}), \tag{C.14}$$

and

$$\max_r X(r) = X(Z_{\max}^{(0)}). \tag{C.15}$$

Now, the main observation here is that for a sufficiently small value of α around zero, we must have that $Z_{\max}^{(\alpha)}$ and $Z_{\max}^{(0)}$ equal because $Z_{\max}^{(0)}$ can take discrete values only, and it is unique. Thus, for $\alpha \cdot \max_r |f(r)| < \min_{i\neq j} |X(z_i) - X(z_j)|$, we have,

$$Z_{\max}^{(\alpha)} = Z_{\max}^{(0)}. \tag{C.16}$$

Combining (C.13)–(C.16) yields,

$$\frac{\mathrm{d}}{\mathrm{d}\alpha} s(\alpha)\bigg|_{\alpha=0} = \lim_{\alpha\to 0} \frac{X(Z_{\max}^{(\alpha)}) + \alpha f(Z_{\max}^{(\alpha)}) - X(Z_{\max}^{(0)})}{\alpha} \tag{C.17}$$

$$= f(Z_{\max}^{(0)}), \tag{C.18}$$

which concludes the proof. □

We are now in a position to prove Lemma C.2. Similarly to the deterministic case, we define the random function,

$$h^{(\alpha)}(r) = \mathsf{S}(r) + \alpha f(r), \tag{C.19}$$

where $\mathsf{S} : \{0, 1, \ldots, d-1\} \to \mathbb{R}$ is a discrete stochastic process, and f is a deterministic function. We assume that S has a continuous probability distribution without any single

point with a measure greater than 0. Using Lemma C.3, for each realization of $\mathsf{S}(r)$, such that $\mathsf{S}(r)$ is injective, we have,

$$f(\hat{\mathsf{R}}) = \frac{\mathrm{d}}{\mathrm{d}\alpha}\, \mathsf{M}_d(\alpha)\Big|_{\alpha=0}. \tag{C.20}$$

Under the assumption above of $\mathsf{S}(r)$, the measure of the set of events that S is not injective is zero. Therefore, the fact that $\frac{\mathsf{M}_d(\alpha)-\mathsf{M}_d(0)}{\alpha}$ is bounded (see, (C.34)) and (C.20), imply that,

$$\begin{aligned}
\mathbb{E}[f(\hat{\mathsf{R}})] &= \int f(\hat{\mathsf{R}})\mathrm{d}\mu \\
&= \int \frac{\mathrm{d}}{\mathrm{d}\alpha}\, \mathsf{M}_d(\alpha)\Big|_{\alpha=0} \mathrm{d}\mu \\
&= \int \lim_{\alpha\to 0}\left[\frac{\mathsf{M}_d(\alpha)-\mathsf{M}_d(0)}{\alpha}\right]\mathrm{d}\mu \\
&= \lim_{\alpha\to 0}\int \left[\frac{\mathsf{M}_d(\alpha)-\mathsf{M}_d(0)}{\alpha}\right]\mathrm{d}\mu \\
&= \frac{\mathrm{d}}{\mathrm{d}\alpha}\, \mathbb{E}\mathsf{M}_d(\alpha)\Big|_{\alpha=0},
\end{aligned} \tag{C.21}$$

which concludes the proof.

C.2 Proof of Proposition C.1

Conditioned on $\mathsf{N}[k]$, the Gaussian vector S (see, (A.19) and (A.20)) can be represented as,

$$\mathsf{S}|\mathsf{N}[k] = \mathsf{Z} + \mu, \tag{C.22}$$

where Z is a zero mean Gaussian random vector with covariance matrix given by (A.20) and μ is given by (A.19). Define,

$$h^{(\alpha)}(r) \triangleq \mathsf{Z}(r) + \mu(r) + \alpha f(r), \tag{C.23}$$

where we use the same notations as in Lemma C.2. Then, using Lemma C.2,

$$\mathbb{E}[f(\hat{\mathsf{R}})] = \frac{\mathrm{d}}{\mathrm{d}\alpha}\mathbb{E}\mathsf{M}_d(\alpha)\Big|_{\alpha=0}, \tag{C.24}$$

where $\mathsf{M}_d(\alpha) = \max_r \{\mathsf{Z}(r) + \mu(r) + \alpha f(r)\}$. Therefore, our goal is now to find the derivative of $\mathbb{E}\mathsf{M}_d(\alpha)$.

Using [29, Theorem 6.2.1], under the assumptions of Proposition C.1, for a sufficiently small value of α such that $\lim_{d\to\infty} |\alpha| \max_r |f(r)| \cdot \sqrt{\log d} = 0$, we have for any $x \geq 0$,

$$\lim_{d\to\infty} \mathbb{P}\left[a_d(\mathsf{M}_d(\alpha) - b_d - m_d^{\star}(\alpha)) \leq x\right] = e^{-e^{(-x)}}, \tag{C.25}$$

where a_d and b_d are given in (C.7) and (C.8), respectively, and

$$m_d^\star(\alpha) \triangleq a_d^{-1} \log \left(d^{-1} \sum_{i=0}^{d-1} e^{a_d(\mu_i + \alpha f(i))} \right). \tag{C.26}$$

For brevity, we denote,

$$\mathsf{T}_d(\alpha) \triangleq a_d \cdot [\mathsf{M}_d(\alpha) - b_d - m_d^\star(\alpha)], \tag{C.27}$$

and we note that,

$$\mathsf{T}_d(\alpha) - \mathsf{T}_d(0) = a_d[(\mathsf{M}_d(\alpha) - m_d^\star(\alpha)) - (\mathsf{M}_d(0) - m_d^\star(0))], \tag{C.28}$$

and so,

$$\Delta_d(\alpha) \triangleq \frac{1}{a_d} \frac{\mathsf{T}_d(\alpha) - \mathsf{T}_d(0)}{\alpha} = \frac{\mathsf{M}_d(\alpha) - \mathsf{M}_d(0)}{\alpha} - \frac{m_d^\star(\alpha) - m_d^\star(0)}{\alpha}, \tag{C.29}$$

for any $\alpha \neq 0$. The following result shows $\Delta_d(\alpha)$ converges zero in the $\mathcal{L}^1$ sense.

Lemma C.4. *For any $\alpha \neq 0$,*

$$\lim_{d \to \infty} |\Delta_d(\alpha)| = 0, \tag{C.30}$$

i.e., $\Delta_d(\alpha) \xrightarrow{\mathcal{L}^1} 0$, as $d \to \infty$.

Proof of Lemma C.4. To prove (C.30), we will first show that $\Delta_d(\alpha)$ converges to zero in probability. Because $\Delta_d(\alpha)$ is uniformly integrable, this is sufficient for the desired $\mathcal{L}^1$ convergence above. Specifically, recall from (C.25) that $\mathsf{T}_d(\alpha)$ converges in distribution to the Gumbel random variable Gum with location zero and unit scale, i.e., $\mathsf{T}_d(\alpha) \xrightarrow{\mathcal{D}} \mathsf{Gum}$, as $d \to \infty$. Furthermore, it is clear that $\frac{1}{a_d} = \frac{1}{\sqrt{2 \log d}} \to 0$, as $d \to \infty$. Thus, Slutsky's theorem [49] implies that,

$$\frac{\mathsf{T}_d(\alpha)}{a_d} \xrightarrow{\mathcal{D}} 0. \tag{C.31}$$

It is known that convergence in distribution to a constant implies also convergence in probability to the same constant [20], and thus,

$$\frac{\mathsf{T}_d(\alpha)}{a_d} \xrightarrow{\mathcal{P}} 0. \tag{C.32}$$

Therefore, the above result together with the continuous mapping theorem [20] implies that,

$$\Delta_d(\alpha) \xrightarrow{\mathcal{P}} 0, \tag{C.33}$$

for every $\alpha \neq 0$.

Next, we show that $\Delta_d(\alpha)$ is bounded with probability one. Indeed, by the definition of $\mathsf{M}_d(\alpha)$ in (C.3), we have,

$$\left|\frac{\mathsf{M}_d(\alpha)-\mathsf{M}_d(0)}{\alpha}\right| \leq \max_{0\leq r\leq d-1}|f(r)| < \mathsf{C} < \infty, \tag{C.34}$$

for some $\mathsf{C} > 0$, where we have used the fact that f is bounded. Furthermore, note that,

$$\frac{\mathrm{d}}{\mathrm{d}\alpha}m_d^\star(\alpha) = \frac{\sum_{i=0}^{d-1} f(i)\exp\{a_d(\mu_i+\alpha f(i))}{\sum_{i=0}^{d-1}\exp\{a_d(\mu_i+\alpha f(i))\}}, \tag{C.35}$$

which is bounded because,

$$\left|\frac{\sum_{i=0}^{d-1} f(i)\exp\{a_d(\mu_i+\alpha f(i))}{\sum_{i=0}^{d-1}\exp\{a_d(\mu_i+\alpha f(i))\}}\right| \leq \max_{0\leq r\leq d-1}|f(r)| < \mathsf{C} < \infty. \tag{C.36}$$

Combining (C.29), (C.34) and (C.36), leads to,

$$|\Delta_d(\alpha)| \leq \left|\frac{\mathsf{M}_d(\alpha)-\mathsf{M}_d(0)}{\alpha}\right| + \left|\frac{m_d^\star(\alpha)-m_d^\star(0)}{\alpha}\right| \leq 2\max_{0\leq r\leq d-1}|f(r)| < \infty. \tag{C.37}$$

Now, since $\Delta_d(\alpha)$ is bounded, it is also uniformly integrable, and thus when combined with (C.33) we may conclude that,

$$\Delta_d(\alpha) \xrightarrow{\mathcal{L}^1} 0, \tag{C.38}$$

as claimed. □

We continue with the proof of Proposition C.1. First, we show that,

$$\lim_{d\to\infty}\lim_{\alpha\to 0}\mathbb{E}\left[\Delta_d(\alpha)\right] = \lim_{\alpha\to 0}\lim_{d\to\infty}\mathbb{E}\left[\Delta_d(\alpha)\right]. \tag{C.39}$$

Indeed, note that,

$$\lim_{d\to\infty}\lim_{\alpha\to 0}\mathbb{E}\left[\Delta_d(\alpha)\right] = \lim_{d\to\infty}\lim_{\alpha\to 0}\int\left[\frac{\mathsf{T}_d(\alpha)-\mathsf{T}_d(0)}{\alpha}\right]\mathrm{d}\mu, \tag{C.40}$$

where $\mathrm{d}\mu$ is the probability measure associated with T_d. From (C.37) we know that $\Delta_d(\alpha)$ is bounded. Thus, applying the dominated convergence theorem, we obtain,

$$\lim_{d\to\infty}\lim_{\alpha\to 0}\int\left[\frac{\mathsf{T}_d(\alpha)-\mathsf{T}_d(0)}{\alpha}\right]\mathrm{d}\mu = \lim_{d\to\infty}\int\lim_{\alpha\to 0}\left[\frac{\mathsf{T}_d(\alpha)-\mathsf{T}_d(0)}{\alpha}\right]\mathrm{d}\mu. \tag{C.41}$$

Since the integral at the right-hand-side of (C.41) is finite and bounded for each value of α, and for each value of d, the order of the limits can be exchanged, thus leading to (C.39). Therefore, from (C.29) and (C.39), we have,

$$\lim_{\alpha\to 0}\lim_{d\to\infty}\mathbb{E}\left[\Delta_d(\alpha)\right] = \lim_{d\to\infty}\lim_{\alpha\to 0}\left[\frac{\mathbb{E}[\mathsf{M}_d(\alpha)-\mathsf{M}_d(0)]}{\alpha} - \frac{[m_d^\star(\alpha)-m_d^\star(0)]}{\alpha}\right] \tag{C.42}$$

$$= \lim_{d\to\infty} \left[\frac{\mathrm{d}}{\mathrm{d}\alpha} \mathbb{E}\mathsf{M}_d(\alpha) - \frac{\mathrm{d}}{\mathrm{d}\alpha} m_d^\star(\alpha) \right] \Bigg|_{\alpha=0}. \tag{C.43}$$

Now, Lemma C.4 implies that the left-hand-side of (C.42) nullifies, and thus,

$$\lim_{d\to\infty} \left[\frac{\mathrm{d}}{\mathrm{d}\alpha} \mathbb{E}\mathsf{M}_d(\alpha) - \frac{\mathrm{d}}{\mathrm{d}\alpha} m_d^\star(\alpha) \right] \Bigg|_{\alpha=0} = 0. \tag{C.44}$$

Finally, combining (C.35) and (C.44), we obtain (C.1), which concludes the proof.

D Proof of Theorem 4.3

Remark on notation. In this section, we omit the dependence on $0 \le i \le M-1$, where this is clear from the context, e.g., $\mathsf{N}_i = \mathsf{N}$ and $\hat{\mathsf{R}}_i = \hat{\mathsf{R}}$. In addition, to streamline notation we often omit the explicit dependence on the signal length and write x (and the associated quantities) in place of $x^{(d)}$ whenever the dimension d is understood from the context.

