The Annals of Statistics
2000, Vol. 28, No. 3, 880–904

ESTIMATING STRUCTURED CORRELATION MATRICES IN SMOOTH GAUSSIAN RANDOM FIELD MODELS

By Wei-Liem Loh[1] and Tao-Kai Lam

National University of Singapore

Dedicated to Charles M. Stein on his eightieth birthday

This article considers the estimation of structured correlation matrices in infinitely differentiable Gaussian random field models. The problem is essentially motivated by the stochastic modeling of smooth deterministic responses in computer experiments. In particular, the log-likelihood function is determined explicitly in closed-form and the sieve maximum likelihood estimators are shown to be strongly consistent under mild conditions.

1. Introduction. In the modeling of computer experiments, it has become quite common practice to approximate the deterministic response as a realization of a stochastic process. In this regard, Sacks, Welch, Mitchell and Wynn (1989) and Sacks, Schiller and Welch (1989) proposed modeling using a Gaussian random field $X(t),\ t \in [0,1]^d$, with a multiplicative covariance function

$$\operatorname{Cov}(X(x), X(y)) = \sigma^2 \prod_{u=1}^{d} \exp(-\theta_u |x_u - y_u|^\gamma) \qquad \forall\ x = (x_1, \dots, x_d)',\ \ y = (y_1, \dots, y_d)' \in [0,1]^d,$$

where $\gamma \in (0, 2]$, $\theta_1, \dots, \theta_d$ and σ^2 are strictly positive parameters. Ying (1991, 1993) investigated the asymptotic properties of the maximum likelihood estimators for the parameters of the covariance function when $\gamma = 1$. In particular, he proved that the estimators are strongly consistent and asymptotically normal under mild conditions. Unfortunately, his proof rests crucially on the Markov property of the Gaussian process when $\gamma = 1$ and it is not clear whether the method can be extended to $\gamma \neq 1$. Recently, van der Vaart (1996) showed that when $\gamma = 1$ and $d = 2$, the maximum likelihood estimators are also asymptotically efficient.

It is of interest to note that with probability 1, the Gaussian random field $X(t),\ t \in [0,1]^d$, is continuous but not mean square differentiable when $\gamma = 1$ and is infinitely mean square differentiable when $\gamma = 2$ [see Stein (1989) and Ying (1993)]. Thus the case $\gamma = 2$ may be especially appropriate when the deterministic response of a computer experiment is known a priori to be smooth.

Received March 1997; revised March 2000.

[1]Supported in part by the National Science and Technology Board of Singapore.

AMS 1991 *subject classifications*. Primary 62D05; secondary 62E20, 62G15.

Key words and phrases. Computer experiment, sieve maximum likelihood estimation, smooth Gaussian random field, strong consistency, structured correlation matrix.

As Ying [(1991), page 295] noted, mainly due to its fixed domain $[0, 1]^d$, the distinctive feature of these problems is that the statistical dependence among the observations is very strong and does not weaken with the asymptotics. In fact the dependence actually becomes stronger as sample size increases. This is more so for the smooth case $\gamma = 2$ than for $\gamma = 1$. In particular Ying obtained the log-likelihood function when $\gamma = 1$ and wrote that it is very unlikely that we can have a similar expression for the log-likelihood function when $\gamma = 2$. He further asked whether in this case the maximum likelihood estimators for $\theta_1, \dots, \theta_d$ and σ^2 are consistent.

This article considers the case $\gamma = 2$ and reports on some partial answers to the questions raised in the previous paragraph. More precisely, we shall focus on the following problem. Let $X(t),\ t \in [0, 1]^d$, denote a zero-mean Gaussian random field with multiplicative covariance function

$$\operatorname{Cov}(X(x), X(y)) = \sigma^2 \prod_{u=1}^{d} \exp[-\theta_u (x_u - y_u)^2] \quad \forall\ x = (x_1, \dots, x_d)',\ \ y = (y_1, \dots, y_d)' \in [0, 1]^d, \tag{1}$$

where $\theta_1, \dots, \theta_d$ and σ^2 are strictly positive unknown parameters. We are concerned with the estimation of $\theta_1, \dots, \theta_d$, the parameters of the correlation function, using observations that are taken from the above random field on a regular lattice, that is,

$$\left\{ X\left(\frac{i_1}{n}, \dots, \frac{i_d}{n}\right) : 1 \le i_u \le n,\ 1 \le u \le d \right\}, \tag{2}$$

where n is a strictly positive integer.

For simplicity, we order the elements of the set in (2) lexicographically as a $n^d \times 1$ column vector $\widetilde{X}_n$. Thus the element $X(i_1/n, \dots, i_d/n)$ precedes the element $X(j_1/n, \dots, j_d/n)$ in $\widetilde{X}_n$ if and only if there exists a $1 \le k \le d$ such that $i_u = j_u$ whenever $1 \le u < k$ and $i_k < j_k$. Then the covariance matrix $\sum_{\theta_1, \dots, \theta_d, \sigma^2, n}$ of $\widetilde{X}_n$ is given by

$$\sum\nolimits_{\theta_1, \dots, \theta_d, \sigma^2, n} = \sigma^2 \bigotimes_{u=1}^{d} R_{\theta_u, n},$$

where the symbol $\bigotimes$ denotes the Kronecker product [see Anderson (1984), page 599] and for each $1 \le u \le d$, $R_{\theta_u, n}$ denotes the $n \times n$ matrix whose (i, j)th element is $\exp[-\theta_u (i - j)^2/n^2]$. The estimation of $\theta_1, \dots, \theta_d$ now reduces to the estimation of the structured correlation matrix $\bigotimes_{u=1}^{d} R_{\theta_u, n}$ (and hence the title of this article).

Since $\widetilde{X}_n \sim N_{n^d}(0, \sum_{\theta_1, \dots, \theta_d, \sigma^2, n})$, the n^d-variate normal distribution with mean 0 and covariance matrix $\sum_{\theta_1, \dots, \theta_d, \sigma^2, n}$, the likelihood function

$L_n(\theta_1, \ldots, \theta_d, \sigma^2)$ is given by

$$L_n(\theta_1, \ldots, \theta_d, \sigma^2) = (2\pi)^{-n^d/2} \left| \sum\nolimits_{\theta_1, \ldots, \theta_d, \sigma^2, n} \right|^{-1/2} \exp\left(- \widetilde{X}_n' \sum\nolimits_{\theta_1, \ldots, \theta_d, \sigma^2, n}^{-1} \widetilde{X}_n / 2 \right)$$

[see, e.g., Anderson (1984)] and consequently the log-likelihood function $l_n(\theta_1, \ldots, \theta_d, \sigma^2)$ satisfies

$$\begin{aligned} 2l_n(\theta_1, \ldots, \theta_d, \sigma^2) = -n^d \log(2\pi) - n^d \log(\sigma^2) - \log \left| \bigotimes_{u=1}^{d} R_{\theta_u, n} \right| \\ - \sigma^{-2} \widetilde{X}_n' \left(\bigotimes\nolimits_{u=1}^{d} R_{\theta_u, n} \right)^{-1} \widetilde{X}_n. \end{aligned} \tag{3}$$

The rest of this article is organized as follows. Section 2 of this article deals with the computation of $|\bigotimes_{u=1}^{d} R_{\theta_u, n}|$ and $(\bigotimes_{u=1}^{d} R_{\theta_u, n})^{-1}$ and the methods used are combinatorial rather than analytical or statistical in nature. The key idea here (see Proposition 1) is the observation that the correlation matrix $\bigotimes_{u=1}^{d} R_{\theta_u, n}$ can be compactly factorized into a product of triangular matrices via the Cholesky decomposition. This dramatically eases the computation of the determinant and inverse of $\bigotimes_{u=1}^{d} R_{\theta_u, n}$ in Corollaries 1 and 2, respectively. These matrix results appear to be new and are especially satisfying as the relatively simple closed-form expressions for the determinant, inverse and Cholesky decomposition of $\bigotimes_{u=1}^{d} R_{\theta_u, n}$ are exact (not approximations). They may also be of independent interest in combinatorics and linear algebra. The explicit log-likelihood function $l_n(\theta_1, \ldots, \theta_d, \sigma^2)$ is then derived and stated in Theorem 1.

Section 3 defines a sieve maximum likelihood estimator $(\hat{\theta}_{1,n}, \ldots, \hat{\theta}_{d,n})$ for $(\theta_1, \ldots, \theta_d)$. Theorem 2 establishes the strong consistency of $(\hat{\theta}_{1,n}, \ldots, \hat{\theta}_{d,n})$ under mild conditions and also provides an upper bound for the rate of strong convergence of $(\hat{\theta}_{1,n}, \ldots, \hat{\theta}_{d,n})$.

Section 4 discusses a number of related issues not addressed in this article such as estimation of the variance σ^2 and "honest" maximum likelihood estimation (unlike sieve maximum likelihood estimation) of the parameters $(\theta_1, \ldots, \theta_d)$ of the correlation function.

The Appendix contains a number of technical lemmas that are needed in the proof of Theorem 2. Finally we remark that many of the calculations in this article are exact and have been checked (either numerically or symbolically) for correctness using the mathematical computation software system *Mathematica* [Wolfram (1996)].

2. Log-likelihood function. In this section we shall derive explicitly the log-likelihood function as given in (3). For simplicity we write for each $1 \le u \le d$,

$$R_{\theta_u, n} = \begin{pmatrix} 1 & w_u & w_u^4 & \cdots & w_u^{(n-1)^2} \\ w_u & 1 & w_u & \cdots & w_u^{(n-2)^2} \\ w_u^4 & w_u & 1 & \cdots & w_u^{(n-3)^2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ w_u^{(n-1)^2} & w_u^{(n-2)^2} & w_u^{(n-3)^2} & \cdots & 1 \end{pmatrix}, \tag{4}$$

where $w_u = \exp(-\theta_u/n^2)$. To compute the Cholesky decomposition, determinant and inverse of $\bigotimes_{u=1}^d R_{\theta_u, n}$, we shall draw on some techniques from number theory and combinatorics.

DEFINITION. The Gaussian polynomial $G(m, n; q)$ is defined by

$$G(m, n; q) = \frac{(1-q^{n-m+1})(1-q^{n-m+2})\cdots(1-q^n)}{(1-q)(1-q^2)\cdots(1-q^m)}$$

if $0 \le m \le n$ and $G(m, n; q) = 0$ otherwise.

REMARK. $G(m, n; q)$ are also known as q-binomial coefficients in the combinatorics literature, an excellent account of which can be found in Chapter 3 of Andrews (1976).

The following lemma will be used frequently in this section.

LEMMA 1. *For $n = 1, 2, \ldots$, we have*

$$\prod_{i=0}^{n-1}(x + q^i z) = \sum_{i=0}^{n} G(i, n; q) q^{i(i-1)/2} z^i x^{n-i} \qquad \forall\ x, q, z \in R.$$

We refer the reader to Goulden and Jackson [(1983), page 101] for the proof of Lemma 1.

