# Probabilistic Logarithmic-Space Algorithms for Laplacian Solvers

**Dean Doron**[*1], **François Le Gall**[†2], **and Amnon Ta-Shma**[‡1]

1 **The Blavatnik School of Computer Science, Tel-Aviv University, Tel Aviv, Israel**
deandoron@mail.tau.ac.il, amnon@tau.ac.il

2 **Graduate School of Informatics, Kyoto University, Kyoto, Japan**
legall@i.kyoto-u.ac.jp

## Abstract

A recent series of breakthroughs initiated by Spielman and Teng culminated in the construction of nearly linear time Laplacian solvers, approximating the solution of a linear system $\mathcal{L}x = b$, where $\mathcal{L}$ is the normalized Laplacian of an undirected graph. In this paper we study the *space complexity* of the problem. Surprisingly we are able to show a probabilistic, *logspace* algorithm solving the problem. We further extend the algorithm to other families of graphs like Eulerian graphs (and directed regular graphs) and graphs that mix in polynomial time.

Our approach is to pseudo-invert the Laplacian, by first "peeling-off" the problematic kernel of the operator, and then to approximate the inverse of the remaining part by using a Taylor series. We approximate the Taylor series using a previous work and the special structure of the problem. For directed graphs we exploit in the analysis the Jordan normal form and results from matrix functions.

**1998 ACM Subject Classification** F.2.1 Numerical Algorithms and Problems

**Keywords and phrases** Laplacian solvers, Randomized logspace, Bounded-space complexity classes, Random walks, Matrix computation

**Digital Object Identifier** 10.4230/LIPIcs.APPROX/RANDOM.2017.41

## 1 Introduction

Approximating the solution of a linear system $\mathcal{L}x = b$, where $\mathcal{L}$ is the normalized *Laplacian* of a graph $G$, is an important algorithmic challenge with multitude of algorithmic applications (see [39] and references therein). In the time-bounded setting this problem has drawn a lot of attention over the past decade. A series of breakthroughs initiated by Spielman and Teng culminated in the construction of almost linear-time algorithms [24, 29, 33, 34, 35, 36].

We are interested in studying the *space* complexity of this problem, and specifically achieving a probabilistic logspace algorithm that approximates a solution to such a system. We show that the class BPL is powerful enough to approximate the solution to a linear

* Supported by the Israel science Foundation grant no. 994/14 and by the United States – Israel Binational Science Foundation grant no. 2010120. This work was done in part while the author was visiting Kyoto University.

† Supported by the Grant-in-Aid for Young Scientists (A) No. 16H05853, the Grant-in-Aid for Scientific Research (A) No. 16H01705, and the Grant-in-Aid for Scientific Research on Innovative Areas No. 24106009 of the Japan Society for the Promotion of Science and the Ministry of Education, Culture, Sports, Science and Technology in Japan.

‡ Supported by the Israel science Foundation grant no. 994/14 and by the United States – Israel Binational Science Foundation grant no. 2010120.

system of equations for a wide and important variety of linear operators, and in particular for Laplacians of undirected graph (which is the focus of the work of Spielman and Teng). In fact we do more and approximate a *generalized inverse* of the Laplacian, i.e., a matrix $\mathcal{L}^\star$ such that $\mathcal{L}\mathcal{L}^\star\mathcal{L} = \mathcal{L}$, which is sufficient for solving such a set of equations. In essence this means that we invert the matrix on the subspace defined by its image, leaving the kernel unchanged. We prove:

▶ **Theorem 1.** *There exists a probabilistic algorithm that gets as input an $n \times n$ stochastic matrix $\mathcal{S}$ that is the transition matrix of an undirected graph and desired accuracy and confidence parameters $\varepsilon, \delta > 0$, and outputs with probability at least $1 - \delta$ an approximation of the generalized inverse $\mathcal{L}^\star = (\mathcal{I} - \mathcal{S})^\star$ to within an $\varepsilon$-accuracy, using*

$$O\left(\log \frac{n}{\varepsilon} + \log\log \frac{1}{\delta}\right)$$

*space.*

We are not aware of any previous space bounded algorithm approximating the solution of Laplacian systems.

It is commonly believed that $\mathsf{BPL} = \mathsf{L}$.[1] There are not too many natural, non-trivial problems in $\mathsf{L}$, with the exception of undirected st-connectivity (and the problems that reduce to it [26, 2]) that was solved by Reingold with an intricate and beautiful algorithm [30] (see also [38]). The situation is similar with $\mathsf{BPL}$. Thus the fact that probabilistic logspace algorithms are capable of approximating a solution to a large class of linear-algebra problems comes as a surprise.[2]

We now proceed to discuss our technique. Our goal is to approximate $f(\mathcal{S})$ where $f$ is the function corresponding to the generalized inverse of $\mathcal{I} - \mathcal{S}$. We begin by considering the simpler case where $f$ has a Taylor expansion.

Let $G$ be a regular undirected graph with an associated transition matrix $\mathcal{S}$. As $G$ is undirected and regular, $\mathcal{S}$ is normal and we can represent it as $\mathcal{S} = V\mathbf{\Sigma}V^\dagger$ where $\mathbf{\Sigma}$ is a diagonal matrix with the eigenvalues of $\mathcal{S}$ lying on the diagonal. Consider a function $f$ with a Taylor expansion $f(x) = \sum_i c_i x^i$. We would like to approximate $f(\mathcal{S}) = \sum_i c_i \mathcal{S}^i = Vf(\mathbf{\Sigma})V^\dagger$.[3] Using Taylor expansion in the space-bounded setting is appealing, as in $\mathsf{BPL}$ we can approximate powers of stochastic matrices (in fact, even matrices with induced $\ell_\infty$ norm of at most 1 [17]). Hence, if the series expansion of $f$ behaves "nicely", we can also approximate $f(\mathcal{S})$ in $\mathsf{BPL}$. Using this approach we can, e.g., approximate the matrix $e^{\mathcal{S}}$ using the Taylor expansion $e^x = \sum_{i=0}^{\infty} \frac{x^i}{i!}$.

We now consider the real problem which is approximating a generalized inverse of the Laplacian $\mathcal{L} = \mathcal{I} - \mathcal{S}$. This means that we want to invert $\mathcal{L} = \mathcal{I} - \mathcal{S}$ on its image, leaving the kernel unchanged. Thus, the function $f$ we want to compute is $\frac{1}{1-x}$ when $x \neq 1$ and 1 otherwise (think of $x$ here as an eigenvalue of $\mathcal{S}$). The function $f$ is not continuous and

---

[1] Some support for this conjecture is given by the following results. Nisan [27] constructed a pseudorandom generator against logspace-bounded non-uniform algorithms that uses seed length $O(\log^2 n)$. Using that he showed $\mathsf{BPL}$ is contained in the class having simultaneously polynomial time and $O(\log^2 n)$ space [28]. Saks and Zhou [31] showed that $\mathsf{BPL}$ is contained in $\mathsf{DSPACE}(\log^{1.5} n)$. Reingold [30] showed undirected st-connectivity can be solved in deterministic logspace. $\mathsf{BPL} = \mathsf{L}$ is also implied by the conjectured existence of certain circuit lower bounds [25].

[2] Note that finding the *exact* inverse of a matrix, as well as many other important problems in linear algebra, is *complete* for the class $\mathsf{DET} \subseteq \mathsf{NC}^2$ – the class of languages that are $\mathsf{NC}^1$ Turing-reducible to computing the determinant of an integer matrix [5, 7, 14].

[3] The fact that $\sum_i c_i \mathcal{S}^i = Vf(\mathbf{\Sigma})V^\dagger$ is a theorem, see, e.g., [23].

so does not have a Taylor series around 1. Also notice that the operator $\mathcal{L}$ always has a non-trivial kernel (1 is always an eigenvalue of $\mathcal{S}$). Thus, we cannot directly employ the Taylor series approach.

Our solution to the problem is to first "peel-off" the 1-eigenspace using the stationary distribution of the corresponding random walk on $G$. We are then left with an invertible operator $\mathcal{I} - \mathcal{A}$ whose eigenvalues are bounded away from 0. We now wish to use the Taylor series approach and approximate $(\mathcal{I} - \mathcal{A})^{-1}$ by $\sum_{i=0}^{\infty} \mathcal{A}^i$, which corresponds to the Taylor series $\frac{1}{1-x} = \sum_{i=0}^{\infty} x^i$. There is yet one obstacle we need to overcome, which is that the operator $\mathcal{A}$ that we get after peeling off the stationary distribution of $G$, is *not* stochastic, and in fact has $\ell_\infty$ norm larger than 1. Thus, offhand, we do not necessarily know how to simulate high powers of it in BPL. Nevertheless, we exploit its unique structure and show it can be simulated in BPL. Finally, by recovering the peeled-off layer, we essentially recover the required operator $\mathcal{L}^\star$.

We now take a step further, and consider *directed* graphs. The directed case poses major challenges, even if just for the mere fact that directed graphs are not necessarily diagonalizable. In fact, even directed graphs with a favorable structure such as vertex-transitive graphs can be non diagonalizable [20]. The directed Laplacian and its application were studied in, e.g., [6, 12, 3]. Recently, Cohen et al. [13] gave faster algorithms for computing fundamental quantities associated with random walks on directed graphs by improving the running time of solving directed Laplacian systems.

Any operator $\mathcal{A}$ can be represented by its singular value decomposition (SVD) $\mathcal{A} = U\mathbf{\Sigma}V$, where $U$ and $V$ are unitary, and $\mathbf{\Sigma}$ is diagonal with the singular values on the diagonal. Another representation of $\mathcal{A}$ is by its *Jordan normal form*, $\mathcal{A} = V\mathbf{A}V^{-1}$, where $V$ is a basis and $\mathbf{A}$ is the matrix of Jordan blocks. The elements on the diagonals of the Jordan blocks are the eigenvalues of $\mathcal{A}$ (with multiplicity as the multiplicity of the roots of its characteristic polynomial). The SVD is the usual representation of choice as it is stable, whereas the Jordan normal form is notoriously unstable to compute (see, e.g., [22, Chapter 7], [15, Chapter 4] and [19]). However, the SVD representation is not convenient when considering BPL algorithms, as $\mathcal{A}$ does not share the same singular vectors with powers of $\mathcal{A}$. Thus, in this paper, we choose to analyze our algorithm using the Jordan normal form. Admittedly, one should expect severe stability problems using such an approach. Surprisingly, we show that under mild conditions we manage to overcome these stability problems.

