

Objective Bayesian analysis for autoregressive models with nugget effects

Cuirong Ren^{a,*}, Dongchu Sun^b

^a Statistical Programs, USPS, Washington DC 20260, USA

^b Department of Statistics, University of Missouri-Columbia, Columbia, MO 65211, USA

ARTICLE INFO

Article history:

Received 22 May 2012

Available online 19 November 2013

AMS subject classifications:

62F

62H

Keywords:

Autoregressive models

CAR model

SAR model

Jeffreys prior

Reference prior

Integrated likelihood

Propriety of posterior

ABSTRACT

The conditional autoregressive (CAR) and simultaneous autoregressive (SAR) models both have been used extensively for the analysis of spatial structure underlying lattice data in many areas, such as epidemiology, demographics, economics, and geography. Default Bayesian analyses have been conducted recently, but the Bayesian approach has not used or explored these two models with nugget effects. In this paper, we consider general autoregressive models including both CAR and SAR models. The Jeffreys-rule, independence Jeffreys, commonly used reference and “exact” reference priors are derived. The propriety of the marginal priors and joint posteriors is studied for a large class of objective priors. Various Jeffreys and reference priors are shown to yield improper posteriors and only the Jeffreys-rule and the “exact” reference priors yield proper posteriors. We make comparisons for these two objective priors using the frequentist coverage probabilities of the credible intervals. An illustration is given using a real spatial data-set.

Published by Elsevier Inc.

1. Introduction

The conditionally autoregressive (CAR) and simultaneously autoregressive (SAR) models are both very popular models used to model the spatial structure underlying lattice data. CAR models were introduced by Besag [2], while SAR models were originally developed as models on the doubly infinite regular lattice beginning with Whittle [21]. Since then, these models have been used to analyze data in diverse areas, such as epidemiology, demographics, economics and geography. Especially, in the last two decades, because of the convenient employment of Gibbs sampling and more general Markov Chain Monte Carlo (MCMC) methods for fitting certain classes of hierarchical spatial models, the study of both CAR and SAR models is resurgent.

The default (automatic) Bayesian analyses for CAR models were considered, perhaps for the first time, by De Oliveira [9], who studied two versions of Jeffreys priors: the Jeffreys-rule and the independence Jeffreys priors, and made comparisons between these two Jeffreys priors, the maximum likelihood estimates, and the uniform prior. He found that the frequentist properties of Bayesian inferences based on the independence Jeffreys prior were similar or better than other priors. Thus, he recommended the independence Jeffreys prior as the default objective prior. However, in the same paper, the results show that the independence Jeffreys prior does not always yield a proper posterior, so he pointed out that the posterior impropriety is a potential problem when the independence Jeffreys prior is used, although it seems unlikely to be encountered in practice. Consequently, Ren and Sun [16] considered the reference priors for the CAR models and concluded

* Corresponding author.

E-mail address: cuirongren@yahoo.com (C. Ren).

that only the Jeffreys-rule and the “exact” reference priors can always yield proper posteriors. They finally recommended two “exact” reference priors as the default prior based on the simulation study. This overcomes the difficulty of improper posterior with the independence Jeffreys and commonly used reference priors.

De Oliveira and Song [12] studied objective Bayesian analyses for SAR models. They also considered two versions of the Jeffreys prior. From the simulation results, they also recommended the independence Jeffreys prior as the default prior. They also identified the same problem in this case, that is, the independence Jeffreys prior does not always yield a proper posterior. Ren [15] proposed the reference priors for SAR models and found the reference priors perform better than the Jeffreys-rule prior, so he finally recommended the reference priors as the default priors.

However, the above models studied do not include the measurement error. As we know, the measurement error/microscale variation is present in many spatial data. In this paper, we will propose objective priors for autoregressive models with the nugget effect, treating both CAR and SAR models as special cases. Although the hierarchical linear mixed models considered by Sun et al. [19] included CAR and SAR models with the nugget effect, they only considered a class of improper noninformative priors and investigated propriety of posteriors under this class of priors. In addition, the class of noninformative priors they considered does not include the objective priors introduced in this paper.

The objective Bayesian approach for spatial data has been studied by many authors. See, for example, [11,14,10,17]. For models with nugget effects, a new parameter was usually introduced in order to conveniently obtain properties and the propriety of posteriors. Ren et al. [17], for example, introduced the noise-to-signal ratio as the new parameter in their paper. For the models we consider in this paper, we also introduce a new parameter, but interestingly, the new parameter is the signal-to-noise ratio instead.

The main object of this paper is to propose default Bayesian analyses for autoregressive models with nugget effects. Jeffreys priors, including Jeffreys-rule and independence Jeffreys priors, the reference priors obtained from asymptotic marginalization or the exact marginalization algorithm, are derived for the parameters in the models and conclusions on propriety of the resulting posteriors for the model parameters are established. Frequentist coverage of the credible intervals for numerical comparisons is described and examples are given based on simulated data to make comparisons between Jeffreys-rule and exact reference priors.

The organization of the paper is as follows. In Section 2, autoregressive models with nugget effects are introduced and the Jeffreys priors and various reference priors are derived. In Section 3, a class of objective priors are introduced, the behavior of integrated likelihood is investigated, and the proprieties of marginal priors and posteriors under all studied objective priors are studied. The two kinds of objective priors, the Jeffreys-rule and “exact” reference priors that always yield the proper posteriors, are compared numerically in terms of frequentist coverage of Bayesian credible intervals, and a real example is used for illustration in Section 4. Section 5 gives a brief summary and comments. Finally, some technical details are given in the Appendix. Several lemmas are present in Appendix A; and the proofs of some main results are present in Appendix B.

2. Autoregressive models and objective priors

2.1. Autoregressive models with nugget effects

We consider a Gaussian Markov random field, where the study area is partitioned into n regions, indexed by integers $1, 2, \dots, n$. For region i , the variable of interest, y_i , is observed through the following model,

$$y_i = \mathbf{x}_i \boldsymbol{\beta} + z_i + e_i, \quad (1)$$

where $\mathbf{x}_i = (x_{i1}, \dots, x_{ip})'$ is pre-specified, $\boldsymbol{\beta} = (\beta_1, \dots, \beta_p)' \in \mathbb{R}^p$ are unknown regression parameters, e_i are iid normal $(0, \delta_0)$ errors, and $(z_1, \dots, z_n)'$ are spatially correlated random effects satisfying

$$(z_i | z_j, j \neq i) \sim N \left(\sum_{j=1}^n W_{ij}(\rho) z_j, \delta_1 \right), \quad (2)$$

where $\boldsymbol{\beta} = (\beta_1, \dots, \beta_p)' \in \mathbb{R}^p$ are unknown regression parameters, $\delta_1 > 0$ and $W_{ij}(\rho)$ are covariance parameters, with $W_{ii}(\rho) = 0$ for all i . It is from [2] that the joint distribution of $\mathbf{z}$ is uniquely determined by the full conditional distributions (2). It is often that $W_{ij}(\rho)$ is a linear function of the weight C_{ij} . That is,

$$W_{ij}(\rho) = \rho C_{ij}, \quad \text{for all } i, j. \quad (3)$$

The matrix $\mathbf{C} = (C_{ij})_{n \times n}$ is often called a *Weight Matrix* or *Proximity Matrix*.

Often the proximity matrix $\mathbf{C}$ is symmetric and treated as known. Common choices of $\mathbf{C}$ are as follows.

- *Adjacency Matrix*: $C_{ij} = 1$ if region i and region j share common boundary.
- *k-neighbor Adjacency Matrix*: $C_{ij} = 1$ if region j is one of the k nearest neighbors of region i .
- *Distance Matrix*: C_{ij} = the distance between centroids of regions i and j .

The case of *Adjacency* matrix was introduced in [3] and the other two cases can be found in [4,18].

Note that (2) and the joint distribution based on (3) are equivalent to

$$\mathbf{z} \sim N_n(\mathbf{0}, \delta_1(\mathbf{I}_n - \rho \mathbf{C})^{-1}). \quad (4)$$

Here ρ is often called a 'spatial parameter'. Let $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n$ be the ordered eigenvalues of $\mathbf{C}$. Because $\text{tr}(\mathbf{C}) = 0$, it is clear that $\lambda_1 < 0 < \lambda_n$. The range of the spatial parameter ρ is $(\lambda_1^{-1}, \lambda_n^{-1})$, including 0 as an interior point.

Alternatively, a simultaneous AR model was proposed by Whittle [21],

$$\mathbf{z} = \rho \mathbf{C} \mathbf{z} + \boldsymbol{\epsilon}, \quad \boldsymbol{\epsilon} \sim N_n(\mathbf{0}, \delta_1 \mathbf{I}_n). \quad (5)$$

See also, [13]. This is equivalent to the model,

$$\mathbf{z} \sim N_n(\mathbf{0}, \delta_1 (\mathbf{I}_n - \rho \mathbf{C})^{-2}), \quad (6)$$

As pointed out in the introduction, since the measurement error is present in many spatial data, we study the model that includes the measurement error in this paper. In addition, we consider unified models including CAR and SAR models as special cases. In general, if we write $\mathbf{y} = (y_1, \dots, y_n)'$ and $\mathbf{X} = (x_{ij})_{n \times p}$, we could consider the following autoregressive model,

$$\mathbf{y} = \mathbf{X} \boldsymbol{\beta} + \mathbf{z} + \mathbf{e}, \quad (7)$$

where

$$\mathbf{z} \sim N_n(\mathbf{0}, \delta_1 (\mathbf{I}_n - \rho \mathbf{C})^{-h}), \quad \text{and} \quad \mathbf{e} \sim N_n(\mathbf{0}, \delta_0 \mathbf{I}_n), \quad (8)$$

where h is a fixed positive constant such as $h = 1, 2$. This is equivalent to considering the above models with a nugget effect for the marginal distribution for $\mathbf{y}$,

$$\mathbf{y} \sim N(\mathbf{X} \boldsymbol{\beta}, \delta_0 \mathbf{I}_n + \delta_1 (\mathbf{I}_n - \rho \mathbf{C})^{-h}). \quad (9)$$

In the geostatistical literature, δ_0 is often called the nugget effect.

Let $\eta = \delta_1 / \delta_0$, the so-called *signal-to-noise ratio*, and $\mathbf{G} = \mathbf{I}_n + \eta (\mathbf{I}_n - \rho \mathbf{C})^{-h}$. Then (9) becomes

$$\mathbf{y} \sim N(\mathbf{X} \boldsymbol{\beta}, \delta_0 \mathbf{G}). \quad (10)$$

Based on the observed data $\mathbf{y}$, the likelihood function of $(\rho, \eta, \delta_0, \boldsymbol{\beta})$ is given by

$$L(\rho, \eta, \delta_0, \boldsymbol{\beta}; \mathbf{y}) = \frac{1}{(2\pi \delta_0)^{n/2} |\mathbf{G}|^{1/2}} \exp \left\{ -\frac{(\mathbf{y} - \mathbf{X} \boldsymbol{\beta})' \mathbf{G}^{-1} (\mathbf{y} - \mathbf{X} \boldsymbol{\beta})}{2\delta_0} \right\}. \quad (11)$$

As we know, a prior is required for a full Bayesian analysis. As was discussed in the introduction, we seek objective priors for the unknown parameter $(\rho, \eta, \delta_0, \boldsymbol{\beta})$. We first give the Jeffreys and usual reference priors, and then derive two exact reference priors. These objective priors are improper and propriety of the posterior is studied.

2.2. Common objective priors

For the model (11) with parameters $(\rho, \eta, \delta_0, \boldsymbol{\beta})$, the Fisher information matrix is given in the following proposition. The proof is given in Appendix B.

Proposition 1. The Fisher information matrix of $(\rho, \eta, \delta_0, \boldsymbol{\beta})$ is of the form,

$$\boldsymbol{\Sigma}(\rho, \eta, \delta_0, \boldsymbol{\beta}) = \begin{pmatrix} \frac{1}{2} \boldsymbol{\Sigma}_3(\rho, \eta, \delta_0) & \mathbf{0} \\ \mathbf{0} & \frac{1}{\delta_0} \mathbf{X}' \mathbf{G}^{-1} \mathbf{X} \end{pmatrix}, \quad (12)$$

where $\boldsymbol{\Sigma}_3(\rho, \eta, \delta_0)$ is equal to

$$\begin{pmatrix} \sum_{i=1}^n \frac{(\lambda_i h \eta)^2}{(1 - \lambda_i \rho)^2 \{\eta + (1 - \lambda_i \rho)^h\}^2} & \sum_{i=1}^n \frac{\lambda_i h \eta}{(1 - \lambda_i \rho) \{\eta + (1 - \lambda_i \rho)^h\}^2} & \frac{1}{\delta_0} \sum_{i=1}^n \frac{\lambda_i h \eta}{(1 - \lambda_i \rho) \{\eta + (1 - \lambda_i \rho)^h\}} \\ \sum_{i=1}^n \frac{\lambda_i h \eta}{(1 - \lambda_i \rho) \{\eta + (1 - \lambda_i \rho)^h\}^2} & \sum_{i=1}^n \frac{1}{\{\eta + (1 - \lambda_i \rho)^h\}^2} & \frac{1}{\delta_0} \sum_{i=1}^n \frac{1}{\eta + (1 - \lambda_i \rho)^h} \\ \frac{1}{\delta_0} \sum_{i=1}^n \frac{\lambda_i h \eta}{(1 - \lambda_i \rho) \{\eta + (1 - \lambda_i \rho)^h\}} & \frac{1}{\delta_0} \sum_{i=1}^n \frac{1}{\eta + (1 - \lambda_i \rho)^h} & \frac{n}{\delta_0^2} \end{pmatrix}. \quad (13)$$

For convenience, we denote the 2×2 submatrix of $\boldsymbol{\Sigma}_3(\rho, \eta, \delta_0)$ by striking out the last row and column by $\boldsymbol{\Sigma}_2(\rho, \eta)$. The following proposition gives the expression of the determinants of $\boldsymbol{\Sigma}_2(\rho, \eta)$ and $\boldsymbol{\Sigma}_3(\rho, \eta, 1)$. The proof is given in Appendix B. In order to simplify the expressions, we denote $1/(1 - \lambda_i \rho)$ and $1/\{\eta + (1 - \lambda_i \rho)^h\}$ by a_i and b_i , $i = 1, \dots, n$, respectively.