PROPOSITION 1. (Cholesky decomposition). *Let $1 \le u \le d$, $w_u = \exp(-\theta_u/n^2)$ and $T_{\theta_u, n}$ be the lower triangular matrix with positive diagonal elements such that $T_{\theta_u, n}T'_{\theta_u, n} = R_{\theta_u, n}$. Then for all $1 \le i,\ j \le n$,*

$$(T_{\theta_u, n})_{i,j} = \begin{cases} w_u^{(i-j)^2}\{\prod_{l=i-j+1}^{i-1}(1-w_u^{2l})\}/\{\prod_{m=1}^{j-1}(1-w_u^{2m})\}^{1/2}, & \text{if } i \ge j, \\ 0, & \text{if } i < j, \end{cases} \tag{5}$$

and

$$(T^{-1}_{\theta_u, n})_{i,j} = \begin{cases} (-w_u)^{i-j} G(j-1, i-1; w_u^2)/\{\prod_{k=1}^{i-1}(1-w_u^{2k})\}^{1/2}, & \text{if } i \ge j, \\ 0, & \text{if } i < j. \end{cases} \tag{6}$$

Furthermore, we have

$$\bigotimes_{u=1}^{d} R_{\theta_u, n} = \left(\bigotimes_{u=1}^{d} T_{\theta_u, n}\right)\left(\bigotimes_{u=1}^{d} T_{\theta_u, n}\right)'.$$

PROOF. To prove (5), we observe from symmetry that it suffices to show that for all $1 \le j \le k \le n$,

$$(R_{\theta_u, n})_{j,k} = \sum_{i=1}^{j} (T_{\theta_u, n})_{j,i} (T_{\theta_u, n})_{k,i},$$

which is equivalent to

$$(7) \quad w_u^{(j-k)^2} = \sum_{i=1}^{j} \frac{w_u^{(i-j)^2+(i-k)^2} \{ \prod_{l=j-i+1}^{j-1} (1-w_u^{2l}) \} \{ \prod_{m=k-i+1}^{k-1} (1-w_u^{2m}) \}}{\prod_{r=1}^{i-1} (1-w_u^{2r})}.$$

Using Lemma 1 repeatedly, we observe that the right-hand side of (7) is equal to

$$\begin{aligned}
& w_u^{(j-k)^2} \sum_{i=0}^{j-1} w_u^{2(j-i-1)(k-i-1)} G(i, j-1; w_u^2) \prod_{m=k-i}^{k-1} (1-w_u^{2m}) \\
&= w_u^{(j-k)^2} \sum_{i=0}^{j-1} w_u^{2(j-i-1)(k-i-1)} G(i, j-1; w_u^2) \\
&\qquad \times \sum_{m=0}^{i} (-1)^m w_u^{m(m-1)+2m(k-i)} G(m, i; w_u^2) \\
&= w_u^{(j-k)^2} \sum_{m=0}^{j-1} (-1)^m w_u^{m(m+1)} G(m, j-1; w_u^2) \\
&\qquad \times \sum_{i=0}^{j-1-m} w_u^{2(j-1-i)(k-1-m-i)} G(i, j-1-m; w_u^2) \\
&= w_u^{(j-k)^2} \sum_{i=0}^{j-1} w_u^{2(j-1-i)(k-1-i)} G(i, j-1; w_u^2) \\
&\qquad \times \sum_{m=0}^{j-1-i} (-1)^m w_u^{m(m+1)} w_u^{-2m(j-1-i)} G(m, j-1-i; w_u^2) \\
&= w_u^{(j-k)^2} \sum_{i=0}^{j-1} w_u^{2(j-1-i)(k-1-i)} G(i, j-1; w_u^2) \prod_{m=0}^{j-2-i} \left(1 - w_u^{2(m+i+2-j)}\right) \\
&= w_u^{(j-k)^2}.
\end{aligned}$$

This proves (5). Next, to prove (6), it suffices to show that for all $1 \le k \le j \le n$,

$$(8) \quad \delta_{j,k} = \sum_{i=k}^{j} \frac{w_u^{(i-j)^2} (-w_u)^{i-k} G(k-1, i-1; w_u^2) \{ \prod_{l=j-i+1}^{j-1} (1-w_u^{2l}) \}}{\{ \prod_{m=1}^{i-1} (1-w_u^{2m}) \}},$$

where $\delta_{j,k}$ denotes the Kronecker delta. Using Lemma 1 again, we observe that the right-hand side of (8) is equal to

$$w_u^{(j-k)^2} G(k-1, j-1; w_u^2) \sum_{i=0}^{j-k} (-1)^i w_u^{i(i-1)+2i(k-j+1)} G(i, j-k; w_u^2)$$

$$= w_u^{(j-k)^2} G(k-1, j-1; w_u^2) \prod_{i=0}^{j-k-1} (1 - w_u^{2(i+k+1-j)})$$

$$= \delta_{j,k}.$$

Finally we observe from Anderson [(1984), page 600] that

$$\bigotimes_{u=1}^{d} R_{\theta_u, n} = \bigotimes_{u=1}^{d} T_{\theta_u, n} T'_{\theta_u, n}$$

$$= \left(\bigotimes_{u=1}^{d} T_{\theta_u, n} \right) \left(\bigotimes_{u=1}^{d} T_{\theta_u, n} \right)'.$$

This proves Proposition 1. □

COROLLARY 1. *The determinant of the matrix* $\bigotimes_{u=1}^{d} R_{\theta_u, n}$ *is given by*

$$\left| \bigotimes_{u=1}^{d} R_{\theta_u, n} \right| = \prod_{u=1}^{d} \left\{ \left\{ \prod_{k=1}^{n-1} [(1 - w_u^{2k})^{n-k}] \right\}^{n^{d-1}} \right\},$$

where $w_u = \exp(-\theta_u/n^2)$ *whenever* $1 \le u \le d$.

PROOF. We observe from Anderson [(1984), page 600] and Proposition 1 that

$$\left| \bigotimes_{u=1}^{d} R_{\theta_u, n} \right| = \prod_{u=1}^{d} \left[|R_{\theta_u, n}|^{n^{d-1}} \right]$$

$$= \prod_{u=1}^{d} \left\{ \left[\prod_{i=1}^{n} (T_{\theta_u, n})_{i,i} \right]^{2n^{d-1}} \right\}$$

$$= \prod_{u=1}^{d} \left\{ \left\{ \prod_{k=1}^{n-1} [(1 - w_u^{2k})^{n-k}] \right\}^{n^{d-1}} \right\}.$$

This proves Corollary 1. □

COROLLARY 2. *Let* $1 \le u \le d$ *and* $w_u = \exp(-\theta_u/n^2)$. *Then for* $1 \le i, \ j \le n$, *the* (i, j)*th element of the inverse of* $R_{\theta_u, n}$ *is given by*

$$(R^{-1}_{\theta_n, n})_{i,j} = (-1)^{i-j} \sum_{m=1}^{i \wedge j} \frac{w_u^{i+j-2m} G(i-m, n-m; w_u^2) G(j-m, n-m; w_u^2)}{\prod_{k=1}^{n-m} (1 - w_u^{2k})}. \tag{9}$$

Furthermore we have

$$\left(\bigotimes_{u=1}^{d} R_{\theta_u, n}\right)^{-1} = \bigotimes_{u=1}^{d} R_{\theta_u, n}^{-1}. \tag{10}$$

PROOF. Equation (9) follows from Proposition 1 and the observation that $R_{\theta_u, n}^{-1} = (T_{\theta_u, n}^{-1})'(T_{\theta_u, n}^{-1})$. Equation (10) is immediate from Anderson [(1984), page 600]. □

We now derive the log-likelihood function for $\theta_1, \ldots, \theta_d$ and σ^2.

THEOREM 1. *Let $\widetilde{X}_n$ be defined as in Section 1 and $w_u = \exp(-\theta_u/n^2)$, $1 \le u \le d$. Then the log-likelihood function $l_n(\theta_1, \ldots, \theta_d, \sigma^2)$ satisfies*

$$\begin{aligned}
&2l_n(\theta_1, \ldots, \theta_d, \sigma^2) \\
&\quad = -n^d \log(2\pi) - n^d \log(\sigma^2) - n^d \sum_{u=1}^{d} \sum_{k=1}^{n-1} \frac{n-k}{n} \log(1 - w_u^{2k}) \\
&\qquad - \frac{1}{\sigma^2} \sum_{i_1, \ldots, i_d,\; j_1, \ldots, j_d = 1}^{n} X\left(\frac{i_1}{n}, \ldots, \frac{i_d}{n}\right) X\left(\frac{j_1}{n}, \ldots, \frac{j_d}{n}\right) \\
&\qquad \times \prod_{u=1}^{d} (-1)^{i_u - j_u} \sum_{m=1}^{i_u \wedge j_u} \frac{w_u^{i_u + j_u - 2m} G(i_u - m, n - m; w_u^2) G(j_u - m, n - m; w_u^2)}{\prod_{r=1}^{n-m}(1 - w_u^{2r})}.
\end{aligned}$$

PROOF. Theorem 1 follows immediately from (3), Corollary 1 and Corollary 2. □

3. Sieve maximum likelihood estimation. Let $\widetilde{X}_n$ be defined as in Section 1 with covariance matrix $\sum_{\theta_1, \ldots, \theta_d, \sigma^2, n}$. This section establishes the strong consistency of a sieve maximum likelihood estimator for $(\theta_1, \ldots, \theta_d)$. For convenience we write for $1 \le u \le d$, $w_u = \exp(-\theta_u/n^2)$, $\tilde{w}_u = \exp(-\tilde{\theta}_u/n^2)$ and $\widetilde{Z}_n = (Z_1, \ldots, Z_{n^d})' = \sigma^{-1}(\bigotimes_{u=1}^{d} T_{\theta_u, n})^{-1} \widetilde{X}_n$ where $T_{\theta_u, n}$ is as in Proposition 1. Since $\widetilde{X}_n \sim N_{n^d}(0, \sigma^2 \bigotimes_{u=1}^{d} R_{\theta_u, n})$ and $T_{\theta_u, n} T'_{\theta_u, n} = R_{\theta_u, n}$, we have $\widetilde{Z}_n \sim N_{n^d}(0, I)$ where I is the $n^d \times n^d$ identity matrix.

LEMMA 2. *With the above notation, we have for all $1 \le j \le i \le n$ with $1 \le u \le d$,*

$$\begin{aligned}
\left(T_{\tilde{\theta}_u, n}^{-1} T_{\theta_u, n}\right)_{i, j} &= \left\{\frac{\prod_{s=1}^{j-1}(1 - w_u^{2s})}{\prod_{r=1}^{i-1}(1 - \tilde{w}_u^{2r})}\right\}^{1/2} \sum_{m=j}^{i} (-\tilde{w}_u)^{i-m} w_u^{(m-j)^2} \\
&\quad \times G(m-1, i-1; \tilde{w}_u^2) G(j-1, m-1; w_u^2).
\end{aligned} \tag{11}$$

Hence

$$\left(T^{-1}_{\tilde\theta_u, n} T_{\theta_u, n}\right)^2_{i,i} = \prod_{k=1}^{i-1}\left(\frac{1-w_u^{2k}}{1-\tilde w_u^{2k}}\right) \ge e^{-\theta_u}(\theta_u/\tilde\theta_u)^{i-1}, \tag{12}$$

for sufficiently large n *uniformly over* $\tilde\theta_u > 0$, $1 \le u \le d$ *and* $1 \le i \le n$.

