Reflections in Hilbert Space III: Eigen-decomposition of Szegedy's operator

James Daniel Whitfield*†

March 30, 2012

By three methods we may learn wisdom: First, by reflection, which is the noblest; second, by imitation, which is the easiest and third by experience which is the bitterest. *–Confucius*

Continuing the discussion from the last time, we turn to analytic properties of the bi-involution operators. The mainstay of the analysis is based on Ref. [1]. Using the eigen-decomposition of W, as applied to reversible Markov chains, we show how a quadratic speed up is obtained. Throughout we will draw comparisons to the anaylsis of Grover's operator given in the first lecture to illustrate how Szegedy's work is a broad generalization of Grover's algorithm.

1 General bi-involution operators

In the first lecture, we considered a bi-reflection operator about two one-dimensional subspaces, $|s\rangle\langle s|$ and $|w\rangle\langle w|$.

1.1 General properties of bi-involutions

Suppose we have two projectors of dimension d_a and dimension d_b,

$$P_a = \sum_{x=1}^{d_a} |\phi_x\rangle\langle\phi_x| \tag{1}$$

$$P_b = \sum_{y=1}^{d_b} |\psi_y\rangle\langle\psi_y| \tag{2}$$

*Columbia University, NEC Labs America, Institute for Scientific Interchange

†*Current address:* Institute for Scientific Interchange, Via Alassio 11/c, 10126 Torino - Italy

The identification with Markov chains was done using transition vectors,

$$|\phi_x\rangle = \sum_y \sqrt{p_{xy}}|x\rangle|y\rangle = |x\rangle|p_x\rangle \tag{3}$$

$$|\psi_y\rangle = \sum_x \sqrt{q_{yx}}|x\rangle|y\rangle = |q_y\rangle|y\rangle \tag{4}$$

These vector imbue the following form for the projector and corresponding reflection (shown for 1 and 3),

$$P_a = \sum_x |\phi_x\rangle\langle\phi_x| = \sum_x |x\rangle\langle x| \otimes |p_x\rangle\langle p_x| \tag{5}$$

$$R_A = \sum_{xyy'} |x\rangle\langle x| \otimes \left(2\sqrt{p_{xy}p_{xy'}} - \delta_{yy'}\right) |y\rangle\langle y'| \tag{6}$$

Note how only the right part of the tensor space is affected while the left part remains invariant. We will return to this point next lecture.

Continuing from last time, we defined the half-projectors,

$$A = \sum_x |\phi_x\rangle\langle x| \tag{7}$$

$$B = \sum_y |\psi_y\rangle\langle y| \tag{8}$$

The half projectors share a common domain $\mathcal{H}_v$ which, for now, is arbitrary. The operators A maps onto $\mathcal{H}_A$=span$[\{\phi_x\}]$ and B maps onto $\mathcal{H}_B$=span$[\{\psi_y\}]$. These half-projectors are related to the projectors via $P_a = AA^\dagger$ and $P_b = BB^\dagger$. The unitary bi-reflection operator (Szegedy's operator) is now given by

$$W = (2AA^\dagger - \mathbf{1})(2BB^\dagger - \mathbf{1}) = R_A R_B. \tag{9}$$

For the remainder of the chapter W will denote the bi-reflection operator and it is parameterized by the two input subspaces of the Hilbert space.

Previously, we gave a geometric decomposition of the Hilbert space depicted in Fig. 1. Now we turn to an more quantitative spectral analysis of the W operator.

First, we consider the dimensionality of the spaces. Let N be the dimension of the edge Hilbert space. Define matrix C as the matrix formed by all the vectors spanning $\mathcal{H}_A$ and $\mathcal{H}_B$, $C = [A|B]$, and then the dimension of the trivial space is given by $N_3 = N - \text{rank}(C)$. Finally, $D \equiv A^\dagger B$, the discriminant matrix used in the next section, has dimension S_p by S_q. Therefore D will have $d = min(S_p, S_q)$ singular values only some of which are non-zero. In the next section, we will calculate $2d$ eigenvectors in the active space. Therefore, there are $\text{rank}(C) - 2d$ remaining vectors in the inactive space.

