

Approximate Diagonalized Covariance Matrix for Signals with Correlated Noise

Bram Dil, Gustaf Hendeby, Fredrik Gustafsson and Bernhard Hoenders

Linköping University Post Print



N.B.: When citing this work, cite the original article.

Original Publication:

Bram Dil, Gustaf Hendeby, Fredrik Gustafsson and Bernhard Hoenders, Approximate Diagonalized Covariance Matrix for Signals with Correlated Noise, 2016, Proceedings of the 19th International Conference of Information Fusion, pages 521-527.

<http://ieeexplore.ieee.org/xpl/articleDetails.jsp?arnumber=7527932>

Copyright: ISIF <http://isif.org>

Postprint available at: Linköping University Electronic Press

<http://urn.kb.se/resolve?urn=urn:nbn:se:liu:diva-130162>

Approximate Diagonalized Covariance Matrix for Signals with Correlated Noise

Bram Dil, Gustaf Hendeby, Fredrik Gustafsson, Bernhard J. Hoenders[†]
Department of Electrical Engineering, Linköping University, Linköping, Sweden
Email: bram.j.dil@liu.se, {hendeby, fredrik}@isy.liu.se

[†] Zernike Institute for Advanced Materials, Groningen University, Groningen, The Netherlands, Email: b.j.hoenders@rug.nl

Abstract—This paper proposes a diagonal covariance matrix approximation for Wide-Sense Stationary (WSS) signals with correlated Gaussian noise. Existing signal models that incorporate correlations often require regularization of the covariance matrix, so that the covariance matrix can be inverted. The disadvantage of this approach is that matrix inversion is computationally intensive and regularization decreases precision. We use Bienaymé’s theorem to approximate the covariance matrix by a diagonal one, so that matrix inversion becomes trivial, even with non-uniform rather than only uniform sampling that was considered in earlier work. This approximation reduces the computational complexity of the estimator and estimation bound significantly. We numerically validate this approximation and compare our approach with the Maximum Likelihood Estimator (MLE) and Cramér-Rao Lower Bound (CRLB) for multivariate Gaussian distributions. Simulations with highly correlated signals show that our approach differs less than 0.1% from this MLE and CRLB when the observation time is large compared to the correlation time. Additionally, simulations show that in case of non-uniform sampling, we increase the performance in comparison to earlier work by an order of magnitude. We limit this study to correlated signals in the time domain, but the results are also applicable in the space domain.

I. INTRODUCTION

This paper proposes a diagonal covariance matrix approximation for Wide-Sense Stationary (WSS) signals sampled over a finite observation time. It allows for efficient approximations of the Maximum Likelihood Estimator (MLE) and the Cramér-Rao Lower Bound (CRLB). The CRLB provides a lower bound on the covariance of unbiased estimators of unknown deterministic parameters. The CRLB implies that the covariance of an unbiased estimator is lower bounded by the inverse of the Fisher Information [1]. Existing analyses on MLE and associated CRLB often assume independent samples. Fisher Information is additive when samples are independent. Hence, the lower bound on the covariance of unbiased estimators decreases with the number of independent samples. This has practical relevance when correlations between samples are negligibly small. Correlations between samples increase with increasing sample density over a finite observation time. [2] shows that when the sampling density increases beyond the Nyquist sampling rate, samples become correlated, and the CRLB and Fisher Information converge asymptotically to their lower and upper bounds, respectively. As in [2], we expect that our method can be used for nonlinear estimators.

We apply Bienaymé’s theorem to approximate the covariance matrix by a diagonal one, so that the computation of

MLE, CRLB and Fisher Information for correlated signals is straightforward. This approach significantly reduces the computational costs compared to computing the MLE and CRLB for multivariate Gaussian distributions. We numerically validate our approach by performing simulations with WSS correlated Gaussian noise, two different autocorrelation functions and a linear estimator. These simulations with highly correlated signals show that our approach differs less than 0.1% from the MLE and CRLB for multivariate Gaussian distributions when the correlation time is small compared to the observation time. When correlations between samples increase, the determinant of the covariance matrix decreases until it becomes singular. The multivariate Gaussian distribution becomes degenerate. This paper uses standard Tikhonov regularization [3] to solve this problem. Hence, the MLE and CRLB for multivariate Gaussian distributions become an approximation due to regularization. We do not quantify the influence of this regularization.

We apply Bienaymé’s theorem to convert an ensemble of correlated samples with a known auto-correlation function to an ensemble of uncorrelated samples. More specifically, we apply Bienaymé’s theorem to quantify the influence of correlations on the covariance of unbiased estimates. In [2], it is assumed that all individual samples are influenced equally by correlations. This approach works well when each sample decreases the covariance of the unbiased estimator with an equal amount. This is not the case with non-uniform sampling. Our new approach quantifies the influence of correlations per sample. In case of uniform sampling, simulations show that both methods provide similar results. In case of non-uniform sampling, simulations show that our new approach increases the performance in comparison to previous work by an order of magnitude.

