Probab. Theory Relat. Fields 102, 123–143 (1995)

Probability Theory and Related Fields

Asymptotical behaviour of several interacting annealing processes

Alain Trouvé

LMENS/DIAM, URA 762, Ecole Normale Superieure, 45 rue d'Ulm, F-75230 Paris Cedex 05, France

Received: 5 October 1993 / In revised form: 3 January 1995

Summary. We prove that the optimal convergence speed exponent for parallel annealing based on periodically interacting multiple searches with time period r is always worse than for independent multiple searches whenever the cost function has only one global minimum. Our proofs will be based on large deviation estimates coming from the theory of generalized simulated annealing (G.S.A).

Mathematics Subject Classification (1991): 60F10, 60J10, 93E25

1 Introduction

Let E be a finite set and consider a family of Markov kernels $(Q_T)_{T>0}$ satisfying for any $(i,j) \in E^2$ the inequalities

$$(1) \qquad \frac{1}{\kappa} q(i,j) e^{-V(i,j)/T} \leq Q_T(i,j) \leq \kappa q(i,j) e^{-V(i,j)/T}$$

where q is an irreducible Markov kernel on E called the communication kernel, $\kappa \geq 1$ and $V : E \times E \to \mathbb{R}_+ \cup \{+\infty\}$ is a non negative real valued function called the communication cost, and satisfying $V(i,j) = +\infty$ iff $q(i,j) = 0$. Assume that the family $(Q_T)_{T>0}$ satisfies the additionnal weak condition (8). Now, considering a non increasing sequence $(T_n)_{n\in\mathbb{N}}$ of positive real valued numbers called the cooling schedule, the theory of generalized simulated annealing is concerned with the behaviour of the Markov chain $(X_n)_{n\in\mathbb{N}}$ on E defined by the one-step transitions $P(X_{n+1} = j \mid X_n = i) = Q_{T_{n+1}}(i,j)$.

Many stochastic optimization algorithms can be properly described as generalized simulated annealing. For instance, for sequential simulated annealing, we restrict ourself to symmetrical communication kernel q and to communication costs V given by $V(i,j) = (U(j) - U(i))^+$ if $q(i,j) > 0$ (and $V(i,j) = +\infty$ otherwise) where U is the function to be minimized. Moreover, many parallelized versions of sequential annealing fall into the scope of generalized simulated annealing. In the general framework, $(X_n)_{n\in\mathbb{N}}$ concentrates, for sufficiently slowly decreasing cooling schedules, on the minima of the virtual energy W defined on E by $W(i) = \lim_{T\to 0} T\ln(\mu_T(i))$ where μ_T is the unique invariant probability measure of Q_T (note that $\min W = 0$). The rate of convergence towards the minima of W has been studied by C.R Hwang and S.J. Sheu in [11] and more recently by the author in [13, 14, 15] (see also T.S Chiang and Y. Chow [7], L. Miclo [12], and J. D. Deuschel and C. Mazza [8] for the continuous time setting). This rate of convergence is characterized by two critical constants H_1 and α which depend on the decomposition of the state space E in cycles according to the large deviation approach developed by Wentzell and Freidlin. In particular, there exist two strictly positive constants K_1 and K_2 such that (see [6, 14])

$$\frac{K_1}{N^\alpha} \leq \sup_{i\in E} \inf_{T_0\geq\cdots\geq T_N} P(\ W(X_N) > 0 \mid X_0 = i\ ) \leq \frac{K_2}{N^\alpha}. \tag{2}$$

In order to speed up the convergence towards the minima of W, one can use a multi-processors computer to parallelize the previous algorithm. Many efforts have been made in this direction with notable success for precise applications [1, 2, 3]. Among these algorithms, we will be interested more particularly in the multiple searches algorithms. Assume that p processors $(P_k)_{1\leq k\leq p}$ are available. Then consider the multidimensional Markov chain $(\mathbf{X}_n)_{n\in\mathbb{N}}$ on E^p defined by the probability transitions

$$P(\mathbf{X}_{n+1} = \mathbf{j} \mid \mathbf{X}_n = \mathbf{i}) = \prod_{k=1}^{p} Q_{T_{n+1}}(\mathbf{i}^k, \mathbf{j}^k)\ ;\ \mathbf{i},\mathbf{j} \in E^p. \tag{3}$$

The variable $\mathbf{X}_n$ is a random vector denoting the value at time n of p independent generalized annealing algorithms under the same cooling schedule. Now, let

$$S(\mathbf{X}_n) = \inf_{1\leq k\leq p} W(\mathbf{X}_n^k).$$

We deduce immediately that

$$\frac{K_1'}{N^{p\alpha}} \leq \sup_{\mathbf{i}\in E^p} \inf_{T_0\geq\cdots\geq T_N} P(S(\mathbf{X}_N) > 0 \mid \mathbf{X}_0 = \mathbf{i}\ ) \leq \frac{K_2'}{N^{p\alpha}}. \tag{4}$$

so that the rate of convergence is improved by a factor of p. This scheme will be called parallel annealing based on non interacting multiple searches. It has been argued by E. Aarts and P. van Laarhoven that interactions between processors should increase the convergence rate of the algorithm. Obviously, many interactions are possible and in [4] R. Azencott and C. Graffigne formalized a simplification of an interacting scheme used by E. Aarts and P. van Laarhoven without mathematical study. It is a natural extension of the previous scheme which is the parallel annealing based on *interacting* multiple searches. Instead of performing independent searches, we can fix an integer $r \geq 1$ that will be the time period between interactions. Now, consider the Markov chain $(\mathbf{Z}_n)_{n\geq 0}$ on

E^p defined through the following algorithmic description. At each time step, each processor P_k performs r usual search-steps from its current configuration $\mathbf{Z}_n^k$ at temperature T_{n+1}. We denote $\mathbf{Y}_{n+1}^k$ the result of these r consecutive search-steps. The new configuration $\mathbf{Z}_{n+1}^k$ is recursively defined by $\mathbf{Z}_{n+1}^1 = \mathbf{Y}_{n+1}^1$ and for $k \geq 2$

$$(5) \qquad \mathbf{Z}_{n+1}^k = \begin{cases} \mathbf{Z}_{n+1}^{k-1} & \text{if } W(\mathbf{Y}_{n+1}^k) \geq W(\mathbf{Z}_{n+1}^{k-1}) \\ \mathbf{Y}_{n+1}^k & \text{otherwise .} \end{cases}$$

As previously, we defined $S(\mathbf{Z}_n) = \inf_{1 \leq k \leq p} W(\mathbf{Z}_n^k) = W(\mathbf{Z}_n^p)$. The main result of this paper will be that under the condition that W has only one global minimum, there exists a strictly positive constant K such that

$$(6) \qquad \frac{K}{N^{\alpha_{int}}} \leq \sup_{\mathbf{i} \in E^p} \inf_{T_0 \geq \cdots \geq T_N} P(S(\mathbf{Z}_N) > 0 \mid \mathbf{Z}_0 = \mathbf{i}\ ) \text{ with } \alpha_{int} \leq p\alpha.$$

If we compare (4) and (6) we deduced that, surprisingly, the interaction between the different processors does not increase the convergence speed towards the global minima of W. Moreover, one can easily exhibit situations where $\alpha_{int} = \alpha$. This result has been conjectured by R. Azencott and C. Graffigne [4] through a mathematical study of particular configuration spaces E of small size for $p = 2$ and confirms the intuition given by experimental studies performed in [10].

The effective computation of critical constants is in many cases a very delicate task. In many concrete situation, even in the sequential annealing framework with a given energy function U, the value of the critical exponent is unknown and its computation is a very hard combinatorial problem. Apparently our problem is simpler since we need here only a comparison between two exponents. However, even if we assume that the critical exponent of an annealing process is known, under a slight modification of the dynamic, the evolution of the critical exponent is often unpredictable and the modification of the energy landscape involved can be very intricate from the combinatorial point of view. For instance, for the non interacting multiple searches algorithm defined by (3), the communication cost function $\mathbf{V}$ is defined by $\mathbf{V}(\mathbf{i},\mathbf{j}) = \sum_{k=1}^p V(\mathbf{i}^k, \mathbf{j}^k)$. If we try to use the combinatorial definition of the virtual energy $\mathbf{W}$ associated with $\mathbf{V}$ as given by Wentzell and Freidlin in [9], we will not be able to prove that $\mathbf{W}(\mathbf{i}) = \sum_{k=1}^p W(\mathbf{i}^k)$ even if this equality is straightforward from a probabilistic point of view. We will encounter a more difficult problem if we try to say something about the cycle decomposition of the product space for $\mathbf{W}$ since this decomposition is related with the computation of the communication altitude which is until now very badly understood even in such a simple case. Now if we consider the problem of periodically interacting generalized annealing processes, we add a difficulty coming from the interactions between the processes. In this case, there is no simple relation between the virtual energy of the interacting processes and the initial virtual energy W. A strictly combinatorial approach for the computation of the critical exponent of the interacting processes seems to be an unreachable issue. As for the case of independent processes, our point of view will be to guess from the observation of the behaviour of the interacting processes in precise situations some of the main elements of the underlying cycle decomposition. Hence, in Sect. 2 we recall briefly the essentials of the theory of simulated annealing and in Sect. 3 we establish some kind of inverse theorems describing how the observation of a generalized simulated annealing process for different constant cooling schedules can give important information on the cycle decomposition. Finally, in the last section, we prove the main result.