As before, we would like to approximate the generalized inverse $\mathcal{L}^\star$. There are two main issues to consider:

1. Peeling-off the 1-subspace. To do so, we need a good approximation of the stationary distribution of the corresponding random walk. In the undirected case, it can be easily inferred (i.e., in L) from the input. Here, we require it as an input to our algorithm.
2. Analyzing the convergence of the Taylor series of $(\mathcal{I} - \mathcal{A})^{-1}$ for a non diagonalizable $\mathcal{A}$. Recall that when a function $f$ acts on a diagonalizable matrix $\mathcal{A}$, it acts on its eigenvalues in the natural way. In the non diagonalizable case, $f$ acts on a Jordan *block*, which might have a large dimension, and although an eigenvalue $\lambda$ on the diagonal is still mapped to an eigenvalue $f(\lambda)$, the structure of the rest of the block is no longer maintained, so we need to give this issue further consideration.

To address the second issue above, we use the theory of matrix functions that tells us exactly what $f(\mathbf{A})$ is. It turns out that there is a direct connection between $f(\mathbf{A})$, the dimension of the Jordan block, and the derivatives of $f$ on the corresponding eigenvalue. Exploiting this connection, we manage to bound the number of terms in the Taylor series that is sufficient for convergence. The caveat here is that two "stability" parameters enter

the picture. First, the spectral gap (whose formal definition we defer), which for directed graphs may no longer be at most polynomially-small and naturally affect the performance of our algorithm. Second, we also need the Jordan basis matrix $V$ of $\mathcal{L}$ to be well-conditioned. We prove:

▶ **Theorem 2** (Informal). *There exists a probabilistic algorithm that gets as input an $n \times n$ stochastic matrix $\mathcal{S}$, desired accuracy and confidence parameters $\varepsilon, \delta > 0$, $\gamma > 0$ which is a lower-bound on the spectral gap of $\mathcal{S}$, $\kappa$ which is an upper bound on the condition number of the Jordan basis of $\mathcal{S}$, and outputs with probability at least $1 - \delta$ an approximation of $\mathcal{L}^\star = (\mathcal{I} - \mathcal{S})^\star$ to within an $\varepsilon$-accuracy, using*

$$O\left(\log \frac{n}{\gamma\varepsilon} + \log\log \frac{\kappa}{\delta}\right)$$

*space.*

Remarkably, the dependency of the space complexity on the condition number of the Jordan basis matrix is *doubly-logarithmic*. This also allows us to show our algorithm operates well on operators for which the eigenvalues are polynomially far apart (see Theorem 29).

Having this theorem we show that in addition to undirected graphs, our approximation algorithm works for well-conditioned regular and Eulerian directed graph (which we know have a non-negligible spectral gap and their stationary distribution is fully-explicit) and general well-conditioned rapidly-mixing directed graphs. We thus see that the algorithm manages to approximate the solution of Laplacian systems over a large (and natural) class of directed graphs.

We conclude with a more philosophical note. In recent years we have seen several results showing that some natural linear-algebraic tasks capture the strength of various space-bounded models of computation. Results along this line are:

1. Ta-Shma [37] showed that it is possible to approximate the SVD of any matrix, and in particular to approximate its inverse, in BQL, with *polynomially-small* accuracy.[4] As no classical analogue is known, this result is one of the very few cases where a natural problem is known to lie in BQL but is not known to be in BPL.
2. Doron et al. [16] gave a BPL algorithm that computes the eigenvalues of *stochastic* matrices having real eigenvalues with *constant* accuracy. Moreover, they gave a linear-algebraic problem which is *complete* for BPL – roughly speaking, approximating, to polynomially-small accuracy, the second eigenvalue of a stochastic matrix (whose eigenvalues are not necessarily real).
3. Fefferman and Lin [18] gave two complete problems for BQL – approximating the inverse and the minimum eigenvalue of positive semi-definite matrices (both to polynomially-small accuracy).

We hence see that the deterministic, probabilistic and quantum space-bounded complexity classes can be roughly characterized by linear-algebraic promise problems, where the difference between the classes lies in the family of operators they can handle, being Hermitian, stochastic or general operators. The *exact* computation can be done in $\mathsf{DET} \subseteq \mathsf{NC}^2 \subseteq \mathsf{DSPACE}(O(\log^2 n))$. Our result is in line with the above, showing that approximating with polynomially-small accuracy the generalized inverse of a large class of *stochastic* matrices is in BPL.

[4] Roughly, BQL stands for the class of languages for which there exists an L-uniform family of quantum circuits solving it with only $O(\log n)$ qubits. It is known that $\mathsf{BQL} \subseteq \mathsf{NC}^2$ [40].

# 2 Preliminaries

## 2.1 Basic facts from linear algebra

For a matrix $\mathcal{A} \in \mathbb{C}^{n\times n}$, $\mathcal{A}^\dagger$ is its conjugate transpose. When it might not be clear from the context, for a vector $v \in \mathbb{C}^n$, we denote $|v\rangle$ as the column vector and $\langle v|$ as the row vector, so $\langle u|\, v\rangle$ is a scalar and $|v\rangle\langle u|$ is a rank-one matrix.

Every matrix $\mathcal{A}$ has a *singular value decomposition* (SVD) $\mathcal{A} = U\mathbf{\Sigma}V^\dagger$, where $U$ and $V$ are unitary and $\mathbf{\Sigma}$ is a diagonal matrix with non-negative entries, known as the singular values of $\mathcal{A}$.

The *spectrum* of a matrix $\mathcal{A}$, denoted $\mathsf{Spec}(\mathcal{A})$, is its set of (complex or real) eigenvalues. The *spectral radius* $\rho(\mathcal{A})$ of $\mathcal{A}$ is the largest absolute value of its eigenvalues. The *operator norm* $\|\mathcal{A}\|$ is $\max_{\|x\|_2=1} \|\mathcal{A}x\|$, which is also the largest singular value of $\mathcal{A}$. Notice that it is possible for $\|\mathcal{A}\|$ to be strictly larger than $\rho(\mathcal{A})$. The operator norm is sub-multiplicative. When $\mathcal{A}$ is invertible, $\kappa(\mathcal{A}) = \|\mathcal{A}\| \left\|\mathcal{A}^{-1}\right\|$ is its *condition number*. Also, we denote $\|\mathcal{A}\|_\infty$ as the induced $\ell_\infty$ norm, that is $\|\mathcal{A}\|_\infty = \max_{i\in[n]} \sum_{j\in[n]} |\mathcal{A}[i,j]|$. It holds that $\|\mathcal{A}\|_\infty \le \sqrt{n}\,\|\mathcal{A}\|$.

For an eigenvalue $\lambda$ of $\mathcal{A}$, a $\lambda$-right-eigenvector (or simply an eigenvector with eigenvalue $\lambda$) is a vector $v$ such that $\mathcal{A}v = \lambda v$. A $\lambda$-left-eigenvector is a vector $v$ such that $v^\dagger\mathcal{A} = \lambda v^\dagger$. We define the spectral gap $\gamma(\mathcal{A}) = 1 - \max_{\lambda\in\mathsf{Spec}(\mathcal{A}),\lambda\neq 1} |\lambda|$. Note that $\gamma(\mathcal{A}) \le \min_{\lambda\in\mathsf{Spec}(\mathcal{A}),\lambda\neq 1} |1-\lambda|$.

We denote by $\mathbf{1}$ the column vector of all ones and similarly $\mathbf{0}$ the column vector of all zeros.

## 2.2 The Perron-Frobenius theorem

The *underlying graph* of a matrix $\mathcal{A}$ has an edge $(i,j)$ iff $\mathcal{A}[i,j] \neq 0$. A matrix $\mathcal{A}$ is *irreducible* if its underlying directed graph is strongly connected. When $\mathcal{A}$ is irreducible, its *period* is the greatest common divisor of the lengths of the closed directed paths in the underlying directed graph of $\mathcal{A}$. We say that $\mathcal{A}$ is *aperiodic* if its period is 1. A matrix $\mathcal{A}$ is *non-negative* if all its entries are non-negative, and it is *stochastic* if it is non-negative and every row sums to 1. We will need the Perron-Frobenius theorem for irreducible non-negative matrices (see, e.g., [21, Chapter 8]).

▶ **Theorem 3.** *Let $\mathcal{A}$ be an irreducible non-negative $n \times n$ matrix with period $h$ and spectral radius $\rho(A) = r$. Then:*

1. *There exists an $r$-right-eigenvector $v_1$ and an $r$-left-eigenvector $u_1$ whose components are all positive.*
2. *$\mathcal{A}$ has exactly $h$ complex eigenvalues with absolute value $r$ and each one of them is a product of $r$ with a different $h$-th root of unity. Consequently, if $\mathcal{A}$ is aperiodic then $r$ is a simple eigenvalue, and all other eigenvalues have absolute value strictly smaller than $r$.*
3. *It holds that $\lim_{k\to\infty} \mathcal{A}^k/r^k = |v_1\rangle\langle u_1|$, where $v_1$ and $u_1$ are normalized so that $\langle v_1|\, u_1\rangle = 1$.*

*If $\mathcal{A}$ is stochastic then $r = 1$. Furthermore, if $\mathcal{A}$ is stochastic, irreducible and aperiodic then $v_1$ is the all-ones vector $\mathbf{1}$ and $u_1 = \pi$ is the* stationary distribution *of the corresponding random walk (all up to normalizations).*

## 2.3 Jordan normal form

▶ **Fact 4.** *Every complex $n \times n$ matrix $\mathcal{A}$ can be expressed in a Jordan normal form $\mathcal{A} = V\mathbf{A}V^{-1}$ where $\mathbf{A} = \mathsf{diag}(\mathbf{A}_1, \ldots, \mathbf{A}_B)$,*

$$\mathbf{A}_b = \mathbf{A}_b(\lambda_b) = \begin{pmatrix} \lambda_b & 1 & & \\ & \lambda_b & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_b \end{pmatrix} \in \mathbb{C}^{\dim_b \times \dim_b},$$

*and $\dim_1 + \ldots + \dim_b = n$. The Jordan matrix $\mathbf{A}$ has the eigenvalues of $\mathcal{A}$ on its diagonal, and is unique up to the ordering of the blocks $\mathbf{A}_b$. For an eigenvalue $\lambda_b$, its algebraic multiplicity is the number of times it appears on the diagonal $\mathbf{A}$ and its geometric multiplicity is the number of blocks having $\lambda_b$ on their diagonal. We say an eigenvalue is* simple *if its algebraic multiplicity is one.*

▶ **Claim 5** ([9], Chapter 3)**.** *Let $\mathcal{A}$ be an $n \times n$ complex matrix and let $\mathcal{A} = V\mathbf{A}V^{-1}$ be the Jordan normal form of $\mathcal{A}$, where $\mathbf{A} = \mathsf{diag}(\mathbf{A}_1, \ldots, \mathbf{A}_B)$. Then, every Jordan block $\mathbf{A}_b$ corresponds to an $\mathcal{A}$-invariant subspace $E_b = \mathsf{Ker}\left((\lambda_b\mathcal{I} - \mathcal{A})^{\dim_b}\right)$ of dimension $\dim_b$. This gives a decomposition $\mathbb{C}^n = \bigoplus_{b=1}^{B} E_b$.*

For a Jordan decomposition $\mathcal{A} = V\mathbf{A}V^{-1}$, we will often write $\mathcal{A} = \sum_{b=1}^{B} V_b\mathbf{A}_bU_b$, where $\mathbf{A}_b$ is the $b$-th Jordan block, $V_b$ are the columns of $V$ that correspond to this block and similarly $U_b$ are the rows of $V^{-1}$ that correspond to this block.