This paper is organized as follows. Sec. II describes the motivation and general estimation framework. Sec. III describes the noise properties and correlation functions employed. Sec. IV describes our approach. Sec. V describes the results of our simulation environment. Sec. VI summarizes the conclusions.

II. MOTIVATION AND PROBLEM FORMULATION

The motivation for this contribution stems from a localization problem in radio networks, where spatial correlations violate the common assumption of independent samples [2].

The problem formulation, however, is broader than this and includes also temporal correlations of the noise. The problem fits into the Gaussian process formulation, which is a well studied area in the machine learning community [4].

A. General Estimation Framework

To formulate a general estimation framework, consider the signal model

$$\mathbf{y} = \mathbf{s}(\boldsymbol{\theta}) + \mathbf{e}, \quad (1)$$

with probability density function

$$p(\mathbf{y}; \boldsymbol{\theta}) = \frac{1}{\sqrt{(2\pi)^n |\mathbf{C}|}} \exp\left(-\frac{1}{2}(\mathbf{y} - \mathbf{s}(\boldsymbol{\theta}))^T \mathbf{C}^{-1}(\mathbf{y} - \mathbf{s}(\boldsymbol{\theta}))\right), \quad (2)$$

where \mathbf{y} denotes the vector of samples

$$\mathbf{y} = (y(t_1) \quad \dots \quad y(t_n)), \quad (3)$$

where $\mathbf{s}(\boldsymbol{\theta})$ denotes the signals

$$\mathbf{s}(\boldsymbol{\theta}) = (s(t_1; \boldsymbol{\theta}) \quad \dots \quad s(t_n; \boldsymbol{\theta})), \quad (4)$$

where \mathbf{e} denotes noise

$$\mathbf{e} = (e(t_1) \quad \dots \quad e(t_n)), \quad (5)$$

and where t denotes time, $|\cdot|$ denotes the determinant, and \mathbf{C} denotes the covariance matrix. Sec. III describes the noise properties and individual elements of the covariance matrix. We assume that \mathbf{C} is constant, i.e. independent of $\boldsymbol{\theta}$, and known, so that the least-squares expression reduces to

$$\hat{\boldsymbol{\theta}} = \arg \min_{\boldsymbol{\theta}} (\mathbf{y} - \mathbf{s}(\boldsymbol{\theta}))^T \mathbf{C}^{-1}(\mathbf{y} - \mathbf{s}(\boldsymbol{\theta})), \quad (6)$$

where $\boldsymbol{\theta}$ denotes the set of unknown parameters. With a Gaussian noise assumption, the least-squares solution coincides with the Maximum Likelihood Estimator (MLE).

The MLE expressed by (6) requires inversion of covariance matrix \mathbf{C} . For uniform sampling, the covariance matrix is a Toeplitz matrix, for which fairly efficient matrix-inversion algorithms exist. For non-uniform sampling, the Toeplitz structure is lost. In any case, the complexity increases almost cubically with the number of samples. This is the main problem we want to tackle:

Our goal is to approximate the covariance matrix \mathbf{C} by a diagonal covariance matrix $\hat{\mathbf{C}}$, so that matrix inversion is reduced to inversion of each diagonal element.

The computation of $\hat{\mathbf{C}}^{-1}$ does not require matrix inversion of \mathbf{C} .

In addition, the multivariate Gaussian distribution only exists when the covariance matrix \mathbf{C} has full rank. When the correlations between samples increase, the covariance matrix becomes singular, at least numerically. In that case, the multivariate Gaussian distribution can be calculated by regularization of the covariance matrix. We use standard L_2 Tikhonov regularization [3] and apply the standard invert function in Matlab [5]. Hence, with regularization, the multivariate Gaussian distribution becomes an approximation with increasing correlations.

B. Cramér-Rao Lower Bound

The CRLB analysis provides a lower bound on the spreads of unbiased estimators of unknown deterministic parameters. This lower bound implies that the covariance of any unbiased estimator, $\text{cov}(\hat{\boldsymbol{\theta}})$, is bounded from below by the inverse of the Fisher Information Matrix (FIM) \mathbf{F} as given by [1]

$$\text{cov}(\hat{\boldsymbol{\theta}}) \geq \mathbf{F}^{-1}(\boldsymbol{\theta}). \quad (7)$$

In (7), The Fisher matrix is evaluated at the true parameters $\boldsymbol{\theta}$, while $\hat{\boldsymbol{\theta}}$ denotes any estimator of these parameters. The elements of the Fisher Information Matrix for multivariate Gaussian distributions are given by [1, §3.9]

$$[\mathbf{F}(\boldsymbol{\theta})]_{a,b} = \left[\frac{\partial \mathbf{s}}{\partial \boldsymbol{\theta}_a} \right]^T \mathbf{C}^{-1} \left[\frac{\partial \mathbf{s}}{\partial \boldsymbol{\theta}_b} \right] + \frac{1}{2} \left[\mathbf{C}^{-1} \frac{\partial \mathbf{C}}{\partial \boldsymbol{\theta}_a} \mathbf{C}^{-1} \frac{\partial \mathbf{C}}{\partial \boldsymbol{\theta}_b} \right]. \quad (8)$$