2 Some basic results about generalized simulated annealing

In this section we want to recall some basic results about the generalized simulated annealing which will be extensively used in the following. All of them are proved in [14] so that we do not mention the proofs. Since we will have to consider different generalized simulated annealing algorithms on different configuration spaces, we will consider for the statement of the basic results a generic configuration space $\mathscr{E}$ on which a family $(\mathscr{Q}_T)_{T>0}$ of Markov kernels is defined, satisfying for any $i,j \in \mathscr{E}$

$$(7) \qquad \frac{1}{\kappa}\mathrm{q}(i,j)e^{-\mathscr{V}(i,j)/T} \leq \mathscr{Q}_T(i,j) \leq \kappa \mathrm{q}(i,j)e^{-\mathscr{V}(i,j)/T}$$

where, as in the introduction, q is an irreducible Markov kernel on $\mathscr{E}$ called the communication kernel, $\kappa \geq 1$ and $\mathscr{V} : E \times E \to \mathbb{R}_+ \cup \{+\infty\}$ is the communication cost satisfying $\mathscr{V}(i,j) = +\infty$ iff $\mathrm{q}(i,j) = 0$. We define the virtual energy $\mathscr{W}$ on $\mathscr{E}$ by $\mathscr{W}(i) = \lim_{T\to 0} T \ln(\mu_T(i))$ where μ_T is the unique invariant probability measure of $\mathscr{Q}_T$ (see [9]) (note that in our definition we have $\min_{i\in\mathscr{E}} \mathscr{W}(i) = 0$). Furthermore, we will assume that for any $i,j \in \mathscr{E}$ there exists a finite family $(a_k^{ij}, b_k^{ij}, c_k^{ij}, d_k^{ij})_{k\in I_k}$ of elements in $\mathbb{R}^4$ such that

$$(8) \qquad \mathscr{Q}_T(i,j) = (\sum_{k\in I_k} a_k^{ij} \exp(b_k^{ij}/T))/(\sum_{k\in I_k} c_k^{ij} \exp(d_k^{ij}/T)),$$

where $\sum_{k\in I_k} c_k^{ij} \exp(d_k^{ij}/T) \neq 0$ for any $T > 0$.

2.1 Critical exponent

We start with some notations (for an extended presentation see [14, 15]).

Notation 1. *Let $B \subset \mathscr{E}$. Let any finite family $g = (g_k)_{0\leq k\leq n_g}$ of elements of B such that $g_0 = i$ and $g_{n_g} = j$ be called a path in B from i to j. The integer n_g (depending on g) is called the length of the path g. Let $Pth_B(i,j)$ denote the set of all paths in B from i to j.*

Definition 1. – *We define the communication altitude from i to j by*

$$\mathscr{A}_s(i,j) = \begin{cases} \inf_{g\in Pth_{\mathscr{E}}(i,j)} \sup_{0\leq k<n_g} (\mathscr{W}(g_k) + \mathscr{V}(g_k, g_{k+1})) & \text{if } i \neq j, \\ \mathscr{W}(i) & \text{if } i = j. \end{cases}$$

- *We say that a non empty subset $\Pi \subset \mathcal{E}$ is a cycle if Π is a singleton or Π satisfies $\sup_{i,j\in\Pi} \mathcal{A}_s(i,j) < \inf_{i\in\Pi, j\in\Pi^c} \mathcal{A}_s(i,j)$. We note $\mathcal{C}(\mathcal{E})$ the set of all the cycles.*
- *For any cycle Π, $H_m(\Pi) = \sup_{i,j\in\Pi} (\mathcal{A}_s(i,j) - \mathcal{W}(i))$ will be called the mixing height of Π and $H_e(\Pi) = \sup_{i\in\Pi} \inf_{j\in\Pi^c} (\mathcal{A}_s(i,j) - \mathcal{W}(i))$ its exit height.*
- *For any $B \subset \mathcal{E}$, we define*

$$\begin{aligned}
\mathcal{M}(B) &= \{ \Pi \in \mathcal{C}(\mathcal{E}) \mid \Pi \subset B \text{ and maximal for inclusion } \} \\
\mathcal{M}_*(B) &= \{ \Pi \in \mathcal{C}(\mathcal{E}) \mid \Pi \subset B, \Pi \neq B \text{ and maximal for inclusion } \} \\
\mathcal{W}(B) &= \inf\{ \mathcal{W}(i) \mid i \in B \} \text{ (potential of } B\text{)}, \\
F(B) &= \{ i \in B \mid \mathcal{W}(i) = \mathcal{W}(B) \} \text{ (bottom of } B\text{)}, \\
H_e(B) &= \sup\{ H_e(\Pi) \mid \Pi \in \mathcal{C}(\mathcal{E}),\ \Pi \subset B \} \text{ (exit height of } B\text{)}.
\end{aligned}$$

Definition 2. – *For any $i \in \mathcal{E}$, let Π_i denote the largest cycle Π such that $i \in F(\Pi)$.*
- *Finally, we define the critical exponent by $\alpha = \inf_{i\notin F(E)} \mathcal{W}(i)/H_e(\Pi_i)$.*

Definition 3. *Let $(\mathcal{X}_n)_{n\in\mathbb{N}}$ be the coordinate process on $\mathcal{E}^{\mathbb{N}}$. For any cooling schedule $\mathcal{T} = (\mathcal{T}_n)_{n\in\mathbb{N}}$ (i.e for any non negative sequence) we denote by $P_{\mathcal{T}}$ the unique probability measure on $\mathcal{E}^{\mathbb{N}}$ with its natural product sigma-algebra such that $(\mathcal{X}_n)_{n\in\mathbb{N}}$ is a Markov chain satisfying:*

- $P_{\mathcal{T}}(\mathcal{X}_{n+1} = j \mid \mathcal{X}_n = i) = \mathcal{Q}_{T_{n+1}}(i,j)$,
- $P_{\mathcal{T}}(\mathcal{X}_0 = i) = \nu_0(i)$ *where ν_0 is a fixed initial probability on $\mathcal{E}$ whose support is $\mathcal{E}$.*

For constant cooling schedules $\mathcal{T}_n \equiv T$, $n \in \mathbb{N}$, we will write P_T instead of $P_{\mathcal{T}}$.

Theorem 1. *[14, 15] Let α be the critical exponent corresponding to the cost function $\mathcal{V}$ and assume that $\alpha < +\infty$. Then there exist two strictly positive constants $K_1 > 0$ and $K_2 > 0$ such that*

$$\frac{K_1}{N^\alpha} \leq \sup_{i\in\mathcal{E}} \inf_{T_0\geq\cdots\geq T_N} P_{\mathcal{T}}(\mathcal{X}_N \notin F(\mathcal{E}) \mid \mathcal{X}_0 = i) \leq \frac{K_2}{N^\alpha}$$

$$\text{and } \frac{K_1}{N^{\alpha_\Pi}} \leq \sup_{i\in\mathcal{E}} \inf_{T_0\geq\cdots\geq T_N} P_{\mathcal{T}}(\mathcal{X}_N \in \Pi \mid \mathcal{X}_0 = i),$$

for any $\Pi \in \mathcal{C}(E)$ such that $\Pi \cap F(E) = \emptyset$, where $\alpha_\Pi = \mathcal{W}(\Pi)/H_e(\Pi)$.

2.2 Main large deviation estimates

In this section, we recall some of the large deviation estimates on exit time and exit point for a generalized annealing algorithm established in [14] and [15]. We consider only the simple case when there is no annealing, thus assuming that the temperature is kept constant.

Definition 4. *Let $B \subset \mathscr{E}$. We define for $i,j \in \mathscr{E}$*

$$\begin{aligned} C_B(i,j) &= (\mathscr{W}(i) + \mathscr{V}(i,j) - \inf_{k \in B^c \cup \{j\}} \mathscr{H}_s(i,k))\mathbf{1}_{i \neq j}, \\ C_B^*(i,j) &= \inf\{ \sum_{k<n_g} C_B(g_k, g_{k+1}) \mid g \in Pth_{\mathscr{E}}(i,j);\ g_k \in B \text{ for } 0 < k < n_g \}. \end{aligned}$$

For i and j in B, $C_B^*(i,j)$ can be interpreted as the communication cost to go from i to j without escaping from B.

Definition 5. *1. Let $B \subset \mathscr{E}$. We define $\tau(B,m) = \inf\{ n > m,\ \mathscr{X}_n \notin B \}$. For $m = 0$, $\tau(B)$ will denote the stopping time $\tau(B,0)$.*
2. Let $B \subset \mathscr{E}$, $G \subset \mathscr{E}$. For $i,j \in \mathscr{E}$ and $m,n \in \mathbb{N}$ we define

$$\begin{aligned} M(B,G)_{i,m}^{j,n} &= P_{\mathscr{T}}(\ \tau(B,m) \geq n,\ \mathscr{X}_{n-1} \notin G,\ \mathscr{X}_n = j \mid \mathscr{X}_m = i\ )\mathbf{1}_{m<n,\ j \in G} \\ L(B,G)_{i,m}^{j,n} &= P_{\mathscr{T}}(\ \tau(B,m) > n,\ \mathscr{X}_n = j \mid \mathscr{X}_m = i\ )\mathbf{1}_{m \leq n,\ j \in G} \end{aligned}$$

Theorem 2. *Let $T > 0$ and assume that $\mathscr{T}_n = T$, $n \geq 0$. There exist $a > 0$, $b \geq 0$, $c > 0$, $d > 0$, $K_1 > 0$ and $K_2 > 0$ depending only on $\mathscr{E}$, q and κ such that:*

(i) For any $B \subset \mathscr{E}$, $B \neq \emptyset$, any $j \in B^c$ and any $n \in \mathbb{N}$

$$K_1 e^{-C_B^*(i,j)/T} \left(1 - (1+b)\exp(-ane^{-H_e(B)/T})\right)^+ \leq \sum_{l=0}^{n} M(B,B^c)_{i,m}^{j,m+l}\ ; i \in B$$

$$\sum_{l \geq 0} M(B,B^c)_{i,m}^{j,m+l} \leq K_2 e^{-C_B^*(i,j)/T}\ ;\ i \in \mathscr{E}.$$

(ii) For any $B \subset \mathscr{E}$, any $n \in \mathbb{N}$ and any $i \in B$

$$P_T(\ \tau(B) > n \mid \mathscr{X}_0 = i\ ) \leq (1+b)\exp(-ane^{-H_e(B)/T}).$$

(iii) For any $\Pi \in \mathscr{C}(\mathscr{E})$, any $n \in \mathbb{N}$ and any $i \in \Pi$

$$P_T(\ \tau(\Pi) > n \mid \mathscr{X}_0 = i\ ) \geq c\exp(-ane^{-H_e(\Pi)/T})\mathbf{1}_{e^{-H_e(\Pi)/T} \leq d}.$$

(iv) For any $\Pi \in \mathscr{C}(\mathscr{E})$, any $f \in F(\Pi)$ and any $j \in \Pi$

$$\sum_{l \geq 0} L(\Pi \setminus F(\Pi), \Pi \setminus F(\Pi))_{f,m}^{j,m+l} \leq K_2 e^{-(\mathscr{W}(j) - \mathscr{W}(\Pi))/T}.$$

Proof. This is an obvious corollary for constant cooling schedule of theorem 4.1 and 4.7 in [15] or theorem 1.43 and 1.46 in [14]. □

3 From energy landscape shape to G.S.A's behaviour...and back

In this section, we will put the emphasis on the relation between generalized simulated annealing at low temperature and the shape of the energy landscape given by $\mathscr{W}$ and $\mathscr{A}_s$. On the one hand, the computation of the communication altitude and of the cycle decomposition is the key to establish the convergence speed of the G.S.A. On the other hand, in many cases, this computation is intractable from a combinatorial point of view. Therefore, we have to play a more delicate game. We have to deduce from partial information on the behaviour in some particular situations of the G.S.A., some information on the energy landscape involved, which in turn gives some new information on the behaviour of the G.S.A. Hence, we need several new results, which are more or less direct consequences of theorem 2. Fortunately, we will need only to study the behaviour at constant temperature so that throughout this section we assume that $T_n = T$ for all $n \in \mathbb{N}$.

