Oded (June 14, 2024): On the analysis of the Zig-Zag product (following Salil Vadhan)

Given a (big) D-regular graph $G=(V,E)$, and a (small) d-regular graph $H=([D],F)$, their Zig-Zag product, denoted $G\textcircled{z}H$, consists of the vertex set $V\times[D]$, which is partitioned to D-vertex clouds such that the cloud that corresponds to vertex $v\in V$ is the set of vertices $C_v=\{(v,i):i\in[D]\}$, and edges that correspond to certain 3-step walks (on $G\textcircled{r}H$, as detailed next).

Actually, it is instructive to first consider the graph, denoted $G\textcircled{r}H$, in which copies of H are placed on the clouds (i.e., for every $v\in V$ and $\{i,j\}\in F$ we place the intra-cloud edge $\{(v,i),(v,j)\}$), and edges of G connect the corresponding clouds by using corresponding edges; that is, if $\{u,v\}\in E$ is the $i^{\rm th}$ (resp., $j^{\rm th}$) edge incident at u (resp., at v), then we place the inter-cloud edge $\{(u,i),(v,j)\}$. Note that each vertex in $G\textcircled{r}H$ has d intra-cloud edges and a single inter-cloud edge. Now, the edges of $G\textcircled{z}H$ correspond to 3-step walks in $G\textcircled{r}H$ that start with an intra-cloud edge, then take the (only available) inter-cloud edge, and lastly take some intra-cloud edge; that is, such a generic walk has the form $(v,i)\to(v,j)\to(w,k)\to(w,\ell)$, where $\{i,j\},\{k,\ell\}\in F$ and $\{(v,j),(w,k)\}$ is an inter-cloud edge in $G\textcircled{r}H$ (i.e., $\{v,w\}\in E$ is the $j^{\rm th}$ edge incident at v and the $k^{\rm th}$ edge incident at w).

We shall assume that both G and H are connected and are not bipartite. In that case it is clear that the graph $G\textcircled{r}H$ is also connected and non-bipartite, and it can be shown that also $G\textcircled{z}H$ has these properties. Showing the latter is simpler when assuming that H has self-loops on each vertex.[1]

Our focus is on upper-bounding the convergence rate of random walks on $G\textcircled{z}H$ (aka *second eigenvalue of the corresponding random walk matrix*) in terms of the corresponding rates of the graphs G and H. (Recall that we refer to the eigenvalues of the corresponding normalized adjacency matrices, where the normalization consists of dividing each entry by the degree of the (regular) graph.) The following result and its proof are adapted from Salil Vadhan's survey on Pseudorandomness.

Theorem 1 (an analysis of the Zig-Zag product (Thm. 4.35 in Vadhan's survey))[2]: *Let $\lambda(X)$ denote the convergence rate of a random walk on the connected and non-bipartite graph X. Then, $(1-\lambda(G\textcircled{z}H))\geq(1-\lambda(G))\cdot(1-\lambda(H)^2)$.*

In other words the spectrum gap of $G\textcircled{z}H$ is lower-bounded by in terms of the spectrum gaps of G and H. In particular, for $\lambda(H)\leq\sqrt{1/2}$, we get $(1-\lambda(G\textcircled{z}H))\geq(1-\lambda(G))/2$, and this is the result that is used in our presentation of the log-space UCONN algorithm (of Omer Reingold).[3]

We note that the proceeding version of RVW (*41st FOCS*, 2000) only claims that $\lambda(G\textcircled{z}H)\leq\lambda(G)+\lambda(H)$, whereas the log-space UCONN algorithm (of Omer Reingold) requires $(1-\lambda(G\textcircled{z}H))=\Omega(1-\lambda(G))$ for a suitable fixed H that satisfies $\lambda(H)<1$. (Specifically, for $\beta<1$, we need a d-regular $d^{O(1/(1-\beta))}$-vertex graph H that satisfies $\lambda(H)\leq\beta$.) We shall use the following lemma, which is of independent interest.

[1] We first show that, for every $v\in V$, if (v,i) and (v,j) are neighbors in $G\textcircled{r}H$, then they are connected by an even length path in $G\textcircled{z}H$. This follows by considering the 3-step walks $(v,i)\to(v,j)\to(w,k)\to(w,\ell)$ and $(w,\ell)\to(w,k)\to(v,j)\to(v,j)$ on $G\textcircled{r}H$, where ℓ is an arbitrary neighbor of k in H. Next, we observe that each path (resp., cycle) is G corresponds to a (not necessarily simple) path (resp., cycle) in $G\textcircled{z}H$, and that the parity of the length of the path (resp., cycle) is preserved.