Here, a and b are elements of the set of unknown parameters $\boldsymbol{\theta}$. We assume that the covariance matrix is independent of unknown parameters $\boldsymbol{\theta}$, so that the second term equals zero. The Fisher Information Matrix expressed by (8) contains n^2 terms and also requires the covariance matrix \mathbf{C} to be inverted. Our approximated diagonal covariance matrix $\hat{\mathbf{C}}$ reduces the number of terms to n and significantly simplifies the matrix inversion. In this case, the inverse of the diagonal elements in the covariance matrix determine the weights.

III. NOISE PROPERTIES AND CORRELATION FUNCTIONS

This section defines the properties of WSS Gaussian noise and describes auto-correlation functions commonly employed in the literature. The measurement noise in (1) is assumed to be a WSS Gaussian process with power spectral density $|E(f)|^2$. Here, f denotes frequency in Hertz, $|\cdot|^2$ denotes the absolute squared value, $E(f)$ denotes the Fourier transform of $e(t)$. The WSS property implies that the mean exists and is independent of t

$$\mathbb{E}[e(t)] = \mu_e, \quad (9)$$

and that the first-order auto-correlation function depends only on the time difference τ

$$\Gamma(\tau) = \mathbb{E}[e(t)e^*(t+\tau)], \quad (10)$$

where $*$ denotes complex conjugate. The auto-covariance function is related to the auto-correlation by

$$C(\tau) = \Gamma(\tau) - \mu_e^2. \quad (11)$$

For simplicity, we assume that the noise has zero mean $\mu_e = 0$, so that $C(\tau) = \Gamma(\tau)$. For WSS processes, the Wiener-Khinchin theorem states that the power spectral density $|E(f)|^2$ and the auto-correlation function $\Gamma(\tau)$ are a Fourier transform pair [6, §3.4]

$$\Gamma(\tau) = \int_{-\infty}^{\infty} |E(f)|^2 e^{-i2\pi f\tau} df, \quad (12)$$

$$|E(f)|^2 = \int_{-\infty}^{\infty} \Gamma(\tau) e^{i2\pi f\tau} d\tau. \quad (13)$$

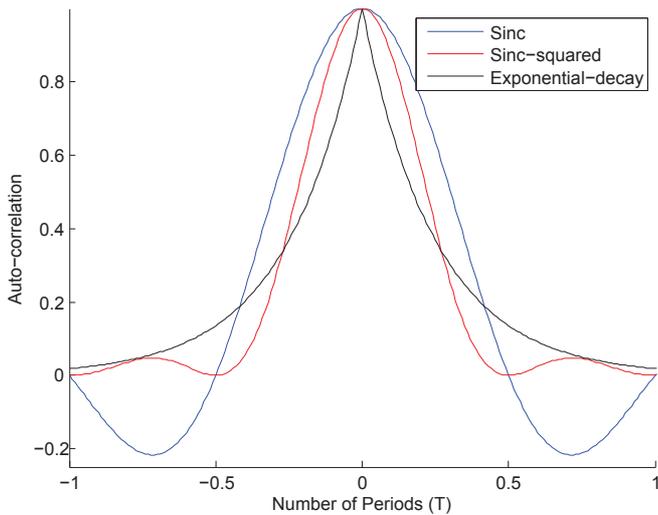


Fig. 1. Different auto-correlation functions as a function of the number of periods in time. The sinc auto-correlation function is expressed by (15). The sinc-squared auto-correlation function is expressed by (16). The exponential-decay auto-correlation function is expressed by (17).

We use the Wiener-Khinchin theorem to relate the bandwidth of the noise to the auto-correlation function of the noise.

The Nyquist-Shannon sampling theorem holds for optimal linear estimators when applied to the auto-covariance function of a bandwidth limited WSS signal [7]. As in [8, §2.4], we use the Nyquist-Shannon sampling theorem to determine the maximum number of independent samples over a finite observation time of WSS signals with a finite bandwidth. We denote the finite observation time by T_{obs} . We denote the finite bandwidth of the respective WSS signal by Δf , which can be calculated from the auto-correlation function as expressed by (13). The Nyquist-Shannon sampling rate equals $2\Delta f$, which is equal to two samples per period T . Hence, defining $T/2$ as the correlation time, the Nyquist sampling rate equals the inverse of the correlation time, which, in turn, equals the bandwidth. The maximum number of independent samples over an observation time T_{obs} equals

$$n_{\text{max}} = 2\Delta f T_{\text{obs}} = \frac{T_{\text{obs}}}{T/2}. \quad (14)$$

The next subsection describes several widely-used auto-correlation functions and their power spectral densities.

