A Bayesian approach to envelope quantile regression

Minji Lee, Saptarshi Chakraborty and Zhihua Su

Edwards Lifesciences

Department of Biostatistics, State University of New York at Buffalo

Department of Statistics, University of Florida

Supplementary Material

In this supplement, we provide definition of some distributions used for the priors and the concept of ergodicity in Section S1. Notes on prior distributions are included in Section S2. Section S3 contains proofs of theorems, and Section S4 derives various full conditional posterior densities used in Algorithm 1 of the main text. Section S5 presents an expectation conditional maximization algorithm for maximum a posteriori estimation in Bayesian quantile envelope model, and Section S6 proposes an extension of Algorithm 1 to handle Tobit censored responses. Additional simulations and data analysis are available in Section S7.

S1 Definitions

S1.1 Definitions of distributions

In this section, we introduce some distributions used in prior construction and augmented data in Section 3.2 of the main article.

Definition 1. Let $\Gamma_m(\cdot)$ be the multivariate gamma function of dimension m, ν be a real number with $\nu > m-1$, and $\boldsymbol{\Psi}$ be a fixed $m \times m$ positive definite matrix. Then a random matrix $\boldsymbol{U} \in \mathbb{R}^{m\times m}$ is said to follow an m-dimensional inverse Wishart distribution with parameters $\boldsymbol{\Psi}$ and ν, denoted by $\boldsymbol{U} \sim \mathrm{IW}_m(\boldsymbol{\Psi},\ \nu)$, if its density function is given by:

$$\frac{|\boldsymbol{\Psi}|^{\nu/2}}{2^{(\nu m)/2}\Gamma_m(\nu/2)}|\boldsymbol{U}|^{-(\nu+m+1)/2}\exp\left(-\frac{1}{2}\,\mathrm{trace}(\boldsymbol{\Psi}\boldsymbol{U}^{-1})\right).$$

Definition 2. A random matrix $\boldsymbol{U} \in \mathbb{R}^{m_1\times m_2}$ is said to follow a matrix normal distribution with mean $\boldsymbol{U}_0$ and covariance matrices $\boldsymbol{B}$ and $\boldsymbol{C}$, denoted by $\boldsymbol{U} \sim \mathrm{MN}_{m_1,m_2}(\boldsymbol{U}_0,\ \boldsymbol{B},\ \boldsymbol{C})$, if $\boldsymbol{U}$ has density

$$\frac{|\boldsymbol{B}|^{-m_2/2}|\boldsymbol{C}|^{-m_1/2}}{(2\pi)^{(m_1m_2)/2}}\exp\left[-\frac{1}{2}\,\mathrm{trace}\left\{\boldsymbol{B}^{-1}(\boldsymbol{U}-\boldsymbol{U}_0)\,\boldsymbol{C}^{-1}(\boldsymbol{U}-\boldsymbol{U}_0)^T\right\}\right],$$

where $\boldsymbol{U}_0 \in \mathbb{R}^{m_1\times m_2}$, and $\boldsymbol{B} \in \mathbb{R}^{m_1\times m_1}$ and $\boldsymbol{C} \in \mathbb{R}^{m_2\times m_2}$ are positive definite matrices.

Definition 3. A random variable U is said to have a generalized inverse Gaussian distribution with parameters $\alpha > 0$, $\beta > 0$ and $\kappa \in \mathbb{R}$, denoted by $U \sim \text{GIG}\,(\alpha,\ \beta,\ \kappa)$, if the density of U is proportional to $u^{(\kappa-1)}\exp\{-(\alpha/u+\beta u)/2\}\mathbb{1}\{u > 0\}$, where $\mathbb{1}(\cdot)$ denotes the indicator function.

S1.2 A brief note on Harris ergodicity

As mentioned in Section 3.3 of the main text, Harris ergodicity of a Markov chain Monte Carlo algorithm guarantees that for any starting point, not just those outside of some pathological set of measure zero, the algorithm converges to its stationarity. Harris ergodicity consists of three properties of a Markov chain, namely, ϕ-irreducibility, Harris recurrence and aperiodicity. Below we define these properties (Meyn and Tweedie, 2012; Roberts and Rosenthal, 2006).

Definition 4. Let $(X_n)_{n=0}^{\infty}$ be a Markov chain with state space $\mathcal{X}$ and n-step transition probabilities $P^n(x, \cdot)$. Given a nonzero measure ϕ on $\mathcal{X}$, the Markov chain $(X_n)_{n=0}^{\infty}$ is called *ϕ-irreducible* if for any point $x \in \mathcal{X}$ and any measurable set A with $\phi(A) > 0$ there exists an integer n such that $P^n(x, A) > 0$.

Definition 5. Let $(X_n)_{n=0}^{\infty}$ be a ϕ-irreducible Markov chain with state space $\mathcal{X}$ and stationary probability distribution π. Then $(X_n)_{n=0}^{\infty}$ is said to be

Harris recurrent if for all $A \subseteq \mathcal{X}$ with $\pi(A) > 0$ and for any starting point $x \in \mathcal{X}$, $P(X_n \in A \text{ infinitely often in } n \mid X_0 = x) = 1$.

Definition 6. Let $(X_n)_{n=0}^{\infty}$ be a ϕ-irreducible Markov chain with transition probabilities $P(x, \cdot)$ on a state space $\mathcal{X}$ with σ-algebra $\mathcal{F}$, and with stationary probability distribution π. The *period* of $(X_n)_{n=0}^{\infty}$ is defined as the largest positive integer D for which there exist disjoint subsets $A_1, A_2, \ldots, A_D \in \mathcal{F}$ with $\pi(A_i) > 0$, such that $P(x, A_{i+1}) = 1$ for all $x \in A_i$ $(1 \leq i \leq D-1)$, and $P(x, A_1) = 1$ for all $x \in A_D$. If $D = 1$, then the chain is *aperiodic*.

S2 Notes on prior distributions

S2.1 Choice of prior distributions

For all parameters except for $\boldsymbol{A}$, we considered standard conjugate prior distributions. The parameter $\boldsymbol{A}$ defines a subspace without resorting to any manifold structure (Cook, Forzani and Su, 2016). For $\boldsymbol{A}$, we use a matrix normal prior, a straightforward generalization of the multivariate normal distribution, for two main reasons. First, the matrix normal family provides simple parametric distributions for unrestricted Euclidean random matrices wherein prior dependencies between rows and between columns of the random matrix can be entirely expressed through the prior covariance matrices. In the examples considered in the paper we specified these prior covariance parameters to be scalar multiples of the identity matrix, which ensures prior independence. Second, the prior mode of the matrix normal distribution is identical to the prior mean, which facilitate straightforward incorporation of subjective prior information to the model. In particular, if *a priori* information on the *most likely* envelope subspace (which can be easier to interpret than the *mean* envelope subspace) is available, one can identify this subspace with a Euclidean matrix $\boldsymbol{A}_0$ following the process described in the paper, and set $\boldsymbol{A}_0$ as the prior mode of the matrix normal prior distribution.

It is of note however that the form of the prior distribution for $\boldsymbol{A}$ does not substantially alter the sampling steps of all the other parameters presented in the proposed MCMC algorithm. The prior density of $\boldsymbol{A}$ only affects the Metropolis step for sampling from the conditional distribution of $\boldsymbol{A}$. The effect is small unless a very strong prior is assumed. Thus, other matrix variate distributions, such as a matrix t-distribution, could in principle be used as an alternative prior for $\boldsymbol{A}$ without substantially affecting the MCMC sampler.

S2.2 Comments on the prior distributions in Khare et al. (2017)

We choose not to use the prior specification in Khare et al. (2017) because it depends on the model structure. The Bayesian envelope model constructed in Khare et al. (2017) was in the context of multivariate (response) linear regression, where the envelope structure was imposed on the multivariate response. This model structure was utilized in the choice of the prior distributions in Khare et al. (2017) such that conjugacy is obtained for all the model parameters except for $(\boldsymbol{\Gamma}, \boldsymbol{\Gamma}_0)$. Moreover, although $(\boldsymbol{\Gamma}, \boldsymbol{\Gamma}_0)$ does not enjoy the conjugacy, its conditional posterior distribution can be sampled through a generalized matrix Bingham distribution. Now we are in a different model, the quantile regression, where the response is univariate and the envelope structure was imposed on the predictors instead of the responses. If we adopt the same prior specification as in Khare et al. (2017), we may lose the conjugacy of the parameters. In addition, the conditional posterior distribution of $(\boldsymbol{\Gamma}, \boldsymbol{\Gamma}_0)$ would be more complicated in our case and could not be sampled through a generalized matrix Bingham distribution.

S3 Proofs of Theorems

Proof of Theorem 1

To prove Theorem 1, we will show that the posterior density is integrable with respect to the Lebesgue measure on the parameter space. We shall denote by $\int f(t)\ dt$ an appropriate Lebesgue integral, and by $\mathbf{1}_n$ an n-component vector with all entries being equal to 1. The complete-data log posterior density is given by

$$
\begin{aligned}
&\log \pi\left(\mu_{\tau,Y}, \sigma, \boldsymbol{\mu}_{\boldsymbol{X}}, \boldsymbol{\eta}, \boldsymbol{A}, \boldsymbol{\Omega}_1, \boldsymbol{\Omega}_2, \mathbb{Z} | \mathbb{X}, \mathbb{W}\right) \\
&= \text{const.} - \frac{n}{2}\log\sigma - \frac{n}{2}\log|\boldsymbol{\Omega}_1| - \frac{n}{2}\log|\boldsymbol{\Omega}_2| \\
&- \frac{1}{2\sigma\gamma^2}\left[\left\{\mathbb{W} - \mu_{\tau,Y}\mathbf{1}_n - \left(\mathbb{X} - \mathbf{1}_n\boldsymbol{\mu}_{\boldsymbol{X}}^T\right)\boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})\boldsymbol{\eta}\right\}^T \boldsymbol{D}^{-1}\right. \\
&\qquad\qquad\left.\left\{\mathbb{W} - \mu_{\tau,Y}\mathbf{1}_n - \left(\mathbb{X} - \mathbf{1}_n\boldsymbol{\mu}_{\boldsymbol{X}}^T\right)\boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})\boldsymbol{\eta}\right\}\right] \\
&- \frac{1}{2}\operatorname{trace}\left\{\left(\mathbb{X} - \mathbf{1}_n\boldsymbol{\mu}_{\boldsymbol{X}}^T\right)\left(\boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})\boldsymbol{\Omega}_1^{-1}\boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})^T + \boldsymbol{\Gamma}_{2\tau}(\boldsymbol{A})\boldsymbol{\Omega}_2^{-1}\boldsymbol{\Gamma}_{2\tau}(\boldsymbol{A})^T\right)\right. \\
&\qquad\qquad\left.\left(\mathbb{X} - \mathbf{1}_n\boldsymbol{\mu}_{\boldsymbol{X}}^T\right)^T\right\} \\
&- \frac{u_\tau}{2}\log\sigma - \frac{1}{2\sigma\gamma^2}\left(\boldsymbol{\eta} - \boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})^T\boldsymbol{e}\right)^T\boldsymbol{M}\left(\boldsymbol{\eta} - \boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})^T\boldsymbol{e}\right) \\
&- (a+1)\log\sigma - b/\sigma - \frac{\nu_1 + u_\tau + 1}{2}\log|\boldsymbol{\Omega}_1| \\
&- \frac{1}{2}\operatorname{trace}\left(\boldsymbol{\Omega}_1^{-1}\boldsymbol{\Psi}_1\right) - \frac{\nu_2 + p - u_\tau + 1}{2}\log|\boldsymbol{\Omega}_2| - \frac{1}{2}\operatorname{trace}\left(\boldsymbol{\Omega}_2^{-1}\boldsymbol{\Psi}_2\right) \\
&- \frac{1}{2}\operatorname{trace}\left\{\boldsymbol{K}^{-1}\left(\boldsymbol{A} - \boldsymbol{A}_1\right)\boldsymbol{L}^{-1}\left(\boldsymbol{A} - \boldsymbol{A}_1\right)^T\right\}
\end{aligned}
$$

$$- n \log \sigma - \frac{1}{\sigma} \sum_{i=1}^{n} Z_i - \frac{1}{2} \sum_{i=1}^{n} \log Z_i. \tag{S3.1}$$

The following equalities will be used in the integral of the posterior density with respect to $\mu_{\tau,Y}$, $\boldsymbol{\mu_X}$, and $\boldsymbol{\eta}$. This first equality is

$$\begin{aligned}
&- \frac{1}{2\sigma\gamma^2} \left[\left\{ \mathbb{W} - \mu_{\tau,Y} \mathbf{1}_n - \left(\mathbb{X} - \mathbf{1}_n \boldsymbol{\mu}_{\boldsymbol{X}}^T \right) \boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A}) \boldsymbol{\eta} \right\}^T \boldsymbol{D}^{-1} \right. \\
&\qquad \left. \left\{ \mathbb{W} - \mu_{\tau,Y} \mathbf{1}_n - \left(\mathbb{X} - \mathbf{1}_n \boldsymbol{\mu}_{\boldsymbol{X}}^T \right) \boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A}) \boldsymbol{\eta} \right\} \right] \\
&= - \frac{1}{2\sigma\gamma^2} \left[\left\{ \mathbb{W} - \mu_{\tau,Y} \mathbf{1}_n - \left(\mathbb{X} - \mathbf{1}_n \boldsymbol{\mu}_{\boldsymbol{X}}^T \right) \boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A}) \boldsymbol{\eta} \right\}^T \boldsymbol{D}^{-1/2} \left(\boldsymbol{P}_{\boldsymbol{D}^{-1/2}\mathbf{1}_n} + \boldsymbol{Q}_{\boldsymbol{D}^{-1/2}\mathbf{1}_n} \right) \right. \\
&\qquad \left. \boldsymbol{D}^{-1/2} \left\{ \mathbb{W} - \mu_{\tau,Y} \mathbf{1}_n - \left(\mathbb{X} - \mathbf{1}_n \boldsymbol{\mu}_{\boldsymbol{X}}^T \right) \boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A}) \boldsymbol{\eta} \right\} \right] \\
&= - \frac{1}{2\sigma\gamma^2} \left[\left\{ \mathbb{W} - \mu_{\tau,Y} \mathbf{1}_n - \left(\mathbb{X} - \mathbf{1}_n \boldsymbol{\mu}_{\boldsymbol{X}}^T \right) \boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A}) \boldsymbol{\eta} \right\}^T \boldsymbol{D}^{-1/2} \boldsymbol{P}_{\boldsymbol{D}^{-1/2}\mathbf{1}_n} \right. \\
&\qquad \left. \boldsymbol{D}^{-1/2} \left\{ \mathbb{W} - \mu_{\tau,Y} \mathbf{1}_n - \left(\mathbb{X} - \mathbf{1}_n \boldsymbol{\mu}_{\boldsymbol{X}}^T \right) \boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A}) \boldsymbol{\eta} \right\} \right] \\
&\quad - \frac{1}{2\sigma\gamma^2} \left[\left\{ \mathbb{W} - \mu_{\tau,Y} \mathbf{1}_n - \left(\mathbb{X} - \mathbf{1}_n \boldsymbol{\mu}_{\boldsymbol{X}}^T \right) \boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A}) \boldsymbol{\eta} \right\}^T \boldsymbol{D}^{-1/2} \boldsymbol{Q}_{\boldsymbol{D}^{-1/2}\mathbf{1}_n} \right. \\
&\qquad \left. \boldsymbol{D}^{-1/2} \left\{ \mathbb{W} - \mu_{\tau,Y} \mathbf{1}_n - \left(\mathbb{X} - \mathbf{1}_n \boldsymbol{\mu}_{\boldsymbol{X}}^T \right) \boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A}) \boldsymbol{\eta} \right\} \right] \\
&= - \frac{1}{2\sigma\gamma^2} \left[\left(\sum_{i=1}^{n} \frac{1}{Z_i} \right) \left\{ \overline{W}_Z - \mu_{\tau,Y} - \boldsymbol{\eta}^T \boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})^T \left(\overline{\boldsymbol{X}}_Z - \boldsymbol{\mu}_{\boldsymbol{X}} \right) \right\}^T \right. \\
&\qquad \left. \left\{ \overline{W}_Z - \mu_{\tau,Y} - \boldsymbol{\eta}^T \boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})^T \left(\overline{\boldsymbol{X}}_Z - \boldsymbol{\mu}_{\boldsymbol{X}} \right) \right\} \right] \\
&\quad - \frac{1}{2\sigma\gamma^2} \left(\mathbb{W}_{zc} - \mathbb{X}_{zc} \boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A}) \boldsymbol{\eta} \right)^T \boldsymbol{D}^{-1} \left(\mathbb{W}_{zc} - \mathbb{X}_{zc} \boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A}) \boldsymbol{\eta} \right),
\end{aligned} \tag{S3.2}$$

where $\boldsymbol{P}_{\boldsymbol{D}^{-1/2}\mathbf{1}_n}$ is a projection matrix onto span$(\boldsymbol{D}^{-1/2}\mathbf{1}_n)$, $\overline{W}_Z = \frac{1}{\sum_{i=1}^{n} 1/Z_i} \sum_{i=1}^{n} \frac{W_i}{Z_i}$, $\overline{\boldsymbol{X}}_Z = \frac{1}{\sum_{i=1}^{n} 1/Z_i} \sum_{i=1}^{n} \frac{1}{Z_i} \boldsymbol{X}_i \in \mathbb{R}^p$, $\mathbb{W}_{zc} = \mathbb{W} - \overline{W}_Z \mathbf{1}_n$ and $\mathbb{X}_{zc} = \mathbb{X} - \mathbf{1}_n \overline{\boldsymbol{X}}_Z^T$.

The second inequality is

$$
\begin{aligned}
&-\frac{1}{2}\operatorname{trace}\left\{\left(\mathbb{X}-\mathbf{1}_n\boldsymbol{\mu}_{\boldsymbol{X}}^T\right)\left(\boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})\boldsymbol{\Omega}_1^{-1}\boldsymbol{\Gamma}_{1\tau}^T+\boldsymbol{\Gamma}_{2\tau}(\boldsymbol{A})\boldsymbol{\Omega}_2^{-1}\boldsymbol{\Gamma}_{2\tau}^T\right)\left(\mathbb{X}-\mathbf{1}_n\boldsymbol{\mu}_{\boldsymbol{X}}^T\right)^T\right\}\\
&=-\frac{1}{2}\operatorname{trace}\left\{n\left(\boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})\boldsymbol{\Omega}_1^{-1}\boldsymbol{\Gamma}_{1\tau}^T+\boldsymbol{\Gamma}_{2\tau}(\boldsymbol{A})\boldsymbol{\Omega}_2^{-1}\boldsymbol{\Gamma}_{2\tau}^T\right)\left(\overline{\boldsymbol{X}}-\boldsymbol{\mu}_{\boldsymbol{X}}\right)\left(\overline{\boldsymbol{X}}-\boldsymbol{\mu}_{\boldsymbol{X}}\right)^T\right\}\\
&\quad-\frac{1}{2}\operatorname{trace}\left(\boldsymbol{\Omega}_1^{-1}\boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})^T\mathbb{X}_c^T\mathbb{X}_c\boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})\right)\\
&\quad-\frac{1}{2}\operatorname{trace}\left(\boldsymbol{\Omega}_2^{-1}\boldsymbol{\Gamma}_{2\tau}(\boldsymbol{A})^T\mathbb{X}_c^T\mathbb{X}_c\boldsymbol{\Gamma}_{2\tau}(\boldsymbol{A})\right),
\end{aligned}
\tag{S3.3}
$$

where $\mathbb{X}_c=\mathbb{X}-\mathbf{1}_n\overline{\boldsymbol{X}}^T$. In addition, we have

$$
\begin{aligned}
&-\frac{1}{2\sigma\gamma^2}\Big\{\left(\mathbb{W}_{zc}-\mathbb{X}_{zc}\boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})\boldsymbol{\eta}\right)^T\boldsymbol{D}^{-1}\left(\mathbb{W}_{zc}-\mathbb{X}_{zc}\boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})\boldsymbol{\eta}\right)\\
&\qquad\qquad+\left(\boldsymbol{\eta}-\boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})^T\boldsymbol{e}\right)^T\boldsymbol{M}\left(\boldsymbol{\eta}-\boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})^T\boldsymbol{e}\right)\Big\}\\
&=-\frac{1}{2}\Big\{\boldsymbol{\eta}^T\left(\frac{1}{\sigma\gamma^2}\boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})^T\mathbb{X}_{zc}\boldsymbol{D}^{-1}\mathbb{X}_{zc}\boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})+\frac{1}{\sigma\gamma^2}\boldsymbol{M}\right)\boldsymbol{\eta}\\
&\qquad\qquad-2\boldsymbol{\eta}^T\left(\frac{1}{\sigma\gamma^2}\boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})^T\mathbb{X}_{zc}^T\boldsymbol{D}^{-1}\mathbb{W}_{zc}+\frac{1}{\sigma\gamma^2}\boldsymbol{M}\boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})^T\boldsymbol{e}\right)\Big\}\\
&\qquad\qquad-\frac{1}{2}\left(\frac{1}{\sigma\gamma^2}\mathbb{W}_{zc}^T\boldsymbol{D}^{-1}\mathbb{W}_{zc}+\frac{1}{\sigma\gamma^2}\boldsymbol{e}^T\boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})\boldsymbol{M}\boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})^T\boldsymbol{e}\right)\\
&=-\frac{1}{2}\left(\boldsymbol{\eta}-\check{\Delta}_{\boldsymbol{\eta}}^{-1}\check{\mu}_{\boldsymbol{\eta}}\right)^T\check{\Delta}_{\boldsymbol{\eta}}\left(\boldsymbol{\eta}-\check{\Delta}_{\boldsymbol{\eta}}^{-1}\check{\mu}_{\boldsymbol{\eta}}\right)\\
&\qquad\qquad-\frac{1}{2}\left(\frac{1}{\sigma\gamma^2}\mathbb{W}_{zc}^T\boldsymbol{D}^{-1}\mathbb{W}_{zc}+\frac{1}{\sigma\gamma^2}\boldsymbol{e}^T\boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})\boldsymbol{M}\boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})^T\boldsymbol{e}-\check{\mu}_{\boldsymbol{\eta}}^T\check{\Delta}_{\boldsymbol{\eta}}^{-1}\check{\mu}_{\boldsymbol{\eta}}\right),
\end{aligned}
\tag{S3.4}
$$

where $\check{\Delta}_{\boldsymbol{\eta}}=\frac{1}{\sigma\gamma^2}\left(\boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})^T\mathbb{X}_{zc}^T\boldsymbol{D}^{-1}\mathbb{X}_{zc}\boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})+\boldsymbol{M}\right)$ and $\check{\mu}_{\boldsymbol{\eta}}=\frac{1}{\sigma\gamma^2}\left(\boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})^T\mathbb{X}_{zc}^T\boldsymbol{D}^{-1}\mathbb{W}_{zc}+\boldsymbol{M}\boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})^T\boldsymbol{e}\right)$.