D.1 Notations and auxiliary results

First, we introduce notation and present several auxiliary results that support Theorem 4.3. Recall the definition of $\widetilde{\mathsf{X}}_k[\ell]$ from (A.21). We define:

$$\sigma_k^2 = \frac{\sigma^2}{2} \cdot \sum_{\ell=0}^{d-1} |\widetilde{\mathsf{X}}_k[\ell]|^2, \tag{D.1}$$

which corresponds to the diagonal entries of $\mathsf{\Sigma}_k[r,s]$ in (A.20). In addition, recall the vector S as defined in (A.5), and introduce the normalized vector:

$$\tilde{\mathsf{S}}_k = \mathsf{S}/\sigma_k, \tag{D.2}$$

where σ_k is given in (D.1). Then, by Lemma A.1, we have:

$$\widetilde{\mathsf{S}}_k | \mathsf{N}[k] \sim \mathcal{N}(\widetilde{\mu}_k, \widetilde{\mathsf{\Sigma}}_k), \tag{D.3}$$

where the mean and covariance are given by:

$$\widetilde{\mu}_k[r] \triangleq 2\sigma_k^{-1} \, |\mathsf{X}[k]| \, |\mathsf{N}_i[k]| \cos\left(\frac{2\pi k r}{d} + \phi_{\mathsf{N}_i}[k] - \phi_{\mathsf{X}}[k] \right) \tag{D.4}$$

for $0 \le r \le d-1$, and

$$\widetilde{\mathsf{\Sigma}}_k[r,s] \triangleq \frac{\sum_{\ell=0}^{d-1} |\widetilde{\mathsf{X}}_k[\ell]|^2 \cos\left(\frac{2\pi\ell}{d}(r-s)\right)}{\sum_{\ell=0}^{d-1} |\widetilde{\mathsf{X}}_k[\ell]|^2}, \tag{D.5}$$

for $0 \le r, s \le d-1$. Note in particular that normalizing S by σ_k ensures that the diagonal entries of $\widetilde{\mathsf{\Sigma}}_k$ are equal to one.

Recall that we assume the template vector is normalized, i.e., $\sum_{\ell=0}^{d-1} |\mathsf{X}[\ell]|^2 = 1$. Under this assumption, we have the following lemma.

Lemma D.1. *Suppose the conditions of Assumption 4.2 hold. Then, for all $k \in \mathbb{N}$, the following limits hold:*

$$\lim_{d\to\infty} \sum_{\ell=0}^{d-1} \left| 2\left|\mathsf{X}[\ell]\right|^2 - \left|\widetilde{\mathsf{X}}_k[\ell]\right|^2 \right| = 0, \tag{D.6}$$

and,

$$\lim_{d\to\infty} \sigma_k^2 = \sigma^2. \tag{D.7}$$

Proof of Lemma D.1. From the definition of $\widetilde{\mathsf{X}}_k[\ell]$ in (A.21), we have:

$$2\left|\mathsf{X}[\ell]\right|^2 - \left|\widetilde{\mathsf{X}}_k[\ell]\right|^2 = \begin{cases} 2\left|\mathsf{X}[\ell]\right|^2 & \text{if} \ell = k, d-k, \\ \left|\mathsf{X}[\ell]\right|^2 & \text{if} \ell = 0, d/2, \\ 0 & \text{otherwise.} \end{cases} \tag{D.8}$$

According to Assumption 4.2, we have,

$$\lim_{d\to\infty} \left\{ \max_{0<k\leq d-1} \left\{\left|\mathsf{X}[k]\right|\right\} \cdot \sqrt{\log d} \right\} = 0, \tag{D.9}$$

and $\mathsf{X}[0] = 0$, Recall that the template vector is normalized, i.e., $\sum_{\ell=0}^{d-1} \left|\mathsf{X}[\ell]\right|^2 = 1$, which implies $\left|\mathsf{X}[\ell]\right|^2 \leq \left|\mathsf{X}[\ell]\right|$ for all $0 \leq \ell \leq d-1$.

Combining (D.8)–(D.9), results,

$$\lim_{d\to\infty} \sum_{\ell=0}^{d-1} \left[2\left|\mathsf{X}[\ell]\right|^2 - \left|\widetilde{\mathsf{X}}_k[\ell]\right|^2 \right] = \lim_{d\to\infty} \left[\left|\mathsf{X}[0]\right|^2 + \left|\mathsf{X}[d/2]\right|^2 + 4\left|\mathsf{X}[k]\right|^2 \right] = 0, \tag{D.10}$$

which proves (D.6).

To establish (D.7), observe that

$$\lim_{d\to\infty} \sigma_k^2 = \frac{\sigma^2}{2} \cdot \lim_{d\to\infty} \sum_{\ell=0}^{d-1} \left|\widetilde{\mathsf{X}}_k[\ell]\right|^2 = \sigma^2, \tag{D.11}$$

where the final equality follows from (D.6). This completes the proof of the lemma. □

We now state a lemma showing that the entries of the covariance matrix $\widetilde{\Sigma}_k[r,s]$ satisfy the conditions of Proposition C.1.

Lemma D.2. *Suppose the conditions of Assumption 4.2 hold. Define,*

$$\rho_{|r-s|} \triangleq \left|\widetilde{\Sigma}_k\left[r,s\right]\right|, \tag{D.12}$$

for $\widetilde{\Sigma}_k\left[r,s\right]$ defined in (D.5). Then, $\rho_0 = 1$, and

$$\rho_n \log\left(n\right) \to 0. \tag{D.13}$$

That is, the covariance matrix $\widetilde{\Sigma}_k[r,s]$ satisfies the conditions required by Proposition C.1.

Proof of Lemma D.2. From the definition of the covariance matrix of $\widetilde{\mathsf{S}}_k|\mathsf{N}[k]$ in (D.5), we observe that it is circulant and fully characterized by its eigenvalues $|\widetilde{\mathsf{X}}_k[\ell]|^2$ (see (A.21)) for $0 \leq \ell \leq d-1$. Due to the normalization by σ_k, the covariance matrix is normalized such that its diagonal entries equal one, i.e., $\rho_0 = 1$.

It remains to show that the off-diagonal elements decay sufficiently fast, namely, $\rho_m \log(m) \to 0$, for $m \to \infty$.

Using the definition of ρ from (D.12), we can write:

$$\rho_m = \widetilde{\Sigma}_k[r, r-m] = \frac{\sum_{\ell=0}^{d-1} |\widetilde{\mathsf{X}}_k[\ell]|^2 \cos\left(\frac{2\pi\ell}{d} m\right)}{\sum_{\ell=0}^{d-1} |\widetilde{\mathsf{X}}_k[\ell]|^2}. \tag{D.14}$$

As $d \to \infty$, the denominator in (D.14) converges to 2 by Lemma D.1. The numerator corresponds to the DFT of the sequence $|\widetilde{\mathsf{X}}_k[\ell]|^2$,

$$\sum_{\ell=0}^{d-1} |\widetilde{\mathsf{X}}_k[\ell]|^2 \cos\left(\frac{2\pi\ell}{d} m\right) = \mathcal{F}\left\{|\widetilde{\mathsf{X}}_k|^2\right\}[m]. \tag{D.15}$$

By Lemma D.1, for each fixed k and for all $m \in \{1, \ldots, d-1\}$,

$$\mathcal{F}\big(|\widetilde{\mathsf{X}}_k|^2\big)[m] = 2\,\mathcal{F}\big(|\mathsf{X}|^2\big)[m] + \Delta_{k,d}[m], \tag{D.16}$$

where the error term $\Delta_{k,d}[m]$ is supported only on the modified frequencies $\ell \in \{0, d/2, k, d-k\}$ and satisfies the uniform bound

$$\sup_{1 \leq m \leq d-1} |\Delta_{k,d}[m]| \leq |\mathsf{X}[0]|^2 + |\mathsf{X}[d/2]|^2 + 4|\mathsf{X}[k]|^2 \xrightarrow[d\to\infty]{} 0, \tag{D.17}$$

using Assumption 4.2(2) and $|\mathsf{X}[0]| = 0$ from Assumption 4.2(3).

Moreover, recalling that $\mathsf{R}_{\mathsf{XX}} = \mathcal{F}\big(|\mathsf{X}|^2\big)$, Assumption 4.2(1) gives

$$\max_{1 \leq m \leq d-1} \log(d) \left|\mathcal{F}\big(|\mathsf{X}|^2\big)[m]\right| = \max_{1 \leq m \leq d-1} \log(d) \left|\mathsf{R}_{\mathsf{XX}}[m]\right| \xrightarrow[d\to\infty]{} 0. \tag{D.18}$$

Combining (D.16)–(D.17) with (D.18) yields

$$\max_{1 \leq m \leq d-1} \log(d) \left|\mathcal{F}\big(|\widetilde{\mathsf{X}}_k|^2\big)[m]\right| \xrightarrow[d\to\infty]{} 0, \tag{D.19}$$

and together with (D.15) this implies the desired decay condition for the off-diagonal correlations ρ_m, completing the proof. □

D.2 High-dimensional limits

We now present a central result, which plays a key role in the proof of Theorem 4.3. To state the result, we first define the functions:

$$f_1(r) \triangleq |\mathsf{N}[k]| \cos\left(\frac{2\pi k}{d} r + \phi_{\mathsf{N}}[k] - \phi_{\mathsf{X}}[k]\right), \tag{D.20}$$

and

$$f_2(r) \triangleq |\mathsf{N}[k]|^2 \sin^2\left(\frac{2\pi k}{d} r + \phi_{\mathsf{N}}[k] - \phi_{\mathsf{X}}[k]\right), \tag{D.21}$$

for $0 \leq r \leq d-1$. Note that f_1 and f_2 correspond to the terms appearing in the expectation in the denominator and numerator of (B.64), respectively.

Proposition D.3 (High-dimensional limits for $f_1(\hat{\mathsf{R}})$ and $f_2(\hat{\mathsf{R}})$). *Assume the template signal x satisfies Assumption 4.2, and that its DFT coefficients are non-vanishing, i.e., $\mathsf{X}[k] \neq 0$, for all $0 < k \leq d-1$. Let $f_1(\hat{\mathsf{R}})$ and $f_2(\hat{\mathsf{R}})$ be defined as in (D.20) and (D.21), with $\hat{\mathsf{R}}$ defined in (2.3). Then, as $d \to \infty$, their expected values satisfy:*

$$\lim_{d\to\infty} \frac{1}{a_d \cdot |\mathsf{X}[k]|} \mathbb{E}[f_1(\hat{\mathsf{R}})] = \sigma, \tag{D.22}$$

and

$$\lim_{d\to\infty} \mathbb{E}[f_2(\hat{\mathsf{R}})] = \frac{\sigma^2}{2}, \tag{D.23}$$

where $a_d \triangleq \sqrt{2\log d}$.

The proof of Proposition D.3 builds on Proposition C.1 and the auxiliary Lemmas D.1–D.2.

Proof of Proposition D.3. Our goal is to prove (D.22) and (D.23). By the law of total expectation, we have,

$$\frac{1}{a_d}\mathbb{E}[f_1(\hat{\mathsf{R}})] = \frac{1}{a_d}\mathbb{E}\left[\mathbb{E}\left[f_1(\hat{\mathsf{R}})\middle|\, \mathsf{N}[k]\right]\right], \tag{D.24}$$

and

$$\mathbb{E}[f_2(\hat{\mathsf{R}})] = \mathbb{E}\left[\mathbb{E}\left[f_2(\hat{\mathsf{R}})\middle|\, \mathsf{N}[k]\right]\right]. \tag{D.25}$$

Accordingly, we will prove

$$\frac{\sigma_k}{a_d \cdot |\mathsf{X}[k]|}\mathbb{E}\left[f_1(\hat{\mathsf{R}})\middle|\, \mathsf{N}[k]\right] \xrightarrow{\mathcal{L}^1} |\mathsf{N}[k]|^2, \tag{D.26}$$

and,

$$\mathbb{E}\left[f_2(\hat{\mathsf{R}})\middle|\, \mathsf{N}[k]\right] \xrightarrow{\mathcal{L}^1} \frac{1}{2}|\mathsf{N}[k]|^2, \tag{D.27}$$

which would yield the desired result.

To proceed, we apply Proposition C.1. Recall the definition of the vector $\tilde{\mathsf{S}}_k$ given in (D.2). Conditioned on $\mathsf{N}[k]$, this vector follows a Gaussian distribution with mean $\widetilde{\mu}_k[r]$ (as defined in (D.4)) and covariance matrix $\widetilde{\Sigma}_k[r,s]$ (defined in (D.5)). We assume that both the mean and the covariance satisfy Assumption 4.2, and we assert that they also meet the criteria of Proposition C.1. Indeed, observe the following:

1. **The mean term.** By Assumption 4.2, we have $|\mathsf{X}[k]|\sqrt{\log d} \to 0$, as $d \to \infty$, for every $0 \le k \le d-1$, implying that $\sqrt{\log d}\max|\widetilde{\mu}_k[r]| \to 0$, where the $|\mathsf{N}[k]|$ term in $\widetilde{\mu}_k[r]$ is finite and independent of d.

2. **The covariance term.** By Lemma D.2, the covariance matrix $\widetilde{\Sigma}_k[r,s]$ satisfies the conditions of Proposition C.1.

