We add a new feature that puts a maximum number of bandwidth K in the resulting estimate. By specifying the value of K, the resulting estimate $\hat{L}$ would be at most K-banded. To achieve this, we first recall that the problem of estimating the r-th row of L

$$\hat{L}_{r,1:r} := \hat{\beta} \in \underset{\beta \in \mathbb{R}^r}{\arg\min} \left\{ -2\log\beta_r + \frac{1}{n}\left\|\mathbf{X}_{1:r}\beta\right\|^2 + \lambda \sum_{\ell=1}^{r-1} \left(\sum_{m=1}^{\ell} w_{\ell m}^2 \beta_m^2 \right)^{1/2} \right\}. \tag{1}$$

Now for $1 \leq r \leq K+1$, the constraint on the bandwidth does not change the optimization problem, and thus is equivalent to its unconstrained version. For $r \geq K+2$, we incorporate the constraint that $\hat{\beta}_{1:(r-K-1)} = 0$. Using the transformation $\ell' = \ell - (r-K-1)$ and $m' = m - (r-K-1)$ for dummy indices ℓ and m in (1), we can equivalently write $\hat{\beta} = \left(\mathbf{0}_{r-K-1}^T, \hat{\theta}^T\right)^T \in \mathbb{R}^r$, where

$$\hat{\theta} \in \underset{\theta \in \mathbb{R}^{K+1}}{\arg\min} \left\{ -2\log\theta_{K+1} + \frac{1}{n}\left\|\mathbf{X}_{(r-K):r}\theta\right\|^2 + \lambda \sum_{\ell'=1}^{K} \left(\sum_{m'=1}^{\ell'} w_{(\ell'+r-K+1)(m'+r-K+1)}^2 \theta_{m'}^2 \right)^{1/2} \right\}. \tag{2}$$

Note that either

$$w_{(\ell'+r-K+1)(m'+r-K+1)} = \frac{1}{(\ell'-m'+1)^2} = w_{\ell'm'} \qquad \text{or} \qquad w_{(\ell'+r-K+1)(m'+r-K+1)} = 1 = w_{\ell'm'}. \tag{3}$$

So (2) is equivalent to solving

$$\hat{\theta} \in \underset{\theta \in \mathbb{R}^{K+1}}{\arg\min} \left\{ -2\log\theta_{K+1} + \frac{1}{n}\left\|\mathbf{X}_{(r-K):r}\theta\right\|^2 + \lambda \sum_{\ell=1}^{K} \left(\sum_{m=1}^{\ell} w_{\ell m}^2 \theta_m^2 \right)^{1/2} \right\},$$

i.e., a standard K-th row problem with the design matrix $\mathbf{X}_{(r-K):r}$.