Using equations S3.2, S3.3 and S3.4, we first integrate the posterior with respect to $\mu_{\tau,Y}$, $\boldsymbol{\mu_X}$ and $\boldsymbol{\eta}$,

$$\int \exp\left[-\frac{1}{2\sigma\gamma^2}\left(\sum_{i=1}^{n}\frac{1}{Z_i}\right)\left\{\mu_{\tau,Y}-\overline{W}_Z-\left(\boldsymbol{\mu_X}-\overline{\boldsymbol{X}}\right)^T\boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})\boldsymbol{\eta}\right\}^T \left\{\mu_{\tau,Y}-\overline{W}_Z-\left(\boldsymbol{\mu_X}-\overline{\boldsymbol{X}}\right)^T\boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})\boldsymbol{\eta}\right\}\right]d\mu_{\tau,Y}$$
$$=\left(2\pi\gamma^2\right)^{1/2}\sigma^{1/2}\left(\sum_{i=1}^{n}\frac{1}{Z_i}\right)^{-1/2},$$
$$\int \exp\left[-\frac{1}{2}n\left(\boldsymbol{\mu_X}-\overline{X}\right)^T\left(\boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})\boldsymbol{\Omega}_1^{-1}\boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})^T+\boldsymbol{\Gamma}_{2\tau}(\boldsymbol{A})\boldsymbol{\Omega}_2^{-1}\boldsymbol{\Gamma}_{2\tau}(\boldsymbol{A})^T\right) \left(\overline{\boldsymbol{X}}-\boldsymbol{\mu_X}\right)\right]d\boldsymbol{\mu_X}$$
$$=(2\pi)^{p/2}|\boldsymbol{\Omega}_1|^{1/2}|\boldsymbol{\Omega}_2|^{1/2}n^{-p/2},$$
$$\int \exp\left[-\frac{1}{2}\left(\boldsymbol{\eta}-\check{\Delta}_{\boldsymbol{\eta}}^{-1}\check{\mu}_{\boldsymbol{\eta}}\right)^T\check{\Delta}_{\boldsymbol{\eta}}\left(\boldsymbol{\eta}-\check{\Delta}_{\boldsymbol{\eta}}^{-1}\check{\mu}_{\boldsymbol{\eta}}\right)\right]d\boldsymbol{\eta}$$
$$=|2\pi\check{\Delta}_{\boldsymbol{\eta}}^{-1}|^{1/2}=(2\pi)^{u_\tau/2}|\check{\Delta}_{\boldsymbol{\eta}}|^{-1/2}$$
$$=(2\pi)^{u_\tau/2}\sigma^{u_\tau/2}\left(\gamma^2\right)^{u_\tau/2}|\boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})^T\mathbb{X}_{zc}^T\boldsymbol{D}^{-1}\mathbb{X}_{zc}\boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})+\boldsymbol{M}|^{-1/2}$$
$$\leq(2\pi)^{u_\tau/2}\left(\gamma^2\right)^{u_\tau/2}\sigma^{u_\tau/2}|\boldsymbol{M}|^{-1/2}.$$

Thus we have

$$\log\left(\iiint\pi\left(\mu_{\tau,Y},\boldsymbol{\mu_X},\boldsymbol{\eta},\sigma,\boldsymbol{\Omega}_1,\boldsymbol{\Omega}_2,\boldsymbol{A},\mathbb{Z}\right)d\mu_{\tau,Y}d\boldsymbol{\mu_X}d\boldsymbol{\eta}\right)$$
$$\leq \text{const.}-\frac{1}{2}\sum_{i=1}^{n}\log Z_i-\frac{1}{2}\log\left(\sum_{i=1}^{n}\frac{1}{Z_i}\right) -\frac{1}{\sigma}\left(\sum_{i=1}^{n}Z_i\right)-\left(\frac{3n-1}{2}+a+1\right)\log\sigma$$

$$
\begin{aligned}
&- \frac{1}{\sigma}\left\{b + \frac{1}{2\gamma^2}\left(\mathbb{W}_{zc}^T \boldsymbol{D}^{-1}\mathbb{W}_{zc} + \boldsymbol{e}^T\boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})\boldsymbol{M}\boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})^T\boldsymbol{e} - \widetilde{\mu}_{\boldsymbol{\eta}}^T\widetilde{\Delta}_{\boldsymbol{\eta}}^{-1}\widetilde{\mu}_{\boldsymbol{\eta}}\right)\right\} \\
&- \frac{\nu_1 + u_\tau}{2}\log|\boldsymbol{\Omega}_1| - \frac{1}{2}\operatorname{trace}\left[\boldsymbol{\Omega}_1^{-1}\left(\boldsymbol{\Psi}_1 + \boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})^T\mathbb{X}_c^T\mathbb{X}_c\boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})\right)\right] \\
&- \frac{\nu_2 + p - u_\tau}{2}\log|\boldsymbol{\Omega}_2| - \frac{1}{2}\operatorname{trace}\left[\boldsymbol{\Omega}_2^{-1}\left(\boldsymbol{\Psi}_2 + \boldsymbol{\Gamma}_{2\tau}(\boldsymbol{A})^T\mathbb{X}_c^T\mathbb{X}_c\boldsymbol{\Gamma}_{2\tau}(\boldsymbol{A})\right)\right] \\
&- \frac{1}{2}\operatorname{trace}\left[\boldsymbol{K}^{-1}\left(\boldsymbol{A} - \boldsymbol{A}_1\right)\boldsymbol{L}^{-1}\left(\boldsymbol{A} - \boldsymbol{A}_1\right)\right]
\end{aligned}
$$

where $\widetilde{\Delta}_{\boldsymbol{\eta}} = \boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})^T\mathbb{X}_{zc}^T\boldsymbol{D}^{-1}\mathbb{X}_{zc}\boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A}) + \boldsymbol{M}$ and $\widetilde{\mu}_{\boldsymbol{\eta}} = \boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})^T\mathbb{X}_{zc}^T\boldsymbol{D}^{-1}\mathbb{W}_{zc} + \boldsymbol{M}\boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})^T\boldsymbol{e}$.

To integrate with respect to Z_i, we have

$$
\begin{aligned}
&\left(\sum_{i=1}^n \frac{1}{Z_i}\right)^{-1/2}\left(\prod_{i=1}^n Z_i\right)^{-1/2}\exp\left(-\frac{1}{\sigma}\sum_{i=1}^n Z_i\right) \\
&= \left(\frac{n}{\sum_{i=1}^n \frac{1}{Z_i}}\right)^{1/2}\left(\prod_{i=1}^n Z_i\right)^{-1/2}\exp\left(-\frac{1}{\sigma}\sum_{i=1}^n Z_i\right) n^{-1/2} \\
&\le \left(\prod_{i=1}^n Z_i\right)^{1/(2n)}\left(\prod_{i=1}^n Z_i\right)^{-1/2}\exp\left(-\frac{1}{\sigma}\sum_{i=1}^n Z_i\right) n^{-1/2} \\
&= \left(\prod_{i=1}^n Z_i\right)^{(1-n)/(2n)}\exp\left(-\frac{1}{\sigma}\sum_{i=1}^n Z_i\right) n^{-1/2}.
\end{aligned}
$$

The inequality holds by the relationship between geometric mean and harmonic mean. For each i,

$$
\int Z_i^{\frac{n+1}{2n}-1}\exp\left(-\frac{1}{\sigma}Z_i\right) dZ_i = \int \sigma^{\frac{n+1}{2n}}\ t^{\frac{n+1}{2n}-1}\ \exp(-t)dt = \sigma^{\frac{n+1}{2n}}\ \Gamma\left(\frac{n+1}{2n}\right).
$$

Therefore, we have

$$\begin{aligned}
&\log\left(\iiiint \pi\left(\mu_{\tau,Y},\boldsymbol{\mu}_{\boldsymbol{X}},\boldsymbol{\eta},\sigma,\boldsymbol{\Omega}_1,\boldsymbol{\Omega}_2,\boldsymbol{A},\mathbb{Z}\right)d\mu_{\tau,Y}d\boldsymbol{\mu}_{\boldsymbol{X}}d\boldsymbol{\eta}d\mathbb{Z}\right)\\
&\leq \text{const.} - (n+a)\log\sigma\\
&\quad -\frac{1}{\sigma}\left\{b+\frac{1}{2\tau^2}\left(\mathbb{W}_{zc}^T\boldsymbol{D}^{-1}\mathbb{W}_{zc}+\boldsymbol{e}^T\boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})\boldsymbol{M}\boldsymbol{\Gamma}_{1\tau}^T\boldsymbol{e}-\widetilde{\mu}_{\boldsymbol{\eta}}^T\widetilde{\Delta}_{\boldsymbol{\eta}}^{-1}\widetilde{\mu}_{\boldsymbol{\eta}}\right)\right\}\\
&\quad -\frac{\nu_1+u_\tau}{2}\log|\boldsymbol{\Omega}_1|-\frac{1}{2}\,\text{trace}\left[\boldsymbol{\Omega}_1^{-1}\left(\boldsymbol{\Psi}_1+\boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})^T\mathbb{X}_c^T\mathbb{X}_c\boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})\right)\right]\\
&\quad -\frac{\nu_2+p-u_\tau}{2}\log|\boldsymbol{\Omega}_2|-\frac{1}{2}\,\text{trace}\left[\boldsymbol{\Omega}_2^{-1}\left(\boldsymbol{\Psi}_2+\boldsymbol{\Gamma}_{2\tau}(\boldsymbol{A})^T\mathbb{X}_c^T\mathbb{X}_c\boldsymbol{\Gamma}_{2\tau}(\boldsymbol{A})\right)\right]\\
&\quad -\frac{1}{2}\,\text{trace}\left[\boldsymbol{K}^{-1}\left(\boldsymbol{A}-\boldsymbol{A}_1\right)\boldsymbol{L}^{-1}\left(\boldsymbol{A}-\boldsymbol{A}_1\right)\right]
\end{aligned}$$

Let $H=\mathbb{W}_{zc}^T\boldsymbol{D}^{-1}\mathbb{W}_{zc}+\boldsymbol{e}^T\boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})\boldsymbol{M}\boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})^T\boldsymbol{e}-\widetilde{\mu}_{\boldsymbol{\eta}}^T\widetilde{\Delta}_{\boldsymbol{\eta}}^{-1}\widetilde{\mu}_{\boldsymbol{\eta}}$. Then the non-negativity of H can be argued as follows (Khare et al., 2017). Define the positive semi-definite matrices

$$\begin{aligned}
\boldsymbol{H}_1 &= \begin{pmatrix} \mathbb{W}_{zc}^T\boldsymbol{D}^{-1}\mathbb{W}_{zc} & \mathbb{W}_{zc}^T\boldsymbol{D}^{-1}\mathbb{X}_{zc}\boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A}) \\ \boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})^T\mathbb{X}_{zc}^T\boldsymbol{D}^{-1}\mathbb{W}_{zc} & \boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})^T\mathbb{X}_{zc}^T\boldsymbol{D}^{-1}\mathbb{X}_{zc}\boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A}) \end{pmatrix}\\
&= \begin{pmatrix} \mathbb{W}_{zc}^T\boldsymbol{D}^{-1/2} \\ \boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})^T\mathbb{X}_{zc}^T\boldsymbol{D}^{-1/2} \end{pmatrix}\begin{pmatrix} \mathbb{W}_{zc}^T\boldsymbol{D}^{-1/2} \\ \boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})^T\mathbb{X}_{zc}^T\boldsymbol{D}^{-1/2} \end{pmatrix}^T \quad \text{and}\\
\boldsymbol{H}_2 &= \begin{pmatrix} \boldsymbol{e}^T\boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})\boldsymbol{M}\boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})^T\boldsymbol{e} & \boldsymbol{e}^T\boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})\boldsymbol{M} \\ \boldsymbol{M}\boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})^T\boldsymbol{e} & \boldsymbol{M} \end{pmatrix}
\end{aligned}$$

$$= \begin{pmatrix} \boldsymbol{e}^T \boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A}) \boldsymbol{M}^{1/2} \boldsymbol{e} \\ \boldsymbol{M}^{1/2} \end{pmatrix} \begin{pmatrix} \boldsymbol{e}^T \boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A}) \boldsymbol{M}^{1/2} \boldsymbol{e} \\ \boldsymbol{M}^{1/2} \end{pmatrix}^T$$

so that

$$\begin{aligned} \boldsymbol{H}_1 + \boldsymbol{H}_2 &= \begin{pmatrix} \mathbb{W}_{zc}^T \boldsymbol{D}^{-1} \mathbb{W}_{zc} + \boldsymbol{e}^T \boldsymbol{\Gamma}_{1\tau} \boldsymbol{M} \boldsymbol{\Gamma}_{1\tau}^T \boldsymbol{e} & \mathbb{W}_{zc}^T \boldsymbol{D}^{-1} \mathbb{X}_{zc} \boldsymbol{\Gamma}_{1\tau} + \boldsymbol{e}^T \boldsymbol{\Gamma}_{1\tau} \boldsymbol{M} \\ \boldsymbol{\Gamma}_{1\tau}^T \mathbb{X}_{zc}^T \boldsymbol{D}^{-1} \mathbb{W}_{zc} + \boldsymbol{M} \boldsymbol{\Gamma}_{1\tau}^T \boldsymbol{e} & \boldsymbol{\Gamma}_{1\tau}^T \mathbb{X}_{zc}^T \boldsymbol{D}^{-1} \mathbb{X}_{zc} \boldsymbol{\Gamma}_{1\tau} + \boldsymbol{M} \end{pmatrix} \\ &= \begin{pmatrix} \mathbb{W}_{zc}^T \boldsymbol{D}^{-1} \mathbb{W}_{zc} + \boldsymbol{e}^T \boldsymbol{\Gamma}_{1\tau} \boldsymbol{M} \boldsymbol{\Gamma}_{1\tau}^T \boldsymbol{e} & \widetilde{\mu}_{\boldsymbol{\eta}}^T \\ \widetilde{\mu}_{\boldsymbol{\eta}} & \widetilde{\Delta}_{\boldsymbol{\eta}}^{-1} \end{pmatrix} \end{aligned}$$

is also positive semi-definite. This implies that the schur complement of $\widetilde{\Delta}_{\boldsymbol{\eta}}^{-1} = \boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})^T \mathbb{X}_{zc}^T \boldsymbol{D}^{-1} \mathbb{X}_{zc} \boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A}) + \boldsymbol{M}$ is non-negative. Thus H is non-negative. Consequently, $b + H/(2\gamma^2)$ is positive, and hence

$$\begin{aligned} &\int \sigma^{-n-a} \exp\left[-\frac{1}{\sigma} \left\{ b + \frac{1}{2\gamma^2} \left(\mathbb{W}_{zc}^T \boldsymbol{D}^{-1} \mathbb{W}_{zc} \right.\right.\right. \\ &\qquad\qquad \left.\left.\left. + \boldsymbol{e}^T \boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A}) \boldsymbol{M} \boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})^T \boldsymbol{e} - \widetilde{\mu}_{\boldsymbol{\eta}}^T \widetilde{\Delta}_{\boldsymbol{\eta}}^{-1} \widetilde{\mu}_{\boldsymbol{\eta}} \right) \right\} \right] d\sigma \\ &= \Gamma\left(n + a - 1\right) \left\{ b + \frac{1}{2\gamma^2} \left(\mathbb{W}_{zc}^T \boldsymbol{D}^{-1} \mathbb{W}_{zc} \right.\right. \\ &\qquad\qquad \left.\left. + \boldsymbol{e}^T \boldsymbol{\Gamma}_{1\tau} \boldsymbol{M} \boldsymbol{\Gamma}_{1\tau}^T \boldsymbol{e} - \widetilde{\mu}_{\boldsymbol{\eta}}^T \widetilde{\Delta}_{\boldsymbol{\eta}}^{-1} \widetilde{\mu}_{\boldsymbol{\eta}} \right) \right\}^{-(n+a-1)} \\ &\le \Gamma\left(n + a - 1\right)\, b^{-(n+a-1)}. \end{aligned}$$

Hence, we have

$$\begin{aligned}
&\log\left(\int\cdots\int\pi\left(\mu_{\tau,Y},\boldsymbol{\mu}_{\boldsymbol{X}},\boldsymbol{\eta},\sigma,\boldsymbol{\Omega}_1,\boldsymbol{\Omega}_2,\boldsymbol{A},\mathbb{Z}\right)d\mu_{\tau,Y}d\boldsymbol{\mu}_{\boldsymbol{X}}d\boldsymbol{\eta}d\mathbb{Z}d\sigma\right)\\
&\leq \text{const.}-\frac{\nu_1+u_\tau}{2}\log|\boldsymbol{\Omega}_1|-\frac{1}{2}\text{trace}\left[\boldsymbol{\Omega}_1^{-1}\left(\boldsymbol{\Psi}_1+\boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})^T\mathbb{X}_c^T\mathbb{X}_c\boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})\right)\right]\\
&\quad-\frac{\nu_2+p-u_\tau}{2}\log|\boldsymbol{\Omega}_2|-\frac{1}{2}\text{trace}\left[\boldsymbol{\Omega}_2^{-1}\left(\boldsymbol{\Psi}_2+\boldsymbol{\Gamma}_{2\tau}(\boldsymbol{A})^T\mathbb{X}_c^T\mathbb{X}_c\boldsymbol{\Gamma}_{2\tau}(\boldsymbol{A})\right)\right]\\
&\quad-\frac{1}{2}\text{trace}\left[\boldsymbol{K}^{-1}\left(\boldsymbol{A}-\boldsymbol{A}_1\right)\boldsymbol{L}^{-1}\left(\boldsymbol{A}-\boldsymbol{A}_1\right)\right].
\end{aligned}$$

The above upper bound, being a constant multiple of the product of inverse Wishart and matrix normal densities, is clearly integrable. This completes the proof. □

Proof of Theorem 2

The data augmentation Metropolis within Gibbs sampler given in Algorithm 1 and its generalization to cases $u_\tau = 0$ and $u_\tau = p$ block-wise updates the $7 + u_\tau$ parameter blocks $\mu_{\tau,Y}, \boldsymbol{\mu}_{\boldsymbol{X}}, \boldsymbol{\eta}$, $\sigma, \boldsymbol{\Omega}_1, \boldsymbol{\Omega}_2$, $\mathbb{Z}, \boldsymbol{a}_1, \cdots, \boldsymbol{a}_{u_\tau}$, where $\boldsymbol{a}_i$ is the ith column of $\boldsymbol{A}$.

We first show ϕ-irreducibility and aperiodicity. Note that the Metropolis step, if present (i.e., if $1 \leq u_\tau \leq p$), has a strictly positive acceptance probability everywhere, since both the proposal and the target conditional posterior densities of the columns $\{\boldsymbol{a}_i\}$ are positive everywhere on the support of $\{\boldsymbol{a}_i\}$. Furthermore, the Gibbs steps, when viewed as Metropolis steps with identical proposal and target densities, also have positive (viz., one) acceptance probabilities. If λ is the Lebesgue measure, B is a Borel set on the parameter space, and $T(\cdot,\cdot)$ is the Markov transition function of the chain for Algorithm 1, then $T(x,B)>0$ whenever $\lambda(B)>0$ for any $x\in\mathbb{R}^1\times\mathbb{R}^p\times\mathbb{R}^{u_\tau}\times\mathbb{R}^1\times\mathbb{S}_+^{u_\tau\times u_\tau}\times\mathbb{S}_+^{(p-u_\tau)\times(p-u_\tau)}\times\mathbb{R}^n\times\mathbb{R}^{(p-u_\tau)\times u_\tau}$ (see, e.g., Geyer, 1998, Section 3.1.9). This ensures that the sampler is λ-irreducible. This also implies that every measurable set with positive Lebesgue measure can be accessed through the Markov chain in a single step from any point, ensuring aperiodicity (see, e.g., Dutta, 2012).

We now prove Harris recurrence. Note that when $u_\tau\in\{0,p\}$, the Metropolis step does not arise, and the MCMC algorithm turns into a Gibbs sampler. This, together with the ϕ-irreducibility of the sampler ensures Harris recurrence when $u_\tau=0$ or p (Roberts and Rosenthal, 2006, Corollary 13). We therefore focus on the cases when $u_\tau\in\{1,\ldots,p-1\}$, i.e., when the algorithm does contain a Metropolis step. Then, we need to show that the unnormalized posterior density of $\mu_{\tau,Y},\boldsymbol{\mu_X},\boldsymbol{\eta},\sigma,\boldsymbol{\Omega}_1,\boldsymbol{\Omega}_2,\boldsymbol{A},\mathbb{Z}$ given $\mathbb{X},\mathbb{Y}$ is Lebesgue integrable with respect to any $1\le k\le 7+u_\tau$ elements of the set $\{\mu_{\tau,Y},\boldsymbol{\mu_X},\ \boldsymbol{\eta},\sigma,\boldsymbol{\Omega}_1,\boldsymbol{\Omega}_2,\mathbb{Z},\boldsymbol{a}_1,\cdots,\boldsymbol{a}_{u_\tau}\}$ (Roberts and Rosenthal, 2006, Corollary 19).

The unnormalized joint posterior density is

$$
\begin{aligned}
& f\left(\mu_{\tau,Y}, \boldsymbol{\eta}, \boldsymbol{\mu}_{\boldsymbol{X}}, \mathbb{Z}, \sigma, \boldsymbol{\Omega}_1, \boldsymbol{\Omega}_2, \boldsymbol{A} | \mathbb{X}, \mathbb{Y}\right) \\
= & \exp\left[-\frac{1}{2\sigma\gamma^2}\left\{\mathbb{W} - \mu_{\tau,Y}\mathbf{1}_n - \left(\mathbb{X} - \mathbf{1}_n\boldsymbol{\mu}_{\boldsymbol{X}}^T\right)\boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})\boldsymbol{\eta}\right\}^T \boldsymbol{D}^{-1}\right. \\
& \qquad \left.\left\{\mathbb{W} - \mu_{\tau,Y}\mathbf{1}_n - \left(\mathbb{X} - \mathbf{1}_n\boldsymbol{\mu}_{\boldsymbol{X}}^T\right)\boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})\boldsymbol{\eta}\right\}\right] \\
& \sigma^{-\left(n + \frac{n}{2} + \frac{u_\tau}{2} + a + 1\right)} \exp\left(-\frac{b}{\sigma}\right) \\
& \exp\left[-\frac{1}{2}\operatorname{trace}\left(\mathbb{X} - \mathbf{1}_n\boldsymbol{\mu}_{\boldsymbol{X}}^T\right)\left(\boldsymbol{\Gamma}_{1\tau}\boldsymbol{\Omega}_1^{-1}\boldsymbol{\Gamma}_{1\tau}^T + \boldsymbol{\Gamma}_{2\tau}\boldsymbol{\Omega}_2^{-1}\boldsymbol{\Gamma}^T\right)\left(\mathbb{X} - \mathbf{1}_n\boldsymbol{\mu}_{\boldsymbol{X}}^T\right)^T\right] \\
& \exp\left[-\frac{1}{2\sigma\gamma^2}\left(\boldsymbol{\eta} - \boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})^T\boldsymbol{e}\right)^T \boldsymbol{M}\left(\boldsymbol{\eta} - \boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})^T\boldsymbol{e}\right)\right] \prod_{i=1}^n Z_i^{-\frac{1}{2}} \exp\left(-\frac{Z_i}{\sigma}\right) \\
& |\boldsymbol{\Omega}_1|^{-\frac{n+\nu_1+u_\tau+1}{2}} \exp\left[-\frac{1}{2}\operatorname{trace}\left(\boldsymbol{\Omega}_1^{-1}\boldsymbol{\Psi}_1\right)\right] \\
& |\boldsymbol{\Omega}_2|^{-\frac{n+\nu_2+p-u_\tau+1}{2}} \exp\left[-\frac{1}{2}\operatorname{trace}\left(\boldsymbol{\Omega}_2^{-1}\boldsymbol{\Psi}_2\right)\right] \\
& \exp\left[-\frac{1}{2}\operatorname{trace}\left\{\boldsymbol{K}^{-1}\left(\boldsymbol{A} - \boldsymbol{A}_0\right)\boldsymbol{L}^{-1}\left(\boldsymbol{A} - \boldsymbol{A}_0\right)^T\right\}\right]. \qquad \text{(S3.5)}
\end{aligned}
$$

Our goal is to prove the integrability of $f\left(\mu_{\tau,Y}, \boldsymbol{\eta}, \boldsymbol{\mu}_{\boldsymbol{X}}, \mathbb{Z}, \sigma, \boldsymbol{\Omega}_1, \boldsymbol{\Omega}_2, \boldsymbol{A} | \mathbb{X}, \mathbb{Y}\right)$ with respect to any $1 \leq k \leq 7 + u_\tau$ parameter blocks. Note that

$$
\begin{aligned}
& -\frac{1}{2\sigma\gamma^2}\left[\left\{\mathbb{W} - \mu_{\tau,Y}\mathbf{1}_n - \left(\mathbb{X} - \mathbf{1}_n\boldsymbol{\mu}_{\boldsymbol{X}}^T\right)\boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})\boldsymbol{\eta}\right\}^T \boldsymbol{D}^{-1}\right. \\
& \qquad \left.\left\{\mathbb{W} - \mu_{\tau,Y}\mathbf{1}_n - \left(\mathbb{X} - \mathbf{1}_n\boldsymbol{\mu}_{\boldsymbol{X}}^T\right)\boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})\boldsymbol{\eta}\right\}\right] \\
= & -\frac{1}{2\sigma\gamma^2}\left(\sum_{i=1}^n \frac{1}{Z_i}\right)\left\{\overline{W} - \mu_{\tau,Y} - \boldsymbol{\eta}^T\boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})^T\left(\overline{\boldsymbol{X}}_Z - \boldsymbol{\mu}_{\boldsymbol{X}}\right)\right\}^2 \\
& \qquad - \frac{1}{2\sigma\gamma^2}\left(\mathbb{W}_{zc} - \mathbb{X}_{zc}\boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})\boldsymbol{\eta}\right)^T \boldsymbol{D}^{-1}\left(\mathbb{W}_{zc} - \mathbb{X}_{zc}\boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})\boldsymbol{\eta}\right)
\end{aligned}
$$

$$\leq -\frac{1}{2\sigma\gamma^2}\left(\sum_{i=1}^{n}\frac{1}{Z_i}\right)\left\{\overline{W} - \mu_{\tau,Y} - \boldsymbol{\eta}^T\boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})^T\left(\overline{\boldsymbol{X}}_Z - \boldsymbol{\mu}_{\boldsymbol{X}}\right)\right\}^2, \qquad \text{(S3.6)}$$

and

$$\begin{aligned}
&-\frac{1}{2}\operatorname{trace}\left\{\left(\mathbb{X} - \mathbf{1}_n\boldsymbol{\mu}_{\boldsymbol{X}}^T\right)\left(\boldsymbol{\Gamma}_{1\tau}\boldsymbol{\Omega}_1^{-1}\boldsymbol{\Gamma}_{1\tau}^T + \boldsymbol{\Gamma}_{2\tau}\boldsymbol{\Omega}_2^{-1}\boldsymbol{\Gamma}_{2\tau}^T\right)\left(\mathbb{X} - \mathbf{1}_n\boldsymbol{\mu}_{\boldsymbol{X}}^T\right)^T\right\} \\
&= -\frac{1}{2}\left\{n\left(\boldsymbol{\mu}_{\boldsymbol{X}} - \overline{\boldsymbol{X}}\right)^T\left(\boldsymbol{\Gamma}_{1\tau}\boldsymbol{\Omega}_1^{-1}\boldsymbol{\Gamma}_{1\tau}^T + \boldsymbol{\Gamma}_{2\tau}\boldsymbol{\Omega}_2^{-1}\boldsymbol{\Gamma}_{2\tau}^T\right)\left(\boldsymbol{\mu}_{\boldsymbol{X}} - \overline{\boldsymbol{X}}\right)\right\} \\
&\qquad - \frac{1}{2}\operatorname{trace}\left(\boldsymbol{\Omega}_1^{-1}\boldsymbol{\Gamma}_{1\tau}^T\mathbb{X}_c^T\mathbb{X}_c\boldsymbol{\Gamma}_{1\tau}\right) - \frac{1}{2}\operatorname{trace}\left(\boldsymbol{\Omega}_2^{-1}\boldsymbol{\Gamma}_{2\tau}^T\mathbb{X}_c^T\mathbb{X}_c\boldsymbol{\Gamma}_{2\tau}\right) \\
&\leq -\frac{1}{2}n\left(\boldsymbol{\mu}_{\boldsymbol{X}} - \overline{\boldsymbol{X}}\right)^T\left(\boldsymbol{\Gamma}_{1\tau}\boldsymbol{\Omega}_1^{-1}\boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})^T + \boldsymbol{\Gamma}_{2\tau}\boldsymbol{\Omega}_2^{-1}\boldsymbol{\Gamma}_{2\tau}(\boldsymbol{A})^T\right)\left(\boldsymbol{\mu}_{\boldsymbol{X}} - \overline{\boldsymbol{X}}\right).
\end{aligned} \qquad \text{(S3.7)}$$

