A BAYESIAN GENERALIZED CAR MODEL FOR CORRELATED SIGNAL DETECTION

D. Andrew Brown, Gauri S. Datta and Nicole A. Lazar

Clemson University and University of Georgia

Supplementary Material

Here the reader may find additional material, including the full conditional distributions for implementing the proposed model, further details of the proof of posterior propriety, and Supplementary Figures discussed in the main text.

S1 Full Conditional Distributions for the Proposed Model

Model (2.4) leads to the following full conditional distributions needed for a Gibbs sampling algorithm:

$$
\begin{aligned}
\mu_j \mid \boldsymbol{\mu}_{(-j)}, \boldsymbol{\gamma}, \sigma^2, \eta, \rho, p, \mathbf{y} &\sim N\left(\frac{\gamma_j y_j + \rho\eta^{-1}\sum_{i\neq j} w_{ji}\mu_i}{\gamma_j + \eta^{-1}(w_j. + d)},\ \sigma^2(\gamma_j + \eta^{-1}(w_j. + d))^{-1}\right), \quad j = 1, \ldots, J \\
\gamma_j \mid \boldsymbol{\mu}, \boldsymbol{\gamma}_{(-j)}, \sigma^2, \eta, \rho, p, \mathbf{y} &\sim \text{Bern}\left(\frac{(1-p)\exp\left(-(y_j-\mu_j)^2/2\sigma^2\right)}{(1-p)\exp\left(-(y_j-\mu_j)^2/2\sigma^2\right) + p\exp\left(-y_j^2/2\sigma^2\right)}\right),\ j = 1, \ldots, J \\
\sigma^2 \mid \boldsymbol{\mu}, \boldsymbol{\gamma}, \eta, \rho, p, \mathbf{y} &\sim \text{InvGam}\left(J,\ \frac{(\mathbf{y}-\boldsymbol{\Gamma}\boldsymbol{\mu})^T(\mathbf{y}-\boldsymbol{\Gamma}\boldsymbol{\mu}) + \eta^{-1}\boldsymbol{\mu}^T(\mathbf{D}_w^* - \rho\mathbf{W})\boldsymbol{\mu}}{2}\right) \\
p \mid \boldsymbol{\mu}, \boldsymbol{\gamma}, \sigma^2, \eta, \rho, \mathbf{y} &\sim \text{Beta}\left(J - \sum_{i=1}^{J}\gamma_i + \alpha, \sum_{i=1}^{J}\gamma_i + 1\right) \\
[\eta \mid \boldsymbol{\mu}, \boldsymbol{\gamma}, \sigma^2, \rho, p, \mathbf{y}] &\propto \eta^{-J/2-2}\exp\left(-\frac{1}{\eta}\left(\frac{\boldsymbol{\mu}^T(\mathbf{D}_w^* - \rho\mathbf{W})\boldsymbol{\mu}}{2\sigma^2}\right)\right)\left(\frac{\eta}{1+\eta}\right)^2,\ \eta > 0 \\
[\rho \mid \boldsymbol{\mu}, \boldsymbol{\gamma}, \sigma^2, \eta, p, \mathbf{y}] &\propto |\mathbf{D}_w^* - \rho\mathbf{W}|^{1/2}\exp\left(-\frac{\boldsymbol{\mu}^T(\mathbf{D}_w^* - \rho\mathbf{W})\boldsymbol{\mu}}{2\eta\sigma^2}\right) I(\nu_1^{-1} < \rho < \nu_J^{-1}),
\end{aligned}
\tag{S1.1}
$$

where $[\,\cdot \mid \cdot\,]$ denotes a conditional density.

S2 Additional Facts

In this Section, we provide a proof that the maximum value of ρ is an increasing function of d in Model (2.4). Also, we detail the simplifications used to obtain (5.1), (5.2), and (5.3) in the proof of posterior propriety.

S2.1 Relationship Between ρ and d

Consider two values $d_1 < d_2$. Let $\mathbf{x}_i^* = \arg\max_{\mathbf{x}} \mathbf{x}^T(\mathbf{D}_w+d_i\mathbf{I})^{-1/2}\mathbf{W}(\mathbf{D}_w+d_i\mathbf{I})^{-1/2}\mathbf{x}/\mathbf{x}^T\mathbf{x},\ i = 1, 2$, and let $\lambda_{J,i} > 0$ be the maximum eigenvalue of $(\mathbf{D}_w + d_i\mathbf{I})^{-1/2}\mathbf{W}(\mathbf{D}_w + d_i\mathbf{I})^{-1/2},\ i = 1, 2$. Then it follows from the properties of the Rayleigh quotient that

