Supplementary Materials for "Bayesian Functional Analysis for Untargeted Metabolomics Data with Matching Uncertainty and Small Sample Sizes"

Guoxuan Ma[1], Jian Kang[1,*], and Tianwei Yu[2,*]

[1]*Department of Biostatistics, University of Michigan, Ann Arbor, MI 48109, USA*

[2]*Shenzhen Research Institute of Big Data, School of Data Science, The Chinese University of Hong Kong - Shenzhen (CUHK-Shenzhen), Shenzhen, Guangdong 518172, China*

[*]*To whom correspondence should be addressed: jiankang@umich.edu, yutianwei@cuhk.edu.cn*

1 Gibbs sampling and full conditionals

We develop a blocked Gibbs sampler for posterior inferences. We provide the full conditionals of each parameter in this section. Let $\mathbf{1}_L$ be an all-one vector of length L. Define $J_0 = \{j : z_j = 0\}$ and $J_1 = \{j : z_j = 1\}$. We can partition $\boldsymbol{\eta}_{1:k}$ into two parts, one containing η_j's with $j \in J_0$ and the other containing η_j's with $j \in J_1$, namely $\boldsymbol{\eta}_{J_0}$ and $\boldsymbol{\eta}_{J_1}$, respectively. We can partition $\boldsymbol{\lambda}_i$ and $\boldsymbol{\Lambda} = (\boldsymbol{\lambda}_1^\top, \ldots, \boldsymbol{\lambda}_p^\top)^\top$ in the same manner, resulting in $\boldsymbol{\lambda}_{i,J_0}$ and $\boldsymbol{\lambda}_{i,J_1}$ for $\boldsymbol{\lambda}_i$, and $\boldsymbol{\Lambda}_{J_0}$ and $\boldsymbol{\Lambda}_{J_1}$ for $\boldsymbol{\Lambda}$.

Full conditional of $\boldsymbol{\lambda}_i$ The full conditional of $\boldsymbol{\lambda}_i$ for $i = 1, \ldots, p$ is

$$p(\boldsymbol{\lambda}_i \mid \boldsymbol{r}, \, \cdot \,) = \text{Multinomial}(\boldsymbol{\lambda}_i \mid 1, \tilde{\boldsymbol{q}}_i),$$

where $\tilde{\boldsymbol{q}}_i = (\tilde{q}_{i,1}, \ldots, \tilde{q}_{i,k})^\top$ with

$$\tilde{q}_{i,j} = c_i \times q_{i,j} \times \text{N}(r_i \mid \eta_j^*, \sigma^2)$$

for $j = 1, \ldots, k$; c_i is a scaling factor such that $\sum_{j=1}^k \tilde{q}_{i,j} = 1$.

Full conditional of η_0   The full conditional of η_0 is

$$p(\eta_0 \mid \boldsymbol{r}, \, \cdot \,) = \text{N}(\eta_0 \mid \eta_{0,mean}, \, s_0^2),$$

where $\eta_{0,mean}$ and s_0^2 are as follows,

$$s_0^2 = \left(\frac{1}{\gamma_0} + \frac{1}{\sigma^2} \sum_{i=1}^{p} \left(\boldsymbol{\lambda}_{i,J_0}^{\top} \mathbf{1}_{|J_0|} \right)^2 \right)^{-1}$$

$$\eta_{0,mean} = \frac{s_0^2}{\sigma^2} \sum_{i=1}^{p} \boldsymbol{\lambda}_{i,J_0}^{\top} \mathbf{1}_{|J_0|} \left(r_i - \boldsymbol{\lambda}_{i,J_1}^{\top} \boldsymbol{\eta}_{J_1} \right).$$

Full conditional of $\boldsymbol{\eta}_{1:k}$ We consider $\boldsymbol{\eta}_{J_0}$ and $\boldsymbol{\eta}_{J_1}$ separately. For $\boldsymbol{\eta}_{J_0}$, each η_j with $j \in J_0$ has a full conditional

$$p(\eta_j \mid \boldsymbol{r}, \cdot) = \mathrm{N}\left(\eta_j \ \Big| \ \sum_{g=1}^{G} \mathrm{I}_{\{K_j=g\}} m_g, \sum_{g=1}^{G} \mathrm{I}_{\{K_j=g\}} \gamma_g \right).$$

