Supp. Material: A structural equation formulation for general quasi-periodic Gaussian processes

Unnati Nigam[1], Radhendushka Srivastava[2], Faezeh Marzbanrad[3], Michael Burke[3]

[1]IITB-Monash Research Academy.
[2]Department of Mathematics, IIT Bombay.
[3]Department of Electrical and Computer Systems Engineering, Monash University .

1 Covariance Function Approximation Error

The exact covariance function, given on the RHS of (9) of the manuscript, depends upon the covariance of the initial building block vector $\mathbf{Y}_1$. Further, the approximate covariance function, given on the RHS of (10) of the manuscript, eliminates the effect of the initial vector $\mathbf{Y}_1$. Let $s \leq t$. By using the periodicity of $\kappa_p(\cdot)$, the ratio between the exact (RHS of (9)) and approximate (RHS of (10)) covariance is given by

$$R = 1 + \omega^{2\lfloor s/p \rfloor} \left[\frac{Cov(Y_{l(t)}, Y_{l(s)})(1-\omega^2)}{\kappa_p(l(t)-l(s))} - 1 \right], \tag{S.1}$$

where $l(t), l(s) \in \{1, 2, \ldots, p\}$(see definition of $l(\cdot)$ in Theorem 1). Since $\omega \in (-1, 1)$, the second term on the RHS of (S.1) converges to 0 as $s \to \infty$. Thus, if s is large, then the ratio R can be arbitrarily close to 1. In particular if $Cov(Y_{l(t), Y_{l(s)}}) = \frac{\kappa_p(l(t)-l(s))}{1-\omega^2}$, then $R = 1$. However, if $Cov(Y_{l(t), Y_{l(s)}}) \neq \frac{\kappa_p(l(t)-l(s))}{1-\omega^2}$, then, given a specified approximation tolerance $\epsilon > 0$, $|R - 1| \leq \epsilon$ when

$$s > p \times \underbrace{\max_{l(t),l(s)\in\{1,\ldots,p\}} \left\lceil \frac{\left(\ln(\epsilon) - \ln\left|\frac{Cov\left(Y_{l(t)},Y_{l(s)}\right)(1-\omega^2)}{\kappa_p(l(t)-l(s))} - 1\right|\right)}{2\ln|\omega|} \right\rceil}_{L(\epsilon)}.$$

Note that if one burns $L(\epsilon)$ initial periodic blocks, then the ratio between the exact and approximate covariance is close to 1 with specified approximation tolerance ϵ uniformly.

Table S.1: Relative Frobenius distance between the exact and approximated covariance matrices.

n	Periodic Initialization	Cold Initialization
600	3.02×10^{-2}	1.21×10^{-1}
3000	1.34×10^{-2}	5.38×10^{-2}
10000	7.35×10^{-3}	2.94×10^{-2}

As discussed above, the exact covariance depends upon the covariance of the initial building block $\mathbf{Y}_1$ used for initialisation. We consider the following two covariance matrices of $\mathbf{Y}_1$ for numerical quantification of approximation error.

- Periodic init. ($Var(\mathbf{Y}_1) = \boldsymbol{\mathcal{K}}$); $L(\epsilon) = \left\lceil \frac{\ln(\epsilon)}{2\ln(|\omega|)} - 1 \right\rceil$
- Cold init. ($Var(\mathbf{Y}_1) = \mathbf{0}$); $L(\epsilon) = \left\lceil \frac{\ln(\epsilon)}{2\ln(|\omega|)} \right\rceil$.

Table S.1 shows the relative Frobenius distance between the exact and approximate covariance matrix under the simulation setup chosen in Section V of the manuscript, i.e., the process $\{Y_t\}_{t=1}^n$ is a QPGP with $p = 10$, $\omega = 0.5$ and κ_p as MacKay's kernel with $\theta = 1$ and $\sigma^2 = 1$. This indicates that the approximation error is very small even with a smaller sample size. The relative Frobenius distance corresponding to periodic initialization is smaller than that of cold initialization. Different QPGP initializations may lead to different degrees of approximation error, so this numerical illustration is specific to the chosen covariance structure of the initial periodic block.

2 Hyperparameter Estimation

An additional optimization step in (24) is used if necessary to enforce a parametric structure of the covariance kernel. When the within-period correlation structure is known, incorporating this information through parametric optimization improves estimation accuracy for both the parameters ω and $\mathcal{K}$. We conducted the estimation and prediction comparison in 1000 simulated samples with $p = 10$, $\omega = 0.5$, and within-period covariance as MacKay's kernel with $\sigma^2 = 1$ and $\theta = 1$. Table S.4 shows that, for the parameter ω, the MSE of estimator (24) is smaller than (20). The average (over 1000 simulation runs) Frobenius distance between $\mathcal{K}$ and the estimate of $\mathcal{K}$ obtained using (24) is also smaller than (20). The Frobenius distance is also viewed as element-wise integrated MSE of the elements of $\mathcal{K}$. This structural refinement leads to a more robust model, resulting in lower prediction MSE (Table S.4).

As discussed in subsection 4.1 of the main manuscript, the hyperparameters can also be estimated in Stage I using grid search (SIGS). Table S.4 shows that SIGS estimates have a smaller MSE (especially for $\mathcal{K}$) at a significantly higher computational cost.

3 Proof of Theoretical results

Proof of Theorem 1: For $t \in \mathbb{Z}^+$, define $i(t) \triangleq \lfloor t/p \rfloor$. Note that $i(t) \in \mathbb{Z}^+ \cup \{0\}$. Since $t = i(t)p + l(t)$, Y_t belongs to the $(i(t)+1)^{\text{th}}$ periodic block (i.e., $\mathbf{Y}_{i(t)+1}$) at its $l(t)^{\text{th}}$ coordinate. Similarly, $s = i(s)p + l(s)$and $Y_s \in \mathbf{Y}_{i(s)+1}$.

For $s \le t$, $i(t)$ and $i(s)$ must satisfy either of the following two scenarios: (a) $i(s) < i(t)$, (b) $i(s) = i(t)$ with $l(s) \le l(t)$. By using the recursive equation (7) repeatedly, we have

$$Y_t = \begin{cases} \sum_{m=0}^{i(t)-i(s)-1} \omega^m \mathsf{Z}_{i(t)+1-m,l(t)} + \omega^{i(t)-i(s)} Y_{i(s)p+l(t)} & \text{if } i(t) > i(s) \\ Y_{i(s)p+l(t)} & \text{if } i(t) = i(s) \end{cases}, \tag{S.2}$$

where $\mathbf{Z}_i \triangleq [\mathsf{Z}_{i,1}, \dots, \mathsf{Z}_{i,p}]^\top$. By using the independence of periodic block $\mathbf{Y}_{i(s)+1}$ and periodic building blocks $(\mathbf{Z}_{i(s)+2}, \mathbf{Z}_{i(s)+3}, \dots, \mathbf{Z}_{i(t)+1})$ and (S.2), we have

$$E(Y_t Y_s) = \begin{cases} \omega^{i(t)-i(s)} E(Y_{i(s)p+l(t)} Y_{i(s)p+l(s)}) & \text{if } i(t) > i(s) \\ E(Y_{i(s)p+l(t)} Y_{i(s)p+l(s)}) & \text{if } i(t) = i(s) \end{cases}. \tag{S.3}$$

Again by using (7) repeatedly, for $l \in \{1, 2, \dots, p\}$, we have

$$Y_{i(s)p+l} = \begin{cases} \sum_{m=0}^{i(s)-1} \omega^m \mathsf{Z}_{i(s)+1-m,l} + \omega^{i(s)} Y_l & \text{if } i(s) \ge 1 \\ Y_l & \text{if } i(s) = 0 \end{cases}. \tag{S.4}$$