PROOF. Equation (11) follows immediately from Proposition 1. Next observe that

$$\begin{aligned} 0 &\le \left(-\frac{2k\tilde\theta_u}{n^2}\right)^{-1}(\exp(-2k\tilde\theta_u/n^2)-1) \\ &= 1+\sum_{r=1}^{\infty}\left(-\frac{2k\tilde\theta_u}{n^2}\right)^r \frac{1}{(r+1)!} \le 1 \qquad \forall\ \tilde\theta_u > 0,\ 1 \le k \le n-1. \end{aligned}$$

Also

$$\begin{aligned} &\log\left\{\prod_{k=1}^{i-1}\left[1+\sum_{r=1}^{\infty}\left(-\frac{2k\theta_u}{n^2}\right)^r\frac{1}{(r+1)!}\right]\right\} \\ &\quad \ge \log\left\{\prod_{k=1}^{n-1}\left[1+\sum_{r=1}^{\infty}\left(-\frac{2k\theta_u}{n^2}\right)^r\frac{1}{(r+1)!}\right]\right\} \ge -\theta_u, \end{aligned}$$

for sufficiently large n uniformly over $1 \le i \le n$. Thus

$$\begin{aligned} \left(T^{-1}_{\tilde\theta_u, n} T_{\theta_u, n}\right)^2_{i,j} &= \prod_{k=1}^{i-1}\left(\frac{1-w_u^{2k}}{1-\tilde w_u^{2k}}\right) \\ &= \left(\frac{\theta_u}{\tilde\theta_u}\right)^{i-1}\left\{\prod_{k=1}^{i-1}\left[1+\sum_{r=1}^{\infty}\left(-\frac{2k\theta_u}{n^2}\right)^r\frac{1}{(r+1)!}\right]\right\} \\ &\qquad \times\left\{\prod_{l=1}^{i-1}\left[1+\sum_{s=1}^{\infty}\left(-\frac{2l\tilde\theta_u}{n^2}\right)^s\frac{1}{(s+1)!}\right]\right\}^{-1} \\ &\ge e^{-\theta_u}(\theta_u/\tilde\theta_u)^{i-1}, \end{aligned}$$

for sufficiently large n uniformly over $\tilde\theta_u > 0$, $1 \le u \le d$ and $1 \le i \le n$. This proves (12). □

Let ν, α_u, β_u, $0 \le u \le d$, be absolute constants such that $0 < \nu < 1$, $0 < \alpha_0 < \sigma^2 < \beta_0 < \infty$ and $0 < \alpha_u < \theta_u < \beta_u < \infty$, $1 \le u \le d$. We define a sieve on the parameter space of $(\theta_1, \dots, \theta_d)$ and Θ_n where

$$\Theta_n = \left\{\left(\frac{i_1}{n^{n^\nu}}, \dots, \frac{i_d}{n^{n^\nu}}\right) : \alpha_u \le \frac{i_u}{n^{n^\nu}} \le \beta_u,\ i_u \text{ integer}, 1 \le u \le d\right\}. \tag{13}$$

The sieve maximum likelihood estimator for $(\theta_1, \dots, \theta_d)$ is that element $(\hat{\theta}_{1,n}, \dots, \hat{\theta}_{d,n}) \in \Theta_n$ such that

$$\sup\{l_n(\hat{\theta}_{1,n}, \dots, \hat{\theta}_{d,n}, \tilde{\sigma}^2): \tilde{\sigma}^2 \in [\alpha_0, \beta_0]\}$$
$$= \sup\{l_n(\tilde{\theta}_1, \dots, \tilde{\theta}_d, \tilde{\sigma}^2): \tilde{\sigma}^2 \in [\alpha_0, \beta_0],\ (\tilde{\theta}_1, \dots, \tilde{\theta}_d) \in \Theta_n\}.$$

PROPOSITION 2. *Let Θ_n be as in (13) and $(\sigma^2, \theta_1, \dots, \theta_d) \in \prod_{u=0}^d (\alpha_u, \beta_u)$. Further let $\widetilde{X}_n$ be defined as in Section 1 with covariance matrix $\sum_{\theta_1,\dots,\theta_d,\sigma^2,n}$. Then for all $0 < \rho < 1 - \nu$ with probability 1,*

$$\frac{1}{n^d}\inf\Bigg\{l_n(\theta_1, \dots, \theta_d, \sigma^2) - l_n(\tilde{\theta}_1, \dots, \tilde{\theta}_d, \tilde{\sigma}^2): \tilde{\sigma}^2 \in [\alpha_0, \beta_0],$$
$$(\tilde{\theta}_1, \dots, \tilde{\theta}_d) \in \Theta_n \Big\backslash \prod_{u=1}^d (\theta_u - d^{-1}\theta_u n^{-\rho}, \theta_u + \theta_u n^{-\rho})\Bigg\} \to \infty,$$

as $n \to \infty$.

PROOF. First let τ and ξ be absolute constants such that $0 < \rho < \tau < \xi < 1$ and $0 < \nu < \xi - \rho$. Then for sufficiently large n,

$$\frac{1}{n^d}\inf\Bigg\{l_n(\theta_1,\dots,\theta_d,\sigma^2) - l_n(\tilde{\theta}_1,\dots,\tilde{\theta}_d,\tilde{\sigma}^2): \tilde{\sigma}^2 \in [\alpha_0,\beta_0],$$
$$(\tilde{\theta}_1,\dots,\tilde{\theta}_d) \in \Theta_n \Big\backslash \prod_{u=1}^d (\theta_u - d^{-1}\theta_u n^{-\rho}, \theta_u + \theta_u n^{-\rho})\Bigg\}$$
$$\geq \frac{1}{n^d}\min\Bigg[\inf\Bigg\{l_n(\theta_1,\dots,\theta_d,\sigma^2) - l_n(\tilde{\theta}_1,\dots,\tilde{\theta}_d,\tilde{\sigma}^2): \tilde{\sigma}^2 \in [\alpha_0,\beta_0],$$
$$(\tilde{\theta}_1,\dots,\tilde{\theta}_d) \in \Theta_n, \prod_{u=1}^d (\tilde{\theta}_u/\theta_u) \geq 1 + n^{-\tau}\Bigg\},$$
$$\inf\Bigg\{l_n(\theta_1,\dots,\theta_d,\sigma^2) - l_n(\tilde{\theta}_1,\dots,\tilde{\theta}_d,\tilde{\sigma}^2): \tilde{\sigma}^2 \in [\alpha_0,\beta_0], (\tilde{\theta}_1,\dots,\tilde{\theta}_d) \in \Theta_n,$$
$$\prod_{u=1}^d (\tilde{\theta}_u/\theta_u) \leq 1 + n^{-\tau}, \tilde{\theta}_v/\theta_v \leq 1 - d^{-1}n^{-\rho} \text{ for some } 1 \leq v \leq d\Bigg\}\Bigg].$$

Thus to prove Proposition 2, it suffices only to show that with probability 1,

$$\frac{1}{n^d}\inf\Bigg\{l_n(\theta_1, \dots, \theta_d, \sigma^2) - l_n(\tilde{\theta}_1, \dots, \tilde{\theta}_d, \tilde{\sigma}^2): \tilde{\sigma}^2 \in [\alpha_0, \beta_0],$$
$$(\tilde{\theta}_1, \dots, \tilde{\theta}_d) \in \Theta_n,\ \prod_{u=1}^d (\tilde{\theta}_u/\theta_u) \geq 1 + n^{-\tau}\Bigg\} \to \infty \tag{14}$$

and

$$\frac{1}{n^d}\inf\Big\{l_n(\theta_1,\dots,\theta_d,\sigma^2)-l_n(\tilde\theta_1,\dots,\tilde\theta_d,\tilde\sigma^2)\colon \tilde\sigma^2\in[\alpha_0,\beta_0],$$

$$(\tilde\theta_1,\dots,\tilde\theta_d)\in\Theta_n\ \prod_{u=1}^d(\tilde\theta_u/\theta_u)\le 1+n^{-\tau},\tag{15}$$

$$\tilde\theta_v/\theta_v\le 1-d^{-1}n^{-\rho}\text{ for some }1\le v\le d\Big\}\to\infty,$$

as $n\to\infty$.

Next we observe from (3) and Corollary 1 that

$$\begin{aligned}&\frac{2}{n^d}\big[l_n(\theta_1,\dots,\theta_d,\sigma^2)-l_n(\tilde\theta_1,\dots,\tilde\theta_d,\tilde\sigma^2)\big]\\&\quad=\log\left(\frac{\tilde\sigma^2}{\sigma^2}\right)+\sum_{u=1}^d\sum_{k=1}^{n-1}\frac{n-k}{n}\log\left(\frac{1-\exp(-2\tilde\theta_u k/n^2)}{1-\exp(-2\theta_u k/n^2)}\right)\\&\qquad+\frac{1}{n^d\tilde\sigma^2}\widetilde X'_n\left(\bigotimes_{u=1}^d R_{\tilde\theta_u,n}\right)^{-1}\widetilde X_n-\frac{1}{n^d\sigma^2}\widetilde X'_n\left(\bigotimes_{u=1}^d R_{\theta_u,n}\right)^{-1}\widetilde X_n.\end{aligned}\tag{16}$$

Using the strong law of large numbers, we further observe that as $n\to\infty$,

$$\frac{1}{n^d\sigma^2}\widetilde X'_n\left(\bigotimes_{u=1}^d R_{\theta_u,n}\right)^{-1}\widetilde X_n\to 1\tag{17}$$

almost surely, and

$$\begin{aligned}&\sum_{u=1}^d\sum_{k=1}^{n-1}\frac{n-k}{n}\log\left(\frac{1-\exp(-2\tilde\theta_u k/n^2)}{1-\exp(-2\theta_u k/n^2)}\right)\\&\quad=\left(\frac{n-1}{2}\right)\sum_{u=1}^d(1+o(1))\log\left(\frac{\tilde\theta_u}{\theta_u}\right),\end{aligned}\tag{18}$$

uniformly over $(\tilde\theta_1,\dots,\tilde\theta_d)\in\prod_{u=1}^d[\alpha_u,\beta_u]$.

To prove (14), we observe from (16), (17) and (18) that

$$\begin{aligned}&\liminf_{n\to\infty}\frac{2}{n^d}\inf\Big\{l_n(\theta_1,\dots,\theta_d,\sigma^2)-l_n(\tilde\theta_1,\dots,\tilde\theta_d,\tilde\sigma^2)\colon\tilde\sigma^2\in[\alpha_0,\beta_0],\\&\qquad\qquad(\tilde\theta_1,\dots,\tilde\theta_d)\in\Theta_n,\ \prod_{u=1}^d(\tilde\theta_u/\theta_u)\ge 1+n^{-\tau}\Big\}\\&\ge\liminf_{n\to\infty}\left\{\inf\left[\log\left(\frac{\alpha_0}{\sigma^2}\right)+\frac{n-1}{2}\sum_{u=1}^d\log\left(\frac{\tilde\theta_u}{\theta_u}\right)-1\colon\prod_{u=1}^d(\tilde\theta_u/\theta_u)\ge 1+n^{-\tau}\right]\right\}\\&=\lim_{n\to\infty}n^{1-\tau}/2=\infty\end{aligned}$$

almost surely. This proves (14).

To prove (15), we observe without loss of generality that it suffices to show that, with probability 1, for all $2 \le p \le d+1$,

$$\frac{1}{n^d}\inf\Big\{ l_n(\theta_1,\dots,\theta_d,\sigma^2) - l_n(\tilde\theta_1,\dots,\tilde\theta_d,\tilde\sigma^2) \colon \tilde\sigma^2 \in [\alpha_0,\beta_0], \tag{19}$$
$$(\tilde\theta_1,\dots,\tilde\theta_d) \in \Theta_n,\ \prod_{u=1}^d (\tilde\theta_u/\theta_u) \le 1+n^{-\tau},\ \tilde\theta_1/\theta_1 \le 1-d^{-1}n^{-\rho},$$
$$\theta_1/\tilde\theta_1 \ge \cdots \ge \theta_{p-1}/\tilde\theta_{p-1} \ge 1 \ge \theta_p/\tilde\theta_p \ge \cdots \ge \theta_d/\tilde\theta_d \Big\} \to \infty,$$

as $n \to \infty$. Let $(\bigotimes_{u=1}^d T^{-1}_{\tilde\theta_u,n}T_{\theta_u,n})_i$ denote the ith row of the $n^d \times n^d$ lower triangular matrix $\bigotimes_{u=1}^d T^{-1}_{\theta_u,n}T_{\tilde\theta_u,n}$ and let $\|\cdot\|$ denote the Euclidean norm in R^{n^d}. We observe from Corollary 2 and Anderson [(1984), page 600] that

$$\begin{aligned}
&\widetilde X_n'\Big(\bigotimes_{u=1}^d R_{\tilde\theta_u,n}\Big)^{-1}\widetilde X_n \\
&\quad= \sigma^2 \widetilde Z_n'\Big(\bigotimes_{u=1}^d T_{\theta_u,n}\Big)'\Big(\bigotimes_{u=1}^d T^{-1}_{\tilde\theta_u,n}\Big)'\Big(\bigotimes_{u=1}^d T^{-1}_{\tilde\theta_u,n}\Big)\Big(\bigotimes_{u=1}^d T_{\theta_u,n}\Big)\widetilde Z_n \\
&\quad= \sigma^2 \widetilde Z_n'\Big(\bigotimes_{u=1}^d T^{-1}_{\tilde\theta_u,n}T_{\theta_u,n}\Big)'\Big(\bigotimes_{u=1}^d T^{-1}_{\tilde\theta_u,n}T_{\theta_u,n}\Big)\widetilde Z_n \qquad (20)\\
&\quad= \sigma^2 \sum_{i=1}^{n^d}\Big\|\Big(\bigotimes_{u=1}^d T^{-1}_{\tilde\theta_u,n}T_{\theta_u,n}\Big)_i'\Big\|^2 \|\widetilde Z_n\|^2 \\
&\qquad\times\Big\{\sum_{j=1}^{n^d}\frac{(\bigotimes_{u=1}^d T^{-1}_{\tilde\theta_u,n}T_{\theta_u,n})_{i,j}}{\|(\bigotimes_{u=1}^d T^{-1}_{\tilde\theta_u,n}T_{\theta_u,n})_i'\|}\frac{Z_j}{\|\widetilde Z_n\|}\Big\}^2.
\end{aligned}$$