The full conditional of $\boldsymbol{\eta}_{J_1}$ is

$$p(\boldsymbol{\eta}_{J_1} \mid \boldsymbol{r}, \cdot) = \mathrm{N}(\boldsymbol{\eta}_{J_1} \mid \boldsymbol{\eta}_{mean}, \boldsymbol{\Sigma}_\eta),$$

where $\boldsymbol{\eta}_{mean}$ and $\boldsymbol{\Sigma}_\eta$ are as follows,

$$\boldsymbol{\Sigma}_\eta = \left(\boldsymbol{\Lambda}_{J_1}^{\top} \mathbf{A}^{-1} \boldsymbol{\Lambda}_{J_1} + \frac{1}{\gamma_1} \mathbf{I}_{|J_1|} \right)^{-1}$$

$$\boldsymbol{\eta}_{mean} = \boldsymbol{\Sigma}_\eta \left\{ \frac{1}{\gamma_1} \boldsymbol{m}_{J_1} + \boldsymbol{\Lambda}_{J_1}^{\top} \mathbf{A}^{-1} \left(\boldsymbol{r} - \eta_0 \boldsymbol{\Lambda}_{J_0} \mathbf{1}_{|J_0|} \right) \right\}.$$

Full conditional of m_g For $g = 1, \ldots, G$, the full conditional of m_g is

$$p(m_g \mid \boldsymbol{r}, \cdot) = \mathrm{N}(m_g \mid m_{g,mean}, s_{m,g}^2),$$

where $m_{g,mean}$ and $s_{m,g}^2$ are as follows,

$$s_{m,g}^2 = \left(\frac{1}{\sigma_g^2} + \frac{N_g}{\gamma_g} \right)^{-1}$$

$$m_{g,mean} = s_{m,g}^2 \left(\frac{\mu_g}{\sigma_g^2} + \frac{\sum_{j:K_i=g} \eta_j}{\gamma_g} \right)$$

with $N_g = \sum_{j=1}^{k} \mathrm{I}_{\{K_i=g\}}$ is the number of j's in the g-th group.

Full conditionals of variance parameters The full conditionals of σ^2, γ_0, γ_g, β_g and σ_g^2 are

$$p(\sigma^2 \mid \boldsymbol{r}, \,\cdot\,) = \text{IG}\left(\sigma^2 \mid a_1 + \frac{1}{2}p,\ b_1 + \frac{1}{2}\sum_{i=1}^{p}\left\{r_i - \sum_{j=1}^{k}\lambda_{ij}\eta_0^{1-z_j}\eta_j^{z_j}\right\}^2\right)$$

$$p(\gamma_0 \mid \boldsymbol{r}, \,\cdot\,) = \text{IG}\left(\gamma_0 \mid a_2 + \frac{1}{2},\ b_2 + \frac{1}{2}\eta_0^2\right)$$

$$p(\gamma_g \mid \boldsymbol{r}, \,\cdot\,) = \text{IG}\left(\gamma_g \mid a_3 + \frac{1}{2}N_g,\ \beta + \frac{1}{2}\sum_{j:K_j=g}(\eta_j - m_g)^2\right)$$

$$p(\beta_g \mid \boldsymbol{r}, \,\cdot\,) = \text{Gamma}\left(\beta \mid a_4 + a_3,\ b_4 + \frac{1}{\gamma_g}\right)$$

$$p(\sigma_g^2 \mid \boldsymbol{r}, \,\cdot\,) = \text{IG}\left(\sigma_g^2 \mid a_5 + \frac{1}{2},\ b_5 + \frac{1}{2}(m_g - \mu_g)^2\right).$$