In the next section, we turn to the eigendecomposition of the W operator and lastly, in Section 3, we will consider two examples: Grover's one-dimensional bi-reflections and reversible Markov chains.

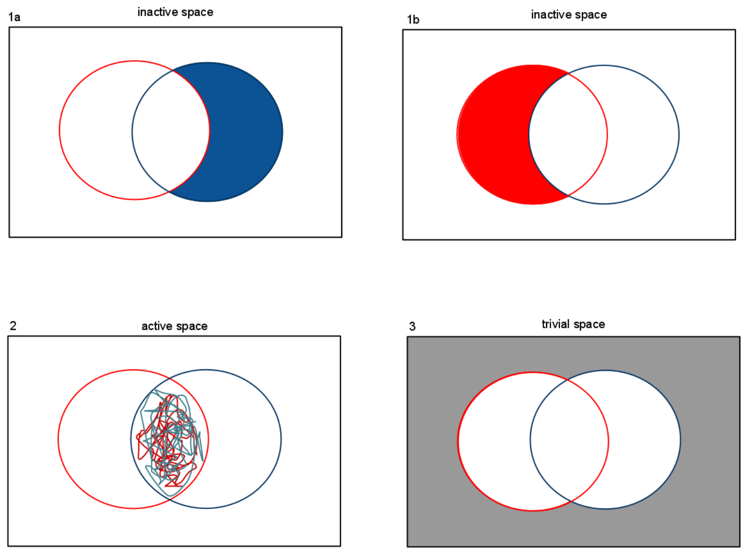

Figure 1: A graphical interpretation of the Hilbert space partitioning determined by the projectors in (1) and (2). Letting $\mathcal{H}_A$ be the red circle and $\mathcal{H}_B$ be the blue circle, figures 1a and 1b depict the inactive space where only sign changes occur, figure 2 is the space where interesting evolution occurs and figure 3 shows the space that is trivial. Using notation from set theory, the inactive space is $\mathcal{H}_A \cup \mathcal{H}_B - \mathcal{H}_A \cap \mathcal{H}_B$, the active space is $\mathcal{H}_A \cap \mathcal{H}_B$ and the trivial space is $(\mathcal{H}_A \cup \mathcal{H}_B)^c$ where A^c denotes the complement of set A.

2 Eigendecomposition of W

2.1 SVD of discriminate matrix

The discriminate matrix is defined as $D \equiv A^\dagger B$ and has a singular value decomposition, $D = U\Sigma V^\dagger$, regardless of the dimension of spaces $\mathcal{H}_A$, $\mathcal{H}_B$, and $\mathcal{H}_v$. For a review of the singular value decomposition see the previous lecture. The left singular vectors $|u_j\rangle$ satisfy,

$$D|v_k\rangle = A^\dagger B|v_k\rangle = \sigma_k|u_k\rangle \tag{10}$$

and the right singular vectors,

$$\langle u_k|D = \langle u_k|A^\dagger B = \langle v_k|\sigma_k, \tag{11}$$

with σ_k the singular value. It is important to point out that $\{v_k\}$ and $\{u_k\}$ are *not in the same Hilbert space* as $|\psi_y\rangle$ and $|\phi_x\rangle$.

The fact that projectors cannot increase length indicates that the singular values of D are less than or equal to unity. Notice that $|Bv| = \sqrt{(B|v\rangle)^\dagger(B|v\rangle)} = \sqrt{\langle v|B^\dagger B|v\rangle} \leq \sqrt{\langle v|v\rangle} = |v|$. We are using a norm based on the inner product of the Hilbert space. By definition $B^\dagger B = \sum_y |y\rangle\langle y|$ is a projector on a space with dimension less than or equal to the full space. Thus, the singular values of $D = A^\dagger B$ must be less than or equal to 1 since $|Bv| \leq |v'|$ and $|A^\dagger v'| \leq |v'|$.

