Counting with Population Protocols

Yves Mocquard, Emmanuelle Anceaume, James Aspnes, Yann Busnel, Bruno Sericola

Counting with Population Protocols

Yves Mocquard*, Emmanuelle Anceaume†,James Aspnes‡,Yann Busnel§ and Bruno Sericola¶
*IRISA / Université de Rennes 1 (France), yves.micquard@irisa.fr
†CNRS / Irisa (France), emmanuelle.anceaume@irisa.fr
‡Yale University (USA), james.aspnes@gmail.com
§Crest / Ensai (France), yann.busnel@ensai.fr
¶INRIA Rennes Atlantique (France), bruno.sericola@inria.fr

***Abstract*—The population protocol model provides theoretical foundations for analyzing the properties emerging from simple and pairwise interactions among a very large number n of anonymous agents. The problem tackled in this paper is the following one: is there an efficient population protocol that exactly counts the difference κ between the number of agents that initially and independently set their state to A and the one that initially set it to B, assuming that each agent only uses a finite set of states ? We propose a solution which guarantees with any high probability that after $O(\log n)$ interactions any agent outputs the exact value of κ. Simulation results illustrate our theoretical analysis.**

***Keywords*-Population protocol; Majority algorithm; Counting problem; Performance evaluation.**

I. Introduction

The population protocol model, introduced by Angluin et al. [2], provides theoretical foundations for analyzing global properties emerging from pairwise interactions among a large number of anonymous agents [5]. In the population protocol model, agents are modeled as identical and deterministic finite state machines, i.e each agent can be in a finite number of states while waiting to execute a transition. When two agents interact, they communicate their local state, and can move from one state to another according to a joint transition function. The patterns of interaction are unpredictable, however they must be fair, in the sense that any interaction that should possibly appear cannot be avoided forever. The ultimate goal of population protocols is for all the agents to converge to a correct value independently of the interaction pattern. Examples of systems whose behavior can be modeled by population protocols range from molecule interactions of a chemical process to sensor networks in which agents, which are small devices embedded on animals, interact each time two animals are in the same radio range.

A lot of work has been devoted to determine the tasks that can be solved in the population protocol model, as well as to study their complexities in terms of memory and convergence time [1], [3], [6], [7], [8]. One of the most studied tasks is the *majority task*. Briefly, solving the majority task amounts for all the agents to eventually output a single value that depends on the initial majority state of each agent of the system, assuming that all the agents set their initial state to one of the two possible initial states, say A or B. Section IV provides an overview of the results recently obtained for the majority task.

In this paper, we focus on an quite important related question. Namely, is there a population protocol that exactly counts the difference κ between the number of agents that initially set their state to A and the one that initially set it to B, and can it be solved in an efficient way, that is with the guarantee that each agent should converge to the exact value of κ after having triggered a sub-linear number of interactions in the size of the system.

We answer this question by the affirmative by presenting a $O(n^{3/2})$-state population protocol that allows each agent to converge to the exact solution by interacting no more than $O(\log n)$ times. The proposed protocol is very simple (as is true for most known population protocols), but is general enough to be used to solve different types of tasks.

Specifically, our algorithm works as follows. Starting from an initial population in which all A agents have a value m, where m is a positive number, and all B agents have a value $-m$, each pair of agents that meet, adopt the average of their values (or as close as they can get when values are restricted to integers, as will be clarified in Section VI). This method preserves the sum of the initial values, and we show that the values of all agents quickly converge to the average initial value. This method is used in Section V and VI to obtain the difference κ between the number of A and B agents.

This mechanism is similar to the averaging phase of the average-and-conquer algorithm of Alistarh et al. [1] for computing the majority value, and indeed our algorithm was in part inspired by this previous work. What distinguishes our algorithm from that of Alistarh et al. [1] is that we generalize the problem to counting in addition to computing majority, demonstrate that the additional mechanisms in their algorithm for propagating the majority value can be omitted without compromising correctness, and give a simpler proof of convergence of the averaging mechanism based on tracking the Euclidean distance between the vector of all agents' values and the uniform vector.

The remainder of this paper is organized as follows. Section II presents the population protocol model. Section III specifies the problem addressed in this work. Section IV provides an overview of the most recent average-based population protocols. A preliminary protocol to compute the difference

This work was partially funded by the French ANR project SocioPlug (ANR-13-INFR-0003), and by the DeSceNt project granted by the Labex CominLabs excellence laboratory (ANR-10-LABX-07-01).

κ together with its analysis are presented in Section V. Note that this protocol does not exactly meet the population protocol model, in that, it does not assume that agents have a finite number of states. However, the motivation behind this presentation is that the proposed argumentation is very similar to the one used in Section VI while enriched to cope with the finite number of states. We have simulated our protocol to illustrate our theoretical analysis. Section VII presents a summary of these simulation results. Finally, Section VIII concludes.

II. Population protocols model

In this section, we present the population protocol model, introduced by Angluin et al. [2]. This model describes the behavior of a collection of agents that interact pairwise. The following definition is from Angluin et al [4]. A population protocol is characterized by a 6-tuple $(Q, \Sigma, Y, \iota, \omega, f)$, over a complete interaction graph linking the set of n agents, where Q is a finite set of states, Σ is a finite set of input symbols, Y is a finite set of output symbols, $\iota : \Sigma \to Q$ is the input function that determines the initial state of an agent, $\omega : Q \to Y$ is the output function that determines the output symbol of an agent, and $f : Q \times Q \to Q \times Q$ is the transition function that describes how two agents interact and update theirs states. Initially all the agents start with a initial symbol from Σ, and upon interactions with agents update their state according to the transition function f. Interactions between agents are orchestrated by a uniform random scheduler: at each discrete time, any two agents are randomly chosen to interact with an uniform law. Note that a uniform random scheduler is fair: any possible interaction cannot be avoided forever.

