Supplementary Materials

Peiyang Li[#], Xiaohui Gao[#], Cunbo Li, Chanlin Yi, Weijie Huang, Yajing Si, Fali Li, Zehong Cao*, Yin Tian* and Peng Xu*

I. SUPPLEMENTARY METHODS

This part describes the detailed model derivation of DLap-GCA. It consists of five categories, namely Appendix. A to Appendix. E. Appendix. A describes the derivation and solution of the target function of DLap-GCA (the formula (26) in the paper). Appendix. B describes the solution of hyperparameter τ_k on the basic of Appendix. A in DLap-GCA (corresponding to formula (27) in the paper). Appendix. C describes the solution of hyperparameter σ_k^2 in DLap-GCA (corresponding to formula (28) in the paper). And Appendix. D describes the solution of hyperparameter ξ_k and β_k in DLap-GCA (corresponding to formula (29) and (30) in the paper). Appendix. E describes the solution process of Lap-GCA which is corresponding to (14) in the paper. Furthermore, the specific derivation and solution process would be described as the following.

Appendix. A

Theoretically, based on Bayes formula

$$p(Y_k,\boldsymbol{g}_k \mid \theta)=p(Y_k \mid \boldsymbol{g}_k,\theta)p(\boldsymbol{g}_k \mid \theta) \tag{A.1}$$

the expected complete data log likelihood $Q(\theta,\theta^{(k)})$ in equation (25) can be rewritten as

$$\begin{aligned} Q(\theta,\theta^{(t)}) &= E_{p(\boldsymbol{g}_k|Y_k,\theta^{(t)})}\ln\left[p(Y_k,\boldsymbol{g}_k \mid \theta)\cdot p(\theta)\right] \\ &= E_{p(\boldsymbol{g}_k|Y_k,\theta^{(t)})}[\ln p(Y_k \mid \boldsymbol{g}_k,\theta)+\ln p(\boldsymbol{g}_k \mid \theta)+\ln p(\theta)] \end{aligned} \tag{A.2}$$

then we have

$$\begin{aligned} Q(\theta,\theta^{(t)}) &= E_{p(\boldsymbol{g}_k|Y_k,\theta^{(t)})}\left[\ln p(Y_k \mid \boldsymbol{g}_k,\theta)\right] \\ &+E_{p(\boldsymbol{g}_k|Y_k,\theta^{(t)})}\left[\ln p(\boldsymbol{g}_k \mid \beta_k,\xi_k)\right]+E_{p(\boldsymbol{g}_k|Y_k,\theta^{(t)})}\left[\ln p(\tau_k)\right] \end{aligned} \tag{A.3}$$

which transformed $Q(\theta,\theta^{(t)})$ into the summation of three expectations. Therefore, $Q(\theta,\theta^{(t)})$ can be expressed as the sum of the abovementioned three expectations, and can be estimated as follows

A. Estimate $E_{p(g_k|Y_k,\theta^{(t)})}\left[\ln p(Y_k \mid \boldsymbol{g}_k,\theta)\right]$

Theoretically, $E_{p(g_k|Y_k,\theta^{(t)})}\left[\ln p(Y_k \mid \boldsymbol{g}_k,\theta)\right]$ can be unfolded as

$$\begin{aligned} &E_{p(\boldsymbol{g}_k|Y_k,\theta^{(t)})}\left[\ln p(Y_k \mid \boldsymbol{g}_k,\theta)\right]=(-\frac{1}{2})\cdot\{(N-q)\cdot\ln 2\pi+\ln|\Sigma_{v_k}| \\ &+Y_k^T\Sigma_{v_k}^{-1}Y_k-2E_{p(\boldsymbol{g}_k|Y_k,\theta^{(t)})}\left[Y_k^T\Sigma_{v_k}^{-1}\boldsymbol{u}\boldsymbol{g}_k\right]+E_{p(\boldsymbol{g}_k|Y_k,\theta^{(t)})}\left[\boldsymbol{g}_k^T\boldsymbol{u}^T\Sigma_{v_k}^{-1}\boldsymbol{u}\boldsymbol{g}_k\right]\} \end{aligned} \tag{A.4}$$

where

$$E_{p(\boldsymbol{g}_k|Y_k,\theta^{(t)})}\left[Y_k^T\Sigma_{v_k}^{-1}\boldsymbol{u}\boldsymbol{g}_k\right]=Y_k^T\Sigma_{v_k}^{-1}\boldsymbol{u}C^{(t)}Y_k \tag{A.5}$$

$$\begin{aligned} &E_{p(\boldsymbol{g}_k|Y_k,\theta^{(t)})}\left[\boldsymbol{g}_k^T\boldsymbol{u}^T\Sigma_{v_k}^{-1}\boldsymbol{u}\boldsymbol{g}_k\right] \\ &=trace\{(\boldsymbol{u}^T\Sigma_{v_k}^{-1}\boldsymbol{u})P^{(t)}\}+Y_k^T(C^{(t)})^T(\boldsymbol{u}^T\Sigma_{v_k}^{-1}\boldsymbol{u})C^{(t)}Y_k \end{aligned} \tag{A.6}$$