[2] Vadhan, *Pseudorandomness*, Foundations and Thrends in Theoretical Computer Sciwnce, Vol. 21 (1–3), 2012.

[3] Recall that $1-\lambda(G\textcircled{z}H)$ is the spectral gap of $G\textcircled{z}H$, whereas $1-\lambda(G)$ is the spectral gap of G. Indeed, we used the hypothesis that $\lambda(H)\leq\sqrt{1/3}$.

Lemma 2 (expanders "behave" like cliques (Lem. 4.19 in Vadhan's survey)): *Let G be a regular n-vertex graph, and W be the corresponding random walk matrix* (i.e., its normalized adjacency matrix)*. Let J be an n-by-n matrix in which all entries equal $1/n$. Letting $E \stackrel{\rm def}{=} (W-\gamma J)/(1-\gamma)$, it holds that $\lambda(G) \leq 1-\gamma$ if and only if $\|E\| \leq 1$, where*

$$\|M\| \stackrel{\rm def}{=} \max_{\overline{x}\in\mathbb{R}^n:\|\overline{x}\|_2=1} \{\|M\overline{x}\|_2\}.$$

Note that J is the random walk matrix of the n-vertex clique (with a self-loop on each vertex). Hence, the forward direction of Lemma 2 asserts that $\lambda(G) \leq 1-\gamma$ implies that a random walk on G is approximated by a random walk on a clique with an error term that is at most $1-\gamma$ (i.e., $W = \gamma J + (1-\gamma)E$ for $\|E\| \leq 1$).

Proof: We start with a few simple observations. Letting $\overline{u} = (1/n, ..., 1/n)^\perp$ denote the uniform (distribution) vector, it follows that

$$E\overline{u} = \frac{W\overline{u} - \gamma J\overline{u}}{1-\gamma} = \frac{\overline{u} - \gamma\overline{u}}{1-\gamma} = \overline{u}.$$

On the other hand, if the vector $\overline{v}$ is orthognal to $\overline{u}$, then $J\overline{v} = 0$ and $E\overline{v} = W\overline{v}/(1-\gamma)$ is also orthogonal to $\overline{u}$.

Now, assuming $\lambda(G) \leq 1-\gamma$ for any $\overline{v}$ that is orthogonal to the uniform vector $\overline{u}$, we get $\|E\overline{v}\| \leq \|\overline{v}\|$ (because $\|W\overline{v}\| \leq \lambda(G)\cdot\|\overline{v}\|$). Hence, decomposing any vector $\overline{x}$ to its uniform and orthognal components, denoted $\overline{x}^{\|}$ and $\overline{x}^{\perp}$ respectively, we get

$$\begin{aligned}\|E\overline{x}\|_2^2 &= \|E(\overline{x}^{\|} + \overline{x}^{\perp}\|_2^2 \\ &= \|E\overline{x}^{\|}\|_2^2 + \|E\overline{x}^{\perp}\|_2^2 \\ &\leq \|\overline{x}^{\|}\|_2^2 + \|\overline{x}^{\perp}\|_2^2\end{aligned}$$

which equals $\|\overline{x}\|$. On the other hand, if $W = \gamma J + (1-\gamma)E$ such that $\|E\| \leq 1$, then for every $\overline{v} \in \mathbb{R}^n$ that is orthogonal to the uniform vector $\overline{u}$, it holds that $\|W\overline{v}\| = \gamma J\overline{v} + (1-\gamma)E\overline{v} \leq 0 + (1-\gamma)\|\overline{v}\|$, which implies $\lambda(G) \leq 1-\gamma$. ■

Proof of Theorem 1

We denote the random walk matrices of G, H and $G \textcircled{z} H$ by W_G, W_H and $W_{G \textcircled{z} H}$ respectively. Letting M denote the matching (in $G \textcircled{z} H$) defined by the inter-cloud edges, we have

$$W_{G \textcircled{z} H} = (I_n \otimes W_H)M(I_n \otimes W_H) \tag{1}$$

where I_n is the n-by-n identity matrix and $A \otimes B$ denotes the tensor product of matrices (i.e., in the resulting matrix each entry of value σ in the matrix A is replaced by a copy of $\sigma \cdot B$). Hence, $I_n \otimes W_H$ describes a random intra-cloud random step, whereas M describes an inter-cloud step.

