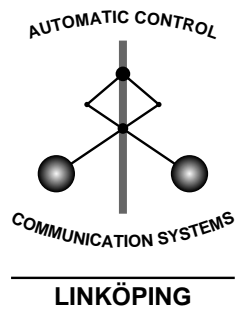


System Identification using Measurements Subject to Stochastic Time Jitter

Frida Eng, Fredrik Gustafsson

Control & Communication
Department of Electrical Engineering
Linköpings universitet, SE-581 83 Linköping, Sweden
WWW: <http://www.control.isy.liu.se>
E-mail: frida@isy.liu.se, fredrik@isy.liu.se

1st October 2004



Report no.: [LiTH-ISY-R-2632](#)

Submitted to IFAC 2005

Technical reports from the Control & Communication group in Linköping are available at <http://www.control.isy.liu.se/publications>.

Abstract

When the sensors readings are perturbed by an unknown stochastic time jitter, classical system identification algorithms based on additive amplitude perturbations will give biased estimates. We here outline the maximum likelihood procedure, for the case of both time and amplitude noise, in the frequency domain, based on the measurement DFT. The method directly applies to output error continuous time models, while a simple sinusoid in noise example is used to illustrate the bias removal of the proposed method.

Keywords: system identification, nonuniform sampling

System Identification using Measurements Subject to Stochastic Time Jitter

Frida Eng* and Fredrik Gustafsson
Department of Electrical Engineering
Linköpings universitet, SE-58183 Linköping, Sweden
Email: frida.eng@isy.liu.se

October 1, 2004

Abstract

When the sensors readings are perturbed by an unknown stochastic time jitter, classical system identification algorithms based on additive amplitude perturbations will give biased estimates. We here outline the maximum likelihood procedure, for the case of both time and amplitude noise, in the frequency domain, based on the measurement DFT. The method directly applies to output error continuous time models, while a simple sinusoid in noise example is used to illustrate the bias removal of the proposed method.

1 Introduction

Nonuniform sampling appears in, e.g., automotive applications, radar imaging, event controlled systems and time critical applications. The non uniformity in the time stamps have different characteristics, depending on the application and in many cases the sampling is almost uniform with sampling noise corrupting the exact timing. In these cases it is not obvious that the true measurement times are known, but instead the stochastic characteristics of the sampling jitter noise is known.

This work discusses the impact of unknown nonuniform jitter sampling on signal estimation. We describe the effect and use this knowledge to improve the maximum likelihood parameter estimation of a single sinusoid. The signal estimation problem is introduced in Section 2 and in Section 3 the effect of noise on the transform is shown. The stochastic properties of the transform is further evaluated Section 4 and an example of parameter estimation is given in Section 5. The work is concluded and future actions are discussed in Section 6.

2 Review and Problem Formulation

The signal estimation problem under consideration can be stated as estimating the signal $s(t; \theta)$ based on discrete time measurements of the signal corrupted

*Mrs. Eng previously published as Miss Gunnarsson

by time and amplitude noise:

$$y_k = s(kT + \tau_k; \theta) + v(kT; \theta). \quad (1)$$

where τ_k is a sequence of independent stochastic variables with probability density function (pdf) $f_T(\tau)$. Both the signal and noise models are parameterized in the unknown parameter vector θ . A system identification example of a signal model is a continuous time output error model

$$s(t; \theta) = g(t; \theta) * u(t) \quad (2)$$

$$S(f; \theta) = G(i2\pi f; \theta)U(f), \quad (3)$$

in the time and frequency domain, respectively.