Since the proofs of the following lemmas are often technical, the reader is invited to skip the proofs at first reading.

3.1 Cumulated mass through a state before escaping from a cycle

We give below an estimation of the mean value of the number of time the process reaches a given state j in a cycle Π before the exit from Π and during a time of order $e^{H/T}$ for $H \geq H_m(\Pi)$. Intuitively, the process reaches a thermal equilibrium in the cycle Π within a time of order $e^{H_m(\Pi)/T}$ so that the probability to be in the state j is about $e^{-(\mathscr{W}(j)-\mathscr{W}(\Pi))/T}$ and the estimation should be of order $e^{(H+\mathscr{W}(\Pi)-\mathscr{W}(j))/T}$. However since the process goes out of the cycle Π in a time of order $e^{H_e(\Pi)/T}$, the above estimation should be restricted to $H \leq H_e(\Pi)$.

Lemma 1. *Let $\Pi \in \mathscr{C}(\mathscr{E})$. For any $j \in \Pi$ and any $R > 0$ there exists $K > 0$ such that for any $H \geq H_m(\Pi)$ we have*

$$\sup_{i\in\Pi} \sum_{0\leq n<[Re^{H/T}]} L(\Pi,\Pi)_{i,0}^{j,n} \leq K(e^{H_e(\Pi)/T} \wedge e^{H/T})e^{(\mathscr{W}(\Pi)-\mathscr{W}(j))/T},$$

where $[x]$ denote the integer part of x.

Proof. We will prove the result by induction on $|\Pi|$. Assume that $\Pi = \{i\}$. Since $L(\Pi,\Pi)_{i,0}^{i,n} = P_T(\ \tau(\Pi) > n \mid \mathscr{X}_0 = i\ )$, we get from theorem 2 (ii) that there exists $K > 0$ such that for any $H \geq 0$

$$\sum_{n=0}^{[Re^{H/T}]-1} L(\Pi,\Pi)_{i,0}^{i,n} \leq (1+b)\frac{1-e^{-aRe^{(H-H_e(\Pi))/T}}}{1-e^{-ae^{-H_e(\Pi)/T}}} \leq K(e^{H/T} \wedge e^{H_e(\Pi)/T}).$$

Assume now that $|\Pi| = n+1$. If $j \in F(\Pi)$ then $L(\Pi,\Pi)_{i,0}^{j,n}$ is upper bounded by $P_T(\ \tau(\Pi) > n \mid \mathscr{X}_0 = i\ )$, and the result follows as for $|\Pi| = 1$. Otherwise

($j \notin F(\Pi)$), let Π' be the unique cycle in $\mathscr{M}(\Pi \setminus F(\Pi))$ such that $\Pi' \ni j$. We have for $q \in \mathbb{N}$

$$\sum_{n=0}^{q} L(\Pi,\Pi)_{i,0}^{j,n} = \sum_{n=0}^{q}\left(L(\Pi',\Pi')_{i,0}^{j,n} + (M(\Pi \setminus F(\Pi),\Pi')L(\Pi',\Pi'))_{i,0}^{j,n}\right)$$
$$+ \sum_{0<m\leq n<q,\ f\in F(\Pi)} L(\Pi,\Pi)_{i,0}^{f,m} L(\Pi \setminus F(\Pi),\Pi \setminus F(\Pi))_{f,m}^{j,n}.$$

From the induction hypothesis we get for $j' \in \Pi'$ and $m' \geq 0$

$$\sum_{n\geq m'} L(\Pi',\Pi'))_{j',m'}^{j,n} \leq Ke^{(\mathscr{W}(\Pi')+H_e(\Pi')-\mathscr{W}(j))/T}.$$

From theorem 2 (iv) we get that for any fixed $m > 0$

$$\sum_{n\geq m}(L(\Pi \setminus F(\Pi),\Pi \setminus F(\Pi))_{f,m}^{j,n} \leq Ke^{-(\mathscr{W}(j)-\mathscr{W}(\Pi))/T}.$$

Since $L(\Pi,\Pi)_{i,0}^{f,m} \leq P_T(\ \tau(\Pi) > m \mid \mathscr{X}_0 = i\ )$ we deduce as for the case $|\Pi| = 1$ that $\sum_{0<m<[Re^{H/T}]} L(\Pi,\Pi)_{i,0}^{f,m} \leq K(e^{H_e(\Pi)/T} \wedge e^{H/T})$. This ends the proof. □

Lemma 2. *Let Π be in $\mathscr{C}(\mathscr{E})$. For any $0 \leq H \leq H_e(\Pi)$, any $f \in F(\Pi)$ and any $R > 0$, there exist $T_0 > 0$ and $K > 0$ such that for any $0 < T \leq T_0$ we have*

$$\sum_{i\in F(\Pi)}\ \sum_{0\leq l\leq Re^{H/T}} L(\Pi,\Pi)_{f,0}^{i,l} \geq Ke^{H/T}.$$

Proof. If $H_e(\Pi) = 0$ then the result is trivial. Therefore, $H_e(\Pi) > 0$ is assumed now. Since for $f \in F(\Pi)$, $L(\Pi,\Pi)_{f,0}^{\cdot,\cdot} = (L(\Pi,F(\Pi))L(\Pi \setminus F(\Pi),\Pi \setminus F(\Pi)))_{f,0}^{\cdot,\cdot}$, using the trivial upper bound $L(\Pi,F(\Pi)) \leq 1$ and theorem 2 (iv), we get

$$\sum_{i\in\Pi\setminus F(\Pi)}\ \sum_{0\leq k\leq Re^{H/T}} L(\Pi,\Pi)_{f,0}^{i,k} \leq Ke^{(H-\delta)/T},$$

where $\delta = \inf\{\ \mathscr{W}(i) - \mathscr{W}(\Pi) \mid i \in \Pi \setminus F(\Pi)\ \} > 0$. However, from theorem 2 (iii) we deduce that there exists $K > 0$ such that for small $T > 0$ we have

$$\sum_{0\leq k\leq Re^{H/T}}\ \sum_{i\in\Pi} L(\Pi,\Pi)_{f,0}^{i,k} \geq \sum_{0\leq k\leq Re^{H/T}} c\exp(-kae^{-H_e(\Pi)/T}) \geq Ke^{H/T}.$$

The lemma follows immediately. □

3.2 Ergodic behaviour within a cycle before exit

It is well known (see [9]) that the exit time of a cycle Π is of order $e^{H_e(\Pi)/T}$. We examine in the following lemma the probability of the deviation ($\tau(\Pi) \leq Re^{H/T}$) for H smaller than $H_e(\Pi)$.

Lemma 3. *Let $\Pi \in \mathscr{C}(\mathscr{E})$. For any $H_m(\Pi) \leq H \leq H_e(\Pi)$ and any $R > 0$ there exists $K > 0$ such that for any $T > 0$ we have for any $i \in \Pi$*

$$P_T(\ \tau(\Pi) \leq Re^{H/T} \mid \mathscr{X}_0 = i\ ) \leq Ke^{(H-H_e(\Pi))/T}.$$

Moreover, if $i \in F(\Pi)$, the result still holds for $0 \leq H < H_m(\Pi)$.

Proof. Let $f \in F(\Pi)$. Note that for $p \in \mathbb{N}$, $P_T(\ \tau(\Pi) \leq p \mid \mathscr{X}_0 = i\ )$ is equal to

$$\sum_{j\in\Pi^c}\Big(\sum_{0\leq m<n\leq p} L(\Pi,\Pi)^{f,m}_{i,0} M(\Pi\setminus\{f\},\Pi^c)^{j,n}_{f,m} + \sum_{0<n\leq p} M(\Pi\setminus\{f\},\Pi^c)^{j,n}_{i,0}\mathbf{1}_{i\neq f}\Big).$$

We get from theorem 2 (i) that

$$\sum_{j\in\Pi^c,\ 0<n} M(\Pi\setminus\{f\},\Pi^c)^{j,n}_{i,0} \leq Ke^{(H_m(\Pi)-H_e(\Pi))/T},$$

since $C^*_{\Pi\setminus\{f\}}(i,j) \geq H_e(\Pi) - H_m(\Pi)$ for $j \in \Pi^c$. Since $C^*_{\Pi\setminus\{f\}}(f,j) \geq H_e(\Pi)$ we deduce that for m fixed $\sum_{n\geq m} M(\Pi\setminus\{f\},\Pi^c)^{j,n}_{f,m} \leq Ke^{-H_e(\Pi)/T}$. Hence, using the trivial upper bound $L(\Pi,\Pi) \leq 1$ we get the result. □

Before escaping from a cycle Π, the annealing process visits all the configurations of the cycle in a time of order $e^{H_m(\Pi)/T}$. This is proved by the following lemma.

