

Published in final edited form as:

*Neural Comput.* 2013 April ; 25(4): 901–921. doi:10.1162/NECO\_a\_00422.

## Some sampling properties of common phase estimators

Kyle Q. Lepage, Mark A. Kramer, and Uri T. Eden

Kyle Q. Lepage: lepage@math.bu.edu; Mark A. Kramer: mak@bu.edu; Uri T. Eden: tzvi@math.bu.edu

### Abstract

The instantaneous phase of neural rhythms is important to many neuroscience related studies. In this short note, the statistical sampling properties of three instantaneous phase estimators commonly employed to analyze neuroscience data are shown to share common features allowing for an analytical investigation into their behaviour. These three phase estimators – the Hilbert, complex Morlet, and discrete Fourier transform – are each shown to maximize the likelihood of the data, assuming the observation of different neural signals. This connection, explored with the use of a geometric argument, is used to describe the bias and variance properties of each of the phase estimators, their temporal dependence, as well as the effect of model miss-specification. This analysis suggests how prior knowledge about a rhythmic signal can be used to improve the accuracy of phase estimates.

### 1 Introduction

Coordination of brain activity across spatial and temporal scales is essential to healthy brain function, (Quyen, 2011). Understanding the mechanisms supporting this coordination remains an active research area. Rhythmic neural activity has been proposed to support dynamic coordination of collective neural behavior and function (Buzsaki, 2006; Buzsáki and Draguhn, 2004; Gray et al., 1989; Schroeder and Lakatos, 2008; Varela et al., 2001). Although many measures exist to characterize neural rhythms (e.g., Pereda et al. (2005); Young and Eggermont (2009)), recent interest has focused on the phase of neural rhythms, (e.g., Quyen and Bragin (2007)). For example, phase synchronization has been proposed as a basic mechanism for integration of brain activity in studies such as Harris et al. (2003); Varela et al. (2001); Womelsdorf et al. (2007), and the phase of low frequency rhythms has been shown to modulate high frequency fluctuations (i.e., cross-frequency coupling) in Canolty et al. (2006); Canolty and Knight (2010); Cohen (2008); Lakatos et al. (2008); Tort et al. (2008).

This brief communication focuses upon three estimators of instantaneous rhythm phase. These estimators are the discrete Fourier transform (DFT) based estimator, an estimator employing the Hilbert transform, and a phase estimator using the complex Morlet wavelet. Studies employing the Hilbert transform based phase estimator to estimate the phase of neurological signals include Buzsaki et al. (2003); Canolty et al. (2006); Hentschke et al. (2007); Palva et al. (2005); Sirota et al. (2008); Tort et al. (2008). The complex Morlet phase estimator is used for the same purpose in studies such as Lakatos et al. (2005); Rizzuto et al. (2006, 2003); Rutishauser et al. (2010). The DFT phase estimator is classically treated in books, for e.g. (Kay, 1993, p. 33).

The relative performance of the DFT, Hilbert and complex Morlet phase estimators has been documented in the neuroscience literature. In Guevara et al. (2005), the effect of reference on phase estimators applied to electroencephalographic data is investigated and in Hurtado

et al. (2003) the phase-locking value is introduced which uses an a priori phase estimate. In these works comparisons are made via simulation without explicitly exploring the effect of parameter choice upon estimator accuracy and the relative performance of phase estimators for data in general. In Aviyente et al. (2011) & Aviyente and Mutlu (2011) a phase estimator based upon a time-frequency distribution is compared to a wavelet phase estimator. In these latter papers the statistical data model associated with the DFT phase estimator is presented.

The accuracy of the Hilbert and complex Morlet phase estimators is compared using both simulated and recorded neural data in Quyen et al. (2001), with a focus on derivative measures of phase synchrony. It is found that little practical difference exists between the Hilbert and complex Morlet phase estimators; but, the effect of parameter choice, time series length, noise level, signal amplitude and signal morphology is not fully addressed and relevant analytic expressions are not provided. In Bruns (2004), phase estimator differences are explored through deterministic calculation and by an analysis of a recording obtained from a subdurally implanted electrode on the cortical surface of a patient with epilepsy. It is found that two phase estimators considered in this scenario, identical to the Hilbert and complex Morlet phase estimators considered within the current work, can each be computed as a convolution of the data with an oscillatory window function. Further, when the oscillatory window functions are specified to have similar effective bandwidths, it is found that the Hilbert and Morlet phase estimators yield similar phase estimates. Unlike these previous works, the current work considers each method as a projection from which insight into the sampling properties of these phase estimators can be analytically derived. In this framework, by explicitly introducing noise and considering signal morphology, phase estimate accuracy is shown to explicitly depend on the number of measurements, on the additive noise variance, on the amplitude of the signal envelope, and upon the dimension of a relevant subspace. Further, it is shown that each of the three phase estimators considered are associated with different a priori information regarding the signal type, and each of these three estimators performs best when the actual signal present conforms to the a priori information. Thus, this paper further quantifies the nature of phase estimate accuracy in the neurosciences providing both an analytical basis for further theoretical advance as well as suggesting the appropriate use of a priori neural signal knowledge when estimating the phase of a neurological rhythm.

In this short note, by an application of classical statistics and a geometrical connection, the DFT, Hilbert, and complex Morlet instantaneous phase estimators are shown to be related; a result consistent with results presented in Bruns (2004) and in Quyen et al. (2001). Each of the three discussed estimators are shown to be maximum likelihood estimators; each estimator maximizing the likelihood associated with a unique, “matched” statistical data model. Facilitated by geometry, asymptotic estimator variance is given for each of the estimators, as well as induced temporal dependencies. All three can be computed as projection operations, where each method has its own projection matrix. The distribution of the phase estimator is determined by the effect of the projection on a signal vector term and upon a noise vector term. The note ends with a discussion on the impact of the current work upon practical instantaneous phase estimation. To the best of the authors’ knowledge this work is the first paper in the neuroscience literature documenting analytical sampling properties of commonly employed estimators of instantaneous phase.

## 2 Phase Estimator Properties

In the following sections, it is shown that the three estimators of instantaneous phase, the DFT, Hilbert and complex Morlet wavelet phase estimators, are each maximum-likelihood estimators associated with specific data models (Section (3)). The results of this paper are summarized in Table (1). Each row of Table (1) describes a specific type of instantaneous

phase estimator. The columns of each row, from left to right, specify the estimator, the signal model for which the estimator is least-squares optimal, the equation in the text defining the model matrix in the linear statistical model for which the phase estimator is maximum likelihood, the dimension,  $p$ , of the span of the columns of the model matrix, and the equivalent prior information that is incorporated into the associated data model. These results, in conjunction with the geometry described in Fig. (1), and the phase estimator probability density function Eqn. (18), describe the essence of this work.

### 3 Classical Theory

The following sections review relevant theory from classical statistics adapted to the problem of instantaneous phase estimation. This material will be used to develop the sampling properties of the three phase estimators discussed in this work.