Now by using the independence of initial periodic block $\mathbf{Y}_1$ and building-blocks $\{\mathbf{Z}_{i(s)}\}_{i(s)>1}$ and (S.4), we have

$$\begin{aligned} E(Y_{i(s)p+l(t)} Y_{j(s)p+l(s)}) = E &\left(\sum_{m=0}^{i(s)-1} \omega^m \mathsf{Z}_{i(s)-m,l(t)} \cdot \sum_{m=0}^{i(s)-1} \omega^m \mathsf{Z}_{i(s)-m,l(s)} \right) \\ &+ \omega^{2i(s)} E\left(Y_{l(t)} Y_{l(s)}\right) \end{aligned}$$

$$= \left[\frac{1-\omega^{2i(s)}}{1-\omega^2}\right] \kappa_p(l(t)-l(s)) + \omega^{2i(s)} E\left(Y_{l(t)} Y_{l(s)}\right). \tag{S.5}$$

The proof is completed by plugging in the expression on the RHS of (S.5) in that of (S.3) and using the periodicity of the covariance kernel κ_p. ■

Proof of Proposition 1: The proof follows by plugging in $E\left(Y_{l(t)}, Y_{l(s)}\right) = \frac{\kappa_p(l(t)-l(s))}{1-\omega^2}$ on the RHS of (S.5). ■

Proof of Theorem 2 (part (1)). Since $\boldsymbol{Y}_n = [Y_1, Y_2, \ldots, Y_n]$ is a QPGP vector with period p and parameters ω_0 and κ_{0p}, by using (7) and the following identity

$$\mathbf{Y}_{i+1} - \omega \mathbf{Y}_i = \mathbf{Z}_{i+1} - (\omega - \omega_0)\mathbf{Y}_i,$$

for all $i = 1, 2, \ldots, k-1$, the reduced likelihood function as defined in (14) can be expressed as follows.

$$\begin{aligned}
\tilde{\ell}_n(\omega, \boldsymbol{\mathcal{K}}) =& \log(|\boldsymbol{\mathcal{K}}|) + \frac{1}{k-1}\sum_{i=1}^{k-1} \mathbf{Z}_{k-i+1}^\top \boldsymbol{\mathcal{K}}^{-1} \mathbf{Z}_{k-i+1} \\
&- 2(\omega-\omega_0)\frac{1}{k-1}\sum_{i=1}^{k-1} \mathbf{Z}_{k-i+1}^\top \boldsymbol{\mathcal{K}}^{-1} \mathbf{Y}_{k-i} \\
&+ (\omega-\omega_0)^2 \frac{1}{k-1}\sum_{i=1}^{k-1} \mathbf{Y}_{k-i}^\top \boldsymbol{\mathcal{K}}^{-1} \mathbf{Y}_{k-i}.
\end{aligned} \tag{S.6}$$

Note that the first term on the RHS of (S.6) does not depend on n whereas the other terms depend on n via $k = n/p$. The limit of $\tilde{\ell}_n$ over $n \to \infty$ is equivalent to the convergence of the RHS of (S.6) as $k \to \infty$ since p is fixed. We establish the probability converges of the second, third and fourth term on the RHS of (S.6) by showing the convergence of mean and variances of these terms continuously.

We begin with the second term on the RHS of (S.6):

$$\begin{aligned}
&E\left[\frac{1}{k-1}\sum_{i=1}^{k-1} \mathbf{Z}_{k-i+1}^\top \boldsymbol{\mathcal{K}}^{-1} \mathbf{Z}_{k-i+1}\right] \\
=&\frac{1}{k-1}\sum_{i=1}^{k-1} E[tr(\boldsymbol{\mathcal{K}}^{-1}\mathbf{Z}_{k-i+1}\mathbf{Z}_{k-i+1}^\top)] = tr(\boldsymbol{\mathcal{K}}^{-1}\boldsymbol{\mathcal{K}}_0).
\end{aligned} \tag{S.7}$$

Using the independence of building blocks $\{\mathbf{Z}_i,\ i \geq 1\}$ and variance of random quadratic form (see equation (50) on pp. 77 for the formula of variance of random quadratic forms in [1]):

$$\mathrm{Var}\left[\frac{1}{k-1}\sum_{i=1}^{k-1} \mathbf{Z}_{k-i+1}^\top \boldsymbol{\mathcal{K}}^{-1} \mathbf{Z}_{k-i+1}\right]$$

$$=\frac{1}{(k-1)^2}\sum_{i=1}^{k-1}\mathrm{Var}(\mathbf{Z}_{k-i+1}^\top \boldsymbol{\mathcal{K}}^{-1}\mathbf{Z}_{k-i+1}) = \frac{2tr((\boldsymbol{\mathcal{K}}^{-1}\boldsymbol{\mathcal{K}}_0)^2)}{k-1}. \tag{S.8}$$

By using (S.7), the expectation of the second term on the RHS of (S.6) converges to $tr(\boldsymbol{\mathcal{K}}^{-1}\boldsymbol{\mathcal{K}}_0)$ continuously over $\boldsymbol{\mathcal{K}} \in \mathfrak{K}$. Similarly, by using (S.8), the variance of the second term on the RHS of (S.6) converges to 0 continuously over $\boldsymbol{\mathcal{K}} \in \mathfrak{K}$ as $k \to \infty$. Therefore, the second term on the RHS of (S.6) converges continuously in probability, i.e., as $k \to \infty$,

$$\frac{1}{k-1}\sum_{i=1}^{k-1}\mathbf{Z}_{k-i+1}^\top \boldsymbol{\mathcal{K}}^{-1}\mathbf{Z}_{k-i+1} \xrightarrow{P} tr(\boldsymbol{\mathcal{K}}^{-1}\boldsymbol{\mathcal{K}}_0) \tag{S.9}$$

Using a similar argument, Lemmas 1 and 2 show that the 3rd and 4th terms on the RHS of (S.6) converge continuously over $\omega \in [-1,1]$ and $\boldsymbol{\mathcal{K}} \in \mathfrak{K}$. By using Lemmas 1 and 2 as $k \to \infty$:

$$\frac{(\omega-\omega_0)}{k-1}\sum_{i=1}^{k-1}\mathbf{Z}_{k-i+1}^\top \boldsymbol{\mathcal{K}}^{-1}\mathbf{Y}_{k-i} \xrightarrow{P} 0 \tag{S.10}$$

$$\frac{(\omega-\omega_0)^2}{k-1}\sum_{i=1}^{k-1}\mathbf{Y}_{k-i}^\top \boldsymbol{\mathcal{K}}^{-1}\mathbf{Y}_{k-i} \xrightarrow{P} \frac{(\omega-\omega_0)^2}{1-\omega_0^2}tr(\boldsymbol{\mathcal{K}}^{-1}\boldsymbol{\mathcal{K}}_0). \tag{S.11}$$

Now by using (S.6), (S.9), (S.10) and (S.11), $\tilde{\ell}_n \xrightarrow{P} \tilde{\ell}$ as $n \to \infty$ continuously over $(\omega, \boldsymbol{\mathcal{K}})$. This completes the proof of part (1).

part (2): By using (25), the first order derivative of the limiting function $\tilde{\ell}$ with respect to ω and $\boldsymbol{\mathcal{K}}$ is given as follows:

$$\frac{\partial \tilde{\ell}}{\partial \omega} = \frac{2(\omega-\omega_0)}{1-\omega_0^2}tr(\boldsymbol{\mathcal{K}}^{-1}\boldsymbol{\mathcal{K}}_0) \tag{S.12}$$

$$\frac{\partial \tilde{\ell}}{\partial \boldsymbol{\mathcal{K}}} = \boldsymbol{\mathcal{K}}^{-1} - \left(1 + \frac{(\omega-\omega_0)^2}{1-\omega^2}\right)\boldsymbol{\mathcal{K}}^{-1}\boldsymbol{\mathcal{K}}_0\boldsymbol{\mathcal{K}}^{-1}. \tag{S.13}$$