A. Auto-Correlation Functions, Bandwidth and Nyquist Sampling Rate

Fig. 1 shows the auto-correlation functions we employ in this paper. For illustrative purposes we include the sinc auto-correlation function

$$C(\tau) = \sigma^2 \text{sinc}\left(\frac{2\pi\tau}{T}\right) = \sigma^2 \frac{\sin\left(\frac{2\pi\tau}{T}\right)}{\left(\frac{2\pi\tau}{T}\right)}. \quad (15)$$

Equation (15) is used by the Nyquist-Shannon sampling theorem [9] and Whittaker-Shannon interpolation formula [8,

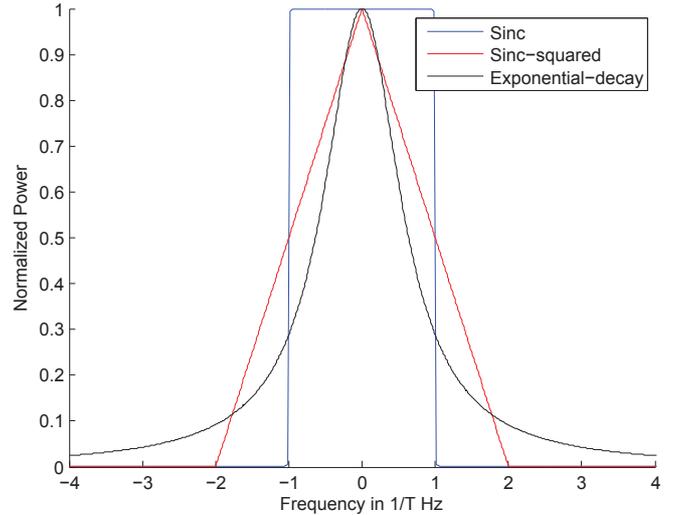


Fig. 2. Power Spectral Density (PSD) of different auto-correlation functions. The power is given as a function of frequency in Hz. The PSD functions are calculated from different auto-correlation functions using (13). The sinc auto-correlation function is expressed by (15). The sinc square auto-correlation function is expressed by (16). The exponential-decay auto-correlation function is expressed by (17).

§2.4]. The first zero crossing of the sinc-function is at half a period, which equals the inverse of the Nyquist sampling rate. When we sample the signal at the Nyquist sampling rate, all samples are uncorrelated. We are particularly interested in situations where we sample the signal faster than the Nyquist sampling rate, hence when samples are correlated.

In many situations of interest, it is the second-order rather than the first-order correlation function that is applicable to the situation [2]

$$C(\tau) = \sigma^2 \text{sinc}^2\left(\frac{2\pi\tau}{T}\right). \quad (16)$$

Fig. 1 shows that the first zero crossing of (16) is also at half a period T . As with the sinc-function, when we sample the signal with the Nyquist sampling rate, all samples are uncorrelated.

We employ a cross-covariance function used in radio localization to model correlation due to shadowing [10] with the frequency dependency proposed in [2]

$$C(\tau) = \sigma^2 \exp\left(\frac{-4\tau}{T}\right). \quad (17)$$

Fig. 1 shows that (17) roughly approaches zero over half a period T . In other words, all auto-correlation functions converge to zero over half a period. Hence, they are assumed to have the same correlation times of half a period. Fig. 2 shows the respective power spectral densities of (15), (16) and (17). Auto-correlation functions (16) and (17) have approximately twice the bandwidth of (15), although they have the same first zero crossings and thus correlation times.

IV. PROPOSED APPROXIMATION

Without loss of generality, our goal to approximate the covariance matrix \mathbf{C} for the multivariate Gaussian distribution

in (2) can be rewritten into estimating the mean in the signal model expressed by (1)

$$y(t_i) = \mu + e(t_i). \quad (18)$$

In vector formalism, this signal model can be written as a linear regression

$$\mathbf{y} = \mathbf{1}^T \mu + \mathbf{e}, \quad (19)$$

where the regression vector is defined by $\mathbf{1} = (1, 1, \dots, 1)^T$. We assume that $y(t_i)$ is a scalar. For the vector case, where y is d -dimensional, we can redefine the regression matrix as $\mathbf{1} = (\mathbf{I}_d, \mathbf{I}_d, \dots, \mathbf{I}_d)^T$. Here, \mathbf{I}_d denotes the d -dimensional identity matrix. The covariance matrix of \mathbf{e} equals \mathbf{C} . The individual elements of covariance matrix \mathbf{C} are defined by substituting (16) and (17) in (11)

$$\mathbf{C}_{i,j} = C(|t_i - t_j|) = C(\tau). \quad (20)$$

Our goal is to approximate the covariance matrix by $\hat{\mathbf{C}}$. Given the signal model in (19), the linear unbiased estimator of the average μ equals

$$\hat{\mu}^{\text{MLE}} = \frac{\mathbf{1}^T \mathbf{C}^{-1}}{\mathbf{1}^T \mathbf{C}^{-1} \mathbf{1}} \mathbf{y}. \quad (21)$$

The need to approximate \mathbf{C} by a diagonal matrix is the same as for the general problem formulation in Sec. II. We use the multivariate Gaussian distribution framework expressed by (8) to calculate the Fisher Information

$$\mathbf{F} = \mathbf{1}^T \mathbf{C}^{-1} \mathbf{1}. \quad (22)$$

This Fisher Information defines the estimation bounds. As we stated before, when correlations between samples increase, the covariance matrix becomes degenerate. Therefore, we use Tikhonov regularization to calculate the inverse of the covariance matrix.