Signal estimation in the time domain, *e.g.*, the least squares (LS) solution

$$\hat{\theta}^{LS} = \arg \min_{\theta} \sum_{k=1}^N (y_k - s(kT; \theta))^2, \quad (4)$$

is less attractive for our case with time noise. For that reason, we focus on frequency domain methods where the signal model is directly expressed as a continuous time model $S(f; \theta)$. The basic idea is to derive the mean $\mu_Y(f; \theta)$ and covariance $C_Y(f; \theta)$ of the measurement DFT $Y(f)$ and use the central limit theorem to motivate that the DFT is asymptotically Gaussian (AsN) distributed,

$$Y(f) \in \text{AsN}(\mu_Y(f; \theta), C_Y(f; \theta)). \quad (5)$$

The asymptotic ML estimate is given by

$$\hat{\theta}^{ML} = \arg \min_{\theta} \sum_f \frac{(Y(f) - \mu_Y(f; \theta))^2}{C_Y(f; \theta)} + \ln(C_Y(f; \theta)), \quad (6)$$

where the sum is preferably over the DFT frequencies $f = n/NT$, or a subset thereof. The minimization is not explicit, so a numeric search algorithm is needed. See Gillberg and Ljung (2005) for a thorough discussion on these issues.

For the well studied special case of amplitude noise only, we have

$$\mu_Y(f; \theta) = S * \Gamma_N(f) \quad (7)$$

$$C_Y(f; \theta) = \Phi_{vv}(f; \theta), \quad (8)$$

where $\Phi_{vv}(f; \theta)$ is the noise spectra and $*$ denotes convolution. The frequency window $\Gamma_N(f)$ includes

- *Leakage*: local behavior around f depending on the number of samples.
- *Aliasing*: $\Gamma_N(f) = \Gamma_N(f + k/T)$ is a periodic function summing up all signal part frequencies being multiples of the sampling frequency $1/T$ according to Poisson's summation formula. That is, the continuous time OE model (3) is summed up using Poisson's summation formula implicitly in this formulation.

These well-known facts have to be modified when time noise is introduced. Aliasing still occurs, but higher frequencies are damped and distorted, and the leakage effect becomes frequency dependent.

Frequency transforms of nonuniformly sampled signals, with known $t_k = t_{k-1} + \tau_k$, was treated earlier in Gunnarsson (2003), where a lot of different transform approximations were investigated. In Gunnarsson et al. (2004), stochastic sampling was introduced into the problem and initial investigations of the stochastic window, W , were made. Basic stochastic properties of this window are further described in Eng and Gustafsson (2005).

This work will concentrate on describing the properties of the window, when the DFT of the sequence y_k is produced, and using this knowledge to perform the ML estimate. The approach is useful when the sampling times are unknown jittered, $t_k = kT + \tau_k$.

Nonuniform sampling, both deterministic and stochastic, is described in Bilinskis and Mikelsons (1992) and Marvasti (2001). The Dirichlet transform is discussed as the accurate way of performing frequency transform of nonuniform time samples in Bilinskis and Mikelsons (1992). The common factor for these works is the complete knowledge of the time stamps, t_k , which make the results unuseful for this presentation.

System identification in the frequency domain has also been treated before, for uniform sampling in the time domain, both theoretical aspects Gillberg and Ljung (2005) and identification in an automotive application Gillberg and Gustafsson (2005). These publications address the problem from a different angle, but are also aiming at frequency domain identification based on nonuniform time domain sampling.

3 Frequency Transform with Noise

To clearly show the impact of the two noise models, we describe the transform for the combined sampling model

$$y_k = z_k + v_k \tag{9}$$

$$z_k = s(kT + \tau_k) \tag{10}$$

$$s(t) = \int S(f) e^{i2\pi ft} df \tag{11}$$

where the measurement noise, v_k , is white and Gaussian with $E[v_k v_l] = \sigma^2 \delta(k - l)$, and the sampling noise, τ_k , is limited to the interval $[-T/2, T/2]$ and $E[\tau_k] = 0$.

This section describes the effect of the noises on the Discrete Fourier Transform (DFT). The stochastics of the transform will differ for the two noise models.

