

## DYNAMIQUES RECUITES DE TYPE FEYNMAN-KAC : RÉSULTATS PRÉCIS ET CONJECTURES

PIERRE DEL MORAL<sup>1</sup> ET LAURENT MICLO<sup>2</sup>

**Résumé.** Soit  $U$  une fonction définie sur un ensemble fini  $E$  muni d'un noyau markovien irréductible  $M$ . L'objectif du papier est de comparer théoriquement deux procédures stochastiques de minimisation globale de  $U$  : le recuit simulé et un algorithme génétique. Pour ceci on se placera dans la situation idéalisée d'une infinité de particules disponibles et nous ferons une hypothèse commode d'existence de suffisamment de symétries du cadre  $(E, M, U)$ . On verra notamment que contrairement au recuit simulé, toute évolution logarithmique de l'inverse de la température conduit asymptotiquement à concentrer les dynamiques de Feynman-Kac recuites sur certains minima globaux de  $U$  (du moins si  $M$  ne fait pas sortir de leur ensemble en un temps presque immédiat). Par contre, ces dernières ont tendance à oublier plus difficilement leur condition initiale et on quantifiera précisément cette propriété. Enfin on présentera les conjectures correspondantes dans une situation plus générale.

**Abstract.** Let  $U$  be a function defined on a finite set  $E$  which is endowed with an irreducible Markov kernel  $M$ . The main purpose of this paper is to theoretically compare two stochastic procedures for the global minimization of  $U$ : simulated annealing and a genetic algorithm. For the latter we simplify the analysis by considering an infinite number of particles. Furthermore, we assume that the setting  $(E, M, U)$  is sufficiently symmetrical and that  $M$  does not force particles to get out of  $\mathcal{U}$  in bounded time, where  $\mathcal{U}$  is the set of global minima of  $U$ . Contrarily to simulated annealing, for all logarithmic inverse temperature evolutions, the considered genetic algorithms concentrate in large time on some particular elements of  $\mathcal{U}$  that we will describe. But it is more difficult for these genetic algorithms to forget their initial condition and we will quantify this property. Corresponding conjectures for more general situations will also be pointed out.

**Classification Mathématique.** 37A30, 46E39, 47D08, 60F10, 60J10, 70K20, 90C27.

Reçu le 23 février 2005. Révisé le 23 août 2005.

### INTRODUCTION

Nous allons ici nous intéresser à des dynamiques non-linéaires à valeurs probabilités de type Feynman-Kac recuit, sur un ensemble fini  $E$  (non vide) et en temps discret. Nous remettons à d'éventuels articles ultérieurs les études correspondantes sur des espaces d'état plus généraux ou en temps continu.

---

*Mots Clés.* Optimisation stochastique, dynamique de Feynman-Kac recuite, recuit simulé généralisé, ergodicité faible, valeur et vecteur propre de Perron-Frobenius, équation de Poisson, temps de retour, probabilité invariante ou fixe, grandes déviations, mesures empiriques, basse température, couplage et inégalités stochastiques, inégalité de Sobolev logarithmique modifiée.

<sup>1</sup> Laboratoire J. A. Dieudonné, UMR 6621, Université de Nice Sophia-Antipolis et CNRS, France ; [delmoral@math.unice.fr](mailto:delmoral@math.unice.fr)

<sup>2</sup> Laboratoire d'Analyse, Topologie, Probabilités, UMR 6632, Université de Provence et CNRS, France ; [miclo@latp.univ-mrs.fr](mailto:miclo@latp.univ-mrs.fr)

Rappelons ce que l'on entend par dynamique recuite de type Feynman-Kac dans le cadre choisi ; elle nécessite la donnée de trois ingrédients de base :

- une matrice  $M = (M(x, y))_{x, y \in E}$  de probabilités de transition sur  $E$  (c'est-à-dire dont les entrées sont positives et qui vérifie pour tout  $x \in E$ ,  $\sum_{y \in E} M(x, y) = 1$ ),
- une fonction  $U = (U(x))_{x \in E}$  définie sur  $E$  à valeurs réelles,
- une suite  $\beta_{\mathbb{N}^*} = (\beta_n)_{n \in \mathbb{N}^*}$  de réels positifs.

Si  $m_0 = (m_0(x))_{x \in E}$  est une probabilité quelconque sur  $E$ , on lui associe alors une suite  $(m_n)_{n \in \mathbb{N}^*}$  de probabilités sur  $E$  construite par la récurrence suivante : pour tout  $n \in \mathbb{N}$ , si  $m_n$  est défini, l'élément suivant  $m_{n+1}$  est donné par

$$\forall x \in E, \quad m_{n+1}(x) := \frac{\sum_{y \in E} m_n(y) M(y, x) \exp(-\beta_{n+1} U(x))}{\sum_{y, z \in E} m_n(y) M(y, z) \exp(-\beta_{n+1} U(z))}. \quad (1)$$

La suite  $(m_n)_{n \in \mathbb{N}}$  est appelée dynamique recuite de Feynman-Kac associée à la condition initiale  $m_0$ , au noyau de mutation (ou d'exploration)  $M$ , au potentiel de sélection  $U$  et au schéma d'inverse de température  $\beta_{\mathbb{N}^*}$ . Cette terminologie provient du fait que ce type de système dynamique à valeurs probabilités sur  $E$  et inhomogène en temps peut s'interpréter comme l'évolution déterministe et limite à une infinité de particules, de l'évolution stochastique des mesures empiriques (spatiales) associées à certains algorithmes génétiques (cf. [9]). Sans vouloir rentrer en détail dans la description de ces derniers, mentionnons simplement que leur évolution est markovienne et alterne des phases de mutation et de sélection, lesquelles se retrouvent dans le passage ci-dessus de  $m_n$  à  $m_{n+1}$ . Plus précisément, l'étape de mutation (relativement au noyau  $M$ ) consiste à transporter une probabilité  $m$  par  $M$ , c'est-à-dire à lui associer la nouvelle probabilité  $mM$  définie par

$$\forall x \in E, \quad mM(x) := \sum_{y \in E} m(y) M(x, y).$$

L'étape de sélection (relativement au potentiel  $U$  et à l'inverse de température  $\beta \geq 0$ ) revient quant à elle à transformer une probabilité  $m$  en la multipliant par une densité proportionnelle à la fonction  $\exp(-\beta U)$ , i.e. à considérer la probabilité  $\Phi^{(\beta)}(m)$  donnée par

$$\forall x \in E, \quad \Phi^{(\beta)}(m)(x) := \frac{\exp(-\beta U(x)) m(x)}{\sum_{y \in E} \exp(-\beta U(y)) m(y)}.$$

À tout instant  $n \in \mathbb{N}$ , la composition de ces deux opérations, d'abord la mutation puis la sélection, à l'inverse de température  $\beta_{n+1}$ , permet de passer de  $m_n$  à  $m_{n+1}$ .

L'une des motivations à l'origine de l'introduction des algorithmes génétiques sous-jacents, surtout quand la température tend à s'annuler en temps grand, est l'optimisation globale de  $\exp(-U)$  : on cherche à trouver (en l'occurrence stochastiquement) les minima globaux de  $U$  (voir par exemple les travaux de Cerf [4, 5] ou [8, 10]). Cependant l'étude de ces processus à nombre fini de particules est beaucoup plus difficile que l'analyse que nous effectuerons ici sur leur limite. Nous espérons toutefois que les techniques et résultats que présenterons aideront d'une certaine manière à la compréhension des algorithmes génétiques à « très grand nombre » de particules.

Si l'on s'attarde un peu sur le mécanisme faisant passer de  $m_n$  à  $m_{n+1}$ , on est effectivement rapidement amené à croire que quand  $\lim_{n \rightarrow \infty} \beta_n = +\infty$ , la probabilité  $m_n$  se concentrera sur les minima globaux de  $U$  en de grands temps  $n \in \mathbb{N}$ . Intuitivement ceci provient du fait que d'une part les mutations ont tendance à étaler les probabilités (envisager par exemple le cas où toutes les entrées de  $M$  sont strictement positives) et d'autre part les sélections ont tendance à reconcentrer les probabilités en les points de leur support où le potentiel  $U$  est minimal. Clairement il existe un conflit entre ces deux tendances, mais comme les  $m_n$ , pour  $n \in \mathbb{N}^*$ , sont obtenues après sélection, on se dit que c'est l'effet de cette dernière qui va l'emporter (la situation aurait été moins favorable si l'on avait inversé l'ordre des opérations : sélection puis mutation).

Cette discussion suggère que l'une des conditions raisonnables à imposer est que le noyau  $M$  permette effectivement de bien visiter  $E$ , c'est-à-dire qu'en utilisant ses transitions, on puisse aller de n'importe quel point de  $E$  à n'importe quel autre. Cette hypothèse s'appelle l'irréductibilité de  $M$  et s'énonce donc rigoureusement par le fait que pour tous  $x, y \in E$ , il existe  $n \in \mathbb{N}^*$  (pouvant dépendre de  $x$  et  $y$ ) et une suite finie  $(x_0, x_1, \dots, x_n)$  d'éléments de  $E$  telle que  $x_0 = x$ ,  $x_n = y$  et pour tous  $0 \leq i < n$ ,  $M(x_i, x_{i+1}) > 0$ . On désignera par le terme de ( $M$ -)chemin allant de  $x$  à  $y$  un tel  $(n+1)$ -uplet et par  $\mathcal{C}(x, y)$  l'ensemble de ces chemins qui ne se recoupent pas (*i.e.* satisfaisant de plus pour tous  $0 \leq i < j \leq n$ ,  $x_i \neq x_j$ , sauf si  $x = y$ , auquel cas cette différence est seulement exigée pour  $0 \leq i < j < n$ ). L'irréductibilité de  $M$  se traduit donc par la condition que pour tous  $x, y \in E$ ,  $\mathcal{C}(x, y) \neq \emptyset$ , que l'on supposera désormais satisfaite.

Une première surprise fut de constater que cette hypothèse est insuffisante pour assurer la concentration voulue, même pour les schémas d'inverse de température les plus convenables que l'on puisse imaginer *a priori*. On en trouvera un contre-exemple dans [10].

Même si dans ce précédent article nous avons donné des conditions sur  $M$  et  $V$  permettant de conclure positivement (pour certains schémas de température, mais dans un cadre d'espace d'état général), il faut bien avouer que nos résultats n'étaient pas très fins. À l'origine du présent manuscrit se trouvait donc la conviction que l'on devait pouvoir faire beaucoup mieux, du moins dans le cas d'ensembles  $E$  finis !

Précisons toutefois que nous n'avons pas réussi à résoudre le cas général. On s'intéressera donc à une situation particulière qui illustre bien les difficultés et propriétés que l'on rencontre lors de ce genre d'étude. Cependant, à chaque fois que nous le pourrons, nous adopterons un point de vue général, en espérant qu'ultérieurement il sera possible d'améliorer la démarche proposée pour aller plus loin. C'est également pourquoi nous présenterons des conjectures sur les comportements que nous pensons être généraux.

Pour décrire l'hypothèse sous laquelle nous avons pu conclure, nous avons besoin de quelques définitions supplémentaires.

Soit  $\emptyset \neq A \subsetneq E$  un sous-ensemble propre de  $E$ , on lui associe la quantité suivante

$$\ln(\lambda(A)) = \lim_{n \rightarrow \infty} \frac{1}{n+1} \ln \left( \sup_{x \in A} \mathbb{P}_x[\tau_{A^c} > n] \right) \quad (2)$$

où  $\tau_{A^c} := \inf\{n \in \mathbb{N} : X_n \notin A\}$  est le premier temps de sortie de  $A$  pour  $(X_n)_{n \in \mathbb{N}}$  qui sous la loi  $\mathbb{P}_x$  est une chaîne de Markov issue de  $x$  (*i.e.*  $X_0 = x$ ,  $\mathbb{P}_x$ -p.s.) et dont les probabilités de transition homogène en temps sont données par la matrice  $M$  (c'est-à-dire qu'à tout instant  $n \in \mathbb{N}$ , une version régulière de loi conditionnelle de  $X_{n+1}$  sachant  $X_n, X_{n-1}, \dots, X_0$  est donnée par la ligne indicée par  $X_n$  de la matrice  $M$ ). On rappellera ultérieurement pourquoi une telle limite existe et on donnera une expression plus analytique de  $\lambda(A) \in ]1, +\infty]$ , qui peut aussi s'interpréter comme la principale valeur propre de Dirichlet de  $A$  relativement à  $M$ . Il est facile de voir que  $\lambda(A)$  ne dépend que de  $M_A := (M(x, y))_{x, y \in A}$ , la restriction de  $M$  à  $A$ , qui est une matrice sous-markovienne (*i.e.* toutes ses entrées sont positives, mais les sommes des lignes sont inférieures à 1). La définition probabiliste a l'avantage de montrer que  $\lambda(A)$  est une mesure de la difficulté que rencontre la chaîne  $(X_n)_{n \in \mathbb{N}}$  pour sortir de  $A$  : plus  $\lambda(A)$  est petit, plus il lui sera facile de sortir de  $A$  (du moins asymptotiquement en de grands temps).

Par ailleurs, munissons  $E$  d'une structure de graphe orienté en disant qu'un point  $x$  conduit à un autre  $y$  si  $M(x, y) > 0$  ou si  $x = y$ . Un sous-ensemble  $A \subset E$  est dit connexe si l'on peut relier tout point de  $A$  à tout autre point de  $A$  par des arêtes orientées restant dans  $A$ . Sauf dans le cas d'un singleton  $A$ , qui par définition est toujours connexe (même si  $M_A = (0)$ ), ceci revient à dire que  $M_A$  est irréductible. Comme d'habitude, un sous-ensemble  $B \subset A$  est appelé une composante connexe de  $A$  s'il est connexe et s'il est maximal (pour l'inclusion) parmi tous les sous-ensembles connexes de  $A$ . Il est facile de voir que  $A$  peut alors s'écrire comme la réunion disjointe de ses composantes connexes.

Ainsi, décomposons l'ensemble des minima globaux de  $U$ ,  $\mathcal{U} := \{x \in E : U(x) = \min_E U\}$ , en ses composantes connexes, disons  $\mathcal{U}_1, \mathcal{U}_2, \dots, \mathcal{U}_r$  pour un certain  $r \in \mathbb{N}^*$ , et on les suppose indicées de telle sorte que

$$\lambda(\mathcal{U}_1) \geq \lambda(\mathcal{U}_2) \geq \dots \geq \lambda(\mathcal{U}_r).$$

Soit  $s = \max\{1 \leq i \leq r : \lambda(\mathcal{U}_i) = \lambda(\mathcal{U}_1)\}$ , on dénote  $\underline{\mathcal{U}} := \mathcal{U}_1 \sqcup \dots \sqcup \mathcal{U}_s$ , « l'ensemble des composantes connexes de minima globaux de  $U$  dont il est difficile de sortir ». On aura aussi besoin de  $\overline{\mathcal{U}}$ , l'ensemble des  $x \in \mathcal{U}$  qui peuvent être facilement atteint à partir de  $\underline{\mathcal{U}}$ , c'est-à-dire tels qu'il existe un chemin  $x_0, \dots, x_n$  allant d'un  $x_0 \in \underline{\mathcal{U}}$  à  $x_n = x$  et tels que pour tout  $0 \leq i \leq n$ ,  $U(x_i) = 0$  (i.e. ce chemin reste dans  $\mathcal{U}$ ).

Enfin, soit  $\mathcal{G}$  le groupe des symétries de notre cadre : il s'agit de l'ensemble des isomorphismes de la structure  $(E, M, U)$ , c'est-à-dire l'ensemble des bijections  $g : E \rightarrow E$  telles que pour tous  $x, y \in E$ , on ait

$$U(g(x)) = U(x) \text{ et } M(g(x), g(y)) = M(x, y).$$

Rappelons que l'orbite d'un point  $x \in E$  sous  $\mathcal{G}$  est l'ensemble  $\mathcal{G}(x) := \{g(x) : g \in \mathcal{G}\}$ .

Nous pouvons maintenant énoncer notre principale hypothèse.

**(H)** : La matrice  $M_{\mathcal{U}_1}$  est irréductible et apériodique : il existe un entier  $n \in \mathbb{N}^*$  tel que la matrice produit  $M_{\mathcal{U}_1}^n$  a toutes ses entrées strictement positives. De plus, il existe  $x_0 \in \mathcal{U}_1$  tel que l'orbite  $\mathcal{G}(x_0)$  intersecte tous les  $\mathcal{U}_i$ , pour  $1 \leq i \leq s$  (i.e.  $\mathcal{G}(x) \cap \mathcal{U}(i) \neq \emptyset$ , pour tout  $2 \leq i \leq s$ ).

Plus prosaïquement, ce dernier point indique que tous les  $\mathcal{U}_i$ , pour  $1 \leq i \leq s$ , sont identiques : le « paysage d'énergie »  $(E, M, U)$  vu de  $x_0$  ou de  $g(x_0)$ , pour  $g \in \mathcal{G}$  est le même. Ceci implique que toutes les matrices  $M_{\mathcal{U}_i}$ , pour  $2 \leq i \leq s$  sont égales à  $M_{\mathcal{U}_1}$ , à une réindexation conjointe près des lignes et des colonnes (notamment elles sont toutes irréductibles et apériodiques). Il apparaît en fait que s'il existe un  $x_0$  comme dans l'hypothèse précédente, alors pour tout  $x \in \underline{\mathcal{U}}$  et tout  $1 \leq i \leq s$ , on a  $\mathcal{G}(x) \cap \mathcal{U}_i \neq \emptyset$ . Cette hypothèse géométrique est évidemment très restrictive, mais il existe une situation où elle est trivialement satisfaite : quand  $s = 1$  (et  $M_{\mathcal{U}_1}$  irréductible apériodique). Le cas le plus simple où (H) est vérifiée avec  $s = 2$  correspond à un ensemble à trois points, disons  $E = \{1, 2, 3\}$ ,

$$M = \begin{pmatrix} a & 1-a & 0 \\ b & 1-2b & b \\ 0 & 1-a & a \end{pmatrix}$$

avec  $0 < a < 1$  ( $a \neq 0$  par « irréductibilité » de  $M_{\{1\}}$  et de  $M_{\{3\}}$ ) et  $0 < b \leq 1/2$ , et  $U(1) = U(3) = 0$ ,  $U(2) > 0$ . Mais l'hypothèse (H) est brisée dès que  $M$  n'admet pas la forme précédente (à une permutation conjointe des lignes et des colonnes, si l'on veut garder  $\text{card}(E) = 3$  et  $s = 2$ ). Il s'agit donc d'une condition extrêmement fragile, très sensible aux variations de  $M$ , même si la matrice d'incidence correspondante  $(\mathbb{1}_{M(x,y)>0})_{x,y \in E}$  n'est pas modifiée (d'ailleurs ceci était déjà vrai pour  $\underline{\mathcal{U}}$ , mais dans une moindre mesure).

Enfin, pour simplifier, on considèrera principalement des schémas d'inverse de température de la forme

$$\forall n \in \mathbb{N}^*, \quad \beta_n := k^{-1} \ln(n) \tag{3}$$

avec  $k > 0$ , car c'est dans cette échelle qu'apparaissent les résultats les plus intéressants, suivant les valeurs prises par  $k$ . On rencontrera souvent des conditions demandant à  $k$  d'être plus grand qu'un certain seuil critique  $k_0 \geq 0$  pour obtenir tel ou tel comportement, lequel continuera souvent à être satisfait sous la seule hypothèse que l'évolution  $\beta_{\mathbb{N}^*}$  est croissante et vérifie  $\sum_{n \in \mathbb{N}^*} \exp(-k_0 \beta_n) = +\infty$ , mais nous n'aborderons pas de telles extensions.

Notre premier résultat affirme que les hypothèses précédentes assurent l'optimisation escomptée :

**Théorème 1.** *Supposons donc que  $M$  soit irréductible, que (H) soit satisfaite et que la suite  $\beta_{\mathbb{N}^*}$  soit de la forme (3) avec  $k > 0$ . Alors pour toute probabilité initiale  $m_0$ , on a*

$$\lim_{n \rightarrow \infty} m_n[\overline{\mathcal{U}}] = 1$$

et en particulier  $\lim_{n \rightarrow \infty} m_n[\mathcal{U}] = 1$ . De plus, tous les points de  $\overline{\mathcal{U}}$  sont effectivement asymptotiquement chargés :

$$\forall x \in \overline{\mathcal{U}}, \quad \liminf_{n \rightarrow \infty} m_n(x) > 0.$$

Cependant, nous pensons que ce résultat n'est qu'un cas particulier :

**Conjecture 2.** Supposons  $M$  irréductible et que  $\lambda(\mathcal{U}) > 0$ , alors pour toute évolution  $\beta_{\mathbb{N}^*}$  croissante et divergente et pour toute loi initiale  $m_0$ , on a

$$\lim_{n \rightarrow \infty} m_n[\overline{\mathcal{U}}] = 1.$$

Par contre, en général  $m_n$  aura tendance pour de grands  $n \in \mathbb{N}$  à se concentrer sur un ensemble plus petit que  $\overline{\mathcal{U}}$  (et même que  $\underline{\mathcal{U}}$ ), c'est-à-dire que l'on pourra avoir pour certains  $x \in \overline{\mathcal{U}}$ ,  $\lim_{n \rightarrow \infty} m_n(x) = 0$ .

On verra que la condition  $\lambda(\mathcal{U}) > 0$  est équivalente à  $\lambda(\mathcal{U}_1) > 0$  (plus précisément, on a toujours  $\lambda(\mathcal{U}) = \max_{1 \leq i \leq r} \lambda(\mathcal{U}_i) = \lambda(\mathcal{U}_1)$ ), qui elle même revient à exiger que  $M_{\mathcal{U}_1}$  est irréductible, ce qui par définition est satisfait sauf éventuellement si  $\mathcal{U}_1$  est un singleton  $\{u_1\}$ , auquel cas il faut que  $M(u_1, u_1) > 0$ . On peut aussi réécrire cette condition  $\lambda(\mathcal{U}) > 0$  en demandant qu'il existe  $x_0 \in \mathcal{U}$  et un chemin « circulaire » de  $\mathcal{C}(x_0, x_0)$  qui reste dans  $\mathcal{U}$ . Sous cette forme, on reconnaît une hypothèse déjà rencontrée dans [10] et le contre-exemple à la convergence  $\lim_{n \rightarrow \infty} m_n[\mathcal{U}] = 1$  fut justement construit en la mettant en défaut sur un ensemble à trois points.

Pour mieux saisir la particularité du théorème 1, il convient de le comparer à son homologue en théorie du recuit simulé, qui est une autre méthode d'optimisation stochastique, dont nous allons maintenant rapidement rappeler le principe. Pour simplifier, outre son irréductibilité, imposons à  $M$  d'être réversible : il existe une probabilité  $\mu$  sur  $E$  telle que

$$\forall x, y \in E, \quad \mu(x)M(x, y) = \mu(y)M(y, x).$$

Il est bien connu que l'irréductibilité implique qu'il ne peut exister qu'une telle probabilité  $\mu$  et que si elle existe, elle charge nécessairement tous les points de  $E$  (i.e. pour tout  $x \in E$ ,  $\mu(x) > 0$ ). Pour tout  $\beta \geq 0$ , introduisons la matrice  $M^{(\beta)}$  de probabilités de transition définie par

$$\forall x, y \in E, \quad M^{(\beta)}(x, y) := \begin{cases} \exp(-\beta(U(y) - U(x))_+)M(x, y), & \text{si } x \neq y \\ 1 - \sum_{z \neq x} \exp(-\beta(U(z) - U(x))_+)M(x, z), & \text{sinon} \end{cases}$$

où rappelons que  $(\cdot)_+$  désigne la partie positive. Plus  $\beta$  est grand, plus  $M^{(\beta)}$  favorise les transitions vers des niveaux de potentiel inférieur, notamment si pour tout  $x \in E$ , on note  $\mathcal{U}(x) = \{x\} \cup \{y \in E : M(x, y) > 0 \text{ et } U(y) \leq U(x)\}$ , alors

$$\lim_{\beta \rightarrow +\infty} \sum_{y \in \mathcal{U}(x)} M^{(\beta)}(x, y) = 1 \tag{4}$$

et cette convergence a lieu exponentiellement rapidement en  $\beta$ .

Comme précédemment, donnons-nous une loi initiale  $p_0$  et un schéma de température  $\beta_{\mathbb{N}^*} = (\beta_n)_{n \in \mathbb{N}^*}$ . On sait construire canoniquement une chaîne de Markov  $Z := (Z_n)_{n \in \mathbb{N}}$  de distribution initiale  $p_0$  (la loi de  $Z_0$  vaut  $p_0$ ) et dont les probabilités de transitions inhomogènes en temps sont données par la famille  $(M^{(\beta_n)})_{n \in \mathbb{N}^*}$  (à tout instant  $n \in \mathbb{N}$ , une version régulière de loi conditionnelle de  $Z_{n+1}$  sachant  $Z_n, Z_{n-1}, \dots, Z_0$  est donnée par la ligne indiquée par  $Z_n$  de la matrice  $M^{(\beta_{n+1})}$ ). La loi de  $Z$  (sur  $E^{\mathbb{N}}$ ) est d'ailleurs uniquement déterminée par ces données. Si  $\lim_{n \rightarrow \infty} \beta_n = +\infty$ , de par (4), la chaîne de Markov  $Z$  peut être vue en temps grand comme une petite perturbation stochastique d'un algorithme discret de « pente descendante » (lui-même un peu aléatoire en général).

Pour tout  $n \in \mathbb{N}$ , notons  $p_n$  la loi de  $Z_n$ . Afin d'effectuer une analogie avec la dynamique de Feynman-Kac recuite, remarquons que le passage de  $p_n$  à  $p_{n+1}$  s'écrit maintenant

$$p_{n+1} = p_n M^{(\beta_{n+1})}.$$

Il s'agit là d'une transformation linéaire (le terme « affine » serait peut-être plus approprié, puisque l'ensemble des probabilités sur  $E$  est seulement convexe), contrairement au mécanisme de passage de  $m_n$  à  $m_{n+1}$  qui de par l'opération de sélection  $\Phi^{(\beta_{n+1})}$  est non-linéaire (il est vrai qu'en apparence cette non-linéarité n'est pas très méchante, puisqu'elle ne fait intervenir qu'un quotient d'expressions linéaires).

De manière plus particulière, considérons un schéma d'inverse de température de la forme (3) pour une certaine constante  $k > 0$ . Dans [17], Hajek a montré qu'il existait une constante  $k_0 \geq 0$  telle que si  $k \geq k_0$ , on est assuré de

$$\lim_{n \rightarrow \infty} p_n[\mathcal{U}] = 1$$

mais que si  $0 < k < k_0$ , alors

$$\limsup_{n \rightarrow \infty} p_n[\mathcal{U}] < 1.$$

De plus, Hajek a décrit explicitement cette constante critique  $k_0$  : pour tous  $x, y \in E$  et  $(x_0, x_1, \dots, x_n) \in \mathcal{C}(x, y)$ , définissons l'élévation (par rapport à  $U$ ) de ce dernier chemin par

$$e_U(x_0, x_1, \dots, x_n) := \max_{0 \leq i \leq n} U(x_i).$$

Ensuite on note  $h_U(x, y)$  la hauteur de communication entre  $x$  et  $y$ , donnée par

$$h_U(x, y) := \min_{(x_0, x_1, \dots, x_n) \in \mathcal{C}(x, y)} e_U(x_0, x_1, \dots, x_n).$$

Enfin on a

$$k_0 := \max_{x \in E} \left( \min_{y \in \mathcal{U}} h_U(x, y) - U(x) \right).$$

Graphiquement,  $k_0$  correspond donc à la hauteur (mesurée sur échelle graduée par  $U$ ) minimale dont il faut s'élever par rapport à sa hauteur de départ pour rejoindre l'ensemble des minima globaux, et ceci dans le cas de la pire position de départ. Ou encore, il s'agit de la plus grande profondeur de puits ne contenant aucun minimum global. Ainsi  $k_0$  est strictement positif si et seulement s'il existe au moins un minimum local qui n'est pas global (rappelons qu'un point  $x \in E$  est appelé un minimum local, sous-entendu relativement à  $U$  et à la structure de graphe orienté sur  $E$  déduite de la matrice d'incidence associée à  $M$ , si pour tout  $y \in E$  satisfaisant  $U(y) < U(x)$ , on a  $h_U(x, y) > U(x)$ ).

Une autre différence entre ces deux procédures réside dans l'observation que la dynamique recuite de Feynman-Kac peut ne pas trouver tous les minima globaux (la conjecture 2 suggère même que l'on ne trouvera jamais les minima globaux de  $\mathcal{U} \setminus \bar{\mathcal{U}}$ , quoique l'on fasse de « raisonnable »). Or l'algorithme de recuit simulé va toujours découvrir tous les minima globaux : Hajek [17] a prouvé que si  $\beta_{\mathbb{N}^*}$  est croissant et divergeant, alors pour tout  $x \in \mathcal{U}$  donné, on ne peut pas avoir  $\lim_{n \rightarrow \infty} p_n(x) = 0$  (pour des évolutions du type (3), on sait montrer plus précisément que  $\lim_{n \rightarrow \infty} p_n(x) > 0$ , mais ce résultat est certainement vérifié pour toute évolution  $\beta_{\mathbb{N}^*}$  comme ci-dessus, bien que nous n'ayons pas trouvé cet énoncé dans la littérature). Insistons également sur le fait que dans ce papier toutes nos affirmations concernant le recuit simulé sont à comprendre en loi et non pas trajectoirellement (ainsi par exemple demander que les trajectoires trouvent tous les minima globaux, c'est-à-dire que

pour tout  $x \in \mathcal{U}$ , p.s.  $\limsup_{n \rightarrow +\infty} \mathbb{1}_{\{x\}}(Z_n) = 1$ , est équivalent pour des évolutions du genre (3) à prendre  $k \geq k_1$ , la constante critique définie ci-dessous en (5), voir par exemple [17]). Car ici ce n'est qu'au niveau des lois qu'une comparaison est possible avec la dynamique recuite de Feynman-Kac, vu que nous n'avons pas introduit les algorithmes stochastiques correspondants. En pratique, ceci signifie que nous sommes en train de parler de résultats concernant un grand nombre de simulations, *i.e.* pour le recuit simulé, de considérer beaucoup de particules évoluant indépendamment comme  $Z$ .

Il ressort de ces considérations qu'en présence de minimum local non global, la dynamique de Feynman-Kac recuite est plus efficace que le recuit simulé pour trouver certains minima globaux, puisqu'il n'y a aucune restriction sur la constante  $k > 0$  apparaissant dans (3). Mais le lecteur attentif aura remarqué que cette affirmation contient une petite escroquerie, car s'il est clair comment en pratique on peut simuler suivant  $p_n$  (il suffit d'utiliser la chaîne  $Z$  jusqu'au temps  $n \in \mathbb{N}$ ), par contre nous avons passé sous silence l'implémentation de l'échantillonnage selon  $m_n$ , lequel doit faire appel aux algorithmes génétiques sous-jacents (succinctement évoqués au début de cette introduction) et se pose alors le problème de la complexité de leur convergence en le nombre de particules (et en le temps  $n$ , voir toutefois [9]). Certes, il existe également une représentation des  $m_n$ , pour  $n \in \mathbb{N}$ , en termes de chaîne de Markov : soit à nouveau  $(X_n)_{n \in \mathbb{N}}$  une chaîne de Markov issue de  $m_0$  et de noyau de transition homogène en temps donné par  $M$ . Alors on a

$$\forall n \in \mathbb{N}, \forall x \in E, \quad m_n(x) := \frac{\mathbb{E}_{m_0} [\exp(-\beta_1 U(X_1) - \beta_2 U(X_2) - \dots - \beta_n U(X_n)) \mathbb{1}_{\{x\}}(X_n)]}{\mathbb{E}_{m_0} [\exp(-\beta_1 U(X_1) - \beta_2 U(X_2) - \dots - \beta_n U(X_n))]}$$

(où l'on a fait la convention usuelle que l'on adoptera partout, d'indiquer en indice de l'espérance la distribution initiale de la chaîne de Markov considérée, on fera aussi un usage systématique du raccourci  $\mathbb{E}_x := \mathbb{E}_{\delta_x}$  comme ce fut déjà le cas en (2)). C'est d'ailleurs ce quotient d'espérances de type Feynman-Kac qui donne son nom aux systèmes dynamiques  $(m_n)_{n \in \mathbb{N}}$  auxquels on s'intéresse dans ce papier. Cependant plusieurs facteurs suggèrent que les « meilleures » approximations d'expressions de ce type reposent justement sur des algorithmes génétiques correspondants (et non pas sur une application de la loi des grands nombres basée sur des échantillons indépendants de la chaîne de Markov  $(X_n)_{n \in \mathbb{N}}$ , voir [9] pour une discussion à ce sujet).

Pour accentuer encore la comparaison entre la dynamique recuite de Feynman-Kac et l'algorithme de recuit simulé classique, rappelons un autre résultat bien connu sur ce dernier, dû à Holley et Stroock [18] (cet article se plaçait dans un contexte de temps continu, mais ses résultats peuvent s'étendre au cas discret). Il concerne l'ergodicité forte des  $p_n$ ,  $n \in \mathbb{N}$ , c'est-à-dire l'oubli en temps grands de la condition initiale  $p_0$ . On peut montrer que du moins pour des schémas de température de la forme (3), avec  $k > 0$ ,  $(p_n)_{n \in \mathbb{N}}$  admet toujours une propriété d'ergodicité faible au sens où  $p_\infty := \lim_{n \rightarrow \infty} p_n$  existe (voir [21] ou [23] pour certaines situations critiques, par exemple si  $k = k_0$ ). Mais il se peut que  $p_\infty$  dépende de  $p_0$  et il existe une constante  $k_1 \geq k_0$  critique pour ce comportement : si  $k \geq k_1$ , on a que  $p_\infty$  est bien déterminé, c'est la renormalisation de la restriction à  $\mathcal{U}$  de la probabilité  $\mu$  réversible pour  $M$ , mais si  $k < k_1$ , l'ensemble des  $p_\infty$ , quand  $p_0$  varie, décrit un ensemble convexe non réduit à un singleton. De plus  $k_1$  peut aussi se décrire « explicitement » :

$$k_1 := \max_{x \in E} \max_{y \in \mathcal{U}} h_U(x, y) - U(x) \quad (5)$$

et il n'est pas difficile de voir que cette expression coïncide aussi avec  $\max_{x \in E} h_U(x, y_0) - U(x)$  pour tout  $y_0 \in \mathcal{U}$  fixé. Graphiquement,  $k_1$  correspond donc à la hauteur minimale dont il faut s'élever par rapport à sa hauteur de départ pour rejoindre un minimum global choisi *a priori* (ou tous les minima globaux), et ceci dans le cas de la pire position de départ. La similitude entre les définitions de  $k_0$  et de  $k_1$  fait que malheureusement, vues de loin, ces deux notions sont souvent confondues (bien sûr dans certains cas on est assuré de l'égalité  $k_0 = k_1$ , par exemple s'il existe un unique minimum global ou un peu plus généralement si  $s = 1$ ). Si l'on ne veut pas mentionner  $p_\infty$  (car nous ne saurons pas prouver que  $\lim_{n \rightarrow \infty} m_n$  existe dans le cadre des dynamiques de Feynman-Kac recuites),  $k_1$  peut aussi se caractériser par la propriété suivante : pour  $k \geq k_1$ , on a pour toutes conditions initiales  $p_0$  et  $p'_0$ ,  $\lim_{n \rightarrow \infty} p_n - p'_n = 0$ , alors que si  $k < k_1$ , on peut trouver deux lois initiales  $p_0$  et  $p'_0$  telles qu'en variation totale,  $\liminf_{n \rightarrow \infty} \|p_n - p'_n\|_{\text{vt}} > 0$ .

Notre second résultat traite de la question correspondante pour les dynamiques recuites de Feynman-Kac, sous la condition supplémentaire que  $\overline{U} = \underline{U}$  (laquelle sera par exemple satisfaite si l'on impose à  $M$  d'être faiblement réversible, au sens où pour tous  $x, y \in E$ ,  $M(x, y) > 0 \Leftrightarrow M(y, x) > 0$ ) et avec une réponse un peu moins précise.

**Théorème 3.** *Supposons  $M$  irréductible, (H) satisfaite, que  $\overline{U} = \underline{U}$  et que le schéma d'inverse de température soit de la forme*

$$\forall n \in \mathbb{N}^*, \quad \beta_n := k^{-1} \ln(A + n)$$

avec  $k > 0$  et  $A \geq 0$ . Il existe une constante critique  $k_2 \geq 0$  telle que : si  $k > k_2$ , on a  $\lim_{n \rightarrow \infty} m_n - m'_n = 0$ , pour tout  $A \geq 0$  et pour toutes conditions initiales  $m_0$  et  $m'_0$ , alors que si  $k < k_2$ , on peut trouver  $A \geq 0$  et deux distributions initiales  $m_0$  et  $m'_0$  tels que  $\liminf_{n \rightarrow \infty} \|m_n - m'_n\|_{\text{vt}} > 0$ .