When the operator is irreducible, aperiodic and stochastic, we can express the Perron-Frobenius theorem in the Jordan terminology and get:

▶ **Claim 6.** *Let $\mathcal{S}$ be an irreducible, aperiodic and stochastic matrix with a stationary distribution $\pi$ so that $\langle\mathbf{1}|\,\pi\rangle = 1$ and let $\mathcal{S} = \sum_{b=1}^{B} V_b\mathbf{S}_bU_b$ be a Jordan decomposition of $\mathcal{S}$. Then,*

- *$\mathbf{S}_1 = (1)$, the $1 \times 1$ matrix with an entry 1.*
- *For all $b \geq 2$, $U_bV_1 = U_b\,|\mathbf{1}\rangle = \mathbf{0}$ and $U_1V_b = \langle\pi|\,V_b = \mathbf{0}^\dagger$. Also, $\sum_{b=1}^{B} V_bU_b = \mathcal{I}$.*
- *$V_1\mathbf{S}_1U_1 = |\mathbf{1}\rangle\langle\pi|$ so $\mathcal{S} = |\mathbf{1}\rangle\langle\pi| + \sum_{b=2}^{B} V_b\mathbf{S}_bU_b$.*

**Proof.** If $v$ is a (right) eigenvector of $\mathcal{S}$ with eigenvalue $\lambda$ then $v \in \mathsf{Im}(\cup_{b:\lambda_b=\lambda}V_b)$. Similarly, if $w$ is a left eigenvector of $\mathcal{S}$, then its eigenvalue is an eigenvalue of $\mathcal{S}$ and $w \in \mathsf{Im}(\cup_{b:\lambda_b=\lambda}U_b)$ (this is because $\mathcal{A}$ and $\mathcal{A}^\dagger$ have the same spectrum, see, e.g., [8, Chapter 9]).

Now, since $\mathcal{S}$ is stochastic, $\mathbf{1}$ is a 1-eigenvector. Also, there is a 1-left-eigenvector that we denote by $\pi$, and we normalize $\pi$ such that $\langle\pi|\,\mathbf{1}\rangle = 1$. Furthermore, by the Perron-Frobenius theorem, the 1-eigenvalue is simple, so $\mathbf{S}_1 = (1)$, $U_1$ is a $1 \times n$ matrix and $V_1$ is a $n \times 1$ matrix. Furthermore, by the above, $\pi \in \mathsf{Im}(U_1)$, and since the dimension of the image is 1, we must have $\mathsf{Im}(U_1) = \mathrm{Span}(\{\pi\})$. Similarly, $\mathsf{Im}(V_1) = \mathrm{Span}(\{\mathbf{1}\})$. This completes the proof of the first item.

For the second item, let $U = V^{-1}$ and observe that since $UV = \mathcal{I}$, $\langle u_i|\,v_j\rangle = \delta_{i,j}$ (where $u_i$ is the $i$-th row of $U$ and $v_j$ is the $j$-th column of $V$). Now, consider $b \neq b'$ and the product $P = U_bV_{b'}$. Every entry of $P$ is of the form $\langle u_{b,i}|\,v_{b',j}\rangle$ where $i \in [\dim_b]$ and $j \in [\dim_{b'}]$. By the previous observation, they are all zeros. Also, $\mathcal{I}$ has a Jordan decomposition $V\mathbf{I}U$, so immediately it is clear that $\sum_{b=1}^{B} V_bU_b = \mathcal{I}$.

For the third item, Suppose $V_1 = \alpha \mathbf{1}$ and $U_1 = \beta \langle \pi |$ for some nonzero $\alpha, \beta \in \mathbb{C}$. We see that $V_1 \mathbf{S}_1 U_1 = \alpha\beta \, |\mathbf{1}\rangle\langle\pi|$. We want to determine $\alpha\beta$. Since $\langle \pi | \, \mathcal{S} = \langle \pi |$ we have that

$$\langle \pi | = \langle \pi | \, \mathcal{S} = \beta^{-1} U_1 \mathcal{S} = \beta^{-1} U_1 \sum_{b=1}^{B} V_b \mathbf{S}_b U_b$$

$$= \beta^{-1} U_1 V_1 \mathbf{I}_1 U_1 + \beta^{-1} \sum_{b=2}^{B} U_1 V_b \mathbf{S}_b U_b = \beta^{-1}\beta\alpha\beta \, \langle \pi | \, \mathbf{1} \rangle \, \langle \pi | \; = \; \alpha\beta \, \langle \pi | \, ,$$

so $\alpha\beta = 1$. Hence, $V_1 \mathbf{S}_1 U_1 = V_1 U_1 = |\mathbf{1}\rangle\langle\pi|$. ◀

## 2.4 Functions of matrices

This subsection follows the book of Higham [23]. In the Jordan basis, each Jordan block is a matrix with some complex value $\lambda$ over the main diagonal and 1 in the diagonal above it. We want to distinguish upper triangular matrices in which elements on the same diagonal have the same value. We note that this class $D$ of matrices is closed under matrix addition and multiplication. We denote:

▶ **Definition 7.** For $0 \leq i \leq n-1$ let $\mathcal{D}_{n,i}$ be the $n \times n$ matrix that has 1 over the $i$-th diagonal and 0 elsewhere, where the 0-th diagonal is the main diagonal and the $i$-th diagonal is the diagonal $i$ elements above it.

Clearly $D = \text{Span}\{\mathcal{D}_{n,0}, \ldots, \mathcal{D}_{n,n-1}\}$ is closed under matrix addition. Also, since

$$\mathcal{D}_{n,i} \cdot \mathcal{D}_{n,j} \; = \; \mathcal{D}_{n,i+j},$$

$D$ is also closed under matrix multiplication.

Suppose $p \in \mathbb{C}[x]$ is a polynomial $p(x) = \sum_{i=0}^{d} c_i x^i$. We can evaluate the polynomial over the ring $M_n(\mathbb{C})$, i.e., given an $n \times n$ matrix $\mathcal{A}$ we let

$$p(\mathcal{A}) = \sum_{i=0}^{d} c_i \mathcal{A}^i.$$

Note that if $\mathcal{A} = V\mathbf{A}V^{-1}$ then $p(\mathcal{A}) = Vp(\mathbf{A})V^{-1}$. Also, if $\mathbf{A} = \mathsf{diag}(\mathbf{A}_1, \ldots, \mathbf{A}_B)$ then $p(\mathbf{A}) = \mathsf{diag}(p(\mathbf{A}_1), \ldots, p(\mathbf{A}_B))$. In the extreme case where $\mathcal{A}$ is diagonalizable and all Jordan blocks have dimension 1, we see that $p$ acts on the eigenvalues of $\mathcal{A}$. In the general case, we need to understand how $p$ acts on a Jordan block $\mathbf{A}_b = \lambda_b \mathbf{I} + \mathcal{D}_{\dim_b,1}$. The answer is quite surprising and holds for arbitrary differentiable functions.

▶ **Lemma 8** ([23], Chapter 1). *Let $f : \mathbb{C} \to \mathbb{C}$ and suppose it is differentiable $n$ times on $\mathsf{Spec}(\mathcal{A})$. Let $\mathbf{A} \in \mathbb{C}^{n \times n}$ be a Jordan block $\mathbf{A} = \lambda\mathbf{I} + \mathcal{D}_{n,1}$. Then,*

$$f(\mathbf{A}) \; = \; \begin{pmatrix} f(\lambda) & f'(\lambda) & \ldots & \frac{f^{(n-1)}(\lambda)}{(n-1)!} \\ & f(\lambda) & \ddots & \vdots \\ & & \ddots & f'(\lambda) \\ & & & f(\lambda) \end{pmatrix} \; = \; \sum_{t=0}^{n-1} \frac{f^{(t)}(\lambda)}{t!} \mathcal{D}_{n,t}.$$

## 2.5 The generalized inverse

Let $\mathcal{A}$ be any complex linear operator. A generalized (reflexive) inverse $\mathcal{A}^+$ of $\mathcal{A}$ is a matrix that satisfies both $\mathcal{A}\mathcal{A}^+\mathcal{A} = \mathcal{A}$ and $\mathcal{A}^+\mathcal{A}\mathcal{A}^+ = \mathcal{A}^+$. A generalized inverse is not unique, however if we further demand that both $\mathcal{A}\mathcal{A}^+$ and $\mathcal{A}^+\mathcal{A}$ are Hermitian, then such an operator is unique, and is called the Moore-Penrose pseudo-inverse and can be computed using the singular values decomposition (SVD). If $\mathcal{A} = U\mathbf{\Sigma}V^\dagger$ is the SVD of $\mathcal{A}$ then the pseudo-inverse is $\mathcal{A}^+ = V\mathbf{\Sigma}^+U^\dagger$ where $\mathbf{\Sigma}^+ = \mathsf{inv}(\mathbf{\Sigma})$ and $\mathsf{inv}(x)$ is the univariate function that is $1/x$ when $x \neq 0$ and 0 otherwise.

We will not work with the SVD but rather with the Jordan canonical form. Let $\mathcal{A} = V\mathbf{A}V^{-1}$ be a Jordan decomposition of a singular matrix $\mathcal{A}$. When the algebraic multiplicity of the eigenvalue 0 is one, the matrix $\mathcal{A}^\star = \mathsf{inv}(A)$, according to Subsection 2.4, is well defined. Namely, $\mathsf{inv}(A) = V\mathbf{A}^{\mathsf{inv}}V^{-1}$ where $\mathbf{A}^{\mathsf{inv}}$ is obtained by inverting every Jordan block that does not correspond to the zero eigenvalue. It is immediate that $\mathcal{A}^\star$ is a generalized inverse, although it does not generally coincide with the pseudo-inverse. From here onward, we denote $\mathcal{A}^\star$ as the generalized inverse $\mathsf{inv}(A)$.

Any generalized inverse $\mathcal{A}^\star$ can be used to determine if a system of linear equations has any solution (and if so, to give them all). More concretely, if the system $\mathcal{A}x = b$ has a solution then all its solution are given by $x = \mathcal{A}^\star b + (I - \mathcal{A}^\star\mathcal{A})w$ for an arbitrary $w$. All of the above claims can be found, e.g., in [4].

