Diffusion process as a computational engine: integer factorisation algorithm

Carlos A. Cadavid, Paulina Hoyos, Jay Jorgenson,
Lejla Smajlović, Juan D. Vélez

Dubrovnik, October 04, 2022.

1. The diffusion as a (theoretical) computational engine

2. Random walks on regular graphs. Diffusion step.

3. Integer factorization

Heat diffusion

Assume that a heat source, such as a flame or a welding torch, is applied to the center of a circular disc of uniform thickness and material composition.

Then two observers who are measuring temperatures at different points on the perimeter will detect a change of temperature at their points of contact at the same rate.

Heat diffusion

Assume that a heat source, such as a flame or a welding torch, is applied to the center of a circular disc of uniform thickness and material composition.

Then two observers who are measuring temperatures at different points on the perimeter will detect a change of temperature at their points of contact at the same rate.

Certain aspects of diffusion are simultaneously, not sequentially, observable!

Diffusion of light

Within a circular ring, imagine a beam of light $\mathcal{B}$ (or some type of focused energy) emanating from a source at a perimeter point $\mathcal{P}_0$. Upon contact with another perimeter point $\mathcal{P}_1$ on the ring, the beam $\mathcal{B}$ splits into M sub-beams of equal magnitude in a prescribed set of directions toward perimeter points $\mathcal{P}_{2,1}, \cdots \mathcal{P}_{2,M}$.

Let each sub-beam upon contact with some $\mathcal{P}_{2,k}$ split in manner as similar to the reflection of $\mathcal{B}$ at $\mathcal{P}_1$, and so on.

Diffusion of light

Within a circular ring, imagine a beam of light $\mathcal{B}$ (or some type of focused energy) emanating from a source at a perimeter point $\mathcal{P}_0$. Upon contact with another perimeter point $\mathcal{P}_1$ on the ring, the beam $\mathcal{B}$ splits into M sub-beams of equal magnitude in a prescribed set of directions toward perimeter points $\mathcal{P}_{2,1}, \cdots \mathcal{P}_{2,M}$.

Let each sub-beam upon contact with some $\mathcal{P}_{2,k}$ split in manner as similar to the reflection of $\mathcal{B}$ at $\mathcal{P}_1$, and so on.

Then after n such splittings, what portion of the original amount energy has returned to $\mathcal{P}_0$?

Diffusion step - intuitive definition

Diffusion step = one instance of contact, reflecting and subsequent splitting.

More precisely: the count will be mathematically captured as one iteration of a symmetric matrix on a finite dimensional vector space.

Diffusion step - intuitive definition

Diffusion step = one instance of contact, reflecting and subsequent splitting.

More precisely: the count will be mathematically captured as one iteration of a symmetric matrix on a finite dimensional vector space.

In the above example: Each contact involves the splitting of a single beam into M sub-beams, then after n diffusion steps, one will have M^n paths of light traversing the ring.

Imagine a ring were 1 kilometer in diameter, and if the beam were to travel at the speed of light, then after 0.01 seconds one would expect to have more than M^{3000} sub-beams crossing various chords of the ring since in almost all circumstances more than 3000 diffusion steps would have taken place.

A natural question

Q: What is the most natural mathematical setup for describing/measuring such diffusion?

A natural question

Q: What is the most natural mathematical setup for describing/measuring such diffusion?

A: A Regular (weighted) graph

Regular weighted graph

X is undirected connected graph with:

- Finite set of vertices V
- Set of edges $E =$ collection of two element subsets of V
- Real valued weight function $w :: V \times V \to \mathbb{R}_{\geq 0}$ which is symmetric and $w(x, y) > 0$ iff $\{x, y \in E\}$.

The degree of a vertex $x \in V$ is

$$d(x) = \sum_{y \in V} w(x, y).$$

A weighted graph X is **regular** of degree d if $d(x) = d$ for all vertices $x \in V$.

Important example: Cayley graph

G a finite abelian group

$S \subseteq G$ a fixed symmetric subset (sometimes one assumes it generates G).

The symmetry condition means that $s \in S$ iff $-s \in S$ (or $s \in S$ iff $s^{-1} \in S$).

For any $\alpha\colon S \to \mathbb{R}^{>0}$ such that $\alpha(s) = \alpha(-s)$, one can construct a weighted Cayley graph $X = \text{Cay}(G, S, \alpha)$ of G with respect to S and α as follows:

- The vertices of X are the elements of G
- Two vertices x and y are connected with an edge if and only if $x - y \in S$ (additive notation)
- The weight $w(x, y)$ of the edge (x, y) is $w(x, y) = \alpha(x - y)$

Then$X = \text{Cay}(G, S, \alpha)$ is a regular weighted graph of degree $d = \sum_{s \in S} \alpha(s)$.