We apply Proposition C.1 and the result in (C.1) to the functions $f_1(\hat{\mathsf{R}})$ and $f_2(\hat{\mathsf{R}})$, with respect to the Gaussian vector $\widetilde{\mathsf{S}}_k|\mathsf{N}[k]$ (D.3). Observe that

$$\max_{0\le r\le d-1} |\widetilde{\mu}_k[r]| = 2\sigma_k^{-1}\,|\mathsf{X}[k]|\,|\mathsf{N}[k]|\,, \tag{D.28}$$

and,

$$f_1(\hat{\mathsf{R}}) = |\mathsf{N}[k]|\cos\left(\frac{2\pi k}{d}\hat{\mathsf{R}} + \phi_{\mathsf{N}}[k] - \phi_{\mathsf{X}}[k]\right) = \frac{\sigma_k}{2\,|\mathsf{X}[k]|}\widetilde{\mu}_k[\hat{\mathsf{R}}]. \tag{D.29}$$

Given that the assumptions of Proposition C.1 are satisfied, it follows that

$$\mathbb{E}\left[f_1(\hat{\mathsf{R}})\middle|\,\mathsf{N}[k]\right] - \frac{\sum_{r=0}^{d-1} f_1(r)e^{\widetilde{\mu}_k[r]a_d}}{\sum_{r=0}^{d-1} e^{\widetilde{\mu}_k[r]a_d}} \xrightarrow{\text{a.s.}} 0, \tag{D.30}$$

and,

$$\mathbb{E}\left[f_2(\hat{\mathsf{R}})\middle|\,\mathsf{N}[k]\right] - \frac{\sum_{r=0}^{d-1} f_2(r)e^{\widetilde{\mu}_k[r]a_d}}{\sum_{r=0}^{d-1} e^{\widetilde{\mu}_k[r]a_d}} \xrightarrow{\text{a.s.}} 0. \tag{D.31}$$

Next, we evaluate the terms at the left-hand-side of (D.30) and (D.31).

Proof of (D.26) and (D.22). We begin by proving that

$$\frac{1}{2a_d}\frac{\sum_{r=0}^{d-1}\widetilde{\mu}_k[r]\exp\{\widetilde{\mu}_k[r]a_d\}}{\sum_{r=0}^{d-1}\exp\{\widetilde{\mu}_k[r]a_d\}}\frac{\sigma_k^2}{|\mathsf{X}[k]|^2} \xrightarrow{\text{a.s.}} |\mathsf{N}[k]|^2\,. \tag{D.32}$$

From the definition of $f_1(r)$, it follows that

$$\sum_{r=0}^{d-1} f_1(r) = 0, \tag{D.33}$$

almost surely. Additionally, from the definition of $\widetilde{\mu}_k[r]$,

$$\widetilde{\mu}_k[r]a_d = 2\sigma_k^{-1}a_d\,|\mathsf{X}[k]|\,|\mathsf{N}[k]|\cos\left(\frac{2\pi k}{d}r + \phi_{\mathsf{N}}[k] - \phi_{\mathsf{X}}[k]\right). \tag{D.34}$$

By Assumption 4.2, we have $a_d\,|\mathsf{X}[k]| \to 0$ as $d \to \infty$. Thus, from (D.34) and the continuous mapping theorem, we obtain

$$\widetilde{\mu}_k[r]a_d \xrightarrow{\text{a.s.}} 0. \tag{D.35}$$

Since $\sum_{r=0}^{d-1} \widetilde{\mu}_k[r] = 0$ almost surely (from (D.33)), and applying the continuous mapping theorem along with (D.35), we deduce that

$$\frac{\sum_{r=0}^{d-1} \widetilde{\mu}_k[r] \exp\{\widetilde{\mu}_k[r] a_d\}}{a_d \sum_{r=0}^{d-1} [\widetilde{\mu}_k[r]]^2} \xrightarrow{\text{a.s.}} 1. \tag{D.36}$$

Similarly, applying the continuous mapping theorem using (D.35), we get

$$\frac{\sum_{r=0}^{d-1} \exp\{\widetilde{\mu}_k[r] a_d\}}{d} \xrightarrow{\text{a.s.}} 1. \tag{D.37}$$

Next, from the definition of $\widetilde{\mu}_k[r]$,

$$\sum_{r=0}^{d-1} [\widetilde{\mu}_k[r]]^2 = 4\sigma_k^{-2} \left(|\mathsf{X}[k]|\, |\mathsf{N}[k]|\right)^2 \sum_{r=0}^{d-1} \cos^2\left(\frac{2\pi k}{d} r + \phi_{\mathsf{N}}[k] - \phi_{\mathsf{X}}[k]\right). \tag{D.38}$$

As $d \to \infty$,

$$\frac{1}{d} \sum_{r=0}^{d-1} \cos^2\left(\frac{2\pi k}{d} r + \phi_{\mathsf{N}}[k] - \phi_{\mathsf{X}}[k]\right) \xrightarrow{\text{a.s.}} \frac{1}{2}. \tag{D.39}$$

Combining (D.38) and (D.39), we obtain

$$\frac{\sigma_k^2}{2\,|\mathsf{X}[k]|^2} \sum_{r=0}^{d-1} [\widetilde{\mu}_k[r]]^2 \xrightarrow{\text{a.s.}} |\mathsf{N}[k]|^2, \tag{D.40}$$

for $d \to \infty$. Thus, combining (D.36)–(D.40), we conclude that

$$\frac{\sigma_k^2}{2\,|\mathsf{X}[k]|^2\, a_d} \frac{\sum_{r=0}^{d-1} \widetilde{\mu}_k[r] e^{\widetilde{\mu}_k[r] a_d}}{\sum_{r=0}^{d-1} e^{\widetilde{\mu}_k[r] a_d}} \xrightarrow{\text{a.s.}} |\mathsf{N}[k]|^2. \tag{D.41}$$

This proves (D.32). Finally, combining (D.29), (D.30), and (D.41), we arrive at

$$\frac{\sigma_k^2}{2\,|\mathsf{X}[k]|^2\, a_d} \mathbb{E}\left[\widetilde{\mu}_k[\hat{\mathsf{R}}] \middle| \mathsf{N}[k]\right] \xrightarrow{\text{a.s.}} |\mathsf{N}[k]|^2. \tag{D.42}$$

Let the term on the left-hand side of (D.42) be denoted by G_d and the term on the right-hand side by G, so that $G_d \xrightarrow{\text{a.s.}} G$. By definition, observe that $|G_d| \leq |\mathsf{N}[k]|^2$, and it is evident that $\mathbb{E}[|\mathsf{N}[k]|^2] = \sigma^2 < \infty$. Thus, by the dominated convergence theorem, we conclude that $G_d \xrightarrow{\mathcal{L}^1} G$, and specifically,

$$\mathbb{E}\left[\frac{\sigma_k^2}{2a_d} \frac{\mathbb{E}[\mu(\hat{\mathsf{R}}) | \mathsf{N}[k]]}{|\mathsf{X}[k]|^2}\right] \to \mathbb{E}\left[|\mathsf{N}[k]|^2\right]. \tag{D.43}$$

By combining (D.43) with (D.29), we get,

$$\frac{\sigma_k}{a_d \cdot |\mathsf{X}[k]|} \mathbb{E}\left[f_1(\hat{\mathsf{R}}) \middle| \mathsf{N}[k]\right] \xrightarrow{\mathcal{L}^1} |\mathsf{N}[k]|^2, \tag{D.44}$$

or equivalently,

$$\mathbb{E}\left[\frac{\sigma_k}{a_d \cdot |\mathsf{X}[k]|}\mathbb{E}\left[f_1(\hat{\mathsf{R}})\middle|\,\mathsf{N}[k]\right]\right] \to \mathbb{E}\left[|\mathsf{N}[k]|^2\right], \tag{D.45}$$

which proves (D.26).

By the law of total expectation, we then obtain

$$\frac{\sigma_k}{a_d \cdot |\mathsf{X}[k]|}\mathbb{E}\left[|\mathsf{N}[k]|\cos\left(\frac{2\pi k}{d}\hat{\mathsf{R}} + \phi_{\mathsf{N}}[k] - \phi_{\mathsf{X}}[k]\right)\right] \to \mathbb{E}\left[|\mathsf{N}[k]|^2\right] = \sigma^2, \tag{D.46}$$

as $d \to \infty$. By Lemma D.1, we know that $\sigma_k^2 \to \sigma^2$ as $d \to \infty$, so from (D.46), we conclude that

$$\frac{1}{a_d \cdot |\mathsf{X}[k]|}\mathbb{E}\left[|\mathsf{N}[k]|\cos\left(\frac{2\pi k}{d}\hat{\mathsf{R}} + \phi_{\mathsf{N}}[k] - \phi_{\mathsf{X}}[k]\right)\right] \to \sigma, \tag{D.47}$$

which proves (D.22).

Proof of (D.27) and (D.23). The numerator in (B.64) converges to,

$$\mathbb{E}[|\mathsf{N}[k]|\sin(\phi_e[k])]^2 = \frac{1}{2}\left[\mathbb{E}\,|\mathsf{N}[k]|^2 - \mathbb{E}[|\mathsf{N}[k]|^2\cos(2\phi_e[k])]\right] \to \frac{1}{2}\mathbb{E}\,|\mathsf{N}[k]|^2, \tag{D.48}$$

as $d \to \infty$, where the last transition is because $\mathbb{E}[\cos(2\phi_e[k])|\mathsf{N}[k]] \xrightarrow{\text{a.s.}} 0$, as $d \to \infty$. Thus,

$$\frac{1}{\sigma^2}\mathbb{E}[|\mathsf{N}[k]|\sin(\phi_e[k])]^2 \to \frac{1}{2\sigma^2}\mathbb{E}\,|\mathsf{N}[k]|^2 = \frac{1}{2}, \tag{D.49}$$

which concludes the proof. □

D.3 Proof of Theorem 4.3

To begin, let us summarize the notation and results from the previous sections as the foundation for the proof. Recall the definition of $\phi_e[k]$ in (A.7), as well as the definitions of $f_1(r)$ and $f_2(r)$ in (D.20)-(D.21), and let $a_d \triangleq \sqrt{2\log d}$. According to Theorem 4.1, the convergence of the Fourier phases is given by:

$$\lim_{M\to\infty}\frac{\mathbb{E}|\phi_{\hat{\mathsf{X}}}[k] - \phi_{\mathsf{X}}[k]|^2}{1/M} = C_k. \tag{D.50}$$

where C_k is given by (B.64):

$$C_k = \frac{\mathbb{E}\left[[|\mathsf{N}[k]|\sin(\phi_e[k])]^2\right]}{\left(\mathbb{E}[|\mathsf{N}[k]|\cos(\phi_e[k])]\right)^2}. \tag{D.51}$$

The constant C_k can be rewritten as:

$$a_d^2 \cdot C_k = a_d^2 \cdot \frac{\mathbb{E}\left[[|\mathsf{N}[k]|\sin(\phi_e[k])]^2\right]}{\left(\mathbb{E}[|\mathsf{N}[k]|\cos(\phi_e[k])]\right)^2} = a_d^2 \cdot \frac{\mathbb{E}[f_2(\hat{\mathsf{R}})]}{\left(\mathbb{E}[f_1(\hat{\mathsf{R}})]\right)^2}, \tag{D.52}$$

where $f_1(r)$ and $f_2(r)$ are defined in (D.20)-(D.21).

The signal x, satisfying the conditions of Theorem 4.3, also satisfies the assumptions of Proposition D.3. By Proposition D.3, we have:

$$\lim_{d\to\infty} \frac{1}{a_d \cdot |\mathsf{X}[k]|}\mathbb{E}[f_1(\hat{\mathsf{R}})] = \sigma, \tag{D.53}$$

and,

$$\lim_{d\to\infty} \mathbb{E}[f_2(\hat{\mathsf{R}})] = \frac{\sigma^2}{2}. \tag{D.54}$$

Now, we are ready to prove the results of the Theorem.

Convergence of the Fourier phases. From (D.52)–(D.54), we obtain:

$$\lim_{d\to\infty} a_d^2 \cdot |\mathsf{X}[k]|^2 \cdot C_k = \frac{1}{2}. \tag{D.55}$$

Combining (D.50), (D.55), results,

$$\lim_{d\to\infty}\lim_{M\to\infty} \frac{\mathbb{E}|\phi_{\hat{\mathsf{X}}}[k] - \phi_{\mathsf{X}}[k]|^2}{1/M} \frac{1}{1/(4\log(d)\,|\mathsf{X}[k]|^2)} = \lim_{d\to\infty} \frac{C_k}{1/(4\log(d)\,|\mathsf{X}[k]|^2)} \tag{D.56}$$

$$= \lim_{d\to\infty} \frac{a_d^2 \cdot |\mathsf{X}[k]|^2 \cdot C_k}{1/2} = 1, \tag{D.57}$$

where (D.56) follows from (D.50), and (D.57) follows from (D.55), proving (4.8).

Convergence of the Fourier magnitudes. Finally, we prove (4.9). By Theorem 4.1 and (4.3), we have:

$$|\hat{\mathsf{X}}[k]| \xrightarrow{\text{a.s.}} \mathbb{E}\left[|\mathsf{N}[k]|\cos\left(\frac{2\pi k}{d}\hat{\mathsf{R}} + \phi_{\mathsf{N}}[k] - \phi_{\mathsf{X}}[k]\right)\right] = \mathbb{E}[f_1(\hat{\mathsf{R}})]. \tag{D.58}$$

Combining (D.53), (D.58) yields,

$$\frac{1}{a_d\sigma}\frac{|\hat{\mathsf{X}}[k]|}{|\mathsf{X}[k]|} \xrightarrow{\text{a.s.}} \frac{1}{a_d\sigma}\frac{\mathbb{E}[f_1(\hat{\mathsf{R}})]}{|\mathsf{X}[k]|} \to 1, \tag{D.59}$$

as $M, d \to \infty$, where the second passage follows from (D.22). As $a_d = \sqrt{2\log(d)}$, this completes the proof of the Theorem.