Consequently,

$$f\left(\mu_{\tau,Y}, \boldsymbol{\eta}, \boldsymbol{\mu}_{\boldsymbol{X}}, \mathbb{Z}, \sigma, \boldsymbol{\Omega}_1, \boldsymbol{\Omega}_2, \boldsymbol{A}|\mathbb{X}, \mathbb{Y}\right) \leq f_0\left(\mu_{\tau,Y}, \boldsymbol{\eta}, \boldsymbol{\mu}_{\boldsymbol{X}}, \mathbb{Z}, \sigma, \boldsymbol{\Omega}_1, \boldsymbol{\Omega}_2, \boldsymbol{A}|\mathbb{X}, \mathbb{Y}\right)$$

where

$$\begin{aligned}
&f_0\left(\mu_{\tau,Y}, \boldsymbol{\eta}, \boldsymbol{\mu}_{\boldsymbol{X}}, \mathbb{Z}, \sigma, \boldsymbol{\Omega}_1, \boldsymbol{\Omega}_2, \boldsymbol{A}|\mathbb{X}, \mathbb{Y}\right) \\
&= \exp\left[-\frac{1}{2\sigma\gamma^2}\left(\sum_{i=1}^{n}\frac{1}{Z_i}\right)\left\{\overline{W}_Z - \mu_{\tau,Y} - \boldsymbol{\eta}^T\boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})^T\left(\overline{\boldsymbol{X}}_Z - \boldsymbol{\mu}_{\boldsymbol{X}}\right)\right\}^2\right] \\
&\quad \exp\left[-\frac{1}{2}n\left(\boldsymbol{\mu}_{\boldsymbol{X}} - \overline{\boldsymbol{X}}\right)^T\left(\boldsymbol{\Gamma}_{1\tau}\boldsymbol{\Omega}_1^{-1}\boldsymbol{\Gamma}_{1\tau}^T + \boldsymbol{\Gamma}_{2\tau}\boldsymbol{\Omega}_2^{-1}\boldsymbol{\Gamma}_{2\tau}^T\right)\left(\boldsymbol{\mu}_{\boldsymbol{X}} - \overline{\boldsymbol{X}}\right)\right] \\
&\quad \exp\left[-\frac{1}{2\sigma\gamma^2}\left(\boldsymbol{\eta} - \boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})^T\boldsymbol{e}\right)^T\boldsymbol{M}\left(\boldsymbol{\eta} - \boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})^T\boldsymbol{e}\right)\right]
\end{aligned}$$

$$\prod_{i=1}^{n} Z_i^{-\frac{1}{2}} \exp\left(-\frac{Z_i}{\sigma}\right) \sigma^{-(n+\frac{n}{2}+\frac{u_\mathcal{T}}{2}+a+1)}$$
$$|\boldsymbol{\Omega}_1|^{-\frac{n+\nu_1+u_\mathcal{T}+1}{2}} \exp\left[-\frac{1}{2}\operatorname{trace}\left(\boldsymbol{\Omega}_1^{-1}\boldsymbol{\Psi}_1\right)\right]$$
$$|\boldsymbol{\Omega}_2|^{-\frac{n+\nu_2+p-u_\mathcal{T}+1}{2}} \exp\left[-\frac{1}{2}\operatorname{trace}\left(\boldsymbol{\Omega}_2^{-1}\boldsymbol{\Psi}_2\right)\right]$$
$$\exp\left[-\frac{1}{2}\operatorname{trace}\left\{\boldsymbol{K}^{-1}\left(\boldsymbol{A}-\boldsymbol{A}_0\right)\boldsymbol{L}^{-1}\left(\boldsymbol{A}-\boldsymbol{A}_0\right)^T\right\}\right]\exp\left(-\frac{b}{\sigma}\right),$$

and it will be enough to show the integrability of $f_0\left(\mu_{\tau,Y}, \boldsymbol{\eta}, \boldsymbol{\mu_X}, \mathbb{Z}, \sigma, \boldsymbol{\Omega}_1, \boldsymbol{\Omega}_2, \boldsymbol{A} | \mathbb{X}, \mathbb{Y}\right)$.

Note that

$$\exp\left[-\frac{1}{2\sigma\gamma^2}\left(\sum_{i=1}^{n}\frac{1}{Z_i}\right)\left\{\overline{W}_Z - \mu_{\tau,Y} - \boldsymbol{\eta}^T\boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})^T\left(\overline{\boldsymbol{X}}_Z - \boldsymbol{\mu_X}\right)\right\}^2\right] \leq 1,$$

and

$$\int \exp\left[-\frac{1}{2\sigma\gamma^2}\left(\sum_{i=1}^{n}\frac{1}{Z_i}\right)\left\{\overline{W}_Z - \mu_{\tau,Y} - \boldsymbol{\eta}^T\boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})^T\left(\overline{\boldsymbol{X}}_Z - \boldsymbol{\mu_X}\right)\right\}^2\right] d\mu_{\tau,Y}$$
$$= \left(2\pi\gamma^2\right)^{\frac{1}{2}} \sigma^{\frac{1}{2}} \left(\sum_{i=1}^{n}\frac{1}{Z_i}\right)^{-\frac{1}{2}}$$
$$\leq \left(2\pi\gamma^2 n^{-1}\right)^{\frac{1}{2}} \sigma^{\frac{1}{2}} \left(\prod_{i=1}^{n} Z_i\right)^{\frac{1}{2n}}.$$

The last inequality is obtained from the relationship between harmonic mean and geometric mean. Therefore, for all $\mu_{\tau,Y}$

$$f_0\left(\mu_{\tau,Y}, \boldsymbol{\eta}, \boldsymbol{\mu_X}, \mathbb{Z}, \sigma, \boldsymbol{\Omega}_1, \boldsymbol{\Omega}_2, \boldsymbol{A} | \mathbb{X}, \mathbb{Y}\right) \leq c_1 \ f_1^{(0)}\left(\boldsymbol{\eta}, \boldsymbol{\mu_X}, \mathbb{Z}, \sigma, \boldsymbol{\Omega}_1, \boldsymbol{\Omega}_2, \boldsymbol{A} | \mathbb{X}, \mathbb{Y}\right)$$

and

$$\begin{aligned}
&\int f_0\left(\mu_{\tau,Y}, \boldsymbol{\eta}, \boldsymbol{\mu_X}, \mathbb{Z}, \sigma, \boldsymbol{\Omega}_1, \boldsymbol{\Omega}_2, \boldsymbol{A} | \mathbb{X}, \mathbb{Y}\right) d\mu_{\tau,Y} \\
&\quad \leq c_1 \ f_1^{(1)}\left(\boldsymbol{\eta}, \boldsymbol{\mu_X}, \mathbb{Z}, \sigma, \boldsymbol{\Omega}_1, \boldsymbol{\Omega}_2, \boldsymbol{A} | \mathbb{X}, \mathbb{Y}\right),
\end{aligned}$$

where $c_1 = 1 + (2\pi\gamma^2 n^{-1})^{\frac{1}{2}}$ and for $\xi = 0, 1$,

$$\begin{aligned}
&f_1^{(\xi)}\left(\boldsymbol{\eta}, \boldsymbol{\mu_X}, \mathbb{Z}, \sigma, \boldsymbol{\Omega}_1, \boldsymbol{\Omega}_2, \boldsymbol{A} | \mathbb{X}, \mathbb{Y}\right) \\
&= \exp\left[-\frac{1}{2\sigma\gamma^2} n \left(\boldsymbol{\mu_X} - \overline{\boldsymbol{X}}\right)^T \left(\boldsymbol{\Gamma}_{1\tau}\boldsymbol{\Omega}_1^{-1}\boldsymbol{\Gamma}_{1\tau}^T + \boldsymbol{\Gamma}_{2\tau}\boldsymbol{\Omega}_2^{-1}\boldsymbol{\Gamma}_{2\tau}^T\right)\left(\boldsymbol{\mu_X} - \overline{\boldsymbol{X}}\right)\right] \\
&\quad \exp\left[-\frac{1}{2\sigma\gamma^2}\left(\boldsymbol{\eta} - \boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})^T\boldsymbol{e}\right)^T \boldsymbol{M}\left(\boldsymbol{\eta} - \boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})^T\boldsymbol{e}\right)\right] \\
&\quad \prod_{i=1}^n Z_i^{-\left(\frac{1}{2} - \frac{\xi}{2n}\right)} \exp\left(-\frac{Z_i}{\sigma}\right) \sigma^{-\left(n + \frac{n}{2} + \frac{u_\tau}{2} + a + 1 - \frac{\xi}{2}\right)} \exp\left(-\frac{b}{\sigma}\right) \\
&\quad |\boldsymbol{\Omega}_1|^{-\frac{n+\nu_1+u_\tau+1}{2}} \exp\left[-\frac{1}{2}\operatorname{trace}\left(\boldsymbol{\Omega}_1^{-1}\boldsymbol{\Psi}_1\right)\right] \\
&\quad |\boldsymbol{\Omega}_2|^{-\frac{n+\nu_2+p-u_\tau+1}{2}} \exp\left[-\frac{1}{2}\operatorname{trace}\left(\boldsymbol{\Omega}_2^{-1}\boldsymbol{\Psi}_2\right)\right] \\
&\quad \exp\left[-\frac{1}{2}\operatorname{trace}\left\{\boldsymbol{K}^{-1}\left(\boldsymbol{A} - \boldsymbol{A}_0\right)\boldsymbol{L}^{-1}\left(\boldsymbol{A} - \boldsymbol{A}_0\right)^T\right\}\right] \quad (< \infty).
\end{aligned}$$

It will therefore suffice to show that $f_1^{(\xi)}\left(\boldsymbol{\eta}, \boldsymbol{\mu_X}, \mathbb{Z}, \sigma, \boldsymbol{\Omega}_1, \boldsymbol{\Omega}_2, \boldsymbol{A} | \mathbb{X}, \mathbb{Y}\right)$ is integrable with respect to any $1 \le k \le 6+u_\tau$ members of the set $\{\boldsymbol{\eta}, \boldsymbol{\mu}_{\boldsymbol{X}}, \mathbb{Z}, \sigma,$ $\boldsymbol{\Omega}_1, \boldsymbol{\Omega}_2, \boldsymbol{a}_1, \cdots, \boldsymbol{a}_{u_\tau}\}$.

Observe that

$$\exp\left[-\frac{1}{2\sigma\gamma^2}\left(\boldsymbol{\eta} - \boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})^T\boldsymbol{e}\right)^T \boldsymbol{M}\left(\boldsymbol{\eta} - \boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})^T\boldsymbol{e}\right)\right] \le 1,$$

and

$$\int \exp\left[-\frac{1}{2\sigma\gamma^2}\left(\boldsymbol{\eta} - \boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})^T\boldsymbol{e}\right)^T \boldsymbol{M}\left(\boldsymbol{\eta} - \boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})^T\boldsymbol{e}\right)\right] d\boldsymbol{\eta}$$
$$= \left((2\pi)^{u_\tau}\gamma^{2u_\tau}|\boldsymbol{M}|^{-1}\right)^{\frac{1}{2}}\sigma^{\frac{u_\tau}{2}}.$$

Hence, we have

$$f_1^{(\xi)}\left(\boldsymbol{\eta}, \boldsymbol{\mu}_{\boldsymbol{X}}, \mathbb{Z}, \sigma, \boldsymbol{\Omega}_1, \boldsymbol{\Omega}_2, \boldsymbol{A}|\mathbb{X}, \mathbb{Y}\right) \le c_2\, f_2^{(\xi,0)}\left(\boldsymbol{\mu}_{\boldsymbol{X}}, \mathbb{Z}, \sigma, \boldsymbol{\Omega}_1, \boldsymbol{\Omega}_2, \boldsymbol{A}|\mathbb{X}, \mathbb{Y}\right),$$

and

$$\int f_1^{(\xi)}\left(\boldsymbol{\eta}, \boldsymbol{\mu}_{\boldsymbol{X}}, \mathbb{Z}, \sigma, \boldsymbol{\Omega}_1, \boldsymbol{\Omega}_2, \boldsymbol{A}|\mathbb{X}, \mathbb{Y}\right) d\boldsymbol{\eta} \le c_2\, f_2^{(\xi,1)}\left(\boldsymbol{\mu}_{\boldsymbol{X}}, \mathbb{Z}, \sigma, \boldsymbol{\Omega}_1, \boldsymbol{\Omega}_2, \boldsymbol{A}|\mathbb{X}, \mathbb{Y}\right),$$

where $c_2 = 1 + \left((2\pi)^{u_\tau}\gamma^{2u_\tau}|\boldsymbol{M}|^{-1}\right)^{\frac{1}{2}}$ and for $\zeta = 0, 1$,

$$f_2^{(\xi,\zeta)}\left(\boldsymbol{\mu}_{\boldsymbol{X}}, \mathbb{Z}, \sigma, \boldsymbol{\Omega}_1, \boldsymbol{\Omega}_2, \boldsymbol{A}|\mathbb{X}, \mathbb{Y}\right)$$

$$= \exp\left[-\frac{1}{2}n\left(\boldsymbol{\mu}_{\boldsymbol{X}} - \overline{\boldsymbol{X}}\right)^T\left(\boldsymbol{\Gamma}_{1\tau}\boldsymbol{\Omega}_1^{-1}\boldsymbol{\Gamma}_{1\tau}^T + \boldsymbol{\Gamma}_{2\tau}\boldsymbol{\Omega}_2^{-1}\boldsymbol{\Gamma}_{2\tau}^T\right)\left(\boldsymbol{\mu}_{\boldsymbol{X}} - \overline{\boldsymbol{X}}\right)\right]$$
$$\prod_{i=1}^{n} Z_i^{-(\frac{1}{2}-\frac{\xi}{2n})}\exp\left(-\frac{Z_i}{\sigma}\right)\sigma^{-(n+\frac{n}{2}+\frac{u_\tau}{2}+a+1-(\frac{\xi}{2}+\frac{\zeta u_\tau}{2}))}\exp\left(-\frac{b}{\sigma}\right)$$
$$|\boldsymbol{\Omega}_1|^{-\frac{n+\nu_1+u_\tau+1}{2}}\exp\left[-\frac{1}{2}\operatorname{trace}\left(\boldsymbol{\Omega}_1^{-1}\boldsymbol{\Psi}_1\right)\right]|\boldsymbol{\Omega}_2|^{-\frac{n+\nu_2+p-u_\tau+1}{2}}$$
$$\exp\left[-\frac{1}{2}\operatorname{trace}\left(\boldsymbol{\Omega}_2^{-1}\boldsymbol{\Psi}_2\right)\right]\exp\left[-\frac{1}{2}\operatorname{trace}\left\{\boldsymbol{K}^{-1}\left(\boldsymbol{A}-\boldsymbol{A}_0\right)\boldsymbol{L}^{-1}\left(\boldsymbol{A}-\boldsymbol{A}_0\right)^T\right\}\right]$$
$$(< \infty).$$

Next, we show that $f_2^{(\xi,\zeta)}\left(\boldsymbol{\mu}_{\boldsymbol{X}}, \mathbb{Z}, \sigma, \boldsymbol{\Omega}_1, \boldsymbol{\Omega}_2, \boldsymbol{A}|\mathbb{X}, \mathbb{Y}\right)$ is integrable with respect to any $1 \le k \le 5+u_\tau$ elements of the set $\{\boldsymbol{\mu}_{\boldsymbol{X}}, \mathbb{Z}, \sigma, \boldsymbol{\Omega}_1, \boldsymbol{\Omega}_2, \boldsymbol{a}_1, \cdots, \boldsymbol{a}_{u_\tau}\}$. We have

$$\exp\left[-\frac{1}{2}n\left(\boldsymbol{\mu}_{\boldsymbol{X}} - \overline{\boldsymbol{X}}\right)^T\left(\boldsymbol{\Gamma}_{1\tau}\boldsymbol{\Omega}_1^{-1}\boldsymbol{\Gamma}_{1\tau}^T + \boldsymbol{\Gamma}_{2\tau}\boldsymbol{\Omega}_2^{-1}\boldsymbol{\Gamma}_{2\tau}^T\right)\left(\boldsymbol{\mu}_{\boldsymbol{X}} - \overline{\boldsymbol{X}}\right)\right] \le 1,$$

and

$$\int \exp\left[-\frac{1}{2}n\left(\boldsymbol{\mu}_{\boldsymbol{X}} - \overline{\boldsymbol{X}}\right)^T\left(\boldsymbol{\Gamma}_{1\tau}\boldsymbol{\Omega}_1^{-1}\boldsymbol{\Gamma}_{1\tau}^T + \boldsymbol{\Gamma}_{2\tau}\boldsymbol{\Omega}_2^{-1}\boldsymbol{\Gamma}_{2\tau}^T\right)\left(\boldsymbol{\mu}_{\boldsymbol{X}} - \overline{\boldsymbol{X}}\right)\right] d\boldsymbol{\mu}_{\boldsymbol{X}}$$
$$= |\boldsymbol{\Omega}_1|^{\frac{1}{2}}|\boldsymbol{\Omega}_2|^{\frac{1}{2}}\left(2\pi n^{-1}\right)^{\frac{p}{2}}.$$

Therefore,

$$f_2^{(\xi,\zeta)}\left(\boldsymbol{\mu}_{\boldsymbol{X}}, \mathbb{Z}, \sigma, \boldsymbol{\Omega}_1, \boldsymbol{\Omega}_2, \boldsymbol{A}|\mathbb{X}, \mathbb{Y}\right) \le c_3\ f_3^{(\xi,\zeta,0)}\left(\mathbb{Z}, \sigma, \boldsymbol{\Omega}_1, \boldsymbol{\Omega}_2, \boldsymbol{A}|\mathbb{X}, \mathbb{Y}\right),$$

and

$$\int f_2^{(\xi,\zeta)}\left(\boldsymbol{\mu}_{\boldsymbol{X}}, \mathbb{Z}, \sigma, \boldsymbol{\Omega}_1, \boldsymbol{\Omega}_2, \boldsymbol{A} | \mathbb{X}, \mathbb{Y}\right) d\boldsymbol{\mu}_{\boldsymbol{X}} \leq c_3 \; f_3^{(\xi,\zeta,1)}\left(\mathbb{Z}, \sigma, \boldsymbol{\Omega}_1, \boldsymbol{\Omega}_2, \boldsymbol{A} | \mathbb{X}, \mathbb{Y}\right),$$

where $c_3 = 1 + (2\pi n^{-1})^{\frac{p}{2}}$ and for $\lambda = 0, 1$,

$$\begin{aligned}
& f_3^{(\xi,\zeta,\lambda)}\left(\mathbb{Z}, \sigma, \boldsymbol{\Omega}_1, \boldsymbol{\Omega}_2, \boldsymbol{A} | \mathbb{X}, \mathbb{Y}\right) \\
&= \prod_{i=1}^{n} Z_i^{-\left(\frac{1}{2}-\frac{\xi}{2n}\right)} \exp\left(-\frac{Z_i}{\sigma}\right) \sigma^{-\left(n+\frac{n}{2}+\frac{u_\tau}{2}+a+1-\left(\frac{\xi}{2}+\frac{\zeta u_\tau}{2}\right)\right)} \exp\left(-\frac{b}{\sigma}\right) \\
& \quad |\boldsymbol{\Omega}_1|^{-\frac{n+\nu_1+u_\tau+1-\lambda}{2}} \exp\left[-\frac{1}{2}\operatorname{trace}\left(\boldsymbol{\Omega}_1^{-1}\boldsymbol{\Psi}_1\right)\right] \\
& \quad |\boldsymbol{\Omega}_2|^{-\frac{n+\nu_2+p-u_\tau+1-\lambda}{2}} \exp\left[-\frac{1}{2}\operatorname{trace}\left(\boldsymbol{\Omega}_2^{-1}\boldsymbol{\Psi}_2\right)\right] \\
& \quad \exp\left[-\frac{1}{2}\operatorname{trace}\left\{\boldsymbol{K}^{-1}\left(\boldsymbol{A}-\boldsymbol{A}_0\right)\boldsymbol{L}^{-1}\left(\boldsymbol{A}-\boldsymbol{A}_0\right)^T\right\}\right] \quad (< \infty)
\end{aligned}$$

It is enough to show that $f_3^{(\xi,\zeta,\lambda)}\left(\mathbb{Z}, \sigma, \boldsymbol{\Omega}_1, \boldsymbol{\Omega}_2, \boldsymbol{A} | \mathbb{X}, \mathbb{Y}\right)$ is integrable with respect to any $1 \leq k \leq 4+u_\tau$ elements of the set $\{\mathbb{Z}, \sigma, \boldsymbol{\Omega}_1, \boldsymbol{\Omega}_2, \boldsymbol{a}_1, \cdots, \boldsymbol{a}_{u_\tau}\}$. Now, for all $\mathbb{Z}$,

$$\prod_{i=1}^{n}\left\{Z_i^{-\left(\frac{1}{2}-\frac{\xi}{2n}\right)} \exp\left(-\frac{Z_i}{\sigma}\right)\right\} \leq \prod_{i=1}^{n} Z_i^{-\left(\frac{1}{2}-\frac{\xi}{2n}\right)},$$

and since $-\left(\frac{1}{2}-\frac{\xi}{2n}\right) = \left(\frac{1}{2}+\frac{\xi}{2n}\right) - 1$ and $\frac{1}{2}+\frac{\xi}{2n} > 0$,

$$\int \prod_{i=1}^{n}\left\{Z_i^{-\left(\frac{1}{2}-\frac{\xi}{2n}\right)} \exp\left(-\frac{Z_i}{\sigma}\right)\right\} d\mathbb{Z}$$

$$= \left\{ \sigma^{\left(\frac{1}{2}+\frac{\xi}{2n}\right)} \Gamma\left(\frac{1}{2}+\frac{\xi}{2n}\right) \right\}^n = \left\{ \Gamma\left(\frac{1}{2}+\frac{\xi}{2n}\right) \right\}^n \sigma^{\left(\frac{n}{2}+\frac{\xi}{2}\right)}.$$

Again we have for $\kappa = 0, 1$,

$$\begin{aligned} & f_3^{(\xi,\zeta,\lambda)}\left(\mathbb{Z}, \sigma, \boldsymbol{\Omega}_1, \boldsymbol{\Omega}_2, \boldsymbol{A} | \mathbb{X}, \mathbb{Y}\right) \\ & \quad \leq c_4 \left(\prod_{i=1}^{n} Z_i\right)^{-\left(\frac{1}{2}-\frac{\xi}{2n}\right)} f_4^{(\xi,\zeta,\lambda,0)}\left(\sigma, \boldsymbol{\Omega}_1, \boldsymbol{\Omega}_2, \boldsymbol{A} | \mathbb{X}, \mathbb{Y}\right), \end{aligned}$$

and

$$\int f_3^{(\xi,\zeta,\lambda)}\left(\mathbb{Z}, \sigma, \boldsymbol{\Omega}_1, \boldsymbol{\Omega}_2, \boldsymbol{A} | \mathbb{X}, \mathbb{Y}\right) d\mathbb{Z} \leq c_4 \; f_4^{(\xi,\zeta,\lambda,1)}\left(\sigma, \boldsymbol{\Omega}_1, \boldsymbol{\Omega}_2, \boldsymbol{A} | \mathbb{X}, \mathbb{Y}\right),$$

where $c_4 = 1 + \left\{\Gamma\left(\frac{1}{2}+\frac{\xi}{2n}\right)\right\}^n$,

$$\begin{aligned} & f_4^{(\xi,\zeta,\lambda,\kappa)}\left(\sigma, \boldsymbol{\Omega}_1, \boldsymbol{\Omega}_2, \boldsymbol{A} | \mathbb{X}, \mathbb{Y}\right) \\ & = \sigma^{-\left(n+\frac{n}{2}+\frac{u_\tau}{2}+a+1-\left(\frac{\xi}{2}+\frac{\zeta u_\tau}{2}+\frac{\kappa n}{2}+\frac{\kappa\xi}{2}\right)\right)} \exp\left(-\frac{b}{\sigma}\right) \\ & \quad |\boldsymbol{\Omega}_1|^{-\frac{n+\nu_1+u_\tau+1-\lambda}{2}} \exp\left[-\frac{1}{2}\,\mathrm{trace}\left(\boldsymbol{\Omega}_1^{-1}\boldsymbol{\Psi}_1\right)\right] \\ & \quad |\boldsymbol{\Omega}_2|^{-\frac{n+\nu_2+p-u_\tau+1-\lambda}{2}} \exp\left[-\frac{1}{2}\,\mathrm{trace}\left(\boldsymbol{\Omega}_2^{-1}\boldsymbol{\Psi}_2\right)\right] \\ & \quad \exp\left[-\frac{1}{2}\,\mathrm{trace}\left\{\boldsymbol{K}^{-1}\left(\boldsymbol{A}-\boldsymbol{A}_0\right)\boldsymbol{L}^{-1}\left(\boldsymbol{A}-\boldsymbol{A}_0\right)^T\right\}\right] \quad (< \infty). \end{aligned}$$

Consequently, it is enough to show that $f_4^{((\xi,\zeta,\lambda,\kappa)}\left(\sigma, \boldsymbol{\Omega}_1, \boldsymbol{\Omega}_2, \boldsymbol{A} | \mathbb{X}, \mathbb{Y}\right)$ is Lebesgue integrable with respect to any $1 \leq k \leq 3+u_\tau$ elements of the set $\{\sigma, \boldsymbol{\Omega}_1, \boldsymbol{\Omega}_2,$

$\boldsymbol{a}_1, \cdots, \boldsymbol{a}_{u_\tau}\}$ for $\xi, \zeta, \lambda, \kappa = 0, 1$.