La constante  $k_2$  admet également une description explicite d'un type similaire à celles de  $k_0$  et  $k_1$ , mais comme elle fait intervenir de nouveaux (quasi-)potentiels, il est préférable de remettre sa définition à la section 4. Indiquons toutefois déjà que  $k_2 = 0$  est équivalent à  $s = 1$  et que si dans (5) le maximum en  $x, y$  est atteint en des éléments de  $\underline{U}$ , alors  $k_2$  sera (souvent bien) plus grand que  $k_1$  (du moins si  $M$  est de plus supposé réversible, puisque nous n'avons envisagé ici le recuit simulé que dans cette situation).

Là encore, nous avons le sentiment qu'un résultat plus général doit être vrai.

**Conjecture 4.** *Supposons  $M$  irréductible,  $\lambda(\mathcal{U}) > 0$  et que le schéma d'inverse de température soit de la forme (3) avec  $k > 0$ . Alors, la limite  $m_\infty := \lim_{n \rightarrow \infty} m_n$  existe toujours et on peut trouver une constante critique  $k_2 \geq 0$  (qui coïncide évidemment avec celle définie dans le théorème 3 si l'hypothèse (H) est satisfaite) telle que si  $k \geq k_2$ , alors  $m_\infty$  ne dépend pas de  $m_0$ , par contre si  $k < k_2$ , l'ensemble des  $m_\infty$ , quand  $m_0$  varie, décrit un ensemble non réduit à un singleton.*

Ainsi pour conclure cette introduction d'une manière un peu informelle, on pourrait dire que la dynamique recuite de Feynman-Kac va trouver plus rapidement des minima globaux de  $U$  que l'algorithme de recuit simulé classique, mais d'une manière moins mélangeante, notamment en pouvant ne pas voir certains minima globaux. Ce manque de stabilité est de mauvaise augure pour l'étude des approximations particulières des dynamiques recuites de Feynman-Kac.

Le plan du papier est le suivant. Dans la section suivante, on va s'intéresser à la probabilité sur  $E$  qui reste fixe par la composition des opérations de mutation et de sélection, notamment on en donnera une représentation probabiliste. Celle-ci nous permettra d'étudier dans la section 2 son comportement à basse température et on sera amené à introduire un quasi-potentiel  $V$  qui jouera un rôle prépondérant dans tout ce manuscrit (le fait que les minima locaux de  $V$  correspondent tous à des minima globaux de  $U$  est une clé importante de toutes les assertions précédentes). Une fois cette étude « statique » effectuée, on pourra s'intéresser à l'aspect dynamique dans la section 3, où l'on se ramènera, par l'application de nouvelles opérations de sélection, presque à un algorithme de recuit simulé généralisé. En étudiant d'une part sa concentration (à l'aide de nouvelles estimations appropriées) et d'autre part ses propriétés d'ergodicité forte (par le biais d'inégalités de Sobolev logarithmiques modifiées), on aboutira respectivement aux théorèmes 1 et 3. Dans la section 4, on s'intéressera à un exemple particulier ne satisfaisant pas l'hypothèse (H), afin d'aller un peu dans la direction des conjectures précédentes. Enfin dans un appendice, on donnera une preuve rapide d'un résultat bien connu qui nous aura été fort utile, du genre principe de Laplace linéaire pour les mesures empiriques associées à des chaînes de Markov finies.

## 1. REPRÉSENTATION DU POINT FIXE

On va s'intéresser dans cette section à la probabilité qui reste fixe par composition des opérations de mutation et de sélection. On se placera uniquement sous l'hypothèse que  $M$  est irréductible et essentiellement la température ne jouera aucun rôle, on pourrait fixer par exemple  $\beta = 1$ .



Commençons par quelques formalités : on note  $\mathcal{P}(E)$  (respectivement  $\mathcal{M}(E)$ ,  $\mathcal{B}(E)$ ) l'ensemble des probabilités sur  $E$  (resp. des mesures signées sur  $E$ , des fonctions réelles définies sur  $E$ ). On a déjà vu qu'un noyau markovien  $M$  agissait sur  $\mathcal{P}(E)$  par le biais de l'application  $\mathcal{P}(E) \ni m \mapsto mM \in \mathcal{P}(E)$  et celle-ci s'étend linéairement sur  $\mathcal{M}(E)$ . On fera aussi agir  $M$  « à droite » sur  $\mathcal{B}(E)$  par  $\mathcal{B}(E) \ni f \mapsto Mf \in \mathcal{B}(E)$ , avec

$$\forall x \in E, \quad Mf(x) := \sum_{y \in E} M(x, y)f(y).$$

Ces deux opérations sont duales, si comme d'habitude on met  $\mathcal{M}(E)$  et  $\mathcal{B}(E)$  en dualité *via* l'intégration :

$$\forall m \in \mathcal{M}(E), \forall f \in \mathcal{B}(E), \quad m[f] := \sum_{x \in E} f(x) m(x).$$

Par ailleurs, pour tout  $\beta \geq 0$ , notons  $T^{(\beta)}$  la composition des opérations de mutation et de sélection à température  $\beta^{-1}$  :

$$T^{(\beta)} : \mathcal{P}(E) \ni m \mapsto \Phi^{(\beta)}(mM) \in \mathcal{P}(E)$$

de sorte qu'à tout instant  $n \in \mathbb{N}$ , (1) se réécrit  $m_{n+1} = T^{(\beta_{n+1})}(m_n)$ .

Afin de se convaincre que pour tout  $\beta \geq 0$  donné, il existe un unique point fixe  $\nu^{(\beta)} \in \mathcal{P}(E)$  pour  $T^{(\beta)}$  (c'est-à-dire vérifiant  $\nu^{(\beta)} = T^{(\beta)}(\nu^{(\beta)})$ ), introduisons l'application linéaire suivante sur  $\mathcal{B}(E)$ ,

$$\begin{aligned} S^{(\beta)} : \mathcal{B}(E) &\rightarrow \mathcal{B}(E) \\ f &\mapsto G^{(\beta)}Mf = (G^{(\beta)}(x)Mf(x))_{x \in E} \end{aligned}$$

où  $G^{(\beta)}$  désignera désormais la fonction  $\exp(-\beta U)$ , pour tout  $\beta \geq 0$ . Dans la base canonique  $(\mathbb{1}_{\{x\}})_{x \in E}$  de  $\mathcal{B}(E)$ , on peut évidemment voir  $S^{(\beta)}$  comme une matrice, et plus précisément, en adoptant la convention habituelle qui représente les éléments de  $\mathcal{B}(E)$  comme des vecteurs colonnes, on a

$$\forall x, y \in E, \quad S^{(\beta)}(x, y) = \exp(-\beta U(x))M(x, y).$$

Les entrées de cette matrice sont positives et elle est « irréductible », le théorème de Perron-Frobenius (voir par exemple [25]) nous assure donc de l'existence de  $\lambda^{(\beta)} > 0$  et d'une fonction  $H^{(\beta)}$  strictement positive (unique à un facteur strictement positif près) tels que

$$S^{(\beta)}H^{(\beta)} = \lambda^{(\beta)}H^{(\beta)}$$

( $\lambda^{(\beta)}$  est en valeur absolue la plus grande valeur propre de  $S^{(\beta)}$  et  $H^{(\beta)}$  est un vecteur propre associé).

Considérons ensuite l'application linéaire suivante sur  $\mathcal{M}(E)$  :

$$\begin{aligned} K^{(\beta)} : \mathcal{M}(E) &\rightarrow \mathcal{M}(E) \\ m &\mapsto mK^{(\beta)} = \left( m \left[ \frac{M[H^{(\beta)}\mathbb{1}_{\{x\}}]}{M[H^{(\beta)}]} \right] \right)_{x \in E} \end{aligned}$$

qui ne dépend pas du choix de la normalisation de  $H^{(\beta)}$ . Dans la base canonique  $(\delta_x)_{x \in E}$  de  $\mathcal{M}(E)$  et en représentant plutôt ses éléments comme des vecteurs lignes, on a

$$\forall x, y \in E, \quad K^{(\beta)}(x, y) = M(x, y)H^{(\beta)}(y)/M[H^{(\beta)}](x). \quad (6)$$

Il apparaît immédiatement que  $K^{(\beta)}$  transforme les probabilités en probabilités, ce qui signifie qu'il s'agit d'un noyau markovien. Il est de plus clair qu'il est irréductible, ce qui admet pour conséquence bien connue qu'il

existe une unique probabilité  $\mu^{(\beta)}$  invariante pour  $K^{(\beta)} : \mu^{(\beta)} = \mu^{(\beta)}K^{(\beta)}$  (ceci découle également du théorème de Perron-Frobenius, notamment  $\mu^{(\beta)}$  charge tous les points de  $E$ ).

On a alors le résultat suivant.

**Proposition 5.** *Sous l'hypothèse que  $M$  est irréductible, pour tout  $\beta \geq 0$ ,  $T^{(\beta)}$  admet un unique point fixe  $\nu^{(\beta)}$  et celui-ci est donné par*

$$\forall x \in E, \quad \nu^{(\beta)}(x) = \frac{(1/MH^{(\beta)}(x))\mu^{(\beta)}(x)}{\mu^{(\beta)}[1/MH^{(\beta)}]}.$$

Désormais, pour indiquer que deux vecteurs sont proportionnels, on utilisera le symbole  $\#$ , ainsi l'égalité précédente s'écrit aussi  $\nu^{(\beta)} \# (1/MH^{(\beta)}) \cdot \mu^{(\beta)}$  (où l'on a aussi fait la convention habituelle que pour  $f \in \mathcal{B}(E)$  et  $m \in \mathcal{M}(E)$ ,  $f \cdot m$  désigne l'élément de  $\mathcal{M}(E)$  qui admet  $f$  pour densité par rapport à  $m$ , le symbole  $\cdot$  permet d'éviter une éventuelle confusion avec le produit matriciel ou le produit de fonctions).

*Démonstration.* Par définition, une probabilité  $\nu$  est un point fixe de  $T^{(\beta)}$  si et seulement si

$$\nu \# \frac{H^{(\beta)}}{MH^{(\beta)}} \cdot (\nu M)$$

ce qui revient aussi à demander que

$$(MH^{(\beta)}) \cdot \nu \# H^{(\beta)} \cdot (\nu M).$$

Posons alors  $\eta := (MH^{(\beta)}) \cdot \nu$ , de sorte que cette équation s'écrive

$$\eta \# H^{(\beta)} \cdot \left( ((1/MH^{(\beta)}) \cdot \eta) M \right)$$

c'est-à-dire

$$\eta \# \eta K^{(\beta)}$$

et on en déduit aisément que  $\eta$  est proportionnel à  $\mu^{(\beta)}$ , la probabilité invariante de  $K^{(\beta)}$ , ce qui conduit au résultat énoncé.  $\square$

**Remarque 6.** Si  $M$  est de plus supposé réversible par rapport à une probabilité  $\mu$ , on peut aller un peu plus loin :

$$\nu^{(\beta)} \# H^{(\beta)} \cdot \mu.$$

En effet, il suffit de voir que

$$\mu^{(\beta)} \# (M[H^{(\beta)}]H^{(\beta)}) \cdot \mu$$

ce qui découle immédiatement du fait que le membre de droite est réversible pour  $K^{(\beta)}$  :

$$\begin{aligned} \forall f, g \in \mathcal{B}(E), \quad ((M[H^{(\beta)}]H^{(\beta)}) \cdot \mu)[fK^{(\beta)}[g]] &= \mu[M[H^{(\beta)}]H^{(\beta)}fK^{(\beta)}[g]] \\ &= \mu[H^{(\beta)}fM[H^{(\beta)}f]] \end{aligned}$$

expression qui est symétrique en  $f, g$ , sous l'hypothèse de réversibilité faite sur  $M$  (on aura noté qu'une probabilité  $\eta$  est réversible par rapport à un noyau  $P$  sur  $E$  si et seulement si pour toutes fonctions  $f, g \in \mathcal{B}(E)$ ,  $\mu[fM[g]] = \mu[gM[f]]$ ).

L'expression de  $\nu^{(\beta)}$  obtenue dans la proposition précédente n'est cependant qu'à moitié satisfaisante, car elle fait appel au théorème « non-constructif » de Perron-Frobenius. On aimerait une formule plus explicite, afin de pouvoir déduire un comportement de type grandes déviations pour  $\nu^{(\beta)}$  quand  $\beta$  est très grand (voir la section suivante). Il est bien connu qu'une telle formulation existe pour  $\mu^{(\beta)}$ , en termes d'arbres couvrants. Rappelons ce résultat dû à Bott et Mayberry [2] (voir aussi le livre de Freidlin et Wentzell [16]). Fixons momentanément  $x \in E$ , on appelle  $x$ -arbre (orienté et recouvrant) un sous-graphe de  $E$  (rappelons qu'il a été naturellement muni d'une structure de graphe orienté dans l'introduction) tel que de tout  $y \neq x$ , il part exactement une arête de  $y$ , aucune arête ne part de  $x$ , mais partant de n'importe quel point  $y \neq x$ , en suivant les arêtes on finit par aboutir en  $x$  (ainsi un tel sous-graphe ne contient pas de cycle et notamment pas de boucle). Un tel  $x$ -arbre  $a$  correspond donc à un certain type d'ensemble d'arêtes et on désigne par  $\mathcal{A}_x$  l'ensemble des  $x$ -arbres. Si  $a \in \mathcal{A}_x$ , on note

$$\pi^{(\beta)}(a) := \prod_{(y,z) \in a} K^{(\beta)}(y,z)$$

puis

$$\pi^{(\beta)}(x) := \sum_{a \in \mathcal{A}_x} \pi^{(\beta)}(a). \quad (7)$$

Alors, pour tout  $\beta \geq 0$ , on a

$$\mu^{(\beta)} \neq \pi^{(\beta)}.$$

Une telle formule semble ne pas exister pour  $H^{(\beta)}$  (ce qui est d'autant plus embêtant que cette fonction entre elle-même dans la définition de  $K^{(\beta)}$ ), mais nous allons maintenant en proposer un ersatz probabiliste.

Considérons  $(X_n)_{n \in \mathbb{N}}$  une chaîne de Markov de noyau de probabilités de transition homogène en temps  $M$ . Fixons  $x_0 \in E$  et désignons par  $\tau_0$  le premier temps de retour en  $x_0$  :

$$\tau_0 := \inf\{n \geq 1 : X_n = x_0\}.$$

On a alors la représentation suivante de  $(\lambda^{(\beta)}, H^{(\beta)})$  :

**Proposition 7.** *Pour tout  $\beta \geq 0$  fixé, la valeur propre  $\lambda^{(\beta)}$  peut être caractérisée comme l'unique réel  $\lambda > 0$  tel que*

$$\mathbb{E}_{x_0} \left[ \prod_{0 \leq n < \tau_0} \frac{G^{(\beta)}(X_n)}{\lambda} \right] = 1. \quad (8)$$

De plus si on impose au vecteur propre  $H^{(\beta)}$  d'être normalisé par la condition  $H^{(\beta)}(x_0) = 1$ , alors on a

$$\forall x \in E, \quad H^{(\beta)}(x) = \mathbb{E}_x \left[ \prod_{0 \leq n < \tau_0} \frac{G^{(\beta)}(X_n)}{\lambda^{(\beta)}} \right].$$

Il est intéressant de noter que la caractérisation de  $\lambda^{(\beta)}$  donnée par (8) ne dépend pas du point  $x_0$  choisi. Elle n'est pas non plus sans rappeler le résultat disant que la probabilité invariante  $\mu^{(\beta)}$  est proportionnelle à

$$\left( \mathbb{E}_{x_0} \left[ \sum_{0 \leq n < \tau_0} \mathbb{1}_{\{x\}}(X_n) \right] \right)_{x \in E}$$

le membre de gauche de (8) pouvant être perçu comme une variante multiplicative de cette formule traditionnelle.

*Démonstration.* On suppose désormais que  $H^{(\beta)}$  a été renormalisé comme ci-dessus. On considère alors le processus  $(M_n)_{n \in \mathbb{N}}$  défini par

$$\forall n \in \mathbb{N}, \quad M_n = \frac{H^{(\beta)}(X_0)}{M[H^{(\beta)}](X_0)} \cdots \frac{H^{(\beta)}(X_{n-1})}{M[H^{(\beta)}](X_{n-1})} H^{(\beta)}(X_n). \quad (9)$$

On voit immédiatement qu'il s'agit d'une martingale strictement positive, quelle que soit la distribution initiale de  $X_0$ . Pour tout  $n \in \mathbb{N}$ , on a donc notamment

$$\begin{aligned} \mathbb{E}_{x_0}[M_{n \wedge \tau_0}] &= \mathbb{E}_{x_0}[M_0] \\ &= \mathbb{E}_{x_0}[H^{(\beta)}(X_0)] \\ &= 1. \end{aligned}$$

Par le lemme de Fatou, si on laisse tendre  $n$  vers l'infini, on obtient, puisque  $\tau_0$  est  $\mathbb{P}_{x_0}$ -p.s. fini,

$$\mathbb{E}_{x_0}[M_{\tau_0}] \leq 1.$$

Mais montrons que l'on a là une véritable égalité. Pour ceci, soit  $(Y_n)_{n \in \mathbb{N}}$  une chaîne de Markov issue de  $x_0$  (*i.e.*  $Y_0 = x_0$  p.s.) et de noyau de transition homogène  $K^{(\beta)}$ . Par définition de ce dernier, on obtient facilement que pour tout  $n \in \mathbb{N}$  et tout événement  $A \subset E^{\{0, \dots, n\}}$ , une formule de changement de probabilité de type Girsanov discret :

$$\begin{aligned} \mathbb{E}_{x_0}[\mathbb{1}_A((Y_p)_{0 \leq p \leq n})] &= \mathbb{E}_{x_0} \left[ \frac{H^{(\beta)}(X_1)}{M[H^{(\beta)}](X_0)} \cdots \frac{H^{(\beta)}(X_n)}{M[H^{(\beta)}](X_{n-1})} \mathbb{1}_A((X_p)_{0 \leq p \leq n}) \right] \\ &= \mathbb{E}_{x_0}[M_n \mathbb{1}_A((X_p)_{0 \leq p \leq n})] \end{aligned}$$

(car  $H^{(\beta)}(x_0) = 1$ ). Ainsi, notant

$$\tau_0(Y) := \inf\{n \geq 1 : Y_n = x_0\}$$

il apparaît que pour tout  $n \in \mathbb{N}$ ,

$$\mathbb{E}_{x_0}[M_n \mathbb{1}_{\{\tau_0 > n\}}] = \mathbb{P}_{x_0}[\tau_0(Y) > n].$$

Or l'irréductibilité de  $K^{(\beta)}$  assure que le membre de droite converge (exponentiellement rapidement) vers 0 quand  $n$  tend vers l'infini, d'où

$$\lim_{n \rightarrow \infty} \mathbb{E}_{x_0}[M_n \mathbb{1}_{\{\tau_0 > n\}}] = 0.$$

Par ailleurs, on a pour tout  $n \in \mathbb{N}$ ,

$$\begin{aligned} 1 &= \mathbb{E}_{x_0}[M_{n \wedge \tau_0}] \\ &= \mathbb{E}_{x_0}[M_{\tau_0} \mathbb{1}_{\{\tau_0 \leq n\}}] + \mathbb{E}_{x_0}[M_n \mathbb{1}_{\{\tau_0 > n\}}] \end{aligned}$$

et on en déduit l'égalité voulue (en invoquant un théorème de convergence monotone) :

$$\begin{aligned} 1 &= \lim_{n \rightarrow \infty} \mathbb{E}_{x_0}[M_{\tau_0} \mathbb{1}_{\{\tau_0 \leq n\}}] \\ &= \mathbb{E}_{x_0}[M_{\tau_0}]. \end{aligned} \quad (10)$$

En utilisant ensuite que  $H^{(\beta)}$  est solution de « l'équation de Poisson multiplicative »

$$\frac{H^{(\beta)}}{M[H^{(\beta)}]} = \frac{G^{(\beta)}}{\lambda^{(\beta)}} \quad (11)$$

on obtient que

$$\begin{aligned} \mathbb{E}_{x_0} \left[ \prod_{0 \leq n < \tau_0} \frac{G^{(\beta)}(X_n)}{\lambda^{(\beta)}} \right] &= \mathbb{E}_{x_0} \left[ \left( \prod_{0 \leq n < \tau_0} \frac{G^{(\beta)}(X_n)}{\lambda^{(\beta)}} \right) H^{(\beta)}(X_{\tau_0}) \right] \\ &= \mathbb{E}_{x_0} [M_{\tau_0}] \\ &= 1. \end{aligned}$$

Cependant, si l'on pose  $\Lambda := \{\lambda > 0 : \mathbb{E}_{x_0} [\prod_{0 \leq n < \tau_0} (G^{(\beta)}(X_n)/\lambda)] < +\infty\}$ , il est clair que sur ce segment, l'application

$$\Lambda \ni \lambda \mapsto \mathbb{E}_{x_0} \left[ \prod_{0 \leq n < \tau_0} \frac{G^{(\beta)}(X_n)}{\lambda} \right]$$

est strictement décroissante, ainsi l'équation (8) caractérise  $\lambda^{(\beta)}$ .

Considérons maintenant  $\tilde{H}^{(\beta)}$  la fonction définie sur  $E$  par

$$\forall x \in E, \quad \tilde{H}^{(\beta)}(x) := \mathbb{E}_x \left[ \prod_{0 \leq n < \tau_0} \frac{G^{(\beta)}(X_n)}{\lambda^{(\beta)}} \right].$$

Commençons par remarquer que les valeurs de cette fonction sont bien finies : pour  $x \in E \setminus \{x_0\}$ , soit  $(x_0, x_1, \dots, x_p) \in \mathcal{C}(x_0, x)$  un chemin ne se recoupant pas allant de  $x_0$  à  $x$  (notamment pour tout  $1 \leq i \leq p$ ,  $x_i \neq x_0$ ). En tenant compte de la propriété de Markov, on peut alors écrire

$$\begin{aligned} 1 &= \mathbb{E}_{x_0} \left[ \prod_{0 \leq n < \tau_0} \frac{G^{(\beta)}(X_n)}{\lambda^{(\beta)}} \right] \\ &\geq \left( \prod_{0 \leq n < p} \frac{G^{(\beta)}(x_n)}{\lambda^{(\beta)}} \right) \mathbb{E}_{x_0} \left[ \mathbb{1}_{\{X_0=x_0, X_1=x_1, \dots, X_p=x_p\}} \prod_{p \leq n < \tau_0} \frac{G^{(\beta)}(X_n)}{\lambda^{(\beta)}} \right] \\ &= \left( \prod_{0 \leq n < p} \frac{G^{(\beta)}(x_n)}{\lambda^{(\beta)}} \right) \mathbb{P}_{x_0}[X_0 = x_0, X_1 = x_1, \dots, X_p = x_p] \mathbb{E}_x \left[ \prod_{0 \leq n < \tau_0} \frac{G^{(\beta)}(X_n)}{\lambda^{(\beta)}} \right] \\ &= \left( \prod_{0 \leq n < p} \frac{G^{(\beta)}(x_n) M(x_n, x_{n+1})}{\lambda^{(\beta)}} \right) \tilde{H}^{(\beta)}(x) \end{aligned}$$

d'où il ressort que  $\tilde{H}^{(\beta)}(x) < +\infty$ . Cependant, pour  $x \neq x_0$ , appliquons la propriété de Markov après une transition :

$$\begin{aligned}\tilde{H}^{(\beta)}(x) &= \frac{G^{(\beta)}(x)}{\lambda^{(\beta)}} \mathbb{E}_x \left[ \prod_{1 \leq n < \tau_0} \frac{G^{(\beta)}(X_n)}{\lambda^{(\beta)}} \right] \\ &= \frac{G^{(\beta)}(x)}{\lambda^{(\beta)}} M[\tilde{H}^{(\beta)}](x)\end{aligned}$$

ainsi  $\tilde{H}^{(\beta)}/M[\tilde{H}^{(\beta)}] = G^{(\beta)}/\lambda^{(\beta)}$  sur  $E \setminus \{x_0\}$ . En  $x_0$ , la même application de la propriété de Markov nous donne plutôt

$$\begin{aligned}\tilde{H}^{(\beta)}(x_0) &= \frac{G^{(\beta)}(x_0)}{\lambda^{(\beta)}} \left( \mathbb{E}_{x_0} [\mathbb{1}_{\{X_1=x_0\}}] + \mathbb{E}_{x_0} [\mathbb{1}_{\{X_1 \neq x_0\}} \tilde{H}^{(\beta)}(X_1)] \right) \\ &= \frac{G^{(\beta)}(x_0)}{\lambda^{(\beta)}} M[\tilde{H}^{(\beta)}](x_0)\end{aligned}$$

où l'on a tenu compte du fait que  $\tilde{H}^{(\beta)}(x_0) = 1$ , d'après (8). En définitive,  $\tilde{H}^{(\beta)}$  vérifie donc (11) partout, *i.e.* est un vecteur propre associé à  $\lambda^{(\beta)}$ , et par unicité (puisque  $\tilde{H}^{(\beta)}(x_0) = 1 = H^{(\beta)}(x_0)$ ), il en découle que comme annoncé,  $H^{(\beta)} = \tilde{H}^{(\beta)}$ .  $\square$

Il existe d'autres caractérisations de la première valeur propre  $\lambda^{(\beta)}$ , présentons maintenant celle bien connue liée à l'étude des grandes déviations satisfaites par les mesures empiriques associées à la chaîne de Markov  $(X_n)_{n \in \mathbb{N}}$ . Pour ceci, rappelons que l'entropie relative de deux probabilités  $m$  et  $\mu$  sur  $E$  est la quantité positive ou infinie définie par

$$\text{Ent}(m|\mu) := \sum_{x \in E} \ln \left( \frac{m(x)}{\mu(x)} \right) m(x) \quad (12)$$

(où l'on a fait la convention habituelle que  $0 \cdot \infty = 0$ ). Pour voir que  $\text{Ent}(m|\mu) \in \bar{\mathbb{R}}_+$ , il vaut mieux réécrire cette expression sous la forme  $\sum_{x \in E} m(x)/\mu(x) \ln(m(x)/\mu(x)) \mu(x)$ , du moins si  $m \ll \mu$  (sinon remarquons que  $\text{Ent}(m|\mu) = +\infty$ ), puis utiliser l'inégalité de Jensen par rapport à la fonction convexe  $\mathbb{R}_+ \ni t \mapsto t \ln(t) \in \mathbb{R}$ .

Si  $\eta \in \mathcal{P}(E)$  est donnée, on désigne par  $\mathcal{K}(\eta)$  l'ensemble des noyaux de probabilités de transition  $P$  sur  $E$  qui admettent  $\eta$  comme probabilité invariante, c'est-à-dire satisfaisant  $\eta = \eta P$ . On définit ensuite une fonctionnelle  $\mathcal{L}$  sur  $\mathcal{P}(E)$  en posant

$$\forall \eta \in \mathcal{P}(E), \quad \mathcal{L}(\eta) := \inf_{P \in \mathcal{K}(\eta)} \sum_{x \in E} \eta(x) \text{Ent}(P(x, \cdot) | M(x, \cdot)). \quad (13)$$

Un résultat célèbre (voir par exemple le livre de Dupuis et Ellis [15]) affirme alors que pour tout  $\beta \geq 0$ , on a

$$\ln(\lambda^{(\beta)}) = - \inf_{\eta \in \mathcal{P}(E)} \beta \eta(U) + \mathcal{L}(\eta). \quad (14)$$

On donnera une preuve très simple de ce résultat dans l'appendice, basée justement sur des arguments similaires à ceux utilisés dans la démonstration précédente.

Pour finir cette section dans le même registre d'idées, remarquons que la martingale introduite en (9) permet d'obtenir d'autres expressions pour  $H^{(\beta)}$ . Supposons par exemple que l'on connaisse *a priori* cette fonction sur un ensemble  $A \subset E$ , disons contenant  $x_0$ . En dénotant

$$\tau := \inf\{n \geq 0 : X_n \in A\} \leq \tau_0$$

on est assuré de

$$\forall x \in E, \quad H^{(\beta)}(x) = \mathbb{E}_x \left[ \left( \prod_{0 \leq n < \tau} \frac{G^{(\beta)}(X_n)}{\lambda^{(\beta)}} \right) H^{(\beta)}(X_\tau) \right]. \quad (15)$$

Cette identité est en effet équivalente à l'assertion

$$\forall x \in E, \quad \mathbb{E}_x[M_\tau] = \mathbb{E}_x[M_0]$$

laquelle admet la même démonstration que (10).

L'égalité (15) sera très utile dans la section suivante pour étudier le comportement à basse température de  $H^{(\beta)}$ , en prenant  $A = \underline{\mathcal{U}}$ .

## 2. COMPORTEMENT À BASSE TEMPÉRATURE

Notre but ici est d'étudier des comportements de type « grandes déviations » à basse température des objets introduits dans la section précédente. Pour ce faire, outre la condition omniprésente d'irréductibilité, nous aurons besoin de l'hypothèse (H). Cependant, remarquons que l'addition d'une constante  $c \in \mathbb{R}$  à  $U$  ne modifie pas vraiment les notions précédentes (mis à part  $\lambda^{(\beta)}$ , qui doit être remplacé par  $\lambda^{(\beta)} \exp(-\beta c)$ ). Ainsi nous ferons également la convention que  $\min_E U = 0$ , de sorte que  $\mathcal{U} = \{x \in E : U(x) = 0\}$ , ce qui simplifiera grandement les écritures à venir.

Commençons donc par nous intéresser à la valeur propre  $\lambda^{(\beta)}$ . Son comportement à basse température n'est pas difficile à obtenir, mais on va déjà y voir poindre l'importance des  $\lambda(\mathcal{U}_i)$  pour  $1 \leq i \leq r$ , ce qui conduira ensuite naturellement à (H).

**Lemme 8.** *Sous l'hypothèse  $\min_E U = 0$ , la limite suivante existe*

$$\lambda^{(\infty)} := \lim_{\beta \rightarrow +\infty} \lambda^{(\beta)}$$

et elle s'exprime aussi comme

$$\lambda^{(\infty)} = \exp \left( - \inf_{\eta \in \mathcal{P}(\mathcal{U})} \mathcal{L}(\eta) \right)$$

(où  $\mathcal{P}(\mathcal{U})$  doit être interprété comme l'ensemble des probabilités  $\eta \in \mathcal{P}(E)$  dont le support est inclus dans  $\mathcal{U}$ ).

*Démonstration.* Rappelons que pour tout  $\beta \geq 0$ , on a

$$\ln(\lambda^{(\beta)}) = - \inf_{\eta \in \mathcal{P}(E)} \beta \eta(U) + \mathcal{L}(\eta). \quad (16)$$

Sur cette expression, il apparaît que  $\lambda^{(\beta)}$  est une quantité décroissante en  $\beta \geq 0$ , elle converge donc pour  $\beta$  grand vers une certaine limite, disons  $\tilde{\lambda}^{(\infty)}$ . *A priori* on a  $\tilde{\lambda}^{(\infty)} \geq \lambda^{(\infty)}$ , car pour tout  $\beta \geq 0$ , on a

$$\inf_{\eta \in \mathcal{P}(E)} \beta \eta(U) + \mathcal{L}(\eta) \leq \inf_{\eta \in \mathcal{P}(\mathcal{U})} \mathcal{L}(\eta).$$

Pour obtenir l'inégalité dans l'autre sens, commençons par considérer le cas où le terme de droite est fini et utilisons que l'on sait que l'infimum dans (16) est atteint. Plus précisément, on montre dans l'appendice que

pour tout  $\beta \geq 0$ ,

$$\begin{aligned} \ln(\lambda^{(\beta)}) &= - \left( \beta \mu^{(\beta)}[U] + \mathcal{L}(\mu^{(\beta)}) \right) \\ &= - \left( \beta \mu^{(\beta)}[U] + \sum_{x \in E} \mu^{(\beta)}(x) \text{Ent}(K^{(\beta)}(x, \cdot) | M(x, \cdot)) \right) \end{aligned}$$

où  $\mu^{(\beta)}$  est la probabilité invariante du noyau  $K^{(\beta)}$  introduit au début de la section précédente. Cependant, l'ensemble des couples  $(\eta, P)$ , formés d'une probabilité  $\eta$  sur  $E$  et d'un noyau  $K$  qui la laisse invariante, est compact (vu par exemple comme un sous-ensemble de  $[0, 1]^{E \sqcup E^2}$ ). On peut donc extraire de la famille  $(\mu^{(\beta)}, K^{(\beta)})_{\beta \geq 0}$  une sous-suite convergente de la forme  $(\mu^{(\beta_n)}, K^{(\beta_n)})_{n \in \mathbb{N}}$ , où  $(\beta_n)_{n \in \mathbb{N}}$  est une suite croissante et divergente de réels positifs. Notons  $(\tilde{\mu}^{(\infty)}, \tilde{K}^{(\infty)})$  sa limite. Il apparaît très simplement (rappelons que l'entropie est une fonctionnelle continue en chacune de ses variables, l'autre étant fixée) que

$$\begin{aligned} \lim_{n \rightarrow \infty} \mathcal{L}(\mu^{(\beta_n)}) &= \lim_{n \rightarrow \infty} \sum_{x \in E} \mu^{(\beta_n)}(x) \text{Ent}(K^{(\beta_n)}(x, \cdot) | M(x, \cdot)) \\ &= \sum_{x \in E} \tilde{\mu}^{(\infty)}(x) \text{Ent}(\tilde{K}^{(\infty)}(x, \cdot) | M(x, \cdot)) \\ &\geq \mathcal{L}(\tilde{\mu}^{(\infty)}) \end{aligned}$$

(on a aussi utilisé dans ce calcul le fait que pour tout  $x \in E$  et tout  $n \in \mathbb{N}$ ,  $K^{(\beta_n)}(x, \cdot) \ll M(x, \cdot)$  et que  $\sup_{\eta \in \mathcal{P}(E), \eta \ll M(x, \cdot)} \text{Ent}(\eta | M(x, \cdot)) < +\infty$ , ce qui implique notamment que pour tout  $x \in E$ ,  $\text{Ent}(\tilde{K}^{(\infty)}(x, \cdot) | M(x, \cdot)) < +\infty$ ). Plus généralement, cet argument permet de voir que la fonctionnelle  $\mathcal{L}$  est s.c.i. sur  $\mathcal{P}(E)$ , ce qui est un résultat bien connu en théorie des grandes déviations (*cf.* par exemple [14, 15]).