Next we observe from the definition of the Kronecker product that there is a bijective mapping $\psi\colon \prod_{u=1}^d\{1,\dots,n\} \to \{1,\dots,n^d\}$ such that $\psi(i_1,\dots,i_d)=i$ if and only if

$$\prod_{u=1}^d \Big(T^{-1}_{\tilde\theta_u,n}T_{\theta_u,n}\Big)_{i_u,i_u} = \Big(\bigotimes_{u=1}^d T^{-1}_{\tilde\theta_u,n}T_{\theta_u,n}\Big)_{i,i}.$$

Hence for $\tau < \xi < 1$ and $2 \le p \le d+1$, it follows from (12) and (20) that

$$\begin{aligned}
&\widetilde X_n'\Big(\bigotimes_{u=1}^d R_{\tilde\theta_u,n}\Big)^{-1}\widetilde X_n \\
&\quad\ge \sigma^2\exp(-(\theta_1+\cdots+\theta_d))\sum_{i_1,\dots,i_d=1}^n\Big\{\prod_{u=1}^d\Big(\frac{\theta_u}{\tilde\theta_u}\Big)^{i_u-1}\Big\}\|\widetilde Z_n\|^2 \\
&\qquad\times\Big\{\sum_{j=1}^{n^d}\frac{(\bigotimes_{u=1}^d T^{-1}_{\tilde\theta_u,n}T_{\theta_u,n})_{\psi(i_i,\dots,i_d),j}}{\|(\bigotimes_{u=1}^d T^{-1}_{\tilde\theta_u,n}T_{\theta_u,n})'_{\psi(i_i,\dots,i_d)}\|}\frac{Z_j}{\|\widetilde Z_n\|}\Big\}^2
\end{aligned}$$

$$
\begin{aligned}
&= \sigma^2 \exp(-(\theta_1 + \cdots + \theta_d)) \sum_{i_1,\ldots,i_d=1}^{n} \left\{ \prod_{u=1}^{d} \left(\frac{\theta_u}{\tilde{\theta}_u} \right)^{i_p - 1} \right\} \left\{ \prod_{u=1}^{p-1} \left(\frac{\theta_u}{\tilde{\theta}_u} \right)^{i_u - i_p} \right\} \\
&\quad \times \left\{ \prod_{u=p+1}^{d} \left(\frac{\tilde{\theta}_u}{\theta_u} \right)^{i_p - i_u} \right\} \| \widetilde{Z}_n \|^2 \left\{ \frac{(\bigotimes_{u=1}^{d} T_{\tilde{\theta}_u, n}^{-1} T_{\theta_u, n})_{\psi(i_i,\ldots,i_d)} \widetilde{Z}_n}{\|(\bigotimes_{u=1}^{d} T_{\tilde{\theta}_u, n}^{-1} T_{\theta_u, n})'_{\psi(i_i,\ldots,i_d)}\| \|\widetilde{Z}_n\|} \right\}^2 \\
&\geq \sigma^2 \exp(-(\theta_1 + \cdots + \theta_d)) \left\{ \prod_{u=1}^{d} \left(\frac{\theta_u}{\tilde{\theta}_u} \right)^{i_p^* - 1} \right\} \left\{ \prod_{u=1}^{p-1} \left(\frac{\theta_u}{\tilde{\theta}_u} \right)^{i_u^* - i_p^*} \right\} \\
&\quad \times \left\{ \prod_{u=p+1}^{d} \left(\frac{\tilde{\theta}_u}{\theta_u} \right)^{i_p^* - i_u^*} \right\} \| \widetilde{Z}_n \|^2 \left\{ \frac{(\bigotimes_{u=1}^{d} T_{\tilde{\theta}_u, n}^{-1} T_{\theta_u, n})_{\psi(i_i^*,\ldots,i_d^*)} \widetilde{Z}_n}{\|(\bigotimes_{u=1}^{d} T_{\tilde{\theta}_u, n}^{-1} T_{\theta_u, n})'_{\psi(i_i^*,\ldots,i_d^*)}\| \|\widetilde{Z}_n\|} \right\}^2,
\end{aligned}
\tag{21}
$$

for sufficiently large n uniformly over $\tilde{\theta}_u > 0,\ 1 \leq u \leq d$, where

$$
i_u^* = \lfloor 2^{d-u} n^{\xi} \rfloor \qquad \forall\ 1 \leq u \leq d.
$$

Here $\lfloor \cdot \rfloor$ denotes the greatest integer function. For simplicity we write

$$
\begin{aligned}
\Theta_n^* = \Bigg\{ (\tilde{\theta}_1, \ldots, \tilde{\theta}^d) \in \Theta_n \colon \prod_{u=1}^{d} (\tilde{\theta}_u / \theta_u) \leq 1 + n^{-\tau}, \tilde{\theta}_1/\theta_1 \leq 1 - d^{-1} n^{-\rho}, \\
\theta_1/\tilde{\theta}_1 \geq \cdots \geq \theta_{p-1}/\tilde{\theta}_{p-1} \geq 1 \geq \theta_p/\tilde{\theta}_p \geq \cdots \geq \theta_d/\tilde{\theta}_d \Bigg\}.
\end{aligned}
$$

Then for sufficiently large n,

$$
\begin{aligned}
&\inf_{(\tilde{\theta}_1,\ldots,\tilde{\theta}_d) \in \Theta_n^*} \left\{ \prod_{u=1}^{d} \left(\frac{\theta_u}{\tilde{\theta}_u} \right)^{i_p^* - 1} \right\} \left\{ \prod_{u=1}^{p-1} \left(\frac{\theta_u}{\tilde{\theta}_u} \right)^{i_u^* - i_p^*} \right\} \left\{ \prod_{u=p+1}^{d} \left(\frac{\tilde{\theta}_u}{\theta_u} \right)^{i_p^* - i_u^*} \right\} \\
&\quad \geq \left(\frac{1}{1 - d^{-1} n^{-\rho}} \right)^{i_1^* - 1} \bigwedge \min_{p \in \{2,\ldots,d\}} \left\{ \left(\frac{1}{1 + n^{-\tau}} \right)^{i_p^*} \left(\frac{1}{1 - d^{-1} n^{-\rho}} \right)^{i_1^* - i_p^*} \right\} \\
&\quad \geq \exp(d^{-1} 2^{d-3} n^{\xi - \rho}).
\end{aligned}
\tag{22}
$$

Consequently we observe from (21) and (22) that for sufficiently large n,

$$
\begin{aligned}
P\Bigg(&\inf_{(\tilde{\theta}_1,\ldots,\tilde{\theta}_d) \in \Theta_n^*} \left\{ \frac{\widetilde{X}'_n (\bigotimes_{u=1}^{d} R_{\tilde{\theta}_u, n})^{-1} \widetilde{X}_n}{\|\widetilde{Z}_n\|^2} \right\} \\
&\quad > \frac{\sigma^2}{\exp((\theta_1 + \cdots + \theta_d))} \exp[d^{-1} 2^{d-4} n^{\xi - \rho}] \Bigg)
\end{aligned}
$$

$$\geq P\Bigg(\inf_{(\tilde\theta_1,\ldots,\tilde\theta_d)\in\Theta_n^*} \Bigg\{ \frac{|(\bigotimes_{u=1}^d T_{\tilde\theta_u,n}^{-1} T_{\theta_u,n})_{\psi(i_1^*,\ldots,i_d^*)} \widetilde{Z}_n|}{\|(\bigotimes_{u=1}^d T_{\tilde\theta_u,n}^{-1} T_{\theta_u,n})'_{\psi(i_1^*,\ldots,i_d^*)}\| \|\widetilde{Z}_n\|} \Bigg\}^2 > \exp[-d^{-1}2^{d-4}n^{\xi-\rho}] \Bigg)$$

$$\geq 1 - \sum_{(\tilde\theta_1,\ldots,\tilde\theta_d)\in\Theta_n} P\Bigg(\frac{|(\bigotimes_{u=1}^d T_{\tilde\theta_u,n}^{-1} T_{\theta_u,n})_{\psi(i_1^*,\ldots,i_d^*)} \widetilde{Z}_n|}{\|(\bigotimes_{u=1}^d T_{\tilde\theta_u,n}^{-1} T_{\theta_u,n})'_{\psi(i_1^*,\ldots,i_d^*)}\| \|\widetilde{Z}_n\|} \leq \exp[-d^{-1}2^{d-5}n^{\xi-\rho}] \Bigg). \tag{23}$$

Next we shall use a geometrical argument to obtain a bound for the second term of the right-hand side of (23). For $m \geq 1$, let ∂_m denote the surface area of the unit hypersphere $S_{m-1} = \{x \in R^m\colon \sum_{i=1}^m x_i^2 = 1\}$. Then $\partial_m = 2\pi^{m/2}/\Gamma(m/2)$, [see, e.g., Anderson (1984), page 280].

Let $\{e_1, \ldots, e_{n^d}\}$ be an orthonormal basis for R^{n^d} such that

$$e_1 = \frac{(\bigotimes_{u=1}^d T_{\tilde\theta_u,n}^{-1} T_{\theta_u,n})'_{\psi(i_1^*,\ldots,i_d^*)}}{\|(\bigotimes_{u=1}^d T_{\tilde\theta_u,n}^{-1} T_{\theta_u,n})'_{\psi(i_1^*,\ldots,i_d^*)}\|},$$

and consider the transformation of this set of rectangular coordinates $\{(y_1, \ldots, y_{n^d}) \neq 0\colon -\infty < y_1, \ldots, y_{n^d} < \infty\}$ to its corresponding polar coordinates $\{(r, \phi_1, \ldots, \phi_{n^d-1})\colon r > 0, -\pi/2 < \phi_1, \ldots, \phi_{n^d-2} \leq \pi/2, \text{ and } -\pi < \phi_{n^d-1} \leq \pi\}$. The Jacobian of such a transformation is given by [see, e.g., Anderson (1984), pages 279–280]

$$r^{n^d-1}\cos^{n^d-2}(\phi_1)\cos^{n^d-3}(\phi_2)\cdots\cos(\phi_{n^d-2}).$$

Since $\widetilde{Z}_n \sim N_{n^d}(0, I)$, $\widetilde{Z}_n/\|\widetilde{Z}_n\|$ is uniformly distributed on the surface of S_{n^d-1} with probability 1 and hence using Stirling's approximation [see, e.g., Feller (1968), page 66],

$$\begin{aligned}
&P\Bigg(\frac{|(\bigotimes_{u=1}^d T_{\tilde\theta_u,n}^{-1} T_{\theta_u,n})_{\psi(i_1^*,\ldots,i_d^*)} \widetilde{Z}_n|}{\|(\bigotimes_{u=1}^d T_{\tilde\theta_u,n}^{-1} T_{\theta_u,n})'_{\psi(i_1^*,\ldots,i_d^*)}\| \|\widetilde{Z}_n\|} \leq \exp[-d^{-1}2^{d-5}n^{\xi-\rho}] \Bigg) \\
&\quad \leq 2\pi(\partial_{n^d})^{-1}\Bigg[\int_{\cos^{-1}\{\exp(-d^{-1}2^{d-5}n^{\xi-\rho})\}}^{\cos^{-1}\{-\exp(-d^{-1}2^{d-5}n^{\xi-\rho})\}} d\phi_1\Bigg] \\
&\qquad \times \Bigg[\int_{-\pi/2}^{\pi/2}\cos^{n^d-3}(\phi_2)\,d\phi_2\Bigg]\cdots\Bigg[\int_{-\pi/2}^{\pi/2}\cos(\phi_{n^d-2})\,d\phi_{n^d-2}\Bigg] \\
&\quad \leq \partial_{n^d-1}(\partial_{n^d})^{-1}3\exp(-d^{-1}2^{d-5}n^{\xi-\rho}) \\
&\quad \leq \exp(-d^{-1}2^{d-6}n^{\xi-\rho}),
\end{aligned} \tag{24}$$