$$\begin{aligned}
\lambda_{J,2} &= \frac{\mathbf{x}_2^{*,T}(\mathbf{D}_w + d_2\mathbf{I})^{-1/2}\mathbf{W}(\mathbf{D}_w + d_2\mathbf{I})^{-1/2}\mathbf{x}_2^*}{\mathbf{x}_2^{*,T}\mathbf{x}_2^*} \\
&= \frac{1}{\mathbf{x}_2^{*,T}\mathbf{x}_2^*}\sum_i\sum_j \frac{w_{ij}x_{i,2}^*x_{j,2}^*}{\sqrt{w_{i\cdot} + d_2}\sqrt{w_{j\cdot} + d_2}} \\
&< \frac{1}{\mathbf{x}_2^{*,T}\mathbf{x}_2^*}\sum_i\sum_j \frac{w_{ij}x_{i,2}^*x_{j,2}^*}{\sqrt{w_{i\cdot} + d_1}\sqrt{w_{j\cdot} + d_1}} \\
&\leq \frac{\mathbf{x}_1^{*,T}(\mathbf{D}_w + d_1\mathbf{I})^{-1/2}\mathbf{W}(\mathbf{D}_w + d_1\mathbf{I})^{-1/2}\mathbf{x}_1^*}{\mathbf{x}_1^{*,T}\mathbf{x}_1^*} \\
&= \lambda_{J,1},
\end{aligned}$$

so $\lambda_{J,1}^{-1} < \lambda_{J,2}^{-1}$. The inequality in the third line follows from the Perron-Frobenius Theorem (e.g., Serre (2002)), which guarantees that the elements of the eigenvector associated to $\lambda_{J,2}$ are all positive. Thus, the maximum possible value for ρ is an increasing function in d.

S2.2 Facts Used in the Proof of Posterior Propriety

The simplifications make use of a simple inequality, which we state as a Lemma.

Lemma 1. *For any positive constants a, b, and c,*

$$\frac{b}{ba+c} < \frac{\max\{b,1\}}{a+c}.$$

Proof. Let $a, c \in \mathbb{R}^+$. Then, for $0 < b < 1$,

$$\begin{aligned}\frac{b}{ba+c} - \frac{1}{a+c} &= \frac{b(a+c) - ba - c}{(ba+c)(a+c)} \\ &= \frac{c(b-1)}{(ba+c)(a+c)} \\ &< 0,\end{aligned}$$

and for $b > 1$,

$$\begin{aligned}\frac{b}{ba+c} - \frac{b}{a+c} &= \frac{b(a+c) - b(ba+c)}{(ba+c)(a+c)} \\ &= \frac{ba(1-b)}{(ba+c)(a+c)} \\ &< 0.\end{aligned}$$

□

Now, the matrix $(w_{(J)} + d)\mathbf{I} - \mathbf{D}_w^* = w_{(J)}\mathbf{I} - \mathbf{D}_w$ is diagonal with nonnegative entries and thus positive semidefinite, which we denote as $(w_{(J)} + d)\mathbf{I} - \mathbf{D}_w^* \geq 0$.

Let $\nu_1 \leq \nu_2 \leq \cdots \leq \nu_J$ be the ordered eigenvalues of $\mathbf{W}^*$. Then the eigenvalues of $\mathbf{I} - \rho\mathbf{W}^*$ are $1 - \rho\nu_j,\ j = 1, \ldots, J$. Hence,

$$\begin{aligned}\rho \in (\nu_1^{-1}, \nu_J^{-1}) &\Rightarrow 1 - \rho\nu_j > 0,\ \forall j \\ &\Rightarrow (\mathbf{I} - \rho\mathbf{W}^*) > 0 \\ &\Rightarrow \eta(\mathbf{I} - \rho\mathbf{W}^*)^{-1} > 0.\end{aligned}$$

By adding and subtracting $\eta(\mathbf{I}-\rho\mathbf{W}^*)^{-1}$, we obtain

$$(w_{(J)}+d)\mathbf{I}-\mathbf{D}_w^* = (w_{(J)}+d)\mathbf{I}+\eta(\mathbf{I}-\rho\mathbf{W}^*)^{-1}-(\mathbf{D}_w^*+\eta(\mathbf{I}-\rho\mathbf{W}^*)^{-1}) \geq 0.$$