Full conditionals of K_j and p_g in the stick-breaking process The full conditionals of K_j for $j = 1, \ldots, k$ in the stick-breaking process [1] is

$$p(K_j \mid \boldsymbol{r}, \,\cdot\,) = \text{Categorical}(K_j \mid \tilde{\boldsymbol{p}}_j),$$

where $\tilde{\boldsymbol{p}}_j = (\tilde{p}_{j,1}, \ldots, \tilde{p}_{j,G})^\top$ with

$$\tilde{p}_g = p_g \times \text{N}(\eta_j \mid m_g, \gamma_g).$$

Then, for $g = 1, \ldots, G - 1$, the full conditional of p_g is

$$p_1 = V_1, \quad p_2 = (1 - V_1)V_2, \quad \ldots, \quad p_g = (1 - V_1)(1 - V_2)\cdots(1 - V_{g-1})V_g,$$

where $V_g \sim \text{Beta}(s_g + N_g, t_g + \sum_{g'>g}^{G} N_{g'})$ and $p_G = 1 - \sum_{g=1}^{G-1} p_g$.

Full conditional of $\boldsymbol{z}$ We assign a weighted Potts prior on the latent binary metabolite class labels $\boldsymbol{z}$. We adopt the Swendsen-Wang algorithm [2, 3], which has been widely used in the Potts model, for the metabolite network partition and efficient group updating of $\boldsymbol{z}$.

Algorithm 1 provides a summary of the group updating scheme of $\boldsymbol{z}$ based on the Swendsen-Wang algorithm. The Swendsen-Wang algorithm consists of two steps, 1) network partition and 2) network relabeling. The network partition aims to cut the metabolite network into subnetworks so that metabolites in each subnetwork share the same class label z. In this step, the algorithm introduces auxiliary variables U_{jl} for $j, l \in \{1, \ldots, k\}$ such that metabolite j and metabolite l are connected on the network. Given the latent class labels for metabolites j and l, i.e., z_j and z_l, U_{jl} follows a uniform distribution between 0 and $\exp\{\rho_{z_j} w_j c_{jl} \text{I}_{\{z_j = z_l\}}\} = \exp\{\rho_{z_j} \text{I}_{\{z_j = z_l\}}\}$ since that metabolite j and metabolite l are connected implies $c_{jl} = 1$ and in the work we fix $w_j = 1$ for all j.

Algorithm 1 Group updating scheme of $\boldsymbol{z}$ by the Swendsen-Wang Algorithm

---

Input: C, $\boldsymbol{z}$, $\boldsymbol{r}$, $\boldsymbol{\rho}$, $\boldsymbol{\pi}$ and all other parameters

Extract vertex set $\mathcal{V}$ and edge set $\mathcal{E}$ from **C**, let $G = \{\mathcal{V}, \mathcal{E}\}$ denote the metabolite network.

Cut G into $G_0 = \{\mathcal{V}_0, \mathcal{E}_0\}$ and $G_1 = \{\mathcal{V}_1, \mathcal{E}_1\}$, where $z_j = c$ for all $j \in \mathcal{V}_c$, for $c = 0, 1$.

Step 1: graph partition

for $c = 0, 1$ **do**

    for $e \in \mathcal{E}_c$ **do**

        Draw U_e from Uniform$(0, \exp\{\rho_c\})$.

        if $U_e < 1$ **then**

            Delete e.

        end if

    end for

    Obtain n_c clusters $G_{c,s}$ for $s = 1, \ldots, n_c$.

end for

Step 2: graph relabeling

for $g \in \{G_{0,1}, \ldots, G_{0,n_0}, G_{1,1}, \ldots, G_{1,n_1}\}$ **do**

    Flip the class label of all metabolites in g, denote $\boldsymbol{z}'$ the new latent class label vector.

    if $\exp\{\sum_{j=1}^{k} \log \pi_{z_j}\} p(\boldsymbol{r}|\boldsymbol{z}, \cdot) < \exp\{\sum_{j=1}^{k} \log \pi_{z'_j}\} p(\boldsymbol{r}|\boldsymbol{z}', \cdot)$ **then**

        Update $\boldsymbol{z}$ by $\boldsymbol{z}'$, i.e., $\boldsymbol{z} \leftarrow \boldsymbol{z}'$.

    end if

    Output: Updated latent class labels $\boldsymbol{z}$.