The key in our approximation is to account for cross-correlations by increasing the covariance of the individual samples. This follows from the standard linear approach as is expressed by Bienaymé's theorem.

A. Bienaymé's Theorem

Suppose we estimate the mean as a weighted sum of the samples

$$\hat{\mu} = \sum_{i=1}^n w_i y(t_i), \quad (23)$$

then Bienaymé's theorem states that its variance equals

$$\text{cov}(\hat{\mu}) = \left(\sum_{i=1}^n w_i^2 \sigma_i^2 \right) + \left(\sum_{i=1}^n \sum_{j \neq i}^n w_i w_j \rho_{i,j} \sigma_i \sigma_j \right). \quad (24)$$

In (24), w_i denotes a weighting factor, σ_i denotes the standard deviation of sample i , and $\rho_{i,j} = \rho(|t_i - t_j|) = \rho(\tau)$ denotes the correlation coefficient between samples i and j . The first term denotes the covariance for uncorrelated samples. The second term denotes the influence of correlated samples on the covariance.

For WSS processes, the variances are equal $\sigma^2 = \sigma_i^2 = \sigma_j^2 = C(0)$. When we assume that there are no correlations and the estimator is unbiased, the weights are set equal to $w_i = w_j = \frac{1}{n}$ and (24) reduces to

$$\begin{aligned} \text{cov}(\hat{\mu}) &= \frac{\sigma^2}{n} + \frac{\sigma^2}{n^2} \sum_{i=1}^n \sum_{j \neq i}^n \rho_{i,j} \\ &= \frac{\mathbf{1}^T (\mathbf{C} \cdot \mathbf{I}) \mathbf{1}}{(\mathbf{1}^T \mathbf{1})^2} + \frac{\mathbf{1}^T (\mathbf{C} - \mathbf{C} \cdot \mathbf{I}) \mathbf{1}}{(\mathbf{1}^T \mathbf{1})^2}, \end{aligned} \quad (25)$$

where (\cdot) denotes the Hadamard product or element-by-element multiplication, \mathbf{I} denotes the identity matrix. That is, $\mathbf{C} \cdot \mathbf{I}$ extracts the diagonal. In (25), the first term denotes the covariance of the unbiased estimator when samples are uncorrelated. The second term denotes the influence of the correlations. The next two subsections distinguish between the global and local influence of correlations.

B. Global Effective Number of Samples

In [2], the influence of correlations in WSS signals is accounted for by rewriting (25) to

$$\text{cov}(\hat{\mu}) = \frac{n}{n_{\text{eff}}} \frac{\sigma^2}{n} = \frac{n}{n_{\text{eff}}} \frac{\mathbf{1}^T (\mathbf{C} \cdot \mathbf{I}) \mathbf{1}}{(\mathbf{1}^T \mathbf{1})^2}, \quad (26)$$

where

$$\frac{n}{n_{\text{eff}}} = 1 + \frac{\sum_{i=1}^n \sum_{j \neq i}^n \rho_{i,j}}{n} = \frac{\mathbf{1}^T \mathbf{C} \mathbf{1}}{\mathbf{1}^T (\mathbf{C} \cdot \mathbf{I}) \mathbf{1}} = \frac{\mathbf{1}^T \mathbf{C} \mathbf{1}}{n \sigma^2}. \quad (27)$$

In (26), the second factor denotes the covariance for uncorrelated samples. The first factor is used to account for correlated samples in the covariance for uncorrelated samples. This factor is equal to the ratio between the number of samples and the effective number of samples n_{eff} . The effective number of samples denotes the number of samples that decreases the covariance of unbiased estimators. When this ratio is approximately one, samples are uncorrelated and the sampling rate is below the Nyquist sampling rate. We use (26) to approximate the covariance matrix by

$$\hat{\mathbf{C}}^{\text{G}} \triangleq \frac{n}{n_{\text{eff}}} \sigma^2 \mathbf{I} = \frac{\mathbf{1}^T \mathbf{C} \mathbf{1}}{n} \mathbf{I}. \quad (28)$$

The diagonal elements of covariance matrix $\hat{\mathbf{C}}^{\text{G}}$ are all equal. The approximation assumes that each sample is equally correlated with other samples. This approximately holds when WSS signals are uniformly sampled and when the correlation time is small compared to the observation time. [2] successfully uses this approximation for a nonlinear estimator. The next subsection accounts for non-uniform sampling.

