On the Number of Binary-Minded Individuals Required To Compute $\sqrt{\frac{1}{2}}$

Olivier Bournez, Guillaume Aupy

On the Number of Binary-Minded Individuals Required To Compute $\sqrt{\frac{1}{2}}$.

Guillaume Aupy[a], Olivier Bournez*,[b]

[a] *Ecole Normale Supérieure de Lyon, 46 Allée d'Italie, 69007 Lyon, FRANCE*
[b] *LIX, Ecole Polytechnique, 91128 Palaiseau Cedex, FRANCE*

Abstract

We recently obtained partial results on the computational power of population protocols when population is assumed to be huge.

We studied in particular a particular protocol that we proved to converge towards $\sqrt{\frac{1}{2}}$, using weak-convergence methods for stochastic processes.

In this note, we prove that this is possible to compute $\sqrt{\frac{1}{2}}$ with precision $\epsilon > 0$ in a time polynomial in $\frac{1}{\epsilon}$ using a number of agents polynomial in $\frac{1}{\epsilon}$, with individuals that can have only two states.

This is established through a general result on approximation of stochastic differential equations by a stochastic Euler like discretization algorithm, of general interest.

1. Introduction

The computational power of networks of finitely many anonymous resource-limited mobile agents has been investigated in several recent papers. In particular, the population protocol model, introduced in [1], consists of a population of finite-state agents that interact in pairs, where each interaction updates the state of both participants according to a transition based on the previous states of the participants. When all agents converge after some finite time to a common value, this value represents the result of the computation.

Their computational power has been investigated under several hypotheses but always when restricted to finite size populations. Predicates stably computable by population protocols in this sense have been characterized as being precisely the semi-linear predicates, that is to say those predicates on counts of input agents definable in first-order Presburger arithmetic [2]. Semi-linearity was shown to be sufficient in [1] and necessary in [3].

*Corresponding author
Email addresses: Guillaume.Aupy@ens-lyon.fr (Guillaume Aupy), Olivier.Bournez@polytechnique.fr (Olivier Bournez)

Refer to [4] for a survey on results obtained for finite size population protocols.

In a recent paper [5], we started to investigate the computational power of population protocols when the size of the population goes to infinity. In particular, we considered the following example. Assume that we have a population of agents that can be either in state + or in state −. Assume that this population is huge, and that at each discrete time step, two agents are paired. These two agents are chosen according to a uniform law (without choosing twice the same). The effect of a pairing is given by the following rules:

$$\begin{cases} ++ & \to & +- \\ +- & \to & ++ \\ -+ & \to & ++ \\ -- & \to & +- \end{cases} \tag{1}$$

These rules must be interpreted as follows: if an agent in state + is paired with an agent in state +, then the second becomes −. If an agent in state + is paired with an agent in state −, then the second becomes +, and symmetrically. If an agent in state − is paired with an agent in state −, then the first becomes in state +.

Suppose that we want to discuss the limit of the proportion $p(k)$ of agents in state + in the population at discrete time k. If $n_+(k)$ denotes the number of agents in state +, and $n_-(k) = n - n_+(k)$ the number of agents in state −,

$$p(k) = \frac{n_+(k)}{n}.$$

Since we are dealing with n indistinguishable agents, the population protocol is completely described by the number of agents in state +. We are then reduced to determine the evolution of the Markov chain

$$(p(k))_{k\in\mathbb{N}} \in \left\{\frac{0}{n}, \frac{1}{n}, \ldots, \frac{n}{n}\right\}.$$

If we put aside the special configuration where all agents are in state − which is immediately left in any next round, any configuration is reachable from any configuration: Hence, $(p(k))$ is an homogeneous irreducible Markov chain in $\{\frac{1}{n}, \ldots, \frac{n}{n}\}$.

A consequence of the ergodic theorem is that the chain $(p(k))$ admits a unique stationary distribution μ.