It will later be evident that when $\mathcal{A} = \mathcal{L} = \mathcal{I} - \mathcal{S}$ is a Laplacian corresponding to an irreducible, aperiodic and stochastic matrix $\mathcal{S}$ with a stationary distribution $\pi$, the expression $\mathcal{I} - \mathcal{A}^\star\mathcal{A}$ is simply $|\mathbf{1}\rangle\langle\pi|$. Thus, if we find $\mathcal{L}^\star$ we can solve any set of equations $\mathcal{L}x = b$ that has a solution. In fact, this also works when we try to solve the system $\mathcal{L}x = b$ for $b$ that does not admit any perfect solution, but is close to a vector in $\mathsf{Im}(\mathcal{L})$. To see that, say $b$ is arbitrary, and on input $b$ and $\mathcal{L}$ we output $z = \mathcal{L}^\star b$. Then $\|\mathcal{L}z - b\| = \|(\mathcal{L}\mathcal{L}^\star - \mathcal{I})b\| = \||\mathbf{1}\rangle\langle\pi|\, b\| = \sqrt{n}\cdot|\langle\pi, b\rangle|$, and so if $b$ is $\delta$ close to being perpendicular to $\pi$ (and so close to being in $\mathsf{Im}(\mathcal{L})$) then the solution $z = \mathcal{L}^\star b$ is such that $\mathcal{L}z$ is $\sqrt{n}\delta$ close to the desired value $b$.

## 2.6 Space-bounded probabilistic computation

### 2.6.1 The model of computation

A space-bounded probabilistic Turing machine has four semi-infinite tapes: a read-only *input tape*, a *work tape*, a read-only uni-directional *random-coins tape* and a write-only uni-directional *output tape*. We say a language is accepted by a probabilistic TM if for every input in the language the acceptance probability is at least 2/3 and for every input not in the language it is at most 1/3. As usual, the acceptance probability can be amplified as long as there is some non-negligible gap between the acceptance probability of yes and no instances.

The complexity class $\mathsf{BPL}$ comprises all languages accepted by a space-bounded probabilistic TM with space complexity $O(\log n)$ and polynomial time.

### 2.6.2 Simulatable matrices

We are often interested in approximating a *value* (e.g., a matrix entry) with probabilistic machines. Assume that for an input $x \in \{0,1\}^n$ there exists a value $u = u(x) \in \mathbb{C}$. We say a probabilistic TM $(\varepsilon, \delta)$-*approximates* $u(x)$ if

$$\forall_{x\in\{0,1\}^n} \Pr_y[|M(x,y) - u(x)| \geq \varepsilon] \;\leq\; \delta.$$

If $u$ is multi-valued (say, a vector) we say a TM $(\varepsilon,\delta)$-approximates $u$ if given an index $i$ it $(\varepsilon,\delta)$-approximates $u[i]$.

▶ **Definition 9.** We say that a family of matrices $\mathfrak{A}$ is *simulatable* if there exists a probabilistic algorithm that on input $\mathcal{A}\in\mathfrak{A}$ of dimension $n$, $k\in\mathbb{N}$, $s,t\in[n]$, $\varepsilon,\delta>0$ runs in space $O(\log\frac{nk}{\varepsilon}+\log\log\frac{1}{\delta})$ and $(\varepsilon,\delta)$-approximates $A^k[s,t]$.

Probabilistic logspace machines can approximate random walks well. In [17], it is shown that:

▶ **Lemma 10.** *The family of stochastic matrices is simulatable.*

We can also conclude:

▶ **Lemma 11** ([17])**.** *Let $A\in\mathbb{C}^{n\times n}$ be a stochastic matrix and let $p=\sum_{i=0}^{d}c_ix^i$ be a complex polynomial such that:*

- *For every $i$, $|c_i|\le M$, and,*
- *The coefficients $c_i$ are explicit in the sense that there exists an algorithm that given $k\le d,\varepsilon,\delta$ outputs an $(\varepsilon,\delta)$-approximation of $c_k$ using $O(\log\frac{nMd\log\frac{1}{\delta}}{\varepsilon})$ space.*

*Then, the entries of $p(A)$ can be $(\varepsilon,\delta)$-approximated using $O(\log\frac{nMd\log\frac{1}{\delta}}{\varepsilon})$ space.*

## 3 Approximating $(\mathcal{I}-\mathcal{A})^{-1}$ by the Taylor series

We start with the simple case of normal matrices, and consider general functions.

▶ **Theorem 12.** *Let $f,p:\mathbb{C}\to\mathbb{C}$ and $\varepsilon>0$. Suppose $\mathcal{A}$ is a normal matrix such that for every $\lambda\in\mathsf{Spec}(A)$, $|f(\lambda)-p(\lambda)|\le\varepsilon$. Then, $\|f(\mathcal{A})-p(\mathcal{A})\|\le\varepsilon$.*

**Proof.** $\mathcal{A}$ is normal, so it is diagonalizable by a unitary matrix, $\mathcal{A}=UDU^\dagger$. Also, $f(\mathcal{A})=Uf(D)U^\dagger$ and $p(\mathcal{A})=Up(D)U^\dagger$. Thus, we have that

$$\|f(\mathcal{A})-p(\mathcal{A})\| \;\le\; \|U\|\,\|U^\dagger\|\,\|f(D)-p(D)\| \;=\; \|f(D)-p(D)\|\,,$$

and $\|f(D)-p(D)\|$ is simply $\max_{\lambda\in\mathsf{Spec}(\mathcal{A})}|f(\lambda)-p(\lambda)|\le\varepsilon$. ◀

With that we can easily see that when $\mathcal{A}$ is normal, $\sum_{i=0}^{T}\mathcal{A}^i$ approximates $(\mathcal{I}-\mathcal{A})^{-1}$ pretty well. Formally,

▶ **Corollary 13.** *Let $\mathcal{A}$ be a normal matrix and suppose $\mathsf{Spec}(\mathcal{A})\subseteq[0,1)$ and in particular $\mathcal{I}-\mathcal{A}$ is invertible. Then,*

$$\left\|(\mathcal{I}-\mathcal{A})^{-1}-\sum_{i=0}^{T}\mathcal{A}^i\right\| \;\le\; \frac{e^{-T\overline{\lambda}(\mathcal{A})}}{\overline{\lambda}(\mathcal{A})}.$$

**Proof.** For $\lambda\in[0,1)$, it holds that

$$\left|\frac{1}{1-\lambda}-\sum_{i=0}^{T}\lambda^i\right| \;\le\; \sum_{T+1}^{\infty}\lambda^i \;=\; \frac{\lambda^{T+1}}{1-\lambda}.$$

The above expression is maximized where $\lambda=1-\gamma(\mathcal{A})$, so we have:

$$\left|\frac{1}{1-\lambda}-\sum_{i=0}^{T}\lambda^i\right| \;\le\; \frac{(1-\gamma(\mathcal{A}))^T}{\gamma(\mathcal{A})} \;\le\; \frac{e^{-T\gamma(\mathcal{A})}}{\gamma(\mathcal{A})},$$

and the corollary follows. ◀

We would like to extend this result to arbitrary operators $\mathcal{A}$. As a first attempt we begin with generalizing Theorem 12 to arbitrary operators. For that we need the representation of $\mathcal{A}$ in its Jordan normal form, and we also need the function $p$ and its derivatives to approximate the target function $f$ and its derivatives well. We prove:

▶ **Theorem 14.** *Let $f, p : \mathbb{C} \to \mathbb{C}$. Suppose $\mathcal{A}$ is an $n \times n$ matrix such that for every $\lambda \in \mathsf{Spec}(A)$ and every $k \leq n$, $|f^{(k)}(\lambda) - p^{(k)}(\lambda)| \leq k! \cdot \varepsilon_k$. Furthermore, assume $\mathcal{A}$ has a Jordan decomposition $\mathcal{A} = V\mathbf{A}V^{-1}$, and the largest Jordan block has dimension $D$. Then, $\|f(\mathcal{A}) - p(\mathcal{A})\| \leq \kappa(V) \cdot \sum_{k=0}^{D-1} \varepsilon_k$.*

**Proof.** Let $\mathbf{A} = \mathbf{A}_1 \oplus \ldots \oplus \mathbf{A}_b$, corresponding to the different Jordan blocks. By Lemma 8, $f(\mathcal{A}) = Vf(\mathbf{A})V^{-1}$ where $f(\mathbf{A}) = f(\mathbf{A}_1) \oplus \ldots \oplus f(\mathbf{A}_b)$,

$$f(\mathbf{A}_i) \;=\; \begin{pmatrix} f(\lambda_i) & f'(\lambda_i) & \ldots & \frac{f^{(\dim_i - 1)}(\lambda_i)}{(\dim_i - 1)!} \\ & f(\lambda_i) & \ddots & \vdots \\ & & \ddots & f'(\lambda_i) \\ & & & f(\lambda_i) \end{pmatrix} \;=\; \sum_{k=0}^{\dim_i - 1} \frac{f^{(k)}(\lambda_i)}{k!} \mathcal{D}_{\dim_i, k},$$

and $\lambda_i$ is the eigenvalue corresponding to the block $\mathbf{A}_i$ of dimension $\dim_i$. The same of course holds for $p$. Thus,

$$\|f(\mathcal{A}) - p(\mathcal{A})\| \;=\; \left\|V(f(\mathbf{A}) - p(\mathbf{A}))V^{-1}\right\| \;\leq\; \kappa(V) \cdot \|f(\mathbf{A}) - p(\mathbf{A})\|.$$

To bound the latter expression, note that

$$\begin{aligned} \|f(\mathbf{A}) - p(\mathbf{A})\| &= \max_{i \in [b]} \|f(\mathbf{A}_i) - p(\mathbf{A}_i)\| \\ &\leq \max_{i \in [b]} \sum_{k=0}^{\dim_i - 1} \left| \frac{f^{(k)}(\lambda_i) - p^{(k)}(\lambda_i)}{k!} \right| \|\mathcal{D}_{\dim_i, k}\| \;\leq\; \sum_{k=0}^{D-1} \varepsilon_k. \end{aligned}$$ ◀

When $\mathcal{A}$ is normal, $\kappa(V) = 1$ and the maximal block length is 1, so we recover Theorem 12. We now check what we get for $(\mathcal{I} - \mathcal{A})^{-1}$ and an arbitrary operator $\mathcal{A}$:

▶ **Corollary 15.** *Suppose $\mathcal{A}$ is an $n \times n$ matrix that has a Jordan decomposition $\mathcal{A} = V\mathbf{A}V^{-1}$. Suppose every eigenvalue $\lambda$ of $\mathcal{A}$ satisfies $|\lambda| < 1$ and in particular $\mathcal{I} - \mathcal{A}$ is invertible. Let $T \in \mathbb{N}$ such that $T \geq \frac{8n^2}{\gamma(\mathcal{A})^2}$, let $f(\mathcal{A}) = (\mathcal{I} - \mathcal{A})^{-1}$ and $p(\mathcal{A}) = \sum_{i=0}^{T} \mathcal{A}^i$. Then,*

$$\|f(\mathcal{A}) - p(\mathcal{A})\| \;\leq\; 2n\kappa(V) \frac{e^{-T\gamma(\mathcal{A})/4}}{\gamma(\mathcal{A})}.$$