Proposition 2. The determinants of $\Sigma_2(\rho, \eta)$ and $\Sigma_3(\rho, \eta, 1)$ are given in the following.

$$|\Sigma_2(\rho, \eta)| = (h\eta)^2 \sum_{i < j} (\lambda_i - \lambda_j)^2 (a_i a_j b_i b_j)^2, \quad (14)$$

$$|\Sigma_3(\rho, \eta, 1)| = \frac{(h\eta)^2}{n} \sum^* (b_i b_j b_k b_l)^2 (\lambda_i - \lambda_j)^2 (\lambda_k - \lambda_l)^2 \{a_i a_j (\eta + c_{ij}) d_{kl} - a_k a_l (\eta + c_{kl}) d_{ij}\}^2, \quad (15)$$

where $\sum^*$ is summation for the indices i, j, k, l satisfying $i < j, k < l, (i - k)n < l - j + i(i + 1)/2 - k(k + 1)/2$, and

$$c_{ij} = \frac{\lambda_i(1 - \lambda_j\rho)^{h+1} - \lambda_j(1 - \lambda_i\rho)^{h+1}}{\lambda_i - \lambda_j},$$

$$d_{ij} = \frac{(1 - \lambda_j\rho)^h - (1 - \lambda_i\rho)^h}{\lambda_i - \lambda_j}.$$

Remark 1. (i) For the CAR model, in this case $h = 1$, we have $c_{ij} = 1 - \lambda_i \lambda_j \rho^2$ and $d_{ij} = \rho$.

(ii) For the SAR model, in this case $h = 2$, $c_{ij} = 1 - 3\lambda_i \lambda_j \rho^2 + \lambda_i \lambda_j (\lambda_i + \lambda_j) \rho^3$ and $d_{ij} = \rho\{2 - (\lambda_i + \lambda_j)\rho\}$.

The most commonly used noninformative prior is the Jeffreys-rule prior. Alternatively, Jeffreys [8] also suggested the use of the independence Jeffreys priors, obtained by assuming that the groups of parameters are independent a priori and computing each marginal prior using Jeffreys-rule when the rest groups of parameters are assumed to be known. For example, the independence Jeffreys prior π^{J1} , obtained by assuming that (ρ, η, δ_0) and β are independent a priori, is proportional to the product of $\sqrt{|\Sigma_3(\rho, \eta, \delta_0)|}$ and a constant because the Jeffreys-rule prior of (ρ, η, δ_0) given β is equal to $\sqrt{|\Sigma_3(\rho, \eta, \delta_0)|}$ and the Jeffreys-rule prior of β given (ρ, η, δ_0) is a constant. Some slightly different computational formulas for Jeffreys and independence Jeffreys priors are given below and can be derived from Proposition 1 directly. The detailed proofs are thus omitted.

Proposition 3. Suppose that the sampling distribution is given by (11).

(a) The Jeffreys-rule prior $\pi^J(\rho, \eta, \delta_0, \beta)$ is given by

$$\pi^J(\rho, \eta, \delta_0, \beta) \propto \frac{1}{\delta_0^{1+p/2}} \sqrt{|\mathbf{X}'\mathbf{G}^{-1}\mathbf{X}| \cdot |\Sigma_3(\rho, \eta, 1)|}. \quad (16)$$

(b) The independence Jeffreys prior π^{J1} , obtained by assuming that (ρ, η, δ_0) and β are independent a priori, is

$$\pi^{J1}(\rho, \eta, \delta_0, \beta) \propto \frac{1}{\delta_0} \sqrt{|\Sigma_3(\rho, \eta, 1)|}. \quad (17)$$

(c) The independence Jeffreys prior π^{J2} , obtained by assuming that $(\rho, \eta), \delta_0$ and β are independent a priori, is

$$\pi^{J2}(\rho, \eta, \delta_0, \beta) \propto \frac{1}{\delta_0} \sqrt{|\Sigma_2(\rho, \eta)|}. \quad (18)$$

(d) The independence Jeffreys prior π^{J3} , obtained by assuming that (ρ, η) and (δ_0, β) are independent a priori, is

$$\pi^{J3}(\rho, \eta, \delta_0, \beta) \propto \frac{1}{\delta_0^{1+p/2}} \sqrt{|\Sigma_2(\rho, \eta)|}. \quad (19)$$

By specifying the order of parameters, we can obtain various reference priors based on the methods in [1]. The ordering of parameters reflects the inferential importance of parameters, in particular, the first parameters are the parameters of interest. For example, the ordering $\{\beta, (\rho, \eta, \delta_0)\}$ means β is the parameter of interest and (ρ, η, δ_0) is of less importance or is a nuisance parameter. Similarly, the ordering $\{(\rho, \eta), \beta, \delta_0\}$ means (ρ, η) is the parameter of interest, β is of less importance, and δ_0 is of least importance. We present these reference priors in the following proposition and the proof is given in Appendix B.

Proposition 4. Suppose that the sampling distribution is given by (11).

(a) The reference priors of $(\rho, \eta, \delta_0, \beta)$ with orderings $\{(\rho, \eta, \delta_0), \beta\}$, $\{\beta, (\rho, \eta, \delta_0)\}$, $\{(\rho, \eta), \delta_0, \beta\}$, $\{(\rho, \eta), \beta, \delta_0\}$, and $\{\beta, (\rho, \eta), \delta_0\}$ are all the same as π^{J1} .

(b) The reference priors of $(\rho, \eta, \delta_0, \beta)$ with orderings $\{\delta_0, (\rho, \eta), \beta\}$, $\{\delta_0, \beta, (\rho, \eta)\}$, $\{\beta, \delta_0, (\rho, \eta)\}$ are all the same as π^{J2} .

(c) The reference prior of $(\rho, \eta, \delta_0, \beta)$ with ordering $\{(\delta_0, \beta), (\rho, \eta)\}$ is the same as π^{J3} .

(d) The reference prior of $(\rho, \eta, \delta_0, \beta)$ with ordering $\{(\rho, \eta), (\delta_0, \beta)\}$ is

$$\pi^{R1}(\rho, \eta, \delta_0, \beta) \propto \frac{1}{\delta_0^{1+p/2}} \sqrt{|\Sigma_3(\rho, \eta, 1)|}. \quad (20)$$

2.3. The exact reference priors

The reference priors in the previous section are derived based on asymptotic marginal distributions. In this section, we consider “exact” reference priors. Note that the conditional Jeffreys-rule (or reference) prior for β is constant when (ρ, η, δ_0) is assumed to be known. Thus, under $\pi(\beta) = 1$, the marginal (integrated) likelihood function of (ρ, η, δ_0) , which we call as Type I marginal likelihood function, is

$$L_*(\rho, \eta, \delta_0; \mathbf{y}) = \int_{\mathbb{R}^p} L(\rho, \eta, \delta_0, \beta; \mathbf{y}) d\beta \propto \frac{1}{\delta_0^{(n-p)/2} |\mathbf{G}|^{1/2} |\mathbf{X}'\mathbf{G}^{-1}\mathbf{X}|^{1/2}} \exp\left(-\frac{S^2}{2\delta_0}\right), \quad (21)$$

where for fixed η and ρ , the generalized residual sum of squares, S^2 , and the generalized least squares estimator of β , $\hat{\beta}$, are given by

$$\hat{\beta} \equiv \hat{\beta}(\rho, \eta) = (\mathbf{X}'\mathbf{G}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{G}^{-1}\mathbf{y}, \quad (22)$$

$$S^2 \equiv S^2(\rho, \eta) = (\mathbf{y} - \mathbf{X}\hat{\beta})'\mathbf{G}^{-1}(\mathbf{y} - \mathbf{X}\hat{\beta}). \quad (23)$$

If we denote $\mathbf{G}^{-1} - \mathbf{G}^{-1}\mathbf{X}(\mathbf{X}'\mathbf{G}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{G}^{-1}$ by $\mathbf{R}_G$, then $S^2 = \mathbf{y}'\mathbf{R}_G\mathbf{y}$.

Furthermore, the reference prior of δ_0 is a scale invariant prior $1/\delta_0$, so we can derive the marginal likelihood of (ρ, η) , which we call as Type II marginal likelihood,

$$L_{**}(\rho, \eta; \mathbf{y}) = \int_0^\infty \int_{\mathbb{R}^p} L(\theta, \eta, \delta_0, \beta; \mathbf{y}) \frac{1}{\delta_0} d\beta d\delta_0 \propto |\mathbf{G}|^{-1/2} |\mathbf{X}'\mathbf{G}^{-1}\mathbf{X}|^{-1/2} (S^2)^{-(n-p)/2}. \quad (24)$$

Note that the constant prior for β is improper. Thus, as functions of the data $\mathbf{y}$, both L_* and L_{**} are improper, but from Theorem 1 (equivalence theory) in [17], $L_*(L_{**})$ is essentially a proper density of a $n-p$ ($n-p-1$) dimensional random variable and it is legitimate to find the Fisher information matrix of (ρ, η, δ_0) ((ρ, η)) based on the Type I(II) marginal likelihood and derive the reference priors, called Type I and II references, respectively.

The Fisher information matrix and the exact reference prior from the Type I likelihood function is given in the following proposition. The proof is similar to Proposition 1, so it is omitted here. In addition, one can verify that the Type II reference prior derived from (24) is the same as Type I's, so we do not state these results for Type II case.

Proposition 5. Suppose that the sampling distribution is given by (11).

(a) The Fisher information matrix of (ρ, η, δ_0) , based on the marginal likelihood function $L_*(\rho, \eta, \delta_0; \mathbf{y})$, is $\frac{1}{2}\Sigma_3^*(\rho, \eta, \delta_0)$, where $\Sigma_3^*(\rho, \eta, \delta_0)$ is equal to

$$\begin{pmatrix} \text{tr}\left(\mathbf{R}_G \frac{\partial}{\partial \rho} \mathbf{G}\right)^2 & \text{tr}\left\{\mathbf{R}_G(\mathbf{I}_n - \rho\mathbf{C})^{-h} \mathbf{R}_G \frac{\partial}{\partial \rho} \mathbf{G}\right\} & \frac{1}{\delta_0} \text{tr}\left(\mathbf{R}_G \frac{\partial}{\partial \rho} \mathbf{G}\right) \\ \text{tr}\left\{\mathbf{R}_G(\mathbf{I}_n - \rho\mathbf{C})^{-h} \mathbf{R}_G \frac{\partial}{\partial \rho} \mathbf{G}\right\} & \text{tr}\{\mathbf{R}_G(\mathbf{I}_n - \rho\mathbf{C})^{-h}\}^2 & \frac{1}{\delta_0} \text{tr}\{\mathbf{R}_G(\mathbf{I}_n - \rho\mathbf{C})^{-h}\} \\ \frac{1}{\delta_0} \text{tr}\left(\mathbf{R}_G \frac{\partial}{\partial \rho} \mathbf{G}\right) & \frac{1}{\delta_0} \text{tr}\{\mathbf{R}_G(\mathbf{I}_n - \rho\mathbf{C})^{-h}\} & \frac{n-p}{\delta_0^2} \end{pmatrix}. \quad (25)$$

(b) Type I exact reference prior with orderings $\{(\rho, \eta, \delta_0), \beta\}$ and $\{(\rho, \eta), \delta_0, \beta\}$ is

$$\pi_*^R(\rho, \eta, \delta_0, \beta) \propto \frac{1}{\delta_0} \sqrt{|\Sigma_3^*(\rho, \eta, 1)|}. \quad (26)$$

3. Propriety of prior and posterior

3.1. A general class of priors

We consider the following class of improper priors for $(\rho, \eta, \delta_0, \beta) \in \Omega = (\lambda_1^{-1}, \lambda_n^{-1}) \times (0, \infty)^2 \times \mathbb{R}^p$ of the form

$$\pi(\rho, \eta, \delta_0, \beta) \propto \frac{\pi(\rho, \eta)}{\delta_0^a}, \quad (\rho, \eta, \delta_0, \beta) \in (\lambda_1^{-1}, \lambda_n^{-1}) \times (0, \infty)^2 \times \mathbb{R}^p, \quad (27)$$

where $a \in \mathbb{R}$ is a hyperparameter and, loosely speaking, $\pi(\rho, \eta)$ is the marginal prior of (ρ, η) . Note that from (27) the priors for (ρ, η) , δ_0 and β are independent, and the marginal prior of β is a constant.

It is easy to see that all objective priors derived in Sections 2.2 and 2.3 belong to (27). The corresponding marginal prior for (ρ, η) is denoted by the same notation as the prior for $(\rho, \eta, \delta_0, \beta)$. For example, from (16) in Proposition 3, the Jeffreys-rule prior can be written as

$$\pi^J(\rho, \eta, \delta_0, \beta) = \frac{\pi^J(\rho, \eta)}{\delta_0^{1+p/2}},$$

where

$$\pi^J(\rho, \eta) \propto \sqrt{|\mathbf{X}'\mathbf{G}^{-1}\mathbf{X}| \cdot |\boldsymbol{\Sigma}_3(\rho, \eta, 1)|}.$$

Under the prior (27), a standard calculation yields

$$\int_0^\infty \int_{\mathbb{R}^p} L(\rho, \eta, \delta_0, \boldsymbol{\beta}; \mathbf{y}) \pi(\rho, \eta, \delta_0, \boldsymbol{\beta}) d\boldsymbol{\beta} d\delta_0 = L_{**a}(\rho, \eta; \mathbf{y}) \pi(\rho, \eta),$$

where

$$L_{**a}(\rho, \eta; \mathbf{y}) \propto \frac{1}{|\mathbf{G}|^{1/2} |\mathbf{X}'\mathbf{G}^{-1}\mathbf{X}|^{1/2} (S^2)^{(n-p)/2+a-1}}. \quad (28)$$

Under the prior (27), the joint posterior distribution of $(\rho, \eta, \delta_0, \boldsymbol{\beta})$ is proper if and only if

$$0 < \int_{\lambda_1^{-1}}^{\lambda_n^{-1}} \int_0^\infty L_{**a}(\rho, \eta; \mathbf{y}) \pi(\rho, \eta) d\rho d\eta < \infty. \quad (29)$$

The next section will give the behavior of the integrated likelihood function $L_{**a}(\rho, \eta; \mathbf{y})$ at λ_1^{-1} and λ_n^{-1} for ρ and at zero and infinity for η , which provides the key to determining the propriety of the posterior of $(\rho, \eta, \delta_0, \boldsymbol{\beta})$ under the class of the prior density (27).