C. Local Effective Number of Samples

This paper extends [2] by calculating $\text{cov}(\hat{\theta})$ as a sum of n individual samples by rewriting (24) to

$$\text{cov}(\hat{\mu}) = \sum_{i=1}^n \left(w_i^2 \sigma_i^2 + \sum_{j \neq i}^n w_i w_j \rho_{i,j} \sigma_i \sigma_j \right), \quad (29)$$

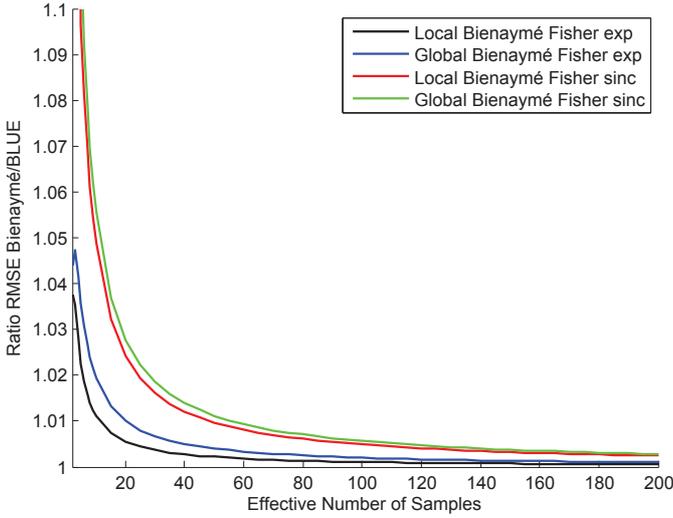


Fig. 3. Ratio between calculated estimation bounds as a function of the effective number of samples. The simulations perform 10 samples per period (T) to ensure optimal performance. The graph shows the average results of 1000 simulation runs. The performance of BLUE is given by (21) and (22). The performance of local Bienaymé is given by substituting (33) into (22) and (21). The performance of global Bienaymé is given by substituting (28) into (22) and (21).

where the first term $w_i^2 \sigma_i^2$ denotes the covariance for uncorrelated sample i . The second term denotes the influence on the covariance of the correlations between sample i and other samples. For WSS processes, $\sigma^2 = \sigma_i^2 = \sigma_j^2 = C(0)$, equal weights $w_i = w_j = \frac{1}{n}$ and unbiased estimators, (29) reduces to

$$\text{cov}(\hat{\mu}) = \sum_{i=1}^n \left(\frac{\sigma^2}{n^2} + \frac{\sigma^2}{n^2} \sum_{j \neq i} \rho_{i,j} \right) = \sum_{i=1}^n \left(\frac{1}{n^2} \mathbf{e}_i^T \mathbf{C} \mathbf{1} \right), \quad (30)$$

where \mathbf{e}_i denotes the standard basis vector. We introduce the notion of the local effective number of samples by rewriting (30) to

$$\text{cov}(\hat{\mu}) = \sum_{i=1}^n \left(\frac{n}{n_{\text{eff},i}} \right) \frac{\sigma^2}{n^2}, \quad (31)$$

where

$$\frac{n}{n_{\text{eff},i}} = 1 + \sum_{j \neq i} \rho_{i,j} = \frac{\mathbf{e}_i^T \mathbf{C} \mathbf{1}}{\sigma^2}. \quad (32)$$

Here, $n_{\text{eff},i}$ denotes the local effective number of samples. The local effective number of samples depends on correlations between sample i and other samples. When the local ratio denoted in (32) of sample i is approximately one, sample i is uncorrelated with other samples. This means that sample i is more than the inverse of the Nyquist sampling rate away, $T/2$, from other samples. When sample i is correlated to other samples, the local effective number of samples decreases as expressed by (32), and the resulting covariance increases as expressed by (31). We use (31) to approximate the covariance matrix by the following diagonal covariance matrix

$$\hat{\mathbf{C}}^L \triangleq \text{diag}(\mathbf{C} \mathbf{1}). \quad (33)$$

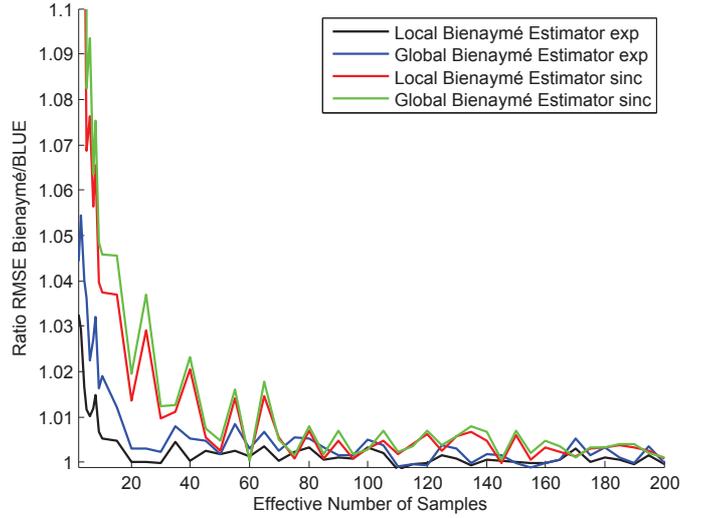


Fig. 4. Ratio between the performance of BLUE and approximations as a function of the effective number of samples. The simulations perform 10 samples per period (T) to ensure optimal performance. The graph shows the average results of 1000 simulation runs. The performance of BLUE is given by (21) and (22). The performance of local Bienaymé is given by substituting (33) into (22) and (21). The performance of global Bienaymé is given by substituting (28) into (22) and (21).

