Joint ML/MAP Estimation of the Frequency and Phase of a Single Sinusoid With Wiener Carrier Phase Noise

Qian Wang, *Member, IEEE*, Zhi Quan, *Senior Member, IEEE*, Suzhi Bi, *Senior Member, IEEE*, and Pooi-Yuen Kam, *Life Fellow, IEEE*

***Abstract*—We address here the issue of jointly estimating the angle parameters of a single sinusoid with Wiener carrier phase noise and observed in additive, white, Gaussian noise (AWGN). We develop the theoretical foundation for time-domain, phase-based, joint maximum likelihood (ML) estimation of the unknown carrier frequency and the initial carrier phase, with simultaneous maximum a posteriori probability (MAP) estimation of the time-varying carrier phase noise. The derivation is based on the amplitude and phase-form of the noisy received signal model together with the use of the best, linearized, additive observation phase noise model due to AWGN. Our newly derived estimators are closed-form expressions, consisting of both the phase and the magnitude of all the received signal samples. More importantly, they all have a low-complexity, sample-by-sample iterative processing structure, which can be implemented iteratively in real-time. As a basis for comparison, the Cramer-Rao lower bound (CRLB) for the ML estimators and the Bayesian CRLB (BCRLB) for the MAP estimator are derived in the presence of carrier phase noise, and the results simply depend on the signal-to-noise ratio (SNR), the observation length and the phase noise variance. It is theoretically shown that the estimates obtained are unbiased, and the mean-square error (MSE) of the estimators attain the CRLB/BCRLB at high SNR. The MSE performance as a function of the SNR, the observation length and the phase noise variance is verified using Monte Carlo simulation, which shows a remarkable improvement in estimation accuracy in large phase noise.**

***Index Terms*—MAP/ML estimation, carrier phase noise, Wiener process, phase-based time-domain frequency/phase estimation, (Bayesian) Cramer-Rao lower bound, MMSE estimation, weighted phase averager.**

I. Introduction

ESTIMATING the parameters, e.g., the carrier frequency and phase of a sine wave with noise is a classic and important issue in communications [1]–[3], biomedical engineering [4]–[6], radar/sonar applications [7], [8], and other areas of signal processing, e.g., power-quality monitoring in the power grid [9], [10]. All these physical applications can resort to almost the same signal model for theoretical performance analysis, where a single sinusoidal carrier wave is used.

Based on the general discrete-time single-sinusoid signal model, a variety of frequency estimation algorithms were developed in additive, white, Gaussian noise (AWGN). Due to the nonlinear dependence of the observed data on angular parameters, most estimators cannot be derived in explicit closed form, and can hardly get a good trade-off between estimation accuracy and computational complexity. The most common approach so far is to model the frequency and phase as unknown and nonrandom parameters, and apply the theory of maximum likelihood (ML) estimation [11]–[14]. One well-known solution is the Fourier-transform-based frequency-domain method, i.e., to locate the peak of the periodogram through a 1-D search, which is computationally intensive [11], [12]. The other popular alternative is called phase-based, time-domain estimation, using the received signal phases as the observation data samples to be fed into the estimator, where the received phase is expressed as the sum of the transmitted signal phase and an additive observation phase noise (AOPN) due to the AWGN [13], [14]. Even though ML estimation has been implemented over the past several decades, the actual structure of the time-domain ML frequency/phase estimator in AWGN is derived explicitly only recently in [3], which makes full use of both the magnitude and the phase of the observations to achieve the best possible estimation accuracy. However, all these prior art on joint frequency and phase estimation is only for the pure AWGN channel, and has not taken into account the detrimental impact of oscillator phase noise.

As is well known, all natural and man-made oscillators (whether optical, electronic, acoustic, atomic, or any other) exhibit phase and frequency instabilities collectively known as phase noise. Oscillator phase noise is one of the fundamental performance impairments in modern communication, radar, spectroscopic, and metrological systems [15], [16]. To be specific, in order to achieve higher spectral efficiency, ultra-high-speed communications nowadays are moving to higher spectral bands. The system performance is often limited by the phase noise introduced by local oscillators used for up/down-conversion, and the higher the carrier frequency used the greater the level of carrier phase noise encountered [17], [18]. Besides, with the advent of coherent optics, the role of laser oscillator phase

Manuscript received May 24, 2021; revised November 18, 2021; accepted December 19, 2021. Date of publication December 23, 2021; date of current version January 18, 2022. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Yik-chung Wu. This work was supported in part by the National Key R&D Program of China under Grant 2019YFB1803305, and in part by the National Natural Science Foundation of China under Grant 61571316. *(Corresponding author: Zhi Quan.)*

Qian Wang, Zhi Quan, and Suzhi Bi are with the College of Electronics and Information Engineering, Shenzhen University, Shenzhen 518060, China (e-mail: wang.qian@szu.edu.cn; zquan@szu.edu.cn; bsz@szu.edu.cn).

Pooi-Yuen Kam is with the School of Science and Engineering, Chinese University of Hong Kong, Shenzhen 518172, China, and also with the Center for Advanced Microelectronic Devices, NUS Research Institute, Suzhou 215123, China (e-mail: pykam@cuhk.edu.cn).

Digital Object Identifier 10.1109/TSP.2021.3137966

noise is becoming well recognized also in the context of optical transmission, and many digital signal processing algorithms are used to compensate for the consequential damage [2], [19]. In radar applications for real-time ranging and imaging, such as frequency modulated continuous wave (FMCW) radar and synthetic aperture radar (SAR) systems, the uncompensated phase noise may cause a time-variant shift, spurious sidelobes, and a broadening of the impulse response, as well as a low-frequency phase modulation of the radar signal, which severely degrades the detection and tracking performance [20], [21]. Moreover, for non-contact biomedical (cardiopulmonary) monitoring, where continuous-wave radar technology is employed to estimate the heart rate frequency, the residual phase noise from an on-chip oscillator has a significant effect on the detection sensitivity of the small phase variations created by heart and respiration motion [22], [23]. Therefore, phase noise is an inevitable factor that should be taken into account in parameter estimation and performance analysis in practical applications.

Considering the time-varying phase noise, the non-stationary Wiener process model is commonly used to characterize the jitter behaviour of practical oscillators in wireless, optical and radar applications [18], [20], [24]. However, in order to simplify theoretical analysis, most algorithms are designed based on the assumption of a quasi-static phase and do not consider the time-varying nature of the carrier phase, and thus results in a significant deterioration of the estimation performance. Moreover, there is only limited published work that uses the Wiener phase noise model directly in joint frequency/phase estimation of a sine wave in noise. For example, [25] analyzed the Cramer-Rao lower bound (CRLB) for frequency estimation in the presence of phase noise, where only an asymptotic approximation of the bound is derived by using a Monte Carlo average, and the estimator designed therein is quite complicated for a receiver implementation. Reference [26] considered the linear minimum-mean-square error (LMMSE) recursive estimation of the Wiener phase noise, and [27] simulated the effect of the Wiener phase noise on an iterative decision-directed frequency and phase noise estimation technique, using the generalized expectation-maximization algorithm on a truncated basis expansion of the phase. Our carrier phase recovery work in [28] for linear, suppressed-carrier digital modulations also modelled the unknown carrier phase as a constant, nonrandom parameter, and applied the ML estimation theory. The estimator can be applied only to the case of very small phase noise in practice. Thus, in [29], we further developed a low-complexity, complex-weighted decision-aided, adaptive complex phasor estimator, which automatically adapts its effective filter length according to the signal-to-noise ratio (SNR), laser-linewidth-symbol-duration product and modulation format. In summary, no systematic approach for deriving the joint ML frequency and phase estimator in Wiener phase noise has been developed, and there is no direct method for analyzing the performance of estimators in the presence of such phase noise in the literature.

Note that using the frequency-domain ML approach first proposed in [12] is not suitable for estimation in the presence of carrier phase noise. The frequency-domain approach normally takes the discrete Fourier transform (DFT) of the received noisy samples and locates the peak of the periodogram. Due to the carrier phase noise, the peak is broadened and the peak location is more prone to errors. Therefore, these methods are very sensitive to noise and cannot be refined in the presence of carrier phase noise. Moreover, the numerical complexity of the DFT-based method is more than ten times larger than the complexity of the time-domain approach using phase angle calculations of the sample sequence [30]. Therefore, the time-domain approach is advantageous and more promising when there is carrier phase noise.

This paper thus derives the explicit structures of phase based, joint ML/MAP estimation of unknown single-sinusoid angle parameters in time-varying phase noise. The joint ML estimation of the unknown but deterministic frequency and initial phase and MAP estimation of the unknown and random phase noise are implemented in time-domain processing, based on the discrete-time received observations. Here, the Wiener phase noise model and the best linearized AOPN model due to the AWGN proposed in our earlier work are used throughout the paper. The main contributions of this paper are summarised as follows:

1) Our newly derived ML/MAP estimators are closed-form, making full use of both the received signal phases and magnitudes. The derivation is based on the fact that using the received signal phases with the instantaneous received amplitude information incorporated in the AOPN model leads to the same ML/MAP estimates of the angle parameters as using the in-phase and quadrature components of the received signals [31]. More importantly, they all have a low-complexity, sample-by-sample iterative processing structure, which avoids the operation of matrix inversion and can be implemented easily in real-time. Moreover, in practical cases with unknown transmitted amplitude, the ML/MAP estimators can be implemented by using the received noisy signal magnitude to replace the transmitted amplitude. Simulations show that the estimation accuracy remains almost the same for a large SNR region.
2) We further provide the theoretical analysis of the estimation error variances of the ML/MAP estimators. As a basis for comparison, the explicit CRLBs for estimating the unknown but deterministic frequency and initial phase, and the exact Bayesian CRLB (BCRLB) for estimating the random phase noise are derived, which indicate the best estimation performance that can be achieved in the presence of phase noise. It is theoretically shown that the ML/MAP estimates obtained are unbiased, and the mean-square error (MSE) of the estimators attain the CRLB/BCRLB at high SNR.
3) Performance analysis of all the estimators are further provided by Monte Carlo simulations, where the improved phase unwrapping algorithm proposed in [31, Sec. IV-C] is used. The specific dependence of the estimation performance on the SNR, the phase noise variance and the number of independent observations is demonstrated in detail by simulation results. It is worth noting that the high-SNR approximation for the derivation will not affect the ML/MAP estimation accuracy, and the threshold SNR can go as low as -4 dB in small phase noise.
4) The linear minimum-MSE (LMMSE) implementation of the weighted phase averager (WPA) estimator is also discussed as an alternative for the frequency estimation in carrier phase noise. Even though it performs worse than the ML estimator, it is easier to implement in practice, since it makes use of the phase differences between the contiguous noisy received signal samples and can avoid the phase unwrapping in most cases.

This paper is organized as follows. Section II introduces the signal model with Wiener carrier phase noise and channel AWGN represented by the best linearized AOPN model. Section III presents the time-domain phase-based frequency/phase ML/MAP estimators. The iterative implementation is discussed in Section IV. Section V analyzes the MSE performance of the estimators and obtains the explicit CRLBs and BCRLB. Section VI discusses the phase unwrapping and the LMMSE implementation of the WPA frequency estimator in phase noise. Numerical and simulation results are given in Section VII. Section VIII concludes the paper.

II. Signal Model in Wiener Phase Noise

Here, we consider a single sinusoid with unknown frequency and initial phase observed in AWGN and carrier phase noise. The complex, baseband received signal $r(k)$ in discrete-time k is given by

$$r(k) = Ae^{j(\omega k+\phi+\theta(k))} + n(k) \tag{1}$$

where A is the transmitted signal amplitude which is assumed to be known when necessary, ω is the unknown nonrandom frequency and ϕ is the unknown nonrandom initial phase. The sequence $\{\theta(k), k = 0, 1, 2\cdots\}$ is the carrier phase noise over the interval $(-\pi, +\pi]$, due to the imperfect oscillators. The random-walk phase noise model used here is the Wiener process, i.e., we have $\theta(k) = \theta(k-1) + \Delta\theta(k)$, where $\theta(0) = 0$ and $\{\Delta\theta(k)\}$ is a sequence of independent, identically distributed (i.i.d.), zero-mean, real Gaussian random variables with variance σ_p^2 (rad^2). The AWGN $\{n(k)\}$ is a sequence of discrete-time, circularly symmetric, zero-mean, i.i.d., complex Gaussian random variables with covariance function $E[n(k)n^*(k-l)] = N_0\delta(l)$. Here, the SNR is defined as A^2/N_0.

The received signal is further rewritten in polar-coordinates form $(|r(k)|, \angle r(k))$ as

$$\begin{aligned} r(k) &= |r(k)|e^{j\angle r(k)} \\ &= |r(k)|e^{j(\omega k+\phi+\theta(k)+\epsilon(k))}. \end{aligned} \tag{2}$$

Here, $|r(k)|$ and $\angle r(k)$ are the received signal amplitude and the unwrapped received signal phase, respectively. Note that $\angle r(k)$ can be obtained from the principal argument of $r(k)$ by phase unwrapping. Term $\epsilon(k)$ is the AOPN due to the AWGN $n(k)$, and [31, Eq. (9)] shows that $\epsilon(k)$ is Tikhonov distributed conditioned on knowing $|r(k)|$. Further, it is shown that the Tikhonov probability density function (pdf) becomes the Gaussian pdf for high SNR,[1] leading to the best linearized AOPN model given by

$$p(\epsilon(k) \mid |r(k)|) = \frac{\exp\left[-\frac{\epsilon(k)^2}{2\alpha}\right]}{\sqrt{2\pi\alpha}}, \alpha = \frac{N_0}{2A|r(k)|}. \tag{3}$$

We will use this Gaussian-distributed AOPN model, i.e., $\epsilon(k) \sim N(0, \alpha)$, throughout this paper.