Note that, by using (S.12) and (S.13), ω_0 and $\boldsymbol{\mathcal{K}}_0$ is a stationary point of $\tilde{\ell}$. Now, the second derivative of $\tilde{\ell}$ with respect to ω and $\boldsymbol{\mathcal{K}}$ is given as follows.

$$\begin{aligned}
\frac{\partial^2 \tilde{\ell}}{\partial \omega^2} &= \frac{tr(\boldsymbol{\mathcal{K}}^{-1}\boldsymbol{\mathcal{K}}_0)}{1-\omega_0^2} \\
\frac{\partial^2 \tilde{\ell}}{\partial \boldsymbol{\mathcal{K}}^2} &= -\boldsymbol{\mathcal{K}}^{-1}\otimes\boldsymbol{\mathcal{K}}^{-1} + \left(1+\frac{(\omega-\omega_0)^2}{1-\omega^2}\right) \\
&\quad \times (\boldsymbol{\mathcal{K}}^{-1}\otimes(\boldsymbol{\mathcal{K}}^{-1}\boldsymbol{\mathcal{K}}_0\boldsymbol{\mathcal{K}}^{-1}) + (\boldsymbol{\mathcal{K}}^{-1}\boldsymbol{\mathcal{K}}_0\boldsymbol{\mathcal{K}}^{-1})\otimes\boldsymbol{\mathcal{K}}^{-1}),
\end{aligned}$$

where $\otimes$ denotes the Kronecker product between matrices. We refer to [2] for differentiation formulae with respect to matrices (in particular, see equations (57), (60) and

(61) on pp. 9–10 of [2]). Further, note that the Hessian matrix of $\tilde{\ell}$ at the stationary point $(\omega_0, \boldsymbol{\mathcal{K}}_0)$ simplifies to

$$H = \begin{bmatrix} \frac{p}{1-\omega_0^2} & 0 \\ 0 & \boldsymbol{\mathcal{K}}_0^{-1} \otimes \boldsymbol{\mathcal{K}}_0^{-1} \end{bmatrix},$$

which is positive-definite. For $\omega \in [-1,1]$ and $\boldsymbol{\mathcal{K}} \in \mathfrak{K}$, define

$$\Delta(\omega, \boldsymbol{\mathcal{K}}) := \tilde{\ell}(\omega, \boldsymbol{\mathcal{K}}) - \tilde{\ell}(\omega_0, \boldsymbol{\mathcal{K}}_0),$$

which simplifies to

$$\Delta(\omega, \boldsymbol{\mathcal{K}}) = \underbrace{-\log|\boldsymbol{\mathcal{K}}^{-1}\boldsymbol{\mathcal{K}}_0| + \mathrm{tr}(\boldsymbol{\mathcal{K}}^{-1}\boldsymbol{\mathcal{K}}_0) - p}_{T_1} + \underbrace{\mathrm{tr}(\boldsymbol{\mathcal{K}}^{-1}\boldsymbol{\mathcal{K}}_0)\frac{(\omega-\omega_0)^2}{1-\omega_0^2}}_{T_2}.$$

Since $\boldsymbol{\mathcal{K}}$ and $\boldsymbol{\mathcal{K}}_0$ are positive-definite, the eigenvalues (say $\lambda_1, \ldots, \lambda_p$) of $\boldsymbol{\mathcal{K}}^{-1}\boldsymbol{\mathcal{K}}_0$ are positive. Thus, $T_1 = \sum_{i=1}^p \big(\lambda_i - \log\lambda_i - 1\big) \geq 0$ and $T_1 = 0$ if and only if $\lambda_1 = \lambda_2 = \cdots = \lambda_p = 1$ which happens if only if $\boldsymbol{\mathcal{K}} = \boldsymbol{\mathcal{K}}_0$. Further, note that $T_2 = (\sum_{i=1}^p \lambda_i)(\omega-\omega_0)^2/(1-\omega_0^2) \geq 0$ and it attains zero value if and only if $\omega = \omega_0$. Thus,

$$\Delta(\omega, \boldsymbol{\mathcal{K}}) = T_1 + T_2 \geq 0,$$

which attains the value 0 if and only if $(\omega, \boldsymbol{\mathcal{K}}) = (\omega_0, \boldsymbol{\mathcal{K}}_0)$. This completes the proof of part (2).

part (3): Since the set $[-1,1]\times\mathfrak{K}$ is a closed bounded subset of $\mathbb{R}^{p^2+1}$, the sequence of continuous functions $\tilde{\ell}_n$ are uniformly continuous. Let a non-random sequence $(\omega_n, \boldsymbol{\mathcal{K}}_n) \to (\omega, \boldsymbol{\mathcal{K}})$ as $n \to \infty$. Then, we have

$$|\tilde{\ell}_n(\omega, \boldsymbol{\mathcal{K}}) - \tilde{\ell}(\omega, \boldsymbol{\mathcal{K}})| \leq |\tilde{\ell}_n(\omega, \boldsymbol{\mathcal{K}}) - \tilde{\ell}_n(\omega_n, \boldsymbol{\mathcal{K}}_n)| + |\tilde{\ell}_n(\omega_n, \boldsymbol{\mathcal{K}}_{\boldsymbol{n}}) - \tilde{\ell}(\omega, \boldsymbol{\mathcal{K}})|. \tag{S.14}$$

By using the uniform continuity of $\tilde{\ell}_n$, we have

$$|\tilde{\ell}_n(\omega, \boldsymbol{\mathcal{K}}) - \tilde{\ell}_n(\omega_n, \boldsymbol{\mathcal{K}}_n)| \leq \sup_{(\omega,\boldsymbol{\mathcal{K}})\in[-1,1]\times\mathfrak{K}} \tilde{\ell}_n(\omega, \boldsymbol{\mathcal{K}}) \times d((\omega_n, \boldsymbol{\mathcal{K}}_n), (\omega, \boldsymbol{\mathcal{K}})), \tag{S.15}$$

where $d(\cdot,\cdot)$ is the Euclidean distance between (p^2+1) dimensional real vectors. By using part (1), and continuity of $\tilde{\ell}$, and compactness of set $[-1,1]\times\mathfrak{K}$, $\tilde{\ell}_n(\omega, \boldsymbol{\mathcal{K}})$ is uniformly bounded in probability. By using the convergence of $(\omega_n, \boldsymbol{\mathcal{K}}_n)$ and uniform boundedness of $\tilde{\ell}_n$, the first term on the RHS of (S.14) converges to 0 with probability 1 uniformly over $(\omega, \boldsymbol{\mathcal{K}}) \in [-1,1]\times\mathfrak{K}$. By using part (1) of Theorem 2, the second term on the RHS of (S.14) converges to 0 in probability as $n \to \infty$. Thus, we have

$$\sup_{(\omega,\boldsymbol{\mathcal{K}})\in[-1,1]\times\mathfrak{K}} |\tilde{\ell}_n(\omega, \boldsymbol{\mathcal{K}}) - \tilde{\ell}(\omega, \boldsymbol{\mathcal{K}})| \xrightarrow{P} 0 \text{ as } n \to \infty. \tag{S.16}$$

Since the stage I estimator $(\tilde{\omega}_n, \tilde{\boldsymbol{\mathcal{K}}}_n)$ is the minimizer of $\tilde{\ell}_n$, by using part(1) of Theorem 2, we have

$$\tilde{\ell}_n(\tilde{\omega}_n, \tilde{\boldsymbol{\mathcal{K}}}_n) \leq \tilde{\ell}_n(\omega_0, \boldsymbol{\mathcal{K}}_0) = \tilde{\ell}(\omega_0, \boldsymbol{\mathcal{K}}_0) + o_P(1) \tag{S.17}$$