The DFT of the sequence y_k is

$$Y(f) = \frac{1}{N} \sum_{k=0}^{N-1} y_k e^{-i2\pi f k T} \quad (12)$$

$$= \frac{1}{N} \sum_{k=0}^{N-1} (z_k + v_k) e^{-i2\pi f k T} \quad (13)$$

$$= Z(f) + \underbrace{\frac{1}{N} \sum_{k=0}^{N-1} v_k e^{-i2\pi f k T}}_{\hat{V}(f)}, \quad (14)$$

where the sampling noise part becomes

$$Z(f) = \frac{1}{N} \sum_{k=0}^{N-1} z_k e^{-i2\pi f k T} \quad (15)$$

$$= \frac{1}{N} \sum_{k=0}^{N-1} \int S(\varphi) e^{i2\pi\varphi(kT+\tau_k)} d\varphi e^{-i2\pi f k T} \quad (16)$$

$$= \int S(\varphi) \underbrace{\frac{1}{N} \sum_{k=0}^{N-1} e^{i2\pi(\varphi-f)kT} e^{i2\pi\varphi\tau_k}}_{\triangleq W(f,\varphi)} d\varphi. \quad (17)$$

Note that this continuous time frequency domain approach is perhaps the only way to explicitly separating the signal and time errors. The assumption of Gaussian noise together with the fact that the sampling noise are independent identically distributed random variables gives the following Lemma:

LEMMA 1 *The Discrete Fourier Transform of the sequence y_k in Eq. (9) is distributed as*

$$Y \in \text{AsN}(E[Z], \text{Cov}(Z) + \text{Cov}(\hat{V})).$$

Proof: We note that τ_k and v_k are independent, which makes Y a sum of independent stochastic variables. Also, $E[v_k] = 0$, which together with the central limit theorem and Eqs. (14) and (17) concludes the proof. ■

4 First Moments with Sample Noise

The moments of the transform in Lemma 1 is given as

$$\begin{aligned} E[Y(f)] &= E[Z(f)] \\ &= \int S(\varphi) E[W(f, \varphi)] d\varphi \end{aligned} \quad (18)$$

and

$$\text{Cov}(Z(f), Z(g)) = \int S(\varphi) \text{Cov}(W(f, \varphi), W(g, \psi)) S(\psi)^* d\varphi d\psi. \quad (19)$$

The addition to the covariance from the measurement noise is simply

$$\text{Cov}(\hat{V}(f), \hat{V}(g)) = \frac{\sigma^2}{N^2} \sum_{k=0}^{N-1} e^{-i2\pi(f-g)kT}.$$

The following Lemmas show the statistics for the stochastic window, W , needed in Eqs. (18) and (19).

LEMMA 2 (MEAN VALUE) *The mean value, with respect to τ_k , of the stochastic window, $W(f, \varphi)$, defined in Eq. (17), is*

$$\begin{aligned} E[W(f, \varphi)] &= \frac{1}{N} \sum_{k=0}^{N-1} e^{i2\pi(\varphi-f)kT} E[e^{i2\pi\varphi\tau_k}] \\ &= \gamma(-\varphi) \underbrace{\frac{1}{N} \frac{1 - e^{i2\pi(\varphi-f)NT}}{1 - e^{i2\pi(\varphi-f)T}}}_{\bar{\Gamma}_N(f-\varphi)}. \end{aligned}$$

The term $\gamma(f) = E[e^{-i2\pi f\tau_k}]$, is damping corresponding to the sampling noise, while $\bar{\Gamma}_N(f)$ is the periodic window arising from the finite sampling,

$$\bar{\Gamma}_N(f) = \frac{1}{N} e^{-i\pi f(N-1)T} \frac{\sin(\pi fNT)}{\sin(\pi fT)}.$$

also known as the Dirichlet function.

Thus, when there is no sampling noise, the expected value is the usual periodic window, $\bar{\Gamma}_N$, since, in that case $\gamma(f) = 1$. Also, in that case, $E[Y(f)] = \int S(\varphi) \bar{\Gamma}_N(f - \varphi) d\varphi$, which is equal to the DFT of $s(kT)$. Note also that $\gamma(f)$ is the moment generating function (MGF) for the distribution of τ_k , and the exact value can be found in textbooks on probability for numerous distributions.