The received signal phases $\{\angle r(k)\}$ up till time $k = N-1$ can be expressed in the vector form as

$$\angle\mathbf{r} = \omega\mathbf{N} + \phi\mathbf{1} + \boldsymbol{\theta} + \boldsymbol{\epsilon} \tag{4}$$

where $\angle\mathbf{r} = [\angle r(0), \angle r(1)\cdots\angle r(N-1)]^T$, $\mathbf{N} = [0, 1\cdots N-1]^T$, $\mathbf{1} = [1, 1\cdots 1]^T$, $\boldsymbol{\epsilon} = [\epsilon(0), \epsilon(1)\cdots\epsilon(N-1)]^T$ and $\boldsymbol{\theta} = [\theta(0), \theta(1)\cdots\theta(N-1)]^T$. All the above vectors are N-dimensional column vectors, and superscript T denotes the vector transpose. Note that $\boldsymbol{\theta}$ and $\boldsymbol{\epsilon}$ are statistically independent of each other. We have $\boldsymbol{\theta} \sim N(\mathbf{0}, \boldsymbol{\Sigma}_{\boldsymbol{\theta}})$ and $\boldsymbol{\epsilon} \sim N(\mathbf{0}, \boldsymbol{\Sigma}_{\boldsymbol{\epsilon}})$, where the covariance matrices $\boldsymbol{\Sigma}_{\boldsymbol{\theta}}$ and $\boldsymbol{\Sigma}_{\boldsymbol{\epsilon}}$ can be determined as

$$\boldsymbol{\Sigma}_{\boldsymbol{\theta}} = \mathrm{Cov}(\boldsymbol{\theta}, \boldsymbol{\theta}) = \begin{bmatrix} 0 & 0 & 0 & 0 & \cdots & 0 \\ 0 & \sigma_p^2 & \sigma_p^2 & \sigma_p^2 & \cdots & \sigma_p^2 \\ 0 & \sigma_p^2 & 2\sigma_p^2 & 2\sigma_p^2 & \cdots & 2\sigma_p^2 \\ 0 & \sigma_p^2 & 2\sigma_p^2 & 3\sigma_p^2 & \cdots & 3\sigma_p^2 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \sigma_p^2 & 2\sigma_p^2 & 3\sigma_p^2 & \cdots & (N-1)\sigma_p^2 \end{bmatrix} \tag{5}$$

and

$$\boldsymbol{\Sigma}_{\boldsymbol{\epsilon}} = \mathrm{Cov}(\boldsymbol{\epsilon}, \boldsymbol{\epsilon}) = \begin{bmatrix} \frac{N_0}{2A|r(0)|} & 0 & \cdots & 0 \\ 0 & \frac{N_0}{2A|r(1)|} & 0 & \vdots \\ \vdots & 0 & \ddots & 0 \\ 0 & \cdots & 0 & \frac{N_0}{2A|r(N-1)|} \end{bmatrix}. \tag{6}$$

Note that $\boldsymbol{\Sigma}_{\boldsymbol{\theta}}$ is a $N \times N$ symmetric matrix, and $\boldsymbol{\Sigma}_{\boldsymbol{\epsilon}}$ is a $N \times N$ diagonal matrix.

III. ML Estimation of ω, ϕ and MAP Estimation of θ

We consider two basic criteria here, namely, the ML criterion and the MAP criterion that are widely used in signal parameter estimation. The ML criterion is for the parameters ω and ϕ that are treated as unknown but deterministic, and the MAP criterion is for the unknown and random parameters $\boldsymbol{\theta}$ here. In this section, we present the joint ML estimation of ω and ϕ and MAP estimation of $\boldsymbol{\theta}$, which is optimal using the observations $\{r(k), k = 0, 1, \ldots, N-1\}$.

In [31, Eqs. (3–4)] for the case of pure AWGN (i.e., $\theta(k) = 0$), we have shown that statistically $|r(k)|$ has no dependence on ω and ϕ, and the pdf term $p(|r(k)| \mid \omega, \phi)$ reduces to $p(|r(k)|)$. Moreover, given ω and ϕ, the only randomness in $\angle r(k)$ is due to $\epsilon(k)$ in pure AWGN, and the conditional pdf $p(\angle r(k) \mid \omega, \phi, |r(k)|)$ becomes $p_{\epsilon(k)}(\angle r(k) - (\omega k + \phi) \mid |r(k)|)$, where $p_{\epsilon(k)}()$ denotes the pdf of the random variable $\epsilon(k)$. Hence, accordingly in the presence of phase noise, the likelihood function $p(r(k) \mid \omega, \phi, \theta(k))$ can be expressed in polar coordinates $(|r(k)|, \angle r(k))$ as

$$\begin{aligned} &p(r(k) \mid \omega, \phi, \theta(k)) \\ &\quad = p_{\epsilon(k)}(\angle r(k) - (\omega k + \phi + \theta(k)) \mid |r(k)|)C_{(|r(k)|)} \end{aligned} \tag{7}$$

where the constant $C_{(|r(k)|)}$ is a function of $|r(k)|$, which has no dependence on the angle parameters. This equivalence means that the magnitude of $r(k)$ can be dropped once it is incorporated into the statistics of the argument of $r(k)$ in the estimation process. This result, therefore, establishes the theoretical basis for phase-based estimation of the single-sinusoid angle parameters. That is, using the received signal phases $\{\angle r(k)\}$ with received signal amplitudes $\{|r(k)|\}$ incorporated into the statistics of the AOPN $\{\epsilon(k)\}$ can lead to the same ML/MAP estimates for ω, ϕ and $\boldsymbol{\theta}$ as using the received signals $\{r(k)\}_{k=0}^{N-1}$.

Based on this phase-based time-domain approach established in our earlier work [31], we will estimate ω, ϕ and $\{\theta(k)\}_{k=0}^{N-1}$

[1] In practice, our simulations in Section VII show that this approximation works for values of SNR as low as 0 dB in ML estimation for a large range of phase noise variances.

using all the N received arguments of $\{r(k)\}_{k=0}^{N-1}$, i.e., the measurements of $\{\angle r(k)\}_{k=0}^{N-1}$. The estimates of ω, ϕ and $\boldsymbol{\theta}$ can be obtained by maximizing the joint pdf $p(\angle\mathbf{r}, \boldsymbol{\theta}|\omega, \phi)$:

$$\begin{aligned}\hat{\omega}^{(N-1)}, \hat{\phi}^{(N-1)}, \hat{\boldsymbol{\theta}}^{(N-1)} &= \arg\max_{\omega,\phi,\boldsymbol{\theta}} p(\angle\mathbf{r}, \boldsymbol{\theta}|\omega, \phi) \\ &= \arg\max_{\omega,\phi,\boldsymbol{\theta}} p(\boldsymbol{\theta}|\angle\mathbf{r}, \omega, \phi) p(\angle\mathbf{r}|\omega, \phi).\end{aligned} \tag{8}$$

Here, $\hat{\omega}^{(N-1)}$ and$\hat{\phi}^{(N-1)}$ denote the ML estimates of the unknown frequency and initial phase at the time point $k = N-1$, respectively, and $\hat{\boldsymbol{\theta}}^{(N-1)}$ denotes the MAP estimates of the phase noise up till time $k = N-1$.

From (8), we first consider the ML estimation of ω and ϕ by maximizing $p(\angle\mathbf{r}|\omega, \phi)$. Conditioning on ω and ϕ, we have $\angle\mathbf{r}|_{\omega,\phi} \sim N(\omega\mathbf{N} + \phi\mathbf{1}, \boldsymbol{\Sigma}_{\mathbf{r}})$ with the covariance matrix $\mathrm{Cov}(\angle\mathbf{r}, \angle\mathbf{r})$ given by $\boldsymbol{\Sigma}_{\mathbf{r}} = \boldsymbol{\Sigma}_{\boldsymbol{\theta}} + \boldsymbol{\Sigma}_{\boldsymbol{\epsilon}}$. That is, due to the independent Gaussian-distributed $\boldsymbol{\theta}$ and $\boldsymbol{\epsilon}$, the likelihood function $p(\angle\mathbf{r}|\omega, \phi)$ can be evaluated as

$$\begin{aligned}p(\angle\mathbf{r}|\omega, \phi) &= (2\pi)^{-\frac{N}{2}} \det(\boldsymbol{\Sigma}_{\mathbf{r}})^{-\frac{1}{2}} \\ &\times \exp\left[-\frac{1}{2}(\angle\mathbf{r} - \omega\mathbf{N} - \phi\mathbf{1})^T \boldsymbol{\Sigma}_{\mathbf{r}}^{-1}(\angle\mathbf{r} - \omega\mathbf{N} - \phi\mathbf{1})\right].\end{aligned} \tag{9}$$

Here, $\det(\boldsymbol{\Sigma}_{\mathbf{r}})^{-\frac{1}{2}}$, where $\det()$ denotes the matrix determinant, is a constant independent of ω and ϕ. Thus, maximizing $p(\angle\mathbf{r}|\omega, \phi)$ is equivalent to minimizing $\Lambda(\omega, \phi) = (\angle\mathbf{r} - \omega\mathbf{N} - \phi\mathbf{1})^T \boldsymbol{\Sigma}_{\mathbf{r}}^{-1}(\angle\mathbf{r} - \omega\mathbf{N} - \phi\mathbf{1})$ with respect to ω and ϕ. That is, the necessary condition to get the ML estimates of ω and ϕ, respectively, is that the corresponding values should satisfy

$$\frac{\partial\Lambda(\omega, \phi)}{\partial\omega} = 0 \text{ and } \frac{\partial\Lambda(\omega, \phi)}{\partial\phi} = 0. \tag{10}$$

By solving the above equations, we can get the ML estimates $\hat{\phi}^{(N-1)}$ and $\hat{\omega}^{(N-1)}$, respectively, given by

$$\hat{\phi}^{(N-1)} = \frac{\mathbf{1}^T \boldsymbol{\Sigma}_{\mathbf{r}}{}^{-1}(\angle\mathbf{r} - \omega\mathbf{N})}{\mathbf{1}^T \boldsymbol{\Sigma}_{\mathbf{r}}{}^{-1}\mathbf{1}} \tag{11}$$

and

$$\hat{\omega}^{(N-1)} = \frac{\mathbf{N}^T \boldsymbol{\Sigma}_{\mathbf{r}}{}^{-1}(\angle\mathbf{r} - \phi\mathbf{1})}{\mathbf{N}^T \boldsymbol{\Sigma}_{\mathbf{r}}{}^{-1}\mathbf{N}}. \tag{12}$$

Further, by jointly solving (11) and (12), the explicit ML estimates of the frequency ω and the initial phase ϕ are finally derived as

$$\begin{aligned}&\hat{\omega}^{(N-1)} \\ &= \frac{\left(\mathbf{N}^T \boldsymbol{\Sigma}_{\mathbf{r}}{}^{-1}\angle\mathbf{r}\right)\left(\mathbf{1}^T \boldsymbol{\Sigma}_{\mathbf{r}}{}^{-1}\mathbf{1}\right) - \left(\mathbf{N}^T \boldsymbol{\Sigma}_{\mathbf{r}}{}^{-1}\mathbf{1}\right)\left(\mathbf{1}^T \boldsymbol{\Sigma}_{\mathbf{r}}{}^{-1}\angle\mathbf{r}\right)}{\left(\mathbf{N}^T \boldsymbol{\Sigma}_{\mathbf{r}}{}^{-1}\mathbf{N}\right)\left(\mathbf{1}^T \boldsymbol{\Sigma}_{\mathbf{r}}{}^{-1}\mathbf{1}\right) - \left(\mathbf{1}^T \boldsymbol{\Sigma}_{\mathbf{r}}{}^{-1}\mathbf{N}\right)^2}\end{aligned} \tag{13}$$

and

$$\begin{aligned}&\hat{\phi}^{(N-1)} \\ &= \frac{\left(\mathbf{1}^T \boldsymbol{\Sigma}_{\mathbf{r}}{}^{-1}\angle\mathbf{r}\right)\left(\mathbf{N}^T \boldsymbol{\Sigma}_{\mathbf{r}}{}^{-1}\mathbf{N}\right) - \left(\mathbf{1}^T \boldsymbol{\Sigma}_{\mathbf{r}}{}^{-1}\mathbf{N}\right)\left(\mathbf{N}^T \boldsymbol{\Sigma}_{\mathbf{r}}{}^{-1}\angle\mathbf{r}\right)}{\left(\mathbf{N}^T \boldsymbol{\Sigma}_{\mathbf{r}}{}^{-1}\mathbf{N}\right)\left(\mathbf{1}^T \boldsymbol{\Sigma}_{\mathbf{r}}{}^{-1}\mathbf{1}\right) - \left(\mathbf{1}^T \boldsymbol{\Sigma}_{\mathbf{r}}{}^{-1}\mathbf{N}\right)^2}.\end{aligned} \tag{14}$$

For the case of no carrier phase noise, i.e., $\sigma_p^2 = 0$, the covariance matrix $\boldsymbol{\Sigma}_{\mathbf{r}}$ in (13) and (14) becomes the diagonal matrix $\boldsymbol{\Sigma}_{\boldsymbol{\epsilon}}$. In this way, we can have

$$\left(\mathbf{N}^T \boldsymbol{\Sigma}_{\mathbf{r}}{}^{-1}\angle\mathbf{r}\right) = \left(\mathbf{N}^T \boldsymbol{\Sigma}_{\boldsymbol{\epsilon}}{}^{-1}\angle\mathbf{r}\right) = \frac{2A}{N_0}\left(\sum_{k=0}^{N-1} k|r(k)|\angle r(k)\right). \tag{15}$$

And all the other product terms in (13) and (14) can reduce to have similar forms of cumulative sums. Thus, the ML estimates in (13) and (14) reduce to be exactly the same as those ML estimates in [3, Eqs. (16–17)] derived only in the presence of AWGN, when there is no carrier phase noise.