Par ailleurs, on a

$$\begin{aligned} +\infty &> \inf_{\eta \in \mathcal{P}(U)} \mathcal{L}(\eta) \\ &\geq \lim_{\beta \rightarrow +\infty} \inf_{\eta \in \mathcal{P}(E)} \beta \eta[U] + \mathcal{L}(\eta) \\ &= \lim_{\beta \rightarrow +\infty} \beta \mu^{(\beta)}[U] + \mathcal{L}(\mu^{(\beta)}) \\ &\geq \limsup_{\beta \rightarrow +\infty} \beta \mu^{(\beta)}[U] \\ &\geq \limsup_{n \rightarrow \infty} \beta_n \mu^{(\beta_n)}[U] \end{aligned}$$

d'où il ressort que nécessairement  $\tilde{\mu}^{(\infty)}$  se concentre sur  $U$ . On en déduit que

$$\begin{aligned} \ln(\tilde{\lambda}^{(\infty)}) &= - \lim_{\beta \rightarrow +\infty} \beta \mu^{(\beta)}[U] + \mathcal{L}(\mu^{(\beta)}) \\ &\leq - \limsup_{\beta \rightarrow +\infty} \mathcal{L}(\mu^{(\beta)}) \\ &\leq - \lim_{n \rightarrow \infty} \mathcal{L}(\mu^{(\beta_n)}) \\ &\leq - \mathcal{L}(\tilde{\mu}^{(\infty)}) \\ &\leq - \inf_{\eta \in \mathcal{P}(U)} \mathcal{L}(\eta) \\ &= \ln(\lambda^{(\infty)}). \end{aligned}$$



Venons-en maintenant au cas où  $\inf_{\eta \in \mathcal{P}(\mathcal{U})} \mathcal{L}(\eta) = +\infty$  (cette situation se rencontre par exemple s'il n'existe aucune arête du graphe orienté sous-jacent allant de  $\mathcal{U}$  vers  $\mathcal{U}$ , car alors pour tout  $\mu \in \mathcal{P}(\mathcal{U})$  et tout  $x \in \mathcal{U}$ , il n'existe pas d'élément  $P$  de  $\mathcal{K}(\mu)$  tel que  $P(x, \cdot)$  soit absolument continu par rapport à  $M(x, \cdot)$ ). Le début des arguments précédents reste valable, sauf que l'on ne peut plus conclure que  $\tilde{\mu}^{(\infty)}(\mathcal{U}) = 1$ , car il est possible que  $\limsup_{n \rightarrow \infty} \beta_n \mu^{(\beta_n)}[U] = +\infty$ . Mais dans ce cas on obtient directement que

$$\lim_{\beta \rightarrow +\infty} \ln(\lambda^{(\beta)}) = -\infty$$

d'où finalement le résultat annoncé.  $\square$

Pour effectuer un lien avec la quantité  $\lambda(\mathcal{U})$  introduite dans l'introduction, munissons tout d'abord  $E$  d'une structure de graphe non-orienté en traçant une arête entre  $x, y \in E$  si et seulement si, soit  $x = y$ , soit  $M(x, y) > 0$ , soit  $M(y, x) > 0$  (il s'agit de la symétrisation de la structure de graphe orienté présentée dans l'introduction, il n'y aura aucune différence essentielle entre ces deux notions si  $M$  est faiblement réversible). Faisons ensuite une observation générale :

**Proposition 9.** *Soit  $\emptyset \neq A \subsetneq E$ , la constante  $\ln(\lambda(A))$  définie en (2) peut aussi s'écrire sous la forme*

$$\ln(\lambda(A)) = - \inf_{\eta \in \mathcal{P}(A)} \mathcal{L}(\eta). \quad (17)$$

De plus, si  $A = A'_1 \sqcup \dots \sqcup A'_{R'}$  est la décomposition de  $A$  en ses composantes connexes pour la structure de graphe non orientée alors

$$\lambda(A) = \max_{1 \leq i \leq R'} \lambda(A'_i). \quad (18)$$

Un peu plus précisément, si pour  $x \in A$  on note  $A(x)$  les éléments de  $A$  que l'on peut atteindre à partir de  $x$  sans sortir de  $A$  (i.e. en suivant cette fois des arêtes du graphe orienté), on a

$$\lim_{n \rightarrow \infty} \frac{1}{n+1} \ln(\mathbb{P}_x[\tau_{A^c} > n]) = \ln(\lambda(A(x))). \quad (19)$$

Traditionnellement, la convergence (2) s'obtient par le biais d'arguments de sous-additivité,  $1/\lambda(A)$  pouvant se réinterpréter comme le rayon spectral de l'opérateur sous-markovien  $\mathbb{1}_A M[\mathbb{1}_A \cdot]$  agissant sur l'espace de Banach obtenu en munissant  $\mathcal{B}(E)$  de la norme uniforme (de manière équivalente,  $1/\lambda(A)$  est le rayon spectral de la matrice  $M_A$ ). Mais nous n'adopterons pas ce point de vue.

*Démonstration.* Formellement, (17) correspond juste à (16) avec  $\beta = 1$  et  $U = (+\infty)\mathbb{1}_{A^c}$ . Pour justifier l'utilisation de ce potentiel prenant la valeur  $+\infty$ , on l'approche pour  $u \geq 0$  grand, par les potentiels  $U_u = u\mathbb{1}_{A^c}$ . Soit  $m_0$  la probabilité équidistribuée sur  $A$ , pour tout  $n \in \mathbb{N}$ , on a

$$\mathbb{P}_{m_0}[\tau_{A^c} > n] = \lim_{u \rightarrow +\infty} \mathbb{E}_{m_0}[\exp(-U_u(X_0) - \dots - U_u(X_n))]$$

(où  $(X_p)_{p \geq 0}$  est une chaîne de Markov homogène de matrice de transition  $M$ ). Or si  $\eta \in \mathcal{P}(A)$  et si  $P \in \mathcal{K}(\eta)$ , on a d'après l'appendice

$$\begin{aligned} & \ln(\mathbb{E}_{m_0}[\exp(-U_u(X_0) - \dots - U_u(X_n))]) \\ & \geq -\text{Ent}(\eta|m_0) - n\text{Ent}(\eta \times P|\eta \times M) - \mathbb{E}_\eta[U_u(Y_0) + \dots + U_u(Y_n)] \\ & = -\text{Ent}(\eta|m_0) - n\text{Ent}(\eta \times P|\eta \times M) \end{aligned}$$

car  $(Y_p)_{p \geq 0}$  est une chaîne de Markov homogène de matrice de transition  $P$  et de loi initiale  $\eta$ , elle reste donc tout le temps dans  $A$ . Si l'on fait tendre  $u$  vers  $+\infty$ , on obtient que

$$\ln(\mathbb{P}_{m_0}[\tau_{A^c} > n]) \geq -\text{Ent}(\eta|m_0) - n\text{Ent}(\eta \times P|\eta \times M)$$

reste alors à diviser par  $n + 1$ , que l'on fait tendre vers  $\infty$ , puis à considérer le supremum sur  $\eta \in \mathcal{P}(A)$  et  $P \in \mathcal{K}(\eta)$  pour se convaincre que

$$\liminf_{n \rightarrow \infty} \frac{1}{n+1} \ln(\mathbb{P}_{m_0}[\tau_{A^c} > n]) \geq - \inf_{\eta \in \mathcal{P}(A)} \mathcal{L}(\eta).$$

Cependant, pour tout  $n \in \mathbb{N}$  et tout  $u \geq 0$ , on a

$$\mathbb{P}_{m_0}[\tau_{A^c} > n] \leq \mathbb{E}_{m_0}[\exp(-U_u(X_0) - \dots - U_u(X_n))].$$

Ainsi en prenant le logarithme de cette inégalité, en divisant par  $n + 1$  que l'on laisse filer à l'infini, il ressort que

$$\limsup_{n \rightarrow \infty} \frac{1}{n+1} \ln(\mathbb{P}_{m_0}[\tau_{A^c} > n]) \leq - \inf_{\eta \in \mathcal{P}(E)} \eta[U_u] + \mathcal{L}(\eta).$$

En utilisant les arguments de la preuve du lemme 8 et en faisant tendre  $u$  vers  $+\infty$ , on obtient

$$\limsup_{n \rightarrow \infty} \frac{1}{n+1} \ln(\mathbb{P}_{m_0}[\tau_{A^c} > n]) \leq - \inf_{\eta \in \mathcal{P}(A)} \mathcal{L}(\eta)$$

puis que le terme de gauche est une véritable limite. Remarquons maintenant que pour tout  $n \in \mathbb{N}$ , on a

$$\begin{aligned} \max_{x \in A} \mathbb{P}_x[\tau_{A^c} > n] / \text{card}(A) &\leq \mathbb{P}_{m_0}[\tau_{A^c} > n] \\ &\leq \max_{x \in A} \mathbb{P}_x[\tau_{A^c} > n] \end{aligned}$$

ce qui permet de conclure à la validité de (17).

Montrons ensuite la dernière convergence présentée dans la proposition. Soit  $x \in A$  fixé, par définition de  $A(x)$ , on a  $\tau_{A^c} = \tau_{(A(x))^c}$  d'où

$$\begin{aligned} \limsup_{n \rightarrow \infty} \frac{1}{n+1} \ln(\mathbb{P}_x[\tau_{A^c} > n]) &= \limsup_{n \rightarrow \infty} \frac{1}{n+1} \ln(\mathbb{P}_x[\tau_{(A(x))^c} > n]) \\ &\leq \limsup_{n \rightarrow \infty} \frac{1}{n+1} \ln\left(\max_{y \in A(x)} \mathbb{P}_y[\tau_{(A(x))^c} > n]\right) \\ &= \ln(\lambda(A(x))) \end{aligned}$$

Cependant, soit  $y \in A(x)$  et  $(x_0, \dots, x_p)$  un chemin de  $\mathcal{C}(x, y)$  qui reste dans  $A$  et donc dans  $A(x)$ , on a pour tout  $n \geq p$ , en utilisant la propriété de Markov,

$$\begin{aligned} \mathbb{P}_x[\tau_{(A(x))^c} > n] &\geq \mathbb{P}_x[X_0 = x_0, \dots, X_p = x_p, \tau_{(A(x))^c} > n] \\ &= M(x_0, x_1) \cdots M(x_{p-1}, x_p) \mathbb{P}_y[\tau_{(A(x))^c} > n - p] \\ &\geq a(y) \mathbb{P}_y[\tau_{(A(x))^c} > n] \end{aligned}$$

avec par définition  $a(y) := M(x_0, x_1) \cdots M(x_{p-1}, x_p)$ . Ainsi en posant  $a := \min_{y \in A(x)} a(y) > 0$  on a notamment,

$$\mathbb{P}_x[\tau_{(A(x))^c} > n] \geq a \max_{y \in A(x)} \mathbb{P}_y[\tau_{(A(x))^c} > n].$$

En prenant le logarithme et en divisant par  $n + 1$  que l'on laisse aller à l'infini, on fait apparaître que

$$\begin{aligned} \liminf_{n \rightarrow \infty} \frac{1}{n+1} \ln(\mathbb{P}_x[\tau_{A^c} > n]) &\geq \lim_{n \rightarrow \infty} \frac{1}{n+1} \ln \left( \max_{y \in A(x)} \mathbb{P}_y[\tau_{A^c} > n] \right) \\ &= \ln(A(x)) \end{aligned}$$

d'où en fin de compte la convergence voulue.

Intéressons-nous à l'égalité (18). Pour cela remarquons que l'application  $\lambda$  définie sur  $\{A : \emptyset \neq A \subsetneq E\}$  à valeurs dans  $\overline{\mathbb{R}}_+$  est croissante :

$$\forall \emptyset \neq A \subset B \subsetneq E, \quad \lambda(A) \leq \lambda(B) \quad (20)$$

comme on le déduit immédiatement de (17). Ainsi on est *a priori* assuré de

$$\lambda(A) \geq \max_{1 \leq i \leq R'} \lambda(A'_i).$$

Par ailleurs, notons que pour tout  $x \in A$ , l'ensemble  $A(x)$  est inclus dans l'un des  $A'_i$ , ce qui permet d'obtenir que  $\lambda(A(x)) \leq \max_{1 \leq i \leq R'} \lambda(A'_i)$ . Reste alors à utiliser les égalités

$$\begin{aligned} \ln(\lambda(A)) &= \lim_{n \rightarrow \infty} \frac{1}{n+1} \max_{x \in A} \ln(\mathbb{P}_x[\tau_{A^c} > n]) \\ &= \max_{x \in A} \lim_{n \rightarrow \infty} \frac{1}{n+1} \ln(\mathbb{P}_x[\tau_{A^c} > n]) \\ &= \max_{x \in A} \ln(\lambda(A(x))) \end{aligned}$$

pour se convaincre de la validité de (18). □

Il est possible d'aller un peu plus loin dans l'analyse de  $\lambda(A)$  :

**Proposition 10.** *Soit  $\emptyset \neq A \subsetneq E$  et  $A_1, \dots, A_R$  ses composantes connexes pour la structure de graphe orientée. On a alors*

$$\lambda(A) = \max_{1 \leq i \leq R} \lambda(A_i).$$

D'après la formule (18), il suffirait de s'intéresser au cas où  $A$  est connexe pour la structure de graphe non-orientée.

*Démonstration.* Par croissance de l'application  $\lambda$  (vue en (20)), on a déjà clairement

$$\lambda(A) \geq \max_{1 \leq i \leq R} \lambda(A_i).$$

Pour se convaincre de l'inégalité inverse, commençons par donner une caractérisation alternative de  $\lambda(A)$  :

$$\lambda(A) = \inf\{l > 0 : \sup_{x \in A} \mathbb{E}_x[\exp(\tau_{A^c}/l)] < +\infty\}.$$

En effet, il s'agit là d'une conséquence immédiate de la définition de cette quantité et de la formule

$$\mathbb{E}[\exp(\tau)] = 1 + \int_0^{+\infty} \mathbb{P}[\tau > u] \exp(u) du$$

valable pour toute variable aléatoire positive  $\tau$ .

Fixons  $x \in A$  et construisons une suite  $(S_n)_{n \in \mathbb{N}}$  de temps d'arrêt de la manière suivante :  $S_0 = 0$  puis par récurrence pour tout  $n \in \mathbb{N}$ ,

$$S_{n+1} = \begin{cases} \tau_{A^c} & , \text{ si } S_n = \tau_{A^c} \\ \inf\{p > S_n : Y(X_p) \neq Y(X_{S_n})\} & , \text{ si } S_n < \tau_{A^c} \end{cases}$$

où  $Y : A \rightarrow \{1, \dots, R\}$  est l'application qui a tout  $y \in A$  associe l'indice  $i \in \{1, \dots, R\}$  tel que  $y \in A_i$ . Définissons également la variable aléatoire  $N = \inf\{n \in \mathbb{N} : S_n = \tau_{A^c}\}$ . Il est clair ( $\mathbb{P}_x$ -p.s.) que si la chaîne de Markov  $(X_n)_{n \in \mathbb{N}}$  sort de l'un des  $A_i$ , avec  $1 \leq i \leq R$  donné, avant  $\tau_{A^c}$ , alors elle ne pourra pas y retourner avant  $\tau_{A^c}$ . On en déduit que  $1 \leq N \leq R$ . Supposons maintenant par l'absurde que l'on ait

$$\lambda(A) > \max_{1 \leq i \leq R} \lambda(A_i)$$

et soit  $\lambda(A) > l > \max_{1 \leq i \leq R} \lambda(A_i)$ . On est alors assuré que

$$K := \max_{1 \leq i \leq R} \max_{y \in A_i} \mathbb{E}_y[\exp(\tau_{A_i^c}/l)] \leq +\infty.$$

En écrivant que  $\tau_{A^c} = \sum_{0 \leq p < R} S_{p+1} - S_p$  et en appliquant successivement la propriété de Markov forte en les  $S_{R-1}, S_{R-2}, \dots, S_1$ , on obtient que

$$\mathbb{E}_x[\exp(\tau_{A^c}/l)] \leq K^R < +\infty$$

ce qui est en contradiction avec le fait que  $l < \lambda(A)$ . D'où finalement la validité de l'égalité recherchée.  $\square$

Il est également possible de prouver le résultat précédent (ainsi que l'égalité (18)) d'une manière plus « analytique », directement à partir de (17), de la convexité de la fonctionnelle  $\mathcal{L}$  et en étudiant les classes d'irréductibilité des noyaux  $K \in \mathcal{K}(\eta)$  avec  $\eta \in \mathcal{P}(A)$  et tels que pour tout  $x \in A$ ,  $K(x, \cdot) \ll M(x, \cdot)$ .

La proposition 10 permet de calculer  $\lambda(\mathcal{U})$ , puisqu'avec les notations de l'introduction, on a

$$\begin{aligned} \lambda(\mathcal{U}) &= \max_{1 \leq i \leq r} \lambda(\mathcal{U}_i) \\ &= \lambda(\mathcal{U}_1). \end{aligned}$$

Ainsi le lemme 8 se réécrit également sous la forme

$$\lim_{\beta \rightarrow +\infty} \lambda^{(\beta)} = \lambda(\mathcal{U}_1) \quad (21)$$

et cette quantité est strictement positive si l'on fait l'hypothèse (H) (ou plus précisément en supposant seulement vérifiée sa première partie).

Il faut maintenant étudier  $H^{(\beta)}$  pour  $\beta \geq 0$  grand. C'est le gros morceau, car ensuite tout en découle relativement facilement (ceci est tout particulièrement clair dans le cas où  $M$  est réversible, comme le montre la remarque 6). Mais avant tout, il faut convenir d'une normalisation pour ce vecteur propre, qui *a priori* n'est défini qu'à une constante multiplicative strictement positive près. Fixons donc un minimum global de  $\underline{\mathcal{U}}$ , disons  $x_0 \in \mathcal{U}_1$ , et imposons désormais que  $H^{(\beta)}(x_0) = 1$ , comme nous l'avons déjà fait dans la proposition 7. Pour décrire le comportement de  $H^{(\beta)}$  à basse température, introduisons la fonction  $V$  suivante, qui jouera par la suite un grand rôle :

$$\forall x \in E, \quad V(x) := \min_{p \in \mathcal{C}(x, \underline{\mathcal{U}})} s_U(p) \quad (22)$$

avec  $\mathcal{C}(x, \underline{U}) := \sqcup_{y \in \underline{U}} \mathcal{C}(x, y)$ , l'ensemble des chemins allant de  $x$  à  $\underline{U}$  sans se recouper et où pour tout  $p = (p_0, p_1, \dots, p_n) \in \mathcal{C}(x, \underline{U})$ , on a noté

$$s_U(p) := \sum_{0 \leq i \leq n} U(p_i) \quad (23)$$

(quantité aussi égale à  $U(x) + \sum_{1 \leq i \leq n-1} U(p_i)$ , puisque  $p_0 = x$  et  $p_n \in \underline{U}$ , ainsi  $U(p_n) = 0$ ).

Observons que pour tout  $x \in E$ ,  $V(x) \geq 0$  et que  $V$  ne peut s'annuler qu'en des points de  $\underline{U}$ . Ainsi tous les minima locaux de  $V$  sont des minima globaux de  $U$ . De plus, on se rend rapidement compte que  $V$  n'a pas de minimum local qui ne soit pas global. Le lecteur est encouragé à faire un dessin pour voir comment  $V$  « lisse » les minima locaux non globaux de  $U$  et même éventuellement certains minima globaux de  $\mathcal{U} \setminus \underline{U}$ , par exemple ceux de  $\mathcal{U}_i$ , avec  $s < i \leq r$ , si les voisins de  $\mathcal{U}_i$  n'appartiennent pas à  $\underline{U}$ . Remarquons que si  $V$  s'annule bien sur  $\underline{U}$ , par contre la réciproque n'est pas vraie en toute généralité : il peut exister des éléments  $x \in \mathcal{U} \setminus \underline{U}$  avec  $V(x) = 0$ , comme le montre l'exemple suivant sur notre espace à trois points favori,  $E = \{1, 2, 3\}$ , avec

$$M = \begin{pmatrix} 1/2 & 1/2 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} \quad \text{et} \quad U = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}.$$

Dans ce cas particulier, on a  $\mathcal{U} = \{1, 3\}$ ,  $\mathcal{U}_1 = \{1\}$ ,  $\mathcal{U}_2 = \{3\}$ ,  $\underline{U} = \mathcal{U}_1$ , (car  $\lambda(\{1\}) = 2 < \lambda(\{3\}) = +\infty$ ) et on calcule que  $V = U$  (par contre, si on part avec  $U(1) = U(2) = 0$  et  $U(3) = 1$ , alors  $V$  est le vecteur transposé de  $(0, 1, 1)$ ).

Plus généralement,  $\mathcal{V} := \{x \in E : V(x) = 0\}$  consiste en les éléments de  $\mathcal{U}$  à partir desquels on peut rejoindre  $\underline{U}$  en suivant des arêtes orientées et sans sortir de  $\mathcal{U}$ . En particulier si  $M$  est faiblement réversible on est assuré de l'égalité  $\mathcal{V} = \underline{U}$ .

Le principal objectif de cette section s'énonce alors

**Proposition 11.** *Supposons  $M$  irréductible et l'hypothèse (H) vérifiée, alors pour tout  $x \in E$ , on a*

$$\lim_{\beta \rightarrow +\infty} \beta^{-1} \ln(H^{(\beta)}(x)) = -V(x).$$

**Remarque 12.** Si de plus  $M$  est supposé réversible, la remarque 6 permet de transférer directement ce comportement sur  $\nu^{(\beta)}$  :

$$\forall x \in E, \quad \lim_{\beta \rightarrow +\infty} \frac{1}{\beta} \ln(\nu^{(\beta)}(x)) = -V(x).$$

Ainsi, d'après l'observation faite avant la proposition précédente, on voit que  $\nu^{(\beta)}$  va se concentrer (exponentiellement vite) à basse température sur  $\underline{U}$ . Ceci sera encore vrai si  $M$  est faiblement réversible (toujours sous l'hypothèse (H)), mais il faudra travailler un peu plus pour l'obtenir ! Par contre, c'est sur  $\overline{U}$  que se concentrera plus généralement  $\nu^{(\beta)}$ .

Malgré la simplicité de l'assertion de la proposition 11, sa preuve va être relativement longue, demandant plusieurs étapes que nous avons décomposées en autant de lemmes.

**Lemme 13.** *Pour tout  $x \in \mathcal{U}_1$ , on a*

$$\lim_{\beta \rightarrow +\infty} \beta^{-1} \ln(H^{(\beta)}(x)) = 0.$$

*Démonstration.* Si  $\mathcal{U}_1$  est formé d'un singleton, ceci est trivial puisqu'alors  $\mathcal{U}_1 = \{x_0\}$ . Sinon par définition,  $M_{\mathcal{U}_1}$  est irréductible, fixons donc  $x \in \mathcal{U}_1$  et  $(x_0, \dots, x_l) \in \mathcal{C}(x_0, x)$  et  $(y_0, \dots, y_k) \in \mathcal{C}(x, x_0)$  deux chemins ne sortant pas de  $\mathcal{U}_1$ . D'après la proposition 7, on a

$$\begin{aligned} H^{(\beta)}(x) &= \mathbb{E}_x \left[ \prod_{0 \leq n < \tau_0} \frac{G^{(\beta)}(X_n)}{\lambda^{(\beta)}} \right] \\ &\geq \prod_{0 \leq n < k} M(y_n, y_{n+1}) \frac{G^{(\beta)}(y_n)}{\lambda^{(\beta)}} \\ &= \left( \prod_{0 \leq n < k} M(y_n, y_{n+1}) \right) (\lambda^{(\beta)})^{-k}. \end{aligned}$$

Or l'irréductibilité de  $M_{\mathcal{U}_1}$  implique également que  $\lambda(\mathcal{U}_1) > 0$ , d'où

$$\lim_{\beta \rightarrow +\infty} \frac{1}{\beta} \ln(\lambda^{(\beta)}) = 0$$

et on en déduit que

$$\liminf_{\beta \rightarrow +\infty} \frac{1}{\beta} \ln(H^{(\beta)}(x)) \geq 0.$$

Par ailleurs, en tenant compte de la propriété de Markov, on a

$$\begin{aligned} 1 &= H^{(\beta)}(x_0) \\ &= \mathbb{E}_{x_0} \left[ \prod_{0 \leq n < \tau_0} \frac{G^{(\beta)}(X_n)}{\lambda^{(\beta)}} \right] \\ &\geq \mathbb{E}_{x_0} \left[ \mathbb{1}_{\{X_1=x_1, \dots, X_l=x_l\}} \prod_{0 \leq n < \tau_0} \frac{G^{(\beta)}(X_n)}{\lambda^{(\beta)}} \right] \\ &= (\lambda^{(\beta)})^{-l} \mathbb{E}_{x_0} \left[ \mathbb{1}_{\{X_1=x_1, \dots, X_l=x_l\}} \mathbb{E}_{X_l} \left[ \prod_{0 \leq n < \tau_0} \frac{G^{(\beta)}(X_n)}{\lambda^{(\beta)}} \right] \right] \\ &= (\lambda^{(\beta)})^{-l} \mathbb{E}_{x_0} \left[ \mathbb{1}_{\{X_1=x_1, \dots, X_l=x_l\}} H^{(\beta)}(x) \right] \\ &= (\lambda^{(\beta)})^{-l} \left( \prod_{0 \leq n < l} M(x_n, x_{n+1}) \right) H^{(\beta)}(x) \end{aligned}$$

d'où

$$\limsup_{\beta \rightarrow +\infty} \frac{1}{\beta} \ln(H^{(\beta)}(x)) \leq 0$$

puis finalement le résultat annoncé.  $\square$

La preuve du résultat suivant est la principale raison d'être de la condition (H).

**Lemme 14.** *Supposons (H) satisfait, alors pour tout  $x \in \underline{\mathcal{U}}$ , on a*

$$\lim_{\beta \rightarrow +\infty} \beta^{-1} \ln(H^{(\beta)}(x)) = 0.$$

*Démonstration.* Soit  $g \in \mathcal{G}$  un isomorphisme de notre cadre. Notons  $\tilde{H}^{(\beta)}$  l'application composée  $H^{(\beta)} \circ g$ . Par invariance de la structure  $(E, M, U)$ ,  $\tilde{H}^{(\beta)}$  reste un vecteur propre relativement à la valeur propre  $\lambda^{(\beta)}$  de l'opérateur  $S^{(\beta)}$  défini au début de la section 2. En effet, pour tous  $x, y \in E$ , on a

$$S^{(\beta)}(g(x), g(y)) = S^{(\beta)}(x, y)$$

ainsi

$$\begin{aligned} \lambda^{(\beta)} \tilde{H}^{(\beta)}(x) &= \lambda^{(\beta)} H^{(\beta)}(g(x)) \\ &= S^{(\beta)}[H^{(\beta)}](g(x)) \\ &= \sum_{y \in E} S^{(\beta)}(g(x), y) H^{(\beta)}(y) \\ &= \sum_{y \in E} S^{(\beta)}(g(x), g(y)) H^{(\beta)}(g(y)) \\ &= \sum_{y \in E} S^{(\beta)}(x, y) \tilde{H}^{(\beta)}(y) \\ &= S^{(\beta)}[\tilde{H}^{(\beta)}](x). \end{aligned}$$

Or par irréductibilité de  $S^{(\beta)}$ , le théorème de Perron-Frobenius nous assure que l'espace propre associé à  $\lambda^{(\beta)}$  est de dimension 1, ainsi il existe  $a_\beta = \tilde{H}^{(\beta)}(x_0) = H^{(\beta)}(g(x_0)) > 0$  tel que  $\tilde{H}^{(\beta)} = a_\beta H^{(\beta)}$ . Cependant comme  $g$  est une bijection de  $E$ , on a  $\max_{x \in E} \tilde{H}^{(\beta)}(x) = \max_{x \in E} H^{(\beta)}(x)$ , d'où  $a_\beta = 1$ . Il en découle notamment que  $H^{(\beta)}(g(x_0)) = 1$ , et par conséquent que pour tout  $x \in \mathcal{G}(x_0)$ ,  $H^{(\beta)}(x) = 1$ . Ainsi l'hypothèse (H) nous permet d'obtenir que pour tout  $2 \leq i \leq s$ , il existe  $x_i \in \mathcal{U}_i$  tel que  $H^{(\beta)}(x_i) = 1$  pour tout  $\beta \geq 0$ , et donc notamment

$$\lim_{\beta \rightarrow +\infty} \beta^{-1} \ln(H^{(\beta)}(x_i)) = 0.$$

Pour conclure au résultat annoncé, on peut utiliser le lemme 13 de deux manières différentes :

- soit en remarquant que pour tout  $2 \leq i \leq s$  et tout  $x \in \mathcal{U}_i$  il existe  $y \in \mathcal{U}_1$  et  $g \in \mathcal{G}$  tels que  $x = g(y)$ ,
- soit en procédant, pour  $2 \leq i \leq s$  donné, à partir de  $x_i$  pour obtenir que pour tout  $x \in \mathcal{U}_i$ ,  $\lim_{\beta \rightarrow +\infty} \beta^{-1} \ln(H^{(\beta)}(x)) = \lim_{\beta \rightarrow +\infty} \beta^{-1} \ln(H^{(\beta)}(x_i)) = 0$ .  $\square$

Ainsi la proposition 11 est déjà prouvée sur  $\underline{\mathcal{U}}$ . Mais le lemme précédent nous a aussi fait faire un grand pas dans le cas général. En effet, si on note

$$\tau := \inf\{n \in \mathbb{N} : X_n \in \underline{\mathcal{U}}\}$$

alors en utilisant la formule (15) avec  $A = \underline{\mathcal{U}}$ , on se rend compte qu'il ne reste plus qu'à obtenir que pour tout  $x \notin \underline{\mathcal{U}}$ ,

$$\lim_{\beta \rightarrow +\infty} \beta^{-1} \ln \left( \mathbb{E}_x \left[ \prod_{0 \leq n < \tau} \frac{G^{(\beta)}(X_n)}{\lambda^{(\beta)}} \right] \right) = -V(x).$$

Pour aller dans cette direction, nous commençons par fixer  $x \notin \underline{\mathcal{U}}$ , puis nous posons, comme dans la proposition 9,  $\underline{\mathcal{U}}^c(x)$  l'ensemble des  $y \notin \underline{\mathcal{U}}$  qui peuvent être atteint à partir de  $x$  en suivant des arêtes orientées sans sortir

de  $\underline{\mathcal{U}}$ . Réduisons ensuite l'espace  $E$  en  $\widehat{E} := \underline{\mathcal{U}}^c(x) \sqcup \{\Delta\}$ , où  $\Delta$  est un nouvel élément (n'appartenant pas à  $E$ ). On munit  $\widehat{E}$  d'un noyau markovien  $\widehat{M}$  en convenant que

$$\forall y, z \in \widehat{E}, \quad \widehat{M}(y, z) := \begin{cases} M(y, z), & \text{si } y, z \in \underline{\mathcal{U}}^c(x) \\ M(y, \underline{\mathcal{U}}), & \text{si } y \in \underline{\mathcal{U}}^c(x) \text{ et } z = \Delta \\ 1, & \text{si } y = \Delta \text{ et } z = x \\ 0, & \text{si } y = \Delta \text{ et } z \neq x. \end{cases}$$

Il est clair que  $\widehat{M}$  est irréductible. Puis on introduit un potentiel  $\widehat{U}$  en posant

$$\forall y \in \widehat{E}, \quad \widehat{U}(y) := \begin{cases} U(y), & \text{si } y \in \underline{\mathcal{U}}^c(x) \\ 0, & \text{si } y = \Delta. \end{cases}$$

Relativement à ce nouveau cadre  $(\widehat{E}, \widehat{M}, \widehat{U})$ , on reconsidère ensuite toutes les notions introduites précédemment (sur lesquelles on mettra évidemment partout des chapeaux). Une observation fondamentale est la suivante :

**Lemme 15.** *Sous la seule hypothèse que  $\lambda(\mathcal{U}) < +\infty$ , on a*

$$\lambda(\widehat{\mathcal{U}}) = \lim_{\beta \rightarrow +\infty} \widehat{\lambda}^{(\beta)} < \lambda(\mathcal{U}).$$

*Démonstration.* La première égalité découle du fait que le lemme 8 et la proposition 9 s'appliquent également à  $(\widehat{E}, \widehat{M}, \widehat{U})$ . Il suffit donc d'analyser  $\widehat{\mathcal{U}}$ . Par définition de  $\widehat{U}$ , il est clair que  $\widehat{\mathcal{U}} = (\mathcal{U} \cap \underline{\mathcal{U}}^c(x)) \sqcup \{\Delta\}$  dont les composantes connexes  $\widehat{\mathcal{U}}_1, \dots, \widehat{\mathcal{U}}_r$  (pour le graphe orienté associé à  $\widehat{M}$ ) sont exactement  $\{\Delta\}$  et les  $\mathcal{U}_i$ , pour  $s < i \leq r$ , inclus dans  $\underline{\mathcal{U}}^c(x)$  (on aura remarqué qu'un tel  $\mathcal{U}_i$  est, soit inclus dans ce dernier, soit disjoint de lui).

Or on constate que  $\lambda(\{\Delta\}) = 0$  ( $\widehat{M}$  fait sortir « immédiatement » de  $\delta$ ), ainsi

$$\begin{aligned} \lambda(\widehat{\mathcal{U}}) &= \max_{1 \leq i \leq r} \lambda(\widehat{\mathcal{U}}_i) \\ &= 0 \vee \max\{\lambda(\mathcal{U}_i) : s < i \leq r \text{ et } \mathcal{U}_i \subset \underline{\mathcal{U}}^c(x)\} \\ &< \lambda(\mathcal{U}). \end{aligned} \quad \square$$

Intéressons-nous maintenant pour  $\beta \geq 0$  au vecteur propre  $\widehat{H}^{(\beta)}$  normalisé par la condition  $\widehat{H}^{(\beta)}(\Delta) = 1$ , et notons qu'il existe une constante  $L \geq 0$  telle que

$$\forall \beta \geq 1, \forall x \in \widehat{E}, \quad \exp(-L\beta) \leq \widehat{H}^{(\beta)}(x) \leq \exp(L\beta).$$

En effet, ces bornes se déduisent facilement par des arguments de chemins des faits suivants :

- $\widehat{S}^{(0)}$  est irréductible et pour tout  $\beta \geq 0$ , la matrice à entrées positives  $\widehat{S}^{(\beta)}$  a même matrice d'incidence que  $\widehat{S}^{(0)}$ ,
- on a

$$\sup_{\beta \geq 1} \max_{x, y \in \widehat{E} : \widehat{S}^{(0)}(x, y) > 0} \left| \frac{1}{\beta} \ln(\widehat{S}^{(\beta)}(x, y)) \right| < +\infty,$$

- pour tout  $\beta \geq 0$ , le vecteur  $\widehat{H}^{(\beta)}$  a ses entrées (strictement) positives et vérifie  $\widehat{H}^{(\beta)}(\Delta) = 1$  et  $\widehat{S}^{(\beta)}[\widehat{H}^{(\beta)}] = \widehat{\lambda}^{(\beta)} \widehat{H}^{(\beta)}$ ,



- on a

$$\sup_{\beta \geq 1} \left| \frac{1}{\beta} \ln(\widehat{\lambda}^{(\beta)}) \right| < +\infty$$

(cette majoration étant elle-même une conséquence des trois précédents points, comme on s'en convainc en considérant dans la relation  $\widehat{S}^{(\beta)}[\widehat{H}^{(\beta)}] = \widehat{\lambda}^{(\beta)} \widehat{H}^{(\beta)}$  des points où  $H^{(\beta)}$  est d'une part maximal et d'autre part minimal).