end for

In the step of network relabeling, the class labels of all the metabolites within each sub-network can either flip simultaneously or remain unchanged. The decision of flipping or not depends on which yields the higher full conditional of $\boldsymbol{z}$ given all the U_{jl}, which is proportional to $\exp\{\sum_{j=1}^{k} \tilde{w}_j \log \pi_{z_j}\} \times p(\boldsymbol{r}|\boldsymbol{z}, \cdot)$ where $p(\boldsymbol{r}|\boldsymbol{z}, \cdot)$ is the data likelihood. Note that $\tilde{w}_j = 1$ as we fix $w_j = 1$ for all $j = 1, \ldots, k$.

2   Real data analysis hyper-parameters

COVID-19 metabolomics data   For the hyper priors in the model, we assign $\sigma^2 \sim \text{IG}(2000, 1000)$ for a good fit of the data and $\gamma_0 \sim \text{IG}(10000, 1)$ under the assumption that the null distribution is tightly centered at 0. We set $\pi_1 = 0.5$ and approximate the Dirichlet Process $\mathcal{DP}(P_0, \tau)$ by a mixture of 50 Gaussian components, where we assign $\mu_g = g$ for and $\sigma_g^2 \sim \text{IG}(50, 50)$. For the variance γ_g of component g, we set $\gamma_g \sim \text{IG}(1000, \beta_g)$ and $\beta_g \sim \text{Gamma}(1000, 1)$. The stick-breaking prior for $\boldsymbol{p}$ is parameterized by $\boldsymbol{s} = \mathbf{10}_{50}$ and $\boldsymbol{t} = \mathbf{10}_{50}$.

Aging mouse brain data We use the same setting of hyper-parameters for the analysis of all three brain regions. We set $\boldsymbol{p} = (0.9, 0.1)^\top$ in the weighted Potts prior. A mixture of 15 Gaussian components is adopted for the approximation of the Dirichlet Process, where we assign Gaussian priors $N(\mu_g, \sigma_g^2)$ to the mean of each component m_g for $g = 1, 2, ..., 15$, where $m_g = g$ and $\sigma_g^2 \sim \text{IG}(100, 100)$. All other hyper-parameters are the same for the COVID-19 metabolomics data analysis.

3 Sensitivity analysis and pathway selection certainty measures

3.1 Sensitivity analysis

We performed sensitivity analysis to verify that our method was robust under mild changes to hyperparameters on both the COVID-19 metabolomics data and the mouse brain data. We varied key hyperparameters by $\pm 20\%$ based on the reported ones in Section S2, which yielded 128 different combinations for each dataset. For the COVID-19 dataset, the pairwise correlation coefficients across the 128 combinations of the metabolite posterior inclusion probabilities had a minimum of 0.954 with interquartile range (IQR) $(0.979, 0.988)$, and the pairwise correlation coefficients across the 128 combinations of the matching uncertainty estimation had a minimum of 0.900 with IQR $(0.954, 0.974)$. For the mouse brain dataset, we here present the result of Thalamus week 3 v.s. week 16, but results for all combinations are available under the pathway_certainty folder at `https://github.com/guoxuan-ma/BAUM`. The pairwise correlation coefficients across all combinations of the metabolite posterior inclusion probabilities was at least 0.885 with IQR $(0.960, 0.986)$, and the pairwise correlation coefficients across all combinations of the matching uncertainty estimation had a minimum of 0.790 with IQR $(0.918, 0.983)$. Results showed that our inference on metabolite importance and matching uncertainty were highly consistent under mild changes on hyperparameters. In addition, Table S1a shows the proportion of the reported metabolic pathways in Figure 2 that were reselected over the 128 combinations, and Table S1b shows the reselection rate for all significant pathways that were found using hyperparameters in Section S2. The selection of pathways was highly consistent for both dataset. This was particularly notable for the Thalamus week 3 v.s. week 16 subset of the mouse brain data because there is very limited observations (36) in this subset, but BAUM achieved a stable performance.