Substituting (A.5) and (A.6) into (A.4), then $E_{p(g_k|Y_k,\theta^{(t)})}\left[\ln p(Y_k \mid g_k,\theta)\right]$ can be further expressed as

$$\begin{aligned} &E_{p(\boldsymbol{g}_k|Y_k,\theta^{(t)})}\left[\ln p(Y_k \mid \boldsymbol{g}_k,\theta)\right] \\ &=(-\frac{1}{2})\cdot\{(N-q)\cdot\ln 2\pi+\ln|\Sigma_{v_k}|+Y_k^T\Sigma_{v_k}^{-1}Y_k-2Y_k^T\Sigma_{v_k}^{-1}\boldsymbol{u}C^{(t)}Y_k \\ &+trace\{(\boldsymbol{u}^T\Sigma_{v_k}^{-1}\boldsymbol{u})P^{(t)}\}+Y_k^T(C^{(t)})^T(\boldsymbol{u}^T\Sigma_{v_k}^{-1}\boldsymbol{u})C^{(t)}Y_k\} \end{aligned} \tag{A.7}$$

B. Estimate $E_{p(\boldsymbol{g}_k|Y_k,\theta^{(t)})}\left[\ln p(\boldsymbol{g}_k \mid \beta_k, \xi_k)\right]$

$E_{p(\boldsymbol{g}_k|Y_k,\theta^{(t)})}\left[\ln p(\boldsymbol{g}_k \mid \beta_k, \xi_k)\right]$ can be rewritten as

$$\begin{aligned} &E_{p(\boldsymbol{g}_k|Y_k,\theta^{(t)})}\left[\ln p(\boldsymbol{g}_k \mid \beta_k, \xi_k)\right] \\ &= (-\frac{1}{2})\cdot\{\left(D\times q\right)\cdot\ln 2\pi - \sum_{m=1}^{D\times q}\ln\frac{1}{2}\sqrt{2\pi\beta_{k,m}\left|\xi_{k,m}\right|}e^{-\frac{1}{2}\beta_{k,m}\left|\xi_{k,m}\right|} \\ &+\ln|\Lambda_k| + E_{p(\boldsymbol{g}_k|Y_k,\theta^{(t)})}\left[\boldsymbol{g}_k^T\Lambda_k^{-1}\boldsymbol{g}_k\right]\} \end{aligned} \tag{A.8}$$

where

$$E_{p(\boldsymbol{g}_k|Y_k,\theta^{(t)})}\left[\boldsymbol{g}_k^T\Lambda_k^{-1}\boldsymbol{g}_k\right] = trace\{\Lambda_k^{-1}P^{(t)}\} + Y_k^T(C^{(t)})^T\Lambda_k^{-1}C^{(t)}Y_k \tag{A.9}$$

Then we have

$$\begin{aligned} &E_{p(\boldsymbol{g}_k|Y_k,\theta^{(t)})}\left[\ln p(\boldsymbol{g}_k \mid \beta_k, \xi_k)\right] \\ &= (-\frac{1}{2})\cdot\{\left(D\times q\right)\cdot\ln 2\pi - \sum_{m=1}^{D\times q}\ln\frac{1}{2}\sqrt{2\pi\beta_{k,m}\left|\xi_{k,m}\right|}e^{-\frac{1}{2}\beta_{k,m}\left|\xi_{k,m}\right|} \\ &+\ln|\Lambda_k| + trace\{\Lambda_k^{-1}P^{(t)}\} + Y_k^T(C^{(t)})^T\Lambda_k^{-1}C^{(t)}Y_k\} \end{aligned} \tag{A.10}$$

C. Estimate $E_{p(\boldsymbol{g}_k|Y_k,\theta^{(t)})}\left[\ln p(\tau_k)\right]$

Here we have

$$E_{p(\boldsymbol{g}_k|Y_k,\theta^{(t)})}\left[\ln p(\tau_k)\right] = \sum_{n=1}^{N-q}\ln p\left(\tau_k\left(n\right)\right) \tag{A.11}$$