for sufficiently large n, say $n \geq n_0$, uniformly over $(\tilde{\theta}_1, \dots, \tilde{\theta}_d) \in \prod_{u=1}^d [\alpha_u, \beta_u]$. Consequently, it follows from (24) that

$$\sum_{n=n_0}^{\infty} \sum_{(\tilde{\theta}_1,\dots,\tilde{\theta}_d)\in\Theta_n} P\left(\frac{|(\bigotimes_{u=1}^d T_{\tilde{\theta}_u,n}^{-1} T_{\theta_u,n})_{\psi(i_1^*,\dots,i_d^*)} \widetilde{Z}_n|}{\|(\bigotimes_{u=1}^d T_{\tilde{\theta}_u,n}^{-1} T_{\theta_u,n})'_{\psi(i_1^*,\dots,i_d^*)}\| \|\widetilde{Z}_n\|} \leq \exp[-d^{-1}2^{d-5}n^{\xi-\rho}] \right)$$

$$\leq \left(\prod_{u=1}^d \beta_u\right) \sum_{n=n_0}^{\infty} n^{dn^{\nu}} \exp(-d^{-1}2^{d-6}n^{\xi-\rho}) \qquad (25)$$

$$\leq \infty,$$

since $\nu < \xi - \rho$. By the Borel–Cantelli lemma [Chung (1974), page 73], we conclude from (23) and (25) that

$$P\left(\inf_{(\tilde{\theta}_1,\dots,\tilde{\theta}_d)\in\Theta_n^*} \left\{ \frac{\widetilde{X}'_n (\bigotimes_{u=1}^d R_{\tilde{\theta}_u,n})^{-1} \widetilde{X}_n}{\|\widetilde{Z}_n\|^2} \right\} \leq \frac{\sigma^2}{(\exp(\theta_1 + \cdots + \theta_d))} \exp[d^{-1}2^{d-4}n^{\xi-\rho}] \text{i.o.} \right) = 0,$$

and hence with probability 1,

$$\liminf_{n\to\infty} \exp(-d^{-1}2^{d-4}n^{\xi-\rho}) \inf_{(\tilde{\theta}_1,\dots,\tilde{\theta}_d)\in\Theta_n^*} \left\{ \frac{\widetilde{X}'_n (\bigotimes_{u=1}^d R_{\tilde{\theta}_u,n})^{-1} \widetilde{X}_n}{\|\widetilde{Z}_n\|^2} \right\} > \frac{\sigma^2}{(\exp(\theta_1 + \cdots + \theta_d))}. \qquad (26)$$

Now it follows from (16), (17), (18) and (26) that with probability 1,

$$\begin{aligned}
\liminf_{n\to\infty} \frac{2}{n^d} \inf \Big\{ & l_n(\theta_1, \dots, \theta_d, \sigma^2) - l_n(\tilde{\theta}_1, \dots, \tilde{\theta}_d, \tilde{\sigma}^2): \\
& \tilde{\sigma}^2 \in [\alpha_0, \beta_0],\ (\tilde{\theta}_1, \dots, \tilde{\theta}_d) \in \Theta_n, \\
& \prod_{u=1}^d (\tilde{\theta}_u/\theta_u) \leq 1 + n^{-\tau},\ \tilde{\theta}_1/\theta_1 \leq 1 - d^{-1}n^{-\rho}, \\
& \theta_1/\tilde{\theta}_1 \geq \cdots \geq \theta_{p-1}/\tilde{\theta}_{p-1} \geq 1 \geq \theta_p/\tilde{\theta}_p \geq \cdots \geq \theta_d/\tilde{\theta}_d \Big\} \\
\geq \liminf_{n\to\infty} \Big\{ & \log\left(\frac{\alpha_0}{\sigma^2}\right) - 1 + \frac{n-1}{2} \log\left(\prod_{u=1}^d \frac{\alpha_u}{\theta_u}\right) \\
& + \frac{\sigma^2}{\beta_0 \exp((\theta_1 + \cdots + \theta_d))} \exp(d^{-1}2^{d-4}n^{\xi-\rho}) \Big\} \\
= \infty. &
\end{aligned}$$

This proves (19) and the proof of Proposition 2 is complete. □

Theorem 2 below shows that the sieve maximum likelihood estimator $(\hat{\theta}_{1,n}, \ldots, \hat{\theta}_{d,n})$ for $(\theta_1, \ldots, \theta_d)$ is strongly consistent and also provides an upper bound for the rate of strong convergence of $(\hat{\theta}_{1,n}, \ldots, \hat{\theta}_{d,n})$ under mild conditions.

THEOREM 2. *Let Θ_n be as in* (13) *with* $4/5 \le \nu < 1$ *and* $(\sigma^2, \theta_1, \ldots, \theta_d) \in \prod_{u=0}^d (\alpha_u, \beta_u)$. *Further let $\widetilde{X}_n$ be defined as in Section* 1 *with covariance matrix* $\sum_{\theta_1,\ldots,\theta_d,\sigma^2,n}$. *Then for all* $0 < \rho < 1 - \nu$ *with probability* 1,

$$(\hat{\theta}_{1,n}, \ldots, \hat{\theta}_{d,n}) \in \prod_{u=1}^d (\theta_u - d^{-1}\theta_u n^{-\rho}, \theta_u + \theta_u n^{-\rho})$$

for sufficiently large n.

PROOF. We observe from Proposition 2 that to prove Theorem 2, it suffices to show that with probability 1,

$$\frac{1}{n^d}\left[l_n(\theta_1, \ldots, \theta_d, \sigma^2) - l_n(\tilde{\theta}_1, \ldots, \tilde{\theta}_d, \sigma^2)\right] \to 0, \tag{27}$$

as $n \to \infty$ uniformly over $(\tilde{\theta}_1, \ldots, \tilde{\theta}_d) \in \prod_{u=1}^d (\theta_u - n^{-n^\nu}, \theta_u + n^{-n^\nu})$. From (3) and Corollary 1, we have

$$\begin{aligned} &\frac{2}{n^d}[l_n(\theta_1, \ldots, \theta_d, \sigma^2) - l_n(\tilde{\theta}_1, \ldots, \tilde{\theta}_d, \sigma^2)] \\ &\quad = \sum_{u=1}^d \sum_{k=1}^{n-1} \frac{n-k}{n} \log\left(\frac{1-\exp(-2\tilde{\theta}_u k/n^2)}{1-\exp(-2\theta_u k/n^2)}\right) \\ &\qquad + \frac{1}{n^d\sigma^2}\widetilde{X}_n'\left(\bigotimes_{u=1}^d R_{\tilde{\theta}_u,n}\right)^{-1}\widetilde{X}_n - \frac{1}{n^d\sigma^2}\widetilde{X}_n'\left(\bigotimes_{u=1}^d R_{\theta_u,n}\right)^{-1}\widetilde{X}_n. \end{aligned} \tag{28}$$

We observe from (18) that

$$\sum_{u=1}^d \sum_{k=1}^{n-1} \frac{n-k}{n} \log\left(\frac{1-\exp(-2\tilde{\theta}_u k/n^2)}{1-\exp(-2\theta_u k/n^2)}\right) \to 0, \tag{29}$$

as $n \to \infty$ uniformly over $(\tilde{\theta}_1, \ldots, \tilde{\theta}_d) \in \prod_{u=1}^d (\theta_u - n^{-n^\nu}, \theta_u + n^{-n^\nu})$. Furthermore using (12), we have

$$\begin{aligned} &\left|\left(T_{\tilde{\theta}_u,n}^{-1} T_{\theta_u,n}\right)_{i,i} - 1\right| \\ &\quad = \left|\left\{\prod_{k=1}^{i-1}\left(1 - \frac{\exp(-2k\tilde{\theta}_u/n^2) - \exp(-2k\theta_u/n^2)}{1-\exp(-2k\theta_u/n^2)}\right)\right\}^{-1/2} - 1\right| \\ &\quad = \left|\left\{\prod_{k=1}^{i-1}\left[1 + \frac{(\tilde{\theta}_u - \theta_u)\exp(-2k\theta_u/n^2)\sum_{r=0}^\infty[-2k(\tilde{\theta}_u-\theta_u)/n^2)]^r[(r+1)!]^{-1}}{\theta_u\sum_{s=0}^\infty(-2k\theta_u/n^2)^s[(s+1)!]^{-1}}\right]\right\}^{-1/2} - 1\right| \\ &\quad \le \max\left\{1 - \left(1 + \frac{2}{n^{n^\nu}\theta_u}\right)^{-(i-1)/2}, \left(1 - \frac{2}{n^{n^\nu}\theta_u}\right)^{-(i-1)/2} - 1\right\} \\ &\quad \le \frac{2n}{n^{n^\nu}\theta_u} \end{aligned} \tag{30}$$

for sufficiently large n uniformly over $1 \le i \le n$ and $\tilde{\theta}_u \in (\theta_u - n^{-n^\nu}, \theta_u + n^{-n^\nu})$. Next we shall show that

$$\left|\left(T^{-1}_{\tilde{\theta}_u, n} T_{\theta_u, n}\right)_{i,j}\right| \le n^{-n^\nu/4}, \tag{31}$$

for sufficiently large n uniformly over $1 \le j < i \le n$ and $\tilde{\theta}_u \in (\theta_u - n^{-n^\nu}, \theta_u + n^{-n^\nu})$. We observe from (11) that for all $1 \le j \le i \le n$ with $1 \le u \le d$,

$$\begin{aligned} &\left\{\frac{[\prod_{r=1}^{i-j}(1-\tilde{w}_u^{2r})][\prod_{t=1}^{j-1}(1-w_u^{2t})]}{\prod_{s=i-j+1}^{i-1}(1-\tilde{w}_u^{2s})}\right\}^{1/2}\left(T^{-1}_{\tilde{\theta}_u, n} T_{\theta_u, n}\right)_{i,j} \\ &\quad = \sum_{m=0}^{i-j}(-\tilde{w}_u)^{i-j-m} w_u^{m^2} G(m, i-j; \tilde{w}_u^2)\left[\prod_{s=m+1}^{m+j-1}\frac{(1-w_u^{2s})}{(1-\tilde{w}_u^{2s})}\right]. \end{aligned} \tag{32}$$

Next define recursively for $k = 0, 1, \dots, i-j$,

$$\begin{aligned} &\eta_{i,j;m,i-j} = 1 \qquad \forall\, 0 \le m \le i-j, \\ &\eta_{i,j;m,i-j-k} = \kappa_{j,m}\eta_{i,j;m+1,i-j-k+1} - \lambda_m \eta_{i,j;m,i-j-k+1} \\ &\qquad\qquad\qquad\qquad\qquad\qquad \forall\, 0 \le m \le i-j-k, \end{aligned} \tag{33}$$

where

$$\begin{aligned} &\kappa_{j,m} = w_u^{2m+1}\left[\frac{\left(1-w_u^{2(m+j)}\right)\left(1-\tilde{w}_u^{2(m+1)}\right)}{\left(1-\tilde{w}_u^{2(m+j)}\right)\left(1-w_u^{2(m+1)}\right)}\right] \qquad \forall\, m = 0, 1, \dots, \\ &\lambda_m = \tilde{w}_u^{2m+1} \qquad \forall\, m = 0, 1, \dots. \end{aligned} \tag{34}$$

