

Graphic Lasso: Clustering Accuracy

Jiixin Hu

March 10, 2021

The precision model is stated as

$$\mathbb{E}[S^k] = \Omega^k = \sum_{l=1}^r u_{kl} \Theta^l, \quad k \in [K].$$

Consider the following penalized optimization problem

$$\max_{\{\mathbf{U}, \Theta^l\}} \mathcal{L}_S(\mathbf{U}, \Theta^l) = - \sum_{k=1}^K \text{tr}(S^k \Omega^k) + \log \det(\Omega^k) + \lambda \left\| \Omega^k \right\|,$$

where \mathbf{U} is a membership matrix, and $\{\Theta^l\}$ are irreducible and invertible.

1 Notations

Notations.

1. $I'_l = \{k : u'_{kl} \neq 0\}$ is the index set for the l -th group based on the membership \mathbf{U}' .
2. δ be the minimal gap between Θ^l . That is

$$\min_{k, l \in [r]} \left\| \Theta^l - \Theta^k \right\|_F^2 = \delta^2.$$

3. Let $l(\mathbf{U}, \Theta^l)$ be the population-based loss function. That is

$$l(\mathbf{U}, \Theta^l) = \mathbb{E}_S[\mathcal{L}_S(\mathbf{U}, \Theta^l)] = - \sum_{k=1}^K \text{tr}(\Sigma^k \Omega^k) + \log \det(\Omega^k) - \lambda \sum_{k=1}^K \left\| \Omega^k \right\|_1.$$

4. Given the membership \mathbf{U}' , let $\hat{\Theta}^l(\mathbf{U}') = \arg \max_{\Theta^l \in \mathcal{P}_{\Theta, \alpha}} \mathcal{L}_S(\mathbf{U}', \Theta)$. Particularly, for each $l \in [r]$, we have

$$\hat{\Theta}^l(\mathbf{U}') = \arg \max_{\Theta} - \sum_{k \in I'_l} \langle S^k, \Theta \rangle + |I'_l| \log \det(\Theta) - \lambda |I'_l| \left\| \Theta \right\|_1,$$

5. Given the membership \mathbf{U}' , let $\tilde{\Theta}^l(\mathbf{U}') = \arg \max_{\mathbf{U}', \Theta^l} \mathcal{L}_S(\mathbf{U}', \Theta^l)$. Particularly, for each $l \in [r]$, we have

$$\tilde{\Theta}^l(\mathbf{U}') = \arg \max_{\Theta} - \sum_{k \in I'_l} \langle S^k, \Theta \rangle + |I'_l| \log \det(\Theta) - \lambda |I'_l| \left\| \Theta \right\|_1.$$

6. Define functions

$$F(\mathbf{U}') = \mathcal{L}_S(\mathbf{U}', \hat{\Theta}^l(\mathbf{U}')), \quad G(\mathbf{U}') = l(\mathbf{U}', \tilde{\Theta}^l(\mathbf{U}')).$$

7. τ be the maximal singular value of the true precision matrix, i.e., $\tau = \max_{l \in [r]} \varphi_{\max}(\Theta^l)$.

8. τ_l be the minimal singular value of the true precision matrix, i.e., $\tau_l = \min_{l \in [r]} \varphi_{\min}(\Theta^l)$.

2 Main Result

Theorem 2.1 (Clustering accuracy). *Let $\{\mathbf{U}, \Theta^l\}$ denote the true parameters. Consider an estimation of membership $\hat{\mathbf{U}}$ such that $F(\hat{\mathbf{U}}) \geq F(\mathbf{U})$. Assume $\lambda \leq \mathcal{O}(n^{-1/2})$. Then, with high probability tends to 1 as $n \rightarrow \infty$, we have the following bound*

$$\mathbb{P}(MCR(\hat{\mathbf{U}}, \mathbf{U}) \geq \epsilon) \leq C_1 \exp \left\{ -C_2 \frac{\epsilon \delta^2 n}{K p^2 \tau^4} \right\},$$

where C_1, C_2 are two positive constants.

Remark 1. Let $\epsilon = \frac{K p^2 \tau^4 \log n}{\delta^2 n}$. Then, with probability $1 - C_1 \exp \{-C_2 \log n\}$, we have

$$MCR(\hat{\mathbf{U}}, \mathbf{U}) \leq \frac{K p^2 \tau^4 \log n}{\delta^2 n} \approx \mathcal{O}(n^{-1}).$$

This result implies that the MCR of MLE decays at the rate of $\mathcal{O}(n^{-1})$.