#### 3.1 Statistical Linear Model

The three phase estimators discussed in this note – the DFT, Hilbert and complex Morlet phase estimators – all involve taking the complex angle of a complex valued time-series. As shown in Section (4), these estimators can be associated with differing signal models in a classical, linear statistical model of the analytic signal,  $\mathbf{d}_a$ , of the observed data,  $\mathbf{d}$ , for a specific frequency of interest  $f_0$ .<sup>1</sup> The analytic signal is a complex-valued representation of the data consisting of only positive frequency components. For real-valued measurements, the discrete Fourier transform of the data,  $\mathbf{d}$ , is redundant, in that all information is contained within the non-negative frequencies. Because of this, the analytic signal of the data,  $\mathbf{d}_a$ , is lossless, possessing all of the information contained within the original time-series of measurements. That is, the measured data,  $\mathbf{d}$  is related to the analytic signal,  $\mathbf{d}_a$ , as Kay (1993),

$$\mathbf{d} = 2\text{Re}\{\mathbf{d}_a\}, \quad (1)$$

where  $\text{Re}\{x\}$  denotes the real part of the complex variable  $x$ .<sup>2</sup> In this work, the analytic signal,  $\mathbf{d}_a$ , of the observed data,  $\mathbf{d}$  is modeled as,

$$\mathbf{d}_a = \mathbf{s}_a + \mathbf{n}_a. \quad (2)$$

Here,  $\mathbf{s}_a$  is the analytic signal representation of a rhythmic signal whose phase is to be estimated, and  $\mathbf{n}_a$  is a zero-mean noise vector with Gaussian distributed and independent elements, each with variance  $\sigma^2$ . Relaxation of the Gaussian assumption is discussed in Section (6). Each of the phase estimators considered in this work correspond to maximum likelihood estimators associated with unique models of the signal  $\mathbf{s}_a$ . The generic phase estimator,  $\hat{\phi}$  is the complex angle of the maximum-likelihood estimator of the signal,  $\hat{\mathbf{s}}_a$ ,

$$\hat{\phi} = \angle\{\hat{\mathbf{s}}_a\}, \quad (3)$$

and thus is also a maximum-likelihood estimator of the instantaneous phase,  $\phi$ , since the transform of a maximum likelihood estimator is itself a maximum likelihood estimator of the transformed quantity (Hogg et al., 2005, p. 316). Here  $\hat{\phi}$  is a vector of instantaneous phase estimators, the  $j^{\text{th}}$  element is associated with the phase ( $\phi_j$ ) of the signal, ( $\mathbf{s}_j$ ), evaluated

<sup>1</sup>In general, we will be interested in frequency intervals about  $f_0$ .

<sup>2</sup>This definition of the analytic signal differs from that obtained via the Hilbert transform by a factor of two but avoids subsequent scaling of the discrete Fourier transform by a factor of two.

at time-index  $j$ . It is useful to model this signal in a linear relationship with an unknown vector of parameters,  $\alpha$ . That is,

$$\mathbf{s}_a = \mathbf{H}\alpha. \quad (4)$$

Here the model matrix,  $\mathbf{H}$  depends on the type of phase estimator - DFT (dft), Hilbert (h), or complex Morlet (cm) - and has  $N$  rows, corresponding to the length of the data, and a number of columns that depends on the type of phase estimator.

### 3.2 Maximum-Likelihood Estimation

As previously discussed,  $\hat{\phi}$  is maximum-likelihood (ML) if  $\hat{\mathbf{s}}_a$  is maximum likelihood. The ML estimator of  $\mathbf{s}_a$  is constructed by first estimating its transform in the frequency domain. The log-likelihood,  $\ell(\alpha)$ , corresponding to the DFT of Eqn. (2) after substituting Eqn. (4), is

$$\ell(\alpha) = \frac{-1}{2\sigma^2} \times (\tilde{\mathbf{d}} - \mathbf{U}^\dagger \mathbf{H} \alpha)^\dagger (\tilde{\mathbf{d}} - \mathbf{U}^\dagger \mathbf{H} \alpha), \quad (5)$$

where  $\mathbf{U}$  is the Fourier matrix specified as,

$$(\mathbf{U})_{j,k} = \frac{e^{i2\pi f_k t_j}}{\sqrt{N}}, j, k = 0, \dots, \frac{N}{2} - 1, \quad (6)$$

where  $\tilde{\mathbf{d}}$  denotes discrete Fourier transformation of  $\mathbf{d}$ ,  $f_k = \frac{k}{N\Delta}$ ,  $t_j = j\Delta$ , and  $\Delta$  is the sample period. Here, without loss of generality,  $N$  is specified to be an integer power of 2. The  $j^{\text{th}}$  time-index and the  $k^{\text{th}}$  Fourier frequency are represented in the  $j^{\text{th}}$  row and the  $k^{\text{th}}$  column of  $\mathbf{U}$ . Thus specified,  $\mathbf{U}^\dagger \mathbf{d}$  computes the DFT,  $\tilde{\mathbf{d}}$  of the vector random data,  $\mathbf{d}$  evaluated at the frequencies,  $f_k$ ,  $k=0, \dots, \frac{N}{2}-1$  and is equal to  $\mathbf{U}^\dagger \mathbf{d}_a$ , due to the restriction of the computation to the non-negative frequencies. Here  $^\dagger$  denotes transpose conjugation. Since  $\mathbf{U}$  is invertible, all information contained in  $\mathbf{d}_a$  is contained in  $\tilde{\mathbf{d}}$ ; the operation is lossless.

Now,

$$\hat{\alpha}_{\text{mle}} = (\mathbf{H}^\dagger \mathbf{U} \mathbf{U}^\dagger \mathbf{H})^{-1} \mathbf{H}^\dagger \tilde{\mathbf{U}} \mathbf{d}, \quad (7)$$

$$= (\mathbf{H}^\dagger \mathbf{H})^{-1} \mathbf{H}^\dagger \mathbf{d}_a. \quad (8)$$

Thus, the projection of the analytic signal of the data onto the space spanned by the columns of  $\mathbf{H}$  yields the maximum-likelihood estimator,  $\hat{\mathbf{s}}_{\text{mle}}$ .<sup>3</sup>

$$\hat{\mathbf{s}}_{\text{mle}} = \mathbf{H} (\mathbf{H}^\dagger \mathbf{H})^{-1} \mathbf{H}^\dagger \mathbf{d}_a, \quad (9)$$

<sup>3</sup>The least-squares estimator of  $\mathbf{s}_a$  is equal to the maximum-likelihood estimator  $\hat{\mathbf{s}}_a$ ; however, the least-squares estimator of  $\phi$  is not the maximum-likelihood estimator  $\hat{\phi}$ , due to the arctangent nonlinearity. Here we choose to discuss maximum likelihood as opposed to least-squares estimators to better connect with the rich, classical theory of maximum likelihood estimation, as well as to unify discussion.

$$=\mathbf{P}_H \mathbf{d}_a, \quad (10)$$

where  $\mathbf{P}_H$  is the projection matrix projecting onto the space spanned by the columns of the model matrix  $\mathbf{H}$ . Thus, as specified in Eqn. (10), the maximum likelihood estimator of  $\mathbf{s}_a$  is the projection of  $\mathbf{d}_a$  onto the span of the columns of  $\mathbf{H}$ . The maximum likelihood estimator,  $\hat{\boldsymbol{\varphi}}$  of  $\boldsymbol{\varphi}$  is,

$$\hat{\boldsymbol{\varphi}}_{\text{mle}} = \tan^{-1} \left( \frac{\text{Im} \{ \hat{\mathbf{s}}_{\text{mle}} \}}{\text{Re} \{ \hat{\mathbf{s}}_{\text{mle}} \}} \right). \quad (11)$$

In Section (4),  $\hat{\boldsymbol{\varphi}}_{\text{mle}}$  will be shown to be equal to the DFT, Hilbert and complex Morlet phase estimators for suitable choices of the model matrix,  $\mathbf{H}$ .

### 3.3 bias $\{\hat{\mathbf{s}}_{\text{mle}}\}$

The bias of the signal estimator depends upon the range of the projection matrix,  $\mathbf{P}_H$ , and hence upon the range, or the span, of the columns of  $\mathbf{H}$ . That is, if  $\mathbf{s}_a$  belongs to the range of  $\mathbf{H}$ , then  $\mathbf{s}_a$  can be represented by a linear combination of the columns of  $\mathbf{H}$ , and  $\mathbf{P}_H \mathbf{s}_a = \mathbf{s}_a$ . Since the noise is zero mean, the projected noise is also zero mean and  $E \{ \mathbf{P}_H \mathbf{d}_a \} = \mathbf{s}_a$ . When  $\mathbf{P}_H \mathbf{s}_a \neq \mathbf{s}_a$ ,  $\hat{\mathbf{s}}_{\text{mle}}$  is biased. This mathematical description can be summarized as follows: when the model matrix  $\mathbf{H}$  is matched to the transformed signal, the signal estimator is unbiased.