Making use of the fact that $\mathbf{B}>0, \mathbf{A}-\mathbf{B}\geq 0 \Rightarrow \mathbf{B}^{-1}-\mathbf{A}^{-1}\geq 0$, it follows that

$$(\mathbf{D}_w^*+\eta(\mathbf{I}-\rho\mathbf{W}^*)^{-1})^{-1}-((w_{(J)}+d)\mathbf{I}+\eta(\mathbf{I}-\rho\mathbf{W}^*)^{-1})^{-1}\geq 0$$

$$\Rightarrow \mathbf{x}^T(\mathbf{D}_w^*+\eta(\mathbf{I}-\rho\mathbf{W}^*)^{-1})^{-1}\mathbf{x} \geq \mathbf{x}^T((w_{(J)}+d)\mathbf{I}+\eta(\mathbf{I}-\rho\mathbf{W}^*)^{-1})^{-1}\mathbf{x}$$

$$\Rightarrow (\mathbf{x}^T(\mathbf{D}_w^*+\eta(\mathbf{I}-\rho\mathbf{W}^*)^{-1})^{-1}\mathbf{x})^{-J/2} \leq (\mathbf{x}^T((w_{(J)}+d)\mathbf{I}+\eta(\mathbf{I}-\rho\mathbf{W}^*)^{-1})^{-1}\mathbf{x})^{-J/2}.$$

Now, $\mathbf{W}^*$ is symmetric, so it has a spectral decomposition of the form $\mathbf{W}^*=\mathbf{PMP}^T$, where $\mathbf{P}$ is the orthogonal matrix of eigenvectors of $\mathbf{W}^*$ and $\mathbf{M}$ is the diagonal matrix of eigenvalues. Let $\mathbf{u}=\mathbf{P}^T\mathbf{x} \Rightarrow \mathbf{x}=\mathbf{Pu}$ so that

$$\begin{aligned}
\mathbf{x}^T((w_{(J)}+d)\mathbf{I}+\eta(\mathbf{I}-\rho\mathbf{W}^*)^{-1})^{-1}\mathbf{x} &= \mathbf{u}^T\mathbf{P}^T((w_{(J)}+d)\mathbf{I}+\eta(\mathbf{I}-\rho\mathbf{W}^*)^{-1})^{-1}\mathbf{Pu} \\
&= \mathbf{u}^T(\mathbf{P}^T(w_{(J)}+d)\mathbf{P}+\eta\mathbf{P}^T(\mathbf{I}-\rho\mathbf{W}^*)^{-1}\mathbf{P})^{-1}\mathbf{u} \\
&= \mathbf{u}^T((w_{(J)}+d)\mathbf{I}+\eta(\mathbf{I}-\rho\mathbf{M})^{-1})^{-1}\mathbf{u} \\
&= \sum_{j=1}^{J}\frac{(1-\rho\nu_j)u_j^2}{(w_{(J)}+d)(1-\rho\nu_j)+\eta}.
\end{aligned}$$

Since $\mathbf{D}_w^*$ is diagonal and the diagonal elements of $\mathbf{W}$ are zero, the diagonal elements of $\mathbf{W}^*=(\mathbf{D}_w^*)^{-1/2}\mathbf{W}(\mathbf{D}_w^*)^{-1/2}$ are zero and thus $\mathrm{tr}(\mathbf{W}^*)=0=\sum_{j=1}^{J}\nu_j$. It must then be true that there are $r_1>0$ negative eigenvalues and $r_2>0$ positive eigenvalues of $\mathbf{W}^*$, since $r:=r_1+r_2=\mathrm{rank}(\mathbf{W}^*)>0$. The summation in the last line can then be separated according to the sign of the eigenvalue in each term as

$$\underbrace{\sum_{j=1}^{r_1}\frac{(1-\rho\nu_j)u_j^2}{(w_{(J)}+d)(1-\rho\nu_j)+\eta}}_{\nu_j<0}+\underbrace{\sum_{j=J-r_2+1}^{J}\frac{(1-\rho\nu_j)u_j^2}{(w_{(J)}+d)(1-\rho\nu_j)+\eta}}_{\nu_j>0}+\underbrace{\sum_{j=r_1+1}^{J-r_2}\frac{u_j^2}{w_{(J)}+d+\eta}}_{\nu_j=0}. \quad \text{(S2.2)}$$