The notion of time in population protocols refers to as the successive steps at which interactions occur, while the parallel time refers to as the successive number of steps each agent executes [5].

Agents do not maintain nor use identifiers (agents are anonymous and cannot determine whether any two interactions have occurred with the same agents or not). However, for ease of presentation the agents are numbered $1, 2, \ldots, n$.

We will represent the trajectory of a population protocol with random interactions as a stochastic process. Let $C = \{C_t,\ t \geq 0\}$ be a stochastic process with state configuration Q^n. For every $t \geq 0$, the configuration at time t of the stochastic process is denoted by $C_t = (C_t^{(1)}, \ldots, C_t^{(n)})$. At each discrete instant t, two distinct indices i and j are chosen among $1, \ldots, n$ with probability $p_{i,j}(t)$. We denote by X_t the random variable representing this choice, that is

$$\mathbb{P}\{X_t = (i, j)\} = p_{i,j}(t).$$

We assume that the random variables X_t and C_t are independent.

We will use in the sequel the Euclidean norm denoted by $\|.\|$ and the infinite norm denoted by $\|.\|_\infty$ defined for all $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$ by

$$\|x\| = \left(\sum_{i=1}^{n} x_i^2\right)^{1/2} \text{ and } \|x\|_\infty = \max_{i=1,\ldots,n} |x_i|$$

It is well-known that these norms satisfy

$$\|x\|_\infty \leq \|x\| \leq \sqrt{n}\|x\|_\infty.$$

III. The Counting Problem

The problem addressed in this work is the following one. We consider a set of n agents, interconnected by a complete graph, that start their execution in one of two distinguished states of $\Sigma = \{A, B\}$. Let n_A be the number of agents whose initial state is A and n_B be the number of agents that start in state B. Let $\kappa = n_A - n_B$, be the quantity referred to as the conserved advantage in the following of the paper. The output set Y is the set of all possible values of κ, that is all integers between $-n$ and n.

A population protocol solves the counting problem within τ steps with probability at least $1 - \delta$, for any $\delta \in (0, 1)$, if for any configuration C_t reachable by the protocol after $t \geq \tau$ steps, it holds that with probability at least $1-\delta$, $\omega(C_t^{(i)}) = \kappa$, for any agent i. As will be shown in the following, κ does not depend on time t, however agents are locally able to compute κ only after a logarithmic number of interactions.

IV. Related Work

The population protocol model was formalized by Angluin et al. [2]. Since then, there has been a lot of work on population protocols, and among them the closest to our work compute majority, which is related to the average problem. For this problem, all the agents start in one of two distinguished states and they eventually converge to 1 if $\kappa > 0$, and to 0 if $\kappa < 0$. In [6], [7], the authors propose a four-state protocol that solves the majority problem with an expected convergence parallel time logarithmic in n. However, the expected convergence time is infinite when κ approaches 0. The authors in [3], [8] propose a three-state protocol that converges with some probability δ, and whose parallel time is logarithmic in n if κ is large enough, i.e $\kappa = O(\sqrt{n} \log n)$. Finally, the closest work to ours is the one of Alistarh et al. [1]. The authors propose a population protocol based on an average-and-conquer method to exactly solve the majority problem. Their algorithm uses two types of interactions, an averaging one and a neutralization interaction. The first type of interactions is close to the one used in our protocol while the second one is used to prevent agents from influencing the final decision. This additional mechanism for propagating the majority value to all the agents makes their proof of convergence intricate. Finally, the authors show that the number of states and convergence parallel time of their algorithms are lower bounded by a $\log n$ factor. In contrast we show that the convergence parallel time of our algorithm is upper bounded by a $\log n$ factor.

V. Counting with a countable state space

We consider in this section the case where the state space Q is not finite but nevertheless countable. It does not correspond strictly to the definition of a population protocol given above but the convergence analysis is stronger and will be extended in the next section to a finite state space. The parameters Q, Σ, Y, ι and ω are application dependent and will be defined at the end of this section where the conserved advantage κ will be computed. The transition function f we use is given by

$$f(a,b) = \left(\frac{a+b}{2}, \frac{a+b}{2}\right).$$

This means that the state space Q is the set of dyadic numbers, the rational numbers with denominators that are powers of 2. Once the pair (i,j) is chosen at time t, the process reaches state C_{t+1}, at time $t+1$, given by

$$C_{t+1}^{(i)} = C_{t+1}^{(j)} = \frac{C_t^{(i)} + C_t^{(j)}}{2} \text{ and } C_{t+1}^{(m)} = C_t^{(m)} \text{ if } m \neq i,j. \tag{1}$$

Lemma *1:* For every $t \geq 0$, we have

$$\sum_{i=1}^{n} C_t^{(i)} = \sum_{i=1}^{n} C_0^{(i)}.$$

Proof: The proof is immediate since the transformation from C_t to C_{t+1} described in Relation (1) does not change the sum of the entries of C_{t+1}. Indeed, from Relation (1), we have $C_{t+1}^{(i)} + C_{t+1}^{(j)} = C_t^{(i)} + C_t^{(j)}$ and the other entries do not change their values. ∎

In the following we denote by ℓ the mean value of the sum of the entries of C_t and by L the row vector of $\mathbb{R}^n$ with all its entries equal to ℓ, that is

$$\ell = \frac{1}{n}\sum_{i=1}^{n} C_t^{(i)} \text{ and } L = (\ell, \ldots, \ell).$$

Theorem *2:* Assuming a uniform choice of the pair (i,j), that is if, for $i \neq j$,

$$p_{i,j}(t) = \frac{1}{n(n-1)},$$

then we have

$$\mathbb{E}\left(\|C_t - L\|^2\right) = \left(1 - \frac{1}{n-1}\right)^t \mathbb{E}\left(\|C_0 - L\|^2\right). \tag{2}$$

Moreover C_t converges almost surely to L, i.e.