Then it follows from (32), (33) and induction that for all $1 \le j \le i \le n$, $0 \le k \le i-j$ with $1 \le u \le d$,

$$\begin{aligned} &\left\{\frac{[\prod_{r=1}^{i-j}(1-\tilde{w}_u^{2r})][\prod_{t=1}^{j-1}(1-w_u^{2t})]}{\prod_{s=i-j+1}^{i-1}(1-\tilde{w}_u^{2s})}\right\}^{1/2}\left(T^{-1}_{\tilde{\theta}_u, n} T_{\theta_u, n}\right)_{i,j} \\ &\quad = \sum_{m=0}^{i-j-k}(-\tilde{w}_u)^{i-j-k-m} w_u^{m^2} G(m, i-j-k; \tilde{w}_u^2) \\ &\qquad \times \left[\prod_{s=m+1}^{m+j-1}\frac{(1-w_u^{2s})}{(1-\tilde{w}_u^{2s})}\right]\eta_{i,j;m,i-j-k} \\ &\quad = \eta_{i,j,0,0}\prod_{s=1}^{j-1}\frac{1-w_u^{2s}}{1-\tilde{w}_u^{2s}}. \end{aligned} \tag{35}$$

To investigate the asymptotics (as $n \to \infty$) of $\eta_{i,\,j;0,0}$ when $j < i$, it is convenient to rewrite (33) in the following way. For each $m = 0, 1, \ldots$, define

$$
\begin{aligned}
\eta_{j+1,\,j,0,0}(m) &= \kappa_{j,m} - \lambda_m \qquad \forall\, 1 \le j \le n,\\
\eta_{i,\,j;0,0}(m) &= (\kappa_{j,m} - \lambda_m)\eta_{i-1,\,j;0,0}(m+1)\\
&\quad + \lambda_m[\eta_{i-1,\,j;0,0}(m+1) - \eta_{i-1,\,j;0,0}(m)]\\
&\qquad\qquad \forall\, 1 \le j < i-1 \le n-1.
\end{aligned}
\tag{36}
$$

Then $\eta_{i,\,j,0,0} = \eta_{i,\,j,0,0}(0)$ whenever $1 \le j < i \le n$. For example, we have

$$
\begin{aligned}
\eta_{j+2,\,j,0,0}(m) &= (\kappa_{j,m} - \lambda_m)(\kappa_{j,m+1} - \lambda_{m+1})\\
&\quad + \lambda_m(\kappa_{j,m+1} - \lambda_{m+1} - \kappa_{j,m} + \lambda_m),\\
\eta_{j+3,\,j,0,0}(m) &= (\kappa_{j,m} - \lambda_m)(\kappa_{j,m+1} - \lambda_{m+1})(\kappa_{j,m+2} - \lambda_{m+2})\\
&\quad + \lambda_{m+1}(\kappa_{j,m} - \lambda_m)(\kappa_{j,m+2} - \lambda_{m+2} - \kappa_{m+1} + \lambda_{m+1})\\
&\quad + \lambda_m(\kappa_{j,m+1} - \lambda_{m+1} - \kappa_{j,m} + \lambda_m)(\kappa_{j,m+2} - \lambda_{m+2})\\
&\quad + \lambda_m(\kappa_{j,m} - \lambda_m)(\kappa_{j,m+2} - \lambda_{m+2} - \kappa_{j,m+1} + \lambda_{m+1})\\
&\quad + \lambda_m(\lambda_{m+1} - \lambda_m)(\kappa_{j,m+2} - \lambda_{m+2} - \kappa_{j,m+1} + \lambda_{m+1})\\
&\quad + \lambda_m^2(\kappa_{j,m+2} - \lambda_{m+2} - \kappa_{j,m+1} + \lambda_{m+1} - \kappa_{j,m+1}\\
&\qquad\qquad + \lambda_{m+1} + \kappa_{j,m} - \lambda_m),
\end{aligned}
\tag{37}
$$

and so on. We observe via induction that $\eta_{i,\,j,0,0}(m)$ can be expressed as a sum of $(i-j)!$ terms where each term is a product of factors of the form λ_k, $\sum_{l=0}^{k}(-1)^l k!\lambda_{m+k-l}[l!(k-l)!]^{-1}$ or $\sum_{l=0}^{k}(-1)^l k!(\kappa_{j,m+k-l} - \lambda_{m+k-l})[l!(k-l)!]^{-1}$. We further observe that each term is also of degree $i-j$ in these factors. Next for each $k = 0, 1, \ldots$, we define the order of $\sum_{l=0}^{k}(-1)^l k!\lambda_{m+k-l}[l!(k-l)!]^{-1}$ to be k and the order of $\sum_{l=0}^{k}(-1)^l k!(\kappa_{j,m+k-l} - \lambda_{m+k-l})[l!(k-l)!]^{-1}$ to be $k+1$. Then each of the $(i-j)!$ terms of $\eta_{i,\,j,0,0}(m)$ has a sum of orders equal to $i-j$.

REMARK. As an illustration, we observe from (37) that $\eta_{j+2,\,j;0,0}$ and $\eta_{j+3,\,j;0,0}$ can be expressed as a sum of 2! and 3! terms, respectively. Furthermore, for example, $(\kappa_{j,m} - \lambda_m)(\kappa_{j,m+1} - \lambda_{m+1})$ is of order 2 and degree 2 whereas $\lambda_m(\lambda_{m+1} - \lambda_m)(\kappa_{j,m+2} - \lambda_{m+2} - \kappa_{j,m+1} + \lambda_{m+1})$ is of order 3 and degree 3.

Finally among the $(i-j)!$ terms of $\eta_{i,\,j;0,0}(m)$, let $a_{i,\,j,\,r,\,s,\,t}$ denote the number of terms having exactly $r-s$ factors of the type $\sum_{l=0}^{k_1}(-1)^l k_1!\lambda_{m+k_1-l}[l!(k_1-l)!]^{-1}$ for some $k_1 = 1, 2, \ldots$, and exactly s factors of the type $\sum_{l=0}^{k_2}(-1)^l k_2!(\kappa_{j,m+k_2-l} - \lambda_{m+k_2-l})[l!(k_2-1)!]^{-1}$ for some $k_2 = 0, 1, \ldots$, where the sum of the orders of these latter s factors is t. Then it follows from (36) that

$$
\begin{aligned}
&a_{j+1,\,j,\,1,\,1,\,1} = 1 \qquad \forall\, 1 \le j \le n,\\
(38)\qquad &a_{i,\,j,\,r,\,s,\,t} = a_{i-1,\,j,\,r-1,\,s-1,\,t-1} + s a_{i-1,\,j,\,r,\,s,\,t-1} + (r-s) a_{i-1,\,j,\,r,\,s,\,t}\\
&\qquad + (i-j-r) a_{i-1,\,j,\,r-1,\,s,\,t} \qquad \forall\, 1 \le j < i-1 \le n-1.
\end{aligned}
$$

Consequently, from Lemmas 3 and 4 (see the Appendix), we have

$$
\begin{aligned}
|\eta_{i,\,j;0,\,0}| &= |\eta_{i,\,j;0,\,0}(0)|\\
&\le \sum_{r,\,s,\,t} (\exp(3\theta_u/n))^{r-s} \left(\frac{4e\theta_u}{n^2}\right)^{i-j-t} \left(\frac{e^3 n}{\theta_u^2 n^{n^\nu}}\right)^{s} \left(\frac{17\theta_u}{n}\right)^{t} a_{i,\,j,\,r,\,s,\,t}\\
&\le \exp(3\theta_u) \left(\frac{17\theta_u}{n}\right)^{i-j} \sum_{r,\,s,\,t} \left(\frac{1}{n}\right)^{i-j-t} \left(\frac{e^3 n}{\theta_u^2 n^{n^\nu}}\right)^{s} a_{i,\,j,\,r,\,s,\,t},
\end{aligned}
$$

and using Lemma 5 (see the Appendix),

$$
\begin{aligned}
|\eta_{i,\,j;0,\,0}| &\le \exp(3\theta_u) \left(\frac{17\theta_u}{n}\right)^{i-j} [(i-j)!]^{1/4} \sum_{r,\,s,\,t=1}^{n} \frac{e^3 4^{i-j}}{\theta_u^2 n^{n^\nu - 1}}\\
(39)\qquad &\le \left(\frac{68\theta_u}{n}\right)^{i-j} [(i-j)!]^{1/4} \frac{\exp(3(\theta_u+1)) n^4}{\theta_u^2 n^{n^\nu}} \qquad \forall\, 1 \le j < i \le n,
\end{aligned}
$$

whenever $n \ge C_{\theta_u} = \max\{C^*_{\theta_u}, C^{**}_{\theta_u}, C^{***}_{\theta_u}\}$ uniformly over $\tilde{\theta}_u \in (\theta_u - n^{-n^\nu}, \theta_u + n^{-n^\nu})$. Since $(i-j)^{i-j} \le e^{i-j}(i-j)!$ [see, e.g., Feller (1968), page 54], it follows from (35) and (39) that

$$
\begin{aligned}
&\left|\left(T^{-1}_{\tilde{\theta}_u,\,n} T_{\theta_u,\,n}\right)_{i,\,j}\right|\\
&\quad = \left|\eta_{i,\,j;0,\,0} \left\{\prod_{s=1}^{j-1} \frac{1-w_u^{2s}}{1-\tilde{w}_u^{2s}}\right\} \left\{\frac{\prod_{s=i-j+1}^{i-1}(1-\tilde{w}_u^{2s})}{[\prod_{r=1}^{i-j}(1-\tilde{w}_u^{2r})][\prod_{t=1}^{j-1}(1-w_u^{2t})]}\right\}^{1/2}\right|\\
(40)\qquad &\quad \le 2|\eta_{i,\,j;0,\,0}| \left\{\frac{(i-1)!}{(j-1)!(i-j)!}\right\}^{1/2} \left\{\frac{n^{i-j}}{(2\theta_u)^{(i-j)/2}[(i-j)!]^{1/2}}\right\}\\
&\quad \le \frac{2\exp(3(\theta_u+1)) n^4 (68\theta_u)^{i-j}}{\theta_u^2 n^{i-j} n^{n^\nu}} \left\{\frac{(i-1)!}{(j-1)!}\right\}^{1/2} \left\{\frac{n^{i-j}}{(2\theta_u)^{(i-j)/2}[(i-j)!]^{3/4}}\right\}\\
&\quad \le \frac{2\exp(3(\theta_u+1)) n^4 n^{(i-j)/2}}{\theta_u^2 n^{n^\nu} (i-j)^{5(i-j)/8}} \exp\left[\frac{e^5 (68\theta_u)^8}{8(2\theta_u)^4}\right],
\end{aligned}
$$

for sufficiently large n uniformly over $1 \le j < i \le n$ and $\tilde{\theta}_u \in (\theta_u - n^{-n^\nu}, \theta_u + n^{-n^\nu})$.

CASE 1. Suppose $1 \le i - j < n^\nu$. Then clearly the right-hand side of (40) is bounded by $n^{-n^\nu/4}$ for sufficiently large n uniformly over $1 \le i - j < n^\nu$ and $\tilde{\theta}_u \in (\theta_u - n^{-n^\nu}, \theta_u + n^{-n^\nu})$.

CASE 2. Suppose $i - j \ge n^\nu$. Since $\nu \ge 4/5$, the right-hand side of (40) is bounded by

$$\frac{2\exp(3(\theta_u + 1))n^4}{\theta_u^2 n^{n^\nu}} \left(\frac{n}{n^{5\nu/4}}\right)^{(i-j)/2} \exp\left[\frac{e^5(68\theta_u)^8}{8(2\theta_u)^4}\right] \le n^{-n^\nu/4},$$

for sufficiently large n uniformly over $n^\nu \le i - j \le n$ and $\tilde{\theta}_u \in (\theta_u - n^{-n^\nu}, \theta_u + n^{-n^\nu})$.