Since $(\omega_0, \boldsymbol{\mathcal{K}}_0)$ is the unique minima of $\tilde{\ell}$, by using (S.17),

$$\begin{aligned}
0 &\leq \tilde{\ell}(\tilde{\omega}_n, \tilde{\boldsymbol{\mathcal{K}}}_n) - \tilde{\ell}(\omega_0, \boldsymbol{\mathcal{K}}_0) \\
&\leq \tilde{\ell}(\tilde{\omega}_n, \tilde{\boldsymbol{\mathcal{K}}}_n) - \tilde{\ell}_n(\tilde{\omega}_n, \tilde{\boldsymbol{\mathcal{K}}}_n) + o_P(1) \\
&\leq \sup_{(\omega, \boldsymbol{\mathcal{K}}) \in [-1,1] \times \mathfrak{K}} |\tilde{\ell}(\omega, \boldsymbol{\mathcal{K}}) - \tilde{\ell}_n(\omega, \boldsymbol{\mathcal{K}})| + o_P(1).
\end{aligned} \tag{S.18}$$

Now by using (S.16) and (S.18), we have

$$0 \leq \tilde{\ell}(\tilde{\omega}_n, \tilde{\boldsymbol{\mathcal{K}}}_n) - \tilde{\ell}(\omega_0, \boldsymbol{\mathcal{K}}_0) \xrightarrow{P} 0 \text{ as } n \to \infty. \tag{S.19}$$

Now, suppose the stage I estimator $(\tilde{\omega}_n, \tilde{\boldsymbol{\mathcal{K}}}_n) \stackrel{P}{\not\to} (\omega_0, \boldsymbol{\mathcal{K}}_0)$, then there exists $\epsilon > 0$ and $\delta > 0$ such that

$$P(d((\tilde{\omega}_n, \tilde{\boldsymbol{\mathcal{K}}}_n), (\omega_0, \boldsymbol{\mathcal{K}}_0)) \geq \varepsilon) > \delta \;\forall\; n. \tag{S.20}$$

Since $(\omega_0, \boldsymbol{\mathcal{K}}_0)$ is the unique minima of $\tilde{\ell}$, $\forall \epsilon > 0$, there exists $\eta > 0$ such that

$$\inf_{\{(\omega, \boldsymbol{\mathcal{K}}): d((\omega, \boldsymbol{\mathcal{K}}), (\omega_0, \boldsymbol{\mathcal{K}}_0)) \geq \epsilon\}} \tilde{\ell}(\omega, \boldsymbol{\mathcal{K}}) > \tilde{\ell}(\omega_0, \boldsymbol{\mathcal{K}}_0) + \eta. \tag{S.21}$$

Now, (S.21) and (S.20) contradicts (S.19). Thus, we have

$$(\tilde{\omega}_n, \tilde{\boldsymbol{\mathcal{K}}}_n) \xrightarrow{P} (\omega_0, \boldsymbol{\mathcal{K}}_0) \text{ as } n \to \infty. \tag{S.22}$$

We now turn to establish the convergence of stage II estimator $(\hat{\omega}, \hat{\boldsymbol{\mathcal{K}}})$ of Algorithm 1. Note that, by using (20), (S.22) and (8), for fixed p and $|t| < p$, we have, as $n \to \infty$

$$\tilde{\kappa}_p(t) \xrightarrow{P} \frac{1}{p - |t|} \sum_{j=1}^{p-|t|} \boldsymbol{\mathcal{K}}_0(j, j + |t|) = \kappa_{0p}(t). \tag{S.23}$$

By using (S.23) and given fixed p, the spectrum of $\tilde{\kappa}_p$ defined as $\tilde{f}(\lambda)$ (for $\lambda \in [-\pi, \pi]$) in (22) converges as follows.

$$\tilde{f}(\lambda) \xrightarrow{P} f_0(\lambda) \triangleq \frac{1}{2\pi} \sum_{|t|<p} \kappa_{0p}(t) e^{-it\lambda} \text{ as } n \to \infty. \tag{S.24}$$

Since $f_0(\lambda) \geq 0$ for all $\lambda \in [-\pi, \pi]$, by using (S.24), we have

$$\max(\tilde{f}(\lambda), 0) \xrightarrow{P} f_0(\lambda) \text{ as } n \to \infty. \tag{S.25}$$

Now, using (21), (22), (S.24), and (S.25), for $|t| < p$, we have

$$\hat{\kappa}_p(t) \xrightarrow{P} \int_{-\pi}^{\pi} e^{it\lambda} f_0(\lambda)\, d\lambda = \kappa_{0p}(t) \text{ as } n \to \infty. \tag{S.26}$$

We now establish the convergence of $\hat{\omega}$ proposed in (23) to complete the proof. By using (7) and (23), we have

$$\hat{\omega} = \omega_0 + \frac{tr(\hat{\boldsymbol{\mathcal{K}}}^{-1} \frac{1}{k-1} \sum_{i=1}^{k-1} \mathbf{Z}_{i+1} \mathbf{Y}_i^\top)}{tr(\hat{\boldsymbol{\mathcal{K}}}^{-1} \frac{1}{k-1} \sum_{i=1}^{k-1} \mathbf{Y}_i \mathbf{Y}_i^\top)}. \tag{S.27}$$

A similar argument to the proof of Lemma 1 and 2 gives

$$\frac{1}{k-1} \sum_{i=1}^{k-1} \mathbf{Z}_{i+1} \mathbf{Y}_i^\top \xrightarrow{P} \mathbf{0} \text{ as } n \to \infty, \tag{S.28}$$

$$\frac{1}{k-1} \sum_{i=1}^{k-1} \mathbf{Y}_i \mathbf{Y}_i^\top \xrightarrow{P} \frac{1}{1-\omega_0^2} \boldsymbol{\mathcal{K}}_0 \text{ as } n \to \infty. \tag{S.29}$$

By using (S.26), (S.27), (S.28) and (S.29), we have

$$\hat{\omega} \xrightarrow{P} \omega_0 \text{ as } n \to \infty. \tag{S.30}$$

This completes the proof. ∎

Lemma 1. *Let $\boldsymbol{Y}_n = [Y_1, \ldots, Y_n]^\top$ be a QPGP vector with period p and parameters ω_0 and κ_{p0}. Suppose $n = kp$ and $\boldsymbol{\mathcal{K}} \in \mathfrak{K}$. Then, as $k \to \infty$, we have*

$$\frac{(\omega - \omega_0)}{k-1} \sum_{i=1}^{k-1} \mathbf{Y}_i^\top \boldsymbol{\mathcal{K}}^{-1} \mathbf{Z}_{i+1} \xrightarrow{P} 0, \tag{S.31}$$

continuously over $\omega \in [-1, 1]$ and $\boldsymbol{\mathcal{K}} \in \mathfrak{K}$.

Proof of Lemma 1: Note that

$$E\left(\frac{(\omega - \omega_0)}{k-1} \sum_{i=1}^{k-1} \mathbf{Y}_i^\top \boldsymbol{\mathcal{K}}^{-1} \mathbf{Z}_{i+1}\right) = \frac{(\omega - \omega_0)}{k-1} \sum_{i=1}^{k-1} E[E(\mathbf{Y}_i^\top \boldsymbol{\mathcal{K}}^{-1} \mathbf{Z}_{i+1} | \mathbf{Y}_i)] = 0. \tag{S.32}$$

Further,

$$\mathrm{Var}\left(\frac{1}{k-1} \sum_{i=1}^{k-1} \mathbf{Y}_i^\top \boldsymbol{\mathcal{K}}^{-1} \mathbf{Z}_{i+1}\right) = \frac{1}{(k-1)^2} \sum_{i=1}^{k-1} \mathrm{Var}\left(\mathbf{Y}_i^\top \boldsymbol{\mathcal{K}}^{-1} \mathbf{Z}_{i+1}\right)$$

$$+\frac{2}{(k-1)^2}\sum_{i=2}^{k-1}\sum_{j=1}^{i-1}\mathrm{Cov}(\mathbf{Y}_i^\top\boldsymbol{\mathcal{K}}^{-1}\mathbf{Z}_{i+1},\mathbf{Y}_j^\top\boldsymbol{\mathcal{K}}^{-1}\mathbf{Z}_{j+1}) \tag{S.33}$$