If $\nu_1^{-1} < \rho < 0$, then $0 < 1-\rho\nu_j < 1$ for $j = 1, \ldots, r_1$ and $1-\rho\nu_j > 1$ for $j = J-r_2+1, \ldots, J$, so

$$\begin{aligned}(S2.2) &\geq \sum_{j=J-r_2+1}^{J} \frac{(1-\rho\nu_j)u_j^2}{(w_{(J)}+d)(1-\rho\nu_j)+\eta} + \sum_{j=r_1+1}^{J-r_2} \frac{u_j^2}{w_{(J)}+d+\eta} \\ &\geq \sum_{j=J-r_2+1}^{J} \frac{u_j^2}{w_{(J)}+d+\eta} + \sum_{j=r_1+1}^{J-r_2} \frac{u_j^2}{w_{(J)}+d+\eta} \\ &= (w_{(J)}+d+\eta)^{-1} \sum_{j=r_1+1}^{J} u_j^2,\end{aligned}$$

where the second line follows from noticing that $(w_{(J)}+d+\eta)^{-1} - (1-\rho\nu_j)((w_{(J)}+d)(1-\rho\nu_j)+\eta)^{-1} < 0$ for $\nu_j > 0$. Similarly, if $0 < \rho < \nu_J^{-1}$, then

$$\begin{aligned}(S2.2) &\geq \sum_{j=1}^{r_1} \frac{(1-\rho\nu_j)u_j^2}{(w_{(J)}+d)(1-\rho\nu_j)+\eta} + \sum_{j=r_1+1}^{J-r_2} \frac{u_j^2}{w_{(J)}+d+\eta} \\ &\geq \sum_{j=1}^{r_1} \frac{u_j^2}{w_{(J)}+d+\eta} + \sum_{j=r_1+1}^{J-r_2} \frac{u_j^2}{w_{(J)}+d+\eta} \\ &= (w_{(J)}+d+\eta)^{-1} \sum_{j=1}^{J-r_2} u_j^2.\end{aligned}$$

Thus, we have that for all $\rho \in (\nu_1^{-1}, \nu_J^{-1})$,

$$\mathbf{x}^T((w_{(J)}+d)\mathbf{I} + \eta(\mathbf{I}-\rho\mathbf{W}^*)^{-1})^{-1}\mathbf{x} \geq k(w_{(J)}+d+\eta)^{-1}$$

$$\Rightarrow (\mathbf{x}^T((w_{(J)}+d)\mathbf{I} + \eta(\mathbf{I}-\rho\mathbf{W}^*)^{-1})^{-1}\mathbf{x})^{-J/2} \leq k'(w_{(J)}+d+\eta)^{J/2}$$

where $0 < k' < \infty$ is constant. This establishes (5.1).

To see that (5.2) holds, note that

$$\begin{aligned} |\mathbf{I}+\eta(\mathbf{D}_w^*-\rho\mathbf{W})^{-1}| &= |(\mathbf{D}_w^*)^{-1/2}||\mathbf{D}_w^*+\eta(\mathbf{I}-\rho\mathbf{W}^*)^{-1}||(\mathbf{D}_w^*)^{-1/2}| \\ &= |(\mathbf{D}_w^*)|^{-1}|\mathbf{D}_w^*+\eta(\mathbf{I}-\rho\mathbf{W}^*)^{-1}| \\ &\equiv k|\mathbf{D}_w^*+\eta(\mathbf{I}-\rho\mathbf{W}^*)^{-1}|. \end{aligned}$$

Letting $w_{_{(1)}} =_{1\le j\le J}^{\min} w_j.$, and again adding and subtracting $\eta(\mathbf{I}-\rho\mathbf{W}^*)^{-1}$, we obtain

$$\mathbf{D}_w^*-(w_{_{(1)}}+d)\mathbf{I}\ge 0 \Rightarrow \mathbf{D}_w^*+\eta(\mathbf{I}-\rho\mathbf{W}^*)^{-1}-((w_{_{(1)}}+d)\mathbf{I}+\eta(\mathbf{I}-\rho\mathbf{W}^*)^{-1})\ge 0.$$

But $\mathbf{B}>0,\ \mathbf{A}-\mathbf{B}\ge 0$ implies $|\mathbf{A}|\ge|\mathbf{B}|$, so we find that

$$|\mathbf{D}_w^*+\eta(\mathbf{I}-\rho\mathbf{W}^*)^{-1}|\ge|(w_{_{(1)}}+d)\mathbf{I}+\eta(\mathbf{I}-\rho\mathbf{W}^*)^{-1}|$$

$$\Rightarrow k|\mathbf{D}_w^*+\eta(\mathbf{I}-\rho\mathbf{W}^*)^{-1}|\ge K|(w_{_{(1)}}+d)\mathbf{I}+\eta(\mathbf{I}-\rho\mathbf{W}^*)^{-1}|.$$