Next, we consider the MAP estimation of the Wiener phase noise $\boldsymbol{\theta}$. Since $\boldsymbol{\theta}$ is *a priori* Gaussian, and $\angle\mathbf{r}|_{\omega,\phi}$ is also *a priori* Gaussian, it follows that $p(\boldsymbol{\theta}|\angle\mathbf{r}, \omega, \phi)$ is a conditional Gaussian pdf. Thus, for given ω and ϕ, the conditional pdf $p(\boldsymbol{\theta}|\angle\mathbf{r}, \omega, \phi)$ is maximized with respect to $\boldsymbol{\theta}$ at the point $\hat{\boldsymbol{\theta}}^{(N-1)}|_{\omega,\phi} = \mathbb{E}[\boldsymbol{\theta}|\angle\mathbf{r}, \omega, \phi]$, where $\mathbb{E}()$ denotes the expectation operation. Also, using Bayes' theorem

$$p(\boldsymbol{\theta}|\angle\mathbf{r}, \omega, \phi) = \frac{p(\angle\mathbf{r}|\boldsymbol{\theta}, \omega, \phi)p(\boldsymbol{\theta}|\omega, \phi)}{p(\angle\mathbf{r}|\omega, \phi)} \tag{16}$$

the MAP estimate $\hat{\boldsymbol{\theta}}^{(N-1)}$ is equivalently obtained as

$$\begin{aligned}\hat{\boldsymbol{\theta}}^{(N-1)}|_{\omega,\phi} &= \arg\max_{\boldsymbol{\theta}} p(\angle\mathbf{r}|\boldsymbol{\theta}, \omega, \phi)p(\boldsymbol{\theta}|\omega, \phi) = \arg\min_{\boldsymbol{\theta}} \\ &\left[(\angle\mathbf{r} - \omega\mathbf{N} - \phi\mathbf{1} - \boldsymbol{\theta})^T \boldsymbol{\Sigma}_{\boldsymbol{\epsilon}}^{-1}(\angle\mathbf{r} - \omega\mathbf{N} - \phi\mathbf{1} - \boldsymbol{\theta})\right. \\ &\left.+\boldsymbol{\theta}^T \boldsymbol{\Sigma}_{\boldsymbol{\theta}}^{-1}\boldsymbol{\theta}\right].\end{aligned} \tag{17}$$

By solving (17), i.e., solving the equation

$$\frac{\partial p(\angle\mathbf{r}|\boldsymbol{\theta}, \omega, \phi)p(\boldsymbol{\theta}|\omega, \phi)}{\partial\boldsymbol{\theta}} = 0 \tag{18}$$

the MAP estimate of $\boldsymbol{\theta}$ given ω and ϕ can be derived as

$$\begin{aligned}\hat{\boldsymbol{\theta}}^{(N-1)}|_{\omega,\phi} &= \boldsymbol{\Sigma}_{\boldsymbol{\theta}}(\boldsymbol{\Sigma}_{\boldsymbol{\theta}} + \boldsymbol{\Sigma}_{\boldsymbol{\epsilon}})^{-1}(\angle\mathbf{r} - \omega\mathbf{N} - \phi\mathbf{1}) \\ &= \boldsymbol{\Sigma}_{\boldsymbol{\theta}} \boldsymbol{\Sigma}_{\mathbf{r}}{}^{-1}(\angle\mathbf{r} - \omega\mathbf{N} - \phi\mathbf{1}).\end{aligned} \tag{19}$$

Finally, substituting the ML estimates $\hat{\omega}^{(N-1)}$ and $\hat{\phi}^{(N-1)}$ in (13) and (14) into (19), we obtain the explicit MAP estimate $\hat{\boldsymbol{\theta}}^{(N-1)}|_{\hat{\omega},\hat{\phi}}$.

In this way, the triple $(\hat{\omega}^{(N-1)}, \hat{\phi}^{(N-1)}, \hat{\boldsymbol{\theta}}^{(N-1)})$ in (13), (14) and (19) give the explicit joint ML estimates of ω and ϕ and MAP estimate of $\boldsymbol{\theta}$. In the next section, we show that one can iteratively update the estimates $\hat{\omega}^{(N-1)}$, $\hat{\phi}^{(N-1)}$ and $\hat{\boldsymbol{\theta}}^{(N-1)}$ as N increases, i.e., as more measurements are available.

IV. Iterative Implementation

Note that all the estimates (13), (14) and (19) can be computed iteratively as the number of samples N increases.

For the derivation of the simple iterative implementation, we first need to efficiently compute the inverse of the matrix $\boldsymbol{\Sigma}_{\mathbf{r}}$ as N increases to $N+1$. Here, we let $\boldsymbol{\Sigma}_{\mathbf{r}(N+1)}$ denote the value of the matrix $\boldsymbol{\Sigma}_{\mathbf{r}}$ at time point $k = N$. The $(N+1)$-square matrix

$\mathbf{\Sigma_r}_{(N+1)}$ can be first expressed in block-matrix form as

$$\mathbf{\Sigma_r}_{(N+1)} = \begin{bmatrix} \mathbf{\Sigma_r}_{(N)} & \mathbf{N}\sigma_p^2 \\ \sigma_p^2\mathbf{N}^T & N\sigma_p^2 + \frac{N_0}{2A|r(N)|} \end{bmatrix}. \quad (20)$$

Then using the lemma of Schur complementation [32], the inverse of $\mathbf{\Sigma_r}_{(N+1)}$ can be derived in terms of the inverse of the N-square matrix $\mathbf{\Sigma_r}(N)$, and thus given in block-matrix form as (21) shown at the bottom of this page, where we have

$$S = N\sigma_p^2 + \frac{N_0}{2A|r(N)|} - \sigma_p^4\mathbf{N}^T\left(\mathbf{\Sigma_r}(N)\right)^{-1}\mathbf{N} \quad (22)$$

which is a scalar updated as N increases. At the time point $k = 1$, i.e., $N = 2$, $\left(\mathbf{\Sigma_r}_{(2)}\right)^{-1}$ is given straightforwardly as

$$\left(\mathbf{\Sigma_r}_{(2)}\right)^{-1} = \begin{bmatrix} (\frac{N_0}{2A|r(0)|})^{-1} & 0 \\ 0 & (\sigma_p^2 + \frac{N_0}{2A|r(1)|})^{-1} \end{bmatrix}. \quad (23)$$

As time k increases, in order to iteratively update the estimates $\hat{\omega}$, $\hat{\phi}$ in (13) and (14), we need to separately update the five terms $(\mathbf{N}^T{\mathbf{\Sigma_r}}^{-1}\angle\mathbf{r})$, $(\mathbf{1}^T{\mathbf{\Sigma_r}}^{-1}\mathbf{1})$, $(\mathbf{N}^T{\mathbf{\Sigma_r}}^{-1})$, $(\mathbf{1}^T{\mathbf{\Sigma_r}}^{-1}\angle\mathbf{r})$ and $(\mathbf{N}^T{\mathbf{\Sigma_r}}^{-1}\mathbf{N})$ inside. For simplicity, at time $k = N-1$, we denote

$$\begin{aligned} A^{(N-1)} &= \left(\mathbf{N}^T{\mathbf{\Sigma_r}}^{-1}\angle\mathbf{r}\right)^{(N-1)}, B^{(N-1)} = \left(\mathbf{1}^T{\mathbf{\Sigma_r}}^{-1}\mathbf{1}\right)^{(N-1)} \\ C^{(N-1)} &= \left(\mathbf{N}^T{\mathbf{\Sigma_r}}^{-1}\mathbf{1}\right)^{(N-1)}, D^{(N-1)} = \left(\mathbf{1}^T{\mathbf{\Sigma_r}}^{-1}\angle\mathbf{r}\right)^{(N-1)} \\ E^{(N-1)} &= \left(\mathbf{N}^T{\mathbf{\Sigma_r}}^{-1}\mathbf{N}\right)^{(N-1)}. \end{aligned} \quad (24)$$

At the starting time $k = 1$, using (23), we can easily have

$$\begin{aligned} A^{(1)} &= \angle r(1)\left(\sigma_p^2 + \frac{N_0}{2A|r(1)|}\right)^{-1}, \\ B^{(1)} &= \left(\frac{N_0}{2A|r(0)|}\right)^{-1} + \left(\sigma_p^2 + \frac{N_0}{2A|r(1)|}\right)^{-1}, \\ C^{(1)} &= \left(\sigma_p^2 + \frac{N_0}{2A|r(1)|}\right)^{-1}, \\ D^{(1)} &= \angle r(0)\left(\frac{N_0}{2A|r(0)|}\right)^{-1} + \angle r(1)\left(\sigma_p^2 + \frac{N_0}{2A|r(1)|}\right)^{-1}, \\ E^{(1)} &= C^{(1)} \end{aligned} \quad (25)$$

which are simple to calculate directly. Substituting (25) into (13) and (14), we simply have $\hat{\omega}^{(1)} = \angle r(1) - \angle r(0)$ and $\hat{\phi}^{(1)} = \angle r(0)$.

Further, applying the recursive matrix inversion result (21) and using the multiplication lemma of block matrices [32], we can derive $A^{(N)}$, $B^{(N)}$, $C^{(N)}$, $D^{(N)}$ and $E^{(N)}$ in terms of $(A^{(N-1)}, B^{(N-1)}, C^{(N-1)}, D^{(N-1)}, E^{(N-1)})$ given as

$$\begin{aligned} A^{(N)} &= (1 - N\sigma_p^2 S^{-1})A^{(N-1)} \\ &\quad + \sigma_p^2 S^{-1}E^{(N-1)}\left(\sigma_p^2 A^{(N-1)} - \angle r(N)\right) + NS^{-1}\angle r(N), \\ B^{(N)} &= B^{(N-1)} + S^{-1}\left(\sigma_p^2 C^{(N-1)} - 1\right)^2, \\ C^{(N)} &= (1 - N\sigma_p^2 S^{-1})C^{(N-1)} \\ &\quad + \sigma_p^2 S^{-1}E^{(N-1)}\left(\sigma_p^2 C^{(N-1)} - 1\right) + NS^{-1}, \\ D^{(N)} &= D^{(N-1)} + \sigma_p^4 S^{-1}A^{(N-1)}C^{(N-1)} \\ &\quad - \sigma_p^2 S^{-1}\left(A^{(N-1)} + \angle r(N)C^{(N-1)}\right) + S^{-1}\angle r(N), \\ E^{(N)} &= E^{(N-1)} + S^{-1}\left(\sigma_p^2 E^{(N-1)} - N\right)^2, \\ S^{(N)} &= S = N\sigma_p^2 + \frac{N_0}{2A|r(N)|} - \sigma_p^4 E^{(N-1)}. \end{aligned} \quad (26)$$

Finally, substituting the updated $(A^{(N)}, B^{(N)}, C^{(N)}, D^{(N)}, E^{(N)})$ into (13) and (14), we can get the updated estimates $\hat{\omega}^{(N)}$ and $\hat{\phi}^{(N)}$ straightforwardly. It is worth emphasizing that starting from (25), the updating process in (26) followed by substitution into (13) and (14) only involves the basic scalar operations, and avoids the matrix inversion and other complex matrix operations. Since the computational complexity of each iteration is $O(1)$, the overall computational complexity of our estimators using (26) is thus $O(N)$. Therefore, the process can be practically implemented in real-time, which makes full use of the new received sample $\{|r(N)|, \angle r(N)\}$ at time $k = N$.

Moreover, substituting the updated estimates $\hat{\omega}^{(N)}$ and $\hat{\phi}^{(N)}$ into (19), we can update $\hat{\boldsymbol{\theta}}$ at time $k = N$, that is

$$\begin{aligned} \hat{\boldsymbol{\theta}}^{(N)}|_{\hat{\omega}^{(N)},\hat{\phi}^{(N)}} &= \mathbf{\Sigma_\theta}^{(N)}\{{\mathbf{\Sigma_r}}^{-1}\}^{(N)} \\ &\left(\angle\mathbf{r}^{(N)} - \hat{\omega}^{(N)}\mathbf{N}^{(N)} - \hat{\phi}^{(N)}\mathbf{1}^{(N)}\right). \end{aligned} \quad (27)$$

Note that $\{{\mathbf{\Sigma_r}}^{-1}\}^{(N)}$ can be recursively obtained by (21), which avoids the matrix inversion operation. Thus, the iterative implementation of (27) requires only the matrix operations of addition, subtraction and multiplication.

Throughout the above analysis, the amplitude A of the sinusoid is assumed known and used in implementation. However, in most practical cases, the amplitude A may be unknown at the receiver. We consider two ways to overcome this issue. One is to replace A by the received noisy signal magnitude $|r(k)|$, as was pointed out and discussed in our earlier work [19], [31]. Simulations presented later will verify that the estimation accuracy is not affected much when using $|r(k)|$ to replace A in the estimators, especially for high SNR when $|r(k)|$ is almost equal to A for most times k. Another way is to first estimate A using our results in [33]. The method proposed therein enables us to estimate A independently of the frequency and phase, i.e., using $|r(k)|$ only. With the estimated A, the accuracy of the proposed estimation algorithms here is not affected much also.

$$\left(\mathbf{\Sigma_r}_{(N+1)}\right)^{-1} = \begin{bmatrix} \left(\mathbf{\Sigma_r}(N)\right)^{-1} + \sigma_p^4 S^{-1}\left(\mathbf{\Sigma_r}(N)\right)^{-1}\mathbf{N}\mathbf{N}^T\left(\mathbf{\Sigma_r}(N)\right)^{-1} & -\sigma_p^2 S^{-1}\left(\mathbf{\Sigma_r}(N)\right)^{-1}\mathbf{N} \\ -\sigma_p^2 S^{-1}\mathbf{N}^T\left(\mathbf{\Sigma_r}(N)\right)^{-1} & S^{-1} \end{bmatrix}. \quad (21)$$

V. Performance Analysis and Lower Bounds

It is well known that the CRLB and the BCRLB are two important performance limits which indicate the best estimation performance that can be achieved with the available observations. The CRLB is a lower bound on the variance of any unbiased estimate for the unknown but deterministic parameters, e.g., ω and ϕ here. The BCRLB provides a lower bound on the MMSE in estimating a random parameter, e.g., $\boldsymbol{\theta}$ here. In this section, we will analyze the MSE performance of the estimators (13), (14) and (19), and derive the unbiased CRLBs for (ω, ϕ) and the BCRLB for $\boldsymbol{\theta}$. Note that all the analysis here assumes that there is no phase unwrapping failures. In the following, we will eliminate the superscript $(N-1)$ on $\hat{\omega}^{(N-1)}$, $\hat{\phi}^{(N-1)}$ and $\hat{\boldsymbol{\theta}}^{(N-1)}$ for simplicity.

A. CRLBs and Performance of ML Estimators for (ω, ϕ)

The quality of the estimates of the frequency ω and the carrier phase ϕ can be measured in terms of their biases and their variances. Here, we use ω_0 and ϕ_0 to denote the actual values assumed by the unknown parameters ω and ϕ, respectively. Generally, it is desirable to obtain unbiased estimates, i.e., we want $\mathbb{E}(\hat{\omega}) = \omega_0$ and $\mathbb{E}(\hat{\phi}) = \phi_0$. Define the estimation errors as $\widetilde{\omega} = \hat{\omega} - \omega_0$ and $\widetilde{\phi} = \hat{\phi} - \phi_0$. Substituting (4) into (13) and (14), the estimation errors can thus be given, respectively, as (28) and (29) shown at the bottom of this page. Due to the fact that $\mathbb{E}(\boldsymbol{\theta} + \boldsymbol{\epsilon}) = 0$, we thus have $\mathbb{E}(\hat{\omega} - \omega_0) = 0$ and $\mathbb{E}(\hat{\phi} - \phi_0) = 0$, which verifies that the ML estimates $\hat{\omega}$ and $\hat{\phi}$ in (13) and (14) are unbiased.