The proof is completed by noting that $f_4^{(\xi,\zeta,\lambda,\kappa)}(\sigma, \boldsymbol{\Omega}_1, \boldsymbol{\Omega}_2, \boldsymbol{A}|\mathbb{X}, \mathbb{Y})$ is a constant multiple of the product of inverse gamma $(\sigma|\widehat{a},\ b)$, inverse Wishart $(\boldsymbol{\Omega}_1|\boldsymbol{\Psi}_1,\ n+\nu_1-\lambda)$, inverse Wishart $(\boldsymbol{\Omega}_2|\boldsymbol{\Psi}_2,\ n+\nu_2-\lambda)$, and matrix normal $(\boldsymbol{A}|\boldsymbol{A}_0,\ \boldsymbol{K}\ \boldsymbol{L})$ densities, where

$$\widehat{a} = n + \frac{n}{2} + \frac{u_\tau}{2} + a - \left[\frac{\xi}{2} + \frac{\zeta u_\tau}{2} + \frac{\kappa n}{2} + \frac{\kappa\xi}{2}\right] > 0.$$

□

Proof of Theorem 3

Arguments similar to those provided in the proof of Theorem 2 establishes ϕ-irreducibility and aperiodicity of the algorithm. To show Harris recurrence, without loss of generality, suppose that the first m responses are censored, i.e., $Y_1, \cdots, Y_m = 0$, and the rest $n-m$ responses are not censored $(n-m \geq 1)$. Let the latent (imputed) data corresponding to $Y_1, \cdots, Y_m$ be $Y_1^*, \cdots, Y_m^*$. Define $(\mathbb{Y}_{1:m}^*)^T = (Y_1^*, \cdots, Y_m^*)^T$, $(\mathbb{Y}_{(m+1):n})^T = (Y_{m+1}, \cdots, Y_n)^T$, $\mathbb{X}_{1:m}^T = (\boldsymbol{X}_1, \cdots \boldsymbol{X}_m)^T \in \mathbb{R}^{p\times m}$, $\mathbb{X}_{(m+1):n}^T = (\boldsymbol{X}_{m+1}, \cdots \boldsymbol{X}_n)^T \in \mathbb{R}^{p\times(n-m)}$, and $\mathbb{X}^T = (\mathbb{X}_{1:m}^T, \mathbb{X}_{(m+1):n}^T)^T \in \mathbb{R}^{p\times n}$. We define $\mathbb{Z}_{1:m}$, $\mathbb{Z}_{(m+1):n}$, $\mathbb{Z}$, $\mathbb{W}_{1:m}^*$, and $\mathbb{W}_{(m+1):n}$ similarly. At each iteration, the MCMC sampler blockwise updates the following $8 + u_\tau$ parameters (or latent data) blocks:

$\{\mathbb{Y}^*_{1:m}, \mathbb{Z}, \mu_{\tau,Y}, \boldsymbol{\mu}_{\boldsymbol{X}}, \boldsymbol{\eta}, \sigma, \boldsymbol{\Omega}_1, \boldsymbol{\Omega}_2, \boldsymbol{a}_1, \cdots, \boldsymbol{a}_{u_\tau}\}$.

The joint unnormalized posterior density of the parameters and the latent data is given by:

$$
\begin{aligned}
& f\left(\mu_{\tau,Y}, \boldsymbol{\mu}_{\boldsymbol{X}}, \boldsymbol{\eta}, \sigma, \boldsymbol{\Omega}_1, \boldsymbol{\Omega}_2, \boldsymbol{A}, \mathbb{Y}^*_{1:m} | \mathbb{X}, \mathbb{Y}_{(m+1):n}\right) \\
= & \exp\left[-\frac{1}{2\sigma\gamma^2}\left\{\mathbb{W}^*_{1:m} - \mu_{\tau,Y}\mathbf{1}_m - \left(\mathbb{X}_{1:m} - \mathbf{1}_m\boldsymbol{\mu}_{\boldsymbol{X}}^T\right)\boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})\boldsymbol{\eta}\right\}^T \boldsymbol{D}^{-1}_{1:m}\right. \\
& \qquad \left.\left\{\mathbb{W}^*_{1:m} - \mu_{\tau,Y}\mathbf{1}_m - \left(\mathbb{X}_{1:m} - \mathbf{1}_m\boldsymbol{\mu}_{\boldsymbol{X}}^T\right)\boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})\boldsymbol{\eta}\right\}\right] \\
& \exp\left[-\frac{1}{2\sigma\gamma^2}\left\{\mathbb{W}_{(m+1):n} - \mu_{\tau,Y}\mathbf{1}_{n-m} - \left(\mathbb{X}_{(m+1):n} - \mathbf{1}_{n-m}\boldsymbol{\mu}_{\boldsymbol{X}}^T\right)\boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})\boldsymbol{\eta}\right\}^T\right. \\
& \qquad \left.\boldsymbol{D}^{-1}_{(m+1):n}\left\{\mathbb{W}_{(m+1):n} - \mu_{\tau,Y}\mathbf{1}_{n-m} - \left(\mathbb{X}_{(m+1):n} - \mathbf{1}_{n-m}\boldsymbol{\mu}_{\boldsymbol{X}}^T\right)\boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})\boldsymbol{\eta}\right\}\right] \\
& \exp\left[-\frac{1}{2}\operatorname{trace}\left(\mathbb{X} - \mathbf{1}_n\boldsymbol{\mu}_{\boldsymbol{X}}^T\right)\left(\boldsymbol{\Gamma}_{1\tau}\boldsymbol{\Omega}_1^{-1}\boldsymbol{\Gamma}_{1\tau}^T + \boldsymbol{\Gamma}_{2\tau}\boldsymbol{\Omega}_2^{-1}\boldsymbol{\Gamma}_{2\tau}^T\right)\left(\mathbb{X} - \mathbf{1}_n\boldsymbol{\mu}_{\boldsymbol{X}}^T\right)^T\right] \\
& \exp\left[-\frac{1}{2\sigma\gamma^2}\left(\boldsymbol{\eta} - \boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})^T\boldsymbol{e}\right)^T \boldsymbol{M}\left(\boldsymbol{\eta} - \boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})^T\boldsymbol{e}\right)\right] \\
& \prod_{i=1}^n Z_i^{-\frac{1}{2}}\exp\left(-\frac{Z_i}{\sigma}\right)\sigma^{-(n+\frac{n}{2}+\frac{u_\tau}{2}+a+1)} \\
& |\boldsymbol{\Omega}_1|^{-\frac{n+\nu_1+u_\tau+1}{2}}\exp\left[-\frac{1}{2}\operatorname{trace}\left(\boldsymbol{\Omega}_1^{-1}\boldsymbol{\Psi}_1\right)\right] \\
& |\boldsymbol{\Omega}_2|^{-\frac{n+\nu_2+p-u_\tau+1}{2}}\exp\left[-\frac{1}{2}\operatorname{trace}\left(\boldsymbol{\Omega}_2^{-1}\boldsymbol{\Psi}_2\right)\right] \\
& \exp\left[-\frac{1}{2}\operatorname{trace}\left\{\boldsymbol{K}^{-1}\left(\boldsymbol{A} - \boldsymbol{A}_0\right)\boldsymbol{L}^{-1}\left(\boldsymbol{A} - \boldsymbol{A}_0\right)^T\right\}\right]\exp\left(-\frac{b}{\sigma}\right),
\end{aligned}
$$

and we will need to show that $f\left(\mu_{\tau,Y}, \boldsymbol{\mu}_{\boldsymbol{X}}, \boldsymbol{\eta}, \sigma, \boldsymbol{\Omega}_1, \boldsymbol{\Omega}_2, \boldsymbol{A}, \mathbb{Y}^*_{1:m} | \mathbb{X}, \mathbb{Y}_{(m+1):n}\right)$ is integrable with respect to any $1 \le k \le (8 + u_\tau)$ elements of the set $\{\mathbb{Y}^*_{1:m},\ \mathbb{Z},\ \mu_{\tau,Y},\ \boldsymbol{\mu}_{\boldsymbol{X}},\ \boldsymbol{\eta},\ \sigma,\ \boldsymbol{\Omega}_1,\ \boldsymbol{\Omega}_2,\ \boldsymbol{a}_1, \cdots, \boldsymbol{a}_{u_\tau}\}$.

Out of these 2^{8+u_τ} integrals, the 2^{7+u_τ} integrals that do not integrate out $\mathbb{Y}^*_{1:m}$ are all finite for each $\mathbb{Y}^*_{1:m} \in \mathbb{R}^m_-$ where $\mathbb{R}^m_- = (-\infty, 0)^m$. (We have already proved this in Theorem 2. Here we can simply take $({\mathbb{W}^*_{1:m}}^T, {\mathbb{W}_{(m+1):n}}^T)^T$ to be $\mathbb{W}$ in (S3.5). The $\boldsymbol{D}$ in (S3.5) is taken to be a diagonal matrix with the first m diagonal elements from $\boldsymbol{D}_{1:m}$ and the next $n-m$ elements from $\boldsymbol{D}_{(m+1):n}$.) So, we need to show that the remaining 2^{7+u_τ} integrals that do integrate out $\mathbb{Y}^*_{1:m}$ are all finite as well. Note that

$$
\begin{aligned}
&\int_{\mathbb{R}^m_-} \exp\left[-\frac{1}{2\sigma\gamma^2}\left\{\mathbb{W}^*_{1:m} - \mu_{\tau,Y}\mathbf{1}_m - \left(\mathbb{X}_{1:m} - \mathbf{1}_m\boldsymbol{\mu}_{\boldsymbol{X}}^T\right)\boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})\boldsymbol{\eta}\right\}^T \boldsymbol{D}^{-1}_{1:m}\right. \\
&\qquad \left.\left\{\mathbb{W}^*_{1:m} - \mu_{\tau,Y}\mathbf{1}_m - \left(\mathbb{X}_{1:m} - \mathbf{1}_m\boldsymbol{\mu}_{\boldsymbol{X}}^T\right)\boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})\boldsymbol{\eta}\right\}\right] d\mathbb{Y}^*_{1:m} \\
&= \int_{\mathbb{R}^m_-} \exp\left[-\sum_{i=1}^m \frac{1}{2\sigma\gamma^2 Z_i}\left(Y^*_i - \theta Z_i - \mu_{\tau,Y} - (\boldsymbol{X}_i - \boldsymbol{\mu}_{\boldsymbol{X}})^T\boldsymbol{\Gamma}_{1\tau}\boldsymbol{\eta}\right)^2\right] d\mathbb{Y}^*_{1:m} \\
&= (2\pi)^{m/2}\gamma^m\sigma^{m/2}\prod_{i=1}^m Z_i^{1/2}\prod_{i=1}^m\left\{1 - \Phi(\theta Z_1 - \mu_{\tau,Y} - (\boldsymbol{X}_i - \boldsymbol{\mu}_{\boldsymbol{X}})^T\boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})\boldsymbol{\eta})\right\} \\
&\leq (2\pi)^{m/2}\gamma^m\sigma^{m/2}\prod_{i=1}^m Z_i^{1/2}. \qquad \text{(S3.8)}
\end{aligned}
$$

Hence, using (S3.8), we have

$$
\begin{aligned}
&\int_{\mathbb{R}^m_-} f\left(\mu_{\tau,Y}, \boldsymbol{\mu}_{\boldsymbol{X}}, \mathbb{Z}, \boldsymbol{\eta}, \sigma, \boldsymbol{\Omega}_1, \boldsymbol{\Omega}_2, \boldsymbol{A}, \mathbb{Y}^*_{1:m}|\mathbb{X}, \mathbb{Y}_{(m+1):n}\right) d\mathbb{Y}^*_{1:m} \\
&\leq (2\pi)^{m/2}\gamma^m \\
&\quad \exp\left[-\frac{1}{2\sigma\gamma^2}\left\{\mathbb{W}_{(m+1):n} - \mu_{\tau,Y}\mathbf{1}_{n-m} - \left(\mathbb{X}_{(m+1):n} - \mathbf{1}_{n-m}\boldsymbol{\mu}_{\boldsymbol{X}}^T\right)\boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})\boldsymbol{\eta}\right\}^T\right.
\end{aligned}
$$

$$\boldsymbol{D}_{(m+1):n}^{-1}\left\{\mathbb{W}_{(m+1):n}-\mu_{\tau,Y}\mathbf{1}_{n-m}-\left(\mathbb{X}_{(m+1):n}-\mathbf{1}_{n-m}\boldsymbol{\mu}_{\boldsymbol{X}}^T\right)\boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})\boldsymbol{\eta}\right\}\Big]$$
$$\exp\left[-\frac{1}{2}\operatorname{trace}\left(\mathbb{X}-\mathbf{1}_n\boldsymbol{\mu}_{\boldsymbol{X}}^T\right)\left(\boldsymbol{\Gamma}_{1\tau}\boldsymbol{\Omega}_1^{-1}\boldsymbol{\Gamma}_{1\tau}^T+\boldsymbol{\Gamma}_{2\tau}\boldsymbol{\Omega}_2^{-1}\boldsymbol{\Gamma}_{2\tau}^T\right)\left(\mathbb{X}-\mathbf{1}_n\boldsymbol{\mu}_{\boldsymbol{X}}^T\right)^T\right]$$
$$\exp\left[-\frac{1}{2\sigma\gamma^2}\left(\boldsymbol{\eta}-\boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})^T\boldsymbol{e}\right)^T\boldsymbol{M}\left(\boldsymbol{\eta}-\boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})^T\boldsymbol{e}\right)\right]$$
$$\prod_{i=1}^m\exp\left(-\frac{Z_i}{\sigma}\right)\prod_{i=m+1}^n Z_i^{-\frac{1}{2}}\exp\left(-\frac{Z_i}{\sigma}\right)\sigma^{-(n+\frac{n-m}{2}+\frac{u_\tau}{2}+a+1)}\exp\left(-\frac{b}{\sigma}\right)$$
$$|\boldsymbol{\Omega}_1|^{-\frac{n+\nu_1+u_\tau+1}{2}}\exp\left[-\frac{1}{2}\operatorname{trace}\left(\boldsymbol{\Omega}_1^{-1}\boldsymbol{\Psi}_1\right)\right]$$
$$|\boldsymbol{\Omega}_2|^{-\frac{n+\nu_2+p-u_\tau+1}{2}}\exp\left[-\frac{1}{2}\operatorname{trace}\left(\boldsymbol{\Omega}_2^{-1}\boldsymbol{\Psi}_2\right)\right]$$
$$\exp\left[-\frac{1}{2}\operatorname{trace}\left\{\boldsymbol{K}^{-1}\left(\boldsymbol{A}-\boldsymbol{A}_0\right)\boldsymbol{L}^{-1}\left(\boldsymbol{A}-\boldsymbol{A}_0\right)^T\right\}\right]$$
$$\triangleq g\left(\mu_{\tau,Y},\boldsymbol{\mu}_{\boldsymbol{X}},\mathbb{Z},\boldsymbol{\eta},\sigma,\boldsymbol{\Omega}_1,\boldsymbol{\Omega}_2,\boldsymbol{A}|\mathbb{X},\mathbb{Y}_{(m+1):n}\right).$$

Using similar results as in (S3.6) and (S3.7),

$$g\left(\mu_{\tau,Y},\boldsymbol{\mu}_{\boldsymbol{X}},\mathbb{Z},\boldsymbol{\eta},\sigma,\boldsymbol{\Omega}_1,\boldsymbol{\Omega}_2,\boldsymbol{A}|\mathbb{X},\mathbb{Y}_{(m+1):n}\right)$$
$$\leq C_0\exp\left[-\frac{1}{2\sigma\gamma^2}\left(\sum_{i=m+1}^n\frac{1}{Z_i}\right)\left\{\overline{W}_{(m+1):n}-\mu_{\tau,Y}\right.\right.$$
$$\left.\left.-\boldsymbol{\eta}^T\boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})^T\left(\overline{\boldsymbol{X}}_{z,(m+1):n}-\boldsymbol{\mu}_{\boldsymbol{X}}\right)\right\}^2\right]$$
$$\exp\left[-\frac{1}{2}n\left(\boldsymbol{\mu}_{\boldsymbol{X}}-\overline{\boldsymbol{X}}\right)^T\left(\boldsymbol{\Gamma}_{1\tau}\boldsymbol{\Omega}_1^{-1}\boldsymbol{\Gamma}_{1\tau}^T+\boldsymbol{\Gamma}_{2\tau}\boldsymbol{\Omega}_2^{-1}\boldsymbol{\Gamma}_{2\tau}^T\right)\left(\boldsymbol{\mu}_{\boldsymbol{X}}-\overline{\boldsymbol{X}}\right)\right]$$
$$\exp\left[-\frac{1}{2\sigma\gamma^2}\left(\boldsymbol{\eta}-\boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})^T\boldsymbol{e}\right)^T\boldsymbol{M}\left(\boldsymbol{\eta}-\boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})^T\boldsymbol{e}\right)\right]$$
$$\prod_{i=1}^m\exp\left(-\frac{Z_i}{\sigma}\right)\prod_{i=m+1}^n Z_i^{-\frac{1}{2}}\exp\left(-\frac{Z_i}{\sigma}\right)\sigma^{-(n+\frac{n-m}{2}+\frac{u_\tau}{2}+a+1)}\exp\left(-\frac{b}{\sigma}\right)$$

$$
\begin{aligned}
&|\boldsymbol{\Omega}_1|^{-\frac{n+\nu_1+u_\tau+1}{2}} \exp\left[-\frac{1}{2}\,\mathrm{trace}\left(\boldsymbol{\Omega}_1^{-1}\boldsymbol{\Psi}_1\right)\right] \\
&|\boldsymbol{\Omega}_2|^{-\frac{n+\nu_2+p-u_\tau+1}{2}} \exp\left[-\frac{1}{2}\,\mathrm{trace}\left(\boldsymbol{\Omega}_2^{-1}\boldsymbol{\Psi}_2\right)\right] \\
&\exp\left[-\frac{1}{2}\,\mathrm{trace}\left\{\boldsymbol{K}^{-1}\left(\boldsymbol{A}-\boldsymbol{A}_0\right)\boldsymbol{L}^{-1}\left(\boldsymbol{A}-\boldsymbol{A}_0\right)^T\right\}\right] \\
\triangleq\; & C_0\; f_0\left(\mu_{\tau,Y}, \boldsymbol{\mu}_{\boldsymbol{X}}, \boldsymbol{\eta}, \sigma, \boldsymbol{\Omega}_1, \boldsymbol{\Omega}_2, \boldsymbol{A}, \mathbb{Z}|\mathbb{X}, \mathbb{Y}_{(m+1):n}\right),
\end{aligned}
$$

where $C_0 = (2\pi)^{m/2}\gamma^m$. Using similar arguments as used in the proof of Theorem 2, we have

$$
\begin{aligned}
&f_0\left(\mu_{\tau,Y}, \boldsymbol{\mu}_{\boldsymbol{X}}, \boldsymbol{\eta}, \mathbb{Z}, \sigma, \boldsymbol{\Omega}_1, \boldsymbol{\Omega}_2, \boldsymbol{A}|\mathbb{X}, \mathbb{Y}_{(m+1):n}\right) \\
&\qquad\qquad \leq C_1\; f_1^{(0)}\left(\boldsymbol{\mu}_{\boldsymbol{X}}, \boldsymbol{\eta}, \mathbb{Z}, \sigma, \boldsymbol{\Omega}_1, \boldsymbol{\Omega}_2, \boldsymbol{A}|\mathbb{X}, \mathbb{Y}_{(m+1):n}\right) \\
&\int f_0\left(\mu_{\tau,Y}, \boldsymbol{\mu}_{\boldsymbol{X}}, \boldsymbol{\eta}, \mathbb{Z}, \sigma, \boldsymbol{\Omega}_1, \boldsymbol{\Omega}_2, \boldsymbol{A}|\mathbb{X}, \mathbb{Y}_{(m+1):n}\right) d\mu_{\tau,Y} \\
&\qquad\qquad \leq C_1\; f_1^{(1)}\left(\boldsymbol{\mu}_{\boldsymbol{X}}, \boldsymbol{\eta}, \mathbb{Z}, \sigma, \boldsymbol{\Omega}_1, \boldsymbol{\Omega}_2, \boldsymbol{A}|\mathbb{X}, \mathbb{Y}_{(m+1):n}\right),
\end{aligned}
$$

where $C_1 = 1 + (2\pi\gamma^2(n-m)^{-1})^{1/2}$, and for $\xi = 0, 1$,

$$
\begin{aligned}
&f_1^{(\xi)}\left(\boldsymbol{\mu}_{\boldsymbol{X}}, \boldsymbol{\eta}, \mathbb{Z}, \sigma, \boldsymbol{\Omega}_1, \boldsymbol{\Omega}_2, \boldsymbol{A}|\mathbb{X}, \mathbb{Y}_{(m+1):n}\right) \\
&= \exp\left[-\frac{1}{2}n\left(\boldsymbol{\mu}_{\boldsymbol{X}} - \overline{\boldsymbol{X}}\right)^T\left(\boldsymbol{\Gamma}_{1\tau}\boldsymbol{\Omega}_1^{-1}\boldsymbol{\Gamma}_{1\tau}^T + \boldsymbol{\Gamma}_{2\tau}\boldsymbol{\Omega}_2^{-1}\boldsymbol{\Gamma}_{2\tau}^T\right)\left(\boldsymbol{\mu}_{\boldsymbol{X}} - \overline{\boldsymbol{X}}\right)\right] \\
&\quad \exp\left[-\frac{1}{2\sigma\gamma^2}\left(\boldsymbol{\eta} - \boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})^T\boldsymbol{e}\right)^T\boldsymbol{M}\left(\boldsymbol{\eta} - \boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})^T\boldsymbol{e}\right)\right] \\
&\quad \prod_{i=1}^{m}\exp\left(-\frac{Z_i}{\sigma}\right)\prod_{i=m+1}^{n} Z_i^{-\left(\frac{1}{2}-\frac{\xi}{2(n-m)}\right)}\exp\left(-\frac{Z_i}{\sigma}\right)
\end{aligned}
$$

$$\sigma^{-(n+\frac{n-m}{2}+\frac{u_\tau}{2}+a+1-\frac{\xi}{2})}\exp\left(-\frac{b}{\sigma}\right)$$
$$|\boldsymbol{\Omega}_1|^{-\frac{n+\nu_1+u_\tau+1}{2}}\exp\left[-\frac{1}{2}\operatorname{trace}\left(\boldsymbol{\Omega}_1^{-1}\boldsymbol{\Psi}_1\right)\right]$$
$$|\boldsymbol{\Omega}_2|^{-\frac{n+\nu_2+p-u_\tau+1}{2}}\exp\left[-\frac{1}{2}\operatorname{trace}\left(\boldsymbol{\Omega}_2^{-1}\boldsymbol{\Psi}_2\right)\right]$$
$$\exp\left[-\frac{1}{2}\operatorname{trace}\left\{\boldsymbol{K}^{-1}\left(\boldsymbol{A}-\boldsymbol{A}_0\right)\boldsymbol{L}^{-1}\left(\boldsymbol{A}-\boldsymbol{A}_0\right)^T\right\}\right].$$