$$\mathbb{P}\{\lim_{t\longrightarrow\infty} C_t = L\} = 1.$$

Proof: Let $x \in \mathbb{R}^n$. For every $i,j = 1,\ldots,n$ with $i \neq j$, the vector y defined by

$$y_i = y_j = \frac{x_i + x_j}{2} \text{ and } y_m = x_m \text{ for } m \neq i,j$$

satisfies

$$\|y-L\|^2 = \|x-L\|^2 - (x_i-\ell)^2 - (x_j-\ell)^2 + 2\left(\frac{x_i+x_j}{2} - \ell\right)^2,$$

which gives

$$\|y-L\|^2 = \|x-L\|^2 - \frac{(x_i-x_j)^2}{2}.$$

Applying this result to the random vectors C_{t+1} and C_t gives, for every $t \geq 0$,

$$\|C_{t+1} - L\|^2 = \|C_t - L\|^2 - \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\left(C_t^{(i)} - C_t^{(j)}\right)^2 1_{\{X_t=(i,j)\}}. \tag{3}$$

By taking the expectations and using the fact that X_t and C_t are independent, we get

$$\mathbb{E}\left(\|C_{t+1} - L\|^2\right) = \mathbb{E}\left(\|C_t - L\|^2\right) - \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\mathbb{E}\left(\left(C_t^{(i)} - C_t^{(j)}\right)^2\right) p_{i,j}(t).$$

Since

$$p_{i,j}(t) = \frac{1}{n(n-1)},$$

we obtain

$$\mathbb{E}\left(\|C_{t+1} - L\|^2\right) = \mathbb{E}\left(\|C_t - L\|^2\right) - \frac{1}{2n(n-1)}\sum_{i=1}^{n}\sum_{j=1}^{n}\mathbb{E}\left(\left(C_t^{(i)} - C_t^{(j)}\right)^2\right).$$

Moreover, we have

$$\begin{aligned}
\sum_{i=1}^{n}\sum_{j=1}^{n}\left(C_t^{(i)} - C_t^{(j)}\right)^2 &= \sum_{i=1}^{n}\sum_{j=1}^{n}\left(\left(C_t^{(i)} - \ell\right) - \left(C_t^{(j)} - \ell\right)\right)^2 \\
&= \sum_{i=1}^{n}\sum_{j=1}^{n}\Big[\left(C_t^{(i)} - \ell\right)^2 + \left(C_t^{(j)} - \ell\right)^2 \\
&\qquad -2\left(C_t^{(i)} - \ell\right)\left(C_t^{(j)} - \ell\right)\Big] \\
&= 2n\|C_t - L\|^2 - 2\sum_{i=1}^{n}\left(C_t^{(i)} - \ell\right)\sum_{j=1}^{n}\left(C_t^{(j)} - \ell\right) \\
&= 2n\|C_t - L\|^2 - 2\sum_{i=1}^{n}\left(C_t^{(i)} - \ell\right)(n\ell - n\ell) \\
&= 2n\|C_t - L\|^2.
\end{aligned}$$

This leads to

$$\mathbb{E}\left(\|C_{t+1} - L\|^2\right) = \left(1 - \frac{1}{n-1}\right)\mathbb{E}\left(\|C_t - L\|^2\right),$$

and thus

$$\mathbb{E}\left(\|C_t - L\|^2\right) = \left(1 - \frac{1}{n-1}\right)^t \mathbb{E}\left(\|C_0 - L\|^2\right).$$

Using the Markov inequality, we get, for all $\varepsilon > 0$

$$\mathbb{P}\{\|C_t - L\|^2 \geq \varepsilon\} \leq \frac{1}{\varepsilon}\left(1 - \frac{1}{n-1}\right)^t \mathbb{E}\left(\|C_0 - L\|^2\right),$$

which means that C_t converges in probability to L when t tends to infinity. Since, for all $\varepsilon > 0$, we have

$$\sum_{t=0}^{\infty} \mathbb{P}\{\|C_t - L\|^2 \geq \varepsilon\} \leq \frac{\mathbb{E}\left(\|C_0 - L\|^2\right)}{\varepsilon} \sum_{t=0}^{\infty} \left(1 - \frac{1}{n-1}\right)^t = \frac{(n-1)\mathbb{E}\left(\|C_0 - L\|^2\right)}{\varepsilon} < \infty,$$

we deduce that C_t converges almost surely to L when t tends to infinity. ∎

The following corollary gives an (ε, δ)-approximation of L by C_t.

Corollary *3:* For all $\varepsilon > 0$ and $\delta \in (0,1)$, if there exists a constant K such that $\mathbb{E}(\|C_0 - L\|_\infty) \leq K$ then, for all $t \geq (n-1)\ln(nK^2/\varepsilon^2\delta)$ we have

$$\mathbb{P}\{\|C_t - L\|_\infty \geq \varepsilon\} \leq \delta.$$

Proof: Let $\tau = (n-1)\ln(nK^2/\varepsilon^2\delta)$. From Relation (3), we see that the random variable $\|C_t - L\|^2$ is a non increasing function of t. So, for all $t \geq \tau$, we have

$$\begin{aligned}\mathbb{P}\{\|C_t - L\|_\infty \geq \varepsilon\} &= \mathbb{P}\{\|C_t - L\|_\infty^2 \geq \varepsilon^2\} \\ &\leq \mathbb{P}\{\|C_t - L\|^2 \geq \varepsilon^2\} \\ &\leq \mathbb{P}\{\|C_\tau - L\|^2 \geq \varepsilon^2\}.\end{aligned}$$

From Theorem 2 and using the Markov inequality, we get

$$\begin{aligned}\mathbb{P}\{\|C_\tau - L\|^2 \geq \varepsilon^2\} &\leq \frac{\mathbb{E}\left(\|C_\tau - L\|^2\right)}{\varepsilon^2} \\ &= \frac{1}{\varepsilon^2}\left(1 - \frac{1}{n-1}\right)^\tau \mathbb{E}\left(\|C_0 - L\|^2\right).\end{aligned}$$