The eigenvalues of $(\mathbf{I}-\rho\mathbf{W}^*)^{-1}$ are $(1-\rho\nu_j)^{-1}$, $j=1,\ldots,J$, so it follows that the eigenvalues of $(w_{_{(1)}}+d)\mathbf{I}+\eta(\mathbf{I}-\rho\mathbf{W}^*)^{-1}$ are $w_{_{(1)}}+d+\eta(1-\rho\nu_j)^{-1}$, $j=1,\ldots,J$. Therefore,

$$\begin{aligned} k|(w_{_{(1)}}+d)\mathbf{I}+\eta(\mathbf{I}-\rho\mathbf{W}^*)^{-1}| &= k\prod_{j=1}^{J}(w_{_{(1)}}+d+\eta(1-\rho\nu_j)^{-1}) \\ &= \frac{k\prod_{j=1}^{J}((1-\rho\nu_j)(w_{_{(1)}}+d)+\eta)}{\prod_{j=1}^{J}(1-\rho\nu_j)}, \end{aligned}$$

and subsequently

$$|\mathbf{I}+\eta(\mathbf{D}_w^*-\rho\mathbf{W})^{-1}|^{-1/2}\le k'\left(\frac{\prod_{j=1}^{J}((1-\rho\nu_j)(w_{_{(1)}}+d)+\eta)}{\prod_{j=1}^{J}(1-\rho\nu_j)}\right)^{-1/2},$$

where $0 < k' < \infty$ is constant. But $1 - \rho\nu_j > 0,$ for all j, so by Lemma 1

$$\frac{1-\rho\nu_j}{(1-\rho\nu_j)(w_{(1)}+d)+\eta} \leq \frac{\max\{1-\rho\nu_j, 1\}}{w_{(1)}+d+\eta}, \quad \forall j \in \{1,\ldots,J\}$$

$$\Rightarrow \frac{\prod_{j=1}^{J}(1-\rho\nu_j)}{\prod_{j=1}^{J}((1-\rho\nu_j)(w_{(1)}+d)+\eta)} \leq \frac{\prod_{j=1}^{J}\max\{1-\rho\nu_j, 1\}}{(w_{(1)}+d+\eta)^J}$$

References

Serre, D. (2002). *Matrices: Theory and applications*. Springer-Verlag, New York.

S3 Supplementary Figures

Figure 1: Comparison of the $|t_2|$ prior on τ (Gelman, 2006) (after transforming to the τ^2 scale) and the prior on $\tau^2 \mid \sigma^2$ suggested by Scott and Berger (2006), denoted by SB. Here, the scale parameter is set to 1 in both densities.

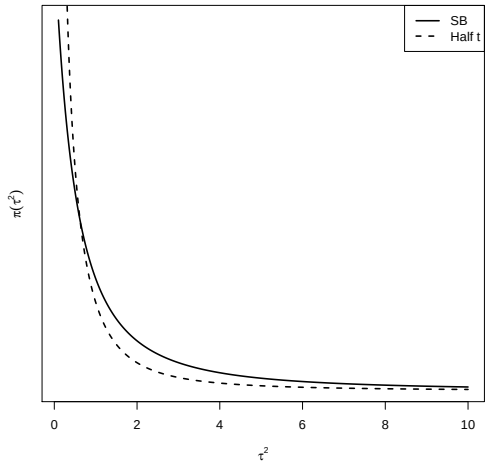

Figure 2: Simulated binary activation pattern drawn from an Ising distribution.

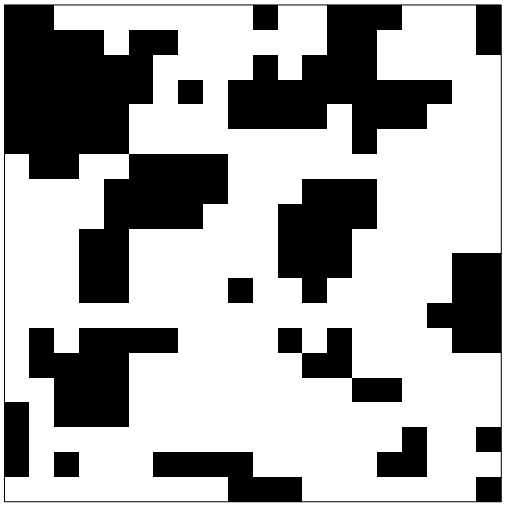

Figure 3: Simulated data using the activation pattern in Figure 2 with non-null mean 3.5.