Further,

$$f_1^{(\xi)}\left(\boldsymbol{\mu_X},\boldsymbol{\eta},\mathbb{Z},\sigma,\boldsymbol{\Omega}_1,\boldsymbol{\Omega}_2,\boldsymbol{A}|\mathbb{X},\mathbb{Y}_{(m+1):n}\right)$$
$$\leq C_2\; f_2^{(\xi,0)}\left(\boldsymbol{\mu_X},\mathbb{Z},\sigma,\boldsymbol{\Omega}_1,\boldsymbol{\Omega}_2,\boldsymbol{A}|\mathbb{X},\mathbb{Y}_{(m+1):n}\right)$$
$$\int f_1^{(\xi)}\left(\boldsymbol{\mu_X},\boldsymbol{\eta},\mathbb{Z},\sigma,\boldsymbol{\Omega}_1,\boldsymbol{\Omega}_2,\boldsymbol{A}|\mathbb{X},\mathbb{Y}_{(m+1):n}\right)d\boldsymbol{\eta}$$
$$\leq C_2\; f_2^{(\xi,1)}\left(\boldsymbol{\mu_X},\mathbb{Z},\sigma,\boldsymbol{\Omega}_1,\boldsymbol{\Omega}_2,\boldsymbol{A}|\mathbb{X},\mathbb{Y}_{(m+1):n}\right),$$

where $C_2 = 1 + ((2\pi)^{u_\tau}\gamma^{2u_\tau}|\boldsymbol{M}|^{-1})^{1/2}$ and for $\zeta = 0, 1$,

$$f_2^{(\xi,\zeta)}\left(\boldsymbol{\mu_X},\mathbb{Z},\sigma,\boldsymbol{\Omega}_1,\boldsymbol{\Omega}_2,\boldsymbol{A}|\mathbb{X},\mathbb{Y}_{(m+1):n}\right)$$
$$=\exp\left[-\frac{1}{2}n\left(\boldsymbol{\mu_X}-\overline{\boldsymbol{X}}\right)^T\left(\boldsymbol{\Gamma}_{1\tau}\boldsymbol{\Omega}_1^{-1}\boldsymbol{\Gamma}_{1\tau}^T+\boldsymbol{\Gamma}_{2\tau}\boldsymbol{\Omega}_2^{-1}\boldsymbol{\Gamma}_{2\tau}^T\right)\left(\boldsymbol{\mu_X}-\overline{\boldsymbol{X}}\right)\right]$$
$$\prod_{i=1}^{m}\exp\left(-\frac{Z_i}{\sigma}\right)\prod_{i=m+1}^{n}Z_i^{-(\frac{1}{2}-\frac{\xi}{2(n-m)})}\exp\left(-\frac{Z_i}{\sigma}\right)$$
$$\sigma^{-(n+\frac{n-m}{2}+\frac{u_\tau}{2}+a+1-(\frac{\xi}{2}+\frac{\zeta u_\tau}{2}))}\exp\left(-\frac{b}{\sigma}\right)$$
$$|\boldsymbol{\Omega}_1|^{-\frac{n+\nu_1+u_\tau+1}{2}}\exp\left[-\frac{1}{2}\operatorname{trace}\left(\boldsymbol{\Omega}_1^{-1}\boldsymbol{\Psi}_1\right)\right]$$

$$|\boldsymbol{\Omega}_2|^{-\frac{n+\nu_2+p-u_\tau+1}{2}} \exp\left[-\frac{1}{2}\operatorname{trace}\left(\boldsymbol{\Omega}_2^{-1}\boldsymbol{\Psi}_2\right)\right]$$
$$\exp\left[-\frac{1}{2}\operatorname{trace}\left\{\boldsymbol{K}^{-1}\left(\boldsymbol{A}-\boldsymbol{A}_0\right)\boldsymbol{L}^{-1}\left(\boldsymbol{A}-\boldsymbol{A}_0\right)^T\right\}\right] < \infty.$$

We keep following this procedure as that in the proof of Theorem 2, and have

$$f_2^{(\xi,\zeta)}\left(\boldsymbol{\mu}_{\boldsymbol{X}}, \mathbb{Z}, \sigma, \boldsymbol{\Omega}_1, \boldsymbol{\Omega}_2, \boldsymbol{A}|\mathbb{X}, \mathbb{Y}_{(m+1):n}\right)$$
$$\leq C_3\ f_3^{(\xi,\zeta,0)}\left(\mathbb{Z}, \sigma, \boldsymbol{\Omega}_1, \boldsymbol{\Omega}_2, \boldsymbol{A}|\mathbb{X}, \mathbb{Y}_{(m+1):n}\right),$$
$$\int f_2^{(\xi,\zeta)}\left(\boldsymbol{\mu}_{\boldsymbol{X}}, \mathbb{Z}, \sigma, \boldsymbol{\Omega}_1, \boldsymbol{\Omega}_2, \boldsymbol{A}|\mathbb{X}, \mathbb{Y}_{(m+1):n}\right) d\boldsymbol{\mu}_{\boldsymbol{X}}$$
$$\leq C_3\ f_3^{(\xi,\zeta,1)}\left(\mathbb{Z}, \sigma, \boldsymbol{\Omega}_1, \boldsymbol{\Omega}_2, \boldsymbol{A}|\mathbb{X}, \mathbb{Y}_{(m+1):n}\right),$$

where $C_3 = 1 + (2\pi n^{-1})^{p/2}$ and for $\lambda = 0, 1$,

$$f_3^{(\xi,\zeta,\lambda)}\left(\mathbb{Z}, \sigma, \boldsymbol{\Omega}_1, \boldsymbol{\Omega}_2, \boldsymbol{A}|\mathbb{X}, \mathbb{Y}_{(m+1):n}\right)$$
$$= \prod_{i=1}^{m} \exp\left(-\frac{Z_i}{\sigma}\right) \prod_{i=m+1}^{n} Z_i^{-\left(\frac{1}{2}-\frac{\xi}{2(n-m)}\right)} \exp\left(-\frac{Z_i}{\sigma}\right)$$
$$\sigma^{-\left(n+\frac{n-m}{2}+\frac{u_\tau}{2}+a+1-\left(\frac{\xi}{2}+\frac{\zeta u_\tau}{2}\right)\right)} \exp\left(-\frac{b}{\sigma}\right)$$
$$|\boldsymbol{\Omega}_1|^{-\frac{n+\nu_1+u_\tau+1-\lambda}{2}} \exp\left[-\frac{1}{2}\operatorname{trace}\left(\boldsymbol{\Omega}_1^{-1}\boldsymbol{\Psi}_1\right)\right]$$
$$|\boldsymbol{\Omega}_2|^{-\frac{n+\nu_2+p-u_\tau+1-\lambda}{2}} \exp\left[-\frac{1}{2}\operatorname{trace}\left(\boldsymbol{\Omega}_2^{-1}\boldsymbol{\Psi}_2\right)\right]$$
$$\exp\left[-\frac{1}{2}\operatorname{trace}\left\{\boldsymbol{K}^{-1}\left(\boldsymbol{A}-\boldsymbol{A}_0\right)\boldsymbol{L}^{-1}\left(\boldsymbol{A}-\boldsymbol{A}_0\right)^T\right\}\right] < \infty.$$

Note that

$$\prod_{i=1}^{m} \exp\left(-\frac{Z_i}{\sigma}\right) \prod_{i=m+1}^{n} Z_i^{-(\frac{1}{2}-\frac{\xi}{2(n-m)})} \exp\left(-\frac{Z_i}{\sigma}\right) \leq \prod_{i=m+1}^{n} Z_i^{-(\frac{1}{2}-\frac{\xi}{2(n-m)})}.$$

Since $-(\frac{1}{2}-\frac{\xi}{2(n-m)}) = \frac{1}{2}+\frac{\xi}{2(n-m)} - 1$ and $\frac{1}{2}+\frac{\xi}{2(n-m)} > 0$, we have

$$\begin{aligned}
&\int \prod_{i=1}^{m} \exp\left(-\frac{Z_i}{\sigma}\right) \prod_{i=m+1}^{n} \left\{ Z_i^{-(\frac{1}{2}-\frac{\xi}{2(n-m)})} \exp\left(-\frac{Z_i}{\sigma}\right) \right\} d\mathbb{Z} \\
&= \sigma^m \left(\sigma^{\frac{1}{2}+\frac{\xi}{2(n-m)}}\right)^{n-m} \left\{ \Gamma\left(\frac{1}{2}+\frac{\xi}{2(n-m)}\right) \right\}^{n-m} \\
&= \left\{ \Gamma\left(\frac{1}{2}+\frac{\xi}{2(n-m)}\right) \right\}^{n-m} \sigma^{m+\frac{n-m}{2}+\frac{\xi}{2}}.
\end{aligned}$$

Then,

$$\begin{aligned}
&f_3^{(\xi,\zeta,\lambda)}\left(\mathbb{Z}, \sigma, \boldsymbol{\Omega}_1, \boldsymbol{\Omega}_2, \boldsymbol{A} | \mathbb{X}, \mathbb{Y}_{(m+1):n}\right) \\
&\qquad \leq C_4 \left(\prod_{i=m+1}^{n} Z_i^{-(\frac{1}{2}-\frac{\xi}{2(n-m)})} \right) f_4^{(\xi,\zeta,\lambda,0)}\left(\sigma, \boldsymbol{\Omega}_1, \boldsymbol{\Omega}_2, \boldsymbol{A} | \mathbb{X}, \mathbb{Y}_{(m+1):n}\right), \\
&\int f_3^{(\xi,\zeta,\lambda)}\left(\mathbb{Z}, \sigma, \boldsymbol{\Omega}_1, \boldsymbol{\Omega}_2, \boldsymbol{A} | \mathbb{X}, \mathbb{Y}_{(m+1):n}\right) d\mathbb{Z} \\
&\qquad \leq C_4 \, f_4^{(\xi,\zeta,\lambda,1)}\left(\sigma, \boldsymbol{\Omega}_1, \boldsymbol{\Omega}_2, \boldsymbol{A} | \mathbb{X}, \mathbb{Y}_{(m+1):n}\right),
\end{aligned}$$

where $C_4 = 1 + \left\{ \Gamma\left(\frac{1}{2}+\frac{\xi}{2(n-m)}\right) \right\}^{n-m}$, and for $\kappa = 0, 1$,

$$f_4^{(\xi,\zeta,\lambda,\kappa)}\left(\sigma, \boldsymbol{\Omega}_1, \boldsymbol{\Omega}_2, \boldsymbol{A} | \mathbb{X}, \mathbb{Y}_{(m+1):n}\right)$$

$$
\begin{aligned}
&= \sigma^{-(n+\frac{n-m}{2}+\frac{u_\mathcal{T}}{2}+a+1-(\frac{\xi}{2}+\frac{\zeta u_\mathcal{T}}{2}+\kappa m+\frac{\kappa(n-m)}{2}+\frac{\kappa\xi}{2}))} \exp\left(-\frac{b}{\sigma}\right) \\
&\quad |\boldsymbol{\Omega}_1|^{-\frac{n+\nu_1+u_\mathcal{T}+1-\lambda}{2}} \exp\left[-\frac{1}{2}\operatorname{trace}\left(\boldsymbol{\Omega}_1^{-1}\boldsymbol{\Psi}_1\right)\right] \\
&\quad |\boldsymbol{\Omega}_2|^{-\frac{n+\nu_2+p-u_\mathcal{T}+1-\lambda}{2}} \exp\left[-\frac{1}{2}\operatorname{trace}\left(\boldsymbol{\Omega}_2^{-1}\boldsymbol{\Psi}_2\right)\right] \\
&\quad \exp\left[-\frac{1}{2}\operatorname{trace}\left\{\boldsymbol{K}^{-1}\left(\boldsymbol{A}-\boldsymbol{A}_0\right)\boldsymbol{L}^{-1}\left(\boldsymbol{A}-\boldsymbol{A}_0\right)^T\right\}\right].
\end{aligned}
$$

Since $(\frac{\xi}{2}+\frac{\zeta u_\mathcal{T}}{2}+\kappa m+\frac{\kappa(n-m)}{2}+\frac{\kappa\xi}{2}) \leq (1+m+\frac{u_\mathcal{T}}{2}+\frac{n-m}{2})$,

$$
\begin{aligned}
&\sigma^{-(n+\frac{n-m}{2}+\frac{u_\mathcal{T}}{2}+a+1-(\frac{\xi}{2}+\frac{\zeta u_\mathcal{T}}{2}+\kappa m+\frac{\kappa(n-m)}{2}+\frac{\kappa\xi}{2}))} \exp\left(-\frac{b}{\sigma}\right) \\
&\qquad \leq \sigma^{-(n-m-1+a+1)} \exp\left(-\frac{b}{\sigma}\right),
\end{aligned}
$$

and the right hand side is a constant multiple of an inverse gamma $(n-m-1+a,\ b)$ density. The remainder of the proof is similar to the proof of Theorem 2. □

S4 Derivations of Full Conditional Distributions

Derivations for various full conditional posterior densities for the parameters in Algorithm 1 are provided as follows.

1. For $\mu_{\tau,Y}$,

$$
\begin{aligned}
& g(\mu_{\tau,Y}|\boldsymbol{\mu}_{\boldsymbol{X}}, \boldsymbol{\eta}, \sigma, \boldsymbol{\Omega}_1, \boldsymbol{\Omega}_2, \boldsymbol{A}) \\
&= \frac{1}{2\sigma\gamma^2} \left\{\mathbb{W} - \mu_{\tau,Y}\mathbf{1}_n - \left(\mathbb{X} - \mathbf{1}_n\boldsymbol{\mu}_{\boldsymbol{X}}^T\right)\boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})\boldsymbol{\eta}\right\}^T \boldsymbol{D}^{-1} \\
&\qquad \left\{\mathbb{W} - \mu_{\tau,Y}\mathbf{1}_n - \left(\mathbb{X} - \mathbf{1}_n\boldsymbol{\mu}_{\boldsymbol{X}}^T\right)\boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})\boldsymbol{\eta}\right\} \\
&= \frac{1}{2\sigma\gamma^2} \left\{\mathbb{W} - \mu_{\tau,Y}\mathbf{1}_n - \left(\mathbb{X} - \mathbf{1}_n\boldsymbol{\mu}_{\boldsymbol{X}}^T\right)\boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})\boldsymbol{\eta}\right\}^T \boldsymbol{D}^{-1/2}\left(\boldsymbol{P}_{\boldsymbol{D}^{-1/2}\mathbf{1}_n} + \boldsymbol{Q}_{\boldsymbol{D}^{-1/2}\mathbf{1}_n}\right) \\
&\qquad \boldsymbol{D}^{-1/2}\left\{\mathbb{W} - \mu_{\tau,Y}\mathbf{1}_n - \left(\mathbb{X} - \mathbf{1}_n\boldsymbol{\mu}_{\boldsymbol{X}}^T\right)\boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})\boldsymbol{\eta}\right\} \\
&= \frac{1}{2\sigma\gamma^2} \left\{\mathbb{W} - \mu_{\tau,Y}\mathbf{1}_n - \left(\mathbb{X} - \mathbf{1}_n\boldsymbol{\mu}_{\boldsymbol{X}}^T\right)\boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})\boldsymbol{\eta}\right\}^T \boldsymbol{D}^{-1/2}\boldsymbol{P}_{\boldsymbol{D}^{-1/2}\mathbf{1}_n} \\
&\qquad \boldsymbol{D}^{-1/2}\left\{\mathbb{W} - \mu_{\tau,Y}\mathbf{1}_n - \left(\mathbb{X} - \mathbf{1}_n\boldsymbol{\mu}_{\boldsymbol{X}}^T\right)\boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})\boldsymbol{\eta}\right\} \\
&\qquad + \text{term indep. of } \mu_{\tau,Y} \\
&= \frac{1}{2\sigma\gamma^2}\left(\sum_{i=1}^{n}\frac{1}{Z_i}\right)\left\{\overline{W}_Z - \mu_{\tau,Y} - \boldsymbol{\eta}^T\boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})^T\left(\overline{\boldsymbol{X}}_Z - \boldsymbol{\mu}_{\boldsymbol{X}}\right)\right\}^T \\
&\qquad \left\{\overline{W}_Z - \mu_{\tau,Y} - \boldsymbol{\eta}^T\boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})^T\left(\overline{\boldsymbol{X}}_Z - \boldsymbol{\mu}_{\boldsymbol{X}}\right)\right\} \\
&\qquad + \text{term indep. of } \mu_{\tau,Y} \\
&= \frac{1}{2\sigma\gamma^2}\left(\sum_{i=1}^{n}\frac{1}{Z_i}\right)\left\{\mu_{\tau,Y} - \overline{W}_Z - \left(\boldsymbol{\mu}_{\boldsymbol{X}} - \overline{\boldsymbol{X}}_Z\right)^T\boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})\boldsymbol{\eta}\right\}^T \\
&\qquad \left\{\mu_{\tau,Y} - \overline{W}_Z - \left(\boldsymbol{\mu}_{\boldsymbol{X}} - \overline{\boldsymbol{X}}_Z\right)^T\boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})\boldsymbol{\eta}\right\} \\
&\qquad + \text{term indep. of } \mu_{\tau,Y},
\end{aligned}
$$

where $\boldsymbol{P}_{\boldsymbol{D}^{-1/2}\mathbf{1}_n}$ is a projection matrix onto $\text{span}(\boldsymbol{D}^{-1/2}\mathbf{1}_n)$, $\overline{W}_Z =$

$\frac{1}{\sum_{i=1}^{n} 1/Z_i} \sum_{i=1}^{n} \frac{W_i}{Z_i}$ and $\overline{\boldsymbol{X}}_Z = \frac{1}{\sum_{i=1}^{n} 1/Z_i} \sum_{i=1}^{n} \frac{1}{Z_i} \boldsymbol{X}_i \in \mathbb{R}^p$. Thus we have

$$\mu_{\tau,Y}|\text{rest} \sim N\left(\overline{W}_Z + \left(\boldsymbol{\mu_X} - \overline{\boldsymbol{X}}_Z\right)^T \boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})\boldsymbol{\eta}, \quad \frac{1}{\sum_{i=1}^{n} 1/Z_i}\sigma\gamma^2\right)$$

2. For $\boldsymbol{\mu_X}$,

$$\begin{aligned}
&g(\boldsymbol{\mu_X}|\mu_{\tau,Y}, \boldsymbol{\eta}, \sigma, \boldsymbol{\Omega}_1, \boldsymbol{\Omega}_2, \boldsymbol{A}) \\
&= -\frac{1}{2\sigma\gamma^2}\left[\left\{\mathbb{W} - \mu_{\tau,Y}\mathbf{1}_n - \left(\mathbb{X} - \mathbf{1}_n\boldsymbol{\mu}_{\boldsymbol{X}}^T\right)\boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})\boldsymbol{\eta}\right\}^T \boldsymbol{D}^{-1}\right. \\
&\qquad\qquad \left.\left\{\mathbb{W} - \mu_{\tau,Y}\mathbf{1}_n - \left(\mathbb{X} - \mathbf{1}_n\boldsymbol{\mu}_{\boldsymbol{X}}^T\right)\boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})\boldsymbol{\eta}\right\}\right] \\
&- \frac{1}{2}\operatorname{trace}\left\{\left(\mathbb{X} - \mathbf{1}_n\boldsymbol{\mu}_{\boldsymbol{X}}^T\right)\left(\boldsymbol{\Gamma}_{1\tau}\boldsymbol{\Omega}_1^{-1}\boldsymbol{\Gamma}_{1\tau}^T + \boldsymbol{\Gamma}_{2\tau}\boldsymbol{\Omega}_2^{-1}\boldsymbol{\Gamma}_{2\tau}^T\right)\left(\mathbb{X} - \mathbf{1}_n\boldsymbol{\mu}_{\boldsymbol{X}}^T\right)^T\right\}.
\end{aligned}$$

For the first term, we have

$$\begin{aligned}
&-\frac{1}{2\sigma\gamma^2}\left(\sum_{i=1}^{n}\frac{1}{Z_i}\right)\left[\left\{\mu_{\tau,Y} - \overline{W}_Z - \boldsymbol{\eta}^T\boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})^T\left(\boldsymbol{\mu_X} - \overline{\boldsymbol{X}}_Z\right)\right\}^T\right. \\
&\qquad \left.\left\{\mu_{\tau,Y} - \overline{W}_Z - \boldsymbol{\eta}^T\boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})^T\left(\boldsymbol{\mu_X} - \overline{\boldsymbol{X}}_Z\right)\right\}\right] + \text{term indep. of } \boldsymbol{\mu_X} \\
&= -\frac{1}{2\sigma\gamma^2}\left(\sum_{i=1}^{n}\frac{1}{Z_i}\right)\left\{\left(\boldsymbol{\mu_X} - \overline{\boldsymbol{X}}_Z\right)^T\boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})\boldsymbol{\eta}\boldsymbol{\eta}^T\boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})^T\left(\boldsymbol{\mu_X} - \overline{\boldsymbol{X}}_Z\right)\right. \\
&\qquad \left.-2\left(\boldsymbol{\mu_X} - \overline{\boldsymbol{X}}_Z\right)^T\boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})\boldsymbol{\eta}\left(\mu_{\tau,Y} - \overline{W}_Z\right)\right\} + \text{term indep. of } \boldsymbol{\mu_X} \\
&= -\frac{1}{2\sigma\gamma^2}\left(\sum_{i=1}^{n}\frac{1}{Z_i}\right)\left\{\boldsymbol{\mu}_{\boldsymbol{X}}^T\boldsymbol{\Gamma}_{1\tau}\boldsymbol{\eta}\boldsymbol{\eta}^T\boldsymbol{\Gamma}_{1\tau}^T\boldsymbol{\mu_X} - 2\boldsymbol{\mu}_{\boldsymbol{X}}^T\boldsymbol{\Gamma}_{1\tau}\boldsymbol{\eta}\boldsymbol{\eta}^T\boldsymbol{\Gamma}_{1\tau}^T\overline{\boldsymbol{X}}_Z\right. \\
&\qquad \left.-2\boldsymbol{\mu}_{\boldsymbol{X}}^T\boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})\boldsymbol{\eta}\left(\mu_{\tau,Y} - \overline{W}_Z\right)\right\} + \text{term indep. of } \boldsymbol{\mu_X}.
\end{aligned}$$

For the second term, we have

$$
\begin{aligned}
&-\frac{1}{2}\,\text{trace}\left\{n\left(\boldsymbol{\Gamma}_{1\tau}\boldsymbol{\Omega}_1^{-1}\boldsymbol{\Gamma}_{1\tau}^T+\boldsymbol{\Gamma}_{2\tau}\boldsymbol{\Omega}_2^{-1}\boldsymbol{\Gamma}_{2\tau}^T\right)\left(\overline{\boldsymbol{X}}-\boldsymbol{\mu}_{\boldsymbol{X}}\right)\left(\overline{\boldsymbol{X}}-\boldsymbol{\mu}_{\boldsymbol{X}}\right)^T\right\}\\
&=-\frac{1}{2}n\left\{\boldsymbol{\mu}_{\boldsymbol{X}}^T\left(\boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})\boldsymbol{\Omega}_1^{-1}\boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})^T+\boldsymbol{\Gamma}_{2\tau}(\boldsymbol{A})\boldsymbol{\Omega}_2^{-1}\boldsymbol{\Gamma}_{2\tau}(\boldsymbol{A})^T\right)\boldsymbol{\mu}_{\boldsymbol{X}}\right.\\
&\qquad\left.-2\boldsymbol{\mu}_{\boldsymbol{X}}^T\left(\boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})\boldsymbol{\Omega}_1^{-1}\boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})^T+\boldsymbol{\Gamma}_{2\tau}(\boldsymbol{A})\boldsymbol{\Omega}_2^{-1}\boldsymbol{\Gamma}_{2\tau}(\boldsymbol{A})^T\right)\overline{\boldsymbol{X}}\right\}\\
&\qquad+\text{term indep. of }\boldsymbol{\mu}_{\boldsymbol{X}}.
\end{aligned}
$$

So,

$$
\begin{aligned}
&\log\pi\left(\boldsymbol{\mu}_{\boldsymbol{X}}|\text{rest}\right)\\
&=-\frac{1}{2}\left[\boldsymbol{\mu}_{\boldsymbol{X}}^T\left\{\frac{1}{\sigma\gamma^2}\left(\sum_{i=1}^n\frac{1}{Z_i}\right)\boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})\boldsymbol{\eta}\boldsymbol{\eta}^T\boldsymbol{\Gamma}^T(\boldsymbol{A})\right.\right.\\
&\qquad\left.+n\boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})\boldsymbol{\Omega}_1^{-1}\boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})^T+n\boldsymbol{\Gamma}_{2\tau}(\boldsymbol{A})\boldsymbol{\Omega}_2^{-1}\boldsymbol{\Gamma}_{2\tau}(\boldsymbol{A})^T\right\}\boldsymbol{\mu}_{\boldsymbol{X}}\\
&\quad-2\boldsymbol{\mu}_{\boldsymbol{X}}^T\left\{\frac{1}{\sigma\gamma^2}\left(\sum_{i=1}^n\frac{1}{Z_i}\right)\boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})\boldsymbol{\eta}\left(\boldsymbol{\eta}^T\boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})^T\overline{\boldsymbol{X}}_Z+\mu_{\tau,Y}-\overline{W}_Z\right)\right.\\
&\qquad\left.\left.+n\left(\boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})\boldsymbol{\Omega}_1^{-1}\boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})^T+\boldsymbol{\Gamma}_{2\tau}(\boldsymbol{A})\boldsymbol{\Omega}_2^{-1}\boldsymbol{\Gamma}_{2\tau}(\boldsymbol{A})^T\right)\overline{\boldsymbol{X}}\right\}\right]\\
&\quad+\text{term indep. of }\boldsymbol{\mu}_{\boldsymbol{X}}.
\end{aligned}
$$

Then,

$$
\begin{aligned}
\boldsymbol{\mu}_{\boldsymbol{X}}|\text{rest}\sim N&\left(\Delta_{\boldsymbol{\mu}_{\boldsymbol{X}}}^{-1}\left\{\frac{1}{\sigma\gamma^2}\left(\sum_{i=1}^n\frac{1}{Z_i}\right)\boldsymbol{\Gamma}_{1\tau}\boldsymbol{\eta}\left(\boldsymbol{\eta}^T\boldsymbol{\Gamma}_{1\tau}^T\overline{\boldsymbol{X}}_Z+\mu_{\tau,Y}-\overline{W}_Z\right)\right.\right.\\
&\qquad\left.\left.+n\left(\boldsymbol{\Gamma}_{1\tau}\boldsymbol{\Omega}_1^{-1}\boldsymbol{\Gamma}_{1\tau}^T+\boldsymbol{\Gamma}_{2\tau}\boldsymbol{\Omega}_2^{-1}\boldsymbol{\Gamma}_{2\tau}^T\right)\overline{\boldsymbol{X}}\right\},\quad\Delta_{\boldsymbol{\mu}_{\boldsymbol{X}}}^{-1}\right),
\end{aligned}
$$

where $\Delta_{\boldsymbol{\mu}_{\boldsymbol{X}}} = \frac{1}{\sigma\gamma^2}\left(\sum_{i=1}^{n}\frac{1}{Z_i}\right)\boldsymbol{\Gamma}_{1\tau}\boldsymbol{\eta}\boldsymbol{\eta}^T\boldsymbol{\Gamma}_{1\tau}^T + n\left(\boldsymbol{\Gamma}_{1\tau}\boldsymbol{\Omega}_1^{-1}\boldsymbol{\Gamma}_{1\tau}^T + \boldsymbol{\Gamma}_{2\tau}\boldsymbol{\Omega}_2^{-1}\boldsymbol{\Gamma}_{2\tau}^T\right)$.