For all $x \in [0,1)$, we have $\ln(1-x) \leq -x$. This leads to

$$\left(1 - \frac{1}{n-1}\right)^\tau \leq e^{-\tau/(n-1)} = \frac{\varepsilon^2\delta}{nK^2}.$$

We then obtain, for $t \geq \tau$,

$$\mathbb{P}\{\|C_t - L\|_\infty \geq \varepsilon\} \leq \frac{\delta}{nK^2}\mathbb{E}\left(\|C_0 - L\|^2\right).$$

Since

$$\|C_0 - L\|^2 \leq n\|C_0 - L\|_\infty^2,$$

we get

$$\mathbb{P}\{\|C_t - L\|_\infty \geq \varepsilon\} \leq \frac{\delta}{K^2}\mathbb{E}\left(\|C_0 - L\|_\infty^2\right) \leq \delta,$$

which completes the proof. ∎

We now apply these results for the computation of the conserved advantage κ. The input set is $\Sigma = \{A, B\}$ and the input function ι is defined by $\iota(A) = m$ and $\iota(B) = -m$, where m is a positive number. This means that, for every $i = 1, \ldots, n$, we have $C_0^{(i)} \in \{-m, m\}$. Recall that n_A (resp. n_B) represents the number of agents whose initial input is A (resp. B), i.e. with initial value m (resp. $-m$) and the conserved advantage κ is equal to $n_A - n_B$. Note that we have $n = n_A + n_B$. Moreover, we have

$$\ell = \frac{1}{n}\sum_{i=1}^{n} C_0^{(i)} = \frac{\kappa m}{n},$$

which shows from Lemma 1 that κ is time independent. The set of output Y is the set of all possible values of κ, i.e. $Y = \{-n, -n+1, \ldots, n-1, n\}$. The set of states Q is the set of all the dyadic numbers which belong to the interval $[-m, m]$. Finally, the output function is, for all $x \in Q$,

$$\omega(x) = \lfloor nx/m + 1/2 \rfloor.$$

Theorem *4:* For all $\delta \in (0,1)$ and $t \geq (n-1)(4\ln 2 + 3\ln n - \ln\delta)$, we have

$$\mathbb{P}\{\omega(C_t^{(i)}) = \kappa, \text{for all } i = 1, \ldots, n\} \geq 1 - \delta.$$

Proof: Since $Q \subset [-m, m]$, we have $\|C_0 - L\|_\infty \leq 2m$. From Corollary 3, we obtain that for all $\varepsilon, \delta > 0$ and $t \geq (n-1)\ln(4nm^2/\varepsilon^2\delta)$, we have $\mathbb{P}\{\|C_t - L\|_\infty \geq \varepsilon\} \leq \delta$ or equivalently

$$\mathbb{P}\{|C_t^{(i)} - \frac{\kappa m}{n}| < \varepsilon, \text{ for all } i = 1, \ldots, n\} \geq 1 - \delta.$$

By taking, $\varepsilon = m/(2n)$, we get, for all $\delta > 0$ and $t \geq (n-1)\ln(16n^3/\delta)$

$$\mathbb{P}\{\kappa < \frac{C_t^{(i)} n}{m} + \frac{1}{2} < \kappa + 1, \text{ for all } i = 1, \ldots, n\} \geq 1 - \delta.$$

This is implies that for all $\delta > 0$ and $t \geq (n-1)(4\ln 2 + 3\ln n - \ln\delta)$, $\mathbb{P}\{\omega(C_t^{(i)}) = \kappa, \text{for all } i = 1, \ldots, n\} \geq 1 - \delta$. ∎

Note that this result is independent from the choice of m, which can be any positive number. The convergence time to get the conserved advantage κ with any high probability is $O(n \log n)$ and thus the parallel convergence time to get κ with any high probability is $O(\log n)$.

VI. Counting with a finite state space

In this section we use the strict definition of population protocols, i.e. we deal with a finite state space Q which is a finite set of integer vectors. Similarly to the previous analysis, the parameters Q, Σ, Y, ι and ω are application dependent and will be defined at the end of the section for the computation of the conserved advantage κ. The transition function f is given by

$$f(a,b) = \begin{cases} \left(\frac{a+b}{2}, \frac{a+b}{2}\right) & \text{if } a+b \text{ is even} \\ \left(\frac{a+b-1}{2}, \frac{a+b+1}{2}\right) & \text{if } a+b \text{ is odd.} \end{cases} \quad (4)$$

Once the couple (i, j) is chosen at time t, the process reaches state C_{t+1}, at time $t+1$, given by

$$\left(C_{t+1}^{(i)}, C_{t+1}^{(j)}\right) = \begin{cases} \left(\frac{C_t^{(i)}+C_t^{(j)}}{2}, \frac{C_t^{(i)}+C_t^{(j)}}{2}\right) & \text{if } C_t^{(i)} + C_t^{(j)} \text{ is even} \\ \left(\frac{C_t^{(i)}+C_t^{(j)}-1}{2}, \frac{C_t^{(i)}+C_t^{(j)}+1}{2}\right) & \text{if } C_t^{(i)} + C_t^{(j)} \text{ is odd} \end{cases}$$

$$\text{and } C_{t+1}^{(m)} = C_t^{(m)} \text{ for } m \neq i, j. \quad (5)$$

Lemma *5:* For every $t \geq 0$, we have

$$\sum_{i=1}^{n} C_t^{(i)} = \sum_{i=1}^{n} C_0^{(i)}.$$

Proof: The proof is immediate since the transformation from C_t to C_{t+1} described in Relation (5) does not change the sum of the entries of C_{t+1}. Indeed, from Relation (5), we have $C_{t+1}^{(i)} + C_{t+1}^{(j)} = C_t^{(i)} + C_t^{(j)}$ and the other entries do not change their values. ∎