3.2. Behavior of integrated likelihood L_{**a}

Proposition 6. Suppose that λ_1 and λ_n are simple eigenvalues. Then as $\rho \rightarrow \lambda_1^{-1}$, we have

$$|\mathbf{G}^{-1}| \propto \frac{(1 - \lambda_1 \rho)^h}{\{\eta + (1 - \lambda_1 \rho)^h\}(\eta + c)^{n-1}}, \quad (30)$$

$$|\mathbf{X}'\mathbf{G}^{-1}\mathbf{X}| \propto \begin{cases} \frac{(1 - \lambda_1 \rho)^h}{\{\eta + (1 - \lambda_1 \rho)^h\}(\eta + c)^{p-1}}, & \mathbf{u}_1 \in \mathcal{C}(\mathbf{X}), \\ \frac{1}{(\eta + c)^p}, & \mathbf{u}_1 \notin \mathcal{C}(\mathbf{X}), \end{cases} \quad (31)$$

$$S^2 \propto \frac{1}{\eta + c}, \quad (32)$$

where c is a positive constant and $\mathcal{C}(\mathbf{X})$ is the column space of $\mathbf{X}$ consisting of all linear combinations of the column vectors of $\mathbf{X}$. The same results hold as $\rho \rightarrow \lambda_n^{-1}$ where λ_1 and $\mathbf{u}_1$ are replaced by λ_n and $\mathbf{u}_n$, respectively.

The proof can be found in Appendix B. Based on the above proposition, we have the following conclusions.

Corollary 1. $L_{**a}(\rho, \eta; \mathbf{y})$ is a continuous function for $(\rho, \eta) \in (\lambda_1^{-1}, \lambda_n^{-1}) \times (0, \infty)$ and has the following limiting behavior as $\rho \rightarrow \lambda_1^{-1}$.

$$L_{**a}(\rho, \eta; \mathbf{y}) \propto \begin{cases} (\eta + c)^{a-1}, & \mathbf{u}_1 \in \mathcal{C}(\mathbf{X}), \\ \frac{(1 - \lambda_1 \rho)^{h/2} (\eta + c)^{a-1/2}}{\{\eta + (1 - \lambda_1 \rho)^h\}^{1/2}}, & \mathbf{u}_1 \notin \mathcal{C}(\mathbf{X}), \end{cases}$$

where c is a positive constant. The same results hold as $\rho \rightarrow \lambda_n^{-1}$ when λ_1 and $\mathbf{u}_1$ are replaced by λ_n and $\mathbf{u}_n$, respectively.

3.3. Properties of marginal priors and posteriors

The following proposition gives the limiting behaviors for $|\boldsymbol{\Sigma}_2(\rho, \eta)|$ and $|\boldsymbol{\Sigma}_3(\rho, \eta, 1)|$.

Proposition 7. Both $|\boldsymbol{\Sigma}_2(\rho, \eta)|$ and $|\boldsymbol{\Sigma}_3(\rho, \eta, 1)|$ are continuous functions for $(\rho, \eta) \in (\lambda_1^{-1}, \lambda_n^{-1}) \times (0, \infty)$ and have the following limiting behavior as $\rho \rightarrow \lambda_1^{-1}$.

$$\sqrt{|\boldsymbol{\Sigma}_2(\rho, \eta)|} \propto \frac{\eta}{(1 - \lambda_1 \rho)\{\eta + (1 - \lambda_1 \rho)^h\}(\eta + c)}, \quad (33)$$

$$\sqrt{|\boldsymbol{\Sigma}_3(\rho, \eta, 1)|} \propto \frac{\eta}{(1 - \lambda_1 \rho)\{\eta + (1 - \lambda_1 \rho)^h\}(\eta + c)^2}, \quad (34)$$

where c is a positive constant. The same results hold as $\rho \rightarrow \lambda_n^{-1}$ when λ_1 is replaced by λ_n .

Proof. Based on Proposition 2, one can easily obtain (33). Without loss of generality, we assume $\lambda_2 \neq \lambda_3$. With some algebra, one can obtain

$$\begin{aligned} |\Sigma_3(\rho, \eta, 1)| &\propto \eta^2 (a_1 b_1^2 b_2 b_3)^2 \{a_2(\eta + c_{12})d_{13} - a_3(\eta + c_{13})d_{12}\}^2 \\ &= \eta^2 (a_1 a_2 a_3 b_1^2 b_2 b_3)^2 \{(\eta + c_{12})d_{13}/a_3 - (\eta + c_{13})d_{12}/a_2\}^2. \end{aligned}$$

From the definitions of a_i , b_i , c_{ij} and d_{ij} , one can obtain the following results:

$$\begin{aligned} \frac{d_{13}}{a_3} - \frac{d_{12}}{a_2} &= \frac{-(\lambda_3 - \lambda_2)/a_1^h + (\lambda_3 - \lambda_1)/a_2^h - (\lambda_2 - \lambda_1)/a_3^h}{(\lambda_2 - \lambda_1)(\lambda_3 - \lambda_1)}, \\ \frac{c_{12}d_{13}}{a_3} - \frac{c_{13}d_{12}}{a_2} &= \left(\frac{d_{13}}{a_3} - \frac{d_{12}}{a_2} \right) / a_1^h. \end{aligned}$$

By these two equations, one can obtain the result in (34). $\square$

Remark 2. By Propositions 6 and 7, we have the limiting behavior for the marginal Jeffreys-rule prior as $\rho \rightarrow \lambda_1^{-1}$:

$$\pi^J(\rho, \eta) \propto \begin{cases} \frac{\eta(1 - \lambda_1 \rho)^{(h-2)/2}}{\{\eta + (1 - \lambda_1 \rho)^h\}^{3/2}(\eta + c)^{(p+3)/2}}, & \mathbf{u}_1 \in \mathcal{C}(\mathbf{X}), \\ \frac{\eta}{(1 - \lambda_1 \rho)\{\eta + (1 - \lambda_1 \rho)^h\}(\eta + c)^{(p+4)/2}}, & \mathbf{u}_1 \notin \mathcal{C}(\mathbf{X}). \end{cases} \quad (35)$$

The same results hold as $\rho \rightarrow \lambda_n^{-1}$ when λ_1 is replaced by λ_n .

Theorem 1. Consider the model with sampling distribution (11).

- (a) The marginal Jeffreys-rule prior $\pi^J(\rho, \eta)$ is proper when both $\mathbf{u}_1$ and $\mathbf{u}_n$ are in $\mathcal{C}(\mathbf{X})$, while it is improper when either $\mathbf{u}_1$ or $\mathbf{u}_n$ is not in $\mathcal{C}(\mathbf{X})$.
- (b) Marginal independence Jeffreys priors $\pi^{J1}(\rho, \eta)$, $\pi^{J2}(\rho, \eta)$ and $\pi^{J3}(\rho, \eta)$, and the marginal reference prior $\pi^{R1}(\rho, \eta)$ are improper.
- (c) The posterior for the Jeffreys-rule prior is proper.
- (d) The posterior for π^{J1} is proper when neither $\mathbf{u}_1$ nor $\mathbf{u}_n$ are in $\mathcal{C}(\mathbf{X})$, while it is improper when either $\mathbf{u}_1$ or $\mathbf{u}_n$ is in $\mathcal{C}(\mathbf{X})$.
- (e) The posteriors for π^{J2} and π^{J3} are always improper.
- (f) When $p = 1$ and both $\mathbf{u}_1$ and $\mathbf{u}_n$ are not in $\mathcal{C}(\mathbf{X})$, the posterior for π^{R1} is proper, while it is improper for other cases.

The proof is given in Appendix B.

Proposition 8. Suppose that λ_1 and λ_n are simple eigenvalues. Then as $\rho \rightarrow \lambda_1^{-1}$, we have

$$\pi_*^R(\rho, \eta) \propto \begin{cases} \frac{1}{(\eta + c)^2} & \text{if } \mathbf{u}_1 \in \mathcal{C}(\mathbf{X}), \\ \frac{\eta}{(1 - \lambda_1 \rho)\{\eta + (1 - \lambda_1 \rho)^h\}(\eta + c)^2} & \text{if } \mathbf{u}_1 \notin \mathcal{C}(\mathbf{X}). \end{cases} \quad (36)$$

The same results hold as $\rho \rightarrow \lambda_n^{-1}$ when λ_1 and $\mathbf{u}_1$ are replaced by λ_n and $\mathbf{u}_n$, respectively.

The proof is given in Appendix B. The following theorem's proof is similar to Theorem 1, so it is omitted.

Theorem 2. Consider the model with sampling distribution (11).

- (a) The marginal exact reference prior $\pi_*^R(\rho, \eta)$ is proper when both $\mathbf{u}_1$ and $\mathbf{u}_n$ are in $\mathcal{C}(\mathbf{X})$, while it is improper when either $\mathbf{u}_1$ or $\mathbf{u}_n$ is not in $\mathcal{C}(\mathbf{X})$.
- (b) The posterior for exact reference prior is proper.

4. Numerical comparisons

4.1. Simulation study

We perform a small simulation experiment to investigate the frequentist coverage of equal-tailed $100(1 - \alpha)\%$ Bayesian credible intervals for two parameters of interest: the spatial parameter ρ and the signal-to-noise ratio η , when one of the two objective priors is used. A 'better' prior is one in which the frequentist coverage is closer to the nominal level.

The setup is similar to [9]. The models are defined on a 10×10 regular lattice with a first order neighborhood system and $\mathbf{C}$ adjacency matrix. Thus ρ must belong to the interval $(-0.260554, 0.260554)$. We consider two different mean functions $\mathbb{E}\{\mathbf{y}(\mathbf{s})\}$, namely the constant ($p = 1$) or $\beta_0 + \beta_1 s_{i1} + \beta_2 s_{i2} + \beta_3 s_{i1} s_{i2} + \beta_4 s_{i1}^2 + \beta_5 s_{i2}^2$ ($p = 6$), and three different values of ρ :

Table 1Frequentist coverage of Bayesian equal-tailed 95% credible intervals for ρ and η by using the CAR model.

		$p = 1$				$p = 6$			
		$\eta = 0.2$	$\eta = 1$	$\eta = 2$	$\eta = 5$	$\eta = 0.2$	$\eta = 1$	$\eta = 2$	$\eta = 5$
		$\rho = 0.05$							
ρ	Reference	0.939	0.938	0.931	0.919	0.973	0.931	0.934	0.930
	Jeffreys-rule	0.953	0.956	0.947	0.936	0.919	0.946	0.952	0.960
η	Reference	0.999	1.000	0.989	0.768	0.998	1.000	0.997	0.913
	Jeffreys-rule	1.000	0.999	0.991	0.744	0.998	1.000	0.998	0.899
		$\rho = 0.12$							
ρ	Reference	0.953	0.909	0.869	0.816	0.977	0.933	0.912	0.881
	Jeffreys-rule	0.963	0.934	0.897	0.848	0.938	0.971	0.978	0.972
η	Reference	0.999	1.000	0.993	0.864	0.997	0.999	0.999	0.955
	Jeffreys-rule	1.000	1.000	0.991	0.852	0.999	1.000	0.998	0.909
		$\rho = 0.25$							
ρ	Reference	0.995	0.991	0.987	0.980	0.987	0.997	0.998	0.998
	Jeffreys-rule	0.994	0.996	0.993	0.988	0.958	0.998	0.999	0.999
η	Reference	0.976	0.996	0.997	0.992	0.988	0.994	0.999	0.998
	Jeffreys-rule	0.984	0.998	0.998	0.990	0.997	1.000	0.999	0.991

Table 2Frequentist coverage of Bayesian equal-tailed 95% credible intervals for ρ and η by using the SAR model.

		$p = 1$				$p = 6$			
		$\eta = 0.2$	$\eta = 1$	$\eta = 2$	$\eta = 5$	$\eta = 0.2$	$\eta = 1$	$\eta = 2$	$\eta = 5$
		$\rho = 0.05$							
ρ	Reference	0.949	0.907	0.869	0.825	0.949	0.907	0.879	0.842
	Jeffreys-rule	0.960	0.937	0.908	0.871	0.940	0.965	0.973	0.969
η	Reference	0.998	0.948	0.819	0.503	1.000	0.964	0.830	0.489
	Jeffreys-rule	0.999	0.964	0.855	0.519	1.000	0.992	0.941	0.593
		$\rho = 0.12$							
ρ	Reference	0.982	0.929	0.873	0.804	0.978	0.913	0.866	0.778
	Jeffreys-rule	0.984	0.960	0.921	0.860	0.962	0.993	0.990	0.979
η	Reference	0.998	0.957	0.924	0.858	0.998	0.966	0.924	0.834
	Jeffreys-rule	0.999	0.970	0.936	0.876	1.000	0.993	0.965	0.849
		$\rho = 0.25$							
ρ	Reference	0.976	0.973	0.973	0.968	0.992	0.985	0.983	0.976
	Jeffreys-rule	0.975	0.971	0.973	0.967	0.987	0.966	0.962	0.962
η	Reference	0.962	0.969	0.976	0.978	0.979	0.980	0.987	0.987
	Jeffreys-rule	0.961	0.964	0.974	0.977	0.988	0.979	0.981	0.980

0.05, 0.12, or 0.25 and four different values of η : 0.2, 1, 2 and 5. 3000 replications are generated for each combined choice of ρ and η and the equal-tailed 95% credible intervals are computed for ρ and η with two different models: CAR and SAR models.

From the simulation results in Table 1, as a CAR model is used, performance for the reference and Jeffreys-rule priors is reasonable for ρ and it is reasonable for η when η is not too large whether $p = 1$ or $p = 6$. When η is large, the performance for both priors for η is sometimes bad. Overall, based on simulated results, the performance of the two priors is similar.

Based on the simulation results in Table 2, when a SAR model is applied, the performance for the reference and Jeffreys-rule priors is reasonable for ρ and it is reasonable for η when η is not too large whether $p = 1$ or $p = 6$. When η is large, the performance for both priors for η is sometimes bad and the performance for the reference prior is worse than the Jeffreys-rule prior.

Another way of making comparisons is to present the coverage probabilities. Here we only make the graphs of the coverage probabilities for both ρ and η when the CAR model is applied. Similarly we can make these graphs under the SAR model. In Fig. 1, both the performance between Jeffreys-rule and reference priors is very similar and the Jeffreys-rule prior is a little better than reference prior when ρ is close to the upper boundary of the interval λ_n^{-1} . In Fig. 2, the performance for both two priors is almost identical, and they do better jobs when η is close to zero or ρ is close to λ_n^{-1} .