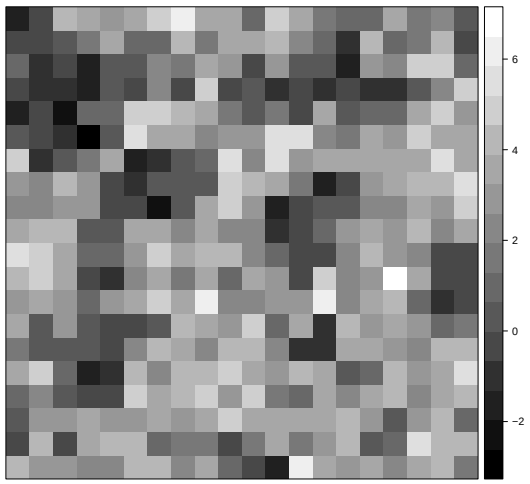

Figure 4: Neighborhood structures and weights used for the Bayesian CAR model in the simulation study. In each illustration, the center square represents the gene of interest, and the numbers are the weights w_{ij} assigned to each gene in the neighborhood.

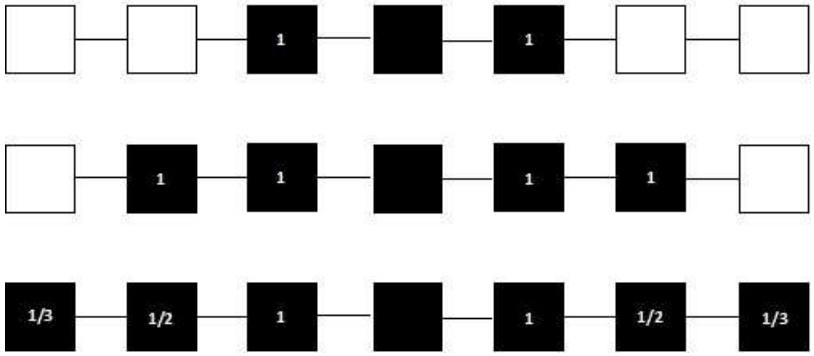

Figure 5: Estimated posterior inclusion probabilities p_i versus test statistics y_i, $i = 1, \ldots, 1000$ for the simulated microarray data with neighborhoods determined through physical adjacency. The dashed (jagged) curve results from the CAR($\mathbf{W}_1$) model, and the dotted (smooth) line results from the SB model. The circled point corresponds to the indicated statistic depicted in Figure 6.

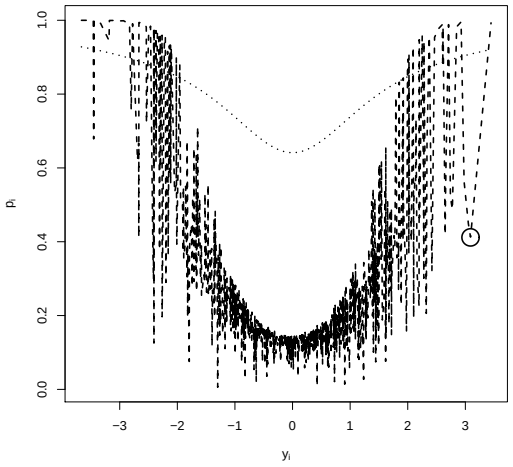

Figure 6: Graphical depiction of test statistics from the simulated microarray data. The test statistics y_i are arranged in order, $i = 1, 2, \ldots, 1000$, going from the lower left to the upper right, row-wise from left to right. The circle indicates the statistic corresponding to the circled (y_i, p_i) point in Figure 5.

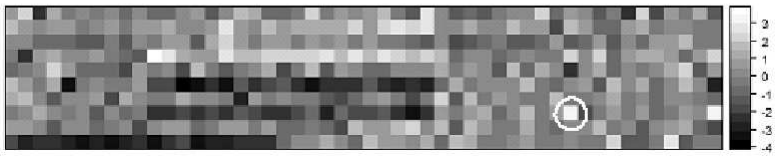

Figure 7: Empirical ROC curves for the testing model using the physical-adjacency CAR model and the generalized CAR using pathways to define neighborhoods with isolated points included.

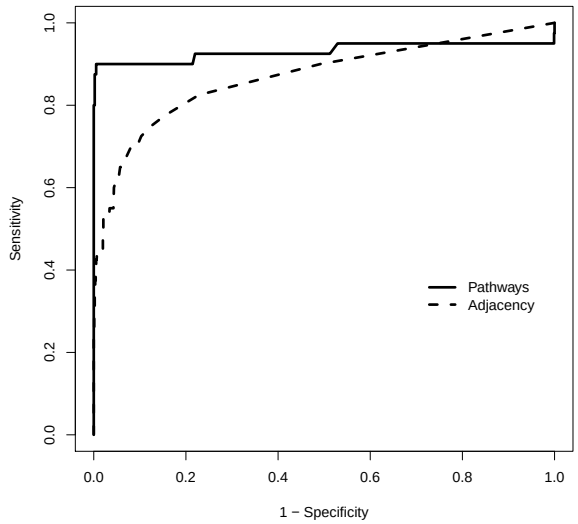

Figure 8: Histogram of test statistics from the simulated pathways example with normal densities superimposed, each with mean zero and standard deviations estimated as $\sqrt{E(\sigma^2 \mid \mathbf{y})}$ from the posterior distributions of both neighborhood structures. The tick marks at the bottom indicate the non-null cases.