Substituting (A.7), (A.10) and (A.11) into (A.3), we get

$$\begin{aligned} &Q(\theta,\theta^{(t)}) = (-\frac{1}{2})\cdot\{\ln|\Sigma_{v_k}| + Y_k^T\Sigma_{v_k}^{-1}Y_k + \ln|\Lambda_k| \\ &-2\sum_{n=1}^{N-q}\ln\frac{1}{\sigma_k^2} + 2\sum_{n=1}^{N-q}\frac{\tau_k\left(n\right)}{\sigma_k^2} + trace\{(\boldsymbol{u}^T\Sigma_{v_k}^{-1}\boldsymbol{u}+\Lambda_k^{-1})P^{(t)}\} + \frac{1}{2}\sum_{m=1}^{D\times q}\beta_{k,m}\left|\xi_{k,m}\right| \\ &-2Y_k^T\Sigma_{v_k}^{-1}\boldsymbol{u}C^{(t)}Y_k - \sum_{m=1}^{D\times q}\ln\sqrt{2\pi\beta_{k,m}\left|\xi_{k,m}\right|} + Y_k^T(C^{(t)})^T(\boldsymbol{u}^T\Sigma_{v_k}^{-1}\boldsymbol{u}+\Lambda_k^{-1})C^{(t)}Y_k\} \end{aligned} \tag{A.12}$$

In essence, (A.12) is mainly consist of three additive operators as

$$\begin{aligned} &Q(\tau_k,\theta^{(t)}) \\ &= (-\frac{1}{2})\cdot\{\mathrm{Y}_k^T[\Sigma_{v_k}^{-1}(I-2\boldsymbol{u}C^{(t)}) + (C^{(t)})^T\boldsymbol{u}^T\Sigma_{v_k}^{-1}\boldsymbol{u}C^{(t)}]Y_k \\ &+\sum_{n=1}^{N-q}\ln\tau_k\left(n\right) + 2\sum_{n=1}^{N-q}\frac{\tau_k\left(n\right)}{\sigma_k^2} + trace\{\boldsymbol{u}^T\Sigma_{v_k}^{-1}\boldsymbol{u}P^{(t)}\}\} \end{aligned} \tag{A.13}$$

$$Q(\sigma_k^2,\theta^{(t)}) = (-\frac{1}{2})\cdot\left\{-2\sum_{n=1}^{N-q}\ln\frac{1}{\sigma_k^2} + 2\sum_{n=1}^{N-q}\frac{\tau_k\left(n\right)}{\sigma_k^2}\right\} \tag{A.14}$$

$$\begin{aligned} &Q(\beta_k,\left|\xi_k\right|,\theta^{(t)}) = (-\frac{1}{2})\cdot\{\ln|\Lambda_k| - \sum_{m=1}^{D\times q}\ln\sqrt{2\pi\beta_{k,m}\left|\xi_{k,m}\right|} \\ &+\frac{1}{2}\sum_{m=1}^{D\times q}\beta_{k,m}\left|\xi_{k,m}\right| + trace\{\Lambda_k^{-1}P^{(t)}\} + Y_k^T(C^{(t)})^T\Lambda_k^{-1}C^{(t)}Y_k\} \end{aligned} \tag{A.15}$$

which are independent from each other, thus it is easy to see that $Q(\theta,\theta^{(t)})$ is in essence proportional to the sum of $Q(\sigma_k^2,\theta^{(t)})$, $Q(\tau_k,\theta^{(t)})$ and $Q(\beta_k,|\xi_k|,\theta^{(t)})$ as

$$Q(\theta,\theta^{(t)}) \propto Q(\sigma_k^2,\theta^{(t)}) + Q(\tau_k,\theta^{(t)}) + Q(\beta_k,|\xi_k|,\theta^{(t)}) \tag{A.16}$$

By approximating these three expectations we can estimate the hyperparameters $\sigma_k^2,\tau_k,\beta_k,|\xi_k|$ directly.