In Fig. 3, the reference prior performs better than the Jeffreys-rule prior when ρ is close to zero and λ_n^{-1} . Especially when ρ is close to zero, the performance of the reference prior is almost perfect, but the performance of Jeffreys-rule prior is better in other cases. In Fig. 4, overall, the reference prior performs better than the Jeffreys-rule prior. On the other hand, when $\rho = 0.05$ and $\eta = 0.2$, its performance is much worse for both ρ and η .

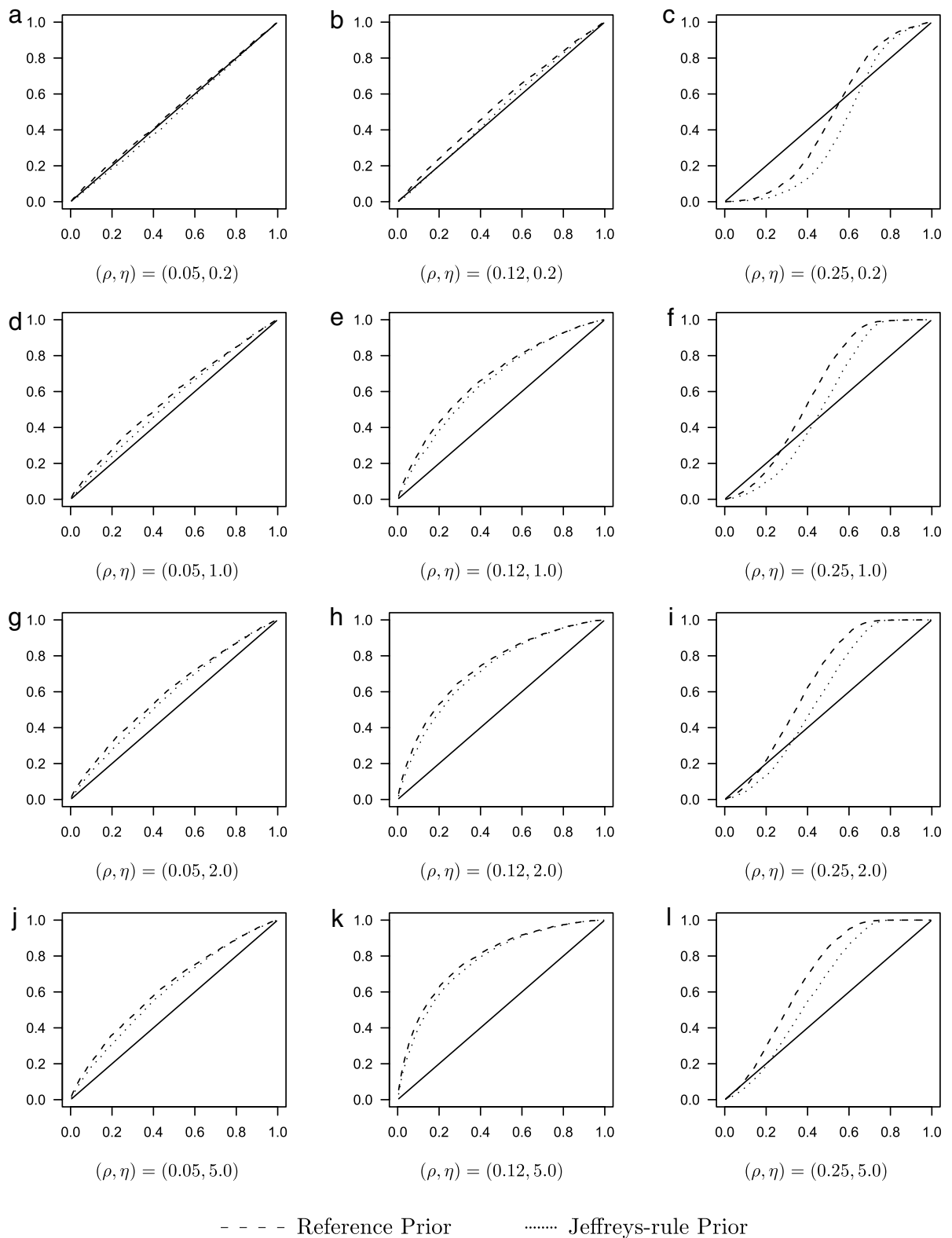


Fig. 1. The coverage probabilities of ρ under the reference and Jeffreys-rule priors when $p = 1$.

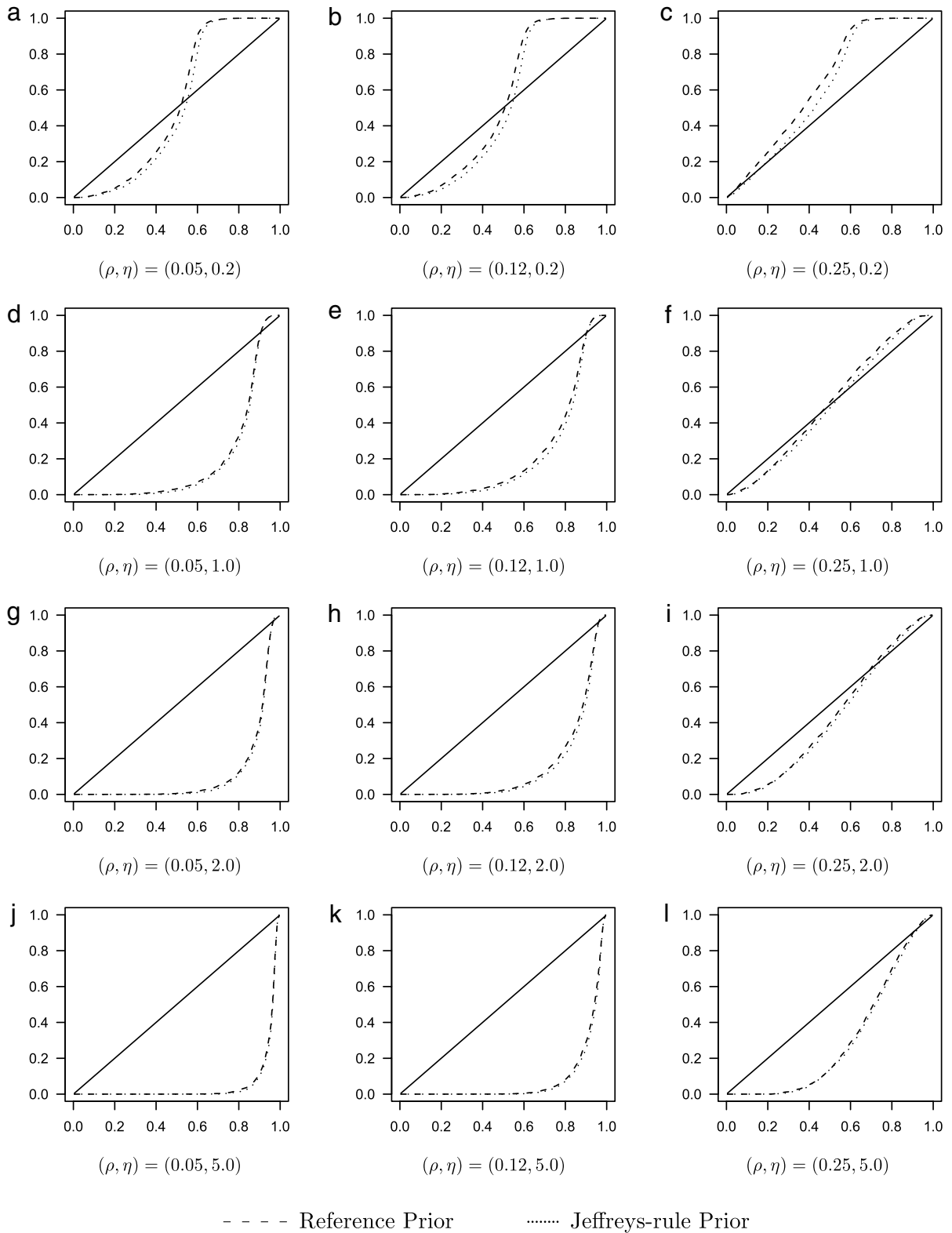


Fig. 2. The coverage probabilities of η under the reference and Jeffreys-rule priors when $p = 1$.

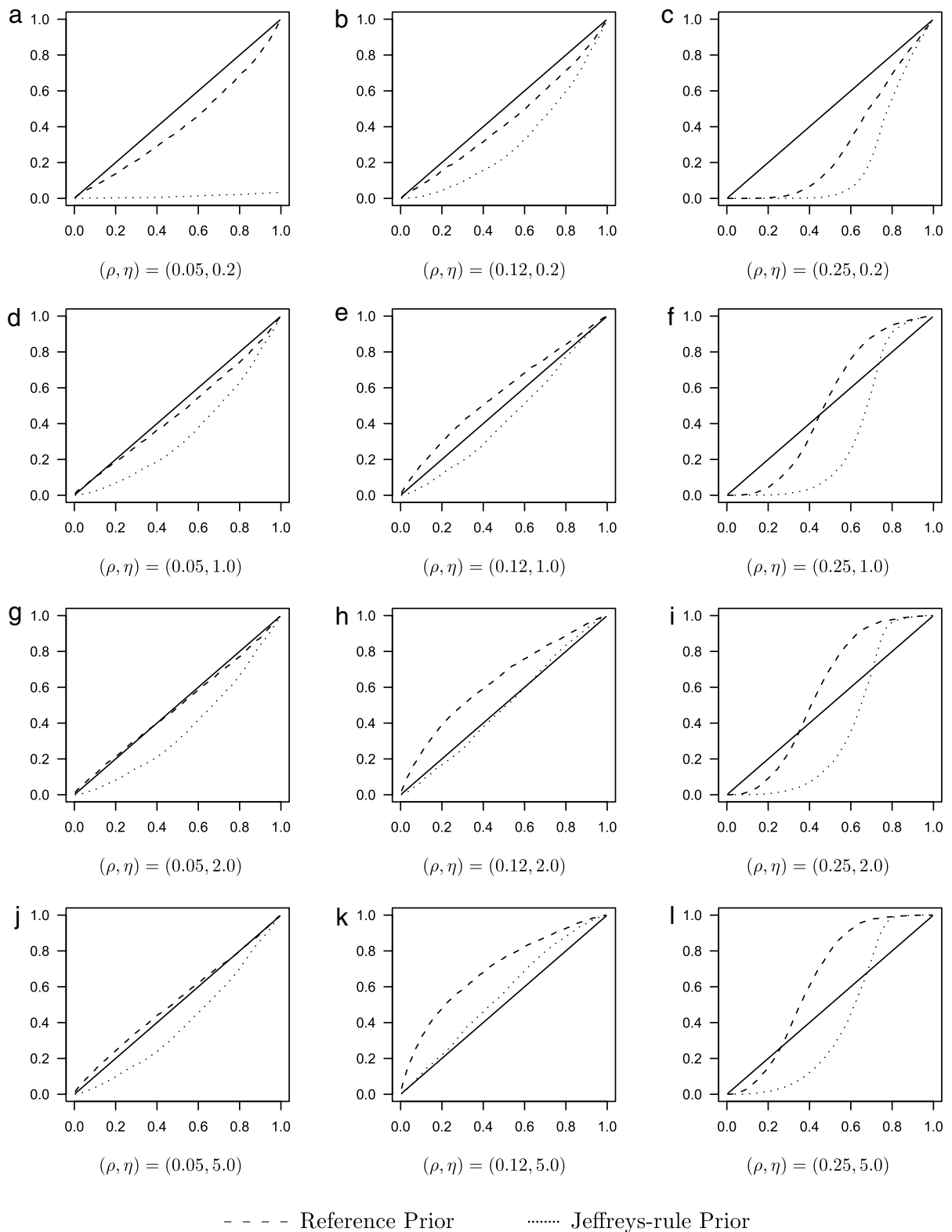


Fig. 3. The coverage probabilities of ρ under the reference and Jeffreys-rule priors when $p = 6$.

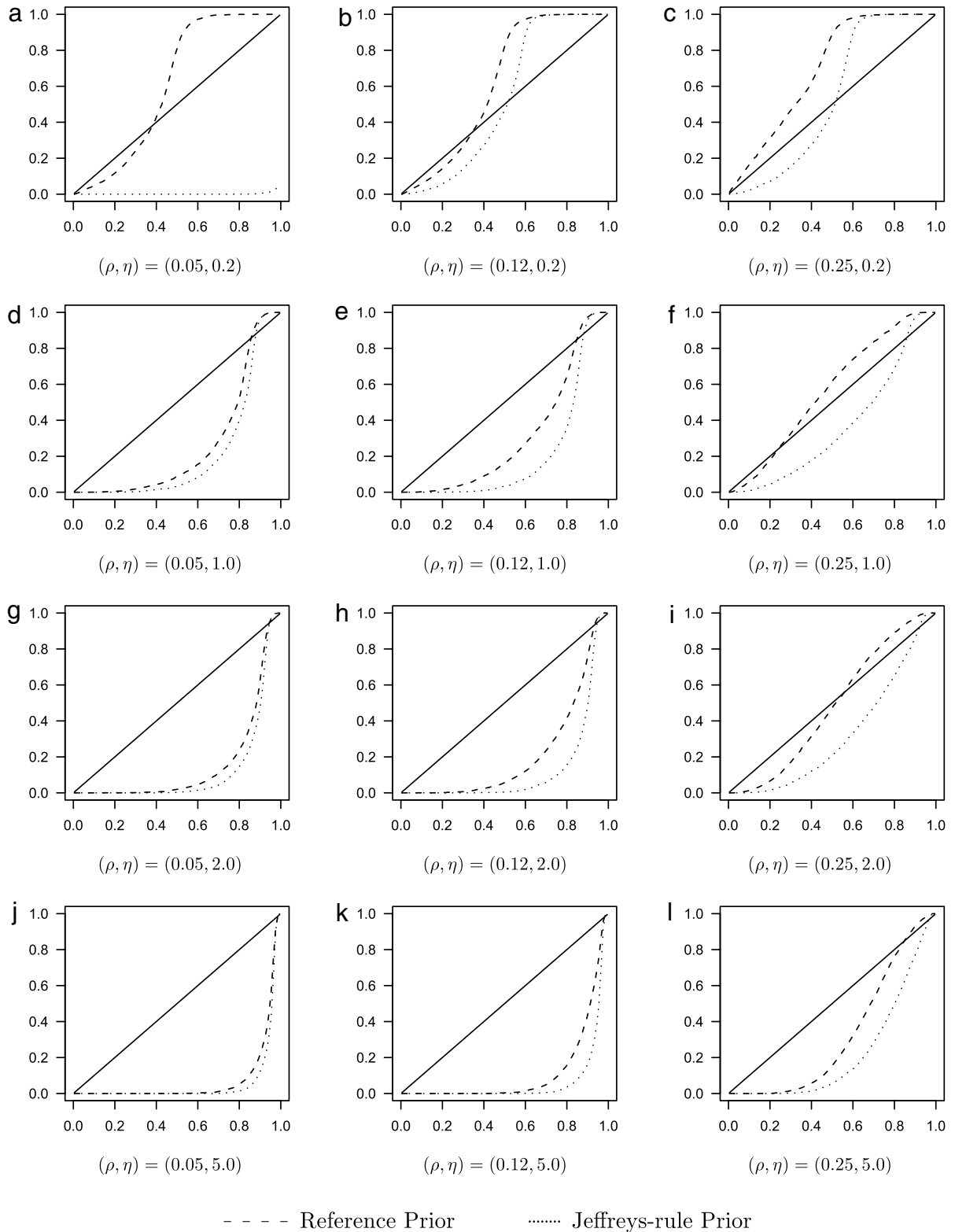


Fig. 4. The coverage probabilities of η under the reference and Jeffreys-rule priors when $p = 6$.