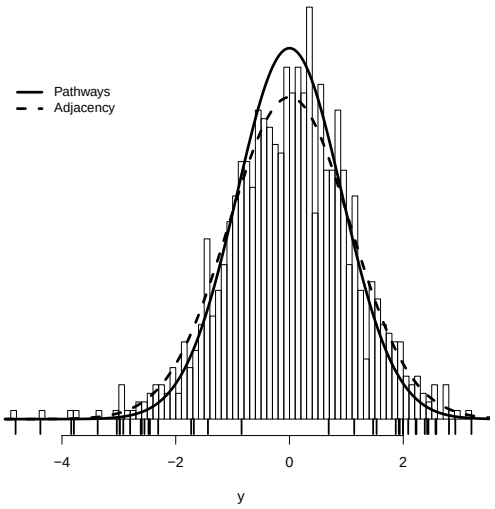

Figure 9: Posterior inclusion probabilities estimated under the CAR testing model with and without the isolated cases. The dashed line is at the 0.95 threshold. Notice that several of the P5 cases have been pulled downward, resulting in more false non-discoveries by excluding the isolated cases.

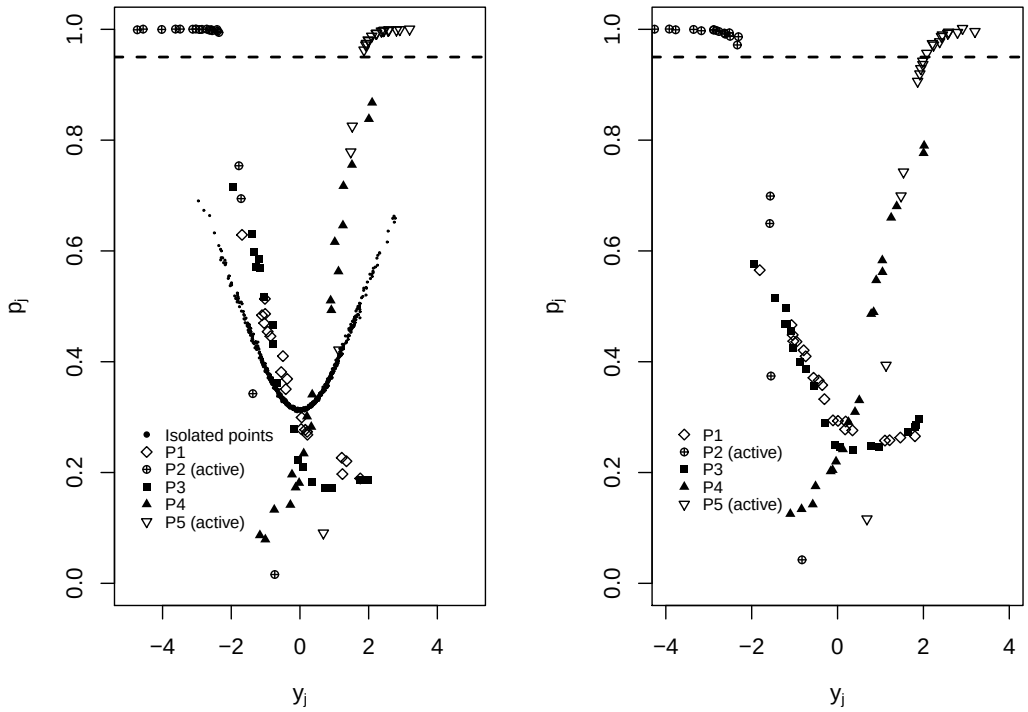

Figure 10: Histograms of realizations of Moran's I calculated from the posterior predictive distribution of the CAR testing model, $p(I(\mathbf{y}^*) \mid \mathbf{y}) = \int_{\boldsymbol{\theta}} p(I(\mathbf{y}^*) \mid \boldsymbol{\theta})\pi(\boldsymbol{\theta} \mid \mathbf{y})d\boldsymbol{\theta}$, under partially incorrect correlation assumptions. The dark vertical line indicates the observed value of I under the assumed neighborhood structure.