Appendix. B

Estimate hyperparameter τ_k

$Q(\tau_k,\theta^{(t)})$ holds the form as described in (A.13), by denoting

$$\begin{aligned} S^{(t)} &= \boldsymbol{u}P^{(t)}\boldsymbol{u}^T \\ K^{(t)} &= \boldsymbol{u}C^{(t)}Y_k \Rightarrow (K^{(t)})^T = Y_k^T(C^{(t)})^T\boldsymbol{u}^T \end{aligned} \tag{B.1}$$

then we have

$$trace\{\boldsymbol{u}^T\Sigma_{v_k}^{-1}\boldsymbol{u}P^{(t)}\} = trace\left\{\mathrm{diag}\left(S^{(t)}\right)\Sigma_{v_k}^{-1}\right\} = \sum_{n=1}^{N-q} S^{(t)}(n,n)\tau_k(n)^{-1} \tag{B.2}$$

$$\begin{aligned} &\mathrm{Y}_k^T[\Sigma_{v_k}^{-1}(I-2\boldsymbol{u}C^{(t)}) + (C^{(t)})^T\boldsymbol{u}^T\Sigma_{v_k}^{-1}\boldsymbol{u}C^{(t)}]Y_k \\ &= Y_k^T\Sigma_{v_k}^{-1}Y_k - 2Y_k^T\Sigma_{v_k}^{-1}\boldsymbol{u}C^{(t)}Y_k + Y_k^T(C^{(t)})^T\boldsymbol{u}^T\Sigma_{v_k}^{-1}\boldsymbol{u}C^{(t)}Y_k \end{aligned} \tag{B.3}$$

Where

$$\begin{aligned} Y_k^T\Sigma_{v_k}^{-1}Y_k &= \sum_{n=1}^{N-q} Y_k(n)^2\tau_k(n)^{-1} \\ Y_k^T\Sigma_{v_k}^{-1}\boldsymbol{u}C^{(t)}Y_k &= \sum_{n=1}^{N-q} Y_k(n)^2 K^{(t)}(n)\tau_k(n)^{-1} \\ Y_k^T(C^{(t)})^T\boldsymbol{u}^T\Sigma_{v_k}^{-1}\boldsymbol{u}C^{(t)}Y_k &= \sum_{n=1}^{N-q}\left(K^{(t)}(n)\right)^2\tau_k(n)^{-1} \end{aligned} \tag{B.4}$$

Thus, we get

$$\begin{aligned} &\mathrm{Y}_k^T[\Sigma_{v_k}^{-1}(I-2\boldsymbol{u}C^{(t)}) + (C^{(t)})^T\boldsymbol{u}^T\Sigma_{v_k}^{-1}\boldsymbol{u}C^{(t)}]Y_k \\ &= \sum_{n=1}^{N-q}[Y_k(n)^2 - 2Y_k(n)K^{(t)}(n) + \left(K^{(t)}(n)\right)^2]\tau_k(n)^{-1} \\ &= \sum_{n=1}^{N-q}\Xi^{(t)}(n)\tau_k(n)^{-1} \end{aligned} \tag{B.5}$$

Where $\Xi^{(t)}$ has described in (31). Substituting (B.2) and (B.5) into (A.13), $Q(\tau_k,\theta^{(t)})$ can rewritten as

$$\begin{aligned} Q(\tau_k,\theta^{(t)}) &= (-\frac{1}{2})\cdot\{\sum_{n=1}^{N-q}\ln\tau_k(n) + 2\sum_{n=1}^{N-q}\frac{\tau_k(n)}{\sigma_k^2} \\ &+ \sum_{n=1}^{N-q}\Xi^{(t)}(n)\tau_k(n) + \sum_{n=1}^{N-q}S^{(t)}(n,n)\tau_k(n)\} \end{aligned} \tag{B.6}$$

The element $\tau_k(n)$ in vector τ_k can be described as

$$Q(\tau_k(n),\theta^{(t)}) = (-\frac{1}{2})\cdot\{\ln\tau_k(n) + 2\frac{\tau_k(n)}{\sigma_k^2} + \frac{\Xi^{(t)}(n) + S^{(t)}(n,n)}{\tau_k(n)}\} \tag{B.7}$$

Take the partial derivation of (B.7) with respect to $\tau_k(n)$, and then the hyperparameter $\tau_{k,t}$ holds the solution as

$$\tau_k^{(t+1)}(n) = \frac{\left(\sigma_k^2\right)^{(t)}}{4}\left(\sqrt{1 + \frac{8\cdot\left\{\Xi^{(t)}(n) + S^{(t)}(n,n)\right\}}{\left(\sigma_k^2\right)^{(t)}}} - 1\right) \tag{B.8}$$

Appendix. C

Estimate σ_k^2

$Q(\sigma_k^2, \theta^{(t)})$ has been expressed in (A.14), by taking the derivation of (A.14) with respect to σ_k^2 and set it to zero, we have

$$\left(\sigma_k^2\right)^{(t+1)} = \frac{\sum_{n=1}^{N-q} \tau_k^{(t)}(n)}{N-q} \tag{C.1}$$