To assess the the stability of the estimation on the null component (g_0) and the alternative component (g_1), we check the coefficient of variation (CV, defined by SD/mean) of the posterior mean of η_0 and η_j's for all j=1,...,k. The CV is calculated over all the hyperparameter combinations and summarized for all real-data

Pathway	Reselection Rate	CM$_1$	CM$_2$
caffeine metabolism	100%	0.90	15.51
de novo fatty acid biosynthesis	100%	0.98	20.16
histidine metabolism	100%	0.99	30.81
linoleic acid metabolism	100%	0.98	42.99
tyrosine metabolism	100%	0.94	17.04
urea cycle/amino group metabolism	100%	0.96	20.30
vitamin b3 (nicotinate and nicotinamide) metabolism	97.7%	0.93	10.84
tryptophan metabolism	92.2%	0.96	14.55
alanine and aspartate metabolism	69.5%	0.96	15.21
arginine and proline metabolism	56.3%	0.94	12.47
phenylalanine and tyrosine metabolism	47.7%	0.92	10.22

(a) Reselection rate and certainty measures of significant pathways for the COVID-19 data reported in Figure 2 over the 128 hyperparameter combinations.

Pathway	Reselection Rate	CM$_1$	CM$_2$
beta oxidation of very long chain fatty acids	100%	0.95	18.19
de novo fatty acid biosynthesis	100%	0.97	22.72
fatty acid activation	100%	0.98	21.22
fatty acid biosynthesis	100%	0.99	37.78
xenobiotics metabolism	100%	0.96	15.40
fatty acid metabolism	82.0%	0.97	19.38
vitamin b3 (nicotinate and nicotinamide) metabolism	75.0%	0.82	8.66
nicotinate and nicotinamide metabolism	71.9%	0.80	6.22
pentose and glucuronate interconversions	36.7%	0.80	6.48
ascorbate and aldarate metabolism	27.3%	0.87	7.85
steroid hormone biosynthesis	19.5%	0.80	5.68
amino sugar and nucleotide sugar metabolism	14.1%	0.88	8.64

(b) Reselection rate and certainty measures of all significant pathways with at least 3 selected metabolites for Thalamus week 3 v.s. week 16 subset (36 observations) of the mouse brain data over the 128 hyperparameter combinations.

Table S1: Reselection rate of significant pathways that are found using the hyperparamters specified in Section S2 over the 128 hyperparameter combinations for (a) the COVID-19 data and (b) the mouse brain data.

analysis. Table S2 shows the summary statistics of CV for different analysis. In rare cases, CV might be large because the mean of posterior mean is close to zero. For all analysis, we overall observe low CV, showing evidence that BAUM is stable on the estimation of g_0 and g_1.

Analysis	η_0	$\eta_1 :$ Mean	$\eta_1 :$ Median	$\eta_1 :$ 95%-Quantile
COVID	15.5%	9.2%	9.2%	10.9%
Hippocampus 3 vs 16	13.0%	25.4%	15.3%	76.5%
Hippocampus 16 vs 59	12.8%	27.6%	15.4%	24.8%
Olfactory Bulb 3 vs 16	16.4%	22.2%	15.7%	26.4%
Olfactory Bulb 16 vs 59	12.4%	22.9%	18.0%	34.8%
Thalamus Bulb 3 vs 16	16.5%	24.1%	16.8%	28.4%
Thalamus Bulb 16 vs 59	17.0%	28.5%	15.3%	23.0%

Table S2: The coefficient of variation (CV) of the posterior mean of η_0 and, the mean, median and 95%-quantile of CV of the posterior mean of η_j's, for the set of seven analysis in this work.

3.2 Pathway selection confidence measures

For each pathway, we define two confidence measures for the selection of the pathway based on its p-value from the hypergeometric test and the metabolite FDR on the pathway subnetwork. Denote $\{\mathrm{FDR}_j\}_{j=1}^{n_s}$ as the set of FDR values for each of the metabolite on the subnetwork in the pathway, where n_s is the number of metabolites in the subnetwork. We define $\mathrm{CM}_1 = (1 - p\text{-value}) \times (1 - \mathrm{median}(\{\mathrm{FDR}_j\}_{j=1}^{n_s})$ and $\mathrm{CM}_2 = (-\log(p\text{-value})) \times (-\log(\mathrm{median}(\{\mathrm{FDR}_j\}_{j=1}^{n_s})))$. The range of CM_1 is 0 to 1 and CM_2 takes positive values. For both measures, a higher value implies a higher confidence level for selecting the pathway.