E Proof of Proposition 5.1

Before proving Proposition 5.1, we first establish the following auxiliary lemma.

Lemma E.1. *Let* $\mathsf{A} = (A_0, A_1, \ldots, A_{d-1})$ *be a* d*-dimensional random vector with* $\mathbb{E}[\mathsf{A}] = 0$*. Then,*

$$\mathbb{E}\left[\max\{A_0, A_1, \ldots, A_{d-1}\}\right] \geq \max_{0\leq r_1, r_2 \leq d-1} \frac{1}{2}\mathbb{E}\left[|A_{r_1} - A_{r_2}|\right]. \tag{E.1}$$

Proof of Lemma E.1. For any two real numbers x and y, we have:

$$\max(x, y) = \frac{1}{2}(x + y + |x - y|). \tag{E.2}$$

Applying this to any pair A_{r_1}, A_{r_2} yields:

$$\mathbb{E}[\max\{A_{r_1}, A_{r_2}\}] = \frac{1}{2}\mathbb{E}[A_{r_1} + A_{r_2} + |A_{r_1} - A_{r_2}|] \tag{E.3}$$

$$= \frac{1}{2}\mathbb{E}[|A_{r_1} - A_{r_2}|], \tag{E.4}$$

where we used the assumption that $\mathbb{E}[A_r] = 0$ for all r.

By the convexity of the max function, it holds that:

$$\mathbb{E}[\max\{A_0, A_1, \ldots, A_{d-1}\}] \geq \mathbb{E}[\max\{A_{r_1}, A_{r_2}\}], \tag{E.5}$$

for every $r_1, r_2 \in \{0, 1, \ldots, d-1\}$. Combining (E.4) and (E.5), we conclude:

$$\mathbb{E}[\max\{A_0, A_1, \ldots, A_{d-1}\}] \geq \max_{0 \leq r_1, r_2 \leq d-1} \frac{1}{2}\mathbb{E}[|A_{r_1} - A_{r_2}|], \tag{E.6}$$

completing the proof. □

Let $n_0, n_1, \ldots, n_{M-1}$ be an i.i.d. sequence of zero-mean random vectors with covariance $\mathbb{E}[n_i n_i^\top] = \Sigma$, which by assumption Σ is positive-definite. Recall the definition of the EfN estimator in (2.4):

$$\hat{x} \triangleq \frac{1}{M} \sum_{i=0}^{M-1} \mathcal{T}_{-\hat{\mathsf{R}}_i} n_i, \tag{E.7}$$

where the estimated shift $\hat{\mathsf{R}}_i$ is given by:

$$\hat{\mathsf{R}}_i \triangleq \arg\max_{0 \leq \ell \leq d-1} \langle n_i, \mathcal{T}_\ell x \rangle. \tag{E.8}$$

Using linearity of the inner product:

$$\langle \hat{x}, x \rangle = \left\langle \frac{1}{M} \sum_{i=0}^{M-1} \mathcal{T}_{-\hat{\mathsf{R}}_i} n_i, x \right\rangle = \frac{1}{M} \sum_{i=0}^{M-1} \langle \mathcal{T}_{-\hat{\mathsf{R}}_i} n_i, x \rangle. \tag{E.9}$$

By SLLN, as $M \to \infty$, we have almost surely:

$$\frac{1}{M} \sum_{i=0}^{M-1} \langle \mathcal{T}_{-\hat{\mathsf{R}}_i} n_i, x \rangle \xrightarrow{\text{a.s.}} \mathbb{E}[\langle \mathcal{T}_{-\hat{\mathsf{R}}_1} n_1, x \rangle]. \tag{E.10}$$

Define for $r \in \{0, 1, \ldots, d-1\}$ the random variables:

$$A_r \triangleq \langle n_1, \mathcal{T}_r x \rangle. \tag{E.11}$$

Then, the right-hand side of (E.10) becomes:

$$\mathbb{E}[\langle \mathcal{T}_{-\hat{\mathsf{R}}_1} n_1, x\rangle] = \mathbb{E}[\max\{A_0, A_1, \ldots, A_{d-1}\}]. \tag{E.12}$$

Applying Lemma E.1, we get:

$$\mathbb{E}[\max\{A_0, A_1, \ldots, A_{d-1}\}] \geq \max_{0\leq r_1, r_2 \leq d-1} \frac{1}{2}\mathbb{E}[|A_{r_1} - A_{r_2}|] \tag{E.13}$$

$$= \max_{0\leq r_1, r_2 \leq d-1} \frac{1}{2}\mathbb{E}[|\langle n_1, \mathcal{T}_{r_1} x - \mathcal{T}_{r_2} x\rangle|]. \tag{E.14}$$

To complete the proof, we show that the lower bound in (E.14) is strictly positive. Since $x \in \mathbb{R}^d$ is nonzero with non-vanishing Fourier components $\mathsf{X}[k] \neq 0$, for every $1 \leq k \leq d-1$, and $\mathcal{T}_r$ is a cyclic shift operator, the set $\{\mathcal{T}_r x : 0 \leq r < d\}$ contains at least $d-1$ distinct vectors. Thus, there exist $r_1, r_2 \in \{0, \ldots, d-1\}$ such that

$$v \triangleq \mathcal{T}_{r_1} x - \mathcal{T}_{r_2} x \neq 0. \tag{E.15}$$

Then the inner product $\langle n_1, v\rangle$ is a real-valued random variable with

$$\mathrm{Var}(\langle n_1, v\rangle) = \mathbb{E}[\langle n_1, v\rangle^2] = v^\top \Sigma v > 0, \tag{E.16}$$

because $v \neq 0$ and Σ is positive definite. Hence, $\langle n_1, v\rangle$ is not almost surely zero, and

$$\mathbb{E}[|\langle n_1, v\rangle|] > 0. \tag{E.17}$$

This implies that

$$\max_{0\leq r_1, r_2 < d} \frac{1}{2}\mathbb{E}[|\langle n_1, \mathcal{T}_{r_1} x - \mathcal{T}_{r_2} x\rangle|] > 0, \tag{E.18}$$

and consequently,

$$\lim_{M\to\infty} \langle \hat{x}, x\rangle = \mathbb{E}[\langle \mathcal{T}_{-\hat{\mathsf{R}}_1} n_1, x\rangle] > 0, \tag{E.19}$$

almost surely. This completes the proof.

F Proof of Theorem 5.2: High-dimensional i.i.d. noise

In this section, we prove Theorem 5.2. The proof relies on the functional central limit theorem for the discrete Fourier transform [37, 14, 13], which we review in Appendix F.1. In Appendices F.2 and F.3, we apply this result to analyze the real and imaginary parts of the EfN estimator under a general i.i.d. noise model, and compare the outcome to the white Gaussian case. Finally, the proof of Theorem 5.2 is deduced in Appendix F.4.

F.1 The functional CLT for DFT

We begin by presenting a functional central limit theorem (CLT) for the DFT, which establishes that the DFT of an i.i.d. real-valued sequence converges in distribution as the dimension $d \to \infty$. This result has been studied in the literature; see, for example, [37, 14, 13]. To formalize this, we state the following functional CLT for DFTs of i.i.d. sequences.

Theorem F.1 (Functional CLT for the DFT)**.** *Let $\{z_n\}_{n\in\mathbb{N}}$ be a sequence of i.i.d. real-valued random variables with zero mean $\mathbb{E}[z_0] = 0$ and finite variance $\mathbb{E}[z_0^2] = \sigma^2 < \infty$. For each integer $d \geq 1$, define the DFT of the finite segment $\{z_0, \dots, z_{d-1}\}$ as*

$$\mathsf{Z}^{(d)}[k] \triangleq \frac{1}{\sqrt{d}} \sum_{\ell=0}^{d-1} z_\ell e^{-2\pi jk\ell/d}, \quad 0 \leq k < d. \tag{F.1}$$

Extend $\mathsf{Z}^{(d)}$ to an infinite sequence by zero-padding outside the index set $\{0, \dots, d-1\}$:

$$\mathsf{Z}^{(d)} = \big(\mathsf{Z}^{(d)}[0], \dots, \mathsf{Z}^{(d)}[d-1], 0, 0, \dots\big) \in \mathbb{C}^{\mathbb{N}}. \tag{F.2}$$

Then, for any fixed finite index set $\{k_1, k_2, \dots, k_m\} \subset \mathbb{N}$, the finite-dimensional vectors $\big(\mathsf{Z}^{(d)}[k_1], \dots, \mathsf{Z}^{(d)}[k_m]\big)$ converge in distribution, as $d \to \infty$,

$$\big(\mathsf{Z}^{(d)}[k_1], \dots, \mathsf{Z}^{(d)}[k_m]\big) \xrightarrow[d\to\infty]{\mathcal{D}} \big(\mathsf{W}_{k_1}, \dots, \mathsf{W}_{k_m}\big), \tag{F.3}$$

where $\mathsf{W} = (\mathsf{W}_k)_{k\in\mathbb{N}}$ is a sequence of i.i.d. circularly symmetric complex Gaussian random variables with $\mathsf{W}_k \sim \mathcal{CN}(0, \sigma^2)$.

This result can be obtained from the multivariate Lindeberg–Feller CLT. This convergence holds jointly over any finite collection of indices, meaning that for every finite subset $I \subset \mathbb{N}$, the finite-dimensional vector $(\mathsf{Z}_k^{(d)})_{k\in I}$ converges in distribution to $(\mathsf{W}_k)_{k\in I}$, where the W_k are i.i.d. circularly symmetric complex Gaussian random variables. The collection of these finite-dimensional distributions is consistent and satisfies the compatibility conditions of Kolmogorov's extension theorem, thereby uniquely determining a probability law on the infinite product space $\mathbb{C}^{\mathbb{N}}$. In this sense, the convergence of $\mathsf{Z}^{(d)}$ to W is fully characterized by convergence of finite-dimensional distributions.

F.2 Notations

General i.i.d noise. Let $z_0, z_1, \dots$ be a sequence of i.i.d random variables with $\mathbb{E}[z_0] = 0$, and $\mathbb{E}[z_0^2] = \sigma^2 < \infty$, and $\mathbb{E}\left[z_0^4\right] < \infty$. Define the (zero-padded) DFT transform of the finite segment $z^{(d)} = (z_0, z_1, \dots, z_{d-1})$ as

$$\mathsf{Z}^{(d)}[k] \triangleq \begin{cases} \dfrac{1}{\sqrt{d}} \displaystyle\sum_{\ell=0}^{d-1} z_\ell e^{-2\pi jk\ell/d}, & 0 \leq k < d, \\ 0, & \text{otherwise.} \end{cases} \tag{F.4}$$

Let $x = (x_0, \ldots, x_{d-1})$ be the deterministic template signal, and denote its DFT by $\mathsf{X}[k]$. Define the maximal correlation shift between x and $z^{(d)}$ in the Fourier domain as

$$\hat{\mathsf{R}}_{\mathsf{Z}}^{(d)} \triangleq \arg\max_{0 \le r < d} \sum_{k=0}^{d-1} |\mathsf{X}[k]|\,|\mathsf{Z}[k]| \cos\left(\frac{2\pi k r}{d} + \phi_{\mathsf{Z}}[k] - \phi_{\mathsf{X}}[k]\right), \tag{F.5}$$

where $\phi_{\mathsf{Z}}[k]$ and $\phi_{\mathsf{X}}[k]$ are the phases of $\mathsf{Z}^{(d)}[k]$ and $\mathsf{X}[k]$, respectively. Define the phase difference after alignment as

$$\phi_{e,\mathsf{Z}}^{(d)}[k] \triangleq \frac{2\pi k \hat{\mathsf{R}}_{\mathsf{Z}}^{(d)}}{d} + \phi_{\mathsf{Z}}[k] - \phi_{\mathsf{X}}[k]. \tag{F.6}$$

Gaussian i.i.d noise. Let $n_0, n_1, \ldots$ be an i.i.d sequence of *Gaussian random variables* with $n_\ell \sim \mathcal{N}(0, \sigma^2)$. Define the DFT of the segment $n^{(d)} = (n_0, \ldots, n_{d-1})$ as

$$\mathsf{N}^{(d)}[k] \triangleq \begin{cases} \dfrac{1}{\sqrt{d}} \displaystyle\sum_{\ell=0}^{d-1} n_\ell e^{-2\pi j k \ell / d}, & 0 \le k < d, \\ 0, & \text{otherwise.} \end{cases} \tag{F.7}$$

Define the corresponding maximal correlation shift:

$$\hat{\mathsf{R}}_{\mathsf{N}}^{(d)} \triangleq \arg\max_{0 \le r < d} \sum_{k=0}^{d-1} |\mathsf{X}[k]|\,|\mathsf{N}[k]| \cos\left(\frac{2\pi k r}{d} + \phi_{\mathsf{N}}[k] - \phi_{\mathsf{X}}[k]\right), \tag{F.8}$$

and define the aligned phase difference as

$$\phi_{e,\mathsf{N}}^{(d)}[k] \triangleq \frac{2\pi k \hat{\mathsf{R}}_{\mathsf{N}}^{(d)}}{d} + \phi_{\mathsf{N}}[k] - \phi_{\mathsf{X}}[k]. \tag{F.9}$$

F.3 Convergence of the real and imaginary parts of the EfN estimator

We now present an auxiliary result that relates the real and imaginary parts of the EfN estimator under both Gaussian i.i.d. and general i.i.d. noise models. This result is a consequence of the functional central limit theorem for the DFT (Theorem F.1).