3. For $\boldsymbol{\eta}$,

$$
\begin{aligned}
& g(\boldsymbol{\eta}|\mu_{\tau,Y}, \boldsymbol{\mu}_{\boldsymbol{X}}, \sigma, \boldsymbol{\Omega}_1, \boldsymbol{\Omega}_2, \boldsymbol{A}) \\
&= -\frac{1}{2\sigma\gamma^2}\left[\left\{\mathbb{W} - \mu_{\tau,Y}\mathbf{1}_n - \left(\mathbb{X} - \mathbf{1}_n\boldsymbol{\mu}_{\boldsymbol{X}}^T\right)\boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})\boldsymbol{\eta}\right\}^T \boldsymbol{D}^{-1}\right. \\
&\qquad \left.\left\{\mathbb{W} - \mu_{\tau,Y}\mathbf{1}_n - \left(\mathbb{X} - \mathbf{1}_n\boldsymbol{\mu}_{\boldsymbol{X}}^T\right)\boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})\boldsymbol{\eta}\right\}\right] \\
&- \frac{1}{2\sigma\gamma^2}\left(\boldsymbol{\eta} - \boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})^T\boldsymbol{e}\right)^T \boldsymbol{M}\left(\boldsymbol{\eta} - \boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})^T\boldsymbol{e}\right) \\
&= -\frac{1}{2\sigma\gamma^2}\left[\boldsymbol{\eta}^T\left\{\boldsymbol{\Gamma}_{1\tau}^T\left(\mathbb{X} - \mathbf{1}_n\boldsymbol{\mu}_{\boldsymbol{X}}^T\right)^T \boldsymbol{D}^{-1}\left(\mathbb{X} - \mathbf{1}_n\boldsymbol{\mu}_{\boldsymbol{X}}^T\right)\boldsymbol{\Gamma}_{1\tau} + \boldsymbol{M}\right\}\boldsymbol{\eta}\right. \\
&\qquad \left. -2\boldsymbol{\eta}^T\left\{\boldsymbol{\Gamma}_{1\tau}^T\left(\mathbb{X} - \mathbf{1}_n\boldsymbol{\mu}_{\boldsymbol{X}}^T\right)^T \boldsymbol{D}^{-1}\left(\mathbb{W} - \mu_{\tau,Y}\mathbf{1}_n\right) + \boldsymbol{M}\boldsymbol{\Gamma}_{1\tau}^T\boldsymbol{e}\right\}\right] \\
&+ \text{term indep. of } \boldsymbol{\eta}.
\end{aligned}
$$

So, we have

$$
\begin{aligned}
& \boldsymbol{\eta}|\text{rest} \sim \\
& N\left(\frac{1}{\sigma\gamma^2}\Delta_\eta^{-1}\left\{\boldsymbol{\Gamma}_{1\tau}^T\left(\mathbb{X} - \mathbf{1}_n\boldsymbol{\mu}_{\boldsymbol{X}}^T\right)\boldsymbol{D}^{-1}\left(\mathbb{W} - \mu_{\tau,Y}\mathbf{1}_n\right) + \boldsymbol{M}\boldsymbol{\Gamma}_{1\tau}^T\boldsymbol{e}\right\},\ \Delta_\eta^{-1}\right),
\end{aligned}
$$

where $\Delta_\eta = \frac{1}{\sigma\gamma^2}\left\{\boldsymbol{\Gamma}_{1\tau}^T\left(\mathbb{X} - \mathbf{1}_n\boldsymbol{\mu}_{\boldsymbol{X}}^T\right)^T \boldsymbol{D}^{-1}\left(\mathbb{X} - \mathbf{1}_n\boldsymbol{\mu}_{\boldsymbol{X}}^T\right)\boldsymbol{\Gamma}_{1\tau} + \boldsymbol{M}\right\}$.

4. For σ,

$$
\log\pi\left(\sigma|\text{rest}\right)
$$

$$= -\left(\frac{3n}{2} + \frac{u_\tau}{2} + a + 1\right) \log \sigma$$
$$- \frac{1}{\sigma}\left[b + \sum_{i=1}^{n} Z_i + \frac{1}{2\tau^2}\left\{\left(\mathbb{W} - \mu_{\tau,Y}\mathbf{1}_n - \left(\mathbb{X} - \mathbf{1}_n\boldsymbol{\mu}_{\boldsymbol{X}}^T\right)\boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})\boldsymbol{\eta}\right)^T\right.\right.$$
$$\boldsymbol{D}^{-1}\left(\mathbb{W} - \mu_{\tau,Y}\mathbf{1}_n - \left(\mathbb{X} - \mathbf{1}_n\boldsymbol{\mu}_{\boldsymbol{X}}^T\right)\boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})\boldsymbol{\eta}\right)$$
$$\left.\left. + \left(\boldsymbol{\eta} - \boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})^T\boldsymbol{e}\right)^T \boldsymbol{M}\left(\boldsymbol{\eta} - \boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})^T\boldsymbol{e}\right)\right\}\right].$$

So, we have

$$\sigma|\text{rest} \sim IG\left(\frac{3n}{2} + \frac{u_\tau}{2} + a, \quad \widehat{b}\right),$$

where

$$\widehat{b} = b + \sum_{i=1}^{n} Z_i + \frac{1}{2\tau^2}\left\{\left(\mathbb{W} - \mu_{\tau,Y}\mathbf{1}_n - \left(\mathbb{X} - \mathbf{1}_n\boldsymbol{\mu}_{\boldsymbol{X}}^T\right)\boldsymbol{\Gamma}_{1\tau}\boldsymbol{\eta}\right)^T \boldsymbol{D}^{-1}\right.$$
$$\left.\left(\mathbb{W} - \mu_{\tau,Y}\mathbf{1}_n - \left(\mathbb{X} - \mathbf{1}_n\boldsymbol{\mu}_{\boldsymbol{X}}^T\right)\boldsymbol{\Gamma}_{1\tau}\boldsymbol{\eta}\right) + \left(\boldsymbol{\eta} - \boldsymbol{\Gamma}_{1\tau}^T\boldsymbol{e}\right)^T \boldsymbol{M}\left(\boldsymbol{\eta} - \boldsymbol{\Gamma}_{1\tau}^T\boldsymbol{e}\right)\right\}$$

5. For $\boldsymbol{\Omega}_1$,

$$\boldsymbol{\Omega}_1|\text{rest} \sim IW_{u_\tau}\left(\boldsymbol{\Psi}_1 + \boldsymbol{\Gamma}_{1\tau}^T\left(\mathbb{X} - \mathbf{1}_n\boldsymbol{\mu}_{\boldsymbol{X}}^T\right)^T\left(\mathbb{X} - \mathbf{1}_n\boldsymbol{\mu}_{\boldsymbol{X}}^T\right)\boldsymbol{\Gamma}_{1\tau},\ \nu_1 + n\right).$$

6. For $\boldsymbol{\Omega}_2$,

$$\boldsymbol{\Omega}_2|\text{rest} \sim IW_{p-u_\tau}\left(\boldsymbol{\Psi}_2 + \boldsymbol{\Gamma}_{2\tau}^T\left(\mathbb{X} - \mathbf{1}_n\boldsymbol{\mu}_{\boldsymbol{X}}^T\right)^T\left(\mathbb{X} - \mathbf{1}_n\boldsymbol{\mu}_{\boldsymbol{X}}^T\right)\boldsymbol{\Gamma}_{2\tau},\ \nu_2 + n\right).$$

7. For $\boldsymbol{A}$,

$$
\begin{aligned}
&\log \pi\left(\boldsymbol{A} \mid \text{rest}\right) \equiv \log H(\boldsymbol{A}) \\
&= -\frac{1}{2\sigma\gamma^2}\left[\left\{\mathbb{W} - \mu_{\tau,Y}\mathbf{1}_n - \left(\mathbb{X} - \mathbf{1}_n\boldsymbol{\mu}_{\boldsymbol{X}}^T\right)\boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})\boldsymbol{\eta}\right\}^T\right. \\
&\qquad\left.\boldsymbol{D}^{-1}\left\{\mathbb{W} - \mu_{\tau,Y}\mathbf{1}_n - \left(\mathbb{X} - \mathbf{1}_n\boldsymbol{\mu}_{\boldsymbol{X}}^T\right)\boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})\boldsymbol{\eta}\right\}\right] \\
&- \frac{1}{2}\,\text{trace}\left\{\left(\mathbb{X} - \mathbf{1}_n\boldsymbol{\mu}_{\boldsymbol{X}}^T\right)\left(\boldsymbol{\Gamma}_{1\tau}\boldsymbol{\Omega}_1^{-1}\boldsymbol{\Gamma}_{1\tau}^T + \boldsymbol{\Gamma}_{2\tau}\boldsymbol{\Omega}_2^{-1}\boldsymbol{\Gamma}_{2\tau}^T\right)\left(\mathbb{X} - \mathbf{1}_n\boldsymbol{\mu}_{\boldsymbol{X}}^T\right)^T\right\} \\
&- \frac{1}{2\sigma\gamma^2}\left(\boldsymbol{\eta} - \boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})^T\boldsymbol{e}\right)^T\boldsymbol{M}\left(\boldsymbol{\eta} - \boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})^T\boldsymbol{e}\right) \\
&- \frac{1}{2}\,\text{trace}\left\{\boldsymbol{K}^{-1}\left(\boldsymbol{A} - \boldsymbol{A}_1\right)\boldsymbol{L}^{-1}\left(\boldsymbol{A} - \boldsymbol{A}_1\right)^T\right\}. \qquad \text{(S4.9)}
\end{aligned}
$$

S5 Algorithm for maximum a posteriori estimation

A conditional expectation maximization algorithm (Meng and Rubin, 1993) can be formulated based on the conditional distributions provided in Algorithm 1 for maximum a posteriori (MAP) estimation of the parameters in the BEQR. At each iteration, the algorithm would first impute the latent data Z_i's by its full conditional mean. Then given the imputed data Z_i, the algorithm would blockwise maximize the model parameters. The full conditional posterior distributions of Z_i's are independent generalized inverse Gaussian whose expectations can be written using modified Bessel functions of the second kind, and thus can be efficiently numerically evaluated.

The full conditional posterior densities of all model parameters except $\boldsymbol{A}$ are standard, and their (conditional) modes are available in closed forms. The full conditional distribution of $\boldsymbol{A}$, although not standard, is smooth, and hence can be maximized via a gradient based optimization technique such as Newton's method.

Algorithm S5.1. Computation of the MAP estimators of the BEQR parameters $\{\mu_{\tau,Y}, \boldsymbol{\mu}_{\boldsymbol{X}}, \boldsymbol{\eta}, \boldsymbol{\Omega}_1, \boldsymbol{\Omega}_2, \sigma, \boldsymbol{A}\}$.

Iterate between the following steps until convergence.

Step-1 For $i = 1, \ldots, n$, impute the latent variables

$$\widehat{Z}_i = \frac{\sqrt{m_1} K_{3/2}\left(\sqrt{m_1 m_2}\right)}{\sqrt{m_2} K_{1/2}\left(\sqrt{m_1 m_2}\right)}$$

where $K_{\cdot}(\cdot)$ is a modified Bessel function,

$$m_1 = \frac{(\theta^2 + 2\gamma^2)}{\widehat{\sigma}\gamma^2}, \quad \text{and} \quad m_2 = \frac{\left\{Y_i - \widehat{\mu}_{\tau,Y} - \widehat{\boldsymbol{\eta}}^T \boldsymbol{\Gamma}_{1\tau}(\widehat{\boldsymbol{A}})^T \left(\boldsymbol{X}_i - \widehat{\boldsymbol{\mu}}_{\boldsymbol{X}}\right)\right\}^2}{\widehat{\sigma}\gamma^2}.$$

Also set $\widehat{W}_i = Y_i - \theta \widehat{Z}_i$.

Step-2 Calculate

$$\widehat{\boldsymbol{A}} = \underset{\mathbb{R}^{(p-u_\tau)\times u_\tau}}{\arg\max} \left\{ -\frac{1}{2}\,\mathrm{trace}\left[\boldsymbol{K}^{-1}\left(\boldsymbol{A}-\boldsymbol{A}_0\right)\boldsymbol{L}^{-1}\left(\boldsymbol{A}-\boldsymbol{A}_0\right)^T\right] \right.$$
$$-\frac{\nu_1+n+u_\tau+1}{2}\log\left|\boldsymbol{\Psi}_1+\boldsymbol{\Gamma}_{1\tau}^T\left(\mathbb{X}-\mathbf{1}_n\boldsymbol{\mu}_{\boldsymbol{X}}^T\right)^T\left(\mathbb{X}-\mathbf{1}_n\boldsymbol{\mu}_{\boldsymbol{X}}^T\right)\boldsymbol{\Gamma}_{1\tau}\right|$$
$$\left. -\frac{\nu_2+n+p-u_\tau+1}{2}\log\left|\boldsymbol{\Psi}_2+\boldsymbol{\Gamma}_{2\tau}^T\left(\mathbb{X}-\mathbf{1}_n\boldsymbol{\mu}_{\boldsymbol{X}}^T\right)^T\left(\mathbb{X}-\mathbf{1}_n\boldsymbol{\mu}_{\boldsymbol{X}}^T\right)\boldsymbol{\Gamma}_{2\tau}\right| \right\},$$

and update $\boldsymbol{C}_{\boldsymbol{A}}$, $\boldsymbol{D}_{\boldsymbol{A}}$, $\boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})$, and $\boldsymbol{\Gamma}_{2\tau}(\boldsymbol{A})$ as follows

$$\boldsymbol{C}_{\widehat{\boldsymbol{A}}} = \begin{pmatrix} \boldsymbol{I}_{u_\tau} \\ \widehat{\boldsymbol{A}} \end{pmatrix},\ \widehat{\boldsymbol{D}}_{\boldsymbol{A}} = \begin{pmatrix} -\widehat{\boldsymbol{A}}^T \\ \boldsymbol{I}_{p-u_\tau} \end{pmatrix},$$
$$\boldsymbol{\Gamma}_{1\tau}(\widehat{\boldsymbol{A}}) = \boldsymbol{C}_{\widehat{\boldsymbol{A}}}\left(\boldsymbol{C}_{\widehat{\boldsymbol{A}}}^T\boldsymbol{C}_{\widehat{\boldsymbol{A}}}\right)^{-1/2},\ \text{and}\ \boldsymbol{\Gamma}_{2\tau}(\widehat{\boldsymbol{A}}) = \widehat{\boldsymbol{D}}_{\boldsymbol{A}}\left(\widehat{\boldsymbol{D}}_{\boldsymbol{A}}^T\widehat{\boldsymbol{D}}_{\boldsymbol{A}}\right)^{-1/2}.$$

Step-3 Compute

(i) $\widehat{\mu}_{\tau,Y} = \overline{W}_Z + \widehat{\boldsymbol{\eta}}^T\boldsymbol{\Gamma}_{1\tau}(\widehat{\boldsymbol{A}})^T\left(\widehat{\boldsymbol{\mu}}_{\boldsymbol{X}} - \overline{\boldsymbol{X}}_Z\right)$.

(ii) $\widehat{\boldsymbol{\mu}}_{\boldsymbol{X}} = \Delta_{\boldsymbol{\mu}_{\boldsymbol{X}}}^{-1}\Xi_{\boldsymbol{\mu}_{\boldsymbol{X}}}$ where

$$\Xi_{\boldsymbol{\mu}_{\boldsymbol{X}}} = \frac{1}{\widehat{\sigma}\gamma^2}\left(\sum_{i=1}^n\frac{1}{\widehat{Z}_i}\right)\boldsymbol{\Gamma}_{1\tau}(\widehat{\boldsymbol{A}})\widehat{\boldsymbol{\eta}}\left(\widehat{\boldsymbol{\eta}}^T\boldsymbol{\Gamma}_{1\tau}^T(\widehat{\boldsymbol{A}})\overline{\boldsymbol{X}}_Z + \widehat{\mu}_{\tau,Y} - \overline{W}_Z\right)$$
$$+\,n\left(\boldsymbol{\Gamma}_{1\tau}(\widehat{\boldsymbol{A}})\widehat{\boldsymbol{\Omega}}_1^{-1}\boldsymbol{\Gamma}_{1\tau}^T(\widehat{\boldsymbol{A}}) + \boldsymbol{\Gamma}_{2\tau}(\widehat{\boldsymbol{A}})\widehat{\boldsymbol{\Omega}}_2^{-1}\boldsymbol{\Gamma}_{2\tau}^T(\widehat{\boldsymbol{A}})\right)\overline{\boldsymbol{X}}$$
$$\Delta_{\boldsymbol{\mu}_{\boldsymbol{X}}} = \frac{1}{\widehat{\sigma}\gamma^2}\left(\sum_{i=1}^n\frac{1}{\widehat{Z}_i}\right)\boldsymbol{\Gamma}_{1\tau}(\widehat{\boldsymbol{A}})\widehat{\boldsymbol{\eta}}\widehat{\boldsymbol{\eta}}^T\boldsymbol{\Gamma}_{1\tau}^T(\widehat{\boldsymbol{A}})$$
$$+\,n\left(\boldsymbol{\Gamma}_{1\tau}(\widehat{\boldsymbol{A}})\widehat{\boldsymbol{\Omega}}_1^{-1}\boldsymbol{\Gamma}_{1\tau}^T(\widehat{\boldsymbol{A}}) + \boldsymbol{\Gamma}_{2\tau}(\widehat{\boldsymbol{A}})\widehat{\boldsymbol{\Omega}}_2^{-1}\boldsymbol{\Gamma}_{2\tau}^T(\widehat{\boldsymbol{A}})\right).$$

(iii) $\widehat{\boldsymbol{\eta}} = \frac{1}{\widehat{\sigma}\gamma^2}\Delta_\eta^{-1}\left\{\widehat{\boldsymbol{\Gamma}}_{1\tau}^T\left(\mathbb{X} - \mathbf{1}_n\widehat{\boldsymbol{\mu}}_{\boldsymbol{X}}^T\right)\boldsymbol{D}^{-1}\left(\widehat{\mathbb{W}} - \widehat{\mu}_{\tau,Y}\mathbf{1}_n\right) + \boldsymbol{M}\widehat{\boldsymbol{\Gamma}}_{1\tau}^T\boldsymbol{e}\right\}$,

where

$$\Delta_\eta = \frac{1}{\widehat{\sigma}\gamma^2}\left\{\boldsymbol{\Gamma}_{1\tau}^T(\widehat{\boldsymbol{A}})\left(\mathbb{X} - \mathbf{1}_n\widehat{\boldsymbol{\mu}}_{\boldsymbol{X}}^T\right)^T\boldsymbol{D}^{-1}\left(\mathbb{X} - \mathbf{1}_n\widehat{\boldsymbol{\mu}}_{\boldsymbol{X}}^T\right)\boldsymbol{\Gamma}_{1\tau}(\widehat{\boldsymbol{A}}) + \boldsymbol{M}\right\}.$$

(iv) $\widehat{\sigma} = \widetilde{b}/(\widetilde{a} - 1)$, where $\widetilde{a} = (3n)/2 + u_\tau/2 + a$ and

$$\begin{aligned}\widetilde{b} = b + \sum_{i=1}^n \widehat{Z}_i + \frac{1}{2\gamma^2}&\left\{\left(\widehat{\mathbb{W}} - \widehat{\mu}_{\tau,Y}\mathbf{1}_n - \widehat{\boldsymbol{\eta}}^T\boldsymbol{\Gamma}_{1\tau}(\widehat{\boldsymbol{A}})^T\left(\mathbb{X} - \mathbf{1}_n\widehat{\boldsymbol{\mu}}_{\boldsymbol{X}}\right)\right)^T\boldsymbol{D}^{-1}\right.\\ &\boldsymbol{D}^{-1}\left(\widehat{\mathbb{W}} - \widehat{\mu}_{\tau,Y}\mathbf{1}_n - \widehat{\boldsymbol{\eta}}^T\boldsymbol{\Gamma}_{1\tau}(\widehat{\boldsymbol{A}})^T\left(\mathbb{X} - \mathbf{1}_n\widehat{\boldsymbol{\mu}}_{\boldsymbol{X}}\right)\right)\\ &\left.+\left(\widehat{\boldsymbol{\eta}} - \boldsymbol{\Gamma}_{1\tau}^T(\widehat{\boldsymbol{A}})\boldsymbol{e}\right)^T\boldsymbol{M}\left(\widehat{\boldsymbol{\eta}} - \boldsymbol{\Gamma}_{1\tau}^T(\widehat{\boldsymbol{A}})\boldsymbol{e}\right)\right\}.\end{aligned}$$

(v) $\widehat{\boldsymbol{\Omega}}_1 = \left\{\boldsymbol{\Psi}_1 + \boldsymbol{\Gamma}_{1\tau}^T(\widehat{\boldsymbol{A}})\left(\mathbb{X} - \mathbf{1}_n\widehat{\boldsymbol{\mu}}_{\boldsymbol{X}}^T\right)^T\left(\mathbb{X} - \mathbf{1}_n\widehat{\boldsymbol{\mu}}_{\boldsymbol{X}}^T\right)\boldsymbol{\Gamma}_{1\tau}(\widehat{\boldsymbol{A}})\right\}/(\nu_1 + n + u_\tau + 1)$.

(vi) $\widehat{\boldsymbol{\Omega}}_2 = \left\{\boldsymbol{\Psi}_2 + \boldsymbol{\Gamma}_{2\tau}^T(\widehat{\boldsymbol{A}})\left(\mathbb{X} - \mathbf{1}_n\widehat{\boldsymbol{\mu}}_{\boldsymbol{X}}^T\right)^T\left(\mathbb{X} - \mathbf{1}_n\widehat{\boldsymbol{\mu}}_{\boldsymbol{X}}^T\right)\boldsymbol{\Gamma}_{2\tau}(\widehat{\boldsymbol{A}})\right\}/(\nu_2 + n + p - u_\tau + 1)$.

S6 MCMC sampler for sampling from the posterior of the BEQR with Tobit censored responses

Algorithm S6.1. One iteration of the data augmentation Metropolis-within-Gibbs sampler.

Step 1 (Censored Data Imputation). For each censored Y_i (i.e., $Y_i = 0$) draw

$$Y_i^* \sim \mathrm{TN}_{(-\infty,0]}\left(\mu_{\tau,Y} + \boldsymbol{\eta}^T \boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})^T \left(\boldsymbol{X}_i - \boldsymbol{\mu}_{\boldsymbol{X}}\right) + \theta Z_i,\ Z_i \sigma\gamma^2\right),$$

and for the remaining i's, set $Y_i^* = Y_i$. Here $\mathrm{TN}_A(\mu, \sigma^2)$ denotes a truncated normal (μ, σ^2) distribution which is truncated to lie on a set $A \subseteq (-\infty, \infty)$.

Step 2 (Data Augmentation). Generate independent $Z_1, \ldots, Z_n$ with

$$Z_i \sim \mathrm{GIG}\left(\frac{\left\{Y_i^* - \mu_{\tau,Y} - \boldsymbol{\eta}^T \boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})^T \left(\boldsymbol{X}_i - \boldsymbol{\mu}_{\boldsymbol{X}}\right)\right\}^2}{\sigma\gamma^2}, \frac{\theta^2 + 2\gamma^2}{\sigma\gamma^2}, \frac{1}{2}\right).$$

Then update $W_i = Y_i^* - \theta Z_i$ for $i = 1, \ldots, n$.