Letting h denote the number of vertices in H and using Lemma 2, we get $W_H = (1-\lambda(H)) \cdot J_h + \lambda(H) \cdot E$ such that $\|E_H\| \leq 1$, where J_h is an h-by-h matrix in which all entries equal $1/h$. Hence,

$$I_n \otimes W_H = (1-\lambda(H)) \cdot I_n \otimes J_h + \lambda(H) \cdot I_n \otimes E_H \tag{2}$$

Combining Eq. (1) and Eq. (2), we get

$$\begin{aligned}
W_{G ⓩ H} &= (1-\lambda(H))^2 \cdot (I_n \otimes J_h) M (I_n \otimes J_h) && (3)\\
&\quad + (1-\lambda(H)) \cdot \lambda(H) \cdot (I_n \otimes J_h) M (I_n \otimes E_H) && (4)\\
&\quad + (1-\lambda(H)) \cdot \lambda(H) \cdot (I_n \otimes E_H) M (I_n \otimes J_h) && (5)\\
&\quad + \lambda(H)^2 \cdot (I_n \otimes E_H) M (I_n \otimes E_H) && (6)
\end{aligned}$$

The following two facts can be readily verified.

1. The matrix product in each of the three last lines (i.e., Eq. (4), Eq. (5), and Eq. (6)) has norm at most 1.

   This uses $\|E_H\| \le 1$ as well as $\|A \otimes B\| = \min(\|A\|, \|B\|)$ and $\|AB\| = \|A\| \cdot \|B\|$.

   Using also $\|A+B\| \le \|A\| + \|B\|$, it follows that the norm of the matrix given by the sum of these three rows is at most $2 \cdot (1-\lambda(H)) \cdot \lambda(H) + \lambda(H)^2$, which equals $1-(1-\lambda(H))^2$.

2. The matrix product in the first line (i.e., $(I_n \otimes J_h) M (I_n \otimes J_h)$) equals $W_G \otimes J_h$. It follows that, for every vector $\overline{v}$ that is orthogonal to the uniform vector, it holds that $\|((I_n \otimes J_h) M (I_n \otimes J_h))\overline{v}\|_2$, which equals $\|(W_G \otimes J_h)\overline{v}\|_2$, is upper-bounded by $\lambda(G) \cdot \|\overline{v}\|_2$.

   To see that $(I_n \otimes J_h) M (I_n \otimes J_h) = W_G \otimes J_h$, we first observe that $W_G \otimes J_h = W_{G ⓩ K_h^+}$, where K_h^+ is the h-vertex clique augmented by self-loops (and $W_{G ⓩ K_h^+}$ is the random walk matrix of the graph $G ⓩ K_h^+$). This can be seen by considering the n^2 disjoint h-by-h submatrices that cover each of these nh-by-nh matrices; specifically, such a submatrix is an all-$1/h$ matrix if and only if it corresponds to an edge of G.

   Next, we show that $(I_n \otimes J_h) M (I_n \otimes J_h) = W_{G ⓩ K_h^+}$ by considering the action of these matrices on each nh-long unit vector (equiv., the distribution obtained by starting in any vertex (u,i) and taking a random step on either graphs). Specifically, the random step described by $(I_n \otimes J_h) M (I_n \otimes J_h)$ maps (u,i) to the uniform distribution on the u^{th} cloud, then maps it to specific vertex in a random cloud that neighbors the u^{th} cloud, and finally maps it to the uniform distribution in the latter cloud. But this is exactly the distribution obtained by starting at (u,i) and taking a random step on the graph $G ⓩ K_h^+$. (Indeed, $(I_n \otimes J_h) M (I_n \otimes J_h) = W_{G ⓩ K_h^+}$ is analogous to Eq. (1).)

Using these two observations, for every vector $\overline{v}$ that is orthogonal to the uniform vector, we have

$$\begin{aligned}
\frac{\|W_{G ⓩ H}\overline{v}\|}{\|\overline{v}\|_2} &\le (1-\lambda(H))^2 \cdot \lambda(G) + (1-(1-\lambda(H))^2)\\
&= 1-(1-\lambda(H))^2 \cdot (1-\lambda(G))
\end{aligned}$$

and the claim follows (i.e., $1-\lambda(G ⓩ H))^2 \ge (1-\lambda(H))^2 \cdot (1-\lambda(G)))$.