Table 3

Summaries of the marginal posterior distributions (2.5% quantile, median, 97.5% quantile) for small-scale-variation parameters (ρ , η , δ_0): 1974–1978 Data by the reference prior.

	$k = 0$	$k = 1$	$k = 2$
ρ	0.142	0.885	−0.990
CI	[−0.322, 0.187]	[−0.989, 0.899]	[−0.999, 0.997]
η	0.220	0.064	0.009
δ_0	954.04	1141.16	1228.87
β_1	1.602	1.601	1.594
β_2	0.036	0.036	0.036

This study is too limited to build that the reference prior overall generates references with a satisfactory frequentist coverage, but our simulation study does provide some evidence that the Jeffreys-rule prior can be very inadequate based on the frequentist coverage.

4.2. Real data analysis

For illustrative purposes, we consider one problem for the 1974–1978 data analyzed by [5]. The dataset consists of the number of SIDS and live births from July 1, 1974 to June 30, 1978 for each of 100 counties of North Carolina and the county-seat locations. SIDS is defined as the sudden death of any infant 12 months old or younger that is unexpected by history and remains inexplicable after an adequate postmortem examination. Because one county (Anson County)'s standardized residual is unacceptably high, it is dropped in our analysis and therefore there are 99 observations (counties), i.e., $n = 99$. The details of the dataset can be found in [5].

Neither [5] nor [16] considered nugget effects in their CAR model for this example. As pointed out, the nugget effects cannot be ignored for spatial data. Therefore, we consider the following CAR model with nugget effects for the transformed data. In this sense, the proposed model generalizes the traditional CAR model by including nugget effects. The model is given by

$$\tilde{\mathbf{y}} = \tilde{\mathbf{X}}\boldsymbol{\beta} + \tilde{\mathbf{e}},$$

where $\tilde{\mathbf{y}} = \mathbf{D}^{-1/2}\mathbf{y}$, $\tilde{\mathbf{X}} = \mathbf{D}^{-1/2}\mathbf{X}$ and $\tilde{\mathbf{e}}$ is assumed to be $N(\mathbf{0}, \delta_0\mathbf{I}_n + \delta_1(\mathbf{I}_n - \rho\tilde{\mathbf{C}})^{-1})$ and $\tilde{\mathbf{C}} = (\tilde{c}_{ij})$ with $\tilde{c}_{ij} = C(k)d_{ij}^k$ if $j \in N_i$ and $\tilde{c}_{ij} = 0$ otherwise. Here d_{ij} is the Euclidean distance between the i th and j th county seats, $k = 0, 1$, or 2 , $C(k) = (\min\{d_{ij} : j \in N_i, i = 1, \dots, n\})^k$ and N_i is the set of neighborhoods for the i th county: $j \in N_i$ if $d_{ij} \leq 30$ miles. $\mathbf{D} = \text{diag}(m_1^{-1}, \dots, m_n^{-1})$ where m_i is the number of live births. $\mathbf{y} = (y_1, \dots, y_n)'$ and

$$y_i = \left(\frac{1000S_i}{m_i}\right)^{1/2} + \left(\frac{1000(S_i + 1)}{m_i}\right)^{1/2},$$

where S_i is the number of SIDS.

$$\mathbf{X} = \begin{pmatrix} 1 & X_{5,1} \\ 1 & X_{5,2} \\ \dots & \dots \\ 1 & X_{5,n} \end{pmatrix}, \quad \boldsymbol{\beta} = \begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix},$$

where $X_{5,i}$ is the (Freeman–Tukey transformed) nonwhite live-birth rate, defined by

$$X_{5,i} = \left(\frac{1000\bar{\omega}_i}{m_i}\right)^{1/2} + \left(\frac{1000(\bar{\omega}_i + 1)}{m_i}\right)^{1/2},$$

and $\bar{\omega}_i$ is the number of nonwhite live births.

In this example, a ratio-of-uniforms method, which can be found in [20], for sampling ρ and η from its marginal posterior distribution, is used for simulation. This method is quite efficient in terms of computation. Because the expected value of parameters could not exist when the reference prior is applied, we report these estimates by quantiles in Table 3.

We conclude that ρ is not significantly different from 0 based on the confidence intervals for three cases. The nugget effects do exist for three cases, so the measurement error cannot be ignored. The estimate of regression coefficients (β_1 and β_2) is very stable, which is similar to the results obtained by Cressie and Chan [5] and Ren and Sun [16].

5. Summary and comments

In this paper, we propose an objective Bayesian analysis method for autoregressive models with nugget effects, which include both CAR and SAR. The reference priors and Jeffreys priors, including the Jeffreys-rule and various independence

Jeffreys priors, are developed for these models. The properties of these priors and resulting posteriors are studied. Based on the simulation study, the exact reference prior is recommended as the default prior.

Berger et al. [11] found that both the independence Jeffreys and reference priors failed to yield proper posteriors, while [14] considered the model in [11] with multi-dimensional separable correlation functions involving several parameters and found that these commonly used objective priors, which include independence Jeffreys and the usual reference priors, all yield proper posteriors under very general assumption. Recently, Ren et al. [17] studied the model for spatially correlated data with measurement errors and found that these commonly used noninformative priors fail to yield proper posteriors. They introduced a noise-to-signal ratio as a new parameter in verifying the posterior propriety. On the other hand, we also need to introduce a signal-to-noise ratio as a new parameter, as detailed in our current paper. All of these results show the complications of using objective priors in spatially correlated data.

Appendix A. Lemmas

Lemma 1. Assume that a_i and b_i , $i = 1, 2, \dots, n (\geq 2)$ are real numbers.

$$n \sum_{i=1}^n a_i b_i - \left(\sum_{i=1}^n a_i \right) \left(\sum_{i=1}^n b_i \right) = \sum_{i < j} (a_i - a_j)(b_i - b_j), \quad (37)$$

$$n \sum_{i=1}^n a_i^2 - \left(\sum_{i=1}^n a_i \right)^2 = \sum_{i < j} (a_i - a_j)^2, \quad (38)$$

$$\sum_{i=1}^n a_i^2 \sum_{i=1}^n b_i^2 - \left(\sum_{i=1}^n a_i b_i \right)^2 = \sum_{i < j} (a_i b_j - a_j b_i)^2. \quad (39)$$

Proof. For $n = 2$, one can easily verify (37) holds. Suppose that (37) holds for $n = k - 1$. For $n = k$, the left side of (37) is equal to

$$\begin{aligned} & (k-1) \sum_{i=1}^{k-1} a_i b_i - \left(\sum_{i=1}^{k-1} a_i \right) \left(\sum_{i=1}^{k-1} b_i \right) + (k-1) a_k b_k + \sum_{i=1}^{k-1} a_i b_i - a_k \sum_{i=1}^{k-1} b_i - b_k \sum_{i=1}^{k-1} a_i \\ &= \sum_{i < j \leq k-1} (a_i - a_j)(b_i - b_j) + \sum_{i=1}^{k-1} (a_i - a_k)(b_i - b_k). \end{aligned}$$

(38) is a special case of (37) when $a_i = b_i$ for all i . (39) can be similarly verified. $\square$

Lemma 2. Suppose that $\mathbf{A}$ is $n \times n$ positive definite and $\mathbf{X}$ is an $n \times p$ matrix with full column rank. Then, there exists an $n \times (n-p)$ matrix $\mathbf{T}$ satisfying

- (i) $\mathbf{T}'\mathbf{X} = \mathbf{O}$ and $\mathbf{T}'\mathbf{T} = \mathbf{I}_{n-p}$.
- (ii) $\mathbf{B} \equiv \mathbf{T}'\mathbf{A}\mathbf{T}$ is $n-p$ diagonal with positive diagonal elements.
- (iii) $\mathbf{T}\mathbf{B}^{-1}\mathbf{T}' = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{X}(\mathbf{X}'\mathbf{A}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{A}^{-1}$.

Lemma 3. For any positive constants a, b, c_1, c_2, c and s_1 , the following integral

$$\int_0^a \int_0^b \left(\frac{x + c_1 y}{x + c_2 y} \right)^{s_1} \frac{1}{(x + cy)^s} dx dy,$$

is finite if and only if $s < 2$.

The following lemma gives another expression of $|\Sigma_3^*(\rho, \eta, 1)|$, which is used in the proof of Proposition 8.

Lemma 4. The determinant of $\Sigma_3^*(\rho, \eta, 1)$ is given by

$$|\Sigma_3^*(\rho, \eta, 1)| = h^2 |\mathbf{H}|, \quad (40)$$

where

$$\mathbf{H} = \begin{vmatrix} \text{tr}\{\mathbf{R}_G(\mathbf{I}_n - \rho\mathbf{C})^{-h-1}\mathbf{C}\}^2 & \text{tr}\{\mathbf{R}_G^2(\mathbf{I}_n - \rho\mathbf{C})^{-h-1}\mathbf{C}\} & \text{tr}\{\mathbf{R}_G(\mathbf{I}_n - \rho\mathbf{C})^{-h-1}\mathbf{C}\} \\ \text{tr}\{\mathbf{R}_G^2(\mathbf{I}_n - \rho\mathbf{C})^{-h-1}\mathbf{C}\} & \text{tr}(\mathbf{R}_G^2) & \text{tr}(\mathbf{R}_G) \\ \text{tr}\{\mathbf{R}_G(\mathbf{I}_n - \rho\mathbf{C})^{-h-1}\mathbf{C}\} & \text{tr}(\mathbf{R}_G) & n - p \end{vmatrix}, \quad (41)$$

and $\mathbf{H}$ is nonnegative definite.

Proof. Noting that $\mathbf{G} = \mathbf{I}_n + \eta(\mathbf{I}_n - \rho\mathbf{C})^{-h}$, we have $\frac{\partial \mathbf{G}}{\partial \rho} = h\eta(\mathbf{I}_n - \rho\mathbf{C})^{-h-1}\mathbf{C}$ and $(\mathbf{I}_n - \rho\mathbf{C})^{-h} = \frac{1}{\eta}(\mathbf{G} - \mathbf{I}_n)$. With some algebra, we have

$$\begin{aligned} \text{tr} \left(\mathbf{R}_G \frac{\partial \mathbf{G}}{\partial \rho} \right)^2 &= h^2 \eta^2 \text{tr} \{ \mathbf{R}_G (\mathbf{I}_n - \rho\mathbf{C})^{-h-1} \mathbf{C} \}^2, \\ \text{tr} \{ \mathbf{R}_G (\mathbf{I}_n - \rho\mathbf{C})^{-h} \}^2 &= \frac{1}{\eta^2} \text{tr} \{ \mathbf{R}_G (\mathbf{G} - \mathbf{I}_n) \}^2, \\ \text{tr} \left\{ \mathbf{R}_G (\mathbf{I}_n - \rho\mathbf{C})^{-h} \mathbf{R}_G \frac{\partial \mathbf{G}}{\partial \rho} \right\} &= h \text{tr} \{ \mathbf{R}_G (\mathbf{G} - \mathbf{I}_n) \mathbf{R}_G (\mathbf{I}_n - \rho\mathbf{C})^{-h-1} \mathbf{C} \}, \\ \text{tr} \left(\mathbf{R}_G \frac{\partial \mathbf{G}}{\partial \rho} \right) &= h \eta \text{tr} \{ \mathbf{R}_G (\mathbf{I}_n - \rho\mathbf{C})^{-h-1} \mathbf{C} \}, \\ \text{tr} \{ \mathbf{R}_G (\mathbf{I}_n - \rho\mathbf{C})^{-h} \} &= \frac{1}{\eta} \text{tr} \{ \mathbf{R}_G (\mathbf{G} - \mathbf{I}_n) \}. \end{aligned}$$

Therefore, $|\Sigma_3^*(\rho, \eta, 1)|$ is equal to

$$h^2 \begin{vmatrix} \text{tr} \{ \mathbf{R}_G (\mathbf{I}_n - \rho\mathbf{C})^{-h-1} \mathbf{C} \}^2 & \text{tr} \{ \mathbf{R}_G (\mathbf{G} - \mathbf{I}_n) \mathbf{R}_G (\mathbf{I}_n - \rho\mathbf{C})^{-h-1} \mathbf{C} \} & \text{tr} \{ \mathbf{R}_G (\mathbf{I}_n - \rho\mathbf{C})^{-h-1} \mathbf{C} \} \\ \star & \text{tr} \{ \mathbf{R}_G (\mathbf{G} - \mathbf{I}_n) \}^2 & \text{tr} \{ \mathbf{R}_G (\mathbf{G} - \mathbf{I}_n) \} \\ \star & \star & n - p \end{vmatrix}.$$

Noting that $\text{tr}(\mathbf{R}_G \mathbf{G}) = n - p$ and $\mathbf{R}_G \mathbf{G} \mathbf{R}_G = \mathbf{R}_G$, one can obtain the expression for $|\Sigma_3^*(\rho, \eta, 1)|$ in (40).