Proof. Since the estimate $\hat{\mathbf{U}}$ satisfies $F(\hat{\mathbf{U}}) \geq F(\mathbf{U})$, we have

$$F(\hat{\mathbf{U}}) - F(\mathbf{U}) = F(\hat{\mathbf{U}}) - G(\hat{\mathbf{U}}) + G(\hat{\mathbf{U}}) - G(\mathbf{U}) + G(\mathbf{U}) - F(\mathbf{U}) \geq 0.$$

By Lemma 4, for any $\epsilon > 0$, we have

$$\begin{aligned} \mathbb{P}(MCR(\hat{\mathbf{U}}, \mathbf{U}) \geq \epsilon) &= \mathbb{P} \left(G(\hat{\mathbf{U}}) - G(\mathbf{U}) \leq \epsilon \delta \left(-\frac{1}{8\tilde{\tau}} \delta + \lambda \sqrt{p} \right) \right) \\ &\leq \mathbb{P} \left(0 \leq \epsilon \delta \left(-\frac{1}{8\tilde{\tau}} \delta + \lambda \sqrt{p} \right) + 2m \right), \end{aligned}$$

where $m = \sup_{\mathbf{U}} |F(\mathbf{U}) - G(\mathbf{U})|$. Let $\tilde{t} = \frac{\epsilon \delta}{2} \left(\frac{1}{8\tilde{\tau}} \delta - \lambda \sqrt{p} \right)$. By Lemma 3, we obtain that

$$\begin{aligned} \mathbb{P}(MCR(\hat{\mathbf{U}}, \mathbf{U}) \geq \epsilon) &\leq \mathbb{P}(m \geq \tilde{t}) \\ &\leq \begin{cases} C_1 \exp \{-C_2 n a(\lambda, \tilde{t})^2\} & |a(\lambda, \tilde{t})| \leq 4\tau_l^{-1} \\ C_1 \exp \{-C_2 n a(\lambda, \tilde{t})\} & |a(\lambda, \tilde{t})| > 4\tau_l^{-1} \end{cases}, \end{aligned}$$

where $a(\lambda, t) = \frac{-(2\lambda+1) + \sqrt{(2\lambda+1)^2 - 4(2\lambda^2 - t/Kp^2\tilde{\tau})}}{2}$.

Consider the Taylor Expansion of $a(\lambda, t)$ around $(0, 0)$. We have

$$\begin{aligned} a(\lambda, t) &\approx \nabla a(\lambda, t)(\lambda, t)^T = \left(-2 + \frac{-4\lambda + 2}{\sqrt{-4\lambda^2 + 4\lambda + 1 + 4t/Kp^2\tilde{\tau}}} \right) \lambda + \left(\frac{2/Kp^2\tilde{\tau}}{\sqrt{-4\lambda^2 + 4\lambda + 1 + 4t/Kp^2\tilde{\tau}}} \right) t \\ &= \mathcal{O}(-\lambda) + \mathcal{O} \left(\sqrt{\frac{t}{Kp^2\tilde{\tau}}} \right). \end{aligned}$$

Note that $\tilde{t} = \mathcal{O}(\frac{\epsilon\delta^2}{\tilde{\tau}} - \epsilon\delta\lambda\sqrt{p})$. Plug \tilde{t} in $a(\lambda, t)$. We have

$$a(\lambda, \tilde{t}) = \mathcal{O}(-\lambda) + \mathcal{O}\left(\sqrt{\frac{\epsilon\delta^2}{Kp^2\tilde{\tau}^2} - \frac{\epsilon\delta\lambda}{Kp^{3/2}\tilde{\tau}}}\right).$$

Note that $\lambda = \mathcal{O}(n^{-1/2})$. Therefore, as $n \rightarrow \infty$, we have

$$a(\lambda, \tilde{t}) \rightarrow \mathcal{O}\left(\sqrt{\frac{\epsilon\delta^2}{Kp^2\tilde{\tau}^2}}\right).$$