**Proof.** Let $\mathcal{A}$ be an $n \times n$ matrix and suppose every eigenvalue $\lambda$ of $\mathcal{A}$ satisfies $|\lambda| < 1$. We consider, again, inverting $\mathcal{I} - \mathcal{A}$ by considering the function $f(\lambda) = \frac{1}{1-\lambda}$ and its power-series expansion $p(\lambda) = \sum_{i=0}^{T} \lambda^i$. For $k \leq n$, one can verify that $\frac{1}{k!} f^{(k)}(\lambda) = \frac{1}{(1-\lambda)^{k+1}}$ and $\frac{1}{k!} p^{(k)}(\lambda) = \sum_{i=0}^{T-k} \binom{k+i}{k} \lambda^i$. Also, $\frac{1}{k!} f^{(k)}(\lambda) = \sum_{i=0}^{\infty} \binom{k+i}{k} \lambda^i$ so we see that

$$\varepsilon_k \;=\; \left| \sum_{i=T-k+1}^{\infty} \binom{k+i}{k} \lambda^i \right|.$$

As $T \geq 4n$, $T - k + 1 \geq T/2$. Also, $\binom{k+i}{k} \leq (k+i)^k \leq (2i)^k$, so $\varepsilon_k \;\leq\; \sum_{i=T/2}^{\infty} (2i)^k \lambda^i$. Now,

we have that $(2i)^k \leq |\lambda|^{-i/2}$, since

$$\begin{aligned}(2i)^k|\lambda|^{i/2} &= e^{k\ln(2i)-(i/2)\ln\frac{1}{|\lambda|}} = e^{\frac{1}{2}\left(2k\ln(2i)-i\ln\frac{1}{|\lambda|}\right)} \leq e^{\frac{1}{2}\left(n\sqrt{i}-i\ln\frac{1}{|\lambda|}\right)} \\ &\leq e^{\frac{\sqrt{i}}{2}\left(n-\sqrt{i}\ln\frac{1}{|\lambda|}\right)} \leq e^{\frac{\sqrt{i}}{2}\left(n-\sqrt{T/2}\cdot\ln\frac{1}{1-\gamma(\mathcal{A})}\right)} \leq e^{\frac{\sqrt{i}}{2}\left(n-\sqrt{T/2}\cdot\gamma(\mathcal{A})\right)} \\ &\leq e^{\frac{\sqrt{i}}{2}(n-2n)} \leq 1.\end{aligned}$$

Plugging it to the above bound for $\varepsilon_k$, we obtain:

$$\varepsilon_k \leq \left|\sum_{i=T/2}^{\infty}\lambda^{i/2}\right| = \left|\frac{\lambda^{T/4}}{1-\sqrt{\lambda}}\right|.$$

To bound $\left|\frac{1}{1-\sqrt{\lambda}}\right|$, we use the fact that:

$$\left|\frac{1}{1-\sqrt{\lambda}}\right| = \frac{|1+\sqrt{\lambda}|}{|1-\lambda|} \leq \frac{2}{\gamma(\mathcal{A})}.$$

Altogether,

$$\varepsilon_k \leq \frac{2}{\gamma(\mathcal{A})}(1-\gamma(\mathcal{A}))^{T/4} \leq \frac{2e^{-T\gamma(\mathcal{A})/4}}{\gamma(\mathcal{A})}.$$

The Corollary follows by applying Theorem 14 and using the fact that $D \leq n$. ◀

## 4 Computing the generalized inverse of the Laplacian

In this section we approximate the generalized inverse of the Laplacian of directed graphs as long as we have a good approximation of its stationary distribution. Formally,

▶ **Theorem 16.** *There exists a probabilistic algorithm that gets as input:*

- *An $n \times n$ irreducible, aperiodic stochastic matrix $\mathcal{S}$,*
- *Two parameters, $\kappa$ and $\gamma$, which describe how stable the input $\mathcal{S}$ is:*
  - *Suppose $\kappa \geq \kappa(V)$, where $\mathcal{S} = V\mathbf{S}V^{-1}$ is any Jordan decomposition of $\mathcal{S}$, and,*
  - *$\gamma(\mathcal{S}) \geq \gamma$.*
- *Desired accuracy and confidence parameters $\varepsilon, \delta > 0$.*
- *An approximation $\tilde{\pi}$ of the stationary distribution $\pi$ of $\mathcal{S}$, where $\|\tilde{\pi} - \pi\| \leq \tau$ and $\tau \leq \frac{\varepsilon}{(T+1)\sqrt{n}}$ for $T = \frac{8n^2}{\gamma^2}\left(1 + \log\frac{n\kappa}{\varepsilon\gamma}\right)$.*

*Let $\mathcal{L}$ denote the Laplacian, $\mathcal{L} = \mathcal{I} - \mathcal{S}$. Then, the algorithm outputs a $(3\varepsilon, \delta)$-approximation of $\mathcal{L}^\star$ using*

$$O\left(\log\frac{n}{\gamma\varepsilon} + \log\log\frac{\kappa}{\delta}\right)$$

*space.*

Intuitively, we would like to employ the following approach. Given a stochastic operator $\mathcal{S}$ with a unique stationary distribution $\pi$, we would like to "peel off" the $1 \times 1$ Jordan block with eigenvalue 1, so that we are left with an operator $\mathcal{A}$ such that $\mathcal{I} - \mathcal{A}$ is invertible. Then, we would like to use Corollary 15 to approximate $(\mathcal{I} - \mathcal{A})^{-1}$ by $\sum_{i=0}^{T}\mathcal{A}^i$, using the fact that we can approximate $\mathcal{A}^i$ well with a BPL algorithm.

There are two obstacles that we need to overcome:

- First, when $\mathcal{S}$ in not normal, we do not have an orthonormal basis, so we need to explain what "peeling off" the stationary distribution means. It turns out that $\mathcal{A} = \mathcal{S} - |\mathbf{1}\rangle\langle\pi|$.
- Second, while $\mathcal{S}$ is stochastic, $\mathcal{A} = \mathcal{S} - |\mathbf{1}\rangle\langle\pi|$ is not, and furthermore, its $\ell_\infty$ norm is usually greater than 1. In particular, we cannot immediately assume that we can approximate high powers of it in BPL. We will show that $\mathcal{A}$ is still simulatable because $|\mathbf{1}\rangle\langle\pi|$ commutes with both $\mathcal{S}$ and $\mathcal{A}$.

We also need to check that the fact that $\tilde{\pi}$ is only close to $\pi$ and not exactly it, does not affect the parameters by too much.

We start the formal exposition with a precise description of the algorithm.

## 4.1 The Algorithm

The algorithm first computes the parameter

$$T \;=\; \left\lceil \frac{8n^2}{\gamma^2}\left(1+\log\frac{n\kappa}{\varepsilon\gamma}\right)\right\rceil.$$

The algorithm then computes an $(\varepsilon,\delta)$-approximation of the matrix

$$\widetilde{Q}_T(\mathcal{S}) \;=\; \left(\sum_{i=0}^{T}\mathcal{S}^i\right) - (T+1)\,|\mathbf{1}\rangle\langle\tilde{\pi}|$$

using Lemma 11 (note that since $\tilde{\pi}$ is given, we *approximate* the power series and compute $(T+1)\,|\mathbf{1}\rangle\langle\tilde{\pi}|$ exactly).

We first argue that the algorithm runs in small space and then analyze correctness.

## 4.2 Efficiency

We observe:

▶ **Lemma 17.** *For every $\varepsilon,\delta>0$ and integer $T$, and any $n\times n$ stochastic matrix $\mathcal{S}$, the entries of $\widetilde{Q}_T(\mathcal{S})$ can be $(\varepsilon,\delta)$-approximated using $O\left(\log\frac{nT\log\frac{1}{\delta}}{\varepsilon}\right)$ space.*

**Proof.** The claim follows directly from Lemma 11 since $\mathcal{S}$ is stochastic. ◀

## 4.3 Correctness

We first do the analysis in the ideal situation that $\tilde{\pi}=\pi$ and see that in this case the algorithm $(2\varepsilon,\delta)$-approximates $\mathcal{L}^\star$. We then show that when $\|\pi-\tilde{\pi}\|\le\tau$ the algorithm $(3\varepsilon,\delta)$-approximates $\mathcal{L}^\star$.

### 4.3.1 Peeling off the 1-eigenspace

Throughout the proof we use the representation of $\mathcal{S}$ guaranteed by Claim 6. Namely, $\mathcal{S}$ can be written as $\mathcal{S}=\sum_{b=1}^{B}V_b\mathbf{S}_bU_b$ where

- $\mathbf{S}_1$ is a $1\times 1$ matrix and $\mathbf{S}_1=(1)$. Also, $V_1U_1=|\mathbf{1}\rangle\langle\pi|$ and $\langle\mathbf{1}|\,\pi\rangle=1$,
- For all $b\ge 2$, $U_b\,|\mathbf{1}\rangle=\mathbf{0}$ and $\langle\pi|\,V_b=\mathbf{0}^\dagger$, and
- $\sum_{b=1}^{B}V_bU_b=\mathcal{I}$.

Our goal is to find the generalized inverse of $\mathcal{L} = \mathcal{I} - \mathcal{S}$. As explained before, our first step is to "peel-off" from $\mathcal{S}$ the 1-eigenspace, and the correct way to do that is by annihilating the $1 \times 1$ Jordan block with eigenvalue 1. We therefore define:

$$\mathcal{A} \;=\; \mathcal{S} - |\mathbf{1}\rangle\langle\pi| \,.$$

We notice that $\mathcal{S}$, $\mathcal{A}$, $\mathcal{L}$ and $\mathcal{L}^\star$ share the same Jordan basis, therefore, if we express $\mathcal{S} = \sum_{b=1}^{B} U_b \mathbf{S}_b V_b$ then

$$\mathcal{L} \;=\; \sum_{b=2}^{B} V_b(\mathbf{I}_b - \mathbf{S}_b)U_b,$$

and,

$$\mathcal{A} \;=\; \sum_{b=2}^{B} V_b \mathbf{S}_b U_b.$$

We denote $\mathbf{L}_b = \mathbf{I}_b - \mathbf{S}_b$ for $b \geq 2$ (and $\mathbf{L}_1$ is the zero matrix). The big advantage of $\mathcal{A}$ over $\mathcal{S}$ is that in $\mathcal{A}$ all eigenvalues have magnitude smaller than 1, as $\mathcal{A} = \sum_{b=2}^{B} V_b \mathbf{S}_b U_b$, and therefore $\mathcal{I} - \mathcal{A}$ is invertible. We still need, however, to relate $\mathcal{L}^\star$ to $(\mathcal{I} - \mathcal{A})^{-1}$. We prove:

▶ **Lemma 18.** $\mathcal{L}^\star = (\mathcal{I} - \mathcal{A})^{-1} - |\mathbf{1}\rangle\langle\pi|$.