By using (S.32) and Definition 1, the summands of the first term on the RHS of (S.33) simplify for $i = 1, 2, \ldots, k-1$:

$$\begin{aligned}
\mathrm{Var}[\mathbf{Y}_i^\top\boldsymbol{\mathcal{K}}^{-1}\mathbf{Z}_{i+1}] &= E(\mathbf{Y}_i^\top\boldsymbol{\mathcal{K}}^{-1}\mathbf{Z}_{i+1})^2\\
&= E[E(tr(\mathbf{Z}_{i+1}\mathbf{Z}_{i+1}^\top\boldsymbol{\mathcal{K}}^{-1}\mathbf{Y}_i\mathbf{Y}_i^\top\boldsymbol{\mathcal{K}}^{-1})|\mathbf{Y}_i)]\\
&= E[tr(\boldsymbol{\mathcal{K}}_{\mathbf{0}}\boldsymbol{\mathcal{K}}^{-1}\mathbf{Y}_i\mathbf{Y}_i^\top\boldsymbol{\mathcal{K}}^{-1}))] = \frac{1}{1-\omega_0^2}tr((\boldsymbol{\mathcal{K}}_0\boldsymbol{\mathcal{K}}^{-1})^2).
\end{aligned} \tag{S.34}$$

Similarly by using (7) and (S.32), the summands of the second term on the RHS of (S.33) is simplified for $j < i$:

$$\begin{aligned}
&\mathrm{Cov}(\mathbf{Y}_i^\top\boldsymbol{\mathcal{K}}^{-1}\mathbf{Z}_{i+1},\mathbf{Y}_j^\top\boldsymbol{\mathcal{K}}^{-1}\mathbf{Z}_{j+1})\\
&= E\left(\left(\omega_0^{i-j}\mathbf{Y}_j + \sum_{m=0}^{i-j-1}\omega_0^m\mathbf{Z}_{i-m}\right)^\top\boldsymbol{\mathcal{K}}^{-1}\mathbf{Z}_{i+1}\mathbf{Y}_j^\top\boldsymbol{\mathcal{K}}^{-1}\mathbf{Z}_{j+1}\right)\\
&= \omega_0^{i-j}E(\mathbf{Y}_j^\top\boldsymbol{\mathcal{K}}^{-1}\mathbf{Z}_{i+1}\mathbf{Y}_j^\top\boldsymbol{\mathcal{K}}^{-1}\mathbf{Z}_{j+1}) + \sum_{m=0}^{i-j-1}\omega_0^m E(\mathbf{Z}_{i-m}^\top\boldsymbol{\mathcal{K}}^{-1}\mathbf{Z}_{i+1}\mathbf{Y}_j^\top\boldsymbol{\mathcal{K}}^{-1}\mathbf{Z}_{j+1})
\end{aligned} \tag{S.35}$$

Since $j < i$ for each summand in the second term on the RHS of (S.33), by using the independence between $\mathbf{Z}_{i+1}$ and $(\mathbf{Y}_j, \mathbf{Z}_{j+1}, \ldots, \mathbf{Z}_i)$ and conditional expectation arguments both the terms on the RHS of (S.35) vanishes. Therefore, by using (S.33) and (S.34), we have

$$\mathrm{Var}\left(\frac{(\omega-\omega_0)}{k-1}\sum_{i=1}^{k-1}\mathbf{Y}_i^\top\boldsymbol{\mathcal{K}}^{-1}\mathbf{Z}_{i+1}\right) = \frac{(\omega-\omega_0)^2 tr(\boldsymbol{\mathcal{K}}_{\mathbf{0}}\boldsymbol{\mathcal{K}}^{-1})^2}{(k-1)(1-\omega_0^2)}. \tag{S.36}$$

Since the RHS of (S.36) converges to 0 continuously over $\omega \in [-1, 1]$ and $\boldsymbol{\mathcal{K}} \in \mathfrak{K}$ as $k \to \infty$, this completes the proof. ■

Lemma 2. *Let $\boldsymbol{Y}_n = [Y_1, \ldots, Y_n]^\top$ be a QPGP vector with period p and parameters ω_0 and κ_{p0}. Suppose $n = kp$ and $\boldsymbol{\mathcal{K}} \in \mathfrak{K}$. Then, as $k \to \infty$, we have*

$$\frac{(\omega-\omega_0)^2}{k-1}\sum_{i=1}^{k-1}\mathbf{Y}_i^\top\boldsymbol{\mathcal{K}}^{-1}\mathbf{Y}_i \xrightarrow{P} \frac{(\omega-\omega_0)^2 tr(\boldsymbol{\mathcal{K}}^{-1}\boldsymbol{\mathcal{K}}_0)}{1-\omega_0^2}, \tag{S.37}$$

continuously over $\omega \in [-1, 1]$ and $\boldsymbol{\mathcal{K}} \in \mathfrak{K}$.

Proof of Lemma 2: By using Proposition 1, we have

$$E\left(\frac{(\omega-\omega_0)^2}{k-1}\sum_{i=1}^{k-1}\mathbf{Y}_i^\top\boldsymbol{\mathcal{K}}^{-1}\mathbf{Y}_i\right)$$
$$=\frac{(\omega-\omega_0)^2}{k-1}\sum_{i=1}^{k-1}tr(\boldsymbol{\mathcal{K}}^{-1}E(\mathbf{Y}_i\mathbf{Y}_i^\top))=\frac{(\omega-\omega_0)^2tr(\boldsymbol{\mathcal{K}}^{-1}\boldsymbol{\mathcal{K}}_0)}{1-\omega_0^2}. \tag{S.38}$$

Note that

$$\mathrm{Var}\left(\frac{1}{k-1}\sum_{i=1}^{k-1}\mathbf{Y}_i^\top\boldsymbol{\mathcal{K}}^{-1}\mathbf{Y}_i\right)=\frac{1}{(k-1)^2}\sum_{i=1}^{k-1}\mathrm{Var}(\mathbf{Y}_i^\top\boldsymbol{\mathcal{K}}^{-1}\mathbf{Y}_i)$$
$$+\frac{2}{(k-1)^2}\sum_{i=2}^{k-1}\sum_{j=1}^{i-1}\mathrm{Cov}(\mathbf{Y}_i^\top\boldsymbol{\mathcal{K}}^{-1}\mathbf{Y}_i,\mathbf{Y}_j^\top\boldsymbol{\mathcal{K}}^{-1}\mathbf{Y}_j) \tag{S.39}$$

By using Proposition 1 and and a similar argument as used in (S.8), the variance of summands of the first term on the RHS of (S.45) is given as follows.

$$\mathrm{Var}(\mathbf{Y}_i^\top\boldsymbol{\mathcal{K}}^{-1}\mathbf{Y}_i)=\frac{2tr(\boldsymbol{\mathcal{K}}^{-1}\boldsymbol{\mathcal{K}}_0)^2}{(1-\omega_0^2)^2} \tag{S.40}$$

for $1\le i\le k-1$. We now turn to simplify the second term on the RHS of (S.45). By using Definition 1 and (7), we have

$$\mathbf{Y}_i=\omega_0^{i-j}\mathbf{Y}_j+\sum_{m=0}^{i-j-1}\omega_0^m\mathbf{Z}_{i-m}\text{ for }i>j. \tag{S.41}$$

Therefore, by using (S.41), we have

$$\mathbf{Y}_i^\top\boldsymbol{\mathcal{K}}^{-1}\mathbf{Y}_i=\omega_0^{2(i-j)}\mathbf{Y}_j^\top\boldsymbol{\mathcal{K}}^{-1}\mathbf{Y}_j+2\sum_{m=0}^{i-j-1}\omega_0^{i-j+m}\mathbf{Y}_j^\top\boldsymbol{\mathcal{K}}^{-1}\mathbf{Z}_{i-m}$$
$$+\sum_{m=0}^{i-j-1}\sum_{m'=0}^{i-j-1}\omega_0^{m+m'}\mathbf{Z}_{i-m}^\top\boldsymbol{\mathcal{K}}^{-1}\mathbf{Z}_{i-m'}. \tag{S.42}$$