Appendix. D

Estimate ξ_k and β_k

$Q(\beta_k, |\xi_k|, \theta^{(t)})$ is defined as the same as (A.15), where

$$trace\{\Lambda_k^{-1} P^{(t)}\} = \sum_{m=1}^{D\times q} \left(P^{(t)}(m,m)\right)\left(\frac{|\xi_{k,m}|}{\beta_{k,m}}\right)^{-1} \tag{D.1}$$

Setting

$$O^{(t)} = \left(C^{(t)}\right) Y_k \Rightarrow (O^{(t)})^T = Y_k^T \left(C^{(t)}\right)^T \tag{D.2}$$

Then $Y_k^T (C^{(t)})^T \Lambda_k^{-1} C^{(t)} Y_k$ in (A.15) can be expressed as

$$Y_k^T (C^{(t)})^T \Lambda_k^{-1} C^{(t)} Y_k = \sum_{m=1}^{D\times q} \left(O^{(t)}(m)\right)^2 \left(\frac{|\xi_{k,m}|}{\beta_{k,m}}\right)^{-1} \tag{D.3}$$

Substituting (D.2) and (D.3) into (A.15), we get

$$\begin{aligned} &Q(\beta_{k,m}, \theta^{(t)}) = Q(|\xi_{k,m}|, \theta^{(t)}) \\ &= (-\frac{1}{2}) \cdot \{\ln\frac{|\xi_{k,m}|}{\beta_{k,m}} - \ln\sqrt{2\pi\beta_{k,m}|\xi_{k,m}|} + \frac{1}{2}\beta_{k,m}|\xi_{k,m}| \\ &+ \left(P^{(t)}(m,m)\right)\left(\frac{|\xi_{k,m}|}{\beta_{k,m}}\right)^{-1} + \left(O^{(t)}(m)\right)^2\left(\frac{|\xi_{k,m}|}{\beta_{k,m}}\right)^{-1}\} \end{aligned} \tag{D.4}$$

Taking the partial derivation of (D.4) with respect to $\beta_{k,m}$ and $|\xi_{k,m}|$ under setting them to zeros, then the parameters can be estimated as

$$\left(\xi_{k,m}^{(t+1)}\right)^2 = \Psi^{(t)}(m) \tag{D.5}$$

$$\frac{1}{\beta_{k,m}^{(t+1)}} = \frac{1}{2}\left(\left|\xi_{k,m}^{(t+1)}\right| + \frac{\Psi^{(t)}(m)}{\left|\xi_{k,m}^{(t+1)}\right|}\right) = \left|\xi_{k,m}^{(t+1)}\right| \tag{D.6}$$

$$\xi_k = \left[\xi_{k,1}^{(t+1)},\ \xi_{k,2}^{(t+1)}, \cdots,\ \xi_{k,D\times q}^{(t+1)}\right] \tag{D.7}$$

$$\beta_k = \left[\beta_{k,1}^{(t+1)},\ \beta_{k,2}^{(t+1)}, \cdots,\ \beta_{k,D\times q}^{(t+1)}\right] \tag{D.8}$$

Where $\Psi^{(t)}$ has been expressed in (31).

Appendix. E

The priors distribution of the weights $\boldsymbol{g}_k$ and the noises model E_k can be described as the following based on (14), i.e.

$$p(\boldsymbol{g}_k \mid \boldsymbol{\beta}_k) = \prod_{m=1}^{D\times q}\left(\frac{2}{\beta_{k,m}}\right)^{-1} e^{-\beta_{k,m}|g_{k,m}|} \tag{E.1}$$

$$p\left(E_k \mid \sigma_k^2\right) = (2\pi)^{-\frac{N-q}{2}} \left|\Sigma_{v_k}\right|^{-\frac{1}{2}} e^{-\frac{1}{2}E_k^T \Sigma_{v_k}^{-1} E_k} \tag{E.2}$$

In order to estimate the marginal likelihood of $p(\boldsymbol{g}_k \mid \boldsymbol{\beta}_k)$ easily, the lower band about $-|g_{k,m}|$ were brought in (E.1), which has been mentioned in formula (20), that is

$$-\left|\boldsymbol{g}_{k,m}\right| \geq -\frac{1}{2}\left(\frac{\boldsymbol{g}_{k,m}^2}{\left|\xi_{k,m}\right|}+\left|\xi_{k,m}\right|\right) \tag{E.3}$$