Le point  $x \in \underline{\mathcal{U}}^c$  étant toujours fixé et avec la constante  $L$  définie ci-dessus, on a l'estimée suivante :

**Lemme 16.** *Si  $\beta \geq 1$  est tel que  $\widehat{\lambda}^{(\sqrt{\beta})} < \lambda^{(\beta)}$ , alors on a*

$$\mathbb{E}_x \left[ \prod_{0 \leq n < \tau} \frac{G^{(\beta)}(X_n)}{\lambda^{(\beta)}} \right] \leq \left( 1 - \frac{\widehat{\lambda}^{(\sqrt{\beta})}}{\lambda^{(\beta)}} \right)^{-1} \exp(2L\sqrt{\beta}) \exp(-(\beta - \sqrt{\beta})V(x)).$$

*Démonstration.* Notons

$$\widehat{\tau} := \inf\{n \geq 0 : \widehat{X}_n = \Delta\}.$$

Puisque  $x \in \underline{\mathcal{U}}^c$ , il est clair d'après nos définitions précédentes que  $(X_n)_{0 \leq n \leq \tau}$  et  $(\widehat{X}_n)_{0 \leq n \leq \widehat{\tau}}$  ont même loi (en particulier  $\tau$  et  $\widehat{\tau}$  ont même distribution), ainsi

$$\mathbb{E}_x \left[ \prod_{0 \leq n < \tau} \frac{G^{(\beta)}(X_n)}{\lambda^{(\beta)}} \right] = \mathbb{E}_x \left[ \prod_{0 \leq n < \widehat{\tau}} \frac{G^{(\beta)}(\widehat{X}_n)}{\lambda^{(\beta)}} \right].$$

Pour évaluer le terme de droite, fixons un temps  $m \in \mathbb{N}^*$  et intéressons-nous à

$$\begin{aligned} & \mathbb{E}_x \left[ \prod_{0 \leq n < m} \frac{G^{(\beta)}(\widehat{X}_n)}{\lambda^{(\beta)}} \mathbb{1}_{\{\widehat{\tau}=m\}} \right] \\ &= \frac{1}{(\lambda^{(\beta)})^m} \mathbb{E}_x \left[ \prod_{0 \leq n < m} G^{(\beta-\sqrt{\beta})}(\widehat{X}_n) \prod_{0 \leq n < m} G^{(\sqrt{\beta})}(\widehat{X}_n) \mathbb{1}_{\{\widehat{\tau}=m\}} \right] \\ &= \frac{(\widehat{\lambda}^{(\sqrt{\beta})})^m}{(\lambda^{(\beta)})^m} \mathbb{E}_x \left[ \prod_{0 \leq n < m} G^{(\beta-\sqrt{\beta})}(\widehat{X}_n) \left( \prod_{0 \leq n < m} \frac{G^{(\sqrt{\beta})}(\widehat{X}_n)}{\widehat{\lambda}^{(\sqrt{\beta})}} \right) \mathbb{1}_{\{\widehat{\tau}=m\}} \right] \\ &= \left( \frac{\widehat{\lambda}^{(\sqrt{\beta})}}{\lambda^{(\beta)}} \right)^m \mathbb{E}_x \left[ \prod_{0 \leq n < m} G^{(\beta-\sqrt{\beta})}(\widehat{X}_n) \left( \prod_{0 \leq n < m} \frac{\widehat{H}^{(\sqrt{\beta})}(\widehat{X}_n)}{\widehat{M}[\widehat{H}^{(\sqrt{\beta})}](\widehat{X}_n)} \right) \mathbb{1}_{\{\widehat{\tau}=m\}} \right]. \end{aligned}$$

Cependant, remarquons que  $(\widehat{X}_n)_{0 \leq n \leq m}$  appartient à  $\mathcal{C}(x, \underline{\mathcal{U}})$  presque sûrement sous la restriction de  $\mathbb{P}_x$  à l'événement  $\{\widehat{\tau} = m\}$ . Ainsi dans la formule ci-dessus, on peut majorer  $\prod_{0 \leq n < m} G^{(\beta-\sqrt{\beta})}(\widehat{X}_n)$  par  $\exp(-(\beta - \sqrt{\beta})V(x))$ . Par ailleurs, soit  $(\widehat{Y}_n)_{n \in \mathbb{N}}$  une chaîne de Markov sur  $\widehat{E}$  homogène en temps, de noyau de probabilités de transition donné par  $\widehat{K}^{(\beta)}$  (défini comme en (6), avec les chapeaux en plus). On peut écrire par changement

de loi

$$\begin{aligned}
\mathbb{E}_x \left[ \prod_{0 \leq n < m} \frac{\widehat{H}^{(\sqrt{\beta})}(\widehat{X}_n)}{\widehat{M}[\widehat{H}^{(\sqrt{\beta})}](\widehat{X}_n)} \mathbb{1}_{\{\widehat{\tau}=m\}} \right] &= \mathbb{E}_x \left[ \frac{\widehat{H}^{(\sqrt{\beta})}(\widehat{Y}_0)}{\widehat{M}[\widehat{H}^{(\sqrt{\beta})}](\widehat{Y}_{m-1})} \mathbb{1}_{\{\widehat{\tau} > m-1\}} \widehat{M}[\mathbb{1}_{\{\delta\}}](\widehat{Y}_{m-1}) \right] \\
&\leq \exp(2L\sqrt{\beta}) \mathbb{E}_x \left[ \mathbb{1}_{\{\widehat{\tau} > m-1\}} \widehat{M}[\mathbb{1}_{\{\delta\}}](\widehat{Y}_{m-1}) \right] \\
&= \exp(2L\sqrt{\beta}) \mathbb{P}_x[\widehat{\tau} = m] \\
&\leq \exp(2L\sqrt{\beta})
\end{aligned}$$

(où l'on a un peu abusé des notations, en désignant de la même manière le temps d'arrêt  $\widehat{\tau}$  associé à  $(\widehat{X}_n)_{n \geq 0}$  et celui relatif à  $(\widehat{Y}_n)_{n \geq 0}$ ). En récapitulant les opérations précédentes, on a donc

$$\mathbb{E}_x \left[ \prod_{0 \leq n < m} \frac{G^{(\beta)}(\widehat{X}_n)}{\lambda^{(\beta)}} \mathbb{1}_{\{\widehat{\tau}=m\}} \right] \leq \exp(2L\sqrt{\beta}) \exp(-(\beta - \sqrt{\beta})V(x)) \left( \frac{\widehat{\lambda}^{(\sqrt{\beta})}}{\lambda^{(\beta)}} \right)^m$$

et il reste alors à sommer en  $m \in \mathbb{N}^*$  pour obtenir le résultat escompté.  $\square$

En tenant compte de la convergence

$$\lim_{\beta \rightarrow +\infty} \frac{\widehat{\lambda}^{(\sqrt{\beta})}}{\lambda^{(\beta)}} = \frac{\lambda(\widehat{\mathcal{U}})}{\lambda(\mathcal{U})} < 1$$

le lemme précédent nous permet d'obtenir la « moitié » de la convergence voulue :

$$\limsup_{\beta \rightarrow +\infty} \beta^{-1} \ln \left( \mathbb{E}_x \left[ \prod_{0 \leq n < \tau} \frac{G^{(\beta)}(X_n)}{\lambda^{(\beta)}} \right] \right) \leq -V(x).$$

Mais il s'agissait là de la partie difficile, l'autre étant beaucoup plus immédiate. En effet, soit  $p = (p_0, p_1, \dots, p_l) \in \mathcal{C}(x, \underline{\mathcal{U}})$  un chemin réalisant le minimum dans la définition de  $V(x)$ ;  $s_U(p) = V(x)$ . On peut alors écrire

$$\begin{aligned}
\mathbb{E}_x \left[ \prod_{0 \leq n < \tau} \frac{G^{(\beta)}(X_n)}{\lambda^{(\beta)}} \right] &\geq \mathbb{E}_x \left[ \mathbb{1}_{\{X_1=p_1, \dots, X_l=p_l\}} \prod_{0 \leq n < \tau} \frac{G^{(\beta)}(X_n)}{\lambda^{(\beta)}} \right] \\
&= \prod_{0 \leq n < l} M(p_n, p_{n+1}) \frac{G^{(\beta)}(p_n)}{\lambda^{(\beta)}} \\
&= \left( \prod_{0 \leq n < l} M(p_n, p_{n+1}) \right) (\lambda^{(\beta)})^{-l} \exp(-\beta V(x))
\end{aligned}$$

d'où

$$\liminf_{\beta \rightarrow +\infty} \beta^{-1} \ln \left( \mathbb{E}_x \left[ \prod_{0 \leq n < \tau} \frac{G^{(\beta)}(X_n)}{\lambda^{(\beta)}} \right] \right) \geq -V(x)$$

ce qui conclut la preuve de la proposition 11.

Ceci nous amène à introduire les quantités suivantes.  
Pour tout  $x \in E$ , soit

$$A(x) := \min_{x \rightarrow y} V(y)$$

où  $x \rightarrow y$  indiquera désormais que  $M(x, y) > 0$  (si  $x \neq y$ , la relation  $x \rightarrow y$  est équivalente au fait que l'arête  $(x, y)$  appartient au graphe orienté sous-jacent, mais ceci n'est plus nécessairement exact si  $x = y$ ). Par ailleurs, si  $a \in \mathcal{A}_x$  est un  $x$ -arbre (dont la définition a été rappelée avant la proposition 7), on note

$$w(a) := \sum_{(y \rightarrow z) \in a} V(z) - A(y),$$

puis

$$\underline{W}(x) := \min_{a \in \mathcal{A}_x} w(a) - \min_{y \in E} \min_{b \in \mathcal{A}_y} w(b). \quad (24)$$

Enfin, on pose

$$\forall x \in E, \quad W(x) := \underline{W}(x) - A(x) - \min_{y \in E} \underline{W}(y) - A(y). \quad (25)$$

On peut maintenant énoncer le

**Corollaire 17.** *Sous les hypothèses de la proposition 11, on a pour tous  $x, y \in E$ ,*

$$\lim_{\beta \rightarrow +\infty} \beta^{-1} \ln(K^{(\beta)}(x, y)) = \begin{cases} -V(y) + A(x), & \text{si } x \rightarrow y \\ -\infty & , \text{ sinon} \end{cases}$$

$$\lim_{\beta \rightarrow +\infty} \beta^{-1} \ln(\mu^{(\beta)}(x)) = -\underline{W}(x)$$

$$\lim_{\beta \rightarrow +\infty} \beta^{-1} \ln(\nu^{(\beta)}(x)) = -W(x).$$

*Démonstration.* Ces convergences sont des conséquences immédiates des représentations de  $K^{(\beta)}$ ,  $\mu^{(\beta)}$  et  $\nu^{(\beta)}$  données dans la section 2. En effet, le comportement asymptotique de  $\mu^{(\beta)}$  découle de sa formulation en termes de  $x$ -arbres et celui de  $\nu^{(\beta)}$  de la description présentée dans la proposition 5.  $\square$

Cependant, l'expression (24) du quasi-potentiel  $\underline{W}$  n'est pas très commode à manipuler. En fait les seules informations que nous ayons réussi à en extraire directement sont les suivantes :

**Lemme 18.** *On a toujours*

$$\min_E \underline{W} = \min_{\underline{U}} \underline{W}$$

et si l'on impose de plus la condition (H), alors

$$\{x \in E : \underline{W}(x) = \min_E \underline{W} = 0\} = \underline{U}.$$

On retrouvera le second résultat ultérieurement d'une manière légèrement différente, notre but ici est juste de montrer comment on peut le déduire de considérations sur les  $x$ -arbres. On se contentera donc d'une démonstration indicative.

*Démonstration.* Soit  $x \notin \underline{\mathcal{U}}$ , et  $a \in \mathcal{A}_x$  un  $x$ -arbre, on cherche à lui associer un  $y$ -arbre  $b \in \mathcal{A}_y$  avec  $y \in \underline{\mathcal{U}}$  et  $w(a) \geq w(b)$ , ce qui permettra de conclure à la première assertion. Pour ceci on commence par considérer l'ensemble  $\{y \in E : x \rightarrow y\}$  et on y choisit  $y_1 \neq x$  tel que  $V(y_1) = A(x)$ . Par construction de  $V$ , ceci est bien possible si  $U(x) > 0$ , car alors  $\min\{V(y) : x \rightarrow y\} = V(x) - U(x) < V(x)$  et un élément minimisant  $y_1$  est donc automatiquement différent de  $x$ . Si  $U(x) = 0$ , il existe au moins un élément  $y' \neq x$  tel que  $x \rightarrow y'$  et  $U(y') = 0$ , il suffit donc de prendre un tel  $y_1 = y'$ . On rajoute alors  $x \rightarrow y_1$  à  $a$  puis on lui enlève l'arête partant de  $y_1$  pour obtenir un  $y_1$ -arbre, disons  $a_1$ . Deux situations sont possibles :

- Si  $y_1 \in \underline{\mathcal{U}}$ , soit  $1 \leq i \leq s$  tel que  $y_1 \in \mathcal{U}_i$ . On efface toutes les arêtes de  $a_1$  qui sont issues d'éléments de  $\mathcal{U}_i$  et on les remplace par celles d'un «  $y_1$ -arbre » inclus dans  $\mathcal{U}_1$  (ceci est toujours possible, car soit  $\mathcal{U}_i$  est le singleton  $\{y_1\}$ , auquel cas le «  $y_1$ -arbre » inclus dans  $\mathcal{U}_1$  est l'ensemble vide, soit  $M_{\mathcal{U}_i}$  est irréductible et alors on peut bien construire dans  $\mathcal{U}_1$  un tel arbre). On note  $b$  le graphe ainsi obtenu, il est clair que  $b \in \mathcal{A}_{y_1}$ .
- Si  $y_1 \notin \underline{\mathcal{U}}$ , on itère la première étape, c'est-à-dire que l'on choisit  $y_2 \neq y_1$  tel que  $V(y_2) = A(x_1)$ , on rajoute  $y_1 \rightarrow y_2$  à  $a_1$  auquel on retire l'arête issue de  $y_2$  pour obtenir un nouvel arbre  $a_2$  etc. On continue jusqu'à ce que  $y_n \in \underline{\mathcal{U}}$  et le premier point ci-dessus permet ensuite de construire  $b \in \mathcal{A}_{y_n}$ . Le lecteur attentif aura remarqué que cette procédure peut éventuellement ne pas aboutir, par exemple si  $U(x) = 0$  et que malencontreusement on soit obligé de prendre  $y_2 = x$ , ce qui conduit apparemment à une boucle! Il faut donc être un peu plus précautionneux et veiller à ce que si l'un des  $y_p$ , pour  $p \in \mathbb{N}$  (par convention  $y_0 = x$ ), vérifie  $U(y_p) = 0$ , alors il faut choisir immédiatement toute une suite finie  $y_{p+1} \rightarrow \dots \rightarrow y_n$  restant dans  $\mathcal{U}$  et telle que  $y_n \in \underline{\mathcal{U}}$  ...

Il apparaît sans difficulté, du fait que  $V(z) - A(y) \geq 0$  pour toute arête orientée  $y \rightarrow z$ , que toutes les opérations précédentes ont fait diminuer la valuation des arbres que l'on a modifiée, ainsi en fin de compte  $w(a) \geq w(b)$ .

Si l'on rajoute la condition (H), on remarque que la quantité  $\min_{a \in \mathcal{A}_y} w(a)$  ne dépend pas de  $y$  tant que ce dernier reste dans  $\underline{\mathcal{U}}$ . D'autre part, notons aussi que sous cette hypothèse on a nécessairement

$$\forall x \in \underline{\mathcal{U}}, \forall y \notin \underline{\mathcal{U}}, \quad x \rightarrow y \Rightarrow V(y) > 0. \quad (26)$$

En effet, si ceci n'était pas vrai, soient  $x, y$  mettant en défaut cette propriété. On pourrait alors trouver un chemin  $p_0 \rightarrow p_1 \rightarrow \dots \rightarrow p_n$ , avec  $p_0 = x, p_1 = y$ , qui resterait dans  $\mathcal{U}$  et tel que  $p_n \in \underline{\mathcal{U}}$ . Soient  $1 \leq i_0, j_0 \leq s$  tels que  $x \in \mathcal{U}_{i_0}$  et  $y \in \mathcal{U}_{j_0}$ . On ne peut pas avoir  $i_0 = j_0$ , sinon ceci contredirait la définition de la composante connexe  $\mathcal{U}_{i_0}$ . Donc  $i_0 \neq j_0$  et dénotons la propriété précédente sous la forme  $i_0 \rightsquigarrow j_0$ . L'hypothèse (H) nous permet de transférer cette relation : pour tout  $1 \leq i \leq s$ , il existe  $1 \leq j \neq i \leq s$  tel que  $i \rightsquigarrow j$ . Il en découle facilement qu'il existe alors un cycle  $i_1 \rightsquigarrow i_2 \rightsquigarrow \dots \rightsquigarrow i_1$  dans  $\{1, \dots, s\}$ , ce qui conduit à nouveau à une contradiction avec la définition des composantes connexes.

Ainsi (26) est satisfait, ce qui permet de voir que lors de la dernière étape de la construction de  $b$  à partir de  $a$ , on a fait décroître strictement la valuation  $w$  de l'arbre, d'où le résultat annoncé.  $\square$

Heureusement, il existe une autre caractérisation du quasi-potentiel  $\underline{W}$  due à Trouvé [27] en termes de symétrie d'élévations, que l'on a réécrit sous forme ensembliste dans [24]. Plus précisément, une fonction  $\widetilde{W} : E \rightarrow \mathbb{R}$  coïncide avec  $\underline{W}$  à une constante additive près si et seulement si pour tout sous-ensemble  $F \subset E$ , on a

$$\min_{x \in F, y \notin F, x \rightarrow y} \widetilde{W}(x) + V(y) - A(x) = \min_{x \notin F, y \in F, x \rightarrow y} \widetilde{W}(x) + V(y) - A(x). \quad (27)$$

Grace à cette formulation, on va pouvoir calculer  $W$  sous l'hypothèse (H) :

**Proposition 19.** *Supposons (H) satisfaite, alors pour tout  $x \in E$ , on a*

$$W(x) = \min_{p \in \mathcal{C}(\underline{\mathcal{U}}, x)} s_U(p)$$

avec  $\mathcal{C}(\underline{\mathcal{U}}, x) := \sqcup_{y \in \underline{\mathcal{U}}} \mathcal{C}(y, x)$  et où la fonctionnelle  $s_U$  sur les chemins a été définie en (23). En quelque sorte, cette formule est duale de celle qui a servi à définir  $V$  en (22).

*Démonstration.* Considérons  $\widehat{W}$  défini comme dans le membre de droite ci-dessus. Pour se convaincre que  $\widehat{W}$  et  $W$  coïncident à une constante additive près, il s'agit de voir que l'application  $\widetilde{W} := \widehat{W} + A$  satisfait bien le critère (27), c'est-à-dire que pour tout sous-ensemble  $F \subset E$ , on a

$$\min_{x \in F, y \notin F, x \rightarrow y} \widehat{W}(x) + V(y) = \min_{x \notin F, y \in F, x \rightarrow y} \widehat{W}(x) + V(y). \quad (28)$$

Pour cela, remarquons que le membre de gauche peut s'écrire

$$\min_{p \in \mathcal{C}(\underline{U}, F, F^c, \underline{U})} s_U(p)$$

où  $\mathcal{C}(\underline{U}, F, F^c, \underline{U})$  désigne l'ensemble des chemins (se recoupant éventuellement) partant de  $\underline{U}$  et arrivant en  $\underline{U}$  et admettant au moins une arête passant de  $F$  à  $F^c$ . Nous allons montrer que

$$\min_{p \in \mathcal{C}(\underline{U}, F, F^c, \underline{U})} s_U(p) \geq \min_{p \in \mathcal{C}(\underline{U}, F^c, F, \underline{U})} s_U(p) \quad (29)$$

ce qui par symétrie permettra de conclure à l'égalité de ces quantités et par suite à la validité de (28). Soit donc  $p \in \mathcal{C}(\underline{U}, F, F^c, \underline{U})$  minimisant pour  $s_U$  et notons  $p_0, p_n \in \underline{U}$  ses points de départ et d'arrivée. Si  $p_0, p_n \in F$ , si  $p_0, p_n \in F^c$  ou si  $p_0 \in F^c, p_n \in F$ , on aura que  $p \in \mathcal{C}(\underline{U}, F^c, F, \underline{U})$  et (29) en découle. Le seul cas problématique correspond donc à la situation où  $p_0 \in F$  et  $p_n \in F^c$ . Mais ceci implique déjà que  $F \cap \underline{U} \neq \emptyset$  et  $F^c \cap \underline{U} \neq \emptyset$ . Observons ensuite que s'il existe  $1 \leq i \leq s$  tel que  $F \cap \mathcal{U}_i \neq \emptyset$  et  $F^c \cap \mathcal{U}_i \neq \emptyset$ , alors les deux membres de (29) sont nuls. On peut donc se restreindre à considérer la situation où chacun des  $\mathcal{U}_i$ , pour  $1 \leq i \leq s$ , est soit inclus dans  $F$ , soit inclus dans  $F^c$ . À une réindexation près de ces ensembles, supposons que  $\mathcal{U}_1, \dots, \mathcal{U}_u \subset F$  et  $\mathcal{U}_{u+1}, \dots, \mathcal{U}_s \subset F^c$ , pour un certain  $1 \leq u < s$ , et que  $p_0 \in \mathcal{U}_1$  et  $p_n \in \mathcal{U}_{u+1}$ . C'est maintenant qu'il faut recourir à la condition (H). Chaque  $g \in \mathcal{G}$  peut être vu comme effectuant une bijection de  $\{\mathcal{U}_1, \dots, \mathcal{U}_s\}$  et notamment on peut lui associer un élément du groupe symétrique  $\mathcal{S}_s$ . On continuera de le dénoter par  $g$  et ceci permet de se représenter  $\mathcal{G}$  comme un sous-groupe de  $\mathcal{S}_s$ . L'hypothèse (H) nous assure alors que l'orbite de 1 par  $\mathcal{G}$  est l'ensemble  $\{1, \dots, s\}$  tout entier, il existe donc  $g \in \mathcal{G}$  tel que  $g(1) = u + 1$ . En considérant l'orbite de 1 par les itérés  $g^q$ , pour  $q \in \mathbb{N}$ , on se rend compte qu'il existe nécessairement un  $q \geq 1$  tel que  $u + 1 \leq g^q(1) \leq s$  et  $1 \leq g^{q+1}(1) \leq u$ . Ceci implique que le chemin image  $g^q(p)$  part de  $\mathcal{U}_{u+1} \sqcup \dots \sqcup \mathcal{U}_s$  pour arriver dans  $\mathcal{U}_1 \sqcup \dots \sqcup \mathcal{U}_u$  et appartient donc à  $\mathcal{C}(\underline{U}, F^c, F, \underline{U})$ . Or par définition du groupe  $\mathcal{G}$ , on a  $s_U(p) = s_U(g^q(p))$  (ceci même pour tout  $q \in \mathbb{Z}$ ), d'où finalement

$$s_U(p) = s_U(g^q(p)) \geq \min_{p \in \mathcal{C}(\underline{U}, F^c, F, \underline{U})} s_U(p).$$

On est donc bien assuré de (29) et par suite de (28).

Pour conclure à l'égalité  $\widehat{W} = W$ , il reste à remarquer que  $\min_E \widehat{W} = 0 = \min_E W$ .  $\square$

**Remarques 20.** a) La première partie du lemme 18 permet de voir que  $\underline{W}(x) = 0$  pour  $x \in \underline{U}$  et comme on a également  $A(x) = 0$  pour de tels points, on voit que  $\underline{W} - A$  s'annule sur  $\underline{U}$ . Or la proposition précédente montre que le minimum de  $\underline{W} - A$  est en particulier atteint sur  $\underline{U}$ , d'où finalement  $\min_E \underline{W} - A = 0$ . Ainsi dans (25), on aurait pu se contenter de définir  $W$  par la formule  $W := \underline{W} - A$ .

b) Si on suppose de plus que  $M$  est faiblement réversible, alors pour tout sous-ensemble  $F \subset E$ , on évidemment une bijection entre  $\mathcal{C}(\underline{U}, F, F^c, \underline{U})$  et  $\mathcal{C}(\underline{U}, F^c, F, \underline{U})$  par retournement du sens de parcours des chemins et la démonstration précédente se simplifie beaucoup. Ce même argument permet d'ailleurs de se convaincre que dans cette situation  $V = W$ , ce qui peut aussi se voir directement dans (28), qui est trivialement satisfait si on y remplace  $\widehat{W}$  par  $V$ , par symétrie de la relation  $\rightarrow$ . Ceci justifie l'une des assertions faites dans la remarque 12.  $\square$

La remarque (a) précédente et la proposition 19 permettent aussi de calculer  $\underline{W}$  :

$$\begin{aligned} \forall x \in E, \quad \underline{W}(x) &= A(x) + W(x) \\ &= \min_{p \in \mathcal{C}(\underline{\mathcal{U}}, x, \underline{\mathcal{U}})} s_U(p) \end{aligned} \quad (30)$$

où  $\mathcal{C}(\underline{\mathcal{U}}, x, \underline{\mathcal{U}})$  est l'ensemble des chemins (se recoupant éventuellement) partant de  $\underline{\mathcal{U}}$ , arrivant en  $\underline{\mathcal{U}}$  et passant par  $x$ . Par des arguments similaires à ceux utilisés à la fin de la démonstration du lemme 18, on se rend facilement compte que  $\underline{W}$  est strictement positif sur  $E \setminus \underline{\mathcal{U}}$ , c'est-à-dire que l'on retrouve que sous la condition (H),

$$\{x \in E : \underline{W}(x) = 0\} = \underline{\mathcal{U}}.$$

Par contre, ceci n'est plus nécessairement vrai pour  $W$  et la proposition 19 montre en fait que

$$\{x \in E : W(x) = 0\} = \overline{\mathcal{U}}$$

où l'ensemble  $\overline{\mathcal{U}}$  a été défini dans l'introduction.

Notamment il ressort qu'à basse température, la probabilité  $\nu^{(\beta)}$  se concentre (exponentiellement rapidement) sur  $\overline{\mathcal{U}}$  :

$$\lim_{\beta \rightarrow +\infty} \nu^{(\beta)}(\overline{\mathcal{U}}) = 1.$$

À la vue de cette convergence, on est en droit de se demander si elle est optimale, *i.e.* si l'on charge effectivement asymptotiquement tous les points de  $\overline{\mathcal{U}}$ . La réponse est affirmative, car on peut préciser les convergences précédentes :

**Proposition 21.** *Sous la condition (H), il existe une constante  $\eta > 0$  telle que*

$$\forall x \in E, \forall \beta \geq 0, \quad \eta \exp(-\beta W(x)) \leq \nu^{(\beta)}(x) \leq \eta^{-1} \exp(-\beta W(x)).$$

*Des encadrement similaires sont aussi valables pour  $H^{(\beta)}$  et  $\mu^{(\beta)}$ , avec des constantes adéquates.*

*Démonstration.* Au vu des formulations de  $\mu^{(\beta)}$  et de  $\nu^{(\beta)}$ , pour  $\beta \geq 0$ , présentées dans la section 2, il suffit de se convaincre qu'il existe  $\eta' > 0$  tel que

$$\forall x \in E, \forall \beta \geq 0, \quad \eta' \exp(-\beta V(x)) \leq H^{(\beta)}(x) \leq (\eta')^{-1} \exp(-\beta V(x)). \quad (31)$$

Mais un réexamen de la preuve des lemmes 13 et 14 montre qu'il existe  $\eta'' > 0$  tel que

$$\forall x \in \underline{\mathcal{U}}, \forall \beta \geq 0, \quad \eta'' \leq H^{(\beta)}(x) \leq (\eta'')^{-1}$$

(où  $H^{(\beta)}$  est renormalisé par la condition  $H^{(\beta)}(x_0) = 1$ , pour un  $x_0 \in \mathcal{U}_1$  fixé), du fait que l'on sait que sous (H),

$$0 < \lim_{\beta \rightarrow +\infty} \lambda^{(\beta)} < +\infty. \quad (32)$$

Enfin, on remarque que dans la preuve du lemme 16, au lieu de considérer  $\sqrt{\beta}$ , on peut se contenter de remplacer cette variable par une constante  $\beta_0 > 0$  suffisamment grande (ceci permet d'obtenir la borne supérieure dans (31), l'autre moitié se déduisant également de (32)).  $\square$

Il est tentant de conjecturer que les limites suivantes existent pour tout  $x \in E$  :

$$\lim_{\beta \rightarrow +\infty} \exp(\beta W(x)) \nu^{(\beta)}(x)$$

mais nous ne chercherons pas ici à atteindre une telle précision. Celle-ci serait toutefois nécessaire, si comme pour le recuit simulé, on voulait prouver que  $\lim_{n \rightarrow \infty} m_n$  existe pour toute évolution de la forme (3) (sous la condition (H) ou plus généralement). Dans le cadre du théorème 3 ou de la conjecture 4, cette limite dépendra ou non de la condition initiale  $m_0$  selon que  $k < k_2$  ou  $k \geq k_2$ .

### 3. ÉVOLUTION

Dans la section précédente, on a étudié le comportement à basse température d'objets « statiques », *i.e.* correspondant à un équilibre, comme le point fixe  $\nu^{(\beta)}$ . Nous allons maintenant appliquer ces résultats à l'étude des dynamiques de Feynman-Kac recuites présentées dans l'introduction (on supposera notamment toujours satisfaite l'hypothèse (H)). Pour ce faire, on se ramènera à des algorithmes de recuit simulé généralisé. Ceux-ci ont déjà été beaucoup étudiés, néanmoins on aura besoin de nouvelles estimées dans un contexte particulier.

Reconsidérons donc  $(m_n)_{n \in \mathbb{N}}$  la suite des probabilités générées par une dynamique de Feynman-Kac recuite associée à un schéma d'inverse de température  $\beta_{\mathbb{N}^*}$ . Pour se ramener à un cadre presque markovien, introduisons d'abord pour tout  $\beta \geq 0$ , une nouvelle opération de sélection

$$R^{(\beta)} : \mathcal{P}(E) \ni m \mapsto \frac{M[H^{(\beta)}] \cdot m}{m[M[H^{(\beta)}]]} \in \mathcal{P}(E)$$

ainsi que sa sélection inverse

$$(R^{(\beta)})^{-1} : \mathcal{P}(E) \ni m \mapsto \frac{1}{m[1/M[H^{(\beta)}]]} \frac{1}{M[H^{(\beta)}]} \cdot m \in \mathcal{P}(E).$$

Ensuite, pour tous  $\beta, \beta' \geq 0$ , soit  $\Psi^{(\beta, \beta')} := R^{(\beta')} \circ (R^{(\beta)})^{-1}$ , qui correspond aussi à une opération de sélection, relativement à la fonction  $M[H^{(\beta')}] / M[H^{(\beta)}]$ .

Enfin, notons pour tout  $n \in \mathbb{N}$ ,

$$\tilde{m}_n := R^{(\beta_n)}(m_n)$$

(prenons par convention  $\beta_0 = 0$ , de sorte que  $\tilde{m}_0 = m_0$ ), puis faisons l'observation fondamentale suivante :

**Lemme 22.** *La suite  $(\tilde{m}_n)_{n \in \mathbb{N}}$  peut aussi être vue comme le flot de la dynamique de Feynman-Kac associée d'une part aux opérations de sélection  $(\Psi^{(\beta_n, \beta_{n+1})})_{n \in \mathbb{N}}$  et d'autre part aux mutations inhomogènes en temps décrites par la suite de noyaux  $(K^{(\beta_{n+1})})_{n \in \mathbb{N}}$ , les sélections étant effectuées en premier, c'est-à-dire que*

$$\forall n \in \mathbb{N}, \quad \tilde{m}_{n+1} = \Psi^{(\beta_n, \beta_{n+1})}(\tilde{m}_n) K^{(\beta_{n+1})}.$$

*Démonstration.* En reprenant les notations de la section 2 et en remarquant que pour tous  $\beta, \beta' \geq 0$ ,  $R^{(\beta)} = \Psi^{(\beta', \beta)} \circ R^{(\beta')}$ , il suffit de montrer que

$$\forall \beta \geq 0, \forall m \in \mathcal{P}(E), \quad R^{(\beta)}(m) K^{(\beta)} = R^{(\beta)}(T^{(\beta)}(m)) \quad (33)$$

(que l'on applique ensuite avec  $m := m_n$  et  $\beta := \beta_{n+1}$ , pour  $n \in \mathbb{N}$ ).

En partant du membre de gauche, on a pour toute fonction test  $f \in \mathcal{B}(E)$ ,

$$\begin{aligned} (R^{(\beta)}(m) K^{(\beta)})[f] &= R^{(\beta)}(m)[K^{(\beta)}[f]] \\ &= \frac{m[M[H^{(\beta)}] K^{(\beta)}[f]]}{m[M[H^{(\beta)}]]} \\ &= \frac{m[M[H^{(\beta)}] f]}{m[M[H^{(\beta)}]]}. \end{aligned}$$

D'autre part, on a

$$\begin{aligned}
R^{(\beta)}(T^{(\beta)}(m))[f] &= \frac{T^{(\beta)}(m)[M[H^{(\beta)}]f]}{T^{(\beta)}(m)[M[H^{(\beta)}]]} \\
&= \frac{(mM)[G^{(\beta)}M[H^{(\beta)}]f]}{(mM)[G^{(\beta)}M[H^{(\beta)}]]} \\
&= \frac{\lambda^{(\beta)}(mM)[H^{(\beta)}f]}{\lambda^{(\beta)}(mM)[H^{(\beta)}]} \\
&= \frac{(mM)[H^{(\beta)}f]}{(mM)[H^{(\beta)}]}
\end{aligned}$$

d'où l'égalité voulue (33). Notamment en y prenant  $m = \nu^{(\beta)}$ , on retrouve que  $R^{(\beta)}(\nu^{(\beta)}) = \mu^{(\beta)}$ , *i.e.* la proposition 5 (les arguments ci-dessus ne sont d'ailleurs qu'une réécriture de sa preuve).  $\square$

**Remarque 23.** Dans les situations homogènes où la température ne dépend pas du temps, le lemme précédent nous ramène à l'étude des marginales temporelles d'une chaîne de Markov. Cette remarque permet d'avoir recours aux propriétés de contraction des noyaux markoviens (voir par exemple [7]) et on peut ainsi sensiblement améliorer les précédentes études de stabilité effectuées dans ce domaine (voir [11–13]).

Cependant nous remettons le développement de ces conséquences à plus tard, puisque nous préférons nous focaliser ici sur des comportements du type recuit (*a priori* plus délicats), où la température a tendance à s'annuler en temps grand.

Puisque l'on va finir par se restreindre à des évolutions  $\beta_{\mathbb{N}^*}$  logarithmiques, on peut espérer que les effets des sélections  $\Psi^{(\beta_n, \beta_{n+1})}$  ne seront pas très importants pour  $n \in \mathbb{N}$  grand, car alors  $\beta_n$  et  $\beta_{n+1}$  ne seront pas très différents, ce qui implique, comme on le précisera ultérieurement, que  $\Psi^{(\beta_n, \beta_{n+1})}$  est proche de l'opérateur identité. Ceci suggère de s'intéresser dans un premier temps aux suites  $(\hat{m}_n)_{n \in \mathbb{N}}$  des marginales temporelles de chaînes de Markov inhomogènes dont les transitions sont dictées par la suite de noyaux  $(K^{(\beta_{n+1})})_{n \in \mathbb{N}}$ , *i.e.* dont les éléments sont des probabilités enchaînées par les relations

$$\forall n \in \mathbb{N}, \quad \hat{m}_{n+1} = \hat{m}_n K^{(\beta_{n+1})}.$$

On se ramène ainsi à considérer un algorithme de recuit simulé généralisé, mais nous aurons besoin d'estimées relativement précises sur ces  $\hat{m}_n$ , pour  $n \in \mathbb{N}$ , pour pouvoir à terme en extraire des renseignements sur la suite  $(m_n)_{n \in \mathbb{N}}$ . L'objectif à atteindre est que

$$\sup_{n \in \mathbb{N}^*} \left\| \frac{\hat{m}_n}{\mu^{(\beta_n)}} \right\|_{\infty} < +\infty. \quad (34)$$

Une telle borne en norme supérieure nécessite généralement plus d'efforts que celles plus faibles, mesurant par exemple un écart au sens de l'entropie relative, utilisées lors d'approches du genre semi-groupe en théorie du recuit simulé. Néanmoins, notre tâche va être grandement facilitée d'une part par le fait que  $\underline{W}$  n'admet pas de minimum local qui ne soit pas global et d'autre part par le caractère « non-tendu » de (34), au sens où cette inégalité n'implique pas une convergence vers 0 de  $\hat{m}_n - \mu^{(\beta_n)}$  en temps grand (d'ailleurs si l'un de ces deux aspects n'avait pas été vérifié, on ne peut pas espérer un tel résultat pour toutes les évolutions logarithmiques de la forme (3) de l'inverse de la température; il faudrait imposer que  $k > 0$  soit assez grand, pour des informations en direction de cette affirmation, on renvoie respectivement à [21] et à [6]).

Puisqu'une grande partie de notre travail dans cette section va justement être de prouver (34) sous des conditions adéquates, on notera désormais par  $(Y_n)_{n \in \mathbb{N}}$  une chaîne de Markov inhomogène sur  $E$  dont les



transitions sont données par les noyaux  $K^{(\beta_{n+1})}$ , à tout instant  $n \in \mathbb{N}$ , relativement à une certaine évolution  $\beta_{\mathbb{N}^*}$ . Pour  $n \in \mathbb{N}$  et  $x \in E$ , on désignera aussi par  $\mathbb{P}_{n,x}$  la loi de  $(Y_{n+p})_{p \in \mathbb{N}}$  conditionnée par  $Y_n = x$ .

Revenons un peu sur l'absence de minimum local qui ne soit pas global mentionnée précédemment. Il s'agit en fait d'une propriété relative à la fonction de coût  $\underline{W}$  associée à la famille  $(K^{(\beta)})_{\beta \geq 0}$  par

$$\begin{aligned} \forall x, y \in E, \quad \underline{W}(x, y) &:= - \lim_{\beta \rightarrow +\infty} \beta^{-1} \ln(K^{(\beta)}(x, y)) \\ &= \begin{cases} V(y) - A(x), & \text{si } x \rightarrow y \\ +\infty, & \text{sinon} \end{cases} \end{aligned}$$

Le quasi-potentiel associé à cette fonction de coût est dénoté de la même manière (mais ne dépend que d'une variable), rappelons qu'il est donné par

$$\forall x \in E, \quad \underline{W}(x) = - \lim_{\beta \rightarrow +\infty} \beta^{-1} \ln(\mu^{(\beta)}(x))$$

où pour tout  $\beta \geq 0$ ,  $\mu^{(\beta)}$  est l'unique probabilité invariante associée à  $K^{(\beta)}$  (le quasi-potentiel peut aussi se décrire uniquement en terme de sa fonction de coût grâce à la caractérisation de Trounev, cf. [24, 27]). On dit alors de manière générale qu'une fonction de coût  $\underline{W}$  n'admet pas de minimum local qui ne soit pas global, si pour tout  $x \in E$ , il existe un chemin  $x = x_0, x_1, \dots, x_p$ , avec  $p \in \mathbb{N}$ , tel que pour tout  $0 \leq i < p$ ,  $\underline{W}(x_i, x_{i+1}) = 0$  et  $\underline{W}(x_p) = 0$ . Ceci est bien le cas dans la situation qui nous concerne, comme on s'en convainc facilement à partir du calcul de  $\underline{W}$  effectué à la fin de la section précédente (sous l'hypothèse (H)). Une conséquence de cette propriété est que pour  $n \in \mathbb{N}$  grand,  $\hat{m}_n$  va avoir tendance à se concentrer sur l'ensemble  $\underline{\mathcal{U}}$  des minima globaux de  $\underline{W}$ . Typiquement, on aura donc qu'en de grands temps,  $Y_n \in \underline{\mathcal{U}}$ , ce qui nous amène à nous intéresser au comportement de  $(Y_{n+p})_{p \in \mathbb{N}}$  sous  $\mathbb{P}_{n,x}$ , pour  $x \in \underline{\mathcal{U}}$ . Notre première estimée s'énonce :

**Lemme 24.** *Supposons que l'évolution  $\beta_{\mathbb{N}^*}$  soit croissante et soit  $L := \text{card}(\{\underline{W}(y) : y \in E\}) \leq \text{card}(E)$ . Il existe une constante  $C > 0$  telle que*

$$\forall n \in \mathbb{N}, \forall x \in \underline{\mathcal{U}}, \forall p \in \mathbb{N}, \forall y \in E, \quad \mathbb{P}_{n,x}[Y_{n+p} = y] \leq Cp^{L-1} \exp(-\beta_n \underline{W}(y)). \quad (35)$$

*Démonstration.* Ordonnons les valeurs de  $\{\underline{W}(y) : y \in E\}$  en  $0 = w_0 < w_1 < \dots < w_{L-1}$ , puis dénotons pour tout  $0 \leq l \leq L-1$ ,  $\mathcal{W}_l := \{y \in E : \underline{W}(y) = w_l\}$ . Comme on l'a rappelé ci-dessus, on a  $\mathcal{W}_0 = \underline{\mathcal{U}}$ . Nous allons prouver la borne (35) par une récurrence sur  $0 \leq l < L$ . Plus précisément, supposons que pour un  $0 \leq l < L-1$ , on soit assuré de l'existence d'une constante  $C_l > 0$  telle que pour tous  $n, p \in \mathbb{N}$  et tout  $y \in \mathcal{W}_0 \sqcup \dots \sqcup \mathcal{W}_l$ , on ait

$$\mathbb{P}_{n,x}[Y_{n+p} = y] \leq C_l p^l \exp(-\beta_n \underline{W}(y))$$

on va en déduire que cette propriété reste satisfaite à l'ordre  $l+1$ .