LEMMA 3 (COVARIANCE) *The missing part for the covariance of $W(f, \varphi)$, is given as*

$$E[W(f, \varphi)W^*(g, \psi)] = \Lambda_\tau^T(\varphi, \psi) \Lambda_N(f - \varphi, g - \psi).$$

The two terms correspond to parts depending on the sampling noise, Λ_τ , and on the finite sampling, Λ_N , and they are

$$\Lambda_\tau(f, g) = \begin{pmatrix} \gamma(-f)\gamma(g) \\ \gamma(g-f) \\ \gamma(-f)\gamma(g) \end{pmatrix}$$

and

$$\Lambda_N(f, g) = \begin{pmatrix} \bar{\Sigma}_N(f, -g) \\ \frac{1}{N} \bar{\Gamma}_N(f - g) \\ \bar{\Sigma}_N(-g, f) \end{pmatrix},$$

respectively. The functions were defined in Lemma 2, except for

$$\begin{aligned}\bar{\Sigma}_N(f, g) &= \frac{1}{N^2} \sum_{k=0}^{N-1} \sum_{l=0}^{k-1} \bar{\gamma}(f)^k \bar{\gamma}(g)^l \\ &= \begin{cases} \frac{1}{N(1-\bar{\gamma}(g))} [\bar{\Gamma}_N(f) - \bar{\Gamma}_N(f+g)], & \bar{\gamma}(g) \neq 1, \\ \frac{\bar{\gamma}(f)^2 \bar{\gamma}(f)^N (-1 + N(\bar{\gamma}(f)^{-1} - \bar{\gamma}(f)^{-2})) + 1}{N^2(1-\bar{\gamma}(f))^2}, & \bar{\gamma}(f) \neq \bar{\gamma}(g) = 1, \\ \frac{N-1}{2N}, & \bar{\gamma}(f) = \bar{\gamma}(g) = 1, \end{cases}\end{aligned}\quad (20)$$

with $\bar{\gamma}(f) = e^{-i2\pi fT}$.

Proof: The calculations are fairly straightforward and can be found in Appendix A. \blacksquare

From construction we must have that

$$\text{Cov}(W) = E[WW^*] - E[W]E[W]^* = 0,$$

when there is no sampling noise.

5 Example

This section describes an example of parameter estimation for the simple signal

$$s(t; \theta) = a_0 \sin(2\pi f_0 t) \quad (21)$$

$$\theta = \begin{pmatrix} a_0 & f_0 \end{pmatrix}^T. \quad (22)$$

The transform of $s(t; \theta)$ is then

$$S(f; \theta) = \frac{a_0}{2i} (\delta(f - f_0) - \delta(f + f_0)). \quad (23)$$

From the expression of the transform it is straightforward to use Eqs. (18) and (18) to calculate the mean value and covariance in the distribution in Lemma 1.

To find the best parameter vector θ the maximum likelihood estimator will be used. The DFT is calculated for the frequencies $f = n/NT$, $n = 0, \dots, N-1$, and will be denoted Y^N . The likelihood function for both cases is then

$$l(\theta) = (Y^N - \mu_{Y^N}(\theta))^* C_{Y^N}(\theta)^{-1} (Y^N - \mu_{Y^N}(\theta)) + \ln \det(C_{Y^N}(\theta)), \quad (24)$$

where $\mu_{Y^N}(\theta)$ is the mean vector with the k th element being

$$(\mu_{Y^N}(\theta))_k = E[Y(\frac{k}{NT}) | \theta],$$

and $C_{Y^N}(\theta)$ is the covariance matrix with element n, m being

$$(C_{Y^N}(\theta))_{n,m} = \text{Cov}(Y(\frac{n}{NT}), Y(\frac{m}{NT}) | \theta).$$

Then,

$$\hat{\theta} = \arg \min_{\theta} l(\theta) \quad (25)$$

is the ML-estimation of the amplitude and frequency of the sinusoid.