When K, p^2 are sufficient large, we can consider $a(\lambda, \tilde{t}) \leq 4\tau_l^{-1}$. Then, we have

$$\mathbb{P}(MCR(\hat{U}, U) \geq \epsilon) \leq C_1 \exp\left\{-C_2 \frac{\epsilon\delta^2 n}{Kp^2\tilde{\tau}^2}\right\},$$

where $\mathbb{P}(MCR(\hat{U}, U) \geq \epsilon) \rightarrow 0$ as $n \rightarrow \infty$. □

3 Useful Lemmas

Lemma 1. Let $Z_i \sim_{i.i.d.} \mathcal{N}(0, \Sigma)$ and $\varphi_{\max}(\Sigma) \leq \tau_0 < \infty$. Let $\Sigma = \llbracket \Sigma_{ij} \rrbracket$, then

$$P\left(\left|\sum_{i=1}^n Z_{ij} Z_{ik} - n\Sigma_{jk}\right| \geq n\nu\right) \leq \begin{cases} c_1 e^{-c_2 n\nu^2}, & |\nu| \leq \delta \\ c_1 e^{-c_2 n\nu}, & |\nu| > \delta \end{cases}$$

where c_1, c_2 depend on τ_0 only, and $\delta = 4\tau_0$.

Proof of Lemma 1. See A.J. Rothman et al. Lemma 1. □

Lemma 2. Let $\tau = \max \varphi_{\max}(\Theta)$. We have

$$\text{vec}(\Delta)^T \int_0^1 (1-v)(\Theta + v\Delta)^{-1} \otimes (\Theta + v\Delta)^{-1} dv \text{vec}(\Delta) \geq \frac{1}{4(\tau^2 + \|\Delta\|_F^2)} \|\Delta\|_F^2.$$

Proof of 2. Consider the integral

$$\begin{aligned} & \text{vec}(\Delta)^T \int_0^1 (1-v)(\Theta + v\Delta)^{-1} \otimes (\Theta + v\Delta)^{-1} dv \text{vec}(\Delta) \\ & \geq \|\Delta\|_F^2 \varphi_{\min} \left(\int_0^1 (1-v)(\Theta + v\Delta)^{-1} \otimes (\Theta + v\Delta)^{-1} dv \right) \\ & \geq \|\Delta\|_F^2 \int_0^1 (1-v) \varphi_{\min}^2((\Theta + v\Delta)^{-1}) dv \\ & \geq \frac{1}{2} \min_{\nu \in [0,1]} \varphi_{\min}^2((\Theta + \nu\Delta)^{-1}). \end{aligned}$$

Note that

$$\min_{\nu \in [0,1]} \varphi_{\min}^2((\Theta + \nu\Delta)^{-1}) \geq \min_{\nu \in [0,1]} \varphi_{\max}^{-2}(\check{\Theta} + \nu\Delta) \geq (\|\Theta\|_2 + \|\Delta\|_2)^{-2} \geq \frac{1}{(\tau^2 + \|\Delta\|_F^2)},$$

where the last inequality follows the fact that $\|\Delta\|_2^2 \leq \|\Delta\|_F^2$. □

Lemma 3 (Estimation error). *Given a membership \mathbf{U}' , assume $\lambda \leq \mathcal{O}(n^{-1/2})$. With high probability, we have the following probability*

$$p(t) = \mathbb{P}(|F(\mathbf{U}') - G(\mathbf{U}')| \geq t) \leq \begin{cases} C_1 \exp\{-C_2 n a(\lambda, t)^2\} & |a(\lambda, t)| \leq 4\tau_l^{-1} \\ C_1 \exp\{-C_2 n a(\lambda, t)\} & |a(\lambda, t)| > 4\tau_l^{-1} \end{cases},$$

where $a(\lambda, t) = \frac{-(2\lambda+1) + \sqrt{(2\lambda+1)^2 - 4(2\lambda^2 - t/Kp^2\tau^2)}}{2}$, C_1, C_2 are two constants, and $p(t) \rightarrow 0$ as $t \rightarrow \infty$.

