Statistica Sinica **34** (2024), 547-565
doi:https://doi.org/10.5705/ss.202020.0311

ON THE EFFICIENCY OF COMPOSITE LIKELIHOOD ESTIMATION FOR GAUSSIAN SPATIAL PROCESSES

Nelson J. Y. Chua*[1,2], Francis K. C. Hui[1] and A. H. Welsh[1]

[1] *The Australian National University and* [2] *Australian Bureau of Statistics*

Abstract: Maximum composite likelihood estimation is an attractive and commonly used alternative to standard maximum likelihood estimation that typically involves sacrificing statistical efficiency for computational efficiency. This statistical efficiency can be quantified by evaluating the sandwich information matrix of the maximum composite likelihood estimator, and then comparing it with the analogous Fisher information matrix for the maximum likelihood estimator. In this paper, we derive new closed-form expressions for the asymptotic relative efficiency of various maximum composite likelihood estimators for a one-dimensional exponential covariance Gaussian process. These expressions are based on a sampling scheme that allows for analyses under three common spatial asymptotic frameworks: expanding domain, infill, and hybrid. Our results demonstrate how the choice of composite likelihood affects the estimation efficiency and consistency, particularly for the infill and hybrid frameworks.

Key words and phrases: Block likelihood, expanding domain, full conditional likelihood, hybrid asymptotics, infill asymptotics, sandwich covariance matrix.

1. Introduction

Maximum composite likelihood estimation is an attractive and commonly used alternative to standard maximum likelihood estimation when the full likelihood is difficult to formulate and/or is computationally intractable (Besag (1974); Lindsay (1988)). Constructing a composite likelihood function involves taking the product of marginal or conditional densities that are individually simpler and collectively quicker to evaluate than the full likelihood (Varin, Reid and Firth (2011)). Although this is beneficial from a computational standpoint, there is usually a loss of statistical efficiency. This loss is often quantified by computing the sandwich covariance matrix (or, at least, an estimate of this quantity), and comparing it with the inverse Fisher information matrix (e.g., see Singhal and Kumar (2019); Xu, Reid and Xu (2024)). However, a theoretical evaluation of these matrices often requires lengthy algebraic manipulation, even for processes with relatively simple dependence structures. One such example is the Gaussian spatial process, which is our focus here.

Studying the asymptotic efficiency for spatial process data is further compli-

*Corresponding author.

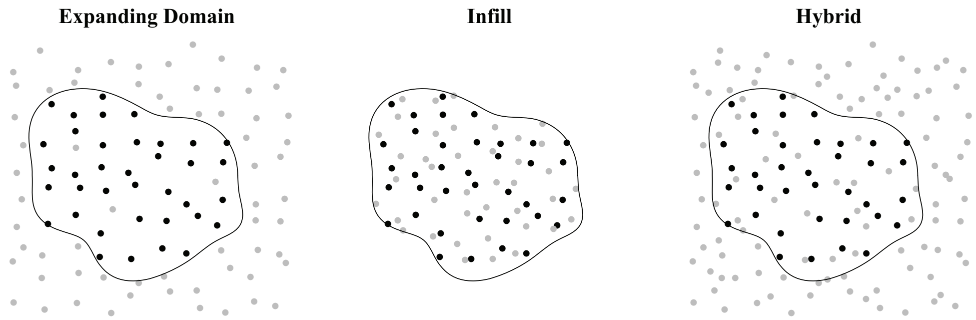

Figure 1. Comparison of asymptotic frameworks for spatial process data. The black-bordered region and black points denote the initial region of collection and the locations of observations, respectively; gray points are further observation locations under the corresponding asymptotic framework.

cated by the choice of asymptotic framework, which is most commonly categorized as expanding domain, infill (or fixed domain), or hybrid (or domain-expanding infill). These asymptotic frameworks have been formalized several times (e.g., see Zheng and Zhu (2012); Lu and Tjøstheim (2014)); their main differences are illustrated in Figure 1. Roughly speaking, under the expanding domain framework, observations are progressively collected in such a way that the domain that contains these observations is unbounded, and the minimum distance between observations is bounded below. In contrast, the spatial domain under infill is bounded and becomes infinitely dense, with observations everywhere within this domain. Finally, as the name suggests, the hybrid framework inherits the "domain unboundedness" and "infinite density" characteristics of the first two frameworks. The choice of framework affects the asymptotic properties of the estimators (Hall and Patil (1994); Zheng and Zhu (2012); Kaufman and Shaby (2013); Chang, Huang and Ing (2017)).

In this paper, we derive sandwich covariance matrices for maximum composite likelihood estimators applied to Gaussian spatial processes, in both general and specific cases. For the general case, we present an expression for the sandwich covariance matrix that unifies existing results specific to particular classes of composite likelihood functions, such as those presented in Stein, Chi and Welty (2004), Eidsvik et al. (2014), and Bevilacqua and Gaetan (2015), among others. We then focus on the widely explored one-dimensional exponential covariance Gaussian process, and develop new results on the asymptotic relative efficiency of maximum composite likelihood estimators under the expanding domain, infill, and hybrid frameworks. In particular, we consider two composite likelihood functions: a composite full conditional likelihood, which extends the work of Bachoc, Lagnoux and Nguyen (2017) under the infill framework, and a composite marginal block likelihood, as studied by Caragea and Smith (2007) under the expanding domain framework. We also demonstrate the effects of the model specification on the asymptotic behavior by comparing results based on which parameters in the exponential covariance model are assumed to be known. Our results indicate that the choice of composite likelihood and model can have major implications for the statistical efficiency and asymptotic behavior of the estimator, particularly under the infill and hybrid asymptotic frameworks.

2. Literature Review

Several general results exist pertaining to maximum likelihood asymptotics for Gaussian spatial processes under the expanding domain, infill, and hybrid frameworks. Mardia and Marshall (1984) showed that for Gaussian spatial regression models, given a few mild regularity conditions, maximum likelihood estimators are consistent and asymptotically normally distributed under the expanding domain framework. Additionally, the asymptotic variance of the maximum likelihood estimator is given by the inverse of the Fisher information matrix $\mathbf{I}(\boldsymbol{\theta}) = E[-\partial^2 \ell(\boldsymbol{\theta};\mathbf{y})/(\partial\boldsymbol{\theta}\,\partial\boldsymbol{\theta}^T)]$, where $\ell(\cdot)$ is the full log-likelihood. For Gaussian spatial autoregressive models, Zheng and Zhu (2012) derived a series of results relating to the maximum likelihood estimator. Specifically, they showed that the estimator is $\sqrt{N}$-consistent under the expanding domain framework, where N is the sample size. It is also consistent under the hybrid framework, though the rate of asymptotic convergence for the spatial covariance parameters is slower than $\sqrt{N}$. However, under infill asymptotics, the maximum likelihood estimator for the spatial covariance parameters may be inconsistent.

Arguably the most widely studied Gaussian spatial process is the two-parameter exponential covariance process, owing to its simplicity and mathematical convenience. In the case of one spatial dimension, the covariance function can be expressed as $\mathrm{cov}(y(s), y(s')) = \sigma^2 \exp(-\alpha|s - s'|)$, for $s, s' \in \mathbb{R}$, where σ^2 is the variance of $y(s)$, and α is the inverse spatial scale parameter. For equally spaced observations, Abt and Welch (1998) and Zhang and Zimmerman (2005) showed the $\sqrt{N}$-consistency of maximum likelihood estimation under the expanding domain framework and derived a simple closed-form expression for $\mathbf{I}(\sigma^2, \alpha)$. In contrast, Ying (1991) showed its inconsistency under the infill framework, attributing this to the lack of asymptotic identifiability of the components in the product $\sigma^2\alpha$ (see also Ibragimov and Rozanov (1978, p. 100)). However, under the hybrid framework, Chang, Huang and Ing (2017) showed that the maximum likelihood estimator is consistent, with a convergence rate that is dictated by the rate of growth of the spatial domain.