Since $\mathbf{R}_G \geq 0$, both $\mathbf{G} - \mathbf{I}_n$ and $\mathbf{A}$ are positive definite and $\mathbf{C}$ is symmetric, and $\mathbf{G} - \mathbf{I}_n$, $\mathbf{A}$ and $\mathbf{C}$ commute, one can obtain $\text{tr} \{ \mathbf{R}_G (\mathbf{G} - \mathbf{I}_n) \mathbf{A} \mathbf{C} \}^2 \geq 0$. By the Schwarz inequality (see, for example, Theorem 6.3.1 at p. 62 of [7]), we have

$$\begin{aligned} \text{tr} \{ \mathbf{R}_G^2 (\mathbf{G} - \mathbf{I}_n) \mathbf{A} \mathbf{C} \} &= \text{tr} \left[\mathbf{R}_G \{ \mathbf{R}_G^{1/2} (\mathbf{G} - \mathbf{I}_n) \mathbf{A} \mathbf{C} \mathbf{R}_G^{1/2} \} \right] \\ &\leq \text{tr}(\mathbf{R}_G^2) \text{tr} \left[\{ \mathbf{R}_G^{1/2} (\mathbf{G} - \mathbf{I}_n) \mathbf{A} \mathbf{C} \mathbf{R}_G^{1/2} \}' \mathbf{R}_G^{1/2} (\mathbf{G} - \mathbf{I}_n) \mathbf{A} \mathbf{C} \mathbf{R}_G^{1/2} \right]. \end{aligned}$$

The last trace in the above inequality is the same as the first element in $\mathbf{H}$. By the relationship between determinants of $\mathbf{H}$ and $\Sigma_3^*(\rho, \eta, 1)$, we have $|\mathbf{H}| \geq 0$. By the equivalent condition for a symmetric matrix to be nonnegative definite, which is that the determinant of every principle submatrix of this matrix is nonnegative (see, for example, Theorem 14.9.11 at p. 252 of [7]), one can obtain that $\mathbf{H}$ is nonnegative definite. $\square$

Appendix B. Proofs

B.1. Proof of Proposition 1

The logarithm of the likelihood function (11) is

$$l = -\frac{n}{2} \log(2\pi) - \frac{n}{2} \log \delta_0 - \frac{1}{2} \log |\mathbf{G}| - \frac{1}{2\delta_0} (\mathbf{y} - \mathbf{X}\boldsymbol{\beta})' \mathbf{G}^{-1} (\mathbf{y} - \mathbf{X}\boldsymbol{\beta}).$$

Denote $(\mathbf{u}_1 \mathbf{u}_2 \dots \mathbf{u}_n)$ by $\mathbf{U}$, where $\mathbf{u}_i$ are the eigenvectors of $\mathbf{C}$. We obtain

$$\mathbf{G} = \mathbf{U} \text{diag} \left(1 + \frac{\eta}{(1 - \lambda_1 \rho)^h}, \dots, 1 + \frac{\eta}{(1 - \lambda_n \rho)^h} \right) \mathbf{U}'.$$

Thus, we have

$$\frac{\partial \mathbf{G}}{\partial \rho} = h \mathbf{U} \text{diag} \left(\frac{\lambda_1 \eta}{(1 - \lambda_1 \rho)^{(h+1)}}, \dots, \frac{\lambda_n \eta}{(1 - \lambda_n \rho)^{(h+1)}} \right) \mathbf{U}'.$$

By using the following facts:

$$\frac{\partial}{\partial \theta} \log |\mathbf{A}_\theta| = \text{tr} \left(\mathbf{A}_\theta^{-1} \frac{\partial}{\partial \theta} \mathbf{A}_\theta \right) \quad \text{and} \quad \frac{\partial}{\partial \theta} \mathbf{A}_\theta^{-1} = -\mathbf{A}_\theta^{-1} \left(\frac{\partial}{\partial \theta} \mathbf{A}_\theta \right) \mathbf{A}_\theta^{-1},$$

where $\mathbf{A}_\theta$ is a positive definite matrix whose elements are differentiable with respect to θ , one can obtain

$$\frac{\partial l}{\partial \rho} = -\frac{1}{2} \text{tr} \left(\mathbf{G}^{-1} \frac{\partial \mathbf{G}}{\partial \rho} \right) + \frac{1}{2\delta_0} (\mathbf{y} - \mathbf{X}\boldsymbol{\beta})' \mathbf{G}^{-1} \frac{\partial \mathbf{G}}{\partial \rho} \mathbf{G}^{-1} (\mathbf{y} - \mathbf{X}\boldsymbol{\beta}).$$

With the following facts:

$$\begin{aligned}\mathbb{E}(\mathbf{X}'\mathbf{A}\mathbf{X}) &= \text{tr}(\mathbf{A}\boldsymbol{\Sigma}) + \boldsymbol{\mu}'\mathbf{A}\boldsymbol{\mu}, \\ \text{cov}(\mathbf{X}'\mathbf{A}\mathbf{X}, \mathbf{X}'\mathbf{B}\mathbf{X}) &= 2\text{tr}(\mathbf{A}\boldsymbol{\Sigma}\mathbf{B}\boldsymbol{\Sigma}) + 4\boldsymbol{\mu}'\mathbf{A}\boldsymbol{\Sigma}\mathbf{B}\boldsymbol{\mu},\end{aligned}$$

where $\mathbf{X} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ and suppose $\mathbf{A}, \mathbf{B}$ are symmetric matrices, one can obtain

$$\mathbb{E}\left(\frac{\partial l}{\partial \rho}\right)^2 = \frac{1}{2}\text{tr}\left(\mathbf{G}^{-1}\frac{\partial \mathbf{G}}{\partial \rho}\right)^2 = \frac{(h\eta)^2}{2} \sum_{i=1}^n \frac{\lambda_i^2}{(1-\lambda_i\rho)^2\{\eta + (1-\lambda_i\rho)^h\}^2}.$$

Similarly,

$$\frac{\partial \mathbf{G}}{\partial \eta} = \mathbf{U} \text{diag}\left(\frac{1}{(1-\lambda_1\rho)^h}, \dots, \frac{1}{(1-\lambda_n\rho)^h}\right) \mathbf{U}'.$$

Thus, we obtain

$$\begin{aligned}\mathbb{E}\left(\frac{\partial l}{\partial \rho} \frac{\partial l}{\partial \eta}\right) &= \frac{1}{2}\text{tr}\left(\mathbf{G}^{-1}\frac{\partial \mathbf{G}}{\partial \rho}\mathbf{G}^{-1}\frac{\partial \mathbf{G}}{\partial \eta}\right) = \frac{h\eta}{2} \sum_{i=1}^n \frac{\lambda_i}{(1-\lambda_i\rho)\{\eta + (1-\lambda_i\rho)^h\}^2}, \\ \mathbb{E}\left(\frac{\partial l}{\partial \eta}\right)^2 &= \frac{1}{2}\text{tr}\left(\mathbf{G}^{-1}\frac{\partial \mathbf{G}}{\partial \eta}\right)^2 = \frac{1}{2} \sum_{i=1}^n \frac{1}{\{\eta + (1-\lambda_i\rho)^h\}^2}.\end{aligned}$$

The rest of the elements in the Fisher information matrix can be found easily or similarly.

B.2. Proof of Proposition 2

By Lemma 1, one can obtain

$$\begin{aligned}|\boldsymbol{\Sigma}_2(\rho, \eta)| &= \left(\sum_{i=1}^n (\lambda_i h \eta a_i b_i)^2\right) \left(\sum_{i=1}^n b_i^2\right) - \left(\sum_{i=1}^n \lambda_i h \eta a_i b_i^2\right)^2 \\ &= (h\eta)^2 \sum_{i < j} (\lambda_i a_i b_i b_j - \lambda_j a_j b_j b_i)^2.\end{aligned}$$

(14) then follows. Note that

$$|\boldsymbol{\Sigma}_3(\rho, \eta, 1)| = (h\eta)^2 \begin{vmatrix} \sum_{i=1}^n (\lambda_i a_i b_i)^2 & \sum_{i=1}^n \lambda_i a_i b_i^2 & \sum_{i=1}^n \lambda_i a_i b_i \\ \sum_{i=1}^n \lambda_i a_i b_i^2 & \sum_{i=1}^n b_i^2 & \sum_{i=1}^n b_i \\ \sum_{i=1}^n \lambda_i a_i b_i & \sum_{i=1}^n b_i & n \end{vmatrix}.$$

If the last column is multiplied by $-\sum_{i=1}^n \lambda_i a_i b_i/n$ and $-\sum_{i=1}^n b_i/n$ is added to the first and second columns, respectively, then with the first two equations in Lemma 1, one can obtain that

$$\begin{aligned}|\boldsymbol{\Sigma}_3(\rho, \eta, 1)| &= \frac{(h\eta)^2}{n} \begin{vmatrix} \sum_{i < j} (\lambda_i a_i b_i - \lambda_j a_j b_j)^2 & \sum_{i < j} (\lambda_i a_i b_i - \lambda_j a_j b_j)(b_i - b_j) \\ \sum_{i < j} (\lambda_i a_i b_i - \lambda_j a_j b_j)(b_i - b_j) & \sum_{i < j} (b_i - b_j)^2 \end{vmatrix} \\ &= \frac{(h\eta)^2}{n} \sum^* \{(\lambda_i a_i b_i - \lambda_j a_j b_j)(b_k - b_l) - (\lambda_k a_k b_k - \lambda_l a_l b_l)(b_i - b_j)\}^2,\end{aligned}$$

where the last equation is obtained by the third equation in Lemma 1. Since

$$\begin{aligned}\lambda_i a_i b_i - \lambda_j a_j b_j &= (\lambda_i - \lambda_j) a_i a_j b_i b_j \left\{ \eta + \frac{\lambda_i(1-\lambda_j\rho)^{h+1} - \lambda_j(1-\lambda_i\rho)^{h+1}}{\lambda_i - \lambda_j} \right\}, \\ b_i - b_j &= (\lambda_i - \lambda_j) \rho b_i b_j \frac{(1-\lambda_j\rho)^h - (1-\lambda_i\rho)^h}{\lambda_i - \lambda_j}.\end{aligned}$$

(15) then holds.

B.3. Proof of Proposition 4

We only prove the result for the ordering $\{\boldsymbol{\beta}, (\rho, \eta), \delta_0\}$ in Part (a). The proofs of the rest are similar. Let $[\rho_k^L, \rho_k^U], [\eta_k^L, \eta_k^U], [\delta_{0k}^L, \delta_{0k}^U]$, and $[\boldsymbol{\beta}_k^L, \boldsymbol{\beta}_k^U]$ be the compact sets so that as $k \rightarrow \infty$, $[\rho_k^L, \rho_k^U] \rightarrow (\lambda_1^{-1}, \lambda_n^{-1})$, $[\eta_k^L, \eta_k^U] \rightarrow (0, \infty)[\delta_{0k}^L, \delta_{0k}^U] \rightarrow (0, \infty)$, and $[\boldsymbol{\beta}_k^L, \boldsymbol{\beta}_k^U] \rightarrow \mathbb{R}^p$. In the following steps, we use the Fisher information matrix of $\boldsymbol{\Sigma}(\rho, \eta, \delta_0, \boldsymbol{\beta})$ given in Eq. (12) and the results in Lemma 2.1 in [6] and follow [1] reference prior algorithm. First, we construct the conditional prior δ_0 given $(\rho, \eta, \boldsymbol{\beta})$ on $[\delta_{0k}^L, \delta_{0k}^U]$,

$$\pi_k(\delta_0 | \rho, \boldsymbol{\beta}) \propto \sqrt{\frac{n}{2\delta_0^2}} \propto \frac{1}{\delta_0}.$$

Next, we can construct the conditional prior for (ρ, η) given $\boldsymbol{\beta}$ on $[\rho_k^L, \rho_k^U] \times [\eta_k^L, \eta_k^U]$,

$$\pi_k(\rho, \eta | \boldsymbol{\beta}) \propto \exp \left\{ \frac{1}{2} \int_{\delta_{0k}^L}^{\delta_{0k}^U} \log \frac{\frac{1}{8} |\boldsymbol{\Sigma}_3(\rho, \eta, \delta_0)|}{\frac{n}{2\delta_0^2}} \pi_k(\delta_0 | \rho, \boldsymbol{\beta}) d\delta_0 \right\} \propto |\boldsymbol{\Sigma}_3(\rho, \eta, 1)|^{1/2},$$

where we use the fact $|\boldsymbol{\Sigma}_3(\rho, \eta, \delta_0)| = |\boldsymbol{\Sigma}_3(\rho, \eta, 1)|/\delta_0^2$.

Last, we can construct the prior of $\boldsymbol{\beta}$ within each compact set:

$$\pi_k(\boldsymbol{\beta}) \propto \exp \left\{ \frac{1}{2} \int_{\rho_k^L}^{\rho_k^U} \int_{\eta_k^L}^{\eta_k^U} \int_{\delta_{0k}^L}^{\delta_{0k}^U} \log \frac{\frac{1}{8\delta_0^2} |\boldsymbol{\Sigma}_3(\rho, \eta, \delta_0)| |\mathbf{X}'\mathbf{G}^{-1}\mathbf{X}|}{\frac{1}{8} |\boldsymbol{\Sigma}_3(\rho, \eta, \delta_0)|} \pi_k(\rho, \eta | \boldsymbol{\beta}) \pi_k(\delta_0 | \rho, \eta, \boldsymbol{\beta}) d\rho d\eta d\delta_0 \right\} \propto 1.$$

Finally, for some interior point of $(\rho_0, \eta_0, \delta_{00}, \boldsymbol{\beta}_0)$ of $(\rho, \eta, \delta_0, \boldsymbol{\beta})$, the joint reference prior is

$$\pi^R(\rho, \eta, \delta_0, \boldsymbol{\beta}) = \lim_{k \rightarrow \infty} \frac{\pi_k(\boldsymbol{\beta}) \pi_k(\rho, \eta | \boldsymbol{\beta}) \pi_k(\delta_0 | \rho, \eta, \boldsymbol{\beta})}{\pi_k(\boldsymbol{\beta}_0) \pi_k(\rho_0, \eta_0 | \boldsymbol{\beta}_0) \pi_k(\delta_{00} | \rho_0, \eta_0, \boldsymbol{\beta}_0)} \propto \frac{1}{\delta_0} |\boldsymbol{\Sigma}_3(\rho, \eta, 1)|^{1/2}.$$

The result then follows.

