

# Notes about convergence rate

Matt Piekenbrock

If the value  $x^*$  that a sequence  $\{x_n\} = x_1, x_2, \dots, x_n$  approaches exists, then  $\{x_n\}$  is called *convergent*. Moreover, if there exist real numbers  $\mu \leq 1$  and  $\alpha \geq 1$  such that:

$$\lim_{n \rightarrow \infty} \frac{|x_{n+1} - x^*|}{|x_n - x^*|^\alpha} = \mu$$

then sequence is convergent, and the value of  $\alpha$  is called the *rate of convergence*. When  $\mu \in (0, 1)$ , the When  $\alpha = 1, 2$ , and  $3$ , we say the sequence exhibits *linear*, *quadratic*, and *cubic* convergence, respectively. A sequence with rate  $1 < \alpha < 2$  is said to converge *superlinearly*.

## 1 Spectral Sparsifiers & Effective Resistance

Given a weighted undirected graph  $G = (V, E, w)$  with  $w : E \rightarrow \mathbb{R}_+$ , denote with  $L_G$  its weighted *graph Laplacian*  $L_G = D - W$ , where  $D = \text{diag}(\deg^w(v_1), \deg^w(v_2), \dots, \deg^w(v_n))$  is a diagonal (weighted) degree matrix and  $W$  is the weighted adjacency matrix satisfying  $W[i, j] = -w_{ij}$  for  $i \neq j$  and  $v_i \sim v_j$ . An intuitive goal is to choose a sparse subgraph  $H = (V, E')$  with  $E' \subseteq E$  satisfying:

$$(1 - \epsilon)L_G \preceq L_H \preceq (1 + \epsilon)L_G \quad \Leftrightarrow \quad (1 - \epsilon)x^T L_G x \leq x^T L_H x \leq (1 + \epsilon)x^T L_G x$$

If we can find a subgraph  $H$  satisfying this, then  $L_H$  is said to be an  $\epsilon$ -*spectral sparsifier* of  $G$ . Such sparsifiers are naturally very appealing in that they satisfying the following properties:

1. (Cut approximation) The weight of every cut  $w_G(E(S, S'))$  is within  $1 \pm \epsilon$  of  $L_H(E(S, S'))$
2. ( $\kappa$ -approximation)  $(1 - \epsilon)^{-1}H$  is a  $(1 + \epsilon)(1 - \epsilon)^{-1}$ -approximation of  $G$ ;  $H$  is called a  $\kappa$ -*approximation* of  $G$  if:

$$L_G \preceq L_H \preceq \kappa L_G \quad \Leftrightarrow \quad \kappa^{-1}L_G^+ \preceq L_H^+ \preceq L_G^+$$

3. (Spectral approximation) If  $\Lambda(L_G) = \{\lambda_1, \lambda_2, \dots, \lambda_n\}$  and  $\Lambda(L_H) = \{\lambda'_1, \lambda'_2, \dots, \lambda'_n\}$ , then  $(1 - \epsilon)\lambda'_k \leq \lambda_k \leq (1 + \epsilon)\lambda'_k$

Since the definition does not explicitly bound the size of  $E'$ , clearly a subgraph  $H \subseteq G$  always exists. It's not immediately clear whether one can obtain sparser graphs  $H$  with e.g. edge sparsities  $|E'| \sim O(n \log(n) \cdot \epsilon^{-1})$ . Surprisingly, Spielman showed a positive existential result towards this direction: every connected, *unweighted* graph  $G = (V, E)$  admits a *weighted* subgraph  $H = (V, E', w')$  that is an  $\epsilon$ -spectral sparsifier, and the number of edges of these sparsifiers  $|E'|$  is the order of  $O(\epsilon^{-2} n \log n)$ . Moreover, they gave a simple randomized sampling algorithm to obtain such sparsifiers.

The idea of their approach is as follows: Let  $L_e$  denote its restriction to edge  $e \in E$ , i.e. the  $n \times n$  matrix with  $L_e[i, i] = L_e[i, j] = 1$  and  $L_e[i, i] = L_e[j, j] = -1$ . This matrix is given by the outer product  $L_e = (\xi_i - \xi_j)(\xi_i - \xi_j)^T$ , where  $\xi_i$  is the characteristic vector with a 1 in the  $i$ -th component and 0 otherwise. Observe that:

$$L_G = \sum_{e \in E} w_e L_e$$

Now, suppose we have edge probabilities  $\{p_e\}_{e \in E}$  satisfying  $\sum_{e \in E} p_e = 1$  and we construct  $H$  by sampling  $k$  edges iid from  $E$  with respect to these probabilities. By definition of expectation, we have:

$$\mathbb{E}[L_H] = \sum_{e \in E} p_e \cdot \frac{1}{k p_e} L_e = \sum_{e \in E} \frac{1}{k} L_e$$

Thus  $L_H$  approximates  $L_G$  via a sum of iid random matrices. By utilizing random matrix theory, one may readily apply concentration inequalities, such as Chernoff bounds, to bound the degree of the approximation. In particular, if  $L_H^{(i)} \preceq \delta \mathbb{E}[L_H^{(i)}]$  with probability 1 for any  $\delta \geq 1$ , then with  $k \sim O(\delta \cdot \epsilon^{-2} \cdot n \log n)$  edges sampled, one has:

$$P(L_H \subseteq_\epsilon L_G) \geq 1 - 2n\epsilon^{-\epsilon^2 k} / 4\delta$$

Computing the effective resistance of a given edge requires—almost by definition—the solution of a linear system on the graph Laplacian.