**Proof.** Recall that $\mathcal{S} = |\mathbf{1}\rangle\langle\pi| + \sum_{b=2}^{B} V_b \mathbf{S}_b U_b$, $\mathcal{A} = \sum_{b=2}^{B} V_b \mathbf{S}_b U_b$ and $\mathcal{I} = \sum_{b=1}^{B} V_b U_b$. Hence,

$$\mathcal{I} - \mathcal{A} \;=\; \sum_{b=1}^{B} V_b U_b - \sum_{b=2}^{B} V_b \mathbf{S}_b U_b \;=\; V_1 U_1 + \sum_{b=2}^{B} V_b(\mathbf{I}_b - \mathbf{S}_b)U_b \;=\; |\mathbf{1}\rangle\langle\pi| + \sum_{b=2}^{B} V_b \mathbf{L}_b U_b.$$

The inverse is thus given by

$$(\mathcal{I} - \mathcal{A})^{-1} \;=\; |\mathbf{1}\rangle\langle\pi| + \sum_{b=2}^{B} V_b \mathbf{L}_b^{-1} U_b \;=\; |\mathbf{1}\rangle\langle\pi| + \mathcal{L}^\star,$$

as desired. ◀

Intuitively, this means that approximating $(\mathcal{I} - \mathcal{A})^{-1}$ suffices for approximating $\mathcal{L}^\star$, and we next consider approximating $(\mathcal{I} - \mathcal{A})^{-1}$.

### 4.3.2 Approximating $(\mathcal{I} - \mathcal{A})^{-1}$

Since all eigenvalues of $\mathcal{A}$ have magnitude smaller than 1, we can apply Corollary 15 and get:

▶ **Lemma 19.**

$$\left\| (\mathcal{I} - \mathcal{A})^{-1} - \sum_{k=0}^{T} \mathcal{A}^k \right\| \;\leq\; \varepsilon.$$

**Proof.** We saw that $\mathcal{A} = \sum_{b=2}^{B} V_b \mathbf{S}_b U_b$, and by the Perron-Frobenius theorem the eigenvalues that are written on $\mathbf{S}_b$ for $b \geq 2$, are at most $1 - \gamma < 1$ in absolute value. Thus, all eigenvalues of $\mathcal{A}$ have absolute value at most $\gamma(\mathcal{S})$. By Corollary 15, for $T \geq \frac{8n^2}{\gamma(\mathcal{S})^2}$,

$$\left\| (\mathcal{I} - \mathcal{A})^{-1} - \sum_{k=0}^{T} \mathcal{A}^k \right\| \;\leq\; 2n\kappa(V)\frac{e^{-T\gamma(\mathcal{S})/4}}{\gamma(\mathcal{S})}.$$

Substituting $T = \left\lceil \frac{8n^2}{\gamma(\mathcal{S})^2} \ln \frac{2n\kappa(V)}{\varepsilon\gamma(\mathcal{S})} \right\rceil$, the desired bound holds. ◀

Thus, the problem now reduces to simulating $\mathcal{A}^i$ in small space. As mentioned before, $\mathcal{A}$ is not stochastic and its $\ell_\infty$ norm is often larger than 1. However $\mathcal{A} = \mathcal{S} - |\mathbf{1}\rangle\langle\pi|$ has a very special form that conforms with the Jordan basis structure, which we now employ:

▶ **Claim 20.** *The matrices $\mathcal{S}$ and $|\mathbf{1}\rangle\langle\pi|$ commute, and furthermore $\mathcal{S}\cdot|\mathbf{1}\rangle\langle\pi| = |\mathbf{1}\rangle\langle\pi|\cdot\mathcal{S} = |\mathbf{1}\rangle\langle\pi|$.*

**Proof.**

$$\mathcal{S}\cdot|\mathbf{1}\rangle\langle\pi| \;=\; |\mathbf{1}\rangle\langle\pi| + \sum_{b=2}^{B} V_b\mathbf{S}_b U_b\cdot|\mathbf{1}\rangle\langle\pi| \;=\; |\mathbf{1}\rangle\langle\pi|\,,$$

and,

$$|\mathbf{1}\rangle\langle\pi|\cdot\mathcal{S} \;=\; |\mathbf{1}\rangle\langle\pi| + \sum_{b=2}^{B} |\mathbf{1}\rangle\langle\pi|\cdot V_b\mathbf{S}_b U_b \;=\; |\mathbf{1}\rangle\langle\pi|\,. \qquad \blacktriangleleft$$

▶ **Claim 21.** *For every $k \geq 1$, $\mathcal{A}^k = \mathcal{S}^k - |\mathbf{1}\rangle\langle\pi|$.*

**Proof.** The proof is by induction on $k$. For $k = 1$ the claim follows by the definition. Assume the statement holds for $k \in \mathbb{N}$, and consider $\mathcal{A}^{k+1}$, so By Claim 20:

$$\begin{aligned}\mathcal{A}^{k+1} &= (\mathcal{S} - |\mathbf{1}\rangle\langle\pi|)\cdot(\mathcal{S}^k - |\mathbf{1}\rangle\langle\pi|)\\ &= \mathcal{S}^{k+1} - \mathcal{S}\cdot|\mathbf{1}\rangle\langle\pi| - |\mathbf{1}\rangle\langle\pi|\cdot\mathcal{S}^k + |\mathbf{1}\rangle\langle\pi|\,\mathbf{1}\rangle\langle\pi|\\ &= \mathcal{S}^{k+1} - |\mathbf{1}\rangle\langle\pi| - |\mathbf{1}\rangle\langle\pi| + |\mathbf{1}\rangle\langle\pi| \;=\; \mathcal{S}^{k+1} - |\mathbf{1}\rangle\langle\pi|\,. \qquad \blacktriangleleft\end{aligned}$$

Thus, $\mathcal{A}$ is simulatable and we can approximate $(\mathcal{I} - \mathcal{A})^{-1}$ in small space.

### 4.3.3 Putting everything together

Define the *ideal* polynomial $Q_T$ by:

$$Q_T(\mathcal{S}) \;=\; \left(\sum_{i=0}^{T}\mathcal{S}^i\right) - (T+1)\,|\mathbf{1}\rangle\langle\pi|\,.$$

▶ **Lemma 22.** $\|\mathcal{L}^\star - Q_T(\mathcal{S})\| \leq \varepsilon$.

**Proof.**

$$\begin{aligned}\|\mathcal{L}^\star - Q_T(\mathcal{S})\| &= \left\|(\mathcal{I}-\mathcal{A})^{-1} - |\mathbf{1}\rangle\langle\pi| - Q_T(\mathcal{S})\right\|\\ &\leq \left\|\left(\sum_{i=0}^{T}\mathcal{A}^i\right) - |\mathbf{1}\rangle\langle\pi| - Q_T(\mathcal{S})\right\| + \varepsilon\\ &= \left\|\mathcal{A}^0 + \sum_{i=1}^{T}\left(\mathcal{S}^i - |\mathbf{1}\rangle\langle\pi|\right) - |\mathbf{1}\rangle\langle\pi| - Q_T(\mathcal{S})\right\| + \varepsilon\\ &= \left\|\left(\sum_{i=0}^{T}\mathcal{S}^i\right) - (T+1)\,|\mathbf{1}\rangle\langle\pi| - Q_T(\mathcal{S})\right\| + \varepsilon \;=\; \varepsilon. \qquad \blacktriangleleft\end{aligned}$$

Finally, we check how the fact that $\tilde{\pi}$ is only close to $\pi$, affects our accuracy. We see that:

▶ **Claim 23.** $\left\|\widetilde{Q}_T(\mathcal{S}) - Q_T(\mathcal{S})\right\| \leq \varepsilon$.

**Proof.** Notice that $\widetilde{Q}_T(\mathcal{S}) - Q_T(\mathcal{S}) = (T+1)\,|\mathbf{1}\rangle\langle\tilde{\pi}-\pi|$. Therefore, $\left\|\widetilde{Q}_T(\mathcal{S}) - Q_T(\mathcal{S})\right\| \leq (T+1)\cdot\|\mathbf{1}\|\cdot\|\tilde{\pi}-\pi\|$. The proof follows because $\|\mathbf{1}\| = \sqrt{n}$ and $\|\tilde{\pi}-\pi\| \leq \tau \leq \frac{\varepsilon}{\sqrt{n}(T+1)}$. ◀

Now, since we $(\varepsilon,\delta)$-approximate $\widetilde{Q}_T(\mathcal{S})$, then except for probability $\delta$ what we output is $\varepsilon$-close to $\widetilde{Q}_T(\mathcal{S})$, and therefore it is $2\varepsilon$-close to $Q_T(\mathcal{S})$ and $3\varepsilon$-close to $\mathcal{L}^\star$, which completes the proof of Theorem 16.

## 5 Some specific families of graphs

Ultimately, we would like to solve in BPL any set of equations $\mathcal{L}x = b$, where $b$ is close to $\mathsf{Im}(\mathcal{L})$, and where $\mathcal{L}$ is the Laplacian of a stochastic matrix $\mathcal{S}$. Theorem 16 is a step towards this goal, but it works only when:

- $\mathcal{S}$ is irreducible, namely, its underlying graph is strongly connected,
- $\mathcal{S}$ is aperiodic,
- We can approximate well the unique stationary distribution $\pi$,
- $\gamma(\mathcal{S}) \geq \frac{1}{n^a}$ for some constant $a$, i.e., all eigenvalues except the largest one, are at most $1-\gamma$ in absolute value, and,
- $\kappa(V) \leq 2^{n^b}$ for some constant $b$, where $\mathcal{S} = V\mathbf{S}V^{-1}$ is a Jordan decomposition and $\kappa(V) = \|V\|\cdot\left\|V^{-1}\right\|$. Notice that here we may tolerate exponential $\kappa(V)$ as the space complexity dependency on $\kappa$ is doubly-logarithmic.

In this section we want to examine which requirements can be relaxed. The section is organized as follows. First, we note that we can get rid of the aperiodicity requirement and we can somewhat relax the spectral gap requirement. Then we show that in some cases we can get rid of the $\kappa(V)$ requirement (when the eigenvalues are polynomially separated). Finally, we give specific results for:

- Undirected graphs,
- Directed Eulerian graphs (which generalize directed regular graphs), and,
- Directed rapidly-mixing graphs.