As we did in the countable case, we denote by ℓ the mean value of the sum of the entries of C_t and by L the row vector of $\mathbb{R}^n$ with all its entries equal to ℓ, that is

$$\ell = \frac{1}{n}\sum_{i=1}^{n} C_t^{(i)} \text{ and } L = (\ell, \ldots, \ell).$$

Theorem *6:* Assuming a uniform choice of the pair (i, j), that is if, for $i \neq j$,

$$p_{i,j}(t) = \frac{1}{n(n-1)},$$

then we have

$$\mathbb{E}\left(\|C_t - L\|^2\right) \leq \left(1 - \frac{1}{n-1}\right)^t \mathbb{E}\left(\|C_0 - L\|^2\right) + \frac{n}{4}. \quad (6)$$

Proof: Let $x \in \mathbb{Z}^n$. For every $i, j = 1, \ldots, n$ with $i \neq j$, the vector y defined by

$$(y_i,\ y_j) = \begin{cases} \left(\frac{x_i+x_j}{2}, \frac{x_i+x_j}{2}\right) & \text{if } x_i + x_j \text{ is even} \\ \left(\frac{x_i+x_j-1}{2}, \frac{x_i+x_j+1}{2}\right) & \text{if } x_i + x_j \text{ is odd} \end{cases}$$

and $y_m = x_m$ for $m \neq i, j$ satisfies

$$\|y - L\|^2 = \|x - L\|^2 - (x_i - \ell)^2 - (x_j - \ell)^2 + \begin{cases} \left(\frac{x_i+x_j}{2} - \ell\right)^2 & \text{if } x_i + x_j \text{ is even} \\ \left(\frac{x_i+x_j+1}{2} - \ell\right)^2 + \left(\frac{x_i+x_j-1}{2} - \ell\right)^2 & \text{if } x_i + x_j \text{ is odd} \end{cases}$$

which gives

$$\|y-L\|^2 = \begin{cases} \|x - L\|^2 - \frac{(x_i-x_j)^2}{2} & \text{if } x_i + x_j \text{ is even} \\ \|x - L\|^2 - \left[\frac{(x_i-x_j)^2}{2} - \frac{1}{2}\right] & \text{if } x_i + x_j \text{ is odd} \end{cases}$$

We introduce the indicator function $1_{\{x \text{ odd}\}}$ defined by $1_{\{x \text{ odd}\}} = 1$ if x is odd and 0 if x is even. We can write

$$\|y - L\|^2 = \|x - L\|^2 - \left[\frac{(x_i - x_j)^2}{2} - \frac{1_{\{x_i+x_j \text{ odd}\}}}{2}\right]$$

Applying this result to the random vectors C_{t+1} and C_t gives, for every $t \geq 0$,

$$\|C_{t+1} - L\|^2 = \|C_t - L\|^2 \quad (7)$$
$$- \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\left[\left(C_t^{(i)} - C_t^{(j)}\right)^2 - 1_{\{C_t^{(i)}+C_t^{(j)} \text{ odd}\}}\right] 1_{\{X_t=(i,j)\}}.$$

By taking the expectations and using the fact that X_t and C_t are independent, we get

$$\mathbb{E}\left(\|C_{t+1} - L\|^2\right) = \mathbb{E}\left(\|C_t - L\|^2\right) - \frac{1}{2}\mathbb{E}\left(\sum_{i=1}^{n}\sum_{j=1}^{n}\left[\left(C_t^{(i)} - C_t^{(j)}\right)^2 - 1_{\{C_t^{(i)}+C_t^{(j)} \text{ odd}\}}\right]\right) p_{i,j}(t).$$

Since

$$p_{i,j}(t) = \frac{1}{n(n-1)},$$

we obtain

$$\mathbb{E}\left(\|C_{t+1} - L\|^2\right) = \mathbb{E}\left(\|C_t - L\|^2\right) \quad (8)$$
$$- \frac{1}{2n(n-1)}\mathbb{E}\left(\sum_{i=1}^{n}\sum_{j=1}^{n}\left[\left(C_t^{(i)} - C_t^{(j)}\right)^2 - 1_{\{C_t^{(i)}+C_t^{(j)} \text{ odd}\}}\right]\right).$$

As in the countable case, we have

$$\sum_{i=1}^{n}\sum_{j=1}^{n}\left(C_t^{(i)} - C_t^{(j)}\right)^2 = 2n\|C_t - L\|^2$$

and if q_t be the number of odd entries of C_t, we have

$$\sum_{i=1}^{n}\sum_{j=1}^{n} 1_{\{C_t^{(i)}+C_t^{(j)} \text{ odd}\}} = 2q_t(n - q_t).$$

It follows that

$$\mathbb{E}\left(\|C_{t+1} - L\|^2\right) = \left(1 - \frac{1}{n-1}\right)\mathbb{E}\left(\|C_t - L\|^2\right) + \frac{\mathbb{E}(q_t(n-q_t))}{n(n-1)}. \quad (9)$$

The function g defined, for $x \in [0, n]$, by $g(x) = x(n-x)$ has its maximum at point $x = n/2$, so we have $0 \leq g(x) \leq n^2/4$. This gives

$$\mathbb{E}\left(\|C_{t+1} - L\|^2\right) \leq \left(1 - \frac{1}{n-1}\right)\mathbb{E}\left(\|C_t - L\|^2\right) + \frac{n}{4(n-1)},$$

that is

$$\mathbb{E}\left(\|C_t - L\|^2\right) \leq \left(1 - \frac{1}{n-1}\right)^t \mathbb{E}\left(\|C_0 - L\|^2\right) + \frac{n}{4(n-1)}\sum_{i=0}^{t-1}\left(1 - \frac{1}{n-1}\right)^i.$$

Since

$$\sum_{i=0}^{t-1}\left(1 - \frac{1}{n-1}\right)^i \leq \sum_{i=0}^{\infty}\left(1 - \frac{1}{n-1}\right)^i = n - 1,$$

we get

$$\mathbb{E}\left(\|C_t - L\|^2\right) \leq \left(1 - \frac{1}{n-1}\right)^t \mathbb{E}\left(\|C_0 - L\|^2\right) + \frac{n}{4},$$

which completes the proof. ∎

Lemma *7:* The sequence $\left(\|C_t - L\|^2\right)_t$ is non increasing.