By using (S.42), the summand of the second term on the RHS of (S.45) is simplified as follows for $i>j$.

$$\mathrm{Cov}(\mathbf{Y}_i^\top\boldsymbol{\mathcal{K}}^{-1}\mathbf{Y}_i,\mathbf{Y}_j^\top\boldsymbol{\mathcal{K}}^{-1}\mathbf{Y}_j)=\omega_0^{2(i-j)}\mathrm{Var}(\mathbf{Y}_j^\top\boldsymbol{\mathcal{K}}^{-1}\mathbf{Y}_j)$$
$$+2\sum_{m=0}^{i-j-1}\omega_0^{i-j+m}\mathrm{Cov}\left(\mathbf{Y}_j^\top\boldsymbol{\mathcal{K}}^{-1}\mathbf{Z}_{i-m},\mathbf{Y}_j^\top\boldsymbol{\mathcal{K}}^{-1}\mathbf{Y}_j\right)$$

$$+\sum_{m=0}^{i-j-1}\sum_{m'=0}^{i-j-1}\omega_0^{m+m'}\mathrm{Cov}\left(\mathbf{Z}_{i-m}^\top\boldsymbol{\mathcal{K}}^{-1}\mathbf{Z}_{i-m'},\mathbf{Y}_j^\top\boldsymbol{\mathcal{K}}^{-1}\mathbf{Y}_j\right) \tag{S.43}$$

By using the similar argument as in (S.40), the first term on the RHS of (S.43) is simplified as follows.

$$\omega_0^{2(i-j)}Var(\mathbf{Y}_j^\top\boldsymbol{\mathcal{K}}^{-1}\mathbf{Y}_j)=\frac{2\omega_0^{2(i-j)}}{(1-\omega_0^2)^2}tr(\boldsymbol{\mathcal{K}}^{-1}\boldsymbol{\mathcal{K}}_0)^2. \tag{S.44}$$

Since $\mathbf{Y}_j$ and $(\mathbf{Z}_{j+1},\dots,\mathbf{Z}_i)$ are independent for $j<i$, by using conditional expectation arguments, the second term on the RHS of (S.43) vanishes. By using the similar argument and the independence between $\mathbf{Y}_j$ and $\mathbf{Z}_{i-m}$ for $j<i$ and $m=0,1,\dots,i-j-1$, the third term on the RHS of (S.43) also vanishes. Therefore, using (S.45), (S.40) and (S.44):

$$\begin{aligned}
&\mathrm{Var}\left(\frac{(\omega-\omega_0)^2}{k-1}\sum_{i=1}^{k-1}\mathbf{Y}_i^\top\boldsymbol{\mathcal{K}}^{-1}\mathbf{Y}_i\right)\\
&=\frac{(\omega-\omega_0)^4}{k-1}\frac{2tr(\boldsymbol{\mathcal{K}}^{-1}\boldsymbol{\mathcal{K}}_0)^2}{(1-\omega_0^2)^2}\left(1+\frac{2}{(k-1)}\sum_{i=2}^{k-1}\sum_{j=1}^{i-1}\omega_0^{2(i-j)}\right)\\
&=\frac{(\omega-\omega_0)^4}{k-1}\frac{2tr(\boldsymbol{\mathcal{K}}^{-1}\boldsymbol{\mathcal{K}}_0)^2}{(1-\omega_0^2)^2}\left[1+\frac{2}{k-1}\frac{[(k-2)\omega_0^2-(k-1)\omega_0^4-\omega_0^{2k}]}{(1-\omega_0^2)^2}\right].
\end{aligned} \tag{S.45}$$

Since the RHS of (S.38) converges to $\frac{(\omega-\omega_0)^2tr(\boldsymbol{\mathcal{K}}^{-1}\boldsymbol{\mathcal{K}}_0)}{1-\omega_0^2}$ and the RHS of (S.45) converges to 0 as $k\to\infty$ continuously over $\omega\in[-1,1]$ and $\boldsymbol{\mathcal{K}}\in\mathfrak{K}$, this completes the proof. ■

Proof of Theorem 3: If $t\le p$, then $Y_t\in\mathbf{Y_1}$ and $\boldsymbol{Y_t}(=\mathbf{Y}_1^{(t)})$ is a zero mean Gaussian vector with covariance matrix $\boldsymbol{\mathcal{K}}_t\triangleq\frac{1}{1-\omega^2}(k_p(i-j))_{1\le i,j\le t}$. Thus, by using the conditional expectation of jointly Gaussian vectors (see (26)), we have

$$\begin{aligned}
\hat{Y}_t=E(Y_t|\boldsymbol{Y_{t-1}})&=\frac{1}{1-\omega^2}\boldsymbol{\mathcal{K}}_{1,l(t)-1}\left(\frac{1}{1-\omega^2}\boldsymbol{\mathcal{K}}_{l(t)-1}^{-1}\right)\mathbf{Y}_1^{(t-1)}\\
&=\boldsymbol{\mathcal{K}}_{1,l(t)-1}\boldsymbol{\mathcal{K}}_{l(t)-1}^{-1}\mathbf{Y}_1^{(t-1)}.
\end{aligned} \tag{S.46}$$

If $t>p$. then $i(t)\ge1$ and $Y_t=Y_{i(t)p+l(t)}$. Thus,

$$\hat{Y}_t=E(Y_t|\boldsymbol{Y_{t-1}})=E[Y_{i(t)p+l(t)}|\mathbf{Y}_{i(t)+1}^{(l(t)-1)},\mathbf{Y}_{i(t)},\dots,\mathbf{Y}_1]. \tag{S.47}$$

Recall the structural equations given in (7), we have

$$\mathbf{Y}_{i(t)+1}^{(l(t))}=\omega\mathbf{Y}_{i(t)}^{(l(t))}+\mathbf{Z}_{i(t)+1}^{(l(t))}. \tag{S.48}$$

By using (S.48), we have

$$\begin{aligned}&E[Y_{i(t)p+l(t)}|\mathbf{Y}_{i(t)+1}^{(l(t)-1)},\mathbf{Y}_{i(t)},\ldots,\mathbf{Y}_1]\\&=\omega Y_{(i(t)-1)p+l(t)}+E[\mathsf{Z}_{i(t)+1,l(t)}|\mathbf{Y}_{i(t)+1}^{(l(t)-1)},\mathbf{Y}_{i(t)},\ldots,\mathbf{Y}_1].\end{aligned}\tag{S.49}$$

By using (S.48) and the independence between $\mathbf{Z}_{i(t)+1}$ and $(\mathbf{Y}_{i(t)},\ldots,\mathbf{Y}_1)$, we have

$$\begin{aligned}&E[\mathsf{Z}_{i(t)+1,l(t)}|\mathbf{Y}_{i(t)+1}^{(l(t)-1)}=\mathbf{y}_{i(t)+1}^{(l(t)-1)},\mathbf{Y}_{i(t)}=\mathbf{y}_{i(t)},\ldots,\mathbf{Y}_1=\mathbf{y}_1]\\&=E\left[\mathsf{Z}_{i(t)+1,l(t)}\middle|\mathbf{Z}_{i(t)+1}^{(l(t)-1)}=\mathbf{y}_{i(t)+1}^{(l(t)-1)}-\omega\mathbf{Y}_{i(t)}^{(l(t)-1)},\mathbf{Y}_{i(t)}=\mathbf{y}_{i(t)},\ldots,\mathbf{Y}_1=\mathbf{y}_1\right]\\&=E\left[\mathsf{Z}_{i(t)+1,l(t)}\middle|\mathbf{Z}_{i(t)+1}^{(l(t)-1)}=\mathbf{y}_{i(t)+1}^{(l(t)-1)}-\omega\mathbf{y}_{i(t)}^{(l(t)-1)}\right]\\&=\boldsymbol{\mathcal{K}}_{1,l(t)-1}\boldsymbol{\mathcal{K}}_{l(t)-1}^{-1}\left(\mathbf{Y}_{i(t)+1}^{l(t)-1}-\omega\mathbf{Y}_{i(t)}^{l(t)-1}\right).\end{aligned}\tag{S.50}$$