On aura remarqué que l'initialisation de cette récurrence est triviale, avec  $C_0 = 1$ . Considérons donc maintenant le passage du niveau  $0 \leq l < L-1$  à  $l+1$ . On note  $f$  la fonction indicatrice de  $\mathcal{W}_{l+1} \sqcup \dots \sqcup \mathcal{W}_{L-1}$ , puis

$$\tau := \inf\{p \in \mathbb{N} : f(Y_{n+p}) = 1\}.$$

Pour tout  $p \in \mathbb{N}$ , on peut écrire

$$f(Y_{n+p}) = f(Y_n) + \sum_{0 \leq q \leq p-1} (K^{(\beta_{n+q+1})} - \text{Id})[f](Y_{n+q}) + M_{n+p}$$

où  $(M_{n+p})_{p \geq 0}$  est une martingale sous  $\mathbb{P}_{n,x}$  (relativement à la filtration naturellement engendrée par  $(Y_{n+p})_{p \in \mathbb{N}}$ ). Fixons  $p \in \mathbb{N}$ , en considérant l'égalité précédente avec  $p$  remplacé par  $\tau \wedge p$  et en utilisant le théorème d'arrêt

de Doob, il apparaît que

$$\begin{aligned} \mathbb{P}_{n,x}[\tau \leq p] &= \mathbb{E}_{n,x}[f(Y_{n+\tau \wedge p})] \\ &= \mathbb{E}_{n,x}[f(Y_n)] + \mathbb{E}_{n,x} \left[ \sum_{0 \leq q \leq \tau \wedge p-1} (K^{(\beta_{n+q+1})} - \text{Id})[f](Y_{n+q}) \right] \\ &= \mathbb{E}_{n,x} \left[ \sum_{0 \leq q \leq \tau \wedge p-1} (K^{(\beta_{n+q+1})} - \text{Id})[f](Y_{n+q}) \right]. \end{aligned}$$

Cependant, on calcule que pour  $q \geq 0$  et  $y \in \mathcal{W}_0 \sqcup \dots \sqcup \mathcal{W}_l$ ,

$$\begin{aligned} (K^{(\beta_{n+q+1})} - \text{Id})[f](y) &= K^{(\beta_{n+q+1})}[f](y) - f(y) \\ &= K^{(\beta_{n+q+1})}[f](y) \\ &= \sum_{z \notin \mathcal{W}_0 \sqcup \dots \sqcup \mathcal{W}_l} K^{(\beta_{n+q+1})}(y, z). \end{aligned}$$

Or, pour  $y \in \mathcal{W}_0 \sqcup \dots \sqcup \mathcal{W}_l$  et  $z \notin \mathcal{W}_0 \sqcup \dots \sqcup \mathcal{W}_l$  vérifiant  $y \rightarrow z$ , on a d'après la formule (30),

$$W(y) + V(z) \geq \underline{W}(z) \geq w_{l+1}$$

d'où

$$\begin{aligned} V(z) - A(y) &\geq w_{l+1} - A(y) - W(y) \\ &= w_{l+1} - \underline{W}(y). \end{aligned}$$

Plus généralement, profitons-en pour signaler que ceci permet aussi de voir que

$$\forall y, z \in E, \quad \underline{W}(y, z) \geq (\underline{W}(z) - \underline{W}(y))_+. \quad (36)$$

Ainsi d'après l'estimée de la proposition 21, il existe une constante  $C > 0$  telle que

$$\begin{aligned} &\left| \mathbb{E}_{n,x} \left[ \sum_{0 \leq q \leq \tau \wedge p-1} (K^{(\beta_{n+q+1})} - \text{Id})[f](Y_{n+q}) \right] \right| \\ &\leq \left| \mathbb{E}_{n,x} \left[ \sum_{0 \leq q \leq p-1} (K^{(\beta_{n+q+1})} - \text{Id})[f](Y_{n+q}) \mathbb{1}_{\mathcal{W}_0 \sqcup \dots \sqcup \mathcal{W}_l}(Y_{n+q}) \right] \right| \\ &\leq C \sum_{0 \leq q \leq p-1} \sum_{y \in \mathcal{W}_0 \sqcup \dots \sqcup \mathcal{W}_l} \exp(-\beta_{n+q+1}(w_{l+1} - \underline{W}(y))) \mathbb{P}_{n,x}[Y_{n+q} = y] \\ &\leq C \sum_{0 \leq q \leq p-1} \sum_{y \in \mathcal{W}_0 \sqcup \dots \sqcup \mathcal{W}_l} \exp(-\beta_n(w_{l+1} - \underline{W}(y))) \mathbb{P}_{n,x}[Y_{n+q} = y]. \end{aligned}$$

En utilisant d'une part la croissance de  $\beta_{\mathbb{N}^*}$  et d'autre part l'hypothèse de récurrence, on peut majorer cette dernière somme par  $CC_l p^l \exp(-\beta_n w_{l+1})$ , puis pour compléter le passage à l'étape  $l+1$  avec  $C_{l+1} = CC_l$ , il suffit d'observer que pour tout  $p \in \mathbb{N}$ ,

$$\mathbb{P}_{n,x}[Y_{n+p} \in \mathcal{W}_{l+1}] \leq \mathbb{P}_{n,x}[\tau \leq p]. \quad \square$$

Ce genre d'estimation reste juste quelque soit le point de « départ »  $x$ , à condition que  $p$  soit au moins de l'ordre de  $\ln(n)$  pour  $n \in \mathbb{N}$  grand, car l'absence de minimum local non global nous assure que de n'importe quel point, on rejoint toujours rapidement  $\underline{\mathcal{U}}$ .

**Lemme 25.** *On reprend les notations du lemme 24, mais on suppose plus précisément que  $\beta_{\mathbb{N}^*}$  est de la forme (3) avec  $k > 0$ . Il existe alors deux constantes  $C, K > 0$  telles que*

$$\forall n \in \mathbb{N}, \forall x \in E, \forall y \in E, \forall p \geq p_n, \quad \mathbb{P}_{n,x}[Y_{n+p} = y] \leq Cp^{L-1} \exp(-\beta_n \underline{W}(y))$$

où l'on a posé  $p_n := \lfloor K \ln(2+n) \rfloor$ .

*Démonstration.* Pour  $n \in \mathbb{N}$ , définissons

$$\tau^{(n)} := \inf\{p \in \mathbb{N} : Y_{n+p} \in \underline{\mathcal{U}}\}$$

Puisque la fonction de coût  $\underline{W}$  n'admet pas de minima local qui ne soit pas global et que l'ensemble des minima globaux est  $\underline{\mathcal{U}}$ , on obtient facilement, en vertu des estimées de la proposition 21, qu'il existe  $\epsilon > 0$  assez petit et  $N \in \mathbb{N}$  assez grand, tels que

$$\forall n \in \mathbb{N}, \forall x \in E, \quad \mathbb{P}_{n,x}[\tau^{(n)} \leq N] \geq \epsilon$$

Ceci est d'ailleurs valable pour toute évolution  $\beta_{\mathbb{N}^*}$ , mais en supposant de plus qu'elle est croissante, on obtient pour tous  $n, p \in \mathbb{N}$  et  $x, y \in E$ ,

$$\begin{aligned} \mathbb{P}_{n,x}[Y_{n+p} = y] &= \mathbb{P}_{n,x}[Y_{n+p} = y, \tau^{(n)} \leq p] + \mathbb{P}_{n,x}[Y_{n+p} = y, \tau^{(n)} > p] \\ &\leq \mathbb{P}_{n,x}[Y_{n+p} = y, \tau^{(n)} \leq p] + \mathbb{P}_{n,x}[\tau^{(n)} > p] \\ &\leq \mathbb{E}_{n,x}[\mathbb{1}_{\{\tau^{(n)} \leq p\}} \mathbb{P}_{n+\tau^{(n)}, Y_{n+\tau^{(n)}}}[Y_{n+p} = y]] + (1-\epsilon)^{\lfloor p/N \rfloor} \\ &\leq C \mathbb{E}_{n,x}[\mathbb{1}_{\{\tau^{(n)} \leq p\}} (p - \tau^{(n)})^{L-1} \exp(-\beta_n \underline{W}(y))] + (1-\epsilon)^{\lfloor p/N \rfloor} \\ &\leq Cp^{L-1} \exp(-\beta_n \underline{W}(y)) + (1-\epsilon)^{\lfloor p/N \rfloor} \end{aligned}$$

où l'avant-dernière inégalité découle d'une application directe du lemme 24, puisque sur l'événement  $\{\tau^{(n)} \leq p\}$ , on a  $Y_{n+\tau^{(n)}} \in \underline{\mathcal{U}}$ . On aboutit au résultat annoncé pour une évolution  $\beta_{\mathbb{N}^*}$  de la forme (3), en prenant  $p \geq \lfloor K \ln(2+n) \rfloor$ , avec  $K > 0$  assez grand (par exemple pour  $K = Nk^{-1} \max_E \underline{W}$ ).  $\square$

Avec les notations précédentes, écrivons pour tout  $n \in \mathbb{N}$ ,  $n' = n + p_n \in \mathbb{N}$ . Des manipulations élémentaires permettent de se rendre compte qu'il existe deux constantes  $n'_0, K' > 0$  telles que

$$n' \geq n'_0 \Rightarrow n \geq n' - \lfloor K' \ln(n' + 2) \rfloor \geq 0.$$

En ayant recours à l'estimée du lemme précédent, on obtient ainsi que pour tout  $n' \geq n'_0$  et tout  $y \in E$ ,

$$\begin{aligned} \widehat{m}_{n'}(y) &= \sum_{x \in E} \widehat{m}_n(x) \mathbb{P}_{n,x}[Y_{n'} = y] \\ &\leq C' \ln^{L-1}(n' + 2) \exp(|\beta_{n'} - \beta_n| \max_E \underline{W}) \exp(-\beta_{n'} \underline{W}(y)) \end{aligned}$$

pour une constante  $C' > 0$  appropriée. Or si  $\beta_{\mathbb{N}^*}$  est de la forme (3), on peut trouver une autre constante  $C'' > 0$  (dépendant de  $k$ ) telle que pour tout  $n' \in \mathbb{N}$  assez grand, disons encore  $n' \geq n'_0$ , on ait

$$\begin{aligned} |\beta_{n'} - \beta_n| &\leq C'' \frac{\ln(n' + 2)}{n'} \\ &\leq C''. \end{aligned}$$

En fin de compte, on obtient donc avec  $C''' := C' \exp(C'' \max_E \underline{W})$

$$\forall n' \geq n'_0, \forall y \in E, \quad \widehat{m}_n(y) \leq C''' \ln^{L-1}(n' + 2)^{L-1} \exp(-\beta_{n'} \underline{W}(y))$$

borne qui peut ensuite s'étendre sans difficulté à tous les  $n' \in \mathbb{N}$ .

En réutilisant la proposition 21, on a donc montré une version un peu moins bonne de l'inégalité voulue (34) : il existe une constante  $C > 0$  telle que

$$\forall n \in \mathbb{N}, \quad \left\| \frac{\widehat{m}_n}{\mu^{(\beta_n)}} \right\|_{\infty} \leq C \ln^{L-1}(2+n). \quad (37)$$

Néanmoins l'avantage des arguments précédents est de montrer que pour obtenir (34), il suffit de se convaincre de la validité du résultat suivant.

**Proposition 26.** *Outre l'hypothèse (H) qui est toujours sous-entendue, supposons que l'évolution  $\beta_{\mathbb{N}^*}$  soit croissante, il existe alors une constante  $C > 0$  telle que*

$$\forall n \in \mathbb{N}, \forall x \in \underline{\mathcal{U}}, \forall p \in \mathbb{N}, \forall y \in E, \quad \mathbb{P}_{n,x}[Y_{n+p} = y] \leq C \exp(-\beta_n \underline{W}(y)).$$

*Démonstration.* On va utiliser des arguments de couplage, souvent plus difficiles à exposer qu'à comprendre. Commençons par rappeler que d'après la proposition 21, il existe deux constantes  $0 < a \leq b$  telles que

$$\forall \beta \geq 0, \forall y, z \in E, \quad a \exp(-\beta \underline{W}(y, z)) \leq K^{(\beta)}(y, z) \leq b \exp(-\beta \underline{W}(y, z)). \quad (38)$$

Puis introduisons les définitions suivantes : pour tout  $\kappa \in \mathbb{N}^*$ , soit

$$\forall y, z \in E, \quad \underline{W}_{\kappa}(y, z) := \min_{y=y_1, y_2, \dots, y_{\kappa}=z} \underline{W}(y_1, y_2) + \dots + \underline{W}(y_{\kappa-1}, y_{\kappa}) \in \bar{\mathbb{R}}_+.$$

De par l'absence de minimum local qui ne soit pas global et du fait que les composantes connexes de  $\underline{\mathcal{U}}$  sont aperiodiques, on peut trouver un  $\kappa \in \mathbb{N}^*$  tel que pour tout  $y \in E$ , il existe  $z \in \underline{\mathcal{U}}$  avec  $\underline{W}_{\kappa}(y, z) = 0$ . On remarque également que si  $\underline{W}(y) < \underline{W}(z)$ , alors  $\underline{W}_{\kappa}(y, z) \geq \underline{W}(y) - \underline{W}(z)$  (ceci découle essentiellement de (36)). En travaillant un peu à partir de (38), on obtient qu'il existe deux constantes  $0 < a_{\kappa} < 1$  et  $b_{\kappa} > 0$ , telles que pour tout  $\beta \geq 0$  et tous  $\beta_1 \geq \beta, \beta_2 \geq \beta, \dots, \beta_{\kappa} \geq \beta$ , on a

$$\forall y \in E, \quad (K^{(\beta_1)} K^{(\beta_2)} \dots K^{(\beta_{\kappa})})(y, \underline{\mathcal{U}}) \geq a_{\kappa} \quad (39)$$

et pour tous  $y, z \in E$  avec  $\underline{W}(y) < \underline{W}(z)$ ,

$$(K^{(\beta_1)} K^{(\beta_2)} \dots K^{(\beta_{\kappa})})(y, z) \leq b_{\kappa} \exp(-\beta(\underline{W}(y) - \underline{W}(z))). \quad (40)$$

Ceci nous amène à considérer pour tout  $\beta \geq 0$ , la matrice  $(P^{(\beta)}(\omega, \omega'))_{\omega, \omega' \in \Omega}$ , où  $\Omega := \{\underline{W}(y) : y \in E\}$ , définie par

$$\forall \omega, \omega' \in \Omega, \quad P^{(\beta)}(\omega, \omega') := \begin{cases} b_{\kappa} \exp(-\beta(\omega' - \omega)), & \text{si } \omega' > \omega \\ a_{\kappa}, & \text{si } \omega \neq 0 \text{ et } \omega' = 0 \\ 1 - a_{\kappa} - b_{\kappa} \sum_{\omega'' > \omega} \exp(-\beta(\omega'' - \omega)), & \text{si } \omega = \omega' \\ 0, & \text{sinon.} \end{cases}$$

Pour  $\beta \geq 0$  assez grand, disons  $\beta \geq B$ ,  $P^{(\beta)}$  sera une matrice markovienne sur  $\Omega$ . L'intérêt de la famille  $(P^{(\beta)})_{\beta \geq B}$  provient des couplages permis par les inégalités (39) et (40). Plus précisément, supposons que la suite  $(\beta_n)_{n \in \mathbb{N}^*}$  prend toutes ses valeurs dans  $[\beta, +\infty[$ , avec  $\beta \geq B$ . Soit  $(\tilde{Y}_0, Z_0)$  une v.a. à valeurs dans  $E \times \Omega$  telle que

$\underline{W}(\tilde{Y}_0) \leq Z_0$ . On peut alors construire une chaîne de Markov  $(\tilde{Y}_n, Z_n)_{n \in \mathbb{N}}$  à valeurs dans  $E \times \Omega$ , telle que  $(\tilde{Y}_n)_{n \in \mathbb{N}}$  soit une chaîne de Markov inhomogène de transitions données par la famille  $(K^{(\beta_{n\kappa+1})} K^{(\beta_{n\kappa+2})} \dots K^{(\beta_{(n+1)\kappa}})_{n \in \mathbb{N}}$  (avec nos notations antérieures,  $(\tilde{Y}_n)_{n \in \mathbb{N}}$  a même loi que  $(Y_{\kappa n})_{n \in \mathbb{N}}$ , si  $Y_0 = \tilde{Y}_0$ ), que  $(Z_n)_{n \in \mathbb{N}}$  soit une chaîne de Markov homogène de matrice de transition  $P^{(\beta)}$  et qu'en tout instant  $n \in \mathbb{N}$ , on soit presque sûrement assuré de  $\underline{W}(\tilde{Y}_n) \leq Z_n$ . L'idée sous-jacente est que partant de l'inégalité précédente, si  $\underline{W}(\tilde{Y}_{n+1})$  passe au-dessus de  $Z_n$ , il est possible grâce à (40) de distribuer une « partie » des  $Z_{n+1}$  au-dessus de  $\underline{W}(\tilde{Y}_{n+1})$  et si  $Z_{n+1} = 0$ , on peut par (39) de distribuer une « partie » des  $\tilde{Y}_{n+1}$  sur  $\underline{U}$ , le tout en respectant les transitions marginales. Nous n'entrerons pas en détail dans cette construction, qui est relativement fastidieuse, mais on renvoie au livre de Lindvall [19] pour une exposition générale de cette méthodologie.

Pour un  $\beta \geq B$  donné, on aura remarqué que la matrice  $P^{(\beta)}$  est irréductible sur  $\Omega$  et elle admet donc une unique probabilité invariante, disons  $\gamma^{(\beta)}$ . La forme des  $P^{(\beta)}$ , pour  $\beta \geq B$ , étant assez particulière, on n'a pas trop de mal à calculer le quasi-potentiel correspondant (en utilisant par exemple la caractérisation de Trouvé [27]) :

$$\forall \omega \in \Omega, \quad \lim_{\beta \rightarrow +\infty} \beta^{-1} \ln(\gamma^{(\beta)}(\omega)) = -\omega.$$

En ayant recours plus précisément la représentation de  $\gamma^{(\beta)}$  en termes d'arbres pointés, il apparaît qu'il existe deux constantes  $a', b' > 0$  telles que

$$\forall \omega \in \Omega, \quad a' \exp(-\beta\omega) \leq \gamma^{(\beta)}(\omega) \leq b' \exp(-\beta\omega).$$

Supposons maintenant que  $\beta_{\mathbb{N}^*}$  soit croissante avec  $\lim_{n \rightarrow \infty} \beta_n > B$  et soit  $n_0 \in \mathbb{N}$  tel que  $\beta_{n_0} \geq B$ . Fixons en second lieu  $x \in \underline{U}$  et  $n \in \mathbb{N}$  avec  $n \geq n_0$ . En appliquant le couplage précédent avec  $\tilde{Y}_0 = x$  et  $Z_0$  distribué comme  $\gamma^{(\beta_n)}$ , on obtient que pour tout  $\omega \in \Omega$  et tout  $p \in \mathbb{N}$ ,

$$\begin{aligned} \mathbb{P}_{n,x}[\underline{W}(Y_{n+\kappa p}) = \omega] &\leq \gamma^{(\beta_n)}(\{\omega' \in \Omega : \omega' \geq \omega\}) \\ &\leq b' \text{card}(E) \exp(-\beta_n \omega). \end{aligned}$$

On a donc montré le résultat voulu, mais seulement à des temps régulièrement espacés de  $\kappa$ . Cependant, en écrivant que pour tous  $p \in \mathbb{N}$ ,  $y \in E$  et  $1 \leq \lambda < \kappa$ , on a

$$\mathbb{P}_{n,x}[Y_{n+\kappa p+\lambda} = y] = \sum_{z \in E} \mathbb{P}_{n,x}[Y_{n+\kappa p} = z] (K^{(\beta_{n+\kappa p+1})} \dots K^{(\beta_{n+\kappa p+\lambda})})(z, y)$$

et en utilisant que d'après la proposition 21, il existe une constante  $C > 0$  telle que pour tous  $n, p \in \mathbb{N}$  et  $1 \leq \lambda < \kappa$ ,

$$\forall y, z \in E, \quad (K^{(\beta_{n+\kappa p+1})} \dots K^{(\beta_{n+\kappa p+\lambda})})(z, y) \leq C \exp(-\beta_n (\underline{W}(z) - \underline{W}(y))_+)$$

on obtient la borne escomptée, du moins pour tous les entiers  $n \in \mathbb{N}$  assez grands,  $n \geq n_0$ . Mais pour  $n < n_0$ , on a  $\beta_n \leq B$  ainsi l'inégalité voulue est trivialement satisfaite avec  $C = \exp(B \max_E \underline{W})$ . Cette observation permet aussi de traiter le cas où  $\lim_{n \rightarrow \infty} \beta_n \leq B$  et termine la preuve de la proposition 26.  $\square$

Pour récapituler où nous en sommes, la borne (34) a donc été prouvée sous la condition (H) et pour toute évolution  $\beta_{\mathbb{N}^*}$  du type (3). Notre objectif suivant va consister à y remplacer le chapeau par un tilde. Pour ceci, remarquons que l'on a en fait prouvé une version un peu plus forte : on a trouvé deux constantes  $n_0 \in \mathbb{N}^*$  et  $K \geq 0$  (dépendant notamment du  $k > 0$  de (3)), telles que

$$\sup_{x \in E} \sup_{n \geq n_0} \left\| \frac{\mathcal{L}(Y_n, \mathbb{P}_{n - \lfloor K \ln(n) \rfloor, x})}{\mu^{(\beta_n)}} \right\| < +\infty \quad (41)$$

où  $\mathcal{L}(Y_n, \mathbb{P}_{n-\lfloor K \ln(n) \rfloor, x})$  désigne la loi de  $Y_n$  sous la probabilité  $\mathbb{P}_{n-\lfloor K \ln(n) \rfloor, x}$ . L'intertemps  $\lfloor K \ln(n) \rfloor$  mesure une sorte d'intervalle de relaxation, inhomogène et « non-tendue » au sens de la norme supérieure. Pour exploiter ceci, remarquons que comme les  $\tilde{m}_n$ , pour  $n \in \mathbb{N}$ , correspondent à un flot de Feynman-Kac, on peut les représenter à l'aide d'intégrales de Feynman-Kac renormalisées : par exemple pour tous  $n \geq n_0$  et  $y \in E$ ,

$$\tilde{m}_n(y) = \frac{\sum_{x \in E} \tilde{m}_{n-\lfloor K \ln(n) \rfloor}(x) \mathbb{E}_{n-\lfloor K \ln(n) \rfloor, x} [\prod_{n-\lfloor K \ln(n) \rfloor \leq q < n} J_{\beta_q, \beta_{q+1}}(Y_q) \mathbb{1}_{\{y\}}(Y_n)]}{\sum_{x \in E} \tilde{m}_{n-\lfloor K \ln(n) \rfloor}(x) \mathbb{E}_{n-\lfloor K \ln(n) \rfloor, x} [\prod_{n-\lfloor K \ln(n) \rfloor \leq q < n} J_{\beta_q, \beta_{q+1}}(Y_q)]} \quad (42)$$

où pour tous  $\beta, \beta' \geq 0$ ,  $J_{\beta, \beta'}$  désigne la fonction  $M[H^{(\beta')}] / M[H^{(\beta)}]$ . Il nous faut donc maintenant estimer ces fonctions, ce qui va nous reconduire à des considérations « statiques ».

Soit  $x_1 \in \mathcal{U}_1$  dorénavant fixé et dénotons par  $\mathcal{G}(x_1) := \{g(x_1) : g \in \mathcal{G}\}$  son orbite sous  $\mathcal{G}$ , dont le cardinal est compris entre  $s$  et  $\text{card}(\mathcal{G})$ . Soit également  $Y = (Y_n)_{n \in \mathbb{N}}$  une chaîne de Markov homogène de transitions données par  $K^{(\beta)}$ , pour  $\beta \geq 0$  également fixé. On pose

$$\tau(Y) := \inf\{n \in \mathbb{N} : Y_n \in \mathcal{G}(x_1)\}.$$

Nous aurons besoin de l'estimation préliminaire suivante, dont la démonstration se trouve déjà en filigrane de la preuve du lemme 25.

**Lemme 27.** *Il existe  $\epsilon > 0$  et  $C > 0$  tels que pour tous  $\beta \geq 0$  et  $x \in E$ ,*

$$\mathbb{E}_x^{(\beta)}[\exp(\epsilon \tau(Y))] \leq C$$

où  $\mathbb{E}_x^{(\beta)}$  désigne l'espérance relative à la probabilité  $\mathbb{P}_x^{(\beta)}$  sous laquelle  $(Y_n)_{n \in \mathbb{N}}$  est une chaîne de Markov homogène de matrice de transition  $K^{(\beta)}$  et issue de  $x$ .

*Démonstration.* En reprenant les notations introduites dans la preuve de la proposition 26, on remarque qu'il existe  $N \in \mathbb{N}^*$  tel que pour tout  $y \in E$ , on puisse trouver  $1 \leq i \leq s$ ,  $x_i \in \mathcal{G}(x_1) \cap \mathcal{U}_i$  et  $0 \leq n \leq N$  tels que  $\underline{W}_n(y, x_i) = 0$ . On déduit de cela et des estimations de la proposition 21 qu'il existe  $\epsilon > 0$  tel que

$$\forall \beta \geq 0, \forall x \in E, \quad \mathbb{P}_x^{(\beta)}[\tau(Y) \leq N] \geq \epsilon$$

d'où pour tout  $p \in \mathbb{N}$ ,

$$\mathbb{P}_x^{(\beta)}[\tau(Y) \geq pN] \leq (1 - \epsilon)^p$$

puis le résultat annoncé. □

Ceci va nous permettre de justifier les calculs suivants.

**Proposition 28.** *Pour tout  $\beta \geq 0$ , considérons le vecteur propre  $H^{(\beta)}$  renormalisé par la condition  $H^{(\beta)}(x_1) = 1$ . Alors pour tout  $x \in E$  fixé, l'application*

$$\mathbb{R}_+ \ni \beta \mapsto H^{(\beta)}(x)$$

*est partout dérivable à droite. De plus, on a pour cette dérivée la borne suivante :*

$$\sup_{x \in E} \sup_{\beta \geq 0} \left| \frac{\partial_{\beta}^+ H^{(\beta)}(x)}{H^{(\beta)}(x)} \right| < +\infty.$$

*Démonstration.* Par la symétrie de notre cadre induite par (H), on a pour tout  $\beta \geq 0$  et tout  $z \in \mathcal{G}(x_1)$ ,  $H^{(\beta)}(z) = 1$ . La fin de la section 2 nous permet alors de voir que pour tout  $\beta \geq 0$ ,

$$\forall x \in E, \quad H^{(\beta)}(x) = \mathbb{E}_x \left[ \prod_{0 \leq n < \tau(X)} \frac{G^{(\beta)}(X_n)}{\lambda^{(\beta)}} \right]$$

(rappelons que  $X = (X_n)_{n \in \mathbb{N}}$  est une chaîne de Markov homogène de matrice de transition  $M$ ).

Pour étudier la dérivabilité en  $\beta \geq 0$  de cette expression, commençons par remarquer que l'application  $\mathbb{R}_+ \ni \beta \mapsto \ln(\lambda^{(\beta)}) \in \mathbb{R}$  est convexe d'après la formule (14). Elle est donc dérivable à droite partout sur  $]0, +\infty[$  et sa dérivée satisfait

$$\begin{aligned} \forall \beta > 0, \quad \left| \partial_\beta^+ \ln(\lambda^{(\beta)}) \right| &\leq \sup_{\eta \in \mathcal{P}(E)} |\eta(U)| \\ &\leq \|U\|_\infty \end{aligned} \tag{43}$$

ce qui permet aussi de voir que la dérivabilité précédente et la borne ci-dessus s'étendent à tout  $\mathbb{R}_+$ . Il s'ensuit que pour toute réalisation  $(X_n)_{n \in \mathbb{N}}$  telle que  $\tau(X) < +\infty$  (ce qui est satisfait  $\mathbb{P}_x$ -p.s.), l'application

$$\mathbb{R}_+ \ni \beta \mapsto \prod_{0 \leq n < \tau(X)} \frac{G^{(\beta)}(X_n)}{\lambda^{(\beta)}}$$

est dérivable à droite, de dérivée donnée en tout  $\beta \geq 0$  par

$$- \prod_{0 \leq n < \tau(X)} \frac{G^{(\beta)}(X_n)}{\lambda^{(\beta)}} \left( \sum_{0 \leq p < \tau(X)} U(X_p) + \partial_\beta^+ \ln(\lambda^{(\beta)}) \right).$$

Or pour tout  $\beta_0 \geq 0$  donné, on peut trouver une constante  $C > 0$  telle que pour tout  $\beta \geq 0$  au voisinage de  $\beta_0$ , l'expression ci-dessus soit bornée en valeur absolue par  $\prod_{0 \leq n < \tau(X)} \frac{G^{(\beta_0)}(X_n)}{\lambda^{(\beta_0)}} \exp(\epsilon \tau(X))$ . Cependant, par un calcul que l'on a déjà souvent effectué dans la section 2, on a

$$\begin{aligned} \mathbb{E}_x \left[ \prod_{0 \leq n < \tau(X)} \frac{G^{(\beta_0)}(X_n)}{\lambda^{(\beta_0)}} \exp(\epsilon \tau(X)) \right] &= \mathbb{E}_x \left[ \prod_{0 \leq n < \tau(X)} \frac{H^{(\beta_0)}(X_n)}{M[H^{(\beta_0)}](X_n)} \exp(\epsilon \tau(X)) \right] \\ &= H^{(\beta_0)}(x) \mathbb{E}_x \left[ \prod_{0 \leq n < \tau(X)} \frac{H^{(\beta_0)}(X_{n+1})}{M[H^{(\beta_0)}](X_n)} \exp(\epsilon \tau(X)) \right] \\ &= H^{(\beta_0)}(x) \mathbb{E}_x^{(\beta_0)} [\exp(\epsilon \tau(Y))] \end{aligned}$$

ainsi d'après le lemme précédent on est en droit d'appliquer un théorème de dérivation (à droite) sous l'espérance pour obtenir que pour tout  $\beta \geq 0$ ,

$$\partial_\beta^+ H^{(\beta)}(x) = H^{(\beta)}(x) \mathbb{E}_x^{(\beta)} \left[ \sum_{0 \leq p < \tau(Y)} U(Y_p) + \partial_\beta^+ \ln(\lambda^{(\beta)}) \right]. \tag{44}$$

Une nouvelle application du lemme 27 et la borne (43) permettent de conclure.  $\square$

Nous disposons désormais de tous les ingrédients nécessaires pour prouver le remplacement des chapeaux par des tildes dans (34) :

**Proposition 29.** *Supposons la condition (H) satisfaite et que l'évolution  $\beta_{\mathbb{N}^*}$  soit de la forme (3) avec  $k > 0$ . On est alors assuré de*

$$\sup_{n \in \mathbb{N}^*} \left\| \frac{\tilde{m}_n}{\mu^{(\beta_n)}} \right\|_{\infty} < +\infty.$$

*Démonstration.* En effet, l'estimée de la proposition 28 implique par intégration qu'il existe une constante  $C > 0$  telle que

$$\forall \beta, \beta' \geq 0, \forall x \in E, \quad \exp(-C |\beta - \beta'|) \leq \frac{H^{(\beta)}(x)}{H^{(\beta')}(x)} \leq \exp(C |\beta - \beta'|)$$

bornes qui se transmettent immédiatement à l'application  $J_{\beta, \beta'}$  :

$$\forall \beta, \beta' \geq 0, \forall x \in E, \quad \exp(-C |\beta - \beta'|) \leq J_{\beta, \beta'}(x) \leq \exp(C |\beta - \beta'|). \quad (45)$$

Il apparaît maintenant clairement que pour  $n \in \mathbb{N}$  grand, l'opération de sélection  $\Psi^{(\beta_n, \beta_{n+1})}$  est très proche de Id. Notamment, on a pour toute trajectoire  $(Y_p)_{p \in \mathbb{N}}$  et  $n \in \mathbb{N}$  assez grand

$$\begin{aligned} \left| \ln \left( \prod_{n - \lfloor K \ln(n) \rfloor \leq q < n} J_{\beta_q, \beta_{q+1}}(Y_q) \right) \right| &\leq \sum_{n - \lfloor K \ln(n) \rfloor \leq q < n} |\ln(\|J_{\beta_q, \beta_{q+1}}\|_{\infty})| \\ &\leq C \sum_{n - \lfloor K \ln(n) \rfloor \leq q < n} |\beta_{q+1} - \beta_q| \\ &\leq C(\beta_n - \beta_{n - \lfloor K \ln(n) \rfloor}) \\ &= Ck^{-1} \ln \left( \frac{n}{n - \lfloor K \ln(n) \rfloor} \right) \\ &\leq 2Ck^{-1} K \frac{\ln(n)}{n} \\ &\leq C' \end{aligned}$$

pour une constante  $C' > 0$  adéquate, indépendante de  $n \in \mathbb{N}$  assez grand.