Proof: It is immediate, from the Equality 7. Indeed, if $C_t^{(i)} + C_t^{(j)}$ is odd then $C_t^{(i)} \neq C_t^{(j)}$ which means, in any case, that $\left(C_t^{(i)} - C_t^{(j)}\right)^2 - 1_{\{C_t^{(i)}+C_t^{(j)} \text{ odd}\}} \geq 0$. ∎

The following corollary gives a δ-approximation for the deviation between C_t and L.

Corollary *8:* For all $\delta \in (0, 1)$, if there exists a constant K such that $\mathbb{E}(\|C_0 - L\|_\infty) \leq K$ then, for all $t \geq (n-1)\ln(4K^2)$ we have

$$\mathbb{P}\{\|C_t - L\|_\infty \geq \sqrt{\frac{n}{2\delta}}\} \leq \delta.$$

Proof: Let $\tau = (n-1)\ln(4K^2)$ and $t \geq \tau$. From Lemma 7, from Theorem 6 and using the Markov inequality, we obtain

$$\mathbb{P}\{\|C_t - L\|_\infty^2 \geq \frac{n}{2\delta}\} \leq \mathbb{P}\{\|C_t - L\|^2 \geq \frac{n}{2\delta}\}$$

$$\leq \mathbb{P}\{\|C_\tau - L\|^2 \geq \frac{n}{2\delta}\} \leq \frac{2\delta}{n}\mathbb{E}\left(\|C_\tau - L\|^2\right)$$

$$\leq \frac{2\delta}{n}\left(1 - \frac{1}{n-1}\right)^\tau \mathbb{E}\left(\|C_0 - L\|^2\right) + \frac{\delta}{2}.$$

For all $x \in [0,1[$, we have $\ln(1-x) \leq -x$. This leads to

$$\left(1 - \frac{1}{n-1}\right)^\tau \leq e^{-\tau/(n-1)} = \frac{1}{4K^2}.$$

We then obtain, for $t \geq \tau$,

$$\mathbb{P}\{\|C_t - L\|_\infty^2 \geq \frac{n}{2\delta}\} \leq \frac{\delta}{2nK^2}\mathbb{E}\left(\|C_0 - L\|^2\right) + \frac{\delta}{2}.$$

Since

$$\|C_0 - L\|^2 \leq n\|C_0 - L\|_\infty^2 \leq nK^2,$$

we obtain

$$\mathbb{P}\{\|C_t - L\|_\infty^2 \geq \frac{n}{2\delta}\} \leq \delta,$$

that is

$$\mathbb{P}\{\|C_t - L\|_\infty \geq \sqrt{\frac{n}{2\delta}}\} \leq \delta,$$

which completes the proof. ∎

We now apply these results for the computation of the conserved advantage κ. Similarly to the countable case, the input set is $\Sigma = \{A, B\}$, and the input function ι is defined by $\iota(A) = m$ and $\iota(B) = -m$, however m is a positive integer. This means that, for every $i = 1, \ldots, n$, we have $C_0^{(i)} \in \{-m, m\}$. We have

$$\ell = \frac{1}{n}\sum_{i=1}^{n} C_0^{(i)} = \frac{\kappa m}{n},$$

which shows from Lemma 5 that κ is time independent. The set of states Q is now the set $\{-m, -m+1, \ldots, m-1, m\}$. The output function is, for all $x \in Q$,

$$\omega(x) = \lfloor nx/m + 1/2 \rfloor.$$

Finally, the output set Y is the set of all possible values of κ, i.e. $Y = \{-n, -n+1, \ldots, n-1, n\}$.

Theorem *9:* For all $\delta \in (0,1)$, $m = \left\lceil \sqrt{2}n^{3/2}/\sqrt{\delta} \right\rceil$ and for all $t \geq (n-1)\left(5\ln 2 + 3\ln n - \ln\delta + \frac{2}{m-1}\right)$, we have

$$\mathbb{P}\{\omega(C_t^{(i)}) = \kappa, \text{ for all } i = 1, \ldots, n\} \geq 1 - \delta.$$

Proof: Since $Q \subset [-m, m]$, we have $\|C_0 - L\|_\infty \leq 2m$. From Corollary 8, we obtain that for all $\delta \in (0,1)$ and $t \geq (n-1)\ln(16m^2)$, we have

$$\mathbb{P}\{\|C_t - L\|_\infty \geq \sqrt{\frac{n}{2\delta}}\} \leq \delta$$

or equivalently

$$\mathbb{P}\{|C_t^{(i)} - \frac{\kappa m}{n}| < \sqrt{\frac{n}{2\delta}}, \text{ for all } i\} \geq 1 - \delta.$$

Introducing the notation $y = n^{3/2}/\sqrt{2\delta}$, we obtain $m = \lceil 2y \rceil$ and thus $2y \leq m < 2y + 1$, which implies that

$$\begin{aligned}
&(n-1)\ln(16m^2)\\
&\leq (n-1)\left(4\ln 2 + 2\ln\left(\sqrt{2}n^{3/2} + \sqrt{\delta}\right) - \ln\delta\right)\\
&= (n-1)\left(4\ln 2 + 2\ln\left(\sqrt{2}n^{3/2}\left(1 + \frac{\sqrt{\delta}}{\sqrt{2n^{3/2}}}\right)\right) - \ln\delta\right)\\
&\leq (n-1)\left(5\ln 2 + 3\ln n - \ln\delta + \frac{\sqrt{2\delta}}{n^{3/2}}\right)\\
&\leq (n-1)\left(5\ln 2 + 3\ln n - \ln\delta + \frac{2}{m-1}\right).
\end{aligned}$$