Turning to the variances of $\hat{\omega}$ and $\hat{\phi}$, i.e., $\sigma^2_{\widetilde{\omega}} = \mathbb{E}[(\hat{\omega} - \omega_0)^2]$ and $\sigma^2_{\widetilde{\phi}} = \mathbb{E}[(\hat{\phi} - \phi_0)^2]$, they can be evaluated straightforwardly as follows. For $\sigma^2_{\widetilde{\omega}}$, we first have (30) shown at the bottom of this page. It further reduces to

$$\sigma^2_{\widetilde{\omega}} = \mathbb{E}_{\{|\mathbf{r}|\}}\left[\frac{\mathbf{1}^T{\boldsymbol{\Sigma}_\mathbf{r}}^{-1}\mathbf{1}}{\left(\mathbf{N}^T{\boldsymbol{\Sigma}_\mathbf{r}}^{-1}\mathbf{N}\right)\left(\mathbf{1}^T{\boldsymbol{\Sigma}_\mathbf{r}}^{-1}\mathbf{1}\right) - \left(\mathbf{1}^T{\boldsymbol{\Sigma}_\mathbf{r}}^{-1}\mathbf{N}\right)^2}\right]. \tag{31}$$

This is due to the facts that $\mathbb{E}[(\boldsymbol{\theta} + \boldsymbol{\epsilon})(\boldsymbol{\theta} + \boldsymbol{\epsilon})^T] = \boldsymbol{\Sigma}_\mathbf{r}$, and $\boldsymbol{\Sigma}_{\boldsymbol{\epsilon}}$ (equivalently $\boldsymbol{\Sigma}_\mathbf{r}$) is random due to its dependence on the set of magnitudes $|\mathbf{r}| = \{|r(k)|\}_{k=0}^{N-1}$ that are random. Similarly, for $\sigma^2_{\widetilde{\phi}}$, we have

$$\sigma^2_{\widetilde{\phi}} = \mathbb{E}_{\{|\mathbf{r}|\}}\left[\frac{\mathbf{N}^T{\boldsymbol{\Sigma}_\mathbf{r}}^{-1}\mathbf{N}}{\left(\mathbf{N}^T{\boldsymbol{\Sigma}_\mathbf{r}}^{-1}\mathbf{N}\right)\left(\mathbf{1}^T{\boldsymbol{\Sigma}_\mathbf{r}}^{-1}\mathbf{1}\right) - \left(\mathbf{1}^T{\boldsymbol{\Sigma}_\mathbf{r}}^{-1}\mathbf{N}\right)^2}\right]. \tag{32}$$

We can see that it is very complex to further evaluate $\sigma^2_{\widetilde{\omega}}$ and $\sigma^2_{\widetilde{\phi}}$ in (31) and (32) using the pdf of $|\mathbf{r}|$. However, using the further high SNR approximation that $|r(k)| \approx A$ for most times k, the covariance matrix $\boldsymbol{\Sigma}_\mathbf{r}$ reduces to be $\boldsymbol{\Sigma}'_\mathbf{r} = \boldsymbol{\Sigma}_{\boldsymbol{\theta}} + \frac{N_0}{2A^2}\mathbf{I}_{N\times N}$, where $\mathbf{I}_{N\times N}$ is the $N \times N$ identity matrix. Therefore, for high SNR, the variances (31) and (32) can be simplified, respectively, as

$$\sigma^2_{\widetilde{\omega}} = \frac{\mathbf{1}^T{\boldsymbol{\Sigma}'_\mathbf{r}}^{-1}\mathbf{1}}{\left(\mathbf{N}^T{\boldsymbol{\Sigma}'_\mathbf{r}}^{-1}\mathbf{N}\right)\left(\mathbf{1}^T{\boldsymbol{\Sigma}'_\mathbf{r}}^{-1}\mathbf{1}\right) - \left(\mathbf{1}^T{\boldsymbol{\Sigma}'_\mathbf{r}}^{-1}\mathbf{N}\right)^2} \tag{33}$$

and

$$\sigma^2_{\widetilde{\phi}} = \frac{\mathbf{N}^T{\boldsymbol{\Sigma}'_\mathbf{r}}^{-1}\mathbf{N}}{\left(\mathbf{N}^T{\boldsymbol{\Sigma}'_\mathbf{r}}^{-1}\mathbf{N}\right)\left(\mathbf{1}^T{\boldsymbol{\Sigma}'_\mathbf{r}}^{-1}\mathbf{1}\right) - \left(\mathbf{1}^T{\boldsymbol{\Sigma}'_\mathbf{r}}^{-1}\mathbf{N}\right)^2}. \tag{34}$$

From (33) and (34), we can see that the estimation error variances at high SNR depend on the SNR A^2/N_0, the phase noise variance σ_p^2 and the number of independent observations N. The higher the SNR, the larger the N or the smaller the σ_p^2, the lower are the estimation error variances. Section IV above implies that the inverse of $\boldsymbol{\Sigma}'_\mathbf{r}$ can be computed recursively as N increases, and this recursive inversion will allow (33) and (34) to be computed recursively in explicit forms. That is, we have $(\mathbf{1}^T{\boldsymbol{\Sigma}'_\mathbf{r}}^{-1}\mathbf{1})^{(N)} = B^{(N)}$, $(\mathbf{N}^T{\boldsymbol{\Sigma}'_\mathbf{r}}^{-1}\mathbf{1})^{(N)} = C^{(N)}$, and $(\mathbf{N}^T{\boldsymbol{\Sigma}'_\mathbf{r}}^{-1}\mathbf{N})^{(N)} = E^{(N)}$ given exactly the same as those in (26), only with the changes:

$$\begin{aligned} S^{(N)} &= N\sigma_p^2 + \frac{N_0}{2A^2} - \sigma_p^4 E^{(N-1)}, \\ E^{(1)} &= C^{(1)} = (\sigma_p^2 + \frac{N_0}{2A^2})^{-1}, B^{(1)} = \frac{2A^2}{N_0} + C^{(1)}. \end{aligned} \tag{35}$$

Next, we consider to derive the unbiased CRLBs for (ω, ϕ). It is well known that the unbiased CRLBs are the diagonal elements of the inverse of the Fisher information matrix $\boldsymbol{J_F}$ [34, Ch. 2]. Given ω and ϕ, the log-likelihood function of $p(\angle\mathbf{r}|\omega, \phi)$

$$\hat{\omega} - \omega_0 = \frac{\left(\mathbf{N}^T{\boldsymbol{\Sigma}_\mathbf{r}}^{-1}(\boldsymbol{\theta} + \boldsymbol{\epsilon})\right)\left(\mathbf{1}^T{\boldsymbol{\Sigma}_\mathbf{r}}^{-1}\mathbf{1}\right) - \left(\mathbf{N}^T{\boldsymbol{\Sigma}_\mathbf{r}}^{-1}\mathbf{1}\right)\left(\mathbf{1}^T{\boldsymbol{\Sigma}_\mathbf{r}}^{-1}(\boldsymbol{\theta} + \boldsymbol{\epsilon})\right)}{\left(\mathbf{N}^T{\boldsymbol{\Sigma}_\mathbf{r}}^{-1}\mathbf{N}\right)\left(\mathbf{1}^T{\boldsymbol{\Sigma}_\mathbf{r}}^{-1}\mathbf{1}\right) - \left(\mathbf{1}^T{\boldsymbol{\Sigma}_\mathbf{r}}^{-1}\mathbf{N}\right)^2}. \tag{28}$$

$$\hat{\phi} - \phi_0 = \frac{\left(\mathbf{1}^T{\boldsymbol{\Sigma}_\mathbf{r}}^{-1}(\boldsymbol{\theta} + \boldsymbol{\epsilon})\right)\left(\mathbf{N}^T{\boldsymbol{\Sigma}_\mathbf{r}}^{-1}\mathbf{N}\right) - \left(\mathbf{1}^T{\boldsymbol{\Sigma}_\mathbf{r}}^{-1}\mathbf{N}\right)\left(\mathbf{N}^T{\boldsymbol{\Sigma}_\mathbf{r}}^{-1}(\boldsymbol{\theta} + \boldsymbol{\epsilon})\right)}{\left(\mathbf{N}^T{\boldsymbol{\Sigma}_\mathbf{r}}^{-1}\mathbf{N}\right)\left(\mathbf{1}^T{\boldsymbol{\Sigma}_\mathbf{r}}^{-1}\mathbf{1}\right) - \left(\mathbf{1}^T{\boldsymbol{\Sigma}_\mathbf{r}}^{-1}\mathbf{N}\right)^2}. \tag{29}$$

$$\begin{aligned} \sigma^2_{\widetilde{\omega}} = \mathbb{E}_{\{\boldsymbol{\theta},\boldsymbol{\epsilon},|\mathbf{r}|\}}\Bigg[&\frac{\left(\mathbf{N}^T{\boldsymbol{\Sigma}_\mathbf{r}}^{-1}(\boldsymbol{\theta} + \boldsymbol{\epsilon})\right)^2\left(\mathbf{1}^T{\boldsymbol{\Sigma}_\mathbf{r}}^{-1}\mathbf{1}\right)^2 + \left(\mathbf{N}^T{\boldsymbol{\Sigma}_\mathbf{r}}^{-1}\mathbf{1}\right)^2\left(\mathbf{1}^T{\boldsymbol{\Sigma}_\mathbf{r}}^{-1}(\boldsymbol{\theta} + \boldsymbol{\epsilon})\right)^2}{\left(\left(\mathbf{N}^T{\boldsymbol{\Sigma}_\mathbf{r}}^{-1}\mathbf{N}\right)\left(\mathbf{1}^T{\boldsymbol{\Sigma}_\mathbf{r}}^{-1}\mathbf{1}\right) - \left(\mathbf{1}^T{\boldsymbol{\Sigma}_\mathbf{r}}^{-1}\mathbf{N}\right)^2\right)^2} \\ &- \frac{2\left(\mathbf{1}^T{\boldsymbol{\Sigma}_\mathbf{r}}^{-1}\mathbf{1}\right)\left(\mathbf{N}^T{\boldsymbol{\Sigma}_\mathbf{r}}^{-1}\mathbf{1}\right)\left(\mathbf{N}^T{\boldsymbol{\Sigma}_\mathbf{r}}^{-1}(\boldsymbol{\theta} + \boldsymbol{\epsilon})(\boldsymbol{\theta} + \boldsymbol{\epsilon})^T{\boldsymbol{\Sigma}_\mathbf{r}}^{-1}\mathbf{1}^T\right)}{\left(\left(\mathbf{N}^T{\boldsymbol{\Sigma}_\mathbf{r}}^{-1}\mathbf{N}\right)\left(\mathbf{1}^T{\boldsymbol{\Sigma}_\mathbf{r}}^{-1}\mathbf{1}\right) - \left(\mathbf{1}^T{\boldsymbol{\Sigma}_\mathbf{r}}^{-1}\mathbf{N}\right)^2\right)^2}\Bigg]. \end{aligned} \tag{30}$$

from (9) is given by

$$\log p(\angle\mathbf{r}|\omega,\phi) = C_{(\mathbf{\Sigma_r})} - \frac{1}{2}(\angle\mathbf{r} - \omega\mathbf{N} - \phi\mathbf{1})^T\mathbf{\Sigma_r}^{-1}(\angle\mathbf{r} - \omega\mathbf{N} - \phi\mathbf{1}) \quad (36)$$

where $C_{(\mathbf{\Sigma_r})}$ is only a function of $\det(\mathbf{\Sigma_r})$ which is independent of ω and ϕ. Then, it is easily shown that for the unknown parameter vector $[\omega\phi]^T$, the Fisher information matrix $\boldsymbol{J_F}$ is derived as

$$\boldsymbol{J_F} = -\mathbb{E}\begin{bmatrix} \frac{\partial^2 \log p(\angle\mathbf{r}|\omega,\phi)}{\partial\omega^2} & \frac{\partial^2 \log p(\angle\mathbf{r}|\omega,\phi)}{\partial\omega\partial\phi} \\ \frac{\partial^2 \log p(\angle\mathbf{r}|\omega,\phi)}{\partial\phi\partial\omega} & \frac{\partial^2 \log p(\angle\mathbf{r}|\omega,\phi)}{\partial\phi^2} \end{bmatrix} = \begin{bmatrix} \mathbf{N}^T\mathbf{\Sigma'_r}^{-1}\mathbf{N} & \mathbf{N}^T\mathbf{\Sigma'_r}^{-1}\mathbf{1} \\ \mathbf{N}^T\mathbf{\Sigma'_r}^{-1}\mathbf{1} & \mathbf{1}^T\mathbf{\Sigma'_r}^{-1}\mathbf{1} \end{bmatrix}. \quad (37)$$

Thus, we have

$$\boldsymbol{J_F}^{-1} = \frac{\begin{bmatrix} \mathbf{1}^T\mathbf{\Sigma'_r}^{-1}\mathbf{1} & -\mathbf{N}^T\mathbf{\Sigma'_r}^{-1}\mathbf{1} \\ -\mathbf{N}^T\mathbf{\Sigma'_r}^{-1}\mathbf{1} & \mathbf{N}^T\mathbf{\Sigma'_r}^{-1}\mathbf{N} \end{bmatrix}}{\left(\mathbf{N}^T\mathbf{\Sigma'_r}^{-1}\mathbf{N}\right)\left(\mathbf{1}^T\mathbf{\Sigma'_r}^{-1}\mathbf{1}\right) - \left(\mathbf{1}^T\mathbf{\Sigma'_r}^{-1}\mathbf{N}\right)^2}. \quad (38)$$

The CRLBs for ω and ϕ, i.e., CRLB_ω and CRLB_ϕ, are the first and second diagonal elements of the matrix $\boldsymbol{J_F}^{-1}$ in (38), respectively. It is worth noting that CRLB_ω and CRLB_ϕ are exactly the same as the variances of $\hat{\omega}$ and $\hat{\phi}$ in (33) and (34). That is, for high SNR, we have

$$\sigma^2_{\tilde{\omega}} = \text{CRLB}_\omega, \text{and} \sigma^2_{\tilde{\phi}} = \text{CRLB}_\phi. \quad (39)$$

The ML estimators (13) and (14) are thus unbiased and asymptotically efficient, since their variances attain the CRLBs at high SNR. Moreover, for the case of no carrier phase noise, i.e., $\sigma_p^2 = 0$, our CRLB_ω and CRLB_ϕ here will reduce to be

$$\text{CRLB}_{\omega|\sigma_p=0} = \frac{N_0}{A^2}\frac{6}{N(N+1)(N-1)}$$
$$\text{CRLB}_{\phi|\sigma_p=0} = \frac{N_0}{A^2}\frac{2N-1}{N(N+1)} \quad (40)$$

which are exactly the same as those results in [3, Eqs. (22–23)]. We can see that CRLB_ϕ is approximately linearly decreasing in N for large N, i.e., $\text{CRLB}_\phi \sim 1/N$. Whereas, CRLB_ω decreases much faster than CRLB_ϕ as N increases, since we almost have $\text{CRLB}_\omega \sim 1/N^3$. From (40), we can deduce that the effects of N on the CRLBs for ω and ϕ in (39) are also different in phase noise, which are similar to the case of no carrier phase noise. This phenomenon is verified by numerical results shown later.