In contrast to maximum likelihood estimations, the large-sample results for maximum composite likelihood estimations of Gaussian spatial processes are often narrower in scope. This is largely because the asymptotics can vary depending on the choice of component densities used in the composite likelihood. Furthermore, even if a particular maximum composite likelihood estimator is shown to be consistent and asymptotically normal, it is often difficult to obtain a closed-form expression for the asymptotic variance given by the sandwich covariance matrix $\mathbf{G}(\boldsymbol{\theta})^{-1} = \mathbf{H}(\boldsymbol{\theta})^{-1}\mathbf{J}(\boldsymbol{\theta})\mathbf{H}(\boldsymbol{\theta})^{-1}$, where $\mathbf{H}(\boldsymbol{\theta}) = E[-\partial^2 c\ell(\boldsymbol{\theta};\mathbf{y})/(\partial\boldsymbol{\theta}\,\partial\boldsymbol{\theta}^T)]$ and $\mathbf{J}(\boldsymbol{\theta}) = E[\{\partial\, c\ell(\boldsymbol{\theta};\mathbf{y})/\partial\boldsymbol{\theta}\}\,\{\partial\, c\ell(\boldsymbol{\theta};\mathbf{y})/\partial\boldsymbol{\theta}\}^T]$, for a composite log-likelihood $c\ell(\cdot)$ (Kent (1982); Lindsay (1988)). For Gaussian processes, calculating $\mathbf{H}(\boldsymbol{\theta})$ requires evaluating second-order moments, whereas $\mathbf{J}(\boldsymbol{\theta})$ is considerably more complicated because it involves fourth-order moments. As such, the majority of available closed-form expressions for asymptotic relative efficiency are derived under the simple exponential covariance Gaussian process. We now review the literature for this specific, but important case.

Caragea and Smith (2007) investigated the so-called "small blocks estimator" (henceforth, the composite marginal block likelihood estimator), where observations along a line are grouped into equally sized blocks, and independence between blocks is deliberately assumed when constructing the composite likelihood function. They focused on the exponential covariance Gaussian process under the analogous equally spaced AR(1) process $y(t+1) = \phi y(t) + \epsilon(t+1)$, where $\{\epsilon(t)\}$ are independent and identically distributed (i.i.d.) normal random variables with mean zero and variance σ_ϵ^2 (thus, $\alpha = -\log\phi$ and $\sigma^2 = \sigma_\epsilon^2/(1-\phi^2)$). Under the assumption that σ_ϵ^2 is known, they derive closed-form expressions for $\mathrm{J}(\phi)$ and $\mathrm{H}(\phi)$ by expressing each $y(t)$ term in the composite likelihood in its causal time series representation. Under this time series (expanding domain) framework, they evaluate the asymptotic relative efficiency ($\lim_{N\to\infty} \mathrm{I}(\phi)^{-1}/\mathrm{G}(\phi)^{-1}$) for various choices of ϕ and block size, finding it to be around 0.9 or higher.

For the same AR(1) process, Davis and Chun (2011) studied composite marginal pairwise likelihood estimators. They showed that the traditional pairwise likelihood, which takes the product of all bivariate densities, yields an estimator that is not fully efficient for most values of ϕ. In contrast, maximizing the consecutive marginal pairwise likelihood (based on the bivariate densities of all adjacent pairs of observations) yields the standard Yule–Walker estimators for ϕ and σ_ϵ^2, and is therefore fully efficient.

Under the infill asymptotic framework, Bachoc, Bevilacqua and Velandia (2019) investigated weighted marginal and conditional pairwise likelihood functions for the exponential covariance Gaussian process. Owing to the aforementioned inconsistency of the maximum likelihood estimators of σ^2 and α under infill (Ying (1991)), they focused instead on estimating the microergodic parameter $\sigma^2\alpha$. They derived the asymptotic distributions of the resulting composite likelihood estimators, showing that full efficiency is achieved by setting weights that correspond to the consecutive marginal pairwise likelihood of Davis and Chun (2011). They also showed that the consistency of the marginal pairwise likelihood estimator is subject to restrictions on the parameter space, unlike the full likelihood and conditional pairwise likelihood estimators.

In the same framework, Bachoc, Lagnoux and Nguyen (2017) derived infill asymptotic distributions for a so-called "cross-validation estimator," obtained from a composite likelihood based on the product of all full conditional distributions; henceforth, the composite full conditional likelihood. They demonstrated the consistency and asymptotic normality of the estimator for $\sigma^2\alpha$, as well as for σ^2 and α separately when the other parameter is fixed. Interestingly, their results show that the asymptotic variance of these estimators depends on the construction of the infill sampling scheme, in contrast to the results of Ying (1991) for the full likelihood.

The primary contribution of this study is the derivation of sandwich covariance matrices that enable a unified analysis of the efficiency of composite likelihood functions under the expanding domain, infill, and hybrid frameworks. In particular, we expand existing results for the exponential covariance Gaussian process, especially those of Bachoc, Lagnoux and Nguyen (2017) for the composite full conditional likelihood, and those of Caragea and Smith (2007) for the composite marginal block likelihood.

3. Sandwich Covariance Matrix for Gaussian Spatial Processes

Let $\{y(\mathbf{s}), \mathbf{s} \in \mathcal{S} \subseteq \mathbb{R}^r\}$ be a zero-mean, r-dimensional Gaussian spatial process with covariance function $\text{cov}(y(\mathbf{s}), y(\mathbf{s}')) \equiv \Sigma(\mathbf{s}, \mathbf{s}'; \boldsymbol{\theta})$ and parameter vector $\boldsymbol{\theta}$. Additionally, consider a finite subset of locations $\{\mathbf{s}_i \in \mathcal{S}, i = 1, \ldots, N\}$, from which observations are taken to form the data vector $\mathbf{y} = [y(\mathbf{s}_1), y(\mathbf{s}_2), \ldots, y(\mathbf{s}_N)]^T$. Note that any expressions for the sandwich covariance matrix on $\boldsymbol{\theta}$ here also apply to a Gaussian process with a nonzero mean function, provided that this function is independent of $\boldsymbol{\theta}$.

The general definition of a composite likelihood function is quite broad and can include component densities based on sums and differences of random variables, among other possibilities (e.g., see Varin, Reid and Firth (2011), Bai, Song and Raghunathan (2012) and Bevilacqua and Gaetan (2015)). Here, we restrict our attention to the common case in which the component densities involve only vectors whose entries are a subset of the entries in the data vector $\mathbf{y}$. We can then write composite marginal log-likelihood functions in the form $c\ell(\boldsymbol{\theta}; \mathbf{y}) = \sum_{m=1}^{M} w_m \log f(\mathbf{y}_m; \boldsymbol{\theta})$, and composite conditional log-likelihood functions as $c\ell(\boldsymbol{\theta}; \mathbf{y}) = \sum_{m=1}^{M} w_m \log f(\mathbf{y}_m^1 | \mathbf{y}_m^2; \boldsymbol{\theta})$, for a set of positive weights w_m and M component densities on data subsets $\mathbf{y}_m$. Note that because $f(\mathbf{y}_m^1 | \mathbf{y}_m^2; \boldsymbol{\theta}) = f(\mathbf{y}_m^1, \mathbf{y}_m^2; \boldsymbol{\theta}) / f(\mathbf{y}_m^2; \boldsymbol{\theta})$, composite conditional log-likelihood functions may also be written in the form of the former expression by allowing the weights to be negative. This result suggests that we can analyze composite conditional likelihood functions using similar methods to those for composite marginal likelihood functions, which is something we use in Section 4.

Because each $\mathbf{y}_m$ follows a (multivariate) Gaussian distribution with some covariance matrix $\mathbf{\Sigma}_m$, the composite log-likelihood can be expressed as $c\ell(\boldsymbol{\theta};\mathbf{y}) = -\sum_{m=1}^{M} w_m(\log\det(\mathbf{\Sigma}_m) + \mathbf{y}_m^T\mathbf{\Sigma}_m^{-1}\mathbf{y}_m)/2$, up to an additive constant. Hence, a straightforward application of calculations from Mardia and Marshall (1984) yields the following result.

Theorem 1. *Let* $\boldsymbol{y}$ *follow a zero-mean Gaussian distribution with covariance parameter vector* $\boldsymbol{\theta}$*. Then,*

$$\{\boldsymbol{H}(\boldsymbol{\theta})\}_{ij} \equiv E\left[-\frac{\partial^2}{\partial\theta_i\partial\theta_j}c\ell(\boldsymbol{\theta};\boldsymbol{y})\right] = \frac{1}{2}\sum_{m=1}^{M} w_m\, tr\left(\mathbf{\Sigma}_m^{-1}\frac{\partial\mathbf{\Sigma}_m}{\partial\theta_i}\mathbf{\Sigma}_m^{-1}\frac{\partial\mathbf{\Sigma}_m}{\partial\theta_j}\right), \quad (3.1)$$

where $tr(\cdot)$ *is the trace operator.*

Obtaining an expression for $\mathbf{J}(\boldsymbol{\theta})$ is more challenging, because it involves taking the expectation of products between pairs of component densities, which introduces fourth-order moments. However, with the assistance of an additional lemma (all proofs are provided in the Supplementary Material), we obtain the following main result.