### 3.4 The distribution of $(\mathbf{n}_p)_j$

The projected noise  $\mathbf{n}_p = \mathbf{P}_H \mathbf{n}_a$ , at time-index  $j$ , is shown in Appendix A to be distributed,<sup>4</sup>

$$(\mathbf{n}_p)_j \sim \mathcal{CN} \left( \mathbf{0}, \frac{p\sigma^2}{N} \right). \quad (12)$$

Here,  $p = \text{tr} \{ \mathbf{P}_H \}$ , where  $\text{tr} \{ \mathbf{X} \}$ , is the trace of the matrix  $\mathbf{X}$ . The projection decomposes the variance of the noise,  $\sigma^2$ , into fluctuations occurring within the span of  $\mathbf{H}$  and fluctuations occurring within the space orthogonal to the span of  $\mathbf{H}$ . Projected noise variance grows linearly with the dimension,  $p$ , that is, with the number of columns of  $\mathbf{H}$ , and hence with the number of model unknowns. The projection is a geometrical way to understand the effect of estimator complexity on inferred signal. As expected, projected noise variance decreases as one over the number of measurements,  $N$ , and increases with model complexity.

### 3.5 var $\{(\hat{\boldsymbol{\varphi}}_{\text{mle}})_j\}$

The situation describing the phase estimator performance at time-index  $j$  is shown in Figure (1) for the situation when  $\hat{\mathbf{s}}_{\text{mle}}$  is unbiased. In this figure, the circularly-symmetric Gaussian projected noise,  $\mathbf{n}_p$ , at time-index  $j$ , adds to the signal vector,  $\mathbf{s}$ , in the complex plane to form the estimated signal vector,  $\hat{\mathbf{s}}_{\text{mle}}$ .

<sup>4</sup>The complex normal density function appearing in Eqn. (12) is a result of taking the discrete Fourier transform of a signal plus noise model of the data. The development so far has assumed Gaussianity, but this is overly restrictive. While Eqn. (12) is exact when  $\mathbf{n}_a$  is assumed Gaussian, Eqn. (12) is also the consequence of a central limit theorem; it is valid in the limit as  $N$  tends to infinity, and is approximately valid for finite lengthed recordings. Convergence to Eqn. (12) can be quite rapid. See Section (6) for further discussion and references.

$$\widehat{\mathbf{s}}_{\text{mle}} = \mathbf{P}_H \mathbf{d}_a, \quad (13)$$

$$= \mathbf{s}_a + \mathbf{n}_p. \quad (14)$$

When the variance of the real and imaginary parts of the projected noise,  $\mathbf{n}_p$  is small relative to signal magnitude, the small angle approximation is valid,  $(\mathbf{s}_a)_j + [(\mathbf{n}_p)_j]_{\parallel} \approx (\mathbf{s}_a)_j$  and the phase estimator,  $(\widehat{\boldsymbol{\varphi}}_{\text{mle}})_j$  equals,

$$(\widehat{\boldsymbol{\varphi}}_{\text{mle}})_j = (\boldsymbol{\varphi}_s)_j + \tan^{-1} \left( \frac{[(\mathbf{n}_p)_j]_{\perp}}{(\mathbf{s}_a)_j + [(\mathbf{n}_p)_j]_{\parallel}} \right), \quad (15)$$

$$\approx (\boldsymbol{\varphi}_s)_j + \frac{[(\mathbf{n}_p)_j]_{\perp}}{|(\mathbf{s}_a)_j|}, \quad (16)$$

$$\stackrel{d}{=} (\boldsymbol{\varphi}_s)_j + \frac{\text{Re} \{(\mathbf{n}_p)_j\}}{|(\mathbf{s}_a)_j|}. \quad (17)$$

The equality in Eqn. (17) is equality in probability distribution. It results by noting that the projected noise,  $(\mathbf{n}_p)_j$ , is isotropic due to circular symmetry; any two orthogonal components of  $(\mathbf{n}_p)_j$  in the complex plane have identical distributions.

Here  $(\boldsymbol{\varphi}_s)_j$  is the actual, unknown phase at time index  $j$ . In this large instantaneous SNR regime the phase estimator is a linear function of a normally distributed random variable; it too is normally distributed.<sup>5</sup> In particular,

$$(\widehat{\boldsymbol{\varphi}}_{\text{mle}})_j \sim \mathcal{N} \left( (\boldsymbol{\varphi}_s)_j, \frac{p\sigma^2}{2N|(\mathbf{s}_a)_j|^2} \right). \quad (18)$$

This result is established formally in the situation of the DFT phase estimator in, again for example, Kay (1993). In this situation,  $p = 1$  and  $|(\mathbf{s}_a)_j| = a$  for  $j = 0, \dots, N - 1$ . Note,  $|(\mathbf{s}_a)_j|$  in the expression for the variance of the phase estimator,  $(\widehat{\boldsymbol{\varphi}}_{\text{mle}})_j$ . The accuracy of phase estimation depends upon amplitude.

While Eqn. (18) is accurate when the small angle approximation is valid, note that this geometry provides the basis to explore low signal to noise ratio regimes. One expects Taylor expansions to be useful in this regard.

### 3.6 Model Miss-Specification

Model miss-specification occurs when  $\widehat{\mathbf{s}}_{\text{mle}}$  is biased. In this situation, Fig. (1), is modified in the following way. Instead of  $(\mathbf{s}_a)_j$  adding to the noise vector  $(\mathbf{n}_p)_j$  to form the signal estimator,  $(\widehat{\mathbf{s}}_{\text{mle}})_j$ , the expected signal,  $(\mathbb{E} \{ \widehat{\mathbf{s}}_{\text{mle}} \})_j$  adds to the noise vector to produce the

<sup>5</sup>As in Eqn. (12), Eqn. (18) is exact when  $\mathbf{n}_a$  is normally distributed and the small-angle approximation is valid. Without the assumption that  $\mathbf{n}_a$  is normally distributed Eqn. (18) is a good approximation for large  $N$  due to a central limit theorem. See Section (6) for further discussion and references.

signal estimator,  $(\hat{\mathbf{s}}_{\text{mle}})_j$ . Due to this replacement,  $(\varphi_s)_j$  is replaced with some other angle, in general differing from the actual angle  $(\varphi_s)_j$ , and angle estimator bias may be introduced. Further, this bias reduces the signal amplitude within the column space of  $\mathbf{H}$ . This reduction increases  $\varphi_n$ , which increases phase angle estimator variance. That is, the signal  $\mathbf{s}_a$ , can be decomposed as a sum of the expected maximum likelihood estimator,  $E\{\hat{\mathbf{S}}_{\text{mle}}\}$  with an error vector,  $\mathbf{b}$ , describing estimator bias,

$$E\{\hat{\mathbf{s}}_{\text{mle}}\} = \mathbf{s}_a + \mathbf{b}. \quad (19)$$

Let  $E\{\hat{\mathbf{s}}_{\text{mle}}\} = \mathbf{s}_0$ . Then,

$$\mathbf{s}_a = \mathbf{s}_0 - \mathbf{b}, \quad (20)$$

$$= E\{\mathbf{P}_H \mathbf{d}_a\} - E\{\mathbf{P}_{H^\perp} \mathbf{d}_a\}, \quad (21)$$

where  $\mathbf{P}_{H^\perp}$  is the projection matrix that projects orthogonal to the column space of  $\mathbf{H}$ . Now, through orthogonality, the signal energy,  $E_s$ , can be decomposed into the sum of two non-negative terms,

$$\begin{aligned} E_s &= \mathbf{s}_a^\dagger \mathbf{s}_a, \\ &= \mathbf{s}_0^\dagger \mathbf{s}_0 + \mathbf{b}^\dagger \mathbf{b}. \end{aligned}$$

The signal energy unaccounted for by the maximum likelihood estimator of  $\mathbf{s}_a$  is due to that part of the signal,  $\mathbf{b}$ , lying outside the column space of  $\mathbf{H}$ . In essence,  $\mathbf{s}_a$  in Fig. (1) is replaced by  $E\{\hat{\mathbf{s}}_a\}$ , which on average has a magnitude less than that of  $\mathbf{s}_a$ . Thus, model miss-specification reduces the average estimated signal energy and tends to increase the variance of  $(\hat{\varphi}_{\text{mle}})_j$ .