If $\sigma_k = \cos\varphi_k$ then φ_k is called the canonical angle as discuss in the previous lecture and characterizes the 'angles' between subspaces $\mathcal{H}_B$ and $\mathcal{H}_A$.

2.2 Eigenvectors on the active space

The singular vectors of D can be used to characterize the eigenvectors of W on the active space. The (normalized) eigenvectors are

$$|\pm k\rangle = \frac{B|v_k\rangle - e^{\pm i\varphi_k}A|u_k\rangle}{\sqrt{2-2\sigma_k^2}} = \frac{B|v_k\rangle - (\sigma_k \pm i\sqrt{1-\sigma_k^2})A|u_k\rangle}{\sqrt{2-2\sigma_k^2}}. \quad (12)$$

Note that $|\pm k\rangle$ is in the edge space ($\mathcal{H}_A \otimes \mathcal{H}_B$) while $|u_k\rangle$ and $|v_k\rangle$ are in the $\mathcal{H}_A$ and $\mathcal{H}_B$ spaces, respectively. The corresponding eigenvalues are:

$$\exp(\pm 2i\varphi_k) = 2\sigma_k^2 - 1 \pm 2i\sigma_k\sqrt{1-\sigma_k^2}. \quad (13)$$

To demonstrate this, two preliminary relations are established. First,

$$\begin{aligned} WB|v_k\rangle &= R_A R_B B|v_k\rangle = R_A B|v_k\rangle & (14)\\ &= 2A(A^\dagger B)|v_k\rangle - B|v_k\rangle & (15)\\ &= 2\sigma_k A|u_k\rangle - B|v_k\rangle & (16) \end{aligned}$$

In (14), since $B|v_k\rangle$ is in $\mathcal{H}_B$ the reflection R_B acts as the identity.

Next, using (16) and the conjugate transpose of (11),

$$\begin{aligned} WA|u_k\rangle &= R_A(2BB^\dagger - \mathbf{1})A|u_k\rangle & (17)\\ &= R_A[2B(B^\dagger A)|u_k\rangle - A|u_k\rangle] & (18)\\ &= R_A[2\sigma_k B|v_k\rangle] - A|u_k\rangle & (19)\\ &= 2\sigma_k(2\sigma_k A|u_k\rangle - B|v_k\rangle) - A|u_k\rangle & (20)\\ &= (4\sigma_k^2 - 1)A|u_k\rangle - 2\sigma_k B|v_k\rangle & (21) \end{aligned}$$

Now returning to (12),

$$\begin{aligned} & WB|v_k\rangle - e^{\pm i\varphi_k}WA|u_k\rangle & (22)\\ =\ & 2\sigma_k A|u_k\rangle - B|v_k\rangle - e^{\pm i\varphi_k}\left[(4\sigma_k^2 - 1)A|u_k\rangle - 2\sigma_k B|v_k\rangle\right] & (23)\\ =\ & \left\{2\sigma_k e^{\pm i\varphi_k} - 1\right\} B|v_k\rangle - e^{\pm i\varphi_k}\left\{4\sigma_k^2 - 2\sigma_k e^{\mp i\varphi_k} - 1\right\} A|u_k\rangle & (24)\\ =\ & (e^{\pm 2i\varphi_k})[B|v_k\rangle - e^{\pm i\varphi_k}A|u_k\rangle] & (25)\\ =\ & e^{\pm 2i\varphi_k}|\pm k\rangle & (26) \end{aligned}$$

Equation (24) requires several algebraic manipulation of basic trigonometric identities. Specifically, note: $e^{\pm 2i\varphi_k} = (\sigma_k \pm i\sqrt{1-\sigma_k^2})^2 = 2\sigma_k^2 \pm 2i\sigma_k\sqrt{1-\sigma_k^2} -$

1. Expanding both terms in the brackets with $e^{\pm i\varphi_k} = \sigma_k \pm i\sqrt{1-\sigma_k^2}$ one arrives at the desired conclusion.