Cases 1 and 2 prove that (31) holds. Finally it follows from (20), (30), (31) and the strong law of large numbers that with probability 1,

$$\frac{1}{n^d \sigma^2} \widetilde{X}_n' \left(\bigoplus_{u=1}^{d} R_{\tilde{\theta}_u, n}\right)^{-1} \widetilde{X}_n \to 1, \tag{41}$$

as $n \to \infty$ uniformly over $(\tilde{\theta}_1, \dots, \tilde{\theta}_d) \in \prod_{u=1}^d (\theta_u - n^{-n^\nu}, \theta_u + n^{-n^\nu})$. Now we conclude that (27) holds using (17), (28), (29) and (41). □

REMARK. It is evident from the proof of Theorem 2 that strong consistency holds for sieve maximum likelihood estimators of $(\theta_1, \dots, \theta_d)$ with coarser sieves than that given by (13) and $4/5 \le \nu < 1$. However we feel that this is somewhat academic since most statisticians would prefer to use a finer (as opposed to a coarser) sieve if both would lead to estimators possessing similar asymptotic properties. This is mainly due to the expectation of better finite sample performance from estimators derived from the finer sieve.

4. Final remarks. This article has shown that sieve maximum likelihood estimation of $(\theta_1, \dots, \theta_d)$ is strongly consistent when $\gamma = 2$. However we observe from the proof of Proposition 2 that it seems likely that the same result should hold for "honest" maximum likelihood estimation as well and that the use of a sieve is probably only a means to avoid further technical complications.

The consistency of the maximum likelihood estimate for the variance σ^2 is still an open question when $\gamma = 2$. In any case, the analysis of the log-likelihood function of $\theta_1, \dots, \theta_d$ and σ^2 indicates that in terms of accuracy, maximum likelihood estimation of σ^2 is at least an order more difficult than sieve maximum likelihood estimation of $(\theta_1, \dots, \theta_d)$.

A more detailed and precise study of the large sample behavior of the elements of the lower triangular matrix $\bigoplus_{u=1}^d T_{\tilde{\theta}_u, n}^{-1} T_{\theta_u, n}$ of Lemma 2 may help resolve the above issues.

APPENDIX

LEMMA 3. *Let λ_m be defined as in* (34). *Then there exists a constant $C^*_{\theta_u}$ (depending only on θ_u) such that*

$$\left|\sum_{l=0}^{k}\frac{(-1)^l k!\lambda_{m+k-l}}{l!(k-l)!}\right| \le \left(\frac{4e\theta_u}{n^2}\right)^k \exp(3\theta_u/n) \quad \forall 0 \le m < m+k \le n,$$

whenever $n \ge C^*_{\theta_u}$ uniformly over $\tilde{\theta} \in (\theta_u - n^{-n^{\nu}}, \theta_u + n^{-n^{\nu}})$.

PROOF. We observe that

$$\begin{aligned}\sum_{l=0}^{k}\frac{(-1)^l k!\lambda_{m+k-l}}{l!(k-l)!} &= \sum_{l=0}^{k}\frac{(-1)^{k-l}k!}{l!(k-l)!}\exp(-\tilde{\theta}_u n^{-2}(2m+1+2l))\\ &= \sum_{r=0}^{\infty}\left(-\frac{\tilde{\theta}_u}{n^2}\right)^r \sum_{s=0}^{r}\frac{(2m+1)^{r-s}2^s}{s!(r-s)!}\sum_{l=0}^{k}\frac{(-1)^{k-l}k!l^s}{l!(k-l)!}\\ &= \sum_{r=k}^{\infty}\left(-\frac{\tilde{\theta}_u}{n^2}\right)^r \sum_{s=k}^{r}\frac{(2m+1)^{r-s}2^s}{s!(r-s)!}\sum_{l=0}^{k}\frac{(-1)^{k-l}k!l^s}{l!(k-l)!}.\end{aligned} \tag{42}$$

Since $k^k \le e^k k!$ [see, e.g., Feller (1968), page 54], it follows from (42) that

$$\begin{aligned}\left|\sum_{l=0}^{k}\frac{(-1)^l k!\lambda_{m+k-l}}{l!(k-l)!}\right| &\le \sum_{r=k}^{\infty}\left(\frac{\tilde{\theta}_u}{n^2}\right)^r \sum_{s=0}^{r-k}\frac{2^k(2m+1)^{r-k-s}2^{s+k}k^{s+k}}{(s+k)!(r-k-s)!}\\ &\le \sum_{r=k}^{\infty}\left(\frac{\tilde{\theta}_u}{n^2}\right)^r \frac{(4e)^k(2m+1+2k)^{r-k}}{(r-k)!}\\ &\le \left(\frac{4e\tilde{\theta}_u}{n^2}\right)^k \sum_{r=0}^{\infty}\left(\frac{\tilde{\theta}_u(2n+1)}{n^2}\right)^r\frac{1}{r!}\\ &\le \left(\frac{4e\tilde{\theta}_u}{n^2}\right)^k \exp(\tilde{\theta}_u(2n+1)/n^2) \qquad \forall\, 0 \le m < m+k \le n.\end{aligned}$$

This proves Lemma 3. □

LEMMA 4. *Let $\kappa_{j,m}$ and λ_m be defined as in* (34). *Then there exists a constant $C^{**}_{\theta_u}$ (depending only on θ_u) such that*

$$\begin{aligned}&\left|\sum_{l=0}^{k}\frac{(-1)^l k!(\kappa_{j,m+k-l}-\lambda_{m+k-l})}{l!(k-l)!}\right|\\ &\qquad \le \frac{e^3 n}{\theta_u^2 n^{n^{\nu}}}\left(\frac{17\theta_u}{n}\right)^{k+1} \qquad \forall\, 0 \le m \le m+k \le n-j,\end{aligned}$$

whenever $n \ge C^{**}_{\theta_u}$ uniformly over $\tilde{\theta}_u \in (\theta_u - n^{-n^{\nu}}, \theta_u + n^{-n^{\nu}})$.

PROOF. It is convenient to expand $(1-w_u^{2(m+1)})^{-1}$ as

$$\left(1-w_u^{2(m+1)}\right)^{-1} = \frac{n^2}{2(m+1)\theta_u}\left[1+\sum_{r=1}^{\infty}\left(-\frac{2(m+1)\theta_u}{n^2}\right)^r\frac{1}{(r+1)!}\right]^{-1}$$
$$= \frac{n^2}{2(m+1)\theta_u}\sum_{s=0}^{\infty}\left(-\frac{2(m+1)\theta_u}{n^2}\right)^s\frac{B_s}{s!},$$

where B_s's are the Bernoulli numbers [see, e.g., Graham, Knuth and Patashnik (1994), pages 283–289]. Hence

$$\frac{1-\tilde{w}_u^{2(m+1)}}{1-w_u^{2(m+1)}} = \frac{\tilde{\theta}_u}{\theta_u}\sum_{r=0}^{\infty}\left(-\frac{2(m+1)}{n^2}\right)^r\sum_{s=0}^{r}\frac{\tilde{\theta}_u^s\theta_u^{r-s}B_{r-s}}{(s+1)!(r-s)!}.$$

Using a similar argument, we have

$$\frac{(1-w_u^{2(m+j)})(1-\tilde{w}_u^{2(m+1)})}{(1-\tilde{w}_u^{2(m+j)})(1-w_u^{2(m+1)})}$$
$$= \sum_{r_2=0}^{\infty}\left(-\frac{1}{n^2}\right)^{r_2}\sum_{r_4=0}^{r_2}(2m+1)^{r_4}$$
$$\times\sum_{r_1=0}^{r_2}\sum_{r_3=0\vee(r_4-r_2+r_1)}^{r_1\wedge r_4}\frac{r_1!(r_2-r_1)!(2j-1)^{r_1-r_3}}{r_3!(r_4-r_3)!(r_1-r_3)!(r_2-r_1+r_3-r_4)!}$$
$$\times\sum_{s_1=0}^{r_1}\frac{\theta_u^{s_1}\tilde{\theta}_u^{r_1-s_1}B_{r_1-s_1}}{(s_1+1)!(r_1-s_1)!}\sum_{s_2=0}^{r_2-r_1}\frac{\tilde{\theta}_u^{s_2}\theta_u^{r_2-r_1-s_2}B_{r_2-r_1-s_2}}{(s_2+1)!(r_2-r_1-s_2)!}. \tag{43}$$

Writing $\theta_u = \tilde{\theta}_u + \varepsilon_u$, it follows from (43) after some rather involved calculations that

$$\kappa_{j,m}-\lambda_m$$
$$= \tilde{w}_u^{2m+1}\left\{\frac{w_u^{2m+1}(1-w_u^{2(m+j)})(1-\tilde{w}_u^{2(m+1)})}{\tilde{w}_u^{2m+1}(1-\tilde{w}_u^{2(m+j)})(1-w_u^{2(m+1)})}-1\right\}$$
$$= \varepsilon_u\sum_{r=1}^{\infty}\left(-\frac{1}{n^2}\right)^r\sum_{r_4=0}^{r}(2m+1)^{r_4}\sum_{s=1\vee(r-r_4)}^{r}\frac{\tilde{\theta}_u^{r-s}}{(r-s)!}$$
$$\times\left\{\left(\sum_{r_2=r-r_4}^{s-1}\frac{\varepsilon_u^{s-1-r_2}}{(s-r_2)!}\sum_{r_1=0}^{r_2}\sum_{r_3=0\vee(r_1-r+r_4)}^{r_1\wedge(r_2-r+r_4)}\right.\right.$$
$$\times\frac{r_1!(r_2-r_1)!(2j-1)^{r_1-r_3}}{r_3!(r_2-r+r_4-r_3)!(r_1-r_3)!(r-r_4-r_1+r_3)!}$$
$$\left.\times\sum_{s_1=0}^{r_1}\frac{\theta_u^{s_1}\tilde{\theta}_u^{r_1-s_1}B_{r_1-s_1}}{(s_1+1)!(r_1-s_1)!}\sum_{s_2=0}^{r_2-r_1}\frac{\tilde{\theta}_u^{s_2}\theta_u^{r_2-r_1-s_2}B_{r_2-r_1-s_2}}{(s_2+1)!(r_2-r_1-s_2)!}\right)$$

$$+\left(\sum_{r_1=1}^{s}\sum_{r_3=0\vee(r_1-r+r_4)}^{r_1\wedge(s-r+r_4)}\frac{r_1!(s-r_1)!(2j-1)^{r_1-r_3}}{r_3!(s-r+r_4-r_3)!(r_1-r_3)!(r-r_4-r_1+r_3)!}\right.$$

$$\left.\times\sum_{s_1=0}^{r_1}\frac{\tilde{\theta}_u^{r_1-s_1}B_{r_1-s_1}}{(s_1+1)!(r_1-s_1)!}\sum_{s_3=1}^{s_1}\frac{s_1!\tilde{\theta}_u^{s_1-s_3}\varepsilon_u^{s_3-1}}{(s_1-s_3)!s_3!}\sum_{s_2=0}^{s-r_1}\frac{\tilde{\theta}_u^{s_2}\theta_u^{s-r_1-s_2}B_{s-r_1-s_2}}{(s_2+1)!(s-r_1-s_2)!}\right)$$

$$\left.+\left(\frac{s!}{(s-r+r_4)!(r-r_4)!}\sum_{s_2=0}^{s}\frac{\tilde{\theta}_u^{s_2}B_{s-s_2}}{(s_2+1)!(s-s_2)!}\sum_{s_3=1}^{s-s_2}\frac{(s-s_2)!\tilde{\theta}_u^{s-s_2-s_3}\varepsilon_u^{s_3-1}}{(s-s_2-s_3)!s_3!}\right)\right\}.$$