Theorem 2. *Let* $\boldsymbol{y}$ *follow a zero-mean Gaussian distribution with covariance parameter vector* $\boldsymbol{\theta}$*. Then,*

$$\begin{aligned}\{\boldsymbol{J}(\boldsymbol{\theta})\}_{ij} &\equiv E\left[\frac{\partial}{\partial\theta_i}c\ell(\boldsymbol{\theta};\boldsymbol{y})\frac{\partial}{\partial\theta_j}c\ell(\boldsymbol{\theta};\boldsymbol{y})\right] \\ &= \frac{1}{2}\sum_{m=1}^{M} w_m^2\, tr\left(\mathbf{\Sigma}_m^{-1}\frac{\partial\mathbf{\Sigma}_m}{\partial\theta_i}\mathbf{\Sigma}_m^{-1}\frac{\partial\mathbf{\Sigma}_m}{\partial\theta_j}\right) \\ &\quad + \frac{1}{2}\sum_{m=1}^{M}\sum_{l\neq m}^{M} w_m w_l\, tr\left(\mathbf{\Sigma}_m^{-1}\frac{\partial\mathbf{\Sigma}_m}{\partial\theta_i}\mathbf{\Sigma}_m^{-1}\mathbf{\Sigma}_{m,l}\mathbf{\Sigma}_l^{-1}\frac{\partial\mathbf{\Sigma}_l}{\partial\theta_j}\mathbf{\Sigma}_l^{-1}\mathbf{\Sigma}_{m,l}^T\right), \quad (3.2)\end{aligned}$$

where $\mathbf{\Sigma}_{m,l} \equiv cov(\boldsymbol{y}_m, \boldsymbol{y}_l)$*.*

Observe the similarity between the first terms in (3.1) and (3.2). In particular, if $w_m = 1$, which is the case for an unweighted composite marginal likelihood, then $\{\mathbf{J}(\boldsymbol{\theta})\}_{ij}$ differs from $\{\mathbf{H}(\boldsymbol{\theta})\}_{ij}$ only by the second term in (3.2). This additional term contributes to the loss of information from using the composite marginal likelihood compared with using the full likelihood.

The use of expressions (3.1) and (3.2) for inferential purposes is often complicated by the computational complexity of (3.2). In particular, the computation time scales with the square of the number of component densities, which offsets the time benefit of using many simpler densities for fast point estimation. Alternative approaches for estimating $\mathbf{J}(\boldsymbol{\theta})$ include window subsampling (Heagerty and Lele (1998)) and the parametric bootstrap (Bai, Kang and Song (2014)). Given the theoretical focus of this study, we leave this as an avenue for future research.

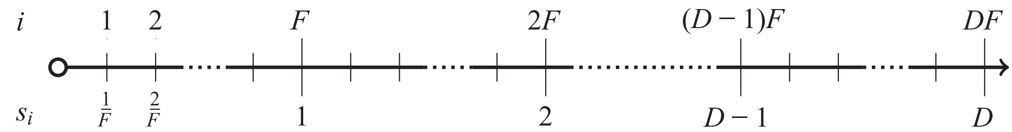

Figure 2. Location of observations under the sampling scheme $\mathcal{S}_{D,F}$. The expansion of the domain is controlled by D, whereas the infill frequency is controlled by F.

4. One-Dimensional Exponential Covariance Gaussian Spatial Process

For the remainder of this article, we focus on the special case of a zero-mean, one-dimensional exponential covariance Gaussian process, where we can obtain analytical large-sample results. We consider unweighted versions of both the composite full conditional likelihood and the composite marginal block likelihood, leaving brief discussion regarding the inclusion of weights to Section 5. We also provide results pertaining to the inclusion of a constant mean parameter μ in the Supplementary Material. An interesting finding related to μ is that the asymptotic behavior of the estimators for μ and σ^2 share similarities, such as similar expressions of asymptotic relative efficiency for the composite marginal block likelihood, and inconsistency under the infill asymptotic framework.

Consider a Gaussian spatial process $\{y(s), s \in (0,\infty)\}$ with mean zero and exponential covariance function $\text{cov}(y(s), y(s')) = \sigma^2 \exp(-|s-s'|\alpha)$. Suppose we use a sampling scheme $\mathcal{S}_{D,F} = \{s_i = i/F, i = 1, \ldots, DF = N\}$, as shown in Figure 2, where D controls the expansion of the spatial domain and F controls the spacing between adjacent observations. Under this setup, the covariance matrix for $\mathbf{y} = [y(s_1), y(s_2), \ldots, y(s_N)]^T$ has entries $\Sigma_{ij} = \sigma^2 \exp(-|i-j|\alpha/F)$, with $\det(\mathbf{\Sigma}) = \sigma^{2N}[1-\exp(-2\alpha/F)]^{N-1}$ and

$$\mathbf{\Sigma}^{-1} = \frac{1}{\sigma^2(1-e^{-2\alpha/F})}\begin{bmatrix} 1 & -e^{-\alpha/F} & 0 & \ldots & 0 & 0 \\ -e^{-\alpha/F} & 1+e^{-2\alpha/F} & -e^{-\alpha/F} & \ldots & 0 & 0 \\ 0 & -e^{-\alpha/F} & 1+e^{-2\alpha/F} & \ddots & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \ddots & 1+e^{-2\alpha/F} & -e^{-\alpha/F} \\ 0 & 0 & 0 & \ldots & -e^{-\alpha/F} & 1 \end{bmatrix}. \tag{4.1}$$

Zhang and Zimmerman (2005) showed that the Fisher information matrix under this setting is given by

$$\mathbf{I}(\sigma^2, \alpha; \mathbf{y}) = \begin{bmatrix} \dfrac{DF}{2\sigma^4} & \dfrac{(DF-1)\rho^2}{\sigma^2 F(1-\rho^2)} \\ \dfrac{(DF-1)\rho^2}{\sigma^2 F(1-\rho^2)} & \dfrac{(DF-1)\rho^2(1+\rho^2)}{F^2(1-\rho^2)^2} \end{bmatrix}, \tag{4.2}$$

where $\rho = \exp(-\alpha/F)$. This serves as a baseline for calculating the relative efficiency of our composite likelihood estimators. In addition, (4.2) allows us to analyze the asymptotic behavior of the maximum likelihood estimation when both σ^2 and α are unknown, as well as when either one is known. In particular, note that $\hat{\sigma}^2_{\mathrm{ML}|\alpha}$ and $\hat{\alpha}_{\mathrm{ML}|\sigma^2}$, the maximum likelihood estimators of σ^2 and α, respectively, when the other is known, are consistent under all three asymptotic frameworks, but $\hat{\sigma}^2_{\mathrm{ML}}$ and $\hat{\alpha}_{\mathrm{ML}}$, the respective maximum likelihood estimators when both parameters are unknown, are only consistent under the expanding domain and hybrid frameworks, as we show indirectly.

In the subsections that follow, we use asymptotic relative efficiency as a metric to compare the composite likelihood estimators. For our model with the parameters $(\sigma^2, \alpha)^T$ (which we temporarily relabel as $(\theta_1, \theta_2)^T$, for generality of the expressions below), we calculate the expanding-domain asymptotic relative efficiency (EDARE) for an individual parameter as $\mathrm{EDARE}(\hat{\theta}_{i,\mathrm{CL}}, \hat{\theta}_{i,\mathrm{ML}}) \equiv \lim_{D\to\infty}\{\mathbf{I}(\theta_1,\theta_2)^{-1}\}_{ii}/\{\mathbf{G}(\theta_1,\theta_2)^{-1}\}_{ii}$, for $i \in \{1,2\}$, when both parameters are unknown, and as $\mathrm{EDARE}(\hat{\theta}_{i,\mathrm{CL}|\theta_j}, \hat{\theta}_{i,\mathrm{ML}|\theta_j}) \equiv \lim_{D\to\infty}\{\mathbf{I}(\theta_1,\theta_2)_{ii}\}^{-1}/[\{\mathbf{H}(\theta_1,\theta_2)_{ii}\}^{-1}\mathbf{J}(\theta_1,\theta_2)_{ii}\{\mathbf{H}(\theta_1,\theta_2)_{ii}\}^{-1}]$, for $i \neq j$, when only θ_j is known. As an overall measure of efficiency, we also calculate $\text{Overall EDARE} \equiv \lim_{D\to\infty}\{\det[\mathbf{I}(\theta_1,\theta_2)^{-1}]/\det[\mathbf{G}(\theta_1,\theta_2)^{-1}]\}^{1/2}$. Where appropriate, we consider the infill asymptotic relative efficiency by replacing $D \to \infty$ with $F \to \infty$, and the hybrid asymptotic relative efficiency by taking both D and F to ∞. Note that in the cases we consider, the hybrid asymptotic relative efficiency is unaffected by the relative growth rates of D and F.