Now we consider the possible singular values. If $0 < \sigma_k < 1$ then there are two corresponding eigenvectors in the active space. However, when $\sigma_k = 1$ there is no active space as $\varphi_k = 0$ and geometrically the spaces are collinear. When $\sigma_k = 0$, the spaces are orthogonal and again the active space cannot contain any non-trivial trajectories.

2.2.1 The generator of evolution

Summarizing the results of this section, any bi-involution operator can be written

$$W = \sum_{k\in\mathcal{H}_A\cap\mathcal{H}_B} e^{\pm 2i\varphi_k}|\pm k\rangle\langle\pm k| + \sum_{z\in\mathcal{H}_A^c\cap\mathcal{H}_B^c} |z\rangle\langle z| - \sum_{c\in\mathcal{H}_A\cup\mathcal{H}_B-\mathcal{H}_A\cap\mathcal{H}_B} |c\rangle\langle c| \quad (27)$$

Since W is unitary,

$$W = e^{-iH_w} = \sum_k e^{-i\lambda}|\lambda_k\rangle\langle\lambda_k| = e^{-i(\sum_k \lambda_k|\lambda_k\rangle\langle\lambda_k|)}, \quad (28)$$

and its Hamiltonian generator is

$$H_w = \sum_{\mathcal{H}_A\cup\mathcal{H}_B-\mathcal{H}_A\cap\mathcal{H}_B} \pi|c\rangle\langle c| + 2\sum_{\mathcal{H}_A\cap\mathcal{H}_B} (\mp 2\varphi_k)|\pm k\rangle\langle\pm k|. \quad (29)$$

3 Basic examples

3.1 Grover's operator

The canonical angle for the Grover's operator is obtained by inspection of Figure 2 from Lecture 1. With the ϕ_r as the angle depicted, the canonical angle for the Grover operator is $\varphi = \pi/2 - \phi_r$. Using (12), the eigenvector of Grover's operator is

$$\begin{aligned}
|\pm\rangle &= \frac{|w\rangle - (\sin\varphi \pm i\cos\varphi)|s\rangle}{\sqrt{2}\cos\varphi} && (30)\\
&= \frac{|w\rangle}{\sqrt{2}\cos\varphi} - \frac{\sin\varphi \pm i\cos\varphi}{\sqrt{2}\cos\varphi}(\cos\varphi|r\rangle + \sin\varphi|w\rangle) && (31)\\
&= \frac{1-\sin^2\varphi \mp i\cos\varphi\sin\varphi}{\sqrt{2}\cos\varphi}|w\rangle \mp i\frac{\cos\varphi \mp i\sin\varphi}{\sqrt{2}}|r\rangle && (32)\\
&= \frac{e^{\mp i\varphi}}{\sqrt{2}}(|w\rangle \pm i|r\rangle) && (33)
\end{aligned}$$

Here the definition of $|r\rangle$ in (L1.11) and identity $\sin^2\theta + \cos^2\theta = 1$ were used. Examining the rotation matrix in (L1.15), it can easily be seen that this is correct by inspection.

3.2 Mixing times and gaps

The mixing time of a Markov matrix is governed by the inverse of the eigenvalue gap between the highest and the second highest eigenvalues. This characterization is especially fruitful if the dynamics are not periodic and a path exists between all sites in space since in this case the equilibrium state is unique and has strictly positive entries.

This relaxation time also governs the speed of Markovian search algorithms [2, 1, 3, 4], the complexity of simulated annealing [5] and Metropolis type algorithms [6, 7]. The eigenvalue gap of a Markov chain can also be related to the eigenvalue gap between the ground state and first excited state along adiabatic paths [8].