By using (S.49) and (S.50), the proof of (27) is completed. Now consider the variance of $\hat{Y}_t$. If $t\le p$, then by using (27):

$$\mathrm{Var}(\hat{Y}_t)=\frac{1}{1-\omega^2}\boldsymbol{\mathcal{K}}_{1,l(t)-1}\boldsymbol{\mathcal{K}}_{l(t)-1}^{-1}\boldsymbol{\mathcal{K}}_{l(t)-1,1}.\tag{S.51}$$

For $t>p$, by using (27), (S.48), and independence between $\mathbf{Y}_{i(t)}$ and $\mathbf{Z}_{i(t)+1}$, we have

$$\mathrm{Var}(\hat{Y}_t)=\ \omega^2\mathrm{Var}\left(Y_{t-p}\right)+\mathrm{Var}\left(\boldsymbol{\mathcal{K}}_{1,l(t)-1}\boldsymbol{\mathcal{K}}_{l(t)-1}^{-1}\left(\mathbf{Y}_{i(t)+1}^{l(t)-1}-\omega\mathbf{Y}_{i(t)}^{l(t)-1}\right)\right).\tag{S.52}$$

Now by using (S.48), the second term on the RHS of (S.52) simplifies as follows.

$$\mathrm{Var}\left(\boldsymbol{\mathcal{K}}_{1,l(t)-1}\boldsymbol{\mathcal{K}}_{l(t)-1}^{-1}\left(\mathbf{Y}_{i(t)+1}^{l(t)-1}-\omega\mathbf{Y}_{i(t)}^{l(t)-1}\right)\right)=\boldsymbol{\mathcal{K}}_{1,l(t)-1}\boldsymbol{\mathcal{K}}_{l(t)-1}^{-1}\boldsymbol{\mathcal{K}}_{l(t)-1,1}.\tag{S.53}$$

This completes the proof.

4 Additional simulations and experiments

Additional estimation performance simulations are included below. The experimental setup is as described in Section 5 of the main paper, with the exception of the period, $p=100$.

Table S.3 shows the RMSE of the proposed estimators and MLE based on grid search from 1000 simulation runs, along with computational time per simulation run. The grid search space remains the same as in subsection 5.2 of the main paper. For $p=100$, we observe similar findings to the $p=10$ case in terms of accuracy as well as computational cost.

Fig. S.1 shows the boxplots of bootstrap standard errors corresponding to QPGP parameters. The width of the boxes is larger for $p=100$ in comparison to $p=10$, but we observe a similar pattern to subsection 5.3 of the main paper.

Table S.5 and Figs. S.2 and S.3 show the average lengths and coverage probability of bootstrap confidence intervals for ω and κ_p, respectively, for $p=100$. The average

Table S.2: Estimation accuracy, prediction performance, and computational cost (per iteration) for (20), (24) & SIGS.

n	Metric	(20)	(24)	SIGS
600	MSE for ω	0.0499	0.0300	0.0296
	Integrated MSE for $\boldsymbol{\mathcal{K}}$	1.0233	1.0037	0.3346
	Prediction MSE	0.1797	0.1492	0.1484
	Time (ms)	1.084	2.360	10.215
3000	MSE for ω	0.0210	0.0123	0.0122
	Integrated MSE for $\boldsymbol{\mathcal{K}}$	0.4601	0.4510	0.1386
	Prediction MSE	0.1692	0.1491	0.1489
	Time (ms)	1.954	3.031	13.641
10000	MSE for ω	0.0113	0.0067	0.0067
	Integrated MSE for $\boldsymbol{\mathcal{K}}$	0.2567	0.2518	0.0775
	Prediction MSE	0.1624	0.1491	0.1491
	Time (ms)	5.417	4.848	17.273

Table S.3: RMSE of proposed estimator and MLE based on 1000 runs, along with respective computational cost

		RMSE			
n	Estimator	ω	θ	σ^2	Time / run (ms)
600	Proposed	0.2267	0.4989	0.4481	8.26
	MLE	0.2201	0.0245	0.2734	29880.29
3000	Proposed	0.0882	0.1860	0.1992	11.83
	MLE	0.0692	0.0140	0.1309	40576.02
10000	Proposed	0.04574	0.0943	0.1057	12.11
	MLE	0.02833	0.0107	0.0847	69650.91

lengths of the bootstrap confidence intervals are larger in comparison to those for $p = 10$ due to a smaller number of blocks k. However, we observe a similar narrowing trend in the average length of CIs as n increases, as in subsection 5.4 of the main paper. We also observe that the empirical coverage probability approaches the nominal level 95%.

Below, we provide the estimates of QPGP parameters corresponding to the chosen periodic kernels for the real data case studies discussed in section 6 of the main paper. Recall, we chose the following periodic covariance kernels: (a) the general kernel, (b) MacKay's kernel, (c) the periodic Matérn kernel with $\nu = 1.5$, and (d) the cosine kernel with $\iota = 1$.

Tables S.6, S.7, and S.8 present the estimates of QPGP parameters corresponding to the selected periodic covariance kernels for the Carbon Dioxide ($p = 12$), Sunspot

Table S.4: Estimation accuracy, prediction performance, and computational cost (per iteration) for (20), (24) & SIGS.

n	Metric	(20)	(24)	SIGS
600	MSE for ω	0.0499	0.0300	0.0296
	Integrated MSE for $\boldsymbol{\mathcal{K}}$	1.0233	1.0037	0.3346
	Prediction MSE	0.1797	0.1492	0.1484
	Time (ms)	1.084	2.360	10.215
3000	MSE for ω	0.0210	0.0123	0.0122
	Integrated MSE for $\boldsymbol{\mathcal{K}}$	0.4601	0.4510	0.1386
	Prediction MSE	0.1692	0.1491	0.1489
	Time (ms)	1.954	3.031	13.641
10000	MSE for ω	0.0113	0.0067	0.0067
	Integrated MSE for $\boldsymbol{\mathcal{K}}$	0.2567	0.2518	0.0775
	Prediction MSE	0.1624	0.1491	0.1491
	Time (ms)	5.417	4.848	17.273

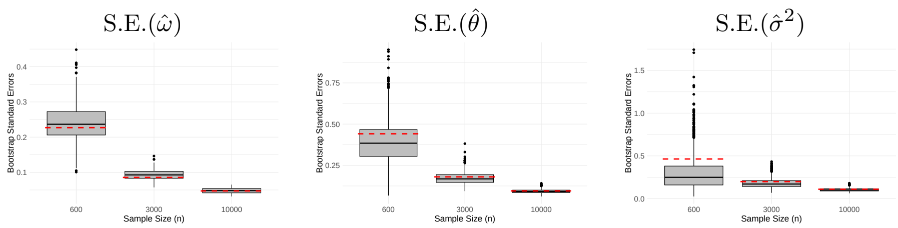

Fig. S.1: Box plots of bootstrap standard errors ($M = 1000$ bootstrap samples) of $\hat{\omega}$, $\hat{\theta}$ and $\hat{\sigma}^2$ for standard QPGP, are shown in left, center and right panel, respectively. Each panel consists of three box plots corresponding to sample sizes 600, 3000 and 10000. The empirical standard error of estimators across simulation runs are shown in dashed horizontal maroon line.

($p = 11$), and Water Level ($p = 148$) datasets, respectively, together with their bootstrap standard errors and 95% confidence intervals. Fig. S.4 shows the estimates of the general covariance kernel κ_p against lag for all three datasets, along with the associated 95% confidence limits and bootstrap standard errors.

