

### Convergence Result of $t$ -LSE

The gradient of  $L(\Phi(V), \varepsilon)$  can be written as the following compact form:

$$\frac{\partial L(\Phi(V), \varepsilon)}{\partial \Phi(V)} = (L_p - L_Q) \Phi(V)$$

where  $L_p$  and  $L_Q$  represent

$$L_p = \text{diag} \left( \sum_j p_{1j}, \sum_j p_{2j}, \dots, \sum_j p_{nj} \right) - P, \quad L_Q = \text{diag} \left( \sum_j q_{1j}, \sum_j q_{2j}, \dots, \sum_j q_{nj} \right) - Q$$

The matrices  $P$  and  $Q$  are defined as

$$p_{ij} = \begin{cases} -4\lambda_{ij} \frac{\partial l_{ij}(\Phi(V), \varepsilon)}{\partial m_{ij}} & e_{ij} \in E \\ 0 & \text{else} \end{cases}, \quad q_{ij} = \begin{cases} 4\lambda_{ij} \frac{\partial l_{ij}(\Phi(V), \varepsilon)}{\partial m_{ij}} & e_{ij} \notin E \\ 0 & \text{else} \end{cases}$$

Here we prove that the search direction  $\Delta = (L_p)^{-1} L_Q \Phi(V) - \Phi(V)$  is a descent direction. The inner-product between  $\Delta$  and  $(L_p - L_Q) \Phi(V)$  is

$$\begin{aligned} \langle \Delta, (L_p - L_Q) \Phi(V) \rangle &= \text{trace}(\Delta^T (L_p - L_Q) \Phi(V)) \\ &= -\text{trace}(\Delta^T L_p \Delta) \end{aligned}$$

Thus we only need to prove that  $L_p$  is positive semi-definite. It is easy to verify that  $L_p$  is symmetric, and for any  $a = [a_1, a_2, \dots, a_n]^T \in \mathbb{R}^n$ ,

$$a^T L_p a = -4 \sum_{i=1}^n \sum_{j=i+1}^n \lambda_{ij} \frac{\partial l_{ij}(\Phi(V), \varepsilon)}{\partial m_{ij}} (a_i - a_j)^2$$

Since  $l_{ij}(\Phi(V), \varepsilon)$  is a decreasing function of  $m_{ij}$ ,  $\lambda_{ij}$  is positive, we have  $a^T L_p a \geq 0$ . Thus we have  $L_p$  is positive semi-definite. Hence, as a result of Zoutendijk's theorem [38], we are guaranteed to converge to a local optimum of  $L(\Phi(V), \varepsilon)$  if we use the search direction in combination with a line-search that satisfies the Wolfe conditions, i.e., a line-search step that simultaneous satisfies the Armijo condition and the curvature condition.