### 3.7 Temporal Dependence of $\hat{\varphi}_{\text{mle}}$

Conditioned upon  $(\hat{\mathbf{s}}_{\text{mle}})_j$ ,  $(\hat{\varphi}_{\text{mle}})_j$  is independent of  $(\hat{\varphi}_{\text{mle}})_k$ , for  $k \neq j$ . Thus, it suffices to study the temporal dependence of  $\hat{\mathbf{s}}_{\text{mle}}$ . As this quantity is multivariate Gaussian, dependency is quantified via the second central moment:

$$\begin{aligned} \text{cov}\{\hat{\mathbf{s}}_{\text{mle}}\} &= E\{\mathbf{P}_H \mathbf{n}_a \mathbf{n}_a^\dagger \mathbf{P}_H^\dagger\}, \\ &= \sigma^2 \mathbf{P}_H. \end{aligned} \quad (22)$$

Note that Eqn. (22) is a matrix with complex values describing the temporal dependency both temporally within the real and imaginary parts, as well as between the real and the imaginary components temporally. When  $\mathbf{H}$  is sinusoidal, temporal dependence is periodic and long-range. When  $\mathbf{H}$  is narrowband, temporal dependence will be inversely proportional to the bandwidth. When  $\mathbf{H}$  has finite-support, then temporal dependence is limited to the domain of support.

## 4 Results

Given the general instantaneous phase estimator properties reviewed in Section (3), all that remains to establish the performance of the DFT, Hilbert and complex Morlet wavelet

estimators is to equate their computation with a specific projection,  $\mathbf{P}_H$ , and matching signal type. The results are summarized in Table (1).

#### 4.1 DFT Phase Estimator

The DFT model is that of a pure sinusoid. The analytic signal of a pure sinusoid at time index,  $j$ , with phase  $\varphi_0$  and frequency  $f_0$  is  $e^{i2\pi f_0 t_j}$ . Thus,

$$(\mathbf{H}^{(dft)})_j = e^{i2\pi f_0 t_j}. \quad (23)$$

#### 4.2 Hilbert Phase Estimator

The DFT model is a reduced rank version of the Hilbert model. The Hilbert transform phase estimator is computed in the following manner. First, the data is band-pass filtered about a frequency of interest to obtain,  $\mathbf{d}_{bp}$ . Then, the band-pass filtered data is Hilbert transformed to obtain  $\mathbf{d}_{hbp}$  and finally an analytic signal representation of the data is constructed,

$$\mathbf{d}_{abp} = \mathbf{d}_{bp} + i\mathbf{d}_{hbp}. \quad (24)$$

The instantaneous Hilbert phase estimator is then the complex angle of the analytic signal representation of the bandpass filtered data,

$$\widehat{\varphi}_h = \angle\{\mathbf{d}_{abp}\}. \quad (25)$$

In the Fourier domain the Hilbert transform introduces a  $90^\circ$  phase shift to the negative frequencies and a  $-90^\circ$  phase shift to the positive frequencies Kay (1993). Hence, the multiplication by  $i$  in Eqn. (24) introduces an overall phase shift of  $180^\circ$  for the data components at negative frequencies and a zero degree phase shift for the data components at positive frequencies. The result is that the complex valued quantity,  $\mathbf{d}_{abp}$  is zero for negative frequencies and twice the discrete Fourier transform of  $\mathbf{d}_{bp}$  for positive frequencies. Thus, the combined effect on  $\mathbf{d}_{abp}$  is, up to non-ideal band-pass filtering effects, the projection of the data  $\mathbf{d}$  onto the space spanned by the complex exponentials at positive frequencies belonging to the pass-band of the band-pass filter. That is,

$$\mathbf{H}^{(H)} = \begin{pmatrix} \vdots & & \vdots \\ e^{i2\pi f_{j_0} t_m} & \dots & e^{i2\pi f_{j_1} t_m} \\ \vdots & & \vdots \end{pmatrix} \quad (26)$$

where  $f_{j_0}$  is the frequency at which the pass-band of frequencies begins, and  $f_{j_1}$  is the frequency at which the pass-band ends. Here  $p$ , the dimension of the projection range, is equal to the number of Fourier components in the pass-band of the band-pass filter. Thus,

$$p = N \times BW \times \Delta, \quad (27)$$

where  $BW$  is the pass-band interval width in Hz.



### 4.3 Complex Morlet Phase Estimator

The complex Morlet phase estimator is typically obtained for a specific scale parameter,  $a$ , shift parameter  $b$ , bandwidth parameter  $f_b$  and center frequency,  $f_c$ , as the complex angle of the wavelet coefficient,  $\Psi_{a,b,f_c,f_b}$ ,

$$\Psi_{a,b,f_c,f_b}(d) = \int_{-\infty}^{\infty} d(t') \psi^*\left(\frac{t'-b}{a}\right) dt', \quad (28)$$

where  $d(t)$  is the ideal, continuous-time data and,

$$\psi^*\left(\frac{t-b}{a}\right) = e^{-\frac{1}{f_b} \frac{(t-b)^2}{a^2}} e^{i2\pi f_c \left(\frac{t-b}{a}\right)}. \quad (29)$$

In practice, the integral in Eqn. (28) is approximated from finite duration data consisting of samples of the continuous-time data  $d(t)$ . Modeling the finite, sampled data in vector form as a realization of the random vector  $\mathbf{d}$ , and approximating the integral in Eqn. (28) with a truncated Riemann sum, one obtains  $\tilde{\Psi}_{a,b,f_c,f_b}$ ,

$$\tilde{\Psi}_{a,b,f_c,f_b}(\mathbf{d}) = \Delta \sum_{n=0}^{N-1} (\mathbf{d})_n e^{-\frac{1}{f_b} \frac{(n\Delta-b)^2}{a^2}} e^{i2\pi f_c \left(\frac{n\Delta-b}{a}\right)}. \quad (30)$$

Taking the complex angle yields an estimator for the instantaneous phase evaluated at time,  $b$ . ie.

$$\widehat{\varphi}_b^{(cm)} = \angle \left\{ \tilde{\Psi}_{a,b,f_c,f_b}(\mathbf{d}) \right\}, \quad (31)$$

$$= \angle \{ \mathbf{P}_H \mathbf{d} \}, \quad (32)$$

where in this case,  $\mathbf{H}$  is a single column with  $m^{\text{th}}$  element equal to,

$$(\mathbf{H}^{(cm)})_m = e^{-\frac{1}{f_b} \frac{(m\Delta-b)^2}{a^2}} e^{i2\pi f_c \left(\frac{m\Delta-b}{a}\right)}. \quad (33)$$

Note that in this situation, the projection matrix depends upon the time-index,  $b$ , and further, the projection introduces real-valued scalings that do not affect the estimated phase.

## 5 Simulation

The three instantaneous phase estimators, for the  $j^{\text{th}}$  time step, are compared in simulation. Each simulation involves 100 iterations in which synthetic data are created, and the phase of a signal estimated using the three phase estimators. These phase estimates are then used, in a kernel-smoothing procedure, to estimate the probability density functions for each of the three phase estimators. The estimated probability density functions, along with a representative time-series, are shown for two signals measured with additive white Gaussian noise (AWGN). In the first simulation the signal is a sinusoid and the resulting estimator probability density functions are shown in Fig. (2), top. In this case, the DFT phase estimator matches the signal and has the minimum column space dimension,  $p$ , of one. This estimator out-performs the other two phase estimators. Here the Hilbert estimator is using a

narrow band of frequencies centred upon the actual sinusoid frequency; it suffers no bias, but has a slightly increased column space dimension,  $p$ , and performs nearly as well as the DFT based phase estimator. The complex Morlet phase estimator possesses a model column space dimension of 1; but, does not match the sinusoid and hence has a biased signal estimate resulting in increased phase estimator variance due to signal mismatch. Next, this simulation is repeated with differing sinusoid amplitudes but with the standard deviation of the additive noise held constant.<sup>6</sup> The standard deviation of the resulting phase estimators is plotted as a function of sinusoid amplitude in the bottom of Fig. (2), along with the theoretical standard deviation obtained from Eqn. (18). The actual phase estimator standard deviations are larger than the theoretical standard deviation with a discrepancy that reduces with signal amplitude. This discrepancy is due to the fact that the small angle approximation, used to obtain an expression for  $\varphi_n$  in Fig. (1), underestimates the actual value.

In the second simulation the signal is a broad-band pulse measured in AWGN. The resulting phase estimator densities are shown in Fig. (3). Now the performance ordering of the phase estimators is reversed. The Morlet estimator matches the signal, has a minimum model column space dimension,  $p$ , of 1, and outperforms the other phase estimators. The Hilbert estimator has less model mismatch than the DFT estimator, but pays for this better signal match in terms of column space dimension and only slightly out-performs the DFT estimator, which is mismatched to the broad-band pulse. Next, this simulation is repeated with differing pulse amplitudes but with the standard deviation of the additive noise held constant.<sup>1</sup> The standard deviation of the resulting phase estimators is plotted as a function of pulse amplitude in the bottom of Fig. (3), along with the theoretical standard deviation obtained from Eqn. (18). As in the first simulation, the actual phase estimator standard deviations are always larger than the theoretical standard deviation with a discrepancy that reduces with signal amplitude. Again, this discrepancy is due to the fact that the small angle approximation, used to obtain an expression for  $\varphi_n$  in Fig. (1), underestimates the actual value.

## 6 Discussion

The variance, bias and asymptotic distribution of the DFT, Hilbert transform based, and complex Morlet instantaneous phase estimators are established. Each of these estimators is shown to be a maximum likelihood estimator based upon a different signal model. Both the bias and variance are shown, in simulation and in theory, to depend upon the extent to which signal present in the data is matched to the phase estimator. The variance of each phase estimator is also shown to depend upon the complexity of the signal model to which it is associated; more complicated signal models yield models with a greater number of unknowns, and thus increased uncertainty, as described by the projection dimension,  $p$ ; which is equal to the number of model unknowns. As represented in the right-most column in Table (1), estimator quality depends upon the quality of the information which its associated model embodies: if specific waveform shapes are known they can be used to form explicit phase estimators with a  $p$  of 1 and minimum variance, provided the actual signal matches the specified waveform. When model miss-specification occurs, a penalty is paid both in terms of bias and variance. This latter effect, due to a reduction in signal amplitude within the column space of the model matrix  $\mathbf{H}$ , magnifies the effect of noise. On the other hand, increasing the model uncertainty, as in the Hilbert type estimator decreases susceptibility to model miss-specification; but, at the price of an increased  $p$ , and increased estimator variance.

---

<sup>6</sup>Because noise variance and squared signal amplitude affect Eqn. (18) as a ratio, it suffices to vary either one while holding the other constant.

In this work, the additive noise vector is assumed to be multivariate Gaussian and the elements of this vector are assumed to be uncorrelated; consistent with temporal independence. The assumption of Gaussianity is overly-restrictive since in the narrowband case (the case of interest),  $\mathbf{n}_a$ , is dominated by a linear combination of a small number of DFT coefficients which are each asymptotically circularly-symmetric Gaussian without an a priori assumption of normality on the pre-Fourier transformed noise, see for example Walker (2000) and (Priestly, 1981, p. 466). The assumption of temporally uncorrelated noise, on the other hand, is important to the current results. If temporal correlation exists, then the real and imaginary parts of  $\mathbf{n}_a$  for a given time index are correlated. When this occurs, the additive randomly oriented noise vector in Fig. (1), is no longer randomly oriented. One expects in this situation that estimators preferentially weighting time-indices where the additive noise vector is expected to lie parallel to the signal will possess advantageous sampling properties.

The classical analysis presented in this paper provides an explicit relation governing the accuracy of commonly employed estimates of phase computed from noisy data. This relation, specifying both the effect of signal amplitude and the role of a priori signal knowledge on phase estimate accuracy, provides neuroscientists with increased awareness when conducting phase based studies, as well as a basis for further understanding derivative quantities, such as measures of phase synchrony.

## Acknowledgments

This work was partially supported by grants from the National Science Foundation [IIS - 0643995] and [DMS - 1042134] and the National Institute of Neurological Disorders and Stroke [R01NS073118] and [R01NS072023]. MAK holds a Career Award at the Scientific Interface from the Burroughs Wellcome Fund.

## References

- Aviyente S, Bernat EM, Evans WS, Sponheim SR. A phase synchrony measure for quantifying dynamic functional integration in the brain. *Human Brain Mapping*. 2011; 32 (1):80–93. [PubMed: 20336687]
- Aviyente S, Mutlu A. A time-frequency-based approach to phase and phase synchrony estimation. *Signal Processing, IEEE Transactions on*. Jul; 2011 59 (7):3086–3098.
- Bruns A. Fourier-, hilbert-and wavelet-based signal analysis: are they really different approaches? *Journal of neuroscience methods*. 2004; 137 (2):321–332. [PubMed: 15262077]
- Buzsáki, G. *Rhythms of the Brain*. Oxford University Press; 2006.
- Buzsáki G, Buhl DL, Harris KD, Csicsvari J, Czeh B, Morozov A. Hippocampal network patterns of activity in a mouse. *Neuroscience*. 2003; 116:201–211. [PubMed: 12535953]
- Buzsáki G, Draguhn A. Neuronal oscillations in cortical networks. *Science*. Jun; 2004 304 (5679): 1926–9. [PubMed: 15218136]
- Canolty RT, Edwards E, Dalal SS, Soltani M, Nagarajan SS, Kirsch HE, Berger MS, Barbaro NM, Knight RT. High gamma power is phase-locked to theta oscillations in human neocortex. *Science*. Sep; 2006 313 (5793):1626–8. [PubMed: 16973878]
- Canolty RT, Knight RT. The functional role of cross-frequency coupling. *Trends Cogn Sci (Regul Ed)*. Nov; 2010 14 (11):506–15. [PubMed: 20932795]
- Cohen MX. Assessing transient cross-frequency coupling in eeg data. *J Neurosci Methods*. Mar; 2008 168 (2):494–9. [PubMed: 18061683]
- Gray CM, König P, Engel AK, Singer W. Oscillatory responses in cat visual cortex exhibit inter-columnar synchronization which reflects global stimulus properties. *Nature*. Mar; 1989 338 (6213):334–7. [PubMed: 2922061]
- Guevara R, Velazquez J, Nenadovic V, Wennberg R, Senjanovi G, Dominguez L. Phase synchronization measurements using electroencephalographic recordings. *Neuroinformatics*. 2005; 3(4):301–313.10.1385/NI:3:4:301 [PubMed: 16284413]

- Harris KD, Csicsvari J, Hirase H, Dragoi G, Buzsáki G. Organization of cell assemblies in the hippocampus. *Nature*. Jul; 2003 424 (6948):552–6. [PubMed: 12891358]
- Hentschke H, Perkins MG, Pearce RA, Banks MI. Muscarinic blockade weakens interaction of gamma with theta rhythms in mouse hippocampus. *European Journal of Neuroscience*. 2007; 26:1642–1656. [PubMed: 17880398]
- Hogg, RV.; McKean, JW.; Craig, AT. *Introduction to Mathematical Statistics*. 6. Pearson Prentice Hall; 2005.
- Hurtado JM, Rubchinsky LL, Sigvardt KA. Statistical method for detection of phase-locking episodes in neural oscillations. *Journal of Neurophysiology*. 2003; 91:1883–1898. [PubMed: 15010498]
- Kay, SM. *Fundamentals of Statistical Signal Processing: Estimation Theory*. Prentice Hall PTR; Upper Saddle River, New Jersey: 1993.
- Lakatos P, Karmos G, Mehta AD, Ulbert I, Schroeder CE. Entrainment of neuronal oscillations as a mechanism of attentional selection. *Science*. Apr; 2008 320 (5872):110–3. [PubMed: 18388295]
- Lakatos P, Shah AS, Knuth KH, Ulbert I, Karmos G, Schroeder CE. An oscillatory hierarchy controlling neuronal excitability and stimulus processing in the auditory cortex. *Journal of Neurophysiology*. 2005; 94:1904–1911. [PubMed: 15901760]
- Palva JM, Palva S, Kaila K. Phase synchrony among neuronal oscillations in the human cortex. *J Neurosci*. Apr; 2005 25 (15):3962–72. [PubMed: 15829648]
- Pereda E, Quiroga R, Bhattacharya J. Nonlinear multivariate analysis of neurophysiological signals. *Progress in Neurobiology*. Jan.2005
- Priestly, MB. *Spectral Analysis and Time Series*. Elsevier Academic Press; 1981.
- Quyen MLV. The brainweb of cross-scale interactions. *New Ideas in Psychology*. 2011; 29 (2):57–63.
- Quyen MLV, Bragin A. Analysis of dynamic brain oscillations: methodological advances. *Trends Neurosci*. Jul; 2007 30 (7):365–73. [PubMed: 17559951]
- Quyen MLV, Foucher J, Lachaux JP, Rodriguez E, Lutz A, Martinerie J, Varela FJ. Comparison of hilbert transform and wavelet methods for the analysis of neuronal synchrony. *Journal of Neuroscience Methods*. 2001; 111 (2):83–98. [PubMed: 11595276]
- Rizzuto DS, Madsen JR, Bromfield E, Schulze-Bonhage A, Kahana MJ. Human neocortical oscillations exhibit theta phase differences between encoding and retrieval. *NeuroImage*. 2006; 31:1352–1358. [PubMed: 16542856]
- Rizzuto DS, Madsen JR, Bromfield EB, Schulze-Bonhage A, Seelig D, Aschenbrenner-Scheibe R, Kahana MJ. Reset of human neocortical oscillations during a working memory task. *Proceedings of the National Academy of Sciences, Neuroscience*. 2003; 100 (13):7931–7936.
- Rutishauser U, Ross IB, Mamelak AN, Schuman EM. Human memory strength is predicted by theta-frequency phase-locking of single neurons. *Nature*. Apr; 2010 464 (7290):903–7. [PubMed: 20336071]
- Scharf, LL. *Statistical Signal Processing: detection, estimation, and time series analysis*. Addison-Wesley Publishing Company, Inc; 1991.
- Schroeder C, Lakatos P. Low-frequency neuronal oscillations as instruments of sensory selection. *Trends Neurosci*. Nov.2008
- Sirota A, Montgomery S, Fujisawa S, Isomura Y, Zugaro M, Buzsaki G. Entrainment of neocortical neurons and gamma oscillations by the hippocampal theta rhythm. *Neuron*. 2008; 60:683–697. [PubMed: 19038224]
- Tort ABL, Kramer MA, Thorn C, Gibson DJ, Kubota Y, Graybiel AM, Kopell NJ. Dynamic cross-frequency couplings of local field potential oscillations in rat striatum and hippocampus during performance of a t-maze task. *PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA*. Dec; 2008 105 (51):20517–22. [PubMed: 19074268]
- Tort ABL, Kramer MA, Thorn C, Gibson DJ, Kubota Y, Graybiel AM, Kopell NJ. Dynamic cross-frequency couplings of local field potential oscillations in rat striatum and hippocampus during performance of a t-maze task. *Proceedings of the National Academy of Sciences of the United States of America*. 2008; 105(51)
- Varela F, Lachaux JP, Rodriguez E, Martinerie J. The brainweb: phase synchronization and large-scale integration. *Nat Rev Neurosci*. Apr; 2001 2 (4):229–39. [PubMed: 11283746]

- Walker A. Some results concerning the asymptotic distribution of sample fourier transforms and periodograms for a discrete-time stationary process with a continuous spectrum. *Journal of Time Series Analysis*. 2000; 21 (1):95–109.
- Womelsdorf T, Schoffelen JM, Oostenveld R, Singer W, Desimone R, Engel AK, Fries P. Modulation of neuronal interactions through neuronal synchronization. *Science*. Jun; 2007 316 (5831):1609–12. [PubMed: 17569862]
- Young CK, Eggermont JJ. Coupling of mesoscopic brain oscillations: recent advances in analytical and theoretical perspectives. *Prog Neurobiol*. Sep; 2009 89 (1):61–78. [PubMed: 19549556]

## A The Distribution of $(\mathbf{n}_a)_j$

The additive, complex valued noise vector,  $\mathbf{n}_a$  in Eqn. (2) is equal to the analytic signal of the additive noise term,  $\mathbf{n}$ , in the data model,  $\mathbf{d} = \mathbf{s} + \mathbf{n}$ , where, for the sake of generality,  $\mathbf{E}\{\mathbf{n}\} = \mathbf{0}$ , and  $\mathbf{E}\{\mathbf{nn}^T\} = \mathbf{R}$  is Toeplitz. In this case, the covariance between elements of the noise vector,  $\mathbf{n}$  depends upon the absolute difference between element indices, consistent with a weak-sense stationary assumption. The analytic signal,  $\mathbf{n}_a$  of the noise,  $\mathbf{n}$ , is distributed mean-zero circularly-symmetric normal. This is shown by examining the moments of the real and imaginary parts of  $\mathbf{n}_a$ , making use of the Fourier decomposition,

$$\mathbf{n}_a = \sum_{k \in F_+} \tilde{\mathbf{n}}_k e^{i2\pi \frac{kt}{N}}. \quad (34)$$

Let  $\tilde{n}_k^{(r)} = \text{Re}\{\tilde{\mathbf{n}}_k\}$ , and let  $\tilde{n}_k^{(i)} = \text{Im}\{\tilde{\mathbf{n}}_k\}$ , denote the real and imaginary parts of the  $k^{\text{th}}$  element of  $\tilde{\mathbf{n}}$ . Then,

$$\begin{aligned} \mathbf{E}\{\text{Re}\{(\mathbf{n}_a)_t\}\text{Im}\{(\mathbf{n}_a)_{t'}\}\} &= \frac{1}{N} \mathbf{E}\left\{ \sum_{k, k' \in F_+} \text{Re}\{\tilde{\mathbf{n}}_k e^{i2\pi \frac{kt}{N}}\} \text{Im}\{\tilde{\mathbf{n}}_{k'} e^{i2\pi \frac{k't'}{N}}\} \right\}, \\ &= \frac{1}{N} \mathbf{E}\left\{ \sum_{k, k' \in F_+} \left[ n_k^{(r)} \cos 2\pi \frac{kt}{N} - n_k^{(i)} \sin 2\pi \frac{kt}{N} \right] \times \left[ n_{k'}^{(r)} \sin 2\pi \frac{k't'}{N} + n_{k'}^{(i)} \cos 2\pi \frac{k't'}{N} \right] \right\}, \quad (35) \\ &= \frac{1}{N} \sum_{k, k' \in F_+} \mathbf{E}\left\{ n_k^{(r)} n_{k'}^{(r)} \right\} \cos 2\pi \frac{kt}{N} \sin 2\pi \frac{k't'}{N} - \mathbf{E}\left\{ n_k^{(i)} n_{k'}^{(i)} \right\} \cos 2\pi \frac{kt}{N} \sin 2\pi \frac{k't'}{N}. \end{aligned}$$

Here  $F_+ = \{k: 0 < \frac{k}{N} \leq \frac{1}{2}\}$  is the set of positive frequencies,  $t, t'$  are integer. Because  $\mathbf{E}\left\{ n_k^{(r)} n_{k'}^{(i)} \right\}$  is equal to zero due to circular symmetry for  $k, k' \in F_+$ , terms involving  $\mathbf{E}\left\{ n_k^{(r)} n_{k'}^{(i)} \right\}$  do not contribute to Eqn. (35), and for clarity have been left out. Continuing,

$$\mathbf{E}\left\{ n_k^{(r)} n_{k'}^{(r)} \right\} = \begin{cases} \frac{S_n(f_k)}{2}, & k = k', k \in F_+ \\ 0, & k \neq k' \end{cases}, \quad (36)$$

where  $S_n(f_k)$  is the spectrum of the additive noise,  $\mathbf{n}$ , evaluated at frequency,  $f_k = \frac{k}{N}$ . When  $\mathbf{n}$  has uncorrelated elements,  $S_n(f_k)$  is equal to  $\sigma^2$  for  $k \in F_+$ . Note that without loss in generality, in this development the sample interval is taken to be 1. The spectrum requires scaling for the situation where the sample interval differs from 1 (so do the frequencies, but this scaling is cancelled by the time terms in the complex exponential).

$$\begin{aligned} E \{ \text{Re} \{ (\mathbf{n}_a)_t \} \text{Im} \{ (\mathbf{n}_a)_{t'} \} \} &= \frac{\sigma^2}{N} \sum_{k \in F_+} \cos 2\pi \frac{kt}{N} \sin 2\pi \frac{kt'}{N} - \cos 2\pi \frac{kt'}{N} \sin 2\pi \frac{kt}{N}, \\ &= 0 \quad (\text{by orthogonality}), \end{aligned} \quad (37)$$

for  $t, t' = 0, 1, \dots, N-1$ . A similar argument yields,

$$E \{ \text{Re} \{ (\mathbf{n}_a)_t \} \text{Re} \{ (\mathbf{n}_a)_{t'} \} \} = \frac{1}{2N} \sum_{k \in F_+} S(f_k) \cos 2\pi \frac{kt}{N} \cos 2\pi \frac{kt'}{N} + S(f_k) \sin 2\pi \frac{kt}{N} \sin 2\pi \frac{kt'}{N}. \quad (38)$$

When the spectrum,  $S(f_k)$  is white,

$$\begin{aligned} E \{ \text{Re} \{ (\mathbf{n}_a)_t \} \text{Re} \{ (\mathbf{n}_a)_{t'} \} \} &= \begin{cases} \frac{1}{2N} \sum_{k \in F_+} S_n(f_k) & t=t' \\ 0 & t \neq t' \end{cases}, \\ &= \begin{cases} \frac{\sigma^2}{2} & t=t' \\ 0 & t \neq t' \end{cases}. \end{aligned} \quad (39)$$

Thus, when the additive noise vector,  $\mathbf{n}_a$  in Eqn. (2) has independent elements, each element is circularly-symmetric Gaussian with real and imaginary components that are independent. Each of these components possess a variance equal to  $\frac{\sigma^2}{2}$ . With an application of Eqn. (37) and Eqn. (39), one obtains,

$$E \{ \mathbf{n}_a \mathbf{n}_a^\dagger \} = \sigma^2 \mathbf{I}, \quad (40)$$

where  $\mathbf{I}$  is the identity matrix. The variance of the projected noise,  $\mathbf{n}_p = \mathbf{P}_H \mathbf{n}_a$ , is computed as follows:

$$\begin{aligned} E \left\{ \sum_j |(\mathbf{n}_p)_j|^2 \right\} &= \sum_j E \{ |(\mathbf{n}_p)_j|^2 \}, \\ N E \{ |(\mathbf{n}_p)_j|^2 \} &= E \{ \mathbf{n}_p^\dagger \mathbf{n}_p \}, \\ N E \{ |(\mathbf{n}_p)_j|^2 \} &= E \left\{ \text{trace} \left\{ \mathbf{P}_H \mathbf{n}_a \mathbf{n}_a^\dagger \mathbf{P}_H^\dagger \right\} \right\}, \\ E \{ |(\mathbf{n}_p)_j|^2 \} &= \frac{\sigma^2 \text{trace} \{ \mathbf{P}_H \}}{N}, \\ E \{ |(\mathbf{n}_p)_j|^2 \} &= \frac{p \sigma^2}{N}. \end{aligned}$$

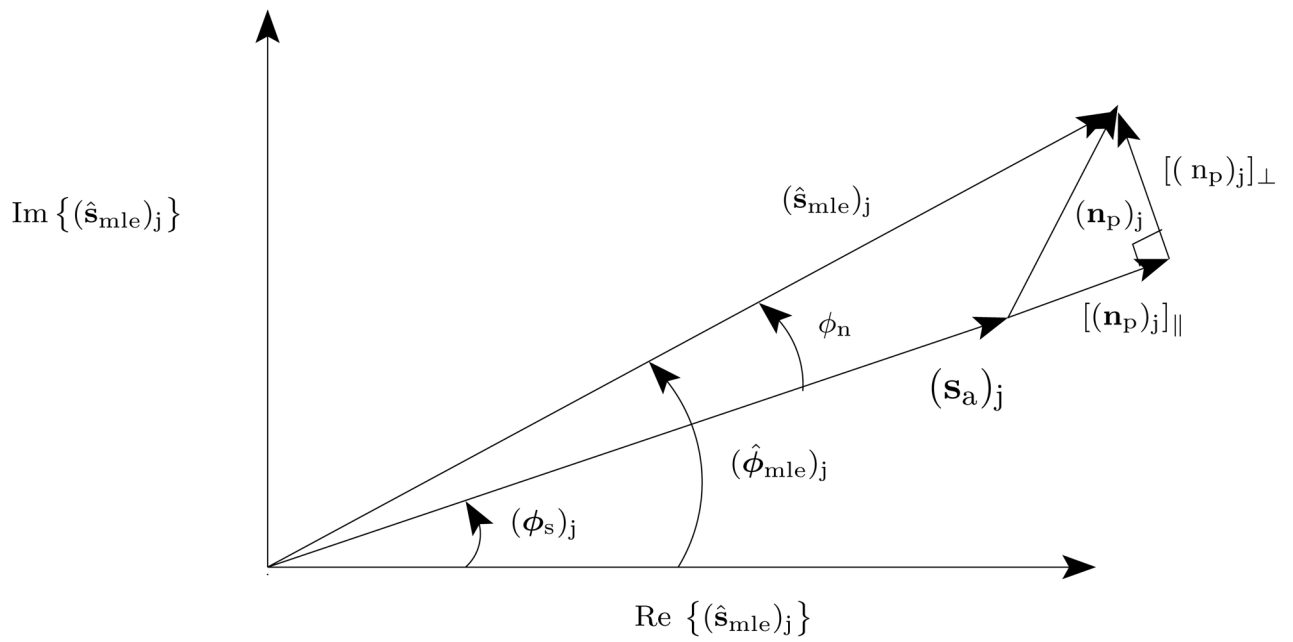
Thus,

$$\begin{aligned} E \{ |(\mathbf{n}_p)_j|^2 \} &= E \{ \text{Re} \{ (\mathbf{n}_p)_j \} \}^2 + E \{ \text{Im} \{ (\mathbf{n}_p)_j \} \}^2, \\ &= \frac{p \sigma^2}{N}, \end{aligned} \quad (41)$$

implying,

$$\begin{aligned} \text{var} \{ \text{Re} \{ (\mathbf{n}_p)_j \} \} &= \frac{p \sigma^2}{2N}, \\ &= \text{var} \{ \text{Im} \{ (\mathbf{n}_p)_j \} \}. \end{aligned} \quad (42)$$

Thus, Eqn. (12) is established.

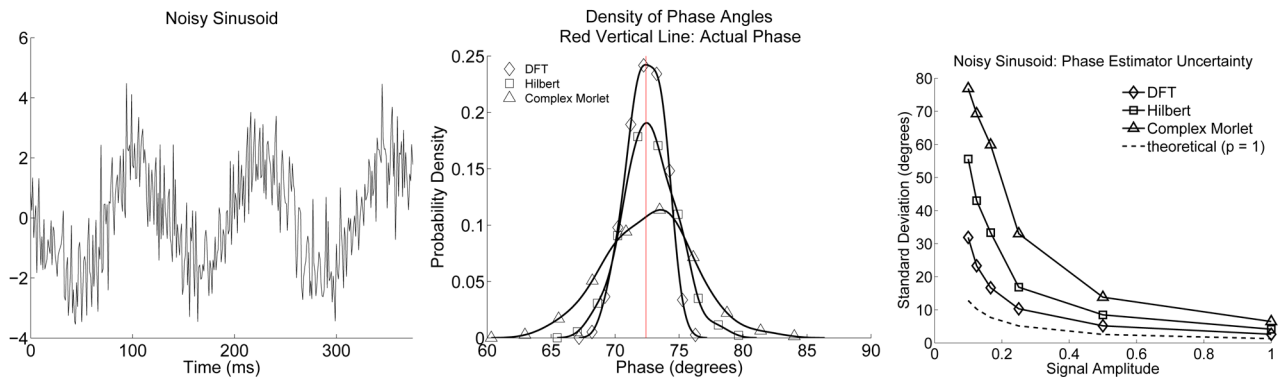


**Figure 1.**

The  $j^{\text{th}}$  element of the MLE signal estimate,  $(\hat{\mathbf{s}}_{\text{mle}})_j$ , is decomposed in the complex plane into a sum of the random noise vector  $(\mathbf{n}_p)_j = (\mathbf{P}_H \mathbf{n}_a)_j$  with the actual complex signal vector,  $(\mathbf{s}_a)_j$ , directed along the actual instantaneous phase angle  $(\phi_s)_j$ . The noise vector is randomly oriented, and when the variance,  $\text{var} \{(\hat{\mathbf{s}}_{\text{mle}})_j\}$  is small relative to  $|(\mathbf{s}_a)_j|^2$ ,  $\phi_n$  is approximately

equal to  $\frac{[(\mathbf{n}_p)_j]_{\perp}}{|(\mathbf{s}_a)_j|}$ .

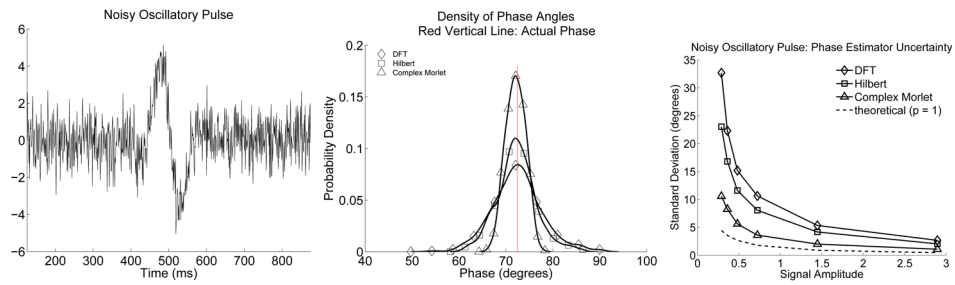




**Figure 2.**

Left plot: A single realization of a sinusoid measured in AWGN. The signal to noise ratio (SNR) is 2, where here SNR is the average signal energy divided by average noise energy or, equivalently, divide by the noise variance. Middle plot: Probability density estimates at a specific time-index for the three instantaneous phase estimators, when used to estimate the phase of the sinusoid. Right plot: Phase estimator standard deviation plotted as a function of amplitude. The estimator employing the DFT has the least variance of the three phase estimators. The complex Morlet phase estimator uses a miss-specified model for this situation, and suffers increased variance and increased bias relative to the variance and bias of the other two phase estimators. The results are qualitatively consistent with theoretical developments, but note that the theoretical standard deviation (bottom plot) lower-bounds the realized standard deviation. This discrepancy decreases with sinusoid amplitude and

results from the small angle approximation used to represent  $\varphi_n$  as  $\frac{[(np)_j]_n}{|(sa)_j|}$ .



**Figure 3.**

Left plot: A single realization of a Gaussian modulated pulse measured in AWGN. The signal to noise ratio is 0.84. Middle plot: Probability density estimates at a specific time-index for the three instantaneous phase estimators, when used to estimate the phase of a sinusoidally modulated Gaussian pulse. Right plot: Phase estimator standard deviation plotted as a function of pulse amplitude. The phase estimator based upon the DFT has the largest variance of the three phase estimators, consistent with the effect of model miss-specification. The complex Morlet phase estimator out-performs the other two estimators, owing to the fact that it uses a projection operator with a range of minimal dimension, while projecting the signal onto itself. The Hilbert phase estimator is not optimal, but rather suffers some model miss-specification and some increased variance due to the dimensionality of its associated model matrix  $\mathbf{H}^{(H)}$ . The results are qualitatively consistent with theoretical developments, but note that the theoretical standard deviation (bottom plot) lower-bounds the realized standard deviation. This discrepancy decreases with pulse amplitude and results

from the small angle approximation used to represent  $\varphi_n$  as  $\frac{(\mathbf{np})_j}{|(\mathbf{sa})_j|}$ .

**Table 1**

## Phase Estimator Summary

Estimator	Matching Signal	Model Matrix	Projected Dimension, p	Prior Information
DFT	$\alpha e^{i2\pi f_0 t}$	Eqn. (23)	<b>1</b>	Sinusoid at $f_0$
Hilbert	$\sum_{k=1}^{N \times BW \times \Delta} \alpha_k e^{i2\pi f_k t}$	Eqn. (26)	$N \times BW \times \Delta$	Frequency interval centered upon $f_0$ .
Complex Morlet	$\alpha e^{-\frac{1}{T_b} \frac{(t-b)^2}{a^2}} e^{i2\pi f_0 t}$	Eqn. (33)	<b>1</b>	Gaussian envelope at time, b, modulated to frequency $f_0$ .