So (E.1) can further infer that $p(\boldsymbol{g}_k \mid \boldsymbol{\beta}_k) \geq p(\boldsymbol{g}_k \mid \boldsymbol{\beta}_k, \xi_k)$, and the posterior distribution of $\boldsymbol{g}_k$ can be rewrite as

$$\begin{aligned} p(\boldsymbol{g}_k \mid Y_k, \sigma_k^2, \boldsymbol{\beta}_k) &= \frac{p(Y_k \mid \boldsymbol{g}_k, \sigma_k^2)\cdot p(\boldsymbol{g}_k \mid \boldsymbol{\beta}_k)}{\int p(Y_k \mid \boldsymbol{g}_k, \sigma_k^2)\cdot p(\boldsymbol{g}_k \mid \boldsymbol{\beta}_k) d\boldsymbol{g}_k} \geq \\ p(\boldsymbol{g}_k \mid Y_k, \sigma_k^2, \boldsymbol{\beta}_k, \xi_k) &= \frac{p(Y_k \mid \boldsymbol{g}_k, \sigma_k^2)\cdot p(\boldsymbol{g}_k \mid \boldsymbol{\beta}_k, \xi_k)}{\int p(Y_k \mid \boldsymbol{g}_k, \sigma_k^2)\cdot p(\boldsymbol{g}_k \mid \boldsymbol{\beta}_k, \xi_k) d\boldsymbol{g}_k} = N\left(\boldsymbol{g}_k \mid CY_k, P\right) \end{aligned} \tag{E.4}$$

where the covariance matrix of $\boldsymbol{g}_k$ is $P=\left(\boldsymbol{u}^T\Sigma_{v_k}^{-1}\boldsymbol{u}+\Lambda_k^{-1}\right)^{-1}$ and the expectation of $\boldsymbol{g}_k$ is $CY_k = P\boldsymbol{u}^T\Sigma_{v_k}^{-1}Y_k$. Therefore, the hyperparameters set can be express as

$$\theta=\left[\beta_{k,1},\beta_{k,2},\cdots,\beta_{k,D\times q},\xi_{k,1},\xi_{k,2},,\xi_{k,D\times q},\sigma_k^2\right] \tag{E.5}$$

Furthermore, by integrating out $\boldsymbol{g}_k$ the expectation-maximization (EM) approach can be calculated as

$$\begin{aligned} &E_{p(\boldsymbol{g}_k \mid Y_k, \theta^{(t)})} ln\left[p(Y_k \mid \boldsymbol{g}_k, \sigma_k^2)\cdot p(\boldsymbol{g}_k \mid \boldsymbol{\beta}_k, \xi_k)\right] = \\ &(-\frac{1}{2})\cdot\{-\sum_{m=1}^{D\times q}\ln\sqrt{2\pi\beta_{k,m}\left|\xi_{k,m}\right|}+\frac{1}{2}\sum_{m=1}^{D\times q}\beta_{k,m}\left|\xi_{k,m}\right| \\ &-2Y_k^T\Sigma_{v_k}^{-1}\boldsymbol{u}C^{(t)}Y_k + trace\{(\boldsymbol{u}^T\Sigma_{v_k}^{-1}\boldsymbol{u}+\Lambda_k^{-1})P^{(t)}\} \\ &+Y_k^T(C^{(t)})^T(\boldsymbol{u}^T\Sigma_{v_k}^{-1}\boldsymbol{u}+\Lambda_k^{-1})C^{(t)}Y_k\} \end{aligned} \tag{E.6}$$

Therefore, by maximizing (E.6) with respect to $\boldsymbol{\beta}_k, \xi_k$ and σ_k^2 respectively, we can obtain an alternating solution as

$$\left(\sigma_k^2\right)^{(t+1)} = \frac{Y_k^TY_k - 2Y_k^T\boldsymbol{u}C^{(t)}Y_k + trace\left\{\boldsymbol{u}\left(P^{(t)}+C^{(t)}Y_kY_k^T\left(C^{(t)}\right)^T\right)\boldsymbol{u}^T\right\}}{N-q} \tag{E.7}$$

$$\left(\xi_{k,m}^{(t+1)}\right)^2 = \Psi^{(t)}(m) \tag{E.8}$$

$$\frac{1}{\beta_{k,m}^{(t+1)}} = \frac{1}{2}\left(\left|\xi_{k,m}^{(t+1)}\right|+\frac{\Psi^{(t)}(m)}{\left|\xi_{k,m}^{(t+1)}\right|}\right) = \left|\xi_{k,m}^{(t+1)}\right| \tag{E.9}$$

where

$$\begin{aligned} \Psi^{(t)} &= diag\left[\left(\boldsymbol{u}^T\left(\Sigma_{v_k}^{(t)}\right)^{-1}\boldsymbol{u}+\Lambda_k^{-1}\right)^{-1}+\left(C^{(t)}Y_k\right)\left(C^{(t)}Y_k\right)^T\right] \\ P^{(t)} &= \left(\boldsymbol{u}^T\left(\Sigma_{v_k}^{(t)}\right)^{-1}\boldsymbol{u}+\left(\Lambda_k^{(t)}\right)^{-1}\right)^{-1} \\ C^{(t)} &= P^{(t)}\boldsymbol{u}^T\left(\Sigma_{v_k}^{(t)}\right)^{-1} \end{aligned} \tag{E. 10}$$