Our approach for deriving the sandwich covariance matrix relies on expressing the sum of the individual Gaussian log-densities in the form $c\ell(\sigma^2,\alpha;\mathbf{y}) = q(\sigma^2,\alpha) - \mathbf{y}^T\mathbf{M}(\alpha)\mathbf{y}/(2\sigma^2)$, for some function $q(\cdot)$ and $N \times N$ symmetric matrix $\mathbf{M}(\alpha)$, which can be decomposed as a linear combination of simpler matrices $\{\mathbf{A}_k\}$, each of which has a well-defined structure; that is, $\mathbf{M}(\alpha) = \sum c_k(\alpha)\mathbf{A}_k$. The objective is to obtain expressions for $\mathrm{tr}(\mathbf{A}_k\boldsymbol{\Sigma})$ and $\mathrm{tr}(\mathbf{A}_k\boldsymbol{\Sigma}\mathbf{A}_l\boldsymbol{\Sigma})$ for all k and l, which can then be used to find $\mathbf{H}(\sigma^2,\alpha)$ and $\mathbf{J}(\sigma^2,\alpha)$. Proofs of the theorems that follow are provided in the Supplementary Material.

4.1. Composite full conditional likelihood

Owing to Markovian dependence in the exponential covariance setting, it is well known (e.g., see p. 170–171 of Cressie and Wikle (2011)) that the distribution of $y(s_i)$ given all of the other observations follows the same distribution as conditioning on the two nearest neighbors on either side of $y(s_i)$ (or on the single nearest neighbor for $y(s_1)$ and $y(s_N)$). In particular, we have $(y(s_1)|y(s_2)) \sim N(\exp(-\alpha/F)y(s_2), \sigma^2[1-\exp(-2\alpha/F)])$, $(y(s_N)|y(s_{N-1})) \sim N(\exp(-\alpha/F)y(s_{N-1}), \sigma^2[1-\exp(-2\alpha/F)])$, and

$$(y(s_i)|y(s_{i-1}), y(s_{i+1})) \sim N\left(\frac{e^{-\alpha/F}[y(s_{i-1}) + y(s_{i+1})]}{1 + e^{-2\alpha/F}}, \sigma^2 \frac{1 - e^{-2\alpha/F}}{1 + e^{-2\alpha/F}}\right),$$

for $1 < i < N$. The composite full conditional log-likelihood is therefore equivalent to a composite conditional "two nearest neighbors" conditional log-likelihood, as given by

$$\begin{aligned} c\ell(\sigma^2, \alpha; \mathbf{y}) &= \log f(y(s_2)|y(s_1); \sigma^2, \alpha) + \log f(y(s_N)|y(s_{N-1}); \sigma^2, \alpha) \\ &\quad + \sum_{m=2}^{N-1} \log f(y(s_m)|y(s_{m-1}), y(s_{m+1}); \sigma^2, \alpha) \\ &= \left(\frac{N}{2} - 1\right) \log\left(1 + e^{-2\alpha/F}\right) - \frac{N}{2} \log\left(2\pi\sigma^2[1 - e^{-2\alpha/F}]\right) \\ &\quad - \frac{1}{2\sigma^2} \mathbf{y}^T \mathbf{M}(\alpha) \mathbf{y}, \end{aligned} \tag{4.3}$$

where $\mathbf{M}(\alpha)$ has a pentadiagonal structure, such that $\mathbf{M}(\alpha)$ can be expressed as a linear combination of five simpler $N \times N$ matrices $\{\mathbf{A}_k, k = 1, 2, .., 5\}$, the entries of which can be defined in terms of the indicator function $I(\cdot)$, as follows: $(\mathbf{A}_1)_{ij} = I(i = j)$, $(\mathbf{A}_2)_{ij} = I(i = j, i \notin \{1, N\})$, $(\mathbf{A}_3)_{ij} = I(i = j, i \notin \{1, 2, N-1, N\})$, $(\mathbf{A}_4)_{ij} = I(|i - j| = 1)$, and $(\mathbf{A}_5)_{ij} = I(|i - j| = 2)$. In particular, we can write

$$\mathbf{M}(\alpha) = \frac{1}{1 - \rho^2}\left(\frac{1 + 2\rho^2}{1 + \rho^2}\mathbf{A}_1 + 2\rho^2\mathbf{A}_2 - \frac{\rho^4}{1 + \rho^2}\mathbf{A}_3 - 2\rho\,\mathbf{A}_4 + \frac{\rho^2}{1 + \rho^2}\mathbf{A}_5\right), \tag{4.4}$$

where $\rho = \exp(-\alpha/F)$. Under this setup, we obtain the following result.

Theorem 3. *For the exponential covariance Gaussian process with composite full conditional log-likelihood defined by* (4.3)*, the matrices comprising the sandwich covariance matrix are given by*

$$\boldsymbol{H}(\sigma^2, \alpha) = \begin{bmatrix} \dfrac{DF}{2\sigma^4} & \dfrac{2\rho^2}{F\sigma^2(1 - \rho^4)}[DF - (1 - \rho^2)] \\ \dfrac{2\rho^2}{F\sigma^2(1 - \rho^4)}[DF - (1 - \rho^2)] & \dfrac{2\rho^2}{F^2(1 - \rho^4)^2}\boldsymbol{h}_1^T(DF, 1)^T \end{bmatrix}$$

and

$$\boldsymbol{J}(\sigma^2, \alpha) = \begin{bmatrix} \dfrac{1}{2\sigma^4(1 + \rho^2)^2}\boldsymbol{j}_1^T(DF, 1)^T & \dfrac{4\rho^2}{\sigma^2 F(1 - \rho^2)(1 + \rho^2)^3}\boldsymbol{j}_2^T(DF, 1)^T \\ \dfrac{4\rho^2}{\sigma^2 F(1 - \rho^2)(1 + \rho^2)^3}\boldsymbol{j}_2^T(DF, 1)^T & \dfrac{4\rho^2}{F^2(1 - \rho^2)^2(1 + \rho^2)^4}\boldsymbol{j}_3^T(DF, 1)^T \end{bmatrix},$$

where

$$\boldsymbol{h}_1 = \begin{bmatrix} 1 + \rho^2 + 3\rho^4 - \rho^6 \\ -1 + \rho^2 - 3\rho^4 + 3\rho^6 \end{bmatrix}, \quad \boldsymbol{j}_1 = \begin{bmatrix} 1 + 4\rho^2 + \rho^4 \\ -2\rho^2 + 4\rho^4 \end{bmatrix},$$

Table 1. EDAREs for the exponential covariance Gaussian process.