Proof of Lemma 3. With given membership \mathbf{U}' , we have estimations $\hat{\Theta}^l(\mathbf{U}')$ and $\tilde{\Theta}^l(\mathbf{U}')$, which we use $\hat{\Theta}^l$ and $\tilde{\Theta}^l$ refer to them for simplicity, respectively. By the definition, we have

$$\begin{aligned} |F(\mathbf{U}') - G(\mathbf{U}')| &= |\mathcal{L}_S(\mathbf{U}', \hat{\Theta}^l) - l(\mathbf{U}', \tilde{\Theta}^l)| \\ &\leq \sum_{l=1}^r |f^l(\hat{\Theta}^l) - g^l(\tilde{\Theta}^l)|, \end{aligned}$$

where

$$f^l(\Theta) = - \sum_{k \in I'_l} \langle S^k, \Theta \rangle + |I'_l| \log \det(\Theta) - \lambda |I'_l| \|\Theta\|_1,$$

and

$$g^l(\Theta) = - \sum_{k \in I'_l} \langle \Sigma^k, \Theta \rangle + |I'_l| \log \det(\Theta) - \lambda |I'_l| \|\Theta\|_1.$$

Note that the functions $f^l(\cdot)$ and $g^l(\cdot)$ for $l \in [r]$ depends on the membership \mathbf{U}' , and $\hat{\Theta}^l, \tilde{\Theta}^l$ are unique maximizers for $f^l(\Theta), g^l(\Theta)$, respectively.

Next, for an arbitrary $l \in [r]$, we try to find the upper bound for $|f^l(\hat{\Theta}^l) - g^l(\tilde{\Theta}^l)|$. For simplicity, we use $f, g, \hat{\Theta}, \tilde{\Theta}$ denote $f^l, g^l, \hat{\Theta}^l$ and $\tilde{\Theta}^l$. Consider a new estimation $\check{\Theta}$ such that

$$\check{\Theta} = \arg \max_{\Theta} - \sum_{k \in I'_l} \langle \Sigma^k, \Theta \rangle + |I'_l| \log \det(\Theta).$$

By a straight calculation, we have the closed form of $\check{\Theta}$, which is equal to

$$\check{\Theta} = \left(\frac{\sum_{k \in I'_l} \Sigma^k}{|I'_l|} \right)^{-1}.$$

Note that $\check{\Theta}$ is the inverse of a combination of true covariance matrices. Thus $\check{\Theta}$ does not violate the singular value constrain.

Then, we have

$$\begin{aligned} |f(\hat{\Theta}) - g(\tilde{\Theta})| &\leq |f(\hat{\Theta}) - f(\check{\Theta})| + |f(\check{\Theta}) - g(\check{\Theta})| + |g(\check{\Theta}) - g(\tilde{\Theta})| \\ &= M_1 + M_2 + M_3. \end{aligned}$$

1. **For** M_1 , we have

$$f(\hat{\Theta}) - f(\check{\Theta}) = \sum_{k \in I'_l} \langle S^k, \check{\Theta} - \hat{\Theta} \rangle + |I'_l| \left(\log \det(\hat{\Theta}) - \log \det(\check{\Theta}) \right) - \lambda |I'_l| \left(\|\hat{\Theta}\|_1 - \|\check{\Theta}\|_1 \right).$$

Define $\Delta_1 = \hat{\Theta} - \check{\Theta}$ and consider the function $m(t) = \log \det(\check{\Theta} + t\Delta_1)$. By Taylor expansion and Lemma 2, we have

$$\begin{aligned} \log \det(\hat{\Theta}) - \log \det(\check{\Theta}) &= m(1) - m(0) \\ &= \langle \check{\Theta}^{-1}, \Delta_1 \rangle - \text{vec}(\Delta_1)^T \int_0^1 (1-v)(\check{\Theta} + v\Delta_1)^{-1} \otimes (\check{\Theta} + v\Delta_1)^{-1} dv \text{vec}(\Delta_1) \\ &\leq \langle \check{\Theta}^{-1}, \Delta_1 \rangle - \frac{1}{2\tau^2 + 2\|\Delta_1\|_F^2} \|\Delta_1\|_F^2, \end{aligned}$$