It is easy to see that this must be an element of $\mathbb{Q}^n$. Hence, its mean $\sum_i \mu(i/n)i/n$ is a rational number, that we denote $p^{(n)}$.

A second consequence of the ergodic theorem is the following convergence :

$$\frac{p(1) + p(2) + ... + p(k)}{k} \overset{k\to\infty}{\to} p^{(n)}, \text{ almost surely.}$$

We proved in [5], that when n goes to infinity, the mean value of $p(k)$ converges to the irrational number $\sqrt{\frac{1}{2}}$. This was obtained through some weak-

convergence methods, based on theoretical results on stochastic processes and their approximation by stochastic discrete time sequences.

By several aspects these results are non-constructive, and not usable to give any bounds on the number of individuals, nor the time required to compute $\sqrt{\frac{1}{2}}$ at precision ϵ.

The purpose of the current note is to circumvent the problem and show that this is possible to compute $\sqrt{\frac{1}{2}}$ with precision $\epsilon > 0$ in a time polynomial in $\frac{1}{\epsilon}$ using a number of agents polynomial in $\frac{1}{\epsilon}$.

This is established through a general result on approximation of stochastic differential equations by a stochastic Euler like discretization algorithm. This provides a convergence result which is more directly exploitable from a computer science point of view (we have bounds on required size and time).

We believe the approach useful in many other contexts. In particular, this works for general population protocols considered in [5], as we will see in next section.

2. $\sqrt{\frac{1}{2}}$ vs General Case

Formally, if $p_n(k)$ denotes the proportion of $+$ in the population of size n at time k, we established in [5] that one can always write

$$p_n(k+1) = p_n(k) + \frac{1}{n}\left(1 - 2\frac{n}{n-1}p_n(k)^2 + \frac{2}{n-1}p_n(k)\right) + \frac{1}{n}\rho_n(k).$$

This can be rewritten

$$p_n(k+1) - p_n(k) = \frac{1}{n}F(p_n(k)) + \frac{1}{n}\epsilon_n(k) + \frac{1}{n}\rho_n(k),$$

where

$$F(x) = 1 - 2x^2,$$

and $\epsilon_n(k) = \frac{2}{n-1}(p_n(k) - p_n(k)^2)$, can be seen as a *deterministic* $\mathcal{O}(\frac{1}{n})$ perturbation, and $\rho_n(k)$ as a *randomized* perturbation, that is to say a random variable, that takes value in interval $[-2, 2]$.

The result then follows from Theorem 1, proved in next section: observe that the ordinary differential equation $\frac{dx}{dt} = F(x)$ is explicitly solvable (we thanks our formal calculus software about reminding us how to solve this type of rational fraction equations): $x(t) = \frac{e^{2\sqrt{2}(t-c_1)}-1}{\sqrt{2}(e^{2\sqrt{2}(t-c_1)}+1)}$, where c_1 is some constant, fixed by the initial condition. Whatever the initial condition is, it converges to $\sqrt{\frac{1}{2}}$. A simple asymptotic development of $x(t) - \sqrt{\frac{1}{2}}$ shows that taking $t = T(\epsilon)$ with $T(\epsilon) > -\frac{1}{2\sqrt{2}}ln(\frac{\epsilon}{\sqrt{2}}) + c_1$ guarantees $|x(t) - \sqrt{\frac{1}{2}}| \leq \epsilon$.

Actually, this works for general population protocols. Following [5], the transition rules of a general population protocol are of the form

$$q\ q' \to \delta_1(q,q')\ \delta_2(q,q')$$

for all $(q_1, q_2) \in Q^2$.