### 5.1 Omitting the aperiodicity requirement using lazy walks

Given a stochastic matrix $\mathcal{S}$ we can convert it to the corresponding lazy walk $\mathcal{S}' = \frac{1}{2}(\mathcal{I}+\mathcal{S})$, that stays in place with probability half. Define:

$$\gamma'(\mathcal{S}) \;=\; \max_{\lambda\in\mathsf{Spec}(\mathcal{S}'),\lambda\neq 1}(1-\Re(\lambda)).$$

The conversion has two benefits. First, the walk is clearly aperiodic. Also, we will be able to replace the condition $\gamma \leq \gamma(S)$, with the milder condition $\gamma \leq \gamma'(\mathcal{S})$. We will also show that we can recover the generalized inverse of the Laplacian of a graph $G$ from that of the lazy walk variant of $G$. We prove:

▶ **Theorem 24.** *There exists a probabilistic algorithm that gets as input:*

- *An $n\times n$ irreducible, stochastic matrix $\mathcal{S}$.*
- *Two parameters, $\kappa$ and $\gamma$, which describe how stable the input $\mathcal{S}$ is:*
  - *Suppose $\kappa \geq \kappa(V)$, where $\mathcal{S} = V\mathbf{S}V^{-1}$ is any Jordan decomposition of $\mathcal{S}$, and,*
  - *$\gamma'(\mathcal{S}) \geq \gamma$.*
- *Desired accuracy and confidence parameters $\varepsilon,\delta > 0$.*
- *An approximation $\tilde{\pi}$ of the stationary distribution $\pi$ of $\mathcal{S}$, where $\|\tilde{\pi}-\pi\| \leq \tau$ and $\tau \leq \frac{\varepsilon}{(T+1)\sqrt{n}}$ for $T = \frac{8n^2}{\gamma^2}\left(1+\log\frac{n\kappa}{\varepsilon\gamma}\right)$.*

*Let $\mathcal{L}$ denote the Laplacian, $\mathcal{L} = \mathcal{I} - \mathcal{S}$. Then, the algorithm outputs a $(3\varepsilon, \delta)$-approximation of $\mathcal{L}^\star$ using*

$$O\left(\log \frac{n}{\gamma\varepsilon} + \log\log \frac{\kappa}{\delta}\right)$$

*space.*

**Proof.** We run the algorithm of Theorem 16 over $\mathcal{S}' = \frac{1}{2}(\mathcal{I} + \mathcal{S})$. It is clear that $\mathcal{S}'$ is stochastic and aperiodic. By assumption, $\mathcal{S}'$ is irreducible (since $\mathcal{S}$ is). Also, $\mathcal{S}$ and $\mathcal{S}'$ have the same $V$ and by assumption $\kappa(V) \leq \kappa$. They also share the same stationary distribution $\pi$, and we are given $\pi'$ which is close to $\pi$.

We will soon prove that $\gamma(\mathcal{S}') \geq \frac{\gamma'(\mathcal{S})}{4}$. Therefore, by Theorem 16, we get a $(3\varepsilon, \delta)$-approximation of $(\mathcal{I} - \mathcal{S}')^\star$. Finally, we will see that $(\mathcal{I} - \mathcal{S}')^\star = 2(\mathcal{I} - \mathcal{S})^\star$ and so we easily get an approximation for $(\mathcal{I} - \mathcal{S})^\star$.

To see that indeed $(\mathcal{I} - \mathcal{S}')^\star = 2(\mathcal{I} - \mathcal{S})^\star$, notice that $\mathcal{I}$ and $\mathcal{S}$ share the same Jordan basis $V$. The first block in $\mathcal{S}'$ and $\mathcal{S}$ is the same, and for $b \geq 2$, if the $b$-th block in $\mathcal{S}$ is $\mathbf{S}_b$, then the $b$-th block in $(\mathcal{I} - \mathcal{S}')^\star$ is $(\mathcal{I} - \frac{1}{2}(I + \mathbf{S}_b))^{-1} = 2(I - \mathbf{S}_b)^{-1}$ and the $b$-th block of $(\mathcal{I} - \mathcal{S})^\star$ is $(\mathbf{I} - \mathbf{S}_b)^{-1}$.

Thus, all that is left is to prove:

▶ **Claim 25.** *It holds that $\gamma(\mathcal{S}') \geq \frac{\gamma'(\mathcal{S})}{4}$.*

**Proof.** Fix $\lambda \in \mathsf{Spec}(\mathcal{S})$, $|\lambda| \leq 1$, and write $\lambda = a + bi$ for $a, b \in \mathbb{R}$. Also, let $\lambda' = \frac{1}{2} + \frac{1}{2}\lambda = \frac{1+a}{2} + \frac{b}{2}i$, which is the corresponding eigenvalue in $\mathcal{S}'$. Thus:

$$|\lambda'|^2 \;=\; \frac{a^2 + b^2 + 2a + 1}{4} \;\leq\; \frac{1 + 2a + 1}{4} \;=\; \frac{1 + \Re(\lambda)}{2},$$

so $1 - |\lambda'| \leq 1 - \sqrt{\frac{1+\Re(\lambda)}{2}}$. The claim follows since for every $R$ such that $|R| \leq 1$, $1 - \sqrt{\frac{1+R}{2}} \geq \frac{1}{4}(1 - R)$. ◀
◀

## 5.2 Undirected graphs

Given an undirected graph we can easily partition it to its connected components using the fact that st-connectivity of undirected graphs is in BPL [1] (in fact, Reingold showed it is in L [30]). Therefore, we can solve the system of equations on each connected component separately.

Now, say we are given an undirected graph $G$ and $\mathcal{A}$ is its adjacency matrix. The stochastic matrix $\mathcal{S}$ associated with $G$ is $D^{-1}\mathcal{A}$, where $D$ is a diagonal matrix with the degree $\deg_i$ of the $i$-th vertex on the $i$-th element of the diagonal. While $\mathcal{A}$ is Hermitian, $\mathcal{S}$ is usually not. Still, $\mathcal{S}$ is similar to a Hermitian matrix in the following form: Express $D^{-1/2}\mathcal{A}D^{-1/2} = V\mathbf{A}V^{-1}$ where $V$ is unitary and $\mathbf{A}$ diagonal with real entries (because $D^{-1/2}\mathcal{A}D^{-1/2}$ is Hermitian), then $\mathcal{S} = (D^{-1/2}V)\mathbf{A}(D^{-1/2}V)^{-1}$. Thus, $\mathcal{S}$ has Jordan normal form $W\mathbf{A}W^{-1}$ with $W = D^{-1/2}V$. We see that

$$\begin{aligned}\kappa(W) &= \left\|D^{-1/2}V\right\| \cdot \left\|VD^{1/2}\right\| \;\leq\; \left\|D^{-1/2}\right\| \left\|D^{1/2}\right\| \|V\| \left\|V^{-1}\right\| \\ &\leq \sqrt{\frac{\lambda_{\max}(D)}{\lambda_{\min}(D)}} \;\leq\; \sqrt{\frac{n}{1}} = \sqrt{n}.\end{aligned}$$

We can therefore always take $\kappa = \sqrt{n}$ in Theorem 24 when we deal with undirected graphs, even when the graph is irregular.

The above discussion shows that $\mathcal{S}$ is similar to the diagonal matrix $\mathbf{A}$ which has a set of real eigenvalues, and therefore so does $\mathcal{S}$. Chung proved that:

▶ **Lemma 26** ([11], Lemma 1.9). *Let $\mathcal{S}$ be a transition matrix of an undirected connected graph with diameter $\Gamma$. Then $\gamma'(\mathcal{S}) \geq \frac{1}{\Gamma \cdot \sum_i \deg_i}$.*

Finally, we need the stationary distribution $\pi$. However, for an undirected graph $G = (V, E)$ the stationary distribution $\pi$ is fully explicit and gives weight $\frac{2\deg_i}{|E|}$ to the vertex $i$. Altogether, we get the theorem for undirected graphs that was stated in the introduction:

▶ **Theorem 27.** *There exists a probabilistic algorithm that gets as input an $n \times n$ stochastic matrix $\mathcal{S}$ that is the transition matrix of an undirected graph and desired accuracy and confidence parameters $\varepsilon, \delta > 0$, outputs a $(\varepsilon, \delta)$-approximation of $\mathcal{L}^\star = (\mathcal{I} - \mathcal{S})^\star$ using*

$$O\left(\log \frac{n}{\varepsilon} + \log\log \frac{1}{\delta}\right)$$

*space.*

We note that the above theorem also holds for *weighted* undirected graphs. To see this, view $\deg_i$ as the sum of weights of the $i$-th vertex, $\deg_i = \sum_j \mathcal{A}[i,j]$, which is also $\lambda_i(D)$. Then, we can take $\kappa = \sqrt{\lambda_{\max}(D)/\lambda_{\min}(D)}$ in Theorem 24. The stationary distribution is again fully explicit. Finally, analogues of Lemma 26 for weighted undirected graph show that $\gamma'(\mathcal{S})$ is at least inverse-polynomially large in the weights of the graph (e.g., Section 5 in [10]).

When $G$ is undirected we can also approximate in BPL the often used *symmetric normalized Laplacian* , which is

$$\mathcal{L}^{\mathsf{sym}} = \mathcal{I} - D^{-1/2}\mathcal{A}D^{-1/2},$$

where $\mathcal{A}$ is the graph's adjacency matrix and $D$ is the diagonal degrees matrix. We have seen that we can approximate $\mathcal{L}^\star = (\mathcal{I} - D^{-1}\mathcal{A})^\star$ in BPL, and

$$(\mathcal{L}^{\mathsf{sym}})^\star = \left(D^{1/2}\mathcal{L}D^{-1/2}\right)^\star = D^{1/2}\mathcal{L}^\star D^{-1/2}.$$

## 5.3 On the parameter $\kappa(V)$

Our algorithm's space complexity has a doubly-logarithmic dependency on $\kappa(V)$ – the minimal condition number of all Jordan bases. When the matrix $\mathcal{S}$ has well-separated eigenvalues (namely, the minimal distance between every two eigenvalues is at least polynomially-small), the dependency can be omitted. This is implied by the following theorem:

▶ **Theorem 28** ([32]). *Let $\mathcal{A}$ be an $n \times n$ matrix with eigenvalues $\lambda_1, \ldots, \lambda_n$ and suppose $\Delta > 0$ is such that $\min_{i \neq j} |\lambda_i - \lambda_j| \geq \Delta$. Also, let $\kappa_A$ be the minimal value of $\kappa(V)$ over all $V$ such that $\mathcal{A} = V\mathbf{A}V^{-1}$ is a Jordan decomposition of $\mathcal{A}$. Then, $\kappa_A \leq n \cdot e^{\frac{\|\mathcal{A}\|^2}{2\Delta^2}}$.*

We can thus conclude:

▶ **Theorem 29.** *There exists a probabilistic algorithm that gets as input:*

- *An $n \times n$ irreducible, stochastic matrix $\mathcal{S}$ and a real parameter $\Delta > 0$ so that it is guaranteed that all the eigenvalues of $\mathcal{S}$ are $\Delta$-separated (that is, $|\lambda_i - \lambda_j| \geq \Delta$ for every distinct $\lambda_i, \lambda_j \in \mathsf{Spec}(\mathcal{S})$).*

- *A parameter* $\gamma$ *such that* $\gamma'(\mathcal{S}) \geq \gamma$.
- *An approximation* $\tilde{\pi}$ *of the stationary distribution* $\pi$ *of* $\mathcal{S}$*, where* $\|\tilde{\pi} - \pi\| \leq \tau$ *and* $\tau \leq \frac{\varepsilon}{(T+1)\sqrt{n}}$ *for* $T = \frac{8n^2}{\gamma^2}\left(1 + \log\frac{n\kappa}{\varepsilon\gamma}\right)$.