	Composite Full Conditional Likelihood	Composite Marginal Block Likelihood (fixed W)
EDARE$(\hat{\sigma}^2_{\mathrm{CL}\mid\alpha}, \hat{\sigma}^2_{\mathrm{ML}\mid\alpha})$	$\frac{(1+\rho^2)^2}{1+4\rho^2+\rho^4}$	$\frac{1}{1+2\rho^2/[W(1-\rho^{2W})]}$
EDARE$(\hat{\alpha}_{\mathrm{CL}\mid\sigma^2}, \hat{\alpha}_{\mathrm{ML}\mid\sigma^2})$	$\frac{(1+\rho^2+3\rho^4-\rho^6)^2}{(1+\rho^2)(1+2\rho^2+6\rho^4+2\rho^6+\rho^8)}$	$\frac{W-1}{W}$
EDARE$(\hat{\sigma}^2_{\mathrm{CL}}, \hat{\sigma}^2_{\mathrm{ML}})$	$\frac{(1-\rho^2)(1+\rho^2)}{1+\rho^4}$	$\frac{(1-\rho^2+2\rho^2/W)/(1-\rho^2)}{1+[2\rho^2(1+\rho^2)]/[W(1-\rho^{2W})(1-\rho^2+2\rho^2/W)]}$
EDARE$(\hat{\alpha}_{\mathrm{CL}}, \hat{\alpha}_{\mathrm{ML}})$	$1-\rho^2$	$\frac{(1-\rho^2+2\rho^2/W)/(1-\rho^2)}{W/(W-1)+4\rho^4/[W(1-\rho^{2W})(1-\rho^2+2\rho^2/W)]}$
Overall EDARE	$(1-\rho^2)^{1/2}$	$\frac{(1-\rho^2+2\rho^2/W)(W-1)^{1/2}/(1-\rho^2)^{1/2}}{\{W-(W-2)\rho^2+2\rho^2(1+\rho^2)/(1-\rho^{2W})\}^{1/2}}$

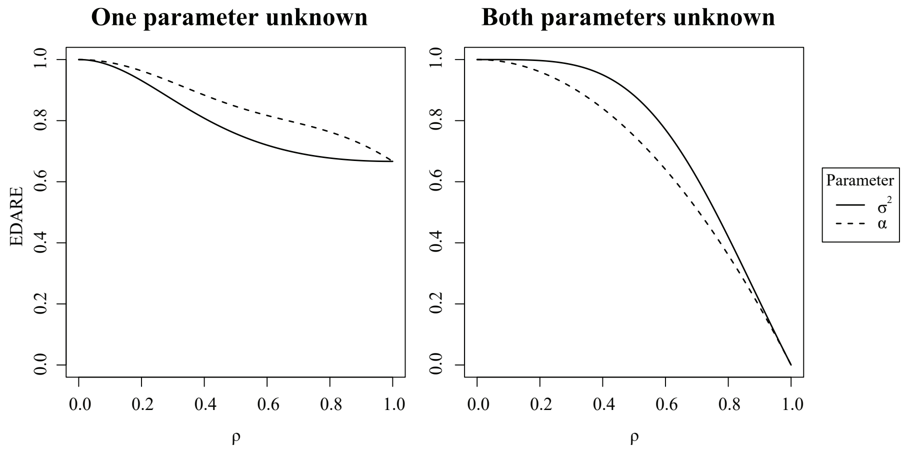

Figure 3. EDARE with respect to $\rho = \exp(-\alpha/F)$ of the composite full conditional likelihood for the exponential covariance Gaussian process.

$$\boldsymbol{j}_2 = \begin{bmatrix} 1+\rho^2+\rho^4 \\ -1+\rho^2+\rho^6 \end{bmatrix}, \quad \boldsymbol{j}_3 = \begin{bmatrix} 1+2\rho^2+6\rho^4+2\rho^6+\rho^8 \\ -1+\rho^2-4\rho^4+8\rho^6+\rho^8-\rho^{10} \end{bmatrix}.$$

Using the above result and (4.2), we can calculate the asymptotic relative efficiency under all three asymptotic frameworks, and when one or both parameters are unknown (a direct expression for $\mathbf{G}(\sigma^2,\alpha)^{-1}$ is provided in the Supplementary Material).

Corollary 1. *The EDAREs of the maximum composite full conditional likelihood estimators for σ^2 and α are presented in the middle column of Table* 1*. When one parameter is known, the (individual) infill and hybrid asymptotic relative efficiency for the other parameter is* 2/3*. However, when both parameters are unknown, the entries in $\boldsymbol{G}(\sigma^2,\alpha)^{-1}$ diverge to ∞ under the infill and hybrid asymptotic frameworks.*

The EDAREs are plotted with respect to $\rho \in (0,1)$ in Figure 3, showing that they are decreasing functions with respect to ρ. This suggests that this composite likelihood estimator performs better when the data are closer to being i.i.d., that is, when the data are spaced far apart relative to the size of the spatial dependence. Conversely, the estimator is less efficient if the dependence between adjacent observations is high, which is inevitable under the infill and hybrid frameworks. Note that the efficiency of 2/3 achieved under these frameworks when one parameter is known agrees with the results of Bachoc, Lagnoux and Nguyen (2017). In contrast, the efficiency decreases to zero when both parameters are unknown, which is caused by the entries in $\mathbf{G}(\sigma^2, \alpha)^{-1}$ diverging to ∞. This highlights an inherent structural issue with the composite full conditional likelihood in this setting.

4.2. Composite marginal block likelihood

When constructing the composite marginal block likelihood, there is ambiguity in how the DF observations are grouped. However, as a natural choice for a one-dimensional equally spaced lattice, and to simplify developments without compromising insight, we follow Caragea and Smith (2007) by grouping observations into B blocks along the number line. We further assume that each block contains exactly W observations, such that $N = DF = BW$. Under this setup, the composite likelihood function is given by

$$\begin{aligned} c\ell(\sigma^2, \alpha; \mathbf{y}) &= \sum_{m=0}^{B-1} \log f(y(s_{Wm+1}), y(s_{Wm+2}), \ldots, y(s_{W(m+1)}); \sigma^2, \alpha) \\ &= -\frac{DF - B}{2} \log\left(1 - e^{-2\alpha/F}\right) - \frac{N}{2} \log\left(2\pi\sigma^2\right) - \frac{1}{2\sigma^2} \mathbf{y}^T \mathbf{M}(\alpha) \mathbf{y}, \end{aligned} \tag{4.5}$$

where $\mathbf{M}(\alpha) = \operatorname{diag}(\mathbf{Q}(\alpha), \mathbf{Q}(\alpha), \ldots, \mathbf{Q}(\alpha))$ for the $W \times W$ matrix

$$\mathbf{Q}(\alpha) = \frac{1}{1 - e^{-2\alpha/F}} \begin{bmatrix} 1 & -e^{-\alpha/F} & 0 & \ldots & 0 & 0 \\ -e^{-\alpha/F} & 1 + e^{-2\alpha/F} & -e^{-\alpha/F} & \ldots & 0 & 0 \\ 0 & -e^{-\alpha/F} & 1 + e^{-2\alpha/F} & \ddots & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \ddots & 1 + e^{-2\alpha/F} & -e^{-\alpha/F} \\ 0 & 0 & 0 & \ldots & -e^{-\alpha/F} & 1 \end{bmatrix}.$$

Using the definitions of the simple matrices from (4.4), but reduced to size $W \times W$, we can also write $\mathbf{Q}(\alpha) = (\mathbf{A}_1 + \rho^2 \mathbf{A}_2 - \rho \mathbf{A}_4)/(1 - \rho^2)$. This leads to the following result:

Theorem 4. *For the exponential covariance Gaussian process with composite marginal block log-likelihood defined by* (4.5), *the matrices comprising the sandwich covariance matrix are given by*

$$\boldsymbol{H}(\sigma^2,\alpha) = \begin{bmatrix} \dfrac{DF}{2\sigma^4} & \dfrac{\rho^2(DF-B)}{F\sigma^2(1-\rho^2)} \\ \dfrac{\rho^2(DF-B)}{F\sigma^2(1-\rho^2)} & \dfrac{\rho^2(1+\rho^2)(DF-B)}{F^2(1-\rho^2)^2} \end{bmatrix}$$

and

$$\begin{aligned}\boldsymbol{J}(\sigma^2,\alpha) &= \begin{bmatrix} \dfrac{1}{2\sigma^4}\Big[DF + \dfrac{2\rho^2}{1-\rho^{2W}}\Big(B - \dfrac{1-\rho^{2DF}}{1-\rho^{2W}}\Big)\Big] & \dfrac{\rho^2}{\sigma^2 F(1-\rho^2)}(DF-B) \\ \dfrac{\rho^2}{\sigma^2 F(1-\rho^2)}(DF-B) & \dfrac{\rho^2(1+\rho^2)}{F^2(1-\rho^2)^2}(DF-B) \end{bmatrix} \\ &= \boldsymbol{H}(\sigma^2,\alpha) + \begin{bmatrix} \dfrac{\rho^2}{\sigma^4(1-\rho^{2W})}\Big(B - \dfrac{1-\rho^{2DF}}{1-\rho^{2W}}\Big) & 0 \\ 0 & 0 \end{bmatrix}.\end{aligned}$$

When analyzing the asymptotic performance of the composite likelihood estimator here, the growth rates of W and B alongside D and F are important. However, from a computational standpoint, we are most interested in the case in which the block size (W) remains fixed and B grows, thus avoiding the computationally expensive evaluation of component densities with many observations. As such, the following results consider W as fixed and B growing in proportion to $N = DF$. Using Theorem 4 and (4.2), we can calculate the asymptotic relative efficiency under all three asymptotic frameworks, as well as when one or both parameters are unknown (a direct expression for $\mathbf{G}(\sigma^2,\alpha)^{-1}$ is provided in the Supplementary Material).