For general protocols, from the proof of Theorem 4 in [5], for $P_n(k)$ describing the vector of $K = [0,1]^Q \subset \mathbb{R}^Q$ whose components are the proportions of agents in the different state at time k for populations of size n, we still have

$$P_n(k+1) = P_n(k+1) + b(P_n(k)) + \frac{1}{n}\epsilon_n(k) + \frac{1}{n}\rho_n(k),$$

where

$$b(x) = \sum_{(q,q')\in Q} x_q x_{q'}(-(e_q + e'_q) + e_{\delta_1(q,q')} + e_{\delta_2(q,q')}),$$

$(e_q)_{q\in Q}$ is the canonical base of $\mathbb{R}^Q$, where $\epsilon_n(k)$ can be seen as a *deterministic* $\mathcal{O}(\frac{1}{n})$ perturbation, and $\rho_n(k)$ is a *randomized* perturbation, that is to say a random variable, that takes value in interval $[-2, 2]$.

It then also follows from Theorem 1, proved in next section, that if the ordinary differential equation

$$\frac{dX}{dt} = b(X),$$

is efficiently globally convergent, then its limit can be approximated at ϵ in a time polynomial in $\frac{1}{\epsilon}$ with a number of agents of size polynomial in $\frac{1}{\epsilon}$.

3. Approximating A Stochastic Differential Equation by An Euler Like Method

Theorem 1. *Assume that $F : K \subset \mathbb{R}^d \to \mathbb{R}^d$ is some $\mathcal{C}^1$ function over some compact K.*

Assume that ordinary differential equation (ODE)

$$\frac{dX}{dt} = F(X) \tag{2}$$

over $K \subset \mathbb{R}^d$ is globally convergent*: there is some $x^* \in K$, such that for all ϵ, there is some $T(\epsilon)$ so that, whatever $X(0)$ is, any solution of the ODE is such that $||X(t) - x^*|| \leq \epsilon$ for $t \geq T(\epsilon)$.*

Assume that it is moreover it is efficiently globally convergent*: we also have that $T(\epsilon)$ is in $\mathcal{O}(\ln 1/\epsilon)$.*

Asume that $(P_n(k))_k$ is a sequence of random variables taking values in compact K, and c and d are two integers so that for all n and k,

- *we have*

$$P_n(k+1) - P_n(k) = \frac{1}{n}F(P_n(k)) + \frac{1}{n}\epsilon_n(k) + \frac{1}{n}\rho_n(k),$$

- *where $\epsilon_n(k)$ is a deterministic term taking value in $[-\frac{d}{n}, \frac{d}{n}]$,*
- *and $\rho_n(k)$ is random variable taking value in interval $[-c, c]$,*

Then for any precision $\epsilon > 0$ arbitrary close to 0 and probability $0 < \mu < 1$ arbitrary close to 1, one can consider some integers n and k that guarantees that whatever the initial condition $P_n(0)$ is, we have

$$||P_n(k) - x^*|| \leq \epsilon.$$

Moreover, whenever μ is fixed, $n = n(\epsilon)$ and $k = k(\epsilon)$ can be taken bounded by a polynomial in $1/\epsilon$.

Proof. Fix precision $\epsilon > 0$ and probability $0 < \mu < 1$.

Let X be a solution of ODE (2). From Taylor-Lagrange on function X, we have for $T = \frac{k}{n}$,

$$X(T + \frac{1}{n}) - X(T) = \frac{1}{n}F(X(T)) + \frac{1}{2n^2}F'(\chi)F(\chi),$$

where $\chi \in [T, T + \frac{1}{n}]$.

Let

$$\tilde{P}_n(\frac{k}{n}) = P_n(k),$$

for all k, n.

We can then write for $T = \frac{k}{n}$,

$$X(T+\frac{1}{n})-\tilde{P}_n(T+\frac{1}{n}) = X(T)-\tilde{P}_n(T)+\frac{1}{n}\left(F(X(T)) - F(\tilde{P}_n(T))\right)-\frac{1}{n}\mu_n(T), \tag{3}$$

where $\mu_n(T) = \epsilon_n(k) + \rho_n(k) - \frac{1}{2n}F(\chi_n)F'(\chi_n)$.