Proposition F.2. *Let $z_0, z_1, \ldots$ be a sequence of i.i.d. real-valued random variables with zero mean, finite variance, $\mathbb{E}[z_0^2] = \sigma^2 < \infty$, and finite fourth moment. Let $n_0, n_1, \ldots$ be an i.i.d. sequence of Gaussian random variables with $n_\ell \sim \mathcal{N}(0, \sigma^2)$. Let $\mathsf{Z}^{(d)}$ and $\mathsf{N}^{(d)}$ denote the DFTs of the sequences $\{z_i\}_{i=0}^{d-1}$ and $\{n_i\}_{i=0}^{d-1}$, respectively. Let $\phi_{e,\mathsf{Z}}^{(d)}[k]$ and $\phi_{e,\mathsf{N}}^{(d)}[k]$ denote the aligned phase differences defined in (F.6) and (F.9), respectively. Then, for each fixed frequency index $k \in \mathbb{N}$,*

$$\lim_{d\to\infty} \left(\mathbb{E}\left[\left|\mathsf{Z}^{(d)}[k]\right| \sin\left(\phi_{e,\mathsf{Z}}^{(d)}[k]\right)\right] - \mathbb{E}\left[\left|\mathsf{N}^{(d)}[k]\right| \sin\left(\phi_{e,\mathsf{N}}^{(d)}[k]\right)\right]\right) = 0, \tag{F.10}$$

$$\lim_{d\to\infty} \left(\mathbb{E}\left[\left|\mathsf{Z}^{(d)}[k]\right| \cos\left(\phi_{e,\mathsf{Z}}^{(d)}[k]\right)\right] - \mathbb{E}\left[\left|\mathsf{N}^{(d)}[k]\right| \cos\left(\phi_{e,\mathsf{N}}^{(d)}[k]\right)\right]\right) = 0, \tag{F.11}$$

and

$$\lim_{d\to\infty}\left(\mathsf{Var}\left[\left|\mathsf{Z}^{(d)}[k]\right|\sin\left(\phi_{e,\mathsf{Z}}^{(d)}[k]\right)\right]-\mathsf{Var}\left[\left|\mathsf{N}^{(d)}[k]\right|\sin\left(\phi_{e,\mathsf{N}}^{(d)}[k]\right)\right]\right)=0, \tag{F.12}$$

$$\lim_{d\to\infty}\left(\mathsf{Var}\left[\left|\mathsf{Z}^{(d)}[k]\right|\cos\left(\phi_{e,\mathsf{Z}}^{(d)}[k]\right)\right]-\mathsf{Var}\left[\left|\mathsf{N}^{(d)}[k]\right|\cos\left(\phi_{e,\mathsf{N}}^{(d)}[k]\right)\right]\right)=0. \tag{F.13}$$

Proof of Proposition F.2. We begin by applying Theorem F.1 to the sequence $\{z_i\}$, which satisfies its assumptions. This in turn implies that,

$$\mathsf{Z}^{(d)} \xrightarrow{\mathcal{D}} \mathsf{W}, \tag{F.14}$$

where $\mathsf{W}_k \sim \mathcal{CN}(0,\sigma^2)$ are i.i.d. complex Gaussian variables. Next, consider the Gaussian noise sequence $n_i \sim \mathcal{N}(0,\sigma^2)$. Its discrete Fourier transform satisfies

$$\mathsf{N}^{(d)} \overset{\mathcal{D}}{=} (\mathsf{W}[0],\ldots,\mathsf{W}[d-1]), \tag{F.15}$$

for every d, since the DFT of an i.i.d. Gaussian sequence remains i.i.d. in distribution with the same complex Gaussian law. Since $\mathbb{E}[z_i^2]=\sigma^2<\infty$, we have

$$\mathbb{E}[|\mathsf{Z}^{(d)}[k]|^2]<\infty, \tag{F.16}$$

for each d and k. It follows that the sequences of random variables,

$$\left\{|\mathsf{Z}^{(d)}[k]|\sin(\phi_{e,\mathsf{Z}}^{(d)}[k])\right\}_{d\in\mathbb{N}^+},\quad \left\{|\mathsf{Z}^{(d)}[k]|\cos(\phi_{e,\mathsf{Z}}^{(d)}[k])\right\}_{d\in\mathbb{N}^+}, \tag{F.17}$$

are uniformly integrable. By Vitali's convergence theorem, we may pass the limit inside the expectation,

$$\lim_{d\to\infty}\left(\mathbb{E}\left[|\mathsf{Z}^{(d)}[k]|\sin(\phi_{e,\mathsf{Z}}^{(d)}[k])\right]-\mathbb{E}\left[|\mathsf{W}[k]|\sin(\phi_{e,\mathsf{W}}[k])\right]\right)=0. \tag{F.18}$$

Furthermore, since $\mathsf{N}^{(d)} \overset{\mathcal{D}}{=} \mathsf{W}$ for all d, and the aligned phase differences $\phi_{e,\mathsf{N}}^{(d)}[k]$ and $\phi_{e,\mathsf{W}}[k]$ are identically distributed, we conclude that

$$\mathbb{E}\left[|\mathsf{N}^{(d)}[k]|\sin(\phi_{e,\mathsf{N}}^{(d)}[k])\right]=\mathbb{E}\left[|\mathsf{W}[k]|\sin(\phi_{e,\mathsf{W}}[k])\right],\quad \forall d\in\mathbb{N}^+. \tag{F.19}$$

Combining (F.18) and (F.19) yields

$$\lim_{d\to\infty}\left(\mathbb{E}\left[|\mathsf{Z}^{(d)}[k]|\sin(\phi_{e,\mathsf{Z}}^{(d)}[k])\right]-\mathbb{E}\left[|\mathsf{N}^{(d)}[k]|\sin(\phi_{e,\mathsf{N}}^{(d)}[k])\right]\right)=0, \tag{F.20}$$

which establishes (F.10). An identical argument with sine replaced by cosine proves (F.11).

For the variance convergence, by assumption $\mathbb{E}[z_i^4]<\infty$, and so,

$$\mathbb{E}[|\mathsf{Z}^{(d)}[k]|^4]<\infty, \tag{F.21}$$

for all d. Therefore, the sequences

$$\left\{|\mathsf{Z}^{(d)}[k]|^2 \sin^2(\phi_{e,\mathsf{Z}}^{(d)}[k])\right\}_{d\in\mathbb{N}^+}, \quad \left\{|\mathsf{Z}^{(d)}[k]|^2 \cos^2(\phi_{e,\mathsf{Z}}^{(d)}[k])\right\}_{d\in\mathbb{N}^+}, \tag{F.22}$$

are uniformly integrable. Applying Vitali's theorem once again, we obtain,

$$\lim_{d\to\infty}\left(\mathsf{Var}\left[|\mathsf{Z}^{(d)}[k]|\sin(\phi_{e,\mathsf{Z}}^{(d)}[k])\right] - \mathsf{Var}\left[|\mathsf{W}[k]|\sin(\phi_{e,\mathsf{W}}[k])\right]\right) = 0. \tag{F.23}$$

As before, since $\mathsf{N}^{(d)} \overset{\mathcal{D}}{=} \mathsf{W}$ and their aligned phase differences are identically distributed, we conclude that,

$$\mathsf{Var}\left[|\mathsf{N}^{(d)}[k]|\sin(\phi_{e,\mathsf{N}}^{(d)}[k])\right] = \mathsf{Var}\left[|\mathsf{W}[k]|\sin(\phi_{e,\mathsf{W}}[k])\right], \quad \forall d\in\mathbb{N}^+. \tag{F.24}$$

Combining (F.23) and (F.24) yields,

$$\lim_{d\to\infty}\left(\mathsf{Var}\left[|\mathsf{Z}^{(d)}[k]|\sin(\phi_{e,\mathsf{Z}}^{(d)}[k])\right] - \mathsf{Var}\left[|\mathsf{N}^{(d)}[k]|\sin(\phi_{e,\mathsf{N}}^{(d)}[k])\right]\right) = 0, \tag{F.25}$$

establishing (F.12). Repeating the same steps with cosine in place of sine proves (F.13). □

F.4 Proof of Theorem 5.2

We are now ready to prove Theorem 5.2. As before, let $\{z_i\}_{i=0}^{M-1}$ be i.i.d. observations, where each $z_i \in \mathbb{R}^d$ has i.i.d. entries with zero mean, finite variance, and bounded fourth moment $\mathbb{E}[(z_i[\ell])^4] < \infty$, for all $\ell \in \{0, 1, \ldots, d-1\}$.

Similarly to (A.13), we analyze the EfN estimator under the noise statistics of $\{z_i\}_{i=0}^{M-1}$. Applying the SLLN, as $M \to \infty$, we obtain:

$$\hat{\mathsf{X}}[k]e^{-j\phi_{\mathsf{X}}[k]} = \frac{1}{M}\sum_{i=0}^{M-1} |\mathsf{Z}_i[k]|\, e^{j\phi_{e,\mathsf{Z}_i}[k]} \tag{F.26}$$

$$\xrightarrow{\text{a.s.}} \mathbb{E}\left[|\mathsf{Z}_1[k]|\cos\left(\phi_{e,\mathsf{Z}_1}[k]\right)\right] + j\mathbb{E}\left[|\mathsf{Z}_1[k]|\sin\left(\phi_{e,\mathsf{Z}_1}[k]\right)\right], \tag{F.27}$$

where the phase difference term is given by

$$\phi_{e,\mathsf{Z}_i}^{(d)}[k] \triangleq \frac{2\pi k \hat{\mathsf{R}}_{\mathsf{Z}_i}^{(d)}}{d} + \phi_{\mathsf{Z}_i}[k] - \phi_{\mathsf{X}}[k]. \tag{F.28}$$

and the corresponding maximal correlation shift $\hat{\mathsf{R}}_{\mathsf{Z}_i}^{(d)}$ is defined by

$$\hat{\mathsf{R}}_{\mathsf{Z}_i}^{(d)} \triangleq \arg\max_{0\le r<d} \sum_{k=0}^{d-1} |\mathsf{X}[k]|\,|\mathsf{Z}_i[k]| \cos\left(\frac{2\pi k r}{d} + \phi_{\mathsf{Z}_i}[k] - \phi_{\mathsf{X}}[k]\right). \tag{F.29}$$

Next, we invoke Proposition F.2, whose assumptions are satisfied in this setting. Let $\mathsf{N}_1^{(d)}$ denotes a noise vector with i.i.d. Gaussian entries that match the first and second moments

of the entries of $\mathsf{Z}_1^{(d)}$ (as defined in Proposition F.2). We note that the results of Theorems 4.1 and 4.3 apply to the case of i.i.d. Gaussian entries $\mathsf{N}_1^{(d)}$. Then, by Proposition F.2, for each fixed frequency index $k \in \mathbb{N}$, the following convergence results hold,

$$\lim_{d\to\infty} \left(\mathbb{E}\left[\left| \mathsf{Z}_1^{(d)}[k] \right| \sin\left(\phi_{e,\mathsf{Z}_1}^{(d)}[k] \right) \right] - \mathbb{E}\left[\left| \mathsf{N}_1^{(d)}[k] \right| \sin\left(\phi_{e,\mathsf{N}_1}^{(d)}[k] \right) \right] \right) = 0, \tag{F.30}$$

$$\lim_{d\to\infty} \left(\mathbb{E}\left[\left| \mathsf{Z}_1^{(d)}[k] \right| \cos\left(\phi_{e,\mathsf{Z}_1}^{(d)}[k] \right) \right] - \mathbb{E}\left[\left| \mathsf{N}_1^{(d)}[k] \right| \cos\left(\phi_{e,\mathsf{N}_1}^{(d)}[k] \right) \right] \right) = 0. \tag{F.31}$$

Moreover, the variances of the corresponding expressions also converge,

$$\lim_{d\to\infty} \left(\mathsf{Var}\left[\left| \mathsf{Z}_1^{(d)}[k] \right| \sin\left(\phi_{e,\mathsf{Z}_1}^{(d)}[k] \right) \right] - \mathsf{Var}\left[\left| \mathsf{N}_1^{(d)}[k] \right| \sin\left(\phi_{e,\mathsf{N}_1}^{(d)}[k] \right) \right] \right) = 0, \tag{F.32}$$

$$\lim_{d\to\infty} \left(\mathsf{Var}\left[\left| \mathsf{Z}_1^{(d)}[k] \right| \cos\left(\phi_{e,\mathsf{Z}_1}^{(d)}[k] \right) \right] - \mathsf{Var}\left[\left| \mathsf{N}^{(d)}[k] \right| \cos\left(\phi_{e,\mathsf{N}_1}^{(d)}[k] \right) \right] \right) = 0. \tag{F.33}$$