B. BCRLB and Performance of MAP Estimator for θ

Similar to the performance analysis on the ML estimators of (ω,ϕ) shown above, the aim in this section is to analytically evaluate the MSE performance of the MAP estimator $\hat{\boldsymbol{\theta}}|_{\omega,\phi}$ in (19), and further derive the BCRLB for the random phase noise $\boldsymbol{\theta}$.

We first define the estimation error vector as $\widetilde{\boldsymbol{\theta}}|_{\omega,\phi} = \hat{\boldsymbol{\theta}}|_{\omega,\phi} - \boldsymbol{\theta}$ and the error covariance matrix as $\mathbf{\Sigma}_{\widetilde{\boldsymbol{\theta}}} = \mathbb{E}[\widetilde{\boldsymbol{\theta}}|_{\omega,\phi}(\widetilde{\boldsymbol{\theta}}|_{\omega,\phi})^T]$. Substituting (4) into (19) with exactly known ω and ϕ, we can easily have

$$\mathbb{E}[\widetilde{\boldsymbol{\theta}}|_{\omega,\phi}] = \mathbb{E}[\mathbf{\Sigma_\theta}\mathbf{\Sigma_r}^{-1}(\boldsymbol{\theta}+\boldsymbol{\epsilon}) - \boldsymbol{\theta}] = \mathbf{0} \quad (41)$$

and

$$\begin{aligned}\mathbf{\Sigma}_{\widetilde{\boldsymbol{\theta}}} &= \mathbb{E}_{\{\boldsymbol{\theta},\boldsymbol{\epsilon},|\mathbf{r}|\}}[(\mathbf{\Sigma_\theta}\mathbf{\Sigma_r}^{-1}(\boldsymbol{\theta}+\boldsymbol{\epsilon}) - \boldsymbol{\theta})(\mathbf{\Sigma_\theta}\mathbf{\Sigma_r}^{-1}(\boldsymbol{\theta}+\boldsymbol{\epsilon}) - \boldsymbol{\theta})^T] \\ &= \mathbb{E}_{\{|\mathbf{r}|\}}\left[\mathbf{\Sigma_\theta}\mathbf{\Sigma_r}^{-1}\mathbf{\Sigma_\epsilon}\right] \end{aligned} \quad (42)$$

due to the fact that $\mathbb{E}[\boldsymbol{\theta}] = \mathbb{E}[\boldsymbol{\epsilon}] = \mathbf{0}$ and $\{\boldsymbol{\theta},\boldsymbol{\epsilon}\}$ are independent of each other. The MAP estimator of $\boldsymbol{\theta}$ is thus unbiased. For high SNR with $|r(k)| \approx A$ for most times k, the covariance matrix $\mathbf{\Sigma}_{\widetilde{\boldsymbol{\theta}}}$ in (42) thus reduces approximately to

$$\begin{aligned}\mathbf{\Sigma}_{\widetilde{\boldsymbol{\theta}}} &= \frac{N_0}{2A^2}\mathbf{\Sigma_\theta}\left[\mathbf{\Sigma_\theta} + \frac{N_0}{2A^2}\mathbf{I}_{N\times N}\right]^{-1} \\ &= \left[\mathbf{\Sigma_\theta}^{-1} + \frac{2A^2}{N_0}\mathbf{I}_{N\times N}\right]^{-1} \end{aligned} \quad (43)$$

where the second identity is due to the matrix inversion lemma [32]: $(\mathbf{I}+\mathbf{A}^{-1})^{-1} = \mathbf{A}(\mathbf{I}+\mathbf{A})^{-1}$. We see that the estimation error variances for $\boldsymbol{\theta}$ at high SNR also depend on the SNR A^2/N_0, the phase noise variance σ_p^2 and the number of independent observations N.

Next, given (ω,ϕ), the BCRLB for $\boldsymbol{\theta}$ is considered. Since $\boldsymbol{\theta}$ is random with known *a priori* Gaussian pdf, the BCRLBs for $\{\theta(k), k = 0, 1\cdots N-1\}$ are the diagonal elements of the inverse of the information matrix $\boldsymbol{J_T}$ given by [34, Ch. 2]

$$\boldsymbol{J_T} = \boldsymbol{J'_F} + \boldsymbol{J_P} \quad (44)$$

where $\boldsymbol{J'_F}$ is the Fisher information matrix representing information obtained from the measurements $\{\angle\mathbf{r}\}$, and $\boldsymbol{J_P}$ represents the *a priori* information of $\boldsymbol{\theta}$. According to the Bayes' theorem (16), the explicit expressions of $\boldsymbol{J'_F}$ and $\boldsymbol{J_P}$ can be derived based on the log-likelihood functions of $p(\angle\mathbf{r}|\boldsymbol{\theta},\omega,\phi)$ and $p(\boldsymbol{\theta}|\omega,\phi)$, respectively, as

$$\begin{aligned}\boldsymbol{J'_F} &= \mathbb{E}\left[\frac{\partial \log p(\angle\mathbf{r}|\boldsymbol{\theta},\omega,\phi)}{\partial\boldsymbol{\theta}}\left(\frac{\partial \log p(\angle\mathbf{r}|\boldsymbol{\theta},\omega,\phi)}{\partial\boldsymbol{\theta}}\right)^T\right] \\ &= \mathbf{\Sigma'_\epsilon}^{-1} \end{aligned} \quad (45)$$

where we have $\mathbf{\Sigma'_\epsilon} = \frac{N_0}{2A^2}\boldsymbol{I}_{N\times N}$, and

$$\begin{aligned}\boldsymbol{J_P} &= \mathbb{E}\left[\frac{\partial \log p(\boldsymbol{\theta}|\omega,\phi)}{\partial\boldsymbol{\theta}}\left(\frac{\partial \log p(\boldsymbol{\theta}|\omega,\phi)}{\partial\boldsymbol{\theta}}\right)^T\right] \\ &= \mathbf{\Sigma_\theta}^{-1}. \end{aligned} \quad (46)$$

Finally, substituting (45) and (46) into (44), we thus have the BCRLB for $\boldsymbol{\theta}$ given by

$$\begin{aligned}\text{BCRLB}_{\boldsymbol{\theta}} &= \boldsymbol{J_T}^{-1} = (\mathbf{\Sigma_\theta}^{-1} + \mathbf{\Sigma'_\epsilon}^{-1})^{-1} \\ &= \mathbf{\Sigma_\theta}(\mathbf{\Sigma_\theta} + \mathbf{\Sigma'_\epsilon})^{-1}\mathbf{\Sigma'_\epsilon} \end{aligned} \quad (47)$$

where the last identity is due to the matrix inversion lemma [32]: $\mathbf{A}^{-1} + \mathbf{B}^{-1} = \mathbf{A}^{-1}(\mathbf{A}+\mathbf{B})\mathbf{B}^{-1}$. It is worth noting that (47) is equivalent to (43), and we thus have $\mathbf{\Sigma}_{\widetilde{\boldsymbol{\theta}}} = \text{BCRLB}_{\boldsymbol{\theta}}$ for high SNR. This means that the estimation error variance of the MAP estimator (19) with known (ω,ϕ) can attain the BCRLB at high SNR, which verifies its asymptotic efficiency.

More importantly, we can see from (47) that $\mathrm{BCRLB}_{\boldsymbol{\theta}}$ mainly depends on the AOPN covariance matrix $\boldsymbol{\Sigma}'_{\boldsymbol{\epsilon}}$ and the phase noise covariance matrix $\boldsymbol{\Sigma}_{\boldsymbol{\theta}}$. That is, the relationship between the inverse SNR $\frac{N_0}{A^2}$ and the phase noise variance σ_p^2 determines the value of $\mathrm{BCRLB}_{\boldsymbol{\theta}}$. For very large phase noise or extremely high SNR, i.e., $\sigma_p^2 \gg \frac{N_0}{A^2}$, $\mathrm{BCRLB}_{\boldsymbol{\theta}}$ approximately becomes $\boldsymbol{\Sigma}'_{\boldsymbol{\epsilon}}$, which means that the AOPN effect dominates in the $\mathrm{BCRLB}_{\boldsymbol{\theta}}$. On the hand, for very small phase noise or low SNR, i.e., $(N-1)\sigma_p^2 \ll \frac{N_0}{A^2}$, $\mathrm{BCRLB}_{\boldsymbol{\theta}}$ asymptotically reduces to be $\boldsymbol{\Sigma}'_{\boldsymbol{\theta}}$, which means that the phase noise effect dominates in the $\mathrm{BCRLB}_{\boldsymbol{\theta}}$.

It is worth noting that these performance analysis results for the BCRLB for $\boldsymbol{\theta}$ are only for the case of perfectly known values of ω and ϕ. For the general case of imperfectly estimated values of ω and ϕ, we can determine the performance of the estimator for $\boldsymbol{\theta}$ only using simulations, and these results will be shown in Section VII later.

VI. Discussions on Phase Unwrapping and LMMSE-WPA Frequency Estimation

In this section, we will first analyze the improved phase unwrapping algorithm used for the implementation of ML/MAP estimators. Then, to avoid unacceptable performance degradation due to phase unwrapping failure in phase noise, the LMMSE-WPA estimator is further discussed.

A. Phase Unwrapping

Note that all the estimators (13), (14) and (19) make use of both the measurement phases $\{\angle r(k)\}_{k=0}^{N-1}$ and the measurement magnitudes $\{|r(k)|\}_{k=0}^{N-1}$ in practical implementation. Since the actual phase data samples are obtained from the principal arguments $\{\arg[r(k)]\}$ which are within the interval $[-\pi, \pi)$, phase unwrapping is thus required to generate the unwrapped received signal phases $\{\angle r(k)\}$. We will use the improved phase unwrapping algorithm proposed in [31, Sec. IV-C] throughout the simulations, which is briefly discussed as follows. To retrieve the actual unwrapped phases $\{\angle r(k)\}$, the phase differences between adjacent data samples are used, and defined as

$$\begin{aligned}\Delta(k) &= \arg[r^*(k-1)r(k)] \\ &= \omega + \Delta\theta(k) + \epsilon(k) - \epsilon(k-1)\end{aligned} \tag{48}$$

where $\Delta(k)$ denotes the phase difference between data samples at time $k-1$ and k, and $*$ denotes the conjugate operation. In practice, $\{\Delta(k)\}$ can be obtained by projecting data sample $r(k)$ onto $r(k-1)$ and then taking the phase of the resulting complex quantity. Specifically, based on the first observation on the phase at time $k=0$, i.e., $\angle r(0)$, a sequence of angle data $\{\varphi(k)\}$ can be computed as

$$\begin{aligned}\varphi(1) &= \angle r(0) + \Delta(1), \\ \varphi(2) &= \varphi(1) + \Delta(2), \\ \varphi(3) &= \varphi(2) + \Delta(3), \\ &\vdots \\ \varphi(N-1) &= \varphi(N-2) + \Delta(N-1), \\ &\vdots\end{aligned} \tag{49}$$

It is obvious that $\{\angle r(k)\}$ can be retrieved from the computed $\{\varphi(k)\}$ as they are mathematically identical. Thus, in practical implementation, the phase of $r(k)$ is collected by unwrapping $\arg[r(k)]$ to within a 2π-interval centered around the computed value $\varphi(k)$, i.e., the value of $\angle r(k)$ chosen is the one lying in the interval $(\varphi(k)-\pi, \varphi(k)+\pi]$. This is simply done by adding multiples of $\pm 2\pi$ to $\arg[r(k)]$, when we have $|\arg[r(k)] - \varphi(k)| > \pi$. Note that this unwrapped phase may be different from the true value of $\angle r(k)$. A phase unwrapping failure occurs when $|\angle r(k) - \varphi(k)| > \pi$ [35].