Using the scheme of Szegedy, many quantum algorithms have been developed which improve upon these applications of Markov chain methods each obtaining a quadratic improvement of the algorithm due to a quadratic opening of the eigenvalue gap of the quantum Markov chain [1, 9, 5, 3, 4]. Briefly, let us review relevant results from Markov theory before giving examples of their quantization.

3.3 Reversible Markov matrices

Let P be a time reversible Markov transition matrix with matrix elements $P_{ij} = prob(s_i \to s_j)$. By reversible we mean that the reversed chain $P^{(*)}$ exists with

$$\xi_i P^{(*)}_{ij} = \xi_j P_{ji} \tag{34}$$

where ξ is the left eigenvector of P with unit eigenvalue (the equilibrium or steady-state distribution). The naming comes from the fact that the reversed Markov chain $P^{(*)}$ preserves the same equilibrium distribution. Since one must divide by the elements of the equilibrium probability vector, it must have all non-zero entries i.e. the chain must be ergodic.

Now, order the eigenvalues $\{\lambda_i\}$ of P in decreasing magnitude and let $P = \lambda_i |n_i\rangle\langle n_i|$. Each eigenvalue has magnitude less than one (assuming the chain is aperiodic) due to the properties of Markov chains (cf. Perron-Frobenius theorem [10]). Moreover, $\lambda_1 = 1$ corresponds to the equilibrium distribution $|n_1\rangle = |\xi\rangle$. By decomposing the Markov chain as

$$P^t_{xy} = \sqrt{\frac{\xi_j}{\xi_i}} \sum \lambda_i^t |u_i(x)\rangle\langle u_i(y)| \tag{35}$$

with $u_i(x) = \sqrt{\xi_x} n_i(x)$ the eigenvectors of $S = \sqrt{\text{diag}(\xi)} P \sqrt{\text{diag}(\xi)^{-1}}$. The decay of the eigenvectors corresponding to eigenvalues less than one is dominated by λ_2. In fact, the convergence time to the equilibrium distribution is proportional to $(1 - \lambda_2)^{-1}$.

3.4 Quantization

Now let us consider the quantization of an ergodic, aperiodic Markov chain P and its reverse Markov chain, $P^{(*)}$. The discriminate matrix is

$$D_{xy} = \sqrt{p_{xy}p_{yx}^{(*)}} = p_{xy}\sqrt{\frac{\xi_x}{\xi_y}}. \tag{36}$$

where we used the definition of the reversed chain (34). Note that the discriminant is equal to $S = \sqrt{\mathrm{diag}(\xi)}P\sqrt{\mathrm{diag}(\xi)^{-1}} = D$ introduced in the last subsection. Since we assumed that P is ergodic, ξ has full support and $\sqrt{\mathrm{diag}(\xi)}$ can be inverted. Notice that $\mathrm{diag}(\xi)$ is a matrix.

The eigenvalue decomposition of $D = \lambda_k|w_k\rangle\langle w_k|$ has the eigenvalues of P and the eigenvectors are given by the re-weighting of the k-th eigenvector of P with ξ:

$$\langle x|w_k\rangle = w_k(x) = \xi_x n_k(x) = \xi_x\langle x|n_k\rangle$$

Since each λ_k is real, the singular values are the same as eigenvalues. Thus the canonical angles are given by $\lambda_k = \cos\varphi_k$.

We will use these canonical angles to show that the Szegedy operator has a quadratically larger eigenvalue gap than the original Markov chain. We need to first show two inequalities then the bound will follow.