Therefore, for $t \geq (n-1)\left(5\ln 2 + 3\ln n - \ln\delta + \frac{2}{m-1}\right)$ we obtain

$$\mathbb{P}\{\kappa + \frac{1}{2} - \frac{y}{m} < \frac{C_t^{(i)} n}{m} + \frac{1}{2} < \kappa + \frac{1}{2} + \frac{y}{m}, \text{ for all } i\} \geq 1 - \delta,$$

Since $y/m \leq 1/2$, this implies

$$\mathbb{P}\{\kappa < \frac{C_t^{(i)} n}{m} + \frac{1}{2} < \kappa + 1, \text{ for all } i\} \geq 1 - \delta.$$

Now this implies that $\mathbb{P}\{\omega(C_t^{(i)}) = \kappa, \text{for all } i\} \geq 1 - \delta$, which completes the proof. ∎

Note that the convergence time to get κ with any high probability is $O(n \log n)$ and thus the parallel convergence time to get κ with any high probability is $O(\log n)$.

VII. Experimental evaluation

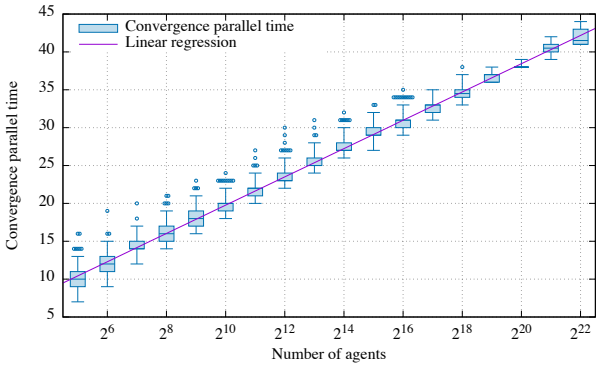
This section describes the main results obtained from the simulation of our protocol.

A. Settings of the simulations

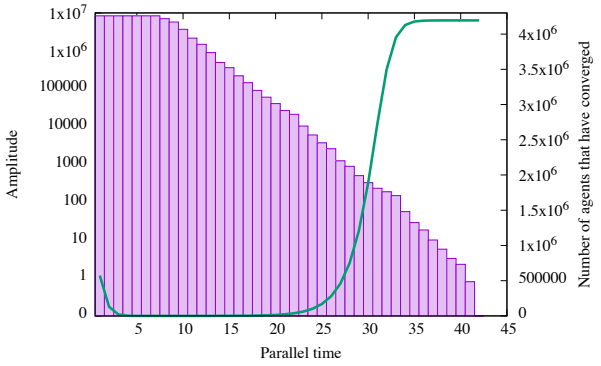
We have run a series of experiments on different sets of initial states and for different parameter settings, in particular the size of the population n to illustrate the behavior of our population protocol described in Section VI. The scheduler of the simulator chooses at each step of the simulation independently and uniformly at random the couple of interacting agents. The number of agents that share the correct output, and the configuration vectors C_t are monitored. Finally, each point on the curves represents the mean of $1,000$ paths.

B. Simulation results

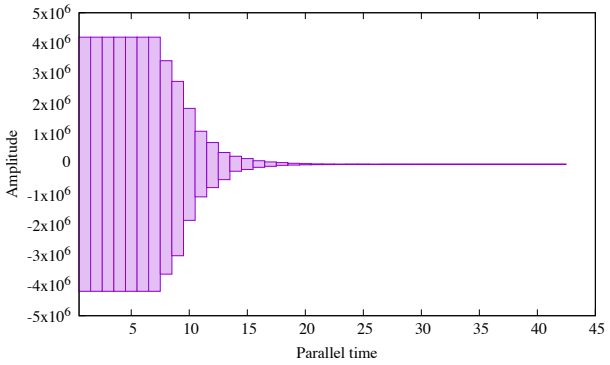
Two metrics have been studied: *(i)* the convergence parallel time to κ as a function of the number of agents and κ, and *(ii)* the amplitude A_t of the configuration vector defined by $A_t = \max_i C_t^{(i)} - \min_i C_t^{(i)}$ as a function of parallel time and κ. The conserved advantage κ has been set to 0 and $3n/5$ to reflect the good behavior of our algorithm whatever the initial gap between the initial majorities.



(a) Convergence parallel time as a function of the population size.

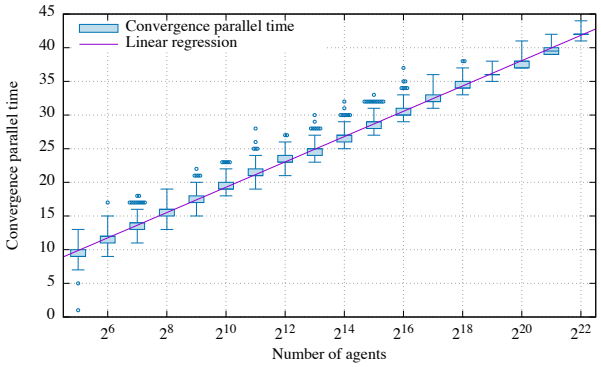


(b) A_t as a function of parallel time t (curve with bars) and number of agents that have converged as a function of parallel time.