The necessary conditions for a good performance of this phase unwrapping operation in phase noise are thus that $\angle r(0)$ is equal to the principal argument of $r(0)$, i.e., we have [31]

$$-\pi \le \angle r(0) = \arg[r(0)] = \phi + \epsilon(0) \le \pi \tag{50}$$

and that $\Delta(k)$ in (48) for any k satisfies

$$-\pi \le \Delta(k) = \omega + \Delta\theta(k) + \epsilon(k) - \epsilon(k-1) \le \pi. \tag{51}$$

Clearly, the validity of these two conditions is expected to depend on the SNR as well as the values of ω and ϕ, since a value of ω and ϕ close to $\pm\pi$ can lead easily to the actual value of $\phi + \epsilon(0)$ in (50) and $\omega + \Delta\theta(k) + \epsilon(k) - \epsilon(k-1)$ in (51) falling outside of the boundaries of the interval, due to a small amount of AOPN. When the conditions (50) and/or (51) are not satisfied, $\arg[r(k)]$ cannot be unwrapped correctly to $\angle r(k)$ and a phase unwrapping failure will occur. The probabilities that $\angle r(0)$ and $\Delta(k)$ lie outside the interval $[-\pi, \pi]$ can be computed simply, using the best linearized model (3). That is, given ω and ϕ, the conditional probabilities $\Pr[|\angle r(0)| \ge \pi \mid \omega, \phi]$ and $\Pr[|\Delta(k)| \ge \pi \mid \omega, \phi]$ are given as

$$\begin{aligned}\Pr[|\angle r(0)| \ge \pi \mid \omega, \phi] &= Q\left(\frac{\pi+\phi}{\sqrt{\frac{N_0}{2A|r(0)|}}}\right) + Q\left(\frac{\pi-\phi}{\sqrt{\frac{N_0}{2A|r(0)|}}}\right) \\ \Pr[|\Delta(k)| \ge \pi \mid \omega, \phi] &= Q\left(\frac{\pi+\omega}{\sqrt{\sigma_p^2 + \frac{N_0}{2A|r(k)|} + \frac{N_0}{2A|r(k-1)|}}}\right) \\ &\quad + Q\left(\frac{\pi-\omega}{\sqrt{\sigma_p^2 + \frac{N_0}{2A|r(k)|} + \frac{N_0}{2A|r(k-1)|}}}\right)\end{aligned} \tag{52}$$

where $Q(x) = \frac{1}{\sqrt{2\pi}}\int_x^\infty \exp(-\frac{u^2}{2})du$ is the Gaussian Q-function. It is worth noting from (49) that the algorithm is limited by the unwrapping failure propagation, since a particular failure at time point j will affect all succeeding points, because $\Delta(j)$ is added to each of $\{\varphi(k)\}_{k=j+1}^{\infty}$. Therefore, the exact phase unwrapping failure probability is difficult to obtain due to the error propagation and the nonlinearity of the modulo-2π operation.

B. LMMSE Implementation of the WPA Frequency Estimator

Phase unwrapping is a very challenging problem in phase extraction, and no perfect phase-unwrapping solution has been obtained even with an enormous research effort [35], [36]. Thus, to avoid phase unwrapping algorithms used in practice, the WPA estimator using the differenced received signal phase is first proposed in [14].

In this section, we will introduce the LMMSE implementation of the WPA estimator in the presence of carrier phase noise, to estimate the unknown frequency ω at time point $k = N-1$ based on all the past and current observations $r(k), k = 0, 1, \ldots, N-1$. As mentioned in [14], the parameter to be estimated here is only the frequency, and the phase is considered to be a nuisance parameter, since the phase differences involved cancel the initial carrier phase ϕ and only keep the incremental phase noise $\Delta\theta(k)$ at each time.

Suppose that a fixed block of N data samples $\{r(k)\}_{k=0}^{N-1}$ is collected for the processing, the inputs to be fed into the estimator are the $N-1$ phase differences $\{\Delta(k)\}_{k=1}^{N-1}$ defined in (48). Therefore, using the vector $\mathbf{\Delta}$ to denote a block of phase differences, we have

$$\mathbf{\Delta} = \mathbf{1}\omega + \mathbf{\Xi} \tag{53}$$

where $\mathbf{\Delta} = [\Delta(1) \quad \Delta(2) \cdots \Delta(N-1)]^T$, $\mathbf{1} = [1 \quad 1 \cdots 1]^T$ with components all equal to 1, and $\mathbf{\Xi} = [\Delta\theta(1) + \epsilon(1) - \epsilon(0), \quad \Delta\theta(2) + \epsilon(2) - \epsilon(1) \cdots \Delta\theta(N-1) + \epsilon(N-1) - \epsilon(N-2)]^T$. Note that $\mathbf{\Delta}$, $\mathbf{1}$ and $\mathbf{\Xi}$ are all $(N-1)$-dimensional column vectors.

The frequency estimate made based on $\mathbf{\Delta}$ in (53) is denoted as $\hat{\omega}_\Delta$ here. The LMMSE implementation of the WPA estimator is a linear function of $\{\Delta(k)\}$, i.e., we have $\hat{\omega}_\Delta = \mathbf{w}^T\mathbf{\Delta}$, where $\mathbf{w}$ is an $(N-1)$-dimensional weighting column vector. By minimizing the MSE $E[(\hat{\omega}_\Delta - \omega)^2]$, the solution is thus given by [31, Eq. (42)]

$$\hat{\omega}_\Delta = \mathbf{w}^T\mathbf{\Delta} = \frac{\mathbf{1}^T\mathbf{C}^{-1}\mathbf{\Delta}}{\mathbf{1}^T\mathbf{C}^{-1}\mathbf{1}}. \tag{54}$$

Here, $\mathbf{C} = [c_{ij}]$ is the covariance matrix of the noise vector $\mathbf{\Xi}$, which has a tridiagonal form and whose components depend on the specific AOPN model. If the best linearized AOPN model (3) is used, we have

$$c_{ij} = \begin{cases} \frac{N_0}{2A|r(i)|} + \frac{N_0}{2A|r(i-1)|} + \sigma_p^2, & i = j \\ -\frac{N_0}{2A|r(j)|}, & i - j = 1 \\ -\frac{N_0}{2A|r(i)|}, & j - i = 1 \\ 0, & |i-j| > 1. \end{cases} \tag{55}$$

We can see that this estimator (54) uses the online received magnitude information $|r(k)|$ to identify those instantaneous received signal phase samples with higher noise variances, and are therefore weighted less. Our AOPN model (3) shows that a smaller received signal magnitude indicates a received signal phase sample with a larger variance. Accordingly, in the region of low SNR when large noise samples are more common, this feature can be exploited to indicate to the estimator the exact positions of the observation data samples which are less reliable, and thus should be weighted less in their contribution to the estimator (54).

Note that in comparison to the ML estimator, it can be easily verified using Monte Carlo simulations that the accuracy of the LMMSE-WPA estimator is not sensitive to the small changes on small values of ω_0. However, for larger values of ω_0, especially values that are near to $+\pi$ or $-\pi$, the estimation accuracy will be affected. Specifically, we can see from (48) that the sum $(\Delta\theta(k) + \epsilon(k) - \epsilon(k-1))$ can more easily drive the phase difference $\Delta(k)$ across the $+\pi$ boundary when ω_0 is near to $+\pi$, meaning that the probability increases as ω_0 increases or as the SNR decreases or as σ_p^2 increases. Based on the condition in (51), when $\Delta(k)$ goes across the $+\pi$ boundary by an amount of, say, $\alpha > 0$, it will be wrapped around and will appear as the value $-\pi + \alpha$. This will make that particular value of $\Delta(k)$ very noisy, and thus degrade the estimation performance.

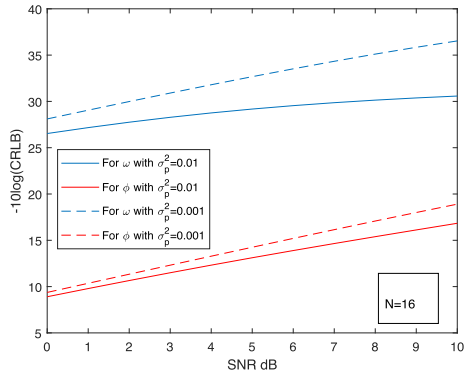


Fig. 1. CRLBs for estimation of ω and ϕ in phase noise as a function of SNR given $N = 16$.

VII. Numerical and Simulation Results

This section presents the numerical and simulation results to illustrate the behavior of the variance of the ML estimators (13) and (14), the MSE of the MAP estimator (19), the MSE of the LMMSE-WPA estimator (54), the CRLB_ω and CRLB_ϕ in (33) and (34) and the $\text{BCRLB}_{\boldsymbol{\theta}}$ in (47) as a function of the various parameters, e.g., the SNR, the number of observations N and the phase noise variance σ_p^2.

The inverse MSE (IMSE) of all the estimators, for example, defined as $\text{IMSE}_\omega = -10\log_{10}\{\sum_{k=1}^{q}[\omega - \hat{\omega}]^2/q\}$ and $\text{IMSE}_\phi = -10\log_{10}\{\sum_{k=1}^{q}[\phi - \hat{\phi}]^2/q\}$, are obtained using Monte Carlo simulations, where q denotes the iteration number for each calculation of IMSE. We have $q = 10^5$ to ensure the accuracy. The inverse CRLBs and BCRLB are also computed in decibels as the benchmarks for numerical comparison. As a figure of merit to characterize the performances of estimators, we define the threshold SNR of an estimator as the value of SNR at which its inverse variance curve dips by 1 dB from the inverse CRLB/BCRLB curve, as is common in the literature [3]. Note that the phase unwrapping algorithm we implement here is the improved phase unwrapping algorithm proposed in [31, Sec. IV-C]. As discussed in Section IV, to attain a good phase unwrapping performance, one must satisfy the necessary conditions (50) and (51).

A. Numerical Analysis of CRLBs and BCRLBs

As a basis for comparison, the inverse CRLBs and BCRLBs are first numerically analysed. As Fig. 1 shows for $N = 16$, the CRLB_ω is much smaller than the CRLB_ϕ through the whole SNR region, which is expected as the result (40) implies. And as σ_p^2 increases, the CRLBs increases inevitably for ω and ϕ, and the CRLB_ω deteriorates faster than the CRLB_ϕ for the same SNR. As Fig. 2 shows, as σ_p^2 increases or the SNR decreases, the $\text{BCRLB}_{\boldsymbol{\theta}}$ for any k increases distinctly. Moreover, as k increases from 1 to $N-1$, the BCRLB for $\theta(k)$ increases obviously, which is expected.

Fig. 3 shows the effect of different N on the CRLBs for ω and ϕ and the BCRLB for $\boldsymbol{\theta}$, respectively. We can see from

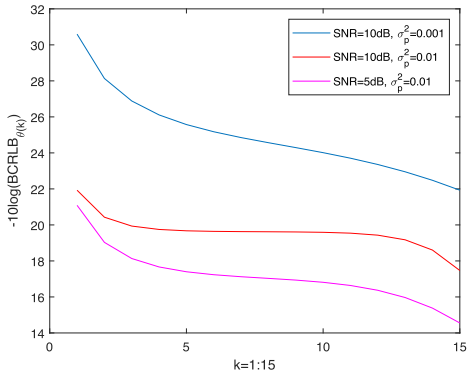
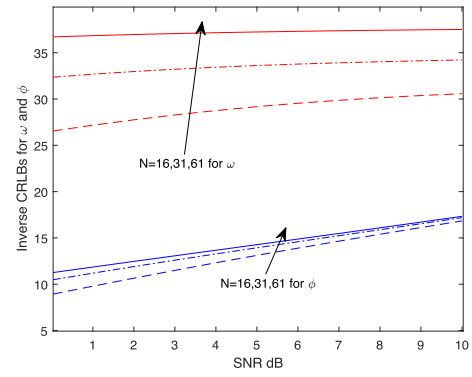
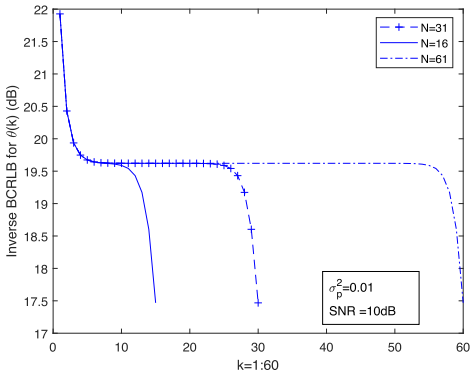


Fig. 2. BCRLBs for estimation of $\{\theta(k), k = 1 : N-1\}$ as a function of k given $N = 16$ and various σ_p^2 and SNR.



(a)



(b)

Fig. 3. (B)CRLBs with different N given $\sigma_p^2 = 10^{-2}\text{rad}^2$. (a) For estimation of ω and ϕ. (b) For estimation of $\{\theta(k), k = 1 : N-1\}$.

Fig. 3(a) that even for large phase noise, the CRLB $_\omega$ decreases much faster than the CRLB$_\phi$ with the increase in N, which is in accordance with the no carrier phase noise case, as Section IV.A explains. Fig. 3(b) shows that for the comparable values of $\frac{1}{2\text{SNR}} = 0.05$ and $\sigma_p^2 = 0.01$ rad^2, the inverse BCRLB remains around 19.7 dB for a larger range of k given larger N.

B. Effect of Parameters on the ML Estimation Performance

Due to the nature of the phase unwrapping failures, the simulated estimation performance is expected to be sensitive

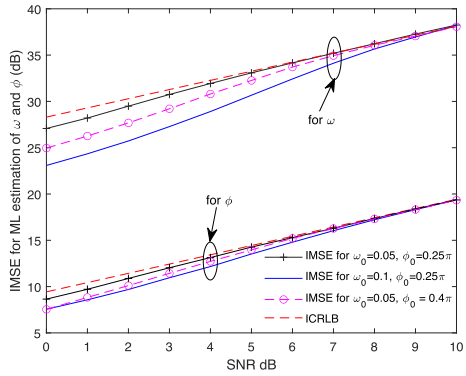


Fig. 4. Simulated IMSE comparison for ML estimation of ω and ϕ with different angle parameters given $N = 16$.

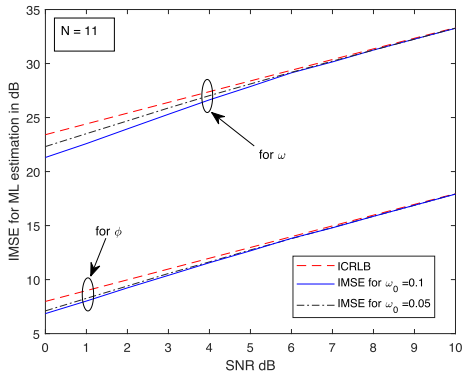


Fig. 5. Simulated IMSE comparison for ML estimation of ω and ϕ with different ω_0 given $N = 11$ and $\phi_0 = 0.25\pi$.

to the actual values of the angle parameters being estimated. For instance, as discussed in [3] for the pure AWGN case, the estimation performance will degrade as the actual frequency ω_0 increases. For the phase noise case with $\sigma_p^2 = 10^{-4}\text{rad}^2$, Fig. 4 shows that the estimation performance degrades as ω_0 or ϕ_0 increases. We can see that as ω_0 increases from 0.05 to 0.1 only, the threshold SNR for ML estimation of ω and ϕ increases from 0 dB to about 7 and 4 dB, respectively. Similarly, for $\omega_0 = 0.05$ and $\phi_0 = 0.4\pi$, the SNR threshold for ML estimation of ω and ϕ becomes 4 dB and 2 dB, respectively.