Note that $f(\hat{\Theta}) - f(\check{\Theta}) \geq 0$, we have

$$\begin{aligned} 0 &\leq \sum_{k \in I'_l} \langle S^k - \Sigma^k, \Delta_1 \rangle - \frac{1}{2\tau^2 + 2\|\Delta_1\|_F^2} |I'_l| \|\Delta_1\|_F^2 + \lambda |I'_l| \|\Delta_1\|_1 \\ &\leq |I'_l| \max_{(i,j), k \in I'_l} |S_{ij}^k - \Sigma_{ij}^k| \|\Delta_1\|_1 - \frac{1}{2\tau^2 + 2\|\Delta_1\|_F^2} |I'_l| \|\Delta_1\|_F^2 + \lambda |I'_l| \|\Delta_1\|_1 \\ &\leq |I'_l| \left(-\frac{1}{2\tau^2 + 2\|\Delta_1\|_F^2} \|\Delta_1\|_F^2 + \left(\lambda + \max_{(i,j), k \in I'_l} |S_{ij}^k - \Sigma_{ij}^k| \right) p \|\Delta_1\|_F \right), \end{aligned}$$

which implies that

$$0 \leq 2\left(\lambda + \max_{(i,j), k \in I'_l} |S_{ij}^k - \Sigma_{ij}^k| \right) p \|\Delta_1\|_F^2 - \|\Delta_1\|_F + 2\left(\lambda + \max_{(i,j), k \in I'_l} |S_{ij}^k - \Sigma_{ij}^k| \right) p \tau^2. \quad (1)$$

By the assumption that $\lambda = \mathcal{O}(n^{-1/2})$ and the Lemma 1, we know have $\lim_{n \rightarrow \infty} \lambda = \lim_{n \rightarrow \infty} \max_{(i,j), k \in I'_l} |S_{ij}^k - \Sigma_{ij}^k| = 0$. Hence, to satisfies the condition (1), we have

$$\lim_{n \rightarrow \infty} \|\Delta_1\|_F = 0.$$

Therefore, for n sufficient large, we have

$$\begin{aligned} |f(\hat{\Theta}) - f(\check{\Theta})| &\leq |I'_l| \left(-\frac{1}{4\tau^2} \|\Delta_1\|_F^2 + \left(\lambda + \max_{(i,j), k \in I'_l} |S_{ij}^k - \Sigma_{ij}^k| \right) p \|\Delta_1\|_F \right) \\ &\leq |I'_l| \tau^2 p^2 \left(\lambda + \max_{(i,j), k \in I'_l} |S_{ij}^k - \Sigma_{ij}^k| \right)^2 \end{aligned}$$

where the third inequality follows by the fact the $\|\Delta\|_1 \leq p \|\Delta\|_F$, and the last inequality follows by the property of quadratic function.

2. **For** M_2 , we have

$$\begin{aligned} |f(\check{\Theta}) - g(\check{\Theta})| &= \left| \sum_{k \in I'_l} \langle S^k - \Sigma^k, \check{\Theta} \rangle \right| \\ &\leq |I'_l| \left\| S^k - \Sigma^k \right\|_2 \|\check{\Theta}\|_2 \\ &\leq p^2 \tau^2 |I'_l| \max_{(i,j), k \in I'_l} |S_{ij}^k - \Sigma_{ij}^k|. \end{aligned}$$

3. **For** M_3 , we have

$$g(\check{\Theta}) - g(\tilde{\Theta}) = \sum_{k \in I'_l} \langle \Sigma^k, \tilde{\Theta} - \check{\Theta} \rangle + |I'_l| \left(\log \det(\check{\Theta}) - \log \det(\tilde{\Theta}) \right) - \lambda |I'_l| (\|\check{\Theta}\|_1 - \|\tilde{\Theta}\|_1).$$

Let $\Delta_2 = \tilde{\Theta} - \check{\Theta}$. By Taylor Expansion and the fact that $g(\check{\Theta}) - g(\tilde{\Theta}) \leq 0$, with similar procedures for M_1 , we have

$$\begin{aligned} g(\check{\Theta}) - g(\tilde{\Theta}) &\geq \sum_{k \in I'_l} \langle \Sigma^k, \Delta_2 \rangle - |I'_l| (\langle \check{\Theta}^{-1}, \Delta_2 \rangle - \frac{1}{4\tau^2} \|\Delta_2\|_F^2) - \lambda |I'_l| \|\Delta_2\|_1 \\ &= \frac{1}{4\tau^2} |I'_l| \|\Delta_2\|_F^2 - \lambda |I'_l| \|\Delta_2\|_1. \end{aligned}$$