Summing (3) from 0 to k, yields

$$X(\frac{k+1}{n})-\tilde{P}_n(\frac{k+1}{n}) = X(0)-\tilde{P}_n(0)+\sum_{i=0}^{k}\frac{1}{n}\left(F(X(\frac{i}{n})) - F(\tilde{P}_n(\frac{i}{n}))\right)-\sum_{i=0}^{k}\frac{1}{n}\mu_n(\frac{i}{n})$$

Since F is $\mathcal{C}^1$ over compact K, it is Λ-Lipschtiz for some Λ.

Using the fact that $X(0) - \tilde{P}_n(0) = 0$, and the fact that F is Λ-Lipschtiz, this gives

$$\left|X(\frac{k+1}{n}) - \tilde{P}_n(\frac{k+1}{n})\right| \leq \sum_{i=0}^{k}\frac{\Lambda}{n}\left|X(\frac{i}{n}) - \tilde{P}_n(\frac{i}{n})\right| + \left|\sum_{i=0}^{k}\frac{1}{n}\mu_n(\frac{i}{n})\right|$$

Introducing

$$\theta_k = \sum_{i=0}^{k}\left|X(\frac{i}{n}) - \tilde{P}_n(\frac{i}{n})\right|,$$

this can be stated as

$$\theta_{k+1} - \theta_k \leq \frac{\Lambda}{n}\theta_k + \left|\sum_{i=0}^{k} \frac{1}{n}\mu_n(\frac{i}{n})\right|$$

Recall

Lemma 2 (Gronwall's Lemma: e.g. [6, page 213]). *Suppose that for some sequences $h_k, \theta_k \geq 0$ and $\epsilon_k \in \mathbb{R}$ we have $\theta_{k+1} \leq (1 + \Lambda h_k)\theta_k + |\epsilon_k|$. Then*

$$\theta_k \leq e^{\Lambda(t_k - t_0)}\theta_0 + \sum_{0 \leq i \leq k-1} e^{\Lambda(t_k - t_{i+1})}|\epsilon_i|,$$

where $t_{k+1} = t_k + h_k$, for all k.

This gives here for $h_k = \frac{1}{n}$, $\epsilon_k = \left|\sum_{i=0}^{k} \frac{1}{n}\mu_n(\frac{i}{n})\right|$

$$\theta_k \leq \sum_{0 \leq i \leq k-1} e^{\frac{\Lambda}{n}(k-i-1)} \left|\sum_{j=0}^{i} \frac{1}{n}\mu_n(\frac{j}{n})\right|,$$

and hence

$$\sup_{0 \leq i \leq k} \left|X(\frac{i}{n}) - \tilde{P}_n(\frac{i}{n})\right| \leq \sum_{0 \leq i \leq k-1} e^{\frac{\Lambda}{n}(k-i-1)} \left|\sum_{j=0}^{i} \frac{1}{n}\mu_n(\frac{j}{n})\right|. \tag{4}$$

This implies

$$\sup_{0 \leq i \leq k} \left|X(\frac{i}{n}) - \tilde{P}_n(\frac{i}{n})\right| \leq \nu(k,n) \sup_{0 \leq i \leq k} \left|\sum_{j=0}^{i} \frac{1}{n}\mu_n(\frac{j}{n})\right|. \tag{5}$$

where

$$\nu(k,n) = \sum_{0 \leq i \leq k-1} e^{\frac{\Lambda i}{n}} = \frac{e^{\frac{\Lambda k}{n}} - 1}{e^{\frac{\Lambda}{n}} - 1} \leq e^{\frac{\Lambda k}{n}} \frac{1 - e^{-\frac{\Lambda k}{n}}}{1 - e^{-\frac{\Lambda}{n}}}$$

which is, doing an asymptotic development, in $\mathcal{O}(e^{\frac{\Lambda k}{n}})$, when n and $T = k/n$ are big enough, say when $n \geq n_1$ and $T \geq T_1$.