Subsequently, the hyperparameter matrices Σ_{v_k} and Λ_k could be updated as

$$\begin{aligned} \Sigma_{v_k}^{(t+1)} &= \left(\sigma_k^2\right)^{(t+1)} I \\ \Lambda_k^{(t+1)} &= diag\left\{\frac{\left|\xi_{k,1}^{(t+1)}\right|}{\beta_{k,1}^{(t+1)}},\frac{\left|\xi_{k,2}^{(t+1)}\right|}{\beta_{k,2}^{(t+1)}},\ldots,\frac{\left|\xi_{k,D\times q}^{(t+1)}\right|}{\beta_{k,D\times q}^{(t+1)}}\right\}. \end{aligned} \tag{E.11}$$

Through updating (E.7)- (E.11) until (E.6) get converged, and the ultimate values of $\boldsymbol{g}_k$ can be further calculated as

$$\boldsymbol{g}_k^* = CY_k \tag{E.12}$$

II. SUPPLEMENTARY RESULTS

A. Data description

BCI Competition IV dataset I (Dataset-2)

This dataset was recorded at 59 EEG channels from 7 healthy participants who performed MI task [1]. It consists of 100 trials per class for each subject with a sample rate of 1000 Hz. More detailed information could be found at https://www.bbci.de/competition/iv/desc_1.html.

BCI Competition III dataset IVa (Dataset-3)

This dataset was recorded at 118 EEG channels from 5 healthy participants who performed MI task [2]. During experiment, each subject conducted 280 trails of motor imagery tasks with 140 trials for right-hand MI, and remaining for foot MI. More detailed information could be found at https://www.bbci.de/competition/iv/desc_1.html.

B. Network characters

As for motor imagery, previous studies have proved that there would be strong ERD/ERS changes over the contralateral sensorimotor area when people are performing related tasks [3-5], which indicates that the network patterns would contain lateralization patterns when the MI EEGs are free of ocular artifacts. Thus, the local efficiency (LE) [4], the accumulation coupling strength (ACS) [6], and the local clustering coefficient (LCC) [4] were adopted as the additional 3 indexes to evaluate the performances of DLap-GCA in capturing inherent lateralization patterns specific to motor imagery.

***Graph index*-1:** local efficiency

local efficiency (LE) measures the network's capacity for regional specialization, by looking at how well connected its sub-networks are. It provides an estimate of fault tolerance, since an abundance of local connections allows the network to route around damage to vertices or edges. Let G_i indicates a sub-network containing all nodes that are neighbors of the *i*-th node, the local efficiency is the averaged efficiencies of sub-network G_i as

$$Le_i = \frac{1}{N_{G_i}\left(N_{G_i}-1\right)}\sum_{j,k\in G_i}\frac{1}{L_{j,k}} \tag{S1}$$

where N_{G_i} denotes the number of nodes in the subgraph G_i. $L_{j,k}$ indicates the length of an edge between node *j* and node *k,* which holds the form as

$$L_{jk} = \begin{cases} 1/c_{jk}, & if \quad c_{jk} \neq 0 \\ 1, & if \quad c_{jk} = 0 \end{cases} \tag{S2}$$

where is the inverse of the defined edge weight between node *j* and node *k*.

***Graph index*-2:** Accumulation coupling strength

Accumulation coupling strength (ACS) measures the total strength that connects one node with another. A higher ACS indicates stronger connections from one node to others, which can be defined as

$$ACS_j = \sum_{i=1}^{N}\left(G_{ij}+G_{ji}\right) \tag{S3}$$

where *G* represents the network, *N* denotes the number of network nodes, *I* and *j* represent node *i* and node *j*, respectively.