Step 3 Generate $\mu_{\tau,Y} \sim \mathrm{N}\left(\overline{W}_Z + \boldsymbol{\eta}^T \boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})^T \left(\boldsymbol{\mu}_{\boldsymbol{X}} - \overline{\boldsymbol{X}}_Z\right),\ \frac{1}{\sum_{i=1}^n \frac{1}{Z_i}} \sigma\gamma^2\right)$, where

$$\overline{W}_Z = \frac{1}{\sum_{i=1}^n \frac{1}{Z_i}} \sum_{i=1}^n \frac{W_i}{Z_i}, \quad \overline{\boldsymbol{X}}_Z = \frac{1}{\sum_{i=1}^n \frac{1}{Z_i}} \sum_{i=1}^n \frac{1}{Z_i} \boldsymbol{X}_i.$$

Step 4 Generate $\boldsymbol{\mu}_{\boldsymbol{X}} \sim \mathrm{N}\left(\Delta_{\boldsymbol{\mu}_{\boldsymbol{X}}}^{-1}\Xi_{\boldsymbol{\mu}_{\boldsymbol{X}}},\ \Delta_{\boldsymbol{\mu}_{\boldsymbol{X}}}^{-1}\right)$, where $\overline{\boldsymbol{X}} = \mathbf{1}_n^T\mathbb{X}/n$,

$$\begin{aligned}\Xi_{\boldsymbol{\mu}_{\boldsymbol{X}}} &= \frac{1}{\sigma\gamma^2}\left(\sum_{i=1}^{n}\frac{1}{Z_i}\right)\boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})\boldsymbol{\eta}\left(\boldsymbol{\eta}^T\boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})^T\overline{\boldsymbol{X}}_Z + \mu_{\tau,Y} - \overline{W}_Z\right)\\ &\quad + n\left(\boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})\boldsymbol{\Omega}_1^{-1}\boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})^T + \boldsymbol{\Gamma}_{2\tau}(\boldsymbol{A})\boldsymbol{\Omega}_2^{-1}\boldsymbol{\Gamma}_{2\tau}^T(\boldsymbol{A})\right)\overline{\boldsymbol{X}},\\ \Delta_{\boldsymbol{\mu}_{\boldsymbol{X}}} &= \frac{1}{\sigma\gamma^2}\left(\sum_{i=1}^{n}\frac{1}{Z_i}\right)\boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})\boldsymbol{\eta}\boldsymbol{\eta}^T\boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})^T\\ &\quad + n\left(\boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})\boldsymbol{\Omega}_1^{-1}\boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})^T + \boldsymbol{\Gamma}_{2\tau}(\boldsymbol{A})\boldsymbol{\Omega}_2^{-1}\boldsymbol{\Gamma}_{2\tau}^T(\boldsymbol{A})\right).\end{aligned}$$

Step 5 Generate $\boldsymbol{\eta} \sim \mathrm{N}\left(\widetilde{\boldsymbol{\eta}}_0,\ \Delta_{\boldsymbol{\eta}}^{-1}\right)$, where

$$\begin{aligned}\Delta_{\boldsymbol{\eta}} &= \frac{1}{\sigma\gamma^2}\left\{\boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})^T\left(\mathbb{X} - \mathbf{1}_n\boldsymbol{\mu}_{\boldsymbol{X}}^T\right)^T\boldsymbol{D}^{-1}\left(\mathbb{X} - \mathbf{1}_n\boldsymbol{\mu}_{\boldsymbol{X}}^T\right)\boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A}) + \boldsymbol{M}\right\},\\ \widetilde{\boldsymbol{\eta}}_0 &= \frac{1}{\sigma\gamma^2}\Delta_{\boldsymbol{\eta}}^{-1}\left\{\boldsymbol{\Gamma}_{1\tau}^T\left(\mathbb{X} - \mathbf{1}_n\boldsymbol{\mu}_{\boldsymbol{X}}^T\right)\boldsymbol{D}^{-1}\left(\mathbb{W} - \mu_{\tau,Y}\mathbf{1}_n\right) + \boldsymbol{M}\boldsymbol{\Gamma}_{1\tau}^T\boldsymbol{e}\right\}.\end{aligned}$$

Step 6 Generate $\sigma \sim \mathrm{IG}\left(\frac{3n}{2} + \frac{u_\tau}{2} + a,\ \widetilde{b}\right)$, where

$$\begin{aligned}\widetilde{b} = b + \sum_{i=1}^{n} Z_i + \frac{1}{2\gamma^2}&\left\{\left(\mathbb{W} - \mu_{\tau,Y}\mathbf{1}_n - (\mathbb{X} - \mathbf{1}_n\boldsymbol{\mu}_{\boldsymbol{X}}^T)\boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})\boldsymbol{\eta}\right)^T\boldsymbol{D}^{-1}\right.\\ &\left.\left(\mathbb{W} - \mu_{\tau,Y}\mathbf{1}_n - (\mathbb{X} - \mathbf{1}_n\boldsymbol{\mu}_{\boldsymbol{X}}^T)\boldsymbol{\Gamma}_{1\tau}\boldsymbol{\eta}\right) + \left(\boldsymbol{\eta} - \boldsymbol{\Gamma}_{1\tau}^T\boldsymbol{e}\right)^T\boldsymbol{M}\left(\boldsymbol{\eta} - \boldsymbol{\Gamma}_{1\tau}^T\boldsymbol{e}\right)\right\}.\end{aligned}$$

Step 7 Generate

$$\boldsymbol{\Omega}_1 \sim \mathrm{IW}_{u_\tau}\left(\boldsymbol{\Psi}_1 + \boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})^T(\mathbb{X} - \mathbf{1}_n\boldsymbol{\mu}_{\boldsymbol{X}}^T)^T(\mathbb{X} - \mathbf{1}_n\boldsymbol{\mu}_{\boldsymbol{X}}^T)\boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A}),\ \nu_1 + n\right).$$

Step 8 Generate

$$\boldsymbol{\Omega}_2 \sim \mathrm{IW}_{p-u_\tau}\left(\boldsymbol{\Psi}_2 + \boldsymbol{\Gamma}_{2\tau}^T(\boldsymbol{A})(\mathbb{X} - \mathbf{1}_n\boldsymbol{\mu}_{\boldsymbol{X}}^T)^T(\mathbb{X} - \mathbf{1}_n\boldsymbol{\mu}_{\boldsymbol{X}}^T)\boldsymbol{\Gamma}_{2\tau}(\boldsymbol{A}),\ \nu_2 + n\right).$$

Step 9 Generate a Markov chain realization for $\boldsymbol{A}$ with stationary density proportional to $H(\boldsymbol{A})$, the full conditional posterior density of $\boldsymbol{A}$ (see (S4.9), with Y_i replaced by Y_i^*).

Let $\boldsymbol{a}_j \in \mathbb{R}^{p-u_\tau}$ denote the j-th column of $\boldsymbol{A}$, $j = 1, \ldots, u_\tau$. Given the tuning parameter $\xi > 0$, for $j = i_1, \ldots, i_u$, where $\{i_1, \ldots, i_{u_\tau}\}$ denotes a random permutation of $\{1, \ldots, u_\tau\}$, perform the following:

(a) Generate $\boldsymbol{a}_j^* \sim \mathrm{N}_{p-u_\tau}\left(\boldsymbol{a}_j,\ \xi^2\boldsymbol{I}_{p-u_\tau}\right)$. Replace the j-th column of $\boldsymbol{A}$ by $\boldsymbol{a}_j^*$ and denote $\boldsymbol{A}^*$ the resulting matrix. Compute $\rho\left(\boldsymbol{A},\ \boldsymbol{A}^*\right) = \exp\left[H(\boldsymbol{A}^*) - H(\boldsymbol{A})\right]$.

(b) Perform a Bernoulli experiment with probability of success $\min(1, \rho\left(\boldsymbol{A},\ \boldsymbol{A}^*\right))$. If a success is obtained, update $\boldsymbol{a}_j^*$ to $\boldsymbol{a}_j$; otherwise retain $\boldsymbol{a}_j$.

(c) After updating $\boldsymbol{A}$, update $\boldsymbol{C}_{\boldsymbol{A}}$, $\boldsymbol{D}_{\boldsymbol{A}}$ and $\boldsymbol{\Sigma}_{\boldsymbol{X}}$.

Remark 1. Algorithm S6.1 can account for the two degenerated cases $u_\tau = 0$ and $u_\tau = p$ as follows: when $u_\tau = 0$, $\boldsymbol{A}$ does not exist and we have $\boldsymbol{\eta} = 0$, $\boldsymbol{\Gamma}_{2\tau}(\boldsymbol{A}) = \boldsymbol{I}_p$ and $\boldsymbol{\Sigma}_{\boldsymbol{X}} = \boldsymbol{\Omega}_2$. Thus the steps involving $\boldsymbol{\eta}$, $\boldsymbol{\Omega}_1$ and $\boldsymbol{A}$ (*Step 5*, *Step 7* and *Step 9* respectively) are not required. On the other hand, when $u_\tau = p$, the BETQR reduces to the BTQR: $\boldsymbol{A}$ does not exist, $\boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A}) = \boldsymbol{I}_p$ and $\boldsymbol{\Sigma}_{\boldsymbol{X}} = \boldsymbol{\Omega}_1$, and the steps involving $\boldsymbol{A}$ and $\boldsymbol{\Omega}_2$ (*Step 8* and

Step 9) are not needed. In each case, Algorithm 1 becomes a full Gibbs sampler.

S7 Additional Simulations and Data Analysis

S7.1 Comparison between the BEQR and `bayesQR` estimators

The estimation variance and MSE of the `bayesQR` estimator are also calculated in the same way as for the BQR or BEQR estimators. Table 1 summaries the ratios of the estimation variance and MSE of the `bayesQR` estimator versus the BEQR estimator for each element in $\boldsymbol{\beta}_\tau$. Compared with Table 1 of the main text, we notice that the ratios are similar, which indicates that `bayesQR` estimator has about the same estimation variance and MSE as the BQR estimators. The BEQR is also more efficient in estimation compared to either the BQR estimator or the `bayesQR` estimator. For illustration, the comparison between the BEQR and `bayesQR` estimators on estimation variance and MSE of the second element in $\boldsymbol{\beta}_\tau$ is displayed in Figure 1.

	$\tau = 0.1$		$\tau = 0.5$	
n	Ratio$_V$	Ratio$_M$	Ratio$_V$	Ratio$_M$
50	2.11 (1.82, 2.92)	3.40 (2.07, 8.34)	2.07 (1.83, 3.23)	3.07 (1.91, 5.64)
100	2.08 (1.67, 2.57)	2.76 (1.97, 7.48)	1.96 (1.35, 2.77)	2.29 (1.60, 4.67)
200	2.02 (1.73, 2.43)	2.47 (1.67, 4.13)	1.90 (1.56, 2.41)	2.04 (1.65, 2.94)
400	1.91 (1.64, 3.02)	2.05 (1.60, 3.40)	1.95 (1.47, 2.65)	1.96 (1.67, 2.80)
800	1.98 (1.79, 3.01)	2.09 (1.75, 2.98)	1.73 (1.51, 2.05)	1.78 (1.57, 2.18)

Table 1: Medians (ranges) of the estimation variance and MSE ratios. Ratio$_V$: Estimation variance ratio of the `bayesQR` estimator versus the BEQR estimator. Ratio$_M$: Mean square error ratio of the `bayesQR` estimator versus the BEQR estimator.

S7.2 Comparison of the frequentist and Bayesian envelope quantile regression estimators

We first adopt the same simulation setting as in Section 5.1 of the main article and generated the data from model (5.1). Then we computed the frequentist envelope quantile regression estimator and Bayesian envelope quantile regression estimator for each of the 200 repetitions. Table 2 summarizes the estimation variances and squared biases for an element in $\boldsymbol{\beta}_\tau$ with $\tau = 0.1$ and 0.5.

Table 2: The estimation variance and squared bias for an element in $\boldsymbol{\beta}_\tau$ with $\tau = 0.1$ and $\tau = 0.5$. The subscript "F" and "B" represent the frequentist's envelope estimator and Bayesian envelope estimator with true dimension, respectively.

	$\tau = 0.1$				$\tau = 0.5$			
n	Var$_B$	Var$_F$	Bias$_B^2$	Bias$_F^2$	Var$_B$	Var$_F$	Bias$_B^2$	Bias$_F^2$
50	0.147	0.079	6.4×10^{-4}	4.1×10^{-3}	0.082	0.067	2.7×10^{-4}	5.8×10^{-4}
100	0.064	0.035	3.2×10^{-5}	2.2×10^{-3}	0.031	0.031	9.5×10^{-5}	6.4×10^{-4}
200	0.033	0.017	9.1×10^{-7}	1.4×10^{-3}	0.021	0.017	4.8×10^{-4}	4.5×10^{-5}
400	0.015	0.008	1.9×10^{-4}	1.8×10^{-3}	0.010	0.008	2.0×10^{-5}	9.6×10^{-5}
800	0.007	0.004	9.4×10^{-5}	1.6×10^{-3}	0.004	0.004	7.7×10^{-6}	5.9×10^{-5}

For both quantile levels, the Bayesian estimates have larger estimation

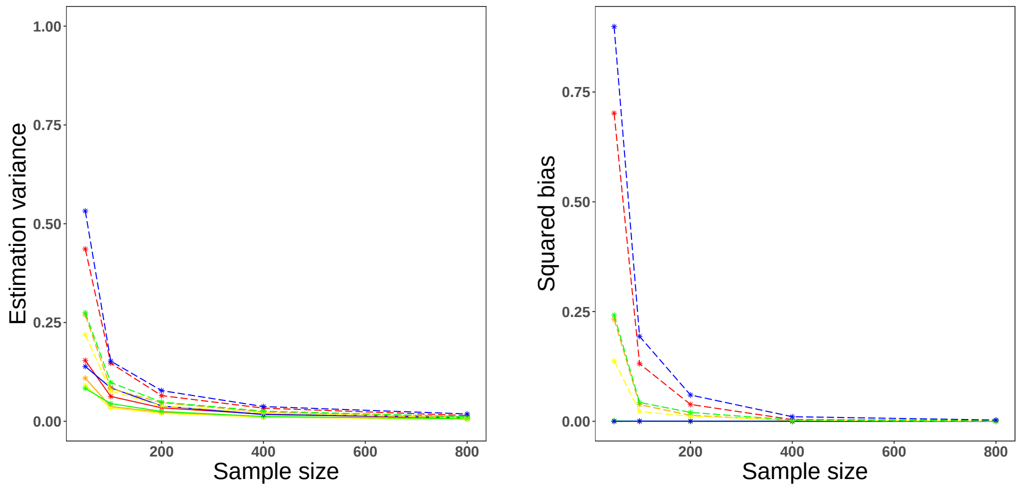

Figure 1: Estimation variances and squared biases of the second element in the BEQR estimator (solid line) and the `bayesQR` estimator (dashed line) with $\tau = 0.1$ (red line), 0.25 (orange line), 0.5 (yellow line), 0.75 (green line), and 0.9 (blue line).

variances and larger MSE than the frequentist estimates. This is caused by model misspecification. The Bayesian envelope quantile regression (3.3) assumes that the error σ_ϵ follows the asymmetric Laplace distribution, while in the data generation (5.1), the error σ_ϵ follows the normal distribution. To address this issue, we now keep all the parameter settings to be the same except that the errors are generated from the the asymmetric Laplace distribution. In other words, the data is generated from the following model: For $i = 1, \ldots, n$,

$$Y_i = \mu_Y + \boldsymbol{\eta}^T \boldsymbol{\Gamma}_{1\tau}^T(\boldsymbol{A})\boldsymbol{X}_i + \epsilon_{i\tau};$$

$$\boldsymbol{X}_i \sim N_p\left(\boldsymbol{\mu}_{\boldsymbol{X}},\ \boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})\boldsymbol{\Omega}_1\boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})^T + \boldsymbol{\Gamma}_{2\tau}(\boldsymbol{A})\boldsymbol{\Omega}_2\boldsymbol{\Gamma}_{2\tau}(\boldsymbol{A})^T\right), \qquad \text{(S7.10)}$$

where the error $\epsilon_{i\tau}$ follows an asymmetric Laplace distribution with density

$$g_{ALD}(\epsilon;\tau) = \tau(1-\tau)\left[e^{(1-\tau)\epsilon} I(\epsilon < 0) + e^{-\tau\epsilon} I(\epsilon > 0)\right].$$

Under this setting, the estimation variances and squared biases for an element in $\boldsymbol{\beta}_\tau$ with $\tau = 0.1$ and 0.5 are displayed in Table 3. The Bayesian envelope estimators have smaller estimation variances and MSE than the frequentist envelope estimators in almost all the cases.

Table 3: The estimation variance and squared bias for an element in $\boldsymbol{\beta}_\tau$ with $\tau = 0.1$ and $\tau = 0.5$. The subscript F and B represent the frequentist's envelope estimator and Bayesian envelope estimator with true dimension, respectively.

	$\tau = 0.1$				$\tau = 0.5$			
n	Var_B	Var_F	Bias^2_B	Bias^2_F	Var_B	Var_F	Bias^2_B	Bias^2_F
50	0.109	0.101	3.0×10^{-5}	3.7×10^{-3}	0.041	0.099	1.9×10^{-3}	4.2×10^{-3}
100	0.041	0.047	4.9×10^{-4}	3.3×10^{-4}	0.017	0.046	1.7×10^{-4}	4.0×10^{-4}
200	0.017	0.024	1.2×10^{-5}	1.4×10^{-7}	0.009	0.023	1.2×10^{-8}	7.9×10^{-6}
400	0.007	0.009	4.0×10^{-5}	4.1×10^{-5}	0.003	0.009	9.2×10^{-6}	2.6×10^{-5}
800	0.004	0.006	7.3×10^{-7}	6.5×10^{-6}	0.002	0.005	1.2×10^{-5}	3.3×10^{-6}

S7.3 Performance of the Bayesian envelope quantile regression estimator under skewed distribution

To assess our approach under skewed error distributions, we generated the data from the same model as (5.1) in the main text, except that ϵ is now generated from a skewed normal distribution with a location parameter 0, a scale parameter 1 and a shape parameter 3 (Azzalini, 2013) instead of the normal distribution. To keep the noise level (i.e., the variance of ϵ) the same as the normal case, we standardized ϵ. The generation of all the other parameters are kept the same.

The performance of LOOIC in dimension selection was reported in Table 4. Comparing the results in Table 4 and those under the normal errors (Table 2 of the main text), it seems that the performance of LOOIC is not affected by the skewness of the distribution. We further computed the ratios of the estimation variance and mean squared error (MSE) of the BQR estimator versus the BEQR estimator, and summarized the ratios in Table 5. Again these ratios are similar to those under the normal errors (Table 1 of the main text), which indicates that the efficiency gains are stable without the normality assumption. It is also observed in the frequentist partial envelope model (Su and Cook, 2011) that the efficiency gains obtained by the envelope approach are robust to the skewness of the error distribution.

Now we investigate the performance of the BEQR based on a skewed distribution of $\boldsymbol{X}$. For fair comparison, we keep the signal level (i.e., the covariance matrix of $\boldsymbol{X}$) the same. Then $\boldsymbol{X}$ is generated as $\boldsymbol{\Sigma}_X^{1/2}\boldsymbol{Z}$, where each element in $\boldsymbol{Z} \in \mathbb{R}^p$ is simulated from skewed normal distribution with a

Table 4: Number of replications (out of 200) for which a given value of u_τ is selected with a skewed distribution on ϵ.

	$\tau = 0.1$					$\tau = 0.5$				
Selected u_τ	1	2	3	4	5	1	2	3	4	5
50	3	102	55	37	3	0	142	29	28	1
100	3	143	35	18	1	0	139	35	25	1
200	0	165	25	10	0	0	145	31	21	1
400	0	173	13	14	0	0	160	18	20	2
800	0	167	15	18	0	0	142	29	28	1

Table 5: Medians (ranges) of the estimation variance and MSE ratios with a skewed distribution on ϵ. Ratio_V: Estimation variance ratio of the BQR estimator versus the BEQR estimator. Ratio_M: MSE ratio of the BQR estimator versus the BEQR estimator.

	$\tau = 0.1$		$\tau = 0.5$	
n	Ratio_V	Ratio_M	Ratio_V	Ratio_M
50	1.59 (1.40, 2.43)	1.59 (1.40, 2.43)	1.73 (1.54, 2.40)	1.74 (1.49, 2.57)
100	1.81 (1.53, 3.36)	1.82 (1.55, 3.38)	1.94 (1.54, 3.11)	1.94 (1.56, 3.09)
200	1.75 (1.49, 2.80)	1.74 (1.51, 2.58)	1.89 (1.58, 2.55)	1.94 (1.57, 2.53)
400	1.89 (1.72, 3.57)	1.89 (1.72, 3.52)	1.92 (1.35, 2.82)	1.97 (1.37, 2.75)
800	1.87 (1.52, 2.96)	1.87 (1.53, 2.95)	1.73 (1.54, 2.40)	1.74 (1.49, 2.57)

location parameter 0, a scale parameter 2 and a shape parameter 3, and then standardized to have variance 1. Generation of other parameters remain the same as the previous setting, except that the error ϵ_i was generated from a standard normal distribution as in the main text.

The dimension selection results of LOOIC were summarized in Table 6 and ratios of the estimation variance and MSE of the BQR estimator versus the BEQR estimator were displayed in Table 7. Compared to the case where $\boldsymbol{X}$ was normally distributed, LOOIC tends to overestimate u_τ for both quantile levels. In contrast to underestimation, overestimation is a smaller issue since it does not provide bias. The only issue with overestimation is that we may not achieve as much efficiency gains as we should have. However, Table 7 shows that the envelope approach still yields substantial efficiency gains, and the gains are similarly to the case where $\boldsymbol{X}$ is normally distributed, which shows that overestimation issue does not cause the loss of much efficiency gains in this case.

Table 6: Number of replications (out of 200) for which a given value of u_τ is selected with a skewed distribution on $\boldsymbol{X}$.

	$\tau = 0.1$					$\tau = 0.5$				
Selected u_τ	1	2	3	4	5	1	2	3	4	5
50	3	77	91	29	0	2	92	80	26	0
100	13	70	98	16	3	1	79	93	24	3
200	12	70	104	14	0	2	90	94	14	0
400	1	68	113	16	2	1	94	82	21	2
800	0	86	101	11	2	0	102	81	16	1

Table 7: Medians (ranges) of the estimation variance and MSE ratios with a skewed distribution on $\boldsymbol{X}$. Ratio$_V$: Estimation variance ratio of the BQR estimator versus the BEQR estimator. Ratio$_M$: MSE ratio of the BQR estimator versus the BEQR estimator.

	$\tau = 0.1$		$\tau = 0.5$	
n	Ratio$_V$	Ratio$_M$	Ratio$_V$	Ratio$_M$
50	1.56 (1.24, 3.73)	1.54 (1.24, 3.54)	2.23 (1.56, 4.27)	2.21 (1.56, 4.16)
100	1.83 (1.36, 2.69)	1.84 (1.36, 2.65)	1.95 (1.43, 3.70)	1.92 (1.44, 3.73)
200	1.85 (1.38, 3.86)	1.80 (1.36, 3.84)	2.28 (1.59, 4.59)	2.27 (1.59, 4.59)
400	2.20 (1.42, 3.85)	2.20 (1.44, 3.73)	2.76 (1.81, 4.87)	2.75 (1.81, 4.86)
800	2.36 (1.76, 4.27)	2.34 (1.73, 4.07)	2.47 (1.80, 3.71)	2.45 (1.84, 3.72)

Finally, we examined the performance of the BEQR with a skewed distribution for both $\boldsymbol{X}$ and ϵ. The error ϵ was generated from a skewed normal distribution in the same way that produced Tables 4 and 5. And $\boldsymbol{X}$ was generated in the same way that produced Tables 6 and 7. The generation of the other parameters was kept the same.

The performance of LOOIC for $\tau = 0.1$ and 0.5 was recorded in Table 8 and the ratios of the estimation variance and MSE of the BQR estimator versus the BEQR estimator were summarized in Table 9. Note that as in the last setting, LOOIC tends to overestimate u_τ. However, the envelope approach still achieves efficiency gains as shown in Table 9, and the efficiency gains are about the same when both ϵ and $\boldsymbol{X}$ are normally distributed.

Table 8: Number of replications (out of 200) for which a given value of u_τ is selected under skewed distributions for both ϵ and $\boldsymbol{X}$.

	$\tau = 0.1$					$\tau = 0.5$				
Selected u_τ	1	2	3	4	5	1	2	3	4	5
50	8	70	100	21	1	5	99	77	13	6
100	9	96	77	17	1	7	80	87	25	1
200	6	103	79	12	0	3	97	80	20	0
400	3	114	70	13	0	0	100	80	17	3
800	0	122	61	17	0	0	97	79	23	1

Table 9: Medians (ranges) of the estimation variance and MSE ratios under skewed distributions for both ϵ and $\boldsymbol{X}$. Ratio_V: Estimation variance ratio of the BQR estimator versus the BEQR estimator. Ratio_M: MSE ratio of the BQR estimator versus the BEQR estimator.

	$\tau = 0.1$		$\tau = 0.5$	
n	Ratio_V	Ratio_M	Ratio_V	Ratio_M
50	1.78 (1.53, 3.43)	1.79 (1.53, 3.34)	2.04 (1.58, 3.71)	2.03 (1.57, 3.60)
100	1.85 (1.63, 3.00)	1.82 (1.61, 3.00)	2.13 (1.49, 3.26)	2.10 (1.49, 3.26)
200	2.25 (1.75, 4.06)	2.26 (1.75, 3.98)	2.16 (1.69, 4.52)	2.11 (1.69, 4.47)
400	2.60 (1.82, 3.87)	2.62 (1.82, 3.87)	3.12 (1.72, 5.26)	2.95 (1.70, 4.76)
800	2.48 (1.91, 4.42)	2.47 (1.92, 4.43)	2.33 (1.71, 3.56)	2.35 (1.60, 3.29)

S7.4 Efficiency gains affected by censored data

In order to investigate the effect of the censored response on the efficiency gains obtained from envelope approach, we generated the data from the same model as in (5.1), i.e., the model for uncensored response. All the model parameters in (5.1) were also generated in the same way. We censored the response at zero, which leads to approximately 20% of responses being censored.