The sampling of the signal, $s(t)$, is done with jitter sampling noise, $z_k = s(kT + \tau_k)$, and we will compare two ways of estimating the parameters. First we model the noise as additive and use the common frequency domain ML method to estimate the parameters. This crude approximation is compared to using the fact that the noise is jitter on the measurement times to produce the ML estimate.

Type 1 Approximate the signal samples as corrupted by measurement noise but not sample noise, $\tau_k = 0$ in Eq. (10):

$$\begin{aligned} E[Y(\frac{k}{NT})] &= \frac{a_0}{2i} (\bar{\Gamma}_N(\frac{k}{NT} - f_0) - \bar{\Gamma}_N(\frac{k}{NT} + f_0)), \\ \text{Cov}(Y(\frac{n}{NT}), Y(\frac{m}{NT})) &= \frac{\sigma^2}{N} \bar{\Gamma}_N(\frac{n-m}{NT}) \\ &= \frac{\sigma^2}{N} \sum_{r=-\infty}^{\infty} \delta(\frac{n-m}{N} - r) \end{aligned}$$

Type 2 Let the signal samples be corrupted by sample noise but not measurement noise, $v_k = 0$ in Eq. (9):

$$E[Y(f)] = \frac{a_0}{2i} (\gamma(-f_0)\bar{\Gamma}_N(f - f_0) - \gamma(f_0)\bar{\Gamma}_N(f + f_0)).$$

The covariance is given from Lemma 3 and Eqs. (23) and (19).

Plots of the likelihood function, Eq. (24), for a certain realization of τ_k , using the sampling noise model (Type 2), are given in Figure 1. The global minimum is clearly shown, but it is obvious that the likelihood function also has local minima. The contour plot shows that the minimum is not exactly centered on the true value pair. This property depends on the realization of the sampling times for this run. Several realizations show that the global minimum moves around the true value.

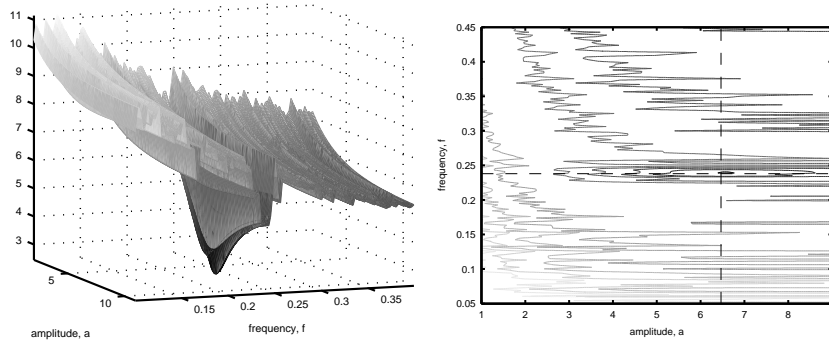


Figure 1: The maximum likelihood function, $l(\theta) = l(a, f)$, Eq. (24), for sampling noise, Type 2. The x-axis is the amplitude, and the y-axis is the frequency. The true parameters are marked with dashed lines in the contour plot on the right. The global minimum is clearly visible in the mesh-plot on the left.

To be able to show the performance of the estimates from Eq. (25), 16 Monte Carlo runs were made for 17 different frequencies and 19 different amplitudes. In Figure 2, the resulting mean errors and the standard deviations are shown. The estimation of the frequency is fairly similar for the two methods, (Type 1 and 2), while in the estimation of the amplitude, the model with measurement noise (Type 2) gives a large bias. This is easily explained with the results from Lemma 2, where it was shown that the difference in mean value, between Type 1 and 2 was an amplitude scaling with the MGF $\gamma(f)$. For both cases the standard deviation increases for larger values of the true parameters, f^0 and a^0 . Though, the standard deviation is significantly smaller for \hat{a} when using the Type 2 model.