By Theorems 4.1 and 4.3, the convergence behavior of the estimator is governed by the variances in (F.32) and (F.33). Therefore, the asymptotic behavior of the estimator for general i.i.d. noise $\{z_i\}$ matches that of the Gaussian i.i.d. case $\{n_i\}$. In particular for (5.3), we have,

$$\phi_{\hat{\mathsf{X}}}[k] - \phi_{\mathsf{X}}[k] = \arctan\left(\frac{\sum_{i=0}^{M-1} \left| \mathsf{Z}_i^{(d)}[k] \right| \sin\left(\phi_{e,\mathsf{Z}_i}[k] \right)}{\sum_{i=0}^{M-1} \left| \mathsf{Z}_i^{(d)}[k] \right| \cos\left(\phi_{e,\mathsf{Z}_i}[k] \right)} \right) \tag{F.34}$$

$$\xrightarrow{\text{a.s.}} \arctan\left(\frac{\mathbb{E}\left[\left| \mathsf{Z}_1^{(d)}[k] \right| \sin\left(\phi_{e,\mathsf{Z}_1}[k] \right) \right]}{\mathbb{E}\left[\left| \mathsf{Z}_1^{(d)}[k] \right| \cos\left(\phi_{e,\mathsf{Z}_1}[k] \right) \right]} \right), \tag{F.35}$$

where (F.35) follows from the SLLN as $M \to \infty$. Applying (F.30) and (F.31) into (F.35), yields,

$$\lim_{d\to\infty} \lim_{M\to\infty} \phi_{\hat{\mathsf{X}}}[k] - \phi_{\mathsf{X}}[k] = \lim_{d\to\infty} \arctan\left(\frac{\mathbb{E}\left[\left| \mathsf{N}_1^{(d)}[k] \right| \sin\left(\phi_{e,\mathsf{N}_1}[k] \right) \right]}{\mathbb{E}\left[\left| \mathsf{N}_1^{(d)}[k] \right| \cos\left(\phi_{e,\mathsf{N}_1}[k] \right) \right]} \right). \tag{F.36}$$

By (B.62), the r.h.s. of (F.36) vanishes for every d, and therefore,

$$\lim_{d\to\infty} \lim_{M\to\infty} \phi_{\hat{\mathsf{X}}}[k] - \phi_{\mathsf{X}}[k] = 0, \tag{F.37}$$

almost surely, which proves (5.3). Similarly, for (5.4), we have,

$$\lim_{M\to\infty} \frac{\mathbb{E}|\phi_{\hat{\mathsf{X}}}[k] - \phi_{\mathsf{X}}[k]|^2}{1/M} = \frac{\mathbb{E}\left(\left[|\mathsf{Z}_1[k]| \sin(\phi_{e,\mathsf{Z}_1}[k]) \right]^2 \right)}{\left(\mathbb{E}[|\mathsf{Z}_1[k]| \cos(\phi_{e,\mathsf{Z}_1}[k])] \right)^2}. \tag{F.38}$$

which is similar to (B.24). Applying (F.31) and (F.32) into (F.38) yields,

$$\lim_{d\to\infty}\lim_{M\to\infty}\frac{\mathbb{E}|\phi_{\hat{\mathsf{X}}}[k]-\phi_{\mathsf{X}}[k]|^2}{1/M}=\lim_{d\to\infty}\frac{\mathbb{E}\left([|\mathsf{N}_1[k]|\sin(\phi_{e,\mathsf{N}_1}[k])]^2\right)}{\left(\mathbb{E}[|\mathsf{N}_1[k]|\cos(\phi_{e,\mathsf{N}_1}[k])]\right)^2}=C_k<\infty. \tag{F.39}$$

Finally, for (5.5), under the assumption that x satisfies Assumption 4.2, then by (4.8), the r.h.s. of (F.39) converges to,

$$\lim_{d\to\infty} a_d^2\cdot|\mathsf{X}[k]|^2\cdot C_k=\frac{1}{2}, \tag{F.40}$$

where $a_d=\sqrt{2\log(d)}$. Thus, substituting (F.40) into the r.h.s. of (F.39) yields,

$$\lim_{d\to\infty}\lim_{M\to\infty}\frac{\mathbb{E}\left[|\phi_{\hat{\mathsf{X}}}[k]-\phi_{\mathsf{X}}[k]|^2\right]}{1/(M\log d)}\cdot\frac{1}{1/(4|\mathsf{X}[k]|^2)}=1, \tag{F.41}$$

which proves (5.5).

G Proof of Proposition 5.4: Circulant Gaussian noise

The proof strategy for Proposition 5.4 closely follows that of the i.i.d. Gaussian case (Theorem 4.1), with appropriate modifications to handle circulant noise. The necessary assumptions and notations are introduced in Appendix G.1. Appendix G.2 establishes the asymptotic convergence of the EfN estimator as $M\to\infty$ under circulant Gaussian noise statistics. In Appendix G.3, we show that conditioning the EfN process on a single Fourier noise coefficient results in a cyclo-stationary process with a cosine trend. Appendix G.4 extends the vanishing imaginary part result from Appendix B.1 to the setting of circulant Gaussian noise. Similarly, Appendix G.5 extends the result of Appendix B.2, showing that the real part remains strictly positive in the circulant case. Finally, Appendix G.6 combines the results of the preceding sections to complete the proof of Proposition 5.4.

G.1 Preliminaries

Let $\{y_i\}_{i=0}^{M-1}\sim\mathcal{N}(0,\Sigma)$, where $\Sigma\in\mathbb{R}^{d\times d}$ is a real, symmetric, and circulant covariance matrix with strictly positive eigenvalues (i.e., Σ is positive-definite). Let $\mathsf{Y}_i=\mathcal{F}\{y_i\}\in\mathbb{C}^d$ denote the DFT of y_i. The random vector Y_i satisfies the following properties:

1. *Diagonalization by the DFT.* Since Σ is circulant, it is diagonalized by the DFT: $\Sigma=F^*\Lambda F$, where F is the DFT matrix, $\Lambda=\mathrm{diag}(\lambda_0,\dots,\lambda_{d-1})$ contains the eigenvalues of Σ, given by the DFT of its first row. As Σ positive-definite, all eigenvalues $\lambda_k\in\mathbb{R}$ and $\lambda_k>0$ for all $k\in\{0,1,\dots d-1\}$.

2. *Distribution of Fourier coefficients.* The vector Y_i is complex Gaussian with distribution $\mathcal{CN}(0,\Lambda)$. Its entries are independent (but not identically distributed) complex Gaussian random variables, satisfying, $\mathbb{E}[\mathsf{Y}_i[k]]=0$, and $\mathbb{E}[\mathsf{Y}_i[k]\overline{\mathsf{Y}_i[\ell]}]=\lambda_k\delta_{k,\ell}$, for every $k,\ell\in\{0,1,\dots d/2\}$.

3. *Fourier phases distribution.* For any k such that $\lambda_k > 0$, the Fourier coefficient $\mathsf{Y}_i[k]$ is a zero-mean, circularly symmetric complex Gaussian random variable. Hence, the phases $\{\phi_{\mathsf{Y}_i}[k]\}_{k=0}^{d/2}$ are i.i.d. and uniformly distributed on $[-\pi, \pi)$ and independent of the magnitude $\{|\mathsf{Y}_i[k]|\}_{k=0}^{d/2}$.

4. *Conjugate symmetry.* Since $y_i \in \mathbb{R}^d$, the DFT satisfies the Hermitian symmetry:

$$\mathsf{Y}_i[d-k] = \overline{\mathsf{Y}_i[k]}, \quad \text{for } 1 \le k \le d-1. \tag{G.1}$$

Thus, only the first $d/2+1$ Fourier coefficients are independent; the rest are determined by conjugate symmetry.

Remark G.1. *To avoid confusion with the i.i.d. Gaussian case, we use the notation y_i and Y_i, rather than n_i and N_i, to denote Gaussian noise with a symmetric circulant covariance matrix.*

G.2 The convergence of the Einstein from Noise estimator

Similar to the derivation in Appendix A.2, the EfN estimator in the setting of circulant Gaussian noise, $\{y_i\}_{i=0}^{M-1} \sim \mathcal{N}(0, \Sigma)$, can be expressed explicitly as:

$$\hat{\mathsf{X}}[k] = \frac{1}{M} \sum_{i=0}^{M-1} |\mathsf{Y}_i[k]| \; e^{j\phi_{\mathsf{Y}_i}[k]} \, e^{j\frac{2\pi k}{d}\hat{\mathsf{R}}_i} \tag{G.2}$$

$$= \frac{e^{j\phi_{\mathsf{X}}[k]}}{M} \sum_{i=0}^{M-1} |\mathsf{Y}_i[k]| \; e^{j\phi_{\mathsf{Y}_i}[k]} \, e^{j\frac{2\pi k}{d}\hat{\mathsf{R}}_i} \, e^{-j\phi_{\mathsf{X}}[k]} \tag{G.3}$$

$$= \frac{e^{j\phi_{\mathsf{X}}[k]}}{M} \sum_{i=0}^{M-1} |\mathsf{Y}_i[k]| \; e^{j\phi_{e,i}[k]}, \tag{G.4}$$

where the shifts $\hat{\mathsf{R}}_i$ are given by

$$\hat{\mathsf{R}}_i \triangleq \arg\max_{0 \le r \le d-1} \langle y_i, \mathcal{T}_r x \rangle \tag{G.5}$$

$$= \arg\max_{0 \le r \le d-1} \langle \mathcal{F}\{y_i\}, \mathcal{F}\{\mathcal{T}_r x\} \rangle \tag{G.6}$$

$$= \arg\max_{0 \le r \le d-1} \sum_{k=0}^{d-1} |\mathsf{X}[k]| \, |\mathsf{Y}_i[k]| \, \cos\left(\frac{2\pi k r}{d} + \phi_{\mathsf{Y}_i}[k] - \phi_{\mathsf{X}}[k]\right). \tag{G.7}$$

and the phase difference is defined as,

$$\phi_{e,i}[k] \triangleq \frac{2\pi k \hat{\mathsf{R}}_i}{d} + \phi_{\mathsf{Y}_i}[k] - \phi_{\mathsf{X}}[k]. \tag{G.8}$$

To simplify notation, define for each $r \in \{0, 1, \ldots, d-1\}$,

$$\mathsf{S}_i[r] \triangleq \sum_{k=0}^{d-1} |\mathsf{X}[k]| \, |\mathsf{Y}_i[k]| \, \cos\left(\frac{2\pi k r}{d} + \phi_{\mathsf{Y}_i}[k] - \phi_{\mathsf{X}}[k]\right), \tag{G.9}$$

so that $\hat{\mathsf{R}}_i = \arg\max_{0\le r\le d-1} \mathsf{S}_i[r]$. We note that for any $0 \le i \le M-1$, the random vector $\mathsf{S}_i \triangleq (\mathsf{S}_i[0], \mathsf{S}_i[1], \ldots, \mathsf{S}_i[d-1])^\top$ is jointly Gaussian with zero mean and a circulant covariance matrix (as it is a Fourier transform of the convolution between y_i and the template x). Hence, S_i forms a cyclo-stationary process. Applying the strong law of large numbers (SLLN), as $M \to \infty$, we have,

$$\hat{\mathsf{X}}[k]\, e^{-j\phi_{\mathsf{X}}[k]} = \frac{1}{M}\sum_{i=0}^{M-1} |\mathsf{Y}_i[k]|\; e^{j\phi_{e,i}[k]} \tag{G.10}$$

$$\xrightarrow{\text{a.s.}} \mathbb{E}\left[|\mathsf{Y}_1[k]| \cos\left(\phi_{e,1}[k]\right)\right] + j\,\mathbb{E}\left[|\mathsf{Y}_1[k]| \sin\left(\phi_{e,1}[k]\right)\right], \tag{G.11}$$

where we have used the fact that the sequences of random variables $\{|\mathsf{Y}_i[k]|\cos(\phi_{e,i}[k])\}_{i=0}^{M-1}$ and $\{|\mathsf{Y}_i[k]|\sin(\phi_{e,i}[k])\}_{i=0}^{M-1}$ are i.i.d. with finite means and variances. Finally, we define for each k,

$$\mu_{\mathsf{A},k} \triangleq \mathbb{E}\left[|\mathsf{Y}_1[k]| \sin\left(\phi_{e,1}[k]\right)\right], \tag{G.12}$$

$$\mu_{\mathsf{B},k} \triangleq \mathbb{E}\left[|\mathsf{Y}_1[k]| \cos\left(\phi_{e,1}[k]\right)\right], \tag{G.13}$$

as the asymptotic imaginary and real parts of $\hat{\mathsf{X}}[k]e^{-j\phi_{\mathsf{X}}[k]}$, respectively.

G.3 Conditioning on the Fourier frequency noise component

We now extend the result of Lemma A.1 to the case where the noise follows a general Gaussian distribution with a real, symmetric, and circulant covariance matrix. That is, we consider observations $\{y_i\}_{i=0}^{M-1} \sim \mathcal{N}(0, \Sigma)$, where $\Sigma \in \mathbb{R}^{d\times d}$ is circulant and symmetric. In this setting, we establish the following result.