We kept the sample sizes same as the uncensored case, and computed the estimation variance of the BTQR and BETQR estimators. The results of estimation variances for an element in $\boldsymbol{\beta}_\tau$ with $\tau = 0.1, 0.25, 0.5, 0.75,$ 0.9 are summarized in Figure 2. The ratios of the estimation variance and MSE of the BTQR estimator versus the BETQR estimator for each element in $\boldsymbol{\beta}_\tau$ with $\tau = 0.1$ and 0.5 are recorded in Table 10. By comparing Table 10 and Table 1 in the main text, we notice that envelope approach achieves similar efficiency gains as with uncensored data. Take the eighth element in $\boldsymbol{\beta}_\tau$ for a close look, with sample size 200 and $\tau = 0.1$, the estimation variance of the BEQR estimator is 0.061 and that of the BQR estimator is 0.144. With censored data, the estimation variance of the BETQR estimator is 0.092 and that of the BTQR estimator is 0.287. This indicates the estimation variance increases with the censored responses for both the BQR and BEQR. However, the envelope approach still preserves the efficiency gains with censored data.

	$\tau = 0.1$		$\tau = 0.5$	
n	Ratio$_V$	Ratio$_M$	Ratio$_V$	Ratio$_M$
50	1.58 (1.23, 2.10)	1.58 (1.29, 2.10)	1.74 (1.53, 2.93)	1.74 (1.52, 2.92)
100	1.66 (1.31, 2.15)	1.63 (1.32, 2.15)	1.91 (1.48, 3.79)	1.91 (1.48, 3.79)
200	1.68 (1.41, 3.05)	1.66 (1.40, 2.92)	1.94 (1.73, 2.67)	1.96 (1.72, 2.69)
400	2.10 (1.68, 3.11)	2.06 (1.66, 3.11)	1.95 (1.68, 2.81)	1.95 (1.68, 2.80)
800	2.19 (1.80, 3.13)	2.11 (1.71, 2.97)	1.94 (1.64, 2.31)	1.94 (1.64, 2.30)

Table 10: Medians (ranges) of the estimation variance and MSE ratios. Ratio$_V$: Estimation variance ratio of the BTQR estimator versus the BETQR estimator. Ratio$_M$: Mean square error ratio of the BTQR estimator versus the BETQR estimator.

S7.5 Efficiency gains affected by censoring levels

To address the effect of the percentage of censored data, we generated the data from the model

$$Y_i = \mu_Y + \boldsymbol{\eta}^T \boldsymbol{\Gamma}_{1\tau}^T(\boldsymbol{A})\boldsymbol{X}_i + (5 + \boldsymbol{\alpha}^T \boldsymbol{X}_i)\epsilon_i;$$

$$\boldsymbol{X}_i \sim N_p\left(\boldsymbol{\mu}_{\boldsymbol{X}},\ \boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})\boldsymbol{\Omega}_1\boldsymbol{\Gamma}_{1\tau}(\boldsymbol{A})^T + \boldsymbol{\Gamma}_{2\tau}(\boldsymbol{A})\boldsymbol{\Omega}_2\boldsymbol{\Gamma}_{2\tau}(\boldsymbol{A})^T\right), \quad i = 1, \ldots, n.$$

We set $p = 8$. The envelope dimension u_τ was fixed at 3 for all τ ($0 \leq \tau \leq 1$). The entries of $\boldsymbol{A}$ were generated from the student's t distribution with 4 degrees of freedom. The matrices $\boldsymbol{\Omega}_1$ and $\boldsymbol{\Omega}_2$ are diagonal matrices. The diagonal elements of $\boldsymbol{\Omega}_1$ were 5, 15, and 30 and those of $\boldsymbol{\Omega}_2$ were 1, 1.1, 1.2, 1.3, and 1.4. All entries in $\boldsymbol{\eta}$ were 3, and each entry of $\boldsymbol{\mu}_{\boldsymbol{X}}$ was 1. The sample size n was varied from 150, 300, 600 and 1200.

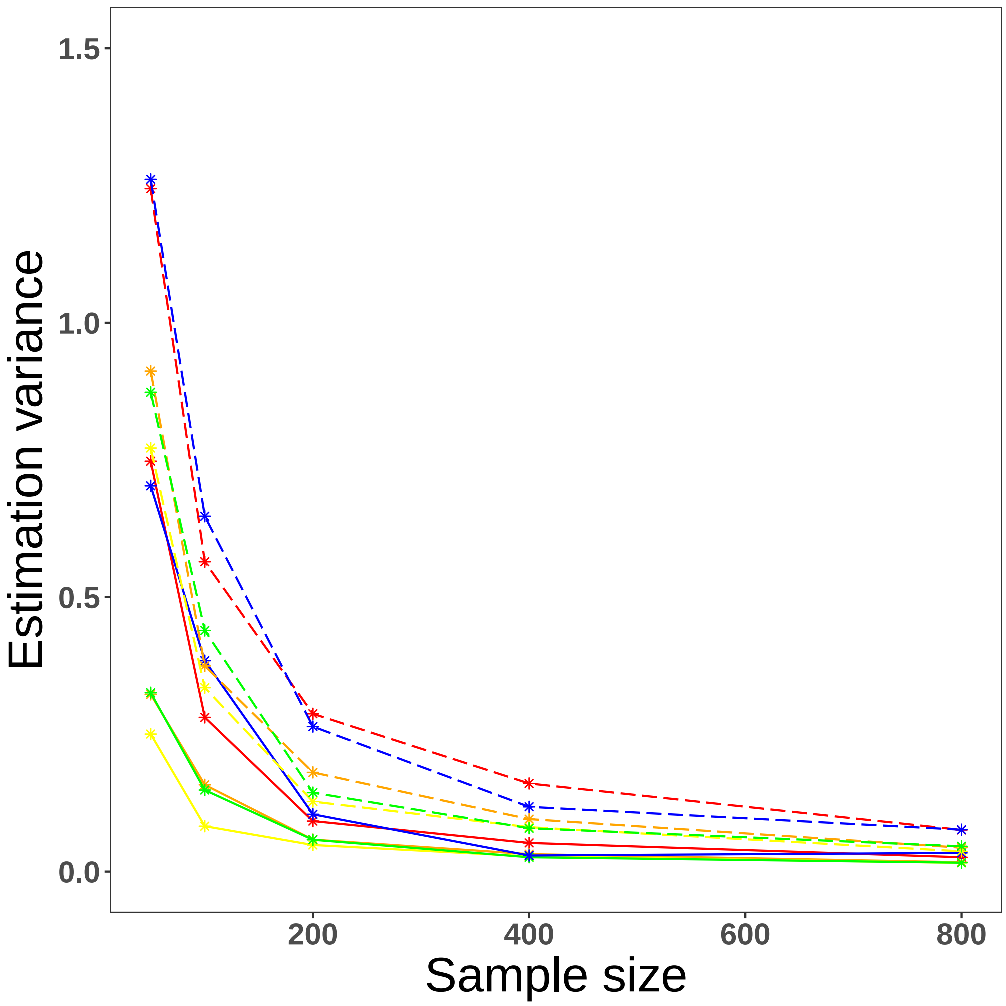

Figure 2: Estimation variances of the BETQR (solid line) and the BTQR (dashed line) estimators with $\tau = 0.1$ (red line), 0.25 (orange line), 0.5 (yellow line), 0.75 (green line), and 0.9 (blue line).

We consider the case where the response is left censored at zero. By adjusting the value of the intercept, we can obtain different censoring percentages. When $\mu_Y = 120$, we have 100% uncensored data. When $\mu_Y = 20$, 20% of the data are censored. When $\mu_Y = 10$, 33% of data are censored. When $\mu_Y = 0$, 50% of data are censored. For each case, we compared the BETQR estimator with true u_τ and the BTQR estimator. The results of the estimation variances of BETQR and BTQR estimators are displayed in Figure 3, and the ratios of the estimation variance and MSE of the BTQR estimator versus the BETQR estimator were summarized in Table 11 and Table 12.

From Figure 3, we notice that the estimation variance increases when the percentage of censoring increases for both the BTQR and BETQR estimator. But the ratios of the estimation variances and MSE are relatively stable across different percentages of censoring, as shown in Tables 11 and 12. For the smallest sample size ($n = 150$), the ratio seems to shrink a little when more data are censored. This maybe because that it is harder to estimate the envelope subspace when the sample size is small. However, the envelope approach still achieve substantial efficiency gains for all cases.

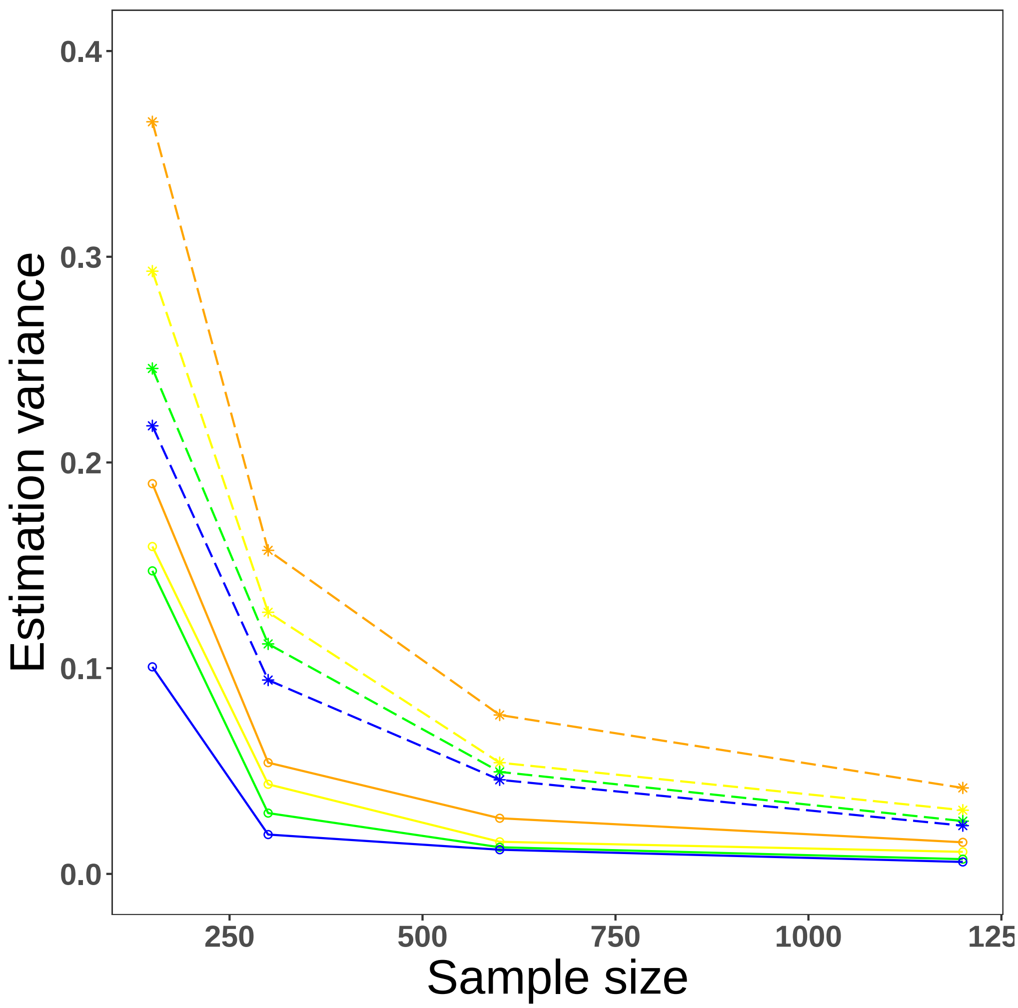

Figure 3: Estimation variances of the BETQR estimator (dashed line) and the BTQR estimator (solid line) with 50% censored responses (orange line), 33% censored responses (yellow line), 22% censored responses (green line), and no censored responses (blue line).

Table 11: Medians (ranges) of the estimation variance and MSE ratios. Ratio_V: Estimation variance ratio of the BTQR versus the BETQR estimator with true u_τ. Ratio_M: MSE ratio of the BTQR versus the BETQR estimator with true u_τ.

	No censored		20% censored	
n	Ratio_V	Ratio_M	Ratio_V	Ratio_M
150	5.35 (3.06, 7.89)	4.97 (2.85, 7.61)	4.70 (2.89, 7.36)	4.62 (2.52, 6.90)
300	5.97 (3.69, 9.22)	5.61 (3.37, 7.99)	6.49 (3.44, 9.14)	6.05 (3.23, 8.21)
600	6.18 (3.40, 8.42)	5.54 (3.13, 6.87)	6.58 (3.59, 9.58)	5.65 (3.35, 8.06)
1200	5.78 (3.85, 9.40)	4.44 (2.94, 7.06)	6.26 (3.51, 9.35)	4.70 (2.71, 7.74)

Table 12: Medians (ranges) of the estimation variance and MSE ratios. Ratio_V: Estimation variance ratio of the standard estimator versus the envelope estimator. Ratio_M: MSE ratio of the standard estimator versus the envelope estimator.

	33% censored		50% censored	
n	Ratio_V	Ratio_M	Ratio_V	Ratio_M
150	4.87 (3.22, 7.08)	4.38 (2.53, 6.81)	3.71 (2.22, 4.91)	3.31 (2.15, 4.90)
300	6.56 (3.33, 9.17)	6.18 (3.16, 8.43)	4.50 (3.23, 7.42)	4.15 (3.10, 7.26)
600	7.12 (3.61, 9.49)	5.94 (3.28, 8.13)	6.95 (3.45, 11.31)	5.88 (3.30, 9.53)
1200	7.01 (3.87, 8.86)	5.04 (2.90, 8.31)	7.64 (3.84, 10.44)	5.60 (2.88, 8.55)

S7.6 Additional simulation with censored data

We generated data from (5.1) with $p = 8$ and $u_\tau = 3$ for all τ $(0 \leq \tau \leq 1)$. The entries of $\boldsymbol{A}$ were generated from a student's t distribution with 4 degrees of freedom. The diagonal elements of $\boldsymbol{\Omega}_1$ were 5, 15, and 30 and those of $\boldsymbol{\Omega}_2$ were 1, 1.1, 1.2, 1.3, and 1.4. The intercept was $\mu_Y = 20$. Entries in $\boldsymbol{\eta}$ and $\boldsymbol{\mu_X}$ were all 3 and all 1 respectively. The sample size n was varied from 150, 300, 600 and 1200. We generated 200 datasets for each sample size. About 20% of the responses had negative values, they were left-censored at zero. The BETQR and Bayesian Tobit quantile regression (BTQR, $u_\tau = p$) point estimators were subsequently computed for each dataset. The results for $\tau = 0.25$, 0.5, 0.75, and 0.9 are shown for illustration. Figure 4 displays the estimation variances and squared biases of the BETQR and BTQR estimators for the second element of $\boldsymbol{\beta}_\tau$. Figure 4 shows BETQR has smaller variances than BTQR, illustrating the estimation efficiency gains achieved by enveloping. The envelope estimator also has smaller squared biases, thus yielding smaller MSEs compared to the BTQR estimator. The ratios of the estimation variances and MSEs of the BTQR to the BETQR estimator, together with the dimension selection performance of LOOIC are summarized in Table 13 for $\tau = 0.9$. The results are similar for other quantile levels. The ratios are strictly bigger than one, which indicates that the BETQR has a smaller estimation variance (and MSE) for all sample sizes than the BTQR. LOOIC performs well with censored data and selects the true u_τ more often as the sample size increases. LOOIC again appears to be conservative, overestimating u_τ more often than underestimating when it fails to select the true dimension.

			Selected u_τ			
n	Ratio$_V$	Ratio$_M$	2	3	4	5
150	1.83 (1.62, 2.59)	1.80 (1.61, 2.58)	13	111	75	1
300	3.76 (2.05, 4.72)	3.63 (2.01, 4.47)	5	158	34	3
600	3.72 (2.36, 5.76)	3.65 (2.29, 5.19)	0	169	31	0
1200	4.06 (2.33, 6.26)	3.67 (2.03, 4.48)	0	168	32	0

Table 13: Medians (ranges) of the estimation variance and MSE ratios and number of replications (out of 200) for which a given value of u_τ is selected with $\tau = 0.9$. Ratio$_V$: Estimation variance ratio of the BTQR estimator versus the BETQR estimator. Ratio$_M$: Mean square error ratio of the BTQR estimator versus the BETQR estimator.

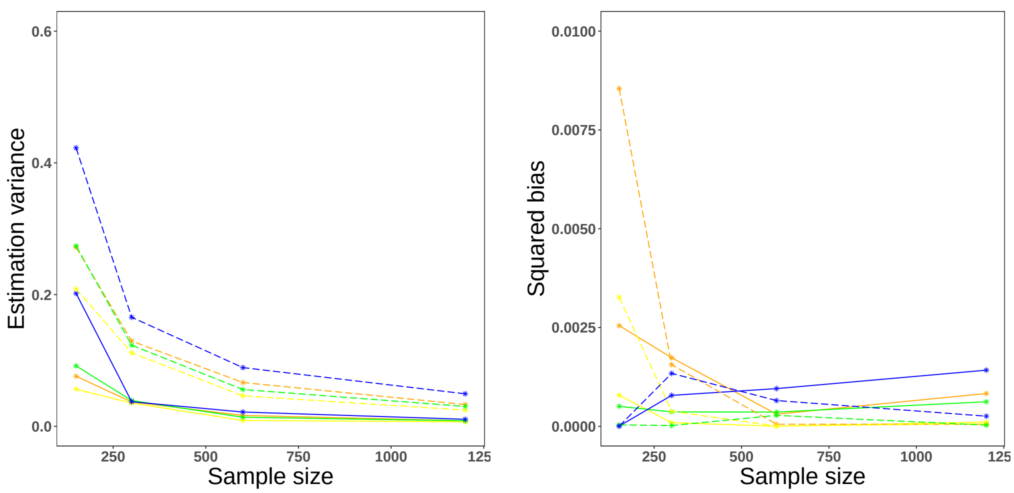

Figure 4: Estimation variances and squared biases of the second element in the BETQR estimator (solid line) and the BTQR estimator (dashed line) of $\boldsymbol{\beta}_\tau$ with $\tau = 0.25$ (orange line), 0.5 (yellow line), 0.75 (green line), and 0.9 (blue line).

S7.7 Data analysis

In this section, we provide additional results from the analysis of the LPGA dataset. Table 14 summarizes the dimension selection results by LOOIC, the ratios of the credible interval length from the BQR estimator versus the BEQR estimator, as well as the ratios of credible interval length from the `bayesQR` estimator versus the BEQR estimator for quantile levels 0.1, 0.25, 0.5, 0.75 and 0.9. From the ratios in Table 14, we notice that the BEQR estimator is more efficient than both the BQR and the `bayesQR` estimator, while the `bayesQR` estimator is the least efficient. The estimates for all predictor coefficients were plotted in Figure 5. Figure 5 confirms the findings from Table 14. The `bayesQR` estimator has especially large variance at extreme quantile levels. A estimation variance can overwhelm signals from all predictors. With the `bayesQR` estimator, all predictors are identified as insignificant. On the other hand, the efficiency gains from the BEQR model facilitate the detection of weaker signals, where some insignificant predictors under the BQR model are identified as significant under BEQR model. A detailed note on the significance of predictors included average drive (AD), percent of fairways hit (PFH), percent of greens reached in regulation (PGRR), average putts per round (APPR), percent of sand saves (PSS), green in regulation putts per hole (GRPH), average percentile in tournaments (APT), rounds completed (RC), and average strokes per round (ASR) at each quantile level was provided in Table 15.

τ	$\widehat{u}_\tau$	Ratio	Ratio_R
0.1	2	6.33 (1.91, 10.4)	59.69 (13.28, 91.15)
0.25	5	1.24 (1.14, 4.56)	5.26 (4.78, 18.93)
0.5	4	1.23 (0.99, 6.13)	3.53 (2.72, 16.94)
0.75	3	2.43 (0.89, 16.3)	6.65 (2.57, 48.95)
0.9	2	2.58 (1.42, 10.5)	9.88 (5.67, 42.93)

Table 14: Dimension selection results by LOOIC and medians (ranges) of length of 95% credible interval ratios of the BQR estimator versus the BEQR estimator. Ratio_R represents ratios of the `bayesQR` estimator versus the BEQR estimator.

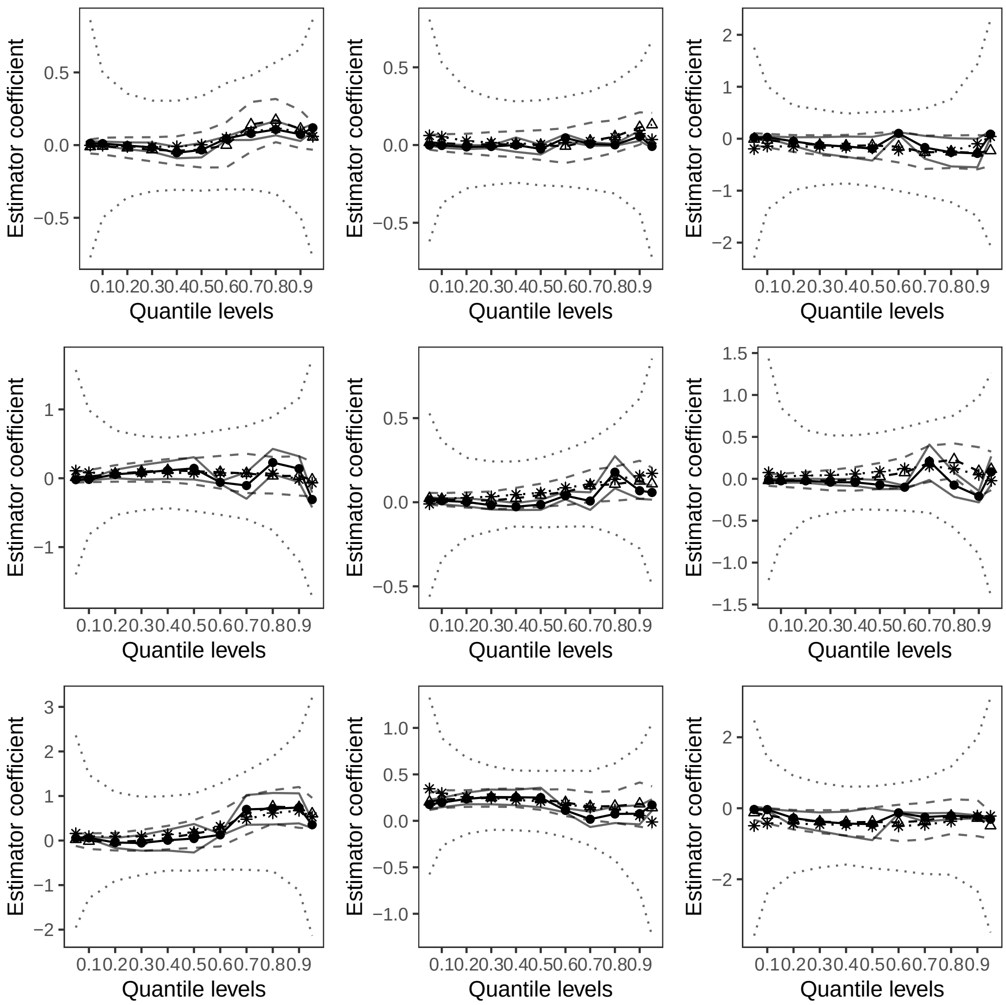

Figure 5: Point and 95% interval estimates of coefficients of all predictors under BEQR (solid lines), BQR (dashed) and `bayesQR` (dotted) models.

	$\tau = 0.1$		$\tau = 0.25$		$\tau = 0.5$		$\tau = 0.75$		$\tau = 0.9$	
Var	BEQR	BQR	BEQR	BQR	BEQR	BQR	BEQR	BQR	BEQR	BQR
AD	none	none	none	none	none	none	(+)	(+)	(+)	none
PFH	none	none	(-)	none	none	none	none	none	(+)	(+)
PGRR	none	none	none	none	none	none	none	none	(-)	none
APPR	(-)	none	(-)	none	none	none	none	none	none	none
PSS	none	none	none	none	none	none	(+)	(+)	(+)	(+)
GRPH	(-)	none	none	none	(-)	none	(+)	(+)	(+)	(+)
APT	(+)	none	none	none	none	none	(+)	(+)	(+)	(+)
RC	(+)	(+)	(+)	(+)	(+)	(+)	(+)	none	none	none
ASR	(-)	none	(-)	(-)	none	none	(-)	none	(-)	none

Table 15: Significance of predictors under BQR and BEQR: "none" mean the predictor is identified as non-significant, "(+)" mean the predictor is significant and has a positive coefficient, and "(-)" means the predictor is significant with a negative coefficient.

Bibliography

Azzalini, A. (2013). *The Skew-Normal and Related Families.* Institute of Mathematical Statistics Monographs. Cambridge University Press.

Dutta, S. (2012). Multiplicative random walk metropolis-hastings on the real line. *Sankhya B*, 74(2):315–342.

Geyer, C. J. (1998). Markov chain Monte Carlo lecture notes.

Khare, K., Pal, S., and Su, Z. (2017). A bayesian approach for envelope models. *The Annals of Statistics*, 45(1):196–222.

Meng, X.-L. and Rubin, D. B. (1993). Maximum likelihood estimation via the ecm algorithm: A general framework. *Biometrika*, 80(2):267–278.

Meyn, S. P. and Tweedie, R. L. (2012). *Markov chains and stochastic stability*. Springer Science & Business Media.

Roberts, G. O. and Rosenthal, J. S. (2006). Harris recurrence of metropolis-within-gibbs and trans-dimensional markov chains. *The Annals of Applied Probability*, pages 2123–2139.

Su, Z. and Cook, R. D. (2011). Partial envelopes for efficient estimation in multivariate linear regression. *Biometrika*, 98(1):133–146.