La borne annoncée en découle immédiatement, si l'on tient compte de (41) et de (42).  $\square$

Nous pouvons maintenant procéder à la

*Preuve du théorème 1.* Par définition, on a

$$\begin{aligned} \forall n \in \mathbb{N}, \quad m_n &= (R^{(\beta_n)})^{-1}(\tilde{m}_n) \\ &= \frac{1}{\tilde{m}_n[1/M[H^{(\beta_n)}]]} \frac{1}{M[H^{(\beta_n)}]} \cdot \tilde{m}_n. \end{aligned} \quad (46)$$

Pour pouvoir exploiter cette formule, il faut tout d'abord estimer  $\tilde{m}_n[1/M[H^{(\beta_n)}]]$ . La minorer n'est pas difficile, car d'après la proposition 21, il existe une constante  $c > 0$  telle que

$$\forall \beta \geq 0, \forall x \in E, \quad \frac{1}{M[H^{(\beta)}](x)} \geq c$$



d'où

$$\forall n \in \mathbb{N}, \quad \tilde{m}_n[1/M[H^{(\beta_n)}]] \geq c.$$

Pour une majoration, outre à la proposition 21, recourons à la proposition précédente pour dénicher une constante  $C > 0$  telle que

$$\begin{aligned} \forall n \in \mathbb{N}, \quad \tilde{m}_n[1/M[H^{(\beta_n)}]] &\leq C\mu^{(\beta_n)}[1/M[H^{(\beta_n)}]] \\ &\leq C' \sum_{x \in E} \exp(\beta_n A(x)) \exp(-\beta_n \underline{W}(x)) \\ &= C' \sum_{x \in E} \exp(-\beta_n W(x)) \\ &\leq C' \text{card}(E) \end{aligned}$$

pour une constante  $C' > 0$  bien choisie. Ainsi, en réutilisant l'identité  $W = A - \underline{W}$ , on obtient à partir de (46) et pour une autre constante  $C > 0$ , la majoration

$$\forall n \in \mathbb{N}, \forall y \in E, \quad m_n(y) \leq C\nu^{(\beta_n)}(y)$$

d'où la propriété de concentration

$$\lim_{n \rightarrow \infty} m_n[\bar{\mathcal{U}}] = 1.$$

Montrer que  $m_n$  charge asymptotiquement pour  $n$  grand tous les points de  $\bar{\mathcal{U}}$  est un peu plus délicat. Revenons pour cela à la proposition 29, car celle-ci montre que

$$\lim_{n \rightarrow \infty} \tilde{m}_n[\underline{\mathcal{U}}] = 1.$$

Supposons de plus que  $m_0 = \tilde{m}_0$  soit invariant par  $\mathcal{G}$ . Puisque toutes nos manipulations ne créent que des objets symétriques par rapport à l'action du groupe  $\mathcal{G}$ , on en déduit qu'à tout instant  $n \in \mathbb{N}$ ,  $\tilde{m}_n$  est préservé par tous les éléments de  $\mathcal{G}$ , d'où

$$\forall 1 \leq i \leq s, \quad \lim_{n \rightarrow \infty} \tilde{m}_n[\mathcal{U}_i] = \frac{1}{s}.$$

Il en découle que pour tout  $n \in \mathbb{N}$  assez grand, on peut trouver  $x_1(n) \in \mathcal{U}_1$  tel que  $\tilde{m}_n(x_1(n)) \geq 1/(2\text{card}(E))$ . Cependant, comme  $M_{\mathcal{U}_1}$  est apériodique, il existe  $p \in \mathbb{N}^*$  tel que pour tous  $x, y \in \mathcal{U}_1$ ,  $(M_{\mathcal{U}_1})^p(x, y) > 0$  et même d'après la proposition 21,

$$\min_{x, y \in \mathcal{U}_1} \inf_{\beta'_1, \dots, \beta'_p \geq 0} (K^{(\beta'_1)} \dots K^{(\beta'_p)})(x, y) > 0.$$

Ainsi en écrivant pour tout  $y \in \mathcal{U}_1$ ,

$$\begin{aligned} \tilde{m}_{n+p}(y) &= \frac{\sum_{x \in E} \tilde{m}_n(x) \mathbb{E}_{n,x} [\prod_{n \leq q < n+p} J_{\beta_q, \beta_{q+1}}(Y_q) \mathbb{1}_{\{y\}}(Y_{n+p})]}{\sum_{x \in E} \tilde{m}_n(x) \mathbb{E}_{n,x} [\prod_{n \leq q < n+p} J_{\beta_q, \beta_{q+1}}(Y_q)]} \\ &\geq \frac{\tilde{m}_n(x_1(n)) \mathbb{E}_{n, x_1(n)} [\prod_{n \leq q < n+p} J_{\beta_q, \beta_{q+1}}(Y_q) \mathbb{1}_{\{y\}}(Y_{n+p})]}{\sum_{x \in E} \tilde{m}_n(x) \mathbb{E}_{n,x} [\prod_{n \leq q < n+p} J_{\beta_q, \beta_{q+1}}(Y_q)]} \end{aligned}$$

et en utilisant les estimées de  $J_{\beta_q, \beta_{q+1}}$ ,  $q \in \mathbb{N}$ , vues précédemment pour une évolution  $\beta_{\mathbb{N}^*}$  de la forme (3), il ressort que l'on peut trouver  $N \in \mathbb{N}$  assez grand de sorte que

$$\min_{x \in \mathcal{U}_1} \inf_{n \geq N} \tilde{m}_n(x) > 0$$

et par symétrie on peut ensuite prendre le minimum sur  $\underline{\mathcal{U}}$ . En se rappelant de (46) et du fait que  $W$  s'annule sur  $\underline{\mathcal{U}}$ , ceci admet pour conséquence que

$$\min_{x \in \underline{\mathcal{U}}} \inf_{n \geq N} m_n(x) > 0.$$

Pour prolonger cette propriété à tout  $\overline{\mathcal{U}}$ , on procède un peu comme ci-dessus : on écrit que pour tous  $n, p \in \mathbb{N}$  tels que  $n - p \geq 0$  et tout  $y \in E$ ,

$$m_n(y) = \frac{\sum_{x \in E} m_{n-p}(x) \mathbb{E}_{n-p, x} [\prod_{n-p < q \leq n} G_{\beta_q}(X_q) \mathbb{1}_{\{y\}}(X_n)]}{\sum_{x \in E} m_{n-p}(x) \mathbb{E}_{n-p, x} [\prod_{n-p < q \leq n} G_{\beta_q}(X_q)]}$$

(où  $\mathbb{E}_{n-p, x}$  correspond évidemment ici aussi à une espérance conditionnée par  $X_{n-p} = x$ ), et on utilise le fait que pour tout  $y \in \overline{\mathcal{U}}$ , on peut trouver un  $M$ -chemin  $y_0 \rightarrow y_1 \rightarrow \dots \rightarrow y_q = y$  restant dans  $\mathcal{U}$  avec  $y_0 \in \underline{\mathcal{U}}$ . On aboutit ainsi à la conclusion que

$$\forall x \in \overline{\mathcal{U}}, \quad \liminf_{n \rightarrow \infty} m_n(x) > 0 \quad (47)$$

du moins sous l'hypothèse que  $m_0$  est invariante sous l'action de  $\mathcal{G}$ .

Pour passer au cas général, on remarque que par irréductibilité de  $M$  et apériodicité de  $M_{\mathcal{U}_1}$ , la matrice  $M$  est elle-même apériodique. On en déduit facilement que l'on peut trouver  $p \in \mathbb{N}$  et une constante  $C \geq 1$  tels que

$$l/C \leq m_p \leq Cl$$

où  $l$  est la probabilité équidistribuée sur  $E$ . Renommons  $l$  en  $l_p$ , puis considérons  $(l_n)_{n \geq p}$  le flot à valeurs dans  $\mathcal{P}(E)$  que l'on obtient en faisant agir la dynamique de Feynman-Kac recuite à partir de l'instant  $p$ . On a donc pour tout  $n \geq p$ ,

$$\forall x \in E, \quad l_n(x) := \frac{\mathbb{E}_{p, l_p} [\exp(-\beta_{p+1}U(X_{p+1}) - \beta_{p+2}U(X_{p+2}) - \dots - \beta U(X_n)) \mathbb{1}_{\{x\}}(X_n)]}{\mathbb{E}_{p, l_p} [\exp(-\beta_{p+1}U(X_{p+1}) - \beta_{p+2}U(X_{p+2}) - \dots - \beta U(X_n))]}$$

et la même formule est valable si l'on y substitue les  $l$  par des  $m$ . Il en découle, par positivité des expressions considérées, que

$$\forall n \geq p, \quad l_n/C^2 \leq m_n \leq C^2 l_n.$$

Or on peut appliquer les résultats précédents à la suite  $(l_n)_{n \geq p}$  et on en déduit donc que (47) est juste quelque soit la condition initiale  $m_0$ .  $\square$

**Remarque 30.** Si l'on avait continué à travailler à partir de l'inégalité (37), on aurait encore pu y remplacer le chapeau par un tilde, par les mêmes manipulations que celles considérées dans la preuve de la proposition 29. Les arguments ci-dessus nous aurait ensuite aussi conduit à  $\sup_{n \in \mathbb{N}} \tilde{m}_n[1/M[H^{(\beta_n)}]] / \ln^{L-1}(2+n) < +\infty$ . Par contre, cette seule information ne permet pas d'aboutir à la conclusion (47).

Intéressons-nous maintenant à l'aspect oublié de la condition initiale envisagé dans le théorème 3. Pour cela, revenons à la famille de noyaux  $(K^{(\beta)})_{\beta \geq 0}$  et aux suites  $(\hat{m}_n)_{n \geq 0}$  correspondantes, associées à certaines évolutions

$\beta_{\mathbb{N}^*}$  d'inverse de la température. Comme nous l'avons déjà rappelé, il s'agit des marginales temporelles d'un algorithme de recuit simulé généralisé. Or pour celles-ci, la propriété d'oubli de la condition initiale est relativement bien connue : il existe une constante critique  $k_2 \geq 0$ , qu'il est possible d'expliciter en termes de la fonction de coût correspondante (mais de manière générale il faut se placer sur  $E^2$  et y considérer la famille de noyaux  $(K^{(\beta)} \otimes K^{(\beta)})_{\beta \geq 0}$ , voir le cours de Catoni [3]), telle que pour des évolutions  $\beta_{\mathbb{N}^*}$  de la forme (3) avec  $k > k_2$ , on a  $\lim_{n \rightarrow \infty} \widehat{m}_n - \widehat{m}'_n = 0$  quelque soit les probabilités initiales  $\widehat{m}_0$  et  $\widehat{m}'_0$ , alors que si  $0 < k < k_2$ , on peut trouver deux telles conditions initiales de sorte que le rapprochement précédent n'ait pas lieu en temps grand.

Dans notre cadre, l'apériodicité supposée de  $M_{\mathcal{U}_1}$  permet de voir que

$$k_2 = \max_{x \in E} \min_{p \in \mathcal{C}(x, x_0)} e_{\underline{W}}(p) - \underline{W}(x) \quad (48)$$

pour tout  $x_0 \in \underline{\mathcal{U}}$  fixé, où  $\mathcal{C}(x, x_0)$  désigne l'ensemble des chemins  $p := (p_0 \rightarrow p_1 \rightarrow \dots \rightarrow p_n)$  allant de  $x$  à  $x_0$ , et où l'élévation  $e_{\underline{W}}(p)$  d'un tel chemin est définie par

$$e_{\underline{W}}(p) := \max_{0 \leq i < n} \underline{W}(p_i) + \underline{W}(p_i, p_{i+1}).$$

La forme particulière de  $\underline{W}$  nous permet d'aller un peu plus loin : rappelons que  $s_U$  est la fonctionnelle définie sur les chemins par  $s_U(p) = U(p_0) + U(p_1) + \dots + U(p_n)$ , si  $p = (p_0 \rightarrow p_1 \rightarrow \dots \rightarrow p_n)$ . Notons pour  $1 \leq i \neq j \leq s$ ,

$$h_{s_U}(i, j) := \min_{p \in \mathcal{C}(\mathcal{U}_i, \mathcal{U}_j)} s_U(p)$$

où  $\mathcal{C}(\mathcal{U}_i, \mathcal{U}_j)$  est l'ensemble des chemins allant d'un point de  $\mathcal{U}_i$  à un point de  $\mathcal{U}_j$ . Fixons ensuite  $1 \leq i_0 \leq s$  et dénotons pour  $1 \leq i \neq i_0 \leq s$ ,  $\mathcal{C}(i_0, i)$  l'ensemble des suites finies d'éléments de  $\{1, \dots, s\}$  ne se recoupant pas,  $\iota := (i_0, i_1, \dots, i_n)$  avec  $n \in \mathbb{N}$  et  $i_n = i$ . Si  $\iota$  est un tel chemin d'indices, on pose

$$h_{s_U}(\iota) := \max_{0 \leq j < n} h_{s_U}(i_j, i_{j+1})$$

puis

$$k'_2 := \max_{1 \leq i \neq i_0 \leq s} \min_{\iota \in \mathcal{C}(i_0, i)} h_{s_U}(\iota).$$

Le résultat suivant permet de caractériser la constante  $k_2$  plus directement en termes du potentiel  $U$  :

**Lemme 31.** *Sous l'hypothèse de symétrie fournie par (H), la constante  $k'_2$  ne dépend pas du choix de  $1 \leq i_0 \leq n$  et on a*

$$k_2 = k'_2.$$

*Démonstration.* Il est classique et facile à retrouver que dans la définition (48), il suffit de considérer un maximum en  $x$  sur les minima locaux associés à la fonctionnelle de coût  $\underline{W}$ , or dans notre situation ceux-ci ne sont autres que les minima globaux de l'application  $\underline{W}$ , dont l'ensemble est  $\underline{\mathcal{U}}$ . Ainsi on a

$$k_2 = \max_{x \in \underline{\mathcal{U}}} \min_{p \in \mathcal{C}(x, x_0)} e_{\underline{W}}(p).$$

(pour tout  $x_0 \in \underline{\mathcal{U}}$  fixé), et il y apparaît notamment que  $k_2 = 0$  si et seulement si  $s = 1$ . Remarquons d'ailleurs que si  $s = 1$ , alors la formule présentée dans la proposition ci-dessus s'annule également (par un abus de langage traditionnel, un « minimum » sur un « ensemble vide de nombre positifs » est pris égal à zéro, pour éviter ce type de convention, nous aurions dû définir d'une manière plus appropriée  $\mathcal{C}(1, 1)$ ). Il suffit donc de considérer le cas où  $s > 1$  et  $k_2 > 0$ . Notons également que si  $1 \leq i \neq j \leq s$  et  $g \in \mathcal{G}$  sont donnés et si  $1 \leq i' \neq j' \leq s$  sont

tels que  $g(\mathcal{U}_i) = \mathcal{U}_{i'}$  et  $g(\mathcal{U}_j) = \mathcal{U}_{j'}$ , alors par symétrie du problème, on est assuré que  $h_{s_U}(i, j) = h_{s_U}(i', j')$ . On en déduit ensuite facilement par l'hypothèse (H) que la constante  $k'_2$  ne dépend pas du choix de  $i_0 \in \{1, \dots, s\}$ .

Après ces quelques préliminaires, commençons par nous convaincre que  $k'_2 \geq k_2$ . En reprenant l'expression de  $k_2$  donnée juste ci-dessus, avec  $x_0 \in \underline{\mathcal{U}}$  fixé, soient  $x \in \underline{\mathcal{U}}$  et  $p \in \mathcal{C}(x_0, x)$  réalisant respectivement le maximum puis le minimum  $y$  apparaissant. Soient ensuite  $1 \leq i_0, i \leq s$  tels que  $x_0 \in \mathcal{U}_{i_0}$  et  $x \in \mathcal{U}_i$ . On a nécessairement  $i_0 \neq i$ , sinon on aurait  $k_2 = 0$ . Par définition, il apparaît que  $k'_2 \geq \min_{\iota \in \mathcal{C}(i_0, i)} h_{s_U}(\iota)$ . Considérons un chemin  $\iota := (i_0, \dots, i_n)$  où est atteint ce dernier minimum, puis pour  $0 \leq l < n$ , soit  $q_l \in \mathcal{C}(\mathcal{U}_{i_l}, \mathcal{U}_{i_{l+1}})$  un chemin tel que  $h_{s_U}(i_l, i_{l+1}) = s_U(q_l)$ . En effectuant des raccords à l'intérieur des  $\mathcal{U}_{i_l}$ , pour  $0 \leq l \leq n$ , entre ces chemins, on peut obtenir un chemin  $q$  allant de  $x_0$  à  $x$  et tel que  $e_W(q) = \max_{0 \leq l < n} e_W(q_l)$  (les raccords ne comptant pas, car ils sont effectués à une élévation nulle). Ainsi, puisque  $e_W(q) \geq e_W(p) = k_2$ , pour obtenir  $k'_2 \geq k_2$  il suffit de vérifier que pour tout chemin  $r := (r_0 \rightarrow \dots \rightarrow r_m) \in \mathcal{C}(\mathcal{U}_j, \mathcal{U}_k)$ , avec  $m \in \mathbb{N}$  et  $1 \leq j, k \leq s$ , on a  $s_U(r) \geq e_W(r)$ . Mais ceci est clair, car pour tout  $0 \leq l < m$ , on a

$$\begin{aligned} \underline{W}(r_l) + \underline{W}(r_l, r_{l+1}) &= A(r_l) + W(r_l) + V(r_{l+1}) - A(r_l) \\ &= W(r_l) + V(r_{l+1}) \\ &\geq (U(r_0) + \dots + U(r_l)) + (U(r_{l+1}) + \dots + U(r_m)). \end{aligned}$$

L'inégalité réciproque  $k_2 \geq k'_2$  est un peu plus délicate à obtenir. Commençons par considérer l'ensemble  $A$  des  $i \in \{1, \dots, s\}$  tels que, soit  $i = 1$ , soit  $\min_{\iota \in \mathcal{C}(1, i)} h_{s_U}(\iota) < k'_2$ . Posons aussi  $B := \{1, \dots, s\} \setminus A$ , qui est non vide par définition de  $k'_2$ . On a donc pour tous  $i \in A$  et  $j \in B$ ,  $h(i, j) \geq k'_2$ . Remarquons qu'inversement pour tous  $i \in B$  et  $j \in A$ ,  $h(i, j) \geq k'_2$ . En effet, soient sinon  $i \in B$  et  $j \in A$  tels que  $h(i, j) < k'_2$ , puis  $g \in \mathcal{G}$  avec  $g(\mathcal{U}_i) = \mathcal{U}_j$ . En s'intéressant, comme à la fin de la preuve de la proposition 19, aux itérés de  $g$ , on finit par trouver  $q \in \mathbb{N}^*$ ,  $i' \in A$  et  $j' \in B$  tels que  $U_{i'} = g^q(U_i)$  et  $U_{j'} = g^q(U_j)$ , ce qui implique de manière contradictoire que  $h_{s_U}(i', j') < k'_2$ . Notons  $\mathcal{U}_A := \sqcup_{a \in A} \mathcal{U}_a$  et  $\mathcal{U}_B := \sqcup_{b \in B} \mathcal{U}_b$  et soit  $p$  un chemin allant d'un point de  $\mathcal{U}_A$  à un point de  $\mathcal{U}_B$  ou inversement. Nous allons vérifier que  $e_W(p) \geq k'_2$ , ce qui fournira l'inégalité  $k_2 \geq k'_2$ , puis permettra de conclure à l'identité annoncée.

Soit donc  $P$  l'ensemble des chemins allant d'un point de  $\mathcal{U}_A$  à un point de  $\mathcal{U}_B$  ou d'un point de  $\mathcal{U}_B$  à un point de  $\mathcal{U}_A$ , puis choisissons dans  $P$  un chemin  $p := (p_0, \dots, p_n)$  d'élévation  $e_W(p)$  minimale. Par symétrie de la situation, supposons par exemple que  $p$  va de  $\mathcal{U}_A$  à  $\mathcal{U}_B$ . Pour  $0 \leq l \leq n$ , soit

$$\begin{aligned} a(l) &:= \min_{q \in \mathcal{C}(p_l, \mathcal{U}_A)} s_U(q) \\ b(l) &:= \min_{q \in \mathcal{C}(p_l, \mathcal{U}_B)} s_U(q) \end{aligned}$$

on a ainsi  $V(p_l) = a(l) \wedge b(l)$ , puis posons

$$L := \min\{0 \leq l < n : b(l+1) \leq a(l+1)\}$$

(on aura remarqué que  $a(0) = 0 < b(0)$  et que  $b(n) = 0 > a(n)$ ). Notre objectif immédiat est de montrer que

$$W(p_L) = \min_{q \in \mathcal{C}(\mathcal{U}_A, p_L)} s_U(q).$$

En effet, si cela ne devait pas être vrai, il existerait  $q' \in \mathcal{C}(\mathcal{U}_B, p_L)$  avec  $s_U(q') = W(p_L)$ . Cependant, par définition de  $L$ , on a  $V(p_L) = a(L) < b(L)$ , or  $b(L)$  est lui-même plus petit que  $\min_{q \in Q} s_U(q)$ , où  $Q$  est l'ensemble des chemins allant de  $p_L$  à  $\mathcal{U}_B$  et dont le second élément est  $p_{L+1}$ . Ceci permet de voir que  $V(p_L) < U(p_L) + V(p_{L+1})$ . Soit maintenant  $q'' \in \mathcal{C}(p_L, \mathcal{U}_A)$  réalisant le minimum dans la définition de  $V(p_L)$  et notons  $q$  le chemin obtenu en concaténant  $q'$  avec  $q''$  (et en prenant soin de supprimer l'un des  $p_L$  « au milieu »).

Il apparaît immédiatement que

$$\begin{aligned}
e_{\underline{W}}(q) &\leq s_U(q) \\
&= s_U(q') + s_U(q'') - U(p_L) \\
&= W(p_L) + V(p_L) - U(p_L) \\
&< W(p_L) + V(p_{L+1}) \\
&\leq e_{\underline{W}}(p)
\end{aligned}$$

ce qui est absurde par choix de  $p$ , car  $q$  est un chemin allant de  $\mathcal{U}_B$  à  $\mathcal{U}_A$ . Ainsi on a nécessairement

$$\begin{aligned}
W(p_L) &= \min_{q \in \mathcal{C}(\mathcal{U}_A, p_L)} s_U(q) \\
V(p_{L+1}) &= \min_{q \in \mathcal{C}(p_{L+1}, \mathcal{U}_B)} s_U(q)
\end{aligned}$$

ce qui nous conduit à considérer les chemins  $q' \in \mathcal{C}(\mathcal{U}_A, p_L)$  et  $q'' \in \mathcal{C}(p_{L+1}, \mathcal{U}_B)$  où sont respectivement atteints ces minima. Soit également  $q \in \mathcal{C}(\mathcal{U}_A, \mathcal{U}_B)$  la concaténation de ces deux chemins, on a en fin de compte

$$k'_2 \leq s_U(q) = W(p_L) + V(p_{L+1}) \leq e_{\underline{W}}(p). \quad \square$$

Nous sommes maintenant en mesure d'amorcer une comparaison entre  $k_2$  et la constante  $k_1$  définie en (5) dans l'introduction, relativement à des algorithmes de recuit simulé réversibles. Soit  $\mathcal{V}$  l'ensemble des minima locaux de  $U$ . Pour tout  $y_0 \in \underline{\mathcal{U}}$  fixé, on peut écrire  $k_1 = k'_1 \vee k''_1$  avec

$$\begin{aligned}
k'_1 &:= \max_{x \in \underline{\mathcal{U}}} h_U(x, y_0) \\
k''_1 &:= \max_{x \in \mathcal{V} \setminus \underline{\mathcal{U}}} h_U(x, y_0) - U(x)
\end{aligned}$$

qui correspondent à l'influence des hauteurs de puits associés respectivement à des éléments de  $\underline{\mathcal{U}}$  et de  $\mathcal{V} \setminus \underline{\mathcal{U}}$ . De manière générale,  $k_2$  sera plus grand (et souvent bien plus grand) que  $k'_1$ , car on y a substitué  $U$  par  $\underline{W}$  (qui domine  $U$ , puisque construit à partir de sommes de valeurs correspondantes de ce potentiel). Ainsi dans la situation où c'est entre des éléments de  $\underline{\mathcal{U}}$  que l'on rencontre la plus grande difficulté de communication (*i.e.*  $k_1 = k'_1$ ), on sera assuré que  $k_2 \geq k_1$ . Dans le cas contraire où  $k_1 = k''_1$ , la comparaison est plus malaisée, par exemple on pourra avoir  $k_1$  bien plus grand que  $k_2$  quand  $s = 1$ !

Notre intérêt pour cette constante  $k_2$  provient de l'impression que nous avons qu'elle devrait jouer le même rôle pour les dynamiques de Feynman-Kac recuites que pour les algorithmes de recuit simulé correspondants :

**Conjecture 32.** Supposons la condition (H) remplie et que l'évolution  $\beta_{\mathbb{N}^*}$  soit de la forme (3) avec  $k > 0$ . La constante  $k_2$  décrite ci-dessus est alors aussi critique pour la propriété de perte de mémoire de la dynamique de Feynman-Kac recuite associée.

Nous nous ne prouverons ce résultat dans certaines situations particulières, ce qui fournira la preuve du théorème 3, remettant le cas général à une éventuelle étude ultérieure.

Commençons par considérer le cas où  $k > k_2$ . Une première étape consiste à transférer la propriété d'oubli de la condition initiale des  $(\hat{m}_n)_{n \in \mathbb{N}}$  vers les  $(\tilde{m}_n)_{n \in \mathbb{N}}$ . Pour cela, on va avoir besoin d'un renforcement de cette propriété, indiquant quantitativement la vitesse d'oubli. Une des manières les plus commodes de la mesurer est d'utiliser des entropies relatives et d'avoir recours à des inégalités de Sobolev logarithmiques modifiées.

Plus précisément, à tout noyau markovien irréductible  $P$  sur  $E$  et à toute probabilité  $\mu \in \mathcal{P}(E)$  chargeant tous les points, on associe  $\alpha(P, \mu) \in \mathbb{R}_+$  la plus petite constante telle que pour toute fonction  $f \in \mathcal{B}(E)$  vérifiant

$\mu[f^2] = 1$ , on ait

$$\text{Ent}(f^2 \cdot \mu | \mu) \leq \frac{1}{\alpha(P, \mu)} \mu[f^2(I - PP^*)[\ln(f^2)]]$$

où  $P^*$  est l'opérateur agissant sur  $\mathcal{B}(E)$  à travers la formule

$$\forall f \in \mathcal{B}(E), \forall x \in E, \quad P^*[f](x) = \frac{(f \cdot \mu)P}{\mu P}(x)$$

(si  $\mu P(x) = 0$ , on convient que  $P^*[f](x) = 1$ , plus généralement, dans toute la suite une densité de Radon-Nikodym discrète sera toujours prise égale à 1 en des points où le dénominateur s'annule). Un des intérêts de cette quantité est que l'on a vu dans [7] qu'il existait une constante universelle  $0 < \rho < 1$  telle que pour toute probabilité  $m \in \mathcal{P}(E)$ , on soit assuré de la propriété de contraction

$$\text{Ent}(mP | \mu P) \leq (1 - \rho\alpha(P, \mu))\text{Ent}(m, \mu).$$

Par ailleurs, dans [20], on a estimé pour  $N \in \mathbb{N}^*$ , des constantes telles que  $\alpha((K^{(\beta)})^N, \mu^{(\beta)})$  à basse température : on a montré que l'on pouvait trouver  $N \in \mathbb{N}^*$  et  $C > 0$  tels que

$$\forall \beta \geq 0, \quad \alpha((K^{(\beta)})^N, \mu^{(\beta)}) \geq C(\beta^{-1} \wedge 1) \exp(-k_2\beta).$$

Cependant, on calcule que pour tout noyau markovien  $P$ , tout  $\mu \in \mathcal{P}(E)$  et tout  $f \in \mathcal{B}(E)$ ,

$$\mu[f^2(I - PP^*)[\ln(f^2)]] = \frac{1}{2} \sum_{x, y \in E} (f^2(y) - f^2(x))(\ln(f(y)) - \ln(f(x))) \sum_{z \in E} P(x, z)P(y, z) \frac{\mu(y)}{\mu P(z)}.$$

Il découle de cette expression et de la proposition 21 que si  $\beta_{\mathbb{N}}$  est une évolution de la forme (3), alors quitte à modifier un peu la constante  $C > 0$ , on a pour l'entier  $N$  fixé comme ci-dessus,

$$\forall n \in \mathbb{N}, \quad \alpha(K^{(\beta_{n+1})} K^{(\beta_{n+2})} \dots K^{(\beta_{n+N})}, \mu^{(\beta_n)}) \geq C(\beta_n^{-1} \wedge 1) \exp(-k_2\beta_n)$$

(par exemple avec la convention  $\beta_0 = 0$ ). Ces rappels nous conduisent au comportement suivant :

**Lemme 33.** *Supposons la condition (H) satisfaite et que l'évolution  $\beta_{\mathbb{N}^*}$  soit de la forme (3), avec  $k > k_2$ , on a alors pour toute condition initiale  $\tilde{m}_0 = m_0$ ,*

$$\lim_{n \rightarrow \infty} \left\| \tilde{m}_n - \mu^{(\beta_n)} \right\|_{\text{vt}} = 0.$$

*Démonstration.* D'après les estimations décrites ci-dessus, il existe une constante  $C > 0$  telle que pour tout  $n \in \mathbb{N}$ ,

$$\text{Ent}(\hat{m}_{n+N} | \mu^{(\beta_n)} K^{(\beta_{n+1})} K^{(\beta_{n+2})} \dots K^{(\beta_{n+N})}) \leq (1 - C(\beta_n^{-1} \wedge 1) \exp(-k_2\beta_n)) \text{Ent}(\hat{m}_n | \mu^{(\beta_n)}).$$

Cependant, d'après la proposition 28 et en utilisant la représentation en termes d'arbres pointés des probabilités invariantes  $\mu^{(\beta)}$ , pour  $\beta \geq 0$ , on peut trouver une autre constante  $C' > 0$  telle que

$$\forall \beta, \beta' \geq 0, \forall x \in E, \quad \left| \ln \left( \frac{\mu^{(\beta)}(x)}{\mu^{(\beta')}(x)} \right) \right| \leq C' |\beta - \beta'|.$$

Ainsi puisque pour tout  $n \in \mathbb{N}$ , on a

$$\begin{aligned}
& \left| \text{Ent}(\widehat{m}_{n+N} | \mu^{(\beta_n)} K^{(\beta_{n+1})} K^{(\beta_{n+2})} \dots K^{(\beta_{n+N})}) - \text{Ent}(\widehat{m}_{n+N} | \mu^{(\beta_{n+N})}) \right| \\
&= \left| \sum_{x \in E} \ln \left( \frac{\mu^{(\beta_{n+N})}(x)}{\mu^{(\beta_n)} K^{(\beta_{n+1})} K^{(\beta_{n+2})} \dots K^{(\beta_{n+N})}(x)} \right) \widehat{m}_{n+N}(x) \right| \\
&\leq \sum_{1 \leq p \leq N} \left| \sum_{x \in E} \ln \left( \frac{\mu^{(\beta_{n+p})} K^{(\beta_{n+p+1})} K^{(\beta_{n+p+2})} \dots K^{(\beta_{n+N})}(x)}{\mu^{(\beta_{n+p-1})} K^{(\beta_{n+p})} K^{(\beta_{n+p+1})} \dots K^{(\beta_{n+N})}(x)} \right) \widehat{m}_{n+N}(x) \right| \\
&= \sum_{1 \leq p \leq N} \left| \sum_{x \in E} \ln \left( \frac{\mu^{(\beta_{n+p})} K^{(\beta_{n+p})} K^{(\beta_{n+p+1})} \dots K^{(\beta_{n+N})}(x)}{\mu^{(\beta_{n+p-1})} K^{(\beta_{n+p})} K^{(\beta_{n+p+1})} \dots K^{(\beta_{n+N})}(x)} \right) \widehat{m}_{n+N}(x) \right| \\
&\leq \sum_{1 \leq p \leq N} \ln \left( \left\| \frac{\mu^{(\beta_{n+p})}}{\mu^{(\beta_{n+p-1})}} \right\|_{\infty} \vee \left\| \frac{\mu^{(\beta_{n+p-1})}}{\mu^{(\beta_{n+p})}} \right\|_{\infty} \right) \\
&\leq C' \sum_{1 \leq p \leq N} |\beta_{n+p} - \beta_{n+p-1}| \\
&= C' |\beta_{n+N} - \beta_n| \\
&\leq C'' \frac{1}{n+N}
\end{aligned}$$

pour une nouvelle constante  $C'' > 0$  bien choisie. Ceci nous amène à poser pour tout  $n \in \mathbb{N}$ ,

$$\begin{aligned}
a_n &:= C(\beta_n^{-1} \wedge 1) \exp(-k_2 \beta_n) \\
b_{n+1} &:= \frac{C''}{n+1}
\end{aligned}$$

de sorte que

$$\text{Ent}(\widehat{m}_{n+N} | \mu^{(\beta_{n+N})}) \leq b_{n+N} + (1 - a_n) \text{Ent}(\widehat{m}_n | \mu^{(\beta_n)})$$

puis par récurrence, pour tous  $n, p \in \mathbb{N}^*$ ,

$$\begin{aligned}
\text{Ent}(\widehat{m}_{n+pN} | \mu^{(\beta_{n+pN})}) &\leq b_{n+pN} + b_{n+(p-1)N} (1 - a_{n+(p-1)N}) + \dots \\
&\quad + \prod_{0 \leq q \leq p-1} (1 - a_{n+qN}) \text{Ent}(\widehat{m}_n | \mu^{(\beta_n)}).
\end{aligned}$$

Cependant, puisque  $\lim_{n \rightarrow \infty} a_n = 0$ , on peut trouver  $n_0$  assez grand tel que pour tous  $n \geq n_0$  et  $p \in \mathbb{N}^*$ ,

$$\begin{aligned}
\prod_{0 \leq q \leq p-1} (1 - a_{n+qN}) &= \exp \left( \sum_{0 \leq q \leq p-1} \ln(1 - a_{n+qN}) \right) \\
&\leq \exp \left( - \sum_{0 \leq q \leq p-1} a_{n+qN} / 2 \right) \\
&\leq \exp \left( -(C/2) p \beta_{n+pN}^{-1} \exp(-k_2 \beta_{n+pN}) \right) \\
&\leq \exp \left( -C^{(3)} p \ln(n + pN) / (n + pN)^a \right)
\end{aligned}$$

pour une constante  $C^{(3)} > 0$  appropriée et avec  $a := k_2/k \in [0, 1[$ . Il apparaît ainsi que

$$\begin{aligned} \text{Ent}(\widehat{m}_{n+pN} | \mu^{(\beta_{n+pN})}) &\leq \frac{C''}{n+N} \left( 1 + (p-1) \exp \left( -C^{(3)} p \ln(n+pN)/(n+pN)^a \right) \right) \\ &\quad + \exp \left( -C^{(3)} p \ln(n+pN)/(n+pN)^a \right) \text{Ent}(\widehat{m}_n | \mu^{(\beta_n)}). \end{aligned}$$

D'autre part, pour tout  $n \in \mathbb{N}$ , par convexité de l'application  $\mathcal{P}(E) \ni m \mapsto \text{Ent}(m | \mu^{(\beta_n)})$ , on a

$$\begin{aligned} \sup_{m \in \mathcal{P}(E)} \text{Ent}(m | \mu^{(\beta_n)}) &= \sup_{x \in E} \text{Ent}(\delta_x | \mu^{(\beta_n)}) \\ &= \left\| \ln(\mu^{(\beta_n)}) \right\|_{\infty} \\ &\leq C^{(4)}(1 \vee \beta_n) \end{aligned}$$

pour une certaine constante  $C^{(4)} > 0$  (à nouveau grâce aux estimées de la Prop. 21).

Il résulte de ces calculs que si l'on pose par exemple pour tout  $n \in \mathbb{N}$ ,  $p_n := \lfloor n^{\frac{a+1}{2}} \rfloor$ , alors

$$\lim_{n \rightarrow \infty} \sup_{m \in \mathcal{P}(E)} \text{Ent}(m K^{(\beta_{n+1})} K^{(\beta_{n+2})} \dots K^{(\beta_{n+p_n N})} | \mu^{(\beta_{n+p_n N})}) = 0.$$

Mais sur l'intervalle de temps  $n, n+1, \dots, n+p_n N$ , les dynamiques de Feynman-Kac « tilde » et markovienne « chapeau » n'ont asymptotiquement pas le temps de différer beaucoup : plus précisément, pour tout  $m \in \mathcal{P}(E)$  on a,

$$\lim_{n \rightarrow \infty} \left\| \frac{m K^{(\beta_{n+1})} K^{(\beta_{n+2})} \dots K^{(\beta_{n+p_n N})}}{\Psi^{(\beta_{n+p_n N-1}, \beta_{n+p_n N})} [\dots \Psi^{(\beta_n, \beta_{n+1})} [m] K^{(\beta_{n+1})}] \dots K^{(\beta_{n+p_n N})}} - 1 \right\|_{\infty} = 0.$$

Ceci provient de la formulation

$$\begin{aligned} &\Psi^{(\beta_{n+p_n N-1}, \beta_{n+p_n N})} [\dots \Psi^{(\beta_n, \beta_{n+1})} [m] K^{(\beta_{n+1})}] \dots K^{(\beta_{n+p_n N})}(x) \\ &= \frac{\mathbb{E}_{n,m} [\prod_{n \leq q < n+p_n N} J_{\beta_q, \beta_{q+1}}(Y_q) \mathbb{1}_{\{x\}}(Y_{n+p_n N})]}{\mathbb{E}_{n,m} [\prod_{n \leq q < n+p_n N} J_{\beta_q, \beta_{q+1}}(Y_q)]} \end{aligned}$$

valable pour tous  $x \in E$  et  $m \in \mathcal{P}(E)$ , des estimées (45) sur les fonctions  $J_{\beta, \beta'}$ , pour  $\beta, \beta' \geq 0$ , déduites de la proposition 28, de la forme (3) de l'évolution  $\beta_{\mathbb{N}^*}$  et enfin de la convergence  $\lim_{n \rightarrow \infty} p_n N/n = 0$ . Il en découle que

$$\lim_{n \rightarrow \infty} \sup_{m \in \mathcal{P}(E)} \text{Ent}(\Psi^{(\beta_{n+p_n N-1}, \beta_{n+p_n N})} [\dots \Psi^{(\beta_n, \beta_{n+1})} [m] K^{(\beta_{n+1})}] \dots K^{(\beta_{n+p_n N})} | \mu^{(\beta_{n+p_n N})}) = 0$$

or pour tout  $n \in \mathbb{N}$ , on a

$$\widetilde{m}_{n+p_n N} = \Psi^{(\beta_{n+p_n N-1}, \beta_{n+p_n N})} [\dots \Psi^{(\beta_n, \beta_{n+1})} [\widetilde{m}_n] K^{(\beta_{n+1})}] \dots K^{(\beta_{n+p_n N})}$$

d'où en fin de compte

$$\lim_{n \rightarrow \infty} \text{Ent}(\widetilde{m}_{n+p_n N} | \mu^{(\beta_{n+p_n N})}) = 0.$$



Mais comme ci-dessus, on vérifie que si  $q : \mathbb{N} \rightarrow \mathbb{N}$  est une application croissante telle que  $\lim_{n \rightarrow \infty} q(n)/n = 0$ , alors on est assuré de

$$\begin{aligned} \lim_{n \rightarrow \infty} \sup_{n \leq p \leq n+q(n)} \left\| \frac{d\tilde{m}_p}{d\tilde{m}_n K^{(\beta_{n+1})} \dots K^{(\beta_p)}} - 1 \right\|_{\mathbb{L}^\infty(\tilde{m}_n K^{(\beta_{n+1})} \dots K^{(\beta_p)})} &= 0 \\ \lim_{n \rightarrow \infty} \sup_{n \leq p \leq n+q(n)} \left\| \frac{d\mu^{(\beta_p)}}{d\mu^{(\beta_n)} K^{(\beta_{n+1})} \dots K^{(\beta_p)}} - 1 \right\|_{\mathbb{L}^\infty(\mu^{(\beta_n)} K^{(\beta_{n+1})} \dots K^{(\beta_p)})} &= 0 \end{aligned}$$

et comme on a toujours  $\text{Ent}(\tilde{m}_n K^{(\beta_{n+1})} \dots K^{(\beta_p)} | \mu^{(\beta_n)} K^{(\beta_{n+1})} \dots K^{(\beta_p)}) \leq \text{Ent}(\tilde{m}_n | \mu^{(\beta_n)})$ , on en déduit facilement une véritable convergence

$$\lim_{n \rightarrow \infty} \text{Ent}(\tilde{m}_n | \mu^{(\beta_n)}) = 0$$

laquelle implique à son tour le résultat annoncé, *via* l'inégalité de Pinsker (voir par exemple [1]).  $\square$

La première assertion du théorème 3 en découle, du moins *a priori* si on y prend  $A = 0$ .