Lemma 4. *Let Π be a cycle such that $H_e(\Pi) > 0$. For any $i,j \in \Pi$ and any $\epsilon > 0$, there exist $R > 0$ and $T_0 > 0$ such that for any $0 < T \leq T_0$ we have*

$$P_T(\ \tau(\Pi\setminus\{j\}) < \tau(\Pi) \wedge Re^{H_m(\Pi)/T} \mid \mathscr{X}_0 = i\ ) \geq 1-\epsilon.$$

Proof. From theorem 2 (ii), since $H_e(\Pi\setminus\{j\}) \leq H_m(\Pi)$), we deduce that there exists $R > 0$ such that for any $i,j \in \Pi$

$$P_T(\ \tau(\Pi\setminus\{j\}) > Re^{H_m(\Pi)/T} \mid \mathscr{X}_0 = i\ ) \leq (1+b)\exp(-aR) \leq \epsilon/2.$$

Moreover, since $H_e(\Pi) > 0$, we have $H_e(\Pi) > H_m(\Pi)$ and we get from lemma 3 that there exists $T_0 > 0$ such that for any $0 < T < T_0$

$$P_T(\ \tau(\Pi) > Re^{H_m(\Pi)/T} \mid X_0 = i) \geq 1-\epsilon/2$$

so that the lemma is proved. □

The next result precise the probability to join two distinct configurations in a time of order $e^{H/T}$ with H smaller than the mixing height $H_m(\Pi)$ without escaping from a cycle Π.

Lemma 5. *Let Π be a cycle such that $|\Pi| > 1$. Then for any $H \geq 0$ such that $\sup_{f \in F(\Pi)} H_e(\Pi \setminus \{f\}) \leq H \leq H_m(\Pi)$ and for any $j \in \Pi$, there exist $T_0 > 0$, $R > 0$ and $K > 0$ such that for any $0 < T \leq T_0$ we have*

$$\inf_{i \in \Pi} P_T(\ \tau(\Pi \setminus \{j\}) < \tau(\Pi) \wedge Re^{H/T} \mid \mathcal{X}_0 = i\ ) \geq Ke^{-(H_m(\Pi)-H)/T}.$$

Proof. Let $f \in F(\Pi)$. Since we have $H_e(\Pi \setminus \{f,j\}) \leq H_e(\Pi \setminus \{f\}) \leq H$, $C^*_{\Pi \setminus \{f,j\}}(f,j) \leq H_m(\Pi)$ and $C^*_{\Pi \setminus \{f\}}(i,f) = 0$ for any $i \in \Pi \setminus \{f\}$, we deduce from theorem 2 (i) and lemma 2 that there exist $R > 0$ and $K > 0$ such that for any $i \in \Pi \setminus \{f\}$

$$\sum_{l < Re^{H/T}} M(\Pi \setminus \{f\}, \{f\})^{f,l}_{i,0} > K, \quad \sum_{l < Re^{H/T}} M(\Pi \setminus \{f,j\}, \{j\})^{j,l}_{f,0} \geq Ke^{-H_m(\Pi)/T}.$$

Moreover $\sum_{0 \leq l < 2Re^{H/T}} L(\Pi, \{f\})^{f,l}_{f,0}$ is greater than

$$\sup\{ \sum_{0 \leq l < Re^{H/T}} L(\Pi, \{f\})^{f,l}_{f,0}, \sum_{l < 2Re^{H/T}} \big(L(\Pi, F(\Pi) \setminus \{f\}) M(\Pi \setminus \{f\}, \{f\})\big)^{f,l}_{f,0}\},$$

and $L(\Pi, F(\Pi)) = L(\Pi, \{f\}) + L(\Pi, F(\Pi) \setminus \{f\})$ so that we get from lemma 2 that $\sum_{0 \leq l < 2Re^{H/T}} L(\Pi, \{f\})^{f,l}_{f,0} \geq Ke^{H/T}$ for a new constant $K > 0$.

Now, since $P_T(\ \tau(\Pi \setminus \{j\}) < \tau(\Pi) \wedge 4Re^{H/T} \mid \mathcal{X}_0 = i\ )$ is greater than

$$\sum_{0 \leq l < 4Re^{H/T}} \Big(\big(I(\Pi \setminus \{f\}) M(\Pi \setminus \{f\}, \{f\}) + I(\{f\})\big) L(\Pi, \{f\}) M(\Pi \setminus \{f,j\}, \{j\})\Big)^{j,l}_{i,0}$$

(where $I(A)^{b,n}_{a,m} = \mathbf{1}_{a=b, m=n, a \in A}$) we deduce easily the result. □

3.3 Returns to a state

In the following proposition, we study the returns to i of a process starting from i. We show that there exists with arbitrary hight probability a passage through i in a time of order $e^{H/T}$ for any $H < H_e(\Pi_i)$. Moreover, if $H_e(\Pi) < H < H_e(\Pi) + \eta$ for an adequate strictly positive η, then the probability that there exits a passage through i in a time of order $e^{H/T}$ tends to 0 with the temperature.

Proposition 1. *Let i be in $\mathcal{E}$.*

(i) Let Π be a cycle such that $F(\Pi) \ni i$. For any $H \geq 0$ such that $H < H_e(\Pi)$, any $\epsilon > 0$ and any $R_0 > 0$, there exist $R_1 > 0$, $K > 0$ and $T_0 > 0$ such that for any $0 < T \leq T_0$ we have

$$P_T(\ \exists n \in \mathbb{N},\ R_0 \leq ne^{-H/T} \leq R_1,\ \mathcal{X}_n = i,\ n \leq \tau(\Pi) \mid \mathcal{X}_0 = i\ ) \geq 1 - \epsilon.$$

(ii) Assume that $i \notin F(\mathscr{E})$. There exists $\eta > 0$ such that for any $H \geq 0$ satisfying $H_e(\Pi_i) < H < H_e(\Pi_i) + \eta$ and for any $0 < R_0 < R_1$ we have

$$\lim_{T \to 0} P_T(\ \exists n \in \mathbb{N},\ R_0 \leq ne^{-H/T} \leq R_1,\ \mathscr{X}_n = i \mid \mathscr{X}_0 = i\ ) = 0.$$

Proof (i). We assume that $H_e(\Pi) > 0$ and we prove this last result by induction on the size $|\Pi|$ of Π. If $|\Pi| = 1$, it is sufficient to notice that for $H < H_e(\Pi)$, lemma 3 says that for any $R > 0$, $P_T(\ \tau(\Pi) > Re^{H/T} \mid \mathscr{X}_0 = i\ ) \leq \epsilon$ for small T. Assume now that $|\Pi| = n + 1$. If $H < H_m(\Pi)$, consider $\Pi' \in \mathscr{M}_*(\Pi)$ such that $F(\Pi') \ni i$ and $H < H_e(\Pi') = H_m(\Pi)$. The result follows then immediately from the induction hypothesis. We assume now that $H \geq H_m(\Pi)$. From the lemma 4 there exist $R > 0$ and $T_0 > 0$ such that we have the inequality $P_T(\ \tau(\Pi \setminus \{i\}) < \tau(\Pi) \wedge Re^{H_m(\Pi)/T} \mid \mathscr{X}_0 = j\ ) \geq 1 - \epsilon/2$ for $j \in \Pi$ and sufficiently small $T > 0$. Now, from the Markov property we get for a fixed $R_0 > 0$ that $P_T(\ \exists n \in \mathbb{N},\ R_0 \leq ne^{-H/T} \leq R_0 + R,\ \mathscr{X}_n = i,\ n \leq \tau(\Pi) \mid \mathscr{X}_0 = i\ )$ is greater than

$$P_T(\ \tau(\Pi) > R_0 e^{H/T} \mid \mathscr{X}_0 = i\ ) \inf_{j \in \Pi} P_T(\ \tau(\Pi \setminus \{i\}) < \tau(\Pi) \wedge Re^{H_m(\Pi)/T} \mid \mathscr{X}_0 = j\ ).$$

Since $P_T(\ \tau(\Pi) > R_0 e^{H/T} \mid \mathscr{X}_0 = i\ ) \geq 1 - \epsilon/2$ for sufficiently small $T > 0$ we get the result. □

Proof (ii). Let Π be the smallest cycle strictly containing Π_i. Let A be the union of all the sub-cycles $\Pi' \in \mathscr{M}_*(\Pi)$ such that $H_e(\Pi') > H_e(\Pi_i)$. From the definition of A, there exists $\eta > 0$ such that for $H > 0$, $H_e(\Pi_i) < H < H_e(\Pi_i) + \eta$ and for any $\Pi' \in \mathscr{M}(A)$ we have $H_e(\Pi') > H$. We have $H_e(\Pi \setminus A) < H < H_e(\Pi)$ so that we deduce from theorem 2 (ii) and lemma 3 that

$$\lim_{T \to 0} P_T(\ \tau(\Pi \setminus A) \leq R_0 e^{H/T} \wedge \tau(\Pi) \mid \mathscr{X}_0 = i\ ) = 1.$$

Moreover, since $H_e(\Pi') > H$ for any $\Pi' \in \mathscr{M}(A)$ we have for $j \in A$

$$\lim_{T \to 0} \inf_{j \in A} P_T(\ \tau(A) > R_1 e^{H/T} \mid \mathscr{X}_0 = j\ ) = 1$$

and the result follows easily. □

3.4 Retrieving communication altitude and exit height

In this part we want to show how to deduce critical quantities of the energy landscape from observation of the annealing process at low temperature. Our first result is deduced from the previous proposition and says that if the process starting from i comes back to i in a time of order $e^{H/T}$ for any $H < H_0$ then i belongs to the bottom of a cycle whose exit height is larger than or equal to H_0.

Proposition 2. *Let $i \in \mathscr{E}$, $H_0 > 0$ and assume that for any $0 < H < H_0$ there exist $R_0 > 0$, $R_1 > 0$ $T_0 > 0$ and $K > 0$ such that for any $0 < T \leq T_0$ we have*

$$P_T(\ \exists n \in \mathbb{N}\ R_0 \leq ne^{-H/T} \leq R_1, \mathscr{X}_n = i \mid \mathscr{X}_0 = i\ ) \geq K. \tag{9}$$

Then we have $H_e(\Pi_i) \geq H_0$.

Proof. This proposition is a direct consequence of proposition 1 (ii). □

Our second result gives an upper bound for the communication altitude from any f in the bottom of $\mathscr{E}$ to any other configuration j.