The next section evaluates numerically the performance of the global and local approximations.

V. EVALUATION AND ANALYSIS LINEAR UNBIASED ESTIMATOR

This section evaluates the simulation results of our approximation in a variety of situations for a linear unbiased estimator of an average. We set the value of the average to $\mu = 100$ and perform 1000 simulation runs to quantify performances. Our simulations employ a Gaussian distribution to simulate the correlated noise, $\mathbf{y} \sim \mathcal{N}(\mu, \mathbf{C})$. We perform simulations with auto-covariance functions (16) and (17).

A. Linear Unbiased Estimator and CRLB of Average

Our simulations employ a Gaussian distribution to simulate the correlated noise. This means that (22) defines the true Fisher Information for our simulation environment and (21) is the Best Linear Unbiased Estimator (BLUE) of the average μ . We compare the true Fisher Information and BLUE to the approximations proposed. We use (33) to approximate the covariance matrix and substitute \mathbf{C} by $\hat{\mathbf{C}}^L$ in (22) and (21) to calculate the performance of our method. We use (28) to approximate the covariance matrix and substitute \mathbf{C} by $\hat{\mathbf{C}}^G$ in (22) and (21) to calculate the performance of the method proposed in [2]. The performance is quantified by the ratio between the root mean squared error (RMSE) calculated by these methods and BLUE.

B. Performance as a Function of the Effective Number of Samples

Fig. 3 shows the ratio between the RMSE calculated by the Fisher Information of BLUE and the approximations as

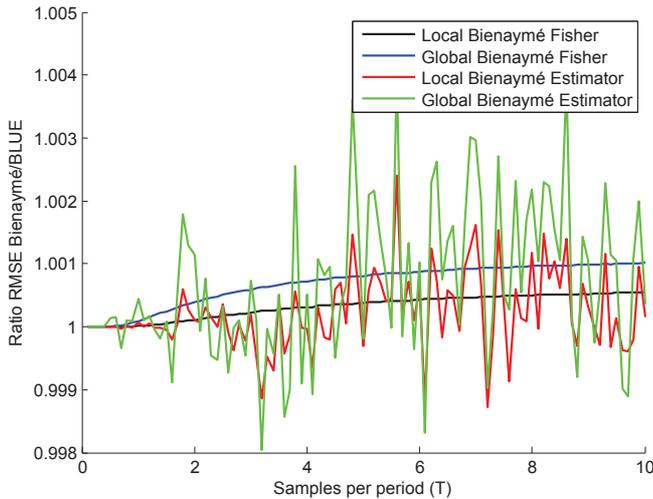


Fig. 5. Ratio between BLUE and approximations as a function of the number samples per period (T). The correlation increases with increasing number of samples per period (T). The graph shows the average results of 1000 simulation runs. The performance of BLUE is given by (21) and (22). The performance of local Bienaymé is given by substituting (33) into (22) and (21). The performance of global Bienaymé is given by substituting (28) into (22) and (21).

a function of the effective number of samples. The effective number of samples is directly related to the observation time, T_{obs} , and the correlation time of half a period, $0.5T$, as given by (14), (27) and (32). When the effective number of samples equals 2, the difference between BLUE and our approximation is the largest. For the exponentially-decaying auto-correlation function, this difference is roughly 5%. For the sinc-squared auto-correlation function, this difference equals roughly 25%. However, this difference drops rapidly below 1% with increasing effective number of samples. Fig. 4 shows similar behavior for the estimators as Fig. 3 for the Fisher Information, except that the difference between BLUE and the approximations decreases faster. Approximations (28) and (33) perform well when the observation time is large compared to the correlation time, hence when the effective number of samples is sufficiently large, $\frac{T_{obs}}{0.5T} \gg 1$.

In comparison with the sinc-squared auto-correlation function, the exponentially-decaying auto-correlation function requires less regularization due to its smoothness. This makes the results for the exponentially-decaying auto-correlation function more reliable, however less realistic [2]. We discuss the results of the sinc-squared auto-correlation function in the respective sections, however due to space constraints we do not show them in the figures.