B.4. Proof of Proposition 6

Let $\mathbf{U} = (\mathbf{u}_1 \mathbf{u}_2 \dots \mathbf{u}_n)$, where $\mathbf{u}_i$ are the eigenvectors of $\mathbf{C}$. Denoting $\text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$ by $\mathbf{A}$, we have $\mathbf{U}'\mathbf{C}\mathbf{U} = \mathbf{A}$, so we obtain

$$\mathbf{U}'\mathbf{G}^{-1}\mathbf{U} = \{\mathbf{I}_n + \eta(\mathbf{I}_n - \rho\mathbf{A})^{-h}\}^{-1} = (\mathbf{I}_n - \rho\mathbf{A})^h \{\eta\mathbf{I}_n + (\mathbf{I}_n - \rho\mathbf{A})^{-h}\}^{-1} \triangleq \mathbf{A},$$

and the conclusion for $|\mathbf{G}^{-1}|$. Note that $\mathbf{X} = \mathbf{Q}\mathbf{R}$ (the so-called QR decomposition), where $\mathbf{Q}$ is an $n \times p$ column orthonormal matrix and $\mathbf{R}$ is a $p \times p$ upper triangular matrix. Thus,

$$|\mathbf{X}'\mathbf{G}^{-1}\mathbf{X}| = O(|\mathbf{Q}'\mathbf{G}^{-1}\mathbf{Q}|) = O(|\mathbf{Q}'\mathbf{U}\mathbf{A}\mathbf{U}'\mathbf{Q}|). \quad (42)$$

If $\mathbf{u}_1 \in \mathcal{C}(\mathbf{X})$, then $\mathbf{u}_1 \in \mathcal{C}(\mathbf{Q})$. One can find $\mathbf{t}_2, \dots, \mathbf{t}_p \in \mathcal{C}(\mathbf{Q})$ such that $\mathbf{u}_1, \mathbf{t}_2, \dots, \mathbf{t}_p$ are orthonormal. Thus, denoting $(\mathbf{u}_1 \mathbf{t}_2 \dots \mathbf{t}_p)$ by $\mathbf{Q}^*$, one can find a nonsingular $p \times p$ matrix $\mathbf{T}$ such that $\mathbf{Q} = \mathbf{Q}^*\mathbf{T}$. With some algebra, we have

$$\mathbf{Q}^*\mathbf{U} = \begin{pmatrix} 1 & \mathbf{0}' \\ \mathbf{0} & \tilde{\mathbf{Q}} \end{pmatrix}.$$

By (42), we have

$$|\mathbf{X}'\mathbf{G}^{-1}\mathbf{X}| = O(|\mathbf{Q}^*\mathbf{U}\mathbf{A}\mathbf{U}'\mathbf{Q}^*|) = O\left(\frac{(1 - \lambda_1\rho)^h}{\eta + (1 - \lambda_1\rho)^h} |\tilde{\mathbf{Q}}'\mathbf{A}_2\tilde{\mathbf{Q}}|\right),$$

where $\mathbf{A}_2 = (\mathbf{I}_{n-1} - \rho\mathbf{A}_2)^h \{\eta\mathbf{I}_{n-1} + (\mathbf{I}_{n-1} - \rho\mathbf{A}_2)^h\}^{-1}$, and $\mathbf{A}_2 = \text{diag}(\lambda_2, \dots, \lambda_n)$. The rest for $|\mathbf{X}'\mathbf{G}^{-1}\mathbf{X}|$ can be easily obtained.

If $\mathbf{u}_1 \notin \mathcal{C}(\mathbf{X})$, the elements involving $\frac{(1 - \lambda_1\rho)^h}{\eta + (1 - \lambda_1\rho)^h}$ are always the linear combination of all $\frac{(1 - \lambda_i\rho)^h}{\eta + (1 - \lambda_i\rho)^h}$, $i = 1, 2, \dots, n$ and at least one coefficient of $\frac{(1 - \lambda_i\rho)^h}{\eta + (1 - \lambda_i\rho)^h}$ is not zero.

When $\rho < 0$, we have

$$\frac{(1 - \lambda_1\rho)^h}{\eta + (1 - \lambda_1\rho)^h} \leq \frac{(1 - \lambda_2\rho)^h}{\eta + (1 - \lambda_2\rho)^h} \leq \dots \leq \frac{(1 - \lambda_n\rho)^h}{\eta + (1 - \lambda_n\rho)^h}.$$

Denoting $\mathbf{U}'\mathbf{X}$ and $\frac{(1 - \lambda_i\rho)^h}{\eta + (1 - \lambda_i\rho)^h}$ by $\tilde{\mathbf{X}}$ and $\tilde{\lambda}_i$, respectively, and considering the function

$$f(\tilde{\lambda}_1, \tilde{\lambda}_2, \dots, \tilde{\lambda}_n) = \mathbf{R}_G,$$

one can find that this function is nondecreasing for each $\tilde{\lambda}_i$. Thus, we have

$$\mathbf{U}\{\tilde{\mathbf{A}}_1 - \tilde{\mathbf{A}}_1\tilde{\mathbf{X}}(\tilde{\mathbf{X}}'\tilde{\mathbf{A}}_1\tilde{\mathbf{X}})^{-1}\tilde{\mathbf{X}}'\tilde{\mathbf{A}}_1\}\mathbf{U}' \leq \mathbf{R}_G \leq \mathbf{U}\{\tilde{\mathbf{A}}_2 - \tilde{\mathbf{A}}_2\tilde{\mathbf{X}}(\tilde{\mathbf{X}}'\tilde{\mathbf{A}}_2\tilde{\mathbf{X}})^{-1}\tilde{\mathbf{X}}'\tilde{\mathbf{A}}_2\}\mathbf{U}', \quad (43)$$

where

$$\tilde{\mathbf{A}}_1 = \text{diag}(\tilde{\lambda}_1, \tilde{\lambda}_2, \dots, \tilde{\lambda}_2), \quad \tilde{\mathbf{A}}_2 = \text{diag}(\tilde{\lambda}_1, \tilde{\lambda}_n, \dots, \tilde{\lambda}_n).$$

If $\mathbf{u}_1 \in \mathcal{C}(\mathbf{X})$, from the proof for $|\mathbf{X}'\mathbf{G}^{-1}\mathbf{X}|$, one can obtain the following

$$\mathbf{U}'\mathbf{R}_G\mathbf{U} = \begin{pmatrix} 0 & \\ & \mathbf{A}_2 - \mathbf{A}_2\tilde{\mathbf{Q}}'(\tilde{\mathbf{Q}}\mathbf{A}_2\tilde{\mathbf{Q}}')^{-1}\tilde{\mathbf{Q}}\mathbf{A}_2 \end{pmatrix}, \quad (44)$$

and (43) becomes

$$\begin{pmatrix} 0 & \\ & \tilde{\lambda}_2(\mathbf{I}_{n-1} - \tilde{\mathbf{Q}}(\tilde{\mathbf{Q}}'\tilde{\mathbf{Q}})^{-1}\tilde{\mathbf{Q}}') \end{pmatrix} \leq \mathbf{U}'\mathbf{R}_G\mathbf{U} \leq \begin{pmatrix} 0 & \\ & \tilde{\lambda}_n(\mathbf{I}_{n-1} - \tilde{\mathbf{Q}}(\tilde{\mathbf{Q}}'\tilde{\mathbf{Q}})^{-1}\tilde{\mathbf{Q}}') \end{pmatrix}.$$

The result then follows.

Let $\tilde{\mathbf{X}}' = (\mathbf{x}_1 \mathbf{X}^{*'})$. If $\mathbf{u}_1 \notin \mathcal{C}(\mathbf{X})$, then $\mathbf{X}^{*'}\mathbf{X}^*$ must be nonsingular. If not, then one can find a $p \times p$ nonsingular matrix $\mathbf{V}$ such that

$$(\mathbf{0} \mathbf{X}^{*'}) = \mathbf{X}^*\mathbf{V}.$$

$\mathbf{X}$ is a full column rank matrix, so is $\tilde{\mathbf{X}}$. Thus, we have the first element in $\tilde{\mathbf{X}}\mathbf{V}$ must be not zero and all other elements in the first column are zero, so we can find a $p \times p$ nonsingular matrix $\mathbf{V}^*$ such that

$$\tilde{\mathbf{X}}\mathbf{V}^* = \begin{pmatrix} 1 & \mathbf{0}' \\ \mathbf{0} & \mathbf{X}^{**} \end{pmatrix}.$$

From this, we will see that $\mathbf{u}_1 \in \mathcal{C}(\mathbf{X})$. This contradicts! One can obtain

$$\begin{aligned} (\tilde{\mathbf{X}}'\tilde{\mathbf{A}}_1\tilde{\mathbf{X}})^{-1} &= (\tilde{\lambda}_1\mathbf{x}_1\mathbf{x}_1' + \tilde{\lambda}_2\mathbf{X}^{*'}\mathbf{X}^*)^{-1} \\ &= \frac{1}{\tilde{\lambda}_2} \left\{ (\mathbf{X}^{*'}\mathbf{X}^*)^{-1} - \frac{\tilde{\lambda}_1(\mathbf{X}^{*'}\mathbf{X}^*)^{-1}\mathbf{x}_1\mathbf{x}_1'(\mathbf{X}^{*'}\mathbf{X}^*)^{-1}}{\tilde{\lambda}_2 + \tilde{\lambda}_1\mathbf{x}_1'(\mathbf{X}^{*'}\mathbf{X}^*)^{-1}\mathbf{x}_1} \right\}. \end{aligned} \quad (45)$$

With some algebra, one can obtain $\mathbf{U}'\mathbf{R}_G\mathbf{U}$ is greater than or equal to

$$\begin{pmatrix} \frac{\tilde{\lambda}_1\tilde{\lambda}_2}{\tilde{\lambda}_2 + \tilde{\lambda}_1\mathbf{x}_1'(\mathbf{X}^{*'}\mathbf{X}^*)^{-1}\mathbf{x}_1} & -\frac{\tilde{\lambda}_1\tilde{\lambda}_2\mathbf{x}_1'(\mathbf{X}^{*'}\mathbf{X}^*)^{-1}\mathbf{X}^{*'}}{\tilde{\lambda}_2 + \tilde{\lambda}_1\mathbf{x}_1'(\mathbf{X}^{*'}\mathbf{X}^*)^{-1}\mathbf{x}_1} \\ -\frac{\tilde{\lambda}_1\tilde{\lambda}_2\mathbf{X}^*(\mathbf{X}^{*'}\mathbf{X}^*)^{-1}\mathbf{x}_1}{\tilde{\lambda}_2 + \tilde{\lambda}_1\mathbf{x}_1'(\mathbf{X}^{*'}\mathbf{X}^*)^{-1}\mathbf{x}_1} & \tilde{\lambda}_2 \left\{ \mathbf{I}_{n-1} - \mathbf{X}^*(\mathbf{X}^{*'}\mathbf{X}^*)^{-1}\mathbf{X}^{*'} + \frac{\tilde{\lambda}_1\mathbf{X}^*(\mathbf{X}^{*'}\mathbf{X}^*)^{-1}\mathbf{x}_1\mathbf{x}_1'(\mathbf{X}^{*'}\mathbf{X}^*)^{-1}\mathbf{X}^{*'}}{\tilde{\lambda}_2 + \tilde{\lambda}_1\mathbf{x}_1'(\mathbf{X}^{*'}\mathbf{X}^*)^{-1}\mathbf{x}_1} \right\} \end{pmatrix}.$$

As $\rho \rightarrow \lambda_1^{-1}$, $\tilde{\lambda}_1/\tilde{\lambda}_2 \rightarrow 0$, so (32) holds. The proof is then complete.

B.5. Proof of Theorem 1

(a) We divide the integral interval into the following four parts:

$$(\lambda_1^{-1}, 0] \times (0, M_1], (\lambda_1^{-1}, 0] \times [M_1, \infty), [0, \lambda_n^{-1}] \times (0, M_1], \text{ and } [0, \lambda_n^{-1}] \times [M_1, \infty).$$

Assume both $\mathbf{u}_1$ and $\mathbf{u}_n$ are in $\mathcal{C}(\mathbf{X})$. Note that the third and fourth integral intervals are similar to the first two, so we only verify whether it is integrable at the first and second integral intervals. By Remark 2, we have the limiting behavior of π^l given in (35). With the transformation

$$t = (1 - \lambda_1\rho)^h, \quad (46)$$

we obtain $d\rho = t^{1/h-1}/(h\lambda_1)dt$. Thus,

$$\begin{aligned} \int \int_{(\lambda_1^{-1}, 0] \times (0, M_1]} \pi^l(\rho, \eta) d\rho d\eta &\propto \int \int_{(\lambda_1^{-1}, 0] \times (0, M_1]} \frac{\eta(1 - \lambda_1\rho)^{(h-2)/2}}{\{\eta + (1 - \lambda_1\rho)^h\}^{3/2}} d\rho d\eta \\ &\propto \int \int_{(0, 1] \times (0, M_1]} \frac{\eta}{t^{1/2}(\eta + t)^{3/2}} dt d\eta \\ &\propto \int \int_{(0, 1] \times (0, M_1]} \left\{ \frac{1}{t^{1/2}(\eta + t)^{1/2}} - \frac{t^{1/2}}{(\eta + t)^{3/2}} \right\} dt d\eta. \end{aligned}$$

For the first integration, we have

$$\int \int_{(0,1] \times (0,M_1]} \frac{1}{t^{1/2}(\eta+t)^{1/2}} dt d\eta = 2 \int_0^1 \frac{(\eta+t)^{1/2}}{t^{1/2}} \Big|_0^{M_1} dt = 2 \int_0^1 \frac{(M_1+t)^{1/2}}{t^{1/2}} dt - 2,$$

is finite since $t^{-1/2}$ is integrable in $(0, 1]$. Since $\frac{t^{1/2}}{(\eta+t)^{3/2}} \leq \frac{1}{(\eta+t)^{3/2}}$, by Lemma 3, the second integration is integrable. Therefore, the Jeffreys-rule prior is integrable for the first integral interval.