Random walks

X a regular weighted graph of degree d

A adjacency matrix (non-negative, symmetric!)

A half-lazy random walk on X is a Markov chain with state space $(V, \mathcal{P}(V))$ with arbitrary initial probability distribution $p_0 \colon V \to \mathbb{R}$, and the transition probability matrix

$$W := \frac{1}{2}\left(I + \frac{1}{d}A\right).$$

After n steps

$$p_n(x) = W^n p_0(x).$$

Random walks - convergence

As expected, probability distribution p_n tends to the uniform distribution as $n \to \infty$.

The rate of convergence is:

$$sup_{x \in V} \left| p_n(x) - \frac{1}{|V|} \right| \leq \lambda_1^n,$$

λ_1 is the largest eigenvalue of W less than 1.

Note that eigenvalues of W are nonnegative with the largest eigenvalue equal to 1 with multiplicity 1.

(Follows easily from the fact that A is symmetric with eigenvalues in $[-d, d]$.)

"Diffusion computabililty"

A diffusion process in $X = \mathrm{Cay}(G, S, \alpha)$ may be regarded as an analog computation on X

Definition

A real-valued function h on G is said to be computable by a diffusion process in X with initial condition $p_0 \colon G \to \mathbb{R}$ if the following holds. Let $\{p_m = W^m p_0\}_{m=0}^{\infty}$ be the sequence of distributions in X with initial probability distribution p_0. Then for any given $\varepsilon > 0$ there exists a positive integer $n = n(\varepsilon)$ such that for all $m > n(\epsilon)$ and all $x \in G$, we have that

$$|p_m(x) - h(x)| < \varepsilon.$$

Example: Projections of any real-valued function on G onto the eigenspaces of the matrix W are diffusion computable

"Diffusion computabililty"

A diffusion process in $X = \mathrm{Cay}(G, S, \alpha)$ may be regarded as an analog computation on X

Definition

A real-valued function h on G is said to be computable by a diffusion process in X with initial condition $p_0 \colon G \to \mathbb{R}$ if the following holds. Let $\{p_m = W^m p_0\}_{m=0}^{\infty}$ be the sequence of distributions in X with initial probability distribution p_0. Then for any given $\varepsilon > 0$ there exists a positive integer $n = n(\varepsilon)$ such that for all $m > n(\epsilon)$ and all $x \in G$, we have that

$$|p_m(x) - h(x)| < \varepsilon.$$

Example: Projections of any real-valued function on G onto the eigenspaces of the matrix W are diffusion computable

Diffusion step

In a diffusion computer, computation begins with a stochastic vector, say ϕ (i.e. with initial probability distribution p_0 on V).

The evolution of ϕ is determined by the symmetric operator W.

One diffusion computation step $=$ one application of W which maps ϕ to $W\phi$.

A measurement is a classical inspection of the vertices of the graph X where the diffusion process takes place.

Quantum versus diffusion steps

The principal manager of the quantum-computing group at Microsoft Research in Redmond, Washington said that

Quantum computing is essentially matrix vector multiplication - it's linear algebra underneath the hood.

A quantum computation entails two different types of operations:

1. The abstract version of the classical evolution equation in quantum mechanics, i.e. a unitary vector ϕ in a Hilbert space $(\mathbb{C}^n, \ell^2)$ evolves into a new vector $\psi = U\phi$ (U is some unitary operator).
2. Measurement of this new state - quantum procedure of "collapsing ψ": each measurement is modeled by the decomposition of $\mathbb{C}^n$ into finite orthogonal subspaces H_i. "Collapsing" means composing ψ with the projection ψ_i onto H_i, and this projection occurs with probability $|\psi_i|^2$.

Elementary considerations

N a positive, odd integer which we write as a product

$$N = \prod_{i=1}^{m} p_i^{e_i}$$

with $m \geq 2$ different odd prime factors with exponents $e_i > 0$.

Assume N is not prime nor a prime power.

To find a factor of N it suffices to find x so that

$$x^2 \equiv 1 \bmod N \quad \textit{and} \quad x \not\equiv \pm 1 \bmod N. \tag{1}$$

Hint:

$$N \mid x^2 - 1 \quad \text{but} \quad N \nmid (x \pm 1).$$

The algorithm

Step 1.

Pick $a \in \mathbb{Z}_N = \{1, \ldots, N\}$ uniformly at random; compute $d = \gcd(a, N)$.