Decomposing $\mu_n(k/n) = \epsilon_n(k) + \rho_n(k) - \frac{1}{2n}F'(\chi_n)F(\chi_n)$, we obtain

$$\left|\sum_{j=0}^{i} \frac{1}{n}\mu_n(\frac{j}{n})\right| \leq \left|\sum_{j=0}^{i} \frac{1}{n}\epsilon_n(j)\right| + \left|\sum_{j=0}^{i} \frac{1}{n}\rho_n(j)\right| + \left|\sum_{j=0}^{i} \frac{1}{2n^2}F'(\chi_n)F(\chi_n)\right|$$

As $\epsilon_n(k)$ was assumed to take values in $[-\frac{d}{n}, \frac{d}{n}]$, the first term can be then bounded as follows

$$\left|\sum_{j=0}^{i} \frac{1}{n}\epsilon_n(j)\right| \leq \frac{d(i+1)}{n^2}.$$

The third term can be bounded as follows

$$\left|\sum_{j=0}^{i} \frac{1}{2n^2} F'(\chi_n) F(\chi_n)\right| \leq \frac{M_1 M_2 (i+1)}{2n^2},$$

given that F is bounded on K, and that $F' = X$ is also bounded on K by respective constants M_1 and M_2.

Equation (5) then allows to write

$$\left|X(\frac{k}{n}) - \tilde{P}_n(\frac{k}{n})\right| \leq \nu(k,n)\left(\frac{d(k+1)}{n^2} + \frac{M_1 M_2 (k+1)}{2n^2}\right) + \frac{\nu(k,n)}{n} \sup_{0\leq i\leq k}\left|\sum_{j=0}^{i} \rho_n(j)\right|,$$

if one prefers

$$\left|X(\frac{k}{n}) - \tilde{P}_n(\frac{k}{n})\right| \leq \mathcal{O}(e^{\Lambda T} T \frac{1}{n}) + \mathcal{O}(e^{\Lambda T} \frac{1}{n}) \sup_{0\leq i\leq k}\left|\sum_{j=0}^{i} \rho_n(j)\right|$$

where $T = \frac{k}{n}$, when $n \geq n_1$ and $T \geq T_1$.

Recall that a sequence of random variables $Z_0, Z_1, \cdots$ is said to be martingale with respect to sequence $X_0, X_1, \cdots$ if, for all $n \geq 0$, we have (i) Z_n is a function from $X_0, X_1, \cdots, X_n$ (ii) $E[|Z_n|] < \infty$ (iii) $E[Z_{n+1}|X_0, \cdots, X_n] = Z_n$. A function is martingale if it is a martingale with respect to itself.

Proposition 1 (Azuma-Hoeffding's Inequality: see e.g. [7]). *Let $Z_1, Z_2, \cdots, Z_n$ a martingale such that*

$$|Z_k - Z_{k-1}| \leq c_k.$$

Then for all $t \geq 0$ and all $\lambda > 0$,

$$Pr(|Z_t - Z_0| \geq \lambda) \leq 2e^{-\lambda^2/(2\sum_{k=1}^{t} c_k^2)}.$$

Now consider $Z_0 = 0$,

$$Z_k = \sum_{j=0}^{k-1} \rho_n(j),$$

for $k > 0$.