Proposition 3. *Let $f \in F(\mathscr{E})$ and $j \in \mathscr{E} \backslash \{f\}$. We assume that there exist $T_0 > 0$, $K > 0$, $H_1 \geq 0$, $R > 0$ and $H_2 \geq 0$ such that for any $0 < T \leq T_0$:*

$$P_T(\ \tau(\mathscr{E} \setminus \{j\}) \leq Re^{H_1/T} \mid \mathscr{X}_0 = f\ ) \geq Ke^{-H_2/T}.$$

Then we have $\mathscr{A}_s(f,j) \leq H_1 + H_2$.

Proof. Assume that $\mathscr{A}_s(f,j) > H_1 + H_2$. Then there exists a cycle Π containing f and not containing j such that $H_e(\Pi) > H_1 + H_2$. With lemma 3 we get

$$\begin{aligned} P_T(\ \tau(\mathscr{E} \setminus \{j\}) \leq Re^{H_1/T} \mid \mathscr{X}_0 = f\ ) &\leq P_T(\ \tau(\Pi) \leq Re^{H_1/T} \mid \mathscr{X}_0 = f\ ) \\ &\leq Ke^{(H_1 - H_e(\Pi))/T} \end{aligned}$$

so that we get the result since $H_e(\Pi) - H_1 > H_2$. □

4 Periodically interacting multiple searches

We will set rigorously in this part our main result on the comparison of the convergence speed exponent of interacting and non interacting parallel generalized annealing processes. Coming back to the notation of the introduction, $(Q_T)_{T>0}$ denotes a family of Markov kernel defining a generalized annealing dynamic on a finite configuration space E with communication cost V and virtual energy W. We call $(\mathbf{Z}_n)_{n\in\mathbb{N}}$ the coordinate process on $(E^p)^{\mathbb{N}}$ where p is the number of processors considered. For clarity, the elements of E^p will be written with bold letters. For any cooling schedule $\mathscr{T} = (T_n)_{n\in\mathbb{N}}$, $P^p_{\mathscr{T}}$ is defined as the unique probability measure (for the natural filtration $\sigma(\ \mathbf{Z}_n \mid n \geq 0))$ such that $(\mathbf{Z}_n)_{n\in\mathbb{N}}$ is under $P^p_{\mathscr{T}}$ a Markov chain satisfying

– $P^p_{\mathscr{T}}(\ \mathbf{Z}_{n+1} = \mathbf{j} \mid \mathbf{Z}_n = \mathbf{i}\ ) = R_{T_{n+1}}(\mathbf{i},\mathbf{j})$, where $R_T(\mathbf{i},\mathbf{j})$ is equal to

$$Q^r_T(\mathbf{i}^1,\mathbf{j}^1)\times \tag{10}$$

$$\prod_{k=2}^{p}\left(Q^r_T(\mathbf{i}^k,\mathbf{j}^k)\mathbf{1}_{[W(\mathbf{j}^k)<W(\mathbf{j}^{k-1})]} + \sum_{W(j')\geq W(\mathbf{j}^{k-1})} Q^r_T(\mathbf{i}^k,j')\mathbf{1}_{[\mathbf{j}^k=\mathbf{j}^{k-1}]}\right)$$

$$\text{where } Q^r_T(i,j) = \sum_{i=i_0,i_1,\cdots,i_r=j}\ \prod_{k=0}^{r-1} Q_T(i_k,i_{k+1}). \tag{11}$$

– $P^p_{\mathcal{T}}(\ \mathbf{Z}_0 = \mathbf{i}\ ) = \nu_0(\mathbf{i})$ where ν_0 is an initial probability measure whose support is the whole space E^p.

Notation 2. *We will often omit the superscript p and say $P_{\mathcal{T}}$ instead of $P^p_{\mathcal{T}}$. Moreover, when $T_n = T$ for any $n \geq 0$, P_T will denote $P_{\mathcal{T}}$.*

Notation 3. *For any $A \subset E$, any $u \in \{1, \cdots, p\}$ and any $m \in \mathbb{N}$, we define the stopping time $\tau_u(A, m) = \inf\{\ n > m\ ,\ \mathbf{Z}^u_n \notin A\ \}$. For $m = 0$, we will often write $\tau_u(A)$ instead of $\tau_u(A, 0)$.*

The kernel R_T is not irreducible on E^p. However, if f_0 is a global minimum of W, then writing $\mathbf{f}_0 = (f_0, \cdots, f_0)$, for any $\mathbf{i} \in E^p$ there exists a path $\mathbf{g} = (\mathbf{g}_k)_{0 \leq k \leq n_\mathbf{g}}$ in $\mathrm{Pth}_{E^p}(\mathbf{i}, \mathbf{f}_0)$ such that $R_T(\mathbf{g}_k, \mathbf{g}_{k+1}) > 0$ for any $0 \leq k < n_\mathbf{g}$ and $T > 0$. Therefore, we can introduce the following notation.

Notation 4. *For any $T > 0$, R_T has one and only one irreducible component called Ω. This component contains $\mathbf{f}_0$ and is independent of T. The restriction to Ω of the Markov kernel family $(R_T)_{T>0}$ defines a generalized simulated annealing algorithm as introduced in (7) (here $\mathcal{E} = \Omega$ and $\mathcal{Q}_T = R_T$). We will call $\mathbf{W}$ the associated virtual energy, $\mathbf{A}_s$ the communication altitude function and $C(\Omega)$ the set of all the cycles of Ω.*

4.1 Main result

Notation 5. *Let α_0 be the critical exponent corresponding to the communication cost V as defined by definition 2.*

Our main result is stated as follow.

Theorem 3. *Assume the virtual energy W has only one global minimum. Then for any time period between interactions r, there exist $K > 0$ and $\alpha_{int} \geq 0$ such that $\alpha_0 \leq \alpha_{int} \leq p\alpha_0$ and*

$$\frac{K}{N^{\alpha_{int}}} \leq \sup_{\mathbf{i} \in E^p} \inf_{T_0 \geq \cdots \geq T_N} P_{\mathcal{T}}(\ W(\mathbf{Z}^p_N) > 0 \mid \mathbf{Z}_0 = \mathbf{i}\ ).$$

Remark 1. Since the exponent $p\alpha_0$ can be reached if we consider p independent generalized simulated annealing algorithms under the same cooling schedule and if we consider at each time the configuration with lowest virtual energy value, the theorem says that the periodic interactions between processors underlying the process $(\mathbf{Z}_n)_{n \in \mathbb{N}}$ do not increase the convergence speed exponent. Moreover, we will see in a small example that one can have $\alpha_{int} = \alpha_0$.

Remark 2. The theorem can be extended to various situations where $|F(E)| > 1$. For instance, let $f_0 \in F(E)$ and $b \in E \setminus F(E)$ be such that the critical exponent α_0 satisfies $\alpha_0 = W(b)/H_e(\Pi_b)$. Then its is not difficult to adapt our proof in order to show that if $H_e(b) > \sup_{f \neq f' \in F(E)} A_s(f, f') - W(f)$, the results hold. Moreover, despite many efforts, we have not been able to find a counterexample with $F(E) > 1$. Hence, we suspect that the condition that $|F(E)| = 1$ could be relaxed completely. However, we have not yet a proof of the theorem in this more general situation.

We will start our proof with the case $r = 1$ and the case $r \geq 2$ will be reduced to this one. Our proof for the case $r = 1$ will be conceptually simple. The lower bound $\alpha_0 \leq \alpha_{int}$ is straightforward if we notice that the process $(\mathbf{Z}_n^1)_{n\in\mathbb{N}}$ is a generalized simulated annealing algorithm with dynamic $(Q_T)_{T>0}$. Hence, the interesting part is the upper bound $\alpha_{int} \leq p\alpha_0$. Let us fisrt intruduce the following definition.

Definition 6. *For any $i \in E$, we define the configuration $d(i)$ on the diagonal of Ω by $d(i) = (i, \cdots, i)$. Moreover, for all $\Pi \in \mathscr{C}(E)$ we note $d(\Pi)$ the subset of Ω defined by $d(\Pi) = \Pi^p \cap \Omega$.*

Let $b \in E \setminus F(E)$ such that $\alpha_0 = W(b)/H_e(\Pi_b)$. We assume that $\alpha_0 < +\infty$ otherwise the result is trivial. We will prove that if $\mathbf{b} = d(b)$ and $\alpha_{\boldsymbol{\Pi}_\mathbf{b}} = \mathbf{W}(\mathbf{b})/H_e(\boldsymbol{\Pi}_\mathbf{b})$ then we have $\alpha_{\boldsymbol{\Pi}_\mathbf{b}} \leq p\alpha_0$ so that we will deduce from theorem 1 that

$$\frac{K}{N^{\alpha_{\boldsymbol{\Pi}_\mathbf{b}}}} \leq \sup_{\mathbf{i}\in E^p} \inf_{T_0\geq\cdots\geq T_N} P_{\mathscr{T}}(\ \mathbf{Z}_N \in \boldsymbol{\Pi}_\mathbf{b} \mid \mathbf{Z}_0 = \mathbf{i}\ ).$$

Moreover, we will prove that $\boldsymbol{\Pi}_\mathbf{b} \subset d(\Pi_b)$ so that $\mathbf{Z}_n \in \boldsymbol{\Pi}_\mathbf{b}$ implies that $\mathbf{W}(Z_n^p) > 0$. Hence the result will be proved. The inequality $\alpha_{\boldsymbol{\Pi}_\mathbf{b}} \leq p\alpha_0$ will be established through the two following steps where we will prove first that

$$H_e(\boldsymbol{\Pi}_\mathbf{b}) = H_e(\Pi_b) \tag{12}$$

and then that

$$\mathbf{W}(\mathbf{b}) \leq pW(b). \tag{13}$$

4.2 Proof (r = 1)-Step1

In this first step we will prove the inequality (12). A strictly combinatorial approach seems to us intractable. We prefer to analyze from a probabilistic point of view the behaviour of parallel annealing at constant but low temperature and to deduce 'a posteriori' from this behaviour the shape of the virtual energy as well as the different communication altitudes between configurations.