If $1 < d < N$, return d. Else, go to Step 2.

Step 2.

Let $M = \lfloor \log_2 N \rfloor + 1$ and compute the set (modular arithmetic)

$$S = \{a^{\pm 2^t} \bmod N : t = 0, \ldots, M\}$$

Lemma (proved in the paper): If there are repetitions in S, then with probability $p(m) = 1 - (m+1)/2^m$ we can find an $x \in \mathbb{Z}_N^*$ satisfying (1) in at most $O(\log_2 N)$ deterministic steps.

What happens in the case when there are no repetitions in S?

Case no repetitions in S

Step 3. Set $b = a^{2^M}$ and run the diffusion computer algorithm to determine the order r_b of b modulo N.

Important: the order r_b must be odd!

Step 4. Compute the smallest integer $k \geq 0$ such that $a^{2^k r_b} \equiv 1 \bmod N$.

Set $r_a = 2^k r_b$ which is the order of a modulo N.

If r_a is even, compute $d = \gcd(a^{r_a/2} - 1, N)$.

This algorithm produces a factor of N with probability $1 - (m+1)/2^m$ after $O((\log N)^2)$ deterministic steps and some number (tbd) of diffusion steps.

Construction of the adopted graph - multiplicative notation

Problem to be solved: Given the number $b = a^{2^M} \in \mathbb{Z}_N^*$ determine the order r_b of b modulo N using diffusion on a suitable graph.

We know that $r_b = r$ is odd and elements of S are distinct.

Our graph is the weighted Cayley graph $X_{N,b} = \mathrm{Cay}(G_{N,b}, S_{N,b}, \alpha_{N,b})$, where

$G_{N,b} = \langle b \rangle \subseteq \mathbb{Z}_N^*$ is the subgroup of $\mathbb{Z}_N^*$ generated by b (it has r elements)

$$S_{N,b} = \{b^{\pm 2^t} : \, t = 0, \ldots, M\}$$

and

$$\alpha_{N,b}(b^{2^t}) := |\{l \in \{0, \ldots, M\} \colon \, b^{2^t} \equiv b^{2^l} \bmod N \, \text{ or } \, b^{2^t} \equiv b^{-2^l} \bmod N\}|,$$

Important remark

In the construction of $X_{N,b}$ we do not know the value of r. All that is required is the value of N since we begin with one point b and, recursively, let the diffusion process develop in $2(M+1)$ possible directions from any given point.

From the beginning, we do not know the entire graph. Nevertheless, since diffusion is local in nature, this allows us to build $X_{N,b}$ one diffusion step at a time.

Important remark

In the construction of $X_{N,b}$ we do not know the value of r. All that is required is the value of N since we begin with one point b and, recursively, let the diffusion process develop in $2(M+1)$ possible directions from any given point.

From the beginning, we do not know the entire graph. Nevertheless, since diffusion is local in nature, this allows us to build $X_{N,b}$ one diffusion step at a time.

We will see that after a number of steps which is polynomial in $\log_2 N$ we will have enough information to approximate the number of vertices of $X_{N,b} = r$.

What makes the process effective is the fact that diffusion occurs simultaneously at all constructed vertices, which provides some form of parallel computation.

Construction of the adopted graph - additive notation

We replace $X_{N,b} = \mathrm{Cay}(G_{N,b}, S_{N,b}, \alpha_{N,b})$ with an equivalent graph (b is fixed, so omitted from notation)

$X_{r,S} = \mathrm{Cay}(C_r, S, \alpha)$ where $C_r = \{0, \dots, r-1\}$,

$$S = \{\pm 2^j : j = 0, \dots, M\} \quad \text{with} \quad M = \lfloor \log_2 N \rfloor + 1.$$

and

$$\alpha(2^j) := |\{l \in \{0, \dots, M\} \colon 2^j \equiv 2^l \bmod r \;\text{ or }\; 2^j \equiv -2^l \bmod r\}|.$$

$X_{r,S}$ has r vertices, and it is regular of degree $|S| = 2(M+1)$

Diffusion on $X_{r,s}$

Fix any point of $X_{r,s}$, say 0 and start diffusion with initial probability $p_0 = (1, 0 \dots, 0)^t$.

Recall that each step amounts to multiplying p_0 by

$$W_{r,s} = \frac{1}{2}\left(1 + \frac{1}{2(M+1)} A_{r,s}\right),$$

where $A_{r,s}$ is the adjacency matrix of $X_{r,s}$

We know that after n steps

$$\left| p_n^{X_{r,s}}(0) - \frac{1}{r} \right| \leq (\lambda_*^{X_{r,s}})^n, \tag{2}$$

where $\lambda_*^{X_{r,s}}$ is the largest eigenvalue of $W_{r,s}$ less than 1.