Corollary 2. *The EDAREs of the maximum composite marginal block likelihood estimators for σ^2 and α are presented in the right column of Table* 1*. Under infill asymptotics, we have*

$$\begin{aligned}\lim_{F\to\infty} var(\hat{\sigma}^2_{CL|\alpha}) &= \lim_{F\to\infty}\{\boldsymbol{H}(\sigma^2,\alpha)_{11}\}^{-1}\boldsymbol{J}(\sigma^2,\alpha)_{11}\{\boldsymbol{H}(\sigma^2,\alpha)_{11}\}^{-1} \\ &= \frac{\sigma^4}{W^2}\left(\frac{2}{\alpha D} - \frac{1-\exp(-2\alpha D)}{\alpha^2 D^2}\right);\end{aligned}$$

thus, $\hat{\sigma}^2_{CL|\alpha}$ is inconsistent under infill (D fixed), and consistent under the hybrid framework ($D, F \to \infty$), with a relative efficiency of zero. In contrast, $\hat{\alpha}_{CL|\sigma^2}$ is consistent under the infill and hybrid asymptotic frameworks, with the same efficiency as in the expanding domain case ($1-1/W$). When both parameters are unknown, $\boldsymbol{G}(\sigma^2,\alpha)^{-1}$ converges to the zero matrix under the hybrid framework, and the estimators achieve full efficiency.

The EDAREs are plotted in Figure 4 with respect to W and ρ. Plots (b) and (d) show that for any given value of ρ, the efficiency of $\hat{\sigma}^2_{\mathrm{CL}|\alpha}$ and $\hat{\alpha}_{\mathrm{CL}|\sigma^2}$ relative to their maximum likelihood counterparts is a monotonically increasing function with respect to W, and approaches one. This is because W controls the extent

Table 2. (Unscaled) asymptotic variance of various estimators based on the Fisher information or sandwich covariance matrix for the exponential covariance Gaussian process under infill.

	Maximum Likelihood	Composite Full Conditional Likelihood	Composite Marginal Block Likelihood (fixed W)
$\hat{\sigma}^2$	> 0	∞	> 0
$\hat{\alpha}$	> 0	∞	> 0
$\hat{\sigma}^2_{\vert\alpha}$	0	0 (Infill ARE $= 2/3$)	> 0
$\hat{\alpha}_{\vert\sigma^2}$	0	0 (Infill ARE $= 2/3$)	0 (Infill ARE $= 1 - 1/W$)
$\widehat{\sigma^2\alpha}$	0	0 (Infill ARE $= 2/3$)	0 (Infill ARE $= 1 - 1/W$)

to which the full likelihood is misspecified, with $W = \infty$ corresponding to the full likelihood.

Plot (a) of Figure 4 shows that as ρ increases, the efficiency for $\hat{\sigma}^2_{\mathrm{CL}|\alpha}$ decreases to zero at all block sizes. This aligns with the finding in Corollary 2 that $\hat{\sigma}^2_{\mathrm{CL}|\alpha}$ is inconsistent under infill ($F \to \infty$), where ρ tends to one. The composite likelihood estimator can be written as $\hat{\sigma}^2_{\mathrm{CL}|\alpha} = (W/N)\sum_{b=1}^{N/W} \hat{\sigma}^2_{b,\mathrm{ML}|\alpha}$, a simple average of the maximum likelihood estimates of σ^2 within each block. Thus, as F increases, the blocks come closer together, which increases the correlation between individual estimates, and hinders the overall reduction in the variance of $\hat{\sigma}^2_{\mathrm{CL}|\alpha}$. This presents an interesting difference to the infill asymptotic behavior of the composite full conditional likelihood, as highlighted in Table 2.

In contrast to $\hat{\sigma}^2_{\mathrm{CL}|\alpha}$, plot (c) of Figure 4 shows that the efficiency of $\hat{\alpha}_{\mathrm{CL}|\sigma^2}$ does not change with respect to ρ, or equivalently, with respect to α or F. Furthermore, plot (d) shows that its efficiency is zero when $W = 1$, where the composite likelihood function treats all observations as independent, and is obviously not a function of α.

When both parameters are unknown, plot (f) of Figure 4 shows that the relative efficiency for σ^2 exhibits a peculiar trough with respect to W for $\rho > 0$, in contrast to the monotonically increasing relationship in plot (b) when α is known. This is partly because we now have full efficiency at both extremes of W, instead of just at $W = \infty$. The improvement in efficiency at $W = 1$ and $W = 2$ for $\hat{\sigma}^2_{\mathrm{CL}}$ compared with that for $\hat{\sigma}^2_{\mathrm{CL}|\alpha}$ is mainly attributable to the maximum likelihood estimator performing comparatively worse when α is unknown, because we know that $\hat{\sigma}^2_{\mathrm{ML}|\alpha}$ is consistent under infill, but $\hat{\sigma}^2_{\mathrm{ML}}$ is not. Between the two extremes, a possible explanation for the loss of efficiency is that at the lower end, σ^2 is being treated as a "between-blocks" variance parameter, whereas at the upper end, it is a "within-blocks" variance parameter. Hence, at the other values of W, $\hat{\sigma}^2_{\mathrm{CL}}$ provides a compromise between these two conflicting extremes.

Interestingly, plots (d) and (h) of Figure 4 show that the relative efficiency of