For the second integral interval, from Remark 2, we have

$$\pi^J(\rho, \eta) \propto \frac{(1 - \lambda_1 \rho)^{h/2-1}}{\eta^{p/2+2}}.$$

It is easy to see that it is integrable. The result then follows. When $\mathbf{u}_1 \notin \mathcal{C}(\mathbf{X})$, from Remark 2, for the first integral interval, we have

$$\pi^J(\rho, \eta) \propto \frac{\eta}{(1 - \lambda_1 \rho)\{\eta + (1 - \lambda_1 \rho)^h\}} = \frac{1}{1 - \lambda_1 \rho} - \frac{(1 - \lambda_1 \rho)^{h-1}}{\eta + (1 - \lambda_1 \rho)^h}.$$

With the transformation given by (46), we have

$$\int \int_{(\lambda_1^{-1}, 0] \times (0, M_1]} \frac{(1 - \lambda_1 \rho)^{h-1}}{\eta + (1 - \lambda_1 \rho)^h} d\rho d\eta = -\frac{1}{\lambda_1 h} \int \int_{(0,1] \times (0, M_1]} \frac{1}{\eta + t} dt d\eta.$$

By Lemma 3 again, the function $\frac{1}{\eta+t}$ is integrable, but $\frac{1}{1-\lambda_1 \rho}$ is not integrable. Therefore, $\pi^J(\rho, \eta)$ is improper when $\mathbf{u}_1 \notin \mathcal{C}(\mathbf{X})$. The same result holds when $\mathbf{u}_n \notin \mathcal{C}(\mathbf{X})$.

(b) From Propositions 3 and 4, we will see that each of these priors is either proportional to $|\Sigma_2(\rho, \eta)|^{1/2}$ or $|\Sigma_3(\rho, \eta, 1)|^{1/2}$. By Proposition 7, the results then follow.

(c) As $\rho \rightarrow \lambda_1^{-1}$, by Corollary 1 and Remark 2, we have

$$L_{**a}(\rho, \eta; \mathbf{y}) \pi^J(\rho, \eta) \propto \frac{\eta(1 - \lambda_1 \rho)^{h/2-1}}{\{\eta + (1 - \lambda_1 \rho)^h\}^{3/2}(\eta + c)^{3/2}}.$$

The rest proof is similar to Part (a). The result then follows. The proofs for the rest parts are similar.

B.6. Proof of Proposition 8

Case 1: $\rho \rightarrow \lambda_1^{-1}$ and $\eta \rightarrow 0^+$ or fixed. By Proposition 5, we get

$$|\Sigma_3^*(\rho, \eta, 1)| \leq (n-p) \text{tr}\{\mathbf{R}_G(\mathbf{I}_n - \rho\mathbf{C})^{-h}\}^2 \text{tr}\left(\mathbf{R}_G \frac{\partial}{\partial \rho} \mathbf{G}\right)^2. \quad (47)$$

If $\mathbf{u}_1 \in \mathcal{C}(\mathbf{X})$, with the expression of $\mathbf{R}_G$ given in (44), we have both $\text{tr}\{\mathbf{R}_G(\mathbf{I}_n - \rho\mathbf{C})^{-h}\}^2$ and $\text{tr}\left(\mathbf{R}_G \frac{\partial}{\partial \rho} \mathbf{G}\right)^2$ are not involved in the element $1 - \lambda_1 \rho$, so the right-hand side of the above inequality is continuous and is bounded by some positive value for all $\rho \in (\lambda_1^{-1}, 0]$.

Next, we assume that $\mathbf{u}_1 \notin \mathcal{C}(\mathbf{X})$. Denote that $\mathbf{A}_2 = \text{diag}(a_2^h, \dots, a_n^h)$ and $\mathbf{B}_2 = \text{diag}(b_2, \dots, b_n)$, where a_i and b_i , $i = 2, \dots, n$ are defined in Proposition 2. Defining

$$\mathbf{R}_\Delta = \mathbf{A}_2^{-1} \mathbf{B}_2 - \mathbf{A}_2^{-1} \mathbf{B}_2 \mathbf{X}^* (\mathbf{X}^{*'} \mathbf{A}_2^{-1} \mathbf{B}_2 \mathbf{X}^*)^{-1} \mathbf{X}^{*'} \mathbf{A}_2^{-1} \mathbf{B}_2,$$

$$c = \mathbf{x}_1' (\mathbf{X}^{*'} \mathbf{A}_2^{-1} \mathbf{B}_2 \mathbf{X}^*)^{-1} \mathbf{x}_1,$$

$$\alpha = \mathbf{A}_2^{-1} \mathbf{B}_2 \mathbf{X}^* (\mathbf{X}^{*'} \mathbf{A}_2^{-1} \mathbf{B}_2 \mathbf{X}^*)^{-1} \mathbf{x}_1,$$

where $\mathbf{X}^*$ is defined in the proof of Proposition 6, one can find that

$$\mathbf{U}' \mathbf{R}_G \mathbf{U} = \frac{b_1}{a_1^h + b_1 c} \begin{pmatrix} 1 & -\alpha' \\ -\alpha & \frac{a_1^h + b_1 c}{b_1} \mathbf{R}_\Delta + \alpha \alpha' \end{pmatrix} \equiv \gamma \mathbf{D},$$

where

$$\gamma = \frac{b_1}{a_1^h + b_1 c}, \quad \mathbf{D} = \begin{pmatrix} 1 & -\alpha' \\ -\alpha & \gamma^{-1} \mathbf{R}_\Delta + \alpha \alpha' \end{pmatrix}.$$

Since $a_1^h \geq b_1 > 0$ and $c > 0$, we obtain $0 < \gamma < 1$. Based on the above expression of $\mathbf{R}_G$, one can obtain that

$$|\Sigma_3^*(\rho, \eta, 1)| = \eta^2 |\sigma_{ij}|,$$

where

$$\begin{aligned}\sigma_{11} &= \lambda_1^2 a_1^{2(h+1)} \gamma^2 + 2\lambda_1 a_1^{h+1} \gamma^2 \alpha' \mathbf{A}_2^{1+1/h} \mathbf{A}_2 \alpha + \text{tr}(\mathbf{R}_\Delta \mathbf{A}_2^{1+1/h} \mathbf{A}_2)^2 \\ &\quad + 2\gamma \alpha' \mathbf{A}_2^{1+1/h} \mathbf{A}_2 \mathbf{R}_\Delta \mathbf{A}_2^{1+1/h} \mathbf{A}_2 \alpha + \gamma^2 (\alpha' \mathbf{A}_2^{1+1/h} \mathbf{A}_2 \alpha)^2, \\ \sigma_{12} &= \lambda_1 a_1^{2h+1} \gamma^2 + \lambda_1 a_1^{h+1} \gamma^2 \alpha' \mathbf{A}_2 \alpha + a_1^h \gamma^2 \alpha' \mathbf{A}_2^{1+1/h} \mathbf{A}_2 \alpha + \text{tr}(\mathbf{R}_\Delta \mathbf{A}_2 \mathbf{R}_\Delta \mathbf{A}_2^{1+1/h} \mathbf{A}_2) \\ &\quad + 2\gamma \alpha' \mathbf{A}_2 \mathbf{R}_\Delta \mathbf{A}_2^{1+1/h} \mathbf{A}_2 \alpha + \gamma^2 \alpha' \mathbf{A}_2 \alpha \alpha' \mathbf{A}_2^{1+1/h} \mathbf{A}_2 \alpha, \\ \sigma_{13} &= \lambda_1 a_1^{h+1} \gamma + \text{tr}(\mathbf{R}_\Delta \mathbf{A}_2^{1+1/h} \mathbf{A}_2) + \gamma \alpha' \mathbf{A}_2^{1+1/h} \mathbf{A}_2 \alpha, \\ \sigma_{22} &= a_1^{2h} \gamma^2 + 2a_1^h \gamma^2 \alpha' \mathbf{A}_2 \alpha + \text{tr}(\mathbf{R}_\Delta \mathbf{A}_2)^2 + 2\gamma \alpha' \mathbf{A}_2 \mathbf{R}_\Delta \mathbf{A}_2 \alpha + \gamma^2 (\alpha' \mathbf{A}_2 \alpha)^2, \\ \sigma_{23} &= a_1^h \gamma + \text{tr}(\mathbf{R}_\Delta \mathbf{A}_2) + \gamma \alpha' \mathbf{A}_2 \alpha, \\ \sigma_{33} &= n - p.\end{aligned}$$

As $\rho \rightarrow \lambda_1^{-1}$ and $\eta \rightarrow 0^+$ or fixed, it can be found that some terms in σ_{ij} are bounded. For example,

$$\sigma_{11} = \lambda_1^2 a_1^{2(h+1)} \gamma^2 + 2\lambda_1 a_1^{h+1} \gamma^2 \alpha' \mathbf{A}_2^{1+1/h} \mathbf{A}_2 \alpha + c_{11},$$

where c_{11} is bounded. Thus, with some algebras, one can obtain that

$$|\Sigma_3^*(\rho, \eta, 1)| \propto (\eta a_1^{h+1} \gamma)^2.$$

Noting that $\gamma = \frac{a_1^{-h}}{\eta + (c+1)a_1^{-h}}$, we have

$$|\Sigma_3^*(\rho, \eta, 1)| \propto \left(\frac{\eta a_1}{\eta + (c+1)a_1^{-h}} \right)^2.$$

Since $\min\left(1, \frac{c_1}{c_2}\right) \leq \frac{x+c_1y}{x+c_2y} \leq \max\left(1, \frac{c_1}{c_2}\right)$, for any positive numbers of c_1 and c_2 , we get

$$|\Sigma_3^*(\rho, \eta, 1)| \propto \frac{\eta^2}{(1 - \lambda_1 \rho)^2 \{\eta + (1 - \lambda_1 \rho)^h\}^2}.$$

Case 2: $\eta \rightarrow \infty$. By Lemma 4, we have

$$|\Sigma_3^*(\rho, \eta, 1)| \leq (n-p) \text{tr} \left\{ \mathbf{R}_G (\mathbf{I}_n - \rho \mathbf{C})^{-h-1} \mathbf{C} \right\}^2 \text{tr}(\mathbf{R}_G^2).$$

Similarly, if $\mathbf{u}_1 \in \mathcal{C}(\mathbf{X})$, the right hand side of the above inequality does not contain the element $1 - \lambda_1 \rho$, so we get a positive constant c' such that

$$\begin{aligned}\text{tr} \left\{ \mathbf{R}_G (\mathbf{I}_n - \rho \mathbf{C})^{-h-1} \mathbf{C} \right\}^2 &\leq c' \text{tr}(\mathbf{B}_2^2), \\ \text{tr}(\mathbf{R}_G^2) &\leq c' \text{tr}(\mathbf{B}_2^2).\end{aligned}$$

Note that $\text{tr}(\mathbf{B}_2^2) \propto \eta^{-2}$. Thus, $|\Sigma_3^*(\rho, \eta, 1)| \propto \frac{1}{\eta^4}$.

Assume that $\mathbf{u}_1 \notin \mathcal{C}(\mathbf{X})$. By

$$\begin{aligned}\text{tr} \left\{ \mathbf{R}_G (\mathbf{I}_n - \rho \mathbf{C})^{-h-1} \mathbf{C} \right\}^2 &\leq \frac{c'}{\eta^2 (1 - \lambda_1 \rho)^2}, \\ \text{tr}(\mathbf{R}_G^2) &\leq c' \text{tr}(\mathbf{G}^{-2}),\end{aligned}$$

and $\text{tr}(\mathbf{G}^{-2}) \propto \eta^{-2}$, one can obtain the result.

References

- [1] J.O. Berger, J.M. Bernardo, On the development of reference priors (disc: p 49–60), in: Proceedings of the Fourth Valencia International Meeting, in: Bayesian Statistics, vol. 4, 1992, pp. 35–49.
- [2] J.G.P. Besag, Spatial interaction and the statistical analysis of lattice systems (with discussion), J. R. Stat. Soc. Ser. B Methodol. 36 (1974) 3–66.
- [3] D. Clayton, J. Kaldor, Empirical Bayes estimates of age-standardized relative risks for use in disease mapping, Biometrics 43 (1987) 671–681.
- [4] N.A.C. Cressie, Statistics for Spatial Data, Rev. ed., Wiley, New York, 1993.
- [5] N.A.C. Cressie, N.H. Chan, Spatial modeling of regional variables, J. Amer. Statist. Assoc. 84 (1989) 393–401.
- [6] G.S. Datta, M. Ghosh, On the invariance of noninformative priors, Ann. Statist. 24 (1996) 141–159.
- [7] D.A. Harville, Matrix From a Statistician's Perspective, Springer-Verlag, New York, 1997.
- [8] H. Jeffreys, Theory of Probability, Oxford University Press, London, 1961.
- [9] V. De Oliveira, Bayesian analysis of conditional autoregressive models, Ann. Inst. Statist. Math. 64 (2012) 107–133.
- [10] V. De Oliveira, Objective Bayesian analysis of spatial data with measurement error, Canad. J. Statist. 35 (2007) 283–301.

- [11] J.O. Berger, V. De Oliveira, B. Sansó, Objective Bayesian analysis of spatially correlated data, *J. Amer. Statist. Assoc.* 96 (2001) 1361–1374.
- [12] V. De Oliveira, J.J. Song, Bayesian analysis of simultaneous autoregressive models, *Sankhya* 70-B (2008) 323–350.
- [13] K. Ord, Estimation methods for models of spatial interaction, *J. Amer. Statist. Assoc.* 70 (1975) 120–126.
- [14] R. Paulo, Default priors for Gaussian processes, *Ann. Statist.* 33 (2005) 556–582.
- [15] C. Ren, Objective Bayesian analysis for SAR models, *Spat. Statist.* 4 (2013) 68–78.
- [16] C. Ren, D. Sun, Objective Bayesian analysis for CAR models, *Ann. Inst. Statist. Math.* 65 (2013) 457–472.
- [17] C. Ren, D. Sun, Z. He, Objective Bayesian analysis for a spatial model with nugget effects, *J. Statist. Plann. Inference* 142 (2012) 1933–1946.
- [18] H. Rue, L. Held, *Gaussian Markov Random fields: Theory and Applications*, Chapman and Hall/CRC, Boca Raton, 2005.
- [19] D. Sun, K. Tsutakawa, Z. He, Propriety of posteriors with improper priors in hierarchical linear mixed models, *Statist. Sinica* 11 (2001) 77–95.
- [20] J.C. Wakefield, A.E. Gelfand, A.F.M. Smith, Efficient generation of random variates via the ratio-of-uniforms method, *Stat. Comput.* 1 (1991) 129–133.
- [21] P. Whittle, On stationary processes in the plane, *Biometrika* 41 (1954) 434–449.