**Proposition 34.** *Supposons  $M$  irréductible,  $(H)$  satisfaite, que  $\overline{\mathcal{U}} = \underline{\mathcal{U}}$  et que le schéma d'inverse de température soit de la forme (3). Alors on a  $\lim_{n \rightarrow \infty} m_n - \nu^{(\beta_n)} = 0$  pour toute condition initiale  $m_0 \in \mathcal{P}(E)$ , dès que  $k > k_2$ .*

*Démonstration.* Pour passer du flot  $(\tilde{m}_n)_{n \in \mathbb{N}}$  à notre dynamique initiale  $(m_n)_{n \in \mathbb{N}}$ , nous allons évidemment réutiliser la formule (46). Il s'agit donc d'évaluer  $\tilde{m}_n[1/M[H^{(\beta_n)}]]$ , en montrant notamment qu'en temps  $n \in \mathbb{N}$  grand, cette quantité oublie la condition initiale  $m_0$ . Dans la preuve du théorème 1, nous avons déjà vu que pour  $x \notin \overline{\mathcal{U}}$ ,

$$\lim_{n \rightarrow \infty} \frac{\tilde{m}_n(x)}{M[H^{(\beta_n)}](x)} = 0.$$

Par ailleurs, d'après le lemme précédent, on a si  $x \in \underline{\mathcal{U}}$ , pour  $n \in \mathbb{N}$  grand,

$$\frac{\tilde{m}_n(x)}{M[H^{(\beta_n)}](x)} \sim \frac{\mu^{(\beta_n)}(x)}{M[H^{(\beta_n)}](x)}$$

expression qui est effectivement indépendante de la condition initiale  $m_0$ . Ainsi en tenant compte de l'hypothèse  $\underline{\mathcal{U}} = \overline{\mathcal{U}}$ , il apparaît que pour toutes conditions initiales  $m_0, m'_0 \in \mathcal{P}(E)$ , pour  $n \in \mathbb{N}$  grand,

$$\tilde{m}_n[1/M[H^{(\beta_n)}]] \sim \tilde{m}'_n[1/M[H^{(\beta_n)}]]$$

puis *via* (46), que

$$\begin{aligned} \forall x \notin \underline{\mathcal{U}}, \quad \lim_{n \rightarrow \infty} m_n(x) &= \lim_{n \rightarrow \infty} m'_n(x) = 0 \\ \forall x \in \underline{\mathcal{U}}, \quad m_n(x) &\sim m'_n(x) \end{aligned}$$

pour  $n \in \mathbb{N}$  grand, d'où la propriété de perte de mémoire annoncée. L'information un peu plus précise suivant laquelle  $\lim_{n \rightarrow \infty} m_n - \nu^{(\beta_n)} = 0$  provient de la proposition 5 et de l'observation que d'après les calculs précédents, on a pour  $n \in \mathbb{N}$  grand,

$$\tilde{m}_n \left[ 1/M \left[ H^{(\beta_n)} \right] \right] \sim \mu^{(\beta_n)} \left[ 1/M \left[ H^{(\beta_n)} \right] \right]. \quad \square$$

Le fait d'introduire des évolutions  $\beta_{\mathbb{N}^*}$  de la forme un peu plus générale envisagée dans le théorème 3 ne change fondamentalement rien aux arguments précédents. Par contre pour montrer que la condition initiale n'est pas entièrement oubliée si  $0 \leq k < k_2$ , nous allons devoir considérer une constante  $A \geq 0$  suffisamment

grande (bien que nous pensions que cette limitation soit uniquement « technique », le résultat devant être juste pour tout  $A \geq 0$ , comme c'est par exemple le cas en théorie du recuit simulé, cf. [21]). Nous aurons aussi besoin du renforcement suivant de la proposition 28 :

**Proposition 35.** *Soit  $x_1 \in \mathcal{U}_1$  fixé et pour  $\beta \geq 0$ , soit  $H^{(\beta)}$  un vecteur propre de Perron-Frobenius comme dans la proposition 28, i.e. normalisé par la condition  $H^{(\beta)}(x_1) = 1$ . Outre (H) vérifiée, supposons que  $\underline{\mathcal{U}} = \overline{\mathcal{U}}$ , alors il existe deux constantes  $\delta > 0$  et  $C > 0$  telles que*

$$\forall x \in \underline{\mathcal{U}}, \forall \beta \geq 0, \quad \left| \partial_\beta^+ \ln(H^{(\beta)}(x)) \right| \leq C \exp(-\delta\beta).$$

Pour prouver ceci, nous allons repartir de (44) et pour transformer cette expression, considérons à  $\beta \geq 0$  fixé, l'équation de Poisson

$$\begin{cases} K^{(\beta)}[\tilde{\phi}^{(\beta)}] - \tilde{\phi}^{(\beta)} = U - \mu^{(\beta)}[U] \\ \mu^{(\beta)}[\tilde{\phi}^{(\beta)}] = 0 \end{cases}$$

que nous savons admettre une unique solution  $\tilde{\phi}^{(\beta)} \in \mathcal{B}(E)$ , du fait de l'irréductibilité de  $K^{(\beta)}$  sur l'ensemble fini  $E$ . Cette unicité et l'invariance de l'équation précédente par le groupe  $\mathcal{G}$  des symétries de notre cadre impliquent que  $\tilde{\phi}^{(\beta)}$  est stable par l'action de  $\mathcal{G}$  :

$$\forall g \in \mathcal{G}, \quad \tilde{\phi}^{(\beta)} \circ g = \tilde{\phi}^{(\beta)}.$$

Considérons ensuite la fonction  $\phi^{(\beta)} := \tilde{\phi}^{(\beta)} - \tilde{\phi}^{(\beta)}(x_1)$  qui est donc l'unique solution de l'équation

$$\begin{cases} K^{(\beta)}[\phi^{(\beta)}] - \phi^{(\beta)} = U - \mu^{(\beta)}[U] \\ \phi^{(\beta)}(x_1) = 0. \end{cases}$$

On est assuré de l'estimée suivante :

**Lemme 36.** *En reprenant les notations en vigueur dans la proposition 28, on a pour tout  $\beta \geq 0$ ,*

$$\forall x \in E, \quad \phi^{(\beta)}(x) = -\mathbb{E}_x^{(\beta)} \left[ \sum_{0 \leq p < \tau(Y)} U(Y_p) - \mu^{(\beta)}[U] \right].$$

Si on suppose de plus que  $\underline{\mathcal{U}} = \overline{\mathcal{U}}$ , alors il existe deux constantes  $\delta > 0$  et  $C > 0$  telles que

$$\forall \beta \geq 0, \forall x \in \underline{\mathcal{U}}, \quad \left| \phi^{(\beta)}(x) \right| \leq C \exp(-\delta\beta).$$

*Démonstration.* En effet, écrivons que pour tout  $n \in \mathbb{N}$ ,

$$\begin{aligned} \sum_{0 \leq p < n} U(Y_p) - \mu^{(\beta)}[U] &= \sum_{0 \leq p < n} K^{(\beta)}[\phi^{(\beta)}](Y_p) - \phi^{(\beta)}(Y_p) \\ &= \phi^{(\beta)}(Y_n) - \phi^{(\beta)}(Y_0) - M_n \end{aligned}$$

avec

$$M_n := \sum_{0 \leq p < n} \phi^{(\beta)}(Y_{p+1}) - K^{(\beta)}[\phi^{(\beta)}](Y_p).$$

Or pour tous  $\beta \geq 0$  et  $x \in E$ , sous  $\mathbb{P}_x^{(\beta)}$ ,  $(Y_p)_{p \in \mathbb{N}}$  est une chaîne de Markov homogène admettant  $K^{(\beta)}$  pour matrice transition, ainsi  $(M_p)_{p \in \mathbb{N}}$  est une martingale issue de 0 sous ces probabilités et par le théorème d'arrêt de Doob, on en déduit que pour tout  $n \in \mathbb{N}$ ,

$$\mathbb{E}_x^{(\beta)}[M_{\tau(Y) \wedge n}] = 0$$

d'où aussi, puisque  $\phi^{(\beta)}$  s'annule sur l'orbite de  $x_1$  par  $\mathcal{G}$ ,

$$\begin{aligned} \mathbb{E}_x^{(\beta)} \left[ \sum_{0 \leq p < \tau(Y) \wedge n} U(Y_p) - \mu^{(\beta)}[U] \right] &= \mathbb{E}_x^{(\beta)} \left[ \phi^{(\beta)}(Y_{\tau(Y) \wedge n}) - \phi^{(\beta)}(Y_0) \right] \\ &= \mathbb{E}_x^{(\beta)} \left[ \phi^{(\beta)}(Y_n) \mathbb{1}_{n < \tau(Y)} \right] - \phi^{(\beta)}(x). \end{aligned}$$

Cependant les estimées du lemme 27 permettent de passer à la limite dans ces égalités quand  $n$  devient grand, pour montrer que

$$\mathbb{E}_x^{(\beta)} \left[ \sum_{0 \leq p < \tau(Y)} U(Y_p) - \mu^{(\beta)}[U] \right] = -\phi^{(\beta)}(x) \quad (49)$$

d'où le premier résultat annoncé.

Pour le second, il suffit de se convaincre qu'il existe deux constantes  $\delta > 0$  et  $C > 0$  telles que

$$\forall \beta \geq 0, \forall x \in \mathcal{U}_1, \quad \left| \phi^{(\beta)}(x) \right| \leq C \exp(-\delta\beta)$$

puisque cette borne s'étendra ensuite à tout  $\underline{U}$  par symétrie. Fixons donc  $x \in \mathcal{U}_1$  et introduisons

$$S_{\mathcal{U}_1} := \inf\{p \in \mathbb{N} : Y_p \notin \mathcal{U}_1\}.$$

Décomposons ensuite  $-\phi^{(\beta)}(x)$  en

$$\mathbb{E}_x^{(\beta)} \left[ \mathbb{1}_{\tau(Y) < S_{\mathcal{U}_1}} \sum_{0 \leq p < \tau(Y)} U(Y_p) - \mu^{(\beta)}[U] \right] + \mathbb{E}_x^{(\beta)} \left[ \mathbb{1}_{\tau(Y) \geq S_{\mathcal{U}_1}} \sum_{0 \leq p < \tau(Y)} U(Y_p) - \mu^{(\beta)}[U] \right].$$

On peut réécrire le premier terme sous la forme

$$-\mathbb{E}_x^{(\beta)} \left[ \mathbb{1}_{\tau(Y) < S_{\mathcal{U}_1}} \mu^{(\beta)}[U] \tau(Y) \right]$$

expression qui se majore en valeur absolue par  $\mu^{(\beta)}[U] \mathbb{E}_x^{(\beta)}[\tau(Y)]$ , elle-même bornée par  $C' \exp(-\delta'\beta)$  pour des constantes  $C', \delta' > 0$  convenablement choisies et indépendantes de  $x \in \mathcal{U}_1$  et de  $\beta \geq 0$  (car  $\underline{W}$  est strictement positif en dehors de  $\mathcal{U}$  et on peut prendre  $\delta' = \min\{\underline{W}(x) : x \notin \mathcal{U}\}$ ).

Intéressons-nous ensuite au second terme de (50) : en utilisant la propriété de Markov forte, il s'écrit aussi

$$\begin{aligned} -\mathbb{E}_x^{(\beta)} \left[ \mathbb{1}_{\tau(Y) \geq S_{\mathcal{U}_1}} \sum_{0 \leq p < S_{\mathcal{U}_1}} \mu^{(\beta)}[U] \right] &+ \mathbb{E}_x^{(\beta)} \left[ \mathbb{1}_{\tau(Y) \geq S_{\mathcal{U}_1}} \sum_{S_{\mathcal{U}_1} \leq p < \tau(Y)} U(Y_p) - \mu^{(\beta)}[U] \right] \\ &= -\mathbb{E}_x^{(\beta)} \left[ \mathbb{1}_{\tau(Y) > S_{\mathcal{U}_1}} \mu^{(\beta)}[U] S_{\mathcal{U}_1} \right] + \mathbb{E}_x^{(\beta)} \left[ \mathbb{1}_{\tau(Y) > S_{\mathcal{U}_1}} \phi^{(\beta)}(Y_{S_{\mathcal{U}_1}}) \right]. \end{aligned}$$

Comme ci-dessus, le premier de ces termes est borné par  $C' \exp(-\delta' \beta)$  avec les mêmes constantes  $C', \delta' > 0$ . Quant au second, on peut le majorer en valeur absolue par  $\mathbb{P}_x^{(\beta)}[\tau(Y) > S_{\mathcal{U}_1}] \|\phi^{(\beta)}\|_\infty$ . Remarquons d'après (49) et le lemme 27 que  $\sup_{\beta \geq 0} \|\phi^{(\beta)}\|_\infty < +\infty$ , il reste donc à se convaincre que la quantité  $\mathbb{P}_x^{(\beta)}[\tau(Y) > S_{\mathcal{U}_1}]$  décroît exponentiellement rapidement en l'inverse de température  $\beta$ . Mais ceci découle d'un calcul traditionnel en théorie des chaînes de Markov, en observant (grâce à la Prop. 21) d'une part qu'à l'intérieur de  $\mathcal{U}_1$ ,  $K^{(\beta)}$  est minoré uniformément en  $\beta \geq 0$  par un noyau sous-markovien irréductible et d'autre part que toutes les transitions  $K^{(\beta)}(x, y)$  avec  $x \in \mathcal{U}_1$  et  $y \notin \mathcal{U}_1$  décroissent exponentiellement rapidement en  $\beta \geq 0$  (ici intervient l'hypothèse que  $\underline{\mathcal{U}} = \overline{\mathcal{U}}$ ).  $\square$

Nous pouvons maintenant passer à la

*Preuve de la proposition 35.* On travaille à un inverse de température  $\beta \geq 0$  fixé. Comme annoncé, on reconsidère l'identité (44), mais on commence par la modifier légèrement : si l'on note le temps d'impact

$$\tilde{\tau}(Y) := \inf\{n \geq 1 : Y_n \in \mathcal{G}(x_1)\}$$

(qui ne diffère du temps d'atteinte  $\tau(Y)$  que si  $Y_0 \in \mathcal{G}(x_1)$ ), on a alors encore

$$\forall x \in E, \quad \partial_\beta^+ H^{(\beta)}(x) = H^{(\beta)}(x) \mathbb{E}_x^{(\beta)} \left[ \sum_{0 \leq p < \tilde{\tau}(Y)} U(Y_p) + \partial_\beta^+ \ln(\lambda^{(\beta)}) \right] \quad (50)$$

(cette égalité n'est d'ailleurs autre que (44) si  $x \notin \mathcal{G}(x_1)$ ). En effet, pour s'en convaincre il suffit de revenir à la fin de la section 2 : l'identité (15) et sa démonstration restent valables si l'on y remplace  $\tau$  par  $\tilde{\tau}$ , défini de manière similaire, sauf que l'infimum est pris sur les entiers strictement positifs. La preuve de (50) s'en suit comme précédemment.

L'intérêt de (50) sur (44) est qu'en y prenant  $x = x_1$ , on obtient une information non triviale :

$$\begin{aligned} 0 &= \partial_\beta^+ H^{(\beta)}(x_1) \\ &= H^{(\beta)}(x_1) \mathbb{E}_{x_1}^{(\beta)} \left[ \sum_{0 \leq p < \tilde{\tau}(Y)} U(Y_p) + \partial_\beta^+ \ln(\lambda^{(\beta)}) \right] \\ &= \mathbb{E}_{x_1}^{(\beta)} \left[ \sum_{0 \leq p < \tilde{\tau}(Y)} U(Y_p) \right] + \partial_\beta^+ \ln(\lambda^{(\beta)}) \mathbb{E}_{x_1}^{(\beta)} [\tilde{\tau}(Y)] \end{aligned}$$

*i.e.*

$$\partial_\beta^+ \ln(\lambda^{(\beta)}) = - \frac{\mathbb{E}_{x_1}^{(\beta)} \left[ \sum_{0 \leq p < \tilde{\tau}(Y)} U(Y_p) \right]}{\mathbb{E}_{x_1}^{(\beta)} [\tilde{\tau}(Y)]}.$$

Or en appliquant la propriété de Markov à un pas, on obtient, puisque  $U(x_1) = 0$ ,

$$\begin{aligned}
\mathbb{E}_{x_1}^{(\beta)} \left[ \sum_{0 \leq p < \tilde{\tau}(Y)} U(Y_p) \right] &= \sum_{y \in E} K^{(\beta)}(x_1, y) \mathbb{E}_y^{(\beta)} \left[ \sum_{0 \leq p < \tau(Y)} U(Y_p) \right] \\
&= \sum_{y \in E} K^{(\beta)}(x_1, y) \mathbb{E}_y^{(\beta)} \left[ \sum_{0 \leq p < \tau(Y)} U(Y_p) - \mu^{(\beta)}[U] \right] \\
&\quad + \mu^{(\beta)}[U] \sum_{y \in E} K^{(\beta)}(x_1, y) \mathbb{E}_y^{(\beta)} [\tau(Y)] \\
&= -K^{(\beta)}[\phi^{(\beta)}](x_1) + \mu^{(\beta)}[U] \sum_{y \in E} K^{(\beta)}(x_1, y) \mathbb{E}_y^{(\beta)} [\tau(Y)] \\
&= -(\phi^{(\beta)}(x_1) + U(x_1) - \mu^{(\beta)}[U]) + \mu^{(\beta)}[U] \sum_{y \in E} K^{(\beta)}(x_1, y) \mathbb{E}_y^{(\beta)} [\tau(Y)] \\
&= \mu^{(\beta)}[U] \mathbb{E}_{x_1}^{(\beta)} [\tilde{\tau}(Y)]
\end{aligned}$$

d'où en fin de compte,

$$\partial_\beta^+ \ln(\lambda^{(\beta)}) = -\mu^{(\beta)}[U] \quad (51)$$

puis

$$\forall x \in E, \quad \partial_\beta^+ \ln(H^{(\beta)}(x)) = -\phi^{(\beta)}(x).$$

La proposition 35 découle alors des estimées de la seconde partie du lemme 36.  $\square$

**Remarque 37.** Il existe diverses manières alternatives de montrer (51). Indiquons-en une qui fait tout quotienter par  $\mathcal{G}$ . Introduisons donc la relation d'équivalence  $\sim$  sur  $E$  définie par

$$\forall x, y \in E, \quad x \sim y \iff \exists g \in \mathcal{G} : y = g(x).$$

Sur l'ensemble quotient  $\check{E} := E / \sim$  (il s'agit de l'ensemble des différentes orbites possibles sous  $\mathcal{G}$ ), considérons le noyau markovien  $\check{M}$  donné par

$$\check{M}(\check{x}, \check{y}) := \sum_{y \in \check{y}} M(x, y)$$

où  $x \in E$  est un représentant quelconque de  $\check{x}$  (on vérifie aisément que la somme du membre de gauche de ne dépend pas de ce choix). De même, on pose  $\check{U}(\check{x}) = U(x)$ , pour tout représentant  $x \in E$  de  $\check{x} \in \check{E}$ .

Dans le nouveau cadre  $(\check{E}, \check{M}, \check{U})$ , on effectue toutes les opérations précédentes et les objets ainsi obtenus sont coiffés d'un  $\check{\phantom{x}}$ . On dispose ainsi pour tout  $\beta \geq 0$  de  $\check{K}^{(\beta)}$ ,  $\check{\mu}^{(\beta)}$  ou  $\check{\nu}^{(\beta)}$  et il est relativement facile de voir que ces derniers peuvent aussi se construire naturellement par passage au quotient à partir de leur homologues  $K^{(\beta)}$ ,  $\mu^{(\beta)}$  ou  $\nu^{(\beta)}$ . Intéressons-nous plus particulièrement à  $\check{\mu}^{(\beta)}$ , la probabilité invariante associée à  $\check{K}^{(\beta)}$ . Comme ce noyau est irréductible sur l'ensemble fini  $\check{E}$ , la chaîne correspondante est récurrente positive et il est bien connu que  $\check{\mu}^{(\beta)}$  peut se représenter de la manière suivante : pour tout  $\check{x} \in \check{E}$  fixé, on a

$$\forall \check{y} \in \check{E}, \quad \check{\mu}^{(\beta)}(\check{y}) = \frac{\mathbb{E}_{\check{x}}[\sum_{0 \leq n < \check{\tau}_{\check{x}}} \mathbb{1}_{\{\check{y}\}}(\check{Y}_n)]}{\mathbb{E}_{\check{x}}[\check{\tau}_{\check{x}}]}$$

où

$$\check{\tau}_{\check{x}} := \inf\{n \geq 1 : \check{Y}_n = \check{x}\}.$$

En revenant à notre cadre initial, cette formule montre, en y prenant  $x = x_1$  (on a alors  $\check{x}_1 = \mathcal{G}(x_1)$ ), que pour toute application  $\varphi \in \mathcal{B}(E)$  invariante par  $\mathcal{G}$  (i.e. telle que pour tout  $g \in \mathcal{G}$ ,  $\varphi \circ g = \varphi$ , c'est-à-dire que  $\varphi$  est constant sur les classes d'équivalence pour  $\sim$ ), on a

$$\mu^{(\beta)}[\varphi] = \frac{\mathbb{E}_{x_1}^{(\beta)} \left[ \sum_{0 \leq p < \check{\tau}(Y)} \varphi(Y_p) \right]}{\mathbb{E}_{x_1}^{(\beta)} [\check{\tau}(Y)]}.$$

En particulier, avec  $\varphi = U$ , on retrouve (51).

L'estimée de la proposition 35 va permettre de limiter fortement l'influence des sélections dans la dynamique des  $(\tilde{m}_n)_{n \in \mathbb{N}}$ . Pour aller dans cette direction, commençons par en déduire le petit résultat suivant.

**Lemme 38.** *Supposons (H) satisfaite et que  $\underline{\mathcal{U}} = \overline{\mathcal{U}}$ . Il existe alors deux constantes  $C, \delta > 0$  telles que*

$$\forall x \in \underline{\mathcal{U}}, \forall \beta + 1 \geq \beta' \geq \beta \geq 0, \quad \left| \frac{M[H^{(\beta')}]}{M[H^{(\beta)}]}(x) - 1 \right| \leq C \exp(-\delta\beta)(\beta' - \beta).$$

*Démonstration.* D'après la proposition 35, on a pour tout  $x \in \underline{\mathcal{U}}$ ,

$$\left| \ln \left( \frac{H^{(\beta')}}{H^{(\beta)}}(x) \right) \right| = \left| \int_{\beta}^{\beta'} \partial_u^+ \ln(H_u(x)) du \right| \leq C \exp(-\delta\beta)(\beta' - \beta)$$

avec les constantes  $C, \delta > 0$  définies là. Or on peut trouver une autre constante  $C_1 \geq 1$  telle que

$$\forall 0 \leq t \leq C, \quad |t - 1| \leq C_1 |\ln(t)|$$

et on en déduit que pour tout  $x \in \underline{\mathcal{U}}$  et tous  $\beta + 1 \geq \beta' \geq \beta \geq 0$ ,

$$\left| \frac{H^{(\beta')}}{H^{(\beta)}}(x) - 1 \right| \leq C_1 C \exp(-\delta\beta)(\beta' - \beta).$$

Par ailleurs, en tenant compte de (51) et du fait que son membre de droite converge vers 0 exponentiellement vite en  $\beta \geq 0$ , on montre de la même manière qu'il existe deux constantes  $C_2, \delta_2 > 0$  telles que pour tout  $x \in \underline{\mathcal{U}}$  et tous  $\beta + 1 \geq \beta' \geq \beta \geq 0$ ,

$$\left| \frac{\lambda^{(\beta')}}{\lambda^{(\beta)}} - 1 \right| \leq C_2 \exp(-\delta_2\beta)(\beta' - \beta).$$

Enfin, en utilisant que pour tous  $\beta \geq 0$  et  $x \in \underline{\mathcal{U}}$ , on a  $G^{(\beta)}(x) = 1$  et donc  $M[H^{(\beta)}](x) = \lambda^{(\beta)} H^{(\beta)}(x)$ , on en déduit facilement le résultat annoncé, avec de nouvelles constantes  $C, \delta > 0$  appropriées.  $\square$

Si par contre  $x \notin \underline{\mathcal{U}}$ , les bornes précédentes ne sont plus nécessairement valables. Cependant, en revenant directement à la proposition 28, il apparaît qu'il existe tout de même une constante  $C' > 0$  telle que

$$\forall x \notin \underline{\mathcal{U}}, \forall \beta + 1 \geq \beta' \geq \beta \geq 0, \quad \left| \frac{M[H^{(\beta')}]}{M[H^{(\beta)}]}(x) - 1 \right| \leq C'(\beta' - \beta). \quad (52)$$

Cette observation va permettre de se convaincre de la validité de la

**Proposition 39.** *Supposons (H) satisfaite et que  $\underline{\mathcal{U}} = \overline{\mathcal{U}}$ . Il existe à nouveau deux constantes  $C, \delta > 0$  telles que*

$$\forall \beta + 1 \geq \beta' \geq \beta \geq 0, \forall m \in \mathcal{P}(E), \quad \left\| \Psi^{(\beta, \beta')}(m) - m \right\|_{\text{vt}} \leq C \exp(-\delta\beta)(\beta' - \beta) \left\| \frac{dm}{d\mu^{(\beta)}} \right\|_{\infty}.$$

*Démonstration.* En effet, d'après le lemme précédent, il existe deux constantes  $C, \delta > 0$  telles que pour tous  $\beta + 1 \geq \beta' \geq \beta \geq 0$  et  $x \in \underline{\mathcal{U}}$ , on ait

$$\begin{aligned} |J_{\beta, \beta'}(x)m(x) - m(x)| &\leq |J_{\beta, \beta'}(x) - 1| \\ &\leq C \exp(-\delta\beta)(\beta' - \beta) \end{aligned}$$

(rappelons que  $J_{\beta, \beta'} = M[H^{(\beta)}]/M[(\beta)']$ ). Pour les  $x \notin \underline{\mathcal{U}}$ , utilisons plutôt (52) et on écrit

$$\begin{aligned} |J_{\beta, \beta'}(x)m(x) - m(x)| &\leq |J_{\beta, \beta'}(x) - 1| \frac{dm}{d\mu^{(\beta)}}(x) \mu^{(\beta)}(x) \\ &\leq C' \left\| \frac{dm}{d\mu^{(\beta)}} \right\|_{\infty} (\beta' - \beta) \mu^{(\beta)}(x) \end{aligned}$$

expression qui peut aussi se majorer par un terme de la forme  $C'' \exp(-\delta''\beta)(\beta' - \beta)$  pour de bons choix des constantes  $C'', \delta'' > 0$  (évidemment indépendantes de  $\beta + 1 \geq \beta' \geq \beta \geq 0$  et de  $m \in \mathcal{P}(E)$ ), car pour  $x \notin \underline{\mathcal{U}}$ , on a  $\underline{W}(x) > 0$ .

Puisque  $\Psi^{(\beta, \beta')}(m) = J_{\beta, \beta'} \cdot m / m[J_{\beta, \beta'}]$ , on en déduit sans difficulté la borne voulue.  $\square$

Nous pouvons maintenant préciser ce que nous entendions par des sélections peu influentes : supposons (H) toujours satisfaite, que  $\underline{\mathcal{U}} = \overline{\mathcal{U}}$  et considérons une évolution  $\beta_{\mathbb{N}^*}$  de la forme (3) avec  $k > 0$ . Les propositions 29 et 39 nous assurent alors que pour toute condition initiale  $\tilde{m}_0 \in \mathcal{P}(E)$ ,

$$\sum_{n \in \mathbb{N}} \left\| \Psi^{(\beta_n, \beta_{n+1})}(\tilde{m}_n) - \tilde{m}_n \right\|_{\text{vt}} < +\infty.$$

En fait, il est possible d'aller un peu plus loin et pour cela considérons justement des évolutions  $\beta_{\mathbb{N}^*}$  paramétrées par  $k > 0$  et  $A \geq 0$  comme dans le théorème 3. Soit  $\tilde{m}_0 \in \mathcal{P}(E)$ , en reprenant la preuve de la proposition 29, on se rend compte que l'on peut trouver une constante  $K > 0$  dépendant de  $k > 0$  mais pas de  $A \geq 0$  ni de  $\tilde{m}_0$ , telle que  $\sup_{n \in \mathbb{N}} \left\| \frac{\tilde{m}_n}{\mu^{(\beta_n)}} \right\|_{\infty} \leq K \left\| \frac{\tilde{m}_0}{\mu^{(\beta_0)}} \right\|_{\infty}$ . Par ailleurs, d'après la proposition 21, on a, si  $\tilde{m}_0 \in \mathcal{P}(\underline{\mathcal{U}})$ , que  $\sup_{A \geq 0} \left\| \frac{\tilde{m}_0}{\mu^{(\beta_0)}} \right\|_{\infty} < +\infty$ . Ainsi la proposition 39 permet d'obtenir que si  $\tilde{m}_0$  est portée par  $\underline{\mathcal{U}}$ , alors

$$\lim_{A \rightarrow +\infty} \sum_{n \in \mathbb{N}} \left\| \Psi^{(\beta_n, \beta_{n+1})}(\tilde{m}_n) - \tilde{m}_n \right\|_{\text{vt}} = 0 \quad (53)$$

(et ce, même uniformément en  $\tilde{m}_0 \in \mathcal{P}(\underline{\mathcal{U}})$ ), c'est-à-dire que pour  $A$  grand, les sélections deviennent effectivement négligeables pour la dynamique  $(\tilde{m}_n)_{n \in \mathbb{N}}$ . C'est heuristiquement la raison pour laquelle on peut dire que les comportements de cette dynamique de Feynman-Kac recuite (et par suite ceux de notre dynamique initiale  $(m_n)_{n \in \mathbb{N}}$ ) sont dictés par ceux de la suite  $(\hat{m}_n)_{n \in \mathbb{N}}$ , qui est plus facile à étudier puisqu'il s'agit des marginales temporelles d'un algorithme de recuit simulé généralisé.

Pour profiter de cette observation importante, plaçons-nous dans le cas où  $k_2 > 0$  (sinon, *i.e.* si  $s = 1$ , la seconde partie du théorème 3 ne contient aucune information) et soit

$$\mathcal{A}_1 := \left\{ x \in E : \min_{p \in \mathcal{C}(x_1, x)} e_{\underline{W}}(p) < k_2 \right\}.$$

Cet ensemble (appelé cycle en théorie du recuit simulé, voir par exemple [22, 27]) contient au moins  $\mathcal{U}_1$ , mais ne recouvre pas tous les  $\mathcal{U}_i$ , pour  $1 \leq i \leq s$ , par définition de  $k_2$ . Il admet aussi la propriété suivante

$$\forall x \in \mathcal{A}_1, \forall y \notin \mathcal{A}_1, \quad \underline{W}(x) + \underline{W}(x, y) \geq k_2 \quad (54)$$

et la caractérisation de Trouvé de  $\underline{W}$  montre que l'on a aussi

$$\forall x \notin \mathcal{A}_1, \forall y \in \mathcal{A}_1, \quad \underline{W}(x) + \underline{W}(x, y) \geq k_2. \quad (55)$$

Ceci va nous permettre de prouver la

**Proposition 40.** *On suppose toujours (H) satisfaite,  $\underline{\mathcal{U}} = \overline{\mathcal{U}}$  et que l'évolution  $\beta_{\mathbb{N}^*}$  est paramétrée par  $k > 0$  et  $A \geq 0$  comme dans le théorème 3, avec  $0 < k < k_2$ . Soit  $\tilde{m}_0 \in \mathcal{P}(\underline{\mathcal{U}} \cap \mathcal{A}_1)$  (par exemple  $\tilde{m}_0 \in \mathcal{P}(\mathcal{U}_1)$ ). On a alors*

$$\lim_{A \rightarrow +\infty} \inf_{n \in \mathbb{N}} \tilde{m}_n(\mathcal{A}_1) = 1.$$

*Démonstration.* On a clairement  $\tilde{m}_0(\mathcal{A}_1) = 1$  et pour se convaincre du résultat ci-dessus, il suffit de montrer que

$$\lim_{A \rightarrow +\infty} \sum_{n \in \mathbb{N}} |\tilde{m}_{n+1}(\mathcal{A}_1) - \tilde{m}_n(\mathcal{A}_1)| = 0.$$

Pour ceci, on majore respectivement chacun des éléments de la somme  $|\tilde{m}_{n+1}(\mathcal{A}_1) - \tilde{m}_n(\mathcal{A}_1)|$  par  $|\tilde{m}_{n+1}(\mathcal{A}_1) - \Psi^{(\beta_n, \beta_{n+1})}(\tilde{m}_n)(\mathcal{A}_1)| + |\Psi^{(\beta_n, \beta_{n+1})}(\tilde{m}_n)(\mathcal{A}_1) - \tilde{m}_n(\mathcal{A}_1)|$ , de sorte que d'après (53), il ne reste plus qu'à voir que

$$\lim_{A \rightarrow +\infty} \sum_{n \in \mathbb{N}} \left| \Psi^{(\beta_n, \beta_{n+1})}(\tilde{m}_n) K^{(\beta_{n+1})}(\mathcal{A}_1) - \Psi^{(\beta_n, \beta_{n+1})}(\tilde{m}_n)(\mathcal{A}_1) \right| = 0.$$

Pour chaque  $n \in \mathbb{N}$ , en effectuant un bilan de ce qui rentre et sort de  $\mathcal{A}_1$  par le biais de  $K^{(\beta_{n+1})}$ , on voit que

$$\begin{aligned} & \left| \Psi^{(\beta_n, \beta_{n+1})}(\tilde{m}_n) K^{(\beta_{n+1})}(\mathcal{A}_1) - \Psi^{(\beta_n, \beta_{n+1})}(\tilde{m}_n)(\mathcal{A}_1) \right| \\ & \leq \sum_{x \in \mathcal{A}_1, y \notin \mathcal{A}_1} \Psi^{(\beta_n, \beta_{n+1})}(\tilde{m}_n)(x) K^{(\beta_{n+1})}(x, y) + \sum_{x \notin \mathcal{A}_1, y \in \mathcal{A}_1} \Psi^{(\beta_n, \beta_{n+1})}(\tilde{m}_n)(x) K^{(\beta_{n+1})}(x, y). \end{aligned}$$

Cependant, en tenant compte des estimées (45), (54) et (55), de la proposition 21 et de la discussion qui précède (53), avec le fait que  $\tilde{m}_0 \in \mathcal{P}(\underline{\mathcal{U}} \cap \mathcal{A}_1)$ , il apparaît qu'il existe une constante  $K > 0$  (dépendant de  $k > 0$ , mais pas de  $A \geq 0$ ), telle

$$\begin{aligned} \left| \Psi^{(\beta_n, \beta_{n+1})}(\tilde{m}_n) K^{(\beta_{n+1})}(\mathcal{A}_1) - \Psi^{(\beta_n, \beta_{n+1})}(\tilde{m}_n)(\mathcal{A}_1) \right| & \leq K \exp(-k_2 \beta_{n+1}) \\ & = \frac{K}{(A + 1 + n)^{k/k_2}}. \end{aligned}$$

L'hypothèse que  $0 < k < k_2$  entre maintenant en jeu en permettant de conclure à la convergence escomptée.  $\square$



Il est aisé d'en déduire l'absence d'oubli de la condition initiale pour la dynamique  $(\tilde{m}_n)_{n \in \mathbb{N}}$ , si  $0 < k < k_2$  et si  $A \geq 0$  est assez grand (sous les conditions (H) et  $\underline{\mathcal{U}} = \overline{\mathcal{U}}$ ). En effet, remarquons que pour tout  $1 \leq i \leq s$ , si  $\mathcal{A}_1 \cap \mathcal{U}_i \neq \emptyset$ , alors  $\mathcal{U}_i \subset \mathcal{A}_1$ . Soit  $t := \text{card}(\{1 \leq i \leq s : \mathcal{A}_1 \cap \mathcal{U}_i \neq \emptyset\})$ , on a  $0 < t < s$ . Si  $\tilde{m}_0 \in \mathcal{P}(E)$  est invariante par  $\mathcal{G}$  (de telles probabilités existent, il suffit de considérer  $\sum_{g \in \mathcal{G}} m \circ g / \text{card}(\mathcal{G})$  avec  $m \in \mathcal{P}(E)$  quelconque), alors pour tout  $n \in \mathbb{N}$ ,  $\tilde{m}_n$  est aussi invariant par  $\mathcal{G}$ . Comme on sait de plus que  $\tilde{m}_n$  finit par se concentrer sur  $\underline{\mathcal{U}}$  en temps grand, la symétrie de notre cadre nous conduit à

$$\lim_{n \rightarrow \infty} \tilde{m}_n(\mathcal{A}_1) = \frac{t}{s},$$

ceci pour tous  $k > 0$  et  $A \geq 0$ . Puis considérons  $\tilde{m}'_0 \in \mathcal{P}(\mathcal{U}_1)$  et choisissons  $A \geq 0$  assez grand de sorte que  $\inf_{n \in \mathbb{N}} \tilde{m}'_n(\mathcal{A}_1) \geq 1/2 + t/(2s)$  ( $0 < k < k_2$  étant fixé). Il apparaît que

$$\begin{aligned} \liminf_{n \rightarrow \infty} \|\tilde{m}_n - \tilde{m}'_n\|_{\text{vt}} &\geq \liminf_{n \rightarrow \infty} \tilde{m}'_n(\mathcal{A}_1) - \tilde{m}_n(\mathcal{A}_1) \\ &\geq \frac{s-t}{2s} \\ &> 0. \end{aligned}$$

Ce résultat peut se retranscrire pour la dynamique de Feynman-Kac recuite  $(m_n)_{n \in \mathbb{N}}$  sans trop de difficulté. Toujours sous (H) et sous les hypothèses  $\underline{\mathcal{U}} = \overline{\mathcal{U}}$  et  $0 < k < k_2$ , on obtient de même que ci-dessus que si  $m_0$  est invariante par  $\mathcal{G}$ , alors

$$\lim_{n \rightarrow \infty} m_n(\mathcal{A}_1) = \frac{t}{s}.$$

Par contre, si  $m_0(\mathcal{U}_1) = 1$ , il apparaît à partir de (46), de la condition  $\underline{\mathcal{U}} = \overline{\mathcal{U}}$  et de la proposition 40 que

$$\lim_{A \rightarrow +\infty} \inf_{n \in \mathbb{N}} m_n(\mathcal{A}_1) = 1$$

ce qui permet de conclure comme ci-dessus et termine donc la preuve du théorème 3.