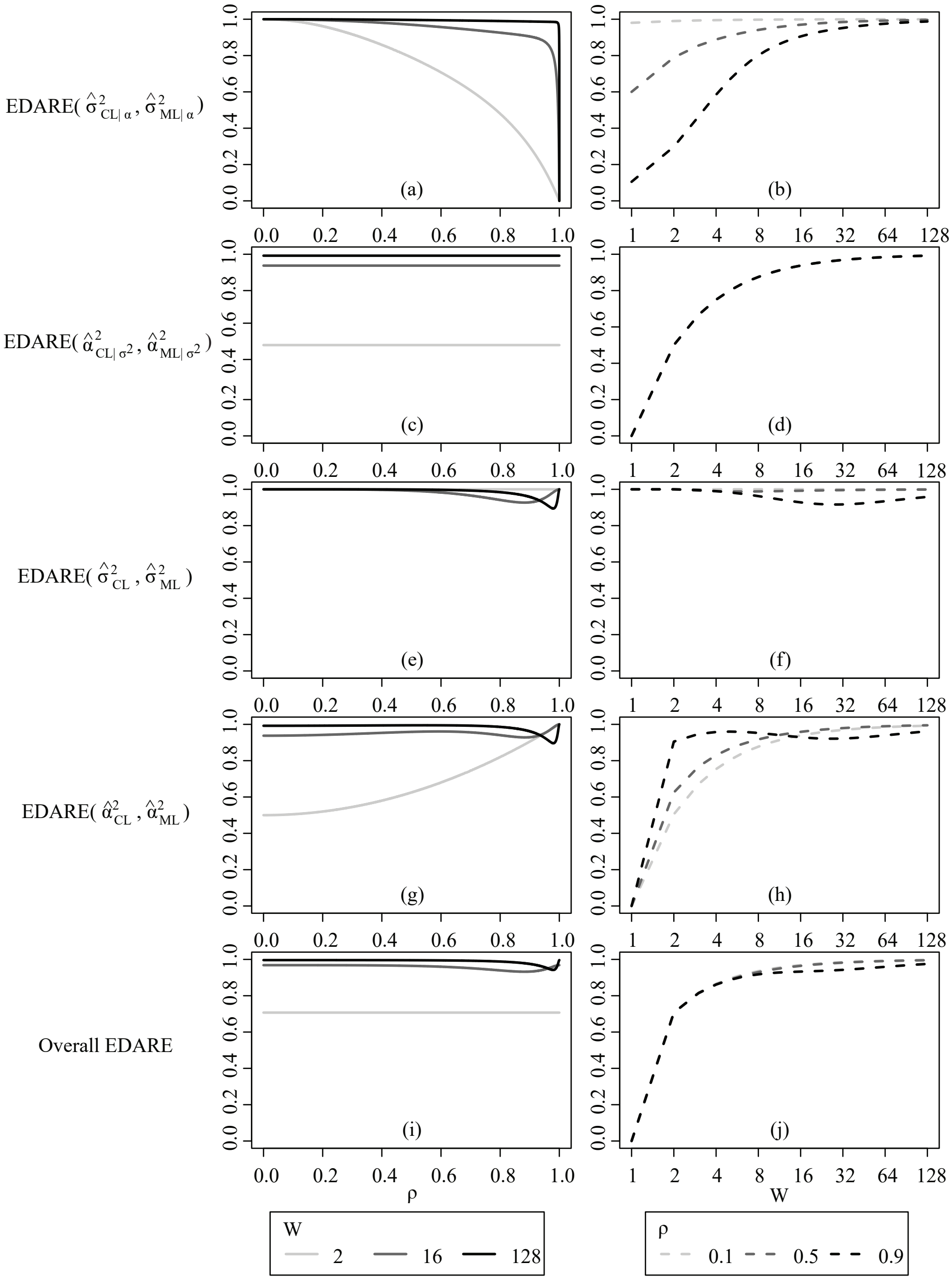

Figure 4. EDARE (with $B \to \infty$) of the composite marginal block likelihood for the exponential covariance Gaussian process. The left column displays ρ on the x-axis, with each curve representing a different value of W, and the right column displays W on the x-axis, with each curve representing a different value of ρ. Note that the EDARE curves in (c) and (d) do not depend on ρ.

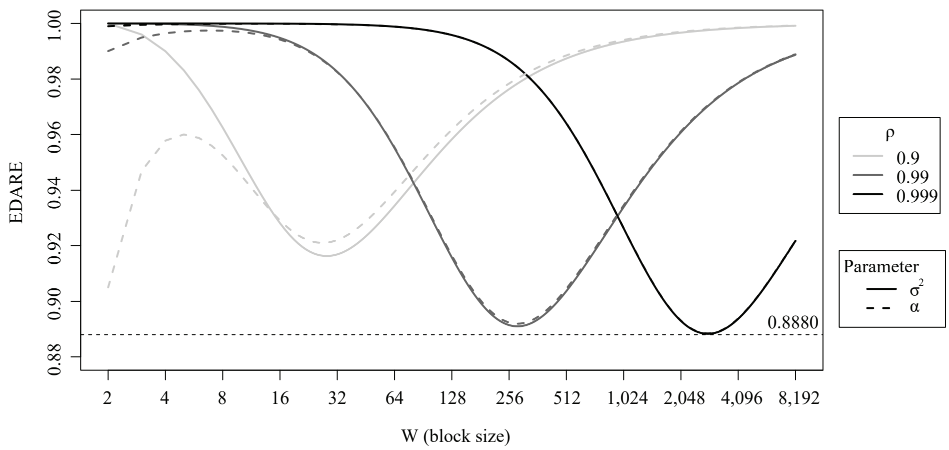

Figure 5. EDARE (with $B \to \infty$) of the composite marginal block likelihood for the exponential covariance Gaussian process for different fixed values of W and when $\rho = \exp(-\alpha/F)$ is near one.

$\hat{\alpha}_{\text{CL}}$ when $\rho \approx 0$ (or $\alpha \approx \infty$) aligns exactly with that of $\hat{\alpha}_{\text{CL}|\sigma^2}$ for all block sizes. However, for other values of ρ, the efficiency of $\hat{\alpha}_{\text{CL}}$ changes, unlike that of $\hat{\alpha}_{\text{CL}|\sigma^2}$, which is likely an effect of the relationship between $\hat{\sigma}^2_{\text{CL}}$ and $\hat{\alpha}_{\text{CL}}$, because they have to be estimated simultaneously. In fact, when the strength of dependence is high, the EDARE of $\hat{\alpha}_{\text{CL}}$ also begins to exhibit a trough. This is exemplified in Figure 5, where, for a fixed $W \geq 2$, we observe that the relative efficiencies of the two parameters converge to the same value as $\rho \to 1$ (or, equivalently, for $\alpha \approx 0$ or $F \to \infty$). Furthermore, as $\rho \to 1$, the minimum of the EDARE curve decreases in value and is attained at a larger block size, though it can be shown numerically that the lowest relative efficiency for any W and $\rho \in (0, 1)$ is 0.8880 (to four decimal places). Owing to this right-shifting behavior, under hybrid asymptotics where both D and F approach infinity, we achieve full efficiency for both σ^2 and α, for any $W \geq 2$. This suggests that, in the limit, there is a negligible information loss from misspecifying independence between blocks under the hybrid asymptotic framework.

In terms of overall efficiency, for $W = 2$ in plots (e), (g), and (i) of Figure 4, the overall EDARE remains constant with respect to ρ (Overall EDARE $=$ $1/\sqrt{2}$), despite the (individual) EDARE for σ^2 being constant (EDARE $= 1$) and the EDARE for α increasing to one as $\rho \to 1$. This contrast between the full individual relative efficiency (as $\rho \to 1$) and the nonfull overall relative efficiency is driven by a stronger information dependence/correlation between $\hat{\sigma}^2_{\text{CL}}$ and $\hat{\alpha}_{\text{CL}}$ compared with that between $\hat{\sigma}^2_{\text{ML}}$ and $\hat{\alpha}_{\text{ML}}$. Note that this would adversely affect the efficiency of an estimator that is a function of both $\hat{\sigma}^2_{\text{CL}}$ and $\hat{\alpha}_{\text{CL}}$ (e.g., $\hat{\sigma}^2_{\text{CL}}\hat{\alpha}_{\text{CL}}$), but is largely inconsequential in the typical situation in which an inference is carried out on each covariance parameter separately.

Given prior results in the literature, we also highlight the efficiency of the estimation for the microergodic parameter $\sigma^2\alpha$ in the context of infill asymptotics presented in Table 2. By applying a transformation of the parameters $\boldsymbol{\phi} = (\phi_1, \phi_2)^T = (\sigma^2\alpha, \alpha)^T$ to the sandwich covariance matrix, our results match those of Bachoc, Lagnoux and Nguyen (2017) that the infill asymptotic relative efficiency of the microergodic parameter for the composite full conditional likelihood is $2/3$. In the same manner, we find that the infill asymptotic relative efficiency of the microergodic parameter for the composite marginal block likelihood is $1 - 1/W$.

5. Conclusion

We have presented a unified case study of asymptotic relative efficiency under the expanding domain, infill, and hybrid frameworks. This was made possible by deriving a closed-form expression for the sandwich covariance matrix under the one-dimensional exponential covariance Gaussian process. We showed that the composite full conditional likelihood performs reasonably when only one parameter is unknown, but is structurally unstable when both parameters are unknown, where it leads to inconsistent estimations under both the infill and the hybrid asymptotic frameworks. On the other hand, the composite marginal block likelihood performs worse, in general, when one parameter is unknown, but exhibits high efficiency, even for small block sizes, when both parameters are unknown, achieving full asymptotic relative efficiency under the hybrid framework. These results highlight the need to carefully consider the large-sample framework when determining which composite likelihood function to use for estimation and inference. Overall, for a one-dimensional exponential covariance Gaussian process with both σ^2 and α unknown, we recommend using the composite marginal block likelihood, but not the composite full conditional likelihood, owing to its aforementioned instability.

For the composite marginal block likelihood, given the various nontrivial relationships between the block size, strength of dependence, and relative efficiencies of the estimators, there is no straightforward optimal choice of block size. However, small block sizes benefit the most from a computational standpoint. Because the efficiency for σ^2 can fall no lower than 0.8880 (see Figure 5), which is still quite high, it would be reasonable to choose W based solely on the efficiency for α in this situation. One approach to achieving this would be to specify a desired level of relative efficiency for α, and then solve for W. Here, a rule of thumb is to consider the worst-case scenario where $\rho \approx 0$, for which $\text{EDARE}(\hat{\alpha}_{\text{CL}}, \hat{\alpha}_{\text{ML}}) \approx (W-1)/W$. Our choice of W based on a desired level of relative efficiency q is then $W = \lceil 1/(1-q) \rceil$, where $\lceil \cdot \rceil$ is the ceiling function.