Proposition 4. *Let f be in $E \setminus F(E)$ and $\mathbf{f} = d(f)$ Then*

$$H_e(\boldsymbol{\Pi}_\mathbf{f}) = H_e(\Pi_f) \text{ and } \boldsymbol{\Pi}_\mathbf{f} \subset d(\Pi_f).$$

Proof. Since $f \in F(\Pi_f)$, for any $H < H_e(\Pi_f)$, any $\epsilon > 0$, we deduce from proposition 1 (i) that there exist $R_0 > 0$, $R_1 > 0$ and $T_0 > 0$ such that for any $0 < T \leq T_0$ we have

$$P_T(\ \exists n \in \mathbb{N}\ R_0 \leq ne^{-H/T} \leq R_1,\ \mathbf{Z}_n^1 = f \mid \mathbf{Z}_0^1 = f\ ) \geq 1 - \epsilon. \tag{14}$$

Moreover, from lemma 6 which is established below we deduce easily using the Markov property the existence of $K > 0$, $T_0 > 0$ such that for any $0 < T \leq T_0$ we have

$$P_T(\ \tau(d(\Pi_f)) \geq e^{H_e(\Pi_f)/T} \mid \mathbf{Z}_0 = \mathbf{f}\ ) \geq K. \tag{15}$$

Since $\mathbf{Z}_n^1 = f$ and $\mathbf{Z}_n \in d(\Pi_f)$ imply that $\mathbf{Z}_n = \mathbf{f}$, it follows from (14) and (15) with $\epsilon \leq K/2$ that $P_T(\ \exists n \in \mathbb{N}\ R_0 \leq ne^{-H/T} \leq R_1,\ \mathbf{Z}_n = \mathbf{f} \mid \mathbf{Z}_0^1 = f\ ) \geq K/2$. Hence from proposition 2, we deduce that $H_e(\boldsymbol{\Pi}_\mathbf{f}) \geq H_e(\Pi_f)$. Moreover, using proposition 1 (ii), we deduce that there exists $\eta > 0$ such that for any $H \geq 0$ satisfying $H_e(\Pi_f) < H < H_e(\Pi_f) + \eta$ and for any $0 < R_0 < R_1$ we have $\lim_{T\to 0} P_T(\ \exists n \in \mathbb{N},\ R_0 \leq ne^{-H/T} \leq R_1,\ \mathbf{Z}_n^1 = f \mid \mathbf{Z}_0 = \mathbf{f}\ )$ so that

$$\lim_{T\to 0} P_T(\ \exists n \in \mathbb{N},\ R_0 \leq ne^{-H/T} \leq R_1,\ \mathbf{Z}_n = \mathbf{f} \mid \mathbf{Z}_0 = \mathbf{f}\ )$$

and $\mathbf{H}_e(\boldsymbol{\Pi}_\mathbf{f}) \leq H_e(\Pi_f)$. This ends the proof of the equality.

Now assume that $\boldsymbol{\Pi}_\mathbf{f} \cap d(\Pi_f)^c \neq \emptyset$ then $|\boldsymbol{\Pi}_\mathbf{f}| > 1$ and $H_e(\boldsymbol{\Pi}_\mathbf{f}) > 0$. Hence, using lemma 4, we deduce that for any $\epsilon > 0$ there exist $R > 0$ and $T_0 > 0$ such that for any $0 < T < T_0$ we have

$$P_T(\ \tau(d(\Pi_f)) < Re^{H_m(\boldsymbol{\Pi}_\mathbf{f})/T} \mid \mathbf{Z}_0 = \mathbf{f}\ ) \geq 1 - \epsilon.$$

Since $H_m(\boldsymbol{\Pi}_\mathbf{f}) < H_e(\boldsymbol{\Pi}_\mathbf{f}) = H_e(\Pi_f)$, we get a contradiction with (15). □

Lemma 6. *Let $\Pi \in \mathscr{C}(E)$. For any $R > 0$ there exists $K > 0$ such that for any $T > 0$ we have*

$$\sup_{\mathbf{i}\in d(\Pi)} P_T(\ \tau(d(\Pi)) \leq Re^{H_m(\Pi)/T} \mid \mathbf{Z}_0 = \mathbf{i}) \leq Ke^{H_m(\Pi)-H_e(\Pi))/T}.$$

Proof. Let $M_k(K) = P_T(\ \tau_u(\Pi) > K\ \forall u \leq k-1,\ \tau_k(\Pi) \leq K \mid \mathbf{Z}_0 = \mathbf{i}\ )$. We have $P_T(\ \tau(d(\Pi)) \leq Re^{H_m(\Pi)/T} \mid \mathbf{Z}_0 = \mathbf{i}\ ) = \sum_{k=1}^{p} M_k(Re^{H_m(\Pi)/T})$ for any $i \in d(\Pi)$. However

$$M_k(K) \leq \sum_{l<K} \sum_{W(j^1)\geq\cdots\geq W(j^k)} P_T(\ \tau_u(\Pi) > l,\ \mathbf{Z}_l^u = j^u\ \forall u \leq k \mid \mathbf{Z}_0 = \mathbf{i}\ )$$
$$\times P_T(\ \tau_k(\Pi, l) = 1 \mid \mathbf{Z}_l^u = j^u \text{ for } 1 \leq u \leq k\ ).$$

Since $P_T(\ \tau_u(\Pi) > l,\ \mathbf{Z}_l^u = j^u\ \forall u \leq k \mid \mathbf{Z}_0 = \mathbf{i}\ )$ is bounded from above by $P_T(\ \tau_1(\Pi) > l,\ \mathbf{Z}_k^1 = j^1 \mid \mathbf{Z}_0 = \mathbf{i}\ )$, and $P_T(\ \tau_k(\Pi) = 1 \mid \mathbf{Z}_0^u = j^u\ \forall u \leq k\ )$ by $K\exp(-(H_e(\Pi) + W(\Pi) - W(j^1))/T)$, we get

$$M_k(Re^{H_m(\Pi)/T}) \leq \sum_{j\in\Pi} \sum_{l<Re^{H_m(\Pi)/T}} KL(\Pi, \Pi)_{i,0}^{j,l} e^{-(H_e(\Pi)+W(\Pi)-W(j))/T}.$$

Since lemma 1 says that $\sum_{l<Re^{H_m(\Pi)/T}} L(\Pi, \Pi)_{i,0}^{j,l} \leq Ke^{(H_m(\Pi)+W(\Pi)-W(j))/T}$, we have proved the result. □

4.3 Proof $(r = 1)$-Step2

Proposition 5. *Assume that there is a unique global minimum f_0 of W on E and that $\alpha_0 < +\infty$. Select any $b \in E \setminus \{f_0\}$ such that $\alpha_0 = W(b)/H_e(\Pi_b)$. Let $\mathbf{b} = d(b)$ and $\mathbf{f}_0 = d(f_0)$. Then we have*

(i) $\mathbf{A}_s(\mathbf{f}_0, \mathbf{b}) \leq A_s(f_0, b) + (p-1)W(b)$.
(ii) $\mathbf{W}(\mathbf{b}) \leq pW(p)$.

Proof (Proposition 5). First, we deduce (ii) from (i). From the definition of b we get $H_e(\Pi_b) = A_s(f_0, b) - W(b)$. Hence from proposition 4 we deduce that $A_s(f_0, b) - W(b) = H_e(\boldsymbol{\Pi}_\mathbf{b}) \leq \mathbf{A}_s(\mathbf{f}_0, \mathbf{b}) - \mathbf{W}(\mathbf{b})$. Hence, using (i) we deduce $A_s(f_0, b) - W(b) + \mathbf{W}(\mathbf{b}) \leq A_s(f_0, b) + (p-1)W(b)$ so that we get (ii).

We start now the proof of (i). Let Π be the smallest cycle in $\mathscr{C}(E)$ containing f_0 and b. We have $\Pi_b \in \mathscr{M}_*(\Pi)$. Assume that we have proved that there exist $T_0 > 0$, $R > 0$, $K > 0$ such that for any $0 < T \leq T_0$,

$$(16) \quad P_T(\tau(\ \Omega \setminus \{\mathbf{b}\}) \leq Re^{H_e(\Pi_b)/T} \mid \mathbf{Z}_0 = \mathbf{f}_0\ ) \geq Ke^{-p(H_m(\Pi) - H_e(\Pi_b))/T}.$$

Then, it follows from proposition 3 that

$$\mathbf{A}_s(\mathbf{f}_0, \mathbf{b}) \leq H_m(\Pi) + (p-1)(H_m(\Pi) - H_e(\Pi_b)) = A_s(f_0, b) + (p-1)W(b),$$

so that part (i) of the proposition can be deduced from inequality (16). To prove (16), let us define for any $0 \leq t \leq p$ the sets B_t and $\widetilde{B}_t$ by

$$B_t = \{\ \mathbf{i} \in d(\Pi) \mid \forall 1 \leq k \leq t\ \mathbf{i}^k = b\ \},\ \widetilde{B}_t = \{\ \mathbf{i} \in \Omega \mid \forall 1 \leq k \leq t\ \mathbf{i}^k = b\ \}.$$

Since $B_0 = d(\Pi)$ and $B_p = \{\mathbf{b}\}$ it will be sufficient to prove that there exist $T_0 > 0$, $R > 0$, $K > 0$ such that for any $0 < T \leq T_0$ and any $n \geq Re^{H_e(\Pi_b)/T}$,

$$\inf_{\mathbf{i} \in B_t} P_T(\tau(\ \Omega \setminus B_{t+1}) \leq n \mid \mathbf{Z}_0 = \mathbf{i}\ ) \geq Ke^{-(H_m(\Pi) - H_e(\Pi_b))/T}.$$

Assume for a while that both of the following lemmas are proved.