**Remarque 41.** L'idée d'utiliser un cycle maximal  $\mathcal{A}_1$  de « hauteur de sortie »  $k_2$  pour mettre en évidence l'absence de perte de mémoire est classique en recuit simulé. A la connaissance des auteurs, son origine remonte à l'article de Holley et Stroock [18]. L'observation que la stabilité de dynamiques de type Feynman-Kac est fortement liée à celles de certains processus markoviens que l'on peut naturellement leur associer se retrouve également dans un article récent de Stannat [26], dans un cadre homogène et continu en temps (mais l'auteur y promet des extensions à des situations inhomogènes).

#### 4. TRAITEMENT D'UN EXEMPLE ET CONJECTURES

Nous allons donner ici quelques indications relatives aux extensions proposées dans l'introduction. Ceci permettra de faire le point sur les propriétés qui nous manquent pour avancer dans leur direction.

Le premier problème qui se pose consiste à étendre la proposition 11, c'est-à-dire à montrer que pour tout  $x \in E$ , la quantité  $\beta^{-1} \ln(H^{(\beta)}(x))$  admet une limite, disons  $-\tilde{V}(x)$ , quand  $\beta$  devient grand. Comme on le verra ci-dessous, on ne sera plus assuré que  $\tilde{V}$  s'annule sur  $\underline{\mathcal{U}}$ , mais nous pensons que l'hypothèse  $\lambda(\mathcal{U}) < +\infty$  suffira à impliquer que l'ensemble des minima locaux de  $\tilde{V}$  est inclus dans  $\mathcal{U}$  (par contre trouver exactement les minima globaux de  $\tilde{V}$  devrait nécessiter des arguments eux-même « globaux », faisant intervenir certaines difficultés de communication entre les différents  $\mathcal{U}_i$ ,  $1 \leq i \leq s$ ). Il en découlera certainement relativement facilement que  $\mu^{(\beta)}$  et  $\nu^{(\beta)}$  se concentrent exponentiellement rapidement respectivement au moins sur  $\underline{\mathcal{U}}$  et  $\overline{\mathcal{U}}$  pour  $\beta \geq 0$  grand (attention, il n'y aura pas en général d'inclusion naturelle entre les véritables lieux de concentration

asymptotique de ces mesures, comme le montre le contre-exemple ci-dessous). Toutefois nous avons l'impression que pour ces deux seuls résultats, il doit exister une manière plus directe de s'en convaincre. Une seconde conséquence, plus intéressante à nos yeux, est que les minima locaux associés à la fonctionnelle de coût relative aux  $(K^{(\beta)})_{\beta \geq 0}$  sont tous inclus dans  $\underline{U}$ , ainsi les algorithmes de recuit correspondant (pour des évolutions de la température de la forme (3) avec  $k > 0$ ) se concentreront asymptotiquement sur  $\underline{U}$ . Pour passer aux dynamiques de Feynman-Kac recuites, il faudra, comme dans les sections précédentes, renforcer les convergences donnant naissance à  $\tilde{V}$ , en montrant par exemple qu'il existe  $C \geq 1$  tel que

$$\forall \beta \geq 1, \forall x \in E, \quad C^{-1} \exp(-\beta \tilde{V}(x)) \leq H^{(\beta)}(x) \leq C \exp(-\beta \tilde{V}(x))$$

et

$$\forall \beta \geq 1, \forall x \in E, \quad \left| \partial_{\beta}^+ \ln(H^{(\beta)}(x)) \right| \leq C$$

(ainsi que des améliorations de cette borne telles que celle présentée dans la Prop. 35). Les conjectures présentés dans l'introduction en découleraient alors en grande partie.

Mais en attendant de pouvoir éventuellement procéder à ces développements, intéressons-nous à un exemple ne satisfaisant pas l'hypothèse (H), pour lequel nous pouvons quand même calculer  $\tilde{V}$  et en déduire les lieux de concentration respectifs de  $\nu^{(\beta)}$  et  $\mu^{(\beta)}$  pour  $\beta \geq 0$  grand.

On prend  $E = \{1, 2, 3, 4\}$ , muni de la matrice markovienne

$$M = \begin{pmatrix} 1/2 & 1/2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1/2 & 1/2 \\ 1 & 0 & 0 & 0 \end{pmatrix}$$

et du potentiel

$$U = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 2 \end{pmatrix}.$$

On a clairement  $\mathcal{U} = \underline{U} = \{1, 3\}$ , avec par exemple  $\mathcal{U}_1 = \{1\}$  et  $\mathcal{U}_2 = \{3\}$ , car  $\lambda(\{1\}) = 1/2 = \lambda(\{1\})$ . Par contre (H) est bien mise en défaut, puisque l'inégalité  $U(2) \neq U(4)$  implique que le groupe  $\mathcal{G}$  des symétries de ce cadre est réduit à l'identité.

En prenant  $x_0 = 1$  dans la proposition 7, commençons par évaluer  $\lambda^{(\beta)}$ . Pour ceci calculons pour  $\beta \geq 0$  et  $\lambda > 1/2$ ,

$$\begin{aligned} \mathbb{E}_1 \left[ \prod_{0 \leq n < \tau_0} \frac{G^{(\beta)}(X_n)}{\lambda} \right] &= \frac{1}{2\lambda} + \frac{1}{2\lambda} \mathbb{E}_2 \left[ \prod_{0 \leq n < \tau_0} \frac{G^{(\beta)}(X_n)}{\lambda} \right] \\ &= \frac{1}{2\lambda} + \frac{1}{2\lambda} \frac{\exp(-3\beta)}{\lambda^2} \sum_{n \in \mathbb{N}^*} \frac{1}{(2\lambda)^n} \\ &= \frac{1}{2\lambda} + \frac{1}{2\lambda} \frac{\exp(-3\beta)}{\lambda^2} \frac{1}{2\lambda - 1} \end{aligned}$$

(le même calcul montre que l'espérance considérée est infinie si  $0 < \lambda \leq 1/2$ ). La valeur propre  $\lambda^{(\beta)}$  est la valeur de  $\lambda > 0$  pour laquelle cette expression vaut 1, *i.e.* est l'unique solution positive de l'équation

$$2\lambda = 1 + \frac{\exp(-3\beta)}{\lambda^2} \frac{1}{2\lambda - 1}.$$

Il y apparaît déjà que si  $\beta \rightarrow +\infty$ , alors  $\lambda^{(\beta)}$  doit nécessairement converger vers  $1/2$  (ce qui peut aussi se déduire de (21)). De plus, en écrivant  $\lambda^{(\beta)} = 1/2 + h^{(\beta)}$  et en effectuant la substitution dans l'égalité ci-dessus, on obtient que pour  $\beta \geq 0$  grand,

$$h^{(\beta)} \sim \exp\left(-\frac{3}{2}\beta\right).$$

Ceci et des calculs d'espérances similaires à celui effectué précédemment permettent de voir que pour  $\beta \geq 0$  grand,

$$H^{(\beta)}(2) \sim 2 \exp\left(-\frac{3}{2}\beta\right)$$

$$H^{(\beta)}(3) \sim \exp\left(-\frac{1}{2}\beta\right)$$

$$H^{(\beta)}(4) \sim 2 \exp(-2\beta)$$

d'où

$$\tilde{V} = \begin{pmatrix} 0 \\ 3/2 \\ 1/2 \\ 2 \end{pmatrix}.$$

On en déduit facilement la fonctionnelle de coût associée à la famille  $(K^{(\beta)})_{\beta \geq 0}$  et on obtient avec une notation évidente

$$\tilde{W} = \begin{pmatrix} 0 \\ 3/2 \\ 0 \\ 3/2 \end{pmatrix}$$

(on aura remarqué que dans cet exemple, pour chaque  $x \in \{1, 2, 3, 4\}$ , il n'y a qu'un seul  $x$ -arbre qui contribue dans (7)). Ainsi pour  $\beta \geq 0$  grand,  $\mu^{(\beta)}$  finit par charger tout  $\mathcal{U}$ . Cependant, des considérations toutes aussi immédiates (voir par exemple la Prop. 5) permettent de voir que

$$\tilde{W} = \begin{pmatrix} 1/2 \\ 3/2 \\ 0 \\ 2 \end{pmatrix}$$

d'où la concentration exponentiellement rapide pour  $\beta \geq 0$  grand de  $\nu^{(\beta)}$  sur le seul point 3.

Pour cet exemple particulier, du fait d'expressions quasi-explicites (on pourrait même tout expliciter, car  $\lambda^{(\beta)}$  est solution d'une équation algébrique du quatrième degré), nous disposons de toutes les estimations nécessaires pour mener à bien le programme décrit au début de cette section et il apparaît que pour toute évolution  $\beta_{N^*}$  de la forme (3) avec  $k > 0$ ,  $m_n$  finit par se concentrer sur 3 en temps grand. Notamment cette dynamique de

Feynman-Kac recuite finit par oublier sa condition initiale, pour tout  $k > 0$ , c'est-à-dire que dans la conjecture 4, on a  $k_2 = 0$ . Par contre remarquons qu'il n'en est pas de même pour les algorithmes de recuit simulé associés, car leur constante critique d'oubli vaut  $3/2$  (ainsi il n'y a pas d'espoir que la conjecture 32 se prolonge en toute généralité). Cependant pour tout  $k > 0$ , ils se concentrent effectivement sur tout  $\underline{U}$  (et ceci reste vrai pour toute valeur de  $U(3)$  prise dans  $\mathbb{R}_+^*$ ). S'agit-il là d'une propriété générale ?

### ANNEXE A. SUR LES MESURES EMPIRIQUES ASSOCIÉES À UNE CHAÎNE DE MARKOV

Nous allons ici retrouver simplement le principe de Laplace linéaire associé aux mesures empiriques d'une chaîne de Markov finie, c'est-à-dire la formule (14). Certains des arguments utilisés peuvent s'étendre à un espace d'état mesurable quelconque, sous des conditions d'irréductibilité appropriées, mais nous ne poursuivrons pas dans cette voie car cela nous ferait sortir des objectifs de ce manuscrit. On se contentera donc de considérer le cas d'une matrice markovienne irréductible  $M$  sur un ensemble fini  $E$ . Comme dans la section 2, la température ne jouera aucun rôle et on laissera tomber les exposants  $\beta$ , en prenant toujours  $\beta = 1$ .

Soit  $m_0 \in \mathcal{P}(E)$  et  $U \in \mathcal{B}(E)$  donnés, on considère une chaîne de Markov  $(X_n)_{n \in \mathbb{N}}$  de distribution initiale  $m_0$  et de probabilités de transition décrites par  $M$ . Le but de cet appendice est de montrer que la limite suivante existe

$$\lim_{n \rightarrow \infty} \frac{1}{n+1} \ln (\mathbb{E}_{m_0} [\exp(-U(X_0) - U(X_1) - \dots - U(X_n))])$$

et de l'identifier avec

$$- \inf_{\eta \in \mathcal{P}(E)} \eta(U) + \mathcal{L}(\eta)$$

où  $\mathcal{L}(\eta)$  a été défini en (13). Pour ceci, commençons par réécrire cette quantité sous la forme

$$\forall \eta \in \mathcal{P}(E), \quad \mathcal{L}(\eta) = \inf_{P \in \mathcal{K}(\eta)} \text{Ent}(\eta \times P | \eta \times M) \tag{56}$$

où pour toute probabilité  $\eta \in \mathcal{P}(E)$  et pour tout noyau markovien  $P$  sur  $E$ ,  $\eta \times P$  désigne la probabilité sur  $E^2$  définie par

$$\forall x, y \in E, \quad (\eta \times P)(x, y) := \eta(x)P(x, y)$$

(rappelons aussi que  $\mathcal{K}(\eta)$  est l'ensemble des matrices markoviennes  $P$  sur  $E$  admettant  $\eta$  comme probabilité invariante). En effet, soit  $\eta \in \mathcal{P}(E)$  et  $P \in \mathcal{K}(\eta)$  tels que  $\eta \times P \ll \eta \times M$  et désignons par  $f$  la densité de Radon-Nikodym correspondante sur  $E^2$  :

$$\forall x, y \in E, \quad f(x, y) := \frac{\eta \times P(x, y)}{\eta \times M(x, y)}$$

(toujours avec la convention  $0/0 = 0$ ). On a alors pour tout  $x \in E$  tel que  $\eta(x) > 0$ ,  $P(x, y) = f(x, y)M(x, y)$  pour tout  $y \in E$ , l'égalité précédente devant être comprise dans  $\mathcal{P}(E)$ . Ainsi  $\eta$ -p.s. en  $x$ , on est assuré de

$$\text{Ent}(P(x, \cdot) | M(x, \cdot)) = \sum_{y \in E} \ln(f(x, y)) P(x, y)$$

et il reste à intégrer ceci en  $x$  par rapport à  $\eta$  pour se convaincre que

$$\text{Ent}(\eta \times P | \eta \times M) = \sum_{x \in E} \text{Ent}(P(x, \cdot) | M(x, \cdot)) \eta(x). \tag{57}$$

D'autre part, si  $\eta \times P$  n'est pas absolument continu par rapport à  $\eta \times M$ , les considérations précédentes montrent aisément que l'on a là encore une égalité, les deux termes étant infinis, d'où finalement la validité de (56).

Cependant il est bien connu que (57) peut se généraliser (voir par exemple les livres de Dupuis et Ellis [15], Theorem C.3.1, ou de Dembo et Zeitouni [14], Theorem D.13) : soient  $\eta, \gamma \in \mathcal{P}(E)$  et  $P, Q$  deux matrices markoviennes sur  $E$ , on a

$$\text{Ent}(\eta \times P | \gamma \times Q) = \text{Ent}(\eta | \gamma) + \sum_{x \in E} \text{Ent}(P(x, \cdot) | Q(x, \cdot)) \eta(x). \quad (58)$$

En effet, on commence par considérer la situation où  $\eta \times P \ll \gamma \times Q$ . Par projection sur le premier facteur, il apparaît que l'on a alors aussi  $\eta \ll \gamma$  et il ressort de manière élémentaire que  $\gamma \times Q$ -p.s. en  $(x, y)$  :

$$\frac{\eta \times P}{\gamma \times Q}(x, y) = \frac{\eta(x) P(x, y)}{\gamma(x) Q(x, y)}$$

ainsi en prenant le logarithme de ces expressions puis en intégrant par rapport à  $\eta \times P$ , on obtient (58). Comme précédemment, cette égalité reste valable même si on n'a pas  $\eta \times P \ll \gamma \times Q$ , les deux membres étant maintenant infinis.

D'ailleurs, les mêmes arguments montrent que cette identité est vraie pour des espaces facteurs différents  $E'$  et  $E$ , c'est-à-dire si  $\eta, \gamma \in \mathcal{P}(E')$  pour un certain ensemble fini  $E'$  et si  $P, Q$  sont deux noyaux markoviens de  $E'$  vers  $E$  (*i.e.* des matrices de probabilités de transition indicées par  $E' \times E$ ).

Cette remarque permet d'étendre par récurrence (58) : pour  $n \in \mathbb{N}$ ,  $\eta \in \mathcal{P}(E)$  et  $P$  un noyau markovien sur  $E$ , notons  $\eta \times P^{\times n}$  la probabilité définie sur  $E^{n+1}$  par

$$\forall (x_0, x_1, \dots, x_n) \in E^{n+1}, \quad (\eta \times P^{\times n})(x_0, x_1, \dots, x_n) := \eta(x_0) P(x_0, x_1) \cdots P(x_{n-1}, x_n).$$

Il s'agit de la loi de  $(Y_0, Y_1, \dots, Y_n)$  quand  $(Y_p)_{p \in \mathbb{N}}$  est une chaîne de Markov de loi initiale  $\eta$  et de probabilités de transition homogènes données par  $P$ . Par contre, on désigne par  $\eta P^n$  la loi de  $Y_n$  correspondante. Pour tout  $n \in \mathbb{N}$ , en utilisant (58) avec  $E' = E^{n+1}$  et  $E$ , il apparaît que pour tous  $\eta, \gamma \in \mathcal{P}(E)$  et tous noyaux markoviens  $P, Q$  sur  $E$ ,

$$\text{Ent}(\eta \times P^{\times(n+1)} | \gamma \times Q^{\times(n+1)}) = \text{Ent}(\eta \times P^{\times n} | \gamma \times Q^{\times n}) + \text{Ent}(\eta P^n \times P | \eta P^n \times Q)$$

puis par itération il ressort que

$$\text{Ent}(\eta \times P^{\times(n+1)} | \gamma \times Q^{\times(n+1)}) = \text{Ent}(\eta | \gamma) + \sum_{0 \leq p \leq n} \text{Ent}(\eta P^p \times P | \eta P^p \times Q).$$

Cette formule se simplifie si l'on suppose que  $\eta$  est invariante par  $P$ , car alors

$$\text{Ent}(\eta \times P^{\times(n+1)} | \gamma \times Q^{\times(n+1)}) = \text{Ent}(\eta | \gamma) + (n+1) \text{Ent}(\eta \times P | \eta \times Q). \quad (59)$$

Pour faire bon usage de cette identité, rappelons une autre propriété de l'entropie :

$$\forall \eta, \gamma \in \mathcal{P}(E), \quad \text{Ent}(\eta | \gamma) = \sup_{f \in \mathcal{B}(E)} \eta(f) - \ln(\gamma[\exp(f)]). \quad (60)$$

En effet, pour se convaincre du fait que pour tout  $f \in \mathcal{B}(E)$ ,

$$\text{Ent}(\eta | \gamma) \geq \eta[f] - \ln(\gamma[\exp(f)]) \quad (61)$$

il suffit de remarquer que

$$\text{Ent}(\eta|\gamma) - \eta[f] + \ln(\gamma[\exp(f)]) = \text{Ent}(\eta|\nu)$$

avec  $\nu := \exp(f) \cdot \gamma / \gamma[\exp(f)]$ . Ceci permet aussi d'obtenir l'égalité dans (60), du moins si  $\eta$  et  $\gamma$  sont équivalentes, en prenant  $f = \ln(d\eta/d\gamma)$ . Si on a seulement  $\eta \ll \gamma$ , il vaut mieux utiliser dans le membre de droite de (60) la suite de fonctions  $(f_n)_{n \in \mathbb{N}}$  de  $\mathcal{B}(E)$  définie par

$$\forall n \in \mathbb{N}, \quad f_n := (-n) \vee \ln(d\eta/d\gamma)$$

pour se convaincre de l'égalité voulue. Enfin si  $\eta \not\ll \gamma$ , soit  $x \in E$  tel que  $\eta(x) > 0$  et  $\gamma(x) = 0$ , en considérant la suite de fonctions  $(n\mathbb{1}_{\{x\}})_{n \in \mathbb{N}}$  dans le membre de droite de (60), on déduit facilement l'identité (60) (à partir du cas où  $\eta \sim \gamma$ , on aurait aussi pu s'en sortir en invoquant des résultats de densité et de continuité, par exemple celle des applications partielles  $\mathcal{P}(E) \ni \eta \mapsto \text{Ent}(\eta|\gamma) \in \bar{\mathbb{R}}_+$  et  $\mathcal{P}(E) \ni \eta \mapsto \text{Ent}(\gamma|\eta) \in \bar{\mathbb{R}}_+$ , pour tout  $\gamma \in \mathcal{P}(E)$  fixé, mais attention, l'entropie relative n'est pas continue conjointement en ses deux arguments, dès que  $E$  n'est pas réduit à un singleton).

Soit  $\eta \in \mathcal{P}(E)$  et  $P \in \mathcal{K}(\eta)$ , en appliquant (61) avec  $E$  remplacé par  $E^{n+1}$ , pour  $n \in \mathbb{N}$ ,  $\eta$  par  $\eta \times P^{\times n}$ ,  $\gamma$  par  $m_0 \times M^{\times n}$  et en considérant pour application  $f$  :

$$E^{n+1} \ni (x_0, x_1, \dots, x_n) \mapsto -U(x_0) - U(x_1) - \dots - U(x_n)$$

on obtient que

$$\begin{aligned} & -\ln(\mathbb{E}_{m_0}[\exp(-U(X_0) - U(X_1) + \dots - U(X_n))]) \\ & \leq \text{Ent}(\eta \times P^{\times n} | m_0 \times M^{\times n}) + \mathbb{E}_\eta[U(Y_0) + U(Y_1) + \dots + U(Y_n)] \end{aligned}$$

où  $(X_p)_{p \in \mathbb{N}}$  et  $(Y_p)_{p \in \mathbb{N}}$  sont des chaînes de Markov de transitions homogènes données respectivement par  $M$  et  $P$ . En tenant compte de l'invariance de  $\eta$  par  $P$  et de (59), il s'ensuit que

$$-\ln(\mathbb{E}_{m_0}[\exp(-U(X_0) - U(X_1) + \dots - U(X_n))]) \leq \text{Ent}(\eta | m_0) + n\text{Ent}(\eta \times P | \eta \times M) + n\eta[U]$$

d'où si  $m_0 \gg \eta$ ,

$$\begin{aligned} \liminf_{n \rightarrow \infty} \frac{1}{n+1} \ln(\mathbb{E}_{m_0}[\exp(-U(X_0) - U(X_1) + \dots - U(X_n))]) & \geq -(\eta[U] + \text{Ent}(\eta \times P | \eta \times M)) \\ & \geq -(\eta[U] + \mathcal{L}(\eta)). \end{aligned}$$

En particulier, si l'on prend  $m_0$  chargeant tous les points de  $E$  (notamment si  $m_0 = m$ , la probabilité invariante pour  $M$ ), on obtient

$$\liminf_{n \rightarrow \infty} \frac{1}{n+1} \ln(\mathbb{E}_{m_0}[\exp(-U(X_0) - U(X_1) + \dots - U(X_n))]) \geq - \inf_{\eta \in \mathcal{P}(E)} \eta[U] + \mathcal{L}(\eta). \quad (62)$$

Réciproquement, soit  $\lambda > 0$  et  $0 < H \in \mathcal{B}(E)$  tels que

$$\frac{H}{M[H]} = \frac{\exp(-U)}{\lambda}$$

(ces deux objets peuvent s'interpréter comme une valeur propre de Perron-Frobenius et un vecteur propre correspondant, voir la section 2). Comme en (6), on leur associe le noyau markovien  $K$  sur  $E$  défini par

$$\forall x \in E, \forall A \subset E, \quad K(x, A) := \frac{M[H\mathbb{1}_A](x)}{M[H](x)}.$$

Cependant on peut écrire que pour tout  $n \in \mathbb{N}$  et tout  $(x_0, x_1, \dots, x_n) \in E^{n+1}$ ,

$$\exp(-U(x_0) - U(x_1) - \dots - U(x_n)) = \lambda^{n+1} \frac{H(x_0)}{M[H](x_0)} \dots \frac{H(x_n)}{M[H](x_n)}.$$

Or, on a vu dans la section 2 que le processus défini par

$$\forall n \in \mathbb{N}, \quad M_n := \frac{H(X_0)}{M[H](X_0)} \dots \frac{H(X_{n-1})}{M[H](X_{n-1})} H(X_n)$$

est une martingale relativement à la probabilité  $\mathbb{P}_{m_0}$ , pour toute distribution initiale  $m_0 \in \mathcal{P}(E)$ . Il en découle que si l'on pose

$$h := \frac{\max_E H}{\min_E H}$$

alors

$$h^{-1} \lambda^{n+1} \leq \mathbb{E}_{m_0}[\exp(-U(x_0) - U(x_1) - \dots - U(x_n))] \leq h \lambda^{n+1}$$

ce qui nous assure notamment de l'existence de la limite suivante

$$\lim_{n \rightarrow \infty} \frac{1}{n+1} \ln(\mathbb{E}_{m_0}[\exp(-U(x_0) - U(x_1) - \dots - U(x_n))]) = \ln(\lambda). \quad (63)$$

Par ailleurs, soit  $\mu$  la probabilité invariante pour  $K$  (qui a hérité de l'irréductibilité de  $M$ ), ainsi on a  $K \in \mathcal{K}(\mu)$ , d'où

$$\begin{aligned} \mathcal{L}(\mu) &\leq \text{Ent}(\mu \times K | \mu \times M) \\ &= \sum_{x \in E} \text{Ent}(K(x, \cdot) | M(x, \cdot)) \mu(x) \\ &= \sum_{x, y \in E} \ln \left( \frac{K(x, y)}{M(x, y)} \right) \mu(x) K(x, y) \\ &= \sum_{x, y \in E} \ln \left( \frac{H(y)}{M[H](x)} \right) \mu(x) K(x, y) \\ &= \sum_{x, y \in E} \ln \left( \frac{\exp(-U(y)) M[H](y)}{\lambda M[H](x)} \right) \mu(x) K(x, y) \\ &= -\ln(\lambda) - \sum_{x, y \in E} U(y) \mu(x) K(x, y) + \sum_{x, y \in E} \ln(M[H](y)) \mu(x) K(x, y) \\ &\quad - \sum_{x, y \in E} \ln(M[H](x)) \mu(x) K(x, y) \\ &= -\ln(\lambda) - \mu[U] + \mu[\ln(M[H])] - \mu[\ln(M[H])] \\ &= -\ln(\lambda) - \mu[U]. \end{aligned}$$

Il en résulte que

$$\begin{aligned} \inf_{\eta \in \mathcal{P}(E)} \eta[U] + \mathcal{L}(\eta) &\leq \mu[U] + \mathcal{L}(\mu) \\ &\leq -\ln(\lambda) \end{aligned}$$

puis que pour toute probabilité  $m_0$  chargeant tous les points de  $E$ , d'après (62),

$$\begin{aligned} & \lim_{n \rightarrow \infty} \frac{1}{n+1} \ln (\mathbb{E}_{m_0} [\exp(-U(x_0) - U(x_1) - \dots - U(x_n))]) \\ &= \ln(\lambda) \\ &\leq - \inf_{\eta \in \mathcal{P}(E)} \eta[U] + \mathcal{L}(\eta) \\ &\leq \lim_{n \rightarrow \infty} \frac{1}{n+1} \ln (\mathbb{E}_{m_0} [\exp(-U(x_0) - U(x_1) - \dots - U(x_n))]). \end{aligned}$$

On en déduit l'égalité de tous ces termes, puis en revenant à (63), l'égalité voulue pour toute distribution initiale  $m_0$  :

$$\lim_{n \rightarrow \infty} \frac{1}{n+1} \ln (\mathbb{E}_{m_0} [\exp(-U(x_0) - U(x_1) - \dots - U(x_n))]) = - \inf_{\eta \in \mathcal{P}(E)} \eta[U] + \mathcal{L}(\eta).$$

En fait, le terme de droite peut aussi s'écrire

$$- \inf_{\eta \in \mathcal{P}(E), P \in \mathcal{K}(\eta)} \eta[U] + \text{Ent}(\eta \times P | \eta \times M)$$

mais on peut remplacer cet infimum par un minimum, puisque ce qui précède montre qu'il est atteint pour  $\eta = \mu$  et  $P = K$ .

**Remarque 42.** En reprenant la terminologie de Dupuis et Ellis [15], on dit que les mesures empiriques définies par

$$\forall n \in \mathbb{N}, \quad L_n := \frac{1}{n+1} \sum_{0 \leq p \leq n} \delta_{X_p}$$

satisfont sous  $\mathbb{P}_{m_0}$  un principe de Laplace, si pour toute application continue  $G : \mathcal{P}(E) \rightarrow \mathbb{R}$  (ici nécessairement bornée, par compacité de  $\mathcal{P}(E)$ ), on a

$$\lim_{n \rightarrow \infty} \frac{1}{n+1} \ln (\mathbb{E}_{m_0} [\exp(G(L_n))]) = - \inf_{\eta \in \mathcal{P}(E)} G(\eta) + \mathcal{L}(\eta)$$

lequel principe est équivalent à un principe de grandes déviations pour les  $(L_n)_{n \in \mathbb{N}}$  sous  $\mathbb{P}_{m_0}$  avec  $\mathcal{L}$  pour fonctionnelle de taux (voir [15]).

Mais la preuve que nous avons donnée ci-dessus n'est valable que pour des applications  $G$  linéaires, c'est-à-dire pour lesquelles il existe  $U \in \mathcal{B}(E)$  tel que

$$\forall \eta \in \mathcal{P}(E), \quad G(\eta) = \eta[U].$$

Evidemment, du moins une fois vérifiée la stricte convexité de  $\mathcal{L}$  sur  $\mathcal{P}(E)$ , le théorème de Gärtner-Ellis (cf. par exemple [14]) permet de retrouver le principe de grandes déviations, et par conséquence le principe de Laplace « total », à partir de ce principe de Laplace seulement linéaire.



## RÉFÉRENCES

- [1] C. Ané, S. Blachère, D. Chafaï, P. Fougères, I. Gentil, F. Malrieu, C. Roberto and G. Scheffer, *Sur les inégalités de Sobolev logarithmiques, Panoramas et Synthèses [Panoramas and Syntheses]*. Société Mathématique de France, **10** Paris (2000). With a preface by Dominique Bakry and Michel Ledoux.
- [2] R. Bott and J.P. Mayberry, Matrices and trees, in *Economic activity analysis*, O. Morgenstern Ed., John Wiley and Sons, Inc., New York (1954) 391–400.
- [3] O. Catoni, Simulated annealing algorithms and Markov chains with rare transitions, in *Séminaire de Probabilités, XXXIII, Lect. Notes Math.* **1709** (1999) 69–119.
- [4] R. Cerf, The dynamics of mutation-selection algorithms with large population sizes. *Ann. Inst. H. Poincaré Probab. Statist.* **32** (1996) 455–508.
- [5] R. Cerf, A new genetic algorithm. *Ann. Appl. Probab.* **6** (1996) 778–817.
- [6] D. Concorde, Estimation de la densité du recuit simulé. *Ann. Inst. H. Poincaré Probab. Statist.* **30** (1994) 265–302.
- [7] P. Del Moral, M. Ledoux and L. Miclo, On contraction properties of Markov kernels. *Probab. Theory Related Fields* **126** (2003) 395–420.
- [8] P. Del Moral and L. Miclo, On the convergence and applications of generalized simulated annealing. *SIAM J. Control Optim.* **37** (1999) 1222–1250 (electronic).
- [9] P. Del Moral and L. Miclo, Branching and interacting particle systems approximations of Feynman-Kac formulae with applications to non-linear filtering, in *Séminaire de Probabilités, XXXIV, Lect. Notes Math.* **1729** (2000) 1–145.
- [10] P. Del Moral and L. Miclo, Annealed Feynman-Kac models. *Comm. Math. Phys.* **235** (2003) 191–214.
- [11] P. Del Moral, *Feynman-Kac formulae*. Probability and its Applications (New York). Springer-Verlag, New York (2004). Genealogical and interacting particle systems with applications.
- [12] P. Del Moral and A. Guionnet, On the stability of interacting processes with applications to filtering and genetic algorithms. *Ann. Inst. H. Poincaré Probab. Statist.* **37** (2001) 155–194.
- [13] P. Del Moral and L. Miclo, On the stability of nonlinear Feynman-Kac semigroups. *Ann. Fac. Sci. Toulouse Math.* **11** (2002) 135–175.
- [14] Amir Dembo and Ofer Zeitouni, *Large deviations techniques and applications, Applications of Mathematics* (New York). Springer-Verlag, New York, second edition **38** (1998).
- [15] P. Dupuis and R.S. Ellis, *A weak convergence approach to the theory of large deviations*. Wiley Series in Probability and Statistics : Probability and Statistics. John Wiley & Sons Inc., New York (1997). A Wiley-Interscience Publication.
- [16] M.I. Freidlin and A.D. Wentzell, *Random perturbations of dynamical systems, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, New York **260** (1984). Translated from the Russian by Joseph Szücs.
- [17] B. Hajek, Cooling schedules for optimal annealing. *Math. Oper. Res.* **13** (1988) 311–329.
- [18] R. Holley and D. Stroock, Simulated annealing via Sobolev inequalities. *Comm. Math. Phys.* **115** (1988) 553–569.
- [19] T. Lindvall, *Lectures on the coupling method*. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons Inc., New York (1992). A Wiley-Interscience Publication.
- [20] L. Miclo, About relaxation time of finite generalized Metropolis algorithms. *Ann. Appl. Probab.* **12** (2002) 1492–1515.
- [21] L. Miclo, Une étude des algorithmes de recuit simulé sous-admissibles. *Ann. Fac. Sci. Toulouse Math.* **4** (1995) 819–877.
- [22] L. Miclo, Sur les problèmes de sortie discrets inhomogènes. *Ann. Appl. Probab.* **6** (1996) 1112–1156.
- [23] L. Miclo, Sur les temps d’occupations des processus de Markov finis inhomogènes à basse température. *Stoch. Stoch. Rep.* **63** (1998) 65–137.
- [24] L. Miclo, Une variante de l’inégalité de Cheeger pour les chaînes de Markov finies. *ESAIM: Probab. Statist.* **2** (1998) 1–21. (electronic).
- [25] E. Seneta, *Nonnegative matrices and Markov chains*. Springer Series in Statistics. Springer-Verlag, New York, second edition, 1981.
- [26] W. Stannat, On the convergence of genetic algorithms – a variational approach. *Probab. Theory Related Fields* **129** (2004) 113–132.
- [27] A. Trounev, Cycle decompositions and simulated annealing. *SIAM J. Control Optim.* **34** (1996) 966–986.