*Let* $\mathcal{L}$ *denote the Laplacian,* $\mathcal{L} = \mathcal{I} - \mathcal{S}$*. Then, the algorithm outputs a* $(3\varepsilon, \delta)$*-approximation of* $\mathcal{L}^\star$ *using*

$$O\left(\log\frac{n}{\Delta\gamma\varepsilon} + \log\log\frac{1}{\delta}\right)$$

*space.*

## 5.4 Eulerian directed graphs

Eulerian graphs are directed graphs where the in-degree and out-degree of each vertex are the same, and so they generalize both regular directed graphs, and general undirected graphs. The stationary distribution is fully explicit (as in undirected graphs that we mentioned before). In this section we note that for Eulerian graphs $\gamma'$ is always non-negligible.

▶ **Claim 30.** *Let* $\mathcal{S}$ *be a transition matrix of a strongly connected Eulerian directed graph with* $m$ *edges. Then,* $\gamma'(\mathcal{S}) \geq \frac{4}{m^2}$.

**Proof.** Chung [12] proved that $\gamma'(\mathcal{S})$ is at least the second smallest eigenvalue $\mu_{n-1}$ (the smallest eigenvalue is 0) of

$$\mathcal{L}_G^{\mathsf{C}} = I - \frac{\Pi^{1/2}\mathcal{S}\Pi^{-1/2} + \Pi^{-1/2}\mathcal{S}^\dagger\Pi^{1/2}}{2},$$

where $\Pi$ is a diagonal matrix with the stationary distribution $\pi$ on the diagonal. Also, in the same paper it is proven that $\mu_{n-1} \geq \frac{4}{m^2}$, which completes the proof. ◀

## 5.5 Rapidly-mixing graphs

Finally, one way to approximate the stationary distribution is by taking a random walk on $G$ until it converges. This follows directly from Lemma 11 and the fact that $\lim_{k\to\infty} P_G^k = |\mathbf{1}\rangle\langle\pi|$ (see Theorem 3). For undirected graphs (and also Eulerian directed graphs) the walk converges in polynomial time, hence, we can approximate the stationary distribution in logarithmic space, except that there is no need to do that because we have an explicit formula for the stationary distribution anyway.

For general directed graphs (even with bounded degree) the convergence rate can be exponentially small and the approach does not work. Nevertheless, there is a whole class of directed graphs, called *rapidly-mixing graphs*, that converge rapidly even though, usually, there is no explicit formula for the stationary distribution. Clearly, for graphs where the walk converges in polynomial time we can *approximate* the stationary distribution $\pi$ in logarithmic space.

## References


1 Romas Aleliunas, Richard M. Karp, Richard J. Lipton, László Lovász, and Charles Rackoff. Random walks, universal traversal sequences, and the complexity of maze problems. In *Proceedings of the 20th Annual Symposium on Foundations of Computer Science*, pages 218–223, 1979.

2 Carme Alvarez and Raymond Greenlaw. A compendium of problems complete for symmetric logarithmic space. *Computational Complexity*, 9(2):123–145, 2000.

3 Frank Bauer. Normalized graph laplacians for directed graphs. *Linear Algebra and its Applications*, 436(11):4193–4222, 2012.

4 Adi Ben-Israel and Thomas N. E. Greville. *Generalized inverses: theory and applications*, volume 15. Springer, 2003.

5 Stuart J. Berkowitz. On computing the determinant in small parallel time using a small number of processors. *Information Processing Letters*, 18(3):147–150, 1984. doi:10.1016/0020-0190(84)90018-8.

6 Anders Björner and László Lovász. Chip-firing games on directed graphs. *Journal of Algebraic Combinatorics*, 1(4):305–328, 1992.

7 Allan Borodin, Joachim von zur Gathen, and John E. Hopcroft. Fast parallel matrix and GCD computations. *Information and Control*, 52(3):241–256, 1982. doi:10.1016/S0019-9958(82)90766-5.

8 Richard Bronson. *Matrix Methods: an Introduction*. Gulf Professional Publishing, 1991.

9 William Clough Brown. *A Second Course in Linear Algebra*. Wiley-Interscience, 1988.

10 Fan R. K. Chung. Laplacian of graphs and cheeger's inequalities. *Combinatorics, Paul Erdos is Eighty*, 2(157-172):13–2, 1996.

11 Fan R. K. Chung. *Spectral Graph Theory*. American Mathematical Society, 1997.

12 Fan R. K. Chung. Laplacians and the cheeger inequality for directed graphs. *Annals of Combinatorics*, 9(1):1–19, 2005.

13 Michael B. Cohen, Jonathan Kelner, John Peebles, Richard Peng, Aaron Sidford, and Adrian Vladu. Faster algorithms for computing the stationary distribution, simulating random walks, and more. In *Proceedings of the 57th Annual Symposium on Foundations of Computer Science*, pages 583–592, 2016.

14 Laszlo Csanky. Fast parallel matrix inversion algorithms. *SIAM Journal of Computing*, 5(6):618–623, 1976.

15 James W. Demmel. *Applied numerical linear algebra*. Society for Industrial and Applied Mathematics, 1997.

16 Dean Doron, Amir Sarid, and Amnon Ta-Shma. On approximating the eigenvalues of stochastic matrices in probabilistic logspace. *Computational Complexity*, pages 1–28, 2016.

17 Dean Doron and Amnon Ta-Shma. On the problem of approximating the eigenvalues of undirected graphs in probabilistic logspace. In *Proceedings of the 42nd International Colloquium on Automata, Languages, and Programming*, pages 419–431, 2015.

18 Bill Fefferman and Cedric Yen-Yu Lin. A complete characterization of unitary quantum space. *arXiv preprint arXiv:1604.01384*, 2016.

19 Peter W. Glynn. Upper bounds on poisson tail probabilities. *Operations Research Letters*, 6(1):9–14, 1987.

20 Chris Godsil. Eigenvalues of graphs and digraphs. *Linear Algebra and its Applications*, 46:43–50, 1982.

21 Chris Godsil and Gordon F. Royle. *Algebraic Graph Theory*. Springer, 2013.

22 Gene H. Golub and Charles F. Van Loan. *Matrix Computations*. Johns Hopkins University Press, Baltimore, MD, USA, 3rd edition, 1996.

23 Nicholas J. Higham. *Functions of Matrices: Theory and Computation*. Society for Industrial and Applied Mathematics, 2008.

24 Jonathan A. Kelner, Lorenzo Orecchia, Aaron Sidford, and Zeyuan Allen Zhu. A simple, combinatorial algorithm for solving SDD systems in nearly-linear time. In *Proceedings of the 45th Annual Symposium on Theory of Computing*, pages 911–920, 2013. doi:10.1145/2488608.2488724.

25 Adam R. Klivans and Dieter Van Melkebeek. Graph nonisomorphism has subexponential size proofs unless the polynomial-time hierarchy collapses. *SIAM Journal on Computing*, 31(5):1501–1526, 2002.

**26** Harry R. Lewis and Christos H. Papadimitriou. Symmetric space-bounded computation. *Theoretical Computer Science*, 19(2):161–187, 1982.

**27** Noam Nisan. Pseudorandom generators for space-bounded computation. *Combinatorica*, 12(4):449–461, 1992.

**28** Noam Nisan. RL⊆SC. *Computational Complexity*, 4(1):1–11, 1994.

**29** Richard Peng and Daniel A. Spielman. An efficient parallel solver for SDD linear systems. In *Proceedings of the 46th Annual Symposium on Theory of Computing*, pages 333–342, 2014. doi:10.1145/2591796.2591832.

**30** Omer Reingold. Undirected connectivity in log-space. *Journal of the ACM*, 55(4), 2008.

**31** Michael E. Saks and Shiyu Zhou. $\mathrm{BP_HSPACE(S)} \subseteq \mathrm{DSPACE}(\mathrm{S}^{3/2})$. *Journal of Computer and System Sciences*, 58(2):376–403, 1999.

**32** Russell A. Smith. The condition numbers of the matrix eigenvalue problem. *Numerische Mathematik*, 10(3):232–240, 1967.

**33** Daniel A. Spielman and Shang-Hua Teng. Nearly-linear time algorithms for graph partitioning, graph sparsification, and solving linear systems. In *Proceedings of the 36th Annual Symposium on Theory of Computing*, pages 81–90, 2004. doi:10.1145/1007352.1007372.

**34** Daniel A. Spielman and Shang-Hua Teng. Spectral sparsification of graphs. *SIAM Journal on Computing*, 40(4):981–1025, 2011. doi:10.1137/08074489X.

**35** Daniel A. Spielman and Shang-Hua Teng. A local clustering algorithm for massive graphs and its application to nearly linear time graph partitioning. *SIAM Journal on Computing*, 42(1):1–26, 2013. doi:10.1137/080744888.

**36** Daniel A. Spielman and Shang-Hua Teng. Nearly linear time algorithms for preconditioning and solving symmetric, diagonally dominant linear systems. *SIAM Journal on Matrix Analysis and Applications*, 35(3):835–885, 2014. doi:10.1137/090771430.

**37** Amnon Ta-Shma. Inverting well conditioned matrices in quantum logspace. In *Proceedings of the 45th Annual Symposium on Theory of Computing*, pages 881–890, 2013. doi:10.1145/2488608.2488720.

**38** Vladimir Trifonov. An $O(\log n \log\log n)$ space algorithm for undirected st-connectivity. *SIAM Journal on Computing*, 38(2):449–483, 2008.

**39** Nisheeth K. Vishnoi. *$Lx = b$ – Laplacian Solvers and their Algorithmic Applications*. Now publishers, 2013.

**40** John Watrous. Space-bounded quantum complexity. *Journal of Computer and System Sciences*, 59(2):281–326, 1999.