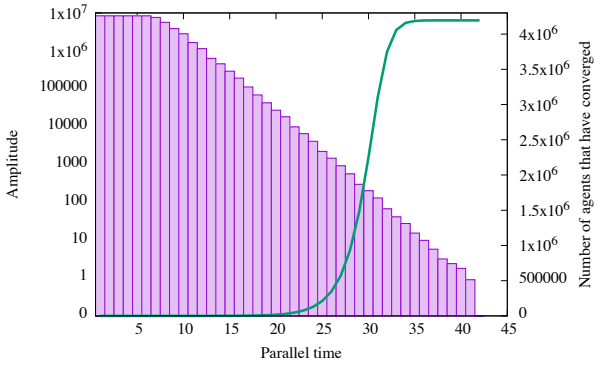


(c) $\max_{i=1,\ldots,n} C_t^{(i)}$ as a function of parallel time t (top curve) and $\min_{i=1,\ldots,n} C_t^{(i)}$ as a function of parallel time t (bottom curve).

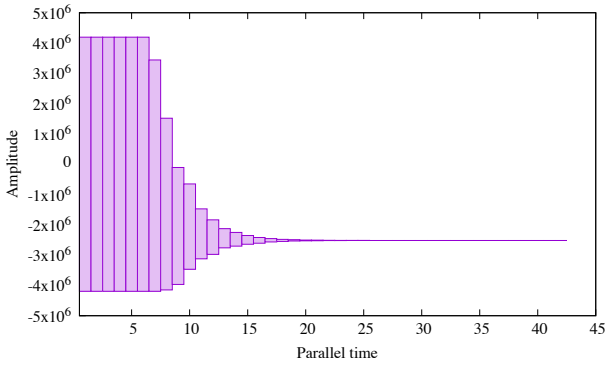
Figure 1. Evolution of the configuration vector for a conserved advantage κ equal to 0. Settings: $n = 2^{22} = 4,19 \times 10^6$.



(a) Convergence parallel time as a function of the population size n.



(b) A_t as a function of parallel time t (curve with bars) and number of agents that have converged as a function of parallel time.



(c) $\max_{i=1,\ldots,n} C_t^{(i)}$ as a function of parallel time t (top curve) and $\min_{i=1,\ldots,n} C_t^{(i)}$ as a function of parallel time t (bottom curve).

Figure 2. Evolution of the configuration vector for a conserved advantage κ equal to $3n/5$. Settings: $n = 2^{22} \approx 4,19 \times 10^6$.

1) Parallel time to convergence: Figures 1(a) and 2(a) represent the convergence parallel time to κ using box-and-whisker diagrams, as a function of the size of the network n.

Specifically, the bottom and top of the box represent respectively the first and third quartiles, and the band inside the box represents the median. The ends of the whiskers represent the 5th percentile and the 95th percentile. Any data not included between the whiskers is plotted as an outlier with small circle. Clearly, these figures perfectly illustrate the theoretical analysis (see Theorem 9), in that each agent interacts a logarithmic number of times in order to converge to κ. Note that the variability of convergence time is very low, as illustrated by the length of the box diagram. Actually, half of the experiments converge with a variability of ± 2 from the median.

2) Temporal evolution of the configuration vector: Figures 1(b) and 2(b) depict, on the curve with bars, the amplitude A_t of the configuration vector as a function of parallel time t, and on the continuous curve (whose the y-axis is on the right), the number of agents that have converged to κ as a function of parallel time. The size of the population is

equal to $n = 2^{22} \approx 4.19$ millions of agents. Both curves illustrate the fast convergence of the agents. One can notice in Figure 1(b) an interesting behavior due to the null initial advantage ($\kappa = 0$): a large number of agents seem to have converged right after their first interaction, and then depart from this converged state during a little bit less than a logarithmic number of interactions, to finally and definitively converge to the exact output value. This phenomenon is easily explained by the fact that a certain number of agents primarily interact with agents exhibiting an opposite value to their own value (m and $-m$) and thus get an average equal to 0 which is also equal to the convergence value κ. Meanwhile, all their subsequent interactions will lead to a non null average value which explains why the number of agents that have converged decrease to 0.

VIII. Conclusion

This paper has presented a population protocol that solves the counting problem, a problem that generalizes the problem to counting in addition to computing majority. Our solution is simple, i.e., it is based on an average transition function, it is efficient, i.e., it guarantees that any agent converges to the exact majority difference κ in $O(\log n)$ interactions with any high probability, and its analysis uses a simple argument that provides a proof of convergence of the averaging mechanism based on tracking the Euclidean distance between the vector of all agents' states and the uniform vector. The present work assumed that the size of the system n is known a priori. As future work, we plan to remove this assumption by combining the present algorithm with a protocol for leader election.

References

[1] Dan Alistarh, Rati Gelashvili, and Milan Vojnovíc. Fast and exact majority in population protocols. Technical Report MSR-TR-2015-13, Microsoft Research, 2015.

[2] Dana Angluin, James Aspnes, Zoë Diamadi, Michael J. Fischer, and René Peralta. Computation in networks of passively mobile finite-state sensors. *Distributed Computing*, 18(4):235–253, 2006.

[3] Dana Angluin, James Aspnes, and David Eisenstat. A simple population protocol for fast robust approximate majority. *Distributed Computing*, 20(4):279–304, 2008.

[4] Dana Angluin, James Aspnes, David Eisenstat, and Eric Ruppert. The computational power of population protocols. *Distributed Computing*, 20(4):279–304, 2007.

[5] James Aspnes and Eric Ruppert. An introduction to population protocols. *Bulletin of the European Association for Theoretical Computer Science, Distributed Computing Column*, 93:98–117, 2007.

[6] Moez Draief and Milan Vojnovic. Convergence speed of binary interval consensus. *SIAM Journal on Control and Optimization*, 50(3):1087–11097, 2012.

[7] George B. Mertzios, Sotiris E. Nikoletseas, Christoforos Raptopoulos, and Paul G. Spirakis. Determining majoirty in networks with local interactions and very small local memory. In *Proceedings of the 41st International Colloquium (ICALP)*, pages 871–882, 2014.

[8] Etienne Perron, Dinkar Vasudevan, and Milan Vojnovic. Using three states for binary consensus on complete graphs. In *Proceedings of the INFOCOM conference*, pages 2527–2435, 2009.