In Table S1a and S1b, we provide the two certainty measures (CM) for the pathway selection in the analysis of the COVID-19 metabolomics data and the mouse brain data respectively. Both confidence measures depended on the p-value of metabolite subnetwork selection and the metabolite-specific FDR on the subnetwork that were obtained using BAUM with hyperparameters specified in Section S2. The first confidence measure (CM_1) has range 0-1 and the second (CM_2) takes positive values. For both measures, a higher value implies a more confident pathway selection. From Table S1, the reselection rate are positively correlated to both confidence measures. A higher confidence measure typically associates with an increased likelihood of reselection, and a pathway with consistent high confidence measures from both CM_1 and CM_2 is highly likely to be reselected frequently. We provide the reselection rate and confidence measures for all pathways in the analysis of both datasets under the pathway_certainty folder at `https://github.com/guoxuan-ma/BAUM`.

4 Additional results

Additional results are available at `https://github.com/guoxuan-ma/BAUM`. We provide pathway plots for the COVID-19 metabolomics data and the mouse brain data under the

pathway_plots folder. We provide all results for sensitivity analysis and pathway selection confidence measures under the pathway_certainty folder. We provide plots of the $p \times k$ posterior matching estimati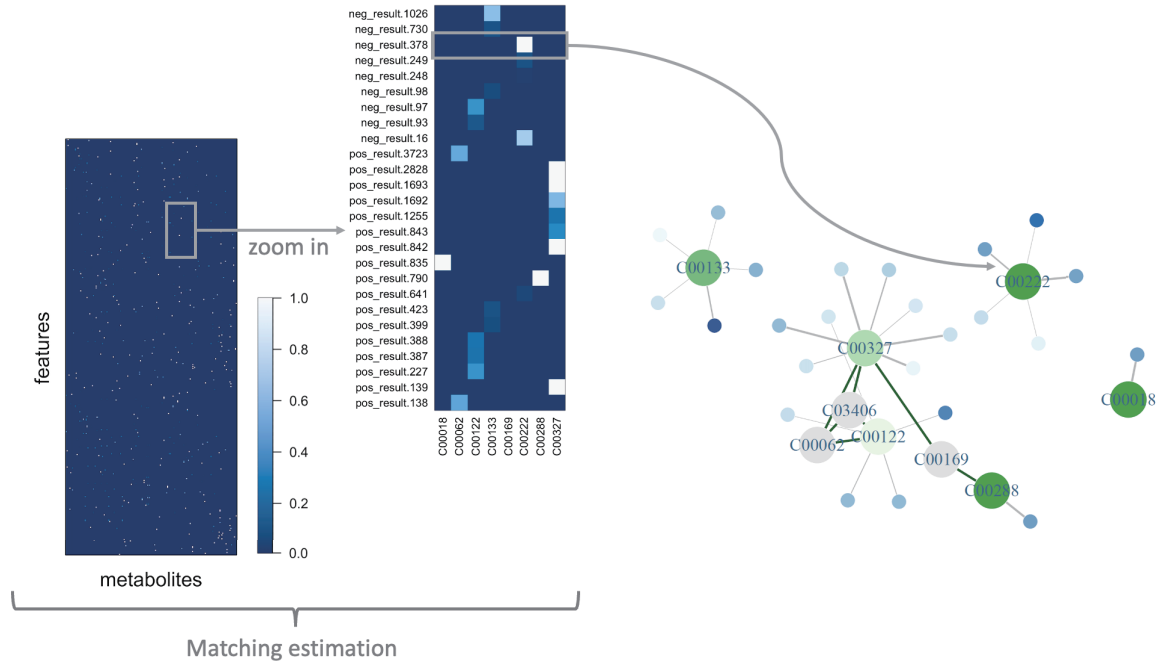on matrix Λ for all analysis in this work under the plot_matching_estimation folder. Figure S1 is an illustration on how to connect the plot of posterior matching estimation matrix Λ to the pathway figures in Section 4 (Figure 2 and Figure 4). Here we use the alanine and aspartate metabolisms (available in Figure 2) in the COIVD-19 metabolism data analysis as an example. On the right is the alanine and aspartate metabolisms where green nodes represent selected metabolites at FDR ≤ 0.2, blue nodes represent features matched to the metabolites and gray nodes represent metabolites that are not selected but connect selected metabolites. Edges between features (blue nodes) and selected metabolites (green nodes) represent the posterior matching estimation. On the left is the posterior matching estimation the whole Λ matrix and a zoom-in view of Λ containing only the selected metabolites and features plotted on the right. The intensity in the matching estimation matrix corresponds to the thickness of the feature-metabolite edges on the right.