C. Performance as a Function of Samples per Period

Fig. 5 shows the ratio between the RMSE of BLUE and our method as a function of the sampling density. The samples are simulated within an observation time of 100 periods ($T_{obs} = 100T$), which amounts to roughly an effective number of samples of 200. This is in accordance with (14). When the density of samples is below two samples per period,

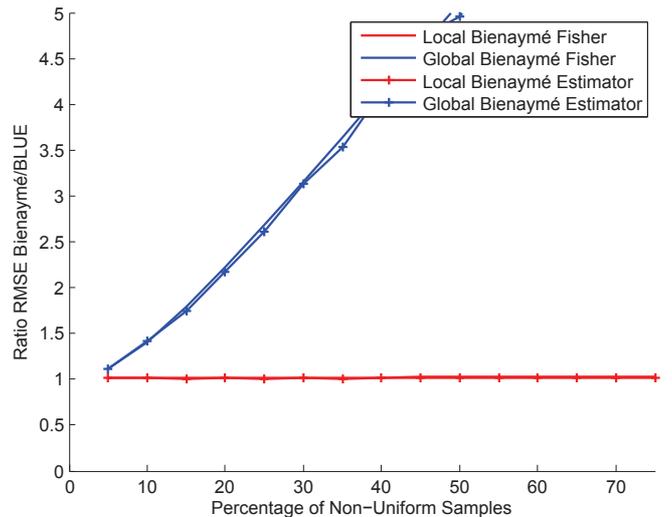


Fig. 6. Ratio between BLUE and approximations as a function of the percentage of non-uniform samples. The simulations perform 5 samples per period to ensure optimal performance. The graph shows the average results of 1000 simulation runs. The performance of BLUE is given by (21) and (22). The performance of local Bienaymé is given by substituting (33) into (22) and (21). The performance of global Bienaymé is given by substituting (28) into (22) and (21).

the samples can be considered as independent. Hence, the CRLB and MLE for multivariate Gaussian distributions provide roughly equal results as with our approximated covariance matrices. When the density of samples increases beyond two samples per period, the correlations between samples increase. In that case, the simulation results show that our approximation differs at most 0.05% and that [2] differs at most 0.1%. The results show that the difference between the estimators is too small to quantify with 1000 simulation runs. In the case of a sinc-squared auto-correlation function, the graphs show similar performance. However, the difference increases to 0.2%. Note that the increase in difference between the methods could be an artifact due to the increase of required regularization of the covariance matrix.

D. Performance as a Function of the Percentage of Non-Uniform Samples

Fig. 6 shows the ratio between the RMSE of BLUE and the approximations as a function of the percentage of non-uniform samples. Samples are simulated within an observation time of 100 periods ($T_{obs} = 100T$), which equals roughly an effective number of samples of 200. The density of samples equals 5 samples per period, which is more than sufficient according to the Nyquist sampling rate. We increase the percentage of non-uniform samples by simulating a percentage of the samples at the same time instant. Note that this percentage of samples has a correlation of one, $C(0) = \sigma^2$. Fig. 6 shows that the difference between BLUE and our local approximation is negligible. The difference between BLUE and the global approximation increases with increasing percentage of non-uniform samples.

VI. CONCLUSION

We propose a method that uses Bienaymé's theorem to approximate the covariance matrix by a diagonal one. The diagonal covariance matrix can be used to approximate the Maximum-Likelihood Estimator (MLE) and Cramér-Rao Lower Bound (CRLB) for multivariate Gaussian distributions. This approximation reduces the computational complexity of the estimator and estimation bound. We numerically validate our approach by performing simulations with WSS correlated Gaussian noise. These simulations show that our approach differs less than 0.1% from the MLE and CRLB for linear estimators when observation time is large compared to the correlation time. Our simulations show that our method works both with uniform and non-uniform sampling.

REFERENCES

- [1] S. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Prentice Hall, 1993.
- [2] B. J. Dil, F. Gustafsson, and B. J. Hoender, "Fundamental bounds on radio localization precision in the far field," Tech. Rep., 2015. [Online]. Available: <http://arxiv.org/abs/1507.07385>
- [3] A. N. Tikhonov and V. Y. Arsenin, *Solutions of Ill-posed problems*. Wiley, 1977.
- [4] C. E. Rasmussen and C. K. I. Williams, *Gaussian Processes for Machine Learning*. MIT Press, 2005.
- [5] 2016. [Online]. Available: <http://nl.mathworks.com/products/matlab/>
- [6] J. W. Goodman, *Statistical Optics*. Wiley, 2015.
- [7] A. V. Balakrishnan, "A note on the sampling principle for continuous signals," *IRE Trans. Inform. Theory*, vol. 3, no. 2, pp. 143–146, Jun. 1957.
- [8] J. W. Goodman, *Introduction to Fourier Optics*. Roberts and Company Publishers, 2004.
- [9] A. Papoulis, "Error analysis in sampling theory," *Proceedings of the IEEE*, vol. 54, no. 7, pp. 947–955, Jul. 1966.
- [10] M. Gudmundson, "Correlation model for shadow fading in mobile radio systems," *IET Electronics Letters*, vol. 7, p. 2145–2146, 1991.