Lemma G.2. *Let S_i be defined as in (G.9), and denote $\mathbb{E}\left[|\mathsf{Y}_i[k]|^2\right] = \lambda_k > 0$ for each $k \in \{0, 1, \ldots, d-1\}$. Then, for every $k \in \left\{1, 2, \ldots, \frac{d}{2}-1, \frac{d}{2}+1, \ldots, d-1\right\}$, the random vector S_i conditioned on $\mathsf{Y}_i[k]$ is Gaussian:*

$$\mathsf{S}_i|\mathsf{Y}_i[k] \sim \mathcal{N}(\mu_{k,i}, \Sigma_{k,i}), \tag{G.14}$$

with mean and covariance given by,

$$\mu_{k,i}[r] \triangleq \mathbb{E}\left[\mathsf{S}_i[r]|\mathsf{Y}_i[k]\right] = 2\,|\mathsf{X}[k]|\,|\mathsf{Y}_i[k]| \cos\left(\frac{2\pi k r}{d} + \phi_{\mathsf{Y}_i}[k] - \phi_{\mathsf{X}}[k]\right), \tag{G.15}$$

for $0 \le r \le d-1$, and

$$\begin{aligned}\Sigma_{k,i}[r,s] &\triangleq \mathbb{E}\left[\left(\mathsf{S}_i[r] - \mathbb{E}\mathsf{S}_i[r]\right)\left(\mathsf{S}_i[s] - \mathbb{E}\mathsf{S}_i[s]\right)|\mathsf{Y}_i[k]\right] \\ &= \sum_{\ell=0}^{d-1} \frac{\lambda_\ell}{2} \cdot |\widetilde{\mathsf{X}}_k[\ell]|^2 \cos\left(\frac{2\pi\ell}{d}(r-s)\right),\end{aligned} \tag{G.16}$$

for $0 \le r, s \le d-1$, where $\widetilde{\mathsf{X}}_k$ is defined by:

$$\widetilde{\mathsf{X}}_k[\ell] \triangleq \begin{cases} 0 & \text{if}\,\ell = k, d-k, \\ \mathsf{X}[\ell] & \text{if}\,\ell = 0, d/2, \\ \sqrt{2}\cdot\mathsf{X}[\ell] & \text{otherwise.}\end{cases} \tag{G.17}$$

Note that the conditional process $\mathsf{S}_i|\mathsf{Y}_i[k]$ is Gaussian because it is given by a linear transform of i.i.d. Gaussian variables. Also, since its covariance matrix is circulant and depends only on the difference between the two indices, i.e., $\Sigma_{k,i}[r,s] = \sigma_{k,i}[|r-s|]$, it is cycle-stationary with a cosine trend. The eigenvalues of this circulant matrix are given by the DFT of its first row, and thus its ℓ-th eigenvalue equals $\lambda_\ell \cdot |\widetilde{\mathsf{X}}_k[\ell]|^2$, for $0 \le \ell \le d-1$.

Remark G.3. *When the noise is i.i.d. Gaussian, that is, $y_i \sim \mathcal{N}(0, \sigma^2 I_{d\times d})$, the eigenvalues of the covariance matrix satisfy $\lambda_\ell = \sigma^2$ for all $\ell \in \{0,1,\ldots,d-1\}$. In this case, the general setting reduces to the one considered in Lemma A.1, thereby recovering its result.*

Proof of Lemma G.2. We recall that if $\{y_i\}_{i=0}^{M-1} \sim \mathcal{N}(0,\Sigma)$, for symmetric circulant matrix Σ, then their DFT coefficients satisfy $\{|\mathsf{Y}_i[k]|\}_{k=0}^{d/2}$, and $\{\phi_{\mathsf{Y}_i}[k]\}_{k=0}^{d/2}$ are independent and $\{\phi_{\mathsf{Y}_i}[k]\}_{k=0}^{d-1} \sim \mathsf{Unif}[-\pi,\pi)$. By definition of S_i (G.9), we have for every $k \neq 0, d/2$,

$$\begin{aligned}\mathsf{S}_i[r]\,|\mathsf{Y}_i[k] =& 2\,|\mathsf{X}[k]|\,|\mathsf{Y}_i[k]|\cos\left(\frac{2\pi k r}{d} + \phi_{\mathsf{Y}_i}[k] - \phi_{\mathsf{X}}[k]\right)\\ &+ \sum_{\ell\neq k,d-k} |\mathsf{X}[\ell]|\,|\mathsf{Y}_i[\ell]|\,\cos\left(\frac{2\pi \ell r}{d} + \phi_{\mathsf{Y}_i}[\ell] - \phi_{\mathsf{X}}[\ell]\right),\end{aligned} \tag{G.18}$$

where we have used the property of $\mathsf{X}[k] = \overline{\mathsf{X}[d-k]}$, $\mathsf{Y}_i[k] = \overline{\mathsf{Y}_i[d-k]}$. Clearly, as $\mathbb{E}\left[\mathsf{Y}_i[\ell]\right] = 0$, for every $0 \le \ell \le d-1$, we have,

$$\mathbb{E}\left[|\mathsf{X}[\ell]|\,|\mathsf{Y}_i[\ell]|\,\cos\left(\frac{2\pi \ell r}{d} + \phi_{\mathsf{Y}_i}[\ell] - \phi_{\mathsf{X}}[\ell]\right)\right] = 0, \tag{G.19}$$

for every $0 \le \ell \le d-1$. Combining (G.18) and (G.19) results,

$$\mu_{k,i}[r] = \mathbb{E}\left[\mathsf{S}_i[r]|\mathsf{Y}_i[k]\right] = 2\,|\mathsf{X}[k]|\,|\mathsf{Y}_i[k]|\cos\left(\frac{2\pi k r}{d} + \phi_{\mathsf{Y}_i}[k] - \phi_{\mathsf{X}}[k]\right), \tag{G.20}$$

proving the first result concerning the means.

The covariance term. In the following, we derive the covariance term,

$$\Sigma_{k,i}[r,s] \triangleq \mathbb{E}\left[\left(\mathsf{S}_i[r] - \mathbb{E}\mathsf{S}_i[r]\right)\left(\mathsf{S}_i[s] - \mathbb{E}\mathsf{S}_i[s]\right)|\mathsf{Y}_i[k]\right]. \tag{G.21}$$

Let,

$$\begin{aligned}\rho_{k,i}[r] &\triangleq \mathsf{S}_i[r] - \mathbb{E}\mathsf{S}_i[r]\\ &= \sum_{\ell\neq k,d-k} |\mathsf{X}[\ell]|\,|\mathsf{Y}_i[\ell]|\,\cos\left(\frac{2\pi \ell r}{d} + \phi_{\mathsf{Y}_i}[\ell] - \phi_{\mathsf{X}}[\ell]\right),\end{aligned} \tag{G.22}$$

and denote,

$$\mathcal{I} = \{1, 2, \ldots k-1, k+1, \ldots, d/2-1\}, \tag{G.23}$$

which is the indices of the Fourier coefficients, excluding $\{0, k, d/2\}$. As the sequences $\{|\mathsf{Y}_i[\ell]|\}_{\ell=0}^{d/2}$ and $\{\phi_{\mathsf{Y}_i}[\ell]\}_{\ell=0}^{d/2}$ are statistically independent, and satisfy $\mathsf{Y}_i[\ell] = \overline{\mathsf{Y}_i[d-\ell]}$ and $\mathsf{X}[\ell] = \overline{\mathsf{X}[d-\ell]}$, we have,

$$\begin{aligned}\rho_{k,i}[r] &= \sum_{\ell \neq k, d-k} |\mathsf{X}[\ell]|\,|\mathsf{Y}_i[\ell]| \cos\left(\frac{2\pi\ell r}{d} + \phi_{\mathsf{Y}_i}[\ell] - \phi_{\mathsf{X}}[\ell]\right) = \\ &= \sum_{\ell \in \{0, d/2\}} |\mathsf{X}[\ell]|\,|\mathsf{Y}_i[\ell]| \cos\left(\frac{2\pi\ell r}{d} + \phi_{\mathsf{Y}_i}[\ell] - \phi_{\mathsf{X}}[\ell]\right) \\ &\quad + 2 \cdot \sum_{\ell \in \mathcal{I}} |\mathsf{X}[\ell]|\,|\mathsf{Y}_i[\ell]| \cos\left(\frac{2\pi\ell r}{d} + \phi_{\mathsf{Y}_i}[\ell] - \phi_{\mathsf{X}}[\ell]\right), \end{aligned} \tag{G.24}$$

where each one of the terms in the sum is independent. Since the terms in the sum on the r.h.s. of (G.24) are independent (i.e., $\mathbb{E}\left[\mathsf{Y}_i[\ell_1]\,\overline{\mathsf{Y}_i[\ell_2]}\right] = \mathbb{E}\left[|\mathsf{Y}_i[\ell_1]|^2\right]\delta_{\ell_1,\ell_2}$), it follows that,

$$\begin{aligned}\Sigma_{k,i}[r,s] &= \mathbb{E}\left[\rho_{k,i}[r]\,\rho_{k,i}[s]\,|\mathsf{Y}_i[k]\right] \\ &= \mathbb{E}\left[\sum_{\ell \in \{0, d/2\}} |\mathsf{X}[\ell]|^2\,|\mathsf{Y}_i[\ell]|^2 \cos\left(\frac{2\pi\ell r}{d} + \phi_{\mathsf{Y}_i}[\ell] - \phi_{\mathsf{X}}[\ell]\right)\cos\left(\frac{2\pi\ell s}{d} + \phi_{\mathsf{Y}_i}[\ell] - \phi_{\mathsf{X}}[\ell]\right)\right] \\ &\quad + 4 \cdot \mathbb{E}\left[\sum_{\ell \in \mathcal{I}} |\mathsf{X}[\ell]|^2\,|\mathsf{Y}_i[\ell]|^2 \cos\left(\frac{2\pi\ell r}{d} + \phi_{\mathsf{Y}_i}[\ell] - \phi_{\mathsf{X}}[\ell]\right)\cos\left(\frac{2\pi\ell s}{d} + \phi_{\mathsf{Y}_i}[\ell] - \phi_{\mathsf{X}}[\ell]\right)\right]. \end{aligned} \tag{G.25}$$

The expectation value in (G.25) is composed of the multiplications of cosines. Applying trigonometric identities, we obtain,

$$\begin{aligned}&\cos\left(\frac{2\pi\ell r}{d} + \phi_{\mathsf{Y}_i}[\ell] - \phi_{\mathsf{X}}[\ell]\right)\cos\left(\frac{2\pi\ell s}{d} + \phi_{\mathsf{Y}_i}[\ell] - \phi_{\mathsf{X}}[\ell]\right) \\ &\quad = \frac{1}{2}\cos\left(\frac{2\pi\ell(r-s)}{d}\right) + \frac{1}{2}\cos\left(\frac{2\pi\ell(r+s)}{d} + 2\left(\phi_{\mathsf{Y}_i}[\ell] - \phi_{\mathsf{X}}[\ell]\right)\right), \end{aligned} \tag{G.26}$$

for every $0 \leq r, s \leq d-1$. Now, since the sequences $\{|\mathsf{Y}_i[\ell]|\}_{\ell=0}^{d/2}$ and $\{\phi_{\mathsf{Y}_i}[\ell]\}_{\ell=0}^{d/2}$ are independent random variables, with $\mathbb{E}\left[|\mathsf{Y}_i[k]|^2\right] = \lambda_k$ and phases $\phi_{\mathsf{Y}_i}[k]$ uniformly distributed over $[-\pi, \pi)$, and by applying the trigonometric identity (G.26), it follows that,

$$\begin{aligned}&\mathbb{E}\left[|\mathsf{Y}_i[\ell]|^2 \cos\left(\frac{2\pi\ell r}{d} + \phi_{\mathsf{Y}_i}[\ell] - \phi_{\mathsf{X}}[\ell]\right)\cos\left(\frac{2\pi\ell s}{d} + \phi_{\mathsf{Y}_i}[\ell] - \phi_{\mathsf{X}}[\ell]\right)\right] \\ &\quad = \frac{1}{2}\mathbb{E}\left[|\mathsf{Y}_i[\ell]|^2\right]\cos\left(\frac{2\pi\ell(r-s)}{d}\right) = \frac{\lambda_\ell}{2}\cos\left(\frac{2\pi\ell(r-s)}{d}\right). \end{aligned} \tag{G.27}$$

Substituting (G.27) into (G.25) leads to,

$$\mathbb{E}\left[\rho_{k,i}[r]\,\rho_{k,i}[s]\,|\mathsf{Y}_i[k]\right] = \sum_{\ell \in \{0, d/2\}} \frac{\lambda_\ell}{2} \cdot |\mathsf{X}[\ell]|^2 \cos\left(\frac{2\pi\ell}{d}(r-s)\right)$$

$$+ 4 \cdot \sum_{\ell \in \mathcal{I}} \frac{\lambda_\ell}{2} \cdot |\mathsf{X}[\ell]|^2 \cos\left(\frac{2\pi\ell}{d}(r-s)\right). \tag{G.28}$$

As for every $\ell \in \mathcal{I}$, $|\mathsf{X}[\ell]| = |\mathsf{X}[d-\ell]|$, we have,

$$4 \cdot \sum_{\ell \in \mathcal{I}} \frac{\lambda_\ell}{2} |\mathsf{X}[\ell]|^2 \cos\left(\frac{2\pi\ell}{d}(r-s)\right) = 2 \sum_{\ell \neq \{0,k,d/2,d-k\}} \frac{\lambda_\ell}{2} |\mathsf{X}[\ell]|^2 \cos\left(\frac{2\pi\ell}{d}(r-s)\right). \tag{G.29}$$

Substituting (G.29) into (G.28), we get,

$$\mathbb{E}\left[\rho_{k,i}[r]\,\rho_{k,i}[s]\,|\mathsf{Y}_i[k]|\right] = \sum_{\ell=0}^{d-1} \frac{\lambda_\ell}{2} \cdot |\widetilde{\mathsf{X}}_k[\ell]|^2 \cos\left(\frac{2\pi\ell}{d}(r-s)\right), \tag{G.30}$$

for $\widetilde{\mathsf{X}}_k[\ell]$ as defined in (G.17), which completes the proof. □

G.4 Convergence of the Fourier phases

Similarly to Appendix B.1 and Lemma B.1, we show here that the imaginary part in (G.11) vanishes. The key observation is that $\{|\mathsf{Y}_i[k]|\}_{k=0}^{d-1}$, and $\{\phi_{\mathsf{Y}_i}[k]\}_{k=0}^{d-1}$ are statistically independent and $\{\phi_{\mathsf{Y}_i}[k]\}_{k=0}^{d-1} \sim \mathsf{Unif}[-\pi,\pi)$.