In future work, we would like to derive the sandwich covariance matrix for weighted versions of the composite likelihood functions. Our findings can be extended to allow for general weights, but the sandwich covariance matrix and subsequent efficiency expressions will likely be less interpretable, except under specific weight configurations, such as binary weights (Bevilacqua and Gaetan (2015)) or optimal weights (Bachoc, Bevilacqua and Velandia (2019); Pace, Salvan and Sartori (2019)). A particular set of weights of interest is the composite full conditional likelihood with alternating binary weights, in line with the work of Besag (1974). Here, initial simulations (not shown) suggest that as the strength of dependence between adjacent observations increases, the variance of the estimator does not diverge to infinity.

It is also of interest to examine the consistency and asymptotic normality of the estimators presented here, particularly under the infill and hybrid asymptotic frameworks. One approach to proving these statistical properties is to combine and extend the proofs presented in Chang, Huang and Ing (2017) and Bachoc, Lagnoux and Nguyen (2017). In particular, for flexibility when analyzing different asymptotic frameworks, it would be preferable to use the sampling scheme of Chang, Huang and Ing (2017), where there are N equally spaced observations in the range $(0, N^{\delta}]$ (with $s_1 = N^{\delta-1}$). In this way, infill occurs when $\delta = 0$, the expanding domain occurs when $\delta = 1$, and any value in between corresponds to a hybrid asymptotic framework.

The efficiency of various composite likelihood functions can also be explored for extensions to the one-dimensional exponential covariance process, such as including a nugget effect (observation measurement error) or a multidimensional separable covariance. This would follow from existing work on these cases for maximum likelihood estimation (Ying (1993); Chen, Simpson and Ying (2000); Chang, Huang and Ing (2017)). The two-dimensional separable exponential covariance case is of particular interest, because the maximum likelihood estimators are consistent under infill, in contrast to the one-dimensional case (Ying (1993)). Caragea and Smith (2007) also study this process for the composite marginal block likelihood, which is still quite efficient, although less so than in the one-dimensional case. Lastly, it would be interesting to consider the effects of “mixed” spatial asymptotics on consistency and efficiency in a multidimensional setting, such as expanding the domain in one direction and infilling in another direction.

Supplementary Material

The online supplementary material covers proofs of Theorems 2, 3 and 4 from this paper, as well as derivations of the Fisher information and sandwich covariance for a constant mean parameter under the exponential covariance model.

Acknowledgments

The authors acknowledge the associate editor and two referees for their helpful comments and suggestions.

References

Abt, M. and Welch, W. J. (1998). Fisher information and maximum-likelihood estimation of covariance parameters in Gaussian stochastic processes. *Canadian Journal of Statistics* **26**, 127–137.

Bachoc, F., Bevilacqua, M. and Velandia, D. (2019). Composite likelihood estimation for a Gaussian process under fixed domain asymptotics. *Journal of Multivariate Analysis* **174**, 104534.

Bachoc, F., Lagnoux, A. and Nguyen, T. M. N. (2017). Cross-validation estimation of covariance parameters under fixed-domain asymptotics. *Journal of Multivariate Analysis* **160**, 42–67.

Bai, Y., Kang, J. and Song, P. X. K. (2014). Efficient pairwise composite likelihood estimation for spatial-clustered data. *Biometrics* **70**, 661–670.

Bai, Y., Song, P. X. K. and Raghunathan, T. E. (2012). Joint composite estimating functions in spatiotemporal models. *Journal of the Royal Statistical Society. Series B (Statistical Methodology)* **74**, 799–824.

Besag, J. (1974). Spatial interaction and the statistical analysis of lattice systems. *Journal of the Royal Statistical Society. Series B (Methodological)* **36**, 192–236.

Bevilacqua, M. and Gaetan, C. (2015). Comparing composite likelihood methods based on pairs for spatial Gaussian random fields. *Statistics and Computing* **25**, 877–892.

Caragea, P. and Smith, R. (2007). Asymptotic properties of computationally efficient alternative estimators for a class of multivariate normal models. *Journal of Multivariate Analysis* **98**, 1417–1440.

Chang, C. H., Huang, H. C. and Ing, C. K. (2017). Mixed domain asymptotics for a stochastic process model with time trend and measurement error. *Bernoulli* **23**, 159–190.

Chen, H. S., Simpson, D. G. and Ying, Z. (2000). Infill asymptotics for a stochastic process model with measurement error. *Statistica Sinica* **10**, 141–156.

Cressie, N. and Wikle, C. K. (2011). *Statistics for Spatio-Temporal Data*. John Wiley and Sons, Hoboken.

Davis, R. A. and Chun, Y. Y. (2011). Comments on pairwise likelihood in time series models. *Statistica Sinica* **21**, 255–277.

Eidsvik, J., Shaby, B. A., Reich, B. J., Wheeler, M. and Niemi, J. (2014). Estimation and prediction in spatial models with block composite likelihoods. *Journal of Computational and Graphical Statistics* **23**, 295–315.

Hall, P. and Patil, P. (1994). Properties of nonparametric estimators of autocovariance for stationary random fields. *Probability Theory and Related Fields* **99**, 399–424.

Heagerty, P. J. and Lele, S. R. (1998). A composite likelihood approach to binary spatial data. *Journal of the American Statistical Association* **93**, 1099–1111.

Ibragimov, I. A. and Rozanov, Y. A. (1978). *Gaussian Random Processes*. Springer, New York.

Kaufman, C. and Shaby, B. (2013). The role of the range parameter for estimation and prediction in geostatistics. *Biometrika* **100**, 473–484.

Kent, J. T. (1982). Robust properties of likelihood ratio tests. *Biometrika* **69**, 19–27.

Lindsay, B. G. (1988). Composite likelihood methods. *Contemporary Mathematics* **80**, 221–239.

Lu, Z. and Tjøstheim, D. (2014). Nonparametric estimation of probability density functions for irregularly observed spatial data. *Journal of the American Statistical Association* **109**, 1546–1564.

Mardia, K. V. and Marshall, R. J. (1984). Maximum likelihood estimation of models for residual covariance in spatial regression. *Biometrika* **71**, 135–146.

Pace, L., Salvan, A. and Sartori, N. (2019). Efficient composite likelihood for a scalar parameter of interest. *Stat* **8**, e222.

Singhal, K. and Kumar, S. (2019). On efficiency and robustness of estimators for a spherical location. *Statistics: A Journal of Theoretical and Applied Statistics* **53**, 601–629.

Stein, M. L., Chi, Z. and Welty, L. J. (2004). Approximating likelihoods for large spatial data sets. *Journal of the Royal Statistical Society. Series B (Statistical Methodology)* **66**, 275–296.

Varin, C., Reid, N. and Firth, D. (2011). An overview of composite likelihood methods. *Statistica Sinica* **21**, 5–42.

Xu, X., Reid, N. and Xu, L. (2024). A note on information bias and efficiency of composite likelihood. *Statistica Sinica* **34**, 523–526.

Ying, Z. (1991). Asymptotic properties of a maximum likelihood estimator with data from a Gaussian process. *Journal of Multivariate Analysis* **36**, 280–296.

Ying, Z. (1993). Maximum likelihood estimation of parameters under a spatial sampling scheme. *The Annals of Statistics* **21**, 1567–1590.

Zhang, H. and Zimmerman, D. L. (2005). Towards reconciling two asymptotic frameworks in spatial statistics. *Biometrika* **92**, 921–936.

Zheng, Y. and Zhu, J. (2012). On the asymptotics of maximum likelihood estimation for spatial linear models on a lattice. *The Indian Journal of Statistics* **74**, 29–56.

Nelson J. Y. Chua

Research School of Finance, Actuarial Studies & Statistics, The Australian National University, Canberra, ACT 2600, Australia.

E-mail: nelson.chua@anu.edu.au

Francis K. C. Hui

Research School of Finance, Actuarial Studies & Statistics, The Australian National University, Canberra, ACT 2600, Australia.

E-mail: francis.hui@anu.edu.au

A. H. Welsh

Research School of Finance, Actuarial Studies & Statistics, The Australian National University, Canberra, ACT 2600, Australia.

E-mail: alan.welsh@anu.edu.au

(Received August 2020; accepted July 2022)