The first inequality we have to demonstrate is

$$2\theta \geq 2\sqrt{1-\cos^2\theta} \tag{37}$$

First we show: $2\theta \geq 2\sqrt{1-\cos^2\theta}$ by bounding the slopes and then showing the end points also satisfy the constraint. The slope of the right and left hand sides are given by (RHS) $= 2\frac{d}{d\theta}\sqrt{1-\cos^2\theta} = 2\cos\theta$ and (LHS) $= 2\frac{d}{d\theta}\theta = 2$. Thus the left hand side grows faster or at the same rate as the right hand side. At the end points of the relevant domain, $[0, \frac{\pi}{2})$, we have at $\theta = 0$: $LHS = RHS = 0$ and at $\frac{\pi}{2}$ we have $LHS = \pi$ and $RHS = 2$. Thus, $RHS \geq LHS$ and we have proven that $2\theta \geq 2\sqrt{1-\cos^2\theta}$.

The second inequality to be shown is

$$\sqrt{1-cos^2\theta} \geq \sqrt{1-\cos\theta} \tag{38}$$

Since $\cos\theta \leq 1$:

$$\begin{aligned} \cos^2\theta &\leq \cos\theta && (39)\\ -\cos^2\theta &\geq -\cos\theta && (40)\\ 1-\cos^2\theta &\geq 1-\cos\theta && (41)\\ \sqrt{1-\cos^2\theta} &\geq \sqrt{1-\cos\theta} && (42) \end{aligned}$$

The gap of the Markov chain P is by $1-\lambda_2 = 1-\cos\varphi_2$. Finally, since the maximum eigenvalue is unity the canonical angle is $\varphi_1 = 0$. From (12), the eigenvalue gap of W is, $|e^{\pm i\varphi_1} - e^{\pm i\varphi_2}| = |1 - e^{\pm i\varphi_2}| = \sqrt{2}\sqrt{1-\cos(\varphi_2)}$ Using the inequalities (37) and (38) just derived,

$$2\varphi_2 \geq 2\sqrt{1-\cos(\varphi_2)} \tag{43}$$

Thus, we see a quadratic improvement in algorithms that depend on the eigenvalue gap of the Markov chain. In the next lecture, we will continue with more interesting applications, examples, and future directions for research.

References

[1] M. Szegedy. Quantum speed-up of Markov chain based algorithms. In *Proc. 45th Ann. IEEE Symp. Found. Comput. Sci.*, pages 32–41, 2004. http://www.cs.rutgers.edu/ szegedy/PUBLICATIONS/walk_focs.pdf.

[2] L. K. Grover. A fast quantum mechanical algorithm for database search. *Proc. 28th ACM Symp. on Theory of Comp. (STOC '96)*, page 212, 1996. Also see arxiv:quant-ph/9605043.

[3] 5th TAMC, LNCS. *Quantum walk based search algorithms*, volume 4978, 2008.

[4] F. Magniez, A. Nayak, J. Roland, and M. Santha. Search via quantum walk. *SIAM J. Comp.*, 40:142–164, 2011.

[5] R. Somma, S. Boixo, H. Barnum, and E. Knill. Quantum simulations of classical annealing processes. *Phys. Rev. Lett.*, 101:130504, 2008.

[6] K. Temme, T. J. Osborne, K. G. Vollbrecht, D. Poulin, and F. Verstraete. Quantum Metropolis sampling. *Nature*, 471:87, 2011.

[7] M.-H. Yung and A. Aspuru-Guzik. A quantum-quantum Metropolis algorithm. *Proc. Natl. Acad. Sci. USA*, 109:754–759, 2012.

[8] D. Aharonov and A. Ta-Shma. Adiabatic quantum state generation. *SIAM J. Comput.*, 37:47–82, 2008. Also see arxiv:quant-ph/0301023.

[9] M. Szegedy. Spectra of quantized walks and a $\sqrt{}$ rule. *arXiv*, 2004. http://arxiv.org/pdf/quant-ph/0401053v1.pdf.

[10] Perron-Forbenius theorem on Wikipedia. http://en.wikipedia.org/wiki/Perron-Frobenius_theorem.