Consequently we observe that

$$\sum_{l=0}^{k}\frac{(-1)^l k!(\kappa_{j,m+k-l}-\lambda_{m+k-l})}{l!(k-l)!}$$

$$=\varepsilon_u\sum_{r=k\vee 1}^{\infty}\left(-\frac{1}{n^2}\right)^r\sum_{r_4=k}^{r}\sum_{r_5=k}^{r_4}\frac{r_4!(2m+1)^{r_4-r_5}2^{r_5}}{r_5!(r_4-r_5)!}\sum_{l=0}^{k}\frac{(-1)^{k-l}k!l^{r_5}}{l!(k-l)!}$$

$$\times\sum_{s=1\vee(r-r_4)}^{r}\frac{\tilde{\theta}_u^{r-s}}{(r-s)!}$$

$$\times\left\{\left(\sum_{r_2=r-r_4}^{s-1}\frac{\varepsilon_u^{s-1-r_2}}{(s-r_2)!}\sum_{r_1=0}^{r_2}\sum_{r_3=0\vee(r_1-r+r_4)}^{r_1\wedge(r_2-r+r_4)}\right.\right.$$

$$\times\frac{r_1!(r_2-r_1)!(2j-1)^{r_1-r_3}}{r_3!(r_2-r+r_4-r_3)!(r_1-r_3)!(r-r_4-r_1+r_3)!}$$

$$\left.\times\sum_{s_1=0}^{r_1}\frac{\theta_u^{s_1}\tilde{\theta}_u^{r_1-s_1}B_{r_1-s_1}}{(s_1+1)!(r_1-s_1)!}\sum_{s_2=0}^{r_2-r_1}\frac{\tilde{\theta}_u^{s_2}\theta_u^{r_2-r_1-s_2}B_{r_2-r_1-s_2}}{(s_2+1)!(r_2-r_1-s_2)!}\right)$$

$$+\left(\sum_{r_1=1}^{s}\sum_{r_3=0\vee(r_1-r+r_4)}^{r_1\wedge(s-r+r_4)}\frac{r_1!(s-r_1)!(2j-1)^{r_1-r_3}}{r_3!(s-r+r_4-r_3)!(r_1-r_3)!(r-r_4-r_1+r_3)!}\right.$$

$$\left.\times\sum_{s_1=0}^{r_1}\frac{\tilde{\theta}_u^{r_1-s_1}B_{r_1-s_1}}{(s_1+1)!(r_1-s_1)!}\sum_{s_3=1}^{s_1}\frac{s_1!\tilde{\theta}_u^{s_1-s_3}\varepsilon_u^{s_3-1}}{(s_1-s_3)!s_3!}\sum_{s_2=0}^{s-r_1}\frac{\tilde{\theta}_u^{s_2}\theta_u^{s-r_1-s_2}B_{s-r_1-s_2}}{(s_2+1)!(s-r_1-s_2)!}\right)$$

$$\left.+\left(\frac{s!}{(s-r+r_4)!(r-r_4)!}\sum_{s_2=0}^{s}\frac{\tilde{\theta}_u^{s_2}B_{s-s_2}}{(s_2+1)!(s-s_2)!}\sum_{s_3=1}^{s-s_2}\frac{(s-s_2)!\tilde{\theta}_u^{s-s_2-s_3}\varepsilon_u^{s_3-1}}{(s-s_2-s_3)!s_3!}\right)\right\}.$$

Since $|B_s/s!|\le 1$ for all $s=0,1,\dots,$ [see, e.g., Graham, Knuth and Patashnik (1994), page 286], for n sufficiently large such that $|\varepsilon_u/\tilde{\theta}_u|\le 1$, $1/2\le\theta_u/\tilde{\theta}_u\le 2$

and $|8\tilde{\theta}_u/n| < 1$, we have for all $0 \le m \le m+k \le n-j$,

$$\left|\sum_{l=0}^{k} \frac{(-1)^l k!(\kappa_{j,m+k-l} - \lambda_{m+k-l})}{l!(k-l)!}\right| \le \frac{2^k|\varepsilon_u|}{\tilde{\theta}_u} \sum_{r=k}^{\infty} 2e^3 \left(\frac{4\tilde{\theta}_u}{n^2}\right)^r (r-k+1)(2n)^r$$
$$+ \frac{2^k|\varepsilon_u|}{\tilde{\theta}_u} \sum_{r=k}^{\infty} e^{11/4} \left(\frac{4\tilde{\theta}_u}{n^2}\right)^r (r-k+1)(2n)^r$$
$$+ \frac{2^k|\varepsilon_u|}{\tilde{\theta}_u} \sum_{r=k}^{\infty} e^{3/4} \left(\frac{4\tilde{\theta}_u}{n^2}\right)^r (r-k+1)$$
$$\times [2(n-j)+1]^r$$
$$\le \frac{4e^3|\varepsilon_u|}{\tilde{\theta}_u} \left(\frac{16\tilde{\theta}_u}{n}\right)^k \sum_{r=k}^{\infty} (r-k+1) \left(\frac{8\tilde{\theta}_u}{n}\right)^{r-k}$$
$$= \frac{4e^3|\varepsilon_u|}{\tilde{\theta}_u} \left(\frac{16\tilde{\theta}_u}{n}\right)^k (1 - 8\tilde{\theta}_u n^{-1})^{-2}.$$

This proves Lemma 4. □

LEMMA 5. *Let $a_{i,j,r,s,t}$ be defined as in* (38). *Then there exists a constant $C^{***}_{\theta_u}$ (depending only on θ_u) such that*

$$(44) \qquad \frac{1}{[(i-j)!]^{1/4}} \left(\frac{e^3 n}{\theta_u^2 n^{n^\nu}}\right)^s \left(\frac{1}{n}\right)^{i-j-t} a_{i,j,r,s,t} \le \frac{e^3 4^{i-j}}{\theta_u^2 n^{n^\nu - 1}} \quad \forall\, 1 \le j < i \le n,$$

whenever $n \ge C^{***}_{\theta_u}$.

PROOF. Clearly, (44) holds when $i = j+1$ for all $n = 1, 2, \dots,$ since $a_{j+1,j,1,1,1} = 1$. Next we assume that $i > j+1$ and we divide the remainder of the proof into two cases.

CASE 1. Suppose that $s \ge (i-j)^{1/4}$. Since $a_{i,j,r,s,t} \le (i-j)!$ and $\nu \ge 4/5$,

$$\frac{1}{[(i-j)!]^{1/4}} \left(\frac{e^3 n}{\theta_u^2 n^{n^\nu}}\right)^s \left(\frac{1}{n}\right)^{i-j-t} a_{i,j,r,s,t} \le \left(\frac{e^3 n}{\theta_u^2 n^{n^\nu}}\right)^{(i-j)^{1/4}} \left(\frac{1}{n}\right)^{i-j-t} [(i-j)!]^{3/4}$$
$$\le \frac{e^3 4^{i-j}}{\theta_u^2 n^{n^\nu - 1}} \qquad \forall\, 1 < j+1 < i \le n,$$

whenever $n \ge C^{\dagger}_{\theta_u}$ for some constant $C^{\dagger}_{\theta_u}$ (depending only on θ_u).

CASE 2. Suppose that $s < (i-j)^{1/4}$. Let $C^{***}_{\theta_u}$ be a constant such that (44) is true for integers less than i, $C^{***}_{\theta_u} \ge C^{\dagger}_{\theta_u}$ and $e^3 \theta_u^{-2} n^{1-n^\nu} < 1$ whenever

$n \geq C^{***}_{\theta_u}$. Using (38) and induction, we have

$$
\begin{aligned}
&\frac{1}{[(i-j)!]^{1/4}}\left(\frac{e^3 n}{\theta_u^2 n^{n^\nu}}\right)^s\left(\frac{1}{n}\right)^{i-j-t} a_{i,\,j,\,r,\,s,\,t}\\
&\quad= \frac{1}{[(i-j)!]^{1/4}}\left(\frac{e^3 n}{\theta_u^2 n^{n^\nu}}\right)^s\left(\frac{1}{n}\right)^{i-j-t} a_{i-1,\,j,\,r-1,\,s-1,\,t-1}\\
&\qquad+ \frac{s}{[(i-j)!]^{1/4}}\left(\frac{e^3 n}{\theta_u^2 n^{n^\nu}}\right)^s\left(\frac{1}{n}\right)^{i-j-t} a_{i-1,\,j,\,r,\,s,\,t-1}\\
&\qquad+ \frac{1}{[(i-j)!]^{1/4}}\left(\frac{e^3 n}{\theta_u^2 n^{n^\nu}}\right)^s\left(\frac{1}{n}\right)^{i-j-t} (r-s)a_{i-1,\,j,\,r,\,s,\,t}\\
&\qquad+ \frac{1}{[(i-j)!]^{1/4}}\left(\frac{e^3 n}{\theta_u^2 n^{n^\nu}}\right)^s\left(\frac{1}{n}\right)^{i-j-t} (i-j-r)a_{i-1,\,j,\,r-1,\,s,\,t}\\
&\quad\leq \frac{1}{[(i-1-j)!]^{1/4}}\left(\frac{e^3 n}{\theta_u^2 n^{n^\nu}}\right)^{s-1}\left(\frac{1}{n}\right)^{i-j-t} a_{i-1,\,j,\,r-1,\,s-1,\,t-1}\\
&\qquad+ \frac{1}{[(i-1-j)!]^{1/4}}\left(\frac{e^3 n}{\theta_u^2 n^{n^\nu}}\right)^s\left(\frac{1}{n}\right)^{i-j-t} a_{i-1,\,j,\,r,\,s,\,t-1}\\
&\qquad+ \frac{1}{[(i-1-j)!]^{1/4}}\left(\frac{e^3 n}{\theta_u^2 n^{n^\nu}}\right)^s\left(\frac{1}{n}\right)^{i-1-j-t} a_{i-1,\,j,\,r,\,s,\,t}\\
&\qquad+ \frac{1}{[(i-1-j)!]^{1/4}}\left(\frac{e^3 n}{\theta_u^2 n^{n^\nu}}\right)^s\left(\frac{1}{n}\right)^{i-1-j-t} a_{i-1,\,j,\,r-1,\,s,\,t}\\
&\quad\leq \frac{4(e^3 4^{i-1-j})}{\theta_u^2 n^{n^\nu-1}} \qquad \forall\, 1 < i-j \leq n-j,
\end{aligned}
$$

whenever $n \geq C^{***}_{\theta_u}$. The last inequality uses the induction hypothesis that (44) is true for integers less than i. This proves Lemma 5. □

Acknowledgments. The authors thank Professor Hans R. Künsch, an Associate Editor and three referees for their detailed and constructive comments which led to a much improved proof of Proposition 2.

REFERENCES

ANDERSON, T. W. (1984). *An Introduction to Statistical Multivariate Analysis*, 2nd ed. Wiley, New York.

ANDREWS, G. E. (1976). *Theory of Partitions*. Addison-Wesley, London.

CHUNG, K. L. (1974). *A Course in Probability Theory*, 2nd ed. Academic Press, New York.

FELLER, W. (1968). *An Introduction to Probability Theory and Its Applications* **1**, 3rd ed. Wiley, New York.

GOULDEN, I. P. and JACKSON, D. M. (1983). *Combinatorial Enumeration*. Wiley, New York.

GRAHAM, R. L., KNUTH, D. E. and PATASHNIK, O. (1994). *Concrete Mathematics*, 2nd ed. Addison-Wesley, New York.

SACKS, J., SCHILLER, S. B. and WELCH, W. J. (1989). Designs for computer experiments. *Technometrics* **31** 41–47.

SACKS, J., WELCH, W. J., MITCHELL, T. J. and WYNN, H. P. (1989). Design and analysis of computer experiments (with discussion). *Statist. Sci.* **4** 409–435.

STEIN, M. L. (1989). Comment on "Design and analysis of computer experiments," by J. Sacks, W. J. Welch, T. J. Mitchell and H. P. Wynn. *Statist. Sci.* **4** 432–433.

VAN DER VAART, A. (1996). Maximum likelihood estimation under a spatial sampling scheme. *Ann. Statist.* **24** 2049–2057.

WOLFRAM, S. (1996). *The Mathematica Book*, 3rd ed. Cambridge Univ. Press.

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

YING, Z. (1993). Maximum likelihood estimation of parameters under a spatial sampling scheme. *Ann. Statist.* **21** 1567–1590.

DEPARTMENT OF STATISTICS
AND APPLIED PROBABILITY
NATIONAL UNIVERSITY OF SINGAPORE
KENT RIGDE
SINGAPORE 117543
REPUBLIC OF SINGAPORE
E-MAIL: wloh@stat.nus.edu.sg

EMC CORPORATION
171 SOUTH STREET
HOPKINTON, MASSACHUSETTS 01748
E-MAIL: tklam@alum.mit.edu