Lemma 7. *Let $\mathbf{i}$ be in B_t. There exist $R > 0$, $K > 0$, $T_0 > 0$ such that for any $n \geq Re^{H_e(\Pi_b)/T}$ and any $0 < T \leq T_0$ we have,*

$$P_T(\ \tau_{t+1}(\Pi \setminus \Pi_b) < \inf_{u \leq t} \tau_u(\Pi_b) \wedge \tau_{t+1}(\Pi) \wedge n \mid \mathbf{Z}_0 = \mathbf{i}\ ) \geq Ke^{-(H_m(\Pi) - H_e(\Pi_b))/T}.$$

Lemma 8. *Let $\mathbf{i}$ be in Ω such that $i^u \in \Pi_b$ for any $u \leq t+1$. There exist $T_0 > 0$, $R > 0$ and $K > 0$ such that for any $0 < T \leq T_0$ we have*

$$P_T(\ \tau(\Omega \setminus \widetilde{B}_{t+1}) \leq Re^{H_m(\Pi_b)/T} \mid \mathbf{Z}_0 = \mathbf{i}\ ) \geq 1/2.$$

From lemma 7 and lemma 8 we get that for any $\mathbf{i} \in B_t$ there exist $T_0 > 0$, $R > 0$ and $K > 0$ such that for any $0 < T \leq T_0$ we have

$$P_T(\ \tau(\Omega \setminus \widetilde{B}_{t+1}) \leq Re^{H_e(\Pi_b)/T} \mid \mathbf{Z}_0 = \mathbf{i}) \geq Ke^{-(H_m(\Pi)-H_e(\Pi_b))/T}.$$

However, from the definition of b and Π, we have $W(i) \geq W(b)$ for $i \notin \Pi$ (notice here that the uniqueness of the global minimum is essential) so that

$$P_T(\ \tau(\Omega \setminus \widetilde{B}_{t+1}) = \tau(\Omega \setminus B_{t+1}) \mid \mathbf{Z}_0 = \mathbf{i}) = 1.$$

Hence, the proof of proposition 5 is ended. □

Proof (Lemma 7). Let $\tau_l = \inf\{ n > 0, W(\mathbf{Z}_n^{t+1}) \geq W(\mathbf{Z}_n^t) \}$ be the crossing time between Z^{t+1} and Z^t, and let $\mathscr{G} = (\mathscr{G}_n)_{n\in\mathbb{N}}$ be the filtration defined by

$$\mathscr{G}_n = \sigma(\mathbf{Z}_k^u \mid k < n,\ u \leq t+1\ ) \vee \sigma(\mathbf{Z}_n^u \mid\ u \leq t) \vee \sigma(\mathbf{1}_{W(\mathbf{Z}_n^{t+1}) \geq W(\mathbf{Z}_n^t)}).$$

As usual $\mathscr{G}_{\tau_{\mathscr{I}}} = \{ \mathscr{A} \in \vee_{n=0}^{\infty} \mathscr{G}_n \mid \forall n \in \mathbb{N}\ \mathscr{A} \cap (\tau_l = n) \in \mathscr{G}_n \}$. Now, let $\widetilde{P}_T = P_T^t \otimes P_T^{p-t}$ where P_T^t is the joint distribution of the t first coordinates. From the definition of $\widetilde{P}_T$ we get that for any A in $\mathscr{G}_{\tau_{\mathscr{I}}}$ we have

$$\widetilde{P}_T(A \mid \mathbf{Z}_0 = \mathbf{j}\ ) = P_T(A \mid \mathbf{Z}_0 = \mathbf{j}\ ).$$

We want to prove now that if $A_0 = \big(\tau_{t+1}(\Pi \setminus \Pi_b) < \inf_{u\leq t} \tau_u(\Pi_b) \wedge \tau_{t+1}(\Pi) \wedge n\big)$, then

(17) $$\widetilde{P}_T(A_0 \mid \mathbf{Z}_0 = \mathbf{i}\ ) \leq P_T(A_0 \mid \mathbf{Z}_0 = \mathbf{i}\ ) \text{ for } \mathbf{i} \in B_t.$$

Assume for a while that (17) has been proved. Then we have

$$\begin{gathered}\widetilde{P}_T(A_0 \mid \mathbf{Z}_0 = \mathbf{i}\ ) \geq \\ \widetilde{P}_T(\ \tau_{t+1}(\Pi \setminus \Pi_b) < \tau_{t+1}(\Pi) \wedge n,\ \inf_{u\leq t} \tau_u(\Pi_b) > n \mid \mathbf{Z}_0 = \mathbf{i}\ ) \geq \\ \widetilde{P}_T(\ \tau_{t+1}(\Pi \setminus \Pi_b) < \tau_{t+1}(\Pi) \wedge n \mid \mathbf{Z}_0 = \mathbf{i}\ )\widetilde{P}_T(\ \inf_{u\leq t} \tau_u(\Pi_b) > n \mid \mathbf{Z}_0 = \mathbf{i}\ ).\end{gathered}$$

since the sigma fields $\sigma(\mathbf{Z}_n^{t+1} \mid n \geq 0)$ and $\sigma(\mathbf{Z}_n^u \mid n \geq 0,\ u \leq t)$ are independent under $\widetilde{P}_T$. Since $\widetilde{P}_T(\ \tau_{t+1}(\Pi \setminus \Pi_b) < \tau_{t+1}(\Pi) \wedge n \mid \mathbf{Z}_n^{t+1} = \mathbf{i}^{t+1}\ )$ is equal to $P_T(\tau_1(\Pi \setminus \Pi_b) < \tau_1(\Pi) \wedge n \mid \mathbf{Z}_0^1 = \mathbf{i}^{t+1})$ and since $\mathbf{i}^{t+1} \in \Pi$, we deduce from lemma 5 that there exist $R_1 > 0$, $K_1 > 0$ and $T_1 > 0$ such that for any $n \geq R_1 e^{H_e(\Pi_b)/T}$ and any $0 < T \leq T_1$ we have

$$\widetilde{P}_T(\ \tau_{t+1}(\Pi \setminus \Pi_b) < \tau_{t+1}(\Pi) \wedge n \mid \mathbf{Z}_n^{t+1} = \mathbf{i}^{t+1}\ ) \geq K_1 e^{-(H_m(\Pi)-H_e(\Pi_b))/T}.$$

Moreover, $\widetilde{P}_T(\ \inf_{u\leq t} \tau_u(\Pi_b) > n \mid \mathbf{Z}_0 = \mathbf{i}\ ) = P_T(\ \inf_{u\leq t} \tau_u(\Pi_b) > n \mid \mathbf{Z}_0 = \mathbf{i}\ )$ where $\mathbf{i}^k \in \Pi_b$ for $k \leq t$. Using the Markov property, we deduce from lemma 6 that there exist $K_2 > 0$ and $T_2 > 0$ such that for any $0 < T \leq T_2$ we have

$$P_T(\ \inf_{u\leq t} \tau_u(\Pi_b) > R_1 e^{H_e(\Pi_b)/T} \mid \mathbf{Z}_0 = \mathbf{i}\ ) \geq K_2.$$

Hence, considering $T_3 = T_1 \wedge T_2$ we get that there exists $K > 0$ such that for any $0 < T \leq T_3$, $\widetilde{P}_T(\ \tau_{t+1}(\Pi \setminus \Pi_b) < \tau_{t+1}(\Pi) \wedge \inf_{u\leq t} \tau_u(\Pi_b) \wedge n \mid \mathbf{Z}_n^{t+1} = \mathbf{i}^{t+1}\ )$ is greater than

$$Ke^{(H_e(\Pi_b)-H_m(\Pi))/T}.$$

This last inequality shows that is sufficient to prove inequality (17) to get the lemma. To prove (17) we introduce the stopping time τ_I. We have

$$\widetilde{P}_T(\ \tau_{t+1}(\Pi \setminus \Pi_b) < \inf_{u\leq t} \tau_u(\Pi_b) \wedge \tau_{t+1}(\Pi) \wedge R_1 e^{H_e(\Pi_b)/T} \mid \mathbf{Z}_0 = \mathbf{i}\ ) = M_1 + M_2$$

$$\text{with } \begin{cases} M_1 = \widetilde{P}_T(\ \tau_I \leq \tau_{t+1}(\Pi \setminus \Pi_b) < \inf_{u\leq t} \tau_u(\Pi_b) \wedge n \mid \mathbf{Z}_0 = \mathbf{i}\ ), \\ M_2 = \widetilde{P}_T(\ \tau_{t+1}(\Pi \setminus \Pi_b) < \inf_{u\leq t} \tau_u(\Pi_b) \wedge \tau_{t+1}(\Pi) \wedge \tau_I \wedge n \mid \mathbf{Z}_0 = \mathbf{i}\ ). \end{cases}$$

Since $\left(\tau_{t+1}(\Pi \setminus \Pi_b) < \inf_{u\leq t} \tau_u(\Pi_b) \wedge \tau_{t+1}(\Pi) \wedge \tau_I \wedge n\right) \in \mathcal{G}_{\tau_I}$, we deduce that

$$M_2 = P_T(\ \tau_{t+1}(\Pi \setminus \Pi_b) < \inf_{u\leq t} \tau_u(\Pi_b) \wedge \tau_{t+1}(\Pi) \wedge \tau_I \wedge n \mid \mathbf{Z}_0 = \mathbf{i}\ ).$$

Furthermore, $M_1 \leq \widetilde{P}_T(\ \tau_I \leq \tau_{t+1}(\Pi \setminus \Pi_b),\ \tau_I < \inf_{u\leq t} \tau_u(\Pi_b) \wedge n \mid \mathbf{Z}_0 = \mathbf{i}\ )$ and $\left(\tau_I \leq \tau_{t+1}(\Pi \setminus \Pi_b),\ \tau_I < \inf_{u\leq t} \tau_u(\Pi_b) \wedge n\right) \in \mathcal{G}_{\tau_I}$ so that

$$\begin{aligned} M_1 &\leq P_T(\ \tau_I \leq \tau_{t+1}(\Pi \setminus \Pi_b),\ \tau_I < \inf_{u\leq t} \tau_u(\Pi_b) \wedge n \mid \mathbf{Z}_0 = \mathbf{i}\ ) \\ &= P_T(\ \tau_I = \tau_{t+1}(\Pi \setminus \Pi_b) < \inf_{u\leq t} \tau_u(\Pi_b) \wedge n \mid \mathbf{Z}_0 = \mathbf{i}\ ) \end{aligned}$$

