Appendix: Sampler Derivations for Algorithm I.

Following Godsill and Rayner (1998), we define the blocks $z_j = (\gamma_j, \beta_j)$ for $j = 1, \ldots, m$. Thus, each z_j can be viewed as a two-dimensional parameter. The idea is to update $(z_1, \ldots, z_m, \sigma^2, \tau^2)$ using a standard Gibbs sampler. For any $j = 1, \ldots, m$, we let $\tilde{\gamma}_{-j} = (\gamma_1, \ldots, \gamma_{j-1}, \gamma_{j+1}, \ldots, \gamma_m)$ and $\tilde{\beta}_{-j} = (\beta_1, \ldots, \beta_{j-1}, \beta_{j+1}, \ldots, \beta_m)'$.

Sampling γ and β. It follows from (2.2) that

$$\begin{aligned}
& p(\beta_j, \gamma_j = 1 | z_1, \ldots, z_{j-1}, z_{j+1}, \ldots, z_m, \sigma, \tau, \mathbf{y}, X) \\
& \propto \left(\frac{1}{\sqrt{2\pi}\sigma}\right)^n \exp\left(-\frac{\|\mathbf{y} - X\beta\|^2}{2\sigma^2}\right) \left(\frac{1}{\sqrt{2\pi}\tau}\right) \exp\left(-\frac{\beta_j^2}{2\tau^2}\right) P(\gamma_j = 1 | \tilde{\gamma}_{-j}) \\
= & \left(\frac{1}{\sqrt{2\pi}\sigma}\right)^n \exp\left(-\frac{\|\mathbf{y} - X_j\beta_j - X_{-j}\tilde{\beta}_{-j}\|^2}{2\sigma^2}\right) \left(\frac{1}{\sqrt{2\pi}\tau}\right) \exp\left(-\frac{\beta_j^2}{2\tau^2}\right) P(\gamma_j = 1 | \tilde{\gamma}_{-j}) \\
= & \left(\frac{1}{\sqrt{2\pi}\sigma}\right)^n \exp\left(-\frac{\beta_j^2 v_j^2 - 2(\mathbf{y} - X_{-j}\tilde{\beta}_{-j})' X_j \beta_j}{2\sigma^2}\right) p(\gamma_j = 1 | \tilde{\gamma}_{-j}) \\
& \cdot \left(\frac{1}{\sqrt{2\pi}\tau}\right) \exp\left(-\frac{\|\mathbf{y} - X_{-j}\tilde{\beta}_{-j}\|^2}{2\sigma^2}\right), \qquad (0.1)
\end{aligned}$$

where $v_j^2 = X_j' X_j + \sigma^2/\tau^2$. Similarly, we have

$$\begin{aligned}
& p(\beta_j, \gamma_j = 0 | z_1, \ldots, z_{j-1}, z_{j+1}, \ldots, z_m, \sigma, \tau, \mathbf{y}, X) \\
\propto & \left(\frac{1}{\sqrt{2\pi}\sigma}\right)^n \delta_0(\beta_j) p(\gamma_j = 0 | \tilde{\beta}_{-j}) \exp\left(-\frac{\|\mathbf{y} - X_{-j}\tilde{\beta}_{-j}\|^2}{2\sigma^2}\right). \qquad (0.2)
\end{aligned}$$

Integrating out β_j in (0.1) and (0.2), we have

$$\begin{aligned}
& p(\gamma_j = 1 | z_1, \ldots, z_{j-1}, z_{j+1}, \ldots, z_m, \sigma, \tau, \mathbf{y}, X) \\
\propto & \left(\frac{1}{\sqrt{2\pi}\sigma}\right)^n \exp\left(\frac{\tilde{u}_j^2}{2\sigma^2 v_j^2}\right) \left(\frac{\sigma}{\tau v_j}\right) p(\gamma_j = 1 | \tilde{\gamma}_{-j}) \exp\left(-\frac{\|\mathbf{y} - X_{-j}\tilde{\beta}_{-j}\|^2}{2\sigma^2}\right),
\end{aligned}$$

and

$$\begin{aligned}
& p(\gamma_j = 0 | z_1, \ldots, z_{j-1}, z_{j+1}, \ldots, z_m, \sigma, \tau, \mathbf{y}, X) \\
\propto & \left(\frac{1}{\sqrt{2\pi}\sigma}\right)^n p(\gamma_j = 0 | \tilde{\gamma}_{-j}) \exp\left(-\frac{\|\mathbf{y} - X_{-j}\tilde{\beta}_{-j}\|^2}{2\sigma^2}\right),
\end{aligned}$$

where $\tilde{u}_j = (\mathbf{y} - X_{-j}\tilde{\beta}_{-j})'X_j$. Therefore,

$$p(\gamma_j = 1|z_1, \ldots, z_{j-1}, z_{j+1}, \ldots, z_m, \sigma, \tau, \mathbf{y}, X) = 1/(1 + \tilde{\rho}_j),$$

where $\tilde{\rho}_j = \frac{p(\gamma_j=0|\tilde{\gamma}_{-j})}{p(\gamma_j=1|\tilde{\gamma}_{-j})} \cdot \frac{\tau v_j}{\sigma} \exp\left(-\frac{\tilde{u}_j^2}{2\sigma^2 v_j^2}\right)$. It follows from (0.1) and (0.2) that

$$p(\beta_j = 0|\gamma_j = 0, z_1, \ldots, z_{j-1}, z_{j+1}, \ldots, z_m, \sigma, \tau, \mathbf{y}, X) = 1$$

and

$$\beta_j|\gamma_j = 1, z_1, \ldots, z_{j-1}, z_{j+1}, \ldots, z_m, \sigma, \tau, \mathbf{y}, X \sim N\left(\frac{\tilde{u}_j}{v_j^2}, \frac{\sigma^2}{v_j^2}\right).$$

Using the above procedure, all the blocks z_js can be updated.

Sampling σ and τ. Step (B) in Algorithm I follows directly from (2.2).