For the COVID-19 metabolism data analysis, we compare the sample correlation among features not linked to the same metabolite to the sample correlation among

8

	cor > 0.15	cor > 0.15	cor > 0.2	cor > 0.3
No link	22.6%	12.8%	3.6%	0.3%
High confidence	31.8%	22.9%	13.7%	5.5%

Table S3: Proportions of the correlations exceeding different thresholds for both feature groups (no link and high confidence).

those with high-confidence links to the same metabolite. Features with high-confidence links show a higher mean correlation than those without links in **68.3%** of the cases, and a higher median correlation in **64.3%** of the cases. The table below summarizes the proportion of the correlations exceeding different thresholds for both groups, generally showing higher correlations for features with high-confidence links to the same metabolites. We also conduct a Kolmogorov–Smirnov test on the correlation distribution among the two feature groups (no link v.s. high confidence). The test has an p-value less than 2.2×10^{-16}, showing evidence for a significant difference. This provides strong statistical evidence of higher correlations among features with high confidence links is higher, and further shows BAUM's matching estimation is meaningful and reliable.

5 Discussions on identifiability

We would like to note that Model (1) is indeed not completely identifiable. However, under the Bayesian framework, we mitigate the issue through two strategies: adding constraints on model parameters and specifying informative/proper priors [4, 5]. Thus, we enable Bayesian learning on the parameters of interest [6].

First, for all $i = 1, \ldots, p$, we require λ_i to be a one-hot binary vector, which is then further constrained by the prior matching confidence q_i in the sense that λ_{ij} must be **0** if $q_{ij} = 0$ and λ_{ij} may either take **0** and **1** if $q_{ij} > 0$. The prior matching confidence q_i is often very sparse in practice, because a feature usually has only a few potentially matched metabolites. In fact, only **8.7%** features have more than **5** potentially matched metabolites in the COVID-19 metabolomics data we have analyzed, and only **3.2%** in the mice brain data. Those strong constraints on λ_{ij} help to mitigate the identifiability issue.

Secondly, we suggest assigning informative priors on some model parameters and have done so in all our analysis. For example, we specify tight prior around **0** on the null component (η_0) to distinguish it from the alternative component. There have been similar discussions in Section **2.3.2**, and we assign several tight priors in

each application (See Section 3 and Section S2).

References

[1] Ishwaran H and James L.F. Gibbs sampling methods for stick-breaking priors. *Journal of the American Statistical Association*, 2001, 96(453):161–173.

[2] Swendsen R.H and Wang J.S. Nonuniversal critical dynamics in monte carlo simulations. *Physical Review Letters*, 1987, 58(2):86.

[3] Jin Z, Kang J, and Yu T. Feature selection and classification over the network with missing node observations. *Statistics in Medicine*, 2022, 41(7):1242–1262.

[4] Lindley D. Bayesian statistics: a review. society for industrial and applied mathematics. *Philadelphia, Pa*, 1972.

[5] Xie Y and Carlin B.P. Measures of bayesian learning and identifiability in hierarchical models. *Journal of Statistical Planning and Inference*, 2006, 136(10):3458–3477.

[6] Gelfand A.E and Sahu S.K. Identifiability, improper priors, and gibbs sampling for generalized linear models. *Journal of the American Statistical Association*, 1999, 94(445):247–253.