Since $P_T(\ \tau_{t+1}(\Pi \setminus \Pi_b) < \inf_{u\leq t} \tau_u(\Pi_b) \wedge \tau_{t+1}(\Pi) \wedge n \mid \mathbf{Z}_0 = \mathbf{i}\ )$ is equal to

$$P_T(\ \tau_I = \tau_{t+1}(\Pi \setminus \Pi_b) < \inf_{u\leq t} \tau_u(\Pi_b) \wedge n \mid \mathbf{Z}_0 = \mathbf{i}\ )$$

$$+P_T(\ \tau_{t+1}(\Pi \setminus \Pi_b) < \inf_{u\leq t} \tau_u(\Pi_b) \wedge \tau_{t+1}(\Pi) \wedge \tau_I \wedge n \mid \mathbf{Z}_0 = \mathbf{i}\ )$$

the inequality (17) is proved. □

Proof (Lemma 8). From lemma 4, there exist $R > 0$, $T_0 > 0$ such that for any $0 < T \leq T_0$ we have $P_T(\ \tau_1(\Pi_b \setminus \{b\}) < \tau_1(\Pi_b) \wedge Re^{H_m(\Pi_b)/T} \mid \mathbf{Z}_0 = \mathbf{i}\ ) \geq 3/4$. Furthermore, from lemma 6, there exist $T_1 > 0$ and $K > 0$ such that for any $0 < T \leq T_1$ we have

$$P_T(\ \inf_{u\leq t+1} \tau_u(\Pi_b) \leq R_1 e^{H_m(\Pi_b)/T} \mid \mathbf{Z}_0 = \mathbf{i}\ ) \leq Ke^{-(H_e(\Pi_b)-H_m(\Pi_b))/T}.$$

Since $H_e(\Pi_b) > 0$ (otherwise $\alpha_0 = +\infty$), we have $H_e(\Pi_b) - H_m(\Pi_b) > 0$ so that for a $T_2 \leq T_1 \wedge T_0$ satisfying $Ke^{-(H_e(\Pi_b)-H_m(\Pi_b))/T} \leq 1/4$, we get

$$P_T(\ \tau_1(\Pi_b \setminus \{b\}) < \inf_{u\leq t+1} \tau_u(\Pi_b) \wedge Re^{H_m(\Pi_b)/T} \mid \mathbf{Z}_0 = \mathbf{i}\ ) \geq 3/4 - 1/4 = 1/2.$$

Now, since $i^u \in \Pi_b$ for $u \leq t+1$ we get that $(\tau(\Omega \setminus \widetilde{B}_{t+1}) \leq Re^{H_m(\Pi_b)/T},\ \mathbf{Z}_0 = \mathbf{i})$ contains $(\tau_1(\Pi_b \setminus \{b\}) < \inf_{u\leq t+1} \tau_u(\Pi_b) \wedge Re^{H_m(\Pi_b)/T},\ \mathbf{Z}_0 = \mathbf{i})$ so that the lemma is proved. □

4.4 Proof ($r \geq 2$)

We assume now that $r \geq 2$ and we will show that the result can be deduced from the case $r = 1$. Let f_0 be the unique global minimum of W. Since $Q_T(f_0,f_0) > 0$ for any $T > 0$ (otherwise we should have another global minimum), Q_T is aperiodic. Hence Q_T^r is irreducible for any $T > 0$ and satisfies (7) and (8) for the new cost

$$V^r(i,j) = \inf_{g \in \mathrm{Pth}(i,j);n_g=r} V(g); i \neq j \in E.$$

Furthermore, the aperiodicity of Q_T implies also that the virtual energy W^r associated with V^r is just the virtual energy W associated with V. Hence our algorithm with $r \geq 2$ can be considered as parallel annealing based on interacting multiple searches with period between interactions 1 for the new family of kernels $(Q_T^r)_{T>0}$. Applying theorem 3 in the case $r = 1$, we get that $\alpha_{int} \leq p\alpha_0'$ where α_0' is the critical exponent associated with V^r. Finally, since a Markov chain with transition matrices in the family $(Q_T^r)_{T>0}$ can be considered as the restriction to times multiple of r of a Markov chain with transition matrices in the family $(Q_T)_{T>0}$, we deduce easily from theorem 1 that $\alpha_0' \leq \alpha_0$. Hence the theorem is proved.

4.5 A small size example

We show here on a small size example that one may have $\alpha_{int} = \alpha_0$ in theorem 3. Assume that $E = \{a, A, B\}$ and that the communication cost function is defined by the following matrix $V(i,j)$

$V(i,j)$	a	A	B
a	0	0	$+\infty$
A	2	0	4
B	$+\infty$	3	0

From the values $V(i,j)$, one can compute the value of the virtual energy and the cycle decomposition with a combinatorial method given in [14] which will not be reported here. However, in this small size example, the cycle decomposition can easily be achieved "by hand" from the definition of cycles given above. We get $W(A) = 0$, $W(B) = 1$ and $W(a) = 2$ so that A is the unique global minimum of W. Concerning the cycle decomposition, we obtain the very simple diagram below (fig. 1) where the horizontal lines represent cycles placed at height $H_e(\Pi)+W(\Pi)$ along a vertical axis. The value of the critical exponent α_0 for reaching the bottom of E is easily deduced: $\alpha_0 = 1/3$. If we consider now the associated parallel annealing on two processors with a time period between interactions reduced to 1 ($p = 2$, $r = 1$), one have $\Omega = \{aa, aA, AA, aB, BA, BB\}$ and the cycle decomposition given in figure 2. Therefore, in this case, we have $(\mathbf{A}_s(\mathbf{f}_0, \mathbf{b}) - \mathbf{W}(\mathbf{b}))/\mathbf{W}(\mathbf{b}) = 3$ for $\mathbf{f}_0 = AA$ and $\mathbf{b} = BB$ so that $\alpha_{int} = \alpha_0 = 1/3$. This example could be extended to the case $r \geq 2$.

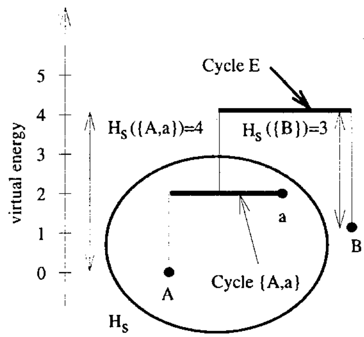

Fig. 1. Decomposition diagram of E

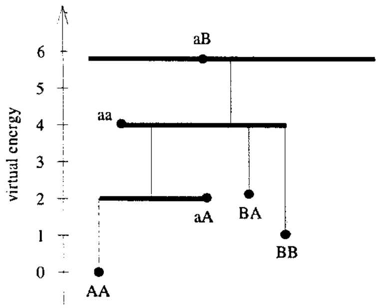

Fig. 2. Decomposition diagram of Ω

5 Conclusion

Coming back to efficiency of the parallel scheme with periodic interactions, we have obtained that the periodic interactions lowered the value of the best convergence speed exponent. In some cases, this exponent is not better than the exponent obtained with a single process. However, from a practical point of view, this is not sufficient to reject definitively this parallel scheme, since we cannot forget the impact of the multiplicative constants appearing in the convergence speed, especially for finite cooling schedules. In our approach, very little is known about these constants and their evaluation still remains an important issue.

We should mention here an other parallel scheme with periodic interactions proposed by Aarts and Laarhoven in [2] and studied experimentally by C. Graffigne in [10] where the processors work at different temperatures. The first processors works with the highest temperature and acts as a rough (or large scale) exploration of the configuration space. The last one uses the lowest temperature and performs an intensive exploration of bottoms of cycles. In this case, the Markov chain is homogeneous in time and in [5], O. Catoni give an estimate of the speed of convergence. According to the experimental study performed in [10], the interactions between processors allow in this modified scheme an interesting acceleration of the sequential convergence speed.

References

1. E. H. L. Aarts, F. M. J. de Bond, J. H. A. Habers, and P. J. M. van Laarhoven. Parallel implementations of statistical cooling algorithms. *Integration VLSI J.*, 4:209–238, 1986.
2. E. H. L. Aarts and P. J. M. van Laarhoven. *Simulated Annealing: Theory and applications*. D. Reidel, Dordrect, 1987.
3. R. Azencott et al., editors. *Simulated annealing: Parallelization techniques*. Willey and Sons, 1992.
4. R. Azencott and C. Graffigne. Parallel annealing by periodically interacting multiple searches: acceleration rates. In R. Azencott et al., editors, *Simulated annealing: Parallelization techniques*, chapter 6, pages 81–90. Willey and Sons, 1992.
5. O. Catoni. Multi-temperature parallelization of simulated annealing. Unpublished paper.

6. O. Catoni. Rough large deviation estimates for simulated annealing. Application to exponential schedules. *Ann. Probab.*, 20:1109–1146, 1992.
7. T.-S. Chiang and Y. Chow. A limit theorem for a class of inhomogeneous Markov processes. *Ann. Probab.*, 1989.
8. J.-D. Deuschel and C. Mazza. L^2 convergence of time non-homogeneous Markov processes: I. spectral estimates. Université de Fribourg, Institut de Mathématiques, preprint (to appear in the Ann. of Appl. Prob.), 1992.
9. M. I. Freidlin and A. D. Wentzell. *Random Perturbations of Dynamical Systems*, volume 260. Springer-Verlag, 1984.
10. C. Graffigne. Parallel annealing by periodically interacting multiple searches: experimental study. In R. Azencott et al., editors, *Simulated annealing: Parallelization techniques*. Willey and Sons, 1992.
11. C.-R. Hwang and S.-J. Sheu. Singular perturbed Markov chains and exact behaviours of simulated annealing process. *J. Theoret. Probab.*, 5(2):223–249, 1992.
12. L. Miclo. *Evolution de l'énergie libre. Applications à l'étude de la convergence des algorithmes du recuit simulé.* PhD thesis, Université d'Orsay, February 1991.
13. A. Trouvé. Optimal convergence rate for generalized simulated annealing. *C. R. Acad. Sci. Paris, Serie I.*, 315:1197–1202, 1992.
14. A. Trouvé. *Parallélisation massive du recuit simulé.* PhD thesis, Université d'Orsay, Jan. 1993.
15. A. Trouvé. Rough large deviation estimates for the optimal convergence speed exponent of generalized simulated annealing algorithms. Rapport de recherche du LMENS (to appear in the Annales de l'I.H.P.), 1994.

This article was processed by the author using the LaTeX style file *pljour1m* from Springer-Verlag.