***Graph index*-3:** local clustering coefficient

The local clustering coefficient (LCC) of a node in a graph quantifies the ratio of its neighbors belonging to a clique, which can be expressed as

$$C_i = \frac{\frac{1}{2}\sum_{j,h\in N}\left(G_{ij}+G_{ji}\right)\left(G_{ih}+G_{hi}\right)\left(G_{jh}+G_{hj}\right)}{\left(\sum_{j\in N}G_{ij}+\sum_{j\in N}G_{ji}\right)\left(\sum_{j\in N}G_{ij}+\sum_{j\in N}G_{ji}-1\right)-2\sum_{j\in N}G_{ij}G_{ji}} \tag{S4}$$

With the aforementioned graph indexes 1-3, the averaged results of the graph indexes across different datasets are reported in the revised manuscript to further illustrate the effectiveness of the proposed DLap-GCA, and the corresponding results are shown in Fig. S1. To eliminate the baseline differences among different indexes, we further conduct min-max normalization to scale the range of their values between 0 and 1. Through these indexes shown in Fig. S1, we can find that the brain networks estimated by DLap-GCA can effectively capture motor imagery related contralateral connection patterns where the brain regions close to C3 play important roles for right-hand motor imagery task, while the brain regions close to C4 are important for left-hand motor imagery task. This lateralized connection pattern and the crucial brain regions are consistent with the observations reported in various studies [3-5, 7].

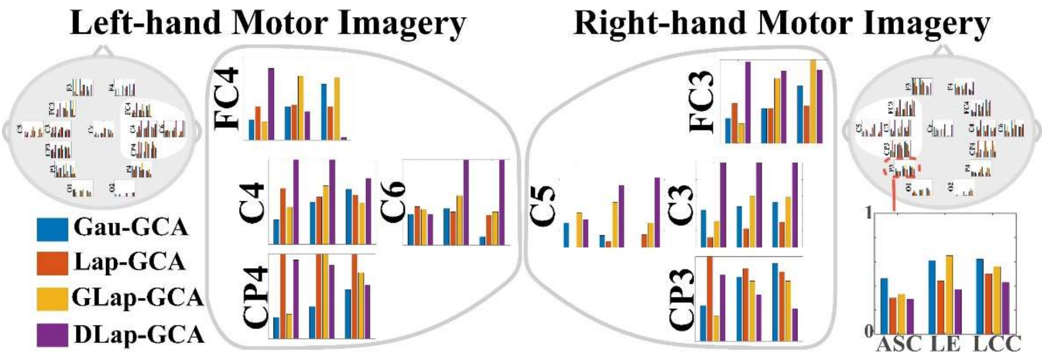

Fig. S1 the average network characters corresponding to left-hand and right-hand MI tasks. The results marked with blue correspond to the brain networks estimated by Gau-GCA. The results marked with red correspond to the brain networks estimated by Lap-GCA. The results marked with yellow correspond to the brain networks estimated by GLap-GCA. The results marked with purple correspond to the brain networks estimated by GLap-GCA.

Through these indices that shown in Fig. S1, we observed that the brain networks estimated by DLap-GCA hold contralateral connection patterns where the brain regions close to C3 play important roles in the right-hand motor imagery task, while the brain regions close to C4 are important in the left-hand motor imagery task.

REFERENCE

[1] B. Blankertz, G. Dornhege, M. Krauledat *et al.*, "The non-invasive Berlin brain–computer interface: fast acquisition of effective performance in untrained subjects," *NeuroImage,* vol. 37, no. 2, pp. 539-550, 2007.

[2] G. Dornhege, B. Blankertz, G. Curio *et al.*, "Boosting bit rates in noninvasive EEG single-trial classifications by feature combination and multiclass paradigms," *IEEE Trans. Biomed. Eng.,* vol. 51, no. 6, pp. 993-1002, 2004.

[3] P. Li, C. Li, J. C. Bore *et al.*, "L1-norm based time-varying brain neural network and its application to dynamic analysis for motor imagery," *J. Neural Eng.,* vol. 19, no. 2, p. 026019, 2022.

[4] F. Li, W. Peng, Y. Jiang *et al.*, "The dynamic brain networks of motor imagery: time-varying causality analysis of scalp EEG," *Int. J. Neural Syst.,* vol. 29, no. 01, p. 1850016, 2019.

[5] X. Gao, W. Huang, Y. Liu *et al.*, "A novel robust Student's t-based Granger causality for EEG based brain network analysis," *Biomed. Signal Proces,* vol. 80, p. 104321, 2023.

[6] A. Gong, J. Liu, S. Chen *et al.*, "Time–frequency cross mutual information analysis of the brain functional networks underlying multiclass motor imagery," *Journal of motor behavior,* vol. 50, no. 3, pp. 254-267, 2018.

[7] P. Li, X. Huang, F. Li *et al.*, "Robust Granger analysis in Lp norm space for directed EEG network analysis," *IEEE Trans. Neural Syst. Rehabil. Eng.,* vol. 25, no. 11, pp. 1959-1969, 2017.