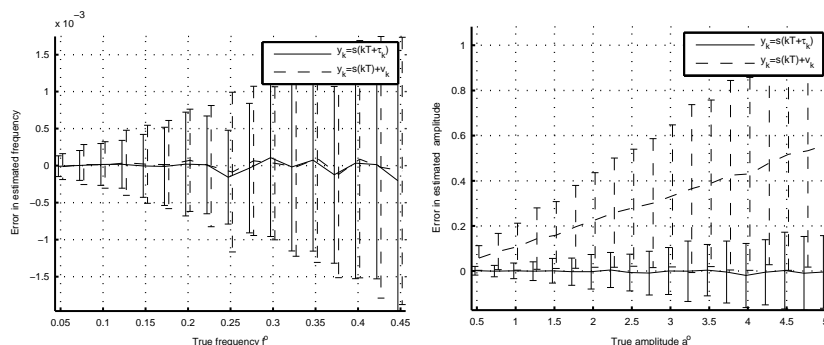


Figure 2: Mean errors and standard deviations in estimated parameters using the two different signal models, Type 1 and 2. The graphs are slightly shifted horizontally to give better visibility. The mean and standard deviation are calculated over all Monte Carlo runs as well as over all amplitudes, a^0 , or all frequencies, f^0 , respectively.

To further explain the increase in the standard deviation we study the case of sampling noise. In Figure 3, the standard deviation is depicted for \hat{f} and \hat{a} as a function of both the true frequency and the true amplitude. These plots show, that it is harder to estimate a higher frequency independent of the amplitude, and this property also affects the ability to estimate the amplitude, when a^0 is large. It is always easy to estimate zero amplitude regardless of its frequency.

6 Conclusions and future work

We have performed signal estimation in the frequency domain for nonuniformly sampled signals, where the actual sample times are perturbed from the nominal uniform values by an unknown realizations of a stochastic process. A continuous time model frequency domain maximum-likelihood approach was taken. It was shown that for estimation of amplitude and frequency in a single sinusoid, our method gives unbiased estimates and a smaller variance compared to a crude amplitude error approximation of the time errors. In future work, we will evaluate our method on output error models.

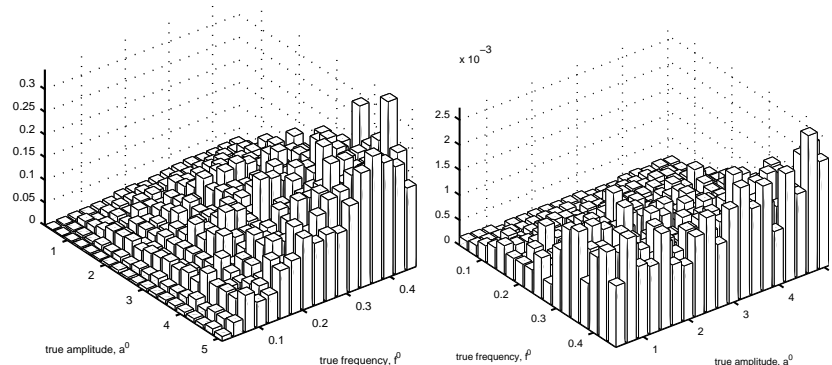


Figure 3: The standard deviation for \hat{a} (left) and \hat{f} (right), using the model with sampling noise, Type 2, as a function of both true amplitude and true frequency.

References

- Ivar Bilinskis and Arnolds Mikelsons. *Randomized Signal Processing*. Prentice Hall, London, 1992.
- Frida Eng and Fredrik Gustafsson. Frequency transforms based on nonuniform sampling — basic stochastic properties. In *Acoustics, Speech, and Signal Processing., IEEE International Conference on*, 2005. Submitted for review.
- Jonas Gillberg and Fredrik Gustafsson. Frequency-domain continuous-time AR modeling using non-uniformly sampled measurements. In *Acoustics, Speech, and Signal Processing., IEEE International Conference on*, 2005. Submitted for review.
- Jonas Gillberg and Lennart Ljung. Frequency-domain identification of continuous-time ARMA models from sampled data. In *16th IFAC World Congress*, 2005. Submitted for review.
- Frida Gunnarsson. On modeling and control of network queue dynamics. Licentiate Thesis, No. 1048, Dep. of Electrical Engineering, Linköpings Universitet, 2003.
- Frida Gunnarsson, Fredrik Gustafson, and Fredrik Gunnarsson. Frequency analysis using nonuniform sampling with applications to adaptive queue management. In *Acoustics, Speech, and Signal Processing., IEEE International Conference on*, May 2004.
- Farokh Marvasti, editor. *Nonuniform Sampling: Theory and Practice*. Kluwer Academic Publishers, 2001.