Lemma G.4. *Recall the definition of $\phi_{e,i}[k]$ in (G.8). Then,*

$$\mu_{\mathsf{A},k} = \mathbb{E}\left[|\mathsf{Y}_1[k]| \sin(\phi_{e,1}[k])\right] = 0, \tag{G.31}$$

for every $0 \le k \le d-1$.

Proof of Lemma G.4. Let $\mathsf{D}[k] \triangleq \phi_{\mathsf{X}}[k] - \phi_{\mathsf{Y}_1}[k]$, and recall the definition of $\hat{\mathsf{R}}_i$ in (A.4), i.e.,

$$\hat{\mathsf{R}}_i = \underset{0 \le r \le d-1}{\arg\max} \sum_{k=0}^{d-1} |\mathsf{X}[k]|\,|\mathsf{Y}_i[k]| \cos\left(\frac{2\pi k r}{d} + \phi_{\mathsf{Y}_i}[k] - \phi_{\mathsf{X}}[k]\right). \tag{G.32}$$

Note that $\hat{\mathsf{R}}_i$ is a function of

$$\hat{\mathsf{R}}_i = \hat{\mathsf{R}}_i\left(\{|\mathsf{Y}_i[k]|\}_{k=0}^{d-1}, \{|\mathsf{X}[k]|\}_{k=0}^{d-1}, \{\phi_{\mathsf{Y}_i}[k]\}_{k=0}^{d-1}, \{\phi_{\mathsf{X}}[k]\}_{k=0}^{d-1}\right), \tag{G.33}$$

and it depends on $\phi_{\mathsf{Y}_i}[k]$ and $\phi_{\mathsf{X}}[k]$ only through $\mathsf{D}[k]$. Accordingly, viewing $\hat{\mathsf{R}}_1$ as a function of $\mathsf{D}[k]$, for fixed $\{|\mathsf{Y}_i[k]|\}_{k=0}^{d-1}, \{|\mathsf{X}[k]|\}_{k=0}^{d-1}$, we have,

$$\hat{\mathsf{R}}_1\left(-\mathsf{D}[0], -\mathsf{D}[1], \ldots, -\mathsf{D}[d-1]\right) = -\hat{\mathsf{R}}_1\left(\mathsf{D}[0], \mathsf{D}[1], \ldots, \mathsf{D}[d-1]\right). \tag{G.34}$$

Namely, from symmetry arguments, by flipping the signs of all the phases, the location of the maximum flips its sign as well. Then, by the law of total expectation,

$$\mu_{\mathsf{A},k} = \mathbb{E}\left[|\mathsf{Y}_1[k]| \sin\left(\frac{2\pi k}{d}\hat{\mathsf{R}}_1 + \phi_{\mathsf{Y}_1}[k] - \phi_{\mathsf{X}}[k]\right)\right]$$

$$= \mathbb{E}\left\{|\mathsf{Y}_1[k]| \cdot \mathbb{E}\left[\sin\left(\frac{2\pi k}{d}\hat{\mathsf{R}}_1 + \phi_{\mathsf{Y}_1}[k] - \phi_{\mathsf{X}}[k]\right)\middle| \{|\mathsf{Y}_1[k]|\}_{k=0}^{d-1}\right]\right\}. \tag{G.35}$$

The inner expectation in (G.35) is taken w.r.t. the uniform randomness of the phases $\{\phi_{\mathsf{Y}_1}[k]\}_{k=0}^{d-1} \in [-\pi, \pi)$. However, due to (G.34), and since the sine function is odd around zero, the integration in (G.35) nullifies. Therefore,

$$\mathbb{E}\left[\sin\left(\frac{2\pi k}{d}\hat{\mathsf{R}}_1 + \phi_{\mathsf{Y}_1}[k] - \phi_{\mathsf{X}}[k]\right)\middle| \{|\mathsf{Y}_1[k]|\}_{k=0}^{d-1}\right] = 0, \tag{G.36}$$

and thus $\mu_{\mathsf{A},k} = 0$. □

G.5 Convergence to non-vanishing signal

In analogy with Appendix B.2 and Proposition B.2, we now establish that the real part of (G.11) does not vanish.

Proposition G.5. *Recall the definition of $\phi_{e,i}[k]$ in (G.8). Fix $d \in \mathbb{N}$, and assume that $\mathsf{X}[k] \neq 0$ for all $0 < k \leq d-1$. Then, for any $0 \leq k \leq d-1$,*

$$\mu_{\mathsf{B},k} \triangleq \mathbb{E}[|\mathsf{Y}_1[k]| \cos(\phi_{e,1}[k])] > 0. \tag{G.37}$$

Proof of Proposition G.5. By the law of total expectation, we have,

$$\begin{aligned}\mathbb{E}[|\mathsf{Y}_1[k]| \cos(\phi_{e,1}[k])] &= \mathbb{E}\left[|\mathsf{Y}_1[k]| \cdot \mathbb{E}\left(\cos(\phi_{e,1}[k])\middle| \mathsf{Y}_1[k]\right)\right] \\ &= \mathbb{E}\left[|\mathsf{Y}_1[k]| \cdot \mathbb{E}\left(\cos\left(\frac{2\pi k\hat{\mathsf{R}}_1}{d} + \phi_{\mathsf{Y}_1}[k] - \phi_{\mathsf{X}}[k]\right)\middle| \mathsf{Y}_1[k]\right)\right].\end{aligned} \tag{G.38}$$

More explicitly, we can write,

$$\begin{aligned}&\mathbb{E}[|\mathsf{Y}_1[k]| \cos(\phi_{e,1}[k])] = \\ &\frac{1}{2\pi}\int_0^\infty \mathrm{d}y\ y f_{|\mathsf{Y}_1[k]|}(y) \int_{-\pi}^{\pi} \mathrm{d}\varphi \mathbb{E}\left[\cos\left(\frac{2\pi k}{d}\hat{\mathsf{R}}_1 + \varphi\right)\middle| |\mathsf{Y}_1[k]| = y, \phi_{\mathsf{Y}_1}[k] = \phi_{\mathsf{X}}[k] + \varphi\right].\end{aligned} \tag{G.39}$$

Now, note that the inner integral can be written as,

$$\begin{aligned}&\int_{-\pi}^{\pi} \mathrm{d}\varphi \mathbb{E}\left[\cos\left(\frac{2\pi k}{d}\hat{\mathsf{R}}_1 + \varphi\right)\middle| |\mathsf{Y}_1[k]| = y, \phi_{\mathsf{Y}_1}[k] = \phi_{\mathsf{X}}[k] + \varphi\right] \\ &\quad = \int_0^{\pi} \mathrm{d}\varphi\ \mathbb{E}\left[\cos\left(\frac{2\pi k}{d}\hat{\mathsf{R}}_1 + \varphi\right)\middle| |\mathsf{Y}_1[k]| = y, \phi_{\mathsf{Y}_1}[k] = \phi_{\mathsf{X}}[k] + \varphi\right] + \\ &\quad + \int_0^{\pi} \mathrm{d}\varphi\ \mathbb{E}\left[\cos\left(\frac{2\pi k}{d}\hat{\mathsf{R}}_1 + \varphi + \pi\right)\middle| |\mathsf{Y}_1[k]| = y, \phi_{\mathsf{Y}_1}[k] = \phi_{\mathsf{X}}[k] + \varphi + \pi\right].\end{aligned} \tag{G.40}$$

Next, we apply Proposition A.5. Using its notation, we define the Gaussian process,

$$\mathsf{S}^{(+)} = \mathsf{S}_1 | \mathsf{Y}_1[k], \tag{G.41}$$

where the r.h.s. follows from (G.14). By (G.15), the mean vector of $\mathsf{S}_1|\mathsf{Y}_1[k]$ has a cosine trend, as assumed in Proposition A.5 in (A.59). Additionally, $\mathsf{S}_1|\mathsf{Y}_1[k]$ is a Gaussian cyclostationary process, as described in (G.16). The final condition to verify is (A.60), which is satisfied by Lemma A.4 (and applies also to circulant Gaussian noise statistics as well).

Since the conditional distribution of $\hat{\mathsf{R}}_1$ given $\{|\mathsf{Y}_1[k]| = y, \phi_{\mathsf{Y}_1}[k] = \phi_{\mathsf{X}}[k] + \varphi\}$ matches that of $\hat{\mathsf{R}}^{(+)}$ in (A.57), and similarly, given $\{|\mathsf{Y}_1[k]| = y, \phi_{\mathsf{Y}_1}[k] = \phi_{\mathsf{X}}[k] + \varphi + \pi\}$, it matches $\hat{\mathsf{R}}^{(-)}$ in (A.58), the sum of the integrands on the right-hand side of (G.40) equals the left-hand side of (A.63). By Proposition A.5, this sum is positive for all $\varphi \in [0, \pi]$. Together with (G.39), this completes the proof of Proposition G.5.

□

G.6 Proof of Proposition 5.4

We are now ready to prove Proposition 5.4. By the definition of the phase difference between the template x and the EfN estimator $\hat{x}$ (as in (A.6)), we have,

$$\phi_{\hat{\mathsf{X}}}[k] - \phi_{\mathsf{X}}[k] = \arctan\left(\frac{\sum_{i=0}^{M-1} |\mathsf{Y}_i[k]| \sin\left(\phi_{e,i}[k]\right)}{\sum_{i=0}^{M-1} |\mathsf{Y}_i[k]| \cos\left(\phi_{e,i}[k]\right)}\right), \tag{G.42}$$

Using the continuous mapping theorem, it is evident that it suffices to prove that,

$$\frac{\sum_{i=0}^{M-1} |\mathsf{Y}_i[k]| \sin\left(\phi_{e,i}[k]\right)}{\sum_{i=0}^{M-1} |\mathsf{Y}_i[k]| \cos\left(\phi_{e,i}[k]\right)} \xrightarrow{\text{a.s.}} 0. \tag{G.43}$$

This, however, follows by applying the SLLN,

$$\frac{\sum_{i=0}^{M-1} |\mathsf{Y}_i[k]| \sin\left(\phi_{e,i}[k]\right)}{\sum_{i=0}^{M-1} |\mathsf{Y}_i[k]| \cos\left(\phi_{e,i}[k]\right)} \xrightarrow{\text{a.s.}} \frac{\mu_{\mathsf{A},k}}{\mu_{\mathsf{B},k}}, \tag{G.44}$$

where $\mu_{\mathsf{A},k} \triangleq \mathbb{E}\left[|\mathsf{Y}_1[k]| \sin(\phi_{e,1}[k])\right]$ and $\mu_{\mathsf{B},k} \triangleq \mathbb{E}\left[|\mathsf{Y}_1[k]| \cos(\phi_{e,1}[k])\right]$, defined in (G.12), and (G.13), respectively. By Lemma G.2, $\mu_{\mathsf{A},k} = 0$, while by Proposition G.5, we have that $\mu_{\mathsf{B},k} > 0$, and thus their ratio converges a.s. to zero by the continuous mapping theorem. Thus, we proved that $\phi_{\hat{\mathsf{X}}}[k] \xrightarrow{\text{a.s.}} \phi_{\mathsf{X}}[k]$. Finally, we prove the convergence rate, given in (5.5). According to Proposition B.4, whose assumptions apply for the case of circulant Gaussian noise statistics as well, we have,

$$\lim_{M\to\infty} \frac{\mathbb{E}|\phi_{\hat{\mathsf{X}}}[k] - \phi_{\mathsf{X}}[k]|^2}{1/M} = \frac{\mathbb{E}\left(\left[|\mathsf{Y}_1[k]| \sin(\phi_{e,1}[k])\right]^2\right)}{\mathbb{E}[|\mathsf{Y}_1[k]| \cos(\phi_{e,1}[k])]^2} < \infty, \tag{G.45}$$

which completes the proof of the Proposition.