Thus, we have

$$\begin{aligned} |g(\check{\Theta}) - g(\tilde{\Theta})| &\leq -\frac{1}{4\tau^2} |I'_l| \|\Delta_2\|_F^2 + \lambda |I'_l| \|\Delta_2\|_1 \\ &\leq -\frac{1}{\tilde{\tau}} |I'_l| \|\Delta_2\|_F^2 + \lambda |I'_l| p \|\Delta_2\|_F \\ &\leq \tau^2 \lambda^2 p^2 |I'_l| \end{aligned}$$

Therefore, we have the upper bound

$$\begin{aligned} |f(\hat{\Theta}) - g(\tilde{\Theta})| &\leq M_1 + M_2 + M_3 \\ &\leq |I'_l| p^2 \tau^2 \left[(\lambda + \max_{(i,j), k \in I'_l} |S_{ij}^k - \Sigma_{ij}^k|)^2 + \max_{(i,j), k \in I'_l} |S_{ij}^k - \Sigma_{ij}^k| + \lambda^2 \right], \end{aligned}$$

and thus we have

$$\begin{aligned} |F(\mathbf{U}') - G(\mathbf{U}')| &\leq \sum_{l=1}^r |f^l(\hat{\Theta}^l) - g^l(\tilde{\Theta}^l)| \\ &\leq K p^2 \tau^2 \left[(\lambda + \max_{(i,j), k \in K} |S_{ij}^k - \Sigma_{ij}^k|)^2 + \max_{(i,j), k \in K} |S_{ij}^k - \Sigma_{ij}^k| + \lambda^2 \right]. \end{aligned}$$

Intuitively, if λ tends to 0, the error only related to the gap between population and sample $\max_{(i,j), k \in K} |S_{ij}^k - \Sigma_{ij}^k|$.

Last, we obtain the probability

$$\begin{aligned} \mathbb{P}(|F(\mathbf{U}') - G(\mathbf{U}')| \geq t) &\leq \mathbb{P} \left((\lambda + \max_{(i,j), k \in K} |S_{ij}^k - \Sigma_{ij}^k|)^2 + \max_{(i,j), k \in K} |S_{ij}^k - \Sigma_{ij}^k| + \lambda^2 \geq \frac{t}{K p^2 \tau^2} \right) \\ &= \mathbb{P} \left(\max_{(i,j), k \in K} |S_{ij}^k - \Sigma_{ij}^k|^2 + (2\lambda + 1) \max_{(i,j), k \in K} |S_{ij}^k - \Sigma_{ij}^k| + 2\lambda^2 - \frac{t}{K p^2 \tau^2} \geq 0 \right) \\ &= \mathbb{P} \left(\max_{(i,j), k \in K} |S_{ij}^k - \Sigma_{ij}^k| \geq \frac{-(2\lambda + 1) + \sqrt{(2\lambda + 1)^2 - 4(2\lambda^2 - t/K p^2 \tau^2)}}{2} \right). \end{aligned}$$

Let $a(\lambda, t) = \frac{-(2\lambda+1) + \sqrt{(2\lambda+1)^2 - 4(2\lambda^2 - t/Kp^2\tau^2)}}{2}$. Since $\lambda = \mathcal{O}(-1/n)$, $a(\lambda, t)$ is well-defined with large n . By the Lemma 1, we have

$$\begin{aligned} p(t) &= \mathbb{P}(|F(\mathbf{U}') - G(\mathbf{U}')| \geq t) \\ &\leq \mathbb{P}\left(\max_{(i,j), k \in K} |S_{ij}^k - \Sigma_{ij}^k| \geq a(\lambda, t)\right) \\ &\leq \begin{cases} C_1 \exp\{-C_2 na(\lambda, t)^2\} & |a(\lambda, t)| \leq 4\tau_l^{-1} \\ C_1 \exp\{-C_2 na(\lambda, t)\} & |a(\lambda, t)| > 4\tau_l^{-1} \end{cases} \end{aligned}$$

□

Lemma 4 (Self-consistency of \mathbf{U}). *Suppose $MCR(\mathbf{U}', \mathbf{U}) \geq \epsilon$ and the minimal gap between $\{\Theta^l\}$ denoted δ is positive. For $\lambda = \mathcal{O}(n^{-1/2})$, we have the perturbation version of the self-consistency.*

$$G(\mathbf{U}') - G(\mathbf{U}) \leq \epsilon\delta \left(-\frac{1}{8\tau^2}\delta + \lambda\sqrt{p}\right) < 0.$$

Proof. See note 0228.

□