A Proof of Lemma 3

The second moment of the stochastic window, W , is by definition

$$\begin{aligned} E[W(f, \varphi)W^*(g, \psi)] &= \frac{1}{N^2} \sum_{k,l} e^{i2\pi(\varphi-f)kT} E[e^{i2\pi(\varphi\tau_k - \psi\tau_l)}] e^{-i2\pi(\psi-g)lT} \\ &= \sum_k \sum_{l < k} + \sum_k \sum_{l=k} + \sum_k \sum_{l > k} \\ &= V_1 + V_2 + V_3. \end{aligned}$$

It will be helpful to note that

$$\begin{aligned} \sum_{k=0}^{N-1} a^k \sum_{l=0}^{k-1} b^l &= \sum_{k=0}^{N-1} a^k \begin{cases} \frac{1-b^k}{1-b}, & b \neq 1 \\ k, & b = 1 \end{cases} \\ &= \begin{cases} \frac{1}{1-b} \left(\frac{1-a^N}{1-a} - \frac{1-(ab)^N}{1-ab} \right), & b \neq 1, a \neq 1, ab \neq 1 \\ \frac{1}{1-b} \left(\frac{1-a^N}{1-a} - N \right), & b \neq 1/a \neq 1 \\ \frac{1}{1-b} \left(N - \frac{1-b^N}{1-b} \right), & b \neq a = 1 \\ \frac{a^2 a^N (-1 + N(a^{-1} - a^{-2})) + 1}{(1-a)^2}, & a \neq b = 1, \\ \frac{N(N-1)}{2}, & a = b = 1 \end{cases} \end{aligned}$$

which immediately gives $\bar{\Sigma}_N(f, g)$ as in Eq. (20). The first term of the second moment above is

$$\begin{aligned} V_1(f, g, \varphi, \psi) &= \frac{1}{N^2} \gamma(-\varphi) \gamma(\psi) \sum_{k=0}^{N-1} e^{i2\pi(\varphi-f)kT} \sum_{l=0}^{k-1} e^{-i2\pi(\psi-g)lT} \\ &= \gamma(-\varphi) \gamma(\psi) \bar{\Sigma}_N(f - \varphi, \psi - g), \end{aligned}$$

since $\gamma(f) = E[e^{-i2\pi f \tau_k}]$ was the MGF of τ_k . The second term is the sum over $l = k$,

$$\begin{aligned} V_2(f, g, \varphi, \psi) &= \frac{1}{N^2} \sum_{k=0}^{N-1} e^{i2\pi(\varphi-f-\psi+g)kT} \gamma(\psi - \varphi) \\ &= \frac{1}{N} \gamma(\psi - \varphi) \bar{\Gamma}_N(-\varphi - g + \psi) \end{aligned}$$

and, finally, the third term is

$$\begin{aligned} V_3(f, g, \varphi, \psi) &= \frac{1}{N^2} \gamma(-\varphi) \gamma(\psi) \sum_{l=0}^{N-1} e^{-i2\pi(\psi-g)lT} \sum_{k=0}^{l-1} e^{i2\pi(\varphi-f)kT} \\ &= \gamma(-\varphi) \gamma(\psi) \bar{\Sigma}_N(\psi - g, f - \varphi). \end{aligned}$$

Identification of terms gives the result of the Lemma.