References

[1] Searle, S.R., Gruber, M.H.J.: Linear Models, 2nd edn. Wiley Series in Probability and Statistics, p. 656. John Wiley & Sons, Inc., Hoboken, NJ, ??? (2017)

[2] Petersen, K.B., Pedersen, M.S.: The Matrix Cookbook. Technical University of Denmark. Version 20121115 (2012)

Table S.5: Average bootstrap CI length with coverage probability for ω

n	**Avg CI Length**	**Coverage Probability**
600	0.9432	0.975
3000	0.3702	0.967
10000	0.1916	0.969

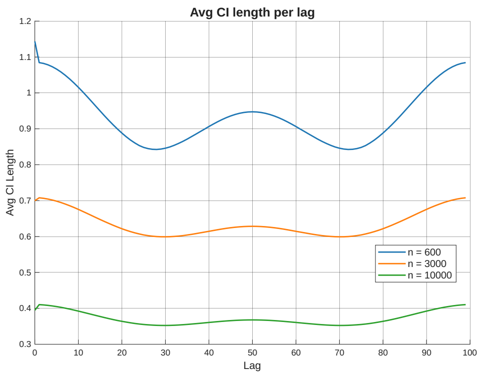

Fig. S.2: Avg CI length for κ_p.

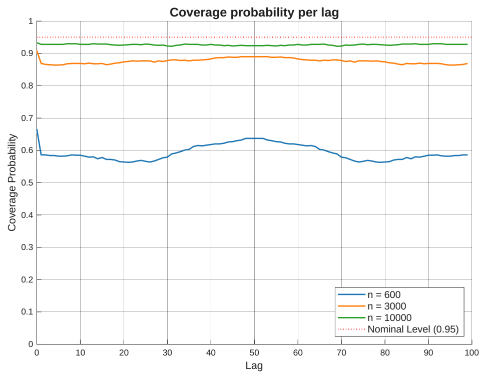

Fig. S.3: Coverage probability for κ_p.

Table S.6: Estimates of QPGP (with $p = 12$) parameters for CO_2 data for different κ_p

Kernel	Parameter	Estimate	Std. Error	Confidence Interval
General kernel	ω	0.9752	0.0085	$(0.9705, 1.0038)$
MacKay (2)	ω	0.9752	0.0085	$(0.9705, 1.0038)$
	σ^2	0.2464	0.0437	$(0.1666, 0.3329)$
	θ	0.8188	0.1251	$(0.5975, 1.0798)$
Matérn ($\nu = 1.5$) (3)	ω	0.9752	0.0085	$(0.9705, 1.0038)$
	σ^2	0.2610	0.0429	$(0.1826, 0.3464)$
	θ	1.9950	0.4053	$(1.3569, 2.9443)$
Cosine (4)	ω	0.9752	0.0085	$(0.9705, 1.0038)$
	σ^2	0.0546	0.0086	$(0.0355, 0.0687)$

Table S.7: Estimates of QPGP (with $p = 11$) parameters for Sunspot data for different κ_p

Kernel	Parameter	Estimate	Std. Error	Confidence Interval
General Kernel	ω	0.7228	0.06375	$(0.5861, 0.8429)$
MacKay (2)	ω	0.7228	0.06375	$(0.5861, 0.8429)$
	σ^2	2335.9067	519.06165	$(1337.9731, 3422.8676)$
	θ	1.8401	0.2085	$(1.6809, 2.4982)$
Matérn ($\nu = 1.5$) (3)	ω	0.7228	0.06375	$(0.5861, 0.8429)$
	σ^2	2568.1523	563.3239	$(1439.5634, 3699.3789)$
	θ	0.7599	0.0814	$(0.5436, 0.8539)$
Cosine (4)	ω	0.7228	0.06375	$(0.5861, 0.8429)$
	σ^2	1254.7285	323.1657	$(709.9667, 2002.5801)$

Table S.8: Estimates of QPGP (with $p = 148$) parameters for Water Level data for different κ_p

Kernel	Parameter	Estimate	Std. Error	Confidence Interval
General kernel	ω	0.9673	0.0102	$(0.9432, 0.9824)$
MacKay (2)	ω	0.9673	0.0102	$(0.9432, 0.9824)$
	σ^2	0.0334	0.0827	$(0.0217, 0.3315)$
	θ	1.7398	1.7659	$(0.8033, 2.3211)$
Matérn ($\nu = 1.5$) (3)	ω	0.9673	0.0102	$(0.9432, 0.9824)$
	σ^2	0.0358	0.0872	$(0.0237, 0.3508)$
	θ	0.8338	0.4892	$(0.5649, 2.1742)$
Cosine (4)	ω	0.9673	0.0102	$(0.9432, 0.9824)$
	σ^2	0.0183	0.0376	$(0.0119, 0.1531)$

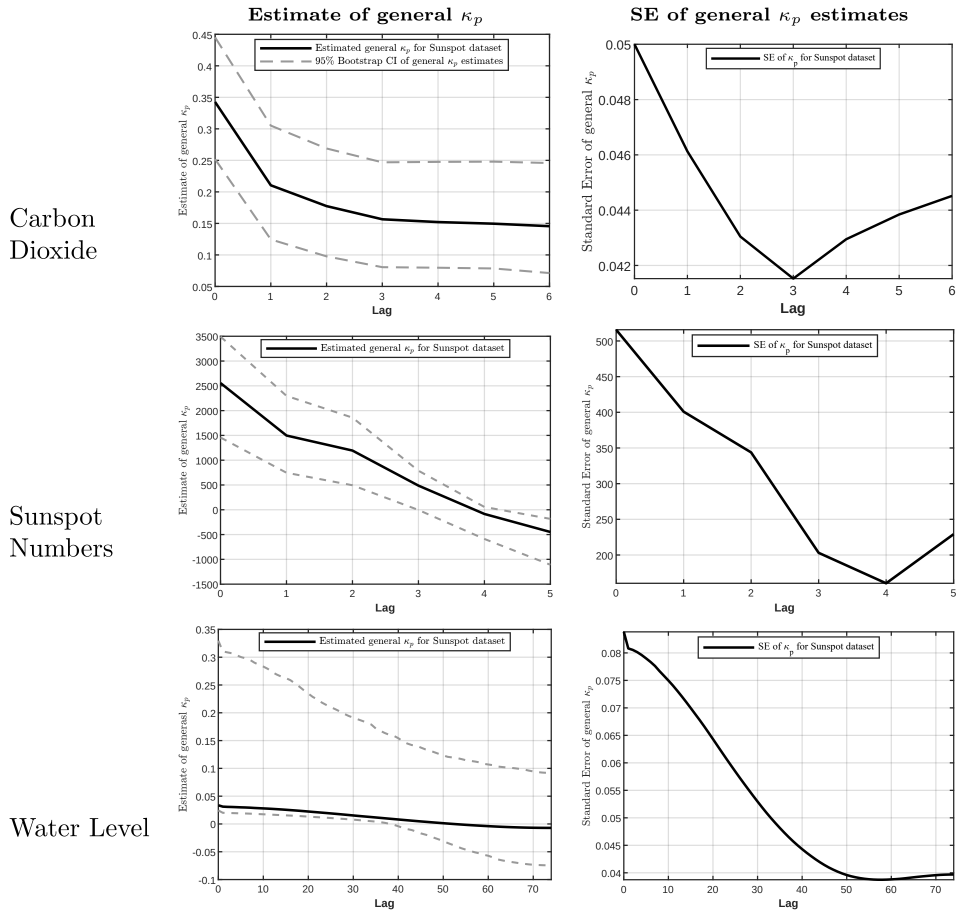

Fig. S.4: Left: estimates of general κ_p (solid black line) with 95% confidence limits (dashed black line). Right: corresponding bootstrap standard errors.