Minimal number of diffusion steps to find r

Therefore, if we conduct n steps where n is such that $(\lambda_*^{X_{r,s}})^n < \frac{1}{N^2}$, the value $p_n^{X_{r,s}}(0)$ which can be regarded as the "amount of heat" at initial point after n steps determines r uniquely through

$$r = \lfloor (p_n^{X_{r,s}}(0))^{-1} \rfloor.$$

Namely, for any two distinct positive integers $m_1, m_2 < N$, the smallest distance between $1/m_1$ and $1/m_2$ is bounded from below by

$$\frac{1}{N} - \frac{1}{N-1} = \frac{1}{N(N-1)} > \frac{1}{N^2}.$$

Thus, if we have $p_n^{X_{r,s}}(e) = 1/r$ within an error of $1/N^2$, we have determined r.

What is $\lambda_*^{X_{r,s}}$

Use known results from the graph theory:

The eigenvalues η_k for $k = 0, \ldots, r-1$ of the adjacency matrix $A_{r,S}$ are: $\eta_0 = 2(M+1)$ and

$$\eta_k = \sum_{x \in S} \alpha(x) e^{\frac{2\pi i}{r} kx} = \sum_{j=0}^{M} e^{\frac{2\pi i}{r} k2^j} + \sum_{j=0}^{M} e^{-\frac{2\pi i}{r} k2^j} \quad \text{for each} \quad 1 \leq k \leq r-1.$$

Therefore, eigenvalues of $W_{r,S}$ are

$$\lambda_k^{X_{r,s}} = \frac{1}{2}\left(1 + \frac{\eta_k}{2(M+1)}\right) \quad \text{for} \quad k = 0, 1, \ldots, r-1.$$

and $\lambda_*^{X_{r,s}}$ is the largest of $\lambda_k^{X_{r,s}}$, $k = 1, \ldots, r-1$.

Bounding $\lambda_*^{X_{r,S}}$

One needs a bound independent of r for the quantity

$$\frac{1}{2(M+1)}\max_{k\in\{1,\dots r-1\}}|\eta_k| = \frac{1}{2(M+1)}\max_{k\in\{1,\dots r-1\}}\left|\sum_{j=0}^{M} e^{\frac{2\pi i}{r}k2^j}+\sum_{j=0}^{M} e^{-\frac{2\pi i}{r}k2^j}\right|.$$

This is a non-trivial task because the trigonometric sum is short.

Using a combination of results one arrives at the bound

$$\frac{1}{2(M+1)}\max_{k\in\{1,\dots r-1\}}|\eta_k| < \frac{2M}{2(M+1)} = 1-\frac{1}{M+1},$$

which yields the bound

$$\lambda_*^{X_{r,S}} \le \frac{1}{2}\left(1+1-\frac{1}{M+1}\right) = 1-\frac{1}{2(M+1)}.$$

Minimal number of diffusion steps

We can now find n so that $(\lambda_*^{X_{r,s}})^n < \frac{1}{N^2}$ by solving

$$\left(1 - \frac{1}{2(M+1)}\right)^n < \frac{1}{N^2}.$$

It suffices to take the smallest integer that is $> 2 \log N(\lfloor \log_2 N \rfloor + 2)$.

Minimal number of diffusion steps

We can now find n so that $(\lambda_*^{X_{r,s}})^n < \frac{1}{N^2}$ by solving

$$\left(1 - \frac{1}{2(M+1)}\right)^n < \frac{1}{N^2}.$$

It suffices to take the smallest integer that is $> 2\log N(\lfloor \log_2 N \rfloor + 2)$.

Conclusion: we proved that after

$$n = \lfloor 2\log N(\lfloor \log_2 N \rfloor + 2) \rfloor + 1$$

heat steps the value r equals $\lfloor (p_n^{X_{r,s}}(0))^{-1} \rfloor$, where $p_n^{X_{r,s}}(0)$ is the heat at initial point.

Minimal number of diffusion steps

We can now find n so that $(\lambda_*^{X_{r,s}})^n < \frac{1}{N^2}$ by solving

$$\left(1 - \frac{1}{2(M+1)}\right)^n < \frac{1}{N^2}.$$

It suffices to take the smallest integer that is $> 2\log N(\lfloor \log_2 N \rfloor + 2)$.