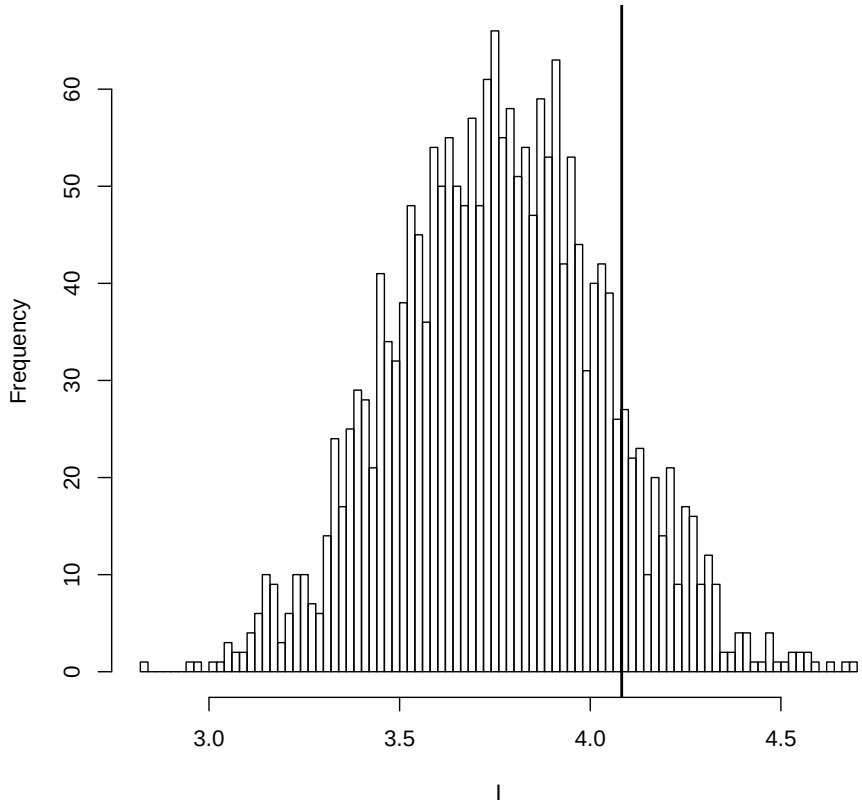

Figure 11: Smoothed posterior densities of p and ρ for the E. Coli data (Xiao, Reilly, and Khodursky, 2009) with different values of α in the prior $p \sim \text{Beta}(\alpha, 1)$.

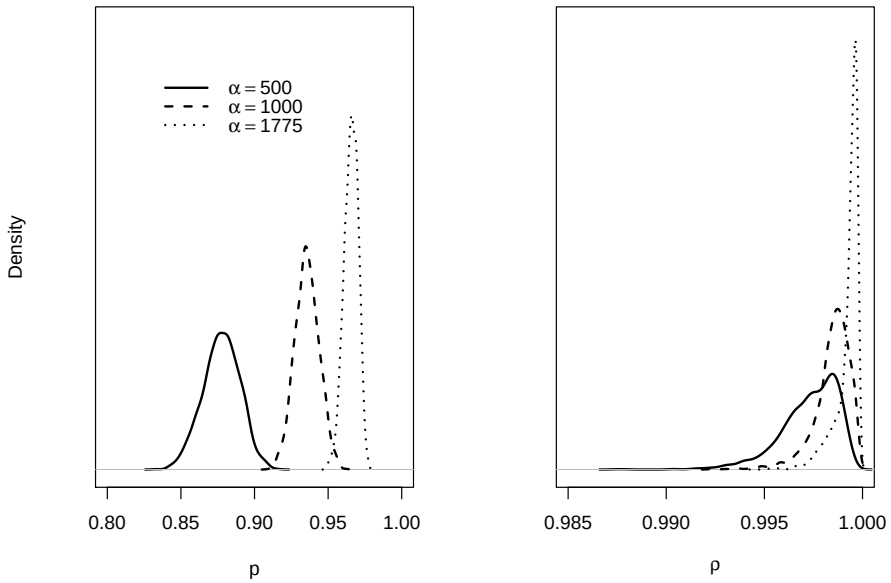

Figure 12: Estimated posterior inclusion probabilities for each of the 22,283 genes in the lymphoblastoid cell data (Subramanian, Tamayo, Mootha, Mukherjee, Ebert, Gillette, Paulovich, Pomeroy, Golub, Lander, and Meslrov, 2005) under both CAR testing models with and without isolated cases. The left panel results from excluding the isolated cases ($d = 0$), the right panel results from including isolated cases ($d = 1$). The horizontal lines represent the 0.99 threshold.