We have by hypothesis

$$|Z_k - Z_{k-1}| \leq c.$$

So, for all $\lambda > 0$,

$$Pr(|Z_k| \geq \lambda) \leq 2e^{-\lambda^2/(2kc^2)}$$

Using some union bounds,

$$Pr(\sup_{0\leq i\leq k} |\sum_{j=0}^{i} \rho_n(j)| > \lambda) \leq P(\bigcup_{0\leq i\leq k} |\sum_{j=0}^{i} \rho_n(j)| > \lambda) \leq \sum_{i=0}^{k} P(|\sum_{j=0}^{i} \rho_n(j)| > \lambda),$$

which is less than

$$\sum_{i=0}^{k} 2e^{-\lambda^2/(2(i+1)c^2)} \leq 2(k+1)e^{-\lambda^2/(2c^2)}.$$

Fix κ so that $1 - 2(k+1)e^{-\kappa^2 T^2} \geq \mu$ whenever $n \geq n_1$ and $k \geq k_1 = T_1 n_1$. Take then $\lambda = \kappa T c\sqrt{2}$. With probability more than μ

$$\sup_{0 \leq i \leq k} |\sum_{j=0}^{i} \rho_n(j)| \leq \kappa c\sqrt{2}T,$$

hence

$$\left| X(\frac{k}{n}) - \tilde{P}_n(\frac{k}{n}) \right| \leq \mathcal{O}(e^{\Lambda T} T \frac{1}{n}) + \mathcal{O}(e^{\Lambda T} T \frac{1}{n}) = \mathcal{O}(e^{\Lambda T} T \frac{1}{n})$$

Take any $T \geq \max(T(\frac{\epsilon}{2}), T_1)$ so that

$$||X(T) - x^*|| \leq \frac{\epsilon}{2}.$$

Then take any $n \geq n_1$ where n is big enough so that $\mathcal{O}(e^{\Lambda T} T \frac{1}{n}) \leq \frac{\epsilon}{2}$: as n_1 is some constant (not depending on ϵ) n can be taken in $\mathcal{O}(\frac{1}{\epsilon} e^{\Lambda T} T) = \mathcal{O}(\frac{1}{\epsilon}(\frac{1}{\epsilon})^{\mathcal{O}(1)} \ln \frac{1}{\epsilon})$, that is to say, polynomial in $\frac{1}{\epsilon}$.

Consider then $k = \max(k_1, Tn)$. We have

$$||P_n(k) - x^*|| = ||\tilde{P}_n(\frac{k}{n}) - x^*|| \leq ||\tilde{P}_n(\frac{k}{n}) - X(\frac{k}{n})|| + ||X(\frac{k}{n}) - x^*|| \leq \epsilon :$$

as k_1 is some constant (not depending on ϵ), and as n is polynomial in $\frac{1}{\epsilon}$, k can also be taken as polynomial in $\frac{1}{\epsilon}$.

References

[1] D. Angluin, J. Aspnes, Z. Diamadi, M. J. Fischer, R. Peralta, Computation in networks of passively mobile finite-state sensors, in: Twenty-Third ACM Symposium on Principles of Distributed Computing, ACM Press, 2004, pp. 290–299.

[2] M. Presburger, Uber die Vollstandig-keit eines gewissen systems der Arithmetik ganzer Zahlen, in welchemdie Addition als einzige Operation hervortritt, Comptes-rendus du I Congres des Mathematicians des Pays Slaves (1929) 92–101.

[3] D. Angluin, J. Aspnes, D. Eisenstat, Stably computable predicates are semilinear, in: PODC '06: Proceedings of the twenty-fifth annual ACM symposium on Principles of distributed computing, ACM Press, New York, NY, USA, 2006, pp. 292–299. doi:http://doi.acm.org/10.1145/1146381.1146425.

[4] J. Aspnes, E. Ruppert, An introduction to population protocols, in: Bulletin of the EATCS, Vol. 93, 2007, pp. 106–125.

[5] O. Bournez, P. Chassaing, J. Cohen, L. Gerin, X. Koegler, On the convergence of population protocols when population goes to infinity, Applied Mathematics and Computation.

[6] J.-P. Demailly, Analyse Numérique et Equations Différentielles, Presses Universitaires de Grenoble, 1991.

[7] M. Mitzenmacher, E. Upfal, Probability and Computing. Randomized Algorithms and Probabilistic Analysis, Cambridge University Press, 2005.