Conclusion: we proved that after

$$n = \lfloor 2\log N(\lfloor \log_2 N \rfloor + 2) \rfloor + 1$$

heat steps the value r equals $\lfloor (p_n^{X_{r,s}}(0))^{-1} \rfloor$, where $p_n^{X_{r,s}}(0)$ is the heat at initial point.

Note that the minimal number of quantum steps to compute the order is $O((\log N)^2 \log\log N)$.

Overview of the algorithm

The algorithm runs Step 1 and either finds a factor or proceeds to Step 2.

At Step 2 we use repeated squaring, so it takes at most $O(2M)$ steps. The algorithm can terminate at Step 2 with no answer; the probability of success if there are repetitions is at least $1 - (m+1)/2^m$.

Overview of the algorithm

The algorithm runs Step 1 and either finds a factor or proceeds to Step 2.

At Step 2 we use repeated squaring, so it takes at most $O(2M)$ steps. The algorithm can terminate at Step 2 with no answer; the probability of success if there are repetitions is at least $1-(m+1)/2^m$.

So, we may have to take another a after Step 2.

Overview of the algorithm

The algorithm runs Step 1 and either finds a factor or proceeds to Step 2.

At Step 2 we use repeated squaring, so it takes at most $O(2M)$ steps. The algorithm can terminate at Step 2 with no answer; the probability of success if there are repetitions is at least $1-(m+1)/2^m$.

So, we may have to take another a after Step 2.

If there are no repetitions in S, Step 3 is run on a diffusion computer and gives the answer r.

Then, Step 4 is run on a classical computer and produces a factor of N with probability at least $1-(m+1)/2^m$.

Overview of the algorithm

The algorithm runs Step 1 and either finds a factor or proceeds to Step 2.

At Step 2 we use repeated squaring, so it takes at most $O(2M)$ steps. The algorithm can terminate at Step 2 with no answer; the probability of success if there are repetitions is at least $1-(m+1)/2^m$.

So, we may have to take another a after Step 2.

If there are no repetitions in S, Step 3 is run on a diffusion computer and gives the answer r.

Then, Step 4 is run on a classical computer and produces a factor of N with probability at least $1-(m+1)/2^m$.

Therefore, the algorithm terminates after at most $O(\log N)^2$ deterministic steps plus at most $O(\log N \log_2 N)$ diffusion steps and finds a factor of N with probability at least $1-(m+1)/2^m$.

Example: $N = 1363$

Step 1. We choose $a = 991$ which is relatively prime to 1363.

Step 2. $M = \lfloor \log_2(N) \rfloor + 1 = 11$, and $S = \{991^{2^0} = 991, 991^{2^1} = 721, \ldots, 991^{2^{11}} = 944, \ldots\} \bmod 1363$. There are no repetitions in S.

Step 3. Set $b = 991^{2^{11}} \equiv 944 \bmod 1363$ and check for repetitions in the set $S_b = \{b^{\pm 2^t} : t = 0, \ldots, 11\}$ finding none. Thus, we run the diffusion computer in order to determine the order r_b of $b = 944$.

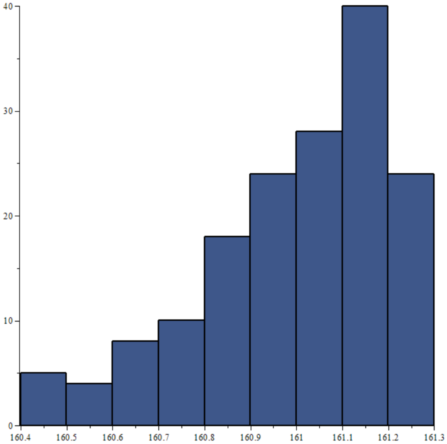
Note: Theoretical bound says that the diffusion computer requires at least $\lfloor 4(M+1)\log(N) \rfloor + 1 = 347$ diffusion steps and one measurement.

Actually, after $n = 25$ iterations, and 11 measurements we were able to conclude that $r_b = 161$.

Step 5. The smallest non negative integer k such that $991^{2^k \times 161} \equiv 1 \bmod 1363$ turns out to be 1. We conclude that $r_a = 322$.
Then we computed $\gcd(991^{161} - 1, 1363) = 47$, thus, $N = 47 \times 29$.

Example: $N = 1363$ cont'

After $n = 25$ times we measured $p_{25}(v)$ for the 11 values of vertices v corresponding to S. For each such v we had that $160 < p_{25}(v)^{-1} < 162$, hence $r_b \in \{160, 161, 162\}$. By trying these values we confirmed that $r_b = 161$.



25 steps.png

The end

Thank you!