Note that the quantitative performance loss with respect to the increase of ω_0 or ϕ_0 also depends on the specific value of N. This is because that the larger N will lead to the larger probability of phase unwrapping failure in the observation window. As Fig. 5 shows for $N = 11$, as ω_0 increases to 0.1, the threshold SNR for estimation of ω increases about 4 dB, whereas the threshold SNR for estimation of ϕ remains about 0 dB. In comparison to the case of $N = 16$ in Fig. 4, the performance loss for $N = 11$ is much smaller as ω_0 increases.

Therefore, to simply illustrate the performance of the newly derived estimators in phase noise, we set small values of $\omega_0 = 0.05$ and $\phi_0 = 0.25\pi$ throughout the following simulations. Especially when ω_0, N and σ_p^2 are small, the unwrapping failure hardly happens and could be ignored as SNR increases. For the cases with large values of ω_0, N or σ_p^2, one should discard those runs in which a phase unwrapping failure occurred, otherwise, the estimation errors would be very large and the simulated MSE

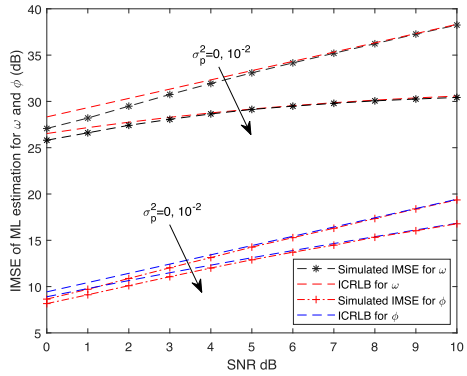


Fig. 6. Simulated IMSE and ICRLBs for ML estimation of ω and ϕ with $N = 16$ and $\sigma_p^2 = 0$ and 10^{-2}rad^2.

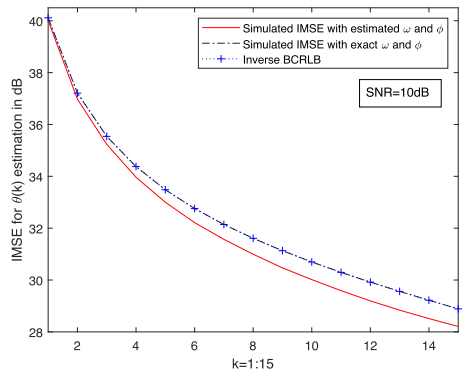
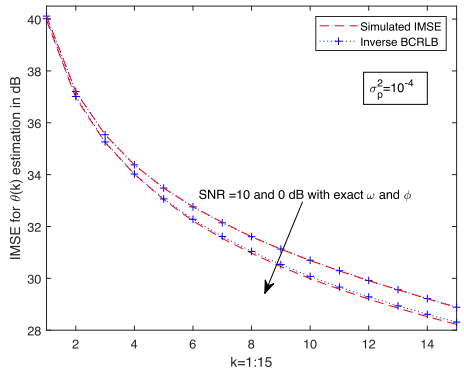
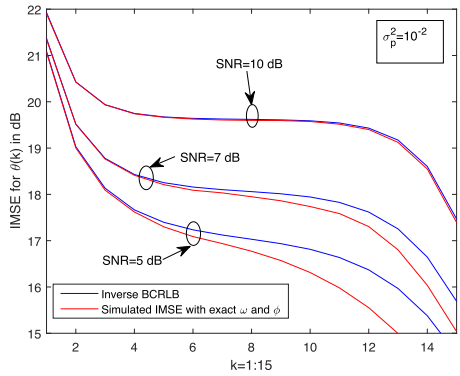


Fig. 7. Simulated IMSE comparison for the MAP estimation of $\{\theta(k), k = 1:15\}$ with $N = 16$ and $\sigma_p^2 = 10^{-4}\text{rad}^2$.



(a)



(b)

Fig. 8. Inverse BCRLBs and simulated IMSE for the MAP estimation of $\{\theta(k), k = 1:15\}$ with $N = 16$. (a) $\sigma_p^2 = 10^{-4}\text{rad}^2$. (b) $\sigma_p^2 = 10^{-2}\ \text{rad}^2$.

will not be compatible with the analytical prediction and the CRLB.

C. Simulated IMSE Analysis of the ML/MAP Estimators

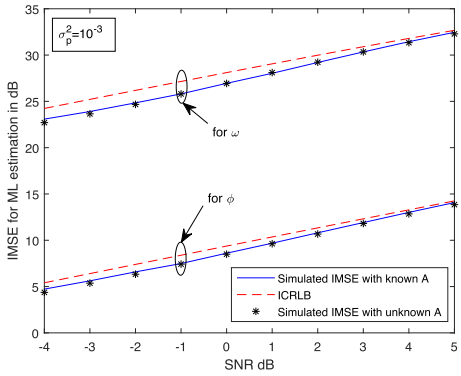
As Fig. 6 shows for both $\sigma_p^2 = 0$ and 10^{-2}rad^2, the simulated MSE quickly attains the CRLB_ω and CRLB_ϕ as SNR increases. The simulation results show that for $N = 16$, the threshold SNRs for the ML estimation of ω and ϕ are all about 0 dB. This implies that a good estimation performance can be achieved at low SNR in various phase noise, even though we assume high SNR when deriving our ML estimators.

The simulated IMSE performance for the MAP estimation of $\{\theta(k), k = 1:N\}$ with $N = 16$ is shown in Figs. 7 and 8. First, Fig. 7 shows that at the same SNR value of 10 dB, the IMSE performance with the exact ω and ϕ coincides with the inverse BCRLB, whereas the performance using the estimated values still dips by about 1 dB from the inverse BCRLB curve. This means that the MAP estimator (19) using the estimated ω and ϕ performs worse than that using the true ω_0 and ϕ_0, which is expected. Then, Fig. 8 shows the detailed simulated IMSE performance given exact ω_0 and ϕ_0. As Fig. 8(a) shows for $\sigma_p^2 = 10^{-4}\text{rad}^2$, the MSE can attain the BCRLB for any k in the SNR range of [0, 10]dB. Whereas, Fig. 8(b) shows that the threshold SNR is about 5 dB for $\sigma_p^2 = 10^{-2}\ \text{rad}^2$, and the MSE of the MAP estimator for any $\theta(k)$ attains the BCRLB at 10 dB. Fig. 8(b) thus shows the asymptotic efficiency of the MAP estimator for $\{\theta(k), k = 1:N\}$ as SNR increases, even in large phase noise.

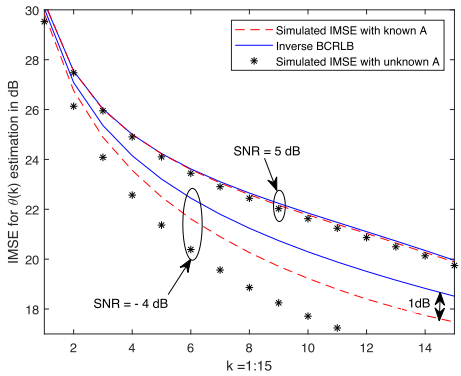
D. Performance Verification for Estimation With Unknown Amplitude

As mentioned in Section IV, to solve the practical issue when the amplitude A may be unknown at the receiver, we can replace A by the received noisy signal magnitude $|r(k)|$ in the implementation of the ML/MAP estimators. When testing the estimation accuracy in the SNR region of $[0 \sim 10]$dB, we will have the same simulation results as in Figs. 6 and 8 for $\sigma_p^2 = 0 \sim 10^{-2}\text{rad}^2$. To illustrate in distinction, we only show the case of estimation in low SNR $[-4 \sim 5]$ dB in Fig. 9 here. As Fig. 9(a) shows, the accuracy of the ML estimators retains in low SNR when replacing A by $|r(k)|$. For the MAP estimation of $\{\theta(k), k = 1:N-1\}$, we can verify that the accuracy remains almost the same in the SNR region above 0 dB. However, the replacement will degrade the estimation accuracy in low SNR below 0 dB, as Fig. 9(b) shows for SNR $= -4$ dB.

Moreover, Fig. 9 also shows that implementing with known amplitude A, the threshold SNR can be as low as -4 dB for both ML estimation and MAP estimation in phase noise of $\sigma_p^2 = 10^{-3}\text{rad}^2$. It emphasizes again that the outstanding estimation performance can be achieved even in low SNR.



(a)



(b)

Fig. 9. Simulated IMSE with/without knowing A given $N = 16$ and $\sigma_p^2 = 10^{-3}\text{rad}^2$. (a) For ML estimation of ω and ϕ. (b) For MAP estimation of $\{\theta(k), k = 1 : N - 1\}$.

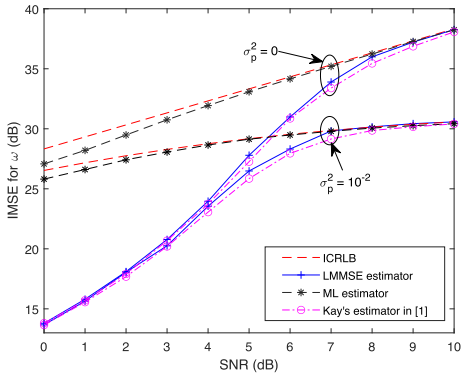
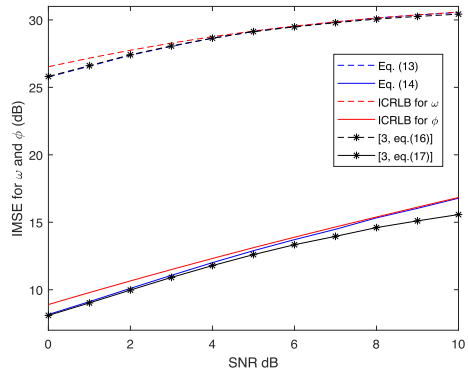


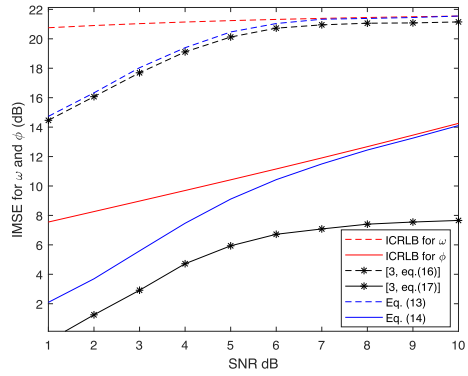
Fig. 10. Simulated IMSE comparison for different frequency estimators given $\omega_0 = 0.05$ and $N = 16$.

E. Comparison with the LMMSE-WPA and Improved Kay's Estimators

First, for the comparison with the LMMSE-WPA frequency estimator (54), Fig. 10 shows that the LMMSE-WPA estimation performance deteriorates much faster than the ML one as SNR decreases. The threshold SNR of the LMMSE-WPA estimation is about 7 dB, and its estimation accuracy is not sensitive to the small changes of the frequency ω_0 as expected.



(a)



(b)

Fig. 11. Simulated IMSE comparison with the ML estimators [3, Eqs. (16–17)] in phase noise. (a) $\sigma_p^2 = 10^{-2}\text{rad}^2$. (b) $\sigma_p^2 = 0.1\ \text{rad}^2$.

Moreover, several classical frequency estimation techniques are mostly considered for designing practical systems [1]. The frequency-domain method is based on the interpolation of Fourier coefficients. One can test using simulations that it requires a huge number of samples for reliable discrete Fourier transform, which can not guarantee the real-time estimation performance. The time-domain method mostly used is the well-known Kay's method with different smoothing parameters. Similar to the LMMSE implementation of the WPA estimator, it is easy to implement in practice. As Fig. 10 shows, the improved Kay's estimator in [1, Eqs. (15–16)] only performs slightly worse than the LMMSE implementation.

Note that in most literatures for parameter estimation in phase noise (e.g., [25]–[27]), the non-linear character of the complicated estimation rules proposed renders ML estimation infeasible. Therefore, for simplicity, most of them finally turn to the approximate solution: least-squares estimation, that is very similar to the Kay's estimator except for the weighting coefficients. From Fig. 10, we believe that all the estimators whose structure is similar to the Kay's estimator will have similar performance, and the ML estimator we derive here will give the optimal solution in both accuracy and efficiency.

F. Comparison with the ML Estimators Designed in Pure AWGN

As Fig. 11(a) shows for $\sigma_p^2 = 10^{-2}\text{rad}^2$, the phase estimator [3, Eq. (17)] has about 1.3 dB performance loss at SNR $=$ 10 dB. Fig. 11(b) shows that the frequency estimator

[3, Eq. (16)] cannot converge to the CRLB_ω at high SNR for $\sigma_p^2 = 0.1$ rad^2. As expected, the performance loss using [3, Eqs. (16–17)] cannot be eliminated, even with the SNR increased in large phase noise. Moreover, Fig. 11(b) demonstrates a remarkable performance improvement using the ML phase estimation (14), compared to using [3, Eq. (17)]. We can see that the MSE performance of [3, Eq. (17)] still dips by 6 dB from the CRLB_ϕ curve at high SNR of 10 dB.

Undoubtedly, the estimation performance in phase noise can be much improved with the ML estimators (13) and (14) used, instead of using the ML estimators [3, Eqs. (16–17)] designed in pure AWGN. In addition, as σ_p^2 increases to 0.1 rd^2, the threshold SNR becomes about 5 dB, as Fig. 11(b) shows. This is to be expected, since all the estimators derived here will have lower performance as the phase noise variance increases, even though the phase noise statistics is taken into account optimally in their design.

VIII. Conclusion

The joint ML estimators for the frequency and initial phase and MAP estimator for the Wiener phase noise are derived in closed form. They are all expressed in weighted linear combinations of the received signal phases, which are simple to be iteratively implemented in practice. At high SNR, the estimators can attain the performance limits specified by the CRLB/BCRLB, which can serve as the benchmarks for the estimation accuracy in phase noise. The LMMSE implementation of the WPA estimator is also given as an alternative for the frequency estimation in phase noise. Numerical comparison and simulation results show in detail how the MSE performs as the SNR, the phase noise variance or the number of observations varies. The estimation performance is remarkably improved when the estimators take into account the carrier phase noise, compared to most existing estimators in both accuracy and efficiency.

References

[1] J. Qiu and S. Kandeepan, "Performance of frequency estimation techniques with phase noise in mm-Wave based 5G systems," in *Proc. Asia Model. Symp.*, 2017, pp. 170–175.

[2] S. Zhang, P. Y. Kam, C. Yu, and J. Chen, "Decision-aided carrier phase estimation for coherent optical communications," *J. Lightw. Technol.*, vol. 28, no. 11, pp. 1597–1607, 2010.

[3] H. Fu and P. Y. Kam, "MAP/ML estimation of the frequency and phase of a single sinusoid in noise," *IEEE Trans. Signal Process.*, vol. 55, no. 3, pp. 834–845, Mar. 2007.

[4] B. E. Paden, "Lower bounds on the frequency estimation error in magnetically coupled MEMS resonant sensors," *IEEE Trans. Biomed. Circuits Syst.*, vol. 10, no. 1, pp. 18–24, Feb. 2016.

[5] G. Morren, P. Lemmerling, and S. Van Huffel, "Decimative subspace-based parameter estimation techniques applied to magnetic resonance spectroscopy signals," in *Proc. 23rd Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, 2001, vol. 3, pp. 2319–2322.

[6] J. Dong *et al.*, "Simultaneous phase unwrapping and removal of chemical shift (SPURS) using graph cuts: Application in quantitative susceptibility mapping," *IEEE Trans. Med. Imag.*, vol. 34, no. 2, pp. 531–540, Feb. 2015.

[7] I. Omelchuk and I. Chyrka, "Maximum a posteriori estimator of the harmonic signal frequency," in *Proc. IEEE Microw., Radar, Remote Sens. Symp.*, 2017, pp. 51–54.

[8] P. M. Baggenstoss, "Recursive decimation/interpolation for ML chirp parameter estimation," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 50, no. 1, pp. 445–455, Jan. 2014.

[9] T. Radil, P. M. Ramos, and A. Cruz Serra, "New spectrum leakage correction algorithm for frequency estimation of power system signals," *IEEE Trans. Instrum. Meas.*, vol. 58, no. 5, pp. 1670–1679, May 2009.

[10] I. Kamwa, M. Leclerc, and D. McNabb, "Performance of demodulation-based frequency measurement algorithms used in typical PMUs," *IEEE Trans. Power Del.*, vol. 19, no. 2, pp. 505–514, Apr. 2004.

[11] L. Palmer, "Coarse frequency estimation using the discrete Fourier transform," *IEEE Trans. Inf. Theory*, vol. IT-20, no. 1, pp. 104–109, Jan. 1974.

[12] D. Rife and R. Boorstyn, "Single tone parameter estimation from discrete-time observations," *IEEE Trans. Inf. Theory*, vol. IT-20, no. 5, pp. 591–598, Sep. 1974.

[13] S. Tretter, "Estimating the frequency of a noisy sinusoid by linear regression (Corresp.)," *IEEE Trans. Inf. Theory*, vol. IT-31, no. 6, pp. 832–835, Nov. 1985.

[14] S. Kay, "A fast and accurate single frequency estimator," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 37, no. 12, pp. 1987–1990, Dec. 1989.

[15] K. S. Lyalin, V. K. Tsvetkov, A. Y. Sheremet, and V. V. Kurganov, "The study of RF PLL phase noise influence to SAR image quality," in *Proc. IEEE Conf. Russian Young Researchers Elect. Electron. Eng.*, 2018, pp. 1670–1672.

[16] K. Siddiq, M. K. Hobden, S. R. Pennock, and R. J. Watson, "Phase noise in FMCW radar systems," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 55, no. 1, pp. 70–81, Feb. 2019.

[17] A. L. Swindlehurst, E. Ayanoglu, P. Heydari, and F. Capolino, "Millimeter-wave massive MIMO: The next wireless revolution?," *IEEE Commun. Mag.*, vol. 52, no. 9, pp. 56–62, Sep. 2014.

[18] R. Wang, H. Mehrpouyan, M. Tao, and Y. Hua, "Channel estimation, carrier recovery, and data detection in the presence of phase noise in OFDM relay systems," *IEEE Trans. Wireless Commun.*, vol. 15, no. 2, pp. 1186–1205, Feb. 2016.

[19] Q. Wang and P. Y. Kam, "Optimum detection of two-dimensional carrier modulations with linear phase noise using received amplitude and phase information and performance analysis," *J. Lightw. Technol.*, vol. 34, no. 10, pp. 2439–2451, May 2016.

[20] F. Herzel and D. Kissinger, "Phase noise analysis of a Homodyne radar system driven by a phase-locked loop," in *Proc. IEEE Int. Symp. Circuits Syst.*, 2017, pp. 1–4.

[21] G. Krieger and M. Younis, "Impact of oscillator noise in bistatic and multistatic SAR," *IEEE Geosci. Remote Sens. Lett.*, vol. 3, no. 3, pp. 424–428, Jul. 2006.

[22] V. L. Petrović, M. M. Janković, A. V. Lupšić, V. R. Mihajlović, and J. S. Popović-Božović, "High-accuracy real-time monitoring of heart rate variability using 24 GHz continuous-wave doppler radar," *IEEE Access*, vol. 7, pp. 74721–74733, 2019.

[23] A. D. Droitcour, O. Boric-Lubecke, V. M. Lubecke, J. Lin, and G. T. A. Kovacs, "Range correlation and I/Q performance benefits in single-chip silicon doppler radars for noncontact cardiopulmonary monitoring," *IEEE Trans. Microw. Theory Techn.*, vol. 52, no. 3, pp. 838–848, Mar. 2004.

[24] E. Ip and J. M. Kahn, "Feedforward carrier recovery for coherent optical communications," *J. Lightw. Technol.*, vol. 25, no. 9, pp. 2675–2692, 2007.

[25] A. Barbieri and G. Colavolpe, "On the Cramer-Rao bound for carrier frequency estimation in the presence of phase noise," *IEEE Trans. Wireless Commun.*, vol. 6, no. 2, pp. 575–582, Feb. 2007.

[26] Y. Su, K. T. Wong, and K. R. Ho, "MMSE recursive estimation of high phase-noise that is wiener non-stationary," in *Proc. IEEE Radar Conf.*, 2009, pp. 1–5.

[27] J. Bhatti, N. Noels, and M. Moeneclaey, "Low-complexity frequency offset and phase noise estimation for burst-mode digital transmission," in *Proc. IEEE 22nd Int. Symp. Pers., Indoor Mobile Radio Commun.*, 2011, pp. 1662–1666.

[28] P. Y. Kam, "Maximum likelihood carrier phase recovery for linear suppressed-carrier digital data modulations," *IEEE Trans. Commun.*, vol. COM-34, no. 6, pp. 522–527, Jun. 1986.

[29] A. Meiyappan, H. Kim, and P. Y. Kam, "A low-complexity, low-cycle-slip-probability, format-independent carrier estimator with adaptive filter length," *J. Lightw. Technol.*, vol. 31, no. 23, pp. 3806–3812, 2013.

[30] P. Handel, "Markov-based single-tone frequency estimation," *IEEE Trans. Circuits Syst. II: Analog Digit. Signal Process.*, vol. 45, no. 2, pp. 230–232, Feb. 1998.

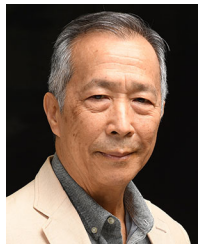
[31] H. Fu and P. Y. Kam, "Phase-based, time-domain estimation of the frequency and phase of a single sinusoid in AWGN—The role and applications of the additive observation phase noise model," *IEEE Trans. Inf. Theory*, vol. 59, no. 5, pp. 3175–3188, May 2013.

[32] K. Petersen and M. Pedersen, *The matrix Cookbook*. 2008, pp. 1–71. [Online]. Available: http://matrixcookbook.com
[33] M.-W. Wu, Y. Jin, Y. Li, T. Song, and P.-Y. Kam, "Maximum-likelihood, magnitude-based, amplitude and noise variance estimation," *IEEE Signal Process. Lett.*, vol. 28, pp. 414–418, 2021.
[34] H. L. Van Trees, *Detection, Estimation and Modulation Theory: Part I*. New York, NY, USA: Wiley, 1968.
[35] Y. Li, H. Fu, and P. Y. Kam, "Improved, approximate, time-domain ML estimators of chirp signal parameters and their performance analysis," *IEEE Trans. Signal Process.*, vol. 57, no. 4, pp. 1260–1272, Apr. 2009.
[36] Z. Xu, T. Lu, and B. Huang, "Fast frequency estimation algorithm by least squares phase unwrapping," *IEEE Signal Process. Lett.*, vol. 23, no. 6, pp. 776–779, Jun. 2016.

Qian Wang (Member, IEEE) received the B.Eng. degree in communications engineering from Harbin Engineering University, Harbin, China, in 2012, and the Ph.D. degree in electrical & computer engineering in 2017 from the National University of Singapore, Singapore, where she received the honor of President's Graduate Fellowship. From 2017 to 2019, she was a Research Engineer with Central Research Institute, Huawei 2012 Lab, where she contributed to IEEE 802.11ad/ay standards. She is currently a Research Associate Professor with the College of Electronics and Information Engineering, Shenzhen University, Shenzhen, China. Her research interests include communication and information theory, signal processing algorithms, and performance analysis/optimization in optical and wireless communications. She is also a Member of Optical Society of America, and a Senior Member of China Communication Society.

Zhi Quan (Senior Member, IEEE) received the B.E. degree in communications engineering from the Beijing University of Posts and Telecommunications, Beijing, China, in 1999, and the Ph.D. degree in electrical engineering from the University of California, Los Angeles (UCLA), Los Angeles, CA, USA, with highest honors in 2009. He is currently a Distinguished Professor with the College of Electronic and Information Engineering, Shenzhen University, Shenzhen, China. He was a Senior System Engineer with Qualcomm Research Center (QRC) of Qualcomm Inc., San Diego, CA, USA, during 2008–2012, and as a Senior RF System Engineer and Architect with Apple Inc. (Cupertino, CA) during 2012–2015. Dr. Quan had contributed to IEEE 802.11ac/ah standards with more than 30 U.S. issued patents and authored or coauthored more than 50 papers in wireless communications and signal processing with more than 5000 citations from Google Scholar. Dr. Quan was the recipient of the UCLA Outstanding Ph.D. Award in 2009, IEEE Signal Processing Society Best Paper Award in 2012, China National Excellent Young Scientist Foundation in 2016, and First Prize Technology Innovation Award by China Institute of Communications in 2020. His current research interests include wireless communication systems, RF system calibration and measurement, data-driven signal processing, and machine learning.

Suzhi Bi (Senior Member, IEEE) received the B.Eng. degree in communications engineering from Zhejiang University, Hangzhou, China, in 2009, and the Ph.D. degree in information engineering from The Chinese University of Hong Kong, Hong Kong, in 2013. From 2013 to 2015, he was a Postdoctoral Research Fellow with the Department of Electrical and Computer Engineering, National University of Singapore, Singapore. Since 2015, he has been with the College of Electronics and Information Engineering, Shenzhen University, Shenzhen, China, where he is currently an Associate Professor. His research interests include optimizations in wireless information and power transfer, mobile computing, and wireless sensing. He was the recipient of the 2019 IEEE ComSoc Asia-Pacific Outstanding Young Researcher Award, 2021 IEEE ComSoc Asia-Pacific Outstanding Paper Award, and best paper awards of IEEE SmartGridComm 2013 and IEEE/CIC ICCC 2021. He is also an Associate Editor for the IEEE WIRELESS COMMUNICATIONS LETTERS.



Pooi-Yuen Kam (Life Fellow, IEEE) was born in Ipoh, Malaysia. He received the S.B., S.M., and Ph.D. degrees in electrical engineering from the Massachusetts Institute of Technology, Cambridge, MA, USA, in 1972, 1973, and 1976, respectively.

From 1976 to 1978, he was a Member of the technical staff with the Bell Telephone Laboratories, Holmdel, NJ, USA, where he was engaged in packet network studies. Since 1978, he has been a Professor with the Department of Electrical and Computer Engineering, National University of Singapore, Singapore, where he was the Deputy Dean of Engineering and the Vice Dean for Academic Affairs of the Faculty of Engineering from 2000 to 2003. He spent the sabbatical year 1987 to 1988, with the Tokyo Institute of Technology, Tokyo, Japan, under the sponsorship of the Hitachi Scholarship Foundation. In 2006, he was invited to the School of Engineering Science, Simon Fraser University, Burnaby, BC, Canada, as the David Bensted Fellow. From April 2015 to March 2017, he was a Distinguished Guest Professor (Global) with the Graduate School of Science and Technology, Keio University, Tokyo. He was a Visiting Professor with the Department of Electronic Engineering, Shanghai Jiao Tong University, Shanghai, China, in 2017, and a Visiting Professor with the University of Electronic Science and Technology of China, Chengdu, China, in 2018. Since 2019, he has been a Professor with the School of Science and Engineering, The Chinese University of Hong Kong in Shenzhen, Shenzhen, China, where he is also the Associate Dean for Student Affairs. Since 2016, he has been with the National University of Singapore Suzhou Research Institute, Suzhou, China, where he is the Principal Investigator of a project on free-space optical communications funded by the National Natural Science Foundation of China. His research interests include communication and information theory and signal processing, and their applications to wireless and optical communications.

Dr. Kam is a Member of the Eta Kappa Nu, Tau Beta Pi, and Sigma Xi. He was elected as a Fellow of the IEEE for his contributions to receiver design and performance analysis for wireless communications. He served on the IEEE VTS awards Committee for years 2019, 2020 and 2021, and on the IEEE ComSoc Fellow Evaluation Committee since 2020. He served on the IEEE Fellow Committee for years 2016, 2017 and 2018, and on the IEEE Fellows Strategic Planning Subcommittee for 2018 and 2019. He was the recipient of the Best Paper Award at the IEEE VTC2004-Fall, at the IEEE VTC2011-Spring, at the IEEE ICC2011, and at the IEEE/CIC ICCC2015. He was the Co-Chair of the Communication Theory Symposium of IEEE Globecom 2014. From 2011 to 2017, he was a Senior Editor of the IEEE WIRELESS COMMUNICATIONS LETTERS. From 1996 to 2011, he was the Editor of the Modulation and Detection for Wireless Systems of IEEE TRANSACTIONS ON COMMUNICATIONS. From 2007 to 2012, he has served on the Editorial Board of